YAP: Yield Modeling and Simulation for Advanced Packaging

Zhichao Chen, Puneet Gupta
Department of Electrical and Computer Engineering, University of California, Los Angeles, Los Angeles, CA, USA
zhichaochen@ucla.edu, puneetg@ucla.edu

***Abstract*—Three-dimensional integration technologies present a promising path forward for extending Moore's law, facilitating high-density interconnects between chips and supporting multi-tier architectural designs. Cu-Cu hybrid bonding has emerged as a favored technique for the integration of chiplets at high interconnect density. This paper introduces YAP, a yield model for wafer-to-wafer (W2W) and die-to-wafer (D2W) hybrid bonding process. The model accounts for key failure mechanisms that contribute to yield loss, including overlay errors, particle defects, Cu recess variations, excessive wafer surface roughness, and Cu density. We also develop an open-source yield simulator and compare the accuracy of the near-analytical yield model with the simulation results. The results demonstrate that YAP achieves virtually identical accuracy while offering over 10,000x faster runtime. YAP enables the co-optimization of packaging technologies, assembly design rules, and overall design methodologies. We used YAP to examine the impact of bonding pitch, compare W2W and D2W hybrid bonding for varying chiplet sizes, and explore the benefits of tighter process controls, such as improved particle defect density.**

***Index Terms*—yield modeling, hybrid bonding, wafer-to-wafer (W2W), die-to-wafer (D2W), 3D integration, chiplet, roughness, pad recess, peeling stress, particle defects, overlay, warpage/bow.**

I. Introduction

As we approach the physical and economic limits of scaling in two-dimensional integrated circuits, Three-Dimensional Integrated Circuits (3D-ICs) have emerged as a promising solution to sustain the momentum of Moore's Law. By stacking multiple layers of devices vertically, 3D-ICs offer significant benefits such as reduced interconnect lengths, enhanced performance, lower power consumption, and higher integration density. A pivotal technology that enables 3D-ICs is hybrid bonding (HB) which facilitates reliable fine pitch interconnections, making it highly suitable for high-bandwidth memory, logic-memory integration, and advanced sensor applications [1]–[4]. The two popular HB variants are Wafer-to-wafer (W2W) which allows for a higher alignment accuracy and die-to-wafer (D2W) bonding which allows for better leveraging of pre-tested known-good-die [1].

Accurate, predictive yield modeling for advanced packaging is essential is key to identifying potential failure mechanisms early in the technology development cycle, enabling system-technology co-optimization, and guiding the development of the chiplet interconnect repair strategy [6]. The overall assembly yield model for advanced integration is determined by multiple components, including chiplet yield, hybrid bonding yield, and through-silicon via (TSV) yield. Though full system yield with simplified underlying models has received a lot of attention [7]–[10], models of the yield of hybrid bonding have largely been overly simplistic. [7], [8] propose system-level yield models specific to 3D stacked ICs, However, [7] does not extend its yield model to include the bonding process, and [8] assumes the bonding yield as a constant, which is also done by [10] when building an assembly yield for a chiplet system. None of these studies considers the complex physical failure mechanisms or derives a concrete yield model for the HB process. Such modeling becomes even more important as the industry strives to rapidly scale down the bonding pitches in the next decade.

Failure mechanisms for hybrid bonding are complex which makes analytical modeling of yield challenging. This paper proposes YAP, a physical mechanism-driven near-analytical yield model. To the best of our knowledge, YAP is the first yield model tailored for the HB process, which provides a detailed analysis and modeling framework to predict and optimize bonding yield. The main contributions are summarized as follows:

- To the best of our knowledge, we present the first yield model specifically for the HB process. This model incorporates critical failure mechanisms contributing to yield loss, including overlay errors, particle defects, Cu recess variations, surface roughness of wafers and dies, and excessive Cu density.
- We construct a yield simulator based on the distribution parameters of various failure mechanisms to validate the proposed yield model and assess its predictive accuracy.
- We perform case studies examining various factors that impact yield, compare the yield performance of W2W and D2W HB approaches, and demonstrate the necessary process control for a high yield performance.

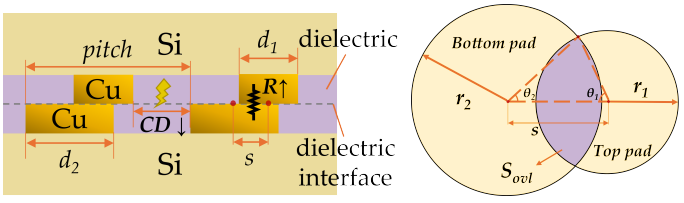
The remainder of the paper is organized as follows: Section II introduces key failure mechanisms of HB processes. Section III proposes modeling methodologies for W2W HB and extends it to D2W HB. Section IV describes the experiment setting, compares the results from the near-analytical model and the simulator, and conducts case studies on critical yield-influencing factors. Finally, Section V concludes the paper and gives suggesting directions for future research.

II. Overview of Failure Mechanisms of Hybrid Bonding

This section provides an overview of the primary failure mechanisms in HB processes, including overlay errors, Cu recess variations, and particle defects. Each mechanism contributes to yield loss when bonding parameters are inadequately controlled, and understanding these mechanisms is critical for optimizing the process yield.

A. Overlay Errors

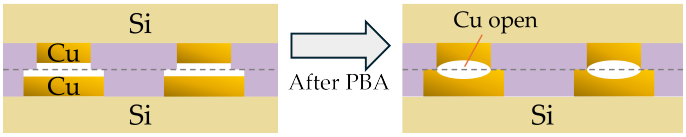
Ensuring the quality of the Cu connection formed during the HB process is critical for the electrical properties of the design. Misalignment of Cu pillars of the top and bottom wafers inevitably exists due to the robot arm calibration error, the warpage induced by thermal stress difference, etc. The misalignment influence on yield loss is increasingly pronounced as we are shrinking the pad size and using the sub-micron pitch [11]. As shown in Fig.1, the excessive misalignment will



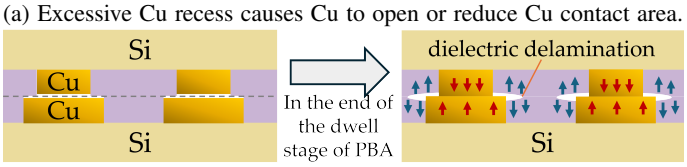
(a) Misalignment leads to resistance increase and dielectric breakdown.

(b) Contact area calculation given the misalignment.

Fig. 1: Failure mechanism of Overlay errors.

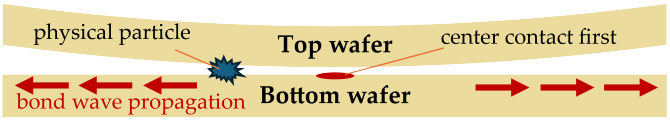


(a) Excessive Cu recess causes Cu to open or reduce Cu contact area.

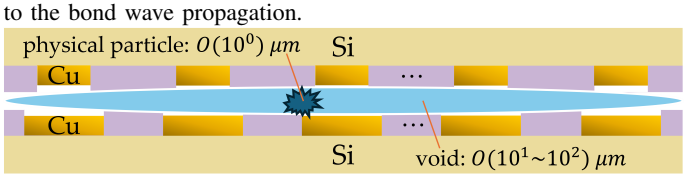


(b) Insufficient Cu recess may cause dielectric delamination, depending on the dielectric bonding strength, Cu density, and annealing temperature, etc.

