Mitigating Process-Induced Wafer Distortion for Enhanced Overlay and Thickness Uniformity

Kangjian Cheng, Jingyan Huang, and Wen Siang Lew, Senior Member, IEEE

***Abstract*—Temperature nonuniformity during rapid thermal processing (RTP) induces localized thermal stress and plastic distortion, leading to lithographic overlay errors and yield degradation. In this work, we report a systematic methodology to detect and mitigate process-induced wafer distortion and residual overlay errors without compromising film thickness uniformity. In-plane distortion (IPD) data are used to quantitatively correlate nonuniform residual wafer shape changes with downstream overlay performance with high accuracy. By optimizing the pyrometer T7 offset, residual overlay is reduced to 24.2% and 20.2% of the original values in the X and Y directions, respectively. While conventional tuning typically compromises wafer edge thickness, the incorporation of a side gas injector enables simultaneous improvement in oxide thickness uniformity from 1.33% to 0.18%. This integrated approach demonstrates a practical method for enhancing RTP control and improving overlay performance in advanced semiconductor manufacturing.**

***Index Terms*—In plane distortion, rapid thermal process, residual overlay, temperature uniformity, thermal oxidation**

I. INTRODUCTION

SINCE the 1970s, in-plane silicon wafer distortion has remained a persistent concern [1, 2]. With continuous device scaling, increasing wafer sizes, and the integration of higher functionality within a single chip, state-of-the-art integrated circuit production processes face growing challenges in meeting the ever-tightening requirements for lithographic focus and overlay [3–9]. Overlay errors, dislocations, and focus errors have become major yield limiters. Overlay-induced misalignment can cause significant functional loss, as shown in Fig. 1, and dislocations degrade transistor performance, device yield, and reliability [5, 8, 10–12].

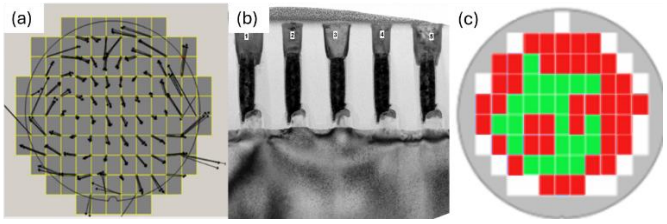

Fig. 1. (a) Overlay map showing poor edge overlay performance on a 300 mm wafer; (b) cross-sectional TEM revealing contact-to-device misalignment; (c) map showing overlay-indued edge yield loss.

This work was supported by the EDB-IPP under Grant No. REQ0377349. (Corresponding author: Wen Siang Lew)
Jingyan Huang is with GlobalFoundries Singapore (jingyan.huang@globalfoundries.com).
Kangjian Cheng and Wen Siang Lew are with the School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore (e-mail: kangjian001@e.ntu.edu.sg; wensiang@ntu.edu.sg).

Although lithography tools and reticle designs continue to advance, enabling accurate correction of low-order magnification, rotation, and skew errors, their capability to correct nonlinear in-plane distortions (IPDs) remains limited [1–9]. Advanced higher-order process correction (HOPC) and correction-per-exposure (CPE) techniques also cannot fully resolve high-order residual overlay errors [10–12]. Even with ideal exposure tools and reticle designs, silicon processing itself induces IPDs that limit overlay performance [2]. As overlay requirements continue to tighten, nonuniform process-induced overlay errors have become an increasing challenge, contributing 25%–30% of total overlay variability and highlighting the need for improved process control [6, 13, 14].

Understanding the fundamental relationship between localized wafer shape variations and lithographic errors is essential, particularly at advanced nodes where feature dimensions continue to shrink [15]. Wafer processing can induce both IPDs and out-of-plane distortions (OPDs) in the wafer surface topology. Typically, high IPDs degrade lithographic focus or cause in-plane deformation that results in overlay misalignment [16–20]. Process-induced overlay errors are primarily driven by wafer distortions originating from nonuniform stress [2, 17, 21, 22]. Such stress nonuniformity arises from within-wafer temperature variations during the heat-up, soak, and cool-down stages of thermal processing, which can lead to IPDs, overlay errors, and even crystallographic slip. Thus, the fundamental challenge is improving within-wafer temperature uniformity. This work focuses on detecting within-wafer temperature variations, improving uniformity, and mitigating the associated process trade-offs.

The conventional approach for studying process-induced overlay requires a long and costly learning cycle. Because overlay errors represent differential positional differences between successive patterning steps, each sample must undergo at least two lithography exposures and all associated process steps. Numerous studies have modeled thermal stress during the rapid thermal processing (RTP) of patterned wafers to improve thermal uniformity [23–27]. Other studies have focused on directly modeling nonuniform residual stress, wafer shape changes, and overlay errors [11, 17, 28–30]. Nevertheless, localized stress is also influenced by specific device patterning and the wafer backside film stack. Developing a dedicated model for each individual product and layer is therefore impractical for a manufacturing fab.

In contrast, patterned wafer geometry (PWG) metrology enables a much faster learning cycle by measuring wafer geometry immediately before and after the process step of interest [2]. It provides quantitative insight into process-induced distortions, and IPD data can be directly correlated with overlay performance [4]. Section III demonstrates how

PWG measurements can be applied to identify process-induced overlay issues.

Meanwhile, rapid thermal processing (RTP) has attracted substantial interest for key VLSI processes such as implant activation annealing and thermal oxidation since the mid-1980s [31]. As device dimensions continue to shrink, the advantages of RTP over batch furnace processes, such as reduced dopant redistribution and tighter thermal control, have become increasingly critical [24, 32, 33]. However, achieving uniform wafer heating remains a major challenge. Nonuniform heating induces thermal plastic deformation, which subsequently causes overlay errors, particularly at the wafer edge [1, 2, 33–39].

