Journal of
Micro/Nanolithography, MEMS, and MOEMS

SPIEDigitalLibrary.org/jm3

# Characterization of wafer geometry and overlay error on silicon wafers with nonuniform stress

Timothy A. Brunner
Vinayan C. Menon
Cheuk Wun Wong
Oleg Gluschenkov
Michael P. Belyansky
Nelson M. Felix
Christopher P. Ausschnitt
Pradeep Vukkadala
Sathish Veeraraghavan
Jaydeep K. Sinha

SPIE

# Characterization of wafer geometry and overlay error on silicon wafers with nonuniform stress

**Timothy A. Brunner**
**Vinayan C. Menon**
**Cheuk Wun Wong**
**Oleg Gluschenkov**
**Michael P. Belyansky**
**Nelson M. Felix**
**Christopher P. Ausschnitt**
IBM SRDC
Zip AE1, 2070 Route 52
Hopewell Junction, New York 12533
E-mail: tbrunner@us.ibm.com

**Pradeep Vukkadala**
**Sathish Veeraraghavan**
**Jaydeep K. Sinha**
KLA-Tencor Corporation
Surfscan-ADE Division
Milpitas, California 95035
E-mail: jaydeep.sinha@kla-tencor.com

**Abstract.** Process-induced overlay errors are a growing problem in meeting the ever-tightening overlay requirements for integrated circuit production. Although uniform process-induced stress is easily corrected, nonuniform stress across the wafer is much more problematic, often resulting in noncorrectable overlay errors. Measurements of the wafer geometry of free, unchucked wafers give a powerful method for characterization of such nonuniform stress-induced wafer distortions. Wafer geometry data can be related to in-plane distortion of the wafer pulled flat by an exposure tool vacuum chuck, which in turn relates to overlay error. This paper will explore the relationship between wafer geometry and overlay error by the use of silicon test wafers with deliberate stress variations, i.e., engineered stress monitor (ESM) wafers. A process will be described that allows the creation of ESM wafers with nonuniform stress and includes many thousands of overlay targets for a detailed characterization of each wafer. Because the spatial character of the stress variation is easily changed, ESM wafers constitute a versatile platform for exploring nonuniform stress. We have fabricated ESM wafers of several different types, e.g., wafers where the center area has much higher stress than the outside area. Wafer geometry is measured with an optical metrology tool. After fabrication of the ESM wafers including alignment marks and first level overlay targets etched into the wafer, we expose a second level resist pattern designed to overlay with the etched targets. After resist patterning, relative overlay error is measured using standard optical methods. An innovative metric from the wafer geometry measurements is able to predict the process-induced overlay error. We conclude that appropriate wafer geometry measurements of in-process wafers have strong potential to characterize and reduce process-induced overlay errors.  [DOI: 10.1117/1.JMM.12.4.043002]

Subject terms: lithography; overlay error; wafer geometry; shape slope residual; nonuniform stress; process-induced wafer distortion.

Paper 13142 received Aug. 8, 2013; revised manuscript received Sep. 19, 2013; accepted for publication Sep. 26, 2013; published online Oct. 25, 2013.

## 1 Introduction

As ground rules shrink, state-of-the-art integrated circuit production processes are always challenged to meet ever-tightening overlay error requirements. In the last several years, various multiple patterning schemes are routinely used to break through the $k_1 = 0.25$ half-pitch limit, and superb overlay error between the individual subpatterns is needed to form a final pattern with acceptable quality. Total on-product overlay budgets are expected[1] to approach the 3-nm regime ($|\text{mean}| + 3\sigma$) by 2016. The unforgiving economics of high volume production require that such tight overlay specifications be met at high throughput exceeding 200 wafers/h. The high energy exposure source and the rapid scanning motions in water create difficult challenges[2] for thermal control, and this has driven sophisticated corrective actions for reticle heating[1] and lens heating.[3] While the recent overlay improvements of exposure tools are most welcome, they are not sufficient. Even if the exposure tools are "perfect," it is possible for silicon processing to distort the wafer in a way that limits overlay capability. Such process-induced overlay errors are an increasingly worrisome component of the overlay error budget. A broad goal of this paper is to explore process-induced overlay errors by measuring wafers with controlled nonuniform stress.

We begin by considering the normal alignment process used by many wafer exposure tools, where alignment marks from a previous pattern layer are used for proper positioning of the resist pattern being exposed. A typical alignment measures the position of many alignment targets across the wafer, and then fits the data to a model. Eq. (1) shows a standard linear model with 6 degrees of freedom

$$\Delta x = T_x - \theta_x y + M_x x, \tag{1}$$

$$\Delta y = T_y + \theta_y x + M_y y. \tag{2}$$

The 3 degrees of freedom from translation $T_x$, $T_y$, and rotation $\theta = (\theta_x + \theta_y)/2$ relate to the misalignment of the wafer. The isotropic magnification error $M = (M_x + M_y)/2$ relates to the wafer size change or wafer expansion due to processing. Wafer processing will include a variety of stressed thin film depositions, hot anneals, and other processes which can change the wafer size, and

a typical process flow might see wafer expansions $M$ vary in the $\pm 2$ ppm (parts per million, i.e., $10^{-6}$) range. To appreciate the significance of this size change for overlay errors, note that $+2$ ppm would correspond to $+600$ nm runout error across a 300-mm wafer, leading to gigantic overlay errors if wafer magnification were not compensated. But in current exposure tools, the magnification $M$ is measured to better than 0.01 ppm, and the overlay error correction is within a few nanometers. Even if the processing creates anisotropic stress such that $M_x \neq M_y$, the six term linear model allows excellent compensation. Note that the key assumption in the linear model of Eq. (1) is that the six parameters of the model are fixed values and do not change across the wafer.

But real processes do not result in perfectly uniform stress for many reasons. Perhaps the most fundamental reason is that designed patterns are seldom perfectly uniform, which can lead to stress variations across each image field. Some portions of the chip design might etch away a larger fraction of a stressed film than other portions. Another source of stress variation is film deposition, which is not perfectly uniform across the wafer. High temperature processes such as rapid thermal anneal (RTA) tools can have nonuniform thermal profiles across the wafer which may lead to thermally induced nonuniform stress. In cases of severe thermal gradients at elevated temperature, silicon crystal planes can plastically yield or "slip," resulting in both crystalline defects as well as overlay errors. These stress variations result in wafer distortions which can limit overlay error capability. We have seen in the previous paragraph that simple wafer magnification changes $M$ are routinely corrected with high precision. Certain types of wafer distortions are amenable to more sophisticated nonlinear alignment schemes, but typically these schemes can only address low-order distortions with relatively slow, smooth variation across the wafer. Higher-order wafer distortions are much more problematic and few practical schemes exist for compensation.

In Sec. 2, we consider the measurement of wafer geometry as a method to characterize nonuniform stress. An optical metrology tool will be described which can simultaneously obtain a detailed surface map of both sides of an unchucked free-standing wafer. Nonuniform stress on one side of the wafer will cause nonuniform curvature of the wafer which will be evident in the surface map. When pulled flat on an exposure tool chuck, the nonuniform stress will cause local magnification changes in the wafer, i.e., higher-order wafer in-plane distortion (IPD). A novel metric from wafer shape data, termed the predicted IPD residual (PIR), will be described which strongly relates to measured overlay errors. The final part of Sec. 2 will use finite-element (FE) models to show how wafer geometry changes relate to IPD, which in turn lead to overlay errors. In Sec. 3, the concept is introduced of an engineered stress monitor (ESM) wafer, where a deliberate nonuniform stress is built into the wafer. The ESM fabrication process allows the spatial character of the stress variation to be varied as desired and will be described in detail. A rich set of alignment marks and overlay targets are included on each ESM wafer for detailed characterization. In Sec. 4, we measure wafer geometry and overlay errors of ESM wafers with several types of stress variation. Strong correlations are observed between the PIR (from wafer shape data) and the measured overlay. Finally, Sec. 5 summarizes this work and briefly considers future directions.

