Nanoindentation of Polymers: An Overview

Mark R. VanLandingham*,†, John S. Villarrubia†, Will F. Guthrie†, and Greg F. Meyers‡

† National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899
‡ The Dow Chemical Company, Analytical Sciences, 1897E Building, Midland, MI 48667

SUMMARY: In this paper, the application of instrumented indentation devices to the measurement of the elastic modulus of polymeric materials is reviewed. This review includes a summary of traditional analyses of load-penetration data and a discussion of associated uncertainties. Also, the use of scanning probe microscopes to measure the nanoscale mechanical response of polymers is discussed, particularly with regard to the associated limitations. The application of these methods to polymers often leads to measurements of elastic modulus that are somewhat high relative to bulk measurements with potentially artificial trends in modulus as a function of penetration depth. Also, power law fits to indentation unloading curves are often a poor representation of the actual data, and the power law exponents tend to fall outside the theoretical range. These problems are likely caused by viscoelasticity, the effects of which have only been studied recently. Advancement of nanoindentation testing toward quantitative characterization of polymer properties will require material-independent calibration procedures, polymer reference materials, advances in instrumentation, and new testing and analysis procedures that account for viscoelastic and viscoplastic polymer behavior.

Introduction

Depth-sensing indentation (DSI) devices allow the amount of penetration of an indenter tip into a material to be measured often using either a constant loading rate or a constant displacement rate1-3). Further, these devices are often capable of producing contact areas and penetration depths characterized by sub-micrometer or even nanometer dimensions for hard materials (e.g., single-crystal silicon, hardness = 14 GPa3)). One objective of using DSI methods is to produce quantitative, absolute measurements of elastic modulus, E. Producing such measurements with nanoscale spatial resolution can be a key to understanding mechanical behavior of technologically important material systems. However, polymeric materials create significant challenges to measuring E accurately using indentation testing. First, many polymers are so soft that the material response cannot be measured at all with DSI devices because the system compliances are too low3). Even for stiffer polymers (E > 1 GPa), producing indents with both lateral and depth dimensions much less than 1 µm is difficult.
This difficulty is related to the load resolution, typically not better than $\pm 100 \text{ nN}$, and the inability of most DSI systems to detect initial contact loads less than $1 \text{ µN}$. Thus, although DSI systems are capable of applying maximum loads on the order of $1 \text{ µN}$, the smallest maximum loads applied in practice are typically tens of micronewtons to reduce the relative uncertainties in load and penetration depth. Finally, current analysis of DSI data is based on elasticity, which when applied to viscoelastic materials could lead to large uncertainties in the calculated values of $E$. Thus, current DSI methods have limited capabilities for studying polymer thin films, polymer composites, and other important polymer systems for which obtaining property information with nanometer spatial resolution is often desired. Further, other calibration and procedural issues persist with regard to modulus measurements that must be addressed prior to the application of these methods to viscoelastic materials. For example, current calibration procedures used in DSI rely on indentation of a reference material of known modulus. Recent studies, including an interlaboratory comparison\textsuperscript{5}, have shown the calibration results to have poor reproducibility and large uncertainties.

Scanning probe microscopy (SPM), and in particular the atomic force microscope (AFM), has been used to study polymeric materials with nanometer spatial resolution. The AFM employs a probe consisting of a sharp tip (nominal tip radius on the order of 10 nm) located near the end of a cantilever beam that is scanned across the sample surface using piezoelectric scanners. The AFM can also be operated in force mode to perform indentation tests. A force curve is produced, which is a plot of tip deflection as a function of the vertical motion of the scanner. This curve can be analyzed to provide information on the local mechanical response\textsuperscript{6-8}. By changing the spring constant of the probe, forces applied during indentation can be varied from a few nanonewtons up to tens or perhaps hundreds of micronewtons. Also, the spring constant of the cantilever probe can be chosen such that small differences in response can be detected for polymers having a certain range of stiffness corresponding to the chosen spring constant\textsuperscript{6}. However, accurate measurement of the probe spring constant can be difficult to achieve. This difficulty combined with a lack of information regarding the tip shape of the AFM probes generally reduces AFM indentation measurements to relative measurements.

A modified version of the AFM, called the interfacial force microscope (IFM)\textsuperscript{9-11}, has also been used to probe the indentation response of polymeric materials. For the IFM, applied force is measured using a self-balancing differential capacitance sensor. An electrochemically sharpened tungsten tip is displaced relative to the sample at a given
displacement rate by piezoelectric scanners, and the measured force is recorded as a function of displacement. Often attempts are made to achieve purely elastic deformation by applying penetration depths of only a few nanometers to tens of nanometers, which is small compared to the tip radius (ranges from 70 nm to 500 nm) of the tungsten probe. In this case, Hertzian contact mechanics can be applied to the loading curve or unloading curve, as has generally been done for indentation testing with the IFM.

In this paper, an overview of nanoindentation and its application to polymeric materials is given. First, traditional analyses of load-penetration data are presented followed by a discussion of uncertainty issues with regard to determining accurate measurements of elastic modulus. The application of depth-sensing indentation to measurements of polymer response is then reviewed. The use of scanning probe microscopes to measure mechanical properties of polymers is then presented, including comparisons to depth-sensing indentation where appropriate.

Overview of Nanoindentation

The Method of Oliver and Pharr

The analysis of indentation load-penetration curves produced by depth-sensing indentation systems is often based on work by Oliver and Pharr\(^1\). Their analysis was in turn based upon relationships developed by Sneddon\(^2\) for the penetration of a flat elastic half space by different probes with particular axisymmetric shapes (e.g., a flat-ended cylindrical punch, a paraboloid of revolution, and a cone). In general, the relationships between penetration depth, \(h\), and load, \(P\), for such indenter geometries can be represented in the form

\[
P = \alpha (h - h_f)^m
\]

where \(\alpha\) contains geometric constants, the sample elastic modulus and Poisson's ratio, and the indenter elastic modulus and Poisson's ratio, \(h_f\) is the final unloading depth, and \(m\) is a power law exponent that is related to the geometry of the indenter; for a flat-ended cylindrical punch, \(m = 1\), for a paraboloid of revolution, \(m = 1.5\), and for a cone, \(m = 2\).

