Testing highly aberrated large optics with a Shack-Hartmann wavefront sensor

Daniel R. Neal, Paul Pulaski, T.D. Raymond, and David A. Neal

WaveFront Sciences, Inc., 14810 Central S.E., Albuquerque, NM 87123

Quandou Wang and Ulf Griesmann,

National Institute of Standards and Technology, Gaithersburg, MD 20899

drneal@wavefrontsciences.com

ABSTRACT

We have adapted a Shack-Hartmann wavefront sensor (SHWFS) to the measurement of highly aberrated large optics. The experiment uses a concave mirror operating at the radius point with a small lens to re-collimate the light onto the wavefront sensor. It is used to test large (300 mm) fused silica wafers in double pass transmission. The optic under test is placed in the intermediate path near the large return mirror. The aberrations of the large mirror, beam splitter and other optics are subtracted by recording a reference set of focal spot on the SHWFS without the wafer. The wavefront error for some of these wafers is nearly 100 waves, yet we are able to make accurate measurements with the wavefront sensor by selecting a sensor with the appropriate combination of focal length and lenslet diameter. The special sensor that we developed uses a megapixel camera with an arrangement of 100 X 100 lenslets. This sensor could achieve several hundred waves of dynamic range with better than 1/20 accuracy. Additional wafer thickness measurements that were made at NIST with the XCALIBIR interferometer corroborate the SHWFS results.

Keywords: wavefront sensor, Shack-Hartmann, Hartmann-Shack, large optics, optical testing, optical metrology, XCALIBIR interferometer

1 INTRODUCTION

There are many methods that are used for testing large optics. These include Fizeau interferometry, Foucault, Ronchi, star test, Smartt interferometry, Hartmann testing, and others.[1] Modern Fizeau interferometers offer a well-established technical art, with high resolution, excellent accuracy, in addition to standard display and output formats. For testing large highly aberrated optics, however, using an interferometer poses substantial challenges. In some implementations, a Fizeau type interferometer has been used with optics to expand the beam to the size of the optics under test. Such interferometers have been built, but they are located at only a few facilities worldwide. Alternatively, a Ritchey-Common test arrangement can be used. In both cases, it is difficult to measure highly aberrated optics, with up to 100 waves of error. These interferometers also may suffer from laboratory turbulence, vibration, limited dynamic range and other effects that make practical application limited.

For some classes of optic, i.e. concave optics, it is possible to use a reference sphere to create an expanding reference wave that can be matched to the surface curvature of the optic under test. Indeed, a variation on this theme is used to test extremely large telescope mirrors on a routine basis.[2] These test methods still suffer from limited difficulty in alignment, turbulence and vibration effects, however.

An alternative technique that has seen growing popularity in recent years is the Shack-Hartman wavefront sensor.[3] These sensors are based, not on interferometry, but on geometric properties of light that allow robust determination of the wavefront slope. They are sensitive, have excellent accuracy, large dynamic range, and are insensitive to vibration.[4] The SHWFS has the same need to expand the test beam to the size of the optic under test, but has some inherent flexibility that facilitates this measurement.

Thus when the authors were faced with the need to quickly test some extremely thin, 300 mm diameter optical flats, a test arrangement using a Shack-Hartmann wavefront sensor was selected as the method of choice.

1.1 Wafer metrology and nanotopography

In silicon wafer metrology, there are a number of strictly geometrical measurements that have become increasingly important for the production of the next generation of integrated circuits. It is important to be able to measure the overall shape, flatness, and nanotopography of the entire surface of standard silicon wafers. These may be up to 300 mm diameter, and are extremely thin (typically 0.7 mm). During the photolithography process, these wafers are secured to a chuck with vacuum or other means. The wafer complies with the chuck shape during the exposure and other processes (in general). However, because the wafer has an inherent stiffness, the wafer does not perfectly assume the shape of the wafer chuck. Local variations in the wafer thickness, surface metrology or chuck flatness cause variations in the surface shape that is exposed. If these variations are significant, they can exceed the depth of focus of the photolithography instrument and create failures in the fabricated circuits.

Previously we have reported on the Columbus instrument for wafer metrology.[5] This instrument uses Shack-Hartmann technology for these extremely sensitive (sub-nanometer) measurements. This is a good example of the scalable nature of this type of sensor. The Columbus instrument makes measurements over a small area of the wafer, and then scans the wafer by moving it on an accurate xy-translation stage. A small overlap in the measurement regions allows these successive measurements to be stitched together to form a full wafer measurement.[6] Using this technique the full wafer bow and warp (shape), flatness and nanotopography can be determined. The instrument has demonstrated repeatability of about 0.6 nm.

One critical calibration step for a flatness measurement is the measurement of the shape of the chuck. This can be achieved by using the Columbus instrument to measure a set of chucked calibration wafers that have known thickness variation. The known variation can be subtracted from the measurement to provide a net chuck shape. The key then, is to provide calibration flats that have the appropriate compliance with an independently measured thickness variation.

