High-speed, non-interferometric nanotopographic characterization of Si wafer surfaces

T.D. Raymond*, D.R. Neal, and D.M. Topa
WaveFront Sciences, Inc.
and
Tony L. Schmitz
Department of Mechanical and Aerospace Engineering
University of Florida

ABSTRACT

We present a high-speed silicon wafer metrology tool capable of resolving surface features in the nanometer height range. This tool uses a high performance Shack-Hartman sensor to analyze the wavefront of a beam of light reflected from a silicon wafer surface. By translating the wafer to analyze small portions of the wafer in each camera frame and then continuously piecing the frames together, we can retain sub-millimeter spatial resolution while rapidly analyzing large apertures. This tool is particularly effective for resolving features near the wafer edge. We will describe the measures required to obtain this level of resolution. We also compare data taken with this device to that taken with the National Institute of Standards and Technology X-ray optics Calibration Interferometer (XCALIBIR). Finally, we show measurements of a variety of typical and atypical 200 mm diameter wafer samples.

Keywords: Shack-Hartmann, nanotopography, silicon wafer, interferometer

1. INTRODUCTION

The semiconductor industry has achieved tremendous improvements in integrated circuit (IC) design over the last several decades, driven primarily by the reduction of the critical dimensions of the minimum circuit feature sizes. Most modern IC foundries today are producing IC's with critical dimensions of 130 nm utilizing sub wavelength photolithography with 193 nm lasers. The industry is striving to achieve IC's with critical dimensions of 90 nm by 2004 and 65 nm by 2007 using either 157 nm lasers or 13 nm light from extreme ultraviolet light from laser induced plasmas. Photolithographic optical systems to achieve the necessary critical dimensions have an optical depth of focus approximately equal to the critical dimension. Naturally, the silicon wafers upon which these circuits are cast must be sufficiently flat over the exposure area to prevent defocusing. Economic considerations dictate that the thickness to diameter ratio of silicon wafers be very small, of order 300, hence the stiffness of the wafer itself is insufficient to maintain this surface figure even over the few millimeter exposure sites for today's photolithography systems. To achieve the site flatness criteria, the wafer must be held rigidly against a wafer chuck during exposure.

Furthermore, as no modern IC's are built on a single layer, the ability to accurately control the thickness of the oxide layers used to insulate the circuit layers is also increasingly important. The dielectric constant and thickness of these layers affect the device speed. Variations in the oxide thickness can lead to reduced device performance or device failures due to dielectric breakdown and incomplete circuit connections. The chemical-mechanical planarization polishing process used to control oxide thickness is sensitive to variations in wafer height. Xu[1] has shown that oxide thickness variations are correlated to the so-called nanotopography of the native silicon wafer surface. Features on the

* TDRaymond@WaveFrontSciences.com, phone 1 505 275 4747, fax 1 505 275 4749; http://www wavefrontsciences.com.

 © 2002 SPIE · 0277-786X/2002/$15.00

wafer surface composed of spatial wavelengths between 0.2 and 20 mm comprise the nanotopography of the surface. This definition excludes the micro roughness of the surface but includes polishing artifacts, surface artifacts due to dopant variations in the crystal, and pits and bumps induced by the mounting processes used to polish the wafer. These features typically have amplitudes on the order of a few nanometers to hundreds of nanometers.

Metrology tools able to measure the flatness and nanotopography of the wafer surface to qualify wafers for IC production are critical to the industry. These tools must be capable of measuring with much greater repeatability and accuracy than the actual industry tolerances to reduce the probability of making errors in the sorting process. They must be capable of nondestructive measurement, without adding micro particles or metal contamination. These tools must also operate at high speed to keep up with the wafer production process.

In this paper we describe the Columbus nanotopography tool, a metrology device that measures the nanotopography of silicon wafer surfaces, and its performance. This tool uses a Shack-Hartmann sensor to optically probe the wafer surface and nondestructively measure nanotopography and site flatness. This method is not based on interferometry, but achieves comparable measurement sensitivity. We compare results of measurements on the Columbus tool to measurements performed on the NIST XCALIBIR system.

