Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by:
Dipl.-Inf. Florian Erdinger
born in Mannheim, Germany
Oral Examination: November 22, 2016

Design of Front End Electronics and a Full Scale 4k Pixel Readout ASIC for the DSSC X-ray Detector at the European XFEL

Referees: Prof. Dr. Peter Fischer
Prof. Dr. Johanna Stachel

Zusammenfassung:

Das Ziel dieser Arbeit war es, einen großformatigen Auslese-ASIC für den 1-Mega Pixel DEPFET Sensor with Signal Compression (DSSC) Detektor zu entwickeln, der am European XFEL (EuXFEL) zum Einsatz kommen wird. Die Anforderungen an den Detektor beinhalten die Auflösung einzelner Photonen bis zu einer minimalen Energie von 0.5 keV kombiniert mit einem großen dynamischen Bereich von bis zu 10000 Photonen bei einer maximalen Bildrate von 4.5 MHz. Die Kernkonzepte des Detektors beinhalten Signalkompression auf Sensorebene, sofortige Digitalisierung und lokale Speicherung im Pixel. Das DSSC System ist ein hybrides System, jedes Sensor Pixel ist mit einem entsprechenden Auslesepixel auf dem ASIC verbunden, der die komplette beschriebene Signalverarbeitungskette enthält. Auf dem ASIC befinden sich 4096 Pixel, weitere Peripherieblöcke und Ansteuerungslogik. Die Entwicklung des ASIC wird beschrieben, seine Komponenten und deren Integration erklärt. Entwicklungen für das analoge Front-End werden speziell hervorgehoben. Der erste vollformatige ASIC wurde, neben zahlreichen weiteren Testchips, im Rahmen dieser Arbeit fertiggestellt. Der EuXFEL und das DSSC Detektor System werden präsentiert um die Rahmenbedingungen für den ASIC darzulegen, der das Kernthema dieser Arbeit darstellt.

Abstract:

The goal of this thesis was to design a large scale readout ASIC for the 1-Mega pixel DEPFET Sensor with Signal Compression (DSSC) detector system which is being developed by an international collaboration for the European XFEL (EuXFEL). Requirements for the DSSC detector include single photon detection down to 0.5 keV combined with a large dynamic range of up to 10000 photons at frame rates of up to 4.5 MHz. The detector core concepts include full parallel readout, signal compression on the sensor or ASIC level, filtering, immediate digitization and local storage within the pixel. The DSSC is a hybrid pixel detector, each sensor pixel mates to a dedicated ASIC pixel, which includes the entire specified signal processing chain along with auxiliary circuits. One ASIC comprises 4096 pixels and a full periphery including biasing and digital control. This thesis presents the design of the ASIC, its components and integration are decribed in detail. Emphasis is put on the design of the analog front-end. The first full format ASIC (F1) has been fabricated within the scope of this thesis along with numerous test chips. Furthermore, the EuXFEL and the DSSC detector system are presented to create the context for the ASIC, which is the core topic of this thesis.

Contents

1

Introduction

The evolution of light sources for scientific experiments continues to challenge technology and designers to provide suitable detector instruments. The European XFEL (X-Ray Free Electron Laser, EuXFEL) is a light source of the 4th generation which will deliver femto second short flashes of monochromatic X-rays down to the Ångstrom wavelength. The brilliance of the radiation generated at the European XFEL will surpass 3rd generation synchrotron sources by orders of magnitude. These exciting properties will provide unprecedented atomic space and time resolution which will revolutionize methods in the field of material science. However, it will also require the development of new 2-D imaging detector concepts.

The experiments at the EuXFEL require challenging detector properties requirements which include good spacial resolution covering large areas, single shot imaging at an event rate of up to 4.5 MHz, single photon detection capability down to 0.5 keV with a signal-to-noise ratio of 1/5, while at the same time preserving a dynamic range of up to 10^4 photons per pixel per image. The electronics noise expressed in equivalent noise charge (ENC) should therefore reach a level of down to $28\,e^-$. The detectors must further be compatible to vacuum operation. The combination of all these properties in large area detectors is unprecedented and requires the development of novel concepts.

The DSSC (**D**EPFET (Depleted Field Effect Transistor) **S**ensor with **S**ignal **C**ompression) is one of the detectors being developed for the needs at the European XFEL. A consortium of several partner institutes including DESY (Hamburg, Germany), Politecnico di Milano (Italy), University of Bergamo (Italy), the European XFEL GmbH (Germany) and Heidelberg University (Germany) to the development of the project. The DSSC is a 1 Mega-pixel camera which covers an area of $21 \times 21\,\text{cm}^2$, an artistic illustration of the final system is shown in figure 1.1. The system is mainly tailored for low energy X-ray experiments, which require very low noise performance. In the DSSC concept, each sensor pixel is bump bonded to a dedicated readout channel on an application specific integrated circuit (ASIC) enabling a full parallel readout of the sensor to cope with the fast event rate (4.5 MHz). The ASIC pixel comprises a signal processing chain of an analog filter, an ADC (analog-to-digital converter) and a digital memory to locally store the signals within the pixels along the burst. The memories are read out during 100 ms gaps between the XFEL bunch trains. A novel DEPFET sensor has been chosen as the central detection device. DEPFET sensors possess extraordinary properties for low noise applications while a matrix of such devices can be read out in parallel. The DEPFET has been expanded with a novel signal compression mechanism - hence the project name DEPFET Sensor with Signal Compression. This mechanism provides the required dynamic range capability by providing a

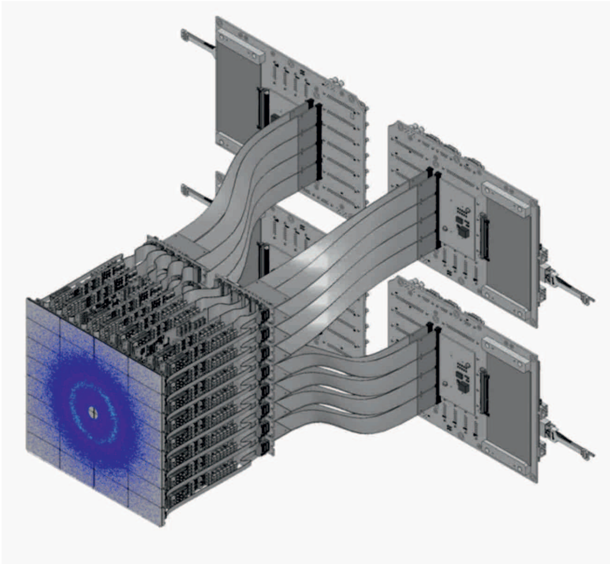


Figure 1.1: Artistic view of the final 1 mega-pixel DSSC system covering an area of $21 \times 21\,\mathrm{cm}^2$. Courtesy of [1].

nonlinear system characteristic, compressing larger signals such that the complete dynamic range can be digitized with 8-9bits in the ASIC. While the low noise performance of linear mode DEPFETs is well known, successful prototypes have shown that it can be combined with a suitable compression mechanism on the sensor level.

This work presents the design of the readout ASIC, the first full format version is shown in figure 1.2 and has been fabricated during the course of this thesis. It is a central element of the DSSC system. The ASIC pixel building blocks of an analog filter and an ADC circuit have been provided by external research groups from Politecnico di Milano and DESY, respectively. Our work, which is described here, has been pixel and full chip integration and verification, starting from small test matrices leading to the first full format 64×64 pixel chip, including the integration of all pixel electronics, the design of a suitable in-pixel memory and the design of a global on-chip digital control block. The DSSC readout ASIC is a *system on chip*, integrating required control logic for the pixel electronics and readout structure are integrated. The interface is made up of a JTAG interface handling slow control and a two wire protocol to handle the dynamic control, while the output data is serialized on a single fast output link. The full format ASIC size is $\approx 15 \times 15\,\mathrm{mm}$, 256 of these ASICs will be mated to 64 sensor dies for the full mega-pixel camera. Several test chip iterations have finally led to the submission and fabrication of an engineering run of the first full format ASIC. It is completely functional and currently being characterized while an extensive R&D phase is in progress for a second version.

Furthermore, this thesis also presents significant work which was contributed by the author to the analog front-end of the pixel electronics for the readout of an alternative sensor. While the DEPFET sensor has been the heart of the system from the early planning phase, a second development track had to be opened approximately one year before the F1 ASIC was scheduled to be submitted for production. The first fabrication of DSSC sensors was projected to take approximately two years, however an unforeseen problem late in the production cycle essentially voided most of this production and furthermore caused doubt about the final availability of sufficient DEPFET sensors. In addition,

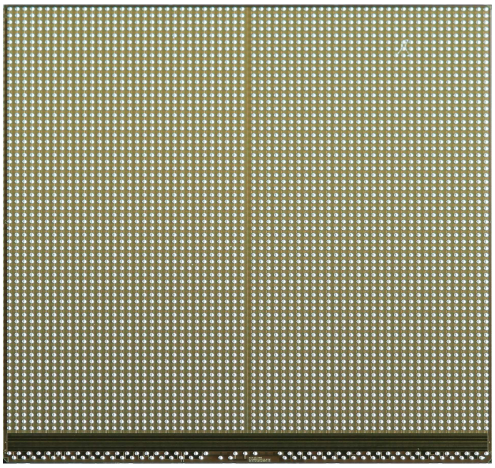


Figure 1.2: Die photograph of the full size $14.9 \times 14.0\,\mathrm{mm}^2$ 4k pixel chip.

the collaboration had to be restructured in this period due to political issues summing up to significant time delays. The unavailability of the DEPFET triggered an extensive R&D phase, which had initially not been foreseen causing a shortage in available manpower. A solution had to be found to provide a functional detector system for the commissioning phase at the European XFEL. The active DEPFET was eventually replaced by a passive linear mini silicon drift detector (MSDD). Consequently, the ASIC had to be equipped with a suitable front-end. Due to personnel shortcomings, the scope of this thesis has been further expanded and also includes analog front end design.

This thesis is structured in a total of seven chapters following this introduction. Chapter two briefly explains the principles of X-ray free electron lasers and introduces the European XFEL to establish the context for the various detection instruments and their requirements. The third chapter contains fundamentals on the two sensor variants used in the DSSC detector system, some important MOSFET properties and signal processing fundamentals which are essential to understand the DSSC system. Chapter four describes the DSSC detector system in detail. The next three chapters focus on the volume contributed to the project: in chapter four, the ASIC design is described in detail while chapter five presents some analog front-end design studies for improvements on the readout of the MSDD. Chapter seven shows some selected measurement on the F1 ASIC and other test structures. Conclusions and an outlook are given in chapter eight.

2

The European XFEL

To put the DSSC detector into context, this chapter deals with the working principle of free electron lasers in general and more specifically introduces the reader to the European XFEL facility. Furthermore, some general background information on the evolution of X-ray sources is provided and the use of X-rays in the scientific world is outlined. This chapter has in general been compiled using the following sources: [2], [3], [4] and [5] while additional sources are referenced in the text.

2.1 Background & Motivation

When Roentgen discovered a new kind of radiation in 1895, he named it X-rays, the *X* representing something unknown as in a mathematical equation. While we have by now a very good understanding of the kind of radiation he discovered, the name has preserved and the *X* has found its way into the name of one of the world's most powerful light sources, the European XFEL. While more than a century has passed since their discovery, the methods of generating and using X-rays has consequently evolved. While they were discovered with a simple discharge tube, today's most advanced sources are huge enterprises, which make use of gigantic electron accelerators and employ hundreds of people. The useful properties of X-rays have been exploited from their discovery, mainly in the field of medical imaging and material science. Today they are an important tool in most fields of applied science. While their short wavelengths is an indispensable tool to probe matter at nanoscale levels, insufficient coherence and rather long pulse durations are a crucial obstacle for many experiments. Combining X-rays with lasing properties has long been a dream of scientists. This feat has been accomplished by the development of X-ray free electron lasers (XFEL). Their general working principle is described in section 2.3.

A very critical figure of merit for many scientific experiment is the so called brilliance (or spectral brightness), which is usually expressed in the units

$$brilliance = \frac{number\ \ of\ \ photons}{s \times mm^2 \times mrad^2 \times 0.1\% BW}$$

The expression $0.1\%BW$ denotes a bandwidth of 10^{-3} around the central frequency ω of the beam. The brilliance is hence maximized if the photon flux, i.e. the number of photons per second in a given bandwidth is large and the photons are emitted very densely from a small spot with little angular divergence.

The combination of extreme brilliance, ultra short ($< 100\,\text{fs}$) pulse length at a repetition rate of up to 4.5 MHz and spacial coherence down to Ångstrom wavelengths will become available at the European X-Ray Free Electron Laser which is being built in Hamburg, Germany. The short pulse duration and fast pulse repetition rate sets the European XFEL apart from similar machines in the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory in Stanford, USA [6] and the SPring-8 Angstrom Compact Free Electron Laser (SACLA) [7] in Hyogo, Japan which are already in operation.

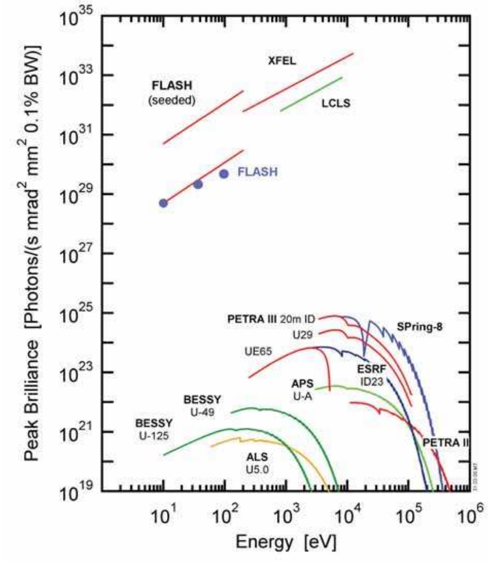


Figure 2.1: Peak brilliance of XFELs versus third generation synchrotron light sources [3].

The main ingredient of an XFEL is a relativistic electron beam of very high quality (the reasons are outlined in section 2.3.1 and section 2.3.2). The construction of highly brilliant XFELs has been made possible by recent progress in particle accelerator and electron injection technologies. Progress in this field has been moderate until the discovery of synchrotron radiation, which has originally been discovered as a parasitic side effect in electron storage rings intended for particle collision. Electrons emit radiation when they are accelerated radially and consequently lose energy. It was however soon realized that this mechanism can serve as an efficient X-ray source, paving the way for new scientific opportunities. These storage rings, which were not dedicated to produce radiation, are referred to as 1st generation synchrotron sources. For the 2nd generation, the properties of the machines were optimized to produce radiation. 3rd generation devices started to employ straight sections to insert wigglers and undulators (see section 2.3.1), which are special devices directing electrons on slalom paths. These devices already boosted the brilliance by four orders of magnitude compared to the bending magnets used in earlier generations. However, the limiting factor to produce even more brilliant radiation is the quality of the electron beam confined in a storage ring. In a storage ring, the geometrical emittance increases with the square for any particular magnet lattice due to quantum fluctuations. In a linear accelerator (linac) however, the normalized emittance is a conserved quantity, i.e. the geometrical emittance decreases linearly with the energy Consequently, linacs started to gain recognition for 4th generation light sources. Progress in linear accelerator technologies has mainly been triggered by collider experiments. In 1992 Pellegrini [8] showed, that with technology available at that time, it was possible to construct a 4 nm FEL, while further improving the wavelength to 0.1 nm with moderate improvements in the electron injection device is possible. His proposal eventually led to the construction of the Linac Coherent Light Source (LCLS). The LCLS uses the Stanford Linear Accelerator (SLAC) equipped with new electron injection devices and special undulators to generate brilliant X-ray radiation. The machine has been commissioned in 2009.

In 2000 the proposal was made to build the Tera-Electronvolt Superconducting Accelerator (TESLA) in Hamburg, Germany, an electron positron collider. Due to the expected superior quality of the electron beam a proposal was included for an X-ray free electron laser laboratory making use of the collider linear accelerator. Several test facilities finally lead to the construction of the soft X-ray FLASH facility (Free Electron Laser in Hamburg) [9] which was commissioned in 2005. The TESLA proposal was finally dismissed, but it was decided to build a dedicated linear accelerator for an X-ray free electron laser, paving the way for the European XFEL (EuXFEL).

2.2 X-rays in science

Since their discovery in 1895, X-rays have proven to be indispensable in most fields of applied science. Five scientific methods have evolved which are based on X-rays:

(1) **X-ray Imaging:**
First demonstrated by Wilhelm Röntgen, X-ray imaging is widely known from its medical applications. More recent experiments have produced high resolution images on the nanometer scale revealing the charge and spin distributions of a sample. Röntgen won a Nobel prize 1901 for the discovery of X-rays.

(2) **X-ray Diffraction:**
The method is widely used to investigate the nature of crystalline structures. Interatomic bonding distances and angles can thereby be revealed. Max von Laue was rewarded with the Nobel prize in 1914 for *his discovery of of the diffraction of X-rays by crystals*. Sir William Henry Bragg and William Lawrence Bragg followed with another Nobel prize related to this method in 1915 *for their service in the analysis of crystal structure by means of X-rays*.

(3) **X-ray Absorption and Emission Spectroscopy:**
These two methods are used to reveal the structure of the electronic shells around the atomic nucleus and electronic energy bands in solids. For absorption experiments, the X-ray energy needs to be finely tuned in order to excite core electrons in the atomic shell. In emission experiments, core electrons are excited by incident radiation and the resulting emission spectrum is recorded. As the excited electrons fall back into the lower energy state, the substance under study emits a spectrum which is characteristic for its nature. Charles Glover Barkla was awarded a Nobel prize in 1917 *for his discovery of the characteristic Röntgen radiation of the elements*.

(4) **Inelastic X-ray Scattering:**
This discipline makes use of the Compton effect, which is named after Arthur Holly Compton who earned the (shared) Nobel Prize in 1927 for its discovery. When photons scatter inelastically off a charged particle, the scattered photons are different in wavelength than the incident photons. This method was later used to measure collective excitations and vibrational elastic properties of matter and the magnetic properties and valence states of ions.

(5) **Photoelectron Spectroscopy:**
The photo electric effect lets a material emit an electron, if the electron acquires more energy

than than is required to bind it in the material by absorbing a photon. Using this method, the structure of bondings in molecules and solids has been revealed. While Robert Milikan was rewarded with the Nobel prize on *his work on the elementary charge of electricity and on the photoelectric field*, Kai Siegbahn pioneered photoelectron spectroscopy starting in 1957 and was awarded a (shared) Nobel prize in 1981 for *his contribution to the development of high-resolution electron spectroscopy*.

A total of 19 Nobel prizes have been awarded to related work since their discovery, undermining their importance for science. All these fields will benefit enormously by the unprecedented brilliance becoming available by XFELs. The European XFEL will further provide unique pulse rates (see figure 2.6), combining the possibility to study tiny structures and processes with an unequaled timing resolution.

2.3 The Free Electron Laser

This section introduces the reader to free electron lasers, including the mechanisms to generate X-rays and the reasons why superior brilliance can be achieved.

2.3.1 General Principle of a High Gain FEL

Although the word *laser* has evolved into a proper word, it is an acronym for *light amplification by stimulated emission of radiation*. In a classical quantum laser, radiation is generated by atomic bound excited electrons transitioning to a lower energy state. The electrons are bound to discrete energy levels in the shell of atoms.

A free electron laser is actually closer related to a vacuum tube, where the electrons emitting radiation are not bound to atoms (hence free electron laser), and the underlying principle of operation is an interaction between an electron beam and electromagnetic radiation. The two main devices comprising an FEL are an electron accelerator and an undulator (depicted in figure 2.2). Electrons are accelerated close to the speed of light and then directed through the undulator, which is a device consisting of a periodic arrangement of alternating magnets. The alternating magnetic field forces the passing electrons on a slalom path. Due to the change of direction on their slalom path, the electrons emit electromagnetic radiation. While this general principle is quite simple, the implementation is difficult especially when targeting the X-ray regime. The properties of the emitted radiation can be tailored through various parameters.

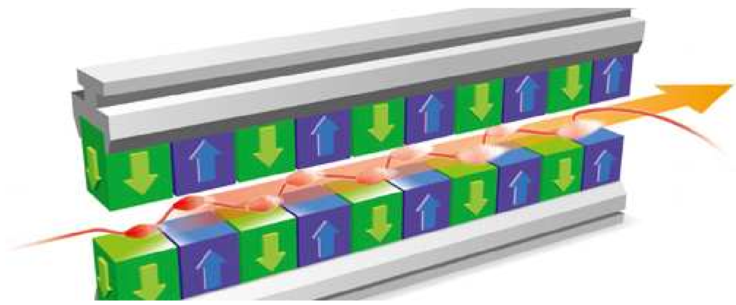


Figure 2.2: An undulator consists of a periodic arrangement of alternating magnets which forces electrons to change their direction and hence emit radiation [2].

The electromagnetic wave travels always faster than the electrons. In the undulator, a resonance condition is fulfilled when the wave slips ahead of the electrons by one wavelength. The resonance frequency of the undulator radiation is given by

$$\lambda = \frac{\lambda_u}{2\gamma^2}(1 + K^2), \tag{2.31}$$

where K is the so called undulator parameter and depends on the composition of the undulator:

$$K = \frac{eB_0\lambda_u}{2\pi mc^2}, \tag{2.32}$$

γ is the Lorentz factor and λ_u is the undulator period. At this frequency, a resonance condition is fulfilled, where a continuous energy from the moving electrons to the wave train takes place.

Since the emitted wavelength depends on the parameters of the undulator and the energy of the electron beam, the resonance frequency can be chosen quite freely and the emitted wavelength hence tuned. Fine tuning can for instance be done by tuning the magnetic gap[1] in the undulator. This is a characteristic which fundamentally distinguishes an FEL from a quantum laser, in which the wavelength of the emitted radiation is defined by the discrete energy transitions in the active medium.

The total intensity seen at the output of the undulator depends on how the waves emitted by the individual electrons in the beam superimpose. Consider a number of N_e electrons at relativistic speed confined in a bunch with length L entering an undulator. A large bunch length L implies a difference in time of when the individual electrons enter the undulator which translates to a phase difference of the emitted radiation. If the bunch length L is long with respect to the radiation wavelength and the longitudinal distribution random with respect to the wavelength, the intensity of the radiation is only proportional to the number of electrons ($I \propto N_e$). If L is however very short, the electron bunch can be treated as a single particle of charge N_e which radiates coherently with an intensity proportional to the square of the number of electrons ($I \propto N_e^2$). Since a typical bunch of electrons can contain on the order of 10^9 to 10^{10} electrons, the difference in intensity is enormous. Even with state of the art technologies, it is not possible to confine such a large number of electrons within a single bunch at X-ray wavelengths. However, a third condition exists, which can approximate the condition imitating a single radiating super particle. Imagine the N_e electrons confined in micro-bunches which are separated in distance by one wavelength of the electromagnetic waves. While this condition cannot produce an intensity of N_e^2, some intermediate power between one and two is feasible. The so-called bunching factor B quantifies this process and is given by

$$B = \sum_{n=0}^{N_e} \frac{1}{N_e} e^{z_{n0} 2\pi i \lambda} \tag{2.33}$$

where z_{n0} is the initial position of an individual electron in the bunch. The intensity of the bunching is calculated by $|B|^2$. $|B|^2 = 0$ implies that the electrons are distributed randomly while for $|B|^2 = 1$ the bunching is at its maximum where either all electrons are grouped very closely, or micro-bunches are separated by λ.

The forming of micro bunches can be reached in an undulator of sufficient length by an interaction of the electromagnetic wave and the electron beam. As radiation is generated in the undulator, it interacts with electrons ahead, accelerating and decelerating them such that they bunch in disks which are one wavelength apart. Considering the following three steps, this mechanism is instable:

[1]distance between alternating magnets perpendicular to the direction of flight of the electrons

(1) As the electromagnetic waves are always faster than the moving electrons, it interacts with electrons ahead in the undulator, modulating their energy.

(2) Due to this energy modulation, the electrons group in longitudinally confined bunches.

(3) The bunches formed in (2) lead to a higher intensity of the electromagnetic field, which feeds positively into (1) and generates an instability.

After some distance in the undulator, the electrons start to bunch at the same phase with respect to the electromagnetic field. The instability provides an exponential amplification of the radiation intensity along the undulator and eventually saturates when the micro bunches are fully formed. In this case, the beam energy has decreased such that the resonance condition is no more satisfied and the intensity growth hence stops. The simplified one-dimensional model gives a very good view of an FEL, the FEL parameter ρ has been introduced which describes all the phenomena of interest. It is given by:

$$\rho = (\frac{K}{4}\frac{\Omega_P}{\omega_u})^{2/3} \tag{2.34}$$

where K is the undulator parameter (2.31), $\Omega_P = 4\pi n_e r_e c^2/\gamma^3$ is the beam plasma frequency, n_e is the electron bunch density and $\omega_u = 2\pi/\lambda_u$. Using this definition, the exponential growth of the radiation power P_L is given by:

$$P_L = P_0 e^{z/L_G} \tag{2.35}$$

where

$$L_G = \lambda_u/4\sqrt{\pi}\rho \tag{2.36}$$

is the gain length and z is the distance along the undulator axis. The saturation power is given by

$$P_S = \rho E I_P \tag{2.37}$$

where E is the electron beam energy and I_P its peak current. In order for the exponential growth to occur, three conditions need to be satisfied:

(1) the energy spread in the electron beam must be smaller than the gain bandwidth,

(2) the radius and angular divergence of the electron and photon beam must match so that an interaction between photons and electrons can take place; in accelerator physics terminology the quantities used here are normalized emittances

(3) the diffraction losses from the radiation beam must be smaller than the FEL gain.

Free electron lasers can generally be divided into two classes, amplifying devices and oscillators. In the amplifier configuration, the electron beam passes the undulator only once, and the device amplifies radiation which is either seeded by an external wave or spontaneously emitted by the electron beam. In FEL oscillators, feedback between the input and the output is applied with an optical resonator. Oscillating FELs have their limit in the ultraviolet regime, primarily limited by the required mirrors. To achieve lasing at higher energies, single-pass amplifying devices need to be used. In the X-ray regime, the self-amplified spontaneous emission phenomenon (SASE, see section 2.3.2) is employed because proper seed radiation is not available.

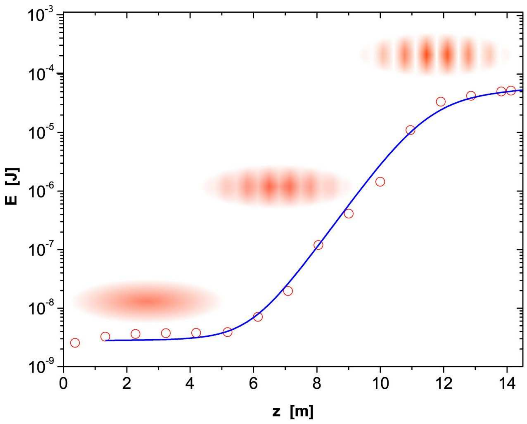


Figure 2.3: Exponential gain growth along an undulator as measured (red circles) at the SASE FEL at DESY [10]. The solid blue line shows the theoretical prediction and the microbunching effect is schematically indicated.

2.3.2 Generation of X-rays in an FEL (SASE)

To generate X-ray radiation, the FEL needs to be configured in a single-pass, high-gain amplifying mode. A critical ingredient to obtain X-rays from an FEL is the self amplification by stimulated emission (SASE) phenomenon first reported in [11]. Due to the lack of sufficient seed radiation at wavelengths in the X-ray regime, a different activation technique for the amplification process is needed. Spontaneous emission in the first part of a long undulator can serve as a seed for the amplification process described in the previous section. This process allows to seed arbitrary wavelengths because no external seed radiation is needed. In order to trigger the instability, some level of density fluctuation is required in the electron beam entering the undulator. The electron bunches coming from the accelerator are initially randomly distributed and inherit a white noise spectrum. Consequently, there is a spectral component which is in the undulator bandwidth and will be amplified as explained in the previous section.

In order to trigger the SASE process, an electron beam of superior quality is required. The beam needs to have a very small cross section, a high charge density and a low energy spread which are characteristics that only linear accelerators can provide.

From 2.31, one may derive, that to move to shorter wavelengths, it suffices to increase the energy of the electron beam. However, the emittance condition ((2)) is very hard to fulfill at small wavelengths. As the Lorentz factor of the electrons is increased, the emittance of the beam shrinks with $1/\gamma$ while the light wavelength shrinks with $1/\gamma^2$. The situation can be eased through tuning the undulator by increasing the undulator period λ_u and / or the undulator parameter K. The side effects are, however, that even larger beam energies are required and the undulator length must be considerably increased because the gain length increases. For the generation of hard X-rays at the LCLS and EuXFEL for instance, very long undulators ($> 100m$) are needed. Any progress in reducing the normalized emittance in the electron beam would simplify the layout of an FEL facility because the beam energies could be reduced to reach the same wavelengths. At the same time, existing powerful linear accelerators could then be used to generate even shorter wavelengths.

2.4 The European XFEL

Figure 2.4: Aerial view of the entire European XFEL facility [2].

2.4.1 The Accelerator

An international collaboration coordinated by the Deutsches Elektronen Synchrotron (DESY) in Hamburg has developed a superconducting linear accelerator technology which was originally intended for an electron-positron linear collider with TeV energy. It was soon realized that this Tera-Electronvolt Superconducting Accelerator (TESLA) has ideal characteristics to build an X-ray free electron laser, leading the way for the construction of the European XFEL in Hamburg, Germany. The superconducting technology makes possible the unique characteristic to generate 2700 flashes trains at a rate of 4.5 MHz. The bunch trains can be repeated at a rate of 10 Hz, summing up to 2700 flashes per second. For comparison, the LCLS does not make use of superconduction in its accelerator and can only produce 120 flashes per second. It is out of scope for this text to explain the physics of the accelerator. However, its capabilities and composition are reported to give the reader an idea of the dimension of the facility.

The accelerating elements are superconducting cavities powered by a radio frequency (RF) system running at a frequency of 1.3 GHz. The RF system is operated in pulsed mode at a repetition rate of 10 Hz due to power limitations. A single RF pulse lasts 600 μs, which is hence also the duration of the flash bunch trains generated in the undulators. Subsequent bunch trains are hence spaced apart in time by nearly 100 ms. Superconduction in the cavities allows for this high duty cycle and a nearly lossless energy transfer from the RF wave to the electron beam.

Takeoff for the electrons begins as they are extracted in bunches of 1 nC from a solid cathode by a laser beam. The laser is operated at a maximum frequency of 4.5 MHz which gives the maximum frequency within a bunch train. Next, the electrons are accelerated by an electron radio frequency gun and injected into the first acceleration stage. Before they enter the main accelerator, they pass through two bunch compression stages, in-between which they traverse a further intermediate acceleration stage.

The compression stages shorten the initially 2 mm long bunches by a factor of 100 and increase the peak current in the bunch to 5 kA. These compression stages are essential to generate the density required to trigger the SASE process to generate radiation in the X-ray regime in the undulators preceding the accelerator. The final and longest acceleration stage takes the electrons from 2GeV to the final nominal energy of 20GeV.

The main accelerator has been designed that it is tunable in a wide range. Since the undulator radiation wavelength depends on the electron energy (see also section 2.3.1), the accelerator is designed such that the acceleration gradient can be changed and hence the radiation wavelength tuned. The crucial first stages including the two compression stages remain untouched to preserve the quality of the electron beam even if the energy is reduced. The dimensions of the undulators for the hard X-ray experiment stations foresee to use a beam energy of 17.5 GeV. It is expected that the system can even produce higher energies than the nominal 20 GeV, there is hence substantial reserve to further shorten the wavelength beyond the target of 0.1 nm. The built in reserve can also be used to increase the repetition rate to up to 30 Hz at 17.5 GeV.

2.4.2 Undulators & Beamlines

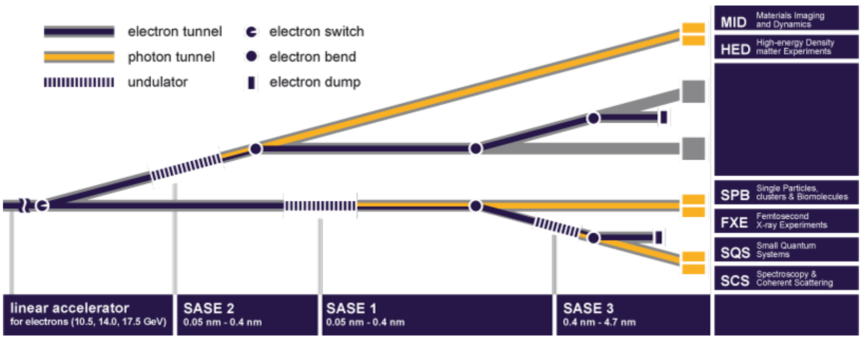


Figure 2.5: Layout of the photon beamlines at the European XFEL [2]. The electron beam from the accelerator can be switched to two alternative paths where they traverse different undulator arrangements to provide photon beams at various energies ranging from 0.25 keV to 12.4 keV.

When the electrons have reached their final energy and leave the accelerator, they are next directed through different undulator arrangements to generate X-rays. There are a total of five beamlines foreseen: three use very long undulators (saturation lengths 81-174m), which generate radiation through the SASE process (SASE1-3) while two further shorted undulators (U1-2) are installed which provide spontaneous synchrotron radiation (no SASE) at energies of 15 keV - 90 keV.

SASE1 is a fixed gap undulator which provides a fixed energy of 12.4 keV with respect to the electron beam energy. SASE2 provides energies of 3.1 keV – 12.4 keV while SASE3 uses the electron beam spent by SASE1 and provides soft X-rays the energy range of 0.25 keV – 3.1 keV.

2.4.3 The Scientific Case

The achieved brilliance of the photon beams at the European XFEL and the ultra fast pulse rate will enable scientist to perform new X-ray techniques. The beamlines have been optimized thoroughly by consulting the scientific community. While it is out of scope here to present all the new possibilities, some general examples are given to highlight the value of the machine for the scientific community.

Nano structures will be able to be studied at unprecedented spacial resolution. Up to today, the structure of no genome has been revealed in high resolution. Details of, for example, the assembly, stability and disassembly of viruses remain ununderstood. These processes can be studied at new levels of precision using the EuXFEL. Experiments here will further allow to reconstruct the genomes of viruses. Viruses which cannot be crystallised as for instance HIV (Human Immunodeficiency Virus) and HSV (Herpes Simplex Virus) are of special interest as diffraction experiments will be possible without the need for crystallisation. Progress is expected in understanding how infections with viruses occur which will hopefully help to find ways to interfere with infections.

The short pulse widths of 100 fs and fast rate of 4.5 MHz combined with ultra short wavelengths allows to make *movies* at unprecedented time resolution providing new means to study ultra fast processes at the nano scale. Filming chemical reactions in real time has long been a dream of scientists, the European XFEL will offer groundbreaking possibilities here. Molecular motion and phonons in a large class of systems will become detectable. Pump and probe experiments are made possible, which reveal the changes in the structure of proteins. New catalysts could be developed by improving the understanding of chemical reactions.

Six scientific instruments (layout depicted in figure 2.5) will be available optimized for different purposes and are hence attributed to different beamlines [2]:

- **SPB/SFX:** (Single Particles, clusters, and Biomolecules and Serial Femtosecond Crystallography): Structure determination of single particles, atomic clusters, biomolecules, virus particles, cells.
- **FXE:** Femtosecond X-ray Experiments, SASE1: Time-resolved investigations of the dynamics of solids, liquids, gases.
- **MID:** Materials Imaging and Dynamics, SASE 2: Structure determination of nanodevices and dynamics at the nanoscale.
- **HED:** High Energy Density Matter, SASE 2: This instrument will be a new, unique platform for experiments combining hard X-ray FEL radiation and the capability to generate matter under extreme conditions of pressure, temperature or electric field using the FEL, high energy optical lasers, or pulsed magnets.
- **SQS:** Small Quantum Systems, SASE 3: Small quantum systems: investigation of atoms, ions, molecules and clusters in intense fields and non-linear phenomena.
- **SCS:** Spectroscopy and Coherent Scattering, SASE 3: Spectroscopy & Coherent Scattering: Electronic and atomic structure and dynamics of nanosystems and of non-reproducible biological objects using soft X-rays.

2.5 Novel Detectors for the European XFEL

While the various scientific instruments will make use of a variety of detectors, the unique timing structure and dynamic range requirements of the various experiments requires novel concepts. After presenting the requirements an overview of the three main detector developments is given.

2.5.1 X-ray Detector Requirements

The requirements for 2D X-ray detectors at the European XFEL have been specified in [12] and are summarized in this section. The scientific experiments mostly require high space resolution detectors, and the capability to detect the number of incident photons in each pixel. For larger photon numbers this requirement is relaxed by Poisson statistics. The detectors need to have a large dynamic range, they should have the capability to detect both single photons and up to 10^4 photons per pixel per image. A single detector system covering the wide energy range at the three different X-ray beamlines of 0.25 keV up to 12.4 keV cannot be optimized sufficiently. Considering the large energy range of a factor of ≈ 50, a system optimized for single photons at 0.25 keV would be completely overloaded with the assignment to process several thousand 12 keV photons. Vacuum compatibility is required for the low energy beamlines because they are operated windowless to preserve the beam quality. The entrance windows on low energy detectors also require special properties such that the X-ray absorption in the surface layer is minimal.

Consequently, three different detector systems are being developed, which focus on different energy ranges. This also eases further constraints, as for instance for the lower energy detector systems, radiation hardness in the processing electronics is of little concern since almost no radiation is propagated through a sensor of standard thickness ($\sim 300\,\mu\mathrm{m} - 500\,\mu\mathrm{m}$). At higher energies, a significant portion of radiation will propagate through the sensor requiring appropriate design techniques in the readout electronics.

In many of the experiments, the sample under study will be completely destroyed by the intense and short XFEL flash. This situation requires single shot integrating detectors, which capture and process the complete image before the next XFEL flash because all photons arrive basically simultaneously. The photon counting technique, which has widely been used to achieve very good signal to noise figures, is not applicable.

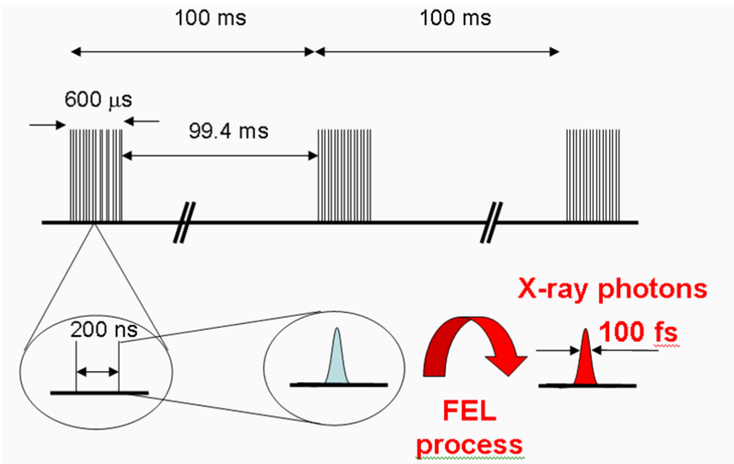


Figure 2.6: Timing pattern of the flashes generated at the European XFEL [3]. The bunch train has a length of 600 µs, the intra-bunch pulse rate is 4.5 MHz. The bunch trains are repeated at 10 Hz.

A very challenging further requirement is to process the maximum pulse rate of 4.5 MHz. Since it is a very unique property of the European XFEL, many experiments will make use of this fast frame rate. Because the single shot imaging technique must be used, only 220 ns are available to process each frame. Current technology does not allow for sending a complete frame off the focal plane of the detector at this frequency. Consequently, the detectors need some mechanism to store the frames temporarily. The fact that the frame rate is not continuous, but interrupted by gaps of almost 100 ms (the reasons are outlined in section 2.4.1), leaves sufficient time in between XFEL bunch trains to send all captured pictures out of the focal plane and further to the data acquisition system.

The timing pattern is depicted in figure 2.6, the flashes sum up to a maximum of 27000 per second. This situation requires novel signal processing concepts in the detector system.

Each detector needs to have a central hole, so that the unscattered beam can pass the device without causing any damage. Beam stopping techniques in front of the detector are not applicable because of the intensity of the beam.

The scientific experiments mostly require large area pixelated detectors with a pixel count of < 1M. The requirements on angular coverage, pixel size and sample to detector distance is inherent to the various experiments and cannot be unified. The range of possible distance of detector to sample under study is given by the dimension of the experimental halls. The experimental halls provide substantial freedom to optimize these parameters for the needs of each respective experiment.

2.5.2 The Adaptive Gain Integrating Pixel Detector (AGIPD)

The Adaptive Gain Integrating Pixel Detector (AGIPD) [13] is foreseen for high energy experiments in the range between 3 keV $-$ 15 keV. The project is developed by a collaboration of institutes led by a group at DESY. A thick 500 µm PIN (p-in-n) Si-sensor segmented into 200 μm^2 pixels provides a very high quantum efficiency for photons of 12.4 keV and above while the entry window was carefully tailored such that the efficiency is still significant down to 3 keV. The sensor is bump bonded to the readout ASIC, which has a dedicated channel for each sensor pixel. Since it is mainly used for high energy experiments, the AGIPD must deal with a radiation dose up to 1GGy over three years. Even when taking into account that the ASIC is shielded by the sensor, it is expected that the ASIC has to cope with doses of 100MGy in the same time span. Radiation hard design methodologies are required in the readout ASIC.

The ASIC signal processing chain is depicted in figure 2.7. It comprises a charge sensitive amplifier (CSA) structure, a double correlated sampling stage and an analog memory to store the images locally. To cope with the dynamic range requirements, the input CSA stage comprises a mechanism to change the gain adaptively. When the output of the amplifier reaches a certain threshold, further feedback capacitance is added to lower the gain significantly. The final analog output voltage of the amplifier along with the gain settings are forwarded to an analog memory stage. The memory can hold up to 352 images and provides a random access mechanism which allows for trigger and veto capabilities. The analog memory cells are a very crucial part of the design because they are very sensitive to leakage currents which are magnified by radiation damage. Extensive irradiation tests have been performed to verify functionality when exposed to the expected doses. These studies resulted in requiring to lower the operating temperature to -20 řC The memory cells are implemented using MOS capacitors and occupy the majority of the pixel area. During the XFEL gaps in between the bunch trains, the stored analog signals are sent off the chip, where they are digitized with commercially available ADCs. Further PCBs assemble the required data streams for the Train Builder interface.

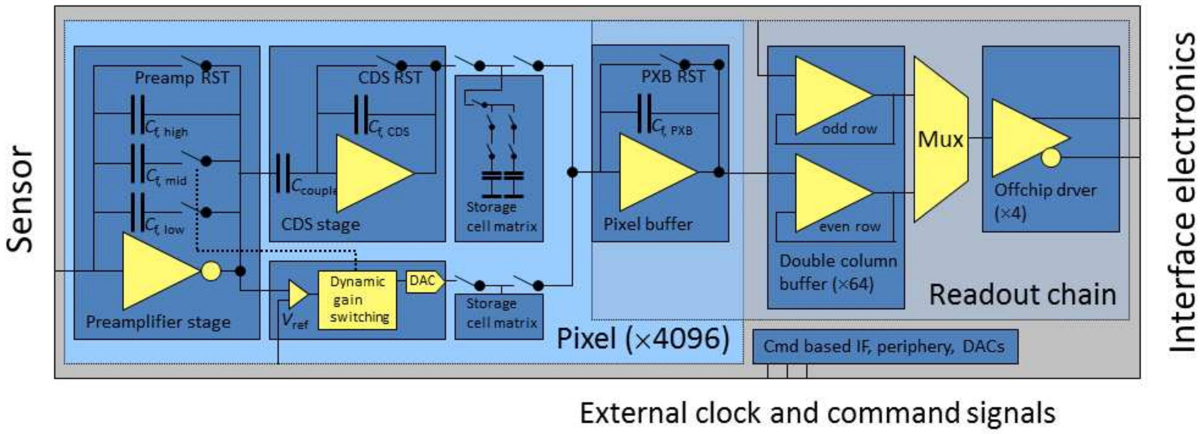


Figure 2.7: Schematic of the signal processing chain in the AGIPD ASIC [13]. The input stage comprises adaptive gain switching to cope with the dynamic range requirements. Analog signals are stored and sent off the chip during the XFEL gaps.

2.5.3 The Large Pixel Detector

The Large Pixel Detector (LPD) [14] is named as such because of its large 500 µm pixel pitch. The sensor is custom made for LPD and made from high resistivity silicon. The sensor is tiled in units of 128 × 32 and the ASIC pixels are substantially smaller than the sensor pixels. The sensor therefore comprises wiring to a denser grid matching the ASIC pixel pitch. Between the ASIC and the sensor dies, there is a silicon interposer with silicon through-vias to connect the ASIC and sensor pixels. Besides the sensor, the interposer represents a second shield protecting the ASIC from damaging radiation, relaxing the design constraints. The interposer is smaller than the ASIC, the regular IO wire bonds to the ASIC can such be hidden behind the sensor. The entire ASIC is such hidden behind the sensor, leaving a gap of active areas equivalent to 4 pixels when two sensor dies are placed adjacently. 16 sensor tiles make up a so called super module, which is the building block for larger areas. The first target configuration is to use 16 super modules to build a 1M pixel detector, while larger systems using more super modules are foreseen.

Figure 2.8 depicts the signal processing chain. The first stage in the ASIC is a charge sensitive amplifier which linearly converts the collected charge from the sensor to a voltage. The feedback capacitance can be selected as 50 pF or 5 pF. With the large feedback capacitor, a dynamic range of up to 10^5 12 keV photons per pixel can be reached. The small feedback capacitor provides substantially less dynamic range but better noise performance because the gain in the preamplifying stage is much higher. Following this preamplification stage, there are three parallel gain stages, which have gains of 1, 10 and 100 respectively. All three amplified signals are forwarded to an analog pipeline memory, which has a depth of 512 entries, and stored for the duration of the XFEL bunch train. The appropriate gain setting is chosen only in the FPGA which reads out the data from the ASIC. The ASIC also comprises 16 successive approximation register (SAR) ADCs, which operate during the gaps in between the XFEL bunch trains and digitize the signals stored in the analog memories. The digitized values are sent off the chip via LVDS wires to the so called Front End Module (FEM) DAQ (data acquisition) board, which also servers to control the chip. The on-chip analog memory is controlled by a command interface

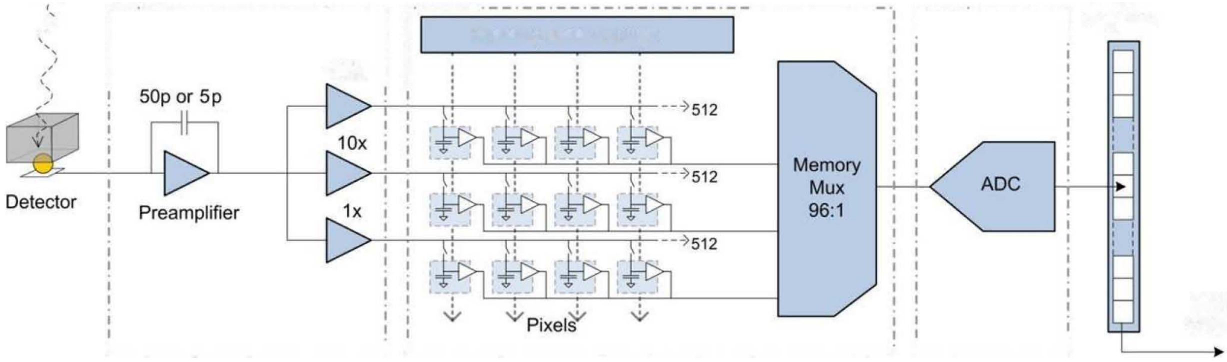


Figure 2.8: Schematic of the LPD signal processing chain [14]. A charge sensitive amplifier is followed by three fold gain stages implementing different gains. All amplified signals are stored in an analog pipeline memory. During the gaps between the bunch trains, the analog information is digitized and sent off the chip.

which also implements a veto mechanism, through which the user can void specific events. The core of the FEM board is a Xilinx Virtex 5 FPGA, which controls and gathers data from 16 super modules. In the FPGA, the correct gain level for each individual pixel is chosen. An attached extra card comprises the 10Gbps optical links which connect to the XFEL DAQ system.

2.5.4 The DEPFET Sensor with Signal Compression (DSSC)

Since this thesis presents work associated to the DEPFET Detector with Signal Compression, a separate chapter (4) has been dedicated to present it in detail. This section therefore provides only a short abstract of the DSSC detector system for completeness.

The DSSC uses an active DEPFET sensor with intrinsic signal compression or a mini silicon drift detector (MSDD), of which each pixel is bump bonded to a dedicated readout channel in an ASIC. The fundamental concepts are analog trapezoidal shaping to achieve for low noise performance, immediate digitization to 8-9bits and subsequent local digital storage within the pixels. The images are accumulated in the in-pixel memories for the duration of the burst and all data is sent off the chip in the gaps in between the bunch trains. The required dynamic range is provided on the sensor level, in case of the DEPFET sensor variant while in the MSDD a suitable mechanism needs to be implemented in the ASIC. The concepts of immediate digitization and signal compression on the sensor level distinguishes the DSSC from the other two detector concepts.

2.5.5 System Control and Data Acquisition

The European XFEL is developing a large software framework named Karabo [15], which will be used to control the beamlines and all further devices required for scientific experiments. The detectors will also be integrated as devices in this software and provide a configuration interface. The framework is very generic and will also handle the data acquisition (DAQ) and calibration and provide data analysis functionality. Physically, the Train Builder System [16] is the common readout interface for all detectors and provides an interface to a PC farm. A schematic of the datapath from the front-end-electronics (detectors) to the back-end PC farms is shown in figure 2.9. Each detector provides its data over multiple 10Gbps links, where different parts of the single images are transferred on different links. The Train Builder [16] accumulates the data and sends out complete images on dedicated 10Gbps lanes such that complete images can be stored on single machines on the back-end PC farm, where data processing and archiving takes place.

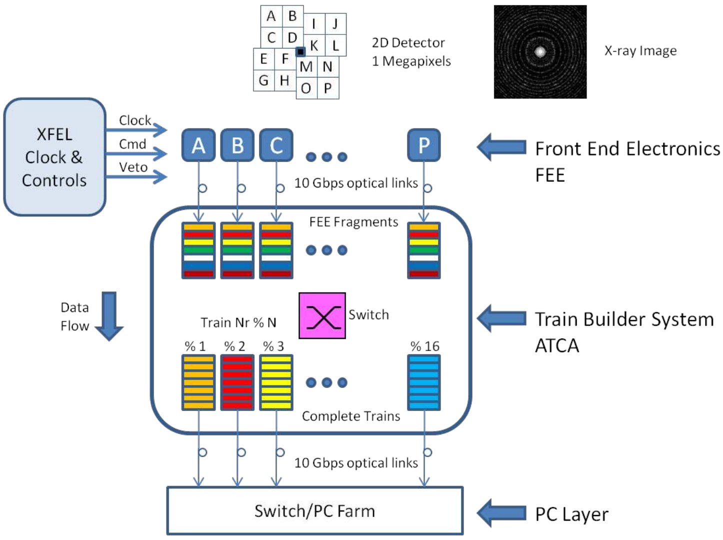


Figure 2.9: The EuXFEL data aquisition system [16]. The Train Builder System is the common interface for all detectors and sends off assembled images to the back-end PC farm.

3

Fundamentals of Silicon Detectors & Signal Shaping

3.1 Overview

This chapter introduces the reader to the underlying principles of the front-end building blocks in the DSSC detector. The first section handles silicon as a detection medium for electromagnetic radiation and the working principles of the sensors used for the DSSC camera. The second section introduces the noise phenomena in MOSFETs and the theory of shaping charge signals from sensors. It is not intended as a general reference but a specific reference for components and principles used in the DSSC system. If not otherwise stated, references [17], [18], [19] and [20] have been used to compile this chapter.

3.2 Silicon Sensors

Their unique properties make semiconductors very suitable for the detection of ionizing radiation. The basic principle makes use of the semiconductive nature of these materials, which means that there is an energy gap between the valence and conduction band. Free charge carriers are generated by incident radiation and can be evaluated by an electronics circuit.

In particular, silicon is very well suited because of its unique properties which are outlined in the next section. Silicon is also the primary choice for integrated electronic circuits. Therefore, the fabrication of silicon detectors has benefited from the electronics industry, as the basic process technologies were already very advanced when silicon has been started to be used as a detection medium. Meanwhile, the integration of detector and processing electronics devices on single silicon dies is possible. In the DSSC project, two different kinds of sensors are under study, the most promising variant is a very sophisticated and unique sensor, where a MOS field effect transistor is integrated on the sensor die providing a signal amplification and compression mechanism. The further readout chain is placed on a dedicated readout ASIC due to its complexity.

3.2.1 Properties and Doping of Silicon

The following properties make silicon a very attractive material for radiation and particle detection [19]:

- The band gap of only 1.12 eV (at room temperature) leads to an average energy of only 3.6 eV to create an electron-hole pair (an order of magnitude smaller than the ionization energy of gas for instance).
- Due to its high density of 2.33 g/ cm^3, the loss of energy per traversed length in silicon is very high (a MIP deposits 3.8 eV/ cm). Consequently, very thin sensors can be built. Furthermore, the generation of few δ-electrons stabilizes the center of gravity of the generated charge cloud. Therefore, very precise position resolution is possible.
- The mobility of both carrier types ($\mu_n = 1450\,\mathrm{cm}^2/\,\mathrm{V\,s}$, $\mu_p = 450\,\mathrm{cm}^2/\,\mathrm{V\,s}$ at room temperature) is very high despite for the large material density, thus allowing for fast charge collection times (~ 10 ns).
- Fixed space charges can be generated by doping the pure silicon (see below). Sophisticated electric field configurations can thus be created leading to a large variety of detector types.

