Wafer alignment mark placement accuracy impact on the layer-to-layer overlay performance

Richard van Haren, Steffen Steinert, Orion Mouraille, Koen D'havé, Leon van Dijk, et al.

Richard van Haren, Steffen Steinert, Orion Mouraille, Koen D'havé, Leon van Dijk, Jan Hermans, Dirk Beyer, "Wafer alignment mark placement accuracy impact on the layer-to-layer overlay performance," Proc. SPIE 11148, Photomask Technology 2019, 1114811 (26 September 2019); doi: 10.1117/12.2536270

SPIE.

Event: SPIE Photomask Technology + EUV Lithography, 2019, Monterey, California, United States

Wafer alignment mark placement accuracy impact on the layer-to-layer overlay performance

Richard van Haren[a], Steffen Steinert[b], Orion Mouraille[a],
Koen D'havé[c], Leon van Dijk[a], Jan Hermans[c], Dirk Beyer[b]
[a]ASML, Flight Forum 1900 (no. 5846), 5657 EZ Eindhoven, The Netherlands
[b]Carl Zeiss SMT GmbH, Carl-Zeiss-Promenade 10, 07745 Jena, Germany
[c]IMEC, Kapeldreef 75, B-30001, Leuven, Belgium

ABSTRACT

It has been demonstrated that the mask-to-mask overlay contribution can be fully characterized by off-line measurements on the PROVE® mask registration tool. This characterization includes the impact of the marks that are used for reticle alignment inside the scanner. This is an important aspect since the scanner is blind to the features inside the image field and intra-field adjustments are only based on measurements of the reticle alignment marks. The off-line determined mask-to-mask overlay was compared with the measured on-wafer results and a perfect correlation ($R^2 > 0.96$) was found. The residual mismatch was around 0.6-nm, which is 30% of the dedicated chuck overlay performance of the scanner that was used. These results enable feed-forward corrections to the scanner to improve the intra-field overlay performance or to predict the intra-field overlay originating from mask writing errors (computational overlay).

We recently extended the work to the layer-to-layer overlay impact by considering the mask writing error of a wafer alignment mark. This wafer alignment mark was exposed in the first layer. Apart from the reticle writing error of the wafer alignment mark itself, the reticle alignment contribution performed on dedicated reticle alignment marks inside the scanner plays an important role as well. The actual position of the selected wafer alignment mark is also impacted by the reticle alignment model corrections at that specific field location. Only when both contributors are considered, the layer-to-layer overlay can be predicted accurately. In this scenario, the layer-to-layer overlay is measured back to the layer in which the alignment marks were defined. This is referred to as the direct alignment use-case.

In this paper, we further investigate the direct alignment use-case in relation to the layer-to-layer overlay. Apart from the reticle writing error and the reticle alignment corrections, the actual placement of the wafer alignment mark during exposure can also be affected by other applied corrections. We will present experimental results of the layer-to-layer overlay as function of the applied automated process corrections on the wafer alignment mark location printed in the first layer. It is shown that the wafer alignment sensor impact should be considered as well in the interpretation of the results. We finally present a strategy to control these kinds of overlay errors.

Keywords: Registration Error, Overlay, Computational Overlay, Reticle, Mask, Direct wafer alignment, Wafer alignment mark, APC, Feed-Forward

1. INTRODUCTION

Different alignment scenarios exist before exposing a wafer inside the scanner. For the first layer that is exposed, a reticle is loaded on the reticle stage, the wafer is placed on the wafer table and brought to the expose side of the scanner. A reticle alignment (RA) takes place based on a wafer stage fiducial to link the reticle position to the wafer stage position. ASML has developed two types of reticle alignment methods. The oldest one is the Transmission Image Sensor (TIS) reticle alignment. This reticle alignment is based on 4 reticle alignment marks that are placed outside the image field. A 6-parameter linear model is used to correct for potential offsets. More recently, a Parallel ILIAS (PARIS) reticle alignment sensor can be used as well. This type of reticle alignment enables a 12-parameter reticle correction. In addition to the 6 linear parameters, 6 higher order polynomial correction terms can be addressed too. The PARIS sensor has been designed to enable reticle alignment as well as capturing thermal effects like lens heating and reticle heating [1]. The first layer exposure field typically contains a wafer alignment (WA) mark to enable wafer alignment inside the scanner for the