## 2 Relationship Between Wafer Geometry and Overlay Error

The geometry of patterned wafers—shape and thickness variation—used in this study was measured using a recently developed product wafer geometry (PWG) tool designed for the metrology of 300-mm patterned wafers. The tool includes a dual-surface metrology system based on the Fizeau interferometry technique where the tool measures both the front and back surfaces simultaneously covering the full surface of the wafer.[4] During measurement, the wafer is held vertically via three point contact at the edges of the wafer, ensuring that the intrinsic free-standing shape of the wafer is maintained with minimal distortion. Wafer shape maps can be calculated as the median surface halfway between the front surface and back surface. Detailed definitions of the wafer geometry of Si wafers are specified by an SEMI standard.[5] Wafer geometry can be classified into components[6] that span different ranges of spatial wavelengths ($\lambda$). For example, roughness of a wafer is defined by very high frequency variations with $\lambda < 0.2$ mm, followed by nanotopography (NT) variations with $\lambda$ ranging from a few tenths of a mm to 20 mm. This paper focuses on wafer shape components with $\lambda > 1$ mm measured by the PWG tool. For this work, and as a practical matter, the median surface is replaced with the back surface data only in order to represent the shape of patterned wafers, thus avoiding the metrology complications of intricate thin film patterns found on the front-side of in-process wafers. In summary, the PWG tool can generate high spatial resolution measurements (>4 million pixels) of shape and thickness variation of process wafers at high throughput suitable for high-volume manufacturing.

Obtaining IPD of chucked wafers from a direct measurement of the out-of-plane deflection, $w$, of free-standing wafers (i.e., wafer shape) is one of the key features of the PWG metrology technique. Hence, we outline underlying physical principles and assumptions behind obtaining IPD vector fields. Silicon (1 0 0) and (1 1 1) wafers possess the in-plane orthotropic symmetry leading to the applicability of isotropic thin plate models to them. Figure 1(a) schematically shows a wafer with a stressed film on the top, along with an unperturbed ideal wafer below. The unperturbed wafer is perfectly flat (uniform thickness and no shape), while the compressively stressed film causes the free-standing wafer to curve into a spherical shape, i.e., the wafer is bowed. Figure 1(b) shows the cross section of a small portion of the wafer illustrating the local height $w(x)$ determined by the PWG tool as the shape of the wafer. The gradient of the local height (slope) gives the angle of the local normal (shown as $A' - B'$ in the diagram), i.e., $dw/dx = -\tan(\phi)$, and this angle leads to the in-plane displacements that we are seeking. Figure 1(c) shows a magnified version of the blue triangle in Fig. 1(b) illustrating how the elastic response of the silicon wafer causes the lateral displacement to depend on the depth within the wafer. The lateral displacement at the top surface, $u^s$, at the median or midsurface $u^0$, and at the neutral surface $u^n$, is related to the geometry of the $A' - B'$ local normal shown in Fig. 1(c). The neutral surface has zero stress by definition, and therefore, the $u^n$

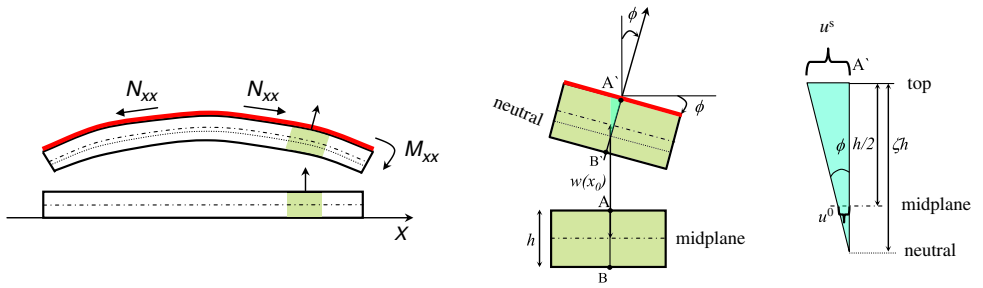



**Fig. 1** Diagram illustrating connection between wafer shape and IPD: (a) schematic of a wafer with a stressed film on the top, (b) cross section of a small portion of the wafer, and (c) magnified version of the blue triangle.

displacements are zero everywhere, i.e., $u^n(x) = 0$. The neutral surface is assumed to be at a depth of $\zeta h$, where $h$ is the wafer thickness and $\zeta$ is a scalar quantity indicating the fractional depth. For a uniform thin film stress on the top of the wafer,[7] the neutral surface is known to be at a depth of $\zeta = 2/3$. The film stress induces a combination of bending and plane-stress deformations to the wafer assuming that vertical shear deformations are small and negligible for typical semiconductor thin film loading. Considering that the wafer displacements pivot about the neutral surface with an angle $\phi$, we can calculate the lateral displacement at the top surface as

$$u^s(x) = \zeta h \tan(\phi) = -[h/2 + (\zeta - 1/2)h]\frac{dw}{dx}$$
$$= -(h/2)\frac{dw}{dx} + u^0(x), \quad (3)$$

$$\text{where } u^0(x) = -(\zeta - 1/2)h\frac{dw}{dx}. \quad (4)$$

With Eq. (4) defining the lateral displacement at the midsurface $u^0$, we can break the top surface displacement into two pieces. The first term $(-(h/2)dw/dx)$ can be identified as the pure bending term,[7,8] representing a pivoting about the wafer midsurface by angle $\phi$. The second term, $u^0$, represents the lateral displacement of the wafer midsurface. So far, all our calculations have been for free, unchucked wafers. But our ultimate goal is to calculate displacements of chucked wafers. An ideal wafer chuck will pull the back surface perfectly flat, and assuming the wafer thickness is uniform, this means that the wafer midsurface will be perfectly flattened. But the pure bending caused by wafer chucking will not affect the $u^0$ displacements caused by the applied thin film stress. For the ideally chucked wafer, all bending terms are eliminated, and we calculate the chucked displacement of the top surface $u^s_{\text{chucked}}$ as

$$u^s_{\text{chucked}}(x) = u^0(x) = -(\zeta - 1/2)h\frac{dw}{dx} \propto -\frac{dw}{dx} \quad (5)$$

indicating that the lateral displacement at the top surface is proportional to the local slope which can be measured by wafer shape metrology. To calculate both $x$- and $y$-components of the IPD, we generalize to a vector displacement $u^s_{\text{chucked}}$, which is proportional to the vector gradient of wafer height $w$,

$$\text{IPD} = u^s_{\text{chucked}} = -c \times \nabla w \propto -\nabla w. \quad (6)$$

For the simple case of uniform thin film loading on wafer front side, the well-known Stoney approximation[7] applies, $\zeta$ is equal to $2/3$, and the slope $c$ in Eq. (6) will be $h/6$, where $h$ is the wafer thickness. But for more general practical applications, we regard the slope $c$ connecting the IPD and the gradient as a variable parameter which can be empirically determined.

Wafers in a semiconductor manufacturing process are subjected to many process steps which can induce stress variations across the wafer, e.g., thin film depositions, RTA processes, etc. These stress variations change the wafer shape, thus rendering them visible to the PWG metrology approach. A real manufacturing process might have many processing steps between layers, so the overlay error would be driven by the accumulated stress changes from all of those processes. For example, a critical overlay error between an active area (AA) patterned layer and a gate patterned layer, would need wafer shape data at AA lithography and also wafer shape data at gate lithography. Based on the preceding considerations, we now describe a novel metric from wafer shape data which can predict IPD of chucked wafers in the lithography scanner, leading to non-correctable overlay errors. Figure 2 schematically outlines

