Numerical and Experimental Investigations on Mechanical Stress in 3D Stacked Integrated Circuits for Imaging Applications

Clément Sart

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Matériaux, Mécanique, Génie Civil, Électrochimie**

Arrêté ministériel : 25 mai 2016

Présentée par

Clement SART

Thèse dirigée par **Rafael ESTEVEZ , Professeur, Université Grenoble Alpes** sous l'encadrement de **Vincent FIORI**, **Sébastien GALLOIS-GARREIGNOT** et **Sandrine LHOSTIS**

préparée au sein du **Laboratoire de Science et Ingénierie des Matériaux et Procédés (SIMaP)**
dans **l'École Doctorale d'Ingénierie - Matériaux, Mécanique, Environnement, Énergétique, Procédés, Production (I-MEP2)**

Numerical and Experimental Investigations on Mechanical Stress in 3D Stacked Integrated Circuits for Imaging Applications

Thèse soutenue publiquement le **13 décembre 2019**,
devant le jury composé de :

M. Olivier THOMAS
Professeur, Aix Marseille Université, Président, Rapporteur
M. Philippe DJEMIA
Professeur, Université Paris 13, Rapporteur
Mme Hélène FRÉMONT
Maître de Conférences HDR, IMS Bordeaux, Examinatrice
M. Rafael ESTEVEZ
Professeur, Université Grenoble Alpes, Examinateur, Directeur de thèse
Mme Sandrine LHOSTIS
Docteur-Ingénieur, STMicroelectronics Crolles, Invitée, Co-encadrante
M. Sébastien GALLOIS-GARREIGNOT
Docteur-Ingénieur, STMicroelectronics Crolles, Invité, Co-encadrant

Contents

Acronyms

3D three-dimensional
ADAS automated driver-assistance systems
ADC analog/digital converter
AFM atomic force microscope
APS active pixel sensor
ARE area release energy
ASIC application-specific integrated circuit
BCB benzocyclobutene
BE back-end
BEoL back-end-of-line
BGA ball-grid-array
BOA bond-over-active
BSI backside-illuminated
C4 controlled-collapse chip connection
CIS CMOS image sensor
CMOS complementary MOS
CMP chemical-mechanical polishing
CO contact
CSP chip-scale packaging
CTE coefficient of thermal expansion
CUP circuit-under-pad
CV constant velocity
CVD chemical vapor deposition
DC direct current
DIB device interface board
DRAM dynamic random-access memory
DRIE deep reactive-ion etching
EBSD electronic backscattering diffraction
ECD electrochemical deposition
EDX energy-dispersive X-ray spectroscopy
EFO electronic flame-off
F2B face-to-back
F2F face-to-face
FAB free air ball
fBEoL far-back-end-of-line
FC flip-chip
FE front-end
FEoL front-end-of-line
FIB focused ion beam
FSG fluorosilicate glass
FSI frontside-illuminated
GPS generalized plane strain
GS global shutter
HAST highly accelerated stress test
HB hybrid bonding
HBM hybrid bonding metal
HBV hybrid bonding via
HDPCVD high-density plasma CVD
HDR high-dynamic-range
HTS high-temperature storage
IC integrated circuit
ILD interlayer dielectric
IMC intermetallic compound
IMD intermetal dielectric
IO input/output
IRDS International Roadmap for Devices and Semiconductors
ISP image signal processing
ITRS International Roadmap for Semiconductors
KOZ keep-out zone
LGA land-grid-array

LPCVD low-pressure CVD
MC molding compound
ML metallization level
MOL middle-of-line
MOS metal-oxide-semiconductor
MOSFET MOS field-effect transistor
MPU microprocessing unit
NRE nodal release energy
NTRS National Roadmap for Semiconductors
PECVD plasma-enhanced CVD
PGA pin-grid-array
PMD pre-metal dielectric
PSG phosphosilicate glass
PTFE polytetrafluoroethylene
PVD physical vapor deposition
RF radiofrequency
RS rolling shutter
RS row select
RST reset
RT room temperature
RUC representative unit cell
SACVD sub-atmospheric CVD
SEM scanning electron microscopy
SF source follower
SiP system-in-package
SLID solid-liquid interdiffusion
SN sensing node
SOI silicon-on-insulator
SPAD single-photon avalanche diode
STI shallow trench isolation
TC thermal cycling
TCR temperature coefficient of resistance
TEM transmission electron microscopy
TEOS tetraethyl orthosilicate
THB thermal-humidity bias
TSV through-silicon via
ULK ultra-low-k
USG undoped silicate glass
UV ultraviolet
VCCT virtual crack closure technique
via vertical interconnect access
WB wire bonding
WLCSP wafer-level chip-scale-packaging

Remerciements

Ces travaux de thèse ont été menés dans le cadre d'un contrat CIFRE à STMicroelectronics Crolles, en collaboration avec le laboratoire SIMaP de l'Université Grenoble-Alpes.

Je tiens d'abord à remercier Rafael Estevez d'avoir accepté de diriger cette thèse, pour son suivi régulier et ses conseils, ses encouragements et sa bonne humeur ! Je remercie également vivement Vincent Fiori, Sandrine Lhostis et Sébastien Gallois-Garreignot pour leur encadrement. Merci de m'avoir fait confiance pour mener à bien ce projet, puis donné tous les moyens pour réussir. J'ai beaucoup apprécié travailler à vos côtés. Merci pour votre soutien, votre rigueur, votre enthousiasme, votre patience aussi...

À MM. Olivier Thomas et Philippe Djémia, j'adresse mes plus sincères remerciements pour leur retour critique et attentif sur ce manuscrit, malgré sa longueur, et d'avoir accepté de participer à mon jury de thèse. Je suis également très reconnaissant envers Mme Hélène Frémont d'avoir examiné ce travail, pour ses remarques constructives et ses encouragements.

Merci à Clément Tavernier et Roberto Gonella de m'avoir accueilli dans leurs services respectifs à STMicroelectronics Crolles, et tout particulièrement à Eric Sabouret, Emmanuel Vincent et Olivier Noblanc pour l'après-thèse.

Une partie de ce travail a été rendue possible grâce au soutien d'une multitude d'experts, notamment dans les équipes R&D Crolles 300 à STMicroelectronics. Merci à Frédéric Lorut pour les observations FIB/SEM, à Vincent Coutellier pour les essais de nanoindentation, à Sébastien Pelissier pour les mesures AFM, à Bart Van Haaren pour le design des structures de test, à Frédéric Ibars et Yann Henrion pour les mesures électriques. Merci également à Grégory Imbert, à l'initiative de la campagne de cratering test.

Je dois beaucoup à Alexis Farcy et Sandrine Lhostis pour leur soutien dans les derniers jours avant la soutenance. Comme des générations de thésards avant moi Alexis, j'ai pu bénéficier de ton coaching en présentation de transparents (en plus de ton expertise sur l'intégration 3D). Tu mérites vraiment ton titre de "gourou" des slides... Quel chemin parcouru entre la première répétition et la soutenance !

Merci à vous avec qui j'ai eu la chance de partager ce moment, de près ou de loin, temporairement ou depuis le tout début. Aux collègues de MTM, Idir, Komi, Axel, Sarah et Marie-Clotilde. Aux collègues de CPI, Chantal, Christophe, Flore, Greg, Lucile, Olivier, Vivien, Yves. Aux collègues de TCAD, Antonin, Benjamin, Chrystel, Denis, Edgar, Edouard, Fabio, Floria, Fred, Gaby, Gaëlle, Guillaume, Jérémy, Marianne, Nico, Pascal, Pierre-Louis, Olivier(s), Romain, Sébastien, Thomas, Yassine, Yvan...

A mes amis et ma famille, pour leurs encouragements tout au long de la thèse (*"Alors cette thèse, ça avance ??"*).

A Xiaoxiao... Pour ton soutien sans faille, tes sacrifices dans les moments difficiles, et ton infinie patience pendant ces "trois ans". Il ne nous reste plus qu'à profiter maintenant !

To Xiaoxiao
To my family

General Introduction

Pursuing the trend towards miniaturization and increased performance of integrated circuits, new processes and architectures are constantly developed. For decades, integrated circuit manufacturing has been driven by a continuous increase of the number of semiconductor devices (typically, transistors) integrated per unit surface. In recent years however, many challenges have arisen in that regard. Indeed, as transistor dimensions reach the ten-nanometer range quantum effects become no longer negligible, leading to increased power consumption and heat generation due to leakage currents. In addition, the need for ever-increasing wiring density between semiconductor devices also results in larger delays in the signal propagation due to the generation of parasitic capacitances.

As an alternative to further device downscaling, a progressive shift towards *3D integration* is currently observed in the semiconductor industry. The aim of this innovative approach is to combine chips of different technologies or different functionalities into a single module. This is typically achieved by vertically stacking integrated circuits instead of tiling them on a plane, enabling considerable gain in several aspects: (i) increased compactness, by leveraging the vertical direction to integrate a larger number of devices at constant chip footprint, (ii) decreased complexity, as all of the different functionalities need not be integrated within a single circuit, (iii) improved versatility, since already developed chips can be combined to form a more complex system, for which the different building blocks can be developed independently.

In particular, 3D integration opens up a host of new possibilities for imaging applications. Indeed, vertical stacking enables to move existing peripheral circuitry around the pixel matrix to a separate chip directly below the image sensor, allowing increased performances through shorter wiring, but also the incorporation of additional functionalities such as advanced image signal processing.

Among various possible strategies to process direct vertical interconnections between the image sensor and its carrier chip (e.g. through-silicon vias, metal pillars), a promising method is Cu/SiO_2 hybrid bonding. This technique consists in the simultaneous direct bonding of metal interconnection pads and their surrounding dielectric surfaces on both sides of the assembly, thereby providing mechanical and electrical connection between the stacked chips, with an interconnection pitch only limited mostly by photolithography resolution and alignment accuracy (i.e. below 1 µm).

In earlier work at STMicroelectronics and CEA-Leti, a Cu/SiO_2 hybrid bonding process has been developed on several test chips, for which the electrical performance, environmental reliability and bonding interface morphology has been extensively studied[1], allowing the transfer from the development stage to the manufacturing stage. In the present work, the integration of a more complex stack, namely a backside-illuminated image sensor on a logic integrated circuit, is considered. Compared to previous studies, a broader scope of process steps is therefore investigated, encompassing image sensor chip fabrication and its encapsulation into a supporting case for interfacing with external devices[2], in addition to the Cu/SiO_2 hybrid bonding process for chip stacking.

The focus of this thesis is on the mechanical robustness of such a 3D integrated imager-on-logic device during its fabrication, aiming to address a number of possible issues for this relatively new technology in semiconductor manufacturing and secure product integration from a thermomechanical perspective. Mechanical stresses building up in the image sensor during chip processing and assembly onto a package are investigated, and the interactions between the different system components analyzed. The mechanical integrity of several key structures is studied, namely (i) interconnection pads at the hybrid bonding interface between the imager/logic chips, (ii) bondpad structures below the wires connecting the imager to the package substrate, and (iii) semiconductor devices in the image sensor, through the evaluation of process-induced mechanical stresses using Si piezoresistive stress sensors. For each item, combined numerical and experimental investigations are carried out, relying on finite element analysis and morphological, mechanical or electrical experimental characterization.

Firstly, we will provide in Chapter I an overview of the fabrication of 3D integrated image sensors, and discuss the main benefits and new perspectives brought by chip stacking. The specific risks and challenges in terms of thermomechanical robustness for this kind of architecture will also be highlighted, and the main objectives of the thesis outlined.

In Chapter II, we will focus on the Cu/SiO_2 hybrid bonding interface between the image sensor and its carrier chip. In a first part, we will evaluate experimentally the influence of various geometries and layouts for the Cu interconnection pads on the bonding surface topography after planarization by chemical-mechanical polishing. In turn, the influence of this initial surface topography on the Cu-Cu bonding interface morphology will be examined. Then, thermomechanical finite element modeling of the hybrid bonding process at the interconnect scale will be carried out, aiming to assess the influence of additional process and design parameters identified as critical for future applications, e.g. interconnect pitch distance or alignment accuracy.

In Chapter III, we are interested in the mechanical robustness of wirebond pads at the backside of the 3D stacked image sensor. An experimental comparison between several pad architectures will be carried out, by failure inspection after the wire bonding process to detect possible cracks or delaminations in the multi-layer interconnection stack. The focus is on the influence of

[1] Taibi (2012) and Beilliard (2015)
[2] Referred to in the semiconductor industry as an “integrated circuit package”.

interconnection layout below the pad, as well as the introduction of a capping layer between the bonding surface and the interconnection stack. Aiming to provide a better understanding of the experimental trends, a multi-scale finite element analysis of a standardized wirebond qualification test, namely the wire pull test, is carried out. This model is then used to investigate the influence of various pad configurations on the mechanical robustness, and thereby propose guidelines to decrease the occurrence of pad mechanical failure.

In Chapter IV, a methodology based on Si piezoresistive stress sensors is proposed for *in situ* monitoring of the stress distribution in the active semiconductor region during the image sensor fabrication sequence, including encapsulation in a package. First, sensor calibration is carried out using a previously developed in-house instrumented four-point bending fixture to determine the piezoresistive coefficients[1]. Inline stress measurements are then carried out both at the wafer-level and the package-level, and compared with thermomechanical finite element modeling of the process sequence. The limitations of this method will be discussed and directions for further work proposed.

In a last part, the main achievements of this work are presented and propositions for further developments outlined.

[1] Ewuame (2016)

Chapter I

3D Stacking of Integrated Circuits for Imaging Applications

The aim of this chapter is to provide an overview of the fabrication processes of integrated circuits for imaging applications (i.e. image sensors), more specifically within the context of 3D stacking, a new integration strategy enabling improved compactness and performance. After an overview of integrated circuit fabrication processes, some of the key challenges currently facing the semiconductor industry are presented. In this context, the main benefits and new perspectives brought by the 3D stacking approach are discussed. For imaging applications, the shift towards 3D architectures gives rise to specific risks and challenges in terms of thermomechanical robustness, that will be highlighted. Finally, the main objectives of the thesis will be outlined.

I.1 Context

During the first half of the 20^{th} century, electronic devices consisted of individual components interconnected by pieces of wire to form discrete circuits. However, shortly after World War II, two major inventions enabled tremendous miniaturization of electronic circuits:

- the discovery of the transistor effect and the realization of the first transistor in 1947 by Bardeen, Brattain and Shockley [1], which provided a more compact replacement for the large and fragile vacuum tubes used at the time;
- the invention of the integrated circuit by Kilby and Noyce [2] in 1958, which enabled further miniaturization by embedding several electronic components together at the surface of the same substrate.

An integrated circuit (IC) can be defined as a collection of electronic devices (mostly transistors, but also diodes, capacitors and resistors) processed and electrically interconnected together onto a flat circular slice of semiconductor material (typically high-purity single-crystal silicon), called

wafer. Among the different types of electronic components present on an IC, transistors are the fundamental building blocks for circuit design. A transistor acts as a voltage-controlled switch and is used to build the logic gates that confer ICs their functionalities.

The first ICs in the beginning of the 1960s comprised only a few dozens of transistors, mostly for aerospace and military applications: a level of integration nowadays referred to as small-scale integration. After the commercialization of the first microprocessors in the early 1970s, IC manufacturing shifted to mass production, with circuits of several thousands of transistors (large-scale integration). It was around that time that Gordon Moore, co-founder of Intel, made his famous observation that the density of components at minimum cost roughly doubles every year[1] [4]. This prediction has verified quite well since then, with ICs comprising hundreds of thousands of transistors in the early 1980s (very-large-scale integration), and the million-transistor milestone[2] crossed in the second half of the 1980s (ultra-large-scale integration).

Nevertheless, in the early 1990s several major difficulties began to arise for keeping the pace of IC miniaturization and performance, requiring the introduction of new fabrication processes and new materials[3]. Since then, IC scaling has therefore been driven by technology roadmaps[4] produced by a group of semiconductor industry experts [5]. These documents assess the current challenges for IC performance improvement, determine directions for research, and set the milestones to be reached by the industry for each IC technology or processing area in the upcoming years (Table I.1).

	1995	1997	1999	2001	2004	2007	2010	2013	2016	2019	2022
Node	0.35 µm	0.25 µm	180 nm	130 nm	90 nm	65 nm	45 nm	28 nm	22 nm	17 nm	13 nm
MLs	5	6	7	7	10	11	12	13	13	14	15

TABLE I.1: Summary of semiconductor technology roadmaps: process nodes for DRAM applications and number of metallization levels (MLs) for MPU applications [6–10].

During the short history of ICs, continuous improvement of circuit performances was made possible through device miniaturization. Many technological advances in terms of the processes and materials used were necessary to sustain this trend. In the next section, the main fabrication processes involved in the manufacturing of image sensors (but also for virtually any other kind of IC) are presented.

[1] A decade later, Moore adjusted this projection to a twofold increase every two years [3].

[2] It can be noted that, rather than transistor count, the commonly used metric to designate IC generations is the *process node* or *technology node*. This terminology refers to the set of fabrication processes and design rules required to reach a given level of miniaturization.

[3] Notable examples include (i) the introduction of chemical-mechanical polishing for planarity improvement in 1990, (ii) the use of copper to replace aluminum interconnections in 1997, (iii) the move to low- and ultra-low-permittivity dielectrics to insulate fine interconnects, respectively for the 90 and 45 nm nodes. These processes and materials will be described in the next section.

[4] First from 1991 with the National Roadmap for Semiconductors (NTRS) in the USA, then from 1998 the International Roadmap for Semiconductors (ITRS), and since 2017 the International Roadmap for Devices and Semiconductors (IRDS).

I.2 Basic processing steps

Many different processes are involved in IC fabrication, each resulting from technological and scientific advances in various research fields. During the fabrication sequence, the *wafers* on which the circuits are processed go back and forth between many processing tools, each devoted to one of the following fundamental operations: film deposition, lithography, etching, oxidation, doping and surface planarization. In the following, these processes are described as they can have an influence on the morphological features, material behavior, or electrical operation of several structures present in the studied image sensor chip, e.g. hybrid bonding interconnects (Chapter II) or piezoresistive stress sensors (Chapter IV).

I.2.1 Film deposition

To form electronic components at the surface of the wafer, to create the metal contacts needed to access them electrically, or simply to provide electrical insulation, thin films of conducting and dielectric materials are deposited successively, layer-upon-layer, onto the substrate. To achieve this, many film deposition techniques have been developed throughout the last decades, each with specific features tailored to adapt to different constraints for IC fabrication. Depending on the mechanism underlying film growth, these deposition processes fall into three broad categories, namely physical vapor deposition (PVD), electrochemical deposition (ECD) and chemical vapor deposition (CVD).

I.2.1.1 Physical vapor deposition

In the semiconductor industry, physical vapor deposition is used mostly to deposit polycrystalline metal thin films. Among the earliest methods is *evaporative deposition*. With this technique, the source material is vaporized in a vacuum chamber (e.g. using a resistive coil or an electron beam) and then condenses at the wafer surface to form a thin solid film. Films obtained by evaporative deposition are highly textured, and with poor step coverage for conformal deposition.

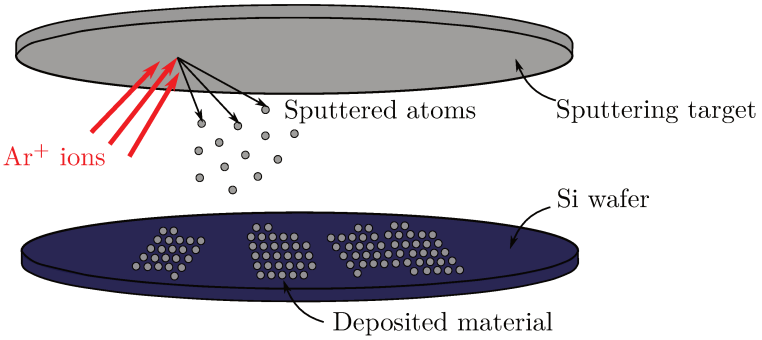


FIGURE I.1: Schematic illustration of thin-film deposition by sputtering.

The most widely used physical vapor deposition (PVD) technique for IC manufacturing is *sputtering* (Figure I.1). The sputtering target, a circular plate covered of source material, is placed in front of the wafer inside a vacuum chamber, thereby forming two parallel electrodes. A gas (typically argon) is then introduced in the chamber while a high voltage is prescribed between the two plates to create a plasma containing positive ions. The sputtering target is maintained at a negative potential, resulting in ion impingement and the ejection of source material away from the plate. The sputtered atoms then form a vapor in the chamber that deposits onto the wafer to form a thin solid film. Due to higher pressure inside the chamber, the latter technique has better step coverage compared to *evaporation*. The obtained films grow without a preferred crystallographic orientation, lead to smaller grain size, and generally contain a larger number of defects and impurities.

I.2.1.2 Electrochemical deposition

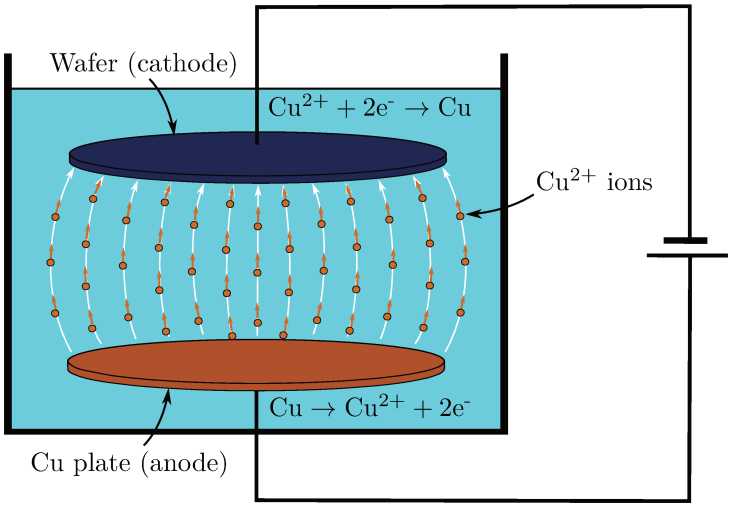


FIGURE I.2: Schematic illustration of thin-film deposition by electroplating.

Metal thin films can also be deposited by *electroplating*, or electrochemical deposition (ECD) (Figure I.2). A plate covered of source material is placed in front of the wafer to form two parallel electrodes. These are immersed in an electrolyte solution containing ions of the metal to be deposited, and connected to a DC current generator. The wafer is maintained at a negative potential to attract the positively charged metal ions. An electrochemical reaction occurs in the electrolytic cell, which turns the ions into atoms. These metal atoms accumulate at the surface of the wafer to form a film of the deposited material. This technique enables to grow thick films, in the range of microns to tens of microns.

I.2.1.3 Chemical vapor deposition

In the semiconductor industry, chemical vapor deposition (CVD) is mostly used to deposit amorphous dielectric thin films. CVD relies on chemical reactions between different carrier gases used to bring all the atoms necessary to form the desired thin film (Figure I.3). The carrier gases are admitted separately into the chamber, where they are adsorbed at the wafer surface. The adatoms and adsorbed molecules then migrate across the wafer surface to react together and grow into a thin film. Adatom mobility is enhanced due to the high temperature inside the chamber. The volatile reaction byproducts on the other hand are desorbed and evacuated out of the reactor. Contrary to sputtering, good film conformality is obtained with this technique.

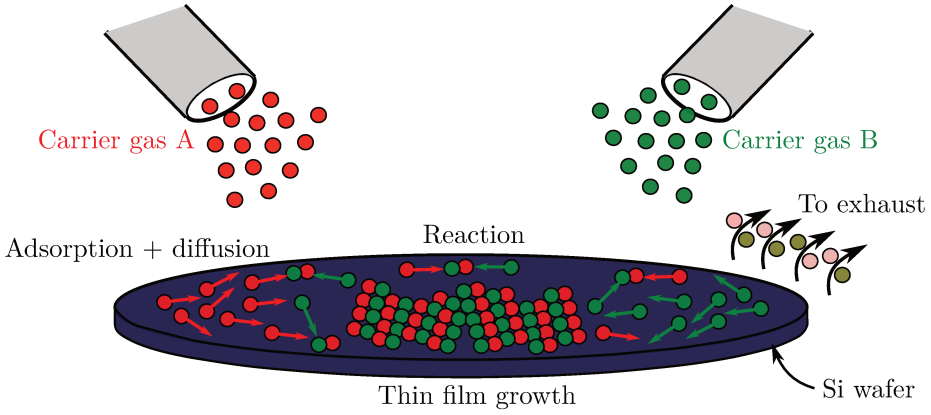


FIGURE I.3: Schematic illustration of thin-film deposition by CVD.

For some applications where low deposition temperatures are required, reactions between carrier gases can be activated using a plasma. With plasma-enhanced CVD (PECVD), an RF current source is used to induce a plasma in the chamber. This plasma enables reactant dissociation in the carrier gases, thereby promoting film deposition at lower temperatures with high growth rates and good film conformality.

A wide range of extensions have been developed in the last few decades to improve the *gapfill* capability of the CVD even further, for instance using different pressure regimes, e.g. with low-pressure CVD (LPCVD) or sub-atmospheric CVD (SACVD). Another example is high-density plasma CVD (HDPCVD), for which in addition to plasma enhancement the surface is bombarded by argon ions to remove the excess material accumulated in the corners[1], thereby leading to better sidewall coverage for high-aspect-ratio features.

[1] Indeed, film deposition around a high-aspect-ratio cavity may lead to accumulation and bridging of the excess material, eventually causing the formation of a void inside the gap to be filled.

I.2.2 Lithography

To fabricate an electronic circuit on the Si substrate, the wafer surface must be processed selectively so that a desired pattern can be formed. This is achieved using a photolithography process. The pattern is first imprinted on a photosensitive resin coating (*photoresist*) using UV light transmitted through the *photomask*, a quartz plate containing opaque chromium patterns (Figure I.4).

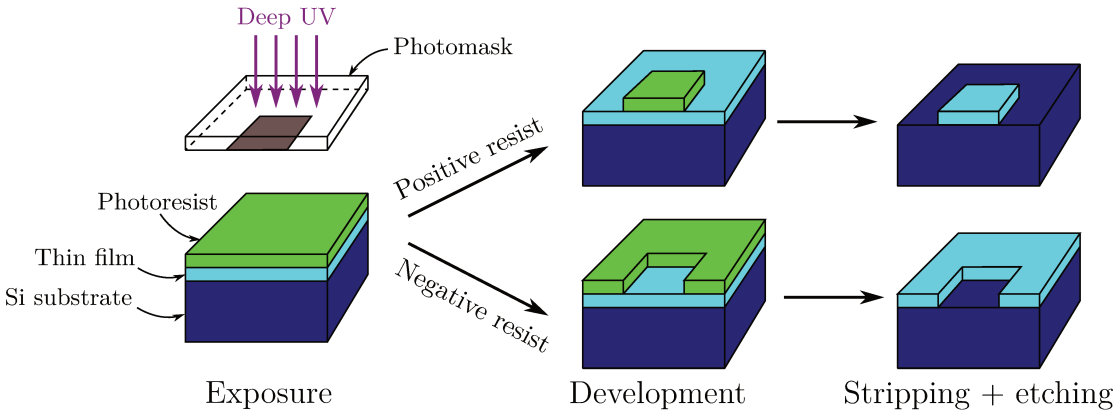


FIGURE I.4: Schematic principle of the photolithography process for layer patterning.

First, liquid photoresist is spin-coated onto the substrate and cured at about 90 °C for solvent evaporation. The wafer is then exposed to UV light through a lens system. In the exposed regions, the photoresist is chemically altered:

- either the resist is polymerized by UV exposure, making subsequent dissolution more difficult in the exposed regions (negative resist),
- or a stabilizing additive is introduced in the resist to prevent dissolution, and this additive breaks down under the action of UV light to make subsequent dissolution easier in the exposed regions (positive resist).

The removal of the exposed photoresist is referred to as the development step. It enables subsequent processing in the patterned regions exclusively, while the rest of the surface is protected under the remaining resin. Upon completion of this subsequent process step, the unexposed photoresist can be removed (or *stripped*) chemically or physically[1].

[1] For instance, using an oxygen plasma

I.2.3 Etching

In many instances during the fabrication of an ICs, it is necessary to selectively remove material, for instance to expose the Si substrate for dopant implantation, or to form trenches to fabricate metal interconnects. In either case, a photolithography step is carried out following which the wafer is coated with a patterned layer of photoresist. There are several ways to transfer this pattern to the underlying layer(s):

wet etching A straightforward method is to dip the wafer in an appropriate etching solution, e.g. buffered hydrofluoric acid for SiO_2 [11]. This approach enables selective etching, i.e. only the layer to be removed is dissolved in the etching solution.

dry etching With this method (also called reactive-ion etching), the wafer is placed in a plasma containing a suitable gas, which reacts with the material to be removed while the energetic ions in the plasma collide with the surface. The formed reaction products are thereby etched off, increasing the chemical reactivity of the underlying material [12].

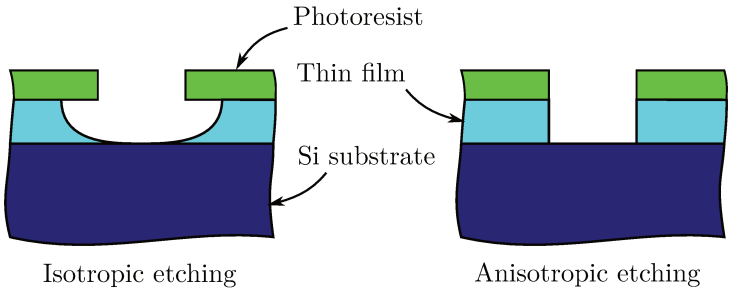


FIGURE I.5: Schematic illustration of the difference between isotropic and anisotropic behavior, obtained by wet and dry etching (respectively).

Wet etching is isotropic, i.e. material removal occurs not only in the vertical direction, but also across the sides of the cavities to be etched (Figure I.5). As a result, the obtained feature size is actually larger compared to the exposed region on the photoresist. For that reason, wet etching is primarily used to remove blanket films on the wafer.

For reactive-ion etching however, the ions able to reach the layer through the patterned photoresist are preferentially those with vertical trajectories. Therefore, anisotropic etching behavior is obtained, with etched cavities having almost vertical sides (Figure I.5). However, despite this advantage, dry etching has generally poorer material selectivity: at the end of the etching process, ions continue to bombard the underlying material[1].

[1]To avoid this, a method referred to as end-point detection is used. The end of the etching process is inferred from the detection by optical emission spectroscopy of a change in the chemical composition of the volatile reaction products etched away by ion impact.

I.2.4 Doping

To form semiconductor devices (i.e. electronic components directly embedded onto the wafer), such as the stress sensors studied in Chapter III, the electrical conductivity of the Si substrate must be enhanced in desired locations. This is achieved by the incorporation of specific doping agents introducing free charge carriers into the crystal.

Depending on whether the desired free carriers are electrons or holes[1], a different type of doping must be used, namely N type or P type (respectively). For the N type, the dopant is an atom with additional electrons compared to Si atoms, while for the P type it is an atom with fewer electrons. The doping atom must be incorporated into Si lattice sites for the free carriers to become available, therefore atoms with an electronic configuration close to that of Si are used. These doping elements are phosphorus or arsenic for N-type doping and boron or indium for P-type doping.

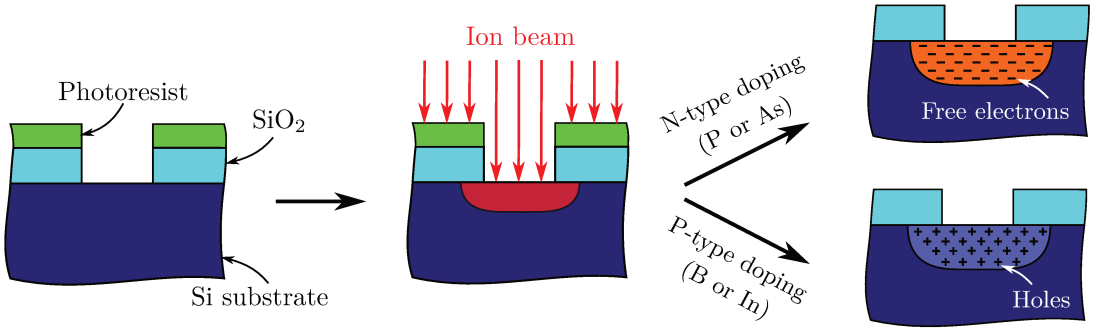


FIGURE I.6: Schematic illustration of the principle of semiconductor doping by ion implantation using a SiO_2 masking layer previously patterned by photolithography.

To introduce these impurities into the Si crystal, the most widely used technique is ion implantation (Figure I.6). With this method, the doping elements are ionized so they can be accelerated towards the wafer at high energies. A photolithography step is always carried out beforehand to protect regions on the wafer where doping is not needed. By scanning the wafer under the ion beam, the dopants are bombarded into the exposed regions of the substrate[2]. By precisely monitoring the energy of the ion beam and implantation duration, a desired implantation dose (i.e. impurity concentration) can be obtained. However, at the end of the doping process the implanted atoms are not all necessarily incorporated to the Si structure. A thermal treatment at high temperatures (900 to 1200 °C) is thus used to *activate* the dopants, i.e. allow diffusion into lattice sites. Incidentally, this activation anneal also enables to heal the defects introduced in the Si crystal due to ion bombardment. The region at the surface of the substrate with activated dopants is referred to as the *active region*.

The temperature and duration of subsequent annealing steps during IC fabrication must be carefully optimized to avoid undesired changes in the implantation profile[3] due to dopant diffusion.

[1] An *electron hole*, or simply *hole*, is an electron vacancy in the valence band of an element, which by convention can be regarded as a positive charge.

[2] To neutralize the implanted ions, the wafer is electrically grounded during implantation.

[3] Dopant concentration distribution across the thickness of the implanted region.

For this purpose, the quantity of interest is the *thermal budget* (i.e. the total amount of thermal energy transferred to the wafer) prescribed to the IC during the process sequence. In particular, for a 3D-IC assembled by hybrid bonding a thermal anneal is necessary for chip stacking, for which the thermal budget must be optimized (Chapter II).

I.2.5 Planarization

In the last thirty years, due to IC miniaturization, more stringent constraints regarding surface planarity of the deposited thin films have become necessary. Indeed, due to ICs becoming more complex, the number of layers deposited onto the substrate has considerably increased, leading cumulatively to larger surface waviness near the end of the fabrication sequence.

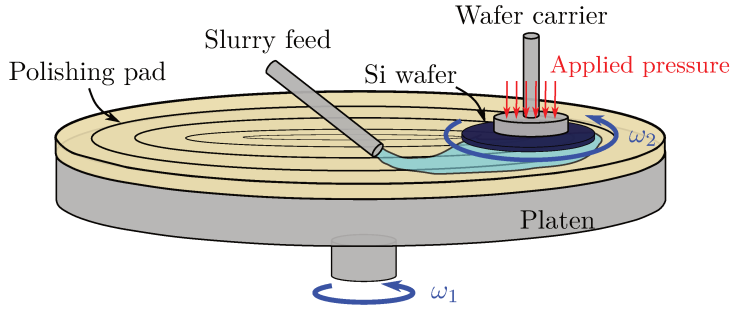


FIGURE I.7: Schematic illustration of surface planarization by CMP.

Chemical-mechanical polishing (CMP), combining chemical etching and mechanical polishing has emerged as the main method for surface planarization of the deposited layers (Figure I.7). The wafer is placed on a rotating carrier and pressed onto a soft pad mounted on a rotating *platen*. Wafer and platen both rotate in the same direction at distinct rotation speeds and thus a relative sliding velocity is prescribed for mechanical polishing. At the same time, a *slurry* comprised of an etching solution and nano-sized abrasive particles (typically silica) is dispensed onto the pad. CMP is a sequential process: the wafer undergoes several polishing steps on different platens, from the most abrasive pad to a buffing pad.

When CMP is carried out on a patterned surface with material heterogeneities, the difference in the abrasion rates may lead to surface topography defects. These defects are discussed further in Chapter II, and their potential detrimental impact on the hybrid bonding process for 3D chip stacking is studied.

I.3 Front-end processing: from bare wafer to IC

IC manufacturing is a long and complex task involving hundreds of processing steps, typically resulting in total fabrication times in the order of months. In the semiconductor industry, these processes are commonly divided into two categories, namely front-end (FE) at the beginning of the process sequence, and back-end (BE) at the end[1]. Starting from a bare Si wafer, ICs are fabricated step-by-step by combining (and repeating) the elementary operations previously described: film deposition, photolithography, etching, doping and surface planarization. Using this approach, all of the electronic components making up the IC can be processed (and interconnected) simultaneously at the substrate surface, layer-upon-layer. This method of fabricating ICs is referred to as the *planar process* or *planar technology*, and falls within the front-end process category.

I.3.1 Semiconductor devices (front-end-of-line)

Planar technology enables to fabricate semiconductor devices directly at the wafer surface. The fabrication of these devices is called *front-end-of-line* (FEoL) processing, and by extension this expression is also refers to the part of the IC multilayer stack where those devices are located. In the following paragraphs and in Appendix A, the main semiconductor devices present in a typical image sensor IC are described, such as transistors and pixel sensors. These semiconductor devices rely on the specific electrical properties obtained at the interface between P-type and N-type doped regions, called P-N junctions, described in Appendix A.

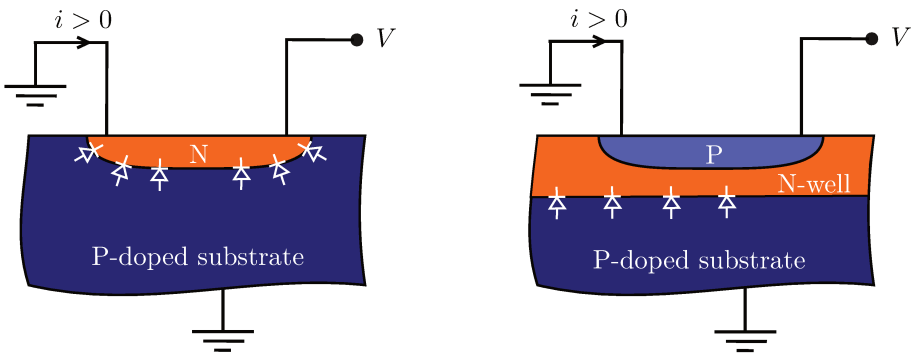
I.3.1.1 Si resistors

Resistors can be fabricated directly onto the Si substrate by selectively doping the wafer surface to form an elongated N-type region embedded within the P-type bulk[2]. If an electrical current originally flowing into the metal interconnects is diverted into this N-doped region, a decrease of conductivity is obtained: the N-doped region thus acts as resistor. The current does not flow into the grounded P bulk due to the formation of a depletion region (i.e. a diode) at the P-N junction, as described in Appendix A (Figure I.8).

This kind of device is referred to as N-type resistor. However, it is also possible to create P-resistors on the P-doped substrate. This can be achieved by first forming a large N region at the surface of the P bulk, referred to as a *well*, and then forming an elongated P-type region embedded within this N-well. In Chapter III, a stress sensor consisting of an array of such N-type

[1] Quite often, these two categories also indicate different geographical locations for manufacturing, with IC fabrication in front-end plants and IC packaging in back-end plants.

[2] Doping agents are already present in bare Si wafers used at the beginning of IC fabrication. The most widely used wafers in the semiconductor industry are P-type substrates. They are incorporated into the melt during the crystal growth process (Czochralski process) to form the silicon ingots from which wafers are then sliced.



N-type resistor P-type resistor

FIGURE I.8: Schematic illustration of the operation of N-type and P-type Si resistors.

and P-type resistors, arranged in a rosette configuration, is used to investigate process-induced stress variations in the active region of the studied 3D image sensor.

I.3.1.2 CMOS active pixel sensor

Using a P-N diode and at least three transistors (see Appendix A), a semiconductor device able to convert photons into a digital signal can be fabricated. This type of sensor is referred to as an *active pixel sensor (APS)* (Figure I.9).

The region of the pixel exposed to light is called the photodiode. In its simplest version, it can obtained by implanting an N-well into the P-type bulk to form a P-N junction. The photodiode enables the conversion of incident photons into free charge carriers (*photogeneration*) inside the depletion region at the P-N junction. Photogeneration occurs due to the photoelectric effect, i.e. the generation of an electron-hole pair upon absorption of a sufficiently energetic photon in the semiconductor substrate. To generate an electron-hole pair, the energy of the photon must overcome the bandgap energy[1] of the semiconductor. For silicon at room temperature, this amounts to an energy $h\nu > E_g = 1.12\,\text{eV}$, corresponding to wavelengths below 1100 nm. Silicon is therefore particularly well-suited for image sensing in the range of visible light.

To preserve the collected signal, recombination of the electron-hole pairs generated by photon absorption in the photodiode must be avoided. Fortunately, under the effect of the internal electric field in the depletion region, electron-hole pairs are effectively dissociated: free electrons are stored in the N well, while holes migrate to the P bulk. It can be noted that although it is also possible to use a P photodiode to collect holes instead of electrons, this is generally not the case due to the higher mobility of electrons compared to holes.

[1] The bandgap energy is the difference between the minimum energy of the conduction band and the maximum energy of the valence band of the considered element (here, silicon). It corresponds to the minimum energy needed for a valence electron to make the transition to the conduction band and become a free electron. As a free electron is created, a hole appears in the valence band and thus an electron-hole pair is generated. The bandgap energy depends on the material considered and photogeneration can be activated both through light (photon energy) or heat (phonon energy).

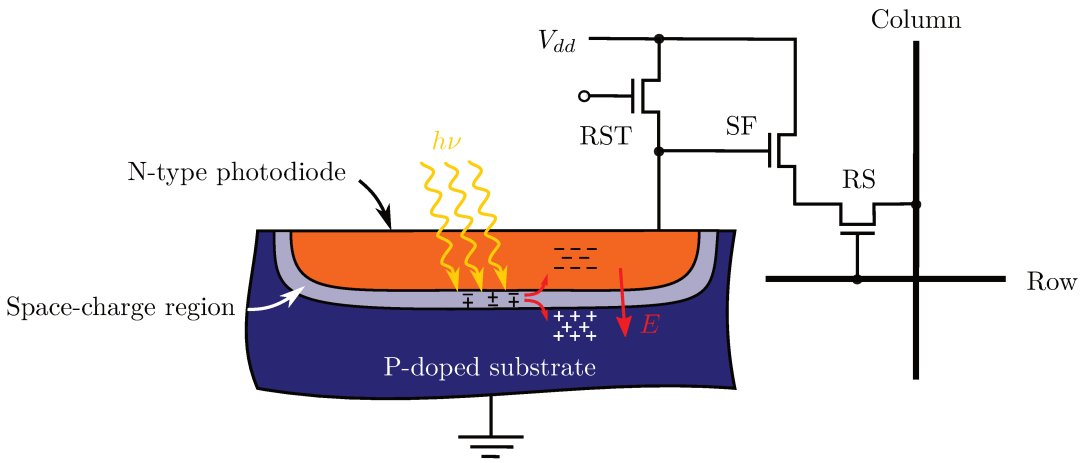


FIGURE I.9: Schematic drawing of a CMOS active pixel sensor.

This kind of device is referred to as a CMOS active pixel sensor (APS), since at least three CMOS transistors are included within the pixel in addition to the photodiode (Figure I.9). Their respective functions are (i) to convert the accumulated charges into a voltage and to amplify it (source follower transistor, SF), (ii) to transmit that voltage to the readout circuit (row select transistor, RS) and (iii) to reset the photodiode into reverse bias for charge collection (reset transistor, RST). The photodiode is directly connected to the source terminal of the reset transistor and the gate terminal of the source follower, via the sensing node (SN).

The integration sequence of the generated signal is as follows:

1. At the beginning of an integration cycle, the reset transistor is switched on to prescribe a potential to the photodiode. This voltage enables to evacuate all charges previously accumulated due to photogeneration.

2. By opening the reset transistor, charge accumulation can start again in the depletion region. The collected signal is converted to a voltage and amplified by the source follower. The row select transistor is used to transmit the amplified signal to the readout circuit.

3. Finally, the reset transistor is switched on again to evacuate all charges previously accumulated and the next integration cycle starts.

Pixels are arranged on the IC to form a matrix. The integration procedure, occurring within each pixel, enables the conversion of accumulated charges into a voltage. By varying the integration time, the exposure time can thus be controlled. The resulting analog signal is then sent to the readout circuit. Generally, the pixel matrix is addressed row by row.

The bandgap energy E_g varies significantly for a strained lattice, as does charge carrier mobility for the CMOS transistors in the pixel. Therefore, a potential adverse effect of mechanical stress on the electro-optical performance of the pixels can be expected [13], which motivates the development of *in situ* stress sensors in the active region of the studied image sensor IC discussed in Chapter III.

I.3.2 Metal-silicon contacts (middle-of-line)

All individual semiconductor devices formed at the wafer surface need to be accessible electrically. This is achieved by fabricating vertical metal plugs embedded in dielectric material (Figure I.10), providing access points to the different terminals of the semiconductor devices, e.g. between a photodiode and its associated transistors. These metal plugs are referred to as *contacts* and the surrounding dielectric matrix is called *pre-metal dielectric (PMD)*.

The pre-metal dielectric layer is generally made of phosphosilicate glass (PSG) (i.e. phosphorus-doped SiO_2), due to the ability of phosphorus atoms to capture ionic contaminants that could otherwise damage semiconductor devices in the underlying active region [14]. This material is sensitive to humidity however, and it is therefore typically capped by a layer of undoped silicate glass (USG). The typical thickness of the PMD layer is in the 100 nm range. Due to the small lateral dimensions of semiconductor devices, metal plugs are generally high-aspect-ratio structures. Therefore, the metal deposition process used to fill contact holes etched in the dielectric layer must be sufficiently conformal. For this reason, tungsten is used to fabricate the metal plugs, because historically it is the easiest metal to deposit by CVD [15]. In addition, diffusion into silicon of the metal must be prevented, and thus most low-resistivity metals such as Al or Cu cannot be used. Tungsten is a refractory metal and offers great benefit in that respect. However, it has low adhesion to SiO_2, and thus the presence of a metal liner (typically TiN) is necessary.

After CVD deposition of the PMD layer, a photolithography step is carried out to etch holes into the layer, which are filled with CVD tungsten. Subsequent chemical-mechanical polishing (CMP) enables to remove the tungsten deposited in excess. This fabrication sequence is sometimes called *middle-of-line* (MOL). Typical dimensions for the tungsten metal plugs are a few 10 nm in diameter for a thickness in the 100 nm range.

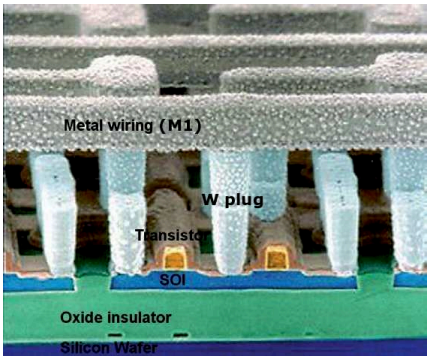


FIGURE I.10: Colorized SEM cross-section view showing, from bottom to top, the active region, the tungsten contacts and the first metallization level (M1), with the dielectric material removed (Source: IBM).

I.3.3 Metal interconnections (back-end-of-line)

In addition to the vertical contacts described above, a horizontal network of wires is also necessary to interconnect the different semiconductor devices and thereby build a circuit. This is achieved by fabricating on top of the PMD layer a network of metal lines embedded in dielectric material (Figure I.11). These metal lines are referred to as *interconnects* and the dielectric material between them as *intermetal dielectric* (IMD).

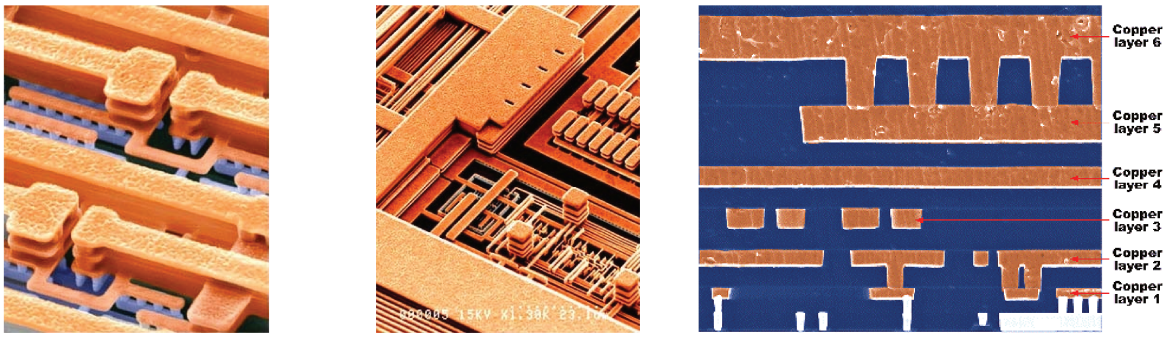
Due to the high density of semiconductor devices at the wafer surface, several tiers are actually necessary for this network of metal lines. Typically, between one and fifteen *metallization levels* (MLs) are processed at the surface of an IC. Altogether, these metallization levels form a stack about 10 µm thick. To interconnect adjacent metallization levels, vertical metal plugs are processed between the horizontal lines. These are referred to as vertical interconnect access, or simply *via*, and the dielectric material between the different metallization levels as *interlayer dielectric* (ILD). For simplicity, metal or via layers are generally referenced as M_i and V_i, with i denoting the metallization level number, in ascending order from the PMD upwards. The fabrication of these interconnects is called *back-end-of-line* (BEoL) processing, and by extension this expression is also used to refer to the interconnect stack.

The BEoL stack also enables a change in scale between the submicron-scale semiconductor devices and the outside world. It is commonly divided into several regions with increasing layer thickness, line width, spacing and length from the PMD upwards (Figure I.11):

- The lowermost layers are the *X-levels* (local interconnects): this network of thin, fine and short metal lines ensures local interconnection between semiconductor devices.
- Absent in some devices, *Y-levels* (semi-global interconnects) consist of layers of intermediate thickness, line width and density, and are used to join groups of semiconductor devices.
- The uppermost are the *Z-levels* (global interconnects): they contain the thickest metal layers, with wide metal lines enabling long distance communication between different regions on the chip, as well as power/ground distribution.

This interconnection network consists of lines with rectangular cross-section, and generally run along orthogonal directions (i.e. no oblique or curved lines), due to limitations inherent to the planar process (especially photolithography). Typical thickness for the different levels are about 100 nm for an X-level, a few 100 nm for a Y-level and about 1 µm for a Z-level. Typical lateral dimensions for the metal lines are in the 0.1-10 µm range, and for the via plugs in the 0.01-1 µm range.

Interconnects were originally fabricated using aluminum for the lines, tungsten for the vias, and SiO_2 for the dielectric layers. After deposition of an Al layer by PVD, a positive photoresist was used during the lithography step to expose the region between the lines, which were removed by dry etching. The dielectric layer was then deposited by CVD to embed the lines and surface waviness of the deposited layer was suppressed by CMP. However, Al lines lead to significant



(a) (b) (c)

FIGURE I.11: Colorized tilted SEM view of (a) local and (b) global metallization levels, and (c) cross-section view of an interconnection stack (Source: IBM).

reliability issues, such as stress voiding and electromigration [16, 17]. These issues were first alleviated using AlCu alloy lines, before the manufacturability of Cu interconnects was demonstrated in 1997 [18]. Cu lines have since progressively been adopted and replaced AlCu alloys in more advanced technology nodes, due to the great benefits of copper in terms of resistivity and electromigration resistance.

Contrary to aluminum, copper cannot be dry-etched to process the lines directly from a blanket film by subtractive etching. To circumvent this issue, a new method referred to as the *damascene process* was developed. After deposition of an etch-stop layer (typically, silicon nitride) and the dielectric layer by PECVD[1], a negative resist is used during the photolithography step to expose the region corresponding to the lines, which is removed by dry etching. A very thin Cu seed layer is then deposited by PVD, and a thick Cu film is grown by ECD to fill the etched pattern. The excess Cu material (*overburden*) is then removed by CMP. However, contrary to aluminum, copper has poor adhesion with SiO_2 and also easily diffuses into the dielectric layer surrounding the line. Therefore, a thin tantalum nitride/tantalum (TaN/Ta) liner is usually deposited by PVD before the Cu seed layer to act as a diffusion barrier and improve adhesion. After deposition of the thick ECD Cu layer, an annealing step is necessary to enable recrystallization and thereby decrease line resistivity.

Historically, the vias between Al lines had been made from tungsten, due to a strong sensitivity to electromigration with aluminum for these high-aspect-ratio structures. However, the introduction of the damascene process has enabled to fabricate Cu vias in addition to the Cu lines. For that reason, a *dual-damascene process* is actually used[2], for which both the lines and the vias are processed during the same process step for a given metallization level[3].

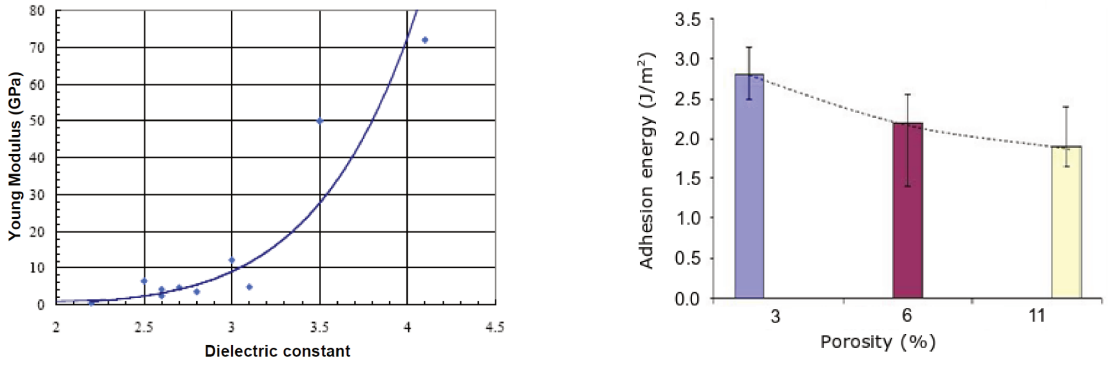
[1] PECVD is the preferred deposition method for dielectric materials in the BEoL due to its low deposition temperatures. Indeed, a great number of layers need to be deposited to form the interconnect stack. Reducing the deposition temperature for these layers enables to alleviate thermal budget constraints.

[2] Except for the M1 level resting on the PMD layer

[3] Several fabrication sequences exist for the dual-damascene process: (i) trench first, via last, (ii) via first, trench last, (iii) self-aligned.

Due to the increase of the number of semiconductor devices at the surface of ICs, finer and denser interconnect structures are necessary, especially for the X levels of the interconnect stack. A direct consequence of interconnect downscaling is the generation of parasitic capacitances between Cu lines (see Section I.5.1). Over time, the contribution of interconnects to signal delay has become significant compared to that of semiconductor devices, resulting in an overall decrease in IC performance. This limitation has led to the introduction of low-permittivity materials (*low-k*[1]) for the X and Y levels instead of conventional USG ($\kappa = 4$). Notable examples include fluorosilicate glass (FSG, $\kappa = 3.7$), i.e. fluorine-doped SiO_2, or SiOC ($\kappa = 3$), i.e. carbon-doped SiO_2. It can also be noted that for these levels, silicon nitride etch-stop layers (SiN, $\kappa = 7$) are also typically replaced with a material with lower permittivity, namely C-doped silicon nitride (SiCN, $\kappa = 5$).

This decrease in the permittivity of dielectric layers has not been without consequences for the mechanical robustness of the interconnect stack. As reported in Figure I.12, low-k dielectrics are brittle materials with low elastic modulus and poor adhesion properties. This makes the fine X-levels more prone to cracking and delamination. More recently, *ultra-low-k (ULK)* dielectrics (SiOC:H, $\kappa = 2.7$) have been developed using porous low-k for further permittivity reduction, thus resulting in even weaker mechanical properties.



(a) (b)

FIGURE I.12: Graphs illustrating (a) the decrease in the Young's modulus of low-k and ultra-low-k dielectric materials [19], and (b) the influence of ultra-low-k dielectric porosity on the adhesion energy of a SiOCH/USG interface [20].

I.3.4 Surface capping (far-back-end-of-line)

At the end of the BEoL processing, a three-dimensional network of Cu lines has been fabricated, interconnecting the different semiconductor devices. A change of scale has also been achieved, from fine local interconnects (X levels) to global interconnects (Z levels), enabling electrical access to these devices for power supply and grounding. However, the largest metal lines in the Z levels are still at most only a few microns wide. Therefore, aluminum plates are formed at the surface of the IC to act as the input/outputs (IOs) of the electronic circuit (Figure I.13 and

[1]Here, the letter κ or k is used to denote the permittivity of the dielectric material.

I.14). These IO pads have dimensions on the order of a hundred microns, much more suitable for electrical probing or IC connection onto a substrate. This is achieved by depositing an Al layer at the surface of the IC by PVD, and then forming the Al pads by subtractive etching.

In addition, a passivation stack is deposited onto the IC surface to provide mechanical and chemical protection from the external environment (Figure I.13). A phosphosilicate glass (PSG) layer is first deposited at the wafer surface by PECVD, due to the ability of phosphorus atoms to trap ionic contaminants [14]. Since PSG is sensitive to humidity however, an additional undoped silicate glass (USG) layer is deposited by PECVD. This passivation stack, very similar to the PMD layer used to protect the active region, has a thickness on the order of one micron. To provide electrical access to the IO pads, the passivation stack is then etched above the Al plates, in a step called *passivation opening*. At the end of the front-end process sequence, a final anneal of the IC is carried out to passivate defects and dangling bonds in the active region.

I.3.5 Pixel array

To guide the incident light signal to the active pixel sensors, arrays of color filters and microlenses are processed at the surface of the IC, above the pixel matrix (Figure I.13 and I.14). The color filter array consists of red, green and blue regions, forming 2×2 unit cell repeated all over the pixel matrix. In the majority of cases a Bayer pattern is used, consisting of a RGGB (red-green-green-blue)[1] cell. This is achieved by depositing pigmented resins which are then developed to form the desired pattern, by successive lithography steps, first for the green cells and then the blue and red cells. An organic layer is deposited onto the color filter array to planarize the surface before forming the microlens matrix. A resin layer is then deposited, from which the microlenses are etched and then melted during an annealing step to obtain a spherical shape.

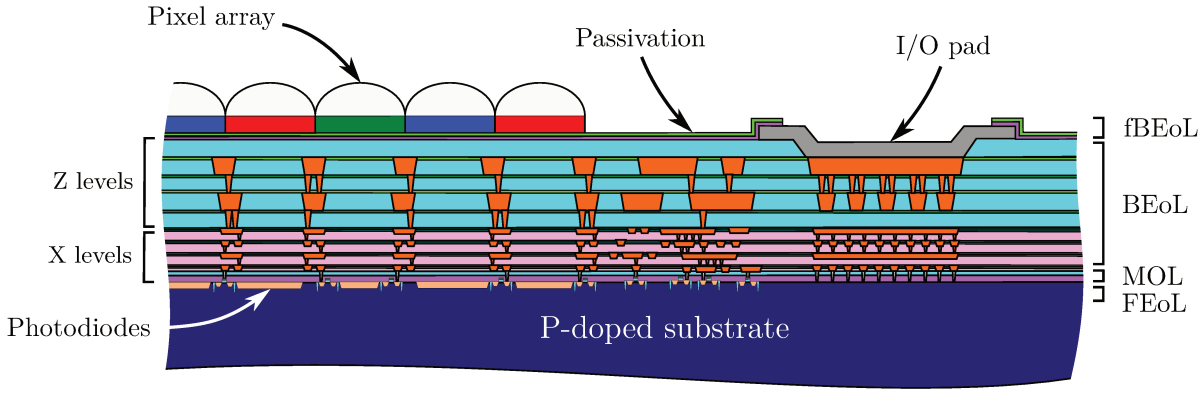
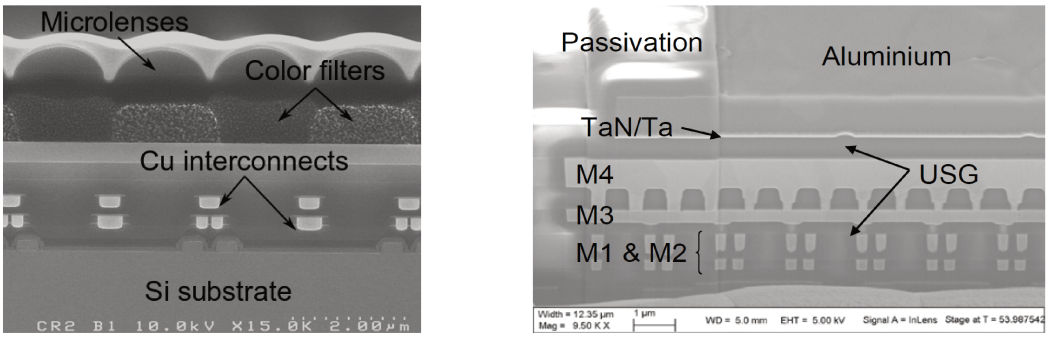


FIGURE I.13: Schematic cross-section of an image sensor (FSI integration).

The most straightforward approach to fabricate image sensors is with the color filter array and microlenses matrix deposited on the frontside of the IC, a configuration referred to as *frontside-illuminated (FSI)*. For FSI integration, the incoming light must therefore be able to travel

[1] The green color is repeated due to a peak of sensitivity of the human eye to those wavelengths, thus requiring maximum signal for this color.



(a) (b)

FIGURE I.14: a) FIB/SEM cross-section view of a FSI 1.75 µm-pitch image sensor [21], and b) FIB/SEM cross-section view of an IO pad of the 120 nm process node [22].

through the BEoL stack to reach the pixels in the active region. This leads to strong constraints on interconnection layout, due to the requirement for a transparent dielectric-only stack above the photodiodes. Although interconnections can be moved to the periphery of the pixels, a non-negligible fraction of the incident light is still reflected on the metal lines, leading to crosstalk between adjacent pixels associated with different colors [23].

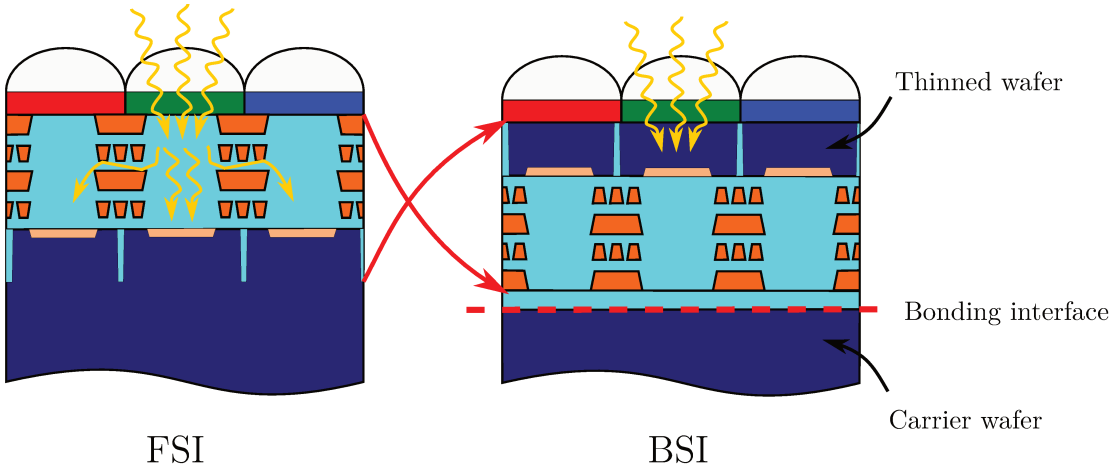


FIGURE I.15: Schematic illustration of the difference between FSI and BSI image sensors.

In the late 2000s, a new technology was introduced, namely backside-illuminated (BSI) image sensors, which has gained much importance in recent years [23] and in particular is mandatory for the 3D integrated image sensor studied in this work. For BSI integration, the colors filters and microlenses are deposited on the backside of the CMOS image sensor (Figure I.15). First, the frontside of the processed wafer is bonded[1] to a bare Si substrate. After thinning of the processed wafer by backgrinding (see Section I.4.1), color filters and the microlens matrix are formed on the backside. In this configuration, the screening effect of the interconnects is avoided, enabling

[1]The process used to attach the two wafers is called *direct bonding*. It relies on the molecular bonding phenomenon occurring between sufficiently flat and smooth surfaces placed in close proximity, due to surface interaction forces. At the end of the BEoL process, a blanket SiO_2 film is first deposited on top of the interconnect stack, which is then subjected to extensive surface preparation by CMP to reach the stringent planarity requirements for molecular bonding. Finally, the bonded assembly is annealed to increase the bonding energy. Molecular bonding is described in more detail in Chapter II.

a significant increase of the optical performances, while at the same time relaxing constraints on BEoL design [23].

In this section, the "front-end" part of the fabrication sequence for the studied image sensor has been presented. In Chapter IV, an approach based on Si piezoresistive stress sensors is proposed to evaluate the mechanical stress induced by these processing steps in the active region of the image sensor. In addition, the contribution of chip packaging to the mechanical stress in the active region will also be investigated. In the following section, the processing steps corresponding to chip packaging, i.e. the "back-end" part of the fabrication sequence, are described.

I.4 Back-end processing: from IC to packaged chip

After completion of the front-end process sequence, an array of ICs has been formed at the surface of the Si substrate. After electrical testing at the wafer level, functional ICs must then be extracted from the wafer to be encapsulated in a supporting case, referred to as a *package*. The role of the package is to protect the die from the external environment (e.g. contamination, shocks) and to act as an electrical interface for direct use in external circuits. The associated process steps fall into the *back-end* (BE) processing category.

I.4.1 Thinning

With a thickness of 775 µm for a diameter of 300 mm, Si wafers used for IC manufacturing provide sufficient mechanical stability throughout the process sequence, e.g. for wafer handling or during various annealing steps. However, this thickness range is generally too large for compact electronic products, such as hand-held and portable devices. Therefore, the backside of processed wafers is typically thinned down to between 400 to 50 µm, depending on the application.

This is achieved using a multi-step grinding process, called *backgrinding* or simply *wafer thinning*. About 90% of the initial thickness is removed by coarse mechanical grinding first, while for the remaining thickness a polishing process and/or chemical etching is used. Indeed, coarse grinding leaves scratches on the wafer backside, potentially detrimental for IC mechanical robustness in the subsequent process steps or product life, e.g. under bending loads.

I.4.2 Singulation

Individual ICs are extracted from the processed wafer in a two-step process called *wafer dicing* or *die singulation*. First, grooves about a dozen micron deep are made at the wafer surface between the ICs, either by mechanical sawing[1], laser cutting, or plasma etching. This step is referred to as *scribing*. The ICs are then singulated by cleaving the Si crystal into small square portions, called *dies* or *chips*, each corresponding to a single circuit. The backgrinding step prior to wafer dicing actually enables easier separation. Thinned dies are quite brittle however and thus during singulation the wafer must be attached to a dicing tape.

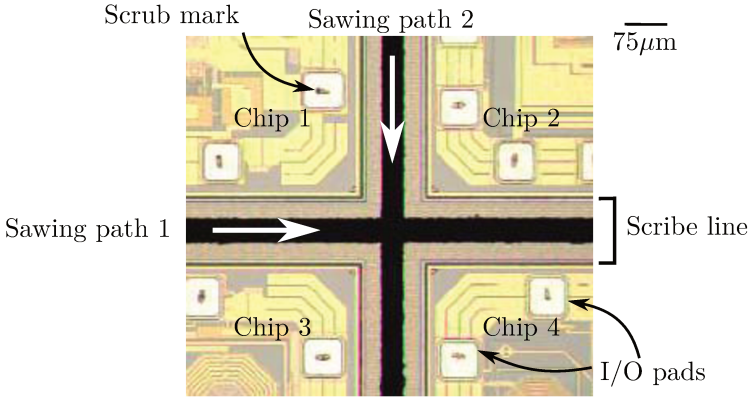


FIGURE I.16: Schematic illustration of the difference between FSI and BSI image sensors [22].

An extra space, called *scribe line* or *sawing street*, is allowed for between ICs in order to leave sufficient room for the dicing process and prevent direct damage to the circuits (Figure I.16). In addition, to avoid chipping damage at the periphery of the dies, a specific interconnect structure is designed near the edges. In this region, a dense and uninterrupted stack of Cu lines and vias is introduced, forming a continuous ring around the die to prevent cracks or delamination emanating from the scribe lines during wafer dicing. In addition, this crack-stop interconnect structure is also used to protect the die from moisture and chemical contamination from the external environment, and is thus referred to as a *seal ring*.

I.4.3 Assembly

The functional chips extracted from the wafer after singulation are assembled onto the *package substrate*, serving as an electrical interface for use in larger devices. Many strategies are available to do so: in this paragraph, the most common approaches are presented, namely wire bonding (WB) and flip-chip (FC) assembly.

[1] Mechanical sawing is performed using a diamond circular blade in presence of cutting fluid.

I.4.3.1 Wire-bonding

Wire bonding has been the chief method for die/substrate assembly long after its introduction in the 1970s [24]. In this approach, the die is attached onto the package substrate from the back side, using an adhesive joint called the *die-attach*. The electrical connection is then obtained by drawing metal wires, thinner than human hair, between the IO pads on the chip side and the *landing pads* on the package substrate side (Figure I.17). Gold is the preferred material to form the wirebonds due to its good bondability with aluminum pads. However, with the surge in the prices of gold the industry is progressively turning to copper.

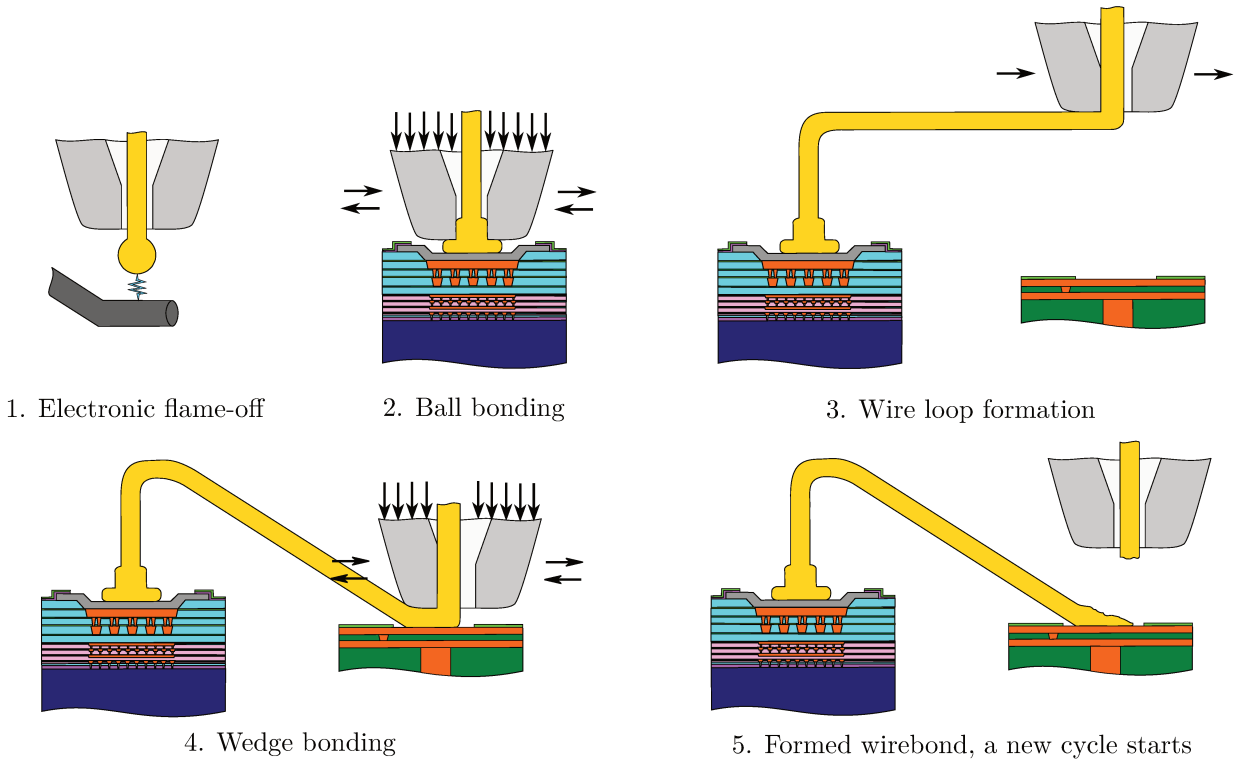


FIGURE I.17: Schematic illustration of the wire-bonding process used to form the electrical connections between the IO pads on the IC (left) and the landing pads on the package substrate (right).

In the first step of the wire bonding process, called the *electronic flame-off (EFO)*, a low energy plasma discharge is used to melt the extremity of the wire held by a capillary tool, and form a metal ball referred to as *free air ball (FAB)*. During ultrasonic bonding, this ball is then pressed onto the IO pad, while a combined lateral ultrasonic scrubbing movement and vertical compressive force is prescribed by the capillary tool. In addition to compressive loading and friction, this process relies on the ultrasonic softening[1] effect to facilitate deformation of the FAB, thereby enabling bond formation at low temperatures for high-melting point metals such as gold or copper. An extension of ultrasonic bonding consists in adding thermal energy during bond formation. This technique is referred to as thermosonic bonding. After bond formation, a

[1] Decrease of the static stress required to induce yield in a metal in which acoustic waves are transmitted, due to the resulting enhanced dislocation mobility [25].

precisely-controlled trajectory is followed by the capillary tool to draw the wire, bonded on the chip side, and form a loop towards the landing pad on the package substrate side. A coarser bond is formed on the landing pads, consisting of a flat tail without formation of a FAB. A specific IO pad layout is generally required for chips assembled by wire-bonding. For these chips, the IO pads are moved towards the periphery of the die to form an *IO ring*. This peripheral layout is well-suited to draw rows of short wires all around the chip with a low pitch distance[1].

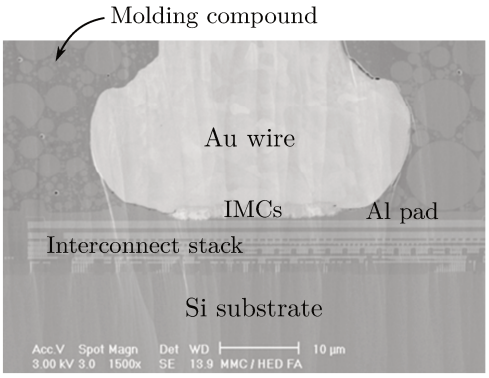


FIGURE I.18: FIB/SEM cross-section view of a gold wire bonded on an IO pad, at the periphery of the IC [26].

The bonding parameters (i.e. compressive force, lateral displacement amplitude and frequency) and the bonding temperature must be carefully validated, so that the desired height and diameter can be obtained for the mashed ball during the first bond, with sufficient robustness and reproducibility. In addition, intermetallic compounds (IMCs) are formed between the gold wire and the Al plate during ultrasonic bonding (Figure I.18). Excessive IMC growth must be avoided to prevent the formation of voids, but the IMC coverage must still be maintained at a sufficient level to provide enough adhesion between the gold wire and the Al pad [27, 28]. Another important factor is the mechanical stress transmitted to the IO pads during ultrasonic bonding. Large deformation of the Al plate occurs during bonding (*Al splash-out*), potentially leading to partial exposure of the interconnect stack below the pad, resulting in poor bond adhesion or even cracking and delamination. With the introduction of Cu wires, this becomes even more critical due to the larger stiffness of the wire material, leading to larger mechanical stress on the underlying circuits [29].

For this reason, an *exclusion region* is often defined below IO pads, where no active components or active circuitry are placed. Instead, a special interconnect structure is set in these regions, referred to as a *pad structure*, generally with a dense layout to improve the mechanical robustness the BEoL stack below the bond. This structure is also present below the IO pads for the other main assembly technique described in the following section, namely flip-chip. However, it can be noted that for some applications with stringent die size requirements, active regions or active circuitry is allowed below the IO pads. These technologies are referred to as *bond-over-active* (BOA) or *circuit-under-pad* (CUP).

[1]While the term *spacing* refers to the clearance between two structures, *pitch* denotes the distance between the respective center of these structures and is generally the quantity of interest to discuss the level of integration of an IC (e.g. transistor pitch, interconnect pitch, and here wire pitch).

I.4.3.2 Flip-chip

The flip-chip approach consists in assembling the die with the frontside directly facing the package substrate, contrary to the wire-bonding approach. The electrical connection is ensured by means of spherical or columnar metal structures called *bumps*, grown onto the IO pads at the wafer level prior to singulation, during a process step called *wafer bumping*. A major advantage of wafer bumping is that it enables high-density interconnection with the package substrate, since the whole surface of the die (not only the periphery) (Figure I.19). Due to device miniaturization, flip-chip assembly has now become mandatory in advanced applications (e.g. high-end microprocessors or graphics processors).

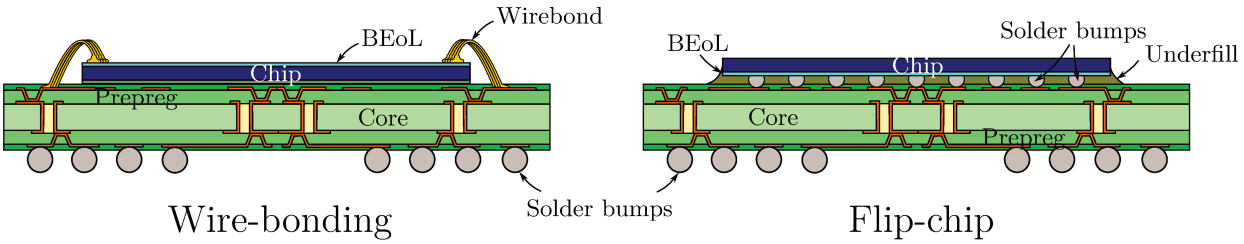


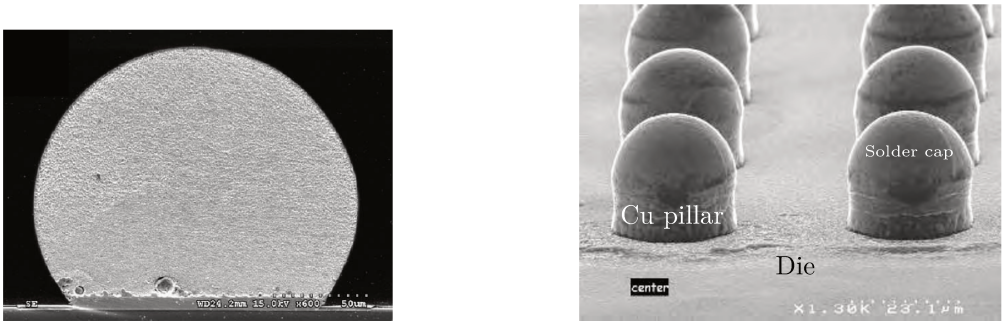
FIGURE I.19: Comparison between wire-bonding and flip-chip assembly.

Flip-chip assembly was first introduced in the 1960s, using a eutectic SnPb alloy to form spherical bumps at the chip surface, referred to as *solder bumps*, or controlled-collapse chip connections (C4s) [30] (Figure I.20a). Indeed, to form the connection between the die and the package substrate, the solder balls are brought to their melting point for about a minute (*mass reflow*). During this process, the collapse of the spherical bump is limited to the IO pad surface, due to different wettability of the Al plate and the surrounding passivation stack. This kind of connection is not suitable for fine-pitch applications however, due to the spherical shape of the bumps, which imposes a limit on the minimum distance between IO pads.

A new type of bump connection has therefore emerged in the early 2000s, namely *Cu pillars* [31]. A Cu pillar structure are grown by ECD at the surface of IO pads to slightly elevate the solder ball, enabling a reduction of the required solder volume[1] (Figure I.20). This approach enables better control over the collapse of the structure as well as smaller bump footprint, making this solution particularly well-suited for fine-pitch applications. However, Cu pillars have much larger stiffness compared to solder bumps, and therefore transfer higher mechanical loads to the interconnect structure below IO pads. To mitigate the load exerted by the bumps on the interconnect stack, it is possible to inject a resin in the inter-bump space, called an *underfill*, enabling better stress redistribution below the bumps.

A series of qualification tests is typically carried out after wire bonding or die bumping to assess the mechanical robustness of the formed bonds. A wide range of standardized testing methods is available to do so, including but not limited to wire (or bump) shear, wire pull or drop impact testing [33–36].

[1] In addition to a reduction of interconnect resistivity.



(a) (b)

FIGURE I.20: (a) SEM cross-section view of a SAC solder bump [32] and (b) SEM tilted view of Cu pillars (Source: CEA-Leti).

For the 3D integrated image sensor studied in this work, wire bonding is used for the die/package assembly. In Chapter III, we will investigate the influence of interconnect layout below the wirebond on the mechanical robustness of the pad structure.

I.4.4 Molding

After assembly, the chip has been electrically connected to the package substrate and must be sealed in a protective case to prevent damage from the external environment, e.g. shocks, contamination and corrosion. The substrate material is generally either a ceramic or a thermoset polymer[1]. To encapsulate the die/substrate assembly, a molding compound (MC) cap is formed by molding or injection, typically consisting of an thermoset epoxy filled with ultra-fine silica particles for CTE reduction. This encapsulation step is then followed by a heat or UV cure for solvent evaporation. The association of the substrate and the encapsulation cap is referred to as the *chip package*. The encapsulation step marks the end of the back-end process[2] and IC fabrication. The packaged IC is then sent for final electrical testing, following which a battery of environmental tests is carried out to assess interconnect reliability, chip/package thermomechanical interaction, humidity or contaminant diffusion into the device. Many different standardized testing methods are available for this purpose, such as thermal cycling (TC), thermal-humidity bias (THB), highly accelerated stress test (HAST) or high-temperature storage (HTS) [37–41], among others.

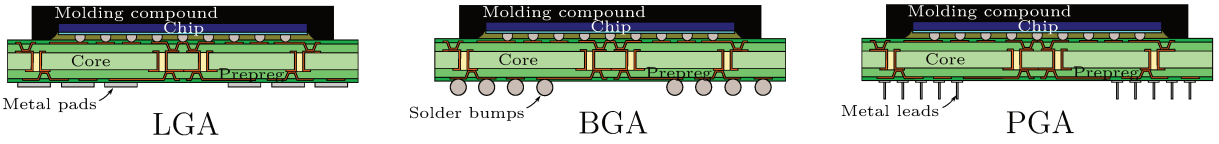


FIGURE I.21: Comparison between several types of IC packages.

[1] It can be noted however that organic packages are more widespread, due to their cost-effectiveness.

[2] The described approach, referred to as chip-scale packaging (CSP), is the most commonly encountered in the semiconductor industry. Another strategy consists in forming the package before singulation, at the wafer level, and then extracting individual packaged ICs. This method is referred to as wafer-level chip-scale-packaging (WLCSP)

The package substrate onto which the IC is assembled, in addition to providing a mechanical support for the thinned die, also contains electrical circuitry and thus enables a change of scale for use on external devices. Organic substrates typically consist of a glass fiber core, laminated on both sides with a multilayer stack of Cu foils and fiber weave, impregnated with a resin bonding agent (*prepreg*). The Cu foils are patterned during the fabrication process to create a pattern of Cu strips forming the electrical circuit. Holes are laser-drilled into the core layer and electroplated with copper for signal transmission between the top and bottom stacks. Landing pads are formed on the top and bottom surfaces, while the remaining part is coated by with a resin called the *solder mask*. Different strategies are available to form the electrical connections on the bottom side. Notable examples include land-grid-array (LGA), ball-grid-array (BGA) or pin-grid-array (PGA) packages (Figure I.21).

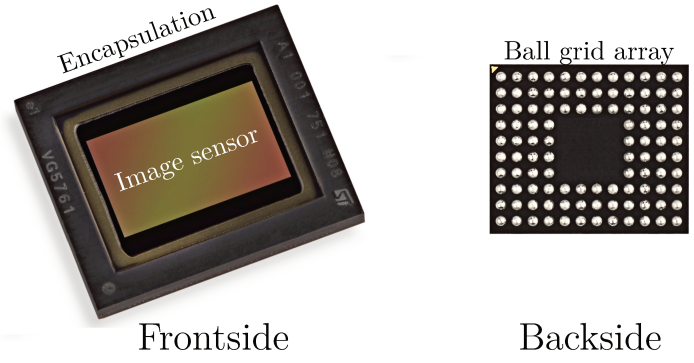


FIGURE I.22: Example of cavity package for image sensing applications (Source: STMicroelectronics).

For imaging applications, naturally the packaged die must still be able to receive incident light. Therefore a square opening is allowed for in the molding cap just above the pixel matrix. To protect the image sensor, a glass lid is placed in the opening and maintained to the cap using a *glass-attach* adhesive. An example of BGA *cavity package* is shown in Figure I.22.

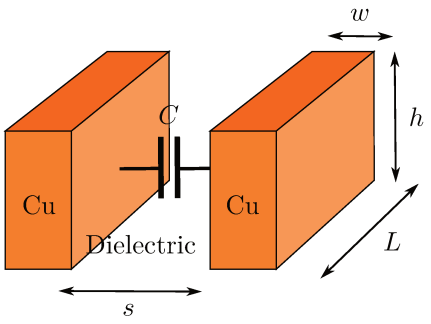
I.5 Current challenges in integrated circuit manufacturing

In the previous section, the classical IC packaging process flow has been described. Next, we present some of the current challenges in the semiconductor industry. New paradigms and innovative integration strategies currently explored to face these challenges and sustain the continuous improvement of IC performances will then be described, with a focus on imaging applications.

I.5.1 More-Moore v. More-than-Moore

A consequence of Moore's law is an increase of fabrication costs for each node. Indeed, a new generation is brought into production every two or three years, each with higher device complexity thus requiring ever-increasing R&D effort, equipment and infrastructures expenses, and manufacturing costs. As of 2018, the 7 nm node (corresponding to circuits integrating tens of billions of transistors) is due for mass-production. Only a handful of manufacturers are now able to follow this path, while others continue to produce earlier nodes, mostly above 28 nm. In addition to prohibitive costs, many challenges arise for such downscaling. Indeed, for nodes in the ten-nanometer range, quantum effects are no longer negligible, leading to increased power consumption and heat generation due to leakage currents in the transistors [42]. Even if those physical limitations are overcome, device miniaturization inevitably leads to a decrease in yield, while larger and larger production volumes become necessary to absorb the increasing fabrication costs mentioned above.

Device miniaturization also requires higher-density interconnection networks, with increasingly long and thin Cu lines. Considering two metal lines with resistivity ρ, separated by a distance s and embedded in an insulating material of dielectric constant κ, conducting currents coming from (or travelling towards) semiconductor devices in the active region of the chip. A coupling phenomenon then arises with the formation of a parasitic capacitance, leading to a delay in the signal propagation [43]:



$$\tau = RC = \rho\frac{L}{wh} \times \kappa\frac{Lh}{s} = \rho\kappa\frac{L^2}{wS} \tag{I.1}$$

with L, w, h and s respectively the length, width, thickness and spacing of the lines.