These calibration flats have several characteristics. They need to have the appropriate compliance so they will properly register against the wafer chuck. And the thickness variation of these flats must be accurately characterized. While it is possible to use an ordinary wafer as a reference flat, this poses a couple of problems. Characterizing a silicon wafer thickness is a difficult process, requiring another instrument similar in complexity to the Columbus machine. To measure the thickness, an infrared interferometer, capable of measuring to extremely high precision over a 300 mm aperture would be needed.

As an alternative to using silicon wafers, we elected to use calibration flats from fused silica. The basic idea is to characterize these flats with visible, optical techniques, and then to coat one side with a reflective layer. The flat would then be attached to the chuck and the coated side measured with Columbus.

The body of this paper describes the sensor technology, the setup required for testing these large flats, and some results. While we have applied this technique to the measurement of these particular 300 mm calibration flats, the measurement technique is much more general, and can be applied to many different types of metrology problems.

2 MEASUREMENT PRINCIPLE

The basic measurement principle is similar to the Hartmann test of the early 1900s. In this test, a mask with holes was placed in front of the lens to be tested. Light passing through the holes was examined at two planes, typically before and after the focal plane. By examining the shift in position of the rays compared to that of an ideal lens, the aberrations, wavefront map, MTF, and other parameters could be determined.

In the late 1960s, Roland Shack proposed first shifting the measurement plane to the pupil plane and then using of a grid of lenslets to sample larger areas, while still providing measurements over a localized area. Figure 1 shows the arrangement of a typical modern Shack-Hartmann sensor. In this case a lenslet array, fabricated using photolithography and etching in fused silica, is used to collect the light and direct it onto a CCD array sensor. The grid of pixels on the CCD array provides an accurate measurement of the focal spot positions.

The lenslet array breaks up the incident wavefront into a large number of small sub-apertures. The key assumption is that over each sub-aperture the only wavefront variation is local tilt. This is readily achieved with sufficiently high-resolution lenslets.[7,8]. The light from each of these samples is collected by the lenslet and focused on the detector. Since the region is small, this usually creates a well-formed focal spot whose position is shifted corresponding to the local wavefront tilt. The CCD detector records this focal spot position, and thus, by comparison against a reference, the local slope can be determined. With a large number of local slope measurements, the wavefront surface can be numerically reconstructed.
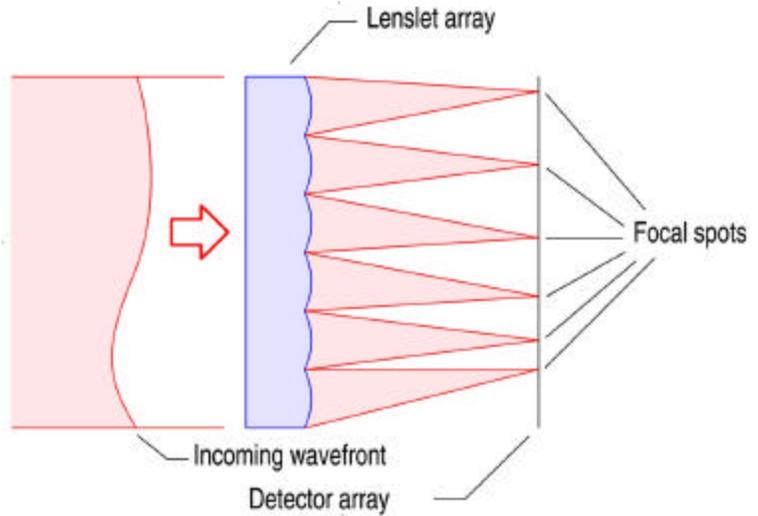

Figure 1 – Basic arrangement of a Shack-Hartmann wavefront sensor

Since the information for all of the focal spots is obtained simultaneously, all of the needed information is obtained in a single CCD frame. With modern CCD camera systems very short exposure times can be used. If there is tilt caused by vibration that occurs between successive frames, it will result in a lateral shift of all the focal spots on the CCD. This is readily identified and subtracted out, or measured if it is useful. The single frame acquisition also means that if the wavefront structures are dynamic (that is, changing rapidly), the instantaneous wavefront will be measured with little error.

The focal spot locations are usually determined by an algorithm called the centroid algorithm:[*]

$$\overline{x}_k = \frac{\sum\limits_{j \in AOI_k} x_j S_j}{\sum\limits_{j \in AOI_k} S_j} \, , \tag{1}$$

where S_j is the modified irradiance distribution over a region AOI_k corresponding to the light from a particular lenslet. A similar equation applies for the y-coordinate of the spot locations. Typically, a threshold algorithm is applied to the irradiance distribution to produce the modified distribution, although other algorithms may apply (deconvolution, for instance).

A reference beam is recorded for use in determining the wavefront gradients from the spot position measurements. Usually this is obtained by recording a plane wave, although the reference may also be calculated numerically. This provides a set of reference centroids x_k^{ref} and y_k^{ref}.

