UNIVERSITA' DEGLI STUDI DI TORINO
DIPARTIMENTO DI FISICA SPERIMENTALE
DOTTORATO DI RICERCA IN FISICA
CICLO XIX

Settore scientifico-disciplinare FIS/01
FISICA SPERIMENTALE

DESIGN, TEST AND SYSTEM INTEGRATION OF FRONT-END ELECTRONICS FOR PARTICLE DETECTION IN HIGH-ENERGY NUCLEAR AND SUBNUCLEAR EXPERIMENTS

Tesi Presentata da:
Victor Sorin Martoiu

Relatori: **Prof. Alberta Marzari-Chiesa**
Dott. Angelo Rivetti
Controrelatore: **Prof. Alberto Aloisio**
Coordinatore: **Prof. Stefano Sciuto**

Anni accademici: 2003/2004–2004/2005–2005/2006

In memory of my father

Table of Contents

Acknowledgements

I would like to thank Alberta Marzari-Chiesa and Angelo Rivetti, my supervisors, for their many suggestions and constant support during this research. I am also thankful to Luciano Ramello for his guidance and support.

I want to thank Giovanni Anneli, Elena Bruna, Sorin Cheran, Daniela Calvo, Ozgur Cobanoglu, Elisabeta Crescio, Paolo De Remigis, Ernesto Filoni, Paolo Giubelino, Flavio Marchetto, Mario-Ivan Martinez, Gianni Mazza, Andrea Mignone, Francesco Rotondo, Sabino Salerno, George Serbanut, Luca Toscano, Flavio Tosello, Richard Wheadon for their collaboration and fruitful discussions. Special thanks to Ozgur Cobanoglu for helping with the layout of the last submission.

I am grateful to my parents and to Ramona for their patience and *love*. Without them this work would never have come into existence (literally).

Finally, I wish to thank the following:

(the *other*) Sorin (for holding the rope tight while climbing); Andrea, Davide, Guido and Stefano (for the painful introduction to ski-alpinism); Alejandra, Alessandro, Alex, Catalin, Dede, Laura and her friends, Luca, Mario-Ivan, Nasim, Ozgur, Sarolta, Sorin, Stefano,... (for all the good and bad times we had together); and Adi, Andrei, Cuky, Dan, Florica, Maca, Mihai, Sile, Sorel, Victor... (for still remembering me after so many years spent away).

Torino, Italy — Sorin Martoiu
February, 2007

Introduction

The ambitious physics goals of the LHC experiments set new frontier in particle detector domain, pushing performance and complexity to unprecedented challenging levels. Silicon-based detectors are an example of particle detection systems that benefited from the extensive research and development work related to the LHC based high-energy physics experiments. Several years ago silicon drift detectors were confined to small scale experiments with relatively low sensitive area like, for instance, medical or spectroscopy applications, while hybrid pixels were an expensive silicon detection system both in terms of design and practical realization. Nowadays, with the extensive use of silicon drift detectors in large scale heavy-ion collision experiments like RHIC and ALICE, and the use of large area high granularity hybrid pixel systems in all four LHC-based experiments, both technologies offer a mature alternative for the implementation of medium or large scale tracking detectors. Non-LHC experiments that are planned to start operation in near future, proposing high interaction rates, and the already foreseen LHC upgrade push on this R&D work. Emerging technologies like CMOS sensors and 3D silicon detectors also gain increasingly interest in the HEP community.

The need for better resolution and higher granularity in tracking detectors calls for higher number of readout channels and more compact front-end electronics. The

material budget, limited by the precision in momentum measurement, imposes for both low power design and the use of new technologies in interconnections, mechanical support, cooling, etc. In these conditions, preserving laboratory performance of individual sub-detectors, while replicating them into large-area detector subsystem, is a demanding task. Eliminating large-system specific features like crosstalk, digital spurious noise, ground level bouncing, to nominate only a few, asks for careful design and extensive production qualification effort. Consequently, system integration in high-energy physics became a research and development domain by itself, owing from domains like electronic and mechanical engineering to materials science.

This work follows the design, realization and test of front-end electronics for particle detection using silicon drift and silicon pixel detectors applied in nuclear and high-energy physics experiments. Reflecting a research experience that started in 2004 within the ALICE SDD group of INFN Torino and the University of Torino, a period when all LHC-based detectors were finalizing the design phase and starting the final integration, this work begins with the system integration aspect of the large scale SDD detection system and continues with the design, prototyping and test of new front-end electronics for silicon detector systems for the PANDA and P326 experiments at GSI and CERN, respectively. The apparent "mis-arrangement" of the design-test-integration flow, apart from being an exact photography of the research experience reported here, also underlines the repetitive and self-consistent nature of the sequence, as all knowledge gathered from completed design-to-integration flow is the starting base for a new one. Specific problematic and technological issues with accent on quality assurance scheme of the integration of the Silicon Drift detector system of the ITS of ALICE experiment is presented together with characterization

of the custom designed VLSI front-end electronics. The work also presents design and prototype characterization of new VLSI front-end electronics for pixel detectors that were realized using the currently commercially available deep-submicron CMOS technology. The migration from silicon drift to pixel detectors is a natural trend driven by newly available VLSI fabrication technologies, in the view of future luminosity upgrade of LHC and other high intensity HEP experiments.

Chapter 1

Experimental challenges in nuclear and particle physics

This chapter gives an overview of the main aspects the modern nuclear and particle physics focuss on, from both theoretical and experimental point of view. The chapter serves as extended introduction to the three experiments that this work contributes to (ALICE, P326, PANDA) setting their scientific environment and, finally, giving specific motivation and description for each of them.

1.1 Short history of experimental nuclear and particle physics

At the end of 19th century the concept of atom formulated by Democritus of Abdera as *the* fundamental indivisible constituent of matter was still in the general belief of the scientists.

In 1898 J. J. Thompson proposed that the long-standing puzzle known as "cathode-rays" were in fact negative charged *corpuscules*, later called *electrons*[1]. This hypothesis was later confirmed by experiments carried out by Thompson and others. Moreover the mass of this *corpuscules* was found to be about one thousand times smaller than the mass of the lightest atom. This extraordinary discovery was the first to contradict the indivisible nature of the atom. It lead to the atomic model known as *the plum pudding model* proposed in 1904 by Thompson, picturing the atom as composed by electrons surrounded by a soup of positive charge to balance the electron's negative charge, like plums surrounded by pudding.

Thirteen years later, in 1911, following the alpha-scattering experiment performed by Hans Geiger and Ernest Marsden, one of the Thompson's students - Ernest Rutherford concludes that the atom has a concentrated positive nucleus. Several years later Rutherford finds the first evidence for a proton (1919). He manages to bombard the nucleus with alpha particles, using nitrogen as target. The nitrogen nucleus absorbs the alpha particle resulting an oxigen nucleus and a free proton. The neutron was first postulated by Rutherford and later in 1932 discovered by James Chadwick. He demonstrated that the "radiation of great penetration power" resulted from the bombardment of beryllium with α-particles was, in fact, "particles of mass 1 and charge 0, or neutrons"[1].

In 1927 Rutherford urged the experimental physics community to develop sources of particles of higher energies than those provided by natural radioactivity, in order to explore deeper into the structure of the atom. John Cockcroft and E. T. S. Walton from Rutherford's Cavendish Laboratory in Cambridge developed a proton

[1]The term *electron* was initially used by G. Johnstone Stoney in 1874 to explain the phenomenon of electrolysis, but they were not regarded as constituents of the atom at the time.

accelerator scheme using a high-voltage transformer and a voltage multiplier. In parallel, Robert Van de Graaff built a particle accelerator based on an electrostatic high-voltage generator. His idea survived through the century being still today the most commonly used potential-drop accelerator.

At the beginning of 1930s Ernest Lawrence, together with its graduate student Stanley Livingston, applied a new concept of particle accelerator, inventing the cyclotron. He used a magnetic field to curve the trajectory of a charged particle, forcing it to cross many times a potential gap in a repetitive fashion, while the potential is switched between positive and negative as the particle crosses the gap in one direction or the other. In this way the particle is accelerated at each gap-crossing until the curvature radius of its trajectory reaches the boundary of the device and the particle escapes. After more than 10 years of hard work and successive prototyping they built a 184-inch cyclotron capable of accelerating various charged particles to energies higher than 100 MeV.

In the mid 1940s, two new types of accelerators were developed: the linear accelerator and the synchrotron, using the newly invented phase stability concept. This inventions paved the way for a new phase of the experimental physics. The number of particle accelerators grew substantially over the years, as well as their complexity and energy range. Rutherford's scattering experiment was reproduced at nucleon scale in 1968 at the Stanford Linear Accelerator Center. This time, the large angle scattering pattern of the high-energy electron-proton collision revealed the quark structure of the proton.

Among other advanced accelerator facilities, the two largest proton synchrotrons, at CERN, near Geneva, and at Fermilab near Chicago, started operation in the mid

1970s. Fixed-target experiments were soon replaced by collision experiments where two particles, both accelerated close to the speed of light, are forced to collide one into another, increasing the center of mass energy, and thus the physics yield. Table 1.1 gives a short summary of some of the collider facilities around the world and their contribution.

Machine	Location	Beam	Energy (GeV)	Radius	Highlight
AdA[1]	Frascati	e^+e^-	$0.25 + 0.25$		
SPEAR	Stanford	e^+e^-	$3 + 3$		c, τ
DORIS	Hamburg	”	$5 + 5$		b
CESR	Cornell	”	$8 + 8$	125 m	”
PEP	Stanford	”	$18 + 18$		b lifetime
PETRA	Hamburg	”	$22 + 22$	300 m	g
SPS	CERN	$\bar{p}p$	$300 + 300$	1 Km	W, Z
TEVATRON	Fermilab	”	$1000 + 1000$		t
SLC	Stanford	e^+e^-	$50 + 50$	–	Z
LEP-I	CERN	”	”	5 Km	Z
LEP-II		”	$100 + 100$	”	W
HERA	Hamburg	ep	$30 + 800$	1 Km	–
LHC	CERN	pp	$7,000 + 7,000$	5 Km	Higgs? SUSY?

[1] In italian, ”Anello di Accumulazione” - the first storage ring build by Bruno Touschek in 1960[2].

Table 1.1: Past and present colliders.

1.2 Overview of the present objectives in nuclear and high-energy physics

1.2.1 Overview of the Standard Model

There are four fundamental interactions in nature.

The Electromagnetic Interaction is the unified representation of the forces at the origin of electric and magnetic phenomena. It causes the attraction felt by particles of opposite electric charges and the repulsion between those that carry the same sign. This interaction therefore holds the atomic structure together in which negatively charged electrons orbit around a positively charged nucleus. The interaction is conveyed by the photon, which acts like a messenger particle between electrically charged particles.

The weak interaction is the force at the origin of physics phenomena like radioactivity which spontaneously transforms a neutron into a proton or the thermonuclear production of energy at the heart of the sun. The conveyors of this interaction are the W and Z bosons discovered at CERN.

The Strong Interaction is the force that binds together the quarks within hadrons, like in protons and neutrons. The interaction is conveyed by the eight gluons. It also keeps protons and neutrons together within the atomic nucleus against the action of the electromagnetic force. The particular property of the Strong force is that its effect increases with distance, making it impossible for the quarks to be observed individually in nature.

The Gravitational Interaction is responsible for the attraction between masses. Its effect in our macroscopic every-day life is widely known, as the fact that any object

on earth is attracted toward its center. It is at astronomical level that it reveals its most subtle effects in phenomena such as the movement of celestial objects or the curvature of the light paths in the vicinity of massive objects. Messenger particles of the gravitational field called "gravitons" is supposed to exist. However there is no experimental evidence of its existence as yet. Gravitational force is so weak that its effects are totally negligible in the subatomic world.

Electroweak theory proposed around 1960 by Sheldon Glashow, Abdus Salam, and Steven Weinberg [3], unifies the electromagnetic and weak interactions. Quantum chromodynamics, on the other side, is the best theoretical quantum-field description of the strong force. The Standard Model is a result of the combination of the electroweak theory (GSW theory) with quantum chromodynamics (QCD), covering all the interactions governing the nature, except the gravitational force.

Within the Standard Model, the elementary constituents of nature are fermions, particles with intrinsic spin of 1/2. There are 12 different matter particles, six quarks (*up, down, strange, charm, bottom, top*) and six leptons (*electron, muon, tau* and their corresponding *neutrinos*).

The six flavors of quark come in three generations of weak isospin doublets that lead to SU(2) symmetry. Each quark has three possible colors. The fact that quarks come in three colors is used as the basis for SU(3) symmetry. The SU(3) color symmetry is exact since the three colored quarks of each flavor are identical in every respect except color. The SU(3) symmetry of colored quarks would automatically require eight ($3^2 - 1 = 8$) gauge bosons. These are the eight gluons characterized by color-anticolor combinations, similar to the colored quark-antiquark pairs used to obtain meson multiplets. The gluons are responsible for mediating the strong

color force that binds hadrons. The Standard Model results when the electroweak theory group represented by $SU(2) \times U(1)$ is extended to include the SU(3) colored quark group which represents QCD. The Standard Model can account for all hadronic (strong interaction) and electroweak phenomena. Meson exchange effects that are still useful in describing low energy nuclear phenomena are residual effects of the strong color force in the same way that molecular forces are consequences of the electromagnetic force.

Most of the fundamental predictions of the Standard Model have been experimentally confirmed. The discovery of the W and Z bosons, as well as the excellent accordance between their measured and predicted properties are very strong confirmations of the electroweak theory. On the other hand QCD has enjoyed the same experimental success. All quarks predicted by the theory were observed in a series of experiments. Moreover, the experiments carried out using the LEP collider at CERN have demonstrated that only three weakly-interacting neutrinos exists, with mass lower than half of the mass of Z boson. This implies the existence of three generations of quarks and leptons, as predicted by theory.

1.2.2 The search for Higgs Boson.

One of the pieces of the puzzle still missing, is the Higgs boson. This particle is an observable component of the Higgs field, which is characterized by a non-zero vacuum expectance value (VEV), being responsible for the mass of every elementary particle, including the Higgs boson. In particular, the Higgs mechanism spontaneously breaks the electroweak gauge symmetry, which otherwise would require that Z and W boson were massless, just as their electromagnetic counterpart, the photon.

Although unconstrained by the theory, available experimental data provide upper and lower bounds on the Higgs mass. Direct searches at LEP2 have ruled it must exceed 114.4 GeV at 95% confidence level [4]. On the other hand, electroweak precision measurements sets an upper bound on the mass of the Higgs boson at 219 GeV with the same confidence level [5]. Given the logarithmic sensitivity of electroweak precision measurements to m_H, the upper bound extends at a factor of a few above this limit. Both ATLAS and CMS collaborations at CERN will search for the Higgs boson in the mass range of $10^2 - 10^3$ GeV.

1.2.3 Inside the hadronic matter

One of the fundamental difference between QCD and Quantum Electrodynamics is the fact that the gluons carry color charges, as opposed to photons which are electrically neutral. The color charge makes the gluons interact amongst themselves, giving birth to a new set of phenomena and key features of QCD such as confinement (the absence of free colour-charged objects in nature) and asymptotic freedom (the fact that quarks and gluons interact weakly at high momenta or short distances).

At distances below 0.1 Fermi, the QCD coupling strength is small, allowing for a perturbative treatment of the strong interaction. On the other hand, at distance scales of the order of 1 Fermi, the coupling become large and a perturbative approach is no longer valid. In this regime, the theoretical treatment involves colorless baryons and mesons rather than quarks and gluons.

Generally there are two methods for experimental exploration of both regimes, perturbative and nonperturbative

- studying the responses of hadronic systems to high-precision probes at various

energy scales,

- creating conditions of high density or temperature in high-energy heavy-ion collisions.

1.2.3.1 Spin of the nucleon

In recent years we have learned that quarks are not the only contributors to the total spin of the nucleon, but also the gluons and possibly the orbital angular momentum of the quarks and gluons do contribute.

The contribution of the quarks to the total spin of the nucleon has been worked out from the longitudinal spin distribution (or helicity) in polarized deep-inelastic scattering experiments, carried out at CERN, DESY and SLAC. The results read that the total quark spin contribution is in the order of 10% to 30%, indicating that additional carriers of angular momentum are needed in the nucleon.

Some recent experiments or proposals try to measure the gluon contribution from photon-gluon fusion events in deep-inelastic scattering experiments. In these processes the virtual photon annihilates with a gluon from the target to produce a quark and its antiquark, while the asymmetry of the process is sensitive to the gluon polarization.

The identification of photon-gluon fusion events through detection of charmed particles or high-p_T pairs of hadrons at COMPASS (CERN) will provide precise measurement of the gluon polarization. Complementary data is expected to come from the RHIC-spin program at BNL using polarized proton beams.

1.2.3.2 Glueballs and hybrids

Another consequence of the gluon-gluon interaction is the existence of gluon-rich bound states known as glueballs, and of mixed states of quarks and gluonic excitations known as hybrids.

The experimental observation of these exotic particles, in particular glueballs, would confirm one of the most important features of QCD. On the other hand, the non-existence of such states would pose a genuine problem for our understanding of hadronic physics in the context of QCD.

One of the best candidate for the ground state glueball has been observed in proton-antiproton annihilation experiments performed at LEAR ($f_0(1500)$)[6] and later by other experiments. Its mass is very close to what Latice QCD predicts for the ground state glueball, while its decay modes and production mechanism appear peculiar in the quark structure hypothesis. Other candidates are the $f_0(1370)$ and $f_0(1710)$ resonances. They appear to be combinations of $q\bar{q}$ states and $non-q\bar{q}$ states, making the analysis of their glueball content a difficult task.

Most of the available experimental data in the glueballs and hybrids sector is limited to the mass range below 2.2 GeV/c^2.

Other gluon-rich processes, like proton-proton collisions or radiative J/ψ decay, that may lead to the production of such exotic states have disadvantages in the high energy range. In the first case production of higher mass states is limited by the fall-off of the cross section with the inverse square of the mass of the state. Radiative J/ψ decay, which could produce gluonic hadrons up to 3 GeV/c^2, lacks the required statistics. For a more complete understanding of the nature of gluonic excitations, a careful study of the spectrum of glueballs and hybrids up to $5GeV/c^2$ is an absolute

necessity[7]. Theory suggests that for masses above 4.3 GeV/c2 the mixing with normal mesons should be suppressed. As a consequence they are predicted to be rather narrow and easy to identify experimentally.

GSI plannes a new ambitious physics program where the search for gluonic excitations in the charmonium sector and the continuing hunt for glueballs, including highly excited states with exotic quantum numbers have a central part. The program uses the PANDA detector at the High-Energy Storage Ring (HESR), and will carry on fixed-target experiments with high quality antiproton beam, including proton-antiproton annihilation with energies up to 15GeV.

Lattice QCD calculations predics the existence of about 15 glueball states, in the mass range that is accessible to the HESR project. Searches for glueballs and hybrids in this energy region can be performed in parallel with studies of charmonium spectroscopy at the proposed PANDA detector. In addition, by comparing different production mechanisms it should be possible to find unambiguous signatures of these exotic states.

1.2.3.3 Charmed quark systems

The charmonium system has a particular importance, compared to the other quarkonia $s\bar{s}$ and $b\bar{b}$, because it lies in the region of intermediate distances where the domains of perturbative and nonperturbative QCD come together. It is for this reason that the charmonium system provides a unique testing ground for QCD. Indeed, the masses and widths of the $c\bar{c}$ states directly reflect the basic $q\bar{q}$ interaction. Moreover, the charmonium spectroscopy may lead to the determination of the gluon condensate of the QCD vacuum.

Charmonium spectroscopy was extensively studied at e^+e^- colliders during 1974 − 1980. However, the technique of studying charmonium via e^+e^- annihilation had important limitations. The masses of several states were well determined but not, in general, their widths.

Experiments R704 at CERN and E760/E835 at Fermilab demonstrated that charmonium formation using $p\bar{p}$ annihilation has two significant advantages compared to e^+e^- annihilation. The first is that, since $p\bar{p}$ annihilation must proceed via two or three intermediate gluons, it can lead to the direct formation of charmonium states with all possible quantum numbers. This means that the precision achievable for all states depends only on the quality of the antiproton beam and not on the detector properties. The second advantage comes from the possibility of cooling antiproton beams (stochastically and/or with electrons) to obtain a momentum resolution of one part in 10^5, which translates directly into improved mass resolution.

Despite these efforts, there remain a number of unresolved fundamental questions concerning the charmonium system. These will be addressed by experiments focused on charmonium spectroscopy, using the PANDA detector system at GSI/HESR. This facility will offer improvements beyond the Fermilab program, by providing higher-energy antiproton beams (15 GeV), higher luminosity, better cooling, and a state-of-the-art hermetic detector for both electromagnetic and charged particles.

When running at full luminosity, HESR will produce a large number of D-meson pairs. Thus, it can also be regarded as a hadronic factory for tagged open charm. The high yield and the well-defined production kinematics of these pairs would allow studies of rare processes in the charm system such as CP-violation or flavour mixing, and determinations of the decay constants of charmed mesons.

1.2.3.4 Hypernuclei

An ordinary nucleus is composed of neutrons and protons. The exchange of one of the *up* or *down* quarks in the nucleus with a *strange* quark leads to a new type of nucleus called hypernucleus. Due to the strangeness conservation in strong and electromagnetic interactions, such bound state lives long enough to have sharp nuclear energy levels. Since the presence of a *strange* quark adds a new quntum number to the nucleus, hypernuclear physics merges isospin and strangeness into the enlarged field of flavour SU(3) many-body dynamics.

Hypernuclei were discovered more than 50 years ago by Marian Danysz and Jerzy Pniewski in a nuclear emulsion exposed to cosmic radiation. Significant progress have been made in the last 15 years, mainly at BNL, KEK and COSY. Recently at DAΦNE in Italy, FINUDA experiment started a research program aiming at production of large number of hypernuclei using low-energy K^- particles. Recently, spectrum measurements of ${}^{12}_{\Lambda}C$, as well as evidence of hypernuclei with high N/Z ratio, ${}^{6}_{\Lambda}H$ and ${}^{7}_{\Lambda}H$, has been reported [8, 9].

More insight in hypernuclear physics will be available also at GSI. Collisions of antiprotons with nuclei is foreseen, which will lead to detaile spectroscopic study of hypernuclei with single and multiple strangeness. Charmed hypernuclei (nuclei containing a charmed baryon) will also be available, opening a new chapter in hypernuclear physics that may lead to new insights new insights into the dynamics related to breaking of SU(4) flavour symmetry by the large mass of the charmed quark.

1.2.3.5 Phases of nuclear matter

Another tantalizing aspect of hadronic world is how do the properties of hadrons and hadronic bound states change due to thermodynamic conditions. Moreover, there is the question whether compressed / heated systems of protons and neutrons exhibit different phases depending on the local nuclear temperature and density, similar to macroscopic matter that may exist in three phases: solid, liquid or gas, and what is the influence of this behaviour at macroscopic level.

In their normal states of lowest energy, nuclei show liquid-like characteristics and have a density of 0.17 $nucleons/fm^3$. At moderately high baryon densities a first-order liquid-gas transition may takes place. In the gas state, the nuclear matter consists of free nucleons and pions.

After the discovery of asymptotic freedom in QCD it was realized by Collins and Perry [10] and independently by Cabibbo and Parisi [11] that a hadronic system at sufficiently high density or temperature should convert into a new state of matter, called quark-gluon plasma (QGP). Lattice QCD calculation predicts that the transition to quark-gluon plasma occurs at a critical temperature $T_c = 175 \pm 15 MeV$.

Above this temperature, the hadrons lose their particle identity, as the colour charges are screened at distances larger than the Debye screening length[12], creating a deconfined state where quarks and gluons can move quasi-freely inside the initial volume.

The high temperature and density conditions for the transition to QCP can be created in laboratory through high-energetic heavy-ion collisions. The scenario of QGP formation and its evolution in highly central nucleus-nucleus collision is as follows:

- following the collision, the nuclei disintegrate into their parton structure, building up a region of high entropy called *fireball*;

- the energy released in the collision begin to thermalize, increasing the pressure and the energy density inside the *fireball* leading to deconfinenment of quarks and gluons and plasma formation;

- as the plasma is cooling down, it expands and then hadronizes when the temperature crosses back T_c

Several experiments have been dedicated to this research field (NA35, NA44, NA45 (CERES), NA50 at the SPS at CERN, STAR at RICH, etc.). A first evidence for the quark-gluon plasma might be the anomalous J/ψ suppression observed in Pb-Pb collisions by NA50 at CERN. The chromoelectric Debye screening effect, characteristic to the deconfined state, cuts off the binding force between the constituent charm and anticharm quarks of J/ψ, and thus signals the occurrence of such state.

The study of QGP formation and its properties may unveil valuable information about hadronic matter and the strong interaction. The scientific importance of this particular state extends to astrophysics and cosmology, as it is generally believed that the universe was in the first instants after the *Big Bang* in a deconfined state with nearly perfect thermal and chemical equilibrium. Good understanding of the laboratory generated QGP and the hadronization phase may lead to better understanding of the universe was formed.

1.3 P326 and CKM mechanism

In 1973 Kobayashi and Maskawa proposed the existence of a third generation of quarks[13] in order to explain the CP-violation first observed in kaon decays in the mid 1960s at the alternating gradient synchrotron at BNL[14]. Their suggestion came before the first experimental evidence of charm, beauty or top quarks. Their theory introduces a 3x3 *complex* and *unitary* matrix called Cabibbo-Kobayashi-Maskawa (CKM) matrix governing the couplings between different quark generations [12] (see figure 1.1):

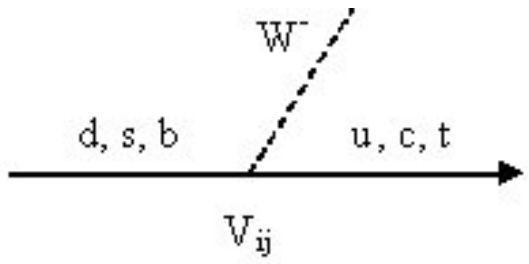


Figure 1.1: Flavour changing diagram example.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.3.1}$$

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.3.2}$$

Using the traditional Wolfenstein parameterization[15] the CKM matrix can be written as follows:

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \partial V_{CKM} \quad (1.3.3)$$

$$\partial V_{CKM} = \begin{pmatrix} 0 & 0 & 0 \\ -iA^2\lambda^5 & 0 & 0 \\ A\lambda^5\frac{\rho+i\eta}{2} & A\lambda^4(1/2-\rho-i\eta) & 0 \end{pmatrix} \quad (1.3.4)$$

Within this theoretical frame, a non-zero value of the parameter η in equation (1.3.3) is responsible for all CP-violating phenomena in flavor changing processes.

The unitarity of the CKM matrix implies

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk} \quad (1.3.5)$$

$$\sum_j V_{ij}V_{kj}^* = \delta_{ik} \quad (1.3.6)$$

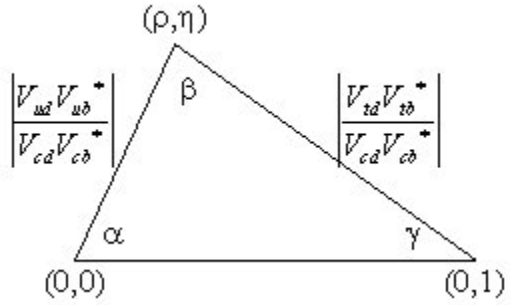


Figure 1.2: Unitarity triangle.

One of the most commonly used unitarity relation and of particular importance for what follows is:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (1.3.7)$$

which can be graphically represented as a triangle in the complex plane (figure 1.2). The apex of the triangle is $(\overline{\rho}, \overline{\eta})$, where

$$\overline{\rho} + i\overline{\eta} = -(V_{ud}V_{ub}^*)/V_{cd}V_{cb}^* \tag{1.3.8}$$

take into account also higher order terms in the Wolfenstein parameterization:

$$\overline{\rho} = \rho(1 - \lambda^2/2 + \ldots) \tag{1.3.9}$$

$$\overline{\eta} = \eta(1 - \lambda^2/2 + \ldots) \tag{1.3.10}$$

This parameterization is extensively used, many of the experiments aiming to constrain CKM elements expressing their results in the $(\overline{\rho}, \overline{\eta})$ plane. Figure 1.3 shows a compilation of the experimental constraints on the apex of the unitarity triangle.

V_{ij}	Value	Error	Rel. error
V_{ud}	0,97377	0,00027	0,03%
V_{us}	0,22570	0,00210	0,93%
V_{ub}	0,00431	0,00030	6,96%
V_{cd}	0,23000	0,01100	4,78%
V_{cs}	0,95700	0,09300	9,72%
V_{cb}	0,04160	0,00060	1,44%
V_{td}	0,00740	0,00080	10,81%
V_{ts}	0,04060	0,00270	6,65%
V_{tb}	0,77000	0,24000	31,17%

Table 1.2: Magnitude of the CKM matrix elements

The elements of the CKM matrix have been extracted from flavor changing reactions involving mostly kaon, D and B decays and their CP-violation behavior. The

values obtained up to date and their relative uncertainties are reported in table 1.2, from [12]. Some of them are not yet independent of theoretical assumptions like, for instance the case of $|V_{td}|$ which assumes $|V_{tb}| = 1$.

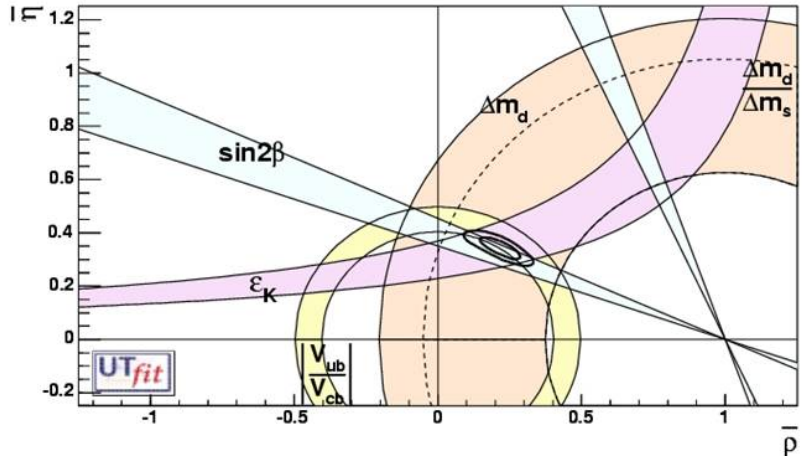


Figure 1.3: Experimental constraints on the apex of the unitarity triangle.

The P326 collaboration proposes to give a 10% accurate estimate of $|V_{td}|$ using the very rare kaon decay $K^+ \to \pi^+\nu\bar{\nu}$ at the CERN SPS [16]. This measurement will improve the knowledge of the CKM matrix and provide a decisive test of Standard Model.

The rare kaon decays of $K \to \pi\nu\nu$ form are of particular interest in the Standard Model framework. Flavor Changing Neutral Current (FCNC) processes are allowed only via loop diagrams that include internal quarks or intermediate bosons. The importance of rare semileptonic FCNC decays, $K_L \to \pi^0\nu\bar{\nu}$ and $K^\pm \to \pi^\pm\nu\bar{\nu}$, is amplified by their clean theoretical character[17]. This particular property comes from the following facts:

- the low energy hadronic matrix elements required are just the matrix elements

of quark currents between hadron states, which can be extracted with good accuracy from non-rare semileptonic decays;

- the main contribution to the FCNC processes comes from the region of very small distances ($\sim 1/m_t,\ 1/m_Z$) where perturbative QCD holds;

- these processes are sensitive to new physics like, for instance, contributions from heavy supersymmetric particles.

The branching ratio for $K^+ \to \pi^+ \nu\bar{\nu}$ can be written as [18]

$$Br(K^+ \to \pi^+ \nu\bar{\nu}) = k_+ \left[\left(\frac{Im\lambda_t}{\lambda^5} X(x_t) \right)^2 + \left(\frac{Re\lambda_t}{\lambda^5} X(x_t) + \frac{Re\lambda_c}{\lambda^5} P_c \right)^2 \right] \quad (1.3.11)$$

$$k_+ = r_{K^+} \frac{3\alpha Br(K^+ \to \pi^0 e^+ \nu)}{2\pi^2 sin^4\theta_W} \lambda^8 = (5.04 \pm 0.17) \times 10^{-11} \left(\frac{\lambda}{0.2248} \right)^8 \quad (1.3.12)$$

where $\lambda = V_{us}$, $\lambda_c = V_{cs}^* V_{cd}$, $\lambda_t = V_{ts}^* V_{td}$, $x_t = m_t^2/m_W^2$ and the $r_{K^+} = 0.901$ is the isospin breaking correction. The coefficients X and P_c are being computed numerically and the most recent Next-to-Next-to-Leading Order χPT calculation predicts[19]:

$$Br(K^+ \to \pi^+ \nu\bar{\nu}) = (8.0 \pm 1.1) \times 10^{-11} \quad (1.3.13)$$

where the error is almost entirely due to uncertainties in m_c and the CKM matrix elements.

