Characterization of random rough surfaces by in-plane light scattering

Y.-P. Zhao, a) Irene Wu, C.-F. Cheng, Ueyn Block, G.-C. Wang, and T.-M. Lu
Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180-3590

(Received 23 December 1997; accepted for publication 8 June 1998)

The reciprocal space structures of Si backside rough surfaces were studied using an in-plane light scattering technique. A diode array detector was used to collect the angular distribution of the scattered intensity. The diffraction profiles are interpreted based on the theory of diffraction from a self-affine rough surface. Roughness parameters such as interface width, lateral correlation length, and the roughness exponent, are extracted from the profiles and are compared to that obtained by real space imaging techniques such as atomic force microscopy and stylus profilometry. Factors that limit the accuracy of light scattering measurements are discussed. © 1998 American Institute of Physics. [S0021-8979(98)08817-3]

I. INTRODUCTION

The characterization of random rough surfaces is of considerable interest in a variety of technical applications. For example, the surface of a thin film is often rough due to thermal fluctuations and lack of surface mobility during the deposition or etching as a result of changing substrate temperature, pressure and growth/etch rate. Experimentally, the most direct way to characterize a random rough surface quantitatively is to measure the surface morphology using real space imaging techniques, such as stylus profilometry (SP) and scanning probe microscopy (SPM). The other popular technique is diffraction. The measurement of scattered radiation as a method for roughness characterization has the advantage that it is nondestructive, and sometimes can be used as a real-time monitoring tool.

Light scattering has been used for a long time to characterize rough surfaces. Most theories assume an isotropic, homogenous, Gaussian height distributed surface, with either a Gaussian or an exponential autocorrelation function. Recently, considerable advancement in the understanding of the nature of rough surfaces has been made based on the concept of self-affinity. A self-affine surface is a class of fractal object that can be described by a ‘‘roughness exponent’’ which is related to the fractal dimension of the surface. Following the diffraction theory developed for the self-affine surface, it has been shown that it is possible to measure the fractal properties of the rough surface, including the determination of the roughness exponent, using light scattering. In this paper, we report a detailed study of the diffraction characteristics of several rough Si backside surfaces using the in-plane light scattering technique. The complete reciprocal space structure of the self-affine rough surface was mapped out by analyzing the angular distribution of the scattered intensity. The surface roughness parameters, including the interface width, lateral correlation length, and the roughness exponent, extracted from the light scattering data were compared with those obtained by real space imaging techniques, atomic force microscopy (AFM) and SP.

II. SELF-AFFINE ROUGH SURFACES

A self-affine, homogeneous and isotropic rough surface can be described using a height–height correlation function defined as

$$H(r) = \langle [h(r)-\bar{h}]^2 \rangle = 2w^2f\left(\frac{r}{\xi}\right).$$

Here $h(r)$ is the surface height at position $r$ on the surface, $\langle \cdots \rangle$ denotes the assembly average. $w = \sqrt{\langle [h(r)-\bar{h}]^2 \rangle}$ is called the interface width, and $\bar{h}$ is the average surface height. The scaling function $f(x) = x^{-\alpha}$ for $x \ll 1$, and $f(x) = 1$ for $x \gg 1$. $\alpha$ is called the roughness exponent (0 $\leqslant \alpha \leqslant 1$), which describes how wiggly the surface is. For a surface with a Gaussian or exponential autocorrelation function [defined as $\langle h(r)h(0) \rangle$], the value of $\alpha$ is 1 and 0.5, respectively. $\xi$ is the lateral correlation length, within which two point surface heights are correlated. In principle, these three parameters are independent, and vary according to the surface manufacturing process. These three parameters, $w$, $\xi$, and $\alpha$, completely characterize the statistical properties of the self-affine surface. For a real space imaging technique, these parameters can be obtained directly from the experimentally measured $H(r)$.

In diffraction, a random rough surface under light illumination can be treated as a continuous rather than a crystalline surface because $\lambda$, the wavelength of incident light, is much larger than the individual atoms and the discrete atomic effect is negligible. The distribution of the scattered intensity is proportional to $S(k)$, the scattering structure factor, where $k$ is the momentum transfer. The scattering structure factor contains information on the morphology of the surface. For an infinitely flat surface, $S(k) \propto \delta(k_\parallel)$, where $k_\parallel$ is the momentum transfer parallel to the surface and it contains two components, $k_x$ and $k_y$. The delta function is a measure of the flatness of a surface in the large scale.

For a rough surface, the scattering profile contains a sharp $\delta$ peak since the surface is flat on the long-range scale.

a)Electronic mail: zhao@rpi.edu
saturate the photodiode array. The sample was mounted on a rotational stage to allow for a variation of incident angle \( \theta \), measured with respect to the surface normal.