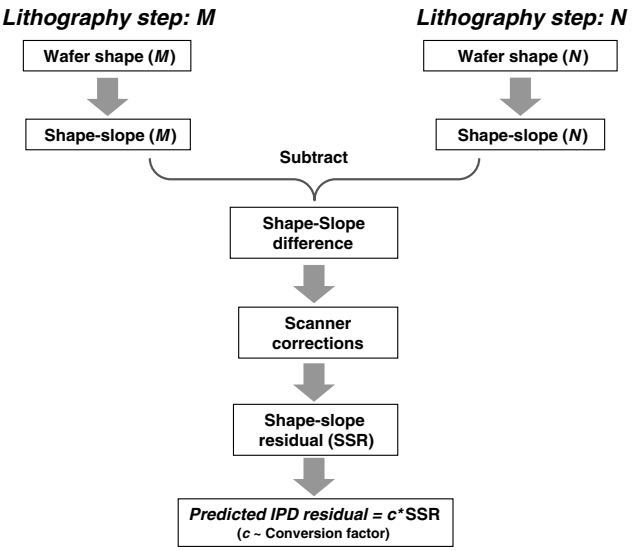



**Fig. 2** Schematic of methodology to calculate predicted IPD residual (PIR) from wafer shape data.

the calculation of this new metric which we call PIR. In order to make predictions of the overlay errors between pattern layer $M$ and pattern layer $N$, we start with wafer shape measurements at both the layers. For each layer, we calculate a wafer shape gradient vector, $\nabla w$ and thus the measured wafer shape map creates a local slope vector map across the wafer. Next, a slope difference map is derived by subtracting the two local slope maps for layer $M$ and $N$. Note that if the shape slope map for layer $M$ is the same as the slope map for layer $N$, then the slope difference map will be zero, and no process-induced overlay error would be predicted. The slope difference map represents the expected IPD from the change in stress that occurred between layers $M$ and $N$, including the change in the uniform stress component assuming that the wafer is pulled perfectly flat by the lithography tool chuck. The chucking performance depends on a combination of the chucking forces and the spatial wavelengths contained in the shape of the wafer being chucked and has been reported elsewhere.[8] But constant magnification components, i.e., simple $M_x$ and $M_y$ components of Eq. (1), will be accurately corrected by the normal scanner alignment process. Therefore, we must subtract such correctable components from the shape slope difference map to obtain a shape slope residual (SSR) vector map, which will correlate most directly to realistic overlay error deviations. The "Scanner Corrections" box of Fig. 2 can mimic any type of alignment process, although most commonly the simple linear models of Eq. (1) are used. If higher-order corrections are applied by a scanner exposure tool then the SSR calculation can be modified to account for these corrections. The final step to predict the PIR is to multiply the SSR by a slope factor "$c$," which may vary for different processes and sources of stress. In general, we recommend determining $c$ from empirical data correlating overlay measurements and SSR data for the specific process under study. Typical $c$ values are the order of $h/6$, meaning that an SSR of 8 nm/mm corresponds to an IPD of 8 nm/mm $\times 0.775$ mm/6 $\approx 1$ nm.

We illustrate this separation of correctable and noncorrectable components in Fig. 3 with IPD (difference) maps from a sample wafer. For simplicity, only the $x$-components of the IPD are shown as color contour maps. By least squares fit, the total IPD map can be easily broken into a linear (correctable) component and a residual (noncorrectable) component. The linear IPD map is dominated by a magnification $M_x$, which is compensated by the scanner corrections as mentioned earlier. The residual component is what we identify as the PIR metric which best predicts the residual overlay errors from wafer stress variation.

In our approach, the PIR is assumed to be proportional to the wafer shape gradient difference. We examine the validity of this assumption by using full scale three-dimensional (3-D) FE models, as described in detail elsewhere.[9,10] Wafer shape measurements were made of silicon wafers at two lithography steps in the process flow. The thickness of the 300-mm wafer was assumed to be uniform with a nominal value of 775 $\mu$m. The FE model includes a vacuum chuck that can apply different chucking pressures as well as accommodate different pin size and spacing values depending on the given vacuum chuck configuration. The wafer shape measurement from the first lithography step was input to the FE model. A nominal chucking pressure of 80 kPa was used, and in a nonlinear 3-D contact simulation, the FE model predicts the IPD at the first lithography step. The FE simulation was repeated for the second lithography step. An IPD difference map was calculated from these two simulated IPD maps from the FE model. Linear scanner corrections, as in Eq. (1), were applied to calculate the residual IPD difference between the lithography steps. Similarly, an SSR map was calculated from the wafer shape map following the procedure of Fig. 2. A comparison between the FE modeled IPD and the PIR from wafer shape data is shown in Fig. 4, showing excellent correlation. Similar comparisons were performed for many different wafer shapes[8] with similar good correlation. For wafers having nominal warp values $\leq$100 $\mu$m, good agreement was observed between FE modeled IPD and the wafer shape derived PIR.

## 3 Processing of ESM

### 3.1 *Fabrication of ESM Wafers and Test Mask Design*

A process is described for building ESM wafers with user-controlled stress variation. We use a silicon nitride film on the top of the wafer as our source of stress. The film is deposited at approximately 40-nm thickness under conditions that create a compressive stress on the order of 3 GPa. When pulled flat on a vacuum chuck, this uniform stress will cause a wafer expansion of approximately 1 ppm, i.e., $\sim 10^{-6}$. ESM wafers are created by patterning the highly stressed nitride film across the wafer such that high pattern density regions retain most of the stress while low pattern density regions relieve the stress. The desired pattern is exposed in photoresist using a scanner exposure tool with wavelength $\lambda = 248$ nm, NA $= 0.8$, and an illumination $\sigma = 0.6$. The

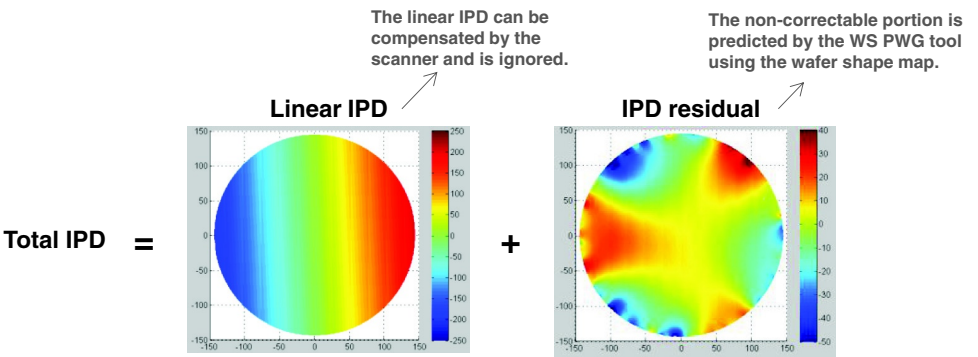



**Fig. 3** Illustration of linear components and higher-order components of $x$-IPD.

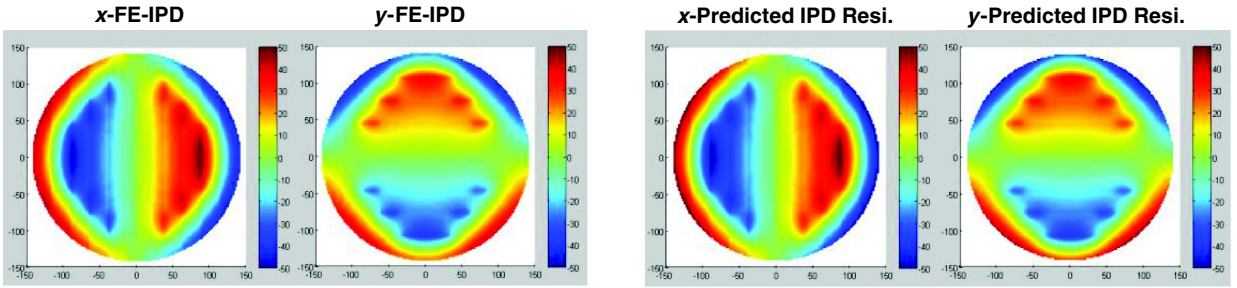



**Fig. 4** Comparison between (a) *x*- and *y*-FE IPD and (b) PIR from wafer shape data. Units in nanometers.

pattern is then transferred into the nitride film via reactive ion etch processing. Details on creating and characterizing highly stressed silicon nitride films have been published elsewhere.[11] With these details and common etching processes, nearly any silicon processing line should be able to build similar ESM wafers.

In order to create controlled pattern density variations across the wafer, we use four different overlay test masks[2] with different pattern densities, as shown in Fig. 5(a). Each mask contains the same set of overlay test targets which sample $13 \times 13$ points across the image field as well as marks for wafer alignment. The 95% pattern density mask can pattern areas which retain most of the stress of the original nitride film, while the 45% density mask and the 20% density mask return more of the wafer to an unstressed state. Unlike the previous masks, all with roughly uniform density across the field, the fourth mask has low pattern density on the left side of the field and high pattern density on the right side. We are able to create ESM wafers with many different types of stress nonuniformities by building up patterns with one or more lithographic exposure passes. In Fig. 5(b), we show a uniform ESM wafer patterned with a simple arrangement of the 45% pattern density images filling the entire wafer. Of course, it is also possible to pattern uniform ESM wafers using the high density or the low density masks. Figure 5(c) shows a wafer patterned with large stress variations in each image field, using the left/right density mask. Finally, Fig. 5(d) shows a radial ESM wafer, where the center area is covered with 95% pattern density images and outer areas use 20% density. Myriad other stress variations are possible to implement by photocomposition of these different mask images across the wafer.