Intrinsic semiconductors are rarely used because sufficient purity of the material is difficult to obtain. Furthermore, the material is mostly intentionally changed by so-called doping with donor (*n-type*) or acceptor atoms (*p-type*). Arsenic can for instance be used as a donor atom. It has an extra valence electron (five) with respect to the electrons required for bonding in the silicon crystal (four). Acceptor atoms, for instance boron, have one vacancy when bonded in the silicon crystal because they only have three valence electrons. This vacancy is called a hole and can be replaced by the electron from a neighbouring atom corresponding to a movement of positive charge.

The basic structure to build a silicon sensor is a diode, which is formed of a $p{-}n$ junction. When an n and a p doped volume are abutted, electrons diffuse into the p region and holes into the n region. An electric field is thus created which counteracts the diffusion and sweeps away any mobile charge carriers, such that a space charge region is formed. Strongly doped regions are indexed with $^+$, moderately doped regions with $^-$. The implantation of sophisticated doping profiles is possible which can be used along with external bias voltages to generate sophisticated electric field configurations to tune the sensor characteristics.

3.2.2 Interaction of Photons with Matter

There are several effects by which photons of various energies can interact with a matter, but only three of them are of practical significance for spectroscopy measurements:

- The absorption of a photon which causes an electron to be ejected from the atomic shell into the conduction band is called the **photoelectric effect**, this effect dominates for low photon energies.
- In a **Compton scattering** interaction, only some of the energy of the photon is transferred to a recoil electron. The amount of energy transferred to the electron depends on the scattering angle. All scattering angles can occur, yielding to a continuum of energies which can be transferred. The maximum energy is transferred for a head-on collision which is called the *Compton edge*. This effect dominates for medium photon energies.
- **Pair production** can happen in the extreme electric field of the absorbing material. The photon disappears and an electron-positron pair is created. Energetically, this process is only possible

when the photon energy exceeds twice the electron rest mass ($2m_0c^2 \approx 1\,\mathrm{MeV}$) In the vicinity of this threshold, the probability for pair production is small but it becomes dominant for several MeV.

For low photon energies (below 100 keV), a photon traversing a sensor is either completely absorbed or it passes the sensor unaffected. For an interaction with photons, a *point like* interaction happens, many $e^- -$ *hole* pairs are generated in a small region. In the case of a monochromatic beam, the beam is therefore not changed in energy but loses intensity. The intensity attenuation by a medium of thickness x is calculated by:

$$I(x) = I_0 e^{-x/L_a} \tag{3.21}$$

where L_a is the absorption length of the medium which depends on the photon energy. At photon energies below 100 keV, the photoelectric effect dominates.

The absorption length is an important parameter for designing a sensor. A large absorption length results in a high probability for a photon to traverse the detection medium without any interaction. If it is short, there is a high probability of ionization to occur at the surface of the medium, where deficiencies might exist causing the loss of signal. These can include surface implants to bias the medium (see section 3.2.4), coverage with an insulation layer (which might be naturally or artificially grown) and deficiencies in the semiconductor lattice.

3.2.3 Energy Resolution

The minimum detectable signal is limited by fluctuations of various nature. In a lot of situations, the minimum energy resolution is determined by noise in the readout electronics. When the boundary conditions are suitable, including for instance no constraints on power dissipation and very long times available for the signal processing, the noise in the readout electronics can become negligible. Nevertheless, there exists a lower limit on the energy resolution, determined by fluctuations of the signal generated in the sensor, which is present even for a fixed energy absorption. In a semiconductor sensor for instance, the incident energy is converted to electron-hole pairs. The number of generated pairs is given by

$$N = \frac{E}{\epsilon} \tag{3.22}$$

where E is the absorbed energy and ϵ the energy required to generate a single electron hole pair. The variance in the signal is given by

$$\sigma^2 = FN \tag{3.23}$$

where F is the Fano factor. The fluctuation of the generated charge carriers is due to a variation in the fraction of energy which causes electron-hole separation. F is therefore always significantly below unity. The difference of the deposited energy ends up in phonons (lattice vibrations) which is eventually dissipated as thermal energy. If all of the deposited energy would cause electron-hole pair separation, their number would be constant (for a fixed energy absorption) and F would then be 0. The Fano factor is a material constant, for silicon F = 0.115. For very low radiation energies in the few- eV range, it is energy dependent.

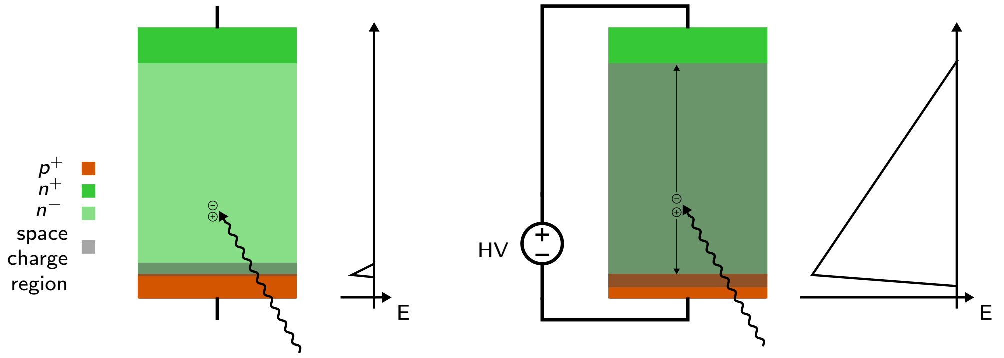


Figure 3.1: Principle of radiation detection with a diode. If no bias is applied (left), the space charge region and thus the sensitive volume is very small. Free charge carriers are not separated if they are created in the non-sensitive volume. A reverse bias increases the sensitive volume and creates a strong electric field to separate free charges.

3.2.4 The Reversely Biased Diode as a Sensor

The basic structure to build a silicon sensor is a diode, which is formed of a $p-n$ junction. The space charge region of the junction (see also section 3.2.1) can be used as for instance a photo sensor. The electric field will separate any free charge carriers which are created by light. This structure is however unsuited in practice because the sensitive volume given by the intrinsic space charge region is very small, the electric field weak and the capacitance of the device small (more on the importance of the capacitance in section 3.4).

By applying a reverse bias to the diode, the sensitive volume can be extended. The capacitance of the device thus drops and it becomes attractive for spectroscopy measurements. Such a structure can for instance be fabricated on a low doped n^- substrate by implanting a p^+ layer on the backside. An n^+ region is typically implanted on the front-side which provides for a good ohmic contact and allows to operate the device in over-depleted mode. The working principle of such a structure is depicted in figure 3.1. In principle, different implanting combinations are possible, a discussion can be found for instance in [17].

To calculate the depletion depth, the electric field, and the potential as a function of the applied voltage, the one-dimensional Poisson equation has to be solved. The doping profile is usually very asymmetric, which has the effect, that the space charge region grows almost exclusively into the lower doped region. Assuming a complete ionization of all donor and acceptor atoms, which is generally valid, the depletion width can be approximated by

$$W \approx \sqrt{\frac{2\epsilon_0 \epsilon_{Si}}{eN_D} V} \tag{3.24}$$

where V is the applied reverse bias, e is the elementary charge and N_D is the donor doping concentration (which is assumed as the lower doped region). Full depletion is reached when the space charge region touches the back side p^+ implant. In this state, the capacitance of the device is minimum. When the bias voltage is further increased, the device enters *overdepletion* in which a constant is added to the electric field.

3.3 Pixelated Sensors

This section introduces three different pixelated sensor types. While they all share the same detection principles of a large depleted volume, the possibility to dope the silicon crystal structure are being exploited to create 2D spatial resolution and equip the sensor with special characteristics. The three presented structures rise in complexity starting with a rather simple n$^+$-in-ndiode and arriving at the DEPFET which incorporates a transistor in the sensor pixel. The structures all collect e$^-$ and build on each other starting from the n$^+$-in-ndevice. The capacitances of the anode are emphasized because it will become clear in section 3.4, that they are key for good noise performance and thus energy resolution.

3.3.1 Pixelated n$^+$-in-nDiode Detectors

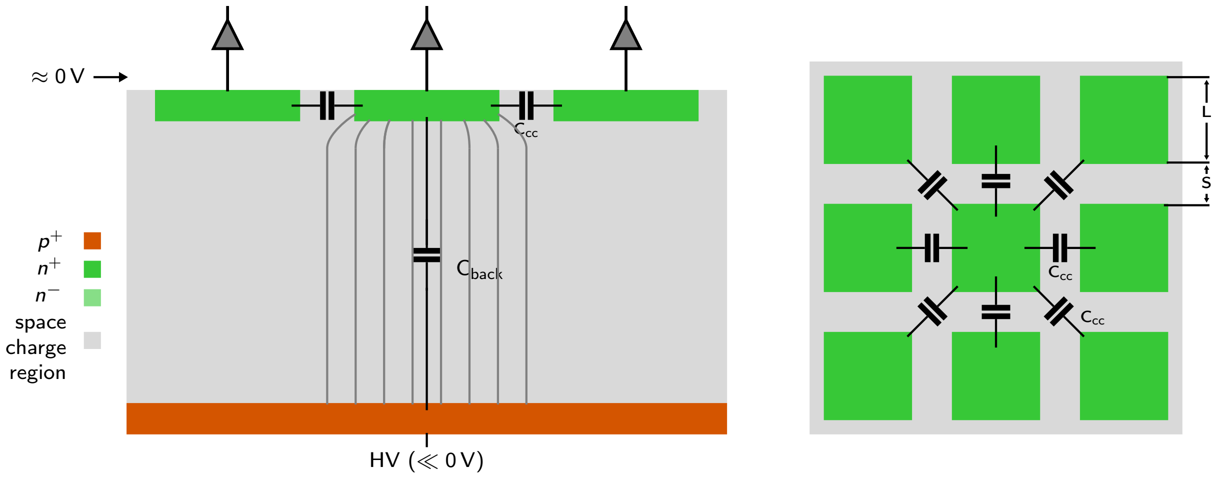


Figure 3.2: In a pixelated n$^+$-in-ndiode array, the capacitance is mainly coupled against the neighbours. The anode (green) needs to span a substantial part of the pixel for good charge collection characteristics.

The front-side (readout side) n^+ implant of the simple diode structure presented in section 3.2.4 can be segmented in the x and y dimensions and connected to a dedicated electronics readout channel to form a pixelated array. The cross section and a top view of such a structure are depicted in figure 3.2. A pixel located in such a 2D array has a capacitance against the back side and coupling capacitance against all neighbouring pixels. The size of the anode dictates the gap to the neighbouring pixel and thus the coupling capacitance. It is essential in this configuration, that the anode spans a large portion of the pixel to achieve a good charge collection characteristic. The electric field lines are bent and there is no lateral field. Therefore, charge can accumulate at the surface in between the pixels which drifts to the anode only slowly.

Depending on the geometry of the pixel (hexagonal, bricks, etc.), a fanout might be necessary to connect to a bump landing pad of a readout ASIC. The associated trace might need to be routed over the active area of a neighbouring pixel, adding to the cross coupling capacitance. The pixel anodes are ohmically connected to the substrate and therefore shorted when the sensor is not biased. The sensor needs to be fully depleted before the anodes are isolated.

To calculate the exact pixel capacitances, the three-dimensional Laplace equation needs to be solved. In [21], analytical expressions are given which fit the accuracy of a numerical solution to $\sim 10\%$. For a sensor with depletion width $W = 500\,\mu\text{m}$ and square pixels with edge length $L = 160\,\mu\text{m}$ and distance $S = 20\,\mu\text{m}$, the total capacitance calculates to 61 fF of which 48 fF are coupled to the neighbours and only 13 fF are decoupled[1] (against the back).

3.3.2 Mini Silicon Drift Detector (MSDD)

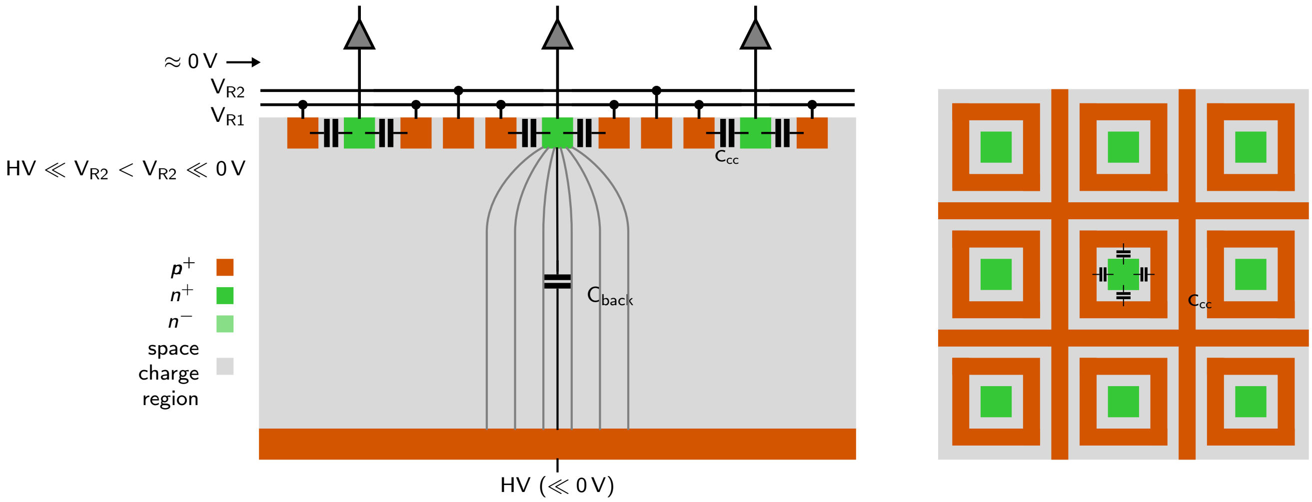


Figure 3.3: In a pixelated MSDD array, the pixel capacitance is grounded because it refers to a p^+ implant which separates the pixel. A lateral field directs signal electrons towards the anode.

To improve the separation of the pixels, p^+ rings can be implanted between the pixels. This creates a structure which is similar to the Silicon Drift Detector invented by [22]. If the p^+ implants are biased negatively against the anode, these implantations have several positive effects on the pixel characteristics:

- The electric field lines are bent, such that signal electrons are directed (through drift) towards the anode. The p^+ rings can be segmented and biased with a gradient to optimize the electric field pointing to the anode. The anode can thus be made very small.
- The required depletion voltage decreases, because a space charge region now also grows from the top of the substrate. This is the principle of sidewards depletion [22].
- The pixels are now shielded against each other, the coupling capacitance against the neighbouring pixel basically vanishes and is replaced by a grounded capacitance.

In this configuration, the anodes are also connected ohmically to the substrate when the sensor is not biased. It suffices however, to bias the p^+ implants sufficiently low to isolate the anodes.

The production is more complex versus the simple n^+-in-nstructure because a second implantation step is needed for the front-side, possibly also requiring extra metal layers for routing.

[1] *decoupled* refers to an AC grounded node

Again, the formulas in [21] can be used to estimate the anode capacitance. For $W = 500\,\mu\text{m}$ and square pixels with an edge length $L = 20\,\mu\text{m}$ and distance $S = 20\,\mu\text{m}$ to the first ring implant, the total capacitance calculates to only 7 fF of which 3.5 fF are coupled to the neighbours and 3.5 fF are decoupled against the first ring and the back.

When redistributing to an ASIC pixel with a different geometry it is now simpler to avoid additional cross coupling because the anode is much smaller and crossing neighbouring pixels can thus be avoided. Using a solder ball for bump bonding, which is larger than the anode, will however add capacitance if it overlaps the ring implants.

3.3.3 The Depleted Field Effect Transistor (DEPFET)

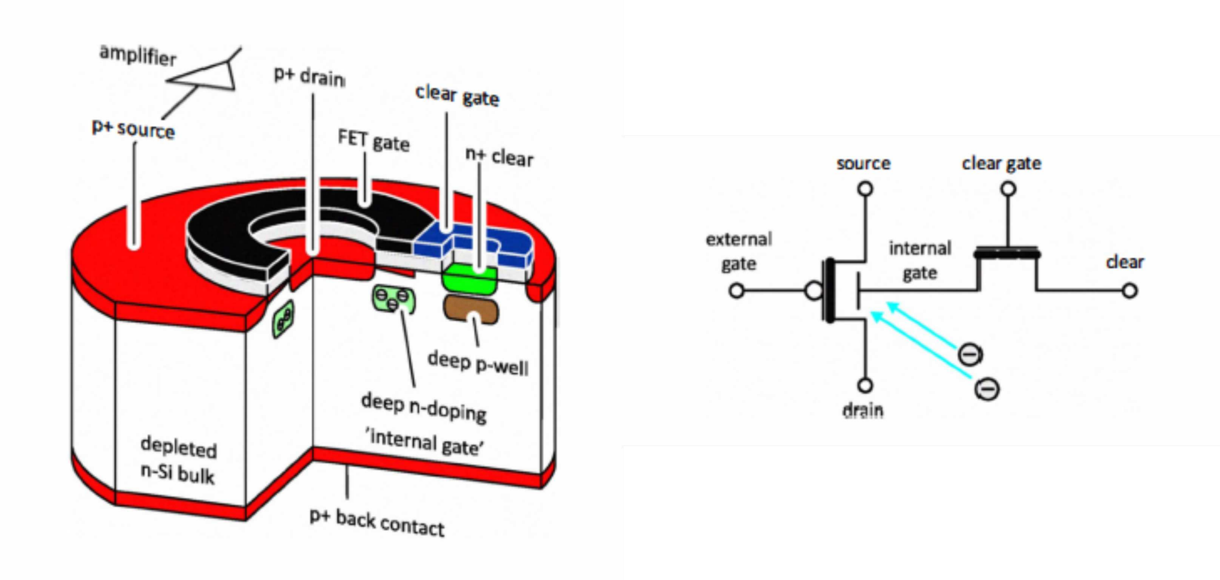


Figure 3.4: Cross section of a DEPFET active pixel sensor and equivalent schematic [23]. The signal charges are collected in the *internal gate* and induce a signal current in the transistor channel. The clear and clear-gate contacts provide the functionality to remove the signal charge from the internal gate.

The principal idea which has led to the DEPFET active pixel sensor has been invented by J.Kemmer and G. Lutz and proposed in [24], a schematic of the device is shown in figure 3.4. A p-channel field effect transistor is placed on a fully depleted high resistivity n-type bulk. The bulk is depleted using the sidewards depletion mechanism [22] (see also section 3.3.2). Beneath the transistor channel, a potential maximum for electrons is created by suitable doping (*deep n-doping*). Free electrons which are caused by incident radiation are collected in this area and induce mirror charges in the transistor channel and hence increase its conductivity. Due to the similar functionality as the regular MOS gate, this area is called the *internal gate* of the DEPFET. The device can collect and store charge in the internal gate without a bias current in the transistor. To evaluate the signal, a bias current in the transistor is required. To save power, the current can be turned on only for the evaluation process if the readout frequency is slow. When electrons are confined in the internal gate, a signal current is superimposed on the bias current which can be evaluated by a suitable electronics circuit. The device hence possesses an intrinsic amplification mechanism and can be classified as an active pixel sensor.

Due to the fact that the charge is collected on a high impedance node (internal gate) and this node does not need to be changed during the readout process, it is non-destructive. It is therefore in principle possible to read out the device multiple times. After the evaluation of the signal, the collected charge needs to be cleared from the internal gate. This is realized through the so called *clear* contact implemented with an n^+ implant located right next to the transistor. By applying a suitably high voltage to the clear contact, electrons are drained from the internal gate to the clear contact by punch through. In order to avoid the loss of signal electrons into the clear n^+ implant, it is shielded by a deep p-well implantation. However, this extra implantation makes the clear process more difficult because it generates a potential barrier between the internal gate and the clear contact. An additional contact, the so-called *clear-gate* is therefore introduced to ease the process. The contact is an SiO_2 contact, which enables an n-channel in the clear shielding p-well. Despite the fact that the n-channel is located at the surface, the punch-through can now reach down to the internal gate. The timing between the clear and clear-gate pulses needs to be such that the clear is enclosed within the clear-gate pulse.

In order to characterize the gain from the internal gate of the DEPFET, the so-called *charge transconductance* g_q has been introduced. Arrording to the analytical DEPFET model [25], it is given by:

$$g_g = \frac{\mathrm{d}i_{ds}}{\mathrm{d}q_{IG}} = \sqrt{\frac{2\mu_p I_D}{WL^3 C_{ox}}} \tag{3.35}$$

where μ_p is the hole mobility in the transistor channel, W is the width and L the length of the channel and $C_o x$ is the oxide capacitance between gate and channel per unit area. The gain is hence dependent on the geometry of the device.

The internal gate however is independent of the pixel size, because as in the case for the MSDD, angular p^+ ring implants create a lateral field directing the signal charge to the internal gate.

The direct measurement of the internal gate capacitance is not possible. In order to calculate the equivalent noise charge (ENC, see section 3.4) of the DEPFET and further signal processing chain with equation 3.47, it is more convenient to refer the signal and noise generators to the external gate of the DEPFET [23]. When we consider a DEPFET configured in source follower configuration[2], and set Δv_s as the change of voltage on the source node due to a signal q in the internal gate, the same Δv_s can also be produced with a voltage Δv_g on the external gate. The equivalent input capacitance can then be defined as:

$$C_{eq} = \frac{q}{\Delta v_g} \tag{3.36}$$

In using this method, an equivalent capacitance of the internal gate of 40 fF and an external transconductance of 83 uS has been measured [23]. Using these values, we can calculate $g_q = 328\,\mathrm{pA}/e^-$.

The series white and 1/f terms in the equation for the equivalent noise charge are a linear function of the input node capacitance. Due to the very low capacitance of the internal gate of the DEPFET, the device is very attractive for low noise applications. Due to the fact that it is enclosed in the depleted bulk of the sensor, it is also very well shielded against pick up from transients in the processing electronics circuits nearby. Charge collecting nodes which are directly read out by an ASIC through bump bonding have much higher capacitance and might be affected by switching noise on for instance power supplies.

[2] drain terminal at ground and biased for instance with a resistor

3.4 Noise and Signal Shaping

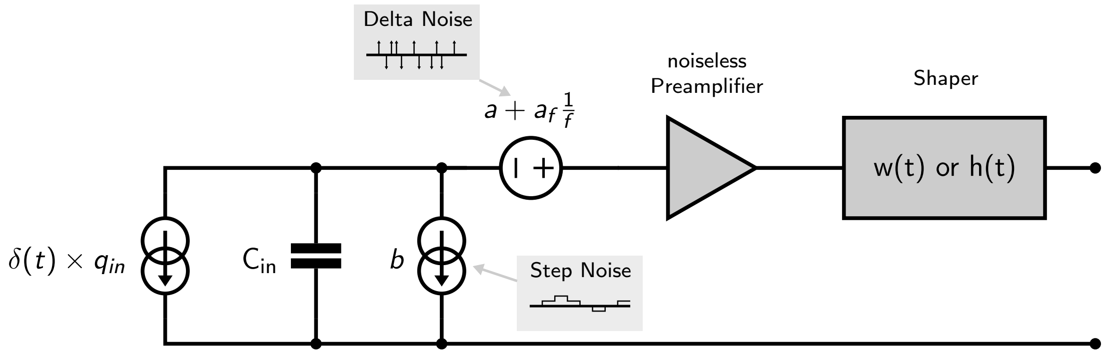


Figure 3.5: Equivalent circuit of a general detector front-end in a voltage sensing configuration with noise sources. The input node has a capacitance C_{in} while the signal is represented as a current pulse. The shaper is modeled using an impulse response $h(t)$ or weighting function $w(t)$.

A general schematic of an analog front-end reading out a sensor is depicted in figure 3.5. The sensor signal is represented by a current pulse $q \times \delta(t)$. The input capacitance C_{in} includes all capacitances shunting the input, like the capacitance of the charge collecting sensor node, the stray capacitance of the interconnect to the readout circuit etc. The discussion in this section is based on a voltage sensitive configuration which very well matches the situation for the DSSC detector (internal gate of the DEPFET or open loop[3] MSDD readout variant) The signal charge is integrated on the input capacitance C_{in} and forms a step voltage signal on the input. The derived expressions are however applicable to most configurations.

The circuit comprises a noiseless preamplifier followed by a filtering, or synonymously, shaping circuit. In reality of course all electronic circuits are noisy, their noise contributions are modeled by referring them to the input in order to relate them to the signal. A parallel current source and a series voltage source account for the total noise on the input. The filtering circuit is used to suppress the noise and thus optimize the signal to noise ratio. It processes all noise sources which precede the filter circuit in the signal processing chain.

Noise is a random process and we can therefore only calculate its statistical characteristics. Noise can for instance be modeled in the frequency domain using power spectral densities. In figure 3.5, the serial noise is modeled with the parameter a and the parallel noise with b. Of practical relevance are serial and parallel noise sources which have a constant power spectral density (white) and serial noise sources with a 1/f spectral power density (pink). Because the signal is a charge, the noise is usually expressed as the rms of the equivalent noise charge (ENC). The ENC is the charge required at the input to generate the rms noise fluctuation at the output of the system. It is calculated by dividing the rms output noise by the signal generated by a unit of charge. The common form used to calculate the ENC for a linear system is given by (for instance [26]):

[3]open loop because a single transistor is used without any feedback

$$ENC^2 = \frac{a}{\tau} C_{in}^2 A_1 + 2\pi a_f C_{in}^2 A_2 + b\tau A_3, \tag{3.47}$$

The parameters in the equation are defined as follows:

- C_{in} is the total capacitance at the charge collecting input node.
- A_1, A_2, A_3 are shaping coefficients given by the implementation of the processing circuit which can be tailored to specific applications, i.e. dominant noise sources in the system. The calculation of the parameters is explained in section 3.4.2 and section 3.4.3.
- τ is a time constant associated to a specific filtering technique, also referred to as the shaping time. Its definition is arbitrary but associated to the filtering technique (shaping coefficients).
- a, a_f and b are power spectral densities modeling the series white, series 1/f and parallel white noise sources in the system, respectively, referred to the input node.

Both the signal and the noise referred to the input are processed with the transfer function of the signal processing electronics. Constraints when designing a filter response for specific applications include for instance: knowledge of the arrival time of the signal, signal settling time and maximum time available for processing. Power consumption, area and complexity also play a decisive role especially for densely integrated circuits. The dominant series white noise in a well designed system is the thermal noise in the preamplifying transistor. The parallel noise source is mostly dominated by the leakage current in the sensor or a shunting bias resistor. Equation 3.47 shows, that long shaping times are favorable to reduce the series white noise but increase the parallel noise contribution. It is also worthwhile to note that the parallel noise contribution does not scale with the input capacitance. This can be understood from the fact that the noise current source is in parallel to the signal source, which means that the input capacitance acts on the signal and the parallel noise in the same fashion and thus cancels out in the calculation of the signal to noise contribution. The 1/f noise contribution is constant with respect to the shaping time and can only be improved by choosing a proper processing scheme.

In general, there are two different types of shapers: time-invariant and time-variant. For a time-invariant system, the output signal is independent on the arrival time of the signal, the response of the system does not change in time. Such systems are generally modeled by their impulse response $h(t)$. For a time-variant system, the output signal does depend on the arrival time of the signal. Time variant systems change their transfer function with respect to time, for instance by switching of capacitors (for instance the FCF filter in the DSSC section 5.3.2.1. They are generally modeled with a weighting function $w(t)$ because the notion of an impulse response is not applicable (see section 3.4.2).

3.4.1 Noise in MOSFETs

Flicker (1/f) Noise

Flicker noise has a power spectral density which is (almost) proportional to the inverse of the frequency and therefore also called 1/f noise. This noise is dominant at low frequencies and can become the limiting noise source for low noise applications. There are several theories for the origin of 1/f noise in MOSFETs [27]. The most commonly found theory in literature attributes it to the random fluctuation of the number of charge carriers in the transistor channel which is outlined in this section. These fluctuations are caused by the trapping and release of charge carriers at the interface of Si (channel) and SiO_2 (gate) of the transistor. Because the silicon crystal ends at this interface, there are dangling bonds, which gives rise to extra energy states. The time constants involved in trap-and-release process are very long. When a lot of such events superimpose, it can be shown that the resulting power spectral density is an inverse function of the frequency. Through the trapping effect, the interface charge is modulated, which in turn modulates the current in the transistor channel. When referred to the gate of a MOSFET, the power spectral density is given by [28]:

$$\overline{v_f^2} = \frac{K}{C_{ox}WL}\frac{1}{f} \tag{3.48}$$

The effect on the current in the channel can be calculated with the transconductance g_m. The parameter K is a process specific constant. There is no dependence on the transistor bias nor the temperature. The inverse relation to the area of the gate can be understood by the fact that for larger devices the effect averages out. Larger devices however come with an increased capacitance which is a burden when used as a preamplifying device sensing a charge signal.

Channel Thermal Noise

The general expression for the thermal current noise in the transistor channel in saturation is given by [28]:

$$\overline{i_n^2} = 4kT\gamma g_m \tag{3.49}$$

where k is Boltzmann's constant, T is the absolute temperature and g_m is the transistors transconductance. γ is an empirical constant which depends on the process and the length of the device. For long channel devices it is on the order of 2/3 while it is substantially larger for sub-micron devices. The noise current is modeled by a current source in parallel with the transistor channel. When it is referred to the gate, the according voltage source has a power spectral density of:

$$\overline{v_n^2} = \frac{4kT\gamma}{g_m} \tag{3.410}$$

The equation shows, that for an amplifying transistor, g_m should be maximized to minimize the noise. For a current source in turn, low noise is achieved by keeping the g_m low.

3.4.2 Time Domain Noise Analysis

For the analysis of filters in the time domain, a nice intuitive approach has been described in [29], which uses an elementary physical picture of noise sources. The method is especially interesting when analyzing time variant filters.

In the paper, parallel (or current) noise is referred to as *step noise* (see figure 3.5) because it results from the discrete electronic nature of current which is flowing in the input of the sensor. Due to the capacitance at the input, the current pulses are integrated resulting in a step signal at the input of the preamplifier. A physical example for white parallel noise is the shot noise in the leakage current of the sensor or the thermal noise of a bias resistor which shunts the input. Series (or voltage) noise is referred to as *delta noise* (see figure 3.5) which is caused by the discrete nature of current flow in for instance the preamplifying transistor. These delta current pulses appear when referred to the input as a delta pulse voltage generator in series to the input. The time domain analysis method presented here is restricted to noise sources which have a white power spectral density. The analysis of 1/f noise in the time domain is more involved, there is no intuitive view and they are better handled in the frequency domain.

For the analysis in the time domain, the weighting function $w(t)$ is used, which has been introduced by Radeka in [30]. The weighting function for a time variant system attributes a weight to a signal as a function of the signal arrival time. The measurement instant T_m is fixed with respect to the operating cycle of the system. Consider a trapezoidal weighting function as shown in figure 3.6. Only the signals which arrive during the flattop of the trapezoid are processed to the full amplitude, while signals arriving during the slopes of the trapezoid only give a fraction of the full amplitude. The weighting function generally does not represent the output waveform of the system. Using this concept, we must apply the weighting function not only to the signal, but also to all noise pulses referred to the input. The final amplitude generated by the system at T_m is given by the real signal plus the effect of all noise pulses which have occurred prior to T_m.

A current pulse $\delta \times q_p$ at the input, caused by a parallel noise source, occurring at time t_0 contributes $q_p \times w(t_0)$ to the output signal. Just like a *proper* charge signal on the input node, it creates a step signal on the input node. When we call $\overline{n_p}$ the average number of these noise pulses occurring in unit time and q_p the corresponding charge, their fluctuation is given by $q_p \times \sqrt{n_p}$. The cumulative effect of all noise pulses on the output can be computed with Campbell's theorem [31] and is given by:

$$\sigma_p^2 = q_p^2 \overline{n_p} \int_{-\infty}^{\infty} [w(t)]^2 \, \mathrm{d}t$$

A delta voltage pulse at the input representing a series noise source can be modeled with a step signal followed by an inverse step Δt later, where Δt is infinitesimal. Physically this corresponds for instance to the flow of an electron through the preamplifying transistor channel, an event which is not integrated when referred to the input in contrast to a charge pulse at the input which is stored on C_{in}. The delta voltage pulse at the input is therefore weighted by:

$$\frac{1}{\Delta t}(w(t) - w(t - \Delta t))$$

For $\Delta t \to 0$ this is the differential of $w(t)$ and we can use the derivative $w'(t)$ to calculate the cumulative effect of all series noise pulses with:

$$\sigma_s^2 = e_s^2 \overline{n_s} C_{in}^2 \int_{-\infty}^{\infty} [w'(t)]^2 \, \mathrm{d}t,$$

where $\delta \times e_s$ is a voltage noise pulse on the input, which correspond to series white noise. The terms $\overline{n_p} q_p^2$ and $\overline{n_s} e_s^2$ are equal to the mathematical noise power spectral densities for parallel and series white noise[4] according to Carson's theorem [31].

In order to provide the form given by equation 3.47 we normalize the weighting function to a reference time interval τ by setting $x = t/\tau$, where τ is now referred to as the shaping time. We thus obtain the following equations for the coefficients processing white series (A_1) and white parallel (A_2) noise sources:

$$A_1 = \int_{-\infty}^{\infty} [w(x)]^2 \mathrm{d}x \tag{3.411}$$

$$A_3 = \int_{-\infty}^{\infty} [w'(x)]^2 \mathrm{d}x \tag{3.412}$$

These calculations show, that series white noise is optimized by minimizing the derivative of the weighting function, which results in slow operation. Any constant part in the weighting function does

[4]which is one half of the physical spectral density

not contribute. At the same time, the parallel white noise is optimized when the total area under the weighting function is minimized. This corresponds to keeping the measurement time short. Missing for equation 3.47 is the coefficient A_2 which processes the 1/f noise. Analysis of the 1/f noise in the time domain is non-intuitive and is better suited for the frequency domain (next section).

3.4.3 Frequency Domain Analysis

The general method to describe a time variant system is to use its impulse response h(t) and the frequency domain equivalent $H(j\omega)$. The concept of a weighting function $w(t)$, however, is also applicable [26]. $h(t)$ gives the signal generated at the output of the system when it is excited with a $\delta(t)$ impulse at the input. $h(t)$ and $w(t)$ are related by:

$$w(t) = h(T_m - t), \tag{3.413}$$

Using the schematic in figure 3.5, the squared equivalent noise charge is given by:

$$ENC^2 = aC_{in}^2 \int_{-\infty}^{\infty} \omega^2 |H(\omega)|^2 \mathrm{d}f + 2\pi a_f C_{in}^2 \int_{-\infty}^{\infty} |\omega||H(\omega)|^2 \mathrm{d}f + b \int_{-\infty}^{\infty} |H(\omega)|^2 \mathrm{d}f \quad ^5 \tag{3.414}$$

For time variant systems, $|H(\omega)|^2$ must be replaced with $|W(\omega)|^2$ which is the Fourier transform of the weighting function $w(t)$ in equation 3.414. In order to bring this equation to the form of 3.47, we set $x = \omega\tau$ which yields for the three noise shaping coefficients:

$$A_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} x^2 \frac{|H(x)|^2}{\tau^2} \mathrm{d}x \tag{3.415}$$

$$A_2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |x| \frac{|H(x)|^2}{\tau^2} \mathrm{d}x \tag{3.416}$$

$$A_3 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|H(x)|^2}{\tau^2} \mathrm{d}x \tag{3.417}$$

[5] $\mathrm{d}f$ results from Parseval's theorem

3.4.4 Trapezoidal Shaping

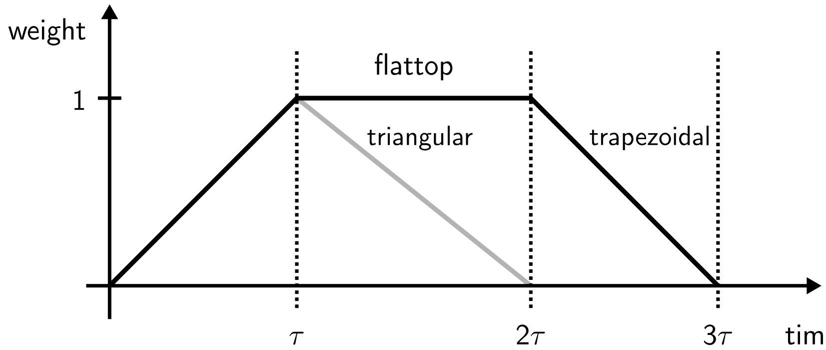


Figure 3.6: Trapezoidal and triangular weighting functions and definition of the shaping time τ. The flattop is necessary to avoid ballistic deficit due to the finite rise time (charge collection time) of the signal and degrades the shaping coefficients for series 1/f (A_2) and parallel noise (A_3).

Triangular and trapezoidal shaping is (theoretically) applicable as a time-variant or time-invariant processing scheme [31]. With the constraints of a finite width weighting function, the triangular weighting function is the optimum filtering scheme [32]. It is however not applicable if the signal has a temporal width. In this case, the triangular shaping would distort the signal, (ballistic deficit). To avoid this effect, a *flattop* can be introduced, resulting in a trapezoid. For the time-invariant case, the system must be timed, such that the signal arrives and settles during the flattop of the trapezoid. The implementation usually features gated integrators which is also employed in a novel circuit in the DSSC ASIC (see section 5.3.2.1).

The shaping parameters for series white and parallel white noise are best calculated using the time domain integrals of equations 3.411 and 3.412. The calculation of the 1/f shaping coefficient is more involved, and needs to be done in the frequency domain using equation 3.417. The parameter for the trapezoidal weighting function as shown in figure 3.6 are thus given by:

$$A_1 = 2,\ A_2 = 1.38,\ A_3 = 1.67$$

3.5 Discussion of the Presented Sensor Types

Equation 3.47 shows that the noise performance is inversely proportional to the input capacitance for systems dominated by series and 1/f noise. In this regard, this section briefly discusses the sensor types presented in section 3.2 and discusses their performance, a summary is shown in table 3.1.

For the n^+-in-ndevice and the MSDD, bonding and the connection of an external electronics circuit is required to read the charge signal from the sensor, the bonding and interconnect usually dominate the capacitance. Besides the inter-pixel capacitance, pick-up for instance from power lines can degrade the signal. While small solder balls are available which keep the interconnect contribution low, the choice is a question of cost and complexity.

In the case of a pixelated n^+-in-ndiode, the capacitance is large because the area of the anode is large and further depends on the pixel size. Most of the capacitance refers to the neighbouring pixels, because there is no shield in between, resulting in the worst cross talk performance.

For the MSDD pixel array, a p^+ implant is added to separate the pixels and introduce a lateral field. The anode can thus be made a lot smaller, reducing its overall capacitance and avoiding cross coupling capacitance to the neighbor.

The DEPFET solution overcomes the interconnect problem since the charge collecting node, the internal gate, is buried in the sensor volume. It is very well shielded against crosstalk, which makes it most attractive for low noise applications. It is however also the most complicated in terms of both production and operation, as for instance switched high voltages are required to clear the charge from the internal gate.

The capacitance of the MSDD and DEPFET devices are independent of the pixel size due to the usage of lateral electric fields to guide the signal charge to the anode.

Sensor Type	**n^+-in-n**	**MSDD**	**DEPFET**
Total Anode Capacitance	61 fF	7 fF	40 fF
Inter Pixel Capacitance	48 fF	negl.	negl.
Anode Interconnect Cap.	$100-400$ fF $\rightarrow$ incl interpx.	$100-400$ fF $\rightarrow$ incl interpx.	0 $\rightarrow$ not needed
Anode Shielding	no	no	yes
Intrinsic Amplification	no	no	yes

Table 3.1: Summary of the estimated pixel anode capacitances for the sensors presented in section 3.2. A pixel pitch of 200 µm has been assumed. For the n^+-in-npixel a pixel gap of 20 µm and the same distance for the gap of anode and first ring for the MSDD pixel.

4

The DSSC Detector

This chapter presents the DSSC detector system [33, 34, 35] in detail, because the design of the sensor readout ASIC is the core part of this thesis. The general requirements for detectors at the EuXFEL have been summarized in section 2.5.1. As already emphasized, the range of energies at the EuXFEL calls for the development of different detectors which are optimized for certain energy ranges. The DSSC detector aims at the lowest X-ray energies, ranging from down to 0.5 keV and up to 6 keV. The development of the detector is divided into several sub-workpackages including sensor design, ASIC design, module design, mechanical design and calibration. The system concept and physical structure of the camera are presented here, while in the last section, the expected performance is summarized.

4.1 Detector System Concept

The most challenging requirement is the combination of single photon resolution at very low energies, the large dynamic range and the maximum frame rate of 4.5 MHz. The simultaneous implementation of these properties goes beyond all existing instruments and requires the development of new concepts and technologies.

A block diagram of the system is depicted in figure 4.1. The main building block of the DSSC system is a DEPFET active pixel sensor which provides the properties of a very low capacitive sensor and amplification mechanism, ideally suited for low noise applications. To cope with the large dynamic range requirement at the EuXFEL, a novel mechanism has been invented, which provides signal compression at the sensor level (see Section 4.1.1). The very fast frame rate of 4.5 MHz is handled by reading out all DEPFET pixels in parallel. The ASIC concept is described in section 4.1.2 while the implementation is presented in detail chapter 5.

When the project was launched, the primary goal was to develop a system based on the DEPFET sensor, which has been expanded by an invention to suit the dynamic range requirements at the European XFEL. While test structures have been fabricated successfully, sufficient large scale matrices are not available yet due to fabrication issues and delays caused by political issues. The fabrication cycles are very long and complex, the development of a simpler version of the system has therefore been started in parallel. For this variant of the system, the active DEPFET sensor is replaced by a simpler passive mini silicon drift detector (MSDD), which is introduced in section 3.3.2. The aim is to deliver this simpler version of the system for the commissioning phase of the EuXFEL. The photon detection mechanisms of both devices are identical, the MSDD does however neither includes the nice

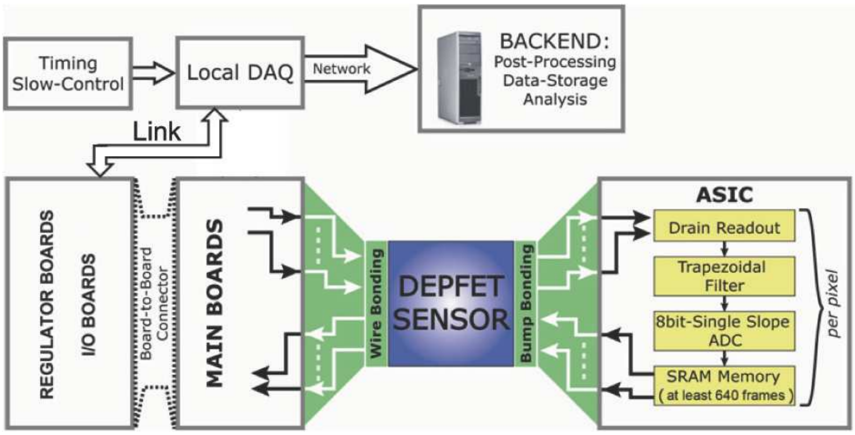


Figure 4.1: Block diagram of the DSSC system [35]. Each sensor pixel is readout by a dedicated ASIC channel, which digitizes the incident event and stores it locally. Several PCBs provide the infrastructure for the sensor and ASIC and establish the connection the EuXFEL DAQ system.

feature of active signal amplification nor a compression mechanism. Furthermore, the MSDD variant has been designed trying to reuse as much of the existing signal processing chain as possible, in order to seamlessly move back to the originally foreseen DEPFET sensor once enough large scale matrices become available. This goal poses constraints on the design of the MSDD front-end.

4.1.1 Sensors

DEPFET Sensor with Signal Compression

The DEPFET is ideally suited for low noise applications because of its very low input capacitance and intrinsic amplification mechanism. The general working principle of the DEPFET has been introduced in section 3.3.3. Considering the dynamic range of up to 10^4 photons of 1 keV, a standard DEPFET with a charge-transconductance g_q of around $350\,\mathrm{pA}/e^-$ and a direct readout of the signal current, the maximum signal current would amount to 1.44 mA.[1]. The minimum signal current for a single 1 keV photon would be given by 97 nA. Statically increasing the input capacitance to scale the signal current would destroy the noise performance according to equation 3.47 while processing the maximum current in the readout electronics is unfeasible when the input capacitance remains unchanged. The dynamic range challenge has been solved by an ingenious new signal compression technique which is possible to be implemented for the DEPFET [36].

The compression mechanism is based on the idea of shaping the internal gate such that the first collected electrons (small signals) have a stronger influence on the transistor current than electrons collected later (larger signals). The internal gate, which is formed by an n^+ doped buried layer, not only covers the transistor channel but also expands across the large area source of the device, these regions are called *overflow regions*. The implantations are made such that the highest potential which attracts the first electrons is located directly under the transistor channel. Signal electrons which are collected

[1]neglecting the fact that a standard linear mode DEPFET is not able to accommodate the according signal charge

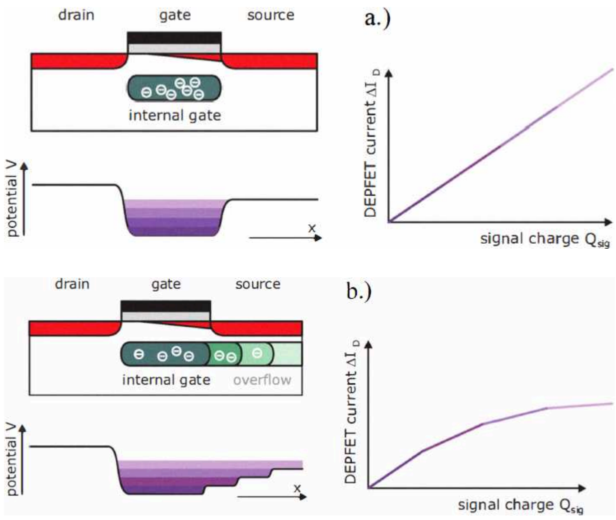


Figure 4.2: Working principle of a standard DEPFET (a) versus the novel DEPFET with signal compression [37]. Electrons collected in the overflow regions of the internal gate have less influence on the transistor current and hence cause a nonlinear characteristic.

here have the strongest influence on the transistor current. Larger signals distribute between the channel and the source region according to the electrostatic potential generated by the special implantation profile. The working principle is illustrated in figure 4.2 and the simulated potentials in the device are shown in figure 4.3. Only the fraction of the electrons located below the channel are effective in modulating the transistor current. Using this principle, a nonlinear characteristic between collected charge and signal current can be achieved. The working principle has been proved by measurements on a first prototypes in [37].

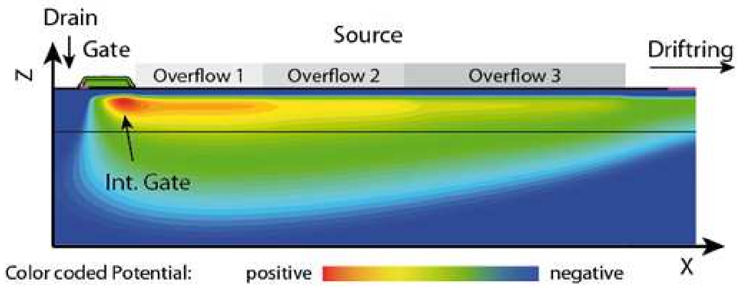


Figure 4.3: Simulated potential of a DEPFET with signal compression [37]

To calibrate the device, a nice mechanism has been proposed in [38], which allows for injecting charge into the internal gate via one of the bias contacts of the device. This mechanism allows for scanning the nonlinear curve of the device without exposing it to real radiation.

A hexagonal layout for the structure has been chosen, a simplified version is depicted in figure 4.5. This geometry provides a more homogeneous drift field versus a conventional square pixel. The DEPFET is located in the center of the cell, the drift field directs free electrons created by incident

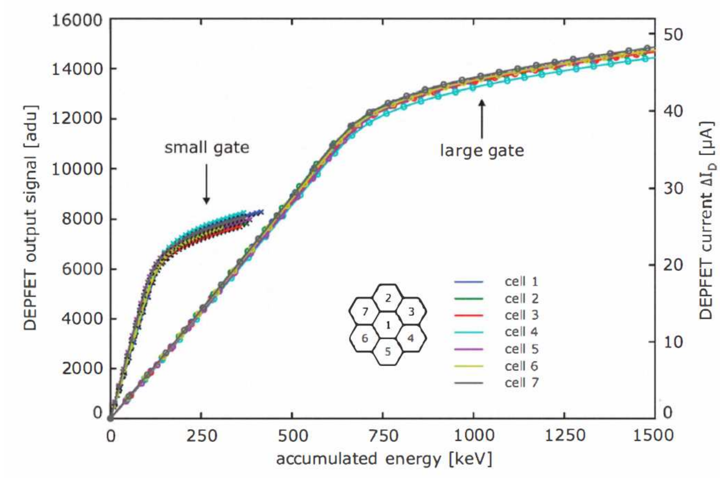


Figure 4.4: Measured gain curves of a prototype DEPFET (7 cell cluster) with intrinsic signal compression [37] (not mated to the DSSC ASIC). The gain depends on the geometry of the transistor.

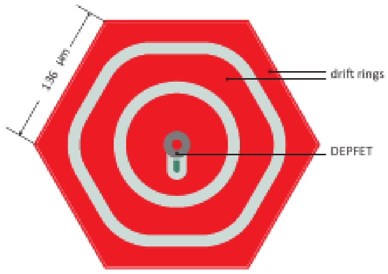


Figure 4.5: Simplified layout of the hexagonal DEPFET pixel [35].

radiation into the internal gate of the DEPFET.

Mini-SDD and ASIC Front-End Configuration

A discussion of some possible sensor variants has been given in section 3.5. To replace the DEPFET, the consortium has eventually decided for a Mini-SDD variant. The type of charge carriers to collect is fixed by the intention to reuse the existing signal processing chain, which needs a negative signal at the input. A simple PIN diode which is more cost efficient due to less processing steps [17] is for instance not applicable because it collects holes. The MSDD structure can be easily derived starting from the DEPFET sensor pixel by removing the transistor and replacing it with a classical electron collecting anode (n^+ implantation). A cross section of the device is depicted in figure 3.3. An argument can be made if the p^+ implants could be omitted in favor of simplicity but simulations have shown that they are indeed necessary for the target collection times of the charge of down to 50 ns which are required for the time variant filtering scheme implemented in the ASIC (section section 5.3.2.1).

In the ASIC, a mechanism is now needed which converts the collected charge to a signal current, as the DEPFET also supplies a signal current. A very simple approach has been followed: the DEPFET transistor has been replaced by a single PMOS transistor in the ASIC, equipped with a signal compression mechanism. The implementation and a discussion of shortcomings of the first version of the front-end are described in section 5.3.2.2.

Properties of DEPFET vs MSDD Readout Topology

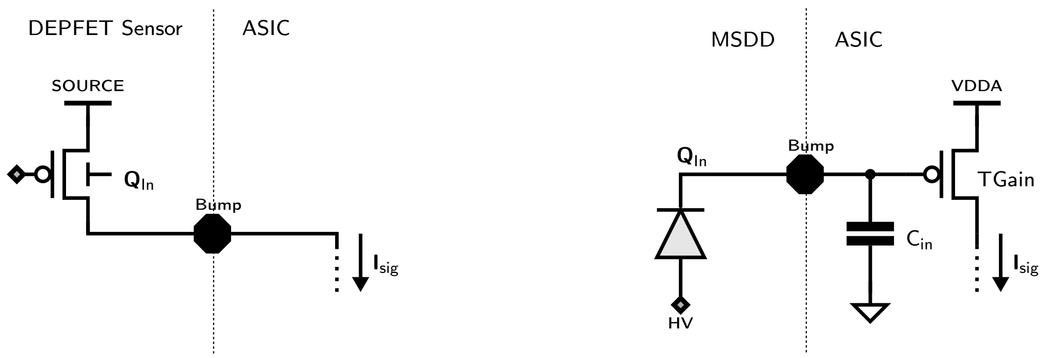


Figure 4.6: DEPFET current readout (left) versus adopted MSDD readout topology (right).

Figure 4.6 shows the DEPFET current readout variant versus the adopted MSDD readout variant. The MSDD readout adopted is not a classical charge sensitive configuration, as there is no virtual ground (*closed loop*) on the input. The simple circuit chosen is an open loop configuration, the charge on the input node is directly converted into a voltage through the total input capacitance on the input node C_{in}, which includes a marginal contribution from the sensor and is mainly dominated by the interconnect (bump + routing in the ASIC) and the gate capacitance of the transistor (TGain) in the ASIC. In principle, one can define a charge transconductance g_q, as it is used to express the gain for the DEPFET, for the MSDD readout topology as well by: $g_q = q_{in}/(C_{in} g_{m,TGain})$. A very large transconductance is required to overcome the larger C_{in} in the MSDD variant which is expected in the order of 400 fF − 500 fF and dominated by the interconnect. C_{in} can be reduced significantly using smaller solder bumps, which are available but introduce complexity in the bumping procedure and add significant extra cost. Due to the good yield with the IBM C4 (Controlled Collapse Chip Connection) process and the fact that for the target DEPFET system the input capacitance does not matter, the decision was made by the collaboration to not change the bumping for the MSDD variant of the system.