Fig. 2: Failure mechanism of Cu recess variations.



(a) The presence of particle will form a main void and a void tail due to the bond wave propagation.



(b) Void formation can fail the dielectric and Cu bonding. A particle of a few microns can form a main void of hundreds of microns [38].

Fig. 3: Failure mechanism of particle defects.

decrease the contact area (S_{ovl}) of the Cu interface, hence increasing the contact resistance and triggering electromigration-related defects [12]. Also, the probability of dielectric breakdown will increase due to the reduced critical distance (CD), causing thinner insulating film between the upper and lower pads of adjacent pillars [2]. It is defined as the pitch being p, the pad being circular, and the top and bottom pad diameters being $d_1 = 2r_1$ and $d_2 = 2r_2$. The critical distance between two perfectly aligned Cu pillars is defined as $CD = p - d_2$, i.e., the distance between two neighboring pads. In some designs, the top pad size is set smaller than bottom pad size to increase the misalignment tolerance [28]. Overlay errors are typically classified as pad-level random errors and systematic errors.

B. Cu Recess Variations

The CMP process will result in Cu recess effects, forming a concave shape on the Cu pad surface. As presented in Fig. 2a, excessive Cu recess will decrease the bonding quality or even incur Cu connection failure after post-bond annealing (PBA) [17], [18]. Meanwhile, Cu protrusion and insufficient Cu recess can degrade the yield. Excessive wafer surface roughness decreases the effective contact area of the dielectric during low-temperature bonding, which lowers the density of covalent bonds after PBA, thereby weakening the bonding strength and energy per unit area at the dielectric interface [18]–[20]. In fine-pitch designs, high Cu density and insufficient Cu recess can incur an undesirable peak peeling stress at the dielectric interface at the end of the annealing dwell stage [21]–[24]. As can be seen in Fig. 2b, if the dielectric interface bonding is unable to withstand the peeling stress, dielectric delamination or cracking can occur, leading to bonding failure [25]–[27]. In summary, to achieve a high yield, especially for a chiplet with a large number of Cu pads, a well-controlled Cu recess variation of the top and bottom pad within a range determined by Cu pattern density, surface roughness, etc. is necessary.

C. Particle Defects

The HB process requires high-standard cleanliness to prevent the presence of physical particles, which will incur void formation at the bonding interface [1], [38]. In addition to physical particles, gas condensation during the bonding process can incur edge voids in the wafer bevel region. The outer edge region is usually removed by a sawing process and silicon dies or chips located inward from the wafer edge will not be influenced, therefore, there will be no yield impact [37]. In the defect model, the emphasis is on the yield loss resulting from the particle-induced void formation. As presented in Fig. 3, in the W2W HB process, due to the bond wave propagation, the presence of a particle at the bonding interface will result in a *main void* with a trailing *void tail* extending radially [38].

III. YAP Yield Model

In this section, we introduce our yield modeling methodology for W2W HB and then extend the model to D2W HB. To validate our derived model, we compare the modeling results with the simulation results across 300 parameter sets. The workflow of simulation and model validation is illustrated in Fig. 4. The detailed experiment setting will be discussed in Section. IV.

A. Overlay Model

Our proposed overlay model estimates the yield loss caused by the Cu pad misalignment. We assume the bonding misalignment is normally distributed with zero mean and process-specific σ_1 [13]. Then, the possibility of survival (POS) of one single pad can be calculated as

$$POS_{ovl,pad} = \frac{1}{\sigma_1\sqrt{2\pi}} \int_{-\delta}^{\delta} e^{-\frac{u^2}{2\sigma_1^2}} \mathrm{d}u \tag{1}$$

where u denotes the random overlay error of the top and bottom pad, and δ represents the maximum allowed overlay error to ensure the pad's survival. δ should be determined by the contact area constraints and the critical distance constraints. We assume s as the systematic overlay error of a Cu connection caused by three distortion components: *translation*, *rotation*, and *magnification* [14]. The translation and rotation mainly originate from the equipment precision limitations, and the magnification mainly comes from the wafer warpage/bow induced by thermal expansion mismatch of multiple materials [21]. We define the translation errors as T_x, T_y in x, y directions, respectively, and the rotation error as α. Typically, the warpage of bonded wafers can range from a few micrometers to over 100 μm, which can be optimized to $\sim$10 μm through run-out compensation [16]. Let B represent the warpage of the bonded wafer. It can be observed that the magnification factor E is linearly correlated to B [15], [16]. We can build a linear model to characterize the magnification factor E as

$$E = k_{mag} \cdot B \tag{2}$$

where k_{mag} is the fitting parameter of the model. The parameter is related to the Cu pad depth, Cu area density, bonding process temperature, etc [21]. We model the systematic misalignment $\Delta x, \Delta y$ in x, y directions respectively by

$$\begin{cases} \Delta x(x,y) = T_x - \alpha \cdot y + E \cdot x, \\ \Delta y(x,y) = T_y + \alpha \cdot x + E \cdot y. \end{cases} \tag{3}$$

The systematic overlay error s at the location (x, y) is by

$$s(x,y) = \sqrt{[\Delta x(x,y)]^2 + [\Delta y(x,y)]^2} \tag{4}$$

As Fig. 1b shows, the contact area of two Cu pads is by

$$S_{ovl} = \begin{cases} \pi r_1^2, & s < r_2 - r_1 \\ \theta_1 r_1^2 + \theta_2 r_2^2 - s r_1 \sin\theta_1, & r_2 - r_1 \le s \le r_1 + r_2. \\ 0, & s > r_1 + r_2 \end{cases} \tag{5}$$

If we assume that, to ensure the pad survival, the contact area should exceed k_{ca} times the surface area of the top pad interface, i.e., $S_{ovl} > k_{ca}\pi r_1^2$, and the critical distance CD should be greater than k_{cd} times the ideal critical distance, i.e.,

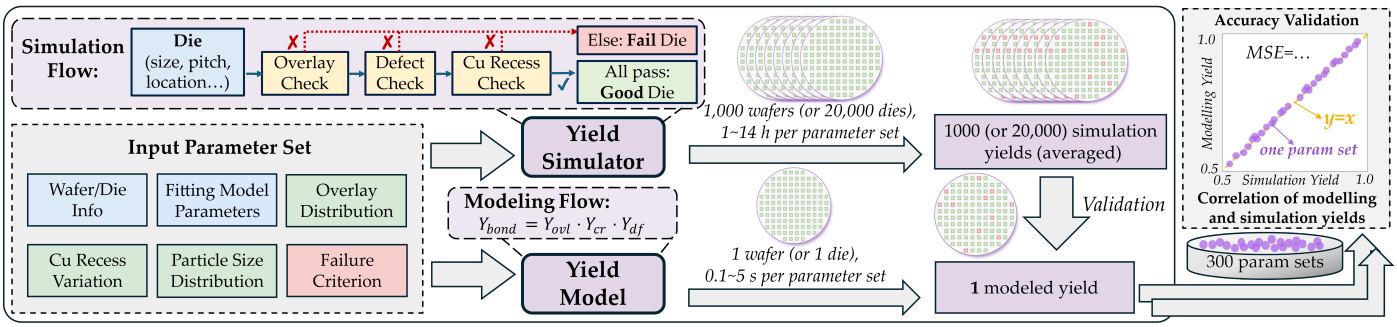


Fig. 4: Simulation workflow and the validation of modeling yield on various input parameter sets.

