Characterization of Extreme Si Thinning Process for Wafer-to-Wafer Stacking

Fumihiro Inoue[1], Anne Jourdain[1], Joeri De Vos[1], Erik Sleeckx[1], Eric Beyne[1]
Jash Patel[2], Oliver Ansell[2], Huma Ashraf[2], Janet Hopkins[2], Dave Thomas[2]
Akira Uedono[3]
[1] imec
Kapeldreef 75, B-3001 Leuven, Belgium
Fumihiro.Inoue@imec.be

[2] SPTS Technologies, Ringland Way, Newport, NP18 2TA, UK
[3] Tsukuba University, Tsukuba, Ibaraki, 305-8573, Japan

Abstract— **Wafer-to-wafer 3D integration has a potential to minimize the Si thickness, which enables us to connect multiple wafers with significantly scaled through-Si vias. In order to achieve this type of 3D structure, backside thinning is a key step. Conventional mechanical grinding is known as the best way to remove bulk Si in terms of cost of ownership (CoO). However, mechanical damage such as induced dislocations needs to be removed after extreme thinning to avoid a serious impact on the device performance. CMP shows the best performance in terms of roughness with a significantly flat surface with only atomic step roughness. Furthermore, the existing mono-vacancies are as low as for a bulk Si substrate. However the total thickness variation (TTV) worsens as more Si is removed. The dry etch process enables a faster etch rate than CMP and wet etching. Furthermore, the mono-vacancy/damage layer after dry etching is equivalent to that achieved when combined with CMP. The combination of CMP and dry etch enables us to achieve extreme thinning of active device wafers (<5 μm) with minimal roughness, no damage layer (mono-vacancy) and no edge delamination.**

Keywords-component; Edge-trim, Grinding, CMP, Dry etch, wafer-to-wafer bonding,

I. INTRODUCTION

Wafer-to-wafer stacking potentially offers a solution for thin wafer handling issues allowing the removal of most of the Si from the top wafer. The extremely thin Si (< 5 μm) can extend the scaling of via-last TSVs, due to the minimized pitch and dimension (pitch < 2 μm, dimension 1 x 5 μm). The via-last TSV formation is significantly influenced by the thinning performance, which brings new challenges in the thinning process [1]. A smooth surface and less damage into the Si are mandatory to achieve a successful backside TSV process. Accurate thickness control and minimal total thickness variation (TTV) across the whole wafer are required for the via etch process [2]. In particular, TTV is much more critical for the range of thin Si. In addition, any mechanical failures occurring at the very wafer edge, an area of the wafer typically excluded from inspection, may also impact multiple wafer-to-wafer bonding. Furthermore, vacancy-type of atomic defects need to be analyzed using depth profilinge for the extreme thinning, since the remaining Si is limited. Thus, a deeper understanding of micro and macro surface characterization is essential for extreme thinning.

In this paper the characterization of different extreme Si wafer thinning processes are discussed on permanently bonded wafers. Several thinning methods, such as CMP and plasma dry etch are compared as subsequent processes of grinding. The wafer scale topography such as TTV, mean thickness control and wafer bowing will be discussed. The wafer inspections were performed up to the wafer bevel. In addition, nano-scale characteristics, such as roughness, mono-vacancies and damage in/to Si will be investigated by using alternative ways of surface detection.

II. EXPERIMENTAL

The thinning processes were performed on the top wafer of permanently bonded wafer stacks. The permanent wafer bonding was performed on 300 mm wafers. Oxide CMP is used to planarize and to smooth the oxide layer. Prior to bonding, the pairing wafers are treated with a combination of plasma activation and a DI water clean. Wafer bonding takes place at room temperature with atmospheric pressure in a clean room ambient. Then an annealing cure was used to improve the adhesion strength. The edge-trim process was performed using a dicing saw tool DFD6860 from Disco. The grinding step was carried out on a Disco DGF8560 series in-feed grinder which consists of rough and fine grinding wheels. The wafer nano-tomography was measured by a wafer automated inspection system. The wafer edge inspections were done in a KLA-tencor CIRCL-AP.

The dry etching for Si thinning was performed by an SPTS Rapier XE system with Near Infra-Red (NIR) interferometer in-situ endpoint detection.

III. RESULTS AND DISCUSSION

A. *Wafer bonding and Edge-trim*

Figure 1 shows CIRCL-AP shadow images taken at the wafer edge, before and after grinding for several edge-trim approaches. As shown in the figure, the wafer edge shape after grinding is different depending on the edge-trim approach used.