This phenomenon is thus amplified as IC miniaturization progresses, resulting in increasing device access times due to signal propagation along the interconnections. Although the move from Al to Cu for interconnects and from USG to low-k dielectric has enabled to mitigate those limitations, the main contributor to signal delay is interconnect length. By the 2000s interconnect delay had already started to overtake transistor response time (Figure I.23). Therefore, as an alternative to following the *more Moore* path, keeping up the pace of transistor miniaturization,

more and more manufacturers are turning to a *more-than-Moore* approach, based on improving IC performance through *heterogeneous integration* [44].

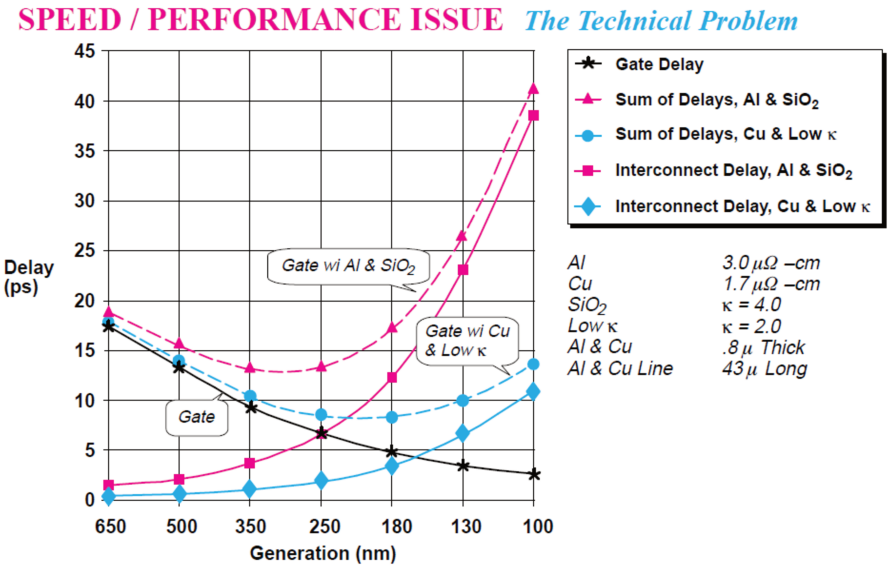


FIGURE I.23: Plot of calculated gate and interconnect delay for the different technology nodes [45].

The heterogeneous integration approach is based on the following observation. ICs are typically divided into two main categories, namely digital (logic, memory) and analog (sensors, radiofrequency, operational amplifiers). In the final application however, it is often desirable to combine several of these functionalities into a single compact, versatile, high-performance device. This is especially true of imaging applications. As presented in Section I.3.1.2, additional circuitry is necessary in CMOS active pixel sensors to gather and convert photogenerated signals from each photodiode in the pixel matrix, and thus extra space has to be dedicated to these circuits at the periphery of the image sensor chip. The aim of heterogeneous integration is to combine chips of different technologies or different functionalities into a single package, enabling considerable gain in several respects:

- smaller die size, leading to larger manufacturing volumes at constant wafer size;
- improved versatility, as chips already available can be combined to form the system and the different building blocks of this system developed independently;
- decreased complexity of individual ICs, since the different functionalities do not need to be integrated within the same chip.

I.5.2 3D packaging

Heterogeneous integration can be achieved by using standard assembly techniques, namely wire-bonding or flip-chip, to integrate several chips with different functionalities inside the same package (Figure I.24). This approach is referred to as *3D packaging*, and such a device is called a *system-in-package* (SiP) [46, 47]. While this approach enables to capitalize on mature and ubiquitous processes, thus requiring fewer equipment and development costs, it has strong limitations in terms of compactness, especially for applications such as hand-held or portable devices.

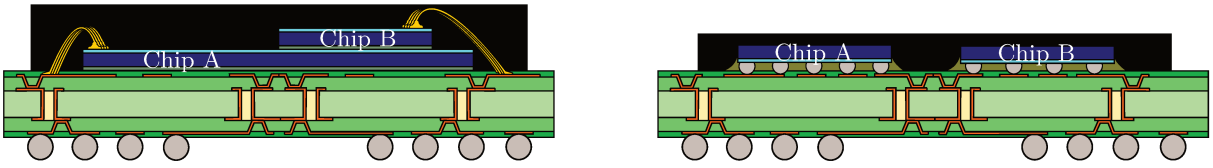


FIGURE I.24: Comparison of wire-bonding (left) and flip-chip (right) assembly for heterogeneous integration (system-in-package).

For instance, wire-bonding can be used to create multi-tier SiPs by stacking chips onto one another from the back side and forming parallel rows of wire bonds connecting each die to the package substrate (Figure I.24). Using this approach, the system form factor can reduced through vertical stacking of the chips. However, dense arrays of wires are necessary in that case, occupying additional space and increasing the risk of adjacent wires contacting (Figure I.25). In addition, similarly to wire-bonding for conventional chips, IO pads are generally located at the periphery of the die, which is a strong limitation for high-density applications.

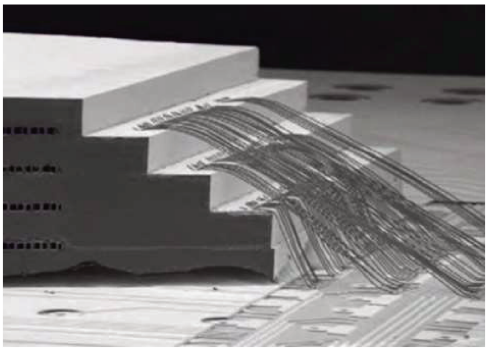


FIGURE I.25: Tilted SEM view of a system-in-package formed by vertical chip stacking using wire-bonding assembly [48].

With flip-chip assembly on the other hand, the entire die surface can be used for the connections with the package substrate, by means of a matrix of bumps. However, the dies are assembled with their frontside facing the substrate and thus cannot be stacked onto one another to form a multi-tier system, contrary to the wire-bonding approach. Instead, the different chips are laid out in the plane of the package substrate, which results in a very large occupied area for the system.

For both of these approaches, the electrical connection between the different chips is indirect, i.e. achievable only via the package substrate. New integrations strategies are thus necessary to achieve heterogeneous integration with reduced form factor, high IO density and direct connection between the chips for increased performance. In the next section, an overview of these innovative architectures is proposed, with a focus on image sensors.

I.5.3 3D integrated circuit

Three-dimensional integration can be defined as the co-integration of ICs, potentially with different functionalities, within the same package by means of direct vertical interconnections [49]. In addition to allowing heterogeneous integration (*more-than-Moore*) in the same way as SiPs, this approach also enables an increase in transistor density (*more Moore*) through chip stacking. High performance can thereby be obtained without the need for device downscaling, provided that direct connections can be formed between the different chips to reduce interconnect length [50] (Figure I.26). The key element for 3D integration is therefore the ability to manufacture vertical interconnects. In the following paragraphs, two types of vertical interconnects are reviewed, namely through-silicon vias (TSVs) and hybrid bonding (HB).

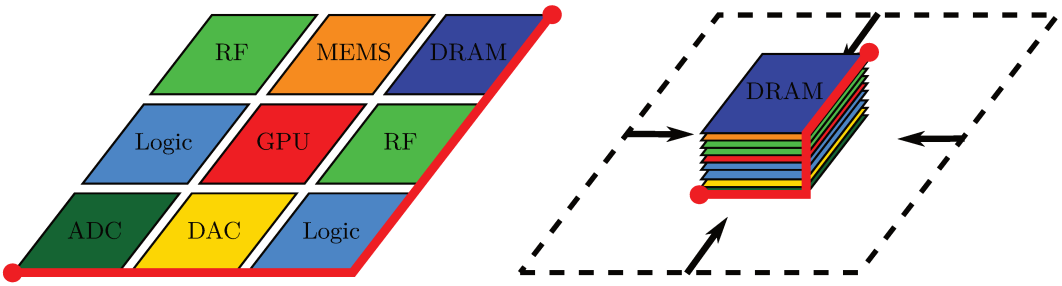


FIGURE I.26: Comparison between the planar and 3D integration approaches for heterogeneous systems.

3D integration opens up a host of new possibilities for imaging applications. Indeed, not only the readout circuits in the image sensor can be moved to a dedicated chip, but also more advanced image signal processing (ISP) functionalities can be incorporated without decreasing system compactness, as will be seen in the following paragraphs.

I.5.3.1 Through-silicon via (TSV)

Although first patented in the 1960s, through-silicon vias (TSVs) were introduced for chip stacking during the 2000s. TSVs are deep, high-aspect-ratio metal plugs etched through the silicon substrate, with nearly vertical sides [51]. They are fabricated using a deep reactive-ion etching (DRIE) process, consisting of a multi-step combined dry/wet etching process, with sidewall passivation[1]. After Cu ECD deposition into the formed cavity and wafer backgrinding, the TSVs are exposed on the backside of the wafer and can serve as a vertical interconnection between two chips.

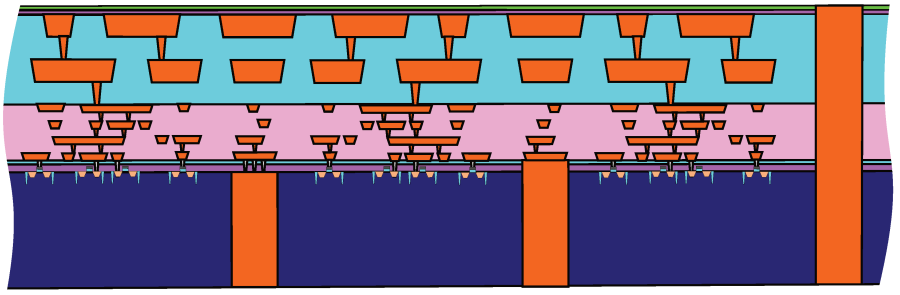


FIGURE I.27: Schematical illustration of the via-first, via-middle and via-last TSV processes [52].

Depending on the time of their fabrication during the process sequence, TSVs can be divided into three categories (Figure I.27):

- *via-first*, before FEoL processing;
- *via-middle*, after MOL but before BEoL;
- *via-last*, after BEoL processing.

A first approach for 3D integration is face-to-back (F2B) stacking, in which the different tiers are stacked with their backside facing the frontside of the underlying device layer (Figure I.28). The mechanical connection between the two tiers can be achieved using direct bonding[2] between inorganic dielectric layers or adhesive bonding with polymers, such as BCB or SU-8 resist [53]. This bonding step is generally carried out collectively at the wafer-level for increased throughput. The electrical connection is then obtained by means of TSVs, tunneling through the top tier substrate. Since the top tier needs to be thinned to be able to etch of the TSVs, a temporary carrier wafer is necessary to handle the thinned wafer during bonding.

[1] This passivation layer enables to prevent lateral etching due to the isotropic wet etch, while allowing etching at the bottom of the cavity by the anisotropic dry etch.
[2] This technique is further described in Chapter II.

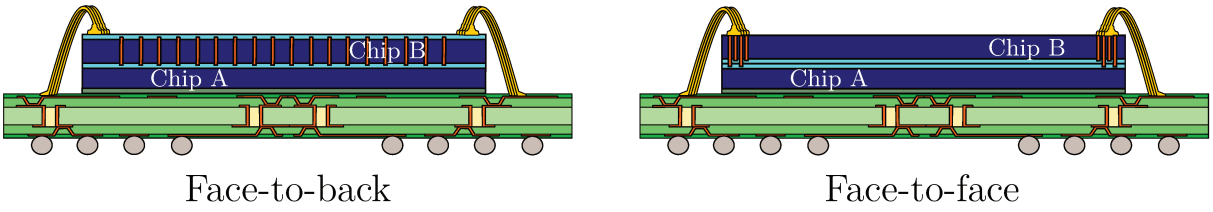


FIGURE I.28: Schematical comparison between the F2B and F2F approaches [54].

However, the F2B approach is not applicable to stack a BSI image sensor onto another IC. Indeed, BSI image sensors are flipped before bonding to their carrier wafer (Section I.4.1) and thus their interconnection levels are not directly accessible on the top surface. In this configuration, referred to as face-to-face (F2F) stacking, TSVs with different depths are required. After wafer thinning, TSV "bridges" are formed between the two device layers by etching two rows of interconnected via-last TSVs, one tunneling through the entire top tier to reach the frontside of the bottom wafer and the other enabling access to the interconnection levels of the top tier (Figure I.28).

Although the integration strategies described in this paragraph have the advantage to rely on mature techniques, namely TSV etching and direct bonding, TSVs are large structures, requiring exclusion regions in the interconnect layout, and with limited pitch. In addition, due to thermo-mechanical stresses around TSVs from the CTE mismatch with the surrounding dielectric layers, a *keep-out zone* (KOZ) is typically enforced near TSVs, where no active devices (e.g. transistors or pixels) are allowed. Another limitation of this integration strategy for image sensing applications is that the TSVs must be at the periphery of the die, due to the presence of the pixel matrix. All of these factors considerably limit the maximum density of interconnection that can be achieved between the two chips.

I.5.3.2 Hybrid bonding (HB)

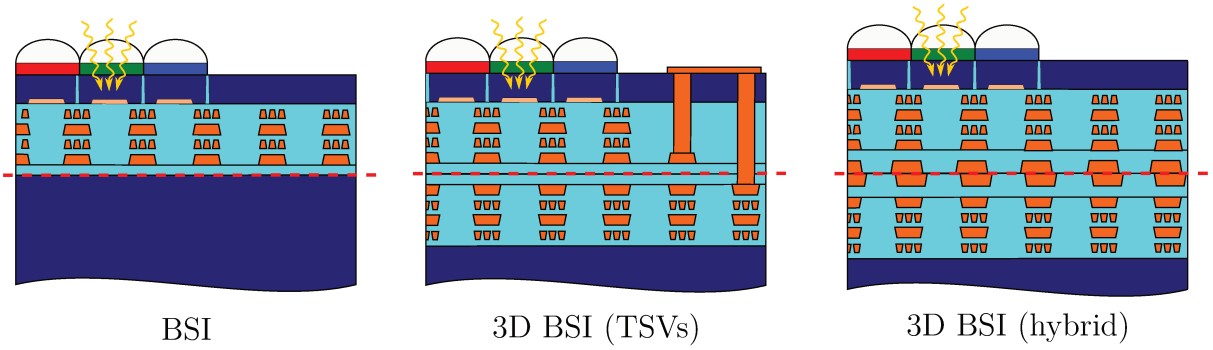


FIGURE I.29: Schematical comparison between the dual-TSV and hybrid bonding approaches for 3D stacking of image sensors.

Another way to achieve 3D integration by F2F stacking is to replace the homogeneous dielectric/dielectric bonding interface between the two tiers with a *hybrid* metal/dielectric interface,

thereby simultaneously providing electrical and mechanical connection[1]. This is achieved by forming an additional metallization level at the frontside of both tiers, with fine-pitch Cu pads immersed in organic or inorganic dielectric material (Figure I.29) [55–57]. By doing so, in the same way as flip-chip assembly, the entire die surface can be used for the connection between the two tiers, thereby enabling ultra-high-density interconnection. After bonding, the top tier is thinned and Cu plugs are processed on the backside to enable electrical connection to the package via the IO pads.

Although the use of polymer dielectric materials at the hybrid bonding interface provides several advantages [49, 58, 59], such as better tolerance to surface waviness during bonding or lower bonding temperature, these materials generally have limited temperature-stability [60] for typical deposition or annealing temperatures in the remainder of the process sequence, i.e. up to 400 °C. In addition, most organic materials are generally not compatible with usual CMOS fabrication processes, for contamination reasons [49]. Conventional SiO_2-based dielectrics on the other hand, such as undoped silicate glass (USG), enable seamless integration of the hybrid bonding step into the usual IC fabrication sequence. Indeed, the additional metallization levels required on both tiers to form the Cu/SiO_2 interface can be processed using the standard damascene process [52].

Surface planarity requirements for Cu/SiO_2 hybrid bonding are quite demanding. As detailed in Chapter II, metal/dielectric hybrid bonding relies on the molecular bonding phenomenon, occurring between extremely smooth and plane surfaces due to surface interaction forces, to make the two tiers adhere to each other at room temperature. Therefore, both device layers to be stacked must undergo extensive cleaning and surface planarization by CMP before bonding. In addition, an annealing step is necessary after bonding to increase the adhesion energy[2] [55, 61], adding up to the total thermal budget already transferred to the IC during the fabrication sequence. Nevertheless, since the hybrid bonding interconnections are fabricated using a standard damascene process, extremely small dimensions are attainable with the current capabilities of deep UV photolithography, potentially below 1 µm. To achieve the electrical connection between the two tiers however, each hybrid bonding interconnect on one side of the assembly must be bonded with its matching counterpart in the other device. Therefore, the main limiting factor for ultra-fine interconnection pitch is the alignment precision attainable with the bonding equipment [62].

Hybrid bonding offers significant advantages compared to the TSV approach for image sensing applications. Indeed, short length pixel-scale vertical connections with the bottom chip can be achieved instead of peripheral connections, thanks to the fine-pitch capability of hybrid bonding interconnects. This increased performance enables direct image signal processing below the

[1] Other technologies exist, such as Cu micropillars, which solely rely on a metal/metal interface to provide the electrical and mechanical connection between the two tiers. This kind of interconnect may be bonded by eutectic bonding, in the same way as Cu pillars in flip-chip assembly, or by solid-liquid interdiffusion (SLID) bonding, enabling low-temperature bonding with a very thin intermediate layer between the Cu micropillars.

[2] As described in Chapter II, bond strengthening occurs due to the formation of covalent bonds at the SiO_2/SiO_2 instead of the weak hydrogen bonds present after bonding at room temperature, and simultaneous reconstruction of the Cu/Cu interface occurs due to grain growth and interdiffusion.

pixel matrix and is thus particularly suitable for application to leading-edge technologies, e.g. automated driver-assistance systems (ADAS), automated face recognition, or global shutter (GS)[1].

The first instances of image sensor chips stacked onto other devices are due to academia in the 2000s, notably the MIT [63–65] and Yale University [66]. In these early devices, the image sensor was stacked using a dual-TSV approach. Although a demonstrator chip with an image sensor stacked by Ni/SiO_2 hybrid bonding was reported by Nikon in 2012 [56], the first commercial 3D BSI image sensor was brought to mass-production by Sony for smartphone cameras in 2013 [67]. A dual-TSV strategy was used to stack the image sensor onto a logic chip, enabling to integrate a high-dynamic-range (HDR) functionality into the device. They were followed in this path by major competitors on the image sensor market, such as Olympus [68–70], TSMC [71, 72] and OmniVision [73]. Among these commercial chips, although some were stacked using the dual-TSV approach similarly to previous generations, others relied on a different technology, namely metal micro-bumping [69, 74]. Simultaneously, image sensors stacked by Cu/SiO_2 or Au/SiO_2 hybrid bonding were also reported among smaller actors[2], such as research institutes or universities [75, 76]. In 2016-2017, 3D BSI image sensors stacked by Cu/SiO_2 hybrid bonding were demonstrated by Sony [77], OmniVision [78], Toshiba [79] and STMicroelectronics [80].

A detailed account of 3D stacking for BSI image sensor applications is presented in Table I.2. An increase in the number of actors and publications pertaining to metal/oxide hybrid bonding for imaging applications can be noticed in the recent period (2015-2018), with a focus reducing the interconnect pitch to improve the density of integration, and thus the performance of the image sensor. The present work in collaboration with STMicroelectronics was initiated within that context. In the next section, the objectives of this thesis are outlined.

I.6 Objectives of the thesis

As discussed in the previous section, chip stacking by Cu/SiO_2 hybrid bonding is a promising technology for high-performance image sensing applications.

In earlier work at STMicroelectronics and CEA-Leti, wafer-to-wafer and chip-to-wafer stacking by Cu/SiO_2 hybrid bonding have been compared, and their feasibility investigated. Taibi [81] has studied interconnect electrical performance, bonding interface morphology and alignment accuracy on dedicated test structures. This work has then been extended by Beilliard [82], who carried out a comprehensive study of the interconnect electrical performance, environmental

[1] As mentioned in Section I.3.1.2, in CMOS image sensors the integration of the signal coming from the pixel matrix is typically done sequentially on a row-to-row basis. Due to this mode of integration, called rolling shutter (RS), significant image distortion may occur when capturing fast-moving objects (*jello effect*). With global shutter (GS) on the other hand, the signal is integrated simultaneously from all pixels through the incorporation of additional memories and transistors, enabling to suppress this kind of artefact.

[2] It cannot be excluded that this technology has also been used by larger semiconductor companies at that time. Indeed, for a number of publications, the technology used for chip stacking is not explicitly described or ambiguous expressions such as “direct connection” or “hybrid-stacking” are used.

reliability and interfacial voiding, using experimental and numerical methods. These two studies focused on the transfer of the Cu/SiO_2 hybrid bonding technology from the development stage to the manufacturing stage, and thus could not cover other aspects involved in the integration of more complex stacks, such as the imager-on-logic device considered here.

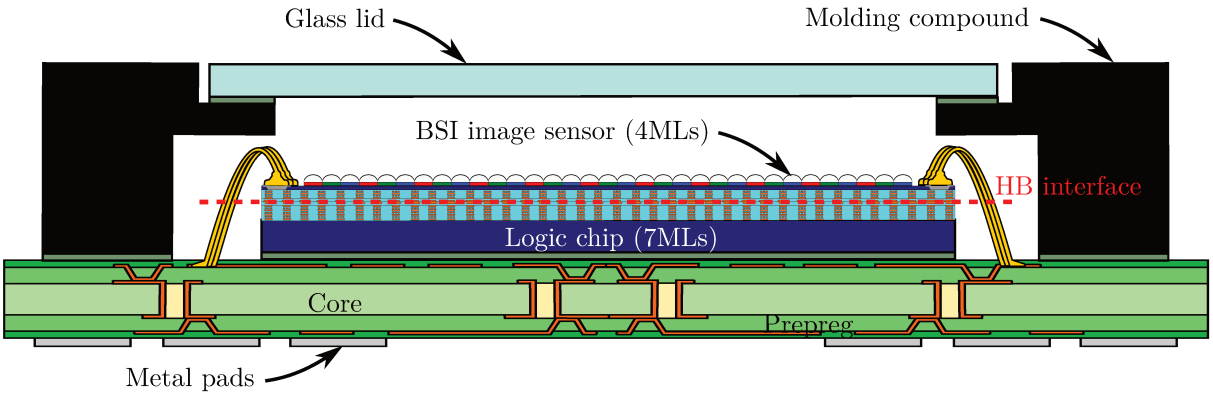


FIGURE I.30: Schematical description of the 3D image sensor studied in this work.

In this work, the investigated test chip consists of a 14Mpixel BSI image sensor with a 1.5 µm pixel pitch stacked by Cu/SiO_2 hybrid bonding onto a 40 nm logic process chip. The 4-ML top die contains exclusively the array of pixels, while the 7-ML bottom die integrates (among others) the pixel readout circuit and image signal processing [83]. This 3D CMOS image sensor is assembled in an organic LGA cavity package by gold wire bonding (Figure I.30). The focus of this study is on the mechanical robustness of such a 3D integrated device during its fabrication. Mechanical stresses building up in the image sensor during chip processing and assembly into a package are investigated. The integrity and mechanical interactions between the structures linking the different system components, namely the hybrid bonding interface between the imager/logic chips and wire bonds between the imager and the package, are also assessed. Compared to previous studies, a broader scope of process steps is therefore investigated, encompassing image sensor chip fabrication and assembly into a package in addition to chip stacking by Cu/SiO_2 hybrid bonding. This work focuses on several key structures in the considered device, such as hybrid bonding pads and IO pads, aiming to address a number of potential mechanical issues described below. For each item, combined numerical and experimental investigations are carried out, relying on finite element analysis and morphological (e.g., 3D-FIB/SEM tomography and AFM[1]), mechanical or electrical characterization techniques.

This study is divided up into three main items:

- **Chapter 2: Experimental and numerical investigations on the Cu/SiO2 hybrid bonding process for chip stacking**
  Extensive surface preparation by CMP is required before the hybrid bonding process to meet the stringent planarity requirements for successful molecular bonding. However, for hybrid bonding both copper and oxide are present on the surfaces to bond. Due to different

[1] Atomic force microscope (AFM)

material removal rates, the polishing process thus typically results in a slight recess of Cu surfaces compared to the surrounding dielectric, with a depth in the nanometric range.

Although it was previously shown that this Cu recess can be expected to be compensated, if sufficiently low, thanks to CTE mismatch between Cu and SiO_2 [84], the maximum acceptable surface topography for successful bonding is not known. In addition, as the hybrid interconnect pitch is targeted to further decrease in the upcoming years, the influence on the bonding behavior of several important design and process parameters is not fully understood. In this chapter, we focus on the hybrid bonding interconnect layer, taken separately, aiming to gain further insight into these aspects.

First, the topography of the surfaces to be bonded is measured after planarization at the interconnect scale by AFM for different sets of CMP conditions, resulting in various observed Cu recess depths. Then, wafers having undergone the same CMP process are bonded and the Cu/Cu interface morphology after bonding and annealing is characterized using SEM/TEM cross-sections or 3D-FIB/SEM, aiming to qualitatively assess interface closure or a possible presence of voids or cavities for each case.

Aiming to correlate these experimental observations, a finite element model for Cu/Cu interface closure during bonding is then proposed, building on earlier work by Beilliard [85]. The influence of various design and process parameters on the bonding behavior is investigated (e.g. interconnect aspect ratio and surface fraction, interconnect overlay or annealing temperature), enabling to secure integration from a thermo-mechanical viewpoint. The role played by (possible) Cu plasticity in assisting or hindering molecular bonding is also assessed.

- **Chapter 3: Study of mechanical failures induced in interconnect layers by the wire bonding process on backside pads**
  Due to the specific integration strategy of the studied 3D BSI image sensor, package assembly is carried out by wire bonding on peripheral IO pads. As a result of the BSI configuration, the interconnect stack is flipped and the wirebond pads are located on the backside of the imager die, just above the local interconnection levels containing thin dielectric layers with fine metal lines. This proximity of local interconnects to the region where wirebond is formed may lead to cracks or delamination in the BEoL multi-layer stack. Indeed, wire-bonding is one of the most critical fabrication steps in terms of mechanical robustness, due to the large stresses generated in the pad structure during the thermosonic bonding process. In this chapter, we will investigate the specificity of such an integration scheme.

  First, the mechanical integrity of several pad structures after wire bonding is compared experimentally. These wirebond pads include different layouts for the metal lines in the underlying interconnect stack (BEoL) and several stacks for the passivation capping (fBEoL). A comparison between these different pad structures is made by failure inspection after thermosonic bonding to detect possible cracks or delamination.

Then, a multi-scale finite element analysis of a standardized wirebond qualification test, namely the wire pull test, is carried out. A stress-based criterion is proposed to assess the mechanical robustness of the different pad architectures, aiming to correlate the trends observed in the experiments. This model is then used to investigate the influence of various design choices on the mechanical robustness, and thereby propose guidelines to decrease the number of pad failure occurrences.

- **Chapter 4: Process-induced Thermomechanical Stresses in a 3D Integrated Circuit assembled by Hybrid Stacking**
  Semiconductor device performance is known to be influenced when subjected to external mechanical load, due to the *strain-induced bandgap narrowing* phenomenon, resulting in a variation of the free carrier mobility in strained Si doped regions. Similarly, the electro-optical performance of image sensors is stress-sensitive [13]. Although no significant issues were reported in that respect for the studied BSI image sensor in its "planar" version, a different stress state can be expected for a 3D BSI image sensor stacked by hybrid bonding, which has a much thicker interconnection stack and includes a hybrid bonding interconnect layer. The stress state for this relatively new integration strategy in manufacturing is not known. It is therefore necessary to have a method enabling *in situ* measurement of the mechanical stress in the active region of the chip during the processing and packaging steps of the 3D image sensor.

  A methodology based on Si piezoresistive stress sensors is explored. For the first time, these sensors are implemented in a 3D BSI image sensor stacked by hybrid bonding. Different locations on the chip are investigated, namely the center and corner regions. By measuring resistance variations throughout the process sequence, the corresponding stress variations can be derived, provided that the piezoresistive coefficients associated with the different orientations have been determined. This calibration procedure is carried out using a previously developed in-house instrumented four-point bending fixture [86]. The stress variations are then evaluated throughout both the chip fabrication and packaging processes, and the result compared with corresponding finite element analyses. Finally, the limitations and applicability of this method for the current technology are discussed, and recommendations are proposed for implementation in future devices.

In a last part, the main achievements of this work, to our knowledge the first thermomechanical stress analysis for a 3D BSI image sensor stacked by hybrid bonding, are presented and further developments suggested.

Year	Stacked layers	Stacking technology	Interconnect pitch	Company/Institution	Reference
2001	64 × 64 12 µm 2MLs APS 0.8 µm 2MLs CMOS ADC	Epoxy adhesive bonding Al dual-TSV	?	MIT	[63]
2005	1Mpixel 8 µm APS 0.35 µm 3MLs CMOS readout	Direct bonding W dual-TSV	8 µm	MIT	[64]
2007	97 × 97 16 µm CIS 0.35 µm CMOS	W dual-TSV	?	Yale Univ.	[66]
2009	1Mpixel 8 µm 3MLs APS 3MLs CMOS readout 5-tier multichip stack	Direct bonding Dual-TSV Au stud bumps	8 µm 500 µm	MIT	[65]
<2012	1.5Mpixel 1.25 µm CIS ?	Ni/SiO_2 HB	?	Kodak	[56]
2013	8Mpixel 4MLs CIS 65 nm 7MLs CMOS logic	Direct bonding Dual-TSV	?	Sony	[67]
2013	704 × 512 4.3 µm 6MLs CIS 0.18 µm 6MLs CMOS process	µpillars	8.6 µm	Olympus	[68]
2013	512 × 832 17 µm SOI X-ray detectors 0.2 µm 5MLsCMOS	µbumps	5 µm	KEK T-Micro	[87]
2014	1.1 µm CIS ASIC	Dual-TSV	?	TSMC	[71]
2014	2056 × 1600 1.1 µm CIS 45 nm ASIC	Dual-TSV or "direct connection"	?	TSMC	[88]
2014	4096 80 µm X-ray APSs 2-tier F2F 0.13 µm 5MLs ASICs	Cu/SiO_2 HB via-middle W-TSVs	4 µm	Fermilab	[75]
2014	64 80 µm APSs 200 µm 3MLs ADC	Pixel-level Au/SiO_2 HB	?	NHK Univ. of Tokyo	[76]
2015	20Mpixel 1.43 µm 4MLs CIS 65 nm 7MLs CMOS logic	?	?	Sony	[89]
2015	16Mpixel 3.8 µm 6MLs CIS 0.13 µm 6MLs CMOS logic	µpillars	7.6 µm	Olympus	[69]
2015	4MLs CIS CMOS logic	Parallel inductive coupling	?	Hokkaido Univ. Keio Univ.	[90]
2015	8Mpixel 1.1 µm CMOS logic	Dual-TSV	?	TSMC	[72]
2015	5Mpixel 1.1 µm CIS on ASIC 16Mpixel 1.12 µm CIS on ASIC 16Mpixel 1 µm CIS on ASIC	?	?	OmniVision	[73]
2015	20Mpixel 1.43 µm 4MLs CIS 65 nm 7MLs CMOS logic	?	?	Sony	[91]
2015	4224 × 240 3.8 µm 6MLs CIS 0.18 µm 6MLs CMOS logic	?	?	Olympus	[70]
2016	8Mpixel 1.12 µm CIS 65 nm CMOS logic	Cu/SiO_2 HB	?	Toshiba	[79]
2016	33Mpixel 4MLs CIS 65 nm 5MLs CMOS logic	"Direct connection"	?	TSMC	[92]
2016	33Mpixel 1.1 µm 4MLs CIS 65 nm 5MLs ASIC	"Hybrid-stacking"	4.4 µm	NHK Brookman Technology TSMC Shizuoka Univ.	[93]
2016	Two 65 nm 9MLs CMOS logic on one 8.3Mpixel 5MLs CIS	µbumps	?	Sony	[74]
2016	6.6 µm 4MLs CIS 65 nm 5MLs CMOS logic	"Direct connection"	6.6 µm	Tohoku Univ. TSMC	[94]
2016	SPAD CIS 40 nm CMOS logic	Cu/SiO_2 HB	7.83 µm	STMicroelectronics Univ. of Edinburgh	[80]
2016	22.5Mpixel 1.1 µm CIS 40 nm CMOS logic	Cu/SiO_2 HB	4 µm	Sony	[77]
2017	20Mpixel 1.22 µm 5MLs CIS 30 nm 3MLs DRAM 40 nm 6MLs CMOS logic	via-last TSVs	?	Sony	[95]
2017	1.27Mpixel 3.5 µm 4MLs CIS 40 nm 7MLs CMOS logic	?	?	Sony Univ. of Tokyo	[96]
2017	8Mpixel 1.1 µm CIS 40 nm CMOS logic	?	?	Qualcomm TSMC	[97]
2017	4.1Mpixel 3.8 µm 4MLs CIS 55 nm 7MLs CMOS logic	?	?	Sony	[98]
2017	8Mpixel 1.1 µm CIS 40 nm CMOS logic	?	?	Qualcomm TSMC	[97]
2017	SPAD CIS 65 nm CMOS logic	F2F "connection layer"	?	EPFL TSMC	[99]
2017	16Mpixel 1 µm CIS on ASIC	Cu/SiO_2 HB	?	OmniVision	[78]
2018	3.9Mpixel 1.5 µm 4MLs CIS 40 nm 6MLs CMOS logic	?	?	Sony	[100]
2018	13.5Mpixel 1.1 µm 4MLs CIS 65 nm 4MLs CMOS logic	"F2F bonding"	?	TSMC	[101]
2018	SPAD CIS 65 nm 5MLs CMOS logic	"Multiple 3D connections per SPAD"	?	Delft Univ. of Technology EPFL TSMC	[102]
2018	8Mpixel 1.5 µm CIS on ASIC	"Stacking technology"	1.5 µm	OmniVision	[103]
2018	14Mpixel 1.5 µm 4MLs CIS 7MLs CMOS logic	Cu/SiO_2 HB	8.8 µm 1.44 µm	STMicroelectronics Univ. of Bordeaux CEA-Leti	[83] This work

TABLE I.2: Overview of 3D stacking for BSI image sensor applications.

References

[1] J. Bardeen, "Research leading to point-contact transistor," Science, vol. 126, no. 3264, pp. 105–112, 1957.

[2] J. S. Kilby, "Invention of the integrated circuit," IEEE Transactions on electron devices, vol. 23, no. 7, pp. 648–654, 1976.

[3] G. E. Moore, "Progress in digital integrated electronics," IEDM Technical Digest, pp. 11–13, 1975.

[4] G. E. Moore, "Cramming more component onto integrated circuits," Electronics, vol. 38, no. 8, pp. 114–117, 1965.

[5] P. A. Gargini, "Roadmap evolution: from NTRS to ITRS, from ITRS 2.0 to IRDS," in International Conference on Extreme Ultraviolet Lithography 2017, vol. 10450, p. 1045018, International Society for Optics and Photonics, 2017.

[6] "International Technology Roadmap for Semiconductors (2005 Edition)," 2005. Available from: http://www.itrs2.net.

[7] "International Technology Roadmap for Semiconductors (2007 Edition)," 2007. Available from: http://www.itrs2.net.

[8] "International Technology Roadmap for Semiconductors (2009 Edition)," 2009. Available from: http://www.itrs2.net.

[9] "International Technology Roadmap for Semiconductors (2011 Edition)," 2011. Available from: http://www.itrs2.net.

[10] "International Technology Roadmap for Semiconductors (2013 Edition)," 2013. Available from: http://www.itrs2.net.

[11] G. Spierings, "Wet chemical etching of silicate glasses in hydrofluoric acid based solutions," Journal of Materials science, vol. 28, no. 23, pp. 6261–6273, 1993.

[12] J. A. Bondur, "Dry process technology (reactive ion etching)," Journal of Vacuum Science and Technology, vol. 13, no. 5, pp. 1023–1029, 1976.

[13] G. C. Dogiamis, B. J. Hosticka, and A. Grabmaier, "Investigations on an ultra-thin bendable monolithic si cmos image sensor," IEEE Sensors Journal, vol. 13, no. 10, pp. 3892–3900, 2013.

[14] R. Smeltzer, "Hole trap creation in sio2 by phosphorus ion penetration of polycrystalline silicon," IEEE Transactions on Nuclear Science, vol. 29, no. 6, pp. 1467–1470, 1982.

[15] P. Ireland, "High aspect ratio contacts: A review of the current tungsten plug process," Thin Solid Films, vol. 304, no. 1-2, pp. 1–12, 1997.

[16] Z. Suo, Q. Ma, and W. Meyer, "Immortal interconnects—prevent cracking and limit void size," MRS Online Proceedings Library Archive, vol. 563, 1999.

[17] S. M. Alam, F. L. Wei, C. L. Gan, C. V. Thompson, and D. E. Troxel, "Electromigration reliability comparison of cu and al interconnects," in Sixth international symposium on quality electronic design (isqed'05), pp. 303–308, IEEE, 2005.

[18] D. Edelstein, J. Heidenreich, R. Goldblatt, W. Cote, C. Uzoh, N. Lustig, P. Roper, T. McDevitt, W. Motsiff, A. Simon, et al., "Full copper wiring in a sub-0.25/spl mu/m cmos ulsi technology," in Electron Devices Meeting, 1997. IEDM'97. Technical Digest., International, pp. 773–776, IEEE, 1997.

[19] T. M. Moore, C. D. Hartfield, J. M. Anthony, B. T. Ahlburn, P. S. Ho, and M. R. Miller, "Mechanical characterization of low-k dielectric materials," in AIP conference proceedings, vol. 550, pp. 431–439, AIP, 2001.

[20] M. Damayanti, Z. Gan, T. Sritharan, S. Mhaisalkar, A. Naman, J. Widodo, and H. Tan, "Effect of porosity and adhesion promoter layer on adhesion energy of nanoporous inorganic low-κ," Thin solid films, vol. 504, no. 1-2, pp. 213–217, 2006.

[21] P. Coudrain, Contribution au développement d'une technologie d'intégration tridimensionnelle pour les capteurs d'images CMOS à pixels actifs. PhD thesis, Toulouse, ISAE, 2009.

[22] S. Gallois-Garreignot, Etude numérique et expérimentale des phénomènes de fissuration dans les architectures avancées de la microélectronique. PhD Thesis, Lyon, INSA, 2010.

[23] F. Roy, A. Tournier, H. Wehbe-Alause, F. Blanchet, P. Boulenc, F. Leverd, L. Favennec, C. Perrot, L. Pinzelli, M. Gatefait, et al., "Challenges in cmos-based images," physica status solidi (c), vol. 11, no. 1, pp. 50–56, 2014.

[24] G. G. Harman, Wire bonding in microelectronics. McGraw-Hill, 2010.

[25] B. Langenecker, "Effects of ultrasound on deformation characteristics of metals," IEEE transactions on sonics and ultrasonics, vol. 13, no. 1, pp. 1–8, 1966.

[26] R. Roucou, Analyse expérimentale et numérique des défaillances mécaniques locales induites dans les interconnexions par les tests paramétriques et les assemblages: optimisation des procédés et des architectures des plots de connexion. PhD thesis, 2010.

[27] C. Cheng, H. Hsiao, S. Chu, Y. Shieh, C. Sun, and C. Peng, "Low cost silver alloy wire bonding with excellent reliability performance," in Electronic Components and Technology Conference (ECTC), 2013 IEEE 63rd, pp. 1569–1573, IEEE, 2013.

[28] C.-C. Lee and L. M. Higgins, "Challenges of cu wire bonding on low-k/cu wafers with boa structures," in Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th, pp. 342–349, IEEE, 2010.

[29] Z. Zhong, "Overview of wire bonding using copper wire or insulated wire," Microelectronics Reliability, vol. 51, no. 1, pp. 4–12, 2011.

[30] K. DeHaven and J. Dietz, "Controlled collapse chip connection (c4)-an enabling technology," in Electronic Components and Technology Conference, 1994. Proceedings., 44th, pp. 1–6, IEEE, 1994.

[31] T. Wang, F. Tung, L. Foo, and V. Dutta, "Studies on a novel flip-chip interconnect structure. pillar bump," in Electronic Components and Technology Conference, 2001. Proceedings., 51st, pp. 945–949, IEEE, 2001.

[32] J.-W. Kim, D.-G. Kim, and S.-B. Jung, "Evaluation of displacement rate effect in shear test of sn–3ag–0.5 cu solder bump for flip chip application," Microelectronics Reliability, vol. 46, no. 2-4, pp. 535–542, 2006.

[33] "Wire Bond Shear Test," standard, JEDEC Solid State Technology Association, May 2017.

[34] "Solder Ball Shear," standard, JEDEC Solid State Technology Association, May 2014.

[35] "Wire Bond Pull Strength," standard, The Institute for Interconnecting and Packaging Electronic Circuits.

[36] "Board Level Drop Test Method of Components for Handheld Electronic products," standard, JEDEC Solid State Technology Association, November 2003.

[37] "Temperature Cycling," standard, JEDEC Solid State Technology Association, October 2014.

[38] "Cycled Temperature Humidity Bias Life Test," standard, JEDEC Solid State Technology Association, July 2013.

[39] "Steady-State Temperature-Humidity Bias Life Test," standard, JEDEC Solid State Technology Association, July 2015.

[40] "Highly Accelerated Temperature and Humidity Stress Test (HAST)," standard, JEDEC Solid State Technology Association, July 2015.

[41] "High Temperature Storage Life," standard, JEDEC Solid State Technology Association, October 2015.

[42] M. M. Waldrop, "The chips are down for moore's law," Nature News, vol. 530, no. 7589, p. 144, 2016.

[43] M. T. Bohr, "Interconnect scaling-the real limiter to high performance ulsi," in Electron Devices Meeting, 1995. IEDM'95., International, pp. 241–244, IEEE, 1995.

[44] M. M. Waldrop, "More than moore," Nature, vol. 530, no. 7589, pp. 144–148, 2016.

[45] "National Technology Roadmap for Semiconductors (1997 Edition)," 1997. Available from: http://www.rennes.supelec.fr/ren/perso/gtourneu/enseignement/roadmap97.pdf.

[46] K. L. Tai, "System-in-package (sip): challenges and opportunities," in Design Automation Conference, 2000. Proceedings of the ASP-DAC 2000. Asia and South Pacific, pp. 191–196, IEEE, 2000.

[47] K. M. Brown, "System in package" the rebirth of sip"," in Custom Integrated Circuits Conference, 2004. Proceedings of the IEEE 2004, pp. 681–686, IEEE, 2004.

[48] B. Vianne, Intégration d'un interposeur actif silicium pour l'élaboration de circuits électroniques complexes. PhD thesis, Université Aix-Marseille, 2016. In French.

[49] P. Garrou, C. Bower, and P. Ramm, Handbook of 3D Integration, Volume 1: Technology and Applications of 3D Integrated Circuits. John Wiley & Sons, 2011.

[50] J. W. Joyner, P. Zarkesh-Ha, and J. D. Meindl, "A stochastic global net-length distribution for a three-dimensional system-on-a-chip (3d-soc)," in ASIC/SOC Conference, 2001. Proceedings. 14th Annual IEEE International, pp. 147–151, IEEE, 2001.

[51] J. P. Gambino, S. A. Adderly, and J. U. Knickerbocker, "An overview of through-silicon-via technology and manufacturing challenges," Microelectronic Engineering, vol. 135, pp. 73–106, 2015.

[52] P. Garrou, J. J.-Q. Lu, and P. Ramm, "Three-dimensional integration," in Handbook of Wafer Bonding (P. Ramm, J. J. Q. Lu, and M. M. V. Taklo, eds.), ch. 15, pp. 261–278, Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, Germany, 2012.

[53] C.-T. Ko and K.-N. Chen, "Wafer-level bonding/stacking technology for 3d integration," Microelectronics reliability, vol. 50, no. 4, pp. 481–488, 2010.

[54] Y. Xie, "Processor architecture design using 3d integration technology," in VLSI Design, 2010. VLSID'10. 23rd International Conference on, pp. 446–451, IEEE, 2010.

[55] P. Garrou, J. J.-Q. Lu, and P. Ramm, "Cu/sio$_2$ hybrid bonding," in Handbook of Wafer Bonding (P. Ramm, J. J. Q. Lu, and M. M. V. Taklo, eds.), ch. 12, pp. 261–278, Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, Germany, 2012.

[56] P. Enquist, "Metal/silicon oxide hybrid bonding," in Handbook of Wafer Bonding (P. Ramm, J. J. Q. Lu, and M. M. V. Taklo, eds.), ch. 13, pp. 261–278, Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, Germany, 2012.

[57] E. Beyne, S.-W. Kim, L. Peng, N. Heylen, J. De Messemaeker, O. O. Okudur, A. Phommahaxay, T.-G. Kim, M. Stucchi, D. Velenis, et al., "Scalable, sub 2μm pitch, cu/sicn to cu/sicn hybrid wafer-to-wafer bonding technology," in Electron Devices Meeting (IEDM), 2017 IEEE International, pp. 32–4, IEEE, 2017.

[58] A. Jourdain, S. Stoukatch, P. De Moor, W. Ruythooren, S. Pargfrieder, B. Swinnen, and E. Beyne, "Simultaneous cu-cu and compliant dielectric bonding for 3d stacking of ics," in 2007 IEEE International Interconnect Technology Conferencee, pp. 207–209, IEEE, 2007.

[59] J. McMahon, E. Chan, S. Lee, R. Gutmann, and J.-Q. Lu, "Bonding interfaces in wafer-level metal/adhesive bonded 3d integration," in 2008 58th Electronic Components and Technology Conference, pp. 871–878, IEEE, 2008.

[60] F. Niklaus, P. Enoksson, E. Kälvesten, and G. Stemme, "Low-temperature full wafer adhesive bonding," Journal of Micromechanics and Microengineering, vol. 11, no. 2, p. 100, 2001.

[61] P. Gueguen, L. Di Cioccio, P. Gergaud, M. Rivoire, D. Scevola, M. Zussy, A. M. Charvet, L. Bally, D. Lafond, and L. Clavelier, "Copper direct-bonding characterization and its interests for 3D integration," Journal of The Electrochemical Society, vol. 156, no. 10, pp. H772–H776, 2009.

[62] B. Rebhan, M. Bernauer, T. Wagenleitner, M. Heilig, F. Kurz, S. Lhostis, E. Deloffre, A. Jouve, V. Balan, and L. Chitu, "$<$ 200 nm wafer-to-wafer overlay accuracy in wafer level cu/sio2 hybrid bonding for bsi cis," in Electronics Packaging and Technology Conference (EPTC), 2015 IEEE 17th, pp. 1–4, IEEE, 2015.

[63] J. Burns, L. McIlrath, C. Keast, C. Lewis, A. Loomis, K. Warner, and P. Wyatt, "Three-dimensional integrated circuits for low-power, high-bandwidth systems on a chip," in Solid-State Circuits Conference, 2001. Digest of Technical Papers. ISSCC. 2001 IEEE International, pp. 268–269, IEEE, 2001.

[64] V. Suntharalingam, R. Berger, J. A. Burns, C. K. Chen, C. Keast, J. Knecht, R. Lambert, K. Newcomb, D. O'Mara, D. Rathman, and others, "Megapixel CMOS image sensor fabricated in three-dimensional integrated circuit technology," in Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International, pp. 356–357, IEEE, 2005.

[65] V. Suntharalingam, R. Berger, S. Clark, J. Knecht, A. Messier, K. Newcomb, D. Rathman, R. Slattery, A. Soares, C. Stevenson, and others, "A 4-side tileable back illuminated 3d-integrated Mpixel CMOS image sensor," Institute of Electrical and Electronics Engineers, 2009.

[66] E. Culurciello and P. Weerakoon, "Three-dimensional photodetectors in 3-D silicon-on-insulator technology," IEEE Electron Device Letters, vol. 28, no. 2, pp. 117–119, 2007.

[67] S. Sukegawa, T. Umebayashi, T. Nakajima, H. Kawanobe, K. Koseki, I. Hirota, T. Haruta, M. Kasai, K. Fukumoto, T. Wakano, and others, "A 1/4-inch 8mpixel back-illuminated stacked CMOS image sensor," in Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International, pp. 484–485, IEEE, 2013.

[68] J. Aoki, Y. Takemoto, K. Kobayashi, N. Sakaguchi, M. Tsukimura, N. Takazawa, H. Kato, T. Kondo, H. Saito, Y. Gomi, and others, "A rolling-shutter distortion-free 3d stacked image sensor with- 160db parasitic light sensitivity in-pixel storage node," in Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International, pp. 482–483, IEEE, 2013.

[69] T. Kondo, Y. Takemoto, K. Kobayashi, M. Tsukimura, N. Takazawa, H. Kato, S. Suzuki, J. Aoki, H. Saito, Y. Gomi, and others, "A 3d stacked CMOS image sensor with 16mpixel global-shutter mode and 2mpixel 10000fps mode using 4 million interconnections," in VLSI Circuits (VLSI Circuits), 2015 Symposium on, pp. C90–C91, IEEE, 2015.

[70] Y. Takemoto, K. Kobayashi, M. Tsukimura, N. Takazawa, H. Kato, S. Suzuki, J. Aoki, T. Kondo, H. Saito, Y. Gomi, and others, "Multi-storied photodiode CMOS image sensor for multiband imaging with 3d technology," in Electron Devices Meeting (IEDM), 2015 IEEE International, pp. 30–1, IEEE, 2015.

[71] J. Liu, D. Yaung, J. Sze, C. Wang, G. Hung, C. Wang, T. Hsu, R. Lin, T. Wang, W. Wang, and others, "Advanced 1.1 um pixel CMOS image sensor with 3d stacked architecture," in VLSI Technology (VLSI-Technology): Digest of Technical Papers, 2014 Symposium on, pp. 1–2, IEEE, 2014.

[72] S.-F. Yeh, K.-Y. Chou, H.-Y. Tu, C. Y.-P. Chao, and F.-L. Hsueh, "A 0.66 e- rms temporal-readout-noise 3d-stacked CMOS image sensor with conditional correlated multiple sampling (CCMS) technique," in VLSI Circuits (VLSI Circuits), 2015 Symposium on, IEEE, 2015.

[73] V. Venezia, C. Shih, W. Yang, B. Zhang, and H. Rhodes, "Stack chip technology: a new direction for CMOS imagers," in IISW conference, 2015.

[74] Y. Oike, K. Akiyama, L. D. Hung, W. Niitsuma, A. Kato, M. Sato, Y. Kato, W. Nakamura, H. Shiroshita, Y. Sakano, and others, "An 8.3 M-pixel 480fps global-shutter CMOS image sensor with gain-adaptive column ADCs and 2-on-1 stacked device structure," in VLSI Circuits (VLSI-Circuits), 2016 IEEE Symposium on, pp. 1–2, IEEE, 2016.

[75] R. Yarema, G. Deptuch, and R. Lipton, "Recent results for 3d pixel integrated circuits using copper-copper and oxide-oxide bonding," (Lake Starnberg, Germany), Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States), 2014.

[76] M. Goto, K. Hagiwara, Y. Iguchi, H. Ohtake, T. Saraya, M. Kobayashi, E. Higurashi, H. Toshiyoshi, and T. Hiramoto, "Three-dimensional integrated CMOS image sensors with pixel-parallel A/D converters fabricated by direct bonding of SOI layers," in Electron Devices Meeting (IEDM), 2014 IEEE International, pp. 2–4, IEEE, 2014.

[77] Y. Kagawa, N. Fujii, K. Aoyagi, Y. Kobayashi, S. Nishi, N. Todaka, S. Takeshita, J. Taura, H. Takahashi, Y. Nishimura, and others, "Novel stacked CMOS image sensor with advanced Cu2cu hybrid bonding," in Electron Devices Meeting (IEDM), 2016 IEEE International, pp. 4–8, IEEE, 2016.

[78] V. Venezia, C. Shih, W. Yang, Y. Zang, Z. Lin, L. Grant, and H. Rhodes, "1.0 μm pixel improvements with hybrid bond stacking technology," in Proceedings of the 2017 International Image Sensor Workshop, Hiroshima, Japan, vol. 30, pp. 8–11, 2017.

[79] K. Shiraishi, Y. Shinozuka, T. Yamashita, K. Sugiura, N. Watanabe, R. Okamoto, T. Ashitani, M. Furuta, and T. Itakura, "A 1.2e- temporal noise 3d-stacked CMOS image sensor with comparator-based multiple-sampling PGA," in 2016 IEEE International Solid-State Circuits Conference (ISSCC), pp. 122–123, Jan. 2016.

[80] T. Al Abbas, N. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3d-stacked CMOS technology," in Electron Devices Meeting (IEDM), 2016 IEEE International, pp. 8–1, IEEE, 2016.

[81] M. Taibi, Intégration 3D haute densité: comportement et fiabilité électrique d'interconnexions métalliques réalisées par collage direct. PhD thesis, Université de Grenoble, 2012.

[82] Y. Beilliard, Etude de l'intégration du collage direct cuivre/oxyde pour l'élaboration d'une architecture 3D-SIC. PhD thesis, Université Grenoble-Alpes, 2015.

[83] J. Jourdon, S. Lhostis, S. Moreau, J. Chossat, M. Arnoux, C. Sart, Y. Henrion, P. Lamontagne, L. Arnaud, N. Bresson, and others, "Hybrid bonding for 3d stacked image sensors: impact of pitch shrinkage on interconnect robustness," in 2018 IEEE International Electron Devices Meeting (IEDM), pp. 7–3, IEEE, 2018.

[84] L. Di Cioccio, P. Gueguen, R. Taibi, D. Landru, G. Gaudin, C. Chappaz, F. Rieutord, F. de Crecy, I. Radu, L. Chapelon, et al., "An overview of patterned metal/dielectric surface bonding: mechanism, alignment and characterization," Journal of The Electrochemical Society, vol. 158, no. 6, pp. P81–P86, 2011.

[85] Y. Beilliard, R. Estevez, G. Parry, P. McGarry, L. Di Cioccio, and P. Coudrain, "Thermomechanical finite element modeling of Cu–SiO_2 direct hybrid bonding with a dishing effect

on Cu surfaces," International Journal of Solids and Structures, vol. 117, pp. 208–220, 2017.

[86] K. A. Ewuame, V. Fiori, K. Inal, P.-O. Bouchard, S. Gallois-Garreignot, S. Lionti, C. Tavernier, and H. Jaouen, "Investigations of thermomechanical stress induced by TSV-middle (through-silicon via) in 3-D ICs by means of CMOS sensors and finite-element method," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 5, no. 8, pp. 1085–1092, 2015.

[87] Y. Arai and M. Motoyoshi, "Application of 3d stacking technology to SOI radiation image sensor," in Electrical Design of Advanced Packaging and Systems Symposium (EDAPS), 2013 IEEE, pp. 5–8, IEEE, 2013.

[88] C. C.-M. Liu, C.-H. Chang, H.-Y. Tu, C. Y.-P. Chao, F.-L. Hsueh, S.-Y. Chen, V. Hsu, J.-C. Liu, D.-N. Yaung, and S.-G. Wuu, "A peripheral switchable 3d stacked CMOS image sensor," in VLSI Circuits Digest of Technical Papers, 2014 Symposium on, pp. 1–2, IEEE, 2014.

[89] A. Suzuki, N. Shimamura, T. Kainuma, N. Kawazu, C. Okada, T. Oka, K. Koiso, A. Masagaki, Y. Yagasaki, S. Gonoi, and others, "A 1/1.7-inch 20mpixel Back-illuminated stacked CMOS image sensor for new imaging applications," in Solid-State Circuits Conference-(ISSCC), 2015 IEEE International, pp. 1–3, IEEE, 2015.

[90] M. Ikebe, D. Uchida, Y. Take, M. Someya, S. Chikuda, K. Matsuyama, T. Asai, T. Kuroda, and M. Motomura, "Image sensor/digital logic 3d stacked module featuring inductive coupling channels for high speed/low-noise image transfer," in VLSI Circuits (VLSI Circuits), 2015 Symposium on, pp. C82–C83, IEEE, 2015.

[91] H. Wakabayashi, A. Suzuki, N. Shimamura, T. Kainuma, K. Koiso, A. Masagaki, Y. Yagasaki, S. Gonoi, M. Mizuno, T. Sugioka, and others, "A 1/1.7-inch 20mpixel Back-illuminated Stacked CMOS Image Sensor with parallel multiple sampling," Proc. 2015 IISW, pp. 44–47, 2015.

[92] C. C. Liu, M. M. Mhala, C. Chang, H. Tu, P. Chou, C. Chao, and F. Hsueh, "A 1.5v 33mpixel 3d-stacked CMOS image sensor with negative substrate bias," in 2016 IEEE International Solid-State Circuits Conference (ISSCC), pp. 124–125, Jan. 2016.

[93] T. Arai, T. Yasue, K. Kitamura, H. Shimamoto, T. Kosugi, S. Jun, S. Aoyama, M. Hsu, Y. Yamashita, H. Sumi, and S. Kawahito, "A 1.1µm 33mpixel 240fps 3d-stacked CMOS image sensor with 3-stage cyclic-based analog-to-digital converters," in 2016 IEEE International Solid-State Circuits Conference (ISSCC), pp. 126–128, Jan. 2016.

[94] H. Sugo, S. Wakashima, R. Kuroda, Y. Yamashita, H. Sumi, T.-J. Wang, P.-S. Chou, M.-C. Hsu, and S. Sugawa, "A dead-time free global shutter CMOS image sensor with in-pixel LOFIC and ADC using pixel-wise connections," in VLSI Circuits (VLSI-Circuits), 2016 IEEE Symposium on, pp. 1–2, IEEE, 2016.

[95] T. Haruta, T. Nakajima, J. Hashizume, T. Umebayashi, H. Takahashi, K. Taniguchi, M. Kuroda, H. Sumihiro, K. Enoki, T. Yamasaki, K. Ikezawa, A. Kitahara, M. Zen, M. Oyama, H. Koga, H. Tsugawa, T. Ogita, T. Nagano, S. Takano, and T. Nomoto, "A 1/2.3inch 20mpixel 3-layer stacked CMOS Image Sensor with DRAM," in 2017 IEEE International Solid-State Circuits Conference (ISSCC), pp. 76–77, Feb. 2017.

[96] T. Yamazaki, H. Katayama, S. Uehara, A. Nose, M. Kobayashi, S. Shida, M. Odahara, K. Takamiya, Y. Hisamatsu, S. Matsumoto, L. Miyashita, Y. Watanabe, T. Izawa, Y. Muramatsu, and M. Ishikawa, "A 1ms high-speed vision chip with 3d-stacked 140gops column-parallel PEs for spatio-temporal image processing," in 2017 IEEE International Solid-State Circuits Conference (ISSCC), pp. 82–83, Feb. 2017.

[97] B.-C. Hseih, S. Khawam, N. Ioannis, M. Muir, K. Le, H. Siddiqui, S. Goma, R. Lin, C.-H. Chang, C. Liu, and others, "A 3d Stacked Programmable Image Processing Engine in a 40 nm Logic Process with a Detector Array in a 45nm CMOS Image Sensor Technologies," in Proceedings of the 2017 International Image Sensor Workshop, Hiroshima, Japan, vol. 30, pp. 4–7, 2017.

[98] T. Takahashi, Y. Kaji, Y. Tsukuda, S. Futami, K. Hanzawa, T. Yamauchi, P. W. Wong, F. Brady, P. Holden, T. Ayers, and others, "A 4.1 Mpix 280fps stacked CMOS image sensor with array-parallel ADC architecture for region control," in VLSI Circuits, 2017 Symposium on, pp. C244–C245, IEEE, 2017.

[99] M.-J. Lee, A. Ximenes, P. Padmanabhan, T. Wang, K. Huang, Y. Yamashita, D. Yaung, and E. Charbon, "A back-illuminated 3d-stacked single-photon avalanche diode in 45nm CMOS technology," in Electron Devices Meeting (IEDM), 2017 IEEE International, pp. 16–6, IEEE, 2017.

[100] O. Kumagai, A. Niwa, K. Hanzawa, H. Kato, S. Futami, T. Ohyama, T. Imoto, M. Nakamizo, H. Murakami, T. Nishino, and others, "A 1/4-inch 3.9 Mpixel low-power event-driven back-illuminated stacked CMOS image sensor," in Solid-State Circuits Conference-(ISSCC), 2018 IEEE International, pp. 86–88, IEEE, 2018.

[101] P.-S. Chou, C.-H. Chang, M. M. Mhala, C. C.-M. Liu, C. Y.-P. Chao, C.-Y. Huang, H. Tu, T. Wu, S.-F. Yeh, S. Takahashi, and others, "A 1.1 μm-Pitch 13.5 Mpixel 3d-stacked CMOS image sensor featuring 230fps full-high-definition and 514fps high-definition videos by reading 2 or 3 rows simultaneously using a column-switching matrix," in Solid-State Circuits Conference-(ISSCC), 2018 IEEE International, pp. 88–90, IEEE, 2018.

[102] A. R. Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D. Yaung, and E. Charbon, "A 256 $\times$ 256 45/65nm 3d-stacked SPAD-based direct TOF image sensor for LiDAR applications with optical polar modulation for up to 18.6 dB interference suppression," in Solid-State Circuits Conference-(ISSCC), 2018 IEEE International, pp. 96–98, IEEE, 2018.

[103] V. C. Venezia, A. C.-W. Hsiung, K. Ai, X. Zhao, Z. Lin, D. Mao, A. Yazdani, E. A. Webster, and L. A. Grant, "1.5mu Dual Conversion Gain, Backside Illuminated Image Sensor Using Stacked Pixel Level Connections with 13ke-Full-Well Capacitance and 0.8 e-Noise," in 2018 IEEE International Electron Devices Meeting (IEDM), pp. 10–1, IEEE, 2018.

Chapter II

Cu/SiO_2 hybrid bonding process for chip stacking

II.1 Introduction

For the 3D integrated circuit considered in this work, with an image sensor stacked onto an image signal processing chip, an electrical connection must be achieved between the two active tiers. To do so, the method used here is hybrid bonding: a novel approach compared to more mature techniques in the semiconductor industry, such as wire-bonding, bumping or TSVs presented in Chapter I. With this method, an additional metal/dielectric patterned layer is processed on top of the interconnect stack in both chips to be bonded. Thanks to specific surface preparation and thermal treatment, these matching surfaces can then be bonded directly, without the need for any intermediate material such as glue or solder. Thus, a very compact interconnection layer can be obtained using hybrid bonding, making this technique particularly suitable for ultra-high-density 3D integration. However, for the considered application, involving Cu/SiO_2 patterned bonding surfaces, surface preparation is key for a successful bonding. Indeed, as will be shown in Section II.3, excellent conformality between the bonding surfaces is mandatory. Therefore, surface topography for the Cu/SiO_2 hybrid bonding layer must be carefully monitored to meet these stringent planarity requirements.

The aim of this chapter is to evaluate the influence of surface topography on the bonding quality, exploring various geometries or layouts for the Cu interconnects of the hybrid bonding layer. The focus is put on the Cu-Cu interface joining matching interconnect pairs on both sides of the assembly. In Section II.2, we start with an overview of the hybrid bonding process for chip stacking. Then, in Section II.3 some key experimental aspects of hybrid bonding are investigated. The influence of Cu interconnect geometry and layout on the resulting surface topography after planarization of the Cu/SiO_2 patterned surface is studied. In turn, the influence of the obtained surface topography on the bonding quality is examined. In Section II.4, finite element modeling of the hybrid bonding process at the interconnect scale is carried out, aiming to assess the

influence of additional process and design parameters that could not be addressed in Section II.3, such as interconnect aspect ratio and surface fraction, metal/dielectric material pair or interconnect misalignment.

II.2 Overview of hybrid bonding for chip stacking

In this section, the hybrid bonding process is described, and an overview of the underlying physical mechanisms is proposed. Then, the main advantages and drawbacks of the technique, as well as the current state-of-the-art are reviewed. Lastly, the present challenges for chip stacking by hybrid bonding, which motivate this study are discussed.

II.2.1 Process variants

The hybrid bonding process consists in the formation of a patterned metal-dielectric interface at the surface of the chips to be stacked, able to provide both electrical connection and mechanical attachment simultaneously, over the whole die surface. Depending on the kind of dielectric material used to insulate the hybrid bonding pads for a given application, the term "hybrid bonding" may actually refer to two very different classes of processes:

metal-polymer hybrid bonding: for those applications in which the hybrid bonding pads are immersed in organic dielectric. Typical examples of polymer dielectrics used for hybrid bonding include benzocyclobutene (BCB) or SU-8.

metal-oxide hybrid bonding: based on inorganic dielectric, as is typically the case in interconnect stack fabrication, with hybrid bonding pads embedded into SiO_2 or in some cases SiCN.

The latter method is used in the considered application to stack the image sensor chip onto the logic processing chip. Therefore, metal-polymer dielectric will not be considered in this work. Instead, the materials involved are TEOS-based PECVD undoped silicate glass (SiO_2) for the dielectric, and electroplated Cu for the hybrid bonding pads. This is mainly because for Cu/SiO_2 hybrid bonding, the patterned bonding layer can be fabricated in a straightforward manner, relying on a mature and ubiquitous technique in the semiconductor industry, namely the damascene process. As seen in Chapter I, this process is also used for interconnect stack fabrication just before hybrid bonding. The hybrid bonding process can therefore be integrated almost seamlessly into the existing manufacturing sequence.

The hybrid bonding process comes right at the end of interconnect stack fabrication (back-end-of-line process), before IO pad opening and IC surface passivation. The integrated circuits to be stacked are processed onto two separate wafers. After the last metallization level has been processed, for each wafer an additional metal/dielectric patterned layer is deposited on top of the interconnect stack to serve as a bonding layer (Figure II.1). In the considered application,

the metallization layout for this patterned bonding layer is the same in both integrated circuits to be bonded: a regular array of metal pads about 5 µm wide and 500 nm thick. While the majority of these *hybrid bonding pads* are electrically active interconnections, dedicated to signal transmission between the stacked chips, a non-negligible fraction (referred to as *dummies*) merely serves a mechanical purpose, aiming to provide better uniformity to the bonding surface and ensure a robust interface.

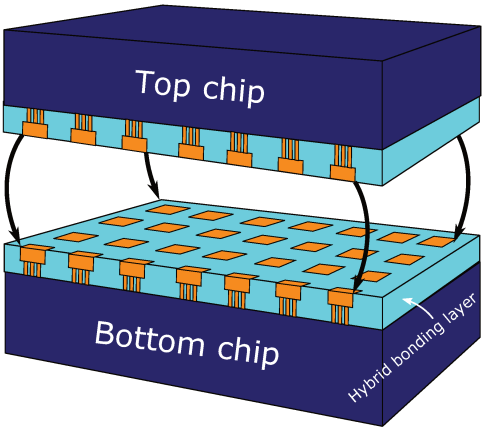


FIGURE II.1: Schematic principle of hybrid bonding layer interconnection between two stacked chips.

To obtain good adhesion during the different steps of the bonding process, surface preparation requirements are quite demanding. Prior to bonding, extensive chemical-mechanical polishing and surface cleaning is necessary, aiming to achieve three objectives: surface planarization, surface decontamination and surface activation.

II.2.2 Surface preparation

II.2.2.1 Planarization

The bonding surfaces must be mirror-polished[1] for adhesion to occur. Indeed, attractive surface forces must be able to overcome repulsive forces associated with the elastic deformation of surface asperities. Therefore, a specific CMP process is used for surface preparation of the hybrid Cu/SiO_2 layer prior to bonding. It can be noted however that low surface roughness alone is not sufficient to ensure good contact over the whole wafer surface. Surface topography must also be monitored over a wide range of length scales, including surface waviness and flatness, as the adhesive forces arising between the smooth bonding surfaces are short-ranged.

[1] Typically, up to about 0.6-0.7 µm RMS roughness for SiO_2/SiO_2 bonding (Moriceau, 2012)

II.2.2.2 Decontamination

For similar reasons, the issue of particle contamination is critical for successful bonding. Indeed, whereas small particles (with diameter less than a few 0.1 µm) can be accommodated through localized elastic deformation of the bonding surface, for larger particles a distinct deformation regime was identified, in which the wafers deform by bending to form a bulge around the particle. This regime typically leads to very large bonding defects, with diameters several order of magnitudes larger than the particle itself. Such particle contamination may be either airborne or process-induced (e.g. left-over slurry particles or polishing residues after surface planarization by CMP). Another crucial aspect is organic contamination, for instance due to volatile organic compounds adsorption during wafer handling or storage in plastic containers. This kind of contamination may lead to the formation of a thin layer of hydrocarbons or silicones at the wafer surface, inhibiting the establishment of surface forces during pre-bonding.

Therefore, extensive surface cleaning is necessary before the bonding process to remove the different contaminants. Typically, this is achieved using a multi-step cleaning process, with (i) organic removal, then (ii) particle removal and (iii) metallic contaminants removal.

II.2.2.3 Activation

Both the chemical composition of the solutions used for surface cleaning, as well as the chronological order of the cleaning sequence are crucial for surface activation, i.e. to create the conditions necessary for the formation of chemical bonds between the two wafer surfaces in contact. For instance, following surface cleaning with typical solutions for organic removal, e.g. sulfuric acid with hydrogen peroxide (SPM solution), the bonding surface acquires a nonzero charge leading to particle attraction. Organic removal is therefore carried out before particle removal. The chemistry used for particle removal, usually the "standard clean 1" solution (SC-1), a mixture of hydrogen ammonium and hydrogen peroxide, enables to activate the bonding surface by rendering it highly hydrophilic, which as will be seen in Section II.2.4 is one of the possible ways to achieve direct bonding. But conversely, this solution etches the surface oxide and may lead to an increase of the surface roughness. Therefore, cleaning time is closely monitored to enable particle removal and surface activation, but limit surface roughening. In addition, the SC-1 clean is also known to induce metallic contamination. A third cleaning step is thus necessary, generally using a hydrochloric acid solution referred to as "standard clean 2" (SC-2). This solution however leads to a degradation of the surface hydrophilicity. Thus, solution concentration, cleaning time, and temperature must be carefully adjusted to mitigate this detrimental influence of the SC-2 clean on the surface activation obtained during the SC-1 clean. Other possibilities for enhanced surface activation include plasma activation and/or bonding in ultra-high-vacuum.

II.2.3 Bonding sequence

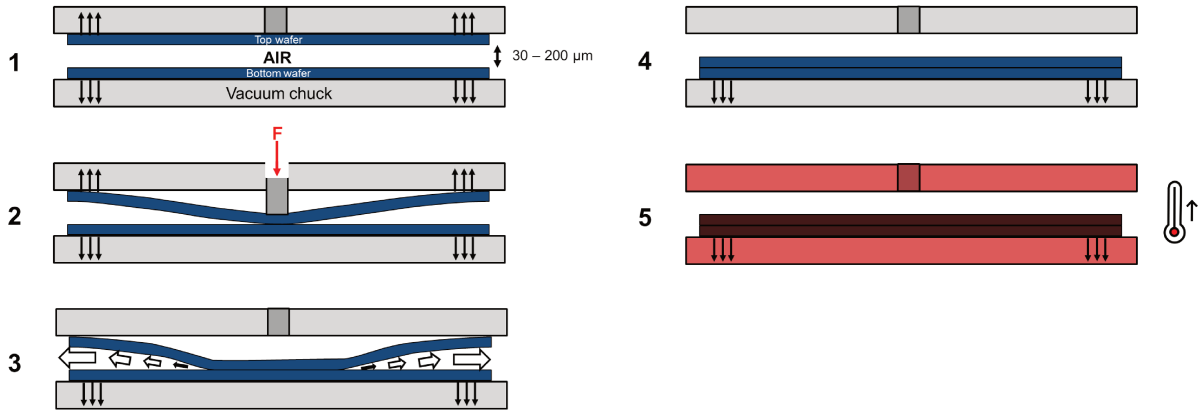


FIGURE II.2: Schematic principle of the wafer bonding process: 1-4) pre-bonding; 5) permanent bonding.

A key feature of metal/oxide hybrid bonding is to enable electrical connection between two chips without the need for metal bumps or any external material, e.g. glue or solder. This is achieved in a two-step process, carried out at the wafer-level in the present study[1]:

Pre-bonding: the wafers to be bonded are positioned with their bonding surfaces facing each other[2]. Then, the matching Cu/SiO$_2$ layouts on both surfaces to be bonded are aligned. To do so, specific Cu metallization structures in the interconnect stack are used as alignment marks for the precision alignment tool. Finally, adhesion is initiated at the center of the wafer by a metal pin[3], pressing the two wafers into contact. Thanks to specific surface preparation, surface forces appear at the periphery of this initial contact region, as will be seen in Section II.2.4. This in turn leads to the spontaneous propagation of a *bonding wave* from the center towards the wafer edge.

Permanent bonding: following pre-bonding at room temperature, the wafers adhere to each other enough to be handled without deteriorating the alignment between the matching Cu/SiO$_2$ layouts on both sides of the bonding interface. However, at this stage the adhesion is quite weak, and the formed bond is actually reversible. While this can represent an opportunity, in case the bonded pair needs to be separated and re-bonded, e.g. to improve alignment or eliminate foreign contaminants, it is not sufficient for robust mechanical attachment of the bonded pair. Therefore, the pre-bonding step is followed by a thermal anneal, which changes the chemical nature of the formed bond, as described in Section II.2.4, resulting an increase of the bonding energy by an order of magnitude. In some

[1] As discussed in Chapter I, hybrid bonding is possible both at the wafer-level or at the chip-level.

[2] Therefore, one wafer has to be flipped, which leads to specific integration issues for subsequent electrical connection of the stacked assembly to the package substrate by wire-bonding: this topic is investigated in Chapter II.

[3] Contact may also be initiated from the edge, as frequently seen in the literature, for direct bonding with homogeneous bonding surfaces. For hybrid bonding however, with patterned surfaces, this may pose a risk in terms of *overlay* between the matching layouts of the Cu/SiO$_2$ patterned bonding surfaces, as detailed in Section II.2.6.

cases, a compressive force may be prescribed in addition on the bonded assembly to further reinforce the bond (*thermo-compression*) [1]. After completion of this *consolidation annealing*, a permanent, mechanically more robust bond is obtained, able to withstand the remainder of the fabrication sequence, especially the critical wafer thinning and wafer sawing steps.

II.2.4 Physical mechanisms

The hybrid bonding process can be seen as a combination of two long-known technologies in the semiconductor industry, involving two distinct bonding mechanisms, namely *direct bonding* between the SiO_2 dielectric portions of the hybrid bonding surface (also referred to as molecular bonding)[1], and *metal bonding* between matching Cu hybrid bonding interconnect pairs (also referred to as diffusion bonding). In the following, drawing upon recent research, the physical mechanisms underlying these two distinct phenomena, both involved in hybrid bonding, are reviewed.

II.2.4.1 Direct bonding

The long-known phenomenon of spontaneous adhesion between flat, clean, smooth glass surfaces [2–4] was first leveraged for industrial purposes in the early 1960s at Philips Research, with the application to the direct bonding of optical devices for laser miniaturization [5]. It was then adapted and developed by the semiconductor industry some twenty-five years later by two research groups independently at IBM and Toshiba for the fabrication of silicon-on-insulator substrates by wafer direct bonding [6, 7]. Since that time, molecular bonding was shown to occur for various kinds of materials (ceramics, semiconductors and metals), bonding pairs (homostructures and heterostructures) and surface properties (hydrophilic and hydrophobic). We focus exclusively on the adhesion between SiO_2 surfaces by the hydrophilic bonding mechanism, which is used in this work for chip stacking. Another type of direct bonding is possible, namely hydrophobic bonding, in the context of semiconductor manufacturing corresponding to adhesion between native oxide stripped Si surfaces. Hydrophobic bonding relies on distinct surface preparation and adhesion mechanisms and will not be described here.

The physical mechanisms and surface parameters controlling the hydrophilic bonding phenomenon have been, and still are, the subject of much research. In the general case, adhesion may arise between the surfaces of two solids, brought in sufficiently close proximity, due to four types of surface interactions [8]:

Capillary forces: due to capillary condensation of water at surface contact sites, depending on the presence or not of a condensable vapor;

[1] It must be noted that direct bonding of Si wafers is also possible in absence of an oxide layer, native or deposited, at the surface before bonding. This type of bonding is referred to as *hydrophobic bonding*.

Coulomb forces: due to the presence of adsorbed electrons or ions on the surfaces, resulting in macroscopic surface charges.

Van der Waals forces: due to the interaction between atomic or molecular dipoles, whether instantaneous, permanent, or induced. Contrary to the former two contributions, these forces are always present, even between neutral atoms or molecules.

Hydrogen bonds due to the interaction of hydrogen with highly electronegative atoms, e.g. oxygen-mediated hydrogen bonds in presence of hydroxyl groups (–OH) on the surfaces. Hydrogen bonds are very energetic, ranging somewhere between a Van der Waals interaction and a covalent bond.

For hydrophilic bonding, the respective contribution of the different surface interactions is determined by surface preparation, bonding environment, and the distance separating the bonding surfaces.

II.2.4.1.1 At room temperature

For wafer bonding, capillary attraction is not the main driving force bringing the two wafers into contact during the early stages of the bonding, except for some specific applications such as direct bonding with a liquid interlayer [9–11]. Indeed, although the bonding surfaces have become highly hydrophilic[1] after surface preparation, due to the basic and oxidizing properties of the surface cleaning solution, only a few water monolayers are adsorbed on the bonding surface. Therefore, it is generally assumed that the volume of adsorbed water is not sufficient to form a meniscus at the bonding front during the contacting phase of the wafer bonding process, and that the radius of curvature of the surface asperities is too large for capillary bridges to appear between the contacting surfaces [13, 14].

In the early stages of wafer bonding, as the wafers are placed in front of each other, a thin cushion of air (about 30 µm-thick [13]) remains trapped in-between, preventing contact between the bonding surfaces. Due to gravity, this thin air layer progressively flows outside of the inter-wafer gap as the top wafer falls into contact with its bottom counterpart. As shown by Navarro [13], this process is very slow however, with about one day's time needed to reach the last 100 nm. Therefore, in practice, bonding is systematically initiated by pressing the two wafers at a given location (in most cases at the center) using tweezers or a metal pin. This first contact point in which adhesion has been initiated then spontaneously extends, leading to the propagation of a "bonding wave" across the whole bonding surface. This method is far more effective, with a bonding front propagation speed of about 1 cm/s [15–17]. It is widely accepted that adhesion is the result of oxygen-mediated hydrogen bonds between the hydroxylated SiO_2 bonding surfaces [18–21]. Thus, rather than capillary forces, hydrogen bonding coupled with air flow outside of the inter-wafer gap is the driving mechanism in the early stages of wafer bonding.

[1] Droplet angles of a few degrees only are obtained in wettability experiments for various surface cleaning chemistries [12].

Wafer bonding experiments were carried out in various atmospheres by Tong and Gösele [18], who observed a strong influence of gas viscosity on the propagation speed of the bonding front. In another study [22], they bonded wafers in ultra-high vacuum and reported instantaneous bonding, by bonding wave velocities in the order of 1 m/s. For a given bonding environment, bonding wave velocity is a good indicator of the quality of the bond [23]. Extensive analytical and numerical work of bonding front propagation has been done to enable the adhesion energy evaluation from bonding wave velocity and assess the influence of various process or design parameters on the bonding quality. Rieutord et al. and other authors [15–17] have developed a quantitative analytical framework based on a Poiseuille flow between two deformed adhesive Kirchoff plates. Turner et al. [24–26] used an analytical energy balance approach and a numerical model based on fracture mechanics (VCCT) to evaluate the equilibrium bonded length for various design or process parameters (e.g. initial wafer geometry, etch patterns, bonding configuration). Kubair et al. [27] and Estevez et al. [28] also proposed a numerical approach, based on cohesive zone modeling.

II.2.4.1.2 During thermal annealing

Immediately after pre-bonding at room temperature, the formed bond is quite weak, with a typical bonding strength of about 100 $mJ\,m^{-2}$ [29]. While sufficient for wafer handling, these adhesion energies are far too low to provide sufficient mechanical robustness during product life, or even the remainder of the fabrication sequence. As mentioned in Section II.2.3, a thermal anneal is thus necessary to strengthen the bonding interface, increasing the bond strength up to about 2 $mJ\,m^{-2}$ [18, 29–31]. Such values, closer to the fracture toughness of bulk silicon [32], enable to form a permanent bonding, considerably mitigating the risk of interfacial fracture.

From a chemical standpoint, the accepted mechanism to explain the large increase of adhesion energy between initial bonding and permanent bonding is a thermally-activated transformation of the comparatively weak hydrogen bonds formed at the bonding interface into covalent bonds. Stengl et al. [33] proposed a conceptual chemical model for the hydrophilic bonding of perfectly flat surfaces. They distinguished three regimes associated with different temperature ranges: between room temperature and 200 °C, formation of hydrogen bonds between two adsorbed water monolayers on each side of the bonding interface, between 200 °C and 700 °C, desorption of the adsorbed water molecules leading to direct contact between the bonding surfaces and formation of hydrogen bonds between silanol groups (Si–OH), and above 700 °C, replacement of the weak hydrogen bonds by covalent siloxane bonds (Si–O–Si). This model was later revised by Tong and Gösele [18], who suggested that in the low temperature regime two or three water monolayers may be present between the bonding surfaces, leading to a more stable configuration. This also lead them to propose new estimations for the separation distance between bonding surfaces in the different regimes. Several works relying on molecular dynamics simulations have since enable to derive more precise estimates for these separation distances [27, 34–37]. Recently, Navarro et al. [21], proposed a mechanistic model for hydrophilic bonding drawing upon these

studies and accounting for realistic surface topography. Their model captures the experimentally-observed hysteretic effect between adhesion and separation [38], which they attributed to the evolving distribution of water along the interface during bonding.

A mechanical approach is also needed to account for surface asperities, in addition to the chemical aspects investigated in the above studies. Indeed, the chemical nature of the formed bond greatly depends on the separation distance between the bonding surfaces at the local scale, and conversely the bond type will in turn determine the average separation distance at the surface scale. Asperity deformation is also a crucial factor deciding whether, for a given surface roughness, excessively large repulsive forces will prevent bond formation, or on the contrary if elastic accommodation of the asperities is possible during bonding. Various studies have investigated these aspects. Yu and Suo [39] have analytically derived a criterion for the bonding of two wafers with a 3D sinusoidal surface profile, confirming some earlier results by Tong and Gösele [40, 41], obtained using thin plate theory. Miki and Spearing [42] have correlated the surface morphology with the resulting apparent bonding energy, using the notion of *bearing ratio*. Turner et al. [43] have investigated the bonding of more realistic surfaces using a mechanics-based numerical model and derived design maps showing acceptable magnitudes of height variations as a function of spatial wavelength. More recently, a number of studies have relied on the Greenwood-Williamson asperity contact model [44], used in conjunction with adhesive contact theory to predict the equilibrium distance between the bonding surfaces. Gui et al. [45] have shown that both the effective bonding energy and the effective bonding area between rough surfaces depend on a dimensionless *surface adhesion parameter*, using the DMT model (Derjaguin-Muller-Toporov [46]), whereas other authors [47, 48] have used the JKR model (Johnson-Kendall-Roberts [49]). Rieutord et al. [47] tested the results obtained using the latter model against X-ray reflectivity and blade insertion measurements carried out on several surfaces with distinct roughness values, and were able to establish a correlation between interface separation and bonding energy.

II.2.4.2 Metal bonding

Spontaneous adhesion between clean and smooth pieces of metal has also been known for a long time. Medieval encyclopedists [50] already highlighted the importance of cleanliness for joining thin gold and silver foils. In the eighteenth century, Desaguliers showed in a series of experiment on friction that pressing clean and smooth lead spheres into contact results in strong attachment [51]. Since the 1970s, *pressure welding* and *diffusion bonding* have been used in the automotive, railway and aeronautical industries for commercial or military applications [52, 53]. In the semiconductor industry, this type of process was also introduced around that time, with the development of *thermosonic wire-bonding*[1] [54]. Another example is the introduction in the 2000s of *thermocompression bonding*, an important technology for 3D integration applications [55, 56].

[1] Mechanical aspects related to this process are investigated for a 3D stacked image sensor in Chap 3.

A distinctive feature of diffusion bonding is the absence of welding material, contrary to eutectic bonding for example between Cu pillars. Instead, adhesion is achieved through the application of pressure and temperature, enabling to (i) achieve intimate contact between the bonding surfaces, and (ii) activate diffusion processes at the interface. Although many different metals can be used for diffusion bonding, we will focus here exclusively on Cu-Cu bonding. While for bonding with dissimilar metals, intermetallic compounds[1] are formed at the interface, the use of a monometallic bonding process allows seamless contact between the bonding surfaces, resulting in the formation of metallic bonds at the interface.

II.2.4.2.1 At room temperature

The physical mechanism involved in the adhesion of Cu surfaces at room temperature are assumed to be quite similar to those described in the previous section for SiO_2 surfaces [57]. Indeed, copper surfaces are easily oxidized, for example during storage in ambient air. The presence of this native oxide during surface cleaning before the bonding process makes the copper surfaces hydrophilic. Gueguen et al. [58] carried out water droplet angle measurements before and after surface preparation, and evidenced a sharp increase of surface hydrophilicity immediately after surface cleaning (from 49° down to 19°[2]). Similarly to the case of SiO_2 surfaces, water adsorption is expected to occur on the Cu surfaces, leading to hydrogen bond formation and thus adhesion during pre-bonding at room temperature. This assumption is supported by a number of experimental studies in which the same process as for SiO_2 hydrophilic bonding was successfully used to achieve direct bonding of Cu blanket wafers at room temperature [58–63]. Another evidence is the possibility to bond materials of very different nature, provided that an oxide has formed at their surface and adequate surface preparation has been carried out. For instance, direct bonding of Cu and SiO_2 blanket wafers at room temperature has been demonstrated [60, 61]. In both cases, namely Cu-Cu and Cu-SiO_2 direct bonding with blanket wafers, blade insertion measurements have revealed bonding energies quite close to those obtained for SiO_2-SiO_2 hydrophilic bonding [60, 61]. It must be noted that the Cu surfaces become hydrophobic again a few hours after surface preparation, leading to a strong time constraint between the surface cleaning and pre-bonding steps [58].

The presence of native oxide on Cu surfaces is not fundamentally detrimental during pre-bonding at room temperature, by enabling adhesion through the formation of hydrogen bonds. Actually, a significant increase of the bonding energy[3] is even observed with storage time after pre-bonding, typically from a few $100\,mJ\,m^{-2}$ to about $2.5\,J\,m^{-2}$ [61, 66]. Di Cioccio et al. [61] characterized the Cu-Cu bonding interface after bonding at room temperature using X-ray reflectivity and showed that the initial surface oxide undergoes significant growth with storage time, from 0.2 nm up to 4 nm for an initial roughness of about 2 nm RMS. This growth is attributed to copper

[1] For example, AuAl for wire bonding or CuSn for Cu pillars.

