Analysis of pixel systematics and space point reconstruction with DEPFET PXD5 matrices using high energy beam test data

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Lars Reuen
aus
Nettetal

Bonn 2011

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Prof. Dr. N. Wermes
2. Gutachter:	Prof. Dr. K. Brinkmann

Tag der Promotion: 21.2.2011

Erscheinungsjahr: 2011

Contents

Introduction

Over the past century our understanding of the foundations of space, time, and the building blocks of energy and matter have been revolutionized by the discoveries made in modern physics. Among them are the advances in the field of particle physics which saw in the last three decades the emergence of what is now known as the *standard model* of particle physics (Fig. 1). This model describes three of the four known fundamental forces: the strong force that binds the atomic nuclei and their constituents, protons and neutrons, together, the electromagnetic force that governs the field of chemistry, and weak force which is responsible for the radioactive beta decay. These forces are mediated between certain particles, the *fermions*, via particles called *bosons*[1]. However, not all particles interact with every force, *leptons* for example do not interact via the strong force. In the standard model the weak with the electromagnetic force are unified to the electroweak force[2] which was experimentally confirmed by the discovery of the neutral current (1973) and the W and Z boson (1983).

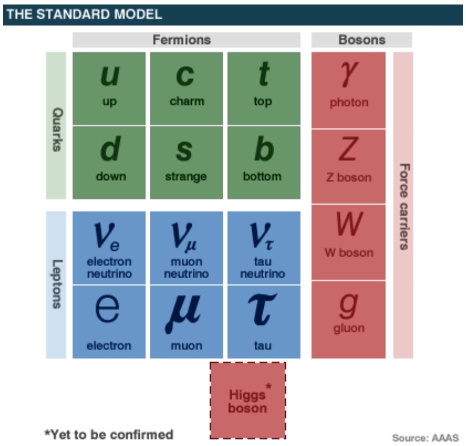


Figure 1: *The constituents of the standard model of particle physics.*

Despite its large success the standard model there are some remaining questions. For example, how do particles acquire their mass? In the standard model the Higgs mechanism explains the generation of the masses. It also predicts the existence of the Higgs particle which has yet to be discovered. Another issue is the so-called hierarchy problem: Why is the weak force 10^{32} times stronger than gravity. On the other hand the ΛCDM[3] model, which is in a certain sense the equivalent of the particle physics standard model in the field of modern cosmology, states that the universe consists to about 74% of so-called dark energy and to $\sim 22\%$ of dark mater and only to $\sim 4\%$ of baryonic matter. The nature of dark matter and dark energy is unknown and the standard model lacks an explanation. New theories that tackle these problems predict the existence of new physics like super symmetric particles or extra dimensions.

---

[1] Fermions have spin $\frac{1}{2}$ and bosons spin 1.

[2] Abdus Salam, Sheldon Glashow and Steven Weinberg received the nobel price in 1979 for their ground breaking theory, Carlo Rubbia and Simon van der Meer received the nobel price in 1983 for their experimental work that led to the discovery of the W and Z bosons

[3] Λ-**C**old **D**ark **M**atter, Λ is the cosmological constant

To address these questions a new generation of particle accelerators is on its way. The most noted is the Large Hadron Collider (LHC) at CERN[4], that is operational since 2009 and is expected to soon run at the targeted 14 TeV energy. To accomplish its goals the new generation of high energy collider experiments needs high precision particle detectors. In case of the LHC the two prominent, general purpose detectors[5] are the ATLAS[6] and the CMS[7] detectors. The innermost central part of these detectors is called vertex detector, which in case of ATLAS and CMS is a silicon pixel detector. The precision with which it measures a particle's position is a key quantity as the identification of a particles decay products depends on this figure of merit. The proper identifications of these decay products is vital for the validation of theories and for the discovery of new physics.

An accelerator experiment complementary to the LHC experiments would be the planned International Linear Collider (ILC), which is an electron positron collider envisioned to work with a center of mass energy of up to 1 TeV and which would provide high precision measurements accompanying the findings of the LHC. The vertex detector of the ILC must fulfill ambitious specifications, including a spatial resolution of better than $5\,\mu$m while contributing with not more than 0.1% of a radiation length per layer to the material budget. These demands have driven the development of new detector technologies like the DEPFET pixel and are central to this thesis. The DEPFET (**DEP**leted **F**ield **E**ffect **T**ransistor) is an active pixel semiconductor detector that integrates a first electronic amplification stage into the sensor material allowing for excellent signal to noise measurements. The first chapter explains the basic principles of a DEPFET semiconductor detector. It also comprises a short review of the underlying physical processes of particle detection with an emphasis on thin detectors.

Based on a DEPFET pixel matrix a vertex detector concept for the ILC has been put forward and a prototype system with a 64x128 pixel matrix has been developed. The DEPFET detector concept and the prototype system are the topic of chapter two. For a complete understanding and evaluation of a detector laboratory measurements of its performance alone are not sufficient. The detector has to be put in a more realistic test environment in the form of a beam test experiment where the detector's response to high energetic particle is measured. The results of such a beam test with the DEPFET prototype system including in-pixel homogeneity measurements will be presented in chapter three.

The quest for higher precision does not only drive the hardware side but also the software side of modern high energy particle experiments. New tools like multivariate analyses are becoming common place in the analysis of particle physics data. However, the position reconstruction method for pixel detectors at large is still the η algorithm, a technique that was originally developed for strip detectors. Therefore another focus of this thesis is the comparative study of new position reconstruction algorithms. These studies will be presented in chapter four. This thesis is completed with a summary of the results given in the last chapter.

---

[4] European Organization for Nuclear Research, originally Conseil Européen pour la Recherche Nucléaire

[5] The other two detector at the LHC, ALICE and LHCb, are more specialized

[6] ATLAS = **A** **T**oroidal **L**HC **A**pparatu**S**

[7] **C**ompact **M**uon **S**olenoid

1

Particle detection with silicon sensors

For any particle physics experiment it is important to understand how the detector interacts with the particles, what kind of physical processes are involved, and what kind of signature to expect inside the detector as a result of such interactions. This chapter will address these issues with respect to the test beam experiment and the position reconstruction studies examined in later chapters of this thesis. The first part will describe the interaction of charged particles in matter: how much energy is deposited in a silicon detector with a given impinging particle and what is the nature of secondary particles, so called *knock-on* or *δ-electrons*. The second part covers the fundamentals of semiconductor particle sensors and the basic working principle of a DEPFET. Throughout this chapter references will be made to the actual experiment, e.g. the sizes of a sensor pixel or the energy of the pions in the test beam:
The beam test was undertaken at CERN with 120 GeV pions. The DEPFET sensor type used for the experiment is subdivided in pixels with different dimensions in X and Y. The X axis of a pixel is 32 μm wide and the Y axis is 24 μm wide. For illustrative purposes the dimensions of three pixels in each axis were added to some figures in this chapter.

1.1 Passage of particles through matter

There are several means by which a charged particle can lose energy in matter, though they can be broadly divided into the collision and the radiation regime. Collision processes are important because for moderately relativistic charged particles heavier than electrons these are the dominant processes responsible for energy loss (section 1.1.1). For electrons on the other hand radiation effects play a much bigger role even at moderate relativistic energies. However, the electrons encountered in the scope of this thesis have energies low enough that collision is the dominant process (section 1.1.2). Figure 1.1 shows the energy loss of muons in copper as a function of the particles energy over several orders of magnitude. As is indicated in the figure the collision loss regime described by the Bethe equation 1.3 is only valid between energies of $\beta\gamma \approx 0.05$ and $\beta\gamma \approx 500$ in case of μ^+ passing through copper. At higher energies, radiative effects begin to be important, while at lower energies the velocity of the particle approaches that of the shell electrons and therefore certain approximations are not valid anymore.

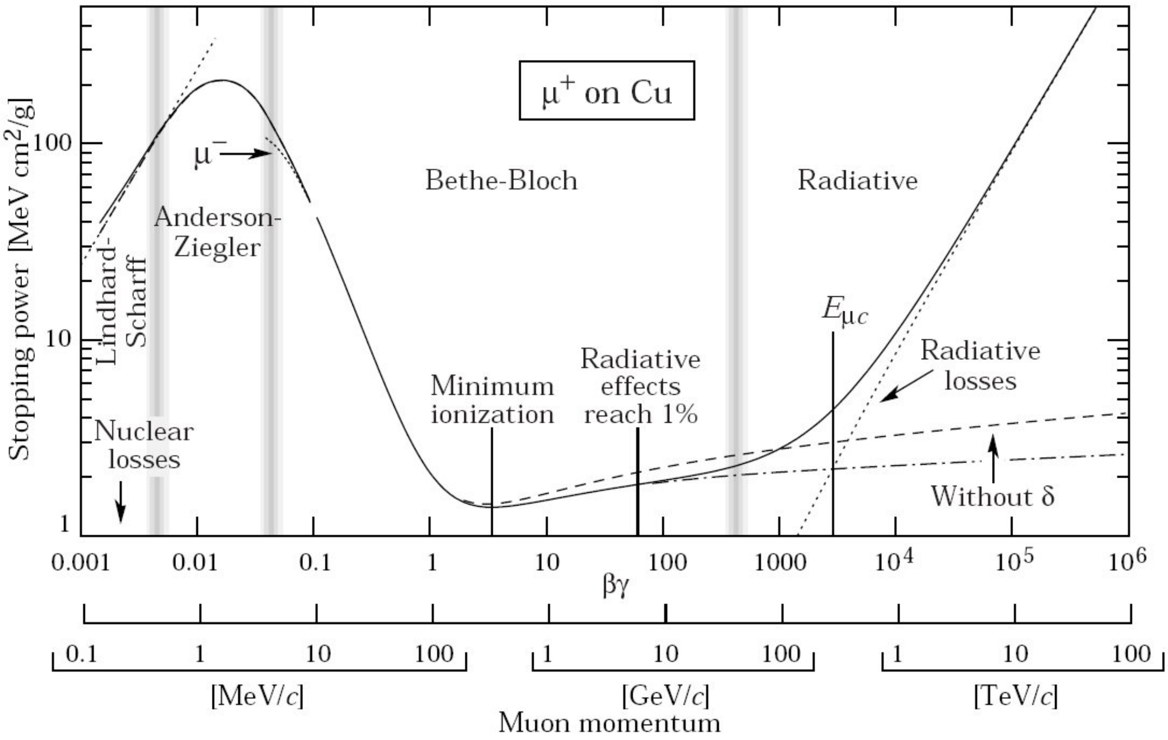


Figure 1.1: *The stopping power $-dE/dx$ for μ^+ particles in copper as a function of $\beta\gamma$ over several decades of energy taken from [1]. The plot is divided in different regimes where the energy loss is dominated by different processes.*

1.1.1 Energy loss of heavy charged particles in matter

For moderately relativistic charged particles heavier than electrons the main energy loss is due to collisions with atoms. The maximum amount of energy T_{max} that can be transferred from an incident particle with rest mass m_0 and a momentum of

$$p = mv = \beta\gamma m_0 \tag{1.1}$$

to an electron (rest mass $m_e = 0.511\,\text{MeV}$) in one collision is [1]

$$T_{max} = \frac{2m_e\beta^2\gamma^2}{1 + 2\gamma\frac{m_e}{m_0} + \left(\frac{m_e}{m_0}\right)^2} \tag{1.2}$$

with the speed of light set to one ($c = 1$) through out this work. The test beam data used in this work was gathered using 120 GeV pions with a rest mass of $m_0^\pi = 139.5\,\text{MeV}$ which yields a $\beta\gamma = 860$ and a maximum energy transfer of $T_{max} \simeq 103.5\,\text{GeV}$. It should be noted that the approximation $T_{max} = 2m_e\beta^2\gamma^2$ sometimes found in older literature is only valid if $\gamma \ll (m_0/m_e)$ and therefore does not apply here. While T_{max} is an upper limit for the energy transfer the lower limit arises from the interaction time between the particle with a speed of β and the shell electron with a binding energy of $\hbar\omega$ and rotation period of $\tau = 1/\omega$. This is described by a material specific ionization constant I. A thorough quantum-mechanical calculation first performed by Bethe [2] states the mean energy loss (*Bethe equation*) as

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\rho\frac{1}{\beta^2}\left[\frac{1}{2}\ln\left(\frac{2m_e\beta^2\gamma^2T_{max}}{I^2}\right) - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right] \tag{1.3}$$

with constants and variables defined as in table 1.2. $\delta(\beta\gamma)$ is the so-called *density correction.*

Density correction

At high energies the particles field extents and as a result of the distant collisions the mean energy loss rises with $\ln(\beta\gamma)$. However, the sensor material also becomes polarized and thus the field extension is limited effectively truncating the logarithmic rise [1]. There are several ways these corrections can be parameterized. For figure 1.2 this parametrization was used [3]:

$$\delta(X) = \begin{cases} 0 & X < X_0 \\ 4.6025X + C + a(X_1 - X)^m & X_0 < X < X_1 \\ 4.6025X + C & X > X_1 \end{cases} \tag{1.4}$$

where $X = \log_{10}(\beta\gamma)$. The constants in this equations are listed in table 1.1. For 120 GeV pions the high energy limit applies, $X(\beta\gamma = 860) = 2.93 < X_1 = 2.87$. At these energies the density correction becomes

$$\delta/2 \to \ln(\hbar\omega_p) + \ln(\beta\gamma) - 1/2 \tag{1.5}$$

and the stopping power growth as $\ln(\beta\gamma)$ instead of $\ln(\beta^2\gamma^2)$. $\hbar\omega_p$ is the plasma energy, which scales as the square root of the electron density [1], hence the name density correction. The remaining relativistic rise is due to the growth of T_{max} which is caused by very rare, large energy transfers to a few electrons (see equation 1.8).

X_0	X_1	m	a	C
0.2014	2.87	3.25	0.1492	4.44

Table 1.1: *Density correction constants for silicon taken from [3]. Note that C can be calculated with $C = \ln(I/\hbar\omega_p)$. These constants are used in equation 1.4 and for figure 1.2.*

For a sufficiently thin detector the effect of the energy loss on the energy of the particle is negligible and the particle's energy can be considered constant. The mean energy loss is then simply $E_{mean} = dE/dx \cdot x$ or

$$E_{mean} = \frac{\xi}{\beta^2}\left[\ln\left(\frac{2m_e\beta^2\gamma^2T_{max}}{I^2}\right) - 2\beta^2 - \delta(\beta\gamma)\right] \tag{1.6}$$

where ξ is a sensor specific constant that will be used frequently in this chapter:

$$\xi = \frac{K}{2}\rho\frac{Z}{A}x \tag{1.7}$$

with x the absorber thickness. Figure 1.2 shows the evolution of the energy loss as a function of pion momentum in $450\,\mu$m silicon with and without density corrections. For low kinetic energies ($\beta\gamma \lesssim 1$) equation 1.3 scales with $dE/dx \approx \beta^{-2}$ until a minimum is reached at $\beta\gamma \approx 3.5$ while for high energies $\beta\gamma > 4$ the logarithmic rise dominates. The position of the minimum of dE/dx is at $\beta\gamma \approx 3$ in good accuracy independent of the particle as can be seen in figure 1.3. The energy loss at this minimum is roughly the same for all materials in this minimum $dE/dx \approx 2\,\mathrm{MeV\,cm^2/g}$ ($dE/dx(Si) = 1.66\,\mathrm{MeV\,cm^2/g}$). Relativistic particles with an energy loss corresponding to this minimum are called "*minimum-ionizing particles*" (MIPs) and their dE/dx is often used as an approximation for particles with higher energies since there is only a (density correction dampened) logarithmic rise in dE/dx.

Restricted energy loss

Sofar no distinction has been made between the energy *loss* of the particle and the actual energy *deposition* inside the sensor. As we shall see later this distinction is important in presents of so called δ-electrons. Although up to T_{max} energy can be transferred form the pion to an electron, but this electron might leave the sensor before depositing all of T_{max}. Therefore one can restrict the maximum energy transfer for *deposition* to T_{cut} and gets [1, 4]

$$E_{mean} = \frac{\xi}{\beta^2}\left[\ln\left(\frac{2m_e\beta^2\gamma^2T_{cut}}{I^2}\right) - 2\beta^2\left(1 + \frac{T_{cut}}{T_{max}}\right) - \delta(\beta\gamma)\right] \tag{1.8}$$

For $T_{cut} \to T_{max}$ this approaches the Bethe equation 1.6. For high energies the energy loss approaches the constant "Fermi plateau", since the $\ln(\beta^2\gamma^2)$ rise in T_{max} is replaced by a constant and the remaining $\ln(\beta^2\gamma^2) = 2\ln(\beta\gamma)$ is canceled out by the density correction. The energy loss in a $450\,\mu$m thick silicon sensor, restricted to $T_{cut} = 2\,\mathrm{MeV}$, is shown in figure 1.2. Its shape in relation to the full Bethe loss is in agreement with curves shown in [4, p.46].

Energy loss fluctuations in thin absorbers

So far only the mean energy loss has been covered. However, the energy loss process in a sensor is of statistical nature. This results in large fluctuations of the energy loss which becomes more crucial in thinner sensors. One important number to characterize how much a sensor is affected by these fluctuations with regard to its thickness is the ratio of the mean energy loss E_{mean} to the maximum energy transfer T_{max}:

$$\kappa = \frac{E_{mean}}{T_{max}} \tag{1.9}$$

The smaller the κ the stronger the fluctuations are, with $\kappa \gtrsim 1$ already approaching an energy distribution that is Gaussian, whereas $\kappa < 0.01$ means that the Landau theory

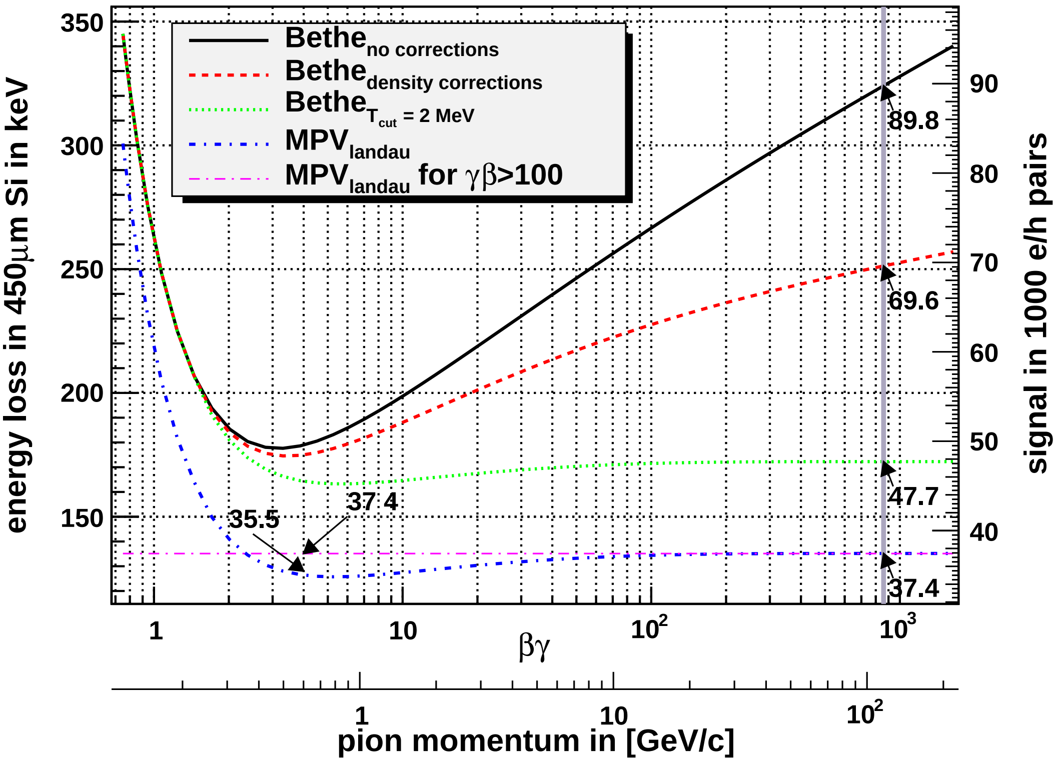


Figure 1.2: The energy loss or, respectively, energy deposition in 450 μm silicon (left axis in keV) and the equivalent number of electron/hole pairs in 1000 as a function of $\beta\gamma$ (and pion momentum). From top to bottom: The first two curves are the Bethe collsion loss with and without density effect corrections (equations 1.6 and 1.4). Below is the restricted energy loss for a upper cut on the deposited energy of $T_{cut} = 2\,MeV$ (equation 1.8) and finally the lowest two are the most probably dE/dx value (equation 1.12) as a function of the pion energy and in the ultra relativistic limit (flat line, equation 1.13). The vertical line at $\beta\gamma = 860$ indicates the energy of the 120 GeV pion used in the DEPFET test beam. The numbers next to the arrows indicate the number of electron/hole pairs in 1000 equivalent to the energy loss for the various curves at $\beta\gamma = 860$. The two other arrows at $\gamma\beta = 4$ show the computed values of the most probable value at this $\gamma\beta$ and in the high energy limit. The difference is $\approx 5\%$.

applies to the energy loss [5, 6, 3]. The latter is also true for 120 GeV pions with $\kappa \ll 10^{-5}$. In this case the energy spectrum can be parameterized by the highly asymmetric Landau distribution which can be a approximated with [7]

$$L(\lambda) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\lambda + e^{-\lambda}\right)\right] \tag{1.10}$$

with λ the deviation from the most probable energy loss:

$$\lambda = \frac{E - E_{MPV}}{\xi} \tag{1.11}$$

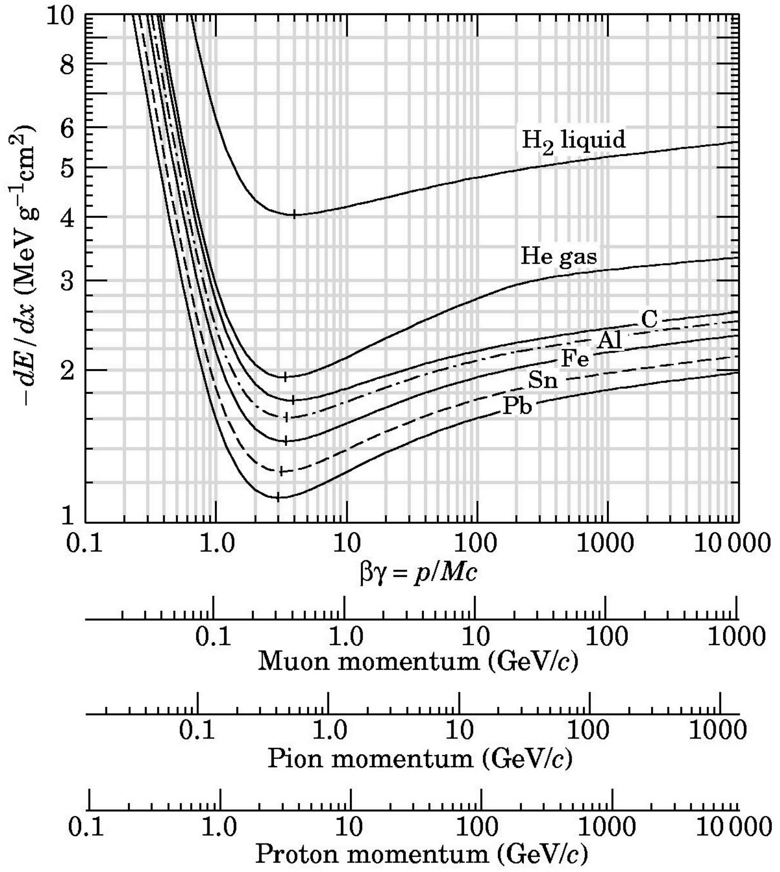


Figure 1.3: *The mean energy loss rate in the Bethe regime for various materials and (heavy) particles. Radiation effects are not considered. (They become significant at $\beta\gamma \gtrsim 1000$.) The minimum of the energy loss rate is virtually independent of the material (when normalized by the materials density) and the particle and roughly at $\beta\gamma = p/mc \approx 3$. [1].*

and E the actual energy loss, ξ defined as in equation 1.7 and E_{MPV} the most probable value of the energy loss [1]:

$$E_{MPV} = \xi \left[\ln \left(\frac{2m_e\beta^2\gamma^2}{I^2}\xi \right) + 0.2 - \beta^2 - \delta(\gamma\beta) \right] \tag{1.12}$$

Since the Landau distribution depicted in figure 1.4 is highly asymmetric with a long tail towards higher energies the mean value, calculated with the Bethe equation (1.3 or 1.8 for restricted energy loss), is not an easy quantity to measure but the peak of the Landau distribution is. This is the most probable energy loss, shortly E_{MPV} for most probable value. The most probable value follows closely the restricted energy loss in its dependence on the impinging particles energy as can be seen in figure 1.2. At high energies the density effect correction $\delta(\gamma\beta)$ (equation 1.6) influences the equation such that

$$E_{MPV} = \xi \left[\ln \left(\frac{2m_e\xi}{(\hbar\omega_p)^2} \right) + 0.2 \right] \tag{1.13}$$

and the most probable value approaches a plateau. For the DEPFET test beam with $x = 450\,\mu\mathrm{m}$ ξ and E_{MPV} are:

$$\xi = 178\,\frac{\mathrm{keV}}{\mathrm{cm}} \cdot x \simeq 8\,\mathrm{keV} \tag{1.14}$$

$$E_{MPV} = 8\,\mathrm{keV} \cdot 15.95 \approx 129\,\mathrm{keV} \tag{1.15}$$

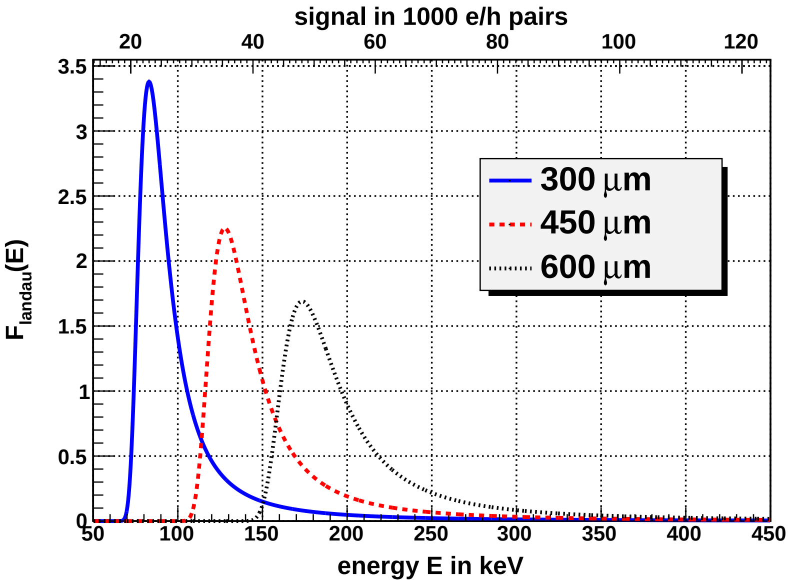


Figure 1.4: *The Landau distribution for three different silicon sensor thicknesses. The distributions were calculated using the ROOT software package from CERN[8]. The values for E_{MPV} and ξ were calculated using the high relativistic limit (equations 1.13 and 1.6 with $\beta = 1$). The bottom axis shows the energy in keV, the upper axis the corresponding of number electron hole pairs in 1000.*

1.1.2 Delta electrons

Until now the treatment of energy loss of a particle when passing through matter has been focused on particles heavier than electrons ($m \gg m_e$). There are two reasons why electrons need to be treated differently:

- The emission probability for bremsstrahlung is proportional to the inverse square of the particles mass. Therefore radiative effects have a much earlier onset in terms of the particles energy for an electron than for heavy particles like pions.

- In the treatment of the collision loss it was assumed that the mass of the particle is large compared to the shell electron mass. The particles are now identical (ignoring the case of positrons).

Atomic number	Z	14 [1]	
Atomic weight	A	28.09 [1]	$\mathrm{g\,mol^{-1}}$
Density	ρ	2.336 [1]	$\mathrm{g/cm^3}$
Mean excitation energy	I	173 [1]	eV
	$K = 4\pi N_A r_e^2 m_e c^2$	0.307 [2]	$\mathrm{MeV\,mol^{-1}\,cm}$
Radiation length	X_0	21.82 [1]	$\mathrm{g\,cm^{-2}}$
plasma energy	$\hbar\omega_p$	31.05 [1]	eV
Critical energy e^- in silicon	E_c	40.19 [1]	MeV
Critical energy μ^- in silicon	E_c^μ	582 [1]	GeV
MIP stopping power in silicon	$dE/dx(MIP)$	3.876 [1]	$\mathrm{MeV\,cm^{-1}}$
pion rest mass	m_0^π	139.5[1]	MeV
electron rest mass	m_e	0.511 [1]	MeV
$\gamma\beta$ for 120 GeV π	$\gamma\beta = p/m_0$	860	
max. energy transfer per collision	T_{max}	103.5	GeV (eqn. 1.2)
Number of δ	F_δ	76.5	$\mathrm{keV\,g^{-1}}$

Table 1.2: *Constants and variables used in this chapter. If not mentioned otherwise material properties reference to silicon and particle properties to a 120 GeV pion. (Sources: [1][1], [2][3]).*

The collision energy loss is of statistical nature and the momentum transfer to shell electrons can sometimes be large enough to give them enough momentum to be treated like an independent particle. These electrons are called δ- or *knock-on electrons*. This section will give a brief overview over both topics: the energy loss of electrons in matter and the characteristics of δ electrons.

The critical energy E_c:

There are two competing energy loss mechanisms, radiative and collision. Their importance for energy loss varies with energy. Radiative losses dominating the high energy regime and collision losses at lower energies. The *critical energy* E_c describes at which energy radiative losses become equal to ionization losses:

$$-\frac{dE}{dx}(E_c)\bigg|_{ionization} = -\frac{dE}{dx}(E_c)\bigg|_{bremsstrahlung} \tag{1.16}$$

There are several parameterizations for E_c, among them [9]

$$E_C = \frac{800}{Z+1.2}[\mathrm{MeV}] \tag{1.17}$$

and more accurate [4]

$$E_C = 2.66\left(\frac{ZX_0^g}{A}\right)^{1.11}[\mathrm{MeV}] \tag{1.18}$$

where X_0^g is the radiation length normalized by material density. For silicon equation (1.18) yields $E_C = 37.6\,\mathrm{MeV}$, other values found in the literature are $E_C = 39\,\mathrm{MeV}$ [3] or $E_C = 40.19\,\mathrm{MeV}$ [1]. The critical energy can also be defined for other particles but is much larger since the radiative losses are proportional to m^{-2}. A rough estimate for muons in silicon would be

$$E_c^{\mu}(Si) \approx E_c^{e}(Si) \left(\frac{m_\mu}{m_e}\right)^2 \approx 1.7\,\mathrm{TeV}, \tag{1.19}$$

however, the critical energy for muons given in [1] is $E_C^{\mu} = 582\,\mathrm{GeV}$. As this number is still considerably larger than the 120 GeV of the pions used in the test beam experiment and the pion mass itself is larger than the muon mass by factor of $(m_\pi/m_\mu)^2 \simeq 1.6$ any radiative losses for the primary particle can be neglected.

Collision energy loss of electrons

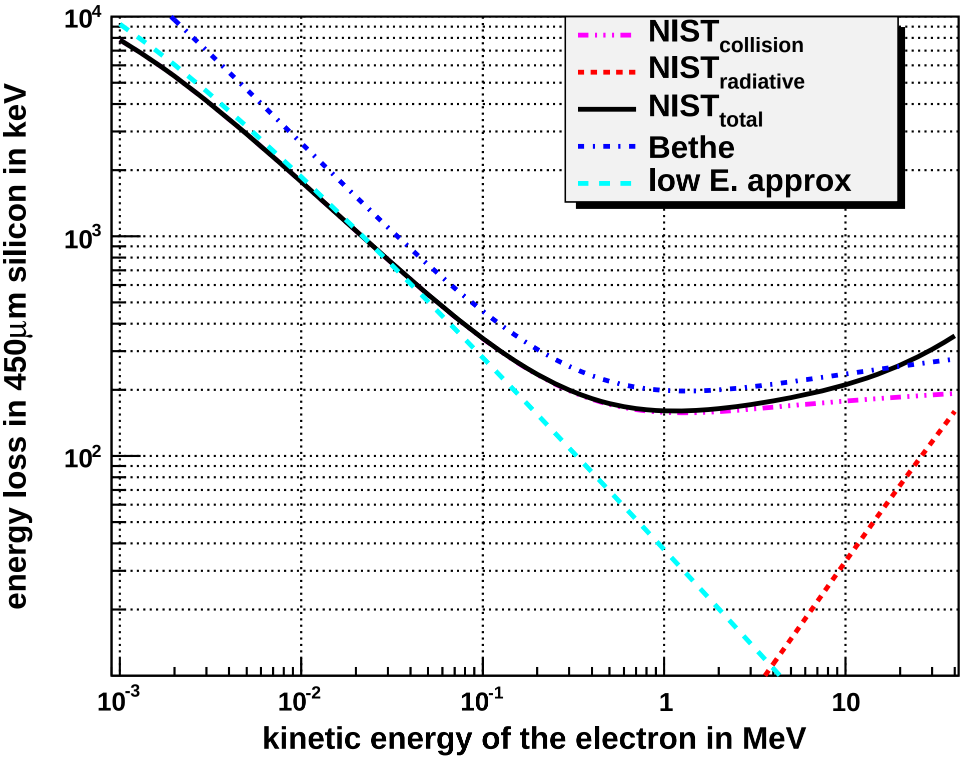


Figure 1.5: *Energy loss of electrons in* $450\,\mu m$ *silicon in keV as a function of the electrons kinetic energy from* $1\,keV$ *to* $E_c \simeq 40\,MeV$ *in a double logarithmic plot. Three of the curves were generated by the NIST online service and show the losses due to radiative and the collision process separated as well as the total energy loss. Furthermore equation 1.20 (Bethe/collision loss for electrons) and equation 1.22 (low energy approximation) are plotted as well. Below 10 MeV the radiative loss becomes negligible small.*

As will be shown later the majority of δ electrons are well below the critical energy and therefore the electron energy loss can be limited to Coulomb interaction. The result is similar to the Bethe equation [4]

$$-\frac{dE}{dx} = \frac{K}{2}\frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{m\beta^2\gamma^2 T_{max}}{I^2}\right) - \beta^2 + f(\gamma) - \delta(\beta\gamma)\right] \tag{1.20}$$

with

$$f(\gamma) = \left(\frac{2\gamma-1}{\gamma^2}\right)\ln 2 - \frac{1}{8}\left(\frac{\gamma-1}{\gamma}\right)^2 \tag{1.21}$$

Equation 1.20 is plotted in figure 1.5 together with data for collision, radiative, and total energy loss generated by the online service of the National Institute of Standards and Technology (NIST). For non-relativistic energies $I \ll\ E \ll m_0$ the drop of the stopping power can be approximated with [10]

$$\frac{dE}{dx} = 7.8 \cdot 10^{10}\frac{Z}{A}\frac{1}{E}\ln\left(1.166\frac{E}{I}\right) \tag{1.22}$$

with E and I in eV and dE/dx in eV g^{-1}/,cm^2. Indeed for kinetic energies below 100 keV equation 1.22 is a sufficient approximation as can be seen in figure 1.5. The general shape of the energy loss shows a minimum just like the energy loss for heavy particles, however for higher energies the energy loss is dominated by radiative process (red line in fig. 1.5) and not by the logarithmic rise of equation 1.20.

The δ electrons energy spectrum and angular distribution

When a charged, heavy particle like a pion passes through matter it transfers energy in a large number of collisions. The differential probability for a certain energy to be transferred from a traversing particle with a speed of β, i.e. to create a free electron with energy T is described by [1]

$$\frac{d^2N}{dxdT} = \frac{1}{2}Kz^2\frac{Z}{A}\frac{1}{\beta^2}\frac{F(T)}{T^2} \tag{1.23}$$

where F describes a spin dependent factor [4]:

$$F(T_\delta) = \begin{cases} 1 - \frac{\beta^2 T_\delta}{T_{max}} & \text{for spin-0 particles,} \\ 1 - \frac{\beta^2 T_\delta}{T_{max}} + \frac{T_\delta^2}{2E^2} & \text{for spin-}\frac{1}{2}\text{ particles} \end{cases} \tag{1.24}$$

For all energies much smaller than the maximum energy transfer possible $T_{\delta e} \ll T_{max}$ the spin factor F can be neglected and since the δ-electron energies are $T_\delta \ll 100\,\mathrm{GeV}$ equation 1.23 becomes $\frac{d^2N}{dxdT} \propto T^{-2}$. Furthermore the differential collision probability given by equation 1.23 is only valid if the energy transfer is much larger than the ionization constant I of the material, $T_{\delta e^-} \gg I = 173\,\mathrm{eV}$. Assuming a thin detector limit, that is that the energy of the primary particle is considered constant (analog to the Landau distribution equation 1.14), one can integrate over the number of electrons per path length x. With ξ given by equation 1.6 one obtains for the average number δ electrons per energy (with $T_{\delta e} \ll T_{max}$ and $\beta = 1$ for 120 GeV pions):

$$\frac{dN}{dT} = \frac{\xi}{T^2}. \tag{1.25}$$

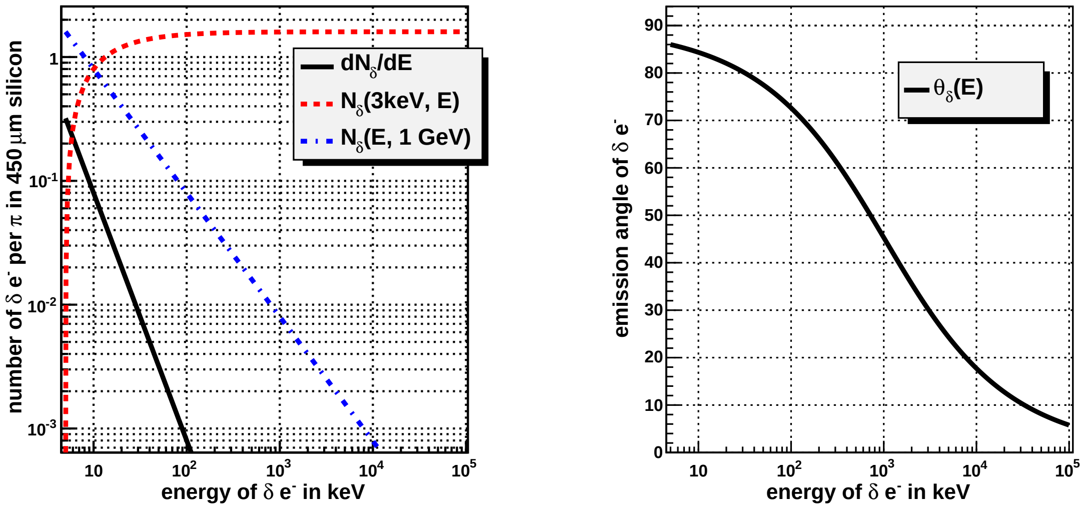


Figure 1.6: ***Left plot:*** *The differential collision probability for a pion in 450 μm silicon as a function of the transferred energy E from $E = 3\, keV$ to $E = 100\, MeV$ in a double logarithmic scale. Furthermore in red the total number of δ electron in a range from $3\, keV$ to E, which is basically the amount of kicked out electrons up to this energy. For a silicon sensor with $450\, \mu m$ most of the δ electrons are below 100 keV. The blue line shows the number of missing electrons, i.e. the number of δ electron in a range from E to $1\, GeV$.* ***Right plot:*** *The emission angle θ_δ of a δ electron with respect to the primary particle track.*

Finally the average number of δ electrons within a given energy range T_{low} to T_{high} is (again assuming $T_{low} \gg I$ and $T_{high} \ll T_{max}$):

$$\frac{dN}{dT} = \xi \left[\frac{1}{T_{low}} - \frac{1}{T_{high}} \right] . \tag{1.26}$$

From equations 1.25 and 1.26 it becomes evident that high energetic δ electrons are rare and that the majority of the collision processes are small scale energy transfers. With $\xi \simeq 8\,\text{keV}$ the probability of a δ electron with an energy higher than 100 keV is $\xi/E = 8\%$ or in other words 92% of all δ electrons are blow 100 keV and 50% are below 16 keV. This is illustrated in figure 1.6.

Due to the kinematics involved there is a fixed relationship between energy and emission angle of a δ electron with kinetic energy T_e [1]

$$\cos \Theta = \frac{T_e}{p_e} \frac{p_{max}}{T_{max}} \simeq \frac{T_e}{p_e} \tag{1.27}$$

and with $p = \sqrt{T_{kin}(T_{kin} + 2m)}$

$$\theta = \arccos \sqrt{\frac{T_\delta}{T_\delta + 2m_e}} \tag{1.28}$$

This means that most δ electrons, since they have a low energy, are emitted in a direction perpendicular to the primary particle track at θ close to 90°.

Range and energy dissipation of electrons

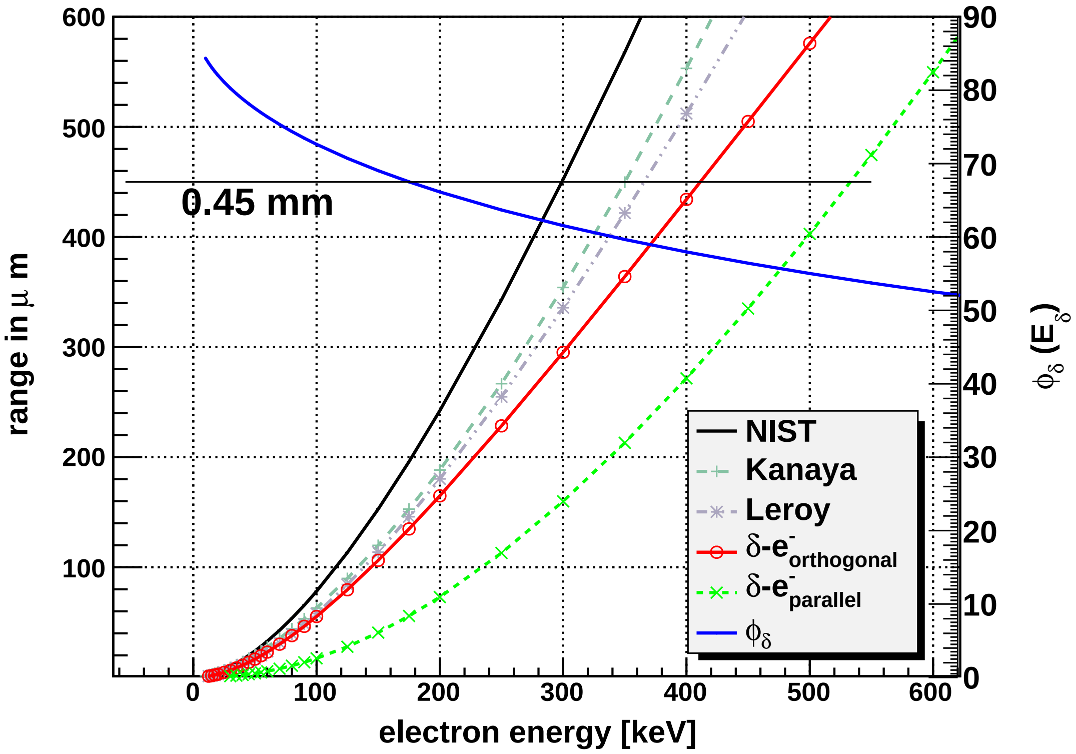


Figure 1.7: The energy range relationship for electrons in silicon plotted for three different approximations for an energy range that sheds light on the energy deposition of δ electrons inside the detector as a whole. The plotted curves are: ***NIST*** *is continuous slowing down approximation taken from the online service of the national institute of standard and technology (NIST),* ***Kanyaka*** *is the Kanaya-Okayama range using equation 1.30,* ***Leroy*** *is the practical range given by equation 1.29,* $\boldsymbol{\delta\text{-}e^-_{orthogonal}}$ *is the projection of a δ electron range onto the plane orthogonal to the primary particle track when using the practical range (equation 1.29) and taking the effects of the angular distribution of δ electrons (equation 1.28) into account.* $\boldsymbol{\delta\text{-}e^-_{parallel}}$ *is parallel complement to* $\delta\text{-}e^-_{orthogonal}$*,* ϕ_δ *is the emission angle of the δ electron according to equation 1.28. From* $\delta\text{-}e^-_{orthogonal}$ *it becomes clear that electrons with more energy than* $E \approx 550\,keV$ *start leaving the sensor and that therefore the deposited energy inside the detector will be different from the total energy lost by collisions (the restricted energy loss, equation 1.8).*

For the primary particles impact point reconstruction it is not only of importance how many δ electrons are created with a certain energy but how far they can travel. There are several definitions for the range of an electron (e.g. [10, p.98]), among them are

- **The CSDA range**: a very close approximation to the average path length traveled by a charged particle as it slows down to rest, calculated in the continuous-slowing-down approximation (CSDA). In this approximation, the rate of energy loss at every

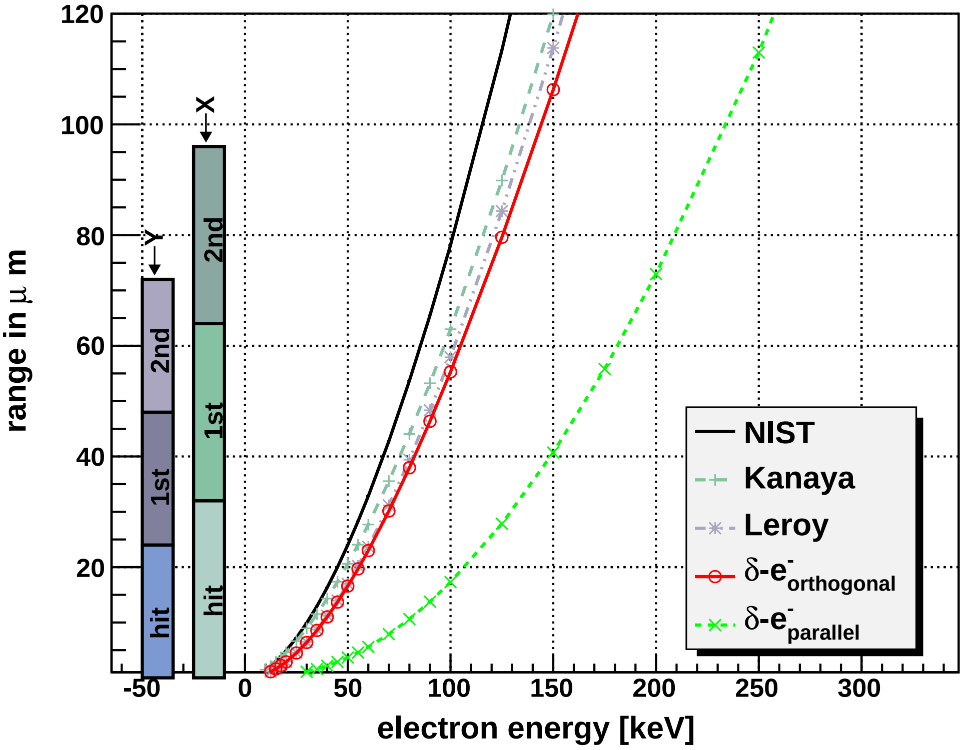


Figure 1.8: *The energy range relationship for electrons in silicon plotted for three different approximations. The plot is a zoom in of plot 1.7 for a smaller energy range. Also the plotting of ϕ_δ has been omitted. The electron ranges are now in the order of the dimensions of the pixels and clusters used in the DEPFET test beam experiment. For illustrative purposes the extent of three DEPFET pixels in X and Y direction are plotted to the left of the curves. "hit" refers to the pixel hit by a particle; "1st" and "2nd" are the according neighbor pixels.*

point along the track is assumed to be equal to the total stopping power. Energy-loss fluctuations are neglected. The CSDA range is obtained by integrating the reciprocal of the total stopping power with respect to energy. In the range energy relationship plot in figures 1.7 and 1.8 the CSDA approximation of the *NIST estar* online service was used.

- **The practical range**: The transmission curves for mono energetic electrons (fig. 1.9) have a long, linear part. The extrapolation of this linear part to zero transmission is the so-called *practical electron range* R_p. In the range energy relationship plot in figures 1.7 and 1.8 the following approximation valid for $0.01 < E < 3\,\text{MeV}$ is used [4]

$$R_p = 0.412E^{1.265-0.0945\ln E} \tag{1.29}$$

with in E in MeV and R_p in g/cm^2.

- **Kanaya-Okayama range**: This range takes elastic and inelastic scattering effects into account [11]

$$\rho R = \frac{2.76 \cdot 10^{-11} A E_0^{5/3}}{Z^{8/9}} \frac{(1 + 0.978 \cdot 10^{-6} E_0)^{5/3}}{(1 + 1.957 \cdot 10^{-6} E_0)^{4/3}} \tag{1.30}$$

with E_0 in **eV** and ρ in g/cm^3

In figure 1.7 these three ranges are plotted together with the emission angle of a δ electron (equation 1.28) and the according horizontal and vertical component of the practical range (equation 1.29). From this plot it becomes clear that δ-electrons with an energy above $E(\delta e^-) \gtrsim 550keV$ will leave the sensor and therefore the deposited energy inside the detector will be different from the total energy lost by collisions. Here the restricted energy loss, equation 1.8, comes into play with - judging from fig. 1.7 - an upper limit between 1 and 2 MeV. A similar plot is shown in figure 1.8 where the electron ranges are compared to the pixel sizes and the read out area of a cluster. A cluster is the read out area of the sensor where the pion has passed through and is 5×5 pixel large and centered at the impact point of the particle. One can see that δ electrons with an energy above $E > 150\,\text{keV}$ will leave the cluster. One way to estimate how much electrons with a given energy will leave the detector or a given read out area are *transmission curves*.

Transmission curves

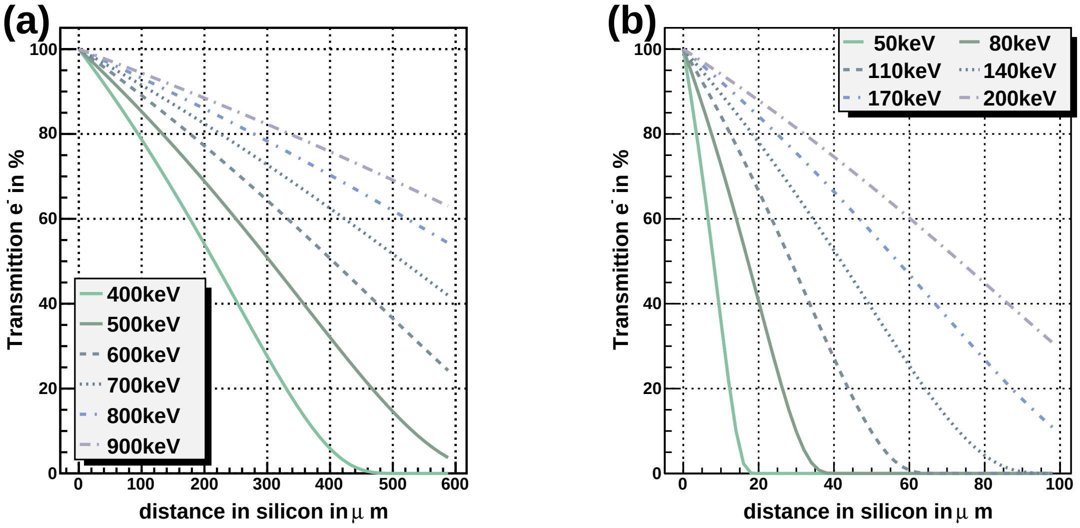


Figure 1.9: *Transmission curves for electrons in silicon according to the Kanaya-Okayama approximation (equation 1.32). On the abscissa is the sensor thickness in µm. The left plot* ***(a)*** *shows electrons at energies where they start leaving the sensor (450 µm). The right plot* ***(b)*** *shows distances and energies in the regime of the pixel dimensions.*

These curves show the fraction of electrons which will leave material of a given thickness. Among others there are these two approximations:

- Plotting T versus x/x_0 with x_0 beeing the mass-thickness corresponding to a transmission of 50% [10]:

$$T(x/x_0) = \exp\left(\frac{x}{x_0}\right)^p \tag{1.31}$$

- A transmission curved based on the Kanaya-Okayama range R_{ko} [11],

$$T(y) = \exp\left(-\frac{\gamma y}{1-y}\right), \tag{1.32}$$

  where $y = x/R_{ko}$ is a reduced depth with R_{ko} from equation 1.30 and $\gamma = 0.187 \cdot Z^{2/3}$ is a factor that takes into account diffusion loss due to multiple collisions for returning electrons and energy retardation due to electronic collisions. For silicon $\gamma_{Si} = 1.086$.

Figure 1.9 shows the latter approximation for the transmission curve **(a)** for distances comparable to 450 μm thick sensor in the left plot and **(b)** for distances comparable to the typical DEPFET pixel cluster size in the right plot. This plot supports the above estimate which gives a limit to the energy deposition by δ electrons of roughly $(1-2)$ MeV.

