A review of hydrophilic silicon wafer bonding

How to cite:

For guidance on citations see FAQs.

© 2014 The Electrochemical Society
Version: Version of Record
Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.1149/2.007403jss
http://jss.ecsdl.org/content/3/4/Q42

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
A Review of Hydrophilic Silicon Wafer Bonding

V. Masteika,a,* J. Kowal,a N. St. J. Braithwaite,a and T. Rogersb

aThe Open University, Milton Keynes MK7 6AA, United Kingdom
bApplied Microengineering Ltd., Harwell Campus, Didcot OX11 0SG, United Kingdom

Hydrophilically activated direct wafer bonding is a technique for gluelessly attaching oxide-coated wafers together. This ability is a vital step in the construction of many microelectronic and microelectromechanical (MEMS) devices. In particular this technique is widely used in the production of 3d interconnected devices due to the lack of interlayer.

This review covers the key papers relating to the theory, techniques and quantification of hydrophilically activated silicon wafer bonding. This begins with a review of the history and development of the art. Bond strength characterization is then reviewed followed first by models of physical deformation and then by plasma and radical activation techniques.

In the interests of clarity and succinctness this review is not an attempt to exhaustively catalog all direct bonding wafer literature. Excluded are hydrophobic and UHV bonding, and the many papers applying bonding techniques to differing combinations of wafer materials. This paper concludes with a summary of the state of the art of direct wafer bonding and a summation of the current wafer-bonding model derived from the papers reviewed.

History and Establishment of the Art

The tendency for optically polished pieces of glass or metal to stick to each other was first noticed around the turn of the 20th century by German craftsmen. Termed Ansprengen or “jumping contact”, the effect was first considered a nuisance and then utilized to stick optical elements such as prisms in place without any interfacial layers. The effect was first studied in detail in 1936 by Lord Rayleigh in “A Study of Glass Surfaces in Optical Contact” where he determined the contact energy to be of the order of 100 mJ m⁻². This phenomenon was subsequently investigated as part of a study on the interaction of surfaces by Tabor et al. He attributed it to Van der Waals bonding between the surfaces, or between adsorbed monolayers of molecules with large dipole moments, such as water. However this knowledge was not applied to wafer bonding for some time.

An early form of wafer bonding, anodic bonding, was first reported in 1969 by Wallis and Pomerantz. This involves the bonding of sodium-containing glass to silicon under an applied voltage and elevated temperature of 300–500 °C. A large electrostatic field arises because most of the potential difference is now dropped across this very narrow depletion zone. It is this field that assists in the formation of the bond. However the use of glass anodic bonding is restricted in semiconductor applications because of the necessary presence of mobile alkali metal ions, which are detrimental to the performance of semiconducting junctions in silicon.

In 1975 Antypas and Edgecumbe first reported “wafer fusion bonding”. This involved bonding a thin film of gallium arsenide to a glass wafer (SiO₂) at an elevated temperature. A mirror polished GaAs wafer was contact bonded to a mirror polished silica handle wafer via Van der Waals interactions. The structure was then heated at above 800 °C to fuse the layers together, hence “wafer fusion”. An aluminum-rich AlGaAs etch-stop layer, implanted in the GaAs wafer prior to bonding to the silica handle wafer, was used to allow the etching back of the GaAs wafer to a thin layer of 1.5 μm. This thin layer could then be transferred via the glass handle wafer. The process is illustrated in Figure 1.

This etch stop technique became important over the next decade. Hydrophilic direct wafer bonding was first reported in 1985/86 when several groups reported the bonding of silicon to silicon at room temperature followed by a high temperature anneal. Mirror polished silicon wafer surfaces were placed into contact after a chemical cleaning stage which covered the surface with OH groups, rendering them hydrophilic. Initial contact resulted in the formation of a weak, room temperature, water-mediated bond that drew the wafers into intimate contact. The bonded wafer pairs were then annealed at over 800 °C to drive out the water interlayer and form strong, permanent covalent bonds (Figure 2).

Lasky et al from IBM covered one or both wafer surfaces with a thick thermal silicon dioxide layer while Shimbo et al. at Toshiba reported this form of bonding without the prior growth of any thermal oxide.

Shimbo was aiming to replace the relatively disordered epitaxially grown crystalline silicon with more optimal single crystalline layers for power applications while Lasky was attempting to create a silicon-on-insulator (SOI) structure.

SOI structures consist of a thin single crystal silicon device layer on top of a disordered silicon handling layer electrically insulated from it by a thermal oxide. SOI was initially developed for its radiation-hard properties but it was quickly realized that it also allowed for operations at lower voltages, offered a higher device yield and enabled the simultaneous use of high and low voltage components within the same device. At the time this structure was created by the Separation by IMplantation of OXygen (SIMOX) method. In this process the silicon wafer was heated to near melting point and bombarded with high-energy oxygen ions, which agglomerated within the silicon, forming the insulating layer. However this method was expensive and created a large number of dislocations and lattice defects because of the ion bombardment.

Lasky’s intent was to develop a SOI production method to replace SIMOX with a cheaper alternative that left a high quality device layer. To this end a single crystal device wafer was bonded to a “handle” wafer covered in thick thermal silicon oxide. After bonding the handle and the device silicon wafers, the device wafer was mostly removed using the etch stop technique to leave a thin device layer.

These papers mark what is traditionally the start of hydrophilic wafer bonding literature. The first use of hydrophilic wafer bonding in a MEMS application can be found in the 1988 paper by Peterson and Berth, describing the creation of pressure sensors.

Although the first explicit direct wafer bonding literature was published in 1985 the theoretical background was established much earlier.

*Electrochemical Society Active Member.
*E-mail: v.masteika@open.ac.uk

Manuscript submitted June 5, 2013; revised manuscript received December 9, 2013. Published February 3, 2014.

© 2014 The Electrochemical Society. [DOI: 10.1149/2.007403jss] All rights reserved.
earlier. The 1969 paper by C. G. Armistead et al. provides a description of how silica (silicon dioxide) and hydroxyl (Si-OH, or OH groups) are arranged on a silicon surface that is still frequently referenced today.

Presenting a mixture of original experimental work and supporting literature, a model is developed of Si wafer surfaces covered with 4.6 ± 0.2 OH groups per 100 Å² at 25°C. In each 100 Å², these consist of 1.4 ± 0.1 single, type A non hydrogen-bonded OH groups and 3.2 ± 0.1 paired, type B hydrogen-bonded OH groups. The surface structures of the two groups are shown in Figure 3.

These two hydroxyl structures exist because of differences in the structure of the silica surface. In order to explain experimental results the paper postulates that silica surfaces are not homogeneous, but instead are composed of numerous polymorphs. The precise density and type of surface hydroxylation then depends on which crystal plane of which silica polymorph makes up each part of the surface.

The paper analyzes the surface concentration and temperature dependence of the two hydroxyl groups, however it does not experimentally determine the crystal structure of the silica surface. Instead an example model of surface crystal structure is advanced that is consistent with the experimental data.

