POLITECNICO DI TORINO

Master of Science in Electronic Engineering

Master Thesis

Realization of a flip chip bonding technique for the assembly of sub-mm neural electrodes

Supervisor
prof. Danilo Demarchi

Candidate
Gabriele Lodi

UC Berkeley Supervisor
prof. Michel Maharbiz

April 2018

Abstract

Brain-Machine Interfaces (BMI) have experienced a huge growth of interest in the international scientific community during the past years; this is certainly due to the revolutionary nature of these technologies, which could have a wide variety of applications, especially in biomedicine, since they can allow to study the brain, treat neurological diseases and even control prosthesis or robotic arms for disabled people. The main goal of these technologies is indeed to sense the electrical signals generated by the neural activity inside the brain and transmit the information outside of the body. To accomplish this goal, some BMI (called "invasive") make use of tiny implantable electrodes to allow for precise electrical recording from a small number of neurons while causing low tissue damage.
The work presented in this thesis regards the realization of a flip chip bonding technique used for the assembly of sub-mm neural electrodes. This bonding technique was preferred to wire bonding since it allows to minimize the device dimensions. The flip chip bonding has been realized using a gold-gold thermocompression technique, which consists in the application of heat and force between the chip and the substrate, both having gold pads with the same layout; thanks to the applied heat and force, the thin gold layers on both sides interdiffuse, thus creating a permanent bond when they cool down.
The realization of this flip chip bonding technique required:

- gold deposition on the aluminum chip pads, for which the electroless ENIG plating process was used
- substrate fabrication using lift-off process
- flip chip bonding with different forces and temperatures
- electrical and mechanical testing of the bonded chips.

Sommario

Le Interfacce Cervello-Macchina hanno sperimentato un'enorme crescita di interesse da parte della comunità scientifica internazionale nel corso degli ultimi anni; questo è sicuramente dovuto alla natura rivoluzionaria di queste tecnologie, che potrebbero avere numerose applicazioni, specialmente in ambito biomedicale, siccome possono permettere di studiare il cervello, trattare patologie neurologiche e addirittura controllare protesi o braccia meccaniche per persone con disabilità. L'obiettivo principale di queste tecnologie è infatti quello di registrare i segnali elettrici generati dall'attività neurale all'interno del cervello e trasmettere l'informazione all'esterno del corpo. Per raggiungere questo obiettivo, alcune BMI (dette "invasive") fanno uso di piccolissimi elettrodi impiantabili che quindi permettano di registrare l'attività elettrica con precisione da un piccolo numero di neuroni, allo stesso tempo causando pochi danni ai tessuti .
Il lavoro presentato in questa tesi riguarda la realizzazione di una tecnica di flip chip bonding per l'assemblaggio di elettrodi neurali micrometrici. La tecnica di flip chip bonding è stata preferita a quella di wire bonding per permettere di minimizzare le dimensioni del dispositivo. Il flip chip bonding è stato realizzato utilizzando la tecnica di termocompressione fra due strati di oro (gold-gold thermocompression): questo consiste nell'applicazione di forza e calore fra il chip ed il substrato, sui quali è necessario acere contatti elettrici in oro con lo stesso layout; grazie al calore ed alla forza applicata, i sottili strati d'oro presenti sui due lati si fondono assieme (inter-diffusione), formando poi un legame permanente dopo il raffreddamento.
La realizzazione di questa tecnica di flip chip bonding ha richiesto:

- deposizione di oro sui contatti metallici in alluminio del chip utilizzando un process di electroless ENIG plating
- fabbricazione del substrato utilizzando un processo di lift-off
- flip chip bonding con diverse forze e temperature
- testing elettrico e meccanico dei chip.

Ai miei genitori

Acknowledgements

First and foremost, I want to thank my parents for having always backed me in my decisions and motivated me during my path, for having supported me both economically and emotionally, for having always believed in me and especially for having raised me according to the important values of respect, honesty and work.

I want to thank my supervisor, prof. Danilo Demarchi, for his help and support, but especially for providing me with the contact of prof. Michel Maharbiz at UC Berkeley for the thesis activity. In regard to this, I want to profoundly thank prof. Michel Maharbiz for offering me the opportunity to join his laboratory and for trusting me and advising me during the six-month period of research I spent there, always treating me as one of his graduate students and making me feel part of a awesome group. Among these students, I want to particularly thank Konlin Shen for allowing me to work on his project and for his help and advices, allowing me to learn a lot of useful things. Moreover, I want to thank all the other graduate students of the group for their help, support and friendship, especially Travis Massey, Tom Zajdel and Bochao Lu for their precious and prompt advices regarding the work in the laboratory.

I want to thank my brother and my closest relatives for encouraging and supporting me; moreover, I thank my best friends, especially the ones at university, for the time spent together during the past two years and for having listened to me and advised me in many situations, in particular Yuri and Martina, with whom I spent most of the time during the master degree and with whom I carried out many labs and projects. Finally, I want to thank all my friends at Berkeley for their support during the toughest periods and for all the unforgettable great times we had and all the trips we took together.

Ringraziamenti

Ringrazio prima di tutto e sopra a tutti i miei genitori per avermi sempre supportato nelle mie scelte e motivato nel mio percorso, per avermi sostenuto sia economicamente che emotivamente, per aver creduto in me in ogni situazione e soprattutto per avermi cresciuto con gli importanti valori del rispetto, dell'onestà e del lavoro.

Ringrazio il mio relatore, il prof. Danilo Demarchi, per il suo aiuto e supporto, ma soprattutto per avermi fornito il contatto del prof. Michel Maharbiz a UC Berkeley per l'attività di tesi. A questo proposito, ringrazio il prof. Michel Maharbiz per avermi offerto l'opportunità di lavorare nel suo laboratorio e per avermi dato fiducia e consigliato durante i sei mesi di ricerca svolti, trattandomi alla pari dei suoi dottorandi e facendomi sentire parte di un gruppo meraviglioso. Fra questi studenti, voglio ringrazio in particolare Konlin Shen per avermi permesso di lavorare sul suo progetto e per avermi aiutato e consigliato, permettendomi di imparare molte cose utili. Inoltre, ringazio tutti gli altri dottorandi del gruppo per l'aiuto, il sostegno e l'amicizia, specialmente Travis Massey, Tom Zajdel e Bochao Lu per i loro preziosi e pronti consigli riguardanti il lavoro nel laboratorio.

Ringrazio mio fratello ed i miei parenti più stretti per l'incoraggiamento e il sostegno che mi hanno dato; inoltre, ringrazio i miei migliori amici, soprattutto quelli dell'università, per il tempo trascorso insieme negli ultimi due anni e per avermi ascoltato e consigliato in tanti momenti, in particolare Yuri e Martina, con i quali ho trascorso gran parte del tempo nel corso della laurea magistrale e con i quali ho svolto molti laboratori e progetti. Infine, ringrazio tutti gli amici di Berkeley per avermi sostenuto nei momenti più duri e per tutti gli indimenticabili momenti trascorsi e i viaggi fatti insieme.

Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

UC	University of California
BMI	Brain-Machine Interface
ENIG	Electroless Nickel-Immersion Gold
BCI	Brain Computer Interface
RMP	Resting Membrane Potential
EEG	Electroencephalography
ECoG	Electrocorticography
MEG	Magnetoencephalography
SQUID	Superconducting Quantum Interference Device
INR	Intracortical Neuron Recordings
fMRI	Functional Magnetic Resonance Imaging
CAT	Computed Axial Tomography
SPECT	Single Photon Emission Tomography
PET	Proton Emission Tomography
EIT	Electrical Impedance Tomography
ND	Neural Dust
EECS	Electrical Engineering and Computer Science
EM	Electro-Magnetic
FDA	Food and Drug Administration
SNR	Signal-to-Noise Ratio
RF	Radio Frequency
ASIC	Application Specific Integrated Circuit
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
NMOS	N-type MOSFET
AC	Alternating Current
PCB	Printed Circuit Board
WB	Wire Bonding
TAB	Tape Automated Bonding
FCB	Flip Chip Bonding
TCP	Tape Carrier Packaging
ILB	Inner Lead Bonding
OLB	Outer Lead Bonding
C4	Controlled Collapse Chip Connection
UBM	Under Bump Metallization
CTE	Coefficient of Thermal Expansion
MCM	Multi-Chip Module
SCM	Single Chip Module
IC	Integrated Circuit
ICA	Isotropically Conductive Adhesive
ICF	Isotropically Conductive Film

ACA	Anisotropically Conductive Adhesive
ACF	Anisotropically Conductive Film
SEM	Scanning Electron Microscopy
IPA	Isopropyl Alcohol
EN	Electroless Nickel
IG	Immersion Gold
TSMC	Taiwan Semiconductor Manufacturing Company
ESD	Electrostatic Discharge
TSP	Trisodium-phosphate
PP	Polypropylene
DI	De-Ionized
RT	Room Temperature
PR	Photoresist
HMDS	Hexamethyldisilazane
UV	Ultraviolet

Chapter 1

Introduction to Brain-Machine Interfaces

Brain-Machine Interfaces (BMI), also known as Brain-Computer Interfaces (BCI), allow humans to interact with the surrounding environment through control signals directly extracted from the brain electrical activity, without exploiting the body muscular system through the nerves. These technologies are particularly interesting for people affected by neuromuscular diseases, since they allow them to directly communicate with a computer, giving them access to everything that is controllable through a computer: information and entertainment (such as books, newspapers, movies, music), communication tools (such as emails and chats), education and research resources and capabilities (such as simulation softwares, text editors, video lectures, scientific journals) and much else [1].
The main tasks performed by a BMI are the acquisition of the neural signal and its processing, which generally requires a pre-processing step (filtering, amplification, digitalization, etc.), followed by a feature extraction and classification step, before the processed information is sent to the final hardware application. The general block diagram of the BMI operations is reported in Figure 1.1.

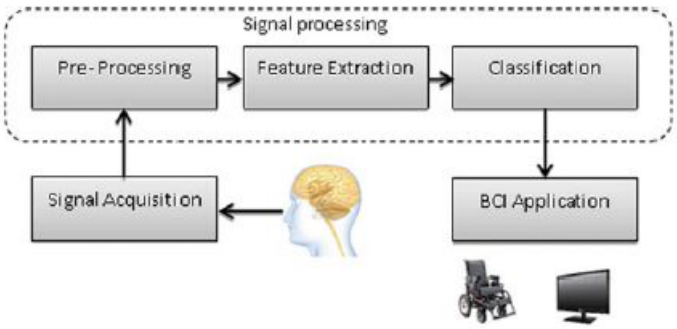


Figure 1.1: General block diagram of a BMI [2]

In order to better understand the brain functioning and how to extract the information on the neural activity, it is necessary to study the brain "architecture" and "wiring" from

a system standpoint, starting from its fundamental components and how they generate and exchange signals.

1.1 The Neuron

The human brain is composed by over 10^{10} neurons [3], which are thus the brain fundamental entities. The neuron is composed of different parts, as shown in Figure 1.2 [4]:

- **dendrites**, which are the inputs of the neuron and are used to collect all the incoming signals from other neurons
- **soma**, which is the core of the neuron, where operations are performed
- **axon**, which is the channel used to transmit the output signal from the neuron to the outputs by means of "action potentials"
- **axon terminals**, which are the outputs of the neuron, since they are connected to dendrites of other neurons through small gaps called "synapses".

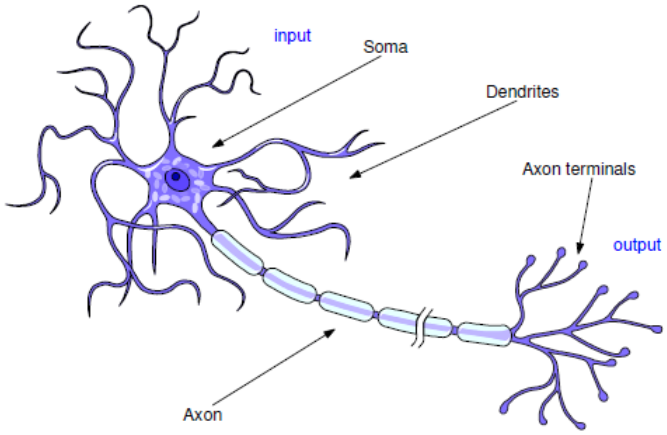


Figure 1.2: The neuron [4]

The neuron may be seen as a sort of transducer which is sensitive to chemical, electrical and mechanical stimuli, as represented in Figure 1.3.

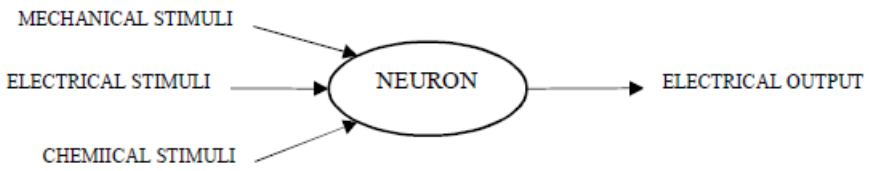


Figure 1.3: System representation of a neuron as a transducer

Actually, the neuron is also able to perform some operations: it would be more appropriate to define it as a system with inputs and outputs [4], as represented in Figure 1.4.

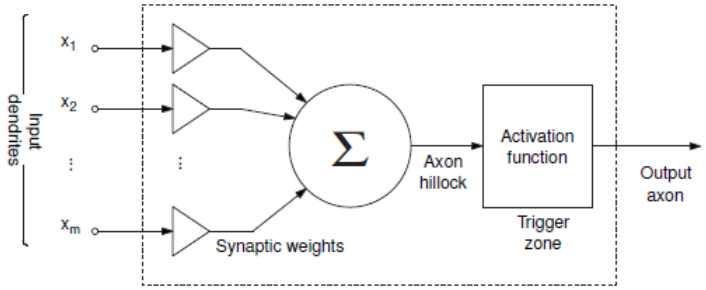


Figure 1.4: Input/output functional model of a neuron [4]

1.1.1 Electrical activity of the neuron

A neuron can be thought of as a sort of cell embedded in a particular ambient: the wall of this cell is a permeable membrane, whose properties may be changed according to electrochemical forces. Inside and outside the cell there are different concentrations of potassium, chloride and sodium ions [4], as shown in Figure 1.5.

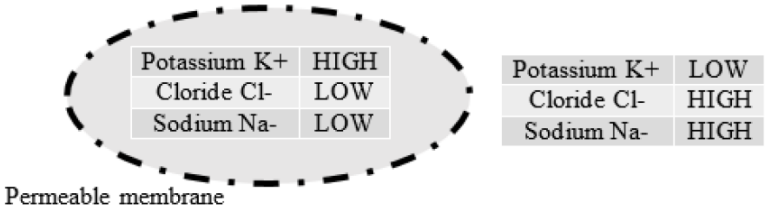


Figure 1.5: Ions concentration inside and outside the neuron

At equilibrium, the electrical potential resulting from the combination of these concentrations is about -70 mV, value known as "Resting Membrane Potential" (RMP). Two different domains are involved in the definition of this equilibrium [4]:

- **electrical**, since electrical forces try to minimize the voltage potential by moving positive ions toward negative ions, going toward the electrical equilibrium
- **chemical**, since chemical forces tend to uniform the concentration by moving ions through the membrane until outer and inner concentrations are equal, going toward the chemical equilibrium.

The final result is a situation of equilibrium between these two domains, which is called "electrochemical equilibrium". The electrochemical gradient across the cell membrane is maintained by a sort of pumping system, called "sodium-potassium pump", which continuously keeps both the inner concentration of potassium ions and the outer concentration of sodium/chloride ions high; indeed, without this pump the gradient would level

out and no information transfer to other neurons would be possible [4]. As already mentioned, neurons are sensitive to mechanical, electrical and chemical stimuli, which affect the membrane permeability, allowing ions to flow in and out the cell [4]:

- if the flow is low, the Na^+/K^+ pump quickly removes the disturbance
- if the flow is high, the Na^+/K^+ pump is not fast enough to remove the disturbance, then the resting potential changes.

If this potential exceeds -55 mV, a special protein opens the membrane to Na^+ ions for a fraction of a millisecond, allowing them to enter into the cell. The consequence is a rapid increase of the membrane potential to about +30 mV, called "depolarization". After a very short time, another protein opens up the membrane to K^+ ions, thus resetting the membrane potential to the RMP level: this phase is called "repolarization" [4]. The electrical impulse generated by this ions flow is called "action potential", which is shown in Figure 1.6.

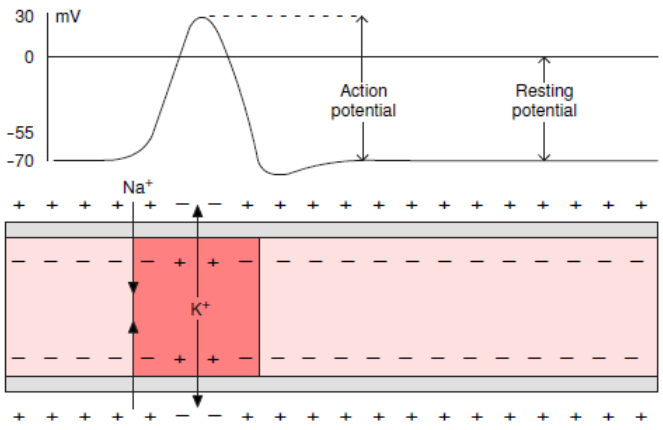


Figure 1.6: Action potential generated by a flow of ions through the cell membrane [4]

During the depolarization state, potassium ions are forced to exit the neuron, thus there is an the excess of sodium ions in the cell that does not provide a stable state, due to electrical repulsion; the excess of positive ions is thus forced to propagate through the axon, creating the information transfer as Na^+/K^+ waves. The representation of the conduction is shown in Figure 1.7.

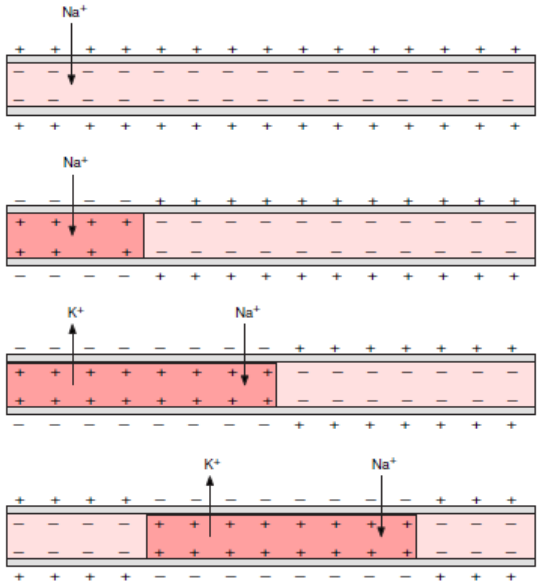


Figure 1.7: Propagation of the action potential along the axon [4]

This type of propagation, which is typical of muscle fibers, is called "continuous propagation", since it has an almost constant velocity in the range of 0.2-2 m/s. The majority of axons in the nerve system are instead characterized by another type of propagation, which is called "saltatory propagation", and it is much faster than the other one, with speeds in the range of 12-120 m/s [4].

1.1.2 Neural signals

Depending on the brain activity state it is possible to individuate different types of brain signals, also called "rhythms" or "waves" [4], which are shown in Figure 1.8:

- **Alpha**: frequency range 8 ÷ 13 Hz and amplitude of about 50 µV. They are associated with a condition of relaxation and can be detected at the back of the head on both sides

- **Beta**: frequency range 13 ÷ 30 Hz and low voltage with maximum amplitude of 5 ÷ 30 µV. They are associated with a condition of thinking and concentration on solving problems

- **Theta**: frequency range of 4 ÷ 8 Hz and amplitude generally less than 100 µV. Theta waves are very slow and are associated with emotional stress and frustration.

- **Delta**: frequency range of 0.5 ÷ 4 Hz, it is the slowest wave with the highest amplitude, sometimes higher than 100 µm; it can be seen both in adults and babies during deep sleep.

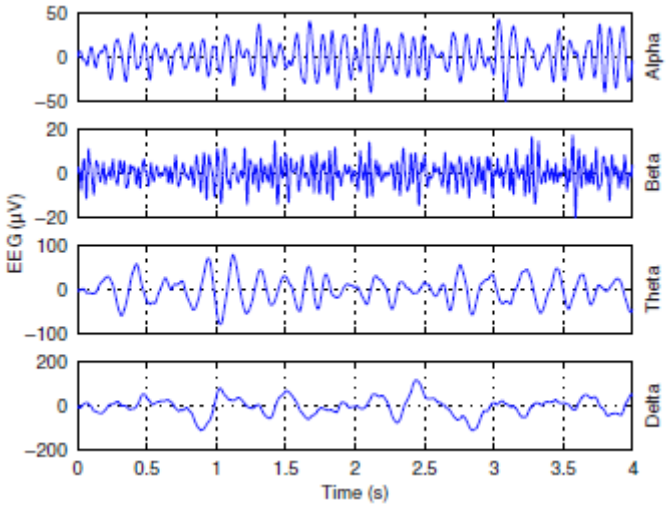


Figure 1.8: Neural signals shapes obtained with EEG [4]

1.1.3 Fundamental principle of neural measurements

In order to figure out how to measure the neural potential, it is useful to consider that a neuron, when triggered, may be seen as an electrical system with two charges (the ions) and thus modelled as an electrical dipole [4], which is represented in Figure 1.9. For an electrical dipole the expression of the potential is:

$$V_P(\underline{r}) = \frac{q}{4\pi\epsilon_0 ||r^3||}(\underline{d} \cdot \underline{r}) = \frac{qdcos\alpha}{4\pi\epsilon_0 r^2}$$

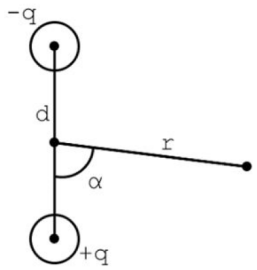


Figure 1.9: Electric dipole

Even if this is just an approximation, since the neuron is not a real electrical dipole and, moreover, in many acquisition techniques between neural sensors and neurons there is the cortex, which is not taken into account in the derivation of the previous voltage expression, nonetheless two important considerations may be derived from the previous formula [4]:

- the cosine term indicates that neural sensors cannot sense the signal coming from neurons placed in particular positions (indeed, if $\theta = \pi + \kappa\pi$ the voltage is null),

which means that many neurons are hidden, in the sense that their potential cannot be measured

- the voltage quadratically depends on the distance, which means that the further is the electrode, the lower will be the potential.

1.2 Neural signals: acquisition techniques

There are many different techniques used to extract information on the brain activity; some of the most important are [5]:

- **Electroencephalography (EEG)**
  EEG consists in the application of recording electrodes on the patient's head, thus allowing to sense brain electrical signals in the form of a polygraph, as shown in Figure 1.10. With respect to other techniques, EEG is a low cost and non-invasive technique; moreover, it is relatively easy to use and it has an high temporal resolution, which means it can resolve short time-scaled events (in the order of milliseconds) [6]. Its main problem is the poor spatial resolution. EEG is mainly used to diagnose brain diseases, such as Alzheimer's disease and epilepsy disease, but it can be used also to detect sleeping, learning, and attentional disorders; moreover, it can be used to monitor the brain responses during brain surgery.

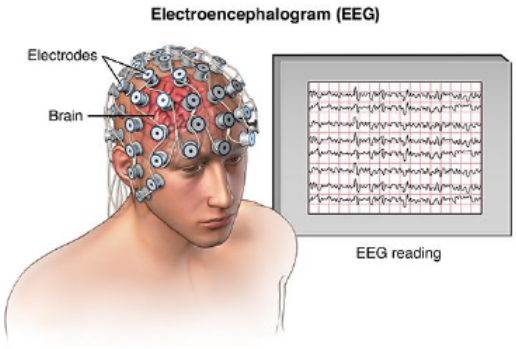


Figure 1.10: An EEG cap equipped with electrodes [7]

- **Electrocorticography (ECoG)**
  ECoG is an invasive technique which allows to record signals directly from the surface of the brain, placing electrodes under the skull, as shown in Figure 1.11. Compared with EEG, ECoG has larger bandwidth, higher spatial resolution and higher output amplitude [8, 9], since the attenuation introduced by the skull is avoided.

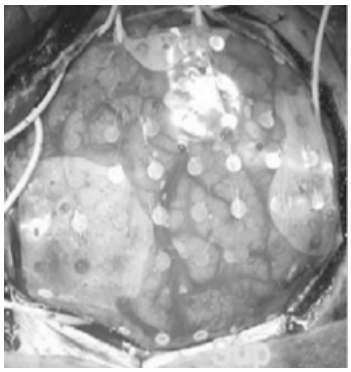


Figure 1.11: Example of electrode placements for an ECoG [10]

- **Magnetoencephalography (MEG)**
  This technique allows to detect the brain activity through the measurement of the magnetic field generated by charged ions, which are excited within neural cells [11]. MEG is a non-invasive technique, which is characterized by high temporal resolution (millisecond resolution) [12]. The device used to detect these magnetic waves is called Superconducting Quantum Interference Device (SQUID) [13]. MEG is used to diagnose neuronal diseases such as Alzheimer's Disease, but it can be used also for children and infants [14], as shown in Figure 1.12, since it is a patient-friendly diagnosis system.

Figure 1.12: MEG system used on a child [15]

- **Intracortical Neuron Recording (INR)**
  Intracortical Neuron Recording is used to record the electrical activity inside the gray matter, using a microelectrode array implanted in the brain cortex, as represented in Figure 1.13. INR is being used in several research groups to try restoring the motor functions in people suffering from some form of paralysis or with limb loss, by establishing a connections between the motor cortex and an external device such as a robotic arm [16].

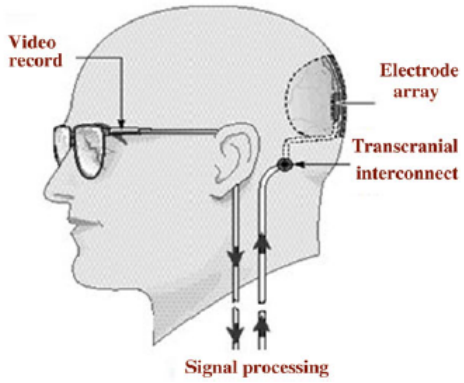


Figure 1.13: Intracortical Neuron Recording system [17]

- **functional Magnetic Resonance Imaging (fMRI)**
  It is a technique based on Magnetic Resonance Imaging (MRI), which is used to extract an activation map indicating the brain activity. This is done by measuring the changes in the blood flow and oxygenation level deriving from neural activity: indeed, when some regions of the brain are involved in a particular mental process, they require more oxygen, thus the blood flow increases in correspondence of these regions in order to meet the increased demand, as shown in Figure 1.14. fMRI is used to study psychological diseases like dyslexia [18] and autism [19].

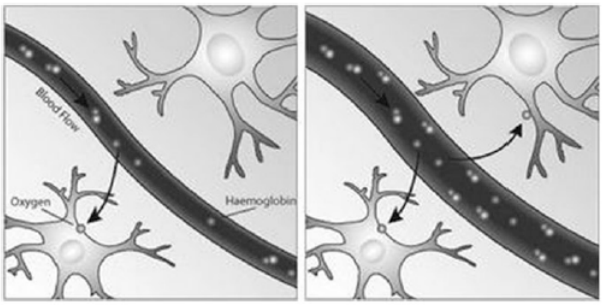


Figure 1.14: Blood flow intensity depends on the brain activity: low level of brain activity (on the left), high level of brain activity (on the right) [1]

- **Tomographic imaging techniques**
  These are non-invasive brain monitoring techniques that allow to obtain a cross-sectional image of the brain by performing different types of measurements [20].
  X-ray Computed Axial Tomography (CAT) [21] uses an X-ray source that rotates around the sample to image; this is placed inside an aperture called "gantry", surrounded by X-ray receptors, so that the X-ray beam passes through it and its attenuated radiation is recorded by the receptors on the opposite side. The spatial distribution of the X-ray attenuation coefficient is then collected and the tridimensional image is reconstructed using an image reconstruction algorithm [22, 23]. The

X-ray CAT mechanism is shown in Figure 1.15.

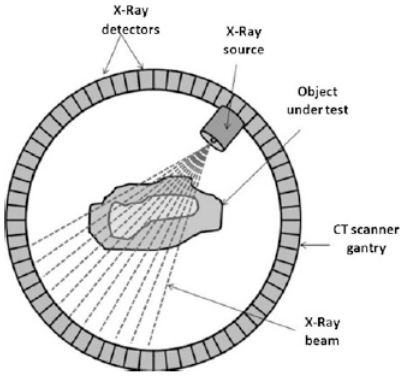


Figure 1.15: X-ray CAT mechanism [20]

Single Photon Emission Computed Tomography (SPECT) [24, 25, 26, 27, 28, 29] is an imaging technique used to study the blood flow inside the brain blood vessels; this technique requires to inject into the patient's body a radioactive pharmaceutical (radioisotope) emitting gamma rays. Then, the radiation is collected by a gamma camera rotating around the patient, which thus allows to reconstruct a tridimensional image.
Positron emission tomography (PET) [30, 31, 32, 29] is based on the same principle of SPECT [28] but using a radioisotope which emits positrons; these immediately annihilate with electrons (less than few millimeters away), thus generating a couple of gamma photons that propagate in opposite directions, as shown in Figure 1.16. The detection of two photons on diametrically opposite receptors allows to identify the position where the positron was emitted [32]. Differently from SPECT, the detected radiation is thus not directly emitted by the radioisotope, but it is generated by the annihilation of the positrons that it emits. The main advantage of PET and SPECT is that they can provide absolute measures of the regional cerebral blood flow, but the downsides are that they are expensive techniques, they require to inject radioactive material into the patient's body and they have a poor temporal resolution [33].

