Universität Bonn
Physikalisches Institut

Low mass hybrid pixel detectors for the high luminosity LHC upgrade

Laura Gonella

Reducing material in silicon trackers is of major importance for a good overall detector performance, and poses severe challenges to the design of the tracking system. To match the low mass constraints for trackers in High Energy Physics experiments at high luminosity, dedicated technological developments are required. This dissertation presents three technologies to design low mass hybrid pixel detectors for the high luminosity upgrades of the LHC. The work targets specifically the reduction of the material from the detector services and modules, with novel powering schemes, flip chip and interconnection technologies. A serial powering scheme is prototyped, featuring a new regulator concept, a control and protection element, and AC-coupled data transmission. A modified flip chip technology is developed for thin, large area Front-End chips, and a via last Through Silicon Via process is demonstrated on existing pixel modules. These technologies, their developments, and the achievable material reduction are discussed using the upgrades of the ATLAS pixel detector as a case study.

Physikalisches Institut der
Universität Bonn
Nußallee 12
D-53115 Bonn

BONN-IR-2013-18
Oktober 2013
ISSN-0172-8741

Universität Bonn
Physikalisches Institut

Low mass hybrid pixel detectors for the high luminosity LHC upgrade

Laura Gonella

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Norbert Wermes
2. Gutachterin:	Prof. Dr. Klaus Desch

Angenommen am:	22.08.2013
Tag der Promotion:	23.09.2013

Contents

1 Introduction 1

2 Experimental environment 5
- 2.1 The Large Hadron Collider . . . 5
  - 2.1.1 LHC operation and upgrades . . . 7
- 2.2 The ATLAS experiment . . . 8
  - 2.2.1 ATLAS operation and upgrades . . . 10
- 2.3 Tracking and vertexing in HEP experiments . . . 12
  - 2.3.1 Radiation length . . . 13
  - 2.3.2 Mutiple scattering . . . 14
  - 2.3.3 Photon conversion . . . 15
  - 2.3.4 Transverse momentum resolution . . . 16
  - 2.3.5 Vertex resolution . . . 17
  - 2.3.6 Challenges of tracking in the LHC environment . . . 19
- 2.4 The ATLAS tracker: from LHC to HL-LHC . . . 22
  - 2.4.1 Pixel detector and IBL . . . 24
  - 2.4.2 SemiConductor Tracker . . . 25
  - 2.4.3 Transition Radiation Tracker . . . 25
  - 2.4.4 Tracker design for HL-LHC . . . 25

3 Pixel detectors for HEP 29
- 3.1 Hybrid pixel detectors . . . 29
  - 3.1.1 Modules . . . 30
  - 3.1.2 Electrical services . . . 33
  - 3.1.3 Mechanics . . . 35
- 3.2 Low material developments . . . 37
  - 3.2.1 Sensor . . . 37
  - 3.2.2 FE-I4 . . . 39
  - 3.2.3 Module concepts . . . 40
  - 3.2.4 Low mass, efficient powering schemes . . . 42
  - 3.2.5 Electro-mechanical structures and cooling . . . 45
- 3.3 Material budget road map . . . 46

4 Serial powering 49
- 4.1 Serial powering for hybrid pixel detectors . . . 49
  - 4.1.1 Serial powering with the FE-I4 readout chip . . . 52

4.2 Serial powering building blocks . . . 55
4.2.1 Voltage regulation . . . 55
4.2.2 SP chain protection (module bypass system) . . . 59
4.2.3 AC-coupled LVDS data transmission . . . 61
4.2.4 HV distribution . . . 62
4.3 Shunt-LDO regulator in FE-I4A . . . 63
4.3.1 Characterization . . . 64
4.4 Serially powered module operation . . . 73
4.4.1 Module performance for different powering options . . . 76
4.4.2 Multi-chip module performance for different temperatures . . . 79
4.5 Serial powering pixel stave prototype . . . 79
4.5.1 Power-up procedure . . . 82
4.5.2 Tuning of the SP stave . . . 82
4.5.3 Timing performance tests . . . 85
4.5.4 Source scan . . . 86
4.5.5 Failure mode studies . . . 86
4.6 SPP chip characterization . . . 88
4.7 Summary and outlook . . . 89

5 Flip chip of thin, large area FE chips **91**
5.1 Flip chip for hybrid pixel detectors . . . 91
5.1.1 IZM AgSn solder bump flip chip process . . . 92
5.1.2 SELEX Indium bump flip chip process . . . 93
5.1.3 Flip chip of thin FE chips . . . 94
5.2 IZM solder flip chip process for thin FE-I4 chips . . . 95
5.3 IBL modules prototype and production with 150 µm thin FE-I4 at IZM . . . 98
5.4 Summary and outlook . . . 100

6 Module concepts with Through Silicon Via technology **103**
6.1 TSV for hybrid pixel detectors . . . 103
6.1.1 Pixel module concepts with via last TSV . . . 104
6.2 IZM via last TSV processing on ATLAS readout wafers . . . 105
6.3 Results of electrical tests on demonstrator TSV modules . . . 108
6.4 Summary and outlook . . . 111

7 Conclusions **113**

Bibliography **115**

List of Figures **121**

List of Tables **125**

Acknowledgements **127**

Chapter 1

Introduction

The Standard Model (SM) of particle physics describes the fundamental structure of matter as we know it today. According to this theory, the universe is made of elementary particles, the fundamental fermions, divided into leptons and quarks. The fundamental bosons mediate the interaction between these particles. The Standard Model describes all interactions between fermions with three of the four known forces, leaving out gravity as it is too weak at the energy scale typical of other forces. These are the strong force, which acts between color charged particles via the gluon, the electromagnetic force, which acts between all charged particles by the exchange of photons, and the weak force, which acts between all fermions (and some bosons) via the W and Z bosons. Particles in the SM acquire mass by interacting with the Higgs field, a scalar background field with a non-zero vacuum expectation value. This mechanism breaks in particular the Electroweak $SU(2) \times U(1)$ symmetry group, by giving mass to the W and Z bosons and leaving the photon mass-less. The breaking requires the existence of a massive scalar boson, the Higgs boson. Figure 1.1 summarizes all ingredients of the SM.

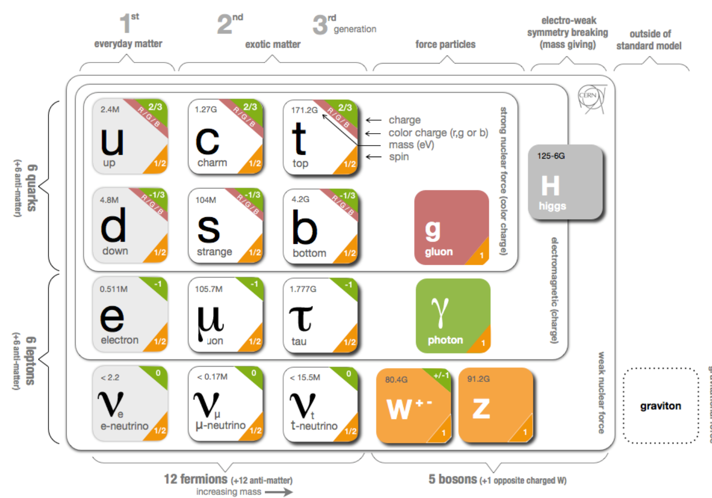


Figure 1.1: The Standard Model of particle physics.

The Standard Model leaves however some open questions, as it does not provide complete answers to, for instance, the hierarchy problem[1], the nature of dark matter, and the matter/anti-

[1] The hierarchy problem is caused by the fact that the electroweak scale is in the 100 GeV range, while the Plank scale is $\approx 10^9$ GeV.

matter asymmetry. To overcome these limitations, a number of possible extensions to the SM are formulated, such as for instance Supersymmetry and extra-dimensions, which are referred to as theories beyond the Standard Model.

The Standard Model has been successfully tested and validated to a very high precision in experiments at, among others, the LEP and Tevatron colliders, and it is currently under scrutiny at the Large Hadron Collider (LHC), where evidence for a new particle compatible with the SM Higgs boson has been recently found. This was possible thanks to the high energy and luminosity of the LHC, designed to probe physics phenomena at the TeV scale. In a 27 km long tunnel, protons will collide with a center of mass energy of $\sqrt{s} = 14\,\text{TeV}$ at a luminosity of $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$. The outcome of the collisions is studied by four experiments located along the accelerator ring, where tracking and vertex detectors play a fundamental role in probing signatures of interesting physics events. The LHC and one of its main experiments, ATLAS, are described in Chapter 2.

In High Energy Physics (HEP) experiments, trackers are commonly located around the interaction point, and are (radially) followed by the calorimeter and muon detector systems. As shown in Figure 1.2, particles from the interaction point traversing such a detector arrangement, leave a trail of electron-hole/ion pairs by ionization in the tracker, and deposit most of their energy via electromagnetic and hadronic showers in the calorimeters. From the reconstructed tracks in the tracking detector, the decay vertices and transverse momentum of charged particles can be obtained, while the energy measurement is performed by the calorimeter. The core of the tracking system is made of specialized vertex detectors to study the secondary decay vertices of heavy flavor quarks and τ-leptons.

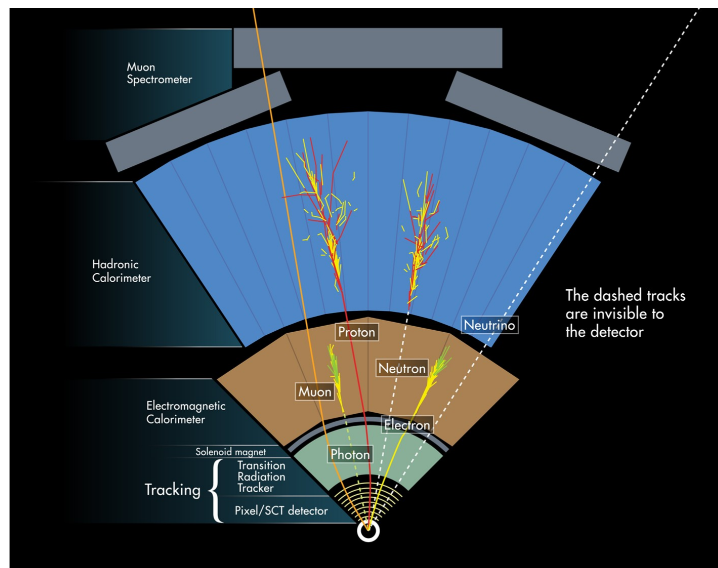


Figure 1.2: Transverse view of $1/8^{th}$ of the ATLAS detector. Particles coming out of the LHC interactions are recorded as tracks of electron-hole/ion pairs in the tracking system and as electromagnetic and hadronic showers in the calorimeters. Muons traverse the entire system and are recorded by the muon spectrometer. Neutrinos are invisible to the entire detector, and are seen as deficiencies in the total momentum balance [1].

Due to their proximity to the IP, trackers at the LHC are exposed to unprecedented rates, radiation levels and particle multiplicities. Efficient tracking and vertexing performance at the LHC thus requires radiation hard trackers, with fast readout, high granularity, all combined with low mass in order not to degrade the measurement capabilities of the overall detector. A low mass tracking detector is in fact of particular importance for both the tracker and calorimeter resolutions. Multiple scattering within the detector material deviates the trajectory of the particles, and thus degrades the tracker resolution. Bremsstrahlung and photon conversion in the tracker, affect the resolution of the electromagnetic calorimeter as many electrons and photons do not reach this detector system as such.

The LHC tracking detectors have demonstrated that these requirements can be met by a combination of silicon pixel and strips detectors at the inner radii, eventually complemented by transition radiation detectors at the outer radii. Pixel detectors are in particular used as vertex detectors. Hybrid pixel detector technology, where sensor and electronics are separate entities, mated in a flip chip process, has greatly improved the vertexing capability of HEP experiments, and is to date the only technology that can satisfy the severe requirements of operating trackers a few cm away from the LHC interaction point. The role and design of trackers for HEP experiments are discussed in Chapter 2, and hybrid pixel detectors are described in detail in Chapter 3.

With the planned High Luminosity LHC (HL-LHC) upgrade, the trackers, and in particular the pixel detectors, will have to cope with even higher rates and radiation levels, translating in more severe constraints for the detector design. Within this framework, the work of this thesis investigates technologies to build low mass hybrid pixel detectors for experiments at the HL-LHC. In particular, the R&D focuses on reducing the material of the detector modules and services, with novel flip chip and interconnection technologies, and powering schemes. The target application for these developments are the upgrades of the ATLAS pixel detector. This is used throughout the thesis as case study to illustrate the investigated technologies, their development, and gain in terms of detector material. The results of this work are however by no means restricted to this application only, and they can be applied more generally to silicon trackers for HEP experiments. The developments carried out in this thesis are described in three separate chapters. Chapter 4 presents the serial powering concept and a possible implementation for the ATLAS pixel detector at the HL-LHC. A dedicated flip chip technology for new generation hybrid pixel detector readout chips is discussed in Chapter 5. Through Silicon Via (TSV) interconnection technology is described in Chapter 6. Finally, a summary of the results and achievements, and an outlook on further development of the proposed technologies is given in Chapter 7.

Chapter 2

Experimental environment

2.1 The Large Hadron Collider

The Large Hadron Collider [2] is the world's largest particle accelerator, built to provide a rich physics program with its unprecedented energy and luminosity. The physics potential offered by the LHC ranges from precise measurements of the Standard Model parameters, to the discovery of new physical phenomena, with focus on the search for the Standard Model Higgs boson.

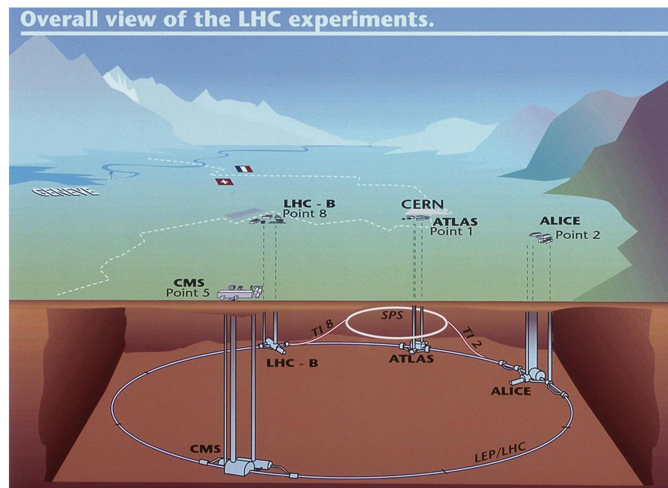


Figure 2.1: Sketch of the LHC ring and the four experiments. The beams are injected in the LHC from the Super Proton Synchrotron (SPS) at an energy of 450 GeV [1].

The LHC is a 27 km ring-shaped accelerator, located at the European Laboratory for Particle Physics (CERN) in Geneva, Switzerland (Figure 2.1). Placed in a tunnel about 100 m underground, the LHC is designed to provide proton-proton (pp) collisions at a center of mass energy of $\sqrt{s} = 14$ TeV. As shown in Figure 2.2, at the TeV energy scale, the cross-section of the interesting physics events, as for instance Higgs production, is in the order of pb, and smaller. To collect the large statistics samples required for precision measurements of these phenomena, colliders must have a high luminosity. The number of events per second generated in the LHC collisions is given by

$$\dot{N}_{event} = L \cdot \sigma_{event} \tag{2.1}$$

where σ_{event} is the cross-section for the process under study, and L is the luminosity of the accelerator. The luminosity is a measure for the number of particles colliding per second, per effective area of the colliding beams. For an intersecting storage ring collider, the luminosity is given as

$$L = \frac{n_1 \cdot n_2 \cdot k \cdot f}{4\pi \cdot \sigma} \tag{2.2}$$

where f is the revolution frequency, k the number of bunches per beam, n_i the number of particles per bunch, and σ the effective area of the beam cross-section. The LHC is designed for a peak luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This value is achieved by filling each of the two rings with 2808 bunches of 10^{11} protons, colliding every 25 ns.

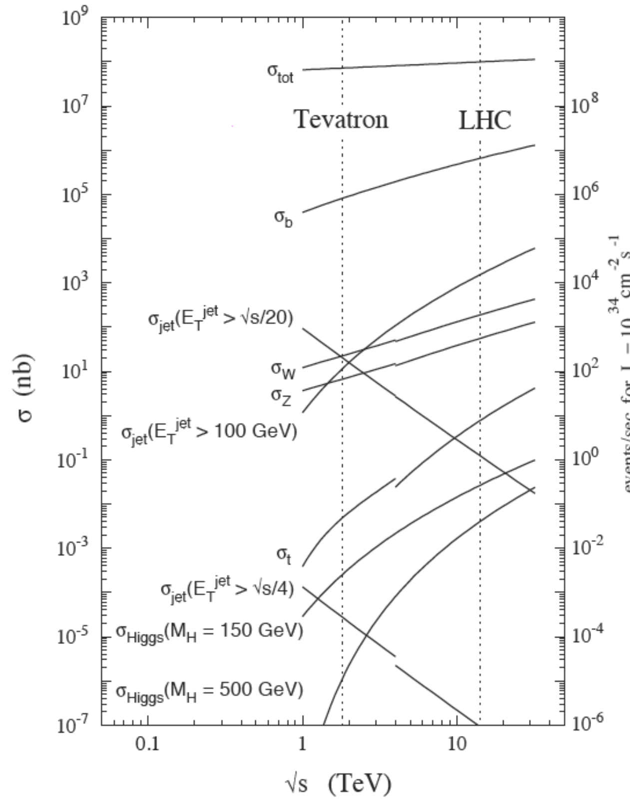


Figure 2.2: Expected hard scattering cross-sections as a function of the center of mass energy for proton-(anti)proton collisions at the LHC (Tevatron). The corresponding event rate for the LHC design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is also depicted [3].

The cross-section for inelastic, non-diffractive pp interactions at the LHC is ~80 mb. At design luminosity the LHC produces a total rate of 10^9 inelastic events/s. This represents a serious experimental difficulty as it implies that every candidate event for new physics is on average superimposed by 23 simultaneous inelastic events per bunch crossing (BX) [4]. These minimum-bias events are referred to as pile-up. The picture is further complicated by elastic and diffractive interactions, producing respectively protons and hadrons scattered at small angles with respect to the beam direction. These interactions are characterized by a low

momentum transfer, and are classified as soft, in contrast to hard interactions involving high momentum transfer, in which particles from processes of interest are scattered at large angles. Disentangling the rare, non-diffractive, high transverse momentum reactions representing interesting physics from the background is a challenge posing stringent requirements on the design of the experiments at the LHC.

To exploit the full physics potential offered by the LHC, four experiments are operated along the accelerator ring in correspondence of the collision points, as shown in Figure 2.1. ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are two general purpose experiments, aiming at a peak luminosity of $L = 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ for proton operation. They offer a rich and diverse physics program with a strong emphasis on probing the origin of the Electroweak Symmetry Breaking (EWSB) mechanism, as predicted by the Standard Model, with major focus on the Higgs boson. ATLAS and CMS also concentrate on the search for new phenomena or new particles predicted by alternative models and theories at the TeV mass scale (supersymmetry (SUSY), extra-dimensions, technicolor, compositeness, etc). The LHCb (Large Hadron Collider Beauty) experiment is made of a single arm forward spectrometer. It aims at $L = 10^{32}\,\mathrm{cm^{-2}s^{-1}}$ to study rare decays of B-hadrons, and CP violation. In addition to the proton beams, the LHC can also be filled with ion beams. Heavy ions, in particular lead nuclei, collide at the LHC with an energy of 2.8 TeV per nucleon, at a design luminosity of $10^{27}\,\mathrm{cm^{-2}s^{-1}}$. The ALICE (A Large Ion Collider Experiment) detector is a dedicated ion experiment to exploit the unprecedented opportunity offered by nucleus-nucleus collisions at the LHC to study the properties of strongly interacting matter at extreme energy density. This includes the possible phase transition to a color-deconfined state: the quark gluon plasma.

2.1.1 LHC operation and upgrades

The LHC began operation in September 2008. A few days after circulation of the first beams, a faulty electrical connection between two magnets resulted in a mechanical damage followed by helium release into the tunnel. The LHC restarted operation a year later, in late summer 2009, after reparation and consolidation of all faulty connections along the accelerator ring. Over the first two years of operation, the machine has delivered almost 6 $\mathrm{fb^{-1}}$ of data to the experiments with luminosity peaking at $3.65\times10^{33}\,\mathrm{cm^{-2}s^{-1}}$, at a center of mass energy of 7 TeV. In 2012, the center of mass energy was increased to 8 TeV, and the LHC delivered over 20 $\mathrm{fb^{-1}}$to the experiments, with a peak luminosity of $\sim8\times10^{33}\,\mathrm{cm^{-2}s^{-1}}$. Figure 2.3 shows the luminosity delivered by the LHC to the ATLAS experiment from 2010 to 2012.

In the next ten years the LHC will undergo an upgrade program in three phases to increase the peak luminosity beyond the initial design goals. The larger luminosity will offer the experiments the possibility to fully understand the EWSB mechanism, through precision measurements of the properties of the Higgs boson, such as its decay modes and couplings, as well as the unique opportunity to extend the mass reach in searches for new physics well into the multi-TeV region. Table 2.1 summarizes the operational schedule and upgrade steps of the LHC. After a first operational phase at design parameters for energy and luminosity, so-called Phase-0, upgrades of the accelerator complex will allow to increase the luminosity up to $5\times10^{34}\,\mathrm{cm^{-2}s^{-1}}$. The experiments will be able to collect large statistics samples, with a total of 3000 $\mathrm{fb^{-1}}$ delivered over ten years during the HL-LHC phase. The very high pile-up environ-

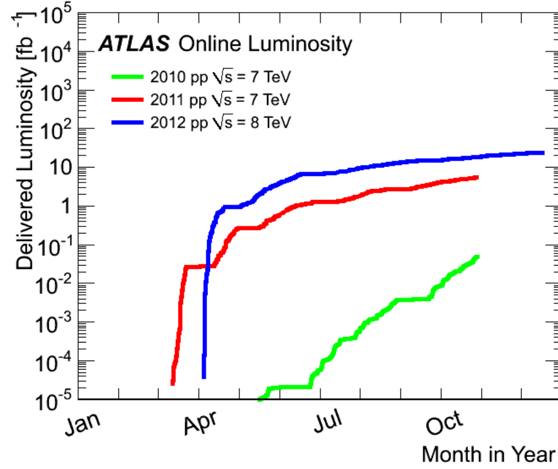


Figure 2.3: Cumulative luminosity versus day delivered to the ATLAS experiment during stable beams and for *pp* collisions, in 2010, 2011, and 2012 [5].

ment in Phase-I and Phase-II, with respectively 55 and 140 pile-up events per bunch crossing, will represent however a major experimental challenge for the detectors in order to maintain their physics performance.

Table 2.1: Planned luminosity upgrades of the LHC. The center of mass energy will reach the design value of $\sqrt{s} = 14$ TeV during Phase-0.

Upgrade	Year	Luminosity $[10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	Integrated luminosity $[\mathrm{fb}^{-1}]$	Pile-up [events/BX]
Phase-0	2015 – 2017	1	50 – 100	23
Phase-I	2019 – 2022	2.2	300 – 400	55
Phase-II (HL-LHC)	2024 – 2034	5	3000	140

2.2 The ATLAS experiment

The ATLAS experiment [4, 6] at the LHC is designed with the primary goal of probing *pp* collisions at high luminosity, providing as many physics signatures as possible, in order to maximize the potential for the discovery of new physics. The LHC represents for the experiments both an unprecedented possibility to study physics at the TeV scale, as well as an extreme experimental environment. To fulfill the requirements for precision measurements, as well as to cope with high interaction rates, radiation doses, particle multiplicities and energies, experiments have to be carefully designed. Large acceptance in pseudorapidity with almost full azimuthal angular coverage is required. High detector granularity is needed to handle the

particle fluxes and to reduce the influence of overlapping events. Detector elements, such as electronics and sensors, should be fast and radiation-hard to survive in the LHC environment during the entire experiment lifetime. Different detector sub-systems should then assure

- good charged particle momentum resolution and track reconstruction efficiency, as well as electron identification, and measurement of secondary vertices for identification of τ-leptons and heavy quarks;
- electromagnetic calorimetry for electron and photon identification and energy measurement, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements;
- good muon identification and momentum resolution, in particular for high-p_T muons;
- highly efficient triggering for particles at low-p_T thresholds, with sufficient background rejection, to achieve an acceptable trigger rate for most physics processes of interest.

These specifications have been followed in the design of the ATLAS experiment, and resulted in the layout shown in Figure 2.4. The ATLAS experiment is a 4π detector, forward-backward symmetric with respect to the interaction point. The detector has a height of 25 m and it is 44 m long, with a total weight of 7000 Tons. The central detector region is called barrel, and the two sides are the end-caps. In the ATLAS coordinate system the z direction is set along the beam axis, while x and y coordinates are in the transverse plane. Polar coordinates are used in the transverse plane, where the R coordinate describes the radial position from the beam axis, and the ϕ coordinate describes the azimuthal angle. The pseudorapidity describes the angular position with respect to the beam axis, and it is defined as $\eta = -\ln\tan(\theta/2)$, where θ is the angle with the beam axis.

The detector is made of different magnet and detector sub-systems. The magnet configuration comprises an inner thin superconducting solenoid surrounding the inner-detector cavity, and three outer large superconducting toroids (one barrel and two end-caps) consisting of independent coils arranged with an eight-fold symmetry outside the calorimeters. This fundamental choice has driven the design of the rest of the detector. The detector sub-systems can be seen as concentric cylinders, centered around the beam pipe. The innermost sub-system is the Inner Detector (ID). Pattern recognition, momentum and vertex measurements, and electron identification are achieved with a combination of three tracking detectors immersed in a solenoidal magnetic field of 2 T. Discrete high-resolution semiconductor pixel and strip detectors are used in the inner part of the tracking volume, and straw-tube tracking detectors with transition radiation detection in its outer part. The ID is described in more detail in Section 2.4. Moving radially outwards, the next sub-detector system is the calorimeter. The overall calorimeter system provides excellent performance in terms of energy and position resolution, using a combination of a liquid-argon electromagnetic sampling calorimeter, followed by a scintillator-tile hadronic calorimeter. The calorimeter is surrounded by the muon spectrometer. The air-core toroid system, with a long barrel and two inserted end-cap magnets, generates strong bending power in a large volume within a light and open structure to minimize multiple scattering effects. Excellent muon momentum resolution is achieved with three layers of high precision

tracking chambers. The muon instrumentation includes, as a key component, trigger chambers with timing resolution of the order of 1.5 −4 ns. A 3-level trigger system is used to convert the 1 GHz interaction rate at design luminosity, to a final data taking rate of approximately 200 Hz. This system provides an overall rejection factor of 5×10^6 against minimum-bias processes, while maintaining maximum efficiency for new physics.

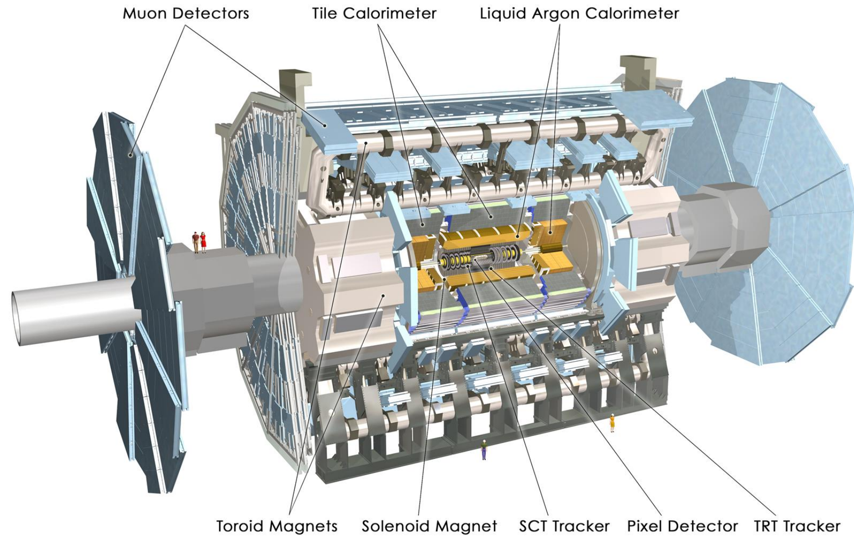


Figure 2.4: Cut away of the overall ATLAS experiment showing the various sub-detectors and the two magnet systems [7].

2.2.1 ATLAS operation and upgrades

The ATLAS detector has been assembled over a period of three years. Commissioning started in the second half of 2007, and data taking in 2009 with the first stable beams from the LHC. The detector performed superbly during the first three years of operation, proving that the chosen design allows to investigate a wide spectrum of physical phenomena at the TeV energy scale, in the harsh LHC environment. With the integrated luminosity recorded by the experiment, the ATLAS collaboration discovered a new boson whose properties confirm the hypothesis of a SM Higgs boson [8]. Searches for the SM Higgs boson have been performed using 5.8 −5.9 fb^{-1} of pp collision data recorded at the beginning of 2012 at a center of mass energy of 8 TeV, and 4.6 −4.8 fb^{-1} recorded in 2011 at a center of mass energy of 7 TeV. Results of the combined analysis are shown in Figure 2.5. The SM Higgs boson is excluded at 95 % confidence level (CL) in the mass range 111 −559 GeV, except for the narrow region 122 −131 GeV. In this region, an excess of events with a significance of 5.9 σ is observed, corresponding to a probability p_0 that the experimental observations are consistent with a background-only hypothesis of 10^{-9}. This proves conclusive evidence for the discovery of the new particle at a mass of ∼126 GeV. The signal strength μ, defined as the ratio between the observed Higgs cross-section and the SM Higgs cross-section, has the value 1.4 ± 0.3 at the fitted mass, which is consistent with the SM Higgs boson hypothesis ($\mu = 1$).

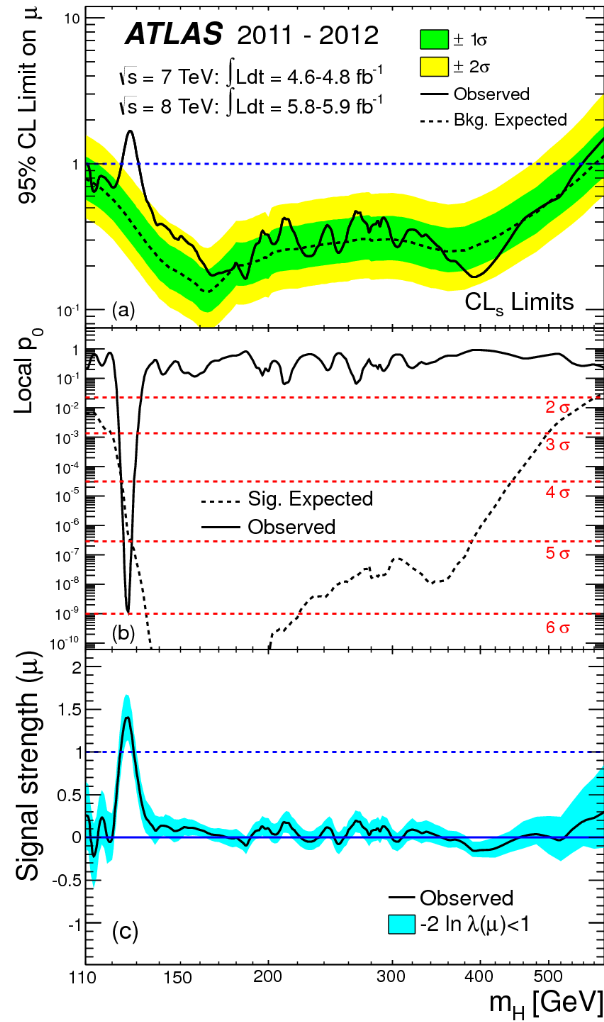


Figure 2.5: Combined search results of several Higgs decay channels: (a) The observed (solid) 95 % CL upper limit on the signal strength as a function of the Higgs mass, m_H, and the expectation (dashed) under the background-only hypothesis. The dark and light shaded bands show the $\pm 1\,\sigma$ and $\pm 2\,\sigma$ uncertainties on the background-only expectation. (b) The observed (solid) local p_0 as a function of m_H and the expectation (dashed) for a SM Higgs boson signal hypothesis ($\mu = 1$) at the given mass. (c) The best-fit signal strength $\hat{\mu}$ as a function of m_H. The band indicates the approximate 68 % CL interval around the fitted value. [8]

The ATLAS collaboration plans a series of upgrades, following the LHC schedule. The detector capability has to be consolidated and improved to maintain the capability of the experiment for precision measurements, and its potential for the discovery of new physics, while meeting the new challenges of operating in a high luminosity environment.

No major intervention to the detector system is planned during LS1, as during Phase-0 the machine will operate at design parameters. Only the pixel detector will undergo the first upgrade project, in view of the Phase-I upgrade, where the LHC will work at a luminosity higher than $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. As explained in Section 2.4.1, in order to preserve tracking robustness and efficiency at higher than design luminosity, a new pixel layer, the Insertable B-Layer (IBL) [9], will be added at a smaller radius inside the present detector. Insertion of the IBL is planned during the first LHC shutdown (2013 - 2014), and operation will start during Phase-0.

The Phase-I upgrade [10] will concern mostly an improvement of the trigger system to cope with luminosities higher than the LHC nominal value. In particular the LV1 trigger upgrade will allow to maintain low-p_T thresholds for isolated leptons, as required by precision measurements of the Higgs boson couplings in the low mass region, as well as by searches for supersymmetric particles in a large region of the SUSY parameter space. Improvement of the high-level trigger will lead to more efficient identification of events with isolated τ-leptons and B-hadrons, improving the selection of Higgs boson decays and the sensitivity to many other physics channels. The detector sub-systems involved in this upgrade will be the calorimeters and the muon spectrometer.

Phase-II will see a major upgrade of the detector systems [11]. In particular, the increased luminosity requires a new ID able to cope with the higher rates, pile-up, and radiation levels. An all new silicon tracker is studied, discussed in more detail in Section 2.4.4. The new tracker readout will include the implementation of a track trigger to improve the ATLAS trigger capabilities. New readout systems are also required to maintain the performance of the calorimeters and of the muon spectrometer. Finally, a new trigger architecture is implemented to exploit the upgrades of the detector readout systems, improving the event selection.

2.3 Tracking and vertexing in HEP experiments

Tracking is a fundamental tool for experiments at the LHC, where tracking detectors are responsible for momentum and vertex measurements. The measurement of the transverse momentum of a charged particle, as well as the reconstruction of the primary vertices of a collision are needed to separate interesting physics events from the pile-up events. In addition to this, the tracker plays a fundamental role in the detection of secondary (detached) vertices to identify jets from hadrons containing heavy quarks, like the b-quark, as well as τ-leptons. Identification of b-jets and τ-leptons is particularly important as they are a distinguishing signature in many physics processes of interest, and dedicated tracking detectors, called vertex detectors, are used for this task. The decay of B-hadrons and τ-leptons is mediated by the weak interaction, thus they do not decay right after production, as shown in Figure 2.10. For example the B^0 meson has a life-time $\tau_{dec} = 1530 \pm 9\,\mathrm{ns}$. The τ-lepton has $\tau_{dec} = 290.6 \pm 1.0\,\mathrm{ns}$. Being produced at high relativistic energy, their decay length l_{dec} is increased by the Lorentz boost to $l_{dec} = \beta\gamma c\tau_{dec}$, so that these particles decay inside the beam pipe, and can be detected by

measuring the position of their decay (secondary) vertex. For instance, a τ-lepton decaying at rest has $l_{dec} \approx 87\,\mu m$. A τ-lepton with a momentum of 10 GeV/c, would travel $l_{dec} \approx 490\,\mu m$ before decaying.

Trackers are made of position sensitive detectors immersed in a magnetic field. As such they can be seen as a magnetic spectrometer, measuring points along the trajectory of the particle. Starting from the measured points, pattern recognition algorithms are used to reconstruct tracks of charged particles. The reconstructed tracks are then combined to find the primary and secondary vertices in the event, and to determine the p_T of the particles coming out of the collision. The required resolution on vertex and momentum measurements translates into a set of specifications for the design of tracking detectors. Accurate measurement of the track coordinates imposes constraints on segmentation, position, and number of detector layers, as well as on the detector mass. Effects dependent on the amount of material traversed by the particles, characterized using the radiation length, can in fact reduce the resolution of the tracker, as well as of the calorimeters. These effects are mainly multiple scattering and photon conversion.

As the focus of the work presented in this thesis is to develop technologies to build low mass pixel detectors, radiation length, multiple scattering, and photon conversion are described in Section 2.3.1, Section 2.3.2, and Section 2.3.3. Vertex and momentum resolution are then discussed in Section 2.3.4 and Section 2.3.5, to derive the requirements on detector design. Further constraints on detector design added by the need of operating in the LHC environment are illustrated in Section 2.3.6. Being the detector closer to the interaction point, the tracker is the most affected by the high pile-up and radiation levels, and dedicated developments are needed, from sensor materials to mechanics, to implement the required geometry.

2.3.1 Radiation length

High energy electrons loose energy in matter primarily via bremsstrahlung. This process describes the radiation of a photon when an electron is decelerated in the electromagnetic field of a nucleus, according to the process shown in Figure 2.6a. The energy loss per unit length due to bremsstrahlung is proportional to the energy of the incident particle according to

$$\frac{dE}{dx} = -\frac{E}{X_0} \tag{2.3}$$

where X_0 is the radiation length. Integrating (2.3) gives

$$E(x) = E_0 \cdot e^{-x/X_0}. \tag{2.4}$$

The radiation length X_0 thus corresponds to the mean distance over which a high energy electron looses $1 - 1/e = 63\,\%$ of its energy by bremsstrahlung.

X_0 is a material property which can be approximated as [12]

$$\rho X_0 = \frac{716.408 gcm^{-2} \cdot A}{Z(Z+1) \cdot \log \frac{287}{\sqrt{Z}}} \tag{2.5}$$

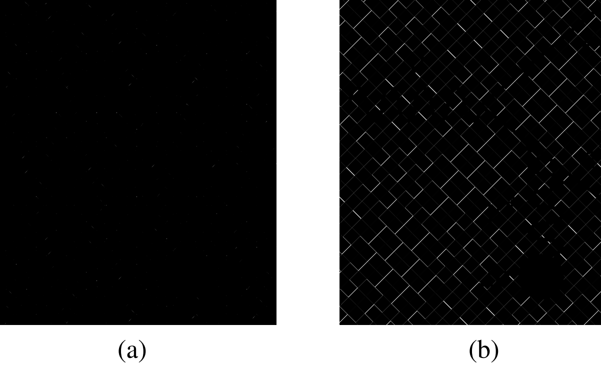
where ρ is the density, A is the mass number, and Z is the atomic number of the material traversed by the electron. The Z^2 dependence describes the coherent scattering with the nucleus,

while the dependence on Z describes the scattering with the shell electrons. In mixtures and compounds the radiation length may be approximated by

$$\frac{1}{X_0} = \sum \frac{w_j}{X_{0j}} \tag{2.6}$$

where w_j and X_{0j} are the fraction by weight and the radiation length of the jth element.

The radiation length is a quantity used to characterize electromagnetic processes that take place in the Coulomb field of the nucleus. They are, in addition to bremsstrahlung, also multiple scattering, and photon conversion. It is also commonly used to quantify the detector mass, also referred in the following as material budget. The detector material is then expressed as a thickness is units of X_0.



(a) (b)

Figure 2.6: Feynman diagrams for (a) bremsstrahlung and (b) photon conversion.

2.3.2 Mutiple scattering

When a charged particle traverses a medium, it is scattered many times at small angles due to interactions with the Coulomb field of the nuclei, an effect called multiple scattering. The particle exits the material at an angle with respect to the incident direction. The distribution of the scattering angle can be described using the Molière theory: it is roughly Gaussian at small angles, but at larger angles it behaves like Rutherford scattering, with larger tails than a Gaussian distribution. Using the Gaussian approximation, the projected distribution of the scattering angle on any plane containing the initial direction $f(\theta_{plane})$ is given as

$$f(\theta_{plane})d\theta_{plane} = \frac{1}{\sqrt{2\pi}\theta_0} \exp\left(\frac{\theta^2_{plane}}{2\theta_0^2}\right) d\theta_{plane}. \tag{2.7}$$

As shown in Figure 2.7, $f(\theta_{plane})$ is centered around zero. The standard deviation θ_0 is the width of the distribution and is given by the so-called "Highland formula" [13, 14]

$$\theta_0 = \frac{13.6\text{MeV}}{\beta c p} z \sqrt{x/X_0}\,[1 + 0.038 \ln(x/X_0)] \approx \frac{13.6\text{MeV}}{\beta p} \sqrt{x/X_0} \tag{2.8}$$

where p, βc and z are the momentum, velocity, and charge number of the incident particle, and x/X_0 is the thickness of the scattering medium in radiation lengths.

The random deflection of the particle trajectory caused by multiple scattering can significantly reduce the accuracy of the track position measurement, and thus degrade the tracker resolution. According to (2.8), multiple scattering is particularly severe for low momentum tracks, and for trackers with high material budget, in which cases it can be the dominant source of error in measuring track coordinates.

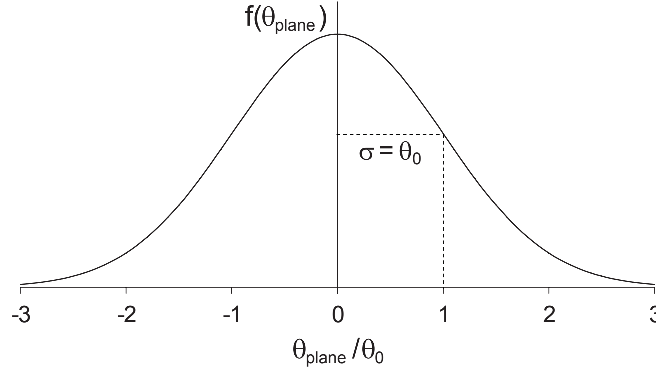


Figure 2.7: Angular distribution of the multiple scattering angle projected on a plane containing the initial direction of the incident particle.

2.3.3 Photon conversion

A high energy photon can covert in an electron-positron (e^-e^+) pair in the Coulomb field of a charge, in a process similar to bremsstrahlung as shown in Figure 2.6b, called photon conversion or pair production. The energy of the photon for pair production must be higher than twice the electron mass plus the recoil energy of the charge. With the exception of the light elements, pair production takes place in the electric field of the nucleus. In this case, the recoil energy can be neglected and the energy threshold for pair production is

$$E_\gamma = 2m_e c^2 \approx 1\,\text{MeV}. \tag{2.9}$$

The photons which are relevant for physics measurements at the LHC have energies in excess of 1 GeV. At these energies, the cross-section for the conversion process is almost completely independent of the energy of the incident photon, and can be expressed in terms of the radiation length as

$$\sigma_{e^-e^+} = \frac{7}{9}\frac{1}{X_0}\frac{A}{N_A \rho} \tag{2.10}$$

from which the absorption length is given by

$$\lambda_{e^-e^+} = \frac{9}{7} X_0. \tag{2.11}$$

For the photon conversion process, the radiation length is thus $\frac{7}{9}$ of the mean free path for pair production. This means that photons with energies above 1 GeV can convert into electron hole pair when interacting with the tracker material. Detection of such energetic photons is of fundamental importance to reconstruct events with high-mass di-photon final states, such as for instance the Higgs boson. Massive trackers thus degrade the measurement of photon properties in subsequent detectors, i.e. the electromagnetic calorimeter, as the photons convert before getting there. Reconstruction of photons measurement via the tracks of the electron-hole pairs is complicated as the momentum of the photon is not shared equally between the electron and the position, meaning that one is produced at low energy. If this energy falls below the threshold required to produce a reconstructable track, the converted photon has only one track and it is difficult to distinguish from a single electron or positron [15].

2.3.4 Transverse momentum resolution

The resolution on momentum measurement is discussed here considering a position sensitive detector made of $N + 1$ layers, corresponding to as many measuring points, equally spaced at radii $r_0, r_2, ..., r_N$ from the interaction point [16]. The detector length is $L = r_N - r_0$. It is immersed in a magnetic field B and its elementary elements have an intrinsic spatial resolution σ.