Investigation of both $K^+ \to \pi^+ \nu\bar{\nu}$ and $K^0 \to \pi^0 \nu\bar{\nu}$ provides an alternative way for determination of the apex of the unitarity triangle. The comparison between this apex and the one obtained from other methods like the B-mesons allows a clean test of the Standard Model.

The study of such rare decay may appear difficult, but recent experiments have demonstrated that it is possible. BNL-AGS-E787 (E949)[20] have identified three such events in a three year data taking period. The follow-up experiment BNL-AGS-E949[21] is expected to increase the statistics, possibly reaching 10 events in the following years. As opposed to P326 proposal, the measurement has been performed with kaon decays at rest. Another initiative at Fermilab called *Charged Kaon at the Main injector* (CKM) proposed in-flight measurement of $K^+ \to \pi^+\nu\bar{\nu}$ branching ratio, targeting around 100 events statistics. Unfortunately the initiative lacked funding and was prematurely terminated.

The value of the branching ratio resulted from BNL-AGS-E787 (E949) is

$$Br(K^+ \to \pi^+\nu\bar{\nu}) = (14.7^{+13.0}_{-8.3}) \times 10^{-11} \tag{1.3.14}$$

The central value in (1.3.14) is higher than the Standard Model expectation in (1.3.13), but the result is compatible with SM if the large errors are accounted for. Figure 1.4 shows the mismatch between the present determination of the apex of the unitary triangle, based on Standard Model prediction (ellipses), and the projection in $(\bar{\rho}, \bar{\eta})$ plane of $Br(K^+ \to \pi^+\nu\bar{\nu})$ due to BNL-AGS-E787 (E949). This deviation might be the indication of new physics beyond the Standard Model, that may differentiate between B-meson and kaon sectors[17].

1.3.1 Background rejection in P326

The main background in studying $K^+ \to \pi^+\nu\bar{\nu}$ comes from the hadronic two-body decay $K^+ \to \pi^+\pi^0$ and the leptonic two-body decay $K^+ \to \mu^+\nu$, with branching ratios 10 orders of magnitude higher than the signal. In particular, the cases where the neutral pion of the hadronic decay is missed, or the muon is mistakenly identified

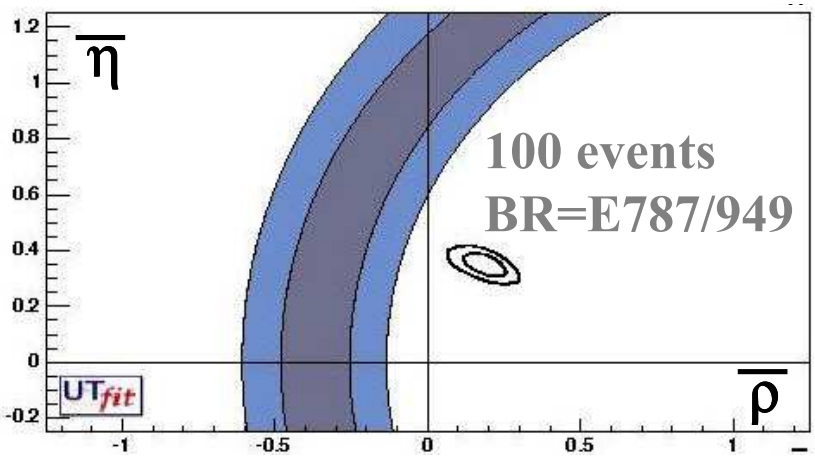


Figure 1.4: Possible new physics from the measurement of $K^+ \to \pi^+\nu\bar{\nu}$. Ellipse represent the present determination of the unitarity triangle apex; bands represent BNL-AGS-E787 (E949)result.[17]

as pion can fake the signal. Thus, a very good photon rejection (from the neutral pion decay) and particle identification (for muon detection) are needed, in conjunction with kinematical constraints. Other decay modes of the kaon are summarized in 1.3, together with the corresponding rejection method.

Figure 1.5 shows the missing mass distribution calculated for the $K^+ \to \pi^+\nu\bar{\nu}$ signal and the most important background decay modes. Two acceptance regions can be identified with missing mass ranging from 0 to the minimum missing mass in $K^+ \to \pi^+\pi^+\pi^-$ decay, separated by a narrow gap around $m^0_{\pi^0}$. The width of the gap is determined by the missing mass resolution of the experiment.

Simulations have shown[16] that signal-to-background ratio better than 10 and signal acceptance higher than 10% can be obtained considering neutral pion rejection inefficiency in the order of 10^{-8} and muon veto inefficiency of 5×10^{-6}, with a missing

Decay Mode	BR	Background Rejection
$K^+ \to \mu^+\nu$	63%	μ PID, kinematics
$K^+ \to \pi^+\pi^0$	21%	photon veto, kinematics
$K^+ \to \pi^+\pi^+\pi^-$	6%	charged particle veto, kinematics
$K^+ \to \pi^+\pi^0\pi^0$	2%	photon veto, kinematics
$K^+ \to \pi^0\mu^+\nu$	3%	photon veto, μ PID
$K^+ \to \pi^0 e^+\nu$	5%	photon veto, E/p

Table 1.3: Background decay modes

mass resolution of

$$(\Delta m)^2 \simeq 8 \times 10^{-3} GeV^2/c^4 \tag{1.3.15}$$

1.3.2 Experimental setup

1.3.2.1 The beam

P326 is using a 400GeV primary proton beam from the SPS at CERN with an intensity of 3×10^{12} protons per $5s$ pulse. The protons hit a beryllium target in order to produce the positive kaons required for the experiment. A series of achromats selects particles with central momentum of 75GeV, with momentum resolution about 1%.

A differential Cerenkov counter called CEDAR is responsible for tagging the positive kaons which represent 6% of the beam content.

1.3.2.2 The detector

Figure 1.6 shows the entire experimental setup of the P326 experiment. It basically consists of two spectrometers, one for the beam and another for the decay products

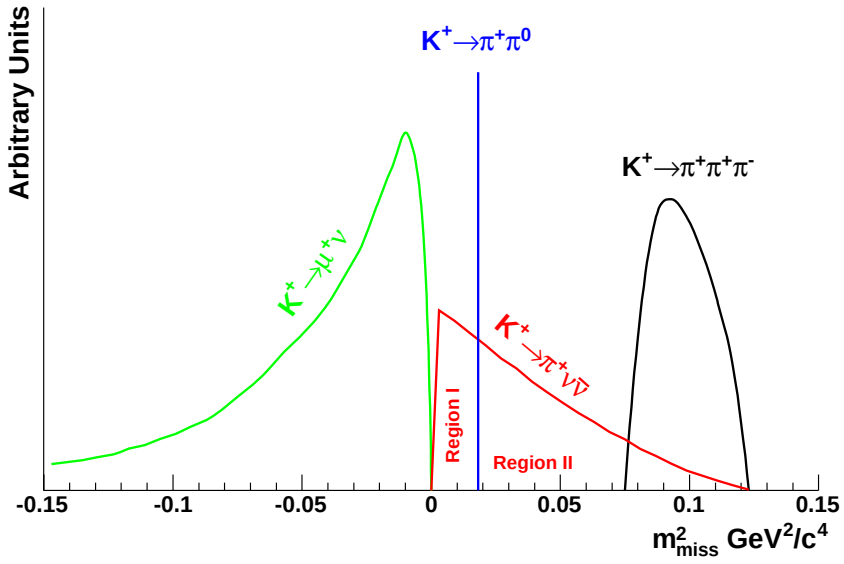


Figure 1.5: Missing mass distribution for $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and the concurrent decay modes

in the forward direction. Particle identification, photon vetoing and track momentum measurements are used with a certain degree of redundancy. The elements of detector together with a brief description are summarized below:

- GIGATRACKER: Three planes of thin silicon pixel detectors perform redundant momentum measurement on the beam components (beam spectrometer).
- ANTI: Ring-shape anti-counters surrounding the vacuum tank provide full coverage for photons originating from the decay region with angles up to 50 mr.
- STRAW TRACKER: A double magnetic spectrometer consisting of chambers of straw tubes measures the direction of the merging pion providing a redundant measurement of its momentum.

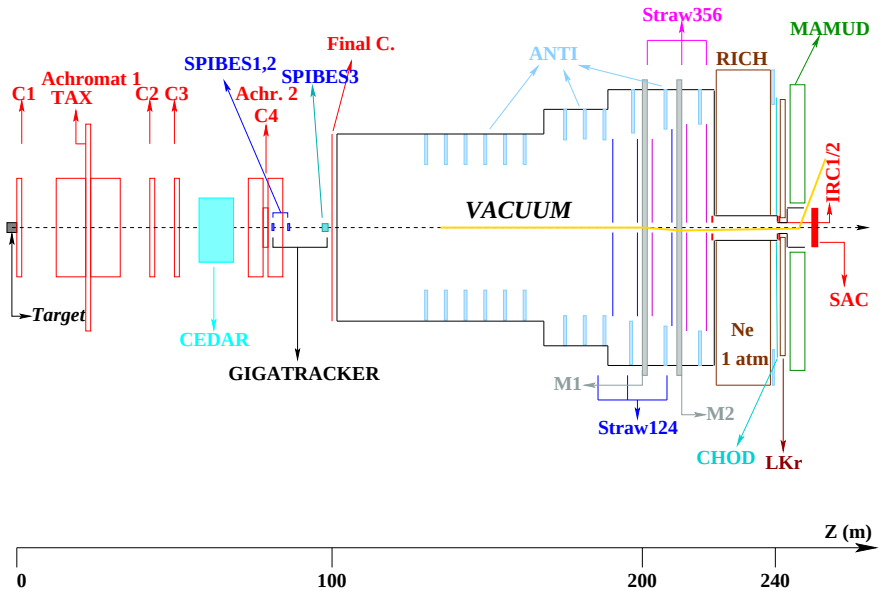


Figure 1.6: P326 experimental setup

- RICH: The ring Imaging Cherenkov counter provides muon and pion identification.

- CHOD: A multi-gap glass RPC hodoscope performs triggering and precise timing of the charged track.

- LKR: A liquid Krypton electromagnetic calorimeter acts as photon veto in the angular region between 1.0 and 15.0 mr.

- MAMUD: A magnetized hadron calorimeter performs high efficiency muon identification.

- IRC1-2, SAC: Intermediate ring and small angle photon veto calorimeters cover the angular regions around and in the beam.

1.3.3 Gigatracker requirements

The missing mass resolution required in order to reject kinematically the background can be obtained with kaon and pion momentum resolutions of 0.3% and 1%, respectively, and an angular resolution of the angle between the kaon and pion directions in the order of $50 - 60\mu rad$ [16].

This specifications define the required performance of the upstream and downstream spectrometer, namely the Gigatracker and the Double Spectrometer.

While the spatial resolution of the beam spectrometer has small influence on the missing mass resolution, allowing for a relatively big pixel size ($300 \times 300\mu m^2$), time resolution has an important influence. Giving the high particle rate crossing the three silicon stations (60 MHz/cm^2 in average), a pion track measured in the downstream spectrometer may be wrongly associated to a kaon candidate in the Gigatracker, so the kinematical rejection power is degraded.

For time resolution of 150ps, simulations show that about 36% of the events have more than one track in the Gigatracker. On the other hand, the probability of wrong matching is about 8% for cases with two tracks in the Gigatracker. The later is dependent on the matching algorithm, therefore may improve when software features are enhanced.

Another consequence is the need for a local time-to-digital converter with bin size as low as 150ps at the front-end level. This enhances the complexity of the chip and puts pressure on the power budget distribution. Several iterations at the collaboration level have set a compromise between the beam geometry and the layout of the silicon station allowing for smaller dimensions of the readout chip than initially proposed. A configuration of 2×5 readout chips with pixel column length of 13.5 mm has been

agreed on. This provides better management of on chip resources and power delivery scheme.

Another important requirement is the material budget. The target is $0.4\%X_0$, with silicon sensor thickness of 200 μm and readout chip thickness of $100\mu m$.

The radiation environment must also be considered. The accumulated design fluence for the Gigatracker sensors is 1.8×10^{14} 1MeV $neutrons/cm^2$ maximum, corresponding to a radiation exposure total dose (estimated for 100 days operation) of about 10 Mrad. Available radiation hard technologies for both sensor and front-end fabrication may not assure operation up to this dose, without substantial degradation in performance. Therefore, it is proposed to replace the silicon stations after a maximum threshold fluence and to operate the sensors at lower temperature (a couple of degrees below zero), in order to extend their lifetime under radiation[22].

1.4 PANDA and QCD

The proposed PANDA experiment will enable the new FAIR facility at GSI to play a significant role in strong interaction physics, providing a link between nuclear physics and hadron physics. The experiment proposes to study interactions of antiprotons with nucleons and nuclei in a fixed target experimental setup.

Antiproton beams in 1-15GeV energy range, stored in the High-Energy Storage Ring (HESR) for in-ring experiments, will provide access to the heavier strange and charm quarks and to copious production of gluons. The proposed program, using resonant antiproton-proton annihilation, is a quantitative and qualitative extension of successful experiments performed recently at the antiproton accumulator at FNAL, USA. Accessible energies are enhanced by a factor of almost 2, while luminosity by a

one order of magnitude reaching $2\times10^{32}cm^{-2}s^{-1}$. At FAIR, advanced antiproton cooling techniques will enable high energy resolution (momentum spread of $10^{-4}-10^{-5}$) and a more versatile detector setup will be employed allowing for the first time a measurement of both electromagnetic and hadronic decays with high precision. The goal is to achieve comprehensive precision spectroscopy of the charmonium system for a detailed study, particularly of the confinement part of the QCD potential. This in turn will help to understand the key aspects of gluon dynamics which are being investigated and quantitatively predicted in the framework of Lattice QCD. The charm quark is sufficiently heavy to lend itself to non-relativistic perturbative treatment far more reliably than the light up, down, and strange quarks. Thus, an optimal testing ground for a quantitative understanding of confinement is provided by charmonium spectroscopy, i.e. the spectroscopy of mesons built of charmed quark-antiquark pairs $(c\bar{c})$.

A central part of the antiproton program presented in this proposal is the first search for gluonic excitations, glueball and hybrids, in the charmonium mass range where they are expected to be less mixed with the multitude of normal mesons. The unambiguous determination of the gluonic modes would establish an important missing link in the confinement problem of hadrons.

The proposed experimental program at the HESR will address the open problem of interactions and in-medium modifications of hadrons with charm quarks in nuclei. On the one hand, this is an extension of the present GSI research program. On the other hand, the program will provide the first insight into the gluonic charmonium-nucleon and charmonium-nucleus interaction. A quantitative knowledge of charmonium-nucleon cross sections is considered to be of crucial importance in the

identification of the formation of the quark-gluon plasma in ultra-relativistic heavy-ion collisions.

Antiproton beams at the proposed facility will allow efficient production of hypernuclei with more than one strange hadron. The program opens new perspectives for nuclear structure studies and is a novel complement to the proposal to study the structure of nuclei with radioactive beams. The nucleon with the strange quark (hyperon) is not restricted in the population of nuclear states as neutrons and protons are. These exotic nuclei offer a variety of new and exciting perspectives in nuclear spectroscopy and for studying the forces among hyperons and nucleons.

The observed dominance of matter in the universe may be attributed to CP violation, an effect directly observed in the decay of neutral kaons and, very recently, in B mesons. CP violation can be studied in the charm meson sector and in hyperon decays, with the HESR storage ring running at full luminosity. An observation of significant CP violation would indicate physics beyond the Standard Model.

1.4.1 Experimental setup

The general layout of PANDA is based on two magnetic spectrometers. A target spectrometer consisting of a solenoid surrounds the interaction region and a forward spectrometer with a dipole magnet provides angular coverage for the most forward angles. The combination of two spectrometers offers full angular coverage, taking into account the wide range of energies and the flexibility requirement. It is foreseen that some of the components of the detector setup be exchanged or added for specific experiments that may require so. Both spectrometers include instrumentation for charged-particle track reconstruction, particle identification, energy and momentum

measurement, as well as photon and neutral particle detection.

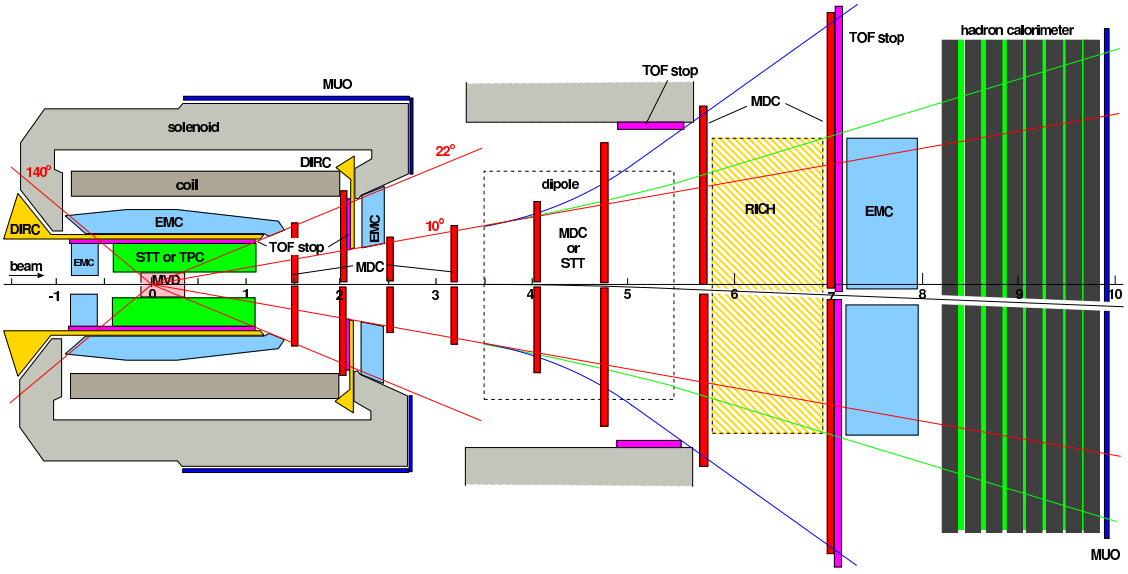


Figure 1.7: Top view of the PANDA detector

The target spectrometer consists of several detectors in a shell-like arrangement. Most of the instruments are contained within the coils of the solenoid:

- a complex silicon micro-vertex detector (MVD) for tracking and particle identification;
- a second tracking detector built out of straw tubes (STT) or a time projection chamber (TPC);
- a ring-imaging Cherenkov (RICH) counter for particle identification;
- two sets of mini drift chambers (MDC);

- another Cherenkov detector based on aerogel RICH or detection of internally reflected Cherenkov light (DIRC);
- an electromagnetic calorimeter (EMC) for photon detection.

Outside the return yoke of the solenoid scintilator bars for muon identification are foreseen (MUO).

The forward spectrometer covers the angular region below 10° and 5° in horizontal and vertical direction, respectively. The current design is based on a 1 m gap dipole and tracking planar devices distributed perpendicular to the beam direction for momentum analysis of charged particles. A shashlyk-type calorimeter consisting of lead scintillator sandwiches (EMC) will detect the photons in the forward range, while other neutral and charged particles with high momentum will be detected in the hadron calorimeter and the muon counters (MUO).

An additional time-of-flight measurement across the two spectrometers is under discussion. It would consist of a start detector within the MVD (possibly one of the layers of the MVD) and stop detectors at the exit window of the target spectrometer within the forward spectrometer.

1.4.2 MVD requirements

Experiments aiming at the identification of open charm and strangeness depend on the identification of secondary decays of particles in displaced vertices. Strange hadrons have decay lengths on the order of cm while charmed mesons and baryons decay within several tens to a few hundred μm. The vertex detector and the tracking capability of the target spectrometer will have to be suited for both. While PANDA will focus on charm production, the identification of kaons will greatly enhance the efficiency for

D mesons with large branchings into channels accompanied by kaons. Thus, a design that can handle both length scales will greatly improve the efficiency of PANDA.

With the interaction rates foreseen for the PANDA experiment, expected in experiments using heavy targets, the innermost detector layers will be exposed to high particle densities. A pixel design is favored for the innermost detectors.

The micro vertex detector (MVD) is foreseen to be mounted very close to the interaction point. It consists of five barrel shaped layers plus five disk-shaped detectors in forward direction. The three innermost layers are composed of pixel detectors to achieve best resolution and to be able to easily detect decay vertices displaced from the interaction point. The outer layers may be composed of microstrip detectors which are easier to handle. The baseline technology chosen for the pixel detectors is the hybrid active pixel sensor for which extensive knowledge is provided by several LHC experiments. The use of already mature designs from those experiments, as well as completely new custom design are the options. As alternatives to silicon pixels, GaAs based detectors are considered as well as much thinner monolithic pixel sensors where the problem of radiation hardness would have to be solved.

PANDA will have to handle significantly slower particles as well. This offers the opportunity to contribute to the particle identification with energy loss measurements. However, the detector thickness will be a crucial parameter for the detector design compared to other applications.

Spatial resolution requirements derived from the secondary vertex resolution favors a $100 \times 100 \mu m^2$ pixel size, which is smaller than the existing LHC designs. The electronics will have to feature high dynamic range amplitude measurement matching the momentum range of the particles crossing MVD. Since the operation of the PANDA

experiment does not foresee any selective trigger, in order to maximize physics yield, the front-end electronics must provide continuous readout. All this arguments favor a custom solution for the MVD pixel design.

1.5 ALICE and quark-gluon plasma

1.5.1 QGP probes in heavy-ion collisions

Traditionally the QGP probes are classified in three classes[23]:

- *soft-probes* - with typical momentum below 1 GeV/c;
- *heavy flavour-probes* - particles containing heavy quarks like c or b quarks;
- *high-p_T probes* particles with momentum range above several tens of GeV/c.

The last two probes are collectively known as hard probes.

Soft-probes

The measurement of global event observables, like the charged-particle multiplicity or the charged-particle pseudo-rapidity distribution, can bring interesting information about the thermodynamical parameters of the reacting matter after collision. The initial energy density ε in the central interaction region is linked to the transverse energy distribution through the Bjorken formula:

$$\varepsilon = \frac{1}{\pi R^2 c\tau_0}\frac{dE_T}{dy} \tag{1.5.1}$$

where R is the radius of the nuclei and τ_0 is the initial equilibration time $\sim 1fm/c$ [24]. The temperature T can be expressed by the average transverse momentum $\langle p_T \rangle$.

This means that $\varepsilon - T$ plot (one possible representation of the nuclear equation of state) can be studied by representing $\langle p_T \rangle$ as function of dE_T/dy.

In reality, the Bjorken formula is model dependent especially through τ_0, so it cannot point out a phase transition alone.

Moreover, the multiplicity information can test hadroproduction models. For instance, the phenomenological approach[25] predicts that the pseudo-rapidity distribution $dN_{ch}/d\eta$ is a linear combination of the number of participants N_{part} and the number of binary collisions N_{coll}.

Leptons can escape from the dense and strongly interacting reaction region, thus providing information about the condition at the time of their creation. In particular, ρ, ω or ϕ vector mesons, with the lifetime comparable to that of the QGP and which can decay into resonant lepton pair ($l\bar{l}$), may change their mass or width when they are produced inside the dense medium. Partial chiral symmetry restoration or final state interaction may be responsable for this phenomena. The latter can be sorted out by comparing the results obtained in the hadronic and leptonic decay modes. Previous nucleus-nucleus collision experiments at SPS at CERN have reported significantly higher dilepton yield in the low mass region (below 600 MeV), for central collisions, than the yield expected from hadron decays.

Another interesting aspect is the difference between the lifetimes of $\rho(\sim 1.32 fm/c)$ and $\omega(\sim 23.4 fm/c)$. Therefore the relative effect on the properties of the two mesons may represent a dynamic probe of the QGP.

Other important phenomena at this energy range is the predicted strangeness enhancement due to a lower threshold for strangeness production in de-confined medium [26], and can flag the QGP formation.

Heavy-flavour probes

Hard partons created in the first instants of the nuclear collision in primary partonic scatterings are expected to lose energy by gluon radiation while crossing the dense medium created after the collision [27]. Due to "dead-cone" effect, the energy loss of massive partons like the charm or beauty quarks will be smaller than the energy loss of "light" quarks. Therefore, comparing the attenuation of D or B mesons with lighter counterparts will allow to test the consistency of the interpretation of quenching effects as due to energy loss in a deconfined medium and to investigate the density of the QGP medium.

Another effect on heavy probes is the quarkonium suppression due to the Debye screening in the deconfined matter. J/ψ suppression has been already observed at SPS[28], but the situation may be complicated at higher energies by possible charmonium recombination in the hot medium. The recombination effect is less likely to affect bottomium, due to its higher mass.

High-p_T probes

High-p_T partons produced in initial parton-parton scatterings fragment into jets that cross the deconfined matter, suffering significant energy losses. This effect has already been observed at RHIC as high-p_T jet suppression[29] and the suppression of back-to-back jet correlations[30].

1.5.2 The ALICE detector

ALICE (A Large Ion Collider Experiment) was defined in its proposal document[31] as a general-purpose heavy-ion experiment designed to study the physics of strongly interacting matter and the quark-gluon plasma (QGP) in nucleus-nucleus collisions at

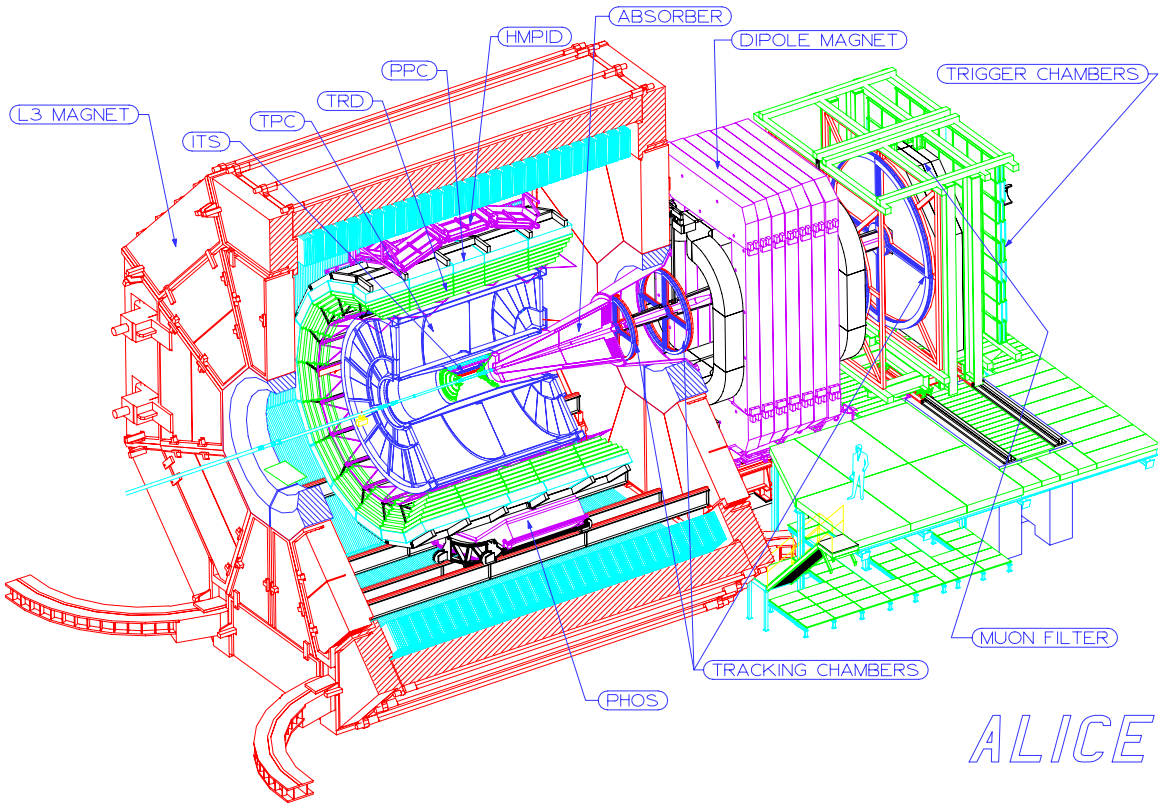


Figure 1.8: The ALICE detector

LHC. In addition to the heavy A-A collision program, the experiment will gather data in p-A and p-p runs in order to provide reference data for the heavy-ion program, but also for genuine p-p and p-A studies. The foreseen collision energy for the heaviest ion is $\sqrt{s} = 5.5TeV$, which is 30 times higher than RHIC.

The detector is divided in central and forward parts (see figure 1.8). The first gathers detectors for hadrons, electrons and photons, covering the mid-rapidity region, while the forward part is a muon spectrometer.

The central part consists of the Inner Tracking System (a silicon based tracking detector), the Time-Projection Chamber (a cylindrical gaseous chamber for tracking and momentum measurement of charged particles), three detctors for particle identification (Time-of-Flight (TOF) detector, Transition-Radiation detector (TRD) and a single-arm ring Cherenkov imaging (HMPID) detector) and a single-arm electromagnetic calorimeter (PHOS). All central detectors are embedded in a large L3 solenoidal magnet. The forward muon spectrometer consists of a massive absorber, a dipole magnet and 14 planes of tracking muon chambers. Zero-Degree Calorimeters (ZDC) cover the beams' rapidity in order to measure the spectator nucleons in heavy-ion collision.

The track reconstruction efficiency is about 90% at $p_T > 1GeV/c$, being dominated by the TPC. At lower momentum the tracks are reconstructed mainly in the ITS. All detectors of the central part, except the electromagnetic calorimeter, have particle identification capabilities. The total momentum coverage for particle identification is from 0.1 GeV/c to a few GeV/c. The choice of the magnetic field (0.5T) is a compromise between the reconstruction efficiency and the resolution in the high-p_T region.

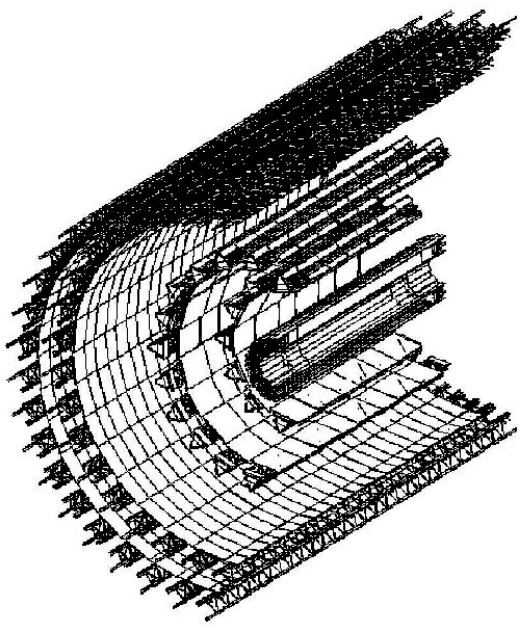


Figure 1.9: Section of the Inner Tracking System

1.5.3 The Inner Tracking System

The ITS is a 6-layer cylindrical tracking detector based on three silicon technologies (figure 1.9: silicon pixel detectors (SPD) for the inner 2 layers, silicon drift detectors (SDD) for the middle 2 layers and silicon microstrip detectors (SSD) for the outer 2 layers.

The SPD detectors provide only spatial resolution and high granularity, being equipped with binary readout. The SDD and SSD have analog readout offering momentum measurement and particle identification capability in the non relativistic region.

Featuring high granularity and spatial resolution, as well as sufficient momentum resolution via energy loss measurement in the lower momentum range, the ITS contributes to the determination of multiplicity distributions and inclusive particle

Parameter	Silicon Pixel	Silicon Drift	Silicon Strip
Spatial precision $r\phi$ (μm)	12	38	20
Spatial precision z (μm)	100	28	830
Two track resolution $r\phi$ (μm)	100	200	300
Two track resolution z (μm)	850	600	2400
Cell size (mm^2)	50×425	150×300	95×40000
Active area per module (mm^2)	12.8×69.6	72.5×75.3	73×40
Readout channels per module	40960	2×256	2×768
Total number of modules	240	260	1698
Total number of readout channels (k)	9835	133	2608
Total number of cells (M)	9.84	23	2.6
Average occupancy (inner layer) (%)	2.1	2.5	4
Average occupancy (outer layer) (%)	0.6	1.0	3.3
Power dissipation in barrel (W)	1500	1060	1100
Power dissipation end-cap (W)	500	1750	1500

Table 1.4: Parameters of the ITS sub-detectors

	SPD	SDD	SSD
Layer 1	1.0	1.1	0.87
Layer 2	1.0	1.1	0.87
Thermal shield/support	0.36	0.29	0.42

Table 1.5: Material budget for each layer of the ITS. Values are given in $\%X_0$

spectra, the study of resonance production (ρ, ω and ϕ), and in particular the behavior of the mass and width of mesons in a dense medium. It also contributes to the mass measurements for heavy-flavour states, improving the signal-to-background ratio in the study of heavy-quarkonia suppression, such as J/ψ and Υ and the observation of jet production and jet quenching, i.e. the energy loss of partons in strongly interacting dense matter.