In applying Equation 1 to the calculation of modulus, Oliver and Pharr\(^1\) made two significant realizations. First, the slope of the unloading curve changes constantly due to constant changes in the contact area. In prior research, the upper portion of the unloading curve was approximated as linear, which incorrectly assumes that the contact area remains constant for
the initial unloading of the material. This practice created a dependence of calculated modulus values on the number of points used in the linear fit\(^1\). Second, if the unloading curve can be fit by a power law expression given by Equation 1, then a derivative, \(dP/dh\), of that expression applied at the maximum loading point \((h_{\text{max}}, P_{\text{max}})\) should yield information about the state of contact at that point. This derivative was termed the contact stiffness, \(S\), and is given by

\[
S = 2aE_r = \frac{2\beta}{\sqrt{\pi}} E_r \sqrt{A}
\]

(2)

where \(a\) is the contact radius and \(A\) is the projected area of tip-sample contact. The reduced modulus, \(E_r\), accounts for deformation of both the indenter and the sample and is given by

\[
\frac{1}{E_r} = \left(\frac{1 - \nu^2}{E}\right) + \left(\frac{1 - \nu_i^2}{E_i}\right)
\]

(3)

where \(E\) and \(\nu\) are the sample elastic modulus and Poisson's ratio, respectively, and \(E_i\) and \(\nu_i\) are the elastic modulus and Poisson's ratio, respectively, of the indenter material. \(\beta\) is used to account for the triangular and square cross sections of many indenters used in nanoindentation studies. For \(\beta = 1\), the cross section of the indenter is assumed to be circular, as the contact radius, \(a\), is replaced by \((A/\pi)^{1/2}\), and Equation 2 (with \(\beta = 1\)) is valid for any indenter that has a shape described by a solid body of revolution of a smooth function. However, the values of \(\beta\), as determined by King\(^13\) using numerical analysis, are only small corrections (e.g., \(\beta = 1.034\) for a triangular punch) and are not often used in practice. Another correction factor has recently been suggested due to unrealistic boundary conditions used by Sneddon\(^14\) (and also Hertz\(^15\)). This correction factor, \(\gamma\), which depends upon the sample Poisson's ratio, \(\nu\), and tip geometry ranges from approximately 1.05 to 1.10 for \(\nu\) between 0.1 and 0.4 and a conical indenter with an opening angle of 70.32°.

In Figure 1, an indentation load-displacement curve is illustrated along with several important parameters used in the Oliver and Pharr analysis. The stiffness, \(S^*\), is the slope of the tangent line to the unloading curve at the maximum loading point \((h_{\text{max}}, P_{\text{max}})\) and is given by

\[
S^* = \left(\frac{dP}{dh}\right)_{(h_{\text{max}}, P_{\text{max}})} = a\alpha n(h_{\text{max}} - h_f)^{n-1}
\]

(4)
where the parenthetic subscript denotes that the derivative is evaluated at the maximum loading point. When the displacement, $h$, is the total measured displacement of the system, $S^*$ is the total system stiffness. After successful calibration and removal of the load-frame compliance, the displacement of the load frame is removed so that $h$ represents only the displacement of the tip into the sample. In this case, $S^* = S$ and the tangent line represents an unloading path for which the contact area does not change. Thus, the contact area, $A$, calculated using $S$ (see Equation 2) should be the actual contact area at maximum load. Also, extrapolating this line down to $P = 0$ yields an intercept value for depth, $h_i$, which should be related to the contact depth, $h_c$, associated with the maximum loading point. However, $h_c$ is related to the deformation behavior of the material and the shape of the indenter, as illustrated in Figure 2. In fact, $h_c = h_{\text{max}} - h_s$, where $h_s$ is defined as the elastic displacement of the surface at the contact perimeter and can be calculated for specific geometries using displacement equations from Sneddon's analyses\textsuperscript{12}). For each of three specific tip shapes (flat-ended punch, paraboloid of revolution, and cone), $h_s = \epsilon P_{\text{max}} / S$ where $\epsilon$ is a function of the particular tip geometry, as summarized in Table 1. Thus, $h_c$ is given by

$$h_c = h_{\text{max}} - \frac{\epsilon P_{\text{max}}}{S} \quad (5)$$

The nanoindentation procedures include calibration of the load-frame compliance, $C_{\text{lf}}$, and the tip shape area function, $A(h_c)$. Prior to the load-frame compliance calibration, the measured displacement, $h_{\text{total}}$, is a combination of displacement of the load frame, $h_{\text{lf}}$, and displacement of the sample, $h_{\text{samp}}$. Treating the system as two springs (the load frame and the sample) in series under a given load, $P$,

$$h_{\text{total}} = h_{\text{lf}} + h_{\text{samp}} \quad (6)$$

Dividing both sides by $P$,

$$C_{\text{total}} = C_{\text{lf}} + C_s = C_{\text{lf}} + \frac{\sqrt{\pi}}{2E_r} \frac{1}{\sqrt{A}} \quad (7)$$

where the total compliance $C_{\text{total}} = 1/S^*$ and the sample compliance $C_s = 1/S$. A number of possible methods exist for determining $C_{\text{lf}}$ using a reference sample that is homogeneous and isotropic and for which both $E$ and $\nu$ are known. Typically, a series of indentation measurements are made on the reference sample. Oliver and Pharr\textsuperscript{1)} suggested using an iterative technique to calibrate both the load-frame compliance and the tip shape with one set
of data from a single reference sample, as both $C_{IF}$ and $A$ are unknowns in Equation 7. While this method has the advantage of not requiring an independent measurement of the area of each indent, its use has been limited, perhaps because it is mathematically intensive.

Fig. 1: An indentation load-displacement curve in which several important parameters used in the Oliver and Pharr analysis are illustrated.

Fig. 2: Illustration of the indentation geometry at maximum load for an ideal conical indenter.