Various methods have been proposed to mitigate RTP-induced wafer distortion, including lowering the temperature ramp rate [34, 36, 40], adding shielding rings [24, 41, 42], optimizing susceptor geometry [43], and implementing wafer backside heating [44]. However, these techniques only partially alleviate overlay issues, and conflicting results persist. Some studies report no correlation between ramp rate and pattern alignment [35, 45], whereas others indicate that lowering the ramp rate compromises RTP's intrinsic advantages [42]. Backside heating also yields inconsistent benefits due to emissivity differences and localized backside cooling [46]. Furthermore, tightening temperature-uniformity requirements from ±5 °C to ±0.5 °C and beyond makes achieving uniform heating increasingly difficult [23, 41, 47]. Currently, zone-specific lamp-power tuning remains the only reliable way to improve overlay, but it compromises oxide-film uniformity and electrical performance [48].

Given these limitations, a controlled benchmark process is required to systematically optimize temperature uniformity without compromising film quality. In this work, rapid thermal oxidation (RTO) is employed as the benchmark process. In the next section, we demonstrate the temperature-tuning approach to improve overlay performance, followed by a method to restore film uniformity while preserving temperature uniformity. These findings provide valuable insight into optimizing both thermal and film uniformity in RTP. Although the results are shown on a specific type of RTP tool, the methodology is broadly applicable to any lamp-heated single-wafer system.

II. Experimental Setup and Methods

A second-generation PWG model (PWG2) from KLA is used for wafer shape characterization. The PWG measurement setup is schematically illustrated in Fig. 2(a). The wafer is held vertically to minimize gravitational distortion and chucking forces, enabling measurement of the intrinsic wafer shape. Both front and back surfaces with 2 mm edge exclusion are measured simultaneously using two Fizeau interferometers operating at a wavelength of 635 nm. Interference fringes are captured by two cameras. The residual PWG-IPD used in this study was derived from differential wafer shape maps measured immediately before and after the process step of interest. The shape changes as shown in Fig. 2(b) was converted to local slope components first, followed by KLA's proprietary 10-term field-correction model to remove linear correctable components. The resulting residual values were used for data analysis.

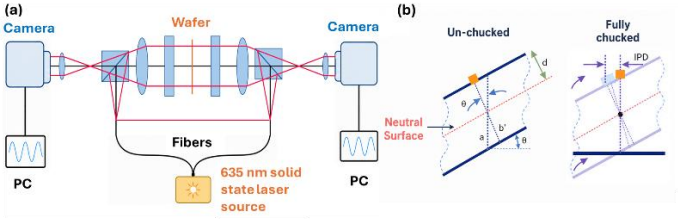

Fig. 2. (a) Schematic of PWG measurement setup; (b) IPD calculation.

The silicon oxide film thickness was measured using a standard KLA Tencor Aleris 8510 model ellipsometer. For patterned production wafers, the outermost measurement pad on scribe line is located at a radius of 144.8 mm, corresponding to approximately 5 mm edge exclusion. For blanket test wafers, a 3 mm edge exclusion was used for thickness measurement.

Multiple 55 nm Bipolar-CMOS-DMOS (BCD) devices were used as test vehicles to investigate the effects of RTP on lithographic overlay performance. All experiments were conducted on a 300 mm RTP system. During ISSG oxidation, wafers were loaded onto a ceramic edge ring rotating at 320 rpm and heated by 409 tungsten halogen lamps located on the front side of the chamber. The lamp array was divided into 15 zones, as illustrated in Fig. 3(a). Wafer temperature was monitored by seven pyrometers located beneath the wafer backside, providing closed-loop feedback control. The pyrometer locations are shown in Fig. 3(b), with Probe 7 positioned near the wafer edge. ISSG oxidation was carried out at 1000 °C under a chamber pressure of 10 Torr. Within-wafer temperature uniformity was optimized by adjusting seven recipe-specific pyrometer temperature offsets. In this study, the offset of pyrometer Probe 7, located at the wafer edge, was systematically adjusted for each experiment.

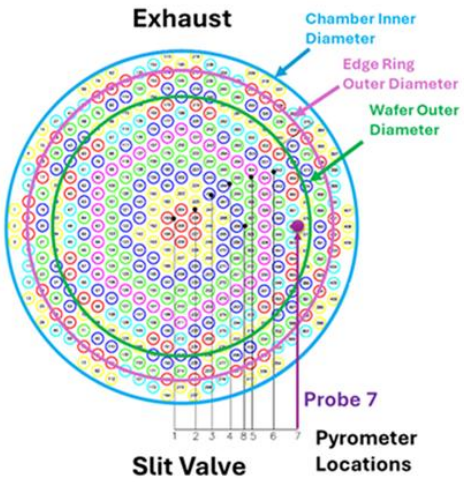

Fig. 3. Schematic illustration of 409 lamp set configuration and locations of seven pyrometer probes, with Probe 7 at the wafer edge.

III. Results and Discussion

In the following analysis, the residual IPD (PWG-IPD) is defined as the non-linear difference between the pre- and post-process PWG measurements. The residual overlay refers to the uncorrectable nonlinear component remaining after standard lithographic corrections.

To establish the quantitative relationship between residual PWG-IPD and downstream lithographic performance, the

temperature of the pyrometer at the wafer edge (T7) during RTP was systematically varied on Device A to generate multiple datasets.

The analysis begins with a qualitative visual assessment. As shown in Fig. 4, all PWG-IPD maps exhibit discontinuous high-magnitude signals along the extreme wafer edge, which are likely associated with edge-related measurement artifacts. Therefore, the subsequent analysis focuses on the wafer-center region, where the measurements are more reliable.