The transverse momentum resolution is given by the quadratic sum of two terms as shown in Figure 2.8

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{\left(\frac{\sigma_{p_T}}{p_T}\right)^2_{point} + \left(\frac{\sigma_{p_T}}{p_T}\right)^2_{MS}}. \tag{2.12}$$

The first term is the resolution to which a track can be measured neglecting multiple scattering, here called point resolution. As shown in [17], in this case the error on the transverse momentum p_T is given by

$$\left(\frac{\sigma_{p_T}}{p_T}\right)_{point} = p_T \cdot \frac{\sigma}{0.3BL^2} \cdot \sqrt{\frac{720N^3}{(N-1)(N+1)(N+2)(N+3)}}. \tag{2.13}$$

According to (2.13), higher resolution is achieved for low-p_T particles. The strong dependence on the spectrometer length favors large detectors to achieve good momentum resolution. Large magnetic fields are also needed to increase the resolution. The dependence on the number of detector layers is on the other hand weak. However a high number of measurement points is important for the robustness of the pattern recognition.

The second term in (2.12) expresses the contribution due to multiple scattering in the detector material, and it is given by

$$\left(\frac{\sigma_{p_T}}{p_T}\right)_{MS} = \frac{1}{0.3B}\frac{0.0136}{\beta}\sqrt{\frac{C_N}{LX_0}} \tag{2.14}$$

where C_N is a coefficient dependent on the number of layers in the detector, equal to 1.3 within a 10 % accuracy [17].

From (2.14), it can be inferred that when multiple scattering dominates the relative momentum resolution does not depend on the momentum and has a weak dependence on the spectrometer length. To achieve a good transverse momentum resolution, the contribution of the multiple scattering term should be lower than the point resolution, which calls for low detector thickness.

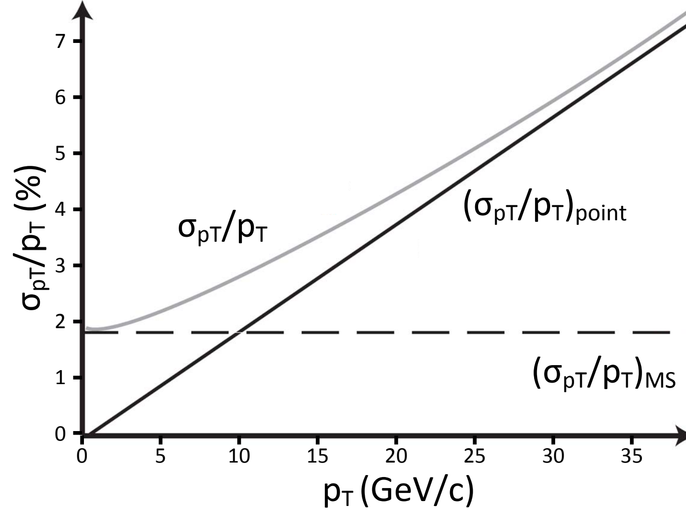


Figure 2.8: The total transverse momentum resolution is given by the quadratic sum of terms. The point resolution increases linearly with the momentum, while multiple scattering within the material of the detector introduces a term which does not depend on the momentum of the particle.

2.3.5 Vertex resolution

As discussed in [18], the vertex resolution can be derived considering a one-dimensional detector arrangement as shown in Figure 2.9. Two detector planes, with resolution σ_1 and σ_2, are placed at a distance r_1 and r_2 respectively from the interaction point, with $r_1 < r_2$. The beam pipe radius is r_0. Assuming first $\sigma_2 = 0$ and $\sigma_1 \neq 0$ (i.e. the position of the track as seen in the second layer is fixed), the resolution σ_{vtx} in the position of the reconstructed vertex is given by

$$\sigma_{vtx1} = \frac{r_2}{r_2 - r_1}\sigma_1 \tag{2.15}$$

for perpendicular incidence. Similarly considering $\sigma_1 = 0$ and $\sigma_2 \neq 0$

$$\sigma_{vtx2} = \frac{r_1}{r_2 - r_1}\sigma_2. \tag{2.16}$$

Taking into account the multiple scattering with the beam pipe material, the spatial resolution of the two layers can be expressed as

$$\sigma_1 = \sigma \oplus (r_1 - r_0)\theta_0 \tag{2.17}$$

$$\sigma_2 = \sigma \oplus (r_2 - r_0)\theta_0 \tag{2.18}$$

where $\oplus$ indicates the sum in quadrature. The same intrinsic spatial resolution σ is assumed for both layers, and multiple scattering in the first detector layer, degrading the resolution of the second, is neglected.

The total vertex resolution is the quadratic sum of σ_{vtx1} and σ_{vtx1}, and the correlation between the multiple scattering terms, which gives

$$\sigma_{vtx} = \sqrt{\left(\frac{r_1}{r_2 - r_1} + 1\right)^2 \sigma^2 + (2r_1 - r_0)^2 \, (13.6\,\mathrm{MeV})^2 \frac{x}{X_0} \frac{1}{p^2}}. \tag{2.19}$$

Again, the last term (see (2.8), with $\beta = 1$) describes the contribution from multiple scattering, which dominates for low momentum particles and for high detector thickness. For high momentum this term is less significant, and the terms arising from the position resolution σ_1 and σ_2 dominate the achievable vertex resolution. A small σ_{vtx} can be reached with

- fine detector segmentation, i.e. small σ;
- low material for the beam pipe and detector layers (low x/X_0);
- large lever arm (large $r_2 - r_1$);
- an innermost layer as close as possible to the beam pipe (small r_1, small $(2r_1 - r_0)$, and large lever arm).

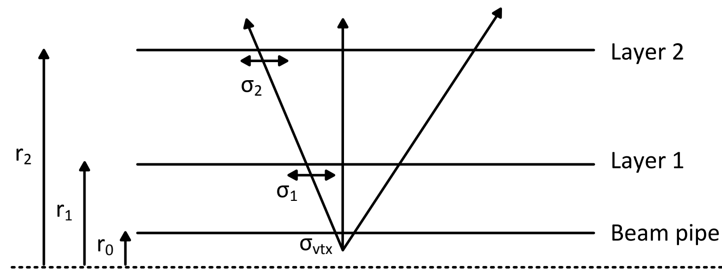


Figure 2.9: Sketch of a one-dimensional vertex detector arrangement, made of two planes with resolutions σ_1 and σ_2, placed at a distance r_1 and r_2, respectively, from the interaction point. The beam pipe radius is r_0.

Alternative to the vertex position, the impact parameter b can be considered for the reconstruction of the vertex. The impact parameter is the distance of closest approach of a track to the primary vertex of the interaction in the plane perpendicular to the beam as shown in Figure 2.10. The error on b determines whether a track can be distinguished from the primary vertex. Assuming $r_1 = r_0$, where r_0 is the beam pipe radius, and a detector resolution higher than the error due to multiple scattering, then the error on the impact parameter is

$$\Delta b = \theta_0 r_0 \tag{2.20}$$

where θ_0 is defined as in (2.8), demanding again a small radius for the innermost detector layer, and low material for the beam pipe and the detector layers.

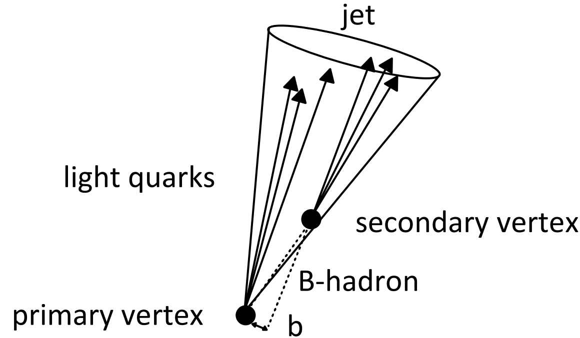


Figure 2.10: Schematic picture of the determination of the impact parameter b of a track generating from a secondary vertex.

2.3.6 Challenges of tracking in the LHC environment

As previously discussed, vertex and momentum measurements in HEP experiments, require large detectors, with fine segmentation, a high number of layers, of which the first as close as possible to the interaction point, immersed in high magnetic fields, with as low a thickness as possible. The harsh environment in which the tracker operates pushes most of these requirements to the technological limit and adds further challenges to the detector design.

Due to pile-up, every 25 ns at the LHC, approximately 1000 particles come out of on average 23 simultaneous vertices, spaced of approximately 1 cm, as the luminous area of the LHC has a Gaussian sigma along the beam direction of ≈8 cm. Two different types of pile-up can be distinguished: in-time and out-of-time pile-up. In-time pile-up is due to additional pp interactions in the same BX of the event of interest. Out-of-time pile-up is instead due to detector signals from collisions which happened one BX before the event of interest, but are reconstructed one BX later due to the long integration time of some detectors. Figure 2.11a shows the luminosity-weighted distribution of the mean number of interactions per crossing, μ, for the LHC 2012 run. It corresponds to the mean of the poisson distribution on the number of interactions per crossing, for each bunch. It is calculated from the instantaneous per bunch luminosity as

$$\mu = \frac{L}{n_{bunch}} \cdot \frac{\sigma_{inel}}{f} \tag{2.21}$$

where L is the instantaneous luminosity, n_{bunch} is the number of colliding bunches, σ_{inel} is the inelastic cross-section (73.0 mb for 8 TeV collisions), f is the LHC revolution frequency. The mean value μ is 20.7, but as shown in Figure 2.11b, the maximum number of events per beam crossing can reach higher values between 30 and 40. This means that most of the events

are accompanied by additional, unwanted interactions. This results in a large number of tracks per BX, that needs to be resolved by the detector, making reconstruction of primary vertices and identification of those associated with a hard scattering event particularly challenging. The challenge of resolving tracks and vertices at the LHC can be understood by looking at Figure 2.11c where a $Z \to \mu\mu$ candidate event is observed in a reaction with 25 reconstructed primary vertices.

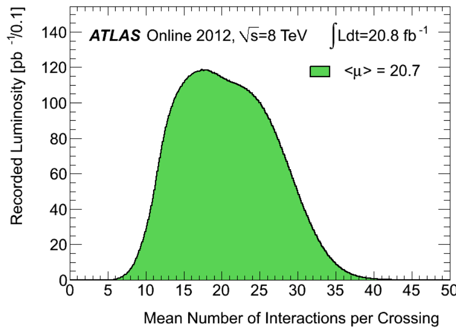
Efficient pattern recognition in such a high pile-up environment requires an occupancy lower than 1 %, where the occupancy is the average number of hits per BX, per elementary detector element. Keeping the occupancy low, reduces the probability that two tracks overlap in the same elementary detector element, and thus the number of ambiguities is small. As shown in [17], the track density from pile-up events, for each bunch crossing, on a detector layer placed at a radius r and at $\eta = 0$ is $\sim 40 \times 1/r^2$, with r in cm. An occupancy < 1 % thus requires detectors with elementary elements smaller than $2.5 \cdot 10^{-4} \times r^2$. Moreover, to minimize the out-of-time pile-up, fast detectors are required, able to process the events in less than 25 ns. Finally, due to the high rates, the tracker is the detector sub-system exposed to the highest radiation levels. As summarized in Table 2.2, during its lifetime the ATLAS ID experiences fluences in a range from 10^{15} $\mathrm{n}_{eq}\mathrm{cm}^{-2}$[1], at the innermost pixel layer, to 10^{12} $\mathrm{n}_{eq}\mathrm{cm}^{-2}$ at the TRT outer radius, and a Total Ionizing Dose (TID) up to ~80 Mrad for the detector closest to the beam pipe.

Table 2.2: The 1 MeV n_{eq} fluence and TID in key areas of the ATLAS ID after 500 fb^{-1} of data (estimated to be approximately seven years of operation) [4].

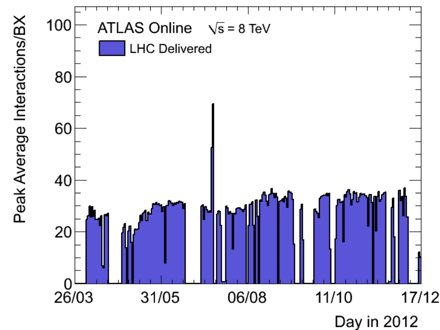
System	Fluence [10^{14} $\mathrm{n}_{eq}\mathrm{cm}^{-2}$]	TID [Mrad]
Pixel layer 0	13.5	79
Strip layer 1	0.8	3.8
Strip disk 9	0.6	2.3
TRT outer radius	0.25	0.35

The required granularity, speed and radiation hardness can be achieved with pixel detectors, which are however too expensive to cover large areas. Pixel detectors are thus used only for the innermost tracking layers. As the requirements in terms of rate and radiation decrease with the distance from the interaction point, pixel detectors are complemented at larger radii by less expensive detector types with lower granularity.

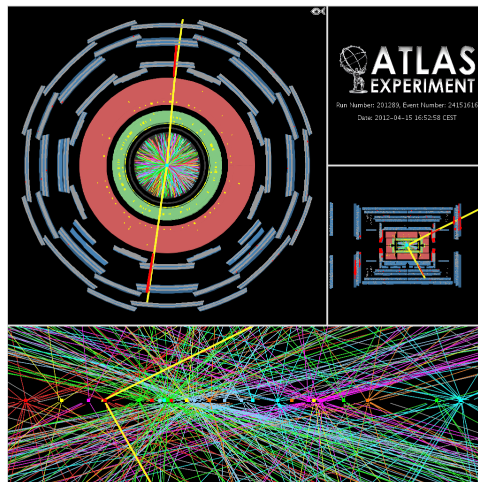
[1] The 1 MeV equivalent neutron fluence, n_{eq}, is the fluence of 1 MeV neutrons producing the same damage in a detector material as induced by an arbitrary particle fluence with a specific energy distribution.



(a) Luminosity-weighted distribution of the mean number of interactions per crossing for 2012. The integrated luminosity and the mean μ values are given in the figure.



(b) Maximum mean number of events per beam crossing versus day. Shown is the average value for all bunch crossings in a lumi-block. Only the maximum value during stable beam periods is shown.



(c) A candidate Z boson event in the di-muon decay channel with 25 reconstructed vertices, demonstrating the high pile-up environment in the LHC 2012 run. The top left picture shows the event in the $R-\phi$ plane of the detector, whilst on the right a cross-section of the detector along z is shown. The blue tiles are the muon chamber, the pink and green section are respectively the hadronic and electromagnetic calorimeters.

Figure 2.11: Number of interactions per crossing for the 2012 LHC run, and a candidate Z boson event with high pile-up [5].

2.4 The ATLAS tracker: from LHC to HL-LHC

The ATLAS tracker or Inner Detector [4, 6] is designed to provide robust and hermetic pattern recognition and precise track reconstruction. Excellent momentum resolution and both primary and secondary vertex measurements are achieved for a nominal p_T threshold of 0.5 GeV, over the pseudorapidity range $|\eta| \leq 2.5$. It also provides electron identification over $|\eta| \leq 2.0$ and for a range of energies between 0.5 GeV and 150 GeV.

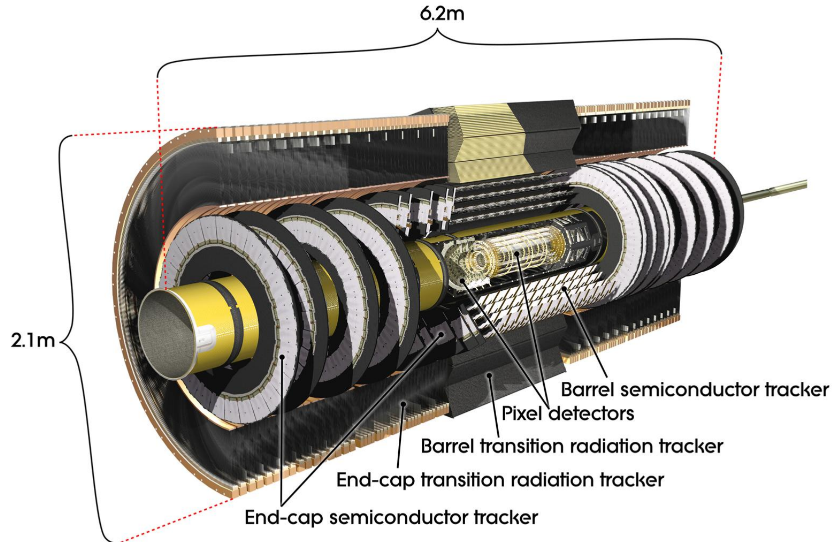
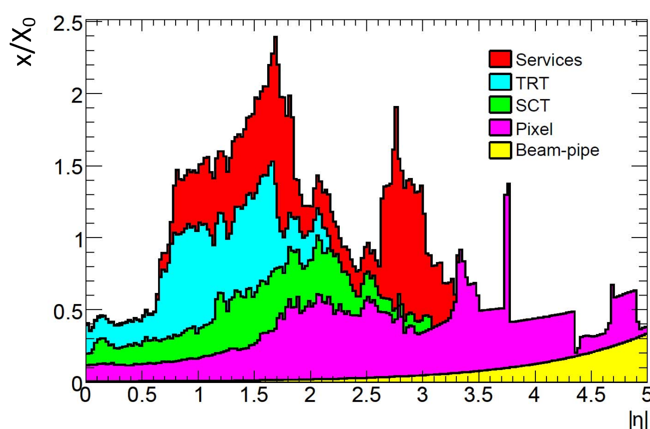


Figure 2.12: Layout of the ATLAS inner detector, showing the barrel and end-cap regions of the three sub-detector systems: pixels, strips, and TRT [7].

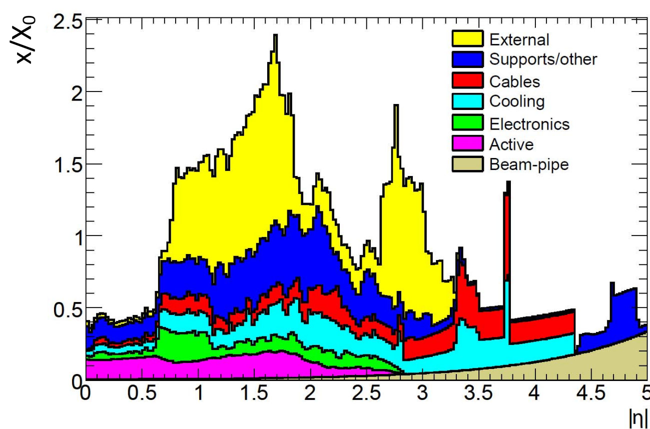
The ID layout is shown in Figure 2.12. It is contained in a cylindrical envelop with a length of ±3512 mm and a radius of 1150 mm, within a solenoidal magnetic field of 2 T. The combination of pixel and strip technology in the innermost part, with straw tube Transition Radiation Tracker (TRT) at the outer radii yields a very robust pattern recognition and high precision in both R - ϕ and z coordinates. In the barrel region the detectors are arranged on concentric cylinders around the beam pipe, while in the end-caps they are located on disks perpendicular to the beam pipe. The detectors have a modular design consisting of sensing elements, electronics readout circuitry, electrical services, and mechanics, i.e. support structure and cooling.

The innermost tracking elements, pixels and strip detectors are made using planar silicon technologies, which allow to design highly segmented detectors to cover large sensing areas. The high radiation hardness required (Table 2.2), in particular for the pixel sensors, can be achieved using oxygen enriched silicon substrates. Readout electronics is fabricated using commercial CMOS technologies, which provide the required speed, bandwidth, and radiation hardness. To maintain an adequate noise performance after radiation damage, as well as to remove the heat generated by the electronics and the leakage current of the sensor, the silicon detectors must be kept at low temperature. The mechanical structure is designed to provide high dimensional stability, in particular for the discrete tracking elements where alignment is critical for precise tracking.

The required high detector granularity with the associated electronics, services, cooling, and good mechanical stability translate in an overall ID weight and material budget much larger than those of previous tracking detectors. Figure 2.13 show the material in units of X_0 traversed by a straight track as a function of $|\eta|$ at the exit of the ID envelop. Services from the strips and TRT detectors, as well as material from the support structure dominate the material budget distribution at the interface of the barrel and end-cap regions. The pixel services leave the detector along the beam pipe and contribute to the material budget at $|\eta| \geq 2.7$. The calorimeter resolution is affected by the large fraction of ID services and structural material external to the ID active volume. As a consequence, many electrons loose most of their energy through bremsstrahlung, and approximately 40 % of the photons convert into electron-positron pairs before reaching the electromagnetic calorimeter.



(a)



(b)

Figure 2.13: Material distribution expressed in units of X_0 as a function of η for the ATLAS ID. The breakdown in (a) indicates the contributions of external services and of individual sub-detectors, including services in their active volume, and in (b) the contributions of different ID components, independent of the sub-detector (adapted from [4]).

2.4.1 Pixel detector and IBL

The pixel detector (Figure 2.14) is used as the first detector around the beam pipe because of its high granularity and radiation hardness. It has approximately 80×10^6 pixels with a size of 50 µm in R - ϕ and 400 µm in z. Three barrel layers are located at radii of 5.05, 8.85, and 12.25 cm. Two sets of three disks, with a mean radius of 17 cm, are placed at a distance of 400.5, 580, and 650 cm from the interaction point. The intrinsic accuracies in the barrel are 10 µm in R - ϕ and 115 µm in z, while in the disks they are 10 µm in R - ϕ and 115 µm in R. The pixel detector provides three points per track over the full acceptance, granting reconstruction of tracks segments independently of the outer detectors. It is designed to provide efficient track reconstruction and vertex measurements close to the interaction point, and it is fundamental to determine the impact parameter resolution of the ID. In particular, the secondary vertex measurement is enhanced with the innermost pixel layer, referred to as B-layer[2], which is placed as close as technically possible to the beam pipe.

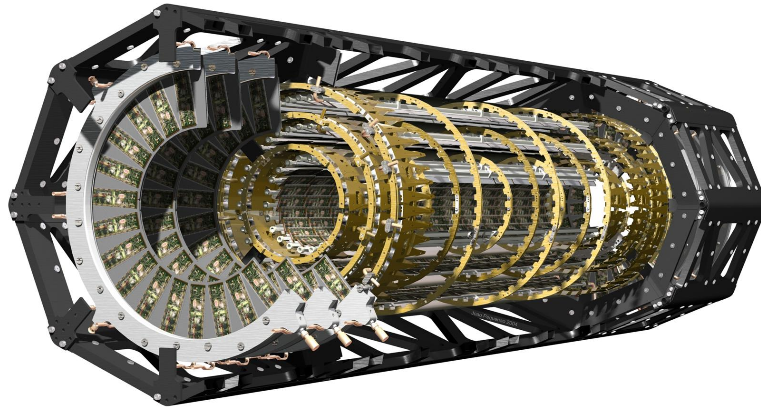


Figure 2.14: Overview of the ATLAS pixel detector [7].

As the B-layer layer is paramount in tagging jets stemming from B-hadrons and τ-leptons, its performance has to be maintained during the entire experiment lifetime. In particular b-tagging and vertexing efficiency have to be consolidated for operation at higher than design luminosity during Phase-I. To achieve this, a new innermost layer, the Insertable B-Layer, is being constructed and will be inserted in 2014 into the present pixel system. By using a new, smaller beam pipe, the IBL can be inserted at a radius of 3.5 cm. The proximity to the interaction point translates into higher rates and radiation levels. To cope with these, the IBL is designed with a smaller pixel size of 50 µm in R - ϕ and 250 µm in z, and features new sensor design and electronics able to withstand a fluence of 3×10^{15} $\mathrm{n}_{eq}\mathrm{cm}^{-2}$ and a TID of 250 Mrad. The thickness of the detector is optimized to reduce multiple scattering that would degrade the performance of the next pixel layers, and to allow insertion of the IBL within the pixel envelop with only 1 mm clearance from the present B-layer. The target thickness for the IBL is 1.5 % X_0. With these features the IBL is able to keep a comparable occupancy at higher luminosity

[2] The name B-layer comes from the fact that the innermost detector layer is of fundamental importance for measurement of secondary vertices of B-hadrons.

as the present B-layer, which is fundamental to maintain vertex resolution in a higher pile-up environment. Also, the IBL can assure full b-tagging efficiency in the case of failures in the present b-layer. Thanks to its closer position to the beam pipe, the IBL actually improves the quality of impact parameter reconstruction for tracks and thereby improves vertexing and b-tagging performance. As a consequence, the detector sensitivity for signals in physics channels involving b-jets is increased. This is for instance important for the low mass SM Higgs in the channel $H \to b\bar{b}$. More details on the ATLAS pixel detector and the IBL are given in Chapter 3.

2.4.2 SemiConductor Tracker

The SemiConductor Tracker (SCT) is designed to provide four precision measurements per track in the intermediate range, contributing to the measurement of momentum, impact parameter and vertex position, as well as providing a good pattern recognition by the use of high granularity. Micro-strip detectors, made of two 64 mm long daisy-chained sensors with a strip pitch of 80 µm, are arranged in four double layers in the barrel and nine disks per end-cap. A set of two micro-strip detectors is used in each barrel layer, as well as in each disk, with a relative stereo angle of 40 mrad. One set of strips is mounted parallel to the beam pipe in the barrel and radially in the disks. This stereoscopic geometry allows to perform three-dimensional position measurements. The SCT has approximately 6.3×10^6 readout channels. The intrinsic accuracy is 17 µm in R - ϕ and 580 µm in z in the barrel. The end-caps provide an accuracy of 17 µm in R - ϕ and 580 µm in R.

2.4.3 Transition Radiation Tracker

The TRT is made of straw tubes with 2 mm radius and equipped with a 30 µm diameter wire. The straws in the barrel are 144 cm long and placed parallel to the beam line. They are divided in two at the center and readout at each end, in order to reduce occupancy. The end-caps are equipped with 37 cm long straws arranged radially, and readout on one side. The total number of TRT readout channels is $\sim 350 \times 10^3$. With an average of 36 points per track, the TRT provides track-following up to $|\eta| = 2.0$, assuring robust pattern recognition. The TRT provides only R - ϕ information, for which it has an intrinsic accuracy of 130 µm per straw. The straw hits contribute significantly to the momentum measurement, since the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length. The TRT also provides electron identification capabilities by detecting transition-radiation photons in the xenon based gas mixture of the straw tubes.

2.4.4 Tracker design for HL-LHC

The ATLAS ID is designed to operate for ten years at a peak luminosity of 10^{34} cm^{-2}s^{-1}, with an average of 23 pile-up events per bunch crossing, and a level-1 trigger rate of 100 kHz. At the end of the LHC program, most of its components will approach the end of their lifetime, namely 700 fb^{-1} for the strips and 400 fb^{-1} for the pixel systems. To withstand to harsher HL-LHC environment a completely new inner detector is needed. At the HL-LHC the peak luminosity will reach 5×10^{34} cm^{-2}s^{-1}, corresponding to 140 pile-up events per bunch crossing.

The delivered total integrated luminosity over ten years will be 3000 fb^{-1}. The higher luminosity will increase significantly the occupancies in both silicon detectors, and the occupancy in the straw tube TRT will reach almost 100 %. To cope with the increased occupancy, finer granularity is needed and the readout of the tracker must provide higher bandwidth than at present time. A multiplicity of more than a 1000 tracks per unit of rapidity is expected in the tracker acceptance. An even higher local track density is expected in the core of high transverse momentum jets, which are a key part of the HL-LHC program. With a longitudinal beam spot standard deviation of $\sigma = 75$ mm, it is a major challenge to associate the tracks with the correct primary or secondary vertices at high efficiency. The high rates translate also in unprecedented radiation levels. As shown in Table 2.3, the fluence will be one order of magnitude larger as for the current detector, and the innermost pixel layer will experience a TID of almost 1 Grad. From a mechanical point of view, greater robustness and redundancy are envisaged, together with a modular detector layout.

Table 2.3: Predictions for the maximum 1 MeV n_{eq} fluence and TID for 3000 fb^{-1} for the ATLAS tracker at the HL-LHC [11].

System	Fluence [10^{15} $n_{eq}cm^{-2}$]	TID [Mrad]
Pixel layer 0	14	770
Pixel layer 3	1.7	90
Pixel disk 1	1.8	90
Short strip layer 1	0.53	22
Long strip layer 4	0.29	63
Strip disk 7	0.81	29

Alongside the requirements dictated from the HL-LHC experimental environment, the guiding principle in designing a tracker for Phase-II is the anticipated physics program. Specifically, the main requirements of the new inner tracker should be

- measurement of the transverse momentum and direction of isolated particles, such as electrons and muons;
- reconstruction of vertices of pile-up events and identify those associated with the hard scattering;
- identification of secondary vertices from b-jets and τ-leptons;
- measurement of tracks in the core of high energy jets;
- reconstruction of tracks associated to converted photons.

To satisfy all the requirements mentioned above, the layout shown in Figure 2.15 is studied [11]. Based on the experience gained with the current tracker, an all-silicon tracker is proposed with pixel sensors at the inner radii to provide pattern recognition and precision position measurements close to the vertex, complemented by strip sensors at larger radii to provides accurate tracking at lower cost and material. The pixel system extends to larger radii and provides more hits in the forward region to increase tracking in dense regions. The momentum resolution is improved by the slightly larger outer radius. In the central region, sensors are arranged on a barrel structure consisting of four pixel layers followed by three short-strip layers and two long strip layers. The forward region is covered by six pixel disks and seven strip disks per side. The first two pixel layers are designed to be replaced without removing the beam pipe. The outer pixel layers and disks should be removable without disturbing the strips. The end-cap strips are retractable to access their services and those at the end of the barrel. Pixel and strip systems will cover an area of 8.2 m^2 and 193 m^2, respectively, instead of the current 2.3 m^2 and 61.1 m^2 [6]. The expected number of channels is ~600 $\times 10^6$ for the pixels and ~70 $\times 10^6$ for the strips.

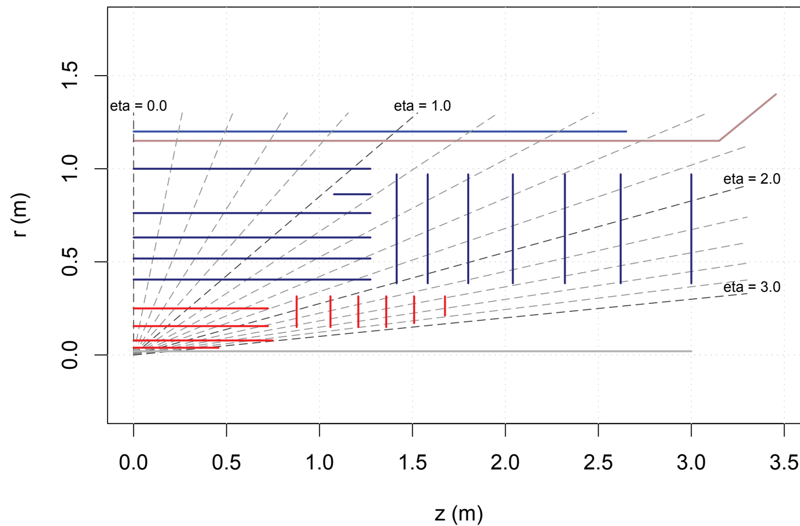


Figure 2.15: Baseline layout of the ATLAS tracker for the HL-LHC, showing the active areas of silicon detectors arranged on cylinders and disks [11].

The two innermost pixel layers are designed with a pixel size of 25 $\times$125 μm^2 and using new sensor and electronics technologies to cope with the high rate and radiation levels. For the outermost pixel layers and disks which have to cover a large area at lower radiation levels and rates, the main requirements is low cost. Their development can profit of the technologies developed for the IBL, which provide the required granularity, i.e. pixel size of 50 $\times$250 μm^2, and radiation hardness at the affordable cost of mature technologies.

The strip layers are double-sided, with axial orientation on one side and sensors rotated by 40 mrad on the other side. Also in this case the innermost layers are equipped with short strips, 23.8 mm, to increase the granularity at small radii, and with long strips, 47.8 mm at larger radii.

The new ID system is designed to minimize the overall tracker material. The expected material distribution is shown in Figure 2.16 in units of X_0. A major improvement is made with respect to the current ID. The main material reduction is achieved by using modern light carbon fiber-based engineered materials for the support structure, and with dedicated design and routing of the detector services. In particular, the services from the pixel detector are routed inwards to small radii and then out along z, to remove them as soon as possible from the tracking region. Material from sensor and electronics is reduced as well. If the current ID contributes $\geq 1.0\ X_0$ for all regions $|\eta| \geq 0.7$, the new tracker remains mostly below $0.7\ X_0$ up to $|\eta| = 2.7$. In this framework, the work presented in this thesis contributes to the reduction of the material budget of the pixel detector with the technologies discussed in Chapter 4, Chapter 5, and Chapter 6 to achieve the target thickness of $\leq 1.5\ \%\ X_0$.

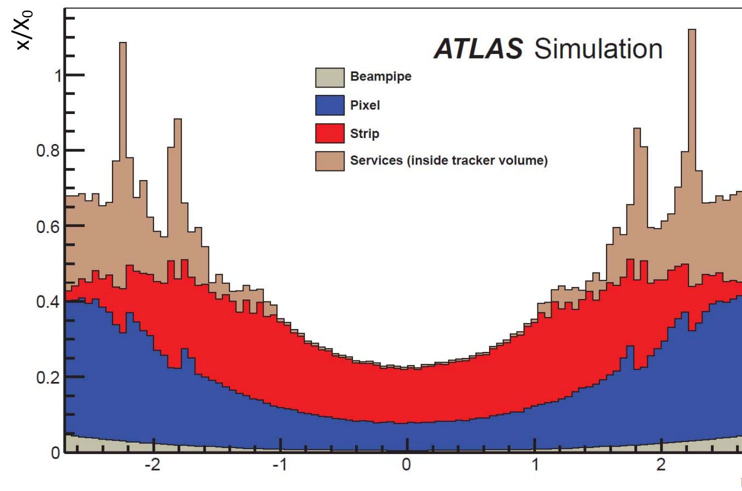


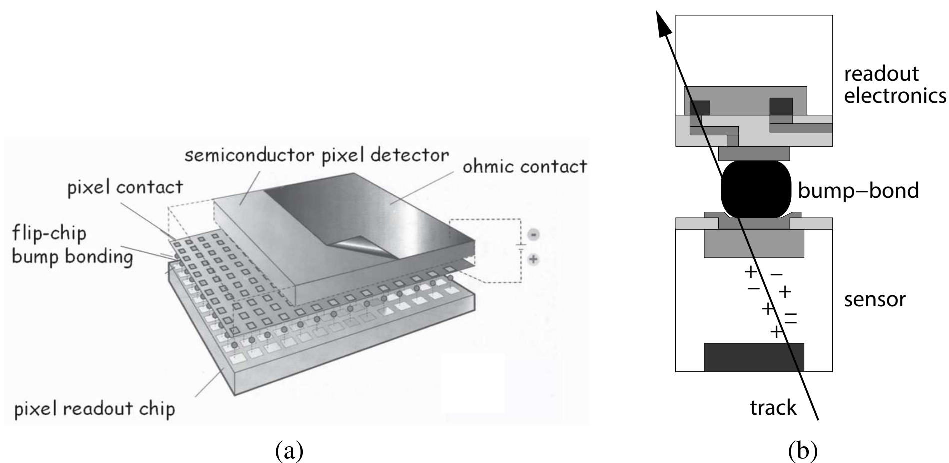
Figure 2.16: Material distribution expressed in units of X_0 as a function of η for the ATLAS tracker at the HL-LHC. The breakdown shows the contribution from the services internal to the tracker volume, the sub-detectors and the beam pipe (adapted from [11]).

Chapter 3

Pixel detectors for HEP

3.1 Hybrid pixel detectors

Pixel detectors are the detectors of choice in HEP experiments as the innermost tracking element [19]. Three of the LHC experiments, ATLAS, CMS and ALICE, employ pixel detectors at the center of their silicon tracker [20, 21, 22]. Pixel detectors feature not only excellent spatial and time resolution, but as well the high granularity and radiation hardness needed at small radii at the LHC. The pixel detector system plays a fundamental role in vertexing, for the identification and reconstruction of secondary as well as primary vertices, in high pile-up environments. It also contributes to pattern recognition, event reconstruction and momentum measurement in combination with the rest of the tracking system.



(a) (b)

Figure 3.1: (a) Generic hybrid pixel detector, showing the 1 :1 connection between sensor elements and readout channels, using flip chip technology [19]. (b) Cross-section of one hybrid pixel cell, consisting of a bump bonding connection between sensor and electronics. The track of an ionizing particle generating electron-hole pairs is also depicted. The charges drift in the depletion region of the sensor under the action of an electric field, producing a signal on the pixel electrode. The signal is amplified and digitized by the electronics readout chip [20].

Pixel detectors for the LHC experiments are made of a large number (in the order of $\sim 10^3$) of building blocks, called modules. A module provides thousands of elementary sensing ele-

ments, the pixels, each one with an associated readout channel, for fast, parallel readout (Figure 3.1). This particular module configuration, in which a high density electrical interconnection is used to mate sensor and electronics, is called a hybrid pixel detector [23]. Each module is an independent multichannel detector, and both sensor and electronics technologies can be optimized to achieve the desired performance. To assure an hermetic, large area coverage, the modules are mounted, with some overlap, on a support structure consisting of a cylindrical barrel region and a forward part on each side of the barrel. In the ATLAS pixel detector 1744 modules are used to cover an area of approximately 2.3 m^2. The support structure needs to be stable, and to provide both electrical isolation and thermal conductivity. The cooling for temperature control is in fact often integrated in the mechanical support. Operation of the pixel detector requires a large volume of service cables to provide power, data and control signals to all detector channels. One of the most difficult challenges in designing a hybrid pixel detector, together with the radiation hardness, is to meet the requirements in terms of material budget. The modules, the cables, and the support structure have to be as thin as technically possible to prevent unwanted particle interactions that would degrade the performance of the entire experiment. Figure 3.2 shows the material distribution of the ATLAS pixel detector in units of X_0, as a function of η. The three ATLAS pixel detector barrel layers have a total thickness of 10.8 % X_0 at normal incidence, i.e. $\eta = 0$, corresponding to a thickness of 3.6 % X_0 per layer [4]. The contributions from the different detector elements are discussed in the following paragraphs. The material budget of the detector is always given per layer and in units of X_0. Unless explicitly referenced, the given numbers are calculated according to the weight of every individual detector component. The total sum agrees with what is quoted in [4].

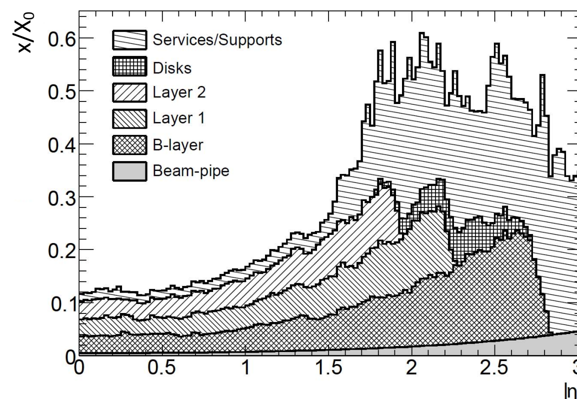


Figure 3.2: Material distribution of the ATLAS pixel detector expressed in X_0, as a function of η. The breakdown shows the contribution from each layer and disk, including services and support structures directly in front of and behind the layer or disk. The remaining services and supports are included in the category "Services/Supports" (adapted from [20]).

3.1.1 Modules

Hybrid pixel detector modules for HEP experiments consist traditionally of a pixelated silicon sensor connected to one or more Front-End (FE) readout chips, using a flip chip process, as

explained in Chapter 5. Connection to the detector services, namely power, data, and control signals, is done using High Density Interconnect (HDI) circuits. Figure 3.3 shows a module of the ATLAS pixel detector, where a multi-chip module configuration is used. The ATLAS Pixel Sensor (APS) [24] features 47 232 pixels with a nominal size of $50\times400\,\mu m^2$, implemented as n^+-implants in a high resistivity n-doped silicon bulk (n^+-in-n). The biasing electrode is implemented as a p^+-implant on the sensor back side. This sensor design is referred to as planar. Radiation hardness for fluences up to $10^{15}\,n_{eq}cm^{-2}$ is reached using an oxygen enriched substrate [25]. The required sensor bias voltage is 150 V before irradiation and increases up to 600 V after irradiation. In order to have a controlled potential drop towards the edge of the sensor during high voltage operation, the biasing electrode on the sensor backside is surrounded by 16 guard rings (Figure 3.7). The sensor is bump bonded to 16 FE-I3 readout chips, arranged in two rows.

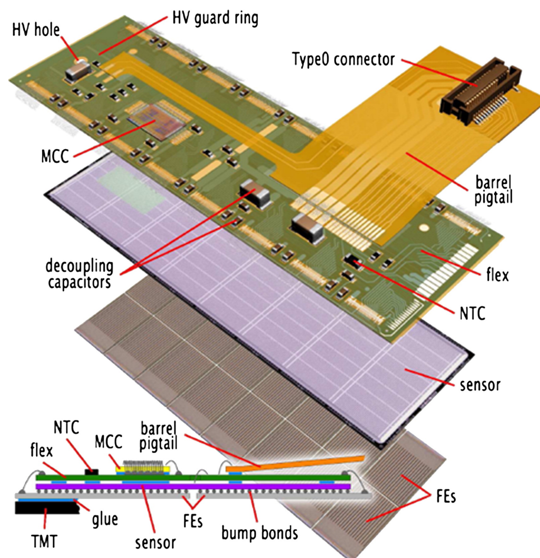


Figure 3.3: View of the different parts of an ATLAS pixel detector module. Starting from the bottom: FE chips, sensor, flex hybrids with passive components. A side view of the module is also shown, where the bump bond connection between electronics and sensor can be seen. The module has a total area of $63\times24.2\,mm^2$ [20].

The FE-I3 [26] is a mixed-mode Integrated Circuit (IC). Analog and digital functionality are integrated in a pixel matrix of 18 columns and 160 rows (Figure 3.9). Wire bond pads, I/O and powering circuitry are located at the bottom of the pixel matrix in a region called periphery. In each FE-I3 readout channel, the sensor charge signal is amplified by a Charge Sensitive Amplifier (CSA), which generates a voltage signal proportional to the deposited charge. The feedback capacitor of the CSA is discharged by a constant current, causing a linear decrease of the signal after a hit. The output of the CSA is then compared to a programmable threshold by a discriminator. Because of the linear discharge of the amplifier, the discriminator output signal, so-called Time Over Threshold (TOT), has a length proportional to the generated charge. The hit information is sent to buffers in the periphery, and stored here until a trigger signal arrives. If no trigger signal is received by the FE in coincidence with the so-called latency time, the

data is erased. For testing purposes, a charge of a definite amount can be injected via a charge injection circuitry integrated in each FE channel. The technology used for this FE chip is a quarter micron (i.e. 250 nm) commercial CMOS technology, which provides the required integration level and timing performance, together with radiation hardness. The FE-I3 has a time resolution smaller than 25 ns to associate hits to the correct LHC bunch crossing. The use of special layout rules makes the FE-I3 tolerant to a TID of 100 Mrad. Operation of the FE-I3 requires an analog (VDDA) and a digital (VDDD) voltage of 1.6 V and 2.0 V respectively. The total chip current consumption for a standard configuration is ~0.11 A.

Each FE is wire bonded to a HDI circuit glued on the sensor back side, implemented as a fine pitch, double-sided flexible printed circuit, referred to as flex hybrid. Copper traces are patterned on both sides of a polyimide substrate, chosen for its radiation hardness, functional reliability and high dielectric strength. The flex hybrid hosts passive components such as power filters and termination resistors, and the Module Controller Chip (MCC) [27]. The MCC is responsible for the serialization of the data from the FEs, and the generation of the control signals. Over the flex, power is distributed to the FEs, data and control signals are routed between the FEs and the MCC. Finally, a second HDI, so-called pigtail, implemented as a kapton strip with copper traces, hosts the connector to plug the module to the detector electrical services (Figure 3.3).

Conventionally the pixel matrix defines the active area of the module, i.e. the part sensitive to interacting radiation, whilst the rest of the module is considered inactive, and it is referred to as passive material. To lower the material budget of the module the inactive fraction should be minimized with respect to the active fraction. The active area of an ATLAS pixel module is 65 % of the total area. The inactive area comes from both the sensor guard rings and the FE periphery. The 16 guard rings of the sensor cover a width of 525 µm (Figure 3.7). Considering also the distance from the last guard ring to the sensor edge, a dead area of 1.1 mm width surrounds the entire sensor, corresponding to 12 % of the total module area. The main contribution to the inactive area comes from the FE chip: the periphery of the 16 FE-I3 corresponds in fact to ~23 % of the total module area.