The wavefront gradient for each location k on the sensor is:

$$\begin{pmatrix} b_x \\ b_y \end{pmatrix}_k = \frac{1}{f} \begin{pmatrix} \overline{x} - x_{REF} \\ \overline{y} - y_{REF} \end{pmatrix}_k \tag{2}$$

where f is the lenslet to detector spacing, which is usually set to the focal length of the lenslet.

The wavefront gradients are connected by the assumption that the wavefront is continuous. While there are some situations where this assumption breaks down, for these very small lenslets it is usually quite realistic. Thus for each point k on the sensor lenslet coordinates (x_k, y_k):

---

[*] This is actually a misnomer. It would more accurately be called a center-of-mass algorithm, since it includes a weighted distribution in the calculation, and not just the shape of the boundary. For connection with the literature in this subject, we've continued to use the term centroid to refer to the determination of these spot positions.

$$\nabla w = \frac{\partial w}{\partial x} \vec{i} + \frac{\partial w}{\partial y} \vec{j} = \boldsymbol{b}_k^x \vec{i} + \boldsymbol{b}_k^y \vec{j} \qquad (3)$$

which is just the definition of the gradient in terms of the scalar field $w(x,y)$, except that we have substituted the measured local gradients $\boldsymbol{b}_x$ and $\boldsymbol{b}_y$.

This equation can be solved for the wavefront $w(x,y)$ in a number of ways. The surface can be described in terms of polynomials, and then a least squares fit routine can be used to find the appropriate coefficients. This is the so-called Modal method. Alternatively, the slope data can be used to solve for a self-consistent set of wavefront heights.

3 LARGE OPTICS TESTING

We have chosen an arrangement for testing the optics in transmission, rather than in reflection (as in the Ritchey-Common test arrangement) so that we can obtain a direct measure of the optics thickness variation.
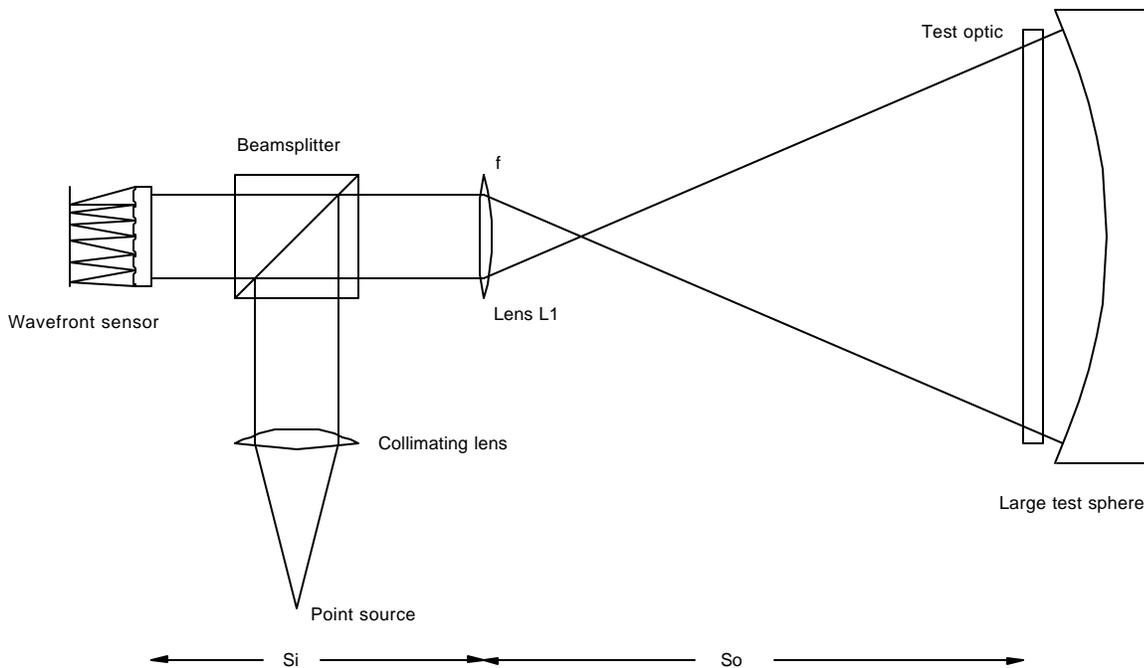

Figure 2 – Test setup for large optics testing

3.1 Test geometry

Figure 2 presents the method that we chose for testing these flats. Since the flats in this case are so thin, we placed them very close to the large spherical mirror. It is often easier to obtain a large spherical mirror than it is to find a large flat. Since the measurement does not take place in collimated space, there is a small correction that must be applied to the measured wavefront. This is derived below. For our measurement, we used a 4.5 m radius of curvature mirror, which results in a rather long optical path. This long optical path required us to set up the experiment across two optical tables on opposite sides of the laboratory. In order to make this measurement, we have taken advantage of several features of the Shack-Hartmann wavefront sensor:

1. Vibration tolerance. Setting up the apparatus across two optical tables would otherwise have posed difficulties for a precision optical measurement. But the SHWFS uses only a single frame for each measurement. So, vibration shows up as just an overall tilt, which is easily subtracted. No vibration isolation is needed.

