PATHS TOWARDS ULTRA HIGH-DENSITY INTERCONNECT 3D SYSTEMS: CONTRIBUTION OF WAFER GEOMETRY FOR ACHIEVING SUB 200NM WAFER-TO-WAFER BONDING OVERLAY

SERENA IACOVO[1], JOERI DE VOS[1], ALAIN PHOMMAHAXAY[1], ANDY MILLER[1], ERIC BEYNE[1]
THOMAS UHRMANN[2], ANDREAS FEHKUHRER[2]
[1] IMEC BELGIUM [2] EVG AUSTRIA

16TH INTERNATIONAL CONFERENCE ON DEVICE PACKAGING 05/03/2020

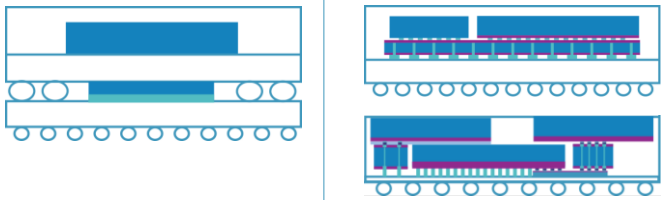
INTRODUCTION

THE 3D INTERCONNECT TECHNOLOGY LANDSCAPE

3D-SIP

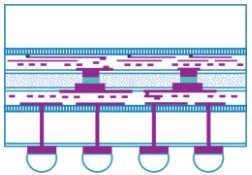
Package stacking

- Muliti-die Packaging
- Interposer "2.5D"
- Embedded Die



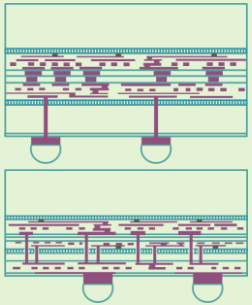
3D-SIC

Die Stacking
µbump

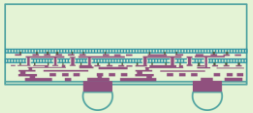


3D-SOC

Wafer-to-Wafer bonding

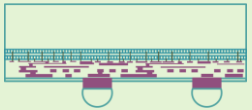


Wafer-to-Wafer Sequential Processing



3D-IC

Transistor Stacking



POR

PTW H2 2019

3D Interconnect Pitch scaling

1mm | 400µm | 100 µm | 40µm | 20µm | 10 µm | 4µm | 2µm | 1µm | 500nm | 100nm

3D Interconnect density (#/mm²)

1 | 10 | 100 | 1000 | 10^4 | 10^5 | 10^6 | 10^7 | 10^8

Package-to-board

Si-to-Package

Si-to-Si (D2D, D2W)

Wafer-to-Wafer (W2W)

Time

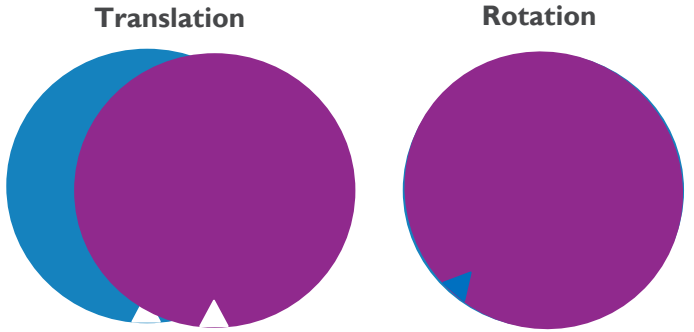
Courtesy of Eric Beyne

imec

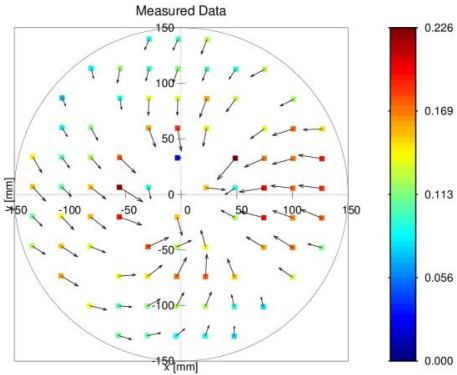
WAFER-TO-WAFER OVERLAY MODEL

Overlay = T + R + S + NL

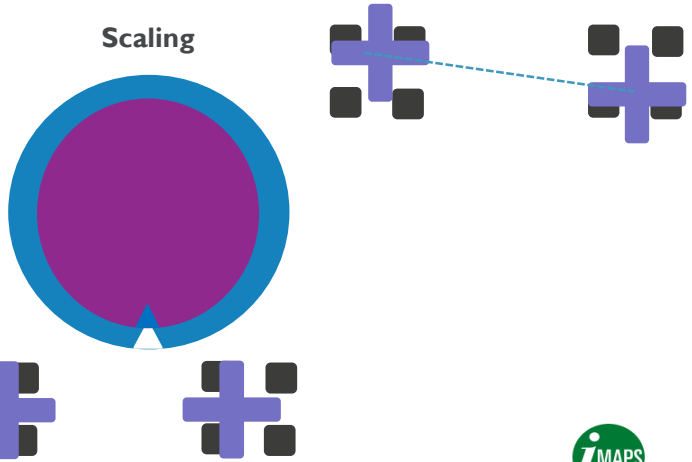
- Translation (X,Y)
- Rotation
- Scaling (ppm)
- Non linear components: Differences between the model and the measured data