### 3.2 *Describe the Exposure of the Second Level Resist Pattern, and Acquiring Alignment and Overlay Data*

A brief description is given for the experimental measurements of ESM wafers reported in Sec. 4. Wafer geometry data is measured using a standard recipe of the PWG tool, with no special alignment marks or measurement targets required on the wafer. As outlined in Fig. 2, the wafer shape is measured before each lithographic layer, and it is the shape slope change which predicts IPD.

Recall that alignment marks and first level overlay targets have already been etched into the nitride film in the ESM wafer build process. For the second level exposure, we use a standard overlay test mask[2] with shapes that interlock with the etched shapes to form measurable overlay targets. To avoid tool-to-tool and chuck-to-chuck overlay mismatch, the same exposure tool and chuck are used for patterning both levels. The wafers are aligned to etched alignment marks located near the center of 83 fields across the wafer. The raw alignment data is analyzed using the standard linear model, the six linear parameters of Eq. (1) are determined by least squares fit, and the resist pattern is exposed for optimal overlay. Once the second level resist is patterned, we measure overlay targets using industry standard optical overlay metrology tools, with precision[2] on the order of 0.3 nm ($3\sigma$). We use both a full-wafer sampling scheme, where we measure 71 image fields at 25 targets per field, as well as a full-field sampling scheme, where we measure 10 fields at the full $13 \times 13$ sampling across the field. Since, all four masks with different pattern densities contain exactly the same overlay targets and alignment marks in standard locations, the same exposure tool recipes, and the same overlay

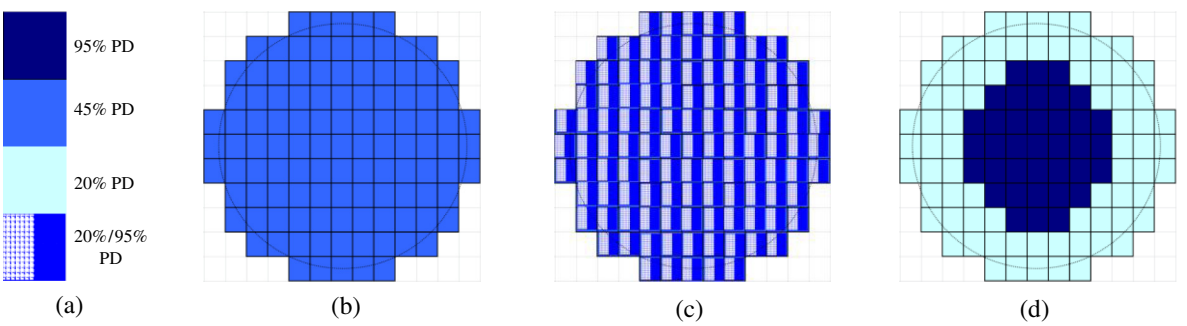



**Fig. 5** ESM pattern layouts: (a) four test masks with varying pattern density, (b) uniform ESM (U45), (c) left/right field ESM (LRF), and (d) radial ESM (RAD).

metrology recipes can be used for ESM wafers of any type, a significant time-saving advantage.

## 4 Correlation of Wafer Shape and Overlay Error

### 4.1 *Wafer Shape Maps for ESM Wafers of Three Different Types*

We now consider experimental measurements from ESM wafers of several different types. We begin by considering the following three exemplary wafers which we name as follows:

- Wafer U45—a wafer patterned uniformly with the 45% pattern density mask, as in Fig. 5(b).
- Wafer LRF—a wafer patterned with the left/right density variation mask, as in Fig. 5(c).
- Wafer RAD—a radial density variation, patterned with the 95% pattern density mask in the center area and 20% pattern density away from the center, as in Fig. 5(d).

All three wafers were measured on the PWG tool just before lithography at pattern level 1 with the nitride layer unpatterned. After the nitride was etched with pattern level 1, thus creating a controlled change in wafer stress, the wafer shape was measured again just before lithography at pattern level 2. The full wafer shape maps from both measurements are shown in Figs. 6(a) and 6(b). For the pattern level 1 data, the stress of the full nitride film causes the wafer to bow with a warp on the order of 100 $\mu$m. After the etch, some of the stress is removed and the warp value is smaller. The changes in the warp of the three wafers are 58, 43, and 32 $\mu$m going from left to right. The differing changes in warp relate to the differences in the level 1 exposure patterns, as in Figs. 5(b)–5(d), which change the amount of film stress removed. The wafer shape data can be used to calculate PIR maps, which are shown in Fig. 7 for the three ESM wafers. Much of the raw shape change visible in Fig. 6 is simple constant curvature, which corresponds to a scanner-correctable magnification change, and is removed in the calculation of PIR. The shape data can be viewed in a completely different way by looking at the NT of the back surface. NT maps[12] are obtained by using a double-Gaussian high-pass filter to extract only high-frequency components of wafer shape with spatial wavelength $\lambda \leq 20$ mm. Figure 8 shows the back surface NT map for all three wafers, which can be directly