978-1-5090-1204-6/16 $31.00 © 2016 IEEE
DOI 10.1109/ECTC.2016.102

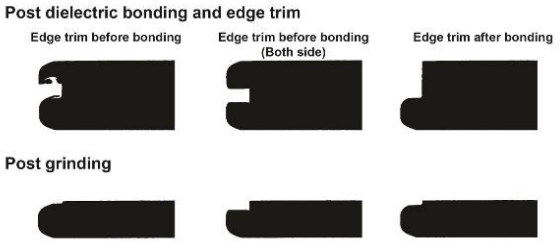

Figure 1. Wafer edge shadow images taken after edge-trim and grinding for different edge-trim approaches

A comparison of the damage into the Si was made between edge-trim before and after bonding. In order to compare the edge-trim impact cross-sectional TEM images were taken at the wafer edge after grinding. As the edge-trim depth after bonding is much deeper than for edge-trim before bonding, a blade with bigger diamonds is suitable for edge-trim after bonding. Figure 2 shows TEM images of dielectric bonded wafers after grinding using edge-trim before bonding (a)-(b) and after bonding (d)-(e). Dislocations and stacking faults in the {111} plane can be seen in the damaged regions in both samples. The presence of acontinuous amorphous Si layer was also observed at the wafer edge. The Si sidewall using edge-trim after bonding has a higher roughness, deeper defects (~500 nm) and thicker and continuous amorphous Si layer. This might be due to the transfer of a significantly big indentation and heat energy caused by the diamond abrasive process. The amorphous Si and different phases of Si are also detectable by using micro-Raman spectroscopy. Figure 2 (c) and (f) shows Raman spectra taken from the sidewall of the edge-trimmed Si. A lot of peaks can be distinguished for the case of an edge-trimmed sidewall. This indicates the temporary occurrence of significant high local stresses during blade dicing [3,4]. This does not create an immediate integration failure but these defects might represent the initiation of a fragile wafer edge.

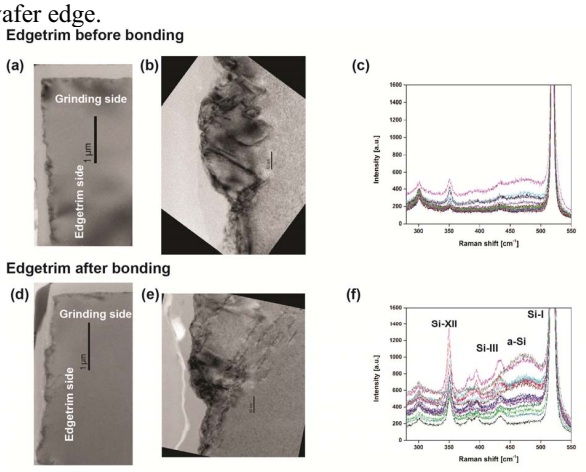

Figure 2. Edge-trim before bonding (a) X-TEM (b) high magnification (c) Raman spectra, and Edge-trim after bonding (d) X-TEM (e) high magnification (f) Raman spectra

B. *Grinding*

Figure 3 shows a whole wafer bright field optical image taken after grinding to 50 µm thickness of the top wafer. After fine grinding, grind marks which are caused by diamond wheel scratches, are distinguishable (see Fig.3 (b)).

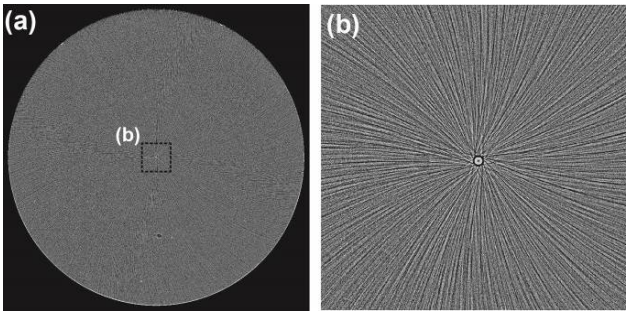

Figure 3. Bright field optical image taken after grinding (a) whole wafer (b) high magnification at wafer center.

Figure 4 shows an AFM image taken around the wafer center after fine grinding. A lot of randomly located lines are visible. The maximum step height (Z range) is around 100 nm, which is not acceptable for subsequent TSV via-last processes.

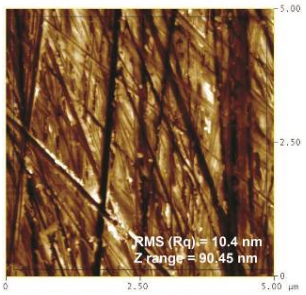

Figure 4. AFM image after fine grinding at the wafer center.

Figure 5 shows cross-sectional TEM images taken at the wafer center and the wafer edge. When the roughness was compared at the wafer center (a) and very edge of the wafer (c), (d), the center has a deeper step height than the wafer edge. In addition, deeper damage (micro-cracks) into the Si was observed underneath the grinding surface (see Fig. 3 (f)). The maximum length of the micro-cracks which can be observed in the TEM specimen was more than 200 nm. At the wafer edge, the micro-cracks were slightly smaller than in the center. However, edge-trim after bonding has slightly deeper micro-cracks at the grinding side compared to using edge-trim before bonding.

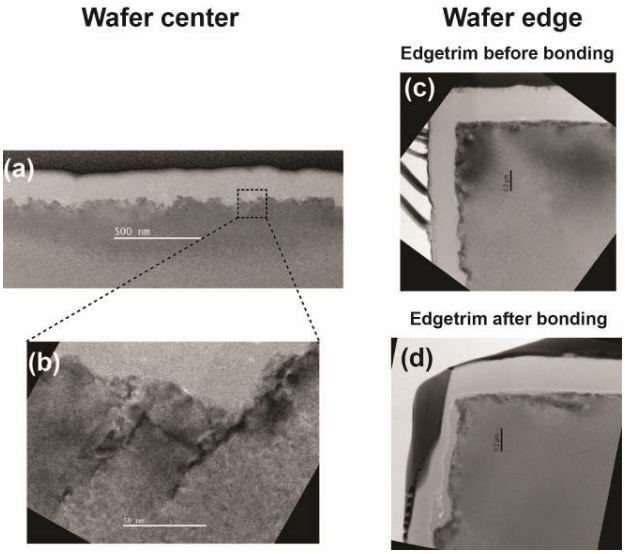

Figure 5. Cross-sectional TEM images after grinding (a) wafer center (b) high magnification (c) wafer edge for edge-trim before bonding (d) waferr edge for edge-trim after bonding