Measured data: dx, dy



$$d_x = T_x + S * x - R * y + NL_x$$
$$d_y = T_y + S * y + R * x + NL_y$$



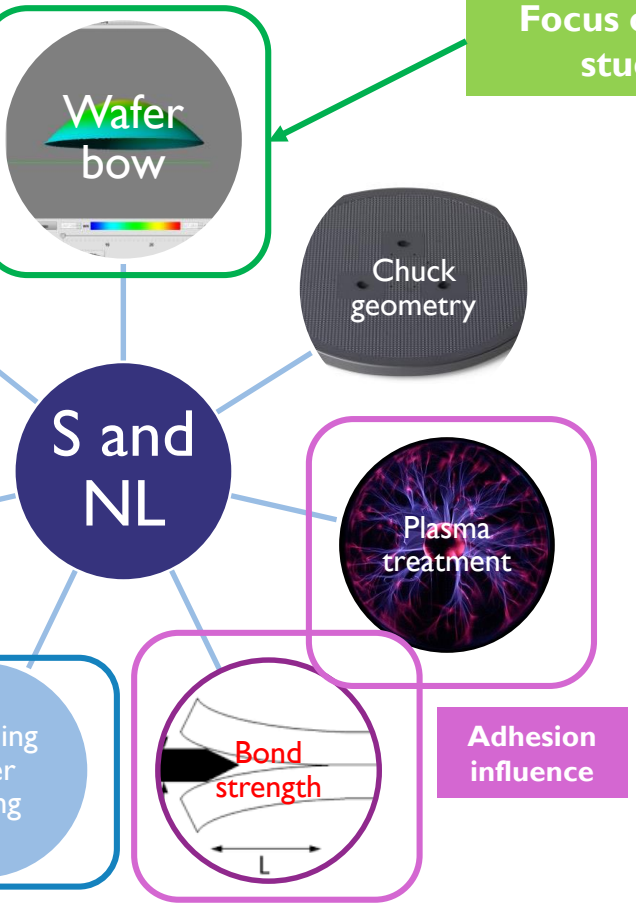
CHALLENGES IN THE ALIGNMENT OPTIMIZATION PROCESS FOR BONDING

Overlay = T + R + S + NL

Optical system

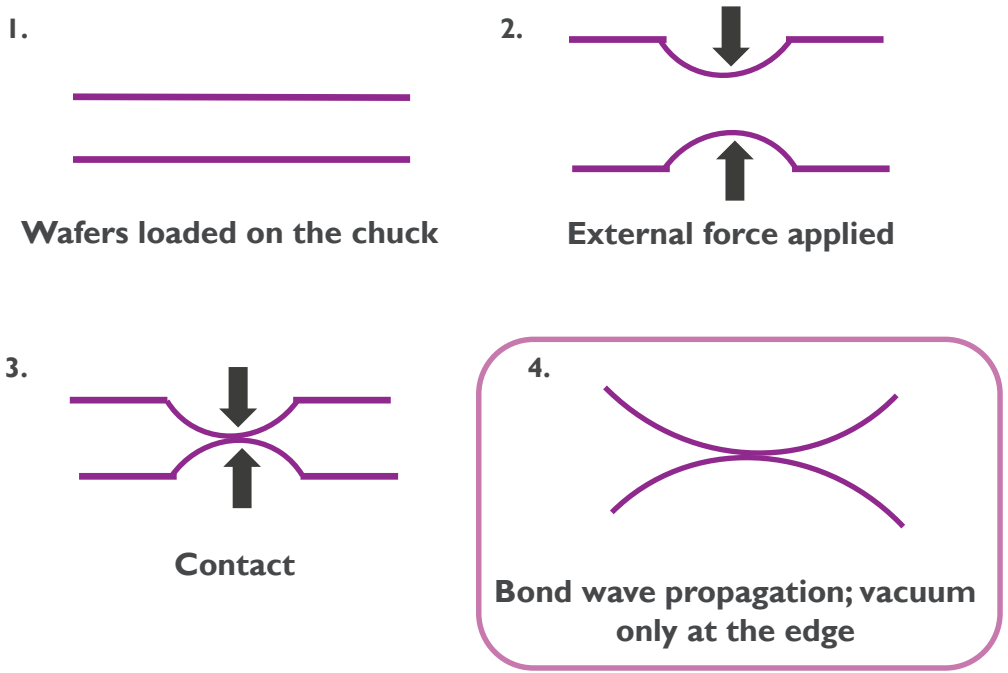
T and R

Alignment method



MOTIVATION

THE BONDING PROPAGATION MODEL



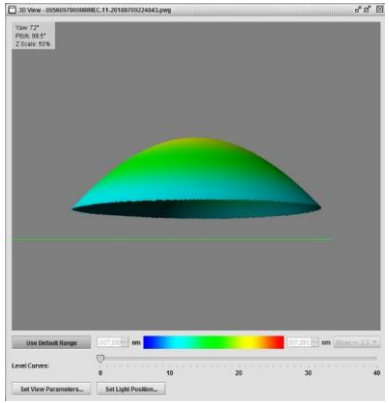
General sequence of direct wafer bonding

- Which impact the wafer bow will have on the bond wave propagation?
- How the scaling after bonding is going to be affected?
- Is there a way to control it and predict it?