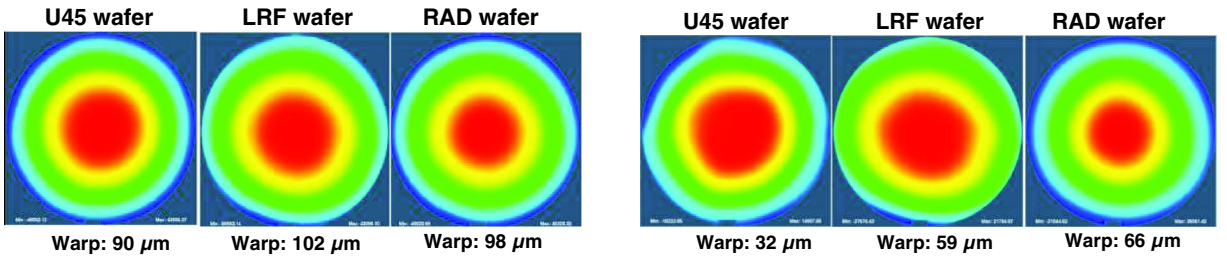



**Fig. 6** PWG measurements of (a) wafer shape at litho. level 1 and (b) shape at litho. level 2.

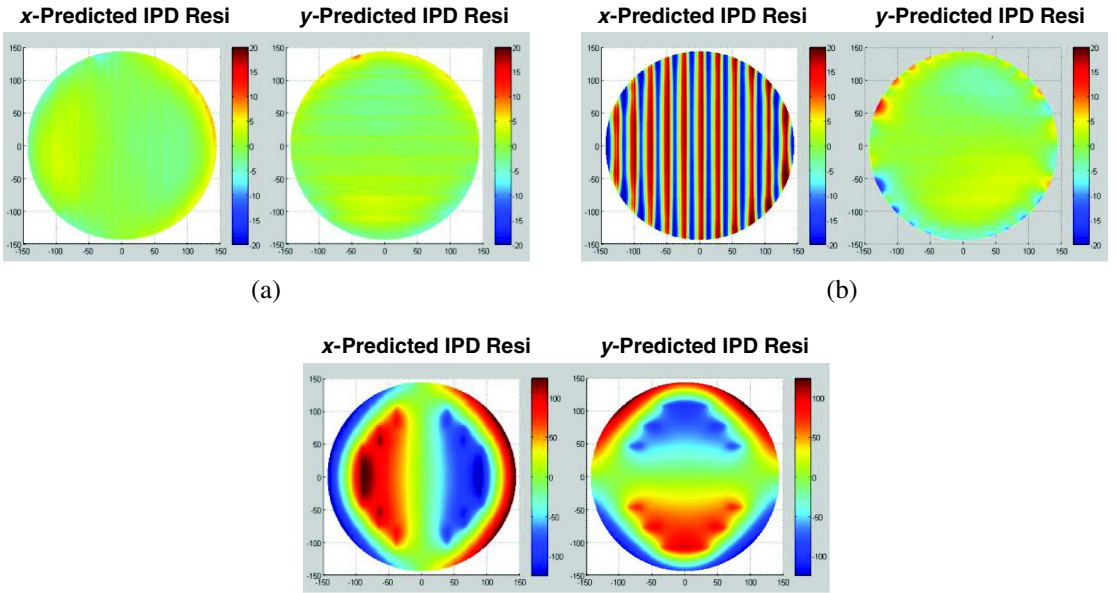



**Fig. 7** PIR maps for the three ESM wafers: (a) U45, (b) LRF, and (c) RAD.

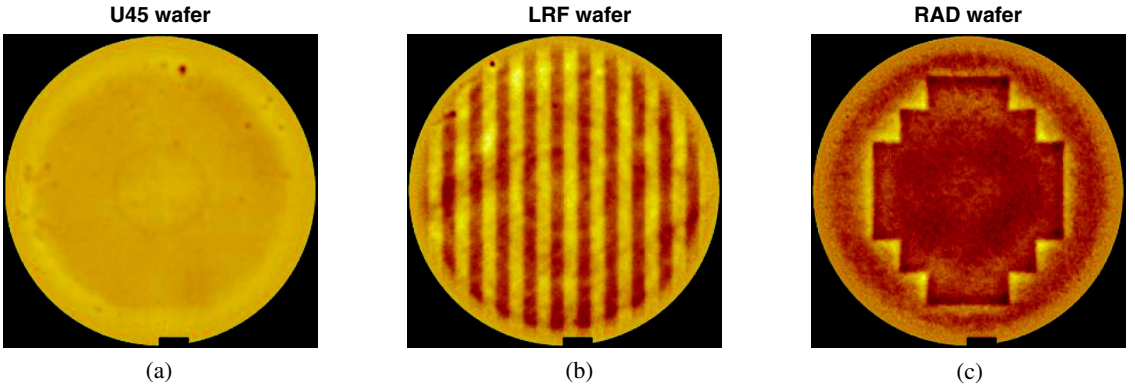



**Fig. 8** Nanotopography (NT) map of back surface for three ESM wafers: (a) U45, (b) LRF, and (c) RAD.

compared to the three ESM styles of Figs. 5(b)–5(d). The NT map highlights the changes in stress designed into these wafers. It is interesting to note that even though stress variations were induced on the front surface of the wafer, the variation propagates through the thickness of the wafer to distort the wafer back surface (by tens of nanometers) on a spatial wavelength $\lambda$ comparable to the wafer thickness.

A comparison of within-wafer 3-sigma values of overlay residuals and PIRs is shown in Fig. 9. The 3-sigma values were computed based on a full wafer sampling plan that includes 1767 valid data points across each wafer. It is seen that there is very good agreement between the overlay and IPD trends. The $x$- and $y$-components of PIRs of the U45 wafer and the $y$-component of that of the LRF wafer are very small since no stress variation was engineered by design. Similarly small $y$-overlay residuals are observed for U45 and LRF wafers. For all three wafers, the overlay residuals are somewhat larger than the PIRs. This is expected because the PIRs arise solely from process-induced wafer distortion, while overlay errors include additional components such as exposure tool errors.

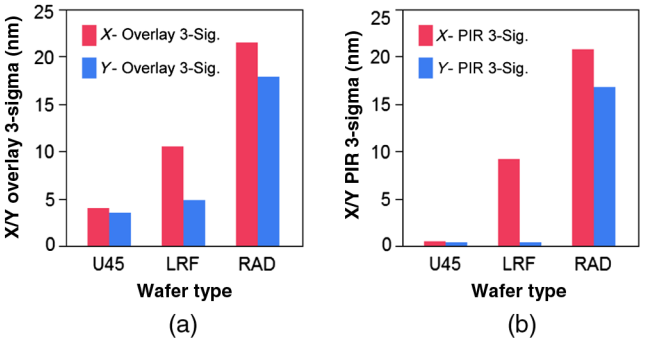



**Fig. 9** Full wafer 3-sigma values of (a) overlay residual and (b) PIR for the three ESM wafers.

## 4.2 *Experimental Data for ESM Wafers with Uniform Stress*

Even though our main interest is in stress variation across the wafer, we now examine wafers with various amounts of uniform stress. These data constitute a "null experiment," where we can directly follow the correction of the magnification errors by the scanner alignment system and also assess experimental noise from factors other than stress variation. We have fabricated three uniform stress wafers using the 95% pattern density test mask (wafer U95), the 45% pattern density test mask (wafer U45), and the 20% pattern density test mask (wafer U20). Figure 10 shows results for uniform ESM wafer U45, where Fig. 10(a) shows the residual

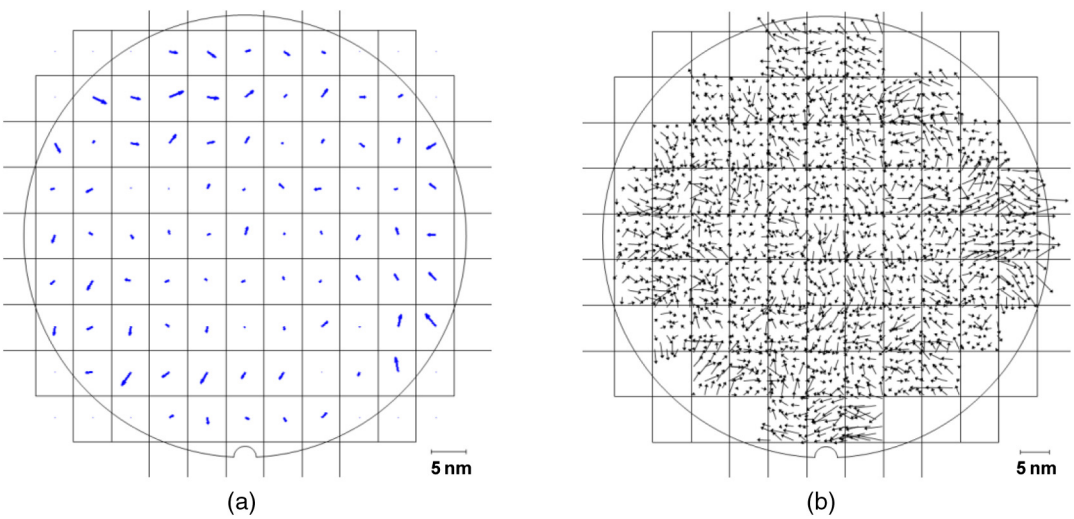



**Fig. 10** Maps of wafer U45 for (a) alignment linear residuals and (b) overlay error linear residuals.

alignment vectors and Fig. 10(b) shows the residual overlay vectors. The alignment residual map, showing one vector per alignment site, demonstrates that the wafer has minimal nonlinear distortion with standard deviation around 2 nm for both $x$- and $y$-components. The overlay residuals are also quite small, on the order of 3 nm ($3\sigma$), with a random character.

Figure 7(a) illustrates the $x$-component and $y$-component of the PIR for wafer U45. Little systematic character is seen in this map, in agreement with the alignment and overlay results of Fig. 10. Yet, we know that wafer U45 had a large bow change between level 1 and level 2, as seen in Fig. 6, which corresponds to a large magnification change of the wafer. Clearly, the alignment process succeeded in removing this large systematic component in the final overlay data, and the scanner correctables removal is essential for the calculation of PIR. All three maps of wafer U45—alignment residuals [Fig. 10(a)], overlay residuals [Fig. 10(b)], and PIR [Fig. 7(a)]—are in agreement that there is little systematic distortion of this wafer.

In Table 1, the wafer warp and the alignment results for five different ESM wafers are shown. The wafer warp just before layer 1 lithography is similar for all wafers, since they all start with a similar blanket nitride coating. However after the nitride etch, the warp is strongly dependent on the pattern density of the mask. The U95 wafer has a relatively small change in the warp because the 95% pattern density nitride film is relatively unchanged. On the other hand, the U20 wafer has a substantially reduced warp, due to etching away roughly 80% of the nitride. The postetch warp of the U45 wafer is intermediate to the low density and high density extremes. The linear magnification parameters $M_x$ and $M_y$ determined by scanner alignment are also reported in Table 1, along with the alignment residuals. All three uniform ESM wafers show that $M_x$ and $M_y$ are closely tracking each other, indicating the expected isotropic magnification change. The U20 wafer has a relatively large change in magnification, since the nitride film has been substantially removed. On the other hand, the U95 wafer has a relatively small change, since the film is only slightly perturbed. Looking at the data for all three uniform ESM, we see that the alignment magnification changes $M_x$ and $M_y$ are approximately proportional to the change in the nitride pattern density. Finally, we observe that for all three uniform ESM wafers, the alignment residuals are relatively small, i.e., better than 2.5 nm [$1\sigma$]. Without even looking at overlay errors, we can already conclude that after correcting for the large magnification changes, the distortion (relative to the reference grid of the scanner alignment system) of all three uniform ESM wafers is quite small.

