Development of a physical haze and microroughness standard

Bradley W. Scheer

VLSI Standards, Incorporated, 3087 North First Street
San Jose, CA 95134-2006

ABSTRACT

Microroughness, or haze, on wafer surfaces can mask the detection of particles by scanning surface inspection systems (SSIS). The ability of silicon dioxide or other films to function efficiently as insulators depends partially on the underlying microroughness of the silicon surface. For thin oxides, breakdown voltages are reduced commensurately with increased levels of microroughness. There are similar effects on film layers deposited in later processing steps, and an effect on bonding for silicon-on-insulator (SOI) applications. The disk drive industry depends on a steady transfer rate of data from the recording medium. Imparted surface texture must be carefully controlled since it is in conflict with the desire to have the head in close proximity to the recording surface; however, too fine of a polish can lead to stick-slip or blocking Additionally, the surface texture must be very uniform across the face of the recording media. The flat panel display industry, with their ubiquitous screens now so common in laptop computers, relies on an orientation layer with a precise amount of microroughness on the substrates.

In these and numerous other applications, a well characterized surface is paramount to high production yields. And yet, the most often used of these measurements, rms microroughness, is grossly misunderstood. *Surface roughness is not a unique number nor is it an intrinsic surface property*. Roughness measurements depend on the parameters of the instrument used for measurement -- whether that be an optical or mechanical profiler, a SSIS used for haze detection, or an atomic force microscope (AFM). Each of these instruments may give very different values from exactly the same surface.

A novel approach to developing a practical haze standard has been employed by photolithographically etching features to as little as 1nm deep into the surface of 150 mm silicon wafers. To prevent a wafer scanner from detecting these features as particles or other light scattering events, a high surface density ($4x10^6$ features/cm²) is produced such that the distance between features is much less than the spatial resolution of current instruments (typically 50μm - 100μm).

Once the wafers have been fabricated, the haze or microroughness level detected by a given scanning instrument may then be calculated from the Power Spectral Density (PSD) function plots generated for each wafer by various techniques (such as angle resolved light scattering or atomic force microscopy). The amount of simulated haze produced by this method is a function of the depth of etch into the silicon surface.

Keywords: atomic force microscopy, bi-directional reflectance distribution function (BRDF), haze, microroughness, power spectral density function (PSD), profilometry, scanning surface inspection system, scatterometry, surface texture, wafer scanner

1. INTRODUCTION

Microroughness is defined as "surface roughness components with spacings between irregularities (spatial wavelength) less than about 100 microns."[1] This definition differentiates microroughness from the larger scale surface variations of bow and warp.* Haze is a condition of a wafer surface resulting in a significant level of light scattering. Haze may result from surface topography (hence, directly related to microroughness) or from dense concentrations of surface or near-surface imperfections. Until recently, there has been no way of calibrating haze and microroughness on various instruments. Perhaps more importantly, there has been no way of correlating seemingly uncorrelatable instruments.

These very small levels of surface texture are becoming more problematical in a number of industries as the complexity of integrated circuits and the amount of information stored on disk drives increases. As an example, geometries in the integrated circuit industry are fast approaching molecular dimensions. The June 1994 report, the National Technology Roadmap for Semiconductors (NTRS), has published a requirement for gate oxides approaching 4.5 nm ± 4%.[2] As a point of reference, the

* Bow and warp are terms associated with silicon wafer specifications. They are typically on a scale of millimeters of spatial wavelength.

0-8194-2250-9/96/$6.00

lattice constant for lightly doped (i.e., nearly pure) silicon is 0.543 nm. The gate dielectric molecule, silicon dioxide, is nominally 0.355 nm "diameter".[3,†] The ability of silicon dioxide or any film layer to function efficiently as an insulator depends partially on the underlying microroughness of the silicon surface. For oxides less than 100Å thick, breakdown voltages are reduced commensurately with increased levels of microroughness. This can be readily understood by envisioning the "peaks" of the microroughness terrain as being much closer to the film surface than the overall average level of the peaks and valleys combined. Additionally, there are similar effects on film layers deposited in later processing steps, and an effect on bonding for silicon-on-insulator (SOI) applications.[4]

Currently, there are several techniques available for measuring microroughness. However, the results tend to be qualitative since there was no physical metrology standard previously available to correlate the accuracy of various instruments.[‡] This becomes especially important when comparing instruments with differing spatial bandwidths, each possessing a unique transfer function. Due to the varying spatial bandwidths, different types of instruments can give rms microroughness values that differ by over an order of magnitude, even when measuring the same surface.[5, 6]

The methodology used to produce the haze and microroughness standard provides a known surface texture on a substrate with feature depths ranging from 1 nm to 10 nm.[§] This is accomplished by precisely etching pits into silicon at known locations on a silicon wafer (figure 1). Two sizes of pits are employed (figure 2) to minimize Airy disk diffraction effects which would otherwise be apparent (see section 4, *Design Considerations*). The diffraction would produce distinct minima on the power spectral density curve that represents the surface.

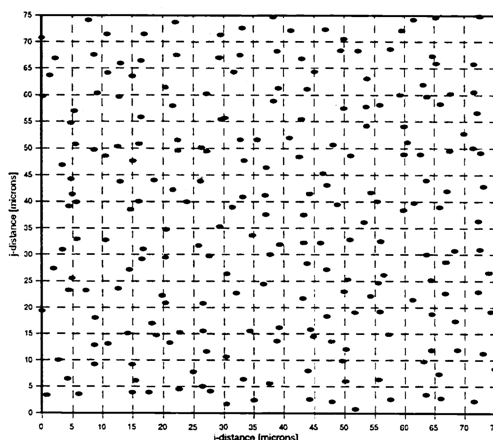

Figure 1. Representation of the placement of etched features on a section of the surface of the wafer. Notice that there is precisely one feature in each imaginary 5 μm square box.

This imparted texture is fully quantifiable by angle resolved light scattering with secondary measurements performed by atomic force microscopy. By having these data available and knowing the spatial bandwidth that any given instrument operates in, it is possible to provide a direct method of quantifying a prescribed surface texture or microroughness value.

In order to characterize surfaces, it is useful to calculate statistical properties from the measurement data obtained from various instruments. The two most widely used of these parameters are the centerline average roughness, R_a, and rms roughness, R_q (equations 1 and 2). Roughness values are calculated from the measured values of height variation, z_i, over a portion of a surface with a given number of sample points, N.[7]

---

† Obviously it is not possible to associate a diameter with an SiO_2 molecule; however, this number is based on a volume ratio which is 2.25 times that of a silicon atom. The linear dimension stated is taken as the cube root of this ratio.

‡ The term "standard" in this sense relates to a physical artifact that is used to verify the accuracy and precision of analytical test equipment. This is in contrast to a consensus, or "paper standard" which may specify a test procedure.

§ There is only one depth present on any given wafer. The range available allows for various levels of microroughness simply defined as Low, Medium, and High (since no single number may be associated with the microroughness value).