Photomask Technology 2019, edited by Jed H. Rankin, Moshe E. Preil, Proc. of SPIE Vol. 11148, 1114811 · © 2019 SPIE · CCC code: 0277-786X/19/$21 · doi: 10.1117/12.2536270

subsequent layers of the device to be made. Ideally, the alignment mark is placed exactly on its designed position. However, the actual alignment mark position can deviate from its designed position for several reasons. One of these is the reticle itself. During the e-beam writing process, image placement errors can be made resulting in positional offsets with respect to the ideal straight grid for all features defined on the mask. These include the reticle alignment marks outside the image field but also the wafer alignment marks that are part of the image field. This is illustrated in Figure 1(a). In this example, we assume that the reticle writing error contribution ($-\Delta x$, 0) on the (PARIS) reticle alignment marks is different from the wafer alignment mark (0, $-\Delta y$). Suppose that the RA marks have a common translation offset in the negative x-direction with respect to the image field as shown in Figure 1(a). This implies that the image field is exposed with a translation offset (Δx, 0) with respect to the grid (in red) that is defined after reticle alignment, as illustrated in Figure 1(b). Although all first layer image fields are exposed with a translation offset (Δx, 0), it is not detected by the wafer alignment sensor. The reason is that the wafer alignment grid is defined based on the design coordinates of the wafer alignment mark. In this example, the wafer alignment mark location is defined at (0,0). This means that the origin ⌀ of the coordinate system after wafer alignment coincides with the actual wafer alignment mark location irrespective of the translation offset introduced by the reticle alignment. This is shown in Figure 1(c). What does matter is the reticle writing error on the wafer alignment mark in the negative y-direction. Since the origin is defined at the actual position of the wafer alignment mark, the image field center will now be at (0, Δy) with respect to the wafer alignment grid. The wafer alignment grid will be used as reference grid for the second layer exposure represented by the green color in Figure 1(d). In case we assume the same reticle writing error contribution on the reticle alignment marks, the second layer will be exposed with an offset of (Δx, 0) with respect to the wafer alignment coordinate system. The resulting overlay between Layer 1 and 2 will be a translation penalty of (Δx, $-\Delta y$) as represented by the orange arrows in Figure 1(d).

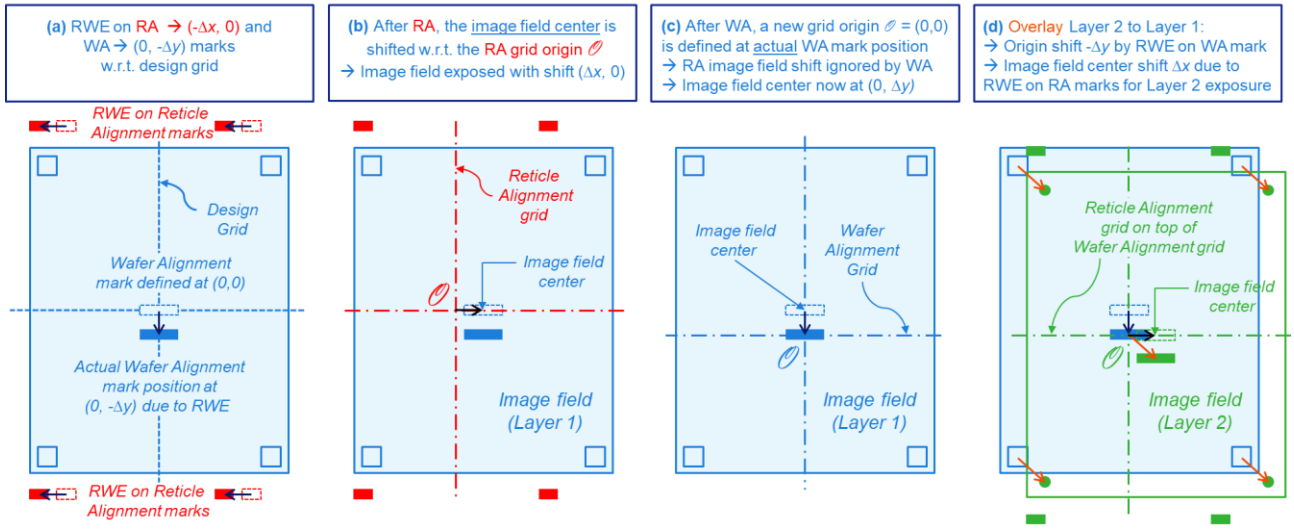

Figure 1: (a) The reticle writing error (RWE) with respect to the design grid on the reticle alignment marks ($-\Delta x$, 0) and the wafer alignment mark (0, $-\Delta y$). (b) Due to the common RWE offset on the reticle alignment marks, the image field center will be at (Δx, 0) with respect to the reticle alignment grid origin. (c) The wafer alignment grid origin (0,0) is defined at the actual wafer alignment mark position. Consequently, the image field center will be at (0, Δy). (d) The second layer image field (in green) is exposed with an offset of (Δx, 0) with respect to the wafer alignment grid origin. This is due to the RWE impact on the reticle alignment marks (like shown in (b)). The resulting layer-to-layer overlay (Δx, $-\Delta y$) is shown by the orange arrows.

All the displacement and resulting overlay errors described above can be predicted very well based on off-line mask registration measurements in combination with the appropriate reticle alignment model (TIS or PARIS) as we have demonstrated in an earlier publication [2].

In the present work, we consider the reticle writing error contribution impact on the intra-field overlay performance as the baseline of the experiment. The baseline overlay performance is represented by 4 reference wafers. The goal of this paper is to extend the previous work by considering the overlay impact as a result of applying additional intra-field exposure corrections to the layer in which the wafer alignment marks are defined. In order to isolate the impact of the

applied intra-field exposure corrections, the baseline overlay performance of the reference wafers is removed from the wafers that received additional exposure corrections. The role of the scanner wafer alignment sensor is considered as well. We will show that the resulting overlay penalties can easily be predicted and corrected by applying appropriate intra-field exposure corrections during the second layer exposure.