$CD > k_{cd}(p - d_2)$, then δ can be given by

$$\delta = \min \left\{ \frac{\theta_1 r_1^2 + \theta_2 r_2^2 - k_{ca}\pi r_1^2}{r_1 \sin\theta_1}, \right. \\ \left. (1 - k_{cd})p - \frac{1}{2}d_1 + \left(k_{cd} - \frac{1}{2}\right) d_2 \right\} \tag{6}$$

Given above, the POS of a die with N pads is given by

$$POS_{ovl,die} = \frac{1}{\sigma_1\sqrt{2\pi}} \min_{i\in[1,N]} \left\{ \int_{-\delta - s_i}^{\delta - s_i} e^{-\frac{u^2}{2\sigma_1^2}} \mathrm{d}u \right\} \tag{7}$$

where s_i denotes the systematic overlay residue of the i-th pad in one die. Assuming one wafer has M dies, the overlay die yield of the wafer can be calculated as

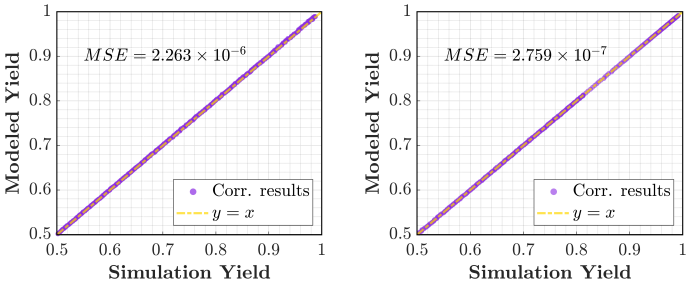
$$Y_{ovl,W2W} = \frac{1}{M} \sum_{j=1}^{M} POS_{ovl,die,j} \tag{8}$$

We vary input parameters for both the model and simulator, including factors such as translation error, rotation error, warpage, die size, and others. The 300 comparison results (purple points) and the mean squared error (MSE) in Fig. 5a demonstrate that our model aligns closely with the simulation data, confirming its reliability and accuracy.

B. Cu Recess Model

We can assume the pad height after the CMP process is normally distributed according to [28], [30]. Let the dielectric surface be zero reference. The height will be negative for the recessed pad and positive for the protruded pad. It is evident that the sum of the top and bottom pad heights also follows a normal distribution. Let h represent the sum of pad heights of two corresponding pads. The mean value of this distribution is denoted by μ_h, and the variance by σ_h^2. It is noticed that h should be clamped within a range of (ζ_-, ζ_+) so that the Cu bonding failure and dielectric delamination can be avoided. Below the calculation of ζ_- and ζ_+ is discussed.

a) Calculation of ζ_-: Based on the observation in [30]–[32], the height difference due to Cu expansion during annealing is linearly correlated to the annealing temperature. The *lower bound* ζ_- of the total Cu heights required to form a qualified Cu bonding area is determined by the sum of Cu expansion after PBA so that the gap between two pads caused by the recess is adequately filled with Cu.



(a) Overlay model. (b) Cu recess model.

Fig. 5: Correlation results of the overlay model and the Cu recess model with the simulation data for W2W HB.

b) Calculation of ζ_+: The critical condition for delamination is that the sum of heights equals ζ_+. It is noticed that surface roughness will decrease the effective contact area of two surfaces. We use the asperity-based roughness model proposed by [19], [33] to calculate the normalized effective contact area $A_b^*(\sigma_z, R_z, E_d, w)$ given roughness and bonding parameters such as the standard deviation of asperity height σ_z, asperity cap radius R_z, Young's modulus of contact surface material E_d, and fully contact bonding energy w [1]. The maximum tolerable peeling stress σ_{tol} to avoid delamination can be given by

$$\sigma_{tol} = A_b^*(\sigma_z, R_z, E_d, w) \times \sqrt{\frac{2E_d w}{t_d}} \tag{9}$$

where t_d denotes the thickness of the surface material [35]. As the ambient temperature changes during the PBA, the thermal expansion mismatch between metal and dielectric will incur various stresses on the bonding interface [27]. Among different interfaces, dielectric-dielectric (e.g. SiO_2-SiO_2) interface is more inclined to delaminate due to relatively lower bonding strength and higher peeling stress at the end of the annealing dwell stage [22], [25]. For simplification, we use a fitting model to evaluate dielectric interface peeling stress dependence by

$$\sigma_{peel} = k_{peel} \cdot D_{Cu} \cdot (h - h_0) \tag{10}$$

where D_{Cu} denotes the Cu pattern density, h_0, k_{peel} are fitting parameters, and k_{peel} is related to the annealing temperature, pad shape, pad arrangement, pad structure, etc. [22], [23], [27]. To prevent delamination, one should have

$$\sigma_{tol} \geq \sigma_{peel} \Rightarrow h \leq h_{peel} \tag{11}$$

Additionally, given the Cu protrusion after CMP will incur delamination, the *upper bound* of the sum of heights is by

$$\zeta_+ = \min\{0, h_{peel}\} \tag{12}$$

To summarize, the POS of this pad during PBA is given by

$$POS_{cr,pad} = \frac{1}{\sqrt{2\pi\sigma_h^2}} \int_{\zeta_-}^{\zeta_+} e^{-\frac{(h-\mu_h)^2}{2\sigma_h^2}} \mathrm{d}h \tag{13}$$

The die yield, as the POS of a die with N pads with respect to the factor of the Cu recess variations is given by

$$Y_{cr,W2W} = POS_{cr,die} = POS_{cr,pad}^N \tag{14}$$

We vary input parameters of Cu recess, pitch, roughness, etc. to validate the Cu recess model. Fig. 5b presents the correlation of the proposed Cu recess model with the simulation data.

C. Defect Model

The relationship between a particle's properties and void size is complex. Due to the significant discrepancy between theoretical and experimental results, it is more practical to develop a fitting model to predict the void size given the particle information. Additionally, for a particle of a specific size and material, it is shown in [3], [38], that the main void size and the void tail length are linearly correlated with the particle's location and the square root of particle thickness. We model

[1]To model the interaction of two rough surfaces, the surface roughness σ_z and Young's modulus E_d need to be normalized [34].