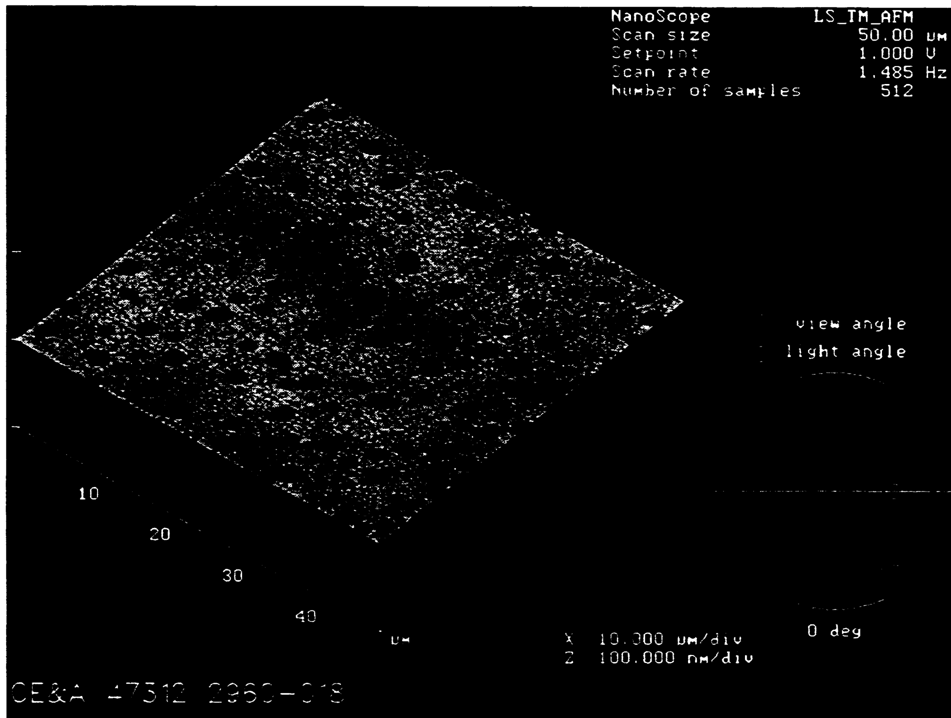

Figure 2. *AFM image of the dual-pit size layout. Notice that the scale of the z-axis is 100x the x-y axis. The actual aspect ratio of the etched pit may be compared to the thickness of a business card 18 inches in diameter!*

$$R_a = \frac{1}{N}\sum_{i=1}^{N} |z_i| \tag{1}$$

$$R_q = \sqrt{\frac{1}{N}\sum_{i=1}^{N} z_i^2} \tag{2}$$

However, these parameters do not uniquely define a given surface. For example, depending on the total traversing length of a given instrument as well as the number of points sampled, surfaces of different textures may yield the same rms roughness value .

A measurement system that samples, say, every 10 microns of a surface may totally miss the high spatial frequency** variations on the sample surface. This produces a microroughness value that is much lower than would be obtained had it sampled every 1 micron.

A more complete surface quantification that takes instrument spatial bandwidth into account is the power spectral density (PSD) curve. The PSD function is the frequency spectrum of the surface roughness measured in units of spatial frequency, typically inverse microns.[8] From the PSD data, readings from various instruments may be correlated by incorporating their spatial bandwidths. This concept is discussed in detail in sections 2, *Instrument Considerations* and 5, *Determination of Specific Haze and Microroughness Values*.

** Spatial frequency is the inverse of spatial wavelength. A surface variation that changes every 10 microns would have a spatial frequency of 0.1 microns.

2. INSTRUMENT CONSIDERATIONS IN DETERMINATION OF HAZE AND MICROROUGHNESS

Consider a surface, magnified to show a section 5 µm in length (figure 3). If this surface is measured by an instrument which samples, say, every 0.5 µm as opposed to every 0.1 µm, then the reported surface tends to look smoother since the high spatial frequencies are missed. This is analogous to the electrical phenomenon of aliasing. In this case however, rather than sampling a time-varying function at too low a rate, a spatially-varying function is sampled too infrequently for a proper representation of the actual surface. In other words, the high end spatial frequency capability (measured in units of inverse microns) of the sampling instrument is too low. On the other extreme, if the warp in a 200 mm silicon wafer is measured by a profilometer with a traversing length of 50 µm, it never senses the warped condition of the surface. Since the warp in the wafer tends to have spatial wavelength components of the order of the wafer diameter, this implies that the spatial wavelength is approximately 100 mm (or, conversely, a spatial frequency of 1×10^{-5} μm^{-1}).

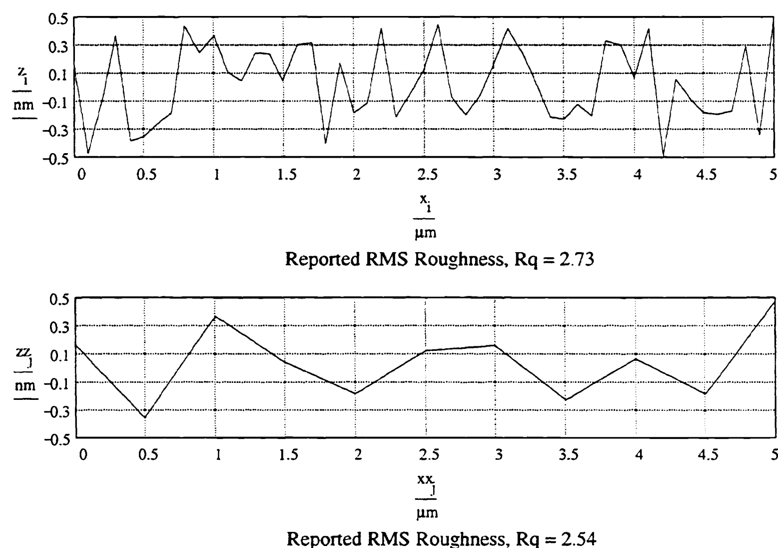

Figure 3. *Reported rms microroughness of the same sample area by two different surface topography instruments.*

The high frequency limit, f_{max}, of an instrument may be calculated directly by knowing the sampling distance, τ_o, where τ_o is greater than or equal to the lateral resolution of the instrument.

$$f_{max} = \frac{1}{2\tau_o} \tag{3}$$

The factor of 2 in the denominator assures that the minimum Nyquist length criterion is met. This helps to minimize aliasing effects.

The low frequency limit, f_{min}, is simply determined from the evaluation length, L, such that

$$f_{min} = \frac{1}{L} \tag{4}$$

where L is less than or equal to the traversing length, L_t. The maximum spatial bandwidth (or bandpass) is then defined by these spatial frequency limits, or may be electronically narrowed (e.g., an electrical cut-off filter on a stylus profilometer).

In the case of a laser based instrument, the bandpass is defined somewhat differently. If the incident monochromatic laser beam, at incidence angle θ_i, specular beam, and scatter beams (at angles θ_s) are all in the same plane, the spatial frequency is related to the scatter angle by the one-dimensional grating equation

$$f = \frac{\sin(\theta_s) - \sin(\theta_i)}{\lambda} \tag{5}$$

where f is the spatial frequency and λ is the laser illumination wavelength.

Ideally, the instrument transfer function is flat in the bandpass region and zero elsewhere. This is never achieved in practice and the spatial bandpass must always be convolved with the transfer function (if available) in order to achieve true inter-instrument comparisons.

Referring back to figure 3, the first profile is measured by an instrument with a high frequency limit of at least 5 μm^{-1} ($f_{\max} = \frac{1}{2\cdot(0.1\mu m)}$). This instrument reports an rms roughness level of 2.73Å. The second instrument (or the same instrument with a low-pass filter) is reporting 2.54Å rms but has a high-end limit of 1 μm^{-1} ($f_{\max} = \frac{1}{2\cdot(0.5\mu m)}$). It is shown later that this difference in reported roughness can be much more drastic with even more diverse instruments.

