

ulm university universität
uulm

Interconnect Technology for Three-Dimensional Chip Integration

Dissertation

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
der Fakultät für Ingenieurwissenschaften und Informatik
der Universität Ulm

von

Andreas Munding

aus Riedlingen a. d. Donau

1. Gutachter: Prof. Dr.-Ing. Erhard Kohn
2. Gutachter: Prof. Carl E. Krill, Ph. D.
Amtierender Dekan: Prof. Dr. rer. nat. Helmuth Partsch

18.04.2007

Contents

List of Figures, Tables, Symbols **vi**

Summary **xi**

1 Introduction **1**
1.1 Motivation . . . 1
1.1.1 Multipurpose Technology Products . . . 1
1.1.2 On-chip wiring demands . . . 2
1.1.3 Three-dimensional integration . . . 3
1.1.4 Biologically inspired neural processing . . . 4
1.2 Interconnect needs for 3D-integration . . . 5
1.2.1 Peripheral Interconnects . . . 5
1.2.2 Area interconnects . . . 6
1.2.3 Through-chip vias . . . 7
1.2.4 Conclusion . . . 9
1.3 A new 3D-chip-integration concept . . . 11
1.3.1 Process chain . . . 11
1.3.2 Critical steps . . . 13
1.4 Subject of this thesis . . . 15
1.4.1 Scope . . . 15
1.4.2 Procedure . . . 17
1.4.3 Outline . . . 18

2 Thinning and wafer handling **19**
2.1 Introduction . . . 19
2.1.1 Concepts for waferbonding . . . 19
2.1.2 Concepts for wafer thinning . . . 20
2.2 Wafer bonding . . . 21
2.2.1 Permanent bond to handling substrate . . . 22
2.2.2 Evaluation of temporary bonding techniques . . . 24
2.2.3 Temporary bonding with surface topology . . . 26
2.2.4 Results . . . 28
2.3 Wafer thinning . . . 29
2.3.1 Mechanical wafer thinning: Lapping . . . 29
2.3.2 Chemical mechanical polishing: CMP . . . 31
2.3.3 Chemical thinning . . . 33
2.3.4 Results . . . 34
2.4 Conclusion . . . 36

3 Patterned electroplating and via filling **37**
3.1 Introduction . . . 37
3.1.1 Challenges . . . 38
3.1.2 Methods of electroplating . . . 39
3.1.3 Plating characteristics on plain substrates . . . 40
3.2 Through-mask plating . . . 41
3.2.1 Plating setup . . . 41
3.2.2 Patterning . . . 41
3.2.3 Seed layer removal . . . 43
3.2.4 Results . . . 44
3.3 Via filling process evaluation . . . 45
3.3.1 Factors influencing feature filling . . . 45
3.3.2 Via preparation . . . 47
3.3.3 Via filling . . . 49
3.3.4 Results . . . 49
3.4 Conclusion . . . 52

4 Bonding and chip-stacking **55**
4.1 Introduction . . . 55
4.1.1 Available metallic bonding techniques . . . 56
4.1.2 Solid liquid interdiffusion . . . 59
4.1.3 Choice of bonding method . . . 61
4.2 The SOLID process . . . 61
4.2.1 Cu—Sn intermetallic system . . . 62
4.2.2 Dynamics of solder layers . . . 64
4.3 SOLID bonding process evaluation . . . 66
4.3.1 Bonding with unstructured metal layers . . . 66
4.3.2 SOLID bonding evaluation with thinned chips . . . 68
4.3.3 Bonding with patterned area metallization . . . 69
4.3.4 Results . . . 70
4.4 Chip-stacking with through-chip vias . . . 73
4.4.1 Results . . . 74
4.5 Conclusion . . . 75

5 Scaling of microjoints **77**
5.1 Introduction . . . 77
5.1.1 Optimization potential in SOLID bonding . . . 77
5.1.2 Prerequisites for scaling of microjoints . . . 80
5.1.3 Kinetic control with metallic separation layers . . . 83
5.2 Phase growth kinetics with metal barriers . . . 85
5.2.1 Material selection criteria . . . 86
5.2.2 Solid state phase growth . . . 86
5.2.3 Results . . . 87

5.3 Wetting of barrier layers 91
5.4 Conclusion 92

6 Outlook **93**

A Experiments and samples **95**
A.1 List of samples and processed substrates 95
A.2 Mask Layouts 96

B Processes **99**
B.1 Process chain for wafer elements 99
B.2 Chip-stacking sequence for flip-chip-bonder 100
B.3 Lithography 101
B.4 Ion-beam sputtering 101
B.5 Lapping sequence 104

C Equipment **105**
C.1 Electroplating setup 105
C.2 Waferbonding presses 105
C.3 Thinning equipment 106
C.4 Flip-chip bonder 107

Bibliography **109**

Acknowledgement **119**

Curriculum vitae **121**

List of Figures

1.1 On-chip wiring problem: Interconnect delay dominates over all delay . . 2
1.2 Interconnect length for multi layer circuits 3
1.3 Biological vision system and its possible hardware replication 4
1.4 Examples of peripheral 3D-interconnect technology 5
1.5 Examples of area interconnect technologies 6
1.6 3D integration at Tohoku University, Japan 7
1.7 3D integration at Tsukuba Research Center 8
1.8 Vertical Chip Integration (VSI) at Fraunhofer IZM 9
1.9 Processes for a novel 3D-integration concept 12
1.10 Interconnect levels in integrated systems 15
1.11 Interconnect life-time and thermal expansion of substrate 16
1.12 Path for implementing high-density 3D-interconnects 17

2.1 Process paths for waferthinning and handling 19
2.2 Waferbonding with BCB 23
2.3 Bulk removal of handling substrate as a replacement for detaching 24
2.4 Handling wafer detachment by a two-layer wafer bonding technique . . . 25
2.5 Dependance of polymerization of BCB on the heating cycle 26
2.6 Weight loss of PMGI during hard curing at elevated temperature 26
2.7 Failed bond, evident after back-lapping 27
2.8 Analysis of failed bond interface 27
2.9 Temperature and Pressure during Waferbonding 28
2.10 Two-layer temporary waferbonding 29
2.11 Surface after lapping and defect depth 30
2.12 Dependence of the lapping rate on grain size and down pressure 31
2.13 Dependence of the lapping rate on the rotational speed 31
2.14 Reactions during chemical mechanical polishing 32
2.15 Etch rate of KOH solutions in boron doped silicon 33
2.16 Visual lapping endpoint detection 35
2.17 Alignment and lithography on the thinned and polished backside 35
2.18 Thinned and polished wafer element after via etching 35
2.19 Surface after chemical wet-etching with an etch-stop layer 36

3.1 Electroplating processes in the process chain 38
3.2 Patterning by through mask plating 42
3.3 Seed layer removal after through mask electroplating 43
3.4 Front and backside metallization patterns on wafer element 44
3.5 Experimental via-etching and inside-via processes 47
3.6 Via holes and resist mask test pattern prior to filling 49
3.7 Filled via and resist structure after Cu electroplating 50

3.8 Plating failure modes 50
3.9 Via-plugs after removing the Si substrate 50
3.10 Resistance of a single via and two vias in series 51
3.11 Leakage current between two vias and from via to front metal 52
3.12 Single via contact and leakage current levels 53
3.13 Backside plated metallization pattern with filled vias 53

4.1 High-melting C4 and low-melting BGA solder joints 57
4.2 Spalling induced dewetting of solder from UBM 58
4.3 Binary metal system for solid liquid interdiffusion 60
4.4 Phase diagram of the Cu—Sn system 63
4.5 Concurring dynamic processes during SOLID bonding 64
4.6 Conditions for SOLID bonding dynamics 65
4.7 Solder layer thickness window for SOLID processing 65
4.8 Cross section of a full area SOLID joint 68
4.9 Errors associated with flip-chip bonding 70
4.10 Determination of Cu-line resistivity and joint resistance 71
4.11 Area mapping of contact resistance 72
4.12 Distribution of joint resistance around the average value 72
4.13 Single joint cross section with tolerable displacement 72
4.14 Cu-lines and bridges after removal of upper Si-substrate 73
4.15 Chip stack with one stacked chip 74
4.16 Electrical characterization of chip-stack interconnects 75
4.17 Isolation properties of 3D-contacts on stacked chips 76
4.18 4-layer and 7-layer stacks of thinned chips 76

5.1 Current state of 3D-chip-stacking with SOLID joints 78
5.2 Passivation of SOLID-pad sidewalls 79
5.3 Asymmetric growth conditions of intermetallics 81
5.4 Simulated intermetallic growth during heating 81
5.5 Initial solid state phase growth for various scenarios 82
5.6 Effect of preheating temperature on intermetallic thickness 83
5.7 Phase growth characteristics for various barrier metals at 200 °C 87
5.8 Determination of the parabolic rate constant 88
5.9 Temperature dependence of the parabolic rate constant 88
5.10 Reference sample with virtually no barrier layer at the interface 89
5.11 Growth characteristics with Ni interlayer 89
5.12 Typical growth behaviour of samples with Ti interlayers 90

6.1 Chip-scale micro-assembly concept with SOLID stacking 94

A.1 Layout for face-to-face SOLID evaluation 96
A.2 Test concept for the characterization of multi-chip stacks 97
A.3 Layout for test structures in multi-chip stacks 97

B.1 Temperature and pressure profile during the process chain 100
B.2 Ion-beam sputter system for etching and deposition of metals 101

C.1 Electroplating cell with horizontal rotating substrate holder 105
C.2 Vacuum bonding with a metallic membrane chamber 106
C.3 Lapping machine PMA2 for mechanical thinning and polishing 106
C.4 Spin-etching setup for chemical thinning of wafer elements 107
C.5 Flip-chip bonding equipment for chip-to-chip stacking 107
C.6 Atmospheric control by bonding in a flow chamber 108

List of Tables

1.1 Examples of present day research on 3D integration 10
1.2 List of process steps and critical issues 14

2.1 Experiments with different adhesive layers on glass substrates 22

3.1 Available etch processes for seed layer removal 44

4.1 Properties of intermetallics in the (Cu/Ni)—Sn system 61
4.2 Advantages and disadvantages of metallic bonding techniques 62
4.3 SOLID experiments with an unstructured metallic system 66

5.1 Solubility of some metals in Cu and Sn 86
5.2 Summary of results from diffusion experiments 91
5.3 Wetting behaviour after heating to 250 °C 92

A.1 List of samples and experiments 95
A.2 Important samples and milestones 96
A.3 Available compositions of epitaxial wafer 96

B.1 Typical process chain for passive chip-stacks 99
B.2 SOLID bonding procedure with “Fineplacer” flip-chip bonder 100
B.3 Lithographic parameters 101
B.4 Parameters for ion-beam sputter etching and deposition 102
B.5 Sputter processes and deposition rates for various targets 102
B.6 Process steps for the preparation of diffusion samples 103
B.7 Reactive ion etching and ashing parameters 103
B.8 Rapid lapping sequence with 12 µm Al_2O_3 grains 104

Nomenclature

ASE Anisotropic silicon etching, advanced silicon etching, page 21
BCB Benzocyclobutene, low-κ resign with good properties for waferbonding, page 9
BEOL back end of line, usually all processes related to interconnects, after transistor fabrication, page 37
CMP Chemical mechanical polishing, the surface is treated by small abrasive particles in a surface altering colloidal solution, page 21
CVD Chemical vapor deposition, page 8
EDP Ethylenediamine pyrocatechol, page 34
FIB focussed ion beam, cross section preparation method, page 9
HARSE High aspect ratio silicon etching, mostly associated with Bosch/STS-type or cryo-type processes, page 21
IP intermetallic phase, intermetallic compound, alloy with a defined composition and crystal lattice, page 60
KOH potassium hydroxide, page 34
MIM Patterns with a metal—insulator—metal vertical structure, page 51
MIS Patterns with a metal—insulator—semiconductor structure, page 51
p^{++} very high dopant concentration, close to saturation, page 12
PEG poly ethylene glycol, water soluble polymer available in a variety of consistencies, according to chain length, page 46
PMGI Photo resist based on polydimethylglutarimide, page 24
RIE Reactive ion etching, chemical dry-etching, page 23
SEM scanning electron microscopy, scanning electron microscope, page 70
SOI silicon on insulator, dielectric substrate with a Si-layer on top, silicon substrate with a buried dielectric layer, page 9
TMAH tetramethylammonium hydroxide, page 34
UBM under bump metallization, metal layers in contact with solder, page 58
WLP Wafer level packaging, page 6
$\Delta t_{\mathrm{delay}}$ Growth delay after temperature step, page 91
Δt_{lin} Duration of barrier induced linear growth, page 91
Δx Total intermetallic thickness, page 80
Δx_{init} Initial intermetallic thickness at $t = 0$, page 91
Δx_{xx} Partial thickness of element or intermetallic compound xx, page 83
ϱ_{xx} Density of element or compound xx, page 83
A Sample area, surface area, contact area, page 31
D Diffusion coefficient, page 85
d Degrees of freedom, number of dimensions of growth, page 85
D_0 Temperature invariant factor of diffusion coefficient, page 85
g = 9.81, gravitational constant, page 31
K Lapping grain size of abrasive media, page 30
k_0 Temperature invariant part of intermetallic growth constant, page 80
k_{heat} Rate constant for linear temperature ramp, page 82

k_{lin} Linear intermetallic growth constant, page 85
k_{P} Preston's constant for material removal, page 31
k_{sqr} Parabolic rate constant for diffusion induced growth processes, page 80
m Growth mode for intermetallic growth, page 85
M_{xx} Molar mass, weight of $n_{\mathrm{A}} = 6.02 \cdot 10^{23}$ atoms of species xx, page 83
n Growth exponent for general growth characterization, page 85
n Rotational speed, revolution rate, page 30
p Pressure, force per area, page 31
Q_0 Activation energy in Jmol^{-1}, page 80
R $= 8.31\,\mathrm{JK}^{-1}\mathrm{mol}^{-1}$, Gas constant, page 80
r Rate, removal rate, etch rate, deposition rate, page 31
$r_{\varepsilon\eta}$ Intermetallic phase composition as a thickness ratio between ε and η phases, page 83
t Time, annealing time, processing time, page 80
T_1 Pre-heating (stand-by) Temperature for SOLID-bonding, page 82
t_{pre} Pre-heating time for SOLID-bonding, page 82
v Average sample speed during lapping, page 31

Summary

Vertical chip integration — also referred to as 3D-integration — is an emerging technology, where a higher interconnect density between chips can be achieved by extending on-chip and inter-chip 2-dimensional planar wiring techniques by the 3rd dimension. Thus, packaged chips need no longer be placed side by side, but can be stacked and interconnected using the chip area, rather than just the periphery. 3D-integration not only allows small footprint circuits, but also can be used to combine subsystems based on different technologies. Several strategies are known to achieve a vertically stacked system, mainly distinguished by scale and yield of the stacking process. This work is based on a chip-scale strategy where selected chips (known good dies) are stacked one at a time, thus achieving a high yield at a low through-put.

Today's most advanced vertical 3D-integration schemes typically rely on three key processes: Via formation, waferthinning and stacking of chips or wafers. These are associated with essential shortcomings which currently limit the possibilities of this technology. When vias are etched from the front-side, the dielectric of the on-chip interconnect metallization is perforated, and the via plugs limit degrees of freedom for routing and wiring. If wafer substrates cannot be thinned down sufficiently, achieving a high through-chip via density is challenging, for high aspect-ratio etching and filling is pushed to its limits.

This thesis addresses feasibility and scalability of a new 3D-integration approach, which allows via-etching from the backside. The proposed approach is based on thin substrates of ca. 10 µm and the backside etched through-chip vias have a moderate aspect-ratio (depth by diameter) of ca. 2. Thus, the demand on via etching and filling processes is low, but high for waferthinning. This high demand on the thinning process is met by introducing a buried epitaxial etch-stop layer which results in a defined substrate thickness and a smooth surface, without the need for polishing. Stacking and interconnecting of the chips is done by employing a solid liquid interdiffusion (SLID) technique, which allows a low bonding temperature below 300 °C and results in a high remelting temperature. The novel approach will allow thinner substrates, shorter vias and higher interconnect densities.

A two-layer chip-stack with a 10 µm thinned chip on top of a bottom chip was implemented and characterized electrically. The 5 µm sized through-chip vias had a minimum series resistance of 0.7 Ω when connected to an inter-chip joint. The resistance of one inter-chip joint was determined to be less than 0.1 Ω. A typical Cu line resistivity was ca. 0.4 Ω/mm at an average joined line cross section of ca. $10 \times 5\,\mu m^2$. Mechanically, chips could be stacked up to a count of 7 metallized layers (1 base layer and 6 stacked chips).

The following process details were addressed: Carrier substrate attach by temporary wafer bonding, with a bond stability under vacuum conditions of up to 300 °C; Two-step mechanical and chemical controlled wafer thinning with etch-stop layers; High-aspect ratio via etching with stable hard masks using a Bosch/STS-type etching process; Confor-

mal or near conformal inside-via deposition processes for isolation and seeding; Via-filling and contact patterning by electroplating; Seedlayer removal; Chip-to-chip alignment and bonding. The minimum features for joints formed using Cu–Sn solid-liquid interdiffusion were in the range of $10 \times 10 \times 10\,\mu m^3$. This was achieved by bonding a Cu/Sn metal pattern on the backside of a thinned chip to a Cu metallization on the front side of a base chip or thinned chip.

Further work addressed concepts for microjoint scaling. The fast intermetallic phase growth reaction between Cu/Sn layers, compared to the slow wetting reaction at an initially separated Cu–Sn interface where identified as factors which limit further joint down scaling. As a method to slow down intermetallic phase growth, and to allow thinner Sn layers, diffusion barrier layers are introduced at the Cu/Sn interface of the backside bond metallization system. The solid state intermetallic phase growth of various barrier metals was analysed using three-layered sputtered and plated metal films on Si substrate. The results at a growth temperature of 200 °C showed that in the presence of Cr (50 nm), Ni (50 nm), CrNi (30 nm), Al/Ni (25/25 nm) and TiN/Ti (13/13 nm) layers, no barrier effect was evident. The effect of a diffusion barrier was observed for time periods ranging from a few seconds to 530 s, depending on barrier material and thickness in layers of Ti, TiW and Ta, while intermetallic growth was suppressed for at least 3000 s with layers of WSi/Ti (13/13 nm), WSiN/Ti (13/13 nm), Ta (50 nm) and TiW (50 nm). Sn film delamination was observed with TaN, TiN and Al barrier layers.

It was demonstrated that it is possible to impose a controlled delay on intermetallic phase growth during solid-liquid interdiffusion bonding (SLID), by using thin layers ($<$ 50 nm) of Ti, TiW and Ta, where the phase growth reaction follows a linear time law for a specific amount of time. The duration of this time period depends on the barrier thickness and could be matched to meet the wetting conditions with thinner Sn layers in a SLID bonding process. This method might thus be suitable to open the ground for sub-micron chip-to-chip interconnect joints in 3D-integration.

Chapter 1

Introduction

Present day electronic systems are able to perform a variety of complex functions, attributed to their inner structure consisting of a network of highly specialized sensing, actuating, emitting, storing and computing devices. These are typically interconnected electrically and mechanically over several levels of hierarchy: Chips are packaged to components, then soldered onto a printed wiring board, which usually is the main systems carrier. Over the years, with rising complexity, the number of signal paths per component has increased, and today, performance is no longer exclusively limited by its components, but by the communication paths on and in-between its components. New solutions in interconnect technology are thus required.

1.1 Motivation

The above statements can be deducted from tendencies observed in present day's integrated circuits. To name but a few, the convergence of technologies and rising demands on on-chip wiring are discussed in the following.

1.1.1 Multipurpose Technology Products

Today — especially in mobile systems — market demands a high level of functionality in a small, portable, or even "hand held" application. On the electronic component level, this means, that as many functions as possible need to be integrated into one chip or into one package, since board space is very limited.

The problem that arises is, that such complex systems employ a variety of technologies, i. e. a signal processing unit, a storage unit and an RF unit, which are normally manufactured in a technology specific way using separate dedicated production processes and optimal host wafer substrates. At present, two main system integration principles are in practice.

One way is to extend the IC manufacturing processes in a way that allows a monolithic chip to host areas of different technologies. This is known as system-on-chip integration, SoC, and has the advantage of fast and efficient on-chip interconnects, as long as performance is not limited by long global interconnect lines on large chips (see next section), or by compromising with other technologies. The interconnects between the different parts of the system are on-chip, and do not need to be addressed further. Clearly such systems need to be designed carefully for their many interfering parameters, which define the combined implementation of multiple technologies. The controversy of such practice

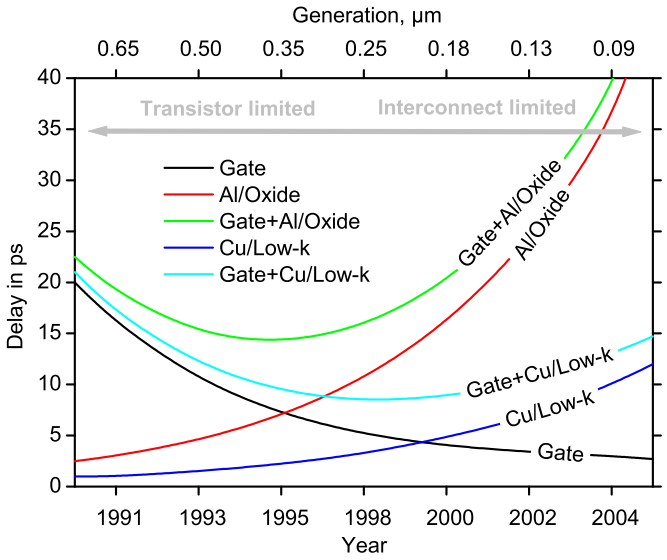


Figure 1.1: On-chip wiring problem: Interconnect delay time dominates over all delay time. Diagramm extracted from Sun et al. [1]

is a high development cost or long time-to-market, and yield sensitivity when increasing the number of process steps.

Another practice is to place chips from different source technologies as close together as possible and combine them into a heterogeneous system. This is generally understood as system-in-package, SiP, which summarizes all possible solutions of interconnecting chips in a hybrid package. This allows to develop and implement the different parts of the system independently or reuse them modularly in a flexible way, resulting in a short time-to-market. Clearly, performance heavily depends on how the chips are placed and interconnected inside the package, and there exists a need to develop efficient methods that result in short chip-to-chip interconnect paths with a low signal delay time.

1.1.2 On-chip wiring demands

In homogeneous chips, the observed tendency of increasing chip areas i. e. for microprocessors and shrinking transistor dimensions have lead to a turning point, where the performance is no longer determined by transistor figures, but by the delay of long (global) interconnect paths.

Figure 1.1 illustrates the projections of Sun et al. [1] from 1997, in which the signal delay time caused by on-chip interconnect parasitics, mainly resistivity and capacitance, will increase in future IC generations, whereas delay time caused by transistor gate parasitics will decrease. The introduction of copper [2, 3] as a replacement for aluminum interconnects and low-κ dielectrics [4] as replacement for SiO_2 has been beneficial to the overall delay time in integrated circuits, but has only postponed the raise of the delay time by a few years (see Figure 1.1). As stated by Tu [5], the RC-delay time is not the only performance limiting factor associated with on-chip interconnects, but also degradation by electromigration in present day copper interconnects.

A potential solution to serve future wiring demands is to shorten the overall intercon-

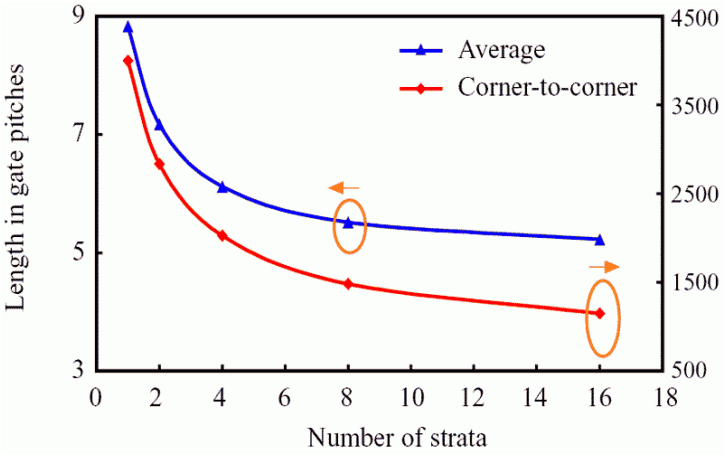


Figure 1.2: Interconnect length for multi layered (multi strata) circuits as modelled by Meindl et al. [6]. The occurrence of long interconnects drops significantly with more than 4 layers

nect length by partitioning a large chip into several smaller chips, which of course then need to be efficiently interconnected off-chip.

1.1.3 Three-dimensional integration

The shortest interconnect path from an arbitrary point on a chip to one on another chip can be achieved by the stacking of chips, thus using the third dimension for interconnecting.

Studies on stochastic models of different scenarios of multilevel (1 to 16 strata; stratum = level, layer) circuits with direct interlevel connection paths [6] lead to the conclusion, that the length of interconnects in only 4 layers of such 3D integrated circuits is already considerably shorter as compared to non 3D architectures. Shown in Figure 1.2 are the results obtained by Meindl et al. [6] for a model of a random logic network with ca. 4 million gates distributed and interconnected over 1 to 16 levels. The curves for average (mapped to the left axis) and corner-to-corner (right axis) interconnects show a similar exponential behaviour. The curves also imply, that 3D-integrating of more than 8 layers will have no substantial benefit on shortening interconnect lengths in random logic networks.

The random logic network model can be considered rather pessimistic for it does not represent a 3D-optimized architecture or does not take into consideration structures like buses and cell arrays which would probably benefit most from 3D-integration for their regular 1D and 2D patterns. Systems, where 3D-integration is considered most beneficial, are thus memory applications and image processing applications.

3D-integration is an emerging technology that has the potential to solve the on-chip interconnect problem by partitioning a large chip into smaller ones, which can then be interconnected vertically using the shortest path. It also has the potential to satisfy technology convergence needs by allowing heterogeneous integration of chips from different sources.

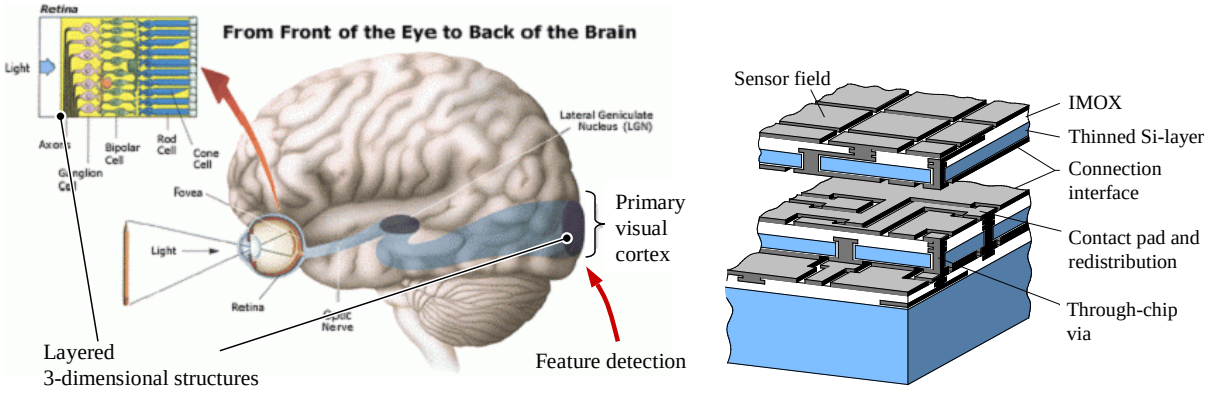


Figure 1.3: Biological vision system and its possible hardware replication as a stack of silicon chips bearing a neural network. Graphics from [9]

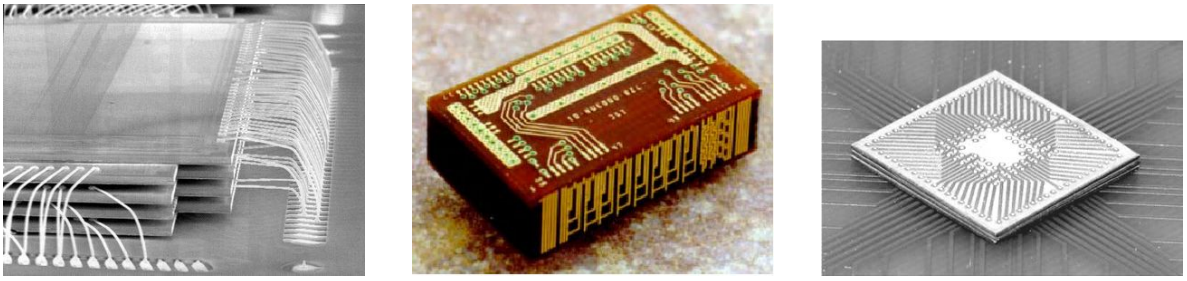
1.1.4 Biologically inspired neural processing

Today's need for 3D-systems mainly originates from two fields of application, memory storage devices and image processing devices. Emerging technologies, such as neural network computing devices have an even higher need for 3D-integration technologies, because as many nearest neighbour neural cells as possible need to be interconnected in order to function efficiently. Pulsed neural networks, as implemented and investigated at Infineon Corporation and Dresden University of Technology [7, 8] operate at a "low" frequency in the kHz range, where clock or signal interference and heat dissipation need not be considered.

The architecture of such a neural imaging system is biologically inspired (see Figure 1.3), and can basically be described as a meshed cube with knots that are in charge of specific functions, as to in which layer they are located. As a replication of the biologic visual system these layers can have the function of the retina (first layer: image sensor), ganglion cells (second and other layers: feature extraction) and cortex (last layers: image processing and storage). Such an architecture is easy to implement into hardware because it is straight forward and highly parallel with repetitive patterns (cells).

Because of its low operation frequency, neural processing is an evolving strategy for low power efficient computing in the domain of signal preprocessing. Most arithmetics in preprocessing and filtering of images consist of convolutional operations [10], for which a neural network with its inter and cross connections is an efficient architectural configuration. Research, related to the analog VLSI implementation of neural cells in an outer context of this work, has been conducted at the University of Dresden [11], supported by simulations and coordinated from Infineon Corporate Research.

From an application point of view, motivation of work related to this thesis is thus to provide a technological platform for the implementation of a 3-dimensional neural network, which replicates a biological vision system. From a technological point of view, motivation is given from a converging technologies background, where a sensor technology



a) b) c)

Figure 1.4: Examples of peripheral 3D-interconnect technologies: a) Peripheral wire bonding by STMicroelectronics [12], b) NEO stacking by Irvine Sensors Corporation [13], and c) Chip-in-polymer by Fraunhofer IZM Berlin [14]

needs to be combined with a signal processing technology, no matter what technology is host to the system's components.

1.2 Interconnect needs for 3D-integration

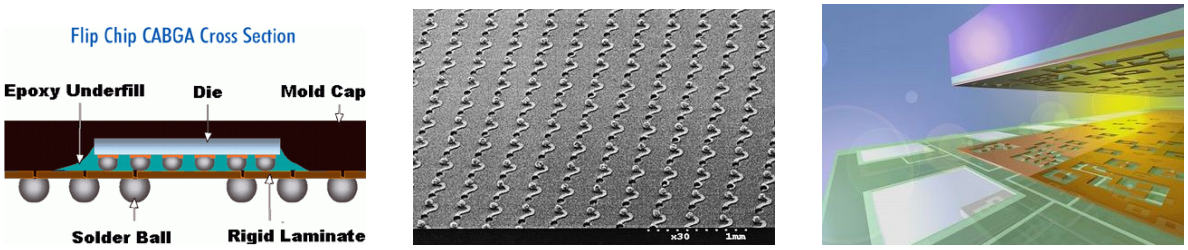
The term interconnect is generally used for defining an electrical connection between two functional entities within a system. It mostly refers to on-chip electrical connections or to chip-to-substrate connection. In evolving 3D-integrated systems, interconnects will also be formed between several levels of stacked chips, thus resembling 3D-interconnects.

3D-interconnects can be classified in various ways. The classification used in this thesis, will be whether interconnects are made within the area of the chips (area-interconnects) or whether the chips are interconnected only at the periphery (peripheral interconnects). A further instance of 3D-interconnects can be made up by interconnects that comprise through-chip vias and inter-chip joints.

1.2.1 Peripheral Interconnects

Classically, chips have been connected to other chips or to a substrate at the chip's periphery, i. e. with bond wires or tape (TAB, tape automated bonding). Recent innovations in 3D-integration have made it possible to produce peripheral wire bond interconnects on thinned and stacked chips, as for example presented by STMicroelectronics [12]. The benefit of such practise is sufficient for many applications, however, in the worst case, a point to point connection from the middle of one chip to the middle of another chip still needs to be as long as a chips horizontal dimension.

There also is an unfavourable linking between the serial formation of wire bonds and cost per chip-stack, when a large number of connections needs to be made. Other techniques thus have been developed, where the peripheral connections are made in a parallel process, for example with metal patterns on the sidewall, as done by Irvine



a) b) c)

Figure 1.5: Examples of area interconnect technologies in order of increasing density: a) Flip-chip package from Amkor [15], b) GeorgiaTech's Sea-of-leads [16], and c) Infineon's SOLID face-to-face technology [17]

Sensors Corporation [13] or by embedding thinned chips into polymers with via feed-throughs, as available i. e. from Fraunhofer IZM in Berlin [14]. Demonstrators of the mentioned peripheral 3D-interconnect principles are depicted in Figure 1.4.

Periphery interconnects may make efficient use of the whole chip or package area by means of contact redistribution, but still result in long interconnect paths and a large package footprint. More crucial however, is the limitation they impose on interconnect density, because then, the one-dimensional periphery turns out to be the bottleneck.

1.2.2 Area interconnects

A higher number of interconnects per chip-area can be achieved, when the whole chip area rather than only the periphery is used for interconnecting. In state of the art packaging, flip-chip techniques are employed, that is, diced chips are bonded face-down onto the package substrate (i. e. a rigid laminate) in order to create chip-to-package interconnects (for example flip-chip-package from Amkor [15], see Figure 1.5a).

In more advanced approaches, interconnect density can be increased by depositing the interconnect metallization system onto the chip surface, when still on the wafer, by means of a batch process. At Georgia Institute of Technology (GeorgiaTech), the sea-of-leads wafer-level packaging (WLP) technology has been developed [16], where metal leads are deposited onto a polymer, the combination of which functions as a mechanical spring for stress relief between chip and substrate.

Infineon's SOLID face-to-face technology is targeted to stack two chips rather than interconnecting a chip to a substrate. This allows even higher interconnect densities, because thermal mismatch is much of a smaller problem between to similar chips, and protection and sealing is taken care of at the package level. Examples of the described area interconnect technologies are shown in Figure 1.5. Infineon's SOLID technology has a particular significance to this theses and will be treated thoroughly in chapter 4 on page 55.

What has been described in this section are high-density area interconnects for con-

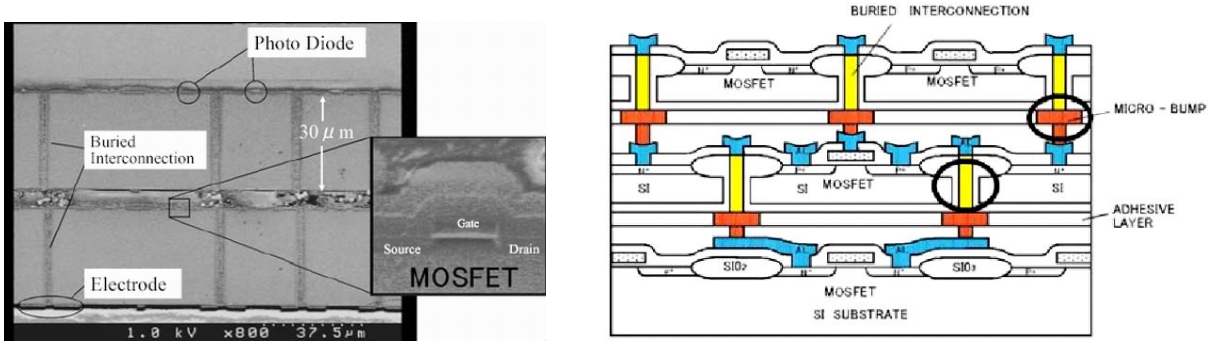


Figure 1.6: 3D integration at Tohoku University, Japan [18]. Small diameter vias with a high aspect ratio were realized using Si deep trench etching and W deposition. Si substrate can then be as thick as 30 µm

ventional 2D-integration, and a stacking technology, with a maximum number of two directly connected chips. For true 3D-integrated systems, with high-density area interconnections between three and more chips, through-chip vias on a micro-scale need to be available. Providing these micro vias in combination with inter-chip joints of similar size has been a subject of research for ca. two decades, and will also be the focus of this work.

1.2.3 Through-chip vias

Through-chip vias provide a connection path from an arbitrary point on the front side of a chip through the bulk silicon of that chip to its backside. These vias, when having a diameter of a few micrometers only, allow the highest density of 3D-interconnects. Methods that provide through-chip vias and allow stacking of three and more chips extend to a rather complex scheme of processing, which mostly includes wafer thinning, DRIE (deep reactive ion etching), via filling, and bonding to a stack of wafers or chips. Laser drilling also can be used to form micro-scale vias, but will not be discussed here for its serial processing nature. Presently known methods of 3D-integration that involve micro scale through-chip vias are introduced shortly in the following.

Tohoku University As demonstrated by the wafer-to-wafer 3D-technology platform developed at Tohoku University, through-silicon vias can be realized with an aspect-ratio (depth to width ratio) of 12:1 (30 by 2.5 µm). The vias were etched 50 µm deep by means of Si deep trench etching and filled with poly-Si and W. First the vias are formed, then the wafer is thinned to 30 µm and bonded to the stack. Thus, when high aspect-ratio via technology is available, the substrate needs only be thinned to 30 – 50 µ, and the thinned wafers need less mechanical support during processing.

The method has been used to fabricate a computing unit with 3D integrated memory [19], an image processing unit and an electronic replication of a biological vision system [18]. Figure 1.6 depicts the vision system integrated into 3 layers of Si chips. The

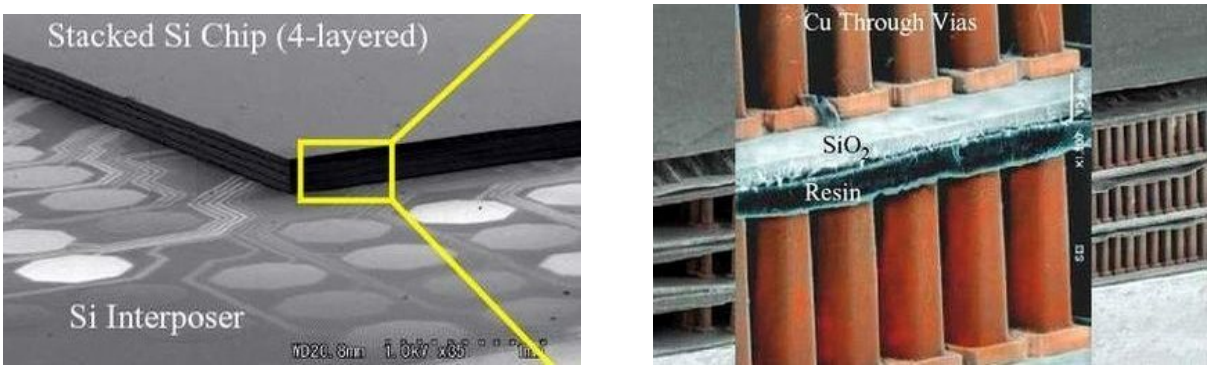


Figure 1.7: 3D integration at ASET, Tsukuba Research Center, Japan [20]. 5-layer stack with 4 thinned chips (l), details after Si removal (r)

contact resistance was measured to be $< 0.1\,\Omega$ and the resistance of the high aspect-ratio W-vias was ca. $7\,\Omega$. Such high resistivity cannot be tolerated by all applications in the field of 3D-integration. Smaller resistance figures have been achieved for example with the technology platform of ASET.

ASET, Association of Super-Advanced Electronic Technologies In terms of via geometry, a similar strategy was followed by ASET: High aspect ratio via features where realized for they consume a small amount of active chip area. At the same time the Si substrate can be made thick enough for not to complicate the thinning process and the handling of thinned wafers.

The vias are etched by means of RIE from the front side of a fully processed substrate, but filled with electroplated Cu. As a seed and barrier layer a conformal CVD-deposit of TiN/Cu was used. However, the maximum tolerable aspect ratio of via-filling using Cu electroplating was reported to be in the range of 7:1 (70 by $10\,\mu m$) and could only be achieved by modifying the chemical composition of the plating electrolyte [20].

With this concept, test chips with contact chains and via chains, as well as active test circuits like ring oscillators (cascaded inverters) spanning up to 4 layers could be produced. The vias had a diameter of $10\,\mu m$ at a pitch of $20\,\mu m$ and a depth of $50\,\mu m$ (corresponds to substrate thickness). In this work, the resistance per via was measured to be ca. $15\,m\Omega$. Figure 1.7 shows a 4-layered chip stack with details on the copper plugs remaining after silicon removal.

Other than the previously described approach, stacking was done here on a chip scale. A provider of both, wafer and chip scale stacking methods is the Fraunhofer Institute IZM in Munich.

Fraunhofer IZM The Vertical Chip Integration (VCI) technology developed at Fraunhofer IZM has evolved from a 1995–1997 joint project "Kubische Integration" with Siemens [21]. This integration process is fully wafer scaled with all processing being done from

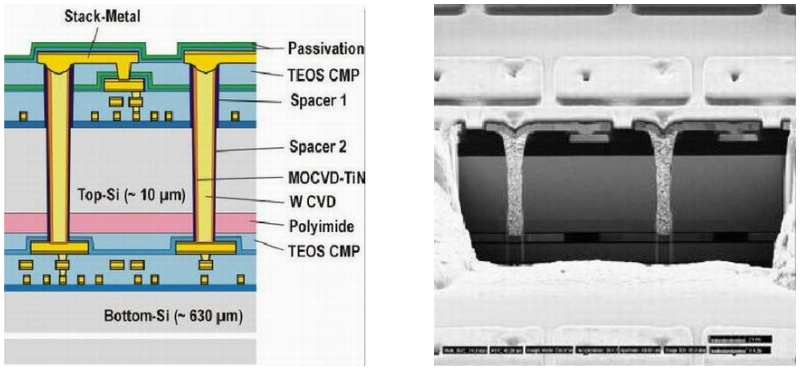


Figure 1.8: Vertical Chip Integration (VSI) with inter-chip vias (ICV) at Fraunhofer IZM in Munich [22]

the front side. The processing sequence is similar to the ones of the previously discussed approaches: Via-forming, thinning, bonding/stacking, with a small variation: The vias were metallized on-stack (CVD-TiN or CVD-W).

The smallest vias had an aspect ratio of 10:1 (14 by 1.4 µm), with a resistance of 2 Ω. Shown in Figure 1.8 is a schematic cross section and a FIB (focussed ion beam milling) cross section of W-filled through-chip vias. Further development evolved into a chip-to-wafer stacking approach [23, 22] in order to address yield sensitive applications.

Rensselaer Polytechnic Institute (RPI) The strategy of 3D-integration at Resselaer Polytechnic Institute differs completely from the ones described so far. Their 3D-integration platform is based on alignment and bonding of SOI (silicon on insulator) wafers with BCB (benzocyclobutene), thus being a fully wafer-level stacking approach. The bonding sequence starts face-to-face and continues face-to-back [24]. The wafers are thinned down to 130 nm with the buried oxide layer serving as a wet-chemical etch stop. Thus, the layer thickness of each level merely consists of the thickness of the circuits metal stack. No handling of thinned wafers is needed, because thinning is performed directly on the stack.