To keep the material budget low, all elements of the module have to be thinned to the minimum allowed thickness. The sensor is produced on 250 µm thin wafers. This thickness is required for mechanical support during all processing steps. The FE electronics is on the other hand thinned after processing. In this case constraints on the minimum allowed thickness come from handling of the FE dies during the hybridization process. The process used for the ATLAS modules allows thinning of the FE-I3 down to 190 µm, as explained in Chapter 5. The MCC is as well thinned below 200 µm. The thickness of the HDI is 100 µm. For a barrel layer, at perpendicular incidence, the material budget of the modules is 1.2 % X_0, which splits between

- sensor: 0.32 % X_0;
- electronics and bump-bonds: 0.27 % X_0;
- flex hybrid with passive components and MCC: 0.39 % X_0;
- pigtail with connector: 0.22 % X_0.

3.1.2 Electrical services

The high granularity of pixel detectors translates into a large number of cables needed to provide power, data and control signals, making service cables one of the most significant contributions to the material budget in trackers. Due to the concentration of cables at the end of the barrel and of the end-caps, the contribution of the services becomes particularly severe at large η, where they are the dominant source of material, and degrade the performance of subsequent detectors (Figure 2.13 and Figure 3.2). In addition to this, the high number of cables makes routing particularly difficult. Services need to be installed over distances that can easily attain 100 m, between the pixel detector at the heart of the experiment and the service room, where power supplies and back-end readout circuitry are located in a radiation free area. Cable channels along this path have a limited cross-section, due to space constraints coming from the need to keep a sealed envelop for environmental control (i.e. temperature and humidity), and from the other detector sub-systems.

To ease routing and integration, services to the ATLAS pixel detector are designed in a modular way, as shown in Figure 3.4, with the possibility of disconnection at intermediate locations between the detector and the service room, using so-called Patch Panels (PPs). The Low Voltage (LV) power for the FE chips and the High Voltage (HV) needed to bias the sensors are provided by power supplies in the service cavern, via routing at the PP4 region. From here, the LV is distributed to the modules over the Type-IV, Type-II, Type-I and Type-0 cables, while the HV is routed on the Type-IV cables all the way to PP1. Data transmission to and from the modules is done electrically until PP0, where the high radiation levels do not allow the use of optical fibers. Optical transceivers are installed at PP0 and data are transmitted optically to/from the service room. Detector monitoring signals and power distribution to the optical items are routed via the PP3 region. Services until PP0 are referred in the following as on-stave services, and those outside the PP0 area as off-stave services.

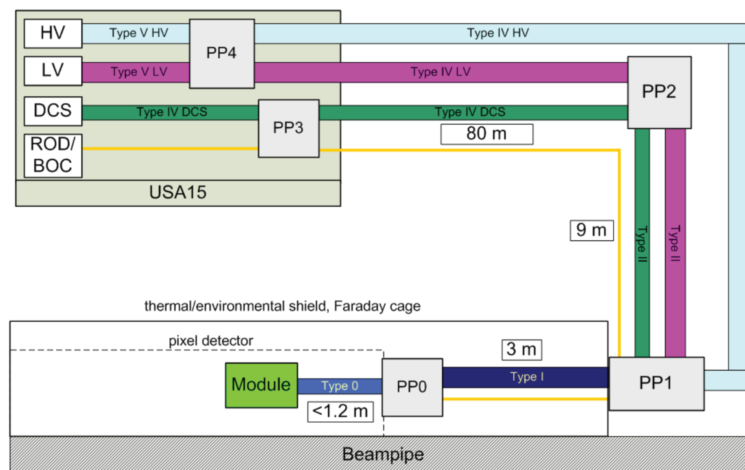


Figure 3.4: Schematic layout of the ATLAS pixel services scheme, between the service room (USA15) and the modules. Different cable types are used for the various services, along the 100 m long path. The naming of the cables is given according to the PP name. For instance, the cables connecting PP0(1) to PP1(2) are called Type-1(2) cables [28].

A large fraction of the services of hybrid pixel detectors is dominated by the *LV* cables to distribute power to the front-end electronics. High speed pixel detectors, with millions of readout channels, designed using sub-micron CMOS technologies, have high current consumption in the order of kA, leading to typical power densities in the order of kWm^{-2}. Power is traditionally distributed using voltage based powering schemes, in which the modules are powered from a constant voltage source. Possible configurations are either point-to-point, in which each module is powered with a set of independent power lines (independent or direct powering), or parallel, where the power lines are shared by a group of modules. These configurations allow control over single detector elements, but due to the large number of modules to be powered, introduce a very high number of cables in the detector. Within the detector active area, low material budget and space constraints require the use of cables with cross-section as small as possible to reduce the cable volume. The cable cross-section A, for a transmitted current I over a length l, and a voltage drop V_{drop} on the cable is given as

$$A = l \cdot \frac{I}{V_{drop}} \cdot \frac{1}{k} \tag{3.1}$$

where k is the electrical conductivity of the cable conducting material. Due to the high current that has to be transmitted, a high voltage drop has to be allowed to reduce A, translating into high power losses on the cables ($V_{drop}I$), which reduce efficiency and produce a lot of heat that needs to be removed by massive cooling systems. In addition to this, the maximum allowed voltage drop between the hybrids and the supply is constrained, to prevent over-voltage across the modules in case of a sudden current drop on the power line. As the value of A cannot be minimized as desirable, the material of the power cables contributes significantly to the material budget of the services, and saturates the cable channels.

In the ATLAS pixel detector, each module has to be powered with the analog and digital voltages needed by the FE, respectively 1.6 V and 2.0 V, and consumes around 2 A. The total current consumption of the pixel detector is quoted to be 3.7 kA, for a total power consumption of 6 kW [4]. The power losses on the cables between the service room and the modules amount to 24 kW ($V_{drop} = 6.4$ V). The power that needs to be provided is thus 30 kW, corresponding to a power efficiency of only 20 %. Regulator stations are installed at PP2 to protect the modules from transients. Sense lines run from PP2 to the modules for voltage regulation. Powering the 1744 ATLAS pixel modules requires a set of four cable pairs for each module, one for VDDA, one for VDDD, and one pair of sense lines for each voltage. The power cables contribute to 76 % of the cable material budget in the active area. The on-stave service cables in fact account for 0.13 % X_0, split between

- *LV* cables (Al): 0.10 % X_0;
- *HV* cables (Cu): 0.02 % X_0;
- data and control (Cu): 0.01 % X_0.

3.1.3 Mechanics

The support structure of a pixel detector serves multiple purposes. It is designed to allow precise module arrangement, assuring hermetic coverage over a defined pseudorapidity range. It should serve as a heat spreader and good thermal conductor for module cooling, while remaining stable over thermal cycling to keep detector alignment. The robustness of the structure needs to be assured by light structures adding as little material as possible to the detector. Light cooling systems are required as well for operation of a pixel detector. Cooling is fundamental for sensor operation after irradiation to avoid reverse annealing and thermal runaway. In addition to this, as already mentioned in Section 3.1.2, the electronics and cables produce a lot of heat that needs to be removed from the detector.

The active area of the ATLAS pixel detector is organized in a barrel region covering a pseudorapidity range $|\eta| \leq 1.5$ and forward regions covering up to $|\eta| = 2.5$. As shown in Figure 3.5, the barrel part is made of three concentric layers with modules loaded on local supports, called staves [29]. Each stave is ~80 cm long and ~2 cm wide, and hosts 13 modules. Due to the inactive module area, full longitudinal and azimuthal coverage is achieved by arranging modules and staves in a way that modules overlap in z along the short edge, and staves overlap in $R - \phi$. The staves are inclined by 20° with respect to the radial direction to increase the possibility of charge sharing between two pixels, which in turn increases the spatial resolution. Figure 3.6a shows the fully equipped innermost layer after assembly into the barrel structure. The forward region is formed by two end-caps with three disks each. Also here modules are mounted on smaller supports called sectors. Each sector hosts three modules per side. Modules on the back side are shifted with respect to those on the front side of the sector to provide full coverage. Eight sectors form a disk (Figure 3.6b). Staves and disks (local support) are made using a Carbon-Carbon (C-C) material with good transverse thermal conductivity, low material budget, excellent stiffness and stability, and low value of the coefficient of thermal expansion (CTE). Barrels and end-caps are supported by a frame (global support) manufactured as well using carbon fiber composite material.

Cooling is integrated in both staves and sectors. An aluminum cooling pipe runs on the backside of each stave and sector, in the latter case with a W-shape. The elementary cooling unit is formed by two barrel staves and two sectors in the disks. The chosen cooling system is a bi-phase evaporative cooling system with C_3F_8 or octafluoropropane [30]. This scheme provides large cooling capacity per unit volume, thanks to the high heat transfer coefficient between the liquid and the elements to be cooled, as well as small temperature gradients along the cooling pipes. The required radius of the cooling tubes is small, leading to a low coolant mass flow, beneficial for material reduction. The C_3F_8 liquid is chosen as it is non-flammable, electrically non-conductive, chemically inert in the event of a leak, and very stable against the high radiation doses expected at the LHC. In order to operate the modules at a temperature of −6 °C, the coolant is kept at a temperature of −25 °C. The absolute pressure of the two-phase fluid as it flows in the cooling structures of the detector is between 1.7 bar and 6 bar, which is reasonably well matched to the requirement of minimal material for the cooling structures.

Mechanics and cooling account for the largest part of the material budget of the pixel detector at normal incidence. As quoted in [4], for one barrel layer, the local support structures with cooling add 1.8 % X_0, and the global support 0.5 % X_0.

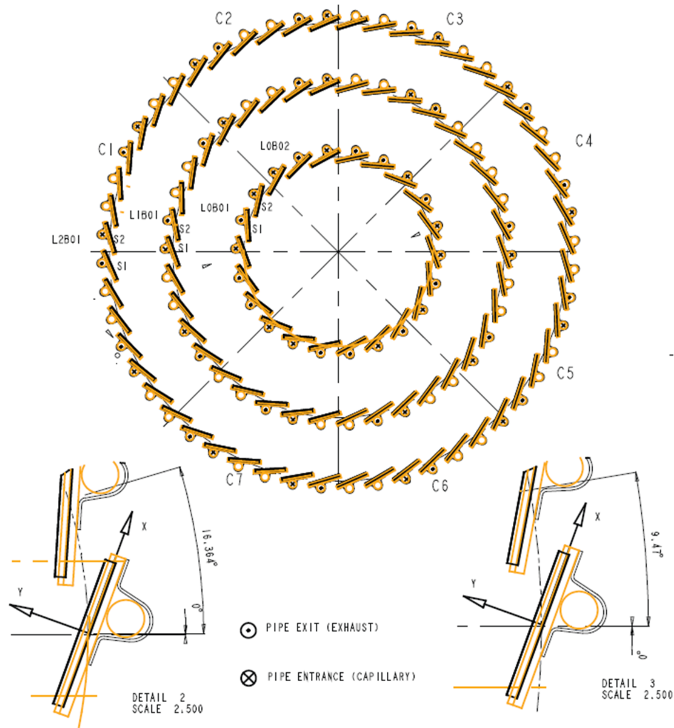
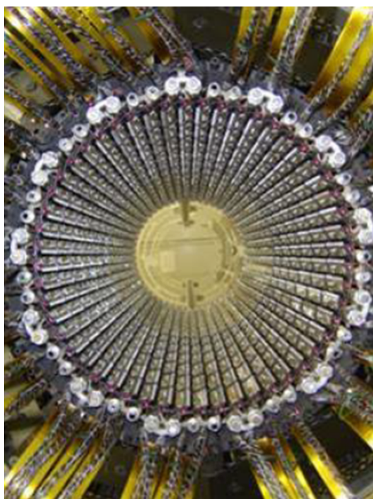
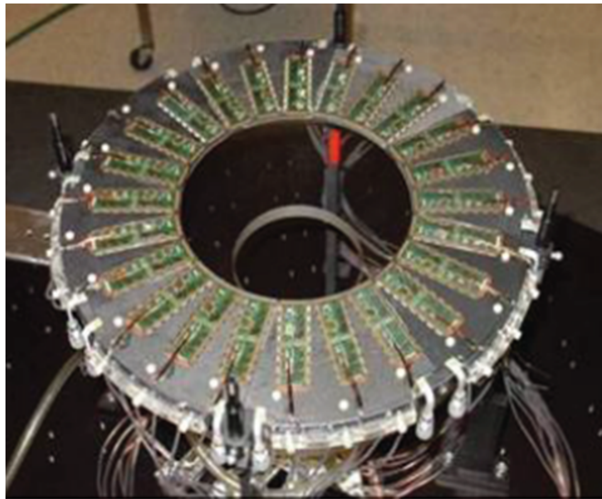


Figure 3.5: Drawing of the stave arrangement to form the three barrel layers. The stave cross-section with modules on the front side and cooling pipe on the back side is also shown [18].



(a)



(b)

Figure 3.6: Photographs of (a) the *b*-layer and of (b) a disk of the ATLAS pixel detector, fully equipped with modules and service cables [18].

3.2 Low material developments

Hybrid pixel detectors are used for the first time as instrumentation for HEP experiments at the LHC. Their design and construction required mastering and adapting semiconductor as well as interconnection technology to the needs of the experiments. The great effort led to detectors that can now fulfill physics requirements in the harsh LHC environment. Nevertheless, the technology available at the time of design, together with the lack of previous experience in building such large scale, high granularity detectors, resulted in some drawbacks. Pixel detectors at the LHC are comparatively massive, having a material budget much higher than what planned during the design phase, and are expensive in fabrication. The use of hybrid pixel detectors for the high luminosity upgrades of the LHC requires the development of new technologies to cope with the higher rate and radiation levels, while reducing detector thickness and cost. Driven by the IBL and Phase-II upgrades, the ATLAS pixel detector collaboration has started R&D programs covering a wide range of detector elements [31], from sensor and FE technology and design, to lighter mechanical structures and powering schemes, as well as new interconnection technologies and module concepts.

Mature R&D projects are presented here, as the work carried out in this thesis has contributed to these developments. In particular, the benefits in terms of material budget are discussed. The described sensor and FE technologies are currently used to build the IBL detector modules, and are chosen as baseline for the outermost pixel detector layers at HL-LHC, given the same requirements in terms of rate, radiation and granularity for the two detector systems. Dedicated electronics developments are required for the innermost pixel layers at the HL-LHC to cope with the high hit rate and radiation levels. R&D is ongoing on 65 nm FE chips and on 3D electronics. In addition to the sensor options described in Section 3.2.1, also diamond is considered as sensor material for the inner layers, given its high radiation hardness. References on these developments can be found at [32, 33, 34].

3.2.1 Sensor

As explained in Section 3.1.1, the rim and thickness of the sensor design introduce inactive material in hybrid pixel modules. Sensor developments for high luminosity upgrades address these sources of inactive material with different design and processing strategies, for the various sensor options considered.

Planar sensors are currently tested for fluences up to $2\times10^{16}\ n_{eq}\text{cm}^{-2}$. Although results at these high fluences are still limited, they indicate that the planar sensor concept can match the requirements for operation in high luminosity environments [35]. In addition to the n^+-in-n configuration as currently used by the ATLAS pixel sensor, also n^+-in-p is considered for the upgrades of the pixel detector [36]. This newly available technology allows electron collection as the n^+-in-n type, while being a cheaper and simpler process (single-sided process).

For both planar sensor types slim edge configurations are tested to reduce the dead area at the rim of the sensor to less than 500 μm for the IBL [37], and down to 100 μm or smaller for the HL-LHC. Sensors with slim edges can be placed on staves without the need of overlap, reducing the detector envelop and thus allowing better detector coverage and positioning at smaller radii. These characteristics are of critical importance for the IBL, which has to be

inserted at a radius of 3.5 cm between the beam pipe and the current *B*-layer, with only 1 mm clearance. They are as well beneficial for the large area detector for Phase-II to provide full detector coverage. The planar n^+-in-n sensor design for the IBL is an example of currently available slim edge planar sensor [38]. This design is based on the APS, but with smaller pixel size for interconnection to the FE-I4 chip ($50 \times 250\,\mu m^2$ instead of $50 \times 400\,\mu m^2$). As shown in Figure 3.7, 13 guard rings over a width of 350 µm are used on the sensor back side, instead of 16. The edge pixels on the front side are longer, and overlap the guard rings to collect charge as close as possible to the physical edge. The total inactive edge is of only 200 µm. While the IBL planar sensor design still features a conservative thickness of 200 µm, the goal for the HL-LHC upgrade of the pixel detector is to have ≤ 150 µm thin sensors. Thinning is beneficial for planar sensors not only to reduce material but also to increase the charge collection efficiency after irradiation. Using dedicated fabrication processes with handle wafers, both planar sensor options have been thinned to 75 µm without loss in performance in terms of bias voltage for full depletion, leakage current, and breakdown behavior [39].

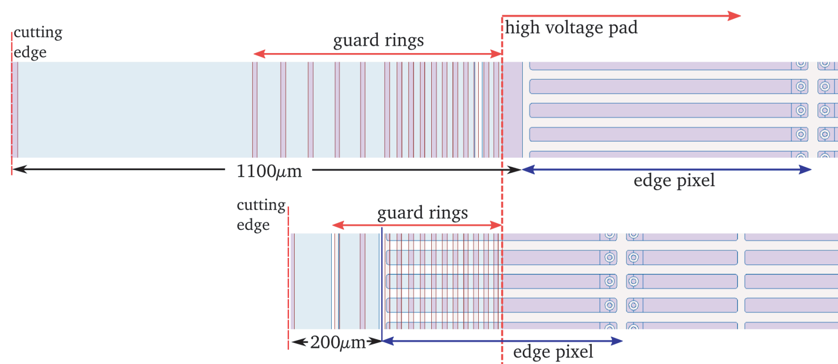


Figure 3.7: Comparison of the edge region of the APS design (top) and the IBL planar sensor design (bottom). The APS has an inactive area width of 1.1 mm, while the IBL planar design reduces it to 0.2 mm, by using less guard rings and extending the edge pixel under the guard ring structure [38].

Alternative to the planar sensor design, 3D sensors are considered for the upgrades of the ATLAS pixel detector. In this case collecting and biasing electrodes are etched as n^+ and p^+ columns through a p-type silicon bulk. Due to the smaller distance between the electrodes, with respect to the planar design, they show better charge collection efficiency after irradiation with lower bias voltage. So far two etching options are studied, shown in Figure 3.8: Full3D with active edges [40] and double-sided 3D with slim fences [41], where the latter is used for the IBL detector. Processing of the Full3D sensors is a single sided process in which columns are etched from the sensor front side. This process integrates as well the formation of active edge electrodes as p^+ walls along the perimeter of the sensor, leaving a maximum of 50 µm wide inactive edge. In the double-sided 3D process electrodes are etched from both sides (n^+ columns from the front side, p^+ columns from the back side). In this case a controlled voltage drop towards the edge requires the use of structures similar to the guard rings in a planar sensor. For the IBL 3D sensor design a guard ring structure is implemented as a fence around the active area with a combination of n^+ and p^+ columns, for a total inactive width of 225 µm. 3D sensors of both types have been produced with a thickness between 200 µm and 250 µm, where 230 µm

is chosen for the IBL.

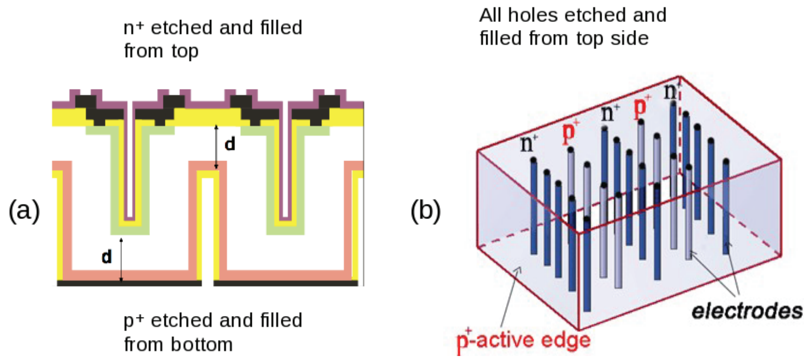


Figure 3.8: Sketch of (a) a double-sided 3D sensor and of (b) a Full3D sensor with active edges [9].

3.2.2 FE-I4

The FE-I4 (Figure 3.9) is the next generation hybrid pixel detector readout chip [42]. It is designed using a commercial 130 nm CMOS technology, featuring a higher density of digital circuitry and higher radiation hardness with respect to the 250 nm process used for the FE-I3. The FE-I4 has 26 880 pixels of $50 \times 250\,\mu m^2$, integrating more functionality than the FE-I3. In particular the FE-I4 integrates as well the functionality of the MCC, so that an external controller chip is not needed anymore. Its radiation hardness is demonstrated up to 250 Mrad. The main difference with respect to FE-I3 is the implementation of a new readout architecture. Hit information is stored locally, within a 4-pixel region, until selected by the trigger. This new readout scheme, made possible by the smaller feature size process, provides the higher rate capability required by operation at the IBL, and can satisfy as well the requirements for operation at the outer pixel layers at the HL-LHC. Two versions of the FE-I4 chip exist, denoted with the letters A and B. The FE-I4A is the full scale prototype chip to test the digital readout logic architecture, and among others different flavors of pixel cells, and powering options, while the FE-I4B is the dedicated IBL production chip.

The FE-I4 is the largest pixel readout IC produced to date. This large size is chosen mainly to reduce the cost of bump bonding, which scales with the number of chips but not with the chip area. A secondary benefit of a larger chip is the increased ratio between active and inactive area, which helps reducing the material budget of the module. In a large area chip in fact, the area of the periphery is reduced with respect to the pixel matrix. While the length of the periphery is fixed by the chosen number of columns, the height of the periphery is defined by the amount of circuitry that needs to be integrated, and by module assembly constraints. The wire bond pads on a pixel FE chip in fact have to be accessed by the wire bonding tool after the sensor is bump bonded to it. As the peripheral logic does not scale with the number of columns, from a certain number of columns on, the height of the periphery is defined only by the module assembly constraints. As shown in Figure 3.9, the FE-I4 has an area of almost $4\,cm^2$, five times bigger than the previous generation FE, with more than four times the number of columns. The

peripheral circuitry is reduced thanks to the new architecture, which moves hit storage from the end of column region to the pixel matrix. The height of the periphery in the FE-I4 is thus fixed only by the wire bonding requirements, and the periphery is partly empty. The ratio between pixel matrix and periphery is increased to 9 :1, with respect to 3 :1 of the FE-I3.

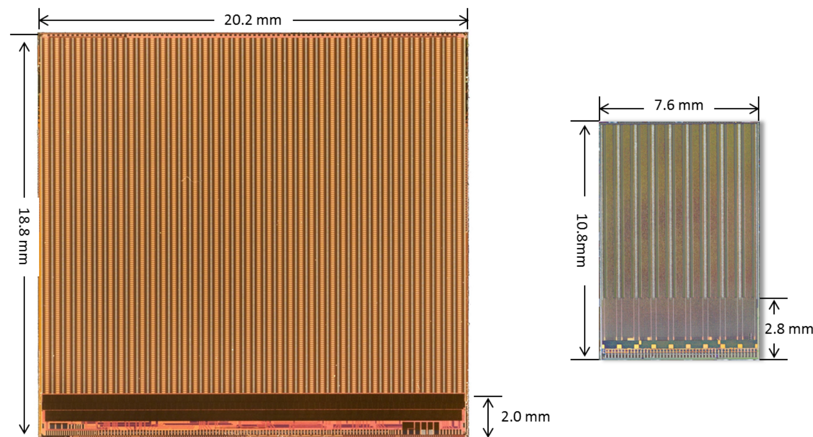


Figure 3.9: Picture of the FE-I4 (left) and FE-I3 (right) pixel readout chips, to scale. They feature respectively 26 880 and 2880 pixels, arranged in matrices of 80 ×336 and 18 ×160. The pixel size, defined by the chosen CMOS process is 50 ×250 μm^2 for FE-I4 and 50 ×400 μm^2 for FE-I3. The FE-I4 area is 3.8 cm^2, and the FE-I3 area is 0.82 cm^2 [38].

The drawback of a large chip, from a material point of view, is the higher thickness required by the flip chip process for connection to the sensor. The required chip thickness for bump bonding with the flip chip process used for the ATLAS pixel detector is 190 µm for the FE-I3, and increases to 450 µm for the FE-I4, due to need of controlling chip planarity during reflow at high temperature. To profit from the benefit associated to larger FE size, dedicated flip chip technologies are thus needed to achieve thin modules. A flip chip process developed within the work of this thesis allows building IBL modules with 150 µm thin FE-I4 chips. This process, its development, prototype results, and gain in terms of detector material, are discussed in Chapter 5.

3.2.3 Module concepts

The IBL modules [38] are built using sensors with reduced guard ring structure and FE-I4 chips, both thinned to the minimum achievable thickness. Two module configurations are used (Figure 3.10). Double-Chip modules (MDC) are built using planar n^+-in-n sensors bump bonded to two FE-I4 readout chips, to equip 75 % of the IBL detector. The remaining 25 % uses Single-Chip modules (MSC) built with a double-sided 3D sensor with slim fence and one FE-I4 chip. Thanks to the reduction of the guard ring structure and of the height of the FE periphery, the total active area of an IBL module is 87 %, instead of 65 % of the current modules. 3 % of the inactive area comes from the inactive edge of the sensor, and 10 % from the FE-I4 periphery. The sensor thickness is lower than for the FE-I3 modules, being 200 µm

for the planar sensor and 230 µm for the 3D sensor. The FE-I4 is 150 µm thin. Bare modules are equipped with a flex hybrid made of two Cu layers on a kapton support, with a thickness of 130 µm. The material budget of an IBL module is 0.70 −0.74 % X_0, almost 40 % less than for an ATLAS pixel module, split between

- sensor: 0.25 −0.29 % X_0;
- FE: 0.18 % X_0;
- flex hybrid with passive components: 0.27 % X_0.

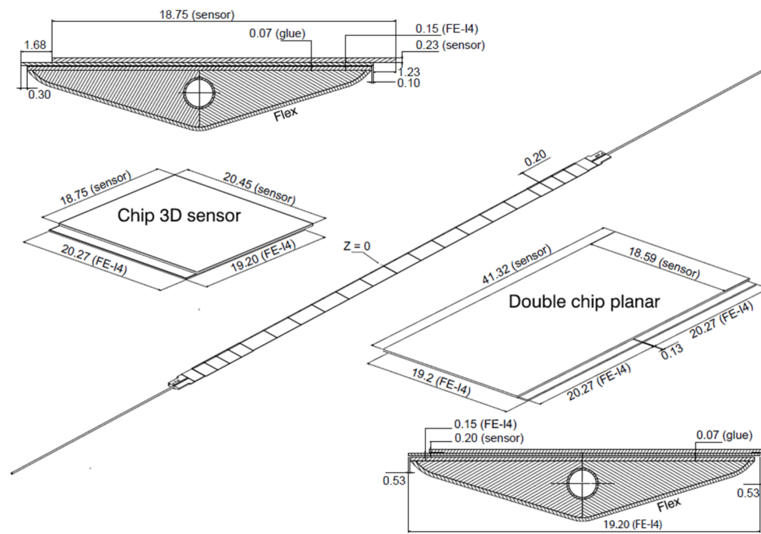


Figure 3.10: Top view of the two IBL module types and cross-section of IBL staves with modules loaded on it. The top view of an IBL stave is also shown, with twelve MDC modules, and eight MSC modules. The modules are placed side by side without overlap, thanks to the slim edge sensor design [38].

For the upgrade of the ATLAS pixel detector at the HL-LHC, mature technologies as the one used for the IBL are further developed to build light, low cost modules for the outer layers. Aggressive thinning of both sensor and electronics is under development with dedicated fabrication processing using handle wafers. The target thickness is ≤150 µm for the sensor and ≤100 µm for the FE chip. Slim edge n^+-in-p planar sensors with inactive edges below 100 µm are the baseline option for the outer layers as they can achieve the desired performance, while keeping cost low. The baseline module configuration for the outer pixel layers at the HL-LHC features one planar sensor connected to four FE-I4 chips, a so-called 4-chip module (Figure 3.11). This module has an active area of 89 %, thanks to the smaller guard ring structure which lowers the contribution of the sensor inactive area to 1 %. It is equipped with a flex hybrid and a pigtail with connector to allow connection to the detector services (Section 3.2.5). The estimated thickness of a 4-chip module is 0.61 % X_0, split between

- sensor: 0.16 % X_0;

- FE: 0.11 % X_0;
- flex: 0.19 % X_0 [11];
- pigtail with connector: 0.15 % X_0.

New interconnection technologies, namely Through Silicon Vias, are investigated that can be applied to the existing hybrid pixel module concept to further reduce its material budget. The usage of TSV, to route the electrical connection to the module on the chip back side, allows to reduce the FE periphery and thickness, and enables a less bulky connection between modules and detector services. Part of the work of this thesis is the validation of a via last TSV process. The process is described in Chapter 6, and its potential for reduction of the material budget of hybrid pixel modules is discussed. A second TSV technology is investigated within the ATLAS pixel collaboration as described in [43].

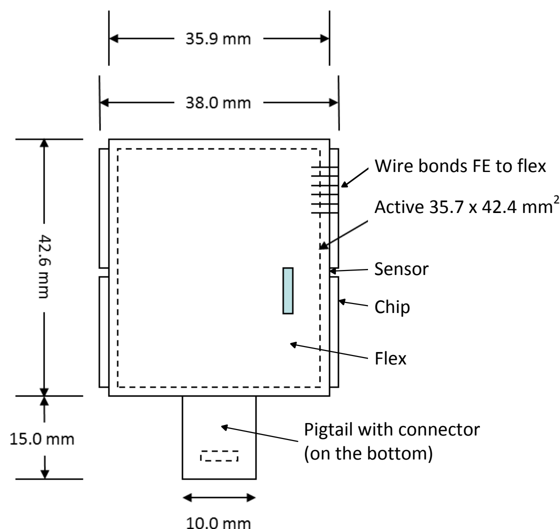


Figure 3.11: Sketch of a 4-chip module for the ATLAS pixel outer layers at the HL-LHC. This conservative design, featuring planar sensor and flex hybrid is chosen as it can be readily produced and used for prototyping activities on sensor, FE, interconnections, as well as detector mechanics and powering (adapted from [11]).

3.2.4 Low mass, efficient powering schemes

As explained in Section 3.1.2, powering schemes for hybrid pixel detectors at LHC are inefficient and massive. At the HL-LHC higher granularity is required to cope with the increased occupancy. Higher current is also expected due to the smaller feature size CMOS process used to design the FE electronics. Independent and parallel powering schemes would result in an efficiency well below 10 %, and in a prohibitive amount of cables in terms of material budget and volume of cables to route through the tight cable channels. New powering concepts have to be investigated which allow to transmit power at low current and high voltage, to achieve

high power efficiency while keeping the material budget low [44]. Two possibilities are considered for the ATLAS pixel detector at the HL-LHC: DC-DC conversion and serial powering (SP) [45].

A powering scheme based on DC-DC conversion is similar to a direct or parallel powering scheme, but with the addition of DC-DC converters on module, as shown in Figure 3.12. Power is distributed at a higher voltage and down converted to the needed supply by the DC-DC converter. Assuming a conversion factor g, the transmitted voltage is gV, where V is the module voltage. The transmitted current is I/g, reducing cable cross-section and the power losses on the cables by a factor g, as listed in Table 3.1.

Serial powering is instead a current based powering scheme, where modules are powered in series by a constant current, as shown in Figure 3.12. The current to voltage conversion is done by regulators on module. The transmitted current is only the current needed by one module, I/n, and the voltage across the chain is nV, where n is the number of modules connected in series. As shown in Table 3.1, depending on the number n of modules powered in series, higher power efficiency and reduced cable volume can be reached. For $n = g$, the same gain is achieved as for DC-DC conversion. However, as the modules are powered with a constant current, the voltage drop V_{drop} is not constrained as in a voltage based powering scheme. It can in principle be chosen only depending on the output voltage capability of the current source and of the allowed power density (i.e. cooling capability). Higher V_{drop} can be allowed in the active area of the detector to reduce the material budget of the cables, while outside the detector the voltage drop can be reduced to lower the power losses. With respect to voltage based powering schemes, serial powering allows more flexibility in the optimization of material and power efficiency. The price to pay is some added complexity at system level.

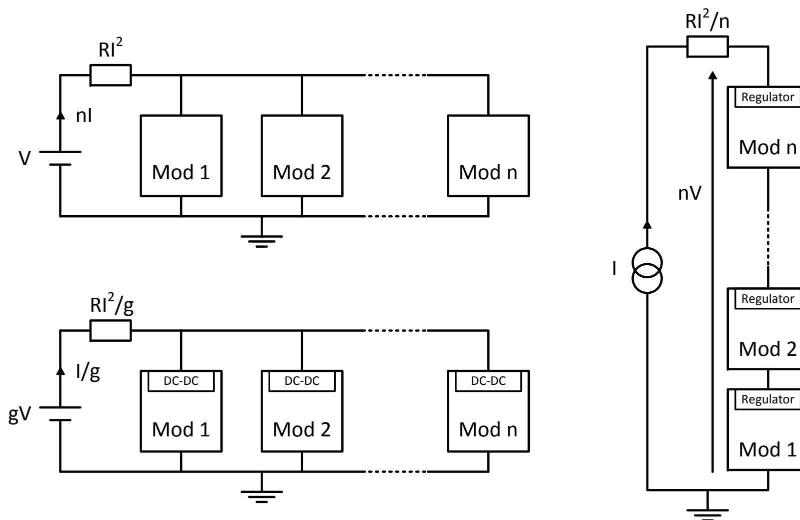


Figure 3.12: Schematic drawings of parallel powering (top left), DC-DC conversion (bottom left) and serial powering (right).

Both approaches require on-stave power conversion. Power converters need to be integrated in the FE to avoid penalty in terms of material. To test both powering options considered for

Table 3.1: Comparison of direct power, DC-DC conversion and serial power. It is assumed that n modules are to be powered. Each module requires a voltage V, and I is the current needed by n modules. For this comparison, the cable cross-section, i.e. resistance, is the same for a given current, for the three powering options.

Power option	Current	Voltage	Cable cross-section (Resistance)	P_{loss}	Power efficiency
Direct	I	V	A (R)	RI^2	$\frac{VI}{VI+RI^2}$
DC-DC	I/g	gV	A/g (gR)	RI^2/g	$\frac{VI}{VI+RI^2/g}$
Serial	I/n	nV	A/n (nR)	RI^2/n	$\frac{VI}{VI+RI^2/n}$

the HL-LHC, the FE-I4A integrates both a ×2 charge pump DC-DC converter, and two Shunt-LDO regulators [46, 47]. The latter is a new regulator concept designed for serially powered detector systems, which combines a shunt and a Low-Dropout (LDO) regulator. The Shunt-LDO working principle is described in detail in Section 4.2.1. To allow studying the different power configurations for the FE, the powering blocks in the FE-I4A are standalone, meaning that they have no internal power or ground connection. External connections are used to select the power configuration for the chip, as shown in Figure 3.13. The FE-I4A can be powered directly, with DC-DC conversion, or with a constant current using the Shunt-LDO regulators.

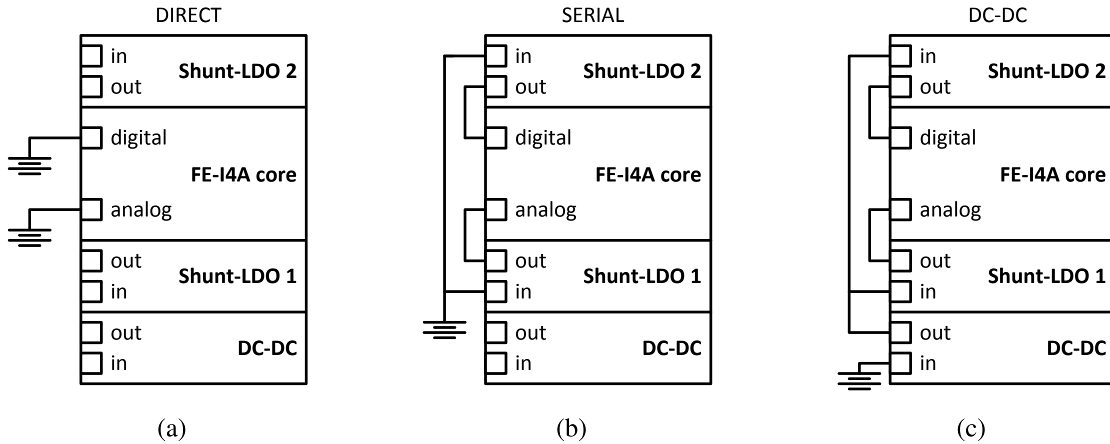


Figure 3.13: Connection of the FE-I4 powering blocks, i.e. DC-DC converter and Shunt-LDO regulators, to power the FE-I4 core, i.e. pixel matrix and peripheral circuitry. Different powering options can be selected externally by means of wire bonds, namely: (a) direct power, (b) serial powering, and (c) DC-DC conversion. In (c) the Shunt-LDO regulators are used in linear mode (Section 4.3.1) to filter the output of the DC-DC converter. The ground connection is not shown.

One focus of the work carried out in this thesis is the development and demonstration of the feasibility of a serial powering scheme for the upgrade of the ATLAS pixel detector HL-LHC, using FE-I4A modules. The SP scheme is described in detail in Chapter 4, where the

advantages in terms of detector thickness and power efficiency are quantified.

3.2.5 Electro-mechanical structures and cooling

Lighter mechanical structures are a key development to reduce the material in the active area of the pixel detector, where they contribute to the largest fraction of the detector material. A new type of all-carbon foam [48] is developed with thermal conductivity two orders of magnitude higher than conventional foams of the same density, and low material budget. This material is chosen for the local support structure of the ATLAS pixel detector upgrades, IBL and Phase-II. The basic concept is to have structures made of a carbon foam core embedded in high stiffness carbon facings to form rigid and lightweight support structures (both stave and disks) for the modules and the electrical services. Cooling is integrated in the structure. Coolant mass and pipe diameter are reduced by using CO_2 [49] instead of C_3F_8. The material of the cooling tube is chosen to be titanium because of its lower material budget.

The IBL is the first pixel detector using this type of structure, and staves are already in fabrication. Figure 3.14 shows a cross-section of the IBL. Similarly to the staves of the current pixel detector, modules are mounted on the flat stave front side, and a cooling pipe runs on the stave backside, embedded in a so-called omega shape part. Modules are placed side by side without the need to overlap thanks to the slim edge sensor design. The on-stave electrical services are in the form of a multilayer cable with Cu traces for data and control signals, and aluminum traces for power distribution. Two multilayer cables serve one stave. The multilayer cables are placed on the stave backside. Connection to the module is done using a wing, part of the multilayer cable. The wing is bent over the stave edge and glued on the flex hybrid. Wing and flex are connected via wire bonds, without the need of a connector. The staves are placed at an angle of 14° with respect to the radial direction, and are ~64 cm long. The thickness of the multilayer stave cable is 0.22 % X_0 [50], split between

- LV lines (Al): 0.07 % X_0;
- HV, data and control lines (Cu): 0.04 % X_0;
- kapton support: 0.08 % X_0;
- wing: 0.03 % X_0.

The material budget for the IBL stave and global support is respectively 0.62 % X_0 and 0.27 % X_0, as quoted in [11].

A box beam stave is prototyped for the outer layers at the HL-LHC. The box beam structure is a flat stave designed to integrate not only cooling but also the on-stave electrical services, as shown in Figure 3.15. Prototypes of box beam staves, 1.4 m long and 4 cm wide, are available with four integrated cables, each designed to serve eight modules. The stave cable is made of a multilayer copper cable for signals and data, and a separate aluminum layer for power distribution. The two layers are connected using tab bonding or ultrasonic soldering techniques. Tabs from the stave cable are glued on the stave surface for connection to the 4-chip modules. 16 modules are mounted on each side of the stave. One of these prototype staves is used for the

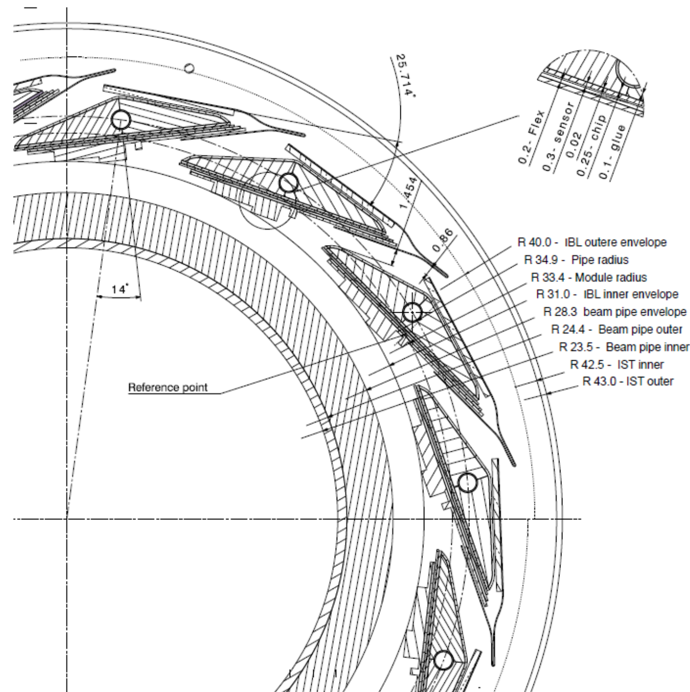


Figure 3.14: Drawing of the IBL layer, showing the cross-section of the staves with modules loaded on it and connected to the services [38].

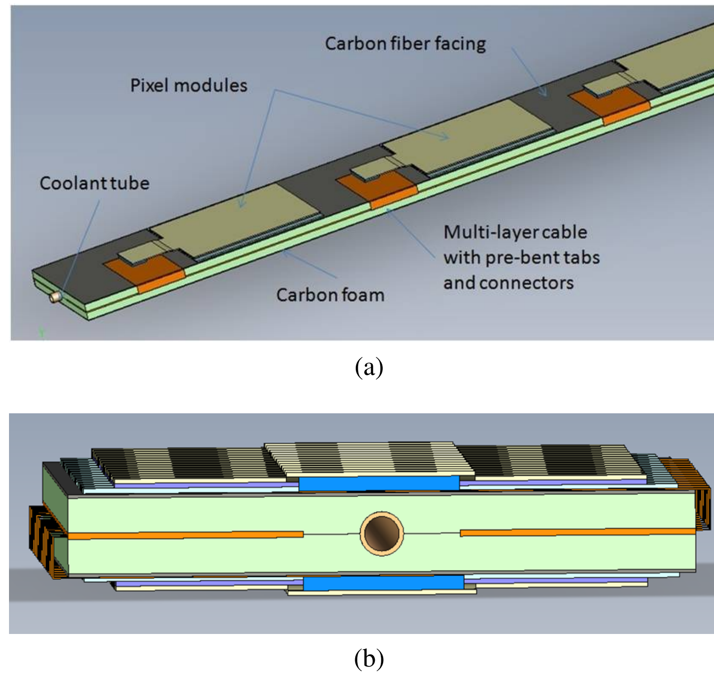
demonstration of a serially powered pixel stave for the HL-LHC presented in Chapter 4. The thickness of the stave cable, assuming serial powering is ~0.09 –0.12 % X_0, split between

- I lines (Al): 0.030 –0.004 % X_0;
- HV, data and control lines (Cu): 0.032 % X_0;
- kapton support: 0.033 % X_0;
- tabs: 0.02 % X_0.

As quoted in [11], the box beam stave has a thickness of 0.64 % X_0, and the global support adds 0.25 % X_0.

3.3 Material budget road map

Table 3.2 summarizes the material budget of the ATLAS pixel detector at the LHC, the IBL, and of the pixel outer layers at the HL-LHC. The development of the hybrid pixel technology using new sensor design, smaller feature size CMOS processes FE electronics, light carbon structures and CO_2 cooling, allowed to build the IBL detector with a thickness of ~1.8 % X_0. For the HL-LHC, these technologies are further developed, and new interconnection technologies and powering schemes are investigated, to build lightweight and cost effective pixel



(a)

(b)

Figure 3.15: (a) Drawing of a section of a box beam stave with modules loaded on it. The connection between modules and multilayer tabs is shown. (b) Drawing of the cross-section of the box beam stave, showing the cooling pipe and the cables integrated in the mechanical structure.

detector modules for the outer layers. An estimate using the mature technology developed for the IBL and assuming SP gives a material budget of $\sim 1.6\,\%\ X_0$ per layer. With respect to the current ATLAS pixel detector, the detector thickness is reduced by about $50-55\,\%$. The target of $\leq 1.5\,\%\ X_0$ is however not yet reached and requires further developments. R&D on thin FE electronics carried out within the work of this thesis, has considerably reduced the IBL detector thickness. Further development of this technology together with an optimized serial powering scheme and TSV could make a significant contribution towards the target material budget for the HL-LHC.