Fig. 4(a) and (b) present the 3D PWG-IPD distributions in the X and Y directions, respectively, at the T7 +5 °C process condition. The maps show a smooth transition from a yellow convex region, through a flat green region, to a blue concave region. The convex and concave domains cover the majority of wafer surface, indicating pronounced and asymmetric shape changes in both X and Y directions induced by the process step. Under this condition, the residual PWG-IPD values are highest, measuring 150.54 nm/mm in the X direction and 152.43 nm/mm in the Y direction. Consequently, as shown in Fig. 4(c), the residual overlay measurements reach 149.66 nm in the X direction and 150.62 nm in the Y direction. The largest residual overlay vectors are observed at the left and bottom regions of the wafer, corresponding to the yellow convex domains in the PWG-IPD maps. Such distortion induces overlay misalignment beyond device specifications, ultimately degrading device yield. When the T7 offset is reduced to +0 °C, the PWG-IPD maps in Fig. 4(d) and (e) show a smaller height difference between the yellow convex and blue concave regions. Additionally, finer surface topography features become more apparent, indicating that process-induced distortion is reduced both in affected area and magnitude. Under the previous condition, these minor local features were obscured by the dominant global distortion. When the T7 offset is further lowered to −5 °C, as shown in Fig. 4(g) and (h), the yellow and blue regions are barely observable, and the 3D PWG-IPD maps appear predominantly green. This suggests the absence of process-induced wafer distortion. Small peaks near the wafer center are observed, which are likely normal topology measurement fluctuations rather than systematic shape changes. Under this condition, the PWG-IPD values are reduced to 8.48 nm/mm in the X direction and 9.34 nm/mm in the Y direction. Consistently, the residual overlay vectors in Fig. 4(i) are nearly eliminated, reaching only 16.37 nm in the X direction and 14.24 nm in the Y direction. Under this condition, process-induced shape distortion is effectively suppressed, thereby mitigating downstream overlay issues.

The analysis then proceeds with a quantitative assessment. In a single measurement, the PWG tool can collect more than four million high-spatial-resolution surface topography data points [49]. Based on these raw data, the PWG system partitions the wafers into multiple sectors of 26 mm × 33 mm. The residual IPD for each sector, defined as the change in slope of the wafer shape, is extracted and presented in Fig. 5(a) and (b) for the T7 +5 °C, T7 +0 °C, and T7 −5 °C conditions. Consistent with previous qualitative observations, process-induced distortion is highest under the T7 +5 °C condition. Sector-level PWG-IPD exhibits mean values ranging from 25.6 to 37.5 nm/mm, with corresponding sigma values of 22.2 to 35.1 nm/mm. As the wafer edge temperature is reduced to +0 °C and −5 °C, both mean and sigma values of sector PWG-IPD decrease significantly. At T7 −5 °C, the mean values range from 0.68 to 1.13 nm/mm, and the sigma values range from 0.61 to 1.59 nm/mm, approximately 3% of the values observed under the worst condition. These results confirm that PWG measurements are sensitive enough to distinguish process-induced wafer distortion, making it feasible to establish robust statistical process control (SPC) charts to monitor the thermal process.

TABLE I
PWG-IPD AND RESIDUAL OVERLAY DATA UNDER T7 OFFSETS

T7 offset (°C)	PWG-IPD X (nm/mm)	PWG-IPD Y (nm/mm)	R-overlay X (nm)	R-overlay Y (nm)
+5	150.54	152.43	149.66	150.62
+0	37.77	39.30	67.56	70.53
−5	8.48	9.34	16.37	14.24

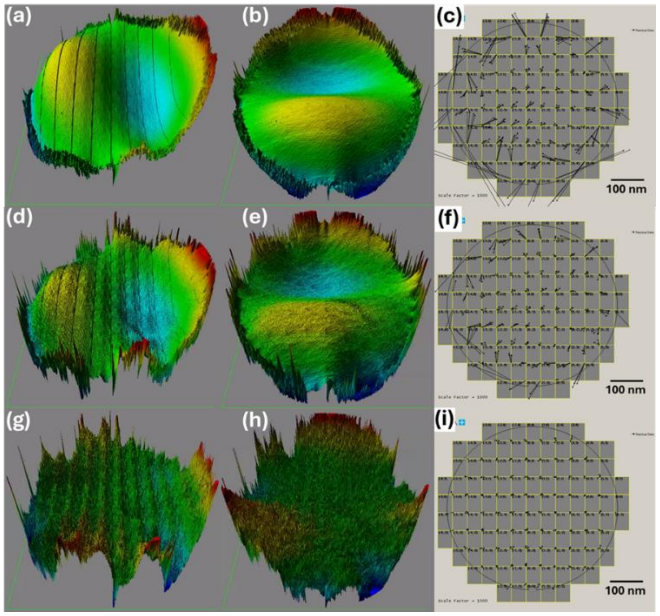

Fig. 4. Residual IPD and residual overlay maps under T7 offsets. T7 +5 °C: (a–b) Strong edge distortion in IPD results in (c) the highest overlay error. T7 +0 °C: (d–e) Reduced IPD distortion yields (f) improved overlay. T7 –5 °C: (g–h) Minimal IPD distortion provides (i) the best overlay performance.

The corresponding residual overlay data was also collected. As shown in Fig. 4(c), each wafer is partitioned into 104 sections based on the reticle size. Residual overlay was measured at the four corners of each section, excluding data from the extreme wafer edge. A total of 300 site-level residual overlay points were collected per wafer. The data are presented in Fig. 5(c) and (d) for the X and Y directions, respectively. Consistent with the PWG-IPD results, the residual overlay sigma values are highest under the T7 +5 °C condition, ranging from 49.9 to 50.5 nm. Under the T7 −5 °C condition, the wafers exhibit much lower sigma values, ranging from 4.75 to 7.42 nm. As expected, the average overlay value for each wafer is approximately 0 nm, as the lithography tool compensates for the global linear components. Despite this, the distribution of residual overlay clearly reflects the same trend observed in PWG-IPD, indicating a correlation between these two measurement methods.