100 plane β–tridymite (Figure 4), a high temperature polymorph of silica, is presented as an example of a substrate for type A groups. The type A group is stable up until approximately 600°C. This is due to the OH groups’ nearly orthogonal orientation to the surface keeping the reactive H end too far from any other bulk silica or hydroxyl oxygen atom to react at lower temperatures.

0001 plane (hexagonal crystal system miller index) β–cristobalite, a higher density silica polymorph (Figure 5) is presented as an example of a substrate for type B groups. Type B groups exist in thermal equilibrium, from 3.2 ± 0.1 at 25°C up until 600°C ± 50°C at which point every group is hydrolysed. Two OH groups are attached to each surface silicon atom, each held at very low angles to the surface (approximately 35°) and close to adjacent hydroxyl oxygen atoms. This favors hydrogen bonding and hydrolysis of adjacent OH groups. A ‘fully’ hydroxylated 0001 β–cristobalite plane contains a maximum of 7.9 OH groups per 100 Å². However the authors determine that a room temperature 0001 β–cristobalite plane is only ~60% hydroxylated.

At room temperature approximately 30% of the surface OH groups on a typical native oxide silica surface are type A. This leads to the conclusion that about 30% of the surface is made up of β–tridymite (or β–tridymite like crystal plane), and 70% is made up of β–cristobalite...
(or β-cristobalite like crystal plane). When fully hydroxylated the remaining 70% of β-cristobalite should contain 5.53 OH groups per 100 Å², rather than the 3.2 OH groups detected at 20–30°C. Therefore at room temperature the native oxide surface of a silicon wafer is only approximately 70% hydroxylated, assuming a similar ratio of β-cristobalite to β-tridymite (or structurally similar planes). This is illustrated in Figure 6.

The most significant conclusion of this paper – that the surface density of OH groups at room temperature is 4.6 per 100 Å² – has since been frequently quoted. However, there have been no subsequent attempts in the wafer bonding literature to investigate the types of OH groups on thermal or native silicon oxide. Intriguingly if surface hydroxylation was controlled during the activation and contact process it could be possible to create bonds with greater than 4.6 interfacial links per 100 Å².

After initial reports in 1985–1986, the first detailed study of direct wafer bonding was published in 1988 by W. P. Maszara et al.11 This was the third paper published discussing direct wafer bonding and the first to exhibit many of the features, techniques and assertions that are frequently repeated in subsequent research. Most notably, it introduces the dual cantilever bond strength test and begins to develop hypotheses on the mechanics of wafer bonding.

A number of wafers were prepared and hydrophilically activated, either by steam or wet ammonium hydroxide (NH₄OH). The activated wafers were then contacted and annealed at a range of temperatures up to 1400°C. Bond strength measurements were made after annealing at various temperatures and the results used to begin developing a bonding model. This model consists of three distinct stages. At room temperature contact, the wafers become weakly attached by a hydrogen bonded water layer. As temperature increases water is removed and covalent oxide bonds form across the interface, strongly bonding the wafers. Finally at high temperatures of 800–1400°C the oxide flows slightly: it plastically deforms. This fills in interfacial microvoids forming a ‘perfect’ bond.

Much of the paper is dedicated to describing and analyzing the dual cantilever bending (DCB) testing methodology. The background theory is presented by P. P. Gillis and J. J. Gilman in 196412 from which the simplified equation subsequently known as the ‘Maszara’ formula, equation 1, is developed (Figure 7). The validity of the equation was then tested by controlling the thickness of the debonding blade and the thicknesses of the wafer samples under test.13–25

$$\gamma = \frac{3E'\nu^2w^3}{8l^4} [1]$$

Where $$E'$$ is the effective Young’s modulus which varies with the orientation of the moment of bending to the crystal lattice. It should be noted that a second, common form of this equation is as follows:

$$\gamma = \frac{3E'\nu^2w^3}{32l^4} [2]$$

Where $$\nu = 2u$$, the entire thickness of the debonding wedge.

A second erroneous conclusion was also introduced in this paper, most likely as a result of the low frame rate IR imaging equipment available at the time. Sub-critical crack growth in a humid atmosphere (i.e. stress corrosion) is mentioned and an attempt is made to detect its influence on DCB testing. However, it is mistakenly assessed as having a minimal effect on the measured bond strength. Many subsequent papers neglect to control or report the elapsed time between blade insertion and crack length measurement, thus allowing stress corrosion to add significantly to the uncertainty of reported measurements.

In 1989 R. Stengl et al. presented one of the first coherent accounts of the modern three stage wafer bonding model.26 This was based upon the authors detailed observations of the bonding wave, initiated at differing temperatures. Combining their own observations with the nascent wafer bonding literature they present a three stage, temperature dependent, bonding model based upon the chemical interactions of OH terminated silicon surfaces.
Publications discussing wafer bonding theory had established the fundamental bonding mechanism in hydrophilic wafer bonding as an initial bond of surface OH groups to a water interlayer.

\[ \text{Si} - \text{OH} + \text{nH}_2\text{O} \rightarrow \text{HO} - \text{Si} \]  \[\text{(a)}\]

During annealing a condensation polymerization reaction occurs to produce covalent high strength bonding.

\[ \text{Si} - \text{OH} + \text{HO} - \text{Si} \leftrightarrow \text{Si} - \text{O} - \text{Si} + \text{H}_2\text{O} \]  \[\text{(b)}\]

As this reaction is reversible up to approximately 400°C, the released water is then able to diffuse through the relatively thin siloxane layer into the bulk silicon, releasing hydrogen as per reaction (c).

\[ \text{Si} - \text{Si} + \text{H}_2\text{O} \rightarrow \text{Si} - \text{O} - \text{Si} - \text{H}_2 \]  \[\text{(c)}\]

This hydrogen gas that is the primary cause of annealing voids. (c)

These chemical reactions were suggested based upon well established silica chemistry literature. However it was desirable to actively probe the chemical states of the treated wafer surfaces and the bond interface, in order to experimentally derive bond densities and the like. Secondary Ion Mass Spectrometry (SIMS) and X-ray photoelectron spectroscopy (XPS) can be used in these roles. However both techniques can only be used to probe either a surface or a subsurface and are not suitable for inspecting the bond interface. Fourier transform infrared spectroscopy (FTIR) proved to be an excellent method for bond interface inspection. First introduced to wafer bonding in the 1994 paper by Chabal et al., infrared absorption spectroscopy is particularly useful for silicon samples due to the materials transparency over much of the IR spectrum. Unfortunately the resultant spectrum is typically of low signal to noise ratio when analyzing thin films. By utilizing the experimental geometry shown in Figure 8, the infrared beam may make as many as a hundred passes through the interface, significantly increasing the signal intensity.

Utilizing FTIR-MIR the Chabal paper determines the presence of chemisorbed OH groups, both at the interface and in the bulk oxide, the presence of 2–3 monolayers of water, and the presence of H₂ gas, trapped at the interface. These results have been repeated several times subsequently utilizing the same technique.

The 1994 publication by Q. Y. Tong et al. marks the next significant stage in the wafer bonding literature. The paper began by studying room temperature bonding through long-term storage of some hundred hours. This led to the observation of the formation of annealing voids and prompted discussion on their nature and attempts to suppress them (Figure 9).