The use of the AFM with indentation measurements provides a method of high-resolution imaging of the plastic impression, which should have approximately the same projected area as the contact area at maximum load, particularly for a highly plastic reference material such
as aluminum. Using this type of an approach, the measured compliance, $C_{\text{total}}$, can be plotted as a function of $1/\sqrt{A}$. A linear curve fit to the data can then be used to determine the load-frame compliance, $C_{\text{lf}}$, which will be the value of the y-intercept. A third method and the one used in the present research is to assume that not only is $E$ independent of penetration depth but also hardness, $H = P_{\text{max}}/A$. Thus, if $H$ is constant, $C_{\text{total}}$ can be plotted as a function of $1/\sqrt{P_{\text{max}}}$, and again the y-intercept of the fitted linear curve yields $C_{\text{lf}}$. In this method, aluminum is often replaced by fused silica, because oxide formation on aluminum can create variations in $E$ and $H$ with penetration depth.

For the load-frame compliance calibration, relatively large indentation loads and depths are applied to a reference material that exhibits significant plastic deformation (e.g. aluminum) so that the contact stiffness is large ($C_s$ is small) and thus $C_{\text{total}}$ is dominated by $C_{\text{lf}}$. For the tip shape calibration, the series of indents applied to a reference material typically covers a larger range of maximum load and maximum penetration depth. The objective of tip shape calibration is to measure the cross-sectional area of the indenter tip as a function of distance from the apex. In Figure 2, the indentation geometry for a conical indenter is illustrated in two dimensions. At a given load, $P$, the contact area, $A$, which is related to the contact radius, $a$, is the cross-sectional area of the indenter tip at a distance, $h_c$ (the contact depth), from the tip apex. From measurements of $h_{\text{max}}$, $P_{\text{max}}$, and $S$, Equations 2 and 5 can be used to calculate $A$ and $h_c$, respectively, for each indentation. A tip shape function, $A(h_c)$, is determined, given a sufficient number of measurements over a range of $h_c$ values, by fitting the $A$ vs. $h_c$ data, typically using a multiterm polynomial fit of the form:

$$A(h_c) = B_0 h_c^2 + B_1 h_c + B_2 h_c^{1/2} + B_3 h_c^{1/4} + \ldots$$

where $B_0, B_1, \ldots, B_n$ are constant coefficients determined by the curve fit. Oliver and Pharr\(^1\) suggested using up to 9 terms ($n=8$) with $B_0 = 24.5$, stating that the area function of a perfect Berkovich indenter, which was the style of indenter they used, is given by

$$A(h_c) = \frac{24.5}{h_c^2}$$

and that all other terms account for deviations from the ideal geometry due to blunting of the tip.

Once the load frame compliance and tip shape calibrations have been performed, either separately or iteratively, measurements of elastic modulus for samples of interest can be made
from indentation data. The unloading curves are again fit to a power law function (see Equation 1), and the fitting parameters are used to calculate $S^*$ (see Equation 4), which is equal to $S$ assuming a correctly determined value of $C_{lf}$. $S$ is then used to calculate $h_c$ (see Equation 5), and $h_c$ is used to calculate $A$ from the tip shape area function. Finally, $S$ and $A$ are used to calculate $E$ using Equation 2.

Uncertainties in Indentation Analysis

Load Frame Compliance

In reviewing the Oliver-Pharr procedure that is typically applied to nanoindentation measurements, the potential for significant uncertainties to arise, particularly in the calibration procedures, and propagate through to the calculation of modulus becomes apparent. First, the uncertainties in the calibration of load-frame compliance ($C_{lf}$), in which a reference sample is indented using high maximum loads, has been shown to be dependent on the properties of the reference material\textsuperscript{5).} For the same DSI system and Berkovich tip, $C_{lf}$ measured by indenting fused silica ($E = 72$ GPa) was over a factor of two lower than that measured by indenting tungsten ($E = 410$ GPa). The uncertainty in $C_{lf}$ should decrease as the ratio of frame compliance to sample compliance, $C_s$, increases. $C_s$ is inversely proportional to the ratio $E/\sqrt{H}$, and this ratio is approximately a factor of 8 larger for tungsten compared to fused silica. Thus, the values of $C_{lf}$ determined using tungsten are likely to have less uncertainty compared to those determined using fused silica, and in fact the variation of $C_{lf}$ with indenter geometry was much smaller for indentation of tungsten than for indentation of fused silica.

In our own study of load frame compliance determination using a fused silica sample and a Berkovich tip, $C_{lf}$ was determined from a set of 24 indentations, 4 each at 6 nominal load levels ranging from 4.0 mN to 5.2 mN, which was the upper range of the DSI system. The power law fits from which $S^*$ was calculated were made using two different curve fitting methods. In the first method, a nonlinear power law fit was made with the DSI system software to a portion of the unloading data ranging from 95% of the maximum load, $P_{max}$, down to 20% of $P_{max}$. In the second method, a nonlinear power law fit was made with commercially available statistics software to nearly the entire unloading curve. Data at the top and bottom portions of the unloading curve were deleted only when the associated residual errors of the fit deviated from the following conventional assumptions\textsuperscript{16):}

1. the errors are normally distributed
2. the errors have zero mean
3. the errors are homogeneously distributed over the range of interest

Using these criteria, data was deleted from the upper portion of the curve for the curve fits of only 4 of the 24 data sets, with the largest remaining load value being greater than 98 % of $P_{\text{max}}$ in each of those 4 cases. Also, the amount of data deleted from the lower portion of the curve was generally between 2 % and 15 % of $P_{\text{max}}$. Typical residual plots corresponding to the final fits are shown for three different data sets in Figure 3, and the resulting values of system compliance (method 2) are compared to those determined using the DSI software fits (method 1) in Figure 4. To determine $C_{\text{lf}}$, $H$ was assumed to be constant with penetration depth such that a linear fit to $C_{\text{total}}$ vs. $1/\sqrt{P_{\text{max}}}$ (see Figure 4) yields a value of $C_{\text{lf}}$ equal to the y-intercept of the fitted linear curve (see Equation 7). For method 1, $C_{\text{lf}} = (3.6 \pm 1.9)$ nm/mN and for method 2, $C_{\text{lf}} = (2.4 \pm 3.0)$ nm/mN. Here and throughout this paper, a numerical value following a "±" symbol is an estimated standard deviation, $\sigma$, unless otherwise stated. Thus for method 2, the value determined for $C_{\text{lf}}$ is not statistically different from zero and for method 1, a value of 0 nm/mN falls in the 95 % confidence range defined by $\pm 2\sigma$. These large uncertainties will significantly affect the tip shape calibration and subsequent modulus measurements. Therefore, given the finite load capabilities of DSI systems, care must be taken to minimize these uncertainties by choosing a reference sample to minimize sample compliance, taking a large number of measurements, and taking measurements over a large range of applied loads, which should include the upper load limit of the system.

