Addressing die displacement in lithography processes for advanced packaging applications

Chia-Ching Huang, Peter van der Krieken, Mikhail Loktev, Sylvain I. Misat*,
Yutai Lee, Jeroen de Boeij, Michiel van der Stam
Kulicke & Soffa Liteq BV,
Hooge Zijde 32, 5626 DC Eindhoven, the Netherlands
*Email: smisat@kns.com

Abstract

To perpetuate Moore's Law, semiconductor manufacturers are adopting advanced packaging technology as the mainstream solution. As a result, today, fan-out wafer level packaging is a key enabler of two- and three-dimensional packaging solutions. Among the manufacturing challenges, die displacement in lithography processes has been widely discussed, but a total solution is yet to come. In this study, we demonstrate die-by-die and global wafer alignment strategies available on-board the Kulicke & Soffa LITEQ 500 projection stepper. The versatility of a die-by-die alignment strategy allows for working with reconstituted wafers without resorting to an external metrology tool. The die displacement residuals can be reduced by three orders of magnitude, typically, from 100um to 100nm. The evaluation of overlay accuracy by implementing a die-by-die and global wafer alignment strategy indicates comparable performances. Considering high volume manufacturing, advanced process control of overlay is also capable of applying to a batch of wafers.

Key words

Advanced packaging, fan-out wafer-level packaging, die displacement, die-by-die alignment, global wafer alignment, reconstituted wafer, overlay, advanced control, high volume manufacturing.

I. Introduction

The front-end chip manufacturing industry is confronted with obstacles of transistor scaling and processes. In the past two decades, several milestones were achieved by novel innovations, such as immersion lithography, high-κ dielectrics, three-dimensional (3D) transistors, etc. However, it has been anticipated that the dimension of the transistor gate will reach the physical limit in coming years [1]. Therefore, semiconductor manufacturers have introduced an alternative solution, the advanced packaging initiative, to sustain the Moore's law [2]. For instance, in one such new approach the traditional monolithic chip is transformed to chiplets which are manufactured in existing individual technology nodes and then reassembled as a polylithic chip. Fan-out wafer level packaging (FOWLP) is one of the key assembly technologies which has been developed for 5th generation mobile network (5G), Antenna in Package (AiP), High Performance Computing (HPC), Artificial Intelligence (AI), Advanced Driver-Assistance System (ADAS), Co-Packaged Optics (CPO) and the Internet of Things (IoT) applications by 2D to 3D chip stacking [3].

In high volume manufacturing (HVM), a primary challenge of FOWLP is die displacement. The physical displacement of the embedded dies, generally, results from die placing, weak attachment, epoxy molding process, local deformation, and so on. There are few manufacturing techniques to mitigate the die displacement in the consecutive process, such as (i) molding dummy dies on the reconstituted wafer to predict corrective parameters; (ii) adding an additional interconnect layer for contact landing; and (iii) placing dies in a micrometer-precision range. These techniques improve the chip yield but alleviate productivity [3].

To address die displacement, the LITEQ 500 projection stepper deploys the die-by-die alignment, which allows to align multiple marks per die. As previously reported, die displacement can then be precisely predicted by up to 4 linear variables per die before wafer exposure. The demonstration is performed on designed wafers. The overlay performances are compared between a die-by-die alignment strategy and a global wafer alignment strategy [4], [5], [6].

II. Experiment

Preliminary

In this study, the die displacement on the reconstituted wafer is reproduced by multiple wafer load-and-unload sequences. Only one die is created on the wafer per wafer load on the stepper. There are 27 dies in total on the wafer

and 96 overlay references in each die. The die size is two times the image size of a front-end reticle. The blank substrate is a 300mm double-sided polished silicon wafer coated with 1.2µm photoresist (THMR-iP3250, TOK). After creating the zero layer (Layer 0) as explained above, four of the marks in each die are assigned for die-by-die alignment strategy in the process recipe, as shown in Fig. 1(a).