Figure 6 shows depth profiles of stress after rough and fine grinding of the top wafer measured by Raman spectroscopy, collected on the cross-section plane. The rough grinding damage on the surface created a large stress inside the Si. The detected stress on the top Si is compressive for both rough and fine grinding. It disappears over about 25 µm for both rough and fine grinding. In other words, this is the limitation of the final thickness of grinding so as to avoid any impact on the top device area and the bonding interface.

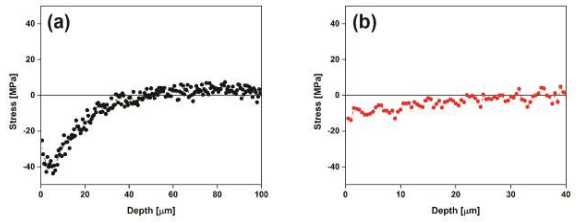

Figure 6. Depth profiles after grinding of top wafer (a) after rough grinding with 120 µm Si thickness (b) after fine grinding with 50 µm Si thickness

C. *Grinding + CMP*

Although the removal rate of CMP for Si is much lower compared to grinding, CMP is known to be an effective stress relief process [5]. Figure 7 shows wafer bright field optical images for different CMP removal amounts after grinding ((a) 0.2, (c) 0.5 and (e) 1 µm, respectively). For the case of 0.2 µm removal, the grinding marks are still distinguishable. It indicates that there is some grinding damage deeper than 200 nm. On the other hand, no grinding marks are visible after 500 nm removal at the wafer center. This indicates that the detectable grinding damage is in the range of 200 to 500 nm. This result corresponds to the dislocation depth which was detected by TEM (Fig 5, maximum more than 200 nm).

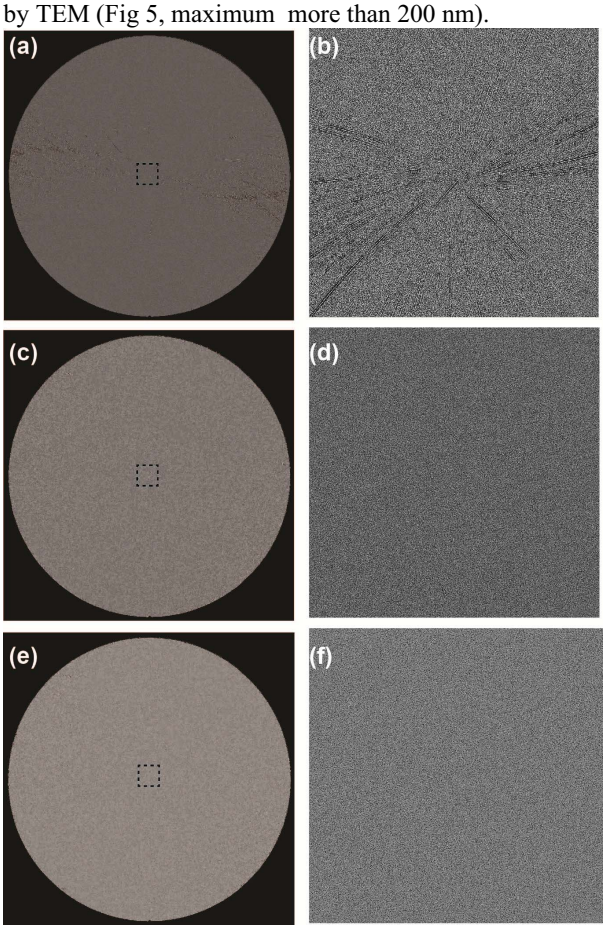

Figure 7. Bright field optical images taken after grinding and CMP for a whole wafer and high magnification at wafer center. (a), (b) 0.2 µm (c), (d) 0.5 µm, (e), (f) 1 µm

Figure 8 (a) shows the Si surface after grinding + 1 µm CMP. The surface roughness which can be seen after grinding has disappeared. The RMS was 0.1 nm, and the Z range was 2.97 nm. These results are almost the same as a bare Si wafer.

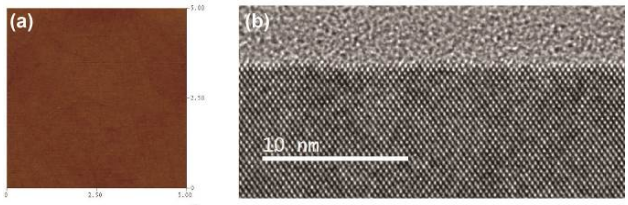

Figure 8. (a) AFM image and (b) cross-sectional TEM image after 1 µm CMP

Figure 8 (b) shows a cross-sectional TEM image taken after grinding + 1 µm CMP. Only atomic roughness steps are visible in the specimen. No dislocations or strain/defect areas are observed on the surface.