Two configurations were used in light scattering as shown in Figs. 1(a) and 1(b). The first one [Fig. 1(a)] shows that the scattered light from the surface passed through a lens with a focal length \( f \) of 17.5 cm, and the detector was set up at the focal plane of the lens so that parallel beams entered the detector. The second setup [Fig. 1(b)] shows that the lens was placed in front of the sample, and the detector was also located at the focal plane of the lens. That is, the sum of distances from the lens to the sample and the sample to the detector in Fig. 1(b) equals the distance from the lens to the detector in Fig. 1(a). In both configurations the detector array was positioned for in-plane scattering measurement [Fig. 1(c)]. The length \( L \) of the diode array was 2.5 cm and it contained 1024 photodiodes. The separation between two adjacent diodes was \( L/1024 \), or approximately 25 \( \mu m \). The resolution of the detector in \( k \) space was \( 25 \times k_0/f \approx 1.4 \times 10^{-3} \mu \text{m}^{-1} \) for the first setup, where \( k_0 = 2 \pi/\lambda \). The resolution of the second setup depends on the distance between the sample and the detector. The advantage of the second setup is that one can change the measured \( k_0 \) range by changing the distance between the lens and the sample. The length of each diode perpendicular to the direction of the array was 0.25 cm, which spans a \( k \) space of \( 1.4 \times 10^{-1} \mu \text{m}^{-1} \). Therefore each diode acts as a slit detector. Since a slit detector was assumed, the measured intensity is a one-dimensional integrated intensity, along the diode array. In Appendix A, we show mathematically the scattering profile of the measurement using a slit detector.

The momentum transfer parallel to the surface, \( k_i = k_x \approx k_0 \tan \gamma \), where \( \gamma \) is the in-plane scattering angle and the \( \gamma \) is small (see the appendix). Also, the momentum transfer perpendicular to the surface, \( k_z = k_0 \cos \theta \gamma + k_0 \cos \theta \approx 2k_0 \cos \theta \), where \( \theta \) is the angle of incidence with respect to the surface normal. The range of \( k_0 \) covered by the diode array is about \( k_0 \times L/f = 1.4 \mu \text{m}^{-1} \). This range determined the size of the window in the profile measurement. The diode array could be moved along the in-plane direction to increase the range of \( k_0 \) and therefore increased the window size. Each scattering profile was obtained within 5 ms using a diode array detector.\(^5\) Fifty thousand accumulations were made for each plot of the angular profile in order to gain a high signal-to-noise ratio.

### III. EXPERIMENTS

The real space characterizations were carried out by both AFM and SP techniques. AFM scans were measured using a Park Scientific Instruments Auto Probe CP with Si$_3$N$_4$ tips. The typical radius and the tip side angle were about 10 nm and 10\(^\circ\), respectively. The SP was an alpha-step profilometer. The tip size of the SP is 5 \( \mu m \). The samples were the back-sides of Si (100) wafers.

In the light scattering (LS) experiment (Fig. 1), the incident light was a helium–neon laser with a wavelength of 632.8 nm. A spatial filter was used in order to improve the quality of the incident beam. Neutral density filters were also used to reduce the intensity of the light so that it would not

and a broad diffuse region reflecting the roughness on the short-range scale. The scattering structure factor can be written as

\[
S(k) = (2\pi)^2 e^{-k_i^2 w^2} S_0(k_i) + S_{\text{diff}}(k_\parallel, k_\perp),
\]

where

\[
S_{\text{diff}}(k_\parallel, k_\perp) = \int \int d^2 r (e^{-k_\parallel^2 H(r)/2} - e^{-k_\parallel^2 w^2}) e^{ik_\parallel \times r},
\]

for a homogeneous and isotropic rough surface with a Gaussian height distribution.