Table 2 collects the overlay error results for the same five ESM wafers, measured with the full wafer sampling shown in Fig. 10(b). All three of the uniform ESM wafers are able to achieve better than 8-nm ($3\sigma$) raw overlay error distributions, and better than 5-nm ($3\sigma$) residual overlay errors. The magnification parameters determined by the linear fit to the overlay error data are all very small, <0.01 ppm in magnitude. These results demonstrate quantitatively the success of the alignment process in correcting for the different magnification changes of the three different ESM wafers. The good overlay data from these three uniform ESM wafers show that large uniform stress changes do not substantially degrade overlay capability.

**Table 1** Warp values, alignment magnification parameters, and alignment residuals for five ESM wafers.

| Wafer | Warp L1 ($\mu$m) | Warp L2 ($\mu$m) | $M_x$ (ppm) | $M_y$ (ppm) | Resid. $x$-($1\sigma$) (nm) | Resid. $y$-($1\sigma$) (nm) |
|---|---|---|---|---|---|---|
| U95 | 95.1 | 88.3 | −0.066 | −0.070 | 1.67 | 1.62 |
| U45 | 83.3 | 29.1 | −0.620 | −0.594 | 2.00 | 1.97 |
| U20 | 84.6 | 17.4 | −0.826 | −0.821 | 2.46 | 2.10 |
| LRF | 102 | 59 | −0.469 | −0.451 | 2.51 | 2.34 |
| RAD | 98 | 66 | −0.413 | −0.408 | 8.12 | 7.88 |

**Table 2** Raw overlay distribution, magnification parameters, and linear residuals for five ESM wafers measured with full wafer sampling.

| Wafer | Raw $x$-($3\sigma$) (nm) | Raw $y$-($3\sigma$) (nm) | $M_x$ (ppm) | $M_y$ (ppm) | Resid $x$-($3\sigma$) (nm) | Resid $y$-($3\sigma$) (nm) |
|---|---|---|---|---|---|---|
| U95 | 4.37 | 5.23 | 0.005 | 0.010 | 2.8 | 3.12 |
| U45 | 4.12 | 4.03 | 0.008 | 0.006 | 3.4 | 3.19 |
| U20 | 6.41 | 7.57 | 0.005 | 0.003 | 4.82 | 4.42 |
| LRF | 11.61 | 5.52 | 0.004 | 0.005 | 10.5 | 4.82 |
| RAD | 23.3 | 19.2 | −0.033 | −0.036 | 21.5 | 17.7 |

### 4.3 *Experimental Data for ESM Wafer LRF with Left/Right Stress Variation Across the Field*

Wafer LRF represents a test case with extreme stress variation across the field, but little variation across the wafer, since the nonuniform field is repeated across the wafer. The stress is low on the left side of the field and high on the right side, as shown in Fig. 5(c), and results in the vertical stripes visible in the PIR map of Fig. 7(b). Note that the $x$-component map shows strong vertical stripes consistent with the designed stress variation across the field. The $y$-component is relatively flat over much of the wafer, but shows some systematic components near the wafer edge. These edge $y$-components are not surprising, since the designed horizontal stress variation results in some vertical stress gradients near the curved wafer edge.

The results of aligning wafer LRF are shown in Table 1, where we see small linear residual errors comparable to the uniform ESM wafers. The designed intrafield stress variation is not visible in the alignment data because only one alignment vector was measured per field. Note that the overall stress change resulted in large alignment $M_x$ and $M_y$ components on the order of $1/2$ ppm. The overlay error results for wafer LRF, in Table 2, show that magnification errors were almost perfectly compensated for by the alignment process. But, the designed left/right stress variation causes the $x$-overlay residual to more than double relative to the uniform ESM wafers. Figure 11 shows overlay error data for wafer LRF using the full-wafer sampling scheme. Because the stress variation is designed across the image field, Fig. 11(a) plots the average intrafield overlay residual vectors, averaged over the 71 image fields measured, along with the PIR (from wafer shape data) from the same points. Both average vector maps show qualitatively similar character with $x$-components varying with the horizontal position in the field. Figure 11(b) shows point-by-point correlation plots for the measured residual overlay error versus the PIR. The $x$-components are strongly correlated with $R^2 = 0.76$ and RMS-error within 2 nm, since the built-in stress variation is in the horizontal direction. The four bunches of data points occur because the overlay data only samples a few horizontal locations in the field. Sampling more locations across the field would smoothly fill in the correlation plot. By contrast, the $y$-components show little correlation and most of the $y$-overlay error comes from factors other than the stress variation.

More details of the stress-induced overlay error across the field were revealed by measuring wafer LRF a second time using full field sampling, where all $13 \times 13$ targets were measured on 10 fields. Figure 12 shows field overlay error maps averaged over the 10 measured fields. Again the PIR is in good qualitative agreement with the measured residual overlay errors. The character of the overlay errors is consistent with the designed stress variation. The left side of the image field has lower compressive stress than the right side, thus the first level pattern on the left side will exhibit less expansion than that of the right side. Since the overlay error represents second level displacements relative to first level, the left side of the field has a positive $M_x$ character where vectors point outward, while the right side shows a negative $M_x$ character with vectors pointing inward. There is high correlation between the PIRs from the wafer shape and the actual measured residual overlay errors, with $R^2 > 0.9$.

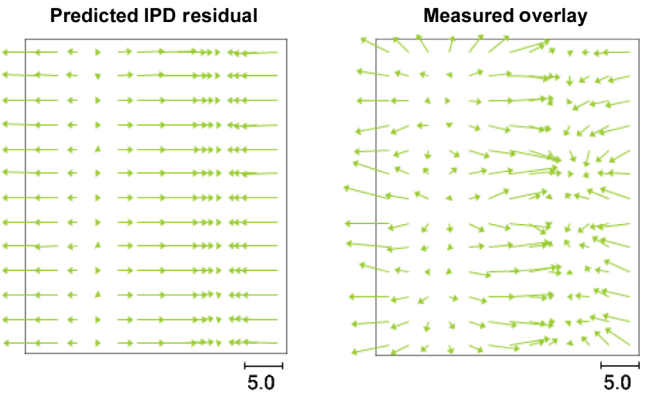
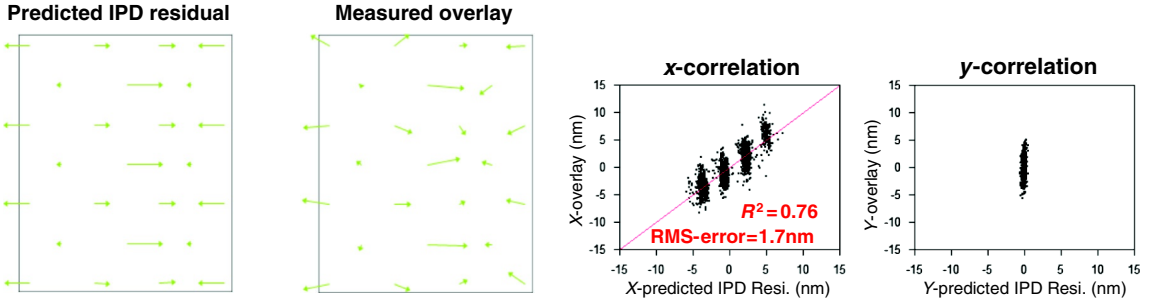



**Fig. 12** Comparison between average-field PIR and measured overlay linear residuals for wafer LRF. The average field was determined by averaging 10 measured fields.

### 4.4 *Experimental Data for ESM Wafer RAD with Radial Stress Variation Across the Wafer*

For wafer RAD, the higher stress in the center area distorts the wafer. This nonlinear distortion is visible to the alignment process, resulting in alignment residuals which are roughly 8 nm ($1\sigma$), much larger than for the other wafers, as shown in Table 1. The large magnification change detected by alignment is almost perfectly corrected, such that the overlay errors summarized in Table 2 show virtually no magnification component. Both the raw overlay errors and the residual overlay errors are on the order of 20 nm ($3\sigma$), indicating that the overlay error is dominated by nonlinear residuals. This is




(a) (b)

**Fig. 11** PIR compared with measured overlay linear residuals for wafer LRF. (a) Vector plots of 25 targets across the field averaged over 71 fields. (b) Point by point correlation plot of $x$- and $y$-components.