The main tasks of the ITS are to reconstruct the primary vertex and the secondary vertices from decays of hyperons, D and B mesons, to track and identify particles with momentum below 100 MeV and to assist the TPC detector improving the momentum and angle resolution for high-p_T particles and covering, with limited precision, the dead regions of the TPC. Table 1.4 summarize the design requirements that individual sub-detectors have to fulfill in order to achieve the above experimental goals.

Another constraint is the maximum allowed amount of material for each layer of detectors (See table 1.5). Especially in the case of particles with small transverse momenta, momentum and impact parameter resolution are dominated by multiple scattering effects in the material of the detector.

1.5.4 The Silicon Drift Detector subsystem

Silicon drift detectors were chosen for the 2 intermediate layers of the ITS of ALICE because they feature very good multi-hit capability doubled by good energy loss resolution, with a small number of readout channels [32]. The SDD subsystem consists of 260 detector modules geometrically disposed in 2 cylindrical layers of 15 and 24 cm radius, for a total detection area of 1.31 m^2.

The single detector is divided in two distinct depletion regions by the high-voltage polarization net, in order to contain the maximum drift time below $5\mu s$ with a reasonable high-voltage. Consequently, the signal is readout via two arrays of 256 anodes at two opposite sides of the detector. The anode pitch is 294 μm. Each drift region is foreseen with three rows of 33 MOS injectors used to monitor the drift velocity during operation.

On each side a complex upilex/Al hybrid structure mounted on a carbon-fiber heat dissipater holds the front-end chips and the few small-size SMD components they need. The support of the detectors and the front-end electronics is entirely realized with non-standard carbon-fiber material in order to meet the tight material budget.

The front-end electronics samples the signal of each anode at a frequency of 40.08 MHz supplied by LHC. The size of the sensitive element (cell) is $294 \times 202\mu m^2$, corresponding to 89.1×10^3 cells per detector. The average spatial resolutions in the drift ($r\phi$) and anode (z) directions are 35 μm and 25 μm respectively. The detection efficiency is larger than 99.5% for amplitude thresholds as high as 10 times the electronic noise.

The front-end tasks are performed by two ASICs. The first is a 64-channel mixed-signal chip (PASCAL) containing low-noise amplifiers, a 256-cell analog memory per channel and 32 SAR-ADC with 10 bit resolution. Two consecutive analog channels are multiplexed to one ADC. The second ASIC is a digital control and de-randomization buffer (AMBRA) which can store up to four events. Four pairs of PASCAL and AMBRA chips are integrated on an upilex hybrid structure, and readout one drift region of the silicon drift detector. Power and data connections are supplied through one flexible upilex flat-cable that runs to the end-cap region of the ITS.

Chapter 2

Semiconductor detectors and front-end electronics in high-energy physics

2.1 Silicon detectors

Early particle detection tools that made possible the discoveries of 20th century were the ionization-chamber and the cloud-chamber invented in early 1900s by Charles Thomson Rees Wilson and Hans Johannes Wilhelm Geiger. Later on, in 1952, Donald Glaser invented the bubble chamber which dominated the fixed target experiments from the late 1950s until early 1970s providing excellent spatial resolution and complete solid angle coverage.

Electronic particle counters, mainly gaseous detectors, were employed only for specific reactions with small number of particles in the final state. Nevertheless, while

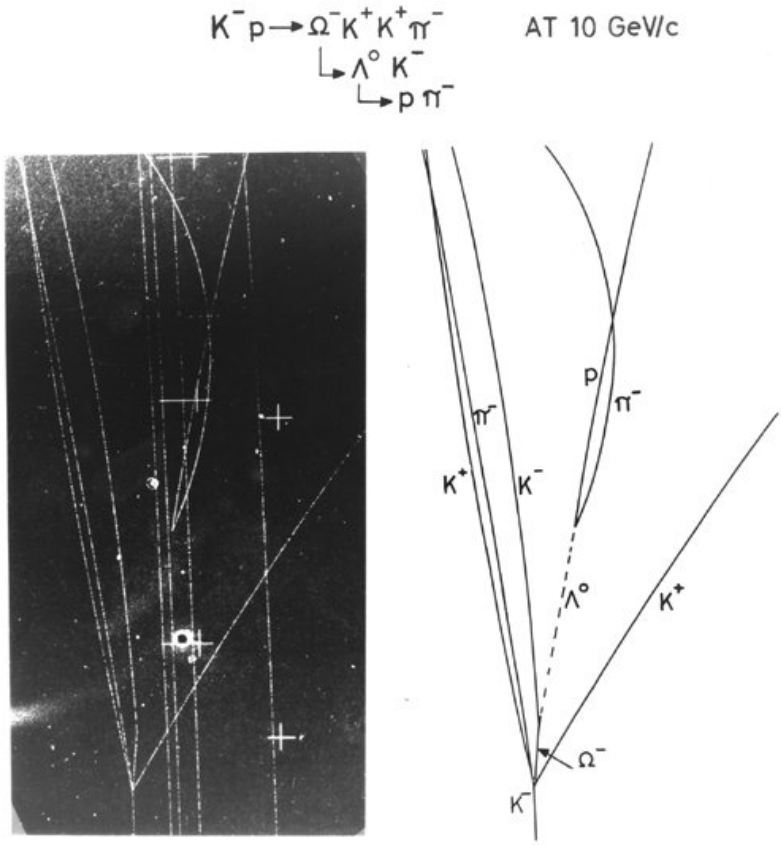


Figure 2.1: Omega minus production in 10 GeV/c K^- p interactions at CERN [33].

the energy and luminosity of the experiments grew higher, event-by-event analysis became impractical with the available tools. Electronic detectors like the position sensitive gaseous detectors (Charpak 1970) came into light offering the possibility to acquire selected data at high rates by triggering the data acquisition system on a characteristic signature of an interesting event and to store the experimental data directly on a computer.

The first semiconductor detector was a employed in 1949 by McKay who managed to measure the polonium α-ray spectrum using a Ge surface barrier detector [34].

Throughout the following decades, though, this type of detector was rather scarcely used in particle physics as it did not provide the required space coverage or the required position resolution at the time. Nuclear physics and x-ray spectroscopy remained the main application fields for semiconductor detectors, taking advantage of the higher ionization yield than in the case of gaseous detectors. Owing to a moderate band gap (1.12 eV in the case of silicon), the energy required to create an electron-hole pair is in the order of 3-4 eV, compared to an ionization energy of 15.75 eV in the case of Argon gas. Germanium detectors with band-gap energy of 0.66 eV are usually used in low-energy spectroscopy applications, where energy resolutions below 1 keV can be obtained. Many semiconductor materials were investigated in both detector technology and transistor industry, but silicon became the standard in both cases.

Moreover, as the industry of integrated circuits was rapidly growing, doubling the number of devices integrated on a chip each year (as observed by Moore - a co-founder of Intel, in 1965), semiconductor detectors increased in size and complexity. Silicon microstrip and silicon drift detectors were born as solid-state analogues of the gaseous multi-wire proportional chamber and drift chambers, offering both good energy resolution and increasingly good spatial resolution. Semiconductor detectors based on genuine charge transport mechanisms like *charge-coupled devices (CCDs)* were also invented.

In 1974, the discovery of the J/ψ meson marked the introduction of a new quantum number called charm and the start of a frenetic search for other particles carrying this quantum number. The experimentalists faced new challenges as the predicted lifetime of charmed particles was in the range of picosecond. Direct observation

implied spatial resolution below 100 μm and high experimental statistics in order to be able to select the rare event of charmed production. While bubble chambers and emulsions did exhibit the required spatial resolution but were impractical from the high statistics point of view and gaseous detectors were already at the limit of their intrinsic performances, it was realized that semiconductor might become the solution.

Basically two ways of tackling the problem emerged. One took advantage of the fact that a large fraction of charmed particles decay into three or more particles, whereas lighter strange particles usually exhibit two body decays. High multiplicity in the final state provided a signature for a possible charmed particle decay. "Active" targets were employed using stacks of fully instrumented silicon layers, either as p-n junctions or surface barrier devices. Providing good linearity and energy resolution these devices could deliver a signal whose height is a measure of the number of the particles crossing them, assuming they are all approximatively minimum ionizing, efficiently discriminating high multiplicity events.

Another approach was to mount detectors with high spatial resolution and high granularity very close to the target and reconstruct off-line the traces of emerging particles. Events where particle tracks are found to originate downstream the target in a secondary vertex are strong candidates for charm decay. The high granularity and good spatial resolution of newly developed semiconductor detectors qualified them to be used as vertex detectors. Moreover, the introduction of planar technology by J. Kemmer in 1980 [35], boosted the use of silicon detectors (mainly microstrip detectors) in high-energy physics The more recent discovery of the radiation hardness properties of oxygenated silicon[36], consequence of an intensive R&D work (see RD50 at CERN), consented their use in the modern high luminosity experiments at LHC

at CERN and other modern facilities [37].

Table 2.1 shows a list of examples of silicon detectors used in high-energy physics experiments. The list is far from being complete, but it shows the explosion of silicon detectors in high-energy physics, in term of total silicon area and number of channels. Some early analysis of the detectors to be employed in the next high-energy Linear Collider foresee a total silicon area as much as 3000 m^2.

Microstrip detectors have received overwhelming interest in the past years due to their simplicity and robustness. Silicon drift detectors have been employed mainly in the heavy ion experiments like NA45/CERES, for example, where very high multiplicity of the final state calls for true two-dimensional detectors with high granularity, and interaction rates are, in general, one or two orders of magnitude lower than those of particle colliders.

Good performance and robustness of the charged-coupled detectors employed in the SLD experiment at the Standford Linear Collider demonstrated their utility as vertex detectors, but this example remained rather singular in high-energy physics. Very high readout time of charged-coupled detectors prohibited their use in most of the large-scale high-energy experiments, with interaction rates of $10^8 s^{-1}$ and beyond. CCD have made their way in other domains like synchrotron radiation experiments, where they became standard instruments.

2.2 Detector technology

From functional point of view, semiconductor detectors can be classified in two main categories [38]: *charged-coupled devices* and *p-n junction devices*. Both types share the same signal formation mechanism but they differentiate through the acquisition

Exp	Loc	Year	Exp. Type	Det. Type	Area (m^2)	Min. pitch (μm)	Technology	ch	Res. μm
NA1	CERN	1982	fixed target	active target			ion implantation + surface barrier		
NA11	CERN	1984	fixed target	microstrips		20	single sided		
WA92	CERN	1992	fixed target	microstrips		10	single sided	8192	1,8
Mark2	SLC	1990	e-e+ collider	microstrips		25	single sided	18432	7,1
SLD	SLC	1991	e-e+ collider	CCD		22		$1.2\,10^8$	5
Delphi	CERN	1990	LEP	microstrips	0,42	25	single sided	$7.4\,10^4$	8
Aleph	CERN	1991	LEP	microstrips	0,25	19	double sided	$7.4\,10^4$	12
OPAL	CERN	1991	LEP	microstrips		25	single sided		6
CDF	Fermilab	1992	Tevatron	microstrips		55	single sided	$4.6\,10^4$	
L3	CERN	1993	LEP	microstrips		19	double sided		
CDF		2000							
D0		2000							
BaBar									
STAR SSD	RICH			microstrips	1		double sided	$5\,10^5$	
STAR SVT	RICH	2001		silicon drift	0,86	250		10^5	
CMS	CERN	2007	LHC	microstrips	210	80	single sided	10^7	
ATLAS	CERN	2007	LHC	microstrips	61	80	tilted single sided	$6.6\,10^6$	

Table 2.1: Examples of silicon detectors used in high-energy physics experiments

mechanism. In both cases, a particle crossing the silicon volume losses a fraction of its energy mainly through *ionization*. In the depleted silicon bulk, the ionization process creates electron-hole pairs. Under the influence of an electric field, electrons and holes travel, respectively, to the positive and negative electrodes, where they are either stored in a potential well (the case of *charge-coupled devices*) or they give rise to a pulse that can be measured in an outer circuit.

Charged-coupled devices (CCDs) are essentially an array of capacitors where the signal charge is trapped close to the surface of the device in a potential well. Later, in readout phase, the charge is transported serially from one capacitor to another. The stream of signals corresponding to each capacitor (*pixel*) is supplied to one amplifier (usually integrated on the same chip) and then fed out to the data acquisition system.

In the case of *p-n junction detector* type, the charge signals flaws directly through the terminals of the diode. Low input-impedance amplifiers are connected to one of the terminals (or both for double-sided microstrip detectors). Readout time is

considerably smaller than in the case of charged-coupled devices, but a large number of electronics channels is needed in order to readout a highly segmented detector.

2.2.1 Microstrip detectors

One of the first segmented semiconductor detectors and the most heavily used in high-energy physics experiments is the microstrip detector. Its concept is similar to that of the gaseous proportional multiwire chamber. Basically, long and narrow diode stripes are formed on the detector surface by ion implantation or diffusion. Typical dimensions of the strips are 8-15 μm wide and 20-50 μm pitch. Each formed diode (*strip*) is readout separately for a good spatial resolution. Several designs use capacitive interpolation in order to reduce the number of electronic channels, reading out every second or more strips (see figure 2.2). The charge cloud diffuses inside the silicon bulk being distributed over multiple strips. Centroid position of the charge can be reconstructed interpolating the signal from adjacent readout channels. One disadvantage of this type of detector is the high capacity of the strips than can reach 100 pF, with negative impact on noise performance.

Double-sided microstrip detectors

Both coordinates of a single hit can be recorded if strips with different orientation are formed on both sides of the silicon bulk. A projective readout has the major drawback in the fact that multiple hits can not be unambiguously reconstructed. If n particles hit the surface of the detector at the same time, there will be n strips firing in x an y directions, respectively, and n^2 possible combinations, or $n^2 - n$ ghost hits. Without any additional information, it is impossible to discriminate the real coordinates of the n hits. Perpendicular relative orientation of the strips maximizes

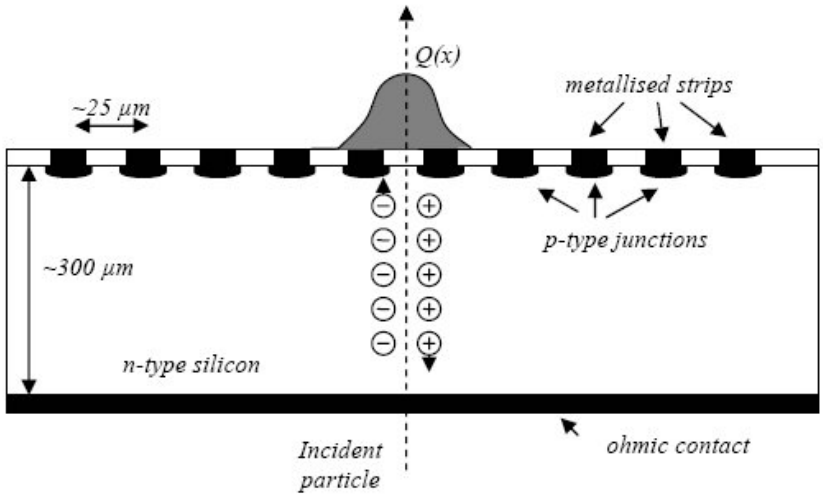


Figure 2.2: Microstrip detector. Capacitive charge division example.

this problem, since each strip corresponding to one coordinate intersects all the strips corresponding to the other coordinate. In practice, the relative angle between strips on the two sides can be small, reducing the number of intersections between strips on opposite sides of the detectors, thus, reducing the probability of "ghost" hits.

Processing both surfaces of the wafer increases the cost of the detector and the extra handling of the wafer may reduce yield. Additional electrical problems occur while processing the second surface. In order to tackle this problems, some of the modern detectors use two single-sided detectors tilted by some small angle.

2.2.2 Silicon drift chambers

Silicon drift detector is a novel semiconductor sensor concept introduced in 1984 by Gatti and Rehak [39]. As suggested by the title of the article announcing its invention, it is the analogue of the gaseous drift chamber. A transversal electric field

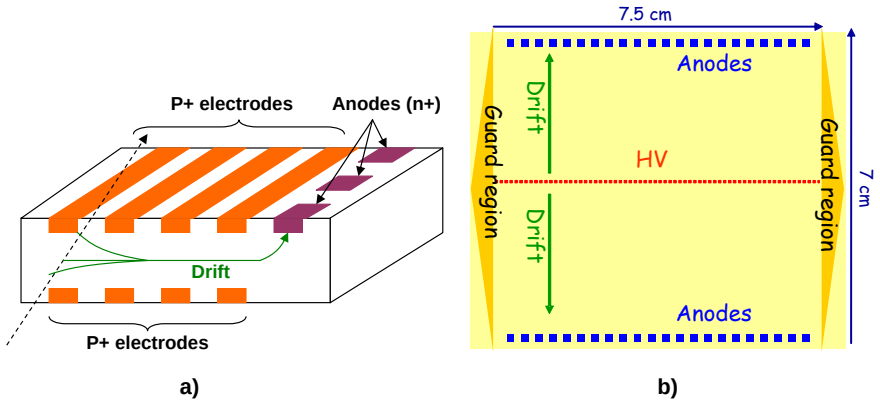


Figure 2.3: a) Principle of a silicon drift sensor; b) The ALICE silicon drift sensor (top view).

is applied to the fully depleted silicon volume. P+ implanted electrodes[1] connected to a resistive ladder (see figure 2.3) are responsible for the electric field. Particles crossing the sensor volume create electron-hole pairs through ionization. While the holes are immediately neutralized at the p+ electrodes, electrons drift at constant velocity under the influence of the transversal electric field, and are collected by an array of n+ anodes disposed perpendicularly to the direction of the drift field at the edge of the sensor. In this way one coordinate of the hit is retrieved from the centroid position of the collected charge, while the other coordinate information is carried by the drift time of the charge cloud. Moreover, the integrated charge corresponding to one hit is directly proportional to the energy lost by the particle passing through the depletion volume of the detector. Low noise level and good energy resolution can be achieved because of the small size of the anodes and their corresponding small capacitance ($\sim 50fF$).

[1]The bulk is supposed to be be n-type.

While drifting inside the silicon bulk for distances in the order of centimeters, the charge cloud spreads much more than in the case of microstrip detectors (where the charges drift just for a small distance, the diffusion being limited by the thickness of the detector - typically 300 μm). For this reason the pitch of the anodes can be considerably larger, without a significant degrade of spatial resolution.

2.2.3 Pixel detectors

Pixel detectors are two-dimensional arrays of integrated diodes operated in the depletion region for charged particles or photon detection. Typical dimensions of a single diode (*pixel*) is in the order of 50 - 300 μm. The early precursor of pixel detectors were called *pad detectors* having a typical pad dimension in the order of millimeters. Besides the excellent spatial resolution, the small dimension of pixel elements also offers great advantages in terms of radiation hardness and signal-to-noise ratio. Many radiation damage problems of silicon can be avoided by the use of small volume detector elements, whereas the small capacity (in the order of tens to hundreds of fF) ensure low-noise operation. True two-dimensional resolution is also a great feature for modern high-energy experiments, since high multiplicity events can be resolved un-ambiguously.

Hybrid pixel detectors

While integrating a high number of pixels on a silicon wafer is (since a lot of time now) not a difficult problem, interfacing them with readout amplifiers constitutes a great difficulty of pixel detector realization. In the last decades, the IC packaging industry performed huge steps towards miniaturization. One of the new emerging technology is the bump-bonding, which permits the realization of vertical contacts

through small size solder bumps with diameter in the order of 10 μm. This technology allows the interconnection of a pixel detector array to a geometrically matching readout electronic array fabricated on a separate wafer.

The detector and the readout electronics are designed separately and use different processes that are optimized for the specific device. The fabrication of the sensor requires a low number of masks and, usually, no lithographic step on the back side.

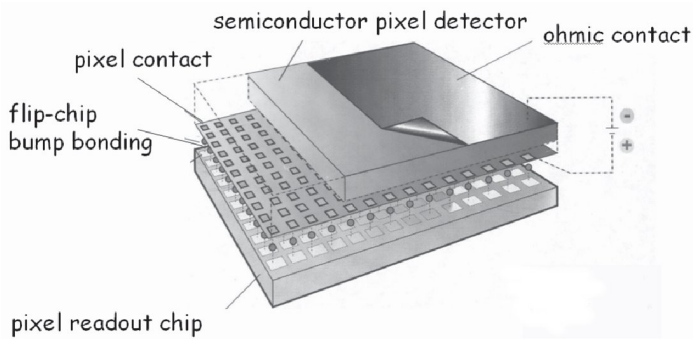


Figure 2.4: Hybrid pixel concept

All four LHC-based experiments use large scale ($\sim m^2$) pixel detectors using hybrid technology, which was matured over the last years in dedicated research and development projects and it is now considered state-of-the-art pixel technology.

2.2.4 Resolution considerations

There are several factors that influence spatial and energy-loss resolutions of semiconductor detectors. One should note that these factors do not individually determine one specific type of resolution but rather contribute in many ways to all of them. On the other hand, detector and readout electronics cannot be separated in this respect. The choices one does for a detector or a readout topology alone generally does

not lead to an optimum situation, but rather the combination of the two has to be considered when all requirements and practical limitations are assessed.

Spatial resolution The first parameter to be taken into account in order to evaluate the spatial resolution is the segmentation of the detector. If binary readout is considered, that is the measurement is blind to signal amplitude, the characteristic distance between two readout electrodes (*pitch*) determines the spatial resolution

$$\sigma = \frac{pitch}{\sqrt{12}} \tag{2.2.1}$$

The factor $\sqrt{12}$ comes from the integration of the random distribution of the *real* hit position with respect to the center of the sensitive electrode in the sensitive space corresponding to the electrode, which is equal to the *pitch*

$$\sigma^2 = \int_{-p/2}^{p/2} \frac{x^2}{p} dx = \frac{p^2}{12} \tag{2.2.2}$$

Another factor with influence on the spatial resolution is the thermal diffusion of carriers inside the semiconductor bulk. The charge cloud spread transversely while drifting toward the collection electrode. The spatial extension of the charge cloud depends on the diffusion coefficient D and the collection time t

$$\sigma_y = \sqrt{2Dt} \tag{2.2.3}$$

Taking into account the Einstein relation that expresses the diffusion coefficient w.r.t. the mobility μ of the carriers, and a linear approximation of the average drift field as the ratio between the bias voltage V_b and the detector thickness d ($\langle E \rangle = V_b/d$), one finds

$$\sigma_y \approx \sqrt{2\frac{kT}{e}\frac{d^2}{V_b}} \tag{2.2.4}$$

which is independent of the carrier type. For a practical situation where $d = 300\mu m$, $T = 300K$ and $V_b = 100V$ the transverse diffusion is $\sigma_y \approx 7\mu m$.

The transversal diffusion may lead to the situation where the charge is collected by two or more neighboring readout electrodes. By using an analog readout (sensible to the amplitude of the signal) one can use the fraction of charge collected by each electrode and interpolate for the true position of the hit. In this way, the resolution is increased beyond the limit given by the geometry of the detector.

In this case, the spatial resolution is limited by the *signal-to-noise* ratio (S/N). Using a simple centroid finding algorithm (which defines the centroid to be the mean of the charge distribution, without any weighting function) the worst-case spatial resolution can be calculated to be [40]:

$$\sigma = p(S/N)^{-1}\sqrt{\frac{N_{ch}(N_{ch}^2+2)}{12}} \approx p(S/N)^{-1}\sqrt{N_{ch}^3} \tag{2.2.5}$$

where N_{ch} is the number of electrodes included in the interpolation. Equation 2.2.5 shows that the best spatial resolution is achieved when the entire signal charge is collected by two electrodes. For this reason, the electrode pitch is usually chosen to match the diffusion width of the charge carriers[41].

Besides the charge diffusion, there are other phenomena that contribute to the width of the collection area. One is the angle of incidence of the particle track on the surface of the detector. The angle that minimizes the resolution is close to $arctan(pitch/thickness)$, which maximizes the probability of two-strip clusters[42].

The second phenomenon is the capacitive coupling of the collecting electrodes through the stray capacitance. This effect can be effectively used to decrease the number of electronic channels with a small resolution penalty, by connecting only every second or third electrode to a readout amplifier. The signal charge collected by

the floating electrodes is coupled capacitively to its neighbors, thus allowing for an accurate position measurement of the center of gravity.

The magnetic field causes the carriers to drift at an angle θ_L with respect to the electric field, given by [43]

$$tan(\theta_L) \sim \mu_H H \tag{2.2.6}$$

where μ_H is the Hall mobility and H is the magnetic field. This causes the charge cloud to spread over a higher area than diffusion does, for high magnetic field.

A more accurate analysis of the spatial resolution with respect to the detector parameters takes into account also the stray and bulk capacitance of the detector elements and in particular their ratio.

Energy resolution. For minimum-ionizing particles, the most probable charge deposition in a 300 μm thick silicon detector is about 3.5 fC which is equivalent to 22000 electron-holes. The statistical spread of this number is

$$\sigma = \sqrt{FN} \tag{2.2.7}$$

where N is the number of charge carriers generated through ionization and F is a correction factor called Fano factor which takes into account the multiple excitation mechanisms that reduce the overall statistical spread. The value of this factor for silicon is approximatively 0.12. This leads to about 50 electrons statistical spread or 0.23% relative resolution of the energy-loss measurement.

In fact, the energy deposited by the ionizing particle crossing the detector volume varies as a Landau distribution with $\sigma_E/E \approx 0.2$ for a 300μm thick detector[44] whose effect on the energy resolution is dominant. The only additional contribution which can be significant is the noise of the readout amplifier. If the ionizing particle crosses

three or more detector layers, one can compute a truncated mean over the multiple energy loss measurements in order to improve the resolution.

Electronic noise.

Silicon detectors are normally operated well below the threshold for avalanche multiplication of carriers and do not provide any intrinsic amplification. The small ionization signal is therefore amplified and shaped into one of a number of standard forms determined by practical considerations, in order to limit the bandwidth of the signal and noise for an optimum signal to noise ratio. There are three major noise sources in semiconductor detector-amplifier systems:

- i_{nd} shot noise from detector leakage currents (parallel),
- i_{nb} noise associated with biasing networks (parallel),
- e_n series noise associated with the amplifier itself.

The spectral density of the parallel noise sources can be expressed as

$$i_{nd}^2 = 2eI_{leak} \tag{2.2.8}$$

$$i_{nb}^2 = \frac{4kT}{R_{bias}} \tag{2.2.9}$$

where I_{leak} is the leakage current of the detector, R_{bias} the value of the resistor that biases the detector and e is the charge of the electron.

Normally the bias resistor R_{bias} is large so its contribution to the overall electronic noise is negligible. On the other hand, high-resistivity silicon bulk is used in order to minimize leakage currents and, thus, its contribution to noise.

The series noise due to the amplifier has one component independent of the frequency (*white noise*) and a frequency-dependent component with $1/f$ spectrum:

$$e_n^2 = S_w + \frac{S_f}{f} \tag{2.2.10}$$

The total noise at the output of the amplifier and signal processing system is obtained by integrating over the full bandwidth of the system. Since the useful amplitude information, in the case of semiconductor detectors, is the charge deposited by an ionizing particle, the noise level is generally expressed in terms of *equivalent noise charge*, defined as the detector charge signal that yields an signal-to-noise ratio equal to one. From a general analysis and neglecting the contribution of the bias network of the detector, the ENC can be expressed as [44]:

$$ENC_{tot}^2 = i_{nd}^2 A_i \tau_s + S_w \frac{A_w}{\tau_s} C^2 + S_f A_f C^2 \tag{2.2.11}$$

where τ_s, A_i, A_w and A_f are coefficients that depend on choice of the pulse shaping filter (τ_s being the characteristic *shaping time* or *peaking time* of the filter) and C is the total capacitance at the input of the amplifier, sum of the detector capacitance, input capacitance of the amplifier and parasitic capacitances. Equation 2.2.11 shows that, depending on the detector parameters and the choice of shaping filter, there is an optimal value of shaping time τ_s that minimizes noise.

2.2.5 Choice of detector topology

It should be noted from equation 2.2.11 that the equivalent noise charge increases with the the detector capacitance (and the related parasitic capacitance). Both bulk and parasitic capacitances are directly related to the dimensions of the detection volume. At first glance, it may seem for instance, that pixel detectors with capacitance in

the order of 100 fF, are better in this respect than strip detectors that typically show capacitance in the order of 1 pF per centimeter of strip. On the other hand, strip detectors need a much lower number of readout channels than pixel detectors do. Thus, for the same power budget allocated to the readout electronics, more power is available in the strip channels that partially compensate the high capacitance disadvantage. It is generally true, though, that strip detectors need higher shaping time than pixel detectors for the same spatial resolution requirement, with negative influence on the rate capability. Typical spatial resolutions obtained with analog readout are 20% to 80% better than the intrinsic geometrical resolution ($p/\sqrt{12}$).

The resolution of silicon drift detectors is subject of more complex analysis. In the anode direction the spatial resolution is governed by the same phenomena as those discussed earlier, with the most notable difference that the carriers drift longitudinally inside the detector bulk for distance up to several centimeters. This makes the diffusion spread of the charge cloud to be much higher than the pixel or strip detector situation, allowing a higher anode pitch for the same resolution requirement. On the other hand, the dimensions of the anode are comparable to the size of a pixel detector with positive effect on signal-to-noise ratio. Generally, with silicon drift detectors the resolution is several times better than the intrinsic geometrical resolution given by the anode pitch.

In drift direction there is no geometrical segmentation of the detector. The spatial resolution is given by the longitudinal spread of the charge cloud which traduces into a temporal broadening of the signal. The best spatial resolution is given by an optimal

filter transfer function given by the following differential equation [39]:

$$-\frac{C^2 e_n^2}{Q_s^2} w''(t) + \frac{2eI_{leak}}{Q_s^2} w(t) = 2\varepsilon_{min}^2 f'(t) \int_{-\infty}^{\infty} f'(\tau) w(\tau) d\tau \tag{2.2.12}$$

where Q_s is the signal charge, $f(t)$ is the normalized form of the signal shape at the anode, $w(t)$ is the filter weighting function and ε_{min}^2 is the minimum value of the time variance, corresponding to the best obtainable spatial resolution in the drift field direction. Electric field nonlinearities and charge recombination phenomena along the drift direction may also limit the resolution. STAR and ALICE have obtained resolutions in the order of $20\mu m$ over a drift path of 3 cm and 3.5 cm, respectively.

In general, in the design of tracking systems based on semiconductor detectors, pixel sensors with high granularity and true two-dimensional capability are preferred for the detectors elements that are in close proximity to the interaction point. At higher distance, the number of traces per unit of area decreases rapidly and strip or silicon drift detectors with lower multiple-track and rate capability can be used. Silicon drift sensors are good candidates when good spatial resolution and high hit multiplicity is targeted, provided that event rates are rather low (several hundreds of hertz).

2.2.6 Trends in semiconductor detectors

2.2.6.1 Hybrid pixels based on 3D integration

A new 3D integration process has been introduced in industry based on non-adhesive die-to-wafer or wafer-to-wafer bonding and interconnect technology, driven by the evolution of mixed-signal electronics and system-on-chip (SoC) concepts. Separate process flows are used to build wafers for analog and digital functions. The two

designs are thinned and placed one on top of the other and interconnections are made through deep vias. The number of layers is not restricted to two. This brings a new degree of freedom into the design, helping to avoid many compromises of the system-on-chip approach.

This technology can be exploited in the case of hybrid pixel integration. Sensor, low-noise amplification, digital conversion and de-randomization circuitry can be implemented in different layers in 3D integration process. The advantages of this approach are better optimization of the design of each layer, robust integration and low material budget, since individual active layers are thinned downto $\sim 7-15\mu m$ before integration. Successful application of this concept in sensor integration domain for industrial applications has been reported (see figure 2.5 [45]). Fermilab has started an R&D work to study this approach and its application in detector instrumentation for physics[46].

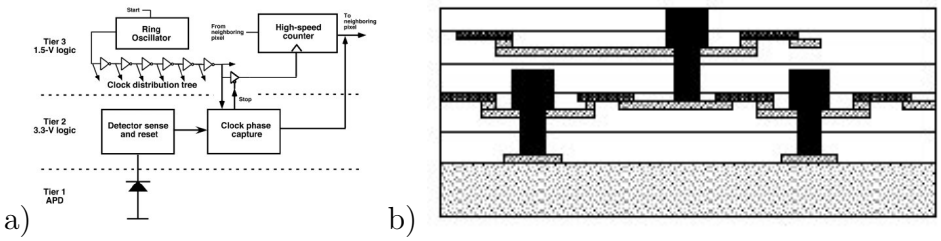


Figure 2.5: Sensor development using 3D integration technology. a) Schematic b) Vertical layout

2.2.6.2 Monolithic pixel detectors

Monolithic pixel detectors are semiconductor sensor that integrate the matrix of sensitive elements together with the amplifier and logic circuitry on the same wafer.

CMOS camera chips that can be found nowadays in consumer digital cameras and cellular phones are one example. The main differences between the consumer products and the sensors designed for X-ray or particle physics experiments are that the latter have larger area and 100% fill factor for efficient particle detection.