Table 3.2: Material budget per layer of the current ATLAS pixel detector, the IBL and the outer layers at the HL-LHC. The thickness of the different detector elements is given in units of X_0.

		Pixels at LHC [X_0]	IBL [X_0]	HL-LHC outer layers [X_0]
Modules	Sensor	0.32 %	0.25 –0.29 %	0.16 %
	FE	0.27 %	0.18 %	0.11 %
	Hybrids	0.61 %	0.27 %	0.34 %
Services		0.13 %	0.22 %	0.09 –0.12 %
Mechanics	Local support	1.80 %	0.62 %	0.64 %
	Global support	0.50 %	0.27 %	0.25 %
Total		3.63 %	1.81 –1.85 %	1.59 –1.62 %

Chapter 4

Serial powering

4.1 Serial powering for hybrid pixel detectors

As explained in Section 3.2.4, serial powering is a current based powering scheme that allows to overcome the drawbacks of the powering schemes currently used at the LHC [44]. In a serial powering scheme, power is supplied to n modules using a constant current source, I, as shown in Figure 4.1. The current flows from module to module via the series connection, where the ground of each module is connected to the power input of the next one. Regulators on-chip perform current to voltage conversion, to generate the analog and digital voltages needed by the FE electronics. The voltage across the module, V_0, is the voltage at the input of the regulator, and the voltage across the chain is nV_0. The transmitted current set at the supply, I, is the maximum current needed by a module, plus the current needed by the regulator.

With respect to a direct (or parallel) powering scheme, where each detector module is powered with an independent power supply and set of cables, the SP configuration allows a great reduction of costs and number of cables: only one power supply and two cables are needed for n modules. As discussed in Section 3.2.4, depending on the number of modules placed in series, serial powering brings a significant reduction of the transmitted current, and thus of cable cross-section and power losses. This results in both a reduced cable volume, and an increased power efficiency.

In addition to this, serial powering has some peculiar features with respect to voltage based powering schemes. First of all, the allowed V_{drop} on the cables is not constrained, as in a direct powering scheme, to the difference between the module operating voltage and the maximum voltage that the module can withstand. As a consequence, it can be chosen to find the optimal trade off between power efficiency and material, in different parts of the detector. Furthermore, voltage regulation is done close to the load. No sense lines are needed to regulate the voltage at the module input, and the voltage regulation can respond faster to changes in the load current. Also, as the current flow to the modules is constant, no transients are expected at the input of the regulator.

In serial powering, the power dissipation of the modules is on average higher than in a direct powering scheme: on top of the power consumed by the regulator, the current flowing to the module needs to be constantly equal to the maximum current needed by the FE chips. This larger module power consumption is however compensated by the very large reduction of power losses on the cables, so that the total detector power consumption decreases by a significant factor. In addition to this, the power overhead of SP modules does not translate

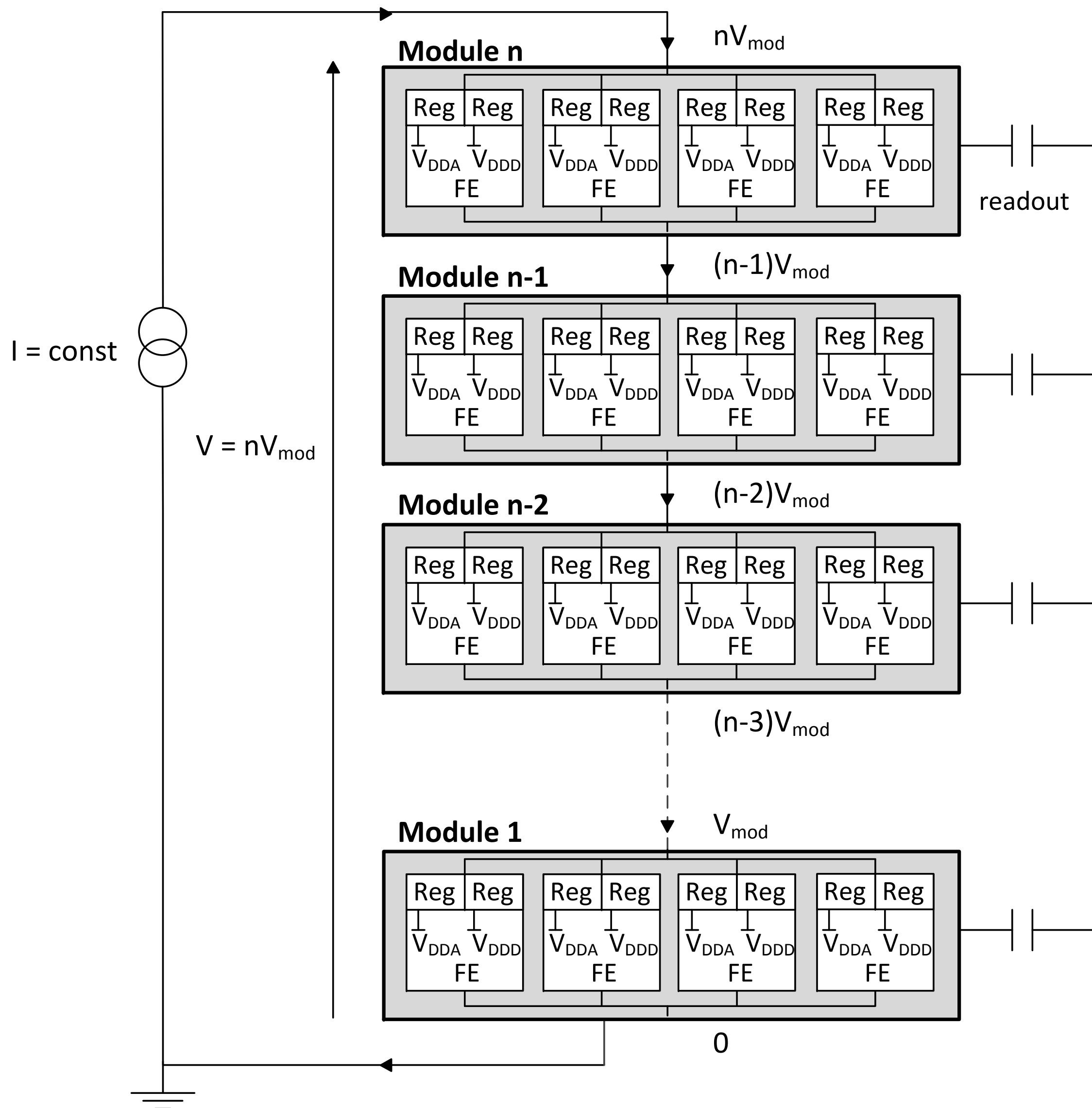


Figure 4.1: Sketch of a SP chain, showing its main features. The modules are connected in series and powered via a constant current I. The voltage across the chain is n times the voltage across one module V_0. The current splits on module between the parallelly connected regulators on the FE chips, which generate the analog and digital voltages, respectively $VDDA$ and $VDDD$. The data communication to and from the modules is AC-coupled.

into different specifications for the detector design. The detector services, namely cables and cooling, are in fact designed for the maximum module power consumption. As the module power consumption in SP is constant, and does not depend on the particular FE configuration, the cooling power does not need to be tuned during detector operation, which simplifies the design of the cooling infrastructure. The reduction of the power losses on the cables is as well beneficial for the cooling system, as warm cables running close to the cooling pipes in the tight service channels, can bring the liquid to its boiling point and decrease the cooling efficiency.

These features lead to a great reduction in cable material (smaller number of cables with smaller cross-section), while increasing the efficiency (smaller resistive losses on the cables), and reducing the costs (reduced total detector power consumption and reduced number of power supply channels). However due to the serial connection of the modules, SP brings some risks and complications at system level. A fault on the power line affects the entire chain. Faulty or noisy modules could potentially influence the performance of other modules in the chain. Modules cannot be powered on/off individually, and all modules, apart from the first one in the chain, are on different ground potentials with respect to the power supplies and readout system. To overcome these drawbacks safety measures and dedicated detector system developments are needed.

The number of modules in a SP chain could in principle be chosen only as a function of the output voltage capability of the current source and of the voltage ratings of the power cables. In practice this number should be chosen as a trade off between material reduction, efficiency and reliability of the chain. Redundancy should be added in the connection of the modules to the power line. Proper filtering and regulation of the supply voltage can be used to filter disturbances on the power line [51]. At module level, the current to voltage conversion should be done using more than one regulator, and by connecting all regulators in parallel. In this way, should one regulator fail, the current flow can still be assured by the other regulators on module. Although these measures assure a very robust design, extra care has to be taken for possible worst case failures, in particular for the case of regulator faults which could cause an over-voltage. This could be addressed using a protection element featuring real time response, i.e. over-voltage protection. This element could as well integrate a slow control feature to allow control over single modules as in the traditionally used powering schemes. Data communication to and from the modules needs AC-coupled data links, and the HV distribution to power the sensors should as well be designed in order to avoid shorting the SP chain.

A SP scheme as described here is demonstrated using modules from the ATLAS pixel detector, as documented in [52, 53, 54, 55]. This R&D culminates in the so-called "Serial Powering Proof of Principle" described in [51]: half-stave of the ATLAS pixel detector, equipped with six FE-I3 modules, is powered in series. This work proves that serial powering is a feasible and reliable scheme to power large area hybrid pixel detectors for HEP experiments, with significant advantages in terms of power efficiency and material budget. Based on the experience gained with the serial powering of FE-I3 modules, an optimized and improved current based power distribution scheme is developed in the framework of this thesis for the next generation hybrid pixel detectors at the HL-LHC [45]. In particular, for the Phase-II upgrade of the ATLAS pixel detector, a dedicated regulator concept, the Shunt-LDO [46], is developed, as well as a bypass scheme. The latter allows to tackle possible failures leading to the loss of an entire chain, and to regain control of single modules as in a direct powering scheme. AC-

coupled data transmission is simplified. These developments are described in Section 4.2. The work presented in this chapter validates all elements of the proposed serial powering scheme, via

- characterization of the Shunt-LDO regulator in the FE-I4A chip (Section 4.3);
- study of the performance of serially powered FE-I4A modules (Section 4.4);
- operation of a serially powered pixel stave, featuring MDC modules from the IBL prototyping with a dedicated flex hybrid, and a prototype box beam stave (Section 4.5);
- demonstration of a module control and protection scheme (Section 4.6).

4.1.1 Serial powering with the FE-I4 readout chip

The serial powering scheme for the ATLAS pixel detector at the HL-LHC presented in this chapter, is based on the pixel outer layer concept described in Section 3.2.5. According to the design of the box beam stave layout, where four cables serve the 32 modules on stave, SP chains of 8 modules are proposed. On module, each FE-I4 chip, integrates two Shunt-LDO regulators to generate the voltages needed by the FE ($VDDA$ = 1.5 V and $VDDD$ = 1.2 V). All the Shunt-LDO regulators on module are operated in parallel. One bypass element is connected in parallel to each module and operated with one line. The advantages of this SP scheme in terms of material budget and power efficiency can be estimated by comparing it to a possible direct powering scheme for the same detector concept.

Table 4.1 summarizes the values of voltages and current, as well as the geometrical parameters of the outermost layers used for the calculations [56]. For the direct powering scheme, it is necessary to define a value for the maximum allowed voltage drop on the power cables. For the cables until PP2 (Type-0, Type-I and Type-II), the assumed voltage drop is the one specified for the design of the IBL powering scheme [57]. The power distribution for the IBL layer is in fact organized in groups of four chips, thus having the same modularity as a direct powering scheme for the outermost layers at the HL-LHC using 4-chip modules. The voltage drop on the Type-IV cables, which are outside the ATLAS detector area, is kept as in the present detector.

As explained in Section 4.1, the V_{drop} in a serial powering scheme can be defined to optimize material budget and power efficiency. For instance, in the detector active area, the voltage drop could be increased to reduce the cable cross-section, and thus the cable material budget. Outside of the detector, where the material budget does not play a role anymore, the voltage drop could be decreased to increase the power efficiency. For the sake of comparison, and to stress the huge potential of SP for material reduction added by the flexibility in the choice of V_{drop}, the calculations are done assuming both the same allowed voltage drop as for the direct powering scheme, and an optimized one. The regulator efficiency is set to 80 %.

Table 4.2 shows the results of the calculations. The length of the on-stave power lines (Al), is reduced by a factor 4.5 and, for the same voltage drop as in direct power, their cross-section, A, is reduced by a factor 8, corresponding to the number of modules placed in series. The material budget of the power lines is reduced from 0.120 % X_0 to 0.030 % X_0. One more control line per module is needed for the protection element, increasing the Cu thickness. However only

Table 4.1: Values used to compare serial and direct powering schemes for an outer layer of the ATLAS pixel detector at the HL-LHC.

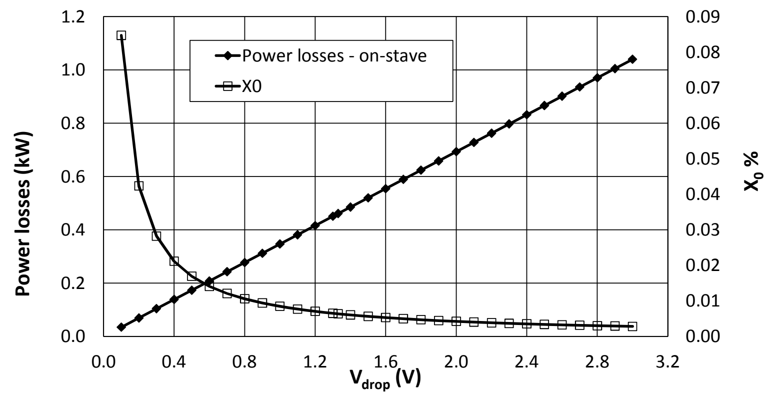
Parameter	Value
VDDA	1.5 V
VDDD	1.2 V
I_{mod}	4 ×0.56 A
V_{drop} - Type-0	0.3 V (roundtrip for $I=I_{mod}$)
V_{drop} - Type-I & Type-II	0.7 V (roundtrip for $I=I_{mod}$)
V_{drop} - Type-IV	2.5 V (roundtrip for $I=I_{mod}$)
Shunt-LDO efficiency	80 %
SP unit	8 modules
# of staves/layer	38
layer radius	20.9 cm
layer length	140 cm

one layer of aluminum for the serial power distribution is needed, instead of two as for direct power, so that the thickness of the kapton support is lower for SP. The total thickness of the stave cable scales by a factor 1.8.

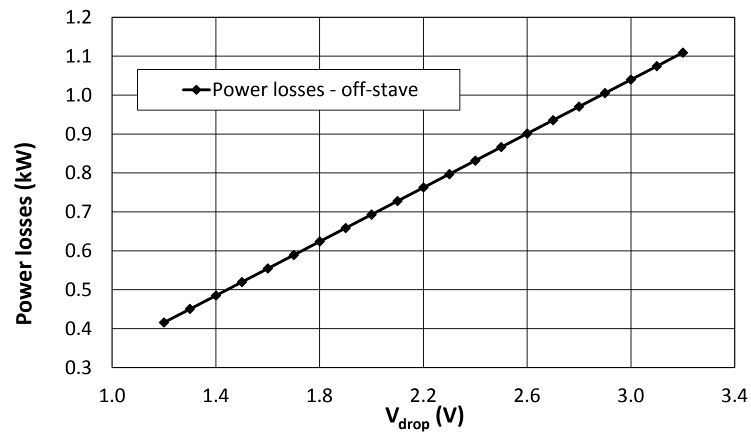
The power losses on the cables are 8 times smaller, both on-stave and off-stave. Although the power consumption of the serially powered modules increases by a factor 1.2, the total on-stave power dissipation stays the same. This results in a total detector power consumption 2.3 times smaller. The power efficiency is 29 % for the direct powering scheme, while it reaches 65 % for serial powering.

Optimizing V_{drop} on the on-stave cable, the material budget can be further reduced, as shown in Figure 4.2a. Also shown is the corresponding increase of the power losses on the cables. Considering for instance the same on-stave power losses as in the direct powering scheme, i.e. 0.8 kW, the material budget of the power cables can be reduced by a factor as high as 30 such that the total cable thickness decreases by a factor 2.3. This translates into an on-stave power consumption increase by 20 %. The total detector power can be kept the same by reducing the V_{drop}, and thus the power losses, on the rest of the cable chain (Figure 4.2b). In the calculation shown here, the V_{drop} on the Type-IV cables is reduced to 0.5 V.

It should be mentioned here for completeness that the extra circuitry connected to serial powering, namely the protection element and the capacitors for the AC-coupled data transmission, would add only 0.006 % X_0 per layer.



(a)



(b)

Figure 4.2: (a) Power losses and material budget of the on-stave power cable as a function of the voltage drop for a SP scheme. (b) Power losses of the off-stave power cables as a function of the voltage drop for a SP scheme.

Table 4.2: Comparison between a direct and a serial powering scheme for the upgrade of the ATLAS pixel detector at the HL-LHC. The values are given for an outer layer. The material budget of the on-stave power cable scales by a factor 4 considering only the current reduction. By optimizing as well the voltage drop on the cables, a reduction by a factor 30 can be achieved. The thickness of the stave cable scales correspondingly by a factor 1.8 and 2.3. The power efficiency of the serial powering scheme is 65 %, 2.2 times higher than for a direct powering scheme.

		Direct power	Serial power	Factor	Serial power optimization	Factor
On-stave cable						
Al length		821 m	182 m	4.5	182 m	4.5
Al cross-section		302 mm^2	38 mm^2	8	5 mm^2	60
	Al	0.120 % X_0	0.030 % X_0	**4**	0.004 % X_0	**30**
	Cu	0.029 % X_0	0.032 % X_0		0.032 % X_0	
Thickness	Kapton	0.036 % X_0	0.033 % X_0		0.033 % X_0	
	Tabs	0.020 % X_0	0.020 % X_0		0.020 % X_0	
	Total	0.205 % X_0	0.115 % X_0	**1.8**	0.089 % X_0	**2.3**
Power losses						
on-stave		0.82 kW	0.10 kW	8	0.80 kW	1
off-stave		8.72 kW	1.11 kW	8	0.42 kW	21
Module power						
		3.82 kW	4.71 kW	0.8	4.71 kW	0.8
Total power						
on-stave		4.64 kW	4.81 kW	1.0	5.51 kW	1.2
detector		13.46 kW	5.92 kW	2.3	5.93 kW	2.3
Power efficiency						
		29%	**65%**	2.2	**65%**	2.2

4.2 Serial powering building blocks

4.2.1 Voltage regulation

As mentioned in Section 4.1, in a SP scheme, on-chip regulators are needed that can operate in parallel, generate different output voltages out of the current supply, and shunt additional current in case of device failure. As shown with FE-I3 modules [51], this can be achieved using a combination of shunt and linear regulators.

Shunt regulators are used to convert a current into a constant voltage. They work similarly to a zener diode in reverse bias. Figure 4.3 shows the typical input characteristics of a shunt regulator. The important properties for a shunt regulator are the input resistance, R_{in}, and the threshold voltage, V_{thres}. R_{in} defines the slope of the I-V characteristics. It is ideally chosen to be small so that, when enough current flows into the regulator, a constant voltage equal to V_{thres} is present at the regulator output. The current which is not taken by the load is shunted

by the regulator.

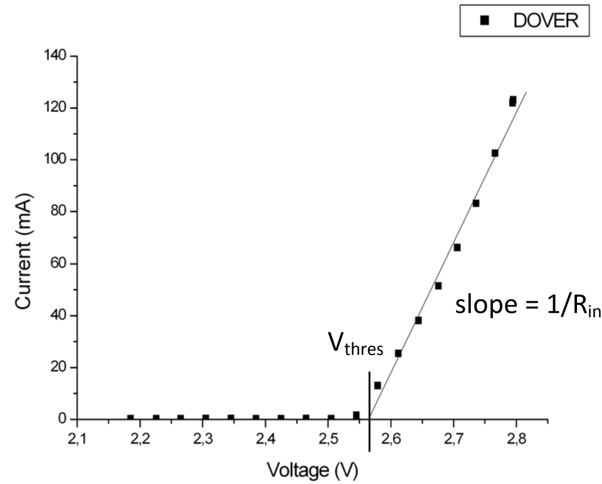


Figure 4.3: I-V characteristics of a shunt regulator integrated in the FE-I3 chip (DOVER). This regulator has $V_{thres} = 2.58$ V and $R_{in} = 1.9\,\Omega$ [51].

A linear regulator can be explained as a variable resistor whose value changes to keep a constant output voltage V_{out} independently of the input voltage V_{in} and the load current I_L. Two important parameters for this type of regulator are thus the line and load regulation, which describe respectively the dependency of the output voltage on variations of V_{in} and of I_L. A minimum difference between V_{in} and V_{out} is needed for regulation. This is defined as the drop-out voltage, $V_{dropout} = V_{in} - V_{out}$. When $V_{dropout}$ is in the order of $0.1 - 0.2$ V, the regulator is called Low-Dropout regulator (LDO). The use of LDO regulators is preferred as they provide higher power efficiency.

Different configurations are possible for the shunt and linear regulators on-chip to generate the FE analog and digital voltages, respectively VDDA and VDDD, as demonstrated with FE-I3 modules [51]. As shown in Figure 4.4a, one shunt regulator per FE is used to convert the input current to a constant voltage. Two linear regulators can then be used to generate the individual levels for $VDDA$ and $VDDD$ out of the output of the shunt. An alternative configuration is to use the shunt regulator to generate the higher voltage needed by the FE, and one linear regulator to derive the second, lower, FE voltage (Figure 4.4b). In both cases, the output voltage of the shunt regulator has to be larger than the output voltage of the LDO regulators plus the $V_{dropout}$.

To add redundancy to the SP chain, all shunt regulators on module are operated in parallel. Parallel connection is beneficial for voltage regulation when using shunt regulators as the input resistances are connected in parallel, lowering the total resistance and improving the voltage stabilization. However, this connection can be critical. Due to mismatch and process variation, the shunt regulators placed in parallel can have different V_{thres}. As the input characteristics of shunt regulators is very steep, the regulator with lower V_{thres} can take all current at start-up and burn. This could eventually lead to a chain reaction and destroy all regulators on module. Safe parallel operation of shunt regulators thus requires to choose the value of the input series resistance according to the spread in threshold voltage values, in order to mitigate

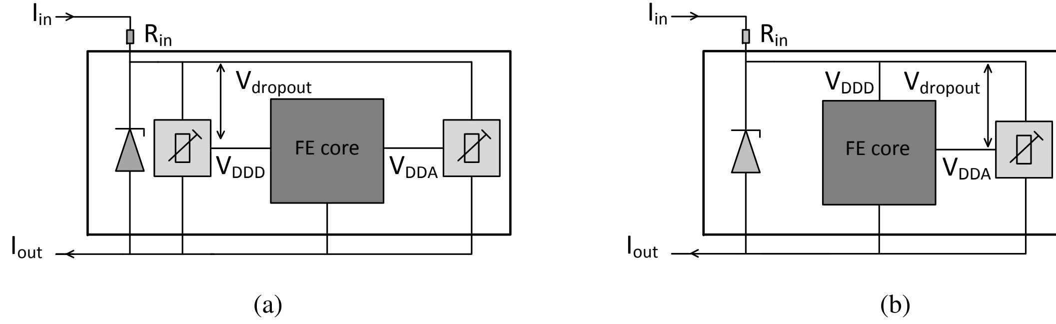


Figure 4.4: Schematic drawings of on-chip shunt and LDO regulators, represented respectively as a zener diode and variable resistors, to convert the input current into two stable voltages. These regulator configurations have been used with FE-I3 modules to generate $VDDD = 2.0\,\text{V}$ and $VDDA = 1.6\,\text{V}$.

the I-V characteristics. For the shunt regulators in FE-I3 the input resistance is chosen in a range between $1.5 - 2.5\,\Omega$ (Figure 4.3). Since the use of this resistance lowers the power efficiency and decreases the voltage stability, a small spread across shunt parameters is desired. This drawback can be overcome by combining the shunt and LDO regulators in a different order, as sketched in Figure 4.5: the LDO serves as the input resistor of the shunt regulator, and the shunt transistor becomes part of the load of the LDO. This combination results in a new regulator concept, the Shunt-LDO regulator, developed specifically to match the needs of serially powered pixel detector systems [46].

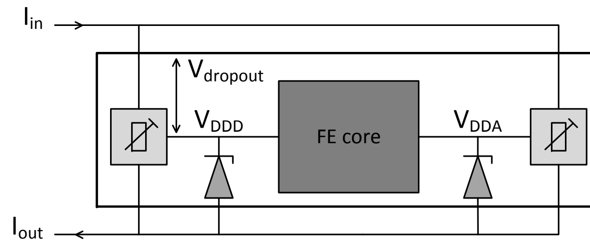


Figure 4.5: Schematic drawing of on-chip current to voltage conversion using the Shunt-LDO principle.

A simplified schematics of the Shunt-LDO regulator is shown in Figure 4.6. The LDO regulator is formed by the error amplifier $A1$, the PMOS pass transistor $M1$, and the resistive divider formed by $R1$ and $R2$. It enforces a voltage regulation loop, responsible to generate a constant output voltage according to

$$V_{out} = 2 \cdot V_{ref} \tag{4.1}$$

where V_{ref} is the reference voltage. The current regulation loop is enforced by the shunt transistor $M4$, the current mirror formed by $M1$ and $M2$, and the differential amplifier $A3$. The amplifier $A2$ and the cascode $M3$ are added to improve the mirroring accuracy. The current regulation loop is responsible to keep the current through the regulator constant, independently

of the load. This is achieved by shunting different amounts of current in $M4$. To properly steer the shunt transistor, two currents are compared in $A3$. A fraction of the input current, defined by the aspect ratio k of the current mirror, is mirrored by $M1$ and $M2$. A reference current, I_{ref}, is set by the resistor $R3$, which depends on the input potential and thus, in turn, on the input current. Should the mirrored current be smaller than I_{ref}, then $M4$ is steered to draw more current and vice versa. The current flowing through the regulator can thus be expressed as

$$I_{in} \approx kI_{ref} \approx k\frac{V_{in} - V_{thM6}}{R3} \tag{4.2}$$

where V_{thM6} is the threshold voltage of transistor M6. The input characteristics of the Shunt-LDO is defined by $R3$ as follows

$$R_{in} \approx \frac{V_{in}}{I_{in}} \approx \frac{R3}{k}. \tag{4.3}$$

From (4.3) it can be inferred that, opposite to a traditional shunt regulator, the Shunt-LDO has an ohmic I-V characteristics, i.e. the regulator behaves as a resistor for the input current. This allows safe operation of Shunt-LDO regulators connected in parallel, even for different values of R_{in}, i.e. different $R3$, as the input current splits between the parallelly placed regulators in the same way as it would split between parallelly connected resistors. This makes the design more robust against process variation and mismatch. Differences in the value of $R3$ lead in fact to different shunt current values but do not destroy the regulator. In addition to this, Shunt-LDO regulators can be placed in parallel even if they generate different output voltages: the difference between their output voltages is compensated by the $V_{dropout}$ across the pass transistor of the LDO regulator, $M1$. Finally, should one of the parallelly placed regulators fail, the extra current can be shunted by the other regulators.

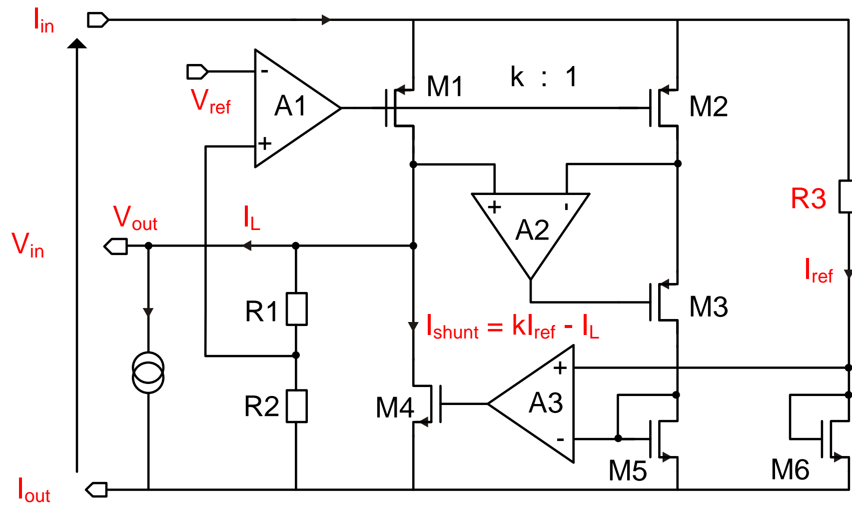


Figure 4.6: Simplified schematics of a Shunt-LDO regulator, explained in the text [46].

The Shunt-LDO regulator can be operated as a standard LDO by switching off the current regulation loop. A third working mode is also possible, called partial shunt mode, in which the

regulator is operated in LDO mode but with enabled current regulation loop. The value of $R3$ is set in order to have a minimum current flow through the regulator in absence of a load. As explained in Section 4.3.1 this configuration can be used to reduce transients on the power line in voltage based powering schemes. The different working modes of the Shunt-LDO regulator can be selected, as summarized in Table 4.3. In particular, the device can be configured in Shunt-LDO mode for application in serially powered systems (current based), or in LDO and partial shunt mode for use in conventional voltage based supply schemes.

Table 4.3: Working modes of the Shunt-LDO regulator. The selection of the working mode is done by choosing either a constant voltage or current supply for the regulator, and by enabling or not the current regulation loop.

	Shunt-LDO mode	LDO mode	Partial shunt mode
Current regulation loop	enabled	disabled	enabled
Regulator input	I	V	V

A complete description of the Shunt-LDO working principle can be found in [46], together with results from simulations and characterization of the prototypes. This work presents results from the Shunt-LDO regulator version integrated in the FE-I4A pixel readout chip in Section 4.3 [47].

4.2.2 SP chain protection (module bypass system)

The main concerns connected to the use of serial powering are the loss of the entire chain of modules in case of a fault on the power line or on one of the modules, and the propagation of a disturbance on the power line. Also, in a serial powering chain, modules cannot be switched on/off individually as in a direct powering scheme. In addition to the redundant connection of modules to the power line, the use of regulators connected in parallel, and voltage filtering, an element is needed to bypass the modules. In this way, faulty or noisy modules can be excluded from the chain by routing the current via the bypass.

The bypass element should have both a slow control feature, which allows switching on/off selected modules, and fast response capability for over-voltage across the module. In a current based scheme over-voltage could occur in case of failure of one or more regulators on the module. In this case, the increased current flowing through the remaining Shunt-LDO regulators could lead to an input voltage higher than the regulator maximum V_{in}. Over-voltage cannot happen as in a voltage based powering scheme as the transmitted current is constant. The protection element should draw no power when not used, and generate a voltage lower than 0.1 V when shunting the entire module current, to not dissipate a considerable amount of power (i.e. heat). The reaction time of the over-voltage protection is specified to be within 100 ns. One bypass element must be connected in parallel to each module, referenced to the local module ground. The bypass element should thus be radiation hard and add as little material as possible. To satisfy these requirements, the bypass should be implemented as a chip in sub-micron CMOS technology, possibly as a one-wire device.

The first implementation of such bypass chip is the so-called Serial Power Protection (SPP) chip [58]. The SPP has been designed as a prototype for the SP voltage regulation and bypassing scheme for the upgrade of the ATLAS strip system for the HL-LHC [59, 60]. On top of the slow control and fast response features, it also provides voltage regulation for the strip modules. It integrates shunt transistors of different sizes, as well as different comparators to sense the voltage across the module. Shunt transistors and comparators can be used in different configurations to activate the regulating and the shunting functionality. The latter can be enabled either upon command or autonomously upon over-voltage. The threshold for over-voltage can be set in a range between 1.3 V and 2.4 V. The SPP is designed as a one-wire device over which power and data can be delivered to the chip. Power can be distributed in parallel to all SPP chips in the serial powering chain, with different bias resistors, by referring the ground of each SPP to the local module ground. Similarly, all SPP chips in a SP chain can be operated with one AC-coupled data line, using an addressing scheme implemented in the chip. Figure 4.7 shows the connection of the SPP to a SP chain of modules. The slow control and over-voltage blocks of the SPP are tested as a proof of concept with serially powered FE-I4 modules. The results are shown in Section 4.6.

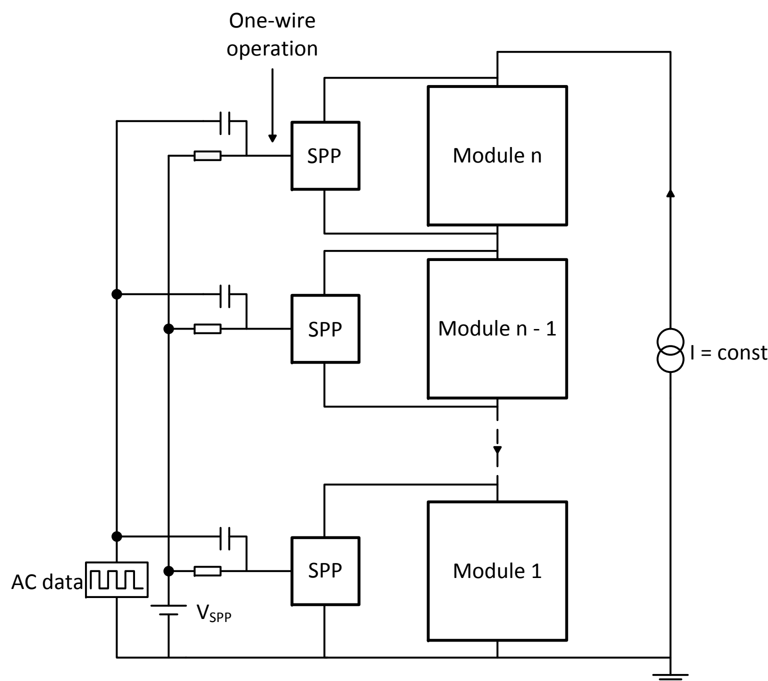


Figure 4.7: Sketch of a SP chain with bypass scheme, implemented using the SPP chip. Power (V_{SPP}) and data (*AC data*) are sent to the chip over one line only. The SPP ground is referenced to the local module ground. To sense and control the voltage across the module, the SPP is connected between the input and output current path, i.e. in parallel to the regulators.

4.2.3 AC-coupled LVDS data transmission

In a serial powering chain, each module is referenced to a local ground potential. Only the first module in the chain is connected to the system ground, where the current supply and the readout system are connected. This configuration introduces an issue for the data transmission. Clock (CLK) and command (CMD) signals to the FEs, as well as data (DO) from the FEs, are sent over LVDS (Low Voltage Differential Signaling) links between the readout and the modules. Differences in the ground potential of transmitter and receiver introduce a shift of the common mode voltage V_{cm}, i.e. of the DC potential of the lines in the LVDS pair. If the V_{cm} moves out of the common mode voltage range at the receiver input, the link does not work properly anymore. The solution to this problem is to use AC-coupled LVDS links.

AC-coupling is a termination scheme which consists in adding capacitors in series with both signals in a differential pair of the LVDS link, as shown in Figure 4.8. As the capacitors block the DC component of the signal over the differential pair, an AC-coupled link helps in guarding against ground differences between transmitter and receiver. The common mode voltage at the receiver input is set by the receiver itself and not by the distant transmitter. An AC-coupled link thus requires self biased receiver inputs to set the V_{cm}. As the current from the transmitter flows through the capacitors, across the termination resistor, only during transitions, frequent transitions have to be guaranteed. This is achieved with DC-balanced signals, where DC-balance refers to the difference between the number of ones and zeros in a set of data. DC-balance is achieved, for the entire data set, for an equal number of zeros and ones. The clock is for instance a perfectly DC-balanced signal. The instantaneous number of consecutive equal bits in a DC-balanced data stream is called the Run Length (RL). The maximum value of RL is particularly important as it defines the lowest frequency component of the signal. As the latter should not be filtered out by the high pass filter made by the AC-coupling termination, the capacitor value has to be chosen according to

$$C = \frac{7.8 \cdot RL \cdot T}{R_{term}} \tag{4.4}$$

where T is the bit period and R_{term} the termination resistor [61]. The capacitor value given by (4.4) ensures a signal attenuation of only 0.25 dB, i.e. 3 %.

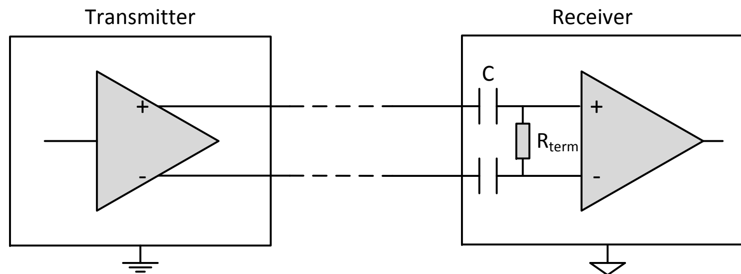


Figure 4.8: Sketch of an AC-coupled LVDS link. Capacitors are placed in series on each line of the differential pair, in front of the termination resistor. This configuration is used with serially powered FE-I4 modules.

In view of the SP implementation, the FE-I4 receiver is designed with a self biasing circuitry.

Biasing circuitry can also be added externally to the inputs of the receiver on the readout side. The data from the FE-I4 is DC-balanced as it uses an 8B/10B encoding protocol, with $RL = 5$ [62]. However, the CMD received by the FE-I4 is not DC-balanced. A possible solution to AC-couple the command link would be to use a link with feedback as for the FE-I3 modules [51]. In an effort to simplify the SP scheme and to minimize the added material, the CMD protocol of the FE-I4 is Manchester encoded to gain DC-balance, and thus use a simple AC-coupling link as in Figure 4.8. The encoding of the CMD protocol is implemented in the firmware of the readout system. Each bit of the CMD stream is sent twice, where the second time the bit is sent inverted, at 80 MHz, i.e. double the sampling speed. The receiver samples the incoming stream at 40 MHz. The sampling clock is aligned with the data so that it samples only the bits of the CMD, and not their inverted counterpart. In this way a DC-balanced signal is obtained which can be understood by the receiver. This scheme is implemented and tested with the serially power modules and discussed in Section 4.5.

4.2.4 HV distribution

Distribution of the HV to bias the sensors in a SP chain has to be done according to the current distribution scheme, to avoid shorting the SP chain. The most straightforward solution would be to use one supply per module, which is however inconvenient due to the large number of modules to power. An easy alternative is to group all modules in a SP chain to one HV power supply, and connect the HV return line to the local grounds of the modules. This could be implemented as shown in Figure 4.9. Between the ground of two modules a high ohmic connection prevents shorting the SP chain. However, due to the particular connection of the HV ground line locally, each module sees a higher bias voltage with respect to the one of the previous module, the difference being the module voltage V_{mod}. The HV value seen by module n is

$$HV_n = HV_{supply} + (n-1)V_{mod} \tag{4.5}$$

where HV_{supply} is the value at the power supply. The total potential difference of the modules in the chain has to be taken into account, to avoid biasing sensors at the top of the chain with a voltage higher than the breakdown voltage, V_{BD}. This scheme can only be implemented if $nV_{mod} < V_{BD} - HV$. For a chain of 8 FE-I4 modules, $nV_{mod} \leq 20$ V.

This is not a concern for planar sensors, both n^+-in-n and n^+-in-p. For the IBL planar sensors for instance, the depletion is reached at −50 V, the operational voltage is −80 V and $V_{BD} - HV \geq$ 50 V. After irradiation $HV = 1000$ V and breakdown does not play a role anymore. Problems can arise with 3D sensors, where the breakdown occurs at a much lower voltage. The 3D sensors for the IBL for instance are operated at −20 V and breakdown is expected already at −30 V. This might require to have more than one HV group per SP unit.

This solution is suited for prototyping and testing activities, and it is implemented for the serial powering pixel modules and stave prototype discussed in Section 4.4 and Section 4.5, respectively. A more sophisticated scheme should however be developed for detector operation. The discussed HV distribution scheme in fact does not allow independent measurement of the sensor leakage current. For this a HV switch is needed to bias sensors in the SP chain

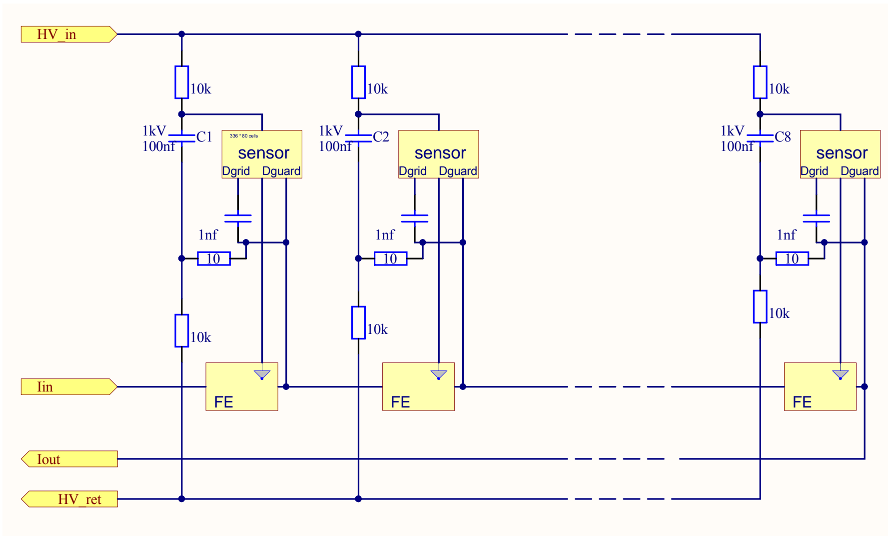


Figure 4.9: Schematics of the *HV* distribution scheme for a serial powering chain. A resistor of 10 kΩ connects the *HV* return line to the local module ground. The other resistors and capacitors shown represent the *HV* filtering scheme.

independently, in order to measure the I-V curve of each module. This measurement is in fact routinely used during detector operation to monitor the radiation damage of the sensor.

4.3 Shunt-LDO regulator in FE-I4A

The Shunt-LDO regulator version in FE-I4A is designed to match the power requirements of the chip. The regulator input can either be a current as high as 0.6 A, or a voltage up to 2.5 V. This voltage rating is possible as every branch of the design is cascoded. The design value of the current mirror aspect ratio k is 1000. The voltage at the output of the regulator is in a range of $1.2-1.5$ V. The reference voltage to define the regulator's output has to be provided externally in this chip version[1], and a 100 nF capacitor to ground is used for filtering. The reference resistor $R3$ can either be internal, with a fixed value of 2 kΩ, or be placed externally to choose a different value. Biasing circuitry for the amplifiers is provided internally, but the biasing node of amplifier $A3$ is available on a pad. By connecting it to ground, the current regulation loop can be switched off.

Two Shunt-LDO regulators are integrated in the FE-I4A chip. No hard-wired connection is done on-chip between the regulators and the FE. Connection of the input, output, and ground ports of the two regulators is done externally via wire bonds, as well as the selection of the working mode. To power the FE-I4, the regulators are connected in parallel, i.e. sharing input and ground connection, as shown in Figure 4.10. The outputs of the regulators are connected to the analog and digital power rails of the chip. The regulators are referred to in the following as analog and digital regulator, and their voltages and currents are labeled with a subscript A and D.

[1] In the FE-I4B the reference voltage generation is integrated in the FE [57].

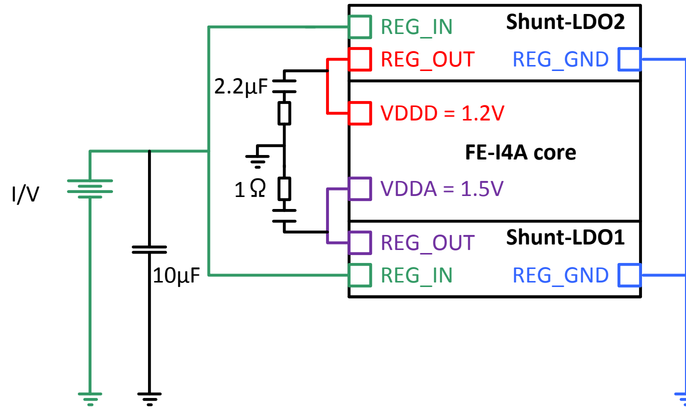


Figure 4.10: Connection of the Shunt-LDO regulators to power the FE-I4 chip. Input and output filters are also shown. The reference voltage input and the selection of the working mode are not shown.