Then, through-chip vias are etched through 130 nm of Si substrate and through typically less than 10 µm of dielectric in the on-chip metal/dielectric stack, which allows small via footprints at rather shallow via aspect ratios of 5:1. In their most recent work, a 2-layered wafer stack has been realized, with 2 metal layers on the stacked wafer and 1 metal layer on the base wafer [25]. Via sizes have been as small as $2 \times 2\,\mu m$ with a combined resistance of $5\,\mu\Omega/cm^2$.

1.2.4 Conclusion

Stacks of chips can be produced either by stacking of whole wafers (wafer-to-wafer), which are then diced, thus producing a large amount of chip-stacks at a time, or by stacking of diced chips (dies) one at a time in a serial process, either onto a single base-chip (chip-to-chip) or onto a base-wafer, which is diced later (chip-to-wafer).

Most of the described 3D-integration methods follow the strategy of wafer-to-wafer stacking (W2W), because its potentially high volume output is attractive to many ap-

Research group	Stacking scale	Process sequence	Via width	Substrate thickness	Via metal	Literature reference
Tohoku University	W2W	v–t–b	2.5 µm	30 µm	poly-Si, W	[26, 27, 28]
ASET	C2C	v–t–b	10 µm	50 µm	Cu	[29, 20]
Fraunhofer IZM Munich	W2W, C2W	v–t–b	1.4 µm	14 µm	TiN, W	[22, 21]
Rensselaer RPI	W2W	b–t–v	2 µm	150 nm	Cu	[25, 30]

Table 1.1: Examples of present day research on 3D integration and references to recent publications. Abbreviations: W2W = wafer-to-wafer, C2C = chip-to-chip, v = via, t = thinning, b = bonding

plications. In terms of economics, it is reasonable to prefer a large-scale batch process, but when stacking whole wafers, yield decreases with the power of the number of added wafers. This results in a high yield sensitivity of the processes, which need to be well established, which they typically are in high-volume production applications. Besides similarly high yield functions, a further requirement for W2W stacking to be beneficial is a homogeneous chip-size.

For other types of applications, where heterogeneous chips need to be stacked (different shapes and varying yield functions) and the various donor wafers are not available in high volume, it seems more reasonable to prefer methods, where chips can be tested and preselected (known good dies) before they are stacked together.

As summarized in Table 1.1, the described methods also differ in their via geometry. When the aim is to have a small via footprint which consumes as low an active chip area as possible, two strategies can be followed. One is to allow the Si substrate to be rather thick in the range of 30 to 50 µm, and to produce narrow vias with aspect-ratios up to 12:1, resulting in a via width in the range of 3 to 5 µm (Tohoku University, Fraunhofer IZM). This is associated with a limited choice of available deposition techniques (mainly high conformal CVD technologies) and materials.

The other possibility is to have vias with a lower aspect-ratio (1:1 to 5:1), which then can be metallized by electroplating of Cu (ASET, Rensselaer RPI). To achieve a small via footprint, more effort has to be put into the process of thinning, which then is required to target a substrate thickness in the range of <1 to 10 µm.

Most of the approaches rely on a processing scheme, that follows the sequence via—thinning—bonding/stacking (v–t–b, see Table 1.1). This implies, that the vias need be etched from the front side of a CMOS wafer through the dielectric of the metal stack, which can sum up to a thickness of 10 µm, and through the silicon substrate. Considering the above implications on filling and via geometries, the aspect-ratio of a 5 µm wide via cannot be lower than 2:1 (5:1 for 2 µm width, as demonstrated at Rensselaer RPI). The approach of Resselaer is an exception, for the vias being etched after thinning (b–t–v) from the backside of the thinned wafer. But the via-connection still needs to be made through the metal stack, because the front-side metallization cannot be used for direct front-back contacts.

All approaches have in common, that the dielectric of the metal stack of a CMOS wafer needs to be through-etched in its full thickness. This is a disadvantage, when high-density through-chip interconnects need to be made as demanded for example by biologically inspired neural processing systems. The vias block routing paths in the on-chip interconnect system.

1.3 A new 3D-chip-integration concept

Considering the conclusions made in the previous section, a novel 3D-integration concept can be sketched, which has the potential to serve best the interconnect needs of a biologically inspired neural vision system or of hybrid systems.

The novel approach is based on the SOLID F2F (face-to-face) technology developed at Infineon Technologies Corporation, which follows the strategy of chip-to-wafer stacking. The core principle is a process called solid liquid interdiffusion bonding, SOLID bonding, which allows thermally stable bonds of high strength where the joint volume merely consists of a single intermetallic compound. Contacts have been realized at a minimum size of 10 µm and a pitch of 20 µm. Resistance of a $15 \times 15\,\mu m^2$ contact was determined to be 7,5 mΩ [31]. Mechanical characterization for an area of $5 \times 5\,mm^2$ revealed a shear strength of 120 N and a peel strength of 40 MPa. The principle of this bonding technique plays a key role in this thesis and will be covered in detail in section 4.2 on page 61.

1.3.1 Process chain

The stacking strategy followed in this new concept, will be of a chip level nature, thus allowing stacking of known good dies originating from heterogeneous technology sources. This is the realistic scenario for hybrid applications the novel approach is targeted to.

In the following, a naming scheme is used, in which the front side of a wafer or chip always refers to the side with the active circuit. Processes can require wafers to be oriented in various ways, with either the front side or the backside receiving the process treatment.

As a through-chip via technology, backside via-etching and via-filling is proposed, which, in combination with the front-side metal interface will be able to provide shortest vias without the need to etch through the on-chip interconnect system of i. e. a CMOS IC. However, a scheme for front-to-back alignment will have to be provided as well. The processing sequence thus will be thinning—via—bonding/stacking (t–v–b), in contrast to the 3D-integration methods summarized in Table 1.1 on the preceding page.

The via geometry is chosen to follow the strategy of low aspect-ratio vias with a target width of 5 µm, in order to have all available degrees of freedom for the choice of the via processes, however, at the expense of providing an advanced thinning scheme with a buried etch-stop layer for a defined target substrate thickness of 10 µm. The resulting processing path is sketched in Figure 1.9 and will be described in the following.

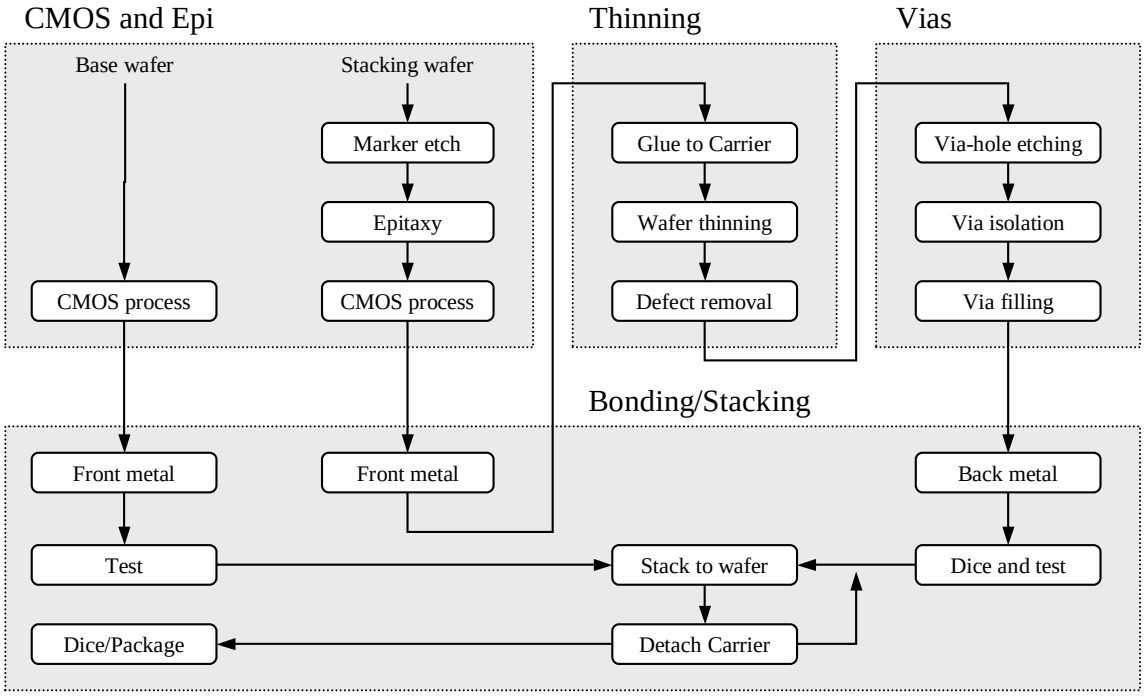


Figure 1.9: Processes for a novel 3D-integration concept with stacking on chip level and a thinning—via—bonding/stacking (t–v–b) process sequence

CMOS and epitaxy Prior to the CMOS process, the stacking wafer substrates need to be preprocessed: Alignment markers are etched into the front side of the blank wafers in order to allow front to back alignment, when flipping to back side processing. These markers can be placed in a way that their geometry does not interfere with the CMOS process. The geometry must be chosen to give enough contrast for the first lithographic mask alignment from both front and backsides.

In a second step of preprocessing a thin highly p-doped epitaxial silicon layer (p^{++}, very high dopant concentration) is grown, followed by a thick ($\approx 10\,\mu m$) epitaxial layer that will define the final thickness of the thinned wafer. This second epitaxial topping layer will host the CMOS circuit and thus needs to be doped according to CMOS technology specifications.

The base wafer and the preprocessed stacking wafers can now be subject to standard CMOS processing. However, the CMOS process should best terminate before applying the final polyimide passivation layer because front side processing is not complete at this point.

Thinning Prior to this step, front side processing needs to be completed by deposition of the interface metallization, which is part of the bonding/stacking functional block.

Now the wafer is flipped and thinned from the backside. In order to handle the ultrathin wafers during backside processing, it needs to be attached to a carrier substrate.

For later detaching a special temporary wafer bonding technique has to be developed. Thinning can be carried out as a two step process, that first mechanically removes material at a high rate (lapping or grinding), then chemically removes defects and stops at the buried p^{++} epitaxial layer [32]. At the end of the thinning block, a smooth surface is required for further processing.

Vias The thinning will reveal the alignment marks, which are the reference to the following via formation processes. Via holes are etched anisotropically into the thinned substrate, stopping at the front side dielectrics with an adjacent landing pad in the substrate nearest metal layer. Next, the vias are isolated with dielectrics and filled with conducting material. Determination of appropriate materials and processes is subject to investigations, however, for via-filling and the backside metallization pattern a combined electroplating process is considered.

Bonding/Stacking Part of this processing block is the front side metallization pattern, as well as the backside metal pattern, which need to be adjacent matching pairs in order to connect back-to-front with other chips. Up to this point, all process steps can be on a wafer level. Then, the wafers are diced into chips.

The chips are now aligned and placed on top of each other. The achievable density of vertical interconnects depends on the alignment accuracy and the maximum tolerable placement mismatch. According to the soldering process, pressure and temperature need to be applied, and care has to be taken to surface preparation in order to achieve agreeable results.

After attaching a chip to the stack, the carrier substrate needs to be removed from the thinned chip, and the front side contacts are exposed. This may require a cleaning process that removes remains of adhesive and restores the contact quality to solderability without destroying other surface structures.

Packaging After chip stacking and removal of the last carrier substrate, the "chip-stacks on a wafer" can be subjected to packaging. Conventional packaging concepts could include forming bonding wires from the diced base wafers bond pads to the package substrate. Advanced packaging concepts could possibly include sealing functionality already at the inter-chip interface during chip-stacking, with the last chip added only being a sealing cap. The stacks could then be packages on a wafer-level (WLP). However, packaging will not be covered in this work, it is included in Figure 1.9 as a reference only.

1.3.2 Critical steps

Apart from its advantages, this new 3D chip-integration approach bears critical steps in its process chain, which represent a challenge to implementation. Table 1.2 lists the processes and critical issues.

Block	Step	Aim, Target	Critical Issues
CMOS & Epi	Preprocessing	Alignment markers, epi	Diffusion in etch stop layer
	Circuit processing	Active circuit	Contrast of alignment markers
Bonding	Contact formation	Front side SOLID contacts	—
Thinning	Wafer bonding	Handling of thinned wafer	Stable at 300 °C, temporary
	Bulk thinning	Mechanical wafer thinning	Substrate damage
	Defect removal	Surface recovery	Etch stop, surface quality
Vias	Oxide deposition	Backside isolation	—
	Anisotropic etch	Via hole	Sidewall angle
	Dielectric depo.	Via isolation	Quality, conformity, leakage
	Anisotropic etch	Establish contact to M1	Remains, damage to via isolation
	Metallization	Barrier and seed layer	Conformity
	Via filling	Through chip contact	Void free filling
Bond./Stack.	Contact formation	backside SOLID contacts	—
	Chip attaching	Chip stack	Alignment, process control
	Carrier removal	Expose front side	Detachment of handling substrate

Table 1.2: List of process steps and critical processing issues

Temporary waferbonding A method of temporary waferbonding that can withstand subsequent temperature processes for oxide deposition and bonding/chip-stacking needs to be identified. Such processes are expected to be carried out at temperatures up to 300 °C. The success of the whole process chain relies on the quality of the implementation of each single step. The temporary wafer bonding step, for being located at the beginning of the process chain, is particularly critical since it can corrupt all other succeeding steps. Most sensitive to the wafer bond are expected to be chemical thinning (alkalic treatment) and deposition of via isolation (temperature).

Front-to-back alignment The etch stop layer is grown first after etching of the alignment markers. All following high temperature processes above ca. 800 °C (growth of active epitaxial layer, front end processes) can cause outdiffusion of dopants and blurring of the contours of the etch stop layer. This can affect the function of the etch stop layer (less selectivity) or the contrast of the alignment markers (edges not visible to the lithography tool). Note that the alignment markers also might change in geometry on the front side, due to the epitaxial overgrowth and dependance of the growth rate on geometric factors (edge orientation, structure depth . . .).

Via processing For sidewall isolation inside the via, a conformal deposition technique needs to be found, with uniform coverage of dielectric across the whole via topology. In a later step, dielectrics need to be removed selectively from the bottom of the via. This combination of sidewall isolation and electrical contact at the bottom of the same confined structure presents a challenge to uniformity of deposition and anisotropy of the

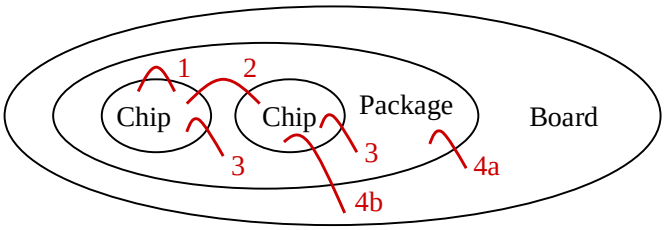


Figure 1.10: Interconnect levels in integrated systems: 1) On-chip; 2) Chip-to-chip; 3) Chip-to-package interconnects; 4) Board level interconnects with high thermal mismatch

etching technology. Failure in these steps can lead to intolerable current leaking and high contact resistance.

Microjoints The SOLID soldering process is known to work with a contact footprint of ca. 20 µm and a joint height of ca. 10 µm. However, for being attractive in neural network applications with highly parallel signal processing paths, it will be necessary to scale contact size and pitch down to the range of a few microns. Contact formation and assembly techniques need to be reconsidered in order to be precise enough under the processing conditions of pressure and temperature.

1.4 Subject of this thesis

This thesis aims to identify an efficient chip-to-chip interconnect technology which is suitable for high-density inter-chip and through-chip connections within a chip-stack of three or more layers, as required in neural image processing applications. It further aims at the implementation and characterization of such interconnect technology using Si as a material platform to deduct its feasibility for CMOS integrated circuits. Therefore, this thesis will focus in particular on the processes of wafer thinning, via filling and chip stacking as part of a the novel 3D-chip-integration concept introduced earlier. Future demands of increasing interconnect density and shrinking joint dimensions are addressed by discussing means of joint scaling.

It is thus intended to provide an interconnect technology platform for 3D-chip-integration on a proof of concept basis, and to present an alternative to already established 3D-integration concepts. Part of other related work, not covered in this theses, is to improve the electric properties of through-chip vias and to demonstrate the versatility of the investigated 3D-integration concept to combining of non-Si technologies with Si-based technologies.

1.4.1 Scope

Figure 1.10 sketches the different levels of system integration and existing interconnects between these levels. The scope of this theses will focus on establishing interconnects on a chip-to-chip level (2), because this level provides the most direct connection path after on-chip interconnects (1).

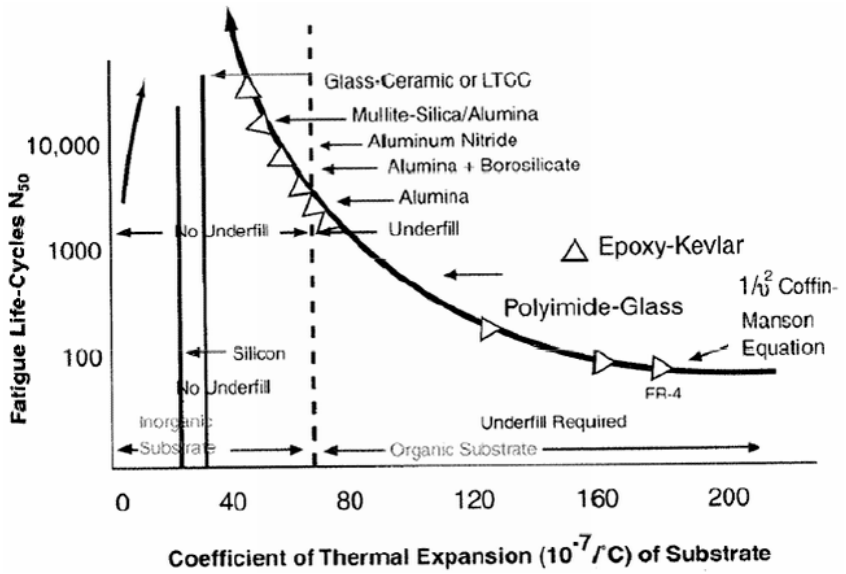


Figure 1.11: Relation between interconnect life-time of Si-based chips and thermal expansion of substrates they are connected to. From Tummala [33]. Longest life-time can be expected from thermally matched substrates, as for example in a chip-stack with direct chip-to-chip interconnects

Chip-to-package (3) and package-to-board (4a) or chip-to-board (4b) interconnect systems are rather outbound connections to dielectric passive substrates, which are made of material compounds with a significantly larger coefficient of thermal expansion (CTE). Thermal and mechanical stress management in these interconnects require either flexible wires (wire bonding) or large ductile joint structures (flip-chip-bonding) that can deform without breaking, combined with underfill polymers that take up part of the mechanical load. This lowers interconnect density, because interconnects are either limited to the periphery, or need a certain joint size to allow for stress relief.

Figure 1.11 illustrates the relation between interconnect reliability and substrate thermal mismatch by plotting fatigue life cycles over the substrate coefficient of thermal expansion (CTE) for Si-based chips [33]. The graph implies, that an underfill polymer needs to be used when interconnecting Si chips to substrates with a CTE above $7 \cdot 10^{-6}/°C$, which is usually not the case when stacking bare chips onto each other.

Thus, packaging related issues can be neglected to a certain extent on a chip-to-chip stacking level, and thus will not be regarded in this theses. However, this work will be dedicated to implementing the shortest interconnect path on a chip-to-chip level, which implies 3D-stacking of chips. For its similarity with flip-chip bonding techniques, and its need for precise alignment, chip-stacking techniques may partly combine known bonding techniques developed for packaging.

As depicted in Figure 1.12, a rough outline of the evaluation concept that leads to an efficient interconnect system for 3D-chip-stacks can now be sketched. It will follow a chip-to-chip stacking strategy, for it is better suited to hybrid integration schemes, where yield usually is more important than throughput. During evaluation experiments it further guaranties sufficient process control within the means of research laboratory equipment, without having to qualify processes for production scale equipment. Area interconnects in a stack with three and more chips require both, through-chip vias and

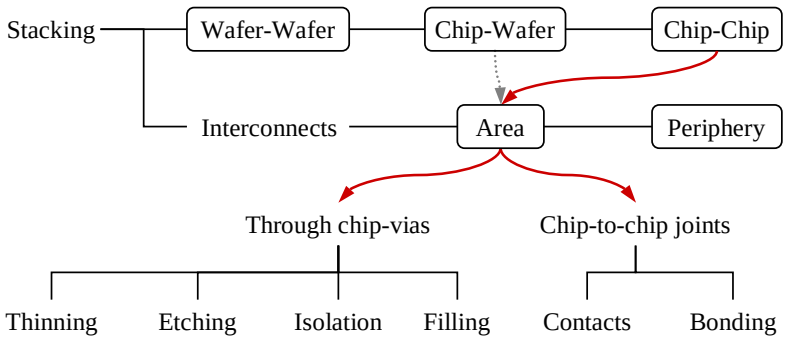


Figure 1.12: Path for implementing high-density 3D-interconnects in a 3D-chip-stack. Area interconnects in a stack with three and more chips require both, through-chip vias and inter-chip connections

inter-chip connections. For the realization of the vias, thinning, via-etching, sidewall isolation and via-filling processes will be evaluated, and for the connections between the chips, a suitable bonding process and contact formation process will be identified and evaluated.

Other related work at the University of Ulm was dedicated to the subject of technology convergence in the investigated novel 3D-integration concept: Hybrid integration with through-chip interconnects was attempted by stacking a diamond-based sensor chip onto a Si base chip [34, 35]; an efficient side-wall isolation method for through-chip vias in silicon was identified, implemented and characterized [36, 37]; and interconnect path testing and signal rerouting concepts where developed and implemented as built-in self-test circuit modules [38, 39].

1.4.2 Procedure

For the proof of concept experiments carried out to evaluate the process steps, samples of ca. $5 \times 5\,\mathrm{cm}^2$ — in the following referred to as *wafer element* — are used in order to represent processes usually carried out on full wafers. This has the advantage of efficient handling and preparation of the samples, and allows good control of the processes within the available equipment and custom made tools. Most work is done on plain Si substrates.

In a first phase, selected processes are evaluated with dedicated experiments and test samples. At this stage, process physics need to be understood in order to optimize parameters and to project them into the target process chain. Equipment needs are identified and additional tools are acquired, either as a commercial product (i. e. flip-chip bonder) or as fully custom designed constructions (i. e. spin etcher, bonding press). By the end of this phase, methods and tools are made available to implement and evaluate the full process chain. In the second phase, the process chain is synthesized back to full length and fine tuned, with the goal to produce a technology platform demonstrator, which can be characterized sufficiently as to provide a proof of concept. A third phase is dedicated to the development of strategies to optimize specific processes. In this thesis, the focus will be on a microjoint technology and its scaling potential when applied to a much denser layout of vertical interconnects with a contact footprint near $1\,\mu\mathrm{m}^2$.

1.4.3 Outline

Chapter 2 is dedicated to wafer thinning and to solving problems associated with it, mainly the problems of defect removal, thickness variations and handling of thinned wafers. Two thinning processes and two waferbonding processes are described, evaluated and compared. In chapter 3, electroplating is evaluated as a method to form a patterned metal interface for chip-to-chip bonding, and as a means to fill through-chip via holes with conductive material. Chapter 4 addresses the process of forming a temperature stable rigid mechanical and electrical connection between two chip faces in a stack. In chapter 5 a method to scale down electrical chip-to-chip interconnections to achieve smaller joint geometries is presented and evaluated. The theses closes with an outlook on further potential applications of the presented interconnect technology in chapter 6.

Chapter 2

Thinning and wafer handling

In packaging, thinning of wafers is often used in order to reduce the length of bonding wires for chip-to-substrate interconnects, or to improve the thermal path of a circuit to the heat sink. Usually the thickness of large wafers lies within ca. 700 µm for 200 mm-wafers and ca. 800 µm for 300 mm-wafers in order to provide enough stability and rigidity for all needed circuit processing.

When thinning down to below ca. 100 µm, large wafers need to be stabilized by a handling substrate. The choice of material used for handling support depends heavily on the processes the thinned wafer is subjected to. Most significant will be thermal conditions and stress build-up during bonding or subsequent processing and wether the handling substrate needs to be detached later or can stay in place permanently.

2.1 Introduction

Thinning below 50 µm is a challenging process, even more for the degrees of freedom in selecting methods of waferbonding and waferthinning is outstandingly limited, as we will see in the following. Figure 2.1 shows the processes associated with thinning, which also includes waferbonding.

2.1.1 Concepts for waferbonding

In this 3D-integration scheme, requirements and conditions for wafer bonding and support of the thinned substrate are very limiting: The bond must be temperatures stable, because subsequent processes (via dielectric deposition + soldering) on the thinned wafer element

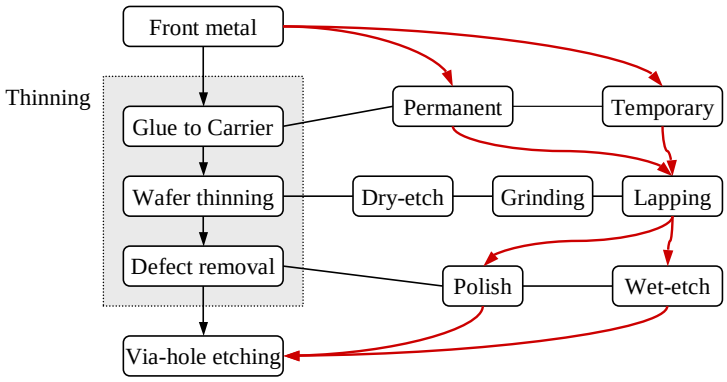


Figure 2.1: Process paths for waferthinning and handling: Two concepts for waferbonding (permanent and temporary), and two concepts for lapping defects removal (polishing and wet-etching with a doped etch-stop layer) will be evaluated

can reach up to 300 °C. Further, a method is needed that can tolerate surface structures, as there will be front side metallization patterns needed for later chip-stacking. Direct wafer bonding relies on high quality surfaces with a flatness on an atomic level [40]. It thus cannot be employed in this case. An adhesive method which is able to embed metallic structures needs to be found. The adhesive volume at the bonding interface will serve as a buffer layer to surface inhomogeneities, which will make this technique robust and tolerant with respect to particle contamination.

When the curing of the bond involves "high" temperatures ($\approx 300\,^{\circ}\mathrm{C}$), wafer and substrate need to have similar CTEs to avoid stress build-up and debonding when cooling down. Therefore, Si is chosen as a handling substrate material. After thinning, the wafer is more flexible and is assumed to respond less critical to (thermal) deformations of the handling substrate. However, the bond needs to be strong as to prevent buckling and delamination.

As depicted in Figure 2.1 two concepts of bonding will be evaluated: permanent adhesive bonding and temporary adhesive bonding. For stacking of the chips, the handling substrate needs to be detachable without remains. However, available temporary sticking agents debond by heating or solvent solution and are not temperature stable and cannot be applied here. Instead it is proposed to use available temperature stable hard curing adhesives in combination with a sacrificial layer. The sacrificial layer can be a metal which is etched wet-chemically, or an organic material that can be dissolved in a solvent. Thus, permanent adhesive waferbonding is part of the concept for temporary waferbonding and will be evaluated first.

Deviations from a parallel alignment of surfaces will result in an error during mechanical thinning and results in a non-uniform thickness of the thinned wafer. It is therefore important to apply the adhesive uniformly, for example by spin-coating. Pastes and thixotropic adhesive materials, or adhesives with filler particles are difficult to apply with a uniform thickness.

The choice of available adhesive materials it thus limited to polymers such as polyimides, epoxys and BCBs (benzocyclobutene).

2.1.2 Concepts for wafer thinning

For wafer thinning down to 10 µm, more than 95 % of Si material need to be removed within a specified short time. This calls for a high removal rate and a high surface planarity, which usually cannot be provided by standard etching equipment. Thus, mechanical thinning techniques like grinding or lapping, known from packaging, will be the methods of choice.

In general, high removal rate processes, particularly mechanical abrasive methods, will result in a rough and defective surface which considerably reduces the mechanical breaking strength of the thinned substrate [41]. Such backside surface no longer suffices lithography and microtechnology processing requirements, and even may induce degradation to front side electronic performance. Crystal layers with cracks and defects need to be removed

and surface smoothness needs to be reestablished.

Therefore, a 2-step thinning technique is proposed. In a first step, the wafers are pre-thinned mechanically by either grinding or lapping. Here, lapping was used for its equipment availability and full access to process conditions. In a second step, the defective surface is restored by means of defect removal. As shown in Figure 2.1, two methods are considered, chemical-mechanical polishing (CMP) and wet-etching.

The thickness variations of starting wafers may be in the range of nanometers over the whole surface, but bending or cupping as a result of internal stresses due to preprocessing can amount to as much as 5 µm [42]. When the wafer is not kept flat during waferbonding, these deviations from flatness will be transferred into thickness variations of the same amplitude during thinning. In order to get a better control of the final thickness, an endpoint detection mechanism — or better — a selective process stop mechanism should be considered.

As described by Seidel et al. [43], the etch rate of doped silicon in various chemical solutions is dependent on the boron concentration. This is attributed to space charge effects at the silicon/etchant interface, which results in a lack of electrons for the etching reaction in highly boron doped silicon. This can be used as an etch-stop during wet-chemical defect removal after lapping. Here, the etch-stop is provided by a buried epitaxial layer, and as an etching solution KOH (potassium hydroxide) is used. Chemical thinning with KOH was not a major part of this work, and thus will be discussed only briefly in section 2.3.3 on page 33.

As an alternative second step for defect removal, plain wafers without an etch stop layer were also subjected to the process chain 2.1. Here, a visual endpoint detection mechanism is used, where the thinning is stopped, when previously front-side-etched markers with a known depth become visible. Defects caused by mechanical thinning are removed by polishing. This technique will be discussed in detail in this chapter, as a replacement for chemical KOH-based thinning, without the need for a buried etch-stop layer.

For endpoint detection during lapping, depth and alignment markers need to be etched from the front side into the silicon surface, employing deep reactive ion etching (DRIE).

2.2 Wafer bonding

Two waferbonding processes, as proposed in section 2.1 on page 19 will be developed and evaluated in the following. One will result in a permanent bond of the wafer element to the handling substrate. For chip-stacking, no feasible solution is available to remove the handling substrate from a thinned chip, however, this type of bond will suffice evaluation experiments on via etching and via filling.

The second bonding process will allow temporary adhesive bonding and will thus suffice all requirements for further processes (as listed in 2.1 on page 19), including handling substrate removal for chip-stacking. First, evaluation experiments are carried

	T_g (min.)	Decomposition	Bonded area	Bond quality
PMGI SF11	190 °C	335 °C	<50 % bubbles	bad, none
Durimide 7510	285 °C	525 °C	100 %	good
Cyclotene 3022-46	350 °C	n. a.	100 %	very good

Table 2.1: First experiments with different adhesive layers on glass substrates. Temperature properties according to the providers [44, 45, 46]

out in order to identify the process, then the process is optimized for embedding of front side metallization structures.

2.2.1 Permanent bond to handling substrate

First experiments were carried out on $20 \times 20\,\mathrm{mm}^2$ glass substrates with different polymer films, and focussed on obtaining a stable, homogeneous bond interface with "high"-temperature stable adhesive media. As explained earlier, at this stage, the carrier detachment feature is not implemented. The curing conditions where basically adopted from Niklaus et al. [47, 48] and led to the following results (summarized in Table 2.1):

- Bonding silicon to thermally mismatched glass (soda-lime glass, CTE $\approx 7.6 \cdot 10^{-6}\,\mathrm{K}^{-1}$) led to debonding during cooling down after curing. Thus, large Si-substrates should only be bonded to thermally matched substrates or Si ($\mathrm{CTE_{Si}} \approx 2.5 \cdot 10^{-6}\,\mathrm{K}^{-1}$).
- Most photo resists (e. i. PMGI), though temperature stable to ca. 300 °C, contain solvents, that outgas during curing. These are trapped at the interface (Table 2.1) and prevent a stable bond.
- Pressure in the range of 1 to 2 bar needs to be applied in order to achieve the required highly parallel alignment and intimate surface contact.
- Soft curing before joining the surfaces avoids excessive squeeze out of adhesive and eliminates tilting and drifting of the aligned substrates during hard curing.

In accordance with Niklaus [47], good results could be achieved with BCB. In the experiments, the BCB Cyclotene 3022-46 from Dow [46] was applied to one surface at a spin speed of 2000 rpm, and a proposed after-cure layer thickness of ca. 3.8 µm. Before joining, the BCB was soft cured for 5 min at 65 °C. The substrate representing the active wafer was slightly smaller as to prevent edge beads to affect the bonding performance.

A pressure of ca. 20 kg was applied on an interface area of ca. $40 \times 40\ \mathrm{mm}^2$ by using a simple cantilever set-up. Shown in Figure 2.2 is an acoustic microscope scan of the bonding interface, obtained from a sample under good curing conditions (left) and a failed bond (right). The scans where provided by the Institute of Micro and Nanomaterials at

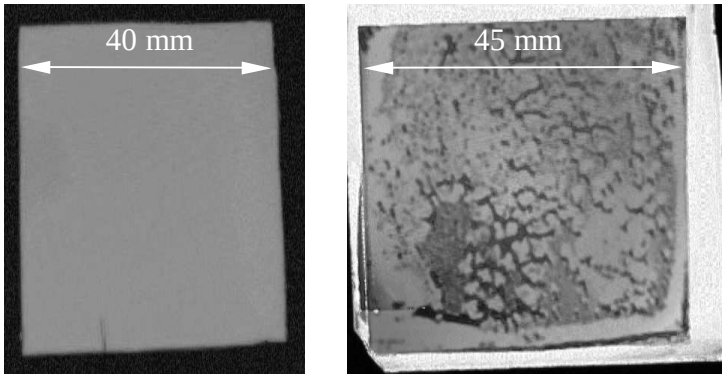


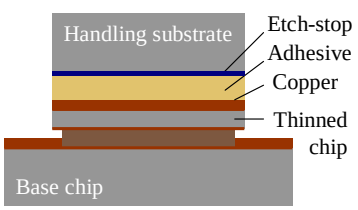
Figure 2.2: Waferbonding results with BCB: Interface scans of an acoustic microscope (SAM) of a successful bond (left), and an interface with large unbonded regions (right)

the University of Ulm. The sample was immersed in water, which served as a coupling media and was scanned by an ultrasonic transducer, generating a graphical interpretation of the reflected acoustic signal, as described in [49]. The bonded regions appear at a lighter tone, whereas unbonded regions, which reflect and scatter parts of the acoustic signal, are represented by a darker tone. The principle of scanning acoustic microscopy has evolved into a powerful nondestructive method to inspect bonded semiconductor devices [50].

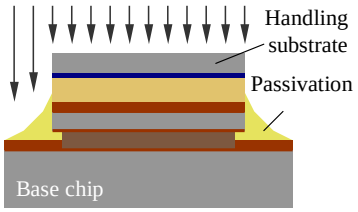
The failed bond, shown to the right in Figure 2.2, resulted from the application of adhesive to both surfaces, with the edge beads of the smaller substrate prematurely sealing the bond, thus trapping air at the interface. It is also assumed, that edge bead thickness variations cause an inhomogeneous pressure distribution and lead to local shifting of the yet soft adhesive material. As curing continues, the trapped air is frozen into the hardened adhesive.

Lapping of the successfully bonded substrates showed, that the wafer element is sufficiently stabilized during thinning, when the material is mechanically stressed. Thus, waferbonding with BCB is a promising method to support thinned substrates for backside processing.

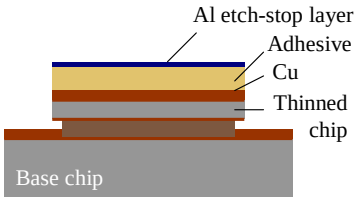
Bulk removal of permanent handling substrate On some samples, a rather unfavourable method was used to remove the handling substrate for a first technology demonstrator. Before bonding the wafer element to the handling substrate, an Al etchstop layer was deposited onto the handling substrate. After dicing and stacking, the substrate was removed by a high-rate silicon RIE process (see Figure 2.3). The base chip and the sidewalls of the stacked chips were protected by a passivation layer, which was applied manually (crystal bond) This was only possible at an early stage of process development, where a plain Cu covered base plate was used as a base chip. To proceed to an advanced stage, however, a working detachable bonding technique needed to be developed.



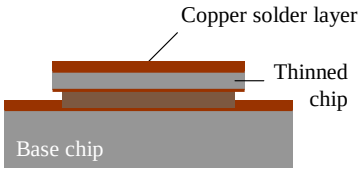
a) Chip on base chip stacking



b) Handlung substrate removal



c) Passivation removal



d) Etch-stop and adhesive removal

Figure 2.3: Bulk removal of handling substrate as a replacement for temporary waferbonding. The base chip surface needs to be protected by a passivation layer

2.2.2 Evaluation of temporary bonding techniques

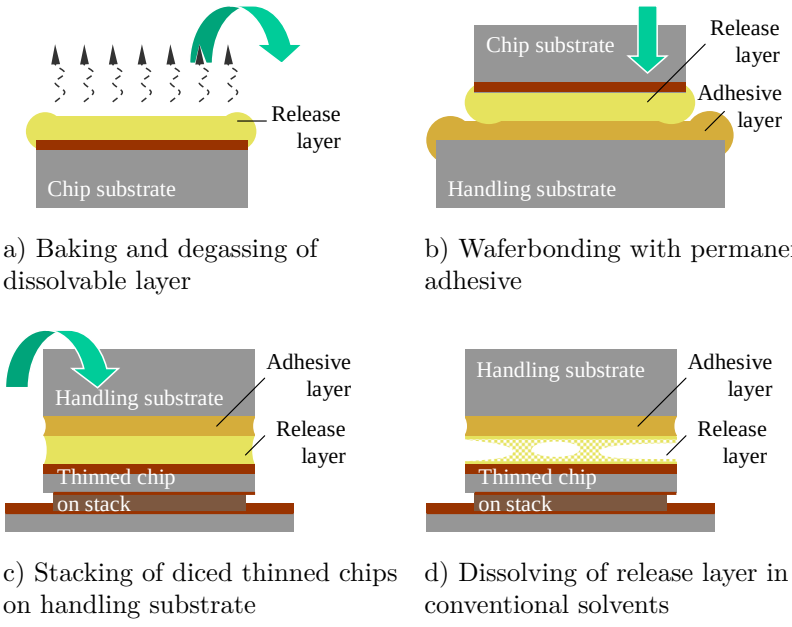
Various principles of substrate detaching were evaluated, namely techniques which involve dissolvable sticking/glueing media by means of standard solvents, and techniques with sacrificial metal layers, where an etching solution exclusively attacks a sacrificial metal layer between the thinned chip and the attached supporting substrate. First evaluation experiments where carried on samples with a flat surface, with no metallization pattern.

Sacrificial metal layers For experiments with sacrificial metal layers, single layers of Al (1000 nm), Ge (500 nm) and Ti (800 nm) were evaporated onto $24 \times 24\,\mathrm{mm}^2$ glass substrates, which then were covered by a layer of BCB. After the BCB was soft cured for 5 min at 70 °C, the substrates were bonded to other glass substrates of the size $20 \times 20\,\mathrm{mm}^2$ in order to avoid interference with edge beads.

The glass substrate allowed for a simple detection of the dissolution progress, when detachability was tested by wet-etching the sacrificial layer. It could be observed, that the reaction rate was strongly reduced by the limited lateral access of the etching solution to the metal layer. Metal layers that can be etched selectively without interfering with other process steps where not available in an appropriate thickness. Thus, sacrificial metal techniques were not considered suitable for this processing scheme.

Resin based release layer The excellent thermal stability of PMGI based resists (see Table 2.1 on page 22) with a decomposition temperature well above 300 °C, and the combination with the above concept of sacrificial layers offer a possible solution to the detachment problem, since PMGI resists can be dissolved in organic solvents without the need for wet-etching.

Previous experiments revealed that waferbonding with solely PMGI is not possible, because of outgassing of solvents during curing. However, when the functionality of layers



a) Baking and degassing of dissolvable layer

b) Waferbonding with permanent adhesive

c) Stacking of diced thinned chips on handling substrate

d) Dissolving of release layer in conventional solvents

Figure 2.4: Principle of handling wafer detachment by a two-layer wafer bonding technique based on PMGI resist as dissolvable layer and BCB as wafer bonding adhesive

is divided into that of a "sacrificial" release and an adhesion layer, PMGI can be applied and fully cured at atmosphere, while sealing and bonding of the PMGI-covered chip substrate to the handing substrate will be possible without outgassing by using BCB.

Figure 2.4 shows the principle of this two-layer wafer bonding technique. First (2.4a), the chip substrate was covered with a PMGI layer of appropriate thickness in order to allow for enough lateral access to the stripper solvent. Then, the resist layer was hard baked at a temperature above the final wafer bonding curing temperature, as to prevent softening or further outgassing by conditioning the resist at harsher conditions.

In order to keep PMGI baking temperature well below its decomposition temperature of 335 °C [44], BCB curing was targeted to a temperature of 250 °, where 100 % of polymerization can be achieved in a reasonable time of approx. 1 hour (see diagram in Figure 2.5 [51]), however, taking into account the polymerization degree reached after soft curing, full curing can be reached in an even shorter time.

For PGMI baking and conditioning, a temperature of 275 °C was used. In an experiment, the weight of a PMGI layer was measured over time and temperature, which revealed, that after reaching this temperature, no considerable weight loss occurred for a time period of ca. 2 hours. Figure 2.6 shows, that weight loss, and thus, outgassing occurs predominantly during the heating-up phase. Thus, it can be assumed, that after this pretreatment, no significant outgassing will occur during the wafer bonding step shown in Figure 2.4b.

In order to address adhesion problems with high viscosity thick film PMGI resist (SF19), first, a thin layer of PMGI-SF11 (700 nm at 4000 rpm) was spun on and soft baked at 150 °C for 10 min. Then PMGI-SF19 was applied at 2000 rpm, which results in a film thickness of 6.5 µm [44] under normal baking conditions.

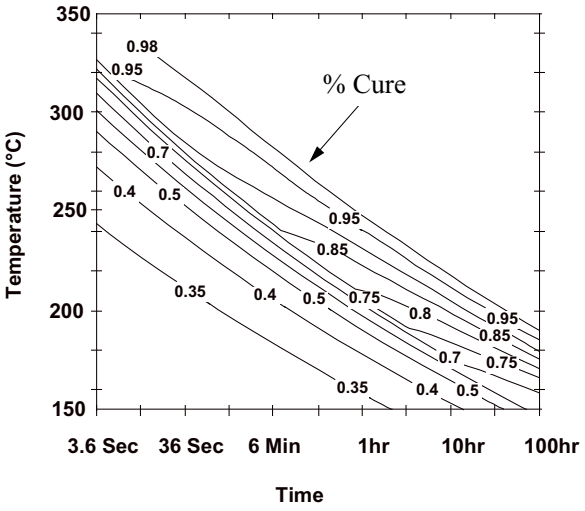


Figure 2.5: Dependance of polymerization of BCB on the heating cycle [51]

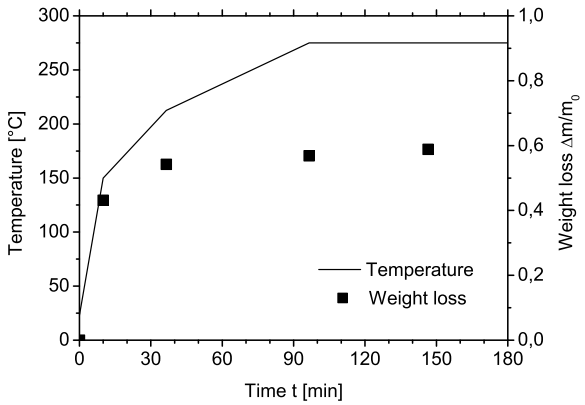


Figure 2.6: Weight loss of PMGI during hard curing at elevated temperature. Most solvent is evaporated after ca. 40 min of ramping up

2.2.3 Temporary bonding with surface topology

The wafer bonding method described in the previous section was applied to substrates with patterned area metallization, using a pneumatic piston for force generation (see Figure C.2 on page 105). The piston was able to generate a downward force of 440 N, which is sufficient to bond a sample area of 32 cm^2 ($\approx 5.5 \times 5.5\,cm^2$) at a bond pressure of 2 bar.