The *CMOS active pixel* approach uses standard CMOS technologies which have a lightly doped epitaxial layer, up to 15μm thick, between the low resistivity silicon bulk and the processing layers. The charge carriers are generated by a passing ionizating particle in the epitaxial layer which are then collected by an n-well/p-epi collection diode (figure 2.6a). The charges reach the collection electrode through thermal diffusion. The sensor volume is fully depleted only in the vicinity of the electrode so the collection efficiency is degraded due to recombination processes. The signal is in the order of some hundred to 1000 electrons.

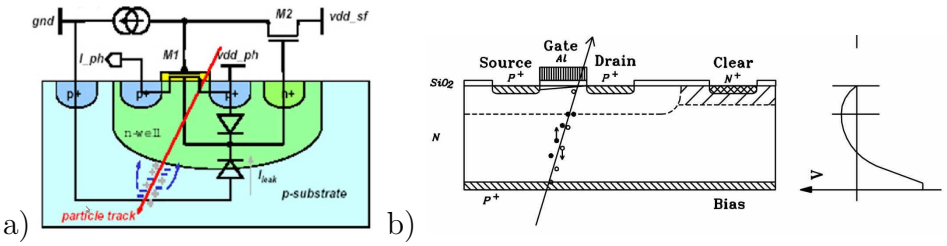


Figure 2.6: Monolithic pixel detectors. a) Photo-FET b) DEPFET

The simplest readout circuit is a 3-transistor matrix readout (reset, source follower stage and line-select). Another approach may include an PMOS transistor in the charge-collecting n-well whose gate voltage is affected by the signal charge, providing amplification. Signal-to-noise ratio of about 20 and spatial resolutions of 5 μm has been obtained, which may qualify them as good candidates for next generation particle experiments.

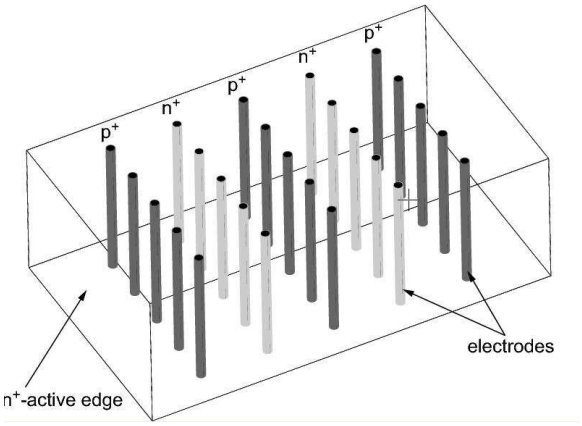


Figure 2.7: 3D detector concept

Another similar approach is the *DEPFET pixel detector* (figure 2.6b). The *depleted field effect transistor* is implanted in each pixel in a sidewards depleted bulk. The geometry of the transistor provides a parabolic potential distribution. The charge carriers generated in the bulk of the sensor are collected in a local minimum of the potential, just underneath the transistor channel, modulating the DEPFET channel current. The amplification of such structure is about $300 - 400 pA/electron$. Very low input capacitance ($\sim$ few fF) and direct amplification leads to noise figures below 10 electrons and good energy and spatial resolution.

2.2.6.3 Active edge 3D detectors

A new promising detector concept has been proposed [47] that replaces the planar geometry with a "3-dimensional" one. The electrodes penetrate the entire thickness of the detector, perpendicular to its surface, rather than being implanted on the surface. As a result, the charge carriers drift inside the bulk parallel to the surface of the detector over the short distance between electrodes (typically $\sim 50 \mu m$). This

means lower depletion voltage, faster collection times and better radiation tolerance. The edge of the sensor can be a collection electrode itself, extending the active area very close to the sensor boundary.

2.3 Front-end electronics

Front-end electronics for modern semiconductor detector systems are *application specific integrated circuits (ASICs)*. They need to be compact and fully optimized to the parameters of the detector. CMOS technologies have gained the predominant position for front-end realization, due to the high integration density and low power consumption. They also give the possibility to combine both digital and analog circuits on the same chip, at a reduced cost.

2.3.1 Noise optimization

The most common amplifier configuration in semiconductor detectors front-ends is the charge sensitive amplifier for its superior noise performance. Sometimes current sensitive mode is used where speed is required and not minimum noise.

Generally the input transistor has the dominant contribution to the amplifier noise term e_n in equation 2.2.11. Considering a classical CMOS transistor model the white noise and $1/f$ noise can be written as

$$S_w = \frac{8}{3}\frac{kT}{g_m} \tag{2.3.1}$$

$$\frac{S_f}{f} = \frac{K_f}{C_{ox}^2 W L f} \tag{2.3.2}$$

$$e_n^2 = S_w + \frac{S_f}{f} \tag{2.3.3}$$

where g_m is the transconductance of the transistor, W and L are respectively the width and length of the MOS channel, C_{ox} is the oxide capacity per unit of area, K_f is the $1/f$ noise coefficient which is fixed by the technology, k is the Boltzmann constant and T the absolute temperature.

Equation 2.2.11 can be rewritten as

$$ENC_{tot}^2 = ENC_w^2 + ENC_f^2 + ENC_i^2 \tag{2.3.4}$$

$$ENC_w^2 = \frac{8}{3}\frac{kT}{g_m}\frac{C^2}{\tau_s}A_w \tag{2.3.5}$$

$$ENC_f^2 = \frac{K_f}{C_{ox}^2 WL}C^2 A_f \tag{2.3.6}$$

$$ENC_i^2 = 2eI_{leak}\tau_s A_i \tag{2.3.7}$$

The only factor depending of the power consumption is the transistor transconductance g_m which has the following form in the classical model in the strong inversion regime

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I_{DS}} \tag{2.3.8}$$

where μ is the carrier mobility (electron mobility for N-channel MOS and hole mobility for P-channel transistor) and I_{DS} is the current in the transistor. Clearly, increasing the current consumption, the white noise contribution decreases until the $1/f$ noise becomes dominant. In this situation the noise is minimized by the so called *optimum capacitance matching* condition. The total capacitance at the input of the amplifier is given by the sum of the detector capacitance C_d, the input capacitance of the first transistor C_{GS}, the feedback capacitor C_f and other parasitic capacitors C_s.

$$C = C_d + C_{GS} + C_f + C_s \tag{2.3.9}$$

Observing that the input capacitance of the transistor is approximately the product

of the specific oxide capacitance C_{ox} and the area of the transistor WL

$$C_{GS} \approx C_{ox} WL \tag{2.3.10}$$

the capacitance matching condition reads

$$C_{GS} = C_d + C_f + C_s \tag{2.3.11}$$

On the other hand, in the case when thermal noise dominates, one can rewrite equation 2.3.8 as

$$g_m = \sqrt{2\mu \frac{C_{GS}}{L^2}} \tag{2.3.12}$$

The thermal equivalent noise charge becomes

$$ENC_w^2 = \frac{8}{3} \frac{kTA_w}{\tau_s} \sqrt{\frac{L^2}{2\mu}} \frac{(C_d + C_f + C_s + C_{GS})^2}{\sqrt{C_{GS}}} \tag{2.3.13}$$

which is minimized when

$$C_{GS} = \frac{C_d + C_f + C_s}{3} \tag{2.3.14}$$

The matching condition when neither of the components are dominant stands in between the matching conditions determined in equations 2.3.11 and 2.3.14. Determination of such condition is generally impossible using an analytical approach, therefore a numerical optimization algorithm must be employed.

One interesting situation is to determine the optimum conditions and the minimum noise level for a given power budget. The analysis in [48] shows that noise decreases roughly as $P^{-0.4}$, until the limit where $1/f$ noise dominates. Above this limit, power has little effect on the noise performance.

In modern technologies, the classical MOS transistor model is not accurate anymore. For highly segmented detectors, the power allocated to the readout electronics

is limited forcing the transistors to work in moderate and weak inversion regime. The level of inversion affect the expressions of the thermal noise, transconductance and input capacitance of the MOS transistor. Moreover, the K_f coefficient of the $1/f$ noise is enhanced for short channel transistors, while the slope of the $1/f$ noise differs for P-channel and N-channel, generally showing values lower than unity for the NMOS transistor[49]. As a consequence, capacitive matching does not apply any more as a general rule of thumb, but rather both drawn dimensions of the MOS channel have to be optimized individually.

2.3.2 Mixed-signal front-ends

One important aspect of ASIC front-ends in CMOS technology is the possibility to deploy both analog and digital circuits on the same readout chip. Traditionally, the term *front-end* refers to the low-noise amplifier and shaper combination, which were also the first building blocks to be integrated on readout ASIC chips. According to [50], the first front-end integrated circuit used in a high-energy physics experiment was the Microplex circuit for the Mark-II silicon vertex detector at SLAC in the early 1980s. The chip contained 128 channels, each consisting of an amplifier, sample-and-hold and analog multiplexing circuitry.

The notable increase of the total number of electronic channels in modern silicon vertexing systems like those of the LHC-based experiments, have brought together low-noise amplifiers, analog-to-digital converters and high-speed data serializers, as well as slow-control, test and calibration blocks, on the same readout chip, mounted in close proximity to the detector. This approach has the important advantage that the information is sent digitally to the upper-level data-acquisition system through

robust and high-speed digital channels, which are much less sensitive to undesirable pick-up noise and crosstalk than analog channels.

On the other hand, through embedded slow-control circuits, the performance of the readout chip can be controlled on-line and eventual nonuniform degradation in performance due to radiation or aging can be compensated.

One example of mixed-signal front-end integrated circuit which will be detailed later in this work is the PASCAL chip employed in the front-end of the silicon drift detectors of ALICE experiment. PASCAL contains 64 channels with low noise amplifiers and analog memories arrays and 32 successive-approximation ADCs, as well as calibration and control circuitry.

As transistor densities increase with lower feature size CMOS technologies, the amount of signal processing and data reduction that can be implemented together with the primary analog amplification and signal shaping blocks, increases considerably. This has a positive effect on the robustness and cost of large detection systems. On the other hand, putting together analog and digital circuitry on the same silicon bulk brings additional problems especially due to the transport through the silicon bulk of spurious switching noise from digital circuits into the sensitive analog building blocks.

2.3.3 Radiation tolerance

When exposed to ionizing radiation, CMOS technologies are affected by charge buildup and defect creation in the silicon oxide layers. These effect lead to MOS threshold voltage shifts, mobility degradation and noise increase. Charge buildup may also open leakage paths either inside the MOS channel or between adjacent n+ diffusions.

Such problems are predominant at the edges of the MOS transistor, where an inversion layer can be induced by charges trapped in the field oxide. Eventually an uncontrolled conductive path may open.

This undesirable effects can be mitigated by using *edgeless transistor geometries*. With such geometries, the gate of the MOS transistor has a *donut-like* shape, enclosing the drain diffusion. This approach eliminates radiation induced leakage currents between the source and drain of the MOS transistor. Additional guardrings surrounding the transistor mitigate the formation of other leakage paths.

Radiation tolerance for doses up to 30Mrad for the 0.25 μm process has been demonstrated with this approach [51]. The principal drawbacks are area penalty and limitations in the transistor form factor.

2.3.4 Scaling effects on the design and performance of front-end electronics

Dennard et al. [52] proposed in 1973 a fabrication methodology of very small MOSFET switching devices suitable for digital integrated circuits, based on ion-implantation. They recognized the fact that MOSFETs can be easily and predictably scaled to lower dimensions and worked out scaling relationships, showing how performance improves with smaller transistors. Dennard's scaling theory and the constant technological progress in microelectronics industry over the last decades have shrinked the minimum feature size of CMOS circuits from $5\mu m$ in the late 1970's to below $90nm$ today. This evolution sustained the evident progress of computer and consumer electronics industry over the last four decades.

Parameter	Expression	Scaling rule
Dimensions	W, L	$1/\alpha$
Voltages	V_{th}, V_{DS}	$1/\alpha$
Electric field	V/L	const
Conductance	G	const
Capacitance	C	$1/\alpha$
Speed	I/CV	α
Switching energy	CV^2	$1/\alpha^3$
Power/gate	CV^2f	$1/\alpha^2$
Transistor density	$1/L^2$	α^2
Power density	V^2G/L^2	const

Table 2.2: Constant-field scaling rules

2.3.4.1 Constant-field scaling

In MOSFETs the current density is proportional to the electric field. Thus, scaling all dimensions and voltages by the same factor, the electric field and the current density remain the same, preserving the DC characteristics. This implies that, to some extent, one design can be easily ported to a new technology by simply shrinking it with the scaling factor α.

As a consequence of scaling, the transistor density and the speed of the digital gates increase with α^2 and α, respectively, while power density stays constant (see Table 2.2). This simple scaling relationship lead the exponential growth in complexity and performance of digital and consumer electronics in the last 30 years, and made CMOS the dominant digital technology.

2.3.4.2 Noise and dynamic range

Let us rewrite the expression of the equivalent noise charge due to series noise in the case of charge-sensitive amplifiers:

$$\begin{aligned} ENC^2 &= ENC_w^2 + ENC_f^2 = \\ &= \frac{8kT}{3}\frac{C^2}{g_m}\frac{A_w}{\tau_s} + \frac{K_f}{C_{ox}^2 WL}C^2 A_f \\ &= \frac{8kT}{3}\frac{C^2}{C_g}\frac{C_g}{g_m}\frac{A_w}{\tau_s} + \frac{K_f}{C_{ox}}\frac{C^2}{C_g}A_f \\ &= \frac{C^2}{C_g}(\frac{8kT}{3}\frac{A_w}{\tau_s}\frac{1}{2\pi f_T} + \frac{K_f}{C_{ox}}A_f) \end{aligned} \tag{2.3.15}$$

where $C_g = C_{ox}WL$ is the input capacitance of the transistor and $f_T = g_m/(2\pi C_g)$ is the intrinsic cutoff frequency.

The latter is process dependent and can be expressed as [53]

$$f_T = \frac{\mu U_T}{2\pi L^2} 2(\sqrt{1+IC} - 1) \tag{2.3.16}$$

where U_T is the thermal voltage and IC is the inversion level. The value of the cutoff frequency scales up while smaller channel length L are accessible. The inversion coefficient is defined as the normalized forward current of the MOSFET:

$$IC = \frac{I_F}{2n\mu C_{ox}(W/L)U_T^2} \tag{2.3.17}$$

where I_F is the absolute drain current, n is the subthreshold slope factor, μ is the mobility and W and L are the effective channel width and length, respectively. A device with an inversion coefficient lower than 0.1 is operated in weak inversion, while for values greater than 10 it is considered to be in strong inversion. For values in between, the operation region is defined as moderate inversion.

Expression 2.3.15 shows that, with the same capacitor matching condition and inversion level and shaping filter, the noise performance improves with scaling.

The power density of the $1/f$ noise can be expressed in terms of process parameters as [54]

$$S_{1/f} = \frac{K_f}{C_{ox}^2 WLf} = \frac{N_t e^2 t_{ox}}{\alpha C_g f} \tag{2.3.18}$$

where N_t is the volumetric trap density in the oxide, e the electron charge, α a process dependent constant, and t_{ox} is the gate-oxide thickness. Therefore the noise parameter $H_f = K_f/C_{ox}$ in the expression of $1/f$ noise scales as t_{ox} which in turn decreases with technology scaling.

On the other hand, the signal at the output of the charge-sensitive amplifier has to stay within V_{DD} which, in turn, decreases for each new technology generation. The dynamic range of the amplifier is roughly given by V_{DD}/ENC ratio. Despite the improvement of noise performance with scaling, the later degrades by as high as 15% per generation[55].

The situation can be solved by encoding the signal into current rather than voltage, with a power consumption penalty, since, generally current-mode circuits tend to dissipate more power for the same bandwidth and linearity as compared to voltage-mode. Another approach is to use time as the amplitude information carrier using linear time-over-threshold circuits. In this case, the signal at the output of the amplifier can as well saturate at V_{DD} as long as the width of the output pulse maintains proportionality with the input charge.

2.3.4.3 Transistor mismatch

Transistor mismatch plays an important role when multiple-channel readout circuits are considered. Nonuniformities may lead to yield or performance loss. The MOSFET threshold voltage mismatch which is dominant in deep sub-micron processes is roughly inversely proportional to the square root of the area of the transistor. The process-dependent proportionality factor A_{VT} generally scales more slowly than the technology scale factor [55].

On the other hand, the power rail scaling has the same negative effect on relative matching $\sigma_{V_{th}}/V_{DD}$ as in the case of dynamic range. In order to preserve this figure, there is no other alternative than the use of larger area devices. Apart from the space penalty, large area devices show higher capacity, with negative impact on the speed/power tradeoff.

One example where relative matching is of fundamental importance is the analog-to-digital conversion. The general trend in this domain is to use a high level of redundancy doubled by digital correction, since digital circuitry comes at a lower price in deep-submicron technologies than analog accuracy does.

2.3.4.4 Radiation tolerance

Modern technologies with lower feature size show better radiation tolerance. This phenomena is linked to the reduced gate oxide thickness in scaled processes. Radiation induced edge leakage in NMOS transistors and loss of isolation between n-doped regions after radiation may still need to be addressed through enclosed geometry design and guard rings in high dose applications. Evaluation of intrinsic radiation

hardness of processes with feature size lower than 0.13 μm is under on-going investigations and lacks of firm conclusions. Nevertheless, first results show better radiation tolerance than 0.25 μm or higher technologies.

Single event upsets (SEU) are of more concern in modern processes, being related to the reduced device size and smaller critical charge. They are successfully addressed through redundancy and error-sensitive or error-correcting logic.

Chapter 3

System integration of the Silicon Drift Detector (SDD) in ALICE

This chapter focuses on the system-level integration of the front-end electronics of the ALICE Silicon Drift Detectors. The layout of the apparatus is described together with an overview of the architecture. The custom test system developed to monitor the quality of the detector during the production flow is detailed.

3.1 SDD detector

Silicon drift detectors (SDDs) were chosen for the two intermediate layers of the Inner Tracking System (ITS) of ALICE experiment because they feature very good multi-hit capability and good energy loss resolution, with a small number of readout channels [31].

As previously discussed, SDDs consist of a fully depleted silicon volume inside which an electric field is created by p+ implanted electrodes connected to a resistive

ladder (see figure 3.1). The electrons originating from ionization processes due to a charged particle crossing the detector volume, drift longitudinally inside the silicon bulk at constant velocity and are collected by an array of n+ anodes disposed perpendicularly to the direction of the drift field at the edge of the sensor.

The coordinate of the hit along the anode line is calculated from the centroid position of the collected charge, while the other coordinate information is inferred by the drift time of the charge. Additionally, the integrated charge corresponding to one hit is directly proportional to the energy lost by the particle passing through the depletion volume of the detector. This fact is exploited to do particle identification via specific energy loss measurements. Low noise level and good energy resolution can be achieved because of the small size of the anodes and their corresponding small capacitance ($\sim 50fF$, see previous chapter).

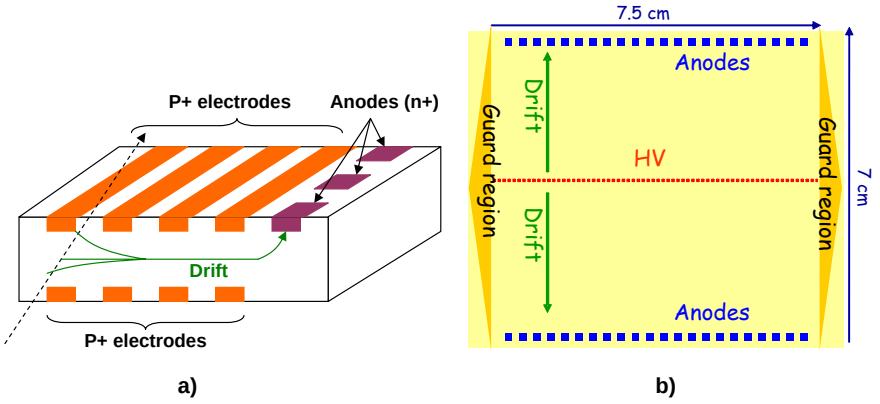


Figure 3.1: a) Principle of a silicon drift sensor; b) The ALICE silicon drift sensor (top view).

The sensor has been built in neutron transmutation-doped (NTD) technology, which provides low doping fluctuations and, thus, good carrier velocity uniformity

[56]. The active area is 70.2 x 75.8 mm^2 and the active-to-total area ratio is 88%. The high-voltage polarization net divides the sensor in two distinct depletion regions, in order to contain the maximum drift time below 5 us with reasonable high-voltage ($\sim 2kV$). Consequently, the signal is readout at two opposite sides of the detector via two 294 um pitch arrays of 256 anodes. The sensor integrates on-board the resistive ladder needed to partition the high voltage distributed to the p+ electrodes and a number of charge injectors used to accurately measure the drift velocity during normal operation (see figure 3.2).

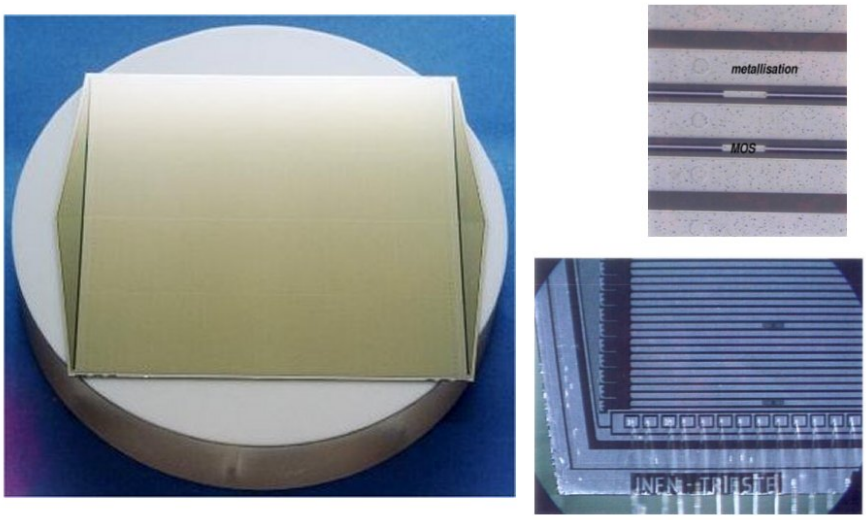


Figure 3.2: Photo of the ALICE SDD detector (left). Zoom of the collection area (right-bottom). Charge injectors (right-top).

The sensor was tested using high momentum pion beam provided by the SPS accelerator at CERN. The average resolution obtained is 33 μm in the drift direction and 27 μm along the anode coordinate, respectively.

3.2 Front-end electronics architecture

Silicon drift sensors require more complex readout and data analysis schemes with respect to other silicon sensors. The front-end electronics must include certain memory volume, analog readout with relatively fast analog-to-digital conversion and, possibly, integrated signal processing. The front-end electronics has to scope the signal within a time window equal or slightly longer than the maximum drift time, in order to retrieve the full information associated to an event.

Previous realization of a similar large sensor system [57] adopted an entirely analog front-end solution where the signal is sampled and stored in an analog memory while the A/D conversion is carried out by dedicated electronics mounted outside the tracker volume.

For the ALICE SDD system the FE tasks were divided between two chips. The first is a 64-channel mixed-signal chip (PASCAL) embedding the analog-to-digital conversion circuitry, while the second is a fully-digital chip (AMBRA) for event derandomization and data reduction. Both chips were designed using rad-hard techniques in 0.25 μm CMOS technology [58].

3.2.1 PASCAL

The architecture of the analog front-end chip is presented in figure 3.3. The current flowing out of the detector is amplified, sampled at 40 Ms/s and fed to a 256-cell circular switched-capacitor analog buffer. When the trigger fires the sampling stops and the content of the analog memory is converted into digital code by 10-bit SAR ADCs. Providing analog-to-digital conversion on-chip avoids the accuracy and noise problems linked to the multiplexing and transport of analog signals on relatively long

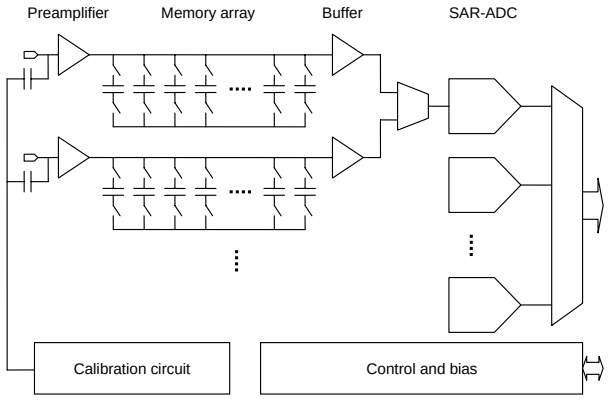


Figure 3.3: Front-end chip architecture (PASCAL)

distances to the DAQ system. The sampling clock as well as the conversion clock is supplied by the LHC bunch crossing clock which operates at 40 MHz. Optionally, the frequencies of the two clock signals can be internally divided by two. A DAC-controlled test pulse generator is integrated for calibration and operation monitoring.

A number of 32 ADCs working at 2 MSPS were integrated on the chip, one ADC being shared by two channels. The power consumption of a single SAR ADC block is as low as 2 mW. The choice of SAR architecture for the A/D conversion proves to be a successful compromise between power consumption, silicon area and performance.

Addressable JTAG protocol is implemented in both front-end chips for control, calibration and test purposes. Using the JTAG link, the user can read and write channel mask registers, preamplifier gain control registers or the test pulse DAC register.

3.2.2 AMBRA

The second chip, called AMBRA[1], is a digital bridge between PASCAL and the data-acquisition system. It performs non-linear 10 to 8 bit conversion and baseline subtraction. It contains 64kb RAM organized as a digital 4-event de-randomization buffer, storing up to 4 events for the case when the data acquisition system is busy processing other events . AMBRA also acts as digital control circuit for PASCAL.

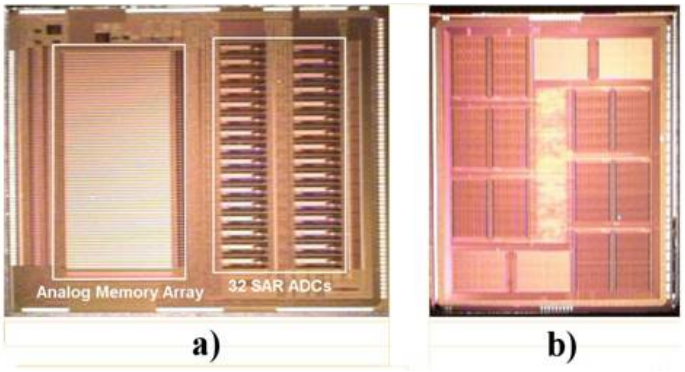


Figure 3.4: Microphotograph of the two front-end chips.

Given the high sensitivity of the preamplifier integrated in PASCAL an important issue of the design was to develop a digital interconnection scheme that minimizes the coupling of spurious digital noise in the analog section with minimum implication on the power budget. LVDS differential digital signaling was preferred for off-chip connections as it provides, in addition, a very good ground separation. LVDS standard requires that a minimum current of 3mA flows through the termination resistor. Considering the number of I/Os between the two chips (approx. 30), this means that the power dissipated by the inter-chip bus is more than 200 mW (3.5 mW/channel), which represent an unnecessary load of the power budget. Since the distance between

[1]"A Multi-Buffer Readout ASIC"

the two chips is very short (2-3 mm) the impedance of the digital lines was increased to $2k\Omega$, decreasing the current to 300 μA and the power consumption by a factor of 10, without losses in the quality of the inter-chip communication.

The JTAG slow-control link is available also for AMBRA, giving access to internal registers for the baseline subtraction and other mask and control purposes. Boundary Scan protocol integrated in AMBRA is of particular importance in the production flow of the sensor system. The Boundary Scan industrial standard provides the possibility to disconnect all digital pins from the internal logic core and control and read their state through the JTAG serial link.

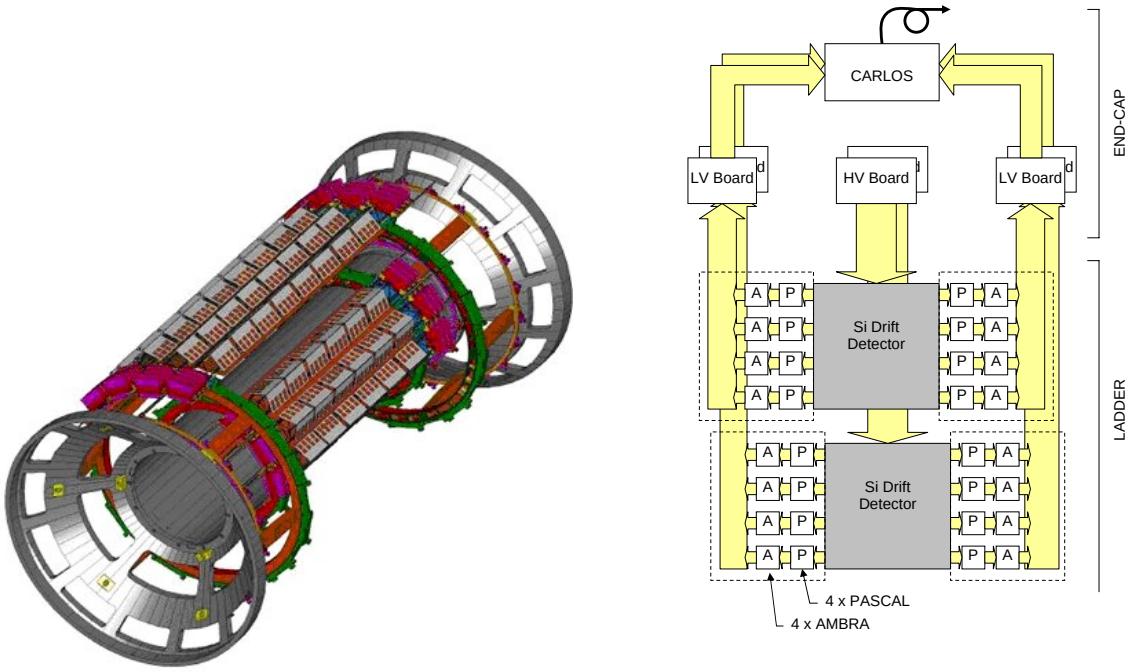
3.3 System integration

3.3.1 Layout of the SDD system

The Silicon Drift Detector subsystem of the Inner Tracking System (ITS) of the ALICE detector consists of 260 detector modules geometrically disposed in 2 cylindrical layers of 15 and 24 cm radius, for a total detection area of 1.31 m^2 and 130,000 electronic channels [31].

The support of the detectors and the front-end electronics is entirely realized with non-standard carbon-fibre material in order to meet the tight material budget. A light space frame called *ladder* hosts 6 or 8 detectors for the inner and outer SDD layer, respectively.

On each side of the detector a complex upilex/Al hybrid structure mounted on a carbon-fibre heat dissipater holds 4 PASCAL+AMBRA doublets and the few small-size SMD components they need (see figure 3.5 b). Flexible microcables of the same



a) b)

Figure 3.5: a) SDD layout overview. b) SDD components.

material as the hybrid structure hosts the data and control bus and the supply connections to the end-cap part of the FE chain. An additional glue thickness between the Aluminum low-pitch traces and the upilex support realizes the controlled impedance needed by the LVDS standard. At the end-cap the information is compressed and send via optical links to the global DAQ by a dedicated ASIC called CARLOS which performs zero-suppression and data compressing.

3.3.2 Production flow

The integration of such a complex and technologically innovative system involves 4 research institutes and also industry participation. A short description of the production flow follows, while a more detailed one can be found in [59].

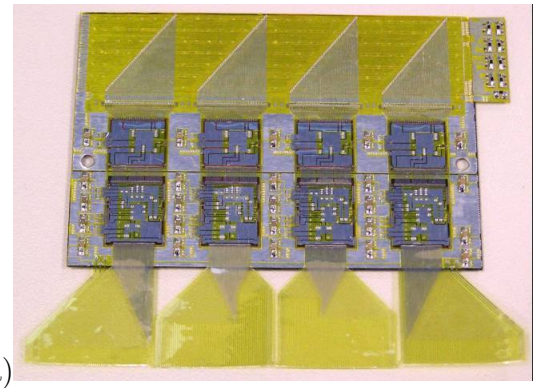
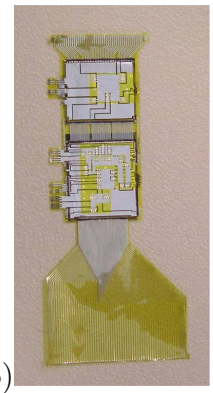
a)  b) 

Figure 3.6: Hybrid (a) and chipcable (b) photographs.

Several steps lead to the realization of the two SDD layers of the ITS. The two chips are first TAB-bonded to a flexible multilayer Al on upilex *chip-cable*. The chip-cable hosts in a continuous structure all the low-pitch connections between the two chips, the power supply connections, as well as the pitch adapters between the detector anodes and the analog inputs of PASCAL, and the pitch adapters between AMBRA and the acquisition bus. The chip-cables come in 4 different shapes corresponding to the 4 positions of the FE chip-pairs on the side of the detector.