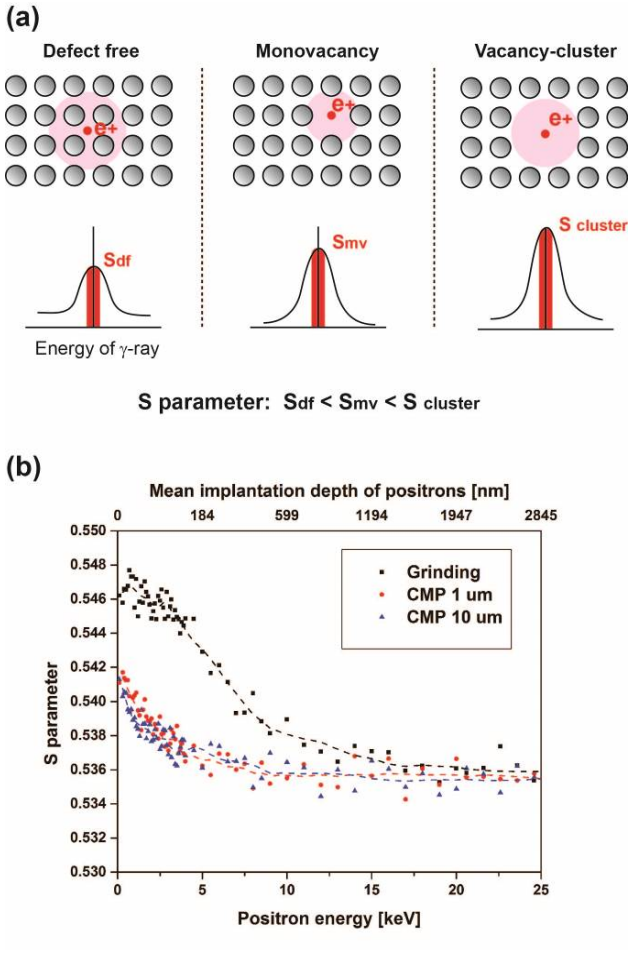

Figure 9. (a) Schematic images of positron annihilation in Si with different types of Si mono-vacancy (b) S parameters as a function of incident positron energy E for the Si after grinding and CMP.

In order to evaluate the damage removal at an atomic scale, the samples were analyzed by positron annihilation spectroscopy (PAS). Figure 9 (a) shows a schematic drawing of a PAS measurement into Si. Positron annihilation is a technique for detecting vacancy-type defects [6]. When a positron is implanted into Si samples, it annihilates with an electron and emits two 511-keV γ–quanta. The energy distribution of the annihilated γ-rays is broadened by the momentum component of the annihilating electron-positron pair. It is parallel to the emitting direction of the γ-rays. A diffusing positron is detectable in a vacancy-type defect because of Coulomb repulsion from the existing ion cores. Due to the momentum distribution of the electrons in such defects differs from that of electrons in the bulk material, these defects can be detected by measuring the Doppler broadening spectra. The resulting changes in the spectra are characterized by the S parameter, which mainly reflects changes due to the annihilation of positron-electron pairs with a low-momentum distribution.

Figure 9 (b) shows the PAS measurement results on grinding, 1 µm CMP and 10 µm CMP. The S parameter on the defect free Si surface is 0.536. For the case of a ground surface, a higher S parameter < 4 keV was detected (~ 150 nm). This is due to the trapping of positrons by vacancy-type defects. In other words, the surface has some vacancy clusters. The maximum dislocation detected in the TEM image in Fig 5 was 200 nm. The TEM results and the PAS data show some correspondence. When a comparison was made on a 1 µm and a 10 µm CMP sample surface no differences were seen. In general, the surface S value is sensitive to the surface condition of the sample. The small surface S parameter value could be due to the electric field near the surface. Therefore, it is indicated that CMP has an excellent ability to remove the grinding damage (dislocations, amorphous-Si and mono-vacancies) with an atomic scale of surface roughness.

Figure 10 shows the evolution of the Si bow after each process step from dielectric CMP to Si CMP. Before bonding (a single wafer), the wafer already has a concave bow due to the film stress of the dielectric adhesive layer. After fusion bonding of the 2 wafers (face to face) and annealing, the bow value increases. Nevertheless, after grinding (total stack thickness 825 µm, top wafer 50 µm), the bow decreases to almost zero. In addition, the bow became + 25 µm after grinding to 25 µm of top wafer Si thickness. This is due to the grinding damage which creates a compressively stressed Si layer [7]. Although this layer is very thin, its high stress value can compensate the tensile stress from the bonding interface. After 1 µm CMP (49 µm Si thickness), the wafer bow is back to concave. This is because the grinding damage was removed by CMP, and the bonding interface (bonded dielectric layer) determines the majority of the origin of the stress.

	Before bonding	Post bonding	Post anneal	Post grinding	Post CMP
Bow - = concave + = convex	- 109 µm	- 114 µm	- 131 µm	1 µm	- 40 µm
Schematic image of wafer shape					

Figure 10. Evolution of the Si bow for permanent bonding to top wafer CMP steps.

Figure 11 shows the depth profile of stress after grinding and CMP measured by Raman spectroscopy, collected on the cross-sectional plane. As discussed in Figure 6, grinding damage creates compressive stress around the

surface. On the other hand, the stress on the surface becomes tensile after CMP. The results show a clear correlation with the bow data in Fig 10. From these results the macro stress relief is accomplished by the CMP of 1 µm Silicon.