**Tip Shape Calibration**

Besides the uncertainty in the load frame compliance, the determination of tip shape from the indentation of a reference sample has a number of additional sources of uncertainty. First, the power law curve fitting procedures can impact the determination of $S$, which is used to calculate both the contact area, $A$, and the contact depth, $h_c$. Within the nanoindentation community, curve fits have been made using only a portion (often less than half) of the unloading data\(^{17}\). Because the power law model of Equation 1 has a physical basis that theoretically should fit the entire unloading curve, a sound basis for using only a portion of the data should be determined. (Note, however, that the relationship between $P$ and $h$ during indentation unloading was found not to follow a power law form in a recent dimensional analysis study\(^{18}\).) Generally, the rapid changes in velocity and direction that occur at the beginning of the unloading curve can affect the first few data points near the maximum load. Also, nonidealities near the tip apex and instrumental uncertainties at low loads can affect
data at the low-load end of the unloading curve. However, these affects are highly dependent on the particular system, tip, and sample used, and thus the most appropriate range of data to use will not be universal. Also, power law fits can be made either by using a linear fit to log(P)-log(h) data or through an iterative nonlinear fit, and in many studies, the type of fit that was used is not clear.

Fig. 3: Typical residual plots corresponding to the final nonlinear power law fits to unloading curves from the load frame compliance calibration data are shown for (a) $P_{\text{max}} = 4.0$ mN, (b) $P_{\text{max}} = 5.2$ mN, and (c) and (d) $P_{\text{max}} = 4.5$ mN. Plots (a), (b), and (c) are the differences between the predicted load, $P'$, and the measured load, $P$, vs. $P$, while plot (d) is a quantile-quantile plot of the residuals in (c).

Of course, the issue is whether or not the curve fit is representative of the data set, particularly in terms of determining an appropriate value of $S$. To study effects of curve fitting on tip shape calibration, an additional set of 27 indentations, 3 to 5 each at 7 nominal load levels ranging from 0.05 mN to 3.0 mN, were made on fused silica with a Berkovich tip and were combined with the 24 indentations used for the compliance calibration. A value of 3.6 nm/mN was used for $C_{\text{fr}}$. The same two curve fitting methods used in our previous study for
load frame compliance were used to fit the unloading curves for the 51 data sets. Again for
method 2, data were deleted only when the corresponding residual errors did not conform to
the conventional assumptions regarding curve fitting. Data was deleted from the upper
portion of the curve for the curve fits of only 5 of the additional 27 data sets, with the largest
remaining load value being greater than 97 % of $P_{\text{max}}$ in each of those 5 cases. Also, except in
a few cases, the amount of data deleted from the lower portion of the curve was generally less
than 20 % of $P_{\text{max}}$, as shown in Figure 5 for all 51 data sets. In fact for several unloading
curves with $P_{\text{max}} < 1 \, \text{mN}$, only a few points at the low load end were not used in the fit.

![Graph showing system compliance for two curve fitting methods](image)

Fig. 4: Values of system compliance for the two curve fitting methods are plotted as a
function of $1/\sqrt{P_{\text{max}}}$. For the high $P_{\text{max}}$ indents used for the compliance calibration, the values of $S$
differ by less than 1 % for the two methods. However, for decreasing values of $P_{\text{max}}$, the differences in $S$
values becomes more significant with a tendency of $S$ for method 2 ($S_2$) to be greater than $S$
for method 1 ($S_1$) by as much as 8 %. This result indicates that, for the lower-load indents,
the value of $S$ is more significantly affected by the arbitrary deletion of data points,
particularly the deletion of points from the high-load region of the unloading curve. Because
$A$ is related to $S^2$, values of $A$ corresponding to $S_2$ are approximately $(5.5 \pm 3.5) \%$ larger than
those corresponding to $S_1$ for $P_{\text{max}} < 0.3 \, \text{mN}$, with a maximum difference of over 16 %. Simultaneously,
values of $h_c$ corresponding to $S_2$ are approximately $(3.5 \pm 3.2) \%$ larger than
those corresponding to $S_1$ for $P_{\text{max}} < 0.3 \, \text{mN}$.
Fig. 5: The data from the lower portion of the unloading curves not used for the nonlinear power law fits is plotted as a percentage of $P_{\text{max}}$ vs. $P_{\text{max}}$.

The method of calculating linear fits to $\log(P)$-$\log(h)$ data was also considered. Taking the logarithm of both sides, Equation 1 becomes

$$\log(P) = \log(\alpha) + m\log(h - h_f)$$  \hspace{1cm} (10)

If $h_f$ were to be used as a fitting parameter as it is in the nonlinear fits, an iterative scheme for this fit would need to be devised. Instead of developing such an algorithm, the values of $h_f$ calculated in the nonlinear fits were used in Equation 10 to fit several of the unloading curves using in the preceding study. Examples of residual plots associated with these fits are shown in Figure 6. In Figure 6a and 6b, the residuals are shown in logarithmic units. In Figure 6c, the residuals are shown in units of load for comparison with Figure 6d, in which the residuals of the corresponding nonlinear fit are shown. Although the fitting parameters are similar for the two fits, the residuals in the units of the log-log fit (Figures 6a and 6b) deviate from the conventional assumptions discussed previously. Therefore, from a statistical standpoint, the nonlinear fit is preferred to a linear fit to logarithmic data.

The uncertainties in $A$ and $h_c$ related to the choice of curve fitting method, at least for this study, do not significantly affect the uncertainties in the tip shape area function, $A(h_c)$. In fact, four-term polynomial fits in the form of Equation 8 to the three sets of ($h_c$,A) data are statistically identical, as shown in Figure 7. Also, for $h_c < 100$ nm, the actual ($h_c$,A) data points calculated using the three different types of fits to unloading curves all fall within $\pm 2\sigma$.
of the $A(h_c)$ fits. These results are not surprising, given that as the measurement of $S$ changes, $A$ and $h_c$ will both either increase or decrease. Thus, differences in $S$ for a given unloading curve due to different curve fits will likely be similar to differences in $S$ for two unloading curves with the same nominal maximum load. However, as previously discussed, the calculated value of $C_{lf}$ appears to be quite sensitive to small changes in $S^*$ and the associated data scatter, at least for cases where the sample compliance is significant, such that curve fitting procedures do affect the determination of $C_{lf}$. Also, for samples of unknown modulus, $S$ is used only to calculate $h_c$, so the associated uncertainties including those related to curve fitting, will lead to uncertainties in the value of $A$ determined from $A(h_c)$ and thus to the value of $E$, particularly for lower load indents.