Energy deposition vs penetration depth

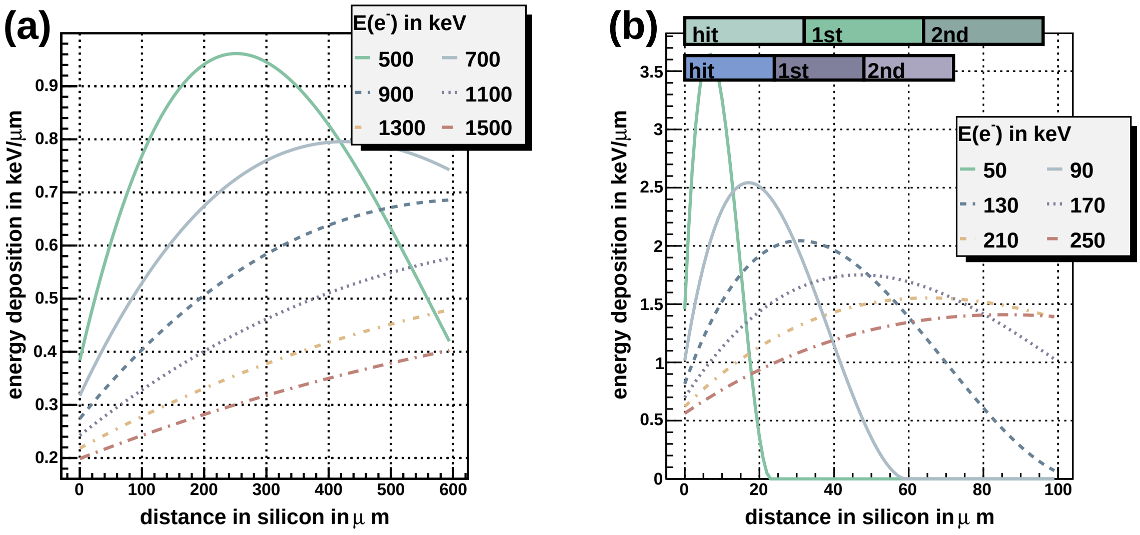


Figure 1.10: *The energy deposition in 450 μm silicon as a function of penetration depth in silicon according to equation 1.33. The Kanaya-Okayama range R_{ko} from equation 1.30 was used for the reduced depth $y = x/R$. The graph on the left **(a)** shows distances comparable to the thickness of a DEPFET sensor (450 μm). The graph on the right **(b)** shows distances comparable to a typical pixel cluster. For illustrative purposes the extent of three DEPFET pixels in X and Y direction is added to the plot on the left. "hit" refers to the pixel hit by a particle; "1st" and "2nd" are the according neighbor pixels.*

Besides the emission probability, angle, and range of a δ electron with a certain energy it is also important to know how the energy deposition looks like as a function of distance traveled or penetration depth. The following approximation comes from low energy electron beam studies [12]:

$$\phi(y)_{Norm} = 0.60 + 6.21y - 12.4y^2 + 5.69y^3 \tag{1.33}$$

where $\phi_{Norm}(y)$ is a normalized approximation for the depth distribution of dissipated energy with $y = x/R$ being the reduced depth. Alternatively a shifted Gaussian can be used as an approximation [10]. Figure 1.10 shows the electron energy deposition $\phi(y)$ in $450\,\mu$m silicon as a function of reduced depth using the Kanaya-Okayama range R_{ko}[1] as denominator. From the left plot **(a)** in this figure it is clear that higher energetic electrons with $E > 500\,\text{keV}$ leave only a fraction of their energy in the sensor thus again arguing for a energy loss restricted to $E \approx 1-2\,\text{MeV}$. The righthand plot **(b)** shows that - similar to the transmission curves - δ electrons below $E \lesssim 50\,\text{keV}$ are absorbed within the seed pixel.

Multiple Coulomb Scattering

As presented before there are several definitions for the range of an electron in matter. This is mainly because there is a profound difference between the total path length a particle travels in matter and the actual path length in one direction (e.g. the depth of the silicon sensor). The difference is caused by scattering processes deflecting the electrons trajectory. So far the interaction between a particle and the nuclei of the traversed medium has been ignored because the cross sections for nuclear interactions is negligibly small in the cases dealt with in this thesis. However, coulomb interaction with nuclei plays a role since this causes the particle trajectory to be deflected. For small angles the distribution of scattering angles can be approximated by a Gaussian distribution with a width of [1]:

$$\theta_0 = \frac{13.6\,\text{MeV}}{p\beta} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln\left(\frac{x}{x_0}\right) \right], \tag{1.34}$$

where p, β, and z are the momentum, velocity, and charge of the incident particle, and x/X_0 is the thickness of the scattering medium in units of X_0. It should be noted that outside the small angle range the distribution of scattering angles becomes non Gaussian. Multiple scattering plays a crucial role for particle detectors, especially silicon vertex detectors. As will be further explained in the section 2.1 vertex detectors consist of several layers each providing a space point for particle track reconstruction. Multiple scattering alters the tracks and impairs the detector's tracking abilities. Hence multiple scattering should be minimized. The same is true for a test beam experiment where the *beam telescope* provides the space point for tracking. As can be deduced from equation 1.34 there are two ways to keep the multiple scattering angle small:

- Increase the particles momentum. This option is only viable in a test beam experiment and was the main reason why the DEPFET test beam activates were moved

---

[1] The Kanaya-Okayama range R_{ko} is plotted in figure 1.7 and figure 1.7.

from the 6 GeV electron beam at DESY, Hamburg to the 120 GeV pion beam at the SPS at CERN, Geneva. As will be explained in the next chapter, a vertex detector should have little effect on the particle trajectory even if it has a small momentum.

- Decrease the material length x/X_0. This basically means thinner sensors and small Z and is important for vertex and tracking detector subsystem in particle experiments. Large parts of the DEPFET vertex detector concept for the ILC, which will be presented in the next chapter, are built reducing the amount of scattering material.

1.2 Semiconductor Detectors

Semiconductors are materials that have an electrical conductivity between that of a conductor and an insulator (fig. 1.11). Crystalline semiconductors like silicon have two important energy bands similar to metallic conductors: the valence and the conduction band. Unlike a metallic conductor, however, these bands do not overlap but are separated by an energy gap which is characteristically for each material. The band gap in silicon is $E_{gap}^{Si} = 1.12\,eV$, which is beneficially low for an utilization as a particle detector material. Due to thermal excitation some electrons will leave the valence band and leave a hole.

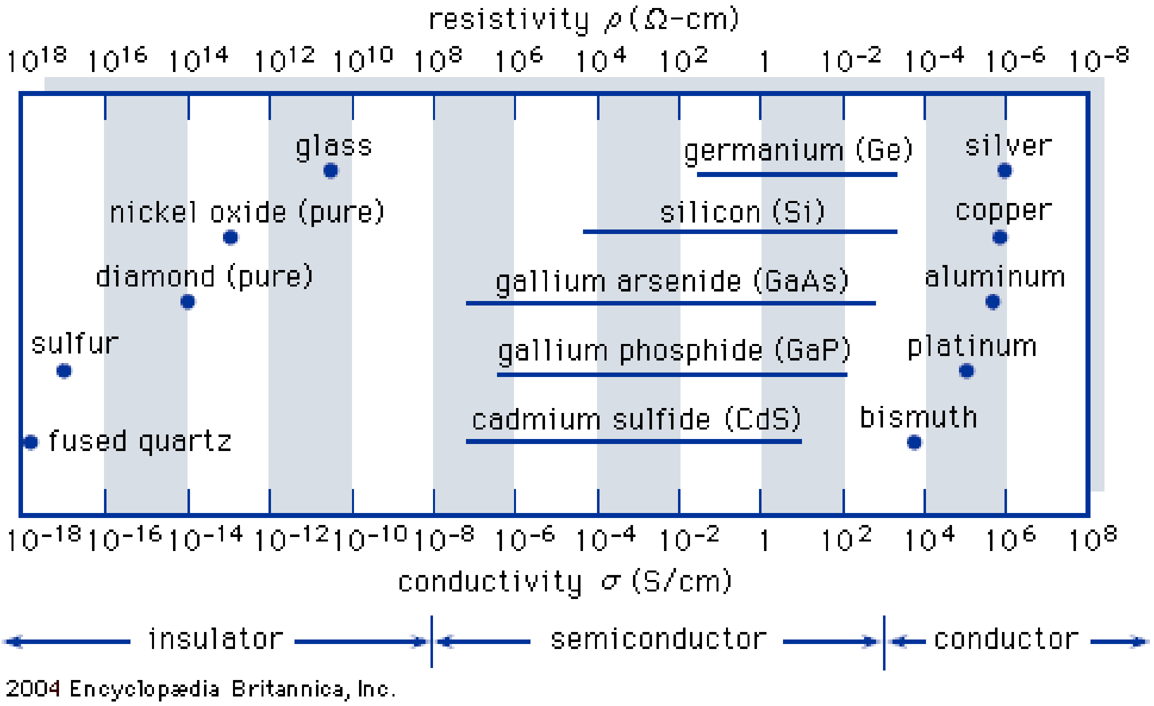


Figure 1.11: *The conductivity of several materials. Semiconductors have a large range of conductivity due to different doping concentrations.*

After some time they will then recombine. The product of electrons n and holes p is constant for a given temperature

$$np = n_{int}^2 \tag{1.35}$$

n_{int} is the intrinsic carrier concentration and for silicon at $T = 300\,\mathrm{K}$, $n_{int} \simeq 1.45 \cdot 10^{10}\,\mathrm{cm}^{-3}$ [4]. The intrinsic carrier concentration has a temperature dependence of

$$n_{int} \approx T^{3/2} \exp\left(\frac{E_{gap}}{2k_bT}\right) \tag{1.36}$$

where $k_b = 8.617 \cdot 10^{-5}\,\mathrm{eV/K}$ is the Boltzmann constant. The mobility for electrons μ_e and holes μ_h is different and therefore they have a different velocities $v = \mu E$ in a given field E. Both mobilities are temperature and field depended.
By introducing impurities in the lattice structure of silicon one can modify the amount of free charge carriers of one type, a process called **doping**. The two types are called n or *donors* type if there is an electron excess and conversely p or *acceptor* if an excess of holes is created. Doping allows to change the properties of silicon, for instance the conductivity [13]

$$\rho = \frac{1}{q_e N \mu} \tag{1.37}$$

where N is the dopant concentration and μ the majority carrier mobility.

P-N junction

The standard silicon detector consists of a p and a n doped region and it is therefore a diode. p and n type charge carriers diffuse along the concentration gradient. While the two carriers types diffuse they recombine with each other and create a border zone void of free charge carriers. However, in this initially neutral zone the depletion of charge carriers will leave (now) electrically charged atoms behind and two space charge regions, a positive and a negative one, are created. This generates a field that, in a state of equilibrium, cancels the migration pressure caused by the concentration gradient (fig. 1.12). The size of the space charge region x_n and x_p is dependent on the donor concentrations N_d and N_a:

$$N_d x_n = N_a x_p \tag{1.38}$$

Assuming an abrupt junction one can approximate the space charge densities $\rho(x)$ with

$$\rho(x) = \begin{cases} q_e N_d \text{ for } 0 \leq x \leq x_n \\ q_e N_a \text{ for } 0 \leq x \leq x_p \end{cases} \tag{1.39}$$

and calculate the contact or build-in voltage V_{bi} and the size of the depletion zone by solving the Poisson equation

$$\frac{\partial^2 \phi(x)}{\partial x^2} = -\frac{\rho(z)}{\epsilon} \tag{1.40}$$

with [4]

$$\epsilon = \epsilon_0 \epsilon_{Si} = 1.054, \mathrm{pF/cm} \tag{1.41}$$

The build-in voltage is then [13]

$$V_{bi} = \frac{kT}{q_e} \ln\left(\frac{N_a N_d}{n_{int}^2}\right) \tag{1.42}$$

where q_e is the elementary charge and the width of the depletion zone is

$$d = \sqrt{\frac{2\epsilon}{q_e}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)(V_{bi})} \tag{1.43}$$

For the detection of particles it is desirable to enlarge the *depletion zone*. This is done by applying a voltage in the same direction as the built-in voltage. This process is called *reverse biasing* and equation 1.43 becomes

$$d = \sqrt{\frac{2\epsilon}{q_e}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)(V_{bi} + V_{ext})} \tag{1.44}$$

Usage as a particle detector

With a sufficiently large reverse bias the diode becomes fully depleted. If an ionizing particle traverses the silicon detector the energy it deposits will create electron hole pairs. However, the energy needed on average to create an electron hole pair in silicon is not $E_{gap}^{Si} = 1.21\,\text{eV}$ but

$$E_{e/h} \simeq 3.6\,\text{eV} \tag{1.45}$$

This is because a part of the energy is used for phonon excitations in the silicon lattice. In the absence of a particle there is still a ”signal” generated by the *dark* or *leakage current* which is caused by the thermal fluctuation of charge carriers. This leakage current has a temperature dependence of

$$I_l \propto T^2 \exp\left(-\frac{E_{gap}}{2k_B T}\right) \tag{1.46}$$

For most silicon sensors the junction is realized by a shallow and highly doped ($N_a > 10^{18}\,\text{cm}^{-3}$) p^+-implant in a low doped ($N_d \approx 10^{12}\,\text{cm}^{-3}$) bulk material [13]. Furthermore the build-in voltage is much smaller than the usual reverse bias voltage $V_{bi} \sim 0.5\,\text{V} \ll V_{dep} \gtrsim 50\,\text{V}$. These conditions simplify equation 1.44 to

$$d = \sqrt{\frac{2\epsilon}{q_e N_d} V} \simeq 0.53\sqrt{\rho_R V}\,\mu\text{m} \tag{1.47}$$

where ρ_R is the resistivity in Ωcm given as [4]

$$\rho_R = \frac{1}{\mu q_e N_D} \tag{1.48}$$

and V in volts. The bias voltage at full depletion of a silicon sensor of thickness D can be expressed by [4, 14]

$$V_{dep} \simeq \frac{D^2}{2\mu\epsilon\rho_R} \approx \frac{q_e N_d D^2}{2\epsilon} \tag{1.49}$$

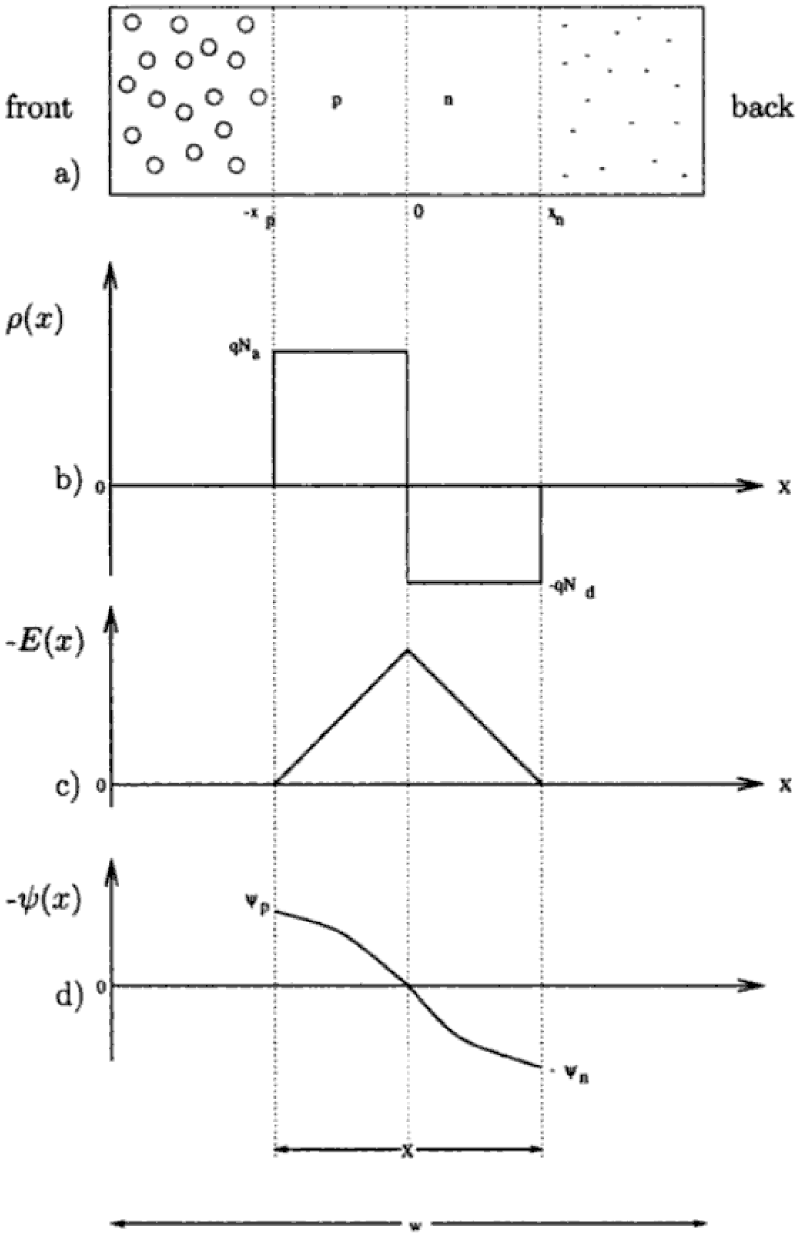


Figure 1.12: *Schematic representation of a p-n junction of thickness w taken from [4, p.455]. The applied bias voltage V_B creates a depletion depth X. (a) represents the junction, (b) the charge distribution $\rho(x)$, (c) the electric field $-E(x)$ and (d) the electrostatic potential $-\psi(x)$.*

The Charge Cloud

The charge generated by an ionizing particle will drift towards the anode with

$$v_{drift}(z) = \frac{dz}{dt} = \pm\mu E(z) \tag{1.50}$$

However, during this drift time the initially small charge cloud will become broader due to diffusion effect. The diffusion process is governed by the diffusion constant D_n and D_p for electrons and holes and the gradients ∇n and ∇p of the electron and hole concentration and the resulting diffusion current density is

$$J_{n,diff} = -D_n \nabla n = -\frac{kT}{q_e}\mu_n \nabla n \tag{1.51}$$

where the Einstein equation

$$\frac{D_N}{\mu_n} = \frac{kT}{q_e} \tag{1.52}$$

was used. $\frac{kT}{q_e}$ is sometimes called the thermal voltage and is $\frac{kT}{q_e} = 0.0259\,\mathrm{V}$ for $T = 300\,\mathrm{K}$. The width of the charge cloud is in first approximation a function of the drift time and the drift time in turn depends on the electric field:

$$E(z) = -\left[\frac{V + V_{dep}}{d} - \frac{2zV_{dep}}{d^2}\right] \tag{1.53}$$

Integration of 1.50 together with 1.53 yields an expression for the drift time [14]

$$t_{drift} = \frac{d^2}{2\mu_e V_{dep}} \ln\left[\frac{V + V_{dep}}{V - V_{dep}}\left(1 - \frac{z}{d}\frac{2Vdep}{V + V_{dep}}\right)\right] \tag{1.54}$$

If one assumes for the initial charge distribution a Dirac δ distribution, the charge cloud shape can be approximated by the fundamental solution to the diffusion equation

$$\phi(x, t_{drift}) = \frac{1}{\sqrt{4\pi D_e t}} \exp\left(-\frac{x - x_0}{4kT}\right) \tag{1.55}$$

which is a normal distribution with a width of

$$\sigma_{cloud} = \sqrt{2D_p t_{drift}} = \sqrt{2\frac{kT}{q}\mu_e t_{drift}} \tag{1.56}$$

Figure 1.13 shows the drift time and the charge cloud width as a function of the sensor depth for different bias and depletion voltages. The actual charge cloud is the superposition of the charge clouds generated along the track of an ionizing particle. This cloud will have a non-Gaussian shape, in this work, however, it will be approximated to first order with a Gaussian with a width which is equal to the average of all Gaussian widths along the path. These are the red lines in the figure 1.13. For a $450\,\mu$m thick sensor the width is $\sigma_{cloud} \approx (6 - 7)\,\mu$m. This is a lower limit as other effects have been neglected:

- The initial charge distribution was approximated with a Dirac δ function. However, a real charge distribution will have a certain extension. According to equation 1.26 on average every track will have a δ electron with an energy of $E_\delta \approx 8\,\mathrm{keV}$ and accordingly more at lower energies, although the condition $E_\delta \gg I = 0.173\,\mathrm{keV}$ starts to be violated. The range of a $E_\delta = 8\,\mathrm{keV}$ electron is roughly $R \approx 1\,\mu$m, which could be used as a first approximation.

- According to [15] the broadening of the initial distribution due to electrostatic repulsion is not negligible. However, their work focuses on electrons in silicon drift detectors and they assume a complete electron hole separation as initial condition.

It is therefore difficult to give a precise width of the charge cloud, however, for a 450 μm thick silicon sensor a width of $\sigma_{cloud} \approx 6-8\,\mu$m can be expected.

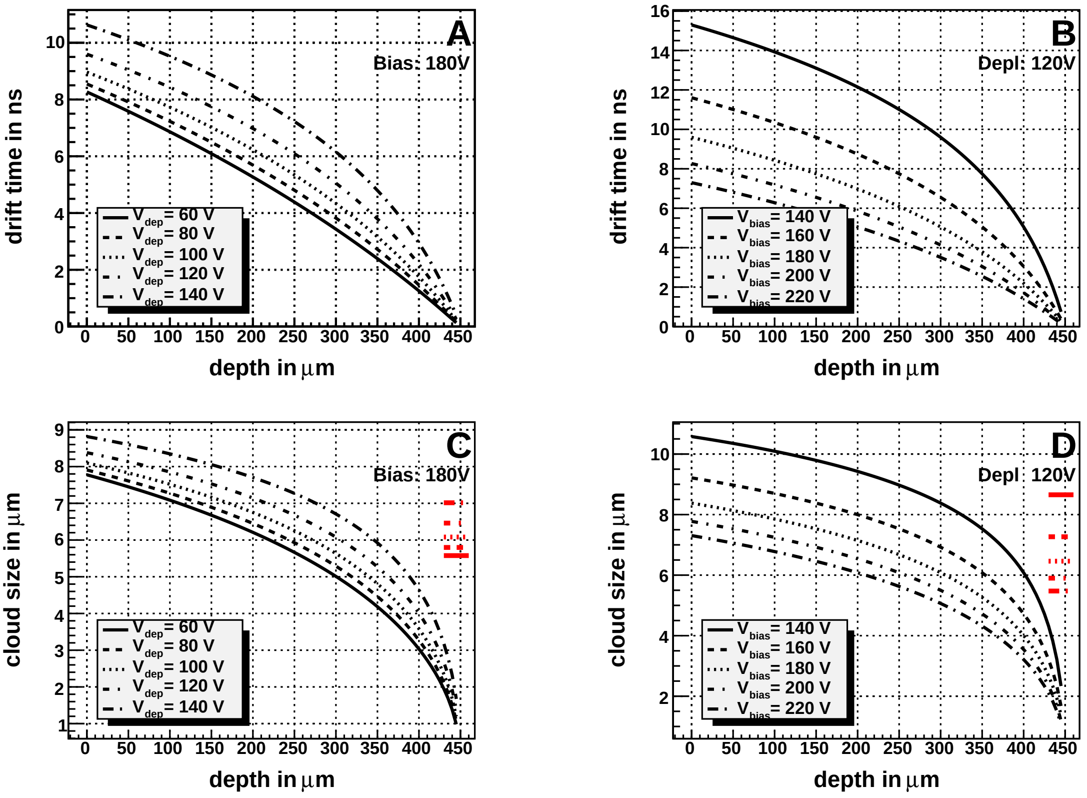


Figure 1.13: *These four plots shows the dependence of the drift time t_{drift} (upper two plots, A and B) and the charge cloud size D_{cloud} (lower two plots, C and D) as a function of the sensor depth for a 450 μm thick silicon sensor. In the plots on the left side the voltage at full depletion V_{dep} is varied and the bias voltage is kept fixed at $V_{bias} = 180$ V. In the plots on the right the voltage at full depletion was kept fixed at $V_{dep} = 120$ V and the bias voltage V_{bias} was changed. The drift time and the cloud size were calculated using equations 1.54 and 1.56. The red lines in the two lower plots mark the average of the widths for each voltage settings.*

1.2.1 The DEPFET sensor

The abbreviation DEPFET - **Dep**leted **F**ield **E**ffect **T**ransistor - contains already the basic principle of this detector type: A FET transistor inside a fully depleted sensor bulk, which is steered by electrons in a potential minimum below its channel. This potential minimum is created by means of sidewards depletion.

Sidewards depletion

The idea of sidewards depletion was first introduced as a way to realize silicon drift detectors [16]. Instead of a simple p-n junction the detector has a p implantation on each side of the n doped bulk material. With an additional $n+$ implantation on the tope side as illustrated in figure 1.14 the bulk is set to a constant potential V_{bulk}, e.g. ground. With increasing reverse bias voltages V_u and V_d on both p contacts the depletion zone of both p-n junctions grows until the sensor is fully depleted. By solving the Poisson equation (1.40) the one dimensional potential distribution perpendicular to the sensor surface is now

$$\phi(z) = \frac{\rho}{2\epsilon} z(d-z) + \frac{z}{d}(V_d - V_u) + V_u \tag{1.57}$$

with a minimum $(d\phi(z)/dz = 0)$ at

$$z_{min} = \frac{d}{2} + \frac{\epsilon}{\rho d}(V_d - V_u) \tag{1.58}$$

For equal bias voltages $V_u = V_d$ the voltage necessary for complete depletion of the sensor is four times lower with sidewards depletion than without as $V_{dep} \propto D^2$ (equation 1.49). As can be seen from equation 1.58 it is possible to move the potential minimum close to one surface by applying correspondingly asymmetric bias voltages, which is essential for the DEPFET sensor. When an ionizing particle creates charge the electrons will drift to this potential minimum. In case of drift sensor a second field moves the electrons towards a read out anode. The DEPFET however works differently.

The DEPFET pixel

The sidewards depletion enables the shaping of a vertical potential minimum, however for the DEPFET also lateral potential minima are needed. This is realized by implanting a small deep-n doped region located under the external gates and therefore the channel of a transistor (fig. 1.15). Electrons created by a radiation will move to this confined area and their presence will alter the potential below the transistor just like a voltage change on the external gate therefore acting like an *internal gate*. The device amplification is the amount of current modulation ∂I_D due to the collected charge ∂Q

$$g_q = \left.\frac{\partial I_d}{\partial Q}\right|_{V_{GS},V_{DS}} \tag{1.59}$$

Current values of g_q are in the order of $g_q \approx 300\,\mathrm{pA/e^-}$. This concept has several benefits over other sensor types:

- Charge collection is ongoing even if the DEPFET transistor is switched off which plays an essential role for the energy dissipation budget in the ILC vertex detector concept.
- The charge is collected in a fully depleted bulk. This allows for a fast and complete charge collection unlike diffusion based concepts like MAPS.

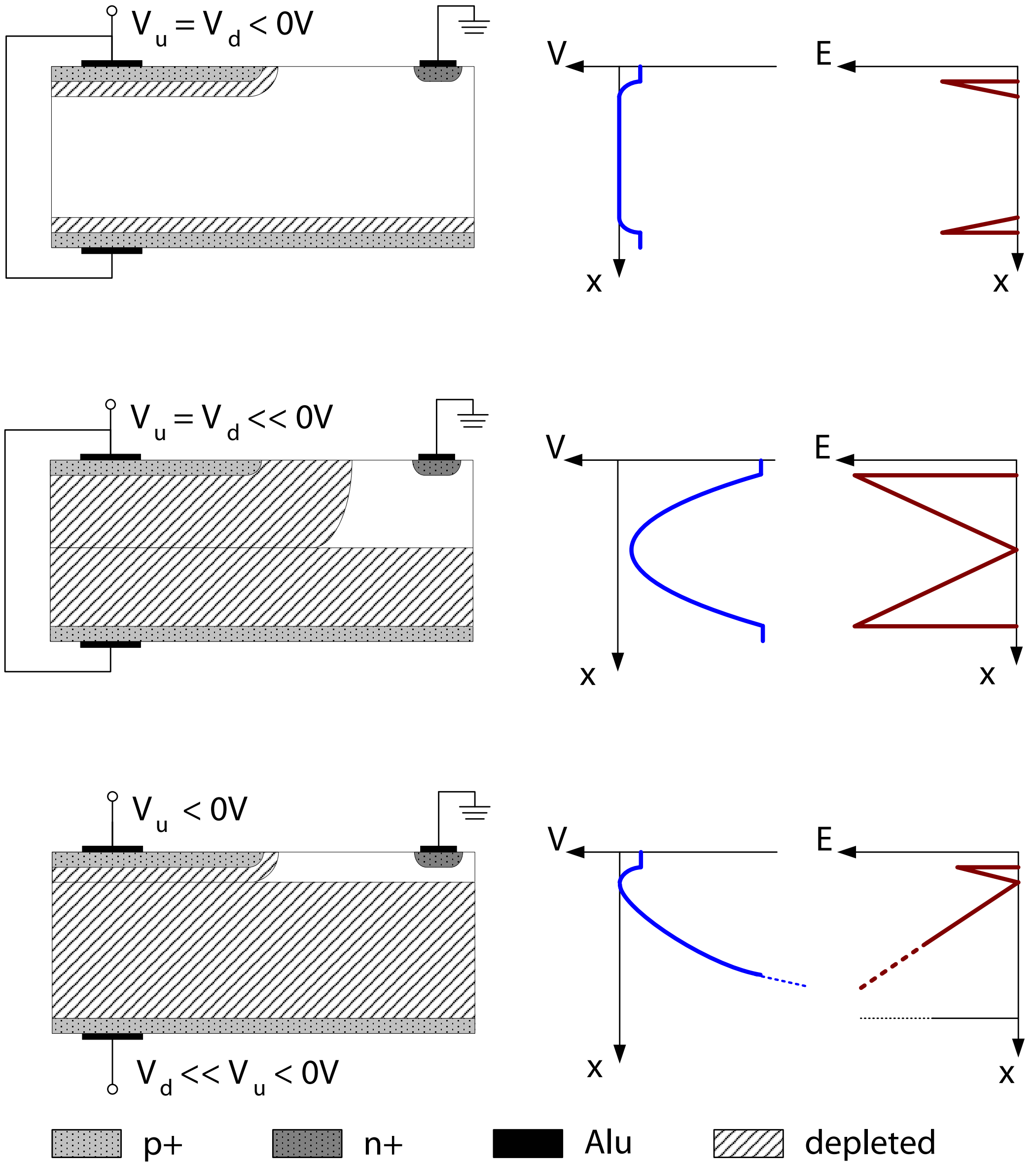


Figure 1.14: *Principle of sidewards depletion: The bulk material is kept at fixed (ground) potential via an n+ contact while the two diodes are depleted from both side with increasing reverse biasing on both p contacts. The position of the potential minimum can be steered by applying different depletion voltages* $V_d \neq V_u$ *to both sides. The resulting potential and electrical field distributions are shown on the right.*

- The input capacitance is very low, in the order of 10 fF. This allows very low noise measurements. Using an ILC type DEPFET with a 10 μm shaping time at room temperature ^{55}Fe measurements with an R.M.S. noise of only $ENC = 1.6\,\mathrm{e}^-$ were achieved [17]

The combination of the last two points yields an excellent signal over noise ratio.

$$\frac{\delta\ \Phi}{\delta} \qquad \frac{}{}$$

ff

The clear process Unlike a silicon drift chamber the charge is not removed by the read out. Therefore an additional *clear contact* is introduced next to the FET gate (see Figure 1.15) to facilitate the removal of the electrons by applying a *clear* pulse. A further contact, the *clear gate*, allows to further control the clear process. This is necessary since clearing of electrons out of the internal gate and collection of electrons inside the internal gate are competing processes. Their interplay is a complicated process and a fined tuned set of bias voltages is needed to achieve the highest signal to noise. Studies can be found in [18, 19, 20, 21]. Originally, the *clear gate* signal was pulsed but the innovations of new DEPFET generations allowed for the introduction of a so called *Common Clear Gate* or **CoCG** device type where only one static *clear gate* voltage is applied. A detailed description of the clear process can be found in [18, 22].

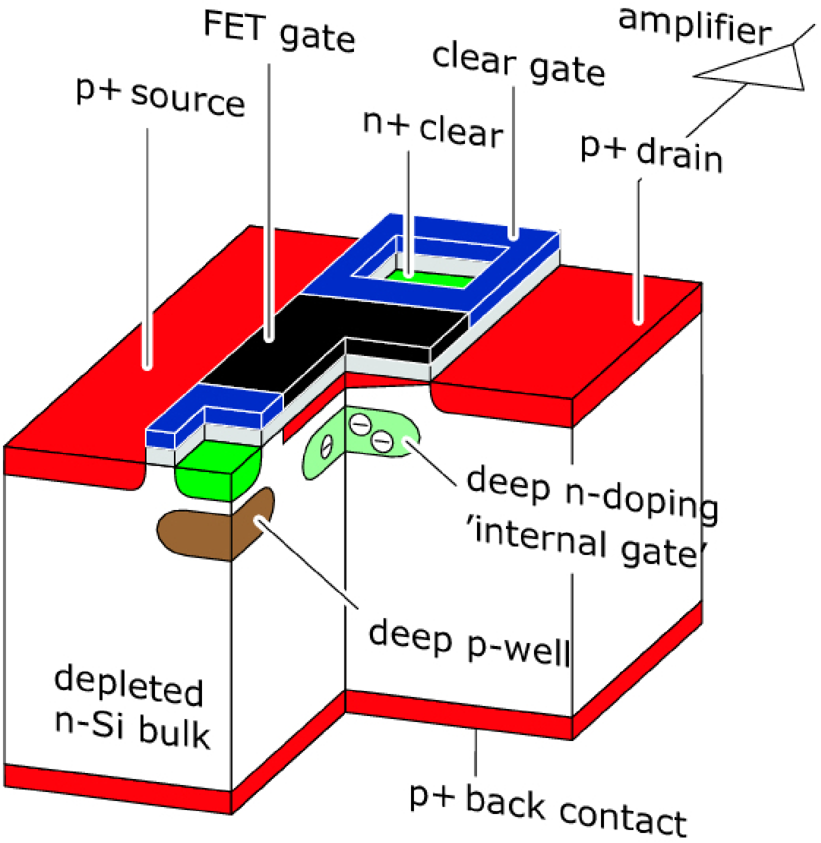


Figure 1.15: *Cross section of a linear DEPFET: The deep n-doping area labeled 'internal gate' is a potential minimum created by means of sidewards depletion and proper doping profiles. Electrons generated by an ionizing particle will move to the internal gate and modulate the transistor current, thus creating a signal. To remove the electrons from the internal gate after read out a clear pulse is applied to the* n^+*-clear contact. The additional clear gate contact allows to steer the competing charge clearing and collecting processes to achieve an optimal signal to noise ratio.*

2

The ILC prototype system

2.1 The vertex detector at the ILC

The physics goals of the international linear collider (ILC) and other modern high energy particle accelerator experiments are manifold. Among them are:

- Precision measurements of the Higgs boson.
- Is there physics beyond the standard model of particle physics like super symmetry?
- How many dimensions does the universe have?

In all these areas, access to the high quality physical measurements relies on heavy flavor identification with high efficiency and purity [23, 24]. The ability to separate b, c, and $udsg$ jets helps for example in suppressing background events.

Flavor tagging

The ability to distinguish between different heavy quark flavors of jets relies on their different properties in terms of mass and decay time. In the interaction point of the experiment short lived particles like the $\gamma(4s)$ decay virtually instantly and the fragments are splitting up in the primary vertex (PV). The lifetime of the B mesons, however, is in the order of 1.6 picoseconds (Table 2.1) and therefore long enough to displace a secondary vertex, created by the B meson decay, from the primary vertex. This is illustrated in figure 2.1. Since the B meson decay is likely to contain a c quark fragment, even tertiary vertices can be found. There are several variables used to tag a yet as a b or c quark type [25]:

- The impact parameter (IP), which is the distance between the reconstructed track of a particle and the primary vertex, i.e. the collision point. This is illustrated in figure 2.1.
- Measurement of secondary (and tertiary) vertices and their distance to the primary vertex.

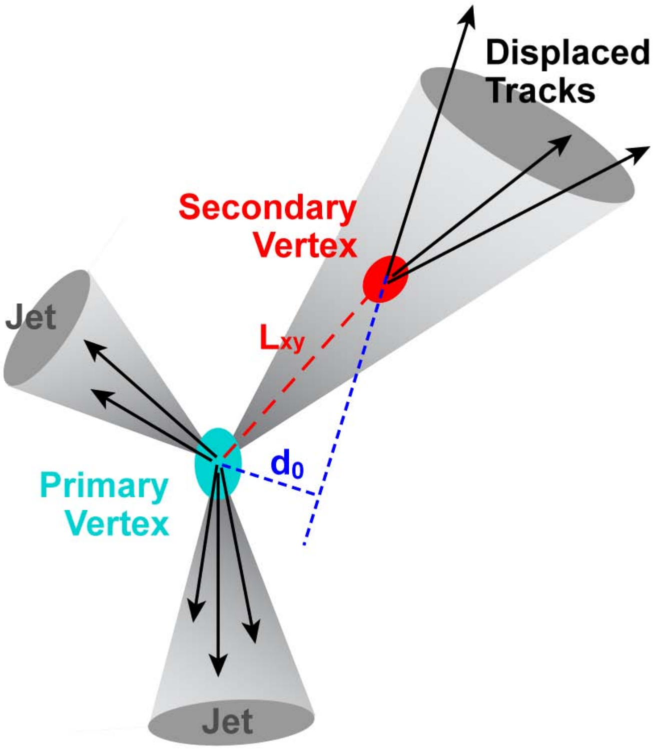


Figure 2.1: *Illustration of heavy flavor tagging by liftime/vertex displacement: A primary particle decays into several fragments of which one is a heavy quark (i.e. with b or c quark content) which travels a certain distance. The heavy quark then decays in several particles leaving displaced (with respect to the primary vertex) tracks inside the detector. The distance between the reconstructed track of a particle and the primary vertex* d_0 *is called the impact parameter and is used as a tagging variable. Furthermore all tracks associated with the secondary vertex can be used for further tagging variables (see figure 2.2).*

- The invariant mass of all tracks associated with a secondary vertex. The mass in c-jets is limited by the D-meson mass.
- The fraction of the charged jet energy included in the secondary vertex. This reflects the different fragmentation properties of different flavors.
- The transverse momentum at a secondary vertex with respect to the b-hadron flight direction. This takes missing particles, e.g. neutrinos from semi-leptonic decays, into account.
- The track rapidity $\varphi = \text{arctanh}\beta$ which differs for b- and c jets due to the higher mass of b mesons.

The first two variables are sometimes referred to as *lifetime tags*, whereas the last four are sometimes called *secondary vertex tags*. Figure 2.2 shows simulated distributions of secondary vertex tag variables for b and c quarks at the DELPHI experiment. Another

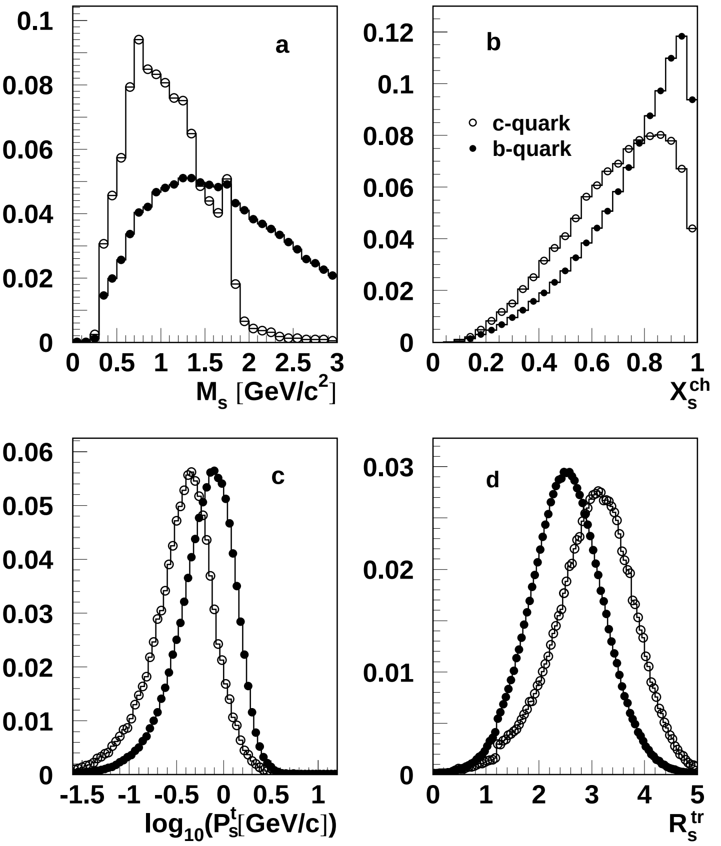


Figure 2.2: *Simulated distribution of four secondary vertex tagging variables for Z hadronic events in the DELPHI experiment taken from [25]. a) is the invariant particle mass, b) The fraction of the charged jet energy included in the secondary vertex, c) the transverse momentum at the secondary vertex, and d) is the rapidity for each SV track.*

variable that can be used is the so-called *soft lepton* tag which is based on semi leptonic decays. However, the low semileptonic branching ratio makes this variable less suitable. Common to all tagging variables is that they need an excellent tracking performance as close as possible to the interaction point. The figure of merit is the *impact parameter resolution*

$$\sigma_{IP} = \sigma_{IP}^{res} \oplus \sigma_{IP}^{ms} \tag{2.1}$$

where $\oplus$ is the quadratic addition of two resolution affecting parameters:

- the position resolution of the vertex detector σ_{IP}^{res}, and
- the multiple scattering σ_{IP}^{ms} of the particle due to the material of the beam pipe and the vertex detector.

For efficient b and c tagging it is important to keep σ_{IP} as low as possible and there are basically three ways to achieve that:

1. The first sensor must be as close as possible to the interaction point, i.e. the radius R_1 of the inner detector layer must be as small as possible,

2. the radiation length x/X_0 of the material of both, beam pipe and sensors, must be kept to a minimum possible, and

3. the intrinsic position resolution of the vertex silicon detector must be as good as possible.

These are, however, conflicting requirements and a realistic set of parameters for a vertex detector constituting a compromise of the technologically feasible and the desire for best physics performance.

particle	quarks	spin	mass [GeV]	life time τ [s]	$c\tau$
π^+	$u\bar{d}$	0	0.140	$2.603 \cdot 10^{-8}$	7.8 m
π^0	$u\bar{u}, d\bar{d}$	0	0.135	$8 \cdot 10^{-17}$	25nm
$K^{+/-}$	$u\bar{s}$	0	0.494	$1.238 \cdot 10^{-8}$	3.7 m
K_L^0	$d\bar{s}$	0	0.497	$5.116 \cdot 10^{-8}$	15.8 m
K_S^0	$d\bar{s}$	0	0.497	$0.89 \cdot 10^{-10}$	2.7cm
D^+	$c\bar{d}$	0	1.869	$1.04 \cdot 10^{-12}$	311.8 μm
D^0	$c\bar{u}$	0	1.865	$0.41 \cdot 10^{-12}$	122.9 μm
B^+	$u\bar{b}$	0	5.279	$1.638 \cdot 10^{-12}$	491.1 μm
B^0	$d\bar{b}$	0	5.279	$1.530 \cdot 10^{-12}$	458.7 μm
J/ψ	$c\bar{c}$	1	3.097	$\lesssim 10^{-20}$	-
$\upsilon(4s)$	$b\bar{b}$	1	10.57	$\lesssim 10^{-20}$	-

Table 2.1: *Some properties of important mesons [1].*

2.1.1 The DEPFET vertex detector concept for the ILC

The constrains and requirements for the ILC vertex detector are well established [26]. The main parameter is the *impact parameter resolution* with

$$\sigma_b(r\Phi, z) \leq 5\,\mu\mathrm{m} \oplus \frac{10\,\mu\mathrm{m\,GeV/c}}{p\,\sin^{3/2}\theta} \tag{2.2}$$

To achieve the resolution, the first inner layer is placed very close to the beam pipe at $r = 15\,\mathrm{mm}$. To achieve this goal several technological challenges for the vertex detector must be met:

- **Pixel sizes**: To achieve the required spatial resolution an accordingly small pixel size in the order of $\approx 25 \text{ x } 25\,\mu\mathrm{m}^2$ is required.

- **Material budget**: The effects of multiple scattering especially for low momentum tracks are only within bearable limits if the material does not amount to more than $x/X_0 \approx 0.1\%$ per layer, which is 1/30 of the radiation length per layer for the ATLAS pixel sensor with $x/X_0 \approx 3\%$ per layer [27]. This can only be achieved with thinned down sensors. However, this means the signal will be scaled down as well. To still achieve the above mentioned spatial resolution an excellent signal to noise performance is required.

- **Power budget**: The above mentioned $x/X_0 \approx 0.1\%$ per layer can only be achieved if no cooling system, which would add additional material, is needed. This means the power dissipation of the ILC vertex detector can only be a fraction of the power dissipation of conventional vertex detectors like e.g. the ATLAS pixel detector with its 18.7 kW [28].

- **Read out speed**: At the radius $r = 15\,\mathrm{mm}$ of the inner most layer the level of the e^+e^- pair production background becomes very high. A hit multiplicity of ≈ 0.05 hits per mm^2 and bunch crossing at $\sqrt{s} = 500GeV$ must be tolerated with a hit occupancy of around 1%.

- **Radiation tolerance**: The requirements on radiation hardness are much less stringent than for the LHC [27]: Electron fluxes of about 1.7×10^{12} per cm^2 and year, corresponding to a total ionizing dose of about 4 kGy for 10 years of operation and a supplementary neutron equivalent flux of 8.5×10^{10} per cm^2 and year. The innermost layer of the ATLAS pixel detector has to endure a dose of 500 krad after only 5 years of operation and a fluence of $10^{15}\,\mathrm{n}_{eq}/\mathrm{cm}^2$ at the same time [29, 27].

All of these requirements are demanding, however a proposal for a DEPFET vertex detector for the ILC was put forward that could meet all the requirements [17]. The envisaged DEPFET vertex detector consists of several barrel shaped layers and each layer is made up of overlapping ladders. Table 2.2 shows the geometrical parameters of this concept and figure 2.3 shows a sketch of an ILC DEPFET module. In the following some features of the DEPFET ILC vertex detector will be addressed with regard to the technological challenges stated above. The radiation tolerance, however, will be skipped. As has been shown in studies the DEPFET sensor itself as well as the (newest generation of) ASICs are capable of working in the radiation environment of the ILC according to their specifications [17].

Layer	Number of ladders	Radius (mm)	Ladder length (mm)	width (mm)
1	8	15.0	100	13
2	8	26.0	2×125	22
3	12	37.0	2×125	22
4	16	48.0	2×125	22
5	20	60.0	2×125	22

Table 2.2: *Default geometrical parameters of the DEPFET based ILC micro-vertex detector [30, 17].*

Position resolution

To achieve the impact parameter resolution of $\sigma_{IP}^{res} = 5\,\mu\mathrm{m}$ a spatial resolution of $\lesssim 4\,\mu\mathrm{m}$ per layer is necessary. The binary position resolution of a 24 μm pitched pixels is $24/\sqrt{12} \approx 7\,\mu\mathrm{m}$. With analog interpolation and an assumed signal to noise ratio of $\sim 30 - 40$ this

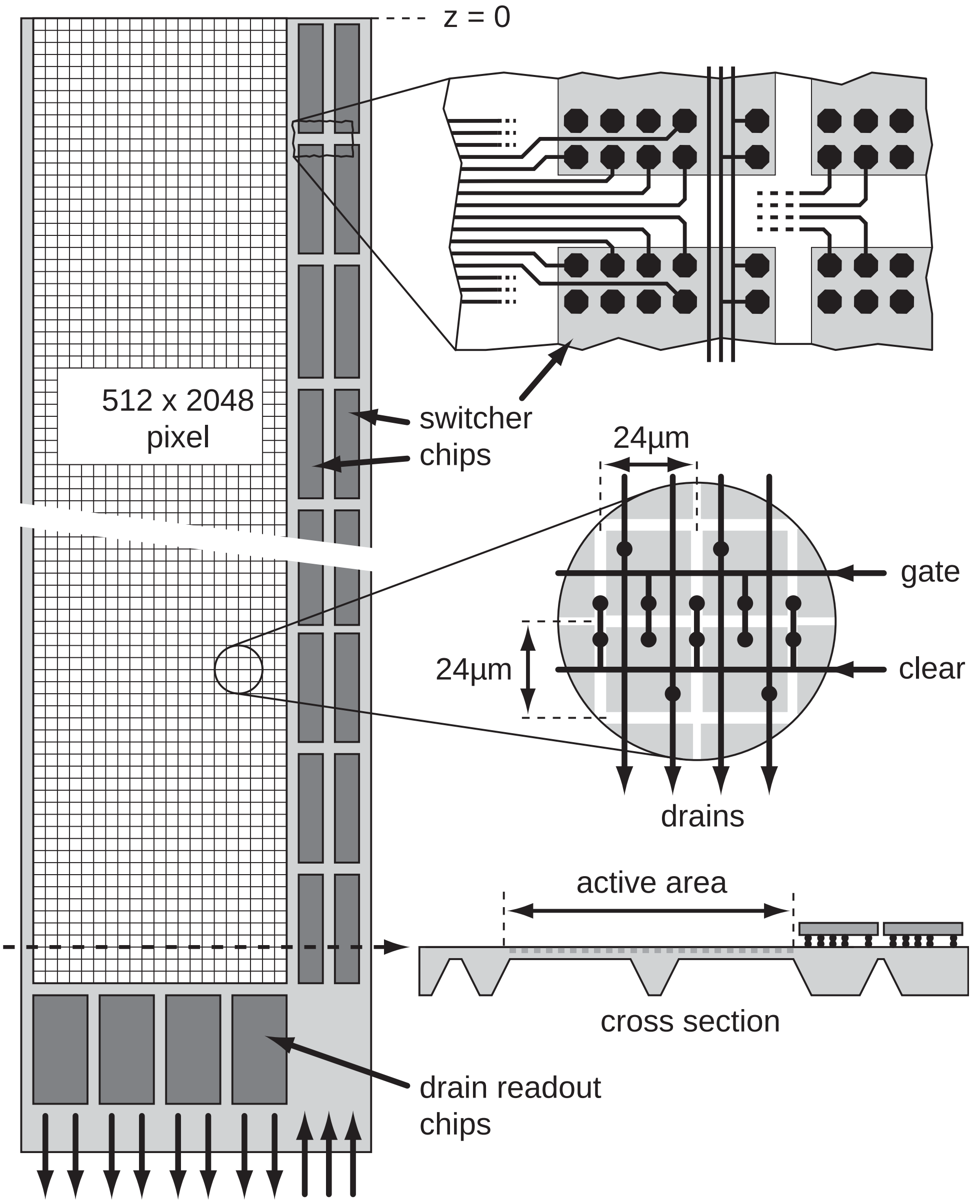


Figure 2.3: *Sketch of an ILC module (half ladder) showing SWITCHER and readout ASICs, the double pixel structure and a cross section [17].*

value can be improved well below the required $4\,\mu$m [17]. As will be shown later the active part of the DEPFET sensor needs to be thinned down to $\approx 50\,\mu$m. An ionizing particle will create a signal (most probable value, eqn. 1.13) of ≈ 3500 electron/hole pairs or roughly 10% of a $450\,\mu$m thick sensor. The noise of the detector is dominated by the readout electronics and not the DEPFET itself [31, 22]. The readout noise is targeted to be on the order of $\approx 100\,e^-$ which would give the above mentioned signal to noise value of $\gtrsim 35$. There are two ways to achieve this goal:

1. Increase the internal gain of the DEPFET g_q. Since the readout electronics is the

dominant noise contributor this translates directly into an improved signal to noise. Current values for ILC pixel types are $g_q \approx 300\,\mathrm{pA/e^-}$, but with changes to the pixel geometry values of $g_q \approx 500\,\mathrm{pA/e^-}$ have been measured

2. Decrease the noise of the readout electronics. This is the topic of ongoing studies including the design and production of a new readout chip (DCD) [32] replacing the CURO readout chip [31]. However, readout speed and power consumption put limits to these efforts. The data used for this thesis was exclusively collected using the CURO readout chip (more below).

Read out scheme and matrix operation

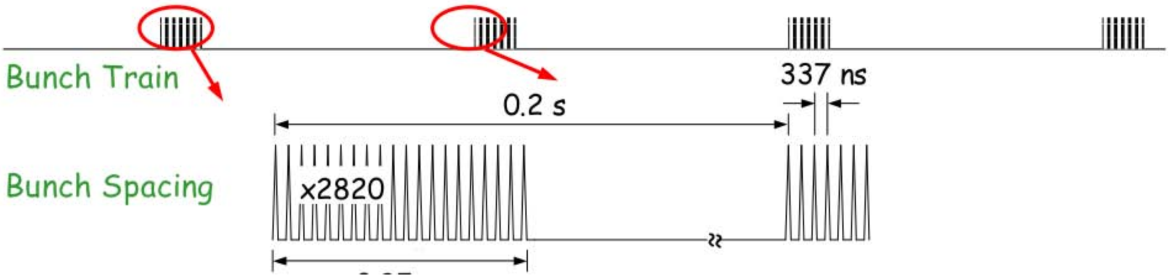


Figure 2.4: *Bunch structure of the TESLA ILC proposal: A bunch train is $\approx 950\,\mu s$ long and consists of 2820 bunch crossings each with a separation of $337\,ns$. Between bunch trains is a $199\,ms$ long break.*

As mentioned above the constrain on the readout for the first layer are severe due to the e^+e^- radiation background. Figure 2.4 shows the time structure of the TESLA ILC accelerator proposal. As the accelerator technology for the ILC will be superconducting the time structure will be close to the TESLA proposal. A readout in between the bunch trains is impossible because the occupancy per pixel in the first layer would be $\approx 8\%$ just with single pixel hits and thus far above the allowed $\approx 1\%$. A frame readout time of $50\,\mu$s would reduce the occupancy by a factor of 20 to $\lesssim 0.5\%$. To achieve this speed the pixels are not read out sequentially but row wise (Fig. 2.5): The (external) gates of all DEPFET pixels in one **row** are connected to a steering ASIC[1] (**SWITCHER**) and the drains of all DEPFET pixels in one **column** are connected to one input channel of the readout chip (**CURO**). The readout chip will then only sample the signals from that row that is switched on by the steering chip. The clear process is similar, all clear contacts of one row are connected to a clear SWITCHER. The volume of raw data generated inside the vertex detector would be too much to be transferred out of the vertex detector. Therefore two things will be implemented in the readout electronics:

1. **Zero suppression**: Signals below a tunable threshold will be discarded. Since hundreds of millions of pixels have to be read out several thousand times a second data reduction immediately after the sampling process is necessary

---

[1] **A**pplication-**S**pecific **I**ntegrated **C**ircuit

2. **In-chip hit storage**: The detected hits will be stored on the readout chip and read out in between the bunch trains. The 0.2 seconds are ample time for data transfer.