4.3.1 Characterization

For characterization, the regulators are operated either stand-alone or in parallel, using an external load, in all three working modes. For these tests, chips are mounted on test boards. Wire bonds from the FE pads to measurement points are used to measure the voltages at the input and output of the regulators. In this way, the voltage drop across the wire bonds and PCB traces is excluded. A current sink is used as load to the regulator. The reference voltage is provided using a voltage supply. The external connection of the reference voltage, requires a correction of the measured regulator output voltage for the Shunt-LDO mode of operation. In this working mode, part of the input current is shunted through the shunt transistor ($M4$). This current flows to the regulator ground and from here, via wire bonds and PCB traces to the board ground, as shown in Figure 4.11. Due to the resistance of this path (R_{gnd}), the regulator ground (GND_{REG}) and the board ground (GND_{PBC}), to which the supplied reference voltage, V_{ref_ext}, is referenced, are not the same. As a consequence, the effective reference voltage seen by the Shunt-LDO regulator (V_{ref}) is lower than the one set at the supply (V_{ref_ext}). By measuring the ground shift across R_{gnd}, the V_{out} can be corrected according to (4.6)

$$V_{out-corr} = 2 \cdot V_{ref_ext} = 2 \cdot (V_{ref} + R_{gnd} I_{shunt}) = V_{out} + 2 R_{gnd} I_{shunt}. \tag{4.6}$$

Only the corrected output voltage value is shown in the following.

The characterization of the Shunt-LDO is performed on a limited number of samples, namely ten regulators. The error on the measurements is negligible, being less than 0.03 % of the measured value. The mean and RMS values of the regulator parameters are obtained from the measured regulator characteristics on the available samples. They are in agreement with the expected spread of parameters at wafer level, as well as with the ones obtained from the measurements at wafer level of the regulators in the FE-I3 [51], and thus considered significant.

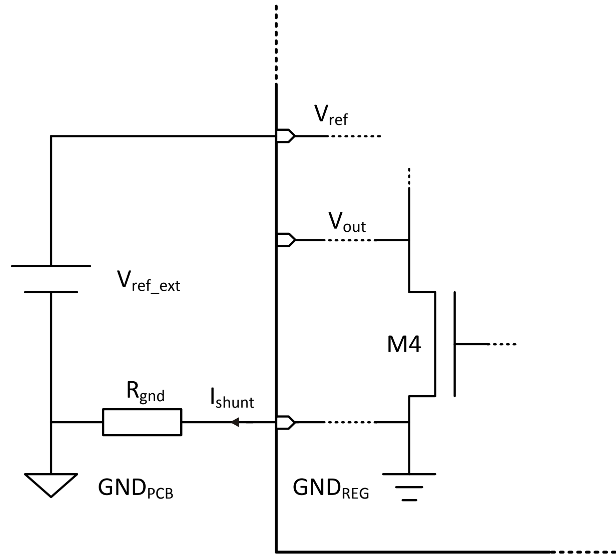


Figure 4.11: Difference between regulator and board grounds for the Shunt-LDO working mode, due to the shunt current path.

Shunt-LDO mode

The Shunt-LDO mode is the dedicated serial powering mode [47]. In this configuration, the current regulation loop is enabled, and the Shunt-LDO regulator is supplied by a constant input current I_{in} (Figure 4.12). The current regulation loop is configured in order to keep I_{in} constant, independently of changes of the load current I_L. Variations of I_L are compensated by shunting different amounts of current, so that at any time $I_{in} = I_L + I_{shunt}$. The regulator is designed in order to be able to shunt the entire input current, i.e. 0.6 A.

Figure 4.13 shows the voltage generation as a function of the input current for the two Shunt-LDO regulators connected in parallel, generating V_{out} = 1.2 V and 1.5 V. After saturation of all regulators in the design (I_{in} = 0.65 A), the V_{in}-I_{in} characteristics is linear, according to (4.3). As the internal reference resistor $R3 = 2\,\mathrm{k\Omega}$ is used, the input resistance of each regulator is 2 Ω. Because of the parallel connection, the expected input resistance is $R_{in} = 1\,\Omega$. The measured value is ~0.89 Ω, with an RMS of 0.087 Ω (Table 4.4), in agreement with the expectation. The output voltages follow the input voltage until the referenced voltage is reached. After this point the V_{out} is regulated and stable. The V_{out} changes of at most 20 mV for a change in I_{in} of 1 A. The measured values are listed in Table 4.4. They lie below the referenced value by at most 0.051 V, corresponding to 3.4 % of the referenced value.

Some considerations should be made on the start-up behavior, i.e. before both outputs reach the referenced value. As observed in the prototypes [46], when $V_{out} < 2 \cdot V_{ref}$, the current mirror aspect ratio k varies. This is due to the offset of the amplifier $A2$, causing a difference in the V_{ds} of the transistors $M1$ and $M2$, which are operated in linear region as long as the output voltage is not regulated. A consequence of the variation of k is that the value of the input impedance changes, according to (4.3). This results in the non-linear slope of V_{in} observed in Figure 4.13,

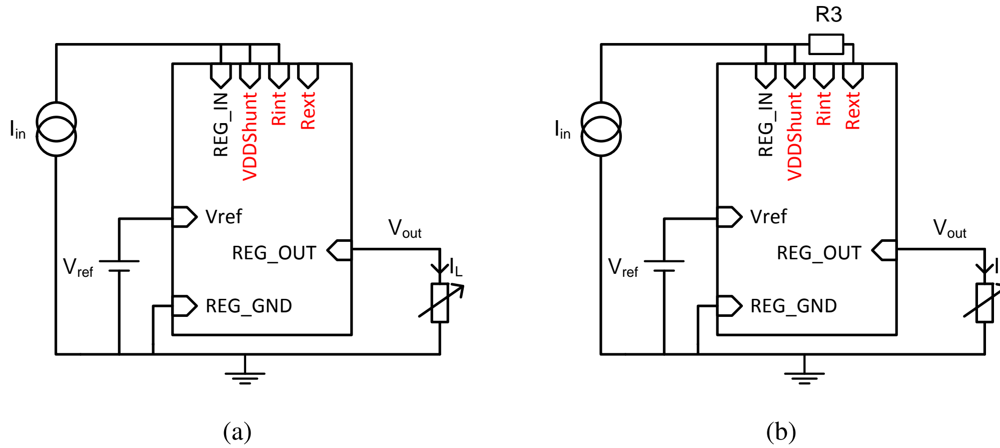


Figure 4.12: Connection of the Shunt-LDO regulator in Shunt-LDO mode with (a) internal or (b) external reference resistor.

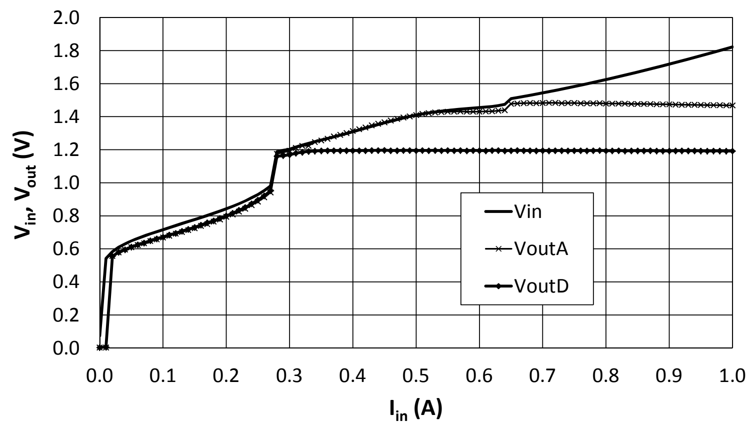


Figure 4.13: I-V characteristics of two regulators in Shunt-LDO mode, connected in parallel, generating different output voltages. The common input voltage of the two regulators, V_{in}, and their output voltages, $V_{outA} = 1.5\,\mathrm{V}$ and $V_{outD} = 1.2\,\mathrm{V}$ are shown. After the referenced output voltages are reached, the characteristics is linear with a slope corresponding to the input resistors of the two regulators placed in parallel. The non-linear behavior before regulation is due to the current splitting between the regulators at start-up, and to the ramp-up time of the input current.

and affects similarly the current splitting between the two regulators. As shown in Figure 4.14 more current flows into the regulator which is already saturated. When the outputs of the two shunt-LDO are regulated, $M1$ and $M2$ are in saturation, and the offset of $A2$ does not play a role anymore. The slope of the input characteristics is linear and the current splits between the two regulators according to the value of $R3$. In Figure 4.14, the current splits equally as the internal resistor of 2 kΩ is used for both regulators, and thus both Shunt-LDO behave as a resistor of 2 Ω for the input current. With respect to the prototypes, this effect has been reduced, with appropriate sizing of the input transistors in $A2$ and by operating them in weak inversion.

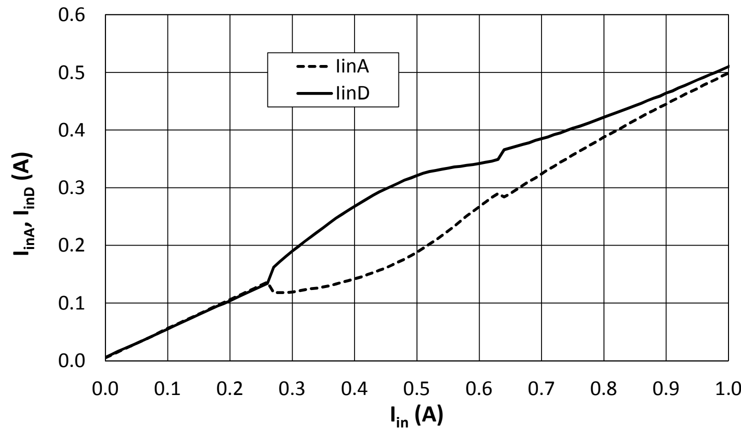


Figure 4.14: Current distribution at start-up between two regulators in Shunt-LDO mode, connected in parallel. The two regulators generate two different output voltages. The internal reference resistor is used. More current flows to the digital regulator which generates the lower output voltage, until the referenced output voltage value is reached also by the analog regulator. The current then splits in equal amount between the two regulators.

In the Shunt-LDO version in FE-I4A oscillations of both input and output voltages are observed as soon as one regulator saturates, a feature not observed in the prototypes. Oscillations last even after saturation of the second regulator. For instance, in Figure 4.13, oscillations start for I_{in} = 0.28 A, and last until I_{in} = 0.65 A. The amplitude of the oscillations is around 0.4 V, with a frequency of ~1 MHz. Measurements show that this occurs when ramping-up the input current, from zero to the desired value in steps. If the current is provided at full value at start-up, with a ramp-up phase ≤ 20 ms, voltages are at the correct value and stable right after saturation of both regulators (I_{in} = 0.6 A for the example in Figure 4.13).

The load regulation of the regulators in parallel is measured by taking an increasing load current at the output of one regulator, while the other regulator sees a constant load. Figure 4.15 shows the example of a measurement in which the load current of the digital regulator is increased. The output voltage changes with the load current, with a measured slope corresponding to an output resistance of 0.14 Ω, with an RMS of 0.009 V (Table 4.4). The analog regulator has a constant load current, and the output voltage does not change, confirming that the two regulators connected in parallel operate independently. The output voltage is maintained as long as at least 0.005 A flow into the shunt regulator, i.e. $I_{shunt} = I_{in} - I_L \geq 0.005$ A. For smaller values of I_{shunt} the input and output voltages collapse.

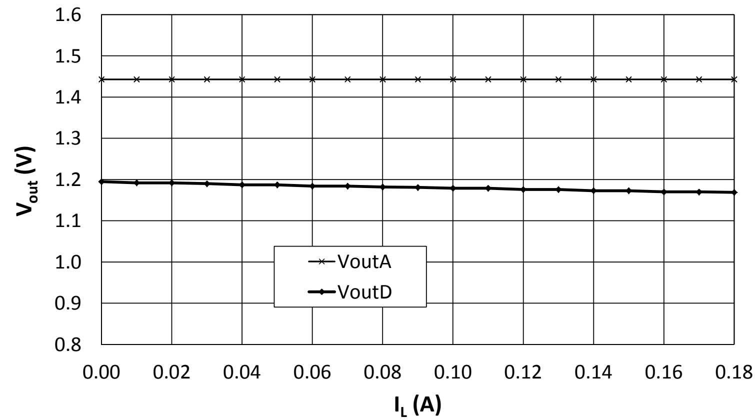


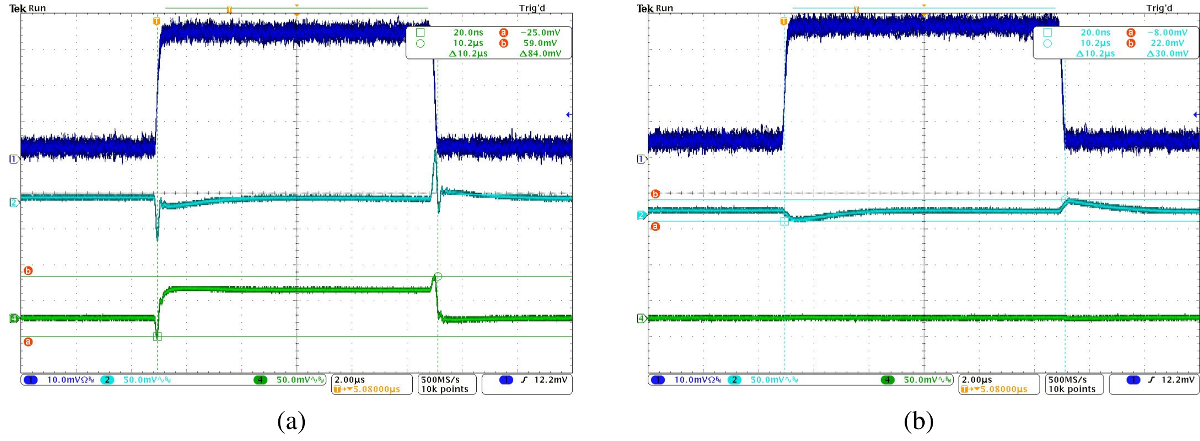
Figure 4.15: Load regulation measurement of two regulators in Shunt-LDO mode, connected in parallel. A load current is taken out of the digital regulator, and its output voltage changes with a slope corresponding to 0.14 Ω. The analog regulator sees a constant current load of 0.38 A.

Table 4.4: Mean and RMS values of the measured regulator parameters in Shunt-LDO mode.

Parameter	Mean	RMS
V_{out} = 1.5 V	1.49 V	0.006 V
V_{out} = 1.2 V	1.18 V	0.008 V
R_{in}	0.89 Ω	0.087 Ω
R_{out}	0.14 Ω	0.009 Ω

The load transient response of the regulator is studied as well. A load current pulse of 0.3 A is applied at the output of the analog regulator. As shown in Figure 4.16a its output voltage (bottom curve) changes of 0.045 V, which corresponds to the measured output impedance of 0.14 Ω. The measured transient peak is of only ~0.02 volt. Also shown in Figure 4.16a is the input voltage of the regulator (middle curve), where transients of ~0.06 V are observed, due to the finite bandwidth of the error amplifier. In this case in fact, opposite to a voltage based powering scheme, transients at the input cannot occur due to changes of the current over the input line as I_{in} is constant. This is confirmed by the fact that there is no level shift of the V_{in}. Also, different values of the input capacitor have been used, and no influence on the V_{in} transients is observed. This measurement shows again the independence of the two regulators working in parallel. In Figure 4.16b the input and output voltages of the digital regulator are shown, where no load current transient is applied. Transients are not observed for the output voltage. Even if the two inputs are connected together, the transients on the V_{in} of the analog regulator affect only minimally the input of the digital regulator, where a transient of only 0.01 V is measured.

Finally, the power efficiency of the regulator is measured. For a single regulator two sources of inefficiency are to be considered: the $V_{dropout}$ across the pass transistor of the LDO, and the I_{shunt}. In this case the power efficiency is given by



(a) (b)

Figure 4.16: Load transient measurement of two regulators in Shunt-LDO mode, connected in parallel. A load transient of 0.3 A is applied to the output of the analog regulator. The top curve in both figures shows the load current transient measured across a 0.1 Ω resistor (0.01 V/div). The middle and bottom curves (0.05 V/div) show respectively the input and output voltages of (a) the analog regulator where the load current transients is applied, and (b) of the digital regulator. The time scale is 2 µs/div.

$$P_{eff} = \frac{V_{out}}{V_{out} + V_{dropout}} \cdot \frac{I_L}{I_L + I_{shunt}}. \tag{4.7}$$

For two regulators connected in parallel, generating different output voltages, a third source of inefficiency needs to be considered, which is the ΔV between the two output voltages. The combined power efficiency is thus calculated as

$$P_{eff-comb} = \left(\frac{V_{outA}}{V_{outA} + V_{dropoutA}} \cdot \frac{I_L}{I_L + I_{shunt}}\right) - \left(\frac{\Delta V}{V_{outA} + V_{dropoutA}} \cdot \frac{I_{LD}}{I_L + I_{shunt}}\right) \tag{4.8}$$

for $I_L = I_{LA} + I_{LD}$, and $I_{shunt} = I_{shuntA} + I_{shuntD}$. Figure 4.17 shows the power efficiency of the two regulators as well as the combined power efficiency, as a function of the shunt current, for $V_{dropout} = 0.2$ V and $\Delta V = 0.3$ V. The regulator generating the higher V_{out} has an efficiency of 85 % for $I_{shunt} = 0.01$ A. The efficiency of the regulator generating $V_{out} = 1.2$ V is lower due to the higher voltage drop across the pass transistor (67 % for $I_{shunt} = 0.01$ A). The combined power efficiency is 79 %. The power efficiency is particularly affected by the shunt current. With increasing shunt current, the power efficiency decreases rapidly, as shown in Figure 4.17. To reach high power efficiency it is of fundamental importance to choose the optimum value of the reference resistor in order to optimize I_{shunt}, and in turn also the I-V characteristics to reach the minimum $V_{dropout}$. For $V_{dropout} = 0.1$ V, $I_{shunt} = 0.005$ A, and $\Delta V = 0.3$ V, the combined power efficiency would be 86 %.

These results confirm that the regulator operation in Shunt-LDO mode is as expected, and thus that the regulator fulfills all specifications to power the FE-I4 chip in a current based powering scheme. The output voltage can be generated according to the referenced value. The measured value of R_{in} agrees with the expected value set via the reference resistor $R3$. The value and spread of R_{out} are sufficiently small such that the resulting difference of output

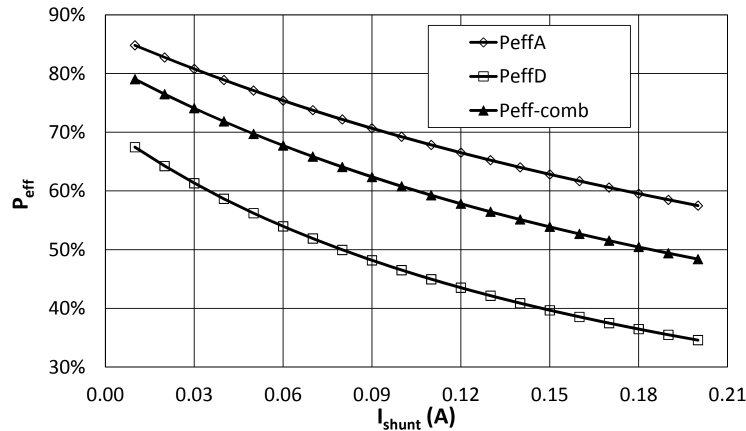


Figure 4.17: Shunt-LDO power efficiency as a function of the shunt current, for two regulators placed in parallel generating different output voltages. The efficiency of each regulator is shown (P_{effA} and P_{effD}), as well as the combined efficiency resulting from the parallel operation ($P_{eff-comb}$).

voltages for different regulators is negligible.

LDO mode

In LDO mode [47], the current regulation loop is disabled and the regulator works as a standard Low-Dropout regulator (Figure 4.18). This working mode is used in the FE-I4A in combination with the DC-DC converter (Figure 3.13).

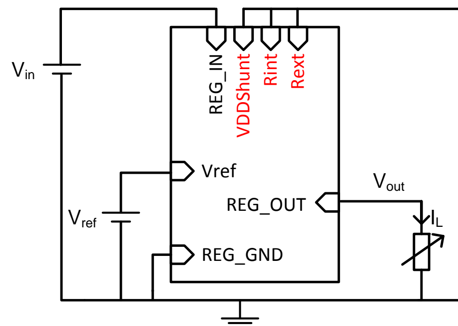


Figure 4.18: Connection of the Shunt-LDO regulator in LDO mode.

The minimum $V_{dropout}$ required for regulation is shown in Table 4.5 as a function of V_{out} and I_L. The measured values are higher than the design value (0.1 V), especially for V_{out} = 1.2 V at $I_L > 0.2$ A. This is however not a concern for the specific application in FE-I4A. For standard configuration, the digital current is below 0.2 A. In addition to this, the regulators share the same input to power the FE-I4A. A V_{in} of about 1.65 V has to be used to provide enough $V_{dropout}$ for the analog regulator. Thus, the $V_{dropout}$ of the digital regulator is ~0.45 V.

Table 4.5: $V_{dropout}$ and line regulation as a function of V_{out} and I_L for the LDO working mode.

	I_L = 0.2 A		I_L = 0.6 A	
	Mean	RMS	Mean	RMS
	$V_{dropout}$ [V]			
V_{out} = 1.2	0.15	0.022	0.27	0.035
V_{out} = 1.5	0.05	0.012	0.15	0.029
	$\Delta V_{out}/\Delta V_{in}$ [mV/V]			
V_{out} = 1.2	2	0.5	4	1.5
V_{out} = 1.5	2	0.5	12	2.5

The line regulation ($\Delta V_{out}/\Delta V_{in}$) is shown Table 4.5. The V_{out} changes of only 12 mV for a ΔV_{in} = 1 V. The measured load regulation corresponds to an output resistance of 0.14 Ω, in agreement with the measurement in Shunt-LDO mode. The quiescent current, i.e. the difference between input and load current, is only 0.001 A.

The power efficiency for single device operation is between 75 % and 95 %, at minimum $V_{dropout}$ and over the entire current range (Figure 4.19). When connected in parallel for $V_{dropout}$ = 0.2 V, and ΔV = 0.3 V, the power efficiency is 79 %.

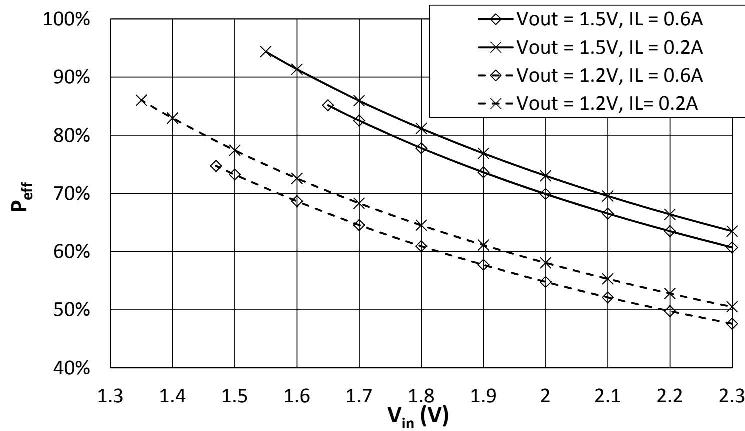


Figure 4.19: Power efficiency of the regulator in LDO mode for different load currents and output voltages as a function of the input voltage.

Partial shunt mode

In partial shunt mode [47], the regulator is operated as in LDO mode, but with enabled current regulation loop (Figure 4.20). This working mode is used for the IBL power distribution scheme [57]. The I_{shunt} is set in order to prevent the input current from falling under a preset value I_{min}, higher than the minimum I_L. Thus

$$I_{in} = I_{min} \text{ for } I_L \leq I_{min} \tag{4.9}$$

$$I_{in} = I_L \text{ for } I_L > I_{min}. \tag{4.10}$$

This configuration can be used to reduce transients in voltage based powering schemes, where changes in the load current translate in changes of the current over the power lines. Transients are thus expected at the input of the regulator. To demonstrate this operation mode, the two Shunt-LDO regulators are connected in parallel with $V_{in} = 1.8V$, and operated in partial shunt mode and LDO mode. The minimum current in partial shunt mode is set to 0.13 A for the analog regulator, and 0.08 A for the digital regulator. A load current pulse from 0.04 A to 0.34 A is applied externally to the analog regulator (top curve in Figure 4.21). When in LDO mode, the entire current transient is present at the regulator output, whilst in partial shunt mode this transient is reduced to 0.21 A because I_{in} cannot be lower than 0.13 A.

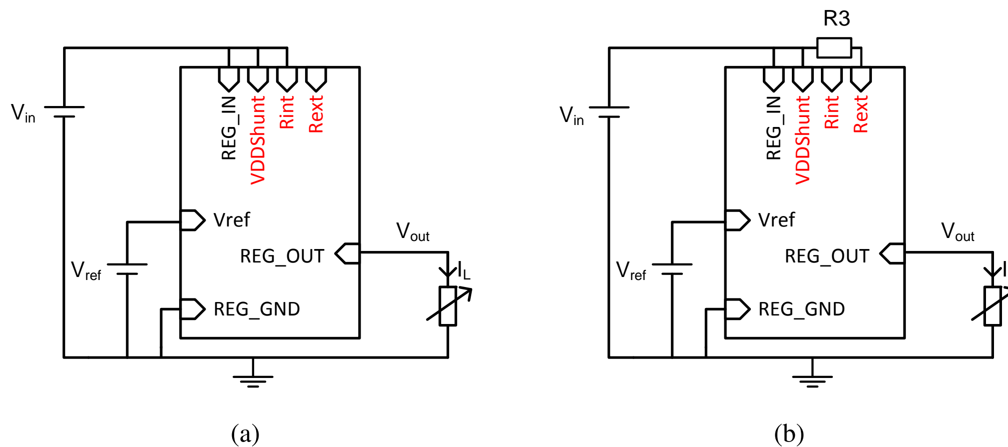
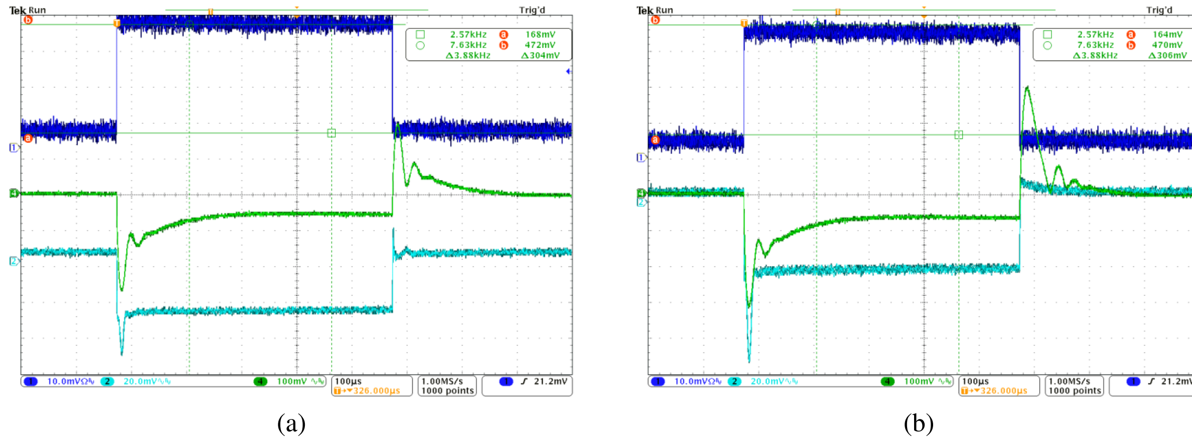


Figure 4.20: Connection of the Shunt-LDO regulator in partial shunt mode with (a) internal or (b) external reference resistor. The connection is the same as in Shunt-LDO mode, but the regulator is powered with a voltage, instead of a current.

As shown in Figure 4.21, the transients on V_{in} (middle curve) are lower in partial shunt mode. Transients of 0.2 V are observed in partial shunt mode (Figure 4.21a), and up to 0.3 V in LDO mode (Figure 4.21b), at the regulator input. The transients at the output of the regulator (bottom curve in Figure 4.21) are mostly due to the fact that V_{in} falls under the minimum required $V_{dropout}$. The output is no longer regulated, and follows the input voltage. A higher V_{in} reduces the transients on V_{out}. The amplitude of the transients observed at the output is at most 0.02 V in partial shunt mode, whilst it reaches 0.05 V in LDO mode. The voltage level shift measured on V_{in} and V_{out} corresponds respectively to the resistance of the lines between supply and regulator input, and to the output resistance of the regulator.



(a) (b)

Figure 4.21: Transients in (a) partial shunt mode and (b) LDO mode. Top curve: current pulse at the regulator output measured across a 0.1 Ω resistor (0.01 V/div). Middle curve: transient at the regulator input (0.1 V/div). Bottom curve: transient at the regulator output (0.02 V/div). The horizontal time scale is 100 µs/div.

4.4 Serially powered module operation

Before setting up a serial powering chain, it is important to study the module performance with FEs powered by a constant current. To this aim, single and double chip modules from the IBL prototyping are used (i.e. 250 µm thick planar n^+-in-n sensor, bump bonded to FE-I4A). In particular two MSC modules and six MDC modules are available for the tests. MSC modules are built with 150 µm thin FE, while the MDC modules are built with 100 µm thin chips. Due to the limited module availability, three MDC modules are built without a sensor, and with FE chips not previously tested at wafer level. They are referred in the following as digital modules and are indicated with an asterisk. These different module configurations are beneficial as they allow to recognize effects on module performance due to the FE electronics or to the sensor.

Single chip modules are tested on PCB boards. This setup allows in fact to easily switch between direct power, and different Shunt-LDO working modes, to compare the module performance for different powering configurations. As double chip modules are foreseen for the SP stave prototype discussed in Section 4.5, they are equipped with a flex hybrid designed for SP distribution only. The flex features a connector compatible with loading on a box beam stave. A MDC module with flex hybrid is shown in Figure 4.22.

To power the FE-I4 with the Shunt-LDO regulators, the connection shown in Figure 4.10 is used. For single chip modules, two regulators are operated in parallel, while for double chip modules, the input current splits between four regulators. The value of the resistor $R3$ is chosen according to the current needs of the FE-I4. For a standard configuration, the FE-I4 power consumption is ~0.35 A (max. 0.38 A) for the analog part and ~0.14 A (max 0.18 A) for the digital part. The reference resistors are chosen to be $R3 = 3\,\mathrm{k\Omega}$ for the analog regulator and $R3 = 9\,\mathrm{k\Omega}$ for the digital regulator. In this way, an input current of 0.5 A per FE, splits into 0.35 A through the analog regulator, and 0.15 A through the digital.

The reference voltages are provided using a commercial adjustable shunt regulator [63],

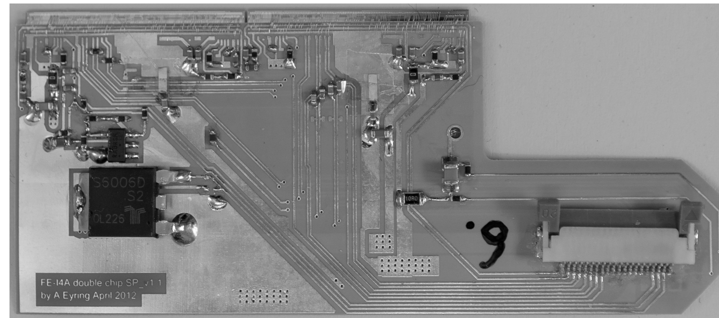


Figure 4.22: Double chip FE-I4A module with dedicated SP flex hybrid. The connector visible on flex top side is used for connection to the test setup. The connector to plug the modules to the stave cable is on the flex back side. The two ICs visible on the left are the commercial over-voltage protection element, and the adjustable shunt to generate the reference voltage. At the top, wire bonds between the FE chips and the flex can be seen. The chip on the left is referred to in the following as FE0, and the one on the right as FE1.

which can be loaded on the module flex. In this way each module can generate its reference voltages independently, without the need of a floating voltage source per module when operated in a chain. This device is connected in parallel to the regulators. For $V_{in} \geq 1.0V$ the adjustable shunt starts regulating and generates the two reference voltages of 0.75 V and 0.6 V. Figure 4.23 shows a comparison of the I-V curve of a Shunt-LDO regulator in shunt-mode, recorded with V_{ref} provided by a voltage supply or with the adjustable shunt. The difference is within the measurement error, proving that the solution with the adjustable shunt regulator can provide a correct and stable reference voltage to the regulators.

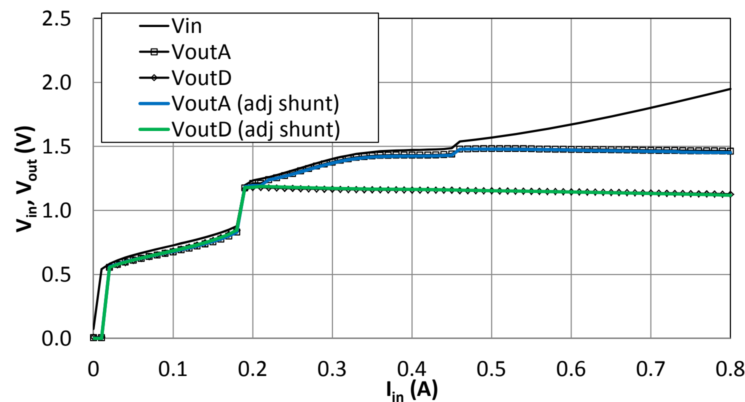
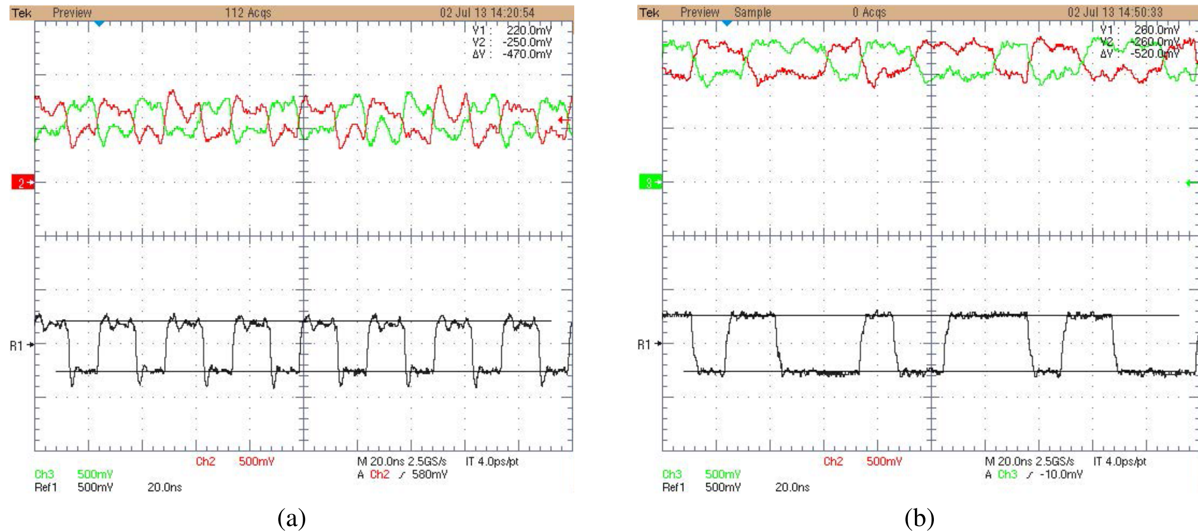


Figure 4.23: Generation of the regulator reference voltage using a voltage supply or a commercial adjustable shunt.

The data communication uses AC-coupled LVDS links (Figure 4.8). According to (4.4), with $R_{term} = 110\,\Omega$, $T = 25\,\text{ns}$, $RL = 1$ for CLK and CMD, and $RL = 5$ for DO, the value of the capacitors should be $\geq 2\,\text{nF}$ for the input signals, and $\geq 9\,\text{nF}$ for the FE output. A value

of 10 nF is used on all lines. The capacitors are mounted at the input of the receiver, i.e. on the flex and PCB for clock and command, and on the readout side for the data. Figure 4.24a and Figure 4.24b show respectively the CLK signal after the capacitors on the flex, and the DO signal after the capacitors on the readout board. Both single-ended and differential signals are shown. The common mode voltage is set by both receivers with dedicated biasing circuitry to approximately half of their bias voltage, and single ended signals are centered around it. The measured common mode voltage is 0.6 V at the input of the FE-I4 receiver, and 1.1 V at the input of the receiver on the readout board. The differential signals have an amplitude of ~0.5 V.



(a) (b)

Figure 4.24: AC-coupled (a) clock and (b) data links. Both single-ended (top curves) and differential (bottom curves) signals are shown. The vertical scale is 0.5 V/div and the horizontal one is 20 ns/div.

The *HV* scheme is implemented as shown in Figure 4.9. Finally, on the module flex, a commercial Silicon Controlled Rectifier (SCR) is used as protection element, to guard the modules from over-voltage [64]. SPP chips were not available at the time the flex hybrid has been designed. The SPP has been tested at a later time with one module (Section 4.6).

Module characterization is done using the same setup and procedures developed for the FE-I4 modules, and in particular for the IBL modules and staves. Tests are done using an USB-based test system developed for the FE-I4 chip, the USBPix [65]. The hardware is based on a FPGA board and different interconnection boards to interface to FE-I4 single ICs and wafers, and FE-I4 modules. Up to four MSC or MDC modules can be operated with the USBPix. A software package provides all scan routines needed for characterization. To check the module performance, first threshold and noise figures are measured for tuned modules. To tune a module, an algorithm is used to set the discriminator threshold to the chosen target value over the entire pixel matrix (globally), and in every pixel (locally). The feedback current of the CSA is as well tuned locally. A threshold scan is then performed. Different amount of charges are injected 100 times into every pixel, and hits are recorded as a function of the injected charge. From the typical obtained S-curve, the threshold and noise are calculated. The main source of error for module measurements comes from the calibration of the injected charge, and it is

calculated to be ∼10 %.

4.4.1 Module performance for different powering options

Table 4.6 shows the results of a tuning on the two MSC modules, powered in three different configurations:

- directly with $VDDA = 1.5$ V and $VDDD = 1.2$ V provided using an external voltage source;
- serially using a current of 0.5 A;
- directly with on-chip voltage regulation, i.e. using the Shunt-LDO regulators in LDO mode ($V_{in} = 1.8$ V).

Table 4.6: Threshold, dispersion and noise of MSC modules for different powering configurations. The target threshold is 3000 e^-, and the target $TOT = 5$ is set for a charge of 20 ke^-.

Module	Power option	Threshold [e^-]	Thr. dispersion [e^-]	Noise [e^-]
MSC1	Direct	2975	40	160
	Serial	2985	43	176
	LDO	2998	45	171
MSC2	Direct	2992	30	143
	Serial	3001	32	167
	LDO	2993	33	164

Independently of the powering scheme, the threshold can be tuned uniformly over the entire pixel matrix (26 880 pixels), with a dispersion of 30 − 40 e^-, typical for the FE-I4 chip [38]. No significant noise increase is observed for modules powered with a constant current. In particular the noise difference between direct and serial powering configurations is in agreement with what is observed in [51] for serially powered FE-I3 modules. The measured noise values are in the expected range for a tuning at 3000 e^-, with a $TOT = 5$ for a charge of 20 ke^- [38]. Threshold and noise maps of the MSC2 module for the different powering configurations are shown in Figure 4.25, Figure 4.26, and Figure 4.27.

The sensor biasing curve is measured as well to validate the chosen HV scheme. Typical curves are measured with modules operated at room temperature, corresponding to a module temperature $T_{mod} \approx 45$ °C. Figure 4.28 shows the measured sensor I-V curve for module MSC2, for the three powering options. For an operating $HV = -100$ V, the leakage current is in the order of a few µA. A slightly higher leakage current is measured for the SP option, due to the higher heat produced on module.

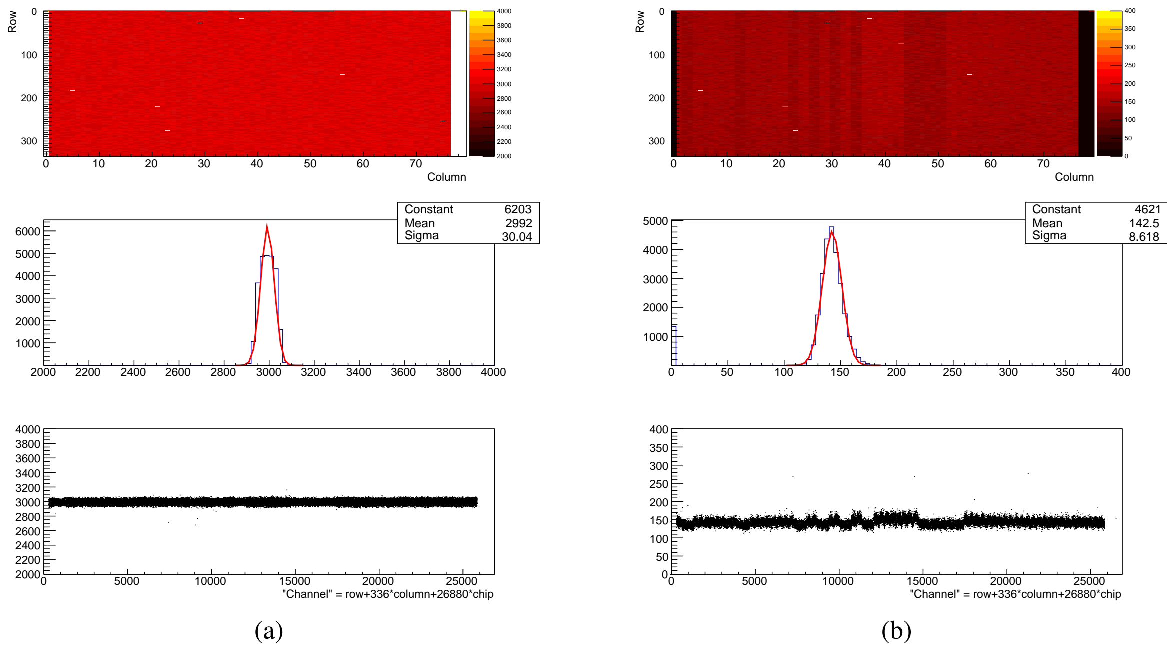


Figure 4.25: (a) Threshold and (b) noise map of the MSC2 module in direct power configuration. The different column flavors of the FE-I4A can be recognized in the noise map.

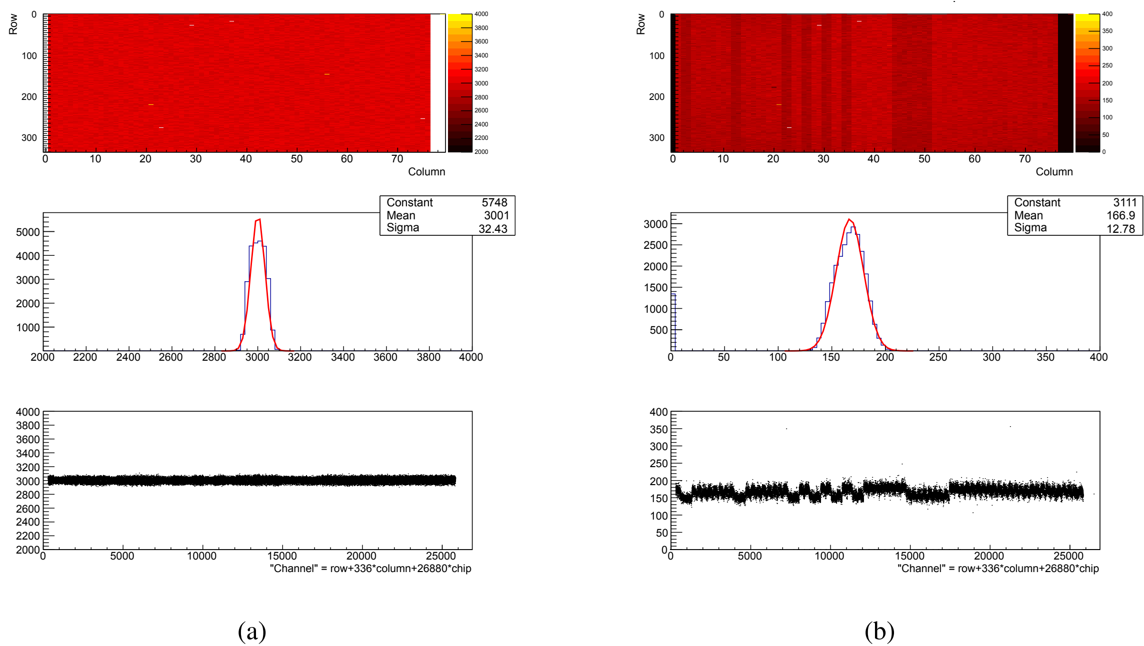


Figure 4.26: (a) Threshold and (b) noise map of the MSC2 module in serial power configuration. The different column flavors of the FE-I4A can be recognized in the noise map.

