Spatial Dimensions in Atomic Force Microscopy: Instruments, Effects, and Measurements

Ronald Dixson*, Ndubuisi Orji

Engineering Physics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899

Ichiko Misumi

Nanoscale Standards Group, National Metrology Institute of Japan, AIST, Tsukuba, Japan

Gaoliang Dai

Nanodimensional metrology department, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany

Abstract

Atomic force microscopes (AFMs) are commonly and broadly regarded as being capable of three-dimensional imaging. However, conventional AFMs suffer from both significant functional constraints and imaging artifacts that render them less than fully three dimensional. To date a widely accepted consensus is still lacking with respect to characterizing the spatial dimensions of various AFM measurements. This paper proposes a framework for describing the dimensional characteristics of AFM images, instruments, and measurements. Particular attention is given to instrumental and measurement effects that result in significant non-equivalence among the three axes in terms of both data characteristics and instrument performance. Fundamentally, our position is that no currently available AFM should be considered fully three dimensional in all relevant aspects.

*Corresponding Author: Please address all correspondence to Ronald Dixson at Ronald.dixson@nist.gov
1. Introduction

Slightly more than 30 years has passed since the invention of the atomic force microscope (AFM) by Binnig et al. [1], but during this time AFM technology has undergone rapid evolution and a proliferation virtually unprecedented in the four-century history of microscopic imaging techniques. AFM is now used in fields ranging from materials science to cell and molecular biology, and for applications ranging from the dimensional metrology of lithographically patterned nano-structures to the localized measurement of the mechanical properties of soft polymers.

The variants of AFM imaging modes and contrast mechanisms are now far too numerous and diverse for most practitioners to have even a modest working knowledge of all. Standardization hasn’t always kept pace with this rapid evolution, and the development of consistent nomenclature has often been lacking. The underlying motivation for this paper is specific absence of any clear conventions for describing instruments, imaging modes, and measurement types in terms of spatial dimensions.

The absence of such consistent conventions and terminology does not necessarily pose much difficulty for those with experience in AFM metrology or instrument design. However, AFMs are now used so widely for such a diversity of applications that many AFM users may lack such experience or have limited knowledge of the underlying physics of AFM image formation and data structure. Nanoparticle metrology or the measurement of three-dimensional integrated circuit features are examples of applications in which the performance limitations or differing information content among the three axes of an AFM could result in confusion or misinterpretation of data. Consequently, a clear taxonomy describing these differences would be useful.

Atomic force microscopes and AFM images are commonly regarded as three dimensional. However, this is frequently stated with little or no elaboration as to the specific meaning or measurement context. At a cursory level this might seem appropriate, since virtually all AFMs are capable of generating three-dimensional images. However, this represents a significant over-simplification since there are more characteristics of both instruments and data that should be considered beyond the apparent dimensions of an image.

Upon closer consideration, it is not really surprising that there are inconsistencies in the dimensional nomenclature used to describe AFM. Even with respect to everyday vocabulary, there is some ambiguity about the meaning of three dimensional. For example, two dimensional drawings can be rendered in perspective in order to create the impression of depth, and such drawings are often described as being three-dimensional despite being rendered in a planar medium.

Human depth perception is also commonly thought of as 3D, but it is fundamentally based on the computational fusion of two distinct fields of view taken from different observation angles. In isolation, each of these fields is essentially a two-dimensional map of color and intensity reaching the retina of an observer’s eye. Thus, to the extent that human vision is appropriately regarded as three-dimensional, this 3D character of the overall process does not reside in the eyes or the images obtained by each eye, but it is the result of additional data processing on multiple inputs – each of which is less than three dimensional in nature. A closely related example, returning to the realm of microscopy, is stereoscopic or 3D scanning electron microscopy (SEM).

In some circumstances, a single SEM image in isolation may afford the visual impression of a topographic surface, but it is actually a map of the electron intensity reaching a detector as a function of the position of the incident electron beam. This map is usually generated using a two-dimensional raster scan of the beam.
location. For a single image, there is no generally applicable method for converting the measured electron intensity at a given pixel into an equivalent surface height. Consequently, at least with respect to spatial dimensions, a single SEM image should be regarded as less than three dimensional. Analogously to depth perception, however, it has been possible since the early 1970s [2-5] to combine multiple SEM images of the same specimen taken under different conditions in order to extract spatial information for an additional axis. 3D SEM methods, often referred to as “tilt SEM,” that involve acquiring multiple images with a known difference in the tilt angle of the incident beam with respect to the sample surface, have been developed more recently.[6-8]

AFMs suffer from both significant functional constraints and imaging artifacts that render them less than fully three dimensional in some respects. Some particularly important examples of this are the limitations resulting from the tip shape and the feedback/servomechanism axis of the tip relative to the sample – which was constrained to the z-axis in early instruments [1]. These limitations meant that it was not possible to measure features with arbitrary geometry (e.g., reentrant structures) or with equivalent performance in all axes (e.g., scan-rate related distortions).

Critical dimension AFMs (CD-AFMs) [9], which are AFMs with dual-axis tip scanning control, did much to improve the situation with the addition of lateral tip-sample position control and feedback in the fast scan axis. This made it possible to image the sidewalls of near-vertical and slightly undercut structures. Such structures are particularly common in the semiconductor industry; since the performance of many devices is directly related to the linewidth of the structures, the term critical dimension (CD) became synonymous with linewidth. In a broader context, CD can also be defined as the smallest measurable dimension that affects performance.

Yet even current state-of-the-art CD-AFM instruments do not have three-axis equivalence in terms of surface sensing, displacement accuracy, or tip position control. A noteworthy example of confusion resulting from the absence of accepted dimensional terminology for AFM is that CD-AFM has been described as being both two and three dimensional by different researchers [9,10]. For the remainder of this paper, we will refer to AFMs other than the CD-AFM as conventional AFM.

To date, there is no widely accepted dimensional nomenclature for AFM data, measurements, and instruments, and this leads to inconsistency and confusion in the AFM user community. The general goal of this paper is to describe the dimensional characteristics of AFM measurements from the different perspectives of scanning and data sampling strategy, probing tip, tip sample interaction, and measurands. In particular, the focus is on instrumental and measurement effects that result in significant non-equivalence among the three spatial dimensions. While our emphasis is on topographic measurements, this framework could be extended to non-topographic contrast, such as in AFM measurements for mechanical, electrical, and magnetic characterizations. Fundamentally, our position is that no current AFM should be considered fully three dimensional in all relevant aspects.

The conventions and nomenclature that we propose in the following sections are also summarized in Table 1. For this purpose, it is necessary to categorize AFM implementations, tip types, and analysis methods as being 1D, 2D, etc. Our intention in this table, however, is less to make rigid declarations than to differentiate between methods, data types, etc. and to underscore that caution must be used in interpreting AFM results in dimensional terms. For this reason, we have chosen to avoid using half-dimension characterizations.
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>1D</th>
<th>1D+</th>
<th>2D</th>
<th>2D+</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property of AFM or data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Data Structure** | --- | --- | Profiles 
(x,z data) | --- | AFM (single height channel or grey scale image) |
| **Data Content** | AFM (single height channel or grey scale image) | --- | AFM (height & lateral channel) | --- | AFM (height & two lateral channels) |
| **Tip Geometry** | AFM (conical, pyramidal) | AFM, CNT tips | Flared CD tips, CNT tips | Flared CD tips | |
| **Tip Position Control/Feedback** | AFM | Step-in mode, balance beam system, LFM, FFM | CD-AFM* | --- | --- |
| **Measurement Equation: Output** | Height, pitch (linear array), pitch (x-y pillar array – single value), linear feature width (single value), side wall angle, \( R_q \) – rms roughness | --- | Edge contours, pitch (x-y pillar array), profile PSD, pillar width | --- | Sidewall angle contours, areal PSD |

Note: CD-AFM* refers to CD-AFM on the feedback system.
There are aspects of physics and mathematics where the concept of a partial dimension can be rigorously defined and therefore has a clear meaning. But with regard to the topics addressed in this paper, we are dealing with instruments and measurements for which there would be neither an exact nor consistent meaning of such designations. Rather, it would function as an indicator of additional dimensional information that is either required or available – or a circumstance in which the performance or constraints of one axis differ dramatically from the others. Consequently, for purposes of discussion in this paper, we have chosen to use the designations 1D+ and 2D+ to indicate such cases, rather than make an arbitrary attempt to define partial dimensions.

The paper is organized as follows: section 2 considers both the content and the structure of AFM data; the section 3 discusses the performance and limitations of various types of AFM instruments – specifically with regard to the tip geometrical, tip sample interaction and the position control/feedback; and finally section 4 focuses the dimensional nature of typical AFM measurements.