The readout is foreseen to be continuous without a trigger. Furthermore, to ease the speed constrains on the readout electronics, two rows are connected to the same switcher channel and are read out at once. This double row concept and the implementation of the matrix operations in the ILC DEPFET prototype system will be discussed below.

Power consumption

As stated above a material limit of $x/X_0 \approx 0.1\%$ means no (active) cooling system can be used inside the vertex detector. However, the power consumption of a DEPFET vertex detector will be very low. The DEPFET sensor itself is the first stage (i.e. the amplifier) of the readout electronics and an active pixel consumes $\approx 500\,\mu$W [17]. A ladder of the inner layer has 2048 active pixels at a time and a resulting power consumption of ≈ 1 W. The switcher ASICs contribute with only 0.85 W since only 2 out of 32 switchers are active and those two have only one row active at a given time [17]. The main power consumer is the readout ASIC. The DCD1[2] needs 5 mW/channel and therefore 10.2 W are needed for all 2048 channels of layer one [17]. The outer layers are slightly wider and longer and produce 21 W per ladder. Table 2.3 gives a summary of these values. Two things are worth noting:

- Power dissipation is largely caused by the readout ASICs. These sit outside the active area and if unforseen developments will drastically increase their power budget, some form of active cooling might still be feasible in the very far forward and backward region of the vertex detector.
- The total power consumption of the entire ILC DEPFET vertex detector as stated in table 2.3 is ≈ 1.3 kW. However, it is foreseen that the detector is operated in a pulsed mode with the majority of the electronics virtually shut down during the 0.2 second long breaks in between the bunch trains (Fig. 2.4). This means that the power dissipation of the entire vertex detector will not exceed a few ten Watts.

layer	ladders	active rows	columns	power/ladder	power/layer
1	8	2	1024	12W	96W
2	8	2	1535	21W	168W
3	12	2	1536	21W	252W
4	16	2	1536	21W	336W
5	20	2	1536	21W	420W
all					1272W

Table 2.3: *Power dissipation per layer. The total power is 1272W, however, with the ILC duty cycle of 1/200 this is reduced to 6.4W [17].*

---

[2] **D**EPFET **C**urrent **D**igitizer

Material budget

With a DEPFET vertex detector operated in a pulsed mode the power dissipation and therefore heat production is low enough to rely on air flow cooling only. However, even without cooling pipes the targeted goal of $x/X_0 \approx 0.1\%$ is very demanding and can only be fullfilled by using thinned down sensors. In the DEPFET concept (the cross section is shown in fig. 2.3) the silicon of the active area is thinned down to $\approx 50\,\mu$m with the area under the steering and readout chips left at $450\,\mu$m silicon. Furthermore a few parts of the wafer will also be left out of the thinning process, thus building a frame giving mechanical stability to the module. The principle feasibility of the thinning process has already been shown with diodes [33]. The switcher ASICs which are also inside the area sensitive to multiple scattering need to be thinned down to $\approx 50\,\mu$m as well. An additional contribution comes from the gold bumps used for the flip chip bump bonding of the ASICs onto the silicon frame. Table 2.4 shows a summary of the contributors to the material budget. With a thinned down sensor the $x/X_0 \approx 0.1\%$ is well in reach.

component	material	X_0 cm	area mm^2	thickness μm	equivalent thickness μm	$\%X_0$
sensor	Si	9.36	13 × 100	50	50	0.05
frame	Si	9.36	2 × 100	450	45	0.05
Switcher	Si	9.36	3 × 100	50	11.5	0.01
gold bumps	Au	0.33			0.46	0.01
all						0.12

Table 2.4: *Material breakdown of the proposed ILC module (innermost layer module). "Equivalent thickness" is the material normalized to the sensitive area of* $13 \times 100\,mm^2$. *The gold bumps (gate and clear per line, service bumps for switcher) have a diameter of* $48\mu m$ *[17].*

2.2 The DEPFET prototype system for the ILC

With the ILC specification in mind the first iteration of a prototype system with a 64 by 128 pixel DEPFET matrix was built in 2005 [34, 35, 36, 37, 38]. Since then the system has been continually improved and tested in both lab and test beams [39, 40, 41]. Furthermore, it has successfully been used in cooperation with the EUDET project [42, 43, 44]. Figure 2.6 shows a picture of the system which consists of two main parts:

- A **hybrid PCB**[3]: This board contains the actual DEPFET sensor matrix, two steering ASICs (SWITCHER), and the readout ASIC (CURO). Furthermore two transimpedance amplifiers convert the currents form the CURO (explained below) to voltages for the ADCs[4] on the DAQ board. The name *hybrid* was chosen since the DEPFET is not a true monolithic active sensor, but needs ASICs to be operated.

---

[3] **P**rinted **C**ircuit **B**oard

[4] **A**nalog-to-**D**igital **C**onverter

- A **DAQ**[5] **PCB**: The DAQ board holds two ADCs, a FPGA[6], a RAM, and a USB to PC communication interface. As already specified by its name, this board is responsible for the data acquisition, i.e. steering the ASICs and (pre)processing incoming data.

In the following the most important components and their interplay with the DEPFET matrix operation will be elucidated.

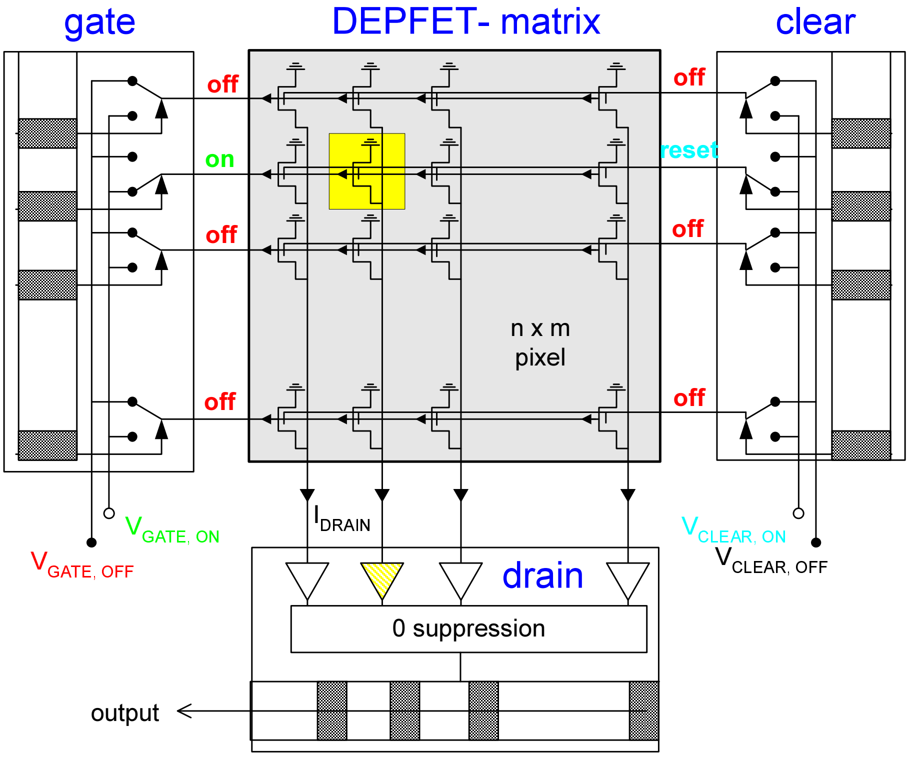


Figure 2.5: *The operation of a DEPFET matrix: The pixels are read out row wise. The gates of all pixels in one* ***row*** *are connected to a channel of a steering ASIC (****GATE SWITCHER****). Similar the clear contacts of this* ***row*** *are connected to another ASIC (****CLEAR SWITCHER****). The drains of each* ***column*** *are connected to a channel of the readout ASIC (****CURO****). The rows are read out sequentially in the following fashion: The according row is switched on by the GATE switcher and the current is sampled by the CURO. Then the CLEAR switcher removes the signal (and leakage current) electrons in this row and the CURO samples the pedestal current of the row.*

DEPFET matrix

The DEPFET matrices used for this thesis are all part of the PXD5 sensor batch. Figure 2.7 shows the wafer layout of this production batch. This is the second of

---

[5] **D**ata **ACq**uisition

[6] **F**ield-**P**rogrammable **G**ate **A**rray

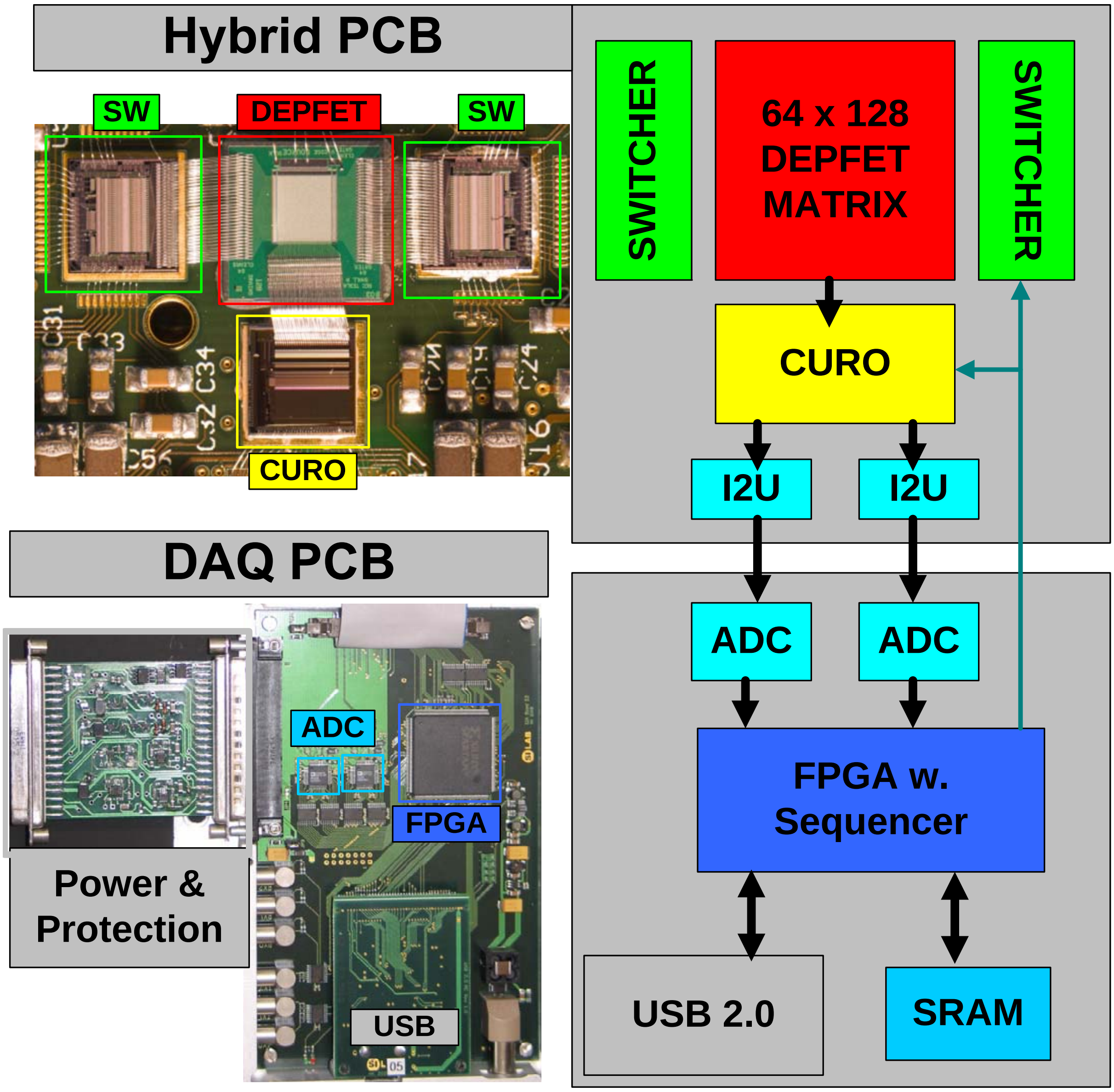


Figure 2.6: *The ”DEPFET for the ILC” prototype system. The upper part shows the detail of the Hybrid PCB with the actual 64 by 128 pixel DEPFET matrix, the two SWITCHERS and the CURO ASIC. Both switcher select the row which is read out in parallel by the CURO (fig. 2.5). The two outputs of the CURO are connected to transimpedance amplifiers, where* I_{sig} *is converted to a voltage. The lower part shows the DAQ board. The two ADCs digitize the voltages from the hybrid and the data is then further processed inside the FPGA. The FPGA is central to the DAQ and holds among other things a sequencer, the control logic for ADCs, ASICS, and RAM, and some preprocessing logic for the incoming data.*

ILC oriented DEPFET pixel design runs. The first run PXD4 has already been operated very successfully both in the laboratory as well as in test beam experiments [34, 35, 36, 37, 38, 39, 40, 41]. Among other things the new PXD5 run contains DEPFET matrices with 64 x 128, 128 x 128, 512 x 512, and 2048 x 128 pixels (length of an ILC

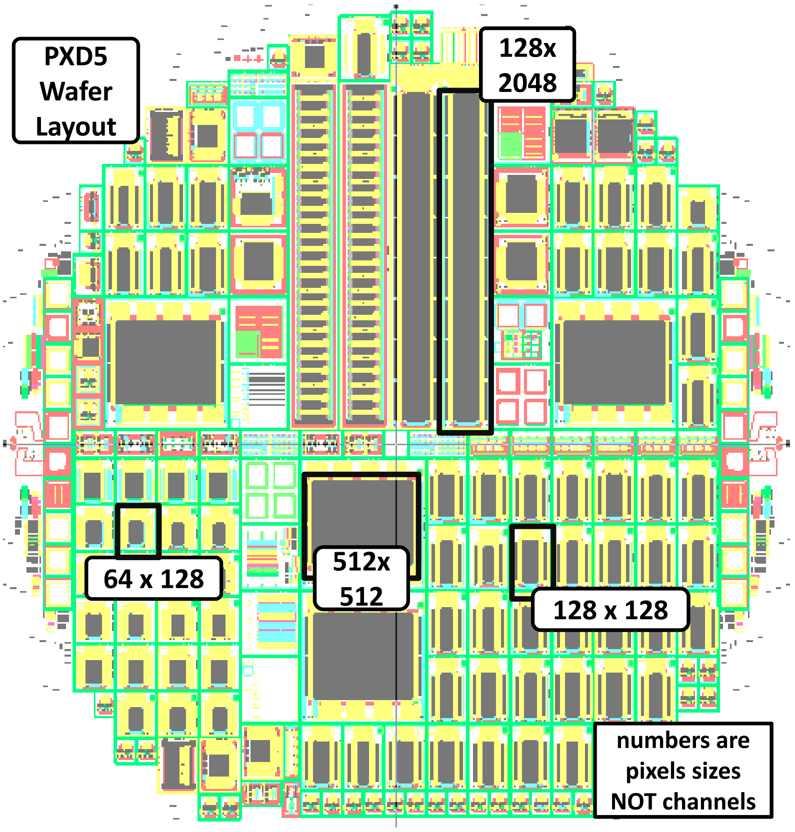


Figure 2.7: *The layout of the PXD5 wafer: This run contains matrices with 64 x 128, 128 x 128, 128 x 2048, and 512 x 512 pixel as indicated in the figure. Other structures are for specific tests, single pixel setups, etc. The data in this thesis is solely gathered with 64 by 128 pixel matrices. However, the newest generation (as of 2010) of DEPFET DAQ and Hybrid boards can handle 128 x 128 pixel matrices.*

module/half ladder). The data in this thesis is solely gathered with 64 by 128 pixel matrices with an older DAQ system. However, the newest generation (as of 2010) of DEPFET DAQ and Hybrid boards can handle 128 x 128 pixel matrices. While there are a few matrices on the PDX5 board which have still the old PXD4 layout ("*REC small*") for comparison measurements, the majority of the pixels have complete new and improved layouts. Some of these layouts will be explained in more detail in the next chapter together with the test beam setup.

Medium and high E implantations: A novelty in the PXD5 production is the so-called medium E implantation. In the last generation of DEPFET pixels, the PXD4 production, a global high energy implantation was applied at the beginning of the manufacturing process. The additional n^+ layer at $z \approx 1.2\,\mu$m depth moves the internal gate deeper to $\approx 1\,\mu$m instead of $\approx 0.6\,\mu$m [18, 45]. This makes the clear process much easier but has the disadvantage of reducing the internal amplification g_q by $\approx 30\%$. For the PXD5 run this high energy implantation was replaced by a medium energy implantation with the idea of finding a compromise between clear efficiency and first stage amplification.

With medium E the n^+ implantation is at a depth of $z \approx 0.6\,\mu$m.

Bias voltages and the sampling sequence

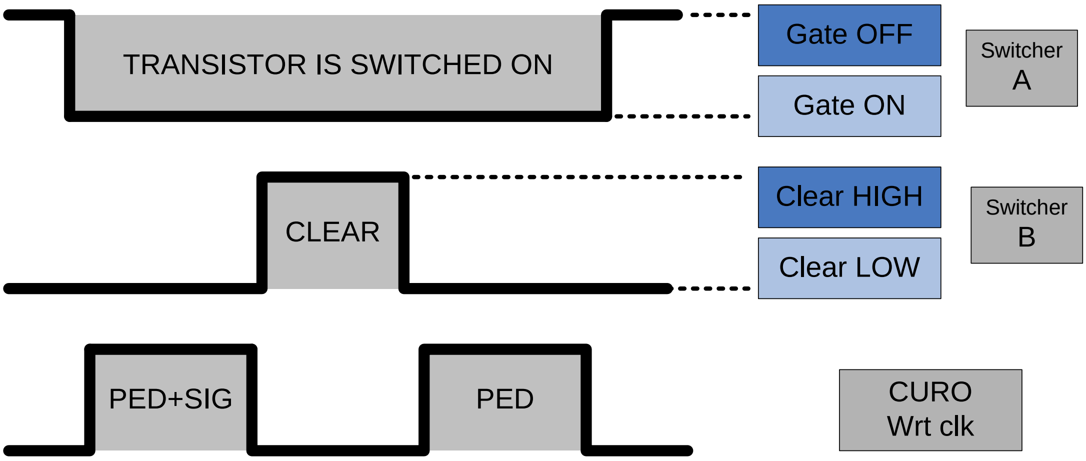


Figure 2.8: *The general sampling sequence for a DEPFET pixel as foreseen for the ILC: The GATE switcher turns all DEPFET transistors in a row on. Then the readout chip samples the signal and the pedestal current $I_{sig} + I_{ped}$. After a clear pulse is applied and the electrons are removed the remaining pedestal current is sampled. Note that in this sampling sequence the leakage current $I_{leakage}$ is part of the signal current I_{sig}. Furthermore the actual sampling sequence for the CURO is somewhat different as will be explained in the CURO description.*

The drain current of DEPFET I_{total} consists not only of a signal I_{sig} but also of a pedestal current I_{ped}. I_{ped} is the current of a DEPFET transistor with a given gate voltage V_{gate}^{ON} and an empty internal gate. In addition there will be a current $I_{leakage}$ caused by electrons from the leakage current integrated since the last clear process. In the ILC concept the current of a row will be sampled twice with a clear pulse in between (Figure 2.8). The first current sampled will thus contain $I_{sig} + I_{ped} + I_{leakage}$ where as the second sample contains only the pedestal I_{ped}. The two dynamic voltages *GATE* and *CLEAR* are switched by the corresponding ASIC between the voltages $Gate_{ON}$ and $Gate_{OFF}$ and $Clear_{HIGH}$ and $Clear_{LOW}$, respectively. The double sampling and the subtraction of pedestal current will be done on-chip inside CURO.

The double pixel and double row structure

The DEPFET pixels have a *double pixel* layout as sketched in figure 2.9. This serves two goals:

- **Double row read out:** To ease the speed constraint on the readout electronics two rows are read out at once, i.e. each SWITCHER channel is connected to two rows. Hence, the drains of one matrix column are alternately connected to two CURO input channels.

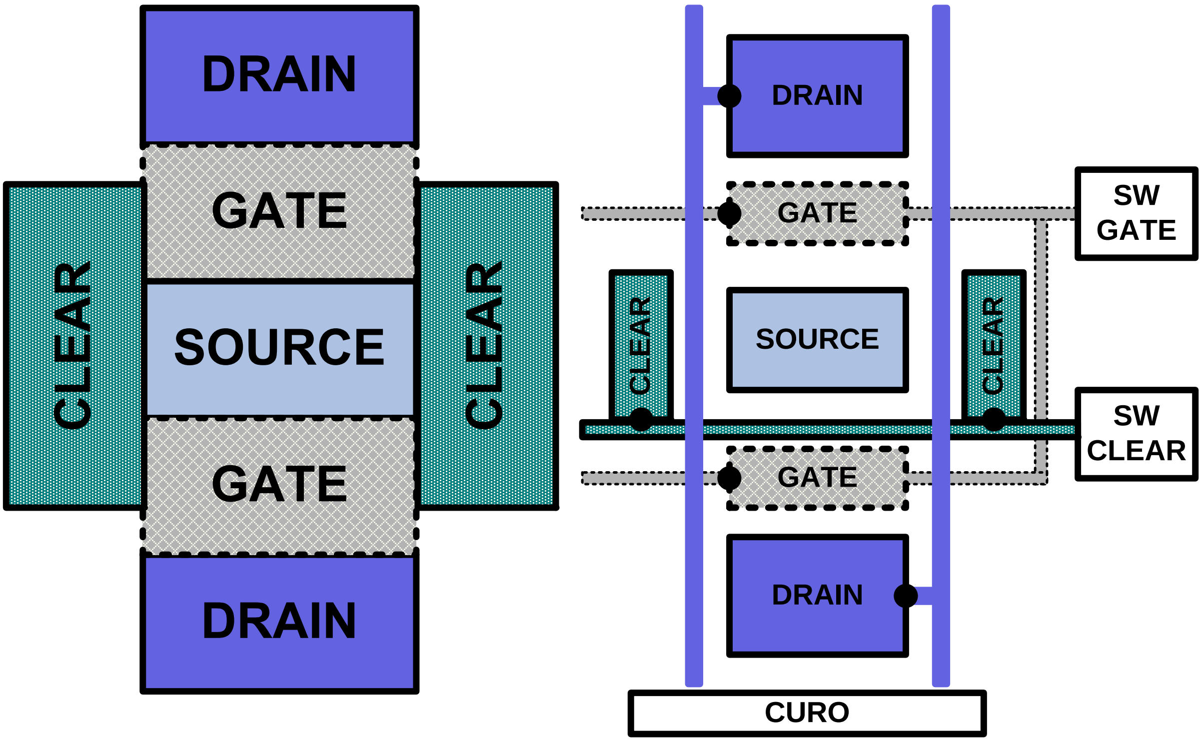


Figure 2.9: *Schematic of a ILC DEPFET double pixel. To increase the readout speed and to reduce the pixel size every SWITCHER channel is connected to two pixels. These pixels share a common source but have two different drains, each connected to a different CURO input channel. This means a* $Gate_{ON}$ *signal from the switcher is applied to a double row, and accordingly the CURO will sample two rows with each sample signal. Hence the readout speed is doubled.*

- **Common source:** The pixel pairs connected to different curo channels share a common source. Thus two pixels are combined to one double pixel cell allowing for a smaller pixel size.

SWITCHER

The **Switcher2** ASIC has been designed in 2002 in 0.8 μm high voltage technology and is able to drive voltage swings of up to 25 V. Each Switcher2 has 2x 64 output channels to drive 64 gate/128 rows. It has a simple on-chip sequencer and can be operated in a daisy chain mode. The chip also allows to bypass the internal sequencer and to steer the channels externally. The ILC prototype system makes exclusive use of the latter option. The range of the voltage swing of the switcher2 is limited by its lower and upper analog supply voltages $AVSS$ and $AVDD$. The digital ground level and the polarity of the switching voltages are adjustable. There are however some caveats when powering up the chip: At no time must any steering voltage (i.e. gate and clear) be outside the boundaries given by $AVSS$ and $AVDD$. To prevent this and other unintended misapplied voltages destroying the chip a *protection board* was added to the prototype system after a few month of operational experience. The Switcher2 was successfully used in all test beam and lab experiments until 2009 when it was replace by the Switcher3. Among other reasons the replacement was necessary, because - unlike Switcher2 - the new Switcher3 is made in a more radiation tolerant technology using thinner gate oxides [32]. A detailed description of the Switcher2 can be found in [46].

CURO

The readout chip CURO 2 (**CU**rrent **R**ead **O**ut) was designed in 2003 [31] and manufactured in a - intrinsically radiation hard - 0.25 μm process. It is specifically targeted at the ILC concept. Figure 2.10 shows an overview. The chip is operated with two signals:

- **Write signal WRT**: The write signal controls the analog current sampling, i.e. when a current is written into current memory cells. The **WRT** is level sensitive: The current $I_{sig}+I_{ped}$ is sampled on logic high into an auxiliary storage cell, whereas I_{ped} is sampled on logic low. During the low phase the I_{ped} is subtracted from $I_{sig}+I_{ped}$ and the remaining signal is stored in one of two alternating buffer cells. With the next **WRT** the content of the current buffer is written to an analog FIFO. The following **WRT** will then override the corresponding current buffer cell, but only in the low phase. During the high phase of the third write clock cycle and while $I_{sig}+I_{ped}$ are stored in the auxiliary buffer, the current of the alternating buffer is compared to a channel specific threshold and the result is written to the digital counterpart of the analog FIFO (on-chip zero suppression). Unfortunately this write clock operation scheme has the disadvantage of neglecting the time needed to remove the signals from the DEPFET cell, i.e. it ignores the duration of the clear pulse. This leads to the problem that the pixel is cleared *during* the I_{ped} sampling (**WRT** is low) and therefore the sampling procedure and correspondingly the readout speed must be much slower than anticipated (see [22] for details).

- **Scan signal SCN**: The scan signal controls the hit scanner, a fast binary back propagating tree search that writes up to two hits per **SCN** including their addresses into a designated on-chip hit ram. If there are more than two hits in one FIFO row correspondingly more **SCN** cycles are needed to process the hits. Therefore for the ILC the **SCN** is targeted to work at a higher speed (40 MHz) than the **WRT** (20 MHz) and in addition a derandomizing FIFO with a depth of up to eight rows is foreseen. Simulations have shown that this combination is sufficient to handle the expected inner layer occupancy of a DEPFET ILC vertex detector without any dead time [31]. However, in the CURO 2 chip the analog FIFO was only implemented as a proof of principle with a depth of only one row. Further the CURO 2 possesses two current outputs. If with a scan clock one or two hits are found the currents stored in the according analog FIFO cells are multiplexed out. Although it has been shown that the zero suppression works in principle [22], the CURO is usually only operated in a fully analog, non zero-suppressed mode, because having access to the full analog information is necessary to study the various DEPFET types in detail. This is achieved by writing a "all hits" test pattern to the hit scanner and ignoring the results of the comparators. With 64 **SCN** cycles the currents of one row are then multiplexed out. Naturally the fully analog readout mode is much slower than

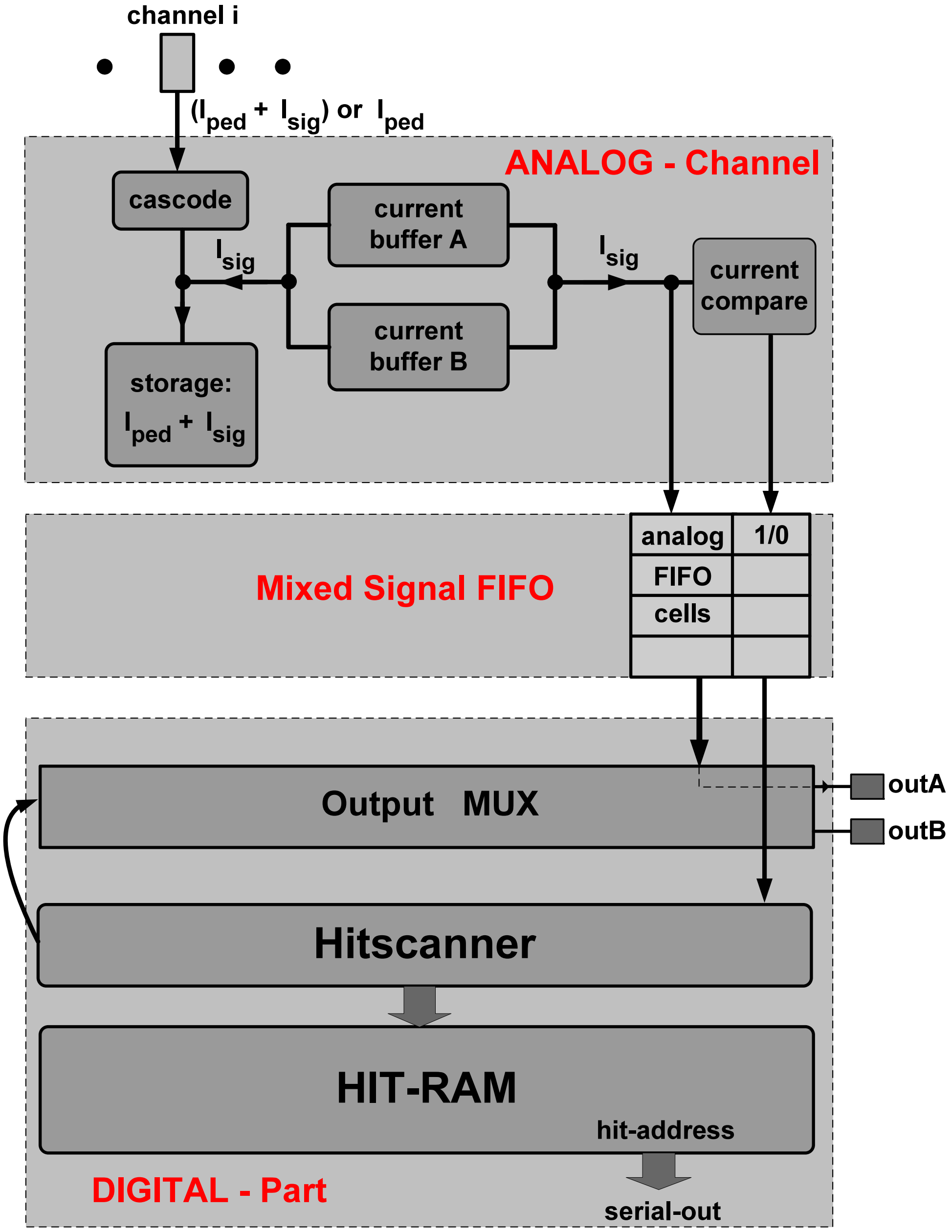


Figure 2.10: *Schematic overview of the CURO building blocks. The analog part of the chip works entirely with current signals. For clarity only one out of 128 input channels is shown. It contains a regulated cascode and a pedestal subtraction unit. Furthermore two alternate current buffers help increase the readout speed. A mixed analog signal and digital hit information FIFO stores the current and a comparator results for each input channel. The depth of the analog part of the mixed signal FIFO is only one for CURO 2, which limits the zero suppression capabilities of the ASIC. The digital part contains a hit finder with parallel tree search structure, a HIT RAM, and a output multiplexer. Extensive details on the chip and its performance can be found in [31, 22].*

the non zero-suppressed mode with the main time consumed by multiplexing each **SCN** two I_{sig} currents out.

A thorough study of the CURO2 performance with and without zero suppression can be found in [22]. In the final ILC design the two signals, **SCN** and **WRT**, are planned to run as a continuous clock, i.e. the system is operated in a triggerless readout mode.

Data acquisition chain

There are two transimpedance amplifiers (AD8015) on the *Hybrid PCB*, one for each analog CURO output where the single ended current from the CURO is converted to a differential voltage signal. To minimize stray capacitance they are in die (non package) form and are - like the DEPFET matrix and the ASICs - wire bonded. The hybrid PCB is connected to a DAQ board with a flat ribbon cable. On the DAQ board are two differential 14 bit ADCs, one for each CURO output channel. Their digital output is fed into a Spartan 3 FPGA which houses the main DAQ control unit, a programmable sequencer, as well as logic for the initialization and configuration of the ASICs, clock management, ADC and SRAM control, processing of CURO hit data, and handling of trigger and busy signal for test beam applications (details in the next chapter). The sequencer is programmed in such a fashion that enables the DAQ to be triggered externally while running in a so-called *rolling shutter* mode. The digitized signals from the ADCs are stored inside a 256 kB SRAM onboard of the DAQ PCB. This is enough memory for 16 consecutive, fully read out matrix frames. Once the memory is full, data taking is paused and the data is transferred to a PC via a USB interface.

Rolling shutter mode:

Since the system was designed as an ILC prototype system it was also designed to be run without a trigger. However, for laser measurements in the laboratory as well as for test beam measurements a triggered readout is desired. Therefore, a triggered, quasi continuous readout scheme, the *rolling shutter* mode was implemented. Figure 2.11 shows a simplified version of the rolling shutter sequence:

- Initially the system waits for a trigger input while continuously clearing the matrix. Without perpetually clearing the matrix the electrons generated by the leakage current would flood the internal gate rendering it insensitive. In this sequencer loop the time between to clears is small (7.7μs) and the DEPFET pixels can be seen void of any signal (and leakage current) electrons.

- When a trigger signal arrives the sequencer jumps to the actual readout sequence. Each row is sampled according to the specification for a non zero-suppressing CURO (i.e., one slow **WRT** signal for sampling and two faster **WRT** pulses for transferring the currents to the FIFO). Then the values for this row which are now stored in the analog FIFO must be multiplexed out of the CURO with 64 **SCN** pulses. Then the next row is processed.

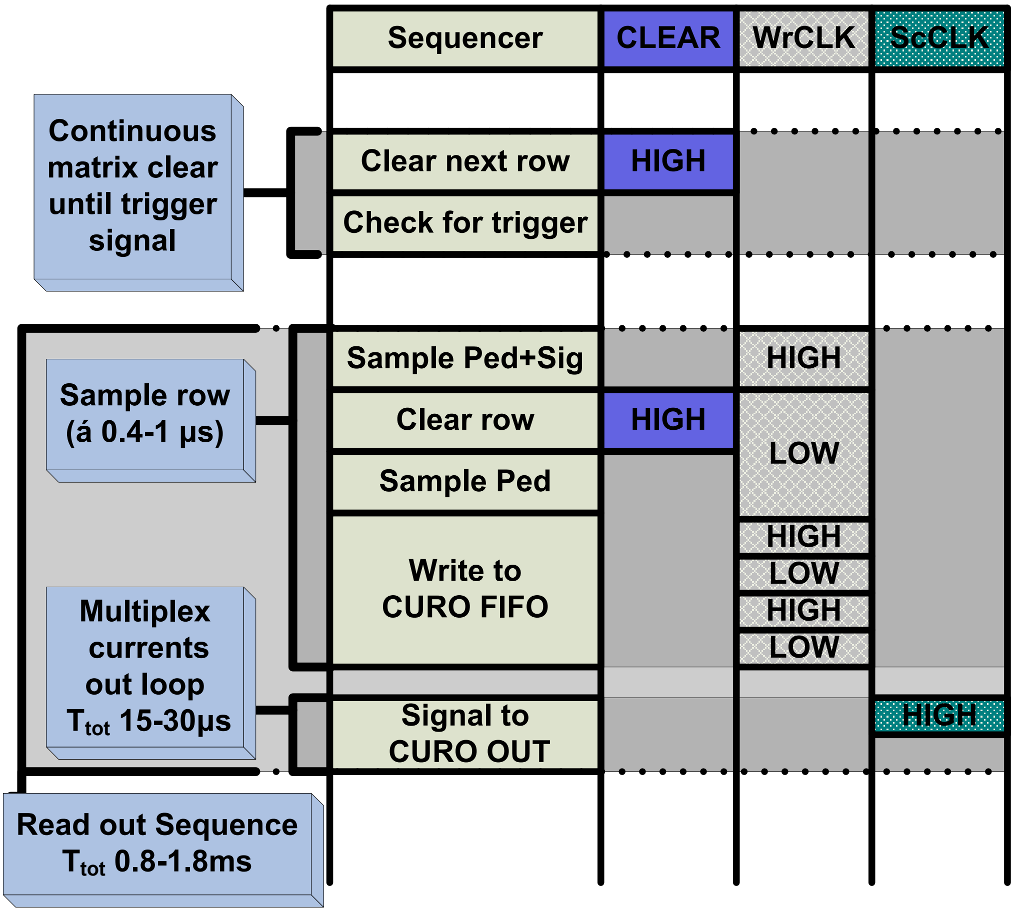


Figure 2.11: *A simplified version of a DEPFET readout program for the sequencer implemented inside the FPGA of the DAQ board. The four columns denote the general activity of the sequencer, the CLEAR pulse (actually a switcher clock), and the write and scan clock of the CURO, the brackets and the according text boxes explain the main sections of the program. While the system is waiting for a readout trigger the matrix is continuously cleared (top most loop). If a trigger is signaled the sequencer jumps to the readout sequence where each row is sampled and than multiplexed out with 64* ***SCN*** *cycles. Since there are 64 double rows this has to happen 64 times. The total time it takes to read out an entire matrix depends on the actual programming and is a compromise between readout speed and signal to noise.*

- When the matrix has been read out completely the sequencer jumps to the beginning of the program thus starting with the continuous clearing of the matrix. If however the RAM on the DAQ board is full, i.e. after 16 triggers, the sequencer is stopped and data transfer to a PC is commenced.

The length of the readout depends on the actual sequence, e.g. on the duration of the CURO current sample, as is indicated in figure 2.11.

name	description	[V]
	SWITCHER 1	
AVDD	analog supply	23
AVSS	analog supply	-0.5
VA HI	Clear HIGH	22
VA LO	Clear LOW	7 - 9
VB HI/LO	common clear gate	6- 8
DVDD*	digital supply	5
	sensor static	
BULK		17
SOURCE		7
BP	backplane depletion	-200

* voltage can be produced by protection board

name	description	[V]
	SWITCHER 2	
AVDD	analog supply	22
AVSS	Analog supply GND	-0.5
VA HI	Gate OFF	15
VA LO	Gate ON	3
VB HI/LO	common clear gate	
DVDD*	digital supply	5
	CURO2	
AVDD*	analog supply	2.5
DVDD*	digital supply	2.5
	Protecion Board	
DVDD_S1		5.5

Figure 2.12: *This is an overview over all voltages needed to operate one DEPFET-ILC prototype system. All voltages are referenced to ground GND.*

Powering scheme

As has been mentioned before the ASICs are sensitive to wrong biasing. Furthermore, a rather large number of voltages is needed to operate the DEPFET prototype system for the ILC vertex detector. Therefore a designated add-on board for the DAQ system has been issued to protect the ASICs from misapplied voltages and to reduce the number of required power supplies. Figure 2.12 shows a table of all voltages needed to operate one DEPFET-ILC prototype system. The protection board clamps the analog steering voltages (GATE and CLEAR) between the analog supply voltages of the SWITCHER2 ($AVSS$ and $AVDD$). Although it is possible to generate some of the supply voltages on the protection board this option is usually dismissed, because unlike most laboratory power supplies the protection board possesses no current limit settings. Setting the right current limit can be crucial for operating the DEPFET system especially for the start up proceedings. When operating multiple DEPFET systems the amount of lab power supplies and the complexity of the system can, especially in test beam environments, be very demanding. A test beam, conducted in 2007, did not produce data partly because of this. As a result of these problems and the experience of the test beams of 2007 and 2008 (see next chapter) a dedicated power supply for the DEPFET ILC prototype system was developed in 2009 [21]. For the test beam in 2008 this power supply however came unfortunately to late.

3

Test Beam Experiment and Analysis

3.1 The Experimental Setup

As has been mentioned in the last chapter, the DEPFET ILC prototype system has been successfully operated in test beam at both CERN[1] and DESY[2]. However, until 2008 none of the PXD5 production matrices had been studied in the test beam[3]. Furthermore no test beam measurements of the DEPFET sensor characteristics were available at that time. Therefore a high resolution test beam campaign with large statistics and PXD5 generation DEPFET pixels was conducted in the summer of 2008. To achieve high resolution the multiple scattering had to be reduced. The SPS[4] facility at CERN near Geneva can deliver 120 GeV pions, which made it the best option for the test beam.

Pixel Types

The PXD5 production features both old (PXD4) as well as completely new designs. The new developments were mainly driven by the demand for smaller (ILC-like) pixel sizes and an improved clear behavior without compromising the internal amplification g_q. Studies had shown that for the latter goal a dynamic *CLEAR-GATE* with a minimal voltage swing enhanced the clear behavior considerably. Therefore among several other test structures these pixel types were realized on the new PXD5 wafers:

- **Rectangular Small** (**REC small**) This is a PXD4 design that was included into the PXD5 run to allow for direct comparison of the different production batch runs. It features a common (i.e. static) clear gate and has a pixel size of 33.0 x 23.75 μm^2.
- **Rectangular Small A** (**REC small A**) Like *REC small* but with a slightly modified pixel size of 36.0 x 22.0 μm^2.

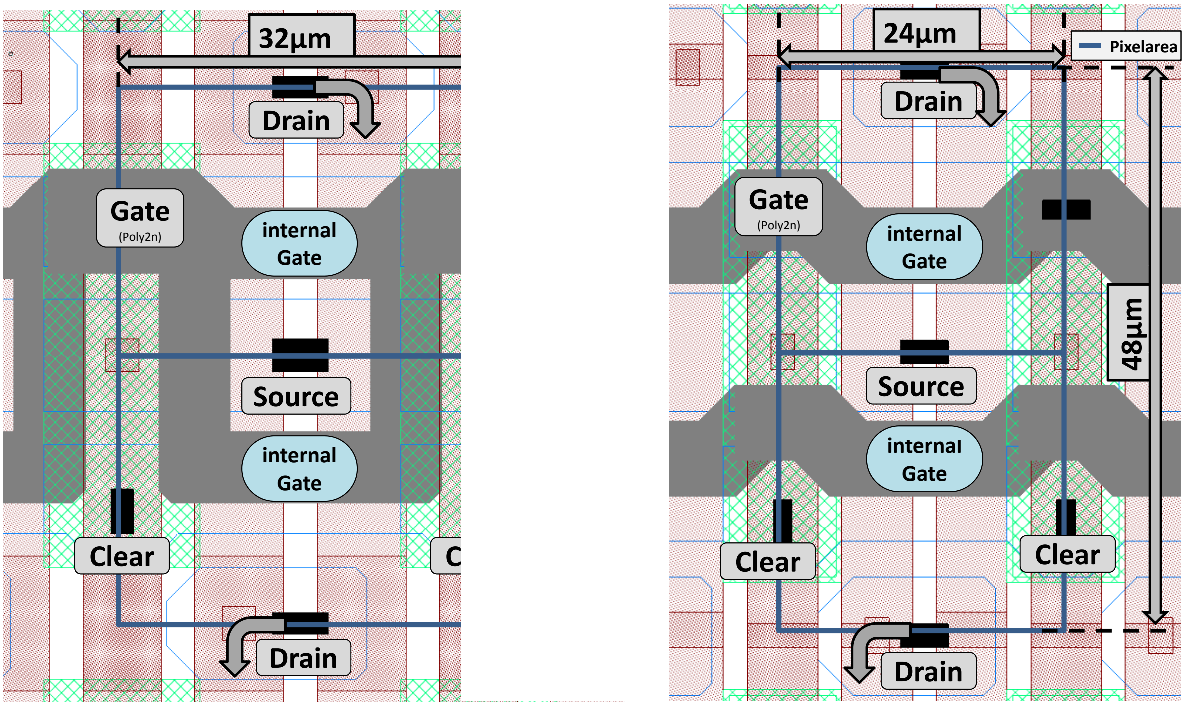
---

[1] Originally: *Conseil Européen pour la Recherche Nucléaire*, now *European Organization for Nuclear Research*

[2] *Deutsches Elektronen Synchrotron*

[3] PXD5 sensor had been studied in a test beam in 2007. However, the data gathered there did not meet the requirements necessary for a sensor evaluation.

[4] Super Proton Synchrotron



(a) CoCG-Large layout (b) CoCG-Small layout

Figure 3.1: *The pixel layouts of (a) the Common Clear-Gate Large (CoCG-L) and (b) Common Clear-Gate Small (CoCG-S) DEPFET pixel designs. The black rectangles indicate the positions of source, drain, and clear. The grey area is the poly silicon of the external gate with the accordingly named ellipses marking the approximate position of the internal gates.*

- **Common Clear Gate Large** (**CoCG-L**) This is a new, improved PXD5-version of the *REC small* layout with a pixel size of 32.0 x 24.0 μm^2 (Figure 3.1).

- **Common Clear Gate Small** (**CoCG-S**) This is the smaller version of *CoCG-L* with a reduced size of 24.0 x 24.0 μm^2. This pixel size is already fulfilling the ILC requirements and the *CoCG-S* design option is the baseline of the PXD5 production run (Figure 3.1).

- **Common Clear Gate Very Small** (**CoCG-VS**)] This is the smallest version of *CoCG* with 20.0 x 20.0 μm^2. This design is at the technological limit of the DEPFET production technology.

- **Simultaneous Clear** (**SIMC**) A new pixel layout where *CLEAR* and *CLEAR-GATE* are applied simultaneous via the same SWITCHER channel. In this case *CLEAR* and *CLEAR-GATE* are hard wired and the required potential difference is adjusted by an additional p-implantation.

- **Capacitative Coupled Clear Gate** (**C3G**) Similar to the SIMC layout, this design features a *CLEAR-GATE* that is clocked by the *CLEAR* signal, however, with a capacitative coupling.

Due to technological difficulties not all of these structures were available in 64 x 128 sized matrices. Furthermore the yield of the production was lower than expected leaving only somewhat "problematic" matrices left for the test beam (more on this below). Before the actual test beam was commenced all sensor candidates were tested in the lab and the bias voltages ($CLEAR$, $GATE$, $CLEAR - GATE$) for an optimal signal to noise were measured.

The beam telescope

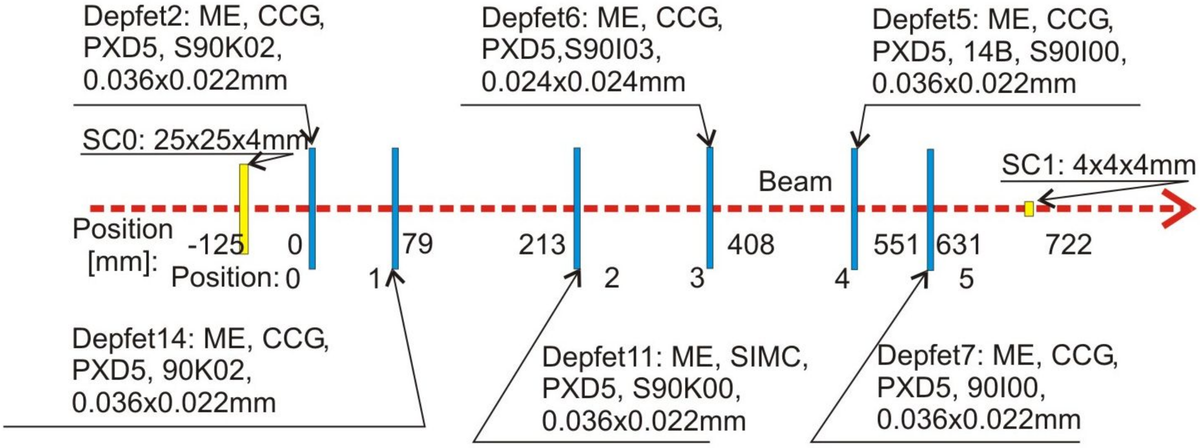


Figure 3.2: *The geometric setup of the test beam 2008 experiment at the SPS at CERN: Six DEPFET sensor planes are triggered by the coincidence of two scintillator triggers SC0 and SC1 with a total trigger area overlap of* 4 *x* $4\,mm^2$.

The goal of a test beam campaign is position sensitive data of the detector's response with respect to an impinging ionizing particle. This includes the spatial position of the particle on the *device under test* (**DUT**). The task of a test beam telescope is to measure space points of a particle from which a designated *alignment and tracking software* will calculate the trajectory of the particle and predict the according entry point on the DUT. Such a test beam telescope is usually made of several sensor planes. They are placed inside the particle beam before and after one or several DUTs. Furthermore the telescope system needs a read out *trigger* signal and (usually) a device *busy* signal.

The *trigger* signal is a coincidence of two scintillators placed before and behind the telescope. The *busy* signal indicates that at least one sensor (telescope or DUT) is busy (e.g. with read out) and unable to respond to a trigger read out request. A central *trigger logic unit* (TLU) blocks incoming trigger signals as long as any device is sending a busy signal. In addition the TLU used in the DEPFET test beam, which is also used in EUDET project [47], sends a unique event number for each trigger to all devices. This increases the data taking stability dramatically. Figure 3.2 shows a schematic of the 2008 test beam setup and table 3.1 lists the used sensors. The two central planes are designated DUTs whereas the outer two pairs were foreseen as telescope planes. Since the *CoCG-L* pixel type showed the most stable and reliable performance it was chosen for the four telescope planes. The two DUTs are of *CoCG small* and of *SIMC* type.

abs. position [μm]	0	79	213	408	551	630
rel. to center [μm]	-315	- 236	-102	93	220	315
module/sensor ID	2	14	11	6	5	7
Wafer ID	S90K02	90K02	S90K00	S90I03	S90I00	90I00
pixel type	CoCG-L	CoCG-L	SIMC	CoCG-S	CoCG-L	CoCG-L
pixel size [μm^2]	32x24	32x24	32x24	24x24	32x24	32x24

Table 3.1: *A table of all DEPFET sensors involved in the test beam 2008 campaign. The wafer Id is referring to the position inside the PXD5 wafer. The pixel types are: CoCG = Common Clear Gate, SIMC = simultaneous clear, and the letters L/S are denoting large and the small pixel types with the according pixel dimensions shown in the row below.*

Performance Issues

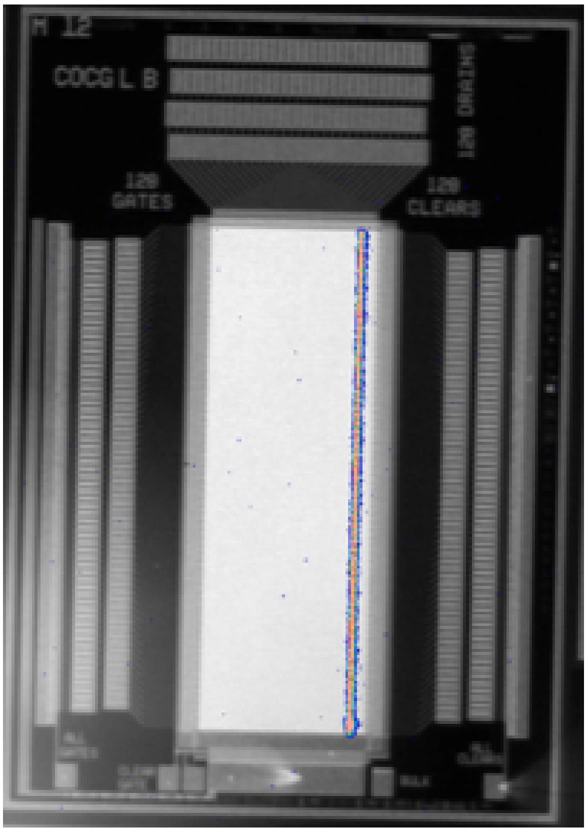


Figure 3.3: *A photograph of the heat emission of a defect DEPFET matrix column taken at the Munich Halbleiterlabor with a Si-CCD cooled to −50° C. This is caused by a short between a drain line to a clear line with a potential of 20 V. During operation this would basically set the input of the CURO (default ≈ 2 V) to the CLEAR HIGH potential and likely destroy one or more ASICs. This is just one example of possible defects which impeded the selection process for possible telescope sensor candidates.*

Although the DEPFET telescope system showed a rather satisfying behavior in the 2008 CERN test beam campaign, some issues remained:

- **Damaged sensors and/or ASICs**: Either the DEPFET matrix itself or some of the ASICs showed some defect. Figure 3.3 shows such a defect. In the 2008 test

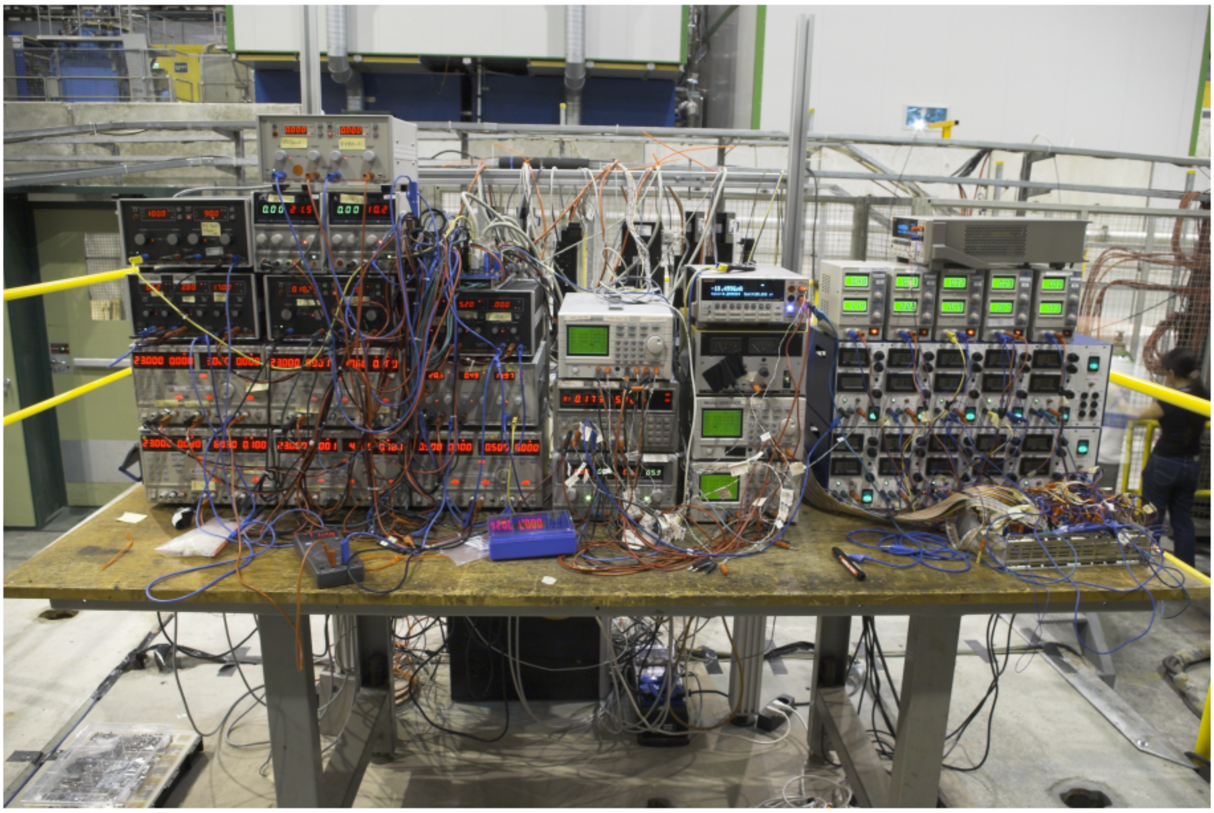


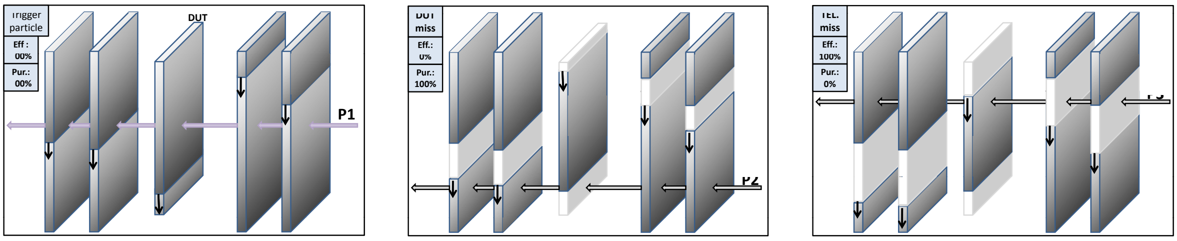
Figure 3.4: *This picture was taken in 2008 at the SPS facility at CERN and shows the arrangement of power supplies needed for the test beam. Despite the obviously large amount of power supply output channels several voltages had to be hooked-up to one channel.*

device type	size in X	size in Y	area	fill factor
Trigger	4 mm	4 mm	$16\,\text{mm}^2$	
CoCG Large	$2048\,\mu\text{m}$	$3072\,\mu\text{m}$	$6.29\,\text{mm}^2$	39.3%
CoCG Small	$1536\,\mu\text{m}$	$3072\,\mu\text{m}$	$4.72\,\text{mm}^2$	20.5%

Table 3.2: *The dimensions of the 64 x 128 CoCG Large and CoCG Small pixel matrices compared with the scintillator trigger overlap. The low fill factor means that, depending on the pixel type, 60 to 80% of all events are triggered by particles NOT interacting with the sensor.*

beam defects were observed on all modules. This caused a need for masking bad sensor areas and in turn reduced the amount of useful data.