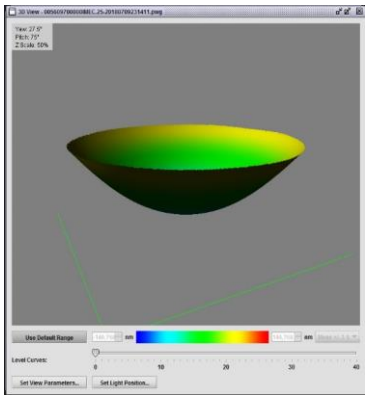
DESCRIPTION OF THE EXPERIMENT

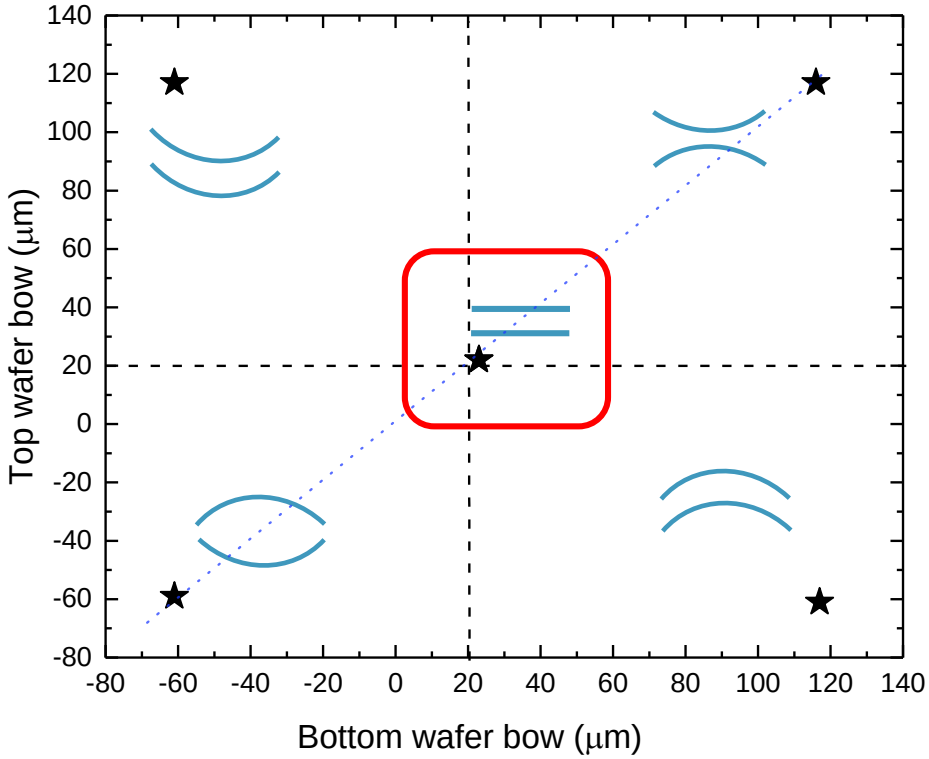
BONDING DOE

BOW>0, tensile



BOW<0, compressive

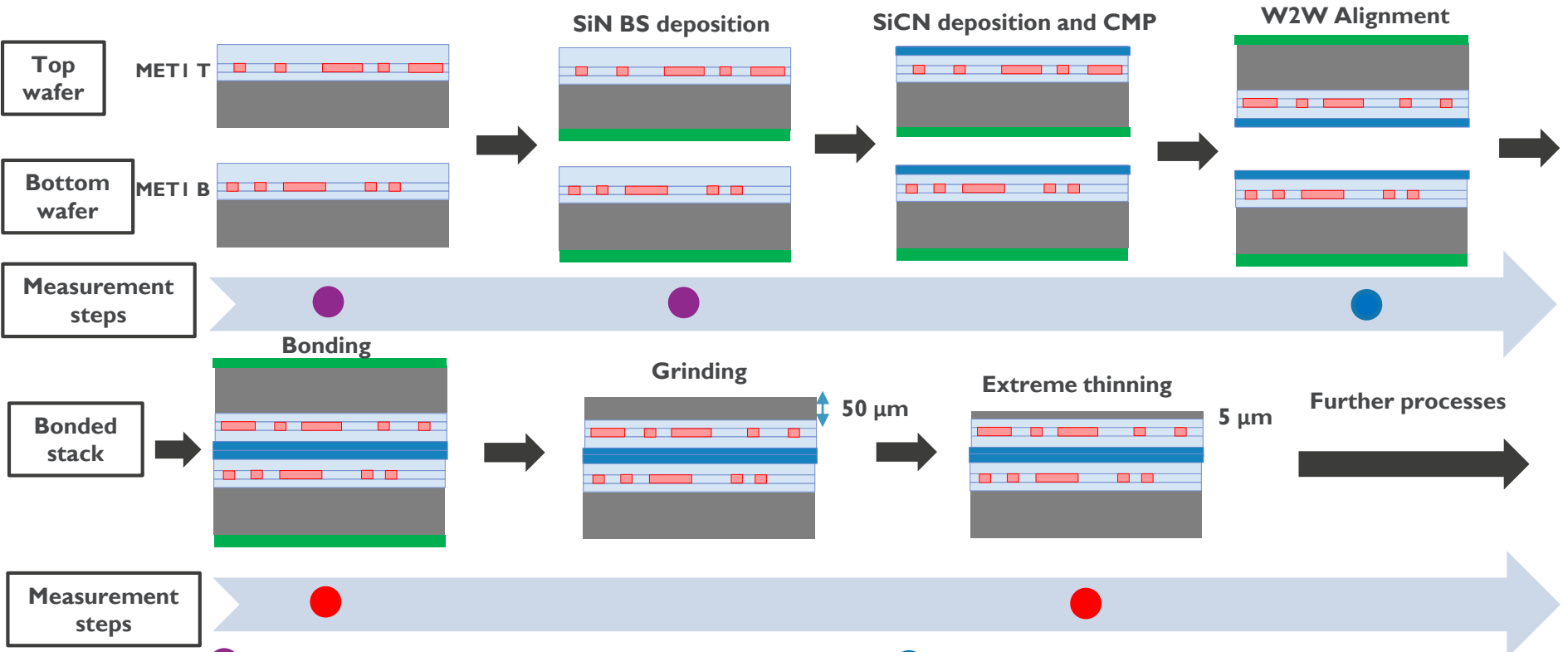




- SiN back side deposition used to tune the bow of the wafers
- Patterned wafers without BS SiN deposition have a natural bow of ~20 µm

- Different bonding test with different bow combinations defined
- Bonding recipe optimized only for the 1st wafer