The four chip-cables are TAB-bonded on a *sub-hybrid* structure hosting all the small size SMD components. The last is beforehand glued on top of a 300 μm carbon fiber heat-dissipater. The FE chips are sandwiched in this way between the chip-cable and sub-hybrid, for an efficient space use. The structure obtained (called *hybrid*) contains all FE electronics needed to serve half a detector. One left-handed and one

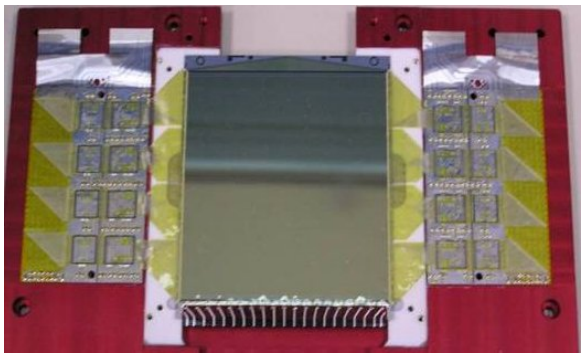


Figure 3.7: Photograph of a semi-module

right-handed hybrid are bonded to the anode pads of the detector on both sides. Microcable for data and power transport are bonded onto the hybrids, as well as the high-voltage cable to the back side of the detector, together with corresponding end-cap low-voltage and high-voltage boards. At this stage of the production flow, the detectors and its dedicated FE electronics (*module*) can be fully characterized under full depletion conditions with IR light or radioactive sources.

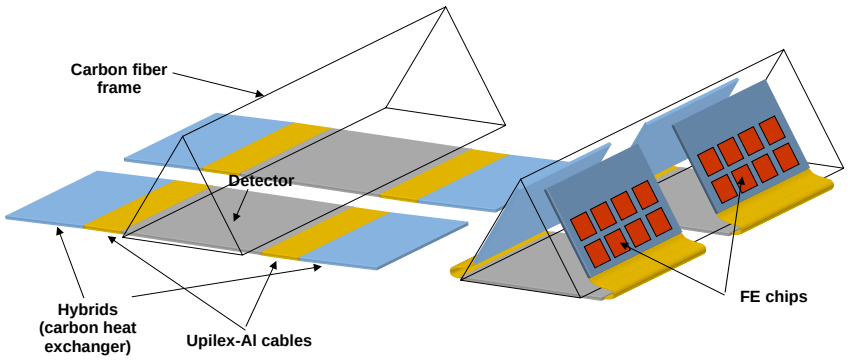


Figure 3.8: Mounting procedure of the modules on the support frame

After assembly, test and mapping of modules, the detectors have to be aligned and

fixed onto their carbon fiber support frame (ladder). This operation is extremely delicate, involving several high precision operations and dedicated high-precision tools. The detectors, slightly overlapping one-another to avoid gaps in the sensitive area, are fixed on to one side of the support frame and aligned with a 10 μm precision. The hybrids, microcables and end-cap boards are wrapped onto the other two sides of the triangle-shape frame (see figure 3.8. The fully equipped ladders are assembled on conical shaped end-cap support structures. The end-cap electronics contains also the optical links and the chips for the Detector Control System (DCS).

3.4 Quality assurance

3.4.1 Overview

Throughout the whole production flow the bonding is severely verified after almost every technological step. Figure 3.9 shows the assembly and test stages up to completion of the modules. All modules of the SDD barrels must be guaranteed to be fully functional with less than 1% dead or defective channels. They have to stand 10 years of operation under experimental conditions with minimum failure rate and performance deterioration.

Functional parameters are measured during tests and the components of the FE chain are classified according to their performance immediately after they are assembled. The pieces that do not meet the performance requirements are discarded. In some cases repair work can be applied. Failed connections on the chip-cable can be manually re-bonded before the gluing process or entire chip-cable can be replaced in a hybrid. Re-bonds between the detector and the FE hybrids are also attempted.

The whole quality assurance scheme is integral part of the production flow. It is designed to limit the failure rate at each production step for an optimized production flow. The components that are to be mounted together must stand the stress of the assembly as well as long-term operation under extreme conditions.

A dedicated test system has been developed and realized at INFN Torino for the test of chip-cables, hybrids and modules, together with the related software. The key features of this system are versatility and user-friendliness. Different tests and functional characterizations are being carried out using this tool, both in-house and in industry commissioned sector (chip-cable production). The system has the capability to detect automatically any failed connection with more than 99% efficiency and to extract functional parameters for a detailed classification of the device under test.

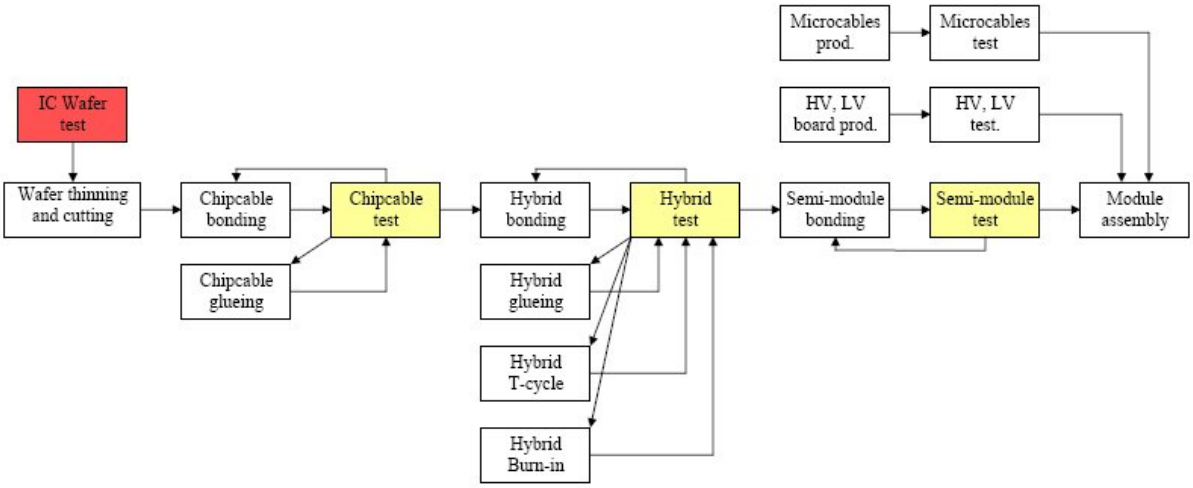


Figure 3.9: Assembly and test flowchart of the SDD module. The boxes filled with yellow color are the test stages performed using the test system described in this chapter.

a) FE Chips

A dedicated wafer probe station at INFN Rome carries out the test of the chips. The test algorithms and the hardware involved were developed at INFN Torino. The entire code set of the ADCs is explored. Analog performance like noise, gain and gain linearity are checked on a channel-by-channel basis, together with detailed operation of the digital circuitry and a full RAM scan. The algorithms were optimized in a series of iterations to assure the completion of the task in a reasonable amount of time. Further details on the chip test can be found in [60].

b) Chip-cables

The chip-cables are foreseen with test extensions that permit to electrically connect the chip doublet to external test hardware via a dedicated test fixture. After all test and post-processing steps are successfully fulfilled the extensions are cut out before bonding the chip-cable to the sub-hybrid. The chip-cable is tested both before and after the TAB-bonds are hardened applying glue.

Some functional tests at this stage proved very useful in detecting the few cases of damaged chips during wafer thinning. Due to the high complexity and high density of traces on the chip-cable, the setup of the dedicated test system encountered practical problems. Although the signal paths were kept at minimum the analog input conditions for lowest noise were not met. Since it was impossible to bring the decoupling capacitors very close to the chips to limit the series resistance, the system is susceptible to oscillations. It was nevertheless possible to overcome these issues by improving the ground potential control and the connection efficiency of the test fixture. Careful algorithm development assured better than 99% fault detection efficiency.

c) Hybrids

Besides the connectivity verification, the hybrids also undergo a series of reliability tests. They must withstand 10 temperature cycles from 25 to 65 degrees and 24 hours continuous operation (burn-in). In this way the foreseen operation conditions are simulated. Any weak bond would be broken by the considerable stress induced in the temperature test cycle. A small sample of hybrids has successfully over passed 200 thermal cycles. The functional tests and characterization of the hybrids shows that the performance of the FE chain has been preserved and in some cases improved with respect to previously reported prototypes [].

d) Semi-modules

As a first step towards the module assembly, the FE hybrids are connected to the detector. The anodes region of the detector can be partially depleted via a grid potential and the resulting capacitor at the input of the preamplifier has systematic influence on noise and on the response to the internal test pulse of the FE chip. In this way the connections and the integrity of the anode region of the detector can be rapidly verified.

e) Module

After this verification the signal and power supply micro-cables are TAB-bonded to both hybrids and the HV cable is connected to the detector, as well as the Low-Voltage and High-Voltage boards. This is the first occasion to operate the detector with its FE chain under high-voltage and with the integrated injectors. After the full assembly of the module the verification tasks are taken over by another test system based on the ITS data-acquisition system. With this system, the detectors are mapped using IR laser light and a precise moving station.

f) Microcables, HV and LV boards

The flexible microcables, the chip-cable and sub-hybrid Al/upilex structures are manufactured at the Scientific Research Technological Institute of Instrument Making Microelectronic Department, Ukraine where they are electrically tested against shorts or open connections. The quality of the controlled-impedance microcables used for digital data transport is checked with 40MHz signals at INFN Torino. The low-voltage and high-voltage boards are assembled at INFN Torino on PCB and ceramic support. They undergo functional and burn-in tests before being integrated in the module.

g) Ladder

The modules go through a lot of mechanical stress while they are aligned and wrapped on to the carbon-fiber support frame. The quality assurance scheme up to this point ensures that the modules can stand this amount of stress with no damage. Nevertheless, a full verification of each module is carried out before the final encapsulation.

3.4.2 Chipcable, hybrid and module test system

A dedicated test system was developed to fulfil the quality assurance requirements of the FE system (HYCC). The test system is based on a commercial FPGA PCI-card manufactured by National Instruments. The convenience of this solution is that custom control pattern and state machines can be implemented inside the FPGA leaving out of the scene additional equipment. The card has a high number of digital I/Os and some analog inputs and outputs which were used to monitor voltages and currents, as well as the temperature of the device under test.

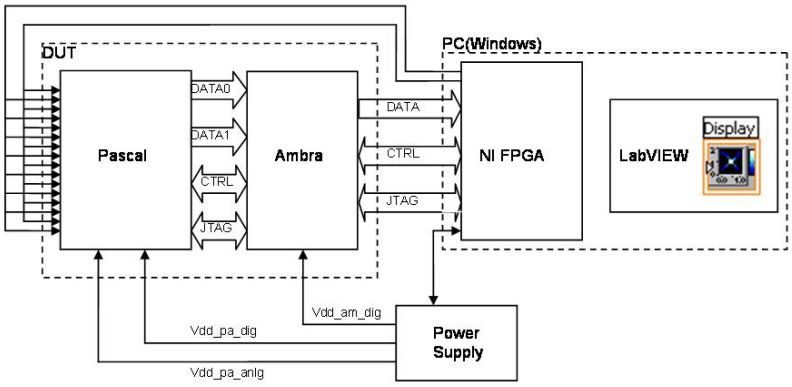


Figure 3.10: Overview of the test system

The programmable logic is configured to fully check the state machines of the FE chip-pair. In addition, a FE emulator that can be used when some of the signals 5 are missing due to, for instance, broken bonds. The operating clock of the FPGA core is set to 80 MHz and it delivers a 40 MHz clock to the device under test. The clock frequency can be scaled down in 4 steps. A JTAG protocol core is implemented in parallel with the control and data acquisition state machine. An interface board performs the CMOS to LVDS conversion and hosts the analog circuitry needed by the voltage and current measurements, as well as the supply regulators for the chip-cables and the hybrids.

The FPGA card also controls the power regulators mounted on the interface board, while the currents absorbed by the device under test are monitored using analog ports available on the FPGA card.

The system is capable to test, control and acquire data from chip-cables, hybrids and with minor modifications, modules. The programmed logic is in transparent software control using the LabVIEW 7.1 FPGA Module package. Data transfer speed

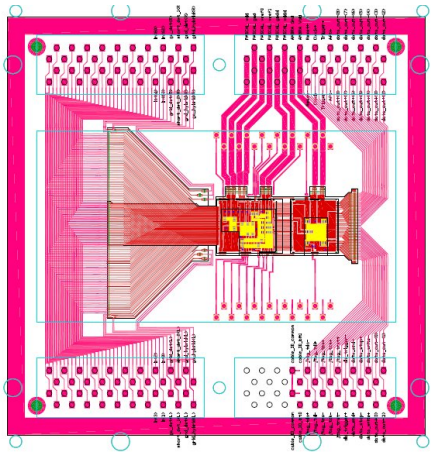
a)  b) 

Figure 3.11: a) Chipcable with test extensions. b) Chipcable test station. Test fixture in prim-plan.

is sufficient for the objective of the system, and the versatility of the software tool is a great advantage.

When a chip-cable is tested, an automated test sequence is performed, which starts with a power supply check and the identification of the chip-cable type, using the JTAG link. Next, the JTAG link on AMBRA is thoroughly tested by writing and reading back the internal registers. When the JTAG link is found to work, a boundary scan sequence is started. The I/O pins of AMBRA are disconnected from the internal digital core and a pattern is written to them through the JTAG, and readout statically using the I/O's of the FPGA-card.

When boundary scans completes, the JTAG link to PASCAL is verified. The functionality of the entire JTAG chain has the highest priority in the test since all other test sequences depend on it.

The digital connections between AMBRA and PASCAL are not accessible from outside and boundary scan circuitry, similar to the one integrated in AMBRA is

not available in the case of PASCAL. For this reason, these connections are verified indirectly, correlating the response of the system with failure patterns corresponding to specific defects. The differential nature of this connections, as well as their non-standard high-impedance makes it difficult to predict the behavior of a faulty connection. A certain degree of redundancy in this tests guaranties the fault detection accuracy needed.

As an example, the connectivity of data busses between the two chips is verified acquiring a high number of events with different settings for the baseline DAC in PASCAL, with or without the internal test pulse active. Bit-by-bit histograms are computed and compared with a set of thresholds in order to detect any faulty connection. The thresholds were calibrated using a set of good and intentionally damaged chip-cables.

When the system is used for the verification of hybrids, each PASCAL-AMBRA chip pair of the four (or each chip-cable) are selected and tested separately. Most of the steps described earlier are repeated for each chip-cable, in order to detect any possible damage that they might have suffered while bonded to the hybrid. The procedure is once again fully automatic.

At the chip-cable level, the inputs of PASCAL can be accessed through the test extensions. For this purpose, two pulse sources are embedded in the trigger and acquisition control state machine implemented in the FPGA card. In the case of hybrid verification, the inputs of PASCAL are open. A strong correlation can be observed between the single channel noise level and the quality of this bonds. The capacitance formed between the traces of the pitch adapters bonded to the inputs of PASCAL and the grounded gig used to manipulate the hybrid under test makes the

noise level to rise. If one of the bonds did not survive the hybrid assembly process, or the temperature-cycle or burn-in tests, the noise level of the corresponding channel stays low.

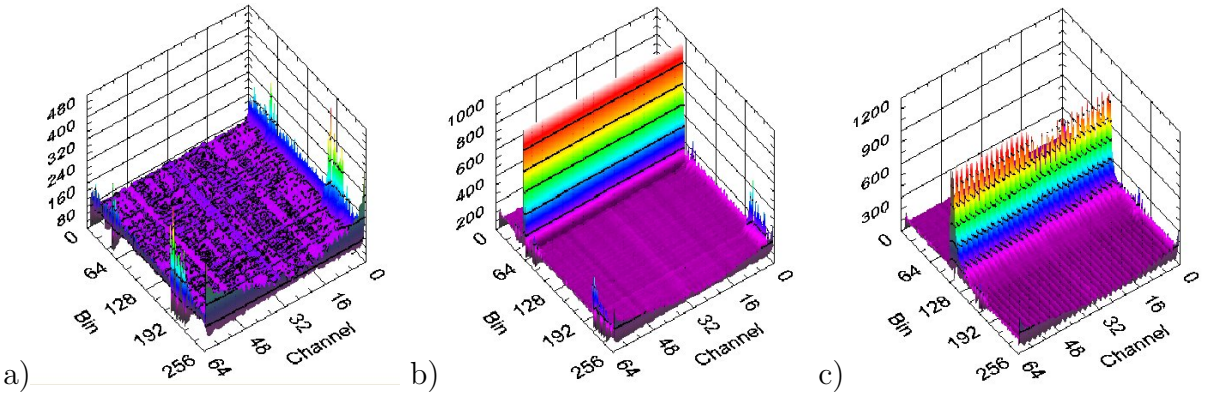
The connection between the detector and hybrid can be verified at semi-module level (i.e. before connecting the high-voltage cable to the detector). Applying the bias voltage to the grid surrounding the anodes, which is connected on the hybrid, changes the noise and the response to the internal test pulse of the channels that are effectively connected to the detector.

3.4.3 Yield and performance measurements

For the chip-cables the first automatic bond process has an yield of about 82% and little more than 50% of defective cables are recovered through subsequent manual re-bonding. In the case of hybrids, the boning is done manually and shows an yield of about 92%. Re-bond or chip-cable replacement solves about 60% of the defect hybrid cases after first assembly step.

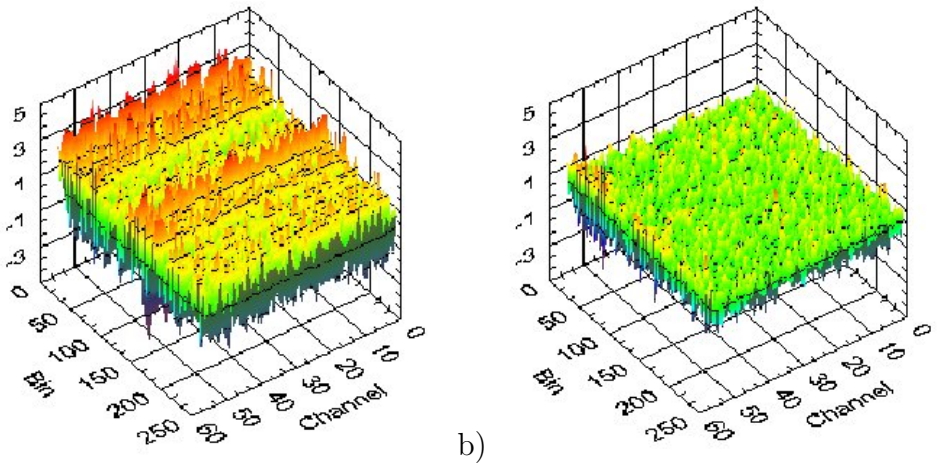
Upon the successful completion of the connectivity tests, a number of events are acquired with and without internal or external test-pulses, in order to characterize the device under test. The main parameters of interest for the selection and classification of the FE components are noise, gain and power consumption. Three examples of raw data events are showed in figure 3.12.

Before measuring the parameters, the raw data needs to be corrected against common-mode effects. One source of this disturbance can be the digital control circuitry integrated in PASCAL, which may induce spurious noise in the analog circuitry.



a) b) c)

Figure 3.12: Chip-cable raw data acquisitions. a) flatfield b) internal test-pulse c) external test-pulse



a) b)

Figure 3.13: a) Noise w/o common-mode correction. b) Noise w/ common-mode correction.

External disturbances like ground bounce or spurious signal pick-up may also contribute. The analog memories in PASCAL sample the output of the preamplifier in parallel and thus are affected by spurious noise sources in the same way. For this reason, a *common-mode* disturbance is present on all channels. Moreover, the dominant common mode component is synchronous with the acquisition trigger so it is similar

for every event. Averaging flat-field events can discriminate this component and the raw data can be corrected, revealing the intrinsic noise performance of device under test. Figure 3.13 shows the difference between a raw and common-mode corrected flat-field.

The noise performance of chip-cables measured during the production process is affected by the practical considerations discussed earlier. Average noise $2-3$ times higher than expected is observed (see figure 3.14).

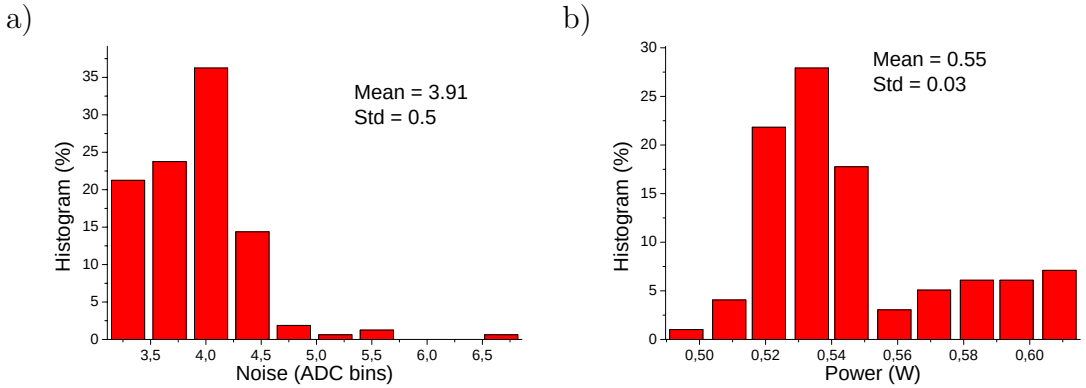


Figure 3.14: a) Noise and b) power consumption of chipcables.

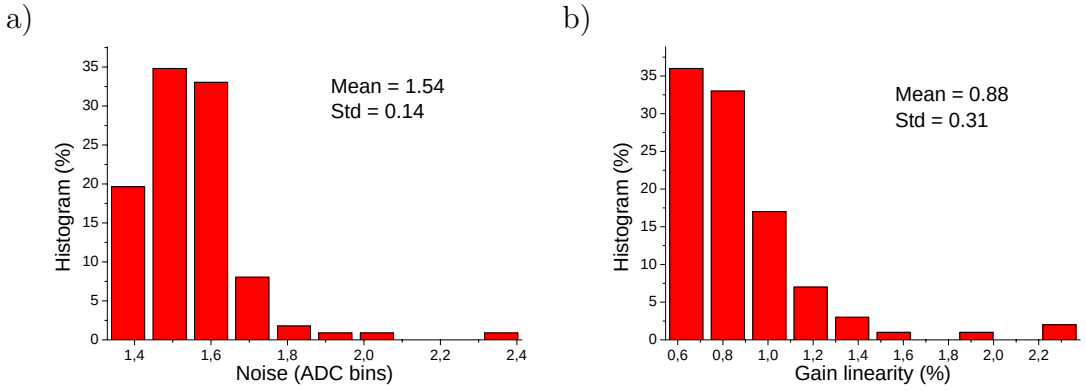


Figure 3.15: a) Open-input noise and b) channel gain variation of hybrids.

The characterization of the hybrids show that the performance of the FE chain has been preserved and in some cases improved with respect to previously PCB-based

prototypes [61]. The noise performance benefited from the lower stray inductance of the TAB-bonds with respect to traditional wire bonds.

The average noise measured in the case of hybrids when the inputs are open is better than 1.6 LSB, corresponding to less than 300 electrons (see figure 3.15a)). In less than 10% of the cases the average noise is higher than this value and most often due to singular noisy channels. Gain non-uniformity between the channels of the same hybrid is below 2% and below 1% in more than 50% of the cases.

Average power consumption is about 2W per hybrid which correspond to about 7.8 mW per channel. Figures 3.16, 3.17 and 3.18 show the measured distribution of static and dynamic power for the hybrid production until 2006.

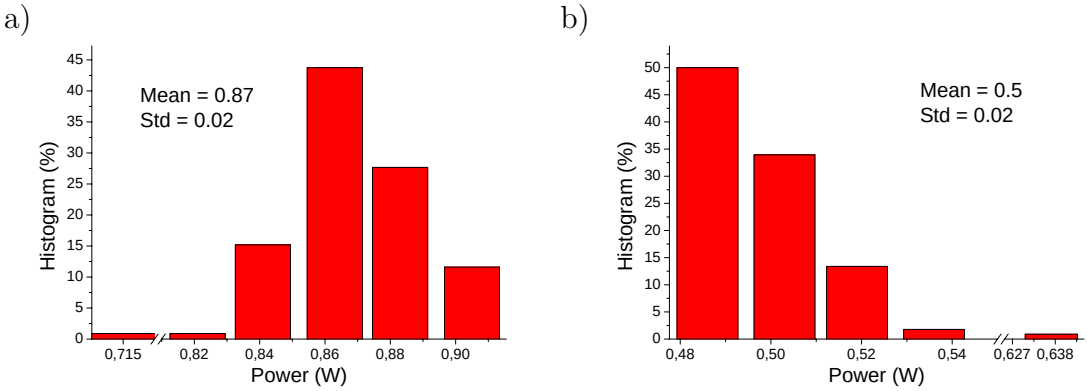


Figure 3.16: Static power consumption of hybrids. a) Analog b) Digital

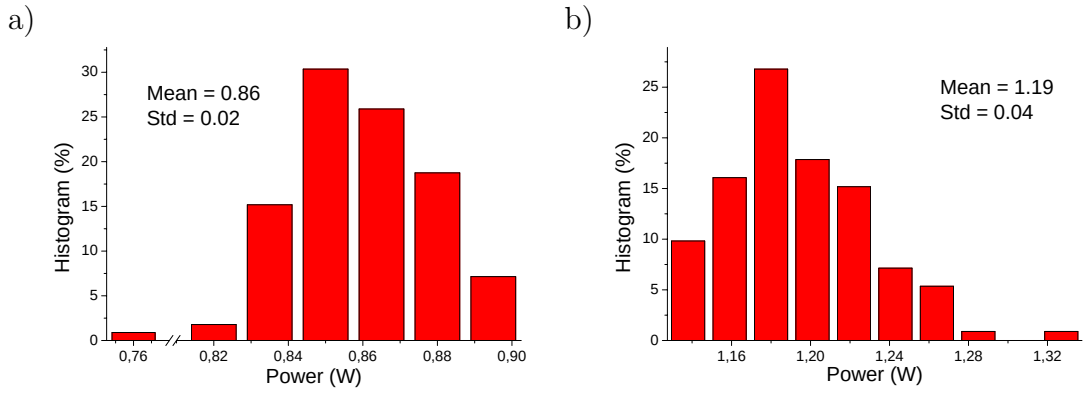


Figure 3.17: Dynamic power consumption of hybrids. a) Analog b) Digital

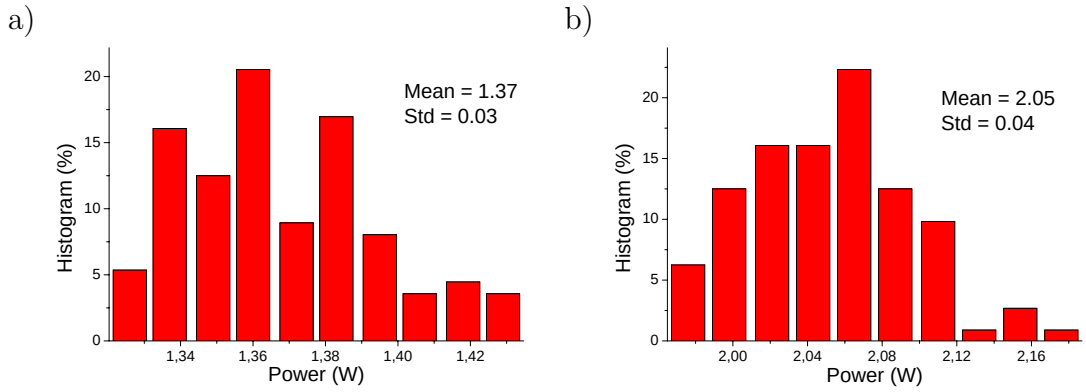


Figure 3.18: Total consumption of hybrids. a) static b) dynamic

Chapter 4

Design and test of front-end electronics for MVD pixel detectors in PANDA

This chapter describes a custom front-end solution for the MVD pixel detectors in PANDA, designed in a commercial 0.13 μm digital CMOS technology. The front-end cell was embedded in a multi-channel prototype chip and tested both before and after irradiation. An analysis of the time resolution capability is given.

4.1 Requirements

We remind in the following the requirements for the pixel detector of the PANDA experiment. A more detailed description of the experiment can be found in chapter 1 of this work:

- pixel size in the order of $100 \times 100 \mu m^2$;
- analog readout with high dynamic range;
- continuous (trigger-less) readout.

The first requirement implies a power budget of few tens of micro-watt per pixel if standard cooling techniques are to be used. The second requirement translates in the need of a higher resolution analog-to-digital conversion at front-end level. Time-over-threshold technique may satisfy both low-power and high analog resolution requirements. The method measures the width of the signal rather than the amplitude itself. A simple digital counter counts the number of clocks the signal stays above some threshold level[1], which gives an indirect measure of the amplitude of the signal. Most of the times, the width of the signal does not depend linearly on the amplitude and a careful calibration must be performed. On the other hand, the resolution is limited by the shaping time and the frequency of the clock signal, thus, for good resolution longer signals are used. This imposes a compromise between analog resolution and rate capability of the front-end.

4.2 Front-end architecture

4.2.1 Preamplifier

The analog front-end cell described here consists of a low-noise preamplifier with capacitive feedback and a discriminator. The feedback loop of the preamplifier is completed with a constant current discharge circuit and a detector leakage current compensation loop (see figure 4.1). The constant current discharge provides a constant slope trailing edge of the preamplifier output signal which, in turn, provides a linear time-over-threshold (TOT) measurement of the detector charge signal. The discharge current source (I_{FB} in figure 4.1) is normally turned off. When the output of

[1] thus the name of the method

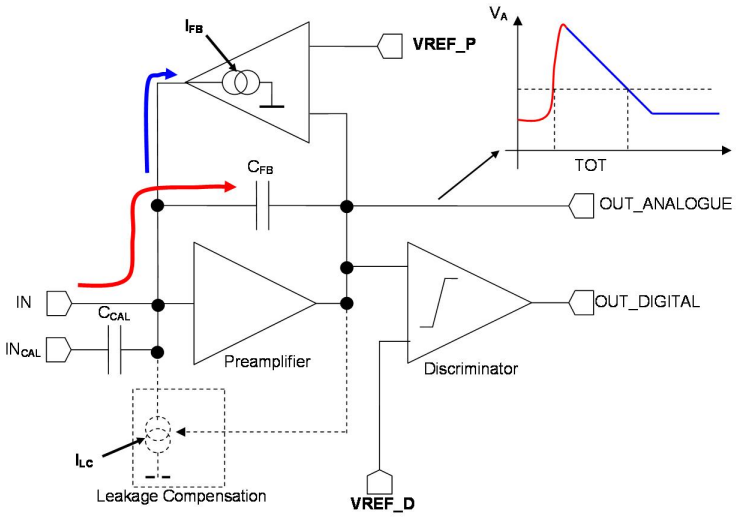


Figure 4.1: Block diagram of the front-end analog pixel cell.

the preamplifier exceeds the threshold voltage (VREF_P) the current source is turned on and the feedback capacitor is discharged with constant rate until equilibrium is restored, and the output level of the amplifier turns back to VREF_P.

The dynamic range is extended by discharging the feedback capacitor at the input node, preserving the linearity of the TOT measure even for higher input signals, when the preamplifier output is saturated. This is because the feedback current control is essentially an on/off switch and does not depend on the absolute amplitude of the signal. The feedback topology considered is similar to the one employed in the ATLAS pixel readout chip[62] or previously in [63].

One side-effect of this topology is the fact that the open-loop gain of the preamplifier drops rapidly when its output node saturates. One consequence of this is the increase of the input impedance of the preamplifier, which is related to the open-loop

gain by an inverse proportionality. The input node of the preamplifier is not anymore a virtual ground and, thus, becomes more sensitive to spurious perturbations or cross-talk.

On the other hand, for high input signals, the voltage at the input node may rise considerably, forcing the feedback circuitry outside the normal operation regime. This effect sets a limit on the dynamic range.

Another drawback of the time-over-threshold technique is the long dead time. This, however, dead-time of few microseconds is acceptable in the case of the MVD of PANDA.

4.2.2 Discriminator

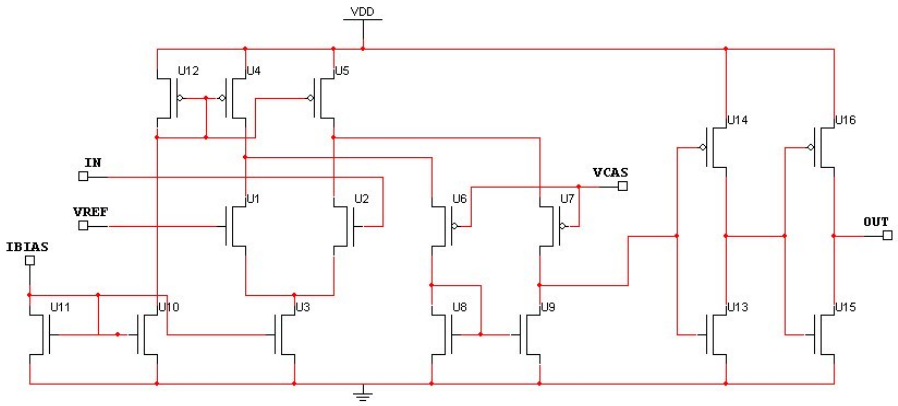


Figure 4.2: Discriminator schematic.