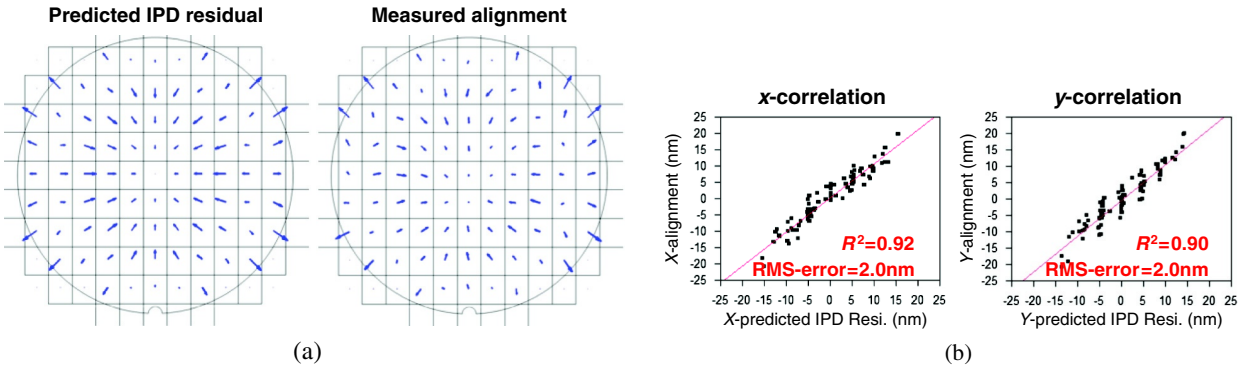



**Fig. 13** Correlation between PIR and alignment linear residuals for ESM wafer RAD. (a) Vector maps showing the 83 alignment sites. (b) Point by point correlation plots for the *x*- and *y*-components.

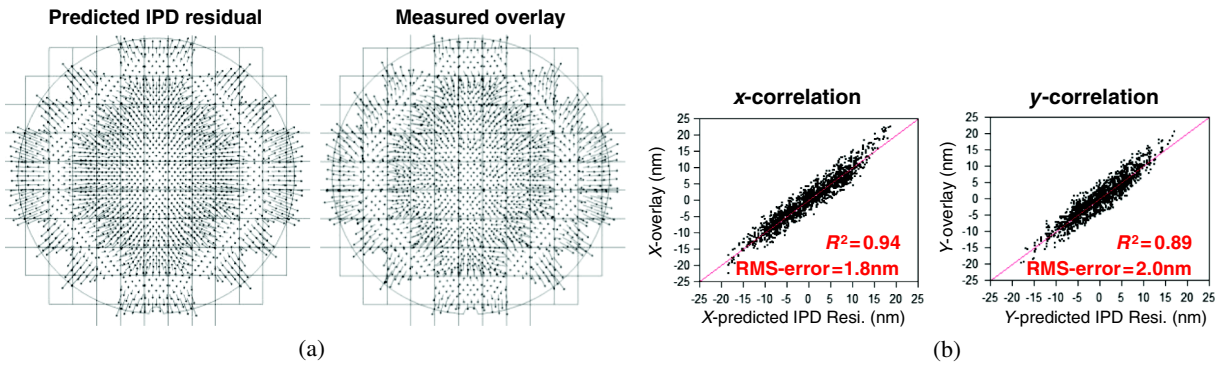



**Fig. 14** Comparison between PIR and overlay linear residuals for ESM wafer RAD. (a) Vector maps at full wafer sampling. (b) Point by point correlation plots for *x*- and *y*-components.

in qualitative agreement with the PIR map of Fig. 7(c) which shows a strong systematic character. By sampling the PIR data at the 83 alignment targets, a more quantitative comparison can be made with the measured alignment errors, shown in Fig. 13. The vector maps, Fig. 13(a), for the PIR are in qualitative agreement with the actual residual alignment data. The point by point linear correlation in Fig. 13(b) is quite strong, with $R^2 > 0.9$ and RMS error within 2 nm. Similarly, we can make comparison to the measured overlay errors by sampling the PIR data at the 1767 overlay targets measured with full wafer sampling. Figure 14(a) shows overlay vector maps comparing the PIR with the measured residual overlay for the full wafer sampling. The agreement is striking, and Fig. 14(b) shows strong correlation with $R^2 > 0.89$ and RMS error within 2 nm. Finally, Fig. 15 shows a similar comparison of PIR and measured overlay at full field sampling, again showing strong correlation.

In summary, the radial stress variation of ESM wafer RAD created significant nonlinear wafer distortion which was well characterized by the PIR from wafer shape data.

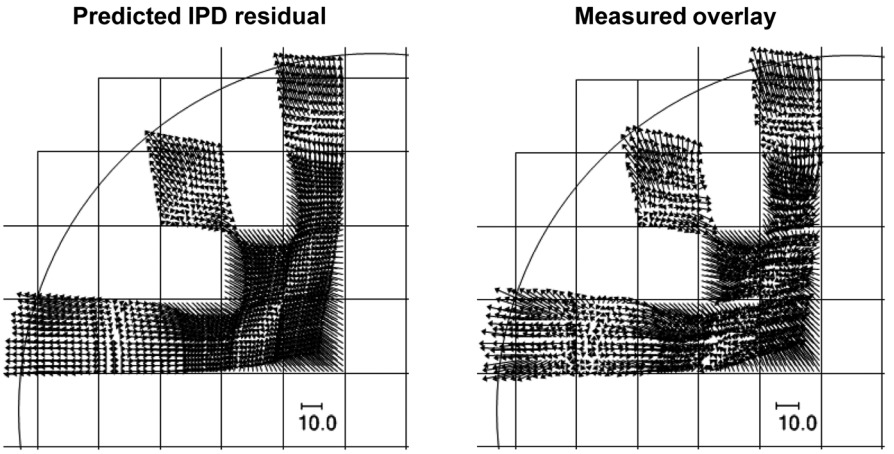



**Fig. 15** Vector maps of PIR compared with measured overlay linear residuals for ESM wafer RAD measured with full field sampling.

This prediction exhibited strong correlations with both alignment residual data as well as measured residual overlay errors. Thus, we demonstrate a capability to quantitatively characterize process-induced wafer distortions via wafer shape metrology.

## 5 Conclusions

Process-induced overlay errors were characterized by measuring the wafer shape of unchucked, free-standing wafers. An optical metrology tool for the measurement of patterned wafer geometry (PWG) was used to obtain high density shape data across full 300-mm wafers. A procedure was described to calculate PIR errors from the wafer shape data, with an explicit subtraction of scanner correctable terms. Special test wafers termed ESM wafers were used to assess the capability of the wafer shape data to predict process-induced overlay errors. A process was described to create the ESM wafers with deliberate stress variations across the wafer. Several specific ESM wafer types were used for the data presented herein. The uniform ESM wafers demonstrated that large magnification errors would be almost perfectly corrected by the scanner alignment process, leaving only small residuals. The nonuniform ESM wafers showed significantly larger overlay residuals due to process-induced overlay, with both the wafer with large stress variation across each image field as well as the ESM wafer with higher stress in the wafer center. For all the ESM wafer data sets, strong correlations were observed between the PIR metric calculated from wafer shape data and the actual measured overlay errors.

We now briefly consider some practical implications of this work and suggest follow-on activities. Traditional investigation of process-induced overlay errors can be a clumsy, expensive, and time-consuming activity. Lithography level $M$ and lithography level $N$ must be printed with suitable overlay metrology targets. Potentially many processing steps must be performed between these two lithography levels and this limits the cycle time for rapid learning. It is also difficult and time consuming to pinpoint which of the processing steps are the root causes of the problem. PWG metrology investigations of process-induced overlay errors have several practical advantages:

- Any processing step can be investigated in detail by measuring wafer shape before and after that particular step.
- Process-induced overlay errors can be predicted with dense sampling across the wafer, without specially designed overlay targets in the lithographic mask designs.
- Data can be obtained at high throughput with a 300-mm wafer scanned in less than a minute.