2 AFM Data: Structure and Content

Virtually all modern AFM instruments typically present data as a seemingly three-dimensional (3D) image, and most systems now include extensive image rendering software. Beyond this apparent capability, however, lies a spectrum of data structures and differences in the spatial information content.

In the earliest AFM systems [1], as well as the scanning tunneling microscopes (STMs) that preceded them, the data sets were presented as a series of profiles and this presentation format is still available in many AFM instruments, even in the state-of-the art CD-AFMs. To illustrate some of its constraints, Fig. 1 shows an example of an image taken using a CD-AFM. The data are shown as a series of profiles and rendered in wire frame format. As is the case in most AFMs, the data were acquired using a type of raster scan, so each profile shown represents a single trace along the fast scan (X) axis taken at a fixed position in the slow scan (Y) axis. Many conventional AFMs are also restricted to square scan areas and pixels of uniform size. CD-AFMs, in order to measure near-vertical sidewalls, do allow for both non-square image sizes and non-uniform pixels in the fast scan axis. In the slow scan axis, however, CD-AFMs are still restricted to a fixed number of scan lines uniformly distributed across the image.

The CD-AFM image in Fig. 1 shows a transition from dense to isolated features. This field-of-view includes a central line feature and end points of the two dense flanking lines. Although the CD-AFM tip and feedback/control system is very effective at imaging sidewalls orthogonal to the fast-scan or x-axis, the limitations are apparent on the sidewalls orthogonal to the slow-scan or y-axis. The fundamental constraint of the system is that the tip is essentially blind to the slow scan axis component of the topographic gradient. For this reason, it is common practice in the CD-AFM metrology to align the features of interest to be orthogonal to the fast scan axis, as was done in this image.

Another example is given in Fig. 2, which is an AFM image of a deep circular etch pit - known in semiconductor manufacturing as a through silicon via (TSV). In such a measurement, the circular edge would be the feature of primary interest. However, most currently available instruments use some form of

<table>
<thead>
<tr>
<th>Measurement Equation: Input Quantities</th>
<th>rms roughness</th>
<th>edge contours, pillar width, profile PSD</th>
</tr>
</thead>
</table>

* Includes CD-AFM and PTB VAP-based method.
a raster scan. Consequently, data points are widely distributed over a rectangular sampling region even if the features of interest are more localized within this area.

The CD-AFM image in Fig. 2 was taken with z-axis travel limited using the control software. The circular profile of the edge in the sample plane underscores the limitations of the instrument response: In those regions where the edge is nearly orthogonal to the fast-scan \((X)\) axis, the edge sampling is robust. However, in those regions where the edge is nearly orthogonal to the slow-scan \((Y)\) axis and the tip-sample feedback is essentially blind, the edge sampling is both sparse and not representative of the actual surface gradient. The overlaid black arrows indicate the lateral direction of the surface gradient for three points. Only along the fast-scan axis does this direction coincide with the axis of tip-sample feedback and position control.

**Figure 1.** CD-AFM image of a transition from dense to isolated features, rendered as a series of profiles in the fast-scan axis (the \(X\) axis in this example). The image shows a central line feature and end points of the two dense flanking lines. In addition to the sampling constraints in the slow-scan \((Y)\) axis, there is a more fundamental limitation of the system feedback and tip control which means that the tip is essentially blind to the slow-scan axis component of the topographic gradient.
Figure 2. CD-AFM image of a circular etch pit in silicon. In those regions where the edge is nearly orthogonal to the fast-scan (X) axis, the edge sampling is robust. However, in those regions where the edge is nearly orthogonal to the slow-scan axis, the edge sampling is both sparse and not representative of the actual surface gradient. The overlaid black arrows indicate the lateral direction of the surface gradient for three points. Only along the fast-scan axis does this direction coincide with the axis of tip-sample feedback and position control.

CD-AFMs do at least allow for a greater density of data points along steep sidewalls and in transition regions. However, the vast majority of conventional AFMs using raster scans also use pixels of uniform lateral spacing. This can result in images where features of interest are sparsely sampled and in which there are large numbers of relatively low value pixels in less important regions of the image. Therefore, careful consideration of such profiles suggests the need for a systematic nomenclature to describe the dimensional characteristics of AFM data.

A profile is often thought of as a set of \((x, z)\) pairs. In this sense, it is reasonably regarded as being two-dimensional (2D) in nature. Although, such data is also frequently thought of in terms of a function \(z = f(x)\) and represented by a one-dimensional (1D) array of stored values, the stored value represents the height value for the vertical or \(Z\) axis, and the lateral or \(X\) axis value is determined by the array index. This form of storage is well suited to those instruments for which the lateral spacing of the data is a fixed value during the scan. However, for those systems in which this is not the case, it is also common to store two 1D arrays: one for the lateral axis and one for the vertical axis – which is equivalent to stored \((x, z)\) pairs. Therefore, we suggest that profile data should be regarded as having a 2D data structure, and we will extend these definitions to the third axis.

The most prevalent AFM data storage format is a 2D array of stored values representing the height (Z axis). This is particularly suitable for instruments for which the values of the lateral axes (X and Y) are not explicitly stored and a uniform XY grid is assumed. However, it is increasingly common for the lateral position channels to be stored in additional 2D arrays, and this is equivalent to storing \((x, y, z)\) triplets. Therefore, consistent with the prevailing intuition and usage, we suggest that the most common AFM data format
should be regarded as having a 3D data structure. However, the possible absence of explicit lateral axis information in such data sets means that it may not always be possible to perform meaningful 3D metrology even when using data sets having a 3D data structure. This highlights the fact that the information content of the data should be considered in addition to the structure, and the content may be of fewer dimensions.

Considering profile data – which has a 2D data structure – we suggest that the dimensions of the information content correspond to the number of 1D arrays stored to represent the data. If only a single array is used, then there is no explicit lateral position data stored. We suggest that such data are 1D in terms of content. Consequently, such data would be appropriate only for 1D metrology (i.e. measurements only requiring accuracy in the Z axis, such as step height.)

For the very common case of AFM data stored as a single 2D array of height values, it is certainly true that such a data set can generate an apparently 3D image, and we regard the data structure as being 3D. However, since no lateral position information is explicitly stored, we suggest that the most prevalent format of stored AFM data be considered 1D with respect to information content.

When considered carefully, the importance of this distinction between data structure and information content may seem more practical rather than theoretical. If the performance of an instrument were ideal, then, in principle, explicitly stored lateral position data could be redundant. Note that this would imply constraints on scanning such that the uniform grid was sampled, but the supplemental arrays of lateral data would not add new information.

As a practical matter, of course, there are no ideal instruments. All real AFMs are subject to scanner motion errors – even instruments with incorporated displacement sensors and closed-loop lateral position control. The resulting lateral position errors are not captured in data sets based on the uniform grid assumption. Images rendered from a single 2D array could thus exhibit significant scanner distortion errors.

Instruments with incorporated displacement sensors, such as those used in some metrology applications, usually permit the storage of data sets with lateral channels in addition to the z-axis data. Some such AFM systems, including those developed by the national metrology institutes (NMIs) of some nations, also have Z axis displacement sensors [11-18]. For these instruments, the content of stored data can consist of true (x, y, z) triplets based on the measured position of the tip or scanner rather than the demand position of the controller. Although not universal among AFMs, three-dimensional data in this sense is common.

3 AFM Instruments: Performance and Limitations

3.1 Tip-Sample Geometric Occlusion

The geometrical interaction of an AFM tip with the sample surface is inherently a three-dimensional process. [19-22] However, the nature and severity of tip-related artifacts in AFM images is not equivalent in all three axes. This can result in different performance limitations in the three axes with respect to the relationship between the actual surface and the image.

Although commonly referred to as a convolution of the tip and sample geometries, the imaging process is most appropriately modeled using mathematical morphology [21,22]. The image thus represents a dilation of the surface with the tip shape. This generally has the effect of broadening lines (i.e., positive going features) and narrowing trenches (i.e., negative going features) - a distortion that can be understood as resulting from de-localization of the tip-sample interaction region due to the finite extent of the tip. As illustrated in Fig. 3, the apparent profile measured is the dilated result of the real structure by the tip shape. For the structure has a larger sidewall slope (\(\Phi\)) than the tip half cone angle (\(\theta\)), there is no localized region
near the apex of the tip which is always in contact with the sample surface during imaging. Consequently, as shown in Fig. 3, the apparent width of the measured features (c) and (d) is a function of the height – even if the sidewalls of the feature are vertical.

\[ Sm = St \oplus Sr \]

**Figure 3.** Schematic drawing shows a conventional AFM tip (having a radius of R and half cone angle of \( \theta \)) interacting with a feature with different sidewall slopes \( \Phi \). For the structure with larger sidewall slopes \( (\Phi > \theta) \), the tip-sample interaction point is not localized at the apex, but extends up the shank to the height of the feature.