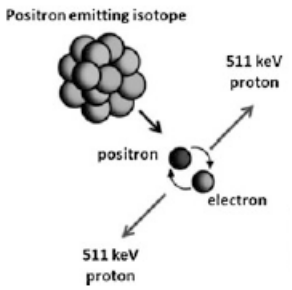


Figure 1.16: Gamma photons generated by positron annihilation [20]

The PET system schematic is reported in Figure 1.17.

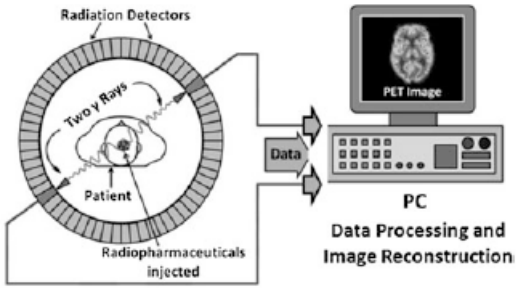


Figure 1.17: PET schematic [20]

Electrical Impedance Tomography (EIT) [34, 35, 36, 37, 38, 39, 40, 41, 42, 43] allows to obtain a 2D or 3D image showing the spatial distribution of the electrical properties of the tested sample, such as conductivity or resistivity [44, 45, 46]. These information are extrapolated from the voltage-current data measured at the domain boundary by an array of surface electrodes using the four electrode method [47, 43], when a constant amplitude, low frequency sinusoidal current is injected through the driving electrode. The EIT system is shown in Figure 1.18. Thus, the recorded data are processed by a computer using an image reconstruction algorithm to obtain the spatial distribution of the electrical properties of the sample under test [48, 49, 50, 51, 52, 53, 54, 55, 56, 57]. Advantages of EIT are the fact that it is a low cost, portable, non-invasive and non-ionizing imaging technique.

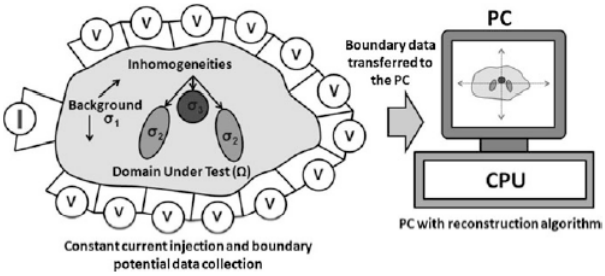


Figure 1.18: EIT schematic [20]

1.2.1 Neural microelectrodes

Microelectrodes are used to record biological electrical potentials, such as action potential, from single cells: for this purpose, they need a thin tip to be inserted into the cell, so they have to be small with respect to a biological cell to avoid damaging it, but also strong enough to penetrate the cell wall [58]. Three types of microelectrodes are typically used [58]:

- **Glass micropipette**

The fabrication process starts from a hollow glass capillary tube (1 mm in diameter), which is heated in the middle using a furnace so that it becomes softer, then it is quickly pulled apart from both ends, thus creating two almost identical microelectrodes having an open tip with a diameter of 0.1 ÷ 10 µm. Then, the largest end of the glass tube is filled with a KCl electrolyte solution and a short piece of Ag/AgCl wire is inserted to provide an electrical contact with the electrolyte solution, as shown in Figure 1.19; in this way, an ionic current starts flowing through the fluid junction at the tip of the microelectrode, thus establishing a closed electrical circuit between the Ag/AgCl wire inside of it and the biological cell. The major issues of glass micropipettes are the complexity of the fabrication process and their high impedance [58].

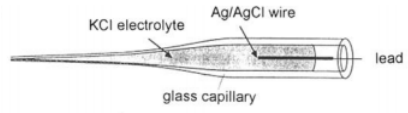


Figure 1.19: Glass micropipette [58]

- **Metal microelectrode**
  It is made of a strong metal wire (e.g., tungsten) with a small diameter; the tip of this microelectrode is usually sharpened down to a diameter of few µm through an electrochemical etching process. Finally, the wire is insulated up to its tip. This type of electrode is very small and robust, but its fabrication is very complex and it has a high source impedance, which is mainly due to the tip impedance [58]. An example of metal microelectrode is represented in Figure 1.20.

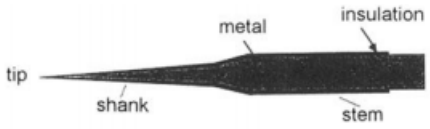


Figure 1.20: Metal microelectrode [58]

- **Solid-state microprobe**
  It consists of a precisely micromachined silicon substrate containing four exposed recording sites, as shown in Figure 1.21, which are usually used for multichannel recordings of biopotentials or for electrical stimulation of neurons [58].

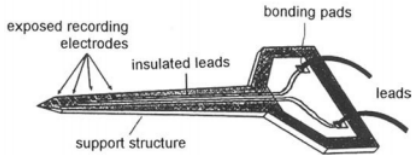


Figure 1.21: Solid-state microprobe [58]

1.3 Neural Dust

The "Neural Dust" project regards the development of an innovative BMI concept, which is carried on in the Maharbiz Lab at the UC Berkeley EECS Department. The aim of this project is to build a wireless chronic implantable technology which allows to record the electrical activity of the nervous system and transmit the information outside the body [59].

In order to address these requirements, it is necessary to consider the principal limiting parameters for implants, which are size and power [59]:

- on the one hand, implants should be as small as possible to minimize tissue damages and to allow for precise recording of the neural activity from a small number of neurons;

- on the other hand, size reduction also means that the recording points become closer to each other, thus measuring smaller voltages; this means that the noise from electronics has to be reduced, requiring higher harvested power, which is in contrast with the fact that smaller devices collect less power.

Electromagnetic waves are the most commonly used method for powering and communication with implants, since they do not require moving parts or thermal/chemical gradients. The main problem of EM is that the propagation of electromagnetic fields causes heating inside tissues: for this reason, the FDA regulations limit the power density to 10 mW/cm^2 [60], which roughly corresponds to the power required to increase a model sample of human tissue by 1 °C, thus avoiding potential negative health effects caused by excessive heating inside the body [61]. As already mentioned, the reduction of dimensions requires a power increase to obtain a reasonable SNR, thus the use of EM poses a limit to the scaling of devices.
For this reason, ultrasounds are used instead, since they are pressure waves, which means they propagate through compression and rarefaction of the propagation medium instead of inducing magnetic and electrical currents as EM waves do. The use of ultrasounds have two main advantages: first, the speed of sound is 10^5x lower than the speed of light in water, leading to much smaller wavelengths at similar frequencies; this yields excellent spatial resolution at these lower frequencies as compared to EM waves [62]. Secondly, acoustic losses in tissues are much smaller than the attenuation of EM waves, as shown in Table 1.1, due to their different propagation mechanisms [59]. This not only results in much higher penetration depths for a given power, but also significantly decreases the amount of unwanted power introduced into the tissue due to scattering or absorption. In fact, for most frequencies and power levels, ultrasound is safe in the human body [62]. The limit imposed by FDA regulations for ultrasound power density is approximately 72x higher (720 mW/cm^2) than the one for EM [60]. This means much higher power can be received at the implants.

	Ultrasound	EM
λ in brain tissue	150 µm @ 10 MHz	5 mm @ 10 GHz
Tissue & path loss at 2 mm	1 dB*	20 dB

Table 1.1: Comparison of power losses in brain tissue for ultrasound and EM waves [61]

Other main innovations that Neural Dust project aims to achieve are:

- **wireless transmission**, which is appealing for avoiding the presence of wires coming out of the body, thus reducing the risk of infections and allowing the patient to move while transmitting data without affecting the signal quality, even if this raises important surgical issues on the correct placement of these devices [59]
- **chronic implantation**, which is important for avoiding the patient to be subjected to frequent surgical operations and providing a lifelong technology, even if this requires the use of perfectly biocompatible materials, not only to avoid the body rejection but also the formation of a scar layer between the device and the tissue, which would degrade the quality of the recorded signal. Moreover, the implanted device has to be perfectly sealed to avoid moisture from entering inside the package, which would risk to deteriorate the device functioning or even break it [61].

1.3.1 ND system configuration

The Neural Dust system is composed by implantable electrodes or sensor nodes, called "motes", and an external transceiver used to establish a bi-directional wireless communication with the implants, both for data transferring and motes powering. This transceiver is also called "interrogator" since it initiates the communication with the implanted motes by sending a wake-up signal (train of pulses), then it waits for a modulated backscattered signal which embeds information on the neural activity [61]; the communication between interrogator and dust motes through the tissue is shown in Figure 1.22.

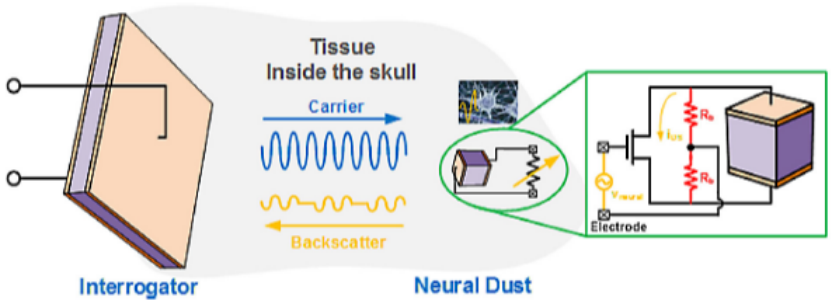


Figure 1.22: Representation of the complete communication cycle between interrogator and ND mote [59]

The system can have two different implementations depending on the application:

- for the **Central Nervous System**, which is essentially the brain, the idea is to use thousands of 10–100 µm scale, free-floating, independent motes, from which the name "neural dust", used to record the local electrical activity. The full ND system is based on a double-link transmission: the external transceiver placed on top of the skull is equipped with battery and memory and it is responsible both for the powering of and data transmission with the sub-dural transceiver via RF through the skull; the sub-dural transceiver, placed directly on the brain cortex to avoid strong signal attenuation due to the skull, is responsible both for the powering of and data transmission with the implanted dust motes via ultrasounds [61]. Figure 1.23 shows the complete system installation in a partial cross section of the head.

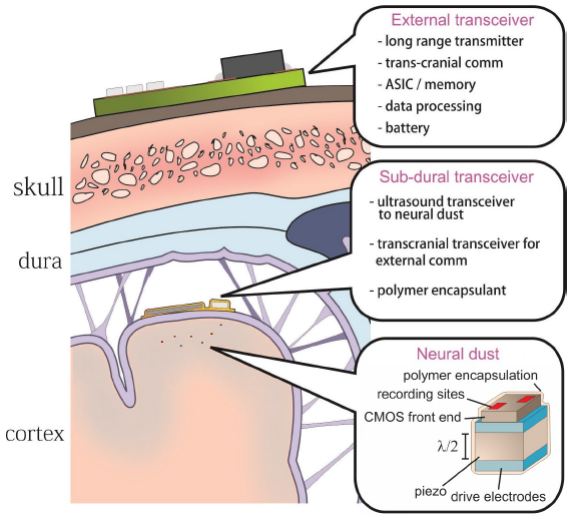


Figure 1.23: ND system complete installation inside the head [61]

- for the **Peripheral Nervous System**, which is composed by all the nerves in the body (bundles of axons), the idea is to use a ND mote attached to a nerve [62], such as the one anchored to the sciatic nerve of a rat that is shown in Figure 1.24.

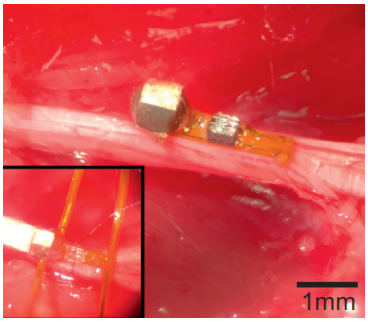


Figure 1.24: ND mote anchored to the sciatic nerve in an anesthetized rat for in-vivo testing [62]

1.3.2 ND motes

Working principle

The Neural Dust motes contain two main components: an ASIC chip and a piezoelectric crystal. The ASIC purpose is to obtain a modification of the dust acoustic impedance related to the neural voltage, so that the backscattered ultrasonic pulse embeds the information of the sensed voltage on the neurons, which can be retrieved through post-processing on the backscattered ultrasonic wave received by the interrogator. The piezoelectric crystal is necessary to harvest the transmitted ultrasonic wave, thus generating a voltage used to power up the ASIC MOSFET. The ASIC layout is reported in Figure 1.25.

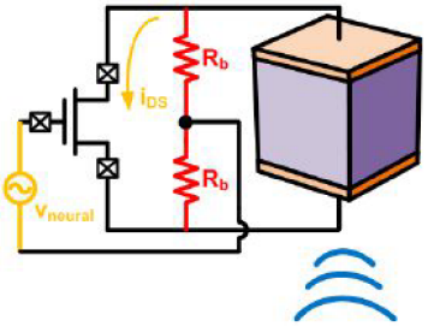


Figure 1.25: ND ASIC schematic [61]

The step-by-step working principle is the following [61]:

1. pulses of ultrasonic energy emitted by the external transducer impinge on the piezoelectric crystal, so that they are partially reflected back and partially cause the piezoelectric crystal to vibrate; thanks to its properties, it converts the mechanical deformation caused by the ultrasonic wave into electrical voltage, which is supplied to the transistor as V_{DS}

2. the sensed neural voltage is applied between the gate of the NMOS and the central point of the resistive branch

3. the change in V_{GS} causes a change in the transistor current I_{DS}, which in turn modifies the charge accumulated on the dust mote plates

4. the modification in the charge on the plates in turn changes the voltage across them, thus affecting the mechanical deformation of the piezoelectric, which is related to its acoustic impedance. This means that the backscattered ultrasonic signal generating from the ultrasound impinging on the piezoelectric has an amplitude which depends on the neural voltage and, more specifically, it is modulated in amplitude through the sensed neural voltage.

The unusual connection shown in Figure 1.25 for the neural voltage is necessary to properly bias the transistor body on the whole cycle of the AC supply voltage, since the

two terminals of the transistor swap sign in every half cycle. Then, since the MOSFET structure is symmetric, source and drain contacts swap their roles, excluding the possibility to physically connect the MOSFET body to either one terminal or the other. Using two resistors with the same value it is possible to modulate the AC voltage during the whole cycle, since the neural voltage is always applied on its central value. The resistor values have to be properly chosen in order to create, together with the parasitic capacitance of the piezoelectric, a low-pass filter able to decouple the low-frequency neural voltage (~kHz) from the high-frequency ND voltage (~MHz) [61].

Assembly

The first mm-sized version of ND mote, already shown implanted in Figure 1.24, is used for recording the electrical activity in the peripheral nervous system [62]. The assembly process is sketched in Figure 1.26: it consists in the attachment of the chip and the piezocrystal to the PCB substrate, which are then wirebonded to provide electrical connection with the PCB metal lines and, ultimately, the whole device is encapsulated in a polymer for biocompatibility [63].

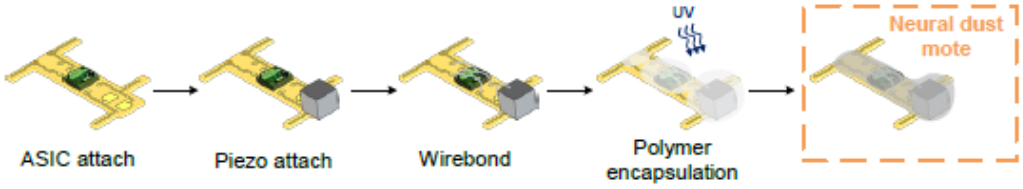


Figure 1.26: ND mote assembly [63]

The final assembled ND mote is sketched in Figure 1.26, while Figure 1.28 shows the real packaged mote from different views, which have dimensions 0.8 mm$times$1 mm$times$3 mm [62].

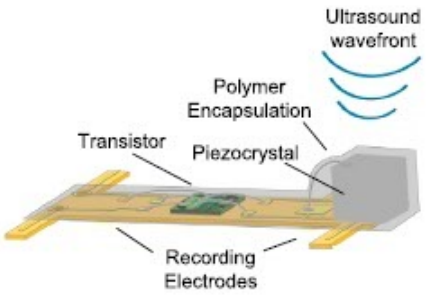


Figure 1.27: Sketch of the assembled ND mote [62]

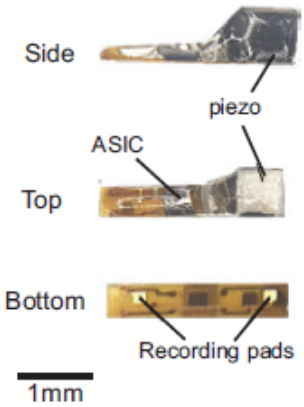


Figure 1.28: Views of the assembled ND mote [62]

Chapter 2

Introduction to chip packaging techniques

The three main techniques used for chip packaging in the microelectronic industry are [64]:

- **Wire Bonding (WB)**

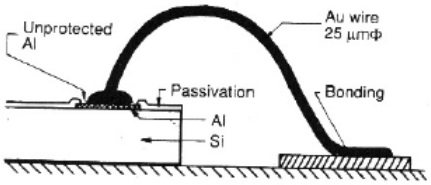


Figure 2.1: Cross-sectional representation of WB [65]

- **Tape Automated Bonding (TAB)**

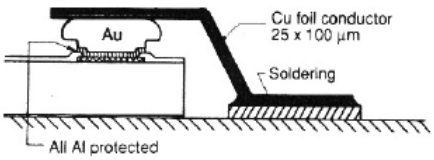


Figure 2.2: Cross-sectional representation of TAB [65]

- **Flip Chip Bonding (FCB)**

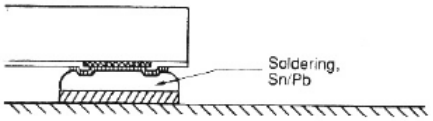


Figure 2.3: Cross-sectional representation of FCB [65]

2.1 Wire Bonding

The dominant technique for chip level interconnect is by far wire bonding. As the name itself suggests, this technique uses a thin metallic wire, either made of gold or aluminum, to connect the chip pads to the bonding pads on the substrate. The bonding mechanism uses a tool with a capillary hole through which the thin metallic wire is threaded and then drawn from one pad to the other. The tool allows to hold the wire with its ending part coming slightly out of the tool tip, which is then moved onto the bonding pad and bonded. After the bond on the chip pad is created, the tool moves to the bonding pad on the substrate, drawing the wire; when the tool detects the right substrate pad, it lands on it and creates another bond, then it breaks the wire. After this, the tool starts over again, bonding together another couple of pads [64, 66].

The bonding methods use pressure, heat and/or ultrasonic energy; the bond shape, bonding tool and bonding parameters are different depending on the wire metal type. The first bond can be either ball bond, if using gold wire, or wedge bond, if using aluminum wire, while the second bond is always a wedge bond, since this shape is created by the fact that the wire is broken in correspondence of it [64]. The three main bonding methods are [64]:

- **thermocompression bonding**, which uses pressure and heat and requires temperatures around 300-400 °C to form a reliable gold-aluminum bond, thus being suitable for ceramic packages but too hot for plastic ones

- **thermosonic bonding**, which uses heat, force and ultrasonic energy and it allows to form a reliable bond at 150-200 °C for applications requiring lower temperature

- **ultrasonic bonding**, which uses pressure and ultrasonic energy and it is generally used for aluminum wires.

A comparison of the three bonding methods is shown in Table 2.1, taking into account the metal wire characteristics, the bonding mechanism used and the shape of the first bond.

Wirebonding method	Normal type of metal wire	Force applied to form the bond	The type of the first bond
Thermocompression	Gold wire *Wire diameter:* 0.0007"-0.002"	Heat and pressure	Ball bond
Thermosonic	Gold wire *Wire diameter:* 0.001" or less	Heat, pressure and ultrasonic energy	Ball bond
Ultrasonic	Aluminum wire *Wire diameter:* 0.0007"-0.002", and 0.005"-0.02" for high current applications (Gold wire can also be used with wire diameter of 0.002" or larger.)	Pressure and ultrasonic energy	Wedge bond

Table 2.1: Comparison of different wire bonding methods [66]

For aluminum wires a wedge bond is formed. The aluminum wire bonding tool and wedge bond shape are shown in Figure 2.4 and Figure 2.5, respectively.

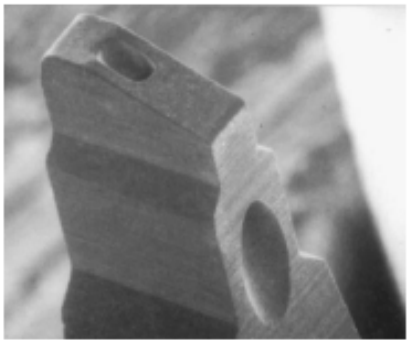


Figure 2.4: Aluminum wire bonding tool [64]

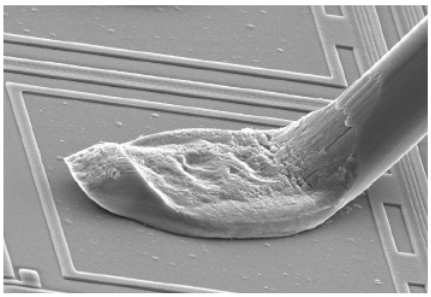


Figure 2.5: Aluminum wedge bond [64]

The process steps for aluminum wire bonding are shown in Figure 2.6.

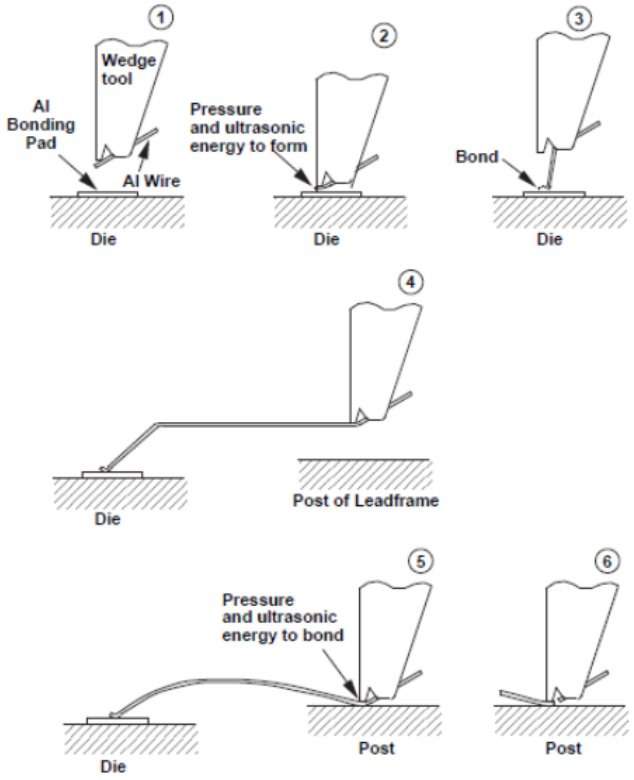


Figure 2.6: Ultrasonic aluminum bonding mechanism [64]

For gold wires a ball bond is created by electrical discharge at the tip of the tool. The gold wire bonding tool and ball bond shape are shown in Figure 2.7 and Figure 2.8, respectively.

Figure 2.7: Gold wire bonding tool [64]

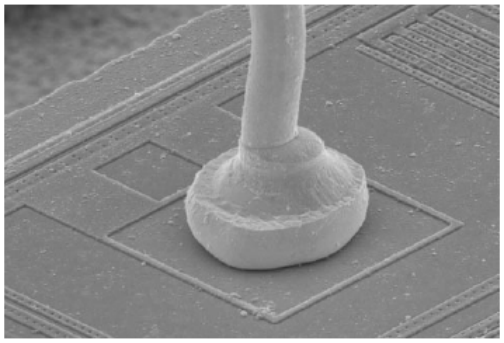


Figure 2.8: Gold ball bond [64]

The process steps for gold wire bonding are shown in Figure 2.9.

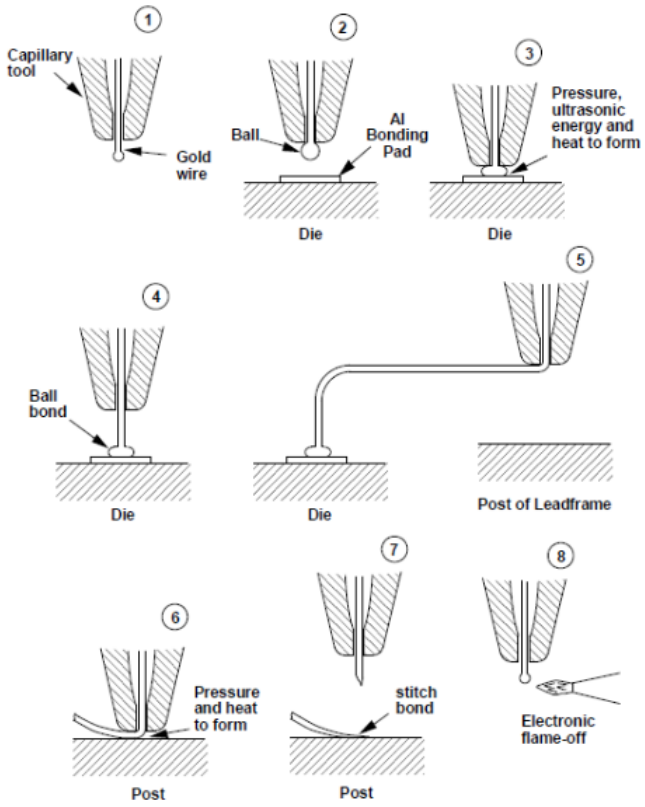


Figure 2.9: Thermocompression/thermosonic gold wirebonding mechanism [64]

In microelectronic industries wire bonding is generally automated for very high number of bonding pads and to obtain high throughput, with high speed bonding machines that can perform up to 10 bonds per second. Gold thermocompression allows to achieve higher bonding speed with respect to aluminum wedge bonding, due to the nature of the motion of tool and substrate; indeed, the first one is 3-5 times faster than the second one, constituting a relevant advantage that often makes it the lower cost alternative [64, 66]. The main advantages of wire bonding are [64, 66]:

- it is a very mature and well-known process
- virtually all dice are compatible with it
- the yield can be very high, typically 100 ppm defects per bond
- it is flexible, meaning that there is no need to change tool for different chips and it can accommodate die shrinks without requiring a package change
- the chip is attached to the substrate with the contact pads on the upper side, thus allowing to check the quality of the formed bonds afterwards.

The main problems and limitations of wire bonding are [64, 66]:

- it only allows for peripheral interconnects, which means a fine pad pitch is required to increase the I/O count, even if very high pad density cannot be achieved (exceeding 1000 I/O counts). An example of pads with fine pitch is shown in Figure 2.10.

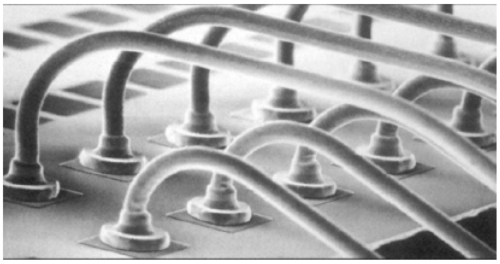


Figure 2.10: Gold wire bonding for fine pitch pads [64]

- the wires introduce high electrical parasitics, resulting in poor electrical performance, which means high propagation delay and noise. A comparison of the parasitics introduced by the three bonding techniques is shown in Table 2.2.

	Resistance Per Length	Inductance Per Length	Typical Lengths	Typical Resistances	Typical Inductances
Wirebond	1 Ohm/inch	25nH/inch	50-100mils	50-100mOhms	1.2-2.5nH
TAB	0.25 Ohm/inch	21nH/inch	100-300mils	25-75mOhms	2.1-6.3nH
Flip Chip	0.08 Ohm/inch	18nH/inch	3-6mils	<1mOhm	<0.1nH

Table 2.2: Electrical parasitics introduced by the three bonding techniques [64]

- it requires quite a large volume due to the presence of the wire loops, which means that the final packaged chip cannot be very small. The cross section of the wire loop connecting two bonded pads is shown in Figure 2.11.