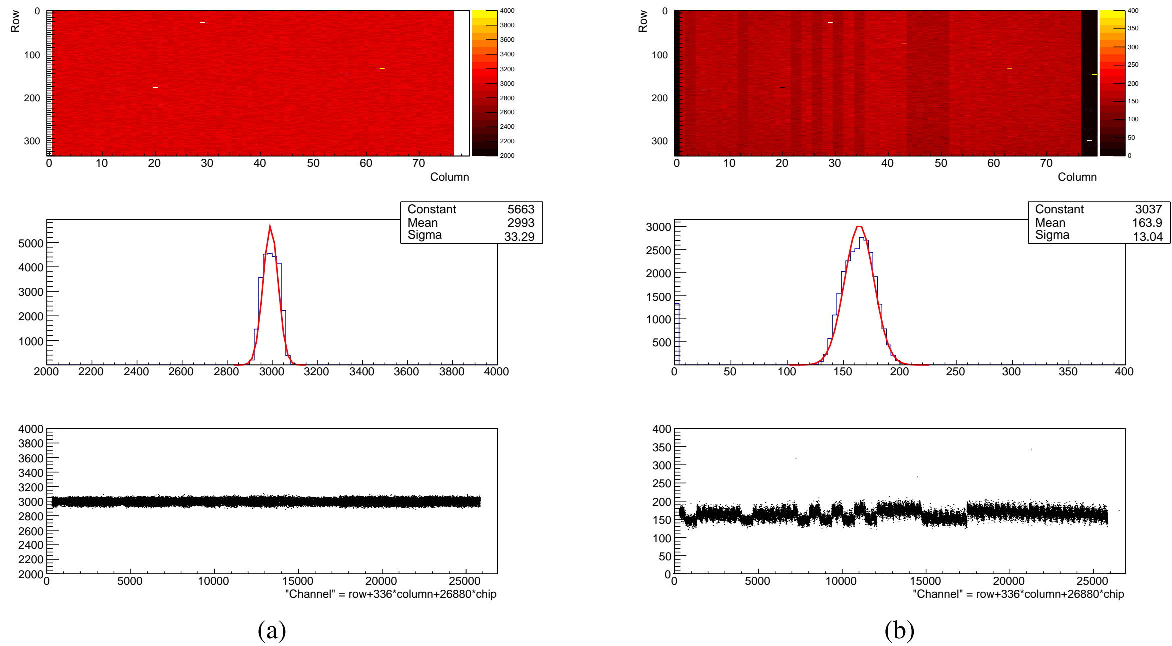


Figure 4.27: (a) Threshold and (b) noise map of the MSC2 module in LDO power configuration. The different column flavors of the FE-I4A can be recognized in the noise map.

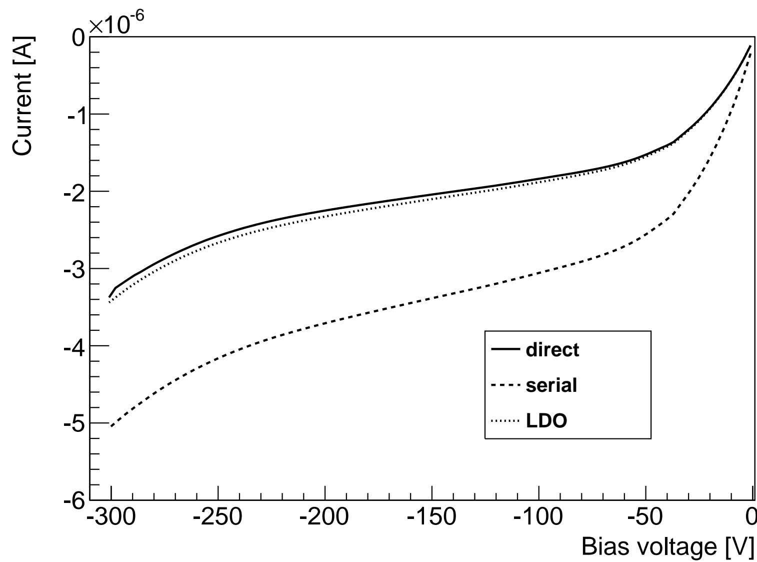


Figure 4.28: Sensor I-V characteristics for MSC modules, operated with different powering schemes, namely direct, serial, and LDO. Modules are operated at room temperature and the FEs are configured. The module temperature is around 45 °C. The higher leakage current measured with the serial powering configuration is due to the higher heat produced on module, approximately 2 °C.

4.4.2 Multi-chip module performance for different temperatures

The same testing procedure is used for MDC modules, where the current splits between Shunt-LDO regulators on different FE chips. In these tests however tuning and threshold scans are performed at a module temperature of −15 °C. These tests allow to check the module performance, including the one of the regulators, at low temperatures as during detector operation. Modules are cooled in a climate chamber, where the environmental temperature is set to −30 °C to reach a module temperature during operation of −15 °C. When operated at room temperature, $T_{mod} \approx 45$ °C. Within this range of module temperature, i.e. −15 °C to 45 °C, the input voltage of the regulator, the analog and digital voltages of the chip are stable within 2 %. The reference voltages change less than 0.1 %. The module performance is thus perfectly maintained at low temperature. As shown in Table 4.7, the tuning is successful, with dispersion and noise values in the expected range for the chosen tuning parameters [38], for both types of modules, i.e. with and without sensor.

Table 4.7: Threshold, dispersion and noise of serially powered MDC modules. MDC5*, MDC6*, and MDC8* are built without sensor. Tuning and scans are performed at a module temperature of −15 °C.

Module	Power option	Threshold [e⁻]	Thr. dispersion [e⁻]	Noise [e⁻]
MDC1	Serial	2994	36	160
MDC2	Serial	2989	44	168
MDC4	Serial	2992	40	151
MDC5*	Serial	2988	43	193
MDC6*	Serial	2990	36	147
MDC8*	Serial	2992	40	177

The sensor I-V characteristics is measured as well during cold operation, which reduces the leakage current to a few nA (Figure 4.29). In particular, for a biasing voltage of −100 V, the measured leakage current is −13 nA, −7 nA and −20 nA respectively for MDC1, MDC2, and MDC4. When operated at room temperature the modules have a leakage current around 10 −20 μA.

4.5 Serial powering pixel stave prototype

As explained in Section 4.1, a serially powered chain of modules requires some modifications at system level, which are not needed in a voltage based powering scheme, such as for instance AC-coupled data transmission. It is thus particularly important to prototype SP chains as early as possible, to address all aspects connected to a current based powering scheme. The results discussed in Section 4.4 show that the chosen regulation, AC-coupling, and *HV* schemes allow operation of serially powered FE-I4 modules without performance degradation. The last

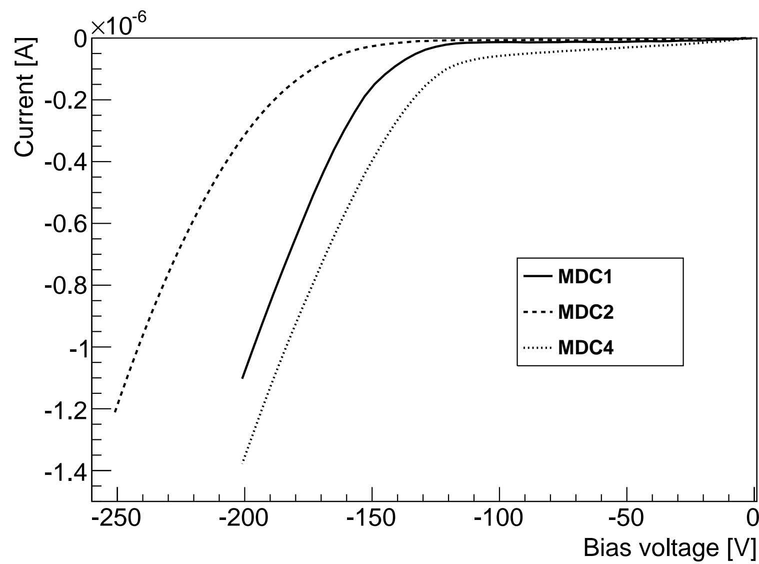


Figure 4.29: Sensor I-V characteristics for serially powered MDC modules. The module temperature is −15 °C.

missing piece of the demonstration is a serially powered chain of modules, where these developments can be brought together and tested at the same time. This is implemented as shown in Figure 4.30.

Double chip FE-I4 modules are mounted on a box beam prototype structure, where one of the four integrated cables is designed for serial current distribution. A dedicated aluminum layer is designed to route the current from one module to the next, and back to the current source. This layer is connected using Tape-Automated Bonding (TAB) to a copper multilayer cable for signals and HV. This first version of the multilayer cable is meant as a dummy to study the thermo-mechanical properties of the box beam stave concept, and as such a bad electrical performance is expected. Due to many open lines on the cable, only four modules instead of eight can be connected in the SP chain. For two of the modules, respectively MDC2 and MDC8*, external CLK and CMD lines are necessary. In MDC1 only one of the two FE can be read back. On the cable tabs where no module is loaded a flex is used to short the current input and output paths to assure current flow to the next module.

The stave is operated at room temperature to avoid further damage that might arise on the cables due to thermal cycling. At the end of the stave a so-called End-Of-Stave (EOS) card is used as a passive interface between the stave on one side, and the current source, HV supply and readout system on the other side. The four MDC modules are connected to the same USBPix system. CLK and CMD are sent to all modules in parallel, meaning that the scan routines run in parallel on all modules. A high digital activity is thus present in the SP chain during testing. As only one module can be read back at a time, all scan routines are performed four times.

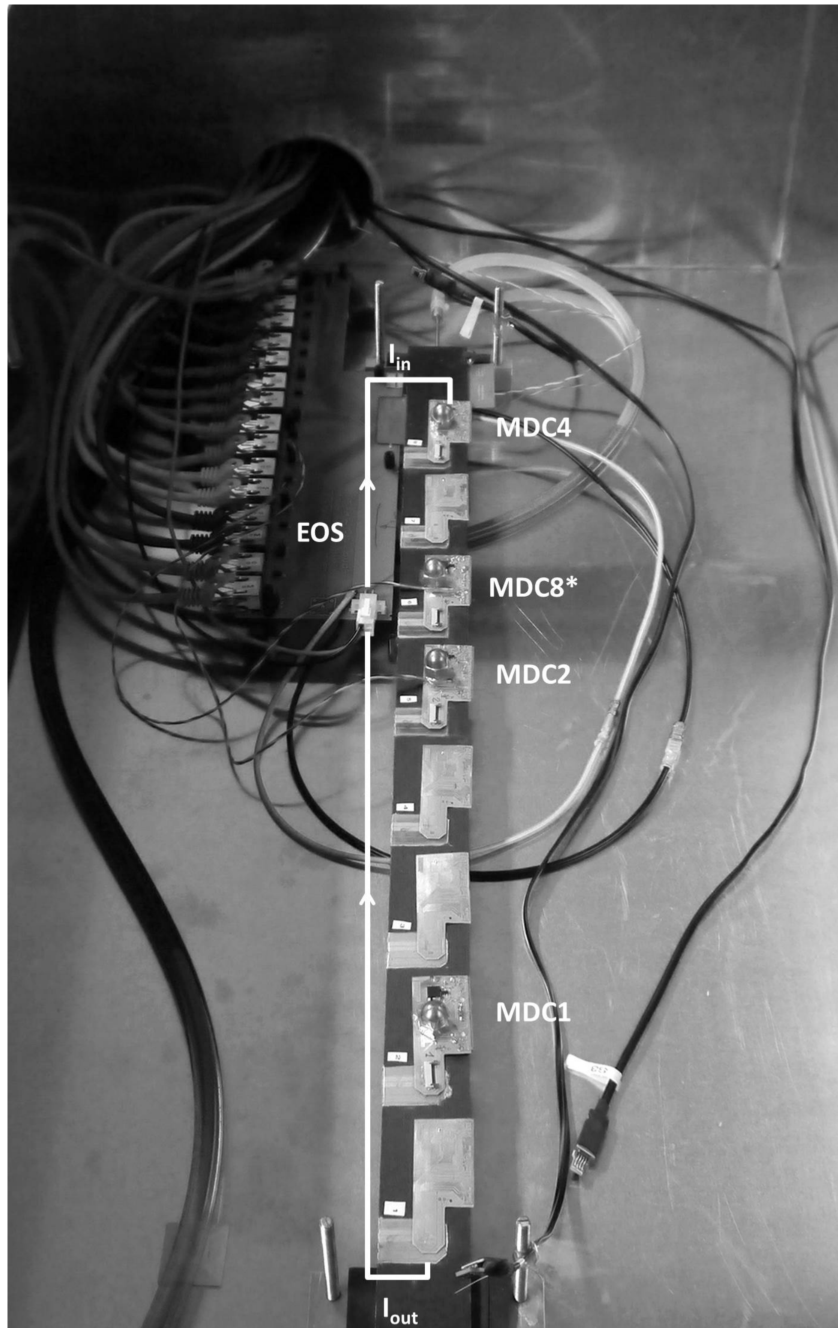


Figure 4.30: SP prototype stave. A full length box beam stave is available, i.e. 1.4 m, with one cable for SP. Only the stave section where SP is implemented is shown in the picture. Four modules are loaded on the stave, where MDC8* is a digital module. The modules are not glued on the stave to allow reworking, and are kept in thermal contact with the stave using weights. The EOS card can also be seen on the top left side of the stave, featuring ethernet connectors and cables for the data lines and the current connector.

4.5.1 Power-up procedure

A current of 1 A is used to power the modules. The current is provided at full value, with a ramp-up speed of less than 20 ms, ensuring no oscillations at power-up. The modules power-up with the correct value of $VDDA$ and $VDDD$, as summarized in Table 4.8. The measured values of $VDDA$ are lower than $2 \cdot V_{ref}$. This is due to the external connection via wire bonds and traces on the module flex between the regulator output and the power rails of the chips. The voltage drop generated by the current flowing to the chip through this path, lowers the $VDDA$ voltage with respect to the voltage generated at the regulator output. In addition to this, the regulator output resistance plays a role as well. This effect is more pronounced for the analog regulator as the analog current at power-up is ~0.05 A. The digital current, in absence of a clock signal, is only a few mA. It should also be mentioned here that with this setup it is not possible to measure the voltage difference between regulator and flex grounds and thus the voltages cannot be corrected for the difference of the provided and effective V_{ref}.

The sensors are biased with a voltage of −80 V. The total leakage current of the three modules is approximately −30 µA, which is expected given the values measured on the MDC modules when tested standalone at room temperature.

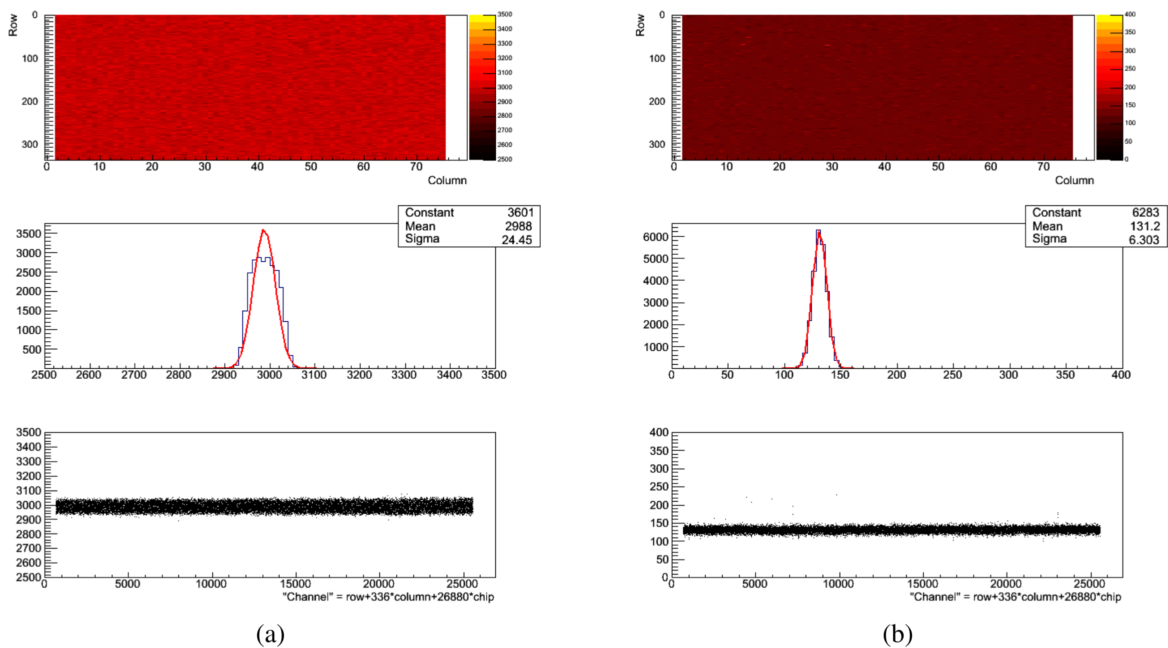
Table 4.8: Voltages of the modules in the SP chain at power-up. $VDDA0$ and $VDD0$ are respectively the analog and digital voltage of FE0, $VDDA1$ and $VDD1$ respectively the analog and digital voltage of FE1.

Module	V_{in}		VDDA0		VDDD0		VDDA1		VDDD1	
	Mean	RMS	Mean	RMS	Mean	RMS	Mean	RMS	Mean	RMS
	[V]									
MDC1	1.90	0.009	1.41	0.006	1.17	0.008	1.44	0.006	1.22	0.008
MDC2	1.92	0.009	1.40	0.006	1.18	0.008	1.46	0.006	1.19	0.008
MDC8*	1.94	0.009	1.44	0.006	1.23	0.008	1.43	0.006	1.15	0.008
MDC4	1.92	0.009	1.44	0.006	1.20	0.008	1.45	0.006	1.20	0.008

4.5.2 Tuning of the SP stave

As a first test the stave modules are tuned. The tuning algorithm is more elaborate than in the tests discussed in Section 4.4, and follows the procedure used for the IBL staves. Both the threshold and the feedback current are tuned over the entire pixel matrix and for each pixel. Modules are tuned to target threshold values of 4000 e^- and 3000 e^-. The TOT is tuned in order to have a value of 9 at 16 ke^-. This choice is made considering that a charge of 16 ke^- corresponds to a Minimum Ionizing Particle (MIP). A TOT = 9 allows a large enough range to measure a MIP, considering charge sharing between neighboring pixels, as well as the possibility to measure charges larger than a MIP.

Table 4.9 shows the threshold, threshold dispersion and noise of the four tuned modules. The modules can be tuned without problems to the selected threshold with dispersion below 30 e^-. The noise values are between 120 e^- and 140 e^- for a threshold of 4000 e^-, and between 130 e^- and 150 e^- for the 3000 e^- threshold. Figure 4.31- 4.33 show the threshold and noise maps of the tuning at 3000 e^- for all the modules. These results, compared to the ones shown in Section 4.4, confirm that the modules operated in a serially powering chain do not influence each other, as observed with FE-I3 modules [55]. The measured values of the threshold dispersion and of the noise are in very good agreement with the values obtained on IBL staves, where FE-I4 modules are powered using a voltage based scheme with on-chip power conversion (i.e. Shunt-LDO regulators in partial shunt mode).



(a) (b)

Figure 4.31: (a) Threshold and (b) noise maps of module MDC1 tuned to 3000 e^-. Only one FE can be read back due to an open line in the LVDS pair. The first two and last four columns of the FE chips are disabled, as they are powered with pads at the top of the FE pixel matrix which are not accessible on module.

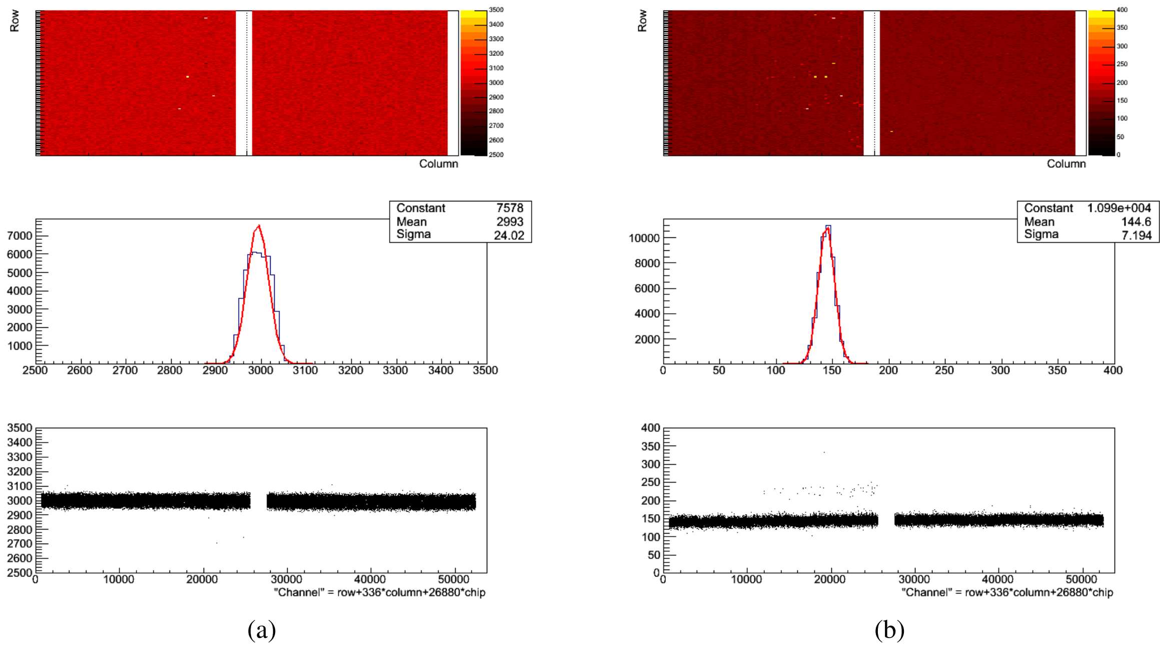


Figure 4.32: (a) Threshold and (b) noise maps of module MDC2 tuned to 3000 e$^-$. The first two and last four columns of the FE chips are disabled, as they are powered with pads at the top of the FE pixel matrix which are not accessible on module.

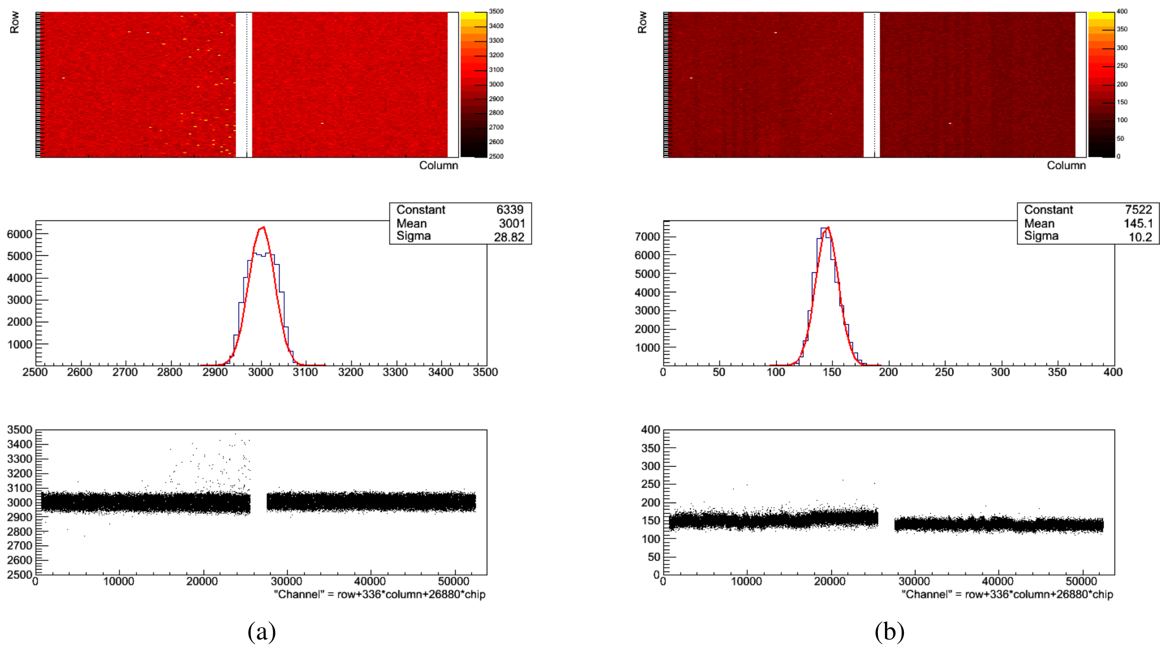


Figure 4.33: (a) Threshold and (b) noise maps of module MDC8* tuned to 3000 e$^-$. The first two and last four columns of the FE chips are disabled, as they are powered with pads at the top of the FE pixel matrix which are not accessible on module.

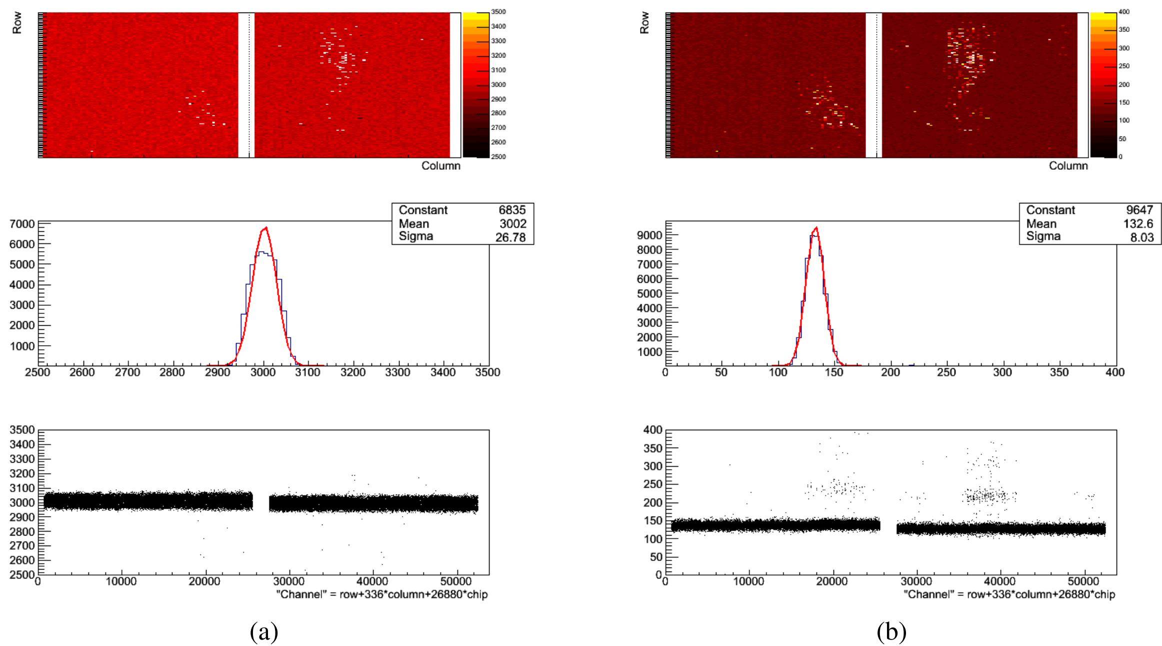


Figure 4.34: (a) Threshold and (b) noise maps of module MDC4 tuned to 3000 e$^-$. The first two and last four columns of the FE chips are disabled, as they are powered with pads at the top of the FE pixel matrix which are not accessible on module. Due to a problem during the flip chip process, some pixels are shorted, and can be recognized as they show higher noise.

4.5.3 Timing performance tests

Further tests can be performed to check the timing performance of the modules. Due to the time walk, two simultaneous hits with a large charge difference are assigned to two different clock cycles of 25 ns each, and thus are not associated with the same trigger signal. The smallest charge that can be detected in time, i.e. assigned to the correct bunch crossing, is equal to the so-called in-time threshold. Its value is measured for the modules on the stave, and used to calculate the overdrive. This is defined as the difference between the in-time threshold and the discriminator threshold, and it corresponds to the amount of charge that needs to be over the threshold to be detected in time. While the in-time threshold depends on the discriminator threshold value (i.e. if two pixels have different threshold, the one with lower threshold has a smaller in-time threshold and seems faster), the value of the overdrive is independent of the threshold and can thus better describe time dependent effects. For the modules on the stave, the overdrive value is in a range of 1000 – 1500 e$^-$ (Table 4.9). Taking into account the FE-I4 time walk correction mechanism, the overdrive for the stave modules is not critical for the timing performance. The FE-I4 would assign hits with TOT ≤ 2 to the previous bunch crossing. Given the chosen tuning parameters (TOT = 9 at 16 ke$^-$), this correspond to an overdrive $\leq$ 3500 e$^-$.

Table 4.9: Results of tests on the SP stave prototype.

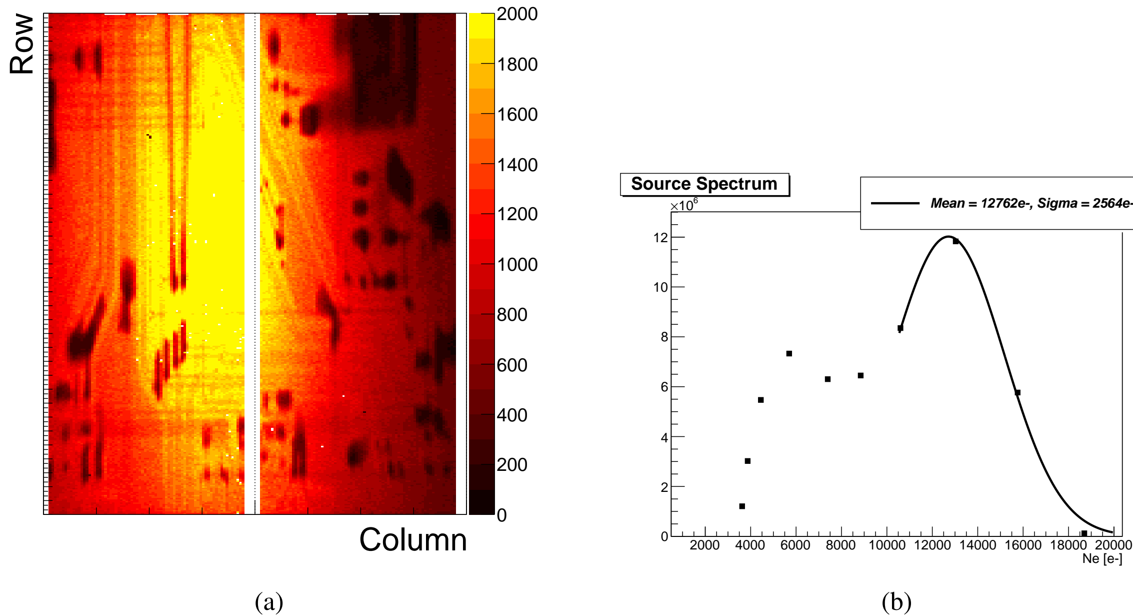
Module	Threshold [e$^-$]	Thr. dispersion [e$^-$]	Noise [e$^-$]	In-time thr. [e$^-$]	Overdrive [e$^-$]
	Tuning at 4000 e$^-$				
MDC1	4038	25	126	-	-
MDC2	4017	23	138	-	-
MDC8*	4018	26	134	-	-
MDC4	4023	25	125	-	-
	Tuning at 3000 e$^-$				
MDC1	2988	24	131	4051	1063
MDC2	2993	24	144	4120	1127
MDC8*	3001	29	145	4390	1389
MDC4	3002	27	133	4237	1235

4.5.4 Source scan

Finally a source scan is performed, which records hits from a radioactive source, with a trigger signal generated by the FE every time a hit is recorded. An ^{241}Am source is used for this test. With this test the full module functionality is checked: charge collection in the sensor, hit processing in the FE and data outputting. Figure 4.35a shows the hit map recorded on a module. Hits are recorded in all pixels. A higher number of hits is recorded in correspondence to the source location. Figure 4.35b shows the source spectrum. A peak in the spectrum is recorded at a charge of (12.8±1.5) ke$^-$, where the error is given by both the error on the calibration of the injected charge and of the TOT. As only four bits are available for the TOT code, the charge resolution of the FE-I4 is limited. This value corresponds to an energy of (45.94±5.53)keV, compatible with the 59.54 keV peak of the ^{241}Am source. The same is obtained on IBL staves. This result proves that all features of the serial powering chain, from voltage generation, to data transmission and sensor biasing work correctly.

4.5.5 Failure mode studies

The results of the tests discussed in the previous paragraphs show that the SP stave performance is excellent in standard working conditions. Next, failure modes have to be addressed, i.e. scenarios in which a disturbance is induced on a module. In this way the effect of a noisy module on the other modules in the SP chain can be studied. One possibility to do so is to make one of the modules noisy by lowering its threshold. For low thresholds, more noise hits are recorded. This translates in a higher digital activity, and thus in a higher digital current than provided (0.15 A per chip). As explained in Section 4.3.1, the input and output voltages collapse if $I_{shunt} < 0.005$ A, thus switching off the module. As a consequence, the module



(a) (b)

Figure 4.35: (a) Source hit map and (b) spectrum recorded with the MDC2 module of the SP stave using an ^{241}Am source. The dark areas in the hit map are due to the passive components mounted on the module flex which attenuate the gamma rays from the ^{241}Am source. The first two and last four columns of the FE chips are disabled, as they are powered with pads at the top of the FE pixel matrix which are not accessible on module.

current consumption decreases, the voltages across the module can rise again, until the module is fully on. At this point, the digital activity is high, the current consumption increases and the voltages collapse, so that an oscillation of the module voltage and current consumption is induced. The other modules in the chain are tested with threshold and source scans while one of the modules oscillates.

This test requires the use of a second USBPix setup, where the noisy module is connected. To induce oscillations on a module, the threshold value is lowered and a noise occupancy scan is performed, which enables all pixels in the matrix and records noise hits, i.e. hits in the FE in absence of an injected charge, with a random trigger signal sent by the readout system. Once the module is brought to oscillate, the other modules are tested. As already explained, commands are sent in parallel to all modules so that the scans run at the same time on all modules. The threshold scan enables only the pixels in one of six rows. If this mask is applied on the oscillating module, its current consumption decreases and the oscillation stops. By connecting this module to a second USBPix setup, the scans do not run on it, and it keeps oscillating.

The threshold of the non oscillating modules is still tuned to the correct value, with dispersion in a range from 25 e^- to 35 e^-. No significant noise increase over the pixel matrix is observed. As summarized in Table 4.10, the noise increase is within 5 e^-. The spectrum of an ^{241}Am is recorded with an energy of (45.80±5.53)keV, as during standard working condition. These results are in agreement with what was observed in the serial powering proof of principle [55], and thus confirm that noise propagation through the power line is not a concern in a

Table 4.10: Difference in noise value with respect to standard operation, when one module in the SP chain is oscillating.

Noisy module	Module			
	MDC1	MDC2	MDC8*	MDC4
	Noise difference wrt standard operation [e^-]			
MDC2	3.4	-	4.9	3.3
MDC8*	3.4	2.8	-	2.8
MDC4	2.5	3.8	4.0	-

serially powered chain of modules.

4.6 SPP chip characterization

The SPP chip is characterized together with a MDC module, connected as shown in Figure 4.7, to demonstrate its working principle. Figure 4.36 shows the SPP chip wire bonded to a test board. The three structures recognizable on the left side of the chip are the shunt transistors. The SPP chip input voltage is 2.4 V, and when not used it consumes 8 mA, i.e. ~20 mW.

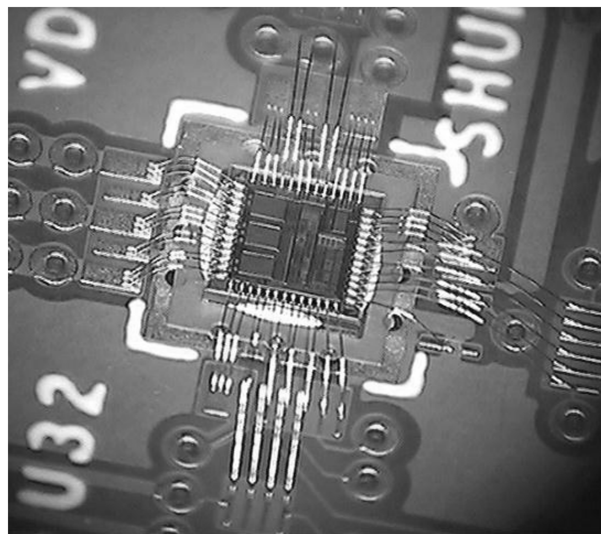


Figure 4.36: Picture of the SPP chip wire bonded on a test board. The three shunt transistors can be seen at the left of the chip [58].

Both the slow control and the over-voltage functionality are checked. The signal to switch on/off the module is sent using an AC-coupled link, over the power line. Figure 4.37a shows the voltage across the chip (top curve) when a command is sent to enable the SPP (bottom

curve). Upon receiving the command, the SPP chip activates its shunting feature, and switches off the module. The input current of 1 A is bypassed by the SPP, and only ~0.02 A flow through the chip. The residual voltage across the module is approximately 0.13 V.

Figure 4.37b shows the voltage across the module in case of over-voltage. The threshold for over-voltage is set to 2.1 V. The current to the module is increased until the voltage at the input of the regulator reaches this value. The SPP activates autonomously, and switches off the module within ~430 ns. Again the module voltage is brought down to ~0.13 V.

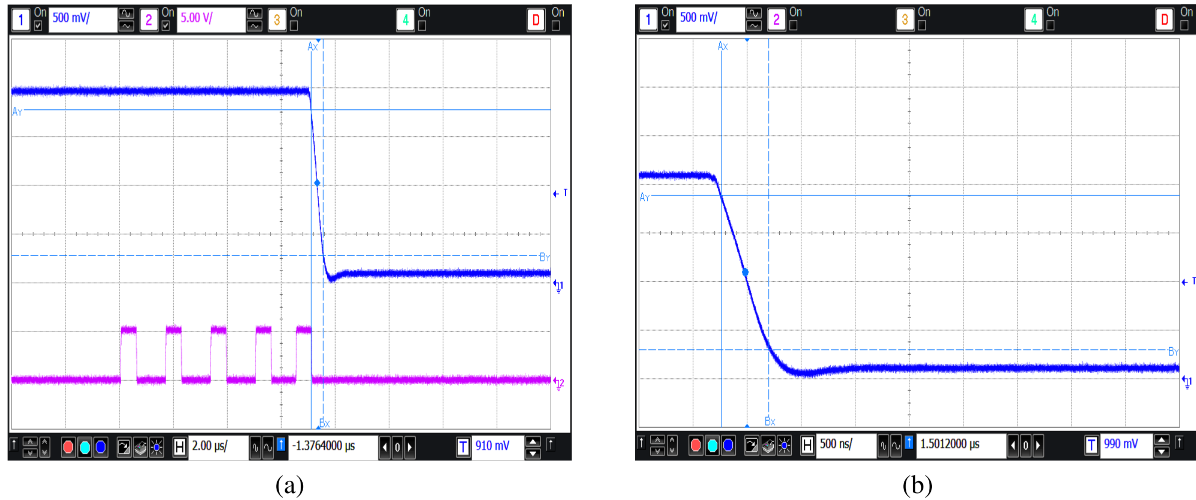


Figure 4.37: (a) Switching off of a serial powering module on command. The top curve is the voltage across the module (0.5 V/div) and the bottom one is the command sent to activate the SPP (5 V/div). The horizontal scale is set to 2 µs/div. (b) Switching off of a serial powering module upon over-voltage. The voltage across the module is shown (0.5 V/div). The horizontal scale is set to 500 ns/div.

To check the module performance when operated in parallel to the SPP, a threshold scan is performed and compared to the result obtained on the module stand-alone. The threshold is tuned to 4000 e^- before connection to the SPP, and the measured threshold, threshold dispersion and noise values are respectively 4014 e^-, 25 e^-, and 136 e^-. When operated with the SPP connected in parallel to the regulators, tuning is maintained, and a threshold of 3933 e^- with a dispersion of 38 e^- is measured. The noise does not change.

These results confirm that the SPP chip working principle satisfies the requirements for a serial powering bypass element. Modules can be controlled independently and protected against over-voltage, and module operation is not disturbed by the bypass element. A dedicated pixel version based on the SPP design can be produced to match the current and voltage specifications of FE-I4 pixel modules.

4.7 Summary and outlook

A serial powering scheme for the ATLAS pixel detector at the HL-LHC is prototyped using FE-I4 modules. On-chip Shunt-LDO regulators provide stable voltages for the module, which

suffer no loss in performance when powered by a constant current, instead of a constant voltage. Excellent performance is obtained with a serial powering stave prototype, where four modules are operated in series. Results of the tests show that modules in a serial powering chain can be operated without faults, with performance as good as for voltage based powering distribution. Despite the series power connection, modules do not influence each other, even in case of a disturbance on one of the modules.

Further tests should be performed on the stave prototype to study the reliability of the SP chain. Failure modes could be studied in more depth, for instance by inducing an oscillation of the voltage across the module with a defined frequency, as performed with FE-I3 modules [51]. A switchable load in parallel to one or more modules can be used to this purpose. In particular the 40 MHz clock frequency, and the 100 kHz frequency of the detector trigger signal should be tested. A different configuration of the AC-coupling scheme should also be validated, in which the decoupling capacitors for CLK and CMD are placed on the readout board at the output of the transmitter, instead that on the flex close to the receiver input. This would be beneficial to reduce further the material in active area.

Although the basic building blocks of a SP chain of modules are developed, further optimization is needed towards a complete serially powered detector:

- the FE chip should provide the capability of measuring the input and output voltages of the regulator, as well as the analog and digital currents, using an on-chip ADC as the one in FE-I4B;
- a trimmable reference resistor for the Shunt-LDO regulator is also desirable. The use of an external resistor, or of an internal resistor with a fixed value, as in the current version, does not provide enough flexibility to adapt to different values of the FE current. A trimmable reference resistor would allow to optimize the regulator I-V curve and shunt current, thus improving the conversion efficiency;
- a dedicated SPP chip should be developed according to the current and voltage specifications of the pixel modules;
- a *HV* switch is needed to provide independent biasing to the sensors in the chain.

The results presented in this chapter show that serial powering is a feasible and reliable solution to power large area hybrid pixel detectors at the HL-LHC, despite the differences and potential point of failures with respect to a voltage based powering scheme. The extra price to pay due to the dedicated developments needed at system level is greatly compensated by the gain in material. Thanks to the lower transmitted current and the flexibility in the choice of the voltage drop on the power lines, a gain in material budget much higher than the number of modules placed in the SP chain can be achieved, making SP the lightest powering scheme. Serial powering can thus bring significant improvement to the resolution of hybrid pixels detectors at the HL-LHC.

Chapter 5

Flip chip of thin, large area FE chips

5.1 Flip chip for hybrid pixel detectors

Flip chip was developed by IBM in the sixties, the so-called Controlled Collapse Chip Connection (C4) [66], to overcome the drawbacks of lateral interconnection technologies, i.e. wire bonding. It is today the most widely used interconnection technology in the semiconductor industry. Flip chip provides a high density vertical connection between a semiconductor device, usually an IC die, and a substrate, such as a Printed Circuit Board (PCB), a carrier, another chip, or a wafer. The connection is achieved using conductive bumps placed on the chip bonding pads. The final assembly sees the chip face-down onto the substrate, opposite to a wire bonding interconnection, where the chip is mounted face-up. Flip chip is a two step process: (1) bump deposition and (2) flip chip assembling. Flip chip processes [67] can be distinguished depending on the bumping and assembling methods used. Most common processes are for instance solder bump, stud bump, plated bump, and adhesive bump flip chip. The different methods offer a wide range of features, from different bump size and pitch to cost, that can cover the needs of a large number of applications.