Although the discussion focused on the long-term storage of wafers, the two significant contributions from this paper were the first graph of temperature-dependent bond strengths of both hydrophilic and hydrophobic bonded wafers, and the first detailed focus on annealing voids.

Annealing voids had been noted from the very early publications and consistently appeared in high densities on activated wafer pairs that had been annealed between 100–900°C, or stored at room temperature for hundreds of hours. It was hypothesized by the authors that the voids were a result of hydrogen released during the bonding reactions and the products of broken-down hydrocarbon surface contaminants.

Two approaches were attempted in order to avoid the formation of annealing voids. The first was to use one or more thermally oxidized wafers as part of the wafer sandwich. It was well understood that excess water in the interface after annealing at >110°C tended to oxidize available bulk silicon or hydrolyse oxide, both actions resulting in a release of excess hydrogen. By using 1 μm thick thermal oxide coated wafers the intention was for the porous thermal oxide layer to adsorb this excess of hydrogen. The second approach was to anneal wafer pairs exhibiting voids at 1000°C for 1h. This allowed the trapped hydrogen to diffuse away from the interface causing the bulk of the voids to close.

The authors also present bond strength measurements taken at regular temperature intervals for wafer pairs being annealed at various temperatures up to 900°C. From this the much-reproduced graph of bond strength as a function of annealing temperature was created, as shown in Figure 10.

Without a detailed discussion on the testing procedure the precise shape of the bond versus temperature curves and bond energy levels achieved cannot be determined. However, irrespective of the

Figure 9. Example of annealing voids reprinted from Zhang and Raskin. This wafer was annealed for 120 h at 400°C.

Figure 10. Bond strength as a function of annealing temperature for hydrophilic and hydrophobic wafers.
absolute values, the general shape has been confirmed through many
subsequent studies and the graph re-printed in many publications.36–42
In order to confirm the hypothesis for the origin of the voids a small
specimen with a single annealing void was isolated.33 The upper wafer
was thinned repeatedly and the void ‘bulge’ measured. Once thinned to
2 μm the bubbles partially collapsed, indicating a release of internal
gas. This was explained as the hydrogen in the mixture diffusing
through the 2 μm silicon, leaving behind less mobile species such as the
hypothesized decomposed hydrocarbons.

The conclusions pertaining to annealing voids in this early pa-
er are supported and expanded upon by the 2010 publication by S.
Vincent et al.34 Their paper presents a detailed study on annealing
voids combining Fourier transform infrared microscopy (FTIR), X-
ray red fluorescence (XRR) analysis, Fourier transform infrared
microscopy (FTIR) and scanning acoustic microscopy (SAM).34 XRR shows the interface growing with
increased annealing temperature as excess water oxidizes bulk silicon.
The density of hydrogen evolved during annealing was measured di-
rectly via NRA and by calculation from the density of water required
for the increase in oxide thickness detected via XRR. This gave a figure
of 2 × 1015 H2 cm−2. SAM shows voids forming during low tem-
perature annealing and then dissolving as the wafers are heated above
900°C. Models are then presented along with experimentally derived
coefficients that describe the formation and dissolution of voids with
increasing annealing temperature. This process is shown to be driven
first by the evolution of hydrogen gas and then the diffusion of that
gas through the bulk silicon.

Most of the early work and analysis on the theory of wafer bonding
was eventually gathered together in the 1996 paper by Q. Y. Tong and
U. Gösele.45

In many respects this paper still remains the state of the art in
terms of hydrophilic bonding theory. The stages initially proposed in
the earlier Maszara paper are expanded upon and theoretical maximum
bond strengths are determined.

As before the bonding stages are characterized by temperature.

Stage 1 This occurs between room temperature and 110°C. It
is dominated by water-mediated hydrogen bonding of one or more
monolayers in the interface. By calculation, a relative humidity (RH)
of 50% leads to surface coverage of more than one monolayer while
at 1.4% RH coverage is 0.24%. Experiments showed that un-annealed
bond strength appears to fall when the humidity during bonding is be-
low 15% RH. The most stable form of adsorbed water in the interlayer
is as triplets, which allows for a total wafer separation of 10 Å to be
accommodated. By calculating from the total number of type A and
 type B OH groups the theoretical maximum interfacial bond energy
(γ) can be determined to be 0.165 J m−2.

Stage 2 Between 110°C and 150°C interfacial water becomes sig-
ificantly more mobile and adsorbed water is outgassed. Both of these
processes lead to a dehydrating of the bond interface and an acceler-
ation of silicon hydroxide to silicon oxide bond conversion. By 150°C
all OH groups in close proximity to other OH groups have converted
and the bond strength stabilizes. Again by similar calculation the
maximum theoretical strength is given as 1.674 J m−2.

Stage 3 Continued annealing from 150°C to 800°C does lit-
tle to increase the bond strength. Trapped water either hydroly-
ses bulk oxide or diffuses into and oxidizes the bulk silicon. The
latter process produces excess hydrogen, which causes annealing
voids to appear and grow during passage through this temperature
range.

Stage 4 Above 800°C bond strength is maximized by oxide flow
to fill interfacial voids. Annealing voids cease growing and shrink
with prolonged or higher temperature annealing. It is commented that
highly strained oxide films and films containing OH groups and water
can begin to flow at 800°C while less dense, more inert films such as
thermal oxide coatings require >1100°C. The theoretical maximum
bond strength is reported as 4.932 J m−2.

These theoretical maxima are dependent on the validity of the
calculated figure of type B OH groups per µm2, which as discussed above is a
thermal equilibrium value at standard room temperature and pressure
as opposed to an absolute. Nevertheless the details of the outlined

Development of Quantification and Testing

Accurate quantification of bond strengths has been a primary con-
cern of many groups and authors. The most widely used test methodol-
gy is the Maszara or dual cantilever bending (DCB) test first reported
in the 1988 Maszara paper.34

As already indicated there were some weaknesses in this original
report and the literature begins to address these by the late 1990s.
The 1997 paper by T. Martini et al.46 is the first coherent critique of
the Maszara test. The formula for γ is expressed as in equation (2), the
implications of stress corrosion noted and an attempt is made to create
a correction model that can remove the effects of stress corrosion on
a recorded crack to reveal the ‘true’ crack length.

In order to illustrate and quantify the effects of stress corrosion
a simple experiment is carried out in which a number of wafers are
bonded and then tested using the DCB method. The expected bond
strength is controlled by varying the annealing time and temperature.
During DCB testing a series of images is taken of the crack front at
various times after insertion. This results in a data set of crack length
as a function of time and bond strength, clearly demonstrating the
rapid rate of subcritical crack growth.

A theoretical treatment is then presented of the crack growth, fitting
it to an exponential function. However the key conclusion of this paper
is the advice either to standardize the time after insertion at which an
image is taken or to conduct the test in vacuum conditions. Unless one
of these conditions is met results are not comparable, either internally
in a data set or to others in the literature.