- **Powering Scheme**: As has been mentioned in the last chapter the DEPFET system is somewhat sensitive to wrong biasing, both the sensor itself and the accompanying chips. Most modules are very sensitive on the *Common Clear Gate* and the $Clear_{LOW}$ potential, where even a small offset of a few 100 mV can corrupt the properties of the sensor. Furthermore, both the initialization process as well as the stable operation of such a telescope system depends on proper current limits.



(a) Trigger particle (b) second particle (c) third particle

Figure 3.5: *Illustration of an unsynchronized rolling shutter mode with integrating DEPFET devices and its effect on efficiency and purity measurements. 3.5(a) A particle $P1$ triggers the system and all device start the readout but at different rows (vertical black arrows). Since the active area is initially at 100% particle is detected in all sensor planes. 3.5(b) If a particle $P2$ traverses the active area of the telescope planes but the already read out part of the DUT the particle will be identified as a miss and lower the efficiency for the DUT. 3.5(c) If a particle $P3$ enters sensitive area only for some telescope modules it will be discarded as a track and a hits the active area of the DUT it will be still as a fake entry.*

Current limits help protect the system in case of an unforeseen malfunction.

Figure 3.4 shows the actual setup of lab power supplies at the SPS test beam. Despite the obviously large amount of power supply output channels, voltages for different modules had to be bundled and hooked-up to one output channel. This had two unintended consequences: first, the current limits were not set properly causing some minor damage and the loss of beam time. However, since the total test beam campaign lasted almost two month the system could be run in a rather stable mode for several weeks. The second problem was, that some very sensitive voltages like *Common Clear Gate* could not be set properly for all modules, leaving some modules in a rather dire state. Especially module 5 and 7 but partially also module 11 suffered from this problem. This will play a major role for the in-pixel studies in section 3.4.

- **Scintillator Trigger Overlap**: The trigger overlap is $16\,\mathrm{mm}^2$ but the DEPFET matrices have areas of $\sim 6.3\,mm^2$ (CoCG-L) and $\sim 4.7\,mm^2$ (CoCG-S), respectively. With the according fill factor only 20% to 40% of all events are triggered by a particle that is actually passing through the sensor. If the DEPFET would not be an integrating device this would mean that 60% to 80% of all events were empty events. However, due to the integrating nature of the DEPFET device also particles unrelated to the trigger will leave signals in the sensor.

- **Unsynchronized rolling shutter mode**: One of the standard measurements for a sensor is the *efficiency* and *purity* of particle detection. Efficiency is the ratio between the number of detected hits and the number of actual, real hits $\epsilon_{eff} = N_{hit}/N_{all}$ and purity is related to the number of fake hits N_{fake} with $p_{pure} = 1 - (N_{fake}/N_{hits})$. This means, a detector with 100 % efficiency detects every particle and in a detector with 100 % purity every detector hit is a real hit, i.e. there are no fake hits like e.g. noise hits. A beam telescope has to provide information about all real hits N_{hits}. However, in the case of an integrating DEPFET telescope in unsynchronized rolling shutter mode these measurements are extremely difficult to undertake. Figure 3.5 illustrates the effects of an unsynchronized rolling shutter mode on efficiency and purity measurements. Lets assume for a moment that the efficiency and purity of all devices is 100%.

If a particle $P1$ triggers the system all devices will start the readout at current row of the rolling shutter mode described in the last chapter. However, due to little offsets in the clocks etc., the rows numbers of the DEPFETs will not be synchronized and the readout will start at a different row for each sensor plane. For the particle $P1$ that triggers the system this has virtually no consequences 3.5(a). Since the active area is initially 100% for all detectors the particle is detected in the DUT with an according track from the telescope system. Thus efficiency and purity are 100%.

The entire read out takes roughly 0.8 ms to 1.6 ms time (depending on the sequence). During this time the active area of each sensor will get continually smaller while the dead part (i.e. all the rows that have already been read out) is growing. If for example a particle $P2$ traverses the active area of the telescope planes it will be produced a track. However, if the same particle passes through the already read out (i.e. dead) part of the DUT, the particle will be identified as a miss and lower the efficiency for the DUT (fig. 3.5(b)). Since this track has been detected by all telescope planes, it is impossible to tell whether particle $P1$ or $P2$ has triggered the system.

On the other hand, if a particle $P3$ enters sensitive area only for some telescope modules it will be discarded as a track. If, however, the same particle hits the active area of the DUT, it will be still counted as a regular, but fake entry and the purity will be lower (fig. 3.5(c)).

Since the telescope system and the DUTs were all run in the rolling shutter mode, the in-pixel studies presented in section 3.4 only focus on the signal homogeneity and its effect on the position reconstruction. However, older measurements performed with the Bonn ATLAS Telescope (BAT, for details see [48]), a double sided silicon strip telescope, showed an efficiency of 99.3% with a purity of more than 96.3% for the DEPFET sensor[5][38].

- **Corrupt start gate/first rows**: There are some issues with the CURO read out chip while running a rolling shutter mode. While the system is in the continuous clear mode (i.e. waiting for a trigger signal) the CURO is idle. With the trigger signal the CURO is activated. The first rows of the frame[6] usually show a defective performance with the effect quickly (i.e. after 2-4 rows) fading away. Details on this issue can be found in [22].

.

---

[5] A seed cut of five times the pixel noise was applied.

[6] Frame means one completely read out of a DEPFET matrix.

3.2 Data Analysis

This section describes the data analysis steps performed before tracking: pedestal and common mode offset correction, gain corrections, masking, and clustering:

- Pedestal and Common Mode: Besides the actually signal generated by an ionizing particle (MIP) S_{MIP} the signal consists of a fixed S_{ped} and a variable offset S_{CM}

  $$S_{tot} = S_{ped} + S_{CM} + S_{MIP} \tag{3.1}$$

  These two components must therefore be removed.

- Gain Corrections: Each row, column, or even pixel can have its own gain. The gain corrections equalize these variations (at least on a row and column scale).

- Masking: Broken or noisy pixels, rows, or columns need to be identified and masked out.

- Clustering: The charge generated by a MIP diffuses in the sensor surface plane while drifting in the orthogonal direction towards the internal gate and therefore has a certain spatial extend. Clustering means first to look for so-called *seed* pixels where MIPs actually went through and then secondly to associate an according area (e.g. 5 x 5 pixels) called *cluster* to this seed pixel.

A general note on the units of deposited energy: In most cases the signal of the DEPFET will be reported in *ADC units* or *ADUs*. A conversion to electron/hole pairs can be found in table 3.3 based on a Landau fit to the signal distribution. However, in most cases only relative energy changes inside a sensor or pixel are relevant for this chapter.

Pedestal Correction

Pedestal or fix pattern correction is the first step in the test beam data processing. It is an offset of a pixel towards a zero baseline (e.g. without any signal) and is stable over a long time period (e.g. hours or days), in contrast to a common mode offset that might vary with every event.

The pedestal $\mu_{ped}^{x,y}$ is calculated pixel wise as the arithmetic mean of a sample of signals without a contribution from a charge generating particle. In addition a noise figure $\sigma_{ped}^{x,y}$ is a calculated as the RMS value of the pedestal data sample. There are two possibilities to ensure, that the pedestals samples are not contaminated with signals from a particle. One can either take designated pedestal data without a particle beam, or otherwise the signals from particles must be removed. For the DEPFET test beam the latter was done using an iterative approach: In a first step a coarse pedestal and noise value for each pixel is gathered, then in a second step signals outside a $\mu_{ped}^{x,y} \pm 3 \cdot \sigma_{ped}^{x,y}$ band are removed from the pedestal calculation. Figure 3.6 shows the results of the pedestal correction for module 6 whereas figure 3.7 shows the remaining pedestals after the correction for all six modules. The residual pedestals after correction are $\lesssim 1\,\mathrm{ADU}$. This is negligible compared to the cluster signals of $\sim 1000 - 2000\,\mathrm{ADUs}$ (table 3.3).

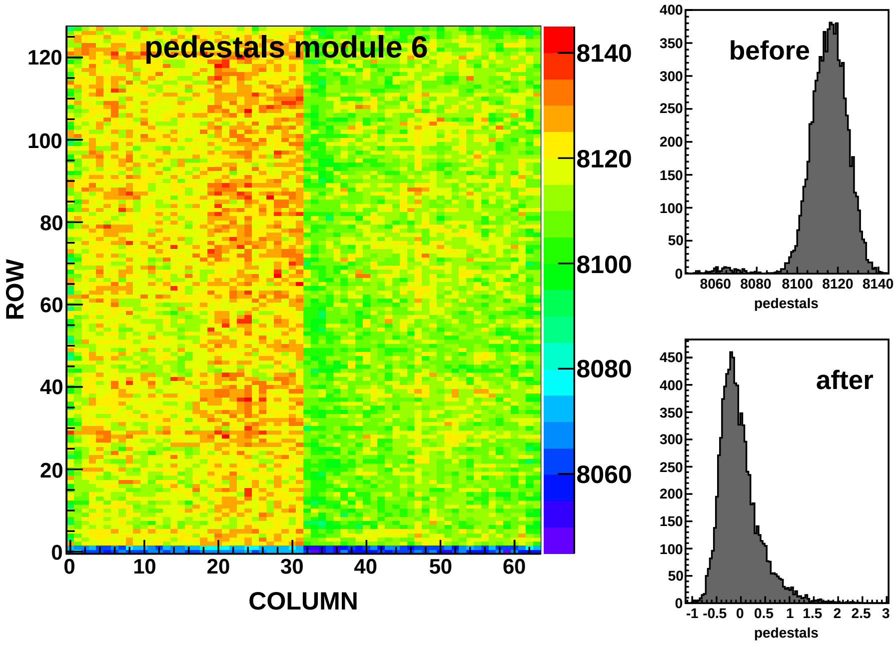


Figure 3.6: *Pedestal offsets for module 6: The left image shows the pedestal offset map, the upper plot on the right shows the pedestal offset distribution and the lower plot on the right shows the residual offset distribution after pedestal correction and common mode correction. The sensor is split by the two readout chains in two halves. This causes the different pedestal values in each sensor half.*

Common mode correction

The common mode correction cancels short time offset fluctuations of a subset of pixels. In the case of the DEPFET it is the common shift of the pedestal currents of the matrix (double) row which is read out entirely at once. A first approach would be subtracting the arithmetic average of a row signal. Yet in case of a particle passing through this sensor area this will lead to additional signal charge and the common mode will be overcompensated for. As has been shown by [20], for DEPFET sensors it is not sufficient to reject pixels on a noise based cut. Therefore an iterative approach was chosen to correct for common mode:

1. In a first step a coarse gate-wise common mode value is calculated using all pixels of a double row and combining both ADC read out chains. These values are then temporary applied to the current event frame.

2. In a second step a pre-clustering is done to block a 5 × 5 pixel area around each seed. For details about the clustering see section 3.2.

3. The third step is a row-wise common mode correction (CMC). The CMC is done separately for the two ADC chains and with a hit rejection based on the clustering from the former step.

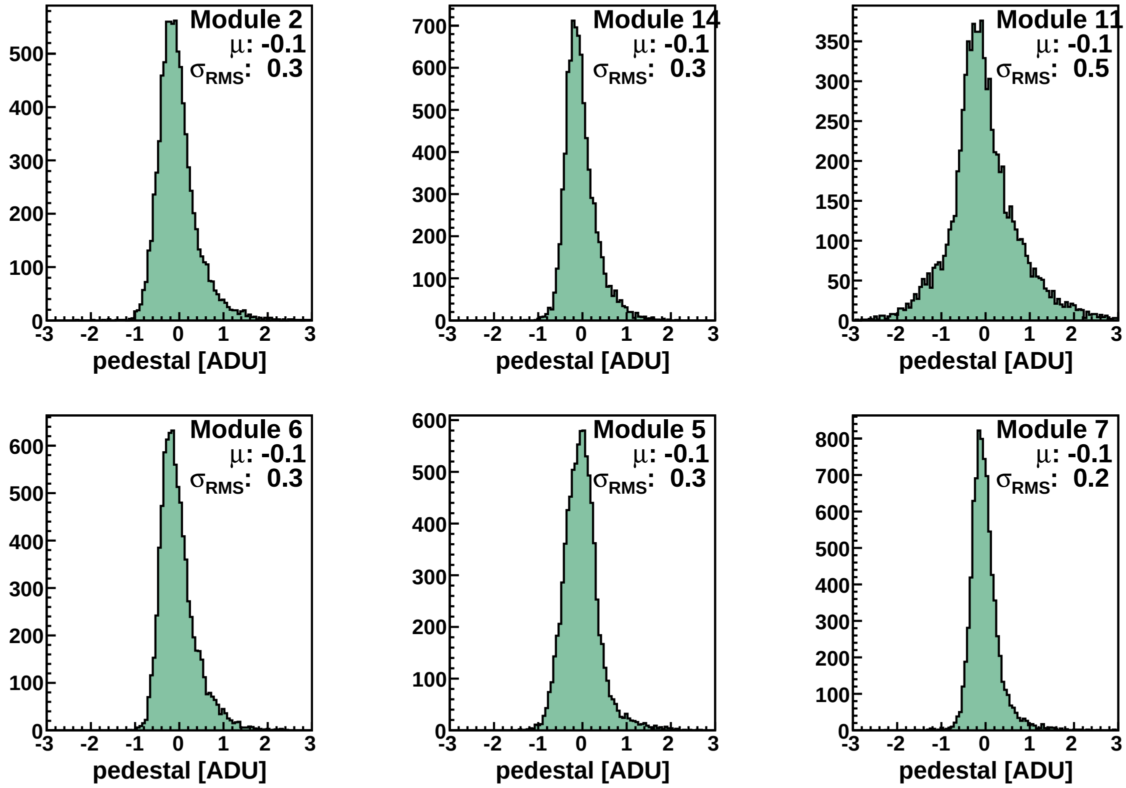


Figure 3.7: *The distribution of pixel pedestals after correction for all six modules. The mean and sigma value of a Gaussian fit to the distribution is printed next to each plot. Compared to the average pixel noise of $\sigma \approx 13$ ADU the remaining pedestals are negligible.*

4. The last step is a column based common mode rejection. To speed up the data analysis this step can be skipped as there is virtually no column based common mode.

Figure 3.8 shows the common mode distribution of the sparisfied data for each module in a logarithmic scale. These common mode values are for rows with hits only, common mode values for empty (i.e. without a hit/cluster) rows are not plotted in figure 3.8. With a few exceptions the common mode values are virtually gaussianly distributed.

Non-Gaussian common mode contributions: These common mode values are rare and contribute less than 0.1% to all common mode values. Figure 3.9 shows the occupancy of CURO channels for these CM values for module 7. As can be seen the distribution for the very high common mode values ($CM > +120$ ADU) shows a systematic bias towards the edges of the read out chip whereas the very low common mode values ($CM < -120$ ADU) are mainly confined to a few read out channels. This behavior is consisted throughout all modules. These findings are also consistent with other lab and test beam results, indicating a sporadic malfunction of channels of the read out chip. An exception to the above is module 11, showing a much poorer performance. In this case this is due

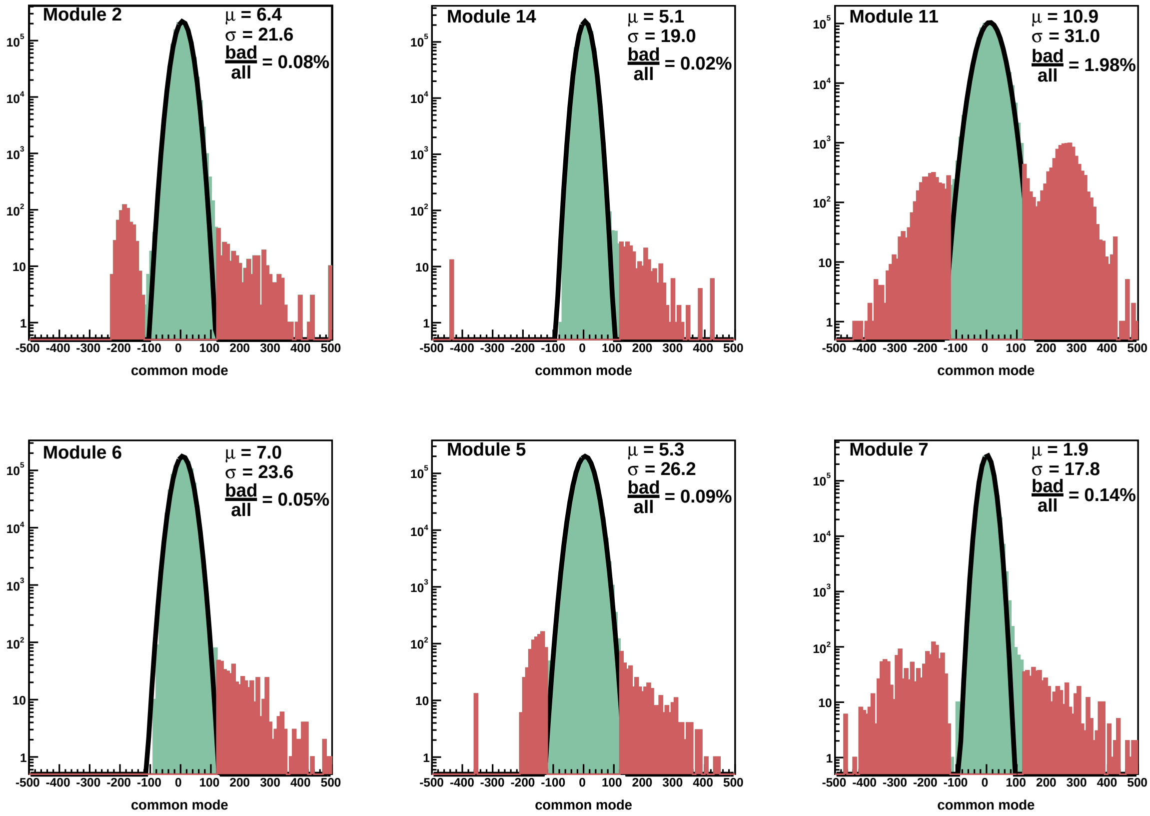


Figure 3.8: *The common mode distributions for each module in a logarithmic scale. The green distribution shows all common mode value within -120 and 120 ADU, the red distribution the common values outside. The mean and sigma value of a Gaussian fit to the common mode distribution as well as the ration between the red and the green clusters are drawn next to each modules distribution. Note that only common mode values from rows with hits are plotted.*

to performance problems of the DEPFET sensor itself.

Leakage Current

One source of the common mode is leakage current: since the sensor is still collecting charge during the readout, the integrated leakage current steadily increases towards the end of the currently read out event frame. This translates into a rising common mode towards the end of a frame. As the start of the readout is random for each event, the row/gate addresses need to be shifted by the start row/gate. Fig. 3.10 shows the common mode distributions on the (Y axis) with respect to the gate in readout order for module 2. The first few rows can be neglected as they are polluted by readout artifacts (see [22] for details). The slope from a straight line to the remaining rows is proportional to the leakage current. With a readout time of $T_{frame} \sim 1.6\,\text{ms}$ and a system gain of $G_{system} = 23e^{-}/ADU$ (see the *clustering* section 3.2) this translates to a leakage current of $I_{Leak} = 111\,\text{fA/pixel}$ or $I_{Leak} = 14.4\,\text{nA/cm}^2$. This measurement is not a very precise quantity

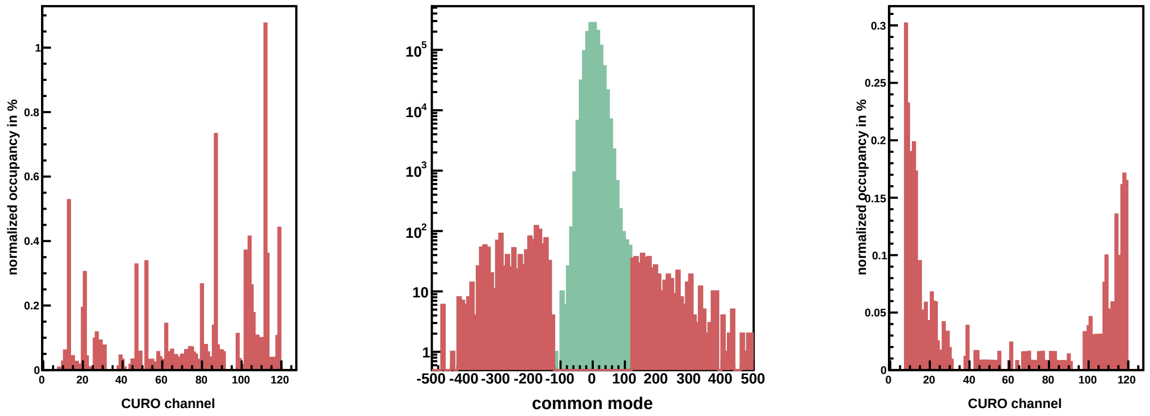


Figure 3.9: *Center image: The common mode distribution for module 7 as shown in fig. 3.8. Left image: The CURO channel occupancies in % for common modes lower than -120 ADU (left) normalized by the total CURO channel occupancies. Right image: the same, but for common mode values higher than +120 ADU.*

since the system gain could only be measured via the most probable value. The values for other telescope modules are between $I_{Leak} = 113\,\mathrm{fA/pixel}$ to $I_{Leak} = 138\,\mathrm{fA/pixel}$ and accordingly $I_{Leak} = 14.7\,\mathrm{nA/cm^2}$ to $I_{Leak} = 18\,\mathrm{nA/cm^2}$. The leakage current is highly temperature dependent. However, a temperature measurement was not implemented in the test beam. Despite the rather unprecise nature of these measurements, they are in agreement with earlier measurements with similar DEPFET sensors (17.3nA/cm^2 [31], 14.8fA/cm^2 [22], 15.3fA/cm^2 [20]).

Noise

After the common mode correction is applied, a final noise figure for each pixel is established. Figure 3.11 displays the pixel noise map and pixel noise distributions before and after common mode correction for module 6. Figure 3.12 shows the pixel noise distributions for all six modules before and after common mode correction including the results of a Gaussian fit. These fit results can also be found in table 3.3. The noise values are basically the same for all modules which is expected as the noise is dominated by the read out chip. However, since the system gain is different the according ENC[7] is different for each module.

Gain Variations

So far it has been assumed that all pixels have the same gain. This is usually not the case due to a variety of reasons, e.g.

- process variations during the production of a DEPFET sensor wafer,
- supply or bias voltage drops inside the read outchip CURO,

---

[7] Equivalent Noise Charge

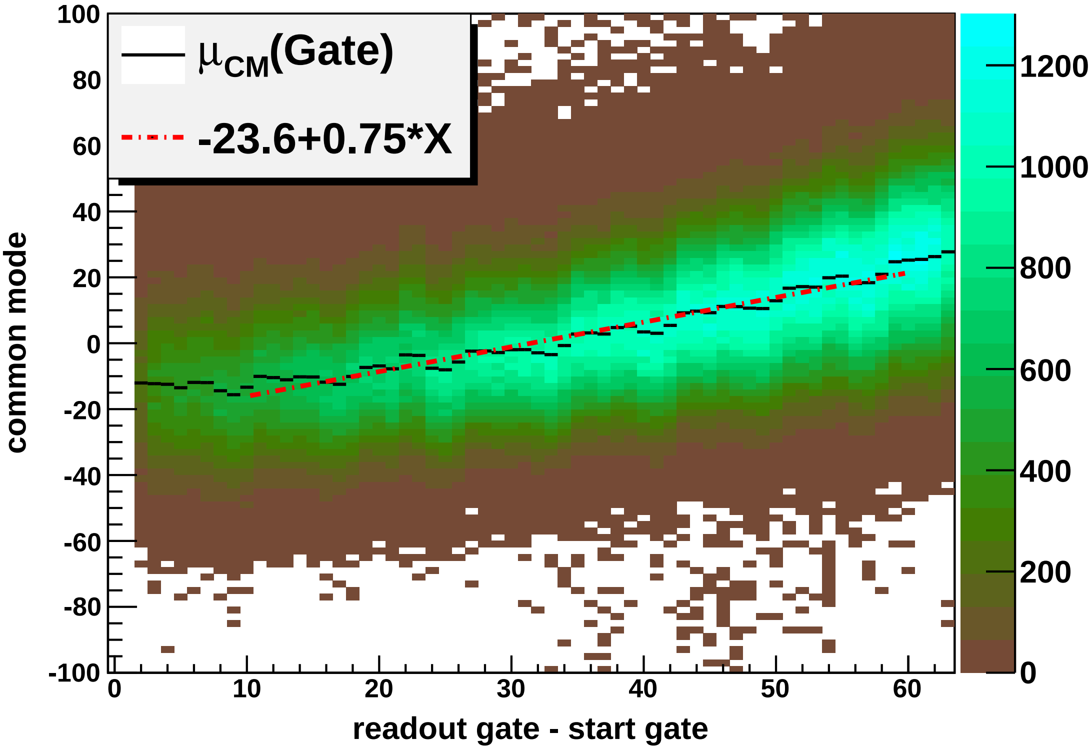


Figure 3.10: *The distribution of common mode values (Y axis) versus the double row/gate number(X axis) for module 2. The gate number is corrected for the random position of the first gate to be read. The overlayed black lines represent the mean value of a Gaussian fit to the common mode distribution for each gate. The red, dashed line indicates the result of a straight line fit to these mean values. The higher the gate number the later it is read out and the more leakage current is integrated. This results in a higher common mode for higher gate numbers.*

- supply or bias voltage drops inside the steering chip SWITCHER,
- critical bias voltages for the DEPFET sensor are slightly off, e.g.common clear gate V_{CCG} or clear low V_{ClLow}.

Row and column vs. gain wise correction: The first choice to tackle gain variations would be a pixel wise correction, but unfortunately the statistics are not high enough to do this for the data at hand. The relative error on the arithmetic mean for example is proportional to $n^{-\frac{1}{2}}$, i.e. for a 1% error $n \sim 10^4$ hits per pixel are needed. With more than 8000 pixels this translates to a total number of $n \sim 10^8$ hits, far outside the scope of any DEPFET test beam. Therefore gain variations were corrected SWITCHER and CURO channel wise. With 128 channels for each axis the necessary number of hits to achieve a 1% error reduces to $n = 128 \cdot 10^4 \sim 10^6$. This is well inside the scope of the underlying test beam data.

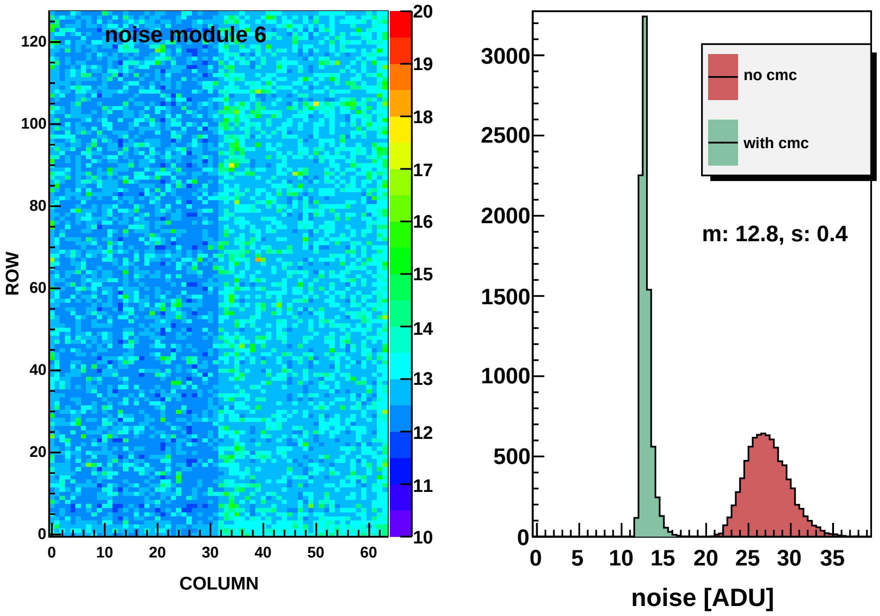


Figure 3.11: *Pixel noise of module 6: The left image shows the common mode corrected pixel noise map, the right plot displays the pixel noise distributions before and after common mode correction. The sensor is split by the two readout chains in two halves. This causes the different noise values in each sensor half.*

Seed vs. cluster wise correction: Taking the seed signal distribution has several benefits over the cluster signal distribution: Only one and not 9 or 25 (depending on the cluster size) different gains are taken into account and the noise is also lower by a factor of 3 or 5, respectively.

Gaussian fit or arithmetic mean correction: The seed pixel signal distribution does not exactly follow a Gaussian distribution. Therefore in addition to a Gaussian distribution fit the arithmetic mean of the seed distribution between 250 and 2000 ADUs was used for gain correction. Although the results are slightly different for some modules the effects on the resolution and signal distributions are negligible. A few sample results are shown in figures 3.13 to 3.16:

- Row wise gain variations: For some modules (e.g. module 5, Fig. 3.13) a distinct row wise dependence of the gain can be observed, whereas in other modules this behavior is much less prominent (e.g module 6, Fig. 3.15). Nevertheless, a matrix wide gain dependence along the y axis can be observed within all modules. These gain variations strongly depend on the biasing conditions of the DEPFET matrix and have also been observed in the lab.

- Column wise gain variations: Column wise gain variations can be found within all modules and are probably attributed to different gains in the CURO read out channel. (Fig. 3.14 and 3.16).

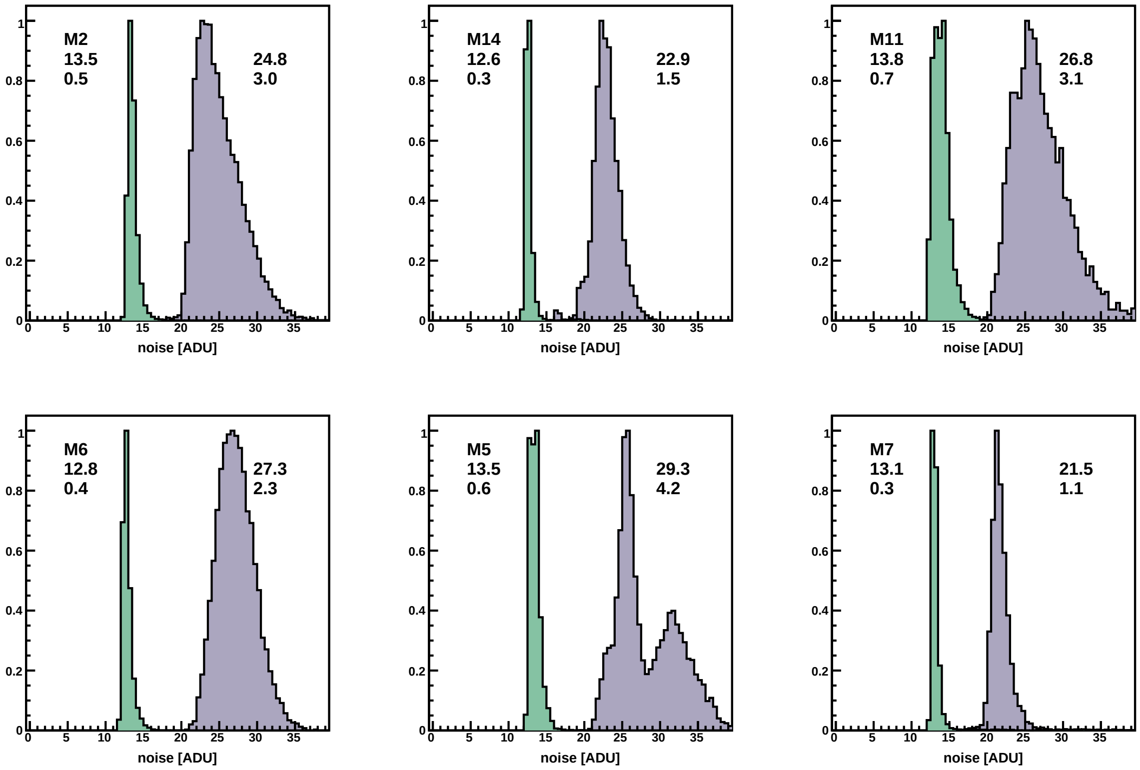


Figure 3.12: *Pixel noise distributions for all modules before and after common mode correction, including the mean and sigma values for Gauss fits to both distributions.*

For the most parts the gain variations have been equalized out (e.g. Fig. 3.14 and 3.16), but in one extreme case, see Fig, 3.14, it was not possible to cancel out the variations completely.

Masking

Noisy or defect parts of the sensor must be masked as is illustrated in figure 3.17. The lower two plots show the seed and the cluster signal distribution. Both distributions have an excess of low pulse hight entries. The upper left picture shows these two variables plotted against each other. Besides the expected majority of entries centered around the most probable value on the Y axis and the mean of the seed signal distribution on the X axis, a large amount of entries can be found in the lower left (i.e. low seed and cluster signal) corner of this picture. This region is included in the area designated **A**. The right picture in this figure shows a hit map of all entries within this marked area **A**. From this picture it becomes clear that the majority of these low pulse hight entries stem from bad readout channels.

Figure 3.17 shows only one possible origin of objectionable signals, other are for example defect rows. These are not shown in figure 3.17 but already masked out together with

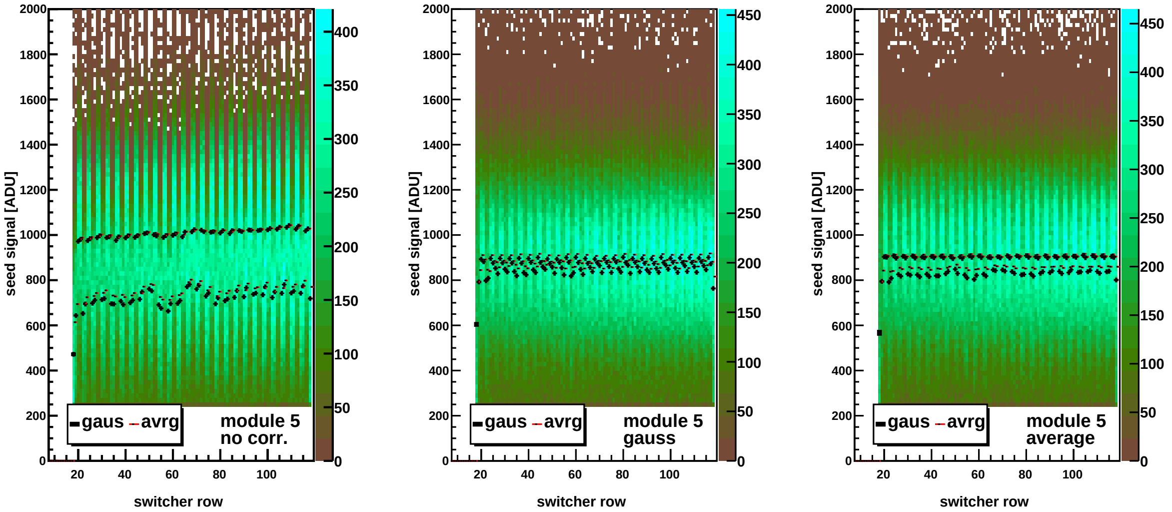


__Figure 3.13:__ Row wise seed signal distribution for module 5 before and after gain corrections: On the x axis is the matrix row and on the y axis the seed signal with the color scale indicating the entries for a certain signal and row, e.g. the color variations along the y axis show the seed signal distribution for each row. Overlayed are the arithmetic and Gaussian mean values. The left plot shows the signals before gain correction, the middle and the right plot show signals after Gaussian and arithmetic mean corrections respectively.

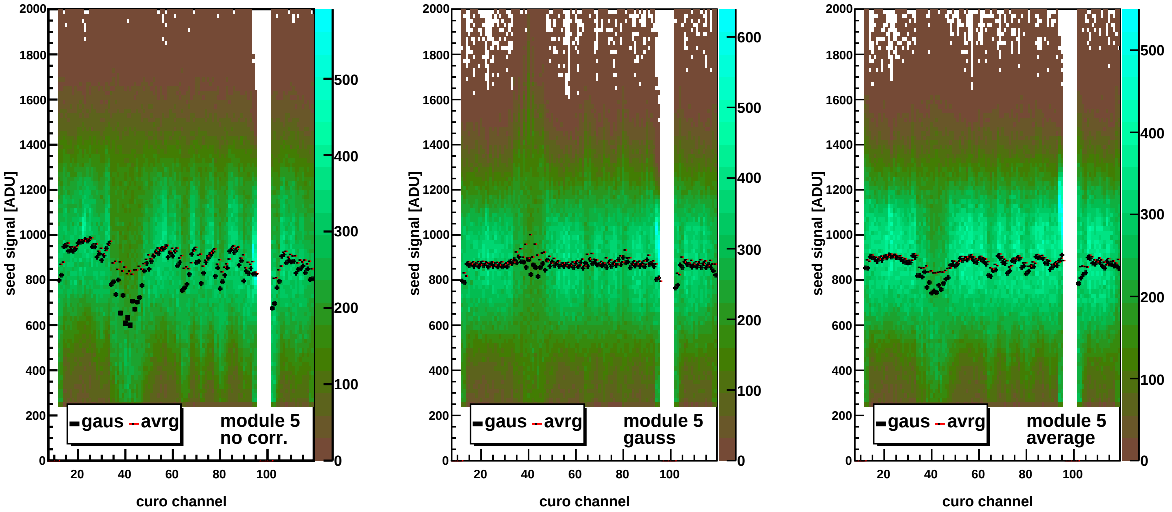


__Figure 3.14:__ Column wise seed signal distribution for module 5 before and after gain corrections. For details see caption of fig. 3.13. Here, the x axis shows the matrix column.

a frame consisting of the outer four rows and columns on each side. In figure 3.17 these masked areas are white (i.e. have no entries). While some modules showed little defects

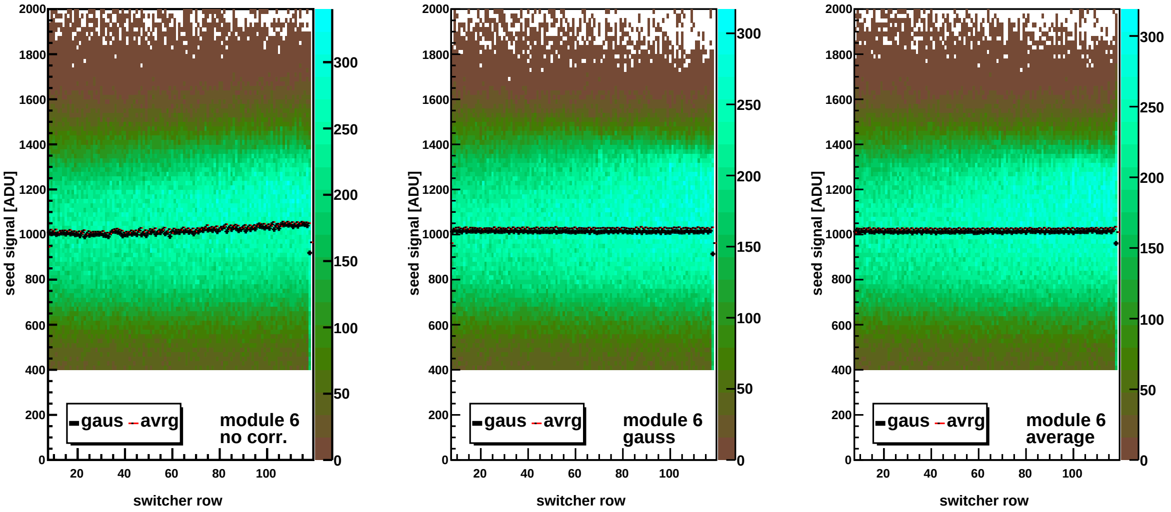


__Figure 3.15:__ Row wise seed signal distribution for module 6 before and after gain corrections. For details see caption of fig. 3.13.

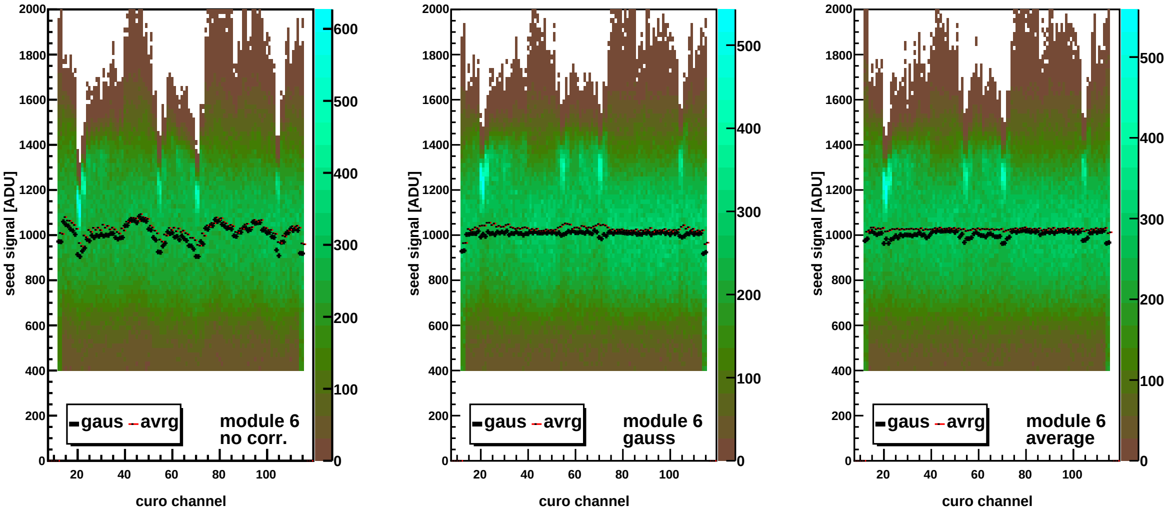


__Figure 3.16:__ Column wise seed signal distribution for module 6 before and after gain corrections. For details see caption of fig. 3.13. Here, the x axis shows the matrix column.

(e.g. module 6 and 14) other modules needed a massive amount of masking (e.g. module 11). The origins of these defects are either defects in the sensor itself, but also defect steering and read out channels caused by a faulty chip or a damaged wire bond. Due to masking the overlapping and useful area of the telescope and DUT planes was significantly reduced. This in turn limited the available statistics.

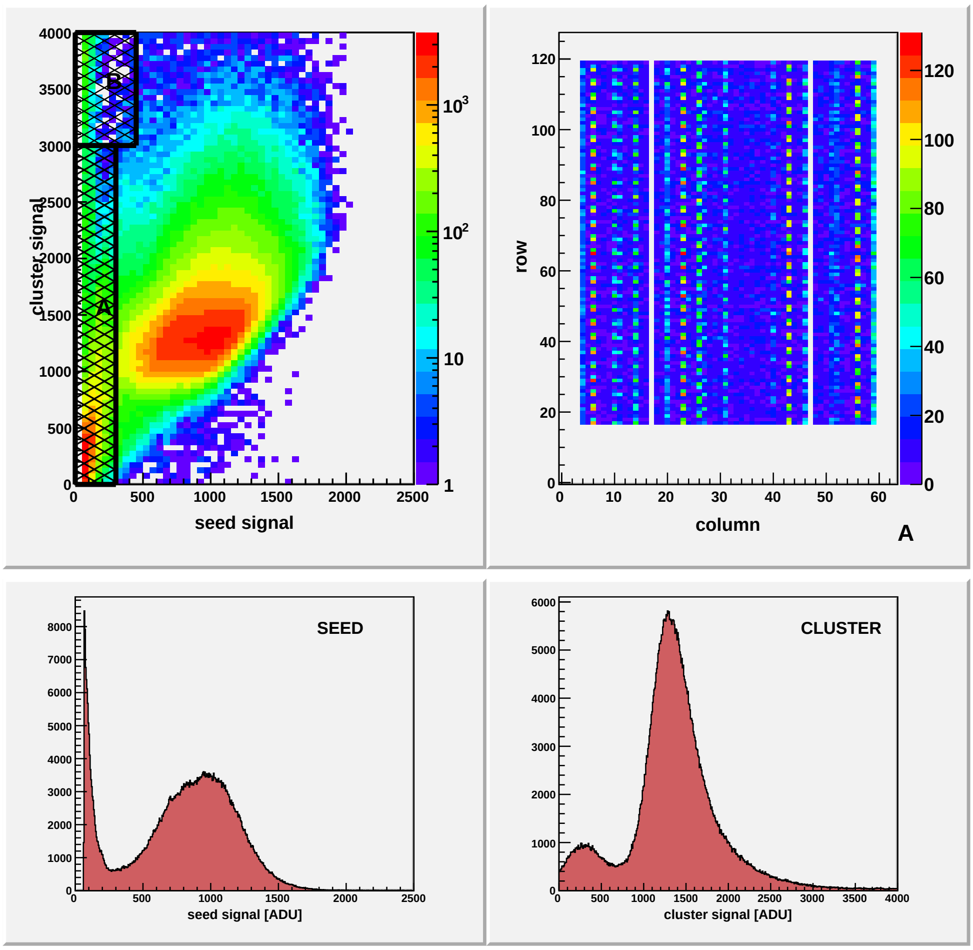


Figure 3.17: *Illustration of masking taking module 7 as an example: The seed and the cluster signal distributions (the lower plots) have low pulse hight entries far exceeding the expected abundance for a well behaving sensor. The upper left plot shows the correlation between seed and cluster signal with two areas marked. The upper right plots shows a hit map of all entries inside the "A" area. Most low pulse entries are confined to noisy read out channels.*

Clustering

The next step in the data processing chain after pedestal and common mode correction is hit finding and clustering. In this step all pixels are rejected except for those hit by a particle, the so called *seed pixel*, and its proximity, the so called *cluster*. The size of the cluster is somewhat arbitrary and a compromise between data reduction and charge cloud coverage. This means, the cluster should be large enough to cover the charge cloud, while at the same time pixels with pure noise are rejected. Therefore this choice depends largely on the size of the charge cloud with respect to the pixel size. In this study a 5×5 pixel area around the seed pixel is associated with the seed. The charge cloud has an estimated width of $\sim 6 - 7\,\mu$m (see chapter 1). Even with a somewhat larger charge cloud of $\sim 12\,\mu$m and an entry point at the seed pixel border, a five pixel wide cluster would still contain 99.997% (corresponding to 4σ) of the total charge. The seed finding

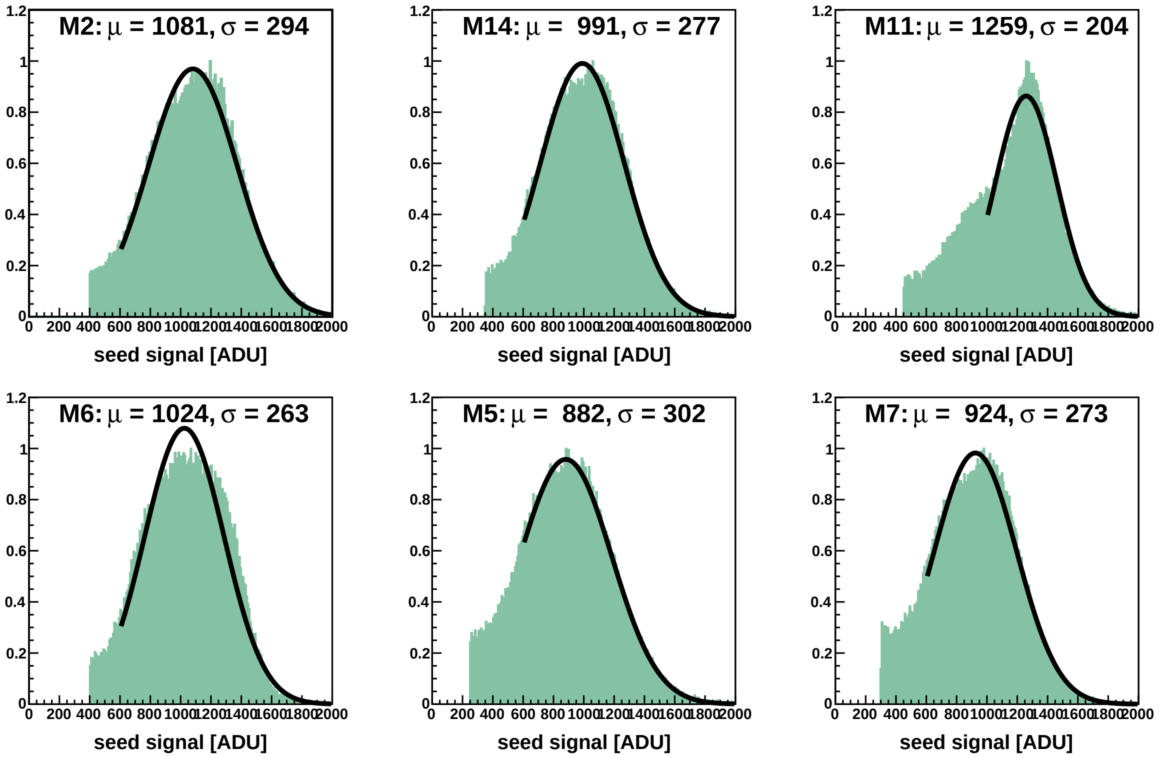


Figure 3.18: *The* ***seed*** *signal histogram for each of the six modules including the fit results of a normal distribution to each histogram.*

and clustering algorithm was implemented in the following fashion:

1. A list is filled with all pixels in an event frame with a signal over noise ratio higher than $S/N \geq 5.0\,\sigma_{noise}$ and sorted by signal height in descending order.
2. Starting with the first pixel (i.e. the highest signal) a list of clusters possessing information about a 5×5 pixel proximity is created. There is no cut on the neighbors.
3. After each cluster was found the pixels associated with these clusters are blocked for the remaining seed search in this event, thus prohibiting false double entries.