Tuning the g_q of the MSDD to the DEPFET g_q is however not enough when considering the noise performance. Equation 3.47 shows that a small capacitance is more beneficial than the g_m of the transistor. The input referred noise of the transistor is given by equation 3.410, it improves only with the root of g_m. Therefore $g_{m,TGain}$ needs to be very large (> 1 ms) to reach the low noise level of the DEPFET. Additionally, the MSDD charge collecting node is much more vulnerable to crosstalk from the dense environment, where again the large solder ball and associated interconnect in the ASIC comes into play. The internal gate is very well shielded due to being buried in the sensor volume, the current signal delivered by the DEPFET is not affected by crosstalk.

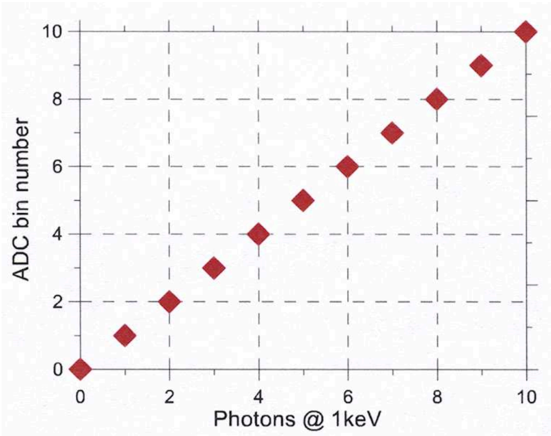
4.1.2 Readout ASIC

Besides the DEPFET sensor, the readout ASIC has the most significant role for the signal quality as it comprises all of the signal processing steps. The DEPFET is configured in drain readout mode, which means that the drain terminal is DC coupled to the ASIC input (figure 4.6, left side). The ASIC consequently has to sink the current needed to bias the DEPFET. A different approach would be the source follower configuration, where a bias resistor needs to be implemented on the sensor and a voltage signal is AC coupled to the DEPFET. AC coupling has the advantage that possible shifts in the threshold voltage of the DEPFET which increase its bias current can be tolerated more easily. The configuration has been studied by the consortium but has been discarded due to worse performance expectations than the drain current readout.

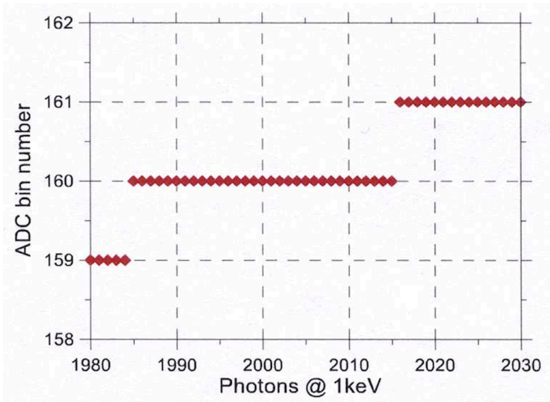
The ASIC input branch of the DEPFET readout mode comprises an adjustable current source which has to account for a possible increase of the DEPFET bias current as the sensor is exposed to large radiation doses in the experiments.[2] The actual signal current from the DEPFET is directly processed by a current mode filer (implementation details in section 5.3.2.1). The filter is of time variant nature and implements a trapezoidal weighting function introduced in section 3.4.4, which is the optimum weighting function at the target readout speed [32]. A triangular form would be slightly better but the flattop is required for signal acquisition and settling. The filter is needed to reach the target signal to noise ratio. After filtering, the signal is directly digitized and stored locally. This approach has been adopted to solve the challenge imposed by the fast maximum frame rate of 4.5 MHz. A precise ADC (implementation details in section 5.3.3) is a complex circuit, especially when implemented in a matrix of 4k pixels. The DSSC concept is therefore to digitize to 8bit only, a 9bit mode is possible for slower operating speeds ($\leq$ 2 MHz). Since the system is required to detect single photons while maximizing the dynamic range, the 8bit ADC is optimally exploited if the first photons (*linear range*) are attributed to single ADC bins. In this attribution scheme, charge sharing effects can not be compensated for but they have shown to be of marginal effect in [33]. For larger signals, the Poisson statistics underlying the signal are exploited. As more and more photons are attributed to single ADC bins through the nonlinear characteristic of the sensor or front-end a quantization error is introduced. The quantization does however not limit the system performance as long as the Poisson noise is dominant (see figure 4.8b).

This attribution scheme poses two very important requirements on the ADC: there must be a mechanism to set the inner bin offset precisely and the non-linearities in the linear region of the system must be very good. Influences of the non-linearities on single photon resolution have been studied in [39]. The signal is Gaussian distributed due to the electronics noise, and must be centered in the ADC bin in order to maximize the probability of proper detection of a single photon and ideally, the bin sizes should be equal. The concept is illustrated in figure 4.8a. Furthermore, the ADC needs a pixel-wise fine gain adjustment mechanism in order to compensate for a gain spread across the DEPFET pixel matrix. The single slope (Wilkinson) ADC is the architecture of choice because the concept scales nicely for a large matrix and both offset and gain adjustment can be implemented conveniently.

[2] The ASIC itself does not need to be radiation tolerant because at the foreseen X-ray energies, virtually nothing reaches the sensitive gates in the ASIC.

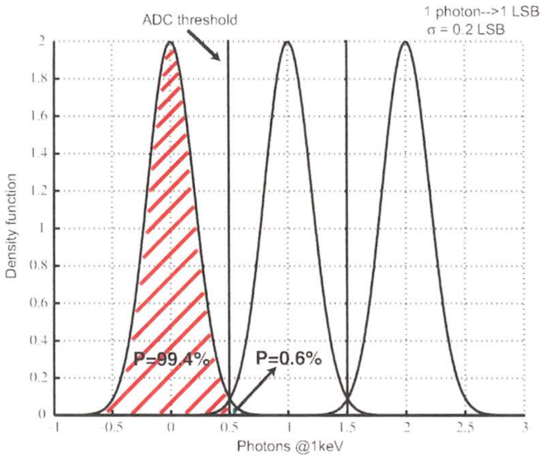


(a)

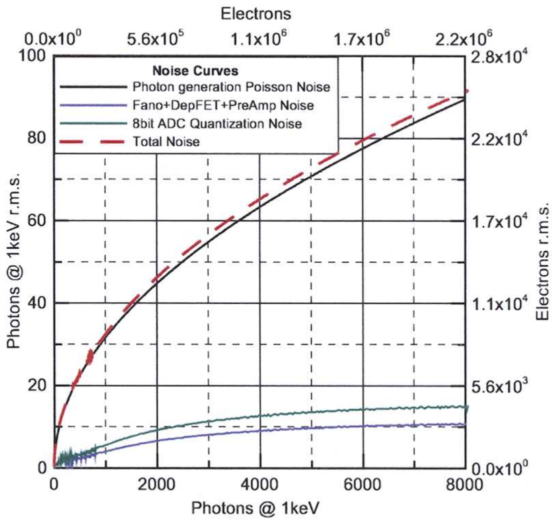


(b)

Figure 4.7: Attribution of photons to ADC bins [33]. For the first photons (a), the allocation of photons to ADC bins is 1:1 while for larger signals (b), more photons are attibuted per ADC bin due the non-linear characteristic of the DEPFET or MSDD front-end.



(a) The electronics noise is Gaussian and the ADC is be calibrated that the first photons fall in the middle of the ADC bins. At an energy of 1 keV and an (exemplary) electronics noise of 55 e^-, this results in a probability of 0.6% to falsely detect 1 photon.



(b) More and more photons are attibuted to single ADC bins as the amount of detected photons increases, increasing the quantization noise. The total noise is however dominated by the Poisson noise of the photon generation process.

Figure 4.8: Noise sources in the system [33].

Due to the attribution of single photons to single ADC bins and the non-linear system characteristic, calibration of the detector is a challenging task. The ASIC features various gain setting possibilities and the ADC offset must be set very precisely. Calibration is a separate work package, the latest progress has been reported in [40].

After the signal is digitized, it is saved in a local memory (implementation details in section 5.3.4), because the frame rate is too fast to transport a full frame off the ASIC in between two subsequent events. The wish is of course to store all 2700 events generated by the EuXFEL, such a capacity is however out of reach. The initial target capacity of 512 could be excelled, which is one of the results of this thesis. Thorough compression of layouts and the implementation of a memory with minimal overhead have lead to a final capacity of 800 events for the first full size ASIC. The memory is implemented as an SRAM (Static Random Access Memory). A mechanism is implemented in the chip which allows selectively discarding uninteresting events (implementation details in section 5.5.4). For the slowest target frame rate at the EuXFEL of 1 MHz (and fixed burst length of 600 µs), all 600 events can be stored.

4.2 Physical Structure of the Detector Head

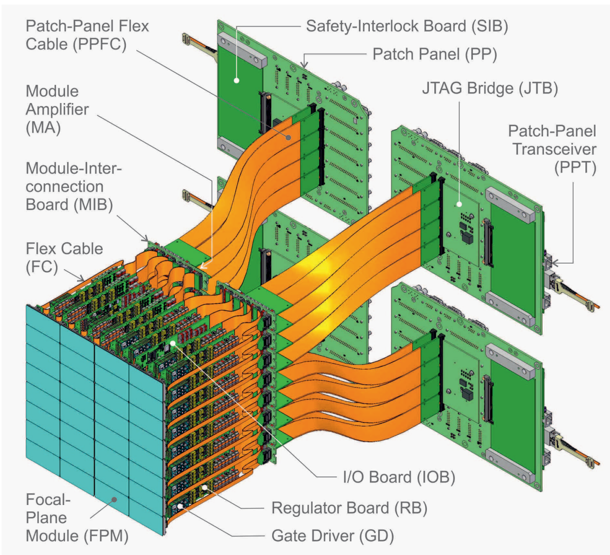


Figure 4.9: 3D view of the camera head [1]. The focal plane is subdivided into four identical quadrants. A quadrant is further subdivided into 16 monolithic sensors, each of which is equipped with 8 readout ASICs comprising 64×64 pixels

The physical structure and building components of the camera head have been published in [41]. A 3D view of the camera head is shown in figure figure 4.9. Physically, the pixel matrix of 1024 *x* 1024 has a size of $21 \times 21\,\mathrm{cm}^2$ and is structured as follows:

- 4 identical quadrants (512×512 pixels each)
- 4 ladders (512×128 pixels each) per quadrant
- 2 sensors (256×128 pixels each) per ladder

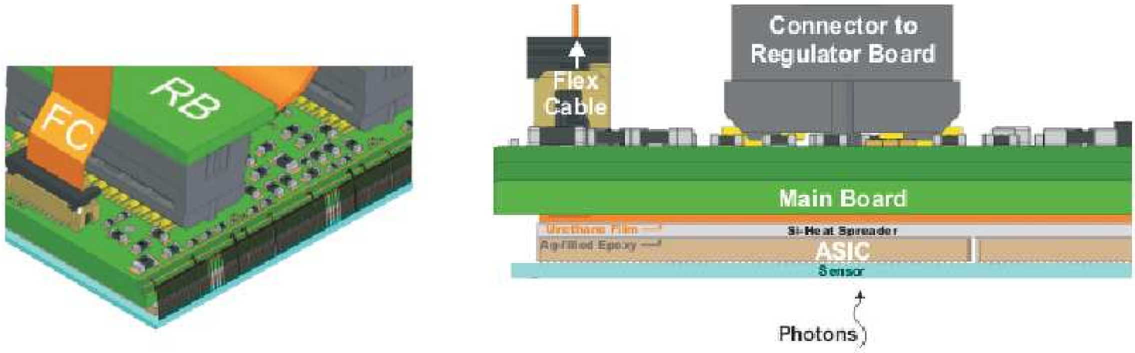


Figure 4.10: 3D views of the focal plane stack [41]. Left: side view showing the wire bonding to the sensor. Right: cross section of the stack of sensor, ASIC, heat spreader and main board.

- 8 ASICs (64 × pixels each) per sensor

The four quadrants can be moved to form a variable central hole through which the unscattered beam can pass without destroying any system components.

The first element in the focal plane are the sensors, which are bump bonded to the sensor. Figure 4.10 shows the focal plane stack: the sensor-ASIC assembly is hosted on a printed circuit board (PCB), the so called Main Board (MB). To carry out the heat generated by the ASICs which peaks during the XFEL burst phase, they are first glued to a Si-heat spreader using Ag-filled epoxy. The heat spreader further is glued to the MB with a hybrid urethane film adhesive. The choice of adhesive is important here because materials with different thermal expansion coefficients are meeting. A low modulus adhesive which stays above the glass transition temperature throughout the operating temperature range needs to be chosen in order to cushion mechanical stress caused by thermal cycling.

The main board is a 20 layer PCB which hosts clock drivers to distribute the 695 MHz ASIC clocks, filters for the sensor bias voltages and decoupling capacitors for the various supply voltages. The ASICs are electrically connected to the MB through the sensor. The sensor extends a little beyond the ASIC IO bump row. On this balcony, wire bond pads are located which serve to connect the sensor operating and bias voltages and fan out to landing pads for the IO bumps of the ASIC. The sensor bias bonds repeats for each ASIC section to form a regular bonding pattern. On the MB side, cavities are carved out to the inner layers of the MB in order to reduce the length of the wire bonds.

Perpendicular to the MB, there are five PCBs, four of which are so-called regulator boards which generate various supply and bias voltages. The fifth is the co called IO board (IOB) hosting an FPGA which generates control signals for the various components on the MB and the ASICs. It further collects the data from 16 ASICs and merges them into four 3.125 Gbit/s streams for the preceding patch panel transceiver.

Due to the large current request and low duty cycle of the signal processing elements in the ASIC, the corresponding supply voltages are only turned on during the XFEL burst phase of $\sim$600 µs (see also section 5.2). Each regulator board hosts twelve regulators to generate these pulsed supply voltages and clear gate drivers for 4×4096 pixels. Because of the high peak currents and the lack of sufficiently fast power supplies[3], the regulators generating the ASIC supply voltages are supplied by large capacitors. These are charged up to 7 V before the start of the XFEL burst and discharge along the burst. The gap in between the XFEL bursts is used for recharging up to 7 V. They have been sized such that the current requests of the ASIC can be satisfied, for the analog supply of a single ASIC for instance 3.2A are required. The total required capacitance for this purpose sums up to 8 mF per regulator board. The ASIC has sense outputs for all positive and negative supply voltages which return to the according regulators in order to cancel the voltage drops due to trace resistances.

The gate drivers serve to generate the required clear and clear gate pulses for the DEPFET sensor. The high and low levels can be adjusted, allowing for pulse amplitudes of 18.5 V and 10.5V respectively. At the maximum operating frequency of 4.5 MHz, only 50 ns are available to clear the DEPFET internal gate. A sophisticated driver has been developed which uses a push-pull output stage able to source currents of 9 A and sink currents of 4.4 A at the required speed. A transmission line transports these pulses to the sensors and is terminated with a 10 Ω resistor next to the wire bond pad.

The DEPFET bias voltages are brought to the main board via the IOB. Similar to the ASIC supply voltages, the source voltage which supplies the bias current of up to 150 µA per pixel needs to be power cycled. This voltage is generated per sensor (32k pixels), requiring to supply a total current of up to 5 A. The approach to generate them with dedicated regulators has been studied. However, the larger voltage requirement of up to 7 V and larger load current[4]. The available area is more efficiently connecting an external (slow) standard power supply and using the available area for capacitors. For a single net of the SOURCE voltage, a total of 37 mF could be placed. The energy required during the burst is thus entirely retrieved from the capacitors since the power supply is too slow to recharge them during the burst. This situation causes a droop of the SOURCE voltage along the XFEL burst. For a the maximum SOURCE current of the voltage has thus drooped by 81 mV at the end of the burst. A thorough investigation was conducted that this droop sustains the DEPFET performance.

These five boards (four regulator and one IO board) are all connected to a module interconnect board. This board mainly contains 28 mF of the capacitance for the DEPFET source voltage. A flex cable directly transports the static DEPFET bias voltages and JTAG signals for the ASIC slow control to the main board.

4.3 Data Acquisition Subsystem

The Data Aquisition (DAQ) subsystem of the detector covers the tasks of providing the captured image data from the focal plane to the XFEL DAQ system and of handling both the dynamic control and configuration (slow control) of the detector system.

The data passes two stages of FPGAs before it is transferred to the Train Builder (see also section 2.5.5) which provides a generic interface for all detectors. The first FPGA stage is located on the IO board (see also section 4.2) and collects serial data streams from 16 ASICs running at 350 Mbit/s. The IOB FPGA merges these data streams and outputs three 3.125 Gbit/s links which are connected to the second stage FPGA located on the Patch Panel Transceiver (PPT). One PPT serves a full detector quadrant and serves 12 3.125 Gbit/s lanes to collect the data of four IOBs. The PPT FPGA is the final device stage before the data is leaving the detector system. Here, standard UDP packets are assembled to comply with the specification of the Train Builder and the data is reordered to chronological order. It does not arrive in chronological order in case there have been VETOs during the respective burst.

[3]standard power supplies react with time constants in the millisecond range

[4]due to more pixels per net compared to the ASIC supply nets

Since the subsequent Train Builder stage requires chronologically ordered data it is necessary to buffer all frames before they are transmitted further. The PPT is the instance receiving the VETO telegrams from the clock and control system and converts the signals into appropriately timed telegrams for the fixed latency VETO mechanism in the ASIC. It implements the same mechanism which is used in the ASIC to keep track of the vetoed events and generates an allocation table of event IDs and memory locations.

The VETO reordering has been implemented and has been proven to be functional [42]. The incoming data from the IOBs is written to an external 800 MHz DDR3 memory as it arrives from the ASICs and read in chronological order using the allocation table generated during the XFEL burst. One PPT has 4 10 Gbit/s serial output links, 16 bit words are sent to the Train Builder (according to specification, payload is only 9 bit) which results in a data rate of 36 Gbit/s, yielding 144 Gbit/s for the whole system.

4.4 Summary of System Properties & Expected System Performance

Table 4.2 lists the system properties of the DSSC detector while table 4.1 lists the expected noise and dynamic range performance for the DEPFET and MSDD variants.

E [keV]	**DEPFET**		**MSDD**	
	Dyn. Range	ENC [e^-]	Dyn. Range	**MSDD**
	# ph	4.5 MHz	# ph	4.5 MHz
0.5	1116	18.4	420	62.6
1	3270	26	2859	71.6
3	> 10000	19.5	1011	58.8

Table 4.1: Expected noise and dynamic range performance of the DEPFET and MSDD system for various photon energies. [43].

Detector Parameter	**Expected performance**
Energy Range	optimized for $0.5 \leq E \leq 6$ keV
Number of Pixels	1024×1024
Sensor Pixel Shape	Hexagonal
Sensor Pixel Pitch	$\approx 204 \times 229\,\mu m^2$
Frame Rate	up to 5 MHz
Stored Frames Per Macro-Bunch	800
Operating Temperature	$-30°C$ optimum, room temperature possible
Operating Condition	Vacuum

Table 4.2: System properties of the DSSC detector. From [35], some numbers are updated.

5

ASIC Design

This chapter presents the design of the readout ASIC (application specific circuit) in detail. The ASIC concept has been published here [44]. The role of the ASIC within the detector concept has been explained in section 4.1.2. The subsequent sections describe the pixel and global chip architecture and implementation, followed by sections handling the integration and verification steps of the final full size 4k pixel chip. The pixels are the central elements of the chip and comprise all the signal processing elements. The MSDD Front-End and Flip Capacitor Filter have been contributed by Politecnico di Milano, Italy, the ADC and reference generation circuit has been contributed by the FEC group at DESY, Hamburg, the LVDS pads and test signal injection circuit have been contributed by the University of Bergamo, Italy. The author's main work has been to integrate all of theses blocks on the schematic level, provide most of the layouts, verify their proper interaction through simulation and provide the required infrastructure including readout and digital control for a 4k pixel matrix. An outline of the test chips submitted since the start of the project leading to the full format version is given in section 5.7.

5.1 Topology & Overview

The full format ASIC has a size of $14.9 \times 14.0\,\text{mm}^2$ and comprises a matrix of 64×64 pixels along with peripheral circuitry. Figure 1.2 shows a photo of the die while figure 5.1 shows the floorplan. The chip is fabricated using an IBM 130nm process, which is widely used in the high energy physics community. The chosen metalization option is a stack of 8 metal layers, which are all very densely used for signal and power routing (details in section 5.3.9). The nominal core and IO voltage is 1.3V, the interface has been designed such that no special higher threshold transistors are required. For the flip chip interconnect, the commercially available C4 (controlled collapse chip connection) process has been chosen using standard solder balls.

The module geometry dictates that all IO pads, including power and control, are located at one side of the die.[1]. It is therefore a little larger than the pixel matrix in one dimension to accommodate the IO pads and peripheral circuits. The topology results in long pixel columns, and supply voltage drops along the pixel columns are thus inevitable.

[1] each 256×128 sensor is equipped with 2×4 ASICs

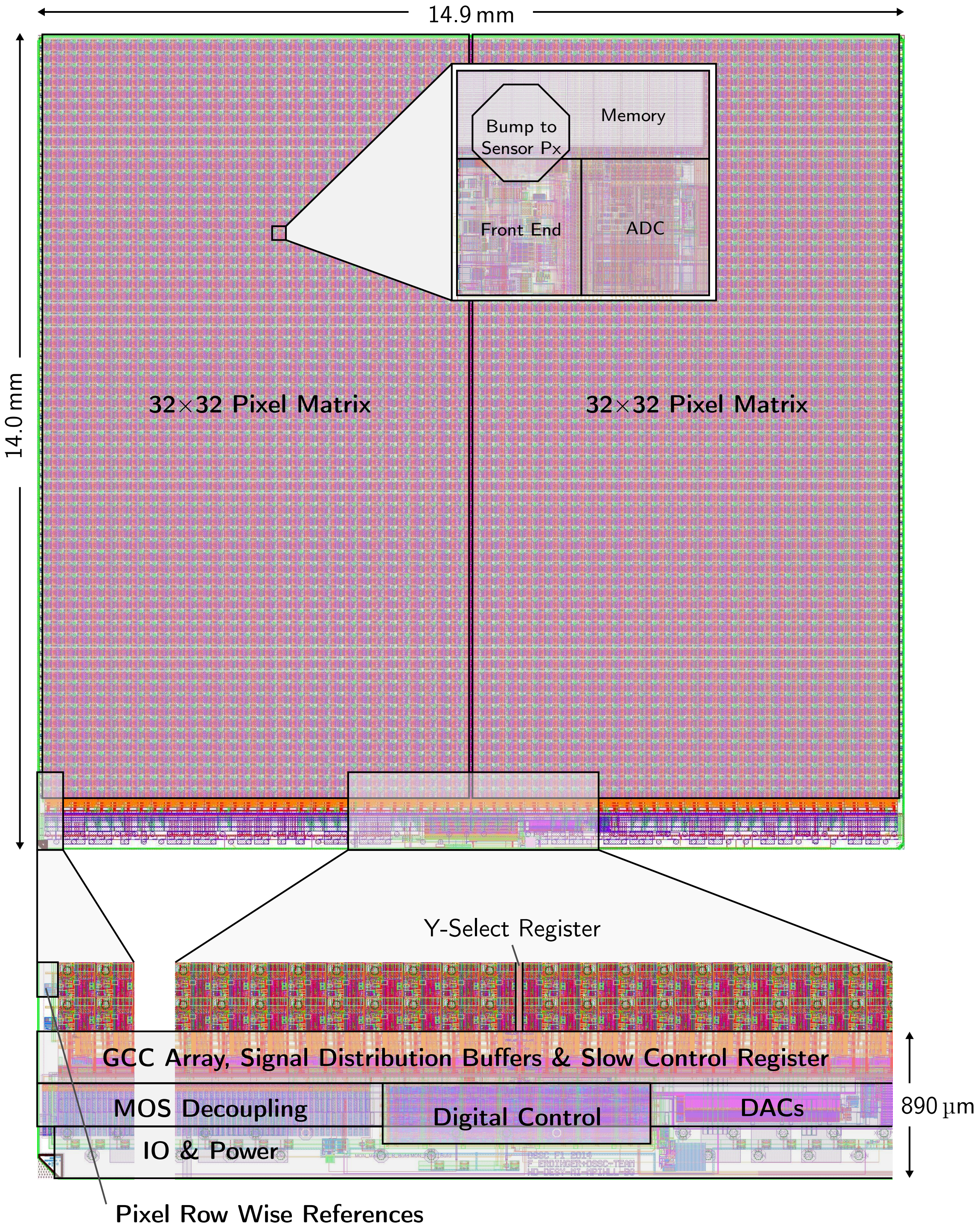


Figure 5.1: Floorplan of the full format ASIC. The die extends beyond the pixel matrix by 890 µm to host the peripheral circuits.

The IO bank features 79 pads, 13 of which are used for the dynamic and slow control interfaces (4×LVDS, 5×CMOS), one is a monitor pad, to which all pixels are connected and all others are used for power supply. No analog bias input is needed, all biases are generated in the chip. The peripheral balcony further hosts the digital control block (section 5.5), several DACs for trimming and calibration purposes, readout structures and a gray coded counter (GCC) per pixel column which is part of the in-pixel ADC. On each side of the pixel matrix, there is a bank of 64 temperature compensated reference circuits, two for each pixel row.

5.2 Operation Principle & Power Cycling

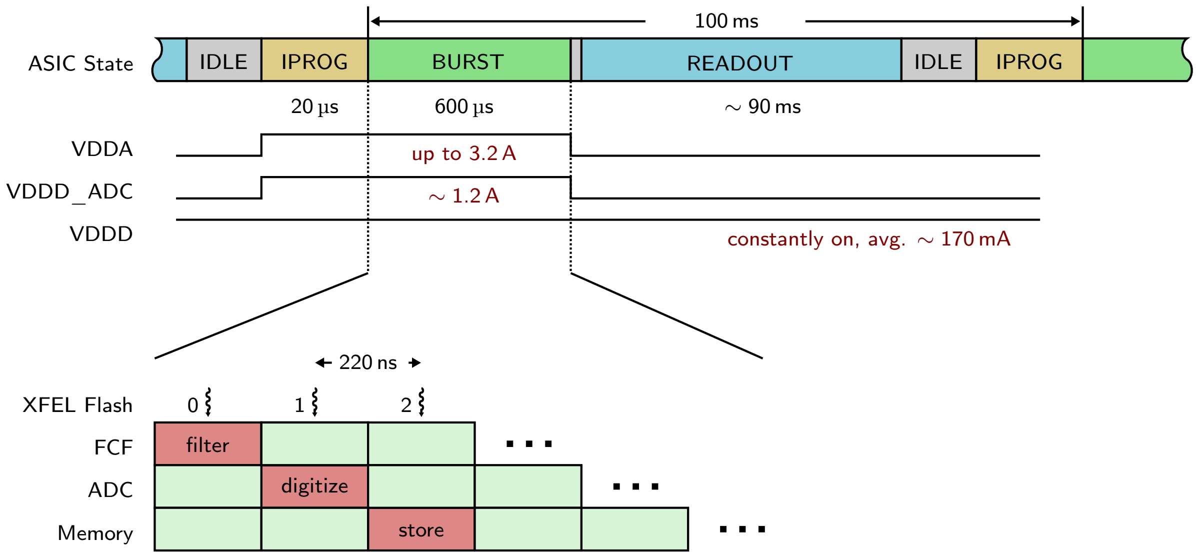


Figure 5.2: General timing diagram of the ASIC operation. `VDDA` and `VDDD_ADC` are pulsed power supplies and only active during the burst due to the high current load. In the burst, the ASIC building blocks operate in pipeline mode.

Before the ASIC is ready for data taking, all slow control registers need to be programmed, including counters which set the operating frequency, the sequencer (section 5.5.3), which controls the analog front-end in the pixels and the 47 bit slow control register (section 5.3.6) in each pixel which stores local gain and offset settings. determined by calibration.

Figure 5.2 shows timing diagrams of the general ASIC operation within an XFEL macro cycle. The macro cycles repeat continuously at a rate of 10 Hz. Each macro cycle contains a burst phase of 600 µs, which contain up to 2770 flashes with a maximum frame rate of 4.5 MHz and is followed by a long phase in which no photons arrive. In the `IDLE` state, the ASIC awaits a command from the controlling PPT FPGA which handles the communication and synchronization with the EuXFEL Clock & Control subsystem. When the ASIC receives the start burst command, it moves to the `IPROG` state which is needed to put the pixel front-end in the proper operating condition (section 5.3.2.3). After the `IPROG` phase, it automatically moves to the `BURST` phase in which the XFEL flashes are processed. During the `BURST` phase, `VETO` commands can be processed (section 5.5.4), which allow to void events

Supply (1.3V)	Pixel		Total	
	I_{sup} [mA]	Power [mW]	I_{sup} [A]	Power [W]
VDDA (analog cycled)	0.791	1.028	3.24	**4.21**
Filter	0.388	0.504		
ADC	0.303	0.394		
MSDD_FE	0.100	0.130		
Reference	0.008	10.4		
per half row (32 pixels)	0.256			
VDDD_ADC (dig. cycled)	0.275	0.358	1.13	**1.46**
GCC & TX	0.018	0.023		
per column (64 pixels)	1.152			
VDDD_GL(dig. static)	0.041	0.053	0.17	**0.22**
Memory	0.034	0.044		
Cntrl	0.007	0.009		
per Chip	30	39		

Table 5.1: Power consumption of the pixel and total chip including 4096 pixels and periphery. The dominant power supplies (VDDA & VDDD _ADC) are cycled and only active during the burst phase.

on the fly in order to keep only the events of interest in the memory. After the burst phase, the ASIC moves back to IDLE. The FPGA initiates the data transfer from the chip by sending the according command telegram. When put in the READOUT state, the ASIC automatically gathers and sends out the content of all pixel memories on one serial output link. The READOUT phase takes up almost all the time between the XFEL burst gaps of 100 ms.

In the DSSC concept, the collected events are digitized immediately and stored digitally. The processing front-end and ADC circuits have a very high power consumption but are actively operating only at a duty cycle of $< 1\%$. Therefore, the power cycling technique is employed, which is challenging for the total power consumption on VDDA of up to 4.21 W, which means that the power is shut down when the according circuits are not needed. Besides separating the digital (VDDD_GL) from the analog supply (VDDA), there is a third line which supplies the digital part of the ADC (VDDD_ADC). VDDD_GL supplies the slow control domain, the IO interface, the digital control block and the in-pixel memories. These circuits need to be powered in order to transfer the data off the chip and for the chip to keep its configuration for the subsequent XFEL burst. The pulsed power lines are enabled shortly before the burst by the IO board FPGA and shut down shortly after the burst. Besides reducing power dissipation, the power cycling technique is needed due to thermal considerations.

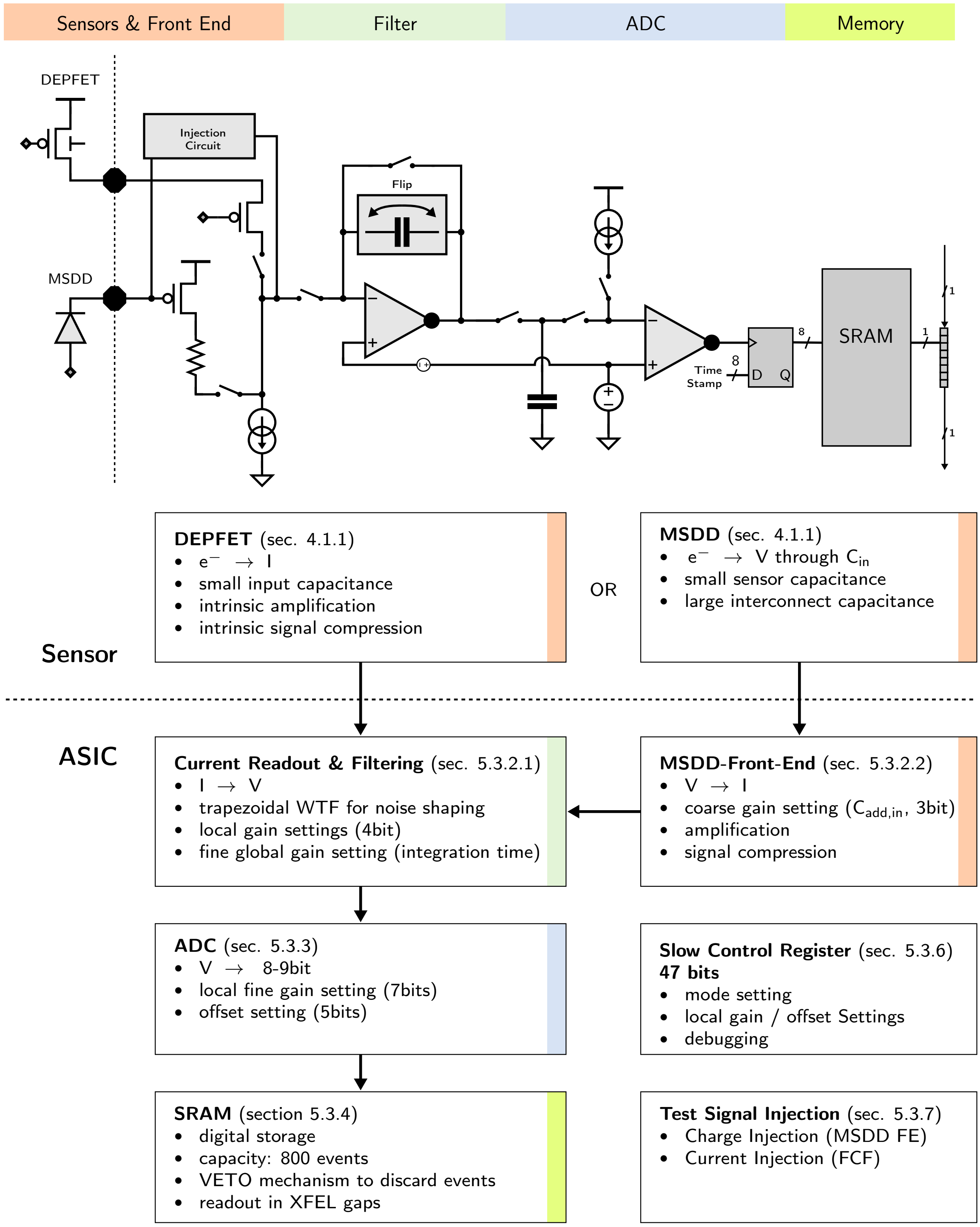


Figure 5.3: Overview of the pixel blocks and their respective core functionalities, the arrows indicate the steps of the signal processing chain. Details are given in the respective sections.

5.3 The ASIC Pixel

5.3.1 Overview

A block diagram comprising the major elements of the pixel and their respective functionalities is depicted in figure 5.3. The ASIC pixel has two modes, which are selected by slow control: a *current readout mode* which is implemented to read out the system specific DEPFET APS and a *charge readout mode* which is implemented to read the charge collected by an MSDD. Details on the sensors are given in section 4.1.1. The two modes differ only in the input stage and share all subsequent processing stages. The second stage is a current mode filter which implements a trapezoidal weighting function. The third stage is an ADC which digitizes the analog output of the filter circuit. The single slope concept has been chosen because it scales nicely to a large matrix since some parts of the required circuits can be shared among pixels. An exemplary transient simulation of the analog domain of the pixel is shown in the verification section (figure 5.31). For local storage, each pixel comprises an SRAM which accumulates the data of an XFEL burst and is read out in the 99.4 ms gaps in between the XFEL bursts.

5.3.2 Front-End Electronics

5.3.2.1 Current Readout Mode (DEPFET) and Flip Capacitor Filter

A schematic of the current readout front-end is depicted in figure 5.4, this is the original topology of the system. In this readout mode, the ASIC processes a signal current and the transfer characteristic of the ASIC circuits is linear since the signal compression is handled by the sensor. The circuit features three major components:

(1) The DSSC APS which converts the collected electrons into a signal current for the ASIC. The DEPFET is configured in drain readout mode where the drain terminal of the DEPFET is directly connected to the ASIC pixel such that the bias current in the transistor needs to be sunk by the ASIC.

(2) The input branch which includes a cascode transistor and a programmable current sink. The cascode ensures fast settling of the signal by hiding the stray capacitance of the bump node and the output resistance of the DEPFET. The current sink cancels the bias current of the DEPFET from the virtual ground node, such that only the signal current flows into the filter. The cancellation is not perfect, a residual current I_{res} remains which has to be compensated by the filter.

(3) The Flip Capacitor Filter (FCF), which is a time variant filter used to optimizes the signal to noise ratio. The current signal is converted to a voltage for the subsequent ADC stage by integration. The filter circuit has a low impedance input node where it provides a virtual ground at the V_{ref} potential.

The flip capacitor filter is based on a new architecture which was proposed in [45]. It is a circuit implementing a trapezoidal weighting function (see section 3.4.4) which provides optimum signal shaping at the target readout speeds of up to 4.5 MHz [32] for series white noise. The circuit performs a double correlated measurement per readout cycle, the slopes of the trapezoid are implemented by current integration. This architecture is an attractive filtering architecture for the DEPFET topology because it already delivers a current signal. The output noise can be calculated by using equation 3.47, when the noise spectral densities are known, the three noise shaping coefficients are given in section 3.4.4. Due to the fast speed, parallel noise is practically absent, 1/f noise contributes very little and series noise dominates. The filter amplifier itself of course also contributes noise and must be designed accordingly.

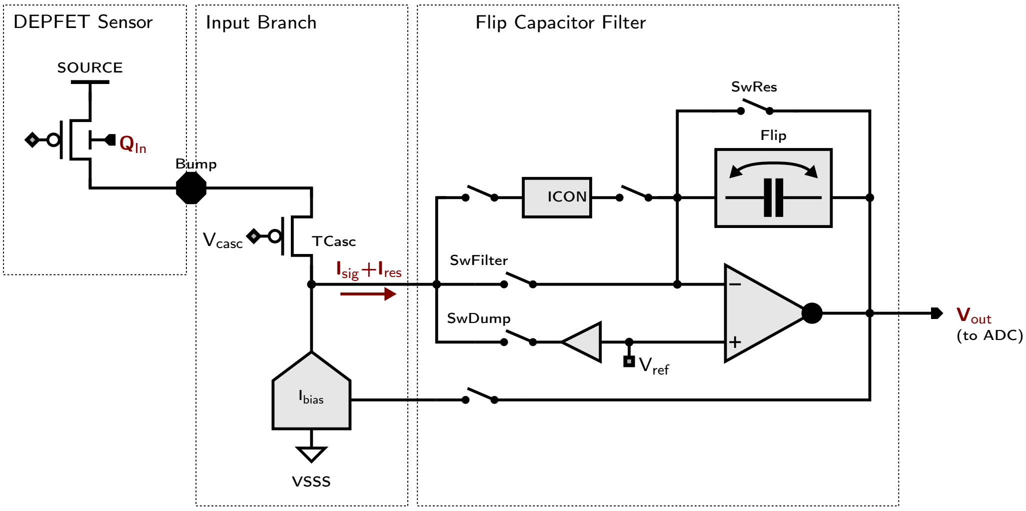


Figure 5.4: Schematic of the current readout mode. The DEPFET provides a signal current which is fed into the flip capacitor filter. The bias current is canceled from the filter input node by a dedicated current sink. This is the original architecture of the system, the charge readout path is not displayed.

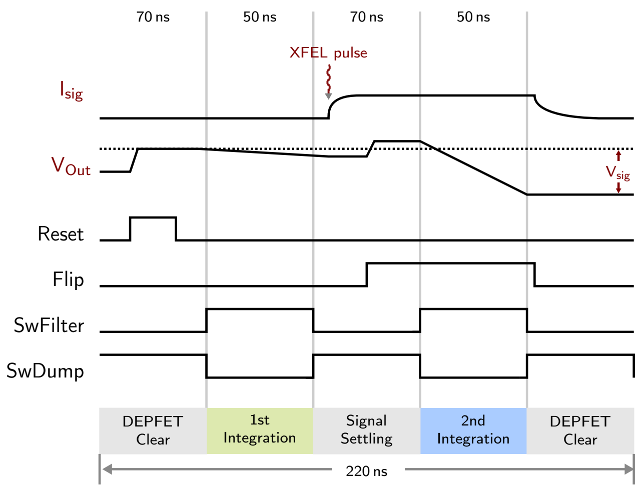


Figure 5.5: Three control signals are required to operate the circuit. The RESET signal shorts the feedback capacitor in the filter to cancel the preceding signal. FILTER and DUMP are complimentary signals which either send the signal current to the integrator or to a dump node. The FLIP signal turns the connection of the feedback capacitor which inverts the polarity of the integrated signal.

As depicted in figure 5.5, a readout cycle consists of four phases:

1. **DEPFET Clear:**
   The internal gate of the DEPFET is cleared to remove the previous signal. Concurrently, the FCF is reset by shorting the feedback capacitor.
2. **1st Integration:**
   The baseline current is integrated. This phase corresponds to the leading slope of the trapezoid. Although, a dedicated circuit subtracts the bias current from the virtual ground node, this is necessary to cancel any fluctuations of the bias current which happen along the burst. Some shift in the gate source voltage of the DEPFET for instance can thus be tolerated. The double correlated sampling also removes electronics noise in the baseline which is much slower than the readout speed (1/f noise).
3. **Signal Settling and Flip (Flattop):**
   The signal settling phase corresponds to the flattop phase in the trapezoidal weighting function. The system needs to be timed (switch sequence in figure 5.5), such that the XFEL pulse is located between the two integrations. Enough time also needs to be reserved that the signal can settle completely to avoid ballistic deficit. The feedback capacitor of the FCF is flipped in this phase to invert the polarity of the first integration. During this phase, the main integrator is gated off, a second auxiliary amplifier is needed to stabilize the virtual ground node as the signal current from the sensor arrives. Disturbances on the virtual ground node during this phase would lead to a non-ideal second integration.
4. **2nd Integration:**
   During the second integration, the baseline plus the signal current is integrated. This phase corresponds to the trailing edge of the trapezoidal weighting function. Since the capacitor was flipped in the flattop phase, the baseline which has been captured on the feedback capacitor during the first integration is canceled in the final output signal.

Figure 5.6 shows a simulated waveform of the filter output. The reference voltage of the circuit is chosen close to the positive rail since the signal is a positive current from the DEPFET flowing into the integrator which consequently produces a negative slope at the output. Some reserve towards the positive rail needs to be left such that the circuit can also cope with an offset current of negative polarity. The circuit was designed to cope with a maximum residual current of $3\,\mu A$ during the first integration for a feedback capacitance of $1.6\,pF$[2]. The presence of a significant residual current has been simulated in figure 5.6. It is evident that it is properly canceled by observing that for the red curve, the first and second integrations are parallel. The red curve simulates a minimum signal and thus returns close to the reference after the second integration. The upper limit of residual current which can be digested is given by the dynamic range of the amplifier. If it is too big, the filter saturates causing nonlinearities in the output signal.[3]

The flip capacitor scheme saves both precious area and power in the pixel by using only a single amplifier to implement the trapezoidal weighting function. The auxiliary amplifier which preserves the virtual ground during the flattop phase consumes very little power because it does not need to be optimized for noise.

[2] This is the minimum feedback capacitance which was initially foreseen for the DEPFET operation mode.

[3] This is true for both polarities of the residual current since the flip of the feedback capacitor inverts the baseline integration.

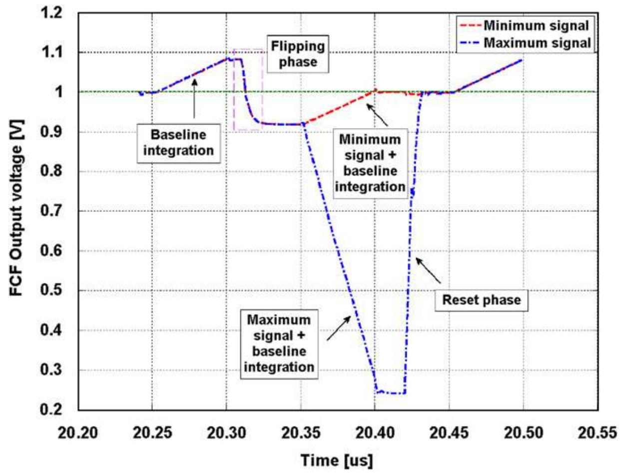


Figure 5.6: Simulated output waveform of the filter for a supply voltage of 1.2V [45]. The circuit can handle both polarities of a baseline current, but after the integration, the output voltage must stay within the dynamic of the amplifier ($\sim$ 100 mV) from the supply rail to avoid nonlinearities.

Initial stand-alone prototypes have been fabricated and characterized [46] before the circuit was included in matrix structures. Measurements using a standard DEPFET with a gain of $\sim$ 350 pA/ e^- bonded to a single channel test chip of the FCF have verified the low noise capability of the circuit. For an integration time of 50 ns corresponding to the 4.5 MHz XFEL pulse frequency, an ENC of 48 e^- was measured giving a signal to noise ratio of 5.8 for 1 keV photons. For an integration time of 400 ns, 13 e^- were measured, allowing single photon resolution for down to 250 eV at an S/N ratio of 5.3. A single integration readout cycle has also been investigated, which allows to increase the integration time to 70 ns for the 5 MHz mode. The weighting function in this case is no longer trapezoidal. Using this method, an ENC of 34 e^- has been measured. This has however been a measurement in an idealized single channel environment and it is unlikely, that the single integration method is applicable for the full matrix. The measurement however shows, that it is really the series white noise which is dominant at this readout speed and series 1/f and white parallel noise sources contribute less.

5.3.2.2 Charge Readout Mode (MSDD)

The MSDD is passive and does not provide signal compression. A suitable mechanism needs to be integrated in the front-end in the ASIC pixel to achieve the required dynamic range. Because the goal is to eventually provide a DEPFET based camera, the focus has been to develop an MSDD front-end which fits seamlessly with the existing processing chain. The baseline solution we have adopted is to use only a single PMOS transistor in the ASIC pixel which basically replaces the DEPFET. A discussion of this topology has been given in section 4.1.1 pointing out the drawbacks versus the DEPFET version. The current readout path is switched off by connecting the gate of the input cascode to the supply rail. Major contributors to the capacitance are: the relatively large solder bump, associated metalization and in-chip interconnect, the input transistor and the DEPFET cascode.

Measurements have shown (see section 7.2.1), that the input capacitance is of the order of 0.9 pF. The additional unexpected capacitance stems from the fact that the n-well of the DEPFET cascode,

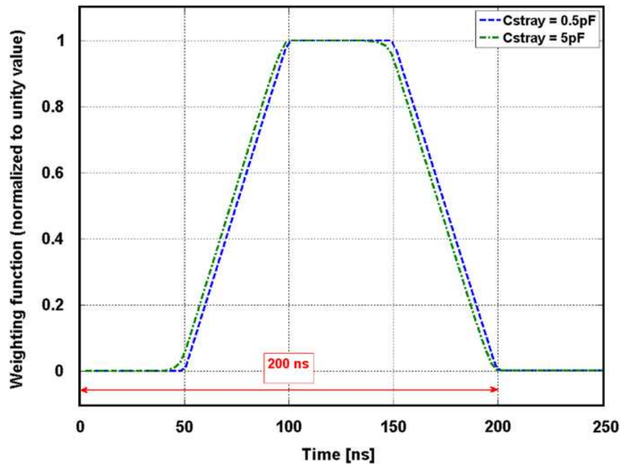


Figure 5.7: Simulated weighting function for two different stray capacitances at the ASIC input node [45]. Even for an exaggerated capacitance of 5 pF the trapezoid is well defined.

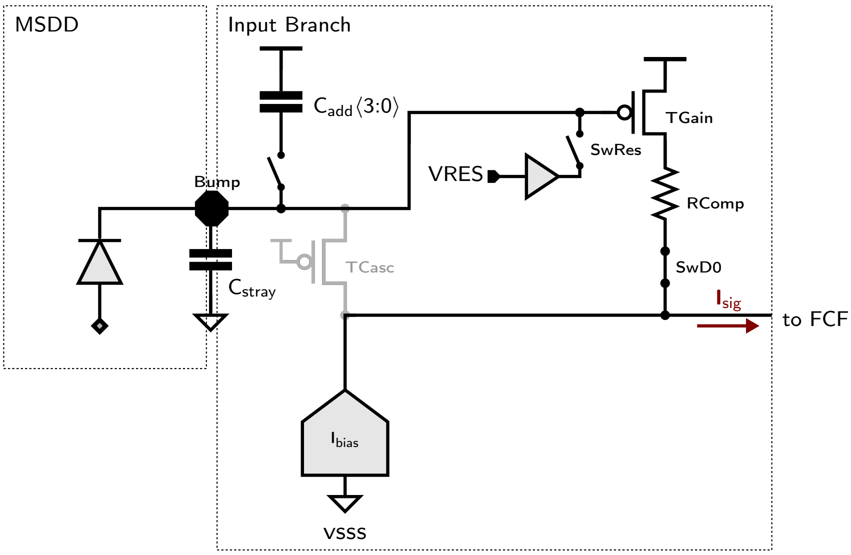


Figure 5.8: Simplified schematic of the charge readout mode. The DEPFET is replaced by a transistor on the chip. The transistor is operated without any feedback (open loop), the gate is biased with a reset voltage and floating during signal acquisition (switch SwRes). The current readout path is switched off.

was also connected to the input node. Connecting the source and bulk of the transistor optimizes its transconductance which yields a better cascode for the DEPFET mode, but of course adds capacitance. This connection could have been avoided.

The bias current in the transistor is adjusted by applying a proper *reset* voltage. The reset voltage needs to have a low impedance because it also serves to clear the signal charge from the input node analogous to the clearing procedure of the DEPFET internal gate. The reset procedure precharges the

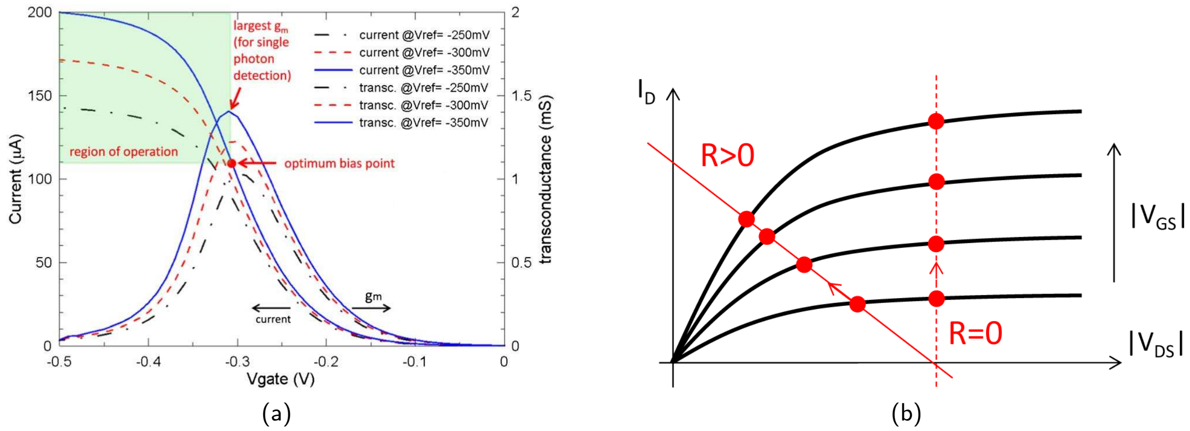


Figure 5.9: Principle of the triode compression [47]. g_m is maximum (a) for the current which has the transistor on the edge of the linear region due to the voltage drop across R. If the current is increased further (Vgate more negative, $|V_{GS}|$ increasing), gm shrinks due to the fact that $|V_{DS}|$ shrinks and the transistor is pushed into the linear region. $|V_{DS}|$ shrinks due to to the fact that the voltage across the resistor increases as the current increases.

input node to a certain potential which generates the desired current in the transistor. Afterwards, the node is left floating. Electrons which are collected on the input node change the potential and create a signal current in the transistor.

The PMOS transistor sensing the charge collecting node is enhanced by the so-called triode compression [47]. The transistor is biased in saturation but on the edge of the linear (or triode) region. A resistive element is connected in series to the transistor leading the current to the virtual ground node. As the current in the transistor and resistor increases with increasing signal amplitude, the voltage across the resistor rises. Consequently, the drain-source voltage of the transistor decreases. This mechanism drives the transistor into the linear region. The transconductance of the transistor consequently decreases. The gain from the input node hence decreases as the signal increases which compresses higher energy signals.

The circuit is very sensitive with respect to its optimum operating point. The working principle is illustrated in figure 5.9. If the PMOS input transistor starts out too deep in saturation (when the reset voltage is too high), the transconductance of the circuit first increases before the compression mechanism can act. This is due to the fact that if the transistor is in saturation and its drain current increases, the transconductance also increases which results in the inverse of a compression. A compressive behavior can only be reached when the transistor is pushed into the triode region where the transconductance decreases. The overall transfer characteristic shows a *point of maximum gain* which is the optimum bias point also in terms of the noise performance. Left of this point in figure 5.9, the transfer characteristic is compressive, while right of this point, the gain increases with increasing signal (negative voltage) on the input gate.