Further issues surrounding the DCB test were highlighted in the
2004 paper by Y. Bertholet et al.47

The paper specifically focuses on the influence of the razor
edge/wafer edge interaction and stress corrosion cracking. Silicon
wafers are hydrophilically activated using oxygen plasma and then
annealed at low temperatures. The samples are then sliced into strips
and tested using the DCB testing procedure in a jig.

Clear evidence is shown of the effect of the razor wedge in the
anomalously high initial bond strengths. Measurements are also made
of bond strength as a function of blade speed and crack growth with
time. As with the previous papers this demonstrates the weaknesses
in DCB testing and possible solutions to them.

All of this work is well reviewed in the authoritative 2005 paper
by O.Vallin et al.48

This review paper clearly and coherently describes all forms of
bond strength testing. Theoretical derivations are provided along with
comprehensive assessments of the strengths and weaknesses of each
technique.

With reference to DCB testing the Maszara formula is derived and
several alternative formulae, which account for lower-order effects
such as torsion and shear, are also presented. The systematic uncer-
tainties due to the razor wedge as well as limitations of the IR optics
are described. Stress corrosion is briefly described and the importance
of avoiding the effect is emphasized.

Since then a small number of papers advancing DCB testing have
been published, the first of which is the 2006 paper by L. Chen et al.49

This paper reprobes much of the earlier work done, noting the razor
bevel effect and reporting on a debonding jig. When a blade is initially
inserted into an interface the resultant crack front is not straight. A
non-straight crack front does not comply with the assumptions of the
Maszara formula and measurements of the maximum perpendicular
distance from blade to crack tip lead to an inaccurate value for the
bond strength (Figure 11).

The paper discusses this effect and formula 3 is derived for mea-
suring bond strength when the crack front is curved.

\[
\gamma = \frac{(5R - 2)E^*a^2w^2}{\lambda(64R)^{1/2}}
\]

[3]
Where R is the radius of wafers and λ is the ratio between an ideal crack and a real crack, as illustrated in Figure 11. Despite this no mention is made anywhere in the subsequent literature of waiting until the crack front has become straight before measuring. Additionally, many images in later papers show clearly non-conforming round fronts,40,50,52 Therefore the formula derived in this paper is potentially of considerable benefit.

A 2008 paper by K. T. Turner et al.53 also addresses this effect. A finite element model was used to analyze the crack front resulting from a wedge insertion. The results clearly show that the bond strength, calculated from a curved crack front using the Maszara formula, differs by +15 to –10% over the range of interface bond energies assessed. This variation is shown to result from the use of full plate wafers and is a good example of the effects of a non-ideal sample shape. The range of bond strengths investigated was well below that achieved by covalent bonding, being 5 mJ m$^{-2}$ to 125 mJ m$^{-2}$. However, by visual extrapolation of the graphs presented for high bond strengths it is clear that an overestimation of up to 25% can be expected for typical edge insertions into full plate wafers.

The influence of stress corrosion on DCB testing was first highlighted in the 1999 publication by M. Petzold et al.54 Two papers further addressing this phenomenon were published in 2012: F. Fournel et al.55 and V. Masteika et al.56

Both of these papers study the effect of ambient humidity on DCB testing. F. Fournel et al. demonstrate the use of an anhydrous nitrogen atmosphere DCB testing apparatus that is able to produce highly repeatable results. Using this setup, bond strength versus annealing temperature graphs are presented for Si - SiO$_2$ and SiO$_2$ - SiO$_2$ wafer bonds. These represent the first such graphs to be explicitly unaffected by stress corrosion. V. Masteika et al. present measurements of the very early crack growth and demonstrate a model of the rate-determining factors at each stage in the development of the crack, and a method derived from it that enables the estimation of initial bond strength values from late-stage crack velocity measurements.

These methods allow for a significant reduction in uncertainty when analyzing DCB tests. One remaining limiting factor is that the IR shadowgraph technique has a minimum void height resolution of one quarter of the wavelength of the IR used. Therefore there will always be part of the DCB crack that is undetected. O. Vallin et al.57 addresses this, presenting a brief analysis of this invisible crack and approximate associated uncertainties of 10% in the final bond strength.

An improved technique for interfacial void detection that could be used to reduce this uncertainty was reported in the 2009 paper by G. Horn et al.57 A new method of infrared imaging, IR photoelasticity, is presented and compared with SAM and IR transmission. It is shown to have the same resolution as SAM while taking approximately 100$^{th}$ of the time. It is also able to detect stress fields associated with voids. The new technique takes advantage of the phenomenon of stress modification of the polarization of light. Circularly polarized light is preferentially oriented depending on the magnitude of the stress, allowing variations to be detected.

As well as the DCB test, testing of samples on standard tensile testing machines is a common practice.17,58–64 For a successful test, the attachment area of the samples must be larger than the bond area, because wafer bonds are frequently stronger than adhesive bonds, even to epoxy resins. This is typically achieved by etching bond islands into one of the substrates prior to bonding.48 Unfortunately there is no way to compare most results reported using this method with DCB tests, as knowledge of the precise testing geometry is required. This can clearly be seen in the wide range of results reported. For example, a ‘well bonded’ sample can have a measured bond strength anywhere from 3 to 24 MPa, depending entirely upon the testing geometry.

Although there are some advantages to this method, it does not necessarily yield more reliable results than DCB testing. In the 2008 paper by K. S. Henriksen et al.55 direct and systematic comparisons are made between pull testing and DCB testing. The conclusion reached is that neither can be relied upon to provide certain, repeatable measurements without careful experimental procedures and enough tests for statistical significance.

The 2008 paper by R. Knechtel66 describes a systematic solution to this issue. It is proposed that any stud pull data set consist of at least 30 samples and that the results can be well understood by plotting on a Weibull plot. It is then shown that this approach produces good repeatability despite significant uncertainties in individual tests.

In order to overcome the problems inherent in both common testing methodologies (which were apparent to regular practitioners well before they were highlighted in the literature) the hybrid, micro chevron test, was developed.

The test was introduced in the 1999 publication by M. Petzold et al.54 Three distinct contributions were made to the field in this paper. It contains the first description of long term stress corrosion in direct bonded wafers; the chevron test is developed, and bond strength variability across the wafer is plotted explicitly for the first time.

The primary concern of this paper was with the failure of direct bonded devices over time due to stress corrosion. Slow, long term, stress corrosion is well known in the materials world and so experiments were conducted to identify the effect in the interface between direct bonded wafers. Two methods were used; DCB and the novel chevron test. DCB testing was used in two ways, firstly to determine initial bond strengths and secondly to determine the rate of crack growth with time.

The chevron test involves etching a chevron pattern into one of the wafer pairs and is illustrated in Figure 12. After bonding, the wafer pair is diced into squares, each containing one chevron. The samples are then pulled apart, with cracks initiating at the chevron tip, where the stress is highest. Data is extracted using the force and displacement measurements from the tensile tester. In addition to the low uncertainties of this method, failed tests are very easily detected from the shape of the load versus time plot. These two factors allow overall uncertainties to be significantly reduced.