Other considerations addressed surface planarity before bonding. Since the hard baking of PMGI leads to a massive loss of volume, it cannot be used to planarize the surface prior to joining. Thus, before bonding, a thin layer of BCB-63 is spun onto the wafer element at 6000 rpm and soft cured for ca. 10 min at 180 °C. Due to a "reflow" effect this has the desired planarization effect, and the metallization pattern is fully embedded into resin. The waferbonding procedure continues by coating of the handling substrate with BCB-63 at ca. 4000 rpm, placing wafer element and handling substrate in the bonding press and apply a pressure of ca. 2 bar at 250 °C for ca. 60 min.

As a solvent to dissolve the PMGI resist layer and to detach the handling substrate, 1-Methyl-2-Pyrrolidone was used successfully.

Air trapping It appears, that the embedding of metal patterns into the adhesive layers negatively influences bond quality: After back-lapping to a remaining thickness of ca. 50 µm, the wafer elements often exhibit shadow regions on the lapped surface (see Figure 2.7). Due to entrapped air, locally unsupported material is bent inwards and subjected to other lapping conditions, thus causing a visible contrast when seen from above. These unbonded regions result in crack formation and breaking of the thinned wafer element during subsequent processes.

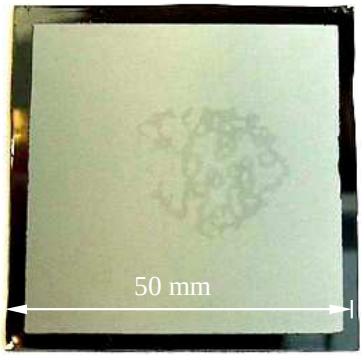


Figure 2.7: Failed bond, evident after back-lapping of a wafer element

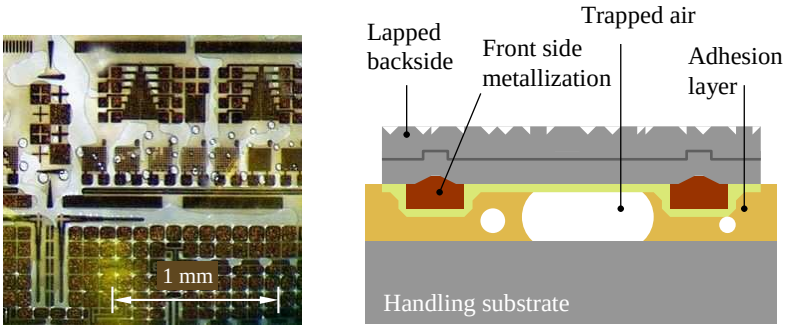


Figure 2.8: Analysis of failed bond interface: a) Through-glass view of bonded glass substrates with trapped air in the interface. b) Schematic cross section of the bonded interface

The bonding process was analysed by investigating the interface of the bond. Therefore, two glass substrates of the size of the wafer element and the carrier substrate where metallized and bonded together using the same described method. The following observations can be made at the now visible bond interface (Figure 2.8):

- During the stage of curing, BCB is temporarily viscous and can adjust its profile according to the local pressure distribution or wetting conditions. The metal pattern represents a barrier to entrapped air bubbles, which then accumulate in areas with no metal pattern.
- The regions with trapped air are predominant in a central region of the wafer elements. This can be caused by the influence of edge beads, which cannot be avoided completely, since coating is applied to both, wafer element and handling substrate

It can be concluded that there are two major effects, that cause air trapping in the bond interface during wafer bonding: Coating thickness variations and wetting.

The first can be attributed to the unfavourable square shape of the wafer elements and the edge beads of the relatively thick coating materials. This will cause air trapping when the coating bulges at the edges first come into contact with the handling wafer, thus forming a seal around a cavity in the center of the wafer element. Although the edge beads are removed manually prior to bonding, the symptom remains persistent, perhaps because of the width of the edge-influenced region.

The second effect can be described as dewetting by the contraction of the heating induced softened coating layers, caused by unfavourable surfacial conditions after spinning on the coatings or joining the wafers. Such dewetting has mainly been observed when applying thick layers of coating.

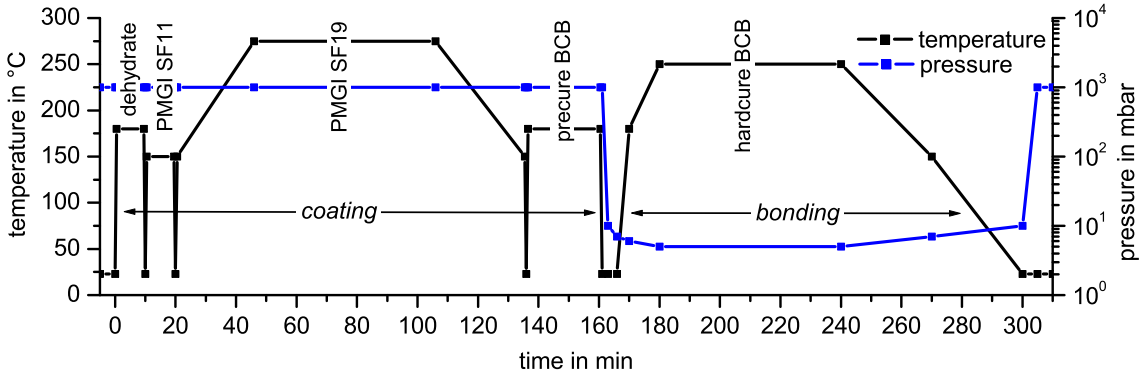


Figure 2.9: Schematic course of temperature and pressure during two-layer temporary waferbonding with PMGI and BCB. Most time is needed to ensure a low outgassing, low solvent containing PMGI release layer, and to ensure sufficient curing of the BCB adhesion layer

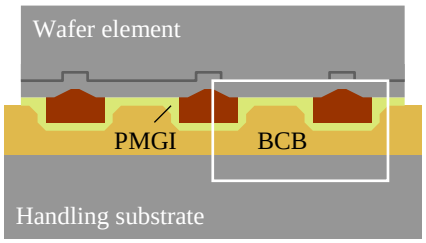
Vacuum bonding As a solution, vacuum bonding was considered most effective for it addresses both of the two described effects associated with air trapping by reducing the amount of gas molecules that can be trapped at the interface.

The bonding equipment was redesigned to bear a heatable chamber with a flexible metallic membrane (see Figure C.2 on page 106). When evacuated, the membrane is applying the bonding force uniformly to the sample. A built-in separation mechanism also allows to hold back the membrane until the vacuum is established to a pressure $< 1\,\text{mbar}$.

2.2.4 Results

Adhesive wafer bonding has been identified as method of choice for its ability to embed the front side metallization pattern. In pre-evaluation experiments with flat surfaces, a method with BCB resin was favoured for permanent waferbonding and a combination of BCB and PMGI resist as sacrificial release layer for temporary waferbonding. These materials have been proven compatible with subsequent processing at temperatures up to 300 °C. For its lack of an effective handling substrate removal procedure, permanent waferbonding with BCB is only considered for evaluation experiments. Therefore, the target process for waferbonding was determined to be the two-layer BCB/PMGI temporary vacuum bonding process developed in the previous section.

A typical temperature and pressure cycle associated with the target two-layer waferbonding sequence is shown in Figure 2.9. The hard baking temperature of 275 °C for the PMGI release layer is kept for 60 min, and the hard curing cycle for BCB waferbonding at 250 °C is also kept for 60 min. When optimized, these cycle times can probably be shortened, for full BCB cure can be reached earlier in pre-cured films, and PMGI outgassing occurs mostly during ramp-up (as was deducted from Figure 2.5 and 2.6 on page 26).



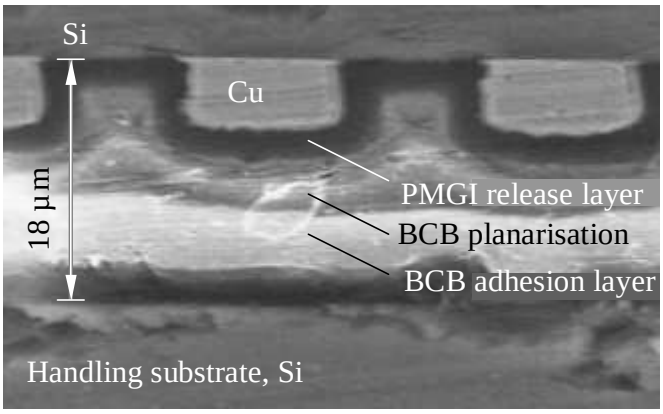


Figure 2.10: Two-layer temporary waferbonding with embedded front-side metallization pattern. Sketch and SEM view

Figure 2.10 shows the bonding results of the improved vacuum bonding process in a micro scale: As can be seen, good embedding of the front side metal pattern into the adhesive coatings has been achieved. The overall adhesive thickness in a waferbonding process as described above, amounts to ca. 18 µm with a PMGI release layer thickness of ca. 2–3 µm.

The release layer thickness of 2–3 µm has been proven to allow sufficient vertical area for solvent attack during chip release: At a temperature of 150 °C diced thinned chips of $10 \times 10\,\mathrm{mm}^2$ can be released from their handling substrate after an immersion time of ca. 30 min. Application of low shear forces supported the release procedure.

2.3 Wafer thinning

An effective thinning process needs to be found, that allows highly precise thinning of Si substrates down to a thickness of 10 µm (see section 2.1). Two alternatives of thinning were proposed in section 2.1 on page 19 and will be evaluated in the following. Both rely on a preceding mechanical bulk material removal step.

2.3.1 Mechanical wafer thinning: Lapping

Lapping is a common technique also known to metallurgy and mineralogy, which provides an efficient means to maintain surface planarity during thinning at a high material removal rate, which usually cannot be achieved by standard Si etching processes.

Lapping differs from the better known technique of grinding in that abrasive particles (mostly aluminum oxide or boron carbide) are used, which are unbound and can move freely over the surfaces. During the process, the water suspended particles are supplied at a constant rate and drained as they are transported to the outside of the lapping plate (see Figure C.3). The sample is pulled over the cast iron lapping plate in rotating and reciprocating movements, while a constant down force is applied. The removal rate depends on the particle size of the abrasive particles, the down force applied to the sample

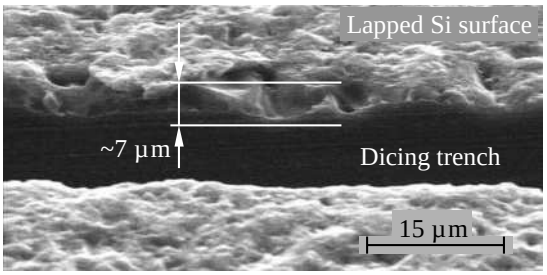


Figure 2.11: Surface after lapping and defect depth of a Si sample lapped with 3 µm Al_2O_3 abrasive particles. Pits as deep as 7 µm have been found in SEM micrographs

and the rotational speed of the lapping plate, and typically reaches $1 - 4\,\mu m/min$.

Samples (bonded wafer and handling substrates) are fixed on the glass holder by a low melting bonding wax. Special care must be taken in this mounting step in order to achieve a highly parallel alignment of surfaces. Abrasive particles were available as calcinated aluminum oxide powder with a maximum grain size of 3, 9 and 12 µm (provided by Struers), which where held in watery suspension by a rotating feeding cylinder.

These particles cause randomly orientated scratches in the substrate surface, which sum up to the desired highly planar surface material removal. However, according to particles size and down force, cracks and crystal defects are generated which affect a considerable volume beneath the surface. From the defect depth of ca. 7 µm, measured after lapping with 3 µm Al_2O_3 grains, it can be estimated, that lapping produces defects that affect a region below the surface which extends to roughly 2 times the grain size of the abrasive particles employed. Studies of defect depth have not been intensified for the numerous parameters associated with lapping.

Thus, the process of lapping needs to be stopped several tens of microns before reaching the final thickness, depending on the grain size used last. In situ, thickness control is only available by means of measuring the piston movement, which is associated with the amount of material removed. For a precise measurement, the process of thinning needs to be interrupted. The glass plate is removed from the sample holder and the total thickness from glass plate to thinned surface is measured.

For the various lapping experiments carried out in the progress of this study, parameters have been documented for process characterization and better removal rate prediction. Thus, the material removal rate r can be calculated at a constant rotational speed n and constant grain size K, and plotted vs. the down pressure $p = mg/A$. As can be seen in Figure 2.12, the following linear dependence can be deducted for this case,

$$r = c_{\mathrm{abr}} \cdot p, \tag{2.1}$$

with an introduced abrasive constant c_{abr}.

Presuming a proportional dependance of average sample speed v and rotational speed n of the lapping plate, another dependence can be deduced: In a first order estimate the pressure independent rate quotient r/p is proportional to the rotational speed n, and thus, to the relative sample speed v (see Figure 2.13). The scattering in the plot can be accounted to the fact, that the plate revolution and the sample holder revolution are coupled by means of surface friction, and thus, the relative speed v heavily depends on

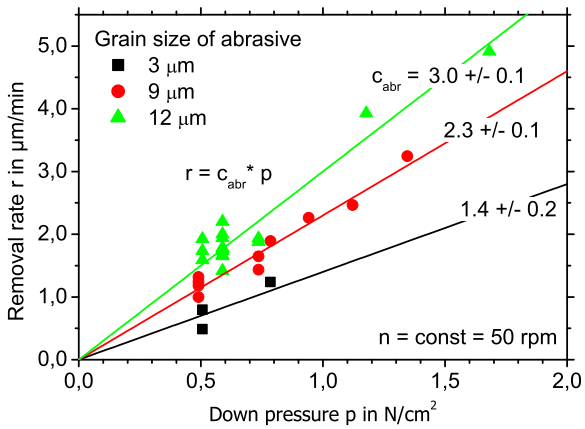


Figure 2.12: Dependence of the lapping rate on grain size and down pressure at a constant rotational speed of 50 rpm

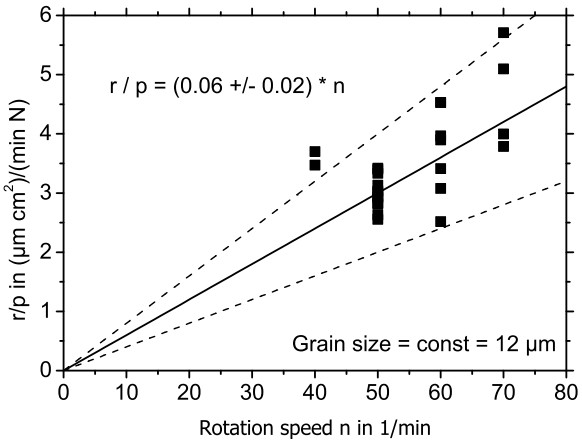


Figure 2.13: Dependence of the lapping rate on the rotational speed with a constant abrasive grain size of 12 µm

the feeding conditions of the abrasive particles, which lacks precision of control in the used configuration.

Thus, the process of lapping shows prestonian behaviour (after Preston [52]), according to which the removal rate r is proportional to both relative sample speed v and down pressure p:

$$r \quad = \quad k_{\mathrm{P}} v p \quad \approx \quad k_{\mathrm{P}} k_{\mathrm{rot}} n \frac{mg}{A}, \tag{2.2}$$

with the preston constant k_{P}, and a rotational constant, which accounts for the translation of the plate rotation into a sample trajectory on the plate's surface.

2.3.2 Chemical mechanical polishing: CMP

As a first of the two proposed thinning processes (see Figure 2.1 on page 19) CMP (chemical mechanical polishing) is evaluated for establishing a well defined defect free backside surface after mechanical thinning. For this method, no epitaxial etch-stop layer is required.

When polishing, favourable starting conditions need to be established already during lapping: The grain size needs to be reduced successively to a final 3 µm when reaching final thickness. This ensures that the defect depth is shallow enough as not to endure polishing unnecessarily. Therefore, a sequence was established, where ca. 60 % of material to be removed are treated at 12 µm graining, another 30 % at 9 µm, until reaching the final thickness plus ≈ 5 µm at 3 µm graining. During polishing removal rate is less than 1 µm per hour using the specified equipment.

Polishing is carried out on the thoroughly cleaned equipment used for lapping. As a polishing plate, a hard, porous polyurethane pad was used in combination with Syton SF1 polishing slurry by Logitech. The slurry is a colloidal solution of silica particles with an

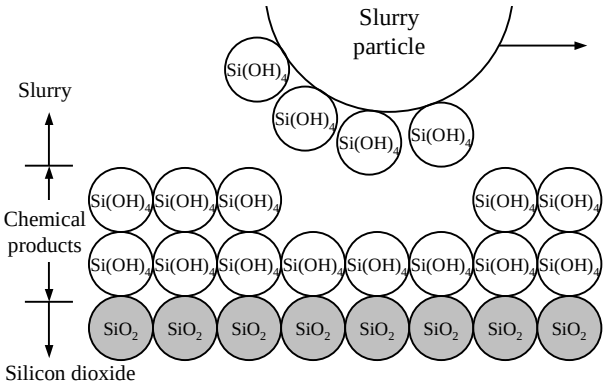


Figure 2.14: Reactions during chemical mechanical polishing of Si/SiO_2. After [53]

average diameter of 0.12 µm, stabilized at an alkalic (addition of NaOH or NH_4OH) pH-factor of ca. 10–11. The solution of these SiO_2 (silica) aerosols in water results in a dispersion of $Si(OH)_4$ molecules (silicic acid dispersion), which can solidify at pH-factors below 10 by initiation of the sol gel reaction. It is thus important to maintain the slurry's pH value.

The principle of CMP relies on two sequential chemical and mechanical processes. The chemical component of the slurry transforms SiO_2 surface molecules of the Si substrate into soluble $Si(OH)_4$ complexes [53],

$$SiO_2 + 2\ H_2O \rightarrow Si(OH)_4. \tag{2.3}$$

Then, the mechanical abrasive component removes these soluble chemical products from the surface and exposes new silicon material to the surface, where it is oxidized into SiO_2 by the chemical agent. Part of this cycle is depicted in Figure 2.14.

When the chemical components of the slurry react with ambient oxygen and with the silica of the sample surface, the pH-buffering agent is consumed. This leads to a decrease in pH-factor and in an initiation of the sol gel process: The Si complexes loose their OH-bindings and slowly polymerize into a solid glassy material. Thus, new slurry needs to be feed at a constant rate in order to provide sufficient buffering. Slurry remains need to be removed from the sample surface immediately after polishing, otherwise dried artefacts will remain.

During polishing, prominent structures are abraded at a faster rate, with the edges being rounded, while a high density of prominent patterns will result in less erosion of the structures in the center of such patterns [54]. This effect of pattern dependent polishing rate is an unwanted characteristic feature of CMP and can be controlled to a certain extent by using hard polishing plates rather than soft cloth-like polishing discs. The polishing rate also depends on the chemical and mechanical properties of the material to be removed, then resulting in a material related pattern dependance referred to as *dishing*. If sufficient control of these effects cannot be achieved by the polishing media, filling patterns have to be introduced at a design level.

The progress of polishing is visible by a higher degree of reflexivity and — at an advanced stage — by a reduction of point defect density as can be seen by optical mi-

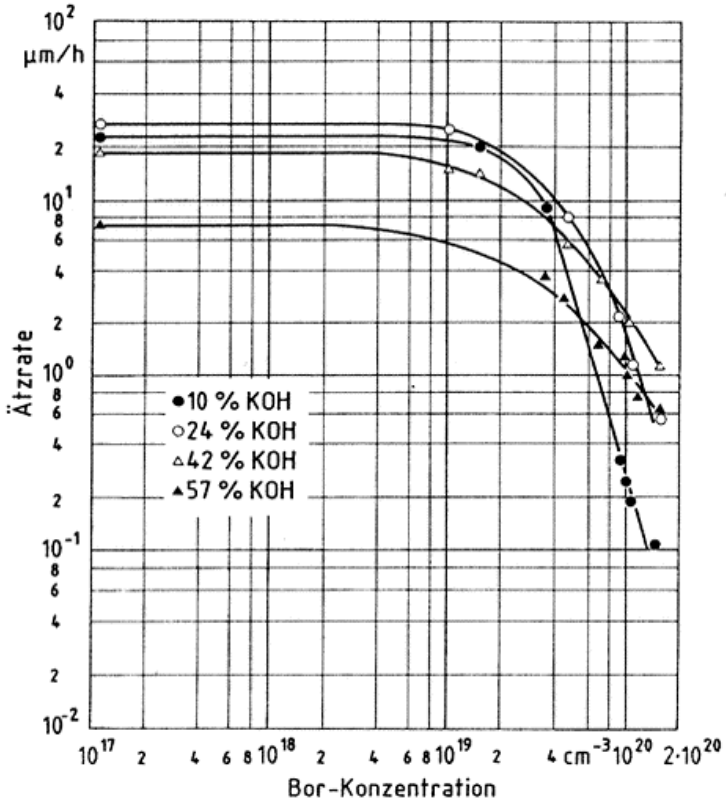


Figure 2.15: Etch rate of KOH solutions in boron doped silicon [43, 32]. Doping starts to affect the etch rate at boron concentrations above ca. 10^{19} cm^{-3}. At concentrations above 10^{20} the physical limit of B solubility in Si is reached. Solutions with 24 % of KOH exhibit the highest etching rate, whereas the highest selectivity at higher boron concentrations can be expected from 10 % KOH solutions

croscopy. Due to the softness of the polishing plate, the surface quality is better near the edges of the sample, where there is a step of about 20 µm in height, resulting from the thickness of the adhesive layer and the thickness of the remaining wafer substrate.

By means of polishing, scratches and defects from lapping can be removed with time, but formation of new scratches during polishing cannot be avoided completely, such that polishing time can be seen as a trade-off between point defect density and scratch formation probability.

2.3.3 Chemical thinning

As a second proposed thinning method (see Figure 2.1 on page 19), defect removal and surface recovery after lapping will be attempted by a selective etching technique, where the defective material is removed in a wet-chemical etching process. In this thesis only the principle is described, the evaluation of this process was part of related work [55].

As described in [32], with alkalic etching solutions based on EDP (ethylenediamine pyrocatechol), TMAH (tetramethylammonium hydroxide) or KOH (potassium hydroxide), the etching rate of Si is doping dependant, allowing an abrupt doping profile to function as an etch-stop. The Selectivity is highest with EDP, but not considered further in this work, for its toxicity and flammability. Among TMAH and KOH, KOH is chosen for its better availability and high doping sensitivity (see Figure 2.15). Clearly, in a

CMOS-application KOH could be replaced by TMAH because of ionic contamination.

The etching solution is applied in a spin-etching device, where the lapped samples are clamped onto a rotating chuck (see Figure C.4 on page 107). For chemical thinning, the lapping sequence can be optimized for fast processing by maintaining grain size at 12 µm during the whole sequence, saving time of the now obsolete sample and equipment cleaning. The wet-etching process is stopped, when the etch-stop layer is exposed at a large area of the wafer element, which is visible by a change of surface reflexivity from matte to reflective.

2.3.4 Results

Two methods of waferthinning have been evaluated, both of which rely on a first mechanical bulk material removal step, and subsequent defect removal. As a method for mechanical material removal, lapping was chosen as target process. From the defect depth of ca. 7 µm, measured after lapping with 3 µm Al_2O_3 grains, it can be estimated, that lapping produces defects that affect a region below the surface which extends to roughly 2 times the grain size of the abrasive particles employed.

Thinning with CMP For process evaluation, plain wafer elements with a size of ca. $45 \times 45\,mm^2$ were preprocessed as proposed in section 2.1 on page 19: Alignment markers are etched from the front side with a depth of ca. 20 µm, also marking the final thickness during lapping. After depositing the front side Cu metallization pattern, the wafer elements are bonded onto a handling substrate of the size $50 \times 50\,mm^2$. Waferbonding is performed using the the established two-layer temporary bonding process described in section 2.2 on page 21. The overall thickness of this configuration sums up to approx. 1375 µm, which includes the wafer element, bonding layers and handling substrate thicknesses. During lapping, this thickness was measured and used for endpoint detection and removal rate prediction.

Lapping was parted into a 3-step sequence, with typical grain size switching points at ca. 850 µm (to 9 µm) and ca. 750 µm (to 3 µm) of overall thickness. Down force and rotation speed were also reduced from 18 N/50 rpm at the beginning (1375 µm) to 10 N/40 rpm at the end of the lapping sequence. The process was stopped, when the first marker pattern was visible from the backside (720 µm), as can be seen in Figure 2.16. The depth of the marker holes then corresponds to the minimum thickness of the wafer element, however, the thickness variation is not known, and thus the maximum thickness cannot be determined precisely. The wafer elements were then subjected to polishing as described in section 2.3.2 on page 31.

Polishing can be stopped, when the defect density has dropped sufficiently in order to allow conformal oxide deposition and defined lithography for backside via etching. As can be seen in Figure 2.17, the surface quality can be recovered to an extent that enables PECCVD of SiO_2 and fine patterned lithography with feature sizes of 5 µm.

When further processed, the target depth of the vias needs to be determined. With

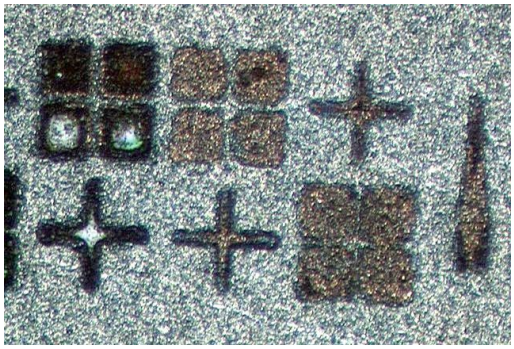


Figure 2.16: Visual lapping endpoint detection: Markers, etched from the front side are visible at the back

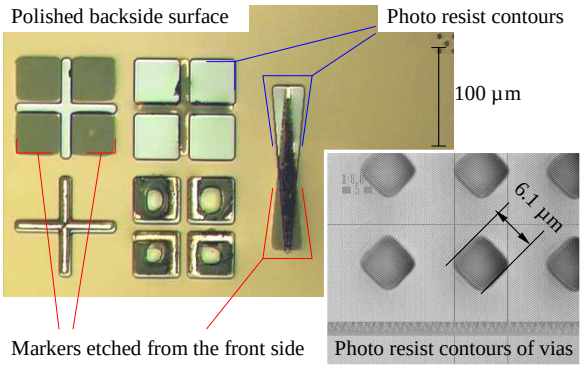


Figure 2.17: Alignment to front side structures and lithography on the thinned and polished backside

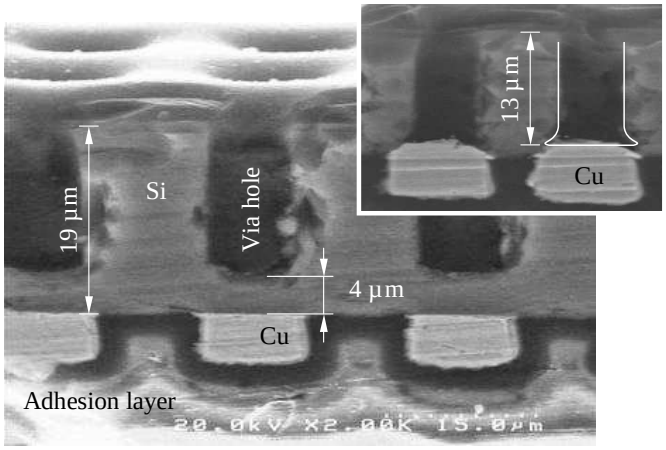


Figure 2.18: Thinned and polished wafer element cross sections after via etching. Tilting induced thickness variations are unpredictable and can lead to over-etched vias (insertion) and too short vias (4 µm of separation between via and Cu-pad in this case) on the same wafer element

high variations in thickness, only a few vias holes will be etched sufficiently, however, most resulting in too short vias or over-etched vias (Figure 2.18). When over-etched, the via profile evolves into a "trumpet"-shape (notching), caused by electrostatic charging of non-etching layers, which deflect the etching ions. This is unwanted, for it makes conformal deposition of sidewall isolation material more difficult.

CMP (chemical mechanical polishing) for defect removal after back lapping of Si wafers is a proper method to recover the surface for subsequent backside processing. However, it lacks a precise method of end-point detection, when the final substrate thickness is reached. It does not tolerate alignment errors and wafer bending during waferbonding, and results in considerable thickness variations after lapping.

Thinning by wet-etching Chemical thinning with a buried epitaxial etch-stop layer can tolerate these thickness non-uniformities to a larger extent and does not rely on the initial surface finish. Evaluation experiments with KOH solutions and boron doped etch-stop layers revealed a larger robustness of this defect removal method against wafer

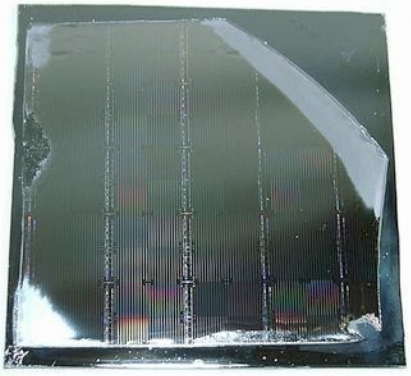


Figure 2.19: Surface of a $50 \times 50\,\mathrm{mm}^2$ wafer element after chemical wet-etching with an epitaxial p^{++}-etch-stop layer. The obtained surface is highly reflective and suffices subsequent PECVD and lithographic processing and via hole etching without the need for polishing

bonding induced alignment errors.

Figure 2.19 shows the surface of a wafer element with a buried boron doped etch-stop layer that has been wet-etched with a KOH solution. Surface quality was found to suffice further processing without the need for polishing. The alignment markers on the surface showed good contrast for precise alignment during the lithographic process for via-hole etching.

2.4 Conclusion

For waferbonding, only temporary waferbonding process is suitable, for the permanent waferbonding process lacks at method to detach or remove the substrate without destroying circuitry. The suitable process relies on a release layer of PMGI resist, which is cured in a way that prevents outgassing in the working temperature range. Adhesion is provided by a layer of BCB, which is also able planarize and embed metallization structures.

Waferthinning can be performed in both ways, provided that sufficient control of flatness is available during bonding and lapping. The tolerable thickness variations depend on the via-etching process and the conditions at the via bottom during etching.

The process employing CMP depends on an additional DRIE processes for etching of alignment and depth markers from the front side. The process can be useful, where buried epitaxial etch-stop layers are not available and where vias cannot be etched from the front side because they would block the CMOS metal layers. The few alignment markers that need to be etched from the front side can be placed in the external non-circuit area typically used for optical alignment of a wafer. In the experiments, a typical target thickness was $20\,\mu\mathrm{m}$ which could vary from $13 - 19\ \mu\mathrm{m}$ when thinned.

The process that uses wet-chemical selective etching in combination with an epitaxial etch-stop layer is tolerant to wafer bending and tilting during wafer bonding, when the vertical alignment error does not exceed a certain limit. It results in a uniform thickness of the thinned wafer element, and easily meets the uniformity requirements of the via etch process. In the experiments, a target thickness of $10\,\mu\mathrm{m}$ was reached with little variations.

Chapter 3

Patterned electroplating and via filling

For evaluation chip-to-chip bonding techniques such as solid liquid interdiffusion bonding (SOLID), front side and backside metallization patterns are required in a layer thickness range of $2-4\,\mu m$ [56]. Sputtered or evaporated films are usually deposited in a thickness of less than $1\,\mu m$, which is attributed to the typically low deposition rates of a few nm/min. Metal patterns of that thickness have a favourable width to height ratio, and can easily be produced by means of lithographic lift-off or back etch techniques.

Most suitable technique for the deposition of $> 2\,\mu m$ metal films, as needed in this 3D-integration scheme, is plating/electroplating. Due to its nature of bottom-up growth rather than top-down deposition, these films can be patterned by passivation of no-growth areas, also referred to as selective growth or through-mask plating. Moreover, electroplating, as instrumented to fill high aspect ratio features in damascene[1] interconnect technology, could also be used to fill the usually larger features necessary to produce through chip interconnects.

3.1 Introduction

Electroplating of Cu and Sn, as used in SOLID bonding, are long known and available processes [57], also used in printed circuit board technology, however, with feature sizes being typically a hundred times larger. During the past 10 years electroplating of Cu was introduced to interconnect technology as copper damascene technology, with electrolyte compositions optimized for small features and high aspect ratios [3, 58]. These newly available high performance electrolytes show potential to meet patterning and via filling requirements of this 3D-integration scheme and thus will be considered for implementation in this chapter.

Figure 3.1 relates the processes where electroplating is most suitable to the process chain. Another advantage of electroplating is its low-temperature capability, which makes it an ideal candidate for back-end processing (processes after active circuit fabrication, usually interconnect and packaging related). The superior electrical conductivity is a further gain, but was not a major concern of this work.

In general, two modes of plated filling/patterning techniques need to be distinguished: Bottom-up filling of isolated structures, as is the case in through-mask plating, and bottom-up filling of conductive structures, as are the conditions for via-filling. For via

[1] Technology to produce copper interconnects by filling and planarization (CMP) of trenches in dielectric material (usually silicon oxide or low-κ material)

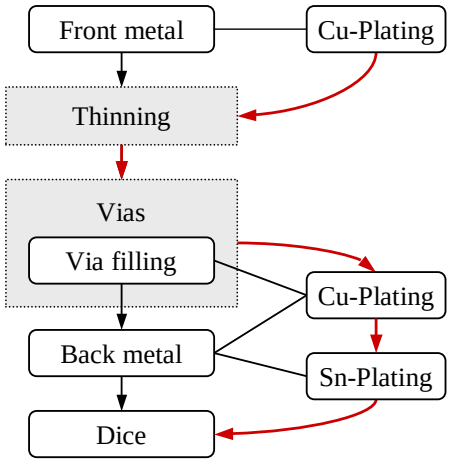


Figure 3.1: Electroplating processes in the process chain. The front and back metal patterns can be produced by means of through-mask (selective) electroplating of Cu and Cu+Sn respectively. Moreover, the processes for via-filling and backside metallization can potentially be combined

filling, a combined process will be evaluated, which results in filled vias and patterned backside metallization for bonding.

3.1.1 Challenges

The implementation of electroplating processes in 3D-integration is associated with the following concerns, which will determine applicability:

Global uniformity: The layer thickness must not vary over the surface of the wafer element as to allow for intimate contact of all soldering pads. This issue may be partly tolerated by the adaptability of liquid solder, but can also lead to squeeze-out of solder. Homogeneous conductance and surface wettability of the seed layer, as well as good bath agitation will be necessary to achieve this.

Surface quality: The surface of the plated films need to be smooth and clean enough to allow a good wetting of liquid solder. This applies mainly to the Cu films, since Sn is deformable and will melt during the bonding process. No local peaks and big grains should be present in the Cu films, which could act as spacer and avoid surface contact or cause tilting during bonding.

Filling capability: The electrolyte should be able to fill via holes, where a voidless conductive path needs to be formed in a high aspect ratio hole. This will also require a conformal deposition of the seeding layer.

Bath maintenance: High performance plating baths have a complex formulation and often need extensive effort to maintain a working composition of the bath. The aim is to find a robust plating formulation with stable components for experimental use.

Electrical performance: The Cu deposit should have a low resistivity as close to bulk Cu as possible. This requires also that contamination by plating additives is as low as possible.

3.1.2 Methods of electroplating

Various techniques of electroplating presently exist [57]. Prerequisite for any plating method is a metallic surfaces for the build-up of an electrochemical potential between electrolyte and electrode, which is the driving force for the ionic transport of deposition material. For the formation of a metallic pattern on a dielectric, i. e. SiO_2 (as in this case), a metallic starting layer, or *seed layer* is necessary.

In electroplating, an electrochemical potential is applied between two immersed electrodes, the anode and the cathode. The sample serves as cathode and needs to be covered by an unpatterned seeding layer, which short-circuits all metal areas on the sample. The seed layer used in this work consisted of an ion-beam sputtered thin films of TiW with a target content of 30 at-% Ti, which served as adhesive layer [59] and diffusion barrier [60, 61] as to prevent diffusion of Cu into Si. The subject of diffusion barriers for Cu will be treated from another perspective in section 5.2.1 on page 86. A variety of plating methods is available, with options as listed in the following:

Chemistry: Acidic, alkalic. Because of Si-etching, alkalic electrolytes usually cannot be used when Si is exposed. Acidic baths are widely used for Cu plating. — **acidic electroplating bath**

Principle: Sacrificial anode, inert (platinum-) anode. Sacrificial anodes are consumed during operation and need dopants (phosphorous copper, anodic copper) as a means to control their consumption rate and ionic concentration in the electrolyte. For permanent operation, inert anodes are used, and the concentration of Cu ions is maintained by adding Cu pellets. For use in an experimental setup, the **sacrificial anode** principle is the first method of choice.

Additives: Surfactants, brightener, leveler, accelerators, inhibitors, polymers, foreign ions, organics. Most formulations will contain surfactants for wetting enhancement. Other additives effect the properties of the deposit (leveler, brightener) or control mobility of ions for certain topology configurations (accelerators, inhibitors). In this work, additives for a **smooth surface**, that enhance growth in a confined hole topology for **via filling** are favoured.

Setup: horizontal, vertical, fixed, rotating. Electrode distance (5 – 10 cm) and electrode area ratio (anode:cathode $\approx$ 3:2) are mostly specified by the electrolyte supplier. For a uniform deposit, sufficient **agitation** of the electrolyte and refreshing with Cu ions is necessary.

Operation: dc, pulsed, reverse pulsed. During pulsed operation, the ionic double layer near the cathode surface is reestablished several times per second. During the pause, the bath can regenerate itself. Pulsing can be used to control the microstructure of the deposit [57].

3.1.3 Plating characteristics on plain substrates

First experience on electroplating of Cu with acidic electrolytes was gained on metallized glass substrates with Cu and Au metallization and spots of photoresist for local passivation and thickness measurement. A setup was used, where the electrolyte was filled into a glass recipient, with the electrodes aligned vertically at a distance of ca. 5 cm. Current was adjusted to ca. 10 mA/cm^2.

Electrolytes with no brightener or leveler additives lead to a matte surface appearance with crystallites $> 1\,\mu$m and showed a tendency towards hillock formation. A considerable degree of aging could be observed already after a storage time of only a few hours, which lead to a transition of color appearance from bright to dark red. It is assumed that the rough surface can bind more ambient humidity which accelerates oxidation of Cu. Deposits from electrolytes with brightener or leveler additives had a close to mirror-like surface finish with crystallites $< 1\,\mu$m in size. These Cu layers where better controllable in uniformity and did not show any storage induced degradation.

An electrolyte which demonstrated the latter behaviour was Slotocoup BV 30 from Schlötter [62], a sulfide based acidic copper bath with additives of sulfuric acid, hydrochloric acid and the additives BV 31, BV 32 and BV 33, which where mixed according to the instructions of the data sheet. Sulfuric acid addition is known to increase conductivity of the electrolyte, and thus overall uniformity, whereas the other components influence brightness (reflectivity), crystallinity and leveling. This bath was designed to be used in PCB applications, where blind board vias need to be filled.

The room temperature deposition rate of Slotocoup BV 30 at a current density of ca. 10 mA/cm^2 was determined to be 0.6 µm/min. The bath was used in combination with anodes of the composition 99.9 % Cu and $0.04 - 0.06$ % P.

For the Sn layers, the electrolyte SAT 20 [63] was chosen, which resulted in a matte surface appearance with crystal grains $> 1\,\mu$m. During work on the SOLID process it was proposed, that a rough Sn surface with larger crystals shortens the time for melting by raising the density of melting centers [56]. SAT 20 is a acidic electrolyte for sacrificial anodes without doping, designed as a metallic resist in PCB board manufacturing. High purity anodic material was used, with a tin content of 99.95 %. The components SAT 22 and SAT 23 are added as specified in the data sheet in order to control ductility and grain size of the deposit.

Deposition rate was measured to be ca. 0.5 µm/min, and the surface roughness (profilometer) R_A was ca. 700 Å at a current density of ca. 10 mA/cm^2. It is known, that Sn films, deposited onto Cu are subject to an aging process referred to as “whisker” growth [64]. However, no such phenomenon was observed on the matte Sn deposit resulting from SAT 20.

3.2 Through-mask plating

For the patterning of the front side metallization a selective growth technique was used rather than back-etching. In an anisotropic wet-etch process the size/height ratio of the features to be realized will result in under-etching and degradation of corners and sidewalls of the structures, while selective growth by through-mask plating will reproduce the vertical and straight sidewall profile given by the isolated resist mask. The resist features are filled in a bottom-up mode, with no deposition taking place at the sidewalls (as compared to the via-filling mode with conductive sidewalls).

In this 3D-integration scheme, the backside metallization pattern is intended to be grown simultaneously during via-filling. This combines the two modes of feature filling and has the advantage, that no planar copper removal process, which could degrade dielectrics, needs to be established. Most of the advanced electroplating formulations are suited for aspect ratios > 3 and footprints $> 1\,\mu m$ [65], thus the hundred times larger features of the metallization pattern, are within a scope were they should not affect filling capability. However, hillock formation can be the result of local overfilling and can lead to a bumpy pad surface.

3.2.1 Plating setup

To obtain a uniform metallization pattern, the plating bath needs an excellent concept of agitation and electrode placement. Such concepts can optimize the cell design by considering flow patterns and electric field shaping by placing secondary electrodes [66] or including ultrasonic actuators [67]. For this patterning and filling attempt, a simple setup configuration will be considered, since filling happens on a micro scale, and macroscopic design issues are considered less influential to microscopic features than plating bath electrochemistry (which will be treated in section 3.3 on page 45). Thus, on a macro scale, uniformity and refreshment cycle issues are addressed only.

For the experiments to be undertaken in this work, an optimal plating cell design is considered to be a rotating disc electrode (RDE) configuration in a horizontal setup, where the sample is placed on top of an upstream filled overflow cup [68, 69]. This configuration provides a high electrolyte refreshing rate and a good uniformity, since the sample constantly changes its position relative to the upstreaming flow of electrolyte and averages local non-uniformities. Such design has been implemented for laboratory use (see Figure C.1), with an adjustable electrode distance between $4 - 5\,cm$ and a circular anode with a diameter of $8\,cm$ at a maximum sample size of ca. $2.5\,inches$.

3.2.2 Patterning

In SOLID bonding, contact pads with a solder layer on top typically will have a height of $3\,\mu m + 2$ (to 3) $\mu m = 5$ (to 6) μm. For through-mask plating, a resist pattern of such thickness, with ideally vertical sidewalls needs to be produced. Options exist within the

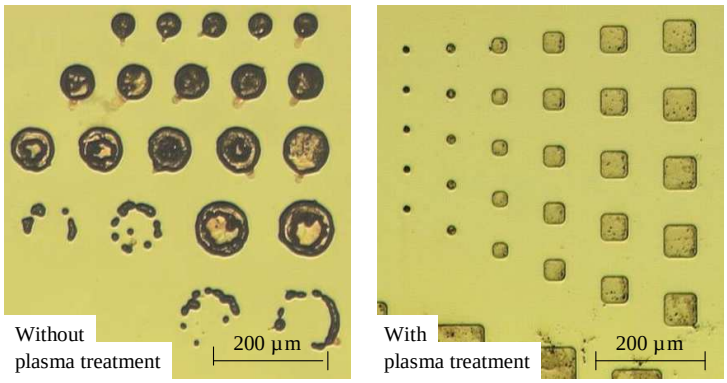


Figure 3.2: Patterning by through mask plating. Left: Insufficient wetting of the resist surface leads to incomplete patterns. Right: Good wetting can be achieved with an oxygen plasma pretreatment

following coating media:

- Omnicoat/SU-8: SU-8 is best cured in a two-step ramped baking procedure and available in a broad range of coating thicknesses (1.5 µm to 250 µm). When cured, it does not dissolve in conventional solvents, and an underlayer of Omnicoat is needed for lifting of SU-8 patterns.
- AZ 5214/PMGI-SF 19: The standard thin film photo resist AZ 5214 E (max. 2 µm) can be used to flood expose the thick film resist PMGI-SF 19 (max. 8 µm). The two-layered undercut profile can be used to shape the contact.
- AZ 9260: Thick positive AZ photoresist (max. 10 µm) which can be dissolved completely in acetone after plating. No specific heat ramping and exposure techniques are necessary.