This article has been accepted for publication in IEEE Transactions on Semiconductor Manufacturing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSM.2026.3710173

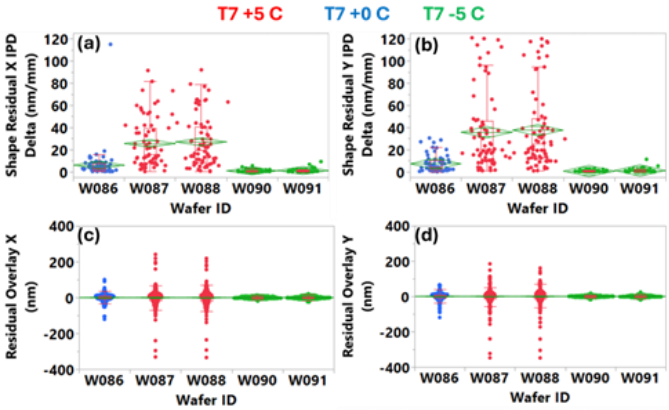

Fig. 5. Residual IPD and overlay distributions under T7 offsets. (a) IPD-X, (b) IPD-Y, (c) overlay-X, (d) overlay-Y.

A strong correlation is observed between the PWG-IPD 3σ values and the measured residual overlay 3σ values in both the X and Y directions. As shown in Fig. 6, PWG-IPD demonstrates clear predictive capability for residual overlay across two different devices, with R^2 values ranging from 0.869 to 0.979. This indicates that the PWG-IPD can effectively detect process-induced wafer topology distortions that lead to lithographic patterning errors, consistent with prior studies [3, 10, 16, 17, 21, 50]. The correlation slopes for Device B (8.45 in X, 9.35 in Y) are an order of magnitude higher than those for Device A (0.871 in X, 0.891 in Y), suggesting that overlay errors in Device B are predominantly driven by process-induced wafer distortion.

While PWG-IPD can reliably predict downstream residual overlay performance, the practical motivation is to evaluate its advantages over conventional metrology. Within-wafer

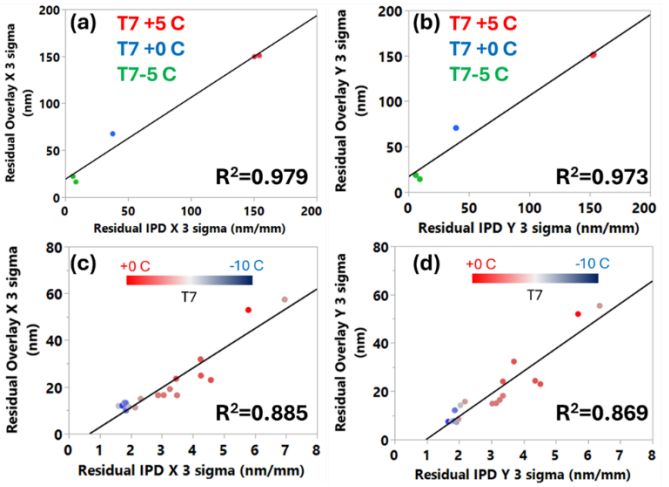

Fig. 6. Linear correlation between wafer-averaged residual overlay and residual IPD in the (a) X and (b) Y directions for Device A, and in the (c) X and (d) Y directions for Device B.

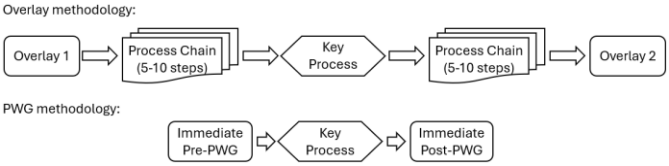

Fig. 7. Comparison of overlay-based and PWG-based methods for detecting process-induced wafer distortion.

temperature variation can be quantified using several approaches, including direct pyrometer measurements and indirect methods based on thermal oxidation rates or implant activation analysis [24, 32, 40, 51, 52]. In these indirect approaches, temperature distribution is inferred from the measured oxide film thickness or sheet resistance distribution. However, pyrometer readings may drift due to degradation or coating accumulation over a preventive maintenance cycle. Oxide growth is sensitive to chamber gas flow dynamics, while sheet resistance is influenced by upstream implant tool-to-tool variations. These coupling effects introduce uncertainties and limit the accuracy of existing methods. To date, overlay measurement remains the most reliable method to detect within-wafer shape distortion. However, as shown in Fig. 7, overlay-based detection requires at least two lithography exposures separated by multiple process steps. This makes the overlay method both costly and time-consuming and exposes more wafers to potential process risk. In contrast, the PWG method enables immediate pre- and post-process monitoring of the key process step, offering comparable accuracy to overlay while providing faster and more efficient feedback.

As demonstrated earlier, T7 tuning directly modulates the within-wafer temperature distribution, consequently affecting both PWG-IPD and overlay performance. Temperature nonuniformity in RTP is widely recognized as the root cause of wafer distortion, and numerous studies have focused on mitigating this limitation to improve overlay performance [34, 38, 39, 47]. Local thermal gradients, particularly in the wafer edge region, generate thermal stress. If this stress exceeds the material yield threshold, plastic deformation occurs, the IPD value increases, and the overlay performance is compromised [1, 36, 53-55]. Maintaining a uniform temperature profile remains critical to mitigating process-induced overlay errors. Classical RTP models describe the wafer edge as intrinsically cooler due to radiative or convective heat loss [24, 33, 40, 42]. In contrast, our results reveal the opposite phenomenon. As shown in Fig. 6, reducing the edge temperature improves both IPD and overlay for both devices, suggesting that under baseline process conditions, the wafer edge is overheated relative to the center. Therefore, negative T7 tuning restores thermal symmetry, consistent with prior studies [47, 56].