The tip dilation usually does not alter the apparent height of most structures – at least those that are wider relative to the width of the tip. An important exception to this occurs in the case of the apparent depth of a feature that is too narrow for the tip to penetrate, as clearly illustrated in figure 3.

Recently, there has been research into the mitigation of such height attenuation, and a method to quantify the magnitude of effect has been developed by Misumi et al [24]. This approach specifies both the geometry of probe characterizer and the method of evaluation, which takes into account the \( R_a \) and the mean length or lateral spacing of the roughness profile elements: \( R_{Sm} \). It is possible to estimate the maximum error and thus the uncertainty in \( R_a \) measurements using this method.

This consideration of roughness also highlights the observation that there are circumstances in which even an isolated step height measurement may be biased due to the tip shape. If the upper and lower surfaces of a step have significantly different roughness at spatial wavelengths near and below the tip width, the resulting difference in relative penetration of the valleys would result in different offsets between the actual and apparent average heights of the two surfaces. This general topic has been considered in more depth by Seah [25].

As a whole, conventional AFM tips – even very sharp ones – are better suited for step height measurements than for width measurements. Such tips are effectively used for step height metrology of features that are wide relative to the tip or for measurement of roughness on smooth surfaces (i.e., surfaces for which the
height variation is small relative to the average spatial wavelength). This is evident by the widespread industrial use of conventional AFM for measurements of film thickness and surface roughness on very smooth silicon wafers or polished optical surfaces [26-30].

In terms of the descriptive framework we are developing in this paper, it makes sense to regard conventional AFM tips as less than three dimensional with respect to functionality. Due to their greater suitability for height measurements, we suggest that most conical and pyramidal tips should be considered as 1D with respect to tip geometry. A cylindrical tip can be thought of as a limiting case of a conventional tip with very sharp taper, and it also represents a logical boundary between 1D tips and CD-AFM tips. For this reason, it makes some sense to regard cylindrical tips – such as carbon nanotube tips (CNT) as being more than 1D.

The situation is qualitatively different once CD-AFM tips are considered. These tips are designed to have relatively sharp points or flares aligned with the lateral axes of the instrument. As such, CD tips are well suited to imaging the vertical surfaces (orthogonal to the sample plane) or sidewalls of features.

Figure 4 shows a two-dimensional illustration of the geometrical interaction between a CD-AFM tip and a surface. In contrast with conical tips, the resulting dilation of a feature by a CD tip is largely independent of height (up to the length of the tip). The apparent width bias can be modeled primarily by a constant offset (the tip width), along with smaller effects that vary with the details of each measurement [31].

\[ W_{\text{apparent}} (h) = W (h) + TW_{\text{zeroth order}} + \varepsilon_{\text{higher order}} \]

**Figure 4.** Simplified schematic of CD-AFM tip interacting with vertical feature. The most basic CD tip parameter is the tip width in a simplified model in which an apparent feature width is primarily increased by an additive offset – the zeroth order width – along with much smaller contributions due to secondary or higher order effects [31].

In addition to sensing both horizontal and vertical surfaces, CD tips are capable of accessing some degree of re-entrance or undercut. Typically, the detectable limit on this reentrance is about 10 nm, although it is different for every tip. Current CD-AFM tips have an essentially similar shape in the third axis (i.e., into the page in Fig. 4). In fact, the majority of CD tips are rotationally symmetric (or very nearly so) about the Z axis.

With respect to geometry, therefore, CD tips might be said to be almost three dimensional – where what is meant by this is that the tip geometry supports an interaction region with the sample that is well-localized on the tip for surfaces aligned with all three axes. Nevertheless, its spatial resolution capability – defined the curvature of the local contact area of the tip – is not equivalent in all three axes. The spatial resolution on vertical sidewalls is often better than on horizontal surfaces since the local contact area at the bottom
part of most CD-AFM tips is relatively blunt – approaching flat. However, as CD-AFM tips continue to shrink, there is less difference between the relative sharpness of the tip flares and the tip bottom.

All CD-AFM tips are limited in the ability to image locally recessed features in all three axes. This is analogous to the apparent depth limitation of 1D tips due to the inability to reach the bottom of narrow trenches. Figure 5 illustrates a hypothetical recessed region on the bottom surface of an overhang. Aside from incidental contact with portions of the lateral flares, there are currently no CD-AFM tips capable of imaging such a surface.

**Figure 5.** Illustration of a flared CD-AFM tip and an inaccessible recessed region of the overhang. No current CD-AFM tip is capable of imaging this region.

We suggest that CD-AFM tips should be thought of as 2D+ with respect to the tip geometry and resultant imaging capability. However, that definition would be more stringent than analogous definitions used in larger scale metrology and so could be considered excessive [32].

### 3.1.1 Geometrical Tip Characterization

Characterization of tip geometry is a closely related topic. Qualitatively, a tip characterizer should have spatial dimensional characteristics equivalent to or exceeding those of the tip that is to be characterized. Sharp spikes, which resemble inverted 1D AFM tips are often used to characterize 1D tips. However, to fully characterize a CD tip, some sort of flared or reentrant structure is required, and such characterizers tend to resemble inverted CD tips.

In actual practice, CD tips are typically characterized using a combination of two types of characterizers. The first type has features with vertical sidewalls, such as the feature illustrated in Fig. 4. This type of characterizer is often known as a vertical parallel structure (VPS). If the width of a feature on the VPS has been traceably calibrated, then it can be used to calibrate the width of a CD-AFM tip. The second type of characterizer has undercut or reentrant features capable of imaging the flares of the CD tips. This characterizer type is frequently referred to as a flared silicon ridge (FSR) – since the first characterizers of this type were etched silicon. FSR characterizers have features similar to those of a hypothetical sample that will be discussed in section 3.2 and illustrated in Fig. 8.
A variety of other characterizer designs have been conceived or tested, however. For example, Tian, Qian, and Villarrubia proposed a design for a CD tip characterizer that resembles exploding petals [33]. Orji and Dixson have collaborated on the evaluation of a variable width comb characterizer and compared the performance with the conventional characterizers [34].

Considerable work has also been done on 'blind reconstruction' techniques for tip characterization [21,22, 33]. Qualitatively, these methods work because the shape of an AFM tip must be consistent with having generated the sharpest apparent features in an image. This makes it possible to estimate an outer bound on the tip shape. Although this is true for any AFM image, these techniques tend to be more effective when used with tip characterizer samples – because tip characterizers are usually known to have very sharp features.

Geometrical tip characterization is very important for AFM measurements, but there are also very important non-geometrical aspects of the tip-sample interaction. This is considered further in the next section.

3.2 Tip-Sample interaction

In addition to the geometrical dilation of the AFM tips on structures, as described above, the tip sample interaction in AFM measurements also sets a fundamental limit with respect to the effective number of dimensions that are involved. Differences in the tip-sample interaction among the axes can result in imaging artifacts, and these in turn can reduce the dimensional content of the image regardless of the data structure or apparent data content.

AFM cantilevers have anisotropic stiffness, and the bending stiffness is typically much lower than the torsional stiffness. For this reason, virtually all dynamic AFM modes involve cantilever oscillation in the vertical axis. In most AFM modes, the measured signal from the cantilever is the vertical deflection or vertical oscillation amplitude. The cantilever response is thus directly sensitive only to the Z-axis component of the tip-sample interaction. From this perspective, an AFM cantilever would be regarded as a 1D force sensor. For measurements of complex 3D nano-structures, which will involve edge transitions and sidewall regions, the tip-sample interaction force and the force gradient will exhibit much larger variations than for measurements of near-planar surfaces. Consequently, there can be non-geometrical, surface-tracking distortions in the apparent 3D form of a surface – particularly in edge regions [35,36]. Another example is that for high-compliance probes, such as carbon nanotube (CNT) tips, bending of the tip when encountering edges is not always negligible, and this can lead to measurement biases [37].