[2] Gondcharton [53] even reports droplet angles below 3°.

[3] Although less marked, a similar increase of the bonding energy with storage time was observed for SiO_2-SiO_2 direct bonding [20, 40, 64], which was attributed to thermally-activated capillary condensation of liquid water bridging the two surfaces in the contact regions [65].

diffusion through the oxide to react with adsorbed water [53, 67]. However, typical durations associated with this kind of interface reinforcement are in the range of days, and in addition the presence of a continuous layer of cuprous oxide at the interface is not desirable from an electrical standpoint. In that respect, the surface oxide is rather regarded as a limiting factor, preventing intimate contact between Cu surfaces, as well as the creation of strong metallic bonds sought after to enhance the bonding energy.

II.2.4.2.2 During thermal annealing

The thermal annealing process carried out to strengthen the SiO_2-SiO_2 interface actually enables to circumvent the detrimental effect of cuprous oxide presence at the Cu-Cu bonding interface. Several experimental studies have been conducted to investigate the evolution of the Cu-Cu bonding interface during annealing, for instance using X-ray reflectivity or TEM cross-sections [58–61, 68, 69]. These have shown that direct contact between Cu surfaces can be obtained, even in the presence of an initial surface oxide, provided that sufficient temperature (and for thermocompression bonding, sufficient pressure) is prescribed. Various mechanisms have been proposed to explain this observation [57–59, 62, 63, 70, 71]:

At room temperature: asperity contact due to surface roughness and interface reinforcement through interface oxide growth;

Between RT and 150 °C**:** fracture/dewetting of the interface oxide, formation of grain joints in regions where the oxide layer is absent, segregation of remaining oxide into nodules along the interface;

Above 150 °C**:** reorganization of the Cu-Cu interface, transformation of initially T-shaped triple junctions into a lower energy 120° configuration.

The loss of continuity of the oxide layer between room temperature and 150 °C may attributed to several factors, among which (i) breakage of the brittle oxide due to dislocation pile-up in the neighboring plastically deforming copper under thermal loading [57, 72], (ii) competition between metal-oxide interface and grain boundary energies resulting in a thermodynamically unstable oxide [57, 63], (iii) reduction of the oxide by the surrounding Cu material with dissolution of oxygen atoms into Cu grains and/or migration into grain boundaries [53, 73–75], (iv) phase transformation of the oxide from Cu_2O to CuO resulting in a volume change of about 12% [53, 76, 77].

The reorganization of the Cu-Cu interface is a result of the apparition of diffusion wedges at triple junction points [62] and/or abnormal grain growth [68]. As a result, the initially flat bonding interface assumes a sawtooth shape and becomes virtually equivalent to bulk-like grain boundaries. A very large increase of the bonding energy ensues, which becomes almost impossible to evaluate using typical destructive methods (blade insertion, four-point bending) due to sample breakage or crack kinking [58, 63, 78]. Thus, it may be inferred that the Cu-Cu bonding energy

becomes comparable to bulk fracture toughness. Radu et al. [60] investigated the influence of annealing temperature on the bonding energy for different types of interface. A very large increase is observed above 150 °C for Cu/Cu full sheet and for Cu/SiO_2 patterned[1] interfaces, with for the latter up to 8 J m^{-2} after annealing at 400 °C. A weak bonding energy was obtained for a Cu/SiO_2 full sheet interface however, with initial values of a few 100 mJ m^{-2} at room temperature, without any significant increase even after annealing at 400 °C.

By combining oxide-oxide direct bonding and metal-metal diffusion bonding within a single heterogeneous interface, the hybrid bonding process enables to achieve simultaneous mechanical and electrical interconnection between the top and bottom chips of the considered 3D stacked image sensor. In the next two paragraphs, the advantages and drawbacks of the Cu/SiO_2 hybrid bonding compared to other methods are reviewed, as well as the current challenges for application of this process to 3D chip stacking.

II.2.5 Advantages and drawbacks

Contrary to 3D assembly methods based on TSVs, with hybrid bonding the connection between the chips is achieved at the bonding interface only, rather than extending through the whole thickness of interconnect stack. Therefore, this technique does not require large FEoL or BEoL exclusion regions, enabling to achieve great compactness and ultra-high density of interconnections. In addition, with hybrid bonding the bonded surface extends laterally over the whole chip area, contrary to metal bonding techniques such as eutectic bonding or solid-liquid interdiffusion (SLID) for which micro-pillars are required at the bonding interface, resulting in a stand-off distance between the stacked dies. Since no solder alloy or intermetallic compound is necessary to form the bond, issues associated with solder bridging or voiding during reflow can also be alleviated.

Among the two hybrid bonding process variants presented in Section II.2.1, namely metal-polymer hybrid bonding and metal-oxide hybrid bonding, the latter can be achieved using similar tools and processes as for interconnect manufacturing, namely the damascene process and chemical-mechanical polishing, thus providing better compatibility with typical CMOS processes. Furthermore, the use of an inorganic dielectric enables to avoid issues typically associated with polymer materials, such as ageing or poor mechanical properties above the glass transition temperature. Another advantage of the Cu/SiO_2 hybrid bonding process considered here is the possibility to achieve strong attachment without necessarily needing to exert pressure on the bond, as for thermocompression bonding.

The above advantages come at the price of extensive surface preparation however, as detailed in Section II.2.2, with stringent requirements on surface flatness. Indeed, both Cu and SiO_2 bonding surfaces are far less compliant compared to polymer dielectrics or eutectic alloys. In addition, thorough surface cleaning is needed to avoid particle, organic and metallic contamination, while

[1] The metal surface fraction for the studied samples was 20% [60].

ensuring that the bonding surfaces remain sufficiently hydrophilic to obtain good adhesion during the pre-bonding step at room temperature. The hybrid bonding process also requires a high thermal budget compared to polymer curing or solder reflow, with temperatures ranging between 200 °C and 400 °C for several hours during the post-bonding annealing step.

II.2.6 Current challenges

Bonding energy measurement

Quantitative evaluation of the bonding energy is a very challenging task. Generally, the obtained energy values are deemed valid only for qualitative comparison, between samples measured using the same technique, with the same protocol, and in a similar environment [23]. As with any interface energy evaluation, the influence of the measurement technique on the obtained value is a critical issue. Indeed, distinct loading configurations with distinct mode mixity lead to quite different results [79]. For SiO_2/SiO_2, another important aspect is the sensitivity of the formed siloxane bond during thermal annealing to environmental effects. A strong influence of water stress corrosion on the measured energy values was evidenced [29, 64, 65, 80]. Furthermore, in presence of copper at the bonding interface, the respective contributions of interface energy and plastic dissipation are difficult to decorrelate [61, 81]. However, those issues are generally encountered in many studies related to fracture mechanics. A more specific issue for wafer bonding is the presence of an hysteresis effect between adhesion and debonding [26, 38]. Several authors stressed the importance of operating a distinction between "bonding energy" and "work of adhesion" [13, 17], and specific measurement methods enabling to obtain the latter were recently developed [64]. Being able to quantitatively evaluate adhesion at the hybrid bonding interface is an important requirement to allow further optimization of the mechanical robustness of the bond formed between the stacked chips. These investigations require specific equipment and sample preparation however, and thus will not be undertaken here.

Bonding-compatible planarization of patterned surfaces

Surface preparation requirements become even more challenging for hybrid bonding, due to the need for simultaneous surface planarization, cleaning and activation of heterogeneous Cu/SiO_2 surfaces. Indeed, Cu and SiO_2 have distinct physical and surface properties (e.g. Young's modulus and material removal rate), causing distinct planarization behaviors during chemical-mechanical polishing [61, 82–84]. As will be seen in the next section, this heterogeneous behavior can lead to topography defects, which in some critical cases may even prevent proper contact between the Cu interconnects at the interface. In addition to a decrease of the electrical contact area, these bonding defects could also result in a degradation of the overall adhesion of the hybrid bonding interface. These issues related to bonding-compatible planarization of Cu/SiO_2 patterned surfaces are investigated experimentally and numerically in the next two sections.

Overlay mismatch between hybrid bonding interconnects

The other main limiting factor for hybrid bonding is actually the alignment precision that can

be reached by the bonding equipment [85]. Indeed, although well-defined hybrid bonding interconnects can be processed below submicrometer dimensions using current photolithography capabilities, the alignment precision during the pre-bonding step must be sufficient for interconnects on one side of the bonding interface can be connected to their counterpart on the other side. This overlay mismatch results to a decrease of the electrical contact area, and the presence of overlapping Cu/SiO_2 regions, which (i) were shown to have significantly lower bonding energy compared to Cu/Cu and SiO_2/SiO_2 even after annealing at 400 °C and (ii) could lead to Cu diffusion into the SiO_2 dielectric. These issues have been addressed in several recent studies [83, 84]. In Section II.3, the consequences of this issue on Cu-Cu interface closure will be assessed, accounting for Cu surface planarization defects resulting from CMP on heterogeneous surfaces.

II.3 Experimental characterization

Note: The characterization work presented in this section was carried out by the Physical Characterization team at STMicroelectronics Crolles. Here, these individual experimental results were brought together, interpreted and contextualized in order to better illustrate the investigated issues and provide a basis for the simulation study.

In the following, test chips assembled by hybrid bonding are characterized before and after bonding, aiming to evaluate the influence of several design parameters, namely interconnect geometry or layout. Among these samples, for each set of process parameters one half is dedicated to surface topography characterization of the Cu/SiO_2 patterned bonding surfaces by atomic force microscopy (AFM) after surface planarization. The remainder undergoes the bonding process and is used for subsequent SEM or TEM inspection of cross-sections of the Cu-Cu bonding interface obtained by ion-milling (FIB), enabling to investigate the influence of various surface topographies on the Cu-Cu interface bonding quality.

II.3.1 Samples description

Two layers of dielectric material separated by an etch-stop layer are deposited sequentially, as follows:

Via layer: SiCN barrier + USG[1] (PECVD: 60 nm + 600 nm)

Metal layer: SiN barrier + TEOS[2] (PECVD: 60 nm + 500 nm)

After photolithography and dry etching to form cavities for the Cu interconnects, a thin diffusion barrier (PVD: TaN/Ta 13 nm) is deposited, and the trenches are filled with copper (PVD: Cu seed 90 nm; ECD: Cu 1 µm). An annealing process is then used to reduce the Cu resistivity

[1] Undoped silicate glass, refers to SiO_2 deposited using a silane-based chemistry
[2] Tetraethylorthosilicate, refers to SiO_2 deposited using a TEOS-based chemistry

and control the microstructural properties. Finally, to remove the excess of metallization (Cu overburden) and obtain a flat and smooth surface, as required for direct bonding, a specific CMP process is used.

II.3.2 Cu/SiO_2 surface topography characterization before bonding

Since Cu and SiO_2 have different removal rates, fine-tuning of the CMP parameters is thus required to avoid obtaining surface topologies that, although very smooth, exhibit significant height variations due to different erosion depths between the interconnects and the surrounding dielectric: in a majority of cases, the Cu lines in BEoL stacks are slightly overpolished compared to the surrounding dielectric surfaces. For hybrid bonding applications, this well-known "Cu dishing" effect is potentially detrimental to the bonding between matching interconnects on both surfaces. As will be seen in the following, how deep a dishing depth is acceptable for efficient hybrid bonding is a central question.

In the following paragraphs, surface topography measurements are carried out at the pattern scale for bonding pads with different geometries or processed using different CMP recipes. The measurements are carried out in the "dummies" regions on several test chips using an AFM (Bruker Dimension Icon, Bruker, Karlsruhe, Germany) in tapping mode with a standard probe (Bruker Tespa-v2) of nominal tip radius 7 nm. Areas of $10 \times 10\,\mu m^2$ centered about a randomly selected pad within the region of interest are scanned with 512×512 pixel resolution at a 1 Hz line scan rate. The measured horizontal and vertical distances are calibrated to ±1%.

In the next paragraph, the surface topography of a Cu hybrid bonding interconnect corresponding to a worst case is used to exemplify some typical planarization defects encountered at the pattern scale.

II.3.2.1 Planarization defects classification

A line scan taken across of a square Cu pattern of width 4.4 µm at the wafer center, after a CMP process identified as leading to a worst case in terms of Cu dishing[1] is presented in Figure II.3. The boundary between Cu and SiO_2 regions is apparent from sharp drops in the measured profile at the periphery of the Cu pattern, referred to in interconnect manufacturing as "fangs". This edge overetching is assumed to result from the high static etch rate of the TaN/Ta diffusion barrier compared to Cu and SiO_2 [86, 87]. A close inspection of the width of the recessed zone therefore shows that for the considered Cu pattern, not only the interconnect, but also a significant fraction of the surrounding oxide is overpolished. The surface topography of a recessed Cu pad may more generally be decomposed into three main features [88]:

[1]For confidentiality reasons, in Figure II.3 (and in the remainder of this chapter) the measured Cu dishing depth is normalized against a reference critical value, whose meaning will be discussed in Section II.3.3.1. The measured pad overpolishing depth is in the 0.5-20 nm range.

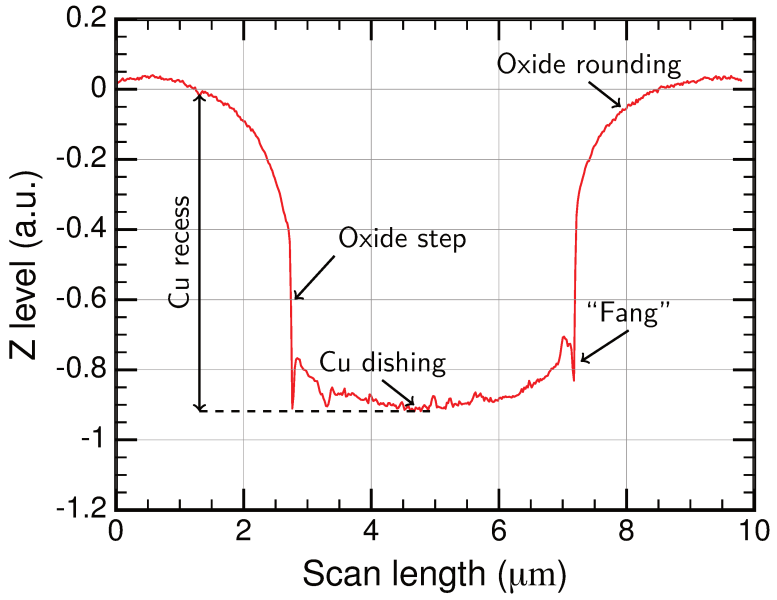


FIGURE II.3: Typical AFM line scan across of a Cu pattern processed using a worst-case CMP recipe. The measured total Cu recess is in the 0.5-20 nm range.

Cu dishing: recess of the pad center with respect to pad edges

Oxide step: recess of the pad edges with respect to the oxide surface

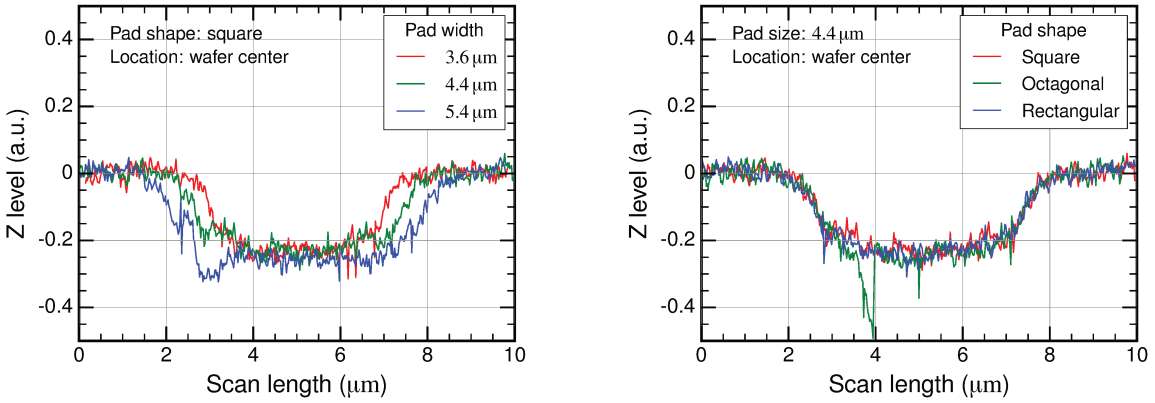
Oxide rounding: recess of the oxide surface at the vicinity of the pad compared to the reference level.

Cu dishing depth, oxide step height, as well as the extent of oxide rounding depends on CMP process parameters (e.g. slurry composition, pad pressure and rotation speed). For instance, a strong influence of CMP process selectivity (material removal rate ratio between Cu and SiO_2) on the topography of patterned surfaces is reported in the literature: dishing depth was shown to increase with slurry selectivity [89, 90].

In the next paragraph, various geometries and layouts are explored for the Cu patterns on the bonding surface. The aim is to assess whether besides process parameters, hybrid bonding interconnect design could also affect the surface topography obtained after CMP, as can be inferred from several numerical and experimental studies in the literature, which evidenced a strong influence of pattern width on the Cu recess [89–91]. Such variations, if any, could in turn have a significant influence on bonding.

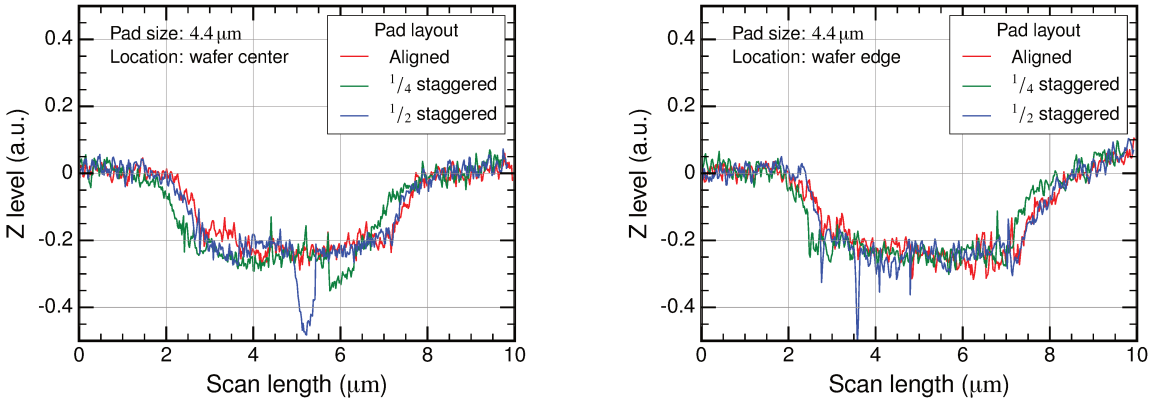
II.3.2.2 Influence of interconnect geometry and layout

Several types of Cu patterns were processed, at constant metal surface fraction, in distinct regions on the surface of a single wafer. Various geometries are explored for these patterns by introducing several geometries, namely square, octagonal, rectangular shapes, and widths of 3.6, 4.4 or 5.4 µm for square and octagonal shapes and $4.4 \times 8.8\,\mu m^2$ for rectangular patterns. In addition, several layouts are also investigated. While the majority of Cu patterns are arranged in a grid, in some regions a staggered pattern was introduced by shifting consecutive rows by a quarter- or a half-period. Line scans acquired across square pads with varying dimensions and processed using recipes B are shown in Figure II.4a.



(a) Influence of pad width (b) Influence of pad shape

FIGURE II.4: AFM line scans across hybrid bonding pads located at the wafer center with recipes B: a) influence of pad width and b) influence of pad shape.



(a) Wafer center (b) Wafer edge

FIGURE II.5: AFM line scans across square pads for different patterning layouts for recipe B a) at the wafer center and b) wafer edge.

For the few measured cases, no significant variation of the surface topography is observed for varying pattern width, shape, nor between aligned and staggered patterns, as apparent from in Figure II.4a and Figure II.5. A comparison was also made between surface topographies measured at the wafer center and wafer edge, showing no significant differences. In addition, the surface roughness for Cu surfaces is stable at around 0.2 nm RMS. In some cases, a few sharp

drops are observed in the Cu surface profiles, which are believed to result from scratches or grain pull-out during CMP.

In conclusion we observe that, for the investigated Cu pattern width range, CMP results in a similar dishing depth. The geometry (square or octagonal) or layout do not influence the magnitude of the dishing either. In the following, a morphological study of the hybrid bonding interface is carried out on wafers having undergone different CMP recipes, resulting in various degrees of Cu recess. The aim is to assess the influence of CMP-induced surface topography on Cu-Cu interface closure after post-bonding annealing.

II.3.3 Cu-Cu interface morphological characterization after annealing

As discussed in the previous section, CMP process selectivity has a strong influence on Cu dishing. To investigate how this planarization defect may in turn affect the bonding, wafers were processed using different slurry selectivity, i.e. various degrees of Cu overpolishing after CMP, and bonded. These wafers are referred to by their respective "batch name", using letters from A to F with increasing slurry selectivity. To characterize the associated dishing, for each batch one wafer was extracted from the manufacturing line to carry out pattern-scale surface topography measurements by AFM, similarly to the previous section. Within each lot, the rest of the samples[1] undergoes the bonding process, followed by consolidation annealing at 400 °C for two hours. Cross-sections are then carried out at the local-scale (a few microns wide) by ion-milling (FIB), allowing detailed characterization of the bonding interface by SEM or TEM.

II.3.3.1 Influence of surface topography on Cu-Cu bonding

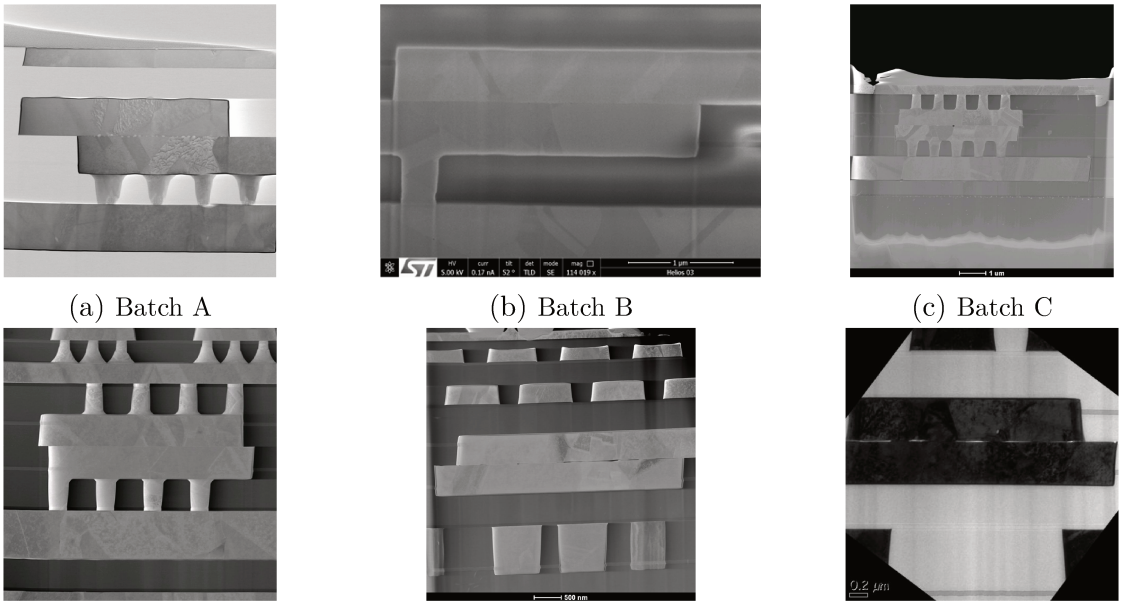
For each batch, FIB/SEM-TEM cross-sections of the bonding interface after post-bonding annealing are presented in Figure II.6:

Lots A to D: within the detection capabilities[2] of the equipment, no bonding defect is visible at the Cu-Cu interface and the pads appear to be completely bonded. The presence of cavities and cuprous oxide nodules with diameters of a few tens of nanometers can be noticed at the Cu-Cu interface however upon closer inspection (Figure II.7a). These nano-sized defects are assumed to arise due to dewetting of the oxide layer at the Cu interface and diffusion-induced voiding [58, 59, 62];

Lots E and F: about 10 nm-thick gaps are present between the Cu surfaces, extending over widths of about 1 µm (Figure II.7b). These large unbonded regions are believed to result from excessive Cu dishing due to the planarization process.

[1] A thin metal layer is deposited at the surface of the samples for the AFM measurements. Therefore, the exact same wafer cannot be used both for surface topography measurements and post-bonding interface characterization.

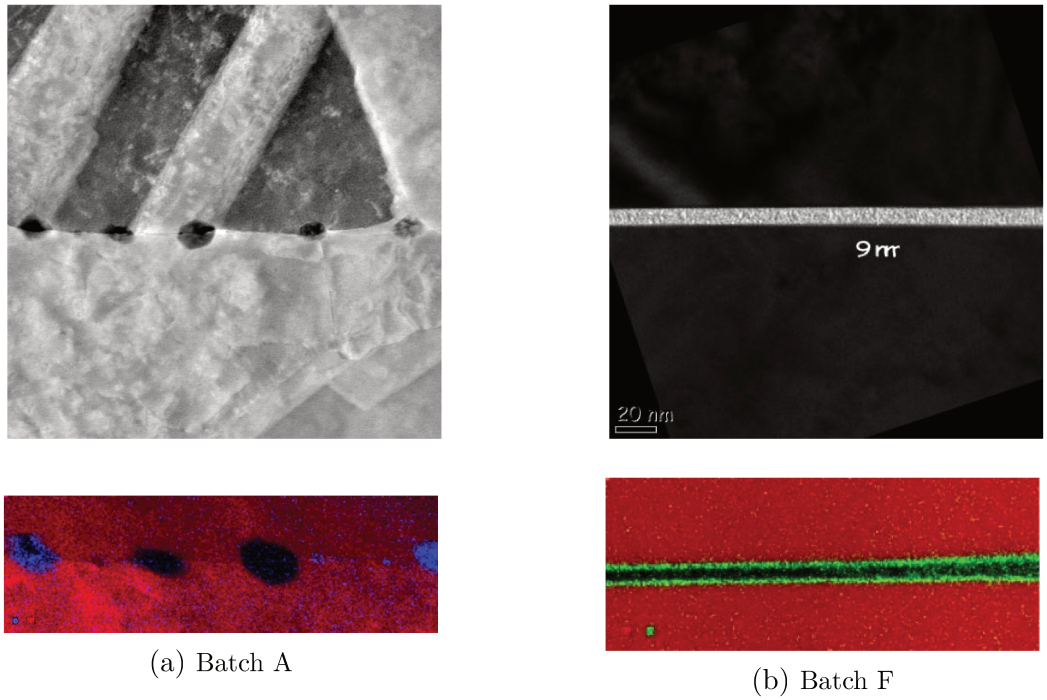
[2] The spatial resolution obtainable with the scanning electron microscope is 2-3 nm, and 0.5 nm for the transmission electron microscope.



(a) Batch A (b) Batch B (c) Batch C

(d) Batch D (e) Batch E (f) Batch F

FIGURE II.6: FIB cross-sections of the Cu-Cu interface between pads of width 4.4 µm after post-bonding annealing for wafers having undergone various CMP processes, with increasing selectivity: (a) scanning TEM; (b) SEM; (c)-(e) scanning TEM; (f) energy-filtered TEM



(a) Batch A (b) Batch F

FIGURE II.7: Zoomed views and corresponding chemical analysis (Cu and O atoms) for two typical bonding defects observed in FIB cross-sections of the Cu-Cu interface after post-bonding annealing for wafers having undergone various CMP processes, with increasing selectivity: (a) scanning TEM and EDX; (b) energy-filtered TEM and EDX.

In some cases, especially for wafers from lots A and B, it can be noted that the hybrid bonding interconnects are slightly misaligned. The influence of this overlay mismatch between recessed

interconnect pairs on Cu/Cu bonding during post-bonding annealing will be investigated numerically in Section II.4.

Aiming to determine typical Cu recess values associated with each of the above six cases, AFM measurements are carried out on wafers extracted from the same lots but not bonded. The results are summarized in Figure II.8, along with a qualitative evaluation of the bonding quality based on the FIB/SEM cross-sections of Figure II.7. Two distinct regimes can be observed:

"Bonded" case: for the smallest Cu recess values, the bonding pads appear to have a completely closed interface after post-bonding annealing (lots A to D).

"Partially bonded" case: conversely, above some critical[1] dishing value, large voids are detected and the Cu-Cu interface is bonded only partially (lots E and F).

These observations may be interpreted as follows: in presence of a sufficiently small dishing, the initial separation between the bonding surfaces can be entirely accommodated during post-bonding annealing thanks to constrained thermal expansion of the Cu pads, having large CTE[2] compared to SiO_2 [61]. However, if the CMP-induced surface topography becomes "too large", thermal expansion cannot compensate the gap between the pads, resulting in partial bonding only.

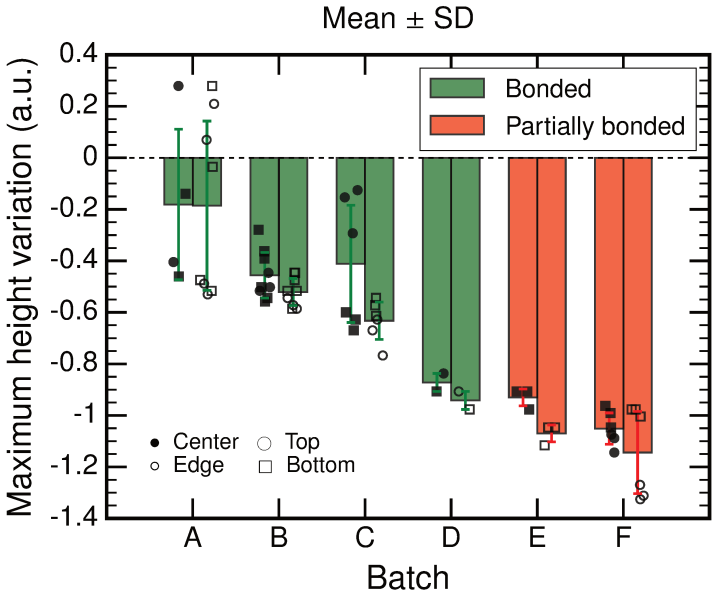


FIGURE II.8: Comparison of hybrid bonding pad surface topography measured on several sites by AFM with bonding quality evaluated qualitatively from FIB/SEM cross-sections on wafers extracted from the same lot.

[1] This is the value that was used as a reference for y-axis scaling in AFM line scans.
[2] $\alpha_{Cu} = 16.5 \times 10^{-6}\,°C^{-1}$, compared to about $\alpha_{SiO_2} = 0.5 \times 10^{-6}\,°C^{-1}$.

II.3.3.2 Bonding defects classification

The high-resolution cross-sections presented above enable a qualitative assessment of the bonding quality ("bonded" or "partially bonded"), and to approximately evaluate the lateral extent of the voids. However, little information is obtained on their spatial distribution across the Cu-Cu interface. Aiming to provide further insight, 3D imaging was carried out via FIB/SEM tomography on bonded pads from lots C and E, corresponding respectively to "bonded" and "partially bonded" cases in Figure II.8.

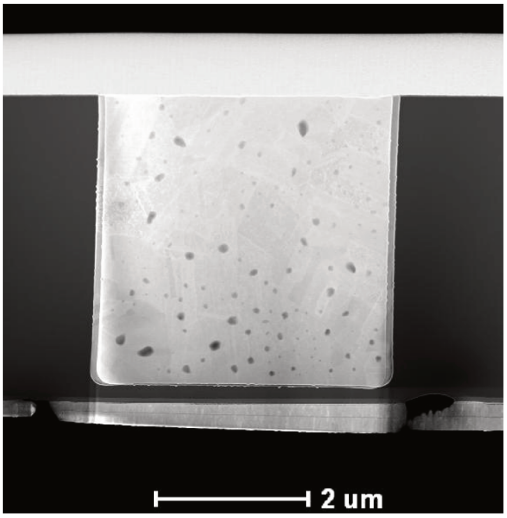
For the lot C sample ("bonded" case), FIB/SEM tomography was carried out by surface milling parallel to the interface plane, with a 15 nm vertical step. A reconstructed view based on the minimum voxel projection for the 133 images obtained is presented in Figure II.9. The Cu-Cu bonding interface can be seen in top view at the center of the picture. It can be noticed that for these bonding pads, the misalignment between Cu surfaces is very low, in the order of 100 nm. With this method, the voids present in the Cu-Cu stack can be observed in the form of dark spots[1]. However, no distinction can be made between the voids located inside the hybrid bonding pads and those present at the bonding interface. Therefore, similar measurements were performed on a different sample exhibiting similar interface morphology ("bonded" case with nano-sized cavities), for which the reconstructed minimum projection was taken perpendicularly to the bonding interface[2]. Two populations can be distinguished for the voids:

- small cavities, with a diameter of about 10 nm, mainly segregated at the bonding interface. This observation is in agreement with earlier results by Beilliard [92].
- larger voids, comparatively scarcer, ranging from 50 to 100 nm, found through the thickness of the Cu pads.

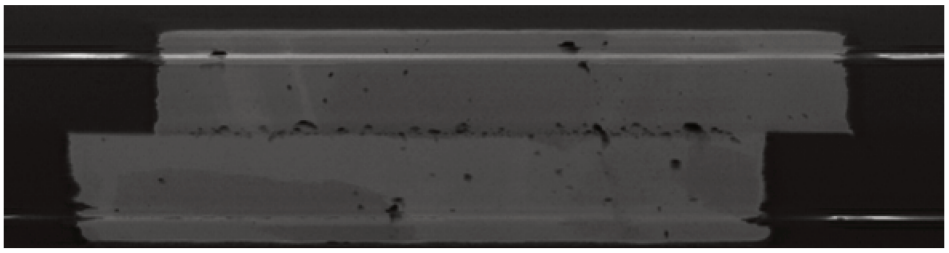
For the lot E sample ("partially bonded"), a different approach was used in order to avoid this issue. Similarly to the previous case, cross-sections of the bonded pads were carried out perpendicularly to the bonding interface, at a 20 nm step. However, for each cross-section the locations where interfacial voids are observed were reproduced "by hand" on a FIB/SEM tomography, corresponding to a projection overlay of the top and bottom pads in top-view, obtained by surface milling parallel to the interface. By doing so, only the interfacial voids are visible, enabling to derive a rough estimate of the effective bonded area. The obtained image is shown in Figure II.10. It can be seen that for this sample having a more marked Cu recess compared to the previous case, large voids are present at the bonding interface, extending over approximately 50% of the bonding surface. No specific symmetry or preferential location is observed for these cavities, nor any correlation detected between the void distribution and the grain structure.

[1] Within the equipment resolution (about 0.5 nm).

[2] While it may seem like grain boundaries are visible in Figure II.9, the reconstructed view is actually obtained by taking the minimum projections for all the cross-sections. Thus, no conclusion can be drawn about any preferential locations for the cavities compared to the grain structure (e.g., grain boundaries or triple junctions)



(a) Longitudinal direction



(b) Transverse direction

FIGURE II.9: 3D FIB/SEM reconstructed view (minimum projection) of the bonding interface after post-bonding annealing: a) parallel to the bonding interface for sample extracted from batch C ("bonded" regime); b) perpendicular to the bonding interface for a different sample, having similar interface morphology ("bonded" regime).

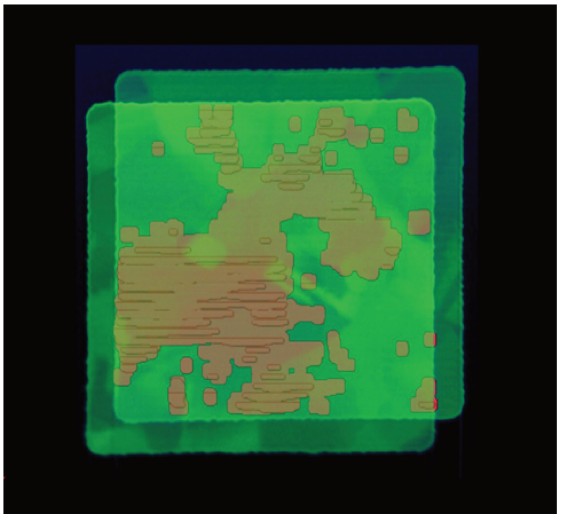


FIGURE II.10: 3D FIB/SEM reconstructed view (projection overlay of top and bottom pads) parallel to the bonding interface after post-bonding annealing for sample extracted from batch E ("partially bonded" regime). Voids observed by FIB/SEM tomography perpendicular to the interface are reproduced in red.

II.3.4 Discussion

In this section, an experimental study of the influence of CMP-induced surface topography defects on Cu-Cu interface bonding quality was carried out. In particular, Cu surfaces are typically in recess compared to the surrounding dielectric materials ("dishing" effect). Cu dishing originates due to the different material removal rates for the metal and dielectric regions on the patterned hybrid bonding surface. Naturally, CMP slurry selectivity (metal to dielectric material removal rate ratio) was therefore seen to have a strong influence on the final metal recess. An increase of metal recess is observed for increasing selectivity. This influence of selectivity on metal recess is well-known and has been reported in both experimental and modeling studies [90, 93].

For fixed CMP process conditions, no significant variations of Cu dishing were measured within the investigated range of pad widths, from about 3 µm to 6 µm. Yet several studies report an increase of metal recess for increasing pattern width [90, 94–97]. In the latter references, two regimes are identified depending on the surface roughness characteristics of the polishing pad:

- for small pattern width compared to typical dimensions of asperities at the surface of the polishing pad, only the smallest asperities can reach the bottom of the recessed metal surface. Larger asperities are filtered, resulting in lower Cu dishing[1].
- for large pattern width, a behavior close to that of a blanket film is obtained, resulting in a saturation of the dishing values.

Possible reasons explaining the mismatch between the latter results in the literature and the observations of this study include: (i) distinct behavior for the considered square Cu patterns, compared to Cu lines typically considered in the literature, (ii) increased asperity filtering due to the very small dimensions considered here compared to the above-mentioned studies, suppressing dynamic material removal in favor of static etching, or (iii) a lack of statistics in our study, with only one pad measured by AFM for each pattern dimension.

The influence on Cu dishing of several other parameters, such as pattern shape (square, octagon, rectangle) or layout (aligned or staggered), was also investigated. Similar to pattern width, no significant variation of Cu recess was obtained. A possible explanation for this observation is the following. In addition to pattern width, the other main design parameter identified in the literature as having the most impact on metal recess is pattern density[2]. For the considered shapes and layouts however, there is no substantial change in either pattern width or pattern density, which could explain a negligible variation of the metal recess in our measurements.

Based on interface morphology characterization by SEM-TEM and 3D FIB/SEM tomography for samples with various typical Cu recess depths, there is evidence that a dishing threshold exists, beyond which only partial Cu-Cu bonding occurs. To estimate its magnitude and which

[1] Even for slightly larger pattern width, polishing pad compliance may also contribute to mitigate Cu dishing, by preventing full contact in the recess regions ("asperity shielding") [90].

[2] i.e. metal volume fraction (or equivalently, surface fraction).

features influence it remains to be clarified. It can also be noted that, in addition to Cu dishing depth, the opening angle formed between two dished pads could also play an important role on the bonding. In that regard, assuming a circular shape for the measured 2D line scans[1], a constant dishing for pads of decreasing width, as observed in the present study within the investigated width range, could be expected to lead to a decrease of the bonding quality, due to larger initial opening angle before bonding. This issue is further discussed in Section II.4.

A central question for hybrid bonding is whether partial bonding due to large Cu dishing may lead to critical electrical performance issues. Lhostis et al. [83] showed that for the type of hybrid bonding interconnect considered here, the most influential factor in terms of the electrical resistivity of the interconnection is the number of vias connecting the pad to the remainder of the interconnect stack, rather than the bonded surface area between the top and bottom hybrid bonding pads. Therefore, partial Cu-Cu bonding, even as critical as depicted in Figure II.10, is not believed to represent a major risk for the electrical performance of the 3D integrated device. Nevertheless, from a mechanical standpoint, partial bonding could lead to serious mechanical reliability issues (i) by decreasing the contact area between the bonding surfaces, leading to debonding at lower loads, and (ii) by generating large initial defects at the interface, present even after consolidation annealing and likely to evolve towards critical dimensions due to fatigue, electromigration or stress voiding phenomena, arising from thermal and electrical stressing during product life.

In the following, numerical modeling of Cu-Cu bonding during post-bonding annealing is carried out, aiming to provide a mechanistic understanding into the influence of several process and design parameters on bonding quality.

Further experimental characterization is required to determine the mechanical properties of the materials involved, thereby providing the necessary input parameters to carry out numerical modeling. In Appendix B, using the same process parameters as for the Cu/SiO_2 hybrid bonding process, the plastic response of an electroplated Cu thin film is derived by nanoindentation using a reverse analysis approach developed by Dao et al. [98].

II.4 Numerical study

In this section, a preliminary study is first carried out to investigate the influence on the results of various possible initial assumptions regarding the stress configuration (2D plane strain/axisymmetric or 3D) and boundary conditions (blocked or free boundaries, periodic or symmetric). Then, we focus on the influence of the assumed initial dishing shape and depth on Cu-Cu bonding. In addition, the influence of thermoelastic constants mismatch between the interconnect and the surrounding matrix is assessed, enabling to compare various metal/dielectric pairs in terms of sensitivity to an initial surface topography defect for successful hybrid bonding. In

[1] This was verified by least-squares fitting, with obtained correlation coefficients above 0.95

another section, the influence of interconnect geometry and layout is investigated. Finally, the scenario of misaligned Cu bonding pads is also investigated.

II.4.1 Problem formulation

II.4.1.1 Geometry

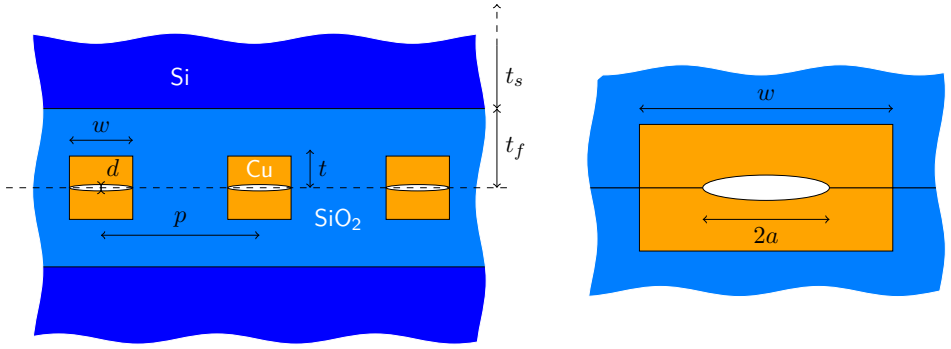


FIGURE II.11: Cross-sectional geometry of the Cu pattern.

We consider the bonding of Cu pads of width w, thickness t and pitch p embedded in a SiO_2 dielectric thin film of thickness $t_f = 2\,\mu m$, as depicted in Figure II.11. A dishing d corresponding to a radius of curvature $R = w^2/8d$ is assumed at the top surface of Cu pads to account for metal recess after CMP. Unless otherwise specified, the Cu interconnects are perfectly aligned with respect to each other. Since the interconnects form a periodic arrangement, we focus on a unit cell of width p. An aspect ratio $t/w = 0.1$ and dishing depth $d/t = 0.01$ are used as reference for the Cu pad geometry. The dielectric film is very thin compared to the Si substrate ($t_s = 775\,\mu m$), and the thickness ratio between the SiO_2 layer and the Cu pads is set to $t_f/t = 2$.

II.4.1.2 Boundary conditions and loading

Assuming a Cartesian coordinate system with the horizontal x_1–axis in the bonding plane and the x_2-axis perpendicular to the thin films, the prescribed boundary conditions are as follows (Figure II.12):

- blocked displacements in the x_1 direction along the left boundary (symmetry condition);
- planar constraint (uniform x_1-displacement) along the right boundary (periodicity condition);
- blocked displacements in the x_2 direction at the intersection between the bottom boundary and the axis of symmetry (rigid-body motion suppression);

A homogeneous thermal load ΔT from ambient to annealing temperature is prescribed. During bonding, unilateral "hard" contact is assumed and "sticking" friction ($\mu = \infty$), using the penalty method.

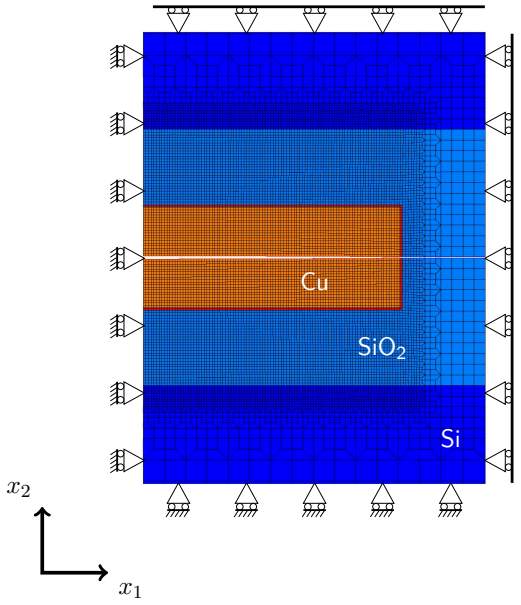


FIGURE II.12: Typical mesh and boundary conditions used in the numerical study (partial view).

II.4.1.3 Material properties

Linear elastic isotropic behavior is assumed for both the Si substrate and the SiO_2 dielectric matrix embedding the Cu interconnects. The corresponding thermoelastic constants are, respectively, $E_f = 130\,\text{GPa}$, $\nu_f = 0.17$, $\alpha_f = 2.6 \times 10^{-6}\,°\text{C}^{-1}$, and $E_s = 70\,\text{GPa}$, $\nu_s = 0.17$, $\alpha_s = 0.5 \times 10^{-6}\,°\text{C}^{-1}$.

Unless otherwise specified, an elastic-plastic response of Cu in the pads is considered. A J_2-flow theory is used for the description of the rate-independent Cu plastic response, with a Ludwik power-law hardening rule:

$$\sigma = \sigma_Y + K{\varepsilon_p}^N \tag{II.1}$$

where $E = 130\,\text{GPa}$, $\nu = 0.34$, $\alpha = 16.5 \times 10^{-6}\,°\text{C}^{-1}$. The assumed values for the yield stress and work hardening exponent are respectively $\sigma_Y = 200\,\text{MPa}$, $N = 0.25$ and $K = 1010\,\text{MPa}$, based on nanoindentation measurements on a blanket Cu thin film (Appendix B).

Numerical solutions are obtained by the finite element method using commercial code ANSYS Mechanical APDL 19.2. The domain is discretized using 8-noded quadrilateral elements (PLANE183) for a 2D generalized plane strain calculation. As depicted in Figure **??**, the mesh in the Cu pad region is uniform and comprised of at least 64×10 square elements of length $l_{\text{el}} \approx 40\,\text{nm}$. The nonlinear, quasi-static problem is solved using a Newton-Raphson iterative scheme. A careful control of the time step is performed to accurately describe contact between Cu surfaces, with at least a hundred load increments in each loading step.

II.4.2 Results

II.4.2.1 Preliminary study: stress configuration and boundary conditions

Stress configuration

In this section, a comparison is first conducted between various stress configurations for the Cu bonding pads, namely 2D axisymmetric, 2D plane strain, 2D generalized plane strain, and 3D with cylindrical Cu pad. The aim is to assess how these simplified descriptions influence the extent of the computed bonding area. Among the considered cases, some configurations merely reflect different geometrical shapes of the Cu pads and are used for comparison between two typical configurations, namely an infinitely long Cu line (2D plane strain case, as investigated by Beilliard et al. [99, 100]) and a cylindrical Cu pad (3D case closer to the application considered here). The other configurations are introduced to assess the validity of several simplifying assumptions. For instance, although the 2D plane strain assumption has been widely used in the literature to investigate thermal stresses in metal interconnects, 2D generalized plane strain (GPS) is in fact a more appropriate configuration for this purpose [101–103]. In addition, we will also assess the validity of the 2D axisymmetric configuration compared to the computationally more expensive case of a 3D periodic cell with a cylindrical Cu pad.

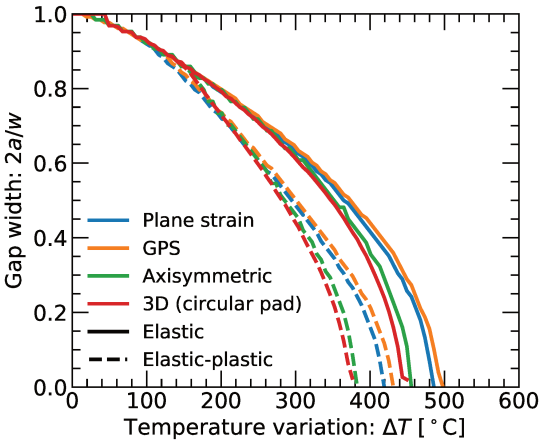


FIGURE II.13: Computed lateral extent of the gap between dished Cu interconnects as a function of temperature, for various stress configurations and for elastic or elastic-plastic Cu behavior.

To do so, we monitor the computed decrease in the lateral extent of the initial gap between dished Cu interconnects, due to contact between the bonding surfaces for increasing temperature. The results for each of the four considered stress configuration cases are plotted in Figure II.13. Several material behaviors are also investigated for the Cu pad, namely elastic and elastic-plastic. A steady increase of the bonded area is observed with increasing temperature, due to constrained thermal deformation of the Cu interconnects. A significant difference is observed between the Cu line and Cu pad cases. For Cu pads (2D axisymmetric or 3D configuration), the threshold temperature increase for full bonding ΔT_{cr} is about 10% lower compared to the

Cu line configuration (plane strain), both for elastic and elastic-plastic behavior. However, for each of these two cases the considered simplifying assumptions yield similar results. For the Cu line case, there is very little difference between plane strain and generalized plane strain for the bonding prediction, and similarly for the Cu pad case, the values obtained with 2D axisymmetric and 3D configurations are very close. Thus, in the following a 2D axisymmetric configuration will be used to simulate the bonding of the Cu pads considered in the experimental study.

Overall, the most significant difference is that between elastic and elastic-plastic Cu behavior. The influence of Cu plastic deformation on bonding becomes noticeable for a temperature increase of about 100 °C, leading to full bonding at a significantly lower critical thermal loading ΔT_{cr} compared to the elastic case, by approximately 15-20%.

Although the difference between the computed threshold temperature variations for full bonding may be regarded as minor between the different cases (below 20%), for a given temperature very large changes in the bonded area are observed. For instance, considering a prescribed temperature increase $\Delta T = 380\,°C$, close to the annealing temperature range used to strengthen the adhesion energy after the bonding process in the experimental study, full bonding is directly obtained for the Cu pad with elastic-plastic behavior, whereas for the Cu line behaving elastically the final lateral extent of the gap between the dished Cu surfaces accounts for about 45% of the pad.

Boundary conditions

Various sets of boundary conditions are also explored for the bonding of Cu pads (2D axisymmetric case):

blocked x_1 and x_2: normal displacements are blocked along the right boundary, and lower and upper edges of the unit cell;

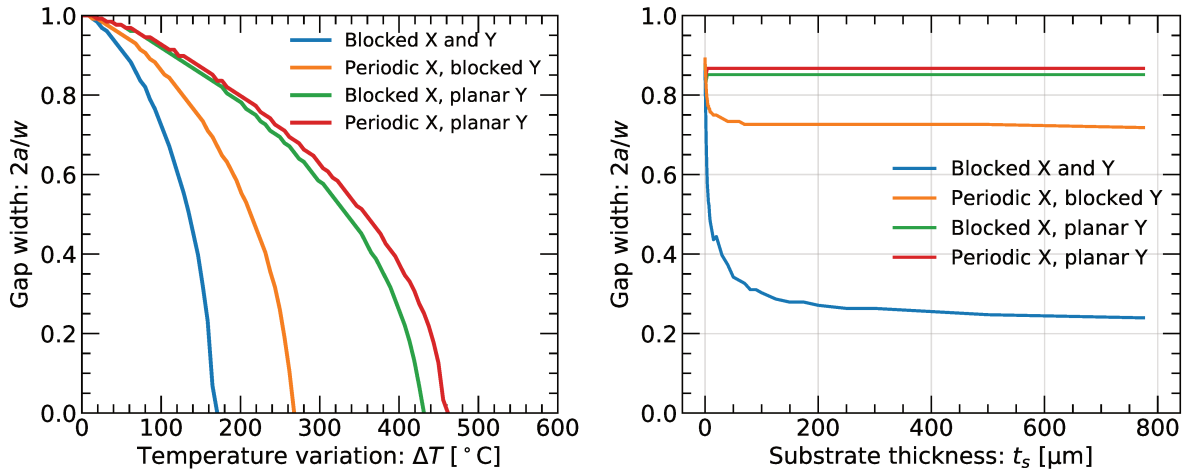
periodic x_1 blocked x_2: a planar constraint is prescribed on the right boundary, and normal displacements are blocked along the lower and upper edges;

These two configurations with blocked x_2 displacements correspond to a situation in which a compression would be exerted on the wafers during post-bonding anneal

blocked x_1 free x_2: normal displacements are blocked along the right boundary, and the upper edge is free. On the lower edge, vertical displacement is prevented for one node to suppress rigid-body motion;

periodic x_1 free x_2: a planar constraint is prescribed on the right boundary, and the upper edge is free. On the lower edge, vertical displacement is prevented for one node to suppress rigid-body motion.

The latter configuration is the reference case since the wafer is free to deform vertically during post-bonding anneal in the bonding process considered in this study.



(a) (b)

FIGURE II.14: Computed lateral extent of the gap between dished Cu interconnects as a function of (a) temperature and (b) Si substrate thickness for different sets of boundary conditions and for elastic or elastic-plastic Cu behavior.

The computed decrease in the lateral extent of the gap between dished Cu interconnects for increasing temperature is plotted in Figure II.14a for the four sets of boundary conditions. A marked influence of the condition prescribed on the upper (or lower) boundary is observed. Compared to cases with a free upper boundary, a much lower critical temperature is obtained for full bonding, by about 40-60%, when vertical displacements are blocked on the upper boundary. In comparison, the difference between free and periodic horizontal displacements is less significant. This difference, although still quite large between both cases with blocked vertical displacement, is only 5% between both cases with free vertical displacement. These trends are the same over the whole temperature range until full bonding.

Because of the very large aspect ratio of the considered unit cell ($2t_s/w = 155$) due to the thick Si substrate, it can be quite difficult to account for the whole thickness while maintaining acceptable computational times. Thus, in addition to the boundary conditions described above, we also investigate the influence of the modeled substrate thickness to simulate bonding between the dished Cu interconnects. The final gap width associated with a temperature increase $\Delta T = 150\,°C$ is plotted in Figure II.14b for the four sets of boundary conditions. For cases with blocked vertical displacement, the computed bonded area is extremely sensitive to the modeled substrate thickness, especially below 100 µm, i.e. 100 times the thickness of the SiO_2 dielectric layer. For cases with free vertical displacement on the other hand, the obtained value is constant whatever the modeled substrate thickness above approximately 10 µm.

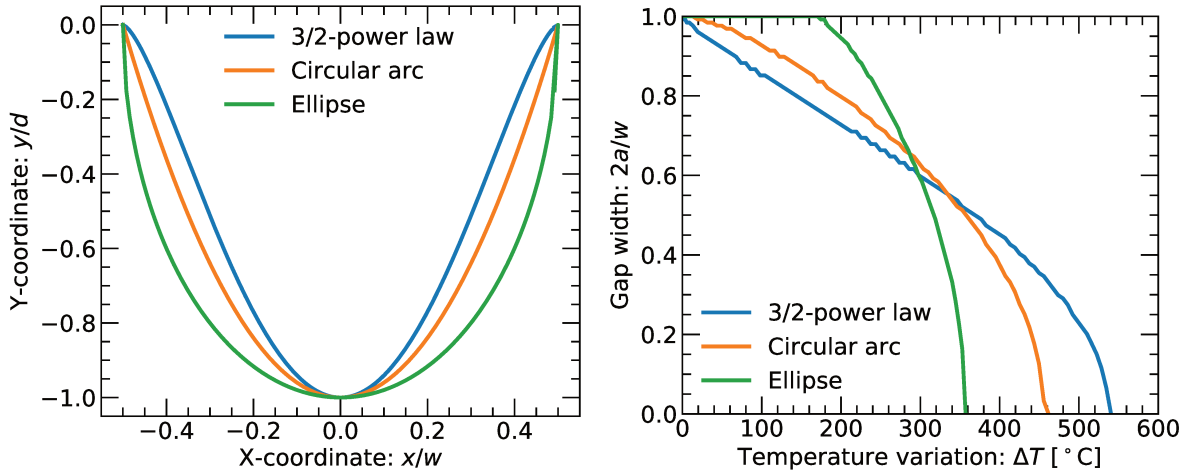
The above preliminary study shows that boundary conditions must be selected carefully for the considered problem. This is due to the nature of the prescribed loading, a homogeneous temperature increase leading to heterogeneous thermal expansion of the different materials (including the thick Si substrate). Thus, configurations with blocked vertical displacements across external boundaries result in additional constraint on thermal expansion of the substrate.

For the considered geometry, the difference between the *blocked* and *periodic* lateral boundary conditions is not significant. However, this is not expected to remain valid with decreasing interconnect pitch. Therefore, in the following, a configuration with periodic conditions for the right-hand edge of the considered half unit cell is adopted to simulate the bonding of Cu pads with initial dishing.

II.4.2.2 Influence of metal recess shape and depth

Metal recess shape

In this paragraph, various shapes are compared for the initial surface topography of the Cu bonding pads at constant metal recess depth, namely 3/2-power law, circular or semi-elliptical. These ideal profiles depicted in Figure II.15a were chosen for their distinct initial slope at the edge of the Cu interconnect, where bonding propagation starts. In that respect, the 3/2-power law and the semi-elliptical shapes are two limit cases, with respectively 0° and 90° initial slopes, while the circular profile on the other hand is an intermediate case, used as reference.



(a) (b)

FIGURE II.15: (a) Initial surface profile for the investigated dishing shapes, and (b) computed lateral extent of the gap between dished Cu interconnects as a function of temperature in the elastic case.

Similarly to the previous paragraphs, the decrease of the gap lateral extent for increasing temperature is plotted in Figure II.15b. Very distinct behaviors are obtained for the three profiles:

- at the beginning of bonding, i.e. for low temperature increase ΔT, the profiles that yield the lowest bonded area are those with the largest initial slope. This is especially noticeable for the semi-elliptical profile with 90° initial slope, for which bonding begins only above $\Delta T > 180°$ while with the other two profiles bonding is initiated right from the beginning of the temperature increase;

- for $\Delta T \approx 300\,°C$, a point is reached where the same gap lateral extent is obtained for the three profiles;
- for $\Delta T > 300\,°C$, i.e. at the end of bonding, the initial trend is reversed with the semi-elliptical profile yielding the largest bonded area.

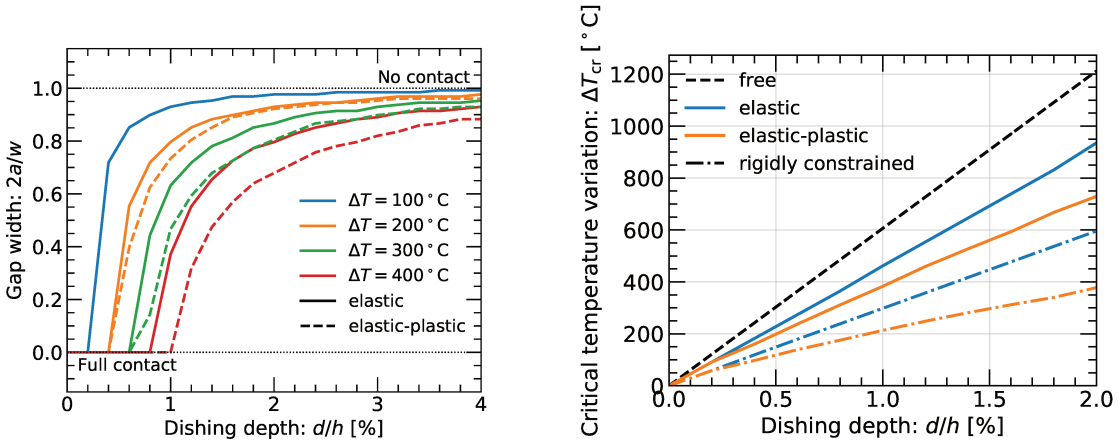
A first observation is that these changes in the rate of bonding with temperature seem to be driven by the local slope or curvature of the surface profile at the current position of the bonding front, with larger slope being detrimental for bonding. However, it can also be noticed that although these three surface profiles have the same initial dishing depth, the obtained critical temperature for full bonding ΔT_{cr} is significantly different in each case. For the semi-elliptical profile, bonding initiation is hindered by the vertical initial slope at the edge of the Cu pad and occurs only for a large temperature increase $\Delta T > 180\,°C$. And yet, compared to the 3/2-power law and circular profiles, the obtained critical temperature for full bonding is significantly lower, respectively by 33% and 15%. Therefore, in addition to local curvature, the initial surface profile is believed to play a significant role on the remainder of the bonding. The assumed underlying mechanism is a compressive stress build-up at the edge of the Cu pad in regions already in contact, exerting a resistance against thermal expansion and in turn making bonding more difficult at the center. This effect would then be expected to be especially acute for profiles with a flat initial slope, as is the case for the 3/2-power law profile, which yields the largest critical temperature for full bonding ΔT_{cr} among the three considered initial surface topographies.

Metal recess depth

In this paragraph, the influence of metal recess depth on Cu bonding for various annealing temperatures is investigated, assuming a circular recess shape. Dishing depths comprised between 0 and 4% of the Cu pad thickness are explored, corresponding to the range of experimentally measured values [61, 83, 84]. The final gap lateral extent computed for temperature increases ranging from 100 to 400 °C is plotted in Figure II.16a, both for the elastic and elastic-plastic Cu case.

A very sharp transition is observed between the *fully bonded* and *no contact* states, which is consistent with the assumption in the experimental study about the existence of a dishing threshold, beyond which only partial bonding occurs. As already observed in the preliminary study, the influence of Cu plasticity becomes noticeable above approximately $\Delta T > 100\,°C$. For plastically deforming Cu pads, an increase of the maximum acceptable dishing for full bonding is obtained, for instance by about 20% for a temperature increase of 400 °C. This effect become more noticeable as the prescribed temperature increase becomes larger.

The computed critical temperature for full bonding ΔT_{cr} is plotted in Figure II.16b as a function of dishing depth (for constant interconnect thickness), both in the elastic and elastic-plastic Cu cases. In the elastic case, a linear relationship is obtained between the two quantities. For comparison, the temperature increase required to compensate metal recess for the limit case of *free thermal expansion* of the Cu pad was also plotted in Figure II.16b. Based on the definition

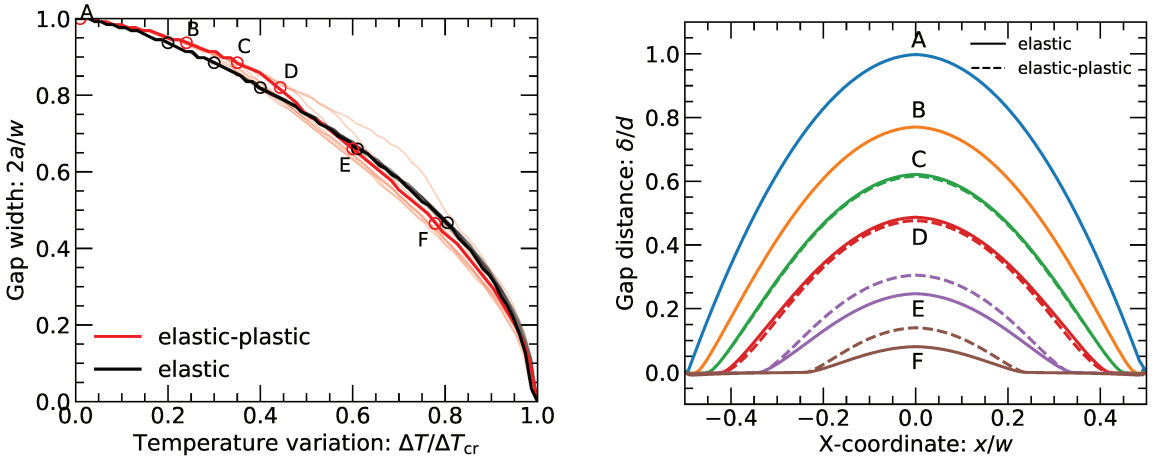


(a) (b)

FIGURE II.16: (a) Computed gap lateral extent as a function of dishing depth for several temperature variations, and (b) threshold temperature increase for full bonding as a function of dishing depth for several configurations.

of thermal deformation, the slope of this linear curve is given by:

$$\Delta T_{cr} = \frac{1}{\alpha}\varepsilon_{cr}^{th} = \frac{1}{\alpha}\frac{d}{h}$$



(a) (b)

FIGURE II.17: (a) Computed gap lateral extent as a function of temperature for various dishing depths in the elastic and elastic-plastic cases, and (b) corresponding gap distance at different stages of the bonding.

Comparison indicates that such a simplified analysis naturally does not enable to estimate the temperature increase needed to compensate the initial dishing. Indeed, the critical temperature ΔT_{cr} is overestimated by about 25% for the considered pad geometry, warranting the need for finite element analysis. When plastic deformation of the Cu pad is accounted for, this disparity rises to nearly one third. Therefore, it can be concluded that thermal expansion at the center of the Cu pad is greatly assisted by:

- constrained deformation in the lateral direction, due to the surrounding dielectric matrix with low CTE;
- plastic deformation of the Cu pad, due to the stresses resulting from constrained deformation.

This strong influence of constrained deformation in assisting dishing compensation becomes further evident by considering the other limit case, namely *rigid thermal expansion* of the Cu pad embedded in an infinitely stiff matrix having zero thermal expansion. This case was also included in Figure II.17a both for elastic or elastic-plastic Cu behavior, and results in a sharp decrease of the critical temperature for full bonding, by 37 and 50% respectively for the elastic and elastic-plastic cases.

In addition to the effects discussed above, another difference between the elastic and elastic-plastic cases can be noticed by plotting the computed decrease in the lateral extent of the initial gap between dished Cu pads for increasing temperature. The obtained values for various dishing depths within the investigated range are superimposed in Figure II.17a, both for the elastic and the elastic-plastic case. To enable proper comparison, the temperature increase ΔT is normalized by the computed critical temperature for full bonding ΔT_{cr}:

- in the elastic case, the curves associated with the different dishing depths match perfectly, in agreement with the previous result that the critical temperature ΔT_{cr} depends linearly on the dishing depth;
- in the elastic-plastic case however, slight departures from the reference elastic curve are observed, with lower "bond advance rate" at the beginning of thermal loading and then slightly larger advance compared to the elastic case. For increasing dishing depth, these departures are lower in magnitude and the *low-bond-advance* phase accounts for a smaller proportion of the total loading sequence.

The obtained variations in the elastic-plastic case are believed to reflect the influence of local plastic deformation near the bonding front, resulting in a different stress distribution for a given temperature, as well as a different deformed shape of the Cu pad, as illustrated in Figure II.17b.

II.4.2.3 Influence of elastic constants mismatch

In the previous section, the thermoelastic mismatch between metal interconnect and dielectric matrix was shown to enable (and enhance) dishing compensation during thermal loading, due to transverse deformation of the constrained Cu pad. However, the respective contributions of Young's modulus and the CTE difference on this ability is not clear. Aiming to provide more information on the influence of elastic constants mismatch, various metal/dielectric pairs are explored for hybrid bonding with an initial dished shape of the metal bonding surface. For simplicity, this study is restricted to elastic behavior, and a single dishing depth ($d = 0.01h$) is considered.

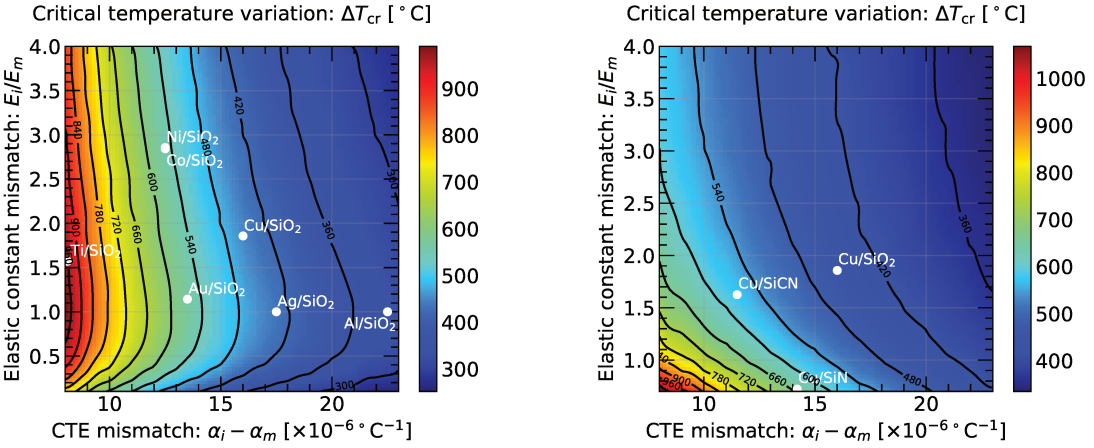


FIGURE II.18: Contour maps of the computed critical temperature for full bonding as a function of elastic constant and CTE mismatch between the interconnect and the dielectric matrix: (a) for an interconnect with varying thermomechanical properties embedded in a SiO_2 dielectric, and (b) for a Cu interconnect embedded in a dielectric with varying thermomechanical properties. Several metal/dielectric candidates are indicated by red circles on the contour map.

Contour maps of the computed critical temperature for full bonding, as a function of elastic constant mismatch E_i/E_m and CTE mismatch $\alpha_i - \alpha_m$ (with the i and m subscripts denoting interconnect and matrix properties) are presented in Figure II.18a and Figure II.18b, respectively for (i) fixed dielectric layer material (SiO_2) and varying metal interconnect properties, and (ii) fixed interconnect material (Cu) and varying dielectric layer properties. It can be noticed that different trends are obtained between these two cases for the same mismatch. This is so because in addition to the properties of the metal interconnect and dielectric layer, the elastic properties of the Si substrate are also believed to play a significant role on the bonding behavior for the considered configuration.

For fixed dielectric layer material (SiO_2) and varying metal interconnect properties (Figure II.18a), the most influential parameter on the threshold temperature increase required for full bonding is clearly the CTE mismatch between the interconnect and the dielectric layer $\alpha_i - \alpha_m$. The elastic constants mismatch on the other hand has little influence for a given CTE mismatch, especially for a stiff interconnect ($E_i/E_m \gg 1$). It can be noted however that having an interconnect with a Young's modulus close to that of the SiO_2 dielectric layer ($E_i/E_m \approx 1$) seems to be detrimental for bonding, which is related in this case to a lack of constraint (see previous section). For fixed interconnect material (Cu) and varying dielectric layer properties (Figure II.18b), the elastic mismatch also plays a minor role on the threshold temperature for complete bonding for large E_i/E_m. However, contrary to the previous case, for low CTE mismatch and intermediate to small E_i/E_m within the investigated range, the influence of the elastic mismatch becomes much more noticeable, with a large increase of the critical temperature for full bonding.

In Figure II.18, the locus of various metal/interconnect pairs has been reported on the elastic mismatch contour map. While some of these material pairs have actually been demonstrated

for hybrid bonding applications, e.g. Ni/SiO_2 [104], Au/SiO_2 [105] or Cu/SiN [106], others are purely prospective and were considered only for comparison purposes. Among the various material pairs, the Cu/SiO_2 configuration used in the present study is found to have a better potential for dishing compensation, for instance compared to Au/SiO_2 with lower CTE mismatch and Cu/SiN with lower elastic mismatch[1]. Although the Al/SiO_2 pair is shown to be very beneficial for dishing compensation, from a practical viewpoint Al is very unlikely to be considered due to its greater sensitivity to electromigration compared to Cu, which would considerably decrease the reliability of such an interconnect for hybrid bonding. Finally, it must also be highlighted that beyond the ability to compensate for surface topography defects, other critical issues for bonding must also be considered for the investigated metal/dielectric pairs, such as their respective CMP selectivity and "bondability" (e.g. oxidation properties), which are beyond the scope of this study.

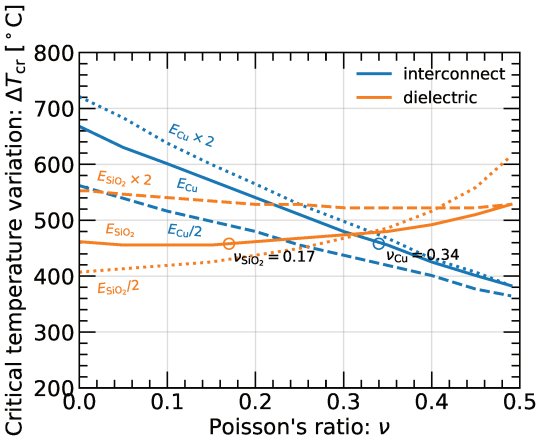


FIGURE II.19: Computed threshold temperature increase for full bonding as a function of interconnect and dielectric Poisson's ratio.

In addition to Young's modulus and CTE, the influence of the Poisson's ratio of the metal interconnect or dielectric layer is also investigated. The threshold temperature increase for full bonding ΔT_{cr} is plotted as a function of Poisson's ratio in Figure II.19, for both the metal interconnect and the dielectric layer. For simplicity, Young's modulus and CTE are maintained constant, equal to the values used so far for a Cu interconnect and SiO_2 dielectric. A significant drop of the critical temperature is observed with decreasing interconnect Poisson's ratio ν_i, in agreement with the assumption of a dishing compensation mechanism by transverse deformation and related constraint effect of the matrix surrounding the Cu pad (see Section II.4.2.2). It can also be noted that a decrease of about 20% is obtained between the reference case ($\nu_i = 0.34$) and the incompressible case ($\nu_i \to 0.5$), which is quite close to the critical temperature decrease observed between the elastic and elastic-plastic cases in Figure II.16b.

[1]For reference [106], it must be pointed out that a radically different configuration to that considered here was used, with protruding interconnects on one side (Cu "nails") and dished interconnects on the other, which can be assumed to enable to somewhat alleviate the influence of surface topography compared to the configuration discussed here.

The Poisson's ratio of the dielectric layer also has an influence on the bonding behavior. To investigate any coupling between the respective impacts of Poisson's ratio and Young's modulus on the critical temperature for full bonding, several ranges of Young's moduli are explored for varying Poisson's ratio. Contrary to the interconnect case, a complex coupling is observed for the dielectric layer. For low Young's modulus, a significant increase of the critical temperature for full bonding is obtained, whereas for high Young's modulus it has only little influence, with a very slight decrease of the critical temperature for increasing Poisson's ratio. This complex behavior could in part explain the observed difference in the contour maps of Figure II.18 between the fixed dielectric material and fixed interconnect material cases. The influence of the dielectric layer elastic properties on the bonding behavior is believed to result from (at least) two effects:

- a constraint effect of the dielectric matrix on the metal interconnect for large CTE mismatch $\alpha_i - \alpha_m$ and large Young's modulus (i.e. $E_i/E_m \ll 1$), assisting the bonding by enabling dishing compensation;
- interaction of the dielectric matrix with the Si substrate, due to imposed displacement from the thick substrate. The resulting lateral expansion prescribed to the dielectric layer under temperature increase in turn leads to a transverse contraction due to the Poisson effect, with an amplitude modulated by a factor of $-\nu_m/E_m$. For large Young's modulus of the dielectric layer ($E_i/E_m \ll 1$), transverse contraction below the interconnect may be detrimental for bonding by pulling the bonding surfaces apart.

The influence of the dielectric layer on the bonding behavior would then be the result of a competition between these two effects. Thus, for fixed dielectric layer material and varying interconnect properties (Figure II.18a), the influence of substrate expansion on the dielectric layer does not vary and the constraint effect is predominant over a wide range of parameters, while for fixed interconnect material and varying dielectric layer properties on the other hand, the interaction effect with the substrate also plays a significant role, leading to a more complex behavior with a strong coupling between the respective impacts of the elastic constants.

II.4.2.4 Influence of interconnect pitch and geometry

Due to circuit miniaturization, the pitch distance between hybrid bonding interconnects is bound to decrease, and the interconnect surface fraction to increase. In the following section, we explore the influence of such trends on the ability to bond the Cu surfaces in presence of an initial dishing. The two main design parameters for hybrid bonding interconnects are investigated (separately), namely interconnect aspect ratio and interconnect surface fraction.

Interconnect aspect ratio

We define the interconnect aspect ratio as the thickness-to-width ratio h/w (Figure II.11). Thus, high aspect ratio refers to interconnects with a slender shape perpendicular to the substrate,

and conversely a low-aspect-ratio interconnect has small thickness compared to the planar dimensions. In practice, to enable ever-increasing electrical connection density, the interconnect width w is more likely to undergo major changes compared to the thickness h. Starting from a reference at $w = 5\,\mu m$, consistent with the test chip investigated in Section II.3.1, the influence of a width decrease (at constant thickness $h = 0.5\,\mu m$) on the threshold temperature increase for full bonding is investigated in Figure II.20. Similarly to the previous paragraphs, both elastic and elastic-plastic behaviors are considered for the Cu interconnects.

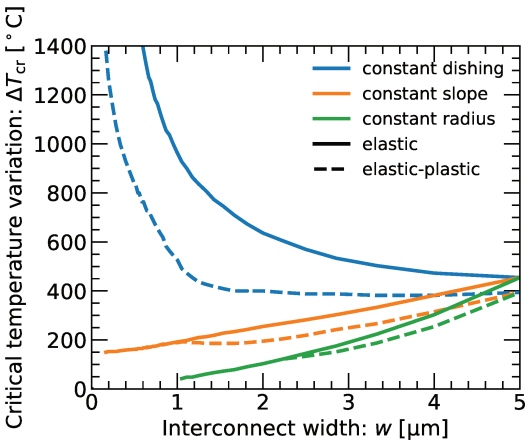


FIGURE II.20: Computed threshold temperature variation for full bonding as a function of interconnect width for several assumed relationships between interconnect width and dishing depth.

Three scenarios are explored to account for several possible relationships between pattern width w and Cu dishing depth d after CMP:

Case 1 (constant depth): the dishing depth remains constant for decreasing interconnect width, in accordance with the experimental observations in Section II.3. An arbitrary value of $d = 5\,nm$ is used.

Case 2 (constant slope): the dishing depth decreases linearly for decreasing interconnect width, or equivalently the slope of the dishing profile at the edge of the bonding surface remains constant, i.e. the ratio d/w is constant. This is closer to the behavior reported in the literature (for parallel lines, and for larger dimensions compared to the investigated range in the present study however, as discussed in Section II.3.4);

Case 3 (constant curvature): the radius of the assumed circular dishing profile remains constant for decreasing interconnect width, i.e. the ratio d/w^2 is constant. This case is included for completeness and comparison purposes.

Very distinct behaviors are obtained between case 1 and cases 2-3. For constant dishing, a very sharp increase of the critical temperature ΔT_{cr} is observed for decreasing interconnect width. Conversely, for constant slope or radius the opposite trend is obtained with a linear or quadratic decrease (respectively).

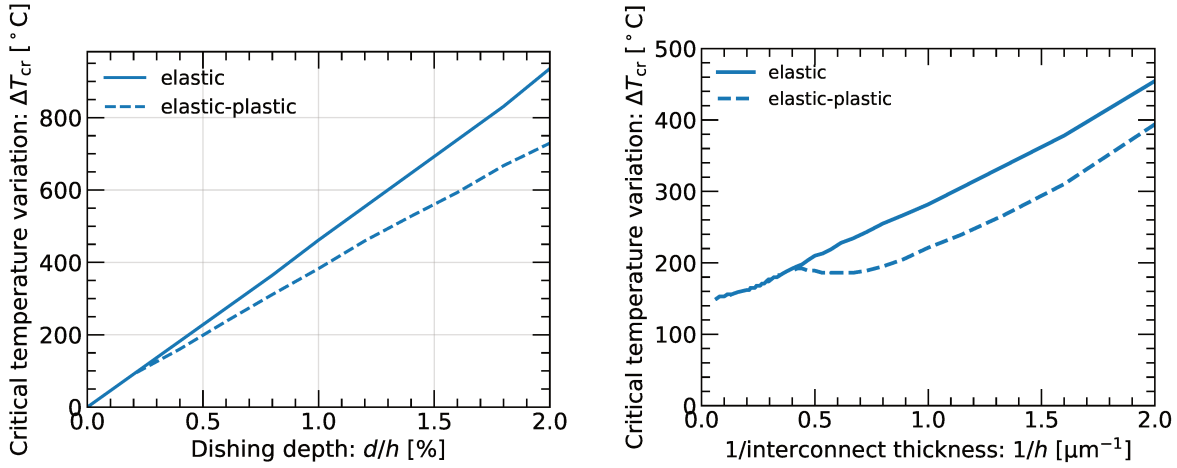


FIGURE II.21: Influence on the computed critical temperature for full bonding of (a) the dishing depth, for constant interconnect width and thickness, and of (b) the interconnect thickness, for constant dishing depth and interconnect width.

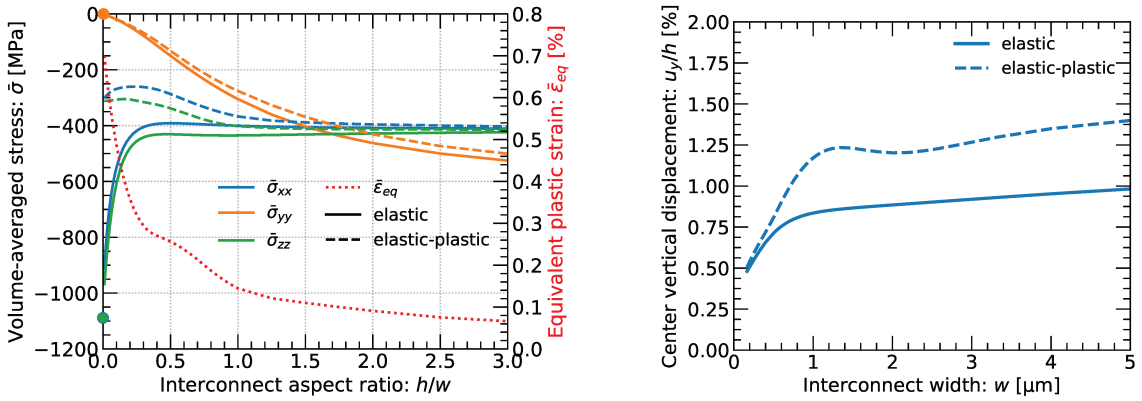
We focus first on cases 2 and 3 (resp. constant slope and constant radius). In Section II.4.2.2, it has been shown that the threshold temperature increase ΔT_{cr} varies linearly with the dishing depth d, for constant interconnect width w and thickness h (Figure II.21a). In addition, here it can also be noticed that ΔT_{cr} is inversely proportional to the interconnect thickness[1] for increasing values from the reference value $h = 0.5\,\mu m$, when d and w are kept constant (Figure II.21b). In conclusion, the critical temperature ΔT_{cr} depends linearly on the dishing-to-thickness ratio d/h.

In Figure II.20, the interconnect thickness was kept constant, and the interconnect width was varied. Thus, for constant slope d/w, the dishing-to-thickness ratio d/h varies linearly with w and as a result ΔT_{cr} is proportional to w. Similarly, for constant curvature d/w^2, the dishing-to-thickness ratio d/h varies as w^2 and thus a quadratic relationship is obtained between the critical temperature ΔT_{cr} and the interconnect width w. For the same reason, it could then be expected for case 1 (constant depth) that the critical temperature ΔT_{cr} be independent from the interconnect width w. However, instead a large increase is observed in Figure II.20 for decreasing interconnect width. This result suggests that one or several additional effects must be at play in this case, besides that of the dishing-to-thickness ratio d/h. Two possible influential parameters may be considered:

- the slope at the edge of the dishing profile d/w, with a larger slope here for decreasing width at constant dishing depth;
- the interconnect aspect ratio h/w, determining the stress magnitude and distribution in the interconnect [107], with a larger aspect ratio here for decreasing width at constant thickness.

[1] It can be noted that a slight deviation from this behavior is observed for large thickness ($h \gg w$).

The first parameter, namely the dishing slope, is not believed to play a significant role on the bonding behavior in that case. Indeed, as discussed in Section II.4.2.2, for constant dishing depth a larger slope at the edge of the interconnect does not necessarily lead to an increase of the critical temperature, and the opposite trend was actually observed in Figure II.11 for the considered metal recess shapes. To investigate the influence of the second parameter (namely interconnect aspect ratio) on the bonding behavior, the volume-averaged stress components and the vertical displacement at the center of the interconnect are plotted in Figure II.22 for the simplified case of a single interconnect layer with a free, plane top surface and a temperature increase of $\Delta T = 400\,°\mathrm{C}$.



(a) (b)

FIGURE II.22: For *partially constrained* configuration: (a) influence of aspect ratio on the volume-averaged stresses in the interconnect, and (b) computed vertical displacement at the center of the interconnect as a function of interconnect width.

For very low aspect ratio, the behavior of the interconnect converges towards that of a thin film (Figure II.22a), with an equibiaxial stress state and a transverse displacement given by:

$$\frac{u_y^0}{h} = -\frac{2\nu}{1-\nu}\Delta\alpha\Delta T \tag{II.2}$$

with $\Delta\alpha = \alpha_s - \alpha$ the CTE difference between the Si substrate and the interconnect.

For high aspect ratio, the nonzero transverse stress component σ_{zz} progressively increases in the interconnect, and the in-plane stress components decrease towards a constant value (Figure II.22a). In addition, as the stress distribution in the interconnect shifts from plane stress to a nearly hydrostatic state, the average plastic deformation progressively vanishes. Due to these changes in the stress configuration in the interconnect, the vertical displacement at the center of the interconnect decreases with increasing aspect ratio (Figure II.22b).

The observed increase of the critical temperature ΔT_{cr} for decreasing interconnect width in Figure II.20 at constant thickness and constant dishing depth (case 1) is believed to result from this influence of interconnect aspect ratio on the stress configuration. For cases 2 and 3 with varying dishing depth on the other hand, the influence of the interconnect aspect ratio h/w on

the bonding behavior is then likely negligible compared to the dishing-to-thickness ratio d/h, and was therefore observed only for case 1 with constant dishing.

Interconnect surface fraction

The above results on interconnect aspect ratio were obtained at constant metal surface fraction ($p/w = 0.25$). However, in addition to changes in the pattern dimensions, the pitch distance p between hybrid bonding interconnects can also be expected to decrease for the upcoming generations, aiming to increase the density of electrical connections. In this paragraph, the influence of the interconnect surface fraction w^2/p^2 on the bonding behavior is investigated, for constant interconnect aspect ratio ($h/w = 0.1$).