INTEGRATION SCHEME: MEASUREMENT STEPS



RESULTS BEFORE BONDING

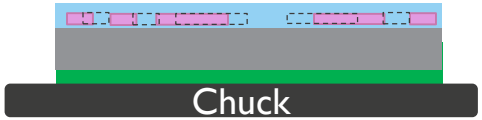
SCALING MEASURED BY THE SCANNER VS BOW

As prepared

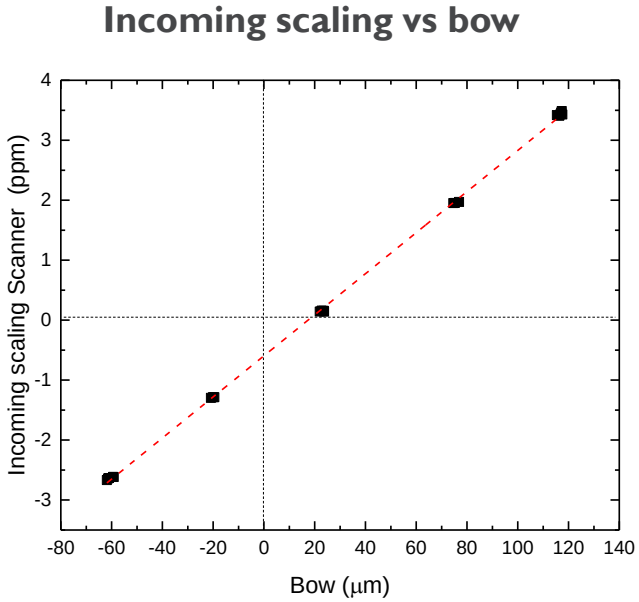
After BS deposition

Metal grid is deformed due to BS deposition

During measurements in the scanner



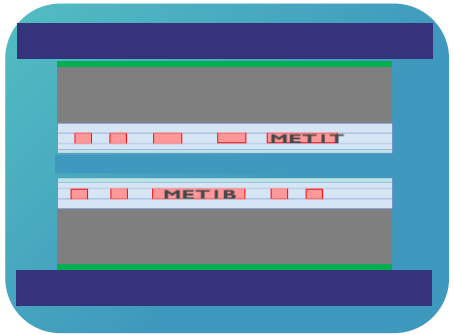
Wafer distortion can be captured by the scanner



- Measurement of the grid scaling on the scanner for each single wafer
- Linear correlation between wafer bow and scaling
  - Per 30 um bow, 1ppm scaling is added

BONDING SIMULATION

Smart View Aligner (SVA)



Wafers held flat once loaded

Distortion of both wafers to be considered!

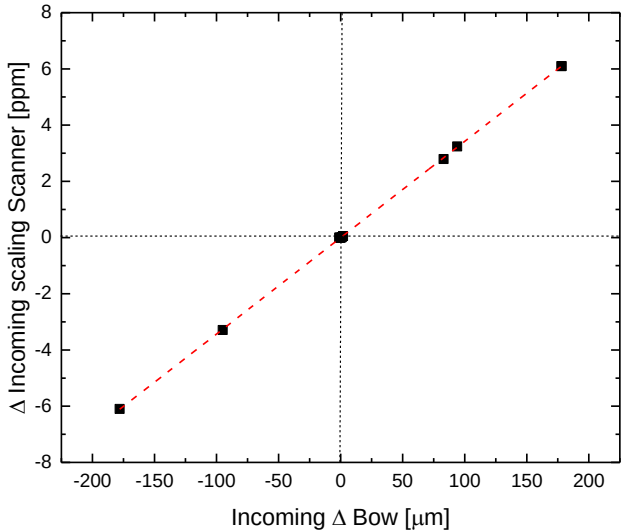
New parameter

Δ = Top – Bottom

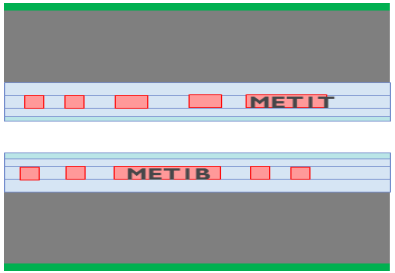
Bonder can give a rough estimation of the Incoming scaling

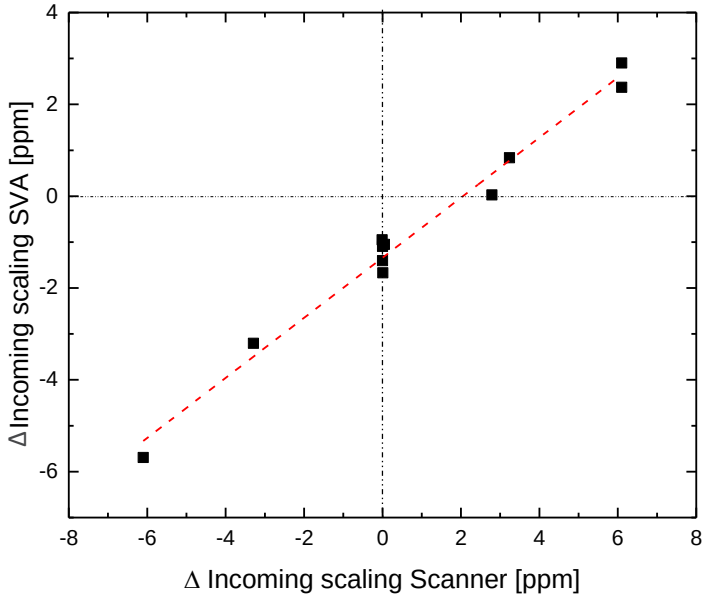
Incoming scaling calculated through the SVA

Δ Incoming scaling vs Δ bow



SCALING MEASURED IN THE SMART VIEW ALIGNER VS SCALING MEASURED IN THE SCANNER





$$\Delta Incoming\ scaling\ SVA = -1.4 + 0.6\ \Delta Incoming\ scaling\ Scanner$$

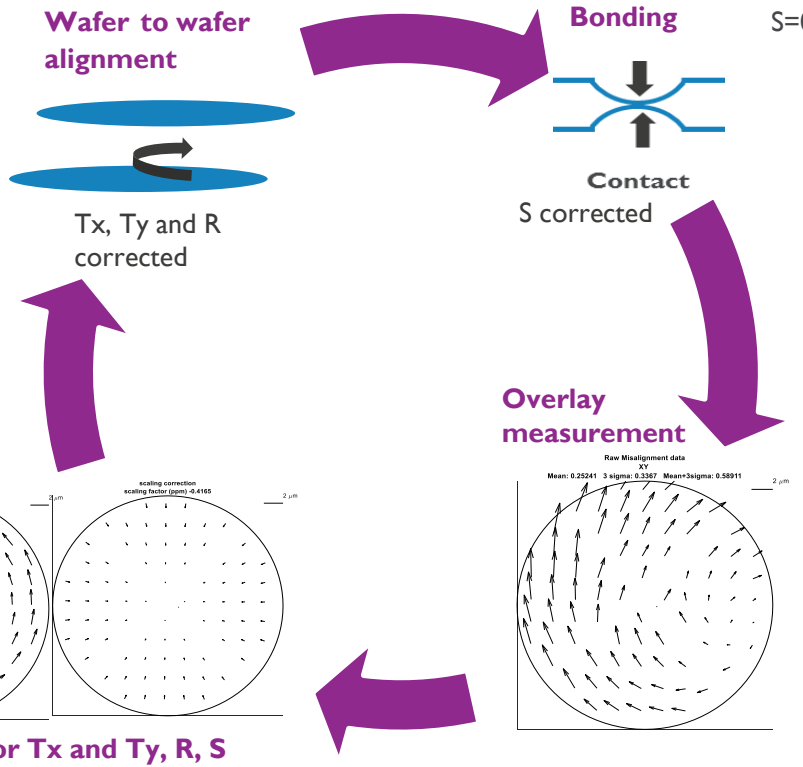
- Linear correlation between incoming scaling calculated in the SVA and the incoming scaling calculated in the scanner