### IV. RESULTS

#### A. Real space measurement

In Fig. 2 we show the AFM surface images (100 \( \mu m \times 100 \mu m \)) and the corresponding power spectra for three different Si(100) backside samples with different interface widths \( w \). These samples are labeled as samples No. 3, No. 5, and No. 10 in Table I. Obviously these three samples have different morphologies. Both sample No. 3 [Fig. 2(a)] and sample No. 5 [Fig. 2(b)] have smaller features compared to sample No. 10 [Fig. 2(c)]. The power spectra show a circular symmetry, which implies that these samples are iso-
tropic (rotational invariance). In order to obtain statistical properties of the surfaces, we plot the height histogram in Fig. 3. The scattered data points with different symbols represent the height histograms from AFM images collected at different randomly chosen positions of the same sample. The solid curves shown in Fig. 3 are the best Gaussian fits for the height distribution. The width of the histogram is proportional to the interface width \( w \) of the surface. Both sample Nos. 3 and 5 are very close to a Gaussian distribution. The height histogram for sample No. 10 deviates from a Gaussian distribution. This may be because the surface feature lateral correlation length is large and the area scan (100 \( \mu \)m \( \times \) 100 \( \mu \)m) is not large enough to include sufficient statistical averaging. Overall, the assumption of a Gaussian height distribution appears to be quite adequate. Also Fig. 3 shows that for the same sample the height histograms measured at different positions are very close to each other. This means that for the same sample statistically, the height distribution does not change from place to place. The height distribution is uniform over the entire sample. Another important test is to check the homogeneity of the samples, i.e., the translational invariance of statistical properties of the samples. For Nos. 3 and 5 are very close to a Gaussian distribution. The height histogram for sample No. 10 deviates from a Gaussian distribution. This may be because the surface feature (lateral correlation length) is large and the area scan (100 \( \mu \)m \( \times \) 100 \( \mu \)m) is not large enough to include sufficient statistical averaging. Overall, the assumption of a Gaussian height distribution appears to be quite adequate. Also Fig. 3 shows that for the same sample the height histograms measured at different positions are very close to each other. This means that for the same sample statistically, the height distribution does not change from place to place. The height distribution is uniform over the entire sample. Another important test is to check the homogeneity of the samples, i.e., the translational invariance of statistical properties of the samples. For
a rough estimate, in Fig. 4 we plot the height–height correlation functions from different sampled images versus the average height–height correlation function for sample Nos. 3, 5, and 10. For sample Nos. 3 and 5, the plots of the sampling height–height correlation functions are very close to the straight line $x = y$, which means statistically these two functions are the same. For sample No. 10, the plots of the sampling height–height correlation functions deviate from the line $x = y$ in various degrees. This could be due to the fact that the scan area is not large enough to include sufficient statistical averaging. The large deviation at large values is expected due to the finite size of the sampling image, where the height–height correlation function at large distance is estimated from a small number of sampling points. Figure 4 shows that the height–height correlation functions of the samples we studied are only functions of the distance between two separate surface points, and not functions of the surface positions, i.e., the samples are homogeneous. Therefore, the diffraction theory mentioned in Sec. II is applied to the rough surface characterization.

As discussed above, the height–height correlation function is a second order statistical measurement of the surface from which the interface width $w$, the lateral correlation length $\xi$, and the roughness exponent $\alpha$, can be determined. We plot the height–height correlation functions for sample Nos. 3, 5, and 10 in Fig. 5. We show that sample No. 10 has the largest interface width and lateral correlation length among these three samples. However, the roughness exponents for all three samples are very close, judging from the slopes of the log–log plots at the small $r$ region. To avoid any subjectivity for determining those roughness parameters, we use a phenomenological scaling function of the form

$$f(x) = \left[1 - e^{-(r/\xi)^{2\alpha}}\right]$$

(4) to fit the height–height correlation data. The results are shown in Table I. Table I also gives the interface width measured by SP. All of these measured roughness exponents fall in between 0.78 and 0.99. This suggests that the dynamics of roughness formation may be quite similar during the manufacturing of the Si wafers.

B. Light scattering and the reciprocal space structure

In Figs. 6, 7, and 8, we plot the angular dependent scattering profiles with different incident angles using the first setup [Fig. 1(a)] for sample Nos. 3, 5, and 10, respectively. One can observe some very noticeable characteristics. (1) For very large incident angles, the scattering profile from each sample contains two parts: a sharp, central ($\delta$-like) peak intensity, and a diffuse profile. (2) For the same sample, as the incident angle decreases, the central sharp intensity gradually decreases until it totally disappears. At the same time, the diffuse profile broadens. (3) The larger the interface width, the larger the incident angle at which the central peak disappears. All these features are actually predicted by the diffraction theory.

One can fit the scattering profiles using a narrow Gaussian peak (which corresponds to the $\delta$ peak intensity convoluted with the instrument broadening) and a broad diffuse intensity using Eq. (3). The full width at half maximum (FWHM) of the diffuse profile as a function of $k_\perp$ can then be obtained by assuming the functional form of Eq. (4). In Fig. 9 we plot the measured reciprocal space structure for these three samples based on the value of the FWHM extracted from angular profiles in Figs. 6, 7, and 8, as a function of $k_\perp$. The circles denote the positions of the half maximum of the diffuse intensity. The heavy lines at $k_\perp = 0$
FIG. 6. The $k_{\parallel}$ dependent light scattering profiles at different incident angles for sample No. 3.

