

Comparison of surface roughness of polished silicon wafers measured by light scattering topography, soft-x-ray scattering, and atomic-force microscopy

C. Teichert, J. F. MacKay, D. E. Savage, and M. G. Lagally
University of Wisconsin-Madison, Madison, Wisconsin 53706

M. Brohl and P. Wagner
Wacker-Chemitronic GmbH, D-84489 Burghausen, Germany

(Received 17 October 1994; accepted for publication 10 February 1995)

The surface roughness of silicon wafers after different stages of chemomechanical polishing was investigated by light scattering topography, soft-x-ray scattering, and atomic-force microscopy. Quantitative values of the rms roughness, the lateral correlation length, and the roughness exponent are extracted. The results suggest deviations from the "ideal" polishing process at large length scales. © 1995 American Institute of Physics.

Chemomechanically polished silicon wafers are used in a wide range of technologies. In many of these applications, the desired properties are influenced by the Si surface morphology. Improvement of polishing methods to increase Si wafer surface quality has been a continuing goal of silicon manufacturers. Among the various methods used for the characterization of microroughness, light scattering topography (LST)¹ is commonly used in industry because it enables a very rapid inspection of the entire wafer. Because LST integrates diffusely scattered light over an instrumentally restricted solid angle, the output of the method, the "haze value," is related only to the rms roughness (in the given bandpass) and no information on lateral scale of the roughness is obtained.² It is known that the lateral frequency of the roughness influences device qualities.³ Lateral scale lengths of roughness can be obtained either from real-space scanned-probe measurements, such as atomic force microscopy (AFM), or from x-ray scattering (XRS) measurements. AFM is direct, is readily interpreted, and has a wide band pass. The disadvantage is that it is very slow. Diffuse x-ray scattering extends the lower bandpass limit of AFM and covers the bandpass of LST. It is rapid, giving an average over a large sample area if desired, but is indirect, requiring modeling for interpretation. In both AFM and XRS, analysis of the height-height correlation and the height-difference functions of surface morphologies^{2,4} provides the lateral correlation length of the roughness as well as the roughness exponent h , which tells how "jagged" a surface with a given rms roughness is.^{5,6}

In this letter we report the determination of values for the rms roughness σ , the lateral correlation length ξ , and the roughness exponent h of Si(001) wafers at different stages of polishing using both soft-x-ray diffuse scattering (SXRS) and AFM. We compare these results with the haze values obtained by LST and discuss them in terms of the band pass of each method. The combination of SXRS, AFM, and LST measurements allows a quantitative overall characterization of the mesoscopic wafer roughness. We investigated samples from three different Si wafers. Wafer A underwent the entire multistep chemomechanical polishing procedure. The final step has a reduced chemical (etching) power in order to

smooth the wafer surface rather than to remove material. Wafers B and C did not experience the final polishing step. They represent the range of variation in roughness that is observed by LST before the final polishing step. Immediately after polishing, all wafers were inspected by LST using a CENSOR ANS 100.¹ This instrument integrates the scattered intensity in a volume range bounded by two cones (5° and 25° with respect to the normal). This detection range together with the wavelengths used ($\lambda=488$ and 514 nm) results in a band pass of $1-6 \mu\text{m}$. Soft-x-ray-scattering measurements were carried out at the multilayer-mirror monochromator⁷ beam line at the Synchrotron Radiation Center, University of Wisconsin-Madison. The monochromator delivers high flux over an energy range of 250 to 2500 eV. The advantage of this source is the very high intensity compared to conventional sources and the higher surface sensitivity afforded by the low energy of the x rays. We used rocking curves to probe the intensity distribution parallel to the sample surface.⁸ For all measurements, a scattering angle of $2\theta=3.4^\circ$ and a wavelength of $\lambda=9.3 \text{ \AA}$ ($E=1330$ eV) were chosen. The instrumental broadening at the experimental conditions⁷ corresponds to an upper limit of the band pass of about $10 \mu\text{m}$, about an order of magnitude better than conventional laboratory x-ray diffractometers. The lower limit of the bandpass, about 1 nm, is determined by the wavelength used. The AFM measurements had a bandpass of 5 nm to $10 \mu\text{m}$, with the upper limit determined by the scan range that was chosen.