The discriminator design follows a low-voltage approach employing a single stage folded-cascode topology followed by two inverters (figure 4.2). The AC currents passing through the input differential copy U1 and U2 are reflected by the PMOS constant

current sources U4 and U5 back to the ground potential through the cascode transistors U6 and U7. NMOS transistors U8 and U9 represent the current-mirror load of the differential pair. The cascodes separates the differential pair from its load, providing better gain. Another advantage is that active transistors are all NMOS transistors which can be designed smaller for the same gain, thus minimizing parasitic capacitances and maximizing speed.

The design provides a DC gain of 160 and gain-bandwidth product of more than 1.4 GHz. Propagation delay is below 10ns and input offset is in the order of 3mV (sigma).

4.2.3 Layout of the analog pixel cell

The layout of the front-end components has been performed using conservative design rules with linear transistor geometry and metallic vertical plate capacitors[2]. Except for the two inverter stages of the discriminator, the minimum gate length of the transistor used was twice the minimum of the technology (120 nm). Particular care has been paid to matching techniques like common centroid design and symmetrical design, especially in the case of the discriminator layout (see figure 4.3 a)), in order to minimize the effect of geometrical process variations. The preamplifier and discriminator occupies $37 \times 51 \mu m^2$ and $12.8 \times 48 \mu m^2$, respectively.

[2]Vertical plate capacitors use the gap between two or more metallic fingers on the same metal level. They take advantage of the high thickness of the metallic layers and the high number of metal layers provided by modern technologies, offering higher capacity density than the traditional parallel plate capacitors.

a)

b)

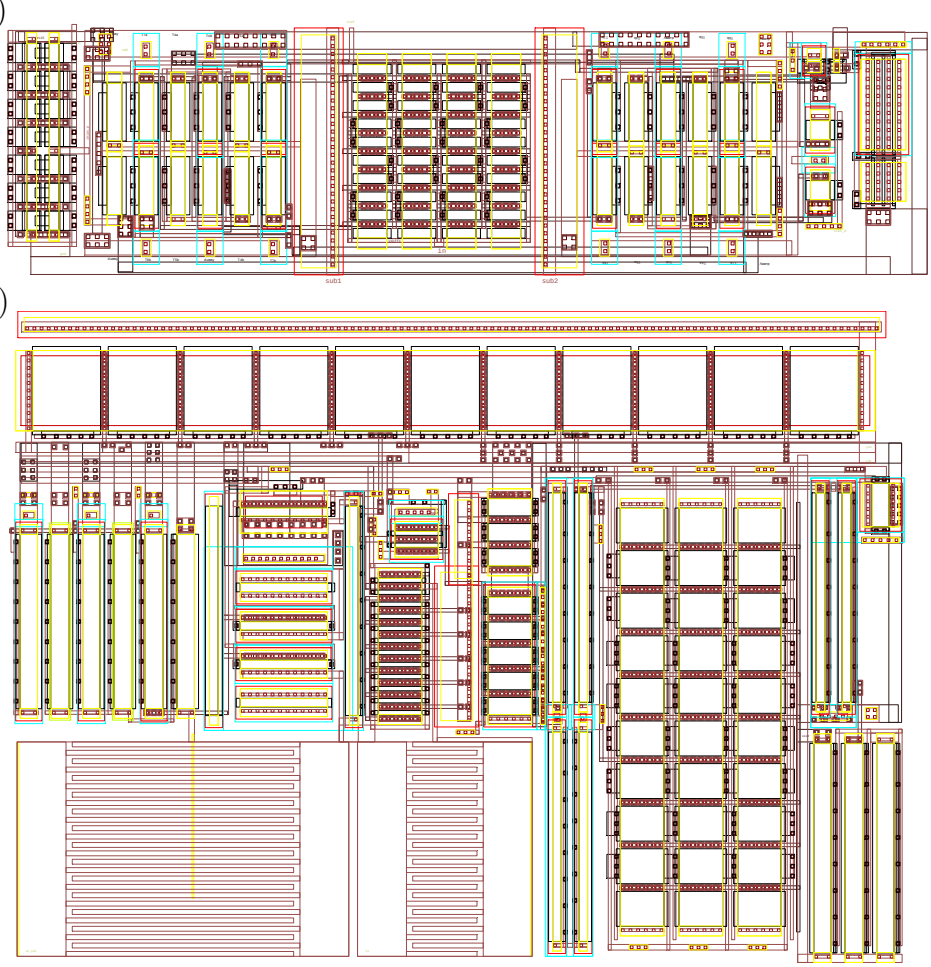


Figure 4.3: Layout of the discriminator (a) and preamplifier (b).

4.3 First prototype chip

4.3.1 Chip architecture

The prototype chip contains 32 front-end cells that share four calibration lines and 8 multiplexed output buses for both the preamplifier and discriminator output (figure 4.4) and a 33rd cell separated from the others. The charge is pulsed in at the input of the preamplifiers through a capacitor of 30fF connected in series with one of the

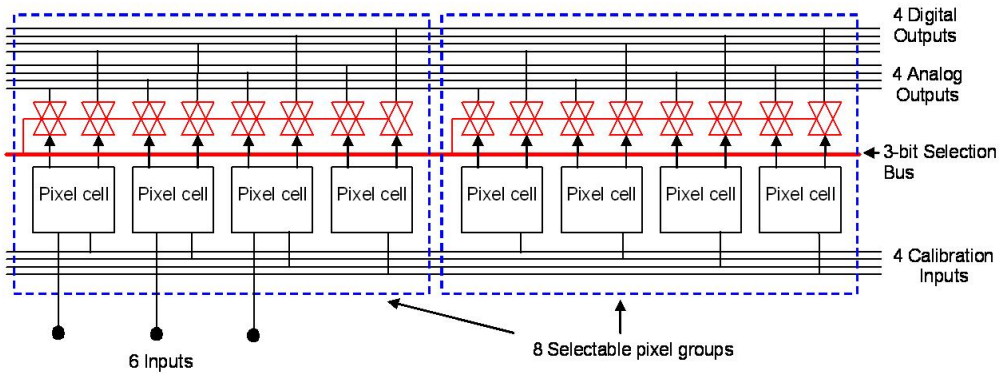


Figure 4.4: Prototype chip architecture.

calibration lines. Six of the 32 front-end cells have direct connection to a dedicated input pad. This configuration gives the possibility to perform cross-talk measurements or check for collective effects in addition to performance measurements on single cells. Four source followers and four digital buffers are used for the multiplexed analog and digital outputs, respectively, in order to be able to drive the relatively high capacity of the output pads and measurement probes ($\sim 1 - 10pF$). Total area of the prototype chip is $2 \times 1mm^2$ (figure 4.5). The chip was fabricated in a commercial $0.13\mu m$ technology through a multiproject wafer run.

4.3.2 Measurements

Figure 4.6 represents the typical pulse at the output of the preamplifier for input charge signals of 0.5 fC, measured with an oscilloscope and a differential probe, for best common mode noise rejection. The measured equivalent noise charge is 98 e^- and the gain in the linear region is 66 mV/fC, while power consumption was $12\mu W$ per pixel cell from 1.2 V power supply. The noise level was measured with no detector or equivalent capacitance connected at the input.

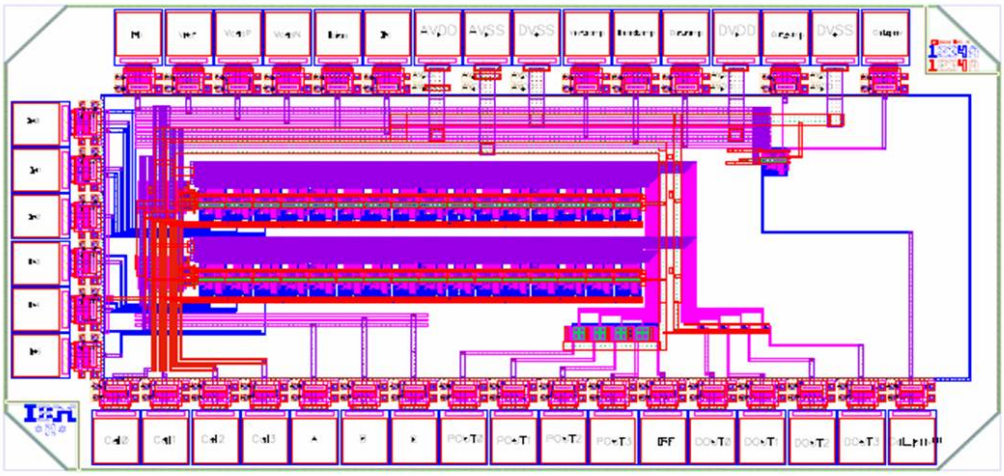


Figure 4.5: Layout of the prototype chip.

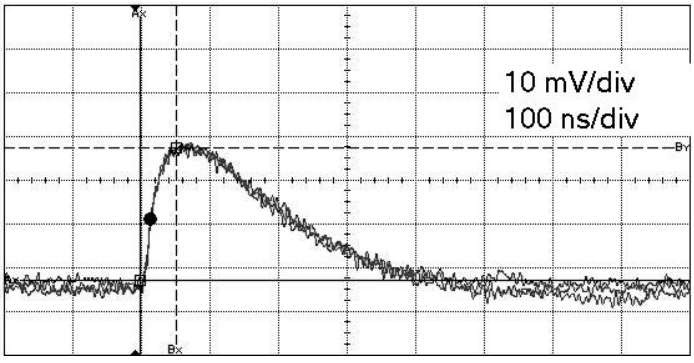


Figure 4.6: Analog signal at the output of the amplifier for 0.5fC input charge

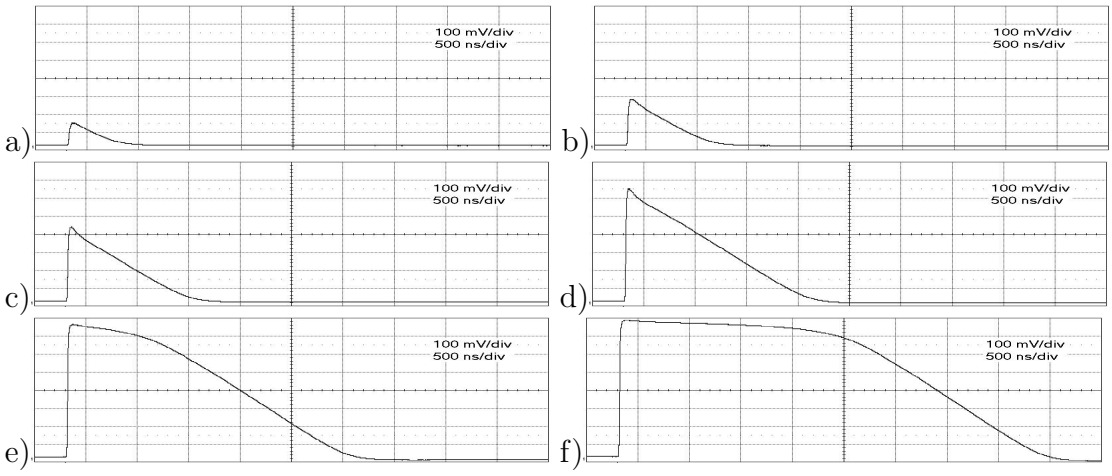


Figure 4.7: Analog signal output for different input charges: a) 2fC b) 3.5fC c) 5fC d) 7.5fC e) 11fC f) 17fC

Figure 4.7 shows the analog signal at the output of the preamplifier for different input charge amplitudes. At high signals the triangular signal shape at the output of the preamplifier, due to the constant current discharge feedback, is clearly visible. At low amplitudes, though, the signal shape is that of a second order response. At this regime the feedback loop does not switch on entirely the discharge current, showing a linear behavior.

Figure 4.8 shows both the analog and digital signals at the output of the preamplifier and discriminator, respectively, for different input charge signals, while figure 4.9 shows the analog amplitude and the time-over-threshold measurement as function of the input charge.

It can be clearly seen that the amplitude of the signal at the output of the amplifier saturates for input charges of the order of 8 fC, while the time-over-threshold remains in linear correspondence with the input charge. A small discontinuity of the TOT

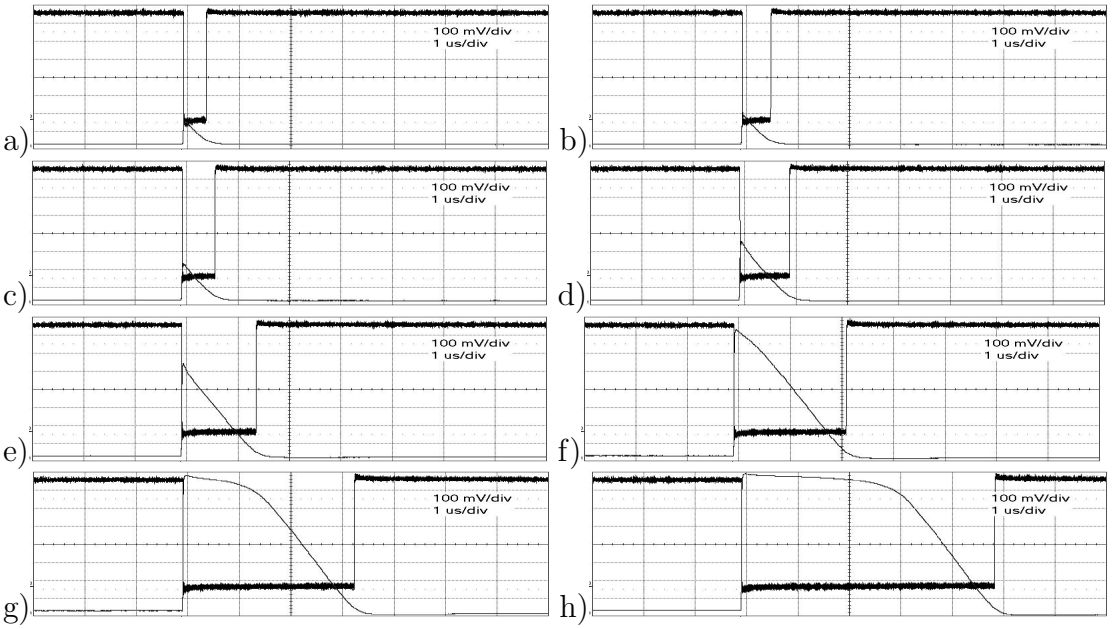


Figure 4.8: Analog and digital signal output for different input charges: a) 2fC b) 2.5fC c) 3fC d) 4.5fC e) 6.5fC f) 10fC g) 15fC h) 22fC

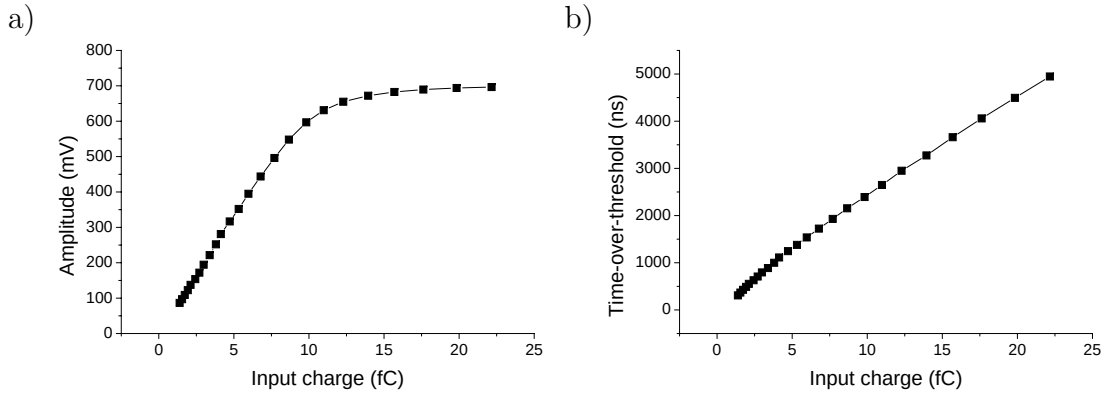


Figure 4.9: Amplitude (a) and time-over-threshold (b).

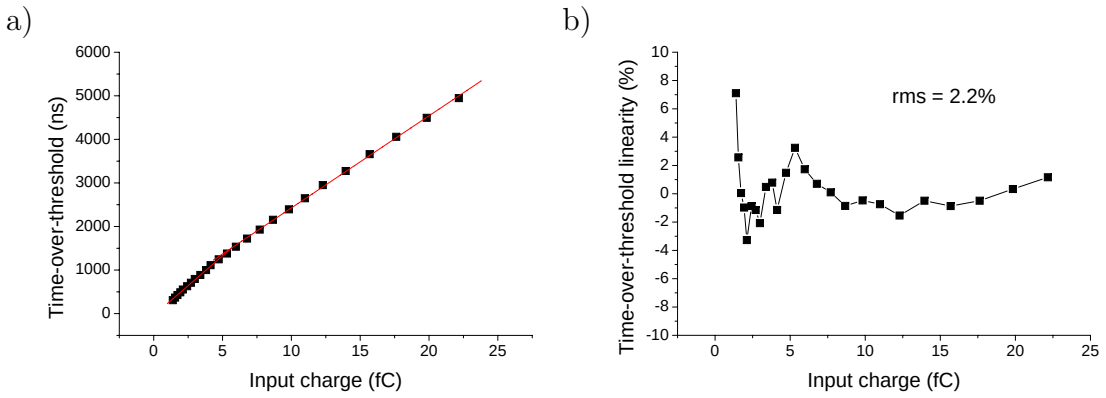


Figure 4.10: a) Bilinear fit of the time-over-threshold b) Time-over-threshold linearity.

slope is visible in the region where the preamplifier saturates. The linearity of the TOT measurement was thus computed using a bilinear fit, accounting for the different slopes for low and high input charge signals. Rms linearity was found to be better than 2.5% (see figure 4.10. At very low charge signals, the linearity is affected by the second order response of the preamplifier discussed earlier.

Analog gain and TOT gain variation across the chip was found to be 5% and 15% respectively. Baseline and offset variations have been measured on all channels of the prototype chip. In order to have the same reference for all measurements, the reference voltage of the preamplifier was kept constant at 300 mV, while the threshold voltage was changed until the trip point of the discriminator was found. The measured distribution of the trip point is shown in figure 4.11. The measure cannot discriminate between the variation of the baseline of the preamplifier and the offset of the discriminator, but it gives the distribution of the sum of the two.

Separate measurement of the offset or baseline variation is limited by the fact that the output signals from the pixel cells do not go through the same output buffer, but each group of eight pixel cells shares one buffer. For that reason a complete

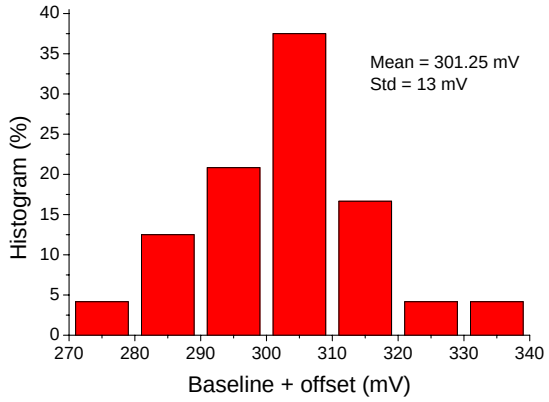


Figure 4.11: Baseline and offset distribution

analysis would have to take into account the relative offset of the output buffers which, unfortunately cannot be measured directly.

The offset and baseline distributions were measured separately in groups of eight pixel cells sharing the same buffer. The average of the mean standard deviation was found to be 3.6 mV and 12 mV, respectively, which are in good accordance with circuital simulations.

4.3.3 Time resolution analysis

A leading-edge discriminator topology such as the one implemented here suffers from high time-walk due to different threshold crossing time for different amplitude of the input signal. The parallel measurement of the amplitude can provide a way to fully compensate this effect off-line, back-projecting the experimental time response at the correct position. This method is limited by the amplitude measurement error (considering both noise and quantization error) and the correct parameterization of the time-walk curve (i.e. time-walk as function of signal amplitude).

The measured time-walk curve as function of the time-over-threshold is shown in

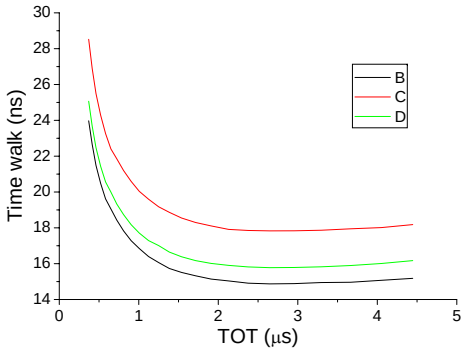


Figure 4.12: Time-walk without correction.

figure 4.12 for three different channels of the same prototype chip. The TOT measure is taken as base of this analysis, as opposed to the charge of the input signal, in order to avoid the influence of the parameterization of the time-over-threshold against the input charge. The measurements were performed with input amplitudes ranging from 1.4 fC to 22 fC in 25 logarithmic-spaced steps. One phenomena worth discussing is the fact that the time-walk slightly decreases for very high input charges. Two causes are believed to contribute to this behavior, namely, the limited slewing rate of the discriminator and the nonlinear influence of the ESD protection circuitry which starts to activate when the calibration pulse reaches amplitude values close to 1.2V of the power rail.

Ideally, the signal shape at the output of the preamplifier is a perfect triangle. In this simple case the equation of the leading edge can be written as:

$$y(t) = \frac{A}{\tau_p} t \tag{4.3.1}$$

where A is the amplitude, τ_p is the peaking time, and t the time variable.

The running time-walk can be defined as the time the signal needs to reach the

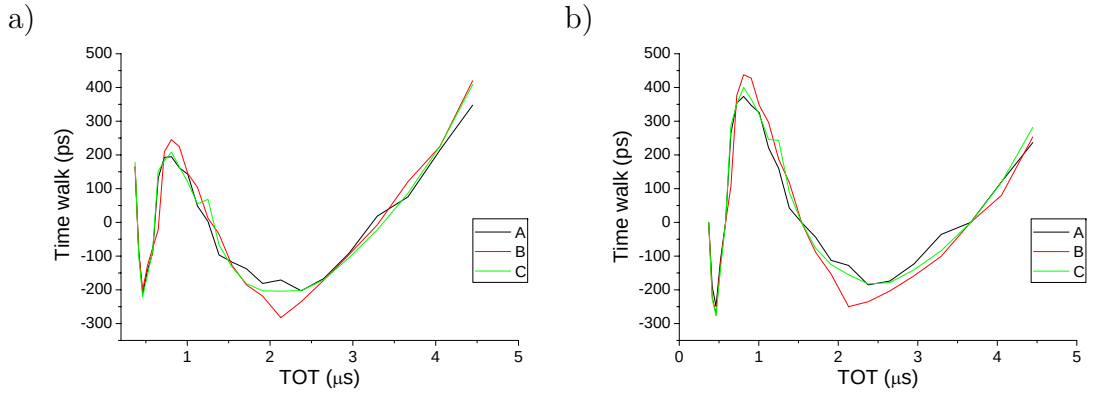


Figure 4.13: Time-walk with amplitude correction using a 3rd order polynomial function. a) 24 calibration points b) 5 calibration points

threshold voltage $y(t_W) = V_{thr}$

$$t_W = \frac{V_{thr}}{A}\tau_p \tag{4.3.2}$$

It follows that, in the ideal case, the running time-walk curve is proportional to the inverse of the amplitude of the signal.

In what follows, the time-walk curve was parameterized using a 3rd order polynomial function of the inverse of the TOT measure:

$$t_w = c_1 + c_2\frac{1}{TOT} + c_3\frac{1}{TOT^2} + c_4\frac{1}{TOT^3} \tag{4.3.3}$$

The constant parameters and the correction was computed considering, on one hand, all 25 measured points, and on other hand, only 5 of them. The last analysis takes into account the fact that the calibration procedure of a high number of pixels is impractical if too many calibration points are considered. Figure 4.13 show the corrected time-walk for the two cases. This demonstrates the fact that the time-walk can be effectively corrected using the time-over-threshold measurement down to 300-400 ps.

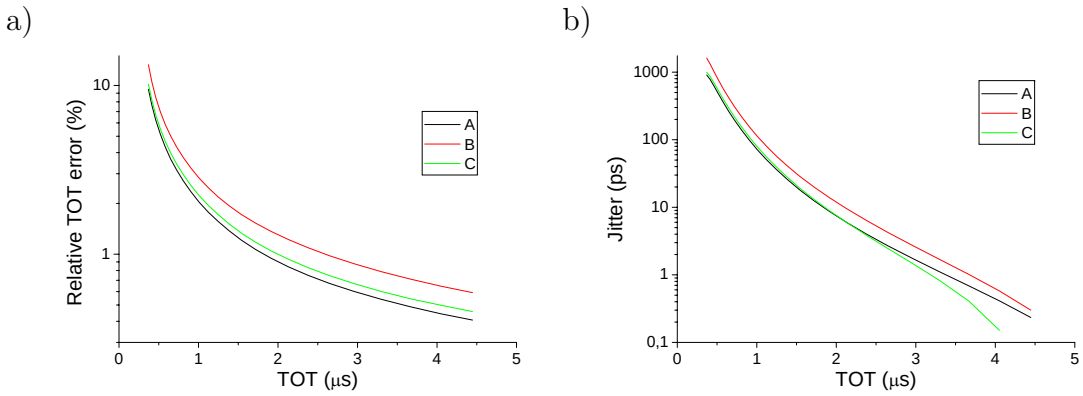


Figure 4.14: a) Measured relative TOT error. b) Contribution of the time-walk correction to jitter.

Another component of the time-resolution is jitter, defined mathematically as the ratio between the rms noise level at the input of the discriminator and the signal slope at the threshold-crossing point. Given the fact that the time-over-threshold measurement is affected by experimental uncertainty, the correction of the time-walk will be affected in accordance. Figure 4.14 shows the standard deviation of the TOT measurement and its contribution to the time resolution.

In figure 4.15 the standard deviation of the leading-edge time (figure 4.15a) and the square root sum of the latter with the contribution of the time-walk correction is given. It can be concluded that the overall time resolution of the system if time-walk correction is considered is well below 2ns.

4.3.4 Radiation hardness

The use of commercial $0.13\mu m$ CMOS technologies in high-energy and nuclear physics did not yet reach maturity. The behavior under radiation of this technologies is still under investigation. One aim of the R&D work described here is to asses the radiation

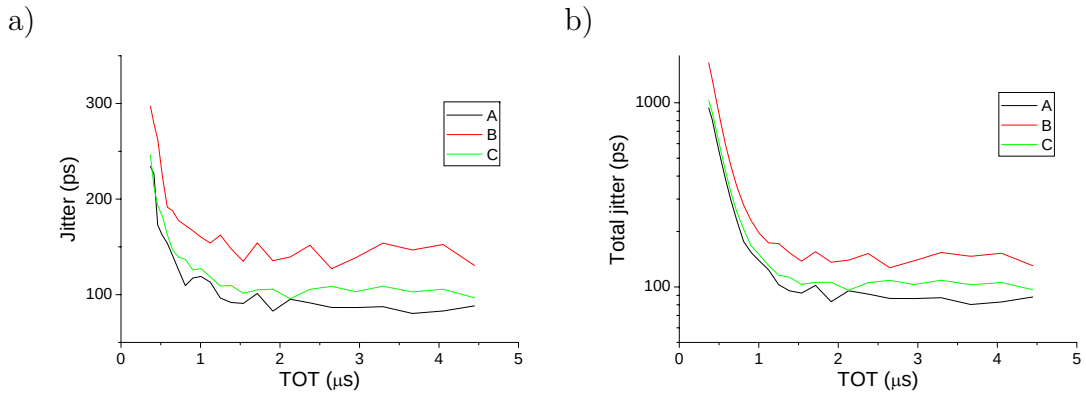


Figure 4.15: a) Measured jitter of the leading discriminator transition. b) Total jitter.

hardness capability of this technology for the radiation fluence foreseen in PANDA. We remind the reader that the design of the analog front-end cell described did not use enclosed geometry transistors but only "classic" linear transistors.

Four prototype chips were exposed to radiation using a X-ray source at CERN. Dose rates of 100 rad/s and 400 rad/s were used. The chips were exposed one day each at total dose up to 19.5 Mrad. After irradiation, the chips were annealed at $25°C \pm 2.5°C$ for several days and subsequently at $100°C \pm 2.5°C$ for one week (see table 4.1).

Chip	Dose rate (rad/s)	Measurement points (Mrad)	Annealing at $25°C$ (days)	Annealing at $100°C$ (days)
A	400	3.5, 5.6, 19.5	4	7
B	100	0.330, 0.771, 6.2	3	7
C	100	0.358, 0.941, 1.7, 2.2, 6.2	1	7
D	100	0.179, 0.358, 5.1	2	7

Table 4.1: Dose rates and annealing time

Measurements were performed after each irradiation and annealing steps. Figure

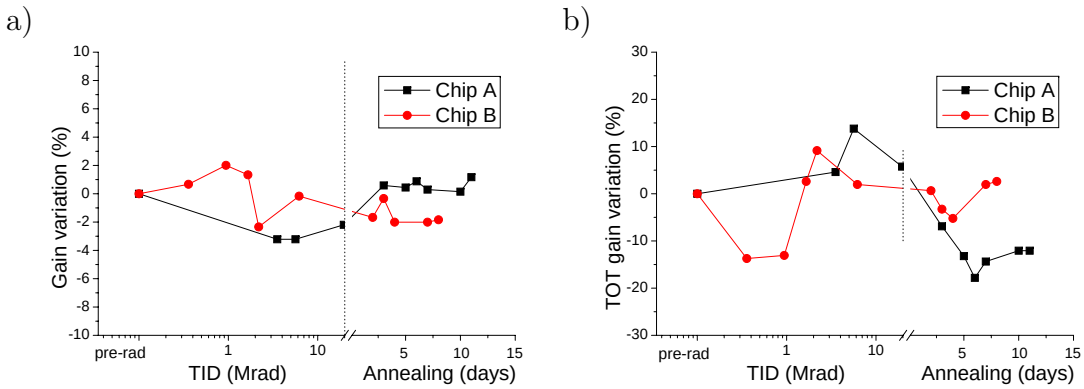


Figure 4.16: Variation of the a) analog gain and b) TOT gain over radiation dose and annealing time.

4.16 shows the variation of the analog gain of the preamplifier and the time-over-threshold gain[3] normalized to the pre-rad value. The relative analog gain variation is within $\pm 3\%$, while the TOT gain variation reaches -15% at worst case. The latter was expected to show higher variation with radiation dose since the value of the discharge current in the preamplifier feedback circuitry is in the order of several nanoamperes, thus comparable to the radiation-induced leakage currents. It is worth observing that in both cases the polarity of the variation is not constant. Until further tests are performed it is unclear weather this is due to the different dose rates in the two cases presented in figure 4.16 or due to technological process variation.

Power consumption was also monitored during irradiation and annealing test. Figure 4.17 shows the variation of the total power consumption with respect to radiation dose and annealing time. The power consumption increases most probably because

[3]The analog and TOT gain were calculated using five measurement points with input charges of 1fC, 3fC, 4fC, 8fC and 12fC. The baseline change due to radiation was compensated by adjusting each time the the threshold voltage to a value of 30 mV above the baseline level, which corresponds to $\sim 0.45 fC$ threshold.

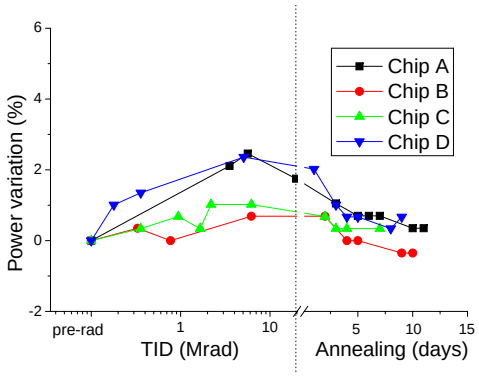


Figure 4.17: Power consumption variation with radiation dose and annealing time.

of radiation-induced leakage currents in lateral transistors which add-up to the static current drawn from the power supply. Nevertheless, the variation is small (below 3%). The maximum power consumption variation is found for radiation doses in $1-6Mrad$ range, which is in close agreement to other radiation tolerance investigations applied to the same technology [64].

As a conclusion, the radiation tolerance of the prototype circuit was found to be satisfactory, despite the exclusive use of linear transistor geometry. The parameter variations induced by radiation can be effectively compensated using static DACs.

Further analysis of the radiation tolerance of the prototype chip is on-going, including neutron beam radiation at the neuron facility in Legnaro, Italy.

Chapter 5

Design of front-end electronics for the Gigatracker silicon pixel beam hodoscope

5.1 Requirements

In P326, backround reduction will be achieved by kinematical selection, highly efficient photon vetoes and particle identification detectors (for more details see first chapter of this work). A beam spectrometer (the *Gigatracker*) is required to do precise momentum, time and angular measurements on all beam tracks.