Clearly, AZ 9260 photo resist has the advantage of simple and fast application at a good pattern resolution. According to the manufacturer's notes [70], the resist can be patterned at a resolution of 3.5 µm of lines and spaces at a two-step spin-on layer thickness of 24 µm. Thus, features with an aspect ratio of $5 - 7$ can be produced with this resist.

During plating of samples masked with AZ 9290, wetting of the Slotocoup BV 30 electrolyte with the seedlayer's and photo resist's surface was not sufficient in order to achieve good results. As can be seen in Figure 3.2, the deposit typically formed near the edge of the resist window, leaving a blank center. This can be explained by bubble trapping in the resist window, where the root area of the bubble keeps the electrolyte from the center of the resist window.

Since the used formulation already contained surfactants (usually ionic tensides, i. e. lauryl sulfate), wetting could only be enhanced by means of changing the resist's surface wetting capability. The samples where subjected to an O_2-plasma treatment of ca. 2 min, which resulted in a more hydrophilic oxygen terminated surface of the photo resist mask. Plated patterns achieved after this pretreatment are shown in Figure 3.2.

Thickness control during plating is available through profilometry: Initially, the height of the photo resist structures is determined. Then, the known characteristic deposition rate is interpolated to a target thickness and plating time, after which the sample is

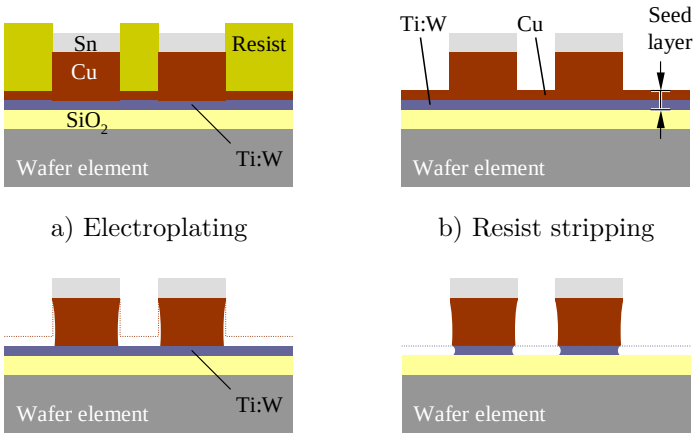


Figure 3.3: Seed layer removal procedure after through mask electroplating. Cu start layer and TiW diffusion barrier layer are removed with different etching chemistry

removed from the plating bath. Plating progress and true deposition rate can now be determined from the differential depth of the resist structures. Eventually, plating can be repeated until the final thickness of ca. $3-4\,\mu m$ for Cu patterns is reached. For backside metallization patterns, $2-3\,\mu m$ of Sn plating follow.

3.2.3 Seed layer removal

After Electroplating and resist stripping, the short-circuited metal structures need to be released by removing the thin conducting seed layer from the previously masked regions, where no metal deposition occurred (see Figure 3.3). For this procedure, the following aspects need to be addressed:

1. The seed layer etchants may attack, but must not degrade the surface of the soldering contacts. Etching processes need to be found, which do not degrade Cu, SiO_2 or Sn surfaces and have a reaction rate that is slow enough to be uniform.

2. No metallic remains on the dielectric surface are tolerable, since this can lead to conductive paths and shorting. However, over-etching will degrade the sidewall profile and can even lead to contact lift-off (as suggested in Figure 3.3d).

Sputter etching with Ar ions, RIE with CF_4, and wet-etching solutions with formulations based on sulfuric acid, ammonium hydroxide and ammonium chloride (see Table 3.1) were subjected to seed layer removal experiments. The experiments revealed, that the process step of seed layer removal is uncritical on the front side. However, on the backside complete removal is a challenge when surface quality could not be recovered to a certain degree of roughness.

	Dry etching		Wet etching		
	Ar Sputter etch	CF_4 RIE	$H_2SO_4 + H_2O_2 + H_2O$ (5:1:200)	$NH_4OH + H_2O_2 + H_2O$ (1:2:200)	$NH_4Cl + H_2O +$ Cu (commercial)
Cu	(+), redeposition	–	+	(ox), tolerable	+
Sn	(+)	–	ox	–	–
TiW	(+)	+	–	+	–
SiO_2	(+)	+	–	–	–

Table 3.1: Available seed layer removal processes. Sputter redepositions are tolerable at high power/short processing time. Sulfuric acid based etching destroys the Sn surface. Symbols: + etched, – not etched by agent, (+) redepositions, ox strong oxide formation

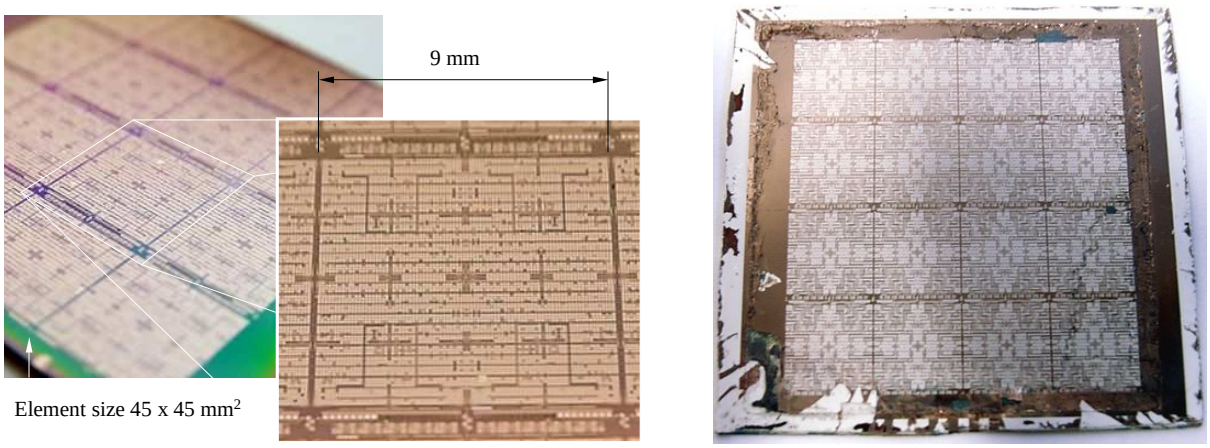


Figure 3.4: Left: front side Cu metallization pattern; right: Backside Cu/Sn through-mask plated wafer element after successful seedlayer removal

3.2.4 Results

On the front side metallization, both wet-chemical and dry-etching seed layer removal techniques could be applied successfully. Sputter redepositions could only be observed during long process times at low power Ar sputtering. The Cu start layer can be optimized for a minimum thickness that still provides plating uniformity. This will reduce removal time and minimize Cu redeposition.

The sulfuric acid based etching solution cannot be used in the presence of Sn, for it transforms the patterned Sn into a porous conglomerate, which can not be removed without destroying the Sn cover. On the backside, the TiW underlayer interlocks with the rough SiO_2 surface and can not be removed by wet-etching within reasonable time. Ti-oxide my have formed at the interface, which needs to be cracked first. If the etching time is prolonged in such a case, lifting of the small footprint structures of $9 \times 9\,\mu m^2$ can occur. On the backside it is thus favourable to remove the seed layer by means of dry-etching.

The sample X 120, shown left in Figure 3.4 was plated using Cubath SC electrolyte [71]. The Cu start layer is high power sputter etched with Ar ions at 1500 W for 10 min. TiW

is then etched with CF_4 reactive ions at 400 W for 2 min. Both processes were carried out at a pressure of 40 mTorr, where etching characteristics shows high anisotropy, as to prevent excessive side wall etching.

Shown on the right in Figure 3.4 on the preceding page is a thinned and completed $45 \times 45\,mm^2$ wafer element front-side-down on a $50 \times 50\,mm^2$ handling substrate. The Cu/Sn metallization was plated as described earlier in this section, and the seed layer was removed by wet-etching. The Cu start layer was removed with a commercial ammonium chloride based etching solution, designed for applications, where Sn is used as a metallic resist for Cu etching. Then, the TiW diffusion barrier layer was removed with a diluted solution of ammonium hydroxide with hydrogen peroxide (see Table 3.1).

3.3 Via filling process evaluation

In the previous section it has been shown that µm-sized pad structures can be generated by means of plating through a dielectric removable mask. It has to be pointed out, that up to now, only structures with isolated side walls could be filled reliably. However, as is the case for the backside etched vias, no access to a short-circuited common conductive layer at the bottom of the vias exists, which would enable a similar filling technique where the resist would be replaced by the thinned and isolated wafer itself.

For via-filling in this 3D-integration scheme, seeding can only take place from the open side of the via hole. Then, a conductive path to the bottom of the via needs to be established, which also includes the isolated sidewalls of the via hole.

This scenario is similar to that of the damascene interconnect formation technology, however, with a bigger feature size, and aspect ratios in the range of 2..3. Work at IBM Research Center has proven Cu filling of small features by electroplating to be reliable and and triggered the replacement of Al evaporation techniques for interconnects in integrated circuits [58, 72]. Such filling technique is an ideal candidate to be implemented in 3D-integration via-filling scenarios.

3.3.1 Factors influencing feature filling

Conventional plating formulations are optimized to give a bright and uniform (leveled) deposit over large features as compared to the ionic diffusion length, which is typically in the range of a few µm. However, when features become smaller, mass transport into micron and submicron range via features is limited and results in a concentration gradient inside the via, leading to a lower deposition rate on the inside surfaces. According to [58] this phenomenon is termed “subconformal” deposition, for snapshots of the via profile show a *decreasing* layer thickness with feature depth.

A common failure mode associated with subconformal via filling are trapped cavities (voids), which can contain electrolyte and gradually destroy the via over time. In the best case, conformal fill can be achieved, where the void is reduced to a one-dimensional seam, which still presents a weak point of the via structure.

For better results, plating formulations have been found, that allow for "superconformal" filling of submicron sized high aspect ratio features. This newly introduced term describes a filling process, where the deposition rate *increases* with feature depth, thus allowing bottom-up filling of high aspect ratio vias.

Plating bath formulation The key to achieve superconformal filling of high aspect ratio via structures is plating bath chemistry. The formulation is extended by additives, which can counteract and even reverse ionic depletion in deep vias:

Accelerators adsorb to the surface and locally increase the deposition rate according to their concentration. They accumulate in features with little electrolyte circulation like corners and via structures. Working accelerator molecules are compounds with pendent sulfur atoms, such as organic mercapto species (C-H chains or rings with a pendent SH group) or thiourea [73, 74].

Suppressors are molecules, that — due to their size and length — predominantly adsorb to the top surface at the via opening and inhibit or decrease Cu deposition there. Mostly polymers, i. e. PEG (poly ethylene glycol) are used, which develop their suppressing property in the presence of chloride ions = co-suppressor [73].

Levelers are suppressing molecules that do not adsorb to the surface, but are transported in the electrolyte by means of convection or diffusion. Thus, they are more present near exposed surfaces like edges or prominent features like bumps, and can avoid local overplating due to accelerator remains after via filling [65].

A commercial acidic sulfur-based electroplating bath with accelerator and suppressor chemistry was available through Enthone's Cubath SC product [71].

Seed layer properties Apart from the plating bath chemistry, seed layer properties are most important factors influencing superconformal deposition into high aspect ratio via features [73]. Among these are: conformity and step coverage, sidewall morphology, Cu oxidation, overhang and asymmetry. Some of these factors are less influential, as deviations from the optimum state are tolerable to some extent.

Seed layer deposition is available by ion beam sputtering onto TiW (for adhesion and barrier layers) and Cu targets with resulting surface layer thicknesses of ca. $10 - 50\,\mathrm{nm}$ for TiW and ca. $50 - 100\,\mathrm{nm}$ for Cu. Usually ion-beam sputtering is directional, however seed re-sputtering from the via-bottom to the lower sidewalls can enhance conformity and lessen shadowing effects. Good coverage of TiW as a barrier layer is essential as to prevent Cu atoms from diffusing into active Si areas where it can act as a charge carrier trap and degrade circuit functionality.

Good wetting and immersion of the seedlayer can only be achieved, when copper oxides which are formed on the surface after seed layer deposition can be dissolved prior to plating. Acidic plating solutions have the capability to dissolve thin layers of CuO and Cu_2O cupric oxides of a thickness of a few atom layers. This property of acidic

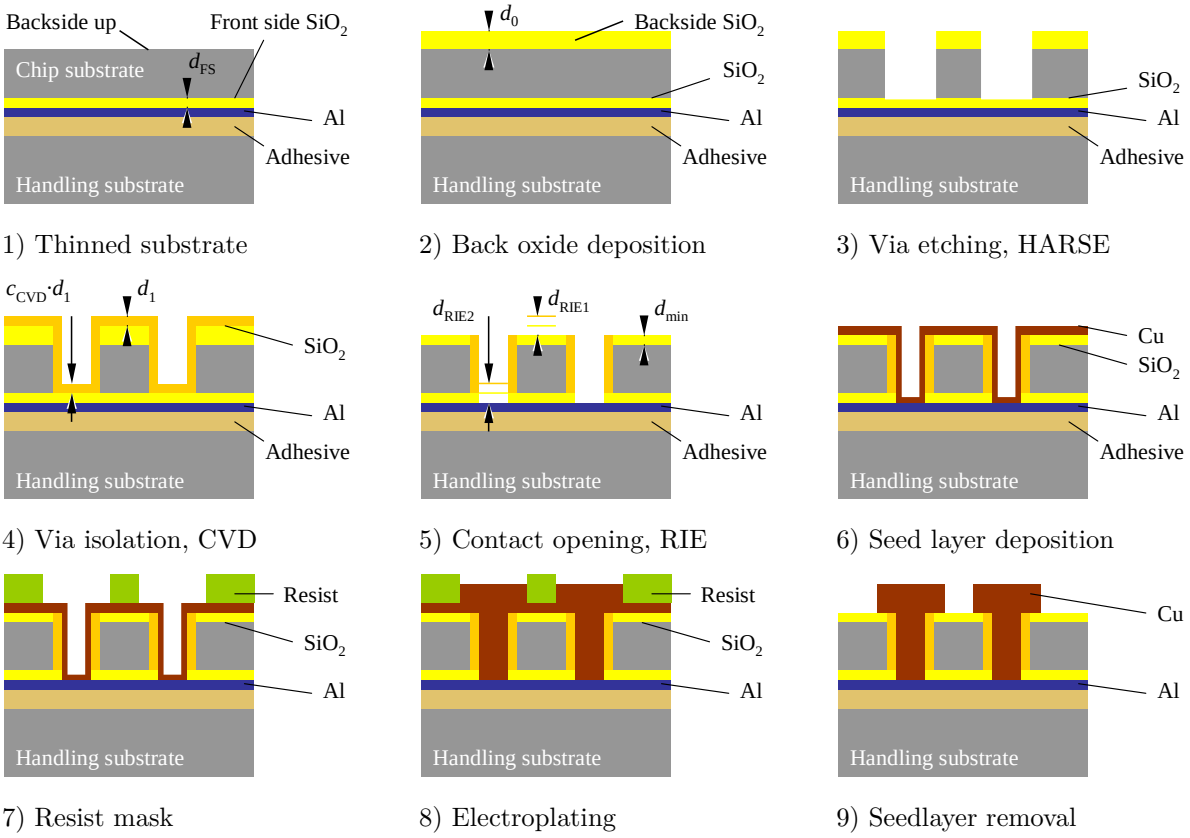


Figure 3.5: Experimental via-etching and inside-via processes for a minimum feature $5 \times 5\,\mu m^2 \times 20\,\mu m$ on permanently bonded and thinned Si chip substrate

electrolytes could be confirmed by observations, where a Cu seedlayer was completely consumed by the electrolyte, when no external plating potential was applied. It could also be observed, that oxygen plasma pretreatment of the resist mask prior to through mask plating did not affect immersion capabilities of the seed layer.

3.3.2 Via preparation

Experiments on via filling were first carried out on etched plain wafers with the desired via topology. A cross sectional view of a trench structure was prepared by means of wafer sawing and SEM, which showed a complete filling of the structure.

For the next stage of evaluation experiments, thinned and polished wafer elements, manufactured as described in chapter 2, with a *permanent* BCB waferbond produced with the pneumatic tool of Figure C.2 on page 105 were available. The substrate was covered with an unpatterned 1 µm thick Al front-side metallization and thinned to a thickness of ca. 20 µm. A lithographic mask was designed, with single via structures and probing pads for electrical and cross sectional evaluation.

As sketched in Figure 3.5, the thinned backside was covered with a 1000 nm thick

layer of silicon oxide, which was structured with CF_4 RIE to serve as electrical isolation and etching mask (see Table B.7 for process details). Two via sizes with mask features of $5 \times 5\,\mu m^2$ and $7 \times 7\,\mu m^2$ were realized. The vias were etched in the Department of Solid State Physics at the University of Ulm, where a two-cycle ASE (anisotropic silicon etching) process was available. This process was based on a SF_6/CHF_3 gas system, a He-cryogenic regime [75, 76] and an inductively coupled plasma (ICP) in order to achieve a high density of reactive ions. For a depth of 20 µm 100 gas exchange cycles were necessary. Mask erosion was critical, as it was determined to be in the range of ca. 2 nm/min at a Si etching rate for 5 µm vias of ca. 160 nm/min (selectivity $S \approx 80\,\%$).

Since part of the SiO_2 mask needs to remain to isolate the soldering metallization against the chip substrate (refer to steps $2-5$ in Figure 3.5), an appropriate thickness is necessary to account for mask erosion and back-etching in the ASE etching cycles:

$$\Big((d_0 - d_{\mathrm{ASE}}) + d_1 - d_{\mathrm{RIE1}}\Big) > d_{\min}, \quad \text{mit} \tag{3.1}$$

$$d_{\mathrm{RIE1}} = \frac{1}{c_{\mathrm{RIE}}}\, d_{\mathrm{RIE2}} \tag{3.2}$$

$$d_{\mathrm{RIE2}} = c_{\mathrm{CVD}}\, d_1 + d_{\mathrm{FS}} \tag{3.3}$$

$$d_0 > d_{\min} + d_{\mathrm{ASE}} + \frac{1}{c_{\mathrm{RIE}}}\, d_{\mathrm{FS}} + d_1 \Big(\frac{c_{\mathrm{CVD}}}{c_{\mathrm{RIE}}} - 1\Big), \tag{3.4}$$

with the minimum backside oxide thickness $d_{\min}$ after processing, the mask material d_{ASE} removed during ASE (mask erosion), the deposited front-side oxide layer thickness d_{FS} between the chip substrate and the first metal layer, and the layer thickness of the via-oxide deposit d_1 at the surface. The geometry constants c_{CVD} and c_{RIE} represent the ratio between reactions at the surface and at the via bottom during oxide deposition (CVD) and etching (RIE). d_{RIE1} denotes the amount of oxide etched at the surface, and d_{RIE2} the amount etched at the bottom of the via. Assuming the following values: $d_{\min} = 250\,\mathrm{nm}$, $d_{\mathrm{FS}} = 250\,\mathrm{nm}$, $c_{\mathrm{CVD}} = c_{\mathrm{RIE}} = 0.5$, and a determined mask erosion $d_{\mathrm{ASE}} = 250\,\mathrm{nm}$, the required backside oxide layer thickness d_0 to be deposited amounts to 1000 nm.

In later experiments, the Bosch-process was available for via-etching. Here, the temperature control and mask erosion ($d_{\mathrm{ASE}} < 50\,\mathrm{nm}$) are not as critical as in a cryogenic ASE-process, and the mask oxide thickness d_0 can be reduced to 800 nm. During etching, the feature depth strongly depends on the opening area, as with all known ASE processes [77]. Thus, in a final layout, all vias should have a uniform size in order to achieve a uniform etching depth.

After etching, the wafer element was diced into $10 \times 10\,\mathrm{mm}^2$ chips to allow more parametric variations to the following yet unknown inside-via deposition processes. CVD and PVD techniques for sidewall dielectrics, and BV 30 and Cubath SC plating baths for via-filling are evaluated. Before metallizing the via inside, the dielectric deposit at the bottom of the via hole needs to be removed for making contact with the front metallization (in this experiment the unpatterned Al layer). Therefore, low pressure RIE with CF_4 (appendix B.4) was considered adequate, since conditions are beneficial as to achieve an anisotropic etching characteristic, where the sidewall oxide can be protected by means of

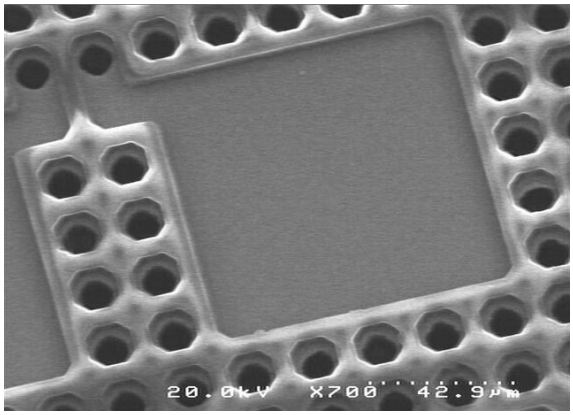


Figure 3.6: Via holes and resist mask test pattern prior to filling. The via features remained free of resist. The large area resist features will form contact pads for measurement with probe tips

shadowing. The seedlayer was deposited using electrode and ion-beam sputter deposition techniques. Setup configuration and process details are described in appendix B.4.

As a plating mask, AZ 9260 resist was used. SEM micrography showed, that after lithography, no resist remains where found inside the small $5 \times 5\mu m^2$ footprint via holes (Figure 3.6). However, remains could be found in the bigger via holes in regions, where alignment resulted in a fractional exposure of resist inside the via. The following effects are thus considered beneficial:

- High viscosity of the photo resist and small via footprint. When the resist is spun on, the surface remains sealed, and ambient air is trapped inside the via, which completely remains resist free.
- Full exposure of possibly resist-soaked vias will help to completely remove the resist during an extended development step.

3.3.3 Via filling

Prior to filling, the samples were subjected to O_2-plasma treatment (as listed in Table B.1) to enhance wettability of the resist surface, then clamped onto the substrate holder of the electroplating setup shown in Figure C.1 on page 105 and immersed in the plating bath. Plating time was between 30 and 60 min at current densities of 7 to 15 mA/cm^2. To obtain a correct value for the current density, the resist-covered surface area was subtracted from the sample area. Substrate rotation speed was adjusted to approx. 60 rpm.

At this stage, Sn plating to produce a solder layer was omitted, since this aspect will be addressed separately in chapter 4. With no Sn present, the seedlayer can be removed with the sulfuric acid and ammonium hydroxide based formulations of Table 3.1 on page 44.

3.3.4 Results

Reasonable filling of ion-beam seeded $7 \times 7\,\mu m^2$ footprint vias could be achieved with Schlötter's Slotocoup BV 30 plating bath, whereas good filling of both 7 µm and 5 µm

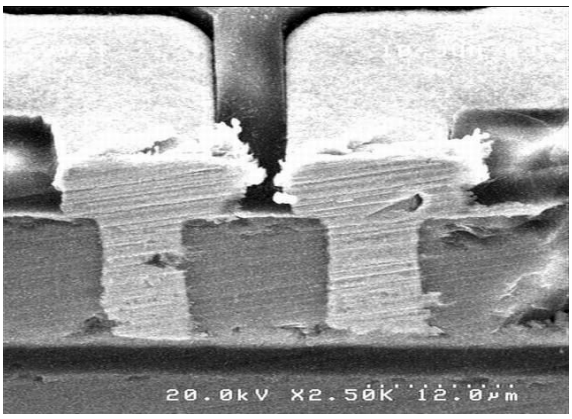


Figure 3.7: Filled via and resist structure after Cu electroplating onto and into a 20 µm thin Si substrate

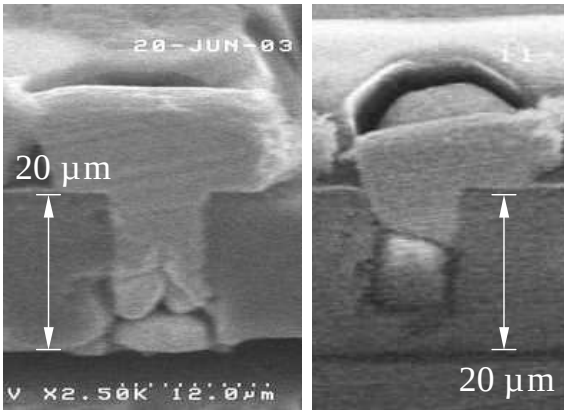


Figure 3.8: Plating failure modes associated with seedlayer discontinuities (l) and resist residues (r) inside the via-hole

feature vias was possible with Enthone's Cubath SC plating bath (see Figure 3.7). This was expected, since Slotocoup is designed to fill typical PCB macro-sized via-holes, and Cubath is based on a more complex formulation to fill features in a submicron range.

Incompletely filled vias were produced in both plating baths, and could mostly be associated with seedlayer discontinuities in the lower sidewall region or with resist remains inside the via-holes due to a misaligned resist pattern (see Figure 3.8).

Shown in the figures are micrographs of raw cut cross sections, obtained after sawing through the $10 \times 10\,\mathrm{mm}^2$ samples with a 30 µm thick hub-type nickel-bond dicing wheel. The regular via pattern helped to achieve a high quote of visible through-cut vias. Another method of evaluating the filling capability of the electroplating process is the selective removal of the complete Si substrate, the copper was plated into. This involves a high-rate Si RIE process with CF_4, which has the property of not attacking Al and Cu surfaces. As can be seen in Figure 3.9, the filled via-structures are visible as free-standing plugs, with the sidewalls fully visible for evaluation.

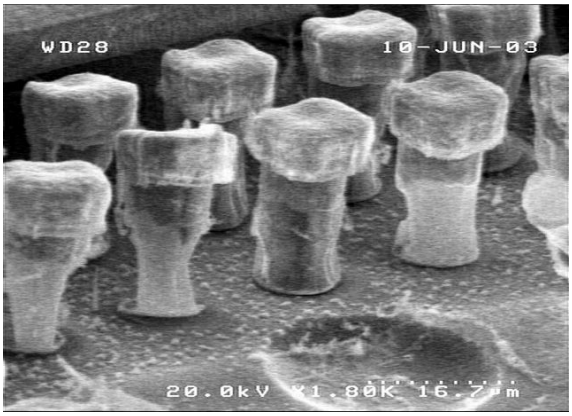


Figure 3.9: Via-plugs after removing the Si substrate during a high-rate dry-etch process from the backside. Etching stops on the front side Al metallization. Most of the vias show continuous sidewall metallization

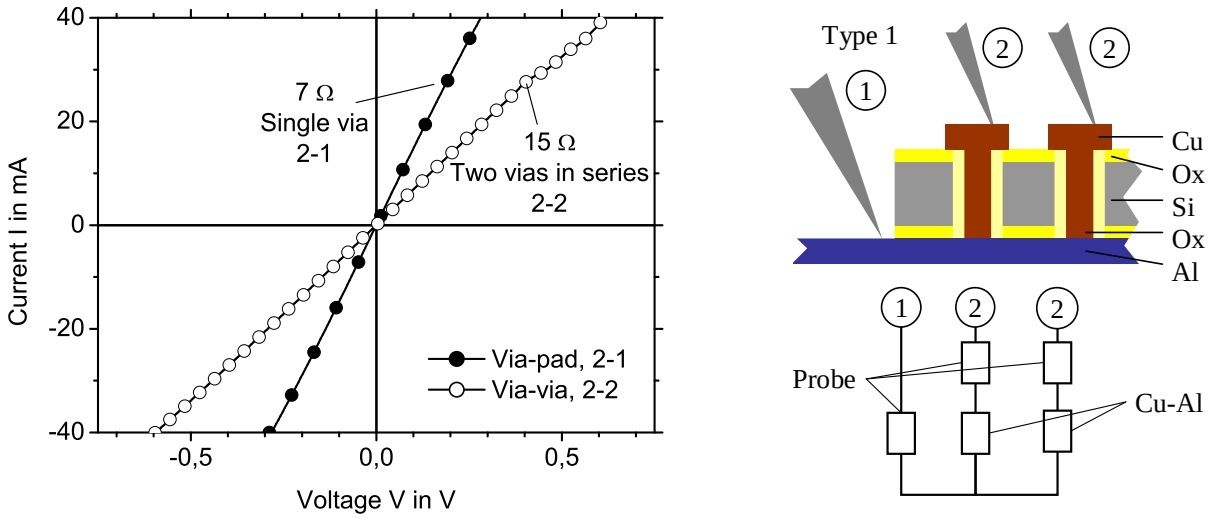


Figure 3.10: Resistance of a single via and two vias in series. The the high value can be explained by dielectric remains in the via

For an electrical characterization, two types of sample structures were manufactured: Type 1 samples where subjected to anisotropic oxide etching to establish a through contact, whereas type 2 samples were left with the oxide layer at the via bottom. Each sample was characterized front-side-down on the carrier substrate by recording IV-curves for electrical connections between two vias, and between measurement pad and front side Al-metallization. Thus, type 1 samples were used to extract overall contact and via resistance (Figure 3.10). When the characteristic resistance of the measurement setup of 1.5 Ω was subtracted, via resistance was determined to be 7 Ω in the case of a single via (contacts 1—2), and 15 Ω for the configuration with two vias in series (2—2).

The purpose of the type 2 structures was solely the measurement of current leaking through the via isolation (Figure 3.11). This leakage current was determined to be in the range of 0.1 nA at a voltage of 1 V for both cases (1—2 and 2—2). The nonlinearities in the curves give raise to suspect that there is actually no true isolation, for the effect could be caused by diodes in a blocking configuration, as can be extracted from the structures equivalent circuit in Figure 3.11. When the oxide thickness is insufficient and pin-holes form, the overall characteristics is dominated by the MIS (metal—insulator—semiconductor) structures, resulting in an asymmetric diode-like iv-characteristics, as evident in the 2—2 configuration. For the leakage through a MIM (metal—insulator—metal) structure, a symmetric characteristics is expected, which follows the leakage mechanism in dielectric materials described by Pool and Frenkel [78].

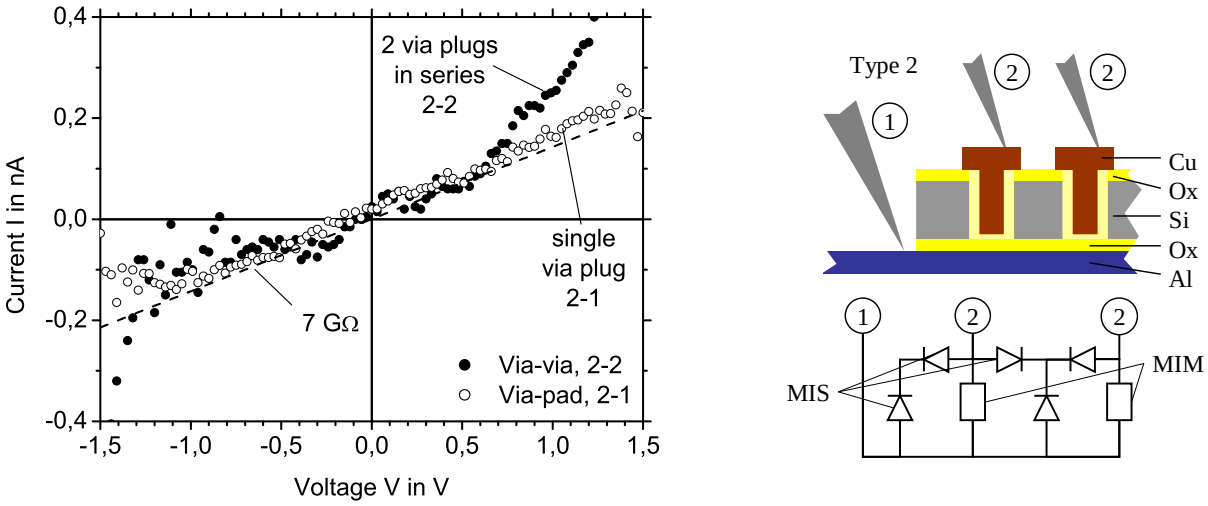


Figure 3.11: Leakage current between two vias and from via to front metal. The nonlinearities could be caused by MIS structures in the leaking path

3.4 Conclusion

A combined Cu electroplating process was evaluated, which serves to simultaneously fill high aspect ratio via features with conducting vertical sidewalls, and to fill thick resist patterns with vertical isolating sidewalls. The method was tailored to produce both, through-chip interconnects and a surface metal pattern, that could be used for the bonding and stacking processes, which will be covered in the next chapter 4.

The structures consisted of a permanently bonded thinned chip on a carrier substrate and had a full area Al metallization as front side metallization, which mimicked the conditions at the via bottom and served as common ground for the measurements (see Figure 3.5 on page 47 for processing details). The Cu-filled via structures exhibited a relatively high via resistance of $7\,\Omega$ for a single via. The estimated value for a geometry of $5 \times 5 \times 10\,\mu m^3$ and a specific resistivity of Cu $\varrho = 1,7\,\mu\Omega$cm, amounts to a significantly smaller value of $42.5\,m\Omega$. Klumpp et al. [22] achieved a single via resistance of ca. $4\,\Omega$, however, on a $1.4 \times 1.4\,\mu m^2$ via footprint with a W-plug.

It is assumed, that the high resistance relates either to an Al oxide layer, which has formed at the bottom of the via during exposure to ambient atmosphere, or is a result of incomplete oxide removal at the via bottom during anisotropic oxide etching. The bottom oxide etch is critical, for it can also degrade the sidewall isolation when etching time is too long. Native oxide layers, however can easily be removed by in-situ sputter cleaning with Ar-ions before depositing the plating seed layer. Such sputter cleaning should also be considered for the same reasons, when depositing the front side metal pattern for chip-to-chip bonding and establishing the electrical contact with the top metal layer of on-chip interconnects.

A further issue that needs to be addressed in on-going work is current leaking, since

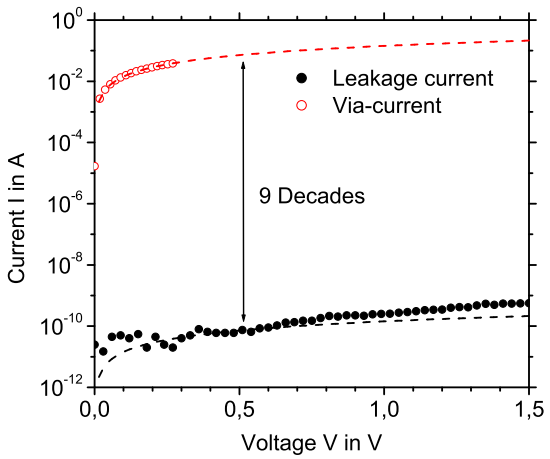


Figure 3.12: Single via contact and leakage current levels

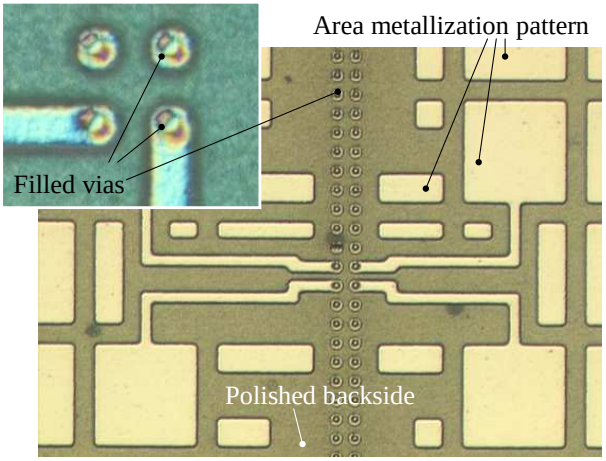


Figure 3.13: Backside plated metallization pattern with filled vias

the current results (see Figure 3.12) are not acceptable for maintaining functionality to a device related to neural processing. A method on how to minimize current leaking is selective curing of oxide defects by anodic oxidation. Studies on the capability of this method are related to other work within this research group and results have been been proven to be promising [37].

The technique developed in this chapter was also applied to temporarily bonded wafer elements, which had a patterned Cu front side metallization and were thinned using the visual endpoint detection method with a subsequent polishing step. Manufacturing of such wafer elements has been described in chapter 2 already. Figure 3.13 shows the backside metallization pattern of a test vehicle (see figures A.2 and A.3) with filled through-chip via-structures on a thinned temporarily bonded wafer element. In chapter 4 we will develop the means bond and stack these structures on a chip-scale.

Chapter 4

Bonding and chip-stacking

Establishing reliable electrical inter-chip connections for thinned and stacked known good dies is a challenging tasks, since technology puts temperature limits on post processing. Up to now, single chips can be tested and sorted out, however, when stacked and connected to other chips, defects are fatal to the system's functionality. Chip-to-chip bonding conditions must not affect the functionality of the circuit.

In order to achieve a high density of vertical chip interconnects, feature sizes of presently available joining techniques need to be reduced drastically. In the past, this has been attempted by implementing adhesive methods (Fraunhofer [21], Tohoku University [28], Rensslaer PI [25]), direct wafer bonding methods (IBM Research Center [79]), and advanced soldering techniques (Infineon [31], Fraunhofer [22]).

An unfavourable effect when using dielectric or conductive polymer based adhesives for chip-joining, is the softening of a previously cured chip interface when approaching glass transition temperature T_g during curing of a newly added level. During softening already established connections can break up or become unreliable.

A rigid joining interface is desired, which can be formed at low temperatures (below 300 °C) and remains rigid when another level is added to the stack, or when the integrated system is mounted onto a board.

Solid liquid interdiffusion bonding (SOLID), shows potential to meet the requirements for small and temperature stable connections: The joint that is formed at the interface is stable up to 600 °C, although formed at a much lower temperature. This allows for repeated heating for subsequent bonding steps without causing damage to already formed joints in the lower layers. In this chapter, SOLID bonding is evaluated, for it seems most suitable for chip-stacking in this 3D-integration approach.

4.1 Introduction

For 3D-integration, an irreversible bonding process is needed. Potentially, this can be achieved by direct bonding techniques, where metallic or dielectric surfaces are pressed and annealed onto each other. These methods usually require a high surface flatness and are very sensitive to particle contamination. Solder based bonding techniques are more robust, but cannot be repeated, as necessary for stacking. Only limited stacking could be possible by means of melting point engineering of solder compositions, where the soldering temperature is successively lowered with each level added to the stack. Moreover, available standard solder bump based technologies cannot provide the required contact density for their "large" amount of solder needed for dimensional control in the

liquid phase.

4.1.1 Available metallic bonding techniques

In the following, available bonding processes, which result in a rigid irreversible electrically conducive connection are discussed. An overview is also given for state of the art soldering in today's advanced packaging, as reviewed Tu et al. [80].

Direct metallic bonding Metallic surfaces with a low degree of roughness are brought into contact at room temperature. This will enable intimate contact at an atomic level, allowing adhesive forces to act across the bond interface. A metallurgical sound contact is formed [81]. This bonding technique is also associated with the terms "surface activated bonding" (SAB) or "diffusion bonding" [82].

All metallic direct bonding techniques other than Au–Au bonding will require a specific surface treatment that completely removes and prevents metallic oxide layers during bonding (surface activation), as well as bonding under vacuum. In most cases this will include precleaning and bonding both under high-vacuum, however without breaking the vacuum. Metal deposition processes need to be optimized in order to produce the necessary surface morphology, if necessary by a polishing step.

Depending on hardness and applied pressure, particles and surface roughness can be tolerated to a certain extent. Among known working combinations are: Cu—Cu, Al—Cu, Al—Au, Al—Si, as well as typical high-melting/low-melting combinations like Pb—Cu, PbSn—Cu, Sn—Cu or SnAg—Cu. Bonding with activated surfaces requires an extraordinary high bonding pressure in the range of $30 - 70\,\mathrm{MPa}$ ($300 - 700\,\mathrm{bar}$) in order to deform the metal layers for intimate contact, and thus is not suitable for stacking of thinned chips.

Thermo-compression bonding Direct metallic surface bonding at elevated temperature (typically ca. $400\,^{\circ}\mathrm{C}$) can reduce the necessary bond pressure to a feasible value. Frictional agitation, i. e. ultrasonic agitation is another method to enhance bond quality at less constrictive conditions. However, for metallic bonding other than Au—Au, accurate heatable vacuum placing equipment with in situ surface cleaning capabilities will have to be provided.

Solder-bump soldering Current two-level packaging structures are realized using high-melting solders for chip-level flip-chip bonding and low-melting solders for board level assembly (see Figure 4.1). During the past 10 years, soldering in packaging technology has undergone several changes, in that peripheral inter-chip connection technologies are increasingly replaced by area arrays of solder bumps, which currently can be pitched at $200\,\mu\mathrm{m}$, allowing 2500 connects per cm^2. In the near future, pitches are expected to half and density increase to $10\,000\ \mathrm{cm}^{-1}$ [80].

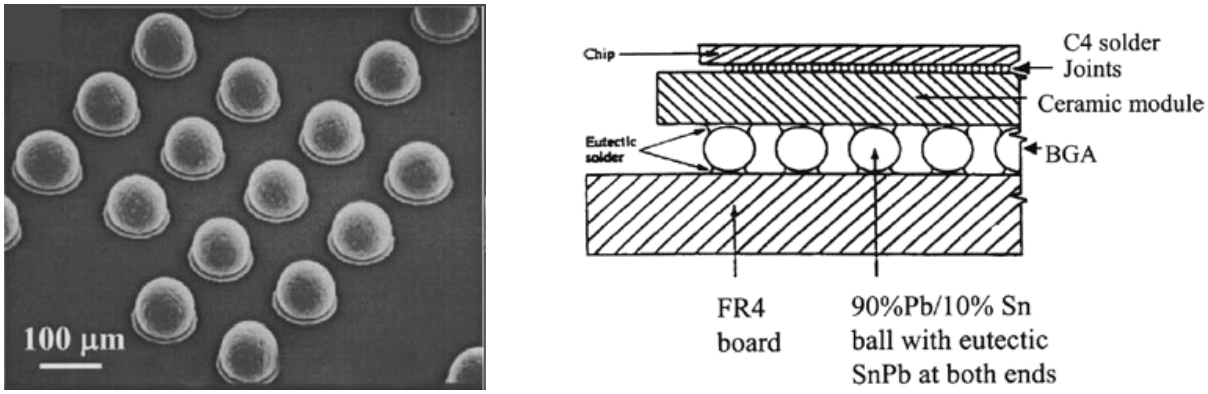


Figure 4.1: High-melting C4 (controlled-collapse-chip-connection) solder joints on the chip level and low-melting BGA (ball grid array) solder joints on the board level. After [80]

Other changes are induced by replacing Pb from soldering for environmental and health implications. Pb as an alloying component in Sn based solders is used for its ability to enhance wetting by decreasing surface tension in the liquid state. The solidus/liquidus transitions can be controlled by the Pb concentration and allow for the design of solders with liquidus temperatures spanning the range from 183 °C for an eutectic composition (38 wt-% Pb and 62 wt-% Sn) to ca. 320 °C for Pb95Sn5 (95 wt-% Pb and 5 wt-% Sn).