The haze values D , i.e., the ratio of the intensity scattered into the detection cone to the incident intensity, obtained by LST for the three samples cover a range of 4 orders of magnitude (see Table I). If it is assumed, as usual, that the height deviations from the mean surface level are Gaussian distributed, the value of D is related to the rms roughness σ by⁹

$$D = R \{ 1 - \exp[-(\sigma 4 \pi / \lambda)^2] \}, \quad (1)$$

where R denotes the total reflectance, i.e., the ratio of the intensity reflected into the entire half space to the incident

TABLE I. Roughness values obtained on three Si(001) wafers by LST, SXRS, and AFM. The haze values are given in ppm of the incident intensity. The bandpass of each technique is also listed.

Sample	LST		Soft x ray			AFM		
	Haze (ppm)	σ (nm)	σ (nm)	ξ (μm)	h	σ (nm)	ξ (μm)	h
A	0.046 ± 0.001	0.013 ± 0.0001	0.078 ± 0.005	1.40 ± 0.1	0.7 ± 0.15	0.115 ± 0.015	1.5 ± 0.3	0.7 ± 0.1
B	12.0 ± 0.6	0.22 ± 0.005	0.47 ± 0.02	1.10 ± 0.2	0.5 ± 0.1	0.46 ± 0.05	1.3 ± 0.3	0.5 ± 0.1
C	83.0 ± 1.0	0.57 ± 0.004	1.25 ± 0.1	1.15 ± 0.15	0.8 ± 0.1	1.10 ± 0.12	1.1 ± 0.3	0.7 ± 0.1
Bandpass	1–6 μm		1 nm–10 μm			5 nm–10 μm		

intensity ($R \approx 0.4$ for Si at $\lambda = 500$ nm and normal incidence). Equation (1) gives a ranking of the rms roughness $\sigma_A < \sigma_B < \sigma_C$.

The measured soft-x-ray rocking curves are shown in Fig. 1. All polishing procedures yield a two-component profile: an instrument limited peak in the specular direction and a broad diffuse component. The existence of such a two-component profile implies a roughness with a long-wavelength cutoff.⁶ For such roughness, the height–height correlation function $C(x)$

$$C(x) = \langle [z(x_0 + x) - \langle z \rangle][z(x_0) - \langle z \rangle] \rangle, \quad (2)$$

where $z(x_0)$ is the height at a point x_0 , and x is the lateral separation between two surfaces points, can be written in terms of σ , ξ , and h as

$$C(x) = \sigma^2 \exp[-(|x|/\xi)^{2h}]. \quad (3)$$

This correlation function treats the surface as self-affine fractal^{4,6} on a short length scale and as smooth on a long length scale. In this approach, the amount of the diffuse component is determined by the rms roughness; the width of this component and its detailed shape give the lateral correlation length and the roughness parameter, respectively. A larger value of the roughness parameter h means more gradual height changes. From the levels of the diffuse components it can easily be seen that wafer A is the smoothest and wafer C is the roughest, in agreement with the LST result. The calculated profiles, also shown in Fig. 1, were obtained using the procedure of Sinha *et al.*⁶ The sharp break between the

specular peak and the diffuse intensity (due to the high resolution of the instrument) allows the determination of unique σ , ξ , h triplets (see Table I). All the samples have approximately the same lateral correlation length, in the one-micrometer range. The roughness exponent ranges from 0.5 to 0.8.

Figure 2 shows as an example an AFM image of wafer C measured over an area of $10 \mu\text{m} \times 10 \mu\text{m}$. All wafers exhibit morphological features with lateral dimensions of about 3–4 μm but of decreasing height for the smoother samples. The σ values listed in Table I are averages of five $10 \mu\text{m} \times 10 \mu\text{m}$ scans at different points on each sample. They are in quantitative agreement with the results of the soft-x-ray analysis. The lateral correlation length of roughness can be obtained from the AFM images by determining the dependence of the height–height correlation $C(x)$ and the height difference function $H(x)$ on the lateral dimension of the analyzed area. To compare with the soft-x-ray analysis (which is one dimensional because of the intensity integration in one direction), we calculated $C(x)$ and $H(x)$ as a function of x and averaged all the line scans. Using Eq. (3), the lateral correlation length ξ is given by the value of x at which the function decays to $1/e$. The resulting values (see Table I) are in agreement with the soft x-ray data. The size of the visible structures mentioned above confirms this length scale.