Figure 4 shows PSD curves of initial haze wafers derived from profiles measured with a Digital Instruments atomic force microscope in Tapping Mode™. The format of this graph is a two-dimensional power spectral density plot which relates surface roughness power per unit of spatial frequency.

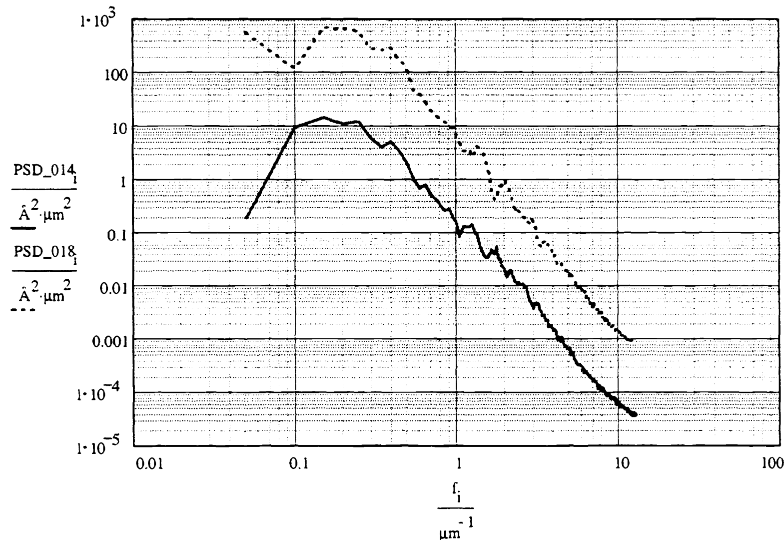

Figure 4. Power spectral density plots from AFM scans of the Haze and Microroughness Standard prototype wafers 2960-014 (low level) and 2960-018(medium level).

The most useful feature of the PSD function is that it relates information about the Fourier transform of the surface into a form that makes it possible to readily compare information generated from various instruments. The rms roughness may be calculated directly as the square root of the integral of the one-dimensional PSD.[††]

The actual scanned area of the wafer was 20 μm x 20 μm with a total of 512 samples per 20 μm scan length. Therefore, the lower spatial frequency limit is

$$f_{\min} = \frac{1}{20\ \mu m} \qquad f_{\min} = 0.05\ \mu m^{-1} \tag{6}$$

and the maximum spatial frequency becomes

$$f_{\max} = \frac{1}{2\left(\frac{20\ \mu m}{512}\right)} \qquad f_{\max} = 12.8\ \mu m^{-1} \tag{7}$$

[††] A full description of the power spectral density is far beyond the scope of this paper but is discussed in more detail in section 5, *Determination of Specific Haze and Microroughness Levels*. The reader interested in a complete description is directed to the Stover reference, *Optical Scattering*.

It is also apparent from figure 4 that this instrument exhibits a pronounced roll-off in response at the low frequency limit (although part of this may be a mathematical artifact). Nonetheless, this emphasizes the importance of knowing the instrument transfer function for a given frequency within its bandwidth.

Finally, figure 5 shows the same sample surfaces as measured by a TMA CASI® angle-resolved light scattering (ARS) instrument. Details of the ARS procedure are presented in depth in section 5, *Determination of Specific Haze and Microroughness Levels*. For now, it is important to realize how the same surface can be measured by completely dissimilar instruments *and* give the same results within the same bandpass.

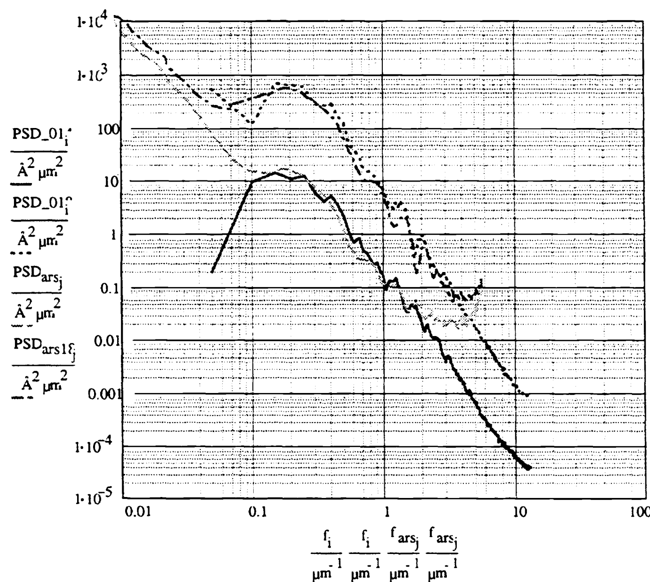

Figure 5. *PSD plots of haze and microroughness standards (2960-014 and 2960-018) employing ARS (70° angles of incidence, λ = 325 nm ending at about 5μm⁻¹) and 20μm x 20μm AFM scans (with a high spatial frequency of 12.8μm⁻¹). Notice that the PSD from ARS may be used to characterize the AFM transfer function on the low end; albeit it is important to distinguish between what is an actual instrument transfer function and what is a mathematical artifact.*

The realization of the graph being on a log-log scale emphasizes the assertions made earlier about the essential importance of knowing the spatial bandwidth limits of a given instrument. Integration between the limits of 0.01 μm^{-1} to 0.1 μm^{-1} (which may be typical for an optical profilometer) results in an rms roughness value over an order of magnitude higher as compared with a 1 μm x 1 μm AFM scan (bandwidth of 1 μm^{-1} to 256 μm^{-1}). Figure 6 depicts typical bandwidths of common surface texture instruments.

Knowing the PSD function of a given surface allows determination of rms roughness in a straightforward manner. For an isotropic surface, the rms roughness, R_q, between spatial frequencies f_{min} and f_{max}, from a one-dimensional power spectral density function, PSD_{iso}, is

$$R_q = \sqrt{\int_{f_{min}}^{f_{max}} PSD_{iso}(f)df} \tag{8}$$

3. HAZE VERSUS MICROROUGHNESS

Up to this point, the terms haze and microroughness have been used fairly interchangeably. According to the SEMATECH Chartered Working Group, *Particle Counting and Microroughness Task Force*, haze may be defined as

> *a condition of a wafer surface resulting in a significant level of light scattering. Haze may result from surface topography (microroughness) or from dense concentrations of*

surface or near-surface imperfections. Haze due to the existence of a collection of imperfections over an observable area is a mass effect; individual imperfections of the type which result in haze cannot be readily distinguished by the eye or other optical detection system without magnification.[9]

For review, microroughness was defined as

surface roughness with spacings between irregularities (spatial wavelength) less than about 100 μm.