ALIGNMENT OPTIMIZATION PROCESS

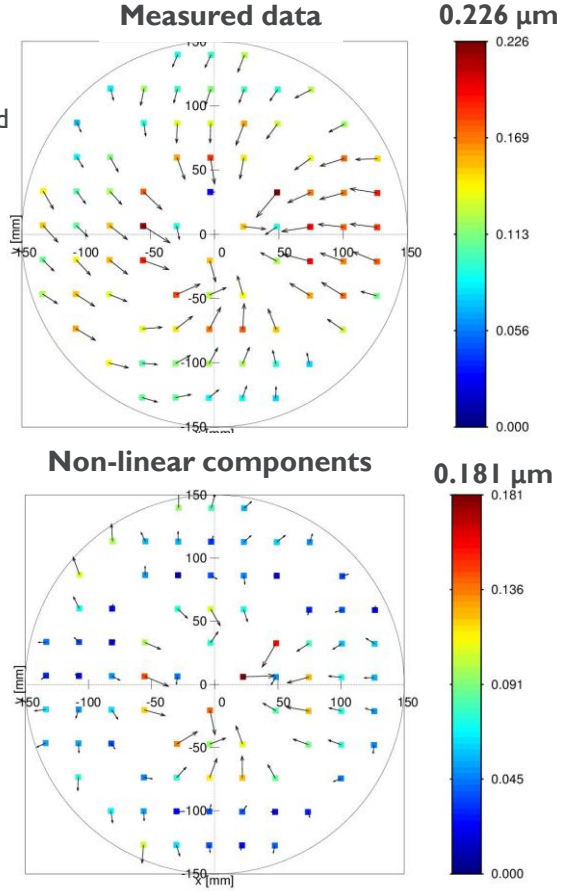
ALIGNMENT OPTIMIZATION PROCESS

RESULTS SETUP PAIR

- Recipe optimized for the setup pair and kept constant for the entire DoE to purely evaluate the bow effect on bonding
- Same offset used for all subsequent wafer pairs



Tx= 1 nm
Ty=28 nm
R= 0.464 µrad
S=0.995 ppm

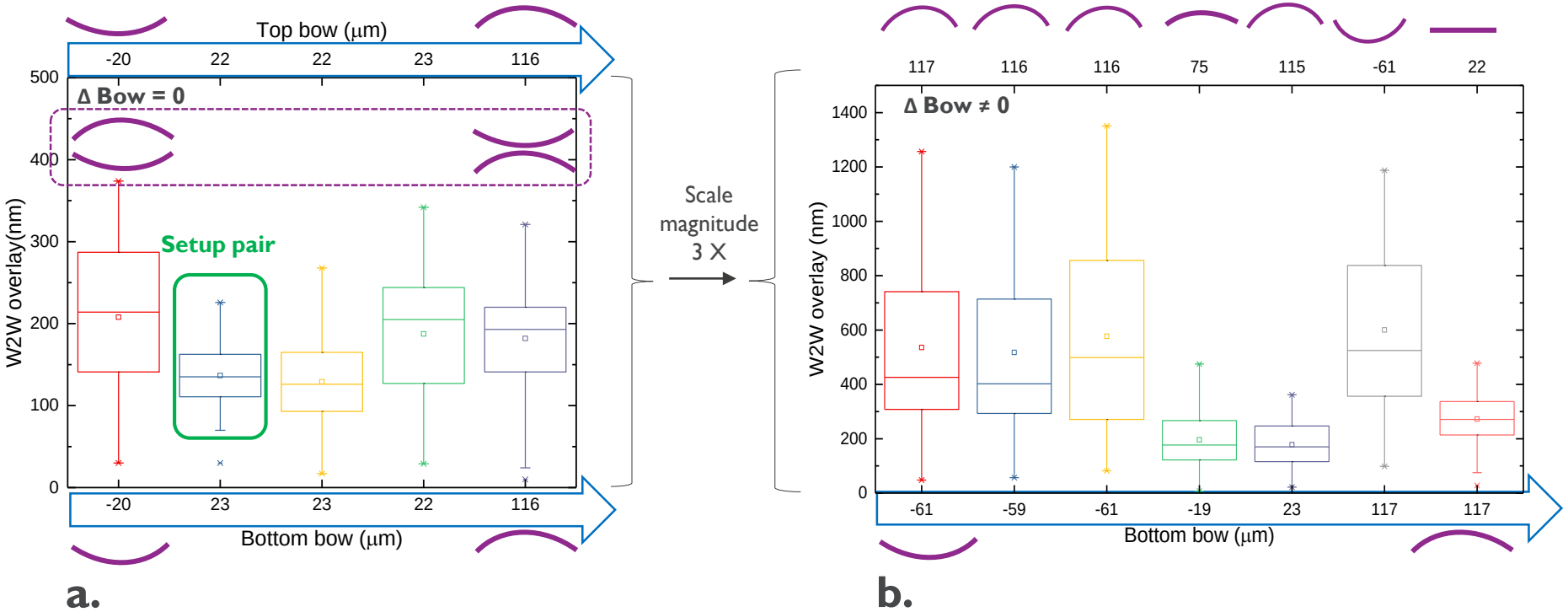


- Non-linear components represents ~80 % of the total misalignment on the max vector