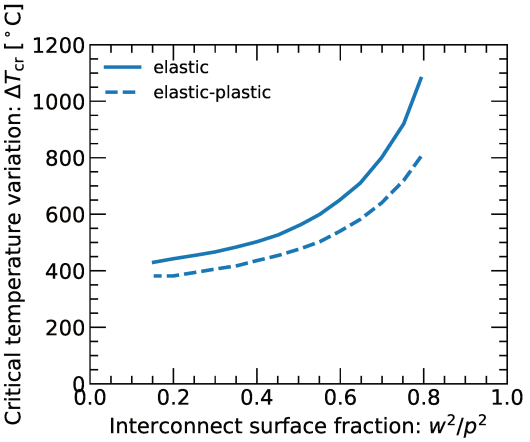


FIGURE II.23: Influence of interconnect surface fraction on the computed threshold temperature increase for full bonding.

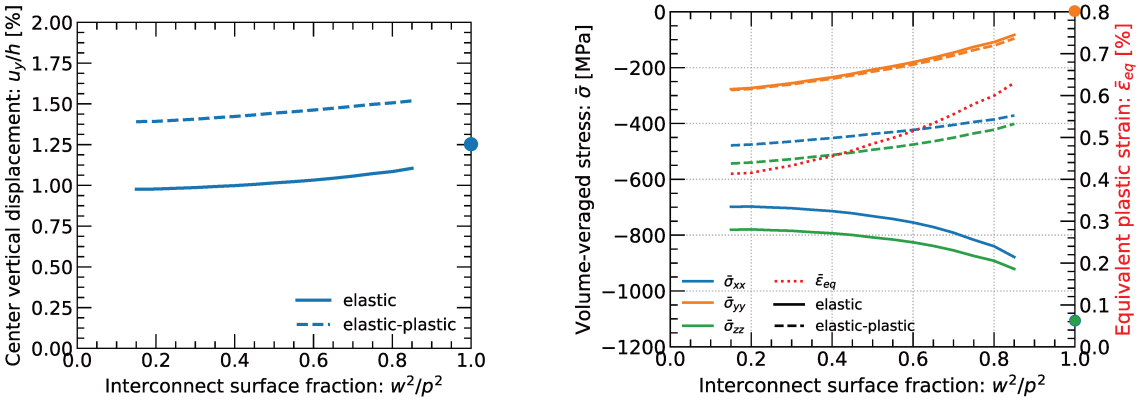
A very large increase of the critical temperature for full bonding ΔT_{cr} is observed for increasing interconnect surface fraction w^2/p^2 (Figure II.23). Considering the limit case of low surface fraction ($p \gg w$), the threshold temperature increase converges towards a constant value of about 400 °C, while for large surface fraction ($p \to w$) it seems to increase asymptotically to infinity. Similarly to the previous paragraphs, a lower critical temperature is systematically obtained in presence of plastic deformation in the Cu interconnect.

To explain the observed trends, two simplified cases are studied, for a flat interconnect surface profile ($d = 0$) and a prescribed thermal loading of $\Delta T = 400$ °C:

Partially constrained expansion: only one side of the bonding assembly is considered, i.e. the top surface of the interconnect layer is a free surface. This case is the same as in Figure II.22 in the previous paragraph.

Fully constrained expansion: the two sides of the bonding assembly are considered, and the interconnect surface profile is flat, i.e. the top surface of the interconnect layer is constrained.

These two limit cases are assumed to account for different stages of the bonding process for interconnect with a dished surface profile, namely the early stages of the bonding for the *partially constrained* case, and the very end of the bonding for the *fully constrained* case.



(a) (b)

FIGURE II.24: (a) For *partially constrained* configuration, computed vertical displacement at the center of the interconnect as a function of interconnect width, and (b) for *fully constrained* configuration, influence of aspect ratio on the volume-averaged stresses in the interconnect.

The vertical displacement at the center of the interconnect is plotted in Figure II.24a as a function of interconnect surface fraction for the *partially constrained* case. A very slight increase of the center displacement is observed for increasing w^2/p^2, converging for $p \to w$ towards the transverse displacement that would be obtained for a Cu thin film. The reason for such a small increase is an already quite low aspect ratio for the reference geometry, with $h/w = 0.1$. The volume-averaged stress components in the interconnect with varying surface fraction for the *fully constrained* case are also plotted in Figure II.24b. Similarly, a modest increase of the in-plane stresses is observed for increasing w^2/p^2 and the behavior converges towards that of a thin film for $p \to w$, with vanishing transverse stress and an equibiaxial stress state in the interconnect.

The latter observations seem to point towards an increase of the bonded area for a given prescribed thermal excursion ΔT, or equivalently a decrease of the critical temperature for full bonding ΔT_{cr} as the interconnect surface fraction w^2/p^2 increases, i.e. as the behavior of the interconnect converges towards that of a thin film. And yet, in Figure II.23 a very sharp increase of ΔT_{cr} is observed for $w^2/p^2 \to 1$. This is because although the interconnect behaves as a thin film, which is beneficial in terms of transverse displacement and plastic deformation, for $w^2/p^2 \to 1$ the remaining dielectric "ligament" between the metal patterns progressively ceases to exert vertical constraint on the interconnects. The consequence is an overall transverse expansion of the interconnect, both on the edge and at the center, thereby no longer enabling dishing compensation and leading to a very large temperature increase necessary to bond the surfaces.

For large surface fraction ($w^2/p^2 \to 1$), the situation is quite different to the case of an interconnect with low aspect ratio ($h/w \to 0$) addressed in the previous section, although for both cases a behavior close to that of a thin film is obtained. This is because for low aspect ratio,

the interconnects form individual thin film islands immersed in a dielectric matrix, whereas for large surface fraction a bulk Cu thin film is obtained as the volume of dielectric vanishes. Conversely, for low surface fraction ($w^2/p^2 \to 0$) the threshold temperature increase ΔT_{cr} converges towards a constant value because the obtained configuration is equivalent, not to a bulk dielectric film, but to a secluded Cu inclusion in the dielectric matrix. As the w^2/p^2 ratio decreases, the interaction between neighboring interconnects in the periodic pattern vanishes. Because this segregated inclusion has a moderate aspect ratio ($h/w = 0.1$), its thermal expansion in the vertical direction is not constrained as much as in the previous case of an interconnect with large aspect ratio ($h \gg w$), and thus ΔT_{cr} does not decrease.

II.4.2.5 Influence of interconnect overlap

In this paragraph, the influence of an overlap between the interconnect pairs is investigated. This type of bonding defect due to misalignment during the pre-bonding phase of the bonding process has already been observed, to different degrees, in the experimental study. For the most critical cases, the overlapping surfaces may account for a significant fraction of the targeted bonding area. In addition to the bonding area decrease that inevitably results from such a defect, we may wonder whether a misalignment between dished Cu surfaces can be detrimental for the bonding of the remaining area available for electrical contact.

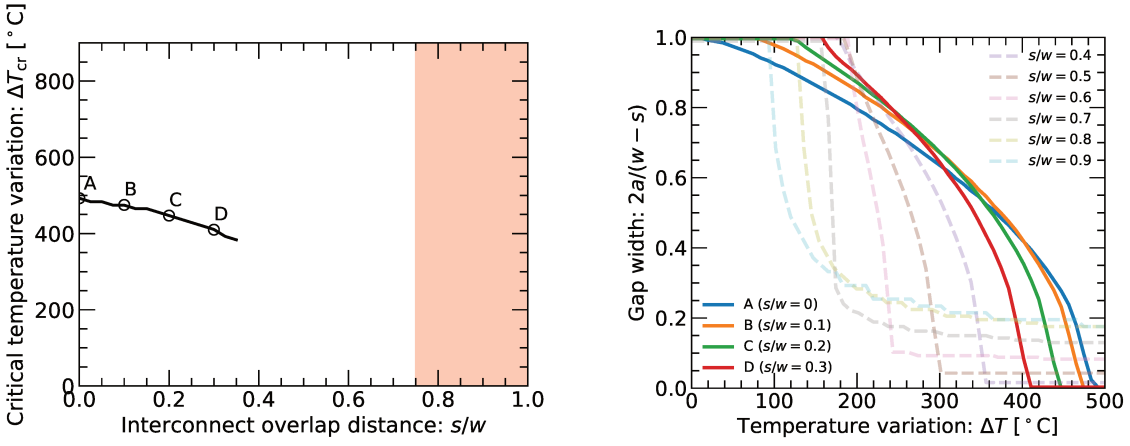


FIGURE II.25: (a) Computed threshold temperature increase for full bonding as a function of interconnect overlap distance, and (b) corresponding gap distance for selected cases.

To monitor the influence of this parameter on the bonding behavior, the critical temperature for full bonding was plotted as a function of the overlap distance s/w (normalized by interconnect width) in Figure II.25a. Misalignment distances accounting for more than 75% of the interconnect width (red domain in Figure II.25a) are not investigated. Indeed, for such a large misalignment distance the contact surface would become small even in comparison to the via/metal contact area, and thus a large increase of the interconnect electrical resistance would be expected [83, 84]. In the remaining interval, two regimes can be distinguished. Below $s/w \approx 0.4$,

a significant decrease of the critical temperature required for full bonding[1] is observed, by about 20%. On the other hand, above $s/w \approx 0.4$, no values are obtained for ΔT_{cr} within the investigated range.

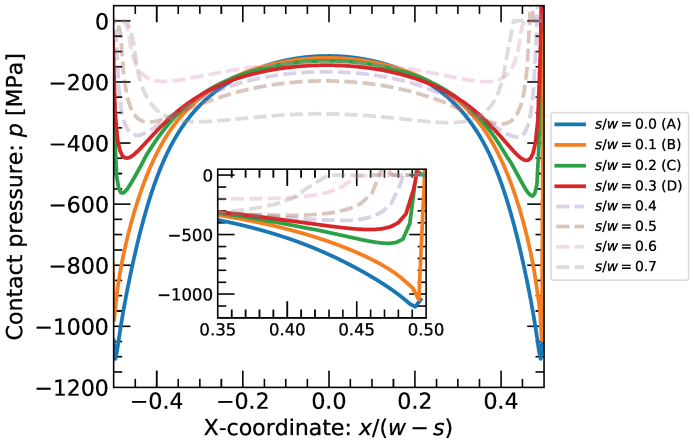


FIGURE II.26: Contact pressure profiles for several interconnect overlap distances.

The gap width decrease for increasing temperature is plotted as a function of the overlap distance in Figure II.25b. From the obtained curves it can be inferred that, for overlap distances above $s/w \approx 0.4$, bonding front propagation stops before full contact can be obtained between the dished Cu surfaces. However, since constant mesh size was used for the analysis, with a lower number of elements at the Cu-Cu interface as the overlapping distance increases, a decrease of accuracy for the computed contact area could explain the observed behavior. To ensure that it was not the case here, the final contact pressure along the bonding surface was plotted in Figure II.26 for various overlapping distances. Above $s/w \approx 0.4$, free surfaces are indeed seen to appear at the periphery of the Cu-Cu contact region. Thus, for too large a misalignment between the dished Cu surfaces full bonding cannot be obtained, even for a very large temperature increase. The latter observation is in agreement with the saturation of the bonded area observed in Figure II.25b, and is explained by the presence of a cavity in the Cu-SiO_2 overlap region, extending to the periphery of the Cu-Cu bonding surface for large misalignment.

II.4.3 Discussion

In the foregoing numerical study, the bonding of Cu hybrid bonding interconnects during thermal annealing has been simulated. After a preliminary study on the sensitivity of the obtained results to the assumed boundary conditions, the influence of various morphological or design parameter on the bonding behavior has been investigated.

A first observation from the preliminary study is that boundary conditions must be selected carefully depending on the investigated configuration. A strong dependence on boundary conditions

[1] Naturally, "full bonding" in that context refers to the effective bonding area, which is actually only a fraction of the targeted bonding area due to interconnect overlapping.

was also observed by Beilliard et al. [99], who simulated two hybrid bonding configurations, with free or blocked vertical displacements along the model top and bottom boundaries. This dependence is due to the nature of the considered problem involving constrained thermal deformation of the interconnect. For this study, the wafers are free to deform vertically during post-bonding anneal, and thus a *free* boundary condition was used. Whether a configuration with blocked vertical displacements is suitable to account for *thermocompression* bonding requires further investigations. Thermocompression experiments were not carried out here, and thus such a study was not addressed in the present work.

The influence of metal recess shape and depth was also investigated. A striking result is the strong result dependence of the temperature increase required for full bonding on the Cu recess shape, even for constant dishing depth. Here, based on AFM measurements of surface profiles, a circle arc profile was deemed a good approximation for our 2D axisymmetric model. An even more important question then is what kind of surface topography could be used in the 3D case, for instance for squares pads which is a far more likely configuration in practice compared to the round pads considered here. Another important question is the influence of other surface topography defects, such as oxide rounding, copper step, as well as that of Cu surface roughness on the bonding behavior. The influence of surface roughness was recently addressed in a simulation study by Wlanis et al. [108], in addition to the effect of Cu creep deformation. In the latter study, the authors also pointed out the discrepancy between the 2D and 3D modeling approaches, when using the current bonded area as a comparison criterion between different configurations. For this reason, in the present study the threshold temperature increase for full bonding was used as a criterion instead.

Regarding the influence of hybrid interconnect design on the bonding behavior, very distinct trends are obtained depending on the initial assumption on the evolution of the dishing depth for varying interconnect width. For constant dishing depth, decreasing the interconnect width is highly detrimental for bonding, whereas for a dishing depth proportional to the interconnect width, smaller interconnects lead to smaller dishing and therefore lower temperature required for full bonding. For the interconnect width range considered here, the actual dependence of the dishing profile on interconnect dimensions is not known. Thus, further experimental or simulation work on the CMP process is needed to gain better understanding into this issue. The influence of metal surface fraction was also studied. A large metal surface fraction was observed to lead to a more difficult bonding, requiring higher temperatures. This is attributed to vanishing vertical constraint at the periphery of the interconnects, attenuating the expansion difference between the center and the edge and thus hindering dishing compensation. Therefore, depending on the relationship between dishing depth and interconnect width, increasing the density of interconnections in future chip generations (by decreasing interconnect width or pitch distance) could potentially lead to bonding issues.

Besides gaining better control on hybrid bonding surface topography after CMP, an influential parameter that could enable to mitigate these potential issues is the choice of materials for the metal/dielectric pair. Naturally, a metal/dielectric having a large CTE mismatch is beneficial for

dishing compensation. For the interconnect material, the critical temperature for full bonding was shown to increase if the Young's modulus close to that of the dielectric matrix. For the dielectric material, a complex coupling with the substrate was observed, with an enhancement of dishing compensation for large Young's modulus compared to both the interconnect and the substrate, as discussed in Section II.4.2.3.

In the latter results, adhesive forces between Cu surfaces were not accounted for. In future work, surface interactions at the bonding interface could be included by using a cohesive zone approach, similarly to a recent study by Beilliard et al. [100]. In particular, the influence of parameters such as metal recess shape and depth, as well as interconnect aspect ratio on the bonding behavior needs to be evaluated in presence of interaction forces.

II.5 Conclusion

In this chapter, the influence of CMP-induced surface topography on the bonding behavior of Cu hybrid bonding interconnects was studied. In addition to Cu recess depth and shape, various interconnect geometries and layouts, or metal/dielectric material pairs have been explored, aiming to mitigate potential bonding issues.

After an overview of the physical mechanisms underlying the hybrid bonding process for chip stacking and the current industry challenges associated with this process, bonding experiments were carried out to investigate the influence of Cu interconnect geometry and layout on the resulting surface topography after planarization of the Cu/SiO_2 patterned surface. For the considered range of interconnect dimensions and layouts, a negligible influence of these parameters on the obtained Cu surface topography is reported. In turn, the influence of the obtained surface topography on the bonding interface quality was examined. Based on interface morphology characterization by FIB/SEM tomography for samples with various Cu recess depths, there is evidence that a dishing threshold exists, beyond which only partial Cu-Cu bonding occurs. To estimate its magnitude and which features influence it remains to be clarified. For most cases, the observed bonding defects are not believed to pose an immediate risk for the electrical performance of the 3D integrated device, for instance due to an increase of the interface resistance. Nevertheless, from a mechanical standpoint, partial bonding could lead to serious mechanical robustness issues, e.g. by leading to debonding at lower loads, or by generating large initial defects at the interface, likely to evolve towards critical dimensions due to fatigue, electromigration or stress voiding phenomena during product life.

Thus, aiming to provide a mechanistic understanding into the influence of several process and design parameters on bonding quality, numerical modeling of Cu-Cu bonding during post-bonding annealing was carried out. A striking result is a strong result dependence of the bonding behavior on the Cu recess shape, even for constant dishing depth, which places a serious constraint on the assumed surface profile for hybrid bonding simulation. In addition, a potentially detrimental impact of an increase in the interconnect aspect ratio or surface fraction was observed, depending

on the relationship between interconnect dimensions and surface topography after planarization. Such a trend, if confirmed, could have negative consequences on dishing compensation for high density of interconnections. This potential detrimental effect could be mitigated by careful selection of the metal/dielectric material pair for hybrid bonding, in addition naturally to better control on CMP-induced topography.

Several important topics may be considered for future work:

- Adhesive forces between Cu surfaces at the hybrid bonding interface were not accounted for in Chapter II, but were previously shown to play a major role on the bonding behavior [100]. The influence of the lateral extent of these interaction forces relative to pad width remains to be clarified. The latter effect, if any, is expected to strongly depend on the assumed shape for the Cu recess profile and more importantly on the relationship between recess depth and pad width, thus warranting further surface topography characterization or CMP process simulation to provide better understanding on these aspects.

- For the most part only 2D Cu recess profiles were considered in the present study, and therefore detailed modeling of the bonding process in the case of a 3D Cu recess surface for square-shaped hybrid bonding pads remains an open topic.

- The interplay between adhesion and plasticity also needs to be further investigated, accounting for the influence of pad aspect ratio on macroscopic plastic deformation, in addition to local plasticity at the bonding front.

References

[1] R. I. Made, C. L. Gan, L. Yan, K. H. B. Kor, H. L. Chia, K. L. Pey, and C. V. Thompson, "Experimental characterization and modeling of the mechanical properties of cu–cu thermocompression bonds for three-dimensional integrated circuits," Acta Materialia, vol. 60, no. 2, pp. 578–587, 2012.

[2] G. Galilei, Discorsi e Dimostrazioni Matematiche Intorno a Due Nuove Scienze. Leiden: Elzevir, 1638. (Giornata Prima, p. 11, The Macmillan Company, New York, 1914).

[3] I. Newton, Opticks. London: Smith and Walford, 1704. (Book Three, Part 1, Observations concerning the Inflexions of the Rays of Light, and the colors made thereby, pp. 388, 389 and 394, Dover, New York, 1952).

[4] R. J. Strutt, "A study of glass surfaces in optical contact," Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences, vol. 156, no. 888, pp. 326–349, 1936.

[5] U. Biermann, A. Van Gorkum, and J. Pals, "Direct bonding: From an optical technology to a broad research topic," Philips Journal of Research, vol. 49, no. 1-2, pp. 1–10, 1995.

[6] J. Lasky, "Wafer bonding for silicon-on-insulator technologies," Applied Physics Letters, vol. 48, no. 1, pp. 78–80, 1986.

[7] M. Shimbo, K. Furukawa, K. Fukuda, and K. Tanzawa, "Silicon-to-silicon direct bonding method," Journal of Applied Physics, vol. 60, no. 8, pp. 2987–2989, 1986.

[8] J. N. Israelachvili, Intermolecular and surface forces. Academic press, 2015.

[9] Z. Liau, "Semiconductor wafer bonding via liquid capillarity," Applied Physics Letters, vol. 77, no. 5, pp. 651–653, 2000.

[10] C. Jerez-Hanckes, D. Qiao, and S. Lau, "A study of si wafer bonding via methanol capillarity," Materials chemistry and physics, vol. 77, no. 3, pp. 751–754, 2003.

[11] J. Berthier, K. Brakke, F. Grossi, L. Sanchez, and L. Di Cioccio, "Self-alignment of silicon chips on wafers: A capillary approach," Journal of Applied Physics, vol. 108, no. 5, p. 054905, 2010.

[12] H. Rayssac, M. O, B. Aspar, and B. Ghyselen, "The bonding energy control: an original way to debondable substrates," in Semiconductor Wafer Bonding VII: Science, Technology, and Applications: Proceedings of the International Symposium, vol. 2003, p. 49, The Electrochemical Society, 2003.

[13] E. Navarro, Direct wafer bonding dynamics. PhD thesis, Université de Grenoble, 2014.

[14] D. Kubair and S. Spearing, "Cohesive zone model for direct silicon wafer bonding," Journal of Physics D: Applied Physics, vol. 40, no. 10, p. 3070, 2007.

[15] F. Rieutord, B. Bataillou, and H. Moriceau, "Dynamics of a bonding front," Physical review letters, vol. 94, no. 23, p. 236101, 2005.

[16] E. Navarro, Y. Bréchet, R. Moreau, T. Pardoen, J.-P. Raskin, A. Barthelemy, and I. Radu, "Direct silicon bonding dynamics: A coupled fluid/structure analysis," Applied Physics Letters, vol. 103, no. 3, p. 034104, 2013.

[17] D. Radisson, F. Fournel, and E. Charlaix, "Modelling of the direct bonding wave," Microsystem Technologies, vol. 21, no. 5, pp. 969–971, 2015.

[18] Q.-Y. Tong and U. Goesele, Semiconductor wafer bonding: science and technology. John Wiley, 1999.

[19] C. Ventosa, F. Rieutord, L. Libralesso, C. Morales, F. Fournel, and H. Moriceau, "Hydrophilic low-temperature direct wafer bonding," Journal of Applied Physics, vol. 104, no. 12, p. 123524, 2008.

[20] T. Plach, K. Hingerl, S. Tollabimazraehno, G. Hesser, V. Dragoi, and M. Wimplinger, "Mechanisms for room temperature direct wafer bonding," Journal of Applied Physics, vol. 113, no. 9, p. 094905, 2013.

[21] E. Navarro, Y. Bréchet, A. Barthelemy, I. Radu, J.-P. Raskin, and T. Pardoen, "Adhesion and separation models for direct hydrophilic bonding," Journal of Applied Physics, vol. 117, no. 8, p. 085305, 2015.

[22] U. Gösele, H. Stenzel, T. Martini, J. Steinkirchner, D. Conrad, and K. Scheerschmidt, "Self-propagating room-temperature silicon wafer bonding in ultrahigh vacuum," Applied physics letters, vol. 67, no. 24, pp. 3614–3616, 1995.

[23] F. Fournel, Applications de l'adhésion moléculaire. Habilitation à diriger des recherches (HDR), Université Grenoble I - Joseph Fourier, 2009. In French.

[24] K. T. Turner and S. M. Spearing, "Modeling of direct wafer bonding: Effect of wafer bow and etch patterns," Journal of Applied Physics, vol. 92, no. 12, pp. 7658–7666, 2002.

[25] K. T. Turner, M. Thouless, and S. Spearing, "Mechanics of wafer bonding: Effect of clamping," Journal of applied physics, vol. 95, no. 1, pp. 349–355, 2004.

[26] K. T. Turner, Wafer bonding: mechanics-based models and experiments. PhD thesis, Massachusetts Institute of Technology, 2004.

[27] D. V. Kubair, D. J. Cole, L. C. Ciacchi, and S. M. Spearing, "Multiscale mechanics modeling of direct silicon wafer bonding," Scripta Materialia, vol. 60, no. 12, pp. 1125–1128, 2009.

[28] R. Estevez, G. Parry, and P. McGarry, "Modelling of molecular bonding with a cohesive zone model strategy," in Proceedings of the Thirteenth International Conference on Fracture (ICF13), 2013.

[29] F. Fournel, L. Continni, C. Morales, J. Da Fonseca, H. Moriceau, F. Rieutord, A. Barthelemy, and I. Radu, "Measurement of bonding energy in an anhydrous nitrogen atmosphere and its application to silicon direct bonding technology," Journal of Applied Physics, vol. 111, no. 10, p. 104907, 2012.

[30] W. Maszara, G. Goetz, A. Caviglia, and J. McKitterick, "Bonding of silicon wafers for silicon-on-insulator," Journal of Applied Physics, vol. 64, no. 10, pp. 4943–4950, 1988.

[31] Y. Bertholet, F. Iker, J.-P. Raskin, and T. Pardoen, "Steady-state measurement of wafer bonding cracking resistance," Sensors and Actuators A: Physical, vol. 110, no. 1-3, pp. 157–163, 2004.

[32] A. Masolin, P.-O. Bouchard, R. Martini, and M. Bernacki, "Thermo-mechanical and fracture properties in single-crystal silicon," Journal of Materials Science, vol. 48, no. 3, pp. 979–988, 2013.

[33] R. Stengl, T. Tan, and U. Gösele, "A model for the silicon wafer bonding process," Japanese Journal of Applied Physics, vol. 28, no. 10R, p. 1735, 1989.

[34] D. Timpel, M. Schaible, and K. Scheerschmidt, "Molecular dynamics studies of silica wafer bonding," Journal of applied physics, vol. 85, no. 5, pp. 2627–2635, 1999.

[35] D. A. Litton and S. H. Garofalini, "Modeling of hydrophilic wafer bonding by molecular dynamics simulations," Journal of Applied Physics, vol. 89, no. 11, pp. 6013–6023, 2001.

[36] D. J. Cole, M. C. Payne, G. Csányi, S. Mark Spearing, and L. Colombi Ciacchi, "Development of a classical force field for the oxidized si surface: Application to hydrophilic wafer bonding," The Journal of chemical physics, vol. 127, no. 20, p. 204704, 2007.

[37] S. Leroch and M. Wendland, "Simulation of forces between humid amorphous silica surfaces: A comparison of empirical atomistic force fields," The Journal of Physical Chemistry C, vol. 116, no. 50, pp. 26247–26261, 2012.

[38] D. Grierson and K. T. Turner, "Characterization of hysteresis of surface energy in room-temperature direct bonding processes," ECS Transactions, vol. 33, no. 4, pp. 573–580, 2010.

[39] H. Yu and Z. Suo, "A model of wafer bonding by elastic accommodation," Journal of the Mechanics and Physics of Solids, vol. 46, no. 5, pp. 829–844, 1998.

[40] Q.-Y. Tong and U. Gösele, "Semiconductor wafer bonding: recent developments," Materials Chemistry and physics, vol. 37, no. 2, pp. 101–127, 1994.

[41] Q.-Y. Tong and U. Gösele, "Thickness considerations in direct silicon wafer bonding," Journal of the Electrochemical Society, vol. 142, no. 11, pp. 3975–3979, 1995.

[42] N. Miki and S. Spearing, "Effect of nanoscale surface roughness on the bonding energy of direct-bonded silicon wafers," Journal of Applied Physics, vol. 94, no. 10, pp. 6800–6806, 2003.

[43] K. Turner, S. Spearing, W. Baylies, M. Robinson, and R. Smythe, "Effect of nanotopography in direct wafer bonding: modeling and measurements," IEEE transactions on semiconductor manufacturing, vol. 18, no. 2, pp. 289–296, 2005.

[44] J. Greenwood and J. P. Williamson, "Contact of nominally flat surfaces," Proceedings of the royal society of London. Series A. Mathematical and physical sciences, vol. 295, no. 1442, pp. 300–319, 1966.

[45] C. Gui, M. Elwenspoek, N. Tas, and J. G. Gardeniers, "The effect of surface roughness on direct wafer bonding," Journal of applied physics, vol. 85, no. 10, pp. 7448–7454, 1999.

[46] B. V. Derjaguin, V. M. Muller, and Y. P. Toporov, "Effect of contact deformations on the adhesion of particles," Journal of Colloid and interface science, vol. 53, no. 2, pp. 314–326, 1975.

[47] F. Rieutord, H. Moriceau, R. Beneyton, L. Capello, C. Morales, and A.-M. Charvet, "Rough surface adhesion mechanisms for wafer bonding," ECS Transactions, vol. 3, no. 6, pp. 205–215, 2006.

[48] G. Liao, T. Shi, X. Lin, and Z. Ma, "Effect of surface characteristic on room-temperature silicon direct bonding," Sensors and Actuators A: Physical, vol. 158, no. 2, pp. 335–341, 2010.

[49] K. L. Johnson, K. Kendall, and A. Roberts, "Surface energy and the contact of elastic solids," Proceedings of the royal society of London. A. mathematical and physical sciences, vol. 324, no. 1558, pp. 301–313, 1971.

[50] B. Anglicus, De genuinis rerum coelestium terrestrium at inferarum proprietatibus rerum. 1240. (Book 16, Chap. 4, p. 718, Wolfgang Richter, Frankfurt, 1601) cf. A. Plößl and G. Kräuter, Mater. Sci. Eng. R25:1 (1999).

[51] J. T. Desaguliers, "Some experiments concerning the cohesion of lead," Philosophical Transactions (1683-1775), vol. 33, pp. 325–347, 1725.

[52] B. Derby and E. Wallach, "Joining methods in space: A theoretical model for diffusion bonding," Acta Astronautica, vol. 7, no. 4-5, pp. 685–698, 1980.

[53] P. Gondcharton, Intégration du collage direct: Couches minces métalliques et évolutions morphologiques. PhD thesis, Grenoble Alpes, 2015. In French.

[54] A. Coucoulas, "Hot work ultrasonic bonding-a method of facilitating metal flow by restoration processes," in Proc. 20 IEEE Electronic Components Conf., pp. 549–556, 1970.

[55] Y. A. Li, R. W. Bower, and I. Bencuya, "Low temperature copper to copper direct bonding," Japanese journal of applied physics, vol. 37, no. 9A, p. L1068, 1998.

[56] C.-T. Ko and K.-N. Chen, "Low temperature bonding technology for 3d integration," Microelectronics reliability, vol. 52, no. 2, pp. 302–311, 2012.

[57] F. Baudin, Contribution à l'étude des phénomènes mis en jeu par le collage direct à basse température de couches métalliques et oxydes métalliques. PhD thesis, Grenoble, 2013.

[58] P. Gueguen, L. Di Cioccio, P. Gergaud, M. Rivoire, D. Scevola, M. Zussy, A. M. Charvet, L. Bally, D. Lafond, and L. Clavelier, "Copper direct-bonding characterization and its interests for 3D integration," Journal of The Electrochemical Society, vol. 156, no. 10, pp. H772–H776, 2009.

[59] K. N. Chen, A. Fan, and R. Reif, "Interfacial morphologies and possible mechanisms of copper wafer bonding," Journal of Materials Science, vol. 37, no. 16, pp. 3441–3446, 2002.

[60] I. Radu, D. Landru, G. Gaudin, G. Riou, C. Tempesta, F. Letertre, L. Di Cioccio, P. Gueguen, T. Signamarcheix, C. Euvrard, et al., "Recent developments of cu-cu non-thermo compression bonding for wafer-to-wafer 3d stacking," in 2010 IEEE International 3D Systems Integration Conference (3DIC), pp. 1–6, IEEE, 2010.

[61] L. Di Cioccio, P. Gueguen, R. Taibi, D. Landru, G. Gaudin, C. Chappaz, F. Rieutord, F. de Crecy, I. Radu, L. Chapelon, et al., "An overview of patterned metal/dielectric surface bonding: mechanism, alignment and characterization," Journal of The Electrochemical Society, vol. 158, no. 6, pp. P81–P86, 2011.

[62] M. Martinez, M. Legros, T. Signamarcheix, L. Bally, S. Verrun, L. Di Cioccio, and C. Deguet, "Mechanisms of copper direct bonding observed by in-situ and quantitative transmission electron microscopy," Thin Solid Films, vol. 530, pp. 96–99, 2013.

[63] L. Di Cioccio, F. Baudin, P. Gergaud, V. Delaye, P.-H. Jouneau, F. Rieutord, and T. Signamarcheix, "Modeling and integration phenomena of metal-metal direct bonding technology," ECS Transactions, vol. 64, no. 5, pp. 339–355, 2014.

[64] F. Fournel, C. Martin-Cocher, D. Radisson, V. Larrey, E. Beche, C. Morales, P. Delean, F. Rieutord, and H. Moriceau, "Water stress corrosion in bonded structures," ECS Journal of Solid State Science and Technology, vol. 4, no. 5, pp. P124–P130, 2015.

[65] D. Radisson, Direct bonding of patterned surfaces. PhD thesis, Université Grenoble-Alpes, 2014.

[66] P. Gondcharton, B. Imbert, L. Benaissa, V. Carron, and M. Verdier, "Kinetics of low temperature direct copper–copper bonding," Microsystem Technologies, vol. 21, no. 5, pp. 995–1001, 2015.

[67] N. Cabrera and N. F. Mott, "Theory of the oxidation of metals," Reports on progress in physics, vol. 12, no. 1, p. 163, 1949.

[68] K. Chen, A. Fan, C. Tan, R. Reif, and C. Wen, "Microstructure evolution and abnormal grain growth during copper wafer bonding," Applied Physics Letters, vol. 81, no. 20, pp. 3774–3776, 2002.

[69] K. Chen, C. Tan, A. Fan, and R. Reif, "Morphology and bond strength of copper wafer bonding," Electrochemical and Solid-State Letters, vol. 7, no. 1, pp. G14–G16, 2004.

[70] P. Gondcharton, B. Imbert, L. Benaissa, and M. Verdier, "Voiding phenomena in copper-copper bonded structures: role of creep," ECS Journal of Solid State Science and Technology, vol. 4, no. 3, pp. P77–P82, 2015.

[71] P. Gondcharton, B. Imbert, L. Benaissa, F. Fournel, and M. Verdier, "Effect of copper–copper direct bonding on voiding in metal thin films," Journal of Electronic Materials, vol. 44, no. 11, pp. 4128–4133, 2015.

[72] H. Mohamed and J. Washburn, "Mechanism of solid state pressure welding," Welding Journal, p. vp, 1975.

[73] H. Wieder and A. Czanderna, "The oxidation of copper films to cuo0. 67," The Journal of Physical Chemistry, vol. 66, no. 5, pp. 816–821, 1962.

[74] M. Lenglet, K. Kartouni, J. Machefert, J. Claude, P. Steinmetz, E. Beauprez, J. Heinrich, and N. Celati, "Low temperature oxidation of copper: The formation of cuo," Materials research bulletin, vol. 30, no. 4, pp. 393–403, 1995.

[75] D. Cocke, G. Chuah, N. Kruse, and J. Block, "Copper oxidation and surface copper oxide stability investigated by pulsed field desorption mass spectrometry," Applied surface science, vol. 84, no. 2, pp. 153–161, 1995.

[76] Z. Munir, "Analytical treatment of the role of surface oxide layers in the sintering of metals," Journal of Materials Science, vol. 14, no. 11, pp. 2733–2740, 1979.

[77] Y. Takahashi, T. Nakamura, and K. Nishiguchi, "Dissolution process of surface oxide film during diffusion bonding of metals," Journal of materials science, vol. 27, no. 2, pp. 485–498, 1992.

[78] P. Gondcharton, B. Imbert, L. Benaissa, and M. Verdier, "Copper-copper direct bonding: Impact of grain size," in 2015 IEEE International Interconnect Technology Conference and

2015 IEEE Materials for Advanced Metallization Conference (IITC/MAM), pp. 229–232, IEEE, 2015.

[79] O. Rayssac, Etude du collage par adhésion moléculaire hydrophile: application au contrôle de l'énergie de collage. PhD thesis, Grenoble INPG, 1999. In French.

[80] Y. Bertholet, Measurement, optimization and multiscale modeling of silicon wafer bonding interface fracture resistance. PhD thesis, Catholic University of Louvain, 2006.

[81] Y. Bertholet, B. Olbrechts, B. Lejeune, J.-P. Raskin, and T. Pardoen, "Molecular bonding aided by dissipative inter-layers," Acta materialia, vol. 55, no. 2, pp. 473–479, 2007.

[82] V. Balan, A. Seignard, D. Scevola, J.-F. Lugand, L. Di Cioccio, and M. Rivoire, "Cmp process optimization for bonding applications," in ICPT 2012-International Conference on Planarization/CMP Technology, pp. 1–7, VDE, 2012.

[83] S. Lhostis, A. Farcy, E. Deloffre, F. Lorut, S. Mermoz, Y. Henrion, L. Berthier, F. Bailly, D. Scevola, F. Guyader, et al., "Reliable 300 mm wafer level hybrid bonding for 3D stacked CMOS image sensors," in Electronic Components and Technology Conference (ECTC), 2016 IEEE 66th, pp. 869–876, IEEE, 2016.

[84] J. Jourdon, S. Lhostis, S. Moreau, J. Chossat, M. Arnoux, C. Sart, Y. Henrion, P. Lamontagne, L. Arnaud, N. Bresson, and others, "Hybrid bonding for 3d stacked image sensors: impact of pitch shrinkage on interconnect robustness," in 2018 IEEE International Electron Devices Meeting (IEDM), pp. 7–3, IEEE, 2018.

[85] B. Rebhan, M. Bernauer, T. Wagenleitner, M. Heilig, F. Kurz, S. Lhostis, E. Deloffre, A. Jouve, V. Balan, and L. Chitu, "$<$ 200 nm wafer-to-wafer overlay accuracy in wafer level cu/sio2 hybrid bonding for bsi cis," in Electronics Packaging and Technology Conference (EPTC), 2015 IEEE 17th, pp. 1–4, IEEE, 2015.

[86] R. Vacassy and Z. Chen, "Edge-over-erosion in tungsten cmp," in Proc. 11th Int. Chem. Mech. Planarization ULSI Multilevel Interconnect. Conf.(CMP-MIC), pp. 632–639, 2006.

[87] G. Banerjee and R. L. Rhoades, "Chemical mechanical planarization historical review and future direction," ECS Transactions, vol. 13, no. 4, pp. 1–19, 2008.

[88] P. Wrschka, J. Hernandez, G. S. Oehrlein, and J. King, "Chemical mechanical planarization of copper damascene structures.," Journal of the Electrochemical Society, vol. 147, no. 2, pp. 706–712, 2000.

[89] G. Fu and A. Chandra, "An analytical dishing and step height reduction model for chemical mechanical planarization (cmp)," IEEE transactions on semiconductor manufacturing, vol. 16, no. 3, pp. 477–485, 2003.

[90] J. J. Vlassak, "A model for chemical-mechanical polishing of a material surface based on contact mechanics.," Journal of the Mechanics and Physics of Solids, vol. 52, pp. 847–873, 2004.

[91] J.-Y. Lai, N. Saka, and J.-H. Chun, "Evolution of copper-oxide damascene structures in chemical mechanical polishing ii. copper dishing and oxide erosion," Journal of The Electrochemical Society, vol. 149, no. 1, pp. G41–G50, 2002.

[92] Y. Beilliard, Etude de l'intégration du collage direct cuivre/oxyde pour l'élaboration d'une architecture 3D-SIC. PhD thesis, Université Grenoble-Alpes, 2015.

[93] N. Elbel, B. Neureither, B. Ebersberger, and P. Lahnor, "Tungsten chemical mechanical polishing," Journal of The Electrochemical Society, vol. 145, no. 5, pp. 1659–1664, 1998.

[94] V. H. Nguyen, R. Daamen, H. van Kranenburg, P. van der Velden, and P. H. Woerlee, "A physical model for dishing during metal cmp," Journal of The Electrochemical Society, vol. 150, no. 11, pp. G689–G693, 2003.

[95] R. Saxena, D. G. Thakurta, R. J. Gutmann, and W. N. Gill, "A feature scale model for chemical mechanical planarization of damascene structures," Thin Solid Films, vol. 449, no. 1-2, pp. 192–206, 2004.

[96] C. Feng, C. Yan, J. Tao, X. Zeng, and W. Cai, "A contact-mechanics-based model for general rough pads in chemical mechanical polishing processes," Journal of The Electrochemical Society, vol. 156, no. 7, pp. H601–H611, 2009.

[97] Q. Xu and L. Chen, "A feature-scale greenwood–williamson model for metal chemical mechanical planarization," Journal of electronic materials, vol. 42, no. 8, pp. 2630–2640, 2013.

[98] M. Dao, N. V. Chollacoop, K. J. Van Vliet, T. A. Venkatesh, and S. Suresh, "Computational modeling of the forward and reverse problems in instrumented sharp indentation," Acta materialia, vol. 49, no. 19, pp. 3899–3918, 2001.

[99] Y. Beilliard, P. Coudrain, L. Di Cioccio, S. Moreau, L. Sanchez, B. Montmayeul, T. Signamarcheix, R. Estevez, and G. Parry, "Chip to wafer copper direct bonding electrical characterization and thermal cycling," in 3D Systems Integration Conference (3DIC), 2013 IEEE International, pp. 1–7, IEEE, 2013.

[100] Y. Beilliard, R. Estevez, G. Parry, P. McGarry, L. Di Cioccio, and P. Coudrain, "Thermo-mechanical finite element modeling of Cu–SiO_2 direct hybrid bonding with a dishing effect on Cu surfaces," International Journal of Solids and Structures, vol. 117, pp. 208–220, 2017.

[101] Y.-L. Shen, S. Suresh, and I. A. Blech, "Stresses, curvatures, and shape changes arising from patterned lines on silicon wafers," Journal of Applied Physics, vol. 80, no. 3, pp. 1388–1398, 1996.

[102] E. Ege and Y.-L. Shen, "Thermomechanical response and stress analysis of copper interconnects," Journal of Electronic Materials, vol. 32, no. 10, pp. 1000–1011, 2003.

[103] T.-S. Park, M. Dao, S. Suresh, A. Rosakis, D. Pantuso, and S. Shankar, "Some practical issues of curvature and thermal stress in realistic multilevel metal interconnect structures," Journal of Electronic Materials, vol. 37, no. 6, pp. 777–791, 2008.

[104] P. Enquist, "Metal/silicon oxide hybrid bonding," in Handbook of Wafer Bonding (P. Ramm, J. J. Q. Lu, and M. M. V. Taklo, eds.), ch. 13, pp. 261–278, Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, Germany, 2012.

[105] T. Kondo, Y. Takemoto, K. Kobayashi, M. Tsukimura, N. Takazawa, H. Kato, S. Suzuki, J. Aoki, H. Saito, Y. Gomi, and others, "A 3d stacked CMOS image sensor with 16mpixel global-shutter mode and 2mpixel 10000fps mode using 4 million interconnections," in VLSI Circuits (VLSI Circuits), 2015 Symposium on, pp. C90–C91, IEEE, 2015.

[106] E. Beyne, S.-W. Kim, L. Peng, N. Heylen, J. De Messemaeker, O. O. Okudur, A. Phommahaxay, T.-G. Kim, M. Stucchi, D. Velenis, et al., "Scalable, sub 2μm pitch, cu/sicn to cu/sicn hybrid wafer-to-wafer bonding technology," in Electron Devices Meeting (IEDM), 2017 IEEE International, pp. 32–4, IEEE, 2017.

[107] C.-H. Hsueh, "Modeling of thermal stresses in passivated interconnects," Journal of applied physics, vol. 92, no. 1, pp. 144–153, 2002.

[108] T. Wlanis, R. Hammer, W. Ecker, S. Lhostis, C. Sart, S. Gallois-Garreignot, B. Rebhan, and G. A. Maier, "Cu-sio2 hybrid bonding simulation including surface roughness and viscoplastic material modeling: A critical comparison of 2d and 3d modeling approach," Microelectronics Reliability, vol. 86, pp. 1–9, 2018.

Chapter III

Investigation of mechanical failures in interconnect layers induced by wire bonding on backside pads

III.1 Introduction

Since the 1990s, the semiconductor industry has been facing a series of challenges in the race for miniaturization, such as the move from Al to Cu interconnects, the introduction of low-κ dielectrics or the search for replacement wire materials due to the surge in the price of gold. These changes have led to increased attention to the mechanical reliability of wirebond pads, especially the integrity of the underlying interconnect structure. The wire bonding process has thus been studied extensively over the past three decades, enabling to gain insight into the physics of thermosonic bonding and the effect of process parameters on bond strength [1].

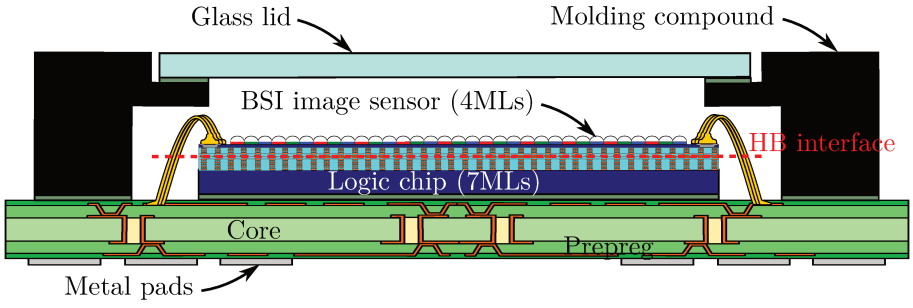


FIGURE III.1: Schematical description of the 3D image sensor studied in this work.
a

In this work, we are interested in the mechanical robustness of wirebond pads in a 3D stacked BSI image sensor (Figure III.1). This advanced architecture gives rise to new challenges. As described in Chapter I, wire bonds are formed on the backside of the flipped BSI image sensor chip (Figure III.2). A direct consequence is the presence of the lowermost metallization levels

(e.g. contact layer, X levels) near the bonding surface, where the loading exerted due to bond compression and ultrasonic vibration is the most critical. These layers contain local interconnect levels, consisting of thinner and narrower metal lines compared to the rest of the interconnect stack. An increased risk of cracking or delamination can therefore be expected in this new configuration.

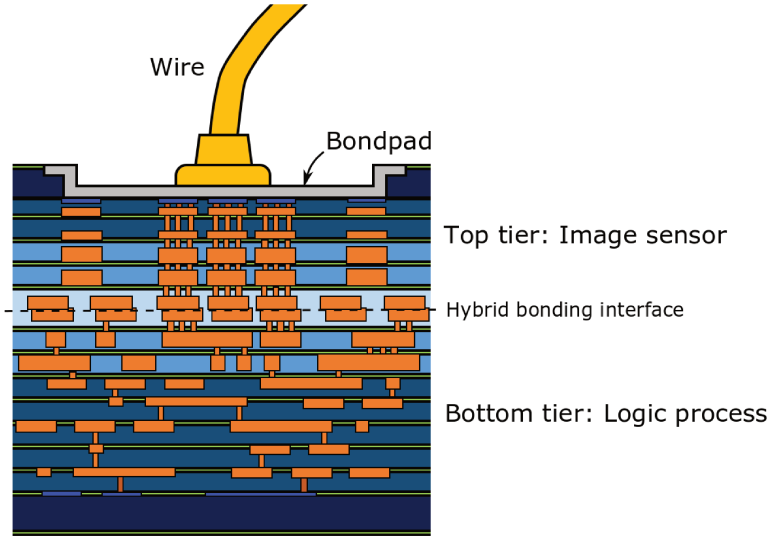


FIGURE III.2: Schematic illustration of bondpad back-end-of-line configuration in the studied IC architecture with chips stacked by hybrid bonding

Extensive experimental testing is required to optimize the mechanical robustness of the wire-bond. This is achieved through fine-tuning of the bonding parameters (e.g. compressive force, ultrasound power, bonding time, temperature) using design-of-experiment methodologies, and robust design of the interconnect structure below the pad. However, very short development cycles are mandatory to meet the industry roadmap's objectives for device miniaturization. Therefore, numerical simulation plays an important role, enabling to gain better understanding of the physics of the bonding process, and providing a cost-effective way to identify the most influent design parameters (layout, materials, dimensions) to strengthen the pad structure using a reduced number of samples.

In this chapter, we start by an overview of simulation studies in the literature regarding the mechanical robustness of the interconnect stack below the bondpad. Then, an experimental comparison between several pad architectures is carried out, based on the number of delamination occurrences observed after wire-bonding. The focus is on the influence of interconnection layout below the pad, as well as the introduction of a capping layer between the bonding surface and the interconnection stack. Aiming to provide a better understanding of the experimental trends, the results are combined with a finite element analysis of a qualification test used to assess bond strength, namely the wire-pull test. This numerical study is then extended to account for additional interconnect layouts and the impact of chip stacking on mechanical stress in the interconnect stack is investigated. Finally, the results and methodology used to compare the different structures are discussed and next steps for future work are proposed.

III.2 Literature review

In this section, simulation studies in the literature regarding the mechanical robustness of the interconnect stack below the bondpad are reviewed. An overview is provided in Table III.7.

III.2.1 Methodology

In general, the issue of bondpad mechanical reliability may be approached from two complementary perspectives:

Bonding-oriented approach the bonding parameters can be optimized, aiming to find a tradeoff between low-intensity bonding conditions inducing minimum deformation below the bond but resulting in poor bond strength [2], or severe bonding conditions potentially leading to high levels of stress in the interconnect structure below the bondpad;

BEoL-oriented approach after finding an optimal set of parameters for the bonding process, if bondpad strength is unacceptably low during qualification testing (e.g., crack propagation or interface delamination in the BEoL), a more robust interconnect layout can be designed below the bonding surface. In addition to the bonding process itself, this optimal architecture must able to withstand subsequent loading, e.g. during qualification, further processing or product life.

Both of these approaches have been tackled in the literature on wire bonding simulation.

III.2.1.1 Bonding-oriented studies

In a majority of studies, the focus is on providing a better understanding of the wire bonding process, aiming to predict the final shape of the bond after ball impact and ultrasonic vibration as well as the resulting state of stress. In addition to the influence of bonding parameters, different materials for the wire or the bondpad are also generally investigated, e.g. the introduction of (ultra-)low-κ dielectric in the interconnect stack or the move to Cu wires. Therefore, models with complex physics and simplified geometry are typically used. Notable examples of such studies are detailed below.

In a series of articles, [3–5] investigated ball deformation by the capillary tool during the impact stage, and the impact of new interconnection materials on bondpad stresses for a 3ML stack using a 2D axisymmetric finite element model. A two-step process was used to model bond formation. First, bond compression on a rigid surface was simulated. A yield stress value was calibrated for the wire by fitting the simulated bond deformation to SEM characterizations of bond shape. Then, the force distribution obtained on the rigid surface was used as a boundary condition on a bondpad model. They showed that a much higher bonding force is required in the case of a Cu wire to obtain the same bond diameter as using gold. Thus, for a given target bond shape,

higher stress is exerted on the underlying pad structure for the case of Cu wires. A comparison of the simulated stress distribution below the pad with Raman spectroscopy measurements was provided in [6], showing very good agreement. However, [7] compared the force distribution on the bonding surface for a realistic interconnect stack and for a rigid substrate in the case of a wire-pull load. They showed that significant mechanical interaction occurs between the ball and the pad structure, and argued that considering a rigid or fully homogenized substrate would not yield sufficiently accurate results to compare the stresses induced in different pad architectures, confirming some of their earlier results on pad structure homogenization methods [8].

Dynamic analysis of Au wire bonding on a 7ML Cu/low-κ pad structure was carried out by [9] and [10, 11]. In their work, both ball impact and ultrasonic vibration were accounted for using a 2D plane strain model, and the influence of the elastic moduli of interconnection materials on bondpad stress was investigated. Higher stress was obtained in the bondpad during the ultrasonic vibration phase of the bonding process, compared to the ball impact phase. [12, 13] also proposed a transient nonlinear dynamic model of wire bonding for a 3ML Al/SiO_2 BOA bond pad, where the cooling step after thermosonic bonding was also considered. Rate-dependent elastic-plastic behavior was assumed for the Au wire and friction between the Au ball and the Al bondpad was accounted for. In addition to guidelines to optimize bond formation, they investigated the influence of pad structure on the stress transmitted to the active region below the bond. They noticed that the stress becomes more uniformly distributed during the cooling process, with a significant increase in magnitude compared to the bond impact and ultrasonic vibration steps. This finding is in agreement with the results of [8], who simulated the wire pull/shear tests and thermal cooling, and also concluded that thermal loading appears to be a major stress contributor. [2] accounted for an increase of the dynamic friction coefficient between the ball and the bondpad due to oxide removal and Au-Al interdiffusion during the ultrasonic vibration phase of the bonding process. In addition, [14] also included a decrease in the wire material yield stress due to ultrasonic energy (ultrasonic softening phenomenon). They argued that the majority of pad failures occur during the ultrasonic stage, consisting of 500-1200 cycles of side-to-side motion exerted on the ball by the capillary tool. They were able to simulate up to 60 cycles for the ultrasonic vibration, after which bond height was found to saturate. In comparison, only one to nine cycles could be accounted for in previous studies [2, 13, 15], due to computational cost limitations. The influence of intermetallic compounds forming at the interface between the wire and the bondpad on the debonding behavior was recently investigated by [16].

Although these bonding-oriented approaches enable to gain more insight into the physics of the wire bonding process, convergence and computational cost issues can become problematic for large parametric studies, in particular for comparing different bondpad architectures. In the majority of cases, two-dimensional models were used and stacks with a low number of interconnection levels or very simplified layouts considered.

III.2.1.2 BEoL-oriented studies

Some authors on the other hand focused on the mechanical robustness of the bondpad interconnect stack, using simplified loading and physics but detailed geometric features. Notable examples include [17, 18], [19, 20] or [21]. They developed 3D models of realistic pad structures based on actual BEoL design rules. Ball deformation was not simulated and instead an a posteriori bond shape was assumed, based on bond shape measurements in SEM views of actual wirebonds. As a result, residual stresses due to the bonding process were not included in the computations, although they are believed to be significant [12, 22]. Instead, static loading was used to simulate the very end of bonding process or subsequent wire pull qualification testing. In addition, the proposed models were based on a multi-scale modeling framework, using equivalent homogeneous orthotropic materials for each layer of the interconnect stack at the global scale, then prescribing the computed displacements as boundary conditions on a detailed model of a representative unit cell at the local scale. As detailed in the following section, energy-based criteria were developed in some of these studies to compare different interconnect layouts. Good agreement was obtained with wire pull or ball shear testing on actual structures, both in terms of the predicted weakest interface in the stack and the most robust pad configuration [17, 21].

III.2.2 Robustness index

III.2.2.1 Stress-based criteria

As can be seen from the literature overview in Table III.7, a majority of simulation studies have been relying on stress-based approaches to investigate the effect of bondpad configuration on mechanical robustness.

[8] developed a 3D finite element model of a bondpad architecture and compared three distinct interconnect layouts under wire pull, wire shear and thermal loading. They used the cumulative volume distribution of the maximum principal stress in the low-κ material to rank the different pad configurations. Overall, good agreement was obtained with qualification testing results on these structures, although distinct trends are obtained depending on whether the average stress or the peak stresses are considered. [23] compared the reliability of six pad options, focusing on the ultrasonic stage. Contrary to most studies, relying on the maximum principal stress or maximum Von Mises stress to compare different pad structures, they proposed a stress criterion based on fatigue theory using the maximum in-plane normal and out-of-plane shear stress range. In some pads, a buffer layer was introduced below the bonding surface, which they showed contributes to considerably attenuate failure risk. Their results were validated by cratering tests[1]. The modified pad interconnect structures showed no significant decrease in terms of mechanical robustness compared to the reference configuration. [22] simulated ball deformation during thermosonic bonding and incorporated an additional step of wire pull just after bond

[1] See Section III.6.1.

formation. A strain-based damage criterion was used with an element erosion procedure to investigate the different failure modes of the pull test. They were able to capture failure mode sensitivity to the elastic properties and to the strength of low-κ intermetal dielectric. [2] used the incremental averaged first principal stress to compare two distinct pad configurations. In their work, incremental stress refers to the stress accumulation during the ultrasonic phase only, subtracting the stress accumulated after the bond compression stage. Volume-averaged stress components were used instead of maximum stresses to avoid mesh dependency. The averaging was based on four elements around the location of interest. Comparison with bond qualification testing indicated that the proposed stress criterion captures well the influence of bond force, but underestimates the effect of ultrasound. [14] proposed a 3D finite element model of the wire bonding process, using a damage criterion developed by [24] to investigate the influence of removing the upper metal levels on pad failure risk.

III.2.2.2 Energy-based criteria

As seen in the previous section, stress-based approaches can lead to reasonable correlation with experimental results. However, in some cases comparing different pad configurations is not so straightforward, either because the influence of interconnect layout is not captured by the stress criterion [25] or because the results are strongly dependent upon the choice of the failure criterion, e.g. averaged or maximum stress [8]. In addition, stress-based quantities are configuration-dependent and therefore cannot be directly associated with intrinsic material quantities (e.g., fracture toughness, adhesion energy). [17] and [7] recommended the use of an energy-based criterion over the monitoring of stress quantities, in particular for its ability to describe delamination failure more accurately.

[17, 18] and [19, 20] developed a 3D finite element model for wire pull on a 6ML Cu/low-κ bond-pad, using a multi-scale approach with a material homogenization procedure. This strategy enabled to simulate local details of the interconnect stack and analyze different pad structures. A novel failure index, namely the area release energy (ARE), was specially developed to compare these structures. This method consists in a two-step process: first, computing the stress distribution on the healthy structure, then selecting a location at a given interface and releasing nodes over a circular area of chosen diameter. Using the computed forces and displacement at these nodes, an energy release can be computed over the selected surface. By repeating this procedure over the whole surface and for each interface, contour mappings for an energy quantity can be computed, even for surfaces with material heterogeneities, thereby providing a simple scalar quantity to compare different pad configurations. A key feature of this method is its ability to determine the most critical interface in a stack, provided that either the considered interfaces all have the same adhesion energy, or that their adhesion energies are distinct but known. [25] observed that the most critical interface in a stack is not necessarily that with the largest interfacial stress. These authors noted however that the ARE, associated with a total energy variation, has no obvious physical meaning. An enhanced ARE method was therefore

developed and validated analytically in [26] to adapt this technique for direct computation of the strain energy release rate.

[7, 21] and [27] also developed an energy-based failure index suitable for mechanical robustness comparisons between different pad architectures, namely the nodal release energy (NRE). Contrary to ARE, in this approach the energy release for a given interface is not computed over the whole surface but directly at the location of maximum stress, enabling to decrease computational cost while still providing a relevant failure index. Indeed, NRE method provided very good agreement with experimental results, in particular the most critical interface according to NRE agreed with experimentally characterized delamination locations.

Contrary to stress-based approaches, from a practical standpoint these methods are more relevant for pad failure by interface delamination rather than crack initiation, since the initial failure location needs to be assumed beforehand. This feature is not a limiting factor however for analyzing commonly encountered "cratering" or "peeling" pad failures, for which delamination in the interconnect stack is observed, especially in bondpads containing low-κ dielectrics with poor adhesion on barrier layers.

A reference length scale must be selected for the energy release calculation, which may influence the results depending on whether it is comparable with typical lengths associated with the metallization pattern along the considered interface. This size-dependence was evidenced by [17], who proposed to use the dimension of the smallest geometric entity in the model (e.g. via plugs) as the reference length scale, or to link it with a characteristic length obtained from experiments and intrinsic to the considered interface.

In addition to failure initiation location, strong assumptions are also required on crack front shape in the interface plane. Recently, several authors used cohesive zone models in simulation studies on wirebond pad mechanical reliability [28–31]. Such models are able to model both crack initiation and propagation, and therefore do not require any assumptions on the shape of the delaminated region. However, a characteristic length is still required to define the traction-separation law associated with the considered interface.

Both stress-based and energy-based methods have been used in the literature to assess the influence of interconnect layout on bondpad mechanical robustness. In the following section, the main results from these studies are reviewed.

III.2.3 Influence of interconnection layout

III.2.3.1 Metal lines

[32] evidenced a new bondpad failure mode, namely cratering damage, related to the presence of low-κ dielectric in the interconnect stack. They showed these failures could be eliminated by modifying the interconnect layout below the bondpad, replacing the Al metal sheets by cross-hatched metal lines. This configuration enables the containment of brittle dielectric inside "reservoirs", while providing mechanical reinforcement to the pad structure. A simplified finite element model of wire bonding confirmed a decrease of maximum tensile and shear stresses in the interconnect stack with this architecture.

The effect of interconnect layout on bondpad mechanical robustness was also investigated by [8]. They found that narrow parallel metal lines present a high percentage of cratering failures during wire pull and wire shear tests, whereas metal sheets with holes or wide parallel lines were within the specifications after qualification testing. These findings were correlated with finite element modeling, which showed higher maximum stress values for the rejected pad architecture.

[17, 18] also simulated the wire pull test and evidenced a beneficial effect of holes in the metal layers on pad robustness, using the ARE method. However, although the results did indicate a decrease of the energy release near the holes, larger values were obtained in other locations. Using the same approach, [25] compared several pad architectures (unfortunately not disclosed) and also observed that the metallization density below the pad is indeed "a factor that affects the mechanical performance". They further argued that "metal density in a single layer does not only locally affect the mechanical performance, but has an impact on the performance of the remaining layers as well".

An energy-based criterion, namely the NRE method, was also used by [7, 21] to compare the robustness of different pad interconnect configurations under bonding and wire pull loading. They found that patterns having wide metal lines (at constant metal surface fraction) lead to lower energy release values compared to narrow metal lines. Comparing two interconnect layouts, namely parallel metal lines and metal sheet with holes, they obtained better robustness with the latter configuration. Good correlation with wire pull testing on these structures was obtained.

III.2.3.2 Via plugs

[33] simulated the stresses in two Cu/SiO_2 bond-over-active (BOA) pad stack versions during wire bonding and thermal cycling, one with minimal metallization density at the edge of the bonded ball (no via above M3), and the other with maximum interconnect density. In their model, a static normal load was used to simulate bond compression and ultrasonic vibration. Results suggest that stress concentration is mainly confined to the uppermost levels of the pad structure near the edge of the bond, therefore with limited impact on the active regions below, and that vias only have a local effect on the stress distribution in the pad. They attributed this dissipation of the bonding forces to the comparable moduli of the copper metallization and SiO_2 dielectric. No failure events were obtained in the experimental study due to excellent mechanical robustness of both pads, thus preventing comparison with numerical results.

Contrary to the previous study, [5] evidenced a significant effect of via presence on pad robustness. Using a 2D axisymmetric model to simulate ball compression by the capillary tool during the impact stage, they found increasing via percentage leads to a large decrease of maximal Von Mises stress in the interlayer dielectric, due to via plugs providing a mechanical supporting structure. They also noticed that this beneficial effect is significantly diminished when the vias start to yield. [12] also simulated ball impact and bond formation, accounting for ultrasonic energy and Au ball viscoplastic behavior. However, they found that higher via density reduces the stress transferred to the active region, but induces larger maximum Von Mises stress in the dielectric layer surrounding the vias.

Using a dynamic finite element model of ball impact and ultrasonic vibration, [34] and [15, 35, 36] investigated the influence of many parameters related to via plugs on pad robustness. They reported the following trends:

- Increasing in the number of via plugs in a given layer results in a decrease of the equivalent tensile stress in the surrounding dielectric.
- As the width of the via plugs is increased, the volume of dielectric material is decreased, which dramatically reduces the equivalent tensile stress in the USG structure.
- Arranging vias in a nested square configuration was reported to provide better stress relief.
- The equivalent tensile stress in the USG dielectric increased with via plug thickness.
- The obtained peak equivalent tensile stress in the via plugs decreased when locating the via array away from the center of the bond, although a higher magnitude was obtained for the stress in the surrounding dielectric material.

In the next section, the strategy for the combined numerical/experimental study on the mechanical robustness of backside wirebond pads for a 3D image sensor stacked by hybrid bonding is presented.

III.3 Strategy

Realistic simulation of ball deformation during the impact stage is a challenging task. Due to the dynamic nature of the wirebonding process, a time-consuming explicit integration scheme must be used and very fine mesh is required at the tool/ball and ball/pad contact regions. Furthermore, frequent remeshing is mandatory to account for large deformation of the Au ball during bond formation. The ultrasonic vibration stage brings further difficulties, considering that for typical bonding parameters hundreds to thousands of cycles must be simulated.

Even beyond computational requirements, some important physical aspects are not easily characterized, let alone incorporated into the model. Examples include the dynamic friction coefficient between the ball and the bonding surface [13, 37], friction-induced heating [38, 39], dynamic effects of the capillary tool [40–42] or force-controlled loading [43]. Thus, detailed modeling of the thermosonic bonding process is beyond the scope of this study and will not be carried out here. Instead, simulation of a qualification test (e.g., wire pull test or wire shear test) used for bondpad mechanical robustness assessment offers a more straightforward alternative, while having been shown adequate for comparing different pad architectures [7, 21, 44].

A universal qualification test in the microelectronics industry is the wire-pull test (ASTM F 459-06). This method consists in pulling off the wire with a hook, generally at mid-span of the wire loop, while the reaction force on the tool is monitored. After wire pull, the failure mode is determined by visual inspection of the bonding surface. For the structure to pass the qualification, the maximum force must be above a certain specification, and fracture located in the bond rather than inside the interconnect stack.

In this work, a 3D finite element model of the wire-pull bond qualification test on backside pads in a 3D image sensor stacked by hybrid bonding is proposed. First, the mechanical robustness of six distinct pad configurations is compared experimentally and numerically. Then, a series of parametric simulation studies are carried out to analyze the influence of several interconnect layouts for the most critical layers in the top tier pad stack.

The investigated 3D IC is comprised of two stacked chips, resulting in a complex interconnect stack with a large number of metallization levels (MLs), as can be seen in Figure III.2. Therefore, a multi-scale approach will be used, with a global model consisting of a stack of equivalent homogeneous orthotropic layers and a local model including detailed interconnect geometry, similar to [17] and [7]. To compare the different pad configurations, a criterion based on the volume-averaged first principal stress will be used, consistent with the risk for brittle fracture in the dielectric material observed in the experimental study.

As discussed in the literature review, the mechanical reliability of bondpad interconnect stacks during and after the wire bonding process has been studied extensively. In particular, there are many numerical studies comparing different pad architectures and materials, or optimizing a given pad configuration. However, to the author's knowledge, so far this kind of work has not

been published for a 3D IC architecture with backside wirebonds, let alone from a simulation perspective. And yet, wire bonding on backside pads can be expected to yield mechanical failures due to the bonding surface being in close proximity to the lowermost metallization levels, containing local interconnects (see Chapter I). Furthermore, a notable feature of the literature reviewed in the previous section is that, apart from a few exceptions [21, 23, 32, 33], there is a lack of combined experimental and numerical studies focusing on the effect of pad configuration on wirebond mechanical robustness. In the following study, for correlation and validation purposes, the numerically-obtained pad architectures ranking will be compared to failure percentages obtained after completion of the wirebonding process on the investigated pad configurations.

In the next two sections, the experimental and numerical methodologies used in this work are detailed.

III.4 Experimental procedure

III.4.1 Cratering test

To verify pad mechanical integrity after completion of the wirebonding process, the cratering test is used. It consists in a deprocessing of the wirebond. Wires are etched off, enabling to check the bonding surfaces for any cracks or delamination due to ball impact, ultrasonic vibration or cooling to room temperature. In this study, the Au bonds were dipped in an aqua regia solution (mixture of nitric acid 69% and hydrochloric acid 36%) with an etching time of 7 minutes followed by a rinse in distilled water. The bonding surface is then visually inspected using an optical microscope. In case of damage in the dielectric layers below the bond, the reported failure mode is termed "pad cratering". The cratering test was carried out on five test chips to investigate the influence of different pad architectures on interconnect stack integrity after wirebonding. For each of the implemented pad configurations, failure percentages (in terms of number of cratering occurrences) were obtained.

III.4.2 Test chip

The investigated test chip is a 3D image sensor comprised of two distinct 8×6 mm dies vertically stacked and interconnected using the hybrid bonding process, as described in Chapter II. The bottom chip has a 7ML BEoL stack, fabricated using a 40 nm CMOS process, while the top chip is a 4ML BSI image sensor with backside wirebond pads capped by a 0.9-micron thick Al layer. Thermosonic bonding is carried out using a commercial bonder with 0.8 mil 4N[1] gold wire. The bonding parameters are summarized in Table III.7. No electrical probing of the pads has been carried out before the wire bonding process. This is to prevent any bias on bond robustness due

[1] Four-nines: 99.99% purity

to by scrub marks on the Al capping during the electrical test, and thus isolate the effect of the wire bonding process.

Time (ms)	Force (mN)	CV (mm/ms)	Temperature (°C)
16	16	0.8	150

TABLE III.1: Bonding parameters for the two bonding regions: bonding time, impact force, capillary tool constant velocity (CV) and temperature

An array of 32×6 pads with various architectures is dedicated to evaluate the influence of interconnect layout on pad robustness. To each of the six columns corresponds a different pad design, with specific capping materials and interconnect layout. These pads have lateral dimensions $75 \times 125\,\mu m^2$ with a pitch distance of $130\,\mu m$.

III.4.3 Test structures

III.4.3.1 Capping

For some of the investigated pads, buffer layers were introduced between the bonding surface and the interconnect stack:

Thin buffer: a 300 nm thick SiO_2 layer.

Thick buffer: a $SiO_2/Si/SiO_2$ multilayer (200 nm/2.8 µm/300 nm).

III.4.3.2 Metal lines

No active circuitry is present below the bonding surface in the top chip, thus affording great freedom for interconnect layout. For the metal levels of the top tier, three types of patterning were introduced, as illustrated in Figure III.3:

- **Pads:** a periodic array of square pads;
- **Grid:** a grid of evenly spaced interlaced perpendicular lines;
- **Plate:** a blanket layer having the same dimensions as the bonding pad.

Conversely, in the bottom chip the interconnection layout is defined by circuit design. Interconnects are forming a complex layout and no fixed configuration can be set.

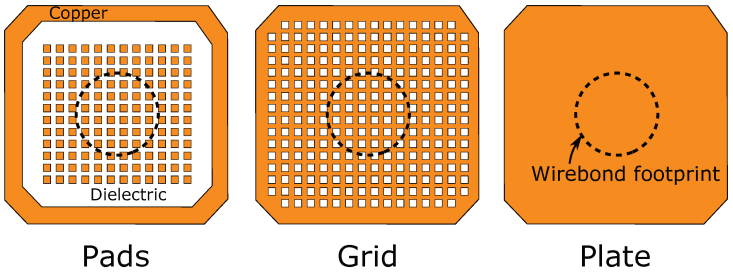


FIGURE III.3: Schematic description of the three types of patterning introduced for the metal interconnect architectures in the top chip.

III.4.3.3 Via plugs

The different metallization levels are vertically interconnected at the regions of overlapping between metal lines using via plugs, which can be arranged in the following two ways, depicted in Figure III.4:

- **Clusters:** a periodic array of vias at lines intersections;
- **Rows:** via lines running below the metal lines, across the edges;

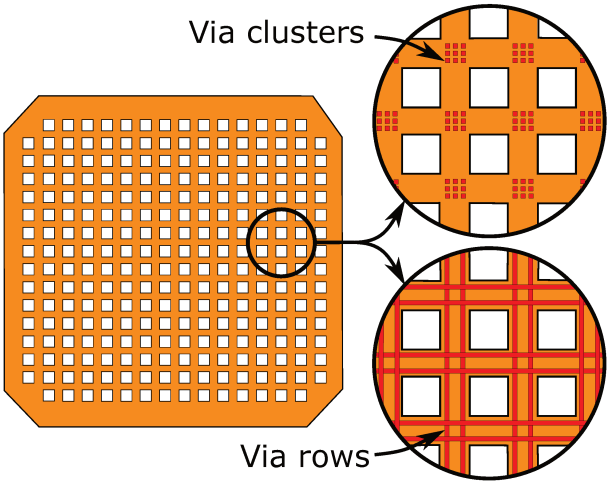


FIGURE III.4: Schematic description of the two types of patterning introduced for the via levels in-between crossing metal lines (for example here in the "grid" configuration).

III.4.3.4 Investigated pad architectures

A total of six distinct pad configurations are investigated, which are detailed in Table III.2 and illustrated in Figure III.5. Different combinations among the latter capping and metallization options are explored.

Pad structures associated with each of the three capping types listed in Section III.4.3.1 are referred to as configurations A, B and C:

Type A: direct contact between Al bonding surface and polycrystalline Si layer;

Type	Capping	Interconnect layout								
		Poly-Si	Contact	M1	Via 1	M2	Via 2	M3	Via 3	M4
A1	No buffer	Plate	Rows	Plate	Rows	Grid	Rows	Grid	Rows	Grid
A2		Plate	None	Pads	None	Grid	Rows	Grid	Rows	Grid
B1	Thin buffer	Plate	Rows	Plate	Rows	Grid	Rows	Grid	Rows	Grid
B2		Plate	None	Pads	None	Grid	Rows	Grid	Rows	Grid
C1	Thick buffer	Plate	Rows	Plate	Rows	Grid	Rows	Grid	Rows	Grid
C2		Plate	None	Pads	None	Grid	Rows	Grid	Rows	Grid

TABLE III.2: Listing of the different pad architectures implemented in the test chip.

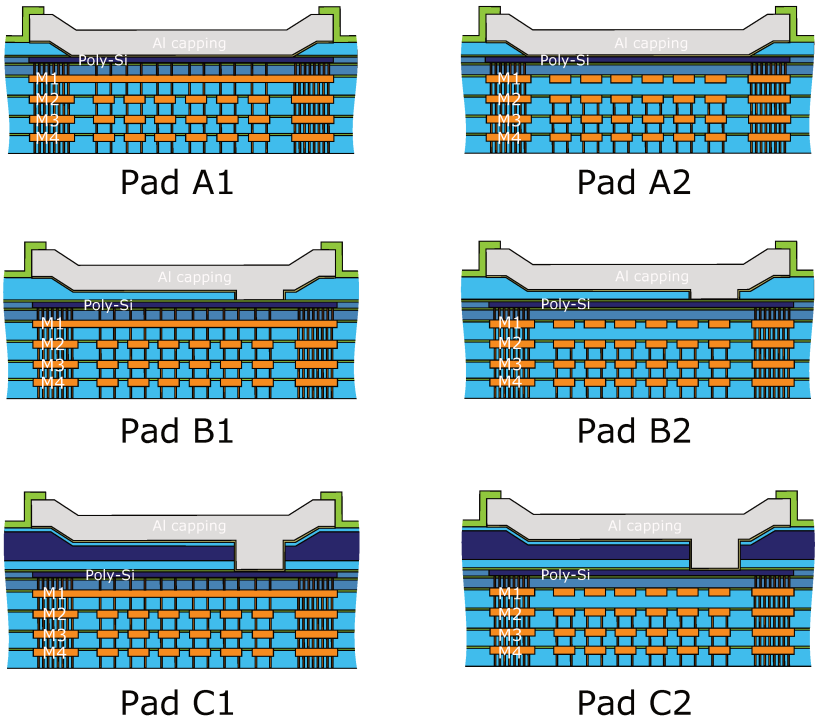


FIGURE III.5: Schematic description of the different pad architecture implemented in the test chip.

Type B: thin buffer layer;

Type C: thick buffer layer.

Regarding the interconnect structure, the focus is put on the first ML (M1), which is expected to be the most critical as mentioned in Section III.3. Two configurations are explored for M1, namely:

Type 1: Plate;

Type 2: Pads.

For the other three MLs in the top chip (M2 to M4), the "grid" configuration is adopted. The via plugs are arranged in the "rows" configuration.

III.5 Modeling methodology

A finite element model of the wire pull test on backside bondpads for a 3D image sensor with chips stacked by hybrid bonding was constructed using finite element code ANSYS Mechanical APDL, Release 17.0 (Ansys Inc., Canonsburg, PA), as depicted in Figure III.6. Similar to several analyses in the literature [7, 17–21, 27], a multi-scale static approach is used to resolve the detailed geometry of the BEoL structure below the bondpad. This feature is key to enable comparison of the mechanical robustness for different pad architectures. The approach is as follows:

- First, a global model with simplified features is used to capture the response of the coarse structure to a prescribed macroscopic load.
- The obtained nodal displacements are then extracted in the region of interest, interpolated, and prescribed as boundary conditions in a second smaller model.
- In this submodel, stress values can be obtained at a smaller scale, accounting for more detailed geometrical and material features.

III.5.1 Global model

The global model features the two-chip IC architecture assembled by hybrid bonding (bottom: 7ML + top: 4ML), with a bonded gold wire (only the extremity is represented). Similarly to [7, 18, 19, 25–27], a force load is prescribed on the wire section at an angle of 20° to mimic the wire pull test. A load of 1 gf is prescribed to the wire section, within the same order of magnitude as typical pull forces encountered in the wire-pull test. The magnitude of the load can be considered arbitrary however, due to the linear elasticity assumption. A half-model is used, accounting for symmetry, and only a fraction of the actual bonding surface is represented in the model since the lateral dimensions of the bondpad are large compared to ball bond diameter. The thick Si substrate is also modeled only partially. Displacements are blocked for nodes on the bottom and lateral boundaries (except the symmetry plane). A parametric study on the model dimensions was carried out to ensure convergence of the results in terms of the wire top displacement. The geometry is meshed using ∼1750000 hexahedral elements with quadratic interpolation. Elastic constants used for the isotropic linear elastic materials in the model are listed in Table III.3 and the mesh is depicted in Figure III.6. Perfect bonding is assumed both at the hybrid bonding interface and at the bonding area between the Au wirebond and the Al capping. Gold aluminide intermetallic compounds known to form during the thermosonic bonding process are not accounted for, due to the difficulty to characterize the mechanical properties and topology for the many AuAl phases present at the interface. The shape of the ball bond is derived from the dimensions of the capillary tool used for thermosonic bonding, as well as FIB/SEM cross-sections after wirebonding. A perfectly flat Au/Al bonding surface is assumed.

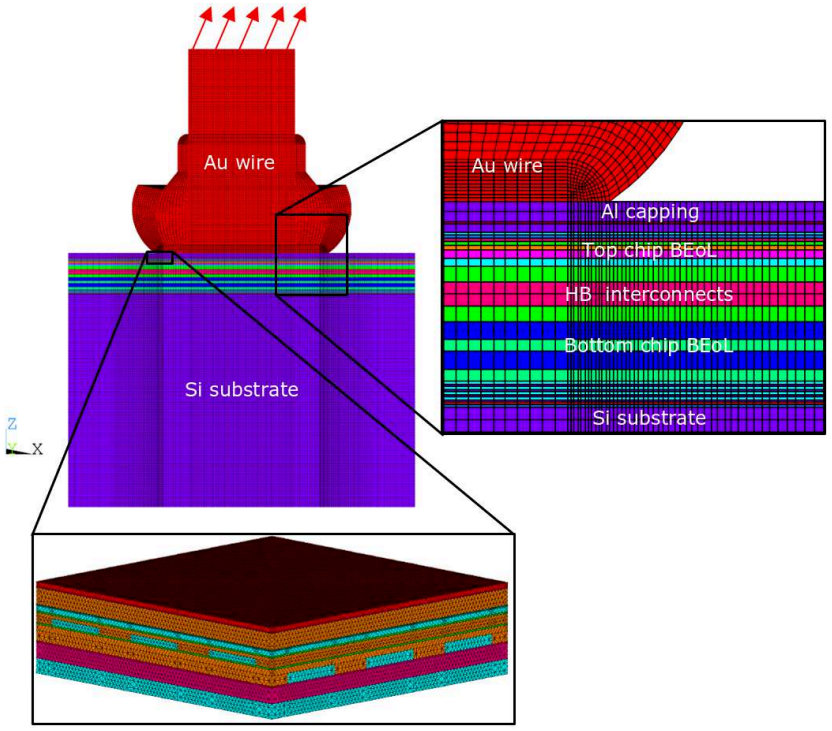


FIGURE III.6: Global and local finite element models used in the multi-scale approach for simulating the wire pull test (with zoomed view on the homogenized layers of the BEoL in the global model).

Material	Young's modulus (GPa)	Poisson's ratio
Cu	130	0.34
Al	70	0.33
Au	40	0.44
W	410	0.26
Si	130	0.28
poly-Si	169	0.22
SiOCH (ULK)	6	0.29
FSG	72	0.25
TEOS	85	0.25
USG	76	0.25
SiN (passivation)	170	0.24
SiN (barrier)	180	0.24
SiCN	79	0.25

TABLE III.3: Material properties used in finite element analysis of the wire pull test.

III.5.2 Homogenization procedure

At the global model scale, details of the BEoL metallization structure cannot be represented accurately, therefore a homogenization procedure was used [8, 17]:

- A representative unit cell (RUC) is isolated from the (periodic) multi-phase material. Periodic boundary conditions are enforced on the faces of the volume. This is done by assigning to each of the faces a reference "master" node, through which the loading is

prescribed. Then, displacements for each pair of facing nodes on opposite faces P and P' are coupled to those of the master node M through the following homogeneous constraints:

$$u_i^P - u_i^{P'} - u_i^M = 0, \quad \text{where } i \in \{x, y, z\}$$

To prevent rigid body motion, displacements are blocked on an "anchor" node.

- Using these boundary conditions, tensile or shear loading can be prescribed to the RUC in each direction.
- The resulting stress and strain can then be determined, either by extracting the nodal displacement and reaction forces on the faces, or by volume-averaging of element stresses and strains. This enables to derive material constants for an equivalent orthotropic linear elastic homogeneous material.

The obtained elastic constants can then be incorporated into the global model. This procedure is repeated for every layer in the BEoL stack below the bondpad, for each of the six investigated pad configurations.

III.5.3 Submodel

To obtain the stress values in the detailed BEoL structure during wire pull, a submodel is introduced in the location of maximum tensile stress below the wire. The volume of interest is a 3×3 repetition of a RUC. Only the 4ML stack in the top tier of the IC, believed to be the most critical location for crack initiation due to its proximity with the bond, is accounted for. The meshed geometry is presented in Figure III.7. More than 3000000 tetrahedral elements with quadratic interpolation are used and isotropic linear elastic material behavior are assumed, with the properties listed in Table III.3. As described above, the boundary conditions for the submodel are the interpolated nodal displacements obtained from the global model solution at the location of interest. Again, each of the six investigated pad architectures was modeled.

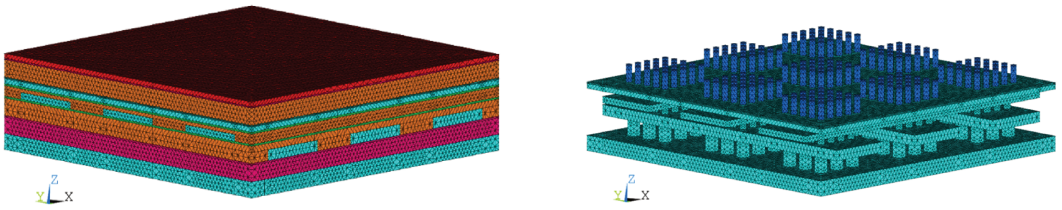


FIGURE III.7: Typical mesh and geometry used for the local model.

III.5.4 Robustness criterion

In this work, the investigated pad architectures are compared using a stress-based approach in the submodel. Consistent with experimental results presented in Section III.6.1, the targeted failure mode is brittle fracture in the dielectric layers close to the bonding region. Thus, the maximum first principal stress in the dielectric material is extracted in the submodel. A volume-averaging procedure is used:

$$< \sigma_1 >_{\mathcal{D}} = \frac{1}{V(\mathcal{D})} \int_{\mathcal{D}} \sigma_1(\mathbf{x}) \mathrm{d}\mathbf{x}$$

The reference domain $\mathcal{D}$ used for averaging is centered about the region of maximum stress for each dielectric layer in the submodel (Figure III.8). Constant dimensions were used for this volume, ensuring that the stress field remains as homogeneous as possible while including a sufficient number of elements to mitigate mesh-dependence issues caused by stress concentrations around the metal structures. A domain meeting those requirements is the $2.5 \times 5\,\mu\mathrm{m}^2$ region located along wire bond footprint, comprising two adjacent periodic unit cells and encompassing the whole thickness of the dielectric layer.

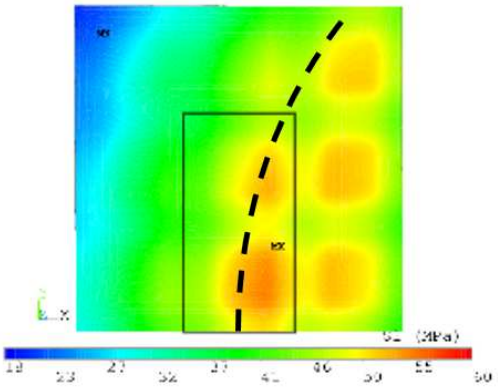


FIGURE III.8: Illustration of the stress-averaging procedure used in the analysis on a typical contour plots of the 1st principal stress for the dielectric material of the CO layer for pad C2 in the submodel (top view). The black dotted line marks the wire bond footprint, while the black plain box represents the domain $\mathcal{D}$ for stress averaging.

Such stress-averaging procedures have been used in the literature for instance:

- to deal with stress singularities in tapered laminate composite materials when using a delamination criterion based on maximum stress [45–47];
- in finite element analyses on solder joint failure reliability [48];
- to study debonding between encapsulation and leadframes in plastic IC packages [49].
- in a more theoretical context, the approach used is similar to some local, stress-controlled failure criteria in fracture mechanics, e.g. Neuber's fictitious notch rounding [50–52] or the RKR model for cleavage fracture [53].

Similarly to energy-based criteria this averaging procedure introduces a notion of characteristic length, corresponding here to the dimension of domain $\mathcal{D}$. As mentioned in Section III.2, [17] proposed to use for this characteristic dimension either a typical length for geometric features in the model (geometry-based) or an intrinsic length for the material obtained from experiments (material-based). For a material-based approach, a first approximation can be obtained based on Dugdale's estimate for the fracture process zone length in linear elastic fracture mechanics [54]:

$$L_{\mathcal{D}} \approx \frac{\pi}{8}\left(\frac{K_{\mathrm{Ic}}}{\sigma_Y}\right)^2$$

For SiO_2 dielectric (USG), with $K_{Ic} \sim 1\,\mathrm{MPa}\sqrt{\mathrm{m}}$ for the fracture toughness[1] and $\sigma_Y \sim 100 - 1000\,\mathrm{MPa}$ for the yield stress, the obtained length $L_{\mathcal{D}}$ ranges between 0.1 and 10 µm.

The dimensions of the considered averaging domain ($2.5 \times 5\,\mu\mathrm{m}^2$ over the thickness of the layer) are inside this estimated range. It can also be noted, for a geometry-based approach, that the estimated value also falls within the range of typical lateral dimensions for metal line interconnects.

In the numerical study, this stress-based criterion relying on a volume-averaging procedure will be used as a robustness index to compare the different pad architectures.

III.6 Results

In the following, first the experimental results obtained for the investigated six pad architectures in the cratering test are presented and correlated with finite element modeling of the wire-pull qualification test on the same structures. Then, these results are complemented with a numerical study focusing on the influence of metal lines and via plugs layout in the metallization level closest to the bonding surface. Finally, the mechanical robustness of the 3D BSI image sensor stacked by hybrid bonding with backside bondpads is compared with the conventional planar BSI image sensor architecture.