In order to develop a more quantitative model for these effects, we consider the AFM measurement scenario illustrated in figure 6 – which involves imaging of an edge and corner region. For AFM measurements in ambient atmospheric conditions, there will typically be a thin film of vapor water and/or hydrocarbon present on both the tip and sample [38]. This is illustrated as a damping layer in figure 6. While the tip sample interaction forces are very complex, all forces must either be conservative or dissipative. This suggests that the first step is to partition the tip sample interaction into a conservative force $\vec{F}_{c,ts}$ and a dissipative force $\vec{F}_{d,ts}$. The conservative force $\vec{F}_{c,ts}$ can be further partitioned into a long range attractive van der Waals force $\vec{F}_{vdW}$ and a contact repulsive force $\vec{F}_{repulsive}$. While some AFMs operated under vacuum conditions are capable of imaging using only the attractive tip sample force, most AFMs for dimensional metrology operate in a regime that is primarily sensitive to the repulsive force.
Figure 6. Model of the tip sample interaction in a 2D edge/corner measurement scenario. The interaction forces consist of a conservative force $F_{c,ts}$ and a dissipative force $F_{d,ts}$ as shown in (a); The conservative force $F_{c,ts}$ depends on the tip sample distance $d_{ts}$, and it is in the normal direction of the local tip sample interaction surface as shown in (b). Only the $z$-component of the force $F_{c,ts}$ directly affects the tip oscillation which is typically in the $z$-direction.

Our next step is to consider a simple case where the tip-sample system may be regarded as a sphere interacting with a flat surface. We model the tip sample interaction force using the Derjaguin-Müller-Toporov (DMT) theory [39], which is suitable for describing systems with low adhesion and small tip radii. This means that the AFM measurement scenario illustrated in figure 6 can be modeled as a spring-mass-damper system [40] and represented by the appropriate equations:

$$\frac{d^2z}{dt^2} + \frac{\omega_0}{Q_{cant}} \frac{dz}{dt} + \omega_0^2 z = \frac{1}{m} (F_d + F_{ts}) \tag{1}$$

and

$$F_{ts,z} = \begin{cases} 
- \frac{HR}{6d_{ts}^2} \cos \theta & (d_{ts} \geq d_i) \\
- \frac{HR}{6d_{ts}^2} \cos \theta - \frac{m_0 \omega_0 \cos^2 \theta + m_0 \omega_0 \sin^2 \theta}{Q_{norm} + Q_{lat}} \frac{dz}{dt} & (d_i > d_{ts} > a_0) \\
- \frac{HR}{6a_0^2} + \frac{4}{3} E^* \sqrt{R} \left( a_0 - d_{ts} \right)^2 \cos \theta & (d_{ts} \leq a_0)
\end{cases}$$

where $Q$ is the quality factor of the cantilever system, including the damping effects both of the cantilever itself and of the squeezed air film; $F_d = kA_d \cos \omega t$ is the driving force insert by the cantilever shaking piezo; $R$ is the tip radius; $H$ is the Hamaker constant; $a_0$ is an intermolecular distance which can be determined as $\sqrt{H/24 \pi \gamma}$ with $\gamma$ being the surface energy; $d_{ts}$ is the tip sample distance; $\theta$ is the inclination of the measured surface; $d_{ls}$ is the thickness of the damping layer; $Q_{norm}$ and $Q_{lat}$ are quality factors applied for describing the energy dissipation due to the water layer; and $E^*$ is the effective young’s modulus of the tip sample system:

$$\frac{1}{E^*} = \frac{(1-v_t^2)}{E_t} + \frac{(1-v_s^2)}{E_s} \tag{2}$$

where $E_t$, $E_s$ are the Young’s moduli, and $v_t$, $v_s$ are the Poisson’s ratios associated with the tip and sample materials, respectively.
Equation 1 can be solved numerically using an Ordinary Differential Equation (ODE) solver. This allows for the simulation of tip sample interaction curves, and examples are shown in Figure 7. In Fig. 7a are two examples of simulated approach curves, applicable to AFM imaging modes using amplitude modulation, which show the cantilever oscillation amplitude (A) as a function of tip-sample distance (d₀). The two simulated curves represent two different measurement configurations: The first represents a tip approaching a horizontal surface (i.e., θ = 0°), and the second represents a tip approaching a sloped sidewall – using θ = 75° as an example. The parameter values used for these simulations were taken either from published results [40,41] or from typical measurement conditions. The sample material is assumed to be SiO₂ with E_s=75 GPa, v_s=0.17, γ=0.031 J/m² and H = 6.4 × 10⁻²⁰ J; the damping layer is assumed to be d_t = 0.5 nm, Q_norm= 10 and Q_lat = 0.1; and the AFM tip is assumed to be silicon with E_t = 130 GPa, v_s=0.27. For the tip-cantilever system, we assume a tip radius of R=10 nm, Q=200 and ω₀=ω=2π × 330 kHz. The free-oscillation amplitude is 11 nm, and the set point amplitude for measurement is taken as 50 % (i.e., 5.5 nm).

As Figure 7(a) illustrates, the approach curves for the two different configurations are significantly different. This can be attributed to differences in the tip-sample interaction force and the force gradient between the case of a horizontal surface and a sloped sidewall. In typical operation, for AFM modes based on amplitude modulation, the amplitude during imaging is held to the set point value, and this defines the detected position of the surface. Differences in the approach curves between the two configurations effectively result in different definitions of surface detection between the two cases. This is illustrated in Figure 7(a) as the offset between the points labeled as P₀ and P₁ – which is almost 5 nm. However, this example is an extreme case and not representative of typical measurement conditions.

To better understand more realistic scenarios, Fig. 7(b) shows the results of additional simulations are represented by a plot of the effective surface detection point as a function of the sidewall angle. The plot, which includes the P₀ and P₁ results, shows results for sidewall angles ranging from 0° to 75° in increments of 5°. For surfaces with sidewall angles less than 20°, the detection point exhibits very little dependence upon the angle. However, for steeper sidewalls, the dependence is more significant. This means for AFM measurements of 3D structures, there is the potential for measurement biases resulting from non-geometrical tip effects – in addition to the purely geometrical dilation of the surface features.

In process metrology for semiconductor manufacturing, the sidewall angles are typically in the range between 80° and 90°, so, following the model developed here, it seems likely the magnitude of the potential bias for such measurements would be even stronger than shown in Fig. 7(b). Nevertheless, these simulations underscore the fact that non-geometrical effects do contribute to the tip-sample interaction, and care should always be taken to perform both system calibrations and application measurements under as similar measurement conditions as possible.

It can be useful to introduce the concept of an effective tip geometry, as has been the subject of prior work of some of the authors [42,43]. The effective tip geometry included both the actual geometry of the tip and the effect of the tip sample interaction, and these two contributions are importantly different. The actual tip geometry is independent of the sample and the measurement conditions, excluding the possible effect of tip wear. In contrast, the tip-sample interaction can depend significantly upon the sample and measurement conditions. Consequently, for measurements of different 3D nanostructures (e.g., different materials, different damping effects, different water layer thickness), the effective tip geometry may be different even if the actual tip geometry has not changed. This is one potential source of so called higher order tip effects[31].
Figure 7. (a) Simulated tip sample interaction curves showing the oscillation amplitude as a function of tip sample distance for two measurement configurations: $\theta = 0^\circ$ and $\theta = 75^\circ$. In typical AFM operation, the amplitude is held to a set point value, such as 50% of the free space amplitude shown by the horizontal dashed line. For the two configurations shown, the effective surface detection occurs at two different tip-sample separations – shown by the points $P_0$ and $P_1$; (b) simulations showing the dependence of the effective detection point as a function of sidewall angle for angles ranging from 0° to 75° in increments of 5°.

While the simulation results shown in Figure 7 are illuminating, the model on which these results were based is highly simplified. Treating the tip-sample system as a sphere-flat interaction is not a realistic geometry in most cases. For practical measurement of 3D nanostructures, the tip-sample geometry will usually be more complex. The illustrations in Figure 8 show examples of more realistic geometries for typical CD-AFM measurements. In these cases, the tip-sample interaction force and force gradient as well as the damping effects of both the squeezed air film and of the vapor water layer will be very different in the three different regimes illustrated. Consequently, there is the potential for differences in the effective surface detection point, and this could result in an unknown variable bias in actual CD-AFM measurements.

Figure 8. A schematic diagram that illustrates the potential for differences in the tip-sample interaction during a CD-AFM measurement of a complex 3D structure at its (a) top; (b) sharp edge-corner; and (c) neck regions.
Based on the results shown in Fig. 7, it seems reasonable to speculate, as suggested by the schematics in Fig. 8, that regions of significant surface curvature (i.e., highly variable sidewall angle) are likely to exhibit the greatest bias due to non-geometric tip effects. From the perspective of tip width calibration, this would be generally favorable because CD-AFM tip width is typically calibrated using VPS-type structures with very uniform vertical sidewalls. However, the tip flare radius and edge height – often known as the vertical edge height (VEH) – are typically characterized using FSR-type or sharp overhang structures with similar geometry to that depicted in Fig. 8. The possibility that such characterizations are significantly influenced by non-geometric tip effects is a question that would benefit from further investigation.