Figure 3.18 shows the seed signal distribution and 3.19 shows the cluster (5×5) signal distribution of each module. To each seed and cluster signal histogram a Gaussian and a Landau distribution, respectively, was fitted. The most probable value of the Landau fit was used to estimate the total gain G_{tot}. The gain is the gain g_q of the DEPFET itself and the system gain g_{sys} consisting of the transimpedance amplifier and ADC gain:

$$G_{tot}\left[\frac{ADU}{e^-}\right] = g_q\left[\frac{nA}{e^-}\right]\cdot g_{sys}\left[\frac{ADU}{nA}\right] \tag{3.2}$$

The expected number of electron hole pairs for $450\,\mu$m silicon is $\sim 36600\,e^-$ resulting in a system gain G_{tot} ranging from $\sim 1/21$ to $\sim 1/30\,\text{ADU}/e^-$. To estimate the actual

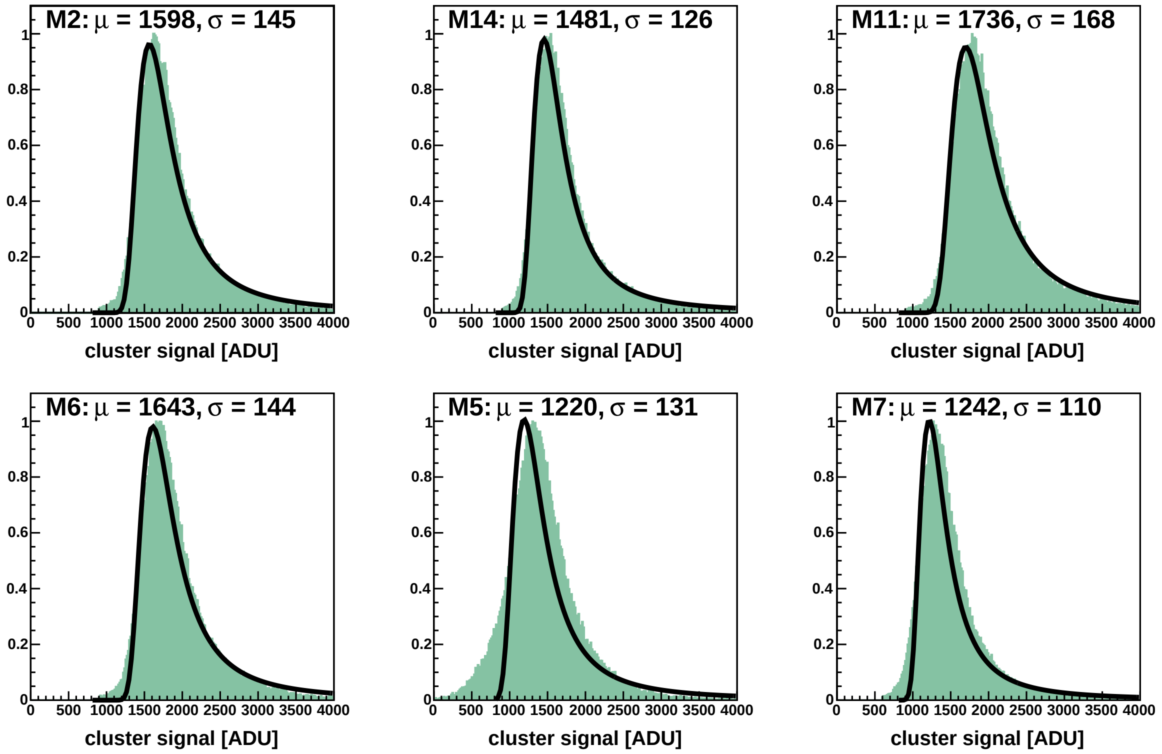


Figure 3.19: *The* 5×5 ***cluster*** *signal histogram for each of the six modules including the fit results of a Landau distribution to each histogram.*

DEPFET pixel gain g_q for each module a system gain of $g_{sys} = 7.7\,nA/ADU$ [22] is used yielding in values of $g_q \sim 0.26 - 0.34\,\mathrm{nA/e^-}$for the `CoCG-L` pixel type, $g_q \sim 0.37\,\mathrm{nA/e^-}$for `SIMC-L`, and $g_q \sim 0.35\,\mathrm{nA/e^-}$for `CoCG-S`. This is comparable to other measurements of g_q(`CoCG-L`) $\sim 0.26 - 0.36\,\mathrm{nA/e^-}$obtained with a radioactive source [20]. Table 3.3 shows a summary of results obtained after the clustering process. The table also shows the ratio of the width parameters (actually ξ of equation 1.7) of the Landau distribution to the most probable value. Theoretical considerations predict a ratio of $\sigma_{MPV}/\mu_{MPV} = 6.2\%$ for $450\,\mu$m. However, the measured values are somewhat higher. This can be contributed to the fact that the individual pixel noise as well as a global gain spread of $\Delta g_q \sim 2.5 - 5\%$ [21] widens the cluster signal. However, for the further course of this work knowledge of the exact gain is not necessary.

3.3 Position reconstruction

In this section the results of standard position reconstruction methods (*center-of-gravity*, η) are presented. For alignment and tracking the SITBEAN package [49] was used. Although the inner two detectors are designated as DUTs in principal every sensor plane can be used as a DUT and the necessary tracking space points are provided by the other sensor planes. Since some of the sensor planes behaved unexpectedly due to the sub-

Module ID	2	14	11	6	5	7
DEPFET type	CoCG-L	CoCG-L	SIMC-L	CoCG-S	CoCG-L	CoCG-L
$\mu_{noise}\,[ADU]$	13.5	12.6	14.1	12.8	13.5	13.1
$\mu_{seed}\,[ADU]$	1081	991	1259	1024	882	924
$\sigma_{seed}\,[e^-]$	294	277	204	263	302	273
$\mu_{MPV}\,[ADU]$	1598	1481	1736	1643	1220	1242
$\sigma_{MPV}\,[ADU]$	145	126	168	144	131	110
$G_{tot}[e^-/ADU]$	22.9	24.7	21.1	22.3	30.0	29.5
$g_q\,[\mathrm{nA/e^-}]$	0.34	0.31	0.37	0.35	0.26	0.26
$\mu_{noise}\,[e^-]$	309	311	297	285	405	386
$\mu_{seed}/\mu_{MPV}[\%]$	68	67	73	62	72	74
$\sigma_{seed}/\mu_{MPV}[\%]$	18	19	12	16	25	22
$\sigma_{MPV}/\mu_{MPV}[\%]$	9	9	10	9	11	9
S/N_{seed}	80	79	89	80	65	71
S/N_{MPV}	18	118	123	128	90	95

Table 3.3: *A summary of sensor properties measured in the test beam. Noise is referring to the average pixel noise.* μ *and* σ *denote a Gaussian fit to the noise, seed and cluster signal distribution.* G_{tot} *is a gain calibration assuming that the most probable value fit* μ_{MPV} *is equivalent to the expected* $\sim 36600\,e^-$ *in* $450\,\mu m$ *silicon and the ENC is calculated using this* G_{tot}*. The signal to noise ratios are* always *referring to the single pixel noise* μ_{noise}*.*

optimal biasing conditions all sensor planes were studied with regard to their position reconstruction performance.

Binary read out

In the binary read out only column and row number of the seed pixel but no analog information is known. Assuming a uniform pixel response and neglecting noise the theoretical resolution depends only on the pixel pitch p:

$$\Delta x = \sqrt{(x - x_{rec})^2} = \sqrt{\frac{1}{p}\int_0^p \left(x - \frac{p}{2}\right)^2 dx} = \frac{p}{\sqrt{12}} \tag{3.3}$$

For the studied DEPFET sensors this yields to $\Delta x = 9.24\,\mu m$ for a pitch of $p = 32\,\mu m$ (COCG Large X-axis) and $\Delta x = 6.93\,\mu m$ for a pitch of $p = 24\,\mu m$ (COCG Large Y-axis and COCG Small X-axis). These theoretical expectations have been confirmed in [22].

Center of Gravity

Just like the homonymous sibling in mechanics this methods yields a weighted average position. Here, the pixel positions is weighted by the pixel charge signal. This is done separately for the x and y axes with the signal sum projected to each axis,

$$X_{CoG} = \sum_x X_x \cdot S_x^{PrjY} = \sum_x \left(X_x \cdot \sum_y S_{x,y} \right) \tag{3.4}$$

$$Y_{CoG} = \sum_y Y_y \cdot S_y^{PrjX} = \sum_y \left(Y_y \cdot \sum_x S_{x,y} \right) \tag{3.5}$$

where $S_{x,y}$ is the signal of pixel x,y and X_x and Y_x are its position projected on the respective axis. For this method the charge sharing process is assumed to be linear and the charge cloud assumed to be box shaped since all charges and positions enter linear into the result. The real shape of the charge cloud, however, is roughly Gaussian leading to a systematic error. Figure 3.20 shows the residual distributions $x_{predicted} - x_{CoG}$ and 3.21 shows the residual distributions $y_{predicted} - y_{CoG}$ for all six modules. In the X axis all modules except for module 6 have a pixel size of 32 μm and module 6 has a pixel pitch of 24 μm. Since the effect of the non linearity depends on the charge cloud to pixel size ratio the module with the smaller pixel dimension is much less affected. The double peak structure seen in figure 3.20 is expected for a certain charge cloud to pixel size ratio. This error becomes apparent in a plot of CoG residuals versus the in-pixel positions using telescope tracks, e.g. module 14 in fig. 3.22. These results are in agreement with a simulated charge cloud of $\approx 6-7\,\mu$m as is shown in figure 3.23 (32 μm pitch) and figure 3.24 (24 μm pitch).

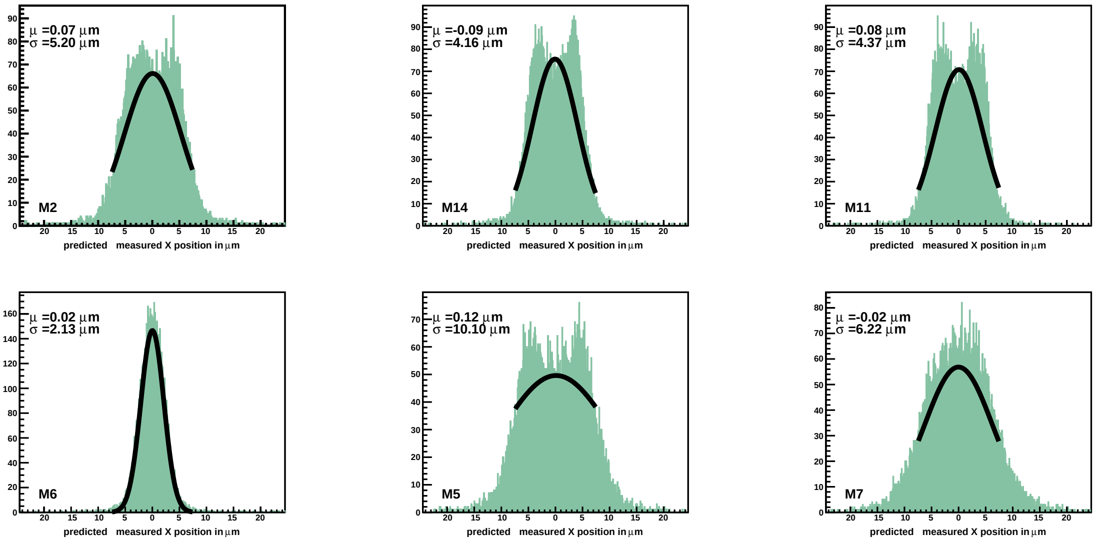


Figure 3.20: *Center of gravity residuals $x_{predicted} - x_{CoG}$ for all six modules. Except for module 6 all modules show a double peak structure due to the systematic error of the center of gravity method. Module 6 has smaller pixel size (24 μm) than the other modules (32 μm) and is therefore less affected.*

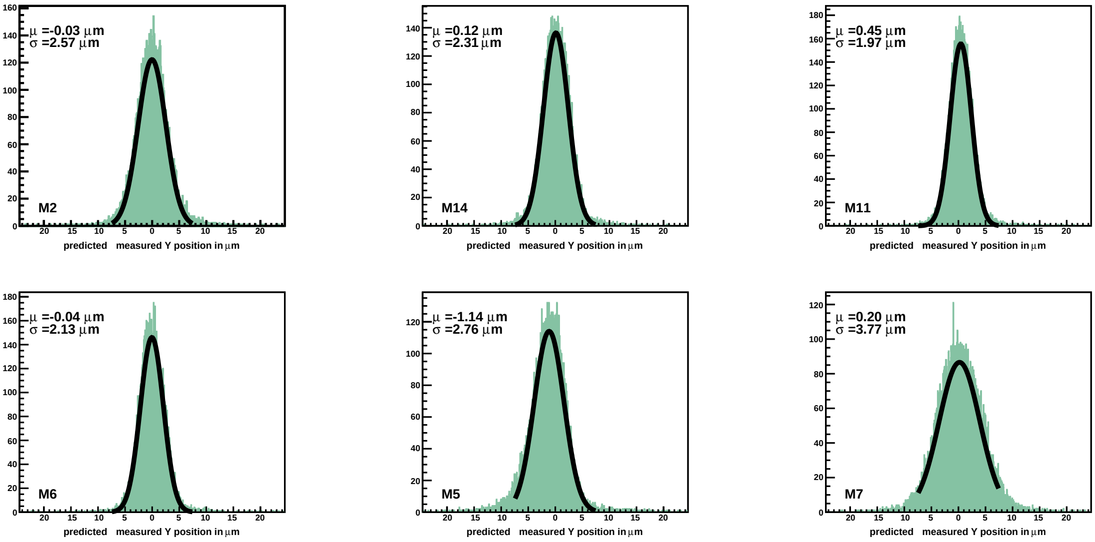


Figure 3.21: *Center of gravity residuals $y_{predicted} - y_{CoG}$ for all six modules. Since all modules have a pixel size of $24\,\mu m$ in this direction the effects of the non linear charge sharing are less pronounced.*

The η algorithm

To avoid the systematic error of the Center-of-Gravity method a new approach, sensitive to the non-linearity of the charge distributions among the pixels, is needed: the η method. The basic principle of the η method is to provide a correction function $\mathcal{F}(\eta)$ for the non-linear charge distribution derived from the empirically obtained charge ratio η between the two pixels with the highest signal [50].

The first of the two central elements of the η method is the variable η itself which describes the charge sharing between two pixels: the seed and its neighbor with the highest signal. With S_L and S_R the signal of the left and right pixel respectively, η is defined as

$$\eta = \frac{S_R}{S_L + S_R} \tag{3.6}$$

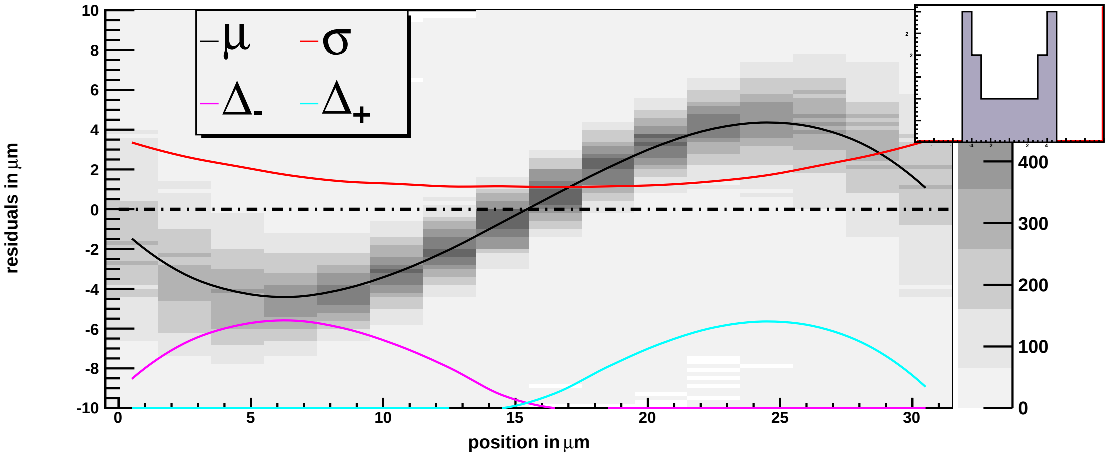
The distribution of η is sampled separately for the x and y axis since no dependence of η_X on η_Y and vice versa has been found.

The η distribution goes from 0 (the whole signal in the left pixel and no signal in the right pixel) to 1 (the whole signal in the right pixel and no signal in the left pixel). The correction function $\mathcal{F}(\eta)$ is derived by integrating over the η distribution:

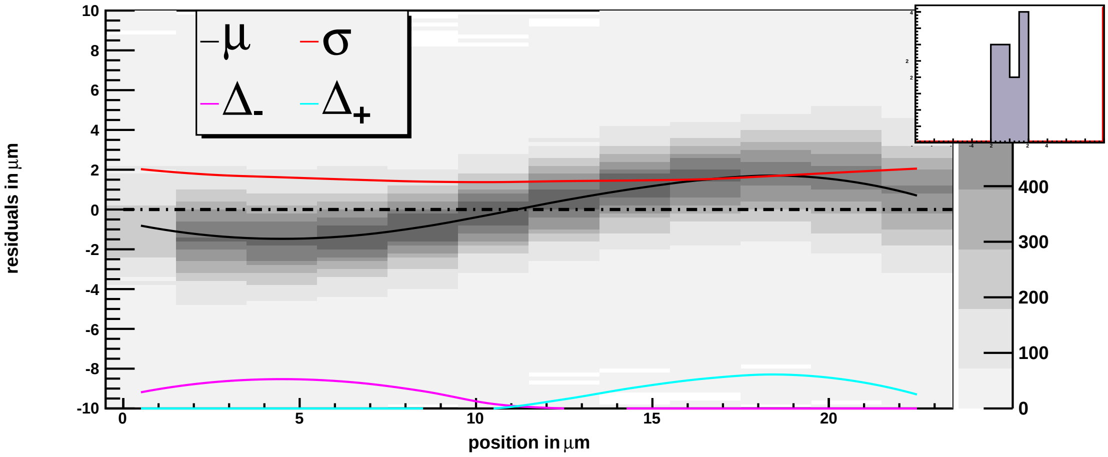
$$\mathcal{F}(\eta) = \frac{1}{N_0} \int_0^{\eta} \frac{dN}{d\tilde{\eta}} d\tilde{\eta} \tag{3.7}$$

Assuming uniform pixel illumination and response behavior, the reconstructed position in x is (for the above definition of η, eqn. 3.6):

$$x_{rec} = x_{left} - p_x \mathcal{F}(\eta_x) \tag{3.8}$$



(a) X vs X



(b) Y vs Y

Figure 3.22: *The X and Y axis residual distribution $x_{predicted} - x_{measured}$ plotted against the in-pixel tracking position $x_{predicted}$ using the center of gravity method for position reconstruction with all units in μm. The black line (μ) represents the mean value μ_{CoG} of a Gaussian fit to the distribution for each 1 μm wide bin. The red line (σ) is the corresponding σ_{CoG} width. The two lines ($\Delta_{\pm}$)on the bottom indicate the deviation of the mean value from an ideal reconstruction method with $\mu_{ideal} = 0$ for all in-pixel positions. The insert in the top right corner of each plot shows the a one dimensional distribution of μ_{CoG}. This figure shows the results for module 14 (CoCG large) with a pixel size of 32 μm in X and 24 μm in Y. As expected this is reflected in the width of the μ_{CoG} deviations from an ideal method $\mu_{ideal} = 0$: The axis with the larger pixel size is much stronger affected also displaying a μ_{CoG} distribution with two distinct peaks. This is reflected in the two peaked residual distributions shown in figure 3.20.*

where p_x is the pixel pitch. Since the η method was originally developed for strip detectors the application towards pixel sensors is not necessarily straight forward. Therefore two

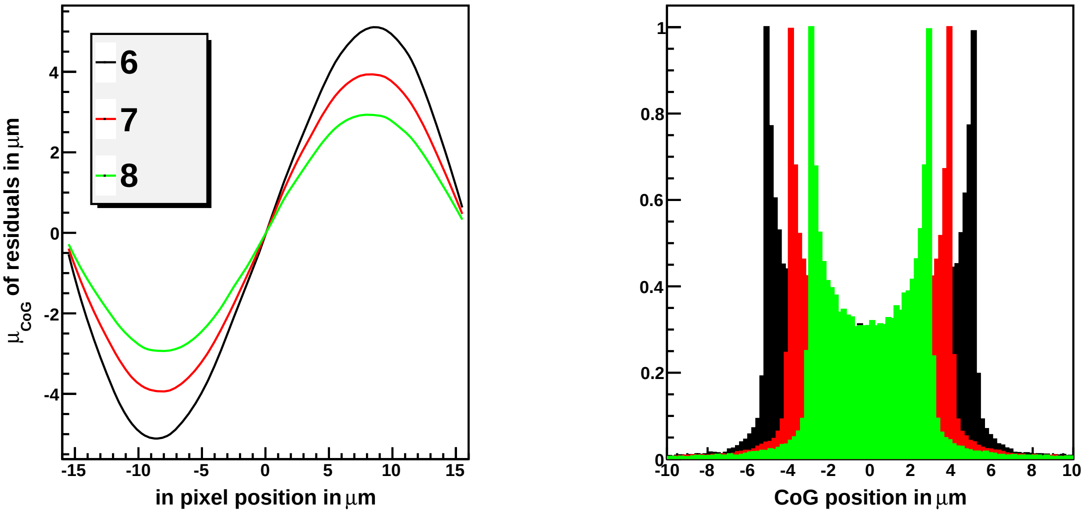


Figure 3.23: *Simulated mean values of center of gravity residuals* $x_{predicted} - x_{CoG}$ *for a pitch of 32 µm and charge cloud size of 6 to 8 µm. The left plot shows the variations as a function of the in-pixel position equivalent to the black line in figure 3.22. The right plot shows the distribution of the mean values equivalent to the small insert in figure 3.22. Note the distinct double peak structure which has also been observed in the data.*

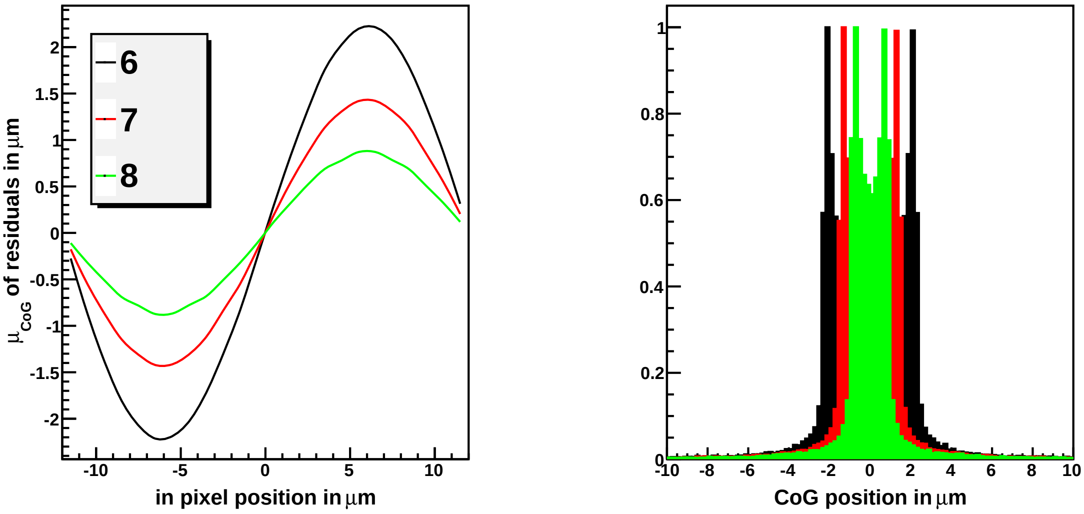


Figure 3.24: *The same as in figure 3.23 but with a pixel pitch of 24 µm. The effects of non linear charge sharing are much less pronounced.*

approaches where chosen:

Strip like and projected η: A first approach (*strip like*) is to only use the direct neighbors of the seed pixel for the η method, e.g. left or right pixel in x and top or bottom pixel in y. The second approach (*projected*) uses the information of a 3×3 pixel cluster around the seed by projecting the signals to each axis. That means that the signal of the left pixel now is the sum of all three pixels to the left of the seed etc. As will be shown later, the *strip like* methods yields somewhat better results.

Individual η distribution for different pixel types: Anticipating later results pixel are found to behave differently according to their row number. For some modules this behavior is based on an even/odd row system, others show a pattern based on the row number modulo 4 (e.g. the patter repeats every fourth row). Therefore η distributions were collected assuming all pixels are behaving equal (*all*), pixel response varies on an even/odd pattern (*double*), and pixel behavior varies on a pattern repeating every fourth row (*quadro*). Anticipating the effects on the residuals, the differences are insignificant and dwarfed by the effect of other corrections. Nevertheless, all seven η distributions are shown in figures 3.25, 3.27, 3.28, and 3.29.

$\mathcal{F}_{FB}(\eta)$ from forward and backward integration of η: There is a caveat when working with the η function. For the correction function $\mathcal{F}(\eta)$ to be derived from equation 3.7 one assumes a uniform response and illumination. Yet a look at the $\mathcal{F}(\eta)$ distributions for module 5 in figure 3.25 reveals that the pixels top and bottom of the seed pixel, equivalent to $\eta < 0.5$ and $\eta > 0.5$, have a different occurrence. This leads to a correction function which does not go through (0.5, 0.5) indicating that the underlying assumption of uniformity is wrong. One way to tackle this problem is to take the average of a forward $\mathcal{F}_F(\eta)$ and backward $\mathcal{F}_B(\eta)$ integration of η with:

$$\begin{aligned} \mathcal{F}_F(\eta) &= \frac{1}{N_0}\int_0^{\eta} \frac{dN}{d\tilde{\eta}} d\tilde{\eta} \qquad (3.9)\\ \mathcal{F}_B(\eta) &= \frac{1}{N_0}\int_{1-\eta}^{1} \frac{dN}{d\tilde{\eta}} d\tilde{\eta} \qquad (3.10)\\ \mathcal{F}_{FB}(\eta) &= \frac{\mathcal{F}_F(\eta) + \mathcal{F}_B(\eta)}{2}. \qquad (3.11) \end{aligned}$$

The filled green curve in fig. 3.25 shows the resulting correction function $\mathcal{F}(\eta)$ and the filled magenta curved the difference between forward and backward integration. It should be noted that this effect is most pronounced in module 5, which is shown in figure 3.25, and other sensors are less severely affected.

Asymmetry in the position of the second highest pixel: For an uniformly behaving system the chance of the second highest pixel being to the left or to the right of the seed pixel should be equal. However, in the test beam some module did not behave uniformly. This becomes evident in figure 3.26 which shows the position of the second highest pixel inside the cluster. The difference in x and y axis is due to the different pixel dimensions along those axes. However, a clear asymmetry can be observed in one and the same axis in almost all cases. The numerical difference in position occupancy results

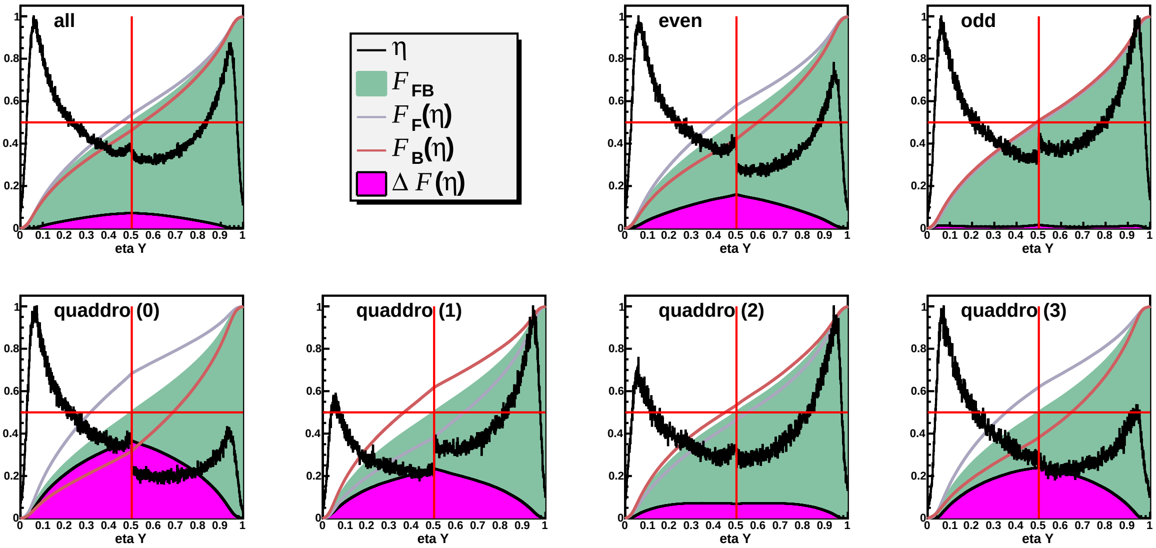


Figure 3.25: *η and $\mathcal{F}(\eta_x)$ for module 5 before any correction of the position occurrence of the second highest pixel. Each of the seven plot shows:*

- *the η distribution itself in black,*
- *the correction function from forward integration of η, $\mathcal{F}_F(\eta)$, in light grey,*
- *the correction function from backward integration $\mathcal{F}_F(\eta)$ in light red,*
- *the average of the forward and backward integration $\mathcal{F}_{FB}(\eta_x)$ (filled, light green curve),*
- *the difference between forward and backward integration $\Delta\mathcal{F}(\eta)$ (filled, magenta curve).*

*The red cross indicates the (0.5, 0.5) point. Shown from top left to bottom right are: one η distribution for **all** pixels, the plot legend, η for **even** and **odd** pixels, and the bottom four plots show η for row number MODULO four, e.g. a pattern that repeats every fourth pixel (see text for details).*

in an erroneous η distribution. Therefore a correction for the left-right pixel occurrence asymmetry should yield in an improved η distribution. Figure 3.27 shows the asymmetry corrected η distributions for module 5 analog to figure 3.25: The differences between forward and backward integration are massively suppressed confirming the validity of this method. The η distribution for the left-right asymmetry corrected module 6 (the central DUT) are shown in figures 3.28 and 3.29.

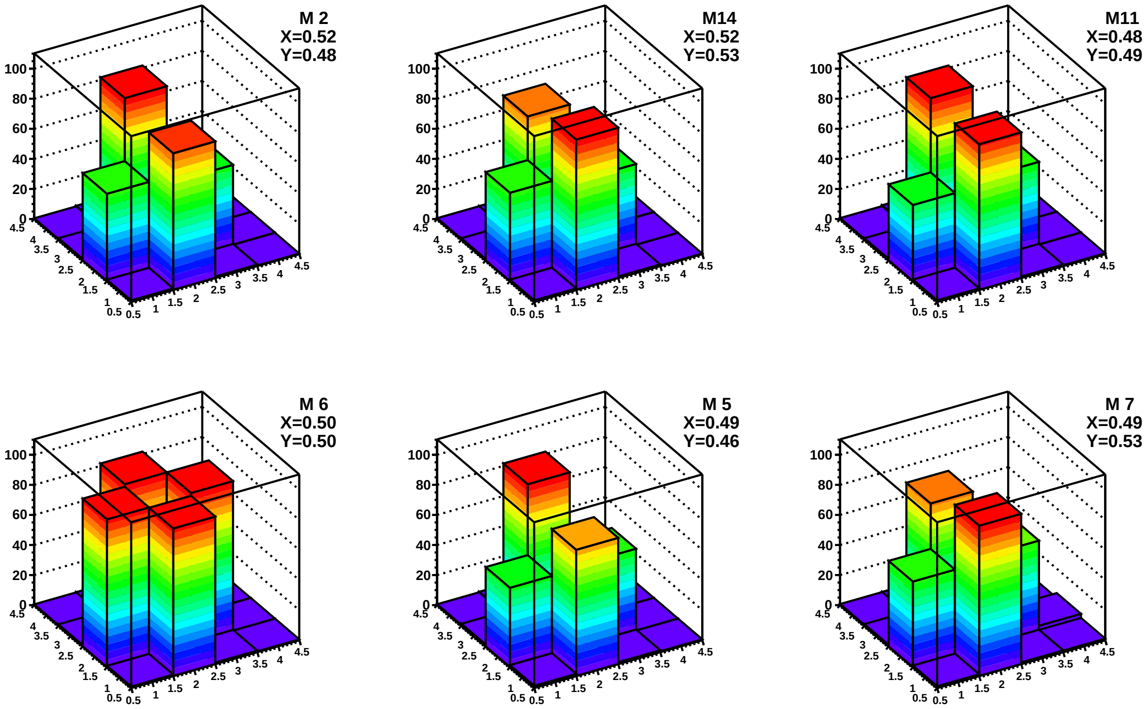


Figure 3.26: *The position of the pixel with the second highest signal inside a cluster for all six modules. Since the x and y dimensions for all modules are different (32 μm and 24 μm) expect for module 6 (both directions 24 μm) the second highest pixel is accordingly more often found along the smaller pixel dimension. On the other hand an ideally responding pixel sensor would show no difference between the pixels in one and the same axis yet some clear asymmetries between left and right seed pixel sides can be found. The asymmetry* $N_R/(N_L + N_R)$ *is plotted inside each plot for both axes.*

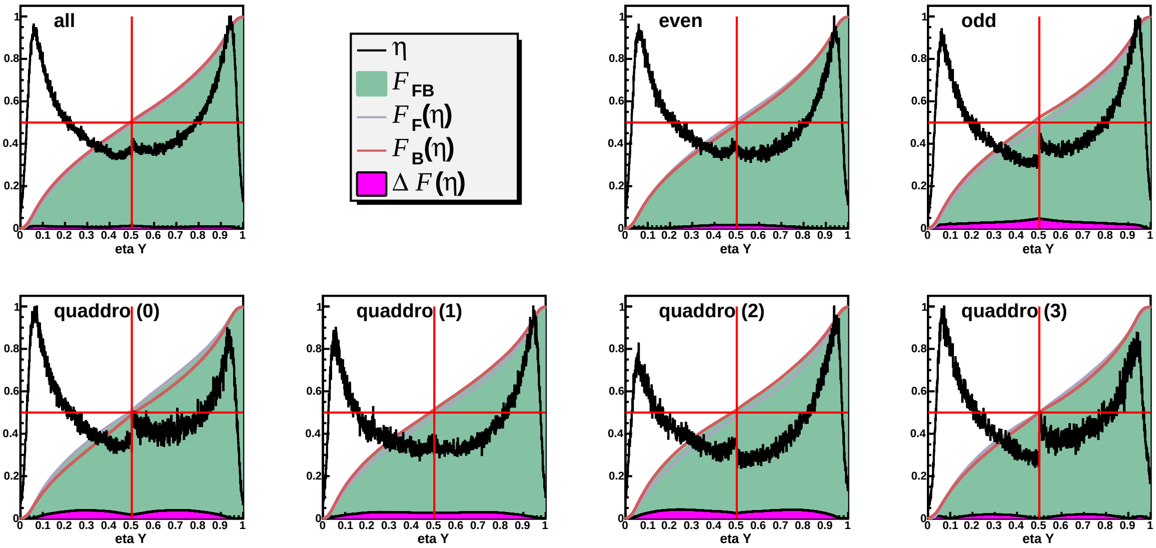


Figure 3.27: *The seven η distributions and $\mathcal{F}(\eta_x)$ functions for the Y axis of module 5 with a correction of the left-right asymmetry shown in figure 3.26. The difference between forward and backward integration (lower magenta histogram) is massively suppressed. For other details see figure 3.25.*

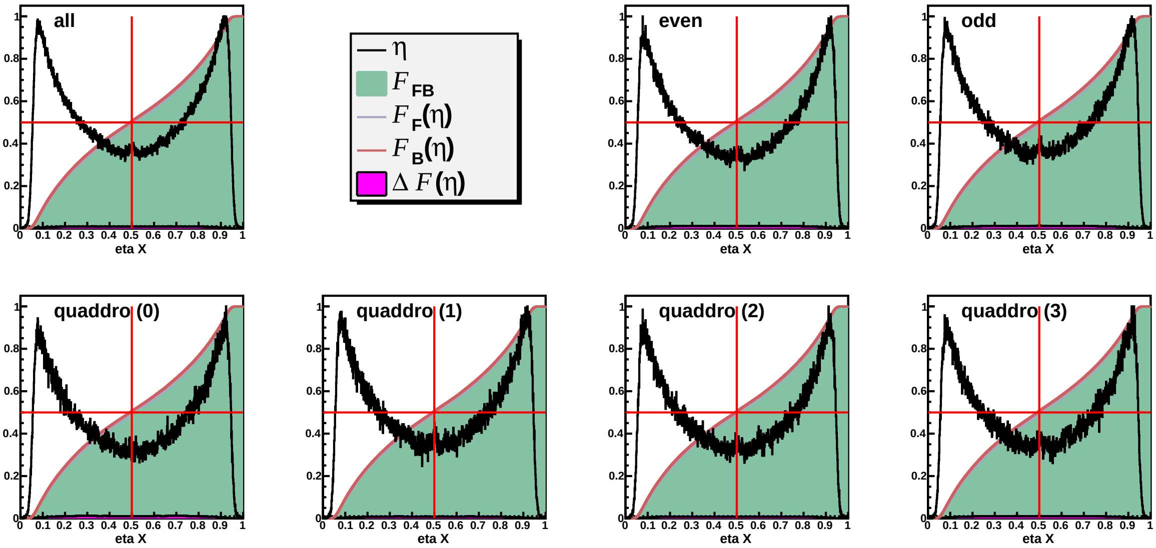


Figure 3.28: *The seven η distributions and $\mathcal{F}(\eta_x)$ functions for the X axis of module 6. For details see figure 3.25.*

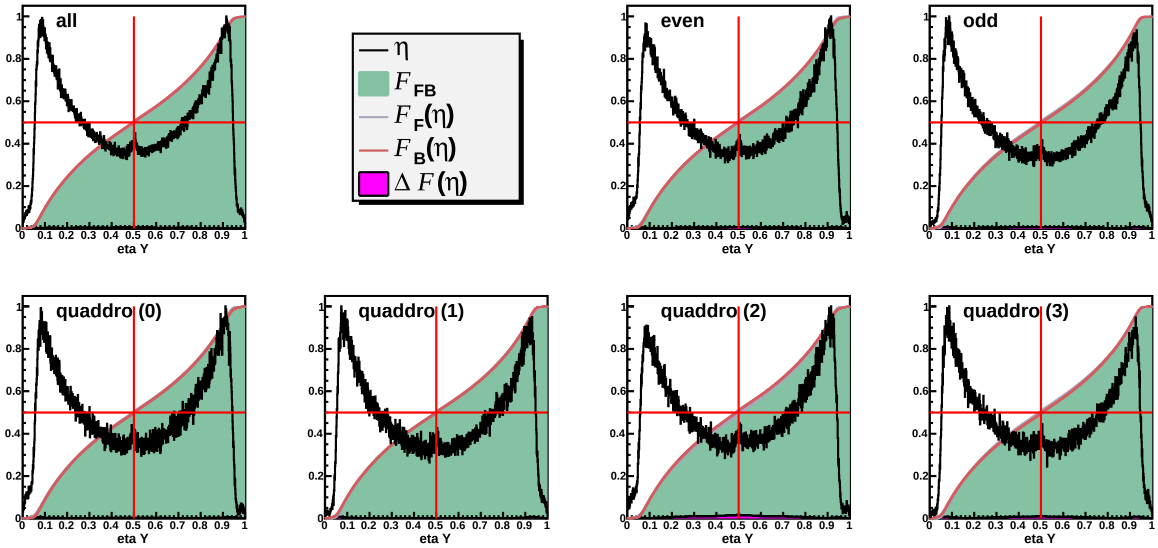


Figure 3.29: *The seven η distributions and $\mathcal{F}(\eta_x)$ functions for the X axis of module 6. For details see figure 3.25.*

3.3.1 Tracking and residual corrections

As shown in the previous section, the ideal condition assumed for the η method, namely uniform pixel illumination and response, are not given in the available data. This will be further investigated within the context of the *in-pixel studies* 3.4. However, also other, additional corrections were applied to further improve the spatial resolution, which is the topic of this section.

Additional Corrections

To further improve the resolution of the DEPFET sensor three additional corrections have been applied:

- gain corrections as described in section 3.2. All residuals shown in this and the previous section are after gain corrections,
- residuals corrections on a sensor scale, and
- residuals corrections based on a row pattern.

Gain Corrections: So far, all the residuals in this section are referring to column and row wise gain corrected data. In table 3.4 standard, single pixel η residuals for uncorrected and for gain corrected data are juxtaposed. Though for the most part the differences are insignificant, module 5 and module 7 show significant improvement in the Y resolution as can be expected from the level of gain correction these modules savored.

Residual corrections on a sensor scale: One common way to check that the alignment of the telescope planes with the tracking software was successful is to plot the residuals against the pixel position. If no error was done the residual distribution should be independent from the pixel position. Although this is the case for most sensor planes module 2 and 14 show a distinct dependence of the residual mean on the Y position (Fig. 3.30 and 3.31). Alignment and pixel size were checked and found to be correct. Furthermore this behavior has been independently observed by other groups using different analysis softwares. To tackle this effect a parabolic function was fitted to the distribution and used as an pixel wise offset correction. The results are shown in the right images of figures 3.30 and 3.31. The improvements for applying such a fit to all sensor planes (both axes) are shown in table 3.4: Only the y axis of module 2 and 14 show a significant improvement.

Residual corrections by row groups: As can be seen in figure 3.25 and will also be shown in more depth in the following section, a distinct pattern repeated either every second or fourth row can be found in the data. Figures 3.32 shows that this is also reflected in the residual distribution with a pattern repeated every fourth row. This correction simply shifts the reconstructed position by the mean value of the residual distribution, based on its quad row membership. Fig. 3.33 shows the results for module 5. It should be noted hat this module shows the worst behavior in terms of this effect and

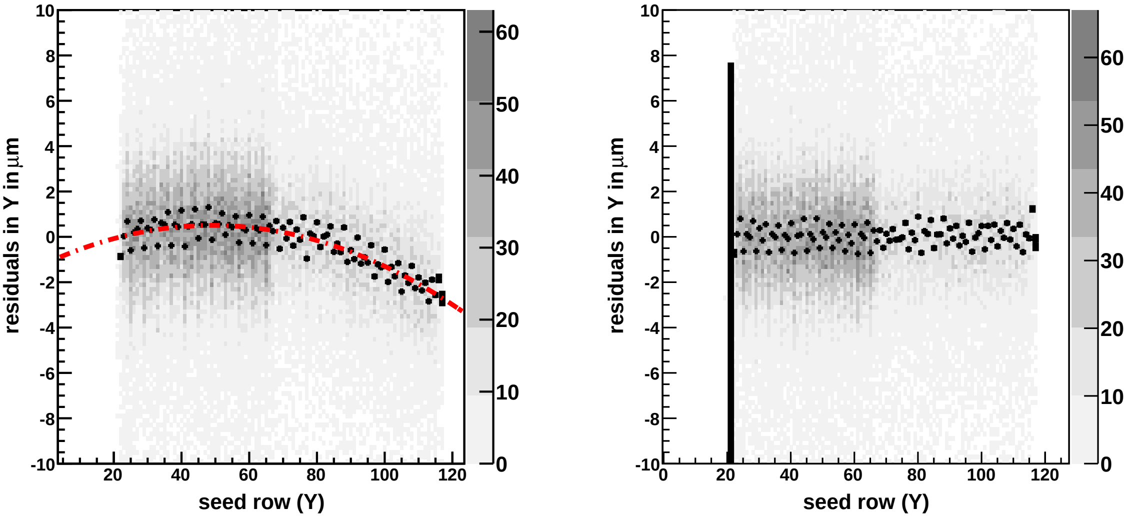


Figure 3.30: *The dependence of the Y axis residuals on the seed row position for module 2. The left picture shows the residuals as a function of the seed row before a correction on sensor scale is done. The black points are the fitted mean values of the residuals, the red line indicates a parabolic fit to the mean values. The right picture shows the same variables after the reconstructed position has been corrected with the parabolic fit shown in the left.*

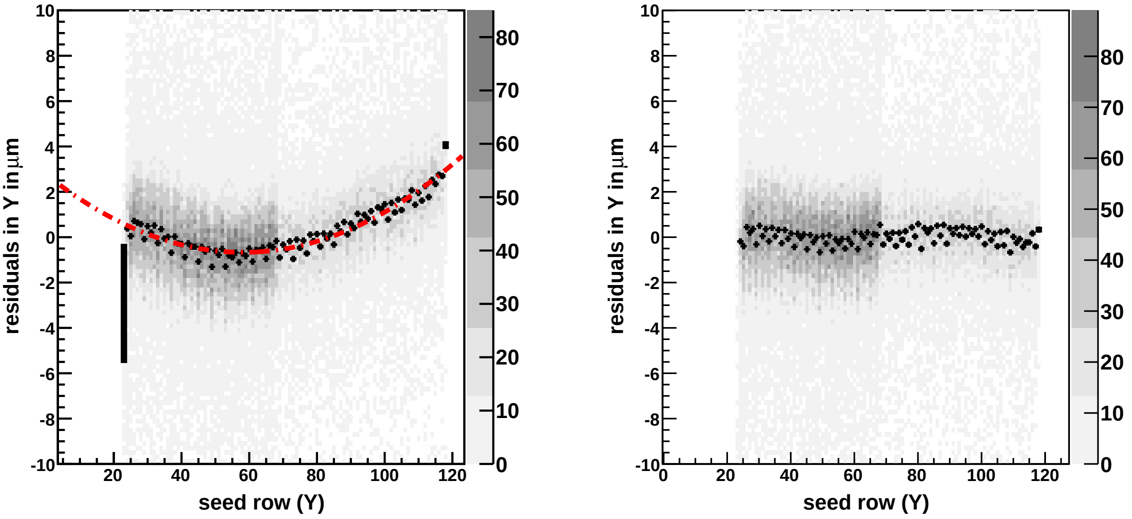


Figure 3.31: *The dependence of the Y axis residuals on the seed row position for module 14. See figure 3.30 for details.*

also the biggest improvement by this correction. The improvements for all other modules are rather marginal.

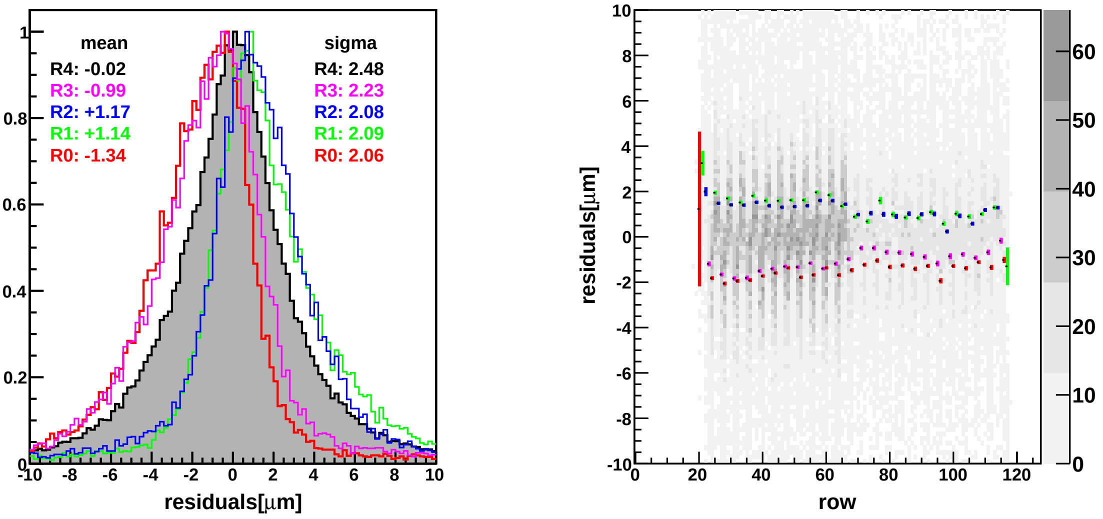


Figure 3.32: *The residual distribution of module 5 separated by rows grouped by their modulo 4 value, e.g. R0 to R3 means* $Y \mod 4 = 0$ *to* $Y \mod 4 = 3$*. The right picture shows the residual distribution as function of the seed row. A distinct pattern becomes apparent. The left plot shows the residual distributions for all rows (solid) and separated by rows grouped by their modulo 4 value, e.g. R0 to R3 means* $Y \mod 4 = 0$ *to* $Y \mod 4 = 3$.

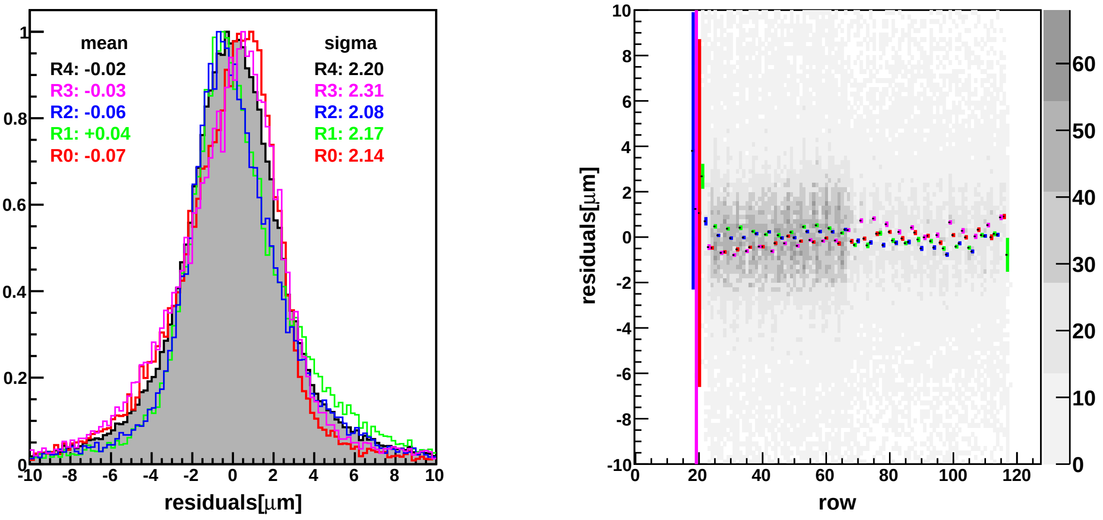


Figure 3.33: *The same as figure 3.32 but now the reconstructed position is shifted by the offset derived from fig. 3.32*

Module	2	14	11	6	5	7
X no corrections	2.63	2.09	2.11	1.96	2.55	2.91
X gain corrections	2.62	2.09	2.12	1.94	2.55	2.93
Y no corrections	2.37	1.88	1.90	1.97	2.84	2.92
Y gain corrections	2.34	1.87	1.87	1.88	**2.45**	**2.76**
Y sensor scale corr.	2.22	**1.63**	1.83	1.88	2.45	2.68
Y group wise corr.	2.19	1.61	1.77	1.78	**2.18**	2.61

Table 3.4: *Summary of the residual and gain corrections on the residual width. All values are in μm and the uncertainty is $\pm 0.05\mu m$.*

3.4 In-pixel studies

One common assumption for silicon (and other) detectors is a homogenous response behavior. The η algorithm for example assumes a homogenous illumination and therefore a homogenous response behavior. A non uniform behavior of a detector both on a larger sensor region as well as on an in-pixel level can be fatal to its usability in a physics experiment. In case of the DEPFET sensor one cannot measure the uniformity of the charge collection efficiency $\epsilon_{charge}(x, y)$ and the gain $g_q(x, y)$ of the DEPFET itself independently. Therefore, the combination of both $S(x, y) = \epsilon_{charge}(x, y) \cdot g_q(x, y)$ is measured and will be referred to as *signal homogeneity.* The uniformity of DEPFET sensors on a large scale has been measured with source measurements to 2.8% to 5% [21]. With the introduction of a gain correction map these variations could be reduced to 1.4% to 2.5%. However, the homogeneity of the signal on an in-pixel scale has not been measured yet. The high statistics of this test beam allowed for the first time to do these measurements.

Seed and cluster signal

In the following the dependence of the seed and cluster signal on the in-pixel position will be examined. *Seed signal* is hereby referring to the mean value μ of a Gaussian fit to the seed signal distribution sampled for a given in-pixel position. *Cluster signal* is referring to the most probable energy loss value of a Landau function fit to the cluster (5×5 pixels) signal distribution sampled for a given position. For statistical reasons X and Y axes will be examined independently, meaning that all signal values are averaged, i.e. projected onto the corresponding axis. The position axes of the histograms have a bin size of $2\,\mu$m with a minimum number of 1450 entries (190k total events and $8 \cdot 16$ bins). The statistic is sufficient for accurate Gaussian and Landau fits. When looking at the seed and the cluster signals on an in-pixel scale two things are to be expected:

1. The seed signal should be high near the pixel center since charge sharing is small and conversely small near the pixel borders where the charge sharing is large. The seed signal should therefore have a peak like behavior with the peak at the pixel center.

2. A 5×5 cluster contains all the charge as has been shown before. Therefore the cluster signal should be independent of the in-pixel position and should be flat.

The DEPFET pixel layout has a double pixel structure and hence the natural choice to study sensor behavior on a μm scale would be the double pixel. However, it turned out that different sensors had different scales of inhomogeneities both in amplitude and in spatial extent. Figure 3.34 shows the seed and cluster signal of sensor 5 as a function of the Y position, i.e. along the SWITCHER axis. At first glance the seed signal behaves as expected peaking at the pixel center. A closer look, however, reveals some form of a pattern repeated every four rows. This pattern becomes very evident when one looks at the cluster signal curve. This sensor has the largest cluster signal inhomogeneities with an R.M.S. of $\sim 8\%$. As has been mentioned above this module showed the worst performance of all modules probably due to biasing conditions.

To confirm that there is indeed a repeating pattern every 8 pixels and that this is not an artifact of the analysis, the same plot was made with an axis of seven pixels (i.e. row number MODULO 7, Fig. 3.35). Indeed here the sensor response looks almost uniform with an R.M.S. of $\sim 1\%$. Other modules show less inhomogeneities (R.M.S. $\sim 3-5\%$), with module 14 showing the best behavior in terms of uniformity (R.M.S. $\sim 1\%$, Fig. 3.38). Also the other modules show a dependence that seems to be connected to the double pixel structure as can be seen in figure 3.37. Along the X axis (CURO side) all modules except for module 5 possess a uniform signal response. Module 5 however shows some non uniformity along this axis based on a double pixel pattern (Figure 3.36). A possible origin for this mysterious X-2 pixels/Y-4 pixels pattern of module 5 will be given in section 3.4 and picture 3.44.

Laser measurements

As mentioned above large scale uniformity measurements of DEPFETs have already been made with a radioactive source. However, this method (at least when using a γ source) lacks the information of the particle's impact point. An alternative method to characterize a sensor uses laser light to create charge in the sensor. There have been several characterization studies of DEPFET matrices using a laser setup in Bonn [18] [19] [20] [21]. This setup allows to shoot a laser beam with position accuracy better than a μm. Hence, complementary to the test beam measurements, in-pixel measurements with this $\lambda = 680\,\text{nm}$ laser were made over a small area of the sensor. However, there are several draw backs with these measurements.