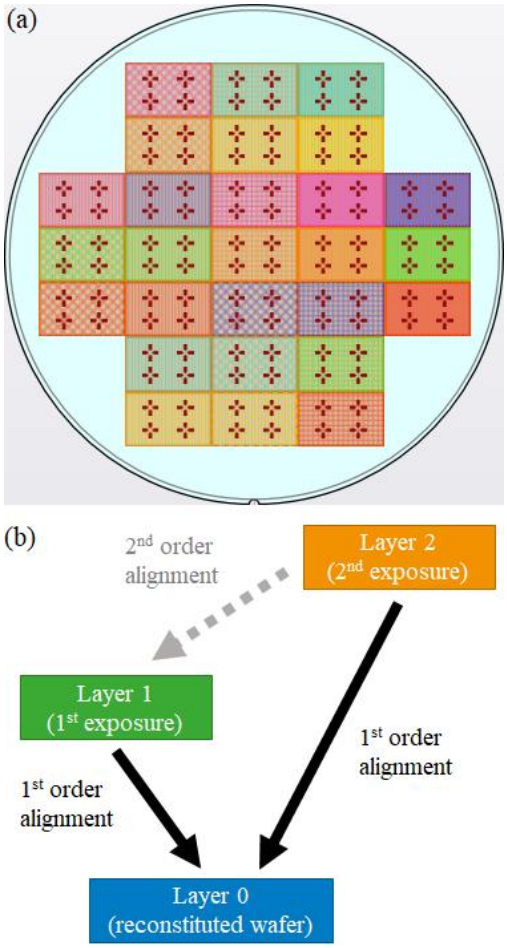

Fig. 1. (a) The designed layout for creating reconstituted wafers (colorful rectangles) and assigned marks (red cross) for die-by-die alignment strategy. (b) The diagram shows the alignment algorithm.

The consecutive processes are building up another two layers, Layer 1 and Layer 2, on the wafer. The subsequent steps consist in exposing Layer 1 and Layer 2 which are each aligned with Layer 0, as shown in Fig 1(b). In other words, two separate images, one for Layer 1 and one for Layer 2 are exposed by aligning to marks on Layer 0. The algorithm employed in this step is called 1st order alignment. A 2nd order alignment algorithm which would focus on aligning Layer 2 to Layer 1 is not treated in the present work.

Prior to each wafer exposure, the metrology system measures the initial set of marks on the Layer 0. By adopting the 1st order alignment algorithm, the die-by-die alignment strategy is measured twice, and the alignment repeatability is examined. Besides, two different schemes of overlay accuracy are considered when the 1st order alignment algorithm is adopted. The overlay accuracy between Layer 1 and Layer 0 or between Layer 2 and Layer 0 is called subsequent layer overlay (SLO). The overlay comparison between Layer 1 and Layer 2 is called single machine overlay (SMO).

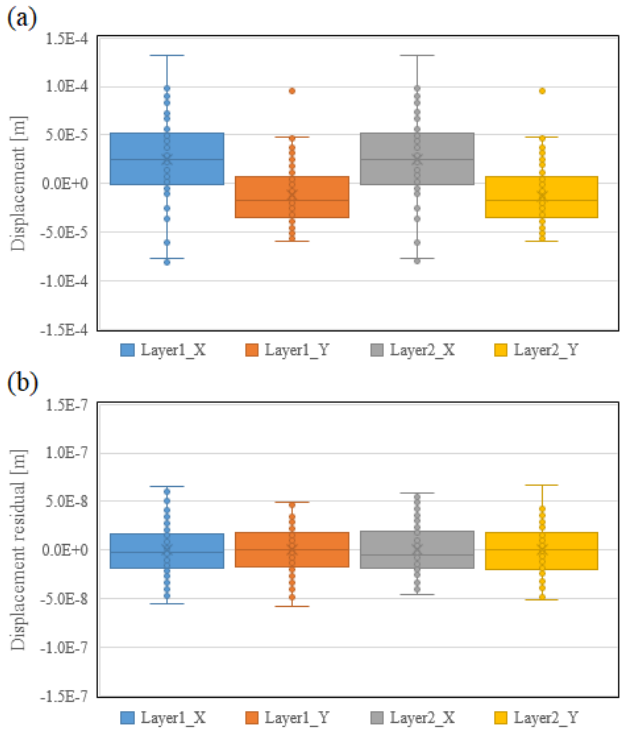

Fig. 2. The box-and-whisker diagrams illustrate the statistics of (a) displacement of mark measurement and (b) displacement residual on Layer 1 and Layer 2 in x- and y-axis, respectively.

Die-by-die alignment strategy

The metrology system in the stepper measures marks assigned in the process recipe. The deviation between set and measured coordinates of the mark is counted as displacement. In Fig. 2(a), the box-and-whisker diagram indicates the displacement distribution in x-axis and y-axis with respect to each layer. The measurement results in absolute mean value plus 3 standard deviations (|mean|+3s) are (131, 96) and (131, 96) in µm for Layer 1 and Layer 2, respectively. These displacement values are of the same order of magnitude as die placing accuracy currently reported for reconstituted wafers.