In the F1 implementation, the reset voltage is generated in the periphery of the chip with an on chip DAC (see section 5.4) and distributed to the pixel via a global wire entering the pixel at the VRES pin in figure 5.8. The voltage is buffered in the pixel with a simple source follower to provide a low impedance node. Static voltage drops and the use of a global reset voltage degrade the reset gate source voltage of the input transistor, which defines its bias current, along the pixel columns. Therefore is it impossible to optimize the transfer characteristic of all pixels in parallel for this first large-scale implementation of the circuit. Fabrication mismatches in the transistor threshold voltages, in particular of the amplifying and reset voltage buffering source follower transistor, add further random inhomogeneities in the matrix. For future variants of the chip, an adaptive mechanism is being investigated which is based on generating the reset voltage pixel-wise to improve the gain homogeneity of the matrix (section 6.2).

The dynamic range of the circuit using a simple resistor is very limited. The curve is rather suboptimal because it has a sharp kink where the gain changes too abruptly. An improved version of the circuit which has been submitted on a dedicated mini matrix on the engineering run uses an NMOS transistor as the series element for the compression. The NMOS changes its resistance throughout the dynamic range which produces a smoother curve. Further parallel branches are also added to preserve a final slope. At the time when F1 has been submitted, the improved version of the circuit has not been available on silicon yet. Therefore, the simpler variant using the simple resistor as the compressing element has been placed on the full scale chip as this variant had been characterized on silicon and was known to work reliably.

The triode compression mechanism is based on a nonlinear characteristic between input signal amplitude and transconductance. A second approach based on a nonlinear characteristic between input capacitance and signal amplitude is also under study. The mechanism is more complex and described in detail in section 6.1.

5.3.2.3 Bias Current Cancellation

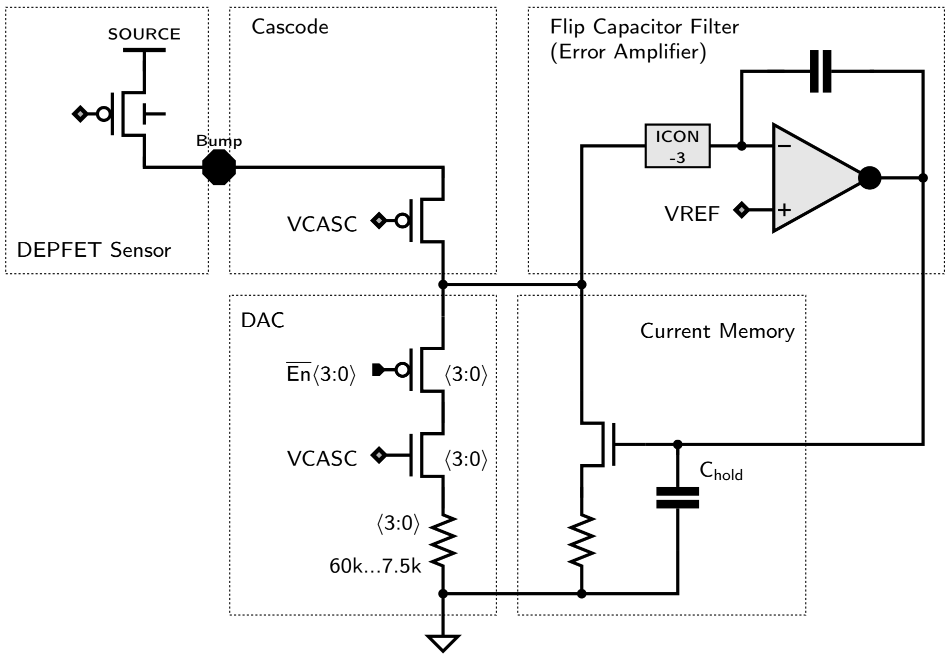


Figure 5.10: Mechanism for the fine programming of the bias current in the input branch, exemplary for the current readout mode. C_{hold} is charged such that the DAC and current memory cell sink I_{bias} and no current is flowing to the filter. The schematic is simplified and shows only the current programming (IPROG) phase, switches are omitted for clarity.

Both readout modes use single transistors converting the collected signal charge to a current which is fed into the filter stage. The bias current is sunk by a common cell, which needs good precision in order to avoid excessive current flowing into the filter. Too much DC current flowing into the filter would cause it to saturate and hence corrupt the output signal.

The implementation of the current sink is depicted in figure 5.10 and has been proposed in [45]. It contains a coarse part (DAC) which is configured by slow control bits and an analog fine tuning branch (current memory) to meet the bias current as close a possible. All of the branches are implemented with cascoded resistors. The cascode gates in the DAC branches are fixed while in the current memory branch, the gate is variable. The resistor in the fine branch is chosen such that ∼2.5 LSB currents of the DAC are covered providing sufficient overhead to assure that all currents can be met. The current in the fine branch is given by the voltage on the VHold capacitor because it defines the voltage across the resistor. The VHold voltage is programmed prior to each XFEL burst, referred to as the *IPROG phase*, which lasts of the order of 50 µs. In this phase, a closed loop is formed by using the main filter amplifier as an error current integrator. The residual current I_{res} of I_{bias} and $I_{subtract}$ is integrated and generates VHold such that the I_{res} eventually vanishes. In order to provide negative feedback along the loop, an inversion is needed, which is implemented by a current converter cell (ICON). Besides inverting, the cell also divides the current sent to the integrator by three to ensure enough phase margin for stability.

The current sink resistors are cascoded to increase the equivalent resistance connected at the virtual ground node, which improves the quality of the virtual ground node. To minimize the series white noise generated by the cell, the resistor values need to be as large as possible. When referred to the input, the noise generated by the resistors appear as a voltage source in series with the input. The maximum possible resistance is given by the current to be sunk and the available voltage headroom. Zero threshold voltage transistors (ZVT) have been chosen for the cascode devices in order to maximize the headroom.

To calculate the noise contribution of the DAC, the thermal noise of the equivalent resistance R_{eq} of the DAC branches given by:

$$\overline{i_n^2} = \frac{4kT}{R_{eq}} \tag{5.31}$$

needs to be referred to the input (via the DEPFET g_q). The series white noise term in equation 3.47 can then be used to calculate the noise contribution of the current sink. Taking into account the trapezoidal filtering scheme for the 4.5 MHz operating mode (corresponding to an integration time τ of 50 ns) and a DEPFET g_q of 600 pA/e^- we get an ENC = $16e^-$ rms as the contribution from the current sink. In the initial phase of the project, the DEPFET amplification was expected to be substantially lower, requiring larger resistors in the DAC to meet the noise target. This can be achieved by using a negative voltage for the ground of the current sink (VSSS in figure 5.10), thus generating more voltage headroom for larger resistors in the DAC.

To verify that the current programming loop has settled properly, the VHold voltage can be evaluated. In order to cancel the current flowing into the ICON cell properly, the VHold voltage needs to settle in the dynamic output range of the filter amplifier. If during the IPROG phase, the amplifier saturates towards the positive supply rail for instance, it means that the sink cannot handle the current. In this case, the digital code in the DAC must be increased to accommodate enough current such that the fine part can handle the remaining current. If the amplifier output voltage is too low, the DAC setting must be decreased. Since the filter amplifier directly charges the VHold capacitor, the voltage is available at the input of the ADC. With proper operation of the switches[4], it is possible to digitize the VHold voltage and store it to the memory to evaluate the mechanism off-line. The proper value for the DAC can thus be found.

5.3.3 Single Slope ADC

The output voltage produced by the FCF is sampled and digitized to 8 − 9 bit within the pixels by a full custom ADC, which has been designed by the DESY FEC group. Several papers about the design have been published: [48], [49] and [39].

The circuit uses the single slope concept, which is based on converting the analog input voltage to a timing information and further converting the timing information to a digital code. The DSSC implementation is depicted in figure 5.11. The conversion process is started by sampling the input voltage signal on one of the capacitors C_{SH}. Next, the capacitor is charged (ramped up) with a constant current until it reaches a defined reference voltage. A comparator monitors the ramping process and fires when the reference voltage is crossed. Since the ramp has a constant slope, the required time is proportional to the input signal. The digital code is obtained from a counter which is started concurrently with the ramping process. The state of the counter is latched with the comparator signal and hence provides a digital code representing the time required for the ramping process. The circuit comprises pixel internal control logic which allows for operating it with a single dynamic control signal referred to as RAMP. A conversion is started by the rising edge of RAMP and ends with its falling edge at which the generated digital word is latched in a register bank. These registers store the value until the next conversion has finished, leaving a minimum of 220 ns for the memory to write the data.

5.3.3.1 The Analog Domain

The core analog building blocks of the ADC are a precise current source which generates the ramp and a comparator to latch the digital time stamp. The overall topology and properties of these circuits needs to be carefully chosen and the preceding FCF stage needs to be taken into account in order to optimize the input dynamic range and noise contribution. Since the output signal polarity of the FCF is negative, the dynamic range is maximized if the entire topology is configured such that the input voltage corresponding to a zero signal is as high as possible. For noise considerations, it is further beneficial, if the comparator reference is close to the zero signal level. This keeps the ramping process short, minimizing the signal dependent noise contribution (see equation 5.32 in section 5.3.3.3) for small signals where it is most crucial (single photon resolution).

These considerations lead to the topology shown in figure 5.11: the reference of the ADC comparator is close to the positive supply and slightly above to the FCF amplifier reference. The current source charges $C_{S\&H}$ in the positive direction and hence sources from the positive supply. An upper limit on V_{ref} is imposed both by the upper output limit of the FCF amplifier and by the headroom requirement of the ramp current source. To provide good linearity a very high output resistance in the current source is required. An actively regulated cascode is implemented which provides this feature while requiring low headroom. Generating the ramp in the pixel is beneficial because it allows for changing the gain of the ADC, i.e. the bin size, pixel by pixel by trimming the current source. The design of the comparator is further simplified with a local ramp and a static reference because it consequently always switches at the same voltage. In the alternate topology of a global ramp and comparing against

[4] which is possible because they can freely be programmed in the sequencer (section 5.5.3)

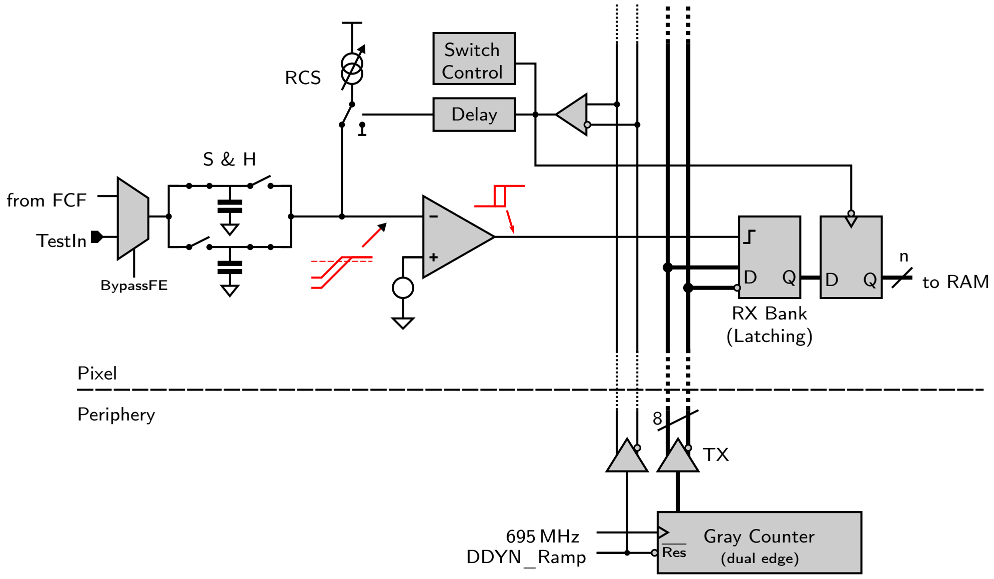


Figure 5.11: Architecture of the in-pixel ADC. The RCS (ramp current source) charges the input voltage which is alternately applied to one of the double buffering sample and hold (S&H) capacitors. A comparator indicates the end of the charging process and latches a time stamp generated by a counter in the periphery of the chip to obtain the digital output.

the input voltage signal, the comparator would need to switch at any possible voltage in the dynamic range causing more stringent common mode rejection requirements.

It is very important that the two S&H capacitors match properly to ensure an equal conversion from both capacitors. In the implementation, the capacitance is 1 pF. This size is imposed by the nominal charging current of 5 µA, which in turn has been optimized to achieve the required noise performance. The ramp current and reference voltage are generated by a dedicated temperature independent reference circuit. Two of these are placed per pixel row (depicted in figure 5.1) to follow the vertical voltage drops along the pixel columns.

At the input of the ADC, a multiplexer has been placed, through which a test signal can be applied instead of the filter output. The TestIn pin is connected to the monitor bus which connects all pixels to the internal voltage DAC and with a pad.

5.3.3.2 Digital Domain

To generate the time stamps, 8 bit counters are located in the periphery of the chip and their state is transmitted to the pixels through coplanar waveguides. A dedicated counter and differential transmitter bank is placed at the foot of each pixel column, serving the full 64 pixel column. The signals propagate on $\sim$ 14 mm long coplanar waveguides to ensure good signal quality. The counter runs continuously throughout the conversion process while it is reset at the start, concurrently with the ramping process in the pixels. The in-pixel comparator triggers a differential receiver bank which latches the counter state.

The counter uses Gray coding, which is essential in this topology to avoid serious bit errors. A Gray code has a Hamming distance of one, i.e. for each transition, only a single bit toggles. The output of the in-pixel comparator is asynchronous, the time stamp can hence be latched at any point in time. When using a counter coding scheme where more than one bit toggles for any transition, it is impossible to assure that in the physical implementation all bits toggle *perfectly* synchronous. Consequently latching intermediate states where only a fraction of the counter bits has toggled is possible which results in errors. Consider for instance an 8 bit binary coded counter: at the transition from `01111111` to `10000000`, all bits toggle and a maximum error of half the entire range is possible. Gray coding avoids this situation because there are no intermediate states at the transitions. In the baseline DSSC topology, the situation is extreme because the bits travel $\sim$ 14 mm along the pixel columns. Mismatch in the bitlines, transmitters and receivers are inevitable, Gray coding is therefore essential. Further benefits of the Gray code are that the bit transition rate is minimal which results in minimum switching noise and dynamic power consumption in the transmitters. The frequency of the LSB is further only a quarter of the equivalent state toggle frequency.

To convert to 8 bit within the given time frame of 220 ns, while providing sufficient overhead to sample the input signal, reset and initialize the involved circuits, temporal bins of 720 ps are required. The resulting clock frequency of 1.4 GHz is relaxed by dual-edge clocking. The duty cycle of the clock hence directly affects the differential non-linearity (DNL) of the circuit. A dedicated circuit in the periphery of the chip allows for correcting the duty cycle.

The receivers in the pixels are based on a double-tail sense amplifier topology which also provides a latching functionality. The circuit does not require any static bias current, it is pre-charged and evaluates the differential input signal when the latch signal transitions. Positive feedback is used to provide a very fast reaction time. After the current conversion has finished, the receivers need to be pre-charged for the next conversion, the digital output is therefore latched again in a set of flip flops to relax the timing constraints of the subsequent memory.

A dedicated logic block in the pixels requires only the RAMP signal to control the entire conversion process. It provides the functionality to safely switch the two sample and hold capacitors at the input, to automatically reset the comparator input node after the conversion has finished and to generate an error code when the conversion process has failed. A failed conversion is detected when the comparator has not fired during the time in which the RAMP signal was asserted. Using an error code instead of storing a dedicated bit signaling a valid conversion saves 10% of memory.

For test purposes, the dynamic digital signals generated by the sequencer block can be overridden using the XDATA line. Reception of telegrams during the burst can be suppressed and the `XDATA` line can be switched to selected sequencer tracks. For very fine granular *external latch* scans of the ADC digital, signals with very high granularity are required. The feature allows for generating the required signal off the chip and feeding it to the pixels instead of the on-chip sequencer signals. When running at frame rates $<$ 4.5 MHz, extended time is available for conversion which allows for generating a 9th bit. The extra bit is generated in the pixels from the Gray code MSB, requiring a second transparent receiver for the MSB. For 9 bit operation, the ramping process is slowed down by a factor of two to accommodate sufficient time for the counter.

5.3.3.3 Noise

Any electronics noise at the input and in the ADC building blocks eventually propagate to the output of the comparator and result in a timing jitter of its output signal. This timing jitter can be expressed also in a voltage noise at the input of the ADC by taking into account the gain (ramp slope) and temporal bin size. The total noise expressed in voltage power at the input of the ADC is given by [39]:

$$\overline{v_n^2} = \overline{v_{ref}^2} + \overline{v_{comp}^2} + \frac{\overline{i_r^2}}{2MC_{S\&H}^2}(V_{ref} - V_{in}) \tag{5.32}$$

It comprises two constant terms: $\overline{v^2}_{ref}$ which is contributed by the reference voltage and $\overline{v^2}_{comp}$ which is contributed by the comparator. It further comprises a signal dependent term which is caused by the integration of the current noise $\overline{i_r^2}$ from the ramp current source. Although the noise increases with larger signal amplitudes it stays negligible when compared to the Poisson noise of the photon generation process. The noise of the comparator is dependent on its bandwidth and the slope of the ramp at the input.

For the 9 bit operation mode for which the ramp current is halved, the ADC noise increases. This is due to the comparator jitter becoming more significant when related to the smaller bin size. Furthermore the flatter ramp also increases the jitter.

5.3.3.4 Gain and Offset Adjustment

The digitization concept of the DSSC system relies on heavily on the ability to tune the offset and gain of the ADC pixel-wise and very precisely (see section 4.1.2). The most attractive means of trimming the gain for the implementation at hand is to change the slope of the ramp by varying the charging current. This can be done for each pixel separately and very fine steps can be implemented which allow to fine tune the system gain to the experiment and cancel out gain variations caused by fabrication. A 6 bit DAC is implemented to adjust the current in 5% steps around the nominal value of 5 µA. For the 9 bit operation mode, bit the ramp current can additionally be divided by two. In principle, $C_{S\&H}$ can also be used to trim the gain on the pixel level but because two capacitors are used, this method is unattractive.

The delay of the comparator varies from pixel to pixel spanning more than a single bin in time. This offset cannot be coped with on-chip and must be canceled offline after readout. The in-pixel offset however must be adjusted on-chip, the system concept requires the offset to be set very precisely such that the first photons fall exactly in the middle of an ADC bin. A programmable delay between the start of the ramp and the start of the counter adjusts the offset of the ADC. 16 delay steps are implemented adjusting the offset in 10% of the bin size, while a second mode is available roughly doubling all delay steps.

5.3.3.5 In-pixel Counting

From the ADC perspective, the most critical characteristic which is required for single photon resolution, besides good noise numbers, is a low DNL, i.e. bins of equal size. The DNL of the global counting architecture is given by the grade of equality of the individual bitlines. Although a Gray code is used, avoiding serious errors, the bitlines themselves are subject to physical mismatch[5] causing signal skews which eventually result in distorting the ADC bins and hence affect the DNL. A measurement of the DNL on F1 is shown in figure 7.13, where it is evident, that the DNL degrades along the pixel columns.

A second approach has been studied in parallel, where the 695 MHz clock is transmitted to the pixel and the counter is located in the pixel. It is possible to safely avoid the Gray coded counter, costs precious area in the pixel, if the comparator output is first synchronized to the clock of the counter. In

[5]in the wires, transmitter and receivers

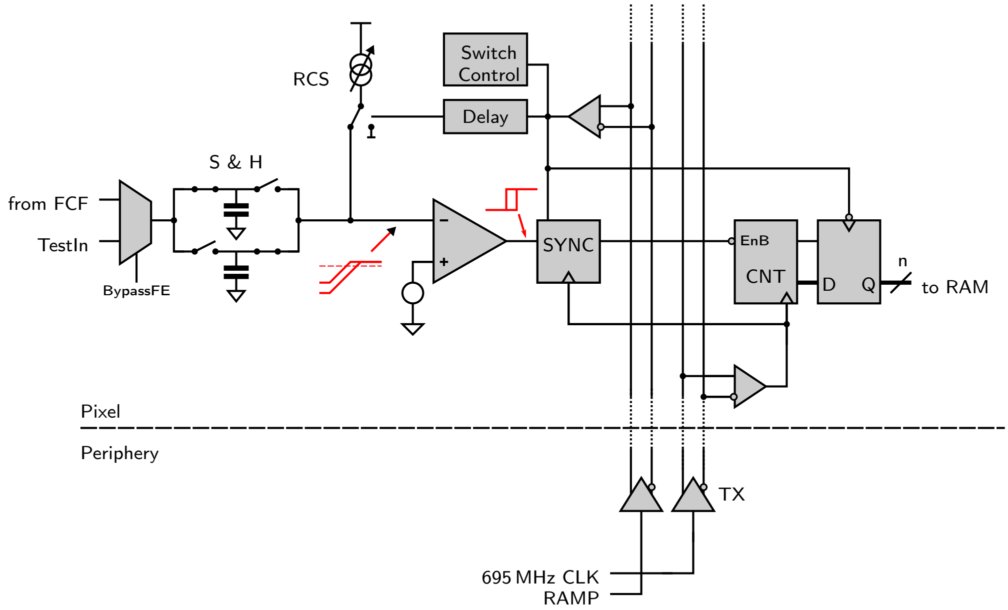


Figure 5.12: Architecture of the in-pixel counting ADC variant. The analog part is identical as in figure 5.11. A very simple ripple-counter can be used in the pixel if the comparator signal is synchronized into the clock domain before it is used to stop the counter.

this design, it is also necessary to employ dual-edge clocking, to achieve the required resolution with 695 MHz[6]. The DNL is then closely related to the duty cycle of the clock as received in the pixel (even-odd binning).

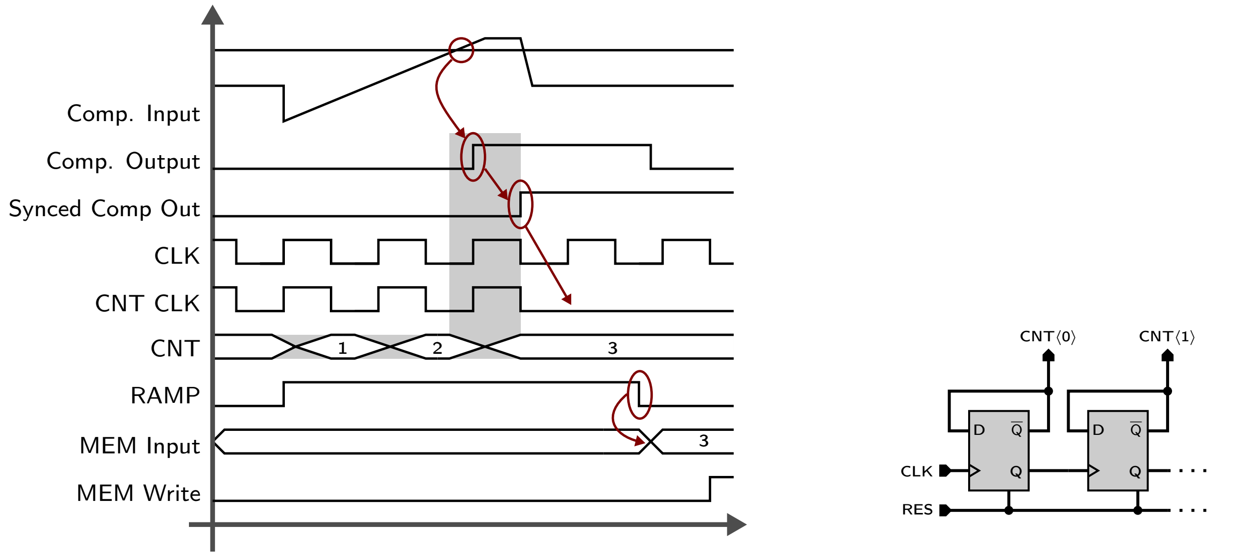
The architecture allows for using a simple ripple counter, for which only the LSB is toggled with the input clock signal, the flip flops of all more significant bits are clocked by their respective preceding bit (see figure 5.13b). The output of the comparator is synchronized to the clock before it stops the ripple counter by gating the clock of the LSB. As only the clock signal of the LSB is gated off when the comparator fires, the MSBs can still reach their proper final value. The value of the stopped ripple counter is transferred into flip-flops at the end of a conversion process. Figure 5.13a depicts this operating principle. In principle, it is possible to directly connect the output of the comparator to the ripple counter because it is basically self-clock gating, as only the LSB is driven with the clock. The synchronizer cell is a chain of two flip flops, which reduces the chance of metastability at the output.[7]

The architecture requires to transmit the clock signal to each pixel. The clock transmitter, receivers and transmission lines have to be adjusted accordingly. For the first implementation (MM3 this work has been carried out by the author starting with the existing circuits for the GCC. For improvements on the L1 test chip, work has been shared with the DESY group. The improvements have mainly targeted to improve the duty cycle of the clock in the pixels. The potential of this architecture is visible in a measued DNL map of L1 shown in figure 7.14.

Measurements on F1 have shown that the DNL degrades along the pixel column (see figure 7.13).

[6] doubling the frequency would complicate transporting the clock to the pixel even more

[7] The delay of two clock cycles produces a constant offset which can be compensated offline.



(a) The comparator output is synchronized to the clock and gates the clock of a simple ripple counter. This way it is impossible to latch wrong states of the counter (shaded grey). The grey area shows the time when the counter ripples into its state.

(b) Principle of a ripple counter. Toggle flip flops create the clock for their respective next bit in the counter.

Figure 5.13: Principle of the in-pixel counting ADC architecture.

The in-pixel counter architecture is presented here because it has the potential to provide superior DNL and remains a viable options for future iterations of the ASIC.

5.3.4 Digital Memory

The digitized data is stored locally in the pixels in a custom static random access memory (SRAM). First memory and readout concepts have been studied in [50], the concepts were further elaborated in this thesis and published here: [51]. The initial proposal was to use a three transistor dynamic random access memory (3T DRAM) cell for the core of the memory. While the layout of the 3T DRAM cell can be made very compact, it consumes much more power. The data is stored on a high impedance node which is inevitably discharged by leakage currents, requiring periodic refresh cycles to keep the data alive. These are bothersome in terms of power consumption mainly because they have to run continuously, also during the readout phase because all of the time in the XFEL gaps is needed to transport the data off the chip. Several test chip iterations have shown that the required refresh frequency becomes bothersome. The DRAM would yield a capacity increase of $\sim 20\%$ versus the final architecture chosen which is based on a dense 6T SRAM cell. We have decided that this gain does not justify taking the extra risk associated to the DRAM and therefore put it aside.

The dense SRAM cell uses special design rules which are proven for the implementation by the foundry. It is almost as compact as the custom optimized layout of the 3T DRAM cell despite for using double the transistor count and both NMOS and PMOS transistors, which imposes n-well spacing rules.

The dense SRAM cell has been used for the core of the memory while it was surrounded by full custom periphery. In principle, memory generators are available to configure the cells in an array and provide all required periphery. They generate however a large overhead by adding address decoders and peripheral circuits which allow for fast access times. This overhead is not required in the DSSC readout ASIC because access times are very relaxed (> 200 ns). Consequently, it is much more efficient in terms of area usage to use a full custom periphery and addressing scheme. While an initial topology has been studied in [50], the final architecture and readout concept has been developed within the scope of the paper at hand. A memory capacity of 800 words (9 bit) per pixel could be achieved using an area of $76 \times 229\,\mu\mathrm{m}^2$ (37% of the pixel).

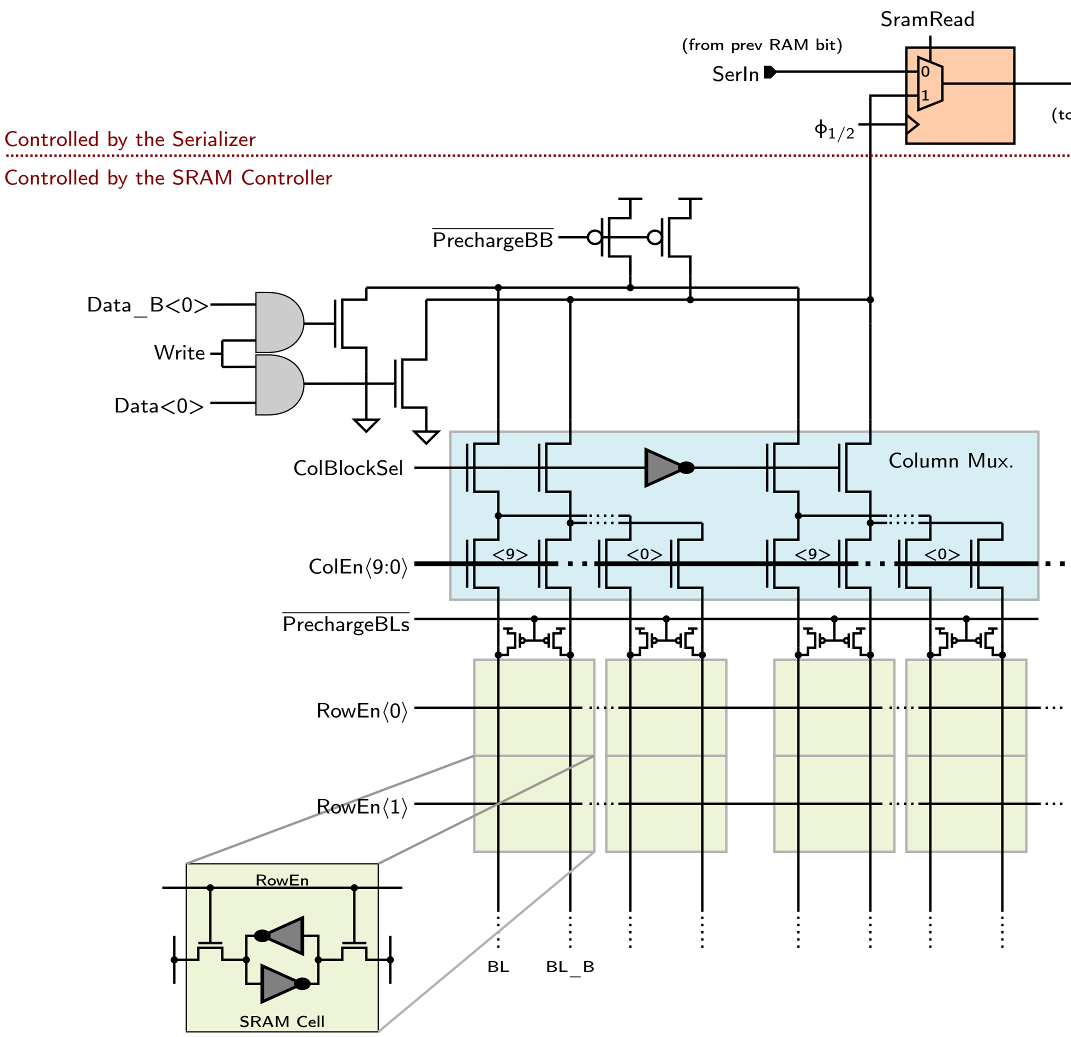


Figure 5.14: Schematic of a BitBlock. The depicted circuitry includes the memory cells (green) the column access multiplexer (blue), readout register (red) and write driver. The block is replicated 9 times to store 9 bit words.

Figure 5.15: Layout of the memory. The memory size is $76 \times 229\,\mu m^2$. The capacity is 800 words of 9 bit, the area of the in-pixel periphery is marginal.

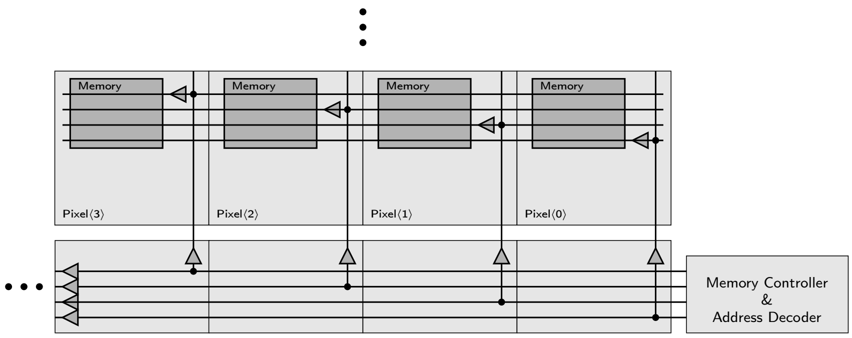


Figure 5.16: Buffering scheme for the address and control signals of the memory, exemplary for only four signals. The memory controller and address decoder are located in the periphery, signal distribution is shared among four pixels to minimize the routing in the pixel columns.

Memory Topology & Addressing Scheme

The pixel memory spans the full width of the pixel, and uses metal layers 1-3 for local routing and 4-5 for the global routing. Constraints for the layout are given in section 5.3.9.

The memory is arranged in sub blocks of so called *BitBlocks*, a schematic of a BitBlock is depicted in figure 5.14. It comprises a 40 rows $\times$ 20 columns bit cell matrix which stores one bit of each data word, a column access multiplexer and the associated peripheral circuits for reading and writing. An individual memory cell is addressed by asserting one RowEn signal to select the row and switching the column multiplexer to the desired column. The column access multiplexer is implemented by two stages of NMOS-only pass gates requiring a ColBlockSel and 10 one hot encoded ColEn signals. Although more stages of pass gates would be possible to reduce the number of ColEn signals, the layout of this topology integrated with the further periphery has given the best silicon and metal area usage. A detailed description of the SRAM working principle can be found for instance in [52]. The cell is read by precharing the BitLines and subsequently connecting the cell of interest through an NMOS pass gate. In standard memories, differential sense amplifiers[8] are used to evaluate the BitLines as fast as possible. For this ASIC, enough time is available and a simple inverter is sufficiently fast. The employed writing mechanism is based on precharge and pull down of one of the BitLines and subsequently connecting the cell. The cell is thus forced into the desired state. In order to reach the full supply rail level, both the BitBus and BitLines are pre-charged to VDD before the SRAM cell is connected.

The BitBlock is replicated 9 times to store full words. They share all control signals while they are serially connected among each other for the readout scheme (SerIn and SerOut in figure 5.14). The pixel therefore has a single serial data input pin and a single serial data output pin, which are connected to the neighboring pixels in the same pixel column. The serial readout scheme is handled in section 5.3.5. The explanation of the reading mechanism is therefore narrowed to loading the data to the serial readout register.

To further safe area, the pixels do not comprise an address decoder and one-hot encoded signals are propagated directly to the pixels. Besides these addressing lines, the memory needs eight further signals for control. The total of 58 signals however is too much to be routed in a single pixel column and is therefore shared among four pixels. This is conveniently possible because the memory spans the entire pixel width and the horizontal signal wires thus run solely above the memory and do not interfere with any other circuits. In each pixel column, one fourth of the control signals is routed which is little enough to avoid congestion. Each signal is buffered once for four pixels to de-load the vertically running signal lines. Each pixel cell therefore comprises 15 buffers, which are connected a level above the pixel cell in the hierarchy. The routing scheme is illustrated exemplary for four control signals in figure 5.16.

The testing mechanism for the memory is described in the section 5.3.5 because it makes use of the readout structures. The memory can be turned off by disabling the peripheral controller in which case only the last word written on the BitBus is held and transferred into the readout register (single frame readout).

5.3.5 Readout

The readout of the ASIC is based entirely on shift register chains. This approach was followed because it requires only minimal control overhead and the required space both silicon and metal wise is very small. As it is visible in figure 5.21, the pixel layout is very dense and the routing layers are congested. Figure 5.17 illustrates the complete readout architecture. Each pixel comprises a 9 bit wide serial shift register (red cells) which is also connected among the pixels in serial fashion. In this way, only a single data line is required along the pixel columns. The readout topology also ensures that the clock speed of the serial register can be very slow, even allowing for two phase clocking to virtually eliminate any timing constraints in the chain. The pixel registers are loaded in parallel with data words from the memory by asserting the SramRead signal and clocking the readout register ($\phi_{1/2}$ figure 5.14). The SRAM reading mechanism is explained in section 5.3.4. The data in the serial chain is next propagated to the pixel column footer cell in word chunks where it is buffered before it travels further. The column footer comprises elements of a word wide parallel shift register (blue cells) which spans across the bottom of a pixel matrix half. The data which was collected serially from the pixel columns is multiplexed into this parallel register row by row. After loading a row, the data is shifted towards

[8]which cost power and area

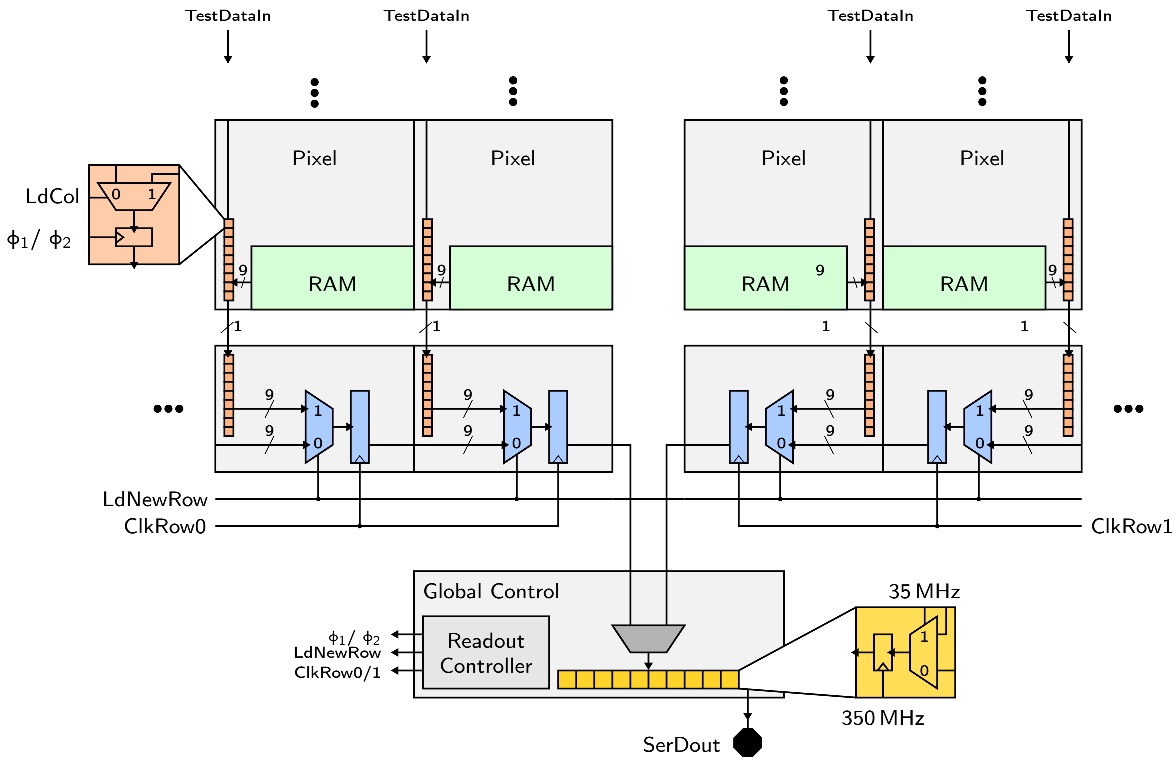


Figure 5.17: Schematic of the datapath from the pixels to the serial output link. Data is shifted serially along the pixel columns (red cells), 9 bit words are shifted along the bottom of the chip towards the center (blue cells) where the synthesized fast output serializer is located (yellow). The control logic is part of the synthesized global digital control block.

the center of the chip where the output serializer (yellow cells) is located. While the row data is being shifted, the column shift registers propagate the next row to the bottom such that it can be multiplexed into the horizontal registers. All cells in the pixel are full custom cells to ensure small size, while standard cells from the synthesis library have been used in the periphery. The output serializer is synthesized along with the global control logic to ensure proper timing closure.

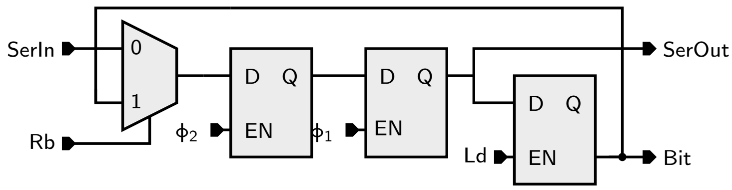
This architecture is very attractive from the control perspective: there is no need to address any pixels separately. Full scale frames are read in parallel from all pixel memories, and shifted out subsequently allowing to apply the exact same control sequence for every pixel and thus share all control signals. The topology is further very scalable because large data busses and tri-stating are avoided. Reading out single pixels or areas of interest on the ASIC with increased speed was not foreseen because it is not required by the application.

Dedicated slow control registers give the possibility to override the output of the memory address generator, memory controller and readout controller in the periphery of the chip. Each state of the controller can thus be produced by JTAG and both the memory and readout shift register chains be operated through slow control. Test data can be assigned to the input of the serial pixel column readout register. The output of this register (red cells) can be multiplexed to the input of the memory, thus allowing to write arbitrary test data to the memory. After data is loaded into the register, either a burst can be issued to fill the memory with completely identical words, or a memory write operation run through slow control. The output data from the two matrix halves (center blue cells in figure 5.17) can be read back with through a slow control register.

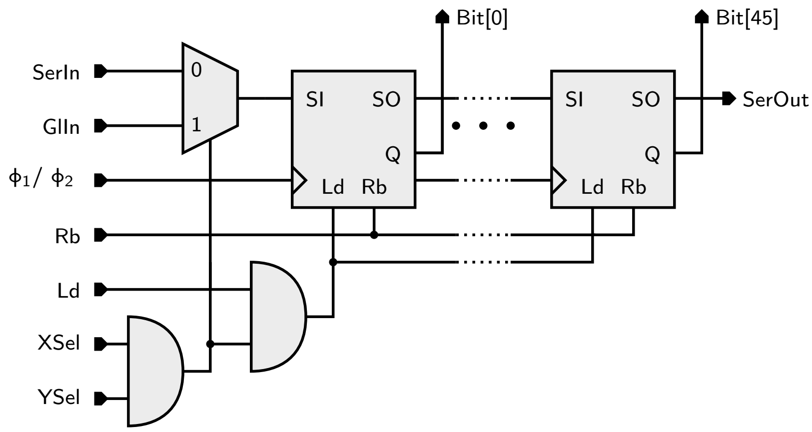
5.3.6 Slow Control

The static configuration for each pixel needs to be adjustable per pixel. Therefore, each pixel contains a 47 bit register storing the individual pixel configuration. It is implemented as a shift chain, a single cell is show in figure 5.18a. The single cell consists of a two-phase flip flop and a latch which stores the programmed data. A schematic of the complete pixel register is shown in figure 5.18b. By storing the data bit in a separate latch, toggling of the static control bits during programming is avoided. Data is first shifted into the whole chain before it is loaded into the data latch. A pixel register can be accessed *directly* or in *pixel chain mode*. A multiplexer at the input feeds data either from the output of the previous pixel in chain mode (SERIN) or from a global wire (GLIN), which is shared among all pixels on the chip, in direct access mode. For direct access, the register has to be selected with the XSEL and YSEL signals which are obtained from two separate registers located at the bottom and center of the pixel matrix (see also section 5.5.6). The two-phase clock signals $\phi_{1/2}$ are derived from TCK in the periphery of the chip, see section 5.25.

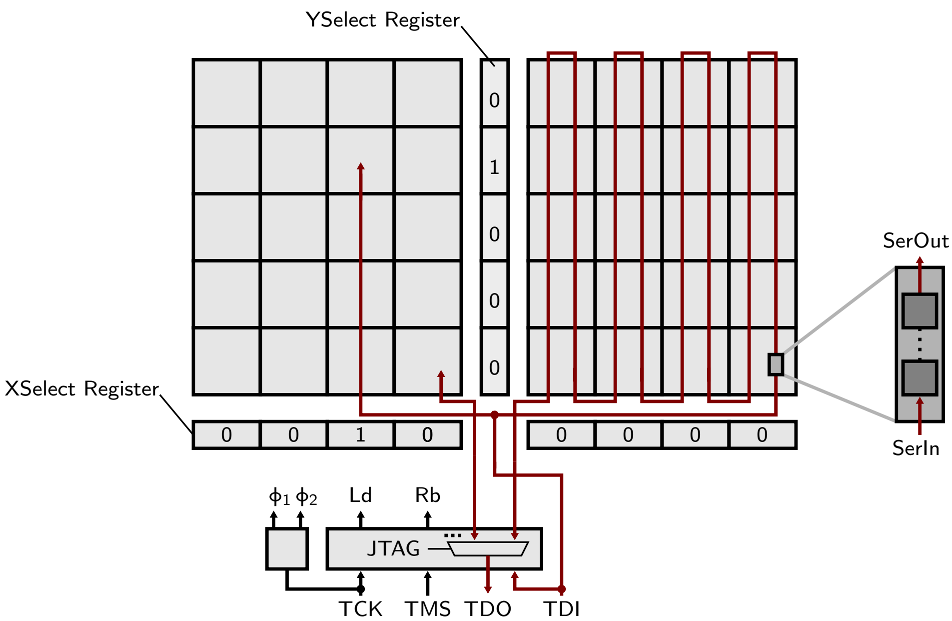
For the direct access mode, the pixel(s) of interest needs to be selected through the x- and y-select registers as shown in figure 5.18c allowing to shift data directly into the selected pixel. Both the x-select and y-select bits have to be asserted to select an individual pixel. If a pixel is not selected, its data input is connected to the preceding pixel, which allows to configure all pixel registers long chain. This is the fastest mode to program all pixels with individual configurations because the x- and y-select registers only have to be programmed once for this mode. The direct access mode is intended for quicker access when for example characterizing a single pixel. The direct access mode can also be used to program all pixels in parallel with the same configuration by loading both the x- and y-select registers with all ones. To read back the pixel register, they need to be configured in a chain, as only the output of the last pixel in the chain is connected to the JTAG TDO multiplexer. Two phase clocking is employed for the full custom register chains to manage hold timing issues in the long register chains. Especially in the long pixel columns when the data has to be transferred from the bottom of the column to the next input at the top (wire length of $\sim 14\,\mathrm{mm}$, $2.3\,\mathrm{pF}$), race conditions between data and clock can lead to hold timing violations. Employing two-phase clocking, a delay between the transparent phases of the two latches can be inserted to relax the timing issues. The two-phase clock signals $\phi_{1/2}$ are generated from TCK, the non-overlap delay is programmable from the digital control block.



(a) Schematic of a slow control register cell.



(b) Schematic of the pixel slow control register.



(c) Slow control architecture for an exemplary 8×4 pixel matrix. In the left matrix half, a single pixel is selected (through the X- and Y- select registers) while all pixels are configured in a chain in the right matrix half.

Figure 5.18: Slow control architecture in the pixel matrix.

5.3.7 Test Signal Injection Circuits

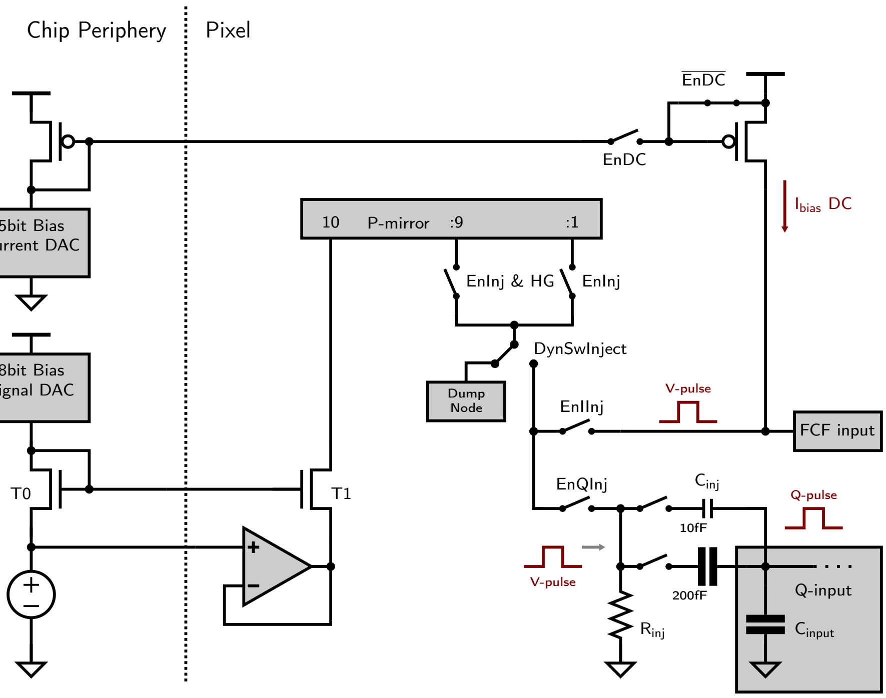


Figure 5.19: Simplified schematic of the pixel test signal injection circuit. The circuit can provide either a current signal which is directly fed into the Flip Capacitor Filter, or a charge signal which is injected into the input node of the charge sensitive front-end. The signal DynSwInject controls a dynamic switch which creates the injection pulse. All further switches are static and controlled by slow control register bits.

Each pixel comprises a circuit to inject a test signal, which has been proposed in [53] The circuit has two modes, suitable to generate a current pulse injected into the filter or a charge pulse for the MSDD front-end. The complete circuit is show in figure 5.19.

Current Injection

The signal current to be injected is generated in the periphery of the chip by an 8 bit current steering DAC. It is mirrored into the pixel using a voltage drop compensating technique. The supply and ground levels are different in each pixel due to voltage drops. In figure 5.19, transistors T0 and T1 mirror the signal current into the pixel. If the sources of these transistors were connected to the local ground lines, the V_{gs} would spread across the matrix due to the different ground levels. This situation can be avoided by referring the mirror transistors to a common reference voltage. This reference voltage is generated in the periphery of the chip and copies into the pixel with an operational amplifier. The reference line is now free of current and stable across the matrix, while the amplifier generates a low impedance copy of the node to sink the mirrored current to the local ground. Remaining mismatches in the mirror are caused by fabrication mismatches of the mirror transistors which include geometric and doping variations and a spread in the offset voltage of the reference generating amplifier.

To inject the signal current into the input node of the filter, it is mirrored again locally (P-mirror in figure 5.19) to obtain the proper polarity. This is required because there is little voltage headroom between the virtual ground input node of the filter and the positive supply (200 mV). A current pulse is generated with a dynamic switch (*DynSwInject* in figure 5.19) which sends the signal current either to the filter or to the *DumpNode*. The high level of this signal needs to span the second integration (signal sampling phase) of the flip capacitor filter. The DumpNode is a copy of the virtual ground provided by the filter (see figure 5.4). In this way, the signal current is sent to the same potential, which minimizes transient effects when switching. To provide a low and high gain mode of the current injection, the ratio of the P-mirror in the pixel can be adjusted between 10:10 and 10:1.

The circuit also includes the possibility to generate a static current (I_{DC} in figure 5.19 which mimics the DEPFET bias current when no sensor is present in lab or wafer level tests. This functionality is intended to verify the proper functionality of the bias current cancellation mechanism described in section 5.3.2.3. A voltage drop compensating technique is not required here because there are no requirements on fine accuracy.

Charge Injection for the Mini-SDD Front-End

The charge injection mode has been designed reusing the current injection circuit to produce negligible overhead in terms of area. Charge is injected into the input of the mini-SDD front-end from a capacitor. The capacitor is directly connected to the signal input node while a negative voltage pulse is generated on its backside to inject signal electrons into the input node. To generate the voltage pulse, the current pulse from the current mode is sent through a resistor. In series with the capacitance of the input (C_{input} node, the injection capacitance forms a capacitive divider, because the input node of our charge sensitive circuit is essentially floating and not a virtual ground node. Therefore not all of the charge on the C_{inj} is injected. The input transistor essentially senses the voltage step on the input node. The injected charge is calculated by:

$$Q_{inj} = \frac{C_{inj} C_{input}}{C_{inj} + C_{input}} R_{inj} I_{inj} \tag{5.33}$$

Two different injection capacitors are implemented with capacitances of 10 fF and 200 fF respectively. Three modes are available: the small capacitor can be combined with both the small and large current pulse to form the low and medium gain modes, while for the large capacitor only the large current pulse can be selected to provide the high gain mode.