FIG. 7. The $k_{\parallel}$ dependent light scattering profiles at different incident angles for sample No. 5.
represent the sharp, central peaks in the intensity profiles. These plots summarize the diffraction characteristics of the rough surfaces. In contrast to a crystalline rough surface, there is no periodic structure along $k'$. For both sample Nos. 3 and 5, in the small $k'$ region, the FWHM is approximately constant, and is independent of $k'$. This behavior was actually predicted from Eq. (3), which can be written as $8-13$.

\[ S_{\text{diff}}(k_\parallel,k_\perp) = 2\pi(k_\perp w \xi)^2 e^{-\xi^2 w^2} F(k_\parallel \xi), \]

for $\Omega = (k_\perp w)^2 \ll 1$, where

\[ F(\gamma) = \int_0^\infty x[1-f(x)]J_0(xy)dx. \]

Here $J_0$ is the zeroth-order Bessel function. The FWHM of the function given by Eq. (5) is proportional to $1/\xi$, independent of $k_\perp$, for sufficiently small values of $k_\perp$. In fact, the right hand side of Eq. (5) is proportional to the power spectrum of the rough surface.\(^9\,\,18\)

For sample No. 10, the FWHM in the small $k_\perp$ range does not have a constant value along $k_\perp$. This is because sample No. 10 has a large interface width $w$. The product $\Omega = (k_\perp w)^2$ is actually quite large ($\sim 1$) even for small values of $k_\perp$. Therefore in general the scattering profile depends not only on the scattering condition (such as the incident angle) but also on the degree of surface roughness (the value of interface width).

The FWHM of the scattering profiles increases as we increase the value of $k_\perp$. For $\Omega \gg 1$, $S_{\text{diff}}(k_\parallel,k_\perp) \propto (k_\perp^{-1/\alpha} \eta)^2 F(k_\parallel k_\perp^{-1/\alpha} \eta)$, where $\eta = w^{1/\alpha}/\xi$ is a very important parameter characterizing the short-range properties of a rough surface, and is proportional to the average terrace size in the case of crystalline rough surfaces.\(^11\) The FWHM increases as a function of $k_\parallel$ and is given by

\[ \text{FWHM} \propto \eta^{-1/k_\parallel^{1/\alpha}}. \]

The shapes of the diffuse profiles given by Eqs. (5) and (7), which represent $\Omega \ll 1$ and $\Omega \gg 1$ cases, respectively, are the same except for the value of the FWHM.

C. Determination of the roughness parameters

1. Interface width $w$

Conventionally the interface width $w$ can be determined through the normalized $\delta$ peak intensity, $R_\delta$,\(^10\)

\[ R_\delta = \frac{\int d^2k_\parallel d^2k_\perp \mathcal{I}(k)}{\int d^2k_\parallel d^2k_\perp \mathcal{I}(k)} = e^{-\Omega} = e^{-(k_\parallel w)^2}. \]

In principle, one scattered intensity profile is sufficient to determine $w$. A more reliable value of $w$ can be obtained by plotting $\ln(R_\delta)$ vs $k_\parallel^2$. The slope is equal to $-w^2$. In Fig. 10 we compare the interface width $w$ calculated from the normalized $\delta$ peak light scattering intensity from one profile with that measured by real space AFM and SP techniques for ten different samples with different values of interface width. The numerical values of roughness parameters are listed in Table I. The dashed line represents the case where the light
scattering and the real space imaging techniques give exactly the same results. Overall, different measurement techniques are seen to give similar results. For small interface widths, the LS results and SP measurement agree very well, but AFM gives a larger value. AFM and SP give somewhat scattered values in the larger interface width regime.

The difference in the value of the interface width obtained from different measurements may originate from different instrumental limitations. Theoretically, \( w \) should be obtained from an infinite sampling area with an infinitely high resolution. However, practically all measurements have certain limitations, including the instrumental resolution and the sampling size. Often, the product of the resolution and the sampling size is a constant. For the AFM measurement, the scan size is 100 \( \mu \)m, and the scan pixel increment is 0.39 \( \mu \)m. These values determine the spatial frequency \( (= 2 \pi / l) \), where \( l \) is a real space distance) range for AFM measurement, from \( 6.28 \times 10^{-2} \) to \( 1.6 \times 10^{1} \) \( \mu \)m\(^{-1}\). However, the tip size of the AFM is very small, the actual spatial resolution is better than 0.39 \( \mu \)m. For the SP measurement, the scan size is 500 \( \mu \)m, and the tip size is 5 \( \mu \)m, which corresponds to a spatial frequency region from \( 1.25 \times 10^{-2} \) to \( 1.25 \) \( \mu \)m\(^{-1}\). For light scattering using a detector array, as discussed in Sec. III, the spatial frequency ranges from \( 1.4 \times 10^{-3} \) to \( 1.4 \) \( \mu \)m\(^{-1}\). Table II summarizes the spatial frequency regions for LS, AFM, and SP, and Fig. 11 illustrates these regions used in our experiments. Usually, as long as the minimum spatial frequency is far below \( 2 \pi / \xi \), the error of the measured \( w \) due to the uncertainty in the lower frequency region can be neglected. This is the case for our sample Nos. 1–6 \((2 \pi / \xi \approx 1.26 \) \( \mu \)m\(^{-1}\)). Therefore the maximum spatial frequency, e.g., the spatial resolution, determines the region where the measurement gives an accurate value of \( w \). Both LS and SP have compatible maximum spatial frequency. However the AFM has an order of magnitude higher frequency than those two. Although the \( w \) values for LS and SP