W2W OVERLAY RESULTS

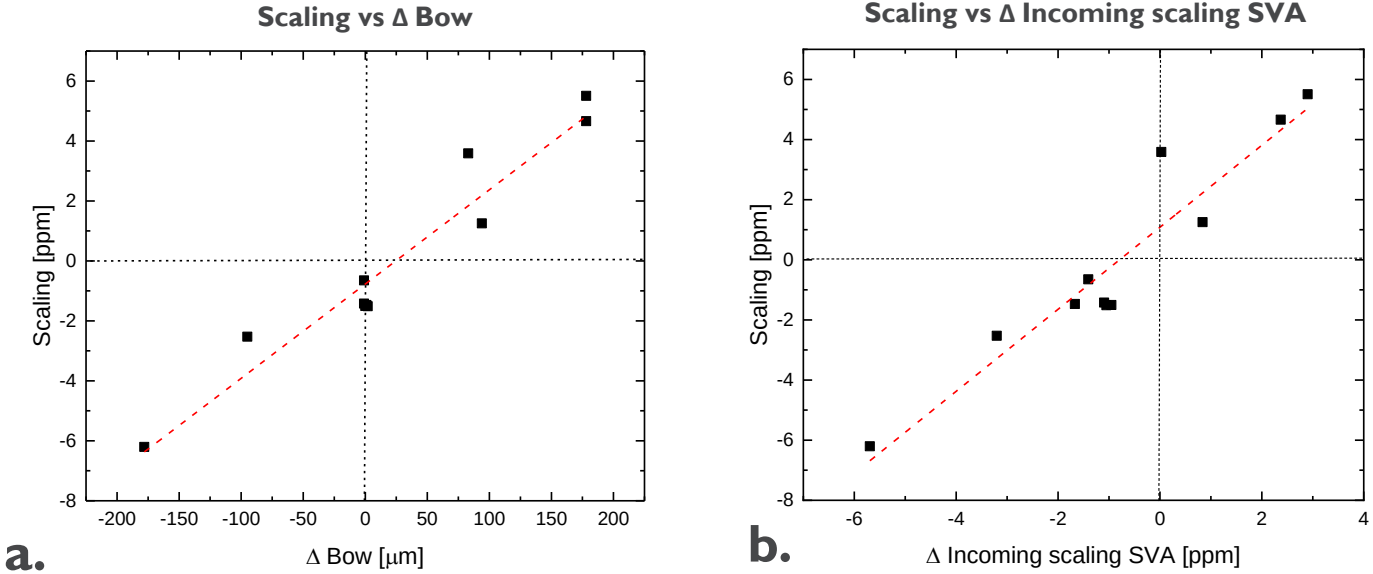
001133

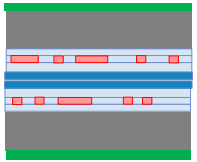
MEASURED OVERLAY



- Similar results obtained when the bow difference between top and bottom is equal to 0 (Δbow=0)
- Largest misalignment obtained for the mixed configurations when |Δ bow| > 90 µm

POST BONDING SCALING VS INPUT PARAMETERS



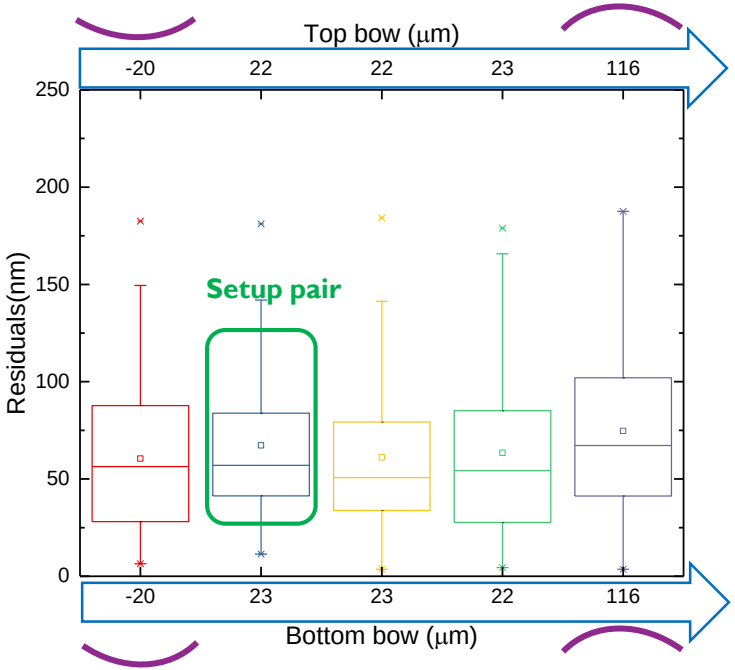


- Incoming scaling SVA linearly correlated to scaling after bonding ➔ Firm predictive power in understanding the scaling after wafer bonding
- Per 30 um bow difference between top and bottom wafer we add 1 ppm scaling at bonding ➔initial deformation kept during bonding