- The energy deposition of a laser is based on different physics than the energy deposition of a charged particle. Light with a wavelength of $\lambda = 680\,\text{nm}$ has a penetration depth of $\approx 4\,\mu\text{m}$ and is virtually almost absorbed in a $450\,\mu\text{m}$ thick silicon sensor. Although this has less an influence on signal uniformity measurements other studies like δ-e$^-$ dependence of position reconstruction methods cannot be done with a laser.

- The charge cloud itself has a somewhat different shape since it is not created along a particle track but rather in one point (at least in a first order approximation for $\lambda = 680\,\text{nm}$ in silicon).

- The laser intensity is not stable and this limits the scannable area.

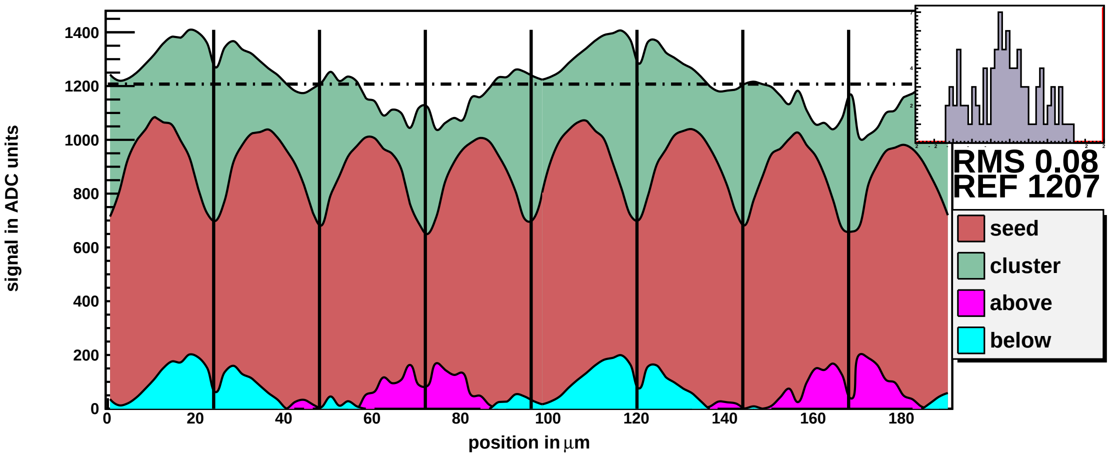


Figure 3.34: *The seed (red) and the cluster signal (green) in ADC units as a function of the tracking position in* ***Y*** *[µm] for* ***sensor 5****. The seed signal is the mean value of a Gaussian fit to the seed pixel signal distribution and the cluster signal is the most probable value of a Landau function fit to the cluster signal distribution. These distributions were sampled and fitted as a function of Y (SWITCHER axis) in steps of two micrometer (bin size* $= 2\,\mu m$*). To increase the statistics all pixels in X (CURO axis) are treated as one. In the Y direction the signals of all pixels with SWITCHER* ***row MODULO 8*** *are overlayed. The vertical black lines indicate the pixel borders of these eight pixels. The horizontal dashed black line is the average of all cluster signals, the turquoise and magenta areas at the bottom of the plot show the deviations of the real cluster signal from this ideal line in absolute ADC unit values. Turquoise means a positive and magenta a negative difference. The small insert at the right shows the one dimensional distribution of the fluctuations of the cluster signals relative to the average value of the cluster signals. The latter is printed below the insert as REF together with the R.M.S. value. As one can see the deviations for sensor 5 are quite substantial (R.M.S.* $\sim 8\%$ *of the ideal cluster signal) and follow a pattern that repeats every four pixels.*

In spite of this drawbacks a very useful measurement could be obtained with a CoCG small sensor with 128 x 128 pixel. The results for seed and cluster signal are shown in figures 3.39 and 3.40. The right plots in figure 3.40 show the seed and cluster signal dependence on the X and Y position similar to figures 3.34 to 3.38. It is evident that there is a cluster signal non-uniformity pattern with a spatial extent of a double pixel. The X (CURO) axis on the other hand seems to be uniform within the uncertainties given by the long time laser stability. (The laser scans line wise from left to right, that means the SWITCHER side is the fast and the CURO side of the sensor is the slow axis).

Effects on residuals

Any inhomogeneities of the pixel gain or charge collection efficiency will lead to an error in the position reconstruction, since both, Center of Gravity as well as η, assume uniform behavior. In the previous sections the row dependence of the residuals has already been shown and even, at least to some extent, corrected for. With cluster signal deviations of up

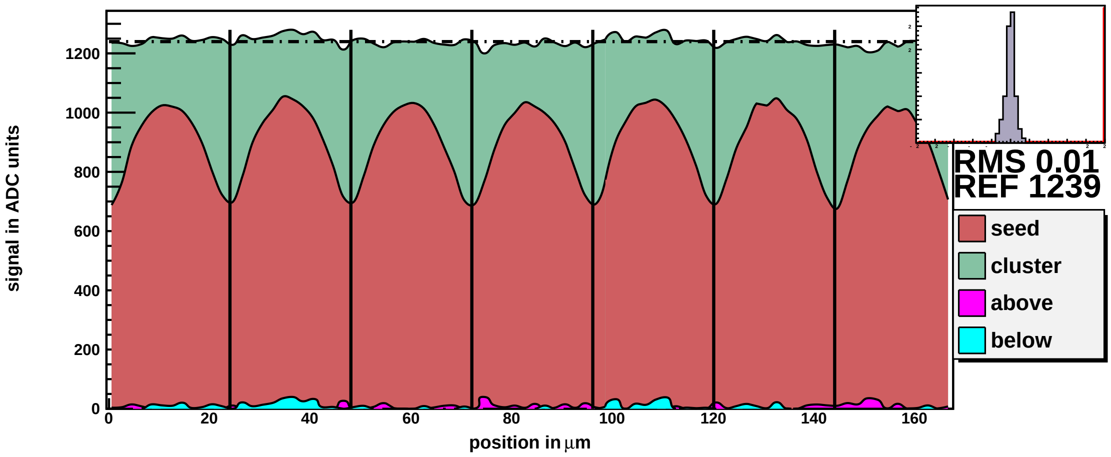


Figure 3.35: *The seed (red) and the cluster signal (green) (ADC units) as a function of the tracking position in* ***Y*** *(µm) for* ***sensor 5*** *similar to figure 3.34. However, this time the seed and cluster signal of pixels with SWITCHER* ***row MODULO 7*** *are overlayed and the vertical black lines indicate the pixel borders of seven pixels. Unlike fig. 3.34 there is no distinct pattern in the run of the cluster signal curve and the fluctuation width is much smaller than in fig. 3.34 (R.M.S. of ~ 1% instead of ~ 8%). This clearly shows that the fluctuations are indeed based on a pattern that repeats every four pixels and is not an artefact of the analysis.*

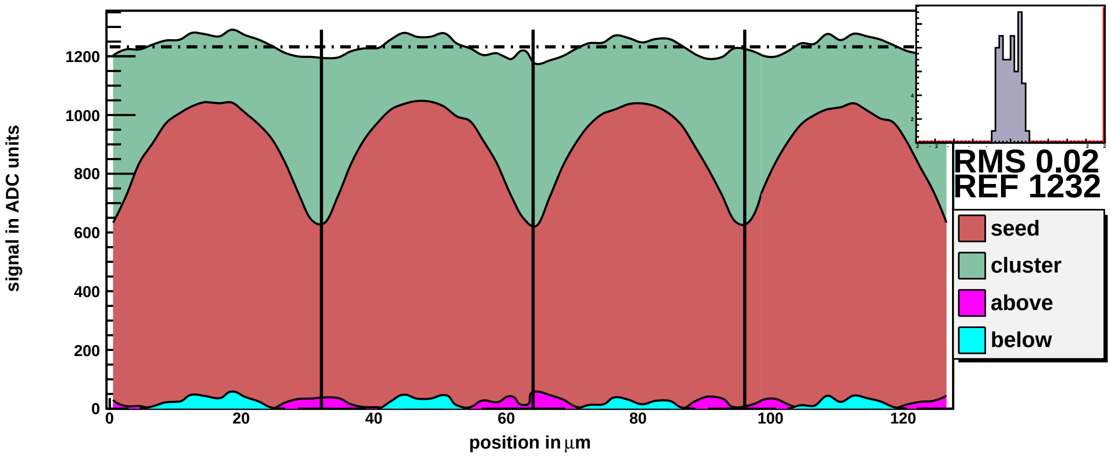


Figure 3.36: *The seed (red) and the cluster signal (green) (ADC units) as a function of the tracking position in* ***X*** *(µm) for* ***sensor 5****. The plot shows basically the same elements that figure 3.34 shows. However, here the seed and cluster signal of pixels are combined on a CURO* ***column MODULO 4*** *base as the signals are plotted against the X axis. The deviations of the cluster signal from an ideal homogeneous behavior is rather small (R.M.S. of ~ 2%) compared to the Y axis of the same sensor (R.M.S. of ~ 8%). No other sensor shows any significant inhomogeneities of the cluster signal along the CURO side of the sensor.*

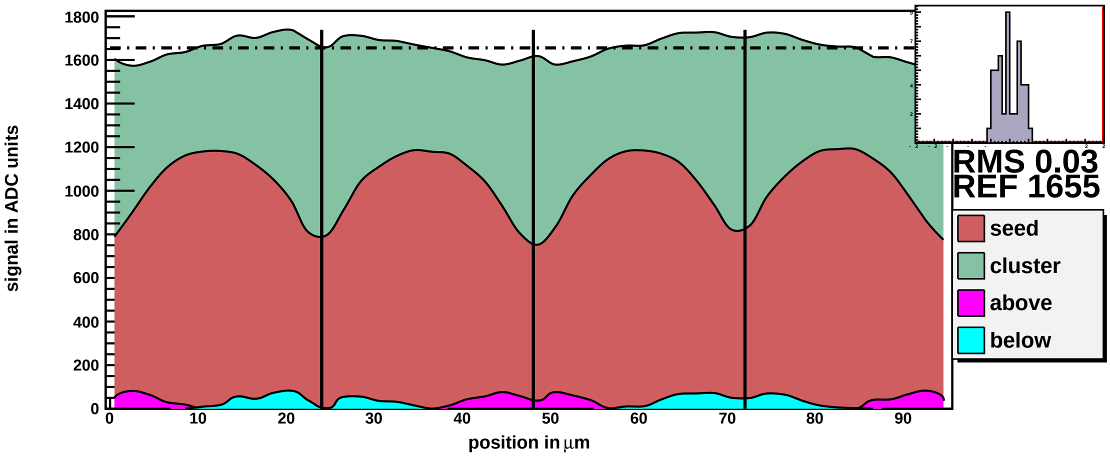


Figure 3.37: *The seed (red) and the cluster signal (green) (ADC units) as a function of the tracking position in* ***Y*** *(µm) for* ***sensor 6*** *similar to figure 3.34. The pixels are combined on a basis of SWITCHER* ***row MODULO 4****.This sensor shows intermediate fluctuations with a R.M.S. value of* $\sim 3\%$ *which has a dependence related to the double pixel sensor layout structure.*

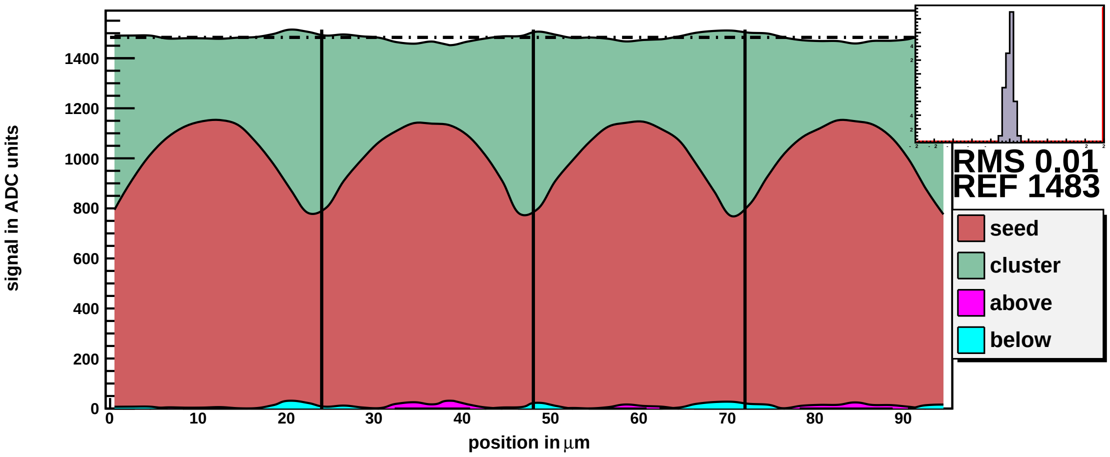


Figure 3.38: *The seed (red) and the cluster signal (green) (ADC units) as a function of the tracking position in* ***Y*** *(µm) for* ***sensor 14*** *similar to figure 3.34. The pixels are combined on a basis of SWITCHER* ***row MODULO 4****. This sensor shows the smallest fluctuations of all sensors with a R.M.S. value of* $\sim 1\%$*. This sensor showed also the most stable performance and possessed the best biasing conditions during the test beam.*

to $\Delta S \sim (10-15)\%$ for sensor 5 residual offsets of up to $\Delta Y_\eta \sim \Delta S \cdot Y_{pitch} \sim 2.5-3.5\,\mu$m can be expected. Figure 3.41 shows the spatial residuals in Y (SWITCHER axis) for sensor 5 as function of the Y (SWITCHER axis) position. Indeed offsets of the mean value from zero of $\Delta\mu \sim 3\,\mu$m can be found. A direct spatial correlation between the residual offsets in figure 3.34 and the cluster signal fluctuations in figure 3.41 cannot be

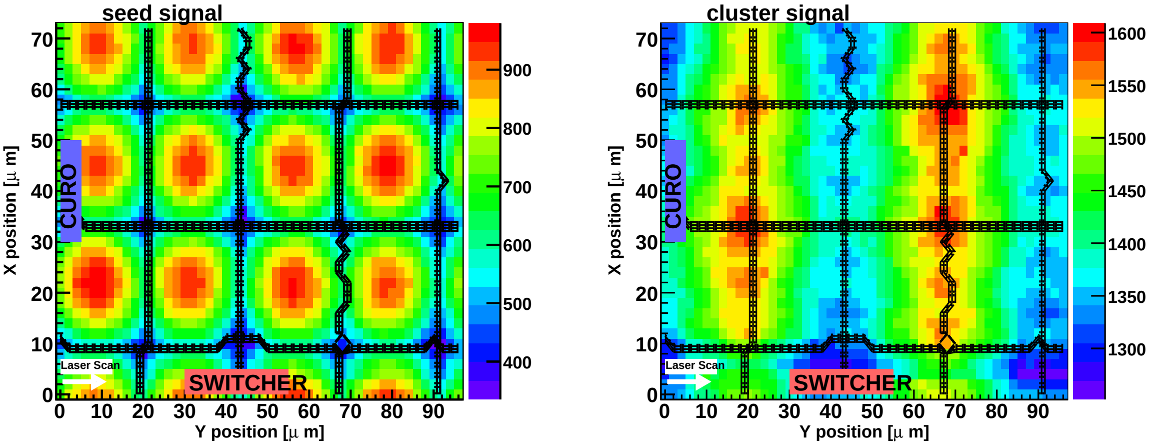


Figure 3.39: *The seed (left) and cluster (right) signal in Z (color scale in units of ADC units) as a function of the position in X and Y (both in μm) for a area of $\sim 3 * 4$ pixels. The signal was created with a $\lambda = 680\,nm$ laser. The sensor has a total of $128 * 128$ pixels and is of the CoCG small layout type ($24 * 24\,\mu m^2$ pixel size) like the sensor 6 of the test beam experiment. The overlayed contours indicate the pixel borders, the white arrow in the bottom left corner the scan direction of the laser beam. The CURO and SWITCHER axis are marked as well. The left image shows clearly the seed maximum at each pixel's center and the minimum at the pixel borders where maximum charge sharing takes place. The right image displays a double pixel feature of the cluster signal dependence similar to figure 3.37.*

expected since the η algorithm uses only two signal pixels and does not map in-pixel inhomogeneities in a straight forward way to in-pixel positions.

However, the residual systematics displayed in figure 3.34 could be due to a software or analysis artefact. To demonstrate that the observed effects are not due to such an artefact, the residuals are plotted against a row group repeated every 7 pixels (i.e. SWITCHER row MODULO 7). The only systematic of the residuals on the in-pixel position left should be on a single pixel scale. The corresponding plot for module 5 is shown in figure 3.35. Indeed, the residual mean value varies on a single pixel scale and the non-uniformity effects on a lager scale are averaged out.
On the other end of the inhomogeneity spectrum is sensor 14. With fluctuations of less than $\Delta S \sim 2\%$ residual mean offsets of $\Delta\mu \lesssim 0.5\,\mu$m can be expected. Figure 3.43 confirms these expectations.

The sensor layout within the DEPFET read out system

Given the deteriorating effect the in-pixel non-uniformities have on the spatial precision of the sensor, it is important to understand the origin of these patterns. While most modules show a double pixel pattern along the Y axis, module 5 showed a pattern that repeats every four rows. At the same time module 5 also possesses a pattern along the X axis that repeats every two columns. In combination with the suboptimal operational

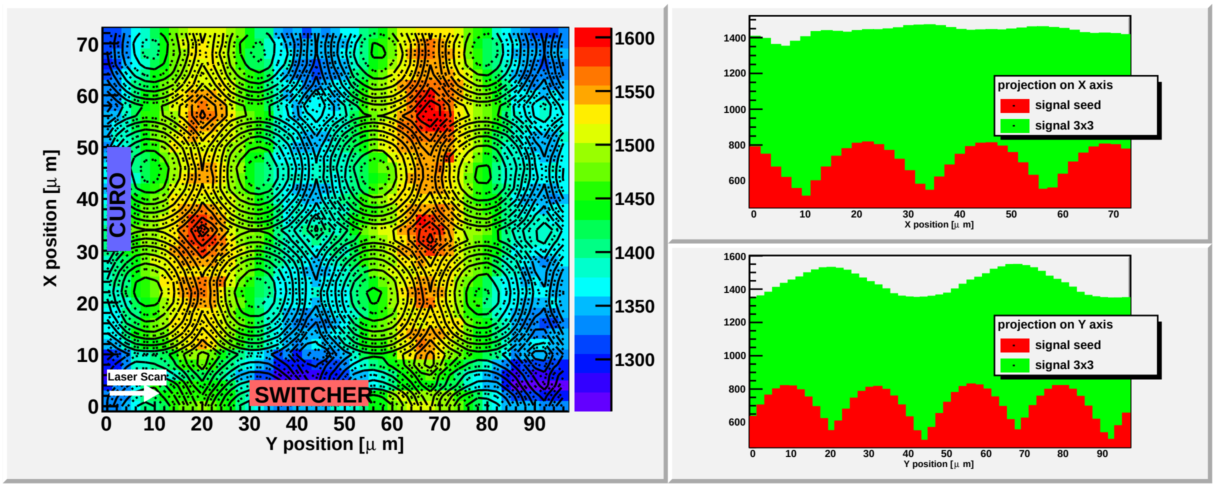


Figure 3.40: *The same laser scan as in figure 3.39. The left picture shows the cluster signal in color scale as a function of the position. Overlayed is the seed signal as contours. In the right images the seed and cluster signals are projected to one axis giving a similar view as fig. 3.34 to 3.38. In the upper pad the seed and cluster signals are plotted against the X position (CURO axis). The fluctuations of the cluster signal along this axis are case by time dependent instabilities of the laser intensity. The fluctuations in the lower pad, which shows the seed and cluster signals vs. the Y position (SWITCHER) confirms the double pixel dependence known from figure 3.37. The R.M.S. value of the cluster signal distribution, if normalized by the average value, is* $\sim 2\%$ *in X and* $\sim 5\%$ *in Y.*

conditions of module 5 during the test beam (including the steering and readout ASICs), picture 3.44 might reveal how this pattern originated:

- There are basically four different pixel types when one also takes read out effects into account. One double pixel comes from the different size of drain and source and therefore different positions of the internal (and external) gate with respect to the geometric pixel center. This structure likely explains the effects seen in all modules expect for module 5.
- Module 5 shows also an X dependence. The connection to a CURO drain is shared by two pixels via the same drain thus giving a quad pixel structure. If this quad pixel structure plays a role in inhomogeneities, so should a CURO column dependence based on a two pixel pattern just as it does in sensor 5.

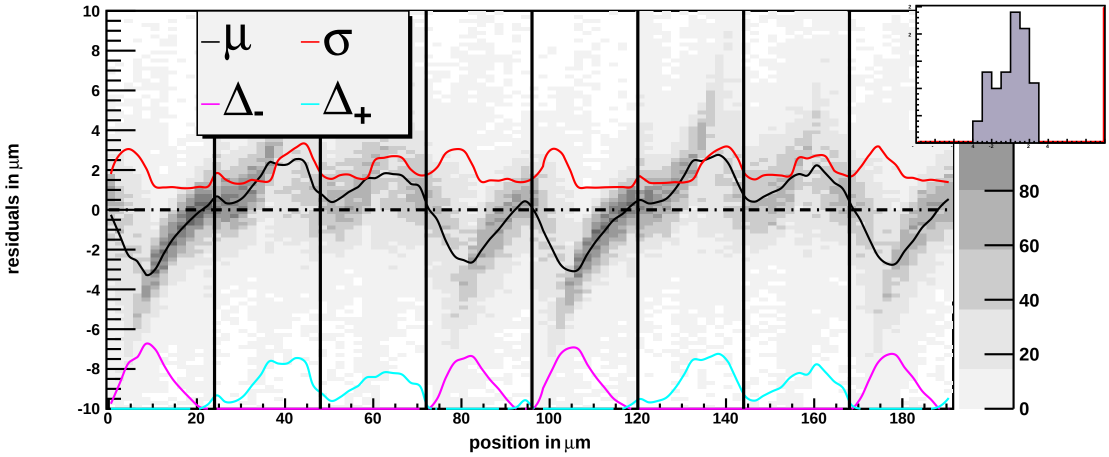


Figure 3.41: *The η residuals as a function of the tracking position in **Y** (μm) for **sensor 5**. In the Y direction the residuals for all pixels with SWITCHER **row MODULO 8** are overlayed. The horizontal black lines mark pixel borders. The grey scale is proportional to the number of entries per residuals $Y_\eta - Y_{tracking}$ (Y axis of this plot) and position along the SWITCHER (Y) axis of the sensor (X axis of this plot, both axes in μm). The black line is the mean value μ and the red line the width σ of a Gaussian fit to the residual distribution of each bin (bin size 1 μm). The horizontal dashed line at zero represents the run of the mean curve for an ideal, perfect position reconstruction algorithm. The magenta and turquoise lines $\Delta_\pm$ at the bottom of the plot indicate the deviation of the mean value from the ideal zero mean line. This deviation shows a distinct pattern repeated ever four pixels analog to the cluster signal value in figure 3.34.*

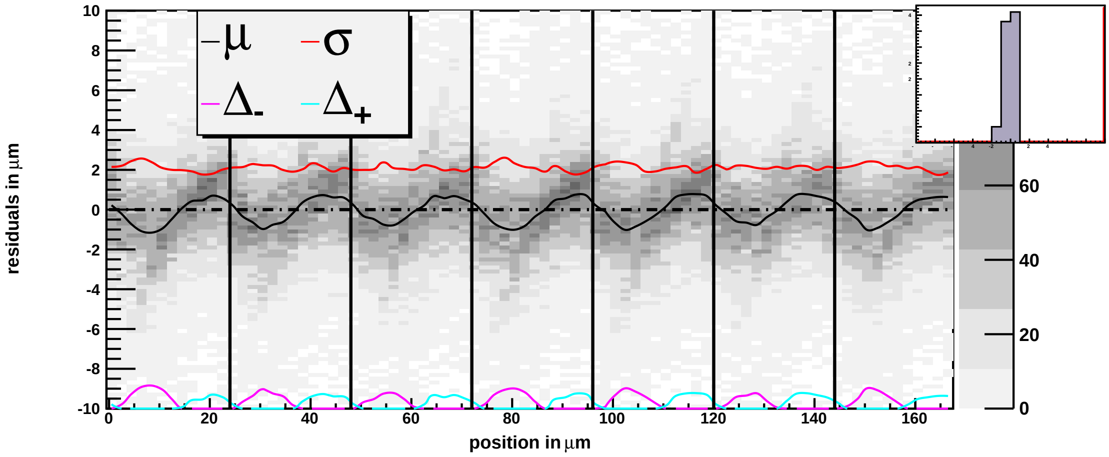


Figure 3.42: *The η residuals as a function of the tracking position in **Y** (μm) for **sensor 5** similar to figure 3.41 but the residual distributions for pixels with SWITCHER **row MODULO 7** are overlayed and the vertical black lines indicate the pixel borders of seven pixels. Now only a single pixel dependence of the mean value deviations (magenta and turquoise lines) can be seen.*

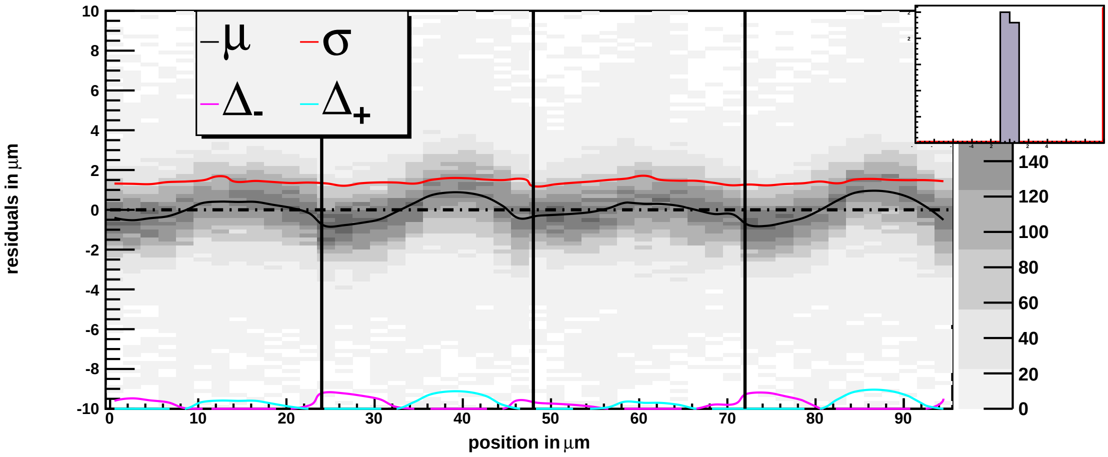


Figure 3.43: *The η residuals as a function of the tracking position in* ***Y*** *(μm) for* ***sensor 14*** *similar to figure 3.41 but the residual distributions for pixels with SWITCHER* ***row MODULO 4*** *are overlayed and the vertical black lines indicate the pixel borders of four pixels. The deviations of the mean of the spatial residuals show a four pixel pattern analog to the cluster signal shown in figure 3.38.*

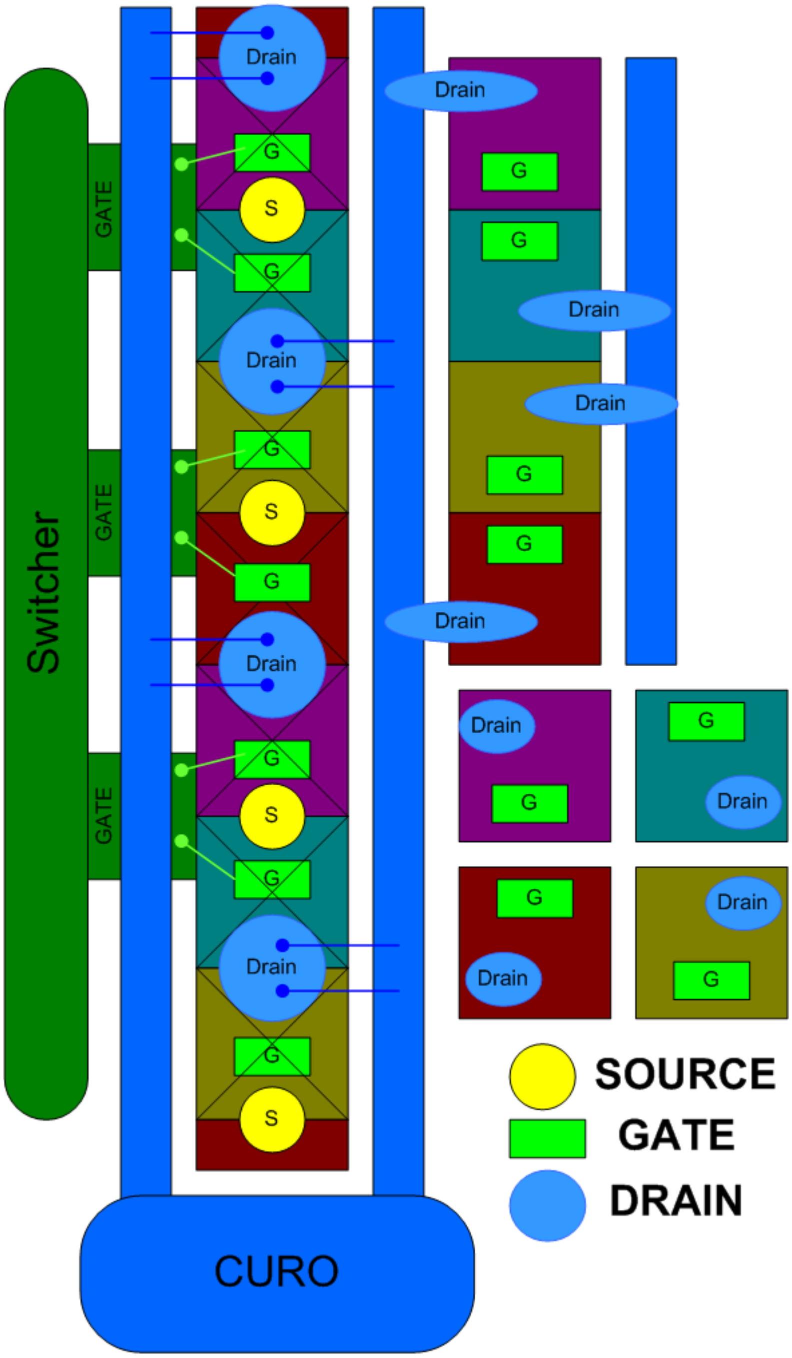


Figure 3.44: *This sketch of the DEPFET sensor layout illustrates, why the in-pixel signal homogeneity and therefore the position reconstruction shows a dependency pattern repeated every four rows: There are basically four different pixel types when one also takes readout effects into account indicated by red, green, blue, and yellow boxes. The differences are the position of the internal gate (magenta circles) with respect to the geometrical pixel center as well as the connection to a common read out (drain). Depending on the bias and general system conditions all four pixel types can show different properties.*

4

Position reconstruction studies

The importance of precise position reconstruction for vertex detectors has been mentioned in chapter 2.1. The DEPFET pixel is a possible answer to the quest for better spatial resolution with a thin vertex detector. However, the quest for better spatial resolution does not only need new ideas in the hardware sector, but also on the software and algorithmic side. This chapter will examine new methods for position reconstruction. An additional emphasis of this study is the influence of δ-electrons on the resolution. As will be shown in the next section the resolution deteriorates with higher energies due to the influence of δ-electrons. This thesis will study possible improvements in position reconstruction precision with three methods:

- **Multiple η distributions** (Sec. 4.1): Instead of using one η distribution separate η distributions are sampled for different energy ranges.
- **A charge cloud based algorithm** (Sec. 4.2): The spatial resolution of the DEPFET telescope allows sampling of the actual shape of the charge cloud with a precision of $\sim 1-2\mu$m. Several analytical approximations to this data will be used for an alternative position reconstruction method.
- **Multivariate methods** (Sec. 4.3): Several multivariate algorithms (neural networks, boosted decision trees, linear discriminant, and probability density estimator) will be used to reconstruct the particle position using a preselected set of input variables.

Two data sets are used for this study. First, the test beam data presented in the last chapter, and second a GEANT-4[1] [51] simulation using the test beam geometry. the latter one has the advantage of not being affected by multiple scattering or telescope resolution effects as the particle position is exactly known. Also, there are no inhomogeneities on neither sensor nor in-pixel scale.

Energy dependence of spatial residuals

The general feasibility of alternative methods for position reconstruction with semiconductor detectors is not the only aspect studied here. The foundations of particle detection

---

[1] Geant4 is a toolkit for simulating the passage of particles through matter.

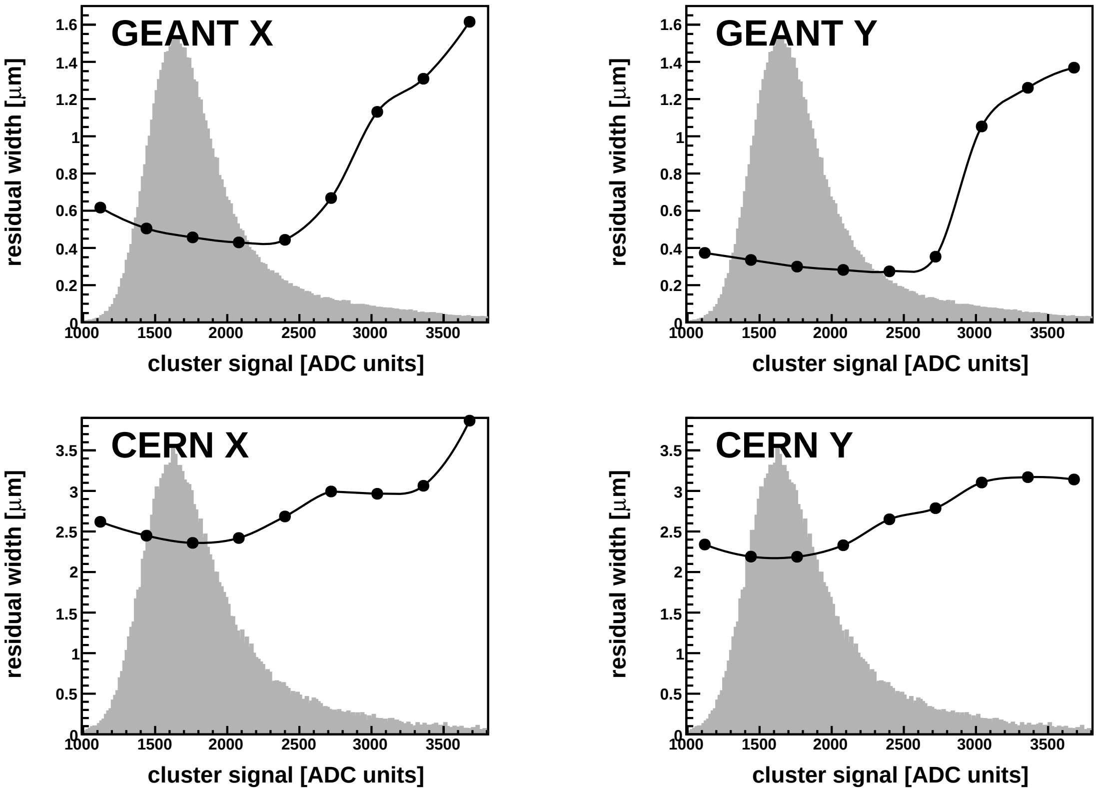
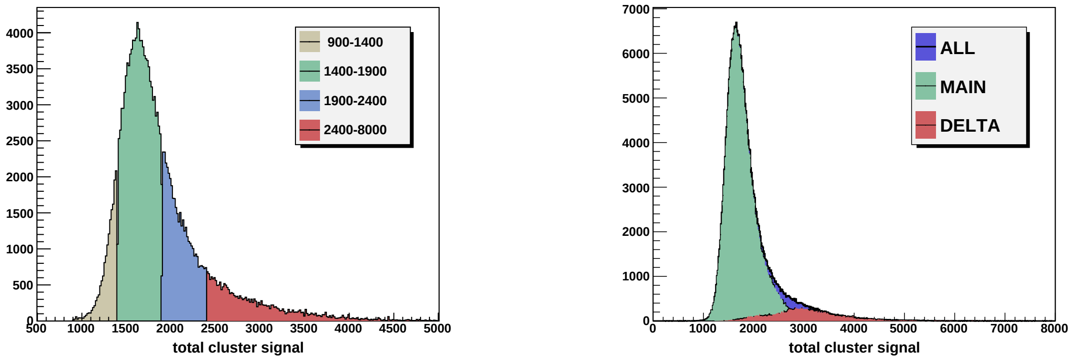


Figure 4.1: *The σ width (Y-axis in μm) of a Gaussian fit to the residual distributions as a function of deposited energy (X-axis in ADC units). The position reconstruction method is the η method as described in the previous chapter. As a references the Landau like energy distribution is plotted in the background. For each* 320 *ADUs a Gaussian was fitted to the residuals. The upper two plots show the results for the GEANT simulation, the lower two plots for the test beam data gathered with module 14 at CERN. As expected the residuals deteriorate considerably for high deposited energies. This is more pronounced for the simulation as the residuals are not affected by multiple scattering and telescope resolution.*

Designation	0	A	B	C	D
energy [ADU]	0-8000	900-1400	1400-1900	1900-2400	2400-8000

Table 4.1: *The division of the cluster energy distribution into four segments (A-D) with the according letter designations.*

with thin semiconductor sensors has been explained in chapter 1. One of the results was that the relative influence of δ-electrons increases with thinner sensors. The conventional way to treat δ-electron fluctuations is a simple cut on the energy deposited inside the sensor, which is usually twice the most probable value. However, since future vertex detectors will be very thin ($\sim 50 - 100\,\mu$m) a simple cut might not be enough if sufficient tracking efficiency is to be retained. Therefore, the other aspect studied in this chapter is the influence of δ-electron fluctuations on the resolution and possible ways to tackle



(a) signal selection

(b) δe- selection

Figure 4.2:* *Left 4.2(a): *The division of the cluster energy distribution into four segments illustrated with test beam data from module 14. The ranges for the cluster signal slices are indicated in the legend.* ***Right 4.2(b):*** *The secondary particle selection criteria for the advanced position reconstruction study. This selection criteria is not available for test beam data.* ***all*** *is refereing to all events,* ***main*** *to events without secondary particles, and* ***delta*** *to events with at least one secondary particle.*

this influence. Figure 4.1 shows the dependence of the residual width on the deposited energy for test beam data and simulated data. As expected the spatial resolution deteriorates with higher energies. To facilitate the investigation the signal spectrum of both, the test beam and the simulated data, is segmented into four energy regimes as displayed in figure 4.2(a). The division has to be limited to four parts to leave enough statistics in each segment for training of the multivariate methods[2]. The GEANT simulation furthermore provides information about the presence of δ-electrons with an energy of $\geq 3\,\text{keV}$[3]. This secondary particle flag is also used to split the simulated data into δ-electron and non-δ-electron events as shown in figure 4.2(b).

4.1 Multiple η distributions

This method uses multiple η functions, each for a designated energy range, to improve the resolution for higher energies. To distinguish this method from the standard single η method, two designators will be used, η_{one} and η_{multi} for a single and for multiple η distributions, respectively. Since the η method takes the ratio of pixel signals, any scaling of a Gaussian shaped charge cloud is canceled, making this a rather robust method with respect to energy fluctuations. However, the smaller the sensor the more influential are fluctuations of the charge cloud due to δ-electrons. One possible way to tackle this problem are multiple η functions, each sampled for a distinct range of the signal spectrum. With

---

[2] The multivariate methods need to be trained first, e.g. a neural network "learns" by being fed with a sufficient amount of training events.

[3] With the parameters of the DEPFET simulation, δ-electron below 3 keV are not simulated as a separate particles. Therefore they are not marked in a GEANT event.

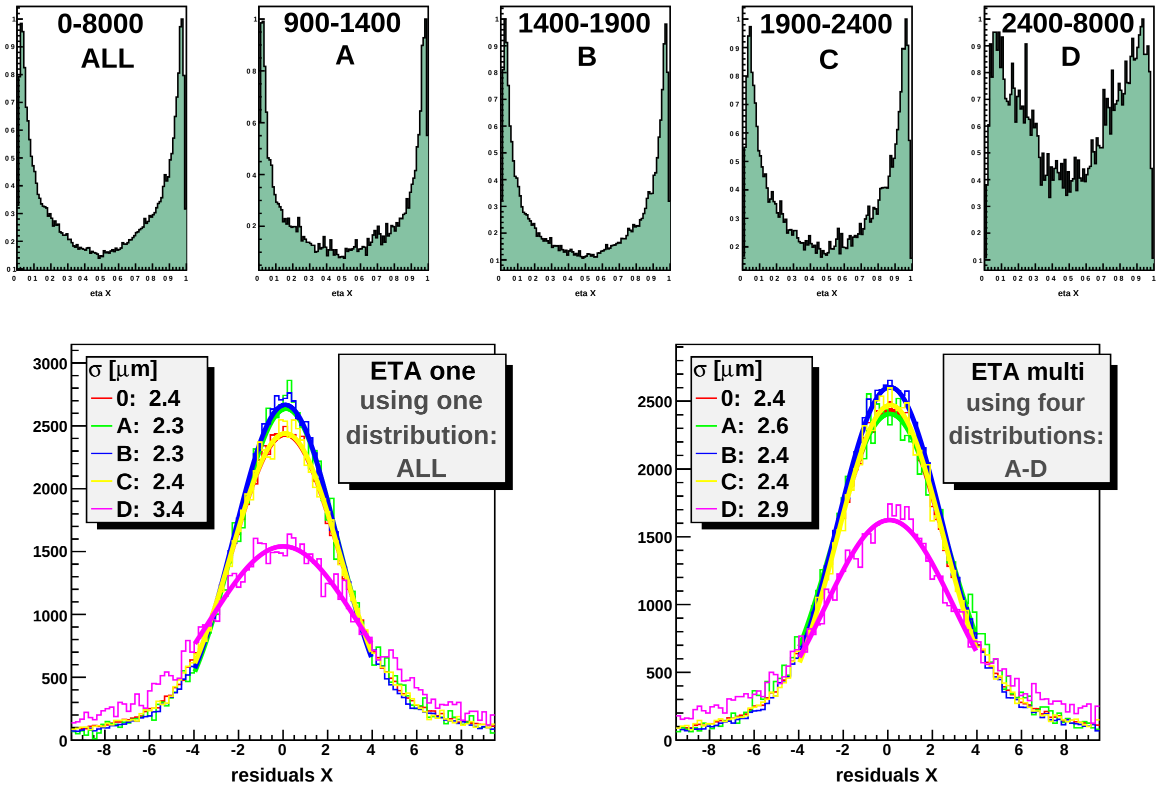


Figure 4.3: *The top row shows the η distributions of module 14 in X (32 μm pixel pitch) for the cluster signal ranges listed in table 4.1. The bottom row shows the residual distributions [μm] in X for these cluster signal range, i.e. five residuals per plot for energy range* ***0*** *to* ***D****. For the left bottom image the standard single η_{one} distribution method was used. In the right bottom plot the η distributions shown above are used for the respective energy ranges, i.e. the η_{multi} method was applied. A Gaussian distribution is fitted to each residual distribution within the range of -4 to 4 μm as indicated by the thick, colored lines. The widths of these fits are shown on the left side of each plot. The residuals are in μm units.*

the energy ranges shown in table 4.1 the η_{multi} method now consists of four separate η distributions η_A to η_D, whereas η_{one} is synonymous to η_0. Figure 4.3 shows the η distributions for module 14 using the energy regimes listed in table 4.1. The energy ranges **A**-**C** that are in the Gaussian like part of the signal distribution (fig. 4.2(a)) have similar η distributions (η_A to η_C). The high energy tail of the Landau distribution **D** on the other hand shows an η_D distribution distinct from the three ranges **A** to **C**. Unlike η_A to η_C it also differs from the η_{one} distribution gathered from all events. This is to be expected as the high energy part of the spectrum is of statically marginal importance. Figure 4.3 also displays the residual distributions for each energy regime, both, for a single η_{one} distribution and η distributions separated by pulse height. A separate η distribution improves the resolution only for the high energetic Landau tail. Since this is also the only η distribution that differs from the general η_{one} distribution this result is expected.

The same study was undertaken with the simulated data set yielding similar results (fig.

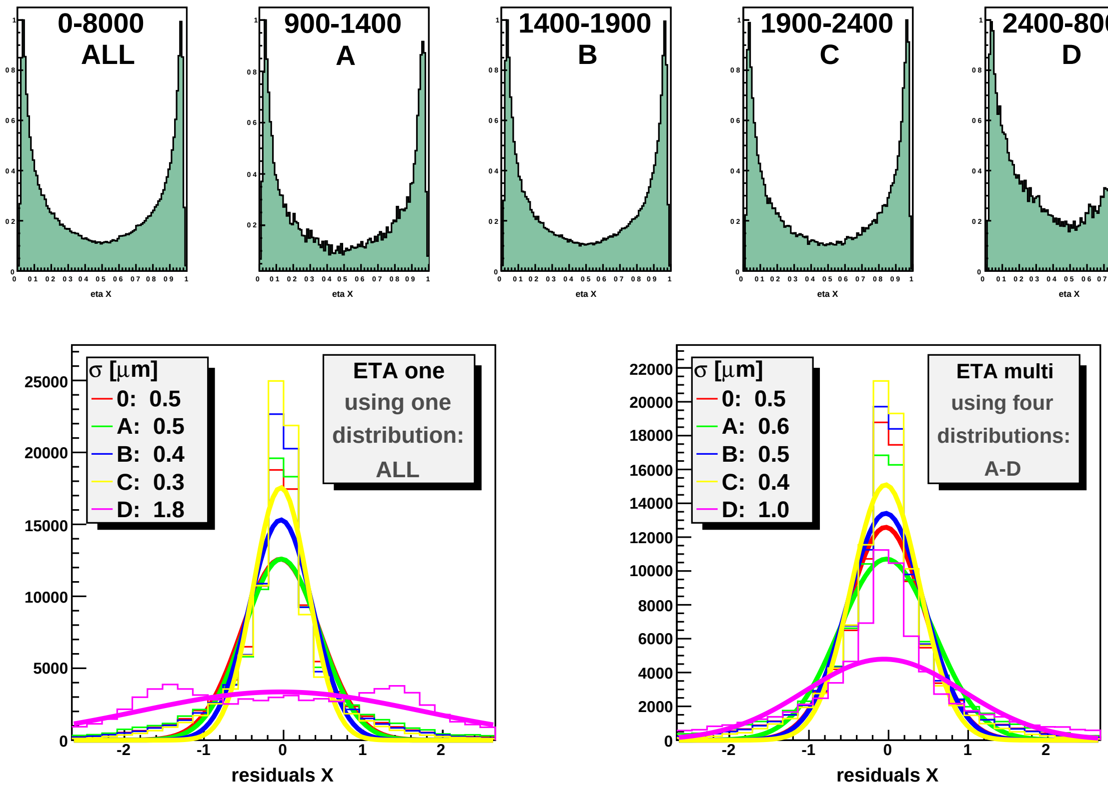


Figure 4.4: *This figure is similar to fig. 4.3. The top rows show the η distributions along the X axis and the two bottom plots show the residual distributions along the X axis. However, the underlying data consists of simulated events for a pixel with CoCG large dimensions (i.e. 32 μm in X). The residuals are in μm units and the error of the residual width fit is $\sim 0.1\,\mu$m.*

4.4): For the Gaussian part centered at the most probable value multiple η_{multi} distributions make no significant improvement or might even worsen the spatial resolution due to lower statistics of the η_{multi} distributions. In the high energy tail, however, a separate η_D function can improve the residuals drastically. Table 4.2 shows the residual widths for all six modules and the GEANT simulations for CoCG small and large sized pixels using a single and multiple η_{multi} distributions, confirming the above said. The improvements along the larger X axis (32 μm) are usually stronger than along the shorter Y axis (24 μm). Two possible explanations for this are:

- The charge cloud dimensions are the same in both directions. The neighbor of the seed pixel along the longer (x) axis contains therefore - relatively seen - less charge and therefore fewer signals than the neighbor along the short axis. The same signal fluctuation has therefore a larger impact on the η distribution along the larger pixel side and a separate η distribution will have a larger impact on the residual width.

- The most probable value of module 14 is $\sim$ 1500 ADU. The energy range **D** with 2400 to 8000 ADU would therefore include δ-electrons with roughly twice to several times the most probable value energy of $\sim$ 130 keV. As has been shown in chapter

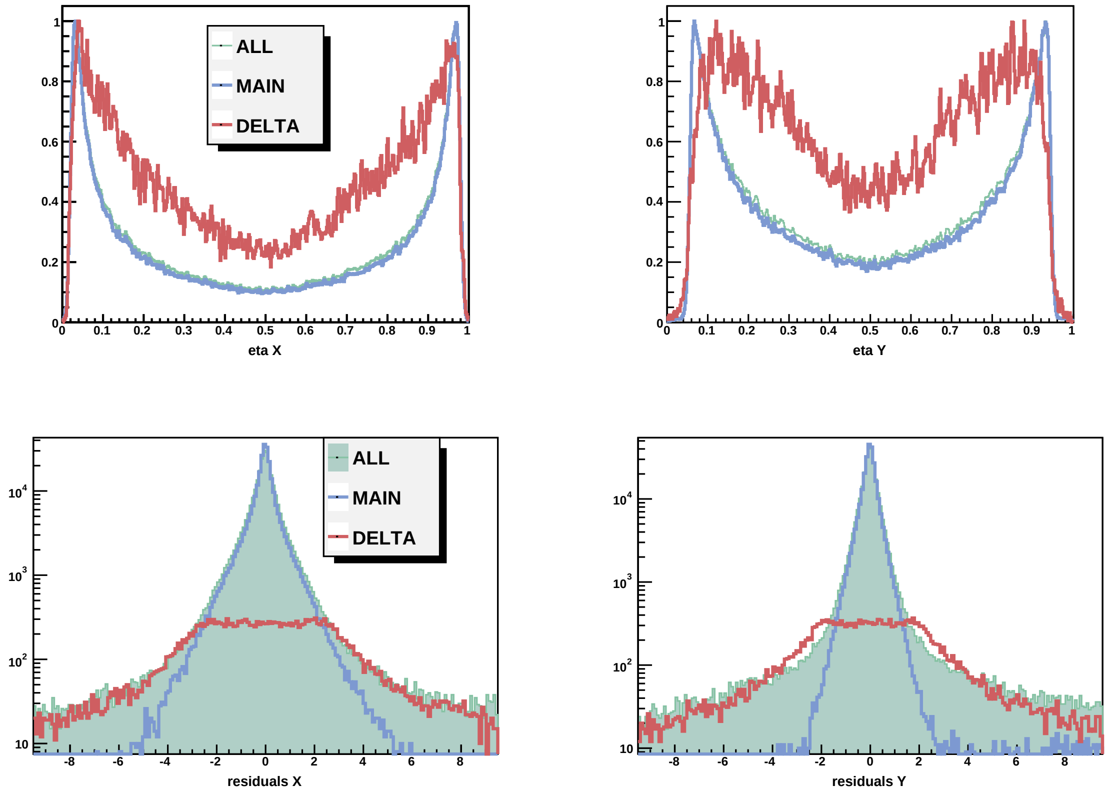


Figure 4.5: *The η distributions (top row, left: X direction, right: Y direction) and the according residuals (bottom row) for the selection criteria only available for simulated events:* ***All*** *is refereing to all events,* ***main*** *to events without secondary particles, and* ***delta*** *to events with at least one secondary particle. The residuals are plotted with a logarithmic scale on the Y axis to illustrate the non Gaussian residual distribution for events with secondary particles (****delta****). The particle position was reconstructed using only a single η_{one} distribution.*

1 (e.g. fig 1.10) that these electrons will deposit more energy in the neighbor with the larger pixel dimensions.

The simulated data also allows the separation of events with and without secondary particles. Figure 4.5 shows the corresponding η distributions and residual width for the 32 μm pitch (X axis) and for the 24 μm pitch (Y axis). The η distributions are distinctly different for the two event types. The residuals of events with secondary particles show a clear non Gaussian distribution (for the residuals only a single η_{one} distribution was used).

The upshot of this study is the recommendation to split the sensor signals into two categories, a Gaussian part around the most probable value and the high energy tail. Although this is common practice for semiconductor detectors, events belonging to the high energetic part are usually discarded. The novelty is to have a separate η distribution which significantly improves the resolution for these events.