2. Short snapshot integration. With an integration time of <0.1 ms, we can freeze most any vibration and path turbulence. In fact, we deliberately induced some turbulence using a small fan and then averaged several hundred frames. This has the effect of averaging out any effects due to turbulence and actually creates a more accurate measurement than any given snapshot.

3. Reference subtraction. The reference was recorded with the test flat removed. A reference wavefront (with induced turbulence and averaging) from the spherical mirror is subtracted from all subsequent data measurements. Since the paths are common except for the test optic, the reference data contains the optical errors in the optics.

In order to make measurements that are focused at the proper reference plane, the distance from the lens L1 to the wavefront sensor (S_i) was adjusted so that it obeyed the Gaussian imaging equation:

$$\frac{1}{S_i} = \frac{1}{f} - \frac{1}{S_o}, \tag{4}$$

where S_o is the distance from the lens L1 to the test optic and f is the focal length of the lens. This was a matter of setting the z-position of the wavefront sensor relative to the other optics.

Even though the wavefront sensor position has been adjusted to the image plane, the test optic is not being tested in collimated light in this configuration. Thus the measurement values have to be adjusted to account for the non-collimation. The correction that must be applied to the wavefront slopes is:

$$\mathbf{g} = \frac{\mathbf{b}}{\mathbf{b}_{Meas}} = 1 - \frac{L}{R} \tag{5}$$
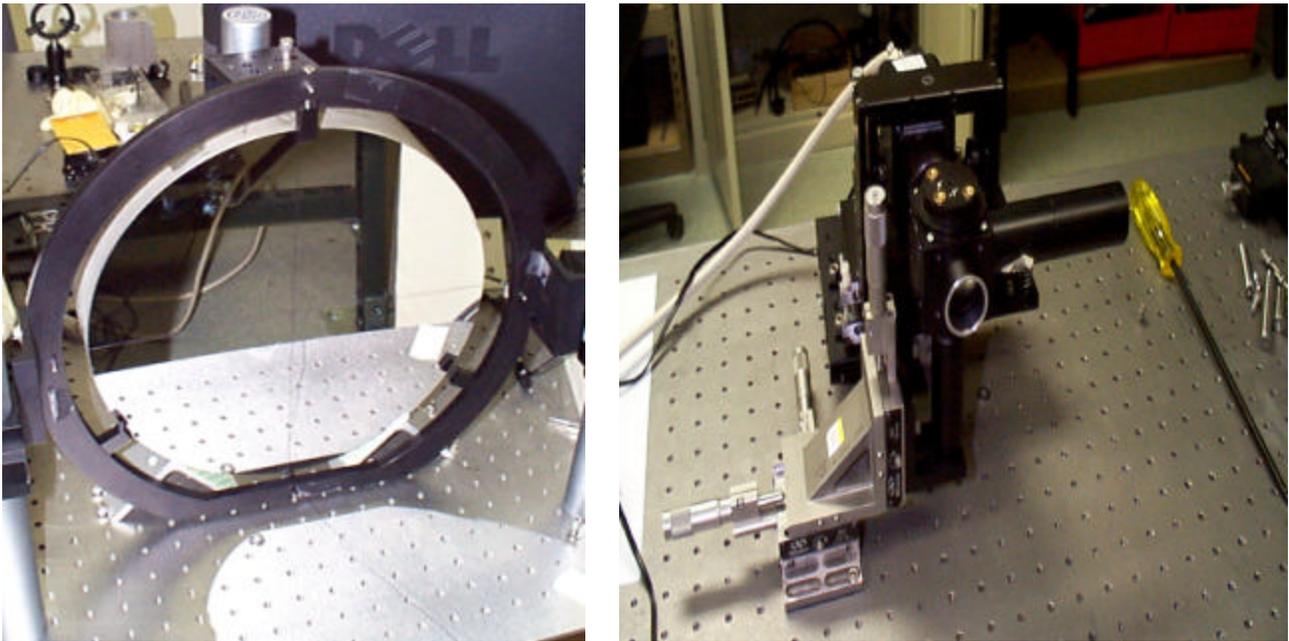
Figure 3 – Photographs of experiment showing the two set up on the two separate tables

Focal length	mm	4.61
Diameter	mm	0.108
Operating Pixel size	μm	7.4
Dynamic range	mrad	5.86
RMS WFE per lenslet	μm	0.00567
4th order Zernike noise floor	μm	0.00567
Focus dyn range (+/-)	μm	21.77

Table 1 – Wavefront sensor design for large optics testing

Figure 4 – Single frame wavefront from the SH LOT of Wafer #1. Note that the image is dominated by turbulence. The p-v wavefront was 12.8 μm.