In addition, by recording the original position of each sample a 2D map of the bond strength of the wafer pair can be produced. This was done in the publication and the presented map demonstrated

---

**Figure 11.** Diagram of a typical real initial crack (a) compared with an idealized initial crack (b). The crack length (l) indicated does not represent the bond strength when analyzed using the Maszara formula.
significant bond strength variability. This type of measurement has not been subsequently repeated.

In order to test long term reliability, stresses well below the critical stress required to initiate debonding were applied to the mounted samples. By measuring time-to-failure with the application of various subcritical stresses, the mean lifetime of samples was estimated and the effects of subcritical crack growth detected.

The reliability of the chevron test has subsequently led to its becoming the ISO standard for bond strength testing. However, its increased complexity (necessitating as it does the prior patterning and etching of one of the wafers) places it as a complementary procedure to DCB rather than a replacement for it.

The highlighting of wafer bond strength variability across a single wafer does present a universal problem in bond quantification that should be taken into account in any testing regime. Zhang et al. acknowledge the phenomenon, presenting uncertainties for DCB testing where a blade is driven across 35 mm of sample. The authors attribute much of this variability to voids but otherwise present few thoughts. It is worth noting that the phenomenon has rarely been mentioned since.

A subsequent paper in 2001 by the same authors reprises many of the same topics but focuses in greater detail on stress corrosion cracking. Plots of the log of crack front speed as a function of applied stress are presented, clearly showing differing regimes, as can be expected from the general stress corrosion literature.

Micro-chevron testing is the current optimal measure of bond strength, but the requirement for patterning of wafers prior to bonding ensures that DCB testing is still widely used. Stud pull testing is also common but suffers from requiring the patterning of micro-chevron testing while being difficult to compare to the wider literature.

Although there is a literature pool of data for DCB testing much of it must be considered unreliable. The combination of stress corrosion, bond strength variability and numerous opportunities for systematic error mean any report should be approached with caution.

Despite these difficulties all three of these testing regimes can be successfully used provided that a robust experimental procedure, including statistically significant sample rates, is adopted.

**Physical Deformation Bonding Model**

The theory of chemical bond formation is only part of the bonding story. Wafers are never perfectly flat and have height variations on all length scales. The chemical bonding provides a force but some of that force has to overcome the stiffness of the wafers, such that their surfaces come into intimate contact across a plane which is the average of the two forms (weighted by the stiffness of each wafer).

The first comprehensive study was made in the 1995 paper Q. Y. Tong and U. Gösele. This summarized and expanded upon earlier efforts by the same authors.

Part of the appeal of direct wafer bonding is the ability of the initial wafer interlayer bonding to pull imperfectly flat wafers into intimate contact without requiring outside forces. This effect is studied and quantified analytically and experimentally. This study focuses only on macro scale voids, with internal heights of the order of micrometers.

Two formulae are derived linking the four parameters of gap height, lateral extent of feature, bond strength and wafer thickness that predict whether a gap will be closed. The sample geometry is illustrated in Figure 13.

Formula 4 describes the conditions for gap closing for gap heights \((u)\) of less than twice the wafer thickness.

\[
    u < 2.6 \left( \frac{l_y}{E^*} \right)^{1/2} \quad [4]
\]

While formula 5 describes the condition for gap closing for gap heights greater than twice the wafer thickness.

\[
    u = \frac{l^2}{\sqrt{1.2Ew^3/\gamma}} \quad [5]
\]

Variations are also presented for wafers of dissimilar thickness and Young’s modulus. Formula 6 is also derived to deal with isolated asperities, such as oxide islands or particle contamination.

\[
    u > 3.7 \left( \frac{w\gamma}{E^*} \right)^{1/2} \quad [6]
\]

In order to test these formulae a number of similar samples of differing thickness are activated hydrophilically and the bond strengths achieved are found to confirm the theoretical predictions.

Similar conclusions are reached in the 1998 paper by H. H. Yu and Z. Suo. By considering the wafer surface to be represented by a 2D sinusoid they present an analysis based on the strain energy field. Instead of separate formulae they derive a single condition dependent on the wavelength, amplitude, and bond strength wherein gaps will be closed.

In 2005 F. Rieutord et al. presented a theoretical paper that describes the bonding wave that propagates between two hydrophilically active wafers when they are contacted. When first brought together, wafers bonded in an environment at atmospheric pressure initially hover on a thin film of air, but when one part of one wafer attains contact with the other wafer via point pressure, a bonding wave initiates that spreads out rapidly over the interfacial plane and brings their entire surfaces into intimate contact. Instead of focusing on the wafer profiles, the study models the speed of the initial bonding wave as a function of bond strength, wafer elasticity and mean free path of any trapped gas.

The ability to theoretically determine the bulk parameters for bonding is important in the successful design of bonded devices. However in practice, variations in surface form over large length scales comparable to the wafer diameter are almost always successfully overcome when bonding 0.5 mm thick unpatterned wafers. Of greater importance to the final bond strength is the wafer nano-topography.

The 2003 paper by N. Miki and S.M. Spearing compares AFM studies of wafer nanotopography with room temperature bonding. Instead of attempting to quantify bondability as a function of the average surface roughness they suggest the use of the bearing area.
This is the area of a sample that lies some defined depth (the bearing depth) beneath the maximum asperity (Figure 14).

A clear relationship between the bearing ratio and the resultant bond strength is determined. The initial water bond of hydrophilically activated wafers was shown to maximize at a bearing depth of 1.4 mm, consistent with the theory posed by Tong et al.45 In addition it was shown that a slight increase in this fine scale roughness can decrease initial bond strength by a factor of up to two.

The same group published a follow-up paper in 200574 studying the effect of nanotopography. Features of the order of 10 nm in height and of 1–10 nm wavelength were measured and their effect on bondability assessed. This is done for three different industrial wafer polishing techniques. A theoretical model is developed to link topography to energy of deformation. Entire wafers are then measured using laser interferometry and the bond energy required to deform the wafers into intimate contact calculated. In all cases the recorded topography with wavelengths down to 1 mm requires less than 1 mJ m$^{-2}$ to deform to fit. By comparison, hydrophilic wafers typically exhibit an initial bond strength of the order of 100 mJ m$^{-2}$. Larger scale wafer effects such as bow and warp require tens of mJ of strain energy to overcome.

A further paper by K.T Turner and S.M Sperling75 published in the next year develops the analytical model the group had published and uses finite element analysis to calculate the bond area achieved when contacting wafers of arbitrary surface geometry. A number of specific cases of wafer geometry pairs are calculated, and then physically fabricated and the computer model results tested experimentally. The results show agreement between predicted and recorded bond strengths of ±10%.

The 2005 paper by L. Nie et al.76 presented an experimental solution to bonding insufficiently flat wafers. Large areas of these wafers were not brought into intimate contact by the room temperature water mediated bond. The poorly bonded pairs were then clamped during annealing which significantly increased the area brought into intimate contact. This allowed for more covalent bonds to form which left the wafer pairs more strongly bonded.