2. DESCRIPTION OF THE METHOD AND INSTRUMENT

2.1. Columbus instrument overview

The major components comprising the Columbus system are illustrated in Figure 1. A fiber coupled 635 nm laser diode light source is collimated and directed onto the wafer surface. The wafer is mounted on a wafer chuck attached to a two-axis translation stage capable of positioning the wafer so that the entire surface can be analyzed one patch at a time. The phase of the reflected light is imprinted with the height variations on the wafer surface. The reflected light is imaged, using a simple 1:1 relay telescope, onto a Shack-Hartmann[2] wavefront sensor for analysis of the phase, and thus the wafer surface.

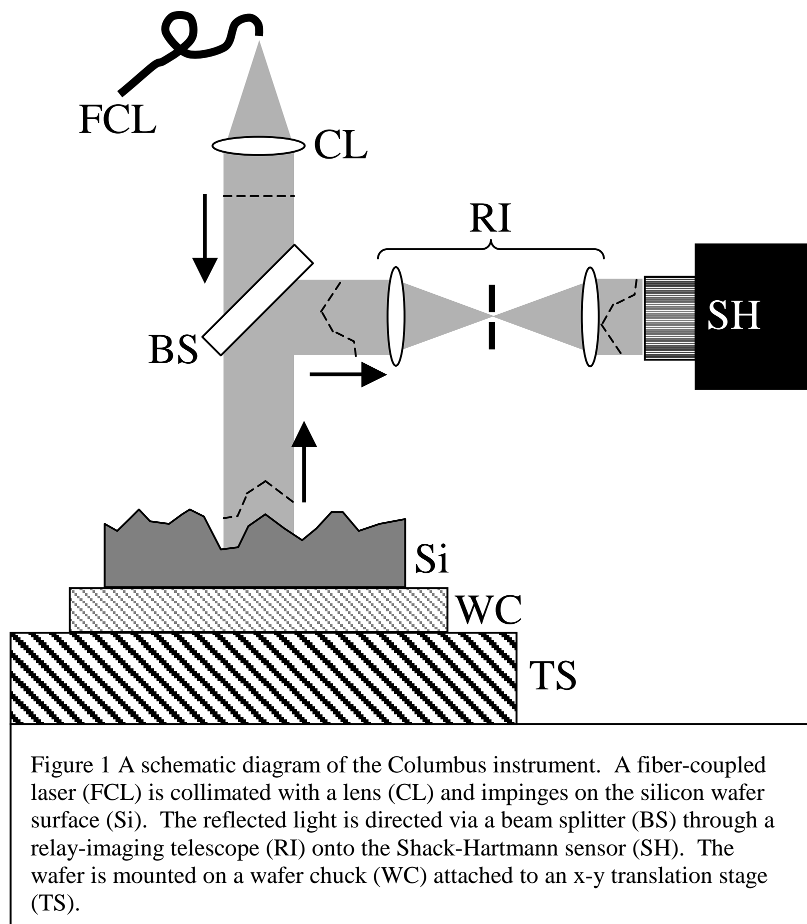

Figure 1 A schematic diagram of the Columbus instrument. A fiber-coupled laser (FCL) is collimated with a lens (CL) and impinges on the silicon wafer surface (Si). The reflected light is directed via a beam splitter (BS) through a relay-imaging telescope (RI) onto the Shack-Hartmann sensor (SH). The wafer is mounted on a wafer chuck (WC) attached to an x-y translation stage (TS).

The Shack-Hartmann sensor is the heart of the instrument. It is shown schematically in Figure 2. The basic elements are a micro lens array, a mounting housing, and a mega-pixel digital charged coupled device (CCD) camera. The micro lens array is produced on a fused silica substrate using gray scale photolithography and plasma etching[3]. Each lenslet in the array has a 0.28 mm square aperture and a focal length of approximately 28 mm. The square array is bonded to a metal fixture that constrains its position relative to the active area of the digital camera. Light incident on each lenslet of the array produces a focal spot