Where R is the radius of the reference mirror, and L is the distance between the mirror vertex and the part under test. *b* is the wavefront slope according to Eq. 2. As the position of the test part approaches the intermediate focus, a smaller and smaller wavefront slope would be measured on the wavefront sensor for the same aberrations of the part. This is the reason that a field lens (a lens at the intermediate focus) does not introduce any aberrations in an optical system.

For the geometry of our particular test setup, the value of this correction factor is 0.997. This factor should be applied prior to reconstruction.

This test setup will work well for thin optics measured in double pass transmission. It could also be applied to other measurements by rearranging the geometry. It would also be applicable for measuring extremely large optics, such as telescope primary mirrors or large turning mirrors.

3.2   Experimental setup

The wavefront sensor was designed based on the need for high spatial resolution, but also high accuracy. To this end we selected a megapixel camera with 7.4 μm pixels and a lenslet array that would give us high spatial resolution. A summary of the wavefront sensor parameters is given in Table 1. The experiment was set up using a 4572 mm radius spherical mirror, 310 mm diameter as shown in Figure 3. The source was a 635 nm diode laser, pigtail coupled to a single mode fiber. This light was collimated with a 100 mm focal length achromat, and then injected into the optical path through a beam splitter cube. A 100 mm focal length relay lens created a focus and then expanded the light to fill the long radius sphere. The calibration wafers were placed within 12 mm of the reference sphere. The reference sphere collected the light and returned it, through the relay lens where it was re-collimated, and then onto the wavefront sensor. The position of the wavefront sensor was adjusted so that the lenslet array was arranged conjugate to the test optic. The apparatus length was dominated by the radius of the reference sphere and was in total, about 5 m in length.
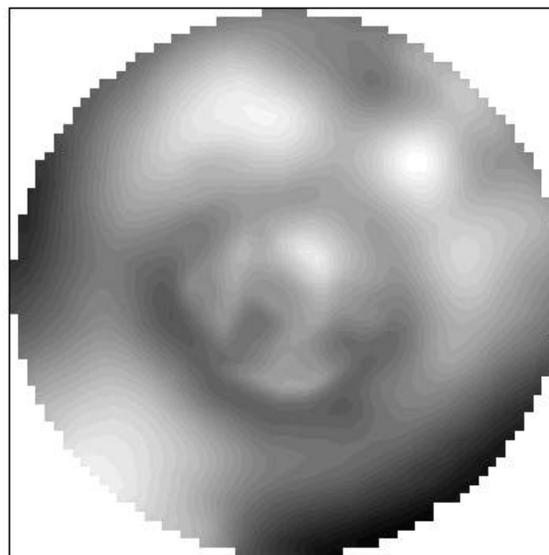
Once the instrument was setup and aligned, a set of

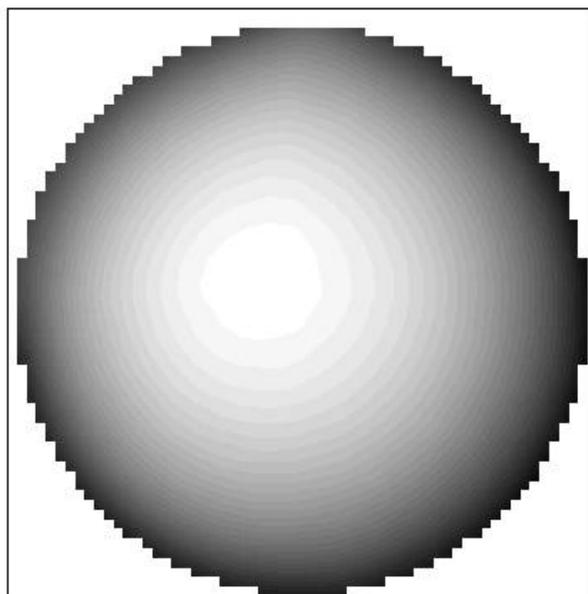**Figure 5 –** 200 frame average of wafer #1 data with induced turbulence. The wavefront p-v was 7.2 μm.

reference focal spot positions was recorded with the test optic removed. 200 frames were averaged to minimize noise, turbulence or other dynamic effects. These centroids were then subtracted from subsequent measurements. This procedure serves to subtract out errors in the reference sphere, relay lens, beam splitter or collimating optics from the test data.

4   RESULTS

Figure 4 show an example wavefront image from one frame of data. The wavefront is clearly dominated by turbulence in the laboratory caused by the long optical path. Rather than try to eliminate this turbulence by controlling air temperature gradients along the path, we took advantage of the short integration time capability of the camera. That is, we deliberately stirred up the air in the optical path with a small fan to induce small scale, rapidly varying turbulence. The sensor was then used to record 200 short exposure images, which were averaged. Since the turbulence induced by the fan had zero mean (verified by repeating the experiment several times), we were able to make a measurement of the average wavefront caused by the imperfections in the part under test. A typical example is presented in Figure 5.