5.3.8 Power Supply Decoupling & Monitoring

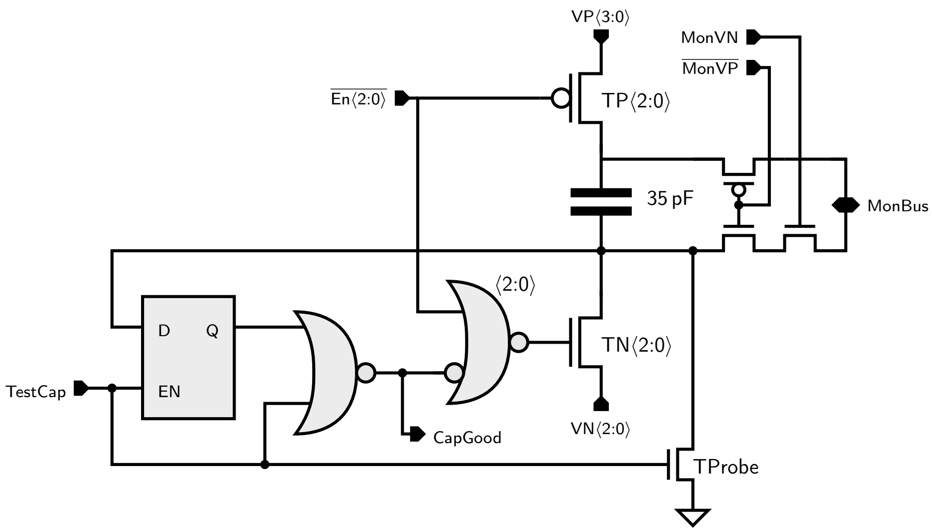
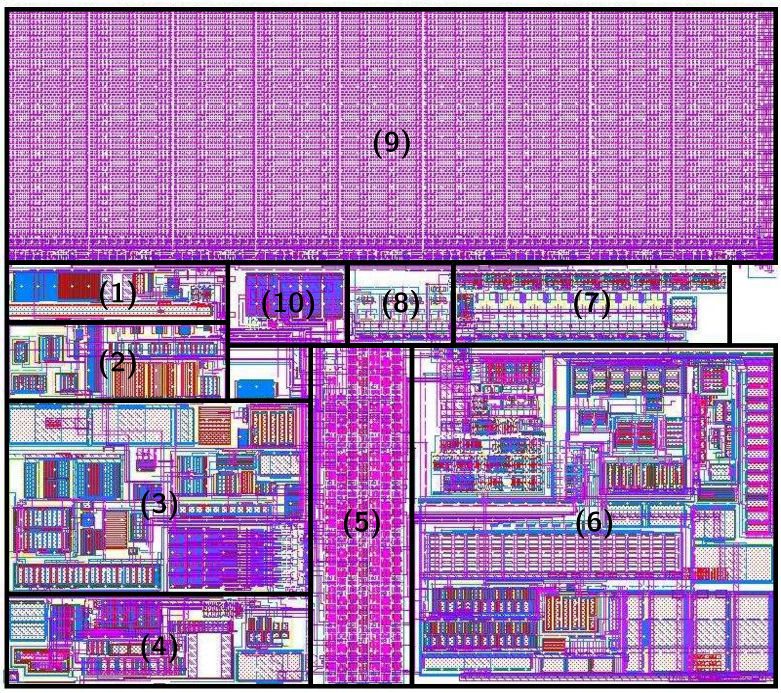


Figure 5.20: Capacitor fault detection circuit. TProbe is dimensioned such that it can sink a very small current to ground when TestCap is active. The bottom of the capacitor is discharged to ground (canceling current in TProbe) if the capacitor is good. If the capacitor is shorted, the weak TProbe cannot discharge the node. The information is latched and can be read back through slow control.

The space above the memory was used to integrate a large decoupling capacitor of $\sim$ 35 pF into the pixel. It is implemented as a metal-insulator-metal (MIM) capacitor. Due to the large area occupied by these capacitors on the chip, fabrication deficits can not be ruled out. A single shorted capacitor connected to a power supply would make an entire chip inoperative. Therefore it is important to detect and disconnect broken capacitors. A circuit which identifies broken capacitors, depicted in figure 5.20, has been integrated in each pixel. When TestCap is active (and at least one En is active), the top of the capacitor is connected to VP through TP while all TN are off. TProbe tries to draw a small current from the bottom side of the capacitor. If the capacitor is good, the bottom of the capacitor is discharged to ground. If it is shorted, the bottom side is pulled to VP. The state of the capacitor is stored by latching the bottom side of the capacitor. The CapGood pin of the circuit is read back through the slow control register. In one of the registers of the chain (shown in figure 5.18a) the read back path is cut to read back the CapGood bit. A pair of three large switches have been added to connect the capacitor to a selected power supply in each pixel. The attribution of decoupling capacitance to power supply is therefore programmable and can be adjusted to optimize performance. To monitor the supplies, a second set of switches has been added to connect the top or bottom of the capacitor to the monitor bus. The voltage drop of each pixel is thus measurable at the monitor pin of the chip.

The same cell was also used with MOS capacitors in the periphery of the chip, where a lot of silicon area was left but all of the upper metal layers were used up for power distribution.

5.3.9 Pixel Layout



(1) MSDD Front-End
(2) Bias Current Cancellation
(3) Flip Capacitor Filter
(4) Test Signal Injection
(5) Slow Control Register
(6) ADC
(7) Time Stamp Receiver
(8) Dynamic Control Receivers
(9) Memory
(10) Decoupling Capacitor Switches

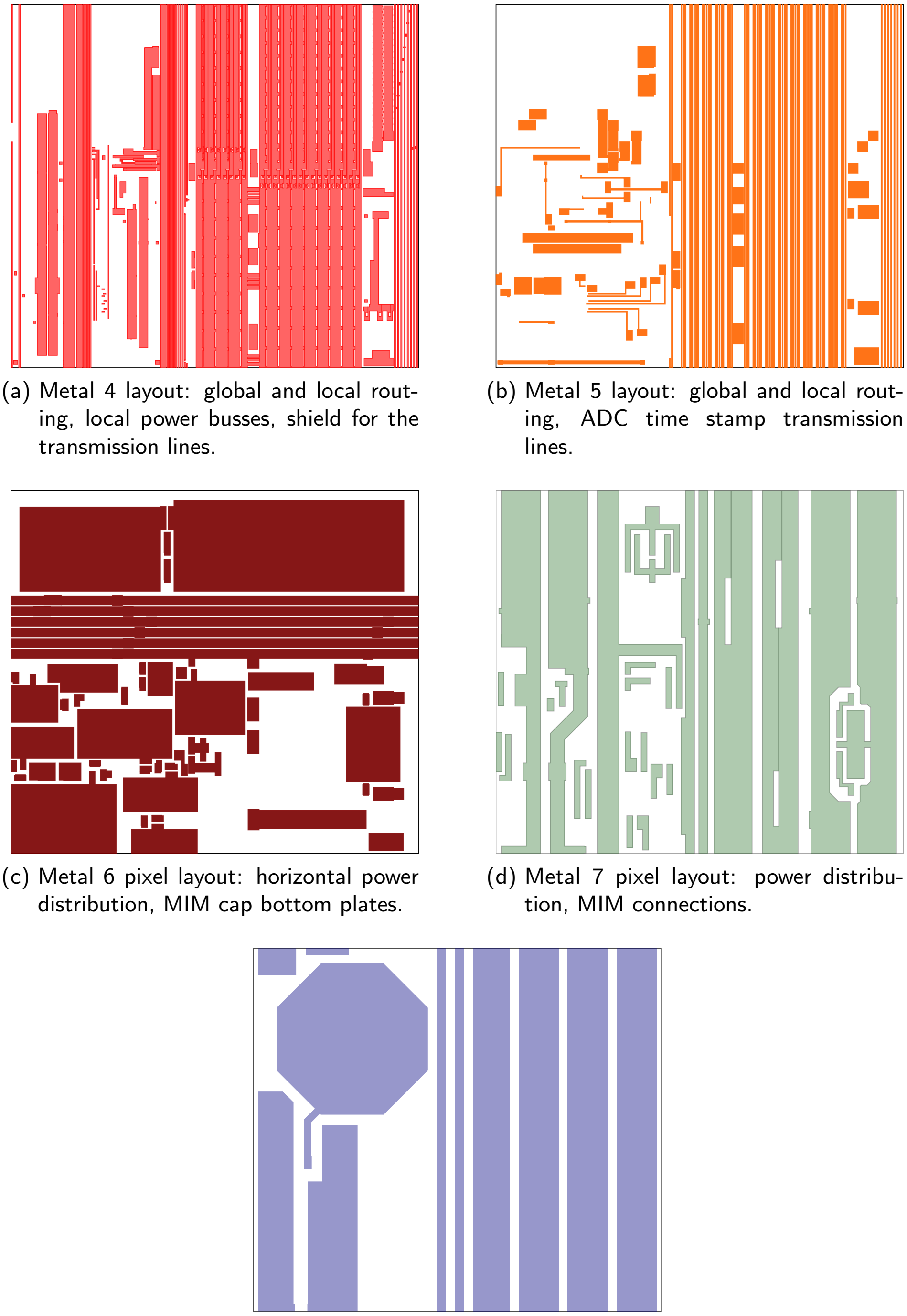
Figure 5.21: Pixel layout in detail including up to metal 3.

The pixel has a size of $204 \times 229\,\mu\text{m}^2$ and has to accommodate the front-end circuits, including test signal injection, pixel part of the ADC and memory. The size of the 47 bit control register is further non-negligible. Figure 5.21 shows the final arrangement of all blocks including local wiring up to metal 3, while figure 5.22 shows metal layers 4-8 which comprise global signal and power routing.

To plan the placement and geometries for the various pixel blocks, we need to take into account the interconnection topology and the usage of MIM caps. The following constraints have led to the final pixel layout:

(1) The width of the power busses needs to be maximized in order to minimize the vertical voltage drops along the pixel columns. As the sheet resistance of the two uppermost metal layers (7 and 8) is substantially lower than for the lower layers, they are predestined to accommodate the power busses. On the top metal layer (8), a large mandatory octagon is required to establish the connection to the solder ball which connects to the sensor pixel. This polygon blocks substantial area from power bus routing.

(2) The transmission lines which transport the ADC time stamps to the pixel use metal 4 for shielding from the bottom and metal 5 for the signal traces blocking these layers for any continuous horizontal wiring and direct access to any MIM caps on top. The transmission lines should not run beneath the bump in order to ease access to the bump and avoid pick-up on the input node. The time stamp receivers should be placed directly under the respective transmission line.



(a) Metal 4 layout: global and local routing, local power busses, shield for the transmission lines.

(b) Metal 5 layout: global and local routing, ADC time stamp transmission lines.

(c) Metal 6 pixel layout: horizontal power distribution, MIM cap bottom plates.

(d) Metal 7 pixel layout: power distribution, MIM connections.

(e) Metal 8 (top) pixel layout: bump landing pad, power distribution.

Figure 5.22: Layout of the upper 5 metal layers.

(3) The MIM caps can only be accessed through metal 7, which is the best layer in terms of sheet resistance. It is best to avoid placing MIM cap connections beneath the bump to preserve the continuity of the power busses. Metal 8 is blocked by the bump and thus metal 7 should be maximized in width here. Since the front-end makes excessive use of MIM caps[9] and the interconnect distance should be minimized, it cannot be placed directly under the bump.

(4) As outlined in section 5.3.4, the memory shares its control signals with the neighboring pixels, therefore requiring horizontal routing. Layers above metal 5 are too coarse in terms of spacing rules for signal routing. Taking into account (1), metal 3 needs to be used for horizontally connecting the memory control. The optimum shape for the memory is therefore such that it spans the full pixel width. This way, the control signals run only on the top of the memory, avoiding interference with other blocks and routing congestion.

(2) leads to the placement of the ADC on the right side of the pixel. (3) voids placing the front-end beneath the bump. Taking into account (4) leads to the memory being placed in the upper part of the pixel. This yields the final pixel layout shown in figure 5.21, which has almost no dead area. The shape of the control register is the most variable as it is a replication of equal cells. The register is placed in between the front-end and ADC in the center of the pixel, in order to ease access to the registers. The free area above the memory is used to place a large decoupling capacitor which is connected in the middle to comply with design rules and not interfere with power bus routing. On top of the front-end, power is routed in metal 7 such that vertical corridors remain to connect the MIM caps. In total, the pixel features 12 MIM capacitors with a total capacitance of 57.4 pF. The area which has to be given up for the corridors is compensated by a partial metal 8 trace which is blocked from continuity by the bump hexagon.

In the initial pixel floorplan one third of the pixel was reserved for each of the major pixel blocks of front-end, ADC and memory. The initial target size of the memory was 512 words, its layout has been arranged such that it is very flexibly sizable. Due to careful initial floorplanning and compression of all layouts, the F1 pixel has a reached capacity of 800 words, despite the late addition of the MSDD front-end.

5.4 13 bit Rail-to-Rail Voltage DAC

The periphery of the ASIC contains a 13 bit rail-to-rail voltage DAC. The output of the DAC is connected to the monitor bus (MonBus) which is in turn a global wire connected to all pixels. The DAC has two functionalities: it provides the reset voltage for the MSDD front-end and it serves as a test input signal for the ADC.

A simplified schematic of the circuit is depicted in figure 5.23. The design is based on a large array of 8192 parallel current sources which are switched on by the digital code (current DAC). The generated current is converted to a voltage by sending it through a resistor. Two modes are available here: the high range (HR) mode and the low range mode (LR). In the low range, the current from the DAC is directly send through a resistor (R1) to ground (green path), the minimum output voltage is thus ground and the maximum cannot reach the positive supply rail in this mode because the current sources need sufficient headroom ($\sim$ 400 mV). In the HR mode, the current is first mirrored and then drawn from the positive supply rail (blue path). Complimentary to the LR mode, the maximum output

[9] In total, the F1 front-end features 9 MIM capacitors with a total capacitance of 20.4 pF.

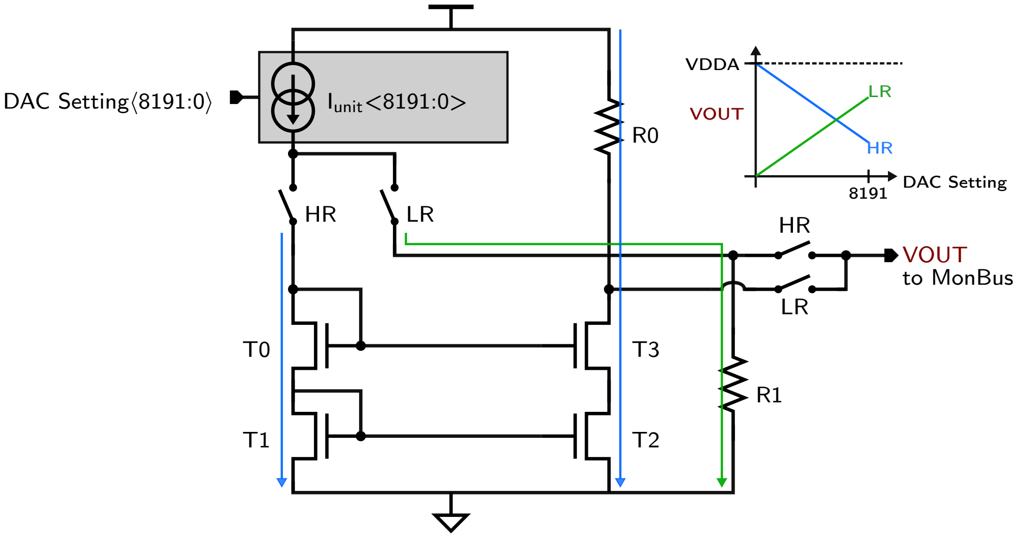


Figure 5.23: Simplified schematic of the DAC. It is based on a current source array, the output voltage is generated by resistors.

voltage is the positive supply rail while the minimum cannot reach ground. The design values for the unit current source is 60 nA, while R0 = R1 = 1.65 kΩ. The reference current is generated internally by a dedicated circuit.[10] The bin size is therefore 99 µV, at the nominal bin size of the pixel ADC of 3.125 mV, this yields 31.57 DAC steps per ADC bin and the output swing for each mode is thus 811 mV. R0 and R1 of course need to have good matching in order for the two modes to match in the bin size.

The core part of the current source array has been available in our research group in the UMC 180 nm technology and ported to the IBM 130 nm used for this ASIC. It comprises 1024 unit current sources (CS) which are switched. Without any further division for more LSBs, these yield 10 bit. The upper 7 MSBs are decoded into a thermometer code, each bit controls a set of 8 CS. The 3 LSBs of the 10 bit binary code code directly control 4 CS, 2 CS and one CS. Using this scheme, a single CS is now remaining, which can be further divided to improve the resolution. A test chip of this 10 bit core structure has however shown, that the structure is already limited at 10 bit resolution due to the physical matching of the CS.

For the F1 chip, this limit has been overcome by using the core cell eight times in parallel. The digital input code is shared for all of the core cells, while the remaining LSB cell for each core is attributed a dedicated digital input signal. The eight LSB cells can thus be used for 3 more DAC bits but only two if they are switched on in strict common centroid fashion.

To characterize the ADC, the ADC input can be switched to the MonBus and thus connected to the DAC. For this direct connection of the DAC, only few pixels can be connected because the output impedance is two high to charge the S&H cap for a lot of pixels in parallel in a short time. A second mode (depicted in figure 5.24) is available where the DAC voltage is buffered in the pixel using the FCF amplifier. It can be put in permanent reset and when the DAC voltage is applied to its positive terminal acts as a buffer for the DAC voltage. This is the mode in which the ADC measurements

[10] The same reference circuit which is located at the side of the pixel matrix is used here.

(section 7.4.2) have been conducted. A measurement of the complete DAC characteristic is shown in section 7.3.

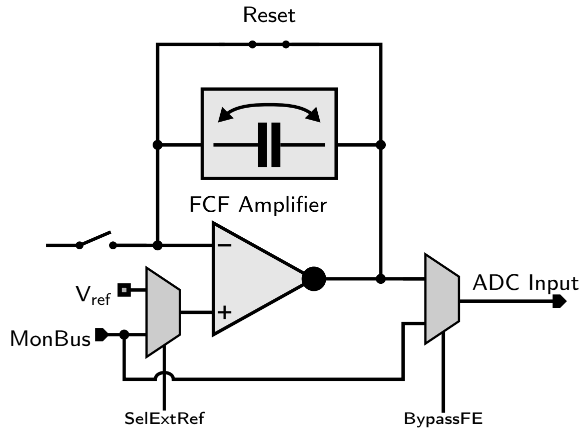


Figure 5.24: The DAC voltage is distributed to the pixels using the global MonBus. It can be directly applied to the ADC input or buffered with the filter amplifier to hide the high output impedance of the DAC.

5.5 Global Digital Control

The global digital control blocks provides dynamic steering signals for all blocks on the chip, including pixel and global circuits, as well as a slow control interface. The dynamic control interface is minimalistic, only two signals along with a further fast clock are needed. These signals are implemented in the LVDS standard to minimize switching noise. Slow control is implemented by a standard JTAG interface requiring four CMOS signals. In the final system, 16 ASICs will be connected in a single JTAG daisy chain. An asynchronous reset is used to initialize all state machines. Each of the core modules comprises a configuration register which is accessed through the JTAG interface. A block diagram of the entire control logic is depicted in figure 5.25. The following subsections present the implementation of each sub-block, while the choice for the different clock speeds is explained in section 5.5.1.

5.5.1 Clocking

The functional elements in the design are all synchronous to the 695 MHz master clock. To simplify the timing constraints and save power, it is divided into several sub clock domains (figure 5.25). The reason for the clock speeds of the individual sub-blocks are:

- 695 MHz: The pixel electronics, in particular the flip capacitor filter, require finely granular control signals. The 695 MHz clock is therefore used undivided to generate these signals. This clock is mainly available because it is needed for the ADC.
- 100 MHz (700 MHz/7): This is the biggest sub-domain and includes all of the logic controlling the burst operation. 50 MHz would be the lower limit because it provides the lowest integer fraction to implement the 1 MHz mode (11 clock cycles). 100 MHz has been chosen to provide some more flexibility and is a reasonable speed for the implementation technology.

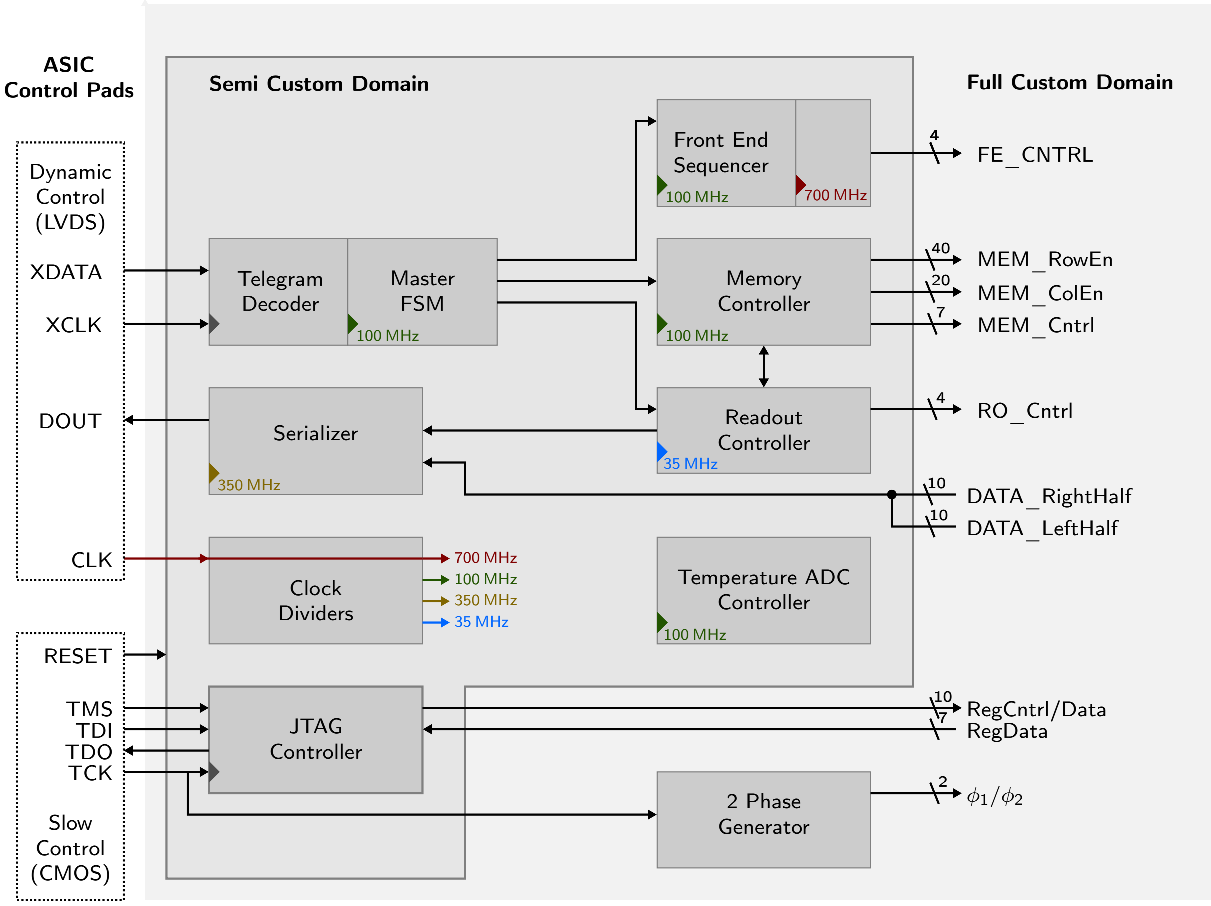


Figure 5.25: Block diagram of the semi-custom on-chip digital control logic. The various clock sub-domains and interactions between the modules are indicated. Each of the core modules comprises their own configuration register accessed through the JTAG interface (not shown for simplicity).

- 350 MHz is the slowest integer fraction of 700 MHz which allows to transmit all data off the chip in < 10 ms through a serial link at single data rate, i.e. one transferred bit per clock cycle.
- The implemented on-chip readout mechanism which transports the data from the pixels to the serial output link allows to use a very slow clock. 35 MHz was chosen because it seamlessly provides a 1/10 serialization ratio for the output data stream and basically eliminates any timing issues.

The telegram decoder is clocked with a separate external clock (`XCLK`), providing flexibility for the further module design because the frequency can hence be chosen rather freely. The decoded command is synchronized with the core 100 MHz clock and forwarded to the Master FSM (Finite State Machine).[11] The JTAG interface is also clocked separately to decouple slow control from the rest of the design. All slow control registers are static for the functional operation mode eliminating any timing constraints to the functional clock domain.

[11] In principle, `XCLK` could be omitted because it is synchronous to 695 MHz.

5.5.2 Dynamic Control

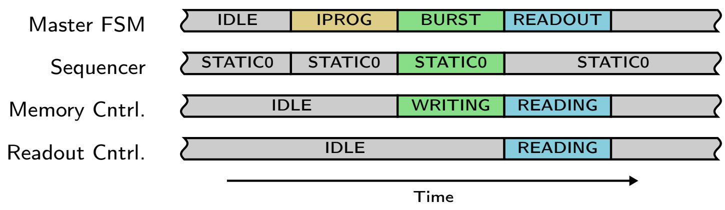


Figure 5.26: State diagram of the blocks in the digital control block (time not to scale).

Dynamic control of the chip is implemented by a simple telegram interface and a finite state machine (Master FSM in figure 5.25) which represents the coarse controlling instance of the chip. The states are depicted in figure 5.26, in addition there is a state in which the chip sends a defined programmable test pattern to calibrate the receiving FPGA to the data stream. Transitions from the IDLE state are triggered through the custom command telegram interface. It provides flexibility using only two LVDS signals - XDATA and XCLK.

The XDATA line is sampled by XCLK, it is low while it is idle while the start of a telegram is signaled by a rising edge on XDATA. The start bit is preceded by 4 telegram bits. The available commands are listed in table table 5.2. When receiving a BURST command, the state machine triggers the bias current programming phase for the pixels before proceeding into the burst state where data is taken. During the burst, it accepts VETO commands which discard selected events from the burst (see section 5.5.4). Slow control registers provide the possibility to adjust the length of the current programming phase, the burst length (number of total taken events), and the measurement cycle length i.e. the frequency of data processing. The measurement cycle length is used to tune the ASIC to the pulse frequency of the XFEL. In fact, it only controls the frequency of the SRAM write operations, all further dynamic control signals required to operate the front-end are generated by a dedicated sequencer block. The sequencer must be tuned to the XFEL operating frequency separately. In the READOUT state, the readout and SRAM controllers are triggered to transmit all data off the chip. When moved to the TEST_PATTERN state, the chip continuously sends a programmable test pattern to calibrate the input delays of the receiving FPGA to the data stream.

The state vector of the state machine can be overridden by slow control, permanently moving the state machine into the desired state. This feature can be used to operate the chip through slow control.

Bits	Command Name	Function
10000	START_BURST	Starts a new burst.
10001	START_READOUT	Starts the readout.
10010	VETO	Vetoes an event with a fixed latency.
10011	SEND_TEST_PATTERN	Starts sending the test data pattern.
10100	STOP_TEST_PATTERN	Stops sending the test pattern.

Table 5.2: Command Telegrams. One bit has been left for reserve.

5.5.3 Front-End Sequencer

The analog front-end and ADC in the pixels require several dynamic digital control signals (see section 5.3). The timing of these signals needs to be finely adjustable to optimally exploit the available time between two events. The sequencer must further be flexible enough to cover the different timing structures of the target applications at XFEL and be able to generate modified patterns for calibration measurements. Several calibration measurements require extending the flattops length[12] to inject test signals. These include for instance calibration measurements using radioactive sources for which the arrival time of the signal is not known and the sensitive time must hence be increased. We have used a very flexible approach here in order to be able to program the sequences very freely as it is not uncommon that use cases arise which were not expected at design time.

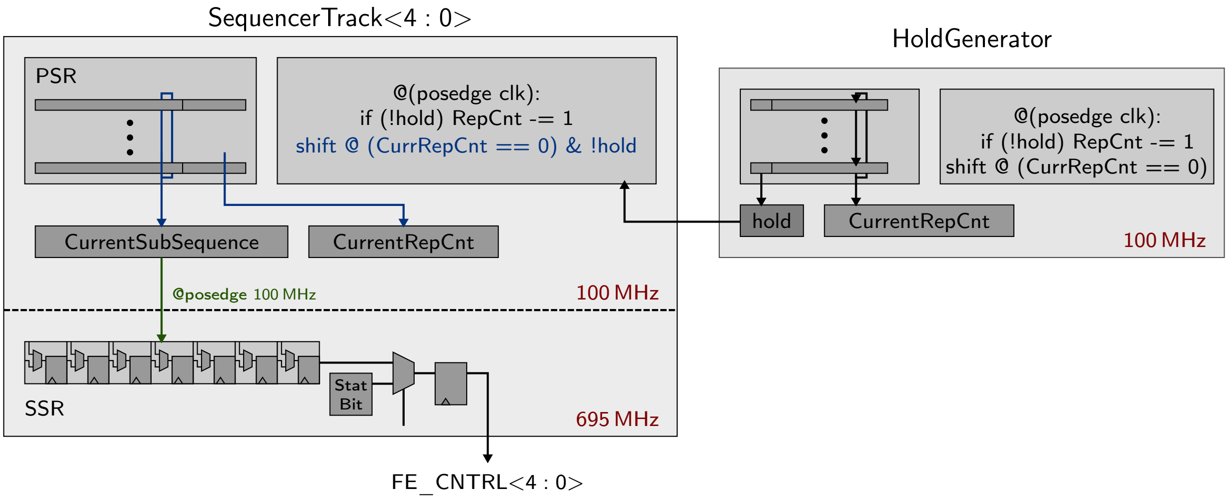


Figure 5.27: Architecture of the sequencer module. There are five SequencerTracks which generate cyclic programmable bit patterns to control the analog front-end and ADC in the pixels. A cyclic fast shift register (FSR) serializes sub-patterns running in a cyclic slow shift register (SSR). All tracks share one HoldGenerator instance.

To provide very fine granularity, the full 695 MHz clock is used in the output stage of the module, while a slower clock is used internally to relax the timing constraints. The block comprises five identical tracks. The general working principle of sequencer is depicted in figure 5.27. The basic building blocks generating the sequence are two shift registers:

- 7 bit wide sub-patterns are cyclically rotated in the **slow parallel shift register (SSR)**, clocked at 100 MHz. Each sub-pattern is associated with a 4 bit repetition count, which specifies for how many clock cycles the current sub-pattern is repeated until the shift register is advanced.
- A **fast serial shift register (FSR)** clocked at the full 695 MHz to serialize the current sub-pattern. The FSR is loaded with the current sub-pattern at 100 MHz. The phase of the load signal is shifted by several 695 MHz clock cycles against the 100 MHz clock to relax the timing constraints.

[12] of the trapezoidal WTF, i.e. the time between baseline and signal sampling of the FCF

In the implementation, the SSR is 14 entries deep which allows to program patterns with a maximum of six completely arbitrary transitions, while further patterns are possible with certain restrictions. A fast pulsing sub-pattern for instance can be generated if the sub-patterns contain multiple transitions.

An additional module, the so-called *HoldGenerator* is shared among all tracks and provides a mechanism to hold the sequence static at certain spots. This module is also based on cyclic shift registers which contain a *hold* bit and an associated counter value. An asserted hold bit stalls the advancement of the SSR and its respective counter state in all of the tracks.[13] A 16 bit counter is implemented providing hold phases of up to 65 µs.

For programming the sequence, all of the registers in the SSR and the HoldGenerator have a shadow configuration register which are themselves configured in a chain and accessed through JTAG. This approach costs a lot of registers which could be saved by configuring all registers in one long chain for programming. The additional shadow registers however simplify the programming mechanism and the required extra space is tolerable.

The sequencer contains an additional feature which allows to generate a pulsing sequence on dedicated tracks during the hold phase. It would have been required for controlling an on-chip inner substrate injection circuit for the DEPFET. The on-chip pulser has become obsolete because the charge injection scheme is not compatible with a large DEPFET matrix. The required control mechanism was added quite simply: by making the polarity of the hold bit programmable per track, they can be configured to run and hold in complimentary fashion. This very simple addition provides the possibility to generate a pulsing sequence on dedicated tracks when the other tracks are in the hold phase.

5.5.4 Memory Controller & VETO Mechanism

The full custom SRAM (static random access memory) block in each pixel is controlled by a dedicated module. A state machine generates the sequence of control signals required to write or read data from the memory. As the capacity of the in-pixel memory is smaller[14] than the length of the XFEL burst, a mechanism is needed to discard uninteresting events on the fly. A fixed latency mechanism has been implemented which is triggered by sending a `VETO` telegram to the chip (see section 5.5.2). The fixed latency is programmable through slow control and has a maximum length of 128 events. When the latency is programmed to 126 for instance, and a `VETO` telegram is sent to the chip at event #226, event #100 will be discarded. The implemented mechanism is based on a shift register with programmable length, which stores the order of used memory locations. The latency length configures the length of the shift register. If no veto is present, the write address is taken from a simple counter which is incremented for each event starting at 0. For each written event, the memory address is shifted into the shift register. Thus, the output of the register always points to the memory location which was written *latency length* events before. This location is to be discarded in case of a `VETO` and is overwritten immediately. The current write address is reinserted into the shift register, which also allows to overwrite this memory location again. Once the veto latency has passed, the event is frozen in the memory and cannot be discarded anymore. Consequently, the order in which the recorded events are located in the in-pixel memories at the end of the burst is not in chronological order. There

[13] In principle, clock gating could have been used to stall the SSR. Safe clock gating however requires special clock gating cells which are not available in the implementation standard cell library and has therefore been avoided. The power savings would anyway only be very marginal.

[14] The memory capacity is smaller for XFEL operating frequencies > 1MHz. At 1MHz, all 600 events can be stored.

is no mechanism in the chip to attribute an event ID to the SRAM entries, the order is reconstructed externally (see section 4.3). For debugging, the chip only stores the number of vetoes processed. This information is forwarded during the readout to the receiving FPGA in the trailer of the data stream to check the proper acceptance of all issued vetoes.

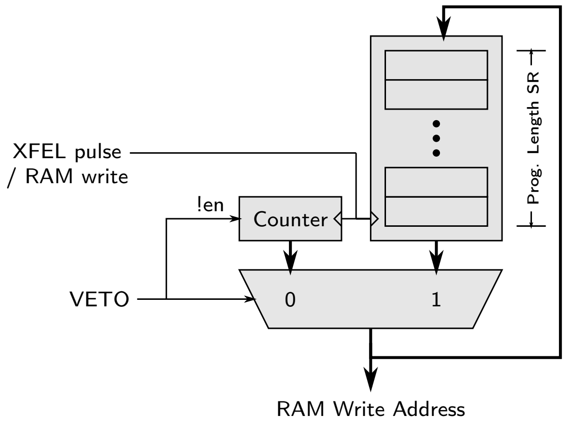


Figure 5.28: Schematic illustration of the fixed latency VETO mechanism. In case of a VETO, the next address for writing is retrieved from a shift register which has the length of the veto latency. To overwrite a past event, the event is selected by sending the VETO telegram at the right time. The written address is always inserted into the register.

5.5.5 Readout Controller and Serializer

The minimum required speed of the serializer is given by the amount of data which has to be transported off the chip during the XFEL burst gaps of 100 ms. The chip accumulates 30 Mbit[15] of data for one burst, hence the output stream of the serializer must be at least 300 Mbit/s. Using half of the 695 MHz ADC clock fits very nicely and allows to add a parity bit to each word yielding a serialization ratio of 1/10. The parallel shift register at the bottom of the pixel matrix therefore needs to be clocked at a modest 35 MHz to continuously deliver data for the output serializer. The 35 MHz clock is also used for the two phase serial shift chain.[16]

The readout controller takes care of steering the data from the pixel memory to the serial data output of the chip through the various shift register chains depicted in figure 5.17. The readout is triggered by sending the chip a Readout telegram. The master state machine then triggers the readout and SRAM controllers. During the readout, the SRAM controller generates the strobes necessary to load the data from the pixel memories to a pixel internal bus. The readout controller then latches the data into the pixel internal readout register which is subsequently switched into a long chain spanning the full pixel columns. The data is shifted serially along the pixel columns in chunks of 9 bit. A 9 bit wide shift register collects the data at the bottom of the pixel matrix and transports it towards the serializer. Two 9 bit wide data streams clocked at 35 MHz arrive in the center of the chip where a checksum bit is added and they are serialized at a ratio of 10-1. Data is sent off the chip at 350 MHz

[15] 4096 pixels $\times$ 800 words/memory $\times$ 9 *bits*/*word*
[16] In fact, four clock cycles are used for the two phase clocking to assure non-overlap and it could still be slower.

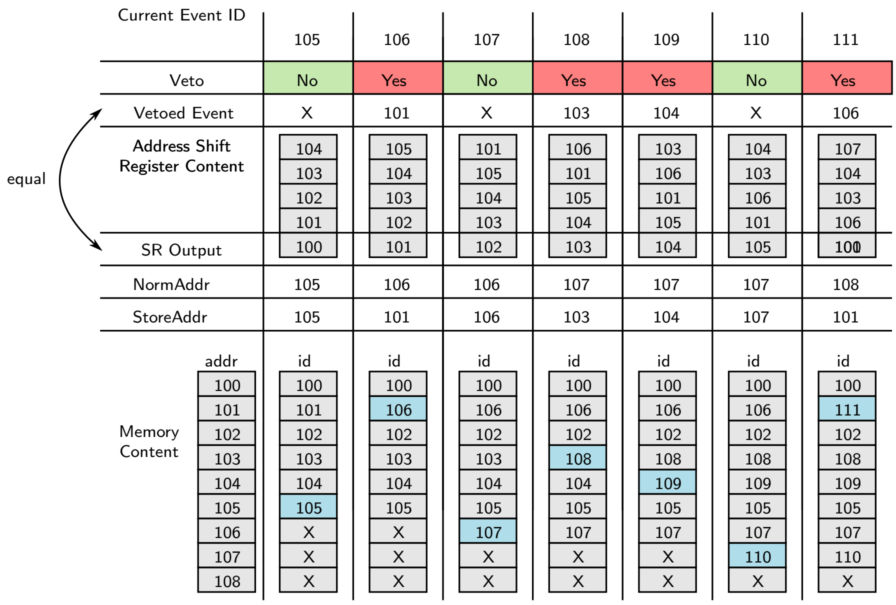


Figure 5.29: Example of the VETO mechanism. The shift register output points to the memory location which is freed and overwritten in case of a VETO. If no veto is issued, data is written to consecutive memory addresses generated by a simple counter.

(see section 5.5.1 for details about the clock frequencies). The start of the data stream is signaled by a rising transition on the data link, it is always low when it is idle. The first transmitted data word is composed of a leading one appended to a known programmable first word, to force a rising transition on the output. The raw data from all pixels is sent next, followed by 7 trailer words. Each 10 bit words carries 9 bit of data from the pixels appended by a checksum bit. The readout is not destructive, the same data can be read out several times by issuing several `READOUT` telegrams. This feature is useful for small lab setups which are not able to transmit the full bandwidth of the ASIC. Several readouts can be done each transmitting different chunks of the full data stream.

5.5.6 Slow Control

All slow control registers are accessed through a standard JTAG (Joint Test Action Group) interface. While other protocols like SPI (Serial Peripheral Interface) or I^2C would have been applicable providing equivalent functionality for this use case, JTAG has been favored without a clear argument. While JTAG is mainly intended as a debugging and testing interface defining several standard mechanisms such as boundary scan it can also elegantly be used as a configuration interface. The JTAG TAP (Test Access Point) controller provides a simple access interface for shift registers using only four chip pads (`TCK`, `TMS`, `TDI` and `TDO`). A daisy chain of 16 chips is formed by connecting `TDI` and `TDO` among chips. `TCK` and `TMS` are broadcast signals. The core of the JTAG module is a state machine which is referred to as the TAP (Test Access Port). It is steered only by `TMS` and `TCK`, and multiplexers to select different registers. `TDI` is sampled on the rising edge of `TCK` while `TDO` is launched with the falling edge of `TCK` according to the JTAG standard. This mechanism allows for resolving both setup and hold issues between chips where the timing delays are rather uncertain by adjusting the clock frequency. The TAP controller manages read back, shifting and loading of the register chains. Each register is allocated with an address. To access a register, the according address is written into the JTAG instruction register. The selected register is then serially fed from the `TDI` pin while its output is multiplexed to the `TDO` pin. Verilog code for TAP controller has been available in our group and could be reused, while the surrounding structures were implemented according to need. Each of the core modules in the digital control block has its own configuration register, which is included in the in the individual block. The JTAG interface was synthesized with a target frequency of 50 MHz, which is fast enough to configure 16 ASICs in a daisy chain in between two bursts at the EuXFEL.

The full custom domain of the chip features seven full-custom two-phase clocked registers. For each of the pixel matrix halves there are three registers: pixel register, x-select register and one global register, while the two halves share a common y-select register. The global register holds the configuration bits which are used at the bottom of the pixel matrix. The pixel register provides the local configuration for the pixels (see section 5.3.6) and can be accessed in *direct* or *pixel chain mode*.

5.5.7 Debugging and Testing Features

A lot of care has been taken, that all of the individual blocks can be tested and characterized separately using on-chip circuits. These features are mandatory for efficient wafer level testing. The according implementation has been added in the sections covering the respective block. Only a list is given here to summarize all of the testing and debugging features which is intended to serve as a reference.

- The Master FSM can be overridden by a slow control register to operate the chip through slow control.
- The memory address generator, memory controller and readout controller can all be overridden by slow control registers, allowing for memory and readout testing through slow control.
- The memory controller can be turned off. In this mode, the chip can only take one event and must be read out afterwards. This mode is intended to evaluate the cross talk of the memory.
- The ADC comparator output can be overridden with a digital signal from the sequencer (*external latch* feature). This way, the digital domain of the ADC can be tested separately. The sequencer tracks can be overridden with the XDATA chip input signal[17] to generate this signal externally with higher granularity.
- There is a charge signal injection circuit for the MSDD front-end and a current injection circuit for the DEPFET front-end.

5.5.8 Implementation

The standard approach to implement digital designs is the so-called *semi custom design flow*. The *semi* residing in this concept stems from the usage of so-called *standard cells* which are used as basic building blocks for such a design. Standard cells can vary in their complexity, ranging from very simple cells such for example a *NAND2* up to complex building blocks such as memories or even full microprocessors. Special software tools, are available which mostly automate the procedure of implementing a design physically starting from a behavioural high level description of the design. The design flow is usually scripted in order to facilitate and reproduce the entire procedure. Proper handling of the tools can be tedious and time consuming but this experience is inevitable for an engineer working in this field. Only

[17] LVDS telegram line, telegram feature can be muted for the duration of the burst

the general flow is therefore presented here shortly while the reader is spared from too much technical detail. This entire process is referred to as the *back-end implementation*.

For the IBM 130nm process we are using, there is a standard cell library from CERN, which has been used for the implementation of the digital control block. Since our group (*Chair of Circuit Design and Simulation*) has access to the latest tools from *Cadence*, these were used for the entire process.

The process of transferring the design from the behavioural description (register-transfer-level or rtl-netlist), most commonly in Verilog or VHDL, to a mapped netlist is called synthesis. The term mapped in this case refers to the fact that the produced netlist consists entirely of standard cells of the target library. The input data for this step includes the rtl-netlist, the target standard cell library including associated timing properties and a set of constraints. The constraint files specify the boundary conditions of the design, including for instance the timing specification for the input and output signals and the clock speed. It is further possible to divide the design here into so called *modes*. A mode is defined by its operating purpose and the different modes exclude each other. The design under consideration was divided into two modes, a slow control (or programming) mode and a functional mode. Each of these modes is associated with a separate set of constraints. The key constraints defining these two modes are the clock signals, which define which part of the design is active. In the slow control mode, only the JTAG clock (TCK) is active and consequently, only the registers driven by this clock can toggle. In the functional mode, only the 695 MHz and `XCLK` clocks are active. Throughout the flow, static timing analysis (STA) is used by the tools to verify correct functionality. STA is a complete and exhaustive method of verification of all timing checks of a design [54]. These checks include for instance setup and hold checks for flip-flops. The setup check ensures that the data arrives early enough at a flip flop and essentially defines the maximum operating frequency. A hold check ensures that the data at the input of a flip flop is held long enough for the flip flop to latch the data. Hold violations are very critical because they cannot be solved externally by decreasing the clock frequency. This would be possible in the case of a setup violation[18].

Static timing analysis is performed numerous times during the entire flow and it is therefore essential to minimize the time required to reach closure to constrain the design properly. For the ASIC at hand for instance, all slow control registers can be regarded as static from the functional domain because they do not toggle. Dividing the design into two modes therefore eliminates a lot of paths from being timed and optimized.

In our case, the synthesis step has been merged with the placement of the design. Traditionally, the placing and routing is done in the same software, but it is in general a design (and tool capability) choice. The RTL Compiler is capable to synthesize to a placed design, which allows physical aware optimizations, in the sense that different logic gates can be used to implement the same functionality. The place and route tool which has been used was Cadence Encounter. It has a lot of build in functionalities like for instance relocating cells, adding buffers or skewing the clock tree to optimize the design, but it cannot change the design logically. Although it was not essential to reach timing closure, the physical aware synthesis has proven to be beneficial. Without the physical aware synthesis for instance, the design needs to be over-constrained, i.e. a shorter clock period needs to be assumed, during the synthesis step to reach the timing specification after placing and routing. In this case, the clock frequency is relaxed when entering the place and route step in order to free time for wire propagation delays caused by routing. In the physical aware synthesis this is not needed because the sufficient estimates of the wire delays are available during synthesis.

The Encounter software has been used to perform the remaining steps of power rail insertion, clock

[18]if there is margin considering the system frequency, which in the case of this project cannot be tuned

tree synthesis, routing, filler cell addition and a final sign-off verification step. STA and optimizations are performed along the way after each step, while the propagation delays are estimated using wire load models before the full routing information becomes available. Only for the sign-off verification the complete required information is available. All of the routing wires are extracted and coupling capacitances between neighbouring wires can be taken into account to verify signal integrity.

5.6 Verification

This section explains the simulation methods used to verify the proper functionality of the full scale ASIC. Simulation is a very important tool and accompanies the designer from start to finish. Most important is the proper interaction of the individual blocks within the full system. Due to the size and complexity of ASIC, it is not feasible to simulated it as a whole. The design has to be abstracted and partitioned to simulate the interaction of the individual parts. The used methodology is outlined while the reader is again spared of too much technical detail.It is essential during the design phase to make sure that it can broken down into smaller pieces which can efficiently be verified.

Analog Simulations

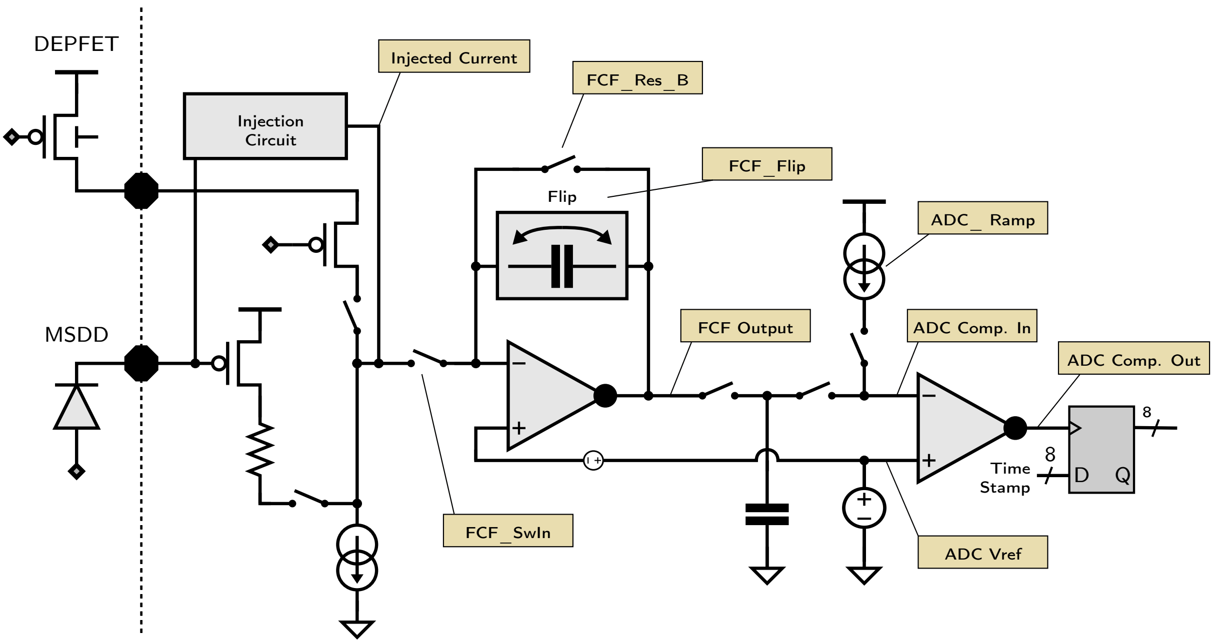


Figure 5.30: Schematic for the simulation depicted in figure 5.31 indicating the displayed signals.

The individual building blocks have been thoroughly simulated on the analog level by the respective design groups including Monte Carlo studies, noise simulations, verification of the trimming and gain setting ranges and-post layout simulations. After the integration of the pixel, the proper interaction of the pixel circuits in the various operation modes was verified with analog simulations to the design. The simulated operation modes include:

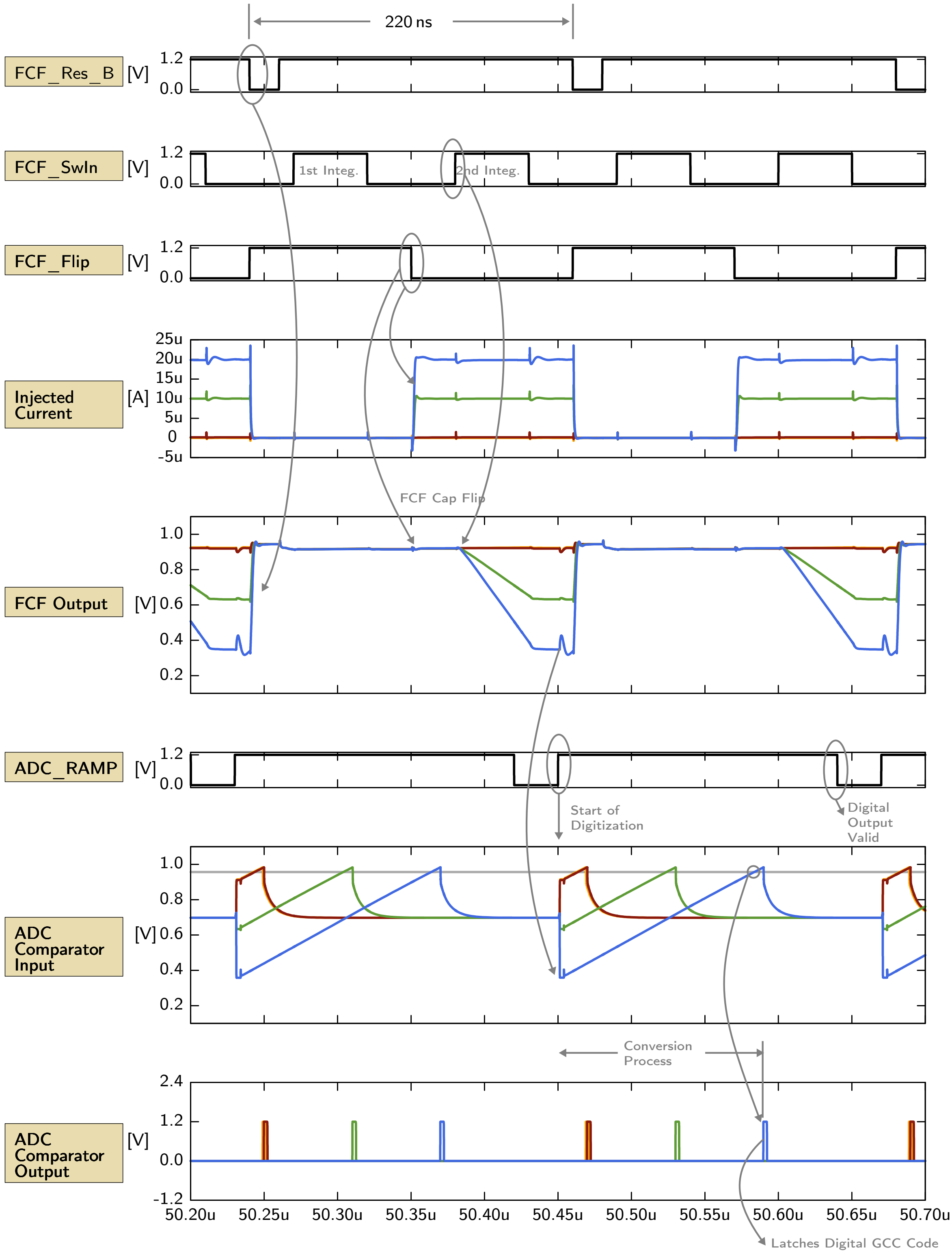


Figure 5.31: Exemplary analog transient simulation of the pixel. Three different current signals (red, green and blue) are injected, showing the analog processing chain in the pixel.

- charge readout mode (ideal charge pulse on the MSDD),
- current readout mode (ideal current pulse from the DEPFET),
- charge injection mode (internal pixel injection),
- current injection mode (internal pixel injection),
- pixel power down mode

With these simulations, the following properties have been verified:

- proper powering up of the circuit,
- proper settling of the current programming phase,
- interaction of the signal processing chain,
- proper power consumption for various supply voltages

The pixel power consumption values given in table 5.1 are based on these simulations. A sample simulation of the complete pixel which includes the pixel injection circuit, filter and ADC is shown in figure 5.31. The memory has been replaced by a black box and the slow control register bits provided by a Verilog-A abstraction to avoid the need for loading the register through simulation. Proper interaction of the ADC and the memory has been simulated. A post layout simulation of the memory and pixel slow control register has been conducted to verify proper operation with real wire loads across the matrix. The parasitics of control wires throughout the matrix have been extracted to properly size all signal buffering structures.