### 3.3 Signal detection, position control and measurement strategy

The function and performance of the tip-sample feedback and position control is partially connected with the tip geometry and the resultant capability of a given tip. At least to some extent, the tip geometry and resulting tip-sample interaction affects the error signal generated through the cantilever deflection and thus the performance of the feedback loop. More importantly, the functionality and performance of the tip position control determines the extent to which an AFM system can exploit the geometric capability of a given tip.

Early AFMs only detected cantilever deflection along the z-axis, and the tip position control and feedback was also limited to the Z-axis. In this sense, such systems were one dimensional. Since the conical and pyramidal tips commonly used on such instruments are geometrically incapable of imaging vertical or reentrant sidewalls, the one-dimensional tip control was not a significant limitation. Consequently, investigators such as Banke et al. [30,44] and Ukraintsev et al. [45,46] have referred to such instruments as 1D-AFM.

Lateral force microscopy (LFM) and friction force microscopy (FFM) augment 1D-AFM with the detection of cantilever deflection in lateral axes [47,48]. However, the feedback and tip position control is still limited to the Z axis. Such methods are more than 1D, but not fully 2D.

Dimensional information from two axes is used in CD-AFM – which incorporates lateral tip position control with tip geometries capable of imaging vertical and reentrant sidewalls. Although CD-AFM is now often referred to as “3D,” the pioneers of the field considered it to be 2D. Martin and Wickramasinghe developed the approach that became the first generation of commercially available CD-AFM. In one of their initial papers describing CD-AFM, the method is also referred to as 2D AFM [9]. It is also worth noting that in the same paper conventional AFM is characterized as 1D AFM – consistent with the nomenclature we are suggesting here. In addition to the technology on which the first commercially available CD-AFM instruments was based, Nyyssonen et al. developed an alternative approach for AFM metrology of steep-sidewall structures at about the same time [49,50]. They also characterized their method as 2D AFM.

Many other investigators have also worked on methods to mitigate the limitations of 1D AFM. For example, Griffith et al. developed a system for metrology of high aspect ratio structures [51-53]. Some of the unique features of their approach were an electrostatic balance beam force sensor and the use of very sharp and near-cylindrical tips. Although this approach improved significantly over the performance of conventional AFM at the time, it was not a two-dimensional method in the same sense as CD-AFM, but it is certainly 1D+.

Subsequently, Morimoto, Watanabe, et al. developed an AFM method for imaging steep sidewalls [54,55]. A central feature of their approach was a scanning algorithm that involved retracting and stepping the tip rather than maintaining continuous contact. The advantage of this method is that the steep-edge artifacts
that result from tip bending and the Z axis feedback loop in conventional AFM are mitigated with the step in approach. Their system was able to use sharp tips – including carbon nanotube (CNT) tips—to maximum advantage – including the ability to both leverage the attractive bending of the CNT toward the sidewall and correct for the resulting distortion [55,56]. Due to its sidewall imaging capability, this system clearly exceeds conventional AFM. At the least, it should be considered 1D+, but could also be regarded as fully 2D in the sense of CD-AFM.

An interesting alternative paradigm for imaging structures with steep sidewalls is to use some sort of tilting method – in which either the AFM head or at least the scan and control axes are rotated with respect to the sample surface. In 2006 Murayama et al. developed a method which used an angled tip and modified scan axes to enable imaging of steep and even undercut sidewalls [57]. Their method relied primarily on software control of the scanner to implement a non-vertical approach to a sidewall, rather than implementing a physical tilt in the scanner. This did require the use of AFM cantilevers which had the tip grown at a non-normal angle (about 65°) with respect to the cantilever, but these are readily available. Although clearance was a challenge in dense trenches and only one sidewall could be imaged at a time, this method represented a major advance, and it could certainly be considered 1D+.

Since 2011, a tilting technique has been developed and commercialized which is based on controlled and measured tilting of an AFM head during imaging [58]. This method does not require flared tips or a non-vertical oscillation of the cantilever as in CD-AFM, and it can use typical AFM cantilevers. However, its accuracy is dependent upon the decoupling of the lateral and vertical scanner axes and the accuracy of the data stitching when the AFM head is at different tilts. In principle, this technique has the potential to play an important role in metrology applications. However, its effective tip-sample position control is 2D as in CD-AFM. The tilting of the head is restricted to a single rotational axis, and so the ability to image sidewalls is limited to sidewalls oriented for the fast scan axis.

Even more recently, researchers at NMIJ developed a metrology AFM with an incorporated tip tilting mechanism [59]. This instrument, called the tilting mAFM (metrology AFM) is unique in its combination of a metrology grade scanner with tip tilting capability. It is also unique in the implemented approach to axis decoupling. The scanner itself does not tilt, only the cantilever is tilted using an integrated rotational stage. The effective angle of the servomechanism is controlled independently using an electronic implementation that combines the raster scan displacement signal with surface sensing signal using the trigonometric functions appropriate to the desired angle of the servomechanism. The effective tip-sample position control is still 2D, however, as is true for CD-AFM and the tilting head design. The tilting of the cantilever in the tilting mAFM is also restricted to a single rotational axis, and so optimal sidewall imaging requires a specific orientation.

The characterization of such techniques as 2D is potentially confusing – since some AFMs have been described as “3D-AFM.” From the perspective that is developed in this paper, however, no method is fully 3D with respect to the tip sample interaction. In our view, the bottom line is that no true 3D-AFM currently exists.

3.3.1 Slow scan axis interaction constraints

Although we have argued in the prior section that current tip geometrical constraints limit all AFMs to less than three dimensions in that respect, the more significant and limiting constraint is usually tip-sample feedback and position control. In all current instruments, the tip-surface interaction is directly sensed and controlled in at most two axes: the Z axis and the fast-scan lateral axis. This constrains such instruments to two-dimensional functionality in the sense of tip-sample interaction and position control.
As initially discussed in section 2, an important consequence of this is that the feedback is essentially ‘blind’ to topographic gradients in the third dimension – along the slow scan axis. However, due to the topographic constraints on real surfaces and the manner in which CD-AFM is typically applied, this limitation presents fewer blatant imaging problems than it might at first appear.

For virtually all physically reasonable surfaces, topographic changes along the slow scan axis will also require an observable topographic change along the fast scan axis. Nevertheless, there are significant resulting limitations on CD-AFM imaging of three-dimensional structures. One important example is the imaging of near-vertical sidewalls – especially reentrance or undercut topography in a plane orthogonal to the slow scan axis.

Subject to the limitations of tip geometry, the surface sensing and tip position control of CD-AFM is effective at tracking reentrant topography along the fast scan axis. As the tip is scanned, at each step the interaction of the tip and surface is monitored and the tip position is adjusted to track the surface. When a sidewall is encountered, the tip position is incremented in both the fast-scan lateral and vertical axes as necessary to track the sidewall.

The situation is completely different if a sidewall is encountered in the slow scan axis. There is no lateral tip dither along the slow scan axis and hence no direct sensing of the surface normal vector component along the slow axis. The tip position in the slow axis is simply incremented by a fixed amount and not adapted to meet changing topographic conditions along the slow axis.

In most circumstances, the surface gradient will have significant components along both the fast and slow scan axes, and so the effect of this limitation is not dramatic. However, for the worst case in which the slow scan axis is normal to the sidewall, the effect is significant. The data on the sidewall surface will be very sparse or entirely absent, and any reentrance will not be observed.

Figures 9 and 10 illustrate the effect of these slow scan axis limitations on CD-AFM images of intersecting features with near-vertical sidewalls. Both images were obtained on a developmental batch of the National Institute of Standards and Technology (NIST) single crystal critical dimension reference material (SCCDRM). This NIST project is aimed at the development of tip width calibration standards and involves the use of lattice-selective etching of crystalline Si (111) to produce nearly vertical sidewalls. Several generations of SCCDRM have been previously developed [60-66].

In figure 9, an image is shown in which one of the two target lines (feature A) is nearly parallel to the slow-scan (Y) axis. This means that the fast-scan (X) axis is nearly orthogonal (in the XY scan plane) to feature A, which is optimal from a sampling perspective. Note that features A and B do not intersect at 90° due to the alignment of the features with the Si crystal planes. Dense sampling of feature A and its sidewalls can be observed, and its full cross-sectional profile is sampled. Feature B is non-parallel with the slow-scan (Y) axis by about 60°, so the fast-scan (X) axis is nearly 60° away from being orthogonal (in the XY scan plane) to feature B. The data density (i.e., the spacing of line scans) on feature B is sparse relative to feature A, but due to a significant fast-scan axis component of the topographic gradient–there is still adequate sampling of the sidewall and the cross-sectional profile.