III.6.1 Cratering test

Configuration	Pad A1	Pad A2	Pad B1	Pad B2	Pad C1	Pad C2
Number of bonds	132	132	132	132	132	132
Cratering failure mode (%)	0.8	12	0	4	0	0

TABLE III.4: Failure rates obtained in cratering tests after completion of the wirebonding process.

The number of cratering occurrences for each pad structure variation is summarized in Table III.4. Failure percentages range between 0 and 12%. Pad A2 has the highest failure rate,

[1] H. Brillet-Rouxel, M. Verdier, M. Dupeux, M. Braccini, S. Orain, *Mater. Res. Soc. Symp. Proc.* **914** (2006).

while no cratering failure modes were detected for pads B1, C1 and C2. Overall, the obtained failure rates are higher for pad type A compared to B and C. Another observation is that within each of these three categories, pad type 1 yields less cratering occurrences than type 2.

Regarding the influence of the capping type, direct contact between the Al bondpad and the interconnect stack (type A) seems to be the worst configuration in terms of mechanical robustness. Conversely, with the introduction of a buffer layer (cases B and C) failure percentages significantly decrease. The effect of capping thickness is also clearly apparent: cratering occurrences are considerably reduced for configuration B2 compared to configuration A2, and completely suppressed for pad C2 having additional layers of larger thickness (more than three times that of the Al pad).

Regarding the influence of interconnect layout, the "plate" configuration for the first ML (M1) below the bonding surface (pads A1, B1 and C1) seems to be more mechanically robust compared to the "pads" configuration. This is evidenced by the low failure percentages obtained in the cratering test for pads A1 and B1 compared to A2 and B2. However, no conclusions can be drawn on the influence of M1 interconnect layout between pads C1 and C2, as the associated failure percentages are both zero.

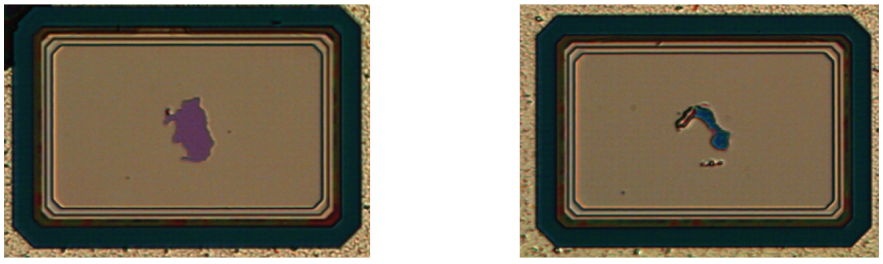
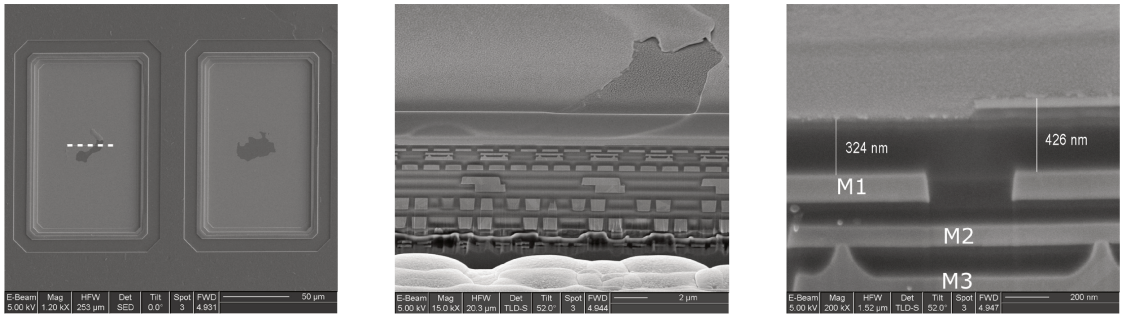


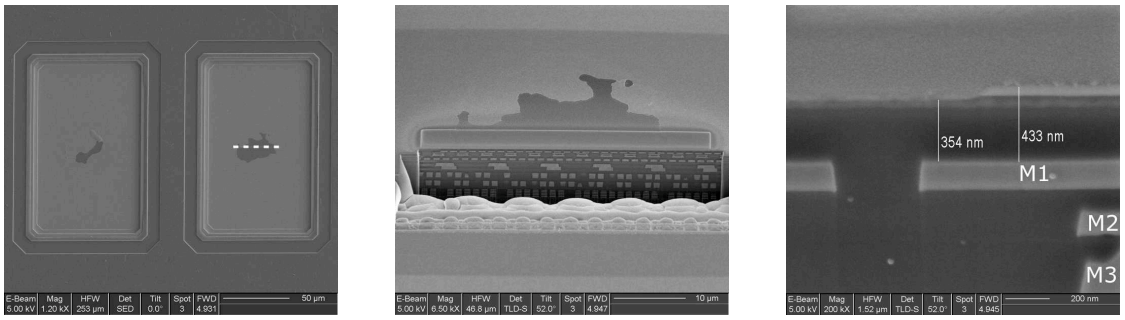
FIGURE III.9: Top-view of some typical cratering failure modes observed on the bonding surface for pad configuration A2 (optical microscope).



(a) (b) (c)

FIGURE III.10: (a): SEM top-view of two damaged pads (type A2); (b): FIB/SEM cross-sectional view showing the interconnect stack of the right-hand pad; (c): Zoomed view revealing cracking in the poly-Si layer. A Pt/Au layer was deposited on the pad surface for sample preparation before FIB milling.

Some typical failures obtained for pad configuration A2 in the cratering test are depicted in Figure III.9. In those failed pads, the Al capping was peeled off and the underlying layers in



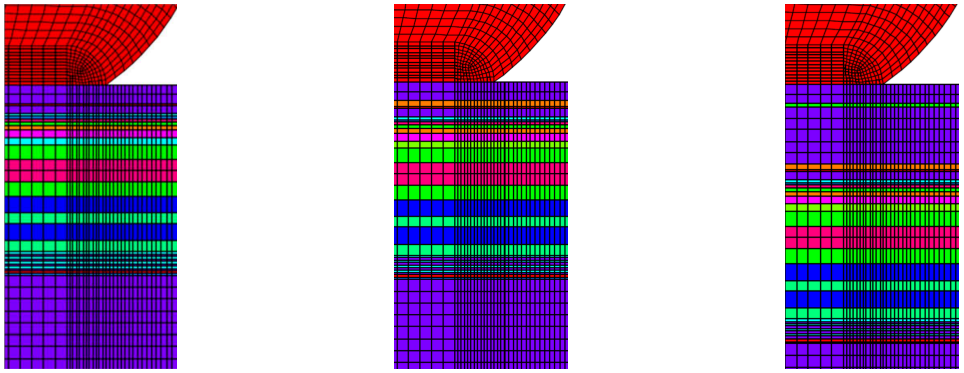
(a) (b) (c)

FIGURE III.11: (a): SEM top-view of two damaged pads (type A2); (b): FIB/SEM cross-sectional view showing the interconnect stack of the left-hand pad; (c): Zoomed view revealing cracking in the poly-Si layer. A Pt/Au layer was deposited on the pad surface for sample preparation before FIB milling.

the interconnect stack are visible. In Figure III.10 and Figure III.11, FIB/SEM cross-sectional views of the fracture surfaces are shown. For these pads, the poly-Si layer was damaged. Overall, visual inspection of failed pads indicated that the poly-Si, CO layer, and in some cases M1 layer were exposed.

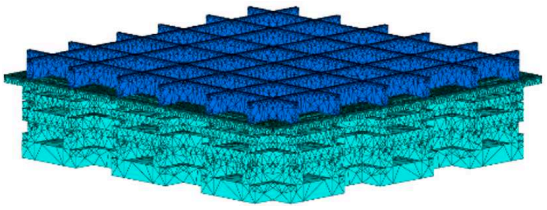
III.6.2 Numerical comparison

In this section, the numerical model and methodology described in Section III.5 are applied to the studied pad configurations (Table III.2), aiming to provide better understanding and correlation with experimental results. The different capping types and interconnect layouts corresponding to the six pad architectures are depicted in Figure III.12 and III.13, respectively for the global model and the submodel.

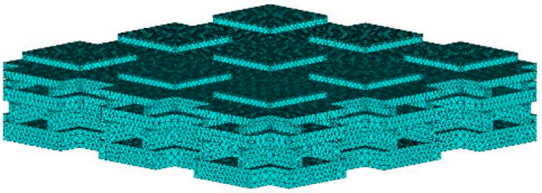


(a) Pads A1, A2 (b) Pads B1, B2 (c) Pads C1, C2

FIGURE III.12: Overview of the different pad cappings associated with each of the studied pad configurations (global model).



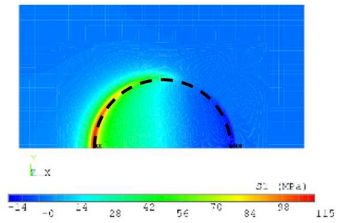
(a) Pads A1, B1, C1



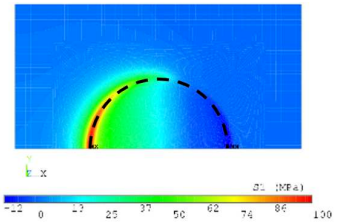
(b) Pads A2, B2, C2

FIGURE III.13: Overview of the different interconnect layouts associated with each of the studied pad configurations (submodel).

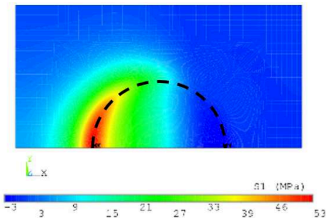
III.6.2.1 Global scale



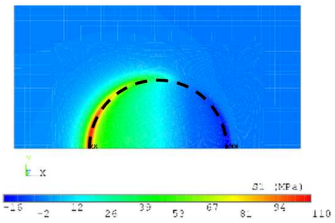
(a) Pad A1
$\max(\sigma_I) = 115\,\text{MPa}$



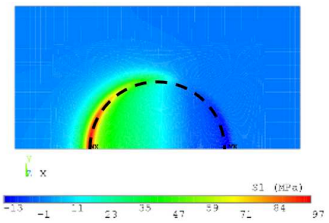
(b) Pad B1
$\max(\sigma_I) = 100\,\text{MPa}$



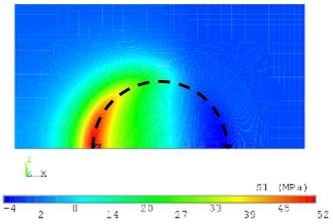
(c) Pad C1
$\max(\sigma_I) = 53\,\text{MPa}$



(d) Pad A2
$\max(\sigma_I) = 110\,\text{MPa}$



(e) Pad B2
$\max(\sigma_I) = 97\,\text{MPa}$



(f) Pad C2
$\max(\sigma_I) = 52\,\text{MPa}$

FIGURE III.14: Contour plots of the 1st principal stress in the homogenized CO layer for the studied pad structures (top view). The black dotted half-circle marks the wire bond footprint.

Contour plots of the 1st principal stress in the homogenized CO layer in the global model for each pad configuration are shown in Figure III.14. The focus was put on this layer since failures in the cratering test were observed in the poly-Si, CO and M1 levels, and because the CO layer is the metallization layer closest to the bonding surface. For clarity, the footprint of the wire bond was also represented on the plots. Two distinct regions are immediately apparent on each side of the bond along the direction of the pull force, one mostly subjected to tensile stress and the other to compressive stress. As discussed in Section III.5, the submodel in the multi-scale approach is located around the region of maximum tensile stress.

Capping type has a strong impact on stress repartition below the pad. From pads A to C, a sharp decrease of the maximum tensile stress is observed in the CO layer. This trend is especially marked between pads B and C, with a reduction in the maximum stress of about 50%. As apparent from the contour plots, this is due to a redistribution of the load on the wider area with increasing capping thickness. However, in these first results only considering the peak tensile stresses in the global model the influence of interconnect layout is much more limited. A 4% variation is obtained at most between pad types 1 and 2, the latter having slightly lower maximum values.

The obtained difference between interconnect layouts 1 and 2 is actually directly correlated to the homogenized equivalent mechanical properties associated with each pad type. Applying the homogenization procedure described in Section III.5 on the whole interconnection layout, it is confirmed that pad type 2 has a slightly lower equivalent transverse Young's modulus E_{zz} compared to type 1 (Figure III.15). While the results obtained with the global model were sufficient to capture the effect of capping thickness on bondpad mechanical robustness, they are not adequate to conclude on the relative mechanical robustness of the investigated interconnection layouts.

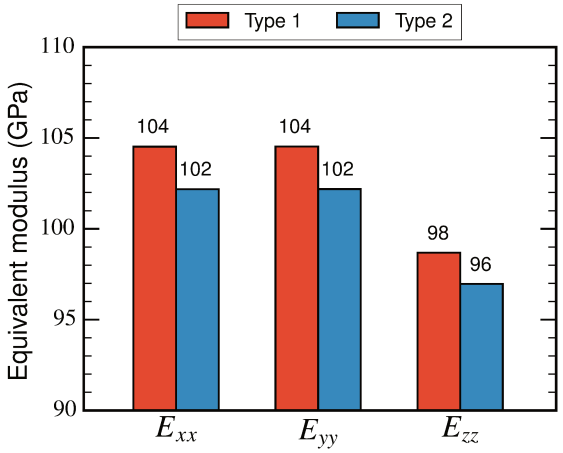
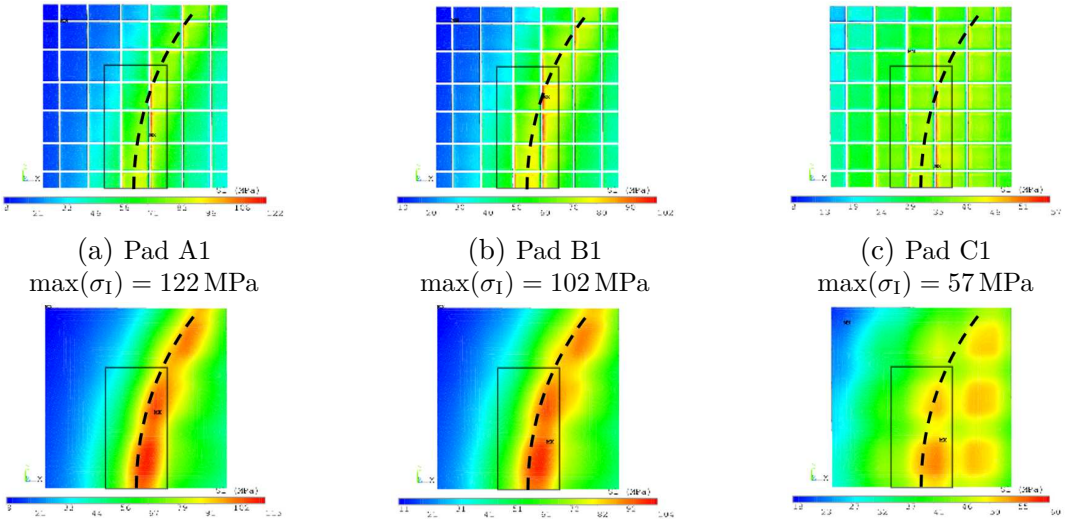


FIGURE III.15: Comparison of equivalent Young's moduli E_{ij} for the whole top tier BEoL stack of pad types 1 and 2 after mechanical properties homogenization.

This shows the critical importance of submodeling, which allows to account for detailed geometric features in the interconnect stack and thus stress concentrations due to material heterogeneities and sharp geometrical features, presumably at the origin of crack initiation in brittle dielectric.

III.6.2.2 Local scale



(a) Pad A1
$\max(\sigma_I) = 122\,\text{MPa}$

(b) Pad B1
$\max(\sigma_I) = 102\,\text{MPa}$

(c) Pad C1
$\max(\sigma_I) = 57\,\text{MPa}$

(d) Pad A2
$\max(\sigma_I) = 115\,\text{MPa}$

(e) Pad B2
$\max(\sigma_I) = 104\,\text{MPa}$

(f) Pad C2
$\max(\sigma_I) = 60\,\text{MPa}$

FIGURE III.16: Contour plots of the 1st principal stress in the dielectric material of the CO layer for the studied pad structures in the first local model (top view). The black dotted line marks the wire bond footprint, while the black plain box represents the domain $\mathcal{D}$ for stress averaging.

Contour plots of the 1st principal stress in the dielectric material of the CO layer in the global model are shown for each pad configuration in Figure III.16. Comparing the peak tensile stress values, again the impact of capping type is immediately apparent with a significant decrease from capping A to C due to a "buffer" effect of the passivation layer with increasing thickness. These values are quite close to those computed with the global model, but the ranking for the six pads is not the same. Similarly to the global model, the difference between the peak stresses for the two types of BEoL architectures (1 and 2) is modest. It can be noted however that the computed stress distributions are quite different for type 1 and 2:

type 1: stress concentration regions are highly localized and sparsely distributed due to the presence of interlaced tungsten contacts rows;

type 2: stress in the dielectric layer spreads more uniformly in regions corresponding to the imprints of the underlying Cu pads in the M1 metallization level. Although stress peaks are eliminated in pad type 2, the magnitude of the uniform stress remains quite high, on the same order as the maximum values obtained in the type 1 pad.

Such features would not be captured if a criterion based on maximum stress values were used. The stress-averaging approach, by introducing a characteristic length, accounts for the lateral spread of the stress peaks and should thus provide a more suitable criterion to study the relative mechanical robustness of the investigated pad configurations.

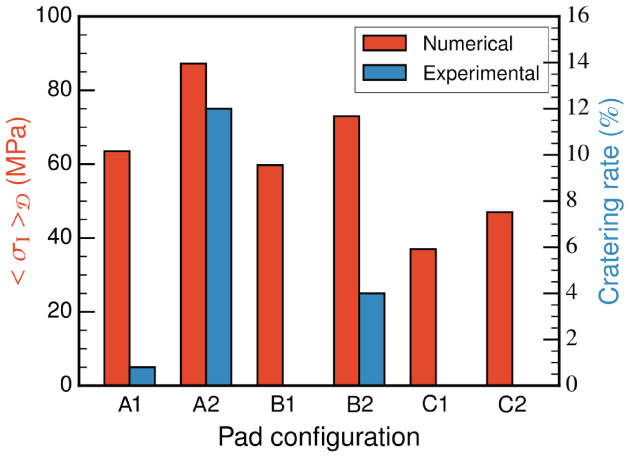


FIGURE III.17: Comparison of volume-averaged maximum 1st principal stress from the first local model and failure rates from the cratering test for the studied pad configurations.

In Figure III.17, the averaged 1st principal stress in the dielectric material of the CO layer is plotted against the failure percentages obtained in the cratering test for each pad structure. There is good agreement between the numerical and the experimental results. The same ranking is obtained in terms of bondpad mechanical robustness: A2, B2 and A1 (by order of decreasing robustness).

Contrary to the maximum stress criterion for which very few differences were seen between type 1 and type 2 pads, the stress-averaging approach seems to provide better sensitivity to interconnect layout. However, numerical and experimental results could not be compared for the pad structures considered most reliable (B1, C1 and C2), precisely because no failures were observed for these three pads in the cratering test.

III.6.3 Parametric study

The numerical model is now applied to investigate the impact of chip architecture on pad mechanical robustness, aiming to determine the most important factors and how they can be combined to optimize pad robustness. The influence of interconnect layout in the top chip is investigated for both metal and via structures.

In this section, the influence of top chip metal and via interconnect layout on bondpad mechanical robustness is studied. Regarding the capping type, direct contact between the Al bondpad and the top chip interconnect stack is used for the whole analysis, as it was found both experimentally and numerically to be the worst case in terms of mechanical robustness (Sections III.6.1 and III.6.2). The focus is on the metallization levels closest to the bonding surface since failure is assumed to occur primarily in these locations, based on experimental results. Using constant meshing[1], several layouts are investigated for the CO and M1 layers. For the remaining metallization levels (M2 to M4), the "grid" layout is used for the metal levels and the "clusters"

[1]Except for the "rows" configuration.

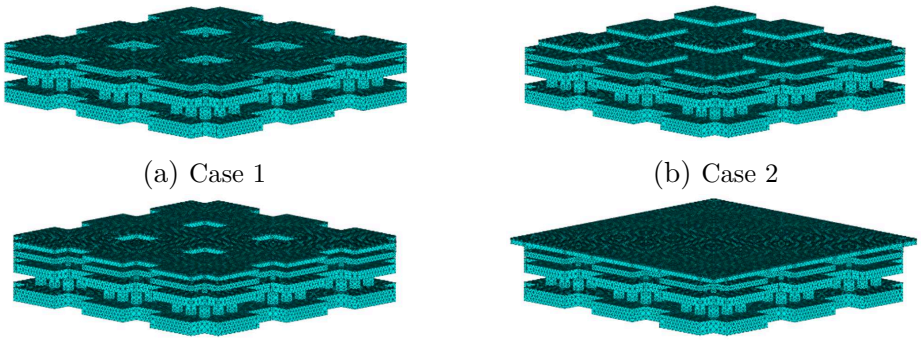
structure for the via, as they are deemed representative of a typical wire bondpad interconnect stack compared respectively to the "pads" and "rows" structures. In this study only the mechanical performance is considered, disregarding electrical aspects and ease-of-processing for the different structures.

III.6.3.1 Effect of metal lines interconnect layout

All three basic metal lines configurations proposed in Section III.4.3, namely "pads", "grid" and "plate", are investigated for the M1 level. The different pads are listed in Table III.5 and the associated submodels showed in Figure III.18. No vias were introduced above or below the M1 layer to isolate the influence of the metal lines layout. The results of wire-pull simulation are plotted in Figure III.19 for the global and local models (respectively the maximum 1st principal stress in the homogenized M1 layer, and the averaged 1st principal stress in the dielectric material of the M1 layer).

	Interconnect layout									M1 Cu surface fraction
	Poly-Si	Contact	M1	Via 1	M2	Via 2	M3	Via 3	M4	
Case 1	Plate	None	None	None	Grid	Clusters	Grid	Clusters	Grid	0%
Case 2			Pads							16%
Case 3			Grid							84%
Case 4			Plate							100%

TABLE III.5: Investigated architectures in the study of the effect of metal lines layout on pad mechanical robustness.

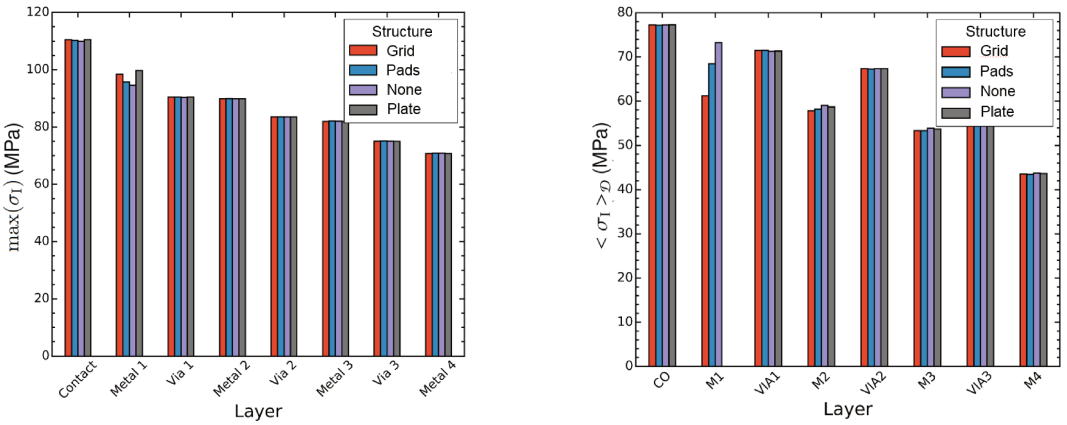


(a) Case 1 (b) Case 2

(c) Case 3 (d) Case 4

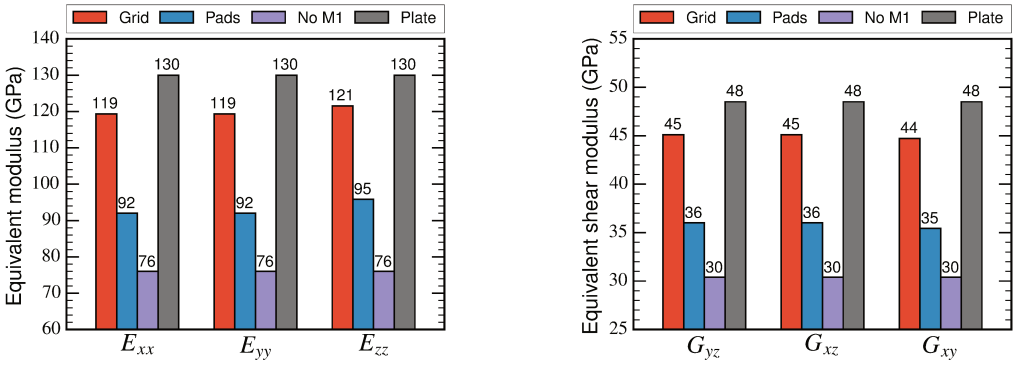
FIGURE III.18: Mesh and geometry (only Cu shown) used in the local models of the four pad variations in the study of the effect of metal level interconnect configuration.

In the global model, stress variations due to changes in the interconnect layout are very local: only the M1 level is affected (Figure III.19a). It can also be noted that for this layer, the obtained maximum tensile stress increases with metallization surface fraction. This is because in the global model, due to material homogenization, the maximum tensile stress in a given layer depends only on the elastic properties of the considered equivalent orthotropic material. Indeed, from case 1 to case 4 the Cu surface fraction in the M1 layer increases from 0 to 100%, with respectively 16% and 84% for case 2 and 3. Since Cu has a significantly larger Young's modulus



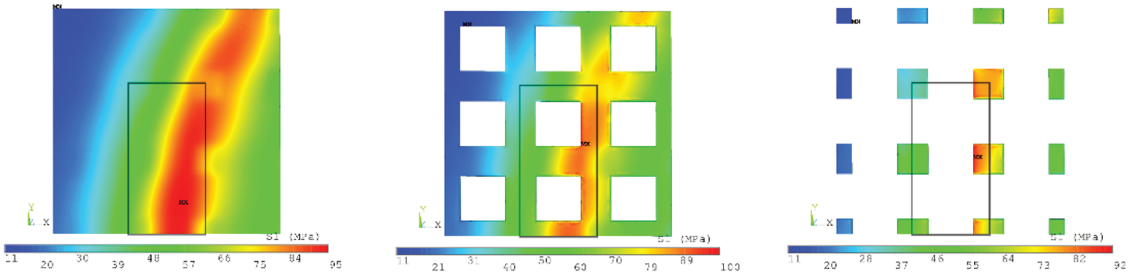
(a) Global model (b) Local model

FIGURE III.19: Comparison of stresses in the different layers of the top chip stack with different metal line configurations for the global model and the submodel. In the submodel, the results for case 4 cannot be compared to the other configurations, since no dielectric is present.



(a) Equivalent Young's moduli (b) Equivalent shear moduli

FIGURE III.20: Comparison of equivalent Young's moduli and shear moduli for the four investigated M1 layer architectures after material homogenization.



(a) Case 1
$\max(\sigma_I) = 95\,\text{MPa}$

(b) Case 2
$\max(\sigma_I) = 100\,\text{MPa}$

(c) Case 3
$\max(\sigma_I) = 92\,\text{MPa}$

FIGURE III.21: Contour plots of the 1st principal stress in the dielectric material of the M1 layer for the studied pad structures in the local model (top view). The black dotted line marks the wire bond footprint, while the black plain box represents the domain $\mathcal{D}$ for stress averaging.

than SiO_2, the equivalent elastic moduli for the homogenized layer significantly increase from case 1 to 4 (Figure III.20). As a result, the load transmitted to the submodel must also increase.

In the submodel however, the obtained ranking is not the same as in the global model (Figure III.19b and Figure III.21). Instead, the largest volume-averaged stress is obtained for case 1, having the lowest equivalent elastic moduli (Figure III.20). This result can be interpreted in terms of a competition at the local scale between two effects:

(i) stress relief in the dielectric material due to the load-bearing capability of the interconnect layout;

(ii) stress concentration due to elastic mismatch and sharp geometric features introduced by metal patterning.

From the submodel results, it seems that the load-bearing effect of the metal interconnect layout is predominant compared to the elastic mismatch. Indeed, for case 3 the averaged dielectric stress is the lowest despite high metallization density (84%), whereas for case 1 the averaged dielectric stress is maximum even though the layer is homogeneous and the applied load from the global model is the lowest of all four cases.

III.6.3.2 Effect of via plugs layout

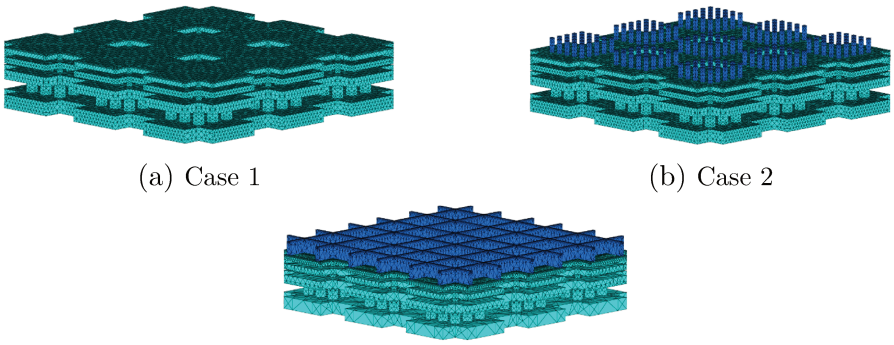
The three basic via plugs configurations proposed in Section III.4.3 are investigated for the CO layer. The resulting pads configuration are listed in Table III.6 and the associated local finite element models illustrated in Figure III.22. For all metal layers the "grid" configuration was used, as it is deemed more representative of a typical wire bondpad interconnect stack compared to the "pads" and "plate" structures. For the same reason, for via layers the "clusters" configuration was used. No vias were introduced between the M1 and M2 layers however, in order to better isolate the effect of via plugs layout.

	Interconnect layout									CO W surface fraction
	Poly-Si	Contact	M1	Via 1	M2	Via 2	M3	Via 3	M4	
Case 1		No CO								0%
Case 2	Plate	Clusters	Grid	No VIA1	Grid	Clusters	Grid	Clusters	Grid	0.6%
Case 3		Rows								5.8%

TABLE III.6: Investigated architectures in the study of the effect of via interconnect layout on pad mechanical robustness.

The results of wire pull simulation are plotted in Figure III.23 for the global and local models (respectively the maximum 1st principal stress in the homogenized CO layer, and the averaged 1st principal stress in the dielectric material of the CO layer). Overall, the results are very similar to the trends obtained regarding the effect of metal level interconnect layout.

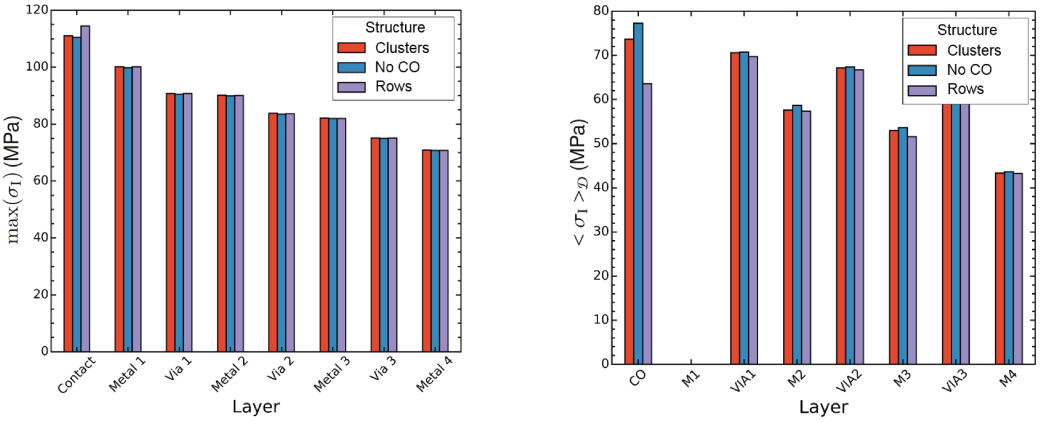
In the global model, the obtained peak stress in the CO layer is directly correlated to metal surface fraction (Figure III.23a). Similarly to the previous section, this is due to the use of an equivalent orthotropic material with homogenized properties. It can also be noted that although the metal surface fraction is very low for the CO layer compared to metal levels, a substantial variation of the equivalent elastic properties is obtained nevertheless between the different via



(a) Case 1 (b) Case 2

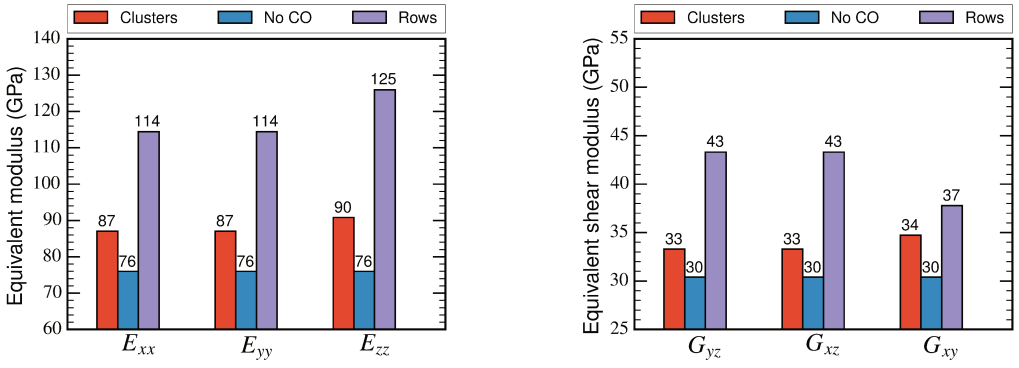
(c) Case 3

FIGURE III.22: Mesh and geometry (only Cu and W shown) used in the local models of the three pad variations in the study of the effect of via level interconnect configuration.



(a) Global model (b) Local model

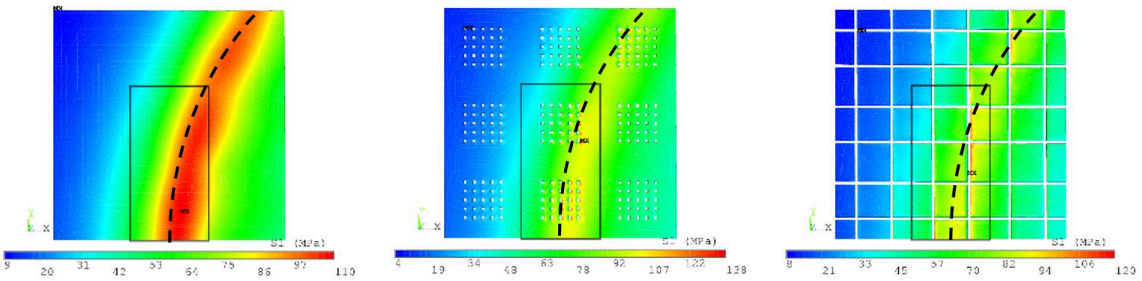
FIGURE III.23: Comparison of stresses in the different layers of the top chip stack with different via configurations for the global and local models.



(a) Equivalent Young's moduli (b) Equivalent shear moduli

FIGURE III.24: Comparison of equivalent Young's moduli and shear moduli for the three investigated CO layer architectures after material homogenization.

layouts. For example, comparing cases 2 and 3 with an increase of metal surface fraction from 0.6 to 5.8%, the equivalent transverse modulus E_{zz} goes up by about 25% (Figure III.24). This is believed to result from two causes:



(a) Case 2
$\max(\sigma_I) = 110\,\text{MPa}$

(b) Case 3
$\max(\sigma_I) = 138\,\text{MPa}$

(c) Case 4
$\max(\sigma_I) = 120\,\text{MPa}$

FIGURE III.25: Contour plots of the 1st principal stress in the dielectric material of the CO layer for the studied pad structures in the local model (top view). The black dotted line marks the wire bond footprint, while the black plain box represents the domain $\mathcal{D}$ for stress averaging.

- the large Young's modulus of tungsten compared to copper[1];
- transverse stiffening due to a structure effect for interlaced via rows in case 3, enhancing the homogenized elastic properties compared to individual vias arranged in clusters for case 2.

Therefore, the load transmitted to the submodel is the largest for case 3 and the lowest for case 1. In the submodel however, similarly to the previous section, the obtained ranking is different. The lowest averaged stress is obtained for case 3 having the largest metal surface fraction, while the highest average stress is obtained for case 1 in which there are no contact plugs (Figure III.23a). This is believed to result from a significant fraction of the total loading being transmitted through the stiff W contact plugs, effectively providing stress relief to the surrounding dielectric material. In addition, it can be noted that even though no via plugs are present between M1 and M2, the stress state in all four metal levels seems to be impacted (although quite modestly) by contact plugs layout. This was not the case in the study on the effect of M1 layout, for which no contact plugs were introduced. Therefore, it may be inferred that contact plugs play an important role in transmitting the load exerted on the wirebond to the underlying interconnect stack.

In order to check for any strong mechanical interactions between CO and M1 interconnect patterns, the above analysis was carried out again with the "grid" structure for the M1 layer, instead of "plate" structure. Even though some mild variations in stress levels are obtained between the two cases (with slightly lower values for the "grid" structure), the results are globally the same. This could indicate that there is little interaction between the effects of contact plugs and M1 lines configuration.

[1]The Young's modulus of Cu is twice larger than that of SiO_2, compared to about five times for tungsten.

III.7 Discussion

III.7.1 Influence of bondpad capping

Among the investigated parameters, bondpad capping is identified as having the most impact on bondpad mechanical robustness, both in experimental and simulation results. For the considered 3D IC with backside bondpads, the introduction of a passivation layer between the Al bonding surface and the underlying interconnect stack provides an effective way to reduce the risk of mechanical failure in the narrow and thin metal lines present nearby the bonding surface. Acting as buffer, it enables to redistribute stresses over a larger area and thus to attenuate the loading exerted by the wirebond on the interconnect stack.

The thickness of the capping layer has a strong impact on the levels of stress exerted on the pad structure. For a thick buffer (type C), as with the Si layer (three times thicker than Al capping, twice larger Young's modulus), the number of failures in the cratering test vanishes regardless of the interconnection layout. However, capping material may have an even stronger influence. Indeed, for a thin buffer (type B), as with the silicon oxide layer (three times thinner than Al capping, similar Young's modulus) a significant decrease of failure percentages is obtained compared to the case of direct contact between Al capping and the interconnect stack.

This result is attractive from a practical standpoint, as SiO_2 layer deposition is a relatively straightforward and ubiquitous process step for typical ICs, whereas the integration of a thick Si layer may require more complex processing. In addition to wirebonding, the introduction of capping layers could also enable to improve mechanical robustness during electrical probing of the bondpad. However, it can be noted that these additional layers may also have potentially detrimental consequences, e.g. in terms of optical performance or interface adhesion.

A question not addressed here is whether increasing Al capping thickness may lead to similar results. In general, increasing Al capping thickness should indeed result in lower stress in the interconnect stack. However, Al thickness plays an important role during bond formation due to mechanical interaction between the wire and the bonding surface. Excessive thickness may actually result in mechanical reliability issues due to larger "splash-out" around the bond after ball impact and ultrasonic vibration, but thinner "Al remnant" below the bond [55, 56].

III.7.2 Influence of interconnect layout

Having observed in the experimental study that the most critical locations for failure in the pad structure are the intermetal (or interlayer) dielectric layers, the influence of interconnect layout on pad mechanical robustness may be interpreted in terms of a competition between two effects:

- At the pad scale, the load-bearing capability of Cu interconnects in a given metallization layer, effectively shielding the dielectric matrix from transverse normal loading;
- At the local scale, the elastic mismatch between Cu metallization and SiO_2 dielectric, and stress concentration around sharp corners or edges.

From experimental results, the load-bearing capability of the interconnect layout seems to have a major impact on pad mechanical robustness. Indeed, among both studied pad configurations, that with fully connected metal lines layout for the M1 layer (type 2) yielded significantly lower failure percentages in the cratering test. In the simulation study, two main contributors to this structure effect of the interconnect stack were identified:

Metallization surface fraction: in a given metallization level, a larger surface fraction for the through-layer Cu interconnects contributes to relieve transverse normal stresses in the surrounding dielectric matrix, by allowing better load redistribution towards the underlying layers.

Pattern connectedness: layouts with connected metal lines (or vias) provide a stiffening effect to the pad due to interconnects acting as reinforcement bars, thereby enabling better redistribution of shear loads and thus stress relief in the dielectric material.

Related to the issue of metallization structure connectedness is the presence or absence of via plugs between the different metal levels. In the experimental study, the presence of via plugs between all metal levels for the type 2 pad could be an additional reason for its better robustness in the cratering test compared to the type 1 pad, with no vias above or below the M1 layer. In the simulation study, considering an isolated M1 layer (no vias above or below), the computed stress variations in the dielectric material in response to an interconnect layout modification remain very local, with few impact on the rest of the interconnect stack. In the presence of stiff W contact plugs, layout changes in the M1 layer lead to slightly larger variations (although still very modest) in the rest of the interconnect stack. However, when the contact layer and the M1 layer are connected, no coupling effect was identified between their respective layouts.

Additional experiments are needed to confirm the effect of via presence on the stress distribution in the different metallization levels of the pad interconnect stack. This effect, if confirmed, could have important consequences in practice for pad structure design. For example, it is expected that in applications with backside bondpads, a pad structure with a via layer between each pair of adjacent metallization levels could be more robust, by enabling stress redistribution from the thin topmost interconnect layers towards the thicker lowermost interconnects. Conversely, in

applications with frontside bondpads, a pad structure without vias in the thin lowermost interconnect layers could help to confine stresses in the thicker topmost interconnects. In particular, such a feature would be attractive for BOA pads, to prevent load transmission to the active region.

The results obtained on the influence of interconnect layout are in agreement with studies by [32] and [8], who investigated different metal lines structures and found the "grid" configuration (type 2) to be the most robust, both in qualification testing and finite element analysis using a stress-based criterion. This finding was confirmed in [21], using an energy-based criterion to perform the same comparison. [32] attributed this robustness to a "mechanical reinforcement" effect with the "grid" configuration.

[25] also identified metallization surface fraction as having an influence on pad mechanical robustness in finite element calculations and argued that "metal density in a single layer [...] has an impact on the performance of the remaining layers". This effect was also observed to some extent in this study for metal lines layers, however with relatively small variations of the stress levels in the remainder of the stack for large variations in the metal surface fraction of the considered layer. For via plugs layers, the results of the present study are in agreement with the observation of [33] that vias only have a local effect on the stress distribution in the pad, with minimum impact on the layers below.

Regarding the local effect of via plugs on the stress distribution, consistent with the results of this study, finite element analyses by [5], [34] and [15, 35, 36] showed a significant effect of via density on the stress levels at the local scale, with a large stress decrease in the dielectric material for increasing via density. [5] attributed this effect to the vias plugs providing a "mechanical supporting structure" to the pad. On the contrary, [12] obtained a stress increase in the dielectric material for increasing via density. This discrepancy may be explained by distinct interconnect layouts being considered in the different studies. It can also be noted that two-dimensional models were used in all the latter, which depending on cases may lead in overestimation or underestimation of the influence of via plugs compared to metal lines. In addition, there is a lack of relevant experimental data for comparison to the finite element results.

In the literature, no clear distinction was made between the respective contributions of the "metallization surface fraction" and "pattern connectedness" mechanisms identified above. In the present study, with the investigated pad configurations these two effects could not be completely separated either, and thus further work is needed. For example, the different interconnect layouts presented here could be compared at constant metallization surface fraction, enabling to isolate the effect of pattern connectedness. It can be noted however that working with constant metal surface fraction may require varying metal pattern width or spacing. However, with the comparison criterion used here, i.e. stress-averaging over a constant domain defined by an intrinsic length associated with the material, an additional bias due to interconnect characteristic length dependence may be introduced.

III.7.3 Methodology

III.7.3.1 Experimental study

The cratering test was used in this study to compare the mechanical robustness of several bondpad configurations. This technique enables *post mortem* observation of the bonding surface after the wirebonding process by chemical etching of the Au wire. Optimized bonding parameters were used for the investigated samples, representative of bonding conditions used in practical applications. Although not detailed here, several verifications were carried out before the test (e.g. bond shape measurements, intermetallic compounds coverage), aiming to rule out extrinsic causes for failure and to ensure the homogeneity of the investigated wirebond population.

The main advantage of this technique is its ease-of-implementation and high throughput. Indeed, all bonds can be etched simultaneously, enabling to inspect a large population. However, with the protocol used here only through-cracks are directly detectable[1]. A possible solution to detect blind cracks is a deprocessing of the bondpad by successively etching the different layers in the interconnect stack, which was not carried out here. In both cases, time-consuming FIB/SEM cross-sections are required to determine the damaged layers or interfaces.

It can be noted that among the investigated bondpads, no cracks were detected for the two configurations with thick capping, thus preventing direct comparison. To trigger more failures and thus obtain more statistics, additional testing could be carried out using slightly more aggressive bonding parameters. However, this is generally not a desirable option, as the bonding conditions would then no longer be representative. Thus, qualification methods such as the ball shear test (JESD22-B116B) and/or the wire pull test (ASTM F 459-06) are an attractive alternative (or supplement) to the cratering test. Because the bond reaction force is monitored against tool displacement during the test, quantitative data can be obtained. In addition, since these qualification procedures include failure mode determination after testing, more comprehensive information is gained for bondpad robustness assessment, or even failure mechanism identification. Indeed, even for failure modes with no apparent cracking of the bonding surface, cracks and delamination inside the interconnect stack may still be revealed by sharp drops in the force-displacement response during the test.

[1] Although in some cases, blisters may indicate delamination inside the interconnect stack, without surface cracks.

III.7.3.2 Simulation study

Good agreement was obtained between experimental and simulation results, although a distinct loading is involved in each case. Failures observed in the cratering test are due to the impact force and ultrasonic lateral vibration of the capillary tool during the wirebonding process, whereas in the simulation study an oblique force loading is prescribed on the wire, corresponding to the wire pull qualification test. Although primarily motivated by the need for a simple and efficient model in an industrial context, wire pull simulation is also deemed relevant for comparison with cratering test results, since in both cases the load includes a normal and a shear component. It can be noted however that a compressive load is prescribed on the pad during the wirebonding process, whereas in the wire-pull test the load is tensile. For verification purposes, the pad architecture comparison carried out in Section III.6.2 was therefore repeated locating the submodel in the maximum compressive stress region, leading to the same results.

The correlation between experimental and simulation results may also be explained by the use of a local stress-averaging method, introducing of a notion of characteristic length in a stress-based criterion and thereby emulating an energetic approach. Although not sufficient to serve directly as a criterion for crack initiation, this approach is believed to provide a suitable reliability index to compare the relative mechanical robustness of different pad configurations while retaining the ease of implementation and computational cost-effectiveness of a stress-based analysis. In addition, this local stress-averaging criterion shares common features with the RKR model for cleavage fracture [53], stating that brittle fracture occurs after a critical stress has been reached across a sufficient length (intrinsic to the material). Therefore, contrary to most studies relying on a maximum stress criterion, it is believed that not only the magnitude of the stress peaks in the brittle dielectric should be considered, but also their extent. For verification purposes, the pad architecture comparison carried out in Section III.6.2 was repeated using a smaller domain[1] for stress averaging in the submodel, leading to the same results.

A comparison between the present method and energetic approaches such as the ARE or the NRE methods, or investigations into the results dependence on the dimensions of the averaging domain may bring more insights into the applicability of the proposed method, for instance comparing the results obtained using a typical dimension of the interconnect pattern against an intrinsic material length evaluated from experimental testing. In addition, the proposed approach could also be applied for other process steps critical for mechanical robustness, such as electrical probing prior to wire bonding, which can lead to cracking or delamination in the interconnect stack due to impact and scrubbing by contact probes [57–63].

[1] Half the length of the domain $\mathcal{D}$ used previously, with the same width, and centered about the peak maximum tensile stress.

III.8 Conclusions

In this chapter, a combined experimental and numerical analysis of the mechanical robustness of wirebond pads in a 3D stacked BSI image sensor was conducted.

In the experimental study, six distinct pad architectures were investigated using the cratering test. The bonding surface was inspected after the wirebonding process to determine the number of failure occurrences for each pad and derive a ranking based on their relative mechanical robustness. A 3D finite element model of the wire-pull qualification test was then proposed, aiming to better understand the experimental results and to investigate in more detail the influence of interconnection layout. A multiscale approach was adopted to account for detailed geometric features of the interconnect stack, and a criterion based on volume-averaged maximum tensile stress in the dielectric material at the local scale was used to compare the mechanical robustness of the different pad configurations.

Based on the results, several guidelines on bondpad interconnect stack architecture were proposed to optimize wirebond pad mechanical robustness. Most failures were observed in cross-sections in the uppermost layers of the top chip (poly-Si layer, contact plugs, first metallization level). The introduction of additional capping layers between the Al bonding surface and the interconnect stack was found to have the most influence on mechanical robustness. By acting as buffer layer, it enables stress redistribution and contributes to attenuate the load exerted on the interconnect stack. Another important factor is interconnection layout in the pad structure. Metallization surface fraction in the interconnect stack has a significant influence, due to the load-bearing capability of metal interconnects, effectively shielding the dielectric matrix from transverse loads. In addition, metallization connectedness in the plane provides further stress relief in the dielectric material, by introducing structural reinforcement to the pad. Via plugs between metal lines also seem to play a significant role on the mechanical robustness. Additional experiments are needed to confirm the effect of via presence on the stress distribution in the different metallization levels of the pad interconnect stack. It is expected in the case of a 3D stacked BSI image sensor that a pad structure with a via layer between each pair of adjacent metallization levels could be more robust, by enabling stress redistribution from the topmost interconnect layers, close to the bonding surface, towards the thicker lowermost interconnects.

Thus, relying on a combined experimental and numerical approach a comprehensive analysis of the most influent factors on wirebond pad mechanical robustness was carried out and guidelines for backside bondpad optimization were derived, aiming to secure integration of 3D stacked BSI image sensor devices.

Several possibilities may be considered for future work:

- Carrying out further experimental testing, in particular the wire-pull and wire-shear qualification tests, aiming to supplement the cratering test data presented here and investigate additional parameters such as the influence of via plugs presence;

- Decorrelate the effect of interconnection layout from that of metallization density by working with constant Cu surface fraction
- Compare the present method with energetic approaches such as the ARE or the NRE methods;
- Further investigate the results dependence on the dimensions of the averaging domain;

Year	Company/Institution	Process	Wire/Pad	BEoL	Configuration	Type	Loading	Criterion	Experiments	Reference
1998	Texas Instruments	WB	Au/Al	Al/low-κ	?	?	?	Stress-based: Maximum tensile/shear stress	Bond shear, ball size IMC coverage	[32]
2003	IMEC	WB	Cu/Cu Au/Al	Cu/SiO_2 Cu/low-κ	2D (axisymmetric)	Static	Normal force Vertical displacement	Stress-based: – Von Mises stress in metal layer – Maximum principal stress in dielectric layer	Ball shape (SEM)	[3–5]
2003	Motorola	WB TC	Au/Al	Cu/SiO_2 Cu/low-κ	2D (axisymmetric)	Static	Normal force	Stress-based: – Maximum principal stress – Maximum compressive stress	Functional testing, MSL, TC	[33]
2003	ASE	WB	Au/Al	Cu/low-κ	2D (plane strain)	Dynamic (explicit)	Normal/Lateral displacement	Stress-based: – Von Mises stress for metal layers – Maximum compressive stress for dielectric layers	n/a	[9] [10, 11] [64]
2004	Melixis/IMEC	WB	Cu/Cu	n/a	2D (axisymmetric)	Static	Vertical displacement Cooling	Stress-based: ?	Ball shape (SEM), EBSD μRaman spectroscopy	[6]
2004	Fairchild	WB	Au/Al	Al/SiO_2	2D (plane strain)	Dynamic (implicit)	Vertical displacement Lateral displacement Cooling	Stress-based: – Maximum principal stress – Shear stress – Von Mises stress	n/a	[12, 13]
2005	STMicroelectronics Philips	Wire shear Wire pull Cooling	Au/Al	Cu/low-κ	3D	Static	Force Temperature variation	Stress-based: – Maximum principal stress – Von Mises stress	n/a	[8]
2005	Philips Eindhoven Univ.	Wire pull	Au/Al	Cu/low-κ	3D	Static	Force	Energy-based: Area release energy	n/a	[17, 18] [19, 20] [25]
2005	A*STAR ASM	WB	Au/Al	Cu/low-κ	2D (plane strain) 3D	Dynamic (explicit)	Normal force Lateral displacement	Stress-based: – Von Mises stress – Shear stress – Compressive stress	n/a	[65] [66]
2006	National Semiconductor	WB	Au/Al	Cu/SiO_2	3D	Dynamic	Normal force Lateral displacement	Stress-based: Maximum in-plane normal stress range	Ball shear test	[23]
2006	ASE	WB Wire pull	Au/Al	Cu/low-κ	2D (plane strain)	Dynamic (explicit)	Normal force Lateral displacement Vertical displacement	Damage-based: Equivalent plastic strain	n/a	[11]
2007	I-Shou Univ. ASE	WB	Au/Al	Cu/low-κ	3D	Dynamic (explicit)	Normal force Lateral displacement	Stress-based: Von Mises stress	Nanoindentation Micro-tensile test AFM	[34] [67] [15, 35, 36, 68, 69] [70–73]
2007	STMicroelectronics Freescale/NXP	WB Wire pull Wire shear	Au/Al	Cu/low-κ	3D	Static	Force	Energy-based: Nodal release energy	Wire pull test	[7, 21] [27]
2007	Eindhoven Univ. Philips	Wire pull	?	Cu/low-κ	2D (plane strain)	Static	Force	Energy-based: Cohesive zone modeling	n/a	[28]
2007	National Semiconductor Kulicke & Soffa	WB	Au/Al	Cu/SiO_2	3D	Static	Normal/lateral displacement	Stress-based: Maximum principal stress	Ball shear	[2]

Year	Company/Institution	Process	Wire/Pad	BEoL	Configuration	Type	Loading	Criterion	Experiments	Reference
2012	UTAC Globalfoundries WorleyParsons Dassault Systèmes	WB Wire pull	Cu/Al	Cu/low-κ	2D (axisymmetric)	Dynamic (explicit)	Normal force Lateral displacement	Stress-based: – Von Mises stress – Peel stress – Maximum principal stress	Ball shear Wire pull	[56, 74]
2012	Idaho State Univ. Brigham Young Univ. ON Semiconductor	WB	Au/Al	Al/SiO_2	3D	?	Lateral displacement	Stress-based: – Maximum principal stress – Von Mises stress	n/a	[75]
2012	Infineon	WB	Cu/Al	Cu/low-κ	2D (axisymmetric)	Dynamic (explicit)	Normal force Lateral displacement	Stress-based: – Maximum principal stress – Maximum shear stress – Von Mises stress	n/a	[76]
2013	Fraunhofer Institute Univ. of Erlangen-Nuremberg AMS	WB	Au/Al	Cu/SiO_2	3D	Dynamic (explicit)	Normal force Lateral displacement	Damage-based: Christensen's criterion [24]	Nanoindentation	[14, 77]
2014	Tsinghua Univ.	WB	Cu/Al	Cu/low-κ	2D	?	Vertical/lateral displacement	Stress-based: – First principal stress – Maximum shear stress	n/a	[78]
2014	Purdue Univ. Binghampton Univ.	WB	Cu/Al	Cu/low-κ	3D	?	Vertical/lateral displacement	Damage-based: Cohesive damage law	n/a	[79, 80]
2015	Fraunhofer Institute Globalfoundries	WB	Cu/Al	Cu/low-κ	3D	?	Vertical/lateral displacement	Damage-based: Cohesive damage law	n/a	[29, 30] [31]
2018	STMicroelectronics Univ. Grenoble-Alpes	Wire pull	Au/Al	Cu/SiO_2	3D	Static	Force	Stress-based: Averaged maximum principal stress	Cratering test	This work

TABLE III.7: Overview of simulation studies in the literature regarding the mechanical robustness of the interconnect stack below the bondpad.

References

[1] G. G. Harman, Wire bonding in microelectronics. McGraw-Hill, 2010.

[2] L. Zhang, V. Gumaste, A. Poddar, L. Nguyen, and G. Schulze, "Analytical and experimental characterization of bonding over active circuitry," Journal of Electronic Packaging, vol. 129, no. 4, pp. 391–399, 2007.

[3] D. Degryse, B. Vandevelde, S. Stoukatch, and E. Beyne, "Mechanical behavior of beol structures containing lowk dielectrics during bonding process," in Electronics Packaging Technology, 2003 5th Conference (EPTC 2003), pp. 815–820, IEEE, 2003.

[4] D. Degryse, B. Vandevelde, and E. Beyne, "FEM study of deformation and stresses in copper wire bonds on cu lowk structures during processing," in Electronic Components and Technology Conference, 2004. Proceedings. 54th, vol. 1, pp. 906–912, IEEE, 2004.

[5] D. Degryse, B. Vandevelde, and E. Beyne, "Mechanical FEM simulation of bonding process on cu lowk wafers," IEEE transactions on components and packaging technologies, vol. 27, no. 4, pp. 643–650, 2004.

[6] J. Chen, D. Degryse, P. Ratchev, and I. De Wolf, "Mechanical issues of Cu-to-Cu wire bonding," IEEE transactions on components and packaging technologies, vol. 27, no. 3, pp. 539–545, 2004.

[7] V. Fiori, L. T. Beng, S. Downey, S. Gallois-Garreignot, and S. Orain, "Gold wire bonding induced peeling in Cu/low-k interconnects: 3D simulation and correlations," in Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems, 2007. EuroSime 2007. International Conference on, pp. 1–9, IEEE, 2007.

[8] V. Fiori and S. Orain, "A multi scale finite element methodology to evaluate wire bond pad architectures," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems, 2005. EuroSimE 2005. Proceedings of the 6th International Conference on, pp. 648–655, IEEE, 2005.

[9] C.-L. Yeh, Y.-S. Lai, and J.-D. Wu, "Dynamic analysis of wirebonding process on cu/low-k wafers," in Electronics Packaging Technology, 2003 5th Conference (EPTC 2003), pp. 282–286, IEEE, 2003.

[10] C.-L. Yeh and Y.-S. Lai, "Transient analysis of the impact stage of wirebonding on cu/low-k wafers," Microelectronics Reliability, vol. 45, no. 2, pp. 371–378, 2005.

[11] C.-L. Yeh and Y.-S. Lai, "Comprehensive dynamic analysis of wirebonding on cu/low-k wafers," IEEE Transactions on advanced packaging, vol. 29, no. 2, pp. 264–270, 2006.

[12] Y. Liu, S. Irving, and T. Luk, "Thermosonic wire bonding process simulation and bond pad over active stress analysis," in 2004 Proceedings. 54th Electronic Components and Technology Conference (IEEE), vol. 1, pp. 383–391, June 2004.

[13] Y. Liu, S. Irving, T. Luk, M. Rioux, and Q. Qian, "Impact of wedge wire bonding and thermal mechanical stress on reliability of bpsg/poly layer of a silicon die," in 2008 58th Electronic Components and Technology Conference, pp. 1714–1718, IEEE, 2008.

[14] A. Wright, S. Koffel, S. Kraft, P. Pichler, J. Cambieri, R. Minixhofer, and E. Wachmann, "Thermo-mechanical ball bonding simulation with elasto-plastic parameters obtained from nanoindentation and atomic force measurements," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2015 16th International Conference on, pp. 1–8, IEEE, 2015.

[15] H.-C. Hsu, H.-S. Chang, S.-C. Tsao, and S.-L. Fu, "Advanced finite element model on copper wire ball bonding," in 2010 34th IEEE/CPMT International Electronic Manufacturing Technology Symposium (IEMT), pp. 1–6, IEEE, 2010.

[16] P. Lall, Y. Luo, and L. Nguyen, "De-bonding simulation of Cu-Al wire bond intermetallic compound layers," in 2016 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 862–870, IEEE, 2016.

[17] M. A. J. van Gils, O. van der Sluis, G. Q. Zhang, J. H. J. Janssen, and R. M. J. Voncken, "Novel damage model for delamination in cu/low-k ic backend structures," in Electronic Components and Technology Conference, 2005. Proceedings. 55th, pp. 988–994, IEEE, 2005.

[18] M. A. J. van Gils, O. van der Sluis, G. Q. Zhang, J. H. J. Janssen, and R. M. J. Voncken, "Analysis of cu/low-k bond pad delamination by using a novel failure index," Microelectronics Reliability, vol. 47, no. 2, pp. 179–186, 2007.

[19] O. van der Sluis, R. A. B. Engelen, W. D. van Driel, M. A. J. van Gils, and R. B. R. van Silfhout, "Efficient damage sensitivity analysis of advanced cu low-k bond pad structures using area release energy," in Thermal, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems, 2006. EuroSime 2006. 7th International Conference on, pp. 1–8, IEEE, 2006.

[20] O. van der Sluis, R. A. B. Engelen, R. B. R. van Silfhout, W. D. van Driel, and M. A. J. van Gils, "Efficient damage sensitivity analysis of advanced cu/low-k bond pad structures by means of the area release energy criterion," Microelectronics Reliability, vol. 47, no. 12, pp. 1975–1982, 2007.

[21] V. Fiori, L. T. Beng, S. Downey, S. Gallois-Garreignot, and S. Orain, "3D multi scale modeling of wire bonding induced peeling in Cu/low-k interconnects: application of an energy based criteria and correlations with experiments," in Electronic Components and Technology Conference, 2007. ECTC'07. Proceedings. 57th, pp. 256–263, IEEE, 2007.

[22] C.-L. Yeh, Y.-S. Lai, and C.-L. Kao, "Transient simulation of wire pull test on cu/low-k wafers," IEEE transactions on advanced packaging, vol. 29, no. 3, pp. 631–638, 2006.

[23] L. Shen, V. Gumaste, A. Poddar, and L. Nguyen, "Effect of pad stacks on dielectric layer failure during wire bonding," in Electronic Components and Technology Conference, 2006. Proceedings. 56th, pp. 6–pp, IEEE, 2006.

[24] R. M. Christensen, "A comprehensive theory of yielding and failure for isotropic materials," Journal of Engineering materials and Technology, vol. 129, no. 2, pp. 173–181, 2007.

[25] R. A. B. Engelen, O. van der Sluis, R. B. R. van Silfhout, W. D. van Driel, and V. Fiori, "Optimization of cu low-k bond pad designs to improve mechanical robustness using the area release energy method," in 2007 International Conference on Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems. EuroSime 2007, pp. 1–4, April 2007.

[26] O. van der Sluis, R. B. R. van Silfhout, R. A. B. Engelen, W. D. Van Driel, G. Q. Zhang, and L. J. Ernst, "A numerical method for efficient failure modelling of three-dimensional bond pad structures," in Electronic Components and Technology Conference, 2007. ECTC'07. Proceedings. 57th, pp. 235–241, IEEE, 2007.

[27] S. Gallois-Garreignot, V. Fiori, S. Orain, and O. van der Sluis, "Numerical analysis of the reliability of cu/low-k bond pad interconnections under wire pull test: application of a 3D energy based failure criterion," in Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems, 2007. EuroSime 2007. International Conference on, pp. 1–7, IEEE, 2007.

[28] B. A. E. van Hal, R. H. J. Peerlings, M. G. D. Geers, and O. van der Sluis, "Cohesive zone modeling for structural integrity analysis of ic interconnects," Microelectronics Reliability, vol. 47, no. 8, pp. 1251–1261, 2007.

[29] J. Auersperg, D. Breuer, K. V. Machani, S. Rzepka, and B. Michel, "FEA study of damage and cracking risks in BEoL structures under copper wirebonding impact," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2015 16th International Conference on, pp. 1–5, IEEE, 2015.

[30] J. Auersperg, D. Breuer, K. V. Machani, S. Rzepka, and B. Michel, "FEA to tackle damage and cracking risks in BEoL structures under copper wire bonding impact," in AMC 2015 – Advanced Metallization Conference, 2015.

[31] E. Kaulfersch, J. Auersperg, D. Breuer, B. Brämer, and S. Rzepka, "Evaluation of pattern scale stress effects of 28nm technology during wire bond and cu pillar flip chip assembly," in Microelectronics Packaging Conference (EMPC), 2015 European, pp. 1–6, IEEE, 2015.

[32] M. Saran, R. Cox, C. Martin, G. Ryan, T. Kudoh, M. Kanasugi, J. Hortaleza, M. Ibnabdeljalil, M. Murtuza, D. Capistrano, et al., "Elimination of bond-pad damage through structural reinforcement of intermetal dielectrics," in Reliability Physics Symposium Proceedings, 1998. 36th Annual. 1998 IEEE International, pp. 225–231, IEEE, 1998.

[33] K. Hess, S. Downey, G. Hall, L. Mercado, J. Miller, W. Ng, and D. Wontor, "Reliability of bond over active pad structures for 0.13-um cmos technology," in Electronic Components and Technology Conference, pp. 1344–1349, IEEE; 1999, 2003.

[34] P.-C. Chin, C.-Y. Hu, H.-C. Hsu, S.-L. Fu, C.-L. Yeh, and Y.-S. Lai, "Characteristic of heat affected zone in thin gold wire and dynamic transient analysis of wire bonding for microstructure of cu/low-k wafer," in Microsystems, Packaging, Assembly and Circuits Technology, 2007. IMPACT 2007. International, pp. 297–300, IEEE, 2007.

[35] H.-C. Hsu, W.-Y. Chang, S.-L. Fu, C.-L. Yeh, and Y.-S. Lai, "Dynamic finite element analysis on underlay microstructure of Cu/low-K wafer during bonding process," in 2007 International Conference on Electronic Materials and Packaging, pp. 1–6, IEEE, 2007.

[36] H. C. Hsu, P. C. Chin, C. Y. Hu, W. Y. Chang, C. L. Yeh, and Y. S. Lai, "Dynamic analysis on underlay microstructure for Cu/low-k wafer during wire bonding," in Key Engineering Materials, vol. 419, pp. 489–492, Trans Tech Publ, 2010.

[37] H.-C. Hsu, W.-Y. Chang, C.-L. Yeh, and Y.-S. Lai, "Characteristic of copper wire and transient analysis on wirebonding process," Microelectronics Reliability, vol. 51, no. 1, pp. 179–186, 2011.

[38] Y. Ding, J.-K. Kim, and P. Tong, "Numerical analysis of ultrasonic wire bonding: Effects of bonding parameters on contact pressure and frictional energy," Mechanics of materials, vol. 38, no. 1, pp. 11–24, 2006.

[39] M. Mayer and A. Zwart, "Ultrasonic friction power in microelectronic wire bonding," in Materials science forum, vol. 539, pp. 3920–3925, Trans Tech Publ, 2007.

[40] I. Jeon, Q. Chung, J. Hong, and K. Byun, "The effect of ultrasonic power on bonding pad and imd layers in ultrasonic wire bonding," in Electronic Materials and Packaging, 2001. EMAP 2001. Advances in, pp. 235–242, IEEE, 2001.

[41] I. Jeon, "The study on failure mechanisms of bond pad metal peeling: Part b—-numerical analysis," Microelectronics reliability, vol. 43, no. 12, pp. 2055–2064, 2003.

[42] Y. Huang, A. Shah, M. Mayer, N. Zhou, and J. Persic, "Effect of ultrasonic capillary dynamics on the mechanics of thermosonic ball bonding," IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol. 57, no. 1, 2010.

[43] C. C. Meng and E. Y. W. Yik, "Forced-time controlled transient dynamic wire bonding simulation for wire bond process development," in Electronics Packaging and Technology Conference (EPTC), 2015 IEEE 17th, pp. 1–4, IEEE, 2015.

[44] E. Spaan, E. Ooms, W. D. Van Driel, C. A. Yuan, D. G. Yang, and G. Q. Zhang, "Wire bonding the future: a combined experimental and numerical approach to improve the cu-wire bonding quality," in Thermal, Mechanical & Multi-Physics Simulation, and Experiments in Microelectronics and Microsystems (EuroSimE), 2010 11th International Conference on, pp. 1–4, IEEE, 2010.

[45] J. C. Fish and S. W. Lee, "Delamination of tapered composite structures," Engineering Fracture Mechanics, vol. 34, no. 1, pp. 43–54, 1989.

[46] A. J. Vizzini, "Influence of realistic ply-drop geometries on interlaminar stresses in tapered laminates," Composite Materials: Fatigue and Fracture, vol. 1230, p. 467, 1995.

[47] W. Cui, M. R. Wisnom, and M. Jones, "New model to predict static strength of tapered laminates," Composites, vol. 26, no. 2, pp. 141–146, 1995.

[48] R. Iannuzzelli, "Predicting solder joint reliability, model validation," in Electronic Components and Technology Conference, 1993. Proceedings., 43rd, pp. 839–851, IEEE, 1993.

[49] S. Yi, L. Shen, J. K. Kim, and C. Y. Yue, "A failure criterion for debonding between encapsulants and leadframes in plastic ic packages," Journal of Adhesion Science and Technology, vol. 14, no. 1, pp. 93–105, 2000.

[50] H. Neuber, "Kerbspannungslehre," Berlin/Heidelberg, Gottingen, 1958. In [52].

[51] D. Radaj, "Generalised neuber concept of fictitious notch rounding," in Advanced Methods of Fatigue Assessment, pp. 1–100, Springer, 2013.

[52] F. Berto, "A brief review of some local approaches for the failure assessment of brittle and quasi-brittle materials," Advances in Materials Science and Engineering, vol. 2014, 2014.

[53] R. O. Ritchie, J. F. Knott, and J. Rice, "On the relationship between critical tensile stress and fracture toughness in mild steel," Journal of the Mechanics and Physics of Solids, vol. 21, no. 6, pp. 395–410, 1973.

[54] T. L. Anderson, Fracture mechanics: fundamentals and applications. CRC press, 2005.

[55] M. Song, G. Gong, J. Yao, S. Xu, S. Lee, M. Han, and B. Yan, "Study of optimum bond pad metallization thickness for copper wire bond process," in 2010 12th Electronics Packaging Technology Conference, pp. 597–602, IEEE, 2010.

[56] J. D. Beleran, Y. Y. Bo, H. G. Robles, A. Milanes, A. Yeo, and C. K. Chong, "Copper wire bond analysis: Pad design effects and process considerations," in Electronic Components and Technology Conference (ECTC), 2012 IEEE 62nd, pp. 1124–1129, IEEE, 2012.

[57] Y. Liu, D. Desbiens, S. Irving, T. Luk, S. Edborg, D. Hahn, and S. Park, "Probe test failure analysis of bond pad over active structure by modeling and experiment," in Electronic Components and Technology Conference, 2005. Proceedings. 55th, pp. 861–866, IEEE, 2005.

[58] Y. Liu, D. Desbiens, T. Luk, and S. Irving, "Parameter optimization for wafer probe using simulation," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems, 2005. EuroSimE 2005. Proceedings of the 6th International Conference on, pp. 156–161, IEEE, 2005.

[59] R. Roucou, V. Fiori, F. Cacho, K. Inal, and X. Boddaert, "Evaluation of probing process parameters and pad designs: experiments and modelling correlations for solving mechanical issues," in Thermal, Mechanical & Multi-Physics Simulation, and Experiments in Microelectronics and Microsystems (EuroSimE), 2010 11th International Conference on, pp. 1–6, IEEE, 2010.

[60] J. Albrecht, G. M. Reuther, J. Brueckner, J. Auersperg, S. Rzepka, and R. Pufall, "Risk assessment of bond pad stacks: Combined utilization of nanoindentation and FE-modeling," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2016 17th International Conference on, pp. 1–4, IEEE, 2016. Fraunhofer ENAS, Infineon Technologies AG.

[61] J. Albrecht, J. Auersperg, G. M. Reuther, P. W. Kudella, J. Brueckner, S. Rzepka, and R. Pufall, "Brittle fracture and damage in bond pad stacks—a study of parameter influences in coupled xfem and delamination simulation of nanoindentation," in Electronics Packaging Technology Conference (EPTC), 2016 IEEE 18th, pp. 724–728, IEEE, 2016.

[62] G. M. Reuther, P. W. Kudella, J. Albrecht, J. Brueckner, J. Auersperg, S. Rzepka, and R. Pufall, "Brittle fracture and damage in bond pad stacks: Novel approaches for simulation-based risk assessment," in Electronic System-Integration Technology Conference (ESTC), 2016 6th, pp. 1–6, IEEE, 2016.

[63] G. M. Reuther, I. Penjovic, A. O. Brezmes, and R. Pufall, "Analytical and simulation-based risk assessment of imprint depth and brittle fracture in bond pad stacks," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2017 18th International Conference on, pp. 1–6, IEEE, 2017.

[64] C.-L. Yeh, Y.-C. Lee, and Y.-S. Lai, "Vibration and bondability analysis of fine-pitch Cu wire bonding," in 2011 12th International Conference on Electronic Packaging Technology and High Density Packaging, pp. 1–7, IEEE, 2011.

[65] A. G. Viswanath, W. Fang, X. Zhang, V. Ganesh, and L. Lim, "Numerical analysis by 3d finite element wire bond simulation on Cu/low-k structures," in 2005 7th Electronic Packaging Technology Conference, vol. 1, pp. 6–pp, IEEE, 2005.

[66] A. G. K. Viswanath, X. Zhang, V. P. Ganesh, and L. Chun, "Numerical study of gold wire bonding process on cu/low-k structures," IEEE Transactions on Advanced Packaging, vol. 30, no. 3, pp. 448–456, 2007.

[67] W.-Y. Chang, H.-C. Hsu, S.-L. Fu, C.-L. Yeh, and Y.-S. Lai, "Characteristic of heat affected zone for ultra thin gold wire/copper wire and advanced finite element wirebonding model," in 2008 10th Electronics Packaging Technology Conference, pp. 419–423, IEEE, 2008.

[68] H. Hsu, C. Hu, W. Chang, C. Yeh, and Y. Lai, Dynamic Finite Element Analysis on Underlay Microstructure of Cu/low-k Wafer during Wirebonding. InTech, 2010.

[69] H.-C. Hsu, W.-Y. Chang, C.-L. Yeh, and Y.-S. Lai, "Characteristic of copper wire and transient analysis on wirebonding process," Microelectronics Reliability, vol. 51, no. 1, pp. 179–186, 2011.

[70] H.-C. Hsu, J.-H. Chien, L.-M. Chu, S.-P. Ju, Y.-T. Feng, and S.-L. Fu, "Nanoscale bondability study on copper-aluminum intermetallic compound using molecular dynamics simulation," in 2012 7th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), pp. 64–67, IEEE, 2012.

[71] H.-C. Hsu, G. Lee, W.-R. Feng, and C.-H. Kuo, "An Investigation on Mechanical Behavior on Pt-Coated Copper Wire from a Nanoscale Perspective," ECS Transactions, vol. 44, no. 1, pp. 907–919, 2012.

[72] H.-C. Hsu, L.-M. Chu, W.-Y. Chang, Y.-F. Chen, J.-H. Chien, S.-L. Fu, S.-P. Ju, and W.-J. Feng, "Further characterization of Pd-coated copper wire on wirebonding process: from a nanoscale perspective," Journal of Materials Science: Materials in Electronics, vol. 24, no. 9, pp. 3594–3602, 2013.

[73] H. C. Hsu, J. H. Chien, J. S. Huang, L. M. Chu, and S. L. Fu, "Nanoscale bondability between Cu-Al intermetallic compound for Cu wirebonding," in 2014 International Conference on Electronics Packaging (ICEP), pp. 618–622, Apr. 2014.

[74] Y. B. Yang, N. Kumar, D. John, R. Hyman, F. Clive, S. Nathapong, and W. Ramroth, "A dynamic study of pad structure impact on bond pad/low-k layer stress in copper wire bond," in Electronics Packaging Technology Conference (EPTC), 2012 IEEE 14th, pp. 654–658, IEEE, 2012.

[75] S. G. Hunter, R. Gohnert, A. J. Dutson, and D. S. Naidu, "Finite element simulations of stresses in cup bond pads of al-si0 2 interconnect," in Electronics Packaging Technology Conference (EPTC), 2012 IEEE 14th, pp. 125–130, IEEE, 2012.

[76] A. Yeo and F. X. Che, "Numerical & experimental analysis of bond pad stack structure for wire bond interconnection," in 2012 35th IEEE/CPMT International Electronics Manufacturing Technology Conference (IEMT), pp. 1–5, Nov 2012.

[77] A. Wright, S. Koffel, P. Pichler, H. Enichlmair, R. Minixhofer, and E. Wachmann, "On the thermo-mechanical modelling of a ball bonding process with ultrasonic softening," in 2013 14th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), pp. 1–8, IEEE, 2013.

[78] X. Bai, B. Chang, and D. Du, "Finite element analysis of influence of passivation layer on cu/low-k structure during thermosonic cu wire bonding," in IOP Conference Series: Materials Science and Engineering, vol. 61, p. 012037, IOP Publishing, 2014.

[79] K. Upreti, H.-Y. Lin, G. Subbarayan, D. Y. Jung, and B. Sammakia, "Simulations of damage and fracture in ulk under pad structures during cu wirebond process," in Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2014 IEEE Intersociety Conference on, pp. 609–615, IEEE, 2014.

[80] K. Upreti, H.-Y. Lin, G. Subbarayan, D. Y. Jung, and B. G. Sammakia, "An assessment of risk of fracture during wirebond over active circuits on ULK dies," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 6, no. 2, pp. 314–325, 2016.

Chapter IV

Process-induced Thermomechanical Stresses in a 3D BSI Image Sensor Stacked by Hybrid Bonding

IV.1 Introduction

In the context of semiconductor manufacturing, with wafer fabrication cycle times on the order of several months, inline monitoring of the different quantities of interest, e.g. mask alignment, layer thickness, surface topography, or device resistance values is mandatory. Such measurements are carried out many times during fabrication, and enable

during product development: to identify processing issues and the associated critical steps without needing the wafers to undergo the whole fabrication sequence;

during mass production: to identify process errors and deviations using statistical process control techniques.

Many issues encountered during integrated circuit processing, such as substrate warping, delamination/cracking in the interconnect stack, or even semiconductor device electrical performance are governed by the stress distribution in the assembly. Having a method able to evaluate stress build-up during the sequence of processing steps is a way to prevent these problems, as well as to correlate and validate the numerical models developed for material or structure optimization.

Several techniques have been developed to measure mechanical stresses in thin films, such as X-ray diffraction [1] or Raman spectroscopy [2]. However, these methods are quite time-consuming in an industrial context and for some applications require extensive (and destructive) sample preparation. A more widespread and cost-effective technique is stress evaluation through wafer curvature measurements, which are routinely carried out during wafer processing and can be

used to determine thin film stress using Stoney's equation [3]. However, this method has strong limitations for complex patterned stacks, since only the average inplane stress is derived.

In an image sensor, the electro-optical performance of the different semiconductor components inside the active pixel sensors (e.g. photodiodes, transistors) is known to be negatively affected by mechanical stress [4]. Aiming to determine whether a significant increase of mechanical stress could be induced in the active region of the image sensor due to the 3D BSI configuration, in this chapter a method based on Si piezoresistive stress sensors is proposed for inline monitoring of the *in situ* stress distribution during the fabrication sequence. Stress measurements are carried out both at the wafer-level and the package-level, and compared with finite element modeling of the sequence of process steps. The limitations of this method will be discussed and directions for further work proposed.

IV.2 *In situ* stress evaluation using Si piezoresistive stress sensors

In this section, after a short overview of the different sensor technologies for *in situ* stress monitoring in integrated circuits, the theoretical framework and general methodology are outlined for Si piezoresistive stress sensors.

IV.2.1 Overview of stress sensors technologies for integrated circuits

Resistor-based strain gauges have long been used in mechanical and civil engineering to derive mechanical stresses in structural components. This is achieved by measuring resistance variations under mechanical loading, after prior calibration to determine the gauge factors, enabling to access the longitudinal, transverse and normal strain components. Using an array of resistors with different orientations (rosette configuration), the 2D strain state can be reconstructed in the location of interest at the sample surface.

This strategy is easily transferrable to microelectronic applications by integrating metal thin film resistors in the interconnect stack of the integrated circuit of interest, in the form of metal serpentines arranged in a rosette configuration. A common calibration method for this strain gauge is to exert a controlled loading using the four-point bending technique on rectangular samples sawed from the Si wafer, while measuring resistance variations in the metal serpentines to determine the gauge factors. Specific design considerations are required to maximize strain gauge sensitivity, while minimizing self-heating[1] of the metal resistors to avoid disturbing the resistance measurement [5]. Contrary to the strain gauges used in structural applications, which are attached at the surface of the stressed body and thus only give access to a 2D stress field,

[1] Due to the Joule effect under the prescribed electrical current during resistance measurements.

in integrated circuits the structure is embedded into the interconnect stack and may therefore be subjected to transverse stresses.

In the general case, strain-induced resistance changes originate from two main sources [6]. This is apparent from writing the resistance of a conductor of length L, rectangular cross-section $W \times t$ and resistivity ρ:

$$R = \frac{\rho L}{W t} \tag{IV.1}$$

Under an incremental length change δL, there is a variation in the transverse dimensions of the conductor δW and δt due to the Poisson effect, but also a resistivity change $\delta \rho$ due to the piezoresistivity effect (as will be detailed in Section IV.2.2). The resulting incremental relative resistance change writes:

$$\frac{\delta R}{R} = \frac{\delta \rho}{\rho} + \frac{\delta L}{L} - \frac{\delta W}{W} - \frac{\delta t}{t} = \frac{\delta \rho}{\rho} + (1 + 2\nu)\frac{\delta L}{L} \tag{IV.2}$$

expressing the transverse strains $\delta W/W$ and $\delta t/t$ as a function of the longitudinal strain $\delta L/L$, through Poisson's ratio ν. With the metal serpentines approach described above, resistance change during a mechanical loading is mostly due to geometrical variations of the deformed resistor (right-hand side term in Equation (IV.2)), whereas the resistivity change is negligible (left-hand side term). This type of sensors typically has poor sensitivity for application to microelectronic devices: a low relative resistance variation is obtained for stress magnitudes typically encountered (on the order of 0.1% resistance variation for 100 MPa). Since these metal serpentines also have very low resistance, for instance compared to semiconductor resistors, accurate stress evaluation using these sensors is challenging [5].

Another approach is to fabricate sensors using doped Si resistors instead of metal serpentines. In these structures, resistance variations under mechanical loading originate from the resistivity change due to the piezoresitive effect, rather than resistor geometry variation[1]. Such piezoresistive sensors have been introduced for practical applications in integrated circuits by researchers at Texas Instruments in the early 1980s [7, 8]. The increase in gauge factors obtained using piezoresistors instead of metal serpentines can be as large as two orders of magnitude, depending on dopant concentration [9].

A fundamental difference between these two types of sensors is also their integration within the die: metal thin film resistors are processed at the BEoL level and thus give access to the stress components inside the interconnect stack, whereas doped Si resistors are fabricated at the FEoL level, enabling to obtain the stress field in the active region. For the 3D BSI image sensor considered in this work, an important question is whether the new integration scheme by chip stacking with hybrid bonding may lead to a significant change in the magnitude or distribution of the stresses exerted on the active components in the pixels, in turn resulting in a degradation of image sensor electro-optical performance. Piezoresistive sensors, located in the active region

[1] Conversely, for metal thin film resistors the resistance change under mechanical loading due to the piezoresistive effect is negligible compared to that induced by resistor geometry variation.

of the chip, are a good candidate for this purpose. Thus, an approach based on doped Si stress sensors[1] is proposed, aiming to enable *in situ* stress evaluation in the active region during the fabrication process of a 3D test chip stacked by hybrid bonding.

IV.2.2 Theoretical background

In this section, the governing equations of silicon piezoresistivity are detailed. These describe the relationship between the normalized change in resistance and local stress field, and will be used in Section IV.2.4 both for the calibration of our stress sensors and the evaluation of the stress variations between two process steps[2]. The notations and developments of [10] are directly followed.

IV.2.2.1 Piezoresistive effect

This piezoresistive effect has been observed in both metals (e.g. Bi, Ni or Co) [11–13] and semiconductors (Si and Ge) [9, 14, 15]. It corresponds to a variation of the components of the resistivity tensor under applied stress. Neglecting second-order piezoresistivity, the resistivity variation writes:

$$\rho_{ij} = \rho_{ij}^0 + \pi_{ijkl}\sigma_{kl} \tag{IV.3}$$

where ρ_{ij}^0 and π_{ijkl} respectively correspond to the Cartesian components of the stress-free resistivity tensor and the piezoresistivity tensor.

Because the resistivity and the stress tensors are both symmetric, this relationship can be considerably simplified, from nine to six equations:

$$\begin{bmatrix} \rho_{11} \\ \rho_{22} \\ \rho_{33} \\ \rho_{23} \\ \rho_{13} \\ \rho_{12} \end{bmatrix} = \begin{bmatrix} \rho_{11}^0 \\ \rho_{22}^0 \\ \rho_{33}^0 \\ \rho_{23}^0 \\ \rho_{13}^0 \\ \rho_{12}^0 \end{bmatrix} + \begin{bmatrix} \pi_{1111} & \pi_{1122} & \pi_{1133} & 2\pi_{1123} & 2\pi_{1113} & 2\pi_{1112} \\ \pi_{2211} & \pi_{2222} & \pi_{2233} & 2\pi_{2223} & 2\pi_{2213} & 2\pi_{2212} \\ \pi_{3311} & \pi_{3322} & \pi_{3333} & 2\pi_{3323} & 2\pi_{3313} & 2\pi_{3312} \\ \pi_{2311} & \pi_{2322} & \pi_{2333} & 2\pi_{2323} & 2\pi_{2313} & 2\pi_{2312} \\ \pi_{1311} & \pi_{1322} & \pi_{1333} & 2\pi_{1323} & 2\pi_{1313} & 2\pi_{1312} \\ \pi_{1211} & \pi_{1222} & \pi_{1233} & 2\pi_{1223} & 2\pi_{1213} & 2\pi_{1212} \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} \tag{IV.4}$$

[1] Metal serpentines rosettes are fundamentally strain gauges, since the measured resistance variations under mechanical loading originate from resistor deformation. If the material in which these sensors are embedded behaves elastically for the load magnitude considered, and the elastic constants of this material are known, then the stress state can be derived from the measured strain components. Since that is the case for the piezoresistive sensors embedded in single-crystal Si, the measured response to mechanical loading can be indifferently expressed in terms of strain or stress, and thus these structures may be equivalently referred to as *strain gauges* or *stress sensors*. Following the literature convention of expressing the governing equation of piezoresistivity in terms of stress, Si piezoresistive sensors will be referred to in the following as *stress sensors*.

[2] Only the *variation* of the different stress components between two process steps is obtainable from the knowledge of the resistance change, not the *absolute* magnitude of stress (contrary to other techniques such as X-ray diffraction or Raman spectroscopy). For simplicity, in the remainder of this chapter we will sometimes simply use the term 'stress', when actually referring to a *variation* of stress between two process steps.