A similar process was reported by L. Y. Huang et al.77 in 2011. Using a NX-2000 Nano Imprint System, a layer of polymer was applied to the top and bottom of the wafer sandwich. This allowed for a uniform pressure to be applied to the entire wafer pair. It was found that an applied pressure could aid in bonding and that a homogeneous approach aided uniformity of bond strength. However it was also found that the bond strength was relatively insensitive to the applied load, with no systematic change reported between 1 and 500 psi (7–3500 kPa).

In 2006 F. Rieutord et al.78 published a paper featuring two interesting topics. The first is a theoretical prediction of bond strength as a function of surface roughness. The second is a description of the use of X-ray reflections to measure the density and thickness of the interface. The surface roughness model uses AFM-derived values for surface peak density, height and radius of curvature and the energy per unit area of bonding to calculate the actual area in contact and the resulting bond strength. This model is found to be in good agreement with density measurements of the interface at room temperature and during annealing and blade insertion tests of bond strength.

The next published model of surface roughness versus bond strength was the 2009 paper by Z. Tang et al.79

This paper presents another model for wafer conformation during bonding based upon a sinusoidal surface topography. A single coefficient that is a function of feature height, length scale and wafer thickness is determined and compared with experimental results. This coefficient is proposed as a predictor of whether the resulting bond will be perfect; void-filled, or wholly unsuccessful. The group published a follow-up paper the next year.80

G. Liao et al81 published a further model in 2011 wherein the surface roughness was described by a Gaussian distribution of asperities. This model includes both elastic and plastic deformation and was used to calculate mean surface roughness parameters that could be closely correlated with experimental results.

A computational model of wafer bonding has been developed by D. V. Kuha et al. between 2006 and 2009.82–84 The initial stage of wafer bonding via water interlayer is modeled using a cohesive zone model. The parameters for this model are derived from first principles using a molecular dynamics simulation of the bond interface. The model itself deals only with flat plane, hydrophilic wafer bonding. However, the authors are confident that further development would lead to a more general model able to predict the effects of wafer surface roughness and patterning on bondability.

**Plasma Activated Bonding**

Initial hydrophilic activation has conventionally been achieved by the use of a chemical dip, the so called wet activation technique. Without a high temperature anneal (>800 °C), this method generally produces low bond strengths.

Although not the first paper to report plasma activation, the 1995 publication by S. N. Farrells et al.85 has particular significance. One of the authors, Roberts, also filed a patent that provides the basis of many commercial plasma-bonding machines.

Native oxide and thermal oxide silicon wafers were activated for 5–10 seconds in a 100 W O$_2$ plasma created in a reactive ion etcher (RIE). RIE systems are not simply plasma generators but accelerate charged species onto the substrate under treatment by virtue of a potential gradient, typically achieved by a bias voltage applied to the plate the substrate rests on. As such, plasma activation using an RIE can be thought of as ion bombardment.

The system had a base pressure of 15 mtorr and an oxygen gas flow rate of 6 sccm, implying a treatment pressure of several tens of mtorr.

Wafers were treated and then contacted ex-situ. No annealing step was undertaken before testing. Quantification was achieved using stud pull testing and was therefore not directly comparable with other reported bond strengths. However, it presented comparisons with wet activation that indicated a significantly improved bond strength. In addition, bond strength was shown to increase steadily over the course of 1000 hours of storage at room temperature. Images of the debonded samples reveal significant ‘pull out’ of material from the bulk wafers, indicating high strength bonding had been achieved.

The improved pre-anneal performance of plasma activation was attributed to a combination of a more active oxide, a buildup of a charged layer, and modifications to surface morphology. TEM studies were performed of the interface, detecting no voids. However, the resolution was of the order of 10 nm and so was incapable of detecting any voids created by nano scale features.

The 2000 paper by P. Amrifeiz, et al.14 provides the first detailed analysis of the effect of plasma activation on oxide surfaces.

This work revisited that done by S. N. Farrells et al.85 characterizing the surface and interface of plasma activated silicon wafers. An AFM was used for topographical analysis and various electrical measurements made of the activated layer. The existence of a highly charged layer detected in the previous paper was disputed and the first evidence for the deleterious effect of overexposure to plasma activation presented.
Samples were prepared using either an oxygen or argon ICP RIE in which charged species were accelerated toward the sample by a voltage on the base platens. Treatment times ranged from 10−240 s.

It was found that AFM measurements were possible only after a de-ionized water dip or storage of 24 h as otherwise the tip stuck to the surface. Those measurements that were possible showed that the surface roughness increased with plasma power. DCB testing of bonded samples activated for increasing periods in the plasma showed that longer activations resulted in inferior bond strengths. Voltage versus capacitance measurements of the interface, performed by depositing an aluminum layer on top of the activated surface, found no large charge in the oxide. This does not support the electric charge model of interfacial water removal proposed in Farrens' paper. The IR beam is channeled through prisms of high refractive index material which are placed in optical contact with the outer wafer surface. This avoids mechanical modification of the bonded wafers, which can prove problematic with poorly bonded specimens. The results correspond well with those of P. Amirfeiz, et al. and the authors are able to show distinct differences between the chemistries of the interface following the four different activation processes. In particular it is shown that H termination of the oxide is present after all treatments, but is reduced after O₂ ambient plasma exposure and further reduced after RIE. However the authors report a lower bond strength after RIE though this is attributed to increased void formation.

H. Moriceau et al. presented a paper in which X-ray reflectivity measurements are made of plasma activated hydrophilic wafer bonds. Notably their specimens achieved strong bonds after room temperature anneal. It was shown that more aggressive ion bombardment (RIE) techniques produce a low density interface layer that fills the gaps caused by surface roughness over the course of tens to hundreds of hours storage or low temperature anneal. This is not the case with chemical activation or ambient plasma, such as microwave generated plasma. It is also shown that RIE bombardment-style plasmas tend to smooth surface roughness while ambient plasma treatment does not.

There are many other papers concerning plasmaactivated wafer bonding, typically reporting on a new plasma generator source or an optimized set of parameters that allow for high strength bonding at low annealing temperatures and with minimal annealing voids. A significant number of these results are applicable only to the precise experimental setup of the reporting authors or are simply unrepeatable. This is partly a result of poor reporting of process and testing parameters but mostly due to poor DCB test methodology. Overall this body of work represents incremental steps in wafer bonding, particularly when reporting the bonding of novel materials to silicon.

The next significant stage in direct wafer bonding was the introduction of radical activation in a number of 2006 papers by M. M. R. Howlader, et al.

In order to minimize the subsurface damage, surface sputtering and bulk heating inherent to RIE exposure, wafer pairs were hydrophilically activated by a two-stage process of exposure to a short O₂ plasma followed by N₂ and N radicals. Wafers were then contacted and stored for 24 hours prior to testing to allow room temperature annealing to take place. This process was reported as improving bond strength and reducing void formation. It is difficult to assess the absolute bond strength achieved because such pull testing was used to make the measurements; however it was clearly shown that radial bond treatment allows for high bond strength with minimal RIE exposure. It was also found that N₂ radical exposure of up to 1200 s does not appreciably roughen the surface.