Two mechanisms may explain this elevated edge temperature. First, intentional edge heating is often employed to compensate for lower oxidation rates at the wafer periphery. In cold-wall chamber processes, precursor gases enter from one side of the chamber and reach the wafer edge before fully activated. Thus, pyrometer or heater power tuning can produce a uniform oxide film while compromising thermal uniformity. Second, RTP platforms incorporate edge ring structures to suppress peripheral heat loss by reflecting thermal radiation back toward the wafer [24, 41, 42, 57]. Although originally intended to improve temperature uniformity, this hardware component can overcompensate, leading to localized edge overheating [58].

Even though previous studies have reported that T7 or T6 tuning can reduce IPD value and improve overlay performance, the practical application is limited. As shown in Fig. 8, lowering the T7 offset from 0 °C to −8 °C results in a substantial reduction in edge thickness. In addition, due to interaction

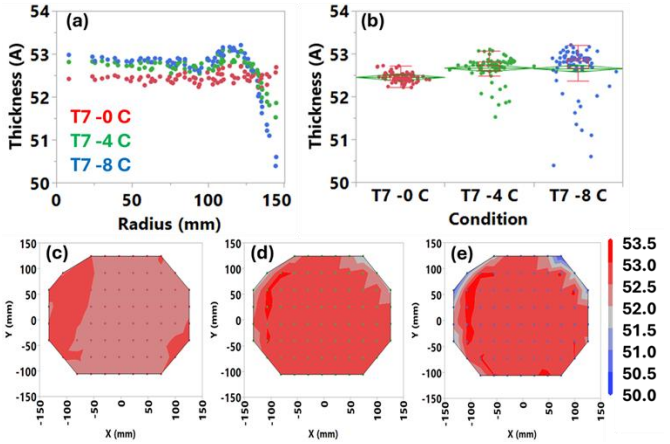

Fig. 8. (a) Radial thickness profile and (b) box plot of SiO_2 thickness under different T7 tuning conditions. Full-wafer thickness contour maps for (c) T7 = 0 °C, (d) T7 = −4 °C, and (e) T7 = −8 °C using same scale.

between neighboring temperature control zones in the multizone lamp system, it is believed that the heater power in the donut region increases in response to the reduced wafer edge temperature. As a result, higher oxide thickness was observed at this region. For critical layers such as gate oxides, this tuning severely degrades film uniformity and may compromise device performance. Consequently, current mitigation strategies rely on carefully selecting conditions that balance both overlay and film thickness uniformity within device specifications.

This dilemma is further illustrated in Fig. 9. As the T7 offset is tuned negatively, the within wafer temperature is more uniform. This results in PWG-IPD in both the X and Y directions decreases and saturates beyond −4 °C. In contrast, film uniformity, defined as the ratio of standard deviation to the mean, continues to degrade monotonically. Consequently, with the existing hardware configuration, it is impossible to simultaneously achieve optimal overlay and film thickness uniformity.

To overcome this limitation, a modified chamber configuration was evaluated. As shown in Fig. 10, a continuous improvement project (CIP) chamber retains the process of record (POR) design, with a 9 o'clock main injector and a 3 o'clock exhaust to maintain laminar gas flow. An additional hydrogen gas injector at 12 o'clock is introduced. This secondary flow is small relative to the total gas flow so that the injected gas is confined to the wafer edge region before merging with the main flow. However, the additional H_2 flow is large compared to the original hydrogen in the chamber. This provides an additional tuning knob to modify the local H_2:O_2 ratio in wafer edge region to enhance oxidation rate, recover oxide thickness uniformity while maintaining temperature uniformity and good overlay performance.

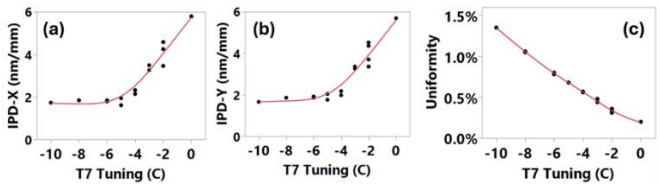

Fig. 9. Variation of (a) IPD-X, (b) IPD-Y, and (c) thickness uniformity with respect to T7 tuning.

TABLE II
OVERLAY AND THICKNESS COMPARISON OF POR AND CIP CHAMBERS AT THE SAME TEMPERATURE SETTING

Model	Residual overlay X (nm)	Residual overlay Y (nm)	Thickness average (Å)	Standard deviation (Å)	Range (Å)	Uniformity (%)
POR	22.27	18.63	42.97	0.57	2.34	1.33
CIP	20.07	16.88	43.35	0.08	0.40	0.18

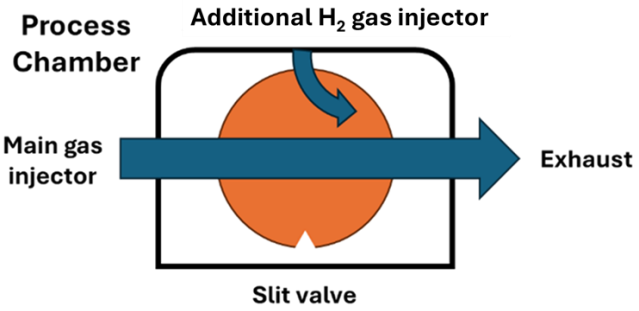

Fig. 10. Schematic illustration of the CIP chamber equipped with an additional gas injector.

Fig. 11 and Table II compare the film thickness results from the original POR chamber and the new CIP chamber under identical temperature settings. The pyrometer was fine-tuned to achieve optimal PWG-IPD and overlay performance. As both chambers operate at the same temperature settings, the overlay results are comparable for POR (22.27 nm in X, 18.63 nm in Y) and CIP (20.07 nm in X, 16.88 nm in Y). However, notable differences are observed in film thickness behavior. In the POR chamber, T7 tuning leads to a pronounced edge thickness loss, with a thickness range of 2.34 Å and a standard deviation of 0.57 Å. In contrast, the CIP chamber exhibits high thickness uniformity across the wafer, with a thickness range of 0.40 Å and a standard deviation of 0.08 Å. The wafer thickness uniformity of 0.18% achieved on the CIP tool is only 14% of the 1.33% obtained on the POR tool. In addition, full loop wafers were fabricated and sent for final electrical testing. Inversion thickness (Tinv) of the gate oxide for both NMOS and PMOS devices is one of the most representative parameters. As shown in Fig. 12, the Tinv data shows that the CIP condition improves the within wafer electrical oxide thickness uniformity for both NMOS and PMOS devices, especially at the wafer edge region. This shows good correlation with the improved inline physical silicon oxide film thickness uniformity measured by ellipsometer. No degradation was observed in the measured electrical test parameters.