![Fig. 6](image.png)

Fig. 6: (a), (b), and (c) are residual plots for linear fits to log($P$)-log($h-h_0$) unloading data for $P_{\text{max}} = 0.5$ mN. The residuals are in logarithmic units for (a) and (b), while in (c) they are in the units of load. (a) and (c) are plots of residuals vs. load, while (b) is a quantile-quantile plot of the residuals in (a). (d) is the residual plot (residuals vs. load) for the corresponding nonlinear fit.

In a number of recent studies, material-independent methods of tip shape calibration using the AFM have been suggested and compared to the method of indenting a reference sample. In
one of these methods, the indenter tip is scanned with an AFM tip to yield direct information regarding the three-dimensional tip shape\textsuperscript{5,19). In another method, the indenter tip is used to scan samples with sharp features (referred to as tip characterizers), generating images that are then used to determine the three-dimensional tip shape (see Figure 8a) using the method of blind reconstruction\textsuperscript{20). For both methods, tip shape area functions generated with the material-independent methods have not agreed with those determined by indenting a reference sample (see Figure 8b).

Fig. 7: Plot of tip shape area function data, $A$ vs. $h_c$, for curve fitting methods 1 and 2 and for the linear fits to log(P)-log(h-h_f). Also shown are curves representing a four-term polynomial fit to each set of $A(h_c)$ data of the form of Equation 8 and ± two estimated standard deviations, $\sigma$, from each fit.

In Figure 8b, the indentation area function data was determined using $C_{lf} = 3.6$ nm/mN. The dashed curves in this figure indicate indentation area functions for $C_{lf} = (3.6 + 2\sigma)$ nm/mN (upper curve) and $C_{lf} = (3.6 - 2\sigma)$ nm/mN (lower curve), where $\sigma = 1.9$ nm/mN (see previous subsection). These dashed curves envelop all of the area function data. Thus, the large uncertainty in the load frame compliance could account for discrepancies between the material-independent and material-dependent tip shape area functions, at least in this case.
Other uncertainties for the indentation method include those related to the detection of a true zero in load and displacement, the load and displacement measurements themselves, and the elastic properties of the sample and indenter materials. Differences between current analysis methods based on elasticity theory and real material behavior, including correction factors $\beta$ and $\gamma$ and the pile-up or sink-in of material around the tip-sample contact, can also be significant sources of uncertainty. Uncertainties related to the material-independent AFM methods include the finite sizes of either the AFM tip or the tip characterizer features, as well as non-tip artifacts such as thermal drift, scanner nonlinearities, feedback-loop response time, and slight deviations from orthogonality of the scanners. Also, the material-independent methods do not account for deformation of the indenter tip, such that the derived tip shape is perhaps a better representation of the actual tip shape than that determined via indentation due to the inclusion of tip deformation. We are currently investigating the impact of these various uncertainties on the total uncertainty of modulus measurements as determined from indentation data.

While no complete uncertainty analysis has yet been performed regarding DSI methods, a few interlaboratory and/or intersystem comparisons have been done, all of which resulted in poor reproducibility in modulus measurements$^{5,21,22}$. In one of these studies that included 8 laboratories and 6 different instrument types, measurements of E for fused silica (nominally $E = 72$ GPa) ranged from $(67.2 \pm 0.4)$ GPa to $(84 \pm 3)$ GPa for $P_{\text{max}} = 100$ mN and from $(68.6 \pm 0.3)$ GPa to $(86.3 \pm 0.3)$ GPa for $P_{\text{max}} = 10$ mN$^5$. In another study with all data presumably taken by the same researchers but using 6 combinations of 3 different DSI instruments and 5 different tips, reduced modulus values for single crystal tungsten ranged from approximately $(210 \pm 50)$ GPa to $(470 \pm 60)$ GPa$^{21}$. In addition, several of the instrument-tip combinations yielded values of reduced modulus that increased with increasing contact depth while other combinations yielded values that increased with decreasing contact depth. The large uncertainties reported for these comparisons are likely due to uncertainties in the calibrations, as has been discussed, as well as the repeatability of the measurements for each instrument$^{22}$.

Finally, the use of ideal tip shapes in indentation calculations can create large uncertainties. In our study, the opening angles, $\beta$, between the normal and the three faces of the two Berkovich tip were calculated from the blind reconstruction data to range from approximately $66.5^\circ$ to $67.5^\circ$. These measured values of $\beta$ are not equal to either the original specification by Berkovich of $65.033^\circ$ degrees or the modified value of $65.3^\circ$ specified by Oliver and Pharr.
However, the measured values are in agreement with values measured for a number of Berkovich indenter tips using a metrological AFM that ranged between 65.7° and 67.4°.  

Fig. 8: (a) Blind reconstruction of a Berkovich indentation tip; and (b) a plot of the area-depth relationship for two Berkovich tips in which blind reconstruction results are compared to results from an indentation tip-shape analysis. In (b), the indentation area function data was determined using $C_{lf} = 3.6$ nm/mN, and the dashed curves approximate indentation area functions for $C_{lf} = (3.6 + 2\sigma)$ nm/mN (upper curve) and $C_{lf} = (3.6 - 2\sigma)$ nm/mN (lower curve), where $\sigma = 1.9$ nm/mN.
Besides the differences in opening angle, all conical and pyramidal indenter tips are likely to have some sort of defect geometry, for example blunting, at the tip apex, as noted by Briscoe and coworkers. From our blind reconstruction results, the two Berkovich tips studied had blunt tips, exhibiting pyramidal geometry only for vertical distances from the apex of 15 nm or greater. Also, the opening angle measured close to the apex (i.e., approaching the blunt region) was between $1^\circ$ and $2^\circ$ greater than the opening angle measured away the apex. As a result of these deviations from ideality, the area values corresponding to a given contact depth will be much different for real tips compared to an ideal tip, particularly for contact depths less than 100 nm as are commonly encountered in nanoindentation.