These were followed by a number of other papers, published between 2009−2011 describing further investigations using the sequential plasma activation method. These papers investigate the relationship between plasma parameters, surface roughness, interface thickness, annealing temperature, surface hydrophilicity and bond strength. It is shown that the optimum plasma parameters correspond well with the surface roughness of the samples. It is shown that a low powered RIE treatment initially reduces surface roughness before increasing it again, and that this minimum corresponds well with a maximum bond strength. One paper focuses on the nucleation and growth of voids activated using the sequential activation method. In particular it identifies nanopores (with approximate depths of 2.2 nm) that correlate well with increased void nucleation and increase in number with increasing plasma power.

In the same year as the initial Howlader reports S. Sood and R. E. Belford, reported purely radical activated wafers. They followed this with a more detailed paper in 2007.

Wafers are activated by the use of O₂, N₂ and Ar radicals created in an Astron remote RF plasma generator. As shown in this second paper the system is arranged with an electrostatic filter to remove any charged particles and allows only radicals to reach the wafer surfaces. The wafers are then contacted ex situ and bonded by a low temperature anneal of 150°C. Bond strength is then determined by DCB.
testing. The primary conclusion is that radical activation produces high strength bonds irrespective of the radical species.

2008 is the year of the first full paper publication detailing the radical (or RAD) ring by T. Rogers and N. Aitken.113 The RAD ring itself – which is a structure designed to allow only the neutral species created in a discharge to reach the wafer surface – is described and initial results are published. These conclude that the RAD ring is capable of producing bonds of equal strength to plasma activated wafers. They show no radial variance in bond strength and this empirical result is supported by a theoretical discussion that indicates that the average rate of diffusion of oxygen radicals in the chamber atmosphere is approximately 1 m m⁻¹.

The 2009 follow up paper by J. Kowal et al.31 repeated much of the background. In order to isolate the effect of radical treatment XPS and AFM studies were carried out. The unannealed void density was also monitored using SAM.

AFM studies clearly showed no increase in surface roughness after RAD activation irrespective of treatment time. By comparison, plasma activated wafers were shown to have a time dependent increase in surface roughness. XPS studies were able only to detect a single change in the surface after radical treatment. A small fluorine peak, possibly a remnant of wafer polishing, was removed by radical activation.

Developing the Bonding Model

Alongside the broad themes this review identifies, there are many papers reporting results and studies on wafer bonding that provide evidence for the models and hypotheses presented in the more significant papers.

The 1995 paper by J. Jao et al.58 uses high resolution TEM to study cross sections of wafers bonded using different technologies. Wet activated hydrophilic and hydrophobic wafers are bonded and annealed at temperatures ranging from 100 to 1200 °C and various techniques are used to analyze bond strength and the nature of the interface. Of particular note is the use of high-resolution transmission electron microscopy to inspect the bond interface after annealing at 300 °C. This provides the first evidence of microvoids.

With the benefit of hindsight, many of the conclusions reached concerning wafer bonding have been shown to be unrepeatable. This is particularly evident in low sample rate studies undertaken using the DCB test due to the unappreciated uncertainties described earlier.

The 1998 paper by Q. Y. Tong et al.37 is a prime example of this. This paper investigates the hypothesis that air, trapped during the ex situ post-activation contacting of wafers, reduces ultimate bond strength. A large number of samples are prepared, some contacted in situ, some ex situ and some contacted in situ then stored or annealed in a vacuum. All samples are then annealed at 150 °C and the bond strength sampled at 20 hour intervals up to 120 hours. Results then show that vacuum contacted or vacuum stored wafers achieve an almost bulk strength bond after approximately 100 hours of annealing. Ex situ bonds however only achieve a quarter of maximum bond strength after annealing. This strongly implies a deleterious effect from trapped residual atmosphere that cannot easily diffuse away through the interface while the wafer is at ambient pressure.

However one year later, in what was in essence a review paper,38 the authors report that the remarkably high bond strength achieved during in situ bonding could not be repeated. This implies either a highly unusual wafer configuration or experimental error.

The 2003 paper by A. Reznicek et al.114 advances the physical details of bonding flow. TEM studies of hydrophilically activated wafer pairs annealed for two hours at 450 °C show the interface to be filled with nanometer scale voids, with lateral dimensions of approximately 15 nm. Approximately 40% of the interface is unbonded in these samples. Subsequent high temperature annealing at 1100 °C causes these voids to coalesce and coarsen, though the total unbounded volume remains constant. At above 1100 °C the voids disappear as the oxide interlayer flows.

From other papers29,115 it is known that wet activation produces a rough island surface geometry substantially different from that produced after plasma activation. Therefore this paper does not directly relate to the interface of plasma activated wafers but does add weight to the hypothesis that the initially weak bonding achieved using wet activation is due to comparatively low contact areas.

Infrared spectroscopy to observe the interfacial chemistry of direct bonded wafers was initially proposed in 1994.39 The 2004 paper by C. Himetshle et al.116 reports a detailed study, using multiple internal transmission and reflection, of wet and plasma activated silicon wafers. After activation and contact, measurements are taken on samples as they are annealed at temperatures ranging from 200 °C to 400 °C.

Wet RCA activation, ‘mild’ oxygen plasma activation and more intense RIE activation were all studied. Of these, oxygen plasma activation leads to the highest bond strengths and the lowest density of annealing voids. The chemical reactions proposed in earlier publications3 can be observed, as can distinct differences between the three treatment types. In particular, in the temperature range observed, the consumption of interfacial water through migration through thin native oxide and reaction with the bulk silicon can be observed to occur at far greater rates for wet and RIE activation.

RIE activation of native oxide - thermal oxide wafer pairs exhibited many of the characteristics previously reported for RIE activation. That is a disordered damage layer able to flow during low temperature or room temperature annealing to maximize contact area. Of particular interest is evidence that such a layer has an OH group density greater than 4.6 nm⁻².

The same group prepared a second paper in 2006.86 This paper had a particular focus on the structure of the oxide interlayer during the annealing process.

As before, the interfacial chemistry of a number of samples is observed using infrared spectroscopy. Growth of the oxide interlayer is observed as well as relaxation of oxide tensions during annealing. Wafers are activated using one of three methods: ‘wet’ chemical activation, O₂ plasma treatment, or NH₃ plasma treatment. After activation the differing chemical evolutions of the three types of bond interface are observed during annealing. These observations reinforce the conclusions reached in the previous paper and highlight the evolution of hydrogen during the annealing process.

Detailed IR and HRTEM studies of the interface showed an approximately 5 nm thick SiO₂ layer bounded by ~0.7 nm thick SiO₂/SiO transition layers in an approximate 40:60 ratio. The SiO₂ layer was then shown to grow with increasing annealing. This layer is then shown to be less strained after this growth.

As the bonding process becomes better understood, increasingly subtle efforts have been made to address the issue of void formation. The research path followed by Tong et al.115-119 attempts to create an oxide layer more conducive to the diffusion of polymerization products.

Fluorine is introduced into a thick oxide layer, reducing its density, and some of the OH groups are replaced with NH₂. The fluorinated oxide allows water and hydrogen to diffuse out more easily from the interface at room temperature, and the NH₂ terminators polymerize into H₂ and Si-N-N-Si interfacial bonds, avoiding the creation of difficult-to-diffuse water.