Digital and System Level Simulations

For verification on the system level, a sophisticated simulation environment has been set up, which is shown in figure 5.32. In this setup, basically all parts of the ASIC can be plugged in. The stimulus of the simulation is generated by the software which controls our lab test setup. It also involves the same FPGA firmware code which is used to control the ASIC in the lab. The software can be set to simulation mode, in which it dumps all data that would normally be sent to the physical FPGA via USB into a file. This file can be read in by the digital simulator, is fed into the input FIFO of the FPGA code. From this point on the data is processed by the FPGA firmware code which further controls the ASIC. This approach allows for configuring all registers in the FPGA and ASIC and controlling the simulation sequence through the software. Complex programming patterns, for instance for the sequencer block can be generated, allowing for sweeping for instance integration times at a higher level of abstraction in C++ code. The software includes a compiler which automates the generation of the programming patterns required for the sequencer. The setup has been very useful also when operating the chip, as the setup can be debugged efficiently through quick simulations.

The actual design under test can vary in this environment. The digital control block can be simulated by itself, in this case the output is checked either manually or with so called Verilog *checker* modules. Proper interaction of the sequencer with the front-end and ADC is basically guaranteed since it can be programmed very freely. Proper physical transmission of the signals to the pixel has been verified in an analog simulation.

The readout structure and pixel memories have been simulated thoroughly on a large scale matrix level to assure that all data shift register chains are controlled properly. The verification includes two steps: a purely logical verification ensuring that all chains and the memory are operated properly through the state machines. Therefore, purely digital models for the memories and readout cells have been developed to abstract the simulation. Test data has been generated in each single pixel and the

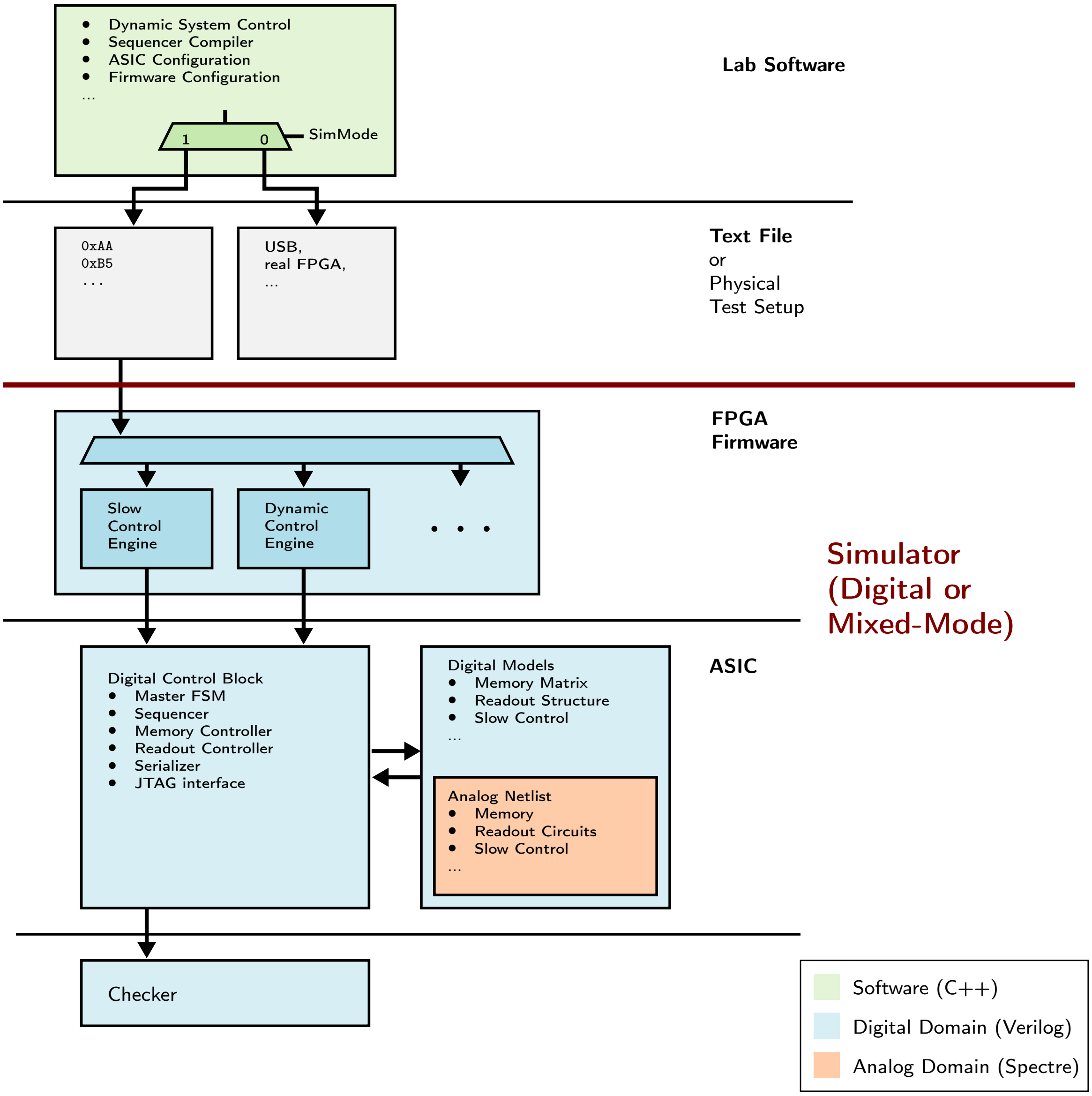


Figure 5.32: The system level simulation setup includes the lab software, the ASIC controlling FPGA on the test setup PCB and the ASIC. Different partitions of the ASIC can be plugged separately for time efficient simulation.

digital output stream checked for correctness using an automated procedure. Physically, the design has been verified by replacing only some of the cells with the analog netlist including extracted parasitics for a mixed-signal simulation. Simulating the whole matrix on the analog level is out of scope on the analog level and not necessary because the design is such that all cells repeat across the matrix. All cells have been designed modularly such that they can be connected in series, buffering all signals on the way. The functionality of two cells in series has been verified, ensuring that the full scale chain is functional on the physical level. On the physical level it suffices for instance to verify that the memory properly interacts with the local readout register chain and the serial register chain among the column is functional, (in figure 5.17: green to red cells and red cells in chain among each other). All of the interfaces have been verified and the buffering properly analyzed. The modular approach has been followed from early on in the design phase also in the test chips to ease scaling to the large matrix.

For the submission of the first test chip including the digital control block, the according firmware and lab software had not been developed. The simulation process was rather tedious because the simulation stimuli had to be provided manually resulting in some unsimulated corner cases. Although the design was completely functional, some workarounds were required for proper operation. The software related simulation environment could be derived from the lab software rather quickly. All that is needed in the software is a simple switch redirecting all data which would normally be set to the FPGA to a text file. This is conveniently implemented in the class which handles the communication with the FPGA firmware. As the lab software was available, all subsequent chips were simulated using this approach.

5.7 Timeline to F1 & Submitted Test Chips

Active development with first test chips has been started in 2009. For the first iteration of test chips, all major pixel elements - filter, ADC, RAM and injection - have been placed on dedicated test chips and characterized by the respective design groups. Successful operation of the test chips led to the integration of the individual blocks in the first 8×8 mini matrix (MM1) test chip submitted in August 2010. The size of 64 pixels has been chosen to mimic a full format 64 pixel column. Wires distributing power and crucial control signals are meandered through the columns to map voltage drops and signal degradation along the long columns of the final chip. A second MM chip with minor improvements has been submitted in August 2011. For MM1 and MM2, only the fast sequencing signals for the front-end have been generated on-chip, all other control signal (mainly for readout, RAM control and slow control) have been generated off-chip. A test chip integrating the digital control block and 4 pixels has been submitted in November 2011. Successful operation of the digital control block led to integration of a pixel matrix and on-chip digital control, essentially forming a mini matrix of the final chip topology in MM3. It has been submitted in May 2012 and is the first chip incorporating the bump bonding interconnect. The submission of a first full format 64×64 chip prototype (F1) on an engineering run has originally been scheduled to be submitted in the summer of 2013. However, the start of the parallel development of an MSDD version of the system requiring additional pixel circuitry and delays in finalizing the purchase agreement for the engineering run have delayed the submission until April 2014. Table 5.3 summarizes all the chips submitted by Heidelberg University during up to date.

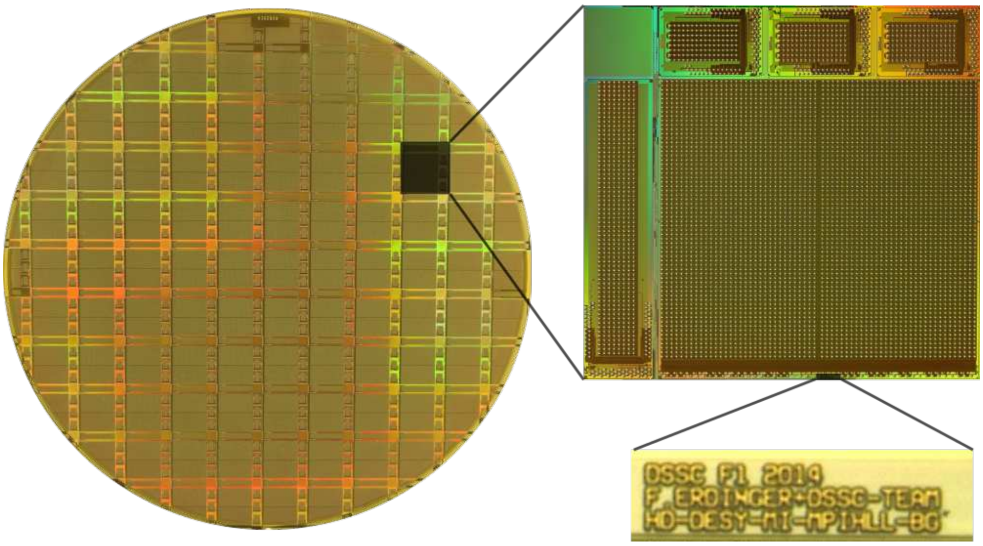
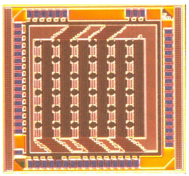
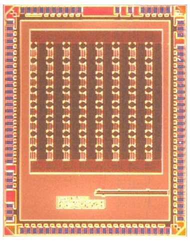
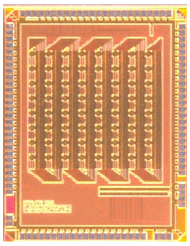
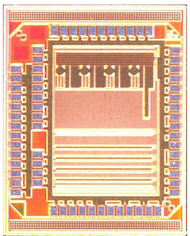
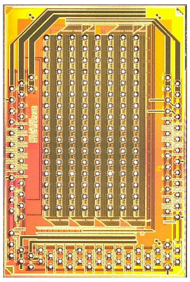
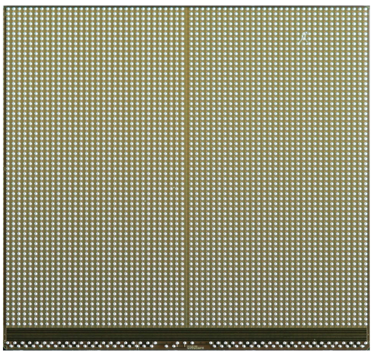
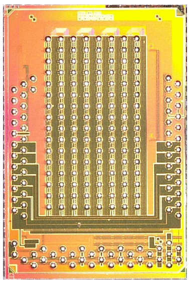
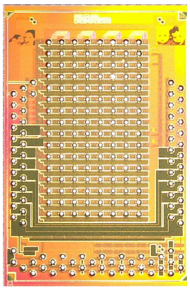
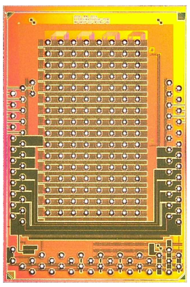
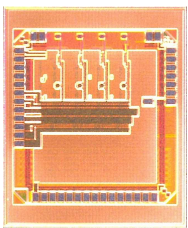


Figure 5.33: Wafer photograph and zoom into the reticle which contains the full scale F1 (4096) and further smaller test chips L1, MM4, MM5 and MM6.

Chip Name	Die Photograph	Subm. Year, Size, Description
DRAM2		2010, $2 \times 2\,\mathrm{mm}^2$, wire bonds Pixelated dynamic memory and read-out structures.
MM1		2010, $2.5 \times 3.2\,\mathrm{mm}^2$, wire bonds 8x8 pixel matrix with first integration of the signal processing chain of filter, ADC, DRAM, no on-chip control only global memory address decoders.

MM2		2011, $2.5 \times 3.2\,\mathrm{mm}^2$, wire bonds 8x8 pixel matrix with first integration the signal processing chain of filter, ADC, SRAM, pixel injection circuit, no on-chip control. Improvements were made on the filter and ADC and the power busses routed in snake fashion to mimic a 64 pixel column. No on-chip control only global memory address decoders.
CNTRL1		2011, $1.6 \times 1.8\,\mathrm{mm}^2$, wire bonds Test chip mainly for the digital control block. 4 pixels were added with minor changes and first test structures of LVDS pads included.
MM3		2012, $3.2 \times 4.8\,\mathrm{mm}^2$, bump bonds First complete mini matrix (8x16) test chip, including all periphery planned for the final 4k pixel chip. The MSDD front-end development has not been started yet at the time of submission. Two full 64 pixel columns in two 4×16 sub-matrices. One half has a global GCC for the ADC, the other an in-pixel counter and the clock is distributed.
F1		2014, $14.9 \times 14.0\,\mathrm{mm}^2$, bump bonds First sull scale engineering run submission. MSDD Front-End and DEPFET front-end are included, the peripheral voltage DAC and a temperature measurement circuit [19] was added.

[19] supplied by FEC DESY

L1		2014, $3.2 \times 14\,\mathrm{mm}^2$, bump bonds Submitted on the F1 engineering run. Test chip for the in-pixel counter ADC with realistic wire loads of a full and straight column. Otherwise the pixel is identical to F1, the periphery is the same as in F1.
MM4		2014, $3.2 \times 4.8\,\mathrm{mm}^2$, bump bonds Submitted on the F1 engineering run. Same pixel (DEPFET + MSDD FE) and periphery circuits as in F1 for comparison with a small matrix.
MM5		2014, $3.2 \times 4.8\,\mathrm{mm}^2$, bump bonds Submitted on the F1 engineering run. Identical to MM4 but without the DEPFET font-end and with the capacitive compression (section 6.1) instead of the triode compression.
MM6		2014, $3.2 \times 4.8\,\mathrm{mm}^2$, bump bonds Submitted on the F1 engineering run. Identical to MM6 but without the DEPFET front-end and with an improved triode compression (F1 includes the conservative baseline design.)
D0M1		2015, 1.7×1.8, wire bonds Test chip for improved MSDD front-ends. The NInput with the pixel wise reset regulation loop (section 6.2) was added as well as two modified pixels of the PInput design.

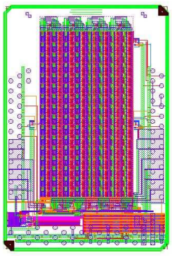
MM7		2016, $3.2 \times 4.8\,\mathrm{mm}^2$, bump bonds 8x16 test chip comprising three MSDD front-end variants: improved NInput, PInput with adaptive reset voltage mechanism and a (linear) charge sensitive amplifier configuration (*classical closed loop solution*, supplied by Politecnico di Milano). In Fabrication.

Table 5.3: Submitted chips during the course of course of this thesis. F1, L1, MM4-6 were all placed on the engineering for F1. They all fit together in the reticle, two full scale chips unfortunately do not fit.

6

Front-End Electronics Design

6.1 A Capacitive Signal Compression Technique

This section presents the study of a capacitive signal compression mechanism which has been carried out in parallel to the triode compression mechanism explained in section 5.3.2.2. The capacitive method described here aims at dampening the voltage swing at the charge collecting node for larger signals to generate a nonlinear system characteristic. The implementation of this mechanism is more involved but offers the advantage of lowering the voltage swing at the input node for large signals. Large voltage swings can couple into neighboring pixels which should still preserve the capability to correctly identify single photons. The described circuit has been implemented on a mini-matrix test chip (MM5) which was fabricated on the F1 engineering run.

6.1.1 The Concept

System Considerations

A conceptual schematic of the proposed circuit is shown in figure 6.1. The compression mechanism integrated in the ASIC should be compatible with the further signal processing chain and hence generate a signal current to be injected into the virtual ground of the filter. The collected charge from the sensor anode is converted to a voltage by the capacitance at the input node of the ASIC. The voltage is further converted to a current by a PMOS transistor. The reset capacitance $C_{res} = C_{dyn,min}+C_{static}$, X-ray energy and target signal to noise ratio dictate the required transconductance g_m from the input node to the filter input (as discussed in section 4.1.1). C_{static} contains the detector capacitance, the capacitance of the input transistor and further stray capacitance C_{stray} contributed by the solder ball and associated interconnect.

The voltage swing at the input generated by a single detected photon essentially defines the maximum allowable voltage swing at the input node, when the characteristic of *voltage* at the input to digital output is *linear*. The estimated sum of all capacitances at the input node is $\approx$ 400 fF (including the initial state of the dynamic capacitance), yielding a voltage swing of 110 µV for a single 1 keV photon. Since this is attributed to the first ADC bin, the maximum allowed swing at the input which fits into the dynamic range of the 8 bit ADC is only 28 mV.

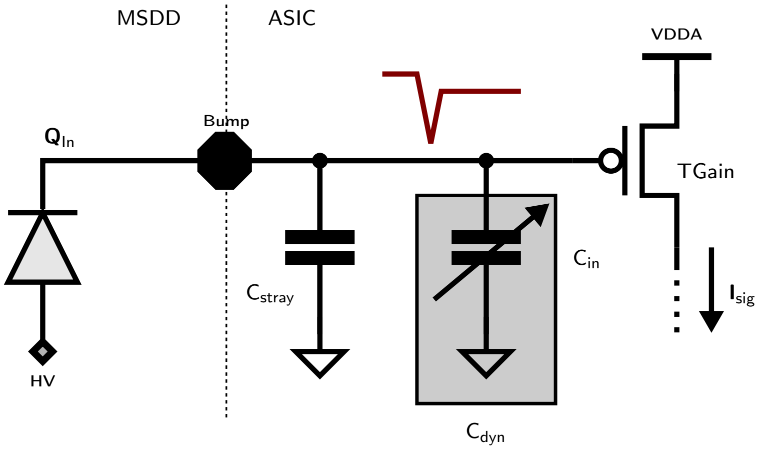


Figure 6.1: Conceptual schematic of Q → I conversion and signal compression. A signal dependent capacitance (C_{dyn}) at the input node of the ASIC node dampens the voltage swing for larger signals.

Considering a target dynamic range of 3270 for 1 keV photons[1] and a static capacitance of 400 fF at the input, the maximum swing would be 363 mV. The capacitive compression concept is to dynamically increase the input capacitance such that this maximum voltage swing reduces to 28 mV to comply with the dynamic range of the ADC.

The system should again be compatible with the DEPFET readout chain, for the first version, the same simple approach was followed as in the F1 readout to use a single transistor[2] and a global reset voltage. This was identified as a sub-optimal solution and is not discussed further here. The problems and a possible solution are explained in section 6.2. This section focuses on a capacitive compression mechanism.

MOS Capacitor Principle

A variable capacitance can be implemented using a MOS (metal-oxide-semiconductor) capacitor. The device has three *macro* states: accumulation, depletion and inversion. The capacitances between the terminals differ significantly for these three states. The mechanism described here uses the depletion and inversion states to develop a capacitance as a function of the applied voltage. The device can be implemented using an NMOS transistor, shorting its source and drain terminals, such a device will be referred to as an NCAP in this section. A cross section is shown in figure 6.2. It thus has three terminals, the *bulk*, *gate* and *drain*. An NMOS transistor is chosen because it has the right *polarity* for the application, it is used in the circuit to *drain* electrons from the input node. The terminal is named accordingly. The bulk terminal is connected to ground in all figures (not explicitly indicated). A very detailed explanation of the capacitances can for instance be found in [27], this section only explains the general behaviour.

The drain capacitance of an NCAP with a fixed gate voltage (v_g) versus the applied drain voltage (v_d) is shown in figure 6.2. If the gate-drain voltage (v_{gd}) is substantially lower than the threshold

[1] dynamic range of the DEPFET in 8 bit mode is 3270

[2] but without the compression resistor

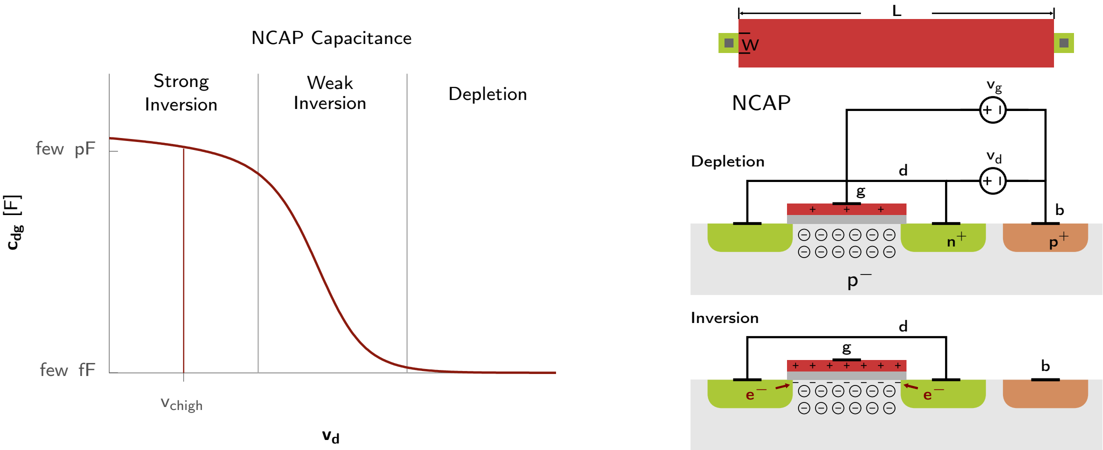


Figure 6.2: Capacitance behavior of an NCAP device with $W \ll L$ for a fixed gate voltage (v_{gate}). The axis are unlabeled due to a nondisclosure agreement. If $v_{gd} \ll v_{threshold}$, the area under the gate is depleted. As the drain voltage decreases, electrons are drawn underneath the gate from the drain regions and *invert* the region under the gate.

voltage (v_{th}), the area under the gate is depleted of free charge carriers. There is no transistor channel, the drain region is ohmically disconnected from the region under the gate. In this state, the capacitance at the drain node is given by the gate-drain overlap capacitances and the junction capacitance. If (the transistor) width (W) of the NCAP is very small, the capacitance amounts to few fF only.

When lowering the drain voltage, the device enters the inversion state, where electrons from the drain are drawn underneath the gate. This happens gradually, from a weak inversion state to strong inversion as v_d is lowered further. In a *regularly* connected transistor this corresponds to the formation of a conductive channel. The final capacitance is given by the area $W \times L$ of the NCAP. In the application at hand, the effect is used as a variable capacitance.

Boosting the MOS Capacitance

If the NCAP is connected to a sensor anode with the drain, and biased such that for the reset level of the input node, the device is in depletion, the capacitance contribution is negligible. Electrons collected on the input node decrease the voltage and thus increase the voltage across the capacitor. However, a large change of voltage is required for the capacitance to increase significantly, when the gate is connected to a fixed voltage. Even the maximum allowed change of voltage at the input node of 28 mV which corresponds to the dynamic range of the ADC barely affects the capacitance.

However, the voltage across the capacitor can be boosted in order for the capacitance to increase already for small to medium signals. Figure 6.3 illustrates this concept: by using an inverting amplifier to sense the drain and control the gate, the capacitance curve is *compressed* in the x-direction ($v_d = v_{in}$). The change of voltage at the drain terminal appears amplified over the capacitor and thus the transition from depletion to inversion happens for smaller changes of the drain voltage. Additionally, the Miller effect (see for instance [28]) boosts the capacitance such that the curve stretches in the y-direction (c_{gd}). By the usage of the boosting amplifier, the red curve, which is equivalent to the curve shown in figure 6.2 can be transformed into the green curve.

This is the basic principle of the circuit, the implementation for a first test chip is described in the

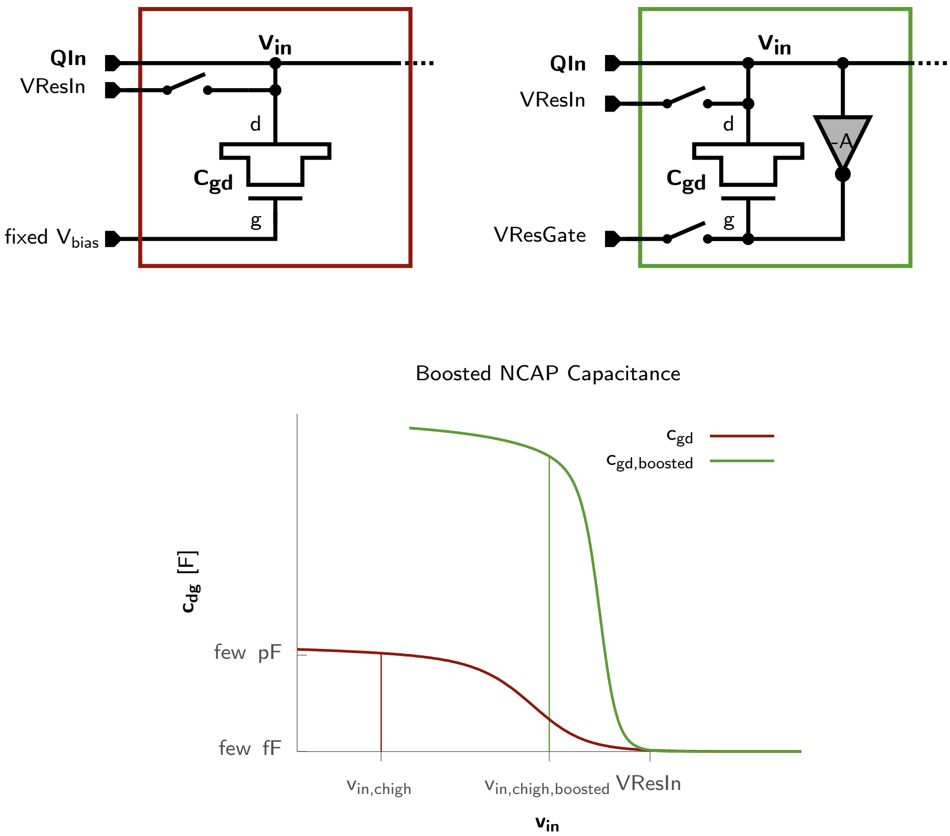


Figure 6.3: Principle of the capacitive compression. Units are omitted in the plot because of a nondisclosure agreement. An NCAP (NMOS transistor) with $W \ll L$ is used to implement a variable capacitance. The plot shows the behavior of the capacitance against the voltage at the input node ($v_{in} = v_d$). The voltage-capacitance behavior of the NCAP is boosted by an amplifier sensing the input node and controlling the gate. The input signal is a charge from the sensor which is converted to a voltage signal (v_{in}). The voltage signal is compressed according to the rise of c_{gd} as the voltage at the input node decreases. The depicted switches are needed to reset the circuit.

next section.

6.1.2 Circuit Implementation

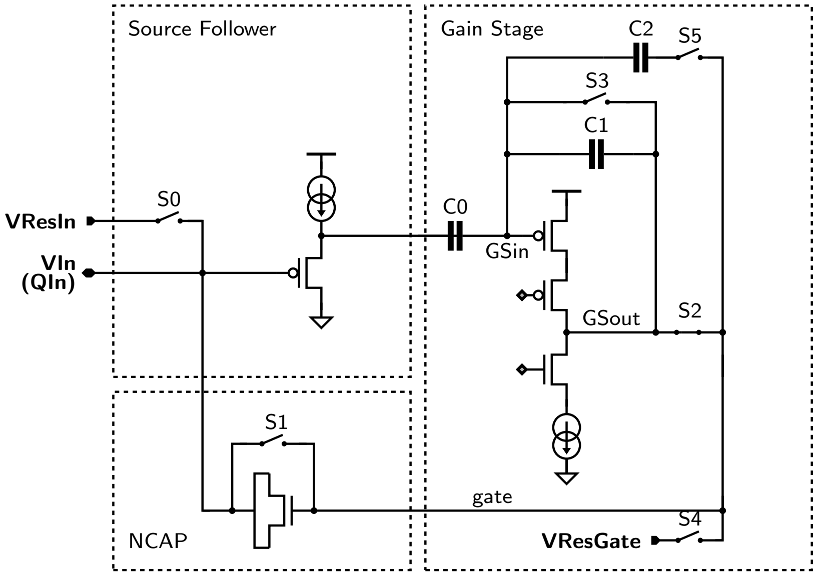


Figure 6.4: Detailed schematic of the non-linear capacitance implementation. To avoid loading the input node in the low capacitive reset state, a source follower senses the input node and drives a preceding amplifier. The amplifier is implemented as a single ended cascoded gain stage, the gain is set by the ratio of C1/C2. The input signal is a charge (QIn) and the *output* is the voltage (VOut) which appears at the same node due to the nonlinear capacitance that the circuit implements.

The amplifier driving the gate of the NCAP has to fulfill the following requirements:

(1) The circuit must load the input node as little as possible to guarantee a very small initial capacitance.

(2) Sufficient speed is required to settle the output during the flattop of the trapezoidal filter function. This property is challenging because the circuit contains a varying feedback capacitor (NCAP) between the input and the output of the amplifier. In the worst case when the incident signal is very large some pF must be driven.

(3) Power consumption must be as low as low as possible since the power budget is already almost exploited.

(4) The reset must allow for imposing a reverse bias on the NCAP.

(5) The amplifier gain must be well defined and tunable.

(6) The dynamic range at the output should reach as positive as possible to provide the best charge handling capacity of the circuit. Once the amplifier saturates, the Miller effect boosting the capacitance vanishes, decreasing the equivalent capacitance at the input node.

(7) Noise is of low importance because in the reset state, the amplifier is coupled to the input only with a very small capacitance, its noise thus has negligible effect.

The full circuit implemented on a first test chip (MM5 in table 5.3) intended for a first proof of concept is shown in figure 6.4.

A source follower is used as the first stage of the amplifier to serve as voltage buffer and avoid capacitive loading of the input. For the gain stage, several different topologies were evaluated, starting with classical two stages differential operational amplifiers. A differential amplifier has the advantage that the rest voltage of the gate can be freely set through the reference voltage of the amplifier. However simulation has shown that it is difficult to obtain sufficient phase margin along the outer loop for all capacitances that can appear through the NCAP between the input and the NCAP gate.

A cascoded single ended gain stage, depicted in figure 6.4 has been evaluated to best fit the requirements listed above. A cascode is used to provide a sufficiently large open loop gain. A straight cascode has been favored against a folded because the folded variant provides less driving strength at the output. The straight cascode, however, cuts into the dynamic output range because two overdrive voltages are now required from the positive supply. The major drawback of the single ended topology is that it requires a complex reset procedure. The idea of the implemented reset mechanism is to first short circuit the gain stage to define the DC voltage at its input and reset the input and the gate of the NCAP separately (switch S2 is open for these steps). C2 is a large capacitor and buffers the voltage difference between the gain stage input and the imposed reset voltage for the gate (S4). For the last step, only S5 remains closed and S2 is reconnected. There is now a large feedback capacitor between the input and the output of the gain stage which disturbs the amplifier equilibrium. C2 imposes a change on the amplifier input such that the output settles near VResIn. The switch sequence is shown in a transient simulation in figure 6.5.

Capacitors are used to set the gain of the amplifying stage and the input source follower. The source follower is thus AC coupled to the amplifier which allows to chose the reset voltage for the gate freely. The gain is set by the ratio of C0/C1 and can be adjusted by implementing a set of statically selectable capacitors for C1.

Figure 6.5 shows transient simulations of the full circuit for various input signals. The circuit settles in < 50 ns which complies with the target speed. The reset is not perfect at the maximum 4.5 MHz speed and needs revision. The capacitive compression works however nicely, the input amplitude for 10000 1 keV photons is only ~ 30 mV.

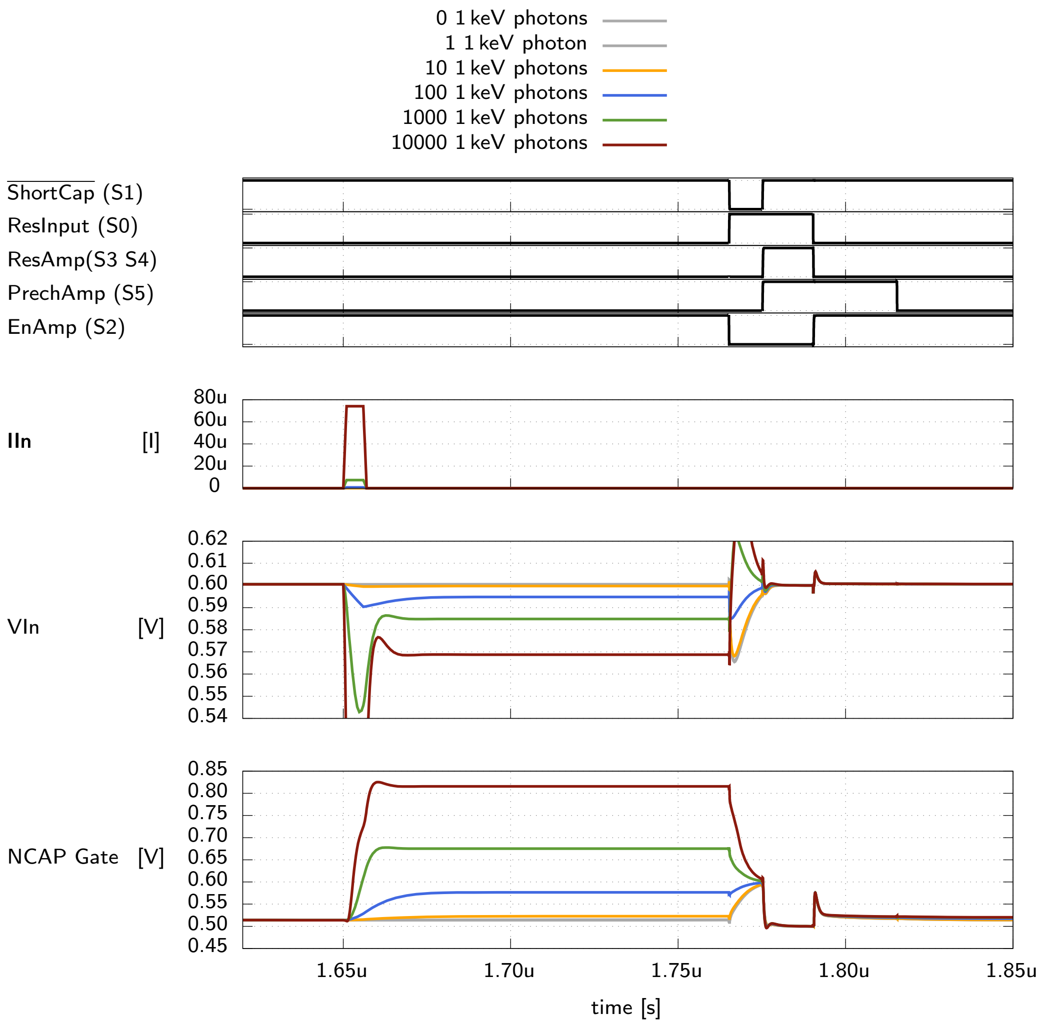


Figure 6.5: Transient simulation of the capacitive compression mechanism for various signal amplitudes. The digital switches are derived in the pixel from two global signals and control the reset phase. In the simulation, the input signal is a current (IIn waveform), the current amplitude is chosen such that the integral equals the charge corresponding to the incident number of photons for each wave. VIn is the charge collecting node, the amplifier steers the NCAP Gate and dampens the voltage change on VIn resulting in a compression of the input charge signal.

6.1.3 Simulated Compression Characteristics

Figure 6.6 shows the simulated compression behaviour. The voltage swing (v_{in}) for signals up to 10000 photons is plotted. A PMOS transistor senses this voltage swing and feeds a current into the Flip Capacitor Filter (Figure 6.1). The characteristics of the compression curve can be tuned by the following parameters:

1. The gain of the amplifier (parameter A for the curves) controls the slope of the capacitive change and the final capacitance for large signals.[3]
2. The shape (W/L) of the NCAP (parameter $NCAP_{low}/NCAP_{high}$ in the curves) controls the initial capacitance (gate-source overlap) and the final capacitance through its total area.
3. The initial bias on the NCAP (V_{bias}) controls the *onset* of the compression i.e. the length of the *linear* range.

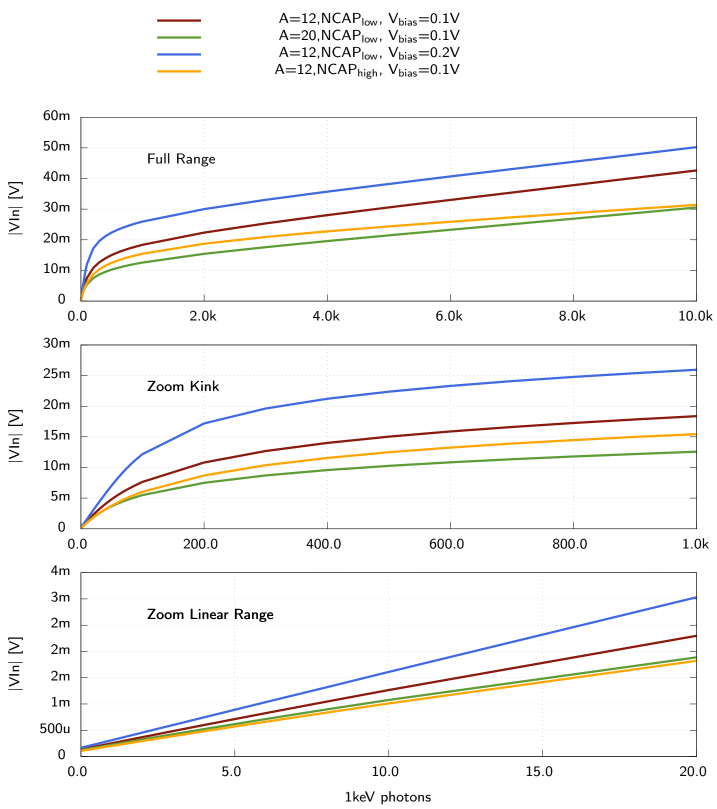


Figure 6.6: Dynamic range of the capacitive compression. The plots show the voltage swing (inverted for clarity) at the input node caused by signals of up to 10000 photons.

[3] as long as the amplifier does not saturate due to a large A, which is happens for the green curve in the full range plot

6.1.4 Test Chip Results & Conclusion

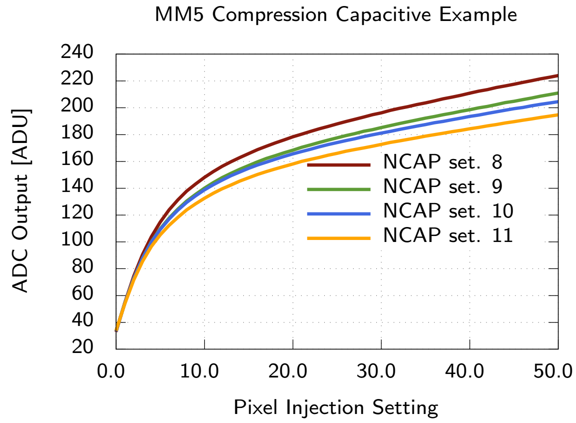


Figure 6.7: A sweep of the internal pixel injection in high gain mode is plotted for different NCAP settings. The injected charge is a function of the input capacitance, which is variable (equation 5.33). No calibration for the x-axis is available yet.

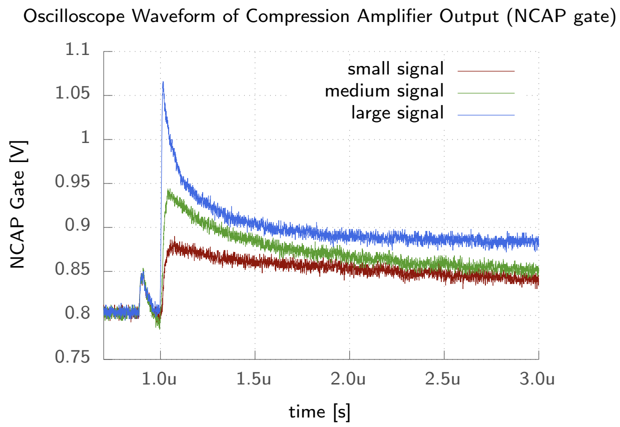


Figure 6.8: The three curves show an oscilloscope measurement of the NCAP gate voltage for three different signals injected on the input node. The circuit needs several microseconds to settle which is too slow for the target opertion speed.

A first variant of the circuit has been submitted for a proof of principle. The circuit comprises three gain settings for the amplifier and four scaled NCAPs with varying W and L = 100 µm (15 total settings).

The compressive behaviour has been measured and verified at slow speed (figure 6.7), the transient response of the circuit is however slower than simulated and varies with the signal amplitude. This is attributed to the large L of 100 µm used in the NCAP devices, which is not properly modeled in simulation. The collection of the charge in the long channel is in reality too slow due to its large channel resistance. Figure 6.8 shows a measurement of the NCAP gate voltage on a monitor pad of the chip. The gate needs several microseconds to settle.

In principle, other geometries are possible which could optimize the transient behavior. An angular structure with a single diffusion contact in the middle would be the best solution. For such a structure, the gate overlap capacitance is small and the area large without requiring a large L. Another test run would be necessary for these studies. The activity on the circuit was however put aside due to more important issues (see section 6.2).

6.2 An Improved Front-End Topology (N-Input)

The main problems of the first versions of the MSDD readout circuits and the reasons therefore are:

- **Pick up of power supply noise:** The source of the preamplifying transistor is directly connected to the supply line, making it as responsive to the changes on the supply as to an input signal. The supply is further shared with all other analog blocks. Despite the fact that all activity is synchronous, the performance suffers enormously from this topology.
- **Signal cross talk through the power supply:** Any semi-static change of current in the supply due to the processing of a signal causes a slight change of voltage on the supply line. These voltage changes are amplified by the input transistor causing cross talk between the pixels.
- **A drastic reduction of gain along the pixel columns:** The F1 design makes use of a global reset voltage which is distributed to all pixels. Static voltage drops along the pixel columns degrade the biasing gate-source voltage of the input transistor decreasing its transconductance.
- **A large gain spread:** Even pixels with the same voltage drop do not deliver the same gain because of fabrication mismatches of the transistors. The fabrication mismatches mainly result in differing threshold voltages which cause a spread in the bias current and hence the transconductance of the transistor.

These issues have triggered an extensive re-design phase of the MSDD readout circuit with the main focus on dealing with these issues. While the initial design phase was driven by pure single channel performance figures and keeping seamless compatibility to the DEPFET readout mode careful studies of robustness and large matrix effects have mostly been undervalued due to stringent timing schedules. There was also no time left for a mini-matrix prototype, requiring a big jump from a single channel test design to the full scale 64×64 pixel matrix. The goal of the re-design is to improve the robustness of the circuit with respect to the mentioned issues, even if it means giving up some performance in terms of noise figures and dynamic range. A dynamic signal compression mechanism has been degraded to lowest priority. A simpler high dynamic range mode can be provided by implementing a mode in which several photons are attributed to each ADC bin, hence sacrificing single photon resolution for dynamic range. As in most analog circuits, trade-offs have to be made in terms of using power for circuit properties such as for instance power supply rejection (PSR). An improved input stage using an NMOS as the main input amplifier has been designed, which is complimentary in terms of the circuit topology to the MSDD readout circuit in F1. The circuit has been evaluated very carefully in theory and simulation and a small test chip has been fabricated. The NMOS input circuit has become a serious candidate for the final implementation on F2. The P-input topology as used in the F1 chip has also been expanded by the features proposed here for the N-input which include the generation of a pixel-wise reset voltage and source stabilization of the amplifying transistor.

While these two variants both comprise an open loop input stage, a third version is under study by the collaboration, in which a charge sensitive amplifier is used as the input stage. While charge sensitive amplifiers are well known and widely used, the caveat for the situation at hand is the current mode interface of the subsequent analog filter stage, which is imposed by the original (DEPFET) topology of the system. A low noise voltage to current conversion is required which matches the dynamic ranges of the CSA and FCF. A good solution is pending.

In the following subsections, a complete detailed elaboration of the proposed N-input front-end circuit is presented. The choice of available properties with respect to investment is discussed. A first test chip has been fabricated successfully, all the implemented features are functional. Some results are presented in section 7.2.3 and section 7.2.4.

6.2.1 Circuit Overview

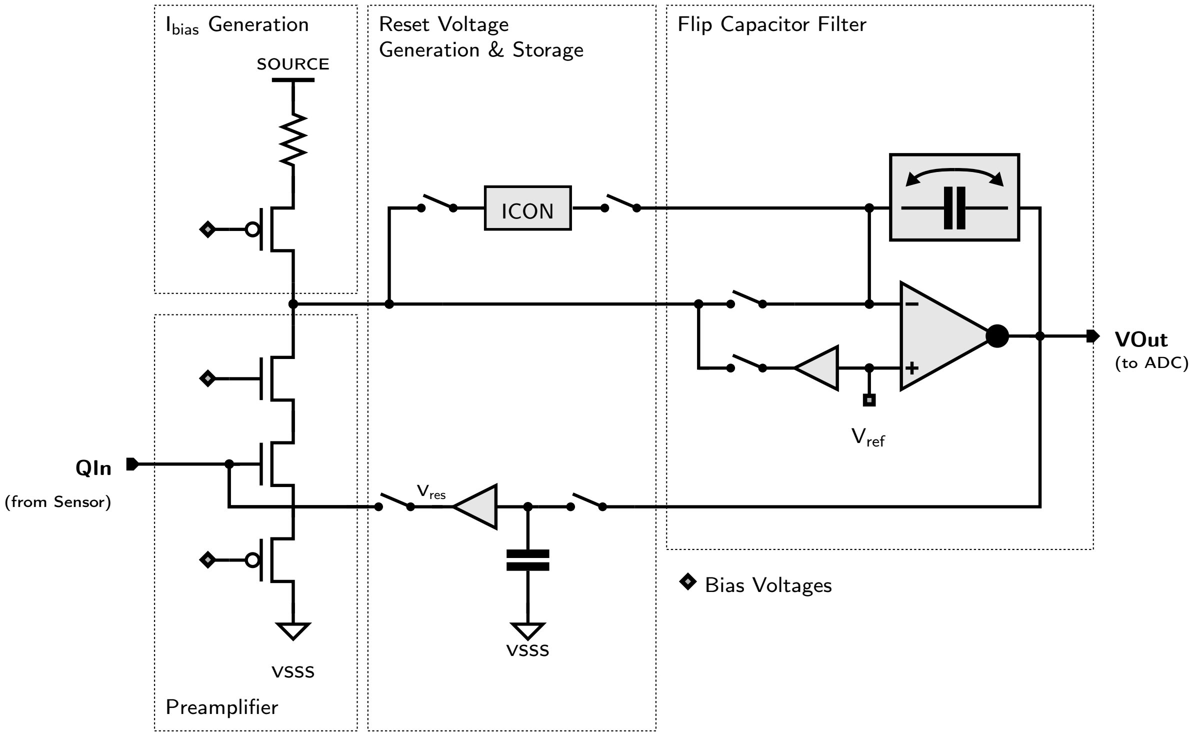


Figure 6.9: Simplified overview schematic of the complete front end based on an NMOS input transistor. The circuit comprises an automatic bias generation mechanism, and re-uses the flip capacitor filter.

A simplified overview schematic of the proposed circuit is shown in Figure 6.9. The main principles and building blocks of the circuit are:

(1) **Separate supply lines:** The input branch is supplied by the SOURCE voltage, which was originally foreseen for the DEPFET design.[4] A separate ground net (VSSS) is used which is connected to the main ground outside of the chip. These two nets are only supplying the input branches of the pixels. SOURCE hence supplies a constant current which does not change due to signal acquisition.

(2) **A preamplifying NMOS transistor** (TGain) which converts the collected charge of the sensor to a signal current. The use of an NMOS transistor avoids an increase of current in the supply line due to an incident signal. It basically redirects bias current as signal current into the filter stage. Note that this also means that the bias current decreases and the transistor consequently loses transconductance with increasing signal magnitudes. The source of the NMOS is decoupled by a transistor in source follower configuration (TStab) to reduce the sensitivity on the supply line (ground). The topology is explained in section 6.2.2.

(3) **Generation of a local reset voltage:** The reset voltage is generated for each pixel separately

[4] Due to the lack of physical connections this was not possible in F1.

such that the preamplifying transistor carries a fixed imposed bias current. This mechanism reduces the spread of input transconductance across the pixel matrix and is explained in detail in section 6.2.3.1.

(4) A **voltage memory cell** which stores the generated reset voltage for the duration of the burst. Leakage is negligible in all process corners by appropriately sizing the capacitor ($> 2\,\mathrm{pF}$) and using a single high-threshold PMOS for writing.

(5) A **current source** which is **globally and locally adjustable** to equalize chip-to-chip and on-chip process variations of the current generator to further equalize the gain across the matrix. Implementation details are given in section 6.2.3.2.

(6) The **flip capacitor filter** implementing a trapezoidal weighting function as it is used in the original DEPFET design. This circuit has been presented in section 5.3.2.1.

An NMOS is used in this topology as the input transistor because this way the current source is above the virtual ground input of the filter which allows to use a higher voltage for supplying the current. The DEPFET SOURCE voltage can be used here which is free in the MSDD variant of the system.

The general sequence of operation remains unchanged with respect to the F1 MSDD readout circuit, (see figure 5.5): Preceding the burst, there is a phase ($\approx 10\,\mu\mathrm{s}$) to generate the reset voltage (similar to the IPROG phase). The event cycle also remains unchanged and consists of the following phases: resetting the input node by pulsing SwRes, baseline integration to cancel any remaining bias current, flattop phase (trapezoidal shaping) where the signal is collected and finally the second integration phase which generates an output voltage for the subsequent ADC stage.

6.2.2 Supply Noise Suppression & Input Referred Noise

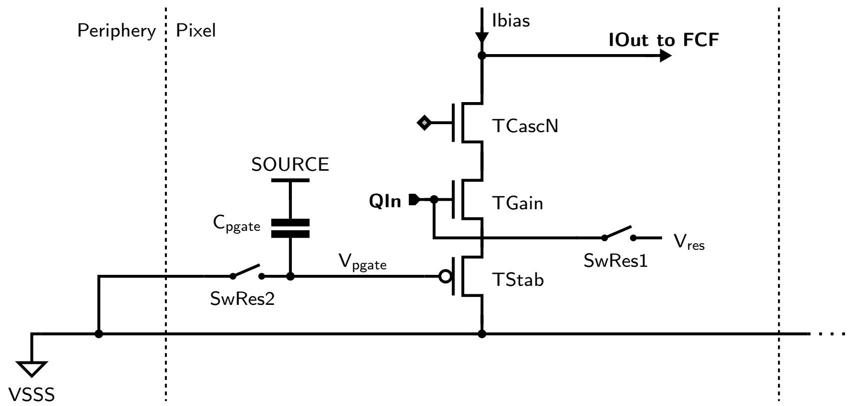


Figure 6.10: To improve the supply rejection, a PMOS transistor (TStab) is added in the source of TGain, which increases the resistance of the circuit seen from the ground line. The gate of the PMOS is very sensitive, the (clean) ground voltage from the chip periphery is sampled in each pixel to avoid cross talk from this line.

The approach to reduce the sensitivity on the power supply is to add a second device in source follower configuration to decouple the source of the amplifying transistor from the supply line. A schematic is shown in figure 6.10. In this configuration, the conductance of the circuit seen from ground is reduced, from the full g_m, N if the source were directly connected to VSSS, to:

$$G_s \approx g_{ds,P} \frac{g_{m,N}}{g_{m,P} + g_{m,N}} \tag{6.21}$$

where $g_{m,N}$ and $g_{m,P}$ denote the transconductances of TGain and TStab, respectively. G_s is introduced here as the *source conductance*. The derivation can be found in appendix A.2.2. The total transconductance G_m from the input node (QIn) to to the filter input is given by

$$G_m \approx \frac{g_{m,N} g_{m,P}}{g_{m,N} + g_{m,P}} = g_{m,N} \parallel g_{m,P} \tag{6.22}$$

the derivation is given in appendix A.2.1. The finite transconductance of TStab degenerates the source of TGain resulting in a loss of overall transconductance. The bias current hence needs to be increased substantially with respect to the single transistor solution to obtain the same overall transconductance. While TStab reduces the sensitivity on the ground line, its gate is as responsive as the real input node (QIn).[5] The quality of the applied bias voltage is hence of utmost importance. The required voltage level for the gate of TStab is ground (VSSS), a clean copy of VSSS is distributed to the pixels and sampled before each signal processing cycle (in figure 6.10 SwRes2 and SwRes1 operate in parallel).