Thus, the PWG metrology approach can bring rapid learning and efficiency to the difficult process of pin-pointing process-induced overlay errors. Once problematic processes are identified, effective process optimizations can be done with immediate feedback. Novel process tool monitoring and tool qualifications become possible. An even more ambitious vision is to feed forward the PIR error to the exposure tool so as to eliminate or at least mitigate these errors. Modern scanners have many degrees of freedom which might be driven by new information about process-induced overlay errors. The newly introduced ESMs are a versatile platform to mimic different process-induced overlay error patterns, quantitatively characterize such errors, and finally to explore various methods to mitigate those errors.

### Acknowledgments

The authors would like to thank Robert Lang for obtaining all PWG metrology results reported in this paper. We also acknowledge the key role of Craig MacNaughton of KLA-Tencor in championing this approach to characterizing process-induced overlay issues.

### References

1. W. de Boeij et al., "Extending immersion lithography down to 1x nm production nodes," *Proc. SPIE* **8683**, 86831L (2013).
2. B. Minghetti et al., "Overlay characterization and matching of immersion photoclusters," *Proc. SPIE* **7640**, 76400W (2010).
3. S. Halle et al., "Lens heating challenges for negative tone develop layers with freeform illumination," *Proc. SPIE* **8326**, 832607 (2012).
4. K. Freischlad, S. Tang, and J. Grenfell, "Interferometry for wafer dimensional metrology," *Proc. SPIE* **6672**, 667202 (2007).
5. Semiconductor Equipment and Materials International, "SEMI M49-0312 Guide for specifying geometry measurement systems for silicon wafers for the 130 nm to 22 nm technology generations," www.semi.org (2012).
6. P. Vukkadala, K. T. Turner, and J. K. Sinha, "Impact of wafer geometry on CMP for advanced nodes," *J. Electrochem. Soc.* **158**(10), H1002 (2011).
7. P. H. Townsend, D. Barnett, and T. A. Brunner, "Elastic relationships in layered composite media with approximation for the case of thin films on a thick substrate," *J. Appl. Phys.* **62**(11), 4438–4444 (1987).
8. K. T. Turner, R. Ramkhalawon, and J. K. Sinha, "The role of wafer geometry in wafer chucking," *J. Micro/Nanolithogr., MEMS, MOEMS* **12**(2), 023007 (2013).
9. K. T. Turner, S. Veeraraghavan, and J. K. Sinha, "Relationship between localized wafer shape changes induced by residual stress and overlay errors," *J. Micro/Nanolithogr., MEMS, MOEMS* **11**(1), 013001 (2012).
10. K. T. Turner, S. Veeraraghavan, and J. K. Sinha, "Predicting distortions and overlay errors due to wafer deformations during chucking on lithography scanners," *J. Micro/Nanolithogr., MEMS, MOEMS* **8**(4), 043015 (2009).
11. M. Belyansky et al., "Methods of producing plasma enhanced chemical vapor deposition silicon nitride thin films with high compressive and tensile stress," *J. Vac. Sci. Technol., A* **26**(3), 517–521 (2008).
12. Semiconductor Equipment and Materials International, "SEMI M43-1109 guide for reporting wafer nanotopography," www.semi.org (2009).

**Timothy A. Brunner** has been working in the area of optical lithography since 1981, with particular interests in advanced image formation, simulation, process control, metrology techniques, and interdisciplinary aspects of lithography. He received his BA from Carleton College in 1975 and his doctorate from MIT in 1980, both in physics. He then worked on the measurement of lithographic tool performance at Perkin-Elmer Corp. and on amorphous silicon TFT processes at Xerox PARC. In 1988, he joined the research staff of IBM, and currently works in the Advanced Imaging and TCAD Department of the IBM SRDC. Dr. Brunner is a member of SPIE, IEEE, and OSA.

**Vinayan C. Menon** has been working in the area of photolithography since 2003, focusing on process development and tool/product controls. He received his BTech in chemical engineering from Indian Institute of Technology Madras in 1990, and his doctorate in materials science and engineering from the Pennsylvania State University in 1996. He then worked on commercializing Aerogels via novel sol-gel processing at NanoPore Inc., and on engineering LPCVD and thermal

anneal processes at Philips Semiconductors. In 2003, he joined the IBM Microelectronics Lithography team, developed and maintained ultra low $k$ stand-alone track processes, implemented and rendered integrated after-develop-inspection (ADI) across fab lithography clusters, developed advanced patterning processes and transferred to manufacturing at multiple technology nodes. He currently works on improving scanner alignment and on-product overlay in the Lithography Tooling and Variance Controls group at IBM Semiconductor Research and Development Center.

**Cheuk Wun Wong** is originally from Hong Kong. He received a BS degree in chemical engineering from the University of California at Berkeley in 2009 and initially worked for energy storage and solar cell start-up companies in the Silicon Valley. He has been working in the area of semi-conductor lithography at IBM Corporation in East Fishkill, New York, since December of 2011. His area of expertise is advanced overlay metrology.

**Oleg Gluschenkov** is a senior scientist at the IBM Semiconductor Research and Development Center, with the primary responsibility for advancing diffusion and thermal processing technology. He received his MS and PhD degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1997 and 1999, respectively, and his MS degree in applied physics and mathematics from the Moscow Institute of Physics and Technology in 1992. In 1999, he joined IBM Microelectronics, where he made technical contributions in the areas of nonplanar DRAM transistors, poly-metal gates, laser annealing, metastable stressors, and process-layout interactions. He is a member of IEEE and holds over 120 US patents.

**Michael P. Belyansky** joined IBM Semiconductor R&D Center in 2000 after finishing his postdoctoral work at SUNY-Albany Nanotech Center. He has been working on thin film process development for several generations of IBM DRAM and logic technologies. He is currently leading the development of dielectric materials and processes for advanced CMOS technologies with particular interests in atomic layer deposition, strain engineering, and process control. He holds 30 patents and has published over 20 papers. He received his BS degree in chemistry from Moscow State University in Russia and a PhD degree in physical chemistry from University of Illinois, Chicago.

**Nelson M. Felix** is currently manager of the development lithography tooling and variance control group in IBM's Semiconductor Research and Development Center. He received his BS from the University of Massachusetts, Amherst (2002) and his PhD from Cornell University (2007), both in chemical engineering. After initial research work in novel photoresist materials and processes, he began working at IBM as a development engineer for overlay and CD metrology applications, where he steadily grew his role in the area to include issues relating to overlay and exposure control.

**Christopher P. Ausschnitt** recently retired from his position as a senior member of the technical staff at the IBM Microelectronics Division. During his career, he specialized in lithography, metrology, and process control. He received his BS degree from Princeton University in 1968, his MS from the University of Pennsylvania in 1970, and his ScD from MIT in 1976, all in electrical engineering. Prior to joining IBM in 1989, he worked as a staff scientist specializing in quantum electronics at Bell Laboratories (1976–1979), a manager of lithography equipment development groups at Perkin-Elmer (1979–1984) and GCA (1984–1986) and the process control systems group at Shipley Company (1986–1989). He is the author of numerous technical publications and has been granted more than 40 US patents.

**Pradeep Vukkadala** received his PhD degree in mechanical engineering from the University of Wisconsin-Madison (UW-Madison) in 2010. Here, his research focused on developing structural-mechanics based correction strategies for overlay errors induced by mask distortions in EUV lithography. He also holds an MS degree in mechanical engineering and industrial engineering from UW-Madison. In 2010, he joined the Advanced Technology Group of the KLA-Tencor Wafer Inspection Division. His current interests include metrology of wafer geometry and its influence on the performance of advanced semiconductor devices.

**Sathish Veeraraghavan** received his BTech degree in mechanical engineering from Indian Institute of Technology (IIT), Madras in 2005, and his MS degree in mechanical engineering from University of Wisconsin-Madison, in 2008. He is an advanced technology engineer in the Wafer Inspection Group, KLA-Tencor Corp. He has more than 5 years of experience in the field of wafer metrology. His primary focus is to understand the relationship between wafer geometry and semiconductor processes and to develop novel wafer metrology metrics for semiconductor process control.

**Jaydeep K. Sinha** received his PhD degree in mechanical engineering from Auburn University. He also holds an MS degree in physics and technology management. He was chief scientist for ADE Corporation and currently is with KLA-Tencor Wafer Inspection Group, where he leads the advanced technology group. His primary interest is semiconductor metrology and he has been involved in this field for over 15 years. He has been named inventor for over 10 patents and has authored over 30 publications. He also co-chairs Semiconductor Equipment and Materials International Standards (SEMI) Advanced Wafer Geometry (AWG).