In Figure 5, the wavefront shows an optical path difference where the center is delayed relative to the edges. For the scale chosen in the plot, this corresponds to a higher wavefront in the middle. The sign convention is chosen so that the wavefront measurement matches the physical thickness variation of the part. Thus the wavefront image, after division by two, can be interpreted as a thickness variation map. It is evident from this figure that the test wafer is considerably thicker in the middle than at the edges. With a conventional polishing process, this is what we would expect. The shape of the test wafer is dominated by the large "defocus" term, or parabolic variation in thickness from center to edge. In order to see other variations, this term must be subtracted from the data.

Figure 6(a) shows the wavefront map from this fused silica test wafer with the parabolic terms subtracted. A lot more detailed structure is evident. There is a sharp feature right at the top of the image. This coincides with the position of a locating notch in the wafer, and with one of the mounting points for the test optic. While the wafer mount was designed to hold these flats with very little stress, we were concerned that we had accidentally induced some figure error in the
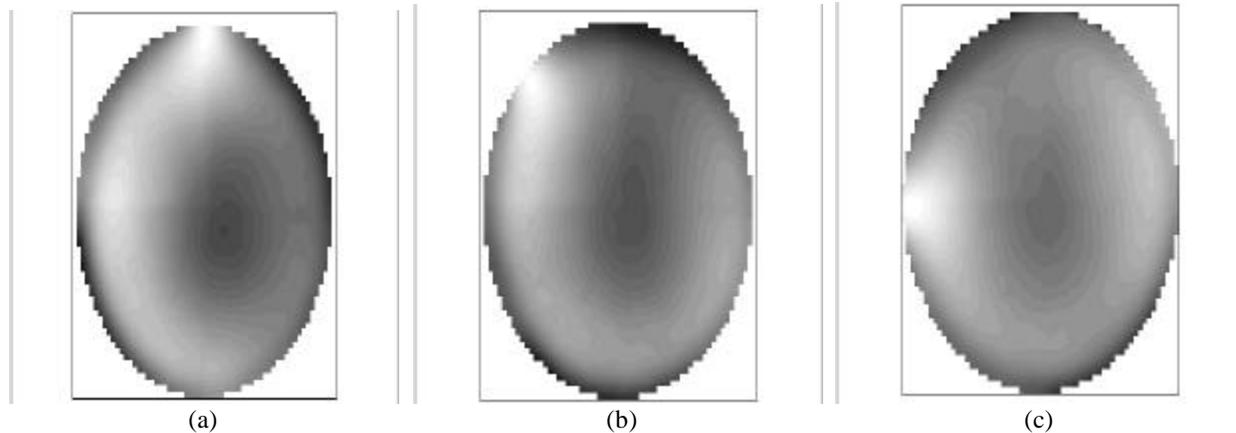

| (a) | (b) | (c) |

Figure 6 – Fused silica wafer #1 measured in previous figure with defocus subtracted. The peak to valley wavefront error was (a) 1.5 µm, (b) 1.7 µm and (c) 2.2 µm. Slight variations in the peak to valley wavefront are due to variations in how the tilt and focus was subtracted from the data.

mounting process. To examine this possibility, we rotated the wafer by 45 and then 90 degrees, and re-measured the data. The corresponding wavefront maps are shown in Figure 6. It is clear from this figure that the wavefront errors rotated with the part, thus eliminating any effects of mounting artifacts from the measurement process.
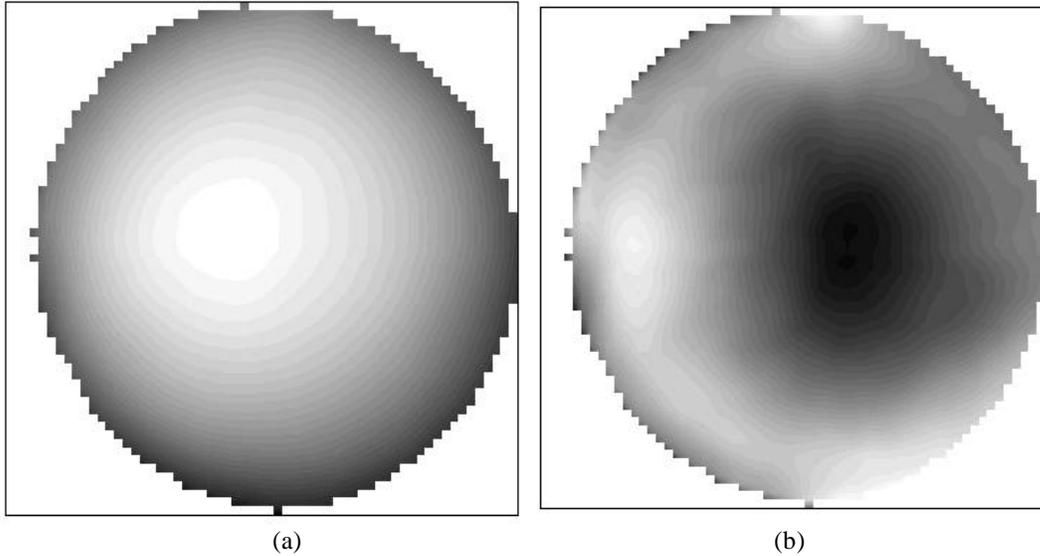
(a)   (b)

Figure 7 – Wafer #2 (a) Full wafer, wf P-V 9.1 µm, (b) focus removed, P-V WFE 1.9 µm.