Pb replacements have been found for low-melting solders (eutectic AgCuSn), however, for high melting solders, the replacements have inferior wetting capabilities. Most important — as we will see in the following — Pb enhances stability of the solder/pad interface during ageing or reflowing, which is not achieved in AgCuSn solders.

Typical methods to form solder bumps are ball placement (on a packaging scale) or layer deposition techniques on a wafer scale (usually plating), followed by a melting and reshaping step, called reflow. Most delicate aspect of bump soldering is the thin layer metallization which is in direct contact with the liquid solder during soldering. Such under bump metallization (UBM) has several functions:

- Provide adhesion to the underlaying metallic or dielectric system. In most cases Al is still used as a top metallization for its compatibility to existing wire bonding technology, and its stability in ambient atmosphere. Usually the wafer has been exposed to air before the under bump metallization system is deposited. Thus, an adhesion layer (typically Cr, Ti, Ni) will be necessary to achieve good adhesion of Cu on Al. Cu on SiO_2 also calls for an adhesion inter layer.

- React with the solder and form a mechanically and thermodynamically stable interface after the solder has solidified.

- Promote the wetting of liquid solder. The wetting behaviour also depends on the solder composition.

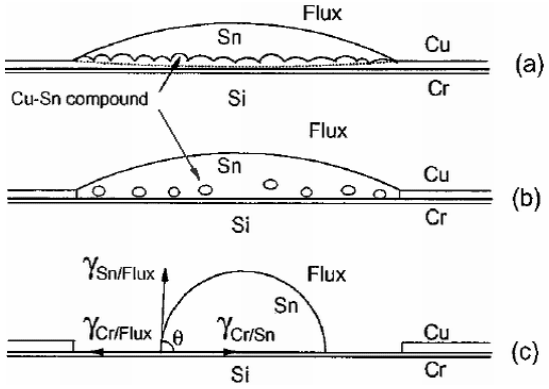


Figure 4.2: Spalling induced dewetting of solder from under bump metallization, from [83]. The Cu film is transformed into scallop shaped intermetallics (a), which detach from the interface and move into the solder volume (b). The solder contracts (c) to adjust to a higher contact angle on Cr

- The solder must not come into contact with Cu damascene wiring, since it will be dissolved and consumed in the Sn solder volume. Damascene Cu must be separated from the solder area by dielectrics (typically the polyimide top passivation layer) or a diffusion barrier layer

Generally, the formation of intermetallics at the interface is an indicator for good wetting and a stable bond, however, during reflow (melting) or aging (solid state reactions) intermetallics can destabilize the joints. *Spalling* is a common phenomenon observed among low-lead eutectic PbSn and lead-free SnAgCu solders during repeated reflowing: The interface of the thin-film under bump metallic system (UBM, typically a sequence of Al-pad/Cr/Cu-Cr/Cu/Au-capping for high-lead solders) becomes instable and dissolves into the molten solder (see Figure 4.2). This results in dewetting and loss of stability of the joint [83].

No spalling was observed after multiple reflow cycles in eutectic PbSn solder on Al/Ni(V)/Cu-UBM[1] .However, spalling in lead-free SnAgCu solders with Ni containing UBM is still a major concern, even when the Ni-layer can withstand 20 reflow cycles [80]. The phenomenon of dissolution or detaching of intermetallics from the interface into the solder volume is not limited to liquid solder, but also occurs at a lower rate during solid state aging, making the joints brittle and unstable. Large volume solder joints never reach thermodynamic stability, thus interface reactions will continue — even at room temperature — until all Cu/Ni or Sn is consumed into intermetallics.

Another shortcoming of conventional soldering techniques are electromigration issues, which here already become relevant at current densities two orders of magnitude lower than those associated with today's copper interconnects [5]. As a measure of the stability of a material against electromigration the critical product $j\Delta x$ can be used:

$$j\,\Delta x = \frac{Y\,\Delta\varepsilon\,\Omega}{Z^*\,e\,\rho}, \tag{4.1}$$

with the current density j, the conductor length Δx, the Young's modulus Y, the effective charge number Z^*, the atomic volume Ω and the resistivity ρ, the electron charge e and

[1]The V content in Ni increases the diamagnetic properties for high rate sputtering deposition of Ni

the elastic limit of the solder material ε. Thus, the smaller Young's modulus and the higher resistivity of most PbSn eutectic solders as compared to Cu result in a smaller critical product and a lower stability against electromigration.

At a feature size of 50 µm controlling the liquid solder volume associated with bumping techniques becomes challenging. Moreover ongoing interfacial reactions during aging can result in an unfavourable heterogeneous joint composition which undermine the stress relief function of such joints. Consider a simple joint model, where a hemispherical bump with diameter d is stretched into a cylinder of height h during soldering. Then, volume consistency will reveal $h = 1/3d$, that is, for a 50 µm diameter joint, only ca. 15 µm remain as a buffer for intermetallic growth from both joint interfaces. According to phase growth experiments conducted by Bader [84], in the presence of a liquid phase, CuSn intermetallics can grow at a rate of 1.25 µm in 5 s. In the micrographs shown by Tu [80], intermetallic scallops on a typical Al/Ni/Cu UBM extend ca. 5 µm into the solder volume after 5 reflow cycles. This already affects ca. 66 % of the entire solder volume, and intermetallic growth no longer can be considered an interfacial phenomenon.

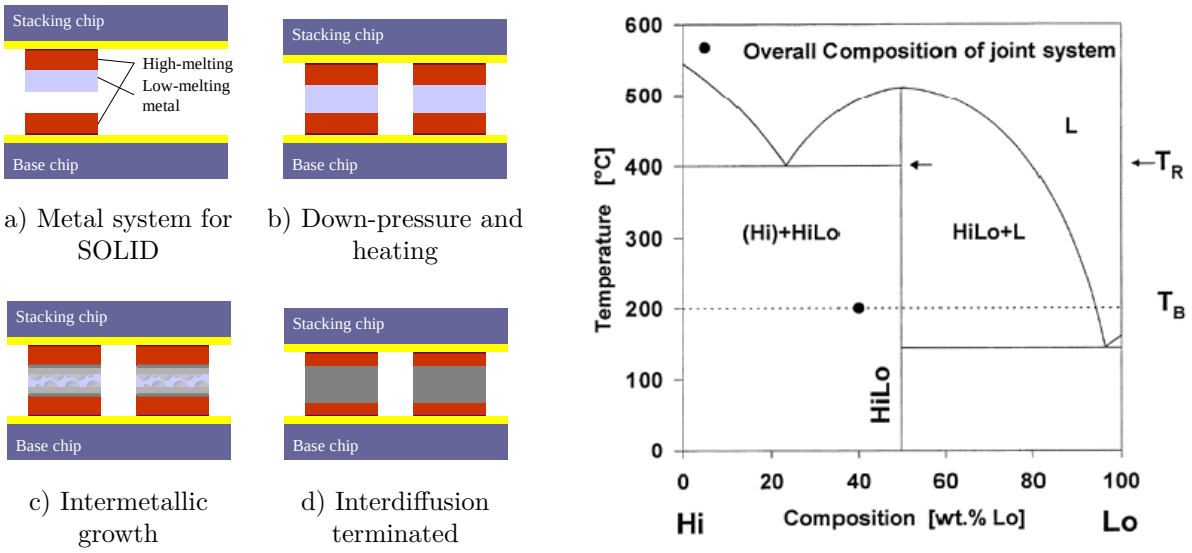
4.1.2 Solid liquid interdiffusion

A soldering technique, where the intermetallics rather than unreacted solder play the key role is solid—liquid interdiffusion. Other terms associated with this bonding technique are transient liquid phase bonding or *isothermal solidification*. Bonding with a transient liquid phase that solidifies already during the soldering process has outstandingly beneficial aspects when applied in 3D-integration.

Principle of joint formation The application of solid—liquid interdiffusion in integrated circuit fabrication was first proposed by Bernstein et al. in 1966 [85]. A low-melting component, Lo — typically pure solder metals or alloys of Sn/In/Ag/Cu binary and ternary systems — is sandwiched in between two high-melting components, Hi (Au, Cu, Ni, Fe or alloys thereof). Then the low-melting component is melted and kept at temperature until the occurring metallurgical transformations are complete and no liquid phase is present. The formed bond then has a higher melting point, which was verified by remelting experiments carried out by Bernstein.

Prerequisites apart from the difference in melting points of the components is the existence of one ore more intermetallic phases with a significantly higher melting point than that of the low melting component. One phase $\mathrm{Hi}_x\mathrm{Lo}_y$ must be the reaction product of the phase transformation processes (interdiffusion) involved. The metal system needs to be designed such, that the overall metal composition in the region affected by diffusion is $\mathrm{Hi}_x\mathrm{Lo}_y$ + Hi, thus providing a surplus of the high melting component [82].

Known suitable binary high-melting/low-melting combinations for solid—liquid interdiffusion are Au—Sn, Ni—Sn, Cu—Sn, Au—In, Ni—In, Cu—In, where the Cu/Ni—Sn system [86, 84] and the Cu—In or Cu—In:Sn systems [87] are of particular interest for their widespread use (Sn based solders on Cu and Ni based UBM) or extraordinary



a) Metal system for SOLID

b) Down-pressure and heating

c) Intermetallic growth

d) Interdiffusion terminated

Figure 4.3: Binary metal system for solid liquid interdiffusion. A surplus of high melting metal is needed in the interface region. Intermetallic phases form by interdiffusion (c) until one stable intermetallic phase Hi_xLo_y remains (c). After [82]

properties (low In melting point of 156 °C).

Superior electrical and mechanical properties The advantage of joints formed by a solid liquid interdiffusion reaction is, that a thermodynamic equilibrium is reached, which results in a long term stability. The joint volume is homogeneous in that it only consists of the high-melting compound and the intermetallic phase, which usually has similar or superior mechanical properties than the metals they are formed of (see Table 4.1). The brittleness of intermetallics associated with failure in conventional solder joints [88] is less critical here, because the more ductile pure low melting component is no longer present in the system.

Studies of intermetallic mechanical and electrical properties in the Cu/Ni—Sn system have been carried out by the National Institute of Standards and Technology (NIST) [90] by means of powder metallurgy and elsewhere by means of nanoindentation [92, 93]. As can be seen in Table 4.1 all intermetallics have a higher Young's modulus, hardness and a lower resistivity than pure Sn and Sn-based solders, which are assumed to have similar mechanical properties. The mechanical properties of the Cu_3Sn intermetallic phase are remarkably close to those of pure Cu or Ni, and the hardness values even outperform those of the pure metal. Kang [92] reports a ductility (plasticity) of Cu_3Sn close to that of Cu, which is beneficial for shear stress tolerance.

Tu et al. [5] proposed that ionic material transport through electromigration mainly depends on the resistivity and Young's modulus of the conducting material (see Equation 4.1). Thus, joints, that mainly consist of intermetallic phases can be subjected to a higher current density than comparable solder based conventional joints. The higher

Property		Cu	Sn	Ni	Cu_6Sn_5	Cu_3Sn	Ni_3Sn_4
Density	g/cm^3	8.9	7.3	8.9	8.28	8.9	8.65
Young's modulus	GPa	117	41	213	85.56	108.3	133.3
Shear modulus	GPa				50.21	42.41	45.0
Vicker's hardness[1]	kg/mm^2	30	10	15	378	343	365
Electrical resistivity	$\mu\Omega$·cm	1.7	11.5	7.8	17.5	8.93	28.5
Thermal conductivity	W/(m·K)	3.98	0.67	0.905	34.1	70.4	19.6
Thermal expansion	$10^{-6}/K$	17.1	23	12.9	16.3	19.0	13.7
Melting point [89]	°C	1083	232		415	670	

Table 4.1: Electrical mechanical and thermal properties of intermetallics in the (Cu/Ni)—Sn system [90, 91]: η-Cu_6Sn_5, ε-Cu_3Sn and Ni_3Sn_4

heat dissipation will be no threat to the joint's stability since remelt temperature is also significantly higher.

Solid—liquid interdiffusion based joints are an excellent replacement of conventional solder joints for high-density high-performance applications.

4.1.3 Choice of bonding method

In Table 4.2 on the next page advantages and disadvantages of metallic bonding techniques are summarized. Solid—liquid interdiffusion bonding is considered most suitable for 3D-integration for it results in small joint features with promising thermal and mechanical stability.

4.2 The SOLID process

The principle of solid—liquid interdiffusion (SLID) was implemented in a commercial product by Hübner et al. [31, 94] at Infineon Technologies Corporation (see also Figure 1.5 on page 6). Phase growth and interdiffusion of the Cu—Sn system has already been described by Tu et al. in 1973 [95] however, the system is complex and has been approached again by Bader [96, 86] and Bartels in 1993 [97], with a more application based focus.

The process developed at Infineon, termed SOLID, is an explicit application of the SLID principle for Cu and Sn. The aim was to develop a cost efficient chip-to-wafer process with reliable short inter-chip connections. One of the surfaces to be joined is metallized with Cu, the other with Cu and a topping Sn layer. When brought together and heated, the sandwich configuration should form a permanent joint of the quality already described.

[1]The Vicker's hardness units are derived from the weight of a proof mass divided by the indentation area

	Advantages	Disadvantages
Soldering	• Tolerates irregular joint dimensions • Easy repair work by desoldering • Very flexible and versatile process; acknowledged as standard technology	• Low thermal joint stability, $T_B > T_R$ • Wetting and spreading of fillers and subsequent excessive erosion can be a problem • Growth of thick intermetallic layers may cause embrittlement
Diffusion Bonding	• High thermal joint stability, $T_B < T_R$ • Homogeneous joints may be produced • Safer material selection with fewer constituents in joint system, no filler	• Extremely clean and flat component surfaces required • High bonding loads or frictional agitation necessary • Long homogenization period due to slow solid state diffusion • Lack of reparability
Isothermal solidification	• High thermal joint stability, $T_B < T_R$ • Tolerates some surface roughness due to transient liquid phase appearance • Low bonding loads • Short bonding times (typically minutes) • Excellent joint filling over large areas • Very thin joints ($< 10\,\mu m$) result in beneficial mechanical properties; filler preparation by thin film processes allow for good dimensional control	• Not as versatile as conventional soldering, flat and mating joint surfaces (within joint thickness) required • Limited to material combinations with favourable phase diagram and reaction kinetics • Lack of reparability

Table 4.2: Advantages and disadvantages of metallic bonding techniques as summarized by Schmid-Fetzer [82]. T_B and T_R are the bonding and remelting temperatures

4.2.1 Cu—Sn intermetallic system

In the Cu—Sn binary system two thermodynamically stable intermetallic phases exist in the temperature range from room temperature up to 350 °C. As can be deducted from the phase diagram provided by Acharya et al. [89] in Figure 4.4, compositions of these phases are at 61 wt-% Sn, which corresponds to Cu_6Sn_5 or η, and 38 wt-% Sn, which corresponds to Cu_3Sn, also termed ε-phase. A third phase, δ, with a composition of 32 wt-% of Sn, is known to be meta stable at room temperature, but can only be composed from a melt of according composition, and thus has no meaning to SOLID bonding, since the Cu always remains solid.

The Sn-rich η-phase already forms at room temperature at the Cu—Sn interface, whereas ε only forms above 50 °C at that same interface [98]. The desired Cu_3Sn-(ε)-phase with its excellent joint properties (Table 4.1) forms at the expense of η and additional Cu. Thus, ε is found at the interface to Cu, and η at the interface to Sn.

Reactions at a Cu—Sn interface heated above 50 °C are as follows (for simplicity, influences of morphology and layer microstructure, as well as temperature dependent diffusivity of the atomic species involved in the reaction, will not be taken into consideration for now): First, Sn-rich η-Cu_6Sn_5 forms at the interface, however, with availability of the η-phase, ε-Cu_3Sn can form between the η-phase and Cu. The growth of the ε-phase is slower than that of the η-phase, such that the two phases coexist as long as there is enough Sn to be transformed into the η-phase.

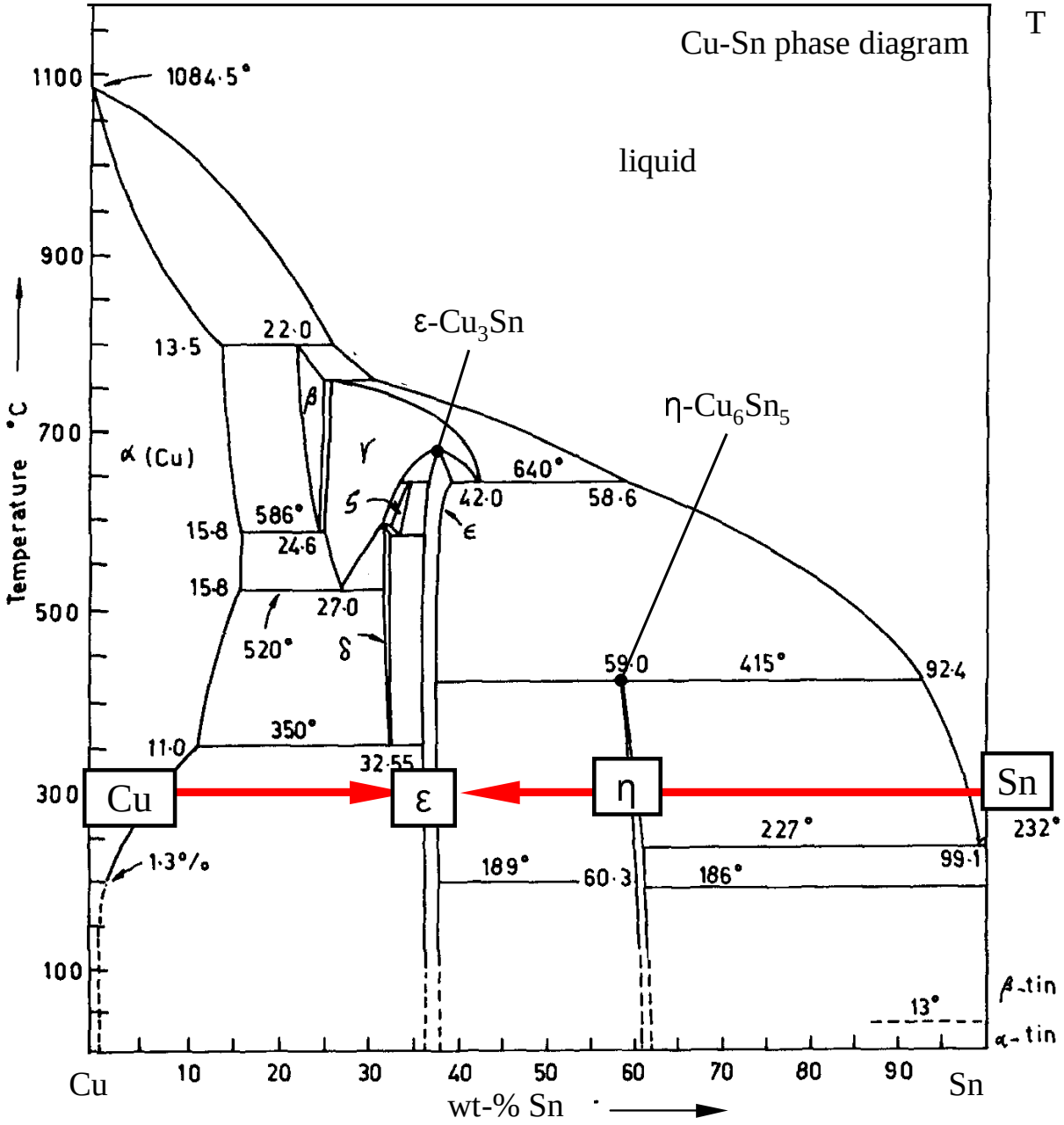


Figure 4.4: Phase diagram of the Cu—Sn system from [89]

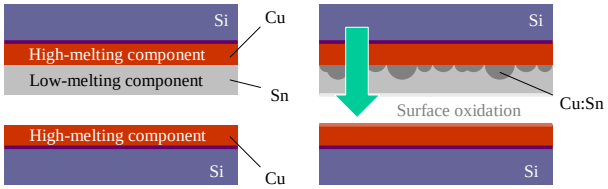


Figure 4.5: Concurring dynamic processes during heating in SOLID bonding: Surface oxidation and intermetallic phase growth

In a sandwich structure, as composed by a SOLID joint, Sn is limited and Cu can be regarded as an unlimited reservoir. Thus, during annealing, all Sn and η will be transformed into ε.

4.2.2 Dynamics of solder layers

In order to enable such reactions to form a stable joint anywhere between chip and wafer, a number of aspects need to be considered [56], as there are time dependant effects,

- Velocity of phase growth from both sides of the joint
- Velocity of oxide formation on both Cu and Sn surfaces
- Wetting inhibition and incubation time
- Reduction or removal of metal oxides
- Time constant for full area melting of the Sn layer,

and structural dependant aspects,

- Tilting of the chip induced by down force application
- Decrease in volume caused by phase reaction
- Lateral contraction and dewetting of liquid Sn on large areas
- Forces induced by the surface tension of liquid Sn.

All listed effects will depend on processing conditions, which have to be optimized to achieve a good result. They are partly concurring processes which can have an inhibiting or accelerating influence on each other. Most important is the oxide induced delayed wetting of one surface, while the other is already in contact with reactive Sn, forming intermetallics, and thus consuming liquid Sn, which in turn is necessary to wet the other surface (Figure 4.5).

During their experiments, Hübner et al. [56] developed criteria which decide wether the bonding attempts fail or succeed, as there are:

1. Metallic oxidation suppression during heating. Most critical are the more stable oxides of the less nobel metals involved. Natural room temperature formed Sn-oxides can be dissolved during melting, however, thick thermal oxide films will form on Sn during heating in atmosphere with a too high partial O_2-pressure. Mechanical agitation to break the oxide during bonding is delicate and not desirable in this case.

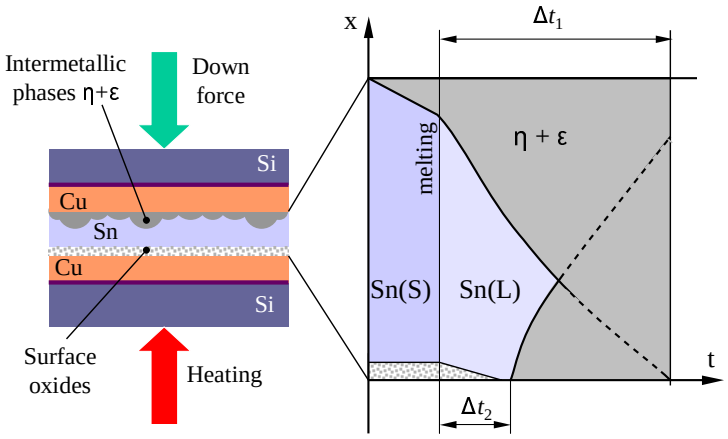


Figure 4.6: Conditions for SOLID bonding dynamics: Liquid Sn may not be consumed before the phase reaction has started from the lower surface $\Delta t_1 > \Delta t_2$

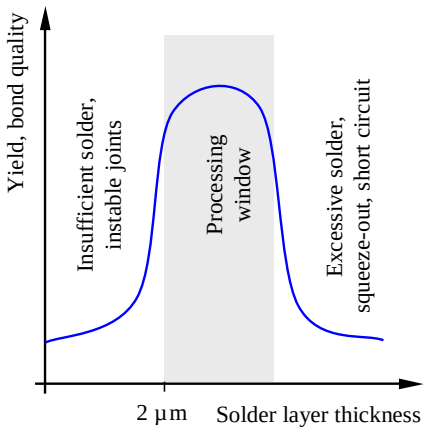


Figure 4.7: Solder layer thickness window for SOLID processing

2. Local control of wetting and phase growth kinetics. Figure 4.6 shows the joint configuration in a cross section and the solidification progress on a time scale. Phase growth will start readily from the Cu layer, onto which Sn is deposited. It will consume all reactive molten Sn after Δt_1. When the adjoining Cu interface can be wetted after $\Delta t_2 < \Delta t_1$, the phase fronts from both interfaces will meet in the middle, wich is a requirement for stable connections.

3. Global control of melting, wetting and phase growth. The local wetting dynamics described in the previous item needs to fulfill the criterion $\Delta t_2 < \Delta t_1$ on any position within the interface area at the same time. On large chips the phenomenon of "melting waves" can be observed, that is, during melting, fronts of liquidisation propagate considerably slow ($1 - 10\,\mathrm{mm/s}$) across the surface. Stable and uniform bonding across the whole chip surface is associated with reaching a state of "floating", where the entire chip floats on a layer of liquid Sn.

4. Control of solder volume to minimize liquid solder movement and displacement

The criteria ensure wettability of the high-melting Cu surfaces with liquid solder (1), enough reactive solder for phase growth from both Cu interfaces into the liquid Sn (2), "freezing" in a tilted surface orientation (3) and avoiding of liquid solder squeeze out (4). This gives a complex parametric system with correlating effects, i. e. enhancing wettability by addition of Pb, Bi or Ag to the Sn-solder will decrease surface tension, but is unfavourable to keeping the solder within its boundaries. Thus, in the following studies, pure Sn is always considered as the base solder. External pressure needs to be applied in order to densify the joint during phase transformation.

The following process details satisfied the criteria and led to good results: Solder layer thickness $d = 2 - 3\,\mu\mathrm{m}$, active flux to reduce CuO and SnO: HCOOH vapor in N_2,

No.	Technique	Sn	Atmosphere	Temp. / Duration	Result
1)	CVD-Chamber, weight	2,5 µm	Vakuum, 10^{-4} Torr	270 ° C / 10 min	no bond
2)	CVD-Chamber, weight	2,5 µm	Vakuum, 10^{-4} Torr	300 ° C / 90 min	no bond
3)	Tube oven, clamp	1,0 µm	N_2	270 ° C / 60 min	Cu ox.
4)	Alloying oven, no force	1,0 µm	NH_4Cl-vapor	250 ° C	no bond
5)	Alloying oven, clamp	2,5 µm	NH_4Cl-vapor	270 ° C	2 % area
6)	Alloying oven, spring	2,5 µm	N_2 + HCOOH	270 °C / 3 min	good
7)	Alloying oven, spring	2,5 µm	N_2 + HCOOH	270 °C / 10 min	2 % area
9)	Alloying oven, spring	2,5 µm	N_2 + HCOOH	300 °C / 15 min	80 % area

Table 4.3: SOLID experiments with an unstructured metallic system on Si-substrates. Sn films of $\approx 2\,\mu m$ and a reactive atmosphere were necessary to achieve stable bonds. The values in % denote the ratio between bonded area and overall interface area

patterned area metallization of ca. $90 \times 90\,\mu m^2$ in regions where the electrical contact density is low. The area pattern will serve as a mechanical joint and provide stability to the small feature solder joints embedded into this area metallization. The patterning also supports fast melting, since the area is partitioned and wavelike melting inhibited. The maximum external pressure that can be applied without the solder joints to collapse amounts to 5 bar for a 2 µm thick pure Sn solder film [56]. Solder film thickness is critical as can be seen in Figure 4.7.

The process sequence settled as follows: After processing and deposition of the solder metallization, diced and tested chips are picked, dipped, aligned and stuck onto a bottom wafer with a sticking wax. An accuracy of 20 µm at a placing speed of 4 s/chip was achieved. Then the wafer was subjected to heat treatment for SOLID in a vacuum oven, applying external pressure to the chips on the wafer [31]. The wafer can then be diced and packaged conventionally.

4.3 SOLID bonding process evaluation

SOLID bonding for this 3D-integration scheme was approached similarly by first producing unstructured full area joints with plated Cu and Sn layers. Therefore, as described in section 3.1.2 on page 39, a seedlayer consisting of a TiW/Cu bilayer was ion-beam sputter-deposited onto Si-substrate, followed by plated films of ca. 6 – 8 µm Cu and ca. 4 – 8 µm Cu with ca. 1 – 2,5 µm Sn on top. The 1 µm Sn films were intended to be joined with each other, thus forming an overall solder layer of 2 µm.

4.3.1 Bonding with unstructured metal layers

Then, the substrates were diced into $10 \times 10\,mm^2$ plates, clamped together and annealed in various heating systems under varying conditions. The results of these experiments are summarized in Table 4.3. It seems that the symmetric joint structure with the 1 µm Sn films already reacted and alloyed with the Cu under-metallization before Sn could

melt and wet its partner. A good bond could be achieved with $\approx 2\,\mu m$ thick one-side Sn films, annealed in an atmosphere of nitrogen with formic acid vapor (HCOOH + N_2) at temperatures of ca. 270 °C. Thus, the results of Hübner et al. [56] could be confirmed on the available equipment.

An oxide free metal surface is crucial for wetting with Sn(l). In conventional soldering it is sufficient to locally break up an oxide film, which then can be superseded by surface tension driven expanding of the liquid solder. In SOLID, mechanical agitation is usually not available, and either active flux or thorough surface cleaning and annealing in an inert atmosphere are necessary to provide sufficient wettability. In semiconductor manufacturing, flux is required to be complectly removable for encapsulated reactive residues are hazardous to electric joints on a long term. Residue-free flux (also associated with the term "flux-free soldering") are available through reactive gas mixtures (forming gas: $5 - 10\,\%$ H_2 in N_2) or evaporable active agents, which are activated at elevated temperatures (usually at $150 - 200\,°C$).

When heated, ammonium chloride decomposes into hydrogen chloride gas, which dissolves CuO into a less disturbing compound:

$$NH_4Cl \longleftrightarrow NH_3 + HCl, \qquad \text{and} \tag{4.2}$$

$$CuO + 2\,HCl \longleftrightarrow CuCl_2 + H_2O. \tag{4.3}$$

The remaining $CuCl_2$ (greenish crystals, brownish when dehydrated) usually is superseded by the liquid solder, however can be disturbing in SOLID, where particle tolerance is limited. Another possibility is the already proposed formic acid vapor (HCOOH in N_2), where the following reactions take place [99]: For temperatures $T > 150\,°C$,

$$MeO + HCOOH \longleftrightarrow Me(COOH)_2 + H_2O, \tag{4.4}$$

and for reaction temperatures $T > 200\,°C$,

$$Me(COOH)_2 \longleftrightarrow Me + CO_2 + H_2, \quad \text{and} \tag{4.5}$$

$$H_2 + MeO \longleftrightarrow Me + H_2O. \tag{4.6}$$

Thus, formic acid vapor not only is able to transform metallic oxides into less problematic compounds, but completely to reduce it to its original metallic stage without residues. HCOOH is also known to reduce the more stable oxides of Sn.

As an experimental setup an alloying oven from Bio-RAD was used, which consisted of a small heating plate, encapsulated by a glass cylinder. A gas inlet and venting outlet provided means of atmospheric control. In the experiments listed in Table 4.3 ammonium chloride powder NH_4Cl was placed onto the surface of the heating plate during annealing. When formic acid was applied, N_2 gas flow at ca. 5 sccm was lead through a bubbler which contained HCOOH in a concentration of $98 - 100\,\%$.

Wetting capability of Sn can be enhanced by elevating temperature [56]. In order to achieve good and reliable results, annealing temperature was raised to 300 °C for all following experiments.

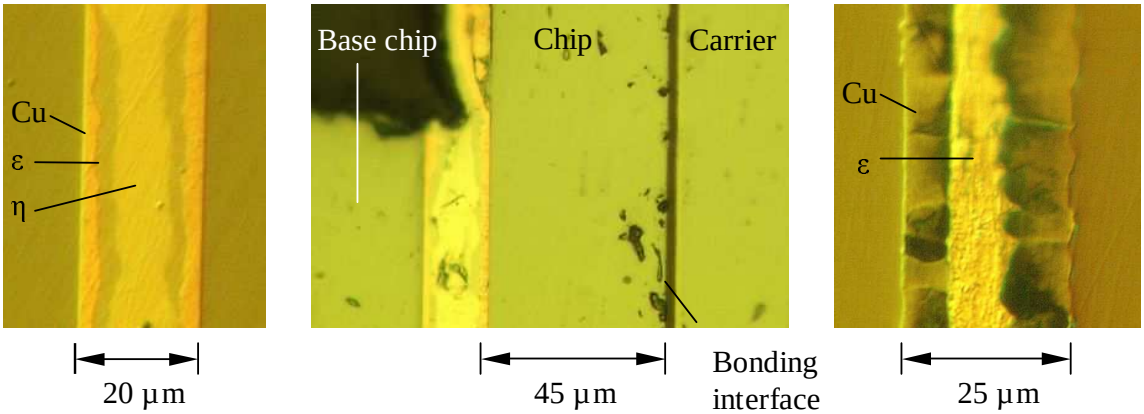


Figure 4.8: Cross section of SOLID bonded interface. Left and middle: Thinned chip on handling substrate bonded to base substrate after 3 min of annealing at 300 °C. Right: Other sample after heating close to remelt temperature

4.3.2 SOLID bonding evaluation with thinned chips

A thinned wafer element, which had already been prepared according to section 2.2.1 on page 22 by permanently bonding it to a handling substrate with BCB, and subsequently lapping and polishing it to a final thickness of ca. 45 µm, was subjected to SOLID. It was equipped with the unstructured SOLID metal system described earlier in this section, and diced into $10 \times 10\,\mathrm{mm}^2$ plates. A base substrate with a Cu top metallization was prepared and the thinned chip with its Cu/Sn surface clamped to it.

The soldered samples were prepared for a cross-sectional analysis: The plates where diced by a wafer saw and embedded into epoxy resign for grinding and polishing. Shown in Figure 4.8 is a sample with the following layer structure (from left to right): Si (base chip), TiW, Cu, Cu_3Sn, Cu_6Sn_5, Cu_3Sn, Cu, TiW, Si (thinned chip), BCB, Ti, Si (handling chip). An 800 nm Ti film on the handling substrate was intended to serve as stripping layer to detach the handling chip after SOLID bonding. As described in chapter 2, this technique was not successful. Here, focus was on the formation of a stable SOLID joint between a thinned chip and a base chip, which could be successfully demonstrated.

Figure 4.8 shows a sample, that was annealed at 300 °C for 3 min. As can be seen, not all Cu and η was transformed into ε during this time period. The amount of remaining copper suggests, that phase transformation would complete after additional annealing, leaving a homogeneous joint with only the ε-phase left in between copper layer residues.

Assuming a linear time law for phase propagation, completion of phase transformations can be estimated to occur after a doubled time period of ca. 6 min. However, Bader [84] has characterized ε-phase growth to follow a square root time law $x_{Cu_3Sn} \sim \sqrt{t}$. When Figure 4.8 is used to determine the average thickness $x_1 = 2.7\,\mu\mathrm{m}$ of the ε-layer after $t = 3\,\mathrm{min}$ of annealing at ca. 270 °C, then a growth constant k of ca. $1.6\,\mu\mathrm{m}\cdot\mathrm{min}^{-\frac{1}{2}}$ can be extracted. Thus, to close the remaining gap of $\Delta x \approx 5,8\,\mu\mathrm{m}$in between the ε-phase

fronts, a total annealing time of

$$t_{\text{heat}} \approx \frac{(x_1 + \Delta x/2)^2}{k^2} \approx 12,3 \text{ min}$$

is necessary. This is only a rough estimation, since growth of the ε-phase in the direction of the Cu-layer was not considered. For following SOLID experiments with a similar layer configuration of ca. 2.5 µm of Sn at an annealing temperature of ca. 270 °C, an overall time of ca. 15 min is considered reasonable. For a complete transformation of Cu and Sn into ε-Cu_3Sn, a layer thickness ratio of

$$\frac{d_{\text{Cu}}}{d_{\text{Sn}}} = \frac{3M_{\text{Cu}}\varrho_{\text{Sn}}}{M_{\text{Sn}}\varrho_{\text{Cu}}} = 1,31 \tag{4.7}$$

is necessary. If not enough copper can be provided, reactive Sn or η remains in the joint, and the under-metallization can be attacked by continued intermetallic reactions. This can result in delamination. The necessary amount of Cu at a specified Sn-layer thickness of 2.5 µm is provided by a minimum thickness of 3.3 µm for the copper layer. With two reactive interfaces, as in this case, the necessary Cu species is provided by two sources, and the single layers can be made thinner. However, melting and wetting non-uniformities need to be taken into account, with phase reactions starting earlier at the Cu-layer already in contact with Sn (underlayer). Then, more Cu needs to be provided by the underlayer. Also, a minimum residual Cu thickness should be considered, since this will protect the sensitive adhesion and diffusion layers from being disrupted by phase reactions. Moreover, it will serve as a buffer against mechanical stress, built up during volume change.

For the following experiments, the minimum Cu-layer thickness is set to 4 µm for both underlayer and wetting layer. Thus, the overall joint height can be expected to be in the range of ca. 10 µm, if volume shrinking and plating non-uniformities are not taken into account.

4.3.3 Bonding with patterned area metallization

In the following, the SOLID process will be evaluated on metallization patterns by face-to-face bonding of $10 \times 10\,\text{mm}^2$ chips. The patterns are designed in order to allow a first electrical joint characterization without the need for thinning of the chips. Equipment and metallic test patterns are developed in order to meet alignment and process specifications needed to form minimum sized SOLID joints. Samples for SOLID bonding were prepared by means of through mask plating as described and evaluated in section 3.1.2 on page 39, with an isolation layer of 500 nm of SiO_2.

For chip-to-chip placement, an accuracy better than 5 µm is desired in order to reach a target vertical interconnect pitch of 15 µm. As a tool that allows such placement a FineTech Pico 145 manual flip-chip bonder was available. The setup and principle of this device is shown in Figure C.5 on page 107. The device allows in-situ placing and

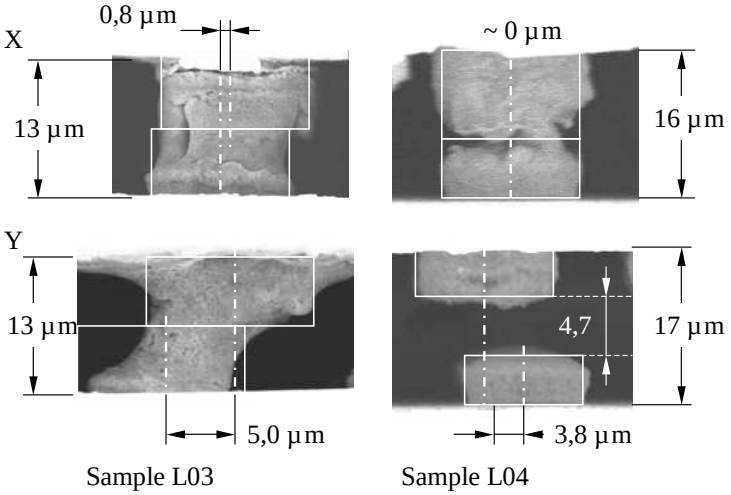


Figure 4.9: Errors associated with flip-chip bonding. Micrographs of the bonding gap, as seen from the edge of bonded chips. Left: Displacement errors in x and y directions; Right: Tilting

annealing without the need of changing equipment. In order to enable flux-less SOLID bonding under a controlled atmosphere, a glass flow box was designed and installed (see Figure C.6).

When heated, the hot plate — though made of invar steel — was subject to thermal expansion in the range of 20 µm at a temperature span of 250 °C. The bonding procedure thus was optimized for a minimum temperature difference between alignment and bonding, with alignment at 150 °C, and bonding at 300 °C. At this alignment temperature, refractive index of heated air already started to impair optical alignment, but the effect was still tolerable.

Knowing this systematic placement errors, the device was calibrated by evaluating cross sectional SEM views of bonded substrates (Figure 4.9) and adjusting the optical beam splitting prism accordingly. Calibration was carried out on $5 \times 5\,\mathrm{mm}^2$ substrates, with a stand-by temperature of 150 °C, a bonding temperature of 300 °C, and a bonding down force of 20 N. The bonding procedure which led to good results and is linked to this calibration is described in Table B.2 on page 100.

4.3.4 Results

After bonding, the samples where characterized electrically. Therefore, a Karl Suss probing station and a HP analyzer were used to contact the peripheral probing pads. These are grouped in pairs and lead to an array with a total of 100 bridge contacts in the inside area of the face-to-face bonded chips (layout in Figure A.1 on page 96).

Figure 4.10 on the next page shows, that the measured resistance generally scales with the Cu line length. The offset resistance at length $l = 0$ can be interpreted as bridge resistance, which is formed of two SOLID joints. The data was obtained from 100 measured pairs, with ca. 50 % being open or high resistivity contacts, and other ca. 10 % classified as shorts. Thus, a filtering has been applied, where resistivity values above a defect line are masked out as defective open circuits, and values below the resistivity of

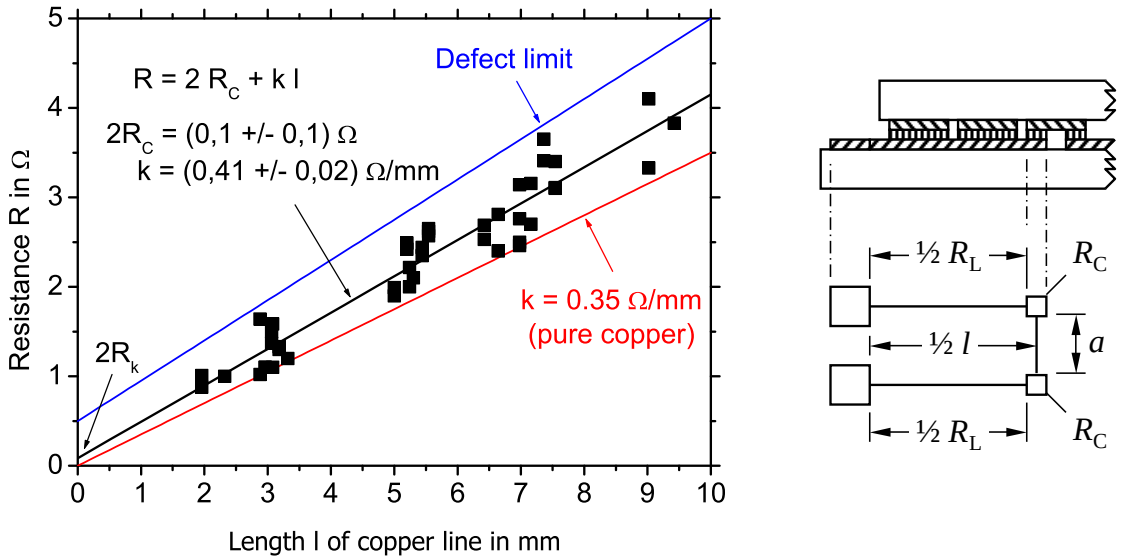


Figure 4.10: Determination of Cu-line resistivity and joint resistance. Also shown is a projection of an ideal Cu line at the given cross section

a pure Cu-line are masked out as shorts (as can be seen in Figure 4.10). The remaining values represent a good match with the model

$$R_{\mathrm{Meas}} = 2R_{\mathrm{C}} + \frac{\varrho_{\mathrm{L}}}{A} l, \tag{4.8}$$

and are used to the determine the the SOLID joint resistance R_{C} and the specific line resistivity ϱ_{L}/A at an average line cross section of $A \approx 10 \times 5\,\mu\mathrm{m}^2$:

$$\begin{aligned} R_{\mathrm{C}} &< 0.1\,\Omega, \\ k = \varrho_{\mathrm{L}}/A &= 0.41\,\Omega/\mathrm{mm}, \\ \varrho_{\mathrm{L}} &\approx 2.0\,\mu\Omega\mathrm{cm} \quad (\varrho_{\mathrm{Cu}} = 1.7\,\mu\Omega\mathrm{cm}). \end{aligned}$$

In Figure 4.11 on the next page the unfiltered data was mapped back to the joint matrix as resistance deviation from the predicted average value, which was obtained by subtracting the respective Cu-line resistance values. This reveals, that the shorts are distributed rather randomly, whereas the open circuits are limited to the top left corner with a diagonal borderline from bottom left to top right. Clearly, this is an indicator for a bonding error typically caused by chip-tilting. Obviously, the surfaces were not aligned in parallel during bonding, thus resulting in a higher gap between the contacts in the upper left corner, which could not be filled by liquid Sn. Such situation can be caused by lever effect of a contamination particle, or by non-uniform pressure application of the bonding tool.