These results demonstrate that the additional side hydrogen gas injector decouples oxidation rate compensation from temperature tuning at wafer edge, enabling the RTO process to simultaneously optimize both overlay and oxide film thickness uniformity.

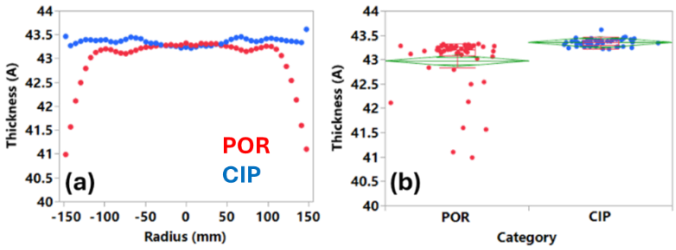

Fig. 11. (a) Linear and (b) box plot comparison of film thickness between POR and CIP chambers.

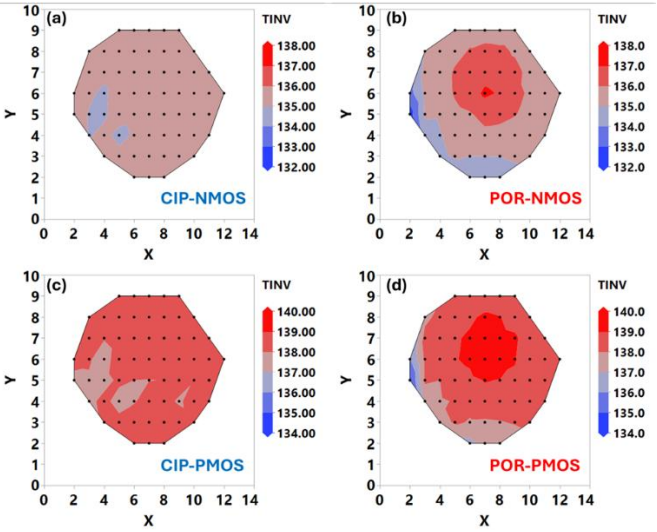

Fig. 12. Comparison of Tinv contour map between NMOS device from (a) CIP chamber and (b) POR chamber, and PMOS device from (c) CIP chamber and (d) POR chamber. In both cases, devices from CIP chambers show a better within wafer uniformity.

IV. Conclusion

This study established a robust PWG-IPD methodology for the accurate detection of process-induced distortion and reliable prediction of downstream overlay performance. By optimizing the T7 offset and incorporating a side gas injector, the RTP process achieved simultaneous improvement in both residual overlay and film thickness uniformity. The demonstrated approach has the potential to enhance RTP performance and support continued device scaling.

References

[1] W. P. de Boeij *et al*., "Extending immersion lithography down to 1× nm production nodes," in *Proc. SPIE*, vol. 8683, 2013, pp. 86931L-1-86931L-9.

[2] T. Brunner *et al*., "Characterization and mitigation of overlay error on silicon wafers with nonuniform stress," in *Proc. SPIE*, vol. 9052, 2014, Art. no. 90520U, doi: 10.1117/12.2045715.

[3] H. Lee *et al*., "Improvement of process control using wafer geometry for enhanced manufacturability of advanced semiconductor devices," in *Proc. SPIE*, vol. 9424, 2015, Art. no. 94240M, doi: 10.1117/12.2085862.

[4] V. R. Nagaswami, "In-plane distortion in silicon wafers induced by a submicron CMOS process," *Microelectron. Eng.*, vol. 9, no. 1, pp. 457–461, 1989, doi: 10.1016/0167-9317(89)90100-7.

[5] K. T. Turner, S. Veeraraghavan, and J. K. Sinha, "Predicting distortions and overlay errors due to wafer deformation during chucking on lithography scanners," *J. Micro/Nanolithogr., MEMS, MOEMS*, vol. 8, no. 4, pp. 043015–043018, 2009, doi: 10.1117/1.3247857.

[6] S. Tran *et al*., "Process induced wafer geometry impact on center and edge lithography performance for sub-2X nm nodes," in *Proc. 26th Annu. SEMI Adv. Semicond. Manufact. Conf. (ASMC)*, 2015, pp. 345–350, doi: 10.1109/ASMC.2015.7164417.

[7] J. Mulkens, P. Hinnen, M. Kubis, A. Padiy, and J. Benschop, "Holistic optimization architecture enabling sub-14-nm projection lithography," *J. Micro/Nanolithogr., MEMS, MOEMS*, vol. 13, no. 1, pp. 011006–1–011006–10, 2014.

[8] B. Minghetti *et al*., "Overlay characterization and matching of immersion photoclusters," in *Proc. SPIE*, vol. 7640, 2010, pp. 76400W-1-76400W-11.

[9] B. J. Lin, "Optical lithography with and without NGL for single-digit nanometer nodes," in *Proc. SPIE*, vol. 9426, 2015, Art. no. 942602.

[10] J. Peterson *et al*., "Lithography overlay control improvement using patterned wafer geometry for sub-22 nm technology nodes," in *Proc. SPIE*, vol. 9424, 2015, Art. no. 94240N, doi: 10.1117/12.2086525.

[11] 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*, vol. 11, no. 1, pp. 013001–1–013001–10, 2012, doi: 10.1117/1.JMM.11.1.013001.

