NANOINDENTATION INDUCED SILICON FRACTURE AND 3D MODELLING

J. Garagorri[1], E. Gorostegui-Colinas[1], M.R. Elizalde[1], D. Allen[2], P. McNally[2]

[1] CEIT and TECNUN (University of Navarra). Manuel de Lardizábal 15, 20018 San Sebastián, Spain
E-mail: jgaragorri@ceit.es, egorostegui@ceit.es

[2] Dublin City University, Research Institute for Networks and Communications Engineering, Dublin 9, Ireland

ABSTRACT

Wafer handling during the manufacturing process introduces micro-cracks and flaws at the wafer edge. Some of these grow into larger cracks during thermal treatment, which can result into wafer breakage, disrupting manufacture. A study of the morphology of the defects and the stress regimes around these micro-cracks is necessary in order to derive quantitative and predictive information to avoid catastrophic failure.
In order to reproduce mechanical damage during wafer handling, controlled damage by nanoindentation has been carried out on initially defect free square Si samples. The samples have been cleaved from 300 mm double side polished silicon wafers and loads between 75 and 150 mN have been applied using a Berkovich diamond tip. Cross-sectional FIBbing has been done in order to characterise the crack systems present under the indents, and Raman measurements have been performed to evaluate the residual stress profile after damage generation.
At the same time, 3D FE modelling has been developed to reproduce the indentation process in a small piece of silicon. Anisotropy has been introduced through elastic constants and isotropic plasticity has been also considered. Two types of simulations have been developed. On one hand, the simplest ones describe just the indentation process, where no crack evolution has been modelled. The Raman spectrum obtained from these simulations is in accordance with the experimental results. On the other hand, cohesive planes have been introduced in the positions where cracks should be developed, according to cross-sectional observations. This implemented model can simulate the evolution of cracks in the material.

KEYWORDS: Silicon, fracture, nanoindentation, FEM, cohesive zone model, FIB, Raman.

1. INTRODUCTION

The up-scaling in the size of silicon wafers in semiconductor industry generates a compromise between the benefits of enhancing the production of devices for micro electro-mechanical (MEM) and electronic applications, and the difficulties related to gravitational effects and thermal stresses on the wafers during manufacturing [1].

Every wafer diameter transition has introduced problems in high-temperature processing. Radial temperature gradients associated with larger wafer diameters cause greater stress in the wafer during heat up and cool down cycles. If the stress exceeds the yield stress of the silicon, slip can form in the wafer [2]. These defects generated during thermal treatments can cause loss of yield performance and reliability in the device. In addition to this, if there already exists surface and sub-surface extended damage in the edge due to wafer handling and shipping, thermal stresses can even induce catastrophic breakage and, subsequently, the need to stop and clean the production line at a great cost, which is estimated in €2.5 M per year for each silicon fabrication line in the world [3].

The aim of this paper is to describe the techniques used to reproduce and characterise the damage in silicon wafers so as to obtain a source of realistic information in order to be able to simulate stress fields and fracture in this material. In particular, FEM simulation technique has been used in order to characterize the stress fields around indentations with a Berkovich tip. Two types of simulations have been considered. The first ones account just for the stress field around the indent, while the second one goes further trying to model how fracture is initiated under the indentation and its evolution during the process.

2. EXPERIMENTAL

In order to reproduce the damage introduced during wafer handling in real processes, controlled damage by nanoindentation has been carried out on initially defect free square Si samples at CEIT. The samples have been cleaved from a (100) oriented 300 mm double-side

polished (DSP) silicon wafer. Loads between 75 and 150 mN have been applied at rates of 10 mN/s, using a Nanoindenter® II (Agilent, formerly Nano Instruments Inc.) with a Berkovich diamond tip. The indents have been oriented so that one of the imprint sides is parallel to [110] direction. The arrangement is as shown in Figure 1.