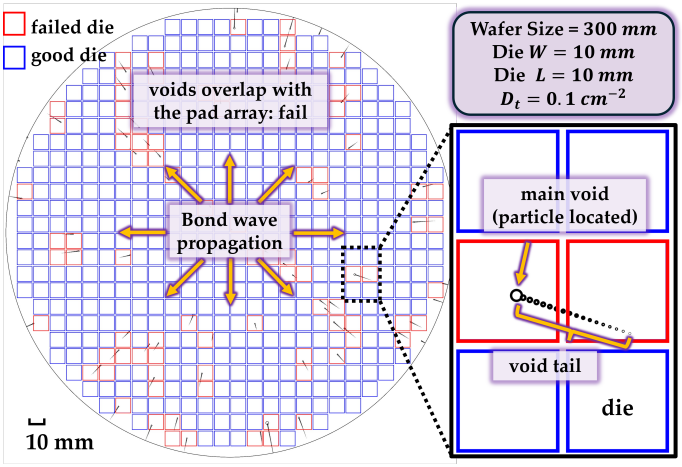


Fig. 6: Visualization of the void formation simulation.

the size r_{mv} of the main void that is located with the distance L $(0 \le L < R)$ from the wafer center by

$$r_{mv} = (k_r L + k_{r_0}) t^{1/2} \tag{15}$$

where t is the particle thickness. Similarly, the void tail length l can be modeled by

$$l = k_l L t^{1/2} \tag{16}$$

where k_r, k_{r_0}, k_l are fitting parameters. Fig. 6 visualizes the simulation of the void formation, which is similar to the scanning acoustic microscopy images of voids in [38]. Since the average void tail length on the wafer can achieve a few millimeters, more than 10 times the scale of the main void size (a few hundred μm in W2W HB, the defect shape can be reasonably simplified as a line as illustrated in Fig. 7. Additionally, a die is considered to have failed if the void tail overlaps the pad array area, as the void size is typically much larger than the pitch (≤ 10 μm).

We assume the thickness distribution of particle defects as $D(t)$. A typical form of $D(t)$ can be [39]

$$D(t) = D_t \cdot \frac{(z-1) \cdot t_0^{z-1}}{t^z}, \qquad t > t_0 \tag{17}$$

where t_0 is the smallest particle thickness, and D_t is the total number of particles of all thicknesses per unit area on the die. The curve shaping parameter z is typically assumed to be between 2 and 3 [40], [41]. By Eq. 16, 17, the distribution of void tail length is given by

$$f_l(l) = \begin{cases} \dfrac{2D_t(z-1)l}{z k_l^2 R^2 t_0}, & l \le k_l R t_0^{1/2} \\ \dfrac{2D_t(z-1)(k_l^2 R^2 t_0)^{z-1}}{z l^{2z-1}}, & l > k_l R t_0^{1/2} \end{cases} \tag{18}$$

where R denotes the radius of the wafer. The comparison of the derived $f_l(l)$ and the simulating distribution is shown in Fig. 8a, confirming the formula's accuracy. The critical area calculation method is presented in Fig. 7. The average critical area regarding the line defect with a length l across all directions, i.e., $0 < \phi < 2\pi$, can be given by

$$A(l) = ab + \frac{2}{\pi}(a+b)l \tag{19}$$

Hence, the average number Λ of particle-induced void tail defects that will fail a die can be given by

$$\Lambda = \int_0^\infty A(l) f_l(l) \mathrm{d}l = D_t ab + \frac{8D_t(z-1)}{3\pi(2z-3)} \cdot (a+b) k_l R t_0^{1/2} \tag{20}$$

Using the Poisson yield model [43], the yield with respect to the particle-induced void formation is given by

$$Y_{df,W2W} = \exp(-\Lambda) \tag{21}$$

We vary input parameters of particle defect density, die size, wafer size, etc. to validate the defect model. Fig. 8b demonstrates the correlation of the defect yield with the simulation results.

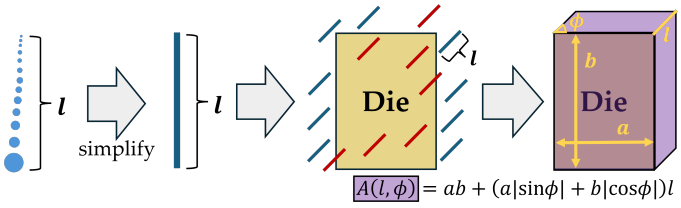
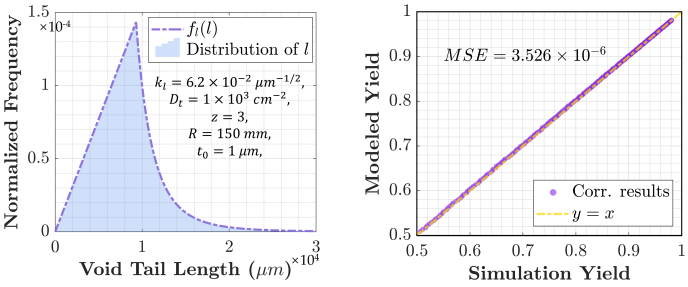


Fig. 7: Critical area calculation. Die size: $a \times b$. Void tail length: l. Void tail direction: ϕ.



(a) Void tail length distribution. (b) Defect model.

Fig. 8: Correlation results of the void tail length distribution and the defect model with the simulation data for W2W HB.

D. Overall Bonding Yield Model

To develop the overall bonding yield model for W2W hybrid bonding, we assume the overlay error, Cu recess variations, and particle defects have independent impacts on the die yield. By combining the Eq. 8, 14, 21, the assembly yield is

$$Y_{W2W} = Y_{ovl,W2W} \cdot Y_{cr,W2W} \cdot Y_{df,W2W} \tag{22}$$

E. D2W Hybrid Bonding Yield Models

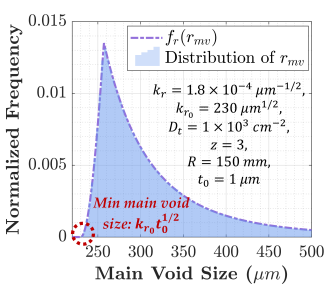
Here we extend the yield model for D2W HB. We assume that the Cu expansion mechanism during PBA in W2W HB remains consistent in the D2W HB case. However, the yield terms regarding the overlay error and particle defects should be revised to account for the differences in bonding approaches.

1) Overlay Model: In D2W hybrid bonding, the systematic overlay error independently happens die-to-die. Given that the die is much smaller than the wafer, the same marker alignment errors at the die edge result in a larger rotation error α and magnification error E than that on average in W2W HB. The overlay yield can be represented by the POS of the die as

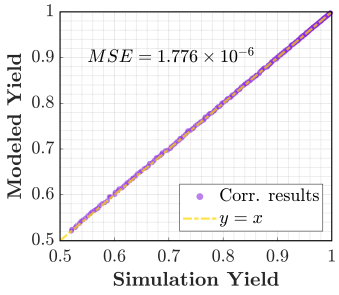
$$Y_{ovl,D2W} = \frac{1}{\sigma_1\sqrt{2\pi}} \min_{i\in[1,N]} \left\{ \int_{-\delta-s_i}^{\delta-s_i} e^{-\frac{u^2}{2\sigma_1^2}} \mathrm{d}u \right\} \tag{23}$$

2) Defect Model: Since the die scale is much smaller than the wafer's, the formation of the void tail is unlikely to occur. Therefore, we only consider the main void-induced failure in the D2W defect model. Combining Eq. 15, 17, the PDF of the main void size r_{mv} can be given by