Hybrid pixel detectors use flip chip technology to connect sensor and readout chip. Two flip chip processes are used for pixel detectors at the LHC, based either on electroplated solder bumping, or on evaporative indium bumping, as described in [68]. These bump bonding processes can provide both the required bump pitch of $50-100\,\mu$m, and density of $\sim$5000 contacts/cm^2. Modules for the ATLAS pixel detector are produced using both flip chip technologies [20]: 45 % of the modules are built using the IZM[1] solder bump process [69] and 55 % with the indium bump process from SELEX[2] [70, 71]. Optical and electrical inspections before and after flip chip, the usage of Known Good Dies (KGD), and the development of reworking procedures allow to obtain a very high yield. Pixel modules are produced with bump defect rates of $\approx 10^{-4}-10^{-5}$ at wafer level and $\approx 10^{-3}-10^{-4}$ after the flip-chip process. Defect modules can be successfully reworked. FE chips can be removed by the application of heat and a pull force, and afterwards a new IC can be flipped onto the sensor, with an almost 100 % probability of reconnecting all pixels. The production yield is between 80 % and 90 %, before any reworking, and well above 90 % after reworking.

[1]Fraunhofer Institute for Reliability and Microintegration.
[2]SELEX Sistemi Integrati (former AMS).

5.1.1 IZM AgSn solder bump flip chip process

The IZM flip chip process [69] used to build modules for the ATLAS pixel detector, uses eutectic solder for the bumps, i.e. AgSn, deposited via electroplating. The process starts with the deposition of an Under Bump Metallization (UBM) on the IC wafer. This step is of fundamental importance to provide good adhesion and long connection lifetime, withstanding mechanical and thermal stresses. The UBM consists of a Ti/W-Cu layer sputtered over the whole wafer surface (plating base), which serves both as diffusion barrier to the Al I/O pads and as a plating base. The electroplating process requires the deposition of a photo-resist mask, structured by lithography to expose the pads, so-called plating mask. The copper base and the eutectic AgSn are then electroplated into the mask. The plating mask and the plating base are removed by stripping and chemical etching. A reflow step follows in which the plated AgSn cylinders are turned into spheres of 25 µm diameter (Figure 5.1). The FE wafer processing concludes with wafer thinning (Section 5.1.3) and FE dicing.

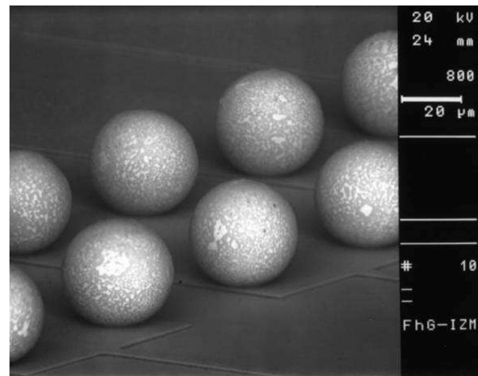


Figure 5.1: Rows of AgSn bump bonds on a FE-I3 wafer after reflow (courtesy of IZM).

The sensor wafer is processed with UBM. The process starts with sputtering of the Ti/W-Cu plating base on the whole wafer, as for the FE electronics. A Cu UBM is then electroplated on the aluminum contact pads using a photo-resist mask, to obtain a solderable metallization of the pixel contacts. Photo-resist and plating base are removed in a wet etching process. After UBM the sensor tiles are diced.

Flip chip assembly is done using a flip chip bonder tool with a pick-and-place accuracy of 3 µm. The chips are tacked to the sensor tiles by solder flux deposited on the sensor surface. A reflow process in a reductive atmosphere mates chip and sensor with thermal cycles reaching a temperature of 260 °C. This high temperature is actually a limiting factor for the solder bump flip chip process, as it may not be acceptable for all mating parts. During reflow the bumps are self-aligned by surface tension, once the parts are brought sufficiently close. To assure bump connectivity, the required chip planarity should be below 15 µm. A plasma cleaning step is performed after flip chip assembly to remove flux residues. The distance between chip and sensor is ~25 µm which minimizes cross talk between electronics and sensor. The connection resistance is smaller than 1 Ω. A cross-section of the AgSn bumps connecting FE and sensor can be seen in Figure 5.2.

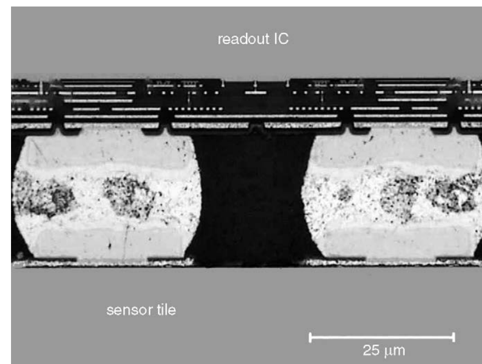
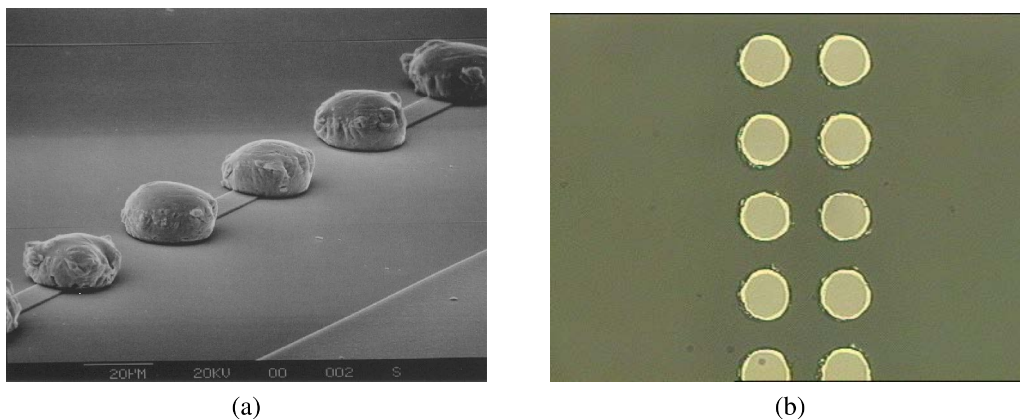


Figure 5.2: Cross-section of an ATLAS pixel module showing the AgSn bump bonds between sensor at the bottom and FE at the top (courtesy of IZM).

5.1.2 SELEX Indium bump flip chip process

SELEX offers a standard indium bumping process[3] [70, 71]. Bumps are deposited on both FE and sensor wafers (Figure 5.3a). The process starts with a wafer cleaning step. A photoresist mask is then deposited on the wafer, leaving the pads exposed to the subsequent process steps: plasma activation of the pads, deposition of a Cr-UBM, and In evaporation. The bump height is limited to 10 µm due to the lift-off process for the removal of the evaporation mask. After bumping, the FE wafer is thinned (Section 5.1.3), and both wafers are diced. Flip chip assembly is done with thermo-compression at a temperature of 80 °C, and a pressure of 25 N/chip (Figure 5.3b). Uniform application of the pressure over the FE area is critical to obtain bump connection. This requires out-of-plane effects given by the parallelism of the dies, the bump uniformity and the planarity of the parts, below 1 µm. The cross-section of the bumps increases by approximately a factor two (20 µm diameter) during flip chip assembly, resulting in a distance between chip and sensor of ~10 µm. The connection resistance is smaller than ~10 Ω.



(a) (b)

Figure 5.3: (a) Indium bump bonds deposited on a glass substrate with a pitch of 50 µm. (b) Flip chip assembly of two glass wafers with In bumps (courtesy of SELEX).

[3]For the CMS pixel detector, a modified indium bumping process is used, as described in [72].

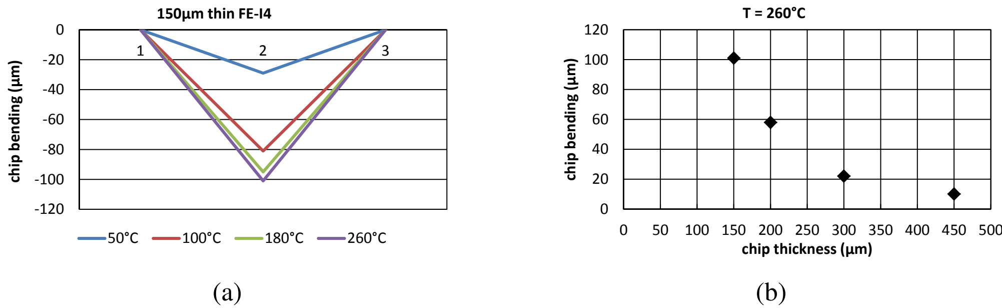
5.1.3 Flip chip of thin FE chips

Thinning of electronics wafers is a standard, well known process in the industry for production of miniaturized electronic components. It is routinely performed with low-cost back side grinding techniques, down to less than 100 µm. Thin FE electronics is particularly important for pixel detectors used in HEP experiments to minimize the material in the path of the particles. The minimum thickness of the FE chip is however constrained by the required handling and processing steps during hybridization. Thin chips are in fact fragile and prone to deformation due to internal stress, arising from the different mechanical and thermal characteristics of the metal layers and the silicon bulk. Techniques such as stress relief and temporary wafer bonding are commonly used in industry during processing of thin wafers or dies to obtain the required planarity.

For the ATLAS modules, the FE-I3 wafers are bumped at full thickness (700 µm) to assure mechanical support during all the handling steps [68]. Wafer thinning is done after bump deposition by back side grinding. To assure that the force applied during thinning is not transmitted to the bumps, a photo-resist layer is used to cover the bumps on the front side for protection. A UV releasing tape holds the wafer in place during thinning. The FE wafers are thinned to about 190 µm and diced immediately afterwards. The choice of the thickness is driven by handling constraints. Thinning and flip chip are not performed in the same facility, involving more manipulations of the thinned dies during shipment. This thickness satisfies as well the requirements of chip planarity to establish bump bond connection, with both the IZM and SELEX processes. The 190 µm thin FE-I3 contributes to 0.27 % X_0 per layer in the ATLAS pixel detector.

To reach the target material budget for the IBL and the ATLAS pixel detector at the HL-LHC, the FE-I4 has to be thinned to lower values than the current FE chip, respectively 150 µm and 100 µm. However, the large area of the FE-I4, five times bigger than the FE-I3 area, complicates further the hybridization process, as larger chips show a higher bending profile when thinned. In addition to this, the FE-I4 has a higher number of metal layers (eight instead of six), a feature that also contributes to a larger chip deformation. The resulting non planarity is particularly critical for the indium bump flip chip process, which demands out-of-plane deformations below 1 µm. For the IZM solder bump process, a bending up to 15 µm can still guarantee full bump connectivity. In this process however, the high temperature reached during reflow (260 °C) increases the chip bowing, due to the different coefficient of thermal expansion of the metal layers and the silicon bulk [73]. A bowing lower than 15 µm at 260 °C is achieved for the FE-I3 with a thickness of ~150 µm. A thickness of 190 µm, is sufficient to keep the bending of the FE-I3 well below 15 µm for the entire temperature range of the reflow process. As shown in Figure 5.4a, the FE-I4 with a thickness of 150 µm bends up to 100 µm during reflow. The necessary planarity for bump connection requires a thickness of 450 µm (Figure 5.4b). The FE-I4 with this thickness would contribute to 0.55 % X_0 per layer, approximately 37 % of the target thickness of the IBL detector.

Dedicated flip chip processes have thus to be developed for the next generation hybrid pixel detectors, in order to enable flip chip of large area FE chips, thinned below 150 µm (FE-I4 and future generation FE chips). In the framework of this thesis, a modified flip chip process is developed in collaboration with IZM. The method is validated for 150 µm FE-I4 chips as



(a) (b)

Figure 5.4: Bending profile of the FE-I4 chip measured on the chip back side at three points along the chip diagonal, as (a) a function of temperature and as (b) a function of thickness. For the AgSn process, constraints on chip planarity are matched with a thickness of 450 µm (courtesy of IZM).

described in Section 5.2. As discussed in Section 5.3, this method is used for the entire IBL module prototyping and production. In addition to IZM, more vendors, as well as alternative interconnection technologies, are under investigation for the HL-LHC upgrade of the ATLAS pixel detector, given the large number of modules to be produced. References on these developments can be found at [74, 43].

5.2 IZM solder flip chip process for thin FE-I4 chips

A temporary handle wafer technique is used to enable flip chip connection of the FE-I4 thinned to $\leq$ 150 µm using the IZM solder bump process [75]. The flip chip process is modified as shown in Figure 5.5. First the electronics wafer is thinned to the desired thickness in a two step process. The silicon bulk is thinned by grinding, and then a stress relief step follows to remove another $8-10$ µm of silicon by dry etching. A support wafer, also called handle or carrier wafer, is bonded on the FE wafer back side to provide mechanical support, and to assure chip planarity during all sequent processing steps. The bonding material is deposited on the support wafer and the connection is achieved with a mix of pressure and temperature. Bump deposition, dicing and flip chip assembly follow the procedure described in Section 5.1.1. After flip chip assembly, as a last step in the process, the support wafer is removed.

The assembly made of electronics wafer, bonding material and handle wafer must resist all process steps, in this case bumping, dicing, flip chip and reflow. The bonding layer needs to provide a voidless bonding interface with good wettability for the support wafer, high thermal stability (maximum temperatures: 150 °C during sputtering, 260 °C during AgSn solder reflow), easy de-bonding process by thermal, chemical or optical release, and leave no residues on the thinned silicon chip. Three different bonding materials are tested with IZM to find the optimal connection between electronics and support wafer [73]. For these investigations, FE-I2[4] pixel wafers are used, thinned to 90 µm, connected to so-called dummy sensors, i.e. a plain silicon tile used as a mechanical replacement of a real sensor. Bump connectivity is checked

[4]The FE-I2 is the prototype of the FE-I3 chip. It has the same functionality, but the design required a further iteration, the FE-I3, to correct some design mistakes.

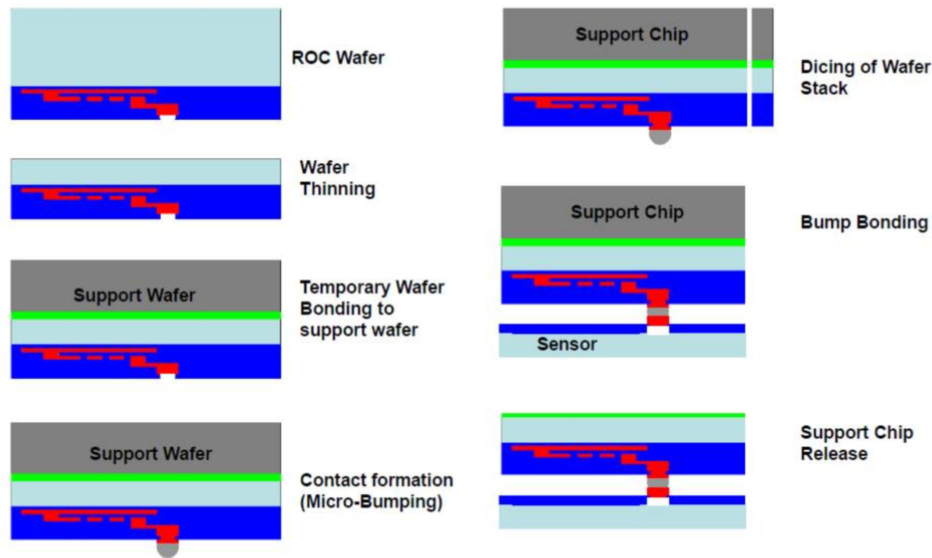


Figure 5.5: Illustration of the IZM modified flip chip process using a temporary wafer bonding technology. The Readout Chip (ROC) wafer is bonded to a support wafer during bump bonding and flip chip. The support wafer is removed only after flip chip assembly (courtesy of IZM).

with optical inspections and electrical tests. To easily recognize unconnected bumps during testing of the assemblies, the Cu plating base is not removed on the dummy sensor. In this way all pixels are shorted together, if connected to the sensor, and their amplifier is in saturation. Only unconnected pixels show a normal behavior.

Two thermo-releasable and one laser-releasable bonding materials are tested. The thermo-releasable materials are a thin film bonding wax and a Brewer WaferBOND HT glue. The handle wafer in both cases is made of silicon. Wafer bonding and de-bonding involve temperatures between 100 °C and 200 °C. Full ATLAS pixel modules are built using 90 µm thin FE-I2 chips with the wax method. With the Brewer method, modules are built with 90 µm thin FE-I2 chips, cut in arrays of 2×1 and 2×2 to mimic a large size FE chip. Unconnected pixels are observed with both methods, suggesting that the FE chips still bends during the reflow process. For the modules built using the wax method, the problem affects bumps along the entire perimeter of the chip. With the Brewer method, as shown in Figure 5.6, unconnected bumps are observed only at the top corners of the chip (opposite the FE periphery), and the problem concerns less than ten bumps in the first two FE columns [73]. The reason for the connectivity problem is found to be the low melting temperature of the bonding material: the bonding interface is not stable during reflow, and the poorly supported, thin FE chip bends.

The tested laser-releasable bonding is polyimide [75]. Wafer bonding is done as for the wax and Brewer methods with a combination of temperature and pressure. The de-bonding requires instead the illumination of the polyimide layer with a laser. As a consequence, the bonding interface between the two wafers does not melt during reflow at 260 °C. The handle wafer in this case is made of glass so that a laser can be scanned over it to dissolve the polyimide layer. Also in this case modules with arrays of 2×1 and 2×2, 90 µm thin FE-I2 chips are built. Observations of the assemblies after flip chip show full bump connectivity in all corners, for

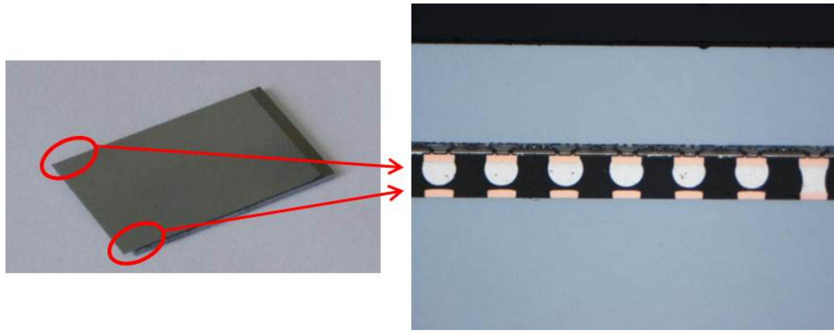


Figure 5.6: Cross-section of a prototype module built using the Brewer WaferBOND HT. The FE chip consists of a 2 ×2 array of FE-I2 chips. A view along the first chip column shows unconnected bumps at the top edges. The round shape of the bumps indicates that connection between FE and sensor is not established during reflow (courtesy of IZM).

both assembly types (Figure 5.7). Full bump connectivity is confirmed as well with electrical tests performed on 2 ×1 modules mounted on rigid PCB boards.

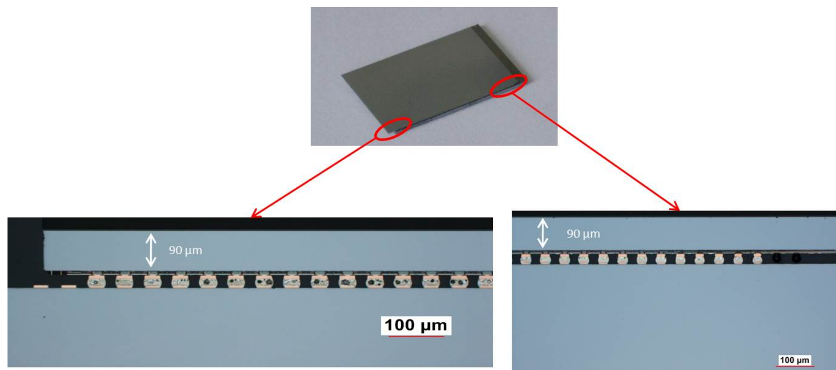
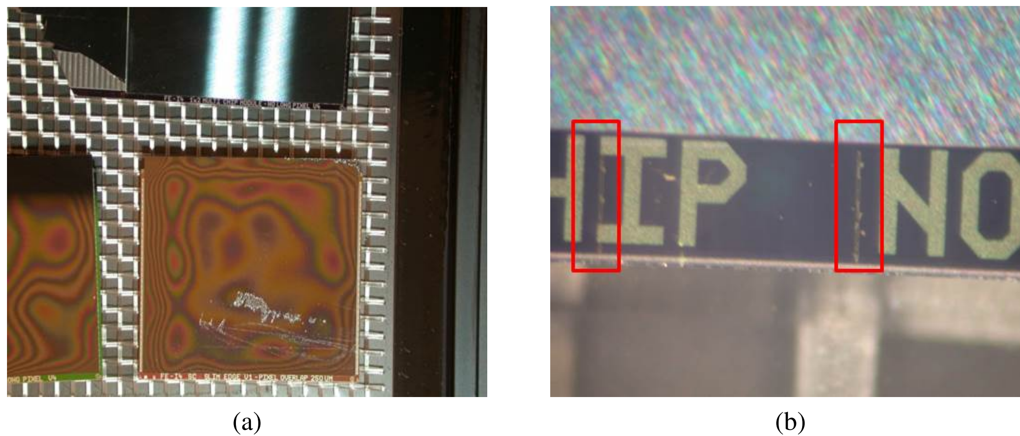


Figure 5.7: Cross-section of a prototype module built using the polyimide method. The FE chip consists of a 2 ×2 array of FE-I2 chips. The zoom-in onto top and bottom bump bonds along the first column shows that bump connection is achieved along the entire chip perimeter (courtesy of IZM).

The outcome of these preliminary investigations indicates that the polyimide bonding material can overcome the drawback of the thermo-releasable bonding materials. A robust connection between FE and handling wafers is provided, keeping the chip flat during all processing steps, from bumping to reflow. It can thus enable flip chip of large ICs using the IZM AgSn process. Nevertheless, process parameters might have to be adjusted depending on the specific application, i.e. depending on the FE area, thickness and metal layer composition. In addition to this, handling of thin chip modules during gluing and wire bonding to a support card, or to a flex hybrid, requires proper mechanical support, optimization of the wire bonding parameters, and dedicated tooling. Experience with 2 ×1 and 2 ×2 thin FE-I2 assemblies shows in fact that wire bonding of thin chips mounted on PCB boards might damage bump bonds close to the periphery of the chip, if the connection between chip and board is not properly done. A uniform layer of hard glue is needed to support the chip during wire bonding, and prevent bending of the peripheral chip area due to the force applied by the wire bonding tool.

5.3 IBL modules prototype and production with 150 µm thin FE-I4 at IZM

The FE-I4, with an area of ~4 cm, is the largest pixel chip in HEP to date, and thus the first to require temporary handling wafer attachment for flip chip, in order to enable thinning to 150 µm and lower. Prototyping and production for the IBL modules are done at IZM with the method described in Section 5.2. Prototypes are built using the FE-I4A chip thinned to 100 µm and 150 µm. Single and double chip planar sensors, and single chip 3D sensors are used. Production modules feature the FE-IB chip thinned to 150 µm connected to MDC planar sensors and MSC 3D sensors. As a real sensor is used, bump connectivity can be checked by illuminating the module with a radioactive source. Unconnected pixels show a lower hit occupancy than connected bumps. For the IBL modules, connected pixels are specified to have a hit occupancy between 0.1 and 4.5 times the mean occupancy over the whole pixel matrix. This test is routinely performed during laboratory tests of prototypes. An alternative test to check bump connectivity is most often used during production and detector integration, where modules are not necessarily accessible for illumination with a source. This method relies on the noise difference measured in threshold scans with and without sensor biasing. A non biased sensor presents a higher capacitance at the input of the readout channel, and thus a higher FE noise is measured. When biasing the sensor, the noise of all connected pixels decreases as the sensor capacitance is lowered. The noise of the unconnected pixels is independent of the sensor bias, and it is lower than for the connected pixels. For the IBL production modules, unconnected pixels are identified as those whose noise changes less then 20 e^- with and without sensor bias.



(a) (b)

Figure 5.8: Visual inspections on IBL module prototypes, showing (a) residues of the polyimide glue on the FE back side, and (b) signs of the laser scanning at the sensor edges.

Visual inspections of the prototype modules show some residues of the polyimide glue on the FE back side (Figure 5.8a), with a thickness of 4-5 µm. Signs of the laser scanning at the edge of the sensor are also visible (Figure 5.8b), which however do not influence the module performance [38]. The measured chip bow is ~10 µm, well below the maximum allowed bow

of 15 µm. Prototype modules are tested electrically, with prototypes MSC modules mounted on test boards, and with all module types assembled with the IBL flex hybrid, to check bump connectivity. Handling during assembling and wire bonding on the test boards or to the flex hybrid can be performed without damage to the modules. For assembling on test boards, the modules are glued on a ceramic support with an uniform glue layer. Dedicated tools are developed for flex assembly and wire bonding, which hold the module flat using a vacuum system. An aluminum plate support with cover is used to support modules during testing.

Results of source scans with an ^{241}Am source on the three prototype module flavors show $<$ 20 disconnected pixels per module. Bump connectivity is confirmed over the entire pixel matrix, confirming successful flip chip and module handling. An example result is shown in Figure 5.9a for a MSC planar module with 150 µm thin FE-I4A. For this module only four unconnected pixels are identified with the source method. To check the mechanical strength of the bump connection, especially at the chip edge, prototype modules undergo thermal cycling at temperatures between −30 °C and 30 °C. A source scan after thermal cycling confirms that no disconnected pixels appear over the whole chip area due to the thermal stress (Figure 5.9b).

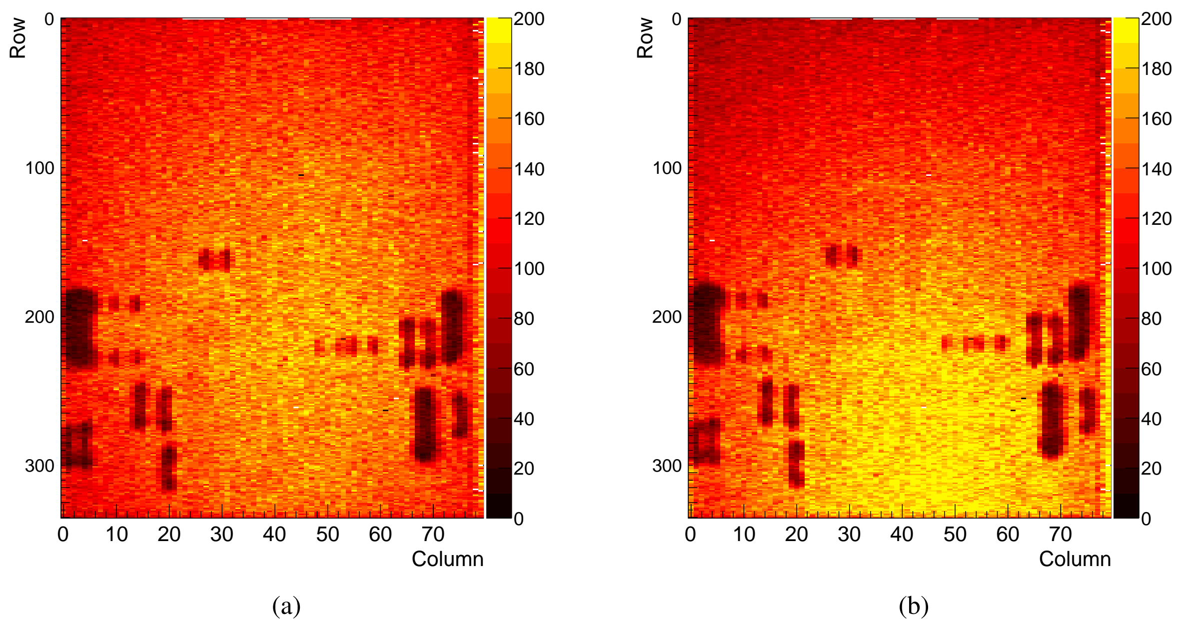
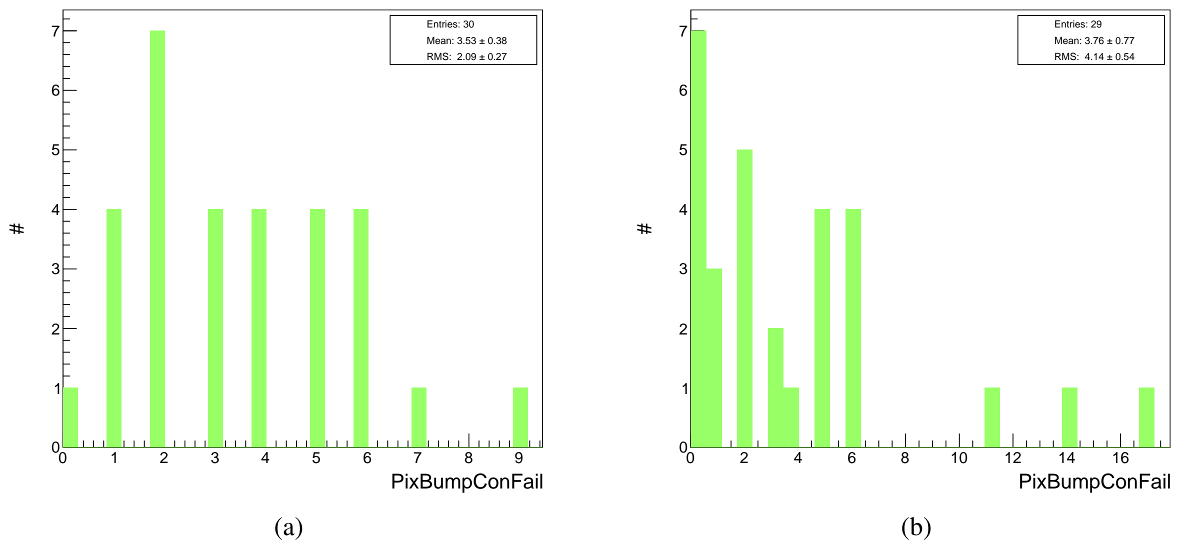


Figure 5.9: Source scan hit map of an IBL MSC planar prototype module with 150 µm thin FE-I4A (a) before thermal cycling and (b) after thermal cycling. Dark areas are due to the passive components mounted on the module flex which attenuate the gamma rays from the ^{241}Am source. The location of the source over the module is visible as an area with higher hit occupancy.

Problems occur with some of the prototype modules with 100 µm thin FE chip. Cracks are visible on the chip back side, possibly due to thermal stress induced by the high laser power on the very thin chip. In some cases the removal of the glass carrier fails. Most likely refraction of the laser light at the diced edges of the glass wafer prevents illumination and dissolution of the polyimide layer. Due to the short IBL development time scale, the investigation of these issues is postponed and it will be part of R&D for Phase-II upgrade. Prototyping is completed with

150 µm thin chips, a thickness matching the IBL requirements, in order to start production on time.

At the time of writing production of IBL modules is ongoing. Approximately 180 MDC planar and 120 MSC 3D modules have been produced, corresponding to flip chip of almost 500 FE-I4B chips. Results of module testing do not show connectivity problems due to the temporary handle wafer technology, i.e. unconnected bumps are not observed along the chip edges on any of the modules. No issues are encountered during laser release. Serious bump problems affect the first three processed module batches, due to an issue with the solder flux. Modules from the last two batches are produced without the use of solder flux during reflow, showing very good bump connectivity. Figure 5.10a shows a histogram of unconnected bumps after module assembly over 30 MDC planar modules, i.e. for 60 FE chips, using the threshold scan method. On average three pixels are disconnected per module. After 10 thermal cycles between −40 °C and 40 °C (Figure 5.10b), keeping the modules for one hour at each temperature, no significant increase of disconnected pixels is observed. The difference in the histogram is due to the different temperatures at which the modules are tested (room temperature after assembly and around −15 °C after thermal cycling), which in turn affects slightly the noise measured during the threshold scan.



(a) (b)

Figure 5.10: Histogram of unconnected bumps over 30 IBL MDC planar modules (a) after module assembling and (b) after burn-in tests, i.e. thermal cycling.

5.4 Summary and outlook

The IZM AgSn flip method for thin ICs using a temporary handle wafer technology is successfully developed and validated. It is currently used for the IBL module production with 150 µm thin FE-I4B chips. Despite its larger area and increased number of metal layers, the FE-I4 is

thinner than the FE-I3 (150 µm vs. 190 µm), and adds only 0.18 % X_0, instead of 0.27 % X_0. Without this development, the FE-I4 thickness would have been 450 µm, introducing 0.55 % X_0. Thus the usage of a temporary handle wafer technique allows to remove 300 µm of silicon, saving 0.37 % X_0 for the IBL layer.

For the HL-LHC upgrade, requiring an even lower FE thickness of 100 µm, the method needs to be tuned further. In particular, the causes for the problems encountered with flip chip of some of the 100 µm thin FE-I4 chips have to be investigated in more detail. Parameters for the bonding wafer release need to be adjusted accordingly. Laser power needs to be tuned, and a different dicing procedure should be tested to improve cutting through the handle wafer to yield clean borders and thus avoid refraction of laser light. In addition to this, a prototyping campaign with the FE-I4 should be performed to find the minimum achievable FE thickness for flip chip with this method.

Chapter 6

Module concepts with Through Silicon Via technology

6.1 TSV for hybrid pixel detectors

The processing of Through Silicon Vias [76] is of fundamental importance for the success of 3D integration, where TSV provide a high density vertical interconnection for the stacking of different electronic tiers. TSV processing through the silicon allows routing the electrical contacts from the front side, to the back side of an IC die. Chips can then be connected vertically, forming a multilayer electronic stack, with high functionality. A standard TSV process is made of four steps: wafer thinning, via etching, via passivation, and via metallization. TSV processes can be differentiated depending on the stage of the CMOS process at which via formation takes place. Via first TSV are formed before FEOL[1], on the unprocessed silicon wafer. Via last TSV, also called post-processing TSV, are formed on the fully processed CMOS wafer, i.e. after BEOL[2].

In the framework of HEP experiments, Through Silicon Via technology can be exploited to develop new module concepts for hybrid pixel detectors. Both via first and via last TSV processes are considered for the upgrades of the ATLAS pixel detector at the HL-LHC [33]. The via first technology is considered to build 3D FE electronics for the innermost detector layers, featuring small pixel size, increased memory and functionality, to cope with the high hit rate. Post-processing TSV can be used to design compact, low mass module concepts using existing hybrid pixel technology for the outer layers. From a technological point of view, via first TSV processes present still more challenges than the via last ones. In this case in fact the process is more complex, as it requires to include via formation in the CMOS process itself, as well as handling and alignment of thin wafers for stacking of the different electronic tiers. Via last TSV, although not yet mature enough for mass production, is already an established technology, offered by a number of packaging companies.

The work presented in this thesis is carried out in collaboration with IZM, and concentrates on the development and validation of a via last TSV process on the ATLAS pixel detector Front-End electronics [77]. A via last TSV process offered by IZM is selected, and prototype

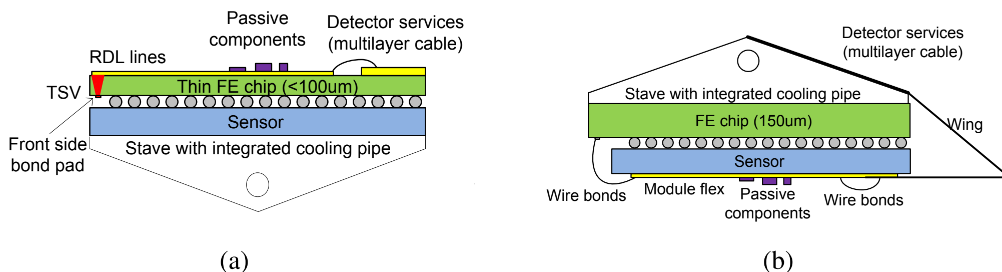
[1] Front End Of Line (FEOL) is the step in the IC fabrication process where transistors are formed on the silicon wafer.

[2] Back End Of Line (BEOL) follows FEOL in the IC fabrication process. In this process step, the individual devices are connected using metal interconnect layers.

modules with TSV are built using 90 µm thin FE-I2 electronics. The process and the results of TSV module characterization are discussed in Section 6.2 and Section 6.3 respectively.

6.1.1 Pixel module concepts with via last TSV

Post-processing TSV can be processed on existing pixel FE electronics. In this application, vias are etched in the peripheral bond pads to bring the electrical contact to the module on the FE back side. Here signals can be routed, using a so-called Redistribution layer (RDL), for connection to the detector services. Passive components can be bonded on the RDL lines. Figure 6.1 shows the sketch of a possible module concept with via last TSV, compared to an IBL module, representing state-of-the-art hybrid pixel detectors. As the front side pads are not used for wire bonding, the height of the periphery can be reduced, leading to a reduction of the inactive area of the chip. The FE thickness is as well reduced as thinning is of fundamental importance for TSV processing, to reduce time and cost of the process. The flex hybrid is not needed anymore, as it is replaced by the RDL. Detector services can run directly on the back side of the modules and be wire bonded to the RDL, resulting in a less bulky and easier interconnection. No wings bending over the stave edge (Figure 6.1b) are needed as in the IBL. Finally, the module is loaded on the stave in a sensor down configuration, beneficial for cooling of the sensing element. One open question is however the connection of the HV to the biasing electrode on the sensor back side. This requires a modification of the sensor design, for instance to route the HV over one edge of the sensor from the front to the back side. Via last TSV can enable compact, low mass hybrid pixel detectors, with minimal modification to the traditional FE electronics design, and using standard CMOS technology. Using a TSV compliant design of the FE chip as well as of the sensor, 4-side abuttable module concepts can be achieved.



(a) (b)

Figure 6.1: Sketch of (a) a possible TSV module concept and (b) of an IBL module.

Modules featuring via last TSV can be used for the outermost pixel layers at the HL-LHC, where tileable modules are beneficial to provide full coverage over the large detector area. In addition to this, as via last TSV can be applied to mature hybrid pixel technologies, the added cost of TSV formation is not prohibitive for the large number of modules to be produced. FE-I4 chips featuring TSV can be used to build 4-chip modules. Using TSV, the height of the FE-I4 periphery could be reduced by 0.6 – 1 mm [78]. The peripheral area would thus decrease to respectively 8 % and 6 % of the FE total area, with respect to the current 11 %. Considering

the 4-chip module concept described in Section 3.2.3, the contribution of the FE periphery to the inactive area of the module would decrease from 10 % to 7 % and 5 %. Together with the required FE thinning down to less than 100 µm, would lead to a FE contribution to the thickness of the detector of less than 0.09 % x/X_0. As lines can be routed directly on the FE back side without the need of a flex, some material can be saved, corresponding to the kapton layers of the flex. As quoted in Section 3.2.3, the flex hybrid for the 4-chip modules at the HL-LHC would account for 0.19 % x/X_0. The contribution from the RDL can be estimated to be ~0.10 % x/X_0. Around 0.15 % x/X_0 can be gained also in the connection to the services, as this can be done using wire bonds, without the need of a flex extension with connector.

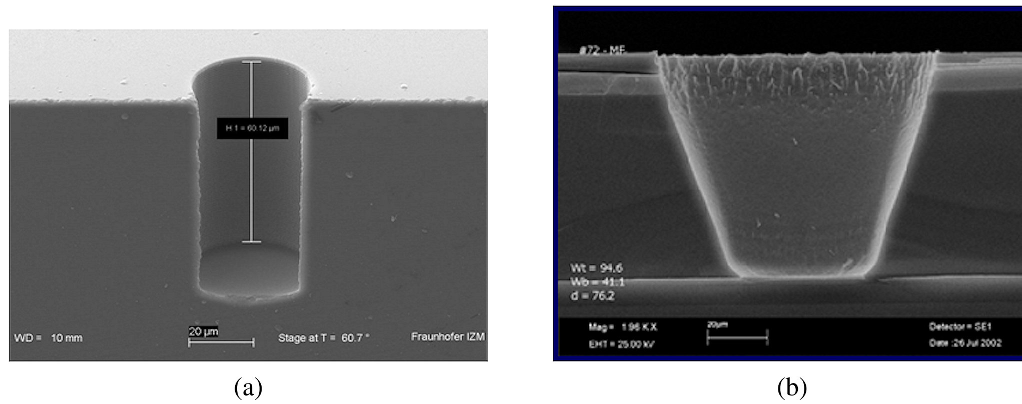
6.2 IZM via last TSV processing on ATLAS readout wafers

Two via last TSV processes developed by IZM are investigated for prototyping with the ATLAS pixel FE electronics, namely straight side wall and tapered side wall TSV. The first makes use of a standard Bosch process [79], in which etching and passivation steps are alternated. Tapered side wall vias are instead etched in one step and passivation is deposited at the end of the etching process. Both processes have been demonstrated on monitor wafers and typical via profiles are shown in Figure 6.2. In the case of via last TSV, the via diameter can be relaxed up to the bond pad size, typically 100 – 150 µm in HEP pixel chips, provided process parameters allow for these via features. The processing time and the available tools determine in fact the achievable quality of the via shape, as well as the homogeneity in the deposition of the isolation layer, on the via side wall and bottom, and of the plating base, across the entire wafer. This in turn sets limits on the aspect ratio that the process can be tuned for. In the straight profile TSV process studied here, the achievable aspect ratio is from 2 : 1 to 5 : 1, where the standard parameters are 20 µm via diameter and 100 µm depth. In the tapered via process, side wall angles of about 60 – 70° can be achieved. With a maximum via top diameter set in this case by the bond pad pitch (approx. 150 µm), the via depth is in the order of 80 – 150 µm, with via bottom diameters of 30 – 40 µm. Current developments at IZM aim at 300 µm depth for a 150 – 200 µm top diameter.

The tapered side wall process is selected to prototype demonstrator TSV modules with the ATLAS FE-I2 electronics, as it allows for a simpler deposition process of the isolation layer with respect to straight side wall TSV, using thin film polymers. A batch of two FE-I2 wafers is processed. The via formation process on these wafers requires both a front side and a back side processing. In this FE design, only the last metal layer is present in the bond pad. Front side processing is thus needed to connect the aluminum bond pad on the front side to the via bottom, as the BEOL SiO_2 layer of 8 µm would be difficult to etch through the via opening (Figure 6.4). A so-called plug is used for the connection. The plug formation is done first, before wafer thinning. The aluminum pad is opened by a wet etching process. The BEOL SiO_2 is then etched up to the Si interface, and the plug is formed by electroplating of copper. Figure 6.3 shows the front side processing steps on the processed ATLAS FE-I2 wafers.

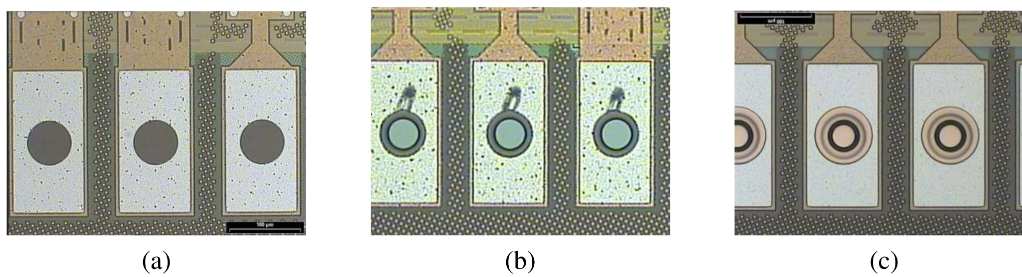
After completion of the plug, the back side processing starts, i.e. the actual process of via formation [80, 81], from wafer thinning to RDL, including:

- carrier bonding on the front side for handling and mechanical support of the thinned



(a) (b)

Figure 6.2: Example of (a) straight and (b) tapered side wall TSV on a monitor wafer. The straight profile TSV has a depth of approximately 60 µm and an opening of 30 µm. For the tapered TSV, bottom and top via diameters are respectively of 41 µm and 95 µm, the side wall angle is 70°, and the wafer thickness is 77 µm (courtesy of IZM).



(a) (b) (c)

Figure 6.3: Front side processing on the ATLAS FE-I2 wafers: (a) Al pad opening, (b) BEOL SiO_2 etch, (c) plug formation (courtesy of IZM).

wafer. For this a perforated glass carrier with polymer bond is used;

- Si back side thinning. The target thickness for the processed wafers is 90 µm;
- Si etch profile for tapered side wall vias using DRIE (Deep Reactive Ion Etching);
- via side wall passivation using a PECVD (Plasma Enhanced Chemical Vapor Deposition) oxide layer and a polymer layer;
- opening of the passivation on the via bottom;
- barrier seed layer deposition (sputter deposition of TiW/Cu plating base);
- via coverage by Cu electroplating;
- patterning of back side RDL and bond pad preparation. The metal used is gold in order to have wire bondable pads;
- bonding of carrier on the wafer back side for further processing (i.e. bumping);
- solvent based de-bonding of the carrier on the front side.