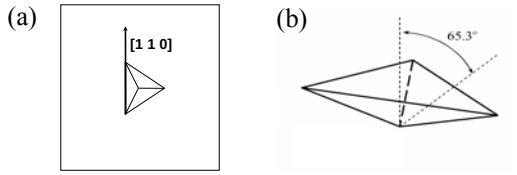

Figure 1. (a) Scheme of the orientation of an indent in a defect free square Si sample. (b) Geometry of the indenter.

A Quanta 3D Dual Beam (FEI) focused ion beam (FIB) has been used for the characterisation of the crack systems under the indents. For this, top-down and cross-section millings have been carried out on indents at different peak loads, and a 3D description of cracks has been obtained from FIB tomography using Amira® 4.0 software.

Micro Raman Spectroscopy (μRS) has been performed at Dublin City University with a Jobin-Yvon LabRam HR800 Raman microscope to evaluate the residual stress profile generated by the indents. This experimental technique is fully explained in references [4] and [5].

3. SIMULATION

This section is dedicated to the modelling of the indentation process. It is divided in two main parts: in the first one a general description of the model is given, while the second part is left to a more detailed description of the fracture modelling fundamentals.

3.1 Description of the modeling technique

The aim of the FEM simulations that have been designed is to describe the experimental process of indentation. All the models have been run with the finite element method based commercial software ABAQUS® (6.8.3 version).

The indenter chosen has been a Berkovich tip. Its geometry is described in Figure 1. It has been implemented as a rigid part, with no material properties. This point does not affect the model since the real tip is made of diamond, whose Young modulus is much larger than the Young modulus of silicon. As for the meshing, it has been made as coarse as possible in order to have few elements. Linear triangular elements have been used in the construction.

For the piece of silicon to be indented two models have been designed depending on the aim of the simulations. The simplest model consists of a 48 x 48 x 96 μm piece. It has been designed using a solid deformable part in which elastic anisotropy and isotropic plasticity have been introduced [6]. The central part is meshed using a structured mesh with an element size of 0.25 μm. This size for the elements is good enough for indents from 150 mN to 250 mN. For higher/lower loads the mesh should be changed to a coarser/finer one, respectively. The rest of the part is meshed much more coarsely. All this can be seen in Figure 2.

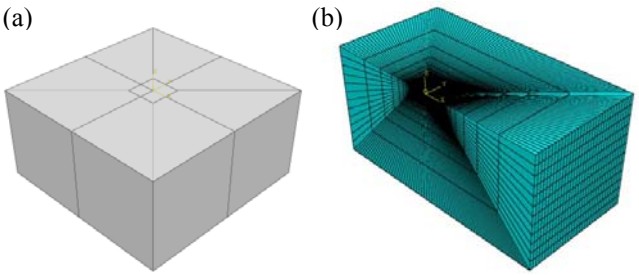

Figure 2. In (a) the geometry of the Si piece to be indented is shown. (b) shows the mesh for for half of the model in the simplest case.

The reason for this election of meshing is simple: the highest deformations appear around the indent. The mesh in this region should be fine enough to account for the correct stress fields and geometry of the indenter imprint. On the rest of the wafer a coarser mesh could be used. As for the election of the size of the Si piece, it is big enough to avoid boundary effects. This type of piece has been used to model just the stress fields during the indentation process.

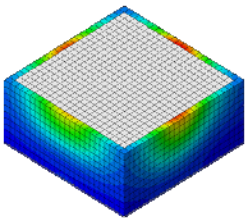

Figure 3. Central part of the big model used to get the correct boundary conditions in the borders for a piece of silicon of this size.

The second type of piece designed has the same geometry but a different technique has been used in order to be able to have a denser mesh with no higher computational exigency. In a first step the model shown in Figure 2(a) is considered with the same type of meshing, but with much bigger elements (being their size of 0.5 μm). In a second step a piece with the size of the central part is used with a refined mesh and the stresses from the first simulation as boundary conditions (see Figure 3). This allows for a much denser mesh in the piece (the element size is lowered to 0.2 μm), a highly important issue in order to be able to see the initiation and evolution of cracks, which is the aim of this second model.