Introducing the Voigt notation (i.e., 11 → 1, 22 → 2, 33 → 3, 23 → 4, 13 → 5, 12 → 6), the above equation becomes:

$$\rho_\alpha = \rho_\alpha^0 + \Pi_{\alpha\beta}\sigma_\beta \tag{IV.5}$$

with $\alpha, \beta = 1, 2, \ldots 6$.

Rewriting this expression in terms of relative resistivity change gives:

$$\frac{\Delta\rho_\alpha}{\bar{\rho}} = \pi_{\alpha\beta}\sigma_\beta \tag{IV.6}$$

where $\Delta\rho_\alpha = \rho_\alpha - \rho_\alpha^0$ and the $\pi_{\alpha\beta} = \frac{\Pi_{\alpha\beta}}{\bar{\rho}}$ are the *piezoresistive coefficients*, with $\bar{\rho} = \frac{\rho_1^0+\rho_2^0+\rho_3^0}{3}$ the *mean unstressed resistivity*.

For silicon, which is a cubic crystal with diamond structure, the number of independent piezoresistive coefficients is reduced to three due to orthotropic symmetry, namely π_{11}, π_{22} and π_{44} [16].

IV.2.2.2 Electrical conduction in the stressed Si crystal

The general conduction equations for an anisotropic ohmic material write:

$$E_i = \rho_{ij}J_j \tag{IV.7}$$

with E_i and J_i the Cartesian components of the electric field and current density vectors.

Substitution of Equation (IV.3) into the above expression gives:

$$E_i = \rho_{ij}^0 J_j + \pi_{ijkl}\sigma_{kl}J_j \tag{IV.8}$$

Again, introducing the Voigt notation:

$$\begin{cases} E_1 = \rho_1 J_1 + \rho_6 J_2 + \rho_5 J_3 \\ E_2 = \rho_6 J_1 + \rho_2 J_2 + \rho_4 J_3 \\ E_3 = \rho_5 J_1 + \rho_4 J_2 + \rho_3 J_3 \end{cases} \tag{IV.9}$$

The conduction equations in reduced index notation then write:

$$\begin{aligned} E_1 &= \left(\rho_1^0 + \bar{\rho}\pi_{1\alpha}\sigma_\alpha\right) J_1 + \left(\rho_6^0 + \bar{\rho}\pi_{6\alpha}\sigma_\alpha\right) J_2 + \left(\rho_5^0 + \bar{\rho}\pi_{5\alpha}\sigma_\alpha\right) J_3 \\ E_2 &= \left(\rho_6^0 + \bar{\rho}\pi_{6\alpha}\sigma_\alpha\right) J_1 + \left(\rho_2^0 + \bar{\rho}\pi_{2\alpha}\sigma_\alpha\right) J_2 + \left(\rho_4^0 + \bar{\rho}\pi_{4\alpha}\sigma_\alpha\right) J_3 \\ E_3 &= \left(\rho_5^0 + \bar{\rho}\pi_{5\alpha}\sigma_\alpha\right) J_1 + \left(\rho_4^0 + \bar{\rho}\pi_{4\alpha}\sigma_\alpha\right) J_2 + \left(\rho_3^0 + \bar{\rho}\pi_{3\alpha}\sigma_\alpha\right) J_3 \end{aligned} \tag{IV.10}$$

IV.2.2.3 Stress-induced resistance change for a single Si piezoresistor in a (100) substrate

IV.2.2.3.1 Coordinate system aligned with the principal symmetry axes of the Si lattice

In a coordinate system (x_1, x_2, x_3) aligned with the principal symmetry axes of the cubic lattice, e.g. with $x_1 = [100]$, $x_2 = [010]$ and $x_3 = [001]$, the unstressed resistivity components and the piezoresistive tensor become, respectively [16]:

$$\begin{aligned} \rho_1^0 = \rho_2^0 = \rho_3^0 = \bar{\rho} \\ \rho_4^0 = \rho_5^0 = \rho_6^0 = 0 \end{aligned} \tag{IV.11}$$

and

$$[\pi_{\alpha\beta}] = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{11} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{bmatrix} \tag{IV.12}$$

by requiring that these quantities remain unchanged under a number of symmetry operations characteristic of a cubic crystal.

Substituting the above expressions into Equation (IV.10) yields for the conduction equations:

$$\begin{aligned} \frac{E_1}{\bar{\rho}} &= \left[1 + \pi_{11}\sigma_{11} + \pi_{12}\left(\sigma_{22} + \sigma_{33}\right)\right] J_1 + \pi_{44}\sigma_{12} J_2 + \pi_{44}\sigma_{13} J_3 \\ \frac{E_2}{\bar{\rho}} &= \left[1 + \pi_{11}\sigma_{22} + \pi_{12}\left(\sigma_{11} + \sigma_{33}\right)\right] J_2 + \pi_{44}\sigma_{12} J_1 + \pi_{44}\sigma_{23} J_3 \\ \frac{E_3}{\bar{\rho}} &= \left[1 + \pi_{11}\sigma_{33} + \pi_{12}\left(\sigma_{11} + \sigma_{22}\right)\right] J_3 + \pi_{44}\sigma_{13} J_1 + \pi_{44}\sigma_{23} J_2 \end{aligned} \tag{IV.13}$$

These equations can then be used to derive the stress-induced resistance change in a uniform filamentary piezoresistor of length L and cross-sectional area A (in the unstressed state) oriented along unit vector $\mathbf{n} = l\mathbf{e_1} + m\mathbf{e_2} + n\mathbf{e_3}$, with l, m and n the direction cosines with respect to axes x_1, x_2 and x_3. Assuming the piezoresistor is subjected to a current density $\mathbf{J} = J\mathbf{n}$, with magnitude $J = I/A$ where I is the current intensity.

Then, Equation (IV.13) becomes:

$$\begin{aligned}
\frac{E_1 A}{\bar{\rho} I} &= \left[1 + \pi_{11}\sigma_{11} + \pi_{12}\left(\sigma_{22} + \sigma_{33}\right)\right] l + \pi_{44}\sigma_{12} m + \pi_{44}\sigma_{13} n \\
\frac{E_2 A}{\bar{\rho} I} &= \left[1 + \pi_{11}\sigma_{22} + \pi_{12}\left(\sigma_{11} + \sigma_{33}\right)\right] m + \pi_{44}\sigma_{12} l + \pi_{44}\sigma_{23} n \\
\frac{E_3 A}{\bar{\rho} I} &= \left[1 + \pi_{11}\sigma_{33} + \pi_{12}\left(\sigma_{11} + \sigma_{22}\right)\right] n + \pi_{44}\sigma_{13} l + \pi_{44}\sigma_{23} m
\end{aligned} \tag{IV.14}$$

Considering that the electric potential difference V along the resistor writes:

$$V = \left(E_1 l + E_2 m + E_3 n\right) L \tag{IV.15}$$

then the resistance of the stressed conductor is given by:

$$\begin{aligned}
R^{\sigma} = \frac{V}{I} = \frac{\bar{\rho} L}{A} [1 &+ \left(\pi_{11}\sigma_{11} + \pi_{12}\left[\sigma_{22} + \sigma_{33}\right]\right) l^2 + \left(\pi_{11}\sigma_{22} + \pi_{12}\left[\sigma_{11} + \sigma_{33}\right]\right) m^2 \\
&+ \left(\pi_{11}\sigma_{33} + \pi_{12}\left[\sigma_{11} + \sigma_{22}\right]\right) n^2 + 2\pi_{44}\left(\sigma_{12} lm + \sigma_{13} ln + \sigma_{23} mn\right)]
\end{aligned} \tag{IV.16}$$

Therefore, neglecting dimensional changes, the relative resistance change under mechanical loading writes:

$$\begin{aligned}
\frac{\Delta R}{R} = \frac{R^{\sigma} - R^0}{R^0} &= \left[\pi_{11}\sigma_{11} + \pi_{12}\left(\sigma_{22} + \sigma_{33}\right)\right] l^2 + \left[\pi_{11}\sigma_{22} + \pi_{12}\left(\sigma_{11} + \sigma_{33}\right)\right] m^2 \\
&+ \left[\pi_{11}\sigma_{33} + \pi_{12}\left(\sigma_{11} + \sigma_{22}\right)\right] n^2 + 2\pi_{44}\left[\sigma_{12} lm + \sigma_{13} ln + \sigma_{23} mn\right]
\end{aligned} \tag{IV.17}$$

with $R^0 = \frac{\bar{\rho} L}{A}$ the unstressed resistance.

In the plane of a (100) Si wafer, as used for the test chip investigated in this work, i.e. setting $n = 0$, Equation (IV.17) becomes:

$$\begin{aligned}
\frac{\Delta R}{R} &= \left[\pi_{11}\sigma_{11} + \pi_{12}\left(\sigma_{22} + \sigma_{33}\right)\right] l^2 + \left[\pi_{11}\sigma_{22} + \pi_{12}\left(\sigma_{11} + \sigma_{33}\right)\right] m^2 \\
&+ 2\pi_{44}\left[\sigma_{12} lm\right]
\end{aligned} \tag{IV.18}$$

with $l = \cos\theta$ and $m = \sin\theta$, where θ is the angle between the resistor orientation and the x_1 axis (Figure IV.1).

IV.2.2.3.2 Off-axis coordinate system

So far all equations have been expressed in coordinate system (x_1, x_2, x_3), aligned with the principal symmetry axes of the crystal, i.e. $x_1 = [100]$, $x_2 = [010]$ and $x_3 = [001]$. However, as will be seen for sensor calibration in Section IV.2.4.2, in some cases it is more convenient to work in an off-axis coordinate system (x'_1, x'_2, x'_3).

Voigt's notation was adopted for simplicity to express the piezoresistivity equations (IV.6) and the conduction equations (IV.9). A consequence is that standard rotation operators cannot be used directly to derive the equivalent expressions in the off-axis coordinate system (x'_1, x'_2, x'_3), as would be the case for the fully expanded equations. The specific transformation relations in Voigt's notation can be obtained by introducing the reduced indices back into the expressions obtained in the off-axis coordinate system in fully expanded form.

Thus, the piezoresistivity equations in reduced index notation:

$$\rho'_\alpha = \rho_\alpha^{0\,\prime} + \pi'_{\alpha\beta}\sigma'_\beta \tag{IV.19}$$

remain valid in the off-axis coordinate system by taking:

$$\begin{aligned} \rho'_\alpha &= T_{\alpha\beta}\rho_\beta \\ \sigma'_\alpha &= T_{\alpha\beta}\sigma_\beta \\ \pi'_{\alpha\beta} &= T_{\alpha\gamma}\pi_{\gamma\delta}T^{-1}_{\delta\beta} \end{aligned} \tag{IV.20}$$

where the transformation operator $[T_{\alpha\beta}]$ writes:

$$[T_{\alpha\beta}] = \begin{bmatrix} l_1^2 & m_1^2 & n_1^2 & 2l_1n_1 & 2m_1n_1 & 2l_1m_1 \\ l_2^2 & m_2^2 & n_2^2 & 2l_2n_2 & 2m_2n_2 & 2l_2m_2 \\ l_3^2 & m_3^2 & n_3^2 & 2l_3n_3 & 2m_3n_3 & 2l_3m_3 \\ l_1l_3 & m_1m_3 & n_1n_3 & l_1n_3 + l_3n_1 & m_1n_3 + m_3n_1 & l_1m_3 + l_3m_1 \\ l_2l_3 & m_2m_3 & n_2n_3 & l_2n_3 + l_3n_2 & m_2n_3 + m_3n_2 & l_2m_3 + l_3m_2 \\ l_1l_2 & m_1m_2 & n_1n_2 & l_1n_2 + l_2n_1 & m_1n_2 + m_2n_1 & l_1m_2 + l_2m_1 \end{bmatrix} \tag{IV.21}$$

with l_i, m_i, n_i the direction cosines for the two coordinate systems (x_1, x_2, x_3) and (x'_1, x'_2, x'_3), defined by:

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{IV.22}$$

with $a_{ij} = \cos(x'_i, x_j)$.

Applying the obtained transformation relations to Equation (IV.10) yields:

$$\begin{aligned} E'_1 &= \left(\rho_1^{0\prime} + \bar{\rho}\pi'_{1\alpha}\sigma'_\alpha\right) J'_1 + \left(\rho_6^{0\prime} + \bar{\rho}\pi'_{6\alpha}\sigma'_\alpha\right) J'_2 + \left(\rho_4^{0\prime} + \bar{\rho}\pi'_{4\alpha}\sigma'_\alpha\right) J'_3 \\ E'_2 &= \left(\rho_6^{0\prime} + \bar{\rho}\pi'_{6\alpha}\sigma'_\alpha\right) J'_1 + \left(\rho_2^{0\prime} + \bar{\rho}\pi'_{2\alpha}\sigma'_\alpha\right) J'_2 + \left(\rho_5^{0\prime} + \bar{\rho}\pi'_{5\alpha}\sigma'_\alpha\right) J'_3 \\ E'_3 &= \left(\rho_4^{0\prime} + \bar{\rho}\pi'_{4\alpha}\sigma'_\alpha\right) J'_1 + \left(\rho_5^{0\prime} + \bar{\rho}\pi'_{5\alpha}\sigma'_\alpha\right) J'_2 + \left(\rho_3^{0\prime} + \bar{\rho}\pi'_{3\alpha}\sigma'_\alpha\right) J'_3 \end{aligned} \tag{IV.23}$$

Observing that the unstressed resistivity components remain unchanged for any coordinate system:

$$\begin{aligned} \rho_1^{0\prime} &= \rho_2^{0\prime} = \rho_3^{0\prime} = \bar{\rho} \\ \rho_4^{0\prime} &= \rho_5^{0\prime} = \rho_6^{0\prime} = 0 \end{aligned} \tag{IV.24}$$

the above expression then writes:

$$\begin{aligned} \frac{E'_1}{\bar{\rho}} &= \left(1 + \pi'_{1\alpha}\sigma'_\alpha\right) J'_1 + \pi'_{6\alpha}\sigma'_\alpha J'_2 + \pi'_{4\alpha}\sigma'_\alpha J'_3 \\ \frac{E'_2}{\bar{\rho}} &= \pi'_{6\alpha}\sigma'_\alpha J'_1 + \left(1 + \pi'_{2\alpha}\sigma'_\alpha\right) J'_2 + \pi'_{5\alpha}\sigma'_\alpha J'_3 \\ \frac{E'_3}{\bar{\rho}} &= \pi'_{4\alpha}\sigma'_\alpha J'_1 + \pi'_{5\alpha}\sigma'_\alpha J'_2 + \left(1 + \pi'_{3\alpha}\sigma'_\alpha\right) J'_3 \end{aligned} \tag{IV.25}$$

Considering the same uniform filamentary piezoresistor subjected to a current density $\mathbf{J} = J\mathbf{n}'$, expressed this time in the off-axis coordinate system, i.e. $\mathbf{n} = l'\mathbf{e_1} + m'\mathbf{e_2} + n'\mathbf{e_3}$, with l', m' and n' the direction cosines with respect to axes x'_1, x'_2 and x'_3, and using the same approach as in the previous paragraph, the stress-induced relative resistance change is readily obtained:

$$\frac{\Delta R}{R} = \pi'_{1\alpha}\sigma'_\alpha l'^2 + \pi'_{2\alpha}\sigma'_\alpha m'^2 + \pi'_{3\alpha}\sigma'_\alpha n'^2 + 2\pi'_{4\alpha}\sigma'_\alpha l'n' + 2\pi'_{5\alpha}\sigma'_\alpha m'n' + 2\pi'_{6\alpha}\sigma'_\alpha l'm' \tag{IV.26}$$

in which the remaining unknown quantities are the piezoresistive coefficients in the off-axis coordinate system. They can be obtained using the transformation relation introduced in Equation (IV.20).

A convenient off-axis coordinate system to express the equations used for sensor calibration is $x'_1 = [110]$, $x'_2 = [\bar{1}10]$ and $x'_3 = [001]$, rotated by 45° with respect to the reference coordinate system aligned with the principal symmetry axes of the crystal, $x_1 = [100]$, $x_2 = [010]$ and $x_3 = [001]$. In this off-axis coordinate system, the piezoresistive coefficients are written:

$$\left[\pi'_{\alpha\beta}\right] = \begin{bmatrix} \frac{\pi_{11}+\pi_{12}+\pi_{44}}{2} & \frac{\pi_{11}+\pi_{12}-\pi_{44}}{2} & \pi_{12} & 0 & 0 & 0 \\ \frac{\pi_{11}+\pi_{12}-\pi_{44}}{2} & \frac{\pi_{11}+\pi_{12}+\pi_{44}}{2} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{11} - \pi_{12} \end{bmatrix} \tag{IV.27}$$

In the plane of a (100) wafer, i.e. setting $n' = 0$, Equation (IV.26) then becomes:

$$\begin{aligned}\frac{\Delta R}{R} = &\left[\left(\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}\right)\sigma'_{11} + \left(\frac{\pi_{11}+\pi_{12}-\pi_{44}}{2}\right)\sigma'_{22}\right] l'^2 \\ &+ \left[\left(\frac{\pi_{11}+\pi_{12}-\pi_{44}}{2}\right)\sigma'_{11} + \left(\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}\right)\sigma'_{22}\right] m'^2 \\ &+\pi_{12}\sigma'_{33} + 2\left(\pi_{11}-\pi_{12}\right)\sigma'_{12} l' m'\end{aligned} \tag{IV.28}$$

with $l' = \cos\phi$ and $m' = \sin\phi$, where ϕ is the angle between the resistor orientation and the x'_1 axis (Figure IV.1).

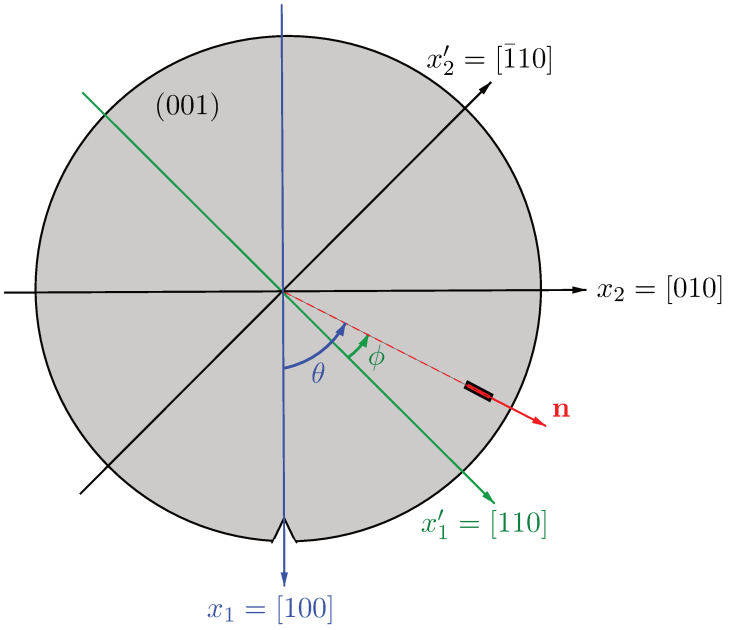


FIGURE IV.1: Unidimensional resistor arbitrarily oriented in the plane of a (100) wafer with respect to the coordinate system formed by the principal crystallographic directions of the Si lattice, along unit vector **n**.

IV.2.2.4 Rosette structures

A notable feature of Equations (IV.18) and (IV.28) is that for any orientation of the piezoresistor in the wafer plane, its resistance does not vary under applied transverse shear stress σ_{23} and σ_{31}[1]. Therefore, only four stress components at best are accessible with the Si piezoresistive stress sensors in a (100) substrate: the normal stresses σ_{11}, σ_{22}, σ_{33} and the inplane shear stress σ_{12}. It can be noted however that all six stress components can be obtained if the Si piezoresistive stress sensors are fabricated on a (111) wafer [10].

Conventional strain gauges in structural applications are attached at the surface of the sample, and are therefore in a plane stress configuration ($\sigma_{33} = 0$). Thus, using a three-element rosette

[1] Although the *resistivity* ρ_i does vary under applied transverse shear stress, as apparent for instance from Equations (IV.6) and (IV.22), the *resistance* R for the considered uniform filamentary resistor under a prescribed current does not.

(typically oriented at 0–45–90° or 0–60–120°), all three inplane stress components σ_{11}, σ_{22} and σ_{12} can be obtained. In microelectronic applications however, in the vast majority of cases the rosette is not at the surface of the sample but embedded within the chip, and therefore the plane stress assumption is not fulfilled ($\sigma_{33} \neq 0$). Since there are four unknown stress components to be determined, it seems reasonable then to introduce four resistors in the rosette (e.g. 0–45–90–135°).

Even adding a fourth resistor however, the four accessible stress components still cannot be obtained. Indeed, as pointed out by [10], writing Equation (IV.18) for four arbitrarily oriented resistors in the plane of a (100) Si wafer leads to the following system of equations, which is not invertible for any orientation[1]:

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} {l_1}^2 & {m_1}^2 & 1 & l_1 m_1 \\ {l_2}^2 & {m_2}^2 & 1 & l_2 m_2 \\ {l_3}^2 & {m_3}^2 & 1 & l_3 m_3 \\ {l_4}^2 & {m_4}^2 & 1 & l_4 m_4 \end{bmatrix} \begin{bmatrix} \pi_{11}\sigma_{11} + \pi_{12}\sigma_{22} \\ \pi_{12}\sigma_{11} + \pi_{11}\sigma_{22} \\ \pi_{12}\sigma_{33} \\ 2\pi_{44}\sigma_{12} \end{bmatrix} \tag{IV.29}$$

with $l_i = \cos\theta_i$ and $m_i = \sin\theta_i$ the direction cosines for the i-th resistor.

This is because although four equations are available for four accessible stress components, the Si crystal has only three independent piezoresistive coefficients. A solution can thus be obtained provided there is at least one resistor with distinct piezoresistive coefficients compared to the three others. Indeed, introducing two distinct sets of piezoresistive coefficients π^a_{ij} and π^b_{ij}, Equation (IV.29) becomes:

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} {l_1}^2\pi^a_{11} + {m_1}^2\pi^a_{12} & {l_1}^2\pi^a_{12} + {m_1}^2\pi^a_{11} & 1 & 2l_1 m_1 \pi^a_{44} \\ {l_2}^2\pi^a_{11} + {m_2}^2\pi^a_{12} & {l_2}^2\pi^a_{12} + {m_2}^2\pi^a_{11} & 1 & 2l_2 m_2 \pi^a_{44} \\ {l_3}^2\pi^a_{11} + {m_3}^2\pi^a_{12} & {l_3}^2\pi^a_{12} + {m_3}^2\pi^a_{11} & 1 & 2l_3 m_3 \pi^a_{44} \\ {l_4}^2\pi^b_{11} + {m_4}^2\pi^b_{12} & {l_4}^2\pi^b_{12} + {m_4}^2\pi^b_{11} & 1 & 2l_4 m_4 \pi^b_{44} \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} \tag{IV.30}$$

which is invertible.

These two distinct sets of piezoresistive coefficients can for instance be obtained by incorporating both N-type and P-type doped Si resistors into the rosette structure[2]. By doing so, all of the four accessible stress components can be derived by measuring the relative resistance changes for all resistors in the four-element rosette. Such piezoresistive sensors were designed and incorporated into the studied 3D hybrid stacked test vehicle, and are presented in the next section.

[1] Due to the direction cosines matrix being singular.

[2] The issue is the same with metal thin film rosettes, and can be solved in a similar way, by introducing two types of metal serpentines (e.g. different section or length) having distinct gauge factors

IV.2.3 Materials and methods

IV.2.3.1 Samples description

The test chip investigated is the same as in Chapter III, with lateral dimensions $9.1 \times 6.4\,\mu m^2$. It has a two-tier 3D integrated architecture, where the top and bottom chips have four and seven interconnect levels respectively and are stacked using the wafer-level Cu/SiO_2 hybrid bonding process.

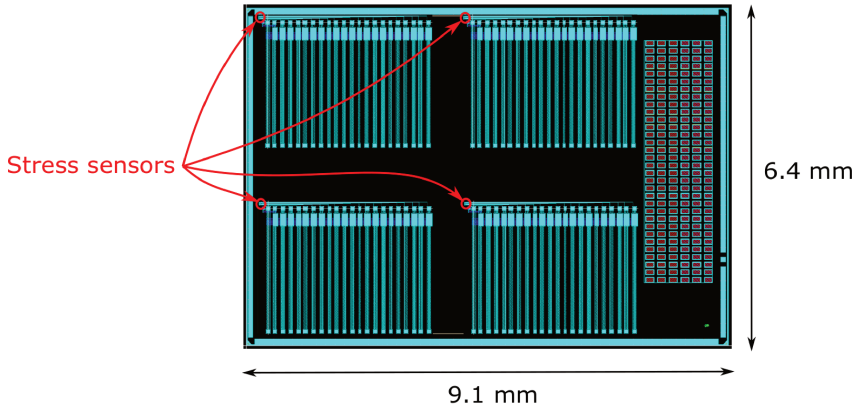


FIGURE IV.2: Design map of the test chip (the piezoresistive stress sensors are located on the upper left corner, top edge, left edge and center regions).

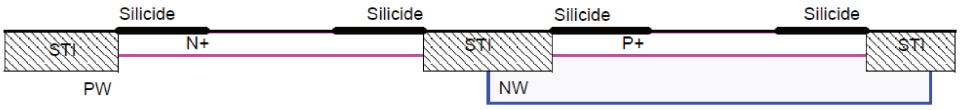


FIGURE IV.3: Schematical cross-section of N-type (left) and P-type (right) Si resistors used in the piezoresistive stress sensors.

The piezoresistive stress sensors designed for this study are integrated in the active region of top tier of the 3D IC (Figure IV.3). Each test chip includes four identical piezoresistive stress sensors, among which two are located on the edge region of the die and the others in the corner and center region (Figure IV.2).

Si piezoresistors in the rosette structure consist of a heavily doped N+ or P+ well with lateral dimensions $4 \times 0.4\,\mu m^2$ embedded in the P doped substrate or a N well (respectively), and are insulated from one another by a silicon oxide shallow trench isolation (STI) (Figure IV.3). Ion implantation is used for the doping of the active region, with impurity surface concentrations of approximately $N_S \approx 1 \times 10^{20}\,cm^{-3}$ across the first 20 nm through the thickness of the active region. The resulting sheet resistance for the P and N type resistors is respectively $245\,\Omega$/sq and $100\,\Omega$/sq.

For the studied chip, the only resistor orientations in conformity with photolithography requirements for the doping process are 0°, ±45° and 90° with respect to the [100] crystallographic direction of the Si substrate. Thus, a 0–45–90–135° rosette configuration was used. In addition,

although four resistors are sufficient to form a Si piezoresistive stress sensor, each four-element rosette was duplicated to enable the integration of two stress sensors within the same region of interest. The formed eight-element rosettes consist of four N-type resistors and four P-type resistors arranged in a rosette structure, each separated by a 45° angle and oriented along the <100> and <110> crystallographic directions of the Si substrate, with resistors of the same doping type placed adjacent to each other (Figure IV.4).

Each stress sensor extends over a surface of $15 \times 15\,\mu m^2$ in the substrate plane. For the measurements to be meaningful, the stress variation induced by the considered process step must be homogeneous across the sensor area. Thus, to avoid large stress gradients in the sensor region, specific design rules were enforced: aside from the interconnect layout required for the sensor to be accessible electrically (from the contact layer to the M2 level), no Cu patterns were allowed in the sensor region, with an exclusion area extending over 1.5 times the sensors width for the M3 and M4 metal levels. However, Cu pads were included at the hybrid bonding interface.

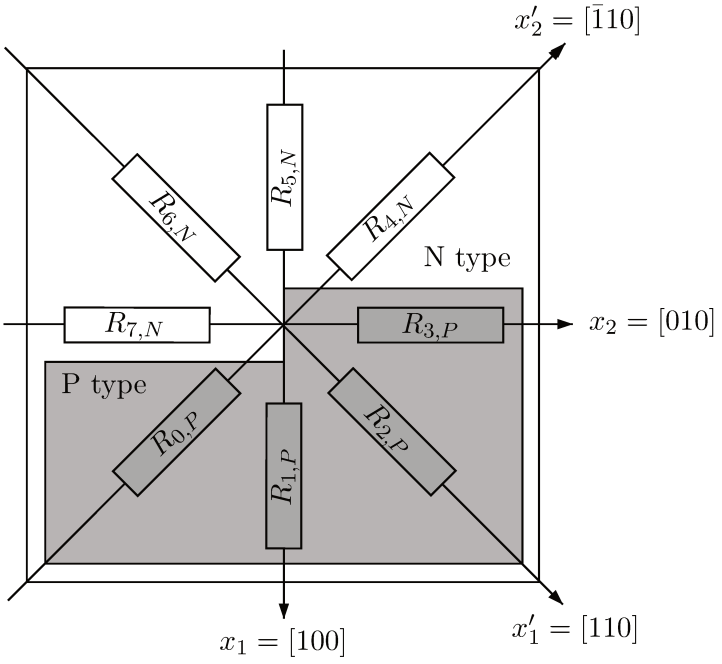


FIGURE IV.4: Eight-element rosette configuration used for the stress sensors implemented in the test chip (top die).

Four wafers from the same lot are used for this study, each including 71 identical integrated circuits. One wafer is dedicated to calibration measurements, whereas the other three are used for inline stress evaluation during the front-end process (before packaging). Among these three wafers, one is used for stress monitoring during the back-end process (during packaging). This wafer is thinned down to 120 µm from the bottom tier side and diced into individual chips, which are then assembled in an organic land-grid-array (LGA) package. During the packaging process, each die is attached to the package substrate using a PTFE-filled non-conductive thermoset die adhesive. The package substrate is a laminate comprised of a double side copper-clad epoxy resin core sandwiched between two pre-impregnated epoxy resin layers, and covered on both sides by a thin layer of solder resist coating. The electrical connection to the board is achieved

using wire bonding, as described in Chapter III. A resin cap, consisting of 30% glass-reinforced liquid crystal polymer cured at 125 °C, is then molded onto the package substrate to encapsulate the wire-bonded chip.

IV.2.3.2 Resistance measurements

IV.2.3.2.1 During wafer processing

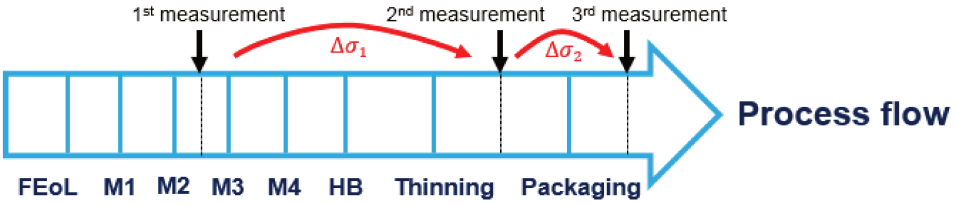


FIGURE IV.5: Time sequence of the inline resistance measurements carried during the wafer processing and chip packaging steps in the fabrication process flow of the 3D BSI image sensor.

Inline wafer-level resistance measurements are carried out as part of routine electrical testing during wafer processing (Figure IV.5):

- Initial resistance values are measured after the M2 step (M2), when the stress sensors are electrically accessible for the first time;
- During the M3 and M4 steps, as a precaution no measurements were carried out. Indeed, scrub marks left on the Cu pads after probing may lead to planarity defects for the hybrid bonding interconnects, which may in turn be detrimental for bonding quality during chip stacking by hybrid bonding.
- After hybrid bonding, wafer thinning and IO pad opening, the sensors are electrically accessible again and a second measurement is performed, corresponding to the final resistance value at the end of wafer processing (PADOPEN step).

The resistance measurements are carried out in a clean room environment at controlled temperature ($T_a = 21.0 \pm 0.5\,°C$) using a fully automated 300mm wafer prober (Accretech UF200, Tokyo Seimitsu Kogu Co. Ltd., Japan) and a parametric test station (Agilent 4070-Series, Agilent Technologies Inc., Palo Alto, CA). Four-terminal (Kelvin) sensing is used, with prescribed currents of 200 µA and 500 µA respectively for P-type and N-type resistors.

IV.2.3.2.2 During chip packaging

For the wafer undergoing the packaging steps (PKG), resistance values are measured a third time after the encapsulation process (Figure IV.5), using a device interface board for automated testing of the individual packaged integrated circuits in a parametric measurement unit (Advantest T2000, Advantest Co., Japan). Again, four-terminal sensing is used, with prescribed currents of 200 µA. A chip identification circuit is implemented in each integrated circuit so that each die can be traced by a unique digital identification tag after encapsulation, enabling proper comparison with previous measurements at the wafer-level.

IV.2.3.3 Calibration procedure

To determine the piezoresistive coefficients of the N-type and P-type resistors in the stress sensors, a known stress field must be prescribed to the sample while simultaneously measuring the resulting resistance variation, as can be seen from Equation (IV.17). Several methods are available for this calibration procedure, such as:

- microindentation [17];
- hydrostatic loading [18, 19];
- three-point bending [20];
- four-point bending [21].

Among these techniques, four-point bending is the most popular method[1], as it is relatively easy to prescribe and monitor the load, and a uniform stress field is obtained.

In this test, a rectangular thin plate sample is placed between two pairs of cylindrical pins and a vertical load is prescribed on the inner pins, placing the sample in a state of bending. With this loading configuration and specimen geometry, a uniform bending stress is prescribed at the center of the sample between the inner pins (Figure IV.3). The sign and magnitude of the uniform bending stress can be varied by adjusting the relative spacing between the inner and outer pins. Indeed, the prescribed flexural stress is given by the following equation:

$$\sigma_f = \frac{3}{2}\frac{L_o - L_i}{bh^2}F \tag{IV.31}$$

where L_o and L_i the width respectively between the outer and inner pins, b the sample width, h the sample thickness, and F the applied load.

A previously developed in-house instrumented four-point bending apparatus, presented in [31] and [30], is used for sensor calibration (Figure IV.6). The rectangular specimen is mounted on a spherical-jointed support, on which several grooves enable to adjust the width between the

[1] see for instance [22–30]

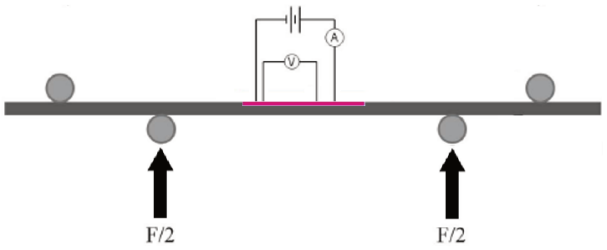


FIGURE IV.6: Schematical principle of instrumented four-point bending method for stress sensor calibration.

cylindrical pins. The other side of the four-point bending fixture is actuated by a piston, and equipped with a force transducer to measure the prescribed load. A magnetic-base stand is used to mount micromanipulators equipped with DC probe needles for the electrical measurements. The bending stress prescribed on the samples was kept reasonably low to avoid sample breakage ($\sigma_f \leq 100\,\text{MPa}$[1]) and prescribed with decreasing magnitude starting from the maximum value, so that undesirable frictional constraints on the sample due to the cylindrical pins at low loads can be minimized. For the sample dimensions considered, the uncertainty in the applied stress is estimated to about 9% [30, 31].

(a) (b)

FIGURE IV.7: Zoomed views of the in-house instrumented four-point bending apparatus used for the calibration procedure showing (a) the micromanipulators enabling precise positioning of the probe tips, and (b) a close-up view showing a sample mounted on the cylindrical pins of the four-point bending fixture and the probe needles used for parallel resistance measurements.

These experiments are carried out in a clean room environment, at controlled temperature ($T_a = 21.0 \pm 0.5\,°\text{C}$), using a semiconductor parametric analyzer (Keithley 4200, Keithley Instruments Inc., Cleveland, OH). The electrical contact with the IO pads is made manually using micromanipulators probes. The resistance values are measured by four-terminal (Kelvin) sensing to eliminate parasitic resistance due to the wires connecting the parametric analyzer to the circuit, as well as the contact resistance due to probing. Direct currents are used with values ranging between 150 µA and 800 µA. Using this setup, the uncertainty in the measured resistance values is on the order of 0.06% [32].

A wafer was diced by mechanical sawing to obtain the rectangular samples for the four-point bending test. The measured sample width after sawing is $b = 11.20 \pm 0.03\,\text{mm}$, for a thickness

[1] In preliminary testing, it was observed that for larger bending stresses, some samples broke in half below either one of the two inner loading pins.

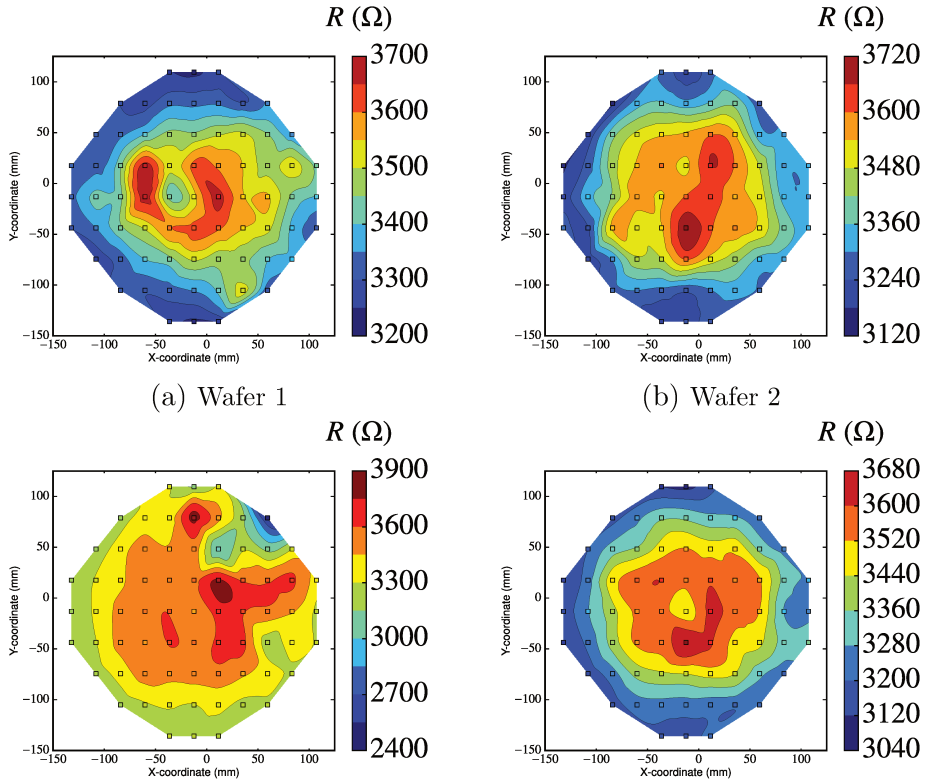
of $h = 778 \pm 20\,\mu m$. The sawing lines are drawn such that the samples have a length larger than 80 mm and that the integrated circuit of interest is located at the center of the obtained rectangular plate. Four samples are extracted from the wafer for the calibration: two oriented along the [100] direction, and the others along the [110] direction. In total, calibration measurements were carried out on twelve sensors for [100] samples, and four sensors for [110] samples. The rectangular samples used for the calibration procedure were extracted at the mid-radius of the wafer, where the measured resistance values are in the intermediate range.

IV.2.4 Results

IV.2.4.1 Initial resistance values

In this section, initial resistance values obtained from inline measurements at the wafer scale after the M2 step are presented. The aim is to:

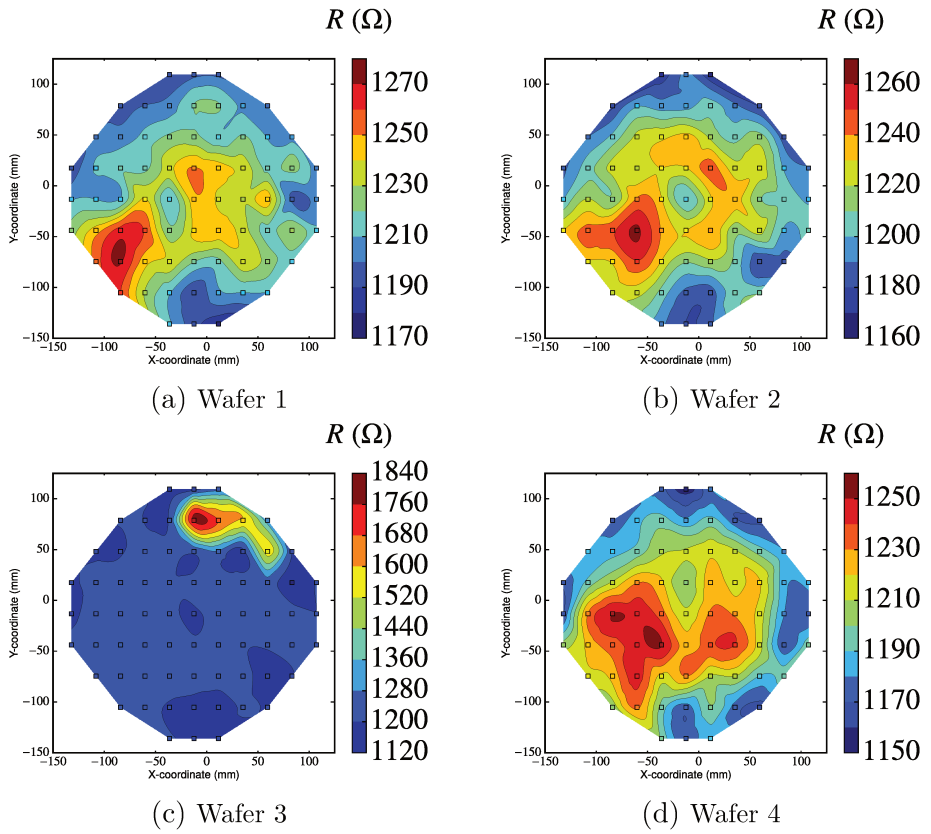
- define reference resistance values for the estimation of the stress variation during the next process steps;
- assess inter-wafer, intra-wafer and intra-sensor variability.



(a) Wafer 1 (b) Wafer 2

(c) Wafer 3 (d) Wafer 4

FIGURE IV.8: Wafer maps of averaged resistance values for P-type resistors in stress sensors located at the die center after the M2 step (outliers were discarded for wafer 3).



(a) Wafer 1 (b) Wafer 2

(c) Wafer 3 (d) Wafer 4

FIGURE IV.9: Wafer maps of averaged resistance values for N-type resistors in stress sensors located at the die center after the M2 step (outliers were discarded for wafer 3).

IV.2.4.1.1 Inter-wafer variability

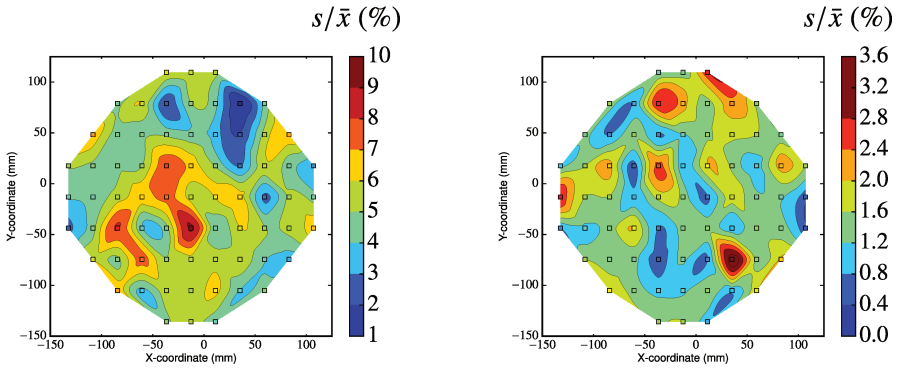
Comparing the same sites for each of the four measured wafers, the resistance variability is found to be quite low: less than 5% for the P-type resistors (Figure IV.8) and 2% for the N-type resistors (Figure IV.9). However, for a number of sites on wafer 3 the measured resistance values are far beyond the internal specifications limits[1] for the implemented resistor technology. This was observed only at the M2 step, but not during the subsequent resistance measurements. Therefore, these fluctuations are not believed to reveal a large inter-wafer variability, but rather to have been caused by poor contact between the probe needles and the IO pads during automated testing. For this wafer, all sites with such outliers at the M2 step were discarded for subsequent measurements (over a significant fraction of the wafer, mostly at the bottom left).

[1] 500 Ω and 8000 Ω.

IV.2.4.1.2 Intra-wafer variability

At the wafer-scale, resistance variability is much more significant on the other hand, with above 20% relative variation for P resistors and 9% for N resistors, with nominal values of about 3400 Ω and 1200 Ω respectively. Regarding the spatial variability of resistance values, a somewhat axisymmetric distribution can be seen for the P-type resistors with the largest values at the wafer center, whereas for the N-type resistors the distribution is more asymmetric, with maximum values found in the lower left quadrant of the wafers.

This larger variability at the wafer-scale is expected, and may be explained by a spatial heterogeneity of the manufacturing processes involved in piezoresistor fabrication, namely ion implantation and activation annealing. For instance, the implant dose may vary during mechanical scanning of the wafer in the implantation process, or the temperature distribution at the wafer surface may be inhomogeneous during heating or cooling in the annealing process (typically, with larger temperature at the wafer center) [33].



(a) P type (wafer 2) (b) N type (wafer 1)

FIGURE IV.10: Wafer maps of intra-sensor variability for P-type and N-type resistors in stress sensors located at the die center measured after the M2 step.

IV.2.4.1.3 Intra-sensor variability

At the sensor-scale, resistance variability is moderately large, with up to 10% relative variation within the rosette structure for P resistors but below 3.6% for N resistors (Figure IV.10). No specific spatial distribution is observed for the intra-sensor variability.

In this section, the reference resistance values at the M2 step have been measured, and their variability and repeatability has been quantified, both at the wafer-scale and the sensor-scale. These values will be used for the estimation of the stress variation during the next process steps. To do so, the piezoresistive coefficients π_{ij} must be evaluated for both N-type and P-type resistors. In the following, calibration results for our stress sensors are presented.

IV.2.4.2 Piezoresistive coefficients

IV.2.4.2.1 Methodology

Wafer 1 from the previous section was diced by mechanical sawing to obtain rectangular thin plate specimens cut along different orientations for the calibration procedure.

For the samples cut along the [100] direction, resistance variations as a function of stress can be obtained from Equation (IV.18). Considering only the resistors oriented along the length (R_1, R_3) and width (R_5, R_7) of the rectangular specimens (Figure IV.4):

$$\begin{aligned}
\frac{\Delta R_1}{R_1} &= \pi_{11}^p \sigma_{11} + \pi_{12}^p \sigma_{22} + \pi_{12}^p \sigma_{33} \\
\frac{\Delta R_3}{R_3} &= \pi_{12}^p \sigma_{11} + \pi_{11}^p \sigma_{22} + \pi_{12}^p \sigma_{33} \\
\frac{\Delta R_5}{R_5} &= \pi_{11}^n \sigma_{11} + \pi_{12}^n \sigma_{22} + \pi_{12}^n \sigma_{33} \\
\frac{\Delta R_7}{R_7} &= \pi_{12}^n \sigma_{11} + \pi_{11}^n \sigma_{22} + \pi_{12}^n \sigma_{33}
\end{aligned} \tag{IV.32}$$

Under uniaxial stress ($\sigma_{11} = \sigma$), a simple relationship is obtained between the resistance change and the prescribed bending stress:

$$\begin{aligned}
\frac{\Delta R_1}{R_1} &= \pi_{11}^p \sigma \\
\frac{\Delta R_3}{R_3} &= \pi_{12}^p \sigma \\
\frac{\Delta R_5}{R_5} &= \pi_{11}^n \sigma \\
\frac{\Delta R_7}{R_7} &= \pi_{12}^n \sigma
\end{aligned} \tag{IV.33}$$

Thus, the π_{11} and π_{12} piezoresistive coefficients can be directly evaluated from the slopes of the calibration curves $\Delta R_i/R_i = f(\sigma)$ for resistors in the longitudinal and transverse directions of the sample, respectively $a_{i,L}$ and $a_{i,T}$:

$$\begin{aligned}
\pi_{11}^p &= a_{1,L} \\
\pi_{12}^p &= a_{3,T} \\
\pi_{11}^n &= a_{5,L} \\
\pi_{12}^n &= a_{7,T}
\end{aligned} \tag{IV.34}$$

For the samples cut along the [110] direction, it is more convenient to express the resistance changes in terms of the stress components in the basis (x_1', x_2', x_3') (Figure IV.1). Then, applying Equation (IV.28) to resistors oriented along the length (R_0, R_4) and width (R_2, R_6) of the

rectangular specimens (Figure IV.4):

$$\begin{aligned}
\frac{\Delta R_0}{R_0} &= \frac{\pi_{11}^p + \pi_{12}^p - \pi_{44}^p}{2}\sigma'_{11} + \frac{\pi_{11}^p + \pi_{12}^p + \pi_{44}^p}{2}\sigma'_{22} + \pi_{12}^p\sigma'_{33} \\
\frac{\Delta R_2}{R_2} &= \frac{\pi_{11}^p + \pi_{12}^p + \pi_{44}^p}{2}\sigma'_{11} + \frac{\pi_{11}^p + \pi_{12}^p - \pi_{44}^p}{2}\sigma'_{22} + \pi_{12}^p\sigma'_{33} \\
\frac{\Delta R_4}{R_4} &= \frac{\pi_{11}^n + \pi_{12}^n - \pi_{44}^n}{2}\sigma'_{11} + \frac{\pi_{11}^n + \pi_{12}^n + \pi_{44}^n}{2}\sigma'_{22} + \pi_{12}^n\sigma'_{33} \\
\frac{\Delta R_6}{R_6} &= \frac{\pi_{11}^n + \pi_{12}^n + \pi_{44}^n}{2}\sigma'_{11} + \frac{\pi_{11}^n + \pi_{12}^n - \pi_{44}^n}{2}\sigma'_{22} + \pi_{12}^n\sigma'_{33}
\end{aligned} \tag{IV.35}$$

Under uniaxial stress ($\sigma'_{11} = \sigma$), the above equation becomes:

$$\begin{aligned}
\frac{\Delta R_0}{R_0} &= \frac{\pi_{11}^p + \pi_{12}^p - \pi_{44}^p}{2}\sigma \\
\frac{\Delta R_2}{R_2} &= \frac{\pi_{11}^p + \pi_{12}^p + \pi_{44}^p}{2}\sigma \\
\frac{\Delta R_4}{R_4} &= \frac{\pi_{11}^n + \pi_{12}^n - \pi_{44}^n}{2}\sigma \\
\frac{\Delta R_6}{R_6} &= \frac{\pi_{11}^n + \pi_{12}^n + \pi_{44}^n}{2}\sigma
\end{aligned} \tag{IV.36}$$

The remaining π_{44} coefficients can then be evaluated from the slopes of the calibration curves $\Delta R_i/R_i = f(\sigma)$ for resistors in the longitudinal or transverse directions of the rectangular specimen, respectively $a_{i,L}$ and $a_{i,T}$:

$$\begin{aligned}
\pi_{44}^p &= \pi_{11}^p + \pi_{12}^p - 2a_{0,T} = 2a_{2,L} - \pi_{11}^p - \pi_{12}^p \\
\pi_{44}^n &= \pi_{11}^n + \pi_{12}^n - 2a_{4,T} = 2a_{6,L} - \pi_{11}^n - \pi_{12}^n
\end{aligned} \tag{IV.37}$$

It can be noted that for Si piezoresistive sensors fabricated on a (001) substrate, the transverse shear piezoresistive coefficient π_{44} can be obtained using the four-point bending method on a sample cut along a <100> direction, although with this type of loading only an inplane uniaxial stress is prescribed[1]. This is due to the *shear-coupling effect* in orthotropic materials, whereby shear strains are induced by a normal load prescribed in any direction other than the symmetry axes of the material, i.e. the <100> directions for the transversely orthotropic Si single-crystal substrate considered here.

[1]For the metal thin film strain gauges presented in Section IV.2.1 however, which are embedded in SiO_2 dielectric, the transverse gauge factor cannot in general be obtained using four-point bending.

Type	π_{11} (TPa^{-1})	π_{12} (TPa^{-1})	π_{44} (TPa^{-1})
P	4.0 ± 2.3	5.7 ± 1.3	492.1 ± 14.6
N	-174.4 ± 1.4	86.0 ± 3.1	-161.1 ± 11.8

TABLE IV.1: Piezoresistive coefficients evaluated by the instrumented four-point bending calibration procedure for P-type and N-type Si resistors used in the stress sensors.

IV.2.4.2.2 Results

The calibration results are shown in Figure IV.11 and Figure IV.12 respectively for P-type and N-type resistors. The measured relative resistance changes are ranging between 1 to 3%/100 MPa, except for resistors R_1 ([100]-oriented, P-type) and R_4 ([$\bar{1}$10]-oriented, N-type) under longitudinal stress, and R_3 ([010]-oriented, P-type) under transverse stress. The latter exhibit very low sensitivity, on the same order as the uncertainty in the resistance measurement (Section IV.2.3.3). For all the other measurements, a linear relationship is obtained between the resistance change and the applied load, confirming the validity of the assumption of linear piezoresistivity in the prescribed stress range.

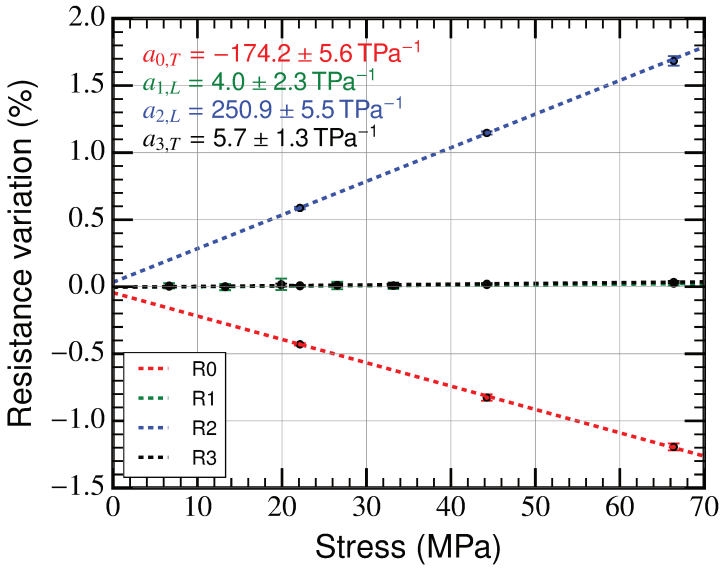


FIGURE IV.11: Relative resistance variation as a function of prescribed stress during the calibration procedure by four-point bending for P-type resistors (mean $\pm$ 3 $\times$ SE).

The piezoresistive coefficients evaluated in the calibration procedure are summarized in Table IV.1. For the π_{44} coefficients, which can be computed using either of the two expressions in Equation (IV.37), the values with the lowest uncertainty was used.

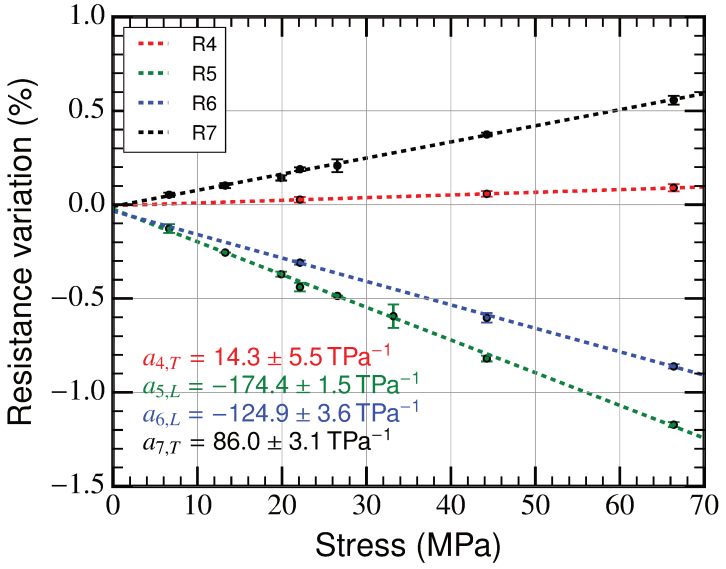


FIGURE IV.12: Relative resistance variation as a function of prescribed stress during the calibration procedure by four-point bending for N-type resistors (mean ± 3 × SE).

IV.2.4.2.3 Discussion

The sign and magnitude of the π_{44}^{p} and π_{11}^{n} coefficients are both in good agreement with the values reported by [34, 35] for highly doped single crystal silicon doped by thermal diffusion. Furthermore, the piezoresistivity theory predicts $\pi_{11}^{n} \approx -2\pi_{12}^{n}$ and $\pi_{11}^{n,p} + 2\pi_{12}^{n,p} \ll \pi_{44}^{n,p}$ [6]. Both of these relationships are verified for the measured piezoresistive coefficients. In addition, internal measurements on piezoresistive stress sensors with the same doping levels yielded similar values [36].

Nevertheless, it can be noted that the magnitude of the calibrated piezoresistive coefficients is more than twice lower than typical values for stress sensor applications [25]. Consequently, the different resistors in the rosette structures exhibit rather low stress sensitivity (≈ 5% for 100 MPa at best), to the extent that the lowest piezoresistive coefficients π_{11}^{p} and π_{12}^{p} could not be evaluated with sufficient accuracy.

In the next section, the obtained piezoresistive coefficients for our Si stress sensors are used to estimate the stress variation in several locations on the investigated test chip from resistance measurements at different steps during the fabrication process.

IV.2.4.3 Stress components estimation

In the following we will assume that the piezoresistive coefficients are identical for all resistors of a given doping type. The stress-induced resistance variations for the eight-element rosettes implemented in the studied test chip (Figure IV.4) can then be written using Equation (IV.17):

$$\begin{aligned}
\frac{\Delta R_0}{R_0} &= \frac{\pi_{11}^p + \pi_{12}^p}{2}\sigma_{11} + \frac{\pi_{11}^p + \pi_{12}^p}{2}\sigma_{22} + \pi_{12}^p\sigma_{33} - \pi_{44}^p\sigma_{12} \\
\frac{\Delta R_1}{R_1} &= \pi_{11}^p\sigma_{11} + \pi_{12}^p\sigma_{22} + \pi_{12}^p\sigma_{33} \\
\frac{\Delta R_2}{R_2} &= \frac{\pi_{11}^p + \pi_{12}^p}{2}\sigma_{11} + \frac{\pi_{11}^p + \pi_{12}^p}{2}\sigma_{22} + \pi_{12}^p\sigma_{33} + \pi_{44}^p\sigma_{12} \\
\frac{\Delta R_3}{R_3} &= \pi_{12}^p\sigma_{11} + \pi_{11}^p\sigma_{22} + \pi_{12}^p\sigma_{33} \\
\frac{\Delta R_4}{R_4} &= \frac{\pi_{11}^n + \pi_{12}^n}{2}\sigma_{11} + \frac{\pi_{11}^n + \pi_{12}^n}{2}\sigma_{22} + \pi_{12}^n\sigma_{33} - \pi_{44}^n\sigma_{12} \\
\frac{\Delta R_5}{R_5} &= \pi_{11}^n\sigma_{11} + \pi_{12}^n\sigma_{22} + \pi_{12}^n\sigma_{33} \\
\frac{\Delta R_6}{R_6} &= \frac{\pi_{11}^n + \pi_{12}^n}{2}\sigma_{11} + \frac{\pi_{11}^n + \pi_{12}^n}{2}\sigma_{22} + \pi_{12}^n\sigma_{33} + \pi_{44}^n\sigma_{12} \\
\frac{\Delta R_7}{R_7} &= \pi_{12}^n\sigma_{11} + \pi_{11}^n\sigma_{22} + \pi_{12}^n\sigma_{33}
\end{aligned} \tag{IV.38}$$

IV.2.4.3.1 Method 1: Least-squares approximation

A generic approach to derive the variation of the different stress components from the resistance changes measured in the eight-element rosette is to compute the least-squares solution for the overconstrained system (IV.38), with eight equations and four unknowns. The main advantage of this method is that it enables to use the information obtained from all eight resistors rather than selecting a specific four-element rosette among many possible combinations. It was proposed and successfully applied for MOS stress sensors by [29, 31, 37, 38].

Expressing Equation (IV.38) in matrix form yields:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{IV.39}$$

for which the associated *normal equation* writes:

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b} \tag{IV.40}$$

Because the system is overdetermined, $\mathbf{A}^T\mathbf{A}$ is invertible and Equation IV.38 can be solved for x:

$$x = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b} = \mathbf{A}^{\dagger}\mathbf{b} \tag{IV.41}$$

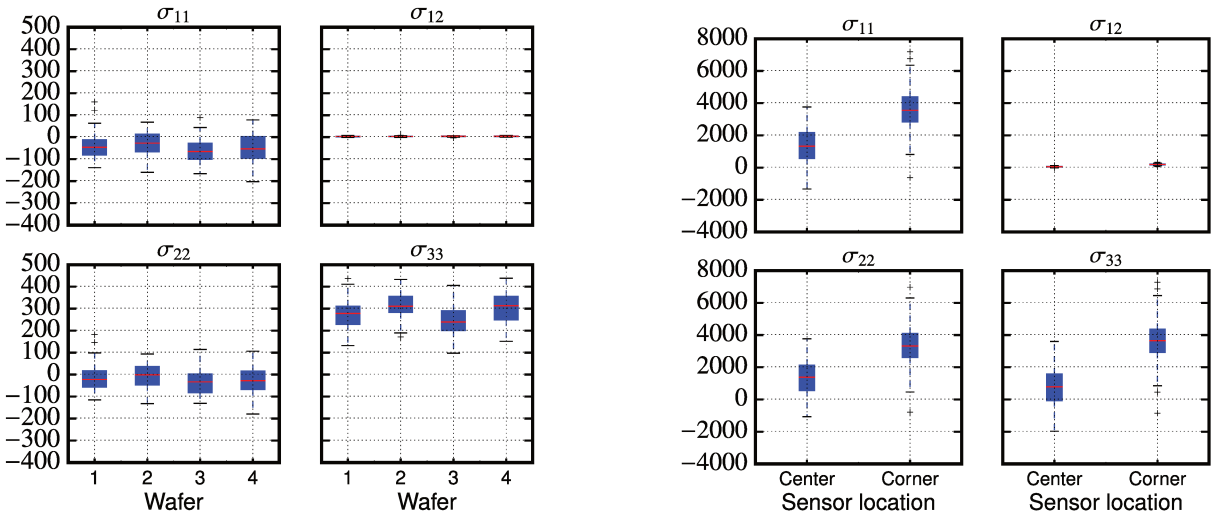
For numerical stability, QR factorization is used to compute the solution [39]. Decomposing the rectangular matrix as $A = QR$ with Q orthogonal and R upper triangular, the pseudo-inverse becomes:

$$\mathbf{A}^{\dagger} = \left(\mathbf{R}^T\mathbf{Q}^T\mathbf{Q}\mathbf{R}\right)^{-1}\mathbf{R}^T\mathbf{Q}^T = \mathbf{R}^{-1}\mathbf{Q} \tag{IV.42}$$

and thus:

$$x = \mathbf{R}^{-1}\mathbf{Q}\mathbf{b} \tag{IV.43}$$

yielding the least-square solution for Equation (IV.38).



(a) M2/PADOPEN (b) PADOPEN/PKG

FIGURE IV.13: Boxplots of the variation of the stress components (in MPa), evaluated at the die center between M2/PADOPEN steps for the four wafers (a) and in wafer 2 between PADOPEN/PKG steps at the die center and corner (b).

Between the M2 and PADOPEN steps, the estimated variations for the stress components using this method are quite consistent between the four investigated wafers. The results at the die center are plotted in Figure IV.13a. The obtained inplane normal stresses σ_{11} and σ_{22} are slightly compressive, with median values ranging between 0 and −70 MPa. For each wafer, the dispersion is very large, with an interquartile range close to 100 MPa, comparable to the estimated stress variations. That is also the case of the estimated out-of-plane normal stress σ_{33}, for which a surprisingly large tensile stress is obtained with median values comprised between 250 and 300 MPa is obtained. This nonzero out-of-plane normal stress, although believed to be greatly overestimated, could be due to a local stress induced by the hybrid bonding pads, which are the only interconnects present above the M2 level. In the substrate plane, the measured shear stresses σ_{12} are very close to zero with very low dispersion (below 10 MPa). No large difference was observed between the estimated stresses at the die center compared to other sensor locations, along the die edges and at the corner.

Between the PADOPEN and PKG steps, the obtained stress variations in the packaged dies (extracted from wafer 2) are on the order of 1 GPa at the die center and 4 GPa at the die corner for the three normal stresses σ_{11}, σ_{22} and σ_{33}, with a large dispersion of about 1 GPa

(Figure IV.13b). These stresses are unrealistically high, almost comparable to the theoretical strength of silicon (on the order of $E/10$). Again, the estimated inplane shear stress σ_{12} vanishes.

With this method, although all eight resistors in the rosette structure can be used altogether for the evaluation of the stress components, the overconstrained system of equations (IV.38) has no exact solution and thus the obtained result in an approximation. This approximate solution may be regarded as an average (in the least-squares sense) and thus be sensitive to outliers, which could play a role in explaining the surprisingly large values obtained for the σ_{33} component between the M2/PADOPEN steps and for σ_{11}, σ_{22} and σ_{33} between the PADOPEN/PKG steps.

In order to check this hypothesis, the variations of the different stress components between the M2/PADOPEN and PADOPEN/PKG steps were estimated again using a second method.

IV.2.4.3.2 Method 2: Direct inversion

Instead of considering the eight resistors in the stress sensor altogether, a different approach is to select two four-element rosettes within the sensor, for which direct inversion of the piezoresistivity equations is possible. In the basis formed by the principal axes of the Si lattice ($x_1 = [100]$, $x_1' = [110]$, $x_2 = [010]$, $x_2' = [\bar{1}10]$), there are twelve possible combinations to form a 0–45–90–135° rosette including both doping types:

(3P1N) rosettes comprising three P-type resistors and a single N-type resistor. Depending on the orientation of the N-type resistor, there are four distinct configurations: 1N3P, 1P1N2P, 2P1N1P and 3P1N;

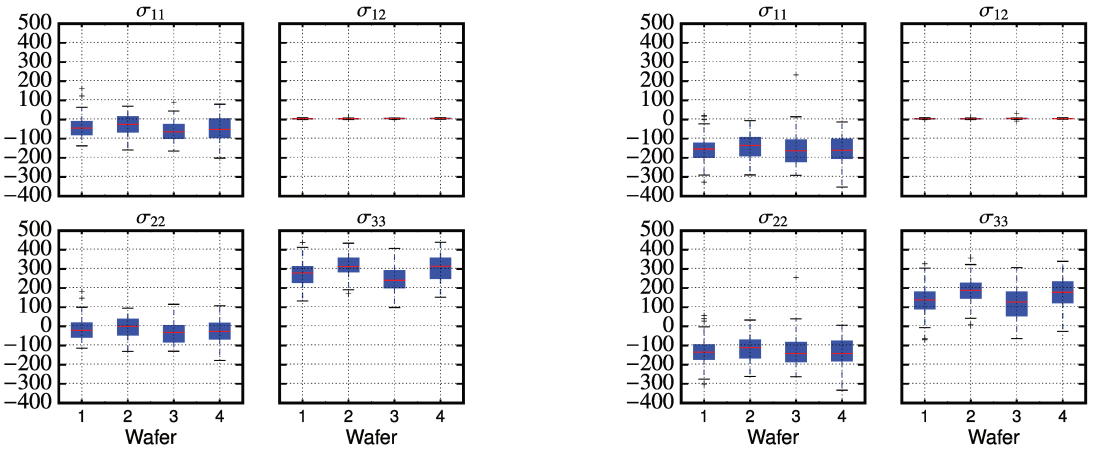
(1P3N) rosettes comprising a single P-type resistor and three N-type resistors. Again, depending on the orientation of the P-type resistor, there are four distinct configurations: 1P3N, 1N1P2N, 2N1P1N and 3N1P;

(2P2N) rosettes comprising two P-type resistors and two N-type resistors. Depending on whether the angle formed between two resistors with the same doping type is 45 or 90°, there are four distinct configurations: 2N2P, 2P2N, $(1N1P)^2$, $(1P1N)^2$.

The corresponding systems of equations are derived and inverted in Appendix C for each group. It can be seen from Equations (C.2-C.12) that among all configurations, those yielding the simplest expression for the stress components $\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}$ are configurations $(1P1N)^2$ and $(1N1P)^2$, respectively associated with rosette A (R_5, R_2, R_7, R_0) and rosette B (R_1, R_6, R_3, R_4) in our stress sensor (Figure IV.4). With the simple equations obtained for these two configurations, involving fewer terms compared to the other possible four-element rosette choices, the estimated stress variations are believed to be less prone to error/uncertainty accumulation. Thus, a possible approach is to derive exact solutions for both selected rosettes by direct inversion of the piezoresistivity equations, and then to compute an average value for the different stress components based on these two solutions. However, for reasons that will become apparent in the next paragraph, configuration $(1P1N)^2$ (rosette A) leads to exceedingly inaccurate results,

and therefore only configuration $(1N1P)^2$ (rosette B) was considered. It is associated with the following system of equations:

$$\begin{cases} \sigma_{11} = \dfrac{\pi_{12}^p\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-\pi_{12}^n\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} + \dfrac{1}{2\left(\pi_{11}^n-\pi_{12}^n\right)}\left[\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right] \\ \sigma_{22} = \dfrac{\pi_{12}^p\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-\pi_{12}^n\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} - \dfrac{1}{2\left(\pi_{11}^n-\pi_{12}^n\right)}\left[\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right] \\ \sigma_{33} = \dfrac{\left(\pi_{11}^n+\pi_{12}^n\right)\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]-\left(\pi_{11}^p+\pi_{12}^p\right)\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} \\ \sigma_{12} = \dfrac{1}{2\pi_{44}^p}\left[\frac{\Delta R_2}{R_2}-\frac{\Delta R_4}{R_4}\right] \end{cases} \tag{IV.44}$$



(a) Least-squares (b) Direct inversion

FIGURE IV.14: Boxplots of the individual stress components variation (in MPa) between the M2/PADOPEN steps, evaluated at the die center for the four wafers by direct inversion using a four-element rosette (a) and by the least-squares method using all eight resistors (b).

Comparing the estimated stress variations obtained by direct inversion and the least-square method between the M2/PADOPEN steps, a notable difference can be seen (Figure IV.14). The inplane normal stresses σ_{11} and σ_{22} obtained using direct inversion are about 100 MPa more compressive compared to those obtained using the least-squares method. A similar difference is observed for the out-of-plane normal stress σ_{33}, which is about 100 MPa less tensile using direct inversion. Inplane shear stress σ_{12} on the other hand is overall insensitive to the chosen method. It can also be noted that the dispersion of the values is larger for direct inversion, with an increase of the interquartile range by about one third. Although not detailed here, similar trends were obtained for the estimated stress variation between the PADOPEN/PKG steps. As with the previous approach, unrealistically large values are obtained for all three normal stress components.

Therefore, the choice of the method used for stress estimation alone cannot explain the observed discrepancies. In the next paragraph, the influence of temperature-induced resistance changes,

which has been neglected so far[1], is investigated for our stress sensors.

IV.2.4.3.3 Method 3: Temperature-compensated stress

Discussion: influence of temperature on the stress estimation

In both of the previous methods, the influence of temperature has been neglected in the equations for the stress-induced resistance changes in the eight-element rosette (IV.38). And yet, a temperature-dependence may arise from at least two factors:

- through the piezoresistive coefficients:

$$\pi_{\alpha\beta} = \pi_{\alpha\beta}(T) \tag{IV.45}$$

- through the resistivity:

$$\rho\,(T) = \rho\,(T_0)\,[1 + \alpha\,(T - T_0)] \tag{IV.46}$$

  neglecting higher-order terms, with α the *temperature coefficient of resistivity*, T_0 a reference temperature and ρ_0 the resistivity at T_0.

Among these two contributions, the temperature-dependence of the piezoresistive coefficients is assumed to be negligible for our stress sensors, based on the measurements of [35]. Their results indicate that the temperature-dependence of the piezoresistivity coefficients vanishes with increasing doping density and becomes almost inexistent for dopant surface concentrations on the order of $10^{20}\,\mathrm{cm}^{-3}$, as is the case in this study.

The temperature-dependence of the resistivity on the other hand may have significant impact on the resistance measurement itself. Indeed, the measured resistance then writes:

$$R\,(T) = R\,(T_0)\,[1 + \alpha\,(T - T_0)] \tag{IV.47}$$

where α is the *temperature coefficient of resistance (TCR)*.

Accounting for the temperature-induced resistance change, the system of equations (IV.38) becomes, for a four-element rosette:

$$\begin{aligned}
\frac{\Delta R_a}{R_a} &= F_1\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}\right) + \alpha_{n,p}\Delta T \\
\frac{\Delta R_b}{R_b} &= F_2\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}\right) + \alpha_{n,p}\Delta T \\
\frac{\Delta R_c}{R_c} &= F_3\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}\right) + \alpha_{n,p}\Delta T \\
\frac{\Delta R_d}{R_d} &= F_4\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}\right) + \alpha_{n,p}\Delta T
\end{aligned} \tag{IV.48}$$

[1] Including in previous studies [30, 31].

with $\{a, b, c, d\}$ the indices of the resistors associated with the selected rosette configuration, and $\Delta T = T - T_0$ the difference between the measurement temperatures for the resistance in the stressed state R^σ and the reference state R^0, respectively T and T_0.

For all possible four-element rosette configurations (R_a, R_b, R_c, R_d), the expressions of the four accessible stress components $(\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12})$ were derived in Appendix C, accounting for the temperature-dependent terms $\alpha_{n,p}\Delta T$. From the obtained equations, it can be seen that additional "temperature-induced" stresses systematically appear in the expressions for the normal stresses $(\sigma_{11}, \sigma_{22}, \sigma_{33})$, whereas the inplane shear stress component σ_{12} is not sensitive to temperature fluctuations. The obtained expressions are of the form:

$$\begin{aligned}
\sigma_{11} &= G_1\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \frac{\Delta R_a}{R_a}, \frac{\Delta R_b}{R_b}, \frac{\Delta R_c}{R_c}, \frac{\Delta R_d}{R_d}\right) - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n}\Delta T \\
\sigma_{22} &= G_2\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \frac{\Delta R_a}{R_a}, \frac{\Delta R_b}{R_b}, \frac{\Delta R_c}{R_c}, \frac{\Delta R_d}{R_d}\right) - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n}\Delta T \\
\sigma_{33} &= G_3\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \frac{\Delta R_a}{R_a}, \frac{\Delta R_b}{R_b}, \frac{\Delta R_c}{R_c}, \frac{\Delta R_d}{R_d}\right) + \frac{\alpha_n \left(\pi_{11}^p + \pi_{12}^p\right) - \alpha_p \left(\pi_{11}^n + \pi_{12}^n\right)}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n}\Delta T \\
\sigma_{33} &= G_4\left(\pi_{11}^{n,p}, \pi_{12}^{n,p}, \pi_{44}^{n,p}, \frac{\Delta R_a}{R_a}, \frac{\Delta R_b}{R_b}, \frac{\Delta R_c}{R_c}, \frac{\Delta R_d}{R_d}\right)
\end{aligned} \tag{IV.49}$$

The temperature-dependent terms are the same for all configurations and depend on the piezoresistive coefficients $\pi_{11}^{n,p}$ and $\pi_{12}^{n,p}$, as well as the temperature coefficients of resistance $\alpha_{n,p}$. Depending on the relative values of these coefficients, the error induced in the estimated stress components may thus be significant.

For the Si piezoresistors considered here, the measured temperature coefficients are $\alpha_n = 1430\,\text{ppm/°C}$ and $\alpha_p = 1450\,\text{ppm/°C}$. Using the piezoresistive coefficients determined in the previous section (Table IV.1), it is estimated that a temperature fluctuation of $\pm 1\,°\text{C}$ may lead to a drift of $\mp 87\,\text{MPa}$ in the measured values of $(\sigma_{11}, \sigma_{22})$, and $\pm 106\,\text{MPa}$ for σ_{33}. Such values are well within the range of the stress variations expected during the fabrication process.

Therefore, temperature fluctuations between two resistance measurements with the previous methods will cause a temperature-induced resistance change that is wrongly attributed to a stress variation. To avoid these spurious stress variations, several strategies may be considered:

- using resistors with different characteristics, leading to a lower temperature-dependence;
- measuring the temperature near the resistors, in order to be able to evaluate the temperature-dependent terms quantitatively for each sensor;
- working with temperature-compensated stress components, as proposed by [40].

The first and second solutions requiring new test chips could not be explored in this work. Therefore, temperature-compensation of the estimated stress variations was investigated. This approach consists in retaining among the four accessible stress components, only those for which the temperature-dependent terms can be suppressed. Noticing that the temperature terms

are the same for the inplane normal stress components σ_{11} and σ_{22} in Equation (IV.48), a temperature-compensated stress component is simply obtained by subtraction, working with $\sigma_{11} - \sigma_{22}$ instead of σ_{11} and σ_{22}. The out-of-plane stress component σ_{33} cannot be exploited because its temperature-dependent term cannot be evaluated, nor compensated. The inplane shear stress σ_{12} on the other hand is intrinsically insensitive to temperature fluctuations. The temperature-independent stress components for the considered four-element sensor configuration are thus $(\sigma_{11} - \sigma_{22}, \sigma_{12})$.

As mentioned in the previous paragraph (Section IV.2.4.3.2), among the twelve sensor configurations detailed in Appendix C, configurations $(1N1P)^2$ and $(1P1N)^2$, respectively associated with rosette A (R_5, R_2, R_7, R_0) and rosette B (R_1, R_6, R_3, R_4), lead to the simplest expressions for the stress components, and are thus deemed particularly advantageous due to an expected lower sensitivity to errors or uncertainties. These expressions are written as follows:

- Rosette A:

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^p \left[\frac{\Delta R_5}{R_5} + \frac{\Delta R_7}{R_7}\right] - \pi_{12}^n \left[\frac{\Delta R_2}{R_2} + \frac{\Delta R_0}{R_0}\right]}{2\left(\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n\right)} + \frac{1}{2\left(\pi_{11}^n - \pi_{12}^n\right)} \left[\frac{\Delta R_5}{R_5} - \frac{\Delta R_7}{R_7}\right] - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{22} = & \frac{\pi_{12}^p \left[\frac{\Delta R_5}{R_5} + \frac{\Delta R_7}{R_7}\right] - \pi_{12}^n \left[\frac{\Delta R_2}{R_2} + \frac{\Delta R_0}{R_0}\right]}{2\left(\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n\right)} - \frac{1}{2\left(\pi_{11}^n - \pi_{12}^n\right)} \left[\frac{\Delta R_5}{R_5} - \frac{\Delta R_7}{R_7}\right] - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{33} = & \frac{\left(\pi_{11}^n + \pi_{12}^n\right)\left[\frac{\Delta R_2}{R_2} + \frac{\Delta R_0}{R_0}\right] - \left(\pi_{11}^p + \pi_{12}^p\right)\left[\frac{\Delta R_5}{R_5} + \frac{\Delta R_7}{R_7}\right]}{2\left(\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n\right)} + \frac{\alpha_n\left(\pi_{11}^p + \pi_{12}^p\right) - \alpha_p\left(\pi_{11}^n + \pi_{12}^n\right)}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^p} \left[\frac{\Delta R_2}{R_2} - \frac{\Delta R_0}{R_0}\right] \end{cases} \tag{IV.50}$$

- Rosette B:

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^n \left[\frac{\Delta R_1}{R_1} + \frac{\Delta R_3}{R_3}\right] - \pi_{12}^p \left[\frac{\Delta R_6}{R_6} + \frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^p \pi_{12}^n - \pi_{11}^n \pi_{12}^p\right)} + \frac{1}{2\left(\pi_{11}^p - \pi_{12}^p\right)} \left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_3}{R_3}\right] - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{22} = & \frac{\pi_{12}^n \left[\frac{\Delta R_1}{R_1} + \frac{\Delta R_3}{R_3}\right] - \pi_{12}^p \left[\frac{\Delta R_6}{R_6} + \frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^p \pi_{12}^n - \pi_{11}^n \pi_{12}^p\right)} - \frac{1}{2\left(\pi_{11}^p - \pi_{12}^p\right)} \left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_3}{R_3}\right] - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{33} = & \frac{\left(\pi_{11}^p + \pi_{12}^p\right)\left[\frac{\Delta R_6}{R_6} + \frac{\Delta R_4}{R_4}\right] - \left(\pi_{11}^n + \pi_{12}^n\right)\left[\frac{\Delta R_1}{R_1} + \frac{\Delta R_3}{R_3}\right]}{2\left(\pi_{11}^p \pi_{12}^n - \pi_{11}^n \pi_{12}^p\right)} + \frac{\alpha_n\left(\pi_{11}^p + \pi_{12}^p\right) - \alpha_p\left(\pi_{11}^n + \pi_{12}^n\right)}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^n} \left[\frac{\Delta R_6}{R_6} - \frac{\Delta R_4}{R_4}\right] \end{cases} \tag{IV.51}$$

It can be noted that these configurations are also the most suitable for temperature compensation of the estimated stress components. Indeed, as can be seen from Equations (IV.50) and (IV.51), equations of very similar form are obtained for the inplane normal stresses σ_{11} and σ_{22}, leaving only one term after subtraction to obtain the temperature-compensated stress $\sigma_{11} - \sigma_{22}$. In these configurations, the expression for the inplane shear component σ_{12} also comprises only one term. For rosette A and B, the temperature-compensated stress components thus write:

- Rosette A:

$$\begin{cases} \sigma_{11} - \sigma_{22} = & \frac{1}{\pi_{11}^n - \pi_{12}^n} \left[\frac{\Delta R_5}{R_5} - \frac{\Delta R_7}{R_7} \right] \\ \sigma_{12} = & \frac{1}{2\pi_{44}^p} \left[\frac{\Delta R_2}{R_2} - \frac{\Delta R_0}{R_0} \right] \end{cases} \tag{IV.52}$$

- Rosette B:

$$\begin{cases} \sigma_{11} - \sigma_{22} = & \frac{1}{\pi_{11}^p - \pi_{12}^p} \left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_3}{R_3} \right] \\ \sigma_{12} = & \frac{1}{2\pi_{44}^n} \left[\frac{\Delta R_6}{R_6} - \frac{\Delta R_4}{R_4} \right] \end{cases}, \tag{IV.53}$$

It can be seen in the above expressions that only three piezoresistive coefficients appear. As a result, these configurations could enable a less time-consuming calibration procedure, since not all the coefficients are required for the stress estimation.

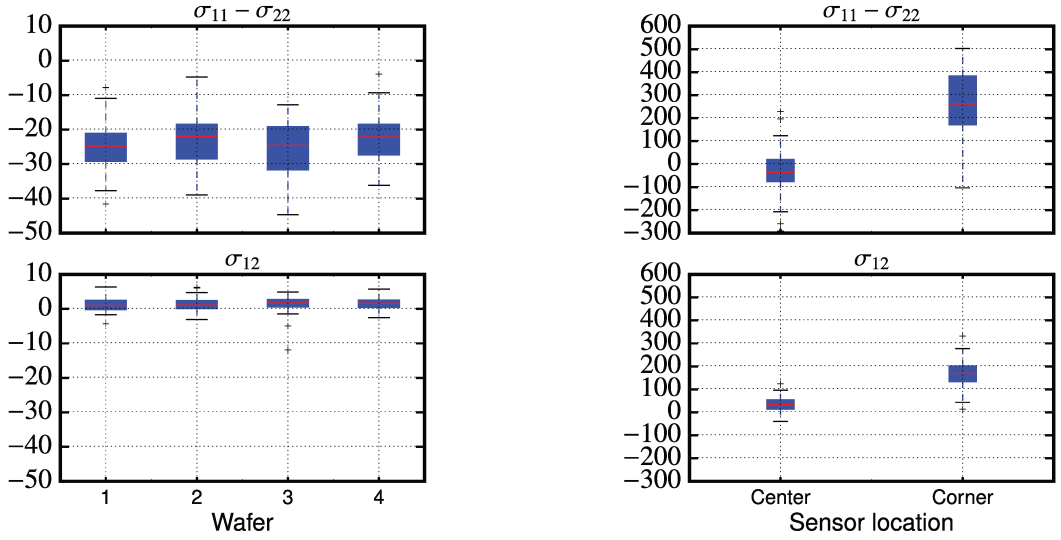
There are still two possible four-element rosette configurations to choose from, however. A major difference between Equation (IV.52) and (IV.53) lies in the involved piezoresistive coefficients: namely, π_{44}^p and $\pi_{11}^n - \pi_{12}^n$ for rosette A, and π_{44}^n and $\pi_{11}^p - \pi_{12}^p$ for rosette B. By selecting the configuration in which the most accurately calibrated piezoresistive coefficients appear, the precision obtained for the estimated temperature-compensated stress components can be optimized [41]. In that regard, based on the previous calibration results (Table IV.1), rosette A is therefore clearly the best configuration for stress estimation. Indeed, its associated piezoresistive coefficients π_{44}^p, π_{11}^n and π_{12}^n are the largest and were therefore determined with the best accuracy. Conversely, for rosette B the coefficients π_{11}^p and π_{12}^p are an order of magnitude lower with very large uncertainties, and more importantly are almost equal. Thus, due to the presence of the term $\pi_{11}^p - \pi_{12}^p$ at the denominator in Equation (IV.53), the temperature-compensated stress component $\sigma_{11} - \sigma_{22}$ should not be evaluated with rosette B[1]. Nevertheless, the inplane shear component σ_{12} can still be obtained since the coefficient π_{44}^n was calibrated with sufficient accuracy.

Results

The temperature-independent stress variations estimated at the die center are compared between the M2/PADOPEN steps and between the PADOPEN/PKG steps for rosette A (Figure IV.15). Again, the results are quite consistent between the four wafers.