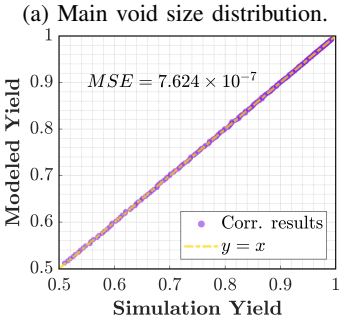
$$f_r(r_{mv}) = \begin{cases} \dfrac{D_t(z-1)t_0^{z-1}}{k_r^2 R^2} \times [\dfrac{2r_{mv}}{z t_0^z} + \dfrac{2k_{r_0}^{2z}}{z(2z-1)r_{mv}^{2z-1}} \\ - \dfrac{2k_{r_0}}{(z-\frac{1}{2})t_0^{z-\frac{1}{2}}}], \quad k_{r_0} t_0^{1/2} < r < (k_r R + k_{r_0}) t_0^{1/2}, \\ \dfrac{2D_t(z-1)t_0^{z-1}(k_r R + k_{r_0})^{2z-2}}{r_{mv}^{2z-1}} - \\ \dfrac{2D_t(z-1)^2 t_0^{z-1}}{k_r^2 R^2 r_{mv}^{2z-1}} \times [\dfrac{(k_r R + k_{r_0})^{2z} - k_{r_0}^{2z}}{z} \\ - \dfrac{2k_{r_0}(k_r R + k_{r_0})^{2z-1} - 2k_{r_0}^{2z}}{z - \frac{1}{2}} + \\ \dfrac{k_{r_0}^2 (k_r R + k_{r_0})^{2z-2} - k_{r_0}^{2z}}{z-1}], \quad r \ge (k_r R + k_{r_0}) t_0^{1/2}. \end{cases} \tag{24}$$



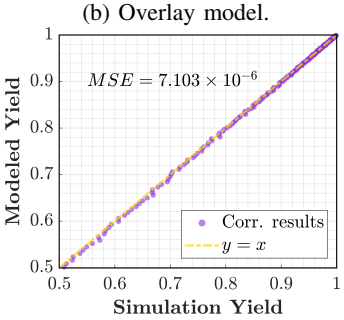
(a) Main void size distribution.



(b) Overlay model.



(c) Cu recess model.



(d) Defect model.

Fig. 9: Correlation results of main void size distribution and the yield model with the simulation data for D2W HB.

where R is the effective radius of the die, i.e., $R = (ab/\pi)^{1/2}$, aiming to remain the average number of particles within the die area. The good alignment between $f_r(r_{mv})$ and the simulated distribution is shown in Fig. 9a. To simplify the critical area calculation, we assume the defect and pad shape are square. The side lengths of voids and top pads are defined as $2r_v, 2r_1$, respectively. The critical area of the main void is given by

$$A(r_v) = \begin{cases} 4N(r_v + r_1)^2, & 2(r_v + r_1) \le p \\ [a + 2(r_v + r_1)] \cdot [b + 2(r_v + r_1)], & 2(r_v + r_1) > p \end{cases} \tag{25}$$

By Eq. 24, 25, the average number Λ of particle-induced main void defects that will fail a die is given by

$$\Lambda = \int_0^\infty A(r_{mv}) f_r(r_{mv}) \mathrm{d}r_{mv} \tag{26}$$

Using the Poisson yield model [43], the yield with respect to the particle-induced void formation for D2W HB is given by

$$Y_{df,D2W} = \exp(-\Lambda) \tag{27}$$

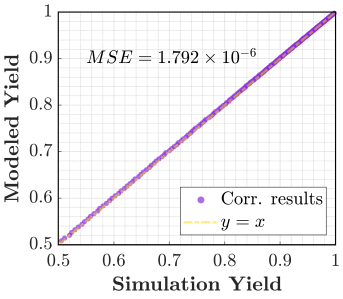
3) Overall Bonding Yield Model: Similarly, we assume the overlay error, Cu recess variations, and particle defects influence the die yield independently for D2W HB. Fig. 9b, 9c, 9d show the correlation results of three yield terms, respectively. By combining Eq. 14, 23, 27, the bonding yield is by

$$Y_{D2W} = Y_{ovl,D2W} \cdot Y_{cr,D2W} \cdot Y_{df,D2W} \tag{28}$$

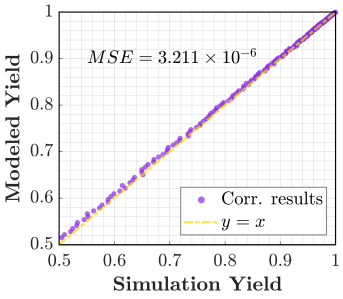
IV. Experimental Results

To validate the derived model, we developed a simulator focusing on Cu-SiO_2 hybrid bonding process. Table I presents the baseline model parameters, with additional details available in our code. These are the parameter values used in our experiments unless otherwise stated.

Fig. 4 outlines the simulation workflow. Overlay error, Cu recess, and particle data are sampled from their respective distributions. Particles are uniformly distributed across the wafer or die, and the void tails are generated based on the fitting model from [38], simulating bond wave propagation (Fig. 6). A die survives only if Cu pillar misalignment is within the safe range (*Overlay Check*), no void overlaps the Cu contact (*Defect Check*), and no dielectric delamination or Cu connection failure occurs (*Cu Recess Check*). The simulation results are closer to the actual conditions with less approximation compared to the model. However, to obtain an accurate yield prediction, simulation requires 1000 wafer samples (20,000 die samples)



(a) Yield of W2W HB.



(b) Yield of D2W HB.

Fig. 10: Correlation of bonding yield with the simulation.

for W2W (D2W) HB, taking 12.2 hours (0.5 hours) on a single CPU (AMD Ryzen 9 8945HS). In contrast, the yield model achieves virtually identical accuracy shown in Fig. 10 in 0.5 s (0.07 s) for W2W HB, offering over *10,000x runtime improvement* and enabling its usage in yield optimization and pathfinding optimization loops.

Below we discuss a few example case studies using the YAP model indicating its strengths in system-technology co-optimization. We vary particle defect density (0.01 cm^{-2}, 0.1 cm^{-2}), pitch (1 μm, 6 μm), and chiplet sizes (10 mm^2, 50 mm^2, 100 mm^2) in the experiment. The yield breakdown and the overall bonding yield are reported in Fig. 11 (W2W setup) and Fig. 12 (D2W setup).

A. Impact of Particle Defect Density

The HB process demands strict elimination of particles at the bonding interface. Fig. 11, 12 show that for the relaxed bonding pitch (6μm), bonding yield is completely defect-limited. Furthermore, W2W HB is more sensitive to particle contamination due to void tail formation during bond wave propagation, resulting in a larger critical area per die. As can be seen from the results, a 10X improvement in defect density (e.g., by following a more stringent cleanroom standard) ensures near-perfect bonding yield for both W2W and D2W for all chiplet sizes.