We only consider the use-case for which the overlay is controlled to the layer in which the alignment marks are defined. This is referred to as the direct alignment scheme. The reason is that the wafer alignment mark positions in Layer 1 are directly impacted by the applied intra-field exposure corrections. The applied intra-field exposure corrections for Layer 1 are systematically varied and the impact on the layer-to-layer overlay is considered. We also show the impact of the Layer 2 intra-field exposure corrections on overlay.

Since the scanner intra-field correction capability keeps on growing over the years, we would like to create more insight and awareness on what happens if the applied corrections are incorrect, instable, or noisy. Most of the time, the corrections are determined by the Manufacturing Execution System and based on overlay measurements of previously exposed lots. The actual applied corrections might be different from the optimal corrections for a given lot.

2. EXPERIMENTAL DETAILS

2.1 Mask details

In order to isolate the impact of the applied intra-field exposure corrections, we make use of one of the masks that was used in the previous studies. This mask was manufactured on an (older generation) e-beam writing tool, the EBM-5000 [3] and is not equipped with a pellicle.

The mask used is a so-called BMMO (Baseliner Matched Machine Overlay) mask and its layout is shown in Figure 2. The mask contains metrology modules in a (13 × 19) layout within the image field. One module is shown in an expanded view. It contains a wafer alignment mark (XPA) that can be read out by the alignment sensor inside the scanner. This makes this mask unique in the sense that all the (13 × 19) positions can be determined with respect to the scanner reference grid.

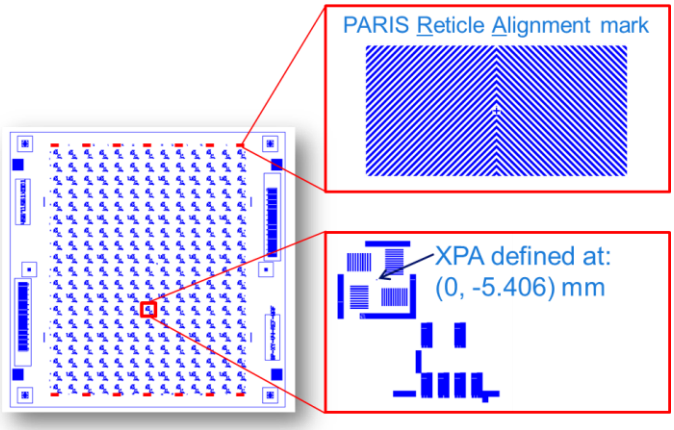

Figure 2: Layout of the mask that was used in this experiment. The image field contains metrology modules in a 13×19 layout. Each module contains an XPA wafer alignment mark that can be read by the scanner alignment sensor. PARIS reticle alignment marks (2×7) at both ends of the mask were used to align the reticle.

The same mask can also be used to expose the second layer and extract overlay information. This is done by exposing the mask with a shift of +520-µm in the *y*-direction with respect to the first layer exposure. By measuring the distance between the two XPA's and subtracting 520-µm, the overlay is obtained. By using the same mask for both Layer 1 and 2, the RWE contribution for the each individual XPA drops out. This enables us to isolate and quantify the impact of the

applied intra-field exposure corrections on the measured overlay in a very clean way. Throughout this work, the XPA wafer alignment mark defined at field coordinates (0, -5.406)-mm is used for wafer alignment when the second layer is exposed.

2.2 Exposure sequence

All experiments in the present work were executed on an ASML NXT:1970Ci (≤ 2-nm single machine overlay, dedicated chuck, full wafer coverage). PARIS reticle alignment was used for all exposures.

We make use of two lots (*A* & *B*); each lot contains 22 wafers. All wafers are coated with photoresist and three layers are exposed sequentially using the mask described in the previous section. The resist was not developed after the zero- and first-layer exposures. Latent alignment mark images were used for wafer alignment. The following corrections are applied for the three layers exposed:

- Zero layer: 4 fields with a -520-µm shift in the *y*-direction are exposed at the top-, bottom-, left-, and right-hand side of the wafer. The XPA in the center of the mask (0,0), now with a -520-µm shift in the *y*-direction, is used for wafer alignment when exposing the first layer.
- First layer: All fields on the wafer are exposed with a 0-µm shift. Wafer alignment is performed on the 4 center XPA alignment marks that were defined in the zero layer. Intra-field exposure corrections are applied while exposing the first layer for both lots. For the first 11 wafers, translation offsets in the *y*-direction (T_y) are applied, and for the subsequent 11 wafers the field magnification in the *y*-direction (M_y) is varied systematically. The XPA alignment mark located at (0, -5.406)-mm, is used for wafer alignment when exposing the second layer. The exposure corrections that are applied in the first layer have an impact on the alignment mark position.
- Second layer: All fields on the wafer are exposed with the BMMO mask but now with a +520-µm shift. For lot *A*, no corrections are applied during the second layer exposure. For lot *B*, T_y corrections for the first 11 wafers and T_y and M_y corrections are applied for the second 11 wafers. More details will be presented in the next chapter.