---

**TABLE II. Spatial frequency regions for LS, AFM, and SP for the present experiments.**

<table>
<thead>
<tr>
<th></th>
<th>Low spatial frequency cutoff (( \mu )m(^{-1}))</th>
<th>High spatial frequency cutoff (( \mu )m(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS</td>
<td>( 1.4 \times 10^{-3} )</td>
<td>1.4</td>
</tr>
<tr>
<td>AFM</td>
<td>( 6.28 \times 10^{-2} )</td>
<td>( 1.6 \times 10^{1} )</td>
</tr>
<tr>
<td>SP</td>
<td>( 1.25 \times 10^{-2} )</td>
<td>1.25</td>
</tr>
</tbody>
</table>
are well matched, AFM gives a higher value. In contrast, as seen from Fig. 10, for large interface widths (sample Nos. 7–10, with $2\pi/\xi=6.28\times10^{-1} \text{m}^{-1}$), SP gives higher values than that of the AFM because in this case the low spatial frequency part contributes more than the high spatial frequency one does. Note that, although LS has a lower minimum spatial frequency than that of SP, the measured $w$ value is much smaller than that measured from SP. This may due to the shadowing effect in the LS measurement, which we will discuss later in Sec. V D.

One can also determine the roughness parameters by the inverse Fourier transform (IFT) of Eq. (3).\textsuperscript{19} Experimentally the height–height correlation function can be determined by taking the Fourier transform of the scattered intensity profiles. Examples will be given later in Sec. IV C 3. In Fig. 12 we compare the interface width determined by the IFT method with that obtained from the normalized $\delta$ peak intensity. There is a finite window size effect associated with this technique and will be discussed in Sec. V A.

2. Lateral correlation length $\xi$

In the light scattering, the lateral correlation length $\xi$ is inversely proportional to the FWHM under the condition $\Omega \ll 1$. In Fig. 13 we show a comparison of the lateral correlation length $\xi$ determined by light scattering and that determined by the AFM technique. Again, data from ten samples were used. The $\xi$ determined through the FWHM of LS appears to agree well with the AFM measurements, but the $\xi$ determined by the IFT technique seems to overestimate the values. The correlation length determined by SP (not shown in Fig. 13) is considerably greater than that obtained by both AFM and LS techniques. We will discuss this discrepancy in Sec. V A.

3. Roughness exponent $\alpha$

One way to determine $\alpha$ from light scattering is to go to $\Omega \gg 1$, and plot the FWHM vs $k_{\perp}$ in the log–log scale ($\text{FWHM} \propto k_{\perp}^{1/\alpha}$).\textsuperscript{10,14} In the present experiment, the largest $k_{\perp}$ that can be reached is $1.9\times10^{1} \text{m}^{-1}$. Therefore the condition $\Omega \gg 1$ can be satisfied only for samples with large interface widths. In Fig. 14 we plot the FWHM vs $k_{\perp}$ for sample No. 10 which has a large interface width close to 0.6 $\mu$m. The $\alpha$ value obtained from the slope of this plot is about 0.89$\pm$0.02, which is consistent with that obtained by AFM, 0.87$\pm$0.02. However, $\alpha$ can be determined by the IFT method without any restriction on the value of $\Omega$.\textsuperscript{19} In Fig. 15 we show the height–height correlation functions determined from the light scattering profiles of sample Nos. 1 and 8. For sample No. 8, the interface width is large, and both AFM and SP give higher $w$ values compared with IFT. But the lateral correlation lengths obtained by AFM and IFT are almost the
same, and are smaller compared with SP. The \( \alpha \) values determined by IFT, AFM, and SP are \( 0.87 \pm 0.02, 0.89 \pm 0.02, \) and \( 0.91 \pm 0.02 \), respectively. The spread of \( \alpha \) is within 1%.

For sample No. 1, the interface width is very small. The \( w \) value determined from light scattering and SP is the same because for large \( r \) the height–height correlation functions overlap. But AFM gives a higher \( w \) value compared with that of LS and SP. Both AFM and light scattering give a similar correlation length, which is smaller than that of the SP measurement. The \( \alpha \) values determined by IFT, AFM, and SP are \( 0.80 \pm 0.02, 0.78 \pm 0.02, \) and \( 0.88 \pm 0.02, \) respectively. In Table I we summarize the roughness parameters determined by different techniques.