The wafer grid for each die is then computed based on the least-square method. The exposure route will follow this "ideal" wafer grid after compensating for the linear

variables. Linear variables, such as translations, rotations, and magnifications, are derived from grouping data into different die bases. Fig. 2(b) displays the displacement residuals where linear variables are subtracted from each mark. The displacement residuals in |mean|+3s are (77, 69) and (76, 72) in nm for Layer 1 and Layer 2, respectively. Ideally, compared to a raw measurement, the die-by-die alignment strategy suppresses the die displacement by about three orders of magnitude.

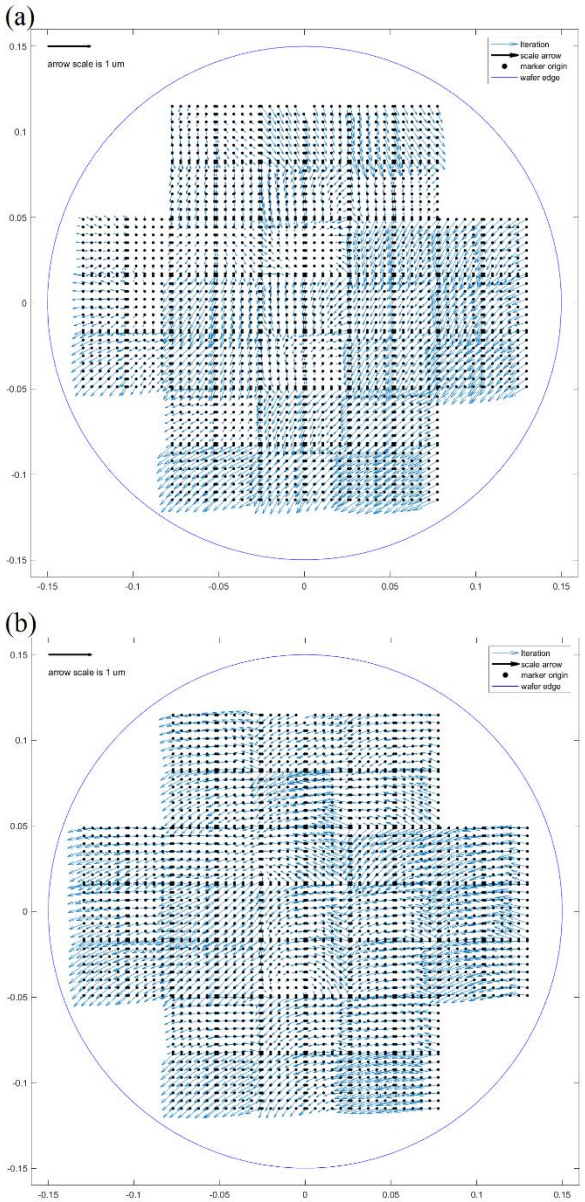

Fig. 3. The SLO fingerprints of (a) Layer 1 and (b) Layer 2.

The statistical prediction of die-by-die alignment strategy firstly substantiates a significant reduction from 100μm to 100nm scale. However, the pattern fidelity, in practice, conjugates with more parameters, such as ambient conditions, materials, thermal effects, post exposure processes, and so on. As a result, the overlay accuracy is one of inline control methodology in manufacturing to validate the pattern. To explore SLO accuracy in detail, a total of 2592 points per iteration on the wafer are measured. The capability exists on-board the stepper, thus removing the need to move the wafer to an external metrology tool. The displacement of each individual point is indicated by a vector, which shows its orientation and magnitude. Fig. 3 illustrates the SLO fingerprints of Layer 1 and Layer 2, in which the fingerprint varies from die to die. To further understand the reproducibility of the die-by-die alignment strategy, the SLO of Layer 1 is subtracted from that of Layer 2. Fig. 4 illustrates the SMO fingerprint between Layer 1 and Layer 2, in which die-to-die variation becomes imperceptible.

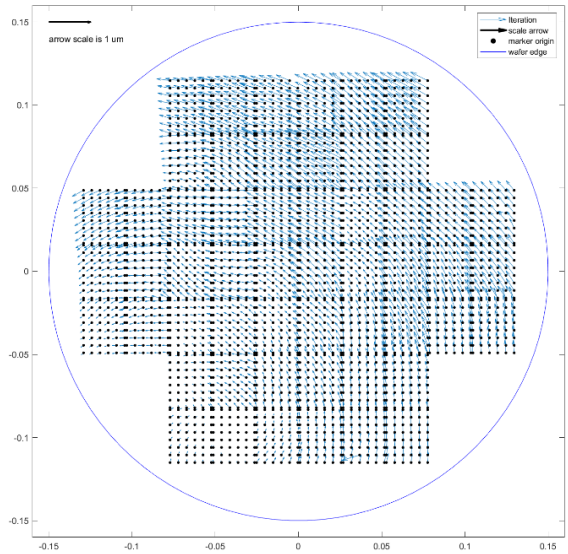

Fig. 4. The SMO fingerprints by subtracting Fig. 3(a) from Fig. 3(b).