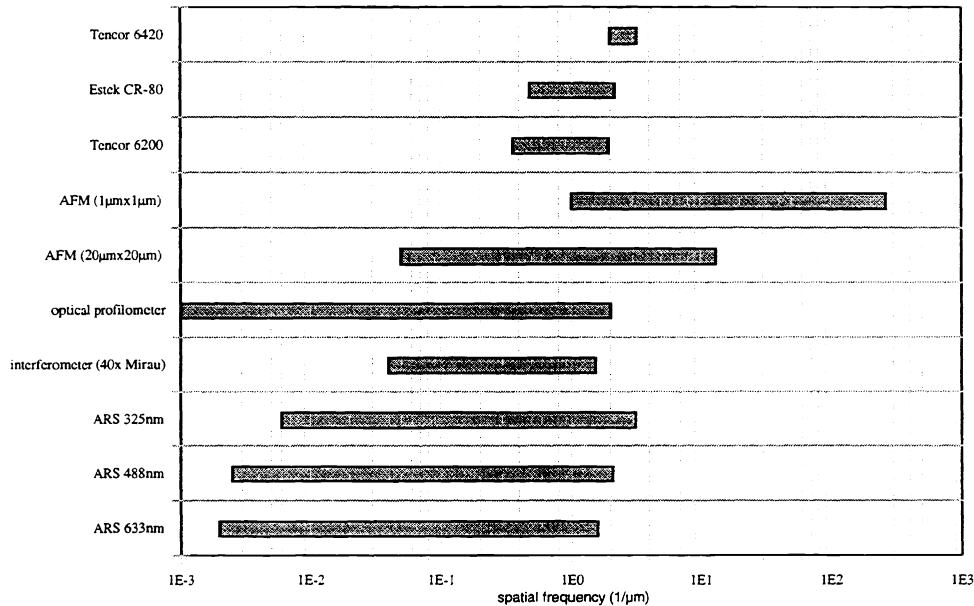

Figure 6. Bandwidths of some surface texture instruments (adapted from reference 4).

Consequently, from these definitions it is possible that effects from both the surface topography of a wafer and *near* the surface of a wafer can contribute to an increase in haze level. Subsurface effects, such as crystal originated particles (COP) and other subsurface defects are detected if within the penetration depth of the laser; the laser scanner being the typical method used for evaluating silicon wafers.

For a wafer scanning instrument, some portion of the radiation scattered from a reflecting surface is collected by the integrating optics. If it were possible to collect all of the scattered radiation over a 2π solid angle, this would be the total integrated scatter. The total integrated scatter, TIS, is the ratio of scattered to specularly reflected power and is then related to the microroughness by the following equation

$$TIS = \frac{\textbf{scattered power}}{\textbf{specularly reflected power}} = \frac{P_s}{RP_i} \cong \left(\frac{R_q 4\pi \cos(\theta_i)}{\lambda} \right)^2 \tag{9}$$

The measurement of haze then may be found directly (in parts per million) as

$$Haze = TIS \cdot R_o \tag{10}$$

where R_o is the specular reflectance as a function of wavelength, polarization, and incidence angle, and is a material property of the substrate.

4. DESIGN CONSIDERATIONS

In order to keep the individual pits (shown in figures 1 and 2) from being identified as separate scattering entities, many features must be contained within a single laser spot as it is being scanned across the surface of a wafer. A high surface density of features is also required for profiler-type systems. Contemporary laser scanners have a spatial resolution (the ability to discern two closely spaced particles) of approximately 50 μm. To allow this standard to function with future generations of scanners, it was (somewhat arbitrarily) decided to space etched features at roughly one-tenth this distance or every five microns. It was further decided to randomly locate the features within the confines of an imaginary 5 μm x 5 μm box (as viewed on the surface of the wafer). This would prevent the possibility of any unintentional pattern being displayed on the scanner. Initially, photomasks were produced in two polarities to produce both raised mesas (bright field mask) and etched pits (dark field mask). Since the results from each wafer type were identical, only the etched pits were eventually used as this was easier to produce cleanly.

Due to the extremely large number of features, it was quickly discovered that the x-y data stream was too large to be handled directly by the photomask facility (approximately 50 million bytes of information). Therefore, a separate FORTRAN routine had to be developed which would directly drive the location coordinate generation for the electron-beam pattern generator used for photomask production.

Once the mask was designed and wafers produced, it was important to verify that the pattern was uniform and isotropic. Table 1 indicates the test matrix performed on an original batch of four prototype wafers. Additionally, a 488 nm wavelength, 100 μm diameter beam at a 5° incidence angle was raster scanned in a 10 mm x 10 mm pattern centered on the wafer. Data were then taken every 200 μm to detect any fine variation in uniformity and isotropy which could not be detected in the data generated from the matrix.

wafer angle	laser angle of incidence	laser wavelength, λ	distance from wafer center	laser beam diameter
0°	5°	488 nm	0 mm	1.0 mm
0°	5°	488 nm	55 mm	1.0 mm
0°	5°	488 nm	0 mm	0.1 mm
0°	5°	633 nm	0 mm	1.0 mm
0°	45°	488 nm	0 mm	1.0 mm
0°	70°	488 nm	0 mm	1.0 mm
0°	45°	633 nm	0 mm	1.0 mm
0°	70°	633 nm	0 mm	1.0 mm
45°	5°	488 nm	0 mm	1.0 mm
90°	5°	488 nm	0 mm	1.0 mm
90°	5°	488 nm	55 mm	1.0 mm
135°	5°	488 nm	0 mm	1.0 mm

Table 1. Matrix of tests performed on four wafers utilizing angle resolved scatterometry.

All of the tests were performed to insure isotropy, uniformity, and wavelength scaling. The laser wavelengths were chosen as they are the most frequently encountered wavelengths used by state-of-the-art wafer scanners.[‡‡] Figures 7 and 8 show that the wafers were indeed uniform and isotropic.

In addition to showing isotropy and uniformity, the angle-resolved scatterometry measurements also detected a problem. The dips in the PSD curves have minima that are exactly at the points where the minima of an Airy disk diffraction pattern are located. Without going through a complete Airy diffraction derivation (which may be found in a good optics text, such as reference 10), consider a circular aperture geometry as shown in figure 9, which is reminiscent of the etched feature on the surface of the wafer. The plane wave is shown coming in from behind the aperture; but realize that since one of the wafer scanners used for testing has normal incidence that this schematic is probably quite close to the actual geometry (recall also that angle of incidence has no bearing on the PSD curve, aside from extending its high spatial frequency range).

[‡‡] There is one company which does use the 514 nm line of the argon-ion laser, but there are relatively few of these scanners in production use.

The aperture has a radius of a while the angle θ may be associated with the scattering angle in the grating equation θ_s. The quantity q/R equals $\sin(\theta)$ and the point P represents the point of the most intense irradiance. The central disk of the most intense irradiance falls to a minimum (the first dark fringe) at a distance that corresponds to the first time the Bessel function $J_1(x) = 0$, or when x = 3.832. If the distance q corresponds to the center of this first dark ring, then the first minima may be found from the expression

$$ka \sin(\theta) = 3.832 \tag{11}$$

where $k = 2\pi / \lambda$ is the wavenumber. Substituting the pit diameter for $2a$ gives the final result