[12] H. Siethoff, P. Haasen, and R. Hasiguti, *Lattice Defects in Semiconductors*. Tokyo, Japan: Univ. Tokyo, 1968, p. 491.

[13] H. Lee *et al*., "Patterned wafer geometry grouping for improved overlay control," in *Proc. SPIE*, vol. 10145, 2017, Art. no. 101450O, doi: 10.1117/12.2257834.

[14] K.-H. Chen *et al*., "Litho process control via optimum metrology sampling while providing cycle time reduction," in *Proc. SPIE*, vol. 7971, 2011, pp. 41–48.

[15] J. M. Trujillo-Sevilla *et al*., "New metrology technique for measuring patterned wafer geometry on a full 300-mm wafer," in *Proc. SPIE*, vol. 12053, 2022, Art. no. 1205303, doi: 10.1117/12.2614320.

[16] T. A. Brunner *et al*., "Patterned wafer geometry (PWG) metrology for improving process-induced overlay and focus problems," in *Proc. SPIE*, vol. 9780, 2016, Art. no. 97800W, doi: 10.1117/12.2218630.

[17] K. T. Turner, P. Vukkadala, and J. K. Sinha, "Models to relate wafer geometry measurements to in-plane distortion of wafers," *J. Micro/Nanolithogr., MEMS, MOEMS*, vol. 15, no. 2, pp. 021404–1–021404–10, 2016, doi: 10.1117/1.JMM.15.2.021404.

[18] J. M. Trujillo-Sevilla *et al*., "New wave front phase sensor used for 3D shape measurements of patterned silicon wafers," in *Proc. SPIE*, vol. 12292, 2022, Art. no. 122920P, doi: 10.1117/12.2642287.

[19] J. F. Valley *et al*., "Approaching new metrics for wafer flatness: an investigation of the lithographic consequences of wafer non-flatness," in *Proc. SPIE*, vol. 5375, 2004, pp. 1098–1108.

[20] R. Brinkhof *et al*., "Wafer geometry and lithography focus performance for advanced nodes—preliminary results," *Int. Adv. Wafer Geometry Task Force*, 2011.

[21] D. M. Owen, "Stress inspection for overlay characterization," in *Proc. SPIE*, vol. 8681, 2013, Art. no. 86812T, doi: 10.1117/12.2025310.

[22] J. Gong, P. Vukkadala, J. K. Sinha, and K. T. Turner, "Determining local residual stresses from high resolution wafer geometry measurements," *J. Vac. Sci. Technol. B*, vol. 31, no. 5, pp. 051205-1-051205-8, 2013.

[23] J. P. Hebb and K. F. Jensen, "The effect of patterns on thermal stress during rapid thermal processing of silicon wafers," *IEEE Trans. Semicond. Manuf.*, vol. 11, no. 1, pp. 99–107, 1998, doi: 10.1109/66.661289.

[24] H. A. Lord, "Thermal and stress analysis of semiconductor wafers in a rapid thermal processing oven," *IEEE Trans. Semicond. Manuf.*, vol. 1, no. 3, pp. 105–114, 1988, doi: 10.1109/66.4383.

[25] C. D. Schaper, M. M. Moslehi, K. C. Saraswat, and T. Kailath, "Modeling, identification, and control of rapid thermal processing systems," *J. Electrochem. Soc.*, vol. 141, no. 11, p. 3200, 1994.

[26] J. Liu and Y.-S. Chen, "Simulation of rapid thermal processing in a distributed computing environment," *Numer. Heat Transfer A*, vol. 38, no. 2, pp. 129–152, 2000.

[27] S. Lin and H.-S. Chu, "Thermal uniformity of 12-in silicon wafer during rapid thermal processing by inverse heat transfer method," *IEEE Trans. Semicond. Manuf.*, vol. 13, no. 4, pp. 448–456, 2000.

[28] Z. Xiang *et al*., "Multi-step mechanical and thermal homogenization for the warpage estimation of silicon wafers," *Micromachines*, vol. 15, no. 3, p. 408, 2024, doi: 10.3390/mi15030408.

[29] F. Che, H. Y. Li, X. Zhang, S. Gao, and K. H. Teo, "Development of wafer-level warpage and stress modeling methodology and its application in process optimization for TSV wafers," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 2, no. 6, pp. 944–955, 2012.

[30] A. Wright *et al*., "Simulating wafer bow for integrated capacitors using a multiscale approach," in *Proc. 17th Int. Conf. Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*, 2016, pp. 1–7.

[31] K. N. Ritz, M. Delfino, C. B. Cooper, and R. A. Powell, "Observation of slip dislocation in (100) silicon wafers after BF_2 ion-implantation and rapid thermal annealing," *J. Appl. Phys.*, vol. 60, no. 2, pp. 800–802, 1986, doi: 10.1063/1.337377.

[32] M. M. Moslehi, "Process uniformity and slip dislocation patterns in linearly ramped-temperature transient rapid thermal processing of silicon," *IEEE Trans. Semicond. Manuf.*, vol. 2, no. 4, pp. 130–140, 1989, doi: 10.1109/66.44616.

[33] C. Ching-Kong, H. Shih-Yu, and Y. Cheng-Ching, "Thermal stress analysis for rapid thermal processor," *IEEE Trans. Semicond. Manuf.*, vol. 16, no. 2, pp. 335–341, 2003, doi: 10.1109/TSM.2003.811884.

[34] J. F. Buller, M. M. Farahani, and S. Garg, "Manufacturing issues related to RTP-induced overlay errors in a global alignment stepper technology," *IEEE Trans. Semicond. Manuf.*, vol. 9, no. 1, pp. 108–114, 1996, doi: 10.1109/66.484290.

[35] B. Feil, M. Drew, and J. Moench, "Patterned-induced pattern misregistration after BPSG RTA reflow," in *Proc. 1st Int. Rapid Thermal Process. Conf.*, 1993, pp. 8–10.