C. Wang and T. Suga also investigated the use of fluorine in a 2012 paper120 with similar conclusions: namely that a fluorinated native oxide is conducive to high strength bonding with reduced voids due to its ease of permeability by water molecules. Although an RIE device is used, the plasma was only ambient, as confirmed by the negligible increase in surface roughness for all treatments.

Attempting to address the same issue, the 2008 paper by C. Ventosa121 studies the effects of annealing a wafer in air post-activation to suppress void formation.

Downloaded on 2014-05-07 to IP 137.108.145.39 address. Redistribution subject to ECS terms of use (see ecsdl.org/site/terms_use) unless CC License in place (see abstract).
Observations are made using X-ray reflectivity and IR absorption data. A post activation anneal converts the native oxide into thick, less dense thermal oxide. As previously described, this provides a sink for reaction products and contains significant numbers of internal type A OH groups. It also prevents the dissociation of water as it cannot diffuse through the thick film. This paper provides direct experimental evidence for these processes.

Unsurprisingly the recorded bond strengths for wafer pairs where the component wafers have been annealed after activation, but prior to initial room temperature bond, are low. It has previously been demonstrated in Ref. 9 that the type B OH groups that make up 70% of surface OH groups at room temperature are in thermal equilibrium and thus will decrease in number if the wafer is heated. Therefore a post-activation anneal at 150 °C will serve to reduce the number of type B OH groups, leading to fewer interfacial bonds and a weaker final bond strength.

One final paper, also published in 2010 by D. S. Grierson and K. T. Turner provides a unique insight into the very initial stages of wafer bonding.

A novel experimental approach is developed to measure the work of adhesion and debonding of hydrophilically bonded wafers. Micromachined silicon beams are allowed to bond to hydrophilic silicon wafers at room temperature. One end of each beam is attached to a mobile stage, allowing the separation to be controlled. From this it is shown that the work of adhesion is significantly less than the work of separation. This is consistent with the initial work of adhesion being done primarily by interfacial water clusters. These serve to bring parts of the wafers into intimate contact, which will over time allow the formation of room temperature siloxane bonds that increase the total bond strength.

Summary and Wafer Bonding Model

A thorough assessment of the wafer bonding literature reveals the following model of hydrophilic wafer bonding:

The silicon dioxide surfaces of silicon wafers are hydrophilically activated by covering them with hydroxyl groups. These are either thermally stable isolated type A groups or thermally unstable, paired type B groups. When two wafers are placed into close contact a water interlayer, forming clusters of three molecules, joins hydroxyl groups up to 10 Å apart. This allows wafers that are not atomically smooth to bond. Wafer surfaces contain features on all length scales. Depending upon the Young’s modulus, thickness of the wafer and ratio of feature length to scale, the energy of the pre-bond may be enough to elastically deform the wafers into conformity; otherwise there will be unbonded regions.

Hydroxyl groups in close proximity will polymerise, producing water, until thermal equilibrium is reached. This process can be accelerated by a thermal anneal. As bonds begin to form across the interface the increase in contact pressure drives the water interlayer out. This allows more bonds to form until all type B and any type A groups brought into close proximity have been converted into siloxane bonds. This occurs at approximately 150–200 °C and results in a maximum bond strength of 1.6 to 2 J m⁻², depending on the number of bridging siloxane bonds.

When a critical number of bonds has formed, water is no longer able to diffuse along the interface. Instead, it either reacts with the bulk oxide to form hydroxyl groups or migrates through the oxide layer to react with the silicon crystal substrate producing siloxane and molecular hydrogen. These two processes occur primarily in the temperature range of 200–800 °C. The hydrogen produced through these processes is unable to diffuse through centimeters of interface to the edge of the wafer and so becomes trapped in micro voids. Micro voids occur where the initial bond energy was insufficient to bring the wafer surfaces into contact. During annealing the voids fill with hydrogen and any other trapped decomposed gas and the pressure forces them to grow. Once they become detectable these voids are known as annealing voids.

When heated above 800 °C the silicon dioxide begins to flow, filling any remaining voids and maximizing bond strength at around 4.9 J m⁻².

Wafers can be hydrophilically activated in a number of ways. Wet activation involves dipping the wafer surfaces in chemicals. However, depending upon the specifics of the chemical treatment, island features are created that may result in approximately 40% of the interface remaining unbonded after an anneal at 200 °C.

Ambient plasma activation is very efficient at covering the wafer surface in hydroxyl groups. However, prolonged exposure produces an atomically roughened surface which cannot be sufficiently deformed or bridged with siloxane bonds.

RIE plasma activation creates a low-density, highly active oxide layer that is covered with a high density of hydroxyl groups. This low-density layer plastically deforms after prolonged room temperature storage or a low temperature anneal to fill micro voids resulting in extremely high strength bonding of more than 2.5 J m⁻². In native oxide covered wafers the high energy bombardment creates a damage layer in the silicon, and this, combined with the low density oxide, allows interfacial moisture to easily oxidize the crystalline silicon, releasing large quantities of hydrogen at annealing temperatures above 200 °C.

Radical activation efficiently covers a siloxane surface with hydroxyl groups but causes almost no topographic changes to the surface. This implies a process more dependent on initial wafer topography than plasma or wet activation.

Annealing voids can be avoided by providing a sink for the hydrogen produced during annealing. This can be achieved through a thick, low density, thermally stable oxide layer. Attempts have been made to remove excess moisture from the interface prior to the formation of hydrogen, but as the final bond strength is a function of the density of water and hydroxyl groups there is limited room for improvement using that method.

Wafer bond strength can be determined by a variety of physical debonding tests. The three most common tests are the DCB test, the micro chevron test and the stud pull test. The DCB test is the most widely used as no pre-patterning of the substrate is required. However care is needed to avoid large uncertainties. The micro chevron test is the most certain. However, it requires a tensile testing machine and pre-patterning of the wafers. The stud pull test has significant uncertainties and requires pre-patterning and a tensile tester and therefore is the least widely used, and is mostly confined to older publications.

Due to the effects of nanotopography, bond strength can vary across a wafer and so care must be taken, irrespective of the testing procedure, to produce a statistically significant number of tests at any data point to reduce and quantify the uncertainties.

Multiple internal reflection infrared absorption spectroscopy can be used to observe interfacial chemistry. Infra-red transmission images can be used to detect interfacial voids, higher than 250 nm. Scanning acoustic microscopy or infrared phase elasticity can be used to detect voids over 50 nm tall. Transmission electron microscopy can be used to study prepared slices of the interface. Ellipsometry can be used to monitor oxide growth on an unbonded sample. Atomic force microscopy or laser interferometry can be used to provide topographic information.

Hydrophilic wafer bonding of surface oxides or nitrides has now been applied to a number of differing materials such as germanium, fused silica and titanium. While the precise parameters vary for each material, hydrophilic silicon wafer bonding provides a good general model for the bonding process. Many of the same bonding technologies and analytical tools are used and many of the same problems remain to be solved.

References