The elements of the first models are all hexahedral structured elements. In the second model cohesive elements have been introduced in a vertical plane. This kind of elements allow for the modelling of the initiation and evolution of fracture according to the chosen cohesive law (described in the following section).

3.2 Cohesive elements and fracture

Cohesive elements are a type of special elements that can account for fracture. This kind of elements must be located in the places where fracture occurs experimentally. In the case of silicon, the correct place seems to be cleavage planes. What is more, not in all these planes fracture initiation and evolution is equally developed. This is because the cleavage tension is different in each case. The planes which present the lowest value of this magnitude are {1 1 1} planes, then {1 1 0} and finally {1 0 0} [6]. In the simulations just {1 1 0} and {1 0 0} planes have been considered.

The simulation of fracture initiation and evolution consists of three stages [7]. First a linear elastic traction separation behavior is considered:

$$\vec{t} = K\varepsilon \qquad (3.2.1)$$

where $\vec{t} = (t_n, t_s, t_t)$ represents the nominal traction stress vector, $\vec{\delta} = (\delta_n, \delta_s, \delta_t)$ the corresponding separations and finally $\varepsilon = (\varepsilon_n, \varepsilon_s, \varepsilon_t) = (\frac{\delta_n}{t_n}, \frac{\delta_s}{t_s}, \frac{\delta_t}{t_t})$ is the nominal strain.

The second step of the simulation begins when

$$quadscrt(t_n, t_t, t_s) = \left(\frac{\langle t_n \rangle}{t_n^0}\right)^2 + \left(\frac{\langle t_t \rangle}{t_t^0}\right)^2 + \left(\frac{\langle t_t \rangle}{t_t^0}\right)^2 \qquad (3.2.2)$$

reaches the value of 1 in at least one of the elements of the cohesive layer. t_n^0, t_s^0 and t_t^0 represent the peak values of the nominal stresses and t_n^0, t_s^0 and t_t^0 the peak values. This is the moment when damage initiates.

Finally, damage evolution, the propagation of the crack, is described by:

$$t_n = \begin{cases} (1-D)\bar{t}_n & t_n \geq 0 \\ \bar{t}_n & \text{otherwise} \end{cases} \qquad (3.2.3)$$

where D is a scalar damage variable that represents the overall damage in the material. The graphical representation of the cohesive law is shown in Figure 4.

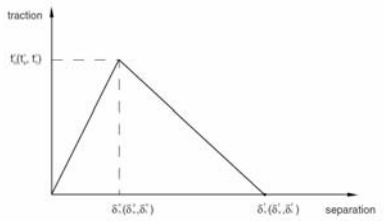

Figure 4. Graphical representation of the cohesive law.

4. RESULTS AND DISCUSSION

4.1 Damage characterization around indents

Direct observation of the crack systems beneath the indents gives an idea of the type and extension of the damage generated by nanoindentation. The micrographs obtained from top-down FIBbing, and shown in Figure 5, suggest a double crack system, i.e., radial and median cracks, coexisting under indents around 100 mN.

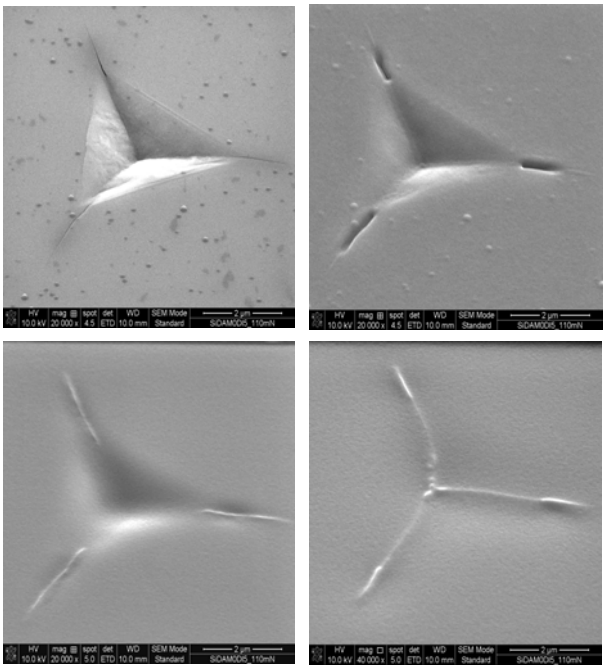

Figure 5. Sequence of FEGSEM micrographs from a top-down FIBbing on a 110 mN Berkovich indent on Si, showing the coexistence of radial (Palmqvist) and median cracks close to the surface, and further development of median cracks under the indenter imprint.