A histogram of the measured joint resistance is shown in Figure 4.12. As can be seen, scattering around the determined average value of 0.1 Ω shows a clear peak, resembling

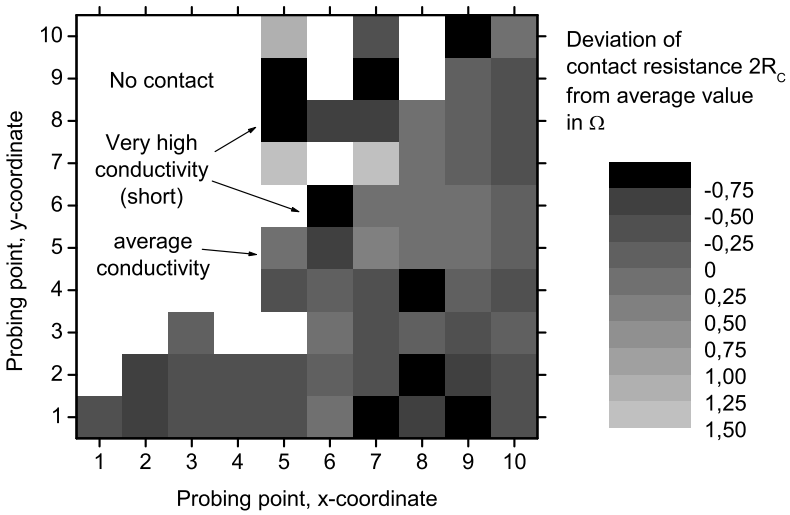


Figure 4.11: Area mapping of contact resistance. Shown is the deviation from the predicted resistance after the model $R = 2R_C + kl$ in Figure 4.10

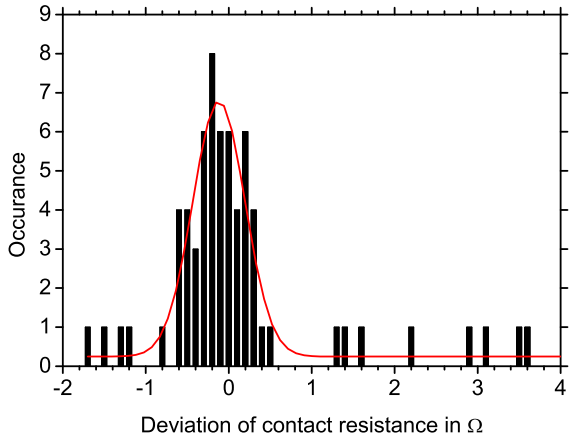


Figure 4.12: Distribution of joint resistance around the average value of 0.1 Ω

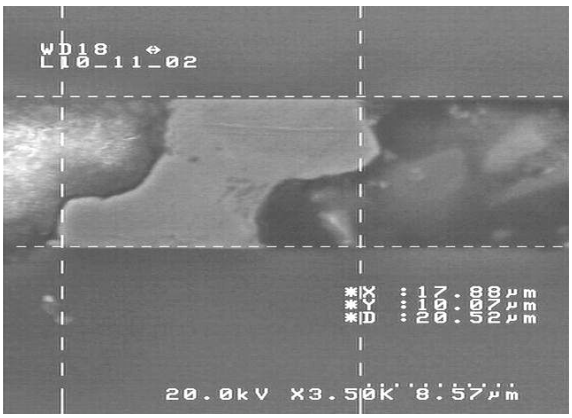


Figure 4.13: Single joint cross section with tolerable displacement

the typical bell shape of gaussian distribution, which confirms the applied line and joint model.

As can be seen, this method of joint characterization produces detailed area mapped electrical information, and can be used to optimize the area uniformity of the bonding process. However, to be applicable on a broader range, probing and measurement need to be automated.

After electrical characterization, the samples where analysed with destructive methods. Cross section cuts were prepared by sawing, embedding, grinding and polishing. Such cross section is shown in Figure 4.13. As can be seen, the joint demonstrates some degree of tolerance against displacement in vertical as well as horizontal directions.

On some samples, the top Si-chip was removed by CF_4 RIE, and the bonded metal pattern was revealed (Figure 4.14 on the facing page). Both, intact structures and structures shorted by excessive Sn where found.

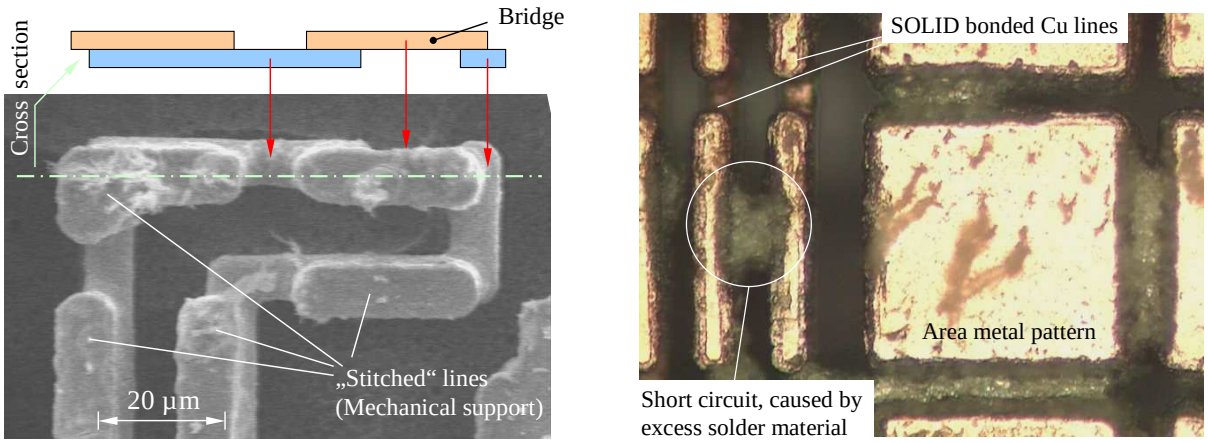


Figure 4.14: Cu-lines and bridges after removal of the upper Si-substrate. Left: Intact SOLID bridge; Right: Short circuited Cu-lines)

4.4 Chip-stacking with through-chip vias

This section will cover the stacking of multiple thinned chips with through-chip vias. All processes evaluated within this 3D-integration approach will be employed: Wafer-thinning with temporary attached handling substrates (chapter 2, via-formation and patterned electroplating for front side, via and backside metallization (chapter 3), and SOLID bonding, as implemented within this chapter.

Samples preparation spans the whole processing scheme, which now consists of 24 steps, as listed in Table B.1. Most SOLID bonding was done at a chip size of $5 \times 5\,\mathrm{mm}^2$.

Wafer element preparation for stacking Two types of samples were available for chip-stacking: First experiments were carried out on thinned and polished samples without a buried etch-stop layer, prepared as described in section 2.3.2 on page 31. For their lack of precise thickness control, via formation was challenging, and the produced stack could only be evaluated mechanically.

With the availability of samples with a buried epitaxial etch-stop layer, as proposed in section 2.3.3 on page 33, thinning was generally successfully with all epitaxial layer variations (see Table A.3), and a greater amount of stackable chips with functional vias could be produced. All vias related to these experiments were filled with CVD-oxide at a deposition temperature of 200 °C. The seed layer was deposited by ion-beam sputtering of TiW and Cu layers, followed by Cu-electrode sputtering to enhance overall conductivity and to avoid seedlayer discontinuities. A two-step process was necessary here, because TiW was not available to electrode sputtering in the desired quality. For details on processing parameters please refer to the tables shown in appendix B.4

For the characterization of stacked chips, via and contact chains over several chip-layers were designed as test patterns. These consisted of opening structures at the bottom chip level, neutral structures on thinned intermediate chip levels, that directly lead

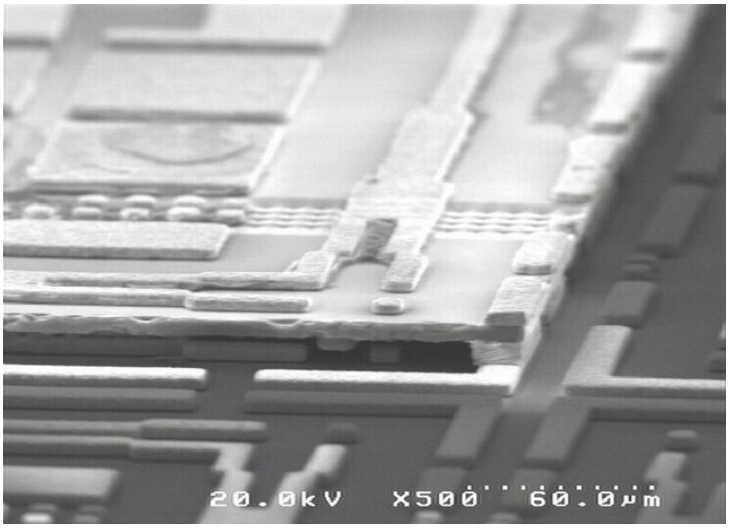


Figure 4.15: Chip stack with one stacked chip, based on thinned wafer substrate with a buried etch stop layer. The small structures are for electrical joints, large metallic patterns are for mechanical joints

through the chip without redistribution, and closing structures on the top chip level. Closing structures in an intermediate level also existed to allow characterization and documentation at any stage of stacking.

For chip-stacking, a detaching and a surface cleaning step were introduced (see also Table B.2). The bonded samples were immersed in NMP at ca. 150 °C for ca. 30 min in order to soften the PMGI release layer. The handling substrate could be removed by lateral off-sliding from the chip-stack. Then, the chip surface and front metallization was cleaned from PMGI residues by means of a O_2-plasma ashing process (Table B.7). During the plasma step oxidized Cu was subsequently removed by diluted HCOOH (formic acid). With the clean and restored front side metallization, the samples where electrically characterized, then subjected to further SOLID bonding cycles.

4.4.1 Results

A Two-level chip-stack with one thinned chip on top of an unthinned bottom chip was assembled with the introduced technique. As can be seen in Figure 4.15, after release of the handling substrate, the metallic surfaces could be cleaned and restored to its former quality, thus enabling further stacking.

The results of an electrical characterization of such structures are shown in Figure 4.16. Note, that in these measurements, the resistance of the measurement setup R_{Cal} and the resistance R_{L} of the Cu-lines are still included. Subtracting these known resistances, the series resistance of 2 vias was determined to be $2R_{\mathrm{Via}} = 1.4\,\Omega$, and the top-bottom resistance through 1 via and 1 solder joint $R_{\mathrm{Via}} + R_{\mathrm{C}}$ was $0.7\,\Omega$. The features where as follows: pad size ca. $9 \times 9\,\mu\mathrm{m}^2$, pitch ca. $15\,\mu\mathrm{m}$, line width ca. $9\,\mu\mathrm{m}$, Cu layer height ca. $4\,\mu\mathrm{m}$.

As revealed by the measurement shown in Figure 4.17, the isolation between two neighbouring contact pads is poor, and the IV-curve even demonstrates a diode characteristic. This indicates, that there is an unwanted metal—semiconductor transition in

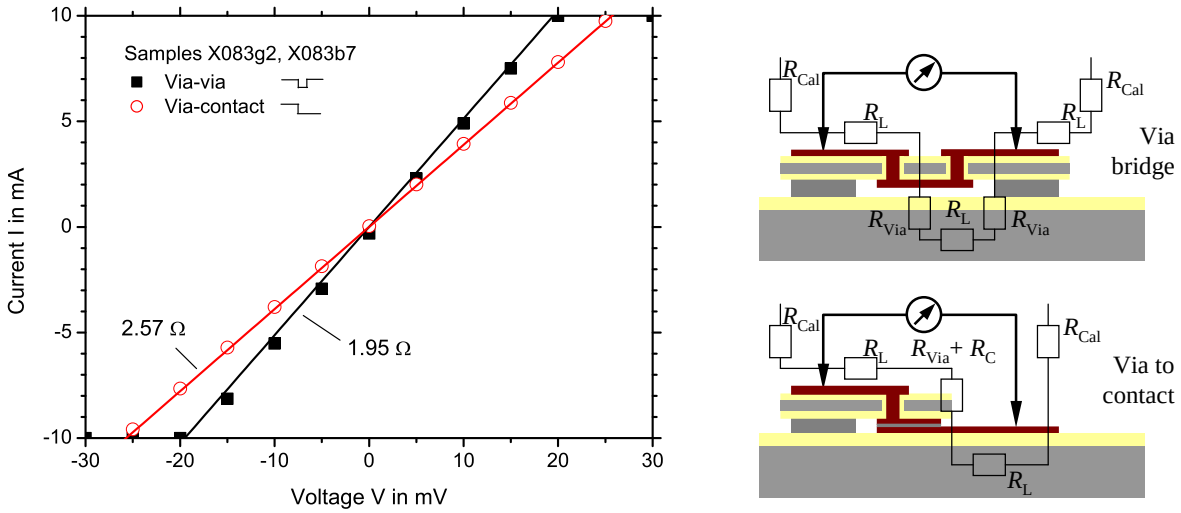


Figure 4.16: Electrical characterization of chip-stack interconnects. Shown are the IV-curves of a via-"bridge" and of a top-bottom connection

the current path giving evidence of a short with the substrate.

Stacking has been accomplished with up to 7 levels with 6 thinned $5 \times 5\,\mathrm{mm}^2$ chips (Figure 4.18). The higher-level chip stacks (4 and 7 levels) could not be characterized electrically at this stage, however the mechanical integrity of the stacks could be proved, in that the chips did not detach or break when the handling substrate was removed and the chip-surface was clean and prepared for the next stacking iteration.

4.5 Conclusion

A complete process chain of a novel 3D integration approach was implemented and evaluated for a passive Cu interconnect network on epitaxial Si-wafers. Thinned and temporarily bonded wafer-elements of $45 \times 45\,\mathrm{mm}^2$ with through-vias and a front and backside metallic interface could successfully be stacked up to the seventh chip. Table B.1 on page 99 lists the 24 process steps necessary to obtain stackable dies.

Electrical connectivity and isolation was characterized on test patterns, that included a maximum of 2 vias and one inter-chip electrical joint. Whereas the electrical conductivity of a through-chip interconnect path was less than $0.7\,\Omega$, isolation resistance between two pads with supposedly no electrical connection was less than $57\,\mathrm{k}\Omega$, and thus not tolerable. These shortcomings are associated with the challenge of high-AR deposition and etching processes inside the via-hole and are related to the backside processing nature of this approach, where thinning is done first (thinning—via—bonding/stacking) and cannot be employed for via-exposing and contact opening, as can be done in a via—thinning—bonding sequence. Therefore, the via contact has to be opened from "inside" by etching the oxide "through" the via using it as a mask. This is a critical step that can attack the sidewall isolation and result in leaking paths through the substrate.

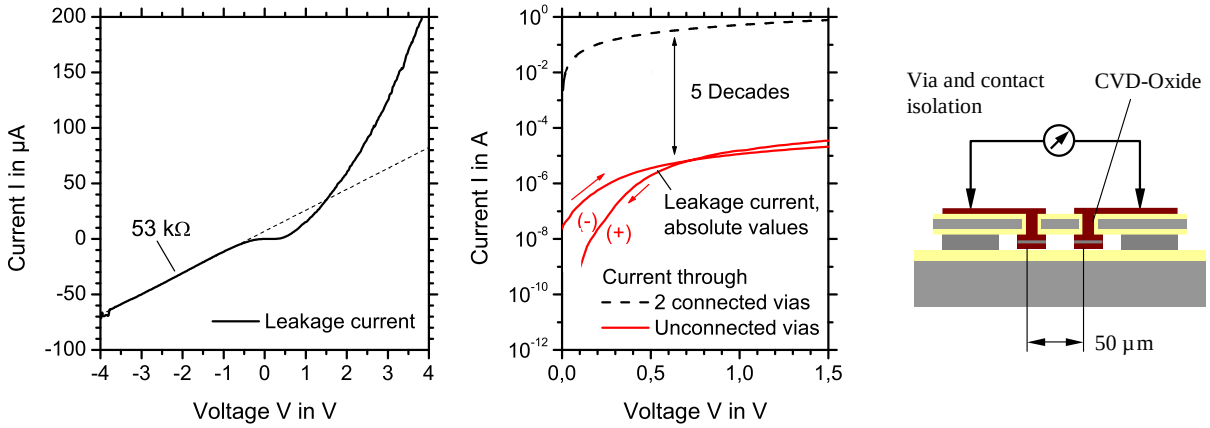


Figure 4.17: Isolation properties of 3D-contacts on stacked chips. Shown are the IV-curves of neighbouring pads at at distance of ca. 50 µm

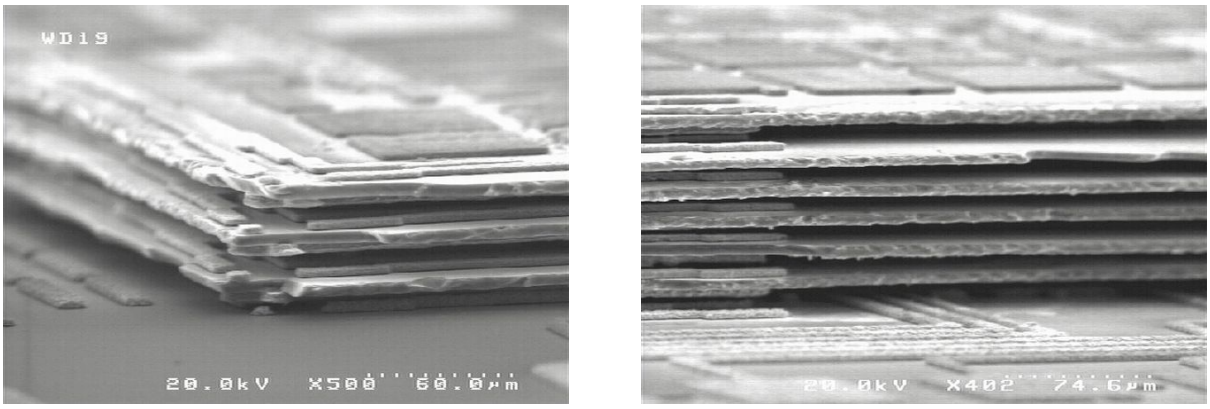


Figure 4.18: 4-layer and 7-layer stacks of thinned chips. One layer is included in the unthinned base chip

A potential method to solve this shortcomings is anodic oxidation. It can use these conductive paths to/from the substrate in order to locally anodize and recover the via isolation, thus closing the leakage paths. Studies on this subject have been undertaken, but are not within the scope of this thesis [37].

Other challenges have been identified among the SOLID bonding technique, where reliable joints could only be provided locally, but not over a typical chip area of $10 \times 10\,\mathrm{mm}^2$. The next chapter will introduce solutions that address excess solder volume and joint scalability in order to avoid shorts by solder squeeze out and to provide an even higher microjoint density.

Chapter 5

Scaling of microjoints

With today's high-aspect ratio silicon etching (HARSE) processes a high through-chip via density can be achieved [100], with feature sizes of down to below 1 µm at a pitch of 2 µm on thinned chip substrate being possible in the near future. In order to make use of this scaling potential in 3D-integration, the microjoints need to be scaled down as well.

As discussed already in chapter 4, the complex dynamic mechanisms during SOLID joining propose a scaling limit to the minimum thickness of Sn solder metallization. As for now, there are generally two solder based joining methods at the lower end of scaling, which can serve the dimensional range of 1000 µm down to 10 µm:

- State of the art *high solder volume* soldering with solder bumps created by reflowing of generally plated solder layer on top of Cu or Ni based under-bump metallization, with features of typically 100 µm, but generally not smaller than 50 µm, and
- SOLID *low solder volume* bonding (also referred to as transient liquid bonding or isothermal solidification), capable of joint features down to 10 µm (see chapter 4.

The topic of this chapter will be the proposal of methods to optimize SOLID based microjoints in terms of solder volume reduction and kinetic control in order to serve the dimensional range from 10 µm down to 1 µm.

5.1 Introduction

First, we will analyze common failure modes, observed within previous joining experiments described in chapter 4, and propose methods to enhance the reliability of microjoints at the 10 µm scale. As for down-scaling, a deeper understanding of phase transformation kinetics during SOLID bonding is necessary, if joint geometries are to be further modified. This section will thus provide the basis to view the process of joint formation from a metallurgical perspective.

5.1.1 Optimization potential in SOLID bonding

The microjoints obtained after the process described in chapter 4 had a footprint of ca. 10×10 µm, embedded into an area pattern with a maximum pattern feature size of 85×85 µm. Figure 5.1 represents this current state of chip-stacking. The metallic system typically had the following structure: 4 µm Cu / 2.5 µm Sn — 4 µm Cu. Although generally successful, the bonds occasionally showed the following shortcomings:

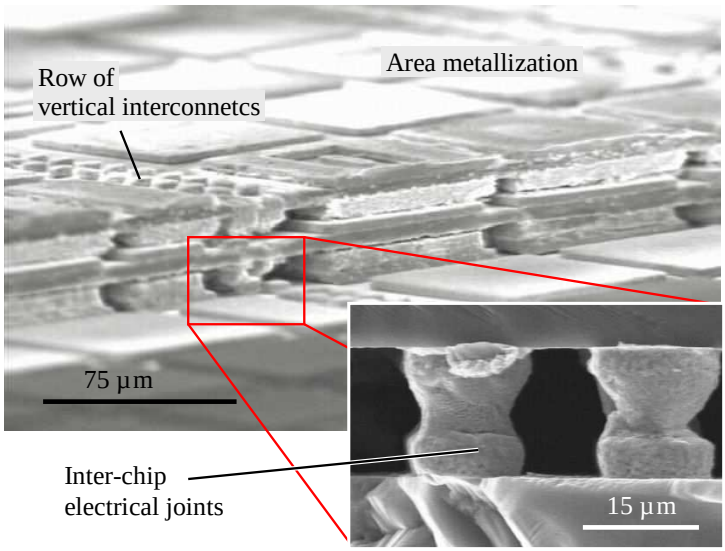
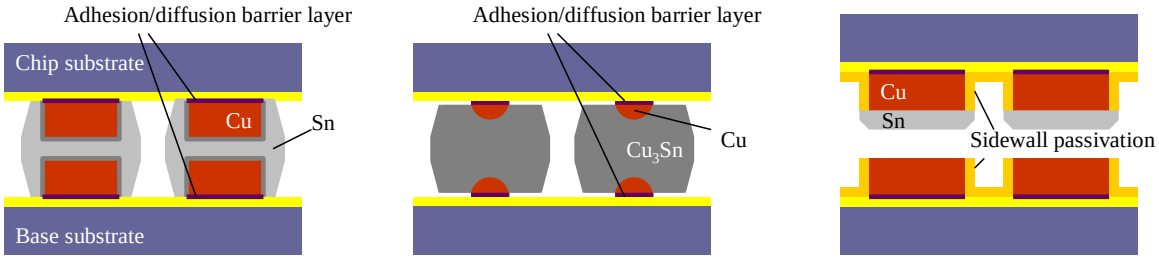


Figure 5.1: Current state of 3D-chip-stacking with 85 µm large area mechanical joints and 10 µm SOLID electrical microjoints. Two thinned and through-connected epitaxial Si-chips on top of a base chip

- Contact lift-off during seed layer etching. The footprint in comparison to the feature height was too small (high aspect ratio), resulting instability and etching undercut is considered the cause for this failure (as suggested in Figure 3.3 on page 43);
- Wetting non-uniformity and surface inhomogeneities caused spreading of solder into "forbidden" regions (see Figure 4.14 on page 73). These inhomogeneities are possibly plating artefacts, e. i. large grains or hillocks in the Cu film, or result from excessive attack during seedlayer etching (porosity);
- Pressure non-uniformity during bonding results in tilting and open connections (see Figure 4.11 on page 72) or shorts by shifting of solder volume. Applying pressure is necessary to prevent voids, but constrains chip movements to adjust itself to a planar position by surface tension;
- Wetting and spreading of solder at the sidewall of contact pads and associated intermetallic growth at the base periphery can impair adhesion of small footprint features (suggested in Figure 5.2).
- Distortion of the mask pattern during lithography. Although limited by subtracting 1 µm from mask patterns during design, and using 7 µm thick resist layers, achieving features < 10 µm was still challenging.

An improvement of surface uniformity of the SOLID bond pad can be achieved by mechanically planarizing the as-plated Cu surface. The resist remains as a filler to prevent particle contamination of the inter-pad space and to provide the base for a following lithography step. This method is quite costly while only addressing one aspect of the observed bonding failures. Solder wetting and spreading can possibly be enhanced, but not constrained effectively to the top surface of the pads.



a) Sidewalls of copper pads are wetted by Sn solder

b) Consumption of adhesion layer by phase reactions

c) Optimization: Passivation of pad sidewalls

Figure 5.2: Consumption of the adhesion layer by uncontrolled wetting of the pad surface with liquid solder. This can be avoided by protecting the sidewalls of the pads.

Sidewall wetting and alloying as depicted in Figure 5.2a can be avoided by deposition of a protective uniform coating. Complete filling of the inter-pad space should be avoided, since it represents a buffer volume that can take up excessive Sn. Any dielectric thin film can be used, however, the seedlayer needs to be removed prior to passivation, and the metal needs to be exposed again at the top surface of the pad by planarization/polishing. In order to avoid further lithography, Sn can be plated with an electroless technique.

Other suggested methods include the deposition of small copper bumps prior to Sn-deposition. These could serve as spacer and provide a well defined distance between the two joint contacts after melting of Sn. This can prevent solder squeeze out and support parallel surface alignment during SOLID bonding.

All of the suggested procedures require at least two more processing steps. Moreover, the major problem of the high aspect-ratio of the joint geometry cannot be solved with these methods, and the goal of higher contact density mentioned at the entrance of this chapter, cannot be met. The joint needs not only be scaled horizontally, but also vertically: Layer thickness needs to be reduced significantly.

Down-scaling Reducing the layer thickness of the joint metallization system is expected to have the following benefits:

- Higher mechanical stability of the joint due to larger cross section in relation to overall surface area, resulting in a small aspect ratio (height to width);
- Sidewall alloying can be neglected at small aspect-ratio features;
- Thin resist patterns can be used to structure the joint metallization. At a contact height of less than ca. 1 µm metal can be pattered by lift-off of PVD-deposited layers;
- Thin metal layers can be deposited with better control of uniformity and less artefacts, since material transport does not involve a liquid phase;

However, as described already in section 4.2 on page 61, solder volume is needed to control the kinetics of surface wetting and spreading and to buffer asymmetric phase growth (see Figure 4.6 on page 65). Thus, a scaling approach has to address at least one of these effects.

5.1.2 Prerequisites for scaling of microjoints

What are the mechanisms, that limit scaling to a minimum solder layer of 2 µm, and how can the SOLID process be modified to enable thinner layers? In section 4.2 time constants are specified and their dependance and linkage to a successful solder connection are explained qualitatively. Two principal key aspects generally determine the geometry of present joints:

1. Wetting dynamics of pads, not previously in contact with solder. This includes the process of melting and spreading of the liquid Sn over the surface. In a more static viewpoint, the behaviour of contact pairs wetted with liquid solder can also be taken into account.

2. Solder consumption by intermetallic phase reactions. This can be divided into phase growth in the solid state and phase growth at the presence of liquid Sn.

Modelling initial non-isothermal conditions We will first assume that under ideal conditions, that is: no surface oxides and a low contact angle, the time liquid Sn needs to wet the surface of a pad can be neglected. Then, phase growth in the presence of liquid Sn from pads already in contact with the Sn layer need not be considered, and the conditions for stable joint formation simplifies to providing enough meltable Sn when reaching 231 °C (see Figure 5.3). In the following it is no longer distinguished between the two phases η and ε, present in the soldering temperature range, since reactive solder is consumed by both.

Clearly, the amount of Sn consumed prior to reaching this temperature depends on the heating history, that is, the temperature over time, and the growth mechanism, which is assumed to be solid state bulk diffusion. Thus, the intermetallic thickness Δx follows a parabolic growth law $\Delta x = k_{\mathrm{sqr}}\sqrt{t}$, which can be expressed in its integral form for non-isothermal phase growth [101],

$$\Delta x \quad = \quad \left(\int_0^t k_{\mathrm{sqr}}^2\Big(T(\tau)\Big)\mathrm{d}\tau \right)^{\frac{1}{2}}, \quad \text{with} \quad k_{\mathrm{sqr}}\Big(T(t)\Big) \quad = \quad k_0 e^{-\frac{Q_0}{RT(t)}} \tag{5.1}$$

where the exponential expression describes the temperature dependance of the parabolic rate constant k_{sqr}, the function $T(t)$ describes the time dependence of temperature, i. e. in a temperature ramp, and R denotes the gas constant. The growth constant k_0 and activation energy Q_0 can be obtained from growth experiments at different temperatures, which confirm the exponential temperature dependance of k_{sqr} (see Figure 5.9).

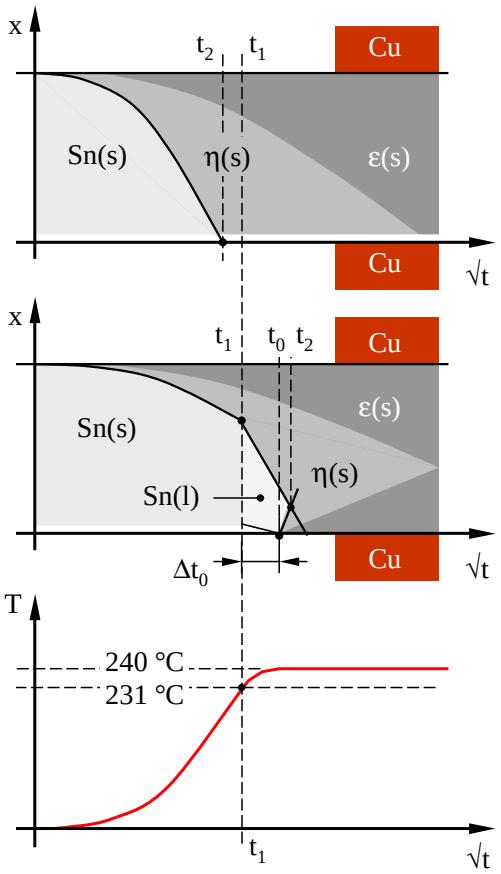


Figure 5.3: Asymmetric growth conditions of intermetallics in SOLID joints

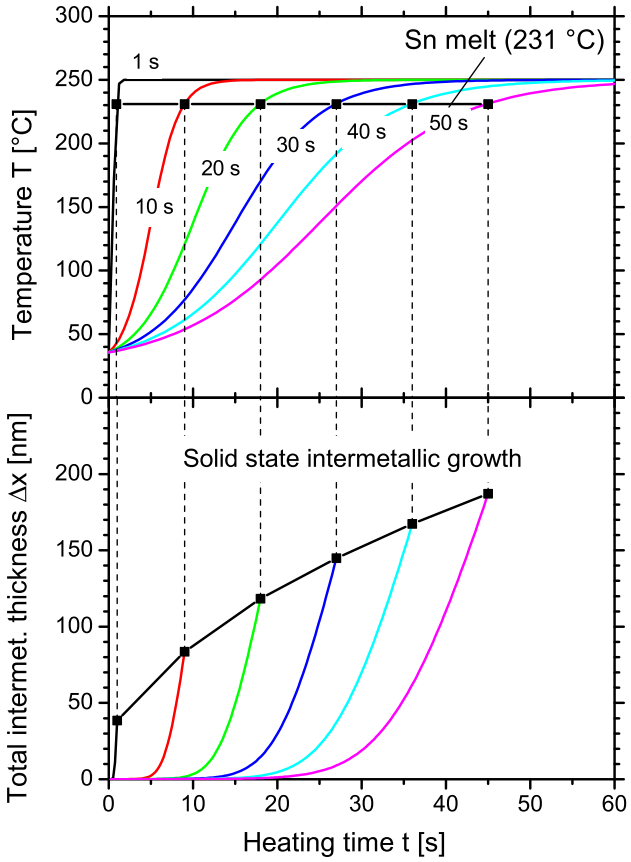
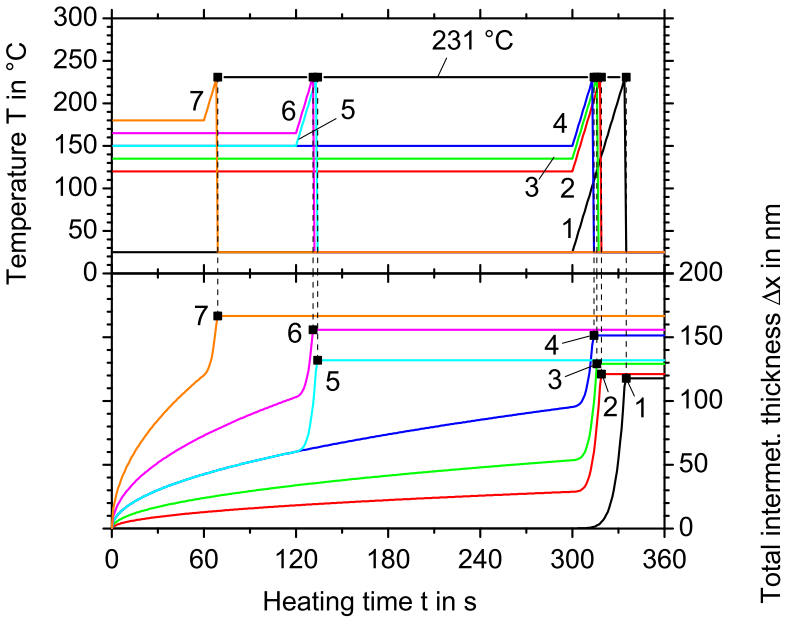


Figure 5.4: Simulated intermetallic growth at the wetted Cu/Sn-interface during heating. T_{Sn} = Melting point of pure Sn

Equation 5.1 is termed temperature integral and has no analytical solution. However, a numerical solution could be provided by "Maple" analytical math software and is shown in Figure 5.4 for a 50 s heating ramp from RT to 250 °C. Note that these curves are only valid in solid state, since growth kinetics will change dramatically in the presence of liquid Sn. As can be seen, the final thickness of intermetallics strongly depends on the heating rate, and in the case where the melting point of Sn is reached after 50 s, ca. 200 nm of intermetallics are formed during solid state diffusion reactions.

As for the SOLID process, applied in this 3D-integration scheme (section 4.2 on page 61), a pre-heating step at a standby temperature of 150 °C and a duration of typically 5 min was introduced to minimize thermal expansion induced misalignment and to activate surfacial metal oxide reduction with HCOOH [99] at a minimum temperature of 150 °C. For an analyzation of different scenarios, a linear ramp model is introduced, that generally holds 3 parameters, that is: Pre-heating time t_{pre}, pre-heating temperature or standby temperature T_1, and the linear heating rate k_{heat}. Figure 5.5 resembles solid state



SOLID process parameters

	T_1	t_{pre}	k_{heat}
1	25 °C	5 min	6 K/s
2	120 °C	5 min	6 K/s
3	135 °C	5 min	6 K/s
4	150 °C	5 min	6 K/s
5	150 °C	2 min	6 K/s
6	165 °C	2 min	6 K/s
7	180 °C	1 min	6 K/s

Figure 5.5: Initial solid state phase growth for various SOLID process scenarios. When reaching Sn melting temperature, heating is switched off in this model

phase growth associated with variations of pre-heating parameters at a typical heating rate of 6 K/s.

Kinetic control by external parameters Least solid state phase growth can be expected when heating directly from RT, as is the case in scenario 1. However, for high precision alignment accuracy, ramping temperature difference should be kept low. Scenarios 4 and 5 reveal, that a shortening of pre-heating time is significantly beneficial to avoid excessive solid state intermetallic growth during "hot alignment", whereas a pre-heating temperature above 150 °C, as in scenarios 6 and 7 accelerate growth to an extent that can not be compensated by a shorter pre-heating time.

Figure 5.6 confirms these relations, showing the intermetallic thickness at the point of liquidus transition (231 °C) in dependance of the pre-heating temperature for various heating rates. Heating rate has also a major influence on solid state intermetallic growth, but more than 10 K/s will not substantially lessen growth.

It has been reported, that phase growth in the Cu—Sn system already starts at room temperature at a linear rate k_{lin} between 3.5 nm and 6 nm per day [98]. In order to guarantee bondability for at least 30 days, an additional Sn buffer of ca. 200 nm will be necessary.

The modelled non-isothermal growth curves predict, that even under ideal wetting conditions, a high heating rate in the range of 10 K/s, and heating from RT without a "hot alignment"-plateau, a minimum of 100 nm of intermetallics will form. Assuming unlimited reservoirs of Cu and Sn, as can be approximated for the case of solid state interdiffusion during heating up, we can define a phase ratio $\xi = \Delta x_\varepsilon / \Delta x_\eta$ as thickness ratio between the cross sectional regions of η and ε intermetallic growth. Then, the total

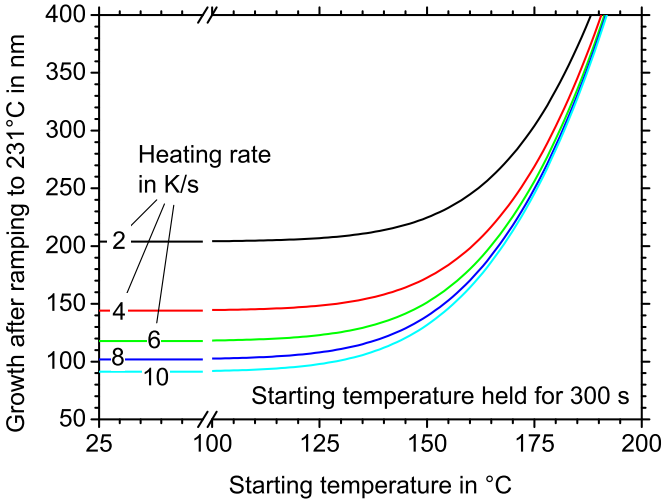


Figure 5.6: Effect of preheating temperature on total intermetallic thickness at liquidus (231 °C) after a preheating time of 300 s. Below 100 °C no significant influence is expected, and above 150 °C intermetallic thickness will consume too high an amount of reactive Sn

intermetallic thickness Δx can be partitioned as follows:

$$\Delta x = \Delta x_\eta + \Delta x_\varepsilon, \quad \text{with} \quad \Delta x_\eta = \Delta x \frac{1}{\xi + 1}, \quad \text{and} \quad \Delta x_\varepsilon = \Delta x \frac{\xi}{\xi + 1}, \tag{5.2}$$

and the amount of consumed Sn during the solid state phase reactions can be expressed as a layer thickness

$$\Delta x_{\mathrm{Sn}} = \Delta x_\eta \frac{5 M_{\mathrm{Sn}} \varrho_\eta}{\varrho_{\mathrm{Sn}} (6 M_{\mathrm{Cu}} + 5 M_{\mathrm{Sn}})} + \Delta x_\varepsilon \frac{M_{\mathrm{Sn}} \varrho_\varepsilon}{\varrho_{\mathrm{Sn}} (3 M_{\mathrm{Cu}} + M_{\mathrm{Sn}})}. \tag{5.3}$$

Using the material constants in Table 4.1 on page 61, the molar masses $M_{\mathrm{Cu}} = 63.5\,\mathrm{g/mol}$ and $M_{\mathrm{Sn}} = 118,7\,\mathrm{g/mol}$, and a phase ratio ξ of 1.0 to 0.5 (as observed in [102]), this amounts to a Δx_{Sn} of 59 nm to 62 nm for $\Delta x = 100\,\mathrm{nm}$.

However, in a SOLID bonding process, additional Sn buffering will be necessary to account for the delay time associated with wetting of the corresponding pad, not previously in contact with Sn. For scaling of microjoints by a factor of 10, down to a solder layer thickness range of 100 nm to 200 nm, external parameter optimization is insufficient. More fundamental structural changes of the solder layer system need to be considered.

5.1.3 Kinetic control with metallic separation layers

In the first scenario sketched in Figure 5.3 on page 81, the solid state phase growth has transformed all Sn into intermetallic phases before even reaching the melting point of Sn at 231 °C. In a bonding process with such conditions, no reliable joint can be formed. In the second scenario, solid state phase growth reactions at the pad with the Sn layer (the upper pad), have been retarded in order to preserve enough Sn for bonding at 250 °C. Then, during bonding, the lower pad can be wetted before solidification. A method to achieve this is the introduction a diffusion barrier layer at the interface between Cu and Sn in the SOLID joints on the backside of the thinned wafer elements. This is expected to have one of the following effects:

1. The diffusion through this separation layer is completely inhibited, such that the complete Sn reservoir is available for reaction with the opposite SOLID contact surface. The diffusion barrier itself is not involved in diffusion and phase reactions;

2. The layer is porous and limits diffusion to a small number of channels, where intermetallic seeds form at a reduced grain density, however roughness of intermetallic layers is expected to increase in this case;

3. The separation layer is involved in intermetallic phase reactions with either Cu, Sn or both. The result of these reactions ideally is a significant delay of phase growth.

The separation layer must not impair adhesion of the layers and stability as well as electrical conductivity of the formed joint. Therefore, only metallic layers are considered in the following. In order to study the effect of such separation layers, again kinetic models need to be developed.

Reaction kinetics Particle transport for growth of intermetallic phases is driven and inhibited by various mechanisms, which take direct influence on the growth rate, and can change in dominance during growth. At a constant temperature, diffusion induced volume increase can be described as follows [101]:

$$V = const \cdot x^d, \quad \text{with} \quad x = \sqrt{Dt} \quad \text{and} \quad D = D_0 \cdot e^{-\frac{Q_0}{RT}}. \tag{5.4}$$

When a phase propagates with a smooth and even front, the diffusion length x can be equaled with the thickness Δx of that phase. D is the diffusion constant, and d represents the number of dimensions, in which an intermetallic volume can expand, i. e. $d = 3$ for porous diffusion barrier layers, where seed grains do not form a closed layer during initialization, and $d = 1$ for closed films of intermetallics. For each growth mechanism, the activation energy Q_0 can take a specific value, making intermetallic growth a complex time, temperature and morphology dependant process regime. In a more general expression, intermetallic volume growth in an interval $[\tau, t]$ can be modelled as

$$V = const \cdot \Big(\int_{\tau}^{t} k \mathrm{d}t \Big)^n, \quad \text{with} \quad n = \frac{d}{m}, \quad \text{and} \quad k = k_0 \cdot e^{-\frac{Q_0}{RT}}, \tag{5.5}$$

where the growth exponent n is composed of the already mentioned dimension constant d and the constant m, that represents the growth mode: $m = 3$ for grain boundary diffusion [102], and $m = 2$ for bulk diffusion, as in Equation 5.4. k represents the temperature dependant rate constant of that growth mode, which can be expressed as an exponential term with a temperature independent factor k_0.

In bulk diffusion, particles need to be transported through the growing intermetallic phase, which, after an initial growth phase becomes the rate controlling barrier to material transport. With increasing intermetallic phase layer thickness, the transport rate through that layer decreases and the phase growth at the interfaces of the intermetallic compounds

to the pure metals is following a parabolic growth law $\Delta x^2 \sim t$, or $\Delta x \sim \sqrt{t}$, with a growth exponent of $n = d/m = 0.5$, and thus $k = D$.

In other growth modes, grain boundaries present an alternative diffusion path with a typically lower activation energy, and thus higher material transport rate as compared to bulk diffusion. Any observed intermetallic phase growth can be modelled as a combination of these growth modes and their corresponding material transport paths.