Figure 6.4 shows the tapered via profile on the FE-I2 wafers, together with the plug on the front side. Top and bottom via diameters are approximately 45 µm and 110 µm.

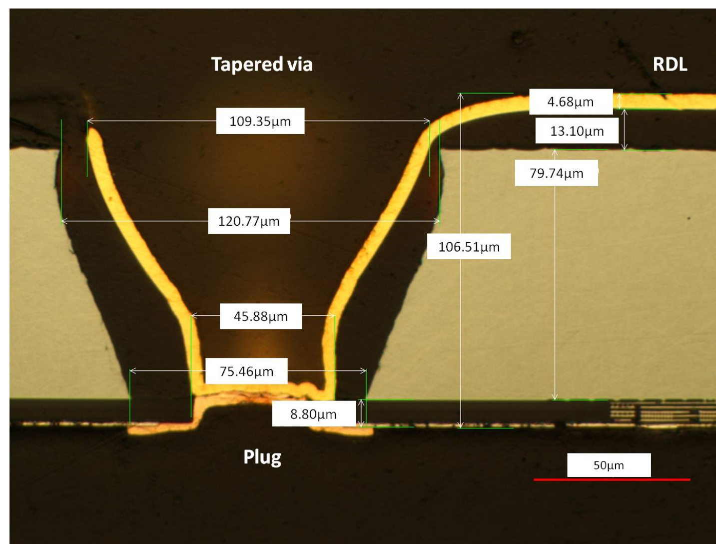


Figure 6.4: SEM (Scanning Electron Microscope) picture of a tapered profile TSV and plug on the FE-I2 wafers (courtesy of IZM).

The RDL and bond pad structure on the FE back side are shown in Figure 6.5. For this first demonstrator TSV modules, a simple RDL pattern is chosen which reproduces the front side bond pads. The bond pads on the back side are located on top of the pixel matrix to be supported by the sensor during wire bonding of the modules on a PCB for testing.

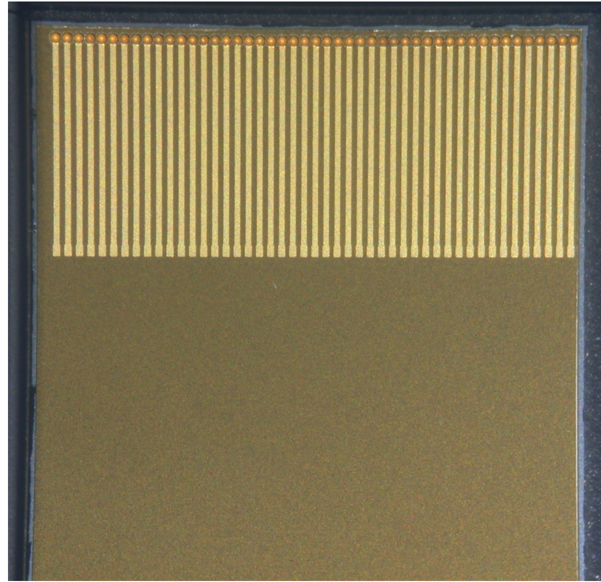


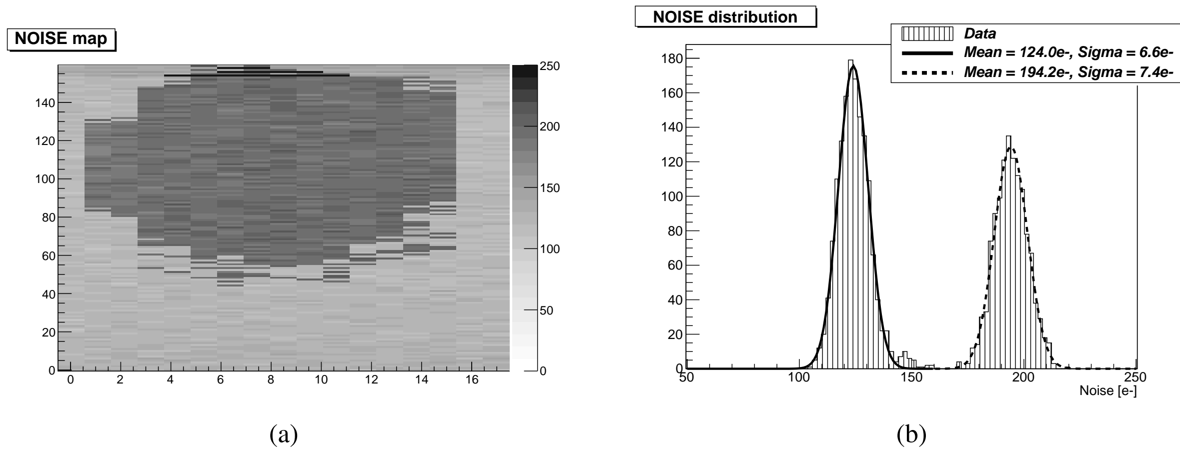
Figure 6.5: RDL and wire bond pads on the FE-I2 back side.

The wafer is then processed for bump bonding to a sensor tile for module assembly. For this TSV proof of principle demonstration, where the main interest is the validation of the TSV process only, the standard IZM flip chip process described in Section 5.1.1 is used, without temporary wafer bonding. It is therefore expected that the thin FE bends up at the corners due to the high temperature used in the reflow process, and that, as a consequence, only an area in the center of the pixel matrix is connected to the sensor. The combination of the two processes, via formation and flip chip of thin ICs using the method described in Section 5.2, requires a further demonstration step. The topography on the TSV chip back side (i.e. the RDL) sets constraints on the thickness and the uniformity of the polyimide bonding layer between the electronics wafer and the handle wafer, possibly requiring an optimization of the process parameters.

6.3 Results of electrical tests on demonstrator TSV modules

The TSV modules are wire bonded onto a dedicated PCB, using the bond pads on the FE back side, i.e. the TSV connection is used to operate the module. Noise maps recorded with threshold scans show two areas with different noise values, indicating connected pixels at the center of the module, and unconnected pixels along the perimeter of the FE, as expected. Figure 6.6a shows a typical noise map obtained on a TSV module. Noise figures for both connected and unconnected pixels are approximately $190-200\,\mathrm{e}^-$ and $120-130\,\mathrm{e}^-$ (Figure 6.6b), respectively, and they compare well with the noise figures from FE-I3 bare chips and modules, operated with the standard front side connection [28, 18].

The untuned threshold values are within $4000-5000\ e^-$ with a dispersion around $700\ e^-$.

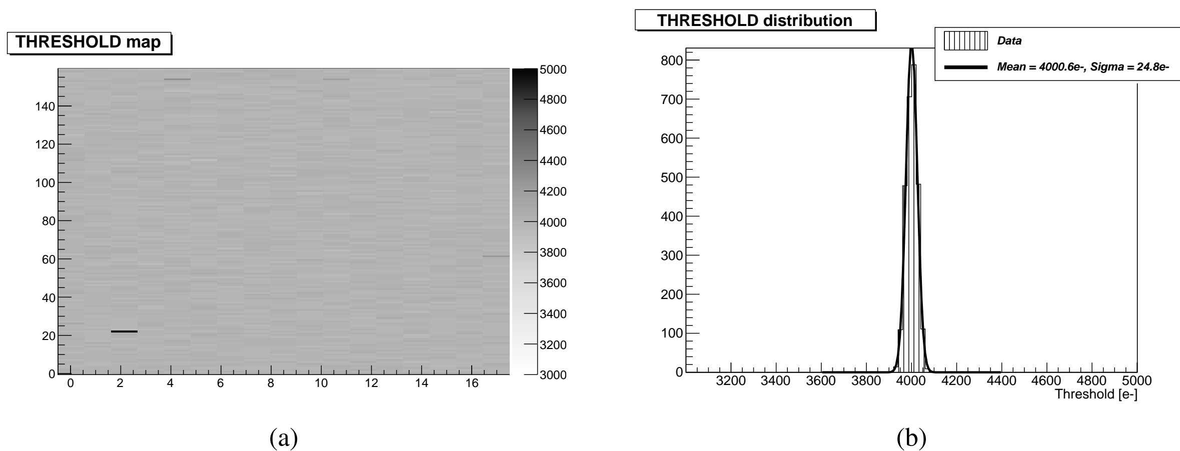


(a) (b)

Figure 6.6: (a) Noise map and (b) noise distribution of a TSV module. Connected pixels, in the center of the pixel matrix, have higher noise as the connection to the sensor increases the input capacitance seen by the amplifier.

The results of a module tuned to 4000 e^{-} is shown in Figure 6.7. After tuning, the threshold value is uniform across the pixel matrix with dispersion reduced to $20-30$ e^{-}. TSV modules can be operated with threshold values down to 2300 e^{-} (Figure 6.8), in agreement with what observed on ATLAS pixel FE-I3 modules, with smaller dispersion (75 e^{-} vs. 125 e^{-} [28]).

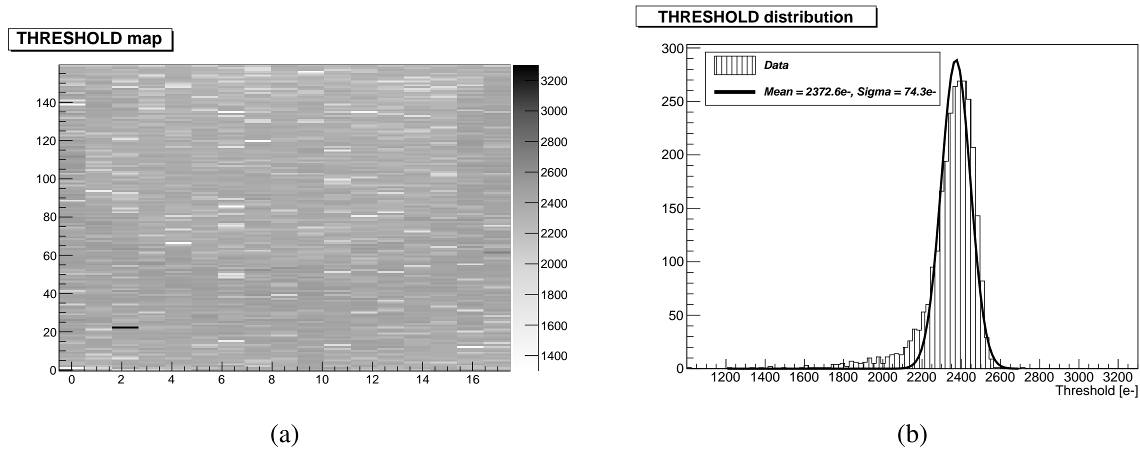
To assess full module operation, the spectrum of a radioactive source (60 keV gamma rays) is recorded. The spectrum is shown in Figure 6.9 demonstrating that the TSV processing is successful throughout, such that the operation of the module suffers no loss in performance. The 59.54 keV peak of the ^{241}Am source is correctly recorded, at a charge of (18.2±1.0) ke^{-}[3], corresponding to an energy (65.52±3.6)keV.



(a) (b)

Figure 6.7: (a) Threshold map and (b) threshold distribution of a TSV module tuned to 4000 e^{-}.

[3]Due to systematics, the error on the measured charge is about 1000 e^{-} [51]



(a) (b)

Figure 6.8: (a) Threshold map and (b) threshold distribution of a TSV module tuned to 2300 e^{-}.

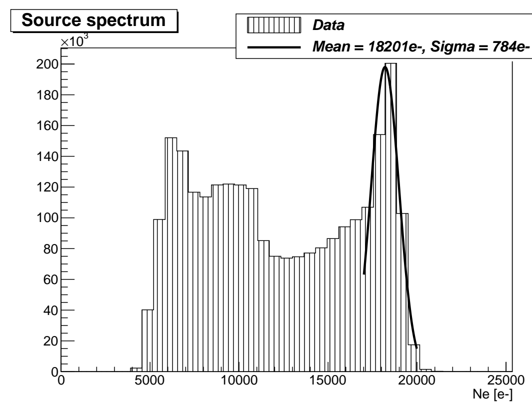


Figure 6.9: Spectrum of an ^{241}Am source recorded with a demonstrator module with TSV (no clustering applied).

6.4 Summary and outlook

A via last TSV process is successfully demonstrated on FE-I2 electronics. The usage of this process can enable the design of compact module concepts with easy interconnection to the detector services, with higher ratio of active over inactive area. Four-chip modules for the outer layers of the ATLAS pixel detectors at the HL-LHC, featuring 90 µm thin FE-I4 chips, with TSV and RDL, and wire bond connection to the detector services, could have ~0.26 % X_0 less than the baseline layout.

Processing of FE-I4B wafers with tapered TSV is ongoing at IZM. For this IC, the front side processing step is not needed, as the full stack of metal layers is available in the pad, and the first metal layer is connected to the last. The few µm of BEOL SiO_2 and polysilicon between the via bottom and the first metal layer can be etched through the via opening. First module assemblies are expected by the end of the year (2013), including the validation of the IZM flip chip method for large, thin ICs with TSV and RDL.

Chapter 7
Conclusions

Efficient momentum and vertex resolution is achieved at the LHC with silicon tracking and vertex detectors. The latter are implemented as hybrid pixel detectors placed as close as possible to the beam pipe, and play a crucial role in reconstructing secondary vertices. Trackers at the LHC are capable of resolving ~1000 tracks per bunch crossing of 25 ns, and associate them to on average 23 vertices distributed over ~16 cm, in a high radiation environment. These achievements required the development of dedicated silicon sensor designs and read-out electronics, as well as services and mechanics, to meet the requirements on granularity, speed, radiation hardness and material budget.

Silicon trackers are foreseen for the high luminosity upgrades of the LHC experiments. The operating condition at higher than design luminosity poses even more stringent requirements on the design of trackers, in particular for the vertex detector. The use of hybrid pixel detectors at the HL-LHC needs a further technological effort to investigate different sensor concepts and material, FE technologies, structural material and services scheme.

In this framework, the R&D presented in this thesis has focused on technologies to reduce the material budget of hybrid pixel detectors for operation during Phase-I and Phase-II, targeting the upgrades of the ATLAS pixel detector. The required material budget reduction goes form 3.6 % X_0 per layer of the current detector, to ≤ 1.5 % X_0 of the IBL and HL-LHC upgrades. An alternative powering distribution scheme, serial powering, is investigated to reduce the thickness of the detector services. Two interconnection technologies, namely solder bump flip chip and Through Silicon Via, have been developed in collaboration with IZM to reduce the material budget of the detector modules.

The thickness of the detector services can be greatly reduced by improving the powering scheme of hybrid pixel detectors. Power distribution at low current and high voltage is mandatory for trackers at HL-LHC not only to meet the requirements in terms of material budget, but as well to match the constraints from cable routing, and to reach high power efficiency. In this context serial powering represents a viable alternative to commonly used voltage based powering schemes. By powering groups of modules with a constant current over a serial connection, the transmitted current can be reduced and the voltage drop on the cables can be optimized, resulting in very high material reduction and increased power efficiency. A serial powering scheme for the outer layers of the ATLAS pixel detector at the HL-LHC is fully prototyped with FE-I4 modules. These development includes, a new regulator concept, a bypass element, and a dedicated AC-coupling scheme. Operation of modules in a serial powering chain is as good as in voltage based powering schemes. No significant influence of the modules on each

other is observed. Thanks to the results of this work, the proposed serial powering scheme is the baseline powering option for all prototyping activities for the Phase-II upgrade of the ATLAS pixel detector based of FE-I4 modules. Compared to a direct powering scheme, the proposed SP scheme would reduce the material budget of the on-stave power cables by at least a factor 4, considering only the reduction of transmitted current. By optimizing the voltage drop on the cables, the reduction factor could be increased up to 30. The corresponding reduction of the thickness of the services in the active area would be of $0.09-0.12\,\%\ X_0$. The detector power efficiency could reach 65 %, almost three times higher than in the present ATLAS pixel detector.

A modified flip chip technology is developed to allow high density interconnection of sensors with the large area, next generation hybrid pixel FE chips, with low thickness. A temporary wafer bonding technology is used during the entire flip chip to control the FE planarity, in particular during reflow at high temperature. The developed technology allows flip chip of $4\,\text{cm}^2$ large FE-I4 chips with a thickness of only 150 µm, and it is used for the entire IBL module prototyping and production. The FE thickness is only $0.18\,\%\ X_0$, approximately 33 % less than in the current ATLAS pixel detector. The amount of material saved using this modified flip chip technology is $0.37\,\%\ X_0$, i.e. 25 % of the IBL target material budget.

The FE material budget can be further reduced employing via last TSV technology. Using this post-processing, vertical interconnection technology to route the electrical connection on the chip back side, both the FE periphery and thickness can be reduced. In addition to this, the connection of the module to the detector services is less bulky and easier than in today's hybrid pixel detectors. Via last TSV interconnection can thus enable the development of low mass, compact hybrid pixel detectors, using standard CMOS technologies. ATLAS pixel FE electronics is successfully processed with a tapered side walls via process. No loss in performance is observed with modules operated using the TSV connection with respect to the standard front side operation. Application of this process to the development of FE-I4 4-chip modules for the ATLAS pixel detector at the HL-LHC would reduce the module thickness from 0.61 % X_0 as in the baseline design to $0.35\,\%\ X_0$.

Although at a very mature stage already, the proposed technologies can be further tuned and optimized towards the HL-LHC upgrade. The technology for flip chip of large area, thin FE chips can potentially allow for lower FE thickness, such as the 100 µm target thickness for the HL-LHC. The full potential of the demonstrated TSV technology could be reached with 4-side abuttable FE and sensor designs, with a dedicated HV bias solution. For serial powering, improvements in the regulator design and AC-coupling scheme would allow to reach even higher power efficiency and lower mass. These developments could then finally converge to a prototype ATLAS pixel outer layer stave for the HL-LHC upgrade featuring serially power distribution, and 4-chip modules with 90 µm thin FE-I4 chip with TSV. Such a detector concept would have, thanks to the R&D of this thesis, a material budget of $\sim 1.3\,\%\ X_0$ per layer, well within the target value, i.e. $\sim 0.8\,\%\ X_0$ less than with direct powering, 450 µm thick FE, and no TSV.

Bibliography

[1] [Online]. Available at: http://www.atlas.ch/photos/.

[2] L. Evans and P. Bryant. 'LHC Machine'. In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[3] S. Catani. *Aspects of QCD, from the Tevatron to the LHC*. 2000. arXiv: hep-ph/0005233 [hep-ph].

[4] ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[5] [Online]. Available at: https://twiki.cern.ch/twiki/bin/view/AtlasPublic.

[6] ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Tech. rep. Geneva: CERN, 1999.

[7] [Online]. Available at: http://cds.cern.ch.

[8] ATLAS Collaboration. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

[9] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. Geneva: CERN, 2010.

[10] ATLAS Collaboration. *Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment*. Tech. rep. CERN-LHCC-2011-012. LHCC-I-020. Geneva: CERN, 2011.

[11] ATLAS Collaboration. *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*. Tech. rep. CERN-LHCC-2012-022. LHCC-I-023. Draft version for comments. Geneva: CERN, 2012.

[12] W.M. Yao, C. Amsler, D. Asner et al. 'Review of Particle Physics'. In: *J. Phys.* G33 (2006), pp. 1–1232. DOI: 10.1088/0954-3899/33/1/001.

[13] V.L. Highland. 'Some Practical Remarks on Multiple Scattering'. In: *Nucl. Instrum. Meth.* 129 (1975), pp. 497–499. DOI: 10.1016/0029-554X(75)90743-0.

[14] G.R. Lynch and O.I. Dahl. 'Approximations to multiple Coulomb scattering'. In: *Nucl. Instrum. Meth.* B58 (1991), pp. 6–10. DOI: 10.1016/0168-583X(91)95671-Y.

[15] ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics*. Geneva: CERN, 2009.

[16] R.L. Gluckstern. 'Uncertainties in track momentum and direction, due to multiple scattering and measurement errors'. In: *Nucl. Instrum. Meth.* 24 (1963), pp. 381–389. DOI: 10.1016/0029-554X(63)90347-1.

[17] F. Ragusa and L. Rolandi. 'Tracking at LHC'. In: *New J. Phys.* 9 (2007), p. 336. DOI: 10.1088/1367-2630/9/9/336.

[18] J. Grosse-Knetter. 'Vertex Measurement at a Hadron Collider - The ATLAS Pixel Detector'. BONN-IR-2008-04. PhD thesis. 2008.

[19] N. Wermes. *Pixel vertex detectors*. 2006. arXiv: physics/0611075 [physics].

[20] G. Aad, M. Ackers, F.A. Alberti et al. 'ATLAS pixel detector electronics and sensors'. In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007.

[21] H.C. Kaestli, W. Bertl, W. Erdmann et al. 'CMS barrel pixel detector overview'. In: *Nucl. Instrum. Meth.* A582 (2007), pp. 724–727. DOI: 10.1016/j.nima.2007.07.058.

[22] A. Kluge, G. Aglieri Rinella, G. Anelli et al. 'The ALICE silicon pixel detector'. In: *Nucl. Instrum. Meth.* A582 (2007), pp. 728–732. DOI: 10.1016/j.nima.2007.07.152.

[23] L. Rossi. 'Pixel detectors hybridisation'. In: *Nucl. Instrum. Meth.* A501 (2003), pp. 239–244. DOI: 10.1016/S0168-9002(02)02041-7.

[24] M.S. Alam, A. Ciocio, K. Einsweiler et al. 'The ATLAS silicon pixel sensors'. In: *Nucl. Instrum. Meth.* A456 (2001), pp. 217–232. DOI: 10.1016/S0168-9002(00)00574-X.

[25] R. Wunstorf. 'Radiation tolerant sensors for the ATLAS pixel detector'. In: *Nucl.Instrum.Meth.* A466 (2001), pp. 327–334. DOI: 10.1016/S0168-9002(01)00568-X.

[26] I. Peric, L. Blanquart, G. Comes et al. 'The FEI3 readout chip for the ATLAS pixel detector'. In: *Nucl. Instrum. Meth.* A565 (2006), pp. 178–187. DOI: 10.1016/j.nima.2006.05.032.

[27] R. Beccherle, G. Darbo, G. Gagliardi et al. 'MCC: The Module Controller Chip for the ATLAS pixel detector'. In: *Nucl. Instrum. Meth.* A492 (2002), pp. 117–133. DOI: 10.1016/S0168-9002(02)01279-2.

[28] J. Weingarten. 'System test and noise performance studies at the ATLAS pixel detector'. CERN-THESIS-2008-033, BONN-IR-2007-10. PhD thesis. 2007.

[29] S. Coelli. 'Mechanics and infrastructure for the ATLAS Pixel Detector'. In: *Nucl. Phys. Proc. Suppl.* 172 (2007), pp. 280–283. DOI: 10.1016/j.nuclphysbps.2007.08.012.

[30] D. Attree, B. Anderson, E.C. Anderssen et al. 'The evaporative cooling system for the ATLAS inner detector'. In: *JINST* 3 (2008), P07003. DOI: 10.1088/1748-0221/3/07/P07003.

[31] M. Garcia-Sciveres. *ATLAS Experiment Pixel Detector Upgrades*. 2011. arXiv: physics.ins-det/1109.4662 [physics.ins-det].

[32] A. Mekkaoui, M. Garcia-Sciveres and D. Gnani. 'Results of 65-nm pixel readout chip demonstrator array'. In: *JINST* 8 (2013), p. C01055. DOI: 10.1088/1748-0221/8/01/C01055.

[33] D. Arutinov, M. Barbero, L. Gonella et al. 'Experience with 3D Integration Technologies in the Framework of the ATLAS Pixel Detector Upgrade for the HL-LHC'. In: *Nucl. Instrum. Meth.* (2013). In Press. DOI: 10.1016/j.nima.2013.04.044.

[34] D. Dobos and H. Pernegger. 'Diamond pixel modules and the ATLAS beam conditions monitor. Collaborations'. In: *Nucl. Instrum. Meth.* A628 (2011), pp. 246–250. DOI: 10.1016/j.nima.2010.06.328.

[35] P. Weigell. *Recent Results of the ATLAS Upgrade Planar Pixel Sensors R&D Project*. 2012. arXiv: physics.ins-det/1210.7661 [physics.ins-det].

[36] A. Macchiolo, L. Andricek, M. Ellenburg et al. *Thin n-in-p pixel sensors and the SLID-ICV vertical integration technology for the ATLAS upgrade at the HL-LHC*. 2012. arXiv: physics.ins-det/1210.7933 [physics.ins-det].

[37] C. Goessling, R. Klingenberg, D. Muenstermann et al. 'Evaluation of the breakdown behaviour of ATLAS silicon pixel sensors after partial guard-ring removal'. In: *Nucl. Instrum. Meth.* A624 (2010), pp. 410–413. DOI: 10.1016/j.nima.2010.06.004.

[38] ATLAS IBL Collaboration. 'Prototype ATLAS IBL Modules using the FE-I4A Front-End Readout Chip'. In: *JINST* 7 (2012), P11010. DOI: 10.1088/1748-0221/7/11/P11010.

[39] L. Andricek, M. Beimforde, A. Macchiolo et al. 'Development of thin sensors and a novel interconnection technology for the upgrade of the ATLAS pixel system'. In: *Nucl. Instrum. Meth.* A636 (2011), S68–S72. DOI: 10.1016/j.nima.2010.04.087.

[40] C. Da Via, E. Bolle, K. Einsweiler et al. '3D active edge silicon sensors with different electrode configurations: Radiation hardness and noise performance'. In: *Nucl. Instrum. Meth.* A604 (2009), pp. 505–511. DOI: 10.1016/j.nima.2009.03.049.

[41] M. Povoli, A. Bagolini, M. Boscardin et al. 'Slim edges in double-sided silicon 3D detectors'. In: *JINST* 7 (2012), p. C01015. DOI: 10.1088/1748-0221/7/01/C01015.

[42] M. Garcia-Sciveres, D. Arutinov, M. Barbero et al. 'The FE-I4 pixel readout integrated circuit'. In: *Nucl. Instrum. Meth.* A636 (2011), S155–S159. DOI: 10.1016/j.nima.2010.04.101.

[43] A. Macchiolo, L. Andricek, H.G. Moser et al. *SLID-ICV Vertical Integration Technology for the ATLAS Pixel Upgrades*. 2012. arXiv: physics.ins-det/1202.6497 [physics.ins-det].

[44] M. Weber. 'Power distribution for SLHC trackers: Challenges and solutions'. In: *Nucl. Instrum. Meth.* A592 (2008), pp. 44–55. DOI: 10.1016/j.nima.2008.03.107.

[45] L. Gonella, D. Arutinov, M. Barbero et al. 'A serial powering scheme for the ATLAS pixel detector at sLHC'. In: *JINST* 5 (2010), p. C12002. DOI: 10.1088/1748-0221/5/12/C12002.

[46] M. Karagounis, D. Arutinov, M. Barbero et al. 'An Integrated Shunt-LDO Regulator for Serial Powered Systems'. In: *Proc. of the European Solid-State Device Conference, ESSCIRC 2009* (2009), pp. 276–279. DOI: 10.1109/ESSCIRC.2009.5325974.

[47] L. Gonella, M. Barbero, F. Huegging et al. 'The shunt-LDO regulator to power the upgraded ATLAS pixel detector'. In: *JINST* 7 (2012), p. C01034. DOI: 10.1088/1748-0221/7/01/C01034.

[48] W.O. Miller and Inc. Allcomp. *Thermally Conductive, Carbon Foam Material for Constructing Silicon-Based Tracking Detectors*. US Dept. of Energy SBIR award 89695. 2008, 2009.

[49] B. Verlaat, M. van Beuzekom and A. van Lysebetten. 'CO2 Cooling for HEP experiments'. In: *Proc. of the Topological Workshop on Electronics for Particle Physics, TWEPP 2008* (2008), pp. 328–336.

[50] C. Gemme. *Private communication*. 2011.

[51] D.B. Ta, T. Stockmanns, F. Huegging et al. 'Serial powering: Proof of principle demonstration of a scheme for the operation of a large pixel detector at the LHC'. In: *Nucl. Instrum. Meth.* A557 (2006), pp. 445–459. DOI: 10.1016/j.nima.2005.11.115.

[52] T. Stockmanns, P. Fischer, F. Huegging et al. 'Serial powering of pixel modules'. In: *Nucl. Instrum. Meth.* A511 (2003), pp. 174–179. DOI: 10.1016/S0168-9002(03)01787-X.

[53] T. Stockmanns. 'Multi-chip module development for the ATLAS pixel detector'. BONN-IR-2004-06. PhD thesis. 2004.

[54] T. Stockmanns, P. Fischer, F. Huegging et al. 'Realization of serial powering of ATLAS pixel modules'. In: *Nucl. Instrum. Meth.* 52 (2005), pp. 2721–2725. DOI: 10.1109/TNS.2005.860163.

[55] D.B. Ta, T. Stockmanns, F. Huegging et al. 'Concept, realization and characterization of serially powered pixel modules (Serial Powering)'. In: *Nucl. Instrum. Meth.* A565 (2006), pp. 113–118. DOI: 10.1016/j.nima.2006.04.074.

[56] [Online]. Available at: http://edms.cern.ch/cern.ch/document/1064078.

[57] M. Backhaus. 'Characterization of the FE-I4B pixel readout chip production run for the ATLAS insertable B-layer upgrade'. In: *JINST* 8 (2013), p. C03013. DOI: 10.1088/1748-0221/8/03/C03013.

[58] M. Newcomer. *Private communication*. 2012.

[59] J. Matheson. 'Progress and advances in serial powering of silicon modules for the ATLAS tracker upgrade'. In: *JINST* 6 (2011), p. C01019. DOI: 10.1088/1748-0221/6/01/C01019.

[60] P.W. Phillips. 'ATLAS strip tracker stavelets'. In: *JINST* 7 (2012), p. C02028. DOI: 10.1088/1748-0221/7/02/C02028.

[61] National Semiconductor. *LVDS Owner's Manual*. Available at: http://www.ti.com/lit/ml/snla187/snla187.pdf.

[62] A. Widmer and P. Franaszek. 'A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code'. In: *IBM, J. Res. Dev.* 27 (1983), pp. 440–451.

[63] [Online]. Available at: http://www.onsemi.com/pub_link/Collateral/NCP100-D.PDF.

[64] [Online]. Available at: http://www.littelfuse.com/.

[65] M. Backhaus, M. Barbero, L. Gonella et al. 'Development of a versatile and modular test system for ATLAS hybrid pixel detectors'. In: *Nucl. Instrum. Meth.* A650 (2011), pp. 37–40. DOI: 10.1016/j.nima.2010.12.087.

[66] L.F. Miller. 'Controlled Collapse Reflow Chip Joining'. In: *IBM, J. Res. Dev.* 13 (1969), pp. 239–250.

[67] J.H. Lau. *Flip Chip Technologies*. McGraw-Hill New York, 1996.

[68] L. Rossi, P. Fischer, T. Rohe et al. *Pixel detectors: From fundamentals to applications*. Springer, 2006.

[69] T. Fritzsch, R. Jordan, M. Toepper et al. 'Experience in fabrication of multichip-modules for the ATLAS pixel detector'. In: *Nucl. Instrum. Meth.* A565 (2006), pp. 309–313. DOI: 10.1016/j.nima.2006.05.010.

[70] C. Gemme, A.M. Fiorello, G. Gagliardi et al. 'Study of indium bumps for the ATLAS pixel detector'. In: *Nucl. Instrum. Meth.* A465 (2000), pp. 200–203. DOI: 10.1016/S0168-9002(01)00390-4.

[71] G. Alimonti, A. Andreazza, A. Bulgheroni et al. 'Analysis of the production of ATLAS indium bonded pixel modules'. In: *Nucl. Instrum. Meth.* A565 (2006), pp. 296–302. DOI: 10.1016/j.nima.2006.05.048.

[72] Ch. Broennimann, F. Glaus, J. Gobrecht et al. 'Development of an indium bump bond process for silicon pixel detectors at PSI'. In: *Nucl. Instrum. Meth.* A565 (2006), pp. 303–308. DOI: 10.1016/j.nima.2006.05.011.

[73] L. Gonella, F. Huegging and N. Wermes. 'Towards minimum material trackers for high energy physics experiments at upgraded luminosities'. In: *Nucl. Instrum. Meth.* A650 (2011), pp. 202–207. DOI: 10.1016/j.nima.2010.11.162.

[74] G. Alimonti, A. Andreazza, G. Corda et al. 'Development of indium bump bonding for the ATLAS Insertable B-Layer (IBL)'. In: *JINST* 8 (2013), P01024. DOI: 10.1088/1748-0221/8/01/P01024.

[75] K. Zoschke, T. Fischer, M. Toepper et al. 'Polyimide based temporary wafer bonding technology for high temperature compliant TSV backside processing and thin device handling'. In: *Proc. of the Electronic Components and Technology Conference, ECTC 2012 IEEE 62nd* (2012), pp. 1054–1061. DOI: 10.1109/ECTC.2012.6248966.

[76] P. Garrou, C. Bower and P. Ramm. *Handbook of 3D Integration*. WILEY-VCH Verlag GmbH & Co, KGaA, 2008.

[77] M. Barbero, T. Fritzsch, L. Gonella et al. 'A via last TSV process applied to ATLAS pixel detector modules: Proof of principle demonstration'. In: *JINST* 7 (2012), P08008. DOI: 10.1088/1748-0221/7/08/P08008.

[78] M. Garcia-Sciveres. *Private communication*. 2012.

[79] F. Laermer and A. Schilp. 'Method of anisotrophic etching of silicon'. Patent US6531068 (US). 2003. URL: http://www.patentlens.net/patentlens/patents.html?patnums=US_6531068&language=&.

[80] K. Zoschke, J. Wolf, C. Lopper et al. 'TSV based silicon interposer technology for wafer level fabrication of 3D SiP modules'. In: *Proc. of the Electronic Components and Technology Conference, ECTC 2011 IEEE 61st* (2011), pp. 836–843. DOI: 10.1109/ECTC.2011.5898608.

[81] K. Zoschke, M. Wegner, M. Wilke et al. 'Evaluation of thin wafer processing using a temporary wafer handling system as key technology for 3D system integration'. In: *Proc. of the Electronic Components and Technology Conference, ECTC 2010 IEEE 60th* (2010), pp. 1385–1392. DOI: 10.1109/ECTC.2010.5490637.

List of Figures

1.1 The Standard Model of particle physics. 1
1.2 Signals left by collision products in the ATLAS detector. 2

2.1 Sketch of the LHC ring and the four experiments. 5
2.2 Expected hard scattering cross-sections at the LHC and Tevatron as a function of the center of mass energy. 6
2.3 Delivered LHC luminosity versus time for 2010, 2011, and 2012 (*pp* data only). 8
2.4 Overview of the ATLAS detector. 10
2.5 ATLAS Higgs discovery plot. 11
2.6 Feynman diagrams for bremsstrahlung and photon conversion. 14
2.7 Angular distribution of the multiple scattering angle. 15
2.8 Transverse momentum resolution. 17
2.9 Sketch of a one-dimensional vertex detector arrangement. 18
2.10 Schematic picture of the determination of the impact parameter b of a track. . . 19
2.11 Number of interactions per crossing for the 2012 LHC run, and a candidate Z boson event with high pile-up. 21
2.12 Layout of the ATLAS inner detector. 22
2.13 Material distribution expressed in units of X_0 as a function of η for the ATLAS ID. 23
2.14 The ATLAS pixel detector. 24
2.15 Baseline layout of the ATLAS tracker for the HL-LHC. 27
2.16 Material distribution expressed in units of X_0 as a function of η for the ATLAS tracker at the HL-LHC. 28

3.1 Hybrid pixel detectors. 29
3.2 Material distribution of the ATLAS pixel detector expressed in X_0, as a function of η. 30
3.3 Sketch of an ATLAS pixel detector module. 31
3.4 Schematic layout of the ATLAS pixel services scheme. 33
3.5 Drawing of the ATLAS pixel detector barrel staves 36
3.6 Pictures of the b-layer and of a disk of the ATLAS pixel detector. 36
3.7 Comparison of the edge region of the APS design and the IBL planar sensor design. 38
3.8 Sketch of a double-sided 3D sensor and of a Full3D sensor with active edges. . 39
3.9 Picture of the FE-I4 and FE-I3 pixel readout chips. 40
3.10 IBL modules. 41
3.11 Sketch of a 4-chip module for the ATLAS pixel outer layers at the HL-LHC. . . 42

3.12 Schematic drawings of a parallel powering, DC-DC conversion and serial powering. . . . 43
3.13 FE-I4 powering options. . . . 44
3.14 Drawing of the IBL layer. . . . 46
3.15 Drawings of a box beam stave. . . . 47

4.1 Sketch of a SP chain. . . . 50
4.2 Cable power losses and material budget as a function of the voltage drop for a SP scheme. . . . 54
4.3 I-V characteristics of a shunt regulator. . . . 56
4.4 Schematic drawing of on-chip shunt and LDO regulators to generate VDDD and VDDA. . . . 57
4.5 Schematic drawing of on-chip current to voltage conversion using the Shunt-LDO principle. . . . 57
4.6 Simplified schematics of a Shunt-LDO regulator. . . . 58
4.7 Sketch of a SP chain with bypass scheme. . . . 60
4.8 Sketch of an AC-coupled LVDS link. . . . 61
4.9 Schematics of the *HV* distribution scheme for a serial powering chain. . . . 63
4.10 Connection of the Shunt-LDO regulators to power the FE-I4 chip. . . . 64
4.11 Difference between regulator and board grounds for the Shunt-LDO working mode. . . . 65
4.12 Connection of the Shunt-LDO regulator in Shunt-LDO mode. . . . 66
4.13 I-V characteristics of two regulators in Shunt-LDO mode, connected in parallel, generating V_{out} = 1.2 V and 1.5 V. . . . 66
4.14 Current distribution at start-up between two regulators in Shunt-LDO mode, connected in parallel. . . . 67
4.15 Load regulation measurement of two regulators in Shunt-LDO mode, connected in parallel. . . . 68
4.16 Load transient measurement of two regulators in Shunt-LDO mode, connected in parallel. . . . 69
4.17 Shunt-LDO power efficiency as a function of the shunt current. . . . 70
4.18 Connection of the Shunt-LDO regulator in LDO mode. . . . 70
4.19 Power efficiency of the regulator in LDO mode. . . . 71
4.20 Connection of the Shunt-LDO regulator in partial shunt mode. . . . 72
4.21 Transients in partial shunt mode and LDO mode. . . . 73
4.22 Picture of a MDC FE-I4A module with dedicated SP flex hybrid. . . . 74
4.23 Generation of the regulator reference voltage using a voltage supply or a commercial adjustable shunt. . . . 74
4.24 AC-coupled clock and data links. . . . 75
4.25 Threshold and noise map of the MSC2 module in direct power configuration. . . 77
4.26 Threshold and noise map of the MSC2 module in serial power configuration. . . 77
4.27 Threshold and noise map of the MSC2 module in LDO power configuration. . . 78
4.28 Sensor I-V characteristics for MSC modules, operated with different powering schemes. . . . 78

4.29 Sensor I-V characteristics for serially powered MDC modules. 80
4.30 Picture of the SP prototype stave. 81
4.31 Threshold and noise maps of module MDC1 tuned to 3000 e^-. 83
4.32 Threshold and noise maps of module MDC2 tuned to 3000 e^-. 84
4.33 Threshold and noise maps of module MDC8* tuned to 3000 e^-. 84
4.34 Threshold and noise maps of module MDC4 tuned to 3000 e^-. 85
4.35 Source hit map and spectrum recorded with a module of the SP stave. 87
4.36 Picture of the SPP chip. 88
4.37 Results of the SPP chip characterization. 89

5.1 AgSn bump bonds on a FE-I3 wafer after reflow. 92
5.2 Cross-section of an ATLAS pixel module showing the AgSn bump bonds. . . . 93
5.3 Indium bump bonds. 93
5.4 Bending profile of the FE-I4 chip as a function of temperature and thickness. . 95
5.5 Illustration of the modified IZM flip chip process using a temporary wafer bonding technology. 96
5.6 Cross-section of a prototype module built using the Brewer WaferBOND HT. . 97
5.7 Cross-section of a prototype module built using the polyimide method. 97
5.8 Visual inspections on IBL module prototypes. 98
5.9 Source scan hit map of an IBL prototype module before and after thermal cycling. 99
5.10 Histogram of unconnected bumps over 30 IBL MDC planar modules. 100

6.1 Comparison of a possible TSV module concept and an IBL module. 104
6.2 Straight and tapered TSV profiles. 106
6.3 Front side processing on the ATLAS FE-I2 wafers. 106
6.4 SEM picture of a tapered profile TSV and plug on the FE-I2 wafers. 107
6.5 RDL and wire bond pads on the FE-I2 back side. 108
6.6 Noise map and noise distribution of a TSV module. 109
6.7 Threshold map and threshold distribution of a TSV module. 109
6.8 Threshold map and threshold distribution of a TSV module. 110
6.9 Spectrum of an ^{241}Am source recorded with a demonstrator module with TSV. . 110

List of Tables

2.1 LHC luminosity upgrades. 8
2.2 Predicted fluence and TID for the ATLAS tracker at the LHC. 20
2.3 Predicted fluence and TID for the ATLAS tracker at the HL-LHC. 26

3.1 Comparison of direct power, DC-DC conversion and serial power. 44
3.2 Material budget per layer of the current ATLAS pixel detector, the IBL and the outer layers at the HL-LHC. 48

4.1 Values used to compare serial and direct powering schemes for an outer layer of the ATLAS pixel detector at the HL-LHC. 53
4.2 Comparison between a direct and a serial powering scheme for the upgrade of the ATLAS pixel detector at the HL-LHC. 55
4.3 Working modes of the Shunt-LDO regulator. 59
4.4 Measured regulator parameters in Shunt-LDO mode. 68
4.5 $V_{dropout}$ and line regulation as a function of V_{out} and I_L for the LDO working mode. 71
4.6 Threshold, dispersion and noise of MSC modules for different powering configurations. 76
4.7 Threshold, dispersion and noise of serially powered MDC modules. 79
4.8 Voltages of the modules in the SP chain at power-up. 82
4.9 Results of tests on the SP stave prototype. 86
4.10 Difference in noise value with respect to standard operation, when one module in the SP chain is oscillating. 88

Acknowledgements

I would like to thank first of all Prof. Dr. Norbert Wermes and Dr. Markus Cristinziani for giving me the possibility to do my Ph.D. within this group. In particular, I would like to express all my gratitude to Prof. Wermes for his trust in my work, for the knowledge, expertise, and wide range of topics within the group, which made my Ph.D. an intense learning experience.

Many thanks to Dr. Fabian Hügging for always answering all of sort of questions I had, at any time, for the infinite discussions on serial powering, and for proof reading this thesis. I am grateful to Dr. Marlon Barbero for his big help, in particular during the shunt-LDO testing, as well as on the non work related issues. Thanks to Dr. Hans Krüger for always providing clues on clueless problems.

Thanking Malte Backhaus for everything he should be thanked for would make this thesis twice as long, and would probably still not be enough. In a nutshell, dear Malte, thanks for helping me with my work, for being a patient office-mate and friend, standing my good and bad moods, especially in the last months, and for all the franzbrötchen.

Thanks to everybody in the SiLab for the lively discussions, coffee breaks, birthday cakes.

A very special, big, thank goes to Andrés, indeed the best tandem partner of "pasarla bien", and to my BFF Giovanna and Emilio, for our valuable skype chats and hanging out "downtown". I would like to thank Francesco for borrowing the oscilloscope probe so often. Hadn't he done so, I would never have met Vale, Simone, Mauricio, Qualli, Azzurra, Alice, Luis, Juan, Calitos, il presidente, Niko, and many more with whom I had a really great time in Bonn, Hamburg, The Hague, Winterberg, Kos. Many thanks to Susanna for being a lot more than just a housemate, and for all the "gute Zeite, schlechte Zeite", chips & gin tonic. Un grazie particolare va anche alla mie "amise", Tabi, Paola, Mary, Vivi e Arra perchè, nonostante la distanza geografica, sono sempre vicine. Ich möchte mich noch sehr herzlich bei Frau Dr. Stumpf-Curio für Ihre Professionalität und Geduld bedanken.

Finally, I would like to thank Andreas Eyring, and not only for everything he taught me about electronics, and the great help and support to my work, but mostly for being my father in Germany. Thanks for always understanding when I was feeling down, thanks for always knowing when I needed a hug, thanks for teaching me what is really important in life.

Il ringraziamento piú grande va a mia mamma, mio fratello, mio papá, e mio zio Gian per avermi supportata e sopportata in tutti questi anni, e anche nei precedenti. Vi voglio bene!