Radial or Palmqvist cracks emerge from the corners of the indenter imprint and have a maximum depth of the order of 0.5 μm. These close-surface cracks have been observed in brittle materials as a result of sharp indenters at low loads [8-11]. Recently, some studies relate the appearance of Palmqvist cracks as the first crack systems to form, reaching their maximum extent upon the unloading [12].

Besides, median cracks -also referred as half-penny cracks when related to Vickers indents- originate

beneath the remaining plastic region under the indenter imprint at higher loads and propagate along the median axis upon the unloading [8].

On the other hand, direct observation through cross-sectional FIBbing allows depicting a different landscape of the damage. In addition to the mentioned crack systems, lateral cracking has been detected under the indenter imprint at peak loads between 120 and 150 mN. According to Figure 6, lateral cracking runs across a region of well oriented dislocations and above the intersection with a median crack. Another example is shown in the 3D reconstruction in Figure 7 where, in addition to the median crack following a vertical (100) plane, a lateral crack is also initiated.

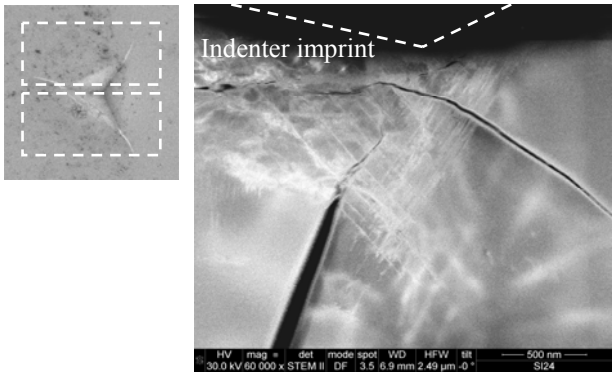

Figure 6. STEM image of a cross-section cut of median and lateral cracks under a 150 mN Berkovich indent on Si.

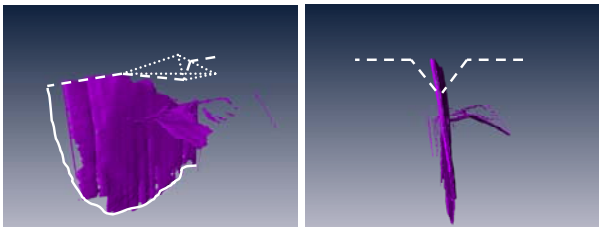

Figure 7. 3D reconstruction from FIB tomography of the median and lateral cracks under a 120 mN Berkovich indent on Si.

4.2 Simulations and comparison with experiments

4.2.1 Nanoindentation Stress Field simulations

The results shown in this section correspond to the simulations run in the first type of silicon piece model. The relative orientation of the Berkovich tip with respect to the material crystal system is shown in Figure 8. Even this simple model accounts well for the stress fields around the indentation. If experimental load-displacement curves are considered, it can be seen (Figure 9) that the predicted loading is equal to the experimental one. Differences appear, however, in the unloading. There are several reasons for these discrepancies. First of all, during the indentation process silicon suffers several phase transformations, which make the material behave differently. In the loading this is reflected as a small pop-in while in the unloading a pop out appears in the curve. Accounting for the transformation in the loading does not seem to be crucial, since good results have been obtained for this part of the indentation process. However, the expansion that the material suffers in the phase transformation occurring at the unloading appears to change the material behaviour completely, as the simulated curve does not agree with the experimental one. Furthermore, this model does not account for cracks which usually grow up in the unloading and may also affect to the shape of the curve.