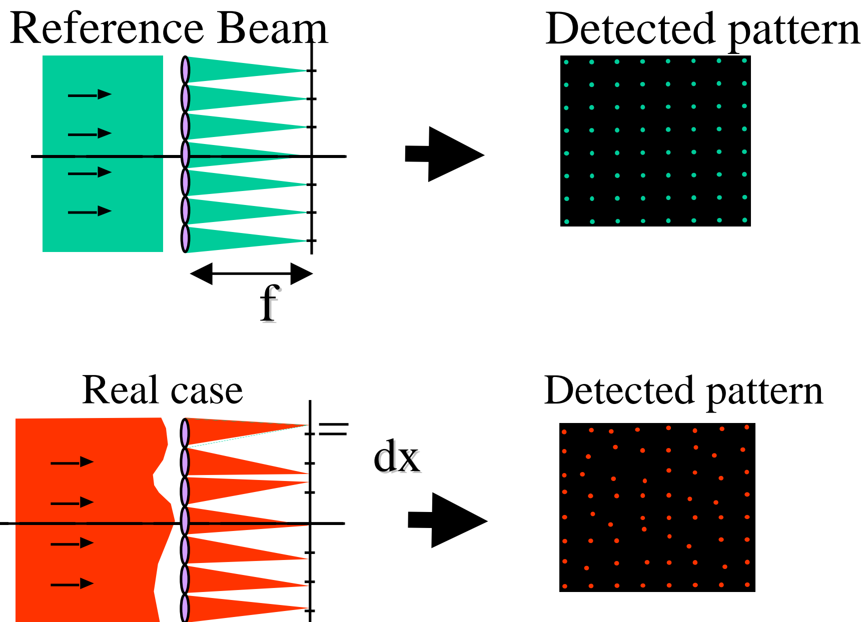

Figure 2 The upper diagram shows the incident collimated light impinging on a lenslet array that creates a set of approximately evenly spaced spots. The lower diagram shows the case for a real wavefront with aberrations.

on the CCD, where the local tilt of the phase front determines the position of the focal spot. By accurately determining the focal spot positions in both directions, the wavefront *gradient* may be deduced. The light incident on the CCD is digitized to 12 bits, pixel-by-pixel and read into a personal computer where advanced centroid or convolution techniques are used to determine the focal spot positions.

During sensor calibration, collimated light is reflected from a 203 mm diameter optical flat (λ/20) and directed through the optical system to the sensor; a reference file containing the focal spot positions is saved for later use. During normal operation, the light reflected from the wafer is also directed through the optical system to the sensor and the deviation of the spot locations in each direction is used to extract the surface gradient of the wafer. Note that this procedure is critical to successful operation; acquiring the reference file through the actual optical system automatically corrects for imperfections in the optical system.

2.2. Measuring large apertures

The data acquisition method described above permits the measurement of the wafer surface gradient over only a 14 mm square region of the reflective surface. To map a large aperture surface, the translation stage is used to position the object (wafer) to successively different locations. The laser is pulsed at each location and the surface gradient is recorded for the corresponding 14 mm square frame. Unfortunately, the average tilt of each frame is affected by the imperfections in the stage bearings. To overcome this limitation, the stage is programmed to permit the measured regions to overlap by a few millimeters as shown in Figure 3. The slope measurements in the overlap regions are used in a linear least squares fitting routine to determine the relative tilts between frames of data. These relative tilts are then used to correct the data and to construct a stitched gradient map for the wafer. Figure 4 shows the uncorrected slope map (a) and the corrected slope map (b) for a sample single side polished wafer. The mechanical stage can introduce more than 100 μrad of tilt that is easily corrected by the fitting routine.

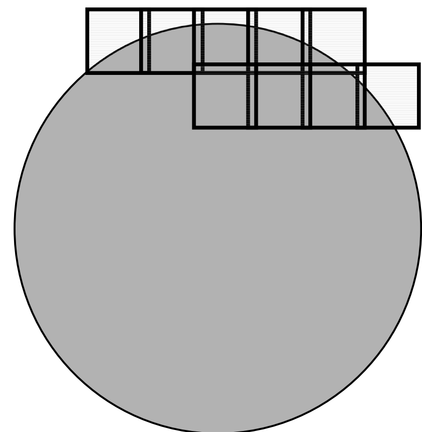

Figure 3 The scan pattern incorporates overlap to stitch the frames.