The roughness exponent h can be determined by evaluating the height difference function. For a self-affine surface,

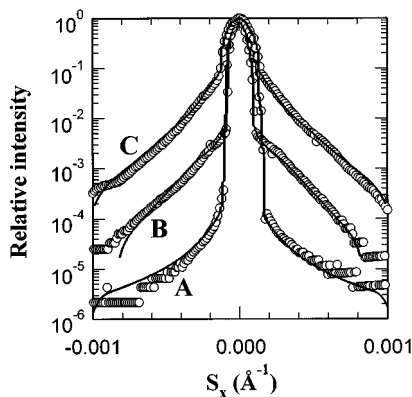


FIG. 1. Soft-x-ray rocking curves (○) along with the calculated intensity distributions (—) for wafers A, B, and C. The fits are made to match the ratio of intensity of the specular to diffuse component. A higher diffuse intensity corresponds to greater rms roughness. The roughness parameters obtained from the fitting are listed in Table I.

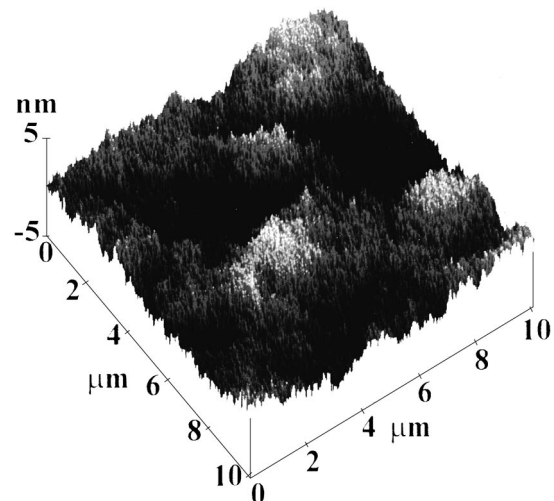


FIG. 2. Typical AFM image, here for wafer C.

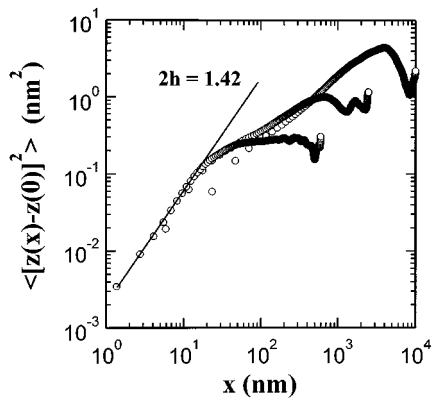


FIG. 3. Log–log plot of the height difference function, $H(x)$, as a function of the lateral range x obtained from AFM images of wafer C. To increase the data density at small distances, curves for $600\text{ nm} \times 600\text{ nm}$ and $2.5\text{ }\mu\text{m} \times 2.5\text{ }\mu\text{m}$ scans are included. From the initial slope in the short-range scale a roughness exponent of $h = 0.71 \pm 0.1$ is determined. Because of finite window-size effects and remaining image curvature, values of $H(x)$ are reliable only up to about $\frac{1}{3}$ of the scan size.

$$H(x) = \langle [z(x + x_0) - z(x_0)]^2 \rangle \quad (4)$$

has at small distances ($x \ll \xi$) an asymptotic power-law form⁴

$$H(x) \sim x^{2h} \quad (x \ll \xi). \quad (5)$$

The initial slope of the log–log plot of $H(x)$ vs x gives h , as shown in Fig. 3 for sample C.

Table I shows that all three measurements yield a ranking of the rms roughness $\sigma_A < \sigma_B < \sigma_C$. The nearly perfect agreement between SXRS and AFM measurements is understandable because both methods look at the surface with almost the same bandpass. The lower values of rms roughness estimated from the LST measurements are mainly due¹⁰ to the lower upper limit of the bandpass of LST, which thus misses the long-range roughness observed with x-ray scattering and AFM. For wafer A, for which the deviation is strongest, the rms values obtained by soft-x-ray scattering and AFM may be somewhat too high because of the effect of noise.¹⁰ The lateral correlation length, about $1\text{ }\mu\text{m}$ for all samples, does not change after the final polishing step, although this step decreases the rms roughness considerably. This value of ξ is much greater than that obtained by others² on silicon wafers after the removal of the oxide. The roughness exponent values indicated that the height changes are gradual.