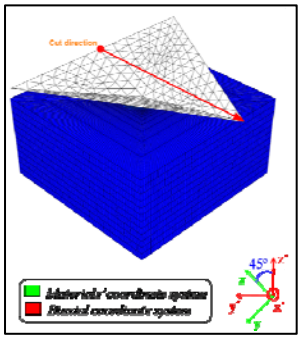

Figure 8. Relative orientation of the crystal axes of silicon and the indenter. The cut direction refers to the line where Raman measurements have been made.

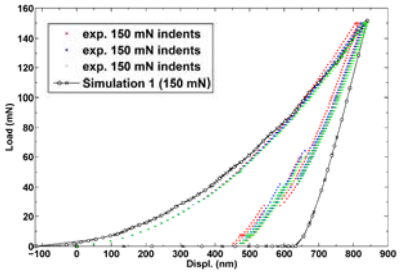

Figure 9. Experimental load displacement curves compared to those obtained with the simple model.

Another way of validating the model is considering the Raman experiments in silicon wafer and comparing the results to the corresponding frequency shifts predicted by the simulations. Details of this experimental technique can be found in references [4] and [5]. First of all, the level of approximation needed for the frequency shift computation must be determined. For this task the stress tensor components in the central uniformly meshed part of the direction line marked in Figure 8 have been plotted in Figure 10. This plot clearly shows that the components σ_{xy}, σ_{yz} and σ_{xz} are negligible, which allows for biaxial approximation. In this approximation the frequency shift observed can be written as in (4.2.1) (stress components must be introduced in MPa) when the considered coordinate system is that of the material and the measurements are made at the top of the sample [5]:

$$\Delta\omega = 1.93 \cdot 10^{-3} \cdot (\sigma_{zz} + \sigma_{xx}) \qquad (4.2.1)$$

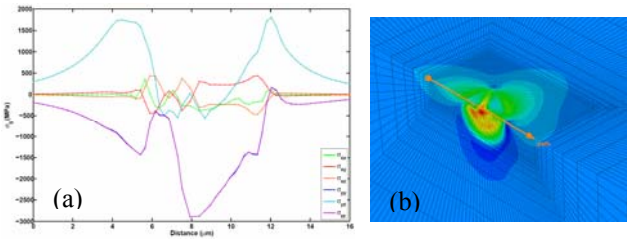

Figure 10. Stress tensor components along the direction marked in Figure 8. Just the central uniformly meshed line highlighted in b has been considered

The results obtained in the simulations agree quite well with the experiments (Figure 11). As happened with the load displacement curves, the discrepancies can be due to the lack of consideration of the phase transformations in the piece of silicon modelled. Furthermore, the resolution of the Raman measurement is limited in the central part of the indent imprint due to the inclined surface.

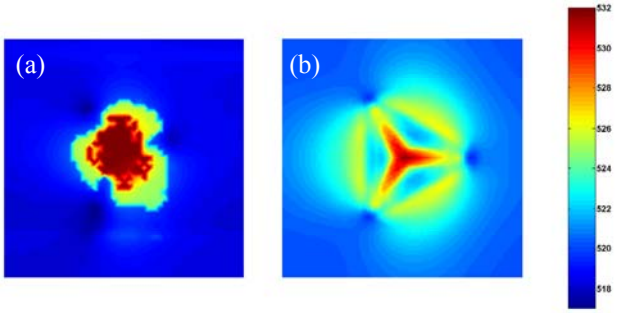

Figure 11. Comparison of the 2D Raman spectrum map (a) obtained experimentally, and (b) the simulation.

4.2.2. Modelling of fracture initiation and evolution in the indentation process

The two cohesive planes considered are a (110) plane (vertical plane) and a (001) plane (horizontal crack) (see Figure 12).