Once the map of the surface gradient is obtained, the surface itself is reconstructed using a wavefront reconstructor algorithm such as that described by Southwell[4]. For these data we have used a successive (or

simultaneous) over relaxation method to numerically solve for the reconstructed surfaces. The reconstructed surface for the above wafer is shown in Figure 5(a). In this case the wafer was simply placed on the wafer chuck and constrained only from transverse motion. Because the wafers are so thin compared to their diameters, they readily distort under their own weight and internal stresses. Note that the peak to valley of this wafer is approximately 3.9 μm but it is not uncommon for acceptable wafers to have a peak to valley exceeding 30 μm.

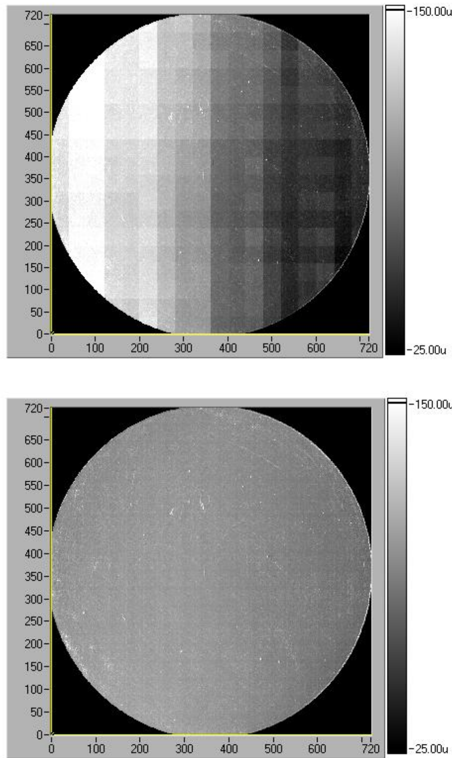

Figure 4 The x slope map for a 203. mm optical flat. a) Shows the raw x slope map before stitching, b) shows the x slope map after stitching. The gray scale varies from – 25 μrad to + 150 μrad.

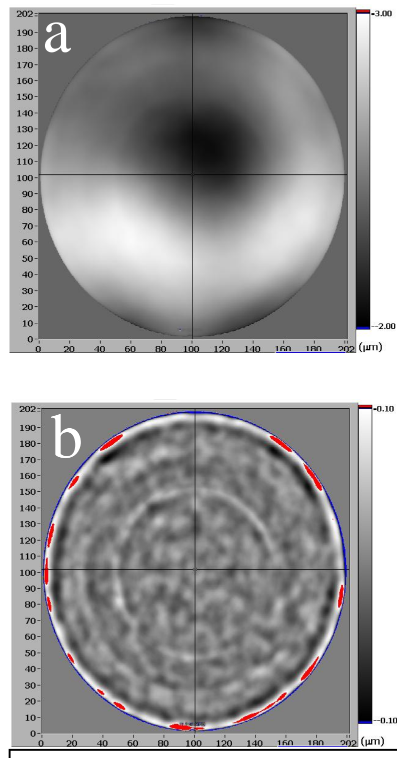

Figure 5 (a) The reconstructed wafer surface shows a peak-to-valley height variation of about 5 microns. (b) Shows the high pass filtered data from the same wafer.

While these low spatial frequency features are of no interest to the IC manufacturers, their large amplitude masks smaller amplitude features of higher spatial frequency, such as the wafer nanotopography. To reveal the nanotopographic features of the surface, the low spatial frequency content of the reconstructed surface must be suppressed. This can be accomplished using a high frequency spatial filter. In this case, we have applied a convolution filter in the form of a double Gaussian high pass filter with 50% cutoff point at spatial frequency 0.05 waves/mm; this filter effectively suppresses features with transverse dimensions greater than about 20 mm. Figure 5(b) shows the high pass filtered data . It is data such as these that are analyzed to determine a wafer's suitability for use in IC manufacture.