The main requirements of the Gigatracker are the time resolution of $\sim$ 140 ps on single track, pixel size of $300 \times 300 \mu m^2$, for an angular resolution of ~ 15 μrad, momentum resolution $\sigma(p)/p \sim 0.4\%$ and material budget below $1\% X_0$.

The total time resolution corresponds to about 200ps time resolution per pixel

station, being the most demanding requirement of the Gigatracker design. This is counterbalanced by a more relaxed space and power limitation. The foreseen pixel size corresponds to a power budget of about 1 mW per pixel for a total power density of $1W/cm^2$.

The high beam rate of more than 800MHz with peak intensity in the hottest spot of the beam of $\simeq$ 150kHz/pixel imposes additional constraints to the design.

5.2 Constant fraction discriminator

Direct comparison of the signal against a constant threshold level, usually called leading-edge discrimination, produces a timing signal that depends on the amplitude of the signal, as high amplitude signals cross the threshold level earlier than low amplitude signals do. This effect is called time-walk and cannot be corrected without the information about signal amplitude. An example of time-walk error and its correction is given in the previous chapter for the time-over-threshold method. However, the long dead-time associated with this approach is not affordable with the high pixel hit-rate expected in P326.

Another method for accurate time measurements is the constant-fraction discriminator. It works by comparing a delayed copy of the input signal with an attenuated one. The result is a bipolar signal with the zero-crossing independent of the amplitude of the input signal, ideally eliminating walk error. The name *constant-fraction* comes from the fact that the timing signal corresponds to the moment when the input signal reaches a fixed fraction of its amplitude, rather than a fixed threshold level.

Considering a triangular signal shape with the leading-edge given by $y(t) = A(t/t_p)$, where A is the amplitude and t_p is the peaking time of the signal, the

crossing time is given by:

$$y(t - t_d) = fy(t) \tag{5.2.1}$$

$$A\frac{t - t_d}{t_p} = fA\frac{t}{t_p} \tag{5.2.2}$$

$$t = \frac{t_d}{1 - f} \tag{5.2.3}$$

where t_d and f are the delay and attenuation parameters of the constant-fraction discriminator.

For other signal shapes the 5.2.1 relation is still a good approximation. The independence against amplitude of the zero-crossing time of the bipolar constant-fraction signal still holds, if the signal shape does not change.

Traditionally, the delay is implemented using a coaxial cable of a particular length. For an integrated version, other shaping methods need to be considered. The simplest is the CR differentiator which requires no fraction circuit, providing a bipolar signal with the zero-crossing time, corresponding to the peak of the signal. One drawback of this method is the fact that the overdrive and the slope of the bipolar signal at zero-crossing depend on the trailing edge of the input signal, with additional timing jitter and walk [65].

Other constant-fraction shaping methods were investigated in [65], like simple RC low-pass filter, distributed RC delay line or Nowlin method [66]. The analysis in the cited work found the RC delay-line as best approach.

A CFD circuit using a distributed polysilicon RC line implemented in a 1.2 μm CMOS technology [67] (see figure 5.1a) obtained $\pm$ 150ps walk and jitter below 150ps for an 100:1 dynamic range with an area consumption of $200 \times 950\mu m^2$ and 15mW of power consumption.

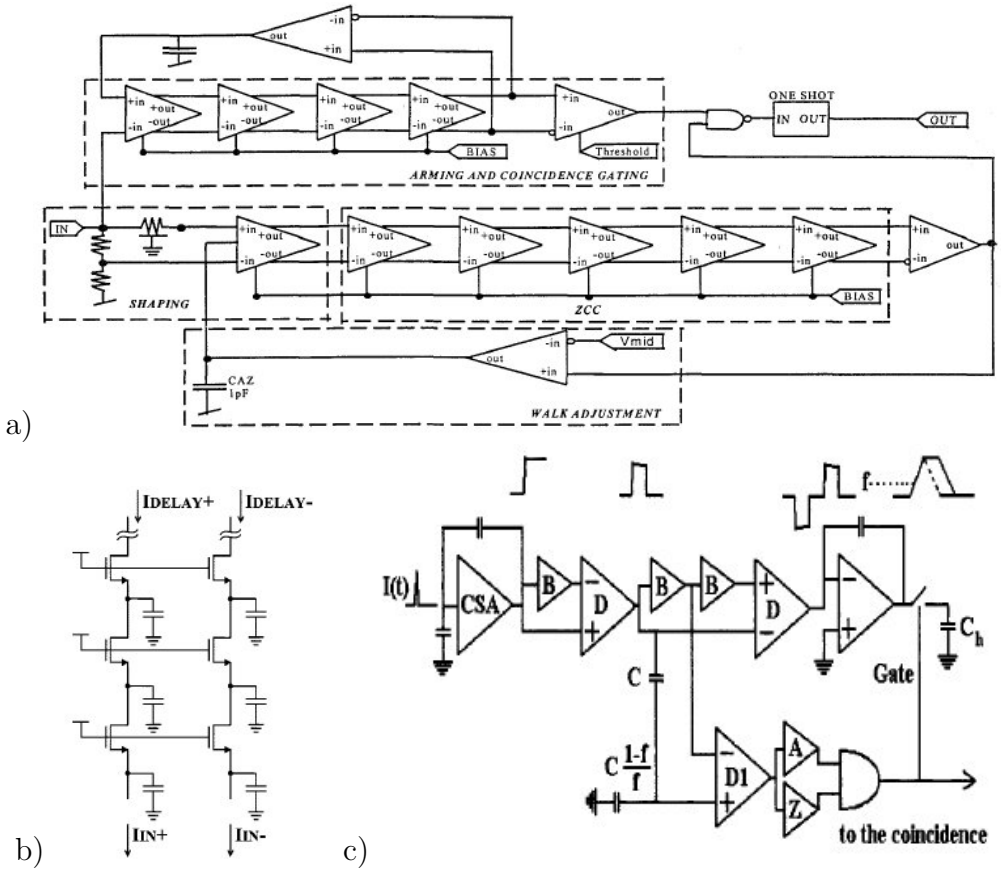


Figure 5.1: a) Schematic of a CFD monolithic implementation using a distributed RC line delay element [67]. b) Current-mode delay filter example [68]. c) CFD circuit using a unity-gain buffer (B) as delay element [54].

Binkley et al. [68] implemented a integrated current-mode 5th order low-pass filter using MOS transistors instead of resistors (see figure 5.1b). While this method provides a good control on the parameters of the delay filter, the DC current supply of the MOS resistors needs to be considerably higher than signal variations for good linearity, thus increasing power consumption. The circuit was implemented in a 0.5 μm CMOS technology and showed intrinsic timing jitter and walk below 100ps for a 3.7:1 dynamic range, with the expense of 50 mW power consumption.

Manfredi et al. [54] used a feedback unity-gain buffer as delay element in the CFD design (figure 5.1c). The results for a front-end circuit optimized for detector capacitance below 200fF were 1ns time-walk and $\sim$ 250ps time jitter, for input charge signals between 2×10^4 and 5×10^5 electrons. The power consumption of the circuit was about 3mW.

For the Gigatracker front-end, given the modest power budget and rather generous silicon space, a passive RC solution is considered. Different optimization aspects are investigated in what follows.

5.2.1 Optimization analysis. Monte-Carlo simulations

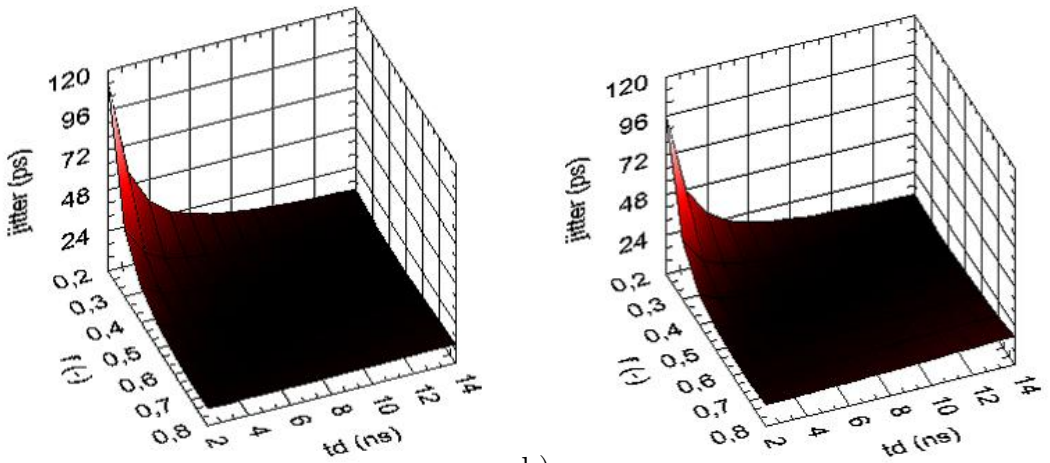
Generally, the optimal performance of a CFD can be found when the zero-crossing time matches the peak of the input signal. Considering again equation 5.2.1 this implies

$$t_d = (1 - f)t_p \tag{5.2.4}$$

This simple result holds for fast input signals (for which the triangular shape approximation stands) and for ideal delay-line and zero-crossing comparator. For a more detailed analysis, one should account for the comparator finite bandwidth and offset, exact signal shape, etc.

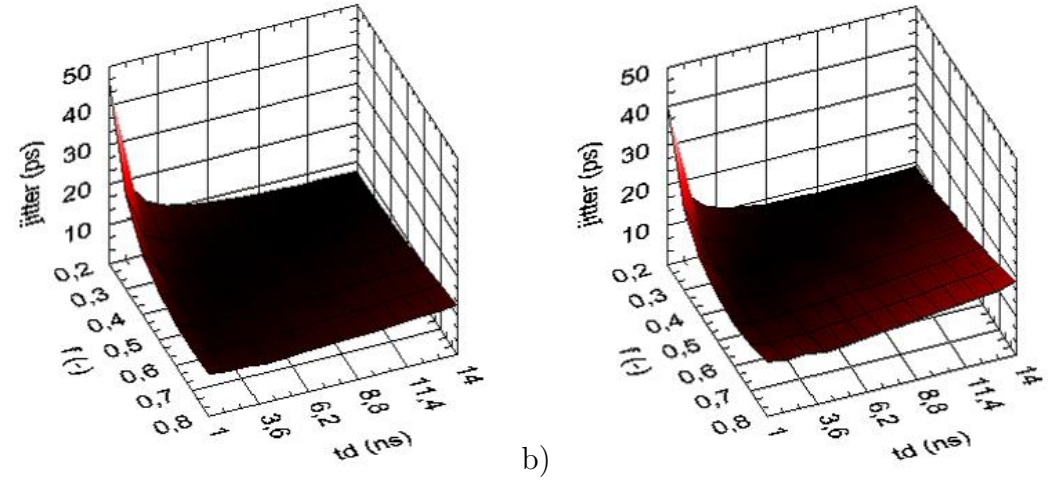
An exact mathematical approach to the optimization problem is rather difficult, especially for higher order delay filters, for which the numerical approach is the only solution.

In an attempt to map the performance against the CFD parameters, a Monte-Carlo simulation was used. Based on detector response simulations carried out for



a) b)

Figure 5.2: Jitter for 1st order delay filter. a) slow comparator b) fast comparator



a) b)

Figure 5.3: Jitter for 2nd order delay filter. a) slow comparator b) fast comparator

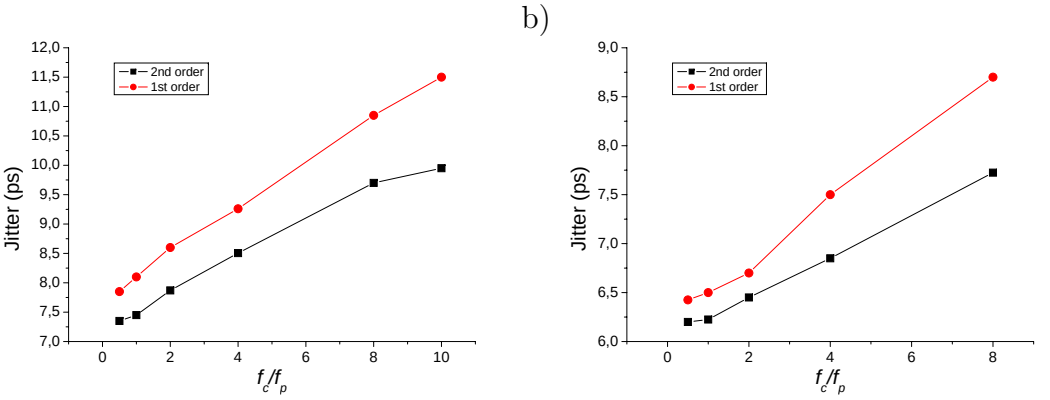


Figure 5.4: Minimum jitter against comparator-to-preamplifier bandwidth ratio for a) 10ns and b) 5ns peaking time

the Gigatracker silicon sensor [69] the detector signal was approximated by a saw-tooth shape signal with 3ns temporal extension. A typical CR-RC preamplifier-shaper response with 10ns peaking time was considered. CFD delay was modelled as a first or second order low-pass filter and the zero-crossing comparator as a first order low-pass filter. The signal was normalized such that at the output of the preamplifier the amplitude can span a 100mV to 1V dynamic range. White noise of 10mV rms amplitude was added at the output of the preamplifier, for a minimum signal-to-noise ratio of 10. The jitter was computed as the variation of the zero-crossing time at the output of the comparator with zero offset, for typically 2000 transient sweeps with 10ps transient step.

One limitation of the model used is the fact that it does not take into account the noise sources associated with the filter components, nor the exact frequency dependence of the noise. Nevertheless, the simulation gives a general idea of how performance change with respect to the CFD parameters, employing a simple model.

Figures 5.2 and 5.3 show the maps of the jitter against fraction and delay time[1] parameters of the CFD for first order and second order delay filter, respectively. Slow and fast comparators considered in this figures have the bandwidth equal or 4 times higher than the preamplifier bandwidth, respectively[2].

For both first order and second order delay filters, the jitter performance is worse for low values of fraction and delay and improves for higher values. In the case of second order delay filter a more pronounced optimum region can be found for values of the CFD parameters in the middle range. On the other hand, the variation of the

[1]Delay time of the low-pass filters is defined as the characteristic RC time-constant

[2]The time constants of the slow and fast comparators is t_p and $t_p/4$ respectively, where t_p is the peaking time of the preamplifier response, 10ns in this case.

performance in this case is smaller than in the case of the first order filter.

Figure 5.4 shows the minimum jitter for optimal combination of CFD parameters, versus the ratio between the bandwidths of the comparator and preamplifier, for 10ns and 5ns preamplifier peaking time. An almost linear dependence can be observed. The explanation is that a slower comparator reduces the overall bandwidth, thus filtering noise.

Time-walk performance was mapped against CFD parameters considering an offset of 1 mV and an amplitude dynamic range of 1:10 (from 100mV to 1V). When investigating the time-walk influence over the time resolution of the system one has to take into account the Landau distribution of the signal amplitude. The time response of the system will be distributed according to the convolution of the Landau with the time-walk curve. In other words, the time-walk corresponding to each amplitude, with respect to an arbitrary moment, must be weighted by the probability associated with that particular amplitude. The time-walk component of the overall time-resolution is the standard variation of this distribution.

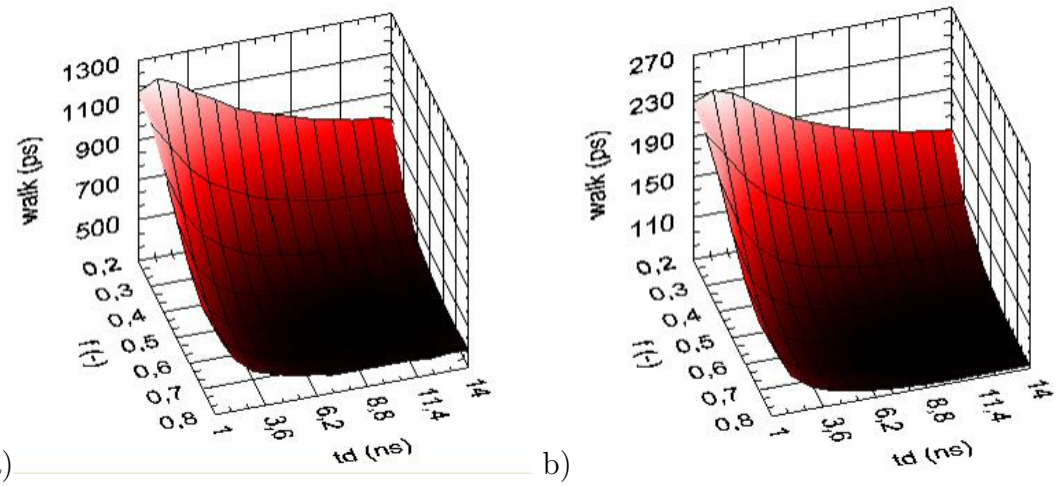
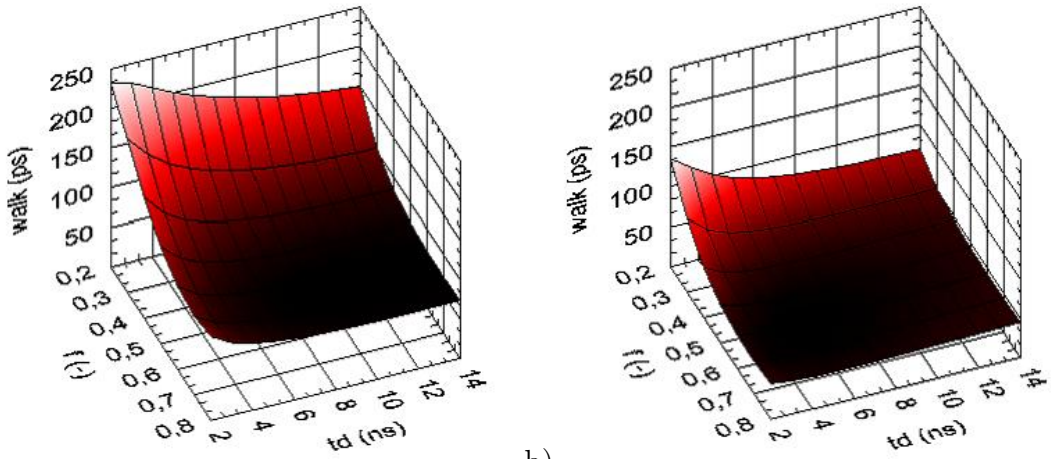


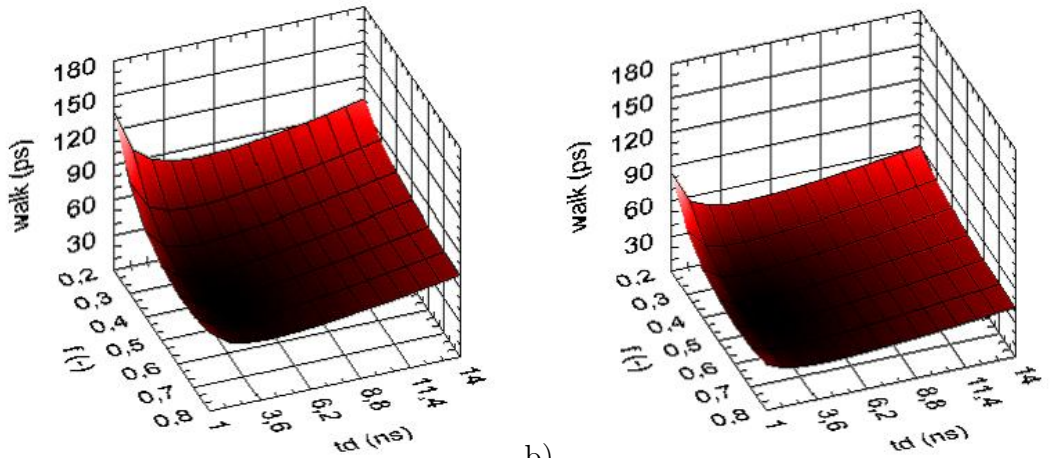
Figure 5.5: Walk for 1st order delay filter. a) 10:1 dynamic range b) Landau corrected

Figure 5.5 shows both the 10:1 dynamic range time-walk and the corresponding Landau corrected walk. The parameters of the Landau distribution considered were the most probable amplitude of 250 mV and a Landau sigma of 75 mV. It can be seen that the corrected walk value is roughly 20% of the raw time-walk.



a) b)

Figure 5.6: Time walk for 1st order delay filter. a) slow comparator b) fast comparator



a) b)

Figure 5.7: Time walk for 2nd order delay filter. a) slow comparator b) fast comparator

Figures 5.6 and 5.7 show the Landau corrected walk against CFD parameters, for

first and second order delay filter. For first order filters best walk performance is obtained for high values of the delay and fraction parameters. In the case of second order filter, a more pronounced optimum region can be found for delay times around 4ns and fraction around $0.6 - 0.7$.

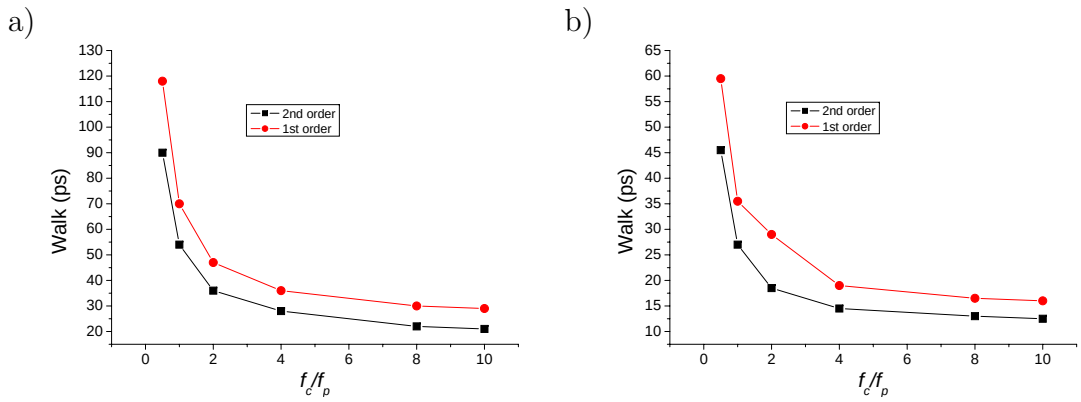


Figure 5.8: Minimum walk against comparator-to-preamplifier bandwidth ratio for a) 10ns and b) 5ns peaking time

In figure 5.8, minimum time-walk is drawn versus the ratio between the comparator and preamplifier bandwidth. In this case the performance greatly improves for faster comparators. This is due to the fact that the slope of the bipolar signal degrades in the case of a slow comparator, while the speed of the comparator has no effect on the offset value. This find suggests that the choice of the speed of the comparator is a trade-off between jitter and walk performance.

An analytical analysis (figure 5.9) shows that for the same zero-crossing time, the slope of the bipolar signal increases for higher delay filter order, and thus, better performance is expected.

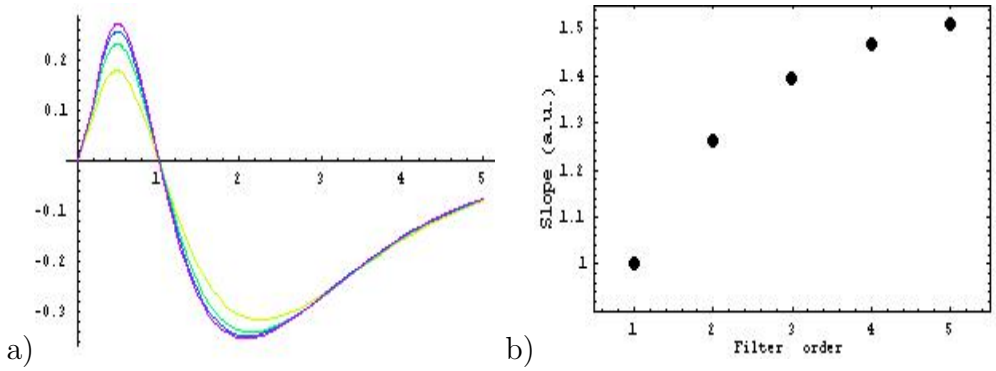


Figure 5.9: a) CFD signal for different delay filter orders (yellow = 1, green = 2, blue = 3, red = 4). b) Zero-crossing slope versus filter order.

5.3 Monolithic realization

5.3.1 CFD

For the practical realization of the CFD shaping-filter we have chosen the passive 4th order RC filter shown in figure 5.10. Resistors $R1 - R10$ and capacitors $C1 - C4$ form the RC 4th order low-pass filter chosen. The resistors $Rf1$ and $Rf2$ connected as inverted pass resistors from input to output form the fraction circuit. The passive elements are implemented using available poly-silicon resistors and vertical plate capacitors. Differential topology was preferred for a better rejection of power supply and other spurious common-mode noise.

The bipolar signal is processed by a zero-crossing amplifier shown in figure 5.11. It consists of two cascaded differential stages with resistive loads. This configuration is used for best speed performance.

The CFD design includes a novel current mode dynamic offset compensation. Figure 5.12 shows the complete schematic of the CFD block. The PMOS transistors $U2$ and $U3$ form a differential amplifier with the tail current source represented by

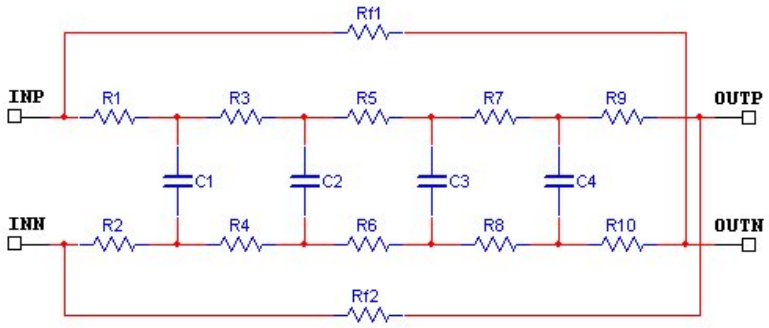


Figure 5.10: CFD filter schematic

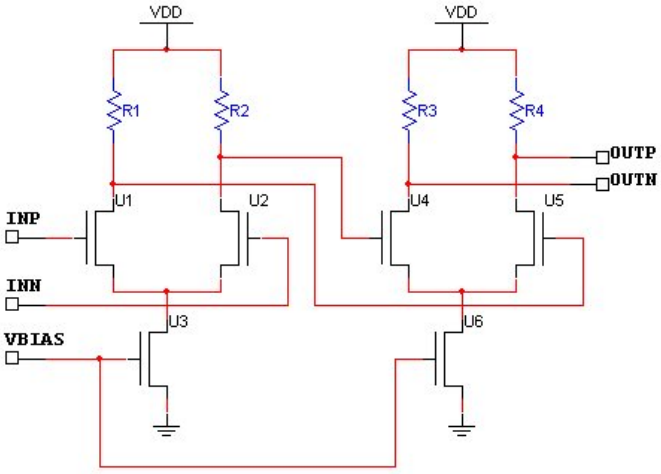


Figure 5.11: CFD zero-crossing discriminator schematic

$U4$. The gates of the differential pair are connected to the output of the discriminator through an low-pass filter with a high time constant. The differential output signal is thus averaged and converted to a differential current that is injected into the resistive ladder of the CFD filter. A slow feedback is obtained which works against offset sources, either at the input of the discriminator or from the front-end stages. The slow feedback provides also a high-pass overall transfer function, reducing the low frequency white and $1/f$ noise.

Figure 5.13 shows the effect of the offset compensation scheme. The signal at the input of the zero-crossing amplifier (figure 5.13a) shows a DC component which is, in fact, generated by the offset compensation circuit in order to counterbalance the offset present at this point. It can be clearly seen that the offset-to-signal ratio is greatly improved.

The time-constant, gain and tail current of the offset compensation circuitry is chosen as a compromise between compensation efficiency and rate capability. A reset circuit that discharges the capacitor $C5$ in figure 5.12 after high amplitude pulses may break this compromise. This reset circuitry was not implemented in this prototype.

In the attempt to limit the noise generated by the filter the resistors $R1 - R10$ in figure 5.10 were chosen as small as possible ($\sim 1k\Omega$). Given the technological size limitations, smaller resistors would need to have higher area, and thus, higher parasitics. The capacitors $C1 - C4$ have a value of 170fC. With this values, the dynamic current drawn for the signal source (i.e. output stage of the preamplifier) shows a peak of about 100 μA. For this reason, a buffer stage with good current-drive capability is needed between preamplifier and CFD filter.

Observing the fact that the resistors closer to the output of the filter ($R9 - R10$)

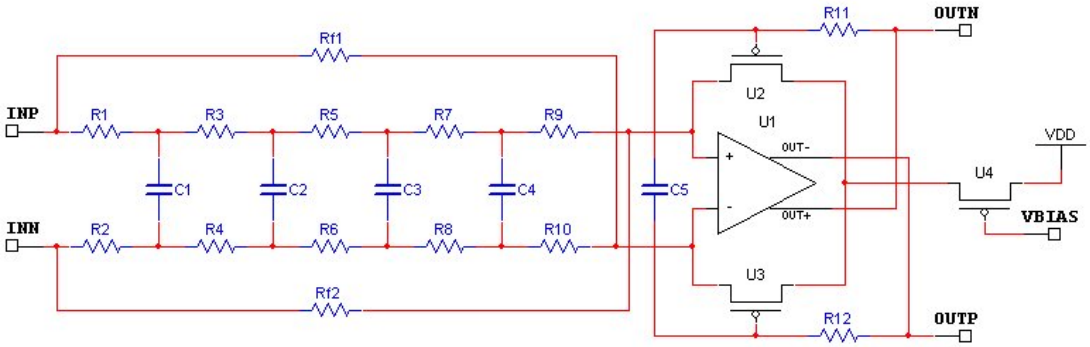


Figure 5.12: Dynamic offset compensation. U1 is the two stage zero-crossing discriminator shown in figure 5.11.

have a higher contribution to noise than the resistors connected at the input nodes of the filter ($R1 - R2$), the values of this resistors can be changed according to their noise contribution, keeping the RC product constant.

Figure 5.14 shows the effect of uniform and non-uniform scaling of the CFD filter components on the peak dynamic current drown from the input and on the total rms noise. Uniform scaling means that all components are scaled by the same factor k, keeping the RC product constant. Non-uniform scaling means that the resistors are scaled down from input to output by a factor kc while capacitors are inversely scaled ($R1 = kc \times R3$, $R3 = kc \times R5$, etc., $C1 = C2/kc$, etc.). One can observe that for the same reduction of the peak dynamic current, the noise penalty is lower in the case of the non-uniform scaling.

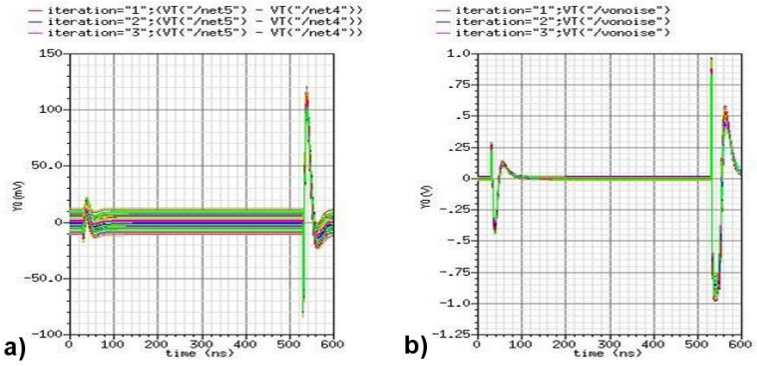


Figure 5.13: SPICE simulation of the dynamic offset scheme. a) Signal at the output of the CFD filter. b) Signal at the output of the zero-crossing discriminator.

5.3.2 Digital output stage with non-symmetrical hysteresis

Traditionally, constant-fraction discriminators use a second arming leading-edge discriminator (see figure 5.15). Due to noise, the output of the zero-crossing discriminator will bounce randomly up and down. For this reason, its output is gated by the output of the arming discriminator, which is at high level when the input signal goes over a certain threshold level. In a multichannel environment, this may cause a severe problem, since the output of the zero-crossing discriminator is still bouncing up and down, generating spurious noise in the power supply rails and in the substrate.

This problem can be avoided, implementing a novel discriminator concept featuring non-symmetrical hysteresis. The threshold level is applied only to the negative falling edge of the bipolar signal (see figure 5.16a). When the bipolar signal crosses this level, the threshold is cancelled and the discriminator acts as a zero-crossing discriminator.