After exposing the second layer for all wafers, the photoresist was developed, and the overlay was measured as described in section 2.1. In order to reduce the measurement time, a (7 × 7) layout was chosen. Wafer 6 and wafer 17 in both lots are used as reference wafers. No exposure corrections were applied for these wafers. The overlay measured on these wafers is used to correct the remaining wafers to isolate the effect of the applied exposure corrections. The overlay measured on the reference wafers was already shown and explained in reference [2].

3. RESULTS

3.1 First layer exposure corrections

In order to keep it relatively simple and comprehensive, we applied two types of exposure corrections, namely Translation *y* (T_y) and a field Magnification in *y* (M_y). While T_y affects all points inside the field equally, this is not the case for M_y. In the latter case each point inside the field with field coordinates (x_F, y_F) will be (x_F, $y_F + \Delta y_F$) after correction. The difference Δy_F is determined by $\Delta y_F = M_y \cdot y_F$. Although we only discuss the results after M_y correction, a similar story can be made for intra-field rotation and/or any other higher order intra-field polynomial correction term. In this sense, the findings presented here are more generically applicable. The exposure corrections are applied via the process correction option in lot operations. This means that each individual wafer is treated as an individual lot for which a unique set of process corrections is applied. In this section, we show the overlay results for lot *A*. For this lot, exposure corrections are only applied during the first layer exposure. All exposure corrections during the second layer exposure are set to zero. A summary of the applied corrections is presented in Table 1.

Table 1: Applied exposure corrections for lot *A*. This lot is targeted to demonstrate the impact of exposure corrections for Layer 1 only. In this layer, an alignment mark is defined at field coordinates (x_F, y_F) = (0, -5.406)-mm that experiences the applied offsets. This alignment mark is used for wafer alignment when the second layer is exposed.

			Layer 1			Layer 2	
		Wafer	T_y [nm]	M_y [ppm]	Displacement in [nm] at wafer alignment mark field location (0, -5.406) mm	T_y [nm]	M_y [ppm]
Lot *A*	Set 1	1	-10	0	-10	0	0
		2	-8	0	-8	0	0
		3	-6	0	-6	0	0
		4	-4	0	-4	0	0
		5	-2	0	-2	0	0
		6	0	0	0	0	0
		7	2	0	2	0	0
		8	4	0	4	0	0
		9	6	0	6	0	0
		10	8	0	8	0	0
		11	10	0	10	0	0
	Set 2	12	0	1.84	-9.95	0	0
		13	0	1.48	-8.00	0	0
		14	0	1.1	-5.95	0	0
		15	0	0.74	-4.00	0	0
		16	0	0.36	-1.95	0	0
		17	0	0	0.00	0	0
		18	0	-0.36	1.95	0	0
		19	0	-0.74	4.00	0	0
		20	0	-1.1	5.95	0	0
		21	0	-1.48	8.00	0	0
		22	0	-1.84	9.95	0	0

Lot *A* contains two times 11 wafers. For the first set of 11 wafers, T_y offsets ranging from -10-nm to +10-nm with steps of 2-nm were applied in Layer 1. For the second set of wafers, M_y was varied systematically. The M_y values were chosen such that at the field location (0, -5.406)-mm the offsets vary from -10-nm to +10-nm with steps of 2-nm similar to the translation offsets. At (0, -5.406)-mm the wafer alignment mark is defined that is used to expose the second layer. Wafers 6 & 17 act as reference wafers. No exposure corrections are applied for Layer 1 and Layer 2. The overlay measured on these wafers is removed from the measured overlay on the wafers that received the exposure corrections.

In Figure 3 we show the measured intra-field overlay results for the first 11 wafers in lot *A*. Examples of averaged intra-field overlay performance are shown for wafers 3 (T_y = -6-nm) and 9 (T_y = +6-nm). These intra-field overlay plots are corrected with the intra-field overlay performance of reference wafer 6. The 6-parameter modelled overlay results expressed in the maximum overlay per field per parameter are shown in the bar plot on the right-hand-side of the figure. What is immediately clear from the results is that the applied T_y offsets in Layer 1 have no impact on the measured overlay. The measured overlay remains well within the overlay specification of the scanner that was used in this experiment. The results can be understood by considering the role of the wafer alignment sensor inside the scanner. The design location (0, -5.406)-mm is assigned to the wafer alignment mark location irrespective of the applied T_y offsets in the first layer. In this respect, the applied T_y offsets resemble the wafer placement variations when the wafer is placed on the wafer table.