[36] G. G. Bentini, L. Correra, and C. Donolato, "Defects introduced in silicon wafers during rapid isothermal annealing: thermoelastic and thermoplastic effects," *J. Appl. Phys.*, vol. 56, no. 10, pp. 2922–2929, 1984, doi: 10.1063/1.333832.

[37] W. Tong, R. Kim, B. Krishnan, S. Kim, O. Vatel, X. Liu, *et al*., "Experimental investigation and manufacturing solution of the rapid thermal process induced overlay residue," in *Proc. SEMI Adv. Semicond. Manuf. Conf. (ASMC)*, 2012, pp. 300–304, doi: 10.1109/ASMC.2012.6212915.

[38] A. Stephens, "Plastic deformation of the silicon wafer," in *Proc. 1st Int. Rapid Thermal Process. Conf.*, 1993, pp. 94–101.

[39] R. Deaton and H. Massoud, "Effect of thermally induced stresses on the rapid-thermal oxidation of silicon," *J. Appl. Phys.*, vol. 70, no. 7, pp. 3588–3592, 1991.

[40] J. Blake, J. C. Gelpey, J. F. Moquin, J. Schlueter, and R. Capodilupo, "Slip free rapid thermal processing," *MRS Proc.*, vol. 92, 1987, doi: 10.1557/PROC-92-265.

[41] S. Hirasawa, T. Suzuki, and S. Maruyama, "Optimal heating control conditions to unify temperature distribution in a wafer during rapid thermal processing with lamp heaters," in *Proc. 9th Int. Conf. Adv. Thermal Process. Semicond. (RTP)*, 2001, pp. 210–216, doi: 10.1109/RTP.2001.1013768.

[42] R. T. Blunt, M. S. M. Lamb, and R. Szweda, "Crystallographic slip in GaAs wafers annealed using incoherent radiation," *Appl. Phys. Lett.*, vol. 47, no. 3, pp. 304–306, 1985, doi: 10.1063/1.96200.

[43] M. Robinson *et al*., "Low dislocation density RF-heated epitaxial silicon," *J. Electrochem. Soc.*, vol. 129, no. 12, p. 2858, 1982.

[44] P. Vandenabeele and K. Maex, "Temperature non-uniformities during rapid thermal processing of patterned wafers," in *Proc. SPIE*, vol. 1189, 1990, pp. 89–105.

[45] N. Akiyama, Y. Inoue, and T. Suzuki, "Critical radial temperature gradient inducing slip dislocations in silicon epitaxy using dual heating of the two surfaces of a wafer," *Jpn. J. Appl. Phys.*, vol. 25, no. 11, pp. 1619–1622, 1986, doi: 10.1143/JJAP.25.1619.

[46] J. P. Hebb and K. F. Jensen, "The effect of multilayer patterns on temperature uniformity during rapid thermal processing," *J. Electrochem. Soc.*, vol. 143, no. 3, p. 1142, 1996.

[47] J. Lv, Q. Wang, Y. Lu, X. Gao, and Q. Sun, "Impacts of RTP pyrometer offsets on wafer overlay residue," in *Proc. China Semicond. Technol. Int. Conf. (CSTIC)*, 2020, pp. 1–3, doi: 10.1109/CSTIC49141.2020.9282587.

[48] W. Renken, "Process diagnostics using wafer temperature mapping," in *Proc. 1st Int. Rapid Thermal Process. Conf.*, 1993, pp. 262–266.

[49] K. Freischlad, S. Tang, and J. Grenfell, "Interferometry for wafer dimensional metrology," in *Proc. SPIE*, vol. 6672, 2007, pp. 667292-1–667202-14.

[50] T. A. Brunner *et al*., "Characterization of wafer geometry and overlay error on silicon wafers with nonuniform stress," *J. Micro/Nanolithogr. MEMS, MOEMS*, vol. 12, no. 4, pp. 043002–1–043002–10, 2013, doi: 10.1117/1.JMM.12.4.043002.

[51] J. C. Gelpey, P. O. Stump, J. Blake, A. Michel, and W. Rausch, "Uniformity characterization of an RTP system," *Nucl. Instrum. Methods Phys. Res. B*, vol. 21, nos. 1–4, pp. 612–617, 1987.

[52] T. Sedgwick, A. E. Michel, S. Cohen, V. Deline, and G. Oehrlein, "Investigation of transient diffusion effects in rapid thermally processed ion implanted arsenic in silicon," *Appl. Phys. Lett.*, vol. 47, no. 8, pp. 848–850, 1985.

[53] A. E. Widmer and W. Rehwald, "Thermoplastic deformation of silicon wafers," *J. Electrochem. Soc.*, vol. 133, no. 11, pp. 2403–2409, 1986, doi: 10.1149/1.2108417.

[54] K. Morizane and P. S. Gleim, "Thermal stress and plastic deformation of thin silicon slices," *J. Appl. Phys.*, vol. 40, no. 10, pp. 4104–4107, 1969.

[55] S. Hu, "Dislocation pinning effect of oxygen atoms in silicon," *Appl. Phys. Lett.*, vol. 31, no. 2, pp. 53–55, 1977.

[56] J. F. Shepard, W. A. Muth, and S. MacNish, "Experimental investigation of the rapid thermal process slip window," in *Proc. 18th Int. Conf. Adv. Thermal Process. Semicond. (RTP)*, 2010, pp. 28–31, doi: 10.1109/RTP.2010.5623803.

[57] G. L. Young and K. A. McDonald, "Effect of radiation shield angle on temperature and stress profiles during rapid thermal annealing," *IEEE Trans. Semicond. Manuf.*, vol. 3, no. 4, pp. 176–182, 1990.

[58] S. A. Campbell and K. L. Knutson, "Transient effects in rapid thermal processing," *IEEE Trans. Semicond. Manuf.*, vol. 5, no. 4, pp. 302–307, 1992.