	X axis					Y axis				
	0	A	B	C	D	0	A	B	C	D
Mod 2 η_{one}	2.4	2.3	2.3	2.4	3.4	2.3	2.3	2.1	2.3	2.9
Mod 2 η_{multi}	2.5	2.6	2.4	2.4	2.9	2.3	2.3	2.2	2.3	2.8
Mod 14 η_{one}	1.9	1.7	1.7	2.2	3.2	1.8	1.7	1.7	2.0	2.6
Mod 14 η_{multi}	1.9	1.9	1.8	2.0	2.4	1.8	1.8	1.7	2.0	2.4
GEANT η_{one}	0.5	0.5	0.4	0.4	1.8	0.3	0.3	0.3	0.2	1.3
GEANT η_{multi}	0.5	0.6	0.5	0.4	0.8	0.3	0.4	0.3	0.3	0.4

Table 4.2: *Residual width in* μm *using one single* η *distribution for all energy ranges* (η_{one}) *and multiple* η *distributions* (η_{multi}). *This means that, for example, the number in cell "Mod 14* η_{one}*, X axis D" describes the residual width for all events in the range of* $(2400-8000)\,ADUs$ *using one* η *distribution collected with all events,* $(0-8000)\,ADUs$. *In contrast, the residual width in cell "Mod 14* η_{multi}*, X axis D" is gathered with a* η *distribution collected exclusively from events in the energy range of* $(2400-8000)\,ADUs$. *All pixels have a X dimension of* $32\,\mu m$ *and a Y dimension of* $24\,\mu m$. *The residuals are evaluated for the energy range indicated in the second row (see table 4.1.* ***The error on the width is*** $0.1\,\mu m$**.**

4.2 A charge cloud based algorithm

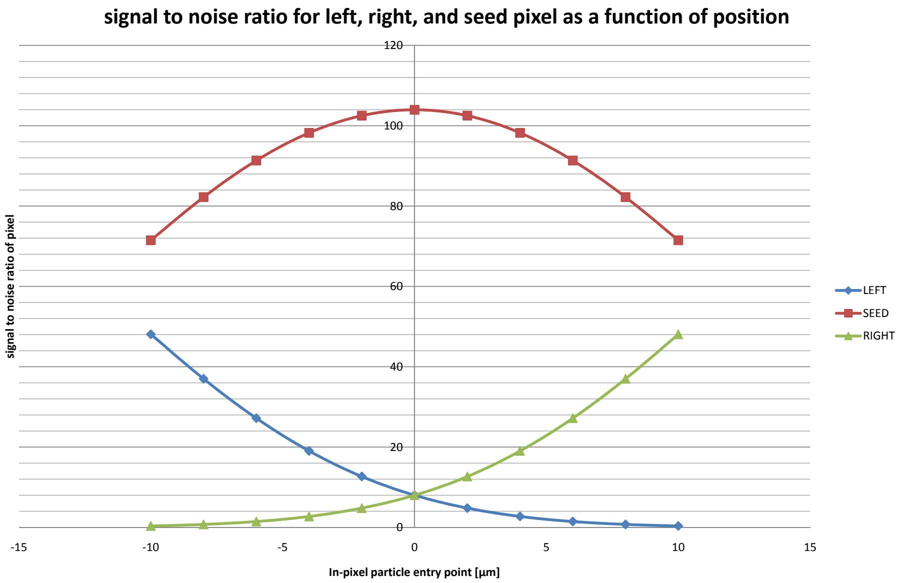


Figure 4.6: *Expected signal to noise ratio of the left, right, and seed pixel as a function of the particle's in-pixel position. The cluster signal to pixel noise ratio is fixed to 120, the pixel pitch is* $24\,\mu m$ *and the charge cloud width is* $8\,\mu m$.

The second method uses an approximation of the total charge cloud and the signals of the three central pixels to fit the particle position. The η method maps the charge sharing between the two strips or pixels with the highest signal and uses the integral of this charge sharing distribution as a correction function. With an estimated charge cloud width of $\sim 7-9\,\mu$m, a pixel pitch of $32\,\mu$m and $24\,\mu$m, respectively, and a cluster signal to pixel noise ratio of $\sim 120-130$, more than two pixels are expected two contain some information about the particle position. Figure 4.6 illustrates the expected signal to noise ratio as a function of the in-pixel position. An additional third pixel can increase the amount of information about the particle's impinging point, especially in the case when the entry point is in the central parts of the seed pixel. The charge cloud based position fit works in three steps:

1. **Sampling the charge cloud shape** (Sec. 4.2.1): A (two dimensional) image of the charge cloud is sampled from the data using the tracking position and the signals from the inner three pixels. Next the image is approximated to a one dimensional distribution by fitting a Gaussian to slices along the Y axis.
2. **Analytical fit** (Sec. 4.2.2): For simplicity and to speed up the algorithm a simple analytical function is fitted to the one dimensional charge cloud distribution.
3. **Position Reconstruction** (Sec. 4.2.3): The function from step two is the basis for a least square fit of the particle position using the signal of the seed and its two neighbor pixels.

This algorithm works with the assumption that a simple charge cloud shape approximation is sufficient. Since the charge cloud shape is expected to be different for higher energies the energy dependency studies shown in the previous section will be carried out as well, i.e. multiple charge cloud approximations will be used.

4.2.1 Sampling the charge cloud shape

The high statistics ($\sim$ 180 k events) and the good telescope resolution ($\sigma_{res} \sim 2\,\mu$m) available allowed precise sampling of the charge cloud. Since the charge is virtually confined to the inner 3×3 pixels of the cluster only these were taken into account. The mapping algorithm was implemented in three steps:

1. **Signal normalization:** The signals of each of the inner 3×3 pixels are normalized by the total signal of these nine pixels.

$$S^{norm}_{x,y} = \frac{S_{x,y}}{\sum_{a=-1,b=-1}^{a=+1,b+1} S_{a,b}} \tag{4.1}$$

2. **Signal - position charting:** For each axis a two dimensional histogram that provides an image of the charge cloud is created. For every event the signals of the seed pixel, S_c, and its two neighbors, S_l and S_r, are charted in the histogram. Their positions are calculated with,

$$p_l = -pitch + pos_{track} \tag{4.2}$$

$$p_c = pos_{track} \tag{4.3}$$

$$p_r = +pitch - pos_{track}, \tag{4.4}$$

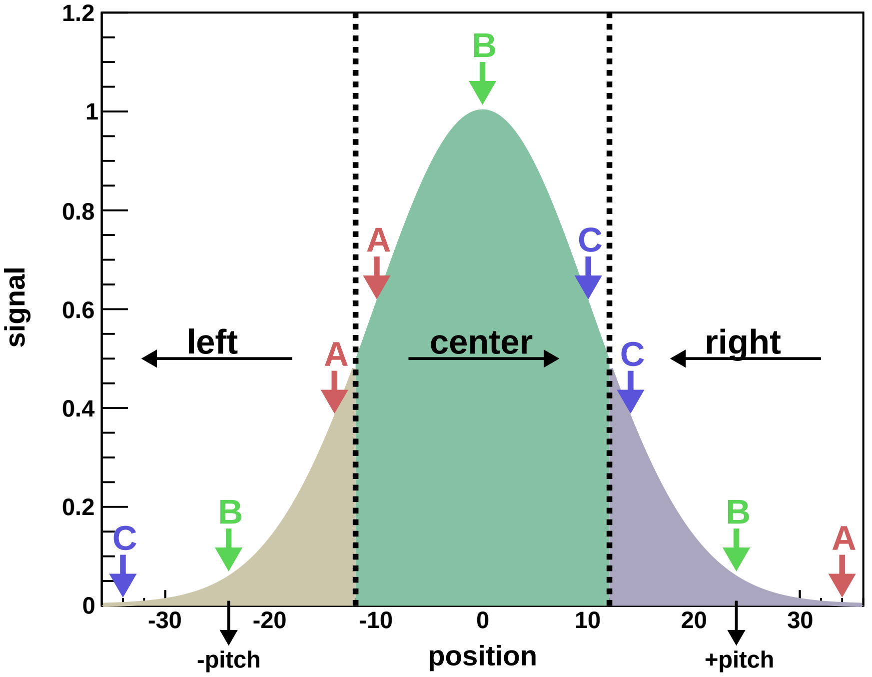


Figure 4.7: *Illustration of the position calculation for the three pixels involved in the charge cloud sampling (eqn. 4.2) using three in-pixel positions A, B, and C. For charting the seed signal A, B, and C can be used without any changes. However, the position in the pixel to the left is not simply the in-pixel position minus the pixel pitch: In this case the signals for position A, B, and C would be charted mirror-inverted. The same is true for the pixel t the right. Hence, the position variation inside the left and the right pixel have the opposite orientation as the central pixel* ($+pitch - pos$)*. However, since the right pixel is at the negative axis side, an additional sign change happens* ($-pitch + pos$)*.*

where $pos_{track} = [\frac{-pitch}{2}...\frac{+pitch}{2}]$ is the in-pixel tracking position and *pitch* the pixel pitch in this direction. This is illustrated in figure 4.7. With sufficient events this generates an image of the charge cloud. The position axis of the histogram has a bin size of $2\,\mu$m.

3. **Gaussian fit:** Each position bin has a one dimensional distribution of the normalized signal. A Gaussian fit is applied to each in position value thus effectively resulting in an average charge cloud approximation in a one dimensional histogram.

The charge cloud was sampled separately for the data selection criteria depicted in picture 4.2. Figure 4.8 and 4.9 show the two dimensional histograms in grey scale and the mean value μ and width σ of the Gaussian fit to each position bin for some of the data selections.

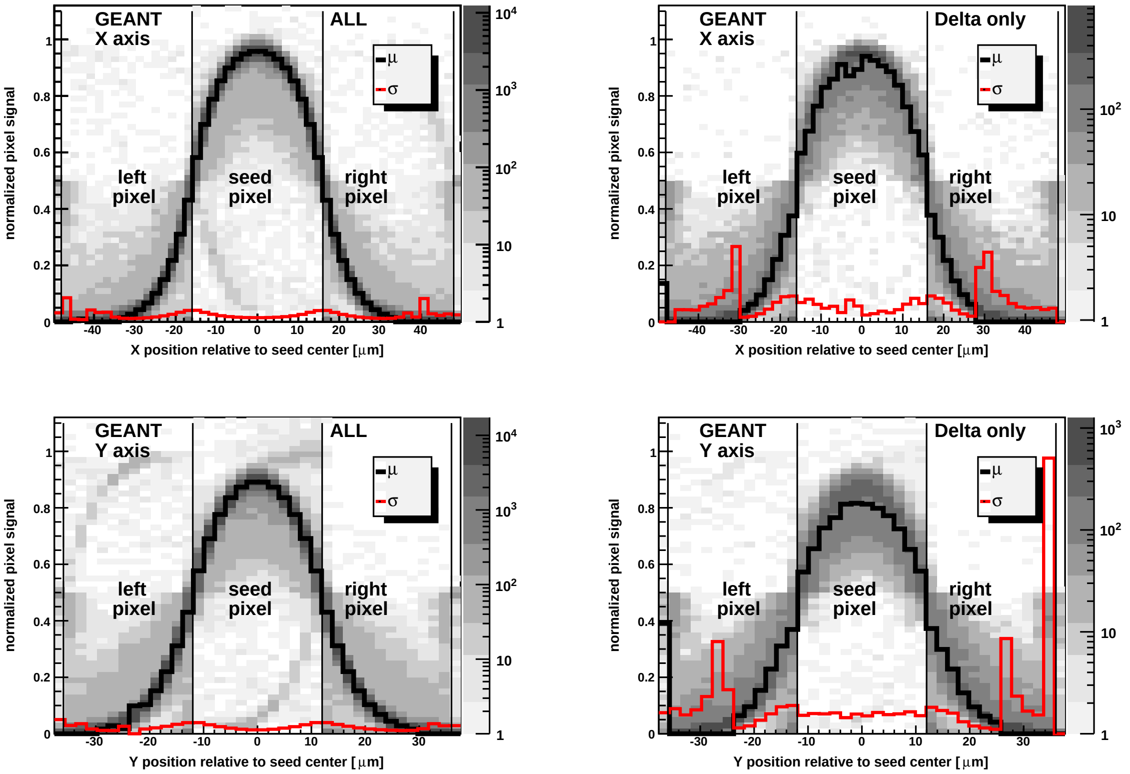


Figure 4.8: *The charge cloud (in grey) sampled for the GEANT simulation, top row in X direction, bottom row in Y direction. The left plots are acquired using all events. In the right plots the δ-electron flag of the GEANT simulation was used. Here, only events with δ-electrons are used. Overlaid are the results of the Gaussian fit to each $1\mu m$ wide bin: The mean value μ in black and the width σ in red. At the outer parts of the charge cloud's wings the fitting routine sometimes encountered problems due to non-gaussian signal distribution in those position bins. In a few instances, when a particle hits very close to a pixel border, the seed pixel assignment can be off by a pixel. This results in the "ghost arcs" seen in the left images. However, these events are rare and this effect is only visible on a logarithmic scale.*

4.2.2 Analytical fit

Although it would be possible to use the sampled charge cloud itself for position reconstruction several analytical functions were fitted to the one dimensional histogram. This was done for algorithmic reasons and to reduce the effect of statistical fluctuations. The position is predicted via a least squares fit of the three pixel signals to the function. The functions to model the charge cloud are generally based on a Gaussian shaped charge cloud and are shown in figure 4.10:

1. **"F1":** A single function following the normal distribution

$$F1(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \tag{4.5}$$

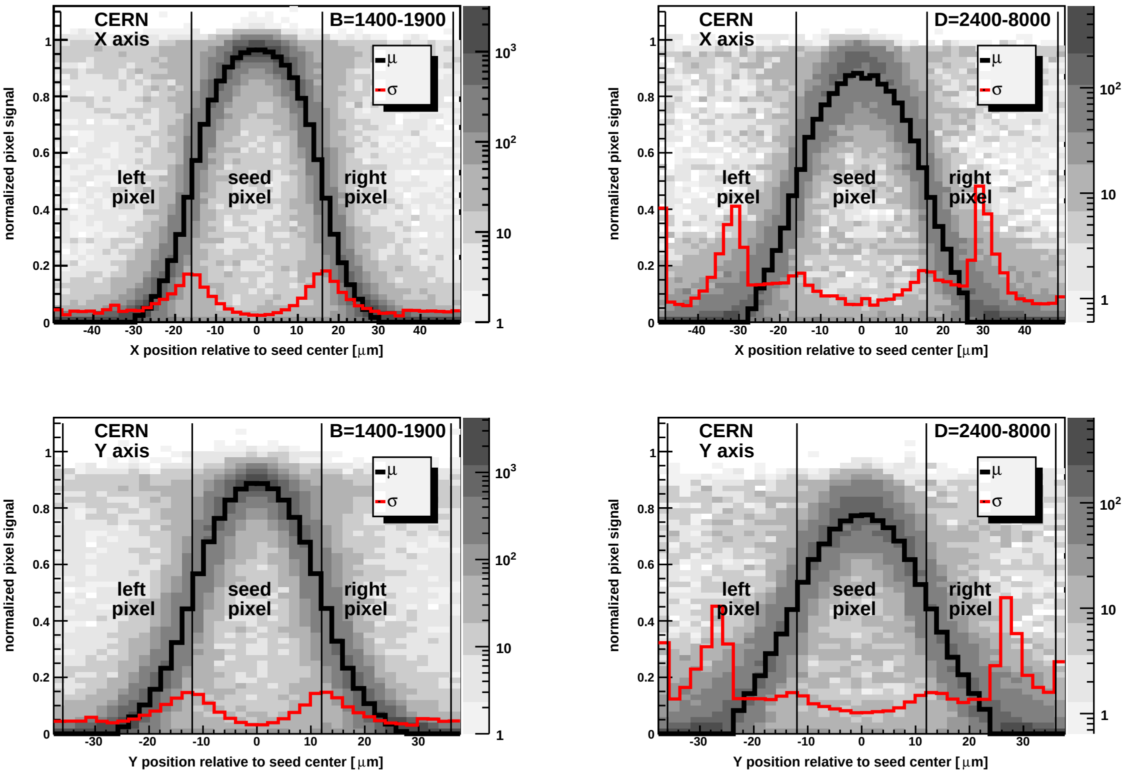


Figure 4.9: *The same as in fig. 4.8 but this time using the test beam data of module 2. The left plots are acquired using the energy cut* $B = 1400 - 1900\,ADU$ *(i.e. around the most probable value peak), the right plots with the energy cut* $D = 2400 - 8000\,ADU$ *(the high energetic Landau tail).*

with two fit ranges denoted by G_{seed} and G_{all}. For G_{seed} only the seed pixel is used for the fit, i.e. the fit range goes from $-pitch$ to $+pitch$. Although this fit is expected to perform badly it is included for completeness. For G_{all} the normal distribution is fitted from $-30\,\mu$m to $+30\,\mu$m. This fit range was limited to the extent of the charge cloud and also yields the best results. However, the approach with a single Gaussian ignores binning effects due to the discrete pixel structure.

2. **"F2":** This function stems from an approximation of the spatial straw tube efficiency distribution[1] and is also fitted from $-30\,\mu$m to $+30\,\mu$m. Besides the charge cloud width (σ) this function has also the pixel size L as a parameter ($W = L/2$):

$$F2(x)|_{-30}^{+30} = \frac{\exp\left(-\frac{|x|-W}{\sigma}\right)}{\left(\exp(-\frac{|x|-W}{\sigma}\right) - \left(\exp(-\frac{|x|-W}{\sigma}\right)^{-1}} \tag{4.6}$$

3. **"F3":** The error function fitted to the negative part and its inverse function, the

---

[1] Private conversation with S. Furletov

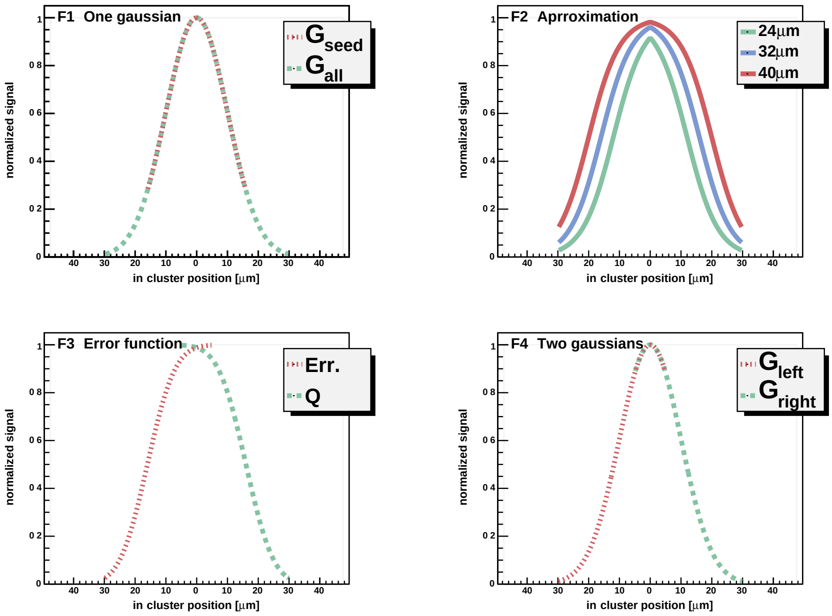


Figure 4.10: *The four charge cloud approximations $F1$ to $F4$ plotted for a charge cloud width parameter of $10\,\mu m$ and a pixel width of $32\,\mu m$, except for $F2$, where the pixel width is varied according to the plots legend $(24, 32, 40\,\mu m)$. When different fit ranges or functions are used this is indicated by two different colors and line styles.*

Q function, fitted to the positive part of the charge cloud shape:

$$F3_A = \mathrm{errf}(t)|_{-30}^{+5} \quad = \quad \frac{2}{\pi}\int_{-\infty}^{t} \exp(-(t^2)\mathrm{d}t \tag{4.7}$$

$$F3_B = Q(t)|_{-5}^{+30} \quad = \quad \frac{1}{2}\left(1 - \mathrm{errf}\left(\frac{t}{\sqrt{2}}\right)\right) \tag{4.8}$$

with $t = \frac{x-W}{\sigma}$ and σ and W as above. The fit limit at the center of the charge cloud is, however, not at 0 but at $+5\,\mu$m and $-5\,\mu$m for the error function and for the Q function respectively. The actually fit ranges are also indicated in equations 4.7 and 4.8. These fit ranges were chosen because they provided the best fit results.

4. **"F4:"** Instead of using one single Gaussian function to approximate the charge cloud like in **F1** this approach uses to Gaussians (G_{left} and G_{right}) fitted to each half separately. If **F1** suffers from binning effects due to the large pixel size compared to the cloud width ($32\,\mu$m and $24\,\mu$m to $\sigma = 10\,\mu$m respectively) this method would ideally perform better, especially for the large pixel dimension. The fit ranges are from $-30\,\mu$m to $+5\,\mu$m (G_{left}) and from $-5\,\mu$m to $+30\,\mu$m (G_{right}).

The approximations of the charge cloud with analytical functions allows also for an estimation of the charge cloud and pixel width. Table 4.3 shows pixel dimensions as a result from those fits which include the pixel dimension as a fit parameter. The difference between fitted and real pixel sizes allows estimating the precision of this method and ranges

		long axis (32 μm)				short axis (24 μm)			
function		$F2$	errf	$Q(X)$	G_{seed}	$F2$	errf	$Q(X)$	G_{seed}
parameter		$W = L/2$			σ	$W = L/2$			σ
GEANT	All (0)	16.3	16.6	16.6	16.4	12.5	13.4	13.4	12.2
	δe^- (2)	16.3	16.6	16.6	15.6	13.0	13.6	13.6	12.8
CERN	All (0)	16.1	16.3	16.4	16.2	12.6	13.3	13.3	11.8
	tail (D)	16.3	16.6	16.6	16.6	12.3	13.4	13.5	11.8

Table 4.3: *The pixel pitch parameter $W = L/2$ derived from fits of equations 4.6, 4.7, and 4.8 to the charge cloud shape. Each number is the average of all modules. The standard deviation for each value is $\sim 0.1\,\mu m$ and much smaller than the deviation from the real pixel dimensions of $\sim 0.2 - 1.6\,\mu m$. GEANT and CERN designate results from simulated and test beam data respectively, (0),(2), and (D) designate the event selection criteria according to figure 4.2. The width σ of G_{seed} is included for completeness.*

from $\sim 0.2 - 1.6\,\mu$m. This table also contains the width σ of the Gaussian restricted to the seed pixel (G_{seed}). Unlike the other functions G_{seed} has two distinctively different widths σ for the X and Y axis which reflects the pixel pitch. Therefore this function is expected to yield very poor results when used for position reconstruction.

		G_{all}	$F2$	errf	$Q(X)$	G_{left}	G_{right}
GEANT	All (0)	11.5(2.0)	7.8(0.1)	8.7(0.3)	8.7(0.3)	10.6(1.1)	10.6(1.1)
	δe^- (2)	10.9(1.0)	7.6(0.2)	8.9(0.4)	8.9(0.2)	10.6(1.1)	10.6(1.1)
CERN	All (0)	11.6(1.4)	7.7(0.5)	8.5(0.7)	8.4(0.6)	10.9(1.1)	10.8(1.1)
	tail (D)	11.0(1.2)	7.8(0.1)	8.6(0.4)	8.8(0.2)	10.5(1.1)	10.5(1.1)

Table 4.4: *The width of the charge cloud resulting from a fit of different functions to the one dimensional charge cloud distribution. Each number is the average of both axes and all modules with the standard deviation in brackets. All numbers are in μm. GEANT and CERN designate results from simulated and test beam data respectively, (0),(2), and (D) designate the event selection criteria according to figure 4.2. No significant differences can be found between these data sets.*

Table 4.4 lists the charge cloud width with a given function. The values disperse considerably and have a tendency to be greater than the expected size of $\sim 6 - 8\,\mu$m (eqn. 1.56). However, the errors on the pixels sizes shown in table 4.3 already indicate that this is not a precision measurement. Furthermore the functions used for the fits do not necessarily model the charge cloud shape in a realistic way. For example, the assumption of a Gaussian shaped cloud generally assumes a point like charge creation whereas the charge is actually continuously created along the particles track. It also assumes an initial Dirac like charge distribution for the diffusion process. In addition the pixel granularity might have an effect on the shape of the charge cloud shaped that is not entirely reflected by the approximations **F1** to **F4**. The error due to such a granularity effect should be greater for the bigger pixel size since the pixel size to cloud width ratio is bigger. Moreover the approach with single Gaussian (**F1**) should show the poorest performance in the presence of such a granularity effect.

4.2.3 Position Reconstruction

The results of the above described fits are used to reconstruct the particle position by means of method of least squares. The in-pixel position of the particle x is varied until the error ΔS is minimal:

$$\Delta S(x)|_{min} = \sum_{-1}^{+1} \left(f_{fit}^{i}(x) - S^{i} \right)^{2} \tag{4.9}$$

where S^i is the normalized signal of one of the three pixels and f_{fit} the corresponding part of the charge cloud function. In the following a short discussion about the general performance of the four charge cloud approximations F1 to F4 is given. In the next section (4.2.4) the performance of the charge cloud method is compare to the η methods presented in the previous sections particularly with regard to the energy dependence shown in section 4.

In-pixel residuals with the charge cloud based algorithm

Figure 4.11 shows the spatial residuals along the large pixel axis (32 μm) as a function of the in-pixel position for charge cloud approximation **F1** to **F4**. To avoid effects from multiple scattering or the limited telescope resolution only the simulated data set is used. The picture shows clearly that the approximations **F1** and **F4** have severe deficits. As mentioned before for the single Gaussian **F1** this is expected, but the two-Gaussian approach **F4** does apparently not predict the position correctly either. The two other methods, **F2** and **F3**, on the other hand have well behaving spatial residual distributions. Although the residual width broadens at the middle of the pixel they are still within $\sigma \sim 2\,\mu$m. More importantly the mid-pixel residuals have no offset, i.e. $\mu \sim -0\,\mu$m, unlike the corresponding residuals for **F1** and **F4**.

Figure 4.12 shows the spatial residuals versus the in-pixel position for the short pixel dimension (24 μm). As expected the smaller granularity of the pixels lead to a smaller effect and the single Gaussian approach **F1** works much better. Notwithstanding **F2** and **F3** give somewhat better results. In contrast, the approach with two Gaussians **F4** has still a very broad pyramid-like base in the total pixel residual distribution which is shown in the insert. This is due to a systematic residual offset seen in the residual vs. in-pixel histogram.

These effects are also observed in the test beam data, however to a much lesser extent as the spatial residuals are dominated by the limited telescope resolution and multiple scattering effects. To sum up, two of the four charge cloud approximations, **F2** and **F3**, show good results, the single Gaussian approach **F1** suffers from pixel granularity effects but might become more usable with smaller pixel size or larger cloud width, and the approach with two Gaussians **F4** is the least suited candidate.

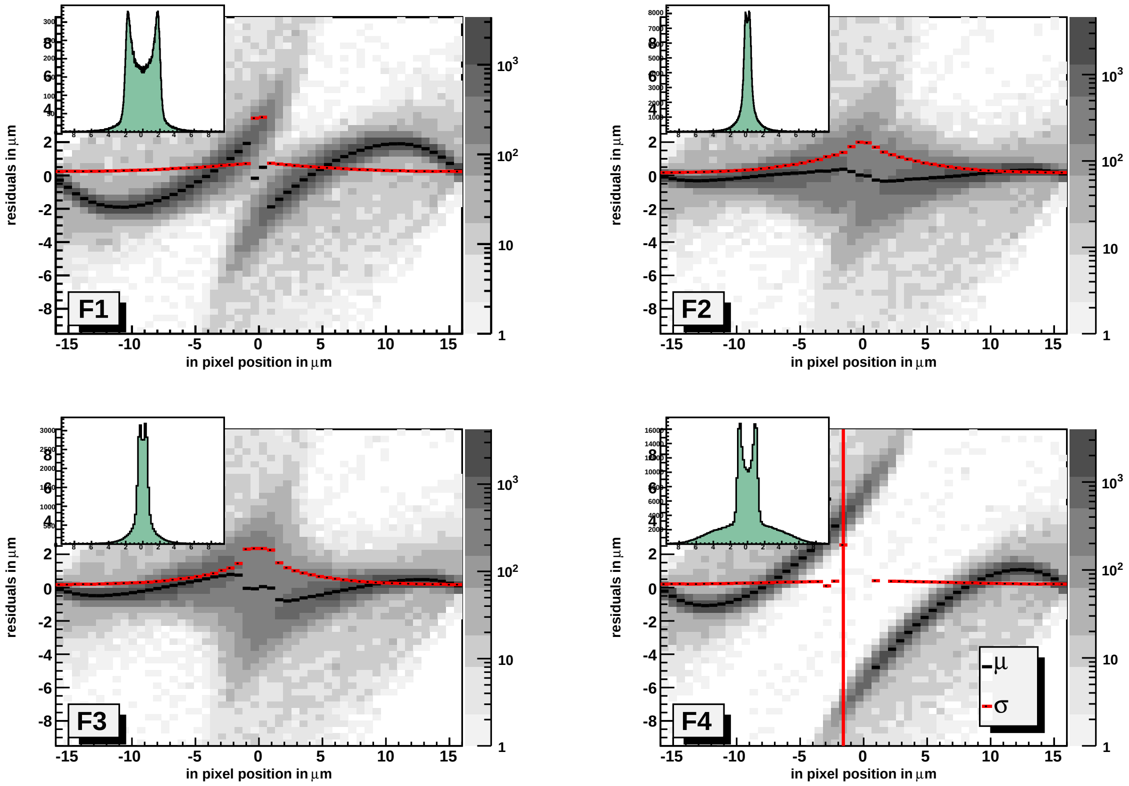


__Figure 4.11:__ The spatial residuals for the pixel X axis (32 μm pixel pitch) as a function of the in-pixel X position using the charge cloud approximation __F1__ to __F4__: For each in-pixel position the corresponding one-dimensional residual distribution is shown in a logarithmic grey scale. The black line shows the mean value μ and the red line shows the width σ of a Gaussian fit to the residual distribution for each in-pixel position bin. The insert in each plot shows the residual distribution for the entire pixel. The data set consists of simulated events without any selection criteria, i.e. all events.

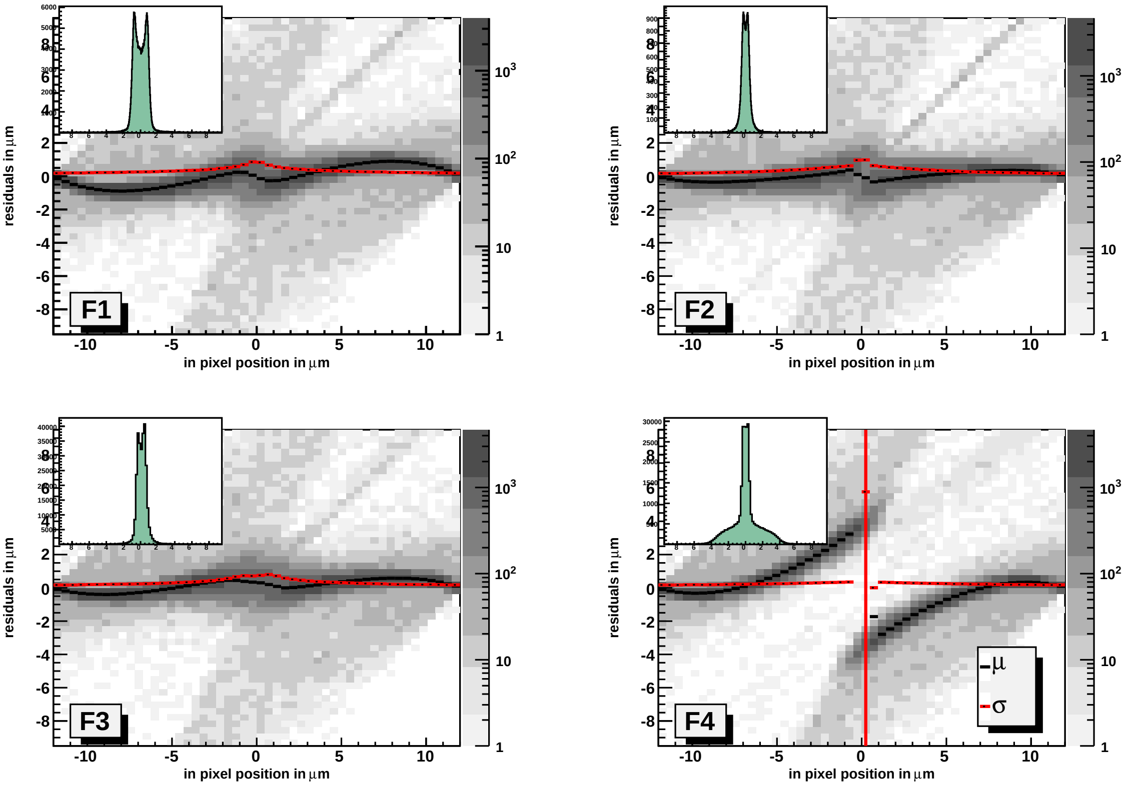


Figure 4.12: *The spatial residuals for the pixel Y axis (24 μm pixel pitch) as a function of the in-pixel Y position using the charge cloud approximation* ***F1*** *to* ***F4****: For each in-pixel position the corresponding one-dimensional residual distribution is shown in a logarithmic grey scale. The black line shows the mean value μ and the red line shows the width σ of a Gaussian fit to the residual distribution for each in-pixel position bin. The insert in each plot shows the residual distribution for the entire pixel. The data set consists of simulated events without any selection criteria, i.e. all events*

4.2.4 Comparison of charge cloud based methods and η methods

This comparison is divided in two parts. First the results for the GEANT simulation and second the results for the test beam data will be presented. In both cases the energy dependence of the residuals will be discussed.

GEANT simulation

Table 4.5 lists the residuals for the position reconstruction methods described so far and for the energy regimes specified in table 4.1. Figure 4.13 visualizes theses residuals. A general result is that the residuals for the energy cuts **A** to **C** are - within the fit error of $\sim 1\,\mu$m - virtually the same as for the entire energy range **0**$= 0 - 8000\,$ADU. The Landau like energy distribution plotted in the background of figure 4.13 makes the statistical dominance of the energy ranges **A** to **C** amongst all events quite evident and therefore this result has to be expected.

The residuals for the four charge cloud shape approximations **F1** to **F4** reflect the findings in the last section: a two-Gaussian approach give the poorest results independent of the pixel axis, whereas a single Gaussian approach gives better results for the smaller Y axis. However, both methods fall behind the approaches **F2** and **F3**. The latter two give similar, yet somewhat poorer results compared to the η methods. The two η approaches, single η_{one} and multiple η_{multi}, perform equally well within the fit uncertainties.

However, the picture changes when looking at the residuals for the high energy tail **D** of the cluster signal distribution. Here the standard η_{one} method clearly fails when compared to alternative methods. The best performance is given by applying a separate η_D distribution for this energy regime, i.e. a tailored multiple η_{multi} approach. Among the charge cloud based methods **F2** gives comparable, yet slightly worse results. The other methods possess clearly poorer position prediction capabilities.

GEANT	X (32 μm)					Y (32 μm)				
CUT	0	A	B	C	D	0	A	B	C	D
η_{one} [μm]	0.5	0.5	0.4	0.4	1.8	0.3	0.3	0.3	0.2	1.3
η_{multi} [μm]	0.5	0.6	0.5	0.4	0.8	0.3	0.4	0.3	0.3	0.5
F1 [μm]	1.6	1.5	1.5	1.5	2.1	0.7	0.7	0.7	0.6	1.2
F2 [μm]	0.5	0.5	0.5	0.5	0.9	0.4	0.4	0.4	0.4	0.6
F3 [μm]	0.6	0.7	0.6	0.6	1.0	0.5	0.5	0.5	0.5	1.1
F4 [μm]	1.2	1.2	1.2	1.2	1.6	1.4	1.3	1.4	1.4	1.5

Table 4.5: *A comparison of residual width in μm for various position reconstruction method and energy cuts using simulated data. See figure 4.13 for a visualization.*

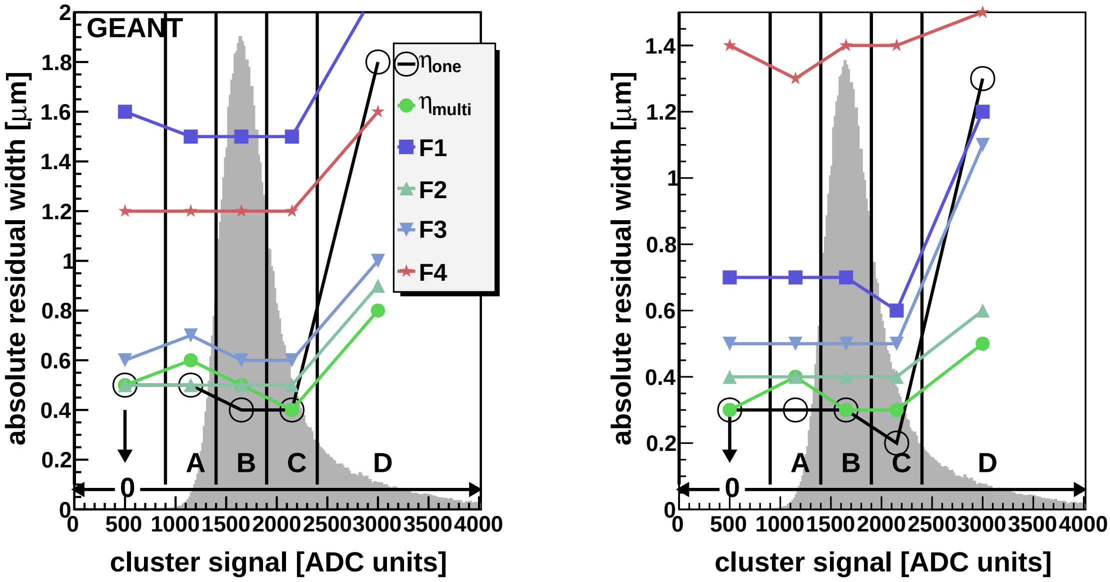
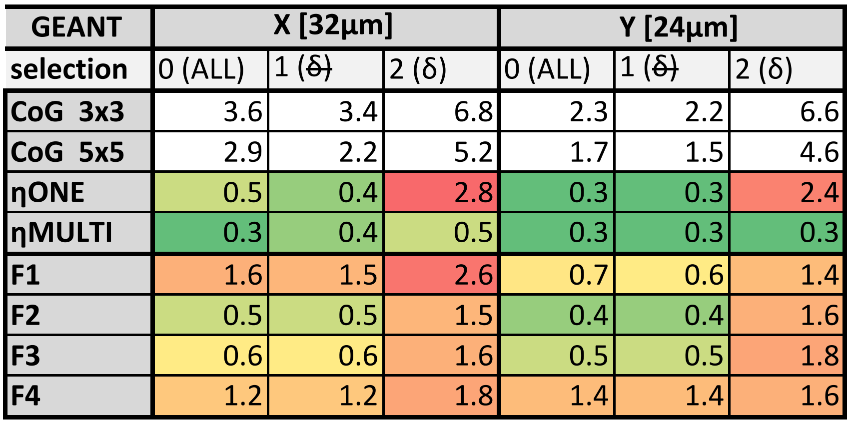


Figure 4.13: *The residual width in μm in X (left) and Y (right) for simulated events of $32 \times 24\,\mu m^2$ sized pixel obtained with a single η_{one} and multiple η_{multi} distributions as well as four charge cloud models **F1** to **F4**. The residuals are plotted as a function of the cluster signal with the cluster signal divided into ranges. The residual values for all events ($0 - 8000\,ADC$ units) are put on the left (at an x-value of 500) and are indicated by the **0**.*



GEANT	X [32μm]			Y [24μm]		
selection	0 (ALL)	1 (~~δ~~)	2 (δ)	0 (ALL)	1 (~~δ~~)	2 (δ)
CoG 3x3	3.6	3.4	6.8	2.3	2.2	6.6
CoG 5x5	2.9	2.2	5.2	1.7	1.5	4.6
ηONE	0.5	0.4	2.8	0.3	0.3	2.4
ηMULTI	0.3	0.4	0.5	0.3	0.3	0.3
F1	1.6	1.5	2.6	0.7	0.6	1.4
F2	0.5	0.5	1.5	0.4	0.4	1.6
F3	0.6	0.6	1.6	0.5	0.5	1.8
F4	1.2	1.2	1.8	1.4	1.4	1.6

Table 4.6: *Residuals in μm for the $32 \times 24\,\mu m^2$ pixel simulation using the δ-electron flag for event classification. The event selection criteria are: All - all events are taken, $\not{\delta}$ - only events without secondary particle are taken, and δ only events with a secondary particle are taken. CoG refers to the Center-of-Gravity method for a 3×3 and a 5×5 pixel cluster resp., the other methods are explained in the text. The color scale indicates the quality of method. Red marks a bad and green a good result.*

δ-electron flag based event selection

The GEANT simulation also allows differentiating events by a δ-electron flag. In this case the multiple η_{multi} method has two separate η distributions, $\eta_{\bar{\delta}}$ for events without

δ-electrons and η_δ for events with δ-electrons. Analogue, the charge cloud shapes are sampled separately for the two events classes. The residuals for this study can be found in table 4.6. For events without δ-electrons the results reflect the findings above: Both η methods work comparably well, closely followed by charge cloud approximations F2 and F3, whereas F1 and F4 show poorer results. For δ-electron only events the charge cloud based methods give better results than the standard single η_{one} method, but the η_{multi} method clearly outperforms any other method. For completeness table 4.6 also shows the Center-of-Gravity residuals for a 3×3 and a 5×5 pixel cluster: In any case the Center-of-Gravity method performs vastly worse than any other method.

Test Beam Data

Akin to the GEANT simulation the multiple η_{multi} and the charge cloud based method have been applied to test beam data.Since some modules showed a suboptimal performance in the test beam this discussion will be limited to module 2 and 14. These modules showed the best behavior on sensor and on in-pixel scale as explained in chapter 3. Also both modules are of the CoCG-Large pixel type with two different axis dimensions, a long X axis (32 μm) and a short Y axis (24 μm). The residuals of module 2 and module 14 are listed in table 4.7 and table 4.8, respectively. These results are visualized in figures 4.14 and 4.15.

The results are similar to those of the GEANT simulation. However, differences are less prominent due to the additional spatial error from the limited telescope resolution and from multiple scattering effects. For events with energy around the most probable value both η methods and F2 and F3 give comparable results. In contrast to the simulation these charge cloud methods give even the best results while the multiple η_{multi} yields slightly worse results. However, these differences are within or close to the residual width fit errors of $\sim 0.1\,\mu$m.

Unlike for the GEANT simulation the spatial resolution is already deteriorating at the energy range **C** which is within twice the most probable value. This can be seen in the plot for the X axis in figure 4.15. Using alternative reconstruction methods like η_{multi} or F2 and F3 can already improve the resolution here. Concordant to the simulation this becomes drastically more important for the high energy tail **D**.

Mod2	X ($32\,\mu$m)					Y ($24\,\mu$m)				
CUT	0	A	B	C	D	0	A	B	C	D
η_{one}	2.4	2.3	2.3	2.4	3.4	2.3	2.3	2.1	2.3	2.9
η_{multi}	2.5	2.6	2.4	2.4	2.9	2.3	2.3	2.2	2.3	2.8
F1	3.2	3.6	3.3	3.0	3.5	2.2	2.6	2.2	2.3	2.9
F2	2.3	2.2	2.2	2.4	3.0	2.2	2.0	2.1	2.3	2.8
F3	2.4	2.3	2.3	2.4	3.0	2.2	2.1	2.1	2.3	2.8
F4	2.6	2.6	2.6	2.6	3.2	2.3	2.7	2.4	2.2	2.7

Table 4.7: *A comparison of residual widths in μm for various position reconstruction method and energy cuts using the test beam data of module 2.*

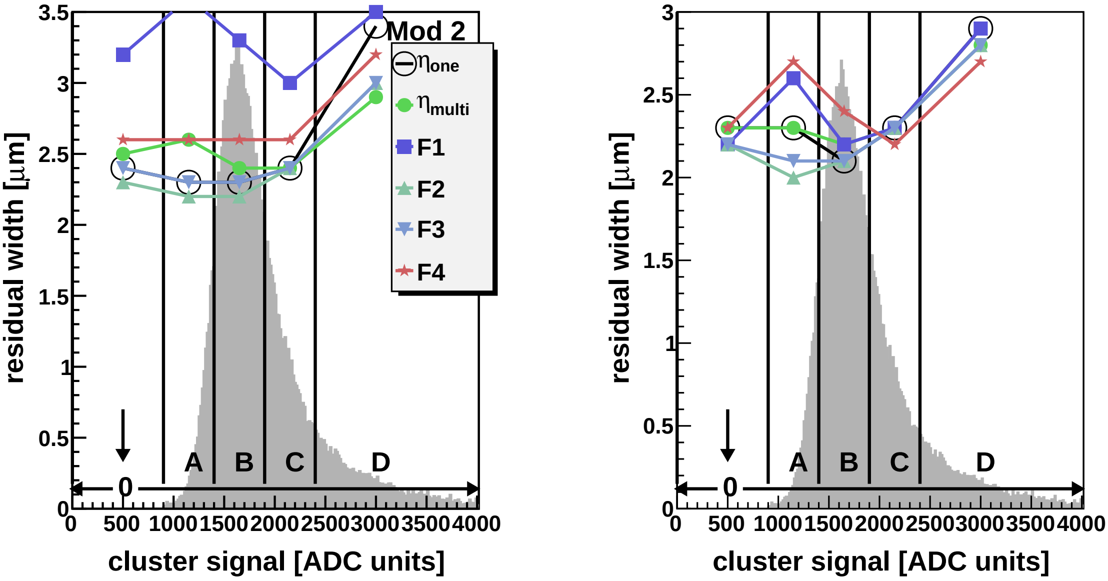


Figure 4.14: *Residual widths in μm in X (left) and Y (right) gathered from test beam data of module 2 and using various position reconstruction methods. The residuals are plotted as a function of the cluster signal with the cluster signal divided into ranges according to table 4.1. The residual values for all events ($0-8000\,ADC$ units) are put on the left (at an x-value of 500) and are indicated by the* ***0****. The pixel dimensions are $32 \times 24\,\mu m^2$.*

Mod14	X (32 μm)					Y (24 μm)				
CUT	0	A	B	C	D	0	A	B	C	D
η_{one}	1.9	1.7	1.7	2.2	3.2	1.8	1.7	1.7	2.0	2.6
η_{multi}	1.9	1.9	1.8	2.0	2.4	1.8	1.8	1.7	2.0	2.4
F1	3.4	3.1	2.4	2.6	3.8	1.7	2.1	1.7	2.0	2.5
F2	1.8	1.6	1.6	1.9	2.5	1.7	1.6	1.6	1.9	2.5
F3	1.7	1.6	1.7	1.9	2.5	1.7	1.6	1.6	1.9	2.5
F4	2.0	2.0	2.0	2.1	2.6	1.7	1.9	1.7	1.9	2.4

Table 4.8: *A comparison of residual width in μm for various position reconstruction method and energy cuts using the test beam data of module 14.*

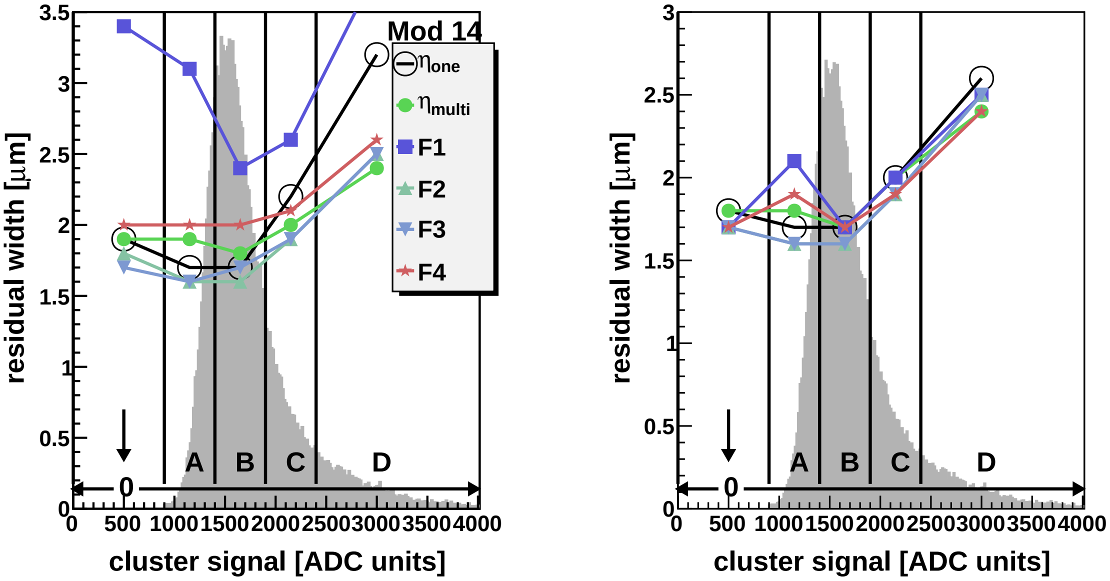


Figure 4.15: *Residual widths in μm in X (left) and Y (right) gathered from test beam data of module 14 and using various position reconstruction methods. The residuals are plotted as a function of the cluster signal with the cluster signal divided into ranges according to table 4.1. The residual values for all events ($0-8000$ ADC units) are put on the left (at an x-value of 500) and are indicated by the* ***0****. The pixel dimensions are $32 \times 24\,\mu m^2$.*

4.3 Multivariate analysis

Multivariate analysis, (**MVA**) or multivariate statistics, studies the dependence of one or several target variables on more than one input variable. MVA encompasses a variety of methods like principal component analysis (PCA), artificial neural networks (ANN), or support vector methods (SVM), to name a few. To employ multivariate analysis for position reconstruction the *Toolkit for Multivariate Analysis with ROOT* (**TMVA**) was chosen, since it is already incorporated into the ROOT analysis frame work. The following four methods were used:

- (multidimensional) linear regression (LD[4], sec. 4.3),
- artificial neural network (MLP, sec. 4.3),
- boosted decision trees (BDT, sec. 4.3),
- and probability density estimators (PDE, 4.3).

Linear regression model (LD)

In the linear regression model the functional dependents of a target variable $y(\mathbf{x})$ on m input variables x_i is linear in its parameters β with

$$y(\mathbf{x}) = \sum_i^m \beta_i x_i = \mathbf{x}^T \beta \tag{4.10}$$

However, it is not required that the input variables x_i are linear. For example one possible input variable would be the square of a pixel signal ($x_1 = S_{seed}^2$). However, for simplicity reasons all input variables (pixel signals and cluster moments as explained later) have a simple linear one to one assignment to the vector x_i. That means for example that a three by three pixel cluster ($x = [-1, 0, +1]$ and $y = [-1, 0, +1]$) with signals $S_{x,y}$ would look like

$$\mathbf{x} = \begin{pmatrix} S_{-1,-1} \\ S_{0,0} \\ S_{+1,+1} \end{pmatrix} \tag{4.11}$$

Assuming there are n training events available the parameter vector β is defined by

$$Y = X\beta = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1m} \\ 2 & x_{21} & \cdots & x_{2m} \\ 3 & x_{31} & \cdots & x_{3m} \\ \vdots & \vdots & \ddots & \vdots \\ n & x_{n1} & \cdots & x_{nm} \end{pmatrix} (\beta_1 \beta_2 \; \cdots \beta_m) \tag{4.12}$$

The application of the least square method with

$$\frac{\partial}{\partial X} \sum_k^n (y_k - \mathbf{x}_k \beta)^2 = 0 \tag{4.13}$$

---

[4] The TMVA notation for linear discriminant and linear regression is in both cases LD. This notation is kept here.

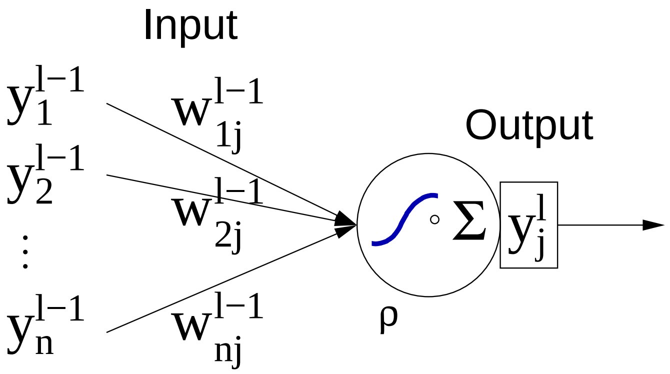


Figure 4.16: *A single neuron j in a layer l with n input connections. The incoming connections carry a weight of $w_{ij}^{(l-1)}$ [53].*

An artificial neural network (**ANN**) is an attempt to mimic a biological network of neurons as found in e.g. a human brain. However, modern ANNs are much more based on statistics and signal processing. Furthermore ANN is actually a class of different network types with different types of interconnectivity and different types of neurons. The ANN as implemented in the TMVA package[5] and used for this study is a multilayer percepton (**MLP**) ANN. Multilayer refers to the structure of the node network and percepton to the node type. Percepton, node and neuron are used interchangeably here. With the exception of the first layer the input of each percepton is the weighted sum of the output of all neurons in the previous layer. The input of neuron j in layer l with n neurons in layer $l-1$ would then be

$$I(\mathbf{y})_j^l = \sum_i^N w_{ij}^{l-1} O_i^{l-1} \tag{4.15}$$

where w_{ij}^{l-1} is the weight for each input O_i^{l-1} (fig. 4.16). These weights contain the knowledge of the neural network and they are the quantities that are trained by a method called *back propagation* (more below). Furthermore every percepton has an *activation function*. Except for the first layer the activation function is non-linear and usually sigmoid

---

[5] The TMVA package offers three different network types. However, it is recommended to use the MLP type.