EFFECT ON NON-LINEAR COMPONENTS

NON-LINEAR COMPONENTS

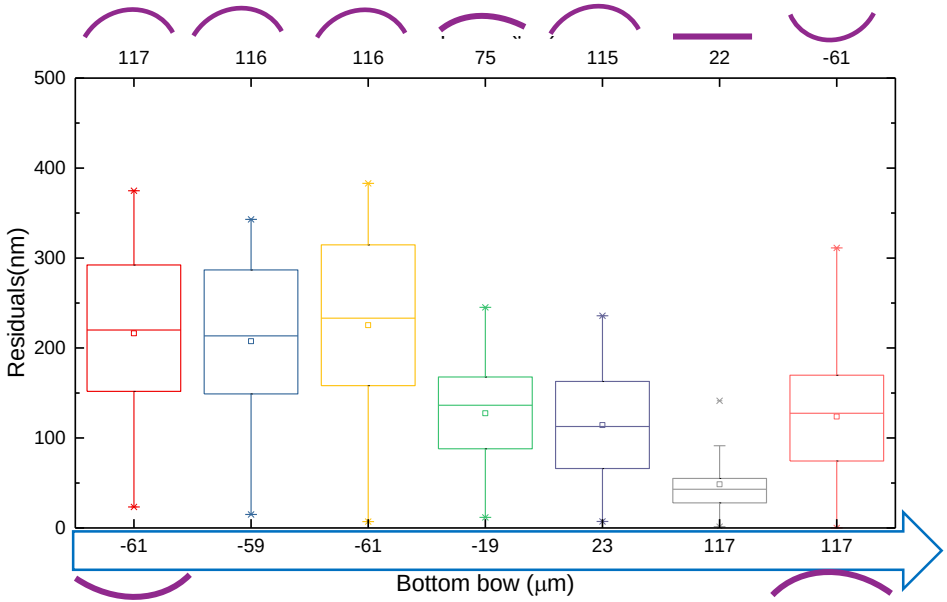
Δ BOW=0



- If Δbow=0 similar non-linear behavior observed

NON-LINEAR COMPONENTS

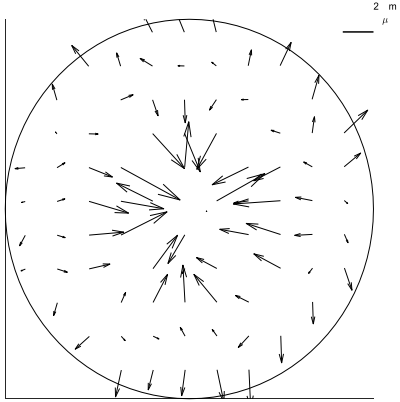
Δ BOW ≠ 0

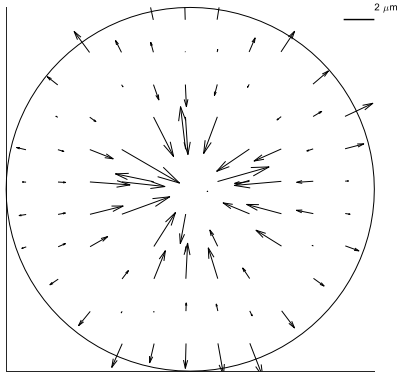


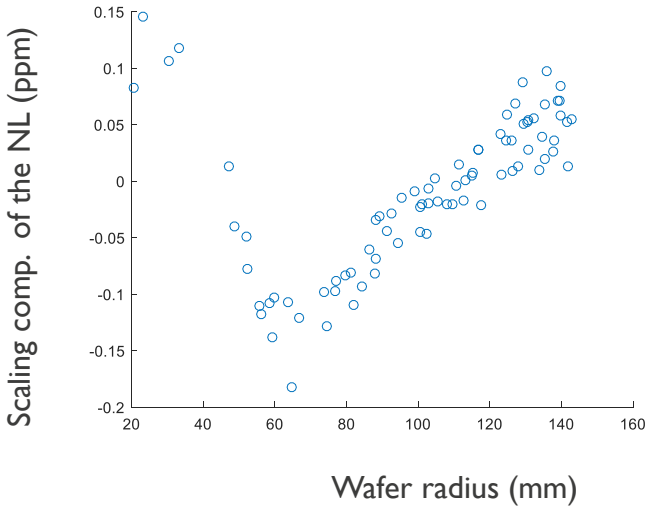
- NL behavior is changing depending on the combination of top and bottom wafer
- Certain combinations are allowing lower values compared to others
  - Best configuration: Top wafer flat and bottom wafer with positive bow

ANALYSIS OF THE RESIDUALS: AN EXAMPLE

Projection of the NL components on
the radial directions
"Scaling components of the NL"



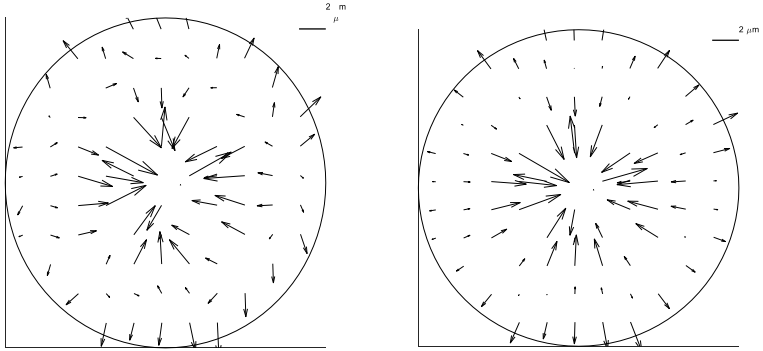




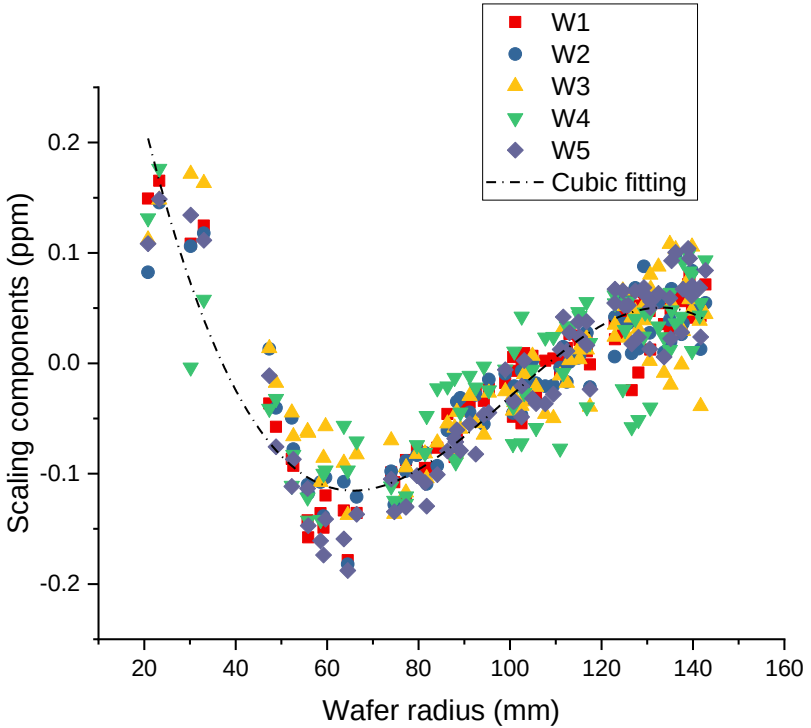
- From the residuals obtained in the experiments we try to extract a new parameter