$$\sin(\theta) = \frac{1.22\lambda}{d} \tag{12}$$

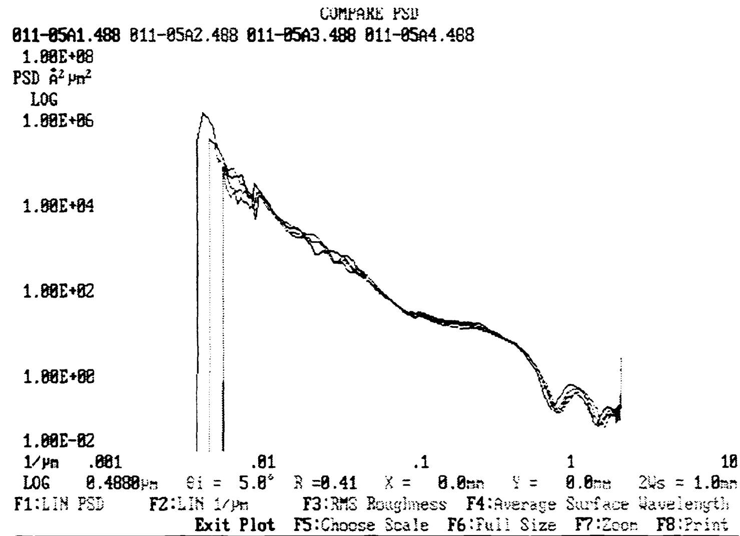

Figure 7. *Wafer number 2960-011, bright field wafer (etched mesas).*

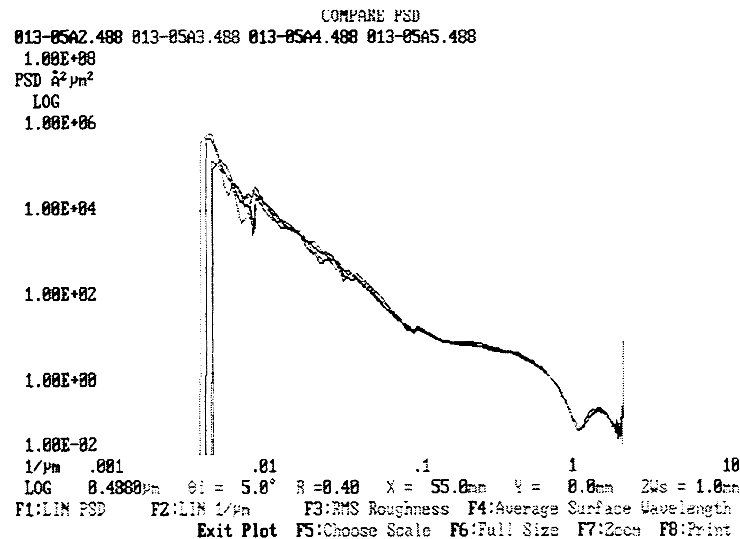

Figure 8. *Wafer number 2960-013, dark field wafer (etched pits). Both wafers were scanned at 5° incidence angle, 488 nm wavelength. The plots are from four different angular positions on the wafers. Notice almost identical response in range of interest (>0.01μm $^{-1}$) indicating isotropy.*

Subsequent minima may be found in the same way by using the appropriate zeros from the first order Bessel function. The rapid decline in irradiance may be seen in figure 10.

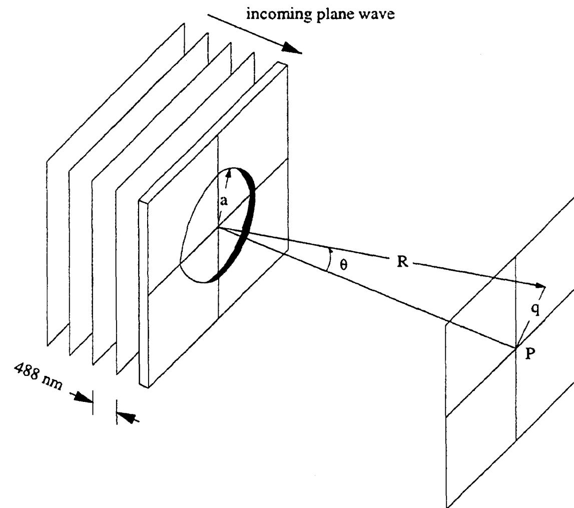

Figure 9. *Circular aperture geometry used to demonstrate the Airy disk diffraction phenomenon with the etched pits in silicon (adapted from reference 10).*

If the angle of incidence is normal to the surface of the wafer ($\theta_i = 0$), then the grating equation is simply

$$f = \frac{\sin\theta_s}{\lambda} \qquad (13)$$

Substituting this into the Airy disk equation allows for easy determination of the actual pit diameter, d, since $d = 1.22/f$. Determining the pit diameter in this way agrees well with direct measurements through atomic force microscopy.

4.1 Remedy for diffraction problem

Using this result, it is possible to emulate a pseudo-BRDF curve as a function of angle, and consequently, model an ideal PSD plot. By doing this, it became apparent that there are two potential ways to minimize the effects of the diffraction problem.

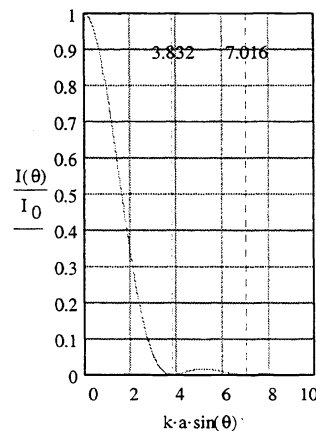

Figure 10. *Relative irradiance within an Airy disk and location of the first two minima.*

One way to reduce the PSD dips would be to simply move the first minima to a high enough spatial frequency so that it would not affect the portion of the curve of interest. Setting the spatial frequency to 3 μm^{-1} would imply a required pit diameter of

1.22/3 μm^{-1}= 0.41 μm. The lower limit of resolution for the photolithographic stepper employed is, at best, 0.8 μm which would place the frequency minima at about only 1.5 μm^{-1}. However, since the diffraction pattern is based on the Bessel function, it is possible to manufacture two pit diameters that correspond to the first roots of the first and second order Bessel functions thereby overlaying the first minima and maxima from each size. (Recall that the first derivative of a Bessel function equals the next higher order, or $dJ_i(x)/ dx = J_{i+1}(x)$). The ratio of these first minima and maxima is 1.342. Therefore, a revision was made in the FORTRAN program that drives the electron beam pattern generator to incorporate a second pit size. The mask generated from this revision produced alternating sizes of pits in the 1.342 ratio on the random vector head. Additionally, the vector head location was reduced to a maximum 4μm x 4μm random placement within the imaginary 5μm x 5μm box to greatly reduce the potential for pit overlap.

The graph in figure 11 shows the theoretical results associated with incorporating a second pit size. A comparison of figures 5 and 8 clearly shows the effect of using the two pit sizes. Figure 5 shows a much smoother PSD curve.

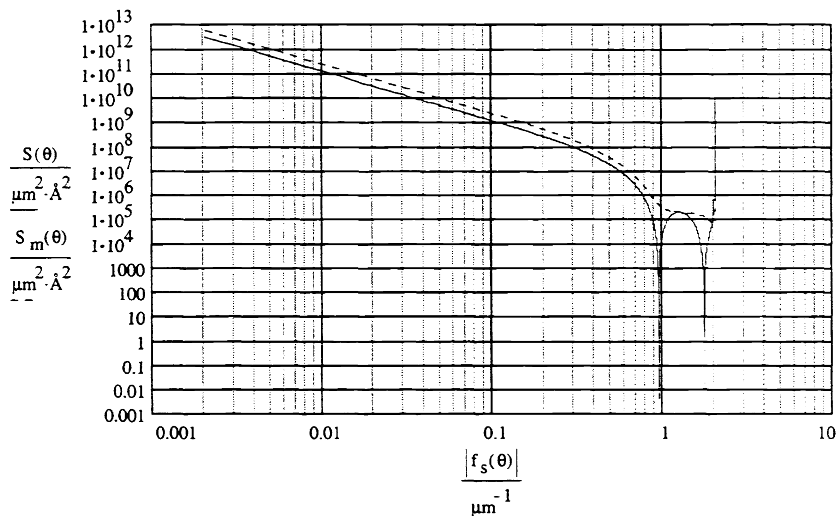

Figure 11. *Simulated response of a two-dimensional PSD from one size of pit, S(θ), versus the response from two pits, S_m(θ). The intensity scale is arbitrary. The relative minima are far more extreme than actual since no standard deviation in pit size is taken into account. The importance of this response though is the smoothing effect by using two pit diameters which are related by the ratios of the first and second order Bessel functions.*

To properly characterize the resulting patterned wafer surface, angle resolved scatterometry was employed to generate a power spectral density plot. This allowed for complete characterization far beyond the spatial frequencies of all contemporary wafer scanners and provides for NIST traceability.[§§] Since the $1/e^2$ beam diameter employed in the ARS system was quite large (typically over 1 mm), it was necessary to design the wafer pattern large enough to eliminate any repetitive pattern geometry within the field-of-view. This resulted in a total surface density of four million features per square centimeter. Figure 12 depicts a section of this layout compared with a typical wafer scanner beam diameter of 100 μm.