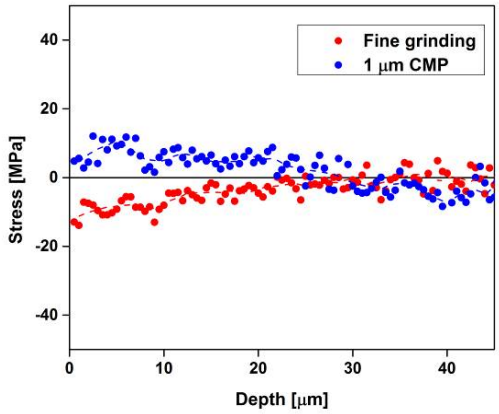

Figure 11. Depth profile of stress after grinding and CMP on the top wafer after thinning.

Figure 12 shows thickness measurement results after grinding (to 50 µm) and 1 µm CMP (to 49 µm). In the radial distribution plot in Fig 12 (a), the wafer shape is as good as for a bulk Si bare wafer. The mean thickness is 49.05 µm and 99.96% of the surface is within a TTV of 1 µm (see Fig 12 (b)).

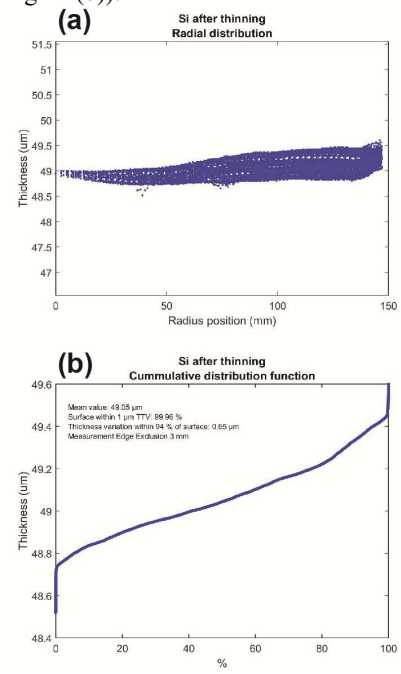

Figure 12. Thickness radial distribustion and the weibull distribution plot after grinding and 1 µm CMP

Figure 13 shows the surface roughness and the ΔTTV after grinding + CMP for different removal amounts by CMP. The ΔTTV gets larger when increasing the removal amount. For the case of more than 20 µm CMP removal , the TTV increased to 2.3 µm (see also Fig 15 (c)). In order to maintain the TTV after grinding + CMP, the CMP removal amount needs to be as small as possible (e.g. 1 µm). Neverthelss, as shown in Figure 6, the grinding stress is propagating towards a depth of 25 µm. In principal, to achieve extreme thinning by grinding and CMP without stress impact to the device, the grinding must stop at least above 25 µm, and the last ~ 20 µm needs to be removed by CMP.

Thus, although grinding + CMP is a suitable process sequence to obtain a nano/atomic scale flat surface without mono-vacancies, there are a lot of challenges for the CMP process to achieve extreme thinning with good TTV and no mechanical impact to the device. Furthermore, the slow removal rate of the CMP process might be another challenge to apply CMP for extreme thinning in terms of CoO control.

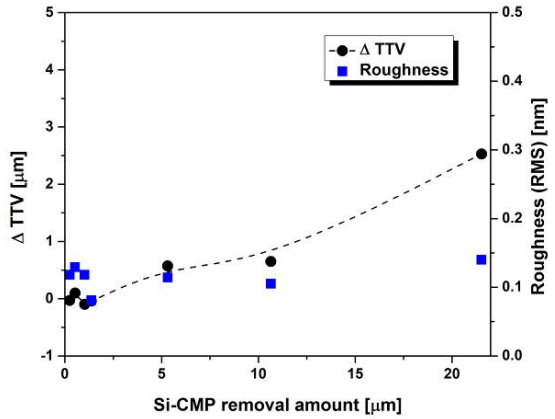

Figure 13. The delta TTV after grinding + CMP for different removal amounts of CMP with the corresponding roughness measured by AFM.

D. *Grinding + Dry etch*

Plasma dry etching of Si is a well-established process for device manufacturing. However, its application for backside Si thinning, in particular for extreme thinning, has not been well investigated and a detailed surface characterization is needed.

The Si thinning by dry etching was done in a Versalis *f*xP integrated system fitted with a Rapier XE etch chamber from SPTS Technologies. This etching module is designed for Si thinning with a high etch rate and good TTV for 300 mm wafers [8]. Furthermore, the in-situ NIR

end-point detection enables the user to to target the desired final Si thickness.

Figure 14 shows a cross-sectional TEM image taken after grinding and dry etching. No grinding induced damage is observed in this TEM image, which indicates that most of the surface damage was removed by the dry etch. In addition, the wafer bow was – 40.2 µm, which is similar to CMP after grinding.