TABLE I: Baseline parameters in yield modeling and simulation

Parameter	Value
Pad pitch [27]	6 μm
Bottom, Top pad size [27]	3 μm, 2 μm
Die size [47]	10 mm $\times$ 10 mm
Wafer size	300 mm
Random misalignment [46]	5 nm (10 nm)*
System x,y translation [46]	5 nm (10 nm)*
System rotation [46]	0.1 μrad (0.05 μrad)*
Bonded wafer warpage [16]	10 μm (3 μm)*
System magnification [16]	0.9 ppm (0.07 ppm)*
Particle defect density [47]	0.1 cm^{-2}
Minimum particle thickness [38]	1 μm
Shaping factor (z) in Eq. 17 [40], [41]	3
Bottom/Pad recess [22], [28]	10.0 nm (1.0 nm)*
Roughness (σ_z) [20], [45]	1 nm
Adhesion energy (SiO_2-SiO_2) [18], [25]	1.2 J/m^2
Young's modulus (SiO_2) [22], [27]	73 GPa
Dielectric thickness [44]	1.5 μm
contact area constraint k_{ca} in Eq. 6 [11]	0.75
critical distance constraint k_{cd} in Eq. 6 [2]	0.75
k_{mag} in Eq. 2 [16]	0.09 m^{-1}
k_{peel} in Eq. 10 [22]	6.55$\times$ 10^{15} N·m^{-3}
h_0 in Eq. 10 [22]	75 nm
k_r in Eq. 15 [38]	1.8$\times 10^{-4}$ $\mu m^{-1/2}$
k_{r0} in Eq. 15 [38]	230 $\mu m^{1/2}$
k_l in Eq. 16 [38]	6.2$\times 10^{-2}$ $\mu m^{-1/2}$

Note: Distribution parameters with * are shown in Mean (Std.).

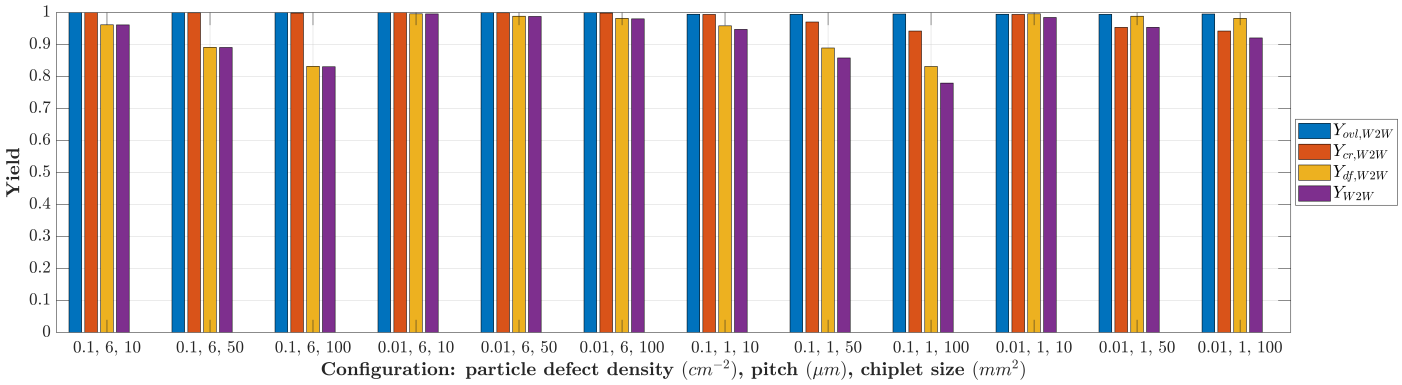


Fig. 11: W2W case studies for various configurations.

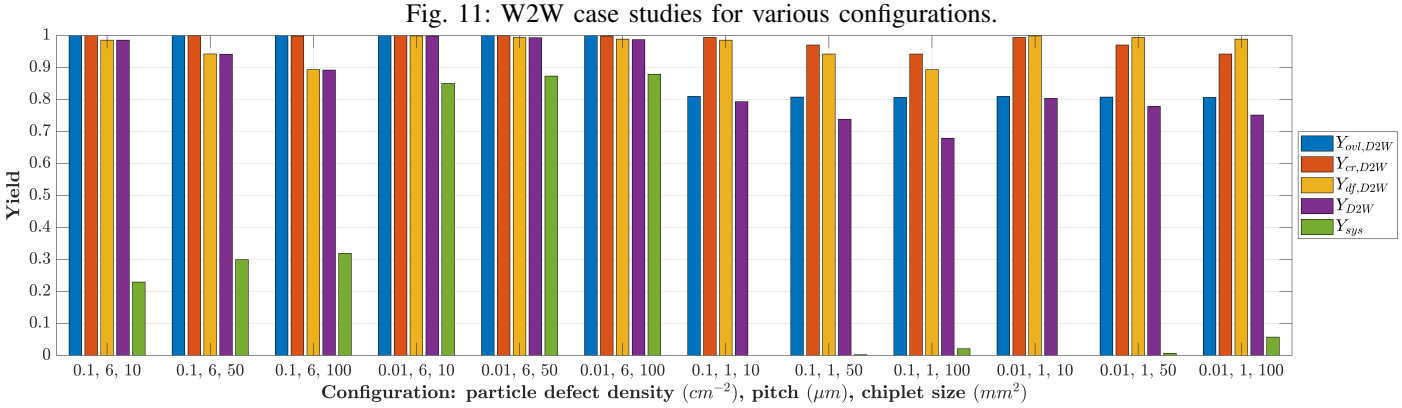


Fig. 12: D2W case studies for various configurations.

B. Impact of Bonding Pitch

In this experiment, the bottom pad size is set to half of the corresponding pitch. As shown in Fig. 11, 12, reducing the pitch from 6 μm to 1 μm leads to a noticeable yield decrease across various chiplet sizes, with the effect being more pronounced in D2W HB. In D2W HB, the primary yield loss stems from increasing overlay errors, as smaller pitches heighten sensitivity to Cu pillar misalignment. For overlay errors, the error α and E resulting from the marker alignment errors scale inversely with the maximum edge distance from the center, which is smaller for a chiplet in D2W HB than for a full wafer. For W2W HB, chiplets closer to the wafer center are more likely to survive, therefore achieving a higher $Y_{ovl,W2W}$. Meanwhile, reducing the pitch significantly increases the number of pads, amplifying the yield's sensitivity to Cu recess variations. For W2W HB increased yield loss is mainly driven by a decrease in Y_{cr} when bonding pitch is decreased. Defect yield, remains largely unaffected, as void size significantly exceeds the pitch, leaving the critical area unchanged.

Another observation is that W2W bonding fares far better than D2W bonding at finer pitches and the difference is even more pronounced at low defect densities. This primarily stems from the ease of high-accuracy alignment for the W2W case.

C. Analyzing Yield Limiters with Varying Chiplet Size

The bonding yield drops with increasing chiplet size for both D2W and W2W bonding, primarily driven by worsened copper recess yield (due to more number of pads per die) and defect yield.

D2W HB can be employed in 2.5D integration to assemble large chiplet systems. Therefore, just looking at a single chiplet yield can be misleading. Though our focus is not full system assembly yield modeling, we add a system yield (Y_{sys} in Fig. 12) calculated simply as $Y_{D2W}^{\#chiplets}$, in the absence of any redundancy, for a nominal system size of 1000 mm^2. This approach reflects the cumulative probability of all chiplets bonding successfully in the system. Increasing the chiplet size, which reduces the number of chiplets required for system assembly, helps mitigate the compounding effect of Y_{D2W} losses. Interestingly, even though Y_{D2W} decreases as the chiplet size increases, Y_{sys} remains higher.[2]

V. Conclusion

This work presents YAP: a comprehensive yield modeling framework for W2W and D2W hybrid bonding in advanced packaging. YAP models overlay errors, particle-induced void defects, and Cu recess variations and is validated against a physics-inspired yield simulator[3]. The proposed YAP yield model accurately predicts bonding yield across various chiplet sizes, pitches, and pad configurations and achieves a 10,000-fold speed improvement over direct simulations while maintaining negligibly small mean square error. YAP enables fast system-technology co-optimization. Case studies using it highlight key differences between W2W and D2W HB, offering insights into yield-limiting failure mechanisms motivating future technology development and system architecture choices.