3. COMPARISON TO INTERFEROMETRIC METHOD

This measurement technique has several advantages over conventional phase shifting interferometry. Unlike an interferometric measurement, this method is insensitive to piston and tilt variations due to platform motion or air turbulence. The piston insensitivity is achieved because the Shack-Hartmann sensor only measures the gradient of the wavefront (or optical surface), and the tilt insensitivity is due to the least squares fitting routine for stitching the data.

This combination of attributes eliminates the primary error sources found in interferometry, and makes this technique uniquely suited to high accuracy measurements without the need to vibration isolate the device or to take extraordinary precautions to minimize temperature gradients in the air or structure. This method is also faster than conventional phase shift interferometry (of equal aperture), because the instrument does not have to settle after each stage motion before the data is acquired. This instrument rapidly acquires the data by continuously scanning the wafer and pulsing the laser to "freeze" the stage motion; an entire 200 mm diameter wafer can be scanned in approximately 36 seconds.

Figure 6 The NIST XCALIBIR interferometer is a state-of-the-art phase shifting interferometer capable of sub nanometer resolution.

To compare the performance of the Columbus instrument to conventional interferometry, we measured a wafer using the X-ray optic CALibration InterferometeR (XCALIBIR) [5], a state-of-the-art instrument developed at NIST. This instrument, designed to measure large optics for extreme ultraviolet photolithography systems with sub nanometer resolution, is shown in Figure 6. This system is a 300 mm diameter phase measuring interferometer, operated at 633 nm, that can be configured in a variety of ways. It is mounted on a massive granite table floating on air cushion legs, within a temperature-controlled room; all sources of heat are carefully managed to reduce temperature gradients in the air to around 0.05C within the optical path. The optical axis is horizontal, which requires that the wafer to be mounted on edge. Once the wafer is mounted, the room temperature stabilizes and vibrations damp out, the data for the entire wafer is acquired within a second.

Since the mounting methods for the wafer were different, the low frequency components of the surface maps from Columbus and XCALIBIR cannot be directly compared; however, the high-frequency filtered data can be compared. Figure 7 shows that these data are strongly correlated. A quantitative comparison of the height profiles, shown in Figure 8, illustrates the close agreement between these systems. The XCALIBIR data does not agree with the Columbus data near the left end of the profile because the phase unwrapping method used to determine the XCALIBIR height map

failed in this region; such limitations are expected where the surface height changes rapidly. The overall and quantitative agreement between the data obtained with these two instruments shows that the Columbus method of data acquisition is comparable to that obtained using interferometry.