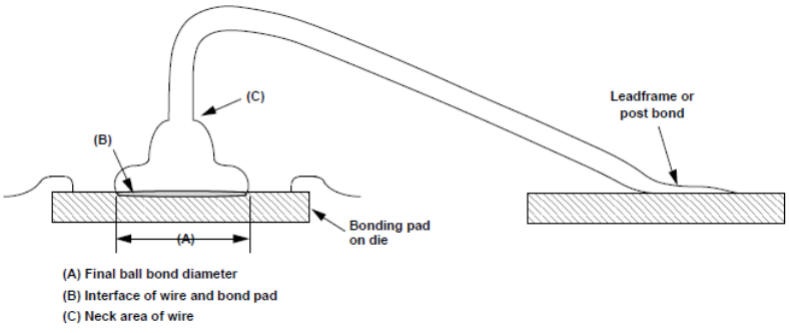


Figure 2.11: Cross section gold bond [64]

2.2 Tape Automated Bonding

Tape Automated Bonding, also known as Tape Carrier Packaging (TCP), is an alternative to wire bonding. As the name itself suggests, this technique involves a properly-designed polyimide tape which is used to hold the interconnection leads [64, 66]. Figure 2.12 shows the tape support fabrication, while Figure 2.13 shows the interconnections assembly from the tape reel onto the substrates.

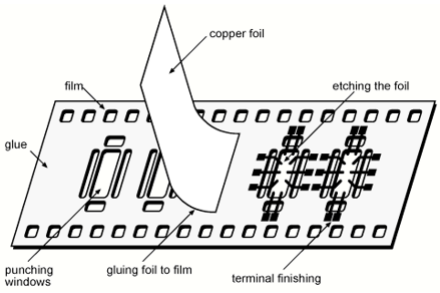


Figure 2.12: TAB tape fabrication [67]

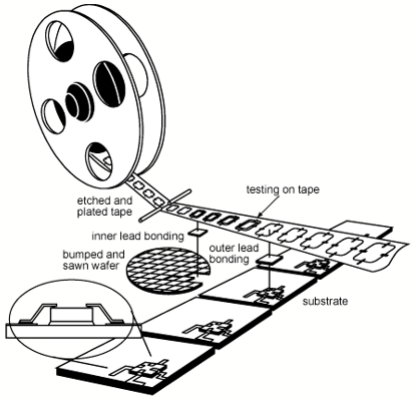


Figure 2.13: TAB reel mounting [67]

The bonding process follows three main steps [66]:

- **Bump formation**: the chip pads are first gold bumped, generally through electroplating, ball bonding or bump transfer.
- **Inner Lead Bonding (ILB)**: after bumping, the copper leads can be aligned with the chip pads on the inner part and bonded, either singularly or all together. The leads can be Au or Sn plated: in the case of gold, the leads are bonded to the chip gold bumps using thermocompression or thermosonic bonding [64]. The most common combinations of chip and tape metallizations are shown in Table 2.3.

Chip Metalization	Tape
Al pads	Gold-plated copper – single-point, thermosonic bonding
Al → gold-plated pads	Gold- or tin-bumped tape, gang bonding
Al → gold-bumped pads	Gold- or tin-plated tape, gang bonding
Al → solder-bumped	Gold-, tin-, or solder-plated tape, gang bonding

Table 2.3: TAB metal combination [64]

In case of thermocompression bonding, the control variables are force, temperature and time; usually, the bonding process uses a properly-sized square metallic bonding head, which pushes onto the inner lead bonding pads and create the bonds with the chip pads [68]. In this case, the inner leads are all bonded at the same time ("gang bonding"), as shown in Figure 2.14, but they can also be singularly bonded ("single-point bonding"), as shown in Figure 2.15.

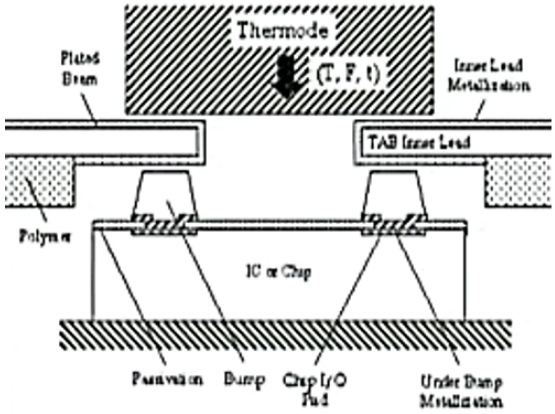


Figure 2.14: ILB thermocompression: gang bonding [68]

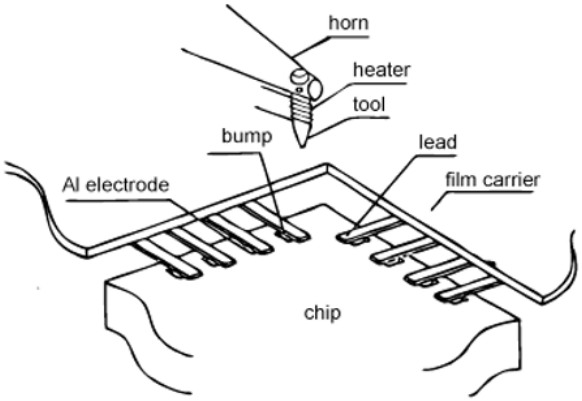


Figure 2.15: ILB thermocompression: single-point bonding [67]

For very small bonding pads even laser bonding can be used. After inner lead bonding is performed, the chip is covered in an epoxy layer to protect it from moisture and scratches [68].

- **Outer Lead Bonding (OLB)**: the chip is removed from the tape by cutting around its boundaries, so that the leads protrude from its edges; then, the chip is attached to the substrate, with the pads either on the upper or lower side, and the leads are bonded or soldered to the substrate metallization; soldering can be used both for Sn and Au plated leads. In case of gold plated leads with gold substrate

metallization, thermocompression or thermosonic bonding are used [64]. The OLB process is represented in Figure 2.16, showing the properly-shaped bonding tool.

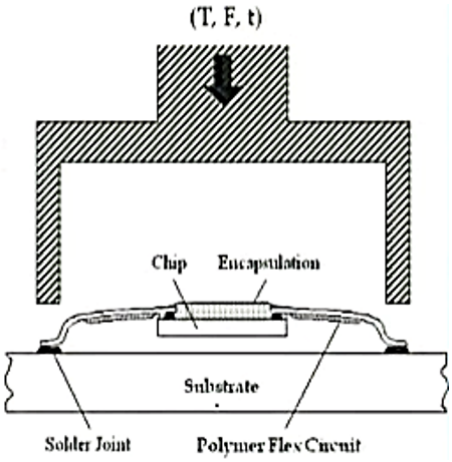


Figure 2.16: OLB thermocompression bonding [68]

Figure 2.17 shows a top and lateral view of a chip bonded with TAB method.

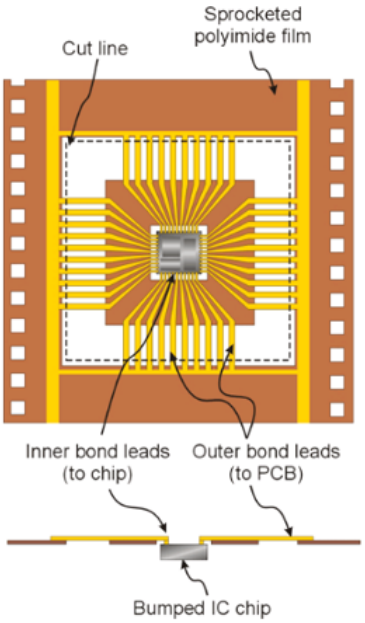


Figure 2.17: Top and lateral view of a chip bonded with TAB method [69]

TAB presents several advantages with respect to wire bonding [66, 68]:

- it allows to bond all the pads together ("gang bonding") after they have been bumped

- after ILB, both test and burn-in can be performed on the chip, thus allowing to proceed with the assembly process only for dice known to be the good ones
- it introduces lower electrical parasitics, as shown in Figure 2.2, thus allowing for better electrical performance
- it allows for smaller bonding pads and smaller on-chip bond pitch
- it allows to decrease the quantity of gold used
- it allows to obtain stronger and more uniform inner bonding
- it allows for higher I/O count density, even if it is limited to 850 pins, since only peripheral bonding pads can be used.

However, TAB also presents some limitations [66, 68]:

- as already mentioned, it only allows for peripheral interconnects, thus a very high I/O counts density cannot be achieved; area pad array could be achieved using multi-layer tape but the process is very expensive
- it requires time and cost to design and fabricate the tape carrier for each particular chip layout and also to bump the devices or the tape; for this reason, this bonding method is suitable only for very high production volumes, in contrast with the wire bonding process flexibility.

2.3 Flip chip bonding

Flip chip bonding was invented by IBM in 1963 and it was called "Controlled Collapse Chip Connection" (C4). In the past decades, the use of flip chip bonding has seen a tremendous growth due to the big advantages it presents with respect to other bonding techniques. This huge interest is mainly due to the trend in electronic industries, which aims to continuously reduce the chip dimensions while increasing its complexity, thus requiring very high pin density. This was not achievable with wire bonding or TAB, since they both allow only for peripheral interconnects; the major advantage of flip chip bonding is indeed the possibility to exploit the whole chip surface for the interconnects, approach called "area array bonding" [64, 68]. As the name itself suggests, flip chip bonding consists in flipping the chip onto the substrate, so that the pads on the two sides face each other and can be directly contacted and bonded together. The process steps of flip chip bonding technique are reported in Table 2.4.

create area array pads	• design the I/O for area array from the beginning • redistribute the traditional peripheral pads to area array pads with one or two metal layers deposited on the wafer
metallization for solder: under bump metallization (UBM)	• adhesion layers: Cr, Ti • solder wetting layer: Ni, Cu • solderable protective surface: Au, solder
build the solder ball	• evaporate, electroplate • solder composition: Pb/Sn 97/3, 95/5, 63/7, 60/40 • reflow
prepare pad on substrate	• screen print solder paste and flux • solder plate and apply flux
align and place flipped chip to substrate	• pick and place SMT equipment
reflow the solder and form the joint	• locally heated • reflow oven
underfill polymer	• snap cured, low viscosity epoxy

Table 2.4: FCB process steps [64]

First of all, an Under Bump Metallurgy (UBM) (described in details in the next subsection) is deposited onto the aluminum pads, then the bumps are formed on top of it. There are many different bumping techniques, but the most used is by far solder bumping with eutectic lead/tin (Pb/Sn) alloy [66]. The cross section of a bumped pad on a die is shown in Figure 2.18.

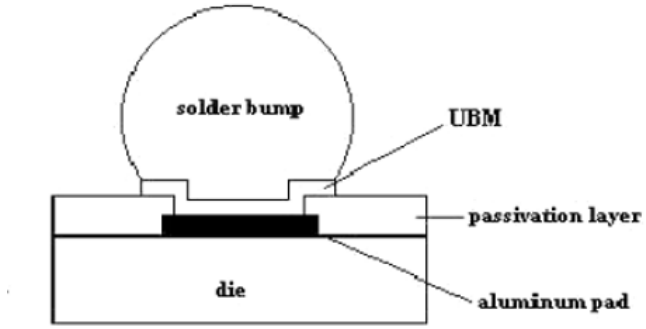


Figure 2.18: Cross-section of solder bump structure [66]

The bump formation is generally performed on the whole wafer as a batch process, since only few techniques allow to operate on a single die. The process of preparing the wafer with the necessary metallization and solder balls is indeed called "wafer bumping". After the bumps are formed, the wafer is diced and the single chips are placed on the substrate with the pads facing down (flipped) and aligned with the bonding pads on the substrate. Before the placement, some flux is applied either on the bumps or on the substrate, keeping the chip in place during the bonding and helping the reflow process. Then, solder bumps are reflowed, forming the bond between chip and substrate pads. Since chip and substrate are generally made of different materials, a Coefficient of Thermal Expansion (CTE) mismatch is present between them, which thus generates mechanical stress on the bonds during thermal cycling, leading to the risk of bond failure. To get rid of this issue, an underfill process is used: a low-viscosity liquid epoxy is dispensed along one or two sides of the die with a syringe, so that the epoxy is drawn under the whole chip area by the capillary forces activated by the very narrow gap between chip and substrate [66]. The underfill epoxy dispensation is shown in Figure 2.19.

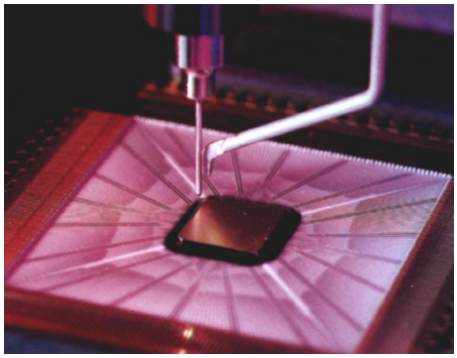


Figure 2.19: Underfill epoxy dispensation [64]

Epoxy underfill helps spreading more uniformly the mechanical stress generated by the CTE mismatch over the whole chip surface, rather than concentrating it at the outermost solder joints, thus preventing bond cracking during thermal cycling; moreover, it has an added function of providing environmental protection and enhancing the thermal

conductivity from the chip to the substrate [64, 66]. The enormous improvement in terms of bond failure percentage during thermal cycling due to the use of underfill epoxy is shown in Figure 2.20, reporting data for different gap heights between chip and substrate.

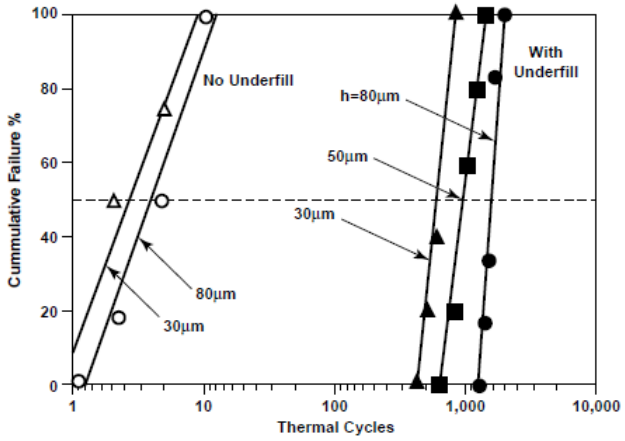


Figure 2.20: Comparison of bond failure percentage with and without underfill epoxy for different gap heights [64]

The sequence of process steps on the bumped chips is shown in Figure 2.21.

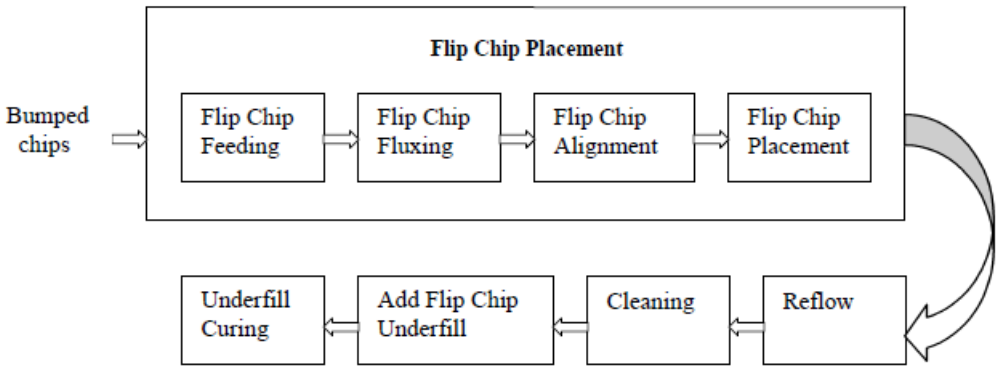


Figure 2.21: Detailed flip chip process steps after chip bumping [66]

The main advantages of flip chip bonding with respect to other bonding techniques are [64, 66, 68]:

- it allows to obtain very high pin density by exploiting the whole chip surface for interconnections (area array pads)
- such as TAB but differently from wire bonding, it allows to bond all the pads at the same time

- it allows for both Multi-Chip Module (MCM) and Single Chip Module (SCM) packaging for application requiring very high pin density
- it allows to enormously reduce the package footprint and volume, since no wires or leads are needed and the overall package footprint only requires a 20% additional space over the bare chip physical size
- it is the assembly technique with the lowest cost for applications requiring high volume production, high I/O density and large dice
- it offers the best electrical and thermal performance, since solder bumps introduce much lower electrical parasitics than wires or leads, as shown in Table 2.2. A solder bump has a parasitic inductance which is less than 10% of the one of a wirebond. The lower parasitics imply better electrical performance, which means shorter propagation delay of signals, thus higher IC operating frequency, and also lower noise. Moreover, the lower resistance also implies less heat generation due to Joule effect.

However, flip chip bonding also has some disadvantages, which are [66]:

- in many cases only wafer bumping can be performed, thus limiting the suitability of this process for single dice
- it is difficult to inspect the bonded pads since the chip covers the bonds
- it implies high costs for low volume production and low I/O count chips.

2.3.1 Under Bump Metallization

Purposes and requirements

UBM consists in a stack of thin metal films, typically made up of three layers with different purposes [66]:

- **adhesion and diffusion barrier layer**, which provides good adhesion to both the bonding pad metallization and the passivation layer, while preventing the diffusion between the bonding pad metallization and the chip solder bumps. Typical metals used are Cr, Ti, Ti/W, Ni, Pd, Mo deposited as thin layer (0.15-0.2 µm)
- **solder wettable layer**, which provides surface for solder bump to adhere to. Typical metals used are Cu, Ni, Pd with thickness of 1-5 µm, which is pretty thick so to generate enough wet surface during the bonding process to obtain a reliable bond
- **oxidation barrier layer**, which prevents UBM from oxidizing and generally consists in a very thin layer of gold with thickness of 50-100 nm, just to cover the underlying metal so to avoid its contact with air.

In addition to the previous list, other requirements for the UBM layer are [66]:

- good ohmic contact to bonding pad metallization, which means that for aluminum pads the thin native oxide layer on top of them has to be removed, either by sputtering or chemical etching, before UBM deposition
- minimum stress on the underlying silicon layer, since an excessive stress could fracture it.

Some of the most used combinations of thin film layers for UBM are Ti/Cu/Au, Ti/Cu, Ti/Cu/Ni, TiW/Cu/Au, Cr/Cu/Au, Ni/Au, Ti/Ni/Pd and Mo/Pd. The UBM structure greatly affects the reliability of the UBM itself and the connections on both its sides: on its bottom with the bonding pad and on its top with the solder bump. For this reason, UBM must be compatible with the solder alloys used for bumping [66].

Deposition techniques

UBM can be deposited by **sputtering**, **evaporation** or **electroless plating** [66]:

- both sputtering and evaporation deposit thin metal layers on the whole wafer, which then have to be patterned using photolithography to leave the metallization only on top of the pads. Another option for evaporation is to use shadow masking, which means that the metal layers are deposited on the wafer through a metal mask that has holes in correspondence of the pads, so that the metal is deposited only onto them. The same mask is also used to evaporate the solder bump
- electroless plating is instead a wet chemical process that deposits metal only on top of the bonding pads by exploiting a redox reaction activated by the pad metallization. This plating technique is used for example to deposit Nickel on top of aluminum, process called "ENIG" (explained in details in the next subsection), which is interesting because it is very cheap, since it does not require any photolithography or expensive vacuum equipment; moreover, it can be used both on whole wafers and single chips.

2.3.2 Solder bumping: types and techniques

Solder bumping is performed in most of the cases using lead/tin (Pb/Sn) alloy, since the formed bump is completely reflowable, thus allowing for self-alignment and collapse capabilities during the reflow process: the self-alignment property reduces the accuracy required in the process of chip alignment and placement onto the substrate, while the collapse property alleviates non-planarity issues. High Pb solders such as 95%Pb/5%Sn or 97%Pb/3%Sn require high reflow temperatures (~330-350 °C), while high Sn solders such as eutectic solder (37%Pb/63%Sn) allow for reflow temperatures ~200 °C, which makes them suitable for applications requiring low operating temperature, such as flip chip on organic substrates. However, since lead is toxic, research is moving toward the use of lead-free solder alloys [66].

Some of the main solder bumping techniques are [66]:

- evaporation
- printing
- electroplating
- solder transfer.

Evaporation

This process generally uses a technique called "shadow masking": both UBM thin films and solder are evaporated through the openings of a metal mask, which is properly designed so that the openings match the substrate pad layout, as shown in Figure 2.22. In this way, solder bumps assume the shape of a cone with height ~100-125 mm depending on the volume of evaporated solder, which is a function of the opening size and the distance between mask and substrate. In the case of Pb/Sn bumps, Pb has higher vapor pressure than Sn, which means it deposits first, thus creating bumps with a non-uniform composition, having lead-rich solder at the bottom and tin-rich solder at the top. This issue is resolved using a reflow step, which allows to homogenize the bump composition and to form the ball-shaped bump, as shown in the right part of Figure 2.22. An alternative to shadow masking is the lift-off process, which basically means substituting the metal mask with a thick photoresist mask spun and patterned on the substrate through photolithography. This mask can be stripped away after UBM and solder evaporation have been performed, thus removing also the metal layer deposited on top of the photoresist.

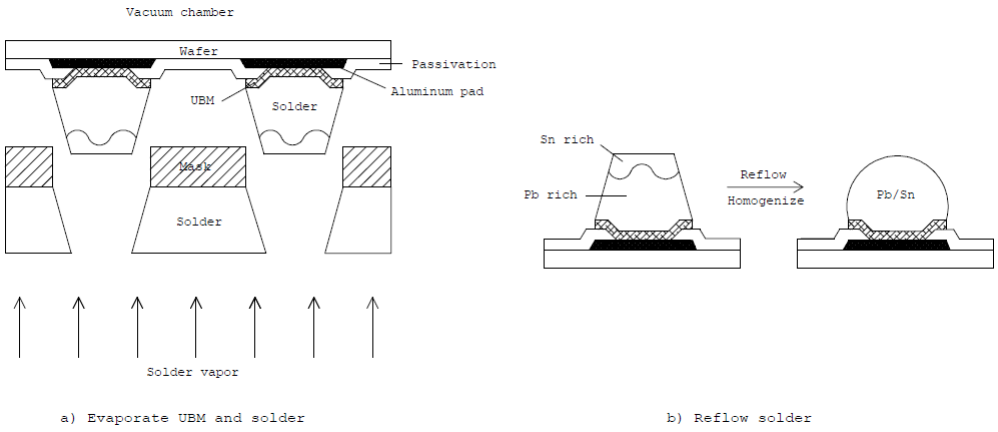


Figure 2.22: Solder evaporation process [66]

Printing

First of all, the UBM stack is deposited and patterned using either shadow masking or lift-off process. Then, solder paste is printed using either screen or stencil printing: some liquid solder is dispensed on top of a stencil having openings that match the substrate

pad layout, then the solder is forced through the openings with a squeegee, as shown in Figure 2.23. The volume and height of the bumps depend on the opening size and stencil thickness: since the first parameter is limited by the pad pitch, the stencil should be as thick as possible to increase the volume of printed paste; however, the higher is the thickness, the larger is the stencil wetted area, which means higher risk of clogging the stencil openings. A proper design has thus to be chosen taking into account this trade-off. As last step, the printed solder is reflowed to form the ball-shaped bump and flux residues are cleaned.

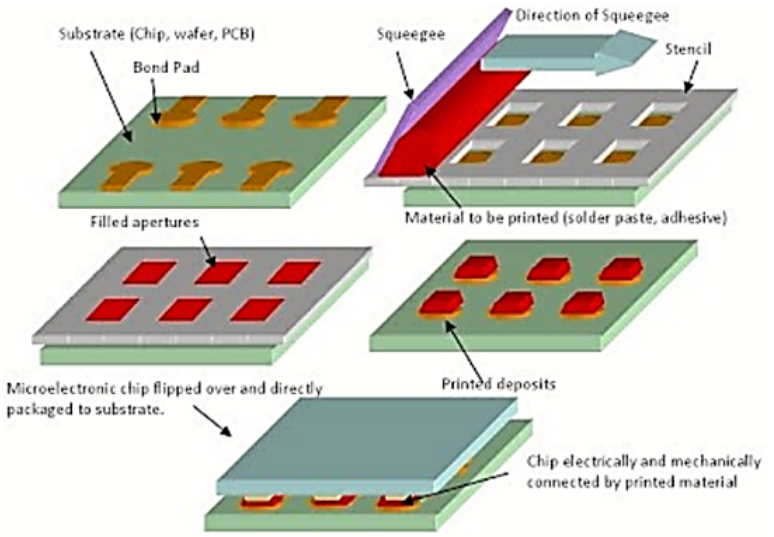


Figure 2.23: Stencil printing process [70]

Electroplating

First of all, UBM is deposited on the whole wafer, since it is used as seed layer for plating. Then, a thick photoresist layer is spun and patterned, determining the shape and height of the plated bumps. Before electroplating, the photoresist residues inside the openings are removed by plasma etching, since they can hinder the adhesion of the bumps, increase contact resistance and lead to an inhomogeneous growth of the metal layers. After cleaning, the wafer is electroplated by immersing it into a plating solution and applying a current, using the wafer as cathode. The obtained plated solder bump has a mushroom-like shape, as shown in Figure 2.24. After the bump formation, photoresist is stripped away and UBM is removed using wet etching by exploiting the plated solder as etching resist to avoid UBM removal underneath the bumps; then, the solder bump is reflowed into a ball shape, as shown in Figure 2.24. Another possibility is to reflow the bump first and then etch the UBM.

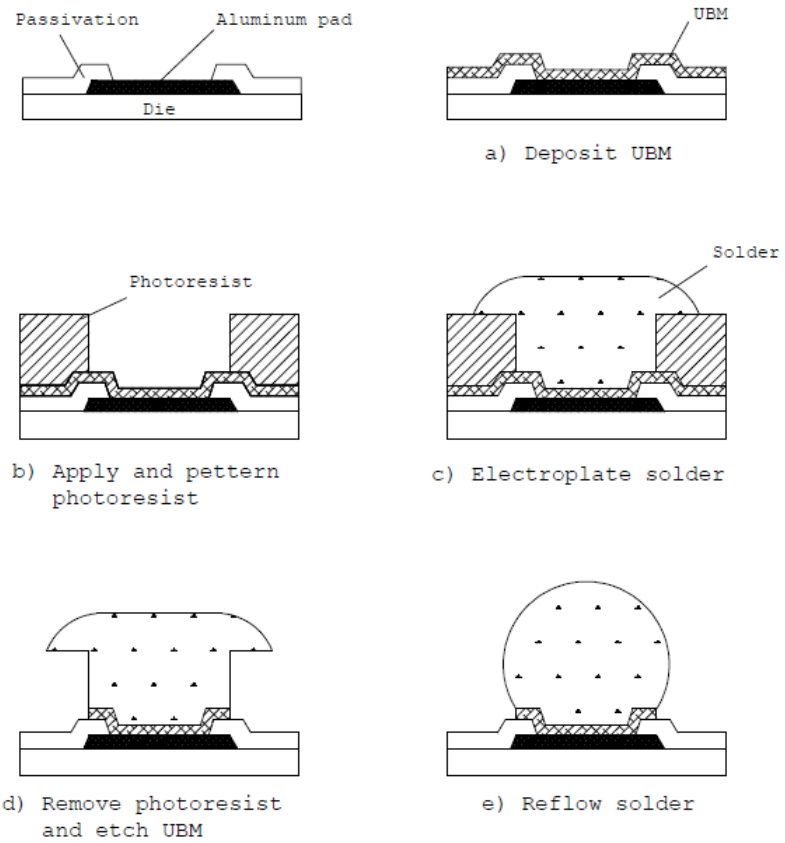


Figure 2.24: Solder electroplating process [66]

Solder transfer

This process requires to form the solder bumps on a carrier and then transfer them onto the bond pads. The carrier has to be made of a material which is non-wettable with solder, such as silicon wafer, so that the bumps can be removed and transferred without damaging them. First of all, the bumps are formed on the carrier with an evaporation process, either through shadow masking or lift-off process, reproducing the substrate pad layout. A thin gold layer (100 nm) is generally evaporated on the carrier before the bump formation in order to improve solder adhesion to it, thus preventing it from detaching during processing; moreover, it also allows to increase the time necessary for the bump to melt during bump transfer, so that it has enough time to wet and stick to the UBM on the chip pads. Then, if necessary, the carrier is diced and the bumps are transferred on fluxed substrate, either on a wafer or a single die. Finally, the bumps are reflowed so that they detach from the carrier and bond to the target pads; since the carrier still sticks to the bumps due to flux residues, they are cleaned and the carrier is finally separated. The main process steps are shown in Figure 2.25.