4.1 Concerns with thin-film effects

In order for this wafer to function as a true standard, any scattered radiation must be due to surface topography variations only. The primary calibration is accomplished utilizing laser-scattering techniques and most film and particle layers tend to show scatter signatures that vary with wavelength. The resulting PSD plot from this surface is invariant with wavelength and is therefore said to wavelength scale. Thin film effects would make wavelength scaling impossible. The reason for this

[§§] Traceability to the National Institute of Standards and Technology, NIST, will be achievable based on the release of a BRDF (Bi-directional Reflectance Distribution Function) Standard Reference Material from NIST currently scheduled for late 1996. Additionally, traceability can be obtained through a first-principles approach based on the wavelength of the laser and certain physical attributes of the system.

becomes apparent when the interaction of the incoming electromagnetic wave with the reflecting surfaces is taken into consideration.

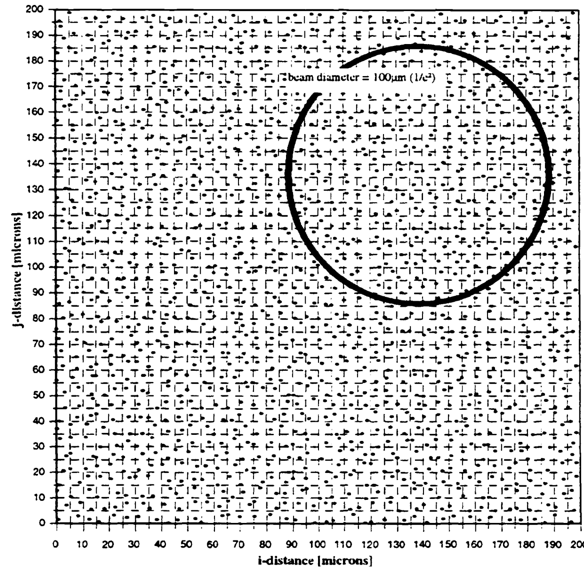

Figure 12. Representation of the placement of etched features on a section of the surface of the wafer, shown in comparison to a 100 μm diameter scanning beam.

Figure 13 illustrates the effect of multiple reflections. The haze resulting from the transparent oxide layer shown as an example in this diagram is not merely due to the surface topography alone, but also due to interference effects from the various optical path lengths and phase shifts (which change with film thickness, in addition to changes in reflectivity and transmissivity due to polarization and angle of incidence).[10,11,12]

The signal change in the measuring instrument must be due to surface topography variations only. Therefore, the haze, or microroughness, level must change but it must do so in a material that has excellent wavelength scaling characteristics.

4.2 Microroughness as a function of etch depth into silicon

Silicon has been shown to wavelength scale from 330 nm to 800 nm.[13] Utilizing the basic design and photomask as shown in figure 12, it was determined that etching the silicon wafer with a proprietary method would give the desired true topographic effect of microroughness. Then, its PSD function would be independent of the technique employed to measure it.

5. DETERMINATIONS OF SPECIFIC MICROROUGHNESS AND HAZE LEVELS

Since all of the data from angle-resolved scatterometry are binned into discrete points, the "integration" required to turn these data into an rms roughness value merely becomes the square root of the summation of the isotropic PSD function, $S_{iso}(f)$, multiplied times the differential frequency step size within the appropriate limits of integration.

$$R_q = \left(\int_{f_l}^{f_h} S_{iso}(f)df \right)^{\frac{1}{2}} \quad \Rightarrow \quad R_q = \sqrt{\sum_{i=f_l}^{f_h} S_{iso} \delta f} \qquad (14)$$

The significance of S_{iso} is discussed later in this section. For now, consider that it is a PSD function for isotropic surfaces.

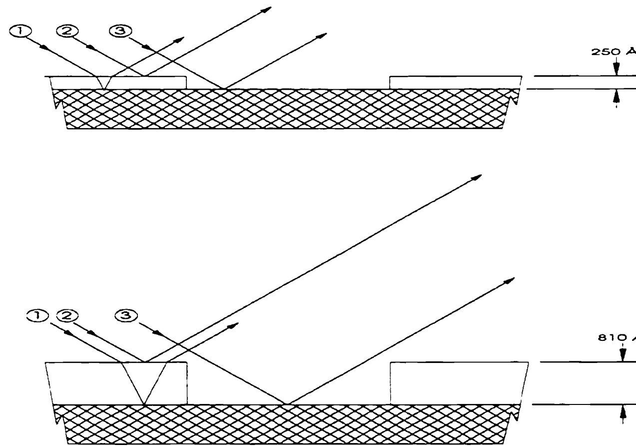

Figure 13. *Simplified drawing of light rays coming into 1) silicon dioxide, refracting, then reflecting from silicon, 2) reflecting from the silicon dioxide directly, and 3) reflecting directly from silicon. The actual optical path length (the summation of the product of refractive index and actual distance traveled) differences are shown by the relative lengths of the outgoing rays from two different thicknesses of SiO_2 (250Å and 810Å). The effects of polarization, phase, and incidence angle are not indicated here. These effects can be extremely important and must also be carefully considered. Drawing is not to scale.*

5.1 Development of the PSD calibration curve

The actual quantity measured through angle-resolved scatterometry is the Bi-directional Reflectance Distribution Function, BRDF

$$\boldsymbol{BRDF} = \frac{P_s / \Omega_s}{P_i \cdot \cos\theta_s} \tag{15}$$

where P_s is the power of the scattered light collected over the solid angle Ω_s as a function of the scatter angle θ_s and P_i is the incident laser power at angle θ_i from the wafer normal.[4] Therefore, BRDF is physically nothing more than the redistributed energy scattered into a given solid angle. Recall that the spatial frequency, f, is related to the scatter angle by the one-dimensional grating equation

$$f = \frac{\sin\theta_s - \sin\theta_i}{\lambda} \tag{16}$$

The PSD function is calculated from the BRDF; it is a measure of the scattered power per unit of spatial frequency, typically in units of Å²μm²

$$S(f_x, f_y) = \frac{10^8 \lambda^4 \boldsymbol{BRDF}}{16\pi^2 \cos\theta_i \cdot \cos\theta_s \, \mathbf{Q}} \tag{17}$$

where, for s-polarization, the polarization factor Q is approximated by the specular reflectance of the wafer surface (see section 5.2 for a more complete description of the polarization factor).*** The reflectance is a function of wavelength, incidence angle, and polarization. If the surface is isotropic, $S(f_x, f_y)$ may be integrated around the azimuthal angle to obtain the isotropic PSD function

*** Other polarization factors have been determined for out-of-plane scatter with various combinations of s- and p-source polarization, different analyzers on the detector, as well as cross-polarization situations for advanced scanning surface inspection systems. This information will be made available once the proper two-dimensional spatial bandwidths are available from the manufacturers.

$$S_{iso}(f) = \int_0^{2\pi} S(f_x, f_y) f \, d\phi_s = 2\pi f \cdot S(f_x, f_y) \tag{18}$$

with units of Å²µm and f equals the root-sum-of-squares of f_x and f_y. It is this value, $S_{iso}(f)$, that is the function from which rms roughness is calculated.