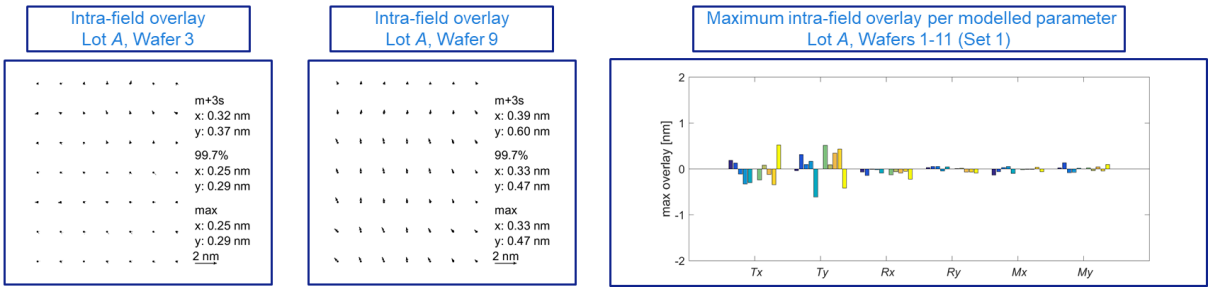

Figure 3: As an example, the intra-field overlay performance for wafers 3 & 9 is shown. These wafers received a T_y translation offset in Layer 1 of -6-nm and +6-nm, respectively. There is no impact on the measured overlay. The same is true for all the other wafers. A summary of all the 11 wafers is shown in the bar plot on the right-hand side of the figure. In order to compare the overlay impact, the maximum overlay inside the field per modelled parameter is presented. All results are obtained after overlay correction measured on reference wafer 6.

In Figure 4 we show the measured intra-field overlay results for the second set of 11 wafers in lot *A*. In analogy with what was shown in Figure 3, wafers 14 & 20 have been selected to show the intra-field overlay performance. For these wafers, M_y corrections of +1.1 ppm and -1.1 ppm were applied when Layer 1 was exposed. These intra-field overlay plots are corrected with the intra-field overlay performance that was measured on reference wafer 17.

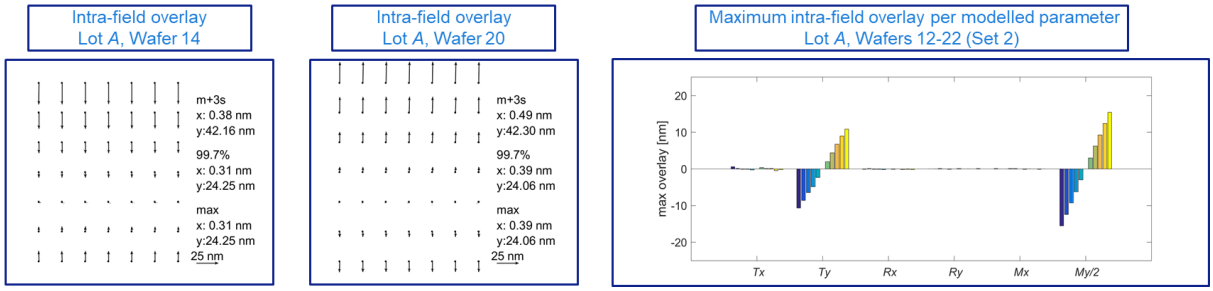

Figure 4: As an example, the intra-field overlay performance for wafers 14 & 20 is shown. These wafers received a M_y offset in Layer 1 of +1.1 ppm and -1.1 ppm, respectively. The measured intra-field overlay shows both a T_y and an M_y signature. The same is observed for all the other wafers. A summary of wafers 12 to 22 is shown in the bar plot on the right-hand side of the figure. In order to compare the overlay impact, the maximum overlay inside the field per modelled parameter is presented. A negative scaling is represented by a negative maximum overlay number. All results are obtained after correction with reference wafer 17.

Interestingly, the resulting intra-field overlay performance does not only show the applied M_y correction with a negative sign, but also contains a T_y component. The T_y component scales with the applied M_y values in Layer 1. To be more specific $T_y = M_y \cdot y_F$, where y_F = -5.406-mm. This coincides with the design location of the wafer alignment mark in Layer 1 that is used for wafer alignment before exposing Layer 2. The observed T_y component is a direct consequence of the fact that the wafer alignment mark that is used for exposing Layer 2 is fixed and defined at (x_F, y_F) = (0, -5.406)-mm. In fact, the field *y*-coordinate is not fixed for this set of wafers. It changes with $\Delta y_F = M_y \cdot y_F$ per applied M_y offset in Layer 1. In case the field coordinates of the wafer alignment mark would have been updated with Δy_F, the T_y component in the measured overlay would not have been present anymore. An alternative would be to utilize this knowledge and apply dedicated T_y and M_y corrections during Layer 2 exposure. We will demonstrate this in section 3.2.