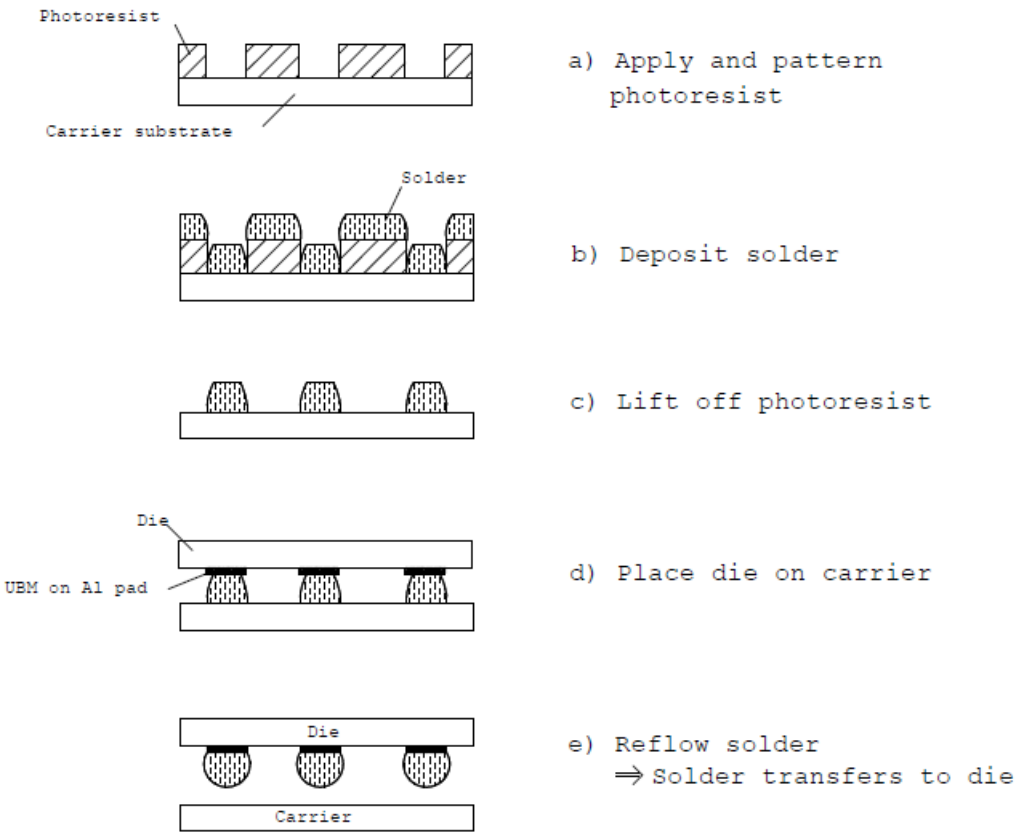


Figure 2.25: Solder transfer process [66]

2.3.3 Unconventional bumping techniques

The most expensive components of the bumping process are photolithography and the use of vacuum equipment such as evaporators. For these reasons, other bumping techniques have been developed to save money [64, 66]:

- gold stud or ball bumping
- ENIG plating.

These two bumping techniques do not require the use of solder, allowing to exploit an adhesive attach approach for the chip assembly, either using isotropically or anisotropically conductive adhesives:

- **Isotropically Conductive Adhesives** are generally epoxies filled with conductive metal particles such as silver, nickel and tin; the concentration of these particles strongly affects the bulk resistivity of the adhesive. Prior to apply the conductive adhesive it is necessary to bump the die pads with gold, either using gold ball bumping or ENIG process, since the aluminum surface does not provide a good

electrical contact with the epoxy [64]. The polymer can be applied either using stencil or screen printing, technique previously described and shown in Figure 2.23. The polymer can be [64]:

- a thermoplastic material, which softens when heated while it hardens after cooling down
- a fully-cured thermoset material, which is a liquid viscous material that can be irreversibly cured using heat
- a B-stage polymer, which can be partially cured in a first step using heat and then, eventually, completely cured in a second step.

In case of a fully-cured thermoset, the substrate pads must be coated with a conductive adhesive, which is the one that actually forms the joint; instead, in case of a thermoset or B-stage polymer, the substrate pads can be uncoated gold pads. Finally, the joint with the bumped die is formed through a thermocompression technique [64]. The result of this process, called Isotropically Conductive Adhesive (ICA) bonding or Isotropically Conductive Film (ICF) bonding, is shown in Figure 2.26.

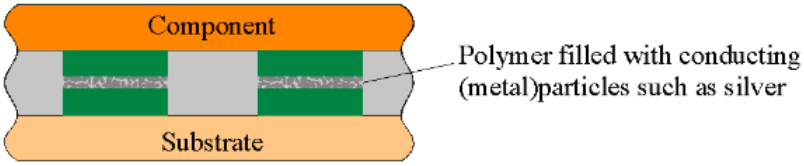


Figure 2.26: ICA bonding [71]

- **Anisotropically Conductive Adhesives** are generally made of an insulating epoxy filled with small and uniformly-sized latex spheres coated in Nickel and gold. The epoxy is applied as a thin film between chip and substrate, then a thermocompression process is used so that some of the spheres are trapped and compressed between the matching pads on the two sides, thus creating an electrical path between them. The density of the spheres inside the insulating epoxy is typically less than 15%, since it just has to be high enough so to have multiple spheres on every pad, but not so high to create a continuous chain of touching spheres between different pads, which would short-circuit them [64]. This bonding process, called Anisotropically Conductive Adhesive (ACA) bonding or Anisotropically Conductive Film (ACF) bonding, is shown in Figure 2.27.

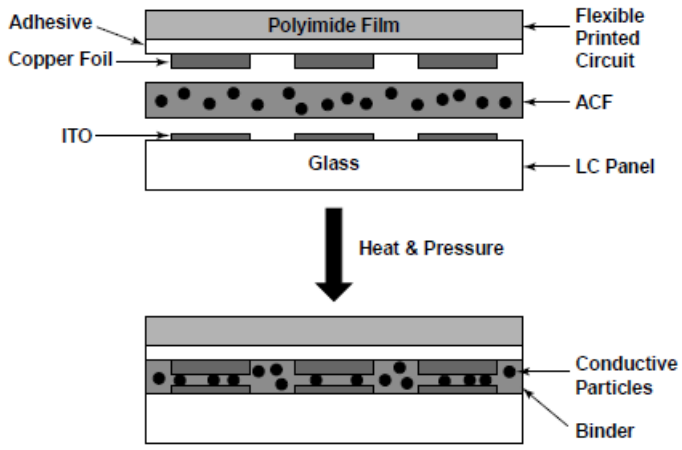


Figure 2.27: ACA bonding sequence [64]

Both of these bonding techniques lead to the formation of joints with relatively high contact resistance, thus implying a limitation in the capacity of carrying current [64].

Gold stud or ball bumping

This method uses the standard wire bonding process to form the bumps, but in this case the wire is broken after it has been bonded to the pad: if the wire is broken right in correspondence of the bond, the process is called "gold ball bumping", since it forms a gold ball on the pad; if a short piece of wire is left on top of the bond, the process is called "gold stud bumping", since it leaves a gold stud of different lengths [64, 66]. The gold ball bumping process is shown in Figure 2.28, while a couple of gold bumps onto die pads are shown in Figure 2.29.

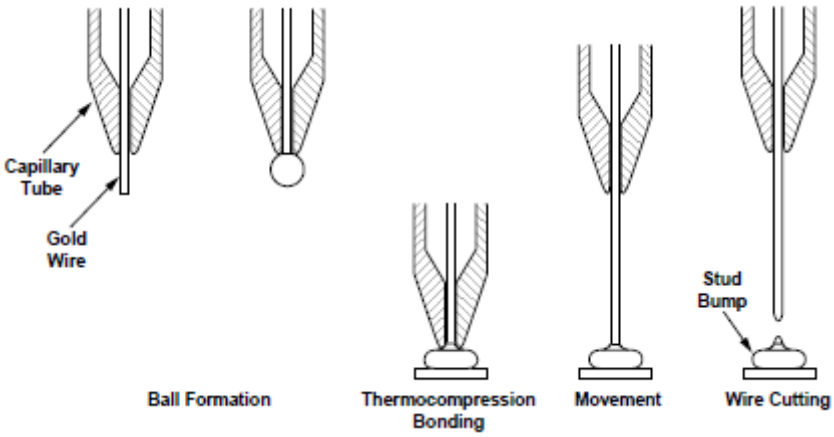


Figure 2.28: Gold ball bumping process [64]

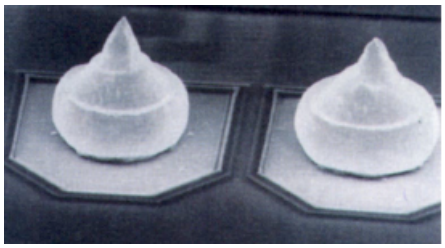


Figure 2.29: SEM picture of gold bumps on a die [64]

After bumping, the gold balls can be flattened, or coined, to result in an array of bumps all having the same height [64], as shown in Figure 2.30.

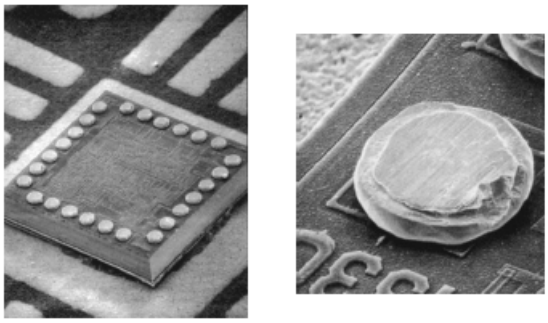


Figure 2.30: Example of coined gold stud bumps [64]

ENIG plating

The ENIG process is a wet-chemistry process used for both UBM deposition and bump formation on the chip aluminum pads. ENIG stands for "Electroless Nickel-Immersion Gold" and it is an electroless plating process: as the name itself suggests, this process does not require any electrical current or voltage, since it exploits a redox reaction between the metal ions in the plating solution and the pad metallization. This means that the deposition verifies on all and only the surfaces made of the metal which activates the redox reaction [72]. On the contrary, in the electroplating process the redox reaction between an anode and a cathode is activated by an electrical current flowing between them, induced by the application of an external voltage: however, it has to be considered that during electroplating the ions concentration in the solution is kept constant by the oxidizing anode, while during electroless plating the ions in the solution are only depleted by the deposition process, thus the solution has to be replenished from time to time in order to keep an adequate concentration which guarantees a deposition rate within a certain range. The main advantages of the ENIG process are [72]:

- it can be used on single dice
- it allows for bonding of fine-pitch bumps (~50 µm), since the bumped pads can be directly bonded, while many other bumping techniques require a solder reflow process that limits the pitch to ~200 µm.

- it is the bumping technique with the lowest cost, as shown in Table 2.5.

Bumping Technology	Cost/8" wafer	Cost/Die
Evaporated High Lead	$200	$1.00
Printed solder bump	$120	$0.60
Electroplated 63Sn/37Pb	$80	$0.40
Au Stud bumping	$70	$0.35
Electroless Ni / Au	$10	0.05

Table 2.5: Comparison of the main bumping techniques in terms of cost per wafer and per single die [72]

An example of ENIG process flowchart is shown in Figure 2.31, together with some pictures showing the pad appearance after the main process steps.

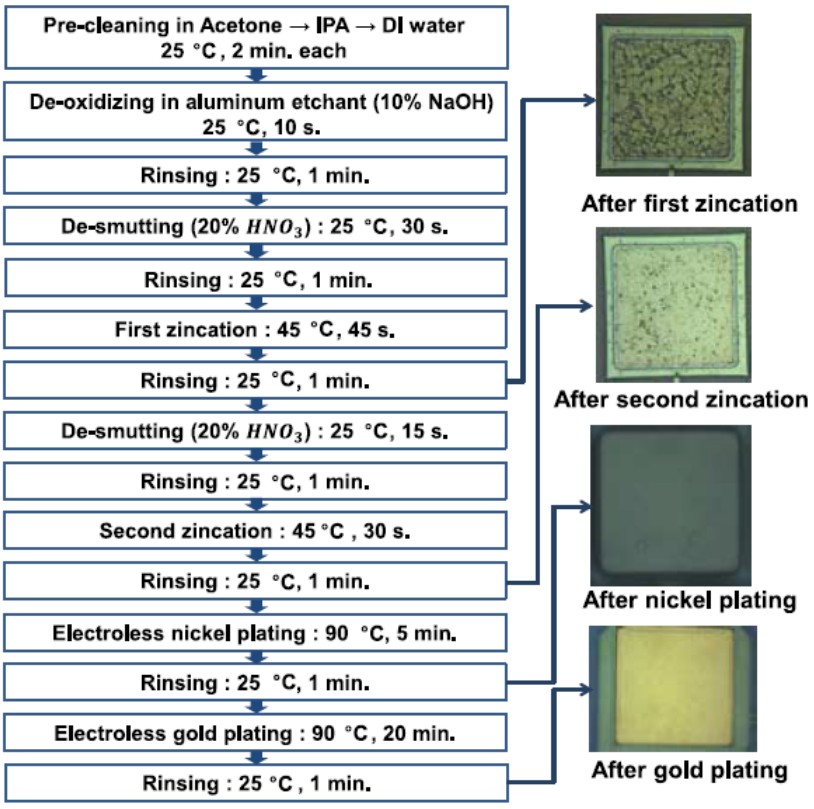


Figure 2.31: ENIG process steps [73]

The ENIG process is composed by several steps which deposit different metal layers, finally depositing gold onto the pads initially made of aluminum: zinc is deposited at first, followed by nickel and gold [72]. Each metal layer is necessary to activate the following redox reaction. All the redox reactions work by removing (displacing) metal atoms from the pad while substituting them with the metal ions present in the plating solution: for this reason, these reactions are called "substitutional" or "displacement" reactions.

Before the first zincation, critical steps are the initial surface cleaning and de-oxidation of the aluminum, in order to remove the thin layer of native oxide. The organic cleaning can be performed using different solvents, like acetone and isopropanol (IPA), or other proprietary metal cleaning chemicals (usually diluted mild acids such as acetic acid). The de-oxidation step is performed using a diluted basic solution, like 10% sodium hydroxide (NaOH), or acid, like diluted phosphoric acid (HPO_3) [74].

The first metallic step is the zincation: the zinc layer acts both as an activation layer for the aluminum layer, preventing it from re-oxidizing after the native oxide removal, and enabling the following Nickel redox reaction. As already said, the zincation is a displacement reaction, in the sense that aluminum atoms are displaced and substituted by zinc atoms during the deposition process [74]; the zincation redox reaction is [75]:

$$\begin{array}{lll} \text{Reduction:} & 2e^- + Zn^{2+} \rightarrow Zn & E^0 = -0.76\ \text{V} \\ \text{Oxidation:} & Al \rightarrow Al^{3+} + 3e^- & E^0 = 2.31\ \text{V} \end{array}$$

A long zincation produces large grains of zinc, while two shorter zincation steps with an intermediate nitric acid stripping allow to obtain better results: indeed, this process allows to smooth down the zinc grains and remove the low-quality deposit resulting from the first zincation (for this reason the intermediate stripping is called "de-smutting"), so that the second zincation can cover all the pads more uniformly and with finer grains. The zincation layer is very important since it is the first metal layer, which means that its roughness and uniformity will directly affect the quality of all the following layers up to the final gold bump [74]. Figure 2.32 allows to compare the zinc deposit uniformity and grain size after one and two zincation steps.

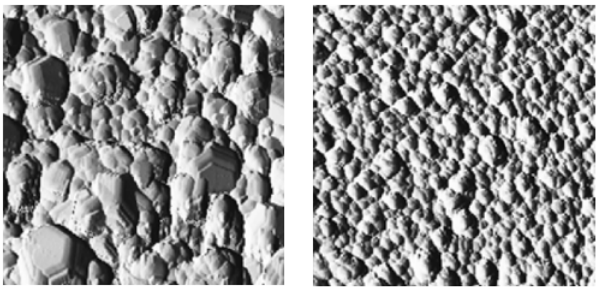


Figure 2.32: Zinc uniformity and grain size after one (left) and 2 (right) zincation steps [76]

The surface roughness at each process step is reported in Table 2.6.

	Initial	Pre.	1st Zinc.	2nd Zinc.
Surface Roughness /nm	22	25	95	24

Table 2.6: Pad surface roughness at each process step [76]

The aluminum layer is consumed both by the zincation substitutional process and the intermediate acid stripping, which means the immersion times have to be carefully calculated depending on the Al pad thickness, in order to avoid the complete removal of the aluminum layer. The zinc grain size and deposition rate can also be improved by adding ultrasonic agitation during the zincation process [77], as shown in Figure 2.33 and Figure 2.34.

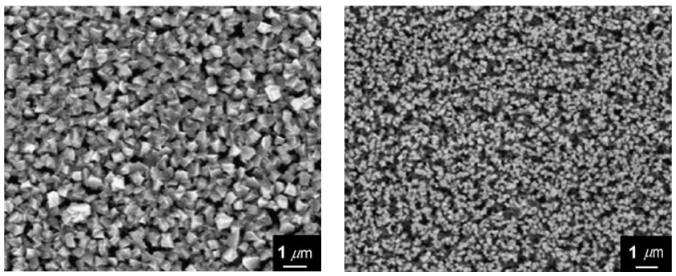


Figure 2.33: Zincation without agitation (left) and with agitation (right) [77]

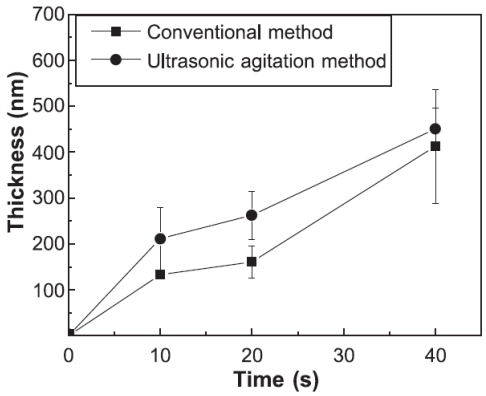


Figure 2.34: Zincation deposition rate improvement using agitation [77]

The Electroless Nickel plating solution is generally a Nickel-phospate based solution using Sodium-hypophosphate as reducing agent ($NaPO_2H_2$): the deposit is thus Ni/P alloy instead of pure Nickel. As already said, this reaction starts as a displacement reaction but then, after the zinc surface is completely covered in Nickel, the deposition

continues as an autocatalytic reaction: this means that the reaction product, in this case Nickel deposit, is also one of the reactants, thus allowing the reaction to sustain itself in a closed loop [74]. The EN redox reaction is [75]:

$$\begin{array}{lll} \text{Reduction:} & 2e^- + Ni^{2+} \rightarrow Ni & E^0 = -0.25\ \text{V} \\ \text{Oxidation:} & Zn \rightarrow Zn^{2+} + 2e^- & E^0 = 0.76\ \text{V} \\ & H_2PO_2^- + H_2O \rightarrow & \\ & H_2PO_3^- + 2H^+ + e^- & E^0 = 0.050\ \text{V} \end{array}$$

The Nickel layer acts both as diffusion barrier for the aluminum surface and solder wettable layer, also providing hardness and mechanical strength to the pad; it is the thickest metal layer in the UBM stack (1-15 µm) and the one that forms the bump shape. It is important to remember that the electroless plating is a maskless process, which thus deposits in all directions with a similar rate: for this reason, the thickness of the deposited metal stack is limited by the pad pitch, more specifically it has to be less than half of the gap between the pads in order to avoid short-circuits [74]. The process steps up to this point are shown in Figure 2.35.

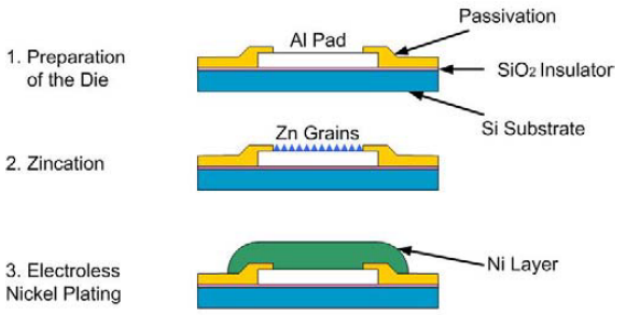


Figure 2.35: ENIG process steps up to the EN [76]

The last step is the Immersion Gold, which deposits a thin layer of gold on top of the Nickel layer to prevent its oxidation due to the contact with air, thus keeping the bump solderable. This is a self-limiting displacement process which deposits only a thin layer of gold: when all the exposed Nickel atoms on the pad surface have been substituted by gold atoms, the redox reaction stops and no more gold is deposited. The gold thickness obtainable with the IG plating solution is indeed limited to few hundreds of nm (0.1-0.3 µm) [74]. The IG redox reaction is [75]:

$$\begin{array}{lll} \text{Reduction:} & e^- + Au^+ \rightarrow Au & E^0 = 0.8\ \text{V} \\ \text{Oxidation:} & Ni \rightarrow Ni^{2+} + 2e^- & E^0 = 0.25\ \text{V} \end{array}$$

In case a thicker gold bump is needed to improve the pad solderability, it is possible to use a plating solution which deposits additional gold onto this thin gold layer: this solution is called "Autocatalytic Gold", which works according to the same principle already explained for the Nickel deposition and can deposit up to several µm of gold

[74]. The entire plating process is quite fast, except for the optional autocatalytic gold plating step, which could take a few hours due to its very low deposition rate (few µm/h).

The final bump structure showing the stack of metal layers is reported in Figure 2.36; the mushroom-like shape is due to the absence of any photoresist mask to give a metal growth directionality.

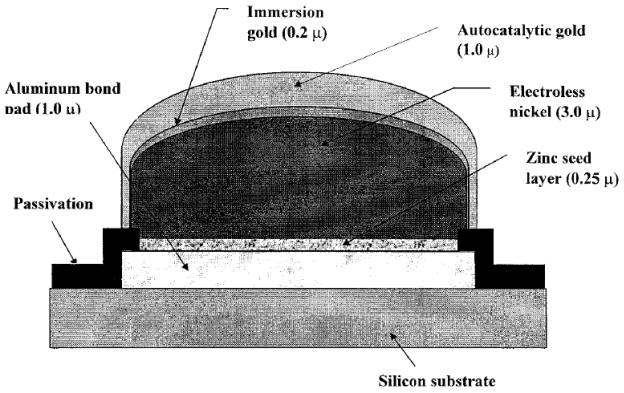


Figure 2.36: ENIG metal layers stack [74]

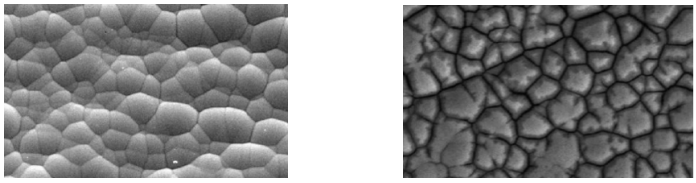
Black pad phenomenon As already said, the Immersion Gold step works through a displacement reaction which substitutes Nickel atoms on the pad surface with gold atoms, according to the following reaction [78]:

$$Ni + 2Au^{+} \rightarrow Ni^{++} + 2Au$$

For this reason, the IG reaction can be defined as a corrosion reaction. However, under certain conditions, the Nickel surface is subjected to a hyper-corrosion, which causes the so called "black pad" phenomenon. In particular, if the Nickel deposit is not uniform and homogeneous, some crevices can be present between the Nickel domains or nodules. Then, the IG solution inside the crevices cannot be replenished at the same rate as the one on the surface of the nodules, thus generating a galvanic cell between the two differentiated solutions; this cell initiates a corrosive side-reaction which generates hydride ions (H^{-}), releasing Nickel atoms without depositing gold [78]:

$$Ni + 2H^{+} \rightarrow Ni^{++} + H_2$$

If the immersion time in the IG solution is quite long and/or the IG solution is strongly acid, this galvanic reaction can extensively compromise the Nickel surface, making it rough, porous, black and hard to solder [78]. Figure 2.37(a) shows the plating result under normal deposition conditions, which is a very fine, uniform, planar, nodular deposit; Figure 2.37(b) shows instead the result of an accelerated deposition rate or reaction components/plating conditions outside of the recommended range, which is a less planar and less uniform surface with extended fissures between the nodules [78].



(a) Uniform and fine grain nodules

(b) Cracks and crevices between the grain nodules

Figure 2.37: SEM image of grain nodules [78]

In order to avoid the black pad issue, two main points must be kept in mind [78]:

- the nickel deposit should be tight and uniform with minimum crevices between the nickel nodules, which can be achieved only through an adequate preparation of the Zn/Al surface
- the gold thickness should not exceed 50-100 nm, as stated in the (IPC)-4552 ENIG specification issued by the Global Trade Association Connecting Electronics Industries in October 2002; indeed, it has been shown that 25-50 nm of gold are already enough to prevent the underlying Nickel surface from oxidizing, thus keeping the bump solderable. Moreover, the gold should be preferably deposited from a non-aggressive plating bath.

Figure 2.38 sums up the main causes of black pad phenomenon for copper pads, which are the same also for aluminum pads.

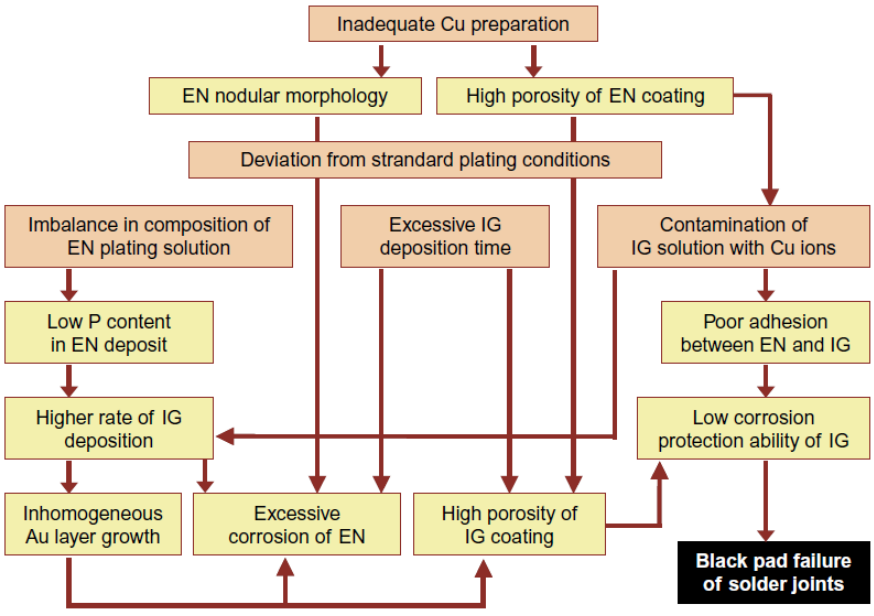


Figure 2.38: Main causes of black pad problem [79]

Chapter 3

Fabrication process

My work has focused on the fabrication of a new sub-mm version of implantable neural electrode for the project called "Neural Dust", already presented in Chapter 1. More specifically, my work has been devoted to the first fabrication step of the sub-mm ND motes, which consists in the ASIC chip attachment: in the previous mm-sized version it was performed by wire bonding, as already shown in Figure 1.26, but for this smaller version the flip chip bonding technique has been preferred in order to minimize the mote dimensions.

The ASIC chips were fabricated by an external fabrication plant (TSMC) and they came in the lab already diced into single chips; indeed, it would have been quite difficult to dice them with the available equipment, since their dimensions are ~400 µm×400 µm×100 µm. The chip layout, which is shown in Figure 3.1, contains five pads of ~70 µm×50 µm each with a minimum pitch of ~100 µm.

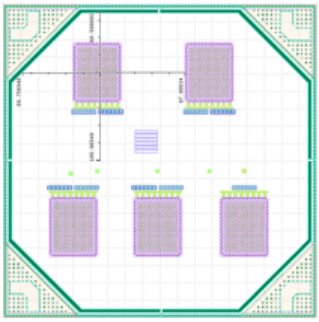


Figure 3.1: Neural Dust ASIC layout

It was decided to perform the chip pad bumping using the ENIG electroless plating process and then bond the chips to the substrates with a gold-gold thermocompression technique. The ENIG plating process (explained in details in Chapter 2) was chosen especially because it is one of the few bumping techniques suitable for single chips; moreover, an electroplating process was not feasible, since it is not possible to attach an electrode

on the µm-sized chip pads, together with the fact that these chips do not have an internal ESD protection, meaning that the use of an external current would have probably damaged the chip circuitry.

3.1 ENIG plating

3.1.1 Process development

The first hurdle was to devise how to hold the tiny chips during the plating process, which required to dip them into several beakers containing the plating solutions. The first idea was to bond the chips to a larger piece of material such as a glass or plastic slide and to use it as a holder. The bonding requirements are:

- resistance to heated acid and basic chemicals (90 °C max temperature for the EN)
- easily removable without residues on the chips at the end of the process.

Different types of glue were tested, such as Crystalbond™ and Loctite® Super Glue, since they are all easily and completely removable with acetone, but for this reason it was not possible to perform the organics cleaning of aluminum pads using acetone; then, Trisodium-phosphate (TSP) was tested as an alternative cleaning solution but it was found to corrode the glue too. The Crystalbond™ adhesive and Loctite® Super Glue are shown in Figure 3.2 and Figure 3.3, respectively.