If the surface is isotropic, BRDF is independent of surface orientation. Also, $S_{iso}(f)$ is a true surface parameter and is therefore independent of incidence angle and incident wavelength.[14] These conditions have all been met and proven applicable for the Haze and Microroughness Standard.

5.2 Out-of-plane scatter and the polarization factor

For two-dimensional calculations, the form of the grating equation varies in the x- and y-directions and is given by the following

$$f_x = \frac{\sin\theta_s \cdot \cos\phi_s - \sin\theta_i}{\lambda} \tag{19}$$

$$f_y = \frac{\sin\theta_s \cdot \sin\phi_s}{\lambda} \tag{20}$$

where f_y is the out-of-incident plane scatter component and ϕ_s is the angle from the incident plane. The combined grating equation is simply

$$f = \sqrt{f_x^2 + f_y^2} \equiv \frac{1}{\lambda} \cdot \sqrt{-2\sin\theta_i \cdot \sin\theta_s \cdot \cos\phi_s + \sin^2\theta_s + \sin^2\theta_i} \tag{21}$$

In general, the dimensionless quantity Q (equation 17) is defined as a reflectivity polarization factor and predicts the effect of the surface material properties on the reflected light. The magnitude of this factor can be rather large and will vary as a function of sample complex refractive index, angles of incidence and scatter, and polarization states of the incoming and detected radiation.[6] Q is the summation of up to four possible combinations of incident and detector polarization states. Therefore, Q becomes $Q = Q_{\alpha\beta}$, where α and β refer to the incident and detector polarization. In practice, in an application with a detector insensitive to polarization, Q becomes

$$Q = \sum_{\beta} Q_{\alpha\beta} \tag{22}$$

Following the nomenclature used by Stover and Church for various source and detector polarization states, where ε is the complex sample dielectric constant ($\varepsilon = n^2$, with n being the complex index of refraction), Q_{ab} takes the form of

$$Q_{ss} = \left| \frac{(\varepsilon - 1)\cos\phi_s}{\left(\cos\theta_i + \sqrt{\varepsilon - \sin^2\theta_i}\right) \cdot \left(\cos\theta_s + \sqrt{\varepsilon - \sin^2\theta_s}\right)} \right|^2 \tag{23}$$

$$Q_{sp} = \left| \frac{(\varepsilon - 1) \cdot \sqrt{\varepsilon - \sin^2\theta_s} \cdot \sin\phi_s}{\left(\cos\theta_i + \sqrt{\varepsilon - \sin^2\theta_i}\right) \cdot \left(\varepsilon\cos\theta_s + \sqrt{\varepsilon - \sin^2\theta_s}\right)} \right|^2 \tag{24}$$

$$Q_{ps} = \left| \frac{(\varepsilon - 1) \cdot \sqrt{\varepsilon - \sin^2\theta_s} \cdot \sin\phi_s}{\left(\varepsilon\cos\theta_i + \sqrt{\varepsilon - \sin^2\theta_i}\right) \cdot \left(\cos\theta_s + \sqrt{\varepsilon - \sin^2\theta_s}\right)} \right|^2 \quad (25)$$

$$Q_{pp} = \left| \frac{(\varepsilon - 1) \cdot \left(\sqrt{\varepsilon - \sin^2\theta_s} \cdot \sqrt{\varepsilon - \sin^2\theta_i} \cdot \cos\phi_s - \varepsilon \sin\theta_i \sin\theta_s\right)}{\left(\varepsilon\cos\theta_i + \sqrt{\varepsilon - \sin^2\theta_i}\right) \cdot \left(\varepsilon\cos\theta_s + \sqrt{\varepsilon - \sin^2\theta_s}\right)} \right|^2 \quad (26)$$

6. OPTICAL SCANNING PARAMETERS OF SOME CONTEMPORARY WAFER SCANNERS

The list in table 2 accounts for many state-of-the-art domestic manufacturers of unpatterned wafer scanners. Two scanners from Tencor[†††], a 6200 and a 6420[‡‡‡], and an Estek CR-80 Wafer Inspection System® from ADE Optical Systems.[§§§] Table 3 lists the reported values of haze from these scanners, along with equivalent haze units from an AFM and the actual value of calculated haze from a CASI® scatterometer. Values are given for two levels of haze; 2960-014 is for a haze wafer etched approximately 1 nm into silicon, 2960-018 is etched 10 nm. As an example, the reported values of haze from the Tencor 6420 were 0.13 ppm and 1.07 ppm for the low and high haze values respectively. The corresponding converted values of microroughness from an atomic force microscope, taken in the same spatial bandwidth, were 0.049 ppm and 1.28 ppm. The values calculated from the PSD curve for this bandwidth were 0.133 ppm and 1.08 ppm.

model	manufacturer	wavelength	incidence angle	spatial bandpass
6200	Tencor	488 nm	0° (normal inc.)	0.36 - 1.9 μm^{-1}
6420	Tencor	488 nm	70°	2.0 - 3.2 μm^{-1}
CR-80	ADE Optical	488 nm	15°	0.49 - 2.1 μm^{-1}

Table 2. *A comparison of the optical parameters of the various surface scanning inspection systems employed.*

7. APPLICATION AND USE OF THE STANDARD

For all systems, whether a profiler used to determine rms microroughness or a surface scanning instrument used for determining haze, there is an accompanying software disk that eliminates tedious calculations. This disk contains all of the data taken by angle-resolved scatterometry for a particular standard. Also, all of the necessary governing equations are included to automatically compute the haze and/or microroughness values for either a given piece of equipment (many of the common, state-of-the-art wafer scanners are listed directly), or a given set of input parameters such as spatial bandwidth limits or traversing length and number of sample points.

7.1 Scanning Surface Inspection Systems

For this type of instrument, it is necessary to set up for a haze scan (as opposed to a particle scan). The value computed by the standard's software will indicate the proper value of haze for a given scanner. Many contemporary scanners may be calibrated directly to display the proper haze value. The equipment manufacturer can provide assistance with this adjustment. Some pieces of equipment, however, may not allow any end-user adjustments to be made. In this case, the factor between the calculated value and the instrument reading can be computed and used as a correction factor. The software will indicate what the proper value of haze should be and provide for an intercomparison between different toolsets.[****]

[†††] Tencor Instruments, 2400 Charleston Road, Mountain View, CA 94043

[‡‡‡] This is a system which has been specifically designed to reduce the effects of surface microroughness on the detection of particles. Its primary application is particle detection in the presence of high levels of haze, such as that found on blanket film layers.

[§§§] ADE Optical Systems, 9625 Southern Pine Boulevard, Charlotte, NC 28273

[****] The optical parameters of some scanners (especially older models), may not be listed on the software disk. The optical parameters may be entered directly once obtained from the equipment manufacturer.

wafer number	6200 reported haze	AFM reported haze	CASI® reported haze
2960-014	13.4 ppm	8.93 ppm	5.82 ppm
2960-018	512 ppm	436 ppm	251 ppm
wafer number	**6420 reported haze**	**AFM reported haze**	**CASI® reported haze**
2960-014	0.13 ppm	0.049 ppm	0.133 ppm
2960-018	1.07 ppm	1.28 ppm	1.08 ppm
wafer number	**CR80 reported haze**	**AFM reported haze**	**CASI® reported haze**
2960-014	12.3 ppm	4.92 ppm	3.60 ppm
2960-018	493 ppm	209 ppm	125 ppm

Table 3. Reported values of haze from various scanning surface inspection systems, AFM's (stated in equivalent haze units) and the true value as calculated from a commercial angle-resolved scatterometer.