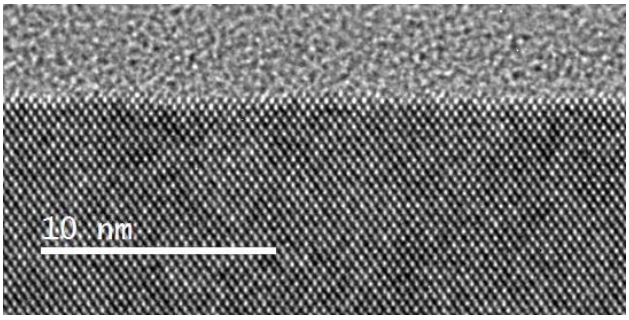

Figure 14. Cross-sectional TEM image after grinding and dry etch.

Figure 15 shows the thickness radial distribution of a 300 mm wafer after (a) 26 µm Si removal by dry etch and (c) 22 µm removal by CMP after grinding. For the case of the dry etch, 99.5 % was kept within 2 µm TTV even after more than 20 µm of Si removal. On the other hand, it was 87.1 % in the case of CMP. The thickness at the wafer edge drastically drops in the case of 22 µm of CMP. Not only is there a benefit of wafer shape control for dry etching but the high etching rate of 9.2 µm/min is approximately 9 times faster than CMP.

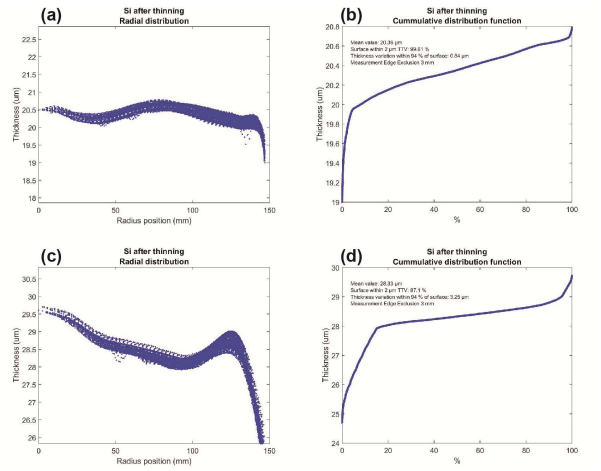

Figure 15. Thickness radial plot and weibull plot after (a), (b) 26 µm Si removal by dry etch after grinding and (c), (d) 22 µm Si removal by CMP after grinding.

The dry etch process shows an excellent performance in terms of surface stress relief and wafer thickness control. However, the surface roughness control was not as good as CMP. Figure 16 shows the wafer bright field optical image taken after grinding + dry etch. The grind mark was still visible even after 10 µm Si dry etch removal. The RMS roughness was 5.4 nm and the step height (Z range) was 47.9 nm, which is a large improvement compared to the value after grinding but it is not acceptable for the via-last process.

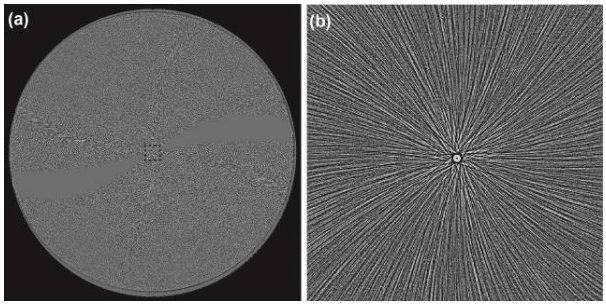

Figure 16. Bright field optical image taken after grinding and dry etch (a) whole wafer (b) high magnification at wafer center

Figure 17 shows PAS measurement results after grinding + dry etching. As a reference, the data after grinding + CMP (1 µm) is also on the graph. As we observed in the TEM image on Figure 14, the surface of grinding + dry etching is as good as CMP, and no vacancies are expected. However, higher S parameters were observed around 7 to 20 keV on the sample. This indicates the presence of mono-vacancies in the 500-2000 nm depth range. This might be due to the propagation of vacancies along with the grinding damage. Therefore, the isotropic dry etching can remove amorphous Si and deformed Si on the top of the surface caused by grinding, but long range dislocations may not be removed andremain in the Si and propagate deeper.

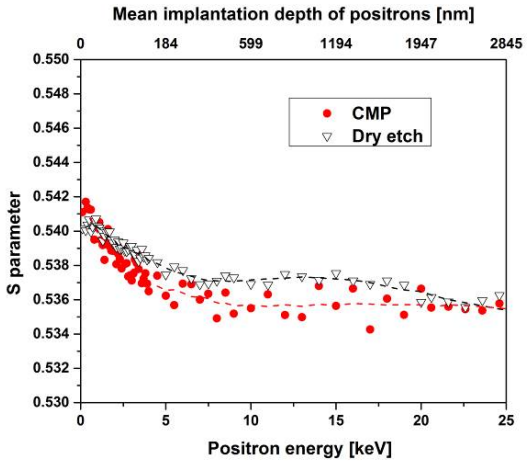

Figure 17. PAS S parameters as a function of incident positron energy E for the Si after grinding and dry etch.

E. Extreme thinning "Grinding + CMP + Dry etch"

From the experimental results discussed in the previous sections, the extreme thinning needs to be accomplished by combining several techniques to achieve a surface without damage and having a good wafer thickness uniformity. Furthermore, the sequence of the processes needs to be taken into account to avoid grinding damage propagation. Figure 18 (a) and (b) show wafer thickness maps and the wafer edge optical images taken after extreme thinning by grinding (to 25 µm) + CMP (to 5 µm). As showed in Figure 15, more than 20 µm Si removal by CMP may create a non-uniform wafer thickness, in particular at the wafer edge. The minimum thickness measured at the wafer edge for the case of grinding + CMP was 3.25 µm. Furthermore wafer edge breakage was seen caused by this process sequence, which might be due to the fragile thin Si edge and the mechanical pressure of CMP.