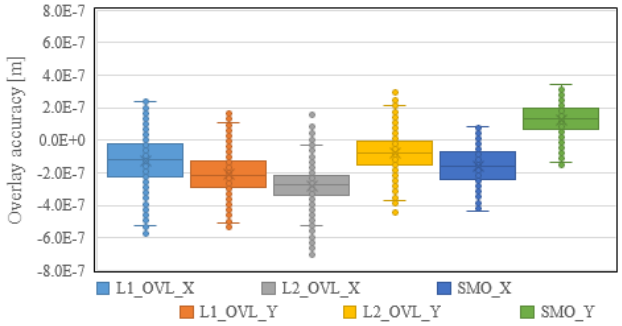

Fig. 5. The box-and-whisker diagram summarizes the |mean|+3s of two overlay schemes in x- and y-axis.

The box-and-whisker diagram in Fig. 5 breaks down SLO and SMO accuracy. The SLO accuracies in |mean|+3s are (568, 546) and (586, 398) in nm of Layer 1 and Layer 2,

respectively. The SMO accuracy in |mean|+3s is (476, 419) in nm which systematic errors contributed to Layer 1 and Layer 2 are mostly removed. Besides, the die-by-die alignment strategy has consistently contributed to the wafer grid around 100nm scale due to displacement residuals. This implies that the die-by-die alignment strategy is highly reproducible.

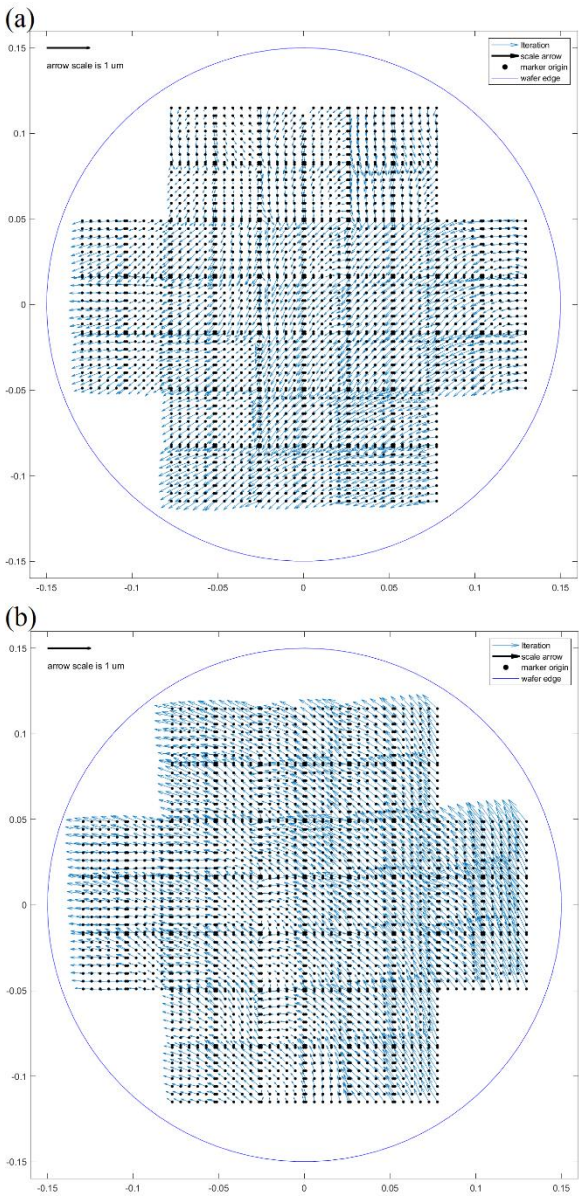

Fig. 6. The SLO fingerprints of two wafer exposures, (a) and (b), by adopting a global wafer alignment strategy.

Global wafer alignment strategy

The global wafer alignment strategy is adopted to determine the wafer grid by aligning multiple marks on a wafer. Here as well, to benchmark the die-by-die alignment strategy, wafers are created by a normal process job. Unlike the creation of a reconstituted wafer in the previous experiment, the wafer is loaded and unloaded one time in the stepper. Two layers are completed by aligning to seven marks on the wafer. The SLO comes with significant wafer-level fingerprints, as shown in Fig. 6. The SLO accuracies in |mean|+3s are (503, 381) and (520, 559) in nm for two layers, respectively. The SLO accuracies by both die-by-die and global wafer alignment strategy are significantly comparable.