This apparatus could be used to screen these test wafers for figure error. An example of a somewhat better test wafer is shown in Figure 7. In this figure, 7(a) shows the complete wavefront, while 7(b) shows the wavefront with focus subtracted. This wafer still has some optical figure errors, but this is expected due to the extremely large diameter to thickness ratio (400:1). We tested several test wafers that were manufactured using a similar process (and from the same manufacturer) and found that the wavefront maps were all quite similar, indicating that the figure errors are due to
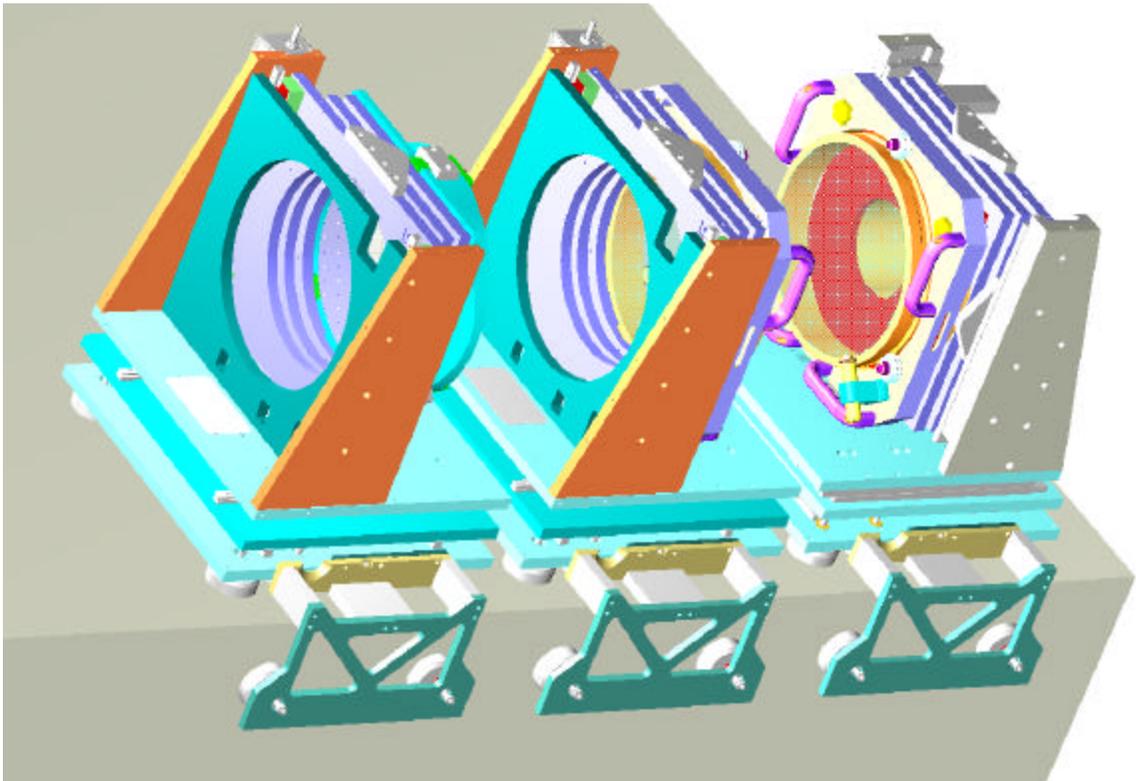

Figure 8 – XCALIBIR 300 mm Fizeau configuration.