The worst-case scenario is shown in figure 10. Feature A is only about 30° out of parallel with the slow-scan (Y) axis, so the fast-scan (X) axis lacks only 30° being orthogonal (in the XY scan plane) with feature A. Data density of sampling of feature A are good. Feature B, however, is very close to being orthogonal (in the XY scan plane) with the slow-scan (Y) axis. This means that the fast-scan (X) axis is very nearly parallel
with feature B. Although the upper surface of the feature is imaged well, there is virtually no sampling of the sidewall or information about the cross-sectional profile along the slow scan axis.

Figure 9. Feature A is nearly parallel to the slow-scan (Y) axis. Dense sampling of the feature sidewall and cross-sectional profile can thus be observed. Feature B is about 30° from being parallel to the fast-scan (X) axis, but there is still good sampling of the sidewall and cross-sectional profile.
Figure 10. Feature A is about 30° from being orthogonal to the slow-scan (Y) axis, but the sampling is still dense and the cross-sectional profile is accurately represented. Feature B, however, is nearly parallel to the fast-scan (X) axis. Although its upper surface is imaged, there is virtually no sampling of the sidewall or information about the cross-sectional profile.

The most common application of CD-AFM is still the width measurement of linear-type features. Consequently, the slow-scan axis ‘blindness’ or two-dimensional tip-sample interaction and control of CD-AFM is often of secondary importance. However, it does greatly limit the applicability of CD-AFM to more complex measurands within the XY scan plane.

An example that is of considerable industrial importance is contour metrology. Although CD-AFM has great potential as a reference tool for other instruments – such as scanning electron microscopes (SEMs) – the use of CD-AFM for contour metrology is greatly hampered by the lack of fully three-dimensional tip-sample interaction sensing and position control.

This limitation of CD-AFM for contour metrology can be effectively illustrated by considering the measurement of pillar-type structures. The specimen we selected was a silicon wafer with multiple large arrays of pillar-type features. These arrays were designed to be scatterometry targets. An image of a portion of such a target, having features that are approximately 100 nm in diameter, is shown in Figure 11a. Due to the lack of tip-sample feedback and sparse sampling along the slow scan axis, there are digitization-like discontinuities in apparent edge position that can be seen around the circumference of the pillars. These discontinuities are particularly apparent near edge boundaries of each pillar along the slow scan axis.

Figure 11. (a) CD-AFM image (top-down rendering) showing a portion of an array of pillars. The fast scan axis is the horizontal axis in this image. There are 50 scan lines distributed across the imaged area in the slow scan axis. Consequently, each pillar is sampled by fewer than 10 line scans. This relatively sparse slow-axis sampling results in the apparent digitization-like edge jumps (i.e., square-like) seen around the perimeter of the pillars. It is most noticeable at the edges of each pillar in the slow scan axis. (b) CD-AFM image using 50 line scans for an individual pillar. This requires a relatively long imaging time for just one feature. However, a much greater fraction of the edge can be sampled using the tip-surface interaction in the fast-scan axis. Qualitatively, the edges of the pillar are well-defined around most of the circumference, although the digitization square-like edges are still noticeable – particularly in the slow axis boundaries of the pillar.
In order to partially mitigate sparse sampling and the blindness of the tip-sample interaction in the slow-scan axis, the image in Figure 11b was zoomed in on a single pillar and a relatively large number of scan lines (50) were taken. This method comes with the disadvantages of a relatively long imaging time and a sample set consisting of a single feature. It is, however, an effective work around since a significant fraction of the edge can be sampled using the tip-surface interaction in the fast-scan axis. Qualitatively, the edges of the pillar seem well-defined around most of circumference of the structure, with only the scan lines crossing near the edge boundaries along the slow-scan axis of the pillar exhibiting the square-like jumps in apparent edge location.

This effect is visible more quantitatively in Figure 12, which shows the result of applying a built-in CD-AFM analysis function for the width of contact holes to the image in Fig. 11b. The results of the analysis are shown as curves in the \((x,y)\) plane that represent the ‘edges’ of the pillar as defined by the top, middle, and bottom widths. As expected, the edge definition seems consistent and stable around most of the pillar. However, near the edge extrema of the pillar along the slow-scan axis, the analysis function is unable to generate results, and there is some indication of instability in the results for the last scan line or so near each extreme.

![Figure 12. Analysis of Fig. 11b image using via function. The results are shown as curves in the \((x,y)\) plane that represent the ‘edges’ of the pillar as defined by the top, middle, and bottom widths. As expected, the edge definition seems consistent and stable around most of the pillar. However, near the ‘top’ and ‘bottom’ along the slow-scan axis, the analysis function is unable to generate results.](image)

Although it is possible to reduce this effect further by taking even more scan lines, this would not eliminate the fundamental problem of tip blindness along the slow-scan axis. The only way to truly eliminate this problem would be to detect the tip-sample interaction and to actively control the tip in both lateral axes. This would help open up the possibility of scanning the tip in a more complex pattern than a raster scan – such as scanning the tip around the edge of a feature to map out the contour rather than scanning line by line across the entire the structure. This would start to more closely resemble the data acquisition and analysis methods used in coordinate measuring machine (CMM) metrology [32], and such a new type of AFM could come much closer to being fully 3D than any current AFM.

Short of developing a new AFM, another approach to the constraint of slow-scan axis blindness is to change the orientation of the fast-scan axis relative to the feature and combine the information from multiple images.
This could be accomplished either by a physical rotation of the sample between images or by changing the fast-scan axis on the instrument. The authors have developed such a method of AFM contour metrology by using two images with orthogonal fast-scan axes. A composite contour is then constructed with information in both lateral axes [67-68]. Conceptually, this method is analogous to tilt SEM or stereo vision in the sense that multiple measurements taken under different observation conditions are combined to increase the effective dimensions of the composite result.

Another related approach to contour metrology is to combine the information from AFM with data from another technique. For example, Foucher et al. [69] have explored the use of CD-AFM in direct combination with CD-SEM as a hybrid metrology technique for contours. In their approach, the strength of CD-AFM in the fast scan axis is used to support the accuracy of CD-SEM across an SEM contour. Each CD-AFM scan line is mapped to a linear slice through the SEM contour along the same axis. Thus each AFM linescan can serve as a reference for the appropriate SEM linescan value. Although this approach mitigates some limitations of CD-AFM, its two-dimensional nature still presents some challenges for sections of the contour that are orthogonal to the slow-scan axis.

3.3.2 A 3D measurement strategy — Vector-Approach Probing (VAP)

The axis-specific AFM imaging constraints described in prior sections are generally related to the fact that the typical AFM measurement strategy is based on a raster scan in which data are acquired using a fast and slow scan axis. These constraints have the effect of rendering the resulting data less than fully 3D in at least some respects. A step closer to a fully 3D AFM measurement would be to implement a specimen probing strategy that more closely resembles the methods typically used in coordinate measuring machines (CMMs). In most CMM measurements the probe can approach the specimen along any axis and the approach axis can be changed as appropriate for every data point. Recently, this type of sampling method was realized in the custom CD-AFM system that the Physikalisch-Technische Bundesanstalt (PTB) has developed [42]. This method is referred to as vector approach probing (VAP). In the VAP method, the structure is probed point-wise by the AFM tip which approaches the surface along a specified 3D vector – which is typically aligned with the local normal vector to the surface. A tip-sample interaction curve is recorded in order to determine the detected surface position for each acquired point.

The VAP method offers considerable flexibility for measurements of nanostructures, since it is not based on a raster scan method and does not ultimately require the use of a fast and slow-scan axis. With the VAP technique, it is possible to probe surfaces in 3 dimensions (XYZ) and measure surfaces in an arbitrary plane (XZ, YZ, XY, etc.). This capability readily allows for direct imaging of sidewalls. An example of this is shown in Figure 13(a). In this case, the vertical sidewall of a line feature is imaged in the xz plane using a flared AFM tip type CDR120 with the VAP approach vector oriented with the y-axis. This alignment with the surface normal also affords optimum sensitivity in the approach curve (i.e., amplitude vs. tip-sample separation). To demonstrate such a measurement, a sidewall surface topography measured on an IVPS (Improved Vertical Parallel Structure, Team Nanotec®) sample is depicted in figure 13(b). Using such a measurement strategy, the point-wise measurement repeatability is typically 0.3 nm (rms), thus offering excellent measurement capability for applications such as the line edge/width roughness (LER/LWR) metrology. The VAP method also has strong potential for applications in contour metrology, as illustrated in figure 13(c). By applying a pre-measurement or using the CAD data of the features, such a VAP measurement strategy can be generated by the software.
Figure 13. (a) Illustration of the measurement strategy for direct sidewall measurements using the VAP method; (b) Measured topography of the sidewall surface of a line feature using the VAP method as shown in (a); (c) Illustration of applicable VAP measurement strategies for measurement of different contour geometries.