**Nanoindentation of Polymers**

*Nanoindentation Using DSI Systems*

A number of problems exist with the current application of nanoindentation techniques to polymeric materials using both DSI and SPM systems. In general, measurements of E using DSI tend to increase with decreasing penetration depth, often referred to as an indentation size effect. This artifact also appears to be a problem for indentation of polymers using DSI. Thus, either $E(\text{surface}) > E(\text{bulk})$ for a large number of materials, which seems unlikely, or these trends result from increased uncertainties for shallow depth indents that are likely due to tip defects near the apex and decreased signal-to-noise ratios at low load and displacement levels. Also, values of E measured for polymers using DSI are significantly higher than values measured using tensile testing or dynamic mechanical analysis (DMA). For example, Lucas et al. reported $E(\text{DSI}) = 1.2 \text{ GPa}$, $E(\text{tensile test}) = 0.4 \text{ GPa}$, and $E(\text{DMA}) = 0.5 \text{ GPa}$ for polytetrafluoroethylene (PTFE). For some studies, modulus values measured using DSI have been compared to handbook or manufacturer values. However, such comparisons can often be misleading, because quoted values of E for many polymer systems can cover a large range due to potential variations in microstructure, semicrystalline morphology, anisotropy, molecular weight, crosslink density, etc. Thus, comparisons of modulus values are most appropriate for polymer samples with identical chemistry, molecular weight, and processing history.

To date, relatively few quasi-static indentation studies of polymers using DSI systems have been published, largely because of the inadequacies of DSI systems and traditional analysis methods as applied to measuring polymer properties. For example, the point taken as representing initial contact ($h = 0$) generally corresponds to a significant applied load, such
that penetration of the indenter tip into a polymer could be significant relative to the maximum displacement, $h_{\text{max}}$. Also, viscoelastic creep during unloading can dramatically affect the slope of the unloading curve, leading to negative values of $S$ in extreme cases\textsuperscript{28,29}). This creep effect has been shown through modeling to be related to a continued increase in the contact radius during the initial portion of the unloading cycle\textsuperscript{30}). Because of this affect, power law fits will generally not be applicable to fitting the unloading response of polymers. In fact, nonlinear power law fits (see Equation 1) often will not reach a convergent solution. However, some commercial DSI system software will produce an answer based on the values of the fitting parameters for the Nth iteration, where N is the total number of iterations allowed by the fitting routine. Perhaps in realizing this problem, some researchers have reverted to determining $S$ using a linear approximation to the upper portion of the unloading curve\textsuperscript{31}), which assumes that the contact area remains constant for the initial unloading of the material. However, this practice creates a dependence of $S$ on the number of points used in the linear fit, as discussed previously. Moreover, it ignores the cause of the deviation from power law behavior, namely viscoelastic creep.

In our recent studies, a benzocyclobutene (BCB) polymer described in detail elsewhere\textsuperscript{20}) was indented with a DSI system using a Berkovich indenter tip. Load frame and tip shape calibrations had been determined using traditional DSI methods, as discussed previously. Additionally, blind reconstruction was used to generate a material-independent tip shape area function\textsuperscript{20}). Eleven load-penetration curves were performed with maximum loads of between 43 $\mu$N and 944 $\mu$N. Several different fitting routines were used to fit the unloading curves. Power law fits using commercially available statistics software also did not converge. Power law fits using the commercial DSI system software also did not converge for any of the data sets, but values for $S$, $h_c$, and $E$ were output based on the values of the fitting parameters for the final iteration. In Figure 9, the resulting curve fits are shown along with fits based on linear fits to logarithmic data, neither of which were good representations of the unloading curves. Note that for nonlinear power law fits, the power law exponents were all greater than 2, ranging from 2.2 to 2.7. Similar results were found by Briscoe and Sebastian. Using the commercially available statistics software, a smooth spline fitting routine was then used, which yielded excellent fits to the unloading data, as shown in Figure 10. Besides producing a good fit to the data, this routine also produced a direct calculation of the first derivative of the fitted curve corresponding to each data point, as shown in Figure 11 as a function of load. For both the higher $P_{\text{max}}$ and lower $P_{\text{max}}$ indents shown in this figure, the slope of the upper
portion of the unloading curve does in fact remain relatively constant. Using the spline fitting routine, however, yields values of $S$ that are not dependent on a human judgement, such as deciding on the number of points to include in a linear fit to the upper portion of the unloading curve.

Fig. 9: An example of a DSI unloading curve for the BCB polymer and two corresponding curve fits: (a) a power law fit using the commercial DSI system software that did not converge and (b) a fit based on linear regression of logarithmic data.

Fig. 10: Examples of DSI unloading curves for the BCB polymer along with the smooth spline curve fits for two different loads.

To compute values of $E$, values of $S$ determined using the smooth spline fits were used to calculate $h_c$ using Equation 5 with $\varepsilon = 0.75$. (Note that because power law fitting was not used, this choice of $\varepsilon$ here is somewhat arbitrary, i.e. it is not related to a power law exponent, $m$ such as in Table 1.) Those values of $h_c$ were then used to determine $A$ from the tip area function determined by indenting fused silica and that determined from blind
reconstruction\textsuperscript{20}). These two sets of modulus values are plotted in Figure 12a as a function of \(h_c\). In both cases, these modulus values do not show any dependence on \(h_c\), in contrast to the \(E\) values plotted in Figure 12b, which correspond to the non-convergent power law fits of the DSI system software. Of course, the uncertainties in the modulus values of Figure 12 are still under investigation, but the use of the non-convergent power law fits appears to be a possible source of artificial increases in \(E\) with decreasing \(h_c\).