Another strategy to determine the value of \( \alpha \) is from the power law behavior at large \( k_i \) in the power spectrum, that is, the light scattering profiles measured under the condition \( \Omega < 1^{9,20} \). In order to obtain the higher \( k_i \) value, or the tail part of the scattering profiles, we used the second setup shown in Fig. 1(b). The detector was moved along the in-plane direction by a one-dimensional (1D) translator with \( \pm 1 \) \( \mu \)m accuracy. In Fig. 16 we show the measured log–log plot of the scattering profile of sample No. 5. In the same graph we also plot the surface power spectrum (1D) calculated from the AFM image for sample No. 5. In order to compare the light scattering profiles and AFM power spectrum, we rescaled the power spectrum to match the intensity of the scattering profile. The scattering profile and the power spectrum appear to be consistent with each other. The slope of the curves at large \( k_i \) gives \(-1/(1 + \alpha)\) (see the appendix). From these plots, the values of \( \alpha \) were extracted to be \( 1.10 \pm 0.02 \) and \( 1.08 \pm 0.02 \) for power spectrum and the light scattering profile, respectively. These numbers are not quite consistent with those obtained by other methods shown above. This discrepancy is due to the inconsistency of the definition of \( \alpha \) derived from the height–height correlation and power spectrum when \( \alpha \) is close to 1.\(^{21} \) It can also be shown that in general the slope of the scattering profiles is equal to \(-1/(d + 2 \alpha)\) for any diffraction angle, or for any value of \( \Omega.^{22,23} \) Here \( d + 1 \) is the dimension of the imbedded space. Therefore one can extract the value of \( \alpha \) from the tail of the profiles obtained at any scattering condition, not just in the small \( \Omega \) regime.

V. DISCUSSIONS
A. Limits on the determination of the roughness parameters

We have discussed how the spatial frequency window affects the determination of surface roughness parameters. For light scattering, the spatial frequency window originates from two sources: one is the detector and the other is the physical limit (the Rayleigh criterion) in the measurement.
The largest translation distance in the measurement (which determines the range of \( k_1 \) covered in the measurement) determines the lower cutoff of the spatial frequency. The distance between the two adjacent detectors in the detector array determined the upper cutoff of the spatial frequency. Both cutoffs are inversely proportional to the focal length \( f \) of the lens. There is another lower spatial frequency limit determined by the incident laser beam size. If we assume that the diameter of the laser beam is \( D \), this would give a lower spatial frequency cutoff at \( 2\pi/D \) as demonstrated by Church.\(^{24}\) Also the higher frequency cutoff cannot be extended to infinity due to the Rayleigh criterion, which states that the optical resolution cannot be less than \( \lambda/2 \). This gives the ultimate high frequency cutoff at \( 4\pi/\lambda \). Therefore, the roughness parameters determined by light scattering must have some limits due to the finite frequency window. If a surface has a correlation length \( \xi \), and \( 2\pi/\xi \) is beyond the frequency window of the light scattering setup, then the measurement would not give correct values of the roughness parameters based on the scattering profile analysis.

For the interface width, the range in which a reliable measurement can be made is determined by the dynamic range given by the detector and the scattering geometry. The interface width can be determined only if there is an obvious \( \delta \) peak that appears in the diffraction profile. For our detector, the dynamic range is \( 1 \sim 60 \times 10^3 \) counts. Assuming that the \( \delta \) peak collected by one pixel is 10 counts higher than the diffuse profile, and the FWHM of the diffuse profile extends to the full range (1024 pixels) of the detector array, then roughly the smallest value of the \( \delta \) peak intensity ratio one can get is \( 60 \times 10^3/(60 \times 10^3+1024 \times 59 \times 990) \approx 10^{-3} \), and the largest is \( 60 \times 10^3/(60 \times 10^3+1024 \times 10) \approx 0.85 \). Substitute these ratios into Eq. (9), the range for \( w \) determined is from \( 0.39k_1 \to 2.63/k_1 \), where \( k_1 \) is determined by the diffraction geometry. For examples, if \( \theta = 70^\circ \), \( w \) ranges from 0.06 to 0.4 \( \mu \)m, and if \( \theta = 45^\circ \), \( w \) ranges from 0.03 to 0.2 \( \mu \)m. Clearly, as the incident angle becomes smaller and smaller, light scattering is suitable for measuring the smaller \( w \). But as \( \theta \) becomes larger, larger \( w \) can be determined.