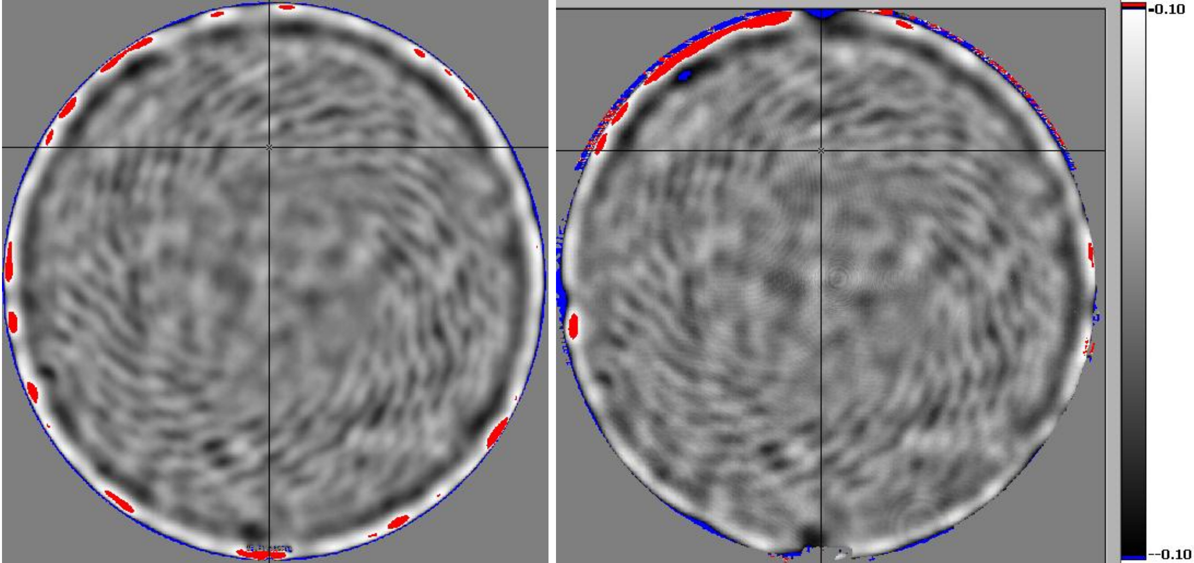

Figure 7 Map a) is the nanotopography of a wafer as measured with Columbus. The map b) is the same wafer measured on XCALIBIR. Both data sets were identically filtered to reveal the nanotopography of the wafer. The gray scale range is ± 0.1 µm.

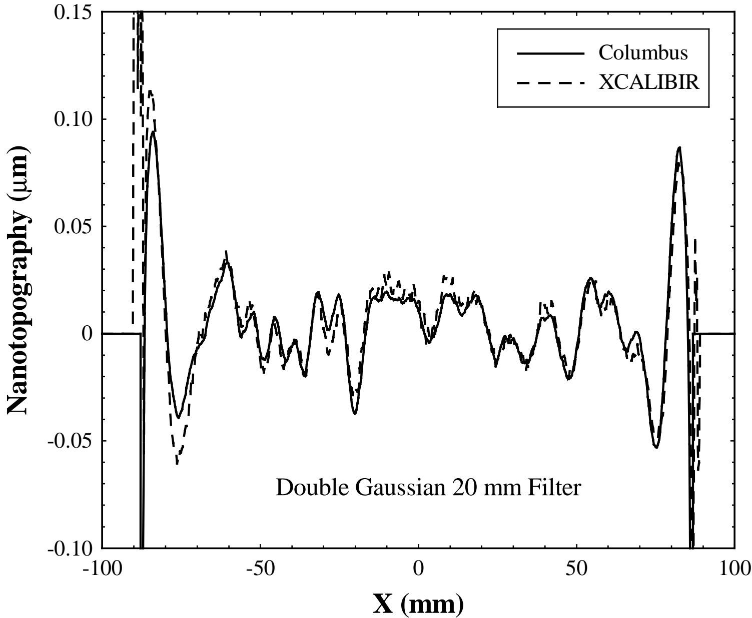

Figure 8 A comparison of the data acquired with Columbus and XCALIBIR. The data is the horizontal profile along the horizontal black line in Figure7.

4. PERFORMANCE AND RESULTS

This section illustrates typical and atypical silicon wafer nanotopography maps highlighting the capabilities of the Columbus tool. The wafers presented here were produced using a variety of methods and exhibit characteristics representative of those manufacturing steps.

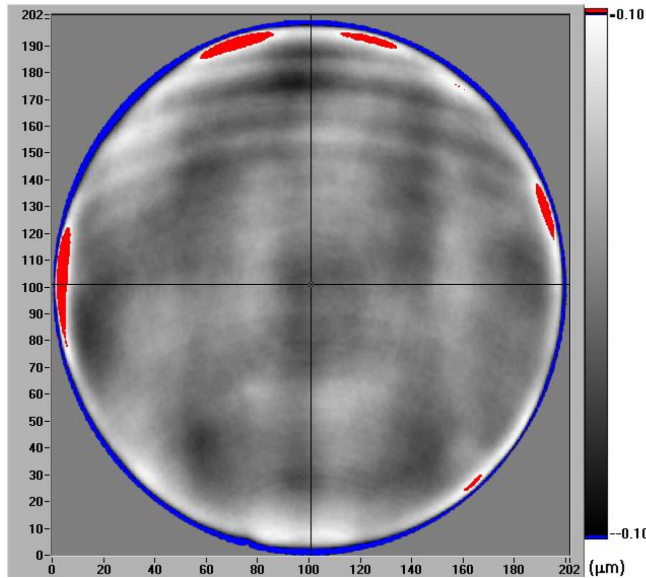

Figure 9 This nanotopography map shows evidence of saw marks.