On the other hand, for the case of grinding (to 50 µm) + CMP (to 49 µm) + Dry etch (to 5 µm), this thickness drop at the edge was not observed (see Fig 18 (c)). The mean thickness after grinding + CMP + dry etch for a 300 mm wafer was 4.96 µm, which is very close to the target thickness. Furthermore, no edge breakage was seen when dry etch was used for the last step of the thinning (Fig 18 (d)).

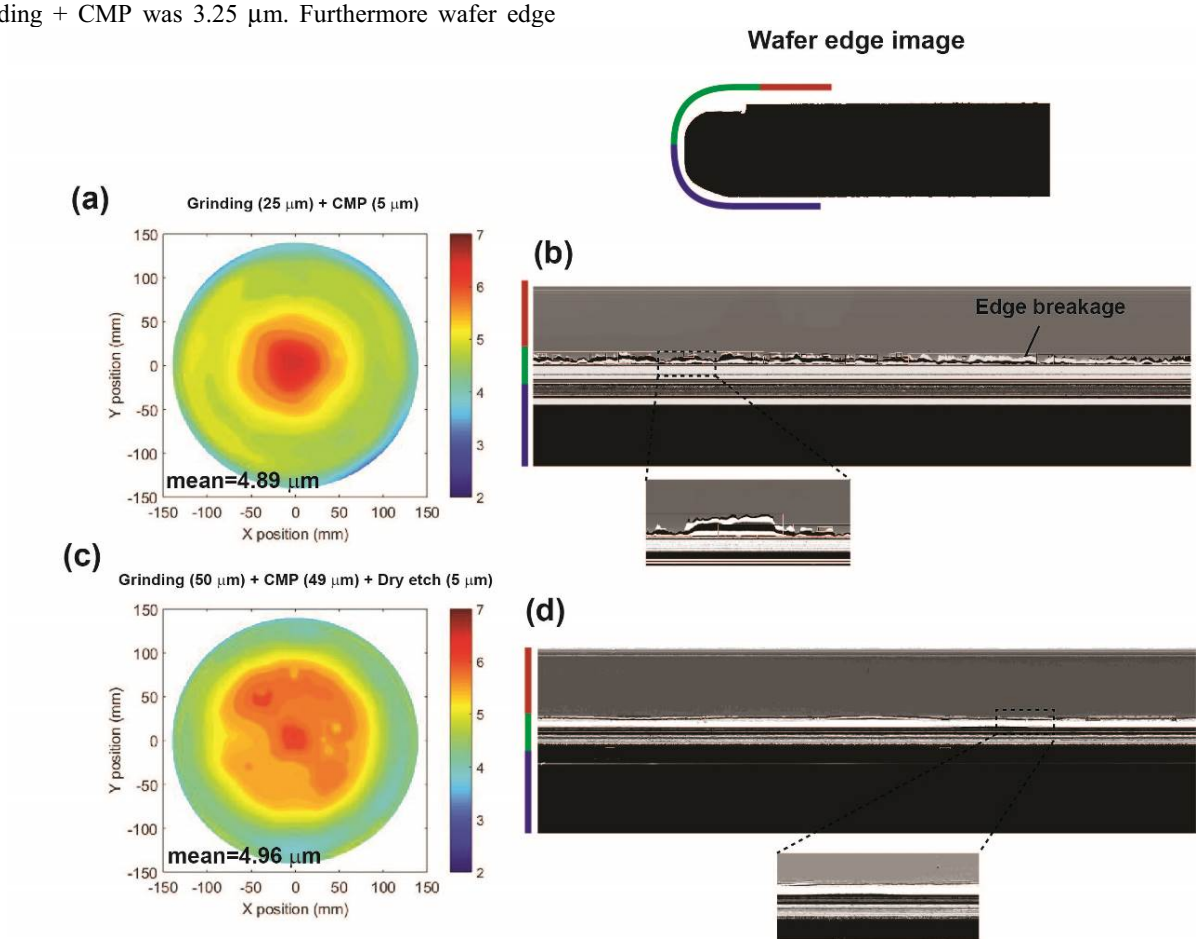

Figure 18. Wafer thickness map and wafer edge inspection image after (a) and (b) grinding and CMP (c) and (d) grinding + CMP + Dry etch.

Figure 19 (a) and (b) shows the cross-sectional TEM images after extreme thinning. No grinding damage remains on the top wafer surface. Figure 18 (c) and (d) show the AFM images after extreme thinning. The Z range for only grinding + dry etch was 47.9 nm, however it improved a lot (8.2 nm) when 1 µm CMP is applied in between the grinding and dry etching.