Spurious switching on noise of the discriminator is thus avoided, as long as the threshold level is well above noise.

a) b)

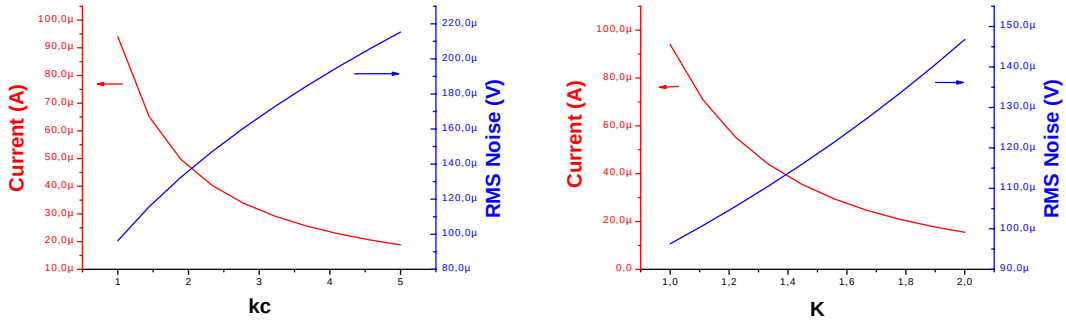


Figure 5.14: a) Uniform scaling of the CFD filter components. b) Non-uniform scaling.

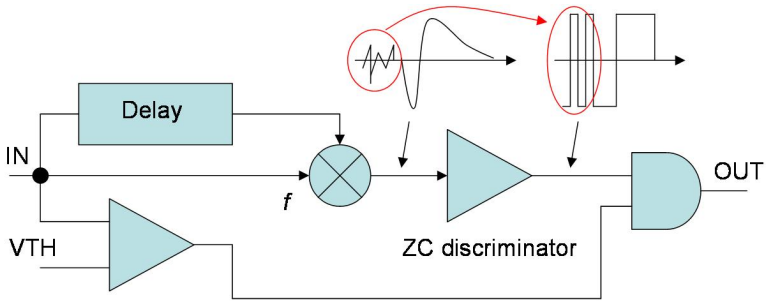


Figure 5.15: Traditional CFD with arming discriminator path.

Figure 5.16b shows the circuital implementation of the concept. It contains two differential pairs sharing the same resistive load. One of the differential pairs ($U1$, $U2$) is connected to the output of the zero-crossing amplifier, while the other ($U3$, $U4$) is controlled by the externally applied hysteresis voltage. Transistors $U7$ and $U8$ act as two switches which short-cut the hysteresis differential pair when the output of the discriminator is at high logic state.

The circuit that controls the hysteresis level needs to be very fast ($2 - 3ns$) such that the hysteresis is cancelled before the zero-crossing of the bipolar signal. For this reason, transistors implementing the circuitry are chosen as small as possible.

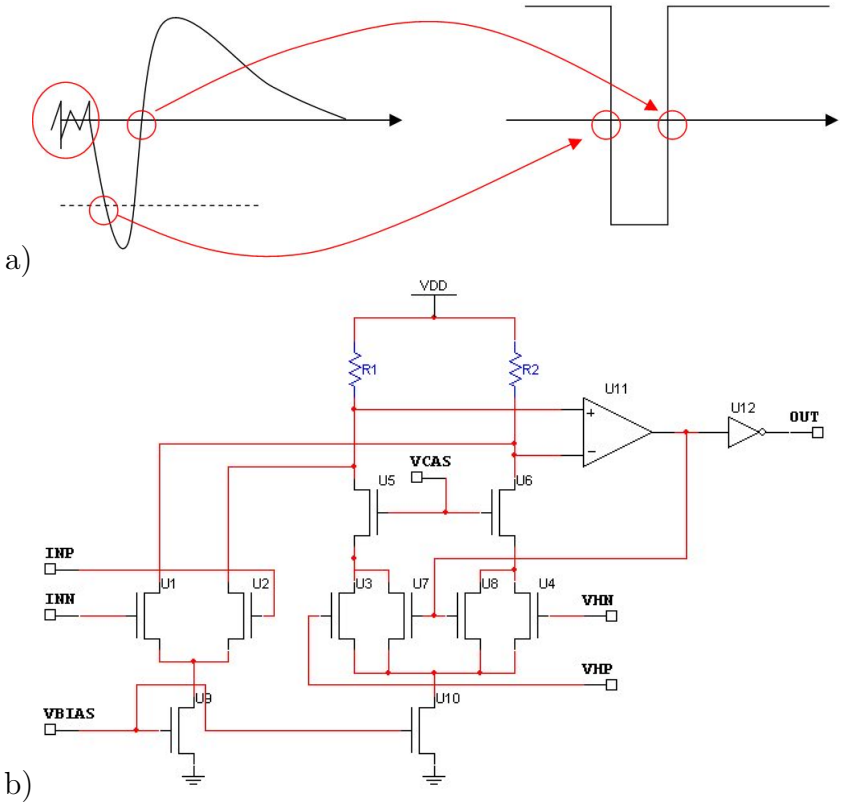


Figure 5.16: a) Non-symmetrical hysteresis concept. b) Schematic of the output buffer implementing the variable hysteresis.

Additional time-walk may occur since small transistor area is a mismatch source. The simulations show that the circuit does not add more than 100ps to the time-walk figure.

5.3.3 Low-noise preamplifier

The schematic of the low-noise preamplifier is shown in figure 5.17. It consists of a cascoded common-source amplifier ($U1$ and $U2$) with current source load ($U3$ and $U4$) and a source follower ($U7$). The feedback circuitry ($R1$ and $C1$) provides the

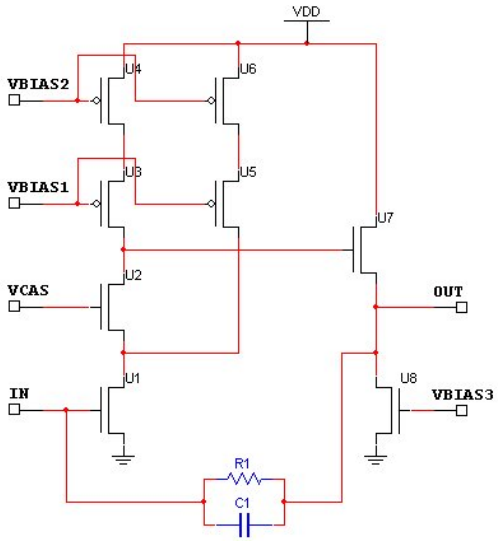


Figure 5.17: Low-noise preamplifier schematic

charge sensitive function and the required signal shaping.

The input referred white noise of the stage can be written as follows, neglecting the contribution of the cascodes[70]:

$$\overline{v_{n,in}^2} \approx 4kT\frac{2}{3}\left(\frac{1}{g_{m1}} + \frac{g_{m4}}{g_{m1}^2}\right) \tag{5.3.1}$$

where g_{m1} and g_{m4} are the transconductances of $U1$ and $U4$.

The current source formed by transistors $U5$ and $U6$ "steels" current from the cascode path. This improves the noise performance, by increasing g_{m1} and decreasing g_{m4} in equation 5.3.1, and thus, reducing the contribution of the current source load.

The preamplifier consumes $70\mu A$ from the power supply, providing $40mV/fC$ gain, $5ns$ peaking time and an equivalent-noise charge of 150 electrons for a 200fF detector capacitance. Simulated linearity (see figure 5.18) is better than 1.5%.

For an input dynamic range from 1fC to 10fC, the preamplifier provides a shaped pulse with amplitude between 40mV and 400mV. The maximum differential signal

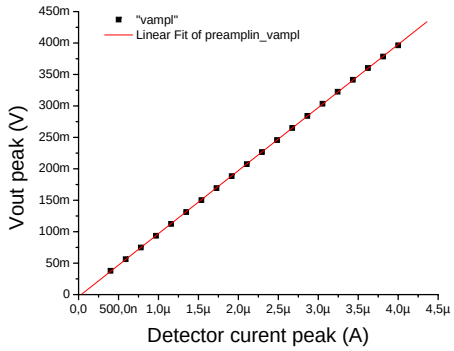


Figure 5.18: Preamplifier linearity

span at the input of the CFD filter is 800mV.

5.3.4 Buffer

As discussed in the subsection regarding the CFD, a buffer stage with high current drive capability is needed between the preamplifier and CFD stage. This stage also performs single-ended to differential conversion. The buffer schematic is shown in figure 5.19a. It consists of a fully differential amplifier stage ($U1 - U4$), two inverter output stages ($U6 - U9$) and a common-mode feedback circuit ($U10 - U13$). The inverter stages use thick oxide transistors. On one hand, they provide lower gate capacity, improving speed and compensation efficiency. On the other hand, they have a higher threshold voltage than normal transistors, leading to a class-AB operation of the stage. The drawback of this implementation is that the class-AB operation is not controlled by any additional circuitry, and thus is subject to process variations that affect both performance and power consumption. Nevertheless, corner simulations showed satisfactory performance.

Figure 5.19b shows the schematic of the single-ended to differential conversion. The inputs $INPREAMP$ and $INREPLICA$ are connected, respectively, to the

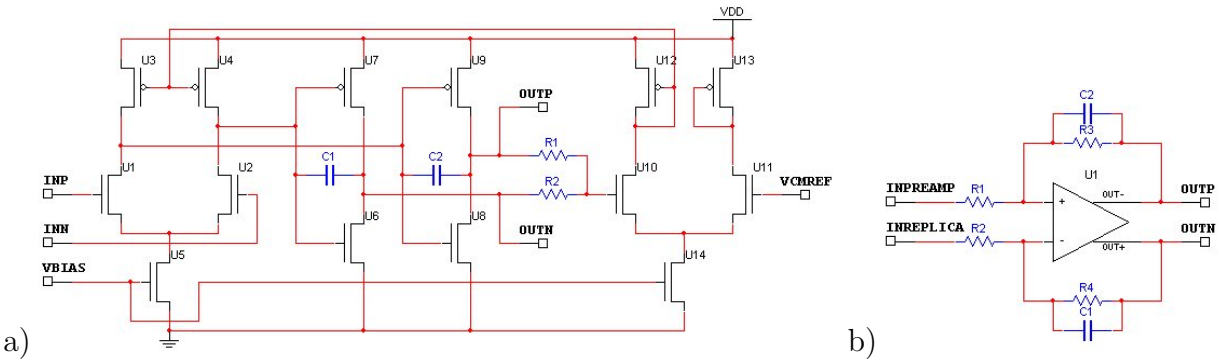


Figure 5.19: a) Buffer amplifier schematic. b) Single-ended to differential configuration.

output of the preamplifier and to the output of a replica circuit of the preamplifier needed to provide the correct DC level at this node. Nested compensation is provided by capacitors $C1$ and $C2$ for better stability.

5.4 Simulation results

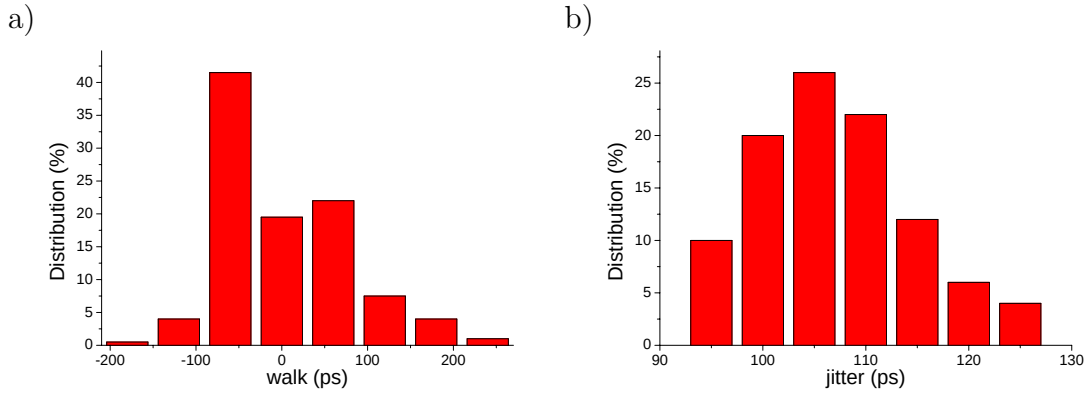


Figure 5.20: Simulated performance variation considering process spread and mismatch. a) Walk b) Jitter

The entire system was simulated using a circuital simulation tool against process

Block	$I_{DD}(\mu A)$	$P@V_{DD} = 1.2V(\mu W)$
Preamp	70	84
Buffer	250	300
CFD	110	132
Output buffer	70	84
Total	**500**	**600**

Table 5.1: Power consumption of the pixel cell

variation[3] and mismatch[4]. Figure 5.20 shows the results of a Monte-Carlo analysis which takes into account both variation sources. Time walk was calculated subtracting the values of the time response of the system for input signals of 1fC and 10fC. Jitter was computed for an input signal of 1fC.

The results show a worst case time-walk and jitter lower than 250ps and 130ps respectively. Time-walk has to be further corrected accounting for the Landau distribution of the input signal. Using the analysis performed in the previous section the time walk contribution to the total time resolution can be approximated to be $250ps \times 20\% = 50ps$. The total power consumption of the front-end chain is 0.6 mW, and the silicon area is $280 \times 80\mu m^2$. Table 5.1 reports the power consumption of all building blocks.

The simulation results indicate a performance with an area and power budget not yet reported in the literature.

[3]Variation of process parameters from wafer to wafer. It affects all components of the design in the same way.

[4]Variation of process parameters within the same wafer. This effect affects the components of the design in a random way, giving birth to mismatch effects like, for example, the offset.

5.5 Prototype chip

A prototype chip has been designed and submitted to fabrication, in order to test the constant-fraction discrimination concept described here. The layout of the pixel cell is shown in figure 5.21. The pixelcell area is $280 \times 80 \mu m^2$.

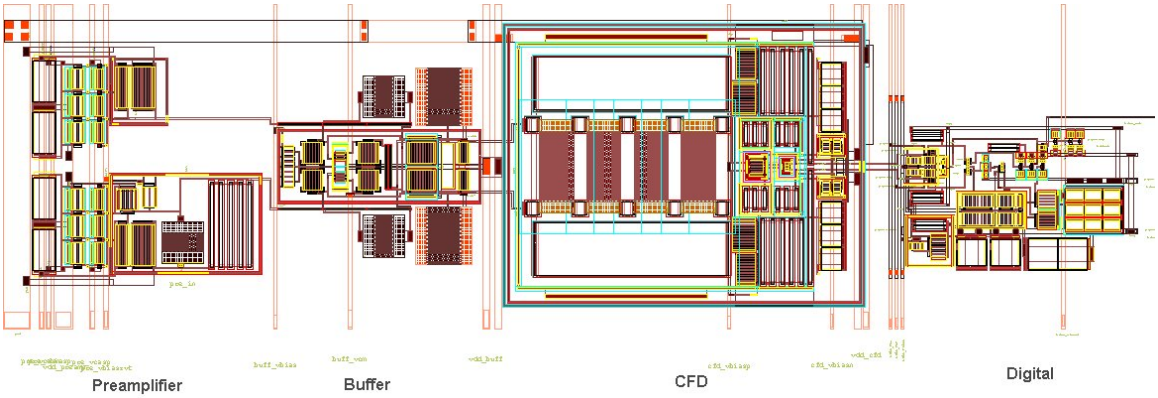


Figure 5.21: Pixel cell layout.

The prototype chip contains three pixel cells (figure 5.22C1,C2,C3), one preamplifier block (figure 5.22B) and one CFD block (figure 5.22A). One calibration capacitor is connected at the input of each of the three pixel cells and at the input of the preamplifier block. Two calibration lines use these capacitors in order to pulse in the signal charge from a step generator outside the chip. One of the pixel cells have the input connected directly to an input pad.

The CFD block is connected directly to input pads and can be stimulated by an external shaper or function generator. The outputs of the CFD block and the output of the preamplifier block are buffered by three operational amplifiers in follower configuration.

On the same die, a 10GHz gated digital voltage-controlled oscillator[5] has been integrated in order to test the immunity of analog blocks to heavy digital noise. The power supply of the VCO has been separated from the power supply of the analog blocks to avoid the coupling of the noise through the power rail.

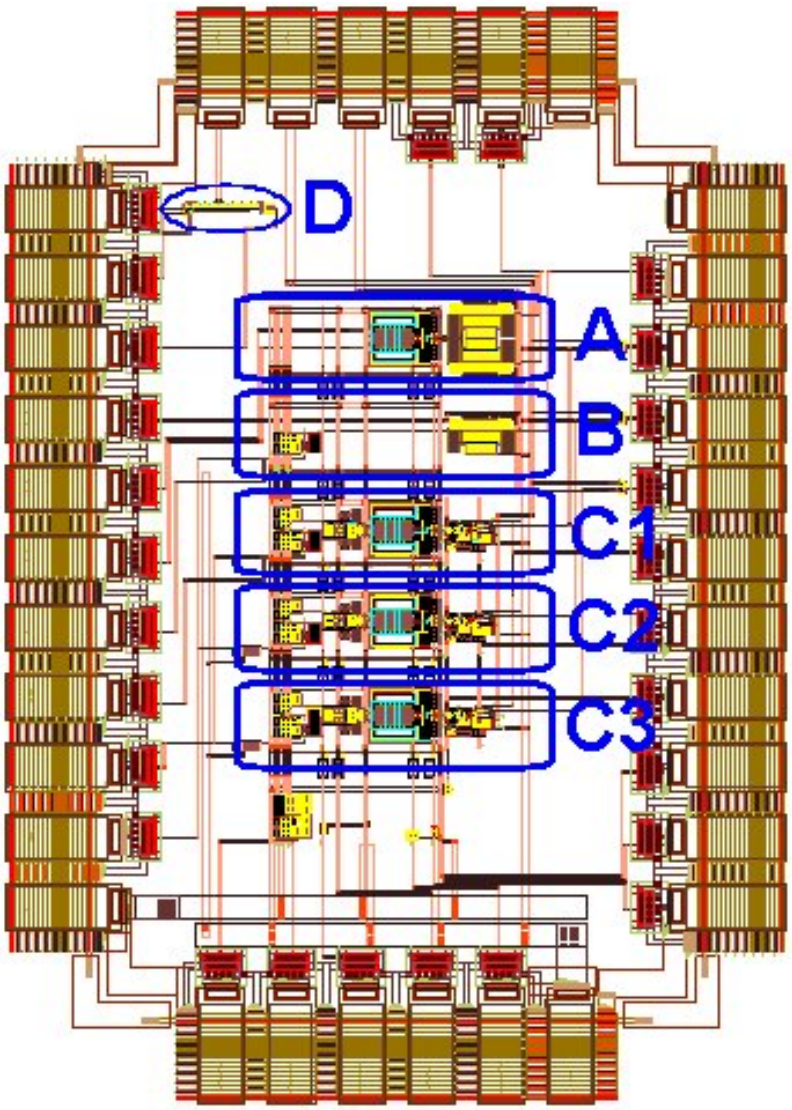


Figure 5.22: Prototype chip layout.

[5]The integrated VCO was designed by Ozgur Cobanoglu from University of Turin

Conclusions

Over the last three decades, silicon detectors have become one of the standard tools for experimental physics, covering a wide range of scientific applications in nuclear and high-energy physics. Extensive research and development work over this years has lead to more versatile and more accurate detector systems, that eventually found applications in a wider range of scientific domains through successful spin-offs towards X-ray spectroscopy, medical physics and imaging.

Constant improvement of performance of silicon detector systems has been assisted by the research and development work regarding the associated front-end electronics. The readout electronics of silicon detectors became faster, smaller and more radiation tolerant. The progress of the microelectronic industry over the last years sustained this trend, opening new possibilities in terms of performance and integration density. On the other hand, new issues emerged with modern deep submicron technologies, opening new R&D fronts.

Tracking detectors in fixed target and collider experiments require large detection area with high position and momentum resolution, imposing severe limits on the power consumption and on the amount of material used for mechanical support, cooling, etc. Consequently, system integration becomes a demanding task, involving

careful design and optimization of the production flow, and extensive quality assurance scheme.

The focus of this work goes on the design and integration of front-end electronic systems for three-dimensional silicon detectors in three modern experiments covering both nuclear and high-energy physics.

The problematic of system integration is covered in the case of the silicon-drift detector (SDD) subsystem of the Inner Tracking System (ITS) of ALICE. A dedicated hardware and software test system has been developed for the integration of the front-end electronics of the silicon-drift detectors. The front-end architecture, the production flow, as well as the results of performance measurements acquired during production using the test bench described are documented here.

The design of a custom analog front-end for the pixel detectors of the Micro Vertex Detector (MVD) of PANDA is described. A prototype chip has been realized and tested. The performance of the front-end is reviewed, as well as the results of radiation tolerance measurements. Additionally, a time resolution analysis has been performed, answering recent concerns in the PANDA collaboration that call for a better time resolution in view of a broader physics program of the experiment.

The time resolution issue in silicon pixel detectors was also treated from the perspective of the planned P326 experiment at CERN. A novel pixel front-end design has been developed using the constant-fraction discriminator technique, that meets the 200ps time resolution requirement of the experiment. The simulation results indicate a performance with an area and power budget not yet reported in the literature. A prototype chip was designed and submitted for fabrication.

The three topics covered here bind together in a detailed, but not exhaustive,

overview of the issues of front-end electronics for silicon detectors in modern nuclear and high-energy physics experiments. They share the use of state-of-the-art semiconductor technologies, with their advantages and limitations, and combine the use of low-noise and low-power design techniques and architectures. The apparent "mis-arrangement" of the design-test-integration flow, apart from being an exact photography of the research experience reported here, also underlines the repetitive and self-consistent nature of the sequence, as all knowledge gathered from a completed design-to-integration flow is the starting base for a new one.

Bibliography

[1] James Chadwick. Possible Existence of a Neutron. *Nature*, page 312, 1932. http://dbhs.wvusd.k12.ca.us/webdocs/Chem-History/Chadwick-neutron-letter.html.

[2] Carlo Bernardini. AdA:The First Electron-Positron Collider. *Phys. Perspect.*, 6:156–183, 2004.

[3] SHELDON LEE GLASHOW. TOWARDS A UNIFIED THEORY - THREADS IN A TAPESTRY. Nobel Lecture, December 1979.

[4] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys. Lett.*, B565:61–75, 2003.

[5] ALEPH Collaboration. Precision electroweak measurements on the Z resonance. *Phys. Rept.*, 427:257, 2006.

[6] Claude Amsler. Proton antiproton annihilation and meson spectroscopy with the Crystal Barrel. *Rev. Mod. Phys.*, 70:1293–1340, 1998.

[7] Mushin Harakeh et al. NuPECC Long Range Plan 2004: Perspectives for Nuclear Physics Research in Europe in the Coming Decade and Beyond. Nupecc report, NuPECC, 2004.

[8] Finuda Collaboration. First results on ${}^{12}_{\Lambda}C$ production at DA@FNE [rapid communication]. *Physics Letters B*, 622:35–44, August 2005.

[9] Finuda Collaboration. Search for ${}^{6}_{\Lambda}H$ and ${}^{7}_{\Lambda}H$ with the $(Kstop^{-}, \pi^{-})$ reaction. *Physics Letters B*, 640:145–149, September 2006.

[10] Collins, J. C. and Perry, M. J. Superdense Matter: Neutrons or Asymptotically Free Quarks? *Phys. Rev. Lett.*, 34(21):1353–1356, May 1975.

[11] N. Cabibbo and G. Parisi. *Phys.Lett.*, B59:67, 1975.

[12] W.-M. Yao et al. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006.

[13] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[14] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.*, 13(4):138–140, Jul 1964.

[15] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[16] G. Anelli et al. Proposal to Measure the Rare Decay $K^{+} \rightarrow \pi^{+}\nu\bar{\nu}$ at the CERN SPS. Technical report, CERN, Geneve, 2005. CERN-SPSC-2005-013.

[17] B. Sciascia. Kaon physics. 2006.

[18] G. Buchalla and A.J. Buras. The rare decays K $\rightarrow$ pi nu anti-nu, B $\rightarrow$ X nu anti-nu and B $\rightarrow$ l+ %l-: An update. *Nucl. Phys. B*, 548:309, 1999. arXiv:hep-ph/9901288.

[19] U. Haisch A. J. Buras, M. Gorbahn and U. Nierste. The Rare Decay $K^{+} \rightarrow \pi^{+}\nu\bar{\nu}$ at the Next-to-Next-to-Leading Order in QCD. *Phys. Rev. Lett.*, 95, 2005. arXiv:hep-ph/0508165.

[20] S. Adler *et al.* [E787 Collaboration]. Further evidence for the decay K+ $\rightarrow$ pi+ nu anti-nu. *Phys. Rev. Lett.*, 88, 2002. arXiv:hep-ex/0111091.

[21] V. V. Anisimovsky *et al.* [E949 Collaboration]. Further study of the decay K+ → pi+ nu anti-nu. *Phys. Rev. Lett.*, 93, 2004. arXiv:hep-ex/0403036.

[22] G. Anelli. A high performance beam hodoscope for the P326 experiment at CERN. Talk at FEE 2006, Perugia, May 2006.

[23] I. Belikov. Physics of the alice experiment. 2006.

[24] Bjorken, J. D. Highly relativistic nucleus-nucleus collisions: The central rapidity region. *Phys. Rev. D*, 27(1):140–151, Jan 1983.

[25] Dmitri Kharzeev and Marzia Nardi. Hadron production in nuclear collisions at rhic and high density qcd. *Phys. Lett.*, B507:121–128, 2001.

[26] Rafelski, Johann and Müller, Berndt . Strangeness Production in the Quark-Gluon Plasma. *Phys. Rev. Lett.*, 48(16):1066–1069, Apr 1982.

[27] R. Baier, Yuri L. Dokshitzer, S. Peigne, and D. Schiff. Induced gluon radiation in a qcd medium. *Phys. Lett.*, B345:277–286, 1995.

[28] B. Alessandro [for the NA50 Collaboration]. A new measurement of J/ψ suppression in Pb-Pb collisions at 158 GeV per nucleon. *Eur. Phys. J.*, C39:335–345, 2005.

[29] Adler, C. et al. [STAR Collaboration]. Centrality Dependence of High-p_T Hadron Suppression in $Au + Au$ Collisions at $\sqrt{s_{NN}} = 130\ GeV$. *Phys. Rev. Lett.*, 89(20):202301, Oct 2002.

[30] Adler, C. et al. [STAR Collaboration]. Disappearance of Back-To-Back High-p_T Hadron Correlations in Central $Au + Au$ Collisions at $\sqrt{s_{NN}} = 200\ GeV$. *Phys. Rev. Lett.*, 90(8):082302, Feb 2003.

[31] ALICE Collaboration. ALICE Technical Design Report. Technical report, CERN/LHCC, Geneve, 12 1999.

[32] ALICE Collaboration. ALICE Technical Design Report. Technical report, CERN/LHCC, Geneve, 12 1999.

[33] CERN. http://www.cern.ch. Web Site.

[34] Alan Owens and A. Peacock. Compound semiconductor radiation detectors. *Nucl. Instr. And Meth.*, A531(1-2):18–37, 2004. Proceedings of the 5th International Workshop on Radiation Imaging Detectors.

[35] J. Kemmer. Fabrication of low noise silicon radiation detectors by the planar process. *Nucl. Instrum. Meth.*, 169:499, 1980.

[36] Andricek, L.; Lutz, G.; Moser, H.G.; Richter, R.H. Radiation hard strip detectors on oxygenated silicon. *Nuclear Science Symposium Conference Record, 2001 IEEE*, 1:573–577, 2001.

[37] Norbert Wermes. Trends in Pixel Detectors: Tracking and Imaging. *IEEE Trans.Nucl.Sci.*, 51:1006–1015, 2004.

[38] G. Hall. Semiconductor particle tracking detectors. *Rep. Prog. Phys*, 57:481–531, 1994.

[39] E. Gatti and P. Rehak. Semiconductor drift chamber - an application of a novel charge transport scheme. *Nucl. Instrum. Meth.*, A225:608–614, 1984.

[40] P. Volkov. Noise in position measurement by centroid calculation. *Nucl. Instrum. Meth.*, A373:206–212, 1996.

[41] H. Dijkstra. Overview of silicon detectors. *Nucl. Instrum. Meth.*, A478:37–45, 2002.

[42] I. Stavitski. The silicon strip detector with very large pitch: Conceptual design and simulation results. *Nucl. Instrum. Meth.*, A473:186–191, 2001.

[43] P. G. Rancoita. Silicon detectors and elementary particle physics. *J. Phys.*, G10:299, 1984.

[44] Helmuth Spieler. *Semiconductor Detector Systems.* Oxford University Press, 2005.

[45] Aull, B. et al. Laser Radar Imager Based on 3D Integration of Geiger-Mode Avalanche Photodiodes with Two SOI Timing Circuit Layers. *Solid-State Circuits, 2006 IEEE International Conference Digest of Technical Papers*, pages 1179– 1188, 2006.

[46] Ray Yarema. 3D Integrated Circuits for HEP. 6th Int. Meeting on Front End Electronics, Perugia, May 2006.

[47] Kenney, C.; Parker, S., Segal; J.; Storment, C. Silicon detectors with 3-D electrode arrays: fabrication and initial test results. *Nuclear Science, IEEE Transactions on*, 46(4):1224–1236, 1999.

[48] P. O'Connor. Analog front ends for highly segmented detectors. *Nucl. Instrum. Meth.*, A522:126–130, 2004.

[49] De Geronimo, G.; O'Connor, P. MOSFET optimization in deep submicron technology for charge amplifiers. *Nuclear Science, IEEE Transactions on*, 52(6):3223–3232, 2005.

[50] G. Hall. Recent progress in front end ASICs for high-energy physics. *Nucl.Instr. and Meth.*, A541:248–258, 2005.

[51] A. Mekkaoui and J. Hoff. 30mrad(sio-2) radiation tolerant pixel front end for the btev experiment. *Nucl. Instrum. Meth.*, A465:166–175, 2000.

[52] Dennard, R.H.; Gaensslen, F.H.; Rideout, V.L.; Bassous, E.; LeBlanc, A.R. Design of ion-implanted MOSFET's with very small physical dimensions. *Solid-State Circuits, IEEE Journal of*, 9(5):256–268, 1974.

[53] Cunha, A.I.A.; Schneider, M.C.; Galup-Montoro, C. An MOS transistor model for analog circuit design. *Solid-State Circuits, IEEE Journal of*, 33(10):1510–1519, 1998.

[54] Manfredi, P.F.; Millaud, J.E.; Sushkov, V.V. . Acquisition and processing of multiparametric information from a pixel matrix . *Nuclear Science, IEEE Transactions on*, 48(4):1244–1248, 2001.

[55] Paul O'Connor. Future trends in microelectronics: Impact on detector readout. *ECONF*, C0604032:0052, 2006.

[56] A. Rashevsky et al. Large area silicon drift detector for the ALICE experiment. *Nucl. Instr. And Meth.*, A485:54–60, 2002.

[57] D. Lynn et al. A 240-channel thick film multi-chip module for readout of silicon drift detectors. *Nucl. Instrum. Meth.*, A439:418–426, 2000.

[58] G. Anelli. *Design and characterization of radiation tolerant integrated circuits in deep submicron CMOS technologies for the LHC experiments*. PhD thesis, Inst. Nat. Polytechnique de Grenoble, France, 2000.

[59] S.Beole et al. Production and assembly of the ALICE silicon drift detectors. 2006. in press.

[60] L.Toscano et al. Mass production testing of the front-end ASICs for the ALICE SDD system. In *Proceedings of 11th Workshop on Electronics for LHC and Future Experiments*, Heidelberg, 2005.

[61] A. Rivetti et al. The front-end system of the silicon drift detectors of ALICE. *Nucl. Instr. And Meth.*, A541:267–273, 2005.

[62] Blanquart, L. and others. FE-I2: A front-end readout chip designed in a commercial 0.25 μm process for the ATLAS pixel detector at LHC. *IEEE Trans. Nucl. Sci.*, 51:1358–1364, 2004.

[63] P. F. Manfredi, A. Leona, E. Mandelli, A. Perazzo, and V. Re. Noise limits in a front-end system based on time-over- threshold signal processing. *Nucl. Instrum. Meth.*, A439:361–367, 2000.

[64] G. Faccio, F.; Cervelli. Radiation-induced edge effects in deep submicron CMOS transistors . *Nuclear Science, IEEE Transactions on*, 52(6):2413–2420, 2005.

[65] Jackson, R.G.; Blalock, T.V.; Simpson, M.L.; Wintenberg, A.L.; Young, G.R. Integrated constant-fraction discriminator shaping techniques for the PHENIX lead-scintillator calorimeter . *Nuclear Science, IEEE Transactions on*, 44(3):303–307, 1997.

[66] C. H. Nowlin. Low-Noise Lumped-Element Timing Filters with Rise-Time Invariant Crossover Times. *Rev. of Sci. Instr.*, 63(4):2322–2326, 1992.

[67] Simpson, M.L.; Young, G.R.; Jackson, R.G.; Xu, M. . A monolithic, constant-fraction discriminator using distributed R-C delay line shaping . *Nuclear Science, IEEE Transactions on*, 43(3):1695–1699, 1996.

[68] D. M. Binkley, B. S. Puckett, B. K. Swann, J. M. Rochelle, and M. S. Musrock. A 10-Mcps, 0.5-μm CMOS Constant-Fraction Discriminator Having Built-In Pulse Tail Cancellation. *Nuclear Science Symposium Conference Record*, 1:151–155, 2001.

[69] Claudio Piemonte. Time response simulation of the NA48 silicon pixel detector. Flattening of the wafers. Gigatracker Working Group, Nov. 2004.

[70] Behzad Razavi. *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, 2001.