Our future work directions include incorporating YAP into a comprehensive system assembly yield model; extending YAP to model other forms of fine-pitch bonding such as thermal-compression bonding; evaluating some of the approximations made in YAP in the context of ultra-fine pitch bonding processes; and developing fault tolerance and yield improvement techniques leveraging YAP.

Acknowledgement

The authors thank Alexander Graening and Krutikesh Sahoo at UCLA and Nicolas Pantano at IMEC for their inputs during the early stages of the work. This work was supported in part by CHIMES, one of the seven centers in JUMP 2.0, a Semiconductor Research Corporation (SRC) program.

References

[1] Elsherbini, A., Jun, K., Vreeland, R., Brezinski, W., Niazi, H., Shi, Y., Yu, Q., Qian, Z., Xu, J., Liff, S. & Others Enabling hybrid bonding on Intel process. *2021 IEEE International Electron Devices Meeting (IEDM)*. pp. 34-3 (2021)

[2]Note that this does not account for any worsened yield of a larger chiplet. A more complete system yield model can be found, for example, in [10] albeit with an oversimplified bonding yield model.

[3]The code of the yield model and simulator is available open-source at `https://github.com/XXX/YAP`.

[2] Ikegami, Y., Onodera, T., Chiyozono, M., Sakamoto, A., Shimizu, K., Kagawa, Y. & Iwamoto, H. Study of Ultra-Fine 0.4 μm Pitch Wafer-to-Wafer Hybrid Bonding and Impact of Bonding Misalignment. *2024 IEEE 74th Electronic Components And Technology Conference (ECTC)*. pp. 299-304 (2024)

[3] Lau, J. Recent Advances and Trends in Cu–Cu Hybrid Bonding. *IEEE Transactions On Components, Packaging And Manufacturing Technology*. **13**, 399-425 (2023)

[4] Lee, S., Jee, Y., Park, S., Lee, S., Hwang, B., Jo, G., Lee, C., Park, J., Jang, A., Jung, H. & Others A study on memory stack process by hybrid copper bonding (HCB) technology. *2022 IEEE 72nd Electronic Components And Technology Conference (ECTC)*. pp. 1085-1089 (2022)

[5] Lujan, A. Cost and Yield Analysis of Die-to-Wafer Hybrid Bonding. *2022 International Conference On Electronics Packaging (ICEP)*. pp. 129-130 (2022)

[6] Marinissen, E., Pancholi, V., Chuang, P. & Keim, M. IEEE Std P3405: New Standard-under-Development for Chiplet Interconnect Test and Repair. *2024 IEEE 42nd VLSI Test Symposium (VTS)*. pp. 1-11 (2024)

[7] Singh, E. Analytical modeling of 3D stacked IC yield from wafer to wafer stacking with radial defect clustering. *2014 27th International Conference On VLSI Design*. pp. 26-31 (2014)

[8] Xu, Q., Jiang, L., Li, H. & Eklow, B. Yield enhancement for 3D-stacked ICs: Recent advances and challenges. *17th Asia And South Pacific Design Automation Conference*. pp. 731-737 (2012)

[9] Campbell, D. Yield modeling of 3D integrated wafer scale assemblies. *2010 Proceedings 60th Electronic Components And Technology Conference (ECTC)*. pp. 1935-1938 (2010)

[10] Graening, A., Pal, S. & Gupta, P. Chiplets: How small is too small?. *2023 60th ACM/IEEE Design Automation Conference (DAC)*. pp. 1-6 (2023)

[11] Zhang, B., Chew, S., Stucchi, M., Dewilde, S., Iacovo, S., Witters, L., Webers, T., Van Sever, K., De Vos, J., Miller, A., Beyer, G. & Beyne, E. Scaling Cu/SiCN Wafer-to-Wafer Hybrid Bonding down to 400 nm interconnect pitch. *2024 IEEE 74th Electronic Components And Technology Conference (ECTC)*. pp. 312-318 (2024)

[12] Moreau, S., Bouchu, D., Jourdon, J., Ayoub, B., Lhostis, S., Frémont, H. & Lamontagne, P. Recent Advances on Electromigration in Cu/SiO_2 to Cu/SiO_2 Hybrid Bonds for 3D Integrated Circuits. *2023 IEEE International Reliability Physics Symposium (IRPS)*. pp. 1-7 (2023)

[13] Ghaida, R., Gupta, M. & Gupta, P. Framework for exploring the interaction between design rules and overlay control. *Journal Of Micro/Nanolithography, MEMS, And MOEMS*. **12**, 033014-033014 (2013)

[14] Armitage Jr, J. & Kirk, J. Analysis of overlay distortion patterns. *Integrated Circuit Metrology, Inspection, And Process Control II*. **921** pp. 207-223 (1988)

[15] Okudur, O., Iacovo, S., Kang, S., Gonzalez, M. & Beyne, E. Simulations of Wafer-to-Wafer Bonding Dynamics and Deformation Mechanisms. *2024 IEEE 10th Electronics System-Integration Technology Conference (ESTC)*. pp. 1-5 (2024)

[16] Kang, S., Iacovo, S., D'havé, K., Van Huylenbroeck, S., Okudur, O., Alexeev, A., Plach, T., Probst, G., Ding, T., Wimplinger, M. & Others Investigation of Distortion in Wafer-to-wafer Bonding with Highly Bowed Wafers. *2024 IEEE 74th Electronic Components And Technology Conference (ECTC)*. pp. 386-393 (2024)

[17] Ren, H. Process Development and Process Window Investigation of Copper-Silicon Dioxide Die-to-Wafer (D2W) Hybrid Bonding. (University of California, Los Angeles,2021)

[18] Chidambaram, V., Leong, Y. & Ren, Q. Wafer Level Fine-Pitch Hybrid Bonding: Challenges and Remedies. *2020 IEEE 22nd Electronics Packaging Technology Conference (EPTC)*. pp. 459-463 (2020)

[19] Gui, C., Elwenspoek, M., Tas, N. & Gardeniers, J. The effect of surface roughness on direct wafer bonding. *Journal Of Applied Physics*. **85**, 7448-7454 (1999)

[20] Dubey, V., Wünsch, D., Gottfried, K., Fischer, T., Helke, C., Hasse, M., Hanisch, A., Hofmann, L., Reuter, D., Wiemer, M. & Others Impact of Dielectric Types on Surface Topography for Wafer-Level Hybrid Bonding. *2024 IEEE 10th Electronics System-Integration Technology Conference (ESTC)*. pp. 1-5 (2024)

[21] Ji, L., Che, F., Ji, H., Li, H. & Kawano, M. Modelling and characterization on wafer to wafer hybrid bonding technology for 3D IC packaging. *2019 IEEE 21st Electronics Packaging Technology Conference (EPTC)*. pp. 87-94 (2019)