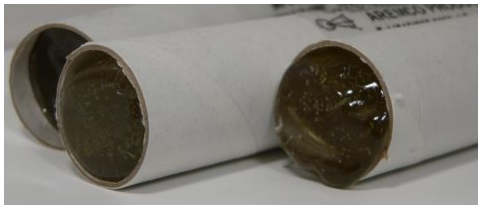


Figure 3.2: Crystalbond™

Figure 3.3: Loctite® Super Glue

The next idea was to build a sort of "basket" or "boat" made of a fine-pitch mesh so that the chips could not escape from it. The material of choice was polypropylene (PP), which is a thermoplastic polymer used for a wide variety of applications, since it is chemically inert to many acids and bases; moreover, PP mesh sheets are commercially available in a wide variety of mesh pitches. In this case, since the chip is 400 µm×400 µm, a 300 µm-pitch mesh was chosen. Then, in order to obtain a "boat" structure, a circular piece of PP mesh was cut in four points and bended so to create vertical walls, which were then heated at their extreme points to make them partially melt, thus allowing to stick them together and create a closed permanent structure after cooling down. Then, another piece of mesh was properly shaped and used as a sort of "lid", in order to prevent

the chips from escaping from the boat and flowing around in the beaker. Indeed, due to their tiny dimensions, the chips would float on the liquid surface when the boat is dipped into the beaker due to the liquid surface tension (it does not happen with solvents like acetone and IPA since they have low surface tension). The use of a "lid" is important since, if the chips exit the boat, it is very difficult to retrieve them from the beaker using tweezers due to their tiny dimensions and, especially, to remove them respecting the proper immersion time in the bath. In Figure 3.4 are shown both the PP boat and a PCB slide used to hold the chips.

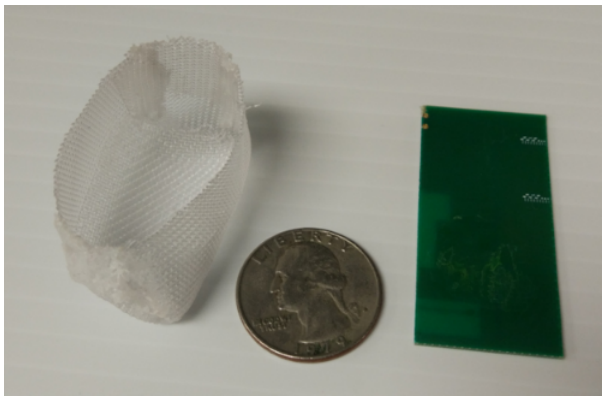


Figure 3.4: Polypropylene boat and PCB slide

This solution for holding the chips was found to be pretty useful and interesting also because it is scalable, meaning that a variable number of chips (large as needed) can be inserted in the same "boat", allowing for better uniformity of the plating results on a batch of chips by plating them all together; moreover, this process does not leave any residue on the chips, since no glues or epoxies are used, and it does not require additional bonding and cleaning steps to be performed on the chips.

The ENIG process flow comprising immersion times, temperatures and concentrations for all the wet steps was developed starting from Figure 2.31. Many tests were performed by changing immersion times and concentrations of the different solutions, trying to optimize uniformity and roughness of the plated metals; in the end, the best results were obtained with the following process:

- Organics cleaning: 2 min acetone + 2 min IPA
- DI rinse
- 1° zincation: 45 sec @RT
- DI rinse
- De-smutting in 20% nitric acid: 10 sec @RT
- DI rinse

- 2° zincation: 30 sec @RT
- DI rinse
- Electroless nickel: 15 min @90 °C
- DI rinse
- Immersion gold: 2 min @75 °C with stirring
- DI rinse

The complete plating setup is shown in Figure 3.5, while Figure 3.6 shows only the beakers containing the chemicals.
The ENIG process was performed inside the cleanroom under a sink vented hood, shown in Figure A.1 of Appendix A. In Figure 3.5 are also present:

- the thermocouple used to check the chemicals temperature (on the right)
- three hot plates to heat up the baths (in the back)
- a timer used to precisely control the immersion times (on the left)
- black plastic tweezers used to grab the chips, since metal tweezers could short-circuit the pads and also scratch the chip
- two squirt bottles containing acetone (red) and IPA (yellow) used for the organics cleaning (on the left)
- a graduated cylinder used to precisely measure the chemical quantities (on the left).

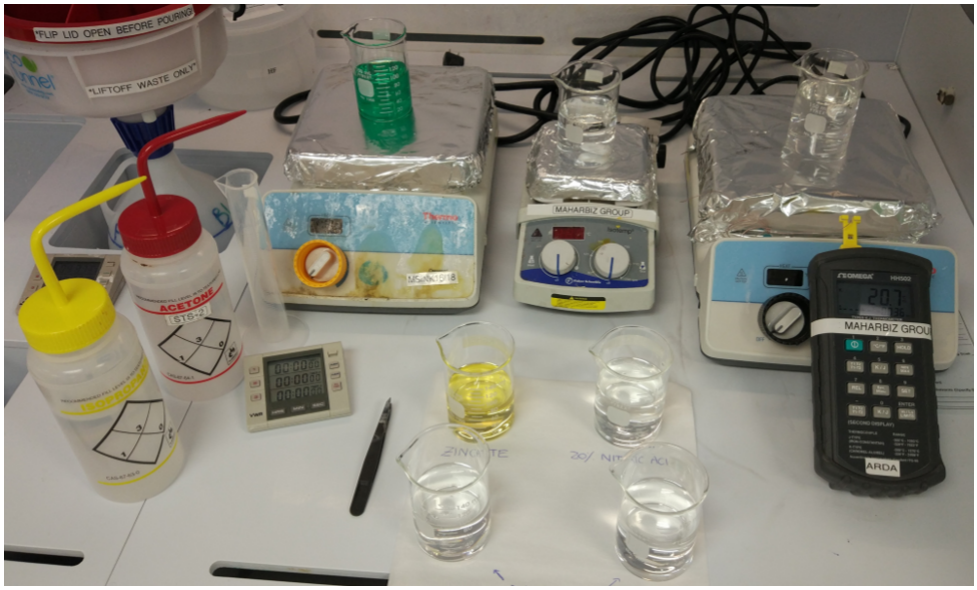


Figure 3.5: ENIG plating complete setup

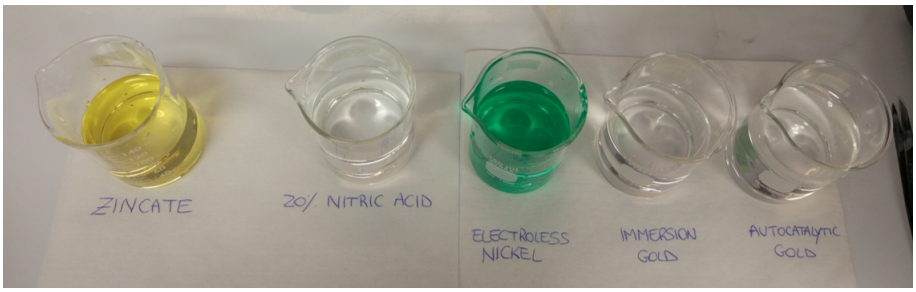


Figure 3.6: ENIG plating beakers setup

After each step there is a de-ionized (DI) water rinse, which is necessary both to abruptly stop the previous reaction and to prevent the oxidation of the deposited metal by keeping it wet, thus avoiding its contact with air. In this process each DI rinse step is a double rinse, using two different water beakers in sequence: this was done because the mesh of the "boat" traps the liquid in each beaker it is dipped into, due to the already mentioned surface tension, then the first DI rinse is just a rapid rinse used to clean the mesh by removing the liquid trapped in its holes, while the second rinse is longer and used to fully clean the chips (the water used for the first DI rinse has to be changed after each rinse since it gets dirty).

All the plating solutions used were commercially available chemicals, which were purchased from different companies: Caswell for the Zincate and the EN, Transene Inc. for IG and Metalor Technologies Inc. for the Autocatalytic Gold:

- the Zincate solution is strongly basic (pH=13-15) due to the high content (~65%) of Sodium-hydroxide (NaOH); the other main component is Zinc Oxide
- the EN solution is mildly acid (pH=3) and it contains mainly Nickel Sulfate and Sodium-hypophosphite (mid-phosphorous 5%-9%)
- the IG solution is basic (pH=9) and it contains Sodium Gold Sulfite
- the Autocatalytic Gold solution is neutral (pH=7) and it contains Gold trisodium disulphite.

Two final notes on the developed ENIG process are:

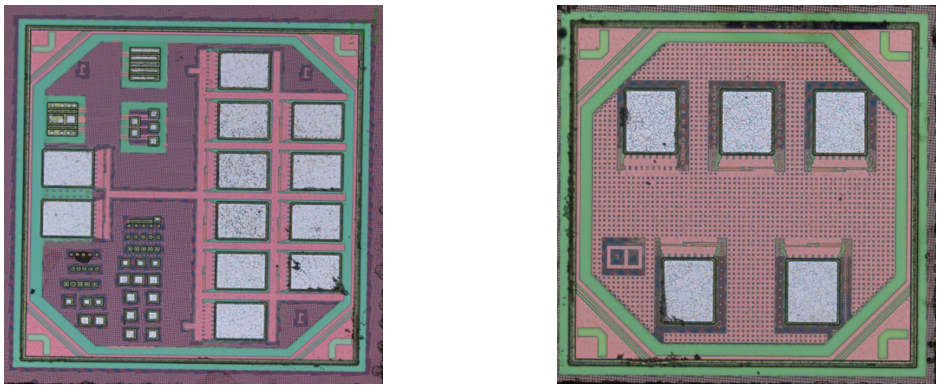
– the Autocatalytic Gold plating solution was inserted in Figure 3.6 even if it was not actually used for the process, since we found the Immersion Gold to be thick enough to allow the chip bonding

– thanks to the high content of NaOH, the zincate solution was meant to both de-oxidize the aluminum and deposit zinc; after several tests, it was indeed found that the ENIG process works also without an additional de-oxidation step. Moreover, the high content of NaOH corrodes the aluminum pad, thus the immersion times

for the double zincation have to be very short to avoid the complete removal of the aluminum layer. For this reason, an additional and unnecessary de-oxidation step (either acid or basic) would only increase the risk of completely removing the aluminum layer.

3.1.2 ENIG plating results

The ENIG process was performed on two types of chips with different pad number and layout, but with the same electrical behavior and function. This was done to verify the reproducibility of the developed plating process on different chips. The Neural Dust chips are shown in Figure 3.7.

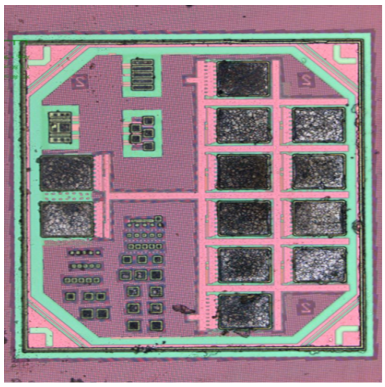


(a) ND chip (Type A) (b) ND chip (Type B)

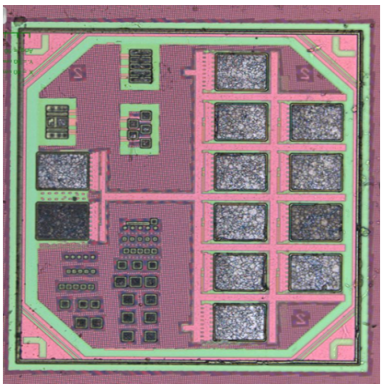
Figure 3.7: ND chips before plating

The chip shown in Figure 3.7(a) is the older version of ND chips (Type A), which are thicker and have a larger footprint with respect to the chip "Type B"; they were used for the assembly of the previous version of ND motes by wire bonding. The chip shown in Figure 3.7(b) (Type B) is instead the one used for the sub-mm version of ND motes and thus the one of interest for the developed flip chip bonding technique presented in this thesis.

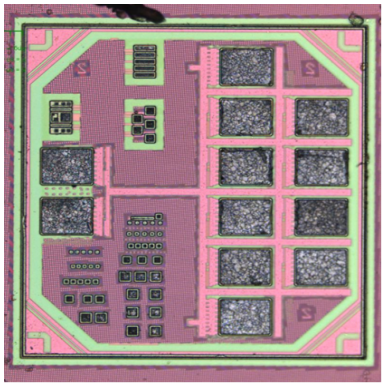
The results of the plating on the ND chip "Type A" are shown in Figure 3.8. The pictures were taken using the Olympus LEXT OLS4000, which is a 3D Laser Measuring Confocal microscope, shown in Figure A.19 of Appendix A.



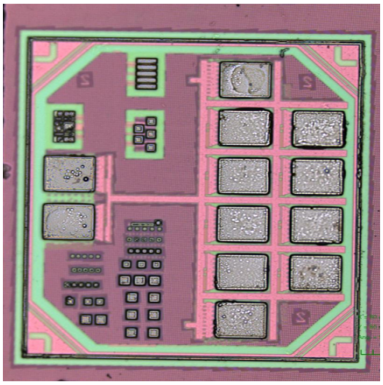
(a) First zincation



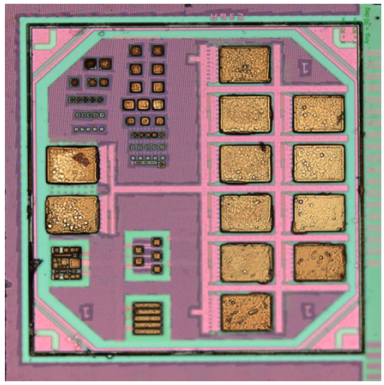
(b) Nitric acid stripping



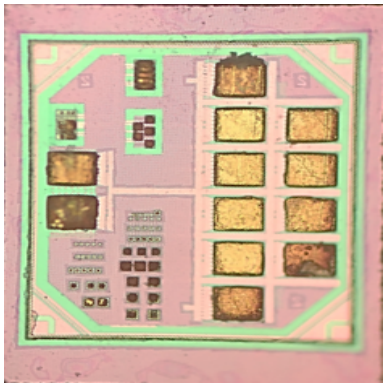
(c) Second zincation



(d) Electroless Nickel



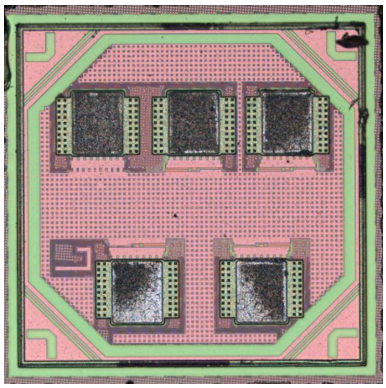
(e) Immersion Gold



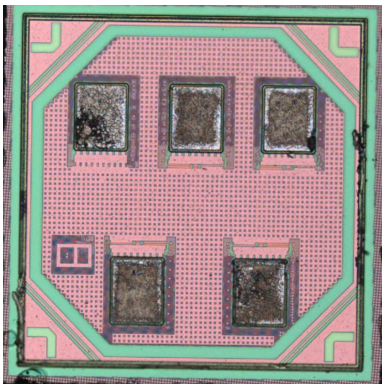
(f) Brownish Gold

Figure 3.8: Plating results for chip "Type A"

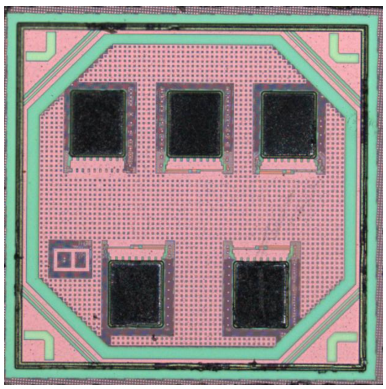
The results of the plating on the ND chip "Type B" are shown in Figure 3.9.



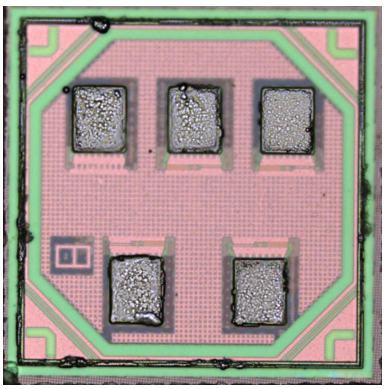
(a) First zincation



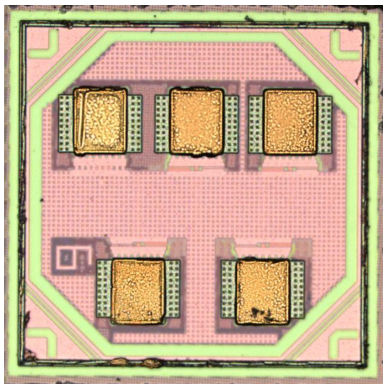
(b) Nitric acid stripping



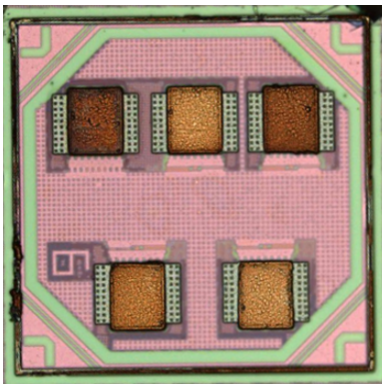
(c) Second zincation



(d) Electroless Nickel



(e) Immersion Gold



(f) Brownish Gold

Figure 3.9: Plating results for chip "Type B"

If the immersion gold plating is too long or the stirring is too fast the plated gold tends to become dark, assuming a brownish color, as shown in Figure 3.8(f) and Figure 3.9(f). This is probably due to the fact that, in these plating conditions, the deposition rate is higher, then the gold layer gets pretty thick and rough even with short immersion times: a rough surface scatters the light coming from above the sample in all directions, while,

on the contrary, a thin and smooth layer of gold acts more similarly to a mirror, reflecting most of the light coming from above in the same direction, thus appearing as a bright and shiny yellow surface. A zoomed view of some brownish gold-plated pads is reported in Figure 3.10.

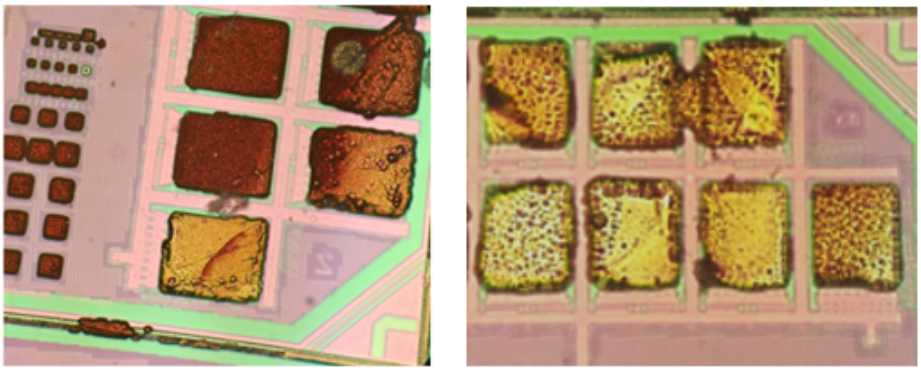
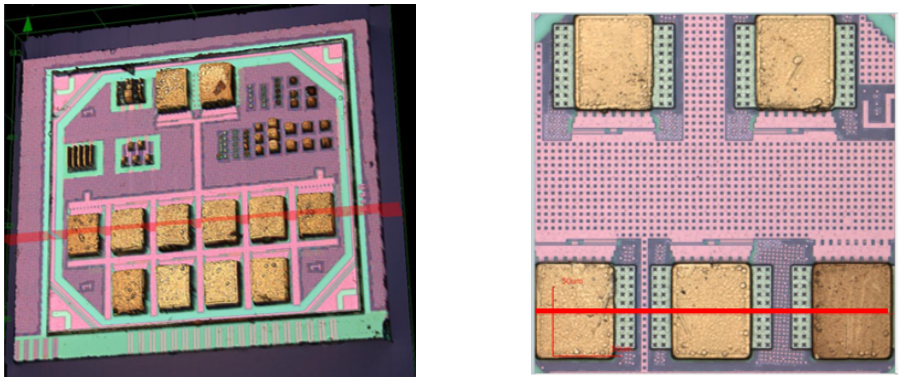


Figure 3.10: Zoomed view of some brownish gold-plated pads

As it is possible to observe in Figure 3.10, when the gold deposit is excessively thick, due to both a too long immersion time and/or too fast stirring, there is a risk of short-circuiting some of the pads, especially in the case of fine-pitch pads. This is due to the characteristics of the ENIG plating process, which, as already mentioned, produces an isotropic metal growth due to the absence of any photolithographic step.

Then, the pad profile at each process step was measured using the same imaging tool along the red lines shown in Figure 3.11.



(a) ND chip (Type A)

(b) ND chip (Type B)

Figure 3.11: Pad profile measurement

Figure 3.12 shows the profile of the six pads in a row for ND chip "Type A". In correspondence of the pad edges the profiles exhibit some peaks, which are probably imaging errors due to the height step.

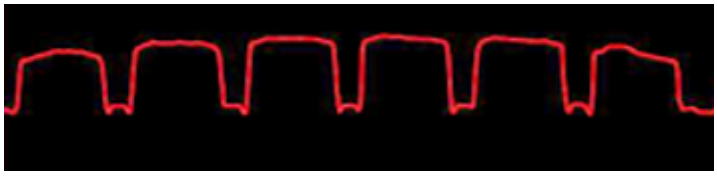
(a) Unplated chip

(b) First zincation

(c) Nitric acid stripping

(d) Second zincation

(e) Electroless Nickel



(f) Immersion Gold

Figure 3.12: Pad profile at each ENIG plating step (Type A)

Even if Figure 3.12 shows a qualitative representation, it is possible to notice that the pad thickness is almost unvaried until the Nickel plating step, which is the one that creates the real bumps; then, the final gold plating contributes to smooth down the bumps surface and make them flatter. Moving to quantitative measurements, the pads thickness and roughness at each process step for the ND chip "Type A" are reported in Table 3.1. The measurements were obtained using the same Olympus LEXT OLS4000 Confocal microscope.

Step	Thickness [µm]	Roughness S_q [µm]	Roughness S_a [µm]
Unplated chip	1.2	0.023	0.014
1° zincation	1.3	0.329	0.138
Nitric stripping	1.5	0.071	0.047
2° zincation	1.4	0.110	0.055
Electroless nickel	6.7	0.226	0.150
Immersion gold	6.9	0.102	0.062

Table 3.1: Thickness and roughness of the chip pads at each ENIG step

As already said, the thickness is almost unchanged during the double zincation process, while the bump is formed with the Nickel plating step, on top of which it is deposited a thin layer of gold. It has to be said that, in this case, the confocal microscope is not very precise for thickness measurements because the bottom layer surrounding the pads is silicon dioxide, which is transparent to light: this means that the laser light used by the microscope to perform the measurements crosses also the oxide layer instead of being reflected back, leading to an overestimation of the pad thickness. However, this issue affects all the measurements performed at each step in the same way, since the silicon oxide layer is unchanged, so these measurements still allow to extrapolated the thickness change at each step by looking at the difference between one value and the previous one.

Much more reliable are instead the roughness measurements: as expected from the theory, the collected data show that the first zincation layer is much rougher than the second one (almost half the roughness), then the Nickel layer increases the pad roughness and the final gold coating brings it back to around 100 nm. The values obtained for the roughness are higher than the ones reported in Table 2.6, but overall within one order of magnitude for the zincation steps, improving after the final gold plating.

Two different types of roughness measurements are reported in Table 3.1:

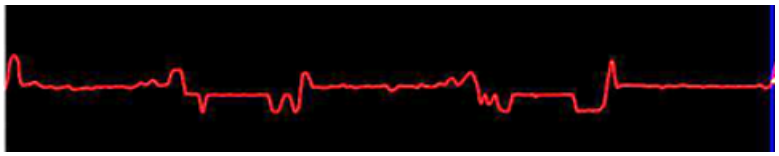
- the first one is the Root Mean Squared Height S_q

$$S_q = \sqrt{\frac{1}{A}\iint_A Z(x,y)\,dx\,dy}$$

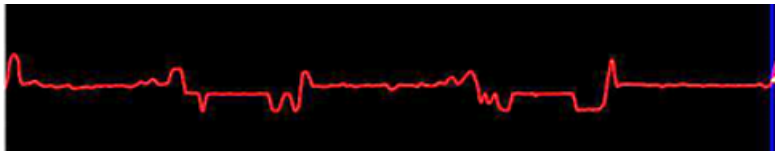
- the second one is the Arithmetic Mean Height S_a

$$S_a = \frac{1}{A}\iint_A |Z(x,y)|\,dx\,dy$$

Figure 3.13 shows the profile of the three pads in a row for ND chip "Type B".



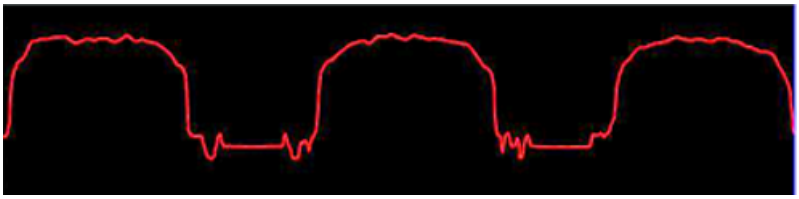
(a) Unplated chip



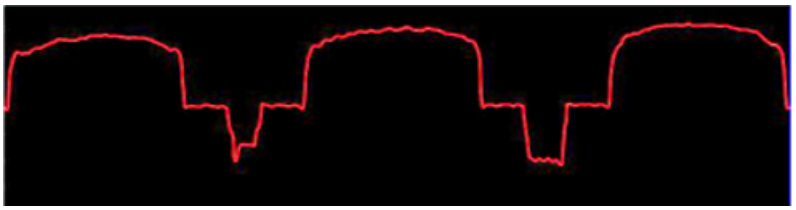
(b) First zincation

(c) Nitric acid stripping

(d) Second zincation



(e) Electroless Nickel



(f) Immersion Gold

Figure 3.13: Pad profile at each ENIG plating step (Type B)

The pads thickness and roughness at each step for the ND chip "Type B" are reported in Table 3.2. The measurements were obtained using the same Olympus LEXT OLS4000

Confocal microscope.

Step	Thickness [µm]	Roughness S_q [µm]	Roughness S_a [µm]
Unplated chip	1.3	0.025	0.018
1° zincation	1.37	0.308	0.200
Nitric stripping	1.26	0.100	0.070
2° zincation	1.25	0.150	0.114
Electroless nickel	6.5	0.284	0.212
Immersion gold	6.9	0.137	0.94

Table 3.2: Thickness and roughness of the chip pads at each ENIG step

Also in this case, the thickness and roughness change according to the same trends already discussed for chips "Type A".

3.2 Fabrication of the substrate

3.2.1 Process development

The substrates were realized using a lift-off process on a 6-inch silicon wafer. In a lift-off process the photoresist is spun and patterned first, so that it is removed only on the areas that has to be metalized; then, the metal is evaporated on the whole wafer and in the end the photoresist is stripped away together with the metal layer on top of it. In order to work properly, this process requires the photoresist to be at least three times thicker than the evaporated metal layer. This option is preferred to evaporate the metal on the whole wafer at first and then pattern it by using photolithography and metal etching techniques on its top. The substrates, also referred to as "target pads" since they are the bond target for the chip pads, have to replicate the same pad layout of the chips. Moreover, some outer test pads are needed to perform the electrical testing after the chip is bonded, since the bonding pads are no longer accessible. The photomask layout for the substrates is shown in Figure 3.14 together with a zoomed view of the central target pads, reporting the dimensions of the main features in µm.

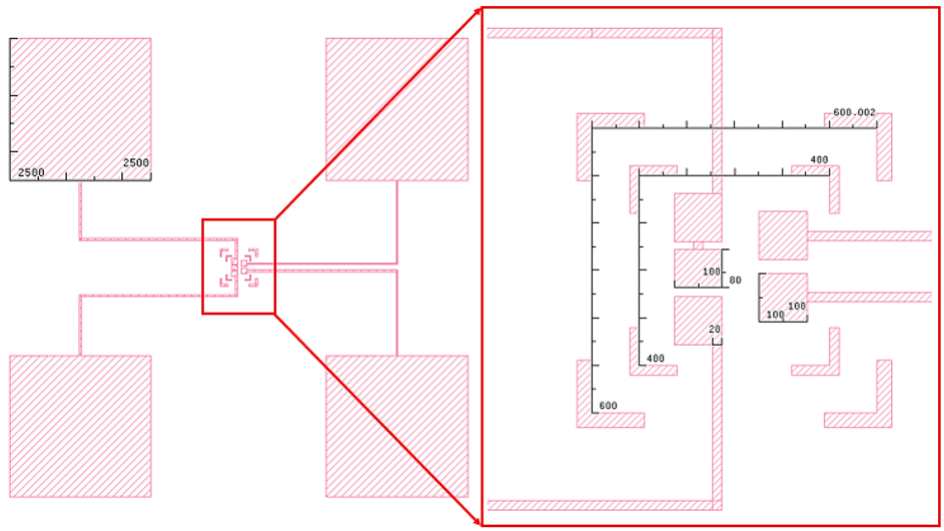


Figure 3.14: Substrate layout and zoomed view of the target pads

As it is possible to observe, the test pads are quite large (2.5 mm×2.5 mm) to allow for easier probe landing during the electrical testing at the probe station. Moreover, the target pads are square and a bit larger (100 µm×100 µm) than the chip pads (70 µm×50 µm) to allow for easier pads contact during flip chip bonding. Two of the target pads are shorted together since it is required by the chip internal circuitry. The two levels of angular features around the target pads have been inserted to facilitate the chip alignment to the substrate before bonding: the inner level has indeed the same dimensions of the chip footprint (400 µm×400 µm), while the outer level has the dimensions of some dummy chips that were fabricated to perform initial bonding tests (600 µm×600 µm).