Figure 9 shows a single side polished (SSP) wafer that illustrates an artifact of the manufacturing process. After the large single crystal boules of silicon are grown, thin wafer blanks are sawed from the boule using a wire saw. As the saw cuts through the boule, it may vibrate. The striking marks in this nanotopography map are the result of the cutting operation and have amplitudes of approximately 20 nm.

Figure 10 shows a typical single side polished (SSP) 200 mm diameter wafer. Wafers such as this are produced by clamping the wafer back to a polishing mandrel that rotates and translates over a polishing pad. The clamping mechanisms vary from manufacturer to manufacturer, but vacuum or wax are typically used. This wafer was mounted using the wax technique. The mottled structure shown in this map is typical of wafers produced in this manner. In contrast, the wafer shown in Figure 5(b) was produced using the vacuum mounting technique.

The wafer shown in Figure 11 is a double side polished (DSP) wafer. It is produced by a center-less polishing method in which the wafer is clamped between two rotating and translating polishing pads and allowed to rotate freely. Both sides of the wafer are polished to produce a high quality wafer with very smooth surface. This wafer shows faint lines near the center of the wafer that are the result of a combination of crystal growing artifact and polishing interaction. Temperature and other process variations can lead to concentration gradients in the boron dopant, used to vary the base electrical resistivity of the wafer, usually incorporated into these crystals. The polishing process reveals these concentration variations by creating nm level amplitude variations. This wafer also has an epitaxial (EPI) layer of silicon that was grown after the polishing. The dot like features at 10 o'clock, 2 o'clock and 6 o'clock are due to the mount used to hold the wafer in the reactor.

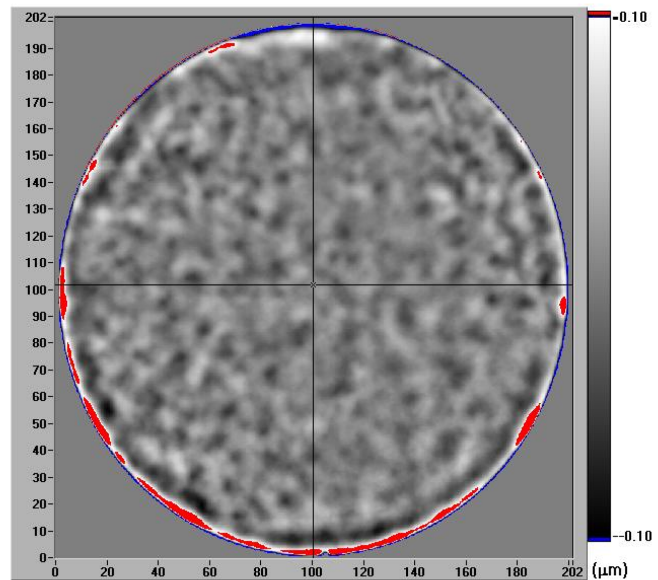

Figure 10 This SSP wafer was produced using wax to adhere it to the polishing head.

Finally, Figure 12 shows a map of an atypical wafer produced by a grinding process. In an attempt to reduce the cost of manufacturing wafers, some producers have experimented with a fast grinding process to reduce the time to produce the wafer. This map is a wafer in which the grind marks have not been completely removed by the final polish. The

grooves are about 10 nm in amplitude and have a period of about 2 mm. The data shown in this plot were generated with a special high pass reconstructor designed to suppress all spatial wavelengths longer than about 4 mm.

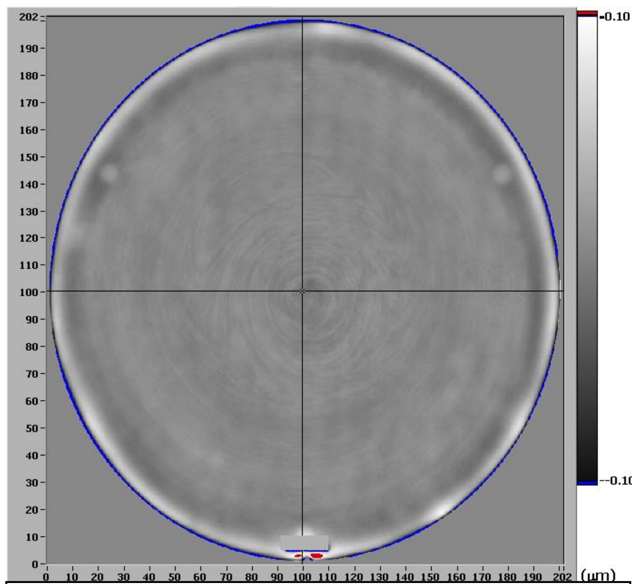

Figure 11DSP wafer showing dopant striation marks, and EPI growth artifact.