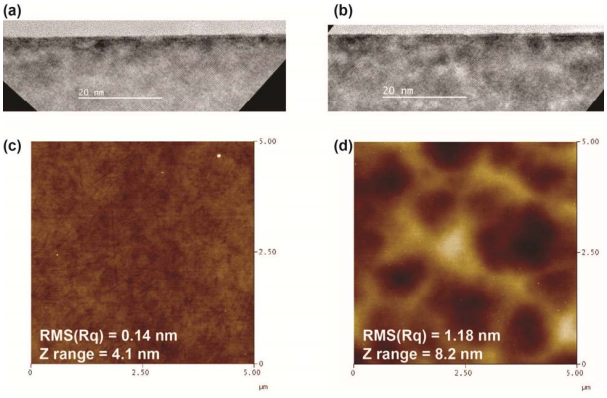

Figure 19. Cross-sectional TEM imagea and AFM imagea after (a) and (c) grinding + CMP (b) and (d) grinding + CMP + Dry etch.

Figure 20 shows the PAS measurement results taken after extereme thinning by grinding + CMP and Grinding + CMP + Dry etch. The S values for these samples are almost the same, suggesting that the concentration of vacancy-type defects is below the detection limit for these samples. Furthermore, the higher S parameter at 7 – 20 keV was not seen for the case of CMP inserted in between grinding and dry etch. This indicates that the removal of grinding damage before dry etch is important to maintain the Si mono-vacancy free.

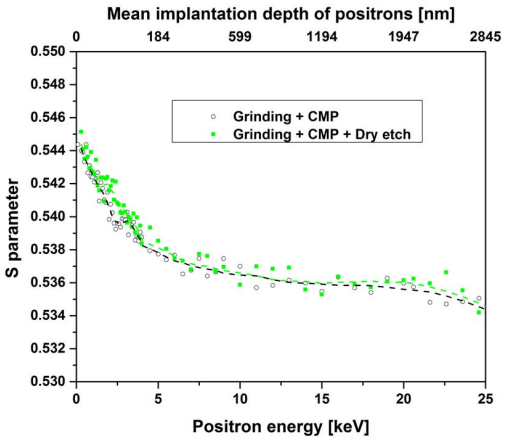

Figure 20. S parameters as a function of incident positron energy E for the Si after grinding + CMP and grinding + CMP + dry etch.

IV. Conclusions

To achieve successful extreme thinning (to a final Si thickness of 5 µm) on permanently bonded wafers, the surface after different thinning processes was characterized from an atomic to a macro scale. After grinding, a damaged layer depth of ~ 200 nm remains with a lot of vacancy clusters. The damaged layer creates a large compressive stress and concavely bows the bonded wafer. CMP can remove the stress and damaged layer resulting in a nano/atomic scale flat surface. The main challenge for CMP is to achieve extreme thinning with a good TTV and no mechanical damage at the wafer edge. Dry etching is better at maintaining the macro wafer shape but the vacancies caused by grinding will propagate along with the damage and remain within a certain Si depth. When CMP is applied in between grinding and dry etch, mono-vacancy free extreme thinning can be achieved with a good wafer shape. Extreme Si thinning processes with no mechanical failures and mono- vacancies have a potential to enable the required scaling of high density via-last TSV's.

Acknowledgment

The authors would like to thank imec's 3D team for their valuable input.

References

[1] Y. S. Kim, S. Kodama, Y. Mizushima, T. Nakamura, N. Maeda, K. Fujimoto, A. Kawai, K. Arai, and T. Ohba "A Robust Wafer Thinning down to 2.6-µm for Bumpless Interconnects and DRAM WOW Applications" IEDM 2015, S08P03

[2] Y. S. Kim, N. Maeda, H. Kitada, K. Fujimoto, S. Kodama, A. Kawai, K. Arai, K. Suzuki, T. Nakamura, T. Ohba "Advanced wafer thinning technology and feasibility test for 3D integration" Microelectronic Engineering 107 (2013) pp 65–71.

[3] A. Kailer, Y. G. Gogotsi, and K. G. Nickel, " Phase transformations of silicon caused by contact loading" J. Appl. Phys. (2002), 91, 2910 3057.

[4] M. J. Smith, M-J. Sher, B. Franta, Y-T. Lin, E. Mazur, and S. Gradečak, "The origins of pressure-induced phase transformations during the surface texturing of silicon using femtosecond laser irradiation" J. Appl. Phys. 112, 083518 (2012); pp083518-1-083518-8

[5] Z.J. Pei, Graham R. Fisher, J. Liu, "Grinding of silicon wafers: A review from historical perspectives" International Journal of Machine Tools & Manufacture 48 (2008) pp1297– 1307.

[6] A. Uedono, Y. Mizushima, Y. Kim, T. Nakamura, T. Ohba, N. Yoshihara, N. Oshima, and R. Suzuki, "Vacancy-type defects induced by grinding of Si wafers studied by monoenergetic positron beams" J. Appl. Phys 116, 134501 (2014)

[7] Y. Yang, K. De Munck, R. Cotrin Teixeira, B. Swinnen, B. Verlinden and I. De Wolf. "Process induced sub-surface damage in mechanically ground silicon wafers" Semicond. Sci. Technol. 23 (2008) 075038

[8] J. De Vos, M. Stucchi, A. Jourdain, E. Beyne, J. Patel, K. Crook, M. Carruthers, J. Hopkins, H. Ashraf, "Impact of Backside Processing on C-V Characteristics of TSV Capacitors in 3D Stacked IC Process Flows" EPTC2015, D1.1 - 191