![Figure 11: Plots of the first derivative of the smooth spline curves in Figure 10 as a function of load for two different loads.](image)

Averaging the modulus values in Figure 12a, \(E = (3.5 \pm 0.3)\) GPa using the blind reconstruction area function, and \(E = (4.1 \pm 0.2)\) GPa using the area function determined from the indentation of fused silica. Both of these values are slightly higher than the room-temperature tensile modulus of 2.9 GPa reported by the manufacturer. As discussed previously, however, the occurrence of creep during unloading affects the determination of \(S\) for polymeric materials. The indentations of the BCB polymer were performed using a load-hold-unload cycle, so an effort was made to characterize the indentation creep response from the 10 s portion of the cycle in which the load was held relatively constant at \(P_{\text{max}}\). The change in tip displacement, \(h\), with time, \(t\), during the hold cycles of each indentation data set was fit using a functional relationship of the form \(\Delta h \sim (\Delta t)^n\), and the exponent, \(n\), ranged from 0.38 to 0.57.

To apply this crude characterize of creep response to the unloading curve, the convolution integral relating the time-dependent strain, \(\varepsilon(t)\), to creep compliance, \(J(t)\), and the time rate of change of stress, \(d\sigma/dt\), was used\textsuperscript{32}. 
\[ \varepsilon(t) = \int J(t - \tau) \frac{d\sigma}{d\tau} d\tau \]  

(11)

A uniform stress, \( \sigma(t) = P(t)/A(t) \), was assumed, \( \varepsilon(t) \) was taken to be related to \( \Delta h/h \), and \( J(t) = k_1 t^n \) was assumed, where \( n \) was determined from the creep response during the hold cycles. Of course, the actual stress and strain distributions are highly non-uniform, and the hold portion of the indentation tests deviates significantly from a conventional creep test. Also, contrary to the Boltzmann Superposition Principle, the integral was not evaluated for all elements of the loading history, but rather just the unloading portion of the indentation loading cycle. Finally, whether indentation behavior of polymers follows the assumptions of linear viscoelasticity is not clear. Thus, this viscoelastic analysis is only a rough estimate of the actual viscoelastic behavior.

Using this approach, the estimated viscoelastic creep was removed from the unloading data for each data set, and the smooth spline fits were again applied. The resulting values of \( S \) were approximately 3% to 6% lower that the \( S \) values determined previously for the total unloading curves. Using the blind reconstruction area function, the corresponding reductions in \( E \) values were between 3% and 5%, resulting in an average value of \( E = (3.4 \pm 0.3) \text{ GPa} \). A more rigorous analytical or numerical approach might result in an even larger influence of creep response on the values of \( E \) determined from indentation unloading curves. Another potential effect could be that of hydrostatic stress, which has been shown to increase the apparent modulus value of some polymers\(^{25}\). Research in these areas is currently being pursued.

More recent applications of DSI to measuring polymer properties have focused on indentation creep tests and also on dynamic indentation testing. Attempts have been made to model indentation creep and stress relaxation data using standard linear solid models from which values of \( E \) can be extracted from fits of the model predictions to the experimental data\(^{33-36}\). Generally, these attempts have met with mixed success, most likely because the experimental aspects do not conform to the assumptions of the analysis. For example, the modulus values measured for instantaneous applications of load or displacement are often much lower than expected, because a finite loading or displacement rate is used that does not simulate a step function. In one case, the contact area was assumed to be constant during viscoelastic creep\(^{36}\), which is not the case as discussed by Unertl\(^{30}\). The simplicity of the models used compared to the actual polymer behavior and assumptions regarding stress and strain could also lead to inaccuracies in the measured polymer properties.
Dynamic indentation testing has recently been developed for several indentation systems\textsuperscript{28,37). This method is similar to dynamic mechanical analysis (DMA) in that a sinusoidal input is applied and the output is monitored. For linear viscoelastic behavior, the output is also sinusoidal but can lag the input signal. The in-phase components of stress and strain can be used to determine the storage modulus, $E'$, and the out-of-phase components are related to the loss modulus, $E''$. The ratio of these moduli, $E''/E'$, is referred to as the loss tangent or $\tan \delta$. 

---

**Fig. 12**: Plots of modulus measured for the BCB polymer from DSI as a function of contact depth. The modulus values in (a) were determined using smooth spline fits and the two different area functions indicated, while the modulus values in (b) were determined using non-convergent power law fits and the indentation area functions.
In DMA, the input oscillation is normally displacement or strain and the output is load or stress. In DSI, an oscillatory load is applied and the displacement is monitored\(^{28}\). The in-phase component of the contact stiffness is then used to calculate \(E'\) and the out-of-phase component is used to calculate \(E''\). The equations used for these calculates are analogous to Equation 2, replacing \(S\) with either the in-phase or out-of-phase component of the contact stiffness and replacing \(E_r\) with either \(E'\) or \(E''\)\(^{37}\). Unlike DMA, however, stress localization and large strain gradients are associated with DSI, and the assumptions of linear viscoelasticity have yet to be checked for dynamic indentation testing. Thus, caution should be used in evaluating the resulting dynamic modulus values.

**Nanoindentation Using SPM Systems**

Three classes of systems can be categorized as SPM nanoindentation systems. In the first type of system, AFM and DSI systems are integrated with the DSI transducer-tip assembly mounted in place of the AFM cantilever probe, and indentation is controlled by the DSI system. Contact mode scanning of the sample is also permitted using the DSI force signal in the AFM feedback loop. However, the contact force applied during scanning is orders of magnitude larger than typical AFM contact mode forces, and thus scanning with these systems can severely deform and damage polymeric materials. Also, indentation tips typically have larger tip radii compared to AFM cantilever tips, thus limiting the resolution of resulting images. However, the ability to scan the surface to find a suitably smooth area or particular surface features is often very important to nanoindentation studies. In terms of indentation testing, the system acts as a traditional DSI system, as has been discussed at length. In fact, this type of system was used in all of our studies of DSI that were previously discussed.