ANALYSIS OF THE RESIDUALS: AN EXAMPLE

Projection of the NL components on the radial directions
"Scaling components of the NL"



Wafers with Δ bow=0

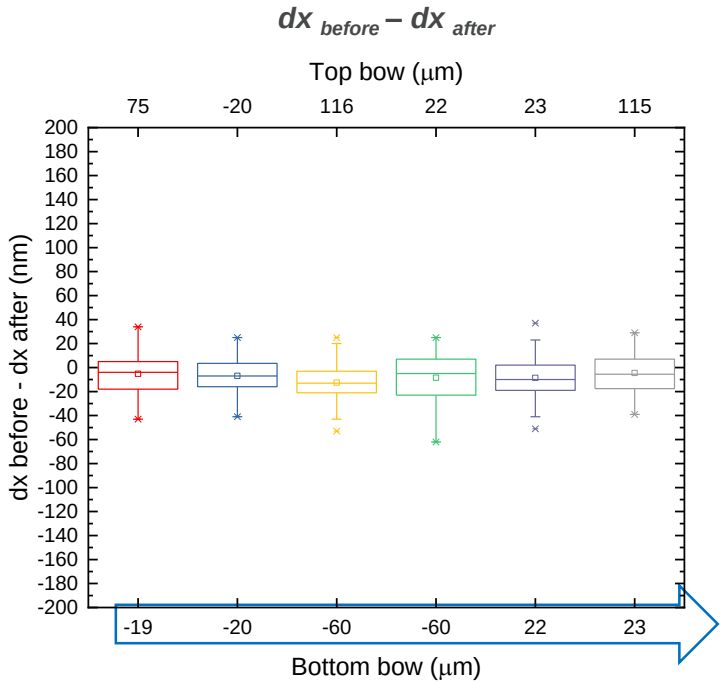


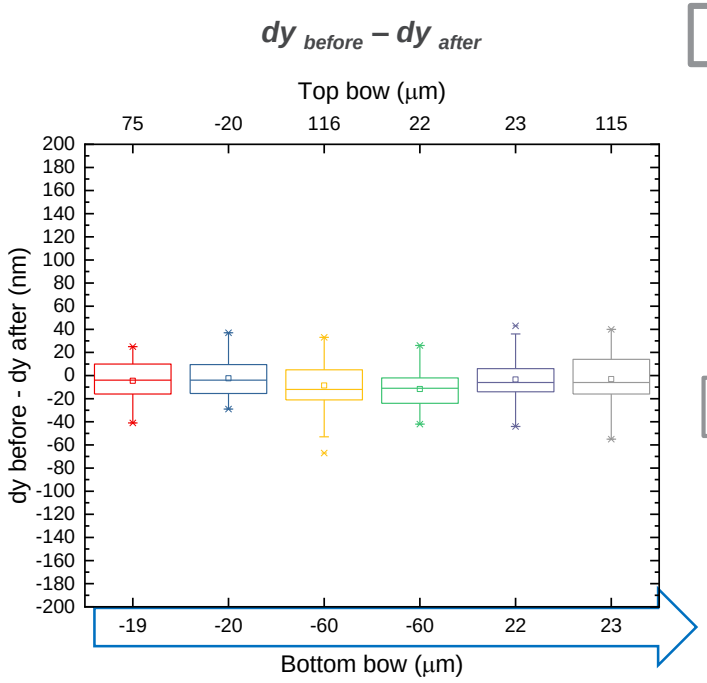
Overlay = T + R + S + NL

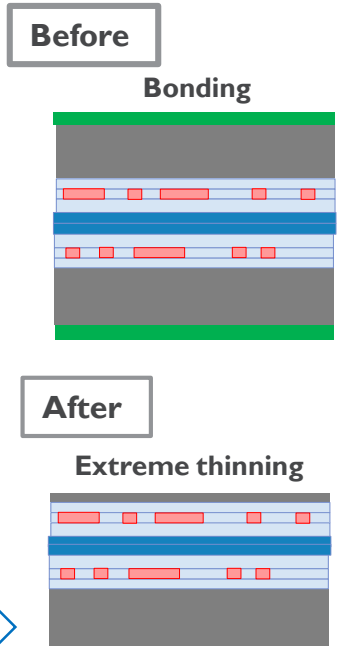
Overlay = T + R + S + S_{NL} + …

- If combination (top-bottom bow) is the same repeatable pattern can be recognized
  - ➔Room for improvement for overlay bonding model

OVERLAY AFTER WAFER THINNING







- Measured data before thinning – Measured data after thinning centered around 0
- Thinning does not have an impact on the alignment established at bonding

CONCLUSIONS

SUMMARY

BOW EFFECT

a setup pair

b *wafer* pair from the same batch of the setup pair

Evaluated for a certain tool configuration

Overlay = T + R + S + NL

$$S_b = S_a + \frac{\Delta \text{bow}}{30}$$

$f\ (incoming\ scaling) \Rightarrow$ *can be estimated through the SVA*

$f\ (top\ bow\ \&\ bottom\ bow)$

- Certain configurations are allowing lower values compared to others
- Results show that this term can be further split into new terms

CONCLUSIONS AND OUTLOOK

- Scaling after bonding linearly correlated to incoming scaling calculated in the smart view aligner.
  - Predictive power in understanding whether a process will lead to misalignment.
- An incoming bow change of 30 um will add 1 ppm scaling error in the final scaling after bonding.
- Non-random residuals measured after bonding
  - Repeatable pattern can be recognized. ➔ Room for improvement for overlay model.
  - Certain configurations are allowing lower values compared to others.
- Results are valid for a specific bonding chuck configuration ➔ Methodology generally applicable.
- Thinning does not have an impact on the alignment established at bonding.
- Different chuck configuration will be tested in the future.