Although the VAP measurement strategy is capable of 3D measurement strategies in principle, the other constraints of CD-AFM related to tip-sample interaction and geometrical occlusion are nevertheless still present. In terms of the definitions and perspective we are developing here, this instrument is most appropriately regarded as 2D or 2D+.

3.3.3 Position drift and slow scan axis noise

The presence of significant drift in an AFM doesn’t fundamentally change the number of spatial dimensions of an instrument or its data in any of the senses we have considered. However, due to the sequential nature of AFM data acquisition, drift-related effects will tend to be of differing magnitudes among the axes of an instrument. Consequently, drift can significantly reduce the usefulness of the dimensional information for a specific axis.

Sources of drift themselves – with stage drift being the most common – are essentially a component of low frequency noise. However, since image data points are acquired as a time series, such drift can produce image artifacts of a non-random and axis-specific nature. The two most common effects are: (1) lateral
image distortion and skewing of features, and (2) apparent slow-scan axis striation due to low frequency z-axis drift – familiar to most users of AFM.

When scanning over a regular feature – such as a grating line – lateral stage drift results in a change of the apparent location of the structure from scan line to scan line. In circumstances for which the drift is monotonic during the timescale of the image, this will change the apparent angle between the instrument axis and the grating lines. Often, an apparent curvature will be induced as well.

Figure 14 shows three CD-AFM images of different but similar targets on three separate NIST SCCDRM samples. Each example illustrates a different impact of drift. The actual features are nearly straight and are aligned nearly orthogonal to the fast scan axis. However, due to lateral stage drift, each image exhibits a different apparent rotation and curvature of the features in x-y plane. The first line of defense against drift is to reduce the image acquisition time. Due to the large scan area and high data density, the images of figure 14 required approximately 20 minutes of scan time each. Consequently, the magnitude of the drift effect is relatively extreme. For shorter imaging times, the effect of lateral stage drift would be less dramatic.

![Figure 14](image.png)

*Figure 14. Images of three different NIST SCCDRM targets exhibiting different apparent rotation and curvature of the features in x-y plane due to lateral stage drift. The target lines are actually very nearly orthogonal to the fast scan axis.*

It is worth noting, however, that even much smaller levels of drift present challenges to applications such as contour metrology. Although the impact of drift on contour measurement is a less serious limitation than the slow axis tip-sample interaction constraint that was previously discussed, it nonetheless would present a challenge to any attempt to perform single-image AFM contour metrology. And even the dual-image methodology developed by the authors is not completely insensitive to drift [68].
The apparent striation of AFM images along the slow-scan axis results from low-frequency stage drift in the z-axis. Because the drift is relatively slow, the adjacent points in a given scan line do not show much effect, and between adjacent scan lines, the drift, though larger, is also relatively low. However, between widely separated scan lines, the cumulative effect of the slow drift on the average z-value of the scan line will be much more noticeable. It is particularly apparent on smooth surfaces with relatively low topographic contrast.

Figure 15 shows a 1D AFM image, rendered in a top-down view, of chromium lines on a silicon substrate. These features are approximately 100 nm high. The apparent horizontal striations are a manifestation of the low frequency drift in the z-axis. Since the x-axis is the fast scan axis, the striations are horizontal in this image. The purpose of the image was to measure the spacing between these features, and data of this type would normally be analyzed line by line. In that case, the slow axis striations don’t present a major problem because a zero-level is usually subtracted from each scan line.

However, to the extent that the striations require line by line analysis or flattening of the image, the relative topographic information along the slow scan axis is diminished or lost. Although the lost information may not be relevant for some applications, it could be regarded as reducing the number of useful dimensions of an AFM image with respect to the data content.

4 AFM Measurands: Input Quantities and the Measurement Equation

The primary emphasis of this paper so far has been the functionality of instruments and data acquisition rather than specific applications or types of measurements. However, a discussion of spatial dimensions in AFM would not be complete without some consideration of measurands. Our treatment here is limited, however, since this topic could easily be the subject of a paper on its own.

The first concept to consider is the definition of a measurand itself. According to the ISO Guide to the Expression of Uncertainty in Measurements [70], the measurand is the quantity intended to be measured. Conceptually, the measurand thus represents the abstract definition of a quantity to be measured. For metrology applications of AFM, this might mean the step height or linewidth of a target feature.

It should be emphasized, as Phillips and coauthors have pointed out, that the measurand is an idealized concept that provides an exact definition that specifies the values to all the conditions that affect the measurand, for example, the temperature at which length is defined is exactly 20 °C [71] and at zero applied force [72]. It is usually impossible to produce a measurement result that is exactly to the specifications of...
the measurand. Any failure to achieve (or perfectly correct to) the conditions of the measurand becomes a source of measurement uncertainty [73]. There is also the challenge that the measurand may be inadequately specified. This is particularly likely to be a problem for applications in which multiple measurement techniques are utilized – nominally to measure the same quantity. The ambiguity of the measurand definition for some of these techniques may not be obvious without careful consideration. The measurement of nanoparticle size by methods ranging from AFM and SEM to differential mobility analysis (DMA) and dynamic light scattering (DLS) is a relevant example of an application where such challenges arise [74]. However, a detailed consideration of this subject is beyond the scope of this paper.

The objective of any actual AFM measurement procedure is to estimate the value of some particular measurand. In many cases, a measurand is not measured directly but is determined from $N$ other quantities through some form of functional relationship. This is often referred to as the measurement equation and represented by:

$$Y = f (X_1, X_2, \ldots X_N)$$  \hspace{1cm} (3)

The $X_i$ are typically referred to as input quantities and might themselves be regarded as measurands and might depend on other quantities. The function $f$ may be represented either analytically or numerically and might not be explicitly known.

For our purposes in this paper, a key point is that number of spatial dimensions of the output of a measurement equation is not necessarily the same as for the input quantities. The output of a measurement equation may be a scalar, vector, or an array, and the same is true for the input quantities. It is therefore necessary to consider the spatial dimensions of both the input quantities and the output of the relevant measurement equation.

Many common measurands in AFM metrology that are abstractly regarded as scalar quantities are thus typically modeled by measurement equations having a 1D output. This includes, among many others, the common measurands of step height, pitch, linewidth, sidewall angle, and root-mean-square roughness. There are a variety of algorithms used to compute these quantities, and the calculation is often line by line. In this case, the results could be presented as an array vector of values – one for each scan line of the image. Such arrays are often an intermediate step in the calculation of a grand average for the entire image, but in some cases an array of values might be regarded as the output. Even in such cases, however, we argue that the most meaningful characterization of the output would be determined by the nature of the array elements rather than the size of the array.

CD-AFM contour metrology is an important example of a measurand for which the output of the measurement equation is 2D. For a given edge detection criterion, a contour ultimately consists of an array or collection of lateral position values – i.e., $(x, y)$ pairs – that represent the outline of all the edges in an image. The total number of such points will be of interest in the context of the application and specific measurement, but it is the two-dimensional character of the basic $(x, y)$ unit that determines the nature of the output.

Consideration of the input quantities to a measurement equation presents a more conceptually complex and interesting landscape. Generally, we propose that the most meaningful metric corresponds to the number of spatial dimensions in the input quantities that directly enter the computation of a spatial dimension in the output.
In the case of most height calculations, for example, only the image information in the \( z \)-axis is directly relevant to the result – which represents a differential value between two defined \( z \)-axis levels. This is straightforwardly seen by considering a histogram algorithm. Other algorithms involve calculating the \( z \)-axis difference between lines fitted to the upper and lower surfaces at defined or extrapolated locations [75]. The subtlety here is that the lateral axes may not be totally irrelevant. For algorithms involving line fitting, there is an implicit assumption of uniform lateral pixel size or perhaps explicit use of data from a lateral position channel. Also, in cases where corrections for form error are being applied, the lateral data may be involved. Nevertheless, the effect of lateral errors on the final step height result is essentially a higher order correction that will be small relative to the impact of \( z \)-axis errors. Overall, it seems most reasonable to regard step height as 1D with respect to the measurement input quantities.

Pitch may be considered either 1D or 2D with respect to both the measurement input quantities and the output of the measurement equation. The conventions we propose are generally consistent with ordinary intuition and common usage of terminology. Typically, a pitch measurement of a linear array is regarded as 1D, and a pitch measurement of an array of pillars is regarded as 2D.