The second type of instrumentation is the interfacial force microscope (IFM) developed at Sandia National Laboratory by Houston and co-workers\(^9\). In terms of nanoindentation of polymers, the IFM has a number of advantages over DSI systems. The main advantage is the ability to indent using rigid displacement control, which allows tip penetration into the sample to be controlled with subnanometer resolution and ensures that the load frame compliance is zero, thus simplifying analysis\(^{11}\). This attribute results in much lower applied forces compared to DSI such that true nanoscale spatial resolution can be achieved. Typically, electrochemically etched tungsten wires have been used with the IFM, and the tip radii have been determined using scanning electron microscopy to range from 70 nm to 500 nm\(^{10,11}\). Penetration of the tip into the sample is normally much less that the probe tip radius, and thus
a Hertzian analysis, which is a special case of Equation 1 with \( m = 1.5 \) and \( \alpha = 4E_r \sqrt{R/3} \), can be used to evaluate the indentation response and calculate sample modulus. In many cases, no hysteresis is observed between loading and unloading curves, indicating pure elastic deformation, such that either the loading curve or unloading curve can be analyzed\(^{11,38}\). Reasonable modulus measurements for polymers using the IFM have been achieved, although similar questions regarding uncertainties as those discussed for DSI remain. Also, in one study, the IFM was used to perform an indentation stress relaxation experiment on PTFE, in which a 15% decrease in the applied load was observed in less than 3 s of the application of a constant displacement\(^{39}\). The current scanning capabilities of the IFM, which include only contact mode, are limited compared to the AFM, but the contact mode scanning forces are significantly less than the integrated DSI-AFM system. Also because of the larger tip radii, image resolution for the IFM is not as good as for AFM.

Quantitative indentation methods have also been developed recently using AFM cantilever probes that can also be used for imaging purposes\(^{6}\). Original attempts to use the AFM as an indentation device were done using a privately built system designed specifically for nanoindentation and having limited scanning capabilities\(^{40}\). More recent efforts, particularly those focussed on indenting polymers and biological materials, have utilized commercial AFM systems and commercial cantilever probes\(^{41-46}\). The advantage of using commercial systems is the potential to combine nanoindentation testing with robust, high resolution imaging capabilities. In fact, through recent developments by one AFM vendor, some commercial AFM systems offer the capability to switch back and forth between tapping mode imaging and indentation mode. However, because commercial AFM systems have not been specifically designed for indentation testing, a number of instrumental uncertainties severely limit their uses as nanoindenters\(^{7}\). Despite these limitations, successful nanoindentation studies of polymers have been reported. These successes include studies of a wide variety of commercial polymeric materials\(^{42}\), interphases in fiber-reinforce polymer materials\(^{8}\), crystalline regions in polyethylene and polyethylene blends\(^{46}\), aerogel powder particles\(^{45}\), polymer thin films\(^{43,44}\), and differences between the surface and interior of a polydimethylsiloxane\(^{6}\). In the AFM indentation studies performed to date, however, the measurements of modulus have been either relative measurements or measurements with large uncertainties due to the use of idealized tip shapes and nominal spring constants. Further, Hertzian or Sneddon models have been utilized in all of these studies, such that viscoelastic behavior has been ignored.
In our recent studies, the BCB polymer used in the previously discussed DSI study was indented using a diamond-tipped stainless steel AFM cantilever probe. The spring constant, \( k_c = (120 \pm 10) \text{ N/m} \), was measured by the manufacturer using a technique in which the force applied to a mica sample is measured by a digital microbalance as a function of measured tip deflection. Blind reconstruction was then used to estimate the tip shape area function as described elsewhere\(^{20}\). Load frame compliance for the AFM cantilever was performed by pushing on a sapphire sample. In this case, because the sample contact stiffness is much larger than the probe spring constant, no tip penetration occurs and the measured force-displacement response is characteristic of the given AFM probe and the particular operating conditions. The probe response can be removed from the force-displacement responses measured on the polymer sample so that only the force-penetration response remains\(^6\)\(^-\)\(^8\). Eight load levels ranging from 1.4 \( \mu \text{N} \) to 13.7 \( \mu \text{N} \) were used to indent the BCB sample. For each load level, 10 force curves were obtained on a smooth sapphire sample, five directly before and five directly after indenting the BCB sample, using the same probe and operating conditions. As with the DSI study of the BCB polymer, smooth spline curves were fit to the unloading curves after unsuccessful attempts to get good fits using nonlinear power law fits and linear fits to logarithmic data.

Results from this study are shown in Figures 13 and 14. In Figure 13a, an AFM image of the plastic impressions remaining in the BCB material after indentation is shown. Representative load-penetration curves corresponding to each of the eight load levels are shown in Figure 13b. The resulting values of modulus are plotted as a function of penetration depth in Figure 14. These values are generally much higher than the bulk modulus value determined from tensile testing, and are also high compared to the values determined in the DSI study. Again, these high modulus values were probably affected by viscoelastic creep, which as previously discussed can increase the initial slopes of the unloading curves. Unlike the DSI study of the BCB polymer, a hold segment was not used between loading and unloading. Also, the stress and strain distributions for the sharper, less ideal AFM tip might have been more severe compared to the Berkovich tip used in the DSI study. Thus, the effects of viscoelastic creep during unloading might have been worse for the AFM indentation study compared to the DSI study, resulting in the higher modulus values.
Conclusions

In this paper, the use of depth-sensing indentation methods to measure the elastic modulus of polymeric materials was reviewed. Included in this review were discussions of traditional analyses of load-penetration data, the use of depth-sensing indenters and scanning probe microscopes to measure the nanoscale mechanical response of polymers, and the associated uncertainties and limitations of various nanoscale indentation measurements. The application of these methods to polymers often leads to inaccurate measurements of elastic modulus. For quasi-static indentation, viscoelastic behavior affects the shape of the unloading curve, resulting in modulus values that are high relative to bulk measurements. Attempts to characterize creep behavior during indentation experiments suffer from system limitations.
(e.g., relatively slow rates of loading compared to the necessary step loading) and potentially inappropriate assumptions regarding the stress and strain distributions under indentation loads. While additional complications might arise, dynamic indentation testing has the potential to alleviate many of the problems associated with quasi-static indentation testing. However, a rigorous analysis of dynamic indentation behavior of polymers, particularly with regard to whether or not linear viscoelasticity holds, has not been reported. Advancement of these methods toward quantitative characterization of polymer properties will require material-independent calibration procedures, polymer reference materials, advances in instrumentation, and new testing and analysis procedures that account for polymer rheological behavior.

![Plot of modulus measured for the BCB polymer from indentation with an AFM cantilever as a function of contact depth.](image)

Fig. 14: Plot of modulus measured for the BCB polymer from indentation with an AFM cantilever as a function of contact depth.

**References**

23. N. M. Jennett, personal communication (June 1, 2000).