For overlay monitoring, both SLO and SMO are performed by standard procedure. In Fig. 7, the average SLO and SMO accuracies in |mean|+3s are (357, 334) and (162, 149) in nm, respectively. The adoption of the overlay scheme depends on the material used, layout design, process control, etc. The implementation of higher order alignment algorithms in the lithography process might lead to higher complexity in overlay accuracy control [1].

Fig.7. The SLO and SMO monitoring by standard procedure.

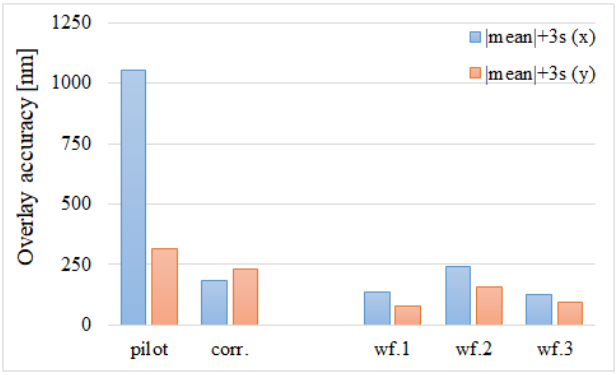

Fig. 8. The overlay accuracy of a batch imitates manufacturing conditions by implementing advanced control techniques. The term “pilot” corresponds to the SLO accuracy measured on the pilot wafer. The term “corr.” refers to the residuals in which wafer-level variables are subtracted from “pilot”.

Advanced process control for volume manufacturing

The use of SLO scheme is close to inevitable in the lithography process. The advanced control by implementing corrective parameters is a well-known technique in manufacturing. In this study, a pilot production wafer is first exposed on the stepper. The SLO accuracy is decomposed down to wafer-level variables and residuals. The statistics of raw (pilot) and residuals (i.e., after extracted wafer-level variables) in |mean|+3s are (1053, 315) and (192, 231) in nm, as shown in the left columns of Fig. 8. The same wafer-level variables are implemented onto a three-wafer batch as corrective parameters within the 3 hours, following the pilot wafer exposure. The average SLO accuracy of the batch in |mean|+3s is (167, 109) in nm which indicates a promising result as expected.

III. Conclusion

The die-by-die alignment strategy unveils the capability to suppress the die displacement from 100μm to 100nm scale. The overlay fingerprints in a dense measurement also indicate that each die depicts its individual behavior in sub-micrometer scale. The SLO accuracy in |mean|+3s is below 600 nm which is comparable to the wafers done by the global wafer alignment strategy. The SMO accuracy by die-by-die alignment strategy also resembles machine specification (below 500 nm) regarding to standard acceptance test.

Considering HVM perspectives, the overlay monitor performs consistently among wafers in either SLO or SMO schemes. For advanced process control, the corrective overlay parameters can first be derived from a pilot production wafer before being implemented on a batch of subsequent wafers. In the study, the average overlay accuracy in a batch is approximately of the same scale as the residual value of the pilot wafer.

The LITEQ 500 projection stepper offers remarkable versatility to deal with die displacement in lithography processes for developing and manufacturing advanced packaging applications.

Acknowledgment

We would appreciate all the team members in Kulicke & Soffa Liteq BV for the support.

References

[1] B. J. Lin, Optical Lithography: Here Is Why, SPIE Press, 2010, DOI: 10.1117/3.821000.
[2] J. H. Lau, Fan-out Wafer Level Packaging, Springer, Singapore, 2018.
[3] Chapter 23, Wafer-Level Packaging (WLP), Heterogeneous Integration Roadmap 2021 Edition, IEEE, 2021.
[4] M. Loktev, S. Misat et al., Overlay Diagnostics of Die-to-die Alignment on the Kulicke and Soffa LITEQ 500 Stepper, 2022 IEEE 24th Electronics Packaging Technology Conference (EPTC).
[5] S. Misat, M. Loktev et al., Die-to-die alignment for lithographic processing of reconstructed wafers; 2022 Smart Systems Integration (SSI) | 978-1-6654-8849-5/22.
[6] J. van der Voort and M. van der Stam, Unlocking the full Potential of Lithography for Advanced Packaging, 2380-4505-2018_1_pp. 000043-000050 (IMAPS 2018 - 51st International Symposium on Microelectronics - Pasadena, CA USA - Oct. 8-11, 2018).