Phases of intermetallic growth Provided enough thermal energy to activate all diffusional paths, 3 different successive modes of growth can be observed on a time scale:

Initial growth: The seeds can expand in all 3 degrees of freedom ($d = 3$), and volume can expand with a growth exponent $n > 1$, until the intermetallic layer forms a closed film. In the Cu—Sn system this growth mode can already take place at RT and be terminated already by the beginning of heating;

Linear growth: The separation layer presents the rate controlling barrier, and intermetallics grow at $n \approx 1$. When transformed, penetrated or dissolved, the growth law transitions to saturated growth

Saturated growth: The growing intermetallic phase takes over the rate control and saturates with $n \approx 0,5$.

The growth of intermetallics thus reflects the growth mode and reveals information on the growth mechanisms involved. Measuring intermetallic growth over time and temperature thus is an excellent method to characterize the effect of metallic separation layers or barriers. An effective kinetic control material is expected to extend the linear growth regime over time, at a low rate k_{lin}, followed by a saturated growth, indicating dissolution or penetration of this layer.

5.2 Phase growth kinetics with metal barriers

As discussed already, the introduction of a metallic barrier layer between the Cu and Sn layers of the backside Cu/Sn-SOLID contact is proposed as a compensation for asymmetric phase growth, which is unfavourable to wetting of front side Cu-SOLID contacts. This barrier layer is expected to slow down the consumption of reactive Sn deposited directly onto it, and thus allows more time for wetting of the mating SOLID contact, or — as in the case of scaling — allows SOLID bonding with significantly thinner metal layers.

In this section, experimental data will be collected in order to evaluate a variety of material combinations for their effect on intermetallic phase growth and wettability with molten Sn. First, material selection criteria are developed and feasible combinations are chosen taking into account their binary phase diagrams with both Sn and Cu as reaction partner.

	Al	Ni	Ti	Si	Cr	Ta
Cu	limited, intermetallics	limited, intermetallics	limited, intermetallics	limited, intermetallics	none	limited, intermetallics
Sn	s: none l: low	limited, intermetallics	limited, intermetallics	s: none l: none	none	limited, intermetallics
Notes	Phase growth betw. Cu & Al	Phase growth	Phase growth	Phase growth betw. Cu & Si	eventual (Cr_2Sn_3), adhesion?	Phase growth

Table 5.1: Solubility of various metals in Cu and Sn according to their binary phase diagrams [103, 104, 105]. s = solid, l = liquid state

5.2.1 Material selection criteria

Assuming Cu to be the dominant diffusing species in the relevant temperature range, efficient barriers against diffusion of Cu need to be found. Since Cu is also the dominant species in the Cu—Si system, common damascene interconnect diffusion barriers can be considered. Typical diffusion barriers in state of the art VLSI circuit technology have been identified by Istratov et al. [61], who also established a rating of efficiency:

$$\text{TaSiN, TiSiN} > \text{TaN, } W_2N\text{, WSiN} > \text{TiN} > \text{Ta} > \text{W, TiW} > \text{Ti.}$$

Among an available subset of these materials, a subset of typical under bump metals used in conventional solder bump packaging is considered as well for the following experiments. Since the diffusion barrier material also need to provide excellent adhesion properties, it should attach well to both adjoining Cu and Sn layers and provide good wetting when Sn is molten. Poor wetting will lead to pad dewetting and eventually to solder squeeze out during bonding.

For this reason alloying behaviour is chosen as a second criteria. Materials which form intermetallics with both Cu and Sn are expected to demonstrate excellent adhesion properties, whereas metal combination that do not mix in solid state are associated with poor adhesion. Various sources of binary phase diagrams have thus been consulted [104, 103, 105, 106] for proper combinations (see Table 5.1). Nitrides of refractive metals show excellent diffusion barrier properties, but might not form intermetallics with Sn for their atomic compound stability. Alloying or combining (alternating layers) such metals with a mediating component (i. e. Ti), can improve adhesion (i. e. TiW [60]).

5.2.2 Solid state phase growth

An electroplated layer of copper was used as a base layer, which was sputter cleaned before deposition of the separation layer. The separation layer was deposited by ion beam sputtering and in-situ covered with an initial part of the tin layer of about 100 nm, without breaking the vacuum. Base pressure was in the range of 10^{-7} Torr. The sputtered Sn layer was then thickened to ca. 2 µm by electroplating, using SAT 20 plating bath.

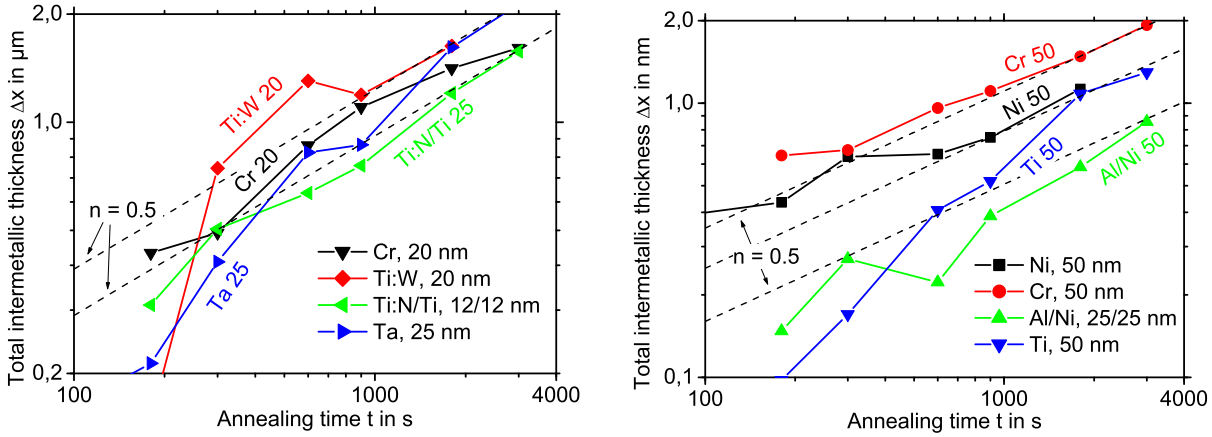


Figure 5.7: Phase growth characteristics at 200 °C for barrier layers with a thickness of ca. 20 nm and 50 nm. For long annealing time the phase growth exponents n ($\Delta x = kt^n$) approach a value of 0.5

This sample preparation sequence is described in detail in Table B.6 and was applied to produce the samples s16, s18 und s25 – s48 with TiW (30 at-% Ti), TiW/Cu, Cr, Ti, TiAl (50 at-% Ti), Al, Al/Ni, Ni, Ta, TaN, TiN/Ti, WSiN/Ti and WSi/Ti as barrier layers. In subsequent diffusion experiments, the influence on growth kinetics of the following parameters have been investigated:

- Layer composition and thickness: typical diffusion barrier materials (TaN, TiW, TiN, WSi, WSiN), and metals which are known to form intermetallics with either Cu or Sn (Al, TiAl, CrNi), or with both (Ni, Ti, Ta). The thickness was generally between 20 nm and 50 nm in order to form a homogeneous deposit;
- Temperature dependence of growth constants in the range of 125 °C to 225 °C, determined by fitting the parabolic growth law for 2 annealing periods;
- Isothermal annealing at a temperature of 200 °C for time periods typically between 30 s and 3000 s. The temperature is considered to reflect close to melt conditions without the presence of a liquid phase, which would distort the applied model of diffusional growth.

Sn was selectively etched from the samples before and after annealing, using a sodium-nitro-phenolate etching solution [84]. The remaining step height of intermetallics was then measured by means of profilometry.

5.2.3 Results

As can be seen in Figure 5.7, the growth exponents of all measured samples approach 0.5 after sufficient annealing time. For the thicker layers it generally takes longer to obtain

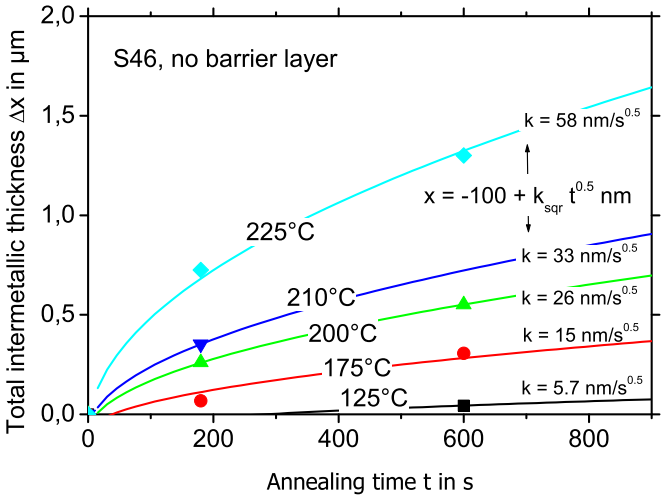


Figure 5.8: Determination of the parabolic rate constant for a temperature range from 125 °C to 225 °C using a set of samples with no diffusion barrier layer

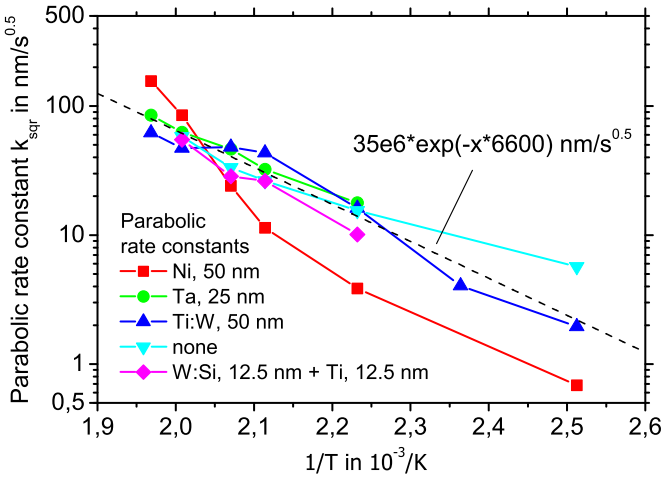


Figure 5.9: Temperature dependence of the parabolic rate constant for various diffusion barrier metals

a parabolic time law. It can be concluded, that for long annealing time, the growth mechanism is governed by bulk diffusion, and for short annealing time the characteristics of the separation layer is predominant, with a time exponent between 0.5 and 1.0, corresponding to a more reaction controlled growth mechanism. Therefore, a mixed growth model with linear and parabolic time regions is used for interpretation of the measured data.

The parabolic rate constant k_{sqr} was measured on sample sets with various diffusion barrier metals between the Cu and Sn layers in the temperature range from 125 °C to 225 °C. The negative thickness value obtained after fitting has no physical meaning, but simplifies the fitting procedure and translates into a growth delay on the time axis, thus taking into account a barrier layer induced growth delay. Figure 5.8 shows the measurements and fitting procedure for the case of no separation layer. Plotting the extracted parabolic rate constants of a choice of materials over $1/T$ reveals the Arrhenius relation

$$k_{\mathrm{sqr}} = 35 \cdot 10^6 \cdot \exp\left(-6600/T[K]\right) \quad \mathrm{nm}/\sqrt{\mathrm{s}},$$

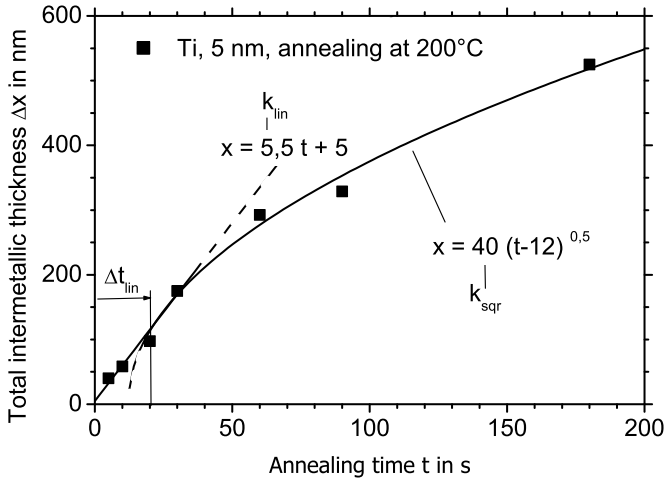


Figure 5.10: Reference sample with virtually no barrier layer at the Cu/Sn interface. The flash-layer of Ti helps to smoothen the intermetallic phase front for more accurate thickness measurement

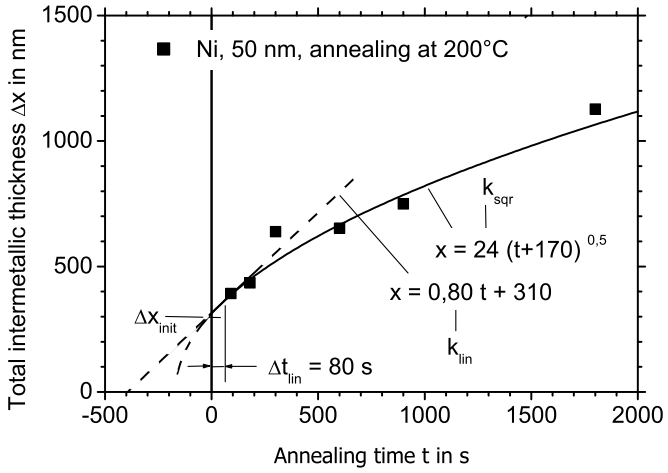


Figure 5.11: Growth characteristics with Ni interlayer. A linear region can not clearly be identified. An initial intermetallic thickness of over 300 nm indicates a high level of room temperature growth

with the activation energy $Q_0 = R \cdot 6600\,\mathrm{K} = 55\,\mathrm{kJ/mol}$ or $Q_0 = k_\mathrm{B} \cdot 6600\,\mathrm{K} = 0.57\,\mathrm{eV}$ for all evaluated samples except for the Ni barrier layer (see Figure 5.9). This means, that in the observed time period between 180 s and 900 s the chosen diffusion barrier layers have generally no influence on the growth rate, their influence is merely limited to a time period $t < 180\,\mathrm{s}$, where they induce a time delay. The previous assumption of linear growth when the barrier layer is rate controlling (see page 85) suggests, that this delay is introduced by a linear growth behaviour at time periods $t < 180\,\mathrm{s}$, where samples with no barrier layer exhibit a parabolic intermetallic growth behaviour (square root law) in that same time period.

In the following experiments, solid state interdiffusion has been characterized at a constant temperature of 200 °C. At this temperature, samples without separation layers could not be evaluated because of roughening of the interface between the intermetallics and Sn. A 5 nm thin flash layer of Ti between Cu and Sn helped to smoothen the interface for accurate reference measurement. The flash layer influences intermetallic growth only in a small time range of a few seconds, which makes it a useful reference (Figure 5.10).

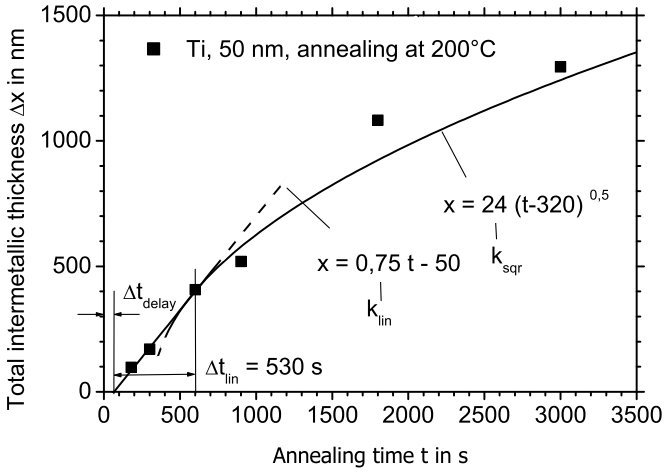


Figure 5.12: Typical growth behaviour of samples with Ti interlayers. The long linear region of 530 s indicates an intact barrier function. Growth of intermetallics at room temperature could also be suppressed

Figure 5.11 and 5.12 represent the two general types of growth behaviour observed among the samples. The 50 nm Ni layer shows an abrupt onset of growth. No evidence of a linear region, and thus no barrier function of the Ni layer was found. The $20-30$ nm layers of Cr, CrNi (50 wt-% Cr) and TiN/Ti, and the 50 nm layers of Cr and Al/Ni showed a similar behaviour. Those layers are not suitable to control the growth of intermetallics in the copper/tin system.

The 50 nm Ti barrier shows a linear region Δt_{lin} of ca. 530 s, during which the barrier function of the separation layer is effective. This type of growth behaviour has also been present in the 25 nm layer of Ta and in 20 nm of TiW (30 at-% Ti). No growth within the observed time interval of 3000 s was observed on 50 nm layers of these materials, and on 25 nm combilayers of WSi/Ti and WSiN/Ti. Thus, Ti, TiW, Ta, WSi/Ti and WSiN/Ti layers with a thickness > 20 nm can be used to control intermetallic growth.

Table 5.2 presents the results of the diffusion experiments, with the parameters fitted to the two region model. The extrapolation of the linear region to $t = 0$ marks the initial thickness Δx_{init}. This value can be interpreted as room temperature intermetallic growth, with no room temperature solder consumption at values of $\Delta x_{\mathrm{init}} = 0$. For the results with high initial thickness, as seen in the case of Ni and Cr, the presence of a linear region can not be deducted from the measurement data. The start of the linear region at $x = 0$ marks a delay time $\Delta t_{\mathrm{delay}}$ during which the temperature of 200 °C is applied, but no intermetallic growth is observed. It is assumed, that this delay is caused by phase nucleation processes. Among these samples, an order of efficiency can be established, depending on the length of the linear region Δt_{lin}:

$$\text{Ti (50 nm)} > \text{Ta (25 nm)} > \text{TiW (20 nm)}.$$

When the thickness of the barrier layers exceeds a certain value, intermetallic growth is suppressed, as can be seen in the case of Ta and TiW layers, where intermetallic growth is observed on the thin Ta (25 nm) and TiW (20 nm) layers, but not on the 50 nm layers of these metals. No intermetallic growth was also observed on the samples with WSi/Ti and WSiN/Ti layers, which had an overall thickness of 25 nm. Considering

Diffusion Barrier	Δx_{init} in nm	k_{lin} in nm/s	k_{sqr} in nm/s$^{0.5}$	Δt_{delay} in s	Δt_{lin} in s
Cr, 50 nm	390	–	34	–	–
Cr, 20 nm	220	–	32	–	–
CrNi, 30 nm	320	–	44	–	–
Ti, 5 nm	5	5.50	40	–	20
Ni, 50 nm	310	–	24	–	–
Al/Ni, 25/25 nm	80	–	16	–	–
Ti, 50 nm	–	0.75	24	67	530
TiW, 20 nm	–	2.50	40	40	130
TiN/Ti, 13/13 nm	120	–	29	–	–
Ta, 25 nm	0	1.40	39	0	400
No growth:	Ta, 50 nm; WSi/Ti, 13/13 nm; WSiN/Ti, 13/13 nm, TiW, 50 nm				
Not measurable:	CuSn, 50 nm; no diff. barrier				
No adhesion:	Al, TiAl, 50 nm; TaN, 20 nm; TiN, 20 nm				

Table 5.2: Summary of results from diffusion experiments. The values in brackets mark experiments, where no barrier function was evident, and thus no linear fitting parameters could be extracted

their thickness, these layers where most effective in suppressing intermetallic compound formation. However, for kinetic control in a solid—liquid interdiffusion (SOLID) bonding process a delay or retardation of phase growth is sufficient. Then, the thickness of the diffusion barrier layer can be adjusted to cause a specific phase reaction delay.

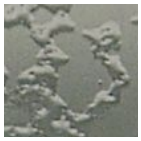
In Al containing layers, Sn did not adhere to the surface, and the films delaminated at the barrier/Sn-interface after electroplating of tin. This was also observed in TaN and TiN layers. These materials only showed good adhesion as part of a combilayer with Ti at the Sn-interface.

5.3 Wetting of barrier layers

As an essential aspect of soldering with separation layers, wettability of the layers with liquid tin is evaluated. The samples are exposed to 250 °C on a hotplate and the surface appearance is classified on a macroscopic scale (see Table 5.3).

Excellent wetting occurred on samples without barrier layers, or layers of evaporated CuSn intermetallics (composition 31:19, wt-%). On other samples, the liquid tin film did not break up, but roughened as a consequence of grain growth. Then, the surface resulted in a matte finish, which was observed on Ni barrier layers. Light dewetting, where the patterns showed a low wetting angle, was present in TiW and CrNi barrier layers, whereas strong dewetting lead to hemispherical droplets on Al/Ni, Ti, Ta and Cr layers.

A change in wettability was achieved in all cases by an isothermal pretreatment at

Surface appearance at 250 °C	 Wetting	 Light dewetting	 Strong dewetting	 Grains	 Micrograins
Without preheating	CuSn	TiW, CrNi	Al/Ni, Ti, Ta, Cr	TaN, WSi/Ti, TiN/Ti, WSiN/Ti	Ni
With preheating	CrNi TiW, Ta	—	—	Cr, WSi/Ti, WSiN/Ti	Ni, Al/Ni, Ti, TaN, TiN/Ti

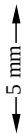


Table 5.3: Wetting behaviour after heating to 250 °C. Dewetting could be observed on all layers without preheating, except CuSn and Ni

200 °C for 5 min. Microscopic observations after selective Sn etching (sodium nitrophenole solution) suggest, that the pretreatment initiates phase nucleation of CuSn intermetallics (Cu_6Sn_5, Cu_3Sn) on the Sn side of the barrier film and promotes wetting at the spots, where intermetallics form. However, the exact composition of this grains, and the presence of compounds with the barrier layer metal could not be determined in this experiment. After the pretreatment, wetting was acceptable, and breaking up of the liquid tin film no longer occurred.

5.4 Conclusion

Scaling of solder joint dimensions in 3D-integration can be achieved by control of intermetallic phase growth kinetics in the solid state. During a typical solder heating cycle of 30 s, ca. 200 nm of solder are consumed, before reaching the liquid state. Solder volume can be saved by lowering the consumption rate with diffusion barriers, which result in an initially slower linear phase growth with time.

A two region model with linear and parabolic time dependence has been fitted to intermetallic growth data. The length of the linear region was then used as a measure of how long a metallic diffusion barrier film actively influences the kinetics of intermetallic growth. For the Cu/Sn-system it has been demonstrated, that with 20 nm layers of TiW (30 at-% Ti), 25 nm of Ta, and 50 nm of Ti, solid state intermetallic growth at 200 °C can be reduced significantly and room temperature growth can be suppressed (see Table 5.2). Growth at 200 °C is suppressed for at least 3000 s with 25 nm barrier layers of WSi/Ti and WSiN/Ti, 50 nm of Ta and 50 nm of TiW (30 at-% Ti). The results are promising to produce joints with an overall joint height $< 1\,\mu m$. The method shows potential for down-scaling of microjoints to a footprint of $1\,\mu m$.

The adhesion of tin with the separation layer metal can be promoted by deposition of combilayers with Ti. Wetting of liquid tin to the separation layer can be enhanced by pretreatment at 200 °C.

Chapter 6

Outlook

In this work, a 3D-integration processing scheme has been implemented and evaluated on a wafer-element basis, which enables through-chip and inter-chip electrical connections. The key techniques covered resemble high precision waferthinning, temporary waferbonding, via formation, and iterative chip-joining. Electrical and mechanical characterization, done at this experimental wafer-element level, already revealed the high potential of this processing scheme in terms of versatility and scalability.

A method of microjoint scaling is proposed, where a diffusion barrier layer is introduced as an under layer to slow down growth of intermetallics. This allows more time for wetting with the mating contact in case of normally scaled SOLID-joints, and significant reduction of Sn layer thickness for advanced microjoint scaling down a feature size in the range of 1 µm. Successful results have been achieved with Ti layers and combination thereof with refractive metal compositions on the Cu side.

Such technology will be the key to higher interconnect and packaging densities and will allow better communication between subsystems of complex applications. The classical 3D-applications are vision systems, for their highly parallel and straight forward interconnect architecture. However, when implemented on other technology platforms, modular combinations of systems based on different technologies will be possible, for it allows separate manufacturing of subsystems. Studies in this context already have been undertaken in the periphery of this thesis [35], with an attempt to transfer this processing scheme into CVD-diamond based UV-sensor technology and to improve via-leakage characteristics by means of anodic oxidation of Si [37], for allowing implementations of this processing scheme in even more advanced highly leakage-sensitive circuit technologies.

Its versatility makes this processing scheme a door opener to a new generation of devices, where all parts of a microsystem can be batch processed and assembled on the same close-to-wafer/chip level, without the need for modules packaging, rather than encapsulating a complete system after being assembled using these new degrees of freedom for wiring. Apart from diamond based UV-sensors, at the University of Ulm SOLID joining has also come into use for assembling of diamond based electro-mechanical microsystems, i. e. microrelays or actuated micro-bridges for chemical probing [107].

As part of a “micro-fab” concept, the key techniques described and implemented in this thesis could be employed to assemble various types of microparts and components of a microsystem. For example, components could be machined on a thinned and temporarily bonded wafer. For assembly, contacts are fabricated onto these components on a wafer level, which are then diced, picked, placed and bonded onto a microsystem. After releasing the carrier substrate the next component can be added to the system (see Figure 6.1).

Most other chip-to-wafer 3D-integration approaches are also capable of combining

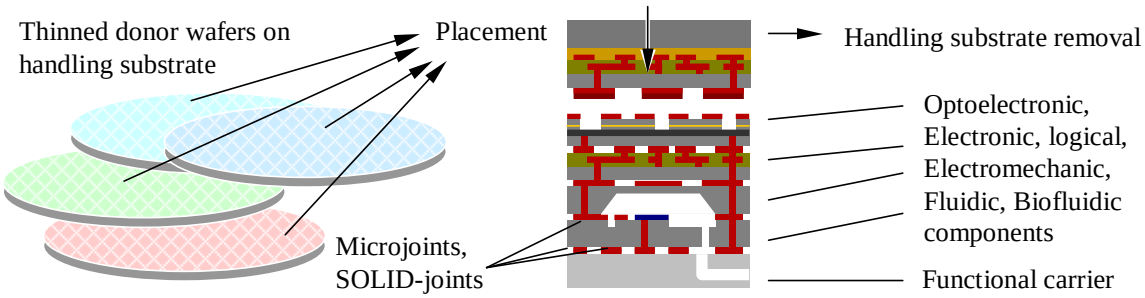


Figure 6.1: Chip-scale micro-assembly and handling concept with SOLID joining and detachable handling substrate

different technologies. However, the uniqueness of this processing scheme is its ability to produce through-chip vias without perforating the on-chip interconnect system. Since the metal stack in common VLSI circuits has a thickness in the range of the thinned chip itself, feature reduction by a factor of 2, and thus increase in area density by a factor of 4 is possible at the same via-AR (aspect ratio). This property is attributed to the novel concept of backside via-etching.

With the developed concept of microjoint scaling with Ti or Ta and combinations thereof with refractive metals as diffusion barriers, the high via densities achievable by current state high aspect-ratio silicon etching can potentially be met by chip-joining technology as well. Thus, methods exist, that allow to overcome physical scaling barriers, and 3D-integration technology may be reliably scaled down to 1 µm.

Chapter A

Experiments and samples

A.1 List of samples and processed substrates

Serie	Date range	Notes, experiment description
1—13	27.06.—11.07.2002	Bonding experiments with ceramic and polymer adhesives
P1—P3	22.07.—20.08.2002	Wafer bonding with Durimide 7510 and BCB 3022 on $5 \times 3\,\mathrm{cm}^2$ Si-substrate
Q1—Q4	22.08.—12.09.2002	CMP planarization of Si/SiO_2 surface structures
P1—P3	02.10.—06.11.2002	Lapping, polishing, electroplating, SOLID
A1—A3	12.11.—13.11.2002	Wafer bonding with BCB, Si substrate and CMOS substrate, ca. 4 $times4\,\mathrm{cm}^2$
B1, B2, C1, C2	26.11.—13.12.2002	"Testchip 0", stack with unstructured SOLID metallization
D1—D4	17.01.2003	Wafer bonding with PMGI SF11 (Single layer)
E1—E3	17.02.—26.02.2003	"Testchip 1", through mask plating with AZ 9260, with and without O_2-plasma
X001x	11.03.—31.07.2003	"Testchip 2", via-filling, via-etching and through mask plating on thinned samples. Milestone: X0014 A1–A7, B1–B7
X002x	10.04.—27.06.2003	Lithographic tests for via-etching, through mask plating samples
X003x	01.08.—25.09.2003	Thinning with epitaxial etch-stop layers
L00—L08	23.10.—12.11.2003	SOLID chip bonding with structured metal layers using the symmetric layout of testchip 2
L09—L11	14.11.—30.03.2004	"Testchip 3", SOLID chip bonding with unthinned substrate for electrical characterization
X04x	21.04.—29.04.2004	Electroplating test
X051—X058	05.05.—29.06.2004	"Testchip 4": Wafer bonding and thinning with plain substrates and surface metallization. BCB/PMGI adhesion/release layers
X060—X065	22.06.—13.08.2004	"Testchip 4": Wafer bonding and thinning with epitaxial etch-stop layer, SOLID chip bonding
X070—X076	13.05.—27.10.2004	Unthinned bottom chips with "Testchip 4" Cu bottom metallization pattern
X080—X087	30.06.—16.09.2004	Thinned chips for passive interconnect chip stacking with X07x, based on "Testchip 4"
X090—X097	15.07.—01.09.2004	"
X100—X103	27.09.—28.10.2004	"
X110, X111	27.09.—27.10.2004	Face-to-face microjoint characterization with "Testchip 3" layout
X120—X123	21.01.—04.02.2005	Thinned chips for passive interconnect chip-stack with X07x, based on "Testchip 4"
BT01—BT03	09.12.—03.02.2005	CMOS Bottom chip front metallization
S01—S44/V1	31.05.—30.11.2005	Phase growth at 200 °C with metallic diffusion barriers
S01—S44/V2,V2	05.08.—27.10.2005	Wetting with metallic diffusion barriers
T01–T17	24.10.—20.11.2005	SLID bonding / diffusion bonding with sputtered Sn layers

Table A.1: List of samples and experiments

	Testchip	Thinned	Vias	Microjoints	Detach	
C1, C2	0	⋆	—	—	(⋆)	Unpatterned SOLID metallization
X014 B1, B6	2	⋆	⋆	—	—	Through-mask plating and via-filling
L10/L11	3	—	—	⋆	—	SOLID bonding with patterned metallization
X065, X083	4	⋆	⋆	⋆	⋆	Complete process chain and characterization

Table A.2: Important samples and milestones. Detaching in testchip 0 is shown in brackets because a bulk Si etch process was used

No.	Marker depth	p^{++}-epi		p^{-}-epi	
	in µm	µm	cm^{-3}	µm	cm^{-3}
4—10	1.0	0.75	$8 \cdot 10^{19}$	9.25	10^{16}
11, 12	2.0	1.2	$5 \cdot 10^{19}$	8.5	10^{16}
13—15	1.0	1.2	$5 \cdot 10^{19}$	8.5	10^{16}
17—20	1.0	0.75	$8 \cdot 10^{19}$	14.25	10^{16}
21, 22	2.0	1.2	$5 \cdot 10^{19}$	13.5	10^{16}
23—25	1.0	1.2	$5 \cdot 10^{19}$	13.5	10^{16}

Table A.3: Available compositions of epitaxial wafer. All wafers with a marker depth of 0.5 µm were defective

A.2 Mask Layouts

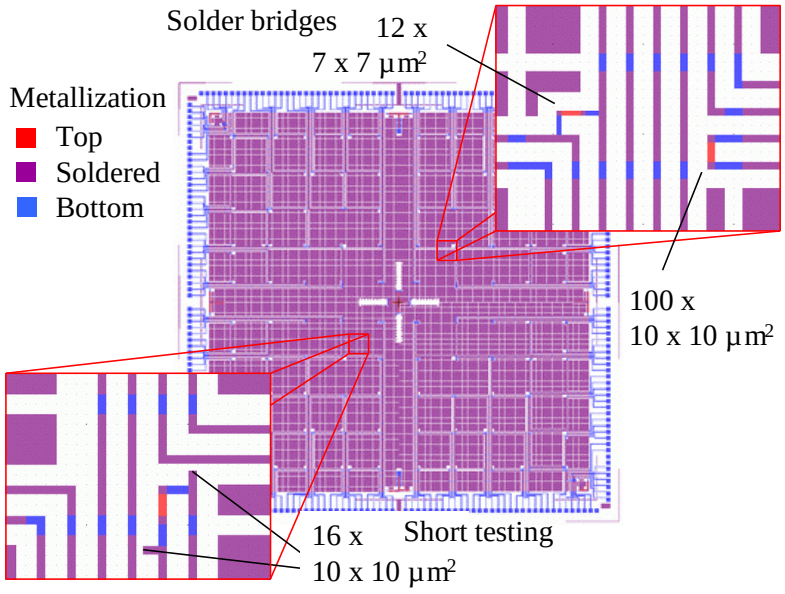


Figure A.1: Layout for face-to-face SOLID evaluation: Interconnect lines and contact bridges on a $9 \times 9\,mm^2$ chip substrate to be joined with a $10 \times 10\,mm^2$ base substrate. The chip can also be evaluated on a $5 \times 5\,mm^2$ base substrate

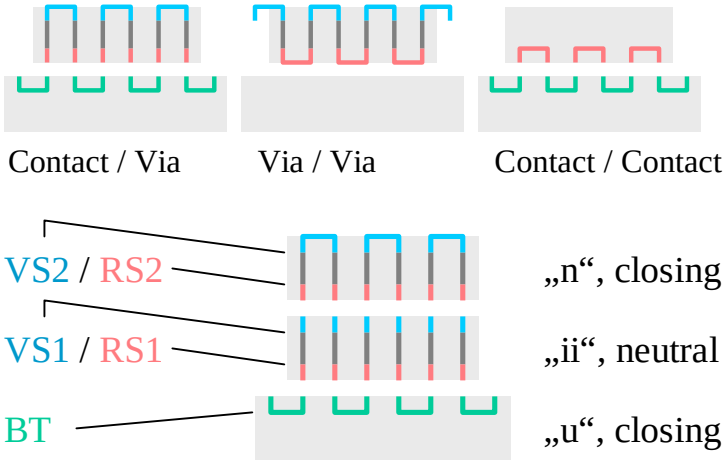


Figure A.2: Test concept for the characterization of multi chip-stacks. A chain is started in a lower level by “u”-type connections, transferred to upper levels by vertical “ii”-type connections, and closed by an “n”-type bridge

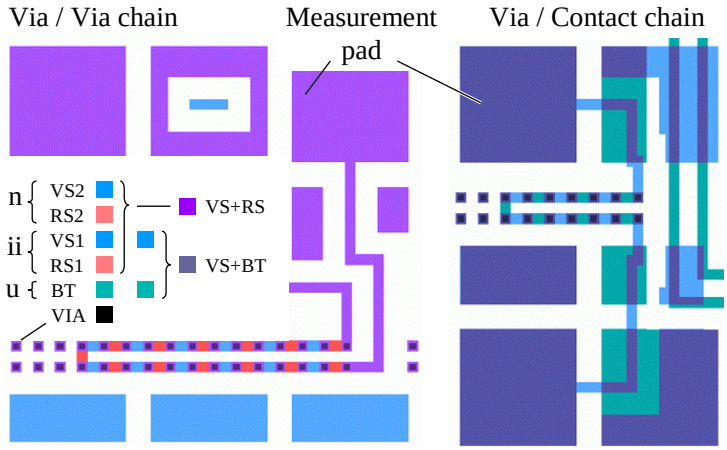


Figure A.3: Layout for test structures in multi-chip stacks

Chapter B

Processes

B.1 Process chain for wafer elements

Starting point of the process chain described in Table B.1 was wafer substrate with the following topology: Alignment marker depth 1.0 µm, 1.25 µm of epitaxy at 900 °C with $p^{++} = 5 \cdot 10^{-19}$ cm$^{-}3$, and 8.5 µm of p^{-}-epitaxy at 1060 °C. The overall epitaxial thickness was ca. 10 µm.

	Process	Details
1.	Substrate preparation	scribe & dice into $45 \times 45\,mm^2$ wafer elements / piranha clean / DI-rinse / N_2-dry
2.	FS oxide deposition	PE-CVD: 250 nm SiO_2, 300 °C, 30 W
3.	FS seedlayer PVD	Ion-beam: ca. 10 nm Ti:W, 10 min / ca. 30 nm Cu, 10 min / $I_{Beam} = 20$ mA
4.	FS lithography "VS"	AZ 9260, 3800 rpm, 60 s / 110 °C, 150 s / 21 mW·cm^{-2}, 62 s / AZ 400 K 1:4, 180 s
5.	FS electroplating Cu	O_2-plasma 100 W, 2 min / BV 30, 12 mA·cm^{-2}, 120 mA, 0.38 V, 20 min / 4.5 µm
6.	FS seedlayer removal	NH_4Cl dip / DI-rinse / NH_4OH 120 s / DI-rinse
7.	Handling substrate	scribe & dice into $50 \times 50\,mm^2$ handing elements / piranha clean / DI-rinse / N_2-dry
8.	FS adhesive coating	180°C, 5 min / PMGI 11, 6 krpm, 60 s, 150 °C, 10 min / PMGI 19, 2 krpm, 60 s 150 °C 10 min / ramp 150↑250 °C (15'), 250 °C (30'), 250↑275 °C (5'), 275 °C (60'), 275↓RT
9.	Wafer bonding	FS: AP 3000, BCB-63 ,6 krpm, 180 °C (20') / handing substrate: AP 3000, BCB-63, 4 krpm, 180 °C (20') / join: vacuum press, < 5 mbar, 180 °C (10'), 180↑250 °C (15'), 250 °C (60'), 250↓RT, vent
10.	BS Lapping	Al-Ox. 12 µm, 160' @ 1.8 kg / 20' @ 1,5 kg / 20' @ 1.2$_e$g / 15' @ 1 kg / – 625 µm
11.	BS KOH wet-etch	KOH 10 %, 500 rpm / $T_{Start} = 72$ °C, $T_{Stop} = 66$ °C, 185 min
12.	BS oxide deposition	PE-CVD: 1000 nm SiO_2, 200 °C, 30 W
13.	Lithography "VIA"	AZ 1:50 7 krpm, 180 °C, 5' / AZ 4533, 2 krpm, 100 °C, 120 s / 21 mW·cm^{-2}, 20 s / AZ 400 K 1:4, 100 s / PEB 100 °C, 20 min / O_2-plasma, 3 min
14.	BS mask patterning	RIE: 45 sccm CF_4, 40 mTorr, 400 W / 20 min, rotate, +15 min
15.	BS via-etching, ASE	Cryo-process with He(L), 100 cycles
16.	BS Surface cleaning	HF 2 %, 30 s / RIE CF_4, 200 W, 75 mTorr 5 min
17.	BS via oxide dep.	PE-CVD: 200 nm SiO_2, 200 °C, 30 W, / 50 sccm He/SiH_4 / 180 sccm N_2O
18.	BS via bottom etching	RIE: 45 sccm CF_4, 40 mTorr, 400 W / 20 min, rotate, +15 min
19.	BS seedlayer PVD	Ion-beam: ca. 10 nm Ti:W, 10 min / ca. 30 nm Cu, 10 min / $I_{Beam} = 20$ mA, alternate tilt
20.	BS lithography "RS"	AZ 9260, 3800 rpm, 60 s / 110 °C, 150 s / 21 mW·cm^{-2}, 62 s / AZ 400 K 1:4, 180 s
21.	Cu-plating, filling	O_2-plasma 100 W, 2 min / BV 30, 12 mA·cm^{-2}, 120 mA, 0.38 V, 30 min / 4.5 µm
22.	Sn-plating	Schlötter SAT 20
23.	BS seedlayer etching	NH_4Cl dip / DI-rinse / NH_4OH 120 s / DI-rinse
24.	Dicing into chips	Wafer saw, 30 µm dicing weel

Table B.1: Typical process chain for passive chip-stacks

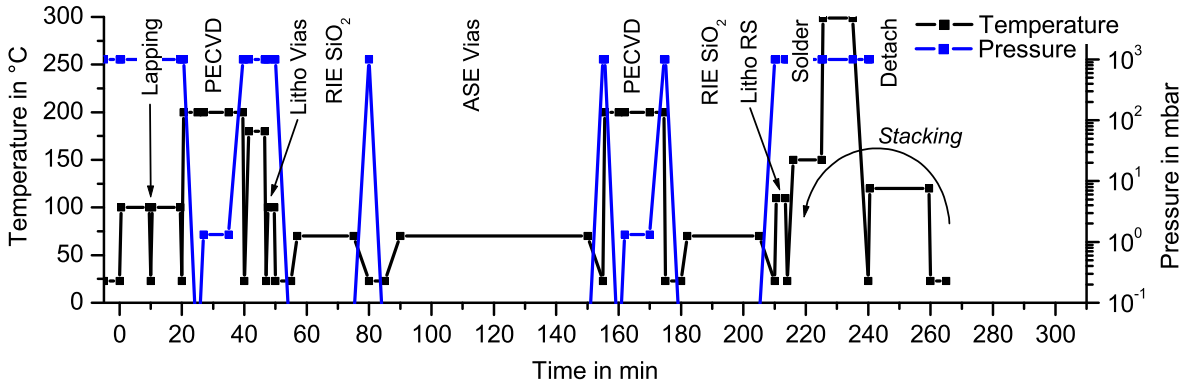


Figure B.1: Temperature and pressure profile during the process chain

B.2 Chip-stacking sequence for flip-chip-bonder

Face-2-face, face-2-back	1.	Pre-alignment of the samples: Height adjustment (z-direction) and θ (rotational) adjustment;
	2.	Heating to 150 °C, lever in lower position (contact, then separate). 2 min of pre-purging with flux (N_2 + HCOOH);
	3.	Lever in upper position, hot alignment: Fine adjustment of x- and y position at 150 °C;
	4.	Lower position, separated: 5 min of purging with flux. Then contact;
	5.	Heating to 300 °C ($\approx$ 35 s). Holding time 10 – 15 min;
	6.	Cooling down to below 100 °C with down force applied, remove sample.
Stacking	7.	Ca. 30 min NMP at ca. 150 °C to soften PMGI release layer. Laterally slide off handling substrate from chip-stack;
	8.	Clean chip surface: O_2-plasma ashing (Table B.7) from PMGI residues until particle free. Reduce oxidized Cu in diluted HCOOH
	9.	Repeat steps 1. to 8.

Table B.2: SOLID bonding procedure with “Fineplacer” flip-chip bonder

B.3 Lithography

Process	Resists	Procedure
Plating mask	AZ 9260	Spin: 4000 rpm / HP: 110 °C, 150 s / Expose: 52 s / Develop: AZ 400 K 4:1, 180 s, 10 % Ultrasonic agitation / (O_2-plasma: 100 W, 2 min) / Thickness: 6.8 µm
Via mask	AZ 4533	Adhesion promoter: AZ dil. 50:1, 7000 rpm / HP: 180 °C, 5 min / Resist: AZ 4533, 2000 rpm / HP: 100 °C, 120 s / Expose: 20 s / Develop: AZ 400 K, 100 s / O_2-plasma: 100 W, 3 min
Lift-off T0x	AZ 5214	1st layer: LOR, 6000 rpm / HP: 180 °C, 15 min / 2nd layer: AZ 5214, 8000 rpm / HP: 100 °C, 90 s / Expose: 6,5 s / Develop: AZ MIF 726, 35 s

Table B.3: Lithographic parameters

B.4 Ion-beam sputtering

Via seed layer sputtering

Shown in Figure B.2 is the ion-beam sputter configuration used. To achieve a good sidewall and step coverage in deep via-holes, the sputter angle was changed during deposition: Ca. 50 % at the standard position, 25 % at 5 °, and other 25 % at −5 °. A Kauffmann-type ion-source with a beam diameter of ca. 20 mm was used. Prior to deposition on the sample, the target was pre-sputtered for a time equivalent of ca. 10 nm minimum. The following tables B.4 and B.5 list typical process parameters for ion-beam sputtering.