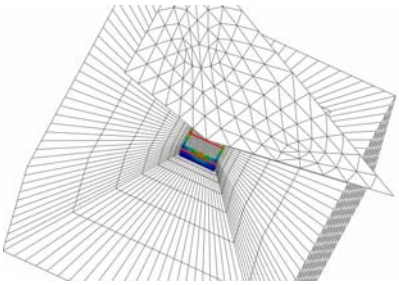

Figure 12. Relative position of the indenter and the vertical and horizontal cohesive layers (the lines are highlighted in red).

First of all, a set of test simulations were performed in the model with a vertical cohesive plane in order to determine the adequate parameters for the cohesive law. According to literature [7], the silicon fracture energy varies between 2 and 4 J/m^2, depending on the cleavage plane considered. For the simulations in this paper a value of 4 J/m^2 has been considered. As for the value of the slope of the cohesive law, it should equal or higher than the average Young modulus of silicon in order to have the same behaviour in the layer and in the rest of the piece. Several simulations have been run with values ranging from 1.87x10^7 MPa to 12.65x10^8 MPa (the average Young Modulus for the type of wafer considered is about 1.80x10^5 MPa). The results indicate that for values above 8.65x10^8 MPa convergence issues arise. This is why finally for all the simulations the value has been set to 1.28x10^8 MPa.

Finally, the peak stress has to be determined. Experimentally, cracks appear even for indents of 30 mN. As simulations must respect this fact, a first value of 2000 MPa was considered, noting that cleavage tension is of this order of magnitude. This value has been decreased until the initiation of damage occurred at 30 mN for a value of the nominal stress of 1200 MPa. Table 1 shows some of the attempts.

Table 1. reflect the behaviour of lowering the nominal stress in the cohesive law.

σc (MPa)	Load when Quadscrt=1
1800	120
1600	75
1300	35
1200	31

Damage initiation occurs when *Quadscrt* (ec. 3.2.2) reaches the value of 1. Figure 13(a) shows the correspondent contour plot for *Quadscrt* at the instant of initiation of damage at a load of 30 mN. In Figure 13(b) the value of *Quadscrt* is shown for the same instant of time but with a nominal stress of 1300 MPa. As can be seen, damage has not been generated yet in this second simulation.

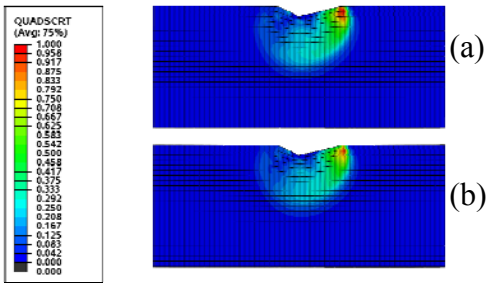

Figure 13. Contour plot of the magnitude quadscrt for simulations with Y=1.28x10^8 , G=4J/m^2. Their difference relies in the value of the nominal stress: 1200 MPa for (a) and 1300 MPa for (b).

Once damage is initiated the damage extent is represented by SDEG, a scalar variable which varies from 0 to 0.8. Every element that reaches this value is deleted, and hence fracture propagation can be easily observed. In Figure 14 a contour plot of this magnitude is shown (a) at maximum indentation load and (b) after the indentation unload. The results indicate that during

unloading of the indentation process fracture continues propagating. The shape of the crack agrees qualitatively with experimental results. It should be noted that the maximum indentation load reached during the simulations is 75 mN. More simulations are needed to fit quantitatively the crack size measured.

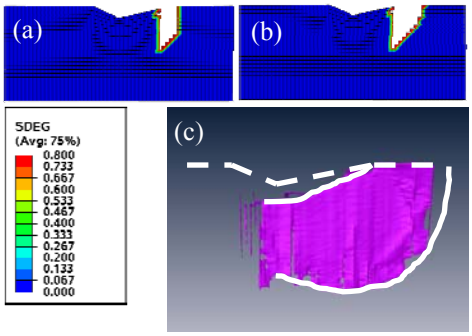

Figure 14. Vertical cohesive plane. The scalar magnitude SDEG, which is a measure of the generated damage, at (a) maximum indentation load and (b) after the indentation process. (c) is a 3D reconstruction of a vertical real crack in the same plane.