Between the M2/PADOPEN steps, the estimated temperature-compensated stress $\sigma_{11} - \sigma_{22}$ is quite low, in the range of -20 to -30 MPa, and the inplane shear stress σ_{12} almost vanishes, with values comprised between 0 and 5 MPa (Figure IV.15a). Although not detailed here, the temperature-independent stress components were also plotted for rosette B. As inferred from Equations (IV.50) and (IV.51), with rosette B the temperature-compensated stress cannot be precisely evaluated, whereas the inplane shear stresses are still exploitable. For the latter, the obtained values range between 0 and 10 MPa, quite

[1] That is the reason why rosette B was not considered in the previous paragraph (Section IV.2.4.3.2)



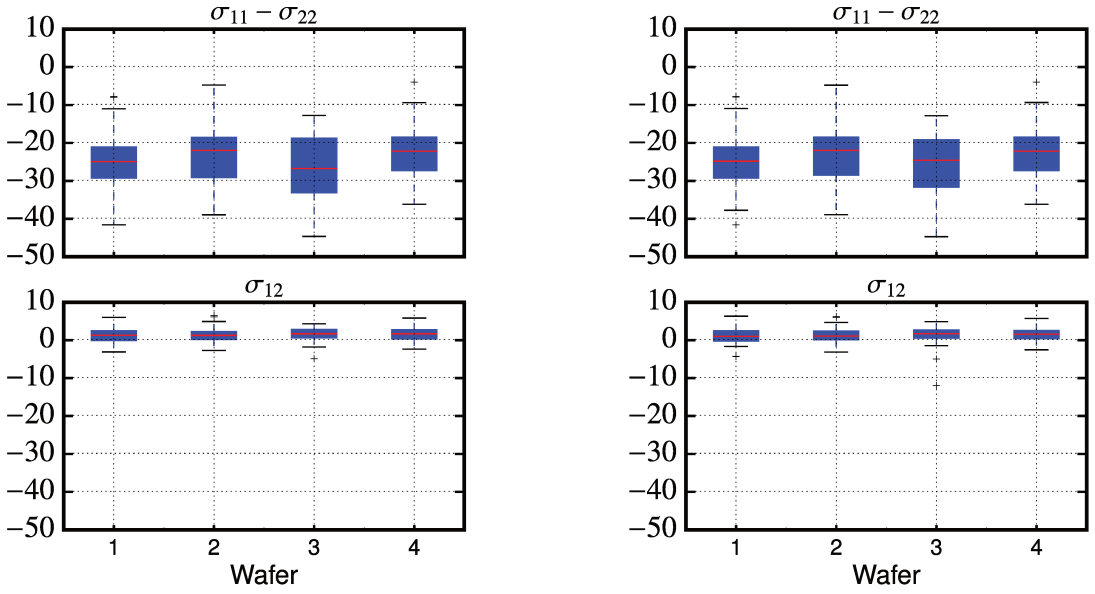
(a) M2/PADOPEN (b) PADOPEN/PKG

FIGURE IV.15: Boxplots of the temperature-independent stress components variation (in MPa) between M2/PADOPEN steps (a) and between PADOPEN/PKG steps (b) evaluated at the die center using rosette A.

close to the shear stresses measured with rosette A, however with a slightly larger dispersion.

Between the PADOPEN/PKG steps, the estimated temperature-compensated stress $\sigma_{11} - \sigma_{22}$ at the die center is comparable to that obtained between the M2/PADOPEN steps, about a few dozen megapascals compressive. The dispersion is significantly higher however, with an interquartile range of approximately 100 MPa. This is in contrast with the two previous methods, for which a significant difference in magnitude between the wafer-level and package-level results was observed. At the die corner however, the temperature-compensated stress is much larger, with values around 200 to 400 MPa tensile (Figure IV.15b). The same trend is observed for the inplane shear stresses σ_{12}, for which values ranging between 0 and 50 MPa are measured at the die center, compared to between 100 to 200 MPa at the die corner.

Comparing the temperature-compensated stress variation $\sigma_{11} - \sigma_{22}$ between the M2/PADOPEN steps evaluated by the direct method and the least-square method, almost identical results are obtained (Figure IV.16), with slightly more dispersion for wafer 3 on which a large number of outliers were seen in the resistance measurements after the M2 step (Section IV.2.4.1). The inplane shear stress σ_{12} is essentially insensitive to the chosen method. Similar trends are obtained for the stress variation between the PADOPEN/PKG steps.



(a) Least-squares (rosette A+B)

(b) Direct method (rosette A)

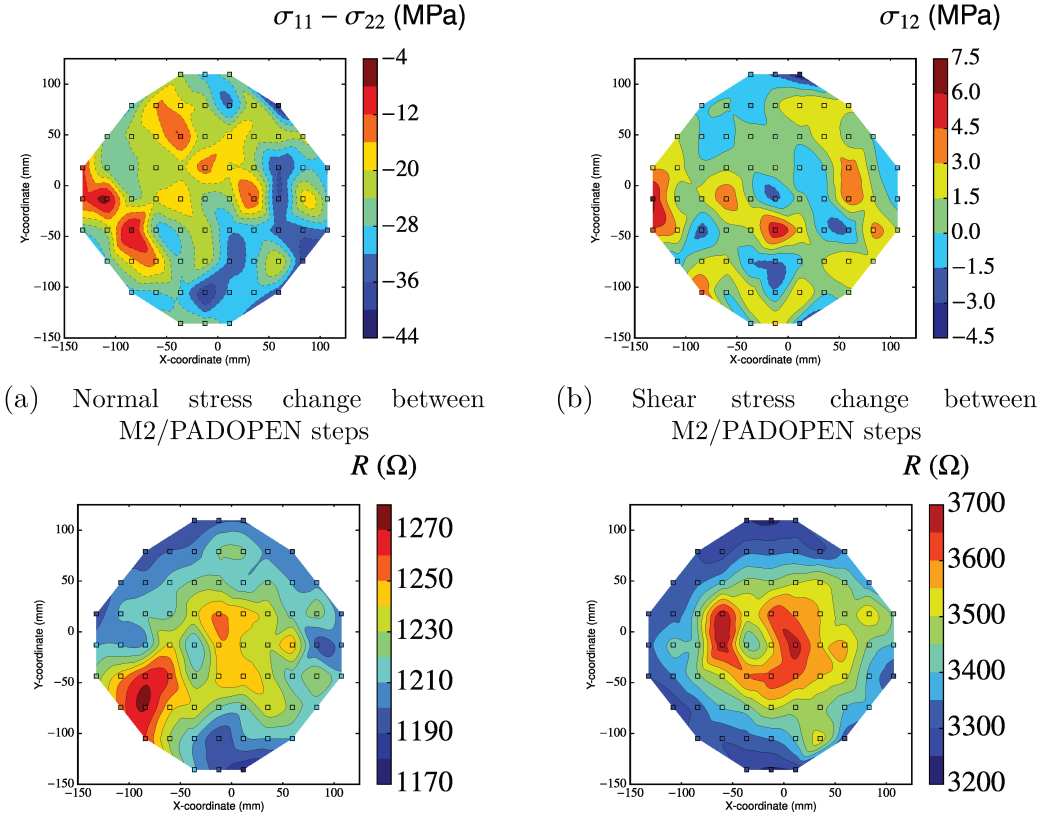
FIGURE IV.16: Boxplots of the temperature-independent stress variations (in MPa) between the M2/PADOPEN steps evaluated at the die center using rosette A (**??**) by direct inversion and rosette A+B (**??**) by the least-squares method.

Discussion: influence of piezoresistive coefficients variability

In addition to the influence of the selected rosette configuration or temperature-induced resistance changes on the evaluated stresses, which were investigated above, another potential error source is an inter-wafer, intra-wafer and intra-sensor variability of the piezoresistive coefficients. Indeed, throughout this analysis the piezoresistive coefficients (but also the temperature coefficients of resistance) have been assumed identical for all resistors of the same doping type, both at the sensor scale, at the die scale, and at the wafer scale. However, the value of the piezoresistive coefficients was shown to depend not only on the doping type, but also on dopant surface concentration [e.g. 34, 35, 42]. In particular, the results of [35] indicate that this sensitivity to dopant concentration is especially acute for high doping levels, as is the case in this study with approximately $10^{20}\,\mathrm{cm}^{-3}$. Because the doping process (here, ion implantation followed by an activation annealing) is known to lead to a variability of the dopant concentration at the wafer scale [33], it can be inferred that this variability may in turn reflect into the wafer-scale distribution of the piezoresistive coefficients.

The variability of the piezoresistive coefficients could not be evaluated directly in this work, due to a limited number of samples and the time-consuming nature of the calibration procedure. Instead, the variability in the resistance measurements at the M2 step, evaluated in Section IV.2.4.1, is used as an indirect indicator for the variability of the doping process. If a strong correlation is observed between the spatial distribution of the initial resistance

values and of the stress variations between the M2/PADOPEN steps, computed assuming constant piezoresistive coefficients, then it could be a sign that the variability of the piezoresistive coefficients indeed plays a significant role on the stress estimation. The results from rosette A are chosen for the comparison, since as it can be seen from Equations (IV.50) the computed stress components in that case depend on only one or two coefficients, either of the N type or the P type, respectively for the temperature-compensated stress component $\sigma_{11} - \sigma_{22}$ and the inplane shear stress σ_{12}. From the wafer maps plotted in Figure IV.17, no common signature is apparent however and thus it could not be shown that there is any significant effect of doping inhomogeneity on stress estimation.



(a) Normal stress change between M2/PADOPEN steps

(b) Shear stress change between M2/PADOPEN steps

(c) N-type resistance intra-wafer variability

(d) P-type resistance intra-wafer variability

FIGURE IV.17: Comparison of the spatial distribution of the evaluated stress changes between M2/PADOPEN steps (a and b) with intra-wafer (c and d) variability of the initial resistance values measured after the M2 step on wafer 1.

In the previous sections, the process-induced stress variations were evaluated experimentally during both front-end processing and back-end processing using Si piezoresistive sensors. The emphasis was put on methodology aspects, aiming to optimize the accuracy of the estimated stress variations through rosette configuration selection, and to obtain a result that is independent from the method selected to derive the stress components, or from external factors such as temperature. However, for Si piezoresistive stress sensors on a (100) wafer only four stress components are accessible, namely $(\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12})$.

This number had to be reduced to two stress components ($\sigma_{11} - \sigma_{22}, \sigma_{12}$), in order to suppress any large spurious stresses expected due to temperature-induced resistance changes between two given process steps. As a result, although the inplane shear stress component σ_{12} could be estimated, the individual normal stress components ($\sigma_{11}, \sigma_{22}, \sigma_{33}$) are no longer accessible. It was thus not possible to achieve the initial aim of this chapter, i.e. assessing whether a significant increase of mechanical stress can be expected in the active region of the image sensor due to the 3D BSI configuration, which may lead a decrease in the electro-optical performance of the different semiconductor components inside the active pixel sensors (e.g. photodiodes, transistors). Therefore, in the next section, finite element modeling of the front-end and back-end fabrication processes is carried out, aiming to better understand the experimental results and complement the measurements with the missing stress components. In addition, very large stresses were systematically obtained at the corner location for the chips measured after the packaging step. Based on this finite element analysis, we will also aim to assess whether such large stresses are indeed likely to occur, or if they could simply result from a measurement error.

IV.3 Finite element modeling of process-induced stress build-up

IV.3.1 Front-end processing

In order to estimate process-induced stress build-up in the active region of the chip during front-end processing, and compare with the stresses evaluated using the piezoresistive stress sensors in the previous section, a 3D finite element model is proposed in which sequential layer deposition, etching and heating steps are simulated using finite element code ANSYS (Mechanical APDL, Release 19.0, Cannonsburg, PA).

IV.3.1.1 Model description

In the top tier of the 3D integrated circuit, where the piezoresistive stress sensors are located, specific design rules are defined, as described in Section IV.2.3.1. In these regions, the 4ML interconnect layout is considerably simplified, with an exclusion area for the M3 and M4 levels, in which no Cu lines or vias allowed, while in the M1 and M2 levels Cu lines and vias as well as W plugs are present to enable electrical access to the sensor. Hybrid bonding pads are still included above the exclusion area however, arranged in a grid pattern with regular spacing for topography homogeneity reasons.

In the bottom tier of the 3D integrated circuit on the other hand, no particular constraint is prescribed in the sensor regions and the standard design rules for the considered technology node are followed. Below the sensors, W plugs and Cu vias are absent from the 7ML

bottom tier interconnect stack, and metal patterning consists of square Cu pads arranged in a staggered grid.

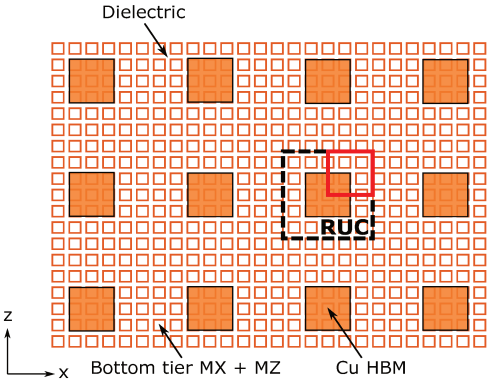
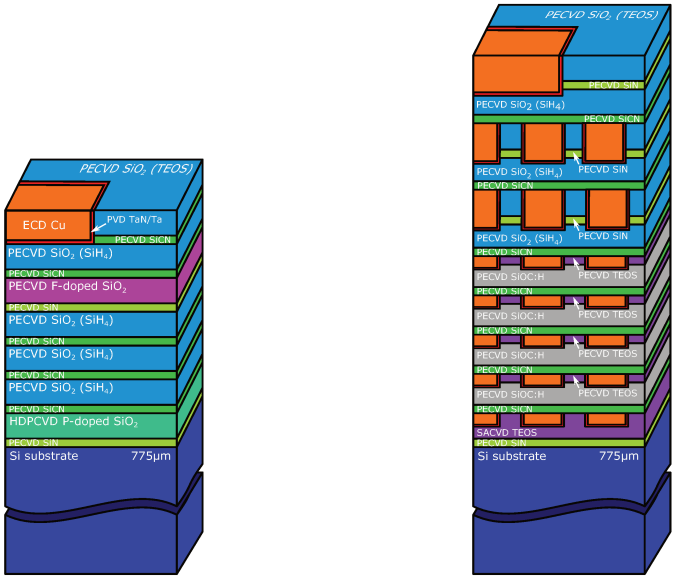


FIGURE IV.18: Schematical top view of the periodic arrangement of the metallic patterns in the top and bottom tiers of the 3D integrated circuit interconnect stack.



(a) Top tier (b) Bottom tier

FIGURE IV.19: Schematical description of the top and bottom tier interconnect stacks in the 3D integrated circuit.

To account for this complex (yet minimal) interconnect structure in the finite element model, a number of simplifying assumptions were made. The stress sensor itself is not accounted for, nor the associated access interconnects present below the M3 level. In addition, a grid pattern with a regular spacing is considered for the Cu pads in the bottom tier interconnect stack, although they are slightly staggered in the actual layout (by less than a half-pitch). This enables to extract a representative unit cell within the formed periodic regular array. Cu hybrid bonding pads have width w and pitch p, and the bottom tier Cu interconnects width $w/3$ and pitch $p/5$ (Figure IV.18). Since the focus is placed on the stress exerted on the active region and the Cu hybrid bonding pads are located in the upper layers of the interconnect stack, the pad octagonal shape can be approximated by a square without significantly altering the results. In addition, due to the symmetries

of the considered unit cell, only a quadrant of width $p/2$ is modeled for each tier. This periodic unit cell has a thickness $h_w + h_i$, the sum of the thicknesses of the wafer and the considered interconnect stack for the top or bottom tier (Figure IV.19). The following boundary conditions are prescribed on the quarter unit cell:

– to enforce the symmetry condition, the displacements along the x and z axes are blocked, respectively on the $x = 0$ and $z = 0$ faces:

$$\begin{cases} u_x(0, y, z) = 0 \\ u_z(x, y, 0) = 0 \end{cases} \tag{IV.54}$$

– to enforce the periodicity condition, the outer boundaries $x = p/2$ and $z = p/2$ are free to move, but constrained to remain planar (rigid face constraint):

$$\begin{cases} u_x(p/2, y, z) = C_1 y + C_2 z + C_3 \\ u_z(x, y, p/2) = C_4 x + C_5 y + C_6 \end{cases} \tag{IV.55}$$

where C_1 to C_6 are real constants.

Hexahedral elements with linear interpolation are used for the mesh, depicted in Figure IV.20. The whole substrate thickness and detailed structure of the stack are accounted for, including thin SiN or SiCN barriers present between the different IMD or ILD layers, as well as TaN/Ta liners separating metal lines from dielectric layers.

The methodology used to simulate the sequential layer deposition process is as follows (Figure IV.26):

Layer deposition : the elements corresponding to the films to be deposited have their stiffness initially suppressed using the so-called *birth-and-death capability* in ANSYS. Then, for each deposited layer the associated elements are re-activated, while a homogeneous temperature load corresponding to the layer deposition temperature is prescribed on the stack (Figure IV.22). The intrinsic stresses present at the deposition temperature, which depend on the deposition process, are accounted for by prescribing a homogeneous equibiaxial initial stress in the elements associated with the deposited layer.

Damascene process : interconnect processing is simulated by first suppressing the stiffness of the elements associated to the patterned regions in the deposited blanket dielectric film (etching step), and then assigning updated mechanical properties for these elements (metal deposition). Metal overburden removal by CMP is not accounted for, as it is assumed not to affect the overall residual stresses significantly [43].

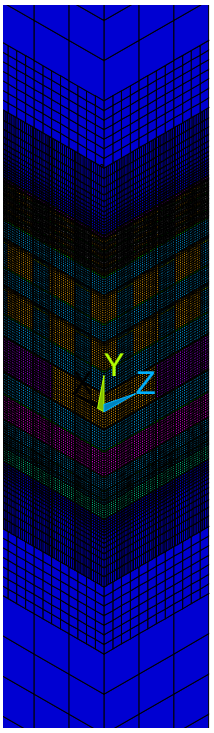


FIGURE IV.20: Mesh used in the finite element analysis for the top and bottom tiers of the 3D integrated circuit (the whole substrate thickness is accounted for in the model, but not represented here).

Wafer bonding : to mimic the bonding step, first the process-induced stresses in the BEoL structure during interconnect processing are computed in two separate models corresponding to the two tiers of the 3DIC. Then, the obtained values are incorporated as initial stresses in the undeformed bonded assembly.

Wafer thinning : wafer backgrinding is accounted for by suppressing the substrate elements associated with the removed material. Here, the final thicknesses for the top and bottom tier are respectively 3 µm and 120 µm.

All materials are considered linear elastic isotropic and the temperature-dependence of the mechanical properties is not considered (Table IV.2), corresponding. Any stress relaxation after deposition due to gas desorption, defect annihilation or recrystallization during the different thermal cycles [44–46] could not be characterized experimentally, and thus was not accounted for.

The intrinsic stresses present at the deposition temperature are estimated from measurements of the residual stresses at room temperature after film deposition, carried out routinely on the fabrication line for process monitoring (Table IV.3):

$$\sigma_I = \sigma_R - \frac{E_f \Delta\alpha \Delta T}{1 - \nu_f} \tag{IV.56}$$

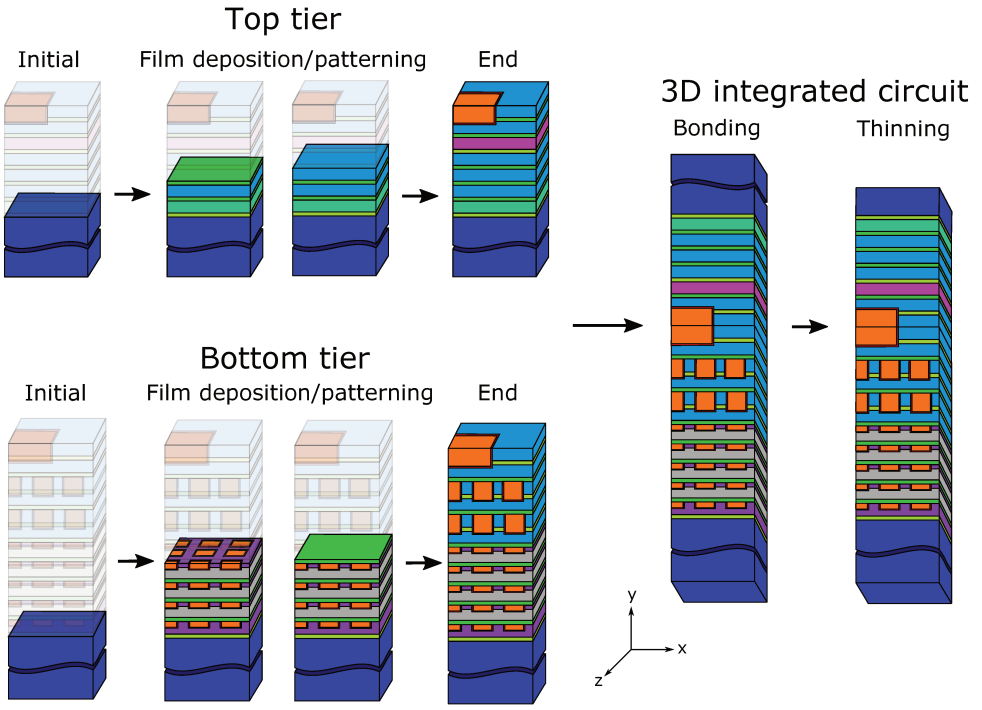
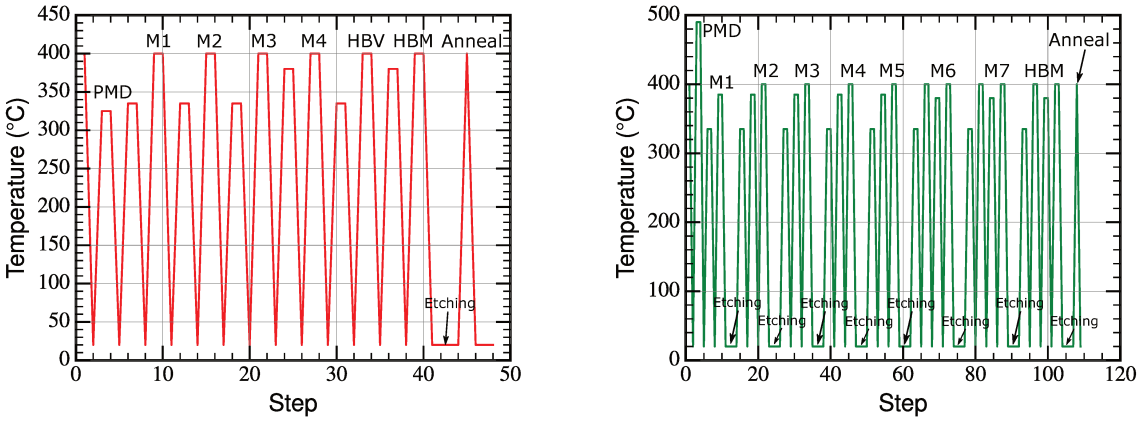


FIGURE IV.21: Schematical description of the different steps in the finite element analysis to account for the whole front-end process sequence.



(a) Top tier

(b) Bottom tier

FIGURE IV.22: Prescribed temperature profiles during the layer deposition sequence for the top (a) and bottom tier (b).

with σ_I and σ_R, respectively the intrinsic stress at the deposition temperature T_0 and residual stress at the current temperature T, E_f and ν_f the film elastic properties, $\Delta\alpha = \alpha_s - \alpha_f$ the CTE mismatch between the substrate and the film, and $\Delta T = T - T_0$ the temperature variation. These residual stresses are obtained indirectly, through wafer curvature measurements by laser scanning on a three-point support before and after deposition. Applying Stoney's equation for thin films on a thick substrate, the measured wafer curvature

TABLE IV.2: Thermoelastic properties used for the different layers of the interconnect stack.

Type	**Material**	**Process**	E (GPa)	ν	α ($10^{-6}\,K^{-1}$)
Passivation	SiN	PECVD	164[a]	0.24[a]	2.2[b]
	SiCN		67[a]	0.19[a]	4.4[b]
Dielectric	SiO_2 (SiH_4)	PECVD	60[c]	0.25[c]	2.67[c]
	SiO_2 (TEOS)	PECVD SACVD[e]	63[a]	0.23[a]	0.5[d]
	PSG	PECVD	52[c]	0.3[c]	1.45[c]
	FSG	PECVD[f]	51[g]	0.19[g]	3[g]
	SiOC:H	PECVD	7.8[a]	0.15[a]	10.8[b]
Metal	TaN/Ta	PVD	130[h]	0.25[h]	6.5[i]
	Cu	ECD	130[j]	0.34[j]	16.5[k]
Semiconductor	Si <100>		130[l]	0.28[l]	2.6[m]

[a] [47] [b] [43] [c] [48] [d] [49]

[e] Unknown, but assumed equivalent to PECVD.

[f] Known only for HDPCVD, but assumed equivalent for PECVD. [g] [50]

[h] [51] [i] [52] [j] [53]

[k] [54] [l] [55] [m] [56]

TABLE IV.3: Residual stresses at 25 °C for the different layers of the top and bottom tier interconnect stack used to evaluate intrinsic stresses generated during the deposition process.

Material	**Process**	T_{dep} (°C)	**Layer**	σ_R **(at 20 °C)** (MPa)
SiN	PECVD	380	IMDZ (bottom) HBM	-120
			IMD4 (top)	-180
		400	PMD	1500
SiCN	PECVD	335	IMD1-3 (top) IMDX IMDZ2 (bottom) HBV	-240
			IMDZ1 (bottom)	-350
SiO_2 (SiH_4)	PECVD	400		-190
SiO_2 (TEOS)	SACVD	490	PMD (bottom)	100
	PECVD	400	HBM	-150
			IMDX IMDZ (bottom)	-240
PSG	HDPCVD	325		-175
FSG	PECVD	400		-190[a]
SiOC:H	PECVD	385[b]		50
TaN/Ta	PVD	20		-1800[c]
Cu	ECD	20		600[c]

[a] Unknown, but assumed to be equivalent to SiO_2 with SiH_4 precursor.

[b] For SiOC:H, the UV curing temperature is considered as the equilibrium temperature.

[c] [44]

change $\Delta\kappa$ can be used to derive an equibiaxial average film stress σ_f [3]:

$$\sigma_f = \frac{E_s h_s^2}{6(1-\nu_s)h_f}\Delta\kappa \tag{IV.57}$$

where E_s, ν_s and h_s the Young's modulus, Poisson's ratio and thickness of the substrate, and h_f film thickness.

IV.3.1.2 Results

The average normal stress and inplane shear stress components at the surface of the Si substrate[1] across the periodic unit cell are plotted in Figure IV.23. Compared to the stress variations between the M2/PADOPEN steps estimated using the Si piezoresistive stress sensors in Section IV.2.4.3, the values obtained from finite element analysis are very low:

- In the initial state (after M2), the average stress exerted on the stress sensor is zero for all components.
- Then, during the remainder of interconnect processing, the average inplane normal stress components $\bar{\sigma}_{xx} = \bar{\sigma}_{<100>}$ and $\bar{\sigma}_{zz} = \bar{\sigma}_{<010>}$ increase gradually, but very modestly, up to about 4 MPa).
- After the HBM step, corresponding to the introduction of the first (and only) patterned layer in the top tier stack, a nonzero average inplane shear stress $\bar{\sigma}_{xz} = \bar{\tau}_{(100)}$ appears, not exceeding 5 MPa (in absolute value). The average out-of-plane normal stress $\bar{\sigma}_{yy} = \bar{\sigma}_{<001>}$ on the other hand remains zero throughout the process sequence.

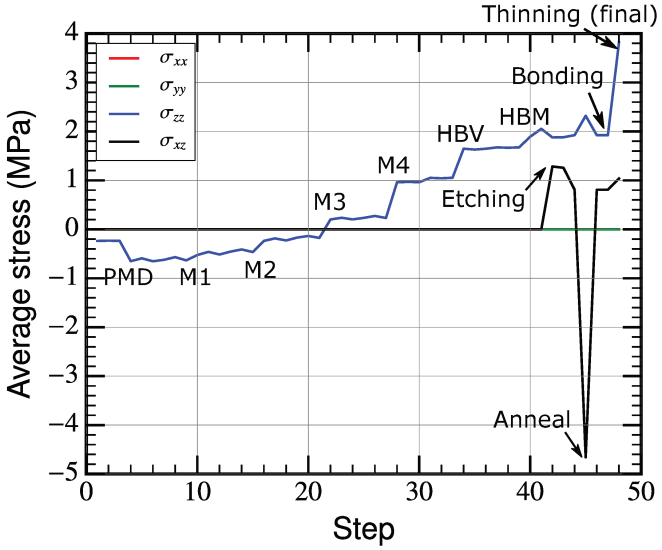
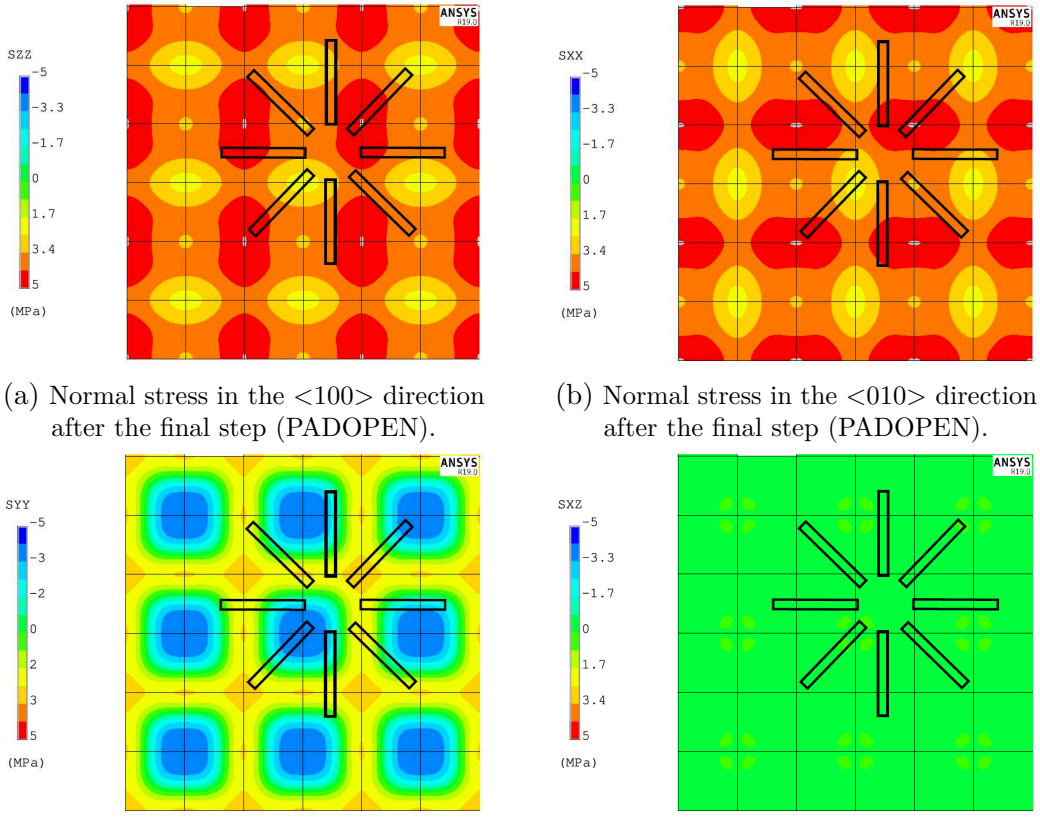


FIGURE IV.23: Computed average normal and inplane shear stress components in the active region of the Si substrate around a stress sensor located at the die center (top tier).

Contour plots of the normal and inplane shear stress components in top view at the surface of the Si substrate are plotted in Figure IV.24, at the end of front-end processing (PADOPEN step). The stress sensor footprint is superimposed on the stress contours (black rectangles), showing the presence of stress gradients across the sensor due to the

[1] A volume-averaging procedure is used across four elements over a thickness comparable to the implantation depth in the test chip, i.e. approximately 300 nm.

hybrid bonding metal pads. A side view of the whole stack at the end of the process is also shown in Figure IV.25. At the beginning of the process sequence, there is no stress at the surface of the Si substrate as the stress sensor structure itself and the associated interconnect layout were not accounted for in the model. This is evidenced by stress contour plots after the M2 step, not detailed here. At the end of the process sequence, after the PADOPEN step, the inplane normal stress components $\sigma_{xx} = \sigma_{<100>}$ and $\sigma_{zz} = \sigma_{<010>}$ reach maximum values of about 5 MPa, tensile, with slightly lower values right below the hybrid bonding pads. The shape of the hybrid bonding pads is clearly visible in the contour plot of the out-of-plane normal stress $\sigma_{yy} = \sigma_{<001>}$, with stresses reaching about 5 MPa, compressive, just below the metal patterns. The inplane shear stress $\sigma_{xz} = \tau_{(001)}$ on the other hand remains almost zero.



(a) Normal stress in the <100> direction after the final step (PADOPEN).

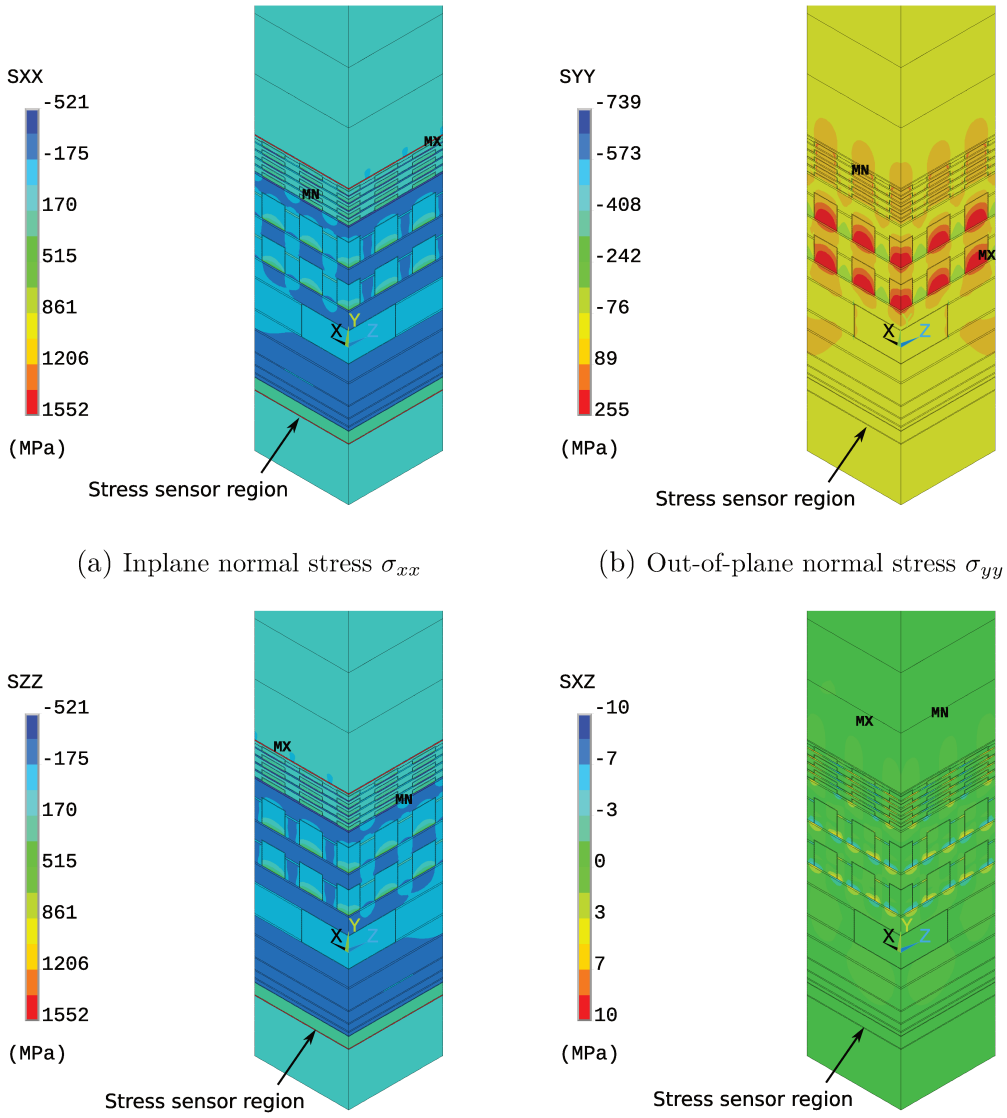
(b) Normal stress in the <010> direction after the final step (PADOPEN).

(c) Normal stress in the <001> direction after the final step (PADOPEN).

(d) Shear stress in the (001) plane after the final step (PADOPEN).

FIGURE IV.24: Contour plots of the normal and inplane shear stress components computed by the finite element method in the active region of the Si substrate around a stress sensor at the end of front-end processing.

The stress variations between the M2/PADOPEN steps obtained by simulation are far below those estimated using the Si piezoresistive stress sensors in Section IV.2.4.3. However, with the proposed method, the bending stresses induced by wafer curvature change during wafer bonding and substrate thinning are not accounted for. Instead, local stresses building up in the sensor region due to CTE mismatch in the interconnect stack are computed separately for the top and bottom tiers, and used as initial stress values in the undeformed



(a) Inplane normal stress σ_{xx}

(b) Out-of-plane normal stress σ_{yy}

(c) Inplane normal stress σ_{zz}

(d) Inplane shear stress σ_{xz}

FIGURE IV.25: Computed normal and inplane shear stress components in the interconnect stack of the assembled 3D integrated circuit after bonding and thinning, around a stress sensor located at the die center.

bonded assembly. These global stresses due to wafer deformation are very low (about 0.5 MPa, compressive), one order of magnitude below the local stresses due to local CTE mismatch obtained from the finite element analysis.

From finite element results, the process-induced stress variations between the M2/PADOPEN steps in the active region during to front-end processing are thus estimated to approximately 5 MPa for the maximum normal stresses ($\sigma_{11}, \sigma_{22}, \sigma_{33}$), and about 0 MPa for the inplane shear stress σ_{12}. These values are far below the stress variations estimated using our Si piezoresistive sensors. An additional observation from the numerical results is the presence of inplane stress gradients across the sensor due to the hybrid bonding metal

pads. The maximum amplitude for the stress distribution is obtained for the out-of-plane normal stress components σ_{33}, with approximately 10 MPa. For the inplane normal stress components $(\sigma_{11}, \sigma_{22})$, the stress amplitude is slightly lower, with about 5 MPa. The temperature-compensated stress component $\sigma_{11} - \sigma_{22}$, on the other hand, is almost zero across the whole surface due to the symmetries of the considered periodic unit cell.

In the next section, a finite element analysis of the die packaging step is carried out, aiming to derive the stress variations in the sensor regions during back-end processing.

IV.3.2 Back-end processing

IV.3.2.1 Model description

In the following analysis, the die to be encapsulated is considered stress-free in the initial state, since we are only interested in the stress variation due to packaging. Since the bending stresses in the top tier Si substrate due to wafer curvature were found in the previous section to be very low (on the order of 0.5 MPa), stress release during die singulation was not taken into account. Due to the large length scale difference, the interconnect stacks in the two-tier 3D integrated circuit are not accounted for. Instead, the die is considered as a homogeneous silicon bulk. For the same reason, metal patterning in the printed circuit board below the die is not detailed and instead an equivalent homogeneous orthotropic multilayer stack is used.

For the considered cavity package, the image sensor is not in direct contact with the epoxy molding compound. Therefore, the encapsulation stress originates for the most part from the CTE mismatch between the die, package substrate, and die-attach glue (with much larger CTE) [57]. Therefore, details of the die attach morphology such as the fillets present at the edge of the chip are accounted for in the analysis [58]. A thickness of 10 µm is assumed for this layer. On the other hand, although the detailed geometry of the package lid and glass cap sealing the chip is taken into account, the lid-attach and glass-attach materials are not modeled because they are not expected to have a significant influence on the encapsulation stress exerted on the chip. The same applies to the gold wires at the periphery of the chip, which does not significantly affect the stress state in the vicinity of the stress sensors. This is the case even at the die corner, since the associated is located sufficiently far away from the bonding pads.

The packaging process comprises many distinct steps, each associated with various processing temperatures (e.g. wire bonding, die-attach and lid-attach curing). For this analysis, the package is assumed stress-free at the die-attach curing temperature (150 °C), as the packaging-induced stresses are expected to originate mostly from the CTE mismatch between the chip, die-attach glue and package substrate. A homogeneous thermal load is

TABLE IV.4: Thermoelastic properties used for the different materials of the packaged integrated circuit.

Material	E (GPa)	ν	α ($10^{-6}\,K^{-1}$)	T_g (°C)
Si	130[a]	0.28[a]	2.5[b]	
Cu	130[c]	0.34[c]	16.5[d]	
Glass[e]	73.6	0.23	3.17	
Solder mask[e]	2.4	0.29	60	100
Prepreg/Core[e,f]	in-plane: 32 out-of-plane: 14	in-plane: 0.11 out-of-plane: 0.39	in-plane: 13 out-of-plane: 25	200-240
Die attach[e]	0.3	0.4	174	-31
Lid[f]	20 °C: 13.1 120 °C: 6.9 200 °C: 4.2	0.4[g]	in-plane: 4 out-of-plane: 87	120

[a] [55] [b] [56] [c] [53]
[d] [54] [e] Available from manufacturer technical datasheet.
[f] [59] [g] Assumed value.

prescribed on the assembly to simulate cooling down to room temperature from the initial stress-free state at 150 °C.

Linear elastic behavior is assumed for all the materials in the analysis (Table IV.4). Mechanical properties for the package materials were obtained from manufacturer datasheets. For the die-attach and package lid, and for most layers of the package substrate ("pre-preg"[1] sheets and core), the glass transition is taken into account, except for the solder mask layer for which this data was not available. Furthermore, for all of the latter materials except the die-attach, transverse orthotropic behavior is assumed. In this static analysis, stress relaxation of the different materials is not taken into account.

Owing to the quarter symmetry of the problem, only one quadrant is simulated. Both hexahedral and tetrahedral elements with quadratic interpolation are used to mesh the structure (Figure IV.26). To mimic the boundary conditions during the resistance measurements in the experimental study, wherein the packaged chip in connected to a DIB, nodes in regions corresponding to landing pads for electrical contact on the bottom surface of the package substrate are pinned.

IV.3.2.2 Results

The computed inplane normal and shear stress components in the top tier of the packaged die are plotted in Figure IV.27. The out-of-plane stress components on the other hand are not considered. Indeed, the packaged image sensor is not immersed in molding compound, but encapsulated in a cavity package. As described in Section IV.3.1, the top chip has been thinned down to 3 µm during the final step, therefore the active region where the stress

[1] Pre-impregnated

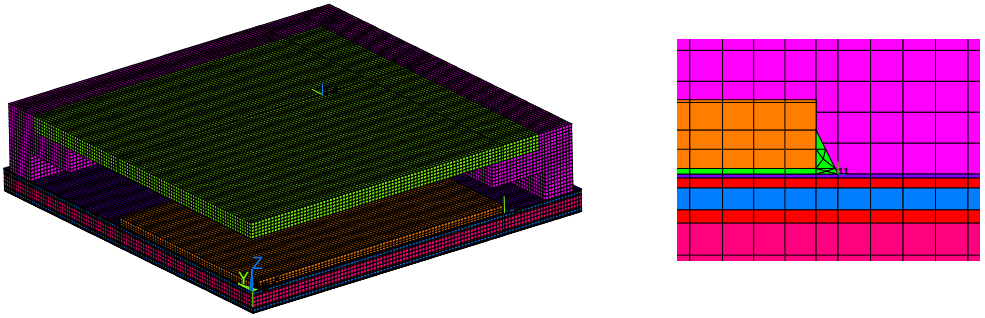


FIGURE IV.26: Mesh used in the finite element analysis of thermomechanical stresses induced by chip encapsulation (quarter-model), with a close-up view of the mesh details near the die-attach fillet.

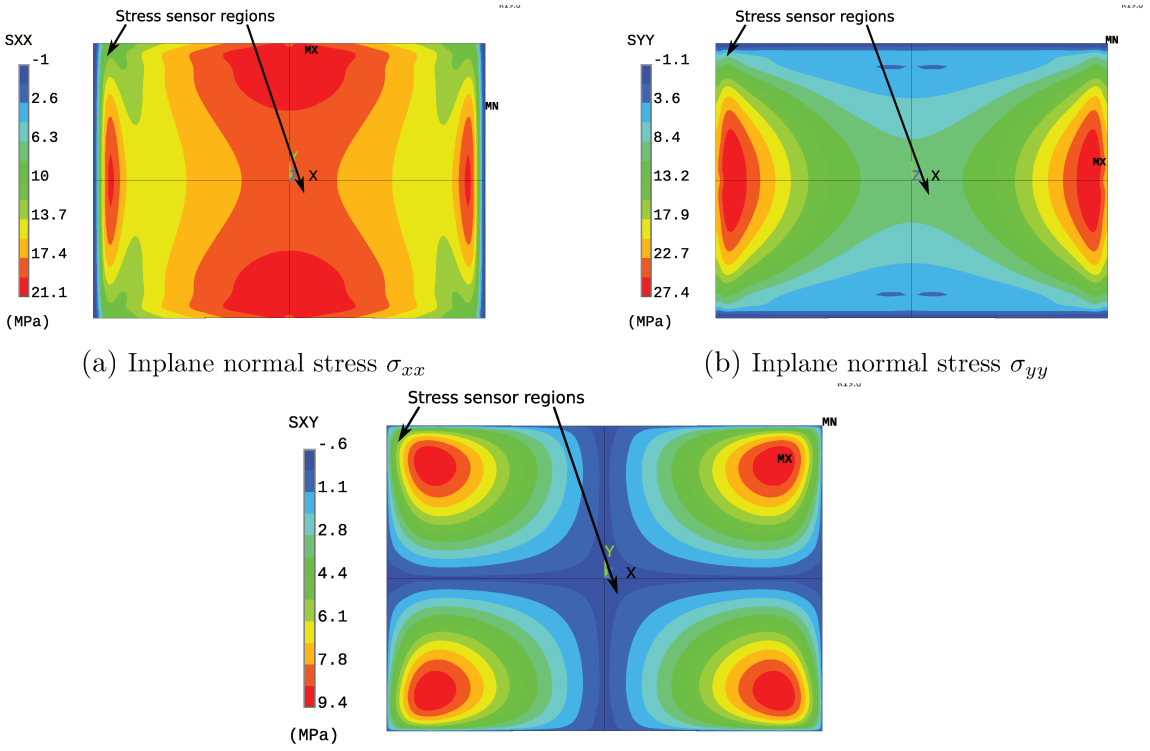
sensors are located is very close to a free surface and the out-of-plane stress components vanish.

The locations of maximum normal stress are found near the edges of the chip, where the highest magnitude is obtained along the width of the rectangular chip, reaching approximately 30 MPa, tensile. Computing the normal stress difference $\sigma_{yy} - \sigma_{xx}$, magnitudes of about -10 MPa and -1 MPa are obtained, respectively for the center and corners of the die. The maximum inplane shear stress on the other hand is located near the corners with values of about 10 MPa, whereas the stress state is equibiaxial at the center of the die, with almost zero shear.

Comparing the numerical results to the estimated stress variations using the Si piezoresistive stress sensors in Section IV.2.4, the computed individual stress components are far below those obtained in the experimental study, by as much as two order of magnitudes for the normal stresses and one for the inplane shear stress (Figure IV.13b). Using the temperature-compensated stress components however (Figure **??**), consistent sign and orders of magnitude are obtained at the die center, whereas at the die corner the computed stress values are an order of magnitude below the experimental values.

A contour plot of the out-of-plane displacements for the packaged die after cooling from 150 °C down to room temperature is shown in Figure IV.22. Due to the use of a cavity package for the image sensor chip, most of the deformation induced by the large CTE mismatch between the different parts is actually taken up by the glass cap, which by acting as a stiffener at the center of the molding compound contributes to mitigate chip warpage. This specific feature of the considered package could explain the relatively low values computed for the packaging-induced stress.

It can be noted that the stresses obtained by finite element modeling are quite low, below 5 MPa for front-end processing and not exceeding 30 MPa for back-end processing.



(a) Inplane normal stress σ_{xx}

(b) Inplane normal stress σ_{yy}

(c) Inplane shear stress σ_{xy}

FIGURE IV.27: Computed inplane normal and shear stress components in the thinned top tier of the packaged die.

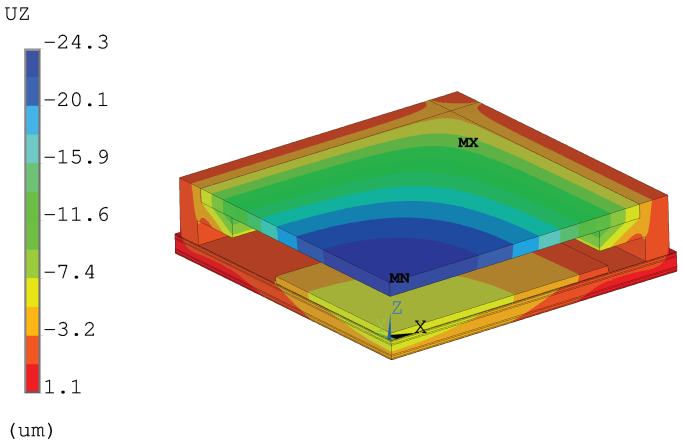


FIGURE IV.28: Computed out-of-plane displacements in the package after cooling from 150 °C to room temperature.

These low stress values seem to indicate a modest influence of the 3D BSI image sensor configuration on pixel operation, from a thermo-mechanical perspective.

The numerical study also confirms that the stress variations estimated using our Si piezoresistive sensors are greatly overestimated. In the following discussion, possible reasons that could explain this mismatch are discussed. Since linear elastic behavior was assumed for

all materials in the finite element analysis, the obtained results are assumed to be quite conservative and thus the focus is put on the stress sensing approach.

IV.4 Discussion

The normal stress variations between the M2 and PADOPEN steps for components σ_{11}, σ_{22} and σ_{33}, estimated using the piezoresistive stress sensors are very large compared to the typical stresses expected during the front-end process steps, especially the out-of-plane normal stress, with up to 300 MPa variation [see, e.g. 43, 60–62]. This discrepancy was confirmed by a finite element analysis of the sequential layer deposition process during front-end manufacturing, for which very low magnitudes were obtained (below 10 MPa). It can be noted that simulation results reveal the presence of an inplane stress gradient across the sensor region due to the resistors spanning laterally over hybrid bonding interconnects.

After die assembly, unrealistically large values are obtained for the normal stress components σ_{11}, σ_{22} and σ_{33}, on the order of 2–3 GPa, much above the expected magnitudes for the packaging process [25, 31, 63], above even the levels of stress that can be taken up by the silicon single-crystal. Again, finite element analysis of the packaging step was carried out, leading to rather low stress magnitudes (below 30 MPa) at the sensor locations. This is most likely due to the specific type of package used here to encapsulate the image sensor, a cavity package, for which the die is not in direct contact with molding compound.

In the following paragraphs, aiming to better understand this mismatch between stress sensing and finite element analysis, the main sources of error associated with this Si piezoresistive stress sensors are discussed, namely:

- the influence of temperature;
- stress sensor sensitivity;
- the measurement procedure.

IV.4.1 Influence of temperature

Among thermal errors, a first possible source is the temperature-dependence of the piezoresistive coefficients, which were considered constant in the analysis. However, as mentioned in IV.2.4.2.3, an experimental study by [34] showed that the temperature-dependence of the piezoresistivity coefficients vanishes with increasing doping density and becomes almost inexistent for impurity concentrations on the order of $1 \times 10^{21}\,\mathrm{cm}^{-3}$, as is the case for the resistors considered here. Therefore, the piezoresistive coefficients for both N- and P-doping should be virtually independent of temperature.

A second kind of thermal error for stress sensing is the temperature dependence of the resistance values through the temperature coefficients of resistance [41]. For each of the twelve possible dual-polarity four-element sensors, the same additional temperature-dependent terms appear in the expressions for the normal stresses (but not the shear stress). For our sensors, a temperature variation of 1 °C between during sensor calibration or between inline resistance measurements is estimated to lead to spurious stress variations comprised between 90 MPa and 110 MPa (in absolute value). These values lie well within the expected magnitude for the process-induced stress variations. Therefore, for accurate stress sensing it is critical to assess the temperature-dependent behavior of the resistors.

Temperature variations between inline resistance measurements may arise from various causes:

- external environment, for instance through fluctuations of the local temperature in the clean room;
- self-heating of the stress sensor due the Joule effect under prescribed current during the resistance measurement.

As mentioned in Section IV.2.3.2.1, the air temperature in the clean room is controlled to ±0.5 °C, and thus fluctuations of 1 °C are indeed possible. Resistor self-heating on the other hand depends on the magnitude of the electrical current used to measure the resistance values. In case of a sufficiently large difference between these currents during the different measurements, temperature-dependent resistance variations could arise. Dedicated electrical measurements would be necessary to assess quantitatively the relative importance of the latter effect [5, 64].

During front-end processing, between the M2/PADOPEN steps, the same measurements conditions and equipment were used. Therefore, temperature-induced stress variations are most likely due to the external environment. Temperature fluctuations of 1 °C during inline resistance measurements, leading to spurious stresses of approximately 100 MPa, are compatible with the magnitude range of the estimated stress variations, on the order of $200 - 300$ MPa.

During back-end processing, between the PADOPEN/PKG steps, different measurement conditions and equipment are used however. The obtained temperature-induced stress variations thus cannot be attributed solely to external temperature fluctuations. The difference in the current intensities used in the resistance measurements before and after packaging could be an additional contributor, through the Joule effect. However, it seems unlikely that such low current variations could cause sufficiently large temperature variations, inducing spurious stress variations on the order of the GPa. Thus, other sources of error must be involved, which are discussed in the next paragraphs.

To avoid parasitic stress variations due to temperature-induced resistance changes, several possibilities may be considered:

- design sensors with piezoresistors having low values of TCRs; the existence of a minimum for the TCRs was evidenced for dopant concentrations around 1×10^{18}–$1 \times 10^{19}\,\mathrm{cm}^{-3}$ for N-doped and P-doped Si resistors [65–68].

- monitor the local temperature in the sensor region during the resistance measurements, enabling to quantify the temperature-dependent terms and correct for their impact in the estimation of the stress components.

- consider only temperature-independent stress components, namely the inplane shear stress σ_{12}, which is intrinsically insensitive to temperature effects, and the inplane normal stress difference $\sigma_{11} - \sigma_{22}$, for which the temperature-dependent terms cancel out [25, 40].

The solution that was adopted here is to work with the temperature-independent stress components. Indeed, among the latter three approaches, that is the only one not requiring to design a new test chip. It can be noted however that not all sensor configurations among the twelve possibilities for dual-polarity rosettes are equivalent in that respect. The (2P2N) rosette type offers several advantages for temperature compensation. In this configuration, the temperature-compensated stress components involve only three out of six independent piezoresistive coefficients, enabling less time-consuming sensor calibration or to obtain more data for the same experimental effort. Among the two possible (2P2N) sensor configurations, that with the N-type resistors aligned with the <100> directions (resp. the P-type resistors with the <110> directions), leads to the best accuracy for stress estimation [40]. Indeed, in this configuration the three piezoresistive coefficients involved are the largest, i.e. those calibrated with the best accuracy. Incidentally, this configuration is also optimal in terms of sensitivity to rosette rotational misalignment [41]. A comparison between the two possible (2P2N) rosette combinations confirmed the optimal character of that particular rosette.

Using temperature-compensated stress components for stress monitoring requires decreasing the number of accessible stress components from four to two, for a (100) Si substrate. Thus, it is recommended to optimize Si piezoresistor doping concentrations to decrease the temperature dependence. If temperature-induced stresses are still non-negligible compared to the expected process-induced stress variations, then implementing Si diode temperature sensors near the stress sensors may be considered in order to quantify the temperature dependence [69, 70].

IV.4.2 Stress sensor sensitivity

IV.4.2.1 Piezoresistive coefficients magnitude

As discussed in Section IV.2.4.2.3, the piezoresistive coefficients evaluated in this work are more than twice lower than typical values for stress sensor applications in the literature. Therefore, piezoresistors in this study exhibit rather low stress sensitivity, between 0.05% and 5% for 100 MPa, compared to between 1% and 10% in the literature [see, e.g. 25, 71]. This is due to the high doping concentrations used, about $10^{20}\,\mathrm{cm}^{-3}$ compared to 10^{17}-$10^{18}\,\mathrm{cm}^{-3}$ in the literature. Indeed, piezoresistive coefficients were shown to decrease sharply in heavily doped silicon for increasing doping concentrations [34, 35, 72]. The high doping levels used in this study were set beforehand by transistor design rules for the considered test chip.

Due to the low stress sensitivity of the sensors, the piezoresistive coefficients with the lowest values π_{11}^p and π_{12}^p could not be evaluated with sufficient accuracy. Indeed, measurement uncertainties for these coefficients were as high as 60% and 20%, respectively. These coefficients appear in the expressions of the three normal stress components and therefore could lead to large errors in the estimation of the process-induced stress variations. That is the case for all twelve possible dual-polarity sensor configurations. However, since in our case the temperature-independent stress components must be used, this issue can be avoided by selecting the rosette with N-type resistors aligned with the <100> directions and P-type resistors along the <110> directions. Indeed, this configuration yields expressions that are independent of π_{11} and π_{12}^p.

Due to the limitations brought by temperature compensation (only two accessible independent stress components instead of four), it is recommended to use moderate doping concentrations, in the $10^{17} - 10^{18}\,\mathrm{cm}^{-3}$ range, to enable improved stress sensitivity and accurate evaluation of all piezoresistive coefficients. It can also be noted that the N-type piezoresistive coefficients are less dissimilar are larger compared to the P-type coefficients. Among the six independent piezoresistive coefficients, π_{11}^p and π_{12}^p are the smallest and π_{44}^p the largest, whereas π_{11}^n, π_{12}^n and π_{44}^n are all within the same intermediate range (in absolute value). Therefore, an attractive possibility would be not only to decrease doping levels, but also to use a single-polarity stress sensor based on N resistors only, as proposed by [73]. By doing so, coefficients π_{11}^p and π_{12}^p in the expressions of the stress components can be replaced by π_{11}^n and π_{12}^n, having larger magnitude and thus calibrated more accurately. For the governing system of equations to remain invertible, two groups of resistors with distinct N-type piezoresistive coefficients[1] would then have to be introduced.

Decreasing piezoresistor doping levels could also help mitigating the influence of temperature. Indeed, it was shown for heavily doped resistors[2] to lead to decrease of the TCRs

[1] This can be achieved for instance by using distinct doping concentrations or different doping atoms.
[2] In the $10^{20}\,\mathrm{cm}^{-3}$ range

[67, 68]. However, if the doping levels become too low, typically below $10^{18}\,cm^{-3}$, the TCRs begin to increase again [67]. Thus, a tradeoff should be found for moderate doping levels, between sufficiently increasing the piezoresistive coefficients to avoid large errors in the estimation of process-induced stress variations and minimizing their temperature dependence. It must also be noted that decreasing the doping levels leads to an increased temperature dependence of the piezoresistive coefficients [34].

IV.4.2.2 Piezoresistive coefficients variability

Another important aspect regarding the robustness of the piezoresistive stress sensor approach is the variability of the piezoresistive coefficients between wafers (*inter-wafer*), but also across the wafer surface (*intra-wafer*). Indeed, piezoresistive coefficient calibration is a time-consuming task and therefore, although ideally each stress sensor should be calibrated individually, the piezoresistive coefficients are typically measured on a few sensors and then used for stress estimation with different sensors, on different wafers.

In this study, limited information on the spatial variability of the piezoresistive coefficients can be obtained using the instrumented four-point bending calibration, which is a destructive technique. Therefore, the distribution of the resistance values is used as an alternative indicator, assumed to reflect at least in part the variability of the doping process, and thus of the piezoresistive coefficients. Indeed, from measurements in the literature it is expected that the piezoresistive coefficients are most sensitive to dopant concentration variations for heavily doped resistors [34, 35, 67, 72].

Preliminary resistance measurements at the wafer level at the beginning of the process sequence revealed low inter-wafer variability (below 5%). Intra-wafer variability, on the other hand, is quite substantial, with relative variations of about 20% for P-type resistors, and about 8% for N-type resistors. Moderate resistance variations (below 10%) were obtained within the eight resistors in each sensor. From this, several assumptions can be made:

- because inter-wafer variability is very low, the fact that the stress sensors were calibrated on only one wafer is not deemed prejudicial;
- intra-wafer variability is less for N-type resistors compared to the P type. This could be a further argument for the use of single-polarity sensors, comprising only N-type resistors, should it be established that the observed intra-wafer variability of the resistance values is indeed associated with substantial variability of the piezoresistive coefficients;
- inter-lot variability could not be assessed in this study. Further measurements would be needed to determine whether sensor calibration is necessary for each lot;

Several possibilities may be considered to optimize the calibration procedure, and thereby obtain more data on the variability of the piezoresistive coefficients:

- an attractive approach is hydrostatic loading calibration [18], enabling to determine the piezoresistive coefficients simultaneously at the wafer-scale, with a nondestructive procedure. However, this method involves the development of a "modified vacuum chuck" test fixture to prescribe a uniform pressure on the wafer, and finite element analysis is required to derive the prescribed stress, accounting for large deflections.
- another approach, relying on the more conventional four-point bending technique, is to introduce rosettes oriented at 22.5° with respect to the principal crystallographic directions of the substrate [74]. With this configuration, a single uniaxial load on samples diced in the same direction is then sufficient to obtain the two independent piezoresistive coefficients required for temperature-independent stress evaluation.

IV.4.3 Measurement procedure

Several differences can be noted between inline resistance measurements carried at the wafer-level and the package-level, which could further contribute to the large difference in the magnitude of the estimated stress components, with unrealistically large values obtained after packaging. During front-end processing, the measurements were performed each time under the same conditions, whereas after the back-end process they were carried out in a different environment, on a distinct equipment, and with lower values for the prescribed currents (200 µA instead of 500 µA for P-type resistors).

The temperature-compensated stresses evaluated after the packaging step (at the die center) are in reasonable agreement with finite element results compared to individual stress components, it may thus be inferred that the very large values obtained for the latter are mostly due to local or external temperature fluctuations. It can be noted however that surprisingly high stresses are still obtained for the sensors located at the corner of the die, even using the temperature-compensated components. Other monitoring structures comprising MOSFETs located near the sensors, namely *ring oscillators*, also yielded unexpectedly large values at the die corner, with oscillation frequencies more than twice those measured at the die center. At the time of redaction of this manuscript, electrical failure analysis is currently underway to investigate the possible causes for these outliers at the corner of the die after packaging.

Several hypotheses can be formulated to explain the discrepancies observed after packaging. In addition to measurements errors, the fact that another equipment was used could generate a shift in the values, due to a different calibration compared to the equipment used at the wafer-level. In addition, the sample was mounted on a device interface board for the resistance measurements, which could exert additional stresses on the die. It must

be noted however that [29] carried out stress sensing experiments on a packaged die (using CMOS sensors) in very similar conditions, but did not encounter such issues.

IV.5 Conclusions

In this chapter, a piezoresistive stress sensor based on doped single-crystal Si resistors was developed to monitor the process-induced stress variations from inline wafer-level and die-level resistance measurements on a hybrid bonded 3D integrated circuit.

After an overview of the general equations of the piezoresistivity theory, applied to the case of a (100) Si substrate, the stress-resistance relationships for dual-polarity four-element rosettes were detailed and the twelve possible configurations compared. Eight-element stress sensors were implemented at several locations on the chip (center and corner) and repeated on four wafers. The calibration procedure for determining the piezoresistive coefficients of the resistors making up the sensors using an in-house instrumented four-point bending fixture was then detailed. Although the coefficients are consistent with the trends observed in the literature, their magnitude is quite low, owing to high impurity concentrations during the ion implantation process used to process the sensors. The variations of the normal stress and inplane shear stress components were evaluated at the wafer-level during die processing and after chip packaging.

In both cases, the obtained stress values are quite large compared to the values reported in the literature for comparable processes. This was confirmed by finite element modeling of the involved processes, namely sequential layer deposition and chip packaging. A potential effect of the temperature-dependence of the resistance value was identified, due to the high doping levels used to process the stress sensors (based on the design rules for the test chip), with estimated stress variations comparable to the typical process-induced stress values expected to arise for very low temperature fluctuations (about 50 MPa in absolute value for 0.5 °C). This assumption led us to work with temperature-compensated stress components, namely the inplane normal stress difference, for which the temperature terms cancel out, and the inplane shear stress, which is intrinsically insensitive to temperature variations. These temperature-compensated stress components are in better agreement with the simulations results, although some unexpectedly high values remained (especially at the die corner after packaging), for which further investigations are ongoing. Two approaches for stress evaluation, namely the least squares method and the direct inversion method were compared. Although both yield the same results for the determination of the temperature-compensated stress components, this is not the case with the individual stress components.

The low stress values (below 30 MPa) obtained in finite element analyses of die processing and packaging tend to indicate a modest influence of the 3D BSI image sensor configuration on pixel operation, from a thermo-mechanical perspective. Although this could not be

achieved here, due to the limitations discussed above, a correlation of these simulation with *in-situ* stress sensing experiments is needed to confirm this finding. The possible sources of errors for the developed stress sensor have been discussed, based on the results of this study and the literature. A few recommendations can thus be proposed for future work on in situ monitoring of processed-induced stress by inline resistance variation measurements using Si piezoresistive sensors:

- Lower doping levels are recommended, in order to increase the magnitude of the piezoresistive coefficients and thus the sensitivity of the sensors to stress variations.
- To assess the variability of the piezoresistive coefficients, a wafer-scale calibration technique such as hydrostatic loading using a modified vacuum chuck could provide a great advantage [18]. If instead four-point bending is used, as in the present study, another interesting approach would be to introduce rosettes oriented at 22.5° with respect to the <100> orientation, enabling to evaluate in a single loading sequence the two independent piezoresistive coefficients needed to obtain the temperature-independent stress components.
- To assess the temperature dependence of the implemented stress sensors, the temperature coefficients of resistance should be determined systematically before calibration and inline stress measurements.
- If temperature-compensated stress components are used, for improved accuracy the dual-polarity rosette configuration with N-type resistors aligned with the <100> directions of the (100) Si substrate and P-type resistors along the <110> directions is recommended [41]. Another interesting approach could be to use single-polarity rosettes with two families of N-type resistors, each having a distinct set of piezoresistive coefficients [73].

References

[1] O. Kraft, M. Hommel, and E. Arzt, "X-ray diffraction as a tool to study the mechanical behaviour of thin films," Materials Science and Engineering: A, vol. 288, no. 2, pp. 209–216, 2000.

[2] I. De Wolf, "Micro-raman spectroscopy to study local mechanical stress in silicon integrated circuits," Semiconductor Science and Technology, vol. 11, no. 2, p. 139, 1996.

[3] G. G. Stoney, "The tension of metallic films deposited by electrolysis," Proc. R. Soc. Lond. A, vol. 82, no. 553, pp. 172–175, 1909.

[4] G. C. Dogiamis, B. J. Hosticka, and A. Grabmaier, "Investigations on an ultra-thin bendable monolithic si cmos image sensor," IEEE Sensors Journal, vol. 13, no. 10, pp. 3892–3900, 2013.

[5] B. Vianne, P. Bar, V. Fiori, S. Gallois-Garreignot, K. A. Ewuame, P. Chausse, S. Escoubas, N. Hotellier, and O. Thomas, "Thermo-mechanical characterization of passive stress sensors in si interposer," Microelectronics Reliability, vol. 55, no. 5, pp. 738–746, 2015.

[6] J. Lau, Thermal stress and strain in microelectronics packaging. Springer Science & Business Media, 2012.

[7] J. L. Spencer, W. F. Schroen, G. A. Bednarz, J. A. Bryan, T. D. Metzgar, R. D. Cleveland, and D. R. Edwards, "New quantitative measurements of ic stress introduced by plastic packages," in Reliability Physics Symposium, 1981. 19th Annual, pp. 74–80, IEEE, 1981.

[8] D. Edwards, G. Heinen, G. Bednarz, and W. Schroen, "Test structure methodology of ic package material characterization," IEEE Transactions on Components, Hybrids, and Manufacturing Technology, vol. 6, no. 4, pp. 560–567, 1983.

[9] C. S. Smith, "Piezoresistance effect in germanium and silicon," Physical review, vol. 94, no. 1, p. 42, 1954.

[10] D. A. Bittle, J. C. Suhling, R. E. Beaty, R. C. Jaeger, and R. W. Johnson, "Piezoresistive stress sensors for structural analysis of electronic packages," Journal of Electronic Packaging, vol. 113, no. 3, pp. 203–215, 1991.

[11] P. W. Bridgman, "The effect of tension on the electrical resistance of certain abnormal metals," in Proceedings of the American Academy of Arts and Sciences, vol. 57, pp. 41–66, JSTOR, 1922.

[12] P. W. Bridgman, "The effect of tension on the transverse and longitudinal resistance of metals," in Proceedings of the American Academy of Arts and Sciences, vol. 60, pp. 423–449, JSTOR, 1925.

[13] P. W. Bridgman, "The effect of pressure on the electrical resistance of single metal crystals at low temperature," in Proceedings of the American Academy of Arts and Sciences, vol. 68, pp. 95–123, JSTOR, 1933.

[14] P. W. Bridgman, "The effect of pressure on the electrical resistance of certain semiconductors," in Proceedings of the American Academy of Arts and Sciences, vol. 79, pp. 127–148, JSTOR, 1951.

[15] W. Paul and G. L. Pearson, "Pressure dependence of the resistivity of silicon," Physical Review, vol. 98, no. 6, p. 1755, 1955.

[16] W. G. Pfann and R. N. Thurston, "Semiconducting stress transducers utilizing the transverse and shear piezoresistance effects," Journal of Applied Physics, vol. 32, no. 10, pp. 2008–2019, 1961.

[17] M. Doelle, C. Peters, P. Ruther, and O. Paul, "Piezo-fet stress-sensor arrays for wirebonding characterization," Journal of Microelectromechanical systems, vol. 15, no. 1, pp. 120–130, 2006.

[18] J. Suhling, R. A. Cordes, Y. L. Kang, and R. C. Jaeger, "Wafer-level calibration of stress sensing test chips," in Electronic Components and Technology Conference, 1994. Proceedings., 44th, pp. 1058–1070, IEEE, 1994.

[19] C.-H. Cho, R. C. Jaeger, J. C. Suhling, Y. Kang, and A. Mian, "Characterization of the temperature dependence of the pressure coefficients of n-and p-type silicon using hydrostatic testing," IEEE Sensors Journal, vol. 8, no. 4, pp. 392–400, 2008.

[20] P.-J. Tzeng, J. H. Lau, M.-J. Dai, S.-T. Wu, H.-C. Chien, Y.-L. Chao, C.-C. Chen, S.-C. Chen, C.-Y. Wu, C.-K. Lee, et al., "Design, fabrication, and calibration of stress sensors embedded in a tsv interposer in a 300mm wafer," in Electronic Components and Technology Conference (ECTC), 2012 IEEE 62nd, pp. 1731–1737, IEEE, 2012.

[21] R. E. Beaty, R. C. Jaeger, J. C. Suhling, R. W. Johnson, and R. D. Butler, "Evaluation of piezoresistive coefficient variation in silicon stress sensors using a four-point

bending test fixture," IEEE Transactions on components, hybrids, and manufacturing technology, vol. 15, no. 5, pp. 904–914, 1992.

[22] H. Miura, A. Nishimura, S. Kawai, and K. Nishi, "Development and application of the stress sensing test chip for ic plastic packages.," TRANS. JAPAN SOC. MECH. ENG.(SER. A)., vol. 493, no. 53, pp. 1826–1832, 1987.

[23] S. A. Gee, V. R. Akylas, and W. F. Van Den Bogert, "The design and calibration of a semiconductor strain gauge array," in Microelectronic Test Structures, 1988. ICMTS. Proceedings of the 1988 IEEE International Conference on, pp. 185–191, IEEE, 1988.

[24] R. E. Beaty, J. C. Suhling, C. A. Moody, D. A. Bittle, R. W. Johnson, R. D. Butler, and R. C. Jaeger, "Calibration considerations for piezoresistive-based stress sensors," in Electronic Components and Technology Conference, 1990.., 40th, pp. 797–806, IEEE, 1990.

[25] J. C. Suhling and R. C. Jaeger, "Silicon piezoresistive stress sensors and their application in electronic packaging," IEEE sensors journal, vol. 1, no. 1, pp. 14–30, 2001.

[26] J. Bartholomeyczik, S. Brugger, P. Ruther, and O. Paul, "Multidimensional cmos in-plane stress sensor," IEEE Sensors Journal, vol. 5, no. 5, pp. 872–882, 2005.

[27] K. Tian, Z. Wang, M. Zhang, and L. Liu, "Design, fabrication, and calibration of a piezoresistive stress sensor on soi wafers for electronic packaging applications," IEEE Transactions on Components and Packaging Technologies, vol. 32, no. 2, pp. 513–520, 2009.

[28] M. Gonzalez, B. Vandevelde, A. Ivankovic, V. Cherman, B. Debecker, M. Lofrano, I. De Wolf, G. Beyer, B. Swinnen, Z. Tokei, et al., "Chip package interaction (cpi): Thermo mechanical challenges in 3d technologies," in Electronics Packaging Technology Conference (EPTC), 2012 IEEE 14th, pp. 547–551, IEEE, 2012.

[29] K. A. Ewuame, V. Fiori, K. Inal, P.-O. Bouchard, S. Gallois-Garreignot, S. Lionti, C. Tavernier, and H. Jaouen, "Investigations of thermomechanical stress induced by TSV-middle (through-silicon via) in 3-D ICs by means of CMOS sensors and finite-element method," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 5, no. 8, pp. 1085–1092, 2015.

[30] B. Vianne, Intégration d'un interposeur actif silicium pour l'élaboration de circuits électroniques complexes. PhD thesis, Université Aix-Marseille, 2016. In French.

[31] K. A. Ewuame, Analyse Expérimentale et Numérique des Contraintes Thermomécaniques Induites lors des Procédés Émergents de Fabrication de Puces Électroniques au moyen des Capteurs Embarqués. PhD thesis, PSL Research University, 2016. In French.

[32] Keithley Instruments, Inc., Keithley 4200-SCS Semiconductor Characterization System QuickStart Manual. Keithley Instruments, Inc.

[33] E. Padonou, Statistical Learning on Circular Domains For Advanced Process Control in Microelectronics. PhD thesis, Université de Lyon, 2016.

[34] O. N. Tufte and E. L. Stelzer, "Piezoresistive properties of silicon diffused layers," Journal of Applied Physics, vol. 34, no. 2, pp. 313–318, 1963.

[35] O. N. Tufte and E. L. Stelzer, "Piezoresistive properties of heavily doped n-type silicon," Physical review, vol. 133, no. 6A, p. A1705, 1964.

[36] K. A. Ewuame. Personal communication. 2017-08-18.

[37] K. A. Ewuame, V. Fiori, K. Inal, P.-O. Bouchard, S. Gallois-Garreignot, S. Lionti, C. Tavernier, and H. Jaouen, "CMOS stress sensor for 3D integrated circuits: Thermo-mechanical effects of through silicon via (TSV) on surrounding silicon," in Thermal, mechanical and multi-physics simulation and experiments in microelectronics and microsystems (eurosime), 2014 15th international conference on, pp. 1–8, IEEE, 2014.

[38] K. A. Ewuame, V. Fiori, K. Inal, P. O. Bouchard, S. Gallois-Garreignot, S. Lionti, C. Tavernier, and H. Jaouen, "Investigation of tsv induced thermo-mechanical stress: Implementation of piezoresistive sensors and correlation with simulation," in Advanced Materials Research, vol. 996, pp. 975–981, Trans Tech Publ, 2014.

[39] W. Gander, M. J. Gander, and F. Kwok, Scientific computing-An introduction using Maple and MATLAB, vol. 11. Springer Science & Business, 2014.

[40] R. C. Jaeger, J. C. Suhling, and R. Ramani, "Errors associated with the design, calibration and application of piezoresistive stress sensors in (100) silicon," IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B, vol. 17, no. 1, pp. 97–107, 1994.

[41] R. C. Jaeger, J. C. Suhling, and A. A. Anderson, "A [100] silicon stress test chip with optimized piezoresistive sensor rosettes," in Electronic Components and Technology Conference, 1994. Proceedings., 44th, pp. 741–749, IEEE, 1994.

[42] D. Kerr and A. Milnes, "Piezoresistance of diffused layers in cubic semiconductors," Journal of Applied Physics, vol. 34, no. 4, pp. 727–731, 1963.

[43] N. Chérault, G. Carlotti, N. Casanova, P. Gergaud, C. Goldberg, and M. Thomas, O .and Verdier, "Mechanical characterization of low-k and barrier dielectric thin films," Microelectronic engineering, vol. 82, no. 3-4, pp. 368–373, 2005.

[44] N. Chérault, Caractérisation et modélisation thermomécanique des couches d'interconnexions dans les circuits sub-microélectroniques. PhD thesis, École Nationale Supérieure des Mines de Paris, 2006. In French.

[45] P. Morin, "Mechanical stress in silicon based materials: Evolution upon annealing and impact on devices performances," in Advanced Thermal Processing of Semiconductors, 2006. RTP'06. 14th IEEE International Conference on, pp. 93–102, IEEE, 2006.

[46] P. Morin, G. Raymond, D. Benoit, P. Maury, and R. Beneyton, "A comparison of the mechanical stability of silicon nitride films deposited with various techniques," Applied surface science, vol. 260, pp. 69–72, 2012.

[47] G. Carlotti, N. Cherault, N. Casanova, C. Goldberg, and G. Socino, "Elastic constants of low-k and barrier dielectric films measured by brillouin light scattering," Thin Solid Films, vol. 493, no. 1-2, pp. 175–178, 2005.

[48] L. Doucet, Caractérisation mécanique et physico-chimique des matériaux métalliques et diélectriques dans des structures d'interconnexions destinées à la microélectronique. PhD thesis, Institut National Polytechnique de Grenoble, 1996. In French.

[49] G. Carlotti, P. Colpani, D. Piccolo, S. Santucci, V. Senez, G. Socino, and L. Verdini, "Measurement of the elastic and viscoelastic properties of dielectric films used in microelectronics," Thin Solid Films, vol. 414, no. 1, pp. 99–104, 2002.

[50] Y.-B. Park and I.-S. Jeon, "Effects of mechanical stress at no current stressed area on electromigration reliability of multilevel interconnects," Microelectronic engineering, vol. 71, no. 1, pp. 76–89, 2004.

[51] M. Grégoire, Propriétés des couches minces et des interconnexions de cuivre. PhD thesis, Institut National Polytechnique de Grenoble, 2006. In French.

[52] G. V. Samsonov, "Mechanical properties of the elements," in Handbook of the Physicochemical Properties of the Elements, pp. 387–446, Springer, 1968.

[53] H. M. Ledbetter and E. R. Naimon, "Elastic properties of metals and alloys. ii. copper," Journal of physical and chemical reference data, vol. 3, no. 4, pp. 897–935, 1974.

[54] F. C. Nix and D. MacNair, "The thermal expansion of pure metals: copper, gold, aluminum, nickel, and iron," Physical Review, vol. 60, no. 8, p. 597, 1941.

[55] M. A. Hopcroft, W. D. Nix, and T. W. Kenny, "What is the young's modulus of silicon?," Journal of microelectromechanical systems, vol. 19, no. 2, pp. 229–238, 2010.

[56] A. Masolin, P.-O. Bouchard, R. Martini, and M. Bernacki, "Thermo-mechanical and fracture properties in single-crystal silicon," Journal of Materials Science, vol. 48, no. 3, pp. 979–988, 2013.

[57] G. Kelly, The simulation of thermomechanically induced stress in plastic encapsulated IC packages. Springer Science & Business Media, 1999.

[58] I. Maus, M. Liebl, C .and Fink, D.-K. Vu, M. Hartung, H. Preu, K. M. B. Jansen, B. Michel, B. Wunderle, and L. Weiss, "Characterization of epoxy based highly filled die attach materials in microelectronics," in Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2017 18th International Conference on, pp. 1–7, IEEE, 2017.

[59] Y. Liu, Power electronic packaging: design, assembly process, reliability and modeling. Springer Science & Business Media, 2012.

[60] E. Ege and Y.-L. Shen, "Thermomechanical response and stress analysis of copper interconnects," Journal of Electronic Materials, vol. 32, no. 10, pp. 1000–1011, 2003.

[61] X. Zhang, A. Kumar, Q. Zhang, Y. Ong, S. Ho, C. Khong, V. Kripesh, J. H. Lau, D.-L. Kwong, V. Sundaram, et al., "Application of piezoresistive stress sensors in ultra thin device handling and characterization," Sensors and Actuators A: Physical, vol. 156, no. 1, pp. 2–7, 2009.

[62] A. Kumar, X. Zhang, Q. X. Zhang, M. C. Jong, G. Huang, L. W. S. Vincent, V. Kripesh, C. Lee, J. H. Lau, D. L. Kwong, et al., "Residual stress analysis in thin device wafer using piezoresistive stress sensor," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 1, no. 6, pp. 841–851, 2011.

[63] Y.-Y. Chang, H. Chung, B.-J. Lwo, and K.-F. Tseng, "In situ stress and reliability monitoring on plastic packaging through piezoresistive stress sensor," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 3, no. 8, pp. 1358–1363, 2013.

[64] L. Siegert and G. Gautier, "Détermination d'une formule analytique prédisant l'effet joule dans les interconnexions en cuivre sur substrat verre," in Journées Nationales du Réseau Doctoral de Microélectronique (JNRDM), vol. 2, pp. 1–3, 2011.

[65] W. M. Bullis, F. H. Brewer, C. D. Kolstad, and L. J. Swartzendruber, "Temperature coefficient of resistivity of silicon and germanium near room temperature," Solid-State Electronics, vol. 11, no. 7, pp. 639–646, 1968.

[66] P. Norton and J. Brandt, "Temperature coefficient of resistance for p-and n-type silicon," Solid-State Electronics, vol. 21, no. 7, pp. 969–974, 1978.

[67] T. H. Tan, D. W. Mcneill, P. Baine, J. H. Montgomery, S. J. N. Mitchell, H. Wadsworth, S. Strahan, and I. Bailie, "Investigation of piezoresistive effect in p-type silicon using four point bending method," 2012.

[68] F. Kerrour, A. Boubakache, and P. Pons, "Modelling of thermal behavior n-doped silicon resistor," Journal of Sensor Technology, vol. 2, no. 3, p. 132, 2012.

[69] J.-Y. Delétage, Etude de la durée de vie d'assemblages microélectroniques par l'utilisation de simulations, de modèles de dégradation et de circuits intégrés spécifiques de test. PhD thesis, Bordeaux 1, 2003.

[70] M. Mansoor, I. Haneef, S. Akhtar, A. De Luca, and F. Udrea, "Silicon diode temperature sensors—a review of applications," Sensors and Actuators A: Physical, vol. 232, pp. 63–74, 2015.

[71] Z. W. Zhong, X. Zhang, B. H. Sim, E. H. Wong, P. S. Teo, and M. K. Iyer, "Calibration of a piezoresistive stress sensor in [100] silicon test chips," in Electronics Packaging Technology Conference, 2002. 4th, pp. 323–326, IEEE, 2002.

[72] J. Richter, M. B. Arnoldus, O. Hansen, and E. V. Thomsen, "Four point bending setup for characterization of semiconductor piezoresistance," Review of scientific instruments, vol. 79, no. 4, p. 044703, 2008.

[73] H. H. Gharib and W. A. Moussa, "A single-polarity piezoresistive three-dimensional stress-sensing rosette," Journal of Microelectromechanical Systems, vol. 20, no. 3, pp. 555–557, 2011.

[74] R. C. Jaeger, J. C. Suhling, M. T. Carey, and R. W. Johnson, "Off-axis sensor rosettes for measurement of the piezoresistive coefficients of silicon," IEEE Transactions on components, hybrids, and manufacturing technology, vol. 16, no. 8, pp. 925–931, 1993.

General Conclusions

In recent years, a number of physical and economical barriers have emerged in the race for miniaturization and speed of integrated circuits. To circumvent these issues, new processes and architectures are continuously developed. In particular, a progressive shift towards *3D integration* is currently observed in the semiconductor industry, as an alternative path to further transistor downscaling. This innovative approach consists in combining chips of different technologies or different functionalities into a single module. A possible strategy to realize such heterogeneous systems is to stack chips on top of each other instead of tiling them on the plane, enabling considerable benefits in terms of compactness and versatility, but also increased performance due to shorter wiring between semiconductor devices. This is especially true for image sensors chips, for which vertical stacking allows the incorporation of additional functionalities such as advanced image signal processing. Among various methods to achieve direct vertical interconnections between, a promising method is Cu/SiO_2 hybrid bonding, enabling simultaneous mechanical and electrical connection between the stacked chips with a submicron interconnection pitch, mostly limited by photolithography resolution and alignment accuracy.

In earlier work at STMicroelectronics and CEA-Leti, a Cu/SiO_2 hybrid bonding process has been developed on several test vehicles, for which the feasibility and reproducibility of chip stacking has been extensively studied[1], allowing the transfer from development stage to manufacturing stage. The main objective of this thesis was to study the mechanical integrity of a more complex stack, namely a 3D BSI image sensor stacked on a logic chip, during integrated circuit processing and chip packaging. In this work, a number of possible issues for this relatively new technology in semiconductor manufacturing have been addressed. The mechanical interactions between the different system components during the fabrication sequence has been analyzed, and the mechanical integrity of several key structures in the image sensor chip investigated, namely (i) interconnection pads at the hybrid bonding interface between the image sensor and logic chip, (ii) wire bondpad structures for image sensor assembly onto the package substrate, and (iii) semiconductor devices in the active region of the image sensor chip, through the evaluation of process-induced mechanical stresses. To do so, a combined numerical and experimental approach

[1] Taibi (2012) and Beilliard (2015)

was adopted, using morphological, mechanical and electrical characterizations, then correlated or extended by thermomechanical finite element analyses, allowing to secure product integration from a thermomechanical perspective.

In the first part of this thesis, the focus was placed on the Cu/SiO_2 hybrid bonding interface between the image sensor and its carrier logic chip, aiming to mitigate potential bonding issues. The influence of various interconnection pad geometries and layouts on the bonding surface topography after planarization by chemical-mechanical polishing was examined. In the investigated range, a negligible influence was reported, and a recessed shape of the Cu pads was typically observed (*Cu dishing*), due to a larger material removal rate compared to the surrounding SiO_2 surfaces. Bonding experiments were then carried out to assess the impact of this initial surface topography on Cu-Cu bonding interface morphology after annealing. Using FIB/SEM tomography on bonded samples with various initial Cu profiles, evidence was found that a dishing threshold exists beyond which only partial Cu-Cu bonding occurs. Although the observed bonding defects are not believed to represent a significant risk for the electrical performance of the 3D integrated circuit, partial bonding could lead to serious mechanical robustness issues.

As hybrid bonding interconnect pitch is targeted to further decrease in the upcoming years, the influence on the bonding behavior of several important design and process parameters is not fully understood. Aiming to provide a mechanistic understanding on the main influent parameters to prevent partial bonding, thermomechanical finite element analysis of Cu-Cu bonding during thermal annealing was carried out. An important result was the observed strong dependence, even for constant dishing depth, of the bonding behavior on Cu recess local curvature. A potentially detrimental impact of an increase in pad aspect ratio or surface fraction was also observed. Such a trend, if confirmed, could mean that dishing compensation is prevented for high density of interconnections. It was shown however that this potential detrimental effect could be mitigated through careful selection of the metal/dielectric material pair to maximize the pad vertical expansion. Finally, the influence of an initial misalignment between facing hybrid bonding pads was also investigated. For sufficiently low misalignement[1], such a defect was not only shown not to cause a decrease of the relative bonded area, but to actually lead to easier bonding for the remaining available Cu-Cu surfaces. This study has thus allowed to identify the main influent parameters on the bonding quality and provide guidelines to limit the occurrences of partial bonding observed in the experimental study.