Relating the source conductance to the input transconductance yields:

$$S = \frac{G_m}{G_d} \approx \frac{g_{m,P}}{g_{ds,P}} \tag{6.23}$$

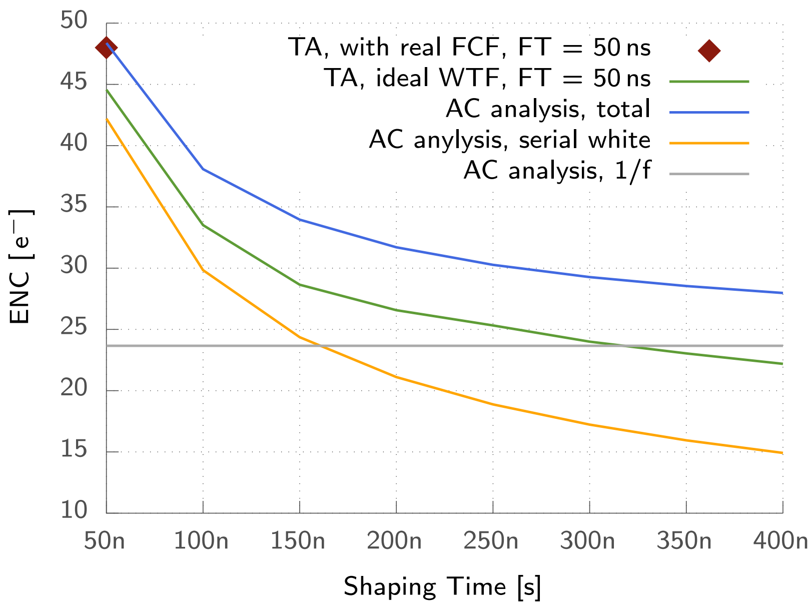


Figure 6.11: Simulated noise figures for the N-Input topology. TA = Transient Analysis.

[5] similar to a *real* differential input

The input referred noise of the circuit is given by

$$\overline{v_{ni}^2} = \sqrt{\frac{\overline{i_{n,N}^2}}{g_{m,N}^2} + \frac{\overline{i_{n,P}^2}}{g_{m,P}^2}} \tag{6.24}$$

For a derivation see appendix A.2.3. Both $g_{m,P}$ and $g_{m,N}$ hence need to be maximized while carrying the same bias current. To exploit the available power best, both transistors are sized such that they are biased in the weak inversion region when conducting the imposed bias current. In weak inversion, the transconductance is not dependent on the channel charge carrier mobility, TStab (PMOS) hence has a similar transconductance as TGain (NMOS) at equal biasing currents.

The noise contribution of the front-end including biasing has been studied for various shaping times using several methods:

- AC simulations to obtain the power spectral densities required for theoretical calculation using equation 3.47
- transient simulations and an ideal trapezoidal weighting function through numerical integration
- a full channel transient simulation for the maximum operating speed

A plot of the results is shown in figure 6.11 For 1 keV photons and 4.5 MHz, the simulated noise using the full analog channel is $49\,e^-$, which yields an SNR ratio of 5.8, meeting the requirements at the XFEL. A better SNR can be obtained for slower operating speeds when the shaping time is increased. The circuit parameters are:

- $G\_m = 1.3\,\mathrm{mS}$
- gain setting: 1 keV/ADC bin
- $C_{in} = 440\,\mathrm{fF}$ (140 fF contributed by the TGain)

With the first prototype of the circuit, a noise level of ENC = $56\,e^-$ was measured (the spectrum is shown in figure 7.6). A supply sensitivity measurement is presented in section 7.2.4.

6.2.3 Biasing & Gain Dispersion Improvement

6.2.3.1 General Principle

To solve the issues of gain spread and voltage drop sensitivity, a new biasing mechanism has been developed. The exploited principle is based on the fact that the transconductance of a MOS transistor is predominantly dependent on its bias current. Therefore a very homogeneous gain distribution can be achieved if the bias current is equal in every pixel. Taking into account a spread of threshold voltages, the reset voltage needs to be different in each pixel because it defines the bias current. The idea is hence to force the bias current and to generate the required biasing reset voltage separately for each pixel.

The implementation of this mechanism is similar as the analog fine tuning of the DEPFET bias cancellation described in section 5.3.2.3 and uses some of the same circuitry. Preceding the burst phase, the circuit is configured in a closed loop configuration, which generates the required reset voltage. The closed loop is depicted in detail in figure 6.12. In this configuration, the main filter amplifier is driving the gate of TGain and acts as an error amplifier, integrating the residual current i_{res} of the bias current i_{bias} and the current in TGain until, it is sufficiently small such that the circuit settles in a stable state. To match the polarities, an inverting stage is required which is implemented

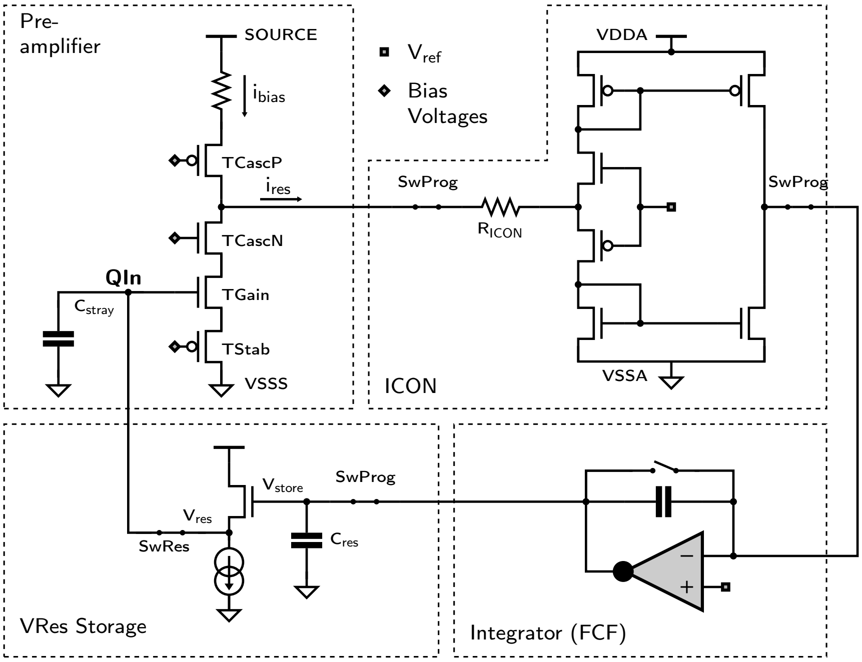


Figure 6.12: Detailed view of the current programming loop. All SProg and the SRes switches are closed to establish the loop while SRes is pulsed to clear a collected signal from the input node.

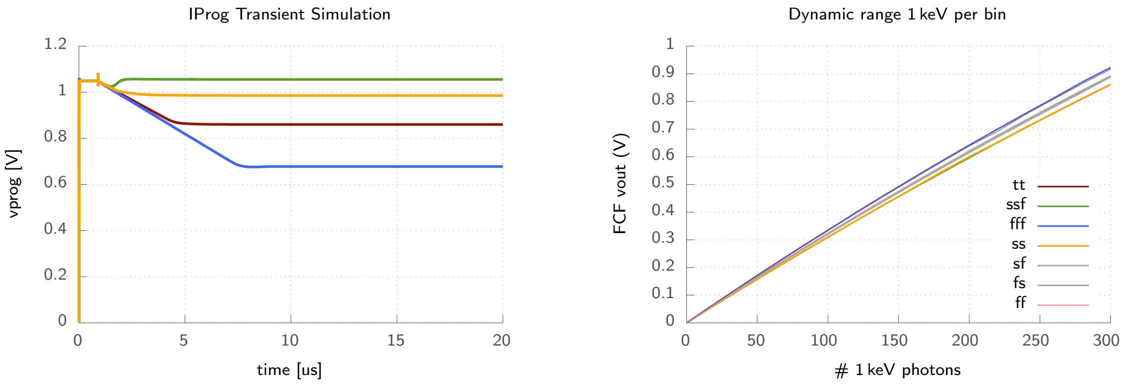


Figure 6.13: Current programming phase of the N-input topology for different process corners. Since the threshold voltages differ significantly across the corners, different reset voltages are programmed. For all simulations, the same bias current is programmed.

by means of a bidirectional current converter cell (ICON, for details see section A.3.1) preceding the filter (current integrator). The reset voltage is written on a large capacitor, which stores it for the duration of the burst. A buffer for this voltage, implemented by a simple source follower, is required to provide a low impedance node which can reset the input node without losing the voltage on the capacitor. When the programming loop has settled, SwRes and SwVprog are opened, the ICON cell is switched into tri-state making the circuit ready for signal acquisition. The implementation of the current source is described in the next section.

A detailed study was carried out to make sure that the programming loop is stable for all process corners. Stability is more of a concern in this loop with respect to the DEPFET variant, because it features the input transistor, the loop gain is thus a lot higher, making it more difficult to keep sufficient phase margin. Details about the loop are given in appendix A.3.2.

The gain dispersion of the circuit was analyzed using Monte Carlo and process corner simulations. Figure 6.13 shows that the circuit finds a correct reset voltage in all process corners for a fixed imposed bias current. The generated reset voltage is highest in the slow-slow corner because both the P and N thresholds are highest here and it is lowest in the fast-fast corner because here both threshold voltages are lowest. For the right plot, several transient simulation were carried out for each corner. After current programming, various signals were deposited on the input, the curve shows the resulting swing at the filter output. The curves deviate only slightly, showing that the gain is kept constant nicely across the process corners. The remaining deviation can be handled easily by trimming the ADC accordingly.

6.2.3.2 The Current Source

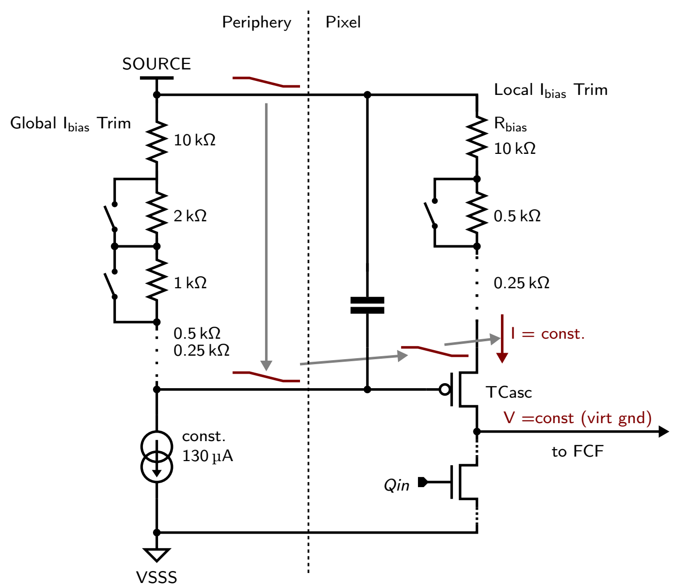


Figure 6.14: I_{bias} generation and trimming mechanism. The used (originally DEPFET) SOURCE voltage droops. The cascode transistor (TCasc) for R_{bias} is biased such that its gate voltage drops in parallel to the SOURCE, keeping bias current constant.

A suitable low noise current can most conveniently be generated by a large resistor. Using the coefficients for the trapezoidal filtering, equation 3.47 can be used to calculate the series white noise contribution by a 10 kΩ resistor in ENC rms at the input of the circuit:

$$ENC = \frac{C_{in}}{G_m e}\sqrt{\frac{4kT}{R_T}} \approx 13e^- \tag{6.25}$$

where k is Boltzmann's constant and T is the absolute temperature, e is the elementary charge and $\tau = 50\,\mathrm{ns}$ is the shaping time corresponding to the operating speed of 4.5 MHz.[6]

Since there is only $\approx 250\,\mathrm{mV}$ of voltage headroom from the input virtual ground node of the filter to VDDA, a higher supply voltage is needed to accommodate the large resistor. The free DEPFET SOURCE voltage which is not required to bias the MSDD can be used here. This power supply has however been designed specifically for the DEPFET. It is unregulated and directly supplied from capacitors making it droop along the burst (see section 4.2). It can however be freely tuned up to 7 V. Using this supply line is further beneficial because it does not affect the power budget of the ASIC which is practically exploited already without the additional MSDD front-end.[7]

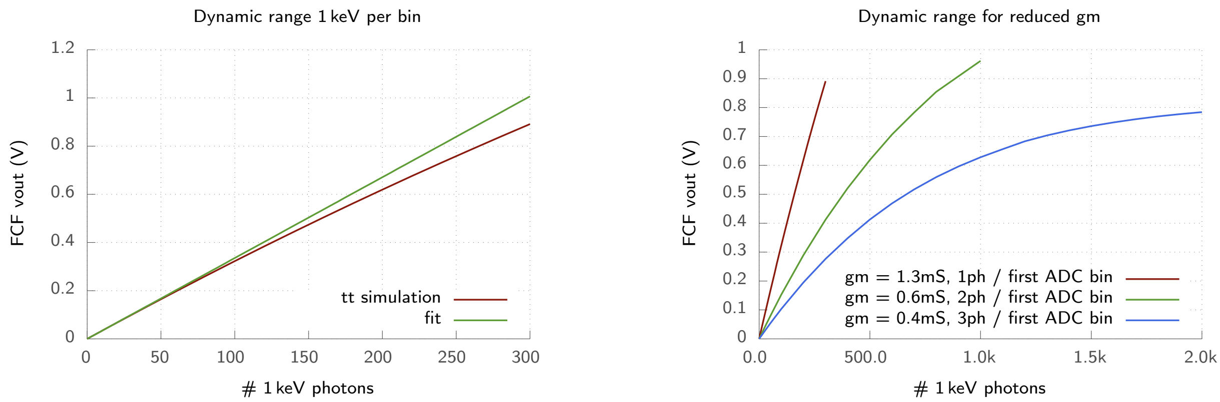
Since SOURCE drops along the burst while the virtual ground node stays constant, connecting the resistor directly to virtual ground node would also decrease the bias current along the burst. This effect can be made negligibly small by adding a properly biased cascode transistor. If the gate of the cascode follows SOURCE, the voltage across the resistor stays (almost) constant. The implementation is depicted in figure 6.14. A constant current is drawn through a resistor connected to SOURCE to generate the gate voltage of TCasc. For an expected $\Delta V_{SRC} \approx 80\,\mathrm{mV}$ this implementation yields $\Delta I_{bias} = 0.2\,\mu\mathrm{A}$. The circuit can tolerate a change of $\Delta I_{bias} = 1\,\mu\mathrm{A}$, leaving substantial reserve which the circuit can compensate[8]. Omitting the cascode would yield $\Delta I_{bias} = 8\,\mu\mathrm{A}$ which the circuit could not tolerate.

Due to chip-to-chip process variations, R_{bias} and hence the bias current spread, resulting in gain dispersion. A suitable mechanism has therefore implemented to tune the bias current globally for each ASIC. Furthermore, on-chip variations need to be expected causing a further spread across the matrix, which is minimized by the addition of a pixel local trim mechanism. The trimming mechanisms are depicted in figure 6.14. The global trim mechanism is based on changing the globally generated cascode gate voltage through a simple resistive DAC. A change of the cascode voltage directly converts into a change of voltage across the resistor and hence into a change of the bias current. The pixel local trim is implemented by additional series resistors which can be bypassed. Monte Carlo simulations have shown that the expected gain dispersion after trimming is less than 1% across the pixel matrix.

[6] for 3.47, the mathematical noise spectral density has to be used which is 1/2 the physical spectral density and $A_2 = 2$

[7] In F1 the power budget is in fact exceeded in MSDD mode, which shortens the maximum burst length due to the energy being supplied by capacitors (see also section 4.2).

[8] ΔI_{bias} is compensated by the filter during the integration phase and hence G_{m} stays constant.



(a) Nominal transfer characteristic of the N-Input front-end (g_m=1.3 mS) determined by transient simulations including the filter circuit. The dynamic input range of the subsequent ADC stage are 800 mV.

(b) The nonlinearity can be exploited as a compression mode when the transconductance is reduced and several photons are attributed to the first bin.

Figure 6.15: Simulated transfer characteristics of the N-Input front-end.

6.2.4 Dynamic Range

The dynamic range of this proposed front-end topology is limited. No suitable compression mechanism has been implemented yet. The triode compression is not applicable because the current in the NMOS transistor decreases due to an incident signal. To push the transistor in the triode region, a rising current would be needed. In principle, the capacitive compression would be applicable, due to stringent timing schedule, this circuit could however not be improved yet to comply with the required 4.5 MHz operating speed.

The transfer characteristic of the circuit is shown in figure 6.15a, including a zoom for the first 30 photons. The displayed fit has been calculated for the first two photons to obtain the residual from a linear system. The transfer characteristic is non-linear due to the fact that the current in the amplifying transistor decreases with the signal and hence loses transconductance. The effect is however too small to be exploited as a compression mechanism when the gain is set to 1 keV per ADC bin. However, decreasing the current and hence the G_m of the transistor and attributing several bins to the first ADC bin, the non-linearity could possibly be exploited as a compression (see figure 6.15b). The high-gain non-linearity and compression mode applicability is under study involving the DSSC calibration group.

7

Selected Measurements

This chapter presents selected measurements on various test chips and the full scale F1 matrix. The measurement setup is presented shortly in section 7.1 to give the reader an idea of the lab system. It has been duplicated and distributed to collaborating groups in Munich (XFEL GmbH), Hamburg (DESY) and Milano (Politecnico di Milano). Final calibration of the system is a dedicated work package in the project. Progress on calibration has been reported for instance in [40]. From the Heidelberg group, F1 matrix measurements are mainly being carried out by Jan Soldat for a separate doctoral thesis [55], while the ADC is being characterized by the DESY FEC group. Some of the results of these characterizations are presented here to demonstrate the full scale functionality of the ASIC section 7.4, the measurements are referenced accordingly. The author of this thesis presented here has mainly been involved in measurements of debugging nature on single pixels and measurements on the smaller test chips. The measurements presented section 7.2 include:

- Capacitance measurements on the input of the MSDD for various assemblies.
- The MSDD front-end characteristic in the F1 ASIC, showing the dispersion to be expected.
- Noise measurements which reveal performance degradation for larger matrices.
- A measurement showing the supply sensitivity of the N-Input topology.

7.1 Measurement Setup

The measurement setup is depicted in Figure 7.1. It is based on a custom FPGA board, which besides the central FPGA contains a USB interface, a circuit to generate the fast 695 MHz clock, connectors which allow to plug in the final system regulator and IO boards and the lab main board. Small regulator board replacements have been designed which allow to feed static supply voltages directly from a power supply and another which feature the real ASIC power cycling circuit. The lab main board contains a connector on which a small carrier with the device under test (DUT) can be plugged. The mainboard is very flexible such that for each of the test chips, there is only a dedicated host carrier needed. The main board features auxiliary circuits for characterizing the ASIC such as an external high precision DAC for characterizing the ADC, a circuit for external current injection for the current readout front-end, etc. PCB design is mainly done by Jan Soldat [55].

A custom C++ software project has been started from early on in the project. It integrates all required features to operate the DUT, including slow control, dynamic control, readout and archiving of the data. Automated measurements are implemented to sweep configuration settings, signal

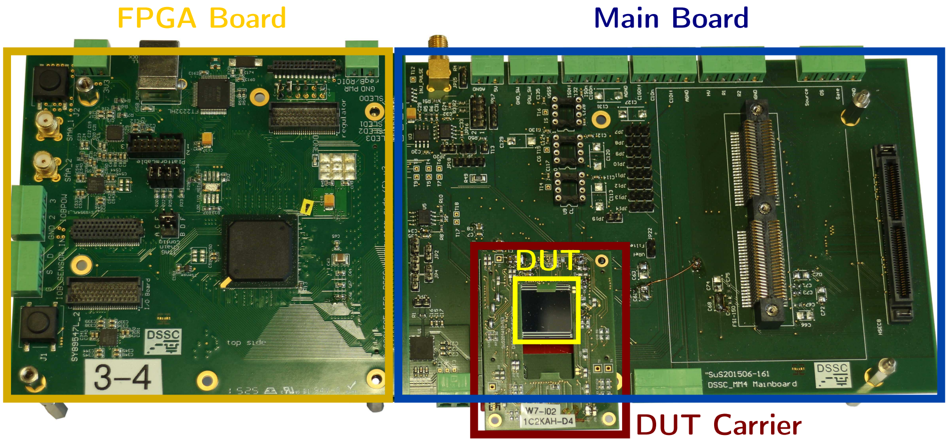


Figure 7.1: Photograph of the lab setup. The device under test (DUT) is hosted on a small carrier. In this way, different DUTs can be tested with the same setup. The PCBs have mainly been designed by Jan Soldat [55].

injection settings etc. A scripting feature has been implemented which allows to run sets of complex measurements, including also a simple loop feature. To archive the data, the software makes use of the ROOT framework [56]. A dedicated software tool has been developed within [42] for data analysis. For the measurement, every single data point is archived, the analysis software comprises for instance the features to plot the recorded data along a burst, histogram data of all settings, computation of the ADC non-linearities, data fits and much more. Details will be available in [42].

7.2 Pixel Characterization Measurements

7.2.1 MSDD Front-End Input Capacitance

The DSSC system runs at a frequency where the white series noise is dominant, the ENC is thus inversely proportional to the input capacitance (equation 3.47) and thus very important. The input capacitance of the MSDD front-end has been measured for various chips and assemblies using the internal charge injection circuit. A measurement procedure has been developed which uses the MIM caps which can be statically added to the input node to estimate the input capacitance. A set of three capacitors are implemented here, these are originally intended to decrease the gain of the front-end for higher photon energies ($> 1\,\mathrm{keV}$).[1]. The procedure is very simple: based on observing the degradation of the gain when adding a known capacitance C_{add}, the input capacitance can be calculated. The pixel injection circuit is used to measure the gain once without any added capacitance and the procedure is repeated with the exact same settings despite for adding a known capacitor to the input. The exact gain for this measurement is irrelevant, the capacitance can be calculated from the relative degradation. The exact value cannot be determined because C_{add} is not exactly know due to process variations but a sufficient estimate can be retrieved. If we call a the nominal gain without added capacitance and a_{addedC} the gain with an added capacitance, we can use the following relations to calculate the capacitance:

[1]at the expense of higher noise, but limiting the required feedback capacitance in the FCF

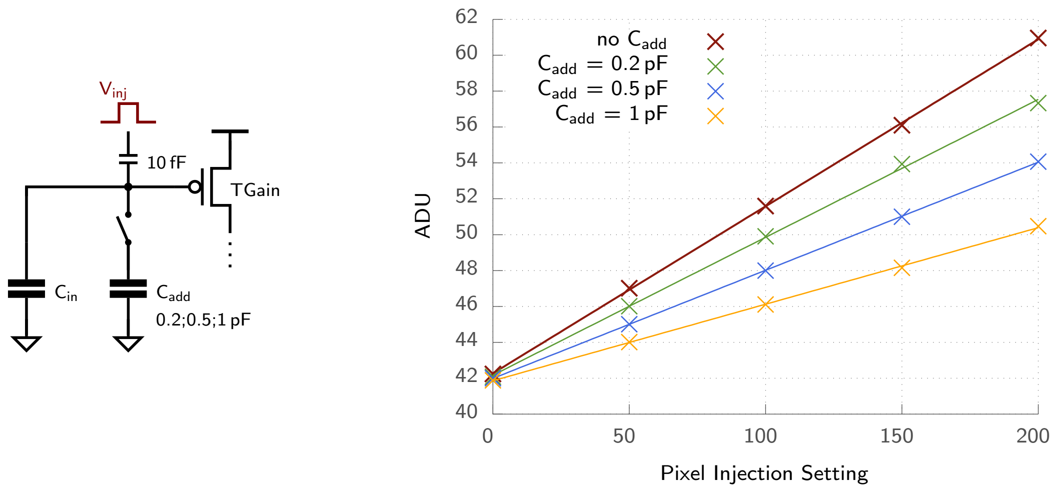


Figure 7.2: Example of the measurement method used to estimate the input capacitance. The gain settings for all measurements are the same despite for added input capacitances. From the slope degradation, the input capacitance can be estimated.

$$a \propto \frac{1}{C} \rightarrow r = \frac{a_{addedC}}{a} = \frac{C_{in}}{C_{in}+C_{add}}$$

The input capacitance can thus be estimated by:

$$C_{in} = \frac{r}{r-1} C_{add} \tag{7.21}$$

According to equation 5.33, the injected charge depends on C_{in}. The low gain injection mode is used for the measurement, because the effect of the small injection capacitance of 10 fF can be neglected for the expected input capacitance of $>$ 200 fF for a sufficiently good estimation.

Figure 7.2 shows an example of the measurement for a single pixel. By adding a 1 pF capacitor to the input, the gain (slope) degrades by a factor of ≈ 2 which leads to a capacitance of $\sim$ 1 pF. Figure 7.3 shows a map and histogram of the F1 chip bump bonded to an MSDD sensor.

The mean capacitance of 900 fF extracted with this measurement from the F1 and MM4 chips is far more than the expected $400 - 500$ fF. The extra capacitance stems from the DEPFET cascode PMOS transistor which is connected to the input. The n-well of this PMOS transistor is unfortunately also connected to the input which could have been avoided. The extra capacitance is mainly attributed to this fact.

The MM6 chip shows a significantly smaller input capacitance of $\sim$ 270 fF. (figure 7.4). The DEPFET front-end is not included, the input transistor is smaller and the landing pad of the bump (last metal in the ASIC) has been made smaller than the recommended size for test purposes. The two peaks in the histogram in figure 7.4 are separated by the capacitance of the sensor anode which is estimated as $\sim$ 60 fF with the applied measurement method. The MSDD mini matrices have a size of (8×8), whereas the MM ASICs have a size of 8×16. The upper half of the sensor matrix only comprises landing pads and no sensor pixels.

For F1, a re-fabrication would be possible, where the DEPFET front-end would be disconnected by patching only one of the metal layers. Additionally, the bump landing pads could be shrunk to decrease the capacitance further.

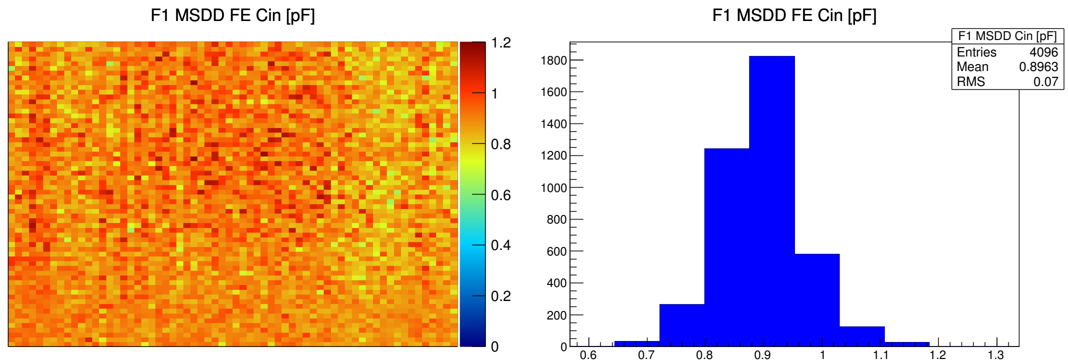


Figure 7.3: MSDD front-end input capacitance measurement for an F1 flipped to an MSDD sensor. The F1 ASIC includes the DEPFET front-end and therefore has an increased input capacitance.

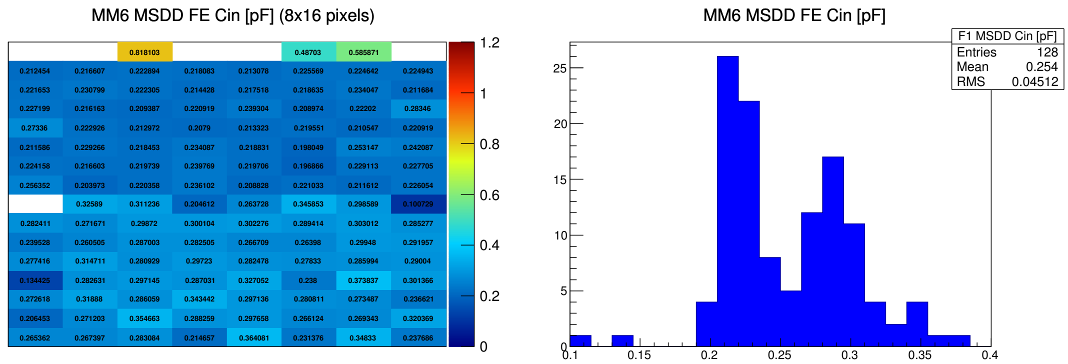


Figure 7.4: MSDD front-end input capacitance measurement for an MM6 chip which does not feature the DEPFET front-end, flipped to an (8×8) MSDD sensor. The upper 64 pixels only have a landing pad on the sensor, there is no sensor pixel. The two peaks in the histogram correspond to the capacitance of the MSDD anode which contributes $\sim 60\,\mathrm{fF}$. The pixels in the map which have a capacitance of $> 400\,\mathrm{fF}$ (including the white pixels) lead to wire bond pads for testing and thus have a higher capacitance.

Assembly	Measured C_{in}	Front-End
F1 on the probe station	~0.92 pF	MSDD + DEPFET
F1 flipped to a 64 × 64 MSDD	~0.90 pF	MSDD + DEPFET
MM4 flipped to an 8 × 8 MSDD	~0.90 pF	MSDD + DEPFET
MM6 flipped to an 8 × 8 MSDD	~0.25 pF	MSDD

Table 7.1: Measured input capacitances. The measurement method relies on a reference MIM cap in the pixel, the fact that the flipped F1 has a lower capacitance than on the probe station can be attributed to fabrication mismatch.

7.2.2 F1 MSDD Front-End Characteristic

The working principle of the F1 MSDD front-end is illustrated in figure 7.5 by means of the internal charge injection. A schematic of the circuit can be found in section 5.3.2.2. The measurement was done on the F1 ASIC, with only a single pixel powered to probe the characteristic of a single pixel without the influence of a large matrix. The measurement can however serve to estimate the gain dispersion across the matrix taking into account the global reset voltage and the voltage drop across the matrix.

The bias point, i.e. current in the input branch is set by means of the reset DAC setting, a larger value corresponds to a larger current. In the lower plot, the reset voltage (reset DAC setting) is plotted against the gain for small signals, extracted by the pixel injection sweep. Moving along the x-axis from the left, the bias current and hence the gain rises until the compression resistor starts to push the transistor into the triode region, increasing the bias current further consequently lowers the gain.

For the best noise performance, the gain should be maximum (red curve) for which it has roughly been set to 1 keV (using the filter and ADC gain settings). For this bias point, the top plot shows, however, that there is no compression. To move to the compression regime, the gain must be lowered (moving to the right on the x-axis in the lower plot) which decreases the noise performance because the transconductance of the transistor is lowered.

In the lower plot, the maximum voltage drop on VDDA across the chip of $\sim$ 50 mV is annotated, a DAC bin corresponds to 110 µV. This voltage drop also represents the maximum spread of the reset V_{gs} for the gain transistor across the matrix because the reset voltage is shared among all pixels. The gain of the front-end is therefore expected to spread by a factor of 3 across the full pixel matrix. Moreover, the transfer characteristic also differs in shape significantly (top plot, green to yellow curve). Possible solutions are under study to used the filter feedback caps and ADC gain to at least equalize the gain across the matrix. The adaptive reset generation which has been proposed by the author (see section 6.2) has been implemented also in this (P-Input) topology by the corresponding designer, to eliminate this spread.

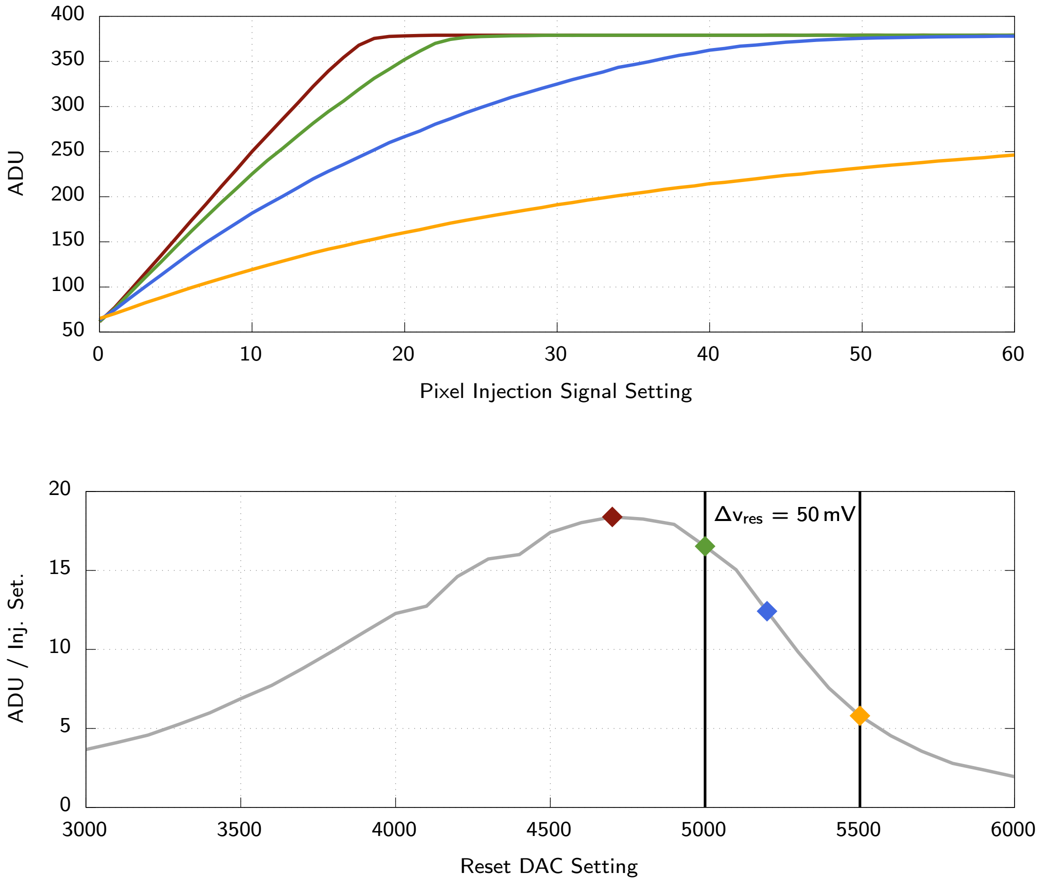


Figure 7.5: The top plot shows a high gain injection sweep, i.e. the front-end transfer characteristic for several reset voltages (reset DAC settings). The pixel injection setting which defines the charge to be injected on the input node is plotted against the output ADU. The red curve corresponds roughly to 1 keV per ADC bin. In the lower figure, the slope of a linear fit for small signals (setting 0-4), i.e. the gain for the first few photons is plotted. $\Delta v_{res} = 50\,\text{mV}$ is the expected spread of the reset voltage across the pixel matrix which causes a gain dispertion of a factor of 3.

7.2.3 Noise

To probe the total noise of the system, the fluctuation of the pedestal needs to be referred to the gain. Therefore, the gain needs to be determined precisely using a known signal source. A radioactive source which has known spectral emission lines can for instance be used. For the presented numbers, an Fe^{55} source was used for the corresponding measurement, which has a $K-\alpha$ emission line at 5.9 keV and a (weaker) $K-\beta$ emission line at 6.5 keV.

Figure 7.6 shows an exemplary spectrum of Fe^{55} which has been measured with the N-Input front-end test chip (D0M1). This small 4×2 test chip has no bumps, the sensor anode is therefore wire-bonded. The wire-bond pad has been optimized to contribute as little as possible. The capacitance measurement described in section 7.2.1 yields ~ 350 fF. The photons from the radioactive source arrive asynchronously to the operation of the time variant Flip Capacitor Filter and are weighted according to their arrival time with the trapezoidal weighting function. The photon peak is therefore less pronounced if the flattop is short with respect to the integration time, because in this case the probability is high that photons arrive during the slopes of the trapezoid. The flattop has been kept short because an unexpected rise in the noise has been observed for long flattops. The cause is under study. A summary of the measured noise figures is given in table 7.2.

ASIC (pixels)	Measured Noise [e^-] (rms)	Measured C_{in}	ASIC Front-End
MM4 (8×16)	130	~ 0.9 pF	DEPFET + MSDD
(single pixel)	101		
F1 (64×64)	not available	~ 0.9 pF	DEPFET + MSDD
(single pixel)	150		
MM5 (8×16)	52	~ 0.28 pF	only MSDD
D0M1 (4×2)	56	~ 0.35 pF	redesigned MSDD

Table 7.2: Measured noise figures for the **MSDD** readout with various ASIC variants. Shaping time is 50 ns for all measurements corresponding to 4.5 MHz operating speed. In single-pixel mode, only one pixel in the matrix is powered.

The given numbers are estimates because some uncertainty remain due to the ADC binning which is not accounted for. From the numbers, the following conclusions can be drawn:

- Operating MM4 in single pixel mode gives reasonable noise numbers when accounting for the large input capacitance.
- For a small scale variant of F1 without the DEPFET front-end and with a separate supply for the input branch (MM5), a good noise level of 52 e^- rms can be reached for the 4.5 MHz operating speed.
- For MM4, a small scale variant of F1, where the input supply is shared with the global analog supply, the noise already degrades when the full chip is operated. This is attributed to the bad power supply rejection of the MSDD front-end on this chip.
- When operating the full scale matrix, the gain cannot be set high enough in order to detect the $Fe^5 5$ $K-alpha$ peak. No noise numbers are available so far.
- The N-Input front-end, which contains the new features to improve the performance in a matrix environment, shows good noise performance on a small test chip (56 e^-). Further improvement of the input of the noise can be expected for a bump-bondable chip with smaller input capacitance.

A larger matrix with some further improvements has been submitted and will yield a better appraisal of the performance in a large matrix.

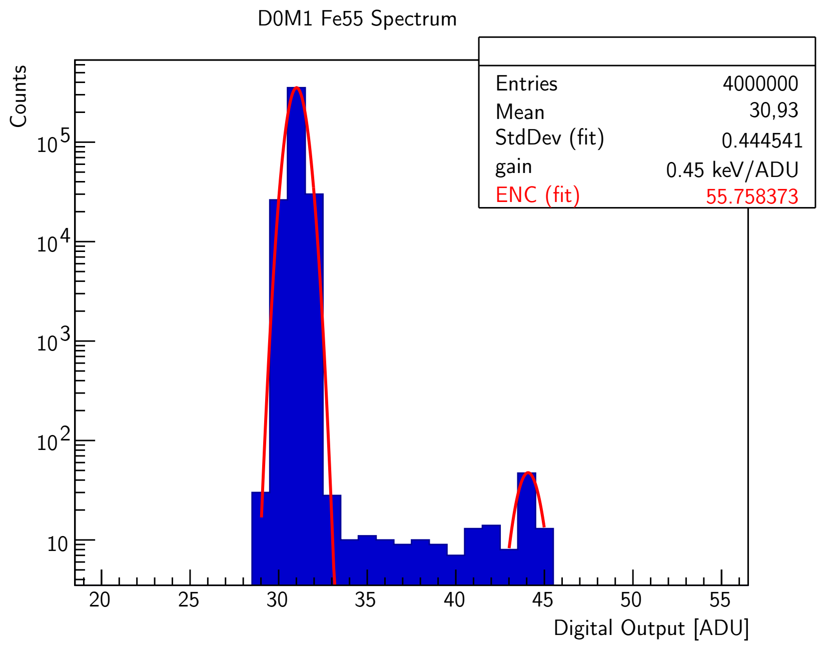


Figure 7.6: Fe^{55} spectrum measured with the N-Input front-end on a wire bonded test chip (D0M1).

7.2.4 NInput Ground Sensitivity

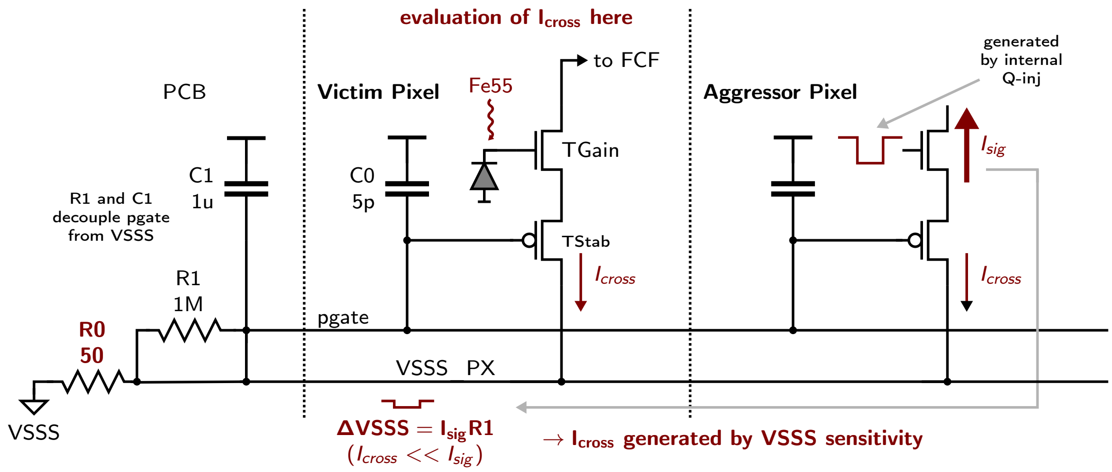


Figure 7.7: Schematic of the supply sensitivity measurement. The VSSS resistance has been exaggerated by a factor of 10 to mimic a long pixel column. An aggressor pixel causes a signal current in the supply line which is evaluated in the victim pixel.

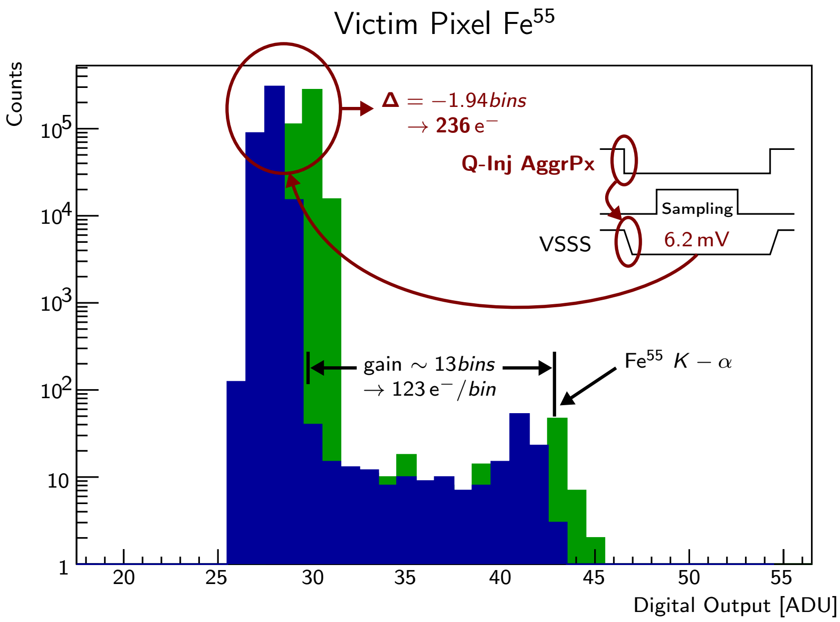


Figure 7.8: Ground sensitivity measurement on the N-Input front-end. The corresponding schematic is shown in figure 7.7. A signal is injected in an aggressor pixel (Q-Inj AggrPx) which generates an exaggerated voltage step on the ground line (VSSS) due to an exaggerated ground line resistance. This moves the signal in the victim pixel and allows to calculate the ground sensitivity of $38\,e^-/\,mV$.

The sensitivity on the ground (VSSS) line of the N-Input charge readout front-end has been measured on the D0M1 test chip using the setup depicted in Figure 7.7. The goal of the measurement is to evaluate the signal induced by the change on the VSSS line in a victim pixel when a large signal is present in an *aggressor* pixel(s) in the same pixel column. The situation is worst if the aggressor pixel is at the top of the column because this pixel sees the largest supply resistance and thus creates the largest step on VSSS due to a signal current. The 2×4 pixels D0M1 test chip only has a column length of 2 pixels, the resistance on the VSSS line is therefore negligible. To mimic a full column, a $50\,\Omega$ resistor is soldered in series to the VSSS pin of the chip. On the full scale chip a resistance of $\sim 5\,\Omega$ is expected, a factor of 10 larger has been chosen to amplify a single aggressor pixel.

In the aggressor pixel, a signal current is generated through the internal charge injection (I_{sig}) and causes a voltage step on the ground line of $I_{sig} \times R0$ (I_{cross} can be neglected here for a good estimation). This voltage step induces a spurious current in all pixels of the same column in a larger matrix according to the pixels ground sensitivity.[2] An Fe^{55} source has been used to roughly set the gain of the victim pixel to $123\,e^-/ADU$ ($\sim 0.5\,keV$). The voltage step on the supply line has been measured using an oscilloscope and the movement of the Fe^{55} spectrum due to the voltage step has been recorded to evaluate the ground line sensitivity. Figure 7.8 shows that the mean of the histogram moves by 1.94 bins due to a voltage step of 6.2 mV on the ground line, which yields a ground sensitivity of $38\,e^-/\,mV$. For a realistic pixel column, a signal which completely starves TGain in the pixel at the top of the column, a voltage step of $\approx 5\,\Omega \times 130\,\mu A = 0.65\,mV$ is generated, which would inject a

[2] The horizontal power busses in the large matrix are only weak and it must also be accounted for several pixels in a row receiving a signal.

signal of $\approx 25\,\mathrm{e}^-$ in all pixels of the same column at the mentioned gain of $127\,\mathrm{e}^-/ADU$. This value is considered a sufficient suppression of signals on the ground line. Exact studies need to take into account the target experiment and pixel hit patterns which is subject of future work.

7.3 13 bit Rail-to-Rail Voltage DAC

Figure 7.9 shows the measured characteristic of the internal 13 bit voltage DAC in the high range mode. The circuit is presented in section 5.4. A zoom is shown for the input voltage range which corresponds to first photons (middle plot). The fit was calculated in the range corresponding to the first 30 ADC bins. The maximum INL in this range is 200 µV corresponding to 6.4% of an ADC bin. The average bin size is 110 µm, which corresponds to 28.4 steps for a nominal ADC bin size of 3.125 mV. This range of the DAC has been used for the measurements presetned in section 7.4.2. The decrease of the INL for DAC settings > 3000 when the NMOS current mirror in figure 5.23 is caused by the NMOS current mirror losing overdrive voltage and is expected by design. The INL in the low range mode shows complimentary behavior (not shown) since the output voltage is generated by sending the DAC current to ground in this mode. The overall INL can thus be optimized by switching in the middle of the dynamic range.

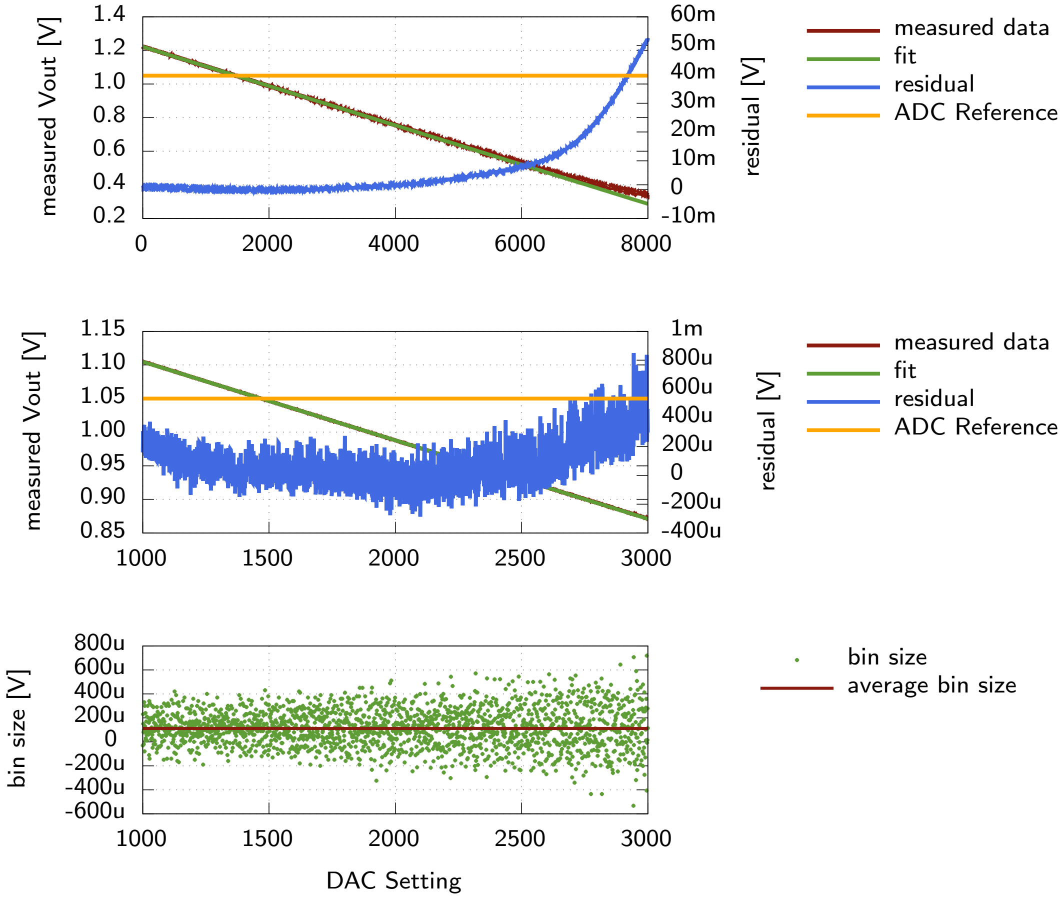


Figure 7.9: Characteristic of the internal 13 bit voltage DAC. Measurement data courtesy of Jan Soldat [55]. The middle plot is a zoom around the ADC reference (100 mV ≈ 30 *ADC bins*).

7.4 Full Scale F1 Matrix Measurements

This section demonstrates the functionality of the full scale F1 matrix. The measurements were carried out by others and the plots were provided. All of the measurements presented in this section were taken at realistic conditions:

- All pixels are operated at the same time.
- Power cycling is employed at a rate of 10 Hz, a single burst has a duration of 600 µs while the power is enabled $\sim$ 100 µs before the burst.
- The full memory depth is used and a full readout of the chip is performed.

7.4.1 MSDD Front-End

Figure 7.10 shows the full scale F1 matrix bump bonded to a 64×64 MSDD sensor covered with an aluminum mask exposed to a pulsed LED. This measurement marks the *imaging* commissioning of the F1 ASIC. It was taken under realistic conditions, the baseline is subtracted. The applied FCF cycle sequence corresponds to the fast 4.5 MHz mode, except for the flattop, which has been extended significantly in order to allow for the LED to deposit significant energy in the sensor. All pixels are operating in parallel and the gain was roughly set to few keV per ADU. Power cycling has been employed at the target speed of 10 Hz. The full memory depth has been used (800 frames), the figure shows in each pixel the mean of one collected burst. The full chain is thus operational.

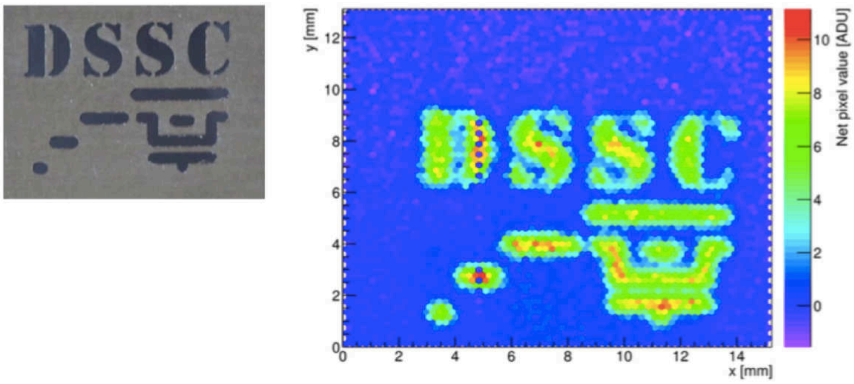


Figure 7.10: Commissioning of a 64×64 pixel MSDD sensor bump bonded the F1 ASIC [55]. The average over one burst is plotted. These are the first real *images* taken with a full format ASIC.

7.4.2 ADC Measurements

The measurements presented in this section have been done by the DESY group and were presented in [57]. A lot of effects which are visible in these measurements stem from horizontal voltage drops. These have been under estimated before the submission and cause problems due to the fact that the reference circuits have been placed row wise at the side of the matrix because a negligible horizontal dependency was expected. The reference circuit generates a reference current which is mirrored into the pixel by distributing the gate voltage of the mirror transistor. Consequently, the difference of reference to the supply line is essential which is not constant along a pixel row. The entire ASIC is therefore relatively sensitive on the supply voltage because the biasing for all amplifiers and reference voltages in the pixel are derived from the reference current.