For the case of a linear array, the data analysis would often be conducted line by line on individual profiles. Since the measurement output is typically given by the lateral distance between defined edges or centroids or perhaps the frequency-domain peak from a 1D fast Fourier transform (FFT), only the data from a single lateral axis (usually the fast-scan axis) is of direct importance. Analogous to the situation with step height, however, the other axes cannot be totally ignored. The vertical axis data does enter the computation since it must be involved in defining the positions of the edges or extrema. However, the sensitivity of the measurement to errors in the \( z \)-axis will be greatly suppressed relative to the fast scan axis. Normally, the slow scan axis would be of very low importance. A pitch measurement of a linear array can thus be regarded as 1D in terms of both the measurement input and output.

In contrast, for the case of a 2D array of pillars, the data analysis would amount to a calculation of differential location involving both lateral axes – regardless of the actual algorithm used. The application of a 2D FFT to determine the average spatial frequency in \( x \) and \( y \) is common. However, spatial domain analysis to determine the relative \((x,y)\) positions of the pillars is also used. In either case, both lateral axes are directly involved in the calculation, while the \( z \)-axis plays a secondary role – essentially as a weighting function. This example also illustrates the significance of and the importance of distinguishing between the measurement output and the input quantities. A pitch analysis of a 2D array of pillars might yield either two pitch values – corresponding to the principal axes – or these might be averaged into a single value intended to represent both axes. In the first case, both the measurement output and the input quantities are two dimensional. However, in the second case the measurement output is 1D while the input quantities are still 2D.

The situation is more complicated for linewidth measurements. A general review of methods for AFM linewidth metrology was recently published by Hussain and coworkers [76]. Their discussion emphasizes instruments and techniques but also touches on measurands.

An interesting new layer is present in linewidth metrology because of the importance of a spatial dimension not related to scanner displacements: the tip width. The dimensions of a linewidth measurement output correspond both with common intuition and the conventions previously described for pitch metrology. In other words, the width measurement of a linear feature yields a 1D output, and the width measurement of a pillar or pit type of feature may have either a 1D or 2D output. With respect to measurement input quantities, however, there is more to consider.
The width measurement of a linear feature must be at least 1D in terms of the measurement input quantities – since data from one lateral axis is involved. Similarly, the width measurement of a pillar type of feature must be at least 2D with respect to measurement input quantities. However, unlike a pitch measurement, the edge definitions relevant to a width measurement are directly sensitive to the lateral tip width in addition to the lateral scanner axis. The view could thus be taken that this additional dependence increases the number of spatial dimensions of the measurement input quantities. For example, one could regard the width measurement of a linear feature – which exhibits a 1D output – as 1D+ or even 2D with respect to the measurement input quantities. The potential advantage of using such descriptions would be that it helps to emphasize to the non-expert that there is a significant fundamental difference between a pitch measurement and a width measurement of linear features. However, since the specific nature of the additional dimensional input is from a different source (the tip instead of the target), the designation of 2D could also be misleading, and we would propose that 1D+ be used to describe this circumstance.

Similar questions of nomenclature also arise for dimensional metrology at larger length scales. At such scales, for which CMMs are the most prevalent metrology tool, a distance between any two parallel faces is always regarded as 1D. A distinction is then drawn between unidirectional and bidirectional measurements [32]. Unidirectional CMM measurements are those for which the probe-related bias (i.e., an effective probe radius) would cancel out. This is analogous to AFM pitch measurements. Bidirectional CMM measurements are those for which the probe-related bias would be doubled. This is equivalent to AFM linewidth measurements. Although the type of terminology used for CMM measurements is clear to a dimensional metrologist working with AFM, our view is that differentiating between these cases as being 1D or 1D+ measurements may be more readily appreciated by a broader audience.

Angular measurements are interesting because they must always involve information from two axes. Therefore, although the measurement output is 1D, the measurement input is 2D. One very common example in CD-AFM is the measurement of the sidewall angles of integrated circuit structures, some of which are near-vertical trapezoidal features [77]. In this case one lateral axis, normally the fast-scan axis, and the vertical axis both enter directly into the slope calculation. Another example would be the measurement of the in-sample-plane non-orthogonality of the principal axes of a 2D grating. Data from both lateral axes enters directly into the calculation.

An interesting possibility emerges at the intersection of contour metrology and angular metrology. Contour metrology, as mentioned previously, is 2D with respect to measurement output. It is also 2D in terms of measurement input, since data from both lateral axes enters directly into the calculation. A conventional contour represents a collection of \((x,y)\) edge locations – given a particular edge definition (e.g., edge location at the half height of the feature). However, it would also be possible to develop a more complex measurement that incorporated additional information – such as the local sidewall angle at each edge position in the contour. Such a contour, consisting of \((x,y,\theta)\) triplets would be 3D in both measurement output and input, since the vertical axis information is directly required for the angle calculation. We are not aware of prior work of this type, but it may become important in the application of CD-AFM to contour metrology.

The final class of measurands we will consider is those pertaining to surface roughness. In the field of classical surface finish metrology, there are a number of commonly used parameters [78]. However, in AFM metrology, the two most common are root-mean-square roughness and the power spectral density (PSD). The designation \(R_q\) is normally used to indicate root-mean-square roughness calculated from a profile while the designation \(S_q\) typically indicates roughness calculated from an image [79].
As previously mentioned, \( R_q \) is a 1D measurement in terms of output. In the surface finish field, \( R_q \) is classified as an amplitude parameter – since it is sensitive only to the z-axis data. There is an implicit assumption of uniform pixel spacing in the typical computation, but it is otherwise insensitive to the lateral axes. We therefore regard these roughness parameters as being 1D in terms of measurement input as well, and this seems appropriate regardless of whether it is calculated line by line as \( R_q \) or using the areal formula for \( S_q \), since both of these are primarily sensitive to the vertical axis.

A consideration of power spectral density also follows readily from classical surface finish. The profile PSD is 2D with respect to both measurement output and inputs – since it is sensitive to both a lateral and vertical axis. The areal PSD (APSD), which is calculated from an entire image, would be classified as 3D with respect to both measurement inputs and output. Although we believe this characterization is appropriate, it is worth noting that it is not fully consistent with common usage. The profile PSD is calculated using a 1D FFT algorithm, and the APSD is calculated by using a 2D FFT. Consequently, these two measurements are sometimes regarded as being 1D and 2D, respectively. We suggest that this is another area in which clarity could benefit from further consideration.

5. Summary

Ultimately, the aim of this paper was to underscore the fact that no truly three-dimensional AFM currently exists. This was emphasized through the development of a framework for describing the spatial dimensions of AFM instruments, data, and measurements.

With respect to AFM instruments and data, it is important to consider the spatial dimensions of the data structure and content, the tip geometry, and the tip-sample interaction and positioning control. The development of the VAP method by PTB represents important progress toward 3D. There has also been progress by others in the development of non-raster AFM scan strategies. These methods include spiral scans\cite{80,81}, cycloid scans\cite{82}, and Lissajous scans\cite{83}.

An important point highlighted here is that the tip-sample feedback and position control are currently the critical factors that limit the technique to being two-dimensional. Furthermore, we suggest that currently available CD-AFM tips would also limit the technique to being less than three dimensional. With respect to AFM measurands, we suggest that the dimensions of both the input quantities and output of the relevant measurement equation should be considered.

The absence of a widely accepted dimensional nomenclature for AFM data, measurements, and instruments is often of limited importance to those with experience in the field. However, there is some risk of misleading newcomers to AFM or those working in related areas. Therefore, our goal in this paper was to propose a framework for describing the dimensional characteristics of AFM images, instruments, and measurements in a manner that can convey important distinctions.

As has been noted, however, there are areas of ambiguity in the framework we have developed. Therefore, this work is intended less to be the final word on the subject than to support the goal of fostering further discussion among AFM users. Greater clarity of meaning in descriptions of the spatial dimensions in AFM instruments, data, and measurements would benefit both AFM users and their collaborators/customers in application-specific fields that utilize AFM measurements.
Acknowledgements

The NIST contribution to this work was supported by the Traceable Scanning Probe Nano-characterization project within the Engineering Physics Division (EPD) of the NIST Physical Measurement Laboratory (PML). The contribution of AIST was supported by the operating support funds of the National Metrology Institute of Japan (NMIJ), AIST. The authors acknowledge valuable discussions with Jack Stone, Steve Phillips, Ted Vorburger, Bill Banke, and Vladimir Ukraintsev, Ryosuke Kizu, Satoshi Gonda, Akiko Hirai, Dorothee Hüser-Espig and Harald Bosse.†

†Certain commercial equipment is identified in this paper to adequately describe the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the equipment identified is necessarily the best available for the purpose.
References