### B. Surface power spectrum and scattering profile

Both the surface power spectrum and the scattering profile are functions in the reciprocal space. Both functions contain information on the surface morphology and are very closely related. Very often researchers treat the diffuse scattering profile as the surface power spectrum.\(^{25}\) This identity is true only under certain scattering conditions. In general, the scattering profile of a surface is not the surface power spectrum or a simple Fourier transform of the surface morphology. It is the Fourier transform of a more complicated function given by Eq. (3). Only when \( \Omega < 1 \), the diffuse profile is proportional to the surface power spectrum.\(^{26}\) For larger values of \( \Omega \), the scattering profiles are no longer proportional to the surface power spectrum.

### C. Non-Gaussian height distribution

Some rough surfaces may not have a Gaussian height distribution. If the height distribution deviates substantially from a Gaussian function, the integrand of Eq. (3) is no longer valid.\(^{27}\) An equivalent of Eq. (8) does not exist and Eq. (8) cannot be used to determine the value of \( \alpha \). However, for small \( \Omega \), the determination of \( w \) and \( \xi \) presented in Sec. IV C is still valid. The IFT method is still effective, but the integral kernel in Eq. (3) is changed due to a different height distribution.

However, it has been shown that as long as the diffraction condition \( \Omega \ll 1 \) is satisfied, all the roughness parameters can be estimated from a diffraction profile without specifying the particular surface height distribution.\(^{27}\) The interface width \( w \) can still be determined by the \( \delta \) peak intensity ratio, with \( \Gamma = 1 - k_1^2w^2 \). The diffuse profile is still proportional to the surface height power spectrum. Therefore, the lateral correlation length \( \xi \) is still inversely proportional to the FWHM of the diffuse profile, and the tail of the diffuse profile would give the roughness exponent \( \alpha \). In fact, as we have shown recently,\(^{23}\) at any diffraction condition (any \( \Omega \)), and for any surface height distribution, the tail of the diffuse profile would always give the roughness exponent \( \alpha \). The major difference is that for large \( \Omega \), one needs to go to the larger \( k_3 \) region in order to extract \( \alpha \).

### D. Other effects: Multiple scattering and shadowing

Multiple scattering was not a major factor in this experiment because the valleys on the surface were not too deep. For \( w \approx 100 \sim 500 \) nm and \( \xi \approx 5 \sim 10 \) \( \mu \)m, the slope of the pits \( \approx w/\xi \approx 0.01 \sim 0.06 \), which corresponds to \( \approx 0.57^\circ \sim 3.4^\circ \). These slopes are not deep enough to cause severe multiple scattering.

When the incident angle with respect to the surface normal becomes very large, the shadowing effect may play a role in the scattering intensity.\(^{28,29}\) Shadowing is the screening of parts of the rough surface by other parts, thus preventing a true profile of the light radiation from being measured. In the present experiment, shadowing might have had an effect on the determination of the roughness parameters for the large interface width case under a large incident angle \( \theta \). Here we can follow Wagner’s formula to estimate the shadowing effect.\(^{29}\) The apparent interface width \( w' \) can be written as

\[
w' = \frac{w^2}{1 + z_0^2/w^2},
\]

where \( z_0 \) is the solution of the following equation:

\[
\frac{z_0}{\sqrt{2}w} = \frac{B}{\sqrt{\pi}} e^{2z_0/2w^2}.
\]

Here

\[
B = \frac{e^{-\nu^2}}{4\sqrt{\pi} \nu} \text{erfc}(\nu),
\]

where \( \text{erfc}(x) \) is the complementary error function, and \( \nu = \xi/(2w tan \theta) \) for the Gaussian autocorrelation function (\( \alpha = 1.0 \)). In Fig. 17, we plot the ratio of \( w'^2/w^2 \) as a function of incident angle for different \( w/\xi \) values. For most of the samples (Nos. 1–6), because \( w/\xi \approx 0.01 \), there is almost
no shadowing effect in our measurement. For sample No. 10 ($w/\xi=0.06$), in the extreme case ($\theta=86^\circ$), the shadowing effect causes the apparent interface width $w'$ to be about 1% less than the real $w$.

VI. CONCLUDING REMARKS

In the present work we have explored the characteristics of the reciprocal space structure of rough Si (backside) surfaces using light scattering for a wide range of diffraction geometry. We found that measurements using in-plane scattering geometry are particularly convenient for the mapping of the reciprocal space characteristics. All relevant roughness parameters such as the interface width $w$, lateral correlation length $\xi$, and the roughness exponent $\alpha$ are quantitatively extracted from these characteristics. Limitations of the LS technique are discussed. We also compared the results of light scattering on the determination of roughness parameters such as AFM and SP.

The roughness parameters determined by the scattering technique are statistical averages of the large area covered by the laser beam, which is in the millimeter range. LS can be applied in a hostile environment such as during chemical vapor deposition or chemical etching of a surface. The temporal resolution is greatly improved using a detector array for collecting data. This scheme would enable researchers to use light scattering as a real-time, in situ monitoring tool for the study of the evolution of rough growth/etch fronts.