3.2 Second layer exposure corrections

In this section, we show the overlay impact of applying exposure corrections when Layer 2 is exposed. A dedicated lot containing 2 times 11 wafers was assigned for this purpose. We refer to this lot as lot *B*. The exposure corrections that were applied when Layer 1 was exposed are identical to what was done for lot *A*. For the first 11 wafers, we just copied the applied T_y corrections from Layer 1 and apply it at Layer 2 as well. This was done to illustrate what happens if exposure corrections are naïvely copied from one litho layer to the next layer without considering the wafer alignment scheme. For the second set of 11 wafers in lot *B*, we consider the location of the wafer alignment mark in Layer 1 and how it was impacted by the Layer 1 exposure corrections. Since this wafer alignment mark is defined in Layer 1 to which the overlay is measured, the applied process corrections in Layer 1 also change the exact location where the alignment mark can be found by the scanner alignment sensor. We know that by applying an M_y exposure correction in Layer 1, the wafer alignment mark at field coordinates (x_F, y_F) = (0, -5.406)-mm, will be displaced by $M_y \cdot y_F$. We therefore not only apply the same M_y correction in Layer 2 as in Layer 1, but also introduce a T_y correction that is equal to $-M_y \cdot y_F$. A summary of all applied exposure corrections for lot *B* is presented in table 2. Wafers 6 & 17 act as reference wafers again. No exposure corrections are applied for these wafers at the time when Layer 1 and 2 were exposed.

Table 2: Applied exposure corrections for lot *B*. This lot is targeted to demonstrate the impact of exposure corrections for Layer 2. The applied exposure corrections for Layer 1 are exactly the same as was done for lot *A*.

			Layer 1			Layer 2	
		Wafer	T_y [nm]	M_y [ppm]	Displacement in [nm] at wafer alignment mark field location (0, -5.406) mm	T_y [nm]	M_y [ppm]
Lot *B*	Set 1	1	-10	0	-10	-10	0
		2	-8	0	-8	-8	0
		3	-6	0	-6	-6	0
		4	-4	0	-4	-4	0
		5	-2	0	-2	-2	0
		6	0	0	0	0	0
		7	2	0	2	2	0
		8	4	0	4	4	0
		9	6	0	6	6	0
		10	8	0	8	8	0
		11	10	0	10	10	0
	Set 2	12	0	1.84	-9.95	10	1.84
		13	0	1.48	-8.00	8	1.48
		14	0	1.1	-5.95	6	1.1
		15	0	0.74	-4.00	4	0.74
		16	0	0.36	-1.95	2	0.36
		17	0	0	0.00	0	0
		18	0	-0.36	1.95	-2	-0.36
		19	0	-0.74	4.00	-4	-0.74
		20	0	-1.1	5.95	-6	-1.1
		21	0	-1.48	8.00	-8	-1.48
		22	0	-1.84	9.95	-10	-1.84

In Figure 5 we show the measured intra-field overlay results for the first 11 wafers in lot *B*. Examples of the averaged intra-field overlay performance are shown for wafers 3 (T_y = -6-nm) and 9 (T_y = +6-nm). In order to isolate the effect of the applied exposure corrections, the intra-field overlay plots shown are corrected with the intra-field overlay performance of reference wafer 6. A clear translation T_y overlay penalty can be observed, which is a direct result from the dialed in T_y exposure correction for Layer 2. We would like to emphasize that for the direct wafer alignment use-case as presented in this work the T_y corrections for Layer 1 do not cancel out the T_y corrections for Layer 2. This is due to the fact that the alignment sensor inside the scanner is unware of the applied T_y corrections in Layer 1, as we already showed in Figure 3. In principle, the awareness can easily be created by dynamically updating the alignment mark field coordinates in the exposure job based on the exposure correction parameters. In case we would have updated the alignment mark field position with the applied T_y offsets in Layer 1, the translation overlay penalties in Figure 5 would have disappeared.

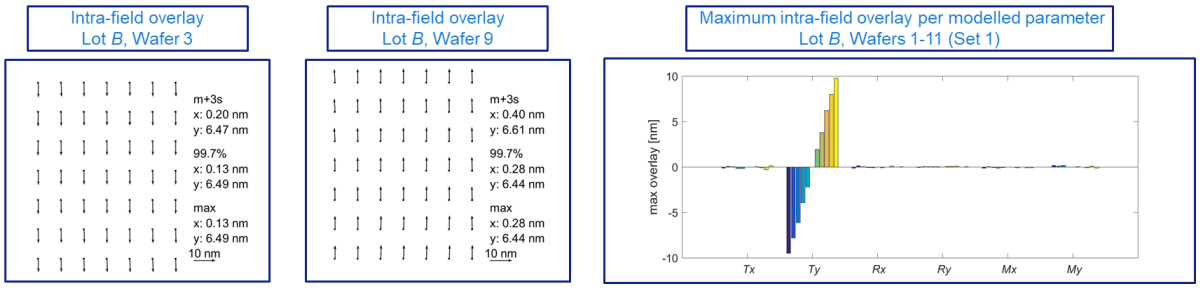

Figure 5: As an example, the intra-field overlay performance for wafers 3 & 9 is shown. These wafers received T_y translation offsets of -6-nm and +6-nm for both Layer 1 and 2, respectively. The T_y offset applied during the second layer exposure directly affects the measured overlay. The same observation can be made for all the other wafers as summarized in the bar plot on the right-hand side of the figure. In order to compare the overlay impact, the maximum overlay inside the field per modelled parameter is presented. All results are obtained after correction with reference wafer 6.