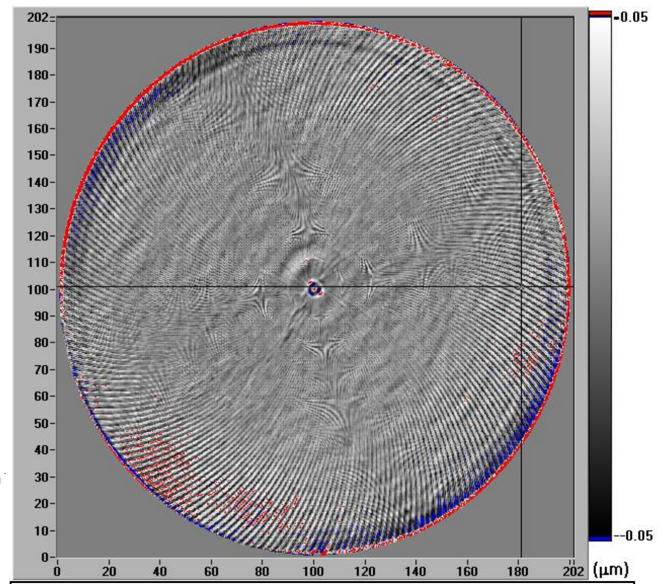

Figure 12 SSP wafer with incompletely polished grind marks.

5.CONCLUSIONS

In conclusion, we have described a new optical metrology tool capable of high-speed measurement of the quality of large aperture optical surfaces. A Shack-Hartmann sensor forms the heart of the instrument and a least squares fit method permits the assembly of sub aperture measurements to permit measurement of large optical surfaces (up to 300 mm). In contrast to interferometric techniques, this method virtually eliminates vibration as a noise source because the raw data measured by the sensor is the surface *gradient*. The instrument has been applied to the measurement of nanotopographic features on silicon wafers for the semiconductor industry. We have demonstrated that this tool is capable of measuring wafers with similar resolution and quality as the state-of-the-art NIST XCALIBIR interferometer. We have also shown a variety of wafers with different topography and have discussed how the processes used to manufacture the wafers have influenced their topography.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Ulf Greismann of NIST who operated the XCALIBIR instrument, and the important technical contributions of Dan Hamrick, Wei Xiong, Tom Rogers and others at WaveFront Sciences as well as Chris Burak and John Adair. Finally, the authors would like to thank the various semiconductor firms with whom we have had numerous stimulating conversations and who may have also contributed wafers used in these studies.

REFERENCES

1. C. Shan Xu, E. Zhao, R. Jairath, and W. Krusell, Electrochemical and Solid-State Lasers, **1** (V4), pg 181 (1998); H. Jeon, T. Katoh, W. Lee, U. Paik, J. Park, H. Suga, Japanese J of Applied Physics, Part 2 Letters, **41**, pg. L443, (2002).

2. R.B. Shack and B.C. Platt, "Production and use of a lenticular Hartmann screen," JOSA **61**, pg. 656 (1971).
3. P. Puloski, Dan Neal, Jim Roller, and Keith Stantz, "Measurement of aberrations in micro lenses using a Shack-Hartmann wavefront sensor," this proceedings (talk 4867-7).
4. W.H. Southwell, "Wave-front estimation from wave-front slope measurements," JOSA **70**(8), pg. 998 (1980).
5. T. L. Schmitz, A. Davies, and C. J. Evans, "Uncertainties in interferometric measurements of radius of curvature," Proc. SPIE Vol. 4451, p. 432-447, Optical Manufacturing and Testing IV, H. Philip Stahl; Ed. (2001).