The second part of the thesis was dedicated to a study of the mechanical robustness of wirebond pads on the backside of the 3D BSI image sensor. The main objective was to optimize the pad structure to prevent potential cracking/delamination after thermosonic bonding. First, several pad architectures were compared using failure inspection, namely the *cratering test*) and FIB/SEM cross-sections to assess mechanical integrity after wire

[1] Typically, below the half-width of the Cu pad.

bonding, and thereby derive a ranking based on the number of observed cracks or delamination. Most failures were observed in the uppermost layers of the interconnect stack, directly below the bond. The introduction of Si/SiO_2 capping layers between the bonding surface and the interconnect stack was seen to considerably improve the pad mechanical robustness, by acting as a buffer layer. The influence of capping thickness was also clearly apparent, as crack occurrences were completely suppressed for the pad configuration with the largest thickness.

Aiming to investigate the influence of interconnection layout below the pad in more detail, a multi-scale finite element analysis of a standardized bond qualification test (*wire-pull test*) was carried out, enabling to account for detailed geometric features in the interconnect stack. A criterion based on the volume-averaged maximum tensile stress in the brittle SiO_2 layers was proposed to compare the mechanical robustness of the different pad configurations. Good correlation with experimental results was obtained, enabling to derive several guidelines to optimize wirebond pad mechanical robustness. Although capping layer presence was found to have the most influence, interconnect layout is also expected to play an important role. Metallization surface fraction has a significant influence in improving the mechanical robustness, due to the load-bearing capability of metal interconnects. In addition, metallization "connectedness" was identified as introducing structural reinforcement to the pad by enabling to relieve mechanical stress exerted on the dielectric material. Similarly, the presence of via plugs between metal lines was also observed to reinforce the mechanical robustness. The thermosonic bonding process is known as one of the most critical fabrication steps in terms of mechanical robustness, due to the large stresses generated in the pad structure. In this work, a comprehensive analysis of the most influent factors in wirebond pad design on the mechanical robustness was provided and guidelines for bondpad optimization were derived, enabling to secure wire-bonding on the backside of a 3D stacked BSI image sensor.

In the final part of this work, a Si piezoresistive stress sensor was developed to monitor process-induced stress variations, both during integrated circuit processing and chip packaging. For the first time, these sensors were implemented in a 3D BSI image sensor stacked by Cu/SiO_2 hybrid bonding. By measuring resistance variations for the doped-Si sensor elements throughout the process sequence, the corresponding stress variations can be derived provided that the different piezoresistive coefficients have been evaluated. This calibration procedure was carried out using a previously developed in-house instrumented four-point bending fixture. Although the relative values of the different coefficients were consistent with the literature, their magnitude was quite low due to the high dopant concentrations used for the image sensor chip (above $1 \times 10^{20}\,\mathrm{cm}^{-3}$). A consequence was a decreased sensitivity to stress variations, causing the measured values to be greatly overestimated. In addition, a potential influence of temperature on the result was identified. Indeed, typical temperature-induced resistance variations were estimated to result in measured stress variations comparable to the expected process-induced stress for the considered chip (i.e.

nearly 50 MPa for 0.5 °C). This suspected strong temperature-dependence has led us to work with temperature-compensated stress components.

To further investigate these potential biases in the stress evaluation, the experimental results were compared with thermomechanical finite element analyses of integrated circuit processing and chip packaging. Much lower values were indeed obtained in the process simulation (below 30 MPa), compared to experimental measurements. In addition, much better agreement between experimental and numerical results was obtained when temperature-compensated stress components were considered, although some unexpectedly high values were still obtained after chip packaging, for which further investigations are required. Thus, lower doping levels are recommended, in order to increase the magnitude of the piezoresistive coefficients and thus the sensitivity of the sensors to stress variations. Nevertheless, low stress values obtained in thermomechanical finite element analyses of both integrated circuit processing and chip packaging tend to indicate a modest influence of the 3D BSI image sensor configuration on pixel operation from a thermo-mechanical viewpoint.

Perspectives

In this thesis, the mechanical integrity of an imager-on-logic 3D integrated circuit during chip processing and packaging was studied. A number of possible issues for this relatively new technology in semiconductor manufacturing have been addressed, allowing to secure product integration from a thermomechanical perspective. Several important questions remain to be investigated however, and a few recommendations may be proposed for future work:

- Adhesive forces between Cu surfaces at the hybrid bonding interface were not accounted for in Chapter II, but were previously shown to play a major role on the bonding behavior [1]. The influence of the lateral extent of these interaction forces relative to pad width remains to be clarified. The latter effect, if any, is expected to strongly depend on the assumed shape for the Cu recess profile and more importantly on the relationship between recess depth and pad width, thus warranting further surface topography characterization or CMP process simulation to provide better understanding on these aspects. In addition, for the most part only 2D Cu recess profiles were considered in the present study, and therefore detailed modeling of the bonding process in the case of a 3D Cu recess surface for square-shaped hybrid bonding pads remains an open topic. Finally, the interplay between adhesion and plasticity needs to be further investigated, accounting for the influence of pad aspect ratio on macroscopic plastic deformation, in addition to local plasticity at the bonding front.

- For several pad configurations studied in Chapter III, no mechanical failures were detected in the cratering test after thermosonic bonding. More comprehensive qualification testing, for instance using wire-pull or wire-shear tests could be carried out to supplement this data. In addition, the proposed simulation approach for pad robustness comparison could be extended by comparing the different pad architectures while maintaining the Cu surface fraction constant, which should enable to better decorrelate the influence of interconnect layout from that of metallization density. Finally, the proposed stress-based criterion for pad robustness ranking could be compared with energetic approaches such as the ARE [2] or the NRE [3] methods, and further investigations could also be carried out to assess the results dependence on the characteristic length used to dimension the averaging domain.

– Aiming to provide a better understanding of the temperature dependence of the stress sensors developed in Chapter IV, the temperature coefficients of resistance of the different sensor elements should be systematically determined, in addition to the piezoresistive coefficients. In addition, a more convenient approach to assess the variability of the piezoresistive coefficients would be to develop a wafer-scale calibration technique such as hydrostatic loading [4], which should provide a great advantage compared to time-consuming and destructive methods such as four-point bending in the present study. If four-point bending is used, a possible improvement could be to introduce calibration-dedicated rosettes oriented at 22.5° with respect to the <100> orientation [5], which enable to obtain in a single loading sequence the two independent piezoresistive coefficients needed to evaluate the temperature-compensated stress components. Regarding stress sensor design, an interesting approach could be to use single-polarity rosettes with two families of N-type resistors[1], each having a distinct set of piezoresistive coefficients [6].

[1] N-type resistors indeed provide overall better sensitivity to stress compared to the P type.

References

[1] Y. Beilliard, R. Estevez, G. Parry, P. McGarry, L. Di Cioccio, and P. Coudrain, "Thermomechanical finite element modeling of Cu–SiO_2 direct hybrid bonding with a dishing effect on Cu surfaces," International Journal of Solids and Structures, vol. 117, pp. 208–220, 2017.

[2] O. van der Sluis, R. A. B. Engelen, W. D. van Driel, M. A. J. van Gils, and R. B. R. van Silfhout, "Efficient damage sensitivity analysis of advanced cu low-k bond pad structures using area release energy," in Thermal, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems, 2006. EuroSime 2006. 7th International Conference on, pp. 1–8, IEEE, 2006.

[3] S. Gallois-Garreignot, V. Fiori, S. Orain, and O. van der Sluis, "Numerical analysis of the reliability of cu/low-k bond pad interconnections under wire pull test: application of a 3D energy based failure criterion," in Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems, 2007. EuroSime 2007. International Conference on, pp. 1–7, IEEE, 2007.

[4] J. Suhling, R. A. Cordes, Y. L. Kang, and R. C. Jaeger, "Wafer-level calibration of stress sensing test chips," in Electronic Components and Technology Conference, 1994. Proceedings., 44th, pp. 1058–1070, IEEE, 1994.

[5] R. C. Jaeger, J. C. Suhling, and A. A. Anderson, "A [100] silicon stress test chip with optimized piezoresistive sensor rosettes," in Electronic Components and Technology Conference, 1994. Proceedings., 44th, pp. 741–749, IEEE, 1994.

[6] H. H. Gharib and W. A. Moussa, "A single-polarity piezoresistive three-dimensional stress-sensing rosette," Journal of Microelectromechanical Systems, vol. 20, no. 3, pp. 555–557, 2011.

Appendix A

Semiconductor devices: P-N diode, MOS capacitor and MOSFET transistor

A.1 P-N diode

Diodes are the most important, yet simplest semiconductor device. They can be obtained by juxtaposing P-type and N-type regions in a semiconductor bulk, thereby forming a P-N junction. In the vicinity of the P-N junction, interdiffusion of majority carriers occurs: holes on the P side start diffusing to the N side, where they recombine with electrons (and conversely for electrons in the N region) (Figure A.1).

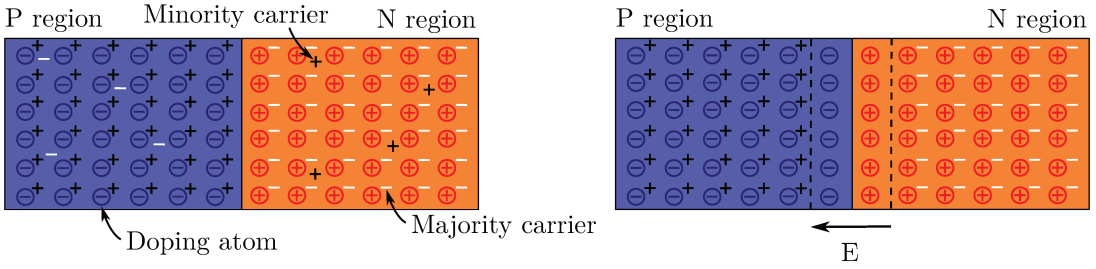


FIGURE A.1: Schematic illustration of the formation of a depletion region at equilibrium due to free carrier interdiffusion and recombination in a P-N junction.

The doping atoms however are incorporated to the silicon lattice and therefore stay in place as majority carriers diffuse through the junction. This leads to the creation of an electric field in the region where doping atoms are exposed, called the depletion region or space-charge region. This electric field opposes majority carrier diffusion (diffusion current) by acting as an intrinsic barrier potential. For Si at room temperature, this barrier corresponds to approximately 0.7 V. At the same time, the electric field enables

minority carriers to cross the depletion region (drift current). Finally, an equilibrium is reached as the diffusion current and drift current oppose each other.

If the P-N junction is connected to a positive terminal on the P side and negative on the N side (forward biasing), the width of the depletion region decreases (Figure A.2), as does the barrier potential opposing diffusion, and therefore the diffusion current increases. For a sufficient potential difference prescribed to the P-N junction ($V > 0.7\,\mathrm{V}$), the intrinsic barrier potential is overcome and an electrical current flows through the P-N junction from the P region to the N region[1].

Conversely, if the P-N junction is connected to a negative terminal on the P side, and positive on the N side (reverse biasing), the majority carriers on each side are attracted to the terminals and the width of the depletion region increases (Figure A.2). More doping atoms become exposed and the electric field in the space-charge region increases, thus decreasing the diffusion current. As a result, both diffusion and drift currents decrease: no current flows through the junction.

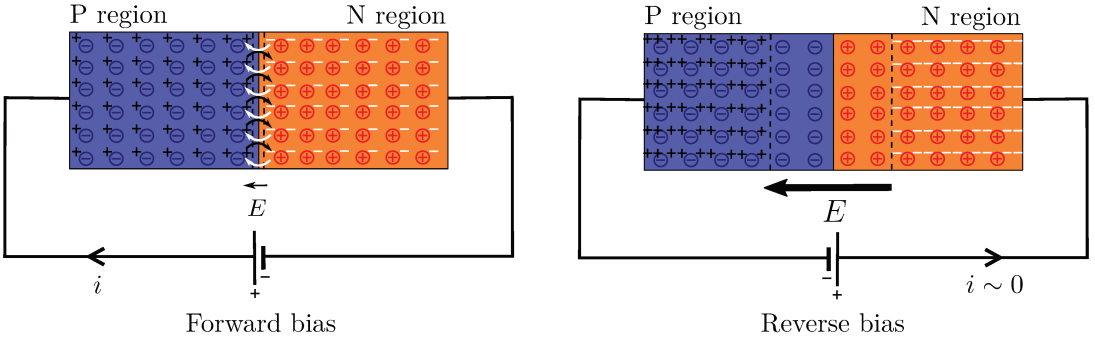


FIGURE A.2: Schematic illustration of the P-N diode operation in forward bias and reverse bias.

Therefore, a P-N junction acts as a diode, conducting current in forward biasing (provided the intrinsic barrier potential of the semiconductor is overcome) and blocking current in reverse biasing[2].

[1] Diffusion current is sustained thanks to recombination of the diffused majority carriers, maintaining a carrier concentration gradient across the junction.

[2] If the reverse bias is too large however, minority carriers are accelerated significantly in the depletion region leading to a temperature increase, in turn increasing the number of minority carriers through thermal generation. A so-called avalanche phenomenon takes place, the diode does not block current anymore, and may even become damaged due to the temperature increase

A.2 MOS capacitor

Capacitors can be formed by depositing a thin insulating oxide layer (*oxide*[1]), sandwiched between a thin conducting layer (*metal gate*[2]) on the P-doped semiconductor substrate (*bulk*). The obtained device is referred to as a metal-oxide-semiconductor (MOS) capacitor.

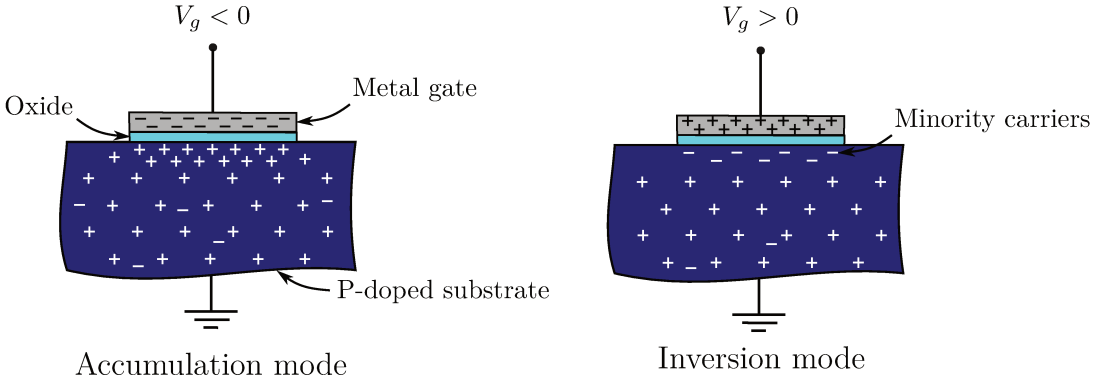


FIGURE A.3: Schematic illustration of the operation of a MOS capacitor.

If a negative potential is maintained on the gate, then its free electrons accumulate near the interface with the oxide (Figure A.3). A vertical electric field results, causing an accumulation of holes on the other side in the P-doped substrate. A capacitor is obtained, and the MOS device is then said to be in *accumulation mode*. Another mode is obtained if a positive potential is prescribed on the gate. In this case, the vertical electric field is in the opposite direction. Holes in the P-doped substrate are repelled, leading to the formation of a depletion region below the gate. This mode of operation is called *depletion mode*.

If the positive voltage at the gate is above a certain threshold ($V_g > V_t$), the accumulation of positive charges in the gate actually causes free electrons, the minority carriers in the P-doped substrate, to be attracted towards the gate (Figure A.3). An N region[3] is therefore created in the bulk below the gate due to effect of the externally-generated electric field. This *field-effect* is at the very basis of the operation of the MOS field-effect transistor (MOSFET), described in the next paragraph. The corresponding mode of operation is called *inversion mode*.

[1] With typical thicknesses in the nanometer range. Although originally it was fabricated using thermal oxide, since 2008 due to device miniaturization, new high-permittivity (*high-k*) materials were introduced, such as hafnium oxide.

[2] Although originally the gate was indeed fabricated using a metal (e.g. aluminum), this is no longer the case. Since 1970, the standard gate material has become polycrystalline silicon, to better accommodate large temperatures and prevent diffusion into the oxide layer.

[3] With the difference that the dopants exposed due to hole depletion are not positive but negative ions

A.3 MOSFET transistor

A MOSFET consists of two P-N junctions flanking a MOS capacitor (Figure A.4). The most straightforward way to achieve this on a P-type wafer is to form two N regions on both sides of the MOS capacitor. In this configuration, if the gate voltage is below the threshold voltage ($V_g < V_t$) then the capacitor is either in accumulation or depletion mode. Therefore, when a potential difference is prescribed between the two N regions, no current flows since in either case no free electrons are available.

However, if the capacitor is placed in the inversion mode ($V_g > V_t$) then the P bulk right below the gate behaves as N-type semiconductor. As a result, a channel of free electrons forms between the N-doped flanks and a current is allowed to flow when a voltage difference is prescribed between both sides. The direction of this current is from the N region with lower potential (*source*) to that with higher potential (*drain*).

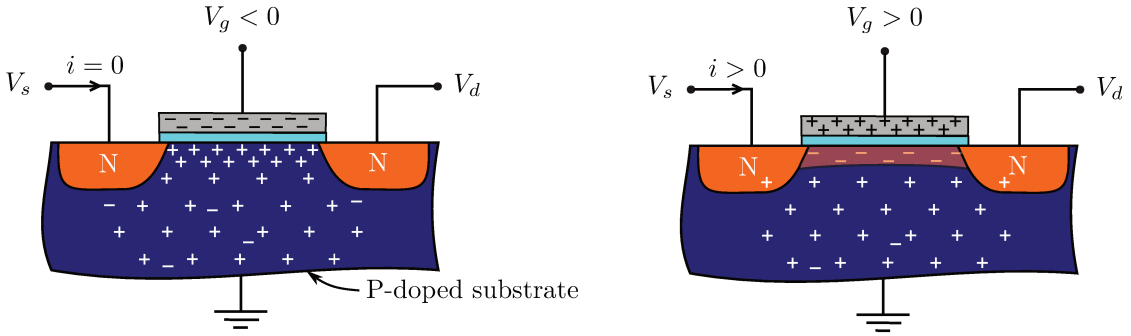


FIGURE A.4: Schematic illustration of the operation of a MOSFET transistor.

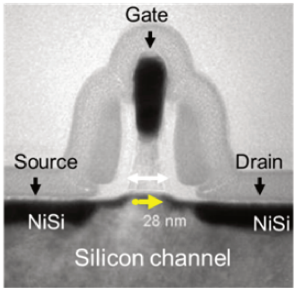


FIGURE A.5: SEM view of a 28 nm MOSFET (Source: STMicroelectronics).

Thus, MOSFETs rely on the field effect in a MOS capacitor to allow (or not) a current to flow between two doped regions. For fixed source and drain voltage, the electrical behavior of the MOSFET is controlled by the gate voltage. In the conducting regime, increasing the gate voltage results in a proportional increase of the source-drain current: the MOSFET operates as an amplifier[1].

[1]However, if the gate voltage is too large, the current saturates and the MOSFET no longer behaves as a linear component.

The transistor described above, a MOS capacitor with N-doped flanks in a P-type bulk, is usually called an NMOS. It is also possible to fabricate a PMOS on a P-type bulk. Similarly to Si resistors, this can be achieved by forming a N-well at the surface of the P bulk, in which the PMOS can be embedded (Figure A.6). To build logic gates for digital ICs, NMOS and PMOS transistors are typically used in association. For instance, the series combination of a PMOS and an NMOS yields an inverter. This technology is referred to as complementary MOS (CMOS).

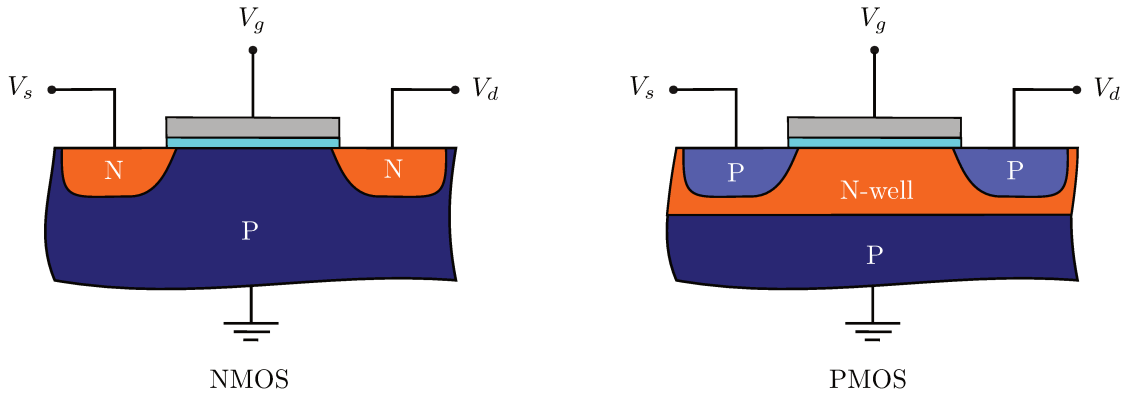


FIGURE A.6: Schematic illustration of the difference between NMOS and PMOS.

Transistors are the most ubiquitous semiconductor devices on ICs (Figure A.5). They are used extensively to create the logic gates in digital ICs, since a transistor operates as a voltage-controlled switch. But transistors can also behave as amplifiers and are therefore used for analogic applications as well.

Appendix B

Estimation of the elastoplastic properties of an electroplated Cu thin film by nanoindentation

B.1 Motivation

Many experimental techniques have been developed to measure the mechanical properties of thin films, such as micro-tensile or micro-cantilever testing for freestanding films [1, 2], or the so-called *bulge test* for a film-on-substrate configuration [3]. While such methods are particularly suitable to investigate the plastic properties of metal thin films, extensive sample preparation is required. Depth-sensing nanoindentation on the other hand can be readily performed on as-deposited blanket thin films, and was thus preferred here owing to its ease of implementation in an industrial context. While the typical use of nanoindentation is to measure the reduced modulus $E_r = E/(1-\nu^2)$ and hardness H of the considered sample, based on loading-unloading curves [4, 5], several studies have extended its range of application to the extraction of plastic properties [6–9].

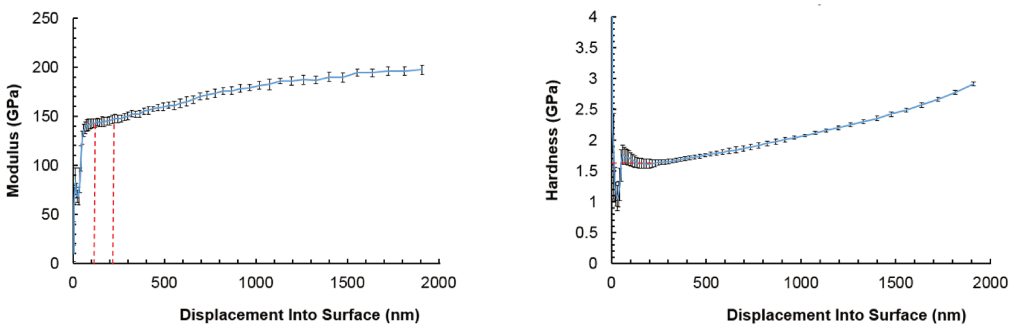
Aiming to provide the necessary input parameters to carry out numerical modeling of the hybrid bonding (Chapter II), the mechanical behavior of an electroplated Cu thin film (similar to those used in the processing of hybrid bonding interconnects) is characterized by nanoindentation experiments. In particular, based on the reverse analysis approach developed by Dao et al. [7] we present an estimation of the plastic properties, assuming a power-law type hardening rule.

B.2 Preliminary study: elastic properties

B.2.1 Materials and methods

The considered sample is a 1.4 µm electroplated Cu thin film deposited on a 775 µm Si wafer. It must be noted that intermediate layers are present between the Cu thin film and the Si thick substrate, consisting of alternating dielectric and barrier layers (SiN + TEOS: PECVD 78 + 750 nm; TaN/Ta: PVD 13nm), as performed during hybrid bonding interconnect processing. The average grain size of the Cu film was estimated in the order of 1 µm by electronic backscattering diffraction (EBSD). Nanoindentation measurements were carried out using a commercial Nanoindentor XP (MTS Corporation, St. Paul, MN) with a diamond Berkovich tip. Fused silica was used as a reference material to determine the compliance of the tool and the area function of the considered tip.

B.2.2 Results



(a) (b)

FIGURE B.1: (a) Reduced modulus and (b) hardness of the Cu thin film as a function of indentation depth by the continuous stiffness measurement method.

As a preliminary study, the reduced modulus and hardness of the thin film were determined by the *continuous stiffness measurement* method, which consists in superimposing a small-amplitude oscillation to the otherwise monotonic indentation load. With this technique, the mechanical properties can be derived as a function of the indentation depth, using the information gained from the partial unloading curves for each load increment [10]. The obtained reduced modulus and hardness of the Cu thin film are respectively 143.0 ± 1.6 GPa and 1.60 ± 0.08 GPa (Figure B.1). These values were evaluated at about 10% of the film thickness to avoid as much as possible any bias due to the influence of the substrate[1] on the measurements. These values are quite close to those found in the literature by nanoindentation for comparable electroplated Cu thin films [12, 13].

[1] For thin films, as a rule of thumb the measurements are generally considered independent from the influence of the substrate for a ratio of indentation depth to coating thickness ranging between 10 to 30% [11]

B.3 Plastic properties

B.3.1 Materials and methods

To characterize the plastic properties of the measured Cu thin film, the reverse analysis method developed by Dao et al. [7] is used, which enables to derive an estimate for the yield stress σ_y and work hardening exponent n from a single loading-unloading nanoindentation curve, assuming a Ludwik power-law hardening rule. Using dimensional analysis, they proposed a set of dimensionless functions relating several characteristics of the $P-h$ force-displacement curve to the elastic-plastic properties of the indented material:

$$\Pi_1\left(\frac{E_r}{\sigma_r}, n\right) = \frac{C}{\sigma_r} \tag{B.1}$$

$$\Pi_2\left(\frac{E_r}{\sigma_r}, n\right) = \frac{1}{E_r h_m} S \tag{B.2}$$

$$\Pi_3\left(\frac{E_r}{\sigma_r}, n\right) = \frac{h_r}{h_m} \tag{B.3}$$

$$\Pi_4\left(\frac{h_r}{h_r}\right) = \frac{H}{E_r} \tag{B.4}$$

$$\Pi_5\left(\frac{h_r}{h_r}\right) = \frac{W_p}{W_t} \tag{B.5}$$

$$\Pi_6 = \frac{1}{E_r\sqrt{A_m}} S = 1.2370 \tag{B.6}$$

where the loading curvature $C = P/h^2$, the unloading slope $S = \frac{\mathrm{d}P_u}{dh}\Big|_{h_m}$, the maximum indentation load and depth $P_{\max}$ and $h_{\max}$, and the residual depth h_r are parameters obtained from the loading-unloading curve (Figure B.4b). The quantity σ_r is the flow stress associated with a *representative plastic strain*[1] of $\epsilon_r = 3.3\%$. This representative flow stress is used in conjunction with the yield stress, both values being estimated from the dimensionless functions, to determine the work hardening exponent n for the assumed power-law hardening plastic behavior:

$$\sigma = \begin{cases} E\varepsilon, & \text{if } \sigma \leq \sigma_Y, \\ R\varepsilon^n, & \text{if } \sigma \geq \sigma_Y \end{cases} \tag{B.7}$$

where the parameter $R = \sigma_Y (E/\sigma_Y)^n$ is the work hardening rate. A J_2-flow theory is used for the description of the rate-independent plastic response, and under multiaxial loading the stress and strain are replaced in Equation (B.7) by the equivalent von Mises stress and strain. This kind of power-law hardening model is the most widely used in nanoindentation

[1] This notion of representative strain was introduced by Tabor [14], who experimentally evidenced a linear relationship between the Vickers hardness H_V (average pressure under a square-based pyramidal diamond indenter) and a representative flow stress σ_r, such that $H_V \approx 3\sigma_r$. The average or *representative* plastic strain ϵ_r associated with this flow stress σ_r during indentation is equivalent to an additional strain of 8% on the tensile stress-strain curve for the indented material, regardless of its initial work-hardening state. In the context of computational modeling of the forward and reverse problem of sharp indentation, Giannakopoulos and Suresh [6] proposed to use a representative strain $\epsilon_r = 29\%$, corresponding to the transition between the innermost plastic "cutting" region below the indenter and its surroundings. In a subsequent work [7], they identified a characteristic plastic strain of $\epsilon_r = 3.3\%$, which allows for the construction of a dimensionless description of the indentation loading response (Π_1 function) that is independent of the work-hardening exponent n. They went on to show that the choice of a representative strain is essentially arbitrary and depends on the form of the Π_i dimensionless functions used, as well as the tip geometry.

mechanics, owing to its simplicity and the fact that it is a good approximation of the post-yield behavior in most metals and alloys for moderate plastic strains [14, 15]. Dao et al. [7] derived the coefficients of the above empirical Π_i functions by carrying out extensive finite element modeling of Berkovich indentation, accounting for a wide range of elastoplastic properties representative of most engineering materials.

The indentation depth must be carefully selected for the measurements. An excessively shallow depth leads to indentation size effects due to surface roughness, indenter tip bluntness or incipient plasticity [16, 17]. In addition, due to the film-on-substrate configuration considered here, if the indentation depth is too large the substrate effect cannot be neglected. To better assess the impact of these undesirable effects, four different maximum indentation depths were explored, corresponding to 5%, 15%, 35% and 65% of the film thickness. For each of these four indentation depths, eighteen loading-unloading curves were acquired.

B.3.2 Results

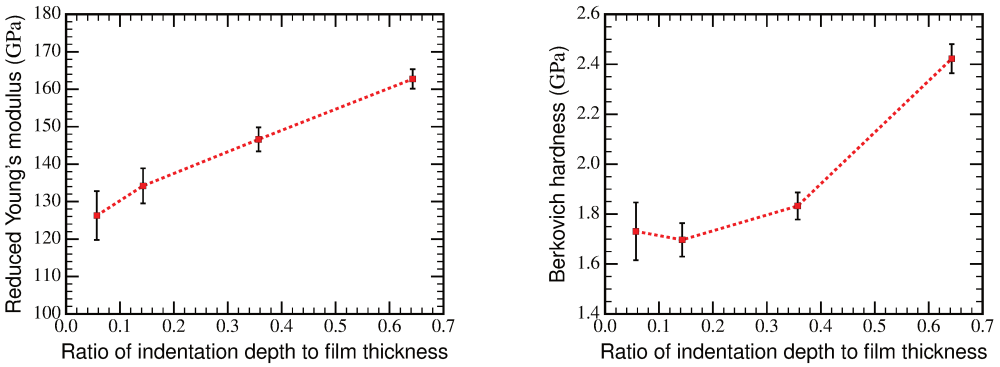
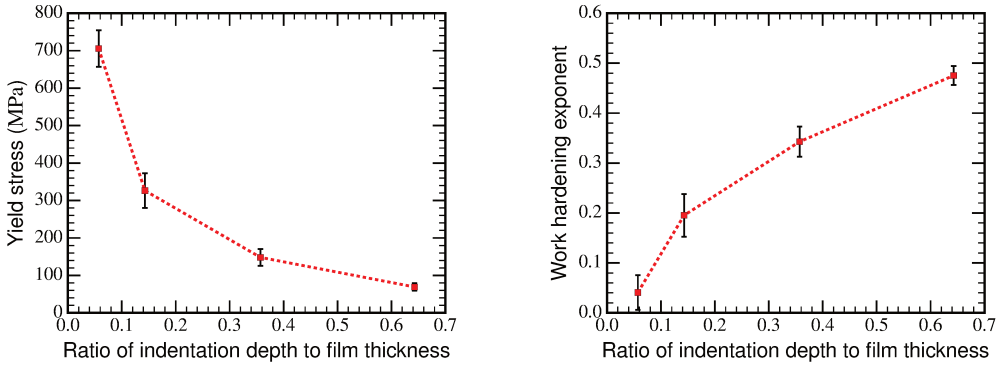


FIGURE B.2: (a) Reduced modulus and (b) hardness of the Cu thin film as a function of indentation depth using the reverse analysis methodology of Dao et al. [7].

The reduced modulus and hardness can also be derived using the reverse analysis methodology described above. The obtained values for each indentation depth are presented in Figure B.2. Both the estimated values and the observed trends for increasing indentation depth are in very good agreement with the continuous stiffness measurements (Figure B.1). As expected, among the four investigated indentation depths, the best agreement with the previous measurements is obtained for intermediate indentation depths, i.e. between 15 and 35% of the film thickness. Within this range, the reduced modulus and hardness are estimated to approximately 140 GPa and 1.8 GPa. Thus, in the following the values obtained for indentation depths ranging between 15 and 35% of the film thickness will be used for the estimation of the plastic properties.

Using this criterion, the estimated yield stress and the work hardening exponent of the Cu thin film are respectively in the range of 200–300 MPa and 0.2–0.3 (Figure B.3). These values are close to those reported by Zhao et al. [15], who performed nanoindentation on a



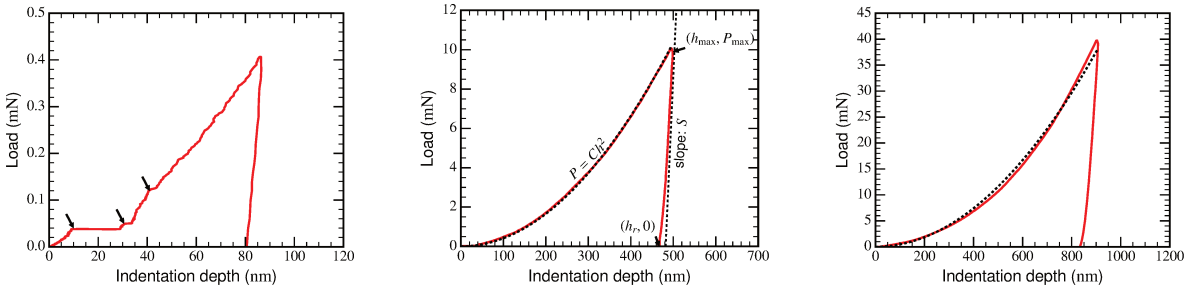
(a) (b)

FIGURE B.3: (a) Yield stress and (b) work hardening exponent of the Cu thin film, obtained as a function of indentation depth from the reverse analysis methodology of Dao et al. [7].

copper film/Si substrate system with 900 nm film thickness using a similar reverse analysis procedure, and found $\sigma_Y = 201$ MPa and $n = 0.205$.

B.4 Discussion

Indentation depth



(a) 5% (b) 35% (c) 65%

FIGURE B.4: Typical loading-unloading curves obtained for indentation depths corresponding to (a) 5%, (b) 35% and (c) 65% of film thickness.

Typical force-displacement curves for 5%, 35% and 65% indentation depths are shown in Figure B.4. The parameters extracted from the curves for the reverse analysis, namely the loading curvature C, the unloading slope S, the maximum indentation load and depth $P_{\max}$ and $h_{\max}$, and the residual depth h_r are also shown on the plots (Figure B.4b). For low indentation depths, the curves exhibit multiple displacement bursts at constant load, making the interpretation of the curves more challenging (Figure B.4a). These are due to the nucleation of geometrically necessary dislocations below the indenter [18, 19]. For high indentation depths, the substrate effect becomes noticeable and the loading curvature is no longer constant, again leading to difficulties in applying the reverse analysis methodology.

The phenomena highlighted for both of these extreme cases are believed to explain the observed departure for the estimated elastic-plastic parameters in the associated indentation ranges.

Uniqueness

The reverse analysis approach adopted here is an attractive technique, as it enables to estimate the elastoplastic properties based on a few simple parameters readily obtained from nanoindentation loading-unloading curves. It has therefore been studied extensively in the literature. Giannakopoulos and Suresh [6] were the first to propose a full set of analytical closed-form functions and the associated methodology for a Berkovich tip. However, these functions were obtained for small deformation cases in their analysis. An extension to large deformation was proposed by Dao et al. [7], who developed a new set of dimensionless functions. These were validated by carrying out the forward analysis, i.e. reproducing experimental force-displacement responses with the elastoplastic properties obtained from the reverse analysis. A major issue however is that of the uniqueness of the solution [20–22]. Dao et al. [7] found their solutions to be unique for low-strength materials, such as copper.

Precision

Uncertainties for the yield stress and work hardening exponent evaluated here using the reverse analysis methodology are quite large. A solution to further improve the precision of the technique could be to use of multiple indenters with different geometries. Chollacoop et al. [23] and Bucaille et al. [8] independently proposed extensions of the work of Dao et al., with universal dimensionless functions dependent on the included tip angle of the indenter. By using several tips with different included tip angles, several flow stress values can be determined, each corresponding to a distinct representative strain ϵ_r, thereby enabling a more precise fit for the power-law equation describing the hardening behavior.

Limitations

In addition to the limitations discussed above, it can be noted that the previous results are obtained for a blanket Cu thin film, whereas for the intended application Cu interconnects are considered. Plastic strain gradient effects are expected for these structures due to their small size and additional lateral constraint, which could result in a significantly different plastic behavior compared the obtained values, e.g. larger yield strength.

B.5 Conclusion

In this appendix, the elastoplastic properties of thin Cu films used in the processing of hybrid bonding pads were estimated by nanoindentation using a reverse analysis method developed by Dao et al. [7]. This method is based on universal dimensionless functions obtained by finite element modeling of Berkovich indentation. This set of empirical functions enables an estimation of both elastic and plastic properties, based solely on a few parameters extracted from experimental loading-unloading curves. The results were compared with the more widespread continuous stiffness measurement technique, as well as other similar measurements in the literature, and good agreement was obtained. The reduced modulus and hardness of the Cu thin film were found close to 140 GPa and 1.6 GPa respectively. Assuming a power-law hardening behavior, the yield stress and work hardening exponents were estimated within around 200–300 MPa and 0.2–0.3 respectively. The uncertainty in the evaluation of the plastic properties is quite large, due to the sensitivity to the chosen indentation depth for the considered thin film-on-substrate configuration. However, it is expected that further precision could be gained by combining measurements carried out using several indenters with different geometries, as was shown by several authors [8, 23].

References

[1] H. Huang and F. Spaepen, "Tensile testing of free-standing cu, ag and al thin films and ag/cu multilayers," Acta Materialia, vol. 48, no. 12, pp. 3261–3269, 2000.

[2] H. Espinosa and B. Prorok, "Size effects on the mechanical behavior of gold thin films," Journal of Materials Science, vol. 38, no. 20, pp. 4125–4128, 2003.

[3] Y. Xiang, X. Chen, and J. J. Vlassak, "The mechanical properties of electroplated cu thin films measured by means of the bulge test technique," MRS Online Proceedings Library Archive, vol. 695, 2001.

[4] M. F. Doerner and W. D. Nix, "A method for interpreting the data from depth-sensing indentation instruments," Journal of Materials research, vol. 1, no. 4, pp. 601–609, 1986.

[5] G. Pharr and W. Oliver, "Measurement of thin film mechanical properties using nanoindentation," Mrs Bulletin, vol. 17, no. 7, pp. 28–33, 1992.

[6] A. Giannakopoulos and S. Suresh, "Determination of elastoplastic properties by instrumented sharp indentation," Scripta materialia, vol. 40, no. 10, pp. 1191–1198, 1999.

[7] M. Dao, N. V. Chollacoop, K. J. Van Vliet, T. A. Venkatesh, and S. Suresh, "Computational modeling of the forward and reverse problems in instrumented sharp indentation," Acta materialia, vol. 49, no. 19, pp. 3899–3918, 2001.

[8] J.-L. Bucaille, S. Stauss, E. Felder, and J. Michler, "Determination of plastic properties of metals by instrumented indentation using different sharp indenters," Acta materialia, vol. 51, no. 6, pp. 1663–1678, 2003.

[9] Z. S. Ma, Y. C. Zhou, S. G. Long, X. L. Zhong, and C. Lu, "Characterization of stress-strain relationships of elastoplastic materials: An improved method with conical and pyramidal indenters," Mechanics of Materials, vol. 54, pp. 113–123, 2012.

[10] W. C. Oliver and G. M. Pharr, "An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments," Journal of Materials Research, vol. 7, no. 6, pp. 1564–1583, 1992.

[11] K. Tunvisut, E. P. Busso, N. P. O'dowd, and H. P. Brantner, "Determination of the mechanical properties of metallic thin films and substrates from indentation tests," Philosophical Magazine A, vol. 82, no. 10, pp. 2013–2029, 2002.

[12] S. H. Hong, K. S. Kim, Y.-M. Kim, J.-H. Hahn, C.-S. Lee, and J.-H. Park, "Characterization of elastic moduli of cu thin films using nanoindentation technique," Composites science and technology, vol. 65, no. 9, pp. 1401–1408, 2005.

[13] S.-Y. Chang and T.-K. Chang, "Grain size effect on nanomechanical properties and deformation behavior of copper under nanoindentation test," Journal of applied physics, vol. 101, no. 3, p. 033507, 2007.

[14] D. Tabor, The hardness of metals. Oxford University Press, 1951.

[15] M. Zhao, Y. Xiang, J. Xu, N. Ogasawara, N. Chiba, and X. Chen, "Determining mechanical properties of thin films from the loading curve of nanoindentation testing," Thin Solid Films, vol. 516, no. 21, pp. 7571–7580, 2008.

[16] W. D. Nix and H. Gao, "Indentation size effects in crystalline materials: a law for strain gradient plasticity," Journal of the Mechanics and Physics of Solids, vol. 46, no. 3, pp. 411–425, 1998.

[17] G. M. Pharr, E. G. Herbert, and Y. Gao, "The indentation size effect: a critical examination of experimental observations and mechanistic interpretations," Annual Review of Materials Research, vol. 40, pp. 271–292, 2010.

[18] S. Suresh, T.-G. Nieh, and B. W. Choi, "Nano-indentation of copper thin films on silicon substrates," Scripta Materialia, vol. 41, no. 9, pp. 951–957, 1999.

[19] A. Gouldstone, H.-J. Koh, K.-Y. Zeng, A. E. Giannakopoulos, and S. Suresh, "Discrete and continuous deformation during nanoindentation of thin films," Acta Materialia, vol. 48, no. 9, pp. 2277–2295, 2000.

[20] Y.-T. Cheng and C.-M. Cheng, "Can stress–strain relationships be obtained from indentation curves using conical and pyramidal indenters?," Journal of Materials Research, vol. 14, no. 9, pp. 3493–3496, 1999.

[21] X. Chen, N. Ogasawara, M. Zhao, and N. Chiba, "On the uniqueness of measuring elastoplastic properties from indentation: the indistinguishable mystical materials," Journal of the Mechanics and Physics of Solids, vol. 55, no. 8, pp. 1618–1660, 2007.

[22] L. Liu, N. Ogasawara, N. Chiba, and X. Chen, "Can indentation technique measure unique elastoplastic properties?," Journal of Materials Research, vol. 24, no. 3, pp. 784–800, 2009.

[23] N. Chollacoop, M. Dao, and S. Suresh, "Depth-sensing instrumented indentation with dual sharp indenters," Acta materialia, vol. 51, no. 13, pp. 3713–3729, 2003.

Appendix C

General equations for the various piezoresistive stress sensor configurations on (100) Si wafers

In this appendix, the general equations relating the stress-induced resistance changes to the stress components are detailed for each rosette configuration.

In the case of piezoresistive stress sensors embedded on a (100) Si substrate, a maximum number of four stress components can be obtained from the measured resistance variations, namely the three normal stresses σ_{11}, σ_{22} and σ_{33}, and the inplane shear stress σ_{12}. Thus, to extract all stress components, four resistors with distinct orientations are required, forming a rosette structure, typically in the 0–45–90–135° configuration. However, since for silicon there are only three independent piezoresistive coefficients (π_{11}, π_{12} and π_{44}), resistors with two different doping types must necessarily be incorporated in the four-element rosette, so that two sets of distinct piezoresistive coefficients are introduced, and the obtained system of equations can be inverted. In the general case, there are twelve possible combinations of N-type and P-type resistors in a four-element rosette enabling to extract the stress components.

The corresponding equations are derived here, accounting for the effect of temperature on the resistance change through the (first-order) thermal coefficient of resistance α. In the following, the four resistors R_1, R_2, R_3 and R_4 are associated respectively with the $x_1 = [100]$, $x_1' = [110]$, $x_2 = [010]$ and $x_2' = [\bar{1}00]$ crystallographic directions (see Figure IV.1). It is assumed that the piezoresistive coefficients and the thermal coefficients of resistance are the same for all the resistors in the rosette structure.

C.1 (3P1N) rosettes

Rosette configuration consisting of 3 P-type resistors and 1 N-type resistors are referred to as (3P1N) rosettes. There are four possible cases, depending on the orientation of the N resistor: 1N3P, 1P1N2P, 2P1N1P, 3P1N.

C.1.1 1N3P rosette: N resistor along [100]

C.1.1.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^n & \pi_{12}^n & \pi_{12}^n & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & \pi_{44}^p \\ \pi_{12}^p & \pi_{11}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & -\pi_{44}^p \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_n \\ \alpha_p \\ \alpha_p \\ \alpha_p \end{bmatrix} \tag{C.1}$$

C.1.1.2 Stress components

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^p \frac{\Delta R_1}{R_1} - \pi_{12}^n \left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_3}{R_3}\right]}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{22} = & \frac{\pi_{12}^p \left(\frac{\Delta R_1}{R_1} - \frac{\pi_{11}^n - \pi_{12}^n}{\pi_{11}^p - \pi_{12}^p} \left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_3}{R_3}\right] \right)}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} + \frac{1}{\pi_{11}^p - \pi_{12}^p} \frac{\Delta R_3}{R_3} - \frac{\alpha_n \pi_{12}^p - \alpha_p \pi_{12}^n}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{33} = & \frac{\frac{\pi_{11}^n \pi_{11}^p - \pi_{12}^n \pi_{12}^p}{\pi_{11}^p - \pi_{12}^p} \left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_3}{R_3}\right] - \left(\pi_{11}^p + \pi_{12}^p\right) \frac{\Delta R_1}{R_1}}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} - \frac{1}{\pi_{11}^p - \pi_{12}^p} \frac{\Delta R_3}{R_3} \\ & + \frac{\alpha_n \left(\pi_{11}^p + \pi_{12}^p\right) - \alpha_p \left(\pi_{11}^n + \pi_{12}^n\right)}{\pi_{11}^n \pi_{12}^p - \pi_{11}^p \pi_{12}^n} \Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^p} \left[\frac{\Delta R_2}{R_2} - \frac{\Delta R_4}{R_4}\right] \end{cases} \tag{C.2}$$

C.1.2 1P1N2P rosette: N resistor along [110]

C.1.2.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^p & \pi_{12}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \pi_{12}^n & \pi_{44}^n \\ \pi_{12}^p & \pi_{11}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & -\pi_{44}^p \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_p \\ \alpha_n \\ \alpha_p \\ \alpha_p \end{bmatrix} \tag{C.3}$$

C.1.2.2 Stress components

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^p \frac{\pi_{44}^n}{\pi_{44}^p}\left[2\frac{\Delta R_4}{R_4}-\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]+2\pi_{12}^p\frac{\Delta R_2}{R_2}-\pi_{12}^n\frac{\Delta R_1}{R_1}-\pi_{12}^p\frac{\pi_{11}^n-\pi_{12}^n}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_3}{R_3}}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} + \frac{1}{2\left(\pi_{11}^p-\pi_{12}^p\right)}\frac{\Delta R_1}{R_1} \\ & -\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{22} = & \frac{\pi_{12}^p \frac{\pi_{44}^n}{\pi_{44}^p}\left[2\frac{\Delta R_4}{R_4}-\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]+2\pi_{12}^p\frac{\Delta R_2}{R_2}-\pi_{12}^n\frac{\Delta R_3}{R_3}-\pi_{12}^p\frac{\pi_{11}^n-\pi_{12}^n}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_1}{R_1}}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} + \frac{1}{2\left(\pi_{11}^p-\pi_{12}^p\right)}\frac{\Delta R_3}{R_3} \\ & -\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{33} = & \frac{\left(\pi_{11}^n+\pi_{12}^n+\frac{\pi_{44}^n}{\pi_{44}^p}\left(\pi_{11}^p+\pi_{12}^p\right)\right)\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-2\left(\pi_{11}^p+\pi_{12}^p\right)\left(\frac{\Delta R_2}{R_2}+\frac{\pi_{44}^n}{\pi_{44}^p}\frac{\Delta R_4}{R_4}\right)}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} \\ & +\frac{\alpha_n\left(\pi_{11}^p+\pi_{12}^p\right)-\alpha_p\left(\pi_{11}^n+\pi_{12}^n\right)}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^p}\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}-2\frac{\Delta R_4}{R_4}\right] \end{cases} \tag{C.4}$$

C.1.3 2P1N1P rosette: N resistor along [010]

C.1.3.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^p & \pi_{12}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & \pi_{44}^p \\ \pi_{12}^n & \pi_{11}^n & \pi_{12}^n & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & -\pi_{44}^p \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_p \\ \alpha_p \\ \alpha_n \\ \alpha_p \end{bmatrix} \tag{C.5}$$

C.1.3.2 Stress components

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^p\left(\frac{\Delta R_3}{R_3}-\frac{\pi_{11}^n-\pi_{12}^n}{\pi_{11}^p-\pi_{12}^p}\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_1}{R_1}\right]\right)}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}+\frac{1}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_1}{R_1}-\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{22} = & \frac{\pi_{12}^p\frac{\Delta R_3}{R_3}-\pi_{12}^n\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_1}{R_1}\right]}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}-\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{33} = & \frac{\frac{\pi_{11}^n\pi_{11}^p-\pi_{12}^n\pi_{12}^p}{\pi_{11}^p-\pi_{12}^p}\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}-\frac{\Delta R_1}{R_1}\right]-\left(\pi_{11}^p+\pi_{12}^p\right)\frac{\Delta R_3}{R_3}}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}-\frac{1}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_1}{R_1} \\ & +\frac{\alpha_n\left(\pi_{11}^p+\pi_{12}^p\right)-\alpha_p\left(\pi_{11}^n+\pi_{12}^n\right)}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^p}\left[\frac{\Delta R_2}{R_2}-\frac{\Delta R_4}{R_4}\right] \end{cases} \tag{C.6}$$

C.1.4 (3P1N) rosette: N resistor along $[\bar{1}10]$

C.1.4.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^p & \pi_{12}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & \pi_{44}^p \\ \pi_{12}^p & \pi_{11}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \pi_{12}^n & -\pi_{44}^n \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_p \\ \alpha_p \\ \alpha_p \\ \alpha_n \end{bmatrix} \tag{C.7}$$

C.1.4.2 Stress components

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^p \frac{\pi_{44}^n}{\pi_{44}^p}\left[2\frac{\Delta R_2}{R_2}-\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]+2\pi_{12}^p\frac{\Delta R_4}{R_4}-\pi_{12}^n\frac{\Delta R_1}{R_1}-\pi_{12}^p\frac{\pi_{11}^n-\pi_{12}^n}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_3}{R_3}}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)}+\frac{1}{2\left(\pi_{11}^p-\pi_{12}^p\right)}\frac{\Delta R_1}{R_1} \\ & -\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{22} = & \frac{\pi_{12}^p \frac{\pi_{44}^n}{\pi_{44}^p}\left[2\frac{\Delta R_2}{R_2}-\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]+2\pi_{12}^p\frac{\Delta R_4}{R_4}-\pi_{12}^n\frac{\Delta R_3}{R_3}-\pi_{12}^p\frac{\pi_{11}^n-\pi_{12}^n}{\pi_{11}^p-\pi_{12}^p}\frac{\Delta R_1}{R_1}}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)}+\frac{1}{2\left(\pi_{11}^p-\pi_{12}^p\right)}\frac{\Delta R_3}{R_3} \\ & -\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{33} = & \frac{\left(\pi_{11}^n+\pi_{12}^n+\frac{\pi_{44}^n}{\pi_{44}^p}\left(\pi_{11}^p+\pi_{12}^p\right)\right)\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-2\left(\pi_{11}^p+\pi_{12}^p\right)\left(\frac{\Delta R_4}{R_4}+\frac{\pi_{44}^n}{\pi_{44}^p}\frac{\Delta R_2}{R_2}\right)}{2\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)} \\ & +\frac{\alpha_n\left(\pi_{11}^p+\pi_{12}^p\right)-\alpha_p\left(\pi_{11}^n+\pi_{12}^n\right)}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\ \sigma_{12} = & -\frac{1}{2\pi_{44}^p}\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}-2\frac{\Delta R_2}{R_2}\right] \end{cases} \tag{C.8}$$

C.2 (1P3N) rosettes

Rosette configurations consisting of 1 P-type resistors and 3 N-type resistors are referred to as (1P3N) rosettes. There are four possible cases, depending on the orientation of the P resistor: 1P3N, 1N1P2N, 2N1P1N, 3N1P. The equations can be obtained by permuting the n and p indices in Section C.

C.3 (2P2N) rosettes

Rosette configurations consisting of 2 P-type resistors and 2 N-type resistors are referred to as (2P2N) rosettes. There are four possible cases: 2N2P, 2P2N, $(1N1P)^2$, $(1P1N)^2$.

C.3.1 2N2P rosette: N resistors along [100] and [110]

C.3.1.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^n & \pi_{12}^n & \pi_{12}^n & 0 \\ \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \frac{\pi_{11}^n}{2}+\frac{\pi_{12}^n}{2} & \pi_{12}^n & \pi_{44}^n \\ \pi_{12}^p & \pi_{11}^p & \pi_{12}^p & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & -\pi_{44}^p \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_n \\ \alpha_p \\ \alpha_p \\ \alpha_p \end{bmatrix} \tag{C.9}$$

C.3.1.2 Stress components

$$\begin{cases}
\sigma_{11} = & \frac{\pi_{12}^n\left(\pi_{11}^p-\pi_{12}^p\right)\left(\pi_{44}^p\left[2\frac{\Delta R_2}{R_2}-\frac{\Delta R_1}{R_1}\right]+\pi_{44}^n\left[2\frac{\Delta R_4}{R_4}-\frac{\Delta R_3}{R_3}\right]\right)-\pi_{44}^n\pi_{12}^p\left(\pi_{11}^p-\pi_{12}^p\right)\frac{\Delta R_1}{R_1}-\pi_{44}^p\pi_{12}^n\left(\pi_{11}^n-\pi_{12}^n\right)\frac{\Delta R_3}{R_3}}{\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)\left[\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)\right]} \\
& +\frac{\pi_{44}^p\frac{\Delta R_1}{R_1}}{\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)}-\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\
\sigma_{22} = & \frac{\pi_{12}^p\left(\pi_{11}^n-\pi_{12}^n\right)\left(\pi_{44}^p\left[2\frac{\Delta R_2}{R_2}-\frac{\Delta R_1}{R_1}\right]+\pi_{44}^n\left[2\frac{\Delta R_4}{R_4}-\frac{\Delta R_3}{R_3}\right]\right)-\pi_{44}^n\pi_{12}^p\left(\pi_{11}^p-\pi_{12}^p\right)\frac{\Delta R_1}{R_1}-\pi_{44}^p\pi_{12}^n\left(\pi_{11}^n-\pi_{12}^n\right)\frac{\Delta R_3}{R_3}}{\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)\left[\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)\right]} \\
& -\frac{\pi_{44}^n\frac{\Delta R_3}{R_3}}{\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)}-\frac{\alpha_n\pi_{12}^p-\alpha_p\pi_{12}^n}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\
\sigma_{33} = & \frac{\left(\pi_{11}^n\pi_{11}^p-\pi_{12}^n\pi_{12}^p\right)\left(\pi_{44}^p\left[\frac{\Delta R_1}{R_1}-2\frac{\Delta R_2}{R_2}\right]+\pi_{44}^n\left[\frac{\Delta R_3}{R_3}-2\frac{\Delta R_4}{R_4}\right]\right)+\pi_{44}^n\left(\pi_{11}^{p\,2}-\pi_{12}^{p\,2}\right)\frac{\Delta R_1}{R_1}+\pi_{44}^p\left(\pi_{11}^{n\,2}-\pi_{12}^{n\,2}\right)\frac{\Delta R_3}{R_3}}{\left(\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n\right)\left[\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)\right]} \\
& +\frac{\pi_{44}^n\frac{\Delta R_3}{R_3}-\pi_{44}^p\frac{\Delta R_1}{R_1}}{\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)}+\frac{\alpha_n\left(\pi_{11}^p+\pi_{12}^p\right)-\alpha_p\left(\pi_{11}^n+\pi_{12}^n\right)}{\pi_{11}^n\pi_{12}^p-\pi_{11}^p\pi_{12}^n}\Delta T \\
\sigma_{12} = & \frac{\left(\pi_{11}^p-\pi_{12}^p\right)\left[\frac{\Delta R_1}{R_1}-\frac{\Delta R_2}{R_2}\right]+\left(\pi_{11}^n-\pi_{12}^n\right)\left[\frac{\Delta R_3}{R_3}-\frac{\Delta R_4}{R_4}\right]}{\pi_{44}^p\left(\pi_{11}^n-\pi_{12}^n\right)-\pi_{44}^n\left(\pi_{11}^p-\pi_{12}^p\right)}
\end{cases} \tag{C.10}$$

C.3.2 2P2N rosette: P resistors along [100] and [110]

The equations can be obtained by permuting the n and p indices in Equations (C.9) and (C.10).

C.3.3 (1N1P)² rosette: N resistors along $x_1 = [100]$ and $x_2 = [010]$

C.3.3.1 Stress-induced resistance changes

$$\begin{bmatrix} \frac{\Delta R_1}{R_1} \\ \frac{\Delta R_2}{R_2} \\ \frac{\Delta R_3}{R_3} \\ \frac{\Delta R_4}{R_4} \end{bmatrix} = \begin{bmatrix} \pi_{11}^n & \pi_{12}^n & \pi_{12}^n & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & \pi_{44}^p \\ \pi_{12}^n & \pi_{11}^n & \pi_{12}^n & 0 \\ \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \frac{\pi_{11}^p}{2}+\frac{\pi_{12}^p}{2} & \pi_{12}^p & -\pi_{44}^p \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \end{bmatrix} + \Delta T \begin{bmatrix} \alpha_n \\ \alpha_p \\ \alpha_n \\ \alpha_p \end{bmatrix} \tag{C.11}$$

C.3.3.2 Stress components

$$\begin{cases} \sigma_{11} = & \frac{\pi_{12}^{p}\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-\pi_{12}^{n}\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}\right)}+\frac{1}{2\left(\pi_{11}^{n}-\pi_{12}^{n}\right)}\left[\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]-\frac{\alpha_n\pi_{12}^{p}-\alpha_p\pi_{12}^{n}}{\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}}\Delta T \\ \sigma_{22} = & \frac{\pi_{12}^{p}\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]-\pi_{12}^{n}\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]}{2\left(\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}\right)}-\frac{1}{2\left(\pi_{11}^{n}-\pi_{12}^{n}\right)}\left[\frac{\Delta R_1}{R_1}-\frac{\Delta R_3}{R_3}\right]-\frac{\alpha_n\pi_{12}^{p}-\alpha_p\pi_{12}^{n}}{\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}}\Delta T \\ \sigma_{33} = & \frac{\left(\pi_{11}^{n}+\pi_{12}^{n}\right)\left[\frac{\Delta R_2}{R_2}+\frac{\Delta R_4}{R_4}\right]-\left(\pi_{11}^{p}+\pi_{12}^{p}\right)\left[\frac{\Delta R_1}{R_1}+\frac{\Delta R_3}{R_3}\right]}{2\left(\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}\right)}+\frac{\alpha_n\left(\pi_{11}^{p}+\pi_{12}^{p}\right)-\alpha_p\left(\pi_{11}^{n}+\pi_{12}^{n}\right)}{\pi_{11}^{n}\pi_{12}^{p}-\pi_{11}^{p}\pi_{12}^{n}}\Delta T \\ \sigma_{12} = & \frac{1}{2\pi_{44}^{p}}\left[\frac{\Delta R_2}{R_2}-\frac{\Delta R_4}{R_4}\right] \end{cases} \tag{C.12}$$

C.3.4 (1P1N)² rosette: P resistors along [100] and [010]

The equations can be obtained by permuting the n and p indices in Equations (C.11) and (C.12).

Résumé

Ces dernières années, un certain nombre de barrières physiques et économiques sont apparues dans la course pour la miniaturisation et l'amélioration des performances des circuits intégrés. En effet, pendant plusieurs décennies, l'industrie des semiconducteurs a été portée par une augmentation continue du nombre de composants semiconducteurs intégrés par unité de surface (typiquement, des transistors MOS). Cependant, alors que des dimensions de longueur de grille inférieures à 10 nm sont atteintes pour les transistors des dernières générations de circuits intégrés, les effets quantiques deviennent non-négligeables et conduisent à une consommation électrique accrue ainsi que des phénomènes d'auto-échauffement, en raison de courants de fuite plus importants. De plus, la nécessité d'une densité d'interconnexions accrue entre des composants semiconducteurs toujours plus nombreux entraine une augmentation des délais de propagation du signal, à cause de la génération de capacités parasites entre les lignes d'interconnexions.

Pour dépasser ces limites, de nouvelles architectures sont continûment développées. En particulier, on observe ces dernières années un tournant dans l'industrie de la microélectronique vers les stratégies d'intégration 3D, comme une alternative à la réduction des dimensions des transistors. Cette approche innovante consiste à combiner en un seul et même module des puces de technologies ou fonctionnalités diverses. Une stratégie possible pour réaliser ces systèmes hétérogènes est d'empiler verticalement les puces les unes sur les autres plutôt que de les juxtaposer dans le plan, permettant ainsi des gains considérables en terme de compacité et de polyvalence des circuits. Ceci vaut en particulier pour les capteurs d'image, pour lesquels l'exploitation de la dimension verticale rend possible l'incorporation de fonctionnalités supplémentaires, notamment pour le traitement d'image. Parmi les nombreuses méthodes existantes pour réaliser des interconnexions verticales directes entre les puces empilées, le collage « hybride » cuivre/oxyde est une approche prometteuse permettant de réaliser simultanément une connexion mécanique et électrique entre les puces, avec un pas d'interconnexion submicronique car limité principalement par la précision d'alignement atteignable entre les plots de collage métalliques au moment de leur mise en contact.

Dans des travaux récents à STMicroelectronics et au CEA-Leti, un procédé de collage hybride cuivre/oxyde (Cu/SiO_2) a été développé à l'aide de plusieurs circuits de test, pour lesquels la performance et la fiabilité électrique, ainsi que la morphologie de l'interface

de collage ont été étudiés en détail[1], rendant possible le transfert de cette technologie de la phase de développement à la phase de production. Dans cette thèse, l'intégration d'un système plus complexe est considérée, à savoir un capteur d'image éclairé par la face arrière empilé sur un circuit logique. En comparaison des études précédentes, un champ plus large de procédés de fabrication est donc étudié, incluant à la fois la fabrication du circuit imageur et son encapsulation dans un boîtier à puce, en plus de l'empilement de puces par le procédé de collage hybride cuivre/oxyde.

Un enjeu majeur pour ce type d'architecture innovante réside dans la tenue mécanique des éléments de connexion électrique. Cette thèse vise à examiner la robustesse mécanique d'un capteur d'image reporté sur un circuit logique par empilement 3D, dans le but de prévenir un certain nombre de problèmes potentiels causés par les contraintes thermomécaniques générées pendant sa fabrication. Dans ce travail de thèse, les contraintes mécaniques générées dans le capteur d'image empilé pendant l'élaboration du circuit et son encapsulation dans un boîtier à puce sont examinées, et les interactions entre les différents composants du système analysées. L'intégrité mécanique de plusieurs structures clés est étudiée, notamment : (i) les plots d'interconnexion à l'interface de collage hybride entre la puce imageur et la puce logique, (ii) les plots d'assemblage filaire faisant le lien entre le capteur d'image empilé et le substrat du boîtier, ainsi que (iii) les composants semiconducteurs dans la zone active du substrat silicium du capteur d'image, à travers l'évaluation *in situ* des contraintes mécaniques induites par les procédés de fabrication grâce à des capteurs de contraintes piézorésistifs à base de silicium dopé. Pour ce faire, une approche combinant caractérisations expérimentales et analyses numériques a été adoptée : les mesures morphologiques, mécaniques et électriques effectuées sont systématiquement corrélées et étendues à l'aide de simulations par la méthode des éléments finis, permettant de garantir la bonne intégration des produits d'imagerie du point de vue thermomécanique.

Dans un premier temps, nous nous sommes concentrés sur l'interface de collage hybride cuivre-oxyde entre le capteur d'image et sa puce-support, afin d'examiner la fermeture de l'interface de collage cuivre-cuivre et les problèmes potentiels associés. Pour l'application considérée dans cette thèse, le procédé de collage hybride est réalisé grâce au dépôt d'une couche supplémentaire structurée cuivre/oxyde à la surface de l'empilement d'interconnexion pour chacune des deux puces à assembler. Grâce à une préparation de surface et un traitement thermique spécifiques, ces surfaces aux motifs assortis peuvent être collées directement sans aucun matériau intermédiaire (ni colle, ni alliage de brasage). Cette technique permet ainsi l'obtention d'une couche d'interconnexion compacte, rendant cette technique particulièrement adaptée pour les applications d'intégration tridimensionnelle à ultra-haute-densité. Toutefois, le procédé de collage hybride cuivre-oxyde requiert une excellente planéité de surface, avec une conformité suffisante pour permettre une bonne fermeture de l'interface hybride. La topographie de surface est donc un facteur clé pour le

[1] Taibi (2012) et Beilliard (2015)

collage hybride cuivre-oxyde, et doit faire l'objet d'un suivi rigoureux afin de remplir les critères de planéité stricts qui lui sont associés.

L'influence de diverses géométries et divers agencements pour les plots métalliques de collage hybride sur l'état de surface après planarisation mécano-chimique a été évaluée expérimentalement. L'impact de cette topographie de surface initiale sur la morphologie de l'interface de collage cuivre-cuivre a ensuite été examiné. Une analyse thermomécanique du procédé de collage a enfin été conduite à l'aide de simulations par la méthode des éléments finis, dans le but de déterminer l'influence de paramètres supplémentaires liés aux règles de dessin ou au procédé, et identifiés comme critiques pour les applications futures, comme par exemple le pas d'interconnexion ou la précision d'alignement.

Pour des paramètres de polissage nominaux, l'analyse de la topographie des surfaces par microscopie à force atomique avant collage a mis en évidence un surpolissage systématique des motifs d'interconnexion de cuivre (phénomène de *dishing*), en raison d'un taux d'érosion plus important par rapport aux surface d'oxyde environnantes. Pour la plage de dimensions considérée, une influence négligeable de la géométrie et de l'agencement des plots d'interconnexion sur la topographie de la surface de collage a été observée. La morphologie de l'interface de collage cuivre-cuivre a ensuite été examinée en imagerie 3D par microscopie électronique à balayage couplée à un faisceau d'ions focalisés (3D FIB/SEM), pour différents profils de surpolissage initiaux des motifs d'interconnexion. Les résultats indiquent l'existence d'une valeur-seuil au-delà de laquelle seul un collage partiel a lieu entre les surfaces métalliques. Même si les défauts de collage observés n'ont pas été identifiés comme pouvant représenter un risque significatif pour la performance électrique du circuit intégré empilé, une fermeture partielle de l'interface cuivre-cuivre peut entraîner des problèmes de robustesse mécanique, comme les précédents travaux l'ont montré (Beilliard, 2015).

Alors qu'une diminution du pas d'interconnexion pour la couche de collage hybride cuivre-oxyde est attendue dans un futur proche, l'influence sur le collage d'un certain nombre de paramètres de conception ou liés au procédé n'est pas encore totalement comprise. Afin d'essayer d'apporter un éclairage sur les principaux paramètres influents pour prévenir le collage partiel, des analyses thermomécaniques du collage cuivre-cuivre pendant le traitement thermique de recuit de consolidation ont été conduites à l'aide de simulations par la méthode des éléments finis. Un résultat important de cette étude est l'observation d'une forte dépendance du collage à la courbure locale du profil de surpolissage, même à profondeur de *dishing* constante. En fonction de la dépendance de la profondeur de surpolissage aux dimensions latérales des motifs, une influence potentiellement néfaste d'une augmentation du facteur de forme des interconnexions, ainsi que de leur densité surfacique, a également été observée. Cette tendance, si elle s'avérait confirmée, pourrait signifier que la compensation thermomécanique du surpolissage des plots d'interconnexions ne serait plus assurée pour une très haute densité d'interconnexions. Les résultats montrent cependant que cet effet néfaste potentiel pourrait être atténué par le choix d'un

couple de matériaux métal/diélectrique maximisant la dilatation verticale du plot de collage. Enfin, l'influence d'un désalignement des plots d'interconnexion au moment de leur mise en contact a également été étudiée. Pour un désalignement suffisamment faible (typiquement, inférieur à la demi-largeur du plot métallique), non seulement une diminution de l'aire relative de contact n'a pas été observée, mais en réalité la fermeture de l'interface cuivre-cuivre se trouve facilitée dans la région de contact restante. Cette étude numérique a donc permis d'identifier les principaux paramètres influents pour la qualité de collage, et de fournir des recommandations pour tenter de limiter les cas de collages partiels observés dans l'étude expérimentale.

Nous nous sommes ensuite intéressés à la robustesse mécanique des plots d'assemblage filaire du capteur d'image empilé, servant à réaliser la connexion électrique entre la puce et le substrat du boîtier dans lequel elle est encapsulée. Ce type d'architecture pour un circuit imageur éclairé par la face arrière donne lieu à de nouveaux défis. En effet, cette configuration impose de former les connexions filaires sur la face arrière de la puce, avec pour conséquence directe la présence des niveaux de métallisation inférieurs à proximité de la surface de connexion. Or, ces couches d'interconnexion locales contiennent les lignes de métal aux dimensions les plus faibles en comparaison du reste de l'empilement d'interconnexions, et l'on peut donc s'attendre dans cette nouvelle configuration à un risque de fissuration ou de délaminage accru. Dans ces travaux de thèse, une comparaison expérimentale de différentes architectures d'empilements d'interconnexion sous les plots d'assemblage filaire est réalisée, au moyen d'inspections de la surface de connexion après le procédé de soudure thermosonique (*cratering test*) ou de coupes de l'empilement d'interconnexion pour détecter de potentielles fissures ou délaminations. Notre attention s'est portée sur l'influence de l'agencement des lignes d'interconnexions sous la connexion filaire, ainsi que de l'introduction d'une couche intermédiaire entre la surface de connexion et l'empilement, sur la tenue mécanique. Dans le but de fournir une meilleure compréhension des résultats expérimentaux, une analyse mécanique multi-échelle d'un test de qualification standard des connexions filaires a été réalisée par la méthode des éléments finis, à savoir le test dit de *wire pull*. Ce modèle a ensuite été étendu dans le but d'étudier l'influence d'un certain nombre de configurations de plot d'assemblage filaire sur la tenue mécanique, et permettre d'établir des recommandations pour diminuer le risque de défaillance mécanique.

La plupart des fissures et délaminations observées dans l'étude expérimentale de l'intégrité mécanique des empilement d'interconnexion après le procédé d'assemblage filaire sont concentrées dans les niveaux de métallisation inférieurs, au plus près de la surface de connexion. L'introduction d'un bicouche épais Si/SiO_2 entre la surface de connexion et l'empilement d'interconnexions a permis d'améliorer considérablement la tenue mécanique, en agissant comme une « couche tampon » absorbant le chargement exercé sur les interconnexions lors du procédé de soudure thermosonique. Une nette influence de l'épaisseur de cette couche de recouvrement a également été observée, avec une suppression complète des défaillances mécaniques pour la configuration correspondant à l'épaisseur la plus importante.

Afin d'examiner plus en détail l'influence de l'agencement des lignes d'interconnexion présentes sous les plots d'interconnexion sur la tenue mécanique, un modèle éléments finis d'un test standard de qualification mécanique des connexions filaires a été développé en suivant une approche multi-échelle, permettant de tenir compte avec une plus grande précision des caractéristiques géométriques de l'empilement d'interconnexions. Un critère fondé sur la moyenne volumique de la première contrainte principale dans les couches d'oxyde fragiles a été proposé pour permettre une comparaison de la robustesse mécanique des différentes architectures de plots de connexion. Une bonne corrélation avec les résultats expérimentaux précédents a été obtenue, permettant de fournir des recommandations pour optimiser la tenue mécanique des plots d'assemblage filaire. Bien qu'il ait été constaté que l'introduction d'une « couche tampon » sous la surface de connexion reste le paramètre le plus influent, l'agencement des interconnexions sous la surface de connexion joue également un rôle important. En premier lieu, l'augmentation de la densité surfacique des motifs métalliques pour les différentes des couches de l'empilement permet d'améliorer la robustesse mécanique, grâce à une reprise de chargement par les lignes d'interconnexion au profit des fragiles couches d'oxyde adjacentes. De plus, la « connexité » des motifs métalliques a été identifiée comme un paramètre permettant un « renforcement structurel » de l'empilement, en atténuant le chargement exercé sur la matrice d'oxyde. De la même manière, la présence de vias métalliques entre les niveaux d'interconnexions contribue également à augmenter la robustesse mécanique. Le procédé de soudure thermosonique génère de forts niveaux de contraintes dans l'empilement d'interconnexions, et fait partie des étapes de fabrication les plus critiques. Dans cette étude, une analyse détaillée des paramètres les plus influents pour la conception d'un plot d'assemblage filaire en terme de robustesse mécanique a été conduite, permettant de fournir des recommandations pour l'optimisation de ces structures, et d'assurer la connexion filaire par la face arrière du capteur d'image empilé.

Dans la dernière partie de cette thèse, une méthodologie reposant sur des capteurs de contrainte piézorésistifs à base de silicium dopé a été proposée pour permettre un suivi *in situ* des contraintes mécaniques dans les zones actives du substrat silicium du capteur d'image empilé, à la fois pendant la fabrication de la puce imageur et son encapsulation dans un boîtier.

Dans le contexte de l'industrie de la microélectronique, avec des temps de fabrication typiques de l'ordre de plusieurs mois, il est nécessaire de mettre en place un contrôle en ligne des différents paramètres d'intérêt pour le circuit intégré considéré, comme par exemple l'alignement des masques de photolithographie, l'épaisseur des couches, la topographie de surface ou encore les résistances électriques des différents composants. Ces paramètres sont régulièrement suivis au cours de la fabrication des circuits intégrés, et permettent pendant la phase de développement d'identifier d'éventuelles anomalies ou des étapes de fabrication critiques sans avoir à attendre la totalité du cycle, et pendant la phase de production d'identifier les dérives des équipements ou de mettre en place un contrôle statistique des

procédés. Cependant, une partie non-négligeable des problèmes potentiellement rencontrés pendant la fabrication des circuits sont en réalité liés à l'apparition de contraintes mécaniques dans les dispositifs. On peut par exemple citer entre autres le gauchissement des plaquettes de silicium, la propagation de fissures ou de délaminations dans les empilement d'interconnexions, ou même des dérives dans les performances électriques ou optiques des composants semiconducteurs, notamment les photodiodes intégrées dans les matrices de pixels des capteurs d'images. Pour toutes ces raisons, il est donc particulièrement souhaitable de pouvoir disposer d'une méthode permettant un suivi de la répartition et de l'accumulation des contraintes mécaniques dans les dispositifs. L'accès à cette mesure permettrait également de valider et calibrer les modèles numériques pour l'optimisation des structures.

De nombreuses techniques ont été développées permettant la mesure du niveau de contrainte mécanique dans des couches minces, comme par exemple la diffraction des rayons X ou encore la spectroscopie Raman. Cependant, ces méthodes sont relativement longues et coûteuses dans un contexte industriel, nécessitant dans certains cas une préparation d'échantillon poussée et destructive. Une solution plus répandue est l'évaluation de la contrainte à travers des mesures de flèche des plaquettes de silicium sur lesquels les circuits sont fabriqués, effectuées régulièrement au cours des procédés de fabrication. Néanmoins, cette méthode ne donne accès qu'à une contrainte moyenne et macroscopique dans le plan du substrat, et peut ainsi s'avérer limitante pour l'étude de couches hétérogènes ou structurées. Dans ce travail de thèse, des capteurs de contrainte piézorésistifs ont été développés à partir de de résistances de silicium dopé disposées sous forme de rosette, avec pour objectif un suivi *in situ* du niveau de contrainte mécanique dans la zone active du substrat silicium où sont intégrés les composants semiconducteurs du circuit imageur. Pour la première fois à notre connaissance, ces capteurs sont implémentés dans un imageur éclairé par la face arrière empilé sur une puce-support par un procédé de collage hybride cuivre-oxyde.

Le principe de ces capteurs de contrainte repose sur la mesure des variations de résistance électrique au cours des différents procédés de fabrication, les variations de contraintes associées pouvant être estimées à condition de disposer des coefficients piézorésistifs des résistances de silicium dopé. Pour réaliser cette étape de calibration, un banc de flexion quatre-point instrumenté précédemment développé à STMicroelectronics a été utilisé. Même si les tendances obtenues pour les différents coefficients sont en accord avec celle rapportées dans la littérature, leur valeur absolue est en revanche bien plus faible en raison des forts niveaux de dopages utilisés pour le circuit imageur considéré, avec des concentrations en impuretés supérieures à $10^{20}\,\mathrm{cm}^{-3}$ pour les résistances du capteur de contrainte. Une conséquence directe de ces forts niveaux de dopage est une sensibilité réduite aux variations de contraintes, et une forte surestimation des valeurs mesurées. De plus, une possible influence significative de la température a également été identifiée. En effet, les variations de résistances dues à des fluctuations de température entre les différentes mesures ont été estimées comme pouvant conduire à elles seules à des variations de contrainte comparables

aux niveaux de chargement attendus pour la fabrication d'un circuit intégré typique, à savoir près de 50 MPa pour 0.5 °C. Cette forte dépendance à la température suspectée pour les capteurs considérés nous a conduit à travailler avec des combinaisons linéaires de composantes de contrainte, afin de permettre une compensation en température.

Pour examiner plus en détail ces biais potentiels dans l'évaluation *in situ* des contrainte mécaniques dans les zones actives de la puce, ces résultats expérimentaux ont été comparés à des simulations thermomécaniques des procédés de fabrication et d'encapsulation du capteur d'image empilé par la méthode des éléments finis. Des valeurs nettement inférieures aux mesures ont en effet été obtenues, de l'ordre de quelques dizaines de mégapascals. De plus, un bien meilleur accord entre les valeurs mesurées et simulées a été observé en considérant les composantes de contrainte compensées en température, même si certaines valeurs inexplicablement élevées ont été constatées en particulier pour l'étape d'encapsulation et nécessitent des analyses complémentaires. Ainsi, des niveaux de dopage modérés à faible sont recommandés afin d'augmenter la valeur des coefficients piézorésistifs, et donc la sensibilité des capteurs aux variations de contraintes. Néanmoins, les valeurs de contrainte relativement faibles obtenues par simulation, à la fois pour la fabrication et l'encapsulation du circuit, tendent à montrer une influence modeste sur le fonctionnement des photodiodes de la configuration *empilement vertical avec éclairage par la face arrière* d'un point de vue thermomécanique pour le capteur d'image considéré.

Dans cette thèse, l'intégrité et la robustesse mécanique d'un capteur d'image éclairé par la face arrière empilé sur un circuit logique par collage hybride cuivre-oxyde a été étudiée. Un certain nombre de défaillances potentielles liées à cette architecture relativement nouvelle pour une application industrielle ont été abordées, permettant d'assurer d'un point de vue thermomécanique l'intégration de ce dispositif innovant.

Titre: Étude expérimentale et numérique des contraintes mécaniques dans les architectures tridimensionnelles sur silicium pour les applications d'imagerie

Mots-clés: *Capteurs d'image CMOS - Collage hybride - Assemblage filaire - Capteurs de contraintes piezorésistifs - Simulation par éléments finis*

Résumé: Ces dernières années, un certain nombre de barrières physiques ou économiques ont fait leur apparition dans la course pour la miniaturisation et l'amélioration des performances des circuits intégrés. Pour dépasser ces limites, de nouvelles architectures sont continûment développées. En particulier, on observe un tournant dans l'industrie de la microélectronique vers les stratégies d'intégration 3D, comme une alternative à la réduction des dimensions des transistors MOS. Cette approche innovante consiste à combiner en un seul et même module des puces de technologies ou fonctionnalités diverses. Une stratégie possible pour réaliser ces systèmes hétérogènes est d'empiler verticalement les puces les unes sur les autres plutôt que de les juxtaposer dans le plan, permettant des gains considérables en terme de compacité et de polyvalence des circuits. Ceci vaut en particulier pour les capteurs d'image, pour lesquels l'exploitation de la dimension verticale rend possible l'incorporation de fonctionnalités supplémentaires, notamment pour le traitement d'image. Parmi les nombreuses méthodes existantes pour réaliser des interconnexions verticales directes entre les puces empilées, le collage « hybride » cuivre/oxyde est une approche prometteuse permettant de réaliser simultanément la connexion mécanique et électrique, avec un pas d'interconnexion submicronique car limité principalement par la précision d'alignement atteignable entre les plots de collage métalliques au moment de leur mise en contact. Un enjeu majeur pour ce type d'architecture innovante est la tenue mécanique des éléments de connexion électrique. Cette thèse vise à examiner la robustesse mécanique d'un capteur d'image reporté sur un circuit logique de technologie plus avancée par empilement 3D, dans le but de prévenir un certain nombre de problèmes potentiels causés par les contraintes thermomécaniques s'accumulant pendant sa fabrication. Dans ce travail de thèse, les contraintes mécaniques générées dans le capteur d'image empilé pendant l'élaboration du circuit et son encapsulation dans un boîtier à puce sont examinées, et les interactions entre les différents composants du système analysées. L'intégrité mécanique de plusieurs structures clés est étudiée, notamment : (i) les plots d'interconnexion à l'interface de collage « hybride » entre la puce imageur et la puce logique, (ii) les plots d'assemblage filaire faisant le lien entre le capteur d'image empilé et le substrat du boîtier, ainsi que (iii) les composants électroniques dans la zone active du substrat silicium du capteur d'image, à travers l'évaluation in-situ des contraintes mécaniques induites par les procédés de fabrication grâce à des capteurs de contraintes piézorésistifs à base de silicium dopé. Pour ce faire, une approche combinant caractérisations expérimentales et analyses numériques a été adoptée : les mesures morphologiques, mécaniques et électriques effectuées sont systématiquement corrélées et étendues à l'aide de simulations par la méthode des éléments finis, permettant de garantir la bonne intégration des produits d'imagerie du point de vue thermomécanique.

Title: Numerical and Experimental Investigations on Mechanical Stress in 3D Stacked Integrated Circuits for Imaging Applications

Keywords: *CMOS Image Sensors - Hybrid Bonding - Wire Bonding - Piezoresistive Stress Sensors - Finite Element Modeling*

Abstract: In recent years, a number of physical and economical barriers have emerged in the race for miniaturization and speed of integrated circuits. To circumvent these issues, new processes and architectures are continuously developed. In particular, a progressive shift towards 3D integration strategies is currently observed in the semiconductor industry as an alternative path to further transistor downscaling. This innovative approach consists in combining chips of different technologies or different functionalities into a single module. A possible strategy to realize such heterogeneous systems is to stack chips on top of each other instead of tiling them on the plane, enabling considerable benefits in terms of compactness and versatility, but also increased performance. This is especially true for image sensor chips, for which vertical stacking allows the incorporation of additional functionalities such as advanced image signal processing. Among various methods to achieve direct vertical interconnections between stacked chips, a promising method is Cu/SiO2 hybrid bonding, enabling simultaneous mechanical and electrical connection with a submicron interconnection pitch mostly limited by photolithography resolution and alignment accuracy. The mechanical integrity of the different electrical connection elements for such a 3D integrated imager-on-logic device is of critical importance. The aim of this thesis is to investigate the mechanical robustness of this relatively new architecture in semiconductor manufacturing during its fabrication, aiming to address a number of possible issues from a thermomechanical perspective. In this work, thermomechanical stresses building up in the image sensor during chip processing and assembly onto a package are investigated, and the interactions between the different system components analyzed. The mechanical integrity of several key structures is studied, namely (i) interconnection pads at the hybrid bonding interface between the imager/logic chips, (ii) bondpad structures below the wires connecting the imager to the package substrate, and (iii) semiconductor devices in the image sensor, through in-situ evaluation of process-induced mechanical stresses using doped Si piezoresistive stress sensors. To do so, for each item a combined numerical and experimental approach was adopted, using morphological, mechanical and electrical characterizations, then correlated or extended by thermomechanical finite element analyses, allowing to secure product integration from a thermomechanical perspective.

Titre: Étude expérimentale et numérique des contraintes mécaniques dans les architectures tridimensionnelles sur silicium pour les applications d'imagerie

Mots-clés: *Capteurs d'image CMOS - Collage hybride - Assemblage filaire - Capteurs de contraintes piezorésistifs - Simulation par éléments finis*

Résumé: Poursuivant la course vers la miniaturisation et la performance des produits de la microélectronique, de nouvelles architectures sont continûment développées. Nous nous intéressons ici à des procédés d'assemblage émergents appliqués à des produits spécifiques. D'une part, les architectures sur silicium tridimensionnelles sont de plus en plus utilisées aux dépens des empilements planaires. D'autre part, afin d'améliorer la performance des imageurs CMOS, ces dernières années la technologie d'éclairage par la face arrière a été développée, devant répondre à des spécifications d'intégration et de fonctionnement particulières. Un procédé innovant consiste dans l'empilement de ce type d'imageur sur un circuit logique de technologie plus avancée, par connexion directe de plots métalliques. Ces travaux portent sur l'évaluation des contraintes thermomécaniques engendrées lors des procédés de fabrication pour cette architecture émergente.

****

Title: Numerical and Experimental Investigations on Mechanical Stress in 3D Stacked Integrated Circuits for Imaging Applications

Keywords: *CMOS Image Sensors - Hybrid Bonding - Wire Bonding - Piezoresistive Stress Sensors - Finite Element Modeling*

Abstract: Pursuing the trend towards miniaturization and increased performance of integrated circuits, new processes and architectures are constantly developed. In this thesis, we are interested in emerging assembly processes applied to specific products. On the one hand, 3D integrated circuits are becoming an increasingly viable approach to enable continuous downscaling, instead of conventional planar integration. On the other hand, to improve the performance of CMOS image sensors, in recent years the backside-illumination approach was developed, involving specific integration and operation requirements. An innovative approach consists in the vertical stacking of such image sensors on a more advanced logic circuit using direct bonding of hybrid metal/oxide surfaces. This work investigates thermomechanical stress build-up during fabrication processes for these emerging architectures.