The process flow for the substrate fabrication is sketched in Figure 3.15.

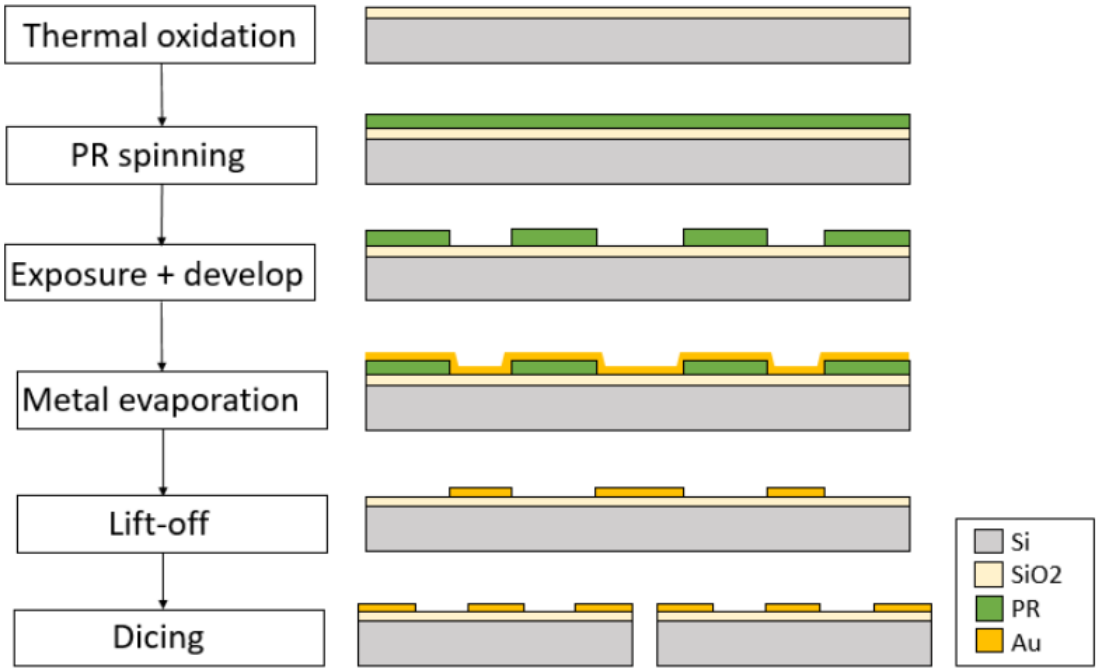


Figure 3.15: Lift-off process

For the fabrication process some 6-inch p-type doped silicon wafers were used. The thickness of the wafers is around 675 µm. The first fabrication step was a wet thermal oxidization performed in an oxidation furnace at 1050 °C to grow 1 µm of SiO_2: this layer is necessary to insulate all the gold-metalized areas resulting in the final layout. The oxidation furnace and some oxidized wafers are shown in Figure A.2 and Figure A.4 of Appendix A, respectively.

Then, the photolithography is performed using a double layer of photoresist: the first one is a 1 µm-thick MICROCHEM® LOR™ (Lift-Off Resist), the second one is a 1.2 µm-thick common i-line PR. The first PR layer has been spun using the tool shown in Figure A.7 of Appendix A, while for the second PR layer it has been used the tool shown in Figure A.8 of Appendix A.

During the exposure and development of the top PR layer, the features inside the bottom LOR layer get larger than the ones on the top layer, providing the undercut necessary for the lift-off process. Indeed, the metal layer evaporated on top of the PR has to be discontinuous in correspondence of the edges of the patterned features, otherwise when metal layer on top of the PR is stripped away together with the PR also the metal inside the features would risk to get damaged or removed. The MICROCHEM® LOR™ process steps are shown in Figure 3.16.

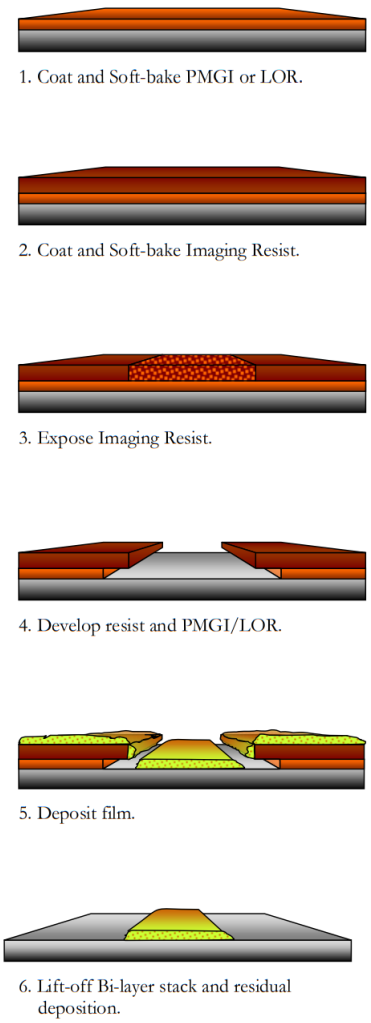


Figure 3.16: LOR process [80]

After the PR development, an O_2-plasma descum is performed for 1 minute at 50 W in order to remove any PR residues inside the patterned region, thus improving the adhesion of the evaporated gold. The tool used for this purpose is shown in Figure A.15 of Appendix A.

Then, 20 nm of chrome as adhesion layer and 150 nm of gold are sequentially evaporated using the e-beam evaporator system shown in Figure A.12 of Appendix A.

After this, the PR is stripped away using a long overnight soak in DOW® MICROPOSIT™ 1165 Remover, leaving the gold layer properly patterned. The whole wafer after gold evaporation and after the lift-off step is shown in Figure 3.17 and Figure 3.18, respectively.

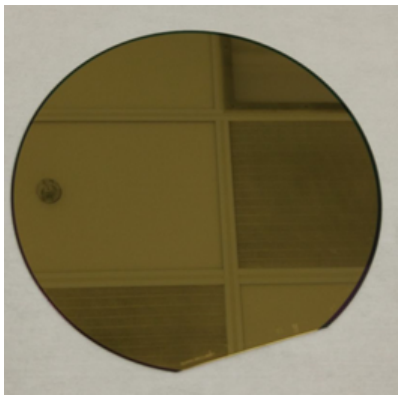


Figure 3.17: Gold coated wafer

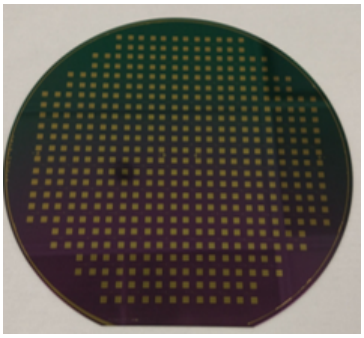


Figure 3.18: Wafer after lift-off

The last step is wafer dicing, performed with the semi-automatic dicing saw shown in Figure A.16 of Appendix A. The substrates obtained for both "Type A" and "Type B" chips after dicing are reported in Figure 3.20.

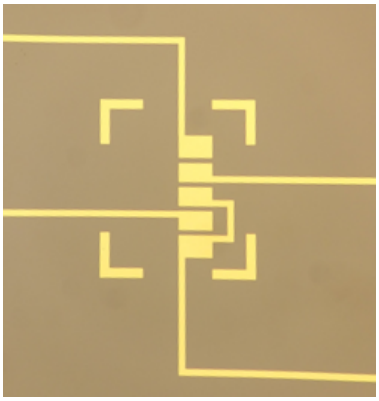


Figure 3.19: Fabricated substrate for chip "Type A"

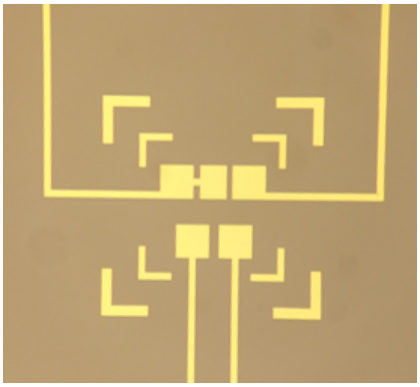


Figure 3.20: Fabricated substrate for chip "Type B"

The same lift-off process was also used to fabricate some dummy chips to perform initial bonding tests, with the only difference that aluminum was evaporated instead of gold to reproduce the real chip pads for chips "Type B"; the photomask used for the dummies fabrication contained different pad layouts with the purpose of finding out which pad configuration would better distribute the force applied during flip chip bonding without cracking the chip; this was done assuming a worst case in which the real chip showed not able to sustain the applied force. The central portion of the photomask used for the dummy chips fabrication is reported in Figure 3.21, since each quarter of the wafer had one of these four pad layouts; the fabricated dummy chip with the same pad layout of chips "Type B" is shown in Figure 3.22. The grid of orthogonal lines shown in Figure 3.21 was inserted to define the chip contours and to help the blade alignment during the final dicing step, since these lines have the same width of the silicon blade kerf (45 µm).

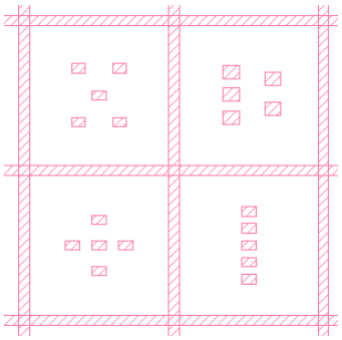


Figure 3.21: Photomask layout for dummy chips fabrication

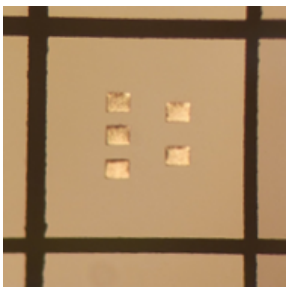


Figure 3.22: Fabricated dummy chip with pad layout if chip "Type B"

3.2.2 Seal-ring issue

A major issue was caused by the chip seal ring, which is an important structure running along the chip perimeter and used in CMOS integrated circuits to avoid introducing mechanical stress and contaminants into the chip circuits during chip dicing. Since the seal ring is also made of aluminum, it gets gold plated as well as the chip pads during the ENIG process. This means that, when the chip is bonded, this ring would short-circuit all the metal lines connecting the inner target pads to the outer test pads, thus preventing the proper chip functioning. The seal ring is the square metal frame comprised in between of the two red lines in Figure 3.23, while in Figure 3.24 it can be seen that the same seal ring has been gold plated (even if not uniformly).

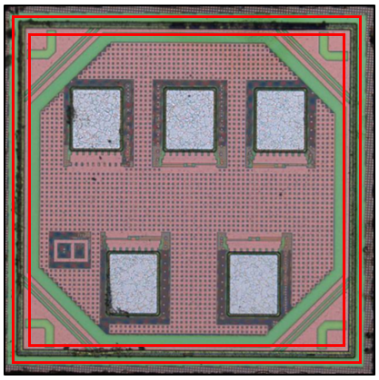


Figure 3.23: Seal-ring structure

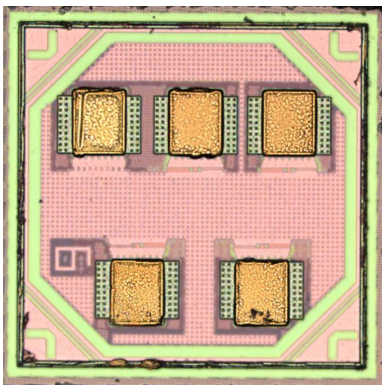


Figure 3.24: Gold plated seal-ring structure

The first idea to solve this problem was to use a seal-ring cover made of photoresist, which means spinning a thin PR layer on the whole substrate and then removing it only on the bonding pads, so that the seal ring would rest onto the PR layer after the bonding, thus preventing its contact with the metal lines and avoiding short-circuit issues. This idea of seal ring cover is shown in Figure 3.25, where the greenish regions are covered in a thin i-line PR layer.

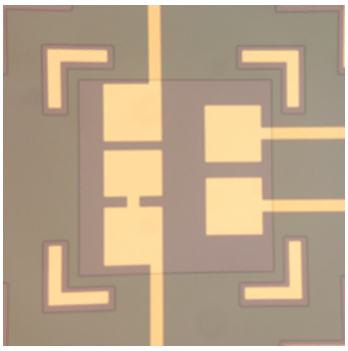


Figure 3.25: Substrate with a thin PR layer as seal-ring cover

This seal-ring cover required an O_2-plasma descum on the substrate before performing the bonding, in order to remove organic contaminants on the target pads, since an acetone cleaning would have removed the PR layer. However, after few bonding tests, this solution did not show good bonding results, probably because the PR layer hindered the contact between chip pads and target pads, thus causing the formation of weak joints or even preventing the joint formation altogether.

The second idea was to raise the target pads with respect to the substrate metallization, so that, during the flip chip bonding, the seal ring could not contact the metal interconnects, thus avoiding short-circuit issues. It was thought to realize this idea by spinning a thick PR layer (~10 µm) on the whole substrate and then removing it exclusively on top of the target pads; in this way, the tall PR features on the target pads could be filled-up by electroplating using the evaporated gold as seed layer, thus creating tall gold pillars only on the pads. The explained concept is sketched in Figure 3.26.

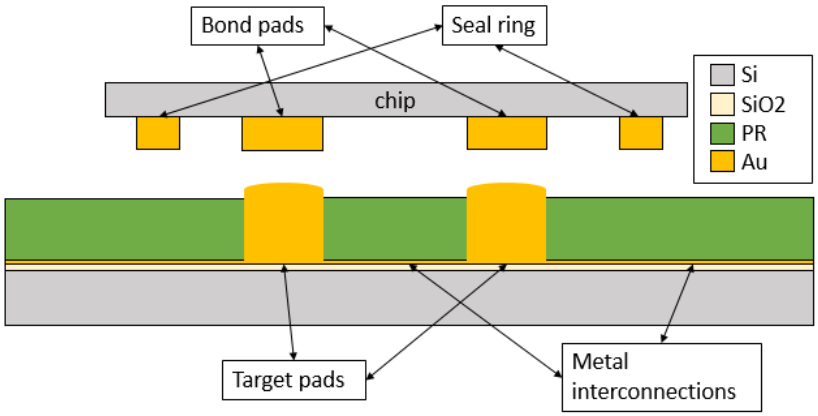


Figure 3.26: Thick photoresist concept

To implement this idea, the whole wafer was taken at the end of the previous lift-off process, it was primed with HMDS to improve PR adhesion using the priming oven shown in Figure A.5 of Appendix A, then the i-line photoresist DOW® MEGAPOSIT™

SPR^{TM}220 was manually spun using the tool reported in Figure A.9 of Appendix A. After this, the PR was soft-baked on a hot-plate and it was left on a table for 2 hours to allow for PR re-hydration before exposure. This step is important since the bulk water concentration in the PR layer drops during baking, but a certain water content is required to allow the photoreaction during exposure; since in this case a pretty thick PR layer is used and, moreover, the air inside the cleanroom has a quite low humidity, it takes some time for the PR to re-hydrate. At the end, the wafer is exposed through a proper photomask and the PR is developed in DOW® MEGAPOSIT^{TM} MF^{TM}-26A Developer. Before dicing, a final O_2-plasma descum was performed to remove PR residues on the pads. Figure 3.27 shows the substrate after the thick PR layer removal on the target pads, where all the regions surrounding the pads look darker due to the presence of the thick PR layer.

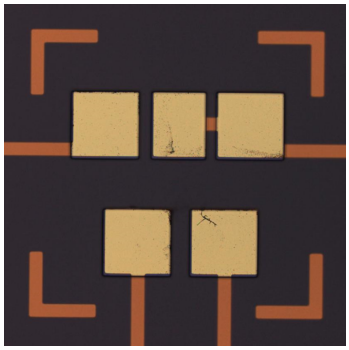


Figure 3.27: Top view of the substrate with thick photoresist

The gold structure left after gold plating and PR stripping in acetone is sketched in Figure 3.28. As it is possible to observe, the seal ring cannot touch the metal interconnections on the substrate, thus preventing short-circuit issues between the target pads.

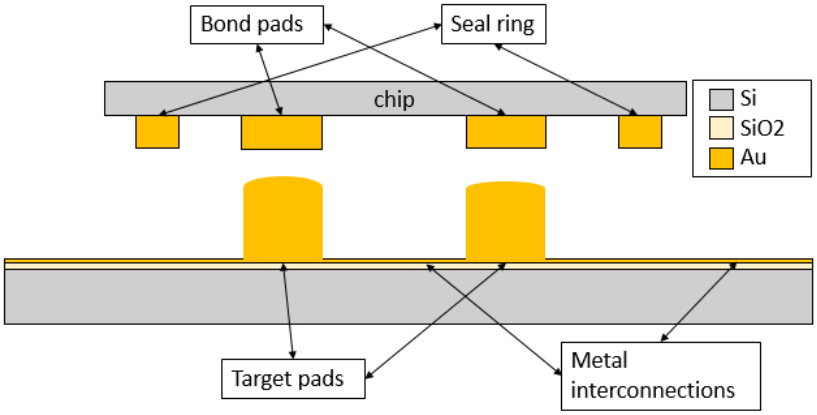


Figure 3.28: Thick photoresist concept after PR removal

The electroplating was performed by using a constant voltage of 1 V, since it was found that higher voltages caused issues due to the generation of bubbles from the target pads, probably due to too high induced currents which were leading to excessive plating rates; indeed, these bubbles were damaging and ultimately removing the PR layer from the substrate regions around the pads, thus allowing the metal growth to short-circuit the target pads. The electroplating setup is reported in Figure 3.29, showing a Tektronix® KeithleytextsuperscriptTM 2400 source-meter on the left and a ThermoFisher® Cimarec^{TM} stirring plate on the right. The electroplating solution was a commercial gold plating chemical from Caswell.

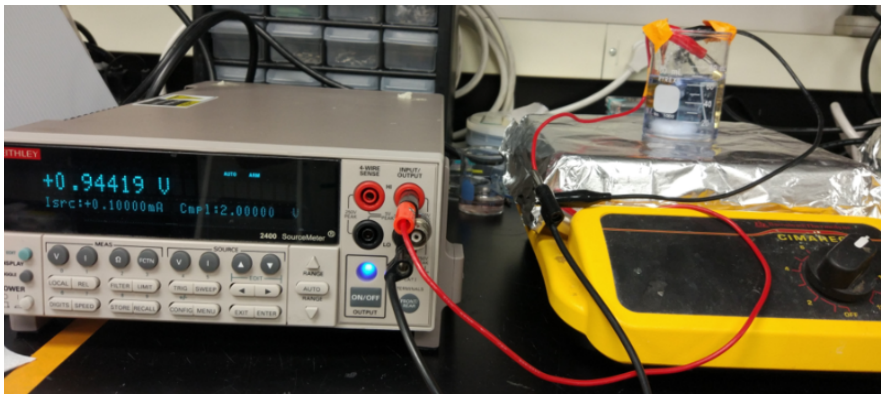


Figure 3.29: Setup for gold electroplating on the substrate pads

The electroplating on the substrate was performed by creating the electrical contact with the outer test pads because they are larger than the target pads: the electrical contact between the cathode electrode and the test pads was provided by some silver wires, removing their plastic sheath and then bonding one of their ending parts to the test pads with some Kapton® tape, so that it was easily removable after the electroplating. In the end, some nail-polish drops were placed on top of the tape and let to cure in order to prevent electrical leakage through the tape, since this would worsen the plating rate and the metal growth.
A Kapton® tape and silver wire spool are shown in Figure 3.30 and Figure 3.31, respectively.

Figure 3.30: Kapton® tape spool

Figure 3.31: Silver wire spool

The substrate preparation on a glass slide with the four protruding silver wires is

shown in Figure 3.32.

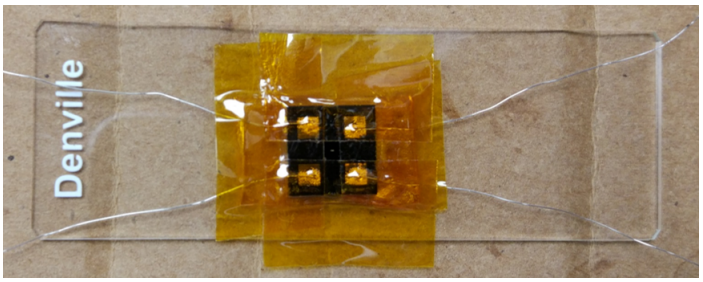
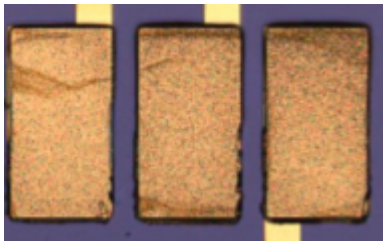


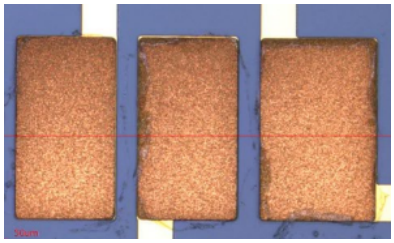
Figure 3.32: Substrate preparation for electroplating

3.2.3 Results for gold electroplating on the substrate

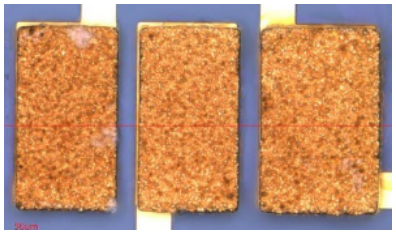
Figure 3.33 shows some zoomed views of the target pads for substrate "Type A" which underwent gold electroplating processes with different plating times.



(a) 0.5 hours electroplating



(b) 1.5 hours electroplating



(c) 3 hours electroplating

Figure 3.33: Substrates pads gold electroplated with different plating times

Figure 3.33 qualitatively shows that, as expected, the plated gold roughness increases with the plating time. Then, the quantitative thickness and roughness measurements were performed on substrate pads electroplated with different plating times, using the Olympus LEXT OLS4000 Confocal microscope. The collected data are reported in Table 3.3, from which it is possible to infer a general trend indicating that the electroplated gold roughness increases with the plating time; then, as already found for the Immersion Gold plating on the chips, the gold roughness increases with the thickness, which also means it looks darker.

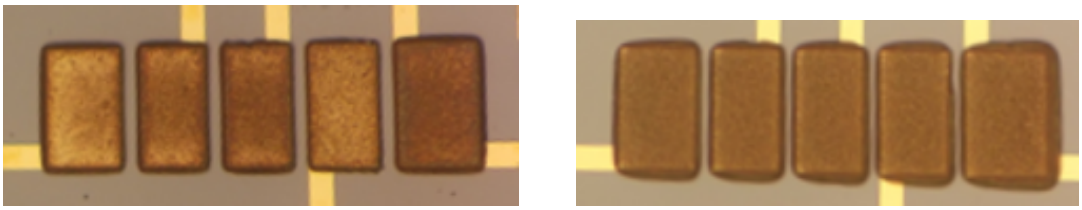
Time [hours]	Thickness [µm]	Roughness S_q [µm]	Roughness S_a [µm]
0.5	1.3	0.056 , 0.056	0.038 , 0.032
1	1.5	0.90 , 0.088	0.061 , 0.058
1.5	1.6	0.119 , 0.133	0.083 , 0.09
2	2.6	0.206 , 0.244	0.144 , 0.176
2.5	2.5	0.483 , 0.513	0.381 , 0.403
3	3.6	0.772 , 0.739	0.607 , 0.579

Table 3.3: Thickness and roughness of the electroplated gold onto the substrate pads

The issues and comments related to the thickness measurements of the gold plated chips are valid also in this case, since the pads have an underlying layer of silicon dioxide.

For the purpose of bonding, there is a trade-off between gold roughness and thickness. On the one hand, lower roughness means larger contact surface between the matching pads on the two sides, thus leading to the formation of stronger bonds; to obtain the same result with a rougher surface it is necessary to apply a much higher pressure and/or heat to smooth out the gold grains. On the other hand, a thicker gold layer allows for a better wettability of the substrate pads, which means a deeper interdiffusion between the gold layers on the two sides, thus leading to the formation of stronger bonds.
After several electroplating tests with different plating times, it was decided that a two-hours plating time was an appropriate trade-off between gold thickness and roughness for all the bonding tests.

In order to get rid of the mentioned problems caused by bubbling it was tried to add stirring to the electroplating process, in order to favor bubbles removal from the bottom of the tall PR features. However, it was found that stirring increases a lot the plating rate and, moreover, the metal growth is not vertical on the pads but it is slanted towards the stirring direction, since the metallic ions are kept in movement inside the plating solution in this direction, thus causing the pads to short-circuit together on top of the PR layer. Figure 3.34 shows a comparison of the metal growth with and without stirring.



(a) Electroplating without stirring (b) Electroplating with stirring

Figure 3.34: Gold-electroplated substrate pads

3.3 Gold-Gold thermocompression

Right before bonding the chips to the substrates, they were both cleaned with acetone and IPA to remove dust and organic contaminants; then, they were flip chip bonded using a gold-gold thermocompression technique by following these steps: the chip is flipped over the substrate, the pads on both sides are aligned, then pressure and heat are applied for a certain time. The temperature has to be high enough to soften the gold layer, so that the application of pressure can causes the interdiffusion of the gold atoms on the two sides and, after cooling down, the formation of a permanent bond. In the thermocompression process, the higher is the temperature used and the lower is the force needed to create the bond; however, the relation between these two parameters is not linear, since the metal interdiffusion process is exponentially dependent on the temperature, while it is linearly dependent on the applied force.

The thermocompression bonding has been performed using the professional flip chip bonder FINETECH® FINEPLACER® lambda, shown in Figure A.20 of Appendix A. The tool is composed of a heated plate, that can hold the substrate with vacuum, and a properly-designed tool used to pick-up the chip from the backside with vacuum; this pick-up tool has a tiny tip, either circular or square, that has to be properly sized depending on the chip dimensions. Moreover, the tool has a lamp with a beam splitter and a camera which allows to see both chip and substrate pads on the lateral monitor, either singularly or superimposed to perform the pad alignment, as shown in Figure 3.35. The software allows to select the temperature of the heated plate, the time for each heating step and the temperature ramp-up/ramp-down speed. The applied force can be set using a movable weight on the lateral arm, which is connected to the pick-up tool holding the chip, within the range 0.5-20 N; when everything is set up, the arm can be lowered down to push the chip onto the substrate and achieve the bonding. The view from a lateral camera during the bonding is shown in Figure 3.36, superimposed on the monitor with the temperature profile.

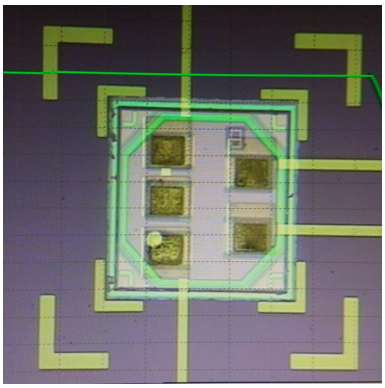


Figure 3.35: Chip alignment

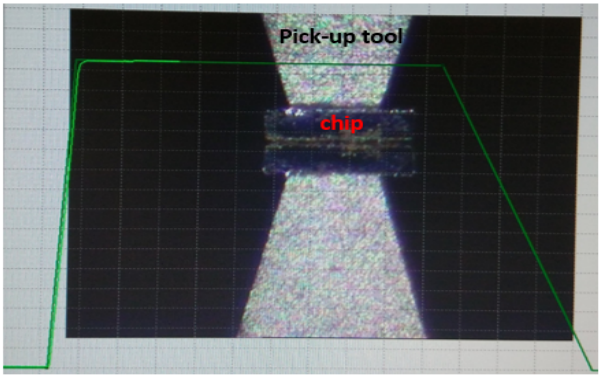


Figure 3.36: View from a lateral camera view during bonding

The heat for the thermocompression process was applied only from the bottom plate

holding the substrate, since the pick-up tool used for these chips did not have an heated element. The temperature profile for the bonding is shown in Figure 3.37: as it can be noticed, the cool-down ramp is much slower (2-3 °C/s) than the heat-up ramp (20 °C/s) in order to avoid chip cracking due to silicon thermal shock. This is particularly important in this case since the chip is very thin (100 µm).

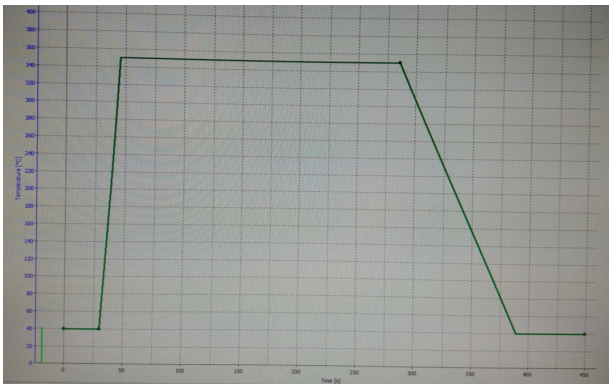


Figure 3.37: Temperature profile for bonding

The bonded chip is shown in Figure 3.38.