The conclusions to be drawn from Fig. 3 are quite intriguing. Thin films grown by physical or chemical vapor deposition typically have roughness that can be described by a correlation function that is self-affine fractal at small length scales but has a long-wavelength cutoff, with magnitude σ^2 , because of the loss of correlation for long length scales $x \gg \xi$.⁴ In addition to smoothing, the aim of polishing is to prevent the enhancement of special structure size. Ideal polishing should produce self-affine fractal roughness on all length scales, i.e., a continuation of the straight line in Fig. 3 to all wavelengths. We see neither a constant, as in grown films, nor a continuation of the initial slope, but a behavior between the two, suggesting that the surface continues to

have correlation to all wavelengths but with weaker correlation than a self-affine fractal surface. There are thus deviations from the “ideal” polishing process.

The samples used here were chosen because of the range of LST haze values they displayed, in order to obtain benchmarks to interpret the haze values in terms of surface roughness, especially for larger correlation lengths. LST orders the samples correctly, but the surface roughness is generally underestimated, mainly because longer-wavelength roughness is excluded or under-represented. Because these samples were specifically chosen for this study, they may not be representative of “typical” quality of commercially available wafers. However, for a wafer “out of the box” from a different supplier we obtained a similar lateral correlation length in the one micrometer range.¹¹

In conclusion, we have monitored the surface roughness evolution during different stages of chemomechanical polishing of silicon wafers. Quantitative data for the rms roughness, the lateral correlation length of roughness, and the roughness exponent were obtained by soft x-ray scattering and AFM. LST haze values, however, generally underestimate the rms roughness. The final polishing procedure applied here results in very smooth surfaces with a rms roughness $\leq 1\text{ }\text{\AA}$ (on a length scale of $10\text{ }\mu\text{m}$). The lateral correlation length of the roughness, introduced by an early stage of polishing, apparently is not influenced by the final polishing step. The persistence of correlation, but below that of a perfectly self-affine surface, suggests deviations from an ideal polishing process. The excellent agreement of the data demonstrates the advantage of combining soft-x-ray scattering and AFM for monitoring the surface roughness evolution; the former produces excellent statistics while the latter produces a local, real-space view that also permits visualization of possible polishing artifacts. This combination of methods allows examination as well of the influence of substrate surface finish on growth-front roughness of thin films or interfacial roughness of multilayer films deposited onto such substrates.^{7,8}

This research was supported by NSF, MRG Grant No. DMR-9121074, and Grant No. DMR92-12658. C.T. acknowledges support of the German Academic Exchange Service. We thank K.-C. Liu and J. Liu for technical assistance.

¹T. Abe, E. F. Stegmeier, W. Hagleitner, and A. J. Pidduck, *Jpn. J. Appl. Phys.* **31**, 721 (1992).

²T. Yoshinobu, A. Iwamoto, and H. Iwasaki, *Jpn. J. Appl. Phys.* **33**, 383 (1994).

³T. Abe and Y. Kato, *J. Appl. Phys.* **32**, 1879 (1993).

⁴H.-N. Yang, G.-C. Wang, and T.-M. Lu, *Diffraction from Rough Surfaces and Dynamic Growth Fronts* (World Scientific, Singapore, 1993), p. 64.

⁵B. B. Mandelbrodt, *The Fractal Geometry of Nature* (Freeman, New York, 1982).

⁶S. K. Sinha, E. B. Sirota, S. Garoff, and H. B. Stanley, *Phys. Rev. B* **38**, 2297 (1988).

⁷D. E. Savage, Y.-H. Phang, J. J. Rownd, J. F. McKay, and M. G. Lagally, *J. Appl. Phys.* **74**, 6158 (1993).

⁸D. E. Savage, J. Kleiner, N. Schimke, Y.-H. Phang, T. Jankowski, J. Jacobs, R. Kariotis, and M. G. Lagally, *J. Appl. Phys.* **69**, 1411 (1991).

⁹H. E. Bennet and J. O. Porteus, *J. Opt. Soc. Am.* **51**, 123 (1961).

¹⁰C. Teichert, J. F. MacKay, D. E. Savage, and M. G. Lagally (unpublished).