[22] Ji, L., Che, F., Ji, H., Li, H. & Kawano, M. Wafer-to-wafer hybrid bonding development by advanced finite element modeling for 3-D IC packages. *IEEE Transactions On Components, Packaging And Manufacturing Technology*. **10**, 2106-2117 (2020)

[23] Wang, H., Chen, H., Xiang, J. & Yang, X. Research on simulation of Cu/SiO_2 hybrid bonding process and interface failure mechanism by Finite Element Analysis. *2023 24th International Conference On Electronic Packaging Technology (ICEPT)*. pp. 1-7 (2023)

[24] Beilliard, Y., Estevez, R., Parry, G., McGarry, P., Di Cioccio, L. & Coudrain, P. Thermomechanical finite element modeling of Cu–SiO_2 direct hybrid bonding with a dishing effect on Cu surfaces. *International Journal Of Solids And Structures*. **117** pp. 208-220 (2017)

[25] Fujii, N., Furuse, S., Yoshioka, H., Ogawa, N., Yamada, T., Hirano, T., Saito, S., Hagimoto, Y. & Iwamoto, H. Bonding Strength of Cu-Cu Hybrid Bonding for 3D Integration Process. *ECS Transactions*. **112**, 3 (2023)

[26] Le, X. & Choa, S. Assessment of the Risk of Crack Formation at a Hybrid Bonding Interface Using Numerical Analysis. *Micromachines*. **15**, 1332 (2024)

[27] Zhao, G., Zeng, Y. & Zhao, Y. Simulation and Experimental Analysis of Thermomechanical Stress Around Interconnects for W2W Hybrid Bonding. *2024 IEEE 10th Electronics System-Integration Technology Conference (ESTC)*. pp. 1-6 (2024)

[28] Kim, S., Fodor, F., Heylen, N., Iacovo, S., De Vos, J., Miller, A., Beyer, G. & Beyne, E. Novel Cu/SiCN surface topography control for 1 μm pitch hybrid wafer-to-wafer bonding. *2020 IEEE 70th Electronic Components And Technology Conference (ECTC)*. pp. 216-222 (2020)

[29] Beyne, E., Kim, S., Peng, L., Heylen, N., De Messemaeker, J., Okudur, O., Phommahaxay, A., Kim, T., Stucchi, M., Velenis, D. & Others Scalable, sub 2μm pitch, Cu/SiCN to Cu/SiCN hybrid wafer-to-wafer bonding technology. *2017 IEEE International Electron Devices Meeting (IEDM)*. pp. 32-4 (2017)

[30] De Messemaeker, J., Witters, L., Zhang, B., Tsau, Y., Fodor, F., De Vos, J., Beyer, G., Croes, K. & Beyne, E. New Cu "Bulge-Out" Mechanism Supporting SubMicron Scaling of Hybrid Wafer-to-Wafer Bonding. *2023 IEEE 73rd Electronic Components And Technology Conference (ECTC)*. pp. 109-113 (2023)

[31] Lin, H., Tran, D., Chiu, W., Chang, H. & Chen, C. In-situ measurement of thermal expansion in Cu/SiO_2 hybrid structures using atomic force microscopy at elevated temperatures. *Applied Surface Science*. **662** pp. 160103 (2024)

[32] Lin, H., Chiu, W., Chang, H. & Chen, C. Observation of Thermal Expansion Behavior of Nanotwinned-Cu/SiO_2 & Regular-Cu/SiO_2 Hybrid Structure via In-Situ Heating AFM. *2024 IEEE 74th Electronic Components And Technology Conference (ECTC)*. pp. 816-820 (2024)

[33] Maugis, D. On the contact and adhesion of rough surfaces. *Journal Of Adhesion Science And Technology*. **10**, 161-175 (1996)

[34] Rieutord, F., Moriceau, H., Beneyton, R., Capello, L., Morales, C. & Charvet, A. Rough surface adhesion mechanisms for wafer bonding. *ECS Transactions*. **3**, 205 (2006)

[35] Hutchinson, J. & Suo, Z. Mixed mode cracking in layered materials. *Advances In Applied Mechanics*. **29** pp. 143-145 (1991)

[36] Nagano, F., Inoue, F., Phommahaxay, A., Peng, L., Chancerel, F., Naser, H., Beyer, G., Uedono, A., Beyne, E., De Gendt, S. & Others Origin of Voids at the SiO_2/SiO_2 and SiCN/SiCN Bonding Interface Using Positron Annihilation Spectroscopy and Electron Spin Resonance. *Ecs Journal Of Solid State Science And Technology*. **12** (2023)

[37] Kim, Y., Nguyen, T. & Choa, S. Enhancement of the bond strength and reduction of wafer edge voids in hybrid bonding. *Micromachines*. **13**, 537 (2022)

[38] Nagano, F., Iacovo, S., Phommahaxay, A., Inoue, F., Chancerel, F., Naser, H., Beyer, G., Beyne, E. & Gendt, S. Void formation mechanism related to particles during wafer-to-wafer direct bonding. *ECS Journal Of Solid State Science And Technology*. **11**, 063012 (2022)

[39] Glang, R. Defect size distribution in VLSI chips. *IEEE Transactions On Semiconductor Manufacturing*. **4**, 265-269 (1991)

[40] Stapper, C. Modeling of defects in integrated circuit photolithographic patterns. *IBM Journal Of Research And Development*. **28**, 461-475 (1984)

[41] Bruls, E. & Others Characterization of defects in integrated circuits: resources, models and applications. (1992)

[42] Lu, T., Hsu, K., Hsu, C., Hsu, C. & Wu, Y. Effect of Cu Film Thickness on Cu Bonding Quality and Bonding Mechanism. *Materials*. **17**, 2150 (2024)

[43] Koren, I. & Koren, Z. Defect tolerance in VLSI circuits: techniques and yield analysis. *Proceedings Of The IEEE*. **86**, 1819-1838 (1998)

[44] Chidambaram, V., Lianto, P., Wang, X., See, G., Wiswell, N. & Kawano, M. Dielectric materials characterization for hybrid bonding. *2021 IEEE 71st Electronic Components And Technology Conference (ECTC)*. pp. 426-431 (2021)

[45] Dubey, V., Wünsch, D., Gottfried, K., Wiemer, M., Fischer, T., Schermer, S., Dittmar, N., Helke, C., Haase, M., Ghosal, S., Hanisch, A., Bonitz, J., Luo-Hofmann, J., Hofmann, L., Lykova, M., Stoll, F., Vogel, K. & Schulz, S. Impact of Dielectric and Copper Via Design on Wafer-to-Wafer Hybrid Bonding. *2023 IEEE 73rd Electronic Components And Technology Conference (ECTC)*. pp. 795-799 (2023)

[46] Mitsuishi, H., Mori, H., Maeda, H., Ushijima, M., Kamashita, A., Okada, M., Aramata, M., Shiomi, T., Sakamoto, S., Takahata, K. & Others 50 nm overlay accuracy for wafer-to-wafer bonding by high-precision alignment technologies. *2023 IEEE 73rd Electronic Components And Technology Conference (ECTC)*. pp. 1664-1671 (2023)

[47] Patel, D., Xie, M., & Koch, J. Hybrid Bonding Process Flow – Advanced Packaging Part 5. https://semianalysis.com/2024/02/09/hybrid-bonding-process-flow-advanced/ (2024)