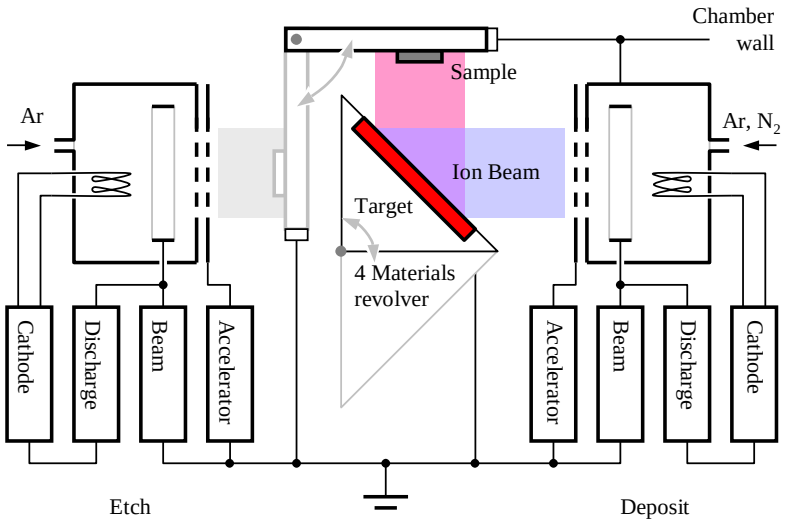


Figure B.2: Ion-beam sputter system for etching and deposition of thin metallic films, i. e. refractive metals, diffusion barriers and seed layers

	Base pressure	Processing press. (Ar)	Beam curr.	Beam voltage
Etch:	$2{\cdot}10^{-7}$ Torr	$6{\cdot}10^{-5}$ Torr	20 mA	500 V
Deposit:	$2{\cdot}10^{-7}$ Torr	$2{\cdot}10^{-5}$ Torr	20 mA	900 V

Table B.4: Parameters for ion-beam sputter etching and deposition

Process	Target (composition)	Density in g/cm^3	Z-Value	Dep. rate in nm/min	Typical duration
Seedlayer	Ti:W (30 at-% Ti)	19.30	1.000	1.13	10 min
	Cu 99.99 %	8.93	0.437	3.15	10 min
Solder underlayers	Ti:W 99.95 %	19.30	1.000	1.13	10 – 45 min
	Cr	7.20	0.305	2.75	10 – 20 min
	Ni	8.91	0.331	2.20	20 min
	Al	2.73	1.080	4.00	5 – 10 min
	Ti	4.54	0.628	1.74	15 – 30 min
	Ta 3N5	16.50		1.56	15 – 30 min
	Ta:N (+N_2)	(10.00)	(0.262)	(2.00)	10 min
	W:Si, :Si:N (+N_2)	(15.60)	(0.151)	(0.90)	15 min
	Ti:N (+N_2)	(5.20)	(1.000)	(0.80)	15 min
Solder layer	Sn	7.30	0.724	2.88	45 – 60 min

Table B.5: Sputter processes and deposition rates for various targets. The values in brackets are rough estimates. The rates and resulting thickness values vary across large sample surfaces

Diffusion barrier sputtering

The sequence described below in Table B.6 was used to prepare the samples S01—S44 and the samples T01—T17 for diffusional experiments related to microjoint scaling (chapter 5). The deposition rates for various target materials and sputter parameters are listed in Table B.4.

1. Cleaning: Ca. 2 min immersion in "piranha" etching solution. Typical sample size: $20 \times 30\,\mathrm{mm}^2$;
2. Seed layer: Ion-beam deposition (see previous section) of Ti or Ti:W (10 min) and Cu (10 min) thin films of usually 10 nm and 40 nm thickness respectively;
3. For S-type samples: Electroplating of a Cu base layer of $1.5 - 2\,\mu\mathrm{m}$ thickness, using standard plating parameters (see plating section);
4. (a) Surface cleaning: Ca. 2 min of ex-situ precleaning in diluted HCOOH, then $10 - 15$ min in-situ ion-beam sputter etching;
   (b) Barrier layer deposition: Various layers of a thickness range of $10 - 50$ nm. First 10 nm with at a sputter angle of +5 ° next 10 nm at an angle of -5 °, and remaining thickness at the standard deposition angle → avoid shadowing due to large Cu grains;
   (c) In-situ cap layer deposition: $80 - 100$ nm of Sn, T-samples: min. 100 nm → barrier layer surface protection.
5. S-type samples: Electroplating of Sn solder layer, ca. 2 textmu m.

Table B.6: Process steps for the preparation of diffusion samples

For reactive sputtering deposition of nitride layers N_2 (2.6 sccm) is mixed to the normal sputter gas Ar (1.3 sccm) and directed to the gas inlet of the ion-source for deposition. The in-situ etching system was composed of an unfocussed hot filament ion source. Switching between etching and deposition was done by activating the respective source and adjusting the sputter angle to the correct position (see Figure B.2 on page 101).

Reactive ion etching (RIE)

Process	Material(s) etched	Material(s) not etched	Gas flow in sccm CF_4	Gas flow in sccm O_2	Gas flow in sccm Ar	Pressure in mTorr	Power in W	Etch rate in nm/min	Typical duration
Bottom oxide etch	SiO_2, Si	Cu, Al	45	—	—	40	400	33	15 min
Release layer residues	Polymers	Metals	—	50	—	150	500		10 min
Polyimide passivation	Polymers	Metals	—	50	—	75	600		30 min
High rate Si etch	Si, SiO_2	Cu	45	50	25	150	1000	—	—
Resist treatment			—	8 %	—	100	100	—	2 min

Table B.7: Reactive ion etching and ashing parameters

B.5 Lapping sequence

Si removed in µm	Thickness		Down force in g	Rotation speed in rpm	Grain size in µm	Time in h	Rate in µm/min
	Total compound in µm	Wafer element in µm					
0	1380	680	—	—	—	—	—
500	880	180	1800	70	12	3:00	2.7
550	830	130	1500	60	12	3:20	2,1
600	780	80	1200	50	12	4:00	1,4
630	750	50	1000	40	12	4:30	1,0

Table B.8: Rapid lapping sequence with 12 µm Al_2O_3 grains

Chapter C

Equipment

C.1 Electroplating setup

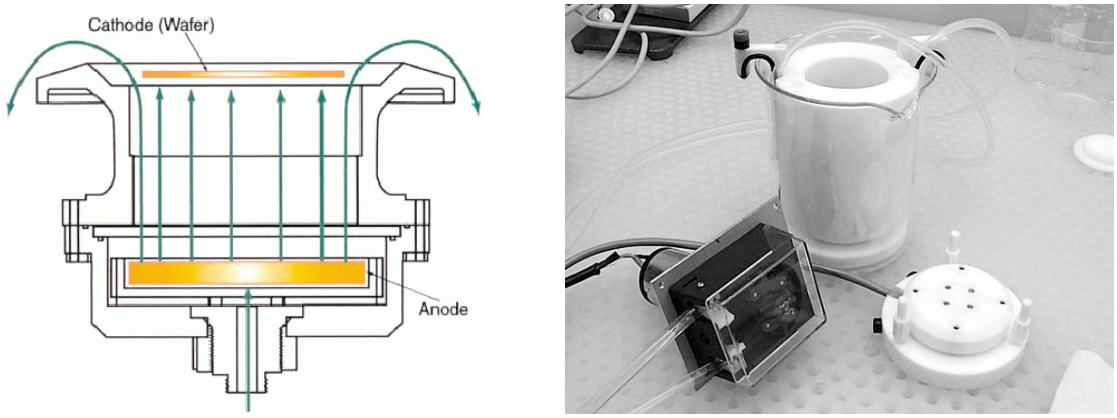
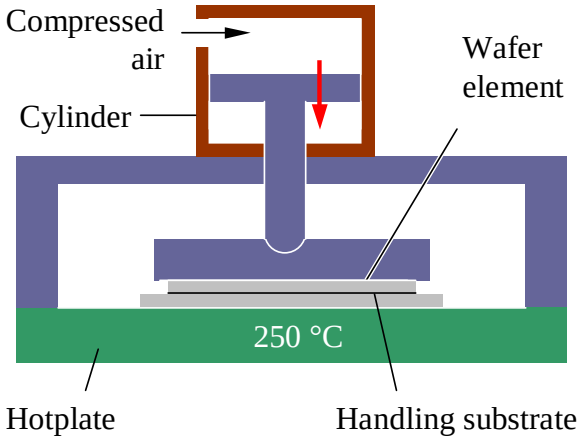


Figure C.1: Electroplating cell with horizontal rotating substrate holder [69]. The sample (cathode) is immersed in upstreaming electrolyte. Right: Setup with electrolyte recipient, peristaltic pump and rotating substrate holder, which is placed face-down onto the electrolyte recipient

C.2 Waferbonding presses



Waferbonding with the down force generated by a pneumatic actuated piston

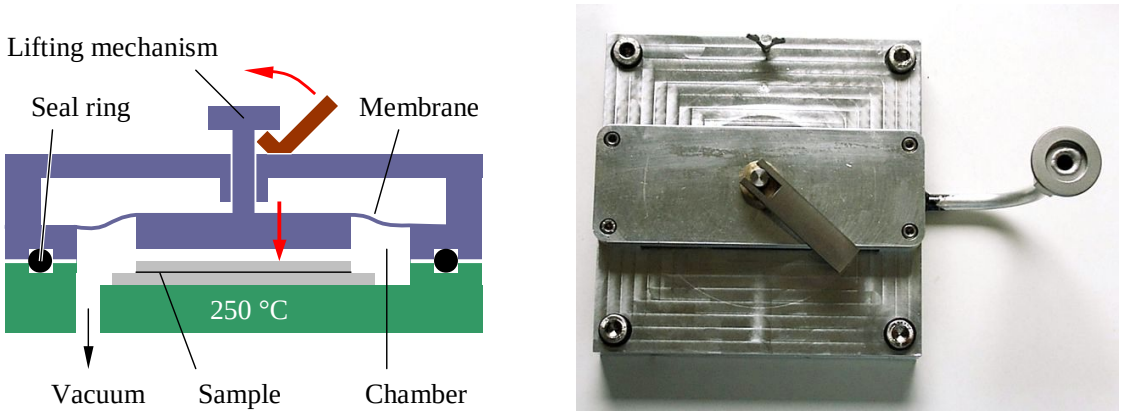


Figure C.2: Vacuum bonding with a metallic membrane in an O-ring sealed chamber. The membrane plate can be lifted for control of bonding force build-up

C.3 Thinning equipment

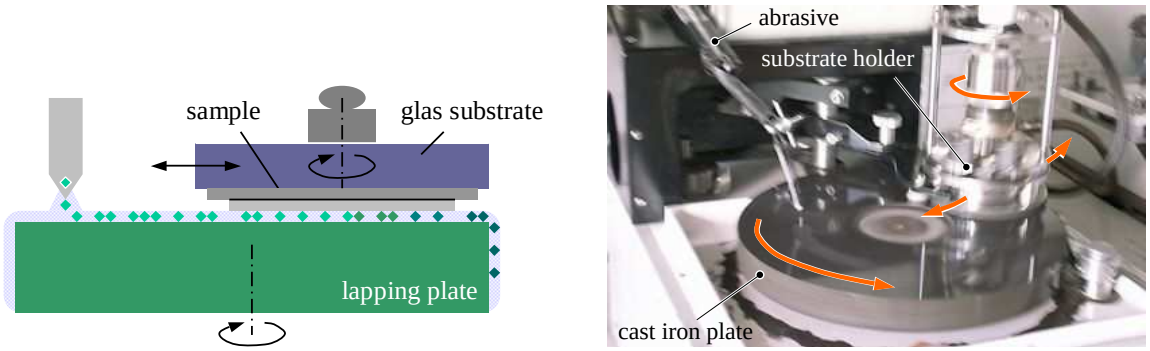


Figure C.3: Lapping machine from Logitech, used for mechanical thinning and polishing experiments. The sample holder accommodates a piston with special bearing, which prevents tilting during operation. The sample can be loaded with a down force between 1 N and 20 N ± 1 N.

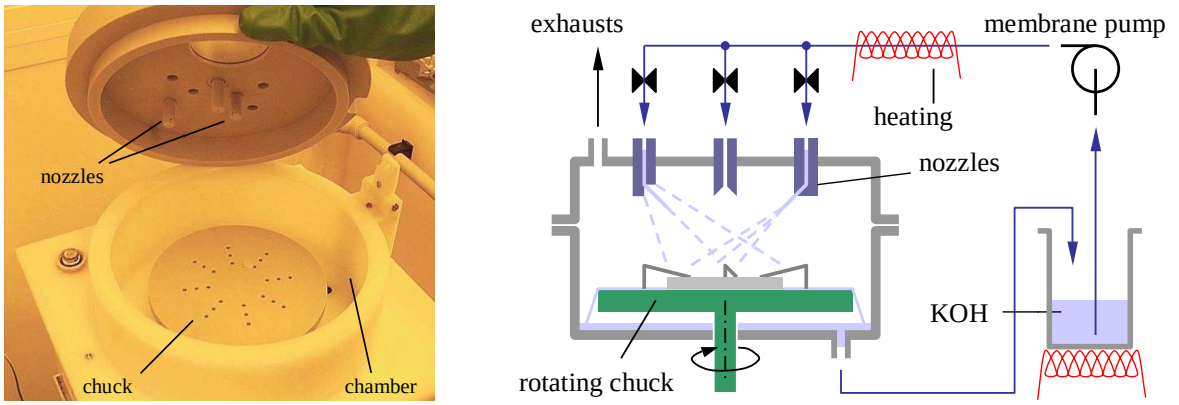


Figure C.4: Spin-etching setup for chemical thinning of wafer elements with KOH. Heated KOH-solution is delivered into the chamber by a membrane pump

C.4 Flip-chip bonder

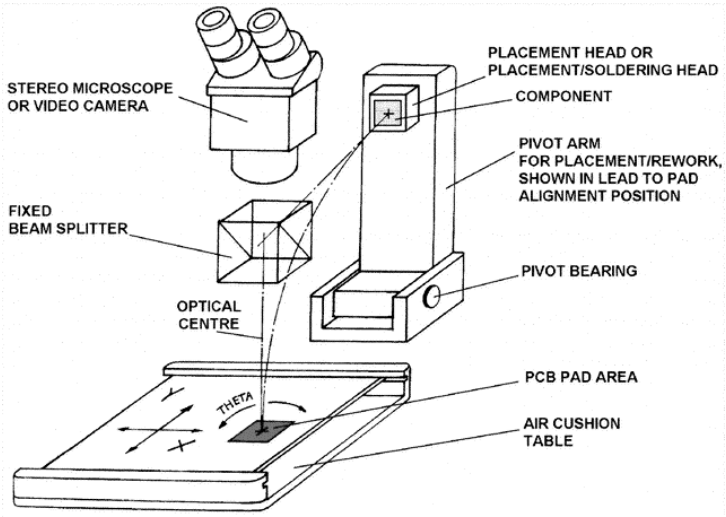


Figure C.5: Flip-chip bonding equipment for chip-to-chip stacking. An accuracy of 5 µm is achieved by a pivoted placement arm in combination with beam spitting optic [108], and a bonding force of up to 20 N can be applied to the chip. The steps of a bonding process suited for SOLID bonding are listed in Table B.2

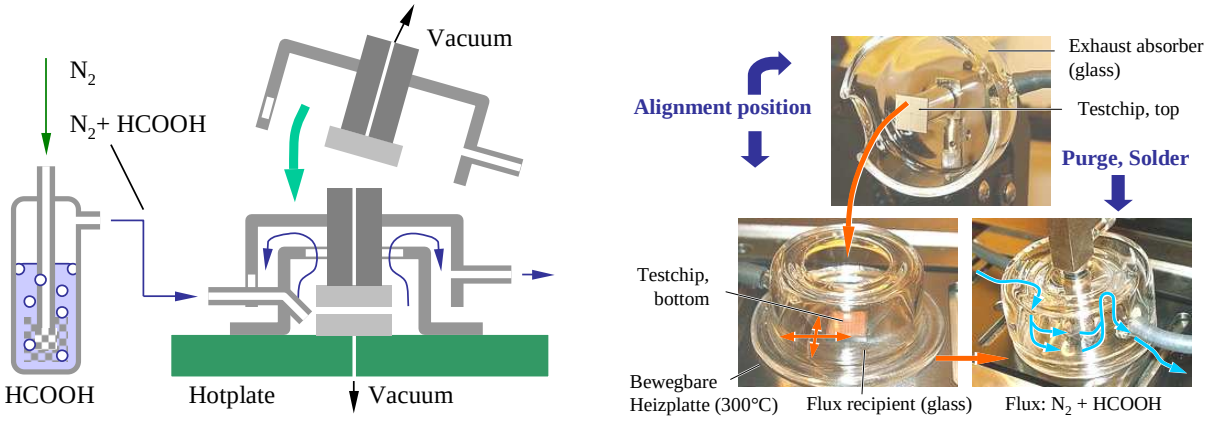


Figure C.6: SOLID bonding under controlled atmosphere in a flow chamber. The chamber is formed when the lid is closed at the lower lever position

Bibliography

[1] S. Sun, "Process Technologies for Advanced Metallization and Interconnect Systems," in *IEDM Technical Digest*, 1997, pp. 765–768.

[2] R. J. Gutmann, T. P. Chow, W. N. Gill, *et al.*, "Copper Metallization Manufacturing Issues for Future ICs," in *Materials Research Society (MRS) Symposium Proceedings*, vol. 337, Apr. 1994, pp. 41–57.

[3] Edelstein *et al.*, "Full Copper Wiring in a Sub-0.25 um CMOS ULSI Technology," *Proc. IEEE IEDM*, pp. 773–776, 1997.

[4] S. Bothra, M. Kellam, and P. Garrou, "Feasability of BCB as an interlevel dielectric in integrated circuits," *Journal of Electronic Materials*, vol. 23, no. 8, pp. 819–825, Aug. 1994.

[5] K. N. Tu, "Recent advances on electromigration in very-large-scale-integration of interconnects," *Journal of Applied Physics*, vol. 94, no. 9, pp. 5451–5473, Nov. 2003.

[6] J. D. Meindl, J. A. Davis, P. Zarkesh-Ha, *et al.*, "Interconnect opportunities for gigascale integration," *IBM Journal of Research and Developement*, vol. 46, no. 2/3, pp. 245–263, Mar./May 2002.

[7] A. Heittmann, D. Matolin, J. Schreiter, R. Schüffny, and U. Ramacher, "An Analog VLSI Pulsed Neural Network for Image Segmentation using Adaptive Connection Weights," in *ICANN 2002*, 2002.

[8] J. Schreiter, U. Ramacher, D. Matolin, A. Heittmann, and R. Schüffny, "Analog Implementation for Networks of Integrate-and-Fire Neurons with Adaptive Local Connectivity," in *International Workshop on Neural Networks for Signal Processing*. Martigny, Switzerland: IEEE, 2002, pp. 657–666.

[9] A. Heittmann, U. Ramacher, and J. Harnisch, "VisionIC: Teilprojekt Intelligente CMOS-Kamera," Dec. 2002, vortrag, Intern.

[10] M. Franz and R. Schüffny, "An Architectural Study of a Massively Parallel Processor for Convolution-Type Operations in Complex Vision Tasks," in *ICANN '96 Proceedings*. Berlin, Heidelberg, New York: Springer-Verlag, 1996, pp. 377–382.

[11] J.-U. Schlüßler, J. Werner, J. Dohndorf, *et al.*, "Current Mode Implementation of a Neural Algorithm for Image Preprocessing," in *International Conference on Microelectronics for Neural Networks and Fuzzy Systems (MicroNeuro)*. Laussane, Schweiz: IEEE, Feb. 1996, pp. 107–110.

[12] STMicroelectronics, "Semiconductor Manufacturing," Magazine, Feb. 2005, p. 25.

[13] K. D. Gann, "Neo-Stacking Technology," Technical Paper, Irvine Sensors Corporation, 3001 Redhill Avenue, Building 3, Costa Mesa, CA 92626-4529, Feb. 2000, http://www.irvine-sensors.com.

[14] K.-F. Becker, E. Jung, A. Ostmann, *et al.*, "Stackable System-On-Packages with Integrated Components," *IEEE Transactions on Advanced Packaging*, vol. 27, no. 2, pp. 268–277, 2004.

[15] Amkor, "Flip Chip Packaging Technology Solution," Internet, Nov. 2006, http://www.amkor.com/enablingtechnologies/FlipChip/.

[16] M. Bakir, H. Reed, H. Thacker, *et al.*, "Sea of Leads (SoL) ultrahigh density wafer-level chip input/output interconnections for gigascale integration (GSI)," *IEEE Trans. Electron Devices*, vol. 50, pp. 2039–2048, Oct. 2003.

[17] H. Hübner, "Face-to-Face Chip Technology Goes Productive at Infineon," Dec. 2003, 3D System Integration International Workshop, Munich.

[18] M. Koyanagi, Y. Nakagawa, K.-W. Lee, *et al.*, "Neuromorphic Vision Chip Fabricated Using Three-Dimensional Integration Technology," in *IEEE International Solid-State Circuits Conference ISSCC*, ser. Digest of Technical Papers, 2001, pp. 270–271.

[19] K. W. Lee, T. Nakamura, T. Ono, *et al.*, "Three-Dimensional Shared Memory Fabricated Using Wafer Stacking Technology," in *Proceedings of IEDM*, 2000, pp. 165–166.

[20] K. Takahashi, "3D chip stacking with copper through-vias," Dec. 2003, international Workshop 3D System Integration, Munich.

[21] P. Ramm, D. Bollmann, R. Braun, *et al.*, "Three dimensional metallization for vertically integrated circuits : Invited lecture," *Microelectronic Engineering*, vol. 37/38, pp. 39–47, Nov. 1997.

[22] A. Klumpp, R. Merkel, P. Ramm, J. Weber, and R. Wieland, "Vertical System Integration by Using Inter-Chip Vias and Solid-Liquid Interdiffusion Bonding," *Japanese Journal of Applied Physics*, vol. 43, no. 7A, pp. L829–L830, 2004.

[23] A. Klumpp, P. Ramm, R. Wieland, and R. Merkel, "Integration Technologies for 3D Systems," Dec. 2003, international Workshop 3D System Integration, Munich.

[24] R. Gutmann, J.-Q. Lu, and E. Rymaszewski, "Three-Dimensional Integrated Circuit Technology: Research Trends and Power Electronics Opportunities," 2001, rensselaer Polytechnic Institute, Troy, USA.

[25] R. J. Gutmann, A. Y. Zeng, S. Devarajan, J.-Q. Lu, and K. Rose, "Wafer-Level Three-Dimensional Monolithic Integration for Intelligent Wireless Terminals," *Journal of Semiconductor Technology and Science*, vol. 4, no. 3, pp. 196–203, Sept. 2004.

[26] M. Koyanagi, "Three-Dimensional Integration Technology by Wafer-to-Wafer Stacking," Dec. 2003, international Workshop 3D System Integration, Munich.

[27] H. Kurino, M. Nakagawa, K. Lee, *et al.*, "Smart Vision Chip Fabricated Using Three Dimensional Integration Technology," 2001, research Report.

[28] T. Matsumoto, M. Satoh, K. Sakuma, *et al.*, "New Three-Dimensional Wafer Bonding Technology Using the Adhesive Injection Method," *Jpn. Journal of Applied Physics*, vol. 37, no. 3B, pp. 1217–1221, Mar. 1998.

[29] K. Takahashi, Y. Taguchi, M. Tomisaka, *et al.*, "Process integration of 3D chip stack with vertical interconnection," in *Electronic Components and Technology 2004 (ECTC '04)*, vol. 1. IEEE, 2004, pp. 601–609, proceedings.

[30] J.-Q. Lu, T. Cale, and R. Gutmann, "3D Hyper-Integration Technology Using Dielectric Adhesive Wafer Bonding and Cu Damascene Inter-Wafer Interconnects," Dec. 2003, international Workshop 3D System Integration, Munich.

[31] H. Hübner, O. Ehrmann, M. Eigner, *et al.*, "Face-to-Face Chip Integration with Full Metal Interface," in *Advanced Metallisation Conference 2002 (AMC 2002)*, Melnick, Cale, Zaima, and Ohta, Eds., vol. V18. San Diego: MRS, 2002, pp. 53–58, iSBN 1-55899-719-9.

[32] H. Seidel, *Mikromechanik*. Berlin, Heidelberg, New York: Springer-Verlag, 1991, ch. 3.2 Nasschemische Tiefenätztechnik, pp. 125–171, iSBN 3-540-1872-9.

[33] Tummala, Ed., *Fundamentals of Microsystems Packaging*. McGraw-Hill, 2001.

[34] D. M. Kück, "Evaluation einer Technologie zur dreidimensionalen Integration von Diamantelementen," Studienarbeit, Fakultät für Ingenieurwissenschaften, Universität Ulm, 2005.

[35] A. Kaiser, D. Kück, P. Benkart, *et al.*, "Concept for diamond 3-D integrated UV sensor," in *Joint International Conference on New Diamond Science and Technology & Applied Diamond Conference (ICNDST & ADC)*, Cary, NC, USA, May 2006.

[36] F. Kalb, "Herstellung und Charakterisierung von Niedertemperatur-Siliziumoxid," Studienarbeit, Fakultät für Ingenieurwissenschaften, Universität Ulm, 2004.

[37] P. Benkart, A. Munding, A. Kaiser, *et al.*, "3-dimensional integration scheme with a thermal budget below 300°C," in *Asia-Pacific Conference of Transducers and Micro-Nano Technology (APCOT)*, Marina Mandarin Hotel, Singapore, June 2006.

[38] M. Bschorr, H.-J. Pfleiderer, P. Benkart, *et al.*, "Eine Test- und Ansteuerschaltung für eine neuartige 3D-Verbindungstechnologie," *Advances in Radio Science*, vol. 3, pp. 305–310, 2005.

[39] M. Bschorr, H.-J. Pfleiderer, P. Benkart, *et al.*, "Yield-Improving Test and Routing Circuits for a Novel 3D Interconnect Technology," *Advances in Radio Science*, vol. 4, pp. 225–229, 2006.

[40] Q.-Y. Tong and U. Gösele, *Semiconductor wafer bonding*, 1st ed., ser. Electrochemical Society Series. New York: John Wiley & Sons, Dec. 1998, iSBN: 0-471-57481-3.

[41] B. H. Yeung, V. Hause, and T.-Y. Lee, "Assessment of backside processes through die strength evaluation," *IEEE Transactions on Advanced Packaging*, vol. 23, no. 3, pp. 582–587, Aug. 2000.

[42] T. A. Hazeldine, "Back lapping semiconductor wafers," Ph.D. dissertation, ULTRA TEC Manufacturing, Inc., Santa Ana, Calif., Mar. 2001, reprint from "Ceramic Industries".

[43] H. Seidel, L. Csepregi, A. Heuberger, and H. Baumgartel, "Anisotropic Etching of Crystalline Silicon in Alkaline Solutions," *Journal of The Electrochemical Society*, vol. 137, no. 11, pp. 3626–3632, 1990.

[44] *NANO PMGI Resists Datasheet*, MicroChem Corp., 1254 Chestnut Street, Newton, MA 02464, 2002, mcc@microchem.com.

[45] *Durimide 7500 Series Polyimide Datasheet*, Arch Chemicals Inc., 400 Massasoit Avenue, East Providence, RI 02914, 2001.

[46] *Processing Procedures for Dry-Etch CYCLOTENE Advanced Electronics Resins (Dry-Etch BCB)*, Dow Chemical Company, 2030 Dow Center, Midland, MI 48674, 1998, cyclotene@dow.com.

[47] F. Niklaus, P. Enoksson, E. Kälvesten, and G. Stemme, "Low-temperature full wafer adhesive bonding," *Journal of Micromechanics and Microengineering*, vol. 11, pp. 100–107, 2001.

[48] F. Niklaus, "Adhesive Wafer Bonding for Microelectronic and Microelectromechanical Systems," Ph. D. Thesis, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden, 2002.

[49] A. Atalar and M. Hoppe, "High-performance acoustic microscope," *Review of Scientific Instruments*, vol. 57, no. 10, pp. 2568–2576, 1986.

[50] S. R. Martell, J. E. Semmens, and L. W. Kessler, "Nondestructive acoustic microimaging (AMI) analysis of MEMS materials, manufacturing, and packaging," *Proceedings of SPIE*, vol. 4558, p. 260, 2001.

[51] *Cyclotene 4000 series advanced electronic resins (Photo BCB)*, Dow Chemical Company, 2030 Dow Center, Midland, MI 48674, Apr. 2001.

[52] F. Preston, *J. Soc. Glass Technol.*, vol. 11, p. 247, 1927.

[53] H. Hocheng, H. Y. Tsai, and Y. T. Su, "Modeling and Experimental Analysis of the Material Removal Rate in the Chemical Mechanical Planarization of Dielectric Films and Bare Silicon Wafers," *Journal of The Electrochemical Society*, vol. 148, no. 10, pp. G581–G586, 2001.

[54] B. Lee, "Modeling of Chemical Mechanical Polishing for Shallow Trench Isolation," Ph. D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Instiute of Technology, 2002.

[55] A. Kaiser, A. Munding, P. Benkart, *et al.*, "3-D integration technology for sensor systems," in *WOCSEMMAD 2005*, Miama, FL, USA, Feb. 2005.

[56] H. Hübner and M. Eigner, "Dynamik ultradünner Lotschichten," Oct. 2002, interner Bericht, Infineon Technologies.

[57] D. Landolt, "Electrodeposition Science and Technology in the Last Quarter of the Twentieth Century," *Journal of The Electrochemical Society*, vol. 149, no. 3, pp. S9–S20, 2002.

[58] P. C. Andricacos, C. Uzoh, J. . Dukovic, J. Horkans, and H. Deligianni, "Damascene copper electroplating for chip interconnections," *IBM Journal of Research and Developement*, vol. 42, no. 5, pp. 567–574, Sept. 1998.

[59] A. Furuya and Y. Ohshita, "Ti concentration effect on adhesive energy at Cu/TiW interface," *Journal of Applied Physics*, vol. 84, no. 9, pp. 4941–4944, 1998.

[60] S.-Q. Wang, S. Suthar, C. Hoeflich, and B. J. Burrow, "Diffusion barrier properties of TiW between Si and Cu," *Journal of Applied Physics*, vol. 73, no. 5, pp. 2301–2320, 1993.

[61] A. A. Istratov and E. R. Weber, "Physics of Copper in Silicon," *Journal of The Electrochemical Society*, vol. 149, pp. G21–G30, 2002.

[62] *Kupferbad SLOTOCOUP BV 30*, Schlötter Galvanotechnik, Talgraben 30, 73312 Geislingen, Germany, 2002, http://www.schloetter.com.

[63] *Mattzinnbad SAT 20*, Schlötter Galvanotechnik, Talgraben 30, 73312 Geislingen, Germany, 2002.

[64] C. Xu, Y. Zhang, C. Fan, and J. A. Abys, "Driving Force for the Formation of Sn Whiskers: Compressive Stress — Pathways for Its Generation and Remedies for Its Elimination and Minimization," *IEEE Transactions on Electronics Packaging Manufacturing*, vol. 28, no. 1, pp. 31–35, Jan. 2005.

[65] J. Reid, "Copper Electrodeposition: Principles and Recent Progress," *Journal of Applied Physics*, vol. 40, pp. 2650–57, 2001.

[66] M. Q. Tran and C. M.-C. Woo, "Plating system with shielded secondary anode for semiconductor manufacturing," US Patent No. 6 402 909, June 2002, advanced Micro Devices Inc.

[67] I. I. Suni, C. Tiwari, A. A. Busnaina, *et al.*, "Copper electroplating ontol silicon wafers using high-frequency acoustic streaming," in *Proc. of the 198th Meeting of the ElectroChemical Society.* Phoenix, AR: ECS, Oct. 2000.

[68] J. Jorne, "Challenges in Copper Interconnect Technology: Macro-Uniformity and Micro-Filling Power in Copper Electroplating of Wafers," *Semiconductor Fabtech*, vol. 11, pp. 267–271, 2001.

[69] Y. Dordi and P. Hey, "Automated Chemical Management for Production Copper Electroplating," *Semiconductor Fabtech*, vol. 11, pp. 273–276, 2001.

[70] *AZ 9200 Photoresist — High-Resolution Thick Resist*, Clariant GmbH, D-65174 Wiesbaden, Germany, 1997.

[71] *Cubath SC Acid copper electroplating process for semiconductor applications*, Enthone-OMI Inc., Concord (Toronto), Ontario, June 1998, http://www.enthone-omi.com.

[72] H. Deligianni *et al.*, "A Model of Superfilling in Damascene Electroplating," *Ext. Abstracts of the 195th Mtg. of the Electrochem. Soc.*, vol. 99-1, no. 267, 1999.

[73] J. Reid, S. Mayer, E. Broadbent, E. Klawuhn, and K. Ashtiani, "Factors Influencing Damascene Feature Fill Using Copper PVD and Electroplating," Ph.D. dissertation, Novellus Systems, 2000, http://www.novellus.com/damascus/tec/tec.asp.

[74] S.-Y. Chiu, J.-M. Shieh, S.-C. Chang, *et al.*, "Characterization of additive systems for damascene Cu electroplating by the superfilling profile monitor," *J. Vac. Sci. Technol. B*, vol. 18, no. 6, pp. 2835–2841, Nov./Dec. 2000.

[75] M. J. Walker, "Comparison of Bosch and cryogenic processes for patterning high aspect ratio features in silicon," Ph.D. dissertation, Oxford Instruments Plasma Technology, 2001.

[76] G. Craciun, M. A. Blauw, E. van der Drift, P. M. Sarro, and P. J. French, "Temperature influence on etching deep holes with SF6/O2 cryogenic plasma," *Journal of Micromechanics and Microengineering*, vol. 12, pp. 390–394, 2002.

[77] C. Chung, H. Lu, and T. Jaw, "High aspect ratio silicon trench fabrication by inductively coupled plasma," *Microsystem Technologies*, vol. 6, pp. 106–108, 2000.

[78] S. M. Sze, *Semiconductors devices, physics and technology.* New York, Chichester, Brisbane: John Wiley Sons, 1985.

[79] K. W. Guarini, A. W. Topol, M. Ieong, *et al.*, "Electrical Integrity of State-of-the-Art 0.13 um SOI CMOS Devices and Circuits Transferred for Three-Dimensional (3D) Integrated Circuit (IC) Fabrication," in *Electron Devices Meeting, 2002. IEDM '02*, Dec. 2002, pp. 943– 945.

[80] K. N. Tu, A. M. Gusak, and M. Li, "Physics and materials challenges for lead-free solders," *Journal of Applied Physics*, vol. 93, no. 3, pp. 1335–1353, 2003.

[81] H. Reichl, "Technologien der Mikrosysteme II," Apr. 2001, vorlesungsskript Technische Universität Berlin.

[82] R. Schmid-Fetzer, "Fundamentals of Bonding by Isothermal Solidification for High Temperature Semiconductor Applications," in *Design fundamentals of high temperature composites, intermetallics, and metal-ceramics systems*, R. Y. Lin *et al.*, Eds., TMS. Warrendale, Pa., Feb. 1996, pp. 75–98, iSBN: 0873393058.

[83] C. Y. Liu, H. K. Kim, and K. N. Tu, "Dewetting of molten Sn on Au/Cu/Cr thin-film metallization," *Applied Physics Letters*, vol. 69, no. 26, pp. 4014–4016, Dec. 1996.

[84] S. Bader, "Erzeugung thermisch stabiler Mikroverbindungen durch isotherme Erstarrung," Dissertation, Fakultät für Chemie, Universität Stuttgart, 1990.

[85] L. Bernstein and H. Bartholomew, "Applications of Solid-Liquid Interdiffusion (SLID) Bonding in Integrated-Circuit Fabrication," *Transactions of the Metallurgical Society of AIME*, vol. 236, pp. 405–412, Mar. 1966.

[86] S. Bader, W. Gust, and H. Hieber, "Entstehung und Aufbau isotherm erstarrter Schichtverbindungen," *DVS-Berichte*, vol. 129, pp. 227–231, 1990.

[87] S. Sommadossi, "Investigation on Diffusion Soldering in Cu/In/Cu and Cu/In-48Sn/Cu Systems," Dissertation, Fakultät für Chemie, Universität Stuttgart, 2002.

[88] D. Frear, J. Jang, J. Lin, and C. Zhang, "Pb-Free Solders for Flip-Chip Interconnects," *JOM*, vol. 53, no. 6, pp. 28–32, 2001.

[89] N. N. Acharya and P. G. Mukunda, "Metallography of copper-tin alloys," *Metallography*, vol. 21, no. 2, pp. 137–150, May 1988.

[90] R. J. Fields and S. R. Low, "Physical and mechanical properties of intermetallic compounds commonly found in solder joints," Research Publication, National Institute of Standards and Technology, Metallurgy Division, Feb. 2002, http://www.metallurgy.nist.gov/mechanical_properties/solder_paper.html.

[91] D. R. Frear, *The Mechanics of Solder Alloy Interconnects*. New York: Van Nostrand Reinhold, 1994, p. 60 ff.

[92] J. S. Kang, R. A. Gagliano, G. Ghosh, and M. E. Fine, "Isothermal Solidification of Cu/Sn Diffusion Couples to Form Thin-Solder Joints," *Journal of Electronic Materials*, vol. 31, no. 11, pp. 1238–1243, 2002.

[93] G.-Y. Jang, J.-W. Lee, and J.-G. Duh, "The Nanoindentation Characteristics of Cu6Sn5, Cu3Sn, and Ni3Sn4 Intermetallic Compounds in the Solder Bump," *Journal of Electronic Materials*, vol. 33, no. 10, p. 1103 ff, 2004.

[94] "Infineon: Chip-Sandwich für größeren Speicher auf Chipkarten," Newsticker, http://www.heise.de/newsticker/meldung/52814, Nov. 2004, Heise Zeitschriften-Verlag.

[95] K. N. Tu, "Interdiffusion and reaction in bimetallic Cu-Sn thin films," *Acta Metallurgica*, vol. 21, no. 4, pp. 347–354, Apr. 1973.

[96] S. Bader, W. Gust, and H. Hieber, "Rapid formation of intermetallic compounds interdiffusion in the Cu—Sn and Ni—Sn systems," *Acta Metallurgica et Materialia*, vol. 43, no. 1, pp. 329–337, Jan. 1995.

[97] F. Bartels, "Mechanisch stabile Hochtemperaturverbindungen aus intermetallischen Phasen," Dissertation, Fakultät für Chemie, Universität Stuttgart, 1993.

[98] K. N. Tu and R. D. Thompson, "Kinetics of interfacial reaction in bimetallic Cu-Sn thin films," *Acta Metallurgica*, vol. 30, no. 5, pp. 947–952, May 1982.

[99] W. Lin and Y. C. Lee, "Study of fluxless soldering using formic acid vapor," *IEEE Transactions on Advanced Packaging*, vol. 22, no. 4, pp. 592–601, 1999.

[100] S. Ok, C. Kim, and D. Baldwin, "High Density, High Aspect Ratio Through-Wafer Electrical Interconnect Vias for MEMS Packaging," *IEEE Transactions on Advanced Packaging*, vol. 26, no. 3, pp. 302–309, Aug. 2003.

[101] A. T. W. Kempen, "Solid State Phase Transformation Kinetics," Ph. D. Thesis, Fakultät für Chemie, Universtität Stuttgart, 2001.

[102] X. Deng, G. Piotrowski, J. J. Williams, and N. Chawla, "Influence of Initial Morphology and Thickness of Cu6Sn5 and Cu3Sn Intermetallics on Growth and Evolution during Thermal Aging of Sn-Ag Solder/Cu Joints," *Journal of Electronic Materials*, vol. 32, no. 12, pp. 1403–1413, 2003.

[103] T. B. Massalski, *Binary alloy phase diagrams*, 2nd ed. Metals Park, Ohio: American Society for Metals, 1990, vol. 1–3, ISBN 0-87170-403-X.

[104] B. Predel, "Landolt-Börnstein," in *Phase Equilibria, Crystallographic and Thermodynamic Data of Binary Alloys*, ser. New Series IV. Springer, 199x, vol. 5a–5j.

[105] "SGTE Phase diagram collection," Internet, 2005, http://web.met.kth.se/dct/pd/.

[106] "Phase Diagrams & Computational Thermodynamics – Solder Systems," Internet, 2005, http://www.metallurgy.nist.gov/phase/solder/solder.html.

[107] J. Kusterer, F. J. Hernandez, S. Haroon, *et al.*, "Bi-stable micro actuator based on stress engineered nano-diamond," *Diamond and Related Materials*, vol. 15, no. 4–5, pp. 773–776, 2006.

[108] *Fineplacer automatischer Flip-Chip Bonder*, Finetech GmbH & Co. KG e-mail: finetech@finetech.de, Wolfener Str. 32/34, Haus 16, D-12681 Berlin, 2002, http://www.finetech.de.

Acknowledgement

I want to thank my family, my uncle Volker Breitfeld, and my girlfriend Mariana for their invaluable support, especially during the summer of 2006. My special thanks I want to dedicate to all people who contributed to this work, the following in particular:

Prof. Dr.-Ing. Erhard Kohn, Head of the Department of Electron Devices and Circuits, for making this thesis possible and for his advice and guidance,

Dr. Holger Hübner, Secure Mobile Solutions, Infineon, München, for providing access to technology resources within Infineon Corporation, and for his invaluable support regarding technological aspects,

Dr. Ulrich Ramacher and Dr. Arne Heittmann, Corporate Research, Infineon, München, for their support on work coordination and access to Infineon resources,

Alexander Kaiser, Peter Benkart and Markus Bschorr, the members of our 3D-research team at the University of Ulm. Countless discussions through all ups and downs of semiconductor technology and its pitfalls welded together this team,

Ralph Müller, Philip Schmid, Joachim Kusterer, Francisco Hernandez-Guillen, Michal Kubovic, Hayssam El-Hajj, Martin Häfele, Tom Zimmermann, Martin Neuburger and Peter Abele, my dear colleagues at the department of Electron Devices and Circuits of the University of Ulm, for countless time spent together, at work and in private,

Franziska Kalb, Daniel Kück and Ulrich Mayr, the students who pushed forward our research activities,

Dr. Wolgang Ebert, Alex Schreiber and Jakiv Men, University of Ulm for maintaining technology resources at the Department of Electron Devices and Circuits at the University of Ulm,

Willi Kogler, Mechanics Workshop, University of Ulm, for manufacturing custom designed devices, namely flip-chip-bonder modifications, wafer-bonder, anodic oxidation setup, and many more,

Dr. Alfred Plettl, Sebastian Fricker Department of Solid State Physics, University of Ulm, for cooperation on anisotropic Si etching (ASE) of via-holes,

Holger Torwesten, Mr. Schäfer and Mirjana Rest Infineon, for providing manufacturing services and support,

Sigrid Hill, Dean's Office, Faculty of Engineering and Computer Science, University of Ulm for helping me submitting the thesis from Norway.

Curriculum vitae

Name:	Munding,	Andreas Hubert
Date of birth:	*10.03.1975*	Riedlingen, Germany
Parents:	Munding, Munding,	Elisabeth Maria, born Breitfeld Eugen Alfons
Education:	*1981 – 1985* *1985 – 1991* *1991 – 1994*	Elementary school in Riedlingen High school (Realschule) in Riedlingen Secondary School (Gymnasium) in Biberach, Germany
Military service:	*1994 – 1995*	Murnau am Staffelsee, Germany
Academic studies:	*1995 – 2002*	Electrical Engineering, Focus on Informations Technology, University of Ulm, Germany
Abroad experience:	*1999 – 2000* *2000*	Student exchange in Canada, Queen's University, Kingston, Ontario Practical training in Puebla, Mexico
Work experience:	*2002 – 2006* *2006 –*	Research and teaching, Department of Electron Devices and Circuits, University of Ulm, Germany Research and development, Institute for Microsystems Technology, Vestfold University College, Horten, Norway

Horten, July 11, 2007