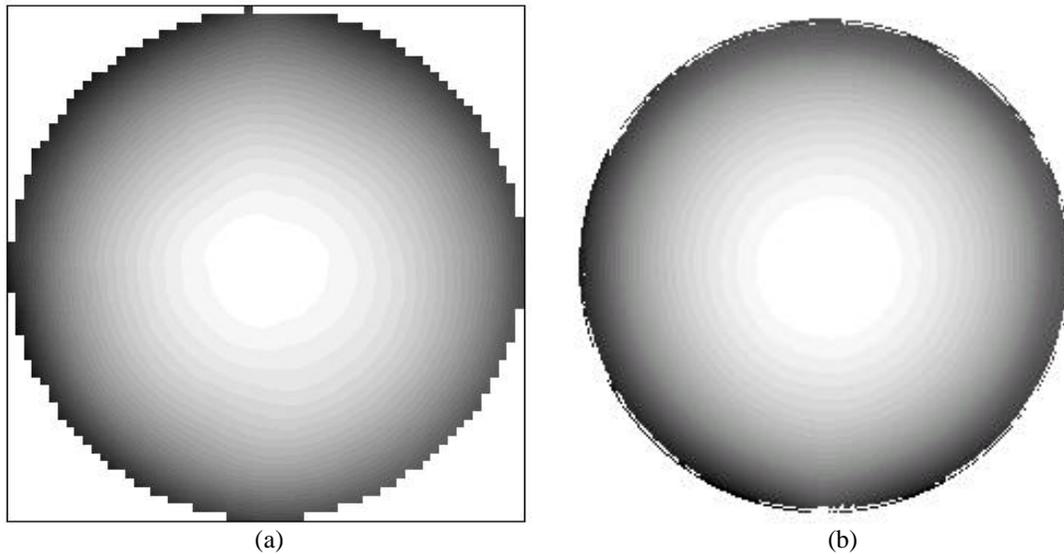|	
(a)	(b)

Figure 9 – Thickness comparison between (a) SHWFS LOT and (b) XCALIBIR measurement for similar wafer (same lot, same manufacturer). In (a) the P-V was 5.75 µm, in (b) 6.2 µm. The measurement uncertainty for the XCALIBER measurement (b) is about λ/100, or 6 nm.

the manufacturing process. We note that wafers produced by a different manufacturer had completely different wavefront distributions. This can be seen in comparing the Wafer #1 wavefront maps (Figure 5 and Figure 6) to the Wafer #2 wavefront maps (Figure 7). While there is some structure that appears similar, we checked that this was not simply due to some error in the optical system by recording a data set with no test wafer in place. In that case, the wavefront is constant to within the measurement system accuracy, which is better than λ/20.

Two of the fused silica wafers from the lot shown in Figure 7 were also measured with the "eXtremely accurate CALIBration InterferometeR" (XCALIBIR) at NIST.   XCALIBIR is a 300 mm aperture multi-configuration interferometer. The glass wafer thickness variation measurements were made in a Fizeau configuration of XCALIBIR. Figure 8 gives a view of the 300 mm collimator lens of XCALIBIR (on the left mount) followed by two mounted 300 mm diameter flats that make up the Fizeau measurement cavity. The fused silica wafer was installed in the measurement cavity of XCALIBIR and the deviation from flatness was measured. This was followed by a second measurement of the errors of the empty measurement cavity. The total thickness variation (TTV) of the wafer is then calculated with the following equation:

$$TTV = \frac{1}{2} \frac{\Delta(wafer) - \Delta(empty)}{n-1} \tag{6}$$

where Δ is the measured optical path difference and n is the refractive index of fused silica at 632.8 nm. Figure 9 shows a comparison of the XCALIBIR measurement with the measurement of Figure 7. The grey scales have been selected to allow direct comparison. Both the peak-to-valley and the distributions match very well between these two. While there are some slight differences, for this experiment we were not able to measure the same wafer, but did measure one from the same manufacturing lot. This level of variation is consistent with the variations that we saw between wafers even in the same lot.

5   CONCLUSION

We have demonstrated a straightforward technique for measuring large, highly aberrated flat optics in transmission. We were able to make rapid measurements of test parts using a Shack-Hartmann wavefront sensor. Comparison of SHWFS measurements with those made using the XCALIBIR interferometer at NIST showed good agreement. This technique,

with minor variations, can readily be extended to measure different optical types such as large spheres, turning flats, or other optics.

6   ACKNOWLEDGEMENTS

7   REFERENCES

1.   D. Malacara, *Optical Shop Testing 2nd Ed.,* John Wiley & Sons, New York.
2.   M. Zecchino, "Ritchey-Common test analyzes uncoated optics," *Photonics Spectra*, p. 34 (July 2003).
3.   D. R. Neal, D. J. Armstrong and W. T. Turner, "Wavefront sensors for control and process monitoring in optics manufacture," *SPIE* **2993**, pp. 211–220, 1997.
4.   D. R. Neal, J. Copland, D.A. Neal, "Shack-Hartmann wavefront sensor precision and accuracy," *SPIE* **4779**, 2002.
5.   T.D. Raymond*, D.R. Neal, D.M. Topa, and T.L. Schmitz, "High-speed, non-interferometric nanotopographic characterization of Si wafer surfaces," *SPIE* **4809**-34, 2002.
6.   D. M. Topa, "Data synthesis using slope and OPD stitching," *SPIE* **5162-18**, 2003.
7.   See for example: D. R. Neal, D. M. Topa, and James Copland, "The effect of lenslet resolution on the accuracy of ocular wavefront measurements," *SPIE* **4245**, 2001.
8.   P. Pulaski, "Making the trade," *OE magazine,* 2003.