For the simulations with a horizontal cohesive plane at a distance of 1µm from the surface, a (001) plane, using the same cohesive law parameters as in the previous simulation the damage generated is almost none: *Quadscrt* is equal to $2x10^{-4}$ at maximum load and to 0.66 after unloading. This result indicates as well that cracks evolve during the unloading of the indentation process. It is noticeable that this result matches the experimental observations, since horizontal cracks are not observed for indents performed below 120 mN. In fact the cleavage tension of (001) planes is much higher, which is an indication of a different response of the {100} planes compared to the {110} family.

Finally, another simulation was made in order to get fracture in this plane. The nominal stress was lowered to 600 MPa. The results of this simulation are shown in Figure 15. In this case at maximum indentation load damage is again negligible but during unloading damage initiates. However, fracture is not achieved. It should be noticed that the position of the horizontal plane is fixed and could not be the most appropriate for damage to occur.

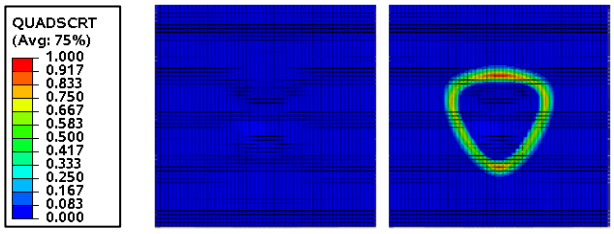

Figure 15. Quadscrt in the horizontal plane at maximum indentation load of 75 mN and after the unloading. The peak stress of the cohesive law was 600 MPa. No cracks appear.

5. CONCLUSIONS

First of all, the simulated stress fields agree well with Raman measurements around the indents, even though no phase transformation nor fracture are implemented in the FEM models.

Fracture simulation first attempts with cohesive elements show that the vertical model seems to be well designed since the initiation and initial evolution agrees with experiments. However, more simulations are needed in order to calibrate the cohesive law parameters to match experimental results quantitatively. In the case of the horizontal crack model the situation of the plane is another parameter to consider. In this case information from 3D FIB tomography reconstruction will be very helpful.

Finally, the fact that small values of nominal stress are required to fit the simulations with the experimental results is probably related with the brittle fracture of silicon and the cleavage stresses.

6. REFERENCES

[1] K.C. Cho, H.T. Jeon, J.G. Park, J. of the Korean Physical Society, Vol. 46, No. 4 (2005), pp. 1001-1006.

[2] W. Murray Bullis, Mat. Sci. and Eng. B72 (2000), pp. 93–98.

[3] International SEMATECH Manufacturing Initiative, Industry Economic Model v8.1ss 2004.

[4] IEEE Transactions on components and packaging technologies, Vol. 28, No. 3, 2005.

[5] D. Allen, J. Wittge, et al. 2009, Nuclear Instruments and Methods in Physics B, DOI: http://dx.doi.org/10.1016/j.nimb.2009.10.174

[6] F. J. Morin, J. P. Maita, Phys. Rev. 96, 1 (1954), pp. 28-35.

[7] International Journal of Fracture 155 (2009), pp. 67-74.

[8] Y. Tang, A. Yonezu, N. Ogasawara, N. Chiba, X. Chen, Proc. R. Soc. A 464 (2008), pp. 2967-2984.

[9] F. Tancret, Scripta mater. 43 (2000), pp. 9–14.

[10] T. Lube, Journal of the European Ceramic Society 21 (2001), pp. 211-218.

[11] M. Manoharan, G. Muralidharan, Journal of Materials Science: Materials in Electronics 13 (2002), pp. 39-41.

[12] K. Nihara, Journal of Materials Science Letters 2 (1983), pp. 221-223.

ACKNOWLEDGEMENTS

The authors want to acknowledge the European Commission for its financial support through SIDAM project (FP7-ICT-216382).