In Figure 6 we show the measured intra-field overlay results for the second set of 11 wafers in lot *B*. The exposure corrections that are assigned to these wafers when Layer 2 is exposed are based on the exposure corrections applied in Layer 1. This means that in addition to an M_y correction, a T_y correction based on the location of the wafer alignment mark in Layer 1 is also required. Only when both correction terms are applied, the resulting intra-field overlay performance is minimized. Figure 6 shows two intra-field overlay plots for wafer 14 & 20 as examples. The maximum overlay per modelled parameter of all wafers is shown in the bar plot on the right-hand side. It basically shows that if the field coordinates of the wafer alignment mark that is used in Layer 1 are known, the displacement at that specific field

location can be calculated and implemented as an overlay translation correction in Layer 2. Although we only considered an M_y correction term during Layer 1 exposure, the same holds true for R_x and R_y and all higher order polynomial intra-field exposure correction terms. These results confirm that copying the applied exposure corrections in Layer 1 to Layer 2 to prevent overlay penalties does not work for the direct wafer alignment use-case. A translation correction is required as well. An alternative would be to update the field coordinates of the alignment mark that is used based on the corrections applied in Layer 1. In that case, copying the applied M_y for Layer 1 to Layer 2 would not have had an overlay impact.

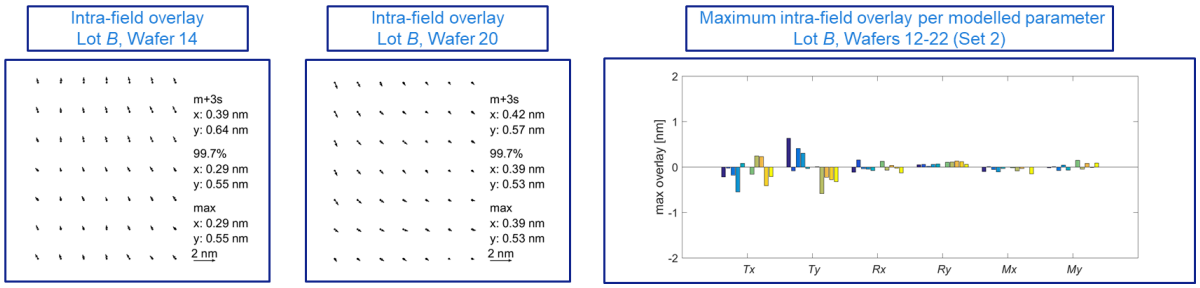

Figure 6: As an example, the intra-field overlay performance for wafers 14 & 20 is shown. These wafers received a M_y offset in Layer 1 of +1.1 ppm and -1.1 ppm, respectively. The measured intra-field overlay shows contains both a T_y and an M_y signature. The same is observed for all the other wafers. A summary of wafers 12 to 22 is shown in the bar plot on the right-hand side of the figure. In order to compare the overlay impact, the maximum overlay inside the field per modelled parameter is presented. All results are obtained after correction with reference wafer 17.

4. DISCUSSION

In the past few years, we have spent quite some effort to correlate the mask-to-mask overlay as measured on the PROVE® [4,5] with the on-wafer measured overlay [6]. The correlation turned out to be excellent enabling us to use off-line mask registration measurements to predict what will be measured on wafer. Initially, the focus was mainly on the mask-to-mask overlay performance [7]. However, it is important to realize that also the wafer alignment marks are exposed together with the product image. They are typically defined in the scribe lines in between the active dies and are also susceptible to reticle writing errors. We decided to extend the work by including the impact of reticle writing errors on the wafer alignment mark and the layer-to-layer overlay. The static contribution was already addressed earlier [2], the present publication addresses the dynamic contribution.

The accuracy of the placement of the alignment mark on the wafer is important when considering the layer-to-layer overlay. First of all, the wafer alignment mark can have a reticle writing error itself. This will cause a deviation from the designed position. In addition, the reticle alignment marks also have writing errors. Since these marks are measured by the scanner, the reticle writing errors on these marks will result in field location dependent offsets. These offsets depend on the number of reticle alignment marks as well as the reticle alignment model that is used. As a result, the wafer alignment mark position gets an additional offset after reticle alignment. The sum of the two contributors makes the actual alignment mark position deviate from its design position.

The wafer alignment marks locations as defined in the exposure job are based on their design coordinates. In case a wafer alignment mark is used for wafer alignment to expose the next layer, the delta between the actual position and the designed position directly ends up as a translational overlay error. This type of errors can be predicted (and hence prevented) by the off-line registration measurements on the PROVE® as we demonstrated in an earlier publication this year [2]. These errors are static by nature and can in principle be eliminated by an exposure correction in the subsequent layer. However, it would be more elegant to apply exposure corrections during the first critical mask exposure to bring the actual positions as close as possible to the design positions. The first (critical) mask exposure is quite unique in this context since all subsequent layers are referenced back to this layer. An off-line dense mask registration measurement on the PROVE® enables this functionality and is therefore recommended.