ACKNOWLEDGMENTS

The project was supported by NSF. The authors thank J. B. Wedding for reading the manuscript. Irene Wu, from Northwestern University, and Ueyn Block, from New Mexico Institute of Mining and Technology, were supported by NSF-REU 1996 and 1997 programs, respectively. C.-F. Cheng is on leave from the Physics Dept., Shandong Normal University, Jinan, China.

APPENDIX

The general two-dimensional scattering profile can be written as

$$S(k_x,k_y) = \int e^{-k^2_0 H(r)/2} e^{-i(k_x x + k_y y)} \, dr.$$  \hfill (A1)

The detail of our diffraction geometry is shown in Fig. 18. Note that in order to illustrate the scattering geometry more clearly, we rotate the original experimental setup presented in Fig. 1 by 180°. The incident plane is in the $yz$ plane, and OR is the reflection direction with a polar angle $\theta$ in the $yz$ plane. The detector plane consisting of a 1024 slit-diode array is perpendicular to the OR direction. OQ points to the center of one of the slit diodes. OQ has a polar angle $\theta'$ ($\neq \theta$). The angle $\gamma$ between OQ and OR is the in-plane scattering angle. The angle $\beta$ is the polar angle from OQ to OS due to the finite size of a slit detector. The angle $\Phi'$ is the azimuthal angle that the projection of OS in the $xy$ plane makes with respect to the $-y$ direction. Using the geometry in Fig. 18, one can write the incident wave vector $k_0 = (0, -k_0 \sin \theta, -k_0 \cos \theta)$, and the scattered wave vector along the OS direction $k_s = [k_0 \sin (\theta' + \beta) \sin \Phi', -k_0 \sin (\theta' + \beta) \cos \Phi', k_0 \cos (\theta' + \beta)]$. One can further determine that $\cos \theta' = \cos \theta \cos \gamma$ and $\tan \Phi' = \tan \gamma \sin \theta$. Therefore, the momentum transfers along both $k_x$ and $k_y$ directions are $k_x = k_0 \tan \gamma / \sqrt{\tan^2 \gamma + \sin^2 \theta}$, $k_y = -k_0 \sin (\theta \tan \gamma / \sqrt{\tan^2 \gamma + \sin^2 \theta})$ and $S(k_x,k_y) = \int k_x \, dk_x \int k_y \, dk_y$. The approximations can be made if both angles $\gamma$ and $\beta$ are small. As we discussed in Sec. III, because of the geometry of the slit detector, the actual intensity profile measured by the detector array is the integration $S(k_x,k_y)$ in both $k_x$ and $k_y$ direction over the range ($k_x - \Delta_x/2, k_x + \Delta_x/2$) and ($-\Delta_y/2, \Delta_y/2$):

$$S(k_x) = \int_{k_x - \Delta_x/2}^{k_x + \Delta_x/2} dk_x \int_{-\Delta_y/2}^{\Delta_y/2} S(k_x,k_y) \, dk_y.$$  \hfill (A2)

From Sec. III, we know that $\Delta_x \approx 1.4 \times 10^{-3}$ $\mu m^{-1}$, and $\Delta_y \approx 1.4 \times 10^{-1}$ $\mu m^{-1}$. Assuming $\Delta_y \gg$ FWHM of the diffuse profile, one can extend the integration over $k_y$ to infinity.
\begin{equation}
S_r(k_x) = \Delta x \int -\infty \rightarrow \infty dk_y \int e^{-k_y^2 H(i)} e^{-i(k_y x + k_y^2)} d\mathbf{r}
\end{equation}
(A3)

\begin{equation}
= \Delta x \int e^{-k_y^2 H(i)/2} e^{-ik_y x} dx,
\end{equation}
(A4)
i.e., the scattering profile becomes the one-dimensional Fourier transform of the function \(e^{-k_y^2 H(i)/2}\). Equation (A4) is the one-dimensional analogue of Eq. (A1). The relations shown in Eqs. (5), (7), (8), and (9) still hold for Eq. (A4). The shape of the diffuse profile obtained from Eq. (A4) differs from that obtained from Eq. (A1). For \(\Omega \ll 1\), the diffuse profile from Eq. (A4) is also proportional to the power spectrum of the surface. But in this case it is the one-dimensional power spectrum which is proportional to \(k_x^{-1-2\alpha}\) instead of \(k_{\parallel}^{-2-2\alpha}\) [the large \(k_{\parallel}\) region from Eq. (A1)].

26. Theoretically, the diffuse profile is proportional to the surface power spectrum when \(\Omega \ll 1\). But experimentally, we found that the condition \(\Omega \ll 1\) can be relaxed to \(\Omega < 1\), for example, please see Fig. 16 in the paper. Also see Ref. 20.