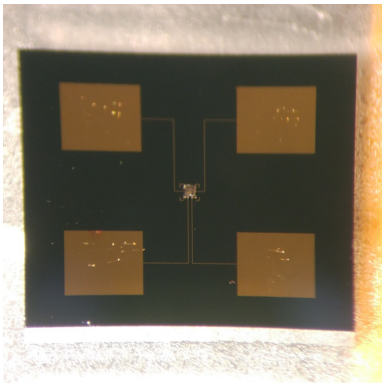


Figure 3.38: Bonded chip onto the substrate

Many bonding tests were performed using different temperatures and pressures, trying to optimize the bonding strength without affecting the chip performance due to the applied heat.

NOTE: as already mentioned in Chapter 2, flip chip bonding generally requires an underfill process to avoid bond cracking due the CTE mismatch on the two sides; however, in this case it was considered not necessary since on both sides there is gold on top of silicon, so the same materials and no CTE mismatch. Nonetheless, an underfill epoxy could be used to improve the bonding strength after verifying that there is complete electrical contact, but in this case it would be necessary to devise how to precisely dispense the underfill epoxy only below such a tiny chip, without involving the surrounding devices.

Chapter 4

Testing results

The testing was performed only on bonded chips "Type B", since they are the only chip of interest for the final assembly of the sub-mm ND motes.

4.1 Electrical testing results

The external pad layout for the chip circuitry schematic shown in Figure 4.1 is reported in Figure 4.2, where "B" is Bulk, "C" is Center (central point of the resistive branch), "G" is Gate, "S" is Source and "D" is Drain.

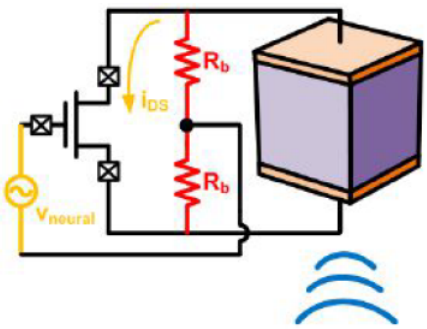


Figure 4.1: ND ASIC schematic [61]

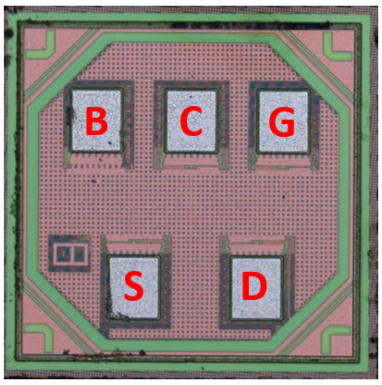


Figure 4.2: ND chip external pads layout

After the bonding, an electrical test was performed using a probe station, which is shown in Figure A.22 of Appendix A, in order to check if all the chip pads were electrically contacting the matching pads on the substrate, thus meaning the bonding was successful.

An example of test at the probe station is shown in Figure 4.3, with the probes touching the four outer test pads on a substrate.

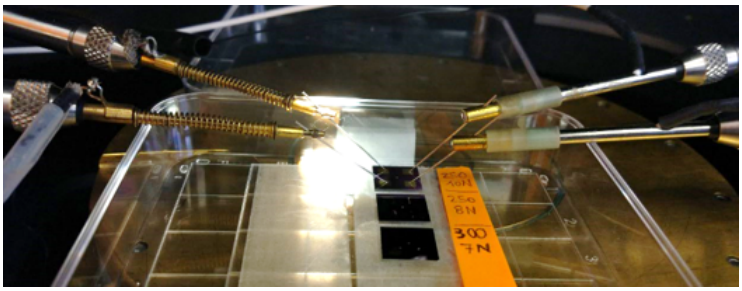


Figure 4.3: Electrical test at the probe station

For the electrical test the pads "G","C" and "S" were all connected to the same ground and V_{DS}=1 V, then V_{GS} was changed to simulate the neural signal while measuring the drain current I_D. The I_D-V_G characteristic measured on the gold plated chip before bonding is reported in Figure 4.4.

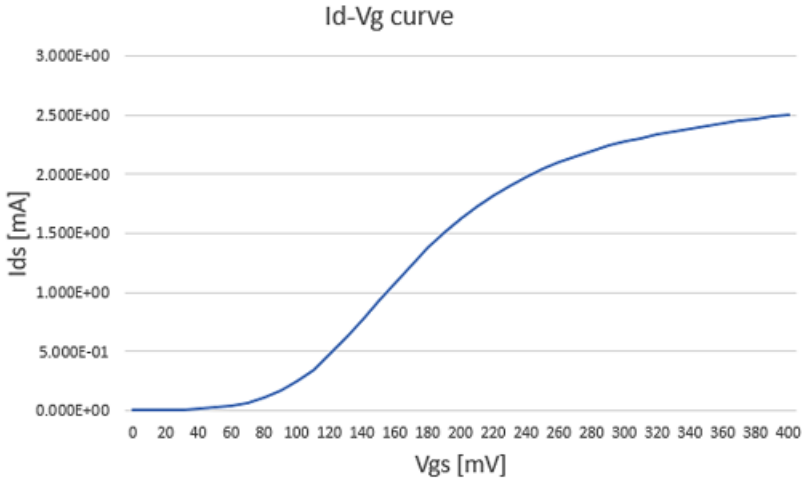


Figure 4.4: I_D-V_G characteristic measured on the chip before bonding

Then, several bonding tests were performed by changing the applied temperature and force, setting the bonding time to 4 minutes, which should be more than enough to form a reliable joint with a gold-gold thermocompression technique. For these tests, it was decided to start with quite high temperature and force, so that it was relatively easy to achieve a complete bonding while, at the same time, gaining practice with the bonding process. Then, the temperature was decreased in steps of 50 °C at a time and starting with a quite high force (10 N). If the formed bond was found to be successful, then the temperature was kept constant while trying to minimize the applied force.

The main aim of these tests was to minimize the bonding temperature, since this would lead to a lower risk of damaging the chip circuitry and it could also allow to reuse the developed bonding process for other materials presenting heat-related concerns, such as piezoelectric materials, which loose their properties beyond a certain temperature

called "Curie temperature".

The bonding parameters for the tests that were found to be successful after the electrical test are reported in Table 4.1: the lowest temperature that allowed to achieve a complete bond was 150 °C, even if a pretty high force (10 N) was needed; on the opposite corner of the table, the lowest bonding force was 2 N, even if it required a quite high temperature (350 °C).

Temp/Force	10N	8N	7N	5N	2N
350°C	10N-350°C-4min	8N-350°C-4min	7N-350°C-4min		2N-350°C-4min
300°C		8N-300°C-4min	7N-300°C-4min	5N-300°C-4min	
250°C	10N-250°C-4min	8N-250°C-4min			
200°C	10N-200°C-4min				
150°C	10N-150°C-4min				

Table 4.1: Successful bonding tests

As it can be noticed in Table 4.1, the tests were conducted with the purpose of improving the minimum temperature and force required to obtain a successful bond. For reasons of time, it was not possible to further optimize the bonding process, which would have require to fill up the blank spots in the table towards its bottom-right corner, by achieving a successful bond with those combinations of parameters.

The I_D-V_G characteristics measured for these successfully-bonded chips are reported in Figure 4.5.

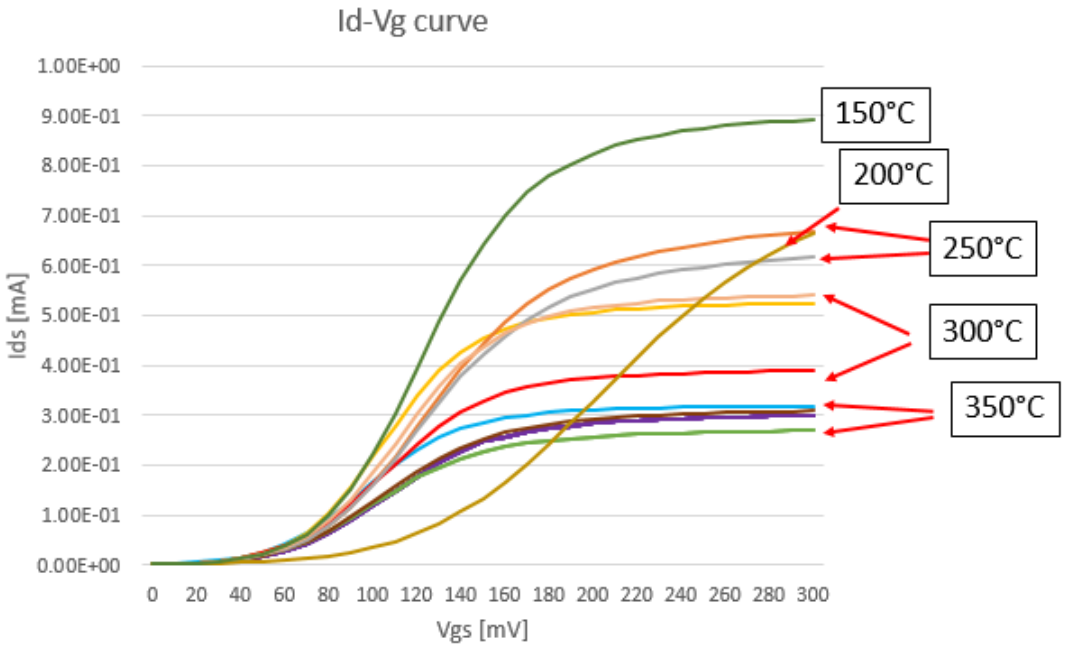


Figure 4.5: I_D-V_G characteristics measured on the successfully-bonded chips

From this graph, it can be noticed that the transistor saturation current increases as the bonding temperature decreases, suggesting the presence of a temperature-related mechanism that affects the transistor performance. Even with the lowest temperature (150 °C) the saturation current measured for the bonded chip (0.9 mA) is almost one third of the saturation current measured for the chip alone before bonding (2.7 mA). The relation between bonding temperature and saturation current is linear, as reported in Figure 4.6 using the values in Table 4.2.

Temp [°C]	Sat_current [mA]
150	9.07E-01
200	8.50E-01
250	0.662
300	0.487
350	3.01E-01

Table 4.2: Saturation current for different bonding temperatures

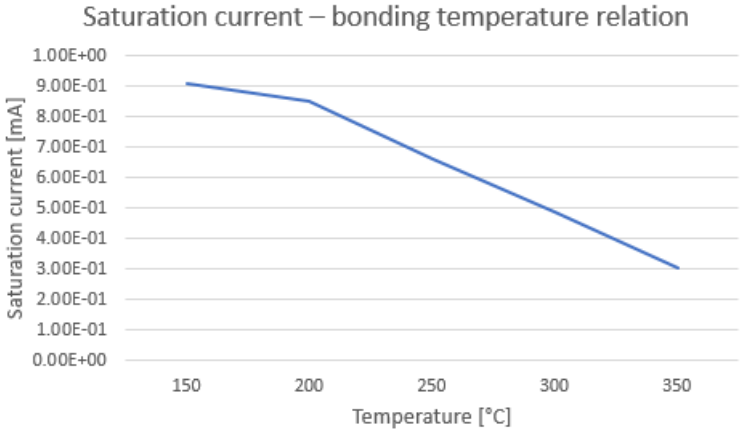


Figure 4.6: Saturation current - bonding temperature relation

The causes of this issue have not been understood since there was not enough time to further investigate them by performing more tests. However, it has to be said that the region of interest for the transistor operation is below 10 mV, since the maximum expected voltage for a neural signal is in the order of few mV (the transistor threshold is in the order of few µV). In this region, all the tests exhibit the same behavior, which is also the one measured on the chip before bonding (I_{DS}~2-3 µA), thus meaning that the developed bonding technique is suitable for this specific application.

Another way to check if the chip was completely bonded was to measure the impedance between transistor source and center pads ("S"-"C" impedance) and between transistor drain and center pads ("D"-"C" impedance) using a multimeter on the substrate test pads; indeed, between these couples of pads there is a resistor with a fixed and know value of 200 kΩ, called R_b in Figure 4.1. In this way, it was also possible to test the connection of single pads in case of non complete bonding: if the impedance was close to 200 kΩ, the pad was bonded; otherwise, if the impedance was very high the pad was not bonded (open-circuit) or, if it was very low, the pad was contacting some other metallic element (short-circuit), but in both cases it was not working properly.

4.2 Shear test results

The last step for the testing was to perform a shear test in order to evaluate the bond strength. For this purpose, a professional bond tester tool was used, which is the Nordson® DAGE 4000 bondtester shown in Figure A.24 of Appendix A. The bond tester has a tool with a metallic tip that pushes the chip on its lateral side with increasing force until the bond breaks, thus reporting the results in a graph like the one shown in Figure 4.7, which shows the applied force as a function of time (but it can be displayed also as a function of the tool spatial displacement). The graph peak corresponds to the maximum applied force, which is the moment when the bond breaks and the chip detaches from the substrate, thus indicating the bond strength; then, the applied force abruptly plunges but it does not go to zero because the tool keeps pushing the chip on the substrate for a certain distance, thus requiring the application of a small force to win the friction. The shear velocity was set to 20 µm/s and the shear height to 50 µm, so that the tool tip was pushing in the center of the chip lateral face. Moreover, the test type was set to "destruct", which means the tool keeps pushing with an increasing force until either the chip detaches from the substrate (the bond breaks) or it reaches the maximum tool force (250 g for the used tool). The force is measured in gram-force (g_f).

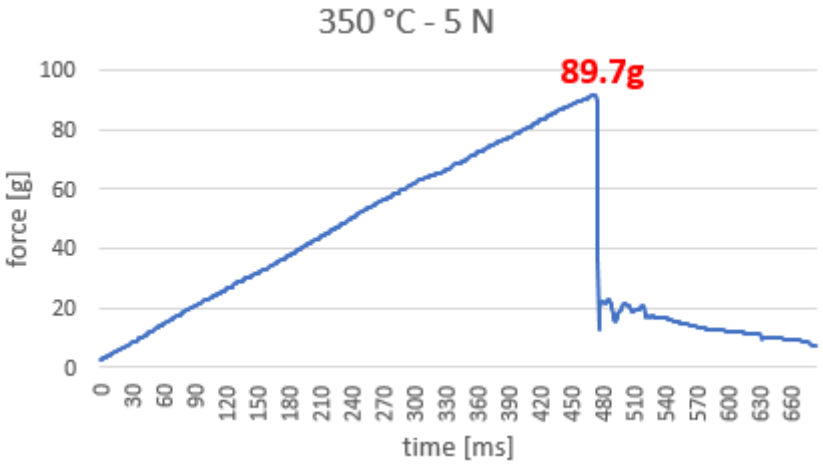
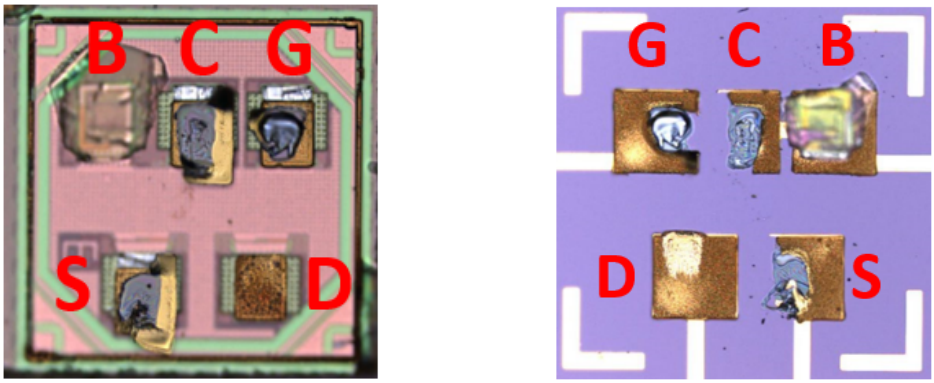


Figure 4.7: Example of graph resulting from a shear test

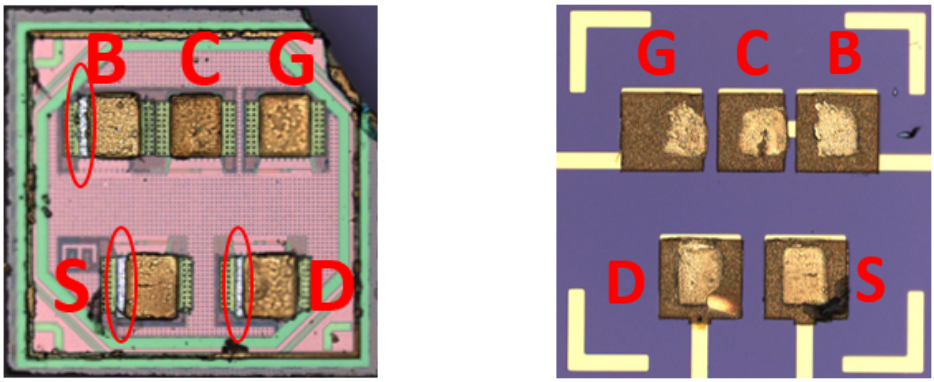
Figure 4.8 and Figure 4.9 show two chip-substrate couples after the shear test (the chip pad configuration is flipped with respect to the one of the substrate pads). As it can be observed in these pictures, in some cases the gold layer on the substrate pad gets removed and it sticks to the corresponding chip pad: this means that the bond between the gold layers on the two sides was stronger than the adhesion force between silicon oxide and evaporated chrome/gold on the substrate side, as shown in Figure 4.8 for the pads "C", "G" and "S". In some other cases it happened the opposite situation: the chip pad metallization detached from the chip and sticked to the matching substrate pad, as shown in Figure 4.8(b) for the "B" pad, where the whole pad metallization was ripped-off the chip together with some of the passivation layer around it.



(a) Chip after shear (b) Substrate after shear

Figure 4.8: Post-shear chip and substrate (300 °C - 8 N)

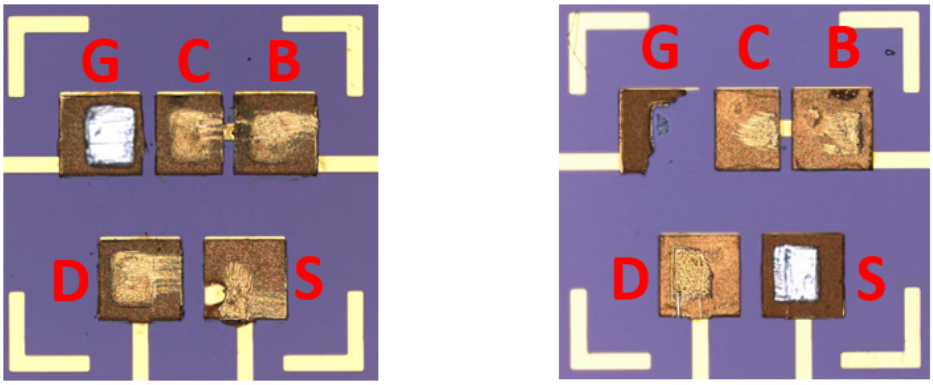
In other cases the bond between the gold layers was weaker and the shear test only caused a displacement of the metal layers plated on the chip pads in the direction of the applied shear force, as shown in Figure 4.9(a) inside the red circles at the edges of the pads.



(a) Chip after shear (b) Substrate after shear

Figure 4.9: Post-shear chip and substrate (350 °C - 7 N)

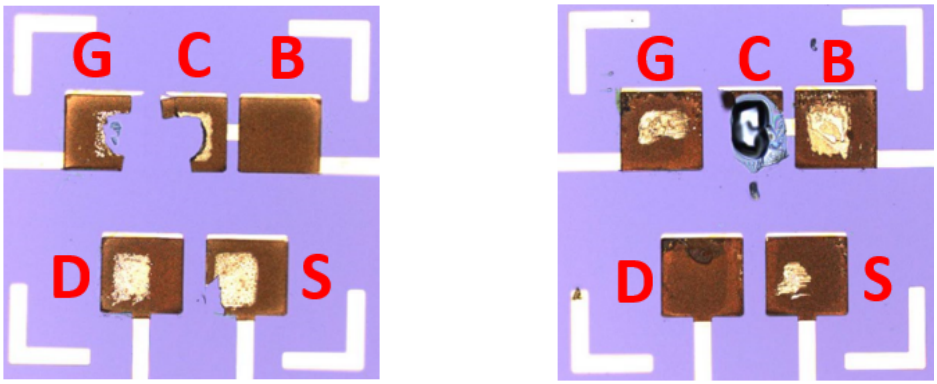
Other cases of chip pads sticking to the substrate pads after the shear test are shown in Figure 4.10, on the "G" pad in Figure 4.10(a) and on the "S" pad in Figure 4.10(b).



(a) (b)

Figure 4.10: Examples of chip pad sticking to the substrate after shearing

More examples of gold layers ripped-off the substrate after the shear test are shown in Figure 4.11, on the "G" and "C" pads in Figure 4.11(a) and on the "C" pad in Figure 4.11(b), where it looks like not only the gold layer is removed but also the underlying silicon dioxide layer is partially ripped off. In some cases there were pads that did not contact at all with the chip pads, as in the case of the "B" pad in Figure 4.11(a), meaning that the bonding was not completely successful.



(a) (b)

Figure 4.11: Examples of gold layers ripped off the substrate after the shear test

The bond strength values resulting from the performed shear tests are reported in Table 4.3.

Bonding parameters	Bond strength [g_f]
350°C-10N	103.5
350°C-8N	103
350°C-7N	70.8
350°C-5N	89.7
350°C-3N	41.8
350°C-2N	5.5
300°C-10N	70.6
300°C-8N	20.6
300°C-7N	43.4
300°C-5N	36.6
250°C-10N	22.3
250°C-5N	25.2

Table 4.3: Bond strength measured values

As it can be seen from the values in the table, the shear test was performed also on some bonded chips for which the electrical test did not result fully successful, which means they were partially bonded; this was done in order to collect more data which could help to better define the bond strength trend. From the data reported in Table 4.3 it can be noticed that, as expected, the bond strength decreases with the bonding force for a fixed temperature; moreover, using the same bonding force range, the bond strength range is shifted towards much lower values when the temperature is reduced in steps of 50 °C.

The obtained results are compatible with the military requirements in MIL-STD-883E, imposing a minimum die shear force of ~100 g for a die of ~2.5×10^{-4} $inches^2$, correspondent to the ~400 µm×400 µm chip used in this work.

Chapter 5

Conclusions and possible improvements

In this master thesis it has been realized a **flip chip bonding technique** using a **gold-gold thermocompression process**, developed for the specific application of assembling sub-mm Neural Dust motes.

A promising idea to hold tiny chips has been devised, using a polypropylene mesh sheet that has been cut and shaped as a "boat" with a covering "lid", with a mesh pitch such that the chips cannot escape while liquids can flow through.

The **ENIG plating process** has been optimized to obtain a smooth and uniform gold layer on the chip pads; moreover, two types of chips with different pad layout and number have been used, demonstrating that the developed plating process is reliable and reproducible on different chips.

The **substrates** have been fabricated and the short-circuit issue related to the chip seal ring has been solved by electroplating tall gold pillars on the substrate target pads.

The **flip chip bonding** has been performed using different combinations of applied force and heating temperature; the **electrical testing** on the bonded chips has proved that the developed bonding technique can work for several combinations of force and temperature. Moreover, the bonding process has been partially optimized, demonstrating that it is possible to achieve complete bonding with temperatures as low as 150 °C and applied forces as low as 2 N. The main concern was the temperature which could damage the chip circuitry, thus the principal goal was to minimize the temperature necessary to achieve a successful bonding, while the applied force should not affect the chip as long as it is not so high to crack it. The **mechanical testing** has shown that it is possible to obtain bond strengths compatible with the military requirements indicated in MIL-STD-883E; moreover, the inspection of the chips and substrates after the shear test allows to notice that in many cases the gold-gold bond between the pads on the two sides was stronger than the gold adhesion force with the underlying layers. These results can

thus be considered completely satisfying and acceptable for this specific application.

Possible **future improvements** for the whole process are:

- fabricate other substrates with a bus bar running along the wafer edge to connect all the pads together, thus allowing to electroplate them all at the same time by attaching an electrode to the bus bar and obtaining a uniform gold thickness
- use a additional autocatalytic gold plating step to obtain a thicker gold layer on the chip pads, which should improve the pad wettability and lead to higher bond strengths
- use thermosonic bonding by adding ultrasonic energy to the thermocompression bonding process, which has been shown in [81] to lower both the bonding temperature and force
- use an heated pick-up tool for the thermocompression bonding, which could allow to shorten the bonding time since it would not be necessary to wait for the chip to heat up through the substrate
- use an underfill epoxy to improve the bond strength, which also requires to figure out how to precisely dispense the epoxy exclusively underneath the chip
- use UV or UV-ozone cleaning to remove the contaminants on the gold surface before bonding, which has been shown in [82, 83] to improve the bond strength
- use argon or argon-oxygen RF-plasma activation of the gold surface before bonding, which has been shown in [84] to lower the minimum bonding force and temperature
- use post-heating of the bonded chips, which in [83] has been demonstrated to improve the bond strength in case of discontinuous metallic bond interfaces by growing them.

Bibliography

[1] M.M. Fouad, K.M. Amin, N. El-Bendary, A.E. Hassanien, *Brain Computer Interface: A Review*, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[2] R.A. Ramadan et al., *Basics of Brain Computer Interface*, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[3] C. Eliasmith and C. H. Anderson. *Neural Engineering*, MIT Press, (2003), in: M. Rossi, *Interfacce cervello-computer: stato dell'arte*, Università di Bologna, Thesis, a.a. 2013-2014.

[4] Pontus Forslund, *A Neural Network Based Brain Computer Interface for Classification of Movement Related EEG*, LiTH-IKP-EX-2107 Linkoping, December 2003.

[5] A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[6] Repovas, G.: *Dealing with noise in EEG recording and data analysis*, Informatica Medica Slovenica J. 15(1), pp. 18–25 (2010), in: A.E. Hassanien, A.T. Azar (eds), Brain-Computer Interfaces, Intelligent Systems Reference Library 74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[7] Olav Krigolson, *That Neuroscience Guy*, available at: http://www.olavkrigolson.com/that-neuroscience-guy, accessed on 22/02/2018.

[8] Freeman, W.J., Holmes, M.D., Burke, B.C., Vanhatalo, S.: *Spatial spectra of scalp EEG and EMG from awake humans*, Clin. Neurophysiol. 114(6), pp. 1053–1068 (2003), in: A.E. Hassanien, A.T. Azar (eds), Brain-Computer Interfaces, Intelligent Systems Reference Library 74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[9] Schwartz, A.B., Cui, X.T., Weber, D.J., Moran, D.W.: *Brain-controlled interfaces: movement restoration with neural prosthetics*, Neuron 52(1), pp. 205–220 (2006), in: A.E. Hassanien, A.T. Azar (eds), Brain-Computer Interfaces, Intelligent Systems

Reference Library 74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[10] Leuthardt, E.C., Schalk, G., Wolpaw, J.R., Ojemann, J.G., Moran, D.W.: *A brain-computer interface using electrocorticographic signals in humans*, J. Neural Eng. 1(2), pp. 63–71 (2004), in: A.E. Hassanien, A.T. Azar (eds), Brain-Computer Interfaces, Intelligent Systems Reference Library 74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[11] Nicolas-Alonso, L.F., Gomez-Gil, J.: *Brain computer interfaces*, a review. Sensors 12(2), 1211–1279 (2012), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[12] Baillet, S., Mosher, J.C., Leahy, R.M.: *Electromagnetic brain mapping*, Sig. Process. Mag. IEEE 18(6), 14–30 (2001), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[13] Cohen, D.: *Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer*, Science 175(4022), pp.664–666 (1972), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[14] Zamrini, E., Maestu, F., Pekkonen, E., Funke, et al, *Magnetoencephalography as a putative biomarker for Alzheimer's disease*, Int. J. Alzheimer's Dis, vol.2011, Article 280289, pp.1-10, (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[15] The National Research Council (NRC): Canada, Laboratory for Clinical Magnetoencephalography, ARCHIVED—magnetoencephalography system, available at: https://www.nrc-cnrc.gc.ca/eng/achievements/highlights/2010/meglab.html, accessed on 22/02/2018.

[16] Homer, M.L., Nurmikko, A.V., Donoghue, J.P., Hochberg, L.R.: *Implants and decoding for intracortical brain computer interfaces*, Annu. Rev. Biomed. Eng. 15, pp.383–405 (2013), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[17] Rousche, P.J.: *Bio-MEMS: designs and applications of cortical interfaces for neuroscience and neuroprosthetics*, Lecture notes, Department of Bioengineering, University of Illinois, Chicago, USA. http://tigger.uic.edu/classes/bioe/bioe200/Rouschelecture.ppt (2014), in: A.E.

Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[18] Richards, T.L., Berninger, V.W.: *Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment*, J. Neurolinguistics 21(4), pp.294–304 (2008), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[19] Wass, S.: *Distortions and disconnections: disrupted brain connectivity in autism*, Brain Cogn. 75(1), pp.18–28 (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[20] Tushar Kanti Bera, *Noninvasive Electromagnetic Methods for Brain Monitoring: A Technical Review*, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[21] Faiz, O., Blackburn, S., Moffat, D.: *Anatomy at a Glance*, 3rd edn. Wiley-Blackwell, Chichester (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[22] Szi-Wen, C., Chang-Yuan, C.: *A comparison of 3D cone-beam Computed Tomography (CT) image reconstruction performance on homogeneous multi-core processor and on other processors*, Measurement 44(10), pp. 2035–2042 (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[23] Li-ming, Wang, Ying-liang, Zhao, Fang-lin, Chen, Han, Y.: *The 3D CT reconstruction algorithm to directly reconstruct multi-characteristic based on EMD*, Measurement 44(10), pp. 2043–2048 (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[24] Brandon, D., Alazraki, A., Halkar, R.K., Alazraki, N.P.: *The role of single-photon emission computed tomography and SPECT/computed tomography in oncologic imaging*, Semin. Oncol. 38(1), pp. 87–108 (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[25] Cho, M.J., Lyoo, I.K., Lee, D.W., Kwon, J.S., Lee, J.S., Lee, D.S., Jung, J.K., Lee, M.C.: *Brain single photon emission computed tomography findings in depressive*

pseudodementia patients, J. Affect. Disord. 69(1–3), pp. 159–166 (2002), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[26] Devous, M.D.: *Single-photon emission computed tomography in neurotherapeutics*, NeuroRx. 2(2), pp. 237–249 PMCID: PMC1064989 (2005), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[27] Imperiale, C., Imperiale, A.: *Some fast calculations simulating measurements from singlephoton emission computed tomography (SPECT) imaging*, Measurement 37(3), pp. 218–240 (2005), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[28] Meeus, I., Galdermans, D., Roland, J., Slabbynck, H., van Schaardenburg, C., Coolen, D.: *Spectacular bone scintigraphy and transverse SPECT images in primary osteosarcoma of the pleura*, Clin. Nucl. Med. 19(8), pp. 738–740 (1994), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[29] Piwnica-Worms, D.: *Clinical molecular imaging today: PET and SPECT*, in: Kufe, D.W., Pollock, R.E., Weichselbaum, R.R., et al. (eds.) Holland-Frei Cancer Medicine, 6th edn. BC Decker, Hamilton (ON) (2003), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[30] Facey, K., Bradbury, I., Laking, G., Payne, E.: *Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers*, Health Technol. Assess. 11(44), iii–iv, xi–267 (2007), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[31] Frey, K.A.: *Positron emission tomography*, in: Siegel, G.J, Agranoff, B.W., Albers, R.W. et al., (eds.), *Basic neurochemistry: molecular, cellular and medical aspects*, Chap. 54, 6th edn. Lippincott-Raven, Philadelphia, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[32] Internet Article: *Technology—positron emission tomography (PET) imaging*, available at: http://www.cellsighttech.com/technology/pet.html, 2014, accessed 22/02/2018, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[33] Ganis, G., Kosslyn, S.M.: *Neuroimaging*, in: Ramachandran V.S. (ed.), *Encyclopedia of the human brain*, pp. 493–505 (2002), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[34] Bayford, R.H.: *Bioimpedance tomography (electrical impedance tomography)*, Ann. Rev. Biomed. Eng. 8, pp. 63–91 (2006), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[35] Bera, T.K., Nagaraju, J.: *Studying the resistivity imaging of chicken tissue phantoms with different current patterns in electrical impedance tomography (EIT)*, Measurement 45(2012), pp. 663–682 (2012), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[36] Bera, T.K.: *Studies on multifrequency multifunction electrical impedance tomography (MfMf-EIT) to improve bio-impedance imaging*, PhD Thesis, IISc, Bangalore, India, Department of Instrumentation and Applied Physics, Indian Institute of Science, Bangalore, India (2013), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[37] Borcea, L.: *Electrical impedance tomography*, Topical Rev. Inverse Probl. 18, R99–R136 (2002), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[38] Cheney, M., Isaacson, D., Newell, J.C.: *Electrical impedance tomography*, SIAM Rev. 41(1), pp. 85–101 (1999), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[39] Denyer, C.W.L.: *Electronics for real-time and three-dimensional electrical impedance tomographs*, PhD Thesis, Oxford Brookes University (1996), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[40] Graham, B.M.: *Enhancements in electrical impedance tomography (EIT) image reconstruction for 3D Lung Imaging*, PhD thesis, University of Ottawa (2007), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[41] Holder, D.S., Hanquan, Y., Rao, A.: *Some practical biological phantoms for calibrating multifrequency electrical impedance tomography*, Physiol. Meas. 17, A167–A177 (1996), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[42] Seo, J.K., Woo, E.J.: *Nonlinear Inverse Problems in Imaging*, 1st edn. Wiley, New York (2012), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[43] Webster, J.G.: *Electrical impedance tomography*, Adam Hilger Series of Biomedical Engineering, Adam Hilger, New York (1990), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[44] Bera, T.K., Nagaraju, J.: *Sensors for electrical impedance tomography*, 2nd edn., in: Webster J.G. (ed.), *The measurement, instrumentation, and sensors handbook*, Chap. 61, pp 61.1–61.30. CRC Press (2014), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[45] Liu, H., Tao, X., Xua, P., Zhang, H., Bai, Z.: *A dynamic measurement system for evaluating dry bio-potential surface electrodes*, Measurement 46, pp. 1904–1913 (2013), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[46] Wang, P., Guo, B., Li, N.: *Multi-index optimization design for electrical resistance tomography sensor*, Measurement 46(8), pp. 2845–2853 (2013), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[47] Bera, T.K., Mohamadou, Y., Lee, K.H., Wi, H., Oh, T.I., Woo, E.J., Soleimani, M., Seo, J.K.: *Electrical impedance spectroscopy for electro-mechanical characterization of conductive fabrics*, Sensors 14, pp. 9738–9754 (2014), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[48] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *Improving conductivity image quality using block matrix-based multiple regularization (BMMR) technique in EIT: a simulation study*, J. Electr. Bioimpedance 2, pp. 33–47 (2011), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[49] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *Improving the image reconstruction in electrical impedance tomography (EIT) with block matrix-based multiple regularization (BMMR): a practical phantom study*, IEEE World Congr. Inform. Commun. Technol. (WICT-2011) India 2011, pp. 1346–1351 (2011b), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[50] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *Improving image quality in electrical impedance tomography (eit) using projection error propagation-based regularization (pepr) technique: a simulation study*, J. Electr. Bioimpedance 2, pp. 2–12 (2011). doi:10.5617/jeb.158, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[51] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *A model based iterative image reconstruction (MoBIIR) algorithm for conductivity imaging in EIT using simulated boundary data*, in: AIP Conference Proceedings, Optics: Phenomena, Materials, Devices, and Characterization: Optics 2011: International Conference on Light, Kerala, (India), pp. 489–491. 23–25 May (2011d), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[52] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *Image reconstruction in electrical impedance tomography (EIT) with projection error propagation-based regularization (PEPR): a practical phantom study*, in: Lect. Notes Comput Sci Springer, 2012, 7135/2012, pp. 95–105, ADCONS 2011 (2012), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[53] Bera, T.K., Biswas, S.K., Rajan, K., Nagaraju, J.: *Projection Error Propagation-based regularization (PEPR) method for resistivity reconstruction in electrical impedance tomography (EIT)*, Measurement 49, pp. 329–350 (2014), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[54] Lionheart, W.R.B.: *EIT reconstruction algorithms: pitfalls, challenges and recent developments*, REVIEW ARTICLE. Physiol. Meas. 25(2004), pp. 125–142 (2004), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[55] Stephenson, D.R., Davidson, J.L., Lionheart, W.R.B., Grieve, B.D., York, T.A.: *Comparison of 3D image reconstruction techniques using real electrical impedance measurement data*, in: 4th World Congress on Industrial Proceedings Tomography, 2005, Japan, pp. 1–8, in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer*

Interfaces, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[56] Yorkey, T.J.: *Comparing reconstruction methods for electrical impedance tomography*, PhD thesis, University of. Wisconsin at Madison, Madison (1986), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[57] 221. Yorkey, T.J., Webster, J.G., Tompkins, W.J.: *Comparing reconstruction algorithms for electrical impedance tomography*, IEEE Trans. Biomed. Eng. BME-34(11), pp. 843–852 (1987), in: A.E. Hassanien, A.T. Azar (eds), *Brain-Computer Interfaces*, Intelligent Systems Reference Library, Vol.74, DOI 10.1007/978-3-319-10978-7_1 (2015).

[58] John Denis Enderle, Joseph D. Bronzino, Susan M. Blanchard, *Introduction to Biomedical Engineering*, second edition, Academic Press series in biomedical engineering (2005).

[59] Dongjin Seo, Jose M. Carmena, Jan M. Rabaey, Michel M. Maharbiz, Elad Alon, *Model validation of untethered, ultrasonic neural dust motes for cortical recording*, Journal of Neuroscience Methods, No.244, pp. 114–122, 2015.

[60] FDA, Information for Manufacturers Seeking Marketing Clearance of Diagnostic Ultrasound Systems and Transducers (2008).

[61] Dongjin Seo, Jose M. Carmena, Jan M. Rabaey, Elad Alon, Michel M. Maharbiz, *Neural Dust: An Ultrasonic, Low Power Solution for Chronic Brain-Machine Interfaces*, arXiv preprint arXiv:1307.2196, available at: https://arxiv.org/pdf/1307.2196.pdf, accessed on 22/02/2018.

[62] Dongjin Seo, Ryan M. Neely, Konlin Shen, Utkarsh Singhal, Elad Alon, Jan M. Rabaey, Jose M. Carmena, and Michel M. Maharbiz, *Wireless Recording in the Peripheral Nervous System with Ultrasonic Neural Dust*, Neuron, Vol. 91, pp. 529–539, August 3, 2016.

[63] Dongjin Seo, Ryan M. Neely, Konlin Shen, Utkarsh Singhal, Elad Alon, Jan M. Rabaey, Jose M. Carmena, and Michel M. Maharbiz, *Supplemental Information, "Wireless Recording in the Peripheral Nervous System with Ultrasonic Neural Dust"*, Neuron, Vol. 91, pp. 529–539, August 3, 2016.

[64] *Chip Bonding at the First Level*, "Roadmaps of packaging technologies", Integrated Circuit Engineering Corporation (ICE), available at: http://smithsonianchips.si.edu/ice/cd/PKG_BK/CHAPT_09.PDF, accessed on 22/02/2018.

[65] Leif Halbo and Per Ohlckers, *Electronics Components, Packaging and Production*, Chapter 2, available at: http://tid.uio.no/kurs/fys4260/chap2.html, accessed on 22/02/2018.

[66] Suwanna Jittinorasett, *UBM Formation on Single Die/Dice for Flip Chip Applications*, Master Thesis, Virginia Polytechnic Institute and State University, August 1999.

[67] Martin Tarr, *Tape Automated Bonding*, archive.today, available at: http://archive.is/54b3, accessed on 22/02/2018.

[68] Prof. G.V. Mahesh, *An Introduction to Electronics Systems Packaging*, Department of Electronic system Engineering, IISc Bangalore, available at: https://www.youtube.com/watch?v=vKj3eXgyC5c&t=357s, accessed on 22/02/2018.

[69] Joseph Fjelstad, Tape Automated Bonding, PCMAG Encyclopedia, available at: https://www.pcmag.com/encyclopedia/term/52555/tape-automated-bonding, accessed on 22/02/2018.

[70] "Stencil printing", Open Source Studio, EGO process, available at https://oss.adm.ntu.edu.sg/linc0013/tag/process/, accessed on 22/02/2018.

[71] Chapter B: Flip-Chip Technology, available at: http://extra.ivf.se/ngl/b-flip-chip/ChapterB1.htm, accessed on 22/02/2018.

[72] CL Wong and James How, *Low Cost Flip Chip Bumping Technologies*, IEEE/CPMT Electronic Packaging Technology Conference, pp. 244-250, 1997.

[73] Kiichi Niitsu, Shoko Ota, Kohei Gamo, Hiroki Kondo, Masaru Hori, and Kazuo Nakazato, *Development of Microelectrode Arrays Using Electroless Plating for CMOS-Based Direct Counting of Bacterial and HeLa Cells*, IEEE Transactions on Biomedical Circuits and Systems, Vol. 9, No. 5, pp. 607-619, October 2015.

[74] Madhumita Datta, Scott A. Merritt and Mario Dagenais, *Electroless Remetallization of Aluminum Bond Pads on CMOS Driver Chip for Flip-Chip Attachment to Vertical Cavity Surface Emitting Lasers (VCSEL's)*, IEEE Transactions on Components and Packaging Technology, Vol. 22, No. 2, pp. 299-306, June 1999.

[75] Esther Wai Ching Yau, Jing Feng Gong, Philip Chan, *Al surface morphology effect on flip-chip solder bump shear strength*, Microelectronics Reliability, No. 44, pp. 323–331, 2004.

[76] Yu Yang, Jian Cai, Shuidi Wang, Songliang Jia, *The Impact of Zincation on the Electroless Nickel UBM for Low Cost Flip Chip Technology*, 6th International Conference on Electronic Packaging Technology, IEEE, 2005.

[77] Jeong-Gi Jin, Sung-Ki Lee, Young-Ho Kim, *Adhesion improvement of electroless plated Ni layer by ultrasonic agitation during zincating process*, Thin Solid Films, Vol. 466, South Korea, pp. 272-278, 2004.

[78] George Milad, *Is "black pad" still an issue for ENIG?*, Circuit World, Vol. 36, No. 1, pp. 10–13, 2010.

[79] R. Ramanauskas, A. Selskis, J. Juodkazyte and V. Jasulaitiene, *PCB failure analysis related to the ENIG black pad problem*, Circuit World, Vol. 39, No. 3, pp. 124–132, 2013.

[80] LOR and PMGI Resists, available at: http://microchem.com/pdf/PMGI-Resists-data-sheetV-rhcedit-102206.pdf, accessed on 22/02/2018.

[81] Thi Thuy Luu, Hoang-Vu Nguyen, Andreas Larsson, Nils Hoivik and Knut E.Aasmundtveit, *Gold to gold thermosonic bonding. Characterization of bonding parameters*, available at: https://brage.bibsys.no/xmlui/bitstream/handle/11250/2431517/SINTEF%2BS17112.pdf?sequence=1&isAllowed=y, accessed on 6/03/2018.

[82] Naoko Unami et al., *Effects of Excimer Irradiation Treatment on Thermocompression Au–Au Bonding*, Jpn. J. Appl. Phys. 49 06GN12, 2010.

[83] James J. Jellison, *Effect of Surface Contamination on the Thermocompression Bondability of Gold*, IEEE Transactions on Parts, Hybrids, and Packaging, Vol. PHP-11, No. 3, pp. 206-211, September 1975.

[84] Eiji Higurashi, Teppei Imamura, Tadatomo Suga, and Renshi Sawada, *Low-Temperature Bonding of Laser Diode Chips on Silicon Substrates Using Plasma Activation of Au Films*, IEEE Photonics Technology Letters, Vol. 19, No. 24, pp. 1994-1996, December 15, 2007.

Appendix A

Cleanroom equipment

Most of the work presented in this thesis has been conducted inside the Marvell Nanofabrication Laboratory at the University of California, Berkeley.
For this work, I used several tools of the lab equipment:

- WAFAB sinks for different wet processes, such as for performing the ENIG plating process, which requires the vented hood of the sink, or other dedicated bath, such as piranha-cleaning or PR-stripping. One of the sinks is shown in Figure A.1.

Figure A.1: Wet processes sink with vented hood

- TYSTAR Silicon Oxidation Furnace, shown in Figure A.2.

Figure A.2: TYSTAR silicon oxidation furnace

The furnace boat loaded with wafers is shown in Figure A.3, while Figure A.4 shows some oxidized wafers on the cooling rack.

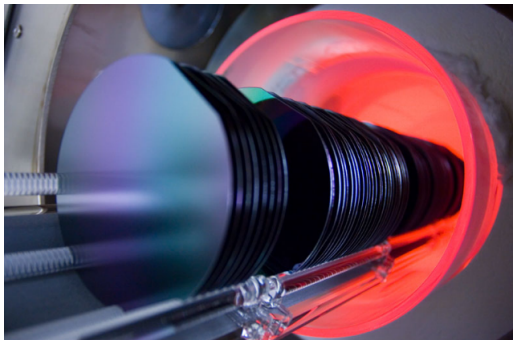


Figure A.3: Furnace boat loaded with wafers

Figure A.4: Oxidized silicon wafers

- HMDS oven, used for wafer priming to improve photoresist adhesion, shown in Figure A.5.

Figure A.5: HMDS priming oven

- Matrix 106 Resist Removal System, used to clean wafers from photoresist, shown in Figure A.6.

Figure A.6: Matrix Resist Removal System

- SVG 8626 model photoresist coater, which automatically performs photoresist spinning and soft-baking according to the selected program, shown in Figure A.7.

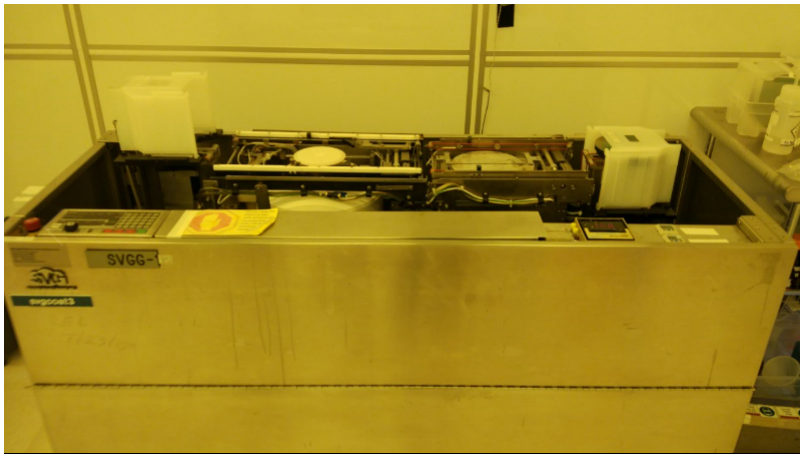


Figure A.7: SVG photoresist coater

- PICOTRACK photoresist coat and development system, with two different tracks that automatically perform: HMDS priming, PR spinning and soft-baking for the first track, post-exposure bake and PR development for the second track. The tool is shown in Figure A.8.

Figure A.8: PICOTRACK photoresist coat and development system

- Headway Manual Load Photoresist Spinner, shown in Figure A.9.

Figure A.9: Manual photoresist spinner

The photoresist dispensing on the wafer before spinning is shown in Figure A.10.

Figure A.10: Photoresist dispensing on the wafer

- Quintel Q4000 Mask Aligner, which allows to perform the mask alignment to the wafer and then the contact exposure with a UV-lamp, shown in Figure A.11.

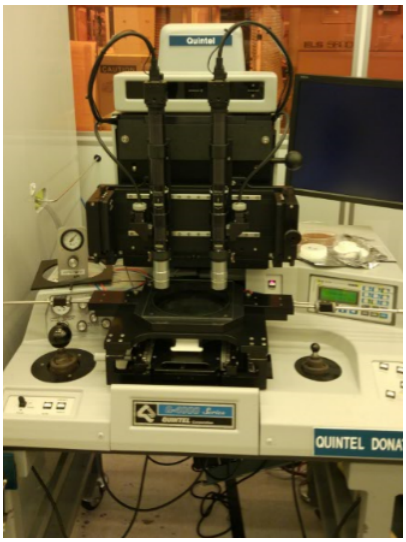


Figure A.11: Quintel mask aligner and UV contact exposure system

- CHA Solution Electron-beam evaporator system, shown in Figure A.12.

Figure A.12: CHA e-beam evaporator system

The inner chamber of the e-beam evaporator with the rotating turret is shown

in Figure A.13, while Figure A.14 shows two graphite crucibles filled with formed metal melt.

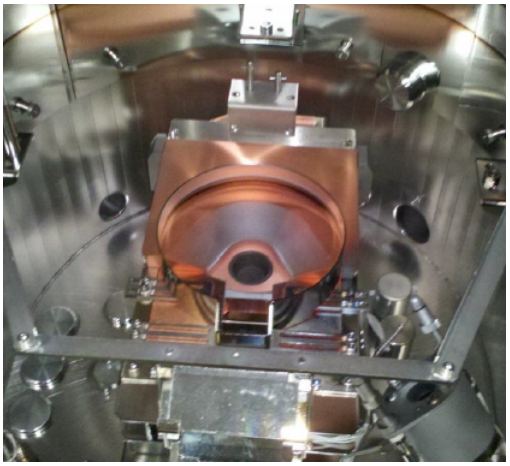


Figure A.13: Inner view of the CHA chamber

Figure A.14: Crucibles filled with formed metal melt

- Technics-C Plasma Etching System used with O2-plasma for photoresist descum, shown in Figure A.15.

Figure A.15: Technics-C Plasma Etching System

- DISCO DAD3240 dicing saw used to dice silicon wafers, shown in Figure A.16.

Figure A.16: DISCO dicing saw

The silicon blade is shown in Figure A.17, while Figure A.18 shows the dicing saw in operation with flowing water to cool down the blade and remove silicon dust from the cuts.

Figure A.17: Silicon blade

Figure A.18: Dicing saw in operation with flowing water

- Olympus LEXT OLS4000 3D Confocal Laser Microscope, shown in Figure A.19.

Figure A.19: Olympus Confocal Laser Microscope

- Finetech FINEPLACER 96 Lambda, wihch is a flip chip bonder tool, shown in Figure A.20.

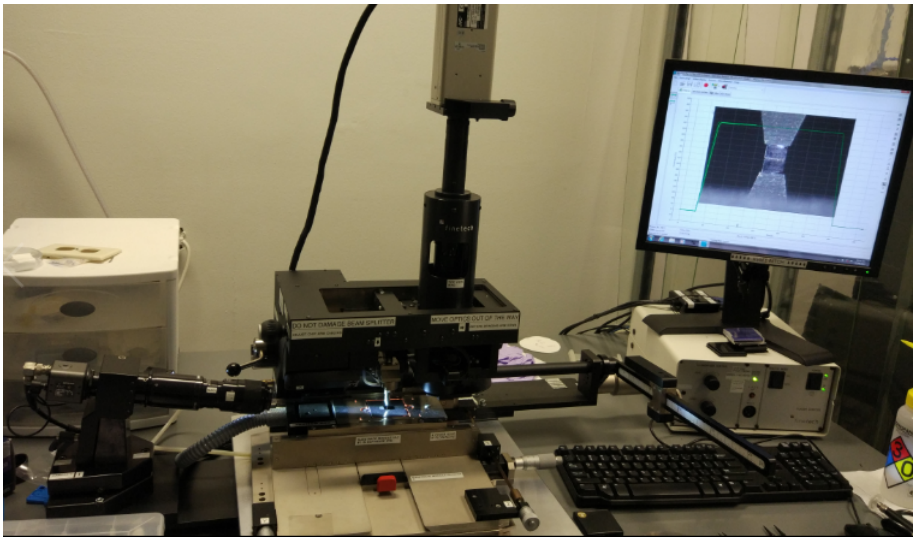


Figure A.20: Finetech flip chip bonder

A zoomed view of the heated plate during the bonding process is shown in Figure A.21.

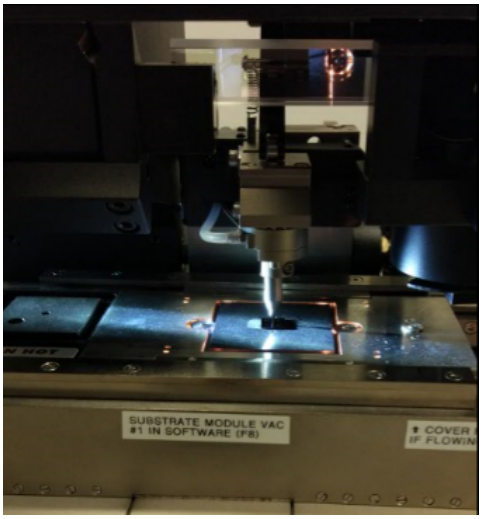


Figure A.21: Finetech flip chip bonder zoom on the heated plate during bonding

- Probe station for electrical testing, shown in Figure A.22 with the complete electronic test-bench setup of source-meters, multimeters and oscilloscopes.

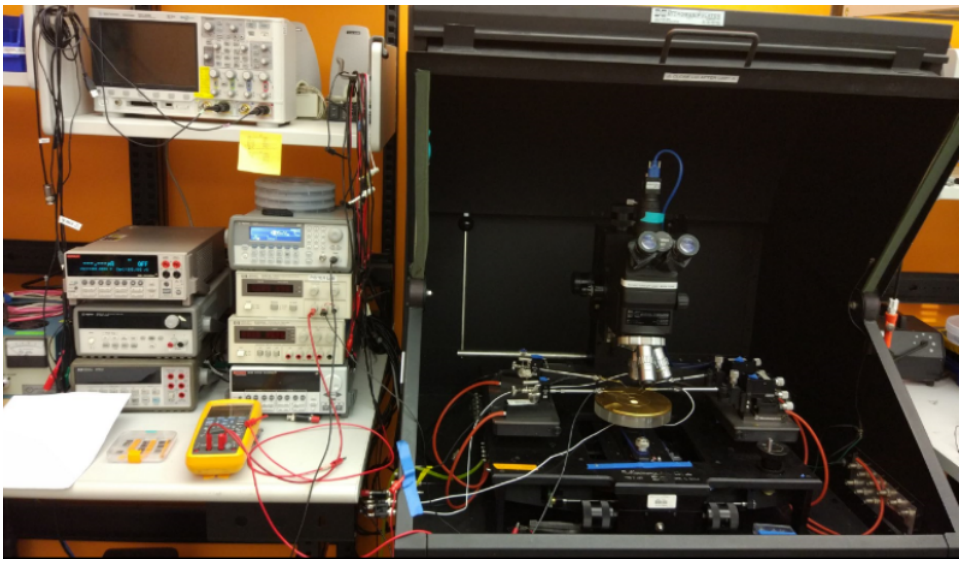


Figure A.22: Probe station complete electronic set-up

A zoomed view of the probe station plate and testing probes is shown in Figure A.23.

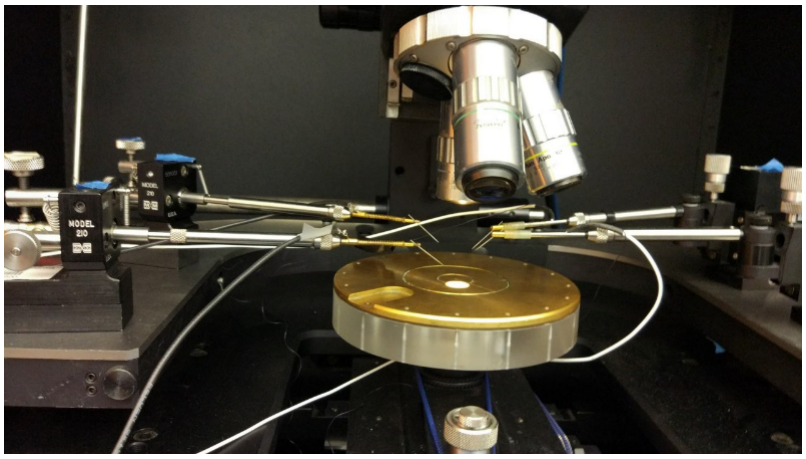


Figure A.23: Zoomed view on the probe station plate and testing probes

- Nordson DAGE 4000 Multipurpose Bondtester, shown in Figure A.24.

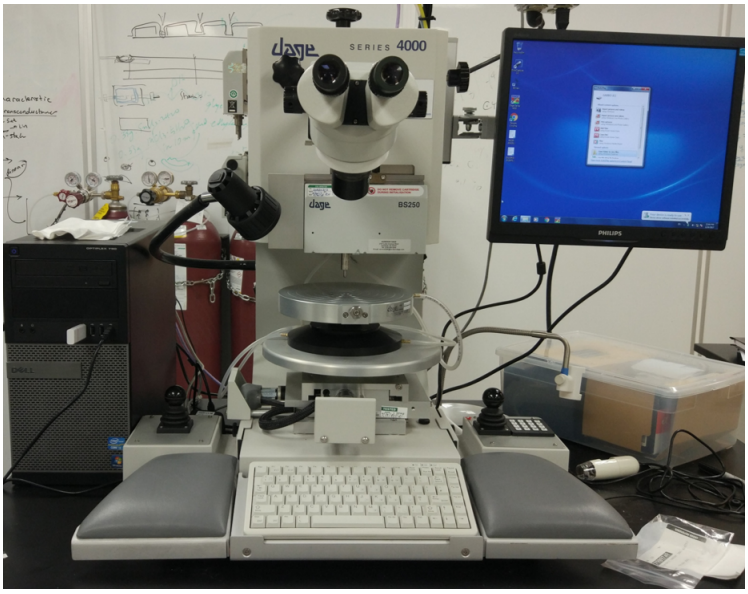


Figure A.24: Nordson DAGE bondtester