Besides the corrections made after reticle alignment, many other additional intra-field corrections can be present as well. They can originate from reticle heating, lens heating, wafer heating, pellicle effects, scanner stability/matching, etc. Applications have been developed to reduce the displacement penalties that are introduced by these error sources. If these applications are not used or if the displacement errors are not fully corrected, the residual errors at the wafer alignment mark locations result in additional translation (overlay) penalties.

Intra-field corrections can also come from outside the scanner. They are often automatically generated by the fab automation system and referred to as automated process corrections (APC). These corrections are based on overlay measurements of earlier exposed lots. In the past, a 6-parameter (linear) intra-field correction model was sufficient to bring the measured overlay within specification. Today, the correction potential per field can be up to 38-parameters.

The overall goal of applying process corrections is to minimize the overlay between the layer that is being exposed with respect to a reference layer. In case the layer that is being exposed contains a wafer alignment mark that will be used in the subsequent layers, the (variations in the) applied corrections will result in an additional overlay translation variation in the subsequent layers. These overlay errors can easily be prevented. Since the applied exposure corrections in the first layer are known, the field coordinates of the alignment mark can in principle be updated with this knowledge. Alternatively, and more common, the exposure corrections for the subsequent layers can be pre-determined and applied during the subsequent layer exposure.

5. CONCLUSIONS

In this work, we have studied the wafer alignment mark placement accuracy impact on the layer-to-layer overlay performance. The wafer alignment mark is defined in the same layer to which the overlay is measured. The actual position of the wafer alignment mark is impacted by both the reticle writing error contribution on the mark itself and on the reticle alignment marks via the reticle alignment model. The delta between the actual position and the design position results in a translation penalty in the measured overlay. This contribution is understood, as was demonstrated in an earlier publication. In this work we have captured this overlay contribution by making use of reference wafers for which no additional exposure corrections are applied. The overlay measured on the reference wafers has been removed from the wafers that received an exposure correction. This enables a clean study on the impact of additional intra-field corrections on the wafer alignment mark accuracy and the impact on the layer-to-layer overlay.

In the current investigation, we have studied the overlay impact after applying exposure corrections to the (first) layer in which the wafer alignment mark is defined. These exposure corrections change the field coordinates of the wafer alignment marks in Layer 1, resulting in translation penalties in the layer-to-layer overlay. We conclude that these overlay penalties can be predicted very well since the applied first layer exposure corrections are known. This also means that we can control the resulting translation penalties very well by applying appropriate translation corrections when exposing Layer 2. An alternative would be to update the field coordinates of the wafer alignment mark based on the applied exposure correction. Although the end overlay performance would be the same, it would enable easy copying of the applied exposure (or process) corrections from one layer to the next.

REFERENCES

[1] Wim de Boeij, Remi Pieternella, Igor Bouchoms, Martijn Leenders, Marjan Hoofman, Roelof de Graaf, Haico Kok, Par Broman, Joost Smits, Jan-Jaap Kuit, Matthew McLaren, "Extending immersion lithography down to 1x nm production nodes," Proc. SPIE 8683, 86831L (2013).

[2] Richard van Haren, Steffen Steinert, Orion Mouraille, Koen D'havé, Leon van Dijk, Jan Hermans, Dirk Beyer, "The impact of the reticle and wafer alignment mark placement accuracy on the intra-field mask-to-mask overlay," Proc. SPIE Vol. 11178 111780R (2019).

[3] Hitoshi Sunaoshi, Yuichi Tachikawa, Hitoshi Higurashi, Tomohiro Iijima, Junichi Suzuki, Takashi Kamikubo, Kenji Ohtoshi, Hirohito Anze, Takehiko Katsumata, Noriaki Nakayamada, Shigehiro Hara, Shuichi Tamamushi and Yoji Ogawa, "EBM-5000: Electron beam mask writer for 45 nm node", Proc. of SPIE Vol. 6283 628306-1 (2006).

[4] Dirk Seidel, Michael Arnz, Dirk Beyer, “In-die photomask registration and overlay metrology with PROVE using 2D correlation methods”, Proc. of the SPIE 8166 (2011).
[5] Steffen Steinert, Hans-Michael Solowan, Jinback Park, Hakseung Han, Dirk Beyer, Thomas Scherübl, “Registration performance on EUV masks using high-resolution registration metrology”, Proc. of SPIE 9985 (2016).
[6] Richard van Haren, Steffen Steinert, Christian Roelofs, Orion Mouraille, Koen D’havé, Leon van Dijk, Dirk Beyer, "Off-line mask-to-mask registration characterization as enabler for computational overlay," Proc. SPIE Vol. 10451, 1045111 (2017).
[7] Richard van Haren, Hakki Ergun Cekli, Xing Lan Liu, Jan Beltman, Anne Pastol, Jean Massin, Emilie Dupre La Tour, Maxime Gatefait, Frank Sundermann, “Impact of reticle writing errors on the on-product overlay performance”, Proc. of SPIE Vol. 9235 923522-1 (2014).