7.2 Surface profiler systems

Typically, the limitations here involve the maximum spatial frequency, the stitching of adjacent die, and the necessity for a square area scan. The first parameter is important due to the restrictions imposed by the shortest wavelength used in the calibration of the Haze and Microroughness Standard. The maximum spatial frequency displayed on the PSD curve is related to the one-dimensional grating equation. This number is approximately $2/\lambda$. For the 325 nm Helium-Cadmium laser used for calibration, the maximum theoretical spatial frequency is about 6 μm^{-1} (due to molecular scattering at high scattering angles, a practical spatial frequency limit is 3.5 μm^{-1}) . Therefore, traversing lengths of 75 µm x 75 µm or larger would typically be employed.[††††] Again, the software determines these numbers automatically based on the traversing length and the number of sample points. If these parameters fall outside of this range, the user is prompted to use new input parameters.

7.2.1 Optical profilometers

Optical profilers require a slightly different set of input parameters: the total scan length used for a given scan, the pixel size employed by the system, and the total magnification.[‡‡‡‡]

The total scan (or traversing) length may be found from the other two parameters. The ratio of the number of pixels multiplied times the pixel size to the total magnification will produce the scan length. This same ratio, when used with the minimum objective and minimum zoom magnifications (reduction),[§§§§] will also yield the maximum traversing length possible for a given system.

$$\textbf{total traversing length} = \frac{\textbf{number of pixels} \times \textbf{pixel size}}{\textbf{total magnification}} \tag{27}$$

The low spatial frequency then is equal to this value.

$$f_l = \frac{1}{\textbf{total traversing length}} \tag{28}$$

The high spatial frequency may be determined similarly, using the Nyquist factor of 2 in the denominator.[*****] Notice that this becomes essentially the same equation used for determining the upper spatial frequency for an AFM.

---

†††† The spatial bandwidth of a 75 µm x 75 µm scan with 512 sample points is 0.013 μm^{-1} to 3.4 μm^{-1}. Smaller scans such as 10 µm x 10 µm (0.1 μm^{-1} to 25.6 μm^{-1}) may be employed on systems that have the capability of moving cursors to a more narrowly defined range with a new resulting rms roughness value. This also eliminates the error due to instrument fall-off at the scan extremes. The instrument transfer function approaches unity when a narrower subset of the scan range is used.

‡‡‡‡ Total magnification comes from the objective magnification multiplied times the image zoom magnification (if present). For example, if a 40x objective lens is used with a 2x zoom magnification, the total magnification will be 80x.

§§§§ As an example, a zoom feature on an optical profilometer may have a range of 0.5x to 2x. The low end of this range is actually a reduction then and will contribute toward allowing the greatest scan length.

***** In some cases, the lower limit of resolution may not be as low as this value would seem to indicate. This is due to the limit of resolution, L_r, of the objective lens and is usually indicated in the specifications for that lens (typically determined by either the Rayleigh or Sparrow criterion). In that case, $1/L_r$ would be the high spatial frequency.

$$f_h = \frac{1}{2 \cdot \left(\frac{\textbf{pixel size}}{\textbf{total magnification}} \right)} \equiv \frac{1}{2 \cdot \left(\frac{\textbf{total traversing length}}{\textbf{number of pixels}} \right)} \quad (29)$$

Be certain that the area scanned is approximately a square area (the same traversing lengths in both x- and y-directions). This assures that the two-dimensional to isotropic PSD function conversion is valid. As with other instruments, make certain to either turn off all filtering, or be certain that no filtering occurs greater than 0.01 μm^{-1} as this would tend to further alter the transfer function.

8. SUMMARY

A practical haze and microroughness standard has been developed that is based on the principles of optical scatterometry for certification. This standard allows for comparison between seemingly uncorrelatable tools, within the proper spatial bandwidth for each toolset. Using this methodology, it is possible to compare haze readings on a scanning surface inspection system to the microroughness readings on an atomic force microscope.

It is expected that an improved correlation between various systems will result after a step height standard in the 1 nm range is developed for AFM z-height calibration (microroughness is proportional to z^2). There is currently no certified step height available from either VLSI Standards or NIST. Also, once scanning systems are calibrated for haze, there is an expected improvement in correlation there as well.

9. ACKNOWLEDGMENTS

Many individuals have contributed directly to the work described in this article. I wish to give special acknowledgment to members of the SEMATECH Particle Counting and Microroughness Task Force; especially Ed Bawolek, Murray Bullis, Randy Goodall, Dan Hirleman, and John Stover for their valuable input over the years. Additionally, Paul DePesa and Ric Diola of Diamon Images were particularly helpful in providing solutions for the mask generation problems. Finally, a special thanks to Jim Greed, Ellen Laird, Darcy Rich, and Craig Scheer of VLSI Standards for assistance in prototype development and characterization.

10. REFERENCES

[1] "Standard Terminology of Silicon Technology", ASTM Designation F 1241, American Society for Testing and Materials, 100 Barr Harbor Drive, West Conshohocken, PA 19428, 1996.

[2] Semiconductor Industry Association, *National Technology Roadmap for Semiconductors*, June 1994.

[3] Fink, Donald G., and Donald Christiansen, eds. *Electrical Engineering Handbook*, Third Edition, McGraw-Hill, New York, 1989.

[4] Bullis, W. Murray, "Microroughness of Silicon Wafers", *Semiconductor Silicon/1994*, PV 94-10, The Elcctrochemical Society, Pennington, NJ, 1994, pgs. 1156-1169.

[5] Bullis, W. Murray, "Characterization of Haze and Microroughness of Silicon Wafers", *MICRO*, January 1996, pgs. 47-53.

[6] Stover, J.C., *Optical Scattering Measurement and Analysis*, SPIE Optical Engineering Press, Bellingham, WA, 1995.

[7] ANSI/ASME Standard B46.1, *Surface Roughness, Waviness, and Lay*, 1993.

[8] Bennett, Jean M. and Lars Mattsson, *Introduction to Surface Roughness and Scattering*, Optical Society of America, Washington, D.C., 1989.

[9] "Glossary of Terms Related to Particle Counting and Surface Microroughness of Silicon Wafers", *SEMATECH Technology Transfer 95082941A-TR*, September 30, 1995.

[10] Hecht, E., *Optics*, Second Edition, Addison Wesley, Reading, Mass., 1987.

[11] Cheng, D.K., *Field and Wave Electromagnetics*, Second Edition, Addison Wesley, Reading, Mass., 1989.

[12] Born, M. and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Sixth Edition, Pergamon Press, Oxford, 1980.

[13] Stover, J. C., "Microroughness and Surface Scatter -- The Scattering Characteristics of Silicon", *Particles, Haze and Microroughness on Silicon Wafers Symposium*, San Jose, CA, September, 1994.

[14] Church, E.L., "Microroughness and Surface Scatter -- Practical Profile Measurements and Finish Parameters", *Particles, Haze and Microroughness on Silicon Wafers Symposium*, San Jose, CA, September, 1994.