Gain Trimming

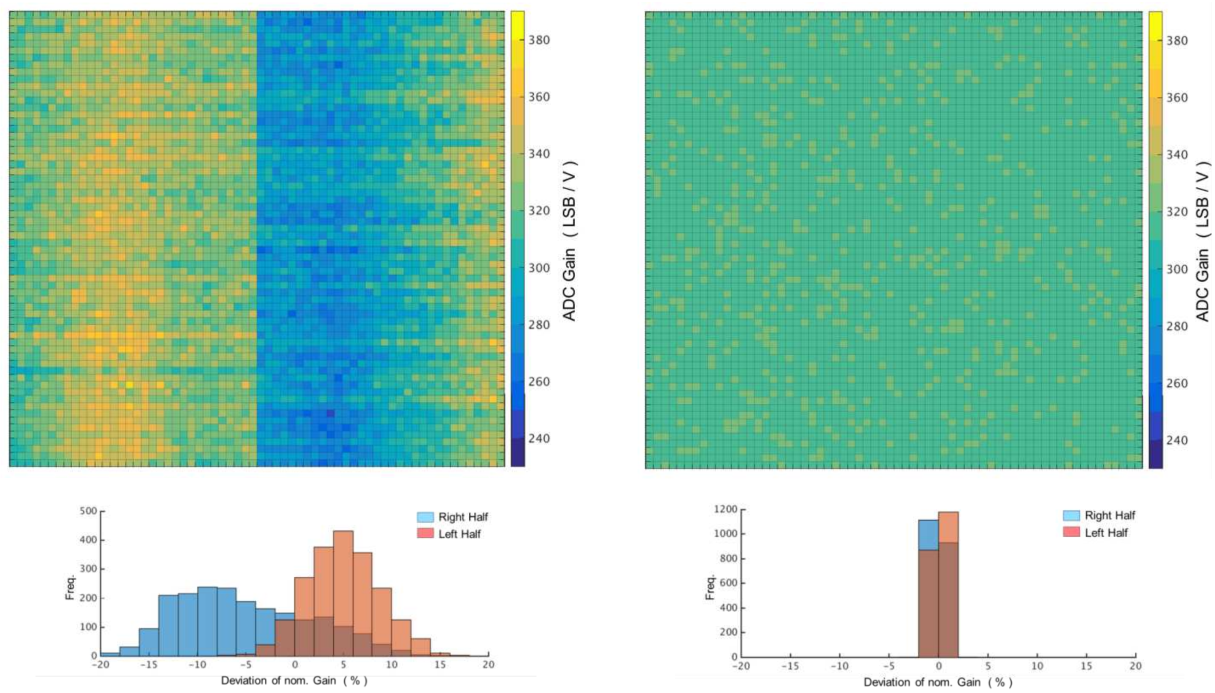


Figure 7.11: Equalization of the ADC gain across the F1 pixel matrix by an automated trimming procedure. The left plots show the situation before the procedure, the right plots show the final result. The ADC gain can be adjusted pixel by pixel using a pixel internal 6 bit DAC. Measurement courtesy of DESY FEC [57].

Figure 7.11 shows that the ADC gain can be adjusted such that the final deviation from the target slope in all of the entire 4096 pixel matrix is in the order of 1%. An automated procedure has been implemented by the Heidelberg group (M. Kirchgessner [42] and J. Soldat [55]) and the DESY group, which uses the internal voltage DAC (section section 5.4). Before the trimming procedure, the matrix shows a deviation of 20% which is mainly attributed to horizontal voltage drops. The references are placed on each side of the pixel matrix, which explains the cut in the middle. Nevertheless, the 6 bit DAC which is implemented in each pixel to change the slope of the ADC ramp is sufficient to equalize the gain across the matrix.

Noise

The noise of the ADC in each pixel has been evaluated for the first signal bin, where it is most important for single photon resolution. The used method is based on first measuring the pixel delay

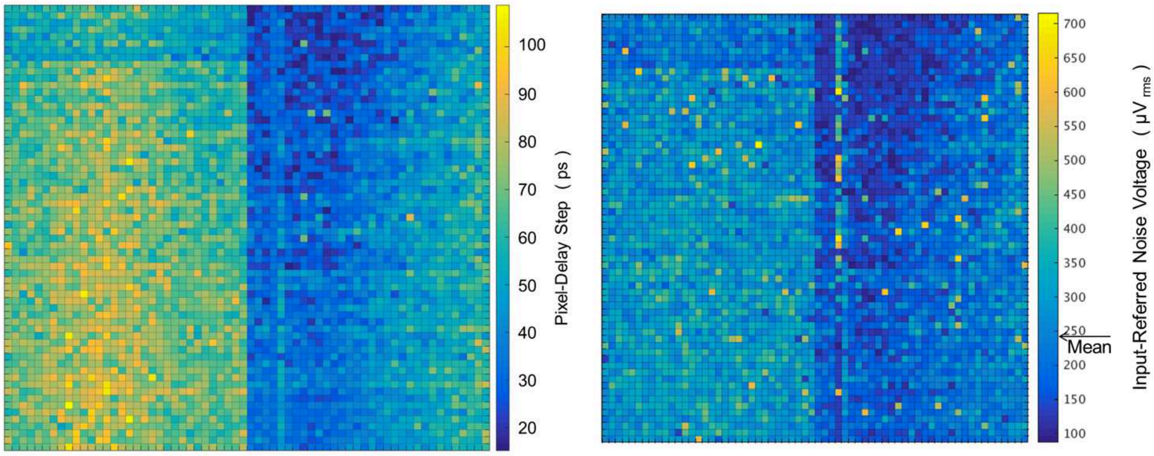


Figure 7.12: Left: Map of the measured mean pixel delay step across the F1 matrix [57]. Right: The pixel delay step can be used to determine the input referred noise voltage of the ADC. The mean across the pixel matrix are ~ 250 µV rms.

in each pixel. This can not be measured directly because it is used to set the inner ADC bin offset. The used methodology was presented in [39]. The ADC characteristic - input voltage against output ADU - is recorded for several gain (ramp slope) settings using the ASIC internal DAC. For each gain setting, a characteristic is recorded for each offset (pixel delay) setting. The pixel delay can hence be calculated in time by evaluating the change of the intercept of the linear regression line for each of these measurements and relating the change to the temporal bin width of 720 ps, which is known from the ADC clock. The ADC characteristic is again measured with the internal voltage DAC (section 7.3).

To evaluate the noise, a second measurement is needed. For a zero input signal, all pixel delay settings are swept. To exclude other noise sources, the reference voltage, which corresponds to a zero signal from the front-end, is digitized. The filter amplifier can be put into permanent reset by shortening the feedback capacitor permanently. This situation avoids any switching effects to characterize the ADC by itself. Sweeping the pixel delay spans ~ 1.5 ADC bins. By evaluating the error function when crossing the ADC bin boundary, the ADC noise can be retrieved in pixel delay steps. The pixel delay which has been measured in the previous step, can in turn be translated to a voltage because the bin width is known in both voltage from the ADC characteristic measurement (DAC sweep) and in time from the ADC clock. These measurements have been carried out on the full matrix by the DESY group, the results for the entire matrix are shown in figure 7.12. The mean input referred noise across the 4k pixel matrix is ~ 250 µV (rms) which is $\sim$8% of a nominal ADC bin and results in a contribution of ENC $= 21\,e^-$ rms for a front-end gain setting of 1 keV[3] per ADC bin and a nominal bin width. It is important to note here, that when determining the ADC contribution to the ENC, the front-end gain has to be taken into account, which determines the amount of charge per ADC bin. For a gain setting of 0.5 keV per ADC, $140\,e^-$ is attributed the first ADC bin and the noise consequently decreases to $11\,e^-$ rms.

[3]in the linear range

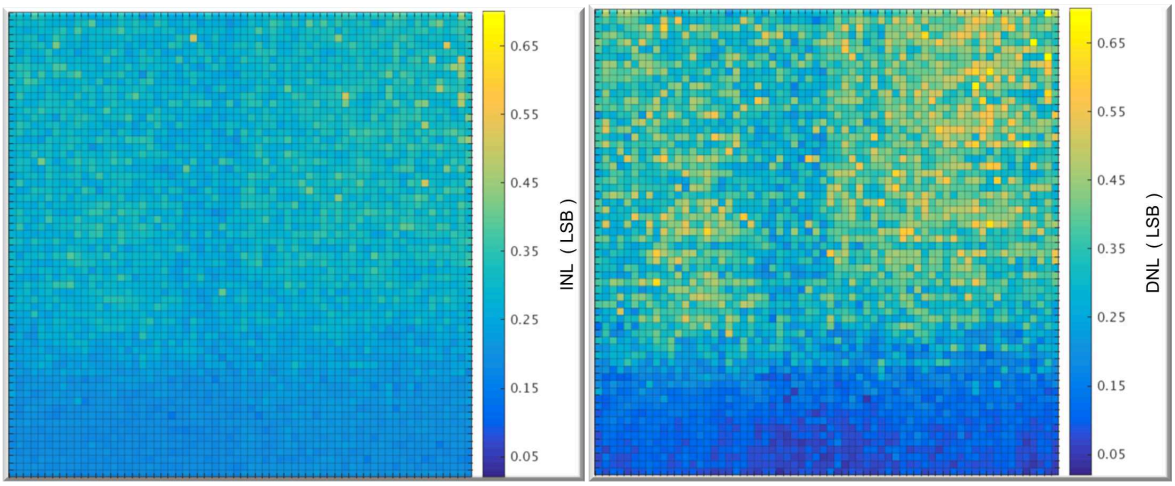


Figure 7.13: INL (left) and DNL (right) of the F1 pixel matrix in RMS LSB [57].

Non-Linearities

The nonlinearities (INL and DNL) of the F1 ADC matrix have been determined for the first 25 bins, where they are most important for single photon resolution. Figure 7.13 shows a maps of the measured values. For 90% of the pixels, the INL remains within ±5%LSB. There is a drop of the DNL which starts at about row 16. This is assumed to result from mismatches in the Gray code transmission lines, the effect is under study. The target is to improve the DNL along the pixel columns.

7.5 In-Pixel Counting ADC

The ADC architecture comprising the in-pixel counter is an alternative topology to the global counting architecture with the aim of improving the DNL. The two architectures are presented in section 5.3.3 and section 5.3.3.5. Test structures have been submitted on two chips so far. The latest, L1, which has been fabricated on the F1 engineering run, has full scale pixel column lengths (8×64 pixels). Figure 7.14 shows a measurement of the DNL on L1. There are some pixels which are not working at all (white pixels) which is due to the fact that the clock signal cannot be properly received in these pixels. The design is very sensitive to the supply voltage level. However, the potential can clearly be seen when comparing it to the F1 map depicted in figure 7.13 (the color code does not match) which uses the global counting ADC architecture.

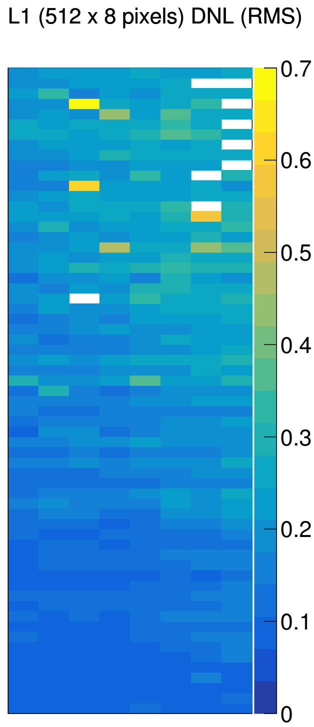


Figure 7.14: Map of the measured DNL of the in-pixel counting ADC architecture (L1 test chip). A very good DNL can be achieved, some pixels however do not receive the clock signal properly and thus count erroneously. Measurement data courtesy of [42].

7.6 Conclusions from the Presented Measurements

The presented measurements are only a small subset of the characterization and calibration efforts which is in progress and involving several groups.

The working principle of the triode compression has been shown with an example measurement. The shown measurements make clear, that a large gain spread in the front-end is expected across the full F1 matrix. The remaining gain settings along the signal processing chain such as the ADC ramp slope can be used to try to homogenize the curves across the matrix, this will however inevitably result in a loss of performance in terms of dynamic range and noise figures. Besides the gain spread, the sensitivity on the supply line is an issue in F1, which has also been addressed in the redesign. An improved topology featuring a voltage drop insensitive biasing technique has been proposed by the author, proven with measurements on a first test chip. A further test chip is in fabrication which comprises three different front-end variants to be evaluated for the next full scale submission.

Furthermore it is evident that the input capacitance needs to be reduced for the MSDD front-end. A discussion is in progress whether the DEPFET front-end needs to be included in F2 because the time when it will become availability is still unknown. A possible solution could be to include the required circuits but not connect them. Dedicated DEPFET or MSDD chips could be fabricated this way by switching a single metal layer during fabrication.

In the ADC, the in-homogeneity is mainly attributed to the combination of horizontal supply voltage drops and the according reference being placed at the side of the matrix. For the F2 ASIC, a simpler reference will therefore be placed directly in each pixel. Furthermore, the pattern of the supply bumps will be optimized to decrease the horizontal dependency. The bump pattern was fixed very early by the design of the module.

Overall, the full functionality of the F1 ASIC has been demonstrated with the measurement and the ASIC can be seen as a success. The shortcomings in the front-end are well understood driving the design phase of F2 which is planned to be submitted late in 2016.

8

Conclusion

8.1 Conclusion

This work has presented the DSSC camera project and in particular the design of the sensor read-out ASIC. The DSSC camera is being developed for low energy experiments at the European XFEL and faces unprecedented challenges as it is required to combine low noise performance down to 0.5 keV with a high dynamic range of up to 10000 photons and very fast readout speeds of 4.5 MHz. These challenges have further been complicated by the unexpected unavailability of the originally foreseen DEPFET sensor due to fabrication issues. New concepts had to be adopted along the way since the DEPFET sensor had been the central element solving the low noise and dynamic range challenges. The advantages of the DEPFET against other sensor types has been discussed in chapter 3.

The core topic of this thesis is the design and integration of a large scale pixelated readout ASIC. The ASIC pixel has been integrated successfully comprising circuits which were contributed in part by collaborating institutes. The author's effort has peaked by submitting an engineering run featuring the first full scale matrix with 4096 and a die size of $14.9 \times 14.0\,\mathrm{mm}^2$ along with further test chips. All ASICs submitted during the course of the years have been completely functional. Tweaks were necessary here and there but overall every submission has been successful. All the designed concepts including the in-pixel memory, digital control logic integration methodology of the large scale matrix have been proven.

The operation of an F1 flipped on an MSDD been started in early 2016. F1 is completely functional which is considered a milestone for the project and can be used to commission the full system. Problems were identified in the analog front-end design, reaching adequate performance is doubtful for the full scale matrix. Again it must be emphasized that this design had to be implemented in a short time frame, F1 is basically the first matrix chip which comprises the MSDD front-end.

Following the discovery of the shortcomings in the F1 front-end, an extensive R & D phase has led to new concepts, where the author has contributed significantly. A prove of concept of these new concepts has been made on a 2×4 test chip. The circuit reaches a noise level of $56\,\mathrm{e}^-$ while implementing concepts to withstand the rigorous environment of a large scale matrix. A further improved 8×16 test chip (MM7) comprising three different front-end variants has been submitted and is currently in fabrication.

8.2 Summary of Own Contributions

This section summarizes the contributions of the author to the DSSC project:

- The design of the pixel memory has been finalized and the matrix readout architecture designed and implemented up to the full scale matrix. This design is full custom despite for the core SRAM cell, which is available from the foundry in a dense layout. Emphasis has been on compactness and suitable control to integrate it in the pixel. A capacity of 800 words has been reached which is a distinguishing property versus the other detectors developed for the EuXFEL.
- The digital control block for the chip has been designed and implemented using a semi-custom design flow and a standard library available for the target process from CERN. The ASIC is basically a system on chip with a minimum command telegram based control interface. The on-chip digital control block includes controllers for the entire ASIC including a JTAG slow control interface, front-end sequencer, memory and readout controllers and data serializer. The architecture is completely own work while the recipe for the physical implementation has been available from another project. However, the implementation scripts had to be adapted significantly for the target design and standard cell library.
- The entire slow control domain of the ASIC has been designed and implemented, including full custom design in the pixel and matrix and the synthesized interface. The pixel slow control register features a direct access mode to program single pixels and a fast chain mode to program the entire matrix. The maximum speed of the slow control domain is 50 MHz which suffices to reprogram the chip between two bursts. Only the JTAG state machine has been available and reused from another project.
- A 13 bit voltage has been designed and implemented. The core has been available in the group and has been ported to the target technology and extended to 13 bit resolution.
- The pixel has been integrated and embedded in a matrix structure up to the full scale 4k matrix. A floorplan of the pixel has been developed taking into account the requirements for all of the circuits. The core parts of ADC, FCF, (F1) MSDD front-end and pixel injection circuits have been provided while most of the layouts are own work. The pixel layout is very dense, the routing and MIM capacitor layers are used extensively both for local and global routing. The challenge here is to optimize the floorplanning such that the memory capacity can be maximized without affecting the functionality of the other blocks.
- Simulations have been carried out on the pixel level and a system level simulation has been set up which includes the lab software and FPGA firmware. Besides verifying the functionality of the ASIC this setup has been proven its value also for the FPGA firmware simulations.
- The ADC concept using the in-pixel counter was implemented, where the 695 MHz clock is transmitted to all pixels. Design work here includes simulation and layout for a test chips including the adaption of the existing transmission lines, transmitter and in-pixel clock receiver for the required clock speed of 695 MHz.
- 11 test chips have been designed and submitted, ranging from dedicated circuit test chips in the case of the memory to small pixel matrices. The schematics and layouts for all chips mentioned in section 5.7, have been drawn and verified (DRC, LVS). The first engineering run of the project

has been submitted. The run contained the first full scale F1 matrix and four further test chips.

- Several novel contributions have been made to the analog-front end design. A novel capacitive signal compression technique based on own ideas has been proposed and implemented on a test chip. An alternative front-end configuration has been proposed with the focus on implementing better robustness with respect to operation in a large matrix. Detailed analog simulations have been performed including noise simulations, Monte Carlo analysis and pixel column level simulations.
- The first test setups have been designed and implemented, including the software and PCBs. The setups, firmware and software have been vastly expanded and improved since this work is also within the scope of two separate doctoral thesis ([55], [42]). Some work has nevertheless been contributed continuously by the author.
- Measurements have been implemented and performed, mostly on single channels and of detective nature to understand the problems in the analog front-end.
- Sensor layouts have been checked and verified for physical compatibility to the ASIC.

8.3 Outlook

A very challenging period is still ahead, as the first ladders (16 ASICs, 512×128 pixels) have been assembled and are presently in the commissioning phase. Strong efforts are undergoing to achieve the best possible performance of the first design. A test beam is planned for the end of 2016. In parallel, we are awaiting the return of the MM7 test chip which includes new front-end variants designed to improve on the F1 shortcomings. Measurements are being prepared in order to compare the new approaches as fast as possible. The second full scale ASIC engineering run to fabricate F2 is being prepared. The main work required here is the rearrangement of the power pads in order to minimize the horizontal supply voltage drops among further minor changes. The digital domain of the chip does not require any changes. The F2 pixel will be finalized when the three front-end variants on MM7 have been evaluated and the final version for F2 is chosen.

A

N-Input Front-End Details

A.1 Small and Large Signal Circuit Modeling

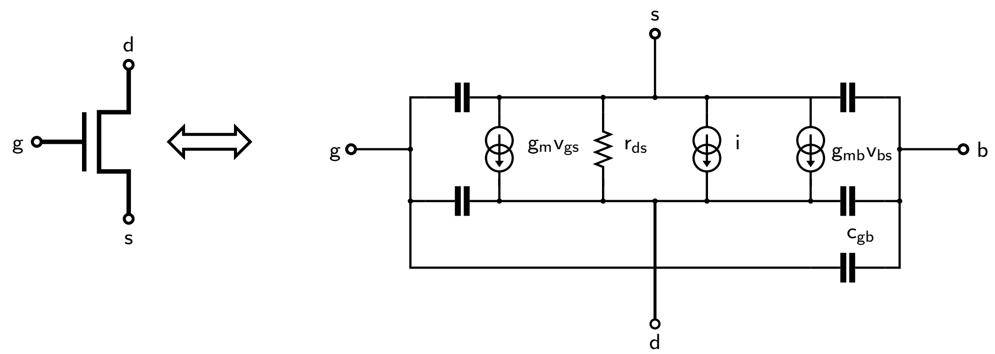


Figure A.1: Small signal equivalent circuit of an NMOS transistor.

In order to analytically and theoretically analyze electronic circuits, so-called small signal equivalent circuits are widely used. The behaviour of more complex components is modeled as a network of ideal components such as current sources, resistors etc. Their values, the so-called *small-signal parameters*, are determined by the operating point, which is a set of DC voltages defining the state of the device. These small-signal parameters are strongly dependent on the operating point. In general, two kinds of different analysis need to be distinguished: *large-signal* and *small-signal* modeling. In a small-signal analysis, the assumption is that the stimulus changes the operating point. In this way, the small-signal parameters are constants, simplifying the analysis significantly. The relation of interest (as for instance the gain of an amplifier) can be obtained by applying Kirchhoff's laws. The inclusion of capacitors and inductors yields differential equations, analysis is consequently performed in the Laplace or Fourier domain. The transfer characteristic of an amplifier for instance can be derived by an AC simulation which simulates the circuit at frequencies of interest in a defined operating point. The gain and phase are calculated in the small signal equivalent of the circuit. In contrast, large-signal modeling takes into account changes of the operating point and hence changes of the component parameters. The start-up behaviour of a circuit needs to be simulated in transient simulation for instance to reveal if and how the circuit reaches the target operating point. It is evident, that a single small-signal analysis is not suitable for such an analysis. In a transient simulation, the circuit is linearized about the operating point for every point in time of interest. Changes of voltages and currents are taken into account to calculate the operating point for the next point in time of the simulation.

A.2 NInput Small Signal Equivalents

A.2.1 Transconductance

To obtain the transconductance G_m from the input node (gate of TGain) to the filter input, we use the small signal equivalent circuit depicted in figure A.2. Intuitively we expect G_m to degrade significantly by the addition of TStab. This is due to the fact that the source of TGain now sees resistance towards the ground connection (at v_x). As a signal on the input node decreases the current in TGain, the v_{gs} of TStab needs to change. v_x consequently needs to drop, which counteracts against the transconductance of TGain.

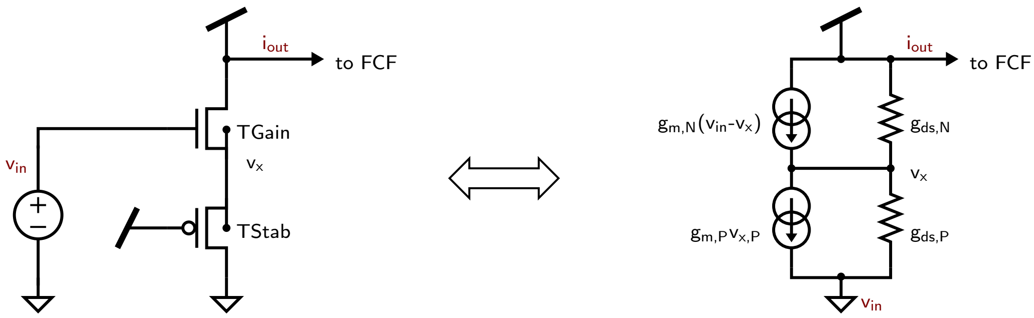


Figure A.2

By summing the currents at v_x, we get:

$$v_x(g_{m,P} + g_{ds,P}) = g_{m,N}(v_{in} - v_x) - g_{ds,N}v_x$$

which gives

$$v_x = \frac{(g_{m,N}v_{in})}{g_{m,N} + gdsn + g_{m,P} + g_{ds,P}}$$

Since i_{out} flows through TStab it can be expressed as:

$$i_{out} = v_x(g_{m,P} + g_{ds,P})$$

which finally yields

$$G_m = \frac{iout}{vin} = \frac{g_{m,P} + g_{ds,P}}{g_{m,P} + g_{ds,P} + g_{m,N} + g_{ds,N}} \tag{A.21}$$

Since $g_{ds} \gg g_m$, we can neglect g_{ds} to get a good approximation with:

$$G_m = \frac{g_{m,N}g_{m,P}}{g_{m,N} + g_{m,P}} = g_{m,N} \parallel g_{m,P} \tag{A.22}$$

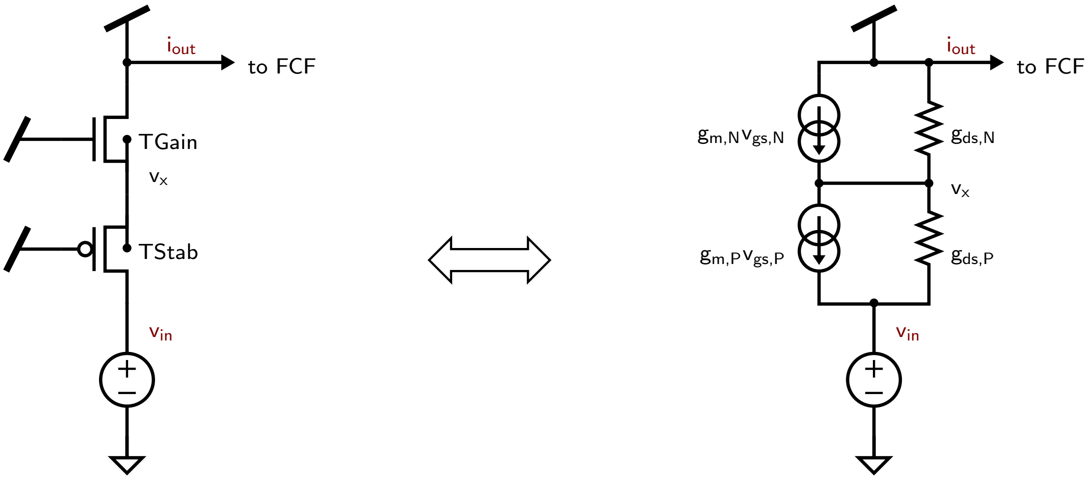


Figure A.3

A.2.2 Ground Sensitivity

The small signal equivalent circuit used to analyze the preamplifier topology for the NInput pixel with respect to sensitivity to a non ideal ground is shown in figure A.3. Capacitances are neglected because we are interested in the low frequency behaviour. Note also that the source of the two transistors are connected to their bulk nodes, eliminating the according current source in the small signal equivalent. The input in this case is the drain of TStab, where a test source (v_{in}) is connected while all other nodes are static (AC ground). Since the signal is a current which is fed into a virtual ground node (current mode filter), we calculate the current which is caused by a change of voltage on the ground line.

Summing the currents at v_{filter} gives:

$$i_{out} = -v_x(g_{m,N} + g_{ds,N})$$

i_{out} also flows through the test source, hence:

$$i_{out} = (v_x - v_{in})g_{ds,P} + v_x g_{m,P}$$

Solving these two equations gives:

$$G_s = \frac{i_{out}}{v_{in}} = g_{ds,P}\frac{g_{m,N} + g_{ds,N}}{g_{m,P} + g_{ds,P} + g_{m,N} + g_{ds,N}} \tag{A.23}$$

where we can again assume $g_m >> g_{ds,P}$ to get a good estimation with

$$G_s = g_{ds,P}\frac{g_{m,N}}{g_{m,P} + g_{m,N}} \tag{A.24}$$

where G_d is introduced as a *source conductance* of the circuit. We can conclude here, that form the point of maximizing suppression of disturbances on the ground, it is best if both the channel resistance and the transconductance of TStab are maximized. This can be explained by the following considerations: A large channel resistance is intuitively good since a resistance is defined by $R =$.

The large transconductance is beneficial because any change of current in TStab by a change of voltage at its drain also needs to pass TGain via its transconductance causing a change of voltage at v_x. Any change of voltage at v_x is however suppressed by a large transconductance of TStab. The transconductance of TStab counteracts against any change of current caused by a change of voltage at v_x.

Taking into account the transconductance at the input node of G_m calculated in the previous section we get a ratio for the sensitivities at the input and ground of:

$$S = \frac{G_m}{G_s} = \frac{g_{m,P}}{g_{ds,P}} \tag{A.25}$$

A.2.3 Input Referred Noise

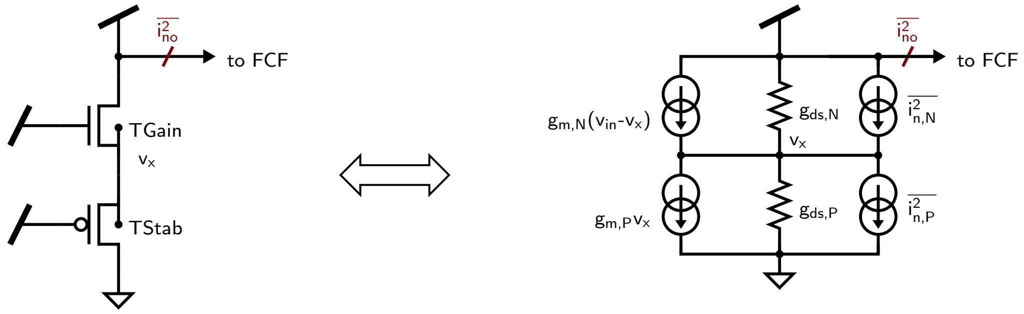


Figure A.4

The presence of noise in the transistor is modeled by a parallel current source. To analyze the effect of the channel noise on the signal, the so-called *input referred noise* needs to be calculated. We therefore first apply an (AC) ground at the input (no signal) and calculate the total noise at the output. The input referred noise is obtained by dividing the output noise by the gain from the input. Figure A.4 shows a small-signal equivalent where the noise current sources are replaced with a single voltage source at the input node to model the input referred noise. In the following, the noise voltage $\overline{v_n}$ will be calculated. Since the two noise sources in the circuit are uncorrelated, their contributions at the output ($\overline{i^2_{no,N}}$ and $\overline{i^2_{no,P}}$) are calculated individually finally summed quadratically:

$$\overline{i^2_{no}} = \overline{i^2_{n,N}} + \overline{i^2_{n,P}}$$

To calculate the contribution of TGain, we set $\overline{i_{n,P}} = 0$. Summing the currents at v_x gives:

$$\overline{i^2_{n,N}} = \overline{v^2_x}(g_{m,N} + g_{ds,N} + g_{m,P} + g_{ds,P})^2$$

while summing at the ground node gives:

$$\overline{i^2_{no,N}} = \overline{v^2_x}(g_{m,P} + g_{ds,P})^2$$

Neglecting g_{ds}, this gives:

$$\overline{i^2_{no,N}} = \left(\frac{g_{m,P}}{g_{m,P} + g_{m,N}}\right)^2 \tag{A.26}$$

For the contribution of TStab, we set $\overline{\mathrm{i}^2_{\mathrm{n,N}}} = 0$. Summing at the FCF input node gives

$$\overline{\mathrm{i}^2_{\mathrm{n,P}}} = -(\overline{\mathrm{v}^2_{\mathrm{x}}}(g_{m,N} + g_{ds,N})^2$$

while summing at the ground node gives

$$\overline{\mathrm{i}_{\mathrm{n,P}}} = \overline{\mathrm{v}^2_{\mathrm{x}}}(g_{m,P} + g_{ds,P})^2$$

again neglecting g_{ds} this yields:

$$\overline{\mathrm{i}^2_{\mathrm{n,P}}} = (\frac{g_{m,N}}{g_{m,P} + g_{m,N}})^2 \tag{A.27}$$

We hence get for the total noise current at the output:

$$\overline{\mathrm{i}^2_{\mathrm{no}}} = \frac{1}{g_{m,P} + g_{m,N}} \sqrt{g^2_{m,N}\overline{\mathrm{i}^2_{\mathrm{n,N}}} + g^2_{m,P}\overline{\mathrm{i}^2_{\mathrm{p,N}}}} \tag{A.28}$$

Which needs to be divided by $\mathrm{G_m}$ to obtain the input referred noise $\overline{\mathrm{v}^2_{\mathrm{ni}}}$:

$$\overline{\mathrm{v}^2_{\mathrm{ni}}} = \frac{\overline{\mathrm{i}^2_{\mathrm{no}}}}{G_m} = \sqrt{\frac{\overline{\mathrm{i}^2_{\mathrm{n,N}}}}{g^2_{m,N}} + \frac{\overline{\mathrm{i}^2_{\mathrm{n,P}}}}{g^2_{m,P}}} \tag{A.29}$$

A.3 Programming Loop

This section present the ICON cell and the stability analysis of the closed programming loop for the N-Input topology.

A.3.1 The ICON Cell

To generate negative DC feedback in the closed loop, an inversion of the signal is needed in the loop, which is implemented by means of a current converter (ICON) cell. The general operating principle of the ICON is depicted in figure A.5. The cell has been reused from the DEPFET design and expanded by an additional resistor at the input to limit the maximum input current which is beneficial for the start-up phase of the circuit.

The target of the programming loop is to eliminate any current intended to bias the preamplifying transistor from flowing into the filter during the burst phase. This condition is achieved in the programming state, when TGain carries exactly the current supplied by $\mathrm{R_{bias}}$ and the ICON input is free of any current. Furthermore, for a perfect cancellation, node $\mathrm{v_x}$ must also settle at exactly the filter reference voltage, any change of this voltage when moving from the programming state to the burst state leads to a remaining offset current caused by the finite resistance at node $\mathrm{v_x}$. Considering the ICON design, the voltage $\mathrm{v_{x,eq}}$ for which the loop is at equilibrium is not defined by the filter amplifier, but essentially by the ratio of TP0 and TN0 and their (shared) biasing gate voltage. This approach has been chosen due to its simplicity as it does not need a sophisticated biasing mechanism A slight jump of $\mathrm{v_x}$ when the ICON is disconnected is therefore unavoidable, leaving some offset current, which has to be cancelled by double correlated sampling.[1]. Because the two gates are shorted, both TP0 and TN0 are in the deep sub-threshold region when the ICON input node is free of current. For a small swing of the input node around $\mathrm{v_{x,eq}}$, the input conductance is hence very small which limits the contribution to the loop gain. The circuit can safely be disconnected during the burst phase because TP0 and TN1 prevent any standing current in the device.

[1]which is employed anyway due to signal shaping (1/f noise)

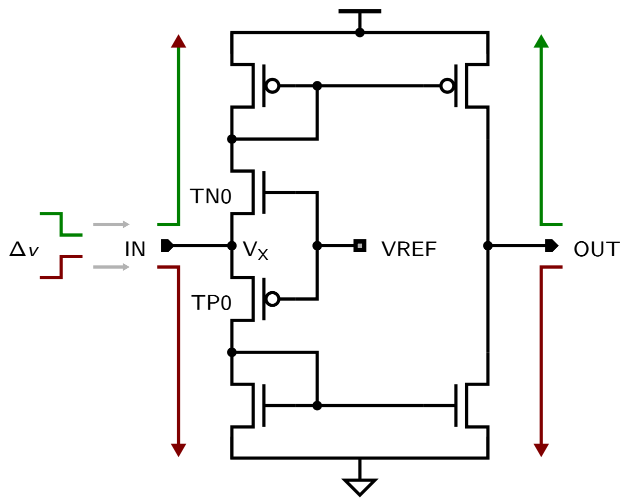


Figure A.5: Working principle of the bidirectional ICON. A set of NMOS and PMOS current mirrors is employed to generated in the output node an inverted and divided copy of the current in the input node. TP0 and TN1 are both in the subthreshold regime, the input has therefore a high impedance.

A.3.2 Stability of the Closed Programming Loop

This section presents the stability analysis for the N-Input current (reset) programming loop described in section 6.2.3.1, the schematic is shown in figure A.6. For a closed loop system, it is essential to make sure that the circuit is stable in the desired operating point. Furthermore it must be made sure that this operating point can be reached safely when powering the circuit. Simulation shows, that a DC operating point can be found for any bias current supplied by R_{bias}. For the operating point to be stable, negative DC feedback must be applied in the loop and the loop gain must have dropped to below unity before the phase of the loop has decreased by an additional 180°. This is intuitive because negative feedback plus a negative phase shift of 180° degrees essentially generates positive feedback at the input if the gain is larger than unity. Positive feedback creates a pile-up effect at the input causing oscillations. The so-called phase margin is defined as the difference of phase at the unity gain crossover frequency and 180°. The system is stable if the phase margin is positive, but it is susceptible to ringing, the closer the phase margin gets to zero. For a large phase margin the system settles very slowly. A phase margin of 60° is generally considered the optimum value [28], delivering negligible peaking while providing fast settling.

The filter input node contributes the most dominant pole of the system which is given by:

$$\omega_0 = \frac{1}{R_{out,ICON}(1 + A)C_f} \tag{A.310}$$

where A is the open loop gain of the filter amplifier, R is the total resistance at the negative input of the amplifier and C_f is the integrator (filter) feedback capacitance. R is given by the output resistance of the ICON cell. The secondary pole is given by:

$$\omega_1 = \frac{1}{r_{on,SwRes}C_{in} + c_{gd,N}g_{m_N}r_{ds,N}} \tag{A.311}$$

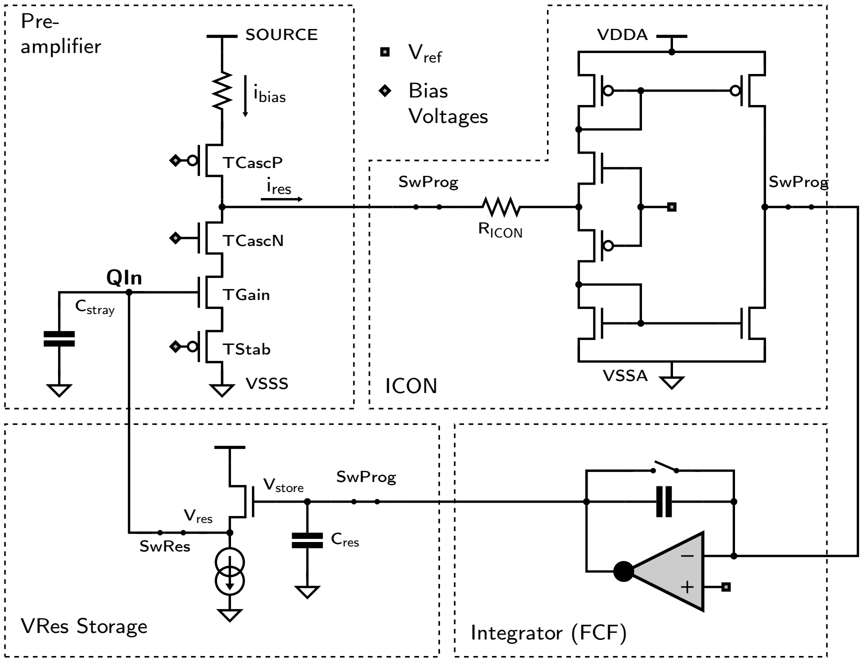


Figure A.6: Detailed view of the current programming loop. All SProg and the SRes switches are closed to establish the loop while SRes is pulsed to clear a collected signal from the input node.

where C_{in} is the total shunt capacitance at the input node ground and the and small signal parameters subscripted by N belong to the input NMOS transistor (TGain). Note that node v_x (see figure A.5) is not a virtual ground node in the programming phase. The lowest resistance at v_x (even without R_{ICON} is given by the channel resistance of TGain ($r_{ds,N}$). This situation configures the input branch as a voltage gain stage with a gain of $A_N = g_{m_N} r_{ds,N}$. The gate-drain capacitance of TGain is therefore amplified by A_N (Miller effect) making it non-negligible. To improve the stability, the most effective measure is to split the poles further apart in frequency. This can be achieved by

(1) Reducing ω_0 with respect to ω_1. This measure starts to dampen the gain at an earlier frequency, such that the gain crossover decreases in frequency increasing the phase margin.

(2) Increasing ω_1 with respect to ω_0. This measure is only effective if ω_1 is pushed out far enough that its phase shift occurs when the loop gain is already sufficiently low.

Due to the enourmous DC loop gain of $\approx 109\,dB$ the most effective means to generate additional pahse margin is to increase C_f and / or $R_{out,ICON}$. In Figure A.8, the phase margin is plotted against the C_f for different process corners. The maximum possible capacitance (due to area constraints) is 13 pF, for the (very extreme) fast-fast-functional corner ('fff') a further increase of $R_{out,ICON}$ is necessary to reach a phase margin of 60°.

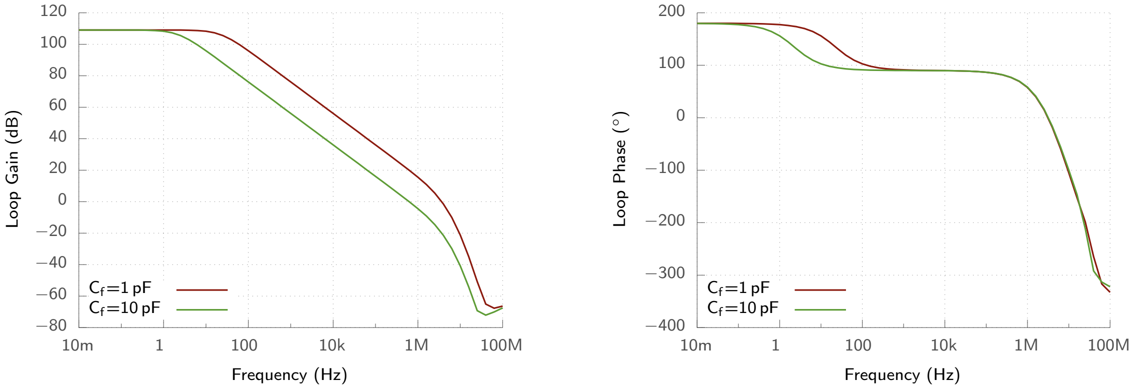


Figure A.7: Loop gain and phase of the closed programming loop. The phase margin can be incrased by increasing the amplifier feedback capacitance.

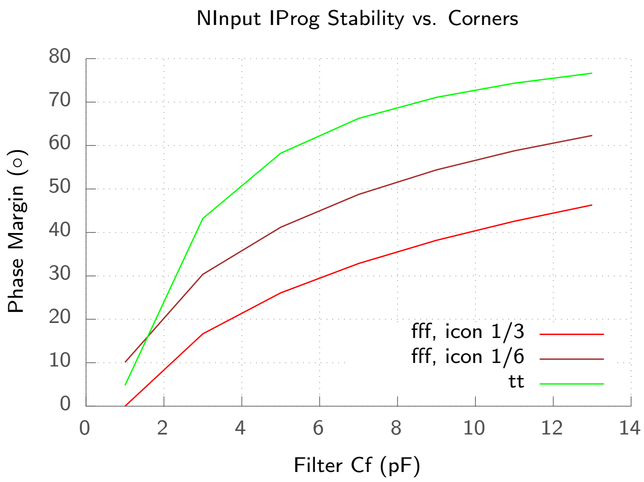


Figure A.8: Phase margin of the loop for all process corners versus integrator feedback capacitance C_f.

Bibliography

[1] Karsten Hansen. Private communication. DESY FEC, Hamburg.

[2] The European XFEL, Hamburg, Germany. http://www.xfel.eu/.

[3] The European X-Ray Free-Electron Laser technical design report. http://xfel.desy.de/technical_information/tdr/tdr/, 2007.

[4] C. Pellegrini. The history of X-ray free electron lasers. *The European Physical Journal H*, 37(5):659–708, Oct. 2012.

[5] Peter Schmüser, Martin Dohlus, and Jörg Rossbach. *Ultraviolet and Soft X-Ray Free-Electron Lasers: Introduction to Physical Principles, Experimental Results, Technological Challenges*. Springer Publishing Company, Incorporated, 1st edition, 2008.

[6] Linac Coherent Light Source. https://lcls.slac.stanford.edu/.

[7] Spring-8 Angstrom Compact Free Electron Laser. http://xfel.riken.jp/eng/index.html.

[8] C PellegrIni. A 4 to 0.1 nm FEL Based on the SLAC Linac. In *Proceedings Workshop on Fourth Generation Light Sources*, page 364, 1992.

[9] Free-electron laser FLASH, Hamburg, Germany. https://flash.desy.de/.

[10] Free-electron laser FLASH: How does it work? http://photon-science.desy.de/facilities/flash/the_free_electron_laser/how_it_works/high_gain_fel/index_eng.html.

[11] A. M. Kondratenko and E. L. Saldin. Generation of coherent radiation by a relativistic electron beam in an ondulator. *Particle Accelerators*, 10:207–216, 1980.

[12] Heinz Graafsma. Requirements for and development of 2 dimensional X-ray detectors for the European X-ray Free Electron Laser in Hamburg. *Journal of Instrumentation*, 4(12):P12011, 2009.

[13] A. Klyuev et al. AGIPD, a high dynamic range fast detector for the European XFEL. *Journal of Instrumentation*, 10(01):C01023, 2015.

[14] A Koch et al. Performance of an LPD prototype detector at MHz frame rates under Synchrotron and FEL radiation. *Journal of Instrumentation*, 8(11):C11001, 2013.

[15] B. Heisen et al. Karabo: An Integrated Software Framework Combining Control, Data Management, and Scientific Computing Tasks. In *14th International Conference on Accelerator & Large Experimental Physics Control Systems, ICALEPCS 2013*, Oct 2013.

[16] J Coughlan, C Day, S Galagedera, and R Halsall. The Train Builder data acquisition system for the European-XFEL. *Journal of Instrumentation*, 6(11):C11029, 2011.

[17] L. Rossi, P. Fischer, T. Rohe, and N. Wermes. *Pixel Detectors : From Fundamentals to Applications*. Springer, 2006.

[18] H. Spieler. *Semiconductor Detector Systems*. Oxford University Press, 2005.

[19] G. Lutz. *Semiconductor Radiation Detectors*. Springer, 2001.

[20] G. Knoll. *Radiation Detection and Measurement*. John Wiley & Sons, Inc., 2010.

[21] Antonio Cerdeira and Magali Estrada. Analytical Expressions for the Calculation of Pixel Detector Capacitances. *IEEE Transactions on Nuclear Science*, 44(1):63–65, February 1997.

[22] Emilio Gatti and Pavel Rehak. Semiconductor drift chamber - An application of a novel charge transport scheme. *Nuclear Instruments and Methods in Physics Research*, 225(3):608 – 614, 1984.

[23] M. Porro et al. Spectroscopic performances of DePMOS detector/amplifier device with respect to different filtering techniques and operating conditions. In *Nuclear Science Symposium Conference Record, 2004 IEEE*, volume 2, pages 724–728, Oct 2004.

[24] J. Kemmer and G. Lutz. New detector concepts. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 253(3):365 – 377, 1987.

[25] J. Kemmer et al. Experimental confirmation of a new semiconductor detector principle. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 288(1):92 – 98, 1990.

[26] E. Gatti, P.F. Manfredi, M. Sampietro, and V. Speziali. Suboptimal filtering of 1/f-noise in detector charge measurements. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 297(3):467 – 478, 1990.

[27] Yannis Tsividis. *Operation and Modeling of the MOS Transistor*. McGraw-Hill, Inc., New York, NY, USA, 1987.

[28] Behzad Razavi. *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, Inc., New York, NY, USA, 2001.

[29] F.S. Goulding. Pulse-shaping in low-noise nuclear amplifiers: A physical approach to noise analysis. *Nuclear Instruments and Methods*, 100(3):493 – 504, 1972.

[30] V. Radeka. Optimum Signal-Processing for Pulse-Amplitude Spectrometry in the Presence of High-Rate Effects and Noise. *IEEE Transactions on Nuclear Science*, 15(3):455–470, June 1968.

[31] E. Gatti and P. F. Manfredi. Processing the signals from solid-state detectors in elementary-particle physics. *La Rivista del Nuovo Cimento (1978-1999)*, 9(1):1–146, 1986.

[32] E. Gatti, M. Sampietro, and P.F. Manfredi. Optimum filters for detector charge measurements in presence of 1/f noise. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 287(3):513 – 520, 1990.

[33] Matteo Porro et al. In *Nuclear Science Symposium Conference Record, 2008. IEEE", title = "Large format X-ray imager with mega-frame readout capability for XFEL, based on the DEPFET active pixel sensor*, pages 1578–1586, Oct 2008.

[34] M. Porro et al. Expected performance of the DEPFET sensor with signal compression: A large format X-ray imager with mega-frame readout capability for the European XFEL. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624(2):509 – 519, 2010.

[35] M. Porro et al. Development of the DEPFET Sensor With Signal Compression: A Large Format X-Ray Imager With Mega-Frame Readout Capability for the European XFEL. *IEEE Transactions on Nuclear Science*, 59(6):3339–3351, Dec 2012.

[36] G. Lutz et al. DEPFET sensor with intrinsic signal compression developed for use at the XFEL free electron laser radiation source. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624(2):528 – 532, 2010.

[37] P. Lechner et al. DEPFET active pixel sensor with non-linear amplification. In *Nuclear Science Symposium Conference Record, 2011. IEEE*, pages 563–568, Oct 2011.

[38] Stefan Aschauer et al. Internal charge injection for the calibration of DEPFETs with non-linear amplification. In *Nuclear Science Symposium Conference Record, 2012. IEEE*, pages 475–481, 2012.

[39] K. Hansen, C. Reckleben, P. Kalavakuru, J. Szymanski, and I. Diehl. 8-bit 5-MS/s Analog-to-Digital Converter for Pixel-Level Integration. *IEEE Transactions on Nuclear Science*, 60(5):3843–3851, Oct 2013.

[40] David Moch et al. Calibration of the Non-Linear System Characteristic of a Prototype of the DSSC Detector for the European XFEL. In *Nuclear Science Symposium Conference Record, 2014. IEEE*, Nov 2014.

[41] K. Hansen, Helmut Klär, and Dirk Müntefering. Camera Head of the DSSC X-Ray Imager. In *Nuclear Science Symposium Conference Record, 2011 IEEE*, pages 668–672, Oct 2011.

[42] Manfred Kirchgessner. PhD thesis in progress at Heidelberg University.

[43] Matteo Porro et al. The Development of the DSSC Detector for the European XFEL: toward the First Ladder Camera. Nuclear Science Symposium Conference, N3-B3-3, 2015. IEEE, Nov 2015.

[44] F. Erdinger et al. The DSSC pixel readout ASIC with amplitude digitization and local storage for DEPFET sensor matrices at the European XFEL. In *Nuclear Science Symposium Conference Record, 2012. IEEE*, pages 591–596, Oct 2012.

[45] L. Bombelli, C. Fiorini, S. Facchinetti, M. Porro, and G. De Vita. A fast current readout strategy for the XFEL DePFET detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624(2):360 – 366, 2010.

[46] S. Facchinetti, L. Bombelli, C. Fiorini, M. Porro, G. De Vita, and F. Erdinger. Characterization of the Flip Capacitor Filter for the XFEL-DSSC Project. *IEEE Transactions on Nuclear Science*, 58(4):2032–2038, Aug 2011.

[47] C. Fiorini et al. A Simple Technique for Signal Compression in High Dynamic Range, High Speed X-ray Pixel Detectors. *IEEE Transactions on Nuclear Science*, 61(5):2595–2600, Oct 2014.

[48] K. Hansen, C. Reckleben, I. Diehl, M. Bach, and P. Kalavakuru. Pixel-level 8-bit 5-MS/s Wilkinson-type digitizer for the DSSC X-ray imager: Concept study. *Nuclear Instruments and Methods in Physics Research Section A*, 629(1):269 – 276, 2011.

[49] C. Reckleben, K. Hansen, P. Kalavakuru, and I. Diehl. 8-bit 5-MS/s per-pixel ADC in an 8-by-8 Matrix. In *Nuclear Science Symposium Conference Record, 2011 IEEE*, pages 1713–1717, Oct 2011.

[50] Florian Erdinger. Design of High Density Memories for the DSSC Pixel Detector at XFEL. Diploma Thesis, Mannheim University, 2009.

[51] F. Erdinger and Peter Fischer. Compact Digital Memory Blocks for the DSSC Pixel Readout ASIC. In *Nuclear Science Symposium Conference Record, 2010. IEEE*, pages 1364–1367, Oct 2010.

[52] Jan M. Rabaey. *Digital Integrated Circuits: A Design Perspective*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[53] E. Quartieri and M. Manghisoni. High precision injection circuit for in-pixel calibration of a large sensor matrix. In *7th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME), 2011*, pages 73–76, July 2011.

[54] J. Bhasker and Rakesh Chadha. *Static Timing Analysis for Nanometer Designs: A Practical Approach*. Springer Publishing Company, Incorporated, 1st edition, 2009.

[55] Jan Soldat. PhD thesis in progress at Heidelberg University.

[56] Rene Brun and Fons Rademakers. ROOT - An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1):81 – 86, 1997.

[57] Christian Reckleben et al. A 64-by-64 Pixel-ADC Matrix. Nuclear Science Symposium Conference, N2AP-60, 2015. IEEE, Nov 2015.

Acknowledgements

First of all I want thank my wife Denise for sharing her life with mine and especially for being so considerate and supportive during submission times. In between the numerous chip submissions for the DSSC project, two babies were born who are all the world to us. I thank my children for being who they are, making me laugh and bringing joy to my life. I thank my parents for raising me to the person I am and helping me to develop a researching character.

I thank my supervisor, Prof. Dr. Peter Fischer, for giving me the opportunity to work on this exciting project and giving me the chance to be creative and let ideas develop. He has always been appreciative of my work and it is a pleasure to work in his group. I want to further thank Prof. Dr. Ivan Peric, KIT, for his support and advices. I am especially grateful that I had the opportunity to present my work at a number of international conferences. All the people at the Chair for Circuit Design and Simulation I want to thank for the friendly and productive working environment.

Working in the DSSC collaboration has been a great pleasure both from the scientific and human perspective. The exchange and discussions especially with the collaborating groups from DESY, led by Dr. Karsten Hansen and from Politecnico di Milano, led by Prof. Dr. Carlo Fiorini have always been fruitful.

I also thank all proof readers and of course the referees for reporting on my thesis. I know I am scratching only the surface here on the people I need to thank, it is out of scope to thank everyone here explicitly. I sincerely hope that nobody feels forgotten.