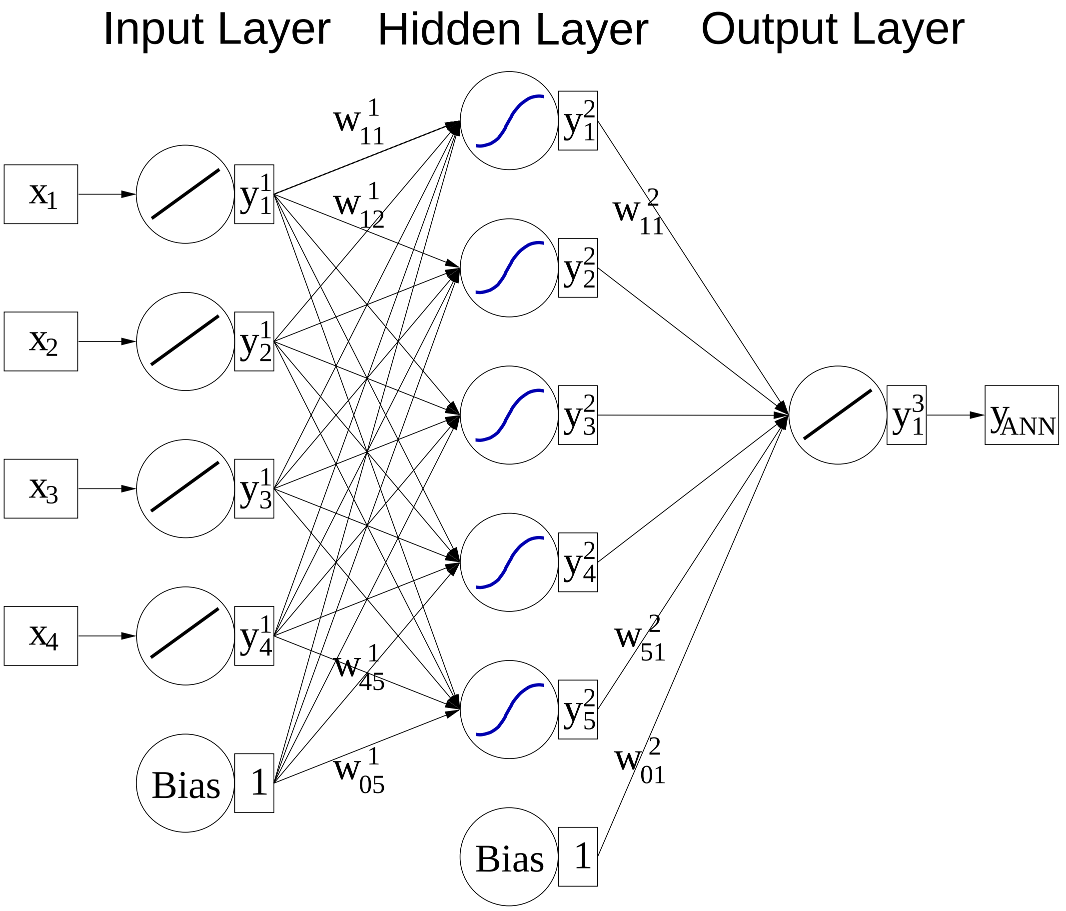


Figure 4.17: *Example of a multilayer percepton artificial neural network with one hidden layer [53].*

like (e.g. tanh, logistic sigmoid or radial). In the scope of this thesis the activation function is always the *hyperbolic tangent* function. The topography of a multilayer percepton network is divided into an input layer, an output layer, and one or several hidden layers between the input and output layer (fig. 4.17). In the input layer one input percepton with a linear activation function is put for each input variable. The output layer has an output percepton for each target variable with a non linear activation function. The number of hidden layers is not fixed as is the number of perceptons per hidden layer. For this study the default values of the TMVA MLP regression neural network were used: Two hidden layers with $N+20$ and $N+15$ neurons, respectively, where N is the number of input variables. The MLP is a feed forward network meaning that there are no feedbacks and no loop connections between neurons.

Back-propagation training. The method used to train the neural network is called back propagation. The idea is to find a set of weights $\mathbf{w}$ that minimizes the error $E(\mathbf{x}|\mathbf{w}) = \frac{1}{2}(Y_{MLP}(\mathbf{x}|\mathbf{w}) - Y_T)^2$. $Y_{MLP}(\mathbf{x}|\mathbf{w})$ is the network output value, Y_T the target value from the training sample, and $\mathbf{x}$ is the set of input variables for a given event. Using the method of the *steepest descent* or *gradient descent* the weights w_{ij}^l connected to layer l are changed according to

$$\Delta w_{ij}^l = -\eta \frac{\partial E(\mathbf{x}|\mathbf{w})}{\partial w_{ij}^l} \tag{4.16}$$

where η is the *learning rate.* The information coming from the input variables is propagated forward through the network. However, the weight changes or *errors* propagates backwards through the network, i.e. first the output layer, then the second, and then the first hidden layers are processed for the error (back) propagation. These steps can be done for all events at once, a technique called *bulk learning* and that requires large amounts of memory. The opposite philosophy, training based on an event by event basis, is called *online learning* and uses less memory but needs more computing power and is the method implemented in TMVA.

Boosted decision trees (BDT)

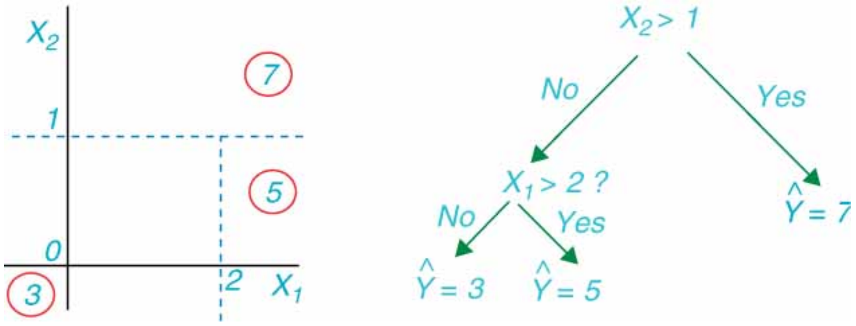


Figure 4.18: *A regression tree is a series of cuts that splits the variable space into several subspaces. Inside each subspace the output is a constant.*

A *decision tree* is a graphical representation of splitting up the variable input space into several subspaces. It consist of a tree like decision node structure as depicted in a simple fashion in figure 4.18. The Boosted Decision Tree (BDT) starts with a root node where the training data is split in two. Then each part of the data goes through the same decision making algorithm resulting in further nodes and smaller subsets of the data. This procedure is repeated until the whole tree is build. Each node has one cut on one variable and the cut is optimized to minimize the *average squared error*:

$$\hat{o} = \frac{\sum^N (y - \hat{y})^2}{N} \tag{4.17}$$

where N is the number of events in the sub sample, y the target variable for each event, and $\hat{y}$ the mean value over all events in that node. The tree is grown until the maximum number of nodes is reached. A tree sufficiently large will give a perfect result for the training data. It is clearly overtrained and will perform suboptimal for test data. This can be avoided by keeping the maximum number of nodes sufficiently low and grow a *forest* of regression trees or by cutting the tree back, a method called *pruning.* The *pruning* method was waived, the maximum number of nodes was set to the default value of 5, and the number of trees was set to 200. Each tree of the forest $f_m(x)$ contributes to the output variable with a weight β_m

$$F(x) = \sum_{m=0}^{M} \beta_m f_m(x). \tag{4.18}$$

The weights are determined by a technique called **Gradient Boost**[54]. It minimizes the error on the prediction which is in case of the TMVA regression BDT defined by the Huber loss function [53]:

$$L(F, y) = \begin{cases} \frac{1}{2}(y - F(\mathbf{x}))^2 & |y - F| \leq \delta, \\ \delta(|y - F| - \delta/2) & |y - F| > \delta. \end{cases} \tag{4.19}$$

This is done applying the *steepest gradient* method to the problem [54].

Probability density estimator (PDE)

This method estimates the most likely output y given a set of n_{var} input parameters $\mathbf{x}$ (*maximum likelihood* method). In a simple approach ("naive Bayes estimator") all input variables are treated as being independent and the likelihood of an output y to occur is

$$\mathfrak{L}(y) = \prod_{k=1}^{n_{var}} P_k(x_k|y) \tag{4.20}$$

where $P_k(x_k|y)$ is the probability distribution function (PDF) for variable x_k to give an outcome of y. However, since the actual analytical behavior of $P_k(x_k|y)$ is unknown, the function needs to be sampled and is therefore an estimation (**PDE**). It is often convenient to work with the logarithm of $\mathfrak{L}$ (log-likelihood):

$$\ln \mathfrak{L}(y) = \sum_{k=1}^{n_{var}} \ln P_k(x_k|y) \tag{4.21}$$

The output y is estimated by maximizing equation 4.21. So far all variables have been treated independently with projected density distributions, but any correlation between the input variables will impair the predictive performance of this method. A true multidimensional approach will not take n_{var} separated histograms but one histogram with $n_{var}+1$ dimensions. However, this multidimensional phase space needs huge training samples to be sufficiently populated. To overcome this problem the target value for a given *test* event i is estimated from an appropriate sized volume inside the multidimensional probability density estimation. The idea behind this is that the probability distribution will not change abruptly. There are different methods to handle this approach, e.g. PDE-Foam, K-nearest neighbors, etc. The method employed here is the range search (**PDERS**) as proposed in [55]. The target value is the weighted average of the target value of all events that are stored within a volume V centered at event i

$$y_{PDE-RS}(i, V) = \frac{\sum_{j \in V} w_j t_j f(dis(i, j))}{\sum_{j \in V} w_j f(dis(i, j))} \tag{4.22}$$

$f(dis(i,j))$ is a gaussian kernel function that weights events according to their distance $dis(i,j)$ from the target event. The volume V itself is selected by an *Adaptive* algorithm[6] which takes the RMS of each dimension as well as the event density in the vicinity of the target event into account.

4.3.1 Information value of input variables

To minimize computation time and have sufficient prediction performance with a given limited training data set it is important to reduce the number of input variables to a sensible volume. Adding an input variable that merely contains noise would just increase the necessary computing power and the uncertainties of the multivariate analysis. Therefore it is important to estimate the amount of information contained within one input variable with regard to the target variable. Three methods have been applied to assess any dependence of the input on the target[7]:

- Correlation factor:

  $$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{4.23}$$

  The correlation factor is symmetric in X and Y and lies within $[-1,1]$. It measures any linear relationship between variable X and Y. A value of 0 denotes that no linear relationship exists, but still a non-linear relationship (quadratic, etc.) might exist.

- Correlation ratio:

  $$\eta^2(Y|X) = \frac{\sigma_{E(Y|X)}}{\sigma_Y} \tag{4.24}$$

  where $E(X|Y)$ is the conditional expectation value of Y given X with the associated conditional probability density function $P(Y|X)$:

  $$E(Y|X) = \int yP(y|x)dy. \tag{4.25}$$

  The correlation ration $\eta^2(Y|X)$ is in general not symmetric and lies between $[0,1]$, depending on how well the data can be fitted with a linear or non linear regression curve.

  In terms of a two dimensional histogram this translates to

  $$\begin{aligned} \sigma_{E(Y|X)} &= \sigma_{\int yP(y|x)dy} && (4.26)\\ &= \sum^{N_x}\left(\int yP(y|x)\,dy\right)(\mu_{Y|X} - \mu_{Y|axis})^2 && (4.27)\\ &= \sum_{i}^{N_x}\left(\int \mathrm{H1Y_i\,dy}\right)(\mu_{\mathrm{Y|X}} - \mu_{\mathrm{Y|axis}})^2 && (4.28)\\ & && (4.29)\end{aligned}$$

---

[6] Adaptive is the TMVA designation of this algorithm.

[7] These methods are also applied by TMVA itself. However, they were employed independently to have a better understanding of the possible predictive power of the input variables.

where $\mathrm{H1Y_i}$ is the one dimensional slice in Y through the 2D histogram at the x bin number i, $\mu_{Y|X}$ is the mean value of this slice, and $\mu_{Y|axis}$ is the mean value of the distribution projected onto the y axis.

- Mutual information:

$$I(X,Y) = \sum_{y \in Y} \sum_{x \in X} P(x,y) \log \left(\frac{P(x,y)}{P(x)P(y)} \right) \tag{4.30}$$

The mutual information measures any mutual dependence of two variables. It is always positive and a value of 0 means that both variables are completely independent.

Two classes of input variables were investigated with these three quantities:

- Pixel signals of the 5 × 5 pixel cluster.
- Statistical moments of the measured spatial charge distribution. This is done for a 3 × 3 and 5 × 5 cluster size.

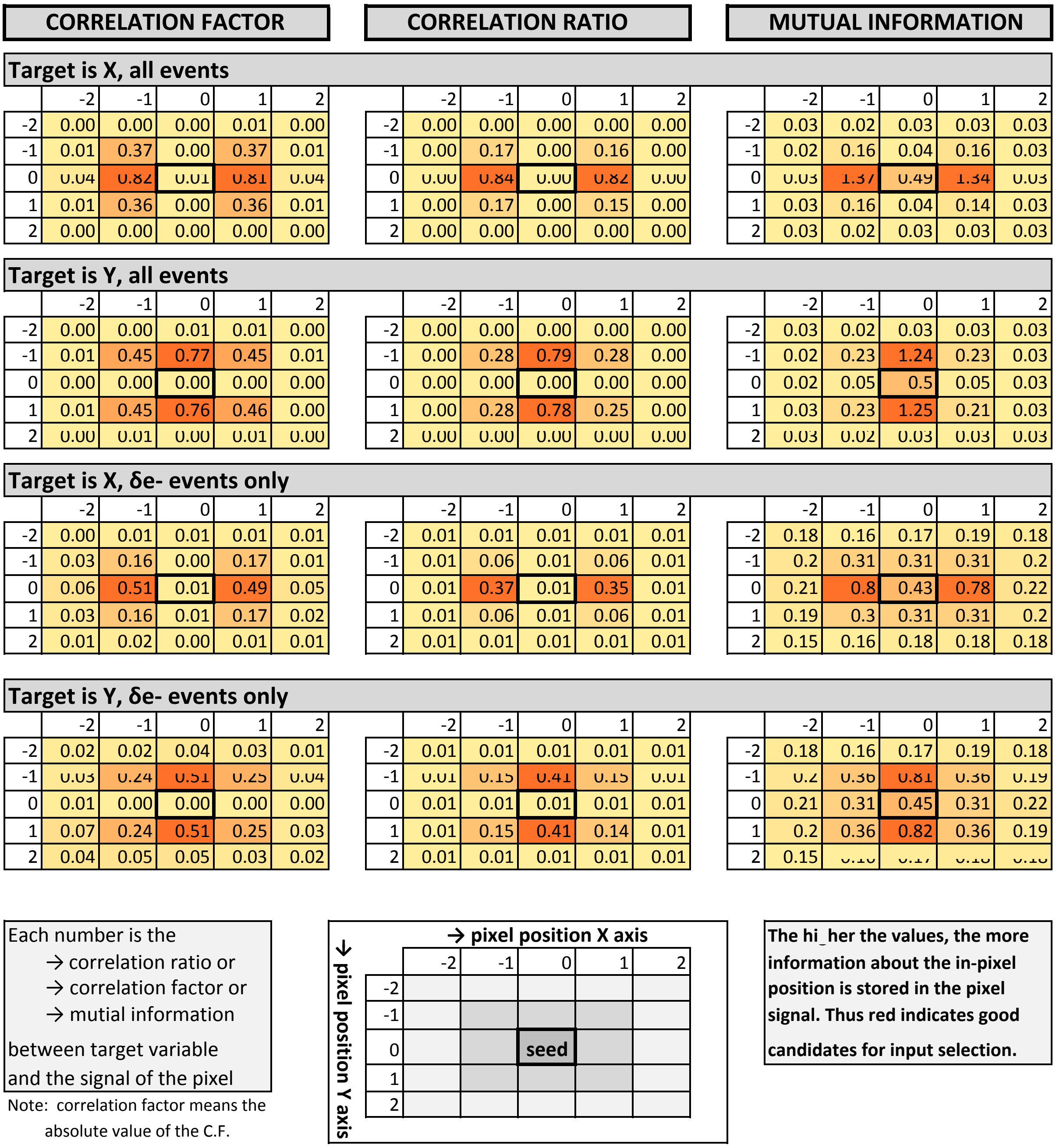
In a pre-stage of this analysis each input variable was projected onto a $[0 \cdots 1]$ interval and a cut excluding the outmost 1% of the events. This step was necessary as the real DEPFET test beam data contains events with outliers that would otherwise dampen the predictive power of a variable.

Pixel signals

All three statistical quantities described in the previous section where calculated for the 5 × 5 pixels of a cluster as input variable and with in-pixel position as target variable. Since the regression mode of TMVA works only on one target variable this was done separately for both directions. The pixel signals where normalized by the total cluster pulse height. Figure 4.19 shows a graphical representation of the pixel signal distribution as a function of the in-pixel position in Y for each of the inner 3 × 3 pixels. Table 4.9 shows the results for the GEANT simulation. These findings are basically the same for all modules in the test beam and the GEANT simulations.

All three methods reveal that outside the area of a 3 × 3 pixel cluster around the seed pixel virtually no information about the in-pixel particle position in either direction can be found. This result is expected since the charge cloud has been measured to be confined to the inner 3 × 3 pixels as shown earlier in this chapter. Furthermore, the two pixels adjacent to the seed in a given axis contain the most information on the in-pixel position along this axis. In contrast, the two seed neighbors in the other direction contain virtually no information about the in-pixel position along a given direction.

When selecting only simulated events with δ-electrons every pixel inside the 3×3 contains information, yet the two direct seed neighbors in the respective axis still contain a dominating amount of information. However, for this specific event selection some information can be found also outside a 3 × 3 pixel area.



CORRELATION FACTOR | **CORRELATION RATIO** | **MUTUAL INFORMATION**

Target is X, all events

	-2	-1	0	1	2		-2	-1	0	1	2		-2	-1	0	1	2
-2	0.00	0.00	0.00	0.01	0.00	-2	0.00	0.00	0.00	0.00	0.00	-2	0.03	0.02	0.03	0.03	0.03
-1	0.01	0.37	0.00	0.37	0.01	-1	0.00	0.17	0.00	0.16	0.00	-1	0.02	0.16	0.04	0.16	0.03
0	0.04	0.82	0.01	0.81	0.04	0	0.00	0.84	0.00	0.82	0.00	0	0.03	1.37	0.49	1.34	0.03
1	0.01	0.36	0.00	0.36	0.01	1	0.00	0.17	0.00	0.15	0.00	1	0.03	0.16	0.04	0.14	0.03
2	0.00	0.00	0.00	0.00	0.00	2	0.00	0.00	0.00	0.00	0.00	2	0.03	0.02	0.03	0.03	0.03

Target is Y, all events

	-2	-1	0	1	2		-2	-1	0	1	2		-2	-1	0	1	2
-2	0.00	0.00	0.01	0.01	0.00	-2	0.00	0.00	0.00	0.00	0.00	-2	0.03	0.02	0.03	0.03	0.03
-1	0.01	0.45	0.77	0.45	0.01	-1	0.00	0.28	0.79	0.28	0.00	-1	0.02	0.23	1.24	0.23	0.03
0	0.00	0.00	0.00	0.00	0.00	0	0.00	0.00	0.00	0.00	0.00	0	0.02	0.05	0.5	0.05	0.03
1	0.01	0.45	0.76	0.46	0.00	1	0.00	0.28	0.78	0.25	0.00	1	0.03	0.23	1.25	0.21	0.03
2	0.00	0.01	0.00	0.01	0.00	2	0.00	0.00	0.00	0.00	0.00	2	0.03	0.02	0.03	0.03	0.03

Target is X, δe- events only

	-2	-1	0	1	2		-2	-1	0	1	2		-2	-1	0	1	2
-2	0.00	0.01	0.01	0.01	0.01	-2	0.01	0.01	0.01	0.01	0.01	-2	0.18	0.16	0.17	0.19	0.18
-1	0.03	0.16	0.00	0.17	0.01	-1	0.01	0.06	0.01	0.06	0.01	-1	0.2	0.31	0.31	0.31	0.2
0	0.06	0.51	0.01	0.49	0.05	0	0.01	0.37	0.01	0.35	0.01	0	0.21	0.8	0.43	0.78	0.22
1	0.03	0.16	0.01	0.17	0.02	1	0.01	0.06	0.01	0.06	0.01	1	0.19	0.3	0.31	0.31	0.2
2	0.01	0.02	0.00	0.01	0.01	2	0.01	0.01	0.01	0.01	0.01	2	0.15	0.16	0.18	0.18	0.18

Target is Y, δe- events only

	-2	-1	0	1	2		-2	-1	0	1	2		-2	-1	0	1	2
-2	0.02	0.02	0.04	0.03	0.01	-2	0.01	0.01	0.01	0.01	0.01	-2	0.18	0.16	0.17	0.19	0.18
-1	0.03	0.24	0.51	0.25	0.04	-1	0.01	0.15	0.41	0.15	0.01	-1	0.2	0.36	0.81	0.36	0.19
0	0.01	0.00	0.00	0.00	0.00	0	0.01	0.01	0.01	0.01	0.01	0	0.21	0.31	0.45	0.31	0.22
1	0.07	0.24	0.51	0.25	0.03	1	0.01	0.15	0.41	0.14	0.01	1	0.2	0.36	0.82	0.36	0.19
2	0.04	0.05	0.05	0.03	0.02	2	0.01	0.01	0.01	0.01	0.01	2	0.15	[illegible]	[illegible]	[illegible]	[illegible]

Each number is the
→ correlation ratio or
→ correlation factor or
→ mutial information

between target variable
and the signal of the pixel

Note: correlation factor means the
absolute value of the C.F.

→ pixel position X axis

↓ pixel position Y axis

	-2	-1	0	1	2
-2					
-1					
0			seed		
1					
2					

The hi_her the values, the more information about the in-pixel position is stored in the pixel signal. Thus red indicates good candidates for input selection.

Table 4.9: Tables of mutual information, correlation ratio, and correlation factor of each ***pixel signal*** *of a* 5×5 *cluster with regard to the X and Y in-pixel position. Each number indicates the usefulness the corresponding pixel signal distribution as an input variable for the multivariate analysis. For the tables in the left, middle, and right column different statistical quantities were used to evaluate the usefulness as an input variable. This study was done for the X and Y axis as a target variable and for two different event classes: All events, and events that are required to have a least one* δ*-electron. This is indicate by the title for each of the four rows of tables. Each table represents the 25 pixels of a* 5×5 *cluster as indicated at the bottom of the table. The data is a GEANT simulation with* $32 \times 24\,\mu m$ *sized pixels. The majority of information is confined to the two direct neighbors along the targeted axis.*

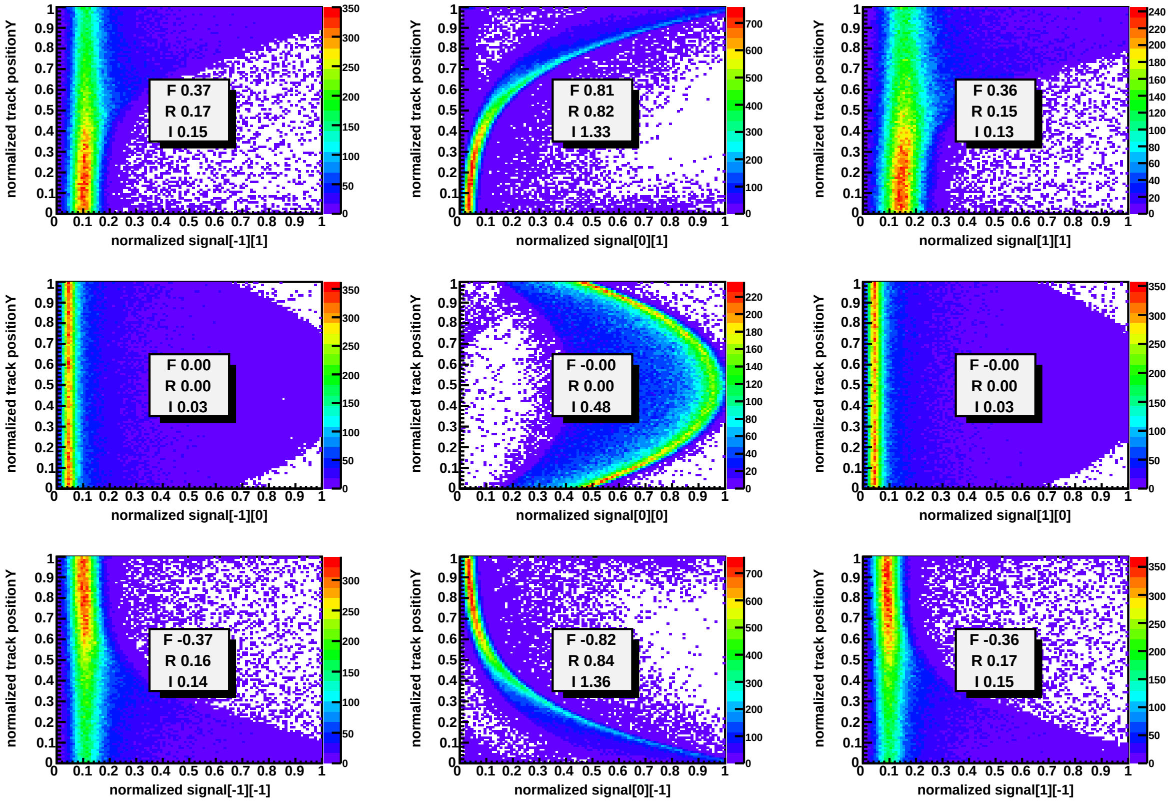


Figure 4.19: *This figures shows nine plots corresponding to the nine pixels in a* 3×3 *cluster, i.e. the central plot corresponds to the seed pixel. In each plot the normalized in-(seed)pixel Y axis position is shown as a function of the normalized pixel signal. Normalized means the variable range has been projected to the range* $[0 \cdots 1]$*. For example, the Y axis of the top, center plot shows the position of the particle inside the seed pixel. The X axis of this plot shows the normalized signal of the upper seed neighbor in the Y direction of the sensor. The colored curve in this plot indicates that this pixel contains more charge (the normalized signal goes to* 1*), if the particle position inside the seed pixel is closer to this neighbor pixel (the position goes to* 1*). These plots were produced using the GEANT simulation (pixel Y axis is* $24\,\mu m$*). In each plot the corresponding correlation factor (F), the correlation ratio (R), and the mutual information (I) calculated for this plot is printed.*

Statistical moments

A moment in mathematics is a quantitative measure of the shape of a distribution. Since the signal distribution of a cluster contains the information about the particle's in-pixel position, a set of statistical moments describing this signal distribution could be used as an input for multivariate methods.
There are two types of moments:

- the algebraic moments or moments about the origin and
- the central moments or moments about the mean.

Note that in the following the moments are normalized by the integral or sum of the distribution respectively. For a probability function this normalization term is one and can be neglected. The following gives a short review of moments (see [56] for details):

The algebraic moment of order n is defined as the expectation value x^n of a distribution, leading to

$$\mu_n^a = \frac{\sum_i x_i^n p_i}{\sum_i p_i} \tag{4.31}$$

for a discrete distribution p and to

$$\mu_n^a = \frac{\int x^n f(x) \mathrm{d}x}{\int f(x) \mathrm{d}x} \tag{4.32}$$

for a continuous distribution f(x). Furthermore the lowest order algebraic moments are equivalent to

$$\mu_0^a = 1 \tag{4.33}$$
$$\mu_1^a = \mu \tag{4.34}$$
$$\mu_2^a = \sigma^2 + \mu^2 \tag{4.35}$$

with σ^2 and μ being the variance and mean value respectively.

The central moment of order n is defined as the expectation value $(x - \mu)^n$ of a distribution with μ being the mean value, leading to

$$\mu_n^a = \frac{\sum_i (x - \mu)_i^n p_i}{\sum_i p_i} \tag{4.36}$$

for a discrete distribution p and to

$$\mu_n^a = \frac{\int (x - \mu)^n f(x) \mathrm{d}x}{\int f(x) \mathrm{d}x} \tag{4.37}$$

for a continuous distribution f(x). For the central moments the lowest orders are [56]

$$\mu_0^a = 1 \tag{4.38}$$
$$\mu_1^a = 0 \tag{4.39}$$
$$\mu_2^a = \sigma^2 \tag{4.40}$$

with σ^2 and μ being the variance and mean value respectively.

Moments of higher order are useful for studies of the behavior of $f(x)$ for large $|x - \mu|$. For a symmetric distribution, e.g. a gaussian distribution, all odd central moments vanish. Any non vanishing odd central moments therefore indicate an asymmetry of the distribution. The first of these moments is the third central moment. As this is not a dimensionless number, an absolute and dimensionless measurement of asymmetry, the *skewness*, is introduced as

$$\gamma = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{E\left[(x-\mu)^3\right]}{\sigma^3} \tag{4.41}$$

where E is the expectation value. For a distribution with a tail to the right γ is positive and with a tail to the left γ is negative [57]. The forth central moment is useful in describing how peaked the distribution is. The *kurtosis* or peakedness is defined as [56, 57]

$$c = \frac{\mu_4}{(\mu_2)^2} - 3 = \frac{E\left[(x-\mu)^4\right]}{\sigma^4} - 3 \tag{4.42}$$

The -3 in the definition is related to the skewness of the gaussian distribution which has - due to this offset - a skewness of zero.

Moments and the characteristic function A complete set of moments determines the probability density function $f(x)$ completely. To illustrate this a characteristic function $\phi_\mu(\mathrm{t})$ is introduced as the fourier transform of $f(x)$:

$$\begin{aligned}
\phi_\mu(t) &= \int_{-\infty}^{+\infty} \exp(it(x-\mu)) f(x) \mathrm{d}x & (4.43)\\
&= E\left[\exp(it(x-\mu))\right] & (4.44)\\
&= E\left[1 + it(x-\mu) + \frac{1}{2!}\left(it(x-\mu)\right)^2 + \frac{1}{3!}\left(it(x-\mu)\right)^3 \cdots\right] & (4.45)\\
&= E\left[\sum_{k=0}^{\infty} it^k \mu_k\right] & (4.46)
\end{aligned}$$

The central moments can then be obtained by

$$\mu_k = \frac{\partial^k \phi(t)}{\partial(it)^k}|_{t=0}. \tag{4.47}$$

Two dimensional distribution: Joint moments For a two dimensional distribution $f(x,y)$ central joint moments are introduced:

$$\mu_{kr} = E\left[(x-\eta_x)^k (y-\eta_y)^r\right] \tag{4.48}$$

where $\eta_x = E[x]$ and $\eta_y = E[y]$ are the mean values for x and y respectively. For a continuous distribution $f_{xy}(x,y)$ equation 4.48 becomes

$$\mu_{kr} = \int\int (x-\eta_x)^k (y-\eta_y)^r f_{xy}(x,y) \mathrm{d}x\mathrm{d}y \tag{4.49}$$

and for a discrete distribution $P(x,y)$ it becomes

$$\mu_{kr} = \frac{\sum_i \sum_j (x-\eta_x)^k (y-\eta_y)^r P(x_i,y_j)}{\sum_i \sum_j P(x_i,y_j)} \tag{4.50}$$

Note that the central joint moments of first order constitute the covariance matrix $V(x,y)$ with elements

$$V_{ij} = E\left[(x-\eta_x)(y-\eta_y)\right] \tag{4.51}$$

The moments of the signal distributions

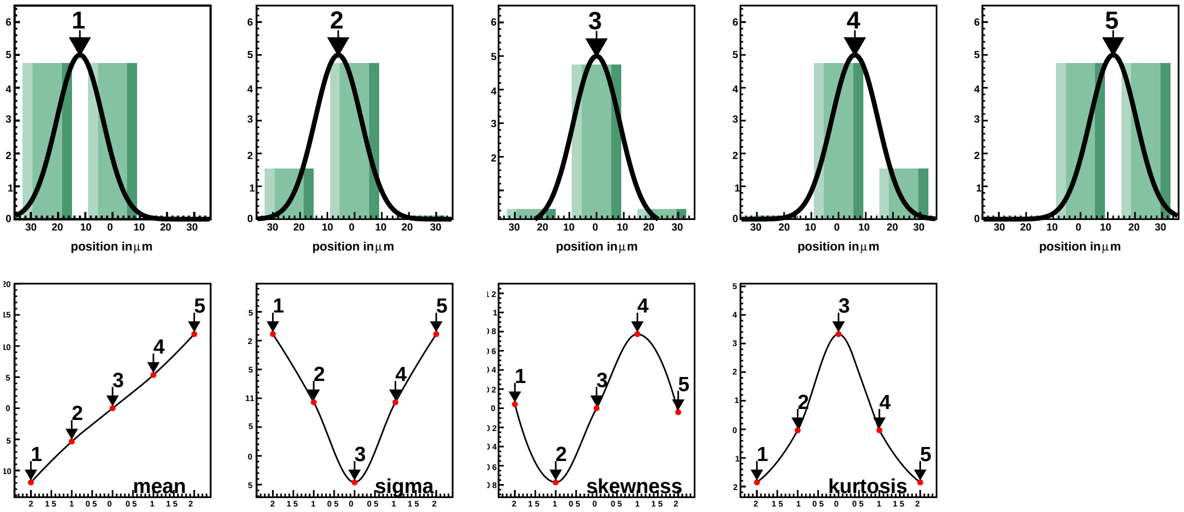


Figure 4.20: An illustration of the expected relationship between the first four moments mean, sigma, skewness, and kurtosis and the in-pixel position. The top row shows the five different impact points ranging from -pitch to +pitch and the resulting pixel signal distributions in arbitrary units for a 24 *μm pitch and a gaussian charge cloud with a width of* $\sigma = 8.5\,\mu m$*. The charge cloud is overlaid but not to scale with the pixel signal. The signal distributions are changing from a flat to a peaked and then to a flat distribution as can be seen in the sigma and kurtosis graph.*

As alternative input variables to the pure pixel signal the first four statistical moments of the spatial signal distribution were introduced. For completeness this includes the joint moments in both pixel axes. In one direction the expected behavior of the first four moments as a function of the in-pixel position is illustrated in figure 4.20 and follows from the Gaussian shaped charge cloud segmented in three pixels:

- The **mean** has an almost linear relationship to the in-pixel position. However, as already explained within the context of the *Center-of-Gravity*, the deviation from a linear relationship depends on the pixel to charge cloud size ratio. Nevertheless this variable is expected to contain the biggest amount of information about the in-pixel position.

- The **sigma** (r.m.s.) has a V shaped behavior and is symmetric with respect to the seed pixel center. If the particle hits the sensor on a pixel boundary the charge sharing is maximized and thus is the spatial signal spread. Therefore the sigma has its highest values here. Conversely, in the seed center it has its lowest value. The suitability of this variable will therefore not show up in the correlation ratio, but in the mutual information.

- The **skewness** dependence on the in-pixel position will have a sine like behavior.

This is due to the fact that on the pixel borders the charge sharing becomes symmetric again.

- The **kurtosis** reflects the "peakness" of the distribution and is therefore highest at the seed pixel center.

Figure 4.21 shows the dependence of the joint moments (3 × 3 cluster) in X and Y up to order four on the in-pixel position along the large pixel axis (X, 32 μm) for a pixel matrix simulated with GEANT. The first four moments along this axis, mean, r.m.s., skewness, and kurtosis, behave exactly as expected (fig. 4.20). Table 4.10 shows the correlation factor, correlation ratio, and mutual information for the joint moments of the spatial signal distribution within a 3×3 pixel cluster. Table 4.11 shows the same for a 5×5 pixel cluster.

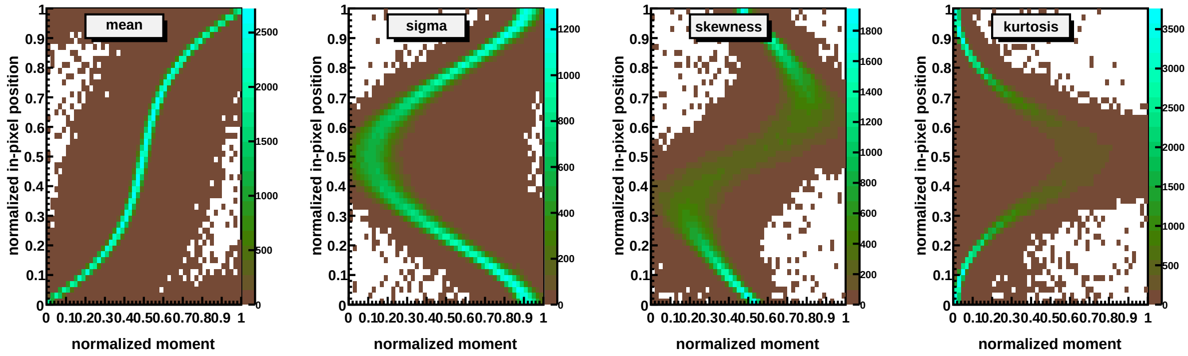
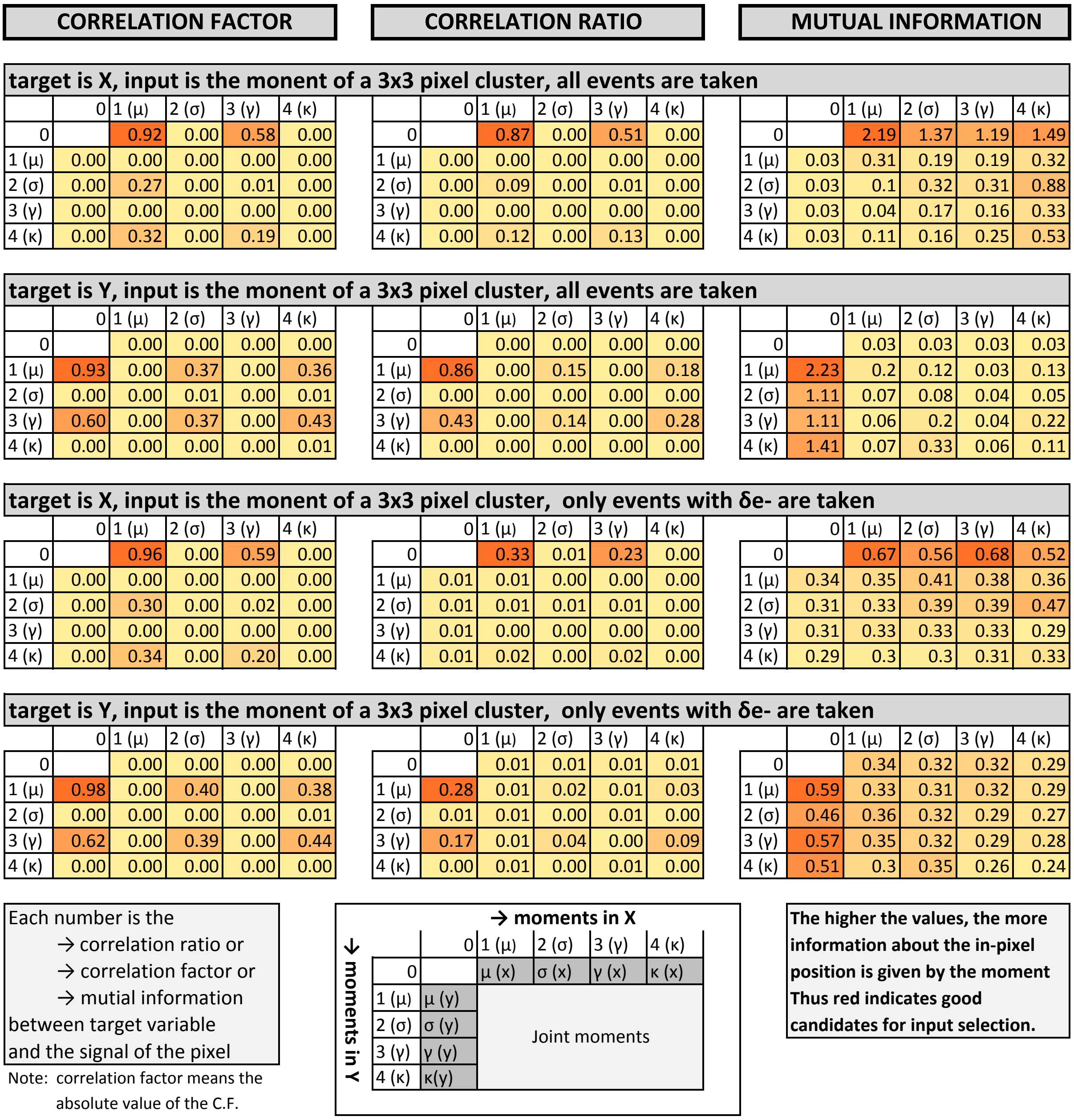


__Figure 4.21:__ This figure shows the in-pixel position as a function of the first four moments of the spatial signal distribution. Both, the in-pixel position and the moments are referring to the pixel X axis (32 μm) and are normalized, i.e. the variable range has been projected to the range $[0 \cdots 1]$. The moments behave exactly as expected (fig. 4.20), however, titled by 90°, because in this plot the input variable (moments) is on the X axis and the target variable (position) is on the Y axis.



CORRELATION FACTOR | **CORRELATION RATIO** | **MUTUAL INFORMATION**

target is X, input is the monent of a 3x3 pixel cluster, all events are taken

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.92	0.00	0.58	0.00	0		0.87	0.00	0.51	0.00	0		2.19	1.37	1.19	1.49
1 (μ)	0.00	0.00	0.00	0.00	0.00	1 (μ)	0.00	0.00	0.00	0.00	0.00	1 (μ)	0.03	0.31	0.19	0.19	0.32
2 (σ)	0.00	0.27	0.00	0.01	0.00	2 (σ)	0.00	0.09	0.00	0.01	0.00	2 (σ)	0.03	0.1	0.32	0.31	0.88
3 (γ)	0.00	0.00	0.00	0.00	0.00	3 (γ)	0.00	0.00	0.00	0.00	0.00	3 (γ)	0.03	0.04	0.17	0.16	0.33
4 (κ)	0.00	0.32	0.00	0.19	0.00	4 (κ)	0.00	0.12	0.00	0.13	0.00	4 (κ)	0.03	0.11	0.16	0.25	0.53

target is Y, input is the monent of a 3x3 pixel cluster, all events are taken

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.00	0.00	0.00	0.00	0		0.00	0.00	0.00	0.00	0		0.03	0.03	0.03	0.03
1 (μ)	0.93	0.00	0.37	0.00	0.36	1 (μ)	0.86	0.00	0.15	0.00	0.18	1 (μ)	2.23	0.2	0.12	0.03	0.13
2 (σ)	0.00	0.00	0.01	0.00	0.01	2 (σ)	0.00	0.00	0.00	0.00	0.00	2 (σ)	1.11	0.07	0.08	0.04	0.05
3 (γ)	0.60	0.00	0.37	0.00	0.43	3 (γ)	0.43	0.00	0.14	0.00	0.28	3 (γ)	1.11	0.06	0.2	0.04	0.22
4 (κ)	0.00	0.00	0.00	0.00	0.01	4 (κ)	0.00	0.00	0.00	0.00	0.00	4 (κ)	1.41	0.07	0.33	0.06	0.11

target is X, input is the monent of a 3x3 pixel cluster, only events with δe- are taken

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.96	0.00	0.59	0.00	0		0.33	0.01	0.23	0.00	0		0.67	0.56	0.68	0.52
1 (μ)	0.00	0.00	0.00	0.00	0.00	1 (μ)	0.01	0.01	0.00	0.00	0.00	1 (μ)	0.34	0.35	0.41	0.38	0.36
2 (σ)	0.00	0.30	0.00	0.02	0.00	2 (σ)	0.01	0.01	0.01	0.01	0.00	2 (σ)	0.31	0.33	0.39	0.39	0.47
3 (γ)	0.00	0.00	0.00	0.00	0.00	3 (γ)	0.01	0.00	0.00	0.00	0.00	3 (γ)	0.31	0.33	0.33	0.33	0.29
4 (κ)	0.00	0.34	0.00	0.20	0.00	4 (κ)	0.01	0.02	0.00	0.02	0.00	4 (κ)	0.29	0.3	0.3	0.31	0.33

target is Y, input is the monent of a 3x3 pixel cluster, only events with δe- are taken

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)		0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.00	0.00	0.00	0.00	0		0.01	0.01	0.01	0.01	0		0.34	0.32	0.32	0.29
1 (μ)	0.98	0.00	0.40	0.00	0.38	1 (μ)	0.28	0.01	0.02	0.01	0.03	1 (μ)	0.59	0.33	0.31	0.32	0.29
2 (σ)	0.00	0.00	0.00	0.00	0.01	2 (σ)	0.01	0.01	0.00	0.01	0.00	2 (σ)	0.46	0.36	0.32	0.29	0.27
3 (γ)	0.62	0.00	0.39	0.00	0.44	3 (γ)	0.17	0.01	0.04	0.00	0.09	3 (γ)	0.57	0.35	0.32	0.29	0.28
4 (κ)	0.00	0.00	0.00	0.00	0.00	4 (κ)	0.00	0.01	0.00	0.01	0.00	4 (κ)	0.51	0.3	0.35	0.26	0.24

Each number is the
→ correlation ratio or
→ correlation factor or
→ mutial information
between target variable
and the signal of the pixel

Note: correlation factor means the absolute value of the C.F.

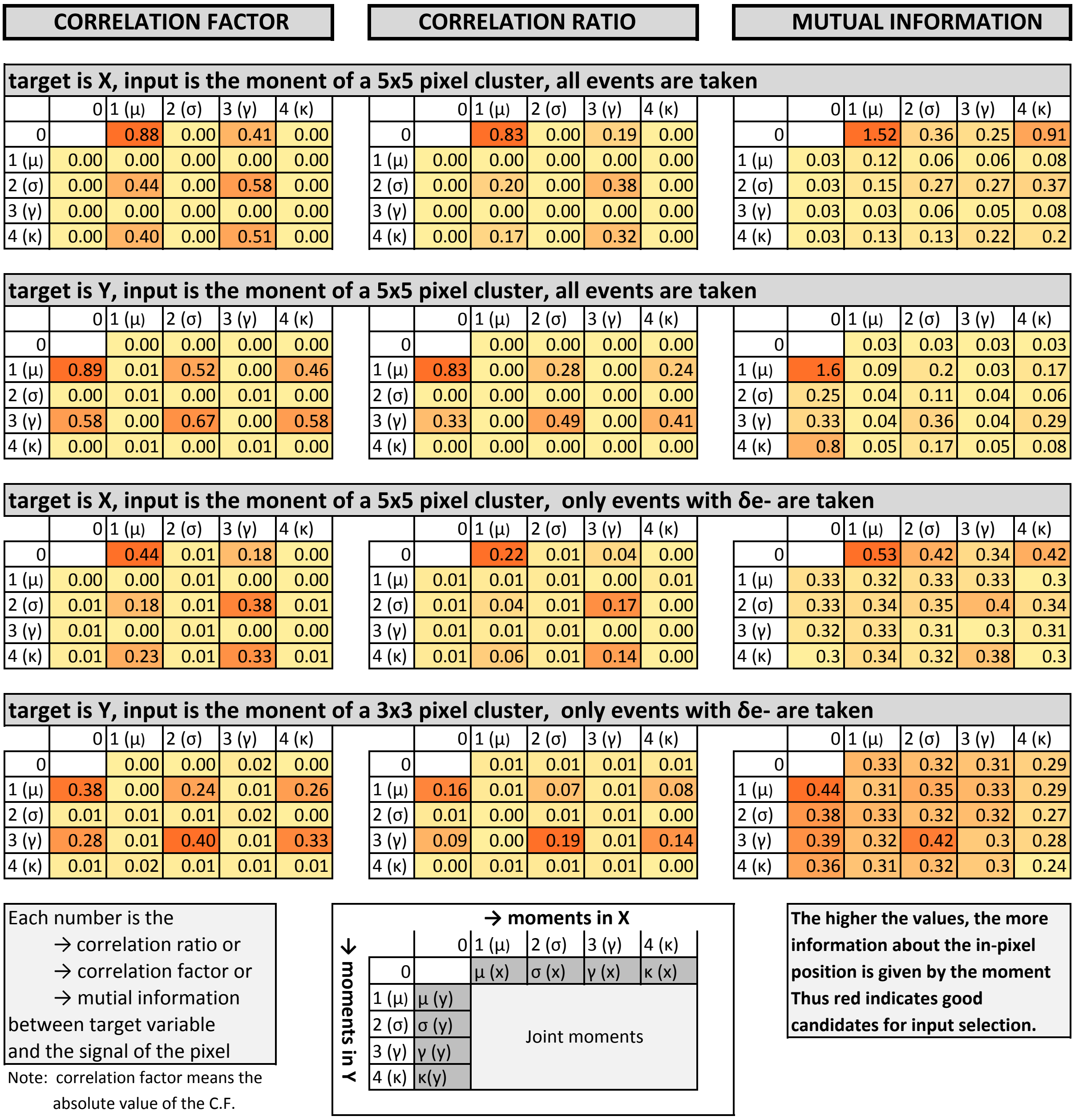
→ moments in X

↓ moments in Y

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		μ (x)	σ (x)	γ (x)	κ (x)
1 (μ)	μ (y)	Joint moments			
2 (σ)	σ (y)				
3 (γ)	γ (y)				
4 (κ)	κ(y)				

The higher the values, the more information about the in-pixel position is given by the moment Thus red indicates good candidates for input selection.

Table 4.10: Tables of mutual information, correlation ratio, and correlation factor for ***statistical moments****. The moments are calculated for the spatial signal distribution of a* 3×3 *pixel clusters. Each table contains the corresponding numbers for the moments up to the fourth order in X and Y direction including joint moments as is indicated at the bottom. Each number indicates the usefulness of the corresponding moment as an input variable for the multivariate analysis. For the tables in the left, middle, and right column different statistical quantities are used to evaluate the usefulness of a moment as an input variable. This study was done for the X and Y axis as target variable and for two different event classes: All events, and events that are required to have a least one δ-electron. The data is a GEANT simulation with* $32 \times 24\,\mu m^2$ *sized pixels.*



CORRELATION FACTOR | **CORRELATION RATIO** | **MUTUAL INFORMATION**

target is X, input is the moment of a 5x5 pixel cluster, all events are taken

Correlation factor:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.88	0.00	0.41	0.00
1 (μ)	0.00	0.00	0.00	0.00	0.00
2 (σ)	0.00	0.44	0.00	0.58	0.00
3 (γ)	0.00	0.00	0.00	0.00	0.00
4 (κ)	0.00	0.40	0.00	0.51	0.00

Correlation ratio:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.83	0.00	0.19	0.00
1 (μ)	0.00	0.00	0.00	0.00	0.00
2 (σ)	0.00	0.20	0.00	0.38	0.00
3 (γ)	0.00	0.00	0.00	0.00	0.00
4 (κ)	0.00	0.17	0.00	0.32	0.00

Mutual information:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		1.52	0.36	0.25	0.91
1 (μ)	0.03	0.12	0.06	0.06	0.08
2 (σ)	0.03	0.15	0.27	0.27	0.37
3 (γ)	0.03	0.03	0.06	0.05	0.08
4 (κ)	0.03	0.13	0.13	0.22	0.2

target is Y, input is the moment of a 5x5 pixel cluster, all events are taken

Correlation factor:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.00	0.00	0.00	0.00
1 (μ)	0.89	0.01	0.52	0.00	0.46
2 (σ)	0.00	0.01	0.00	0.01	0.00
3 (γ)	0.58	0.00	0.67	0.00	0.58
4 (κ)	0.00	0.01	0.00	0.01	0.00

Correlation ratio:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.00	0.00	0.00	0.00
1 (μ)	0.83	0.00	0.28	0.00	0.24
2 (σ)	0.00	0.00	0.00	0.00	0.00
3 (γ)	0.33	0.00	0.49	0.00	0.41
4 (κ)	0.00	0.00	0.00	0.00	0.00

Mutual information:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.03	0.03	0.03	0.03
1 (μ)	1.6	0.09	0.2	0.03	0.17
2 (σ)	0.25	0.04	0.11	0.04	0.06
3 (γ)	0.33	0.04	0.36	0.04	0.29
4 (κ)	0.8	0.05	0.17	0.05	0.08

target is X, input is the moment of a 5x5 pixel cluster, only events with δe- are taken

Correlation factor:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.44	0.01	0.18	0.00
1 (μ)	0.00	0.00	0.01	0.00	0.00
2 (σ)	0.01	0.18	0.01	0.38	0.01
3 (γ)	0.01	0.00	0.01	0.00	0.00
4 (κ)	0.01	0.23	0.01	0.33	0.01

Correlation ratio:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.22	0.01	0.04	0.00
1 (μ)	0.01	0.01	0.01	0.00	0.01
2 (σ)	0.01	0.04	0.01	0.17	0.00
3 (γ)	0.01	0.01	0.01	0.00	0.00
4 (κ)	0.01	0.06	0.01	0.14	0.00

Mutual information:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.53	0.42	0.34	0.42
1 (μ)	0.33	0.32	0.33	0.33	0.3
2 (σ)	0.33	0.34	0.35	0.4	0.34
3 (γ)	0.32	0.33	0.31	0.3	0.31
4 (κ)	0.3	0.34	0.32	0.38	0.3

target is Y, input is the moment of a 3x3 pixel cluster, only events with δe- are taken

Correlation factor:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.00	0.00	0.02	0.00
1 (μ)	0.38	0.00	0.24	0.01	0.26
2 (σ)	0.01	0.01	0.01	0.02	0.00
3 (γ)	0.28	0.01	0.40	0.01	0.33
4 (κ)	0.01	0.02	0.01	0.01	0.01

Correlation ratio:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.01	0.01	0.01	0.01
1 (μ)	0.16	0.01	0.07	0.01	0.08
2 (σ)	0.01	0.00	0.01	0.01	0.00
3 (γ)	0.09	0.00	0.19	0.01	0.14
4 (κ)	0.00	0.01	0.01	0.01	0.00

Mutual information:

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		0.33	0.32	0.31	0.29
1 (μ)	0.44	0.31	0.35	0.33	0.29
2 (σ)	0.38	0.33	0.32	0.32	0.27
3 (γ)	0.39	0.32	0.42	0.3	0.28
4 (κ)	0.36	0.31	0.32	0.3	0.24

Each number is the
→ correlation ratio or
→ correlation factor or
→ mutial information
between target variable
and the signal of the pixel

Note: correlation factor means the absolute value of the C.F.

→ **moments in X**

↓ **moments in Y**

	0	1 (μ)	2 (σ)	3 (γ)	4 (κ)
0		μ (x)	σ (x)	γ (x)	κ (x)
1 (μ)	μ (y)	Joint moments			
2 (σ)	σ (y)				
3 (γ)	γ (y)				
4 (κ)	κ(y)				

The higher the values, the more information about the in-pixel position is given by the moment Thus red indicates good candidates for input selection.

Table 4.11: *Tables of mutual information, correlation ratio, and correlation factor for* ***statistical moments****. The moments are calculated for the spatial signal distribution of a* 5×5 *pixel clusters. Each table contains the corresponding numbers for the moments up to the fourth order in X and Y direction including joint moments as is indicated at the bottom. Each number indicates the usefulness of the corresponding moment as an input variable for the multivariate analysis. For the tables in the left, middle, and right column different statistical quantities are used to evaluate the usefulness of a moment as an input variable. This study was done for the X and Y axis as target variable and for two different event classes: All events, and events that are required to have a least one δ-electron. The data is a GEANT simulation with* $32 \times 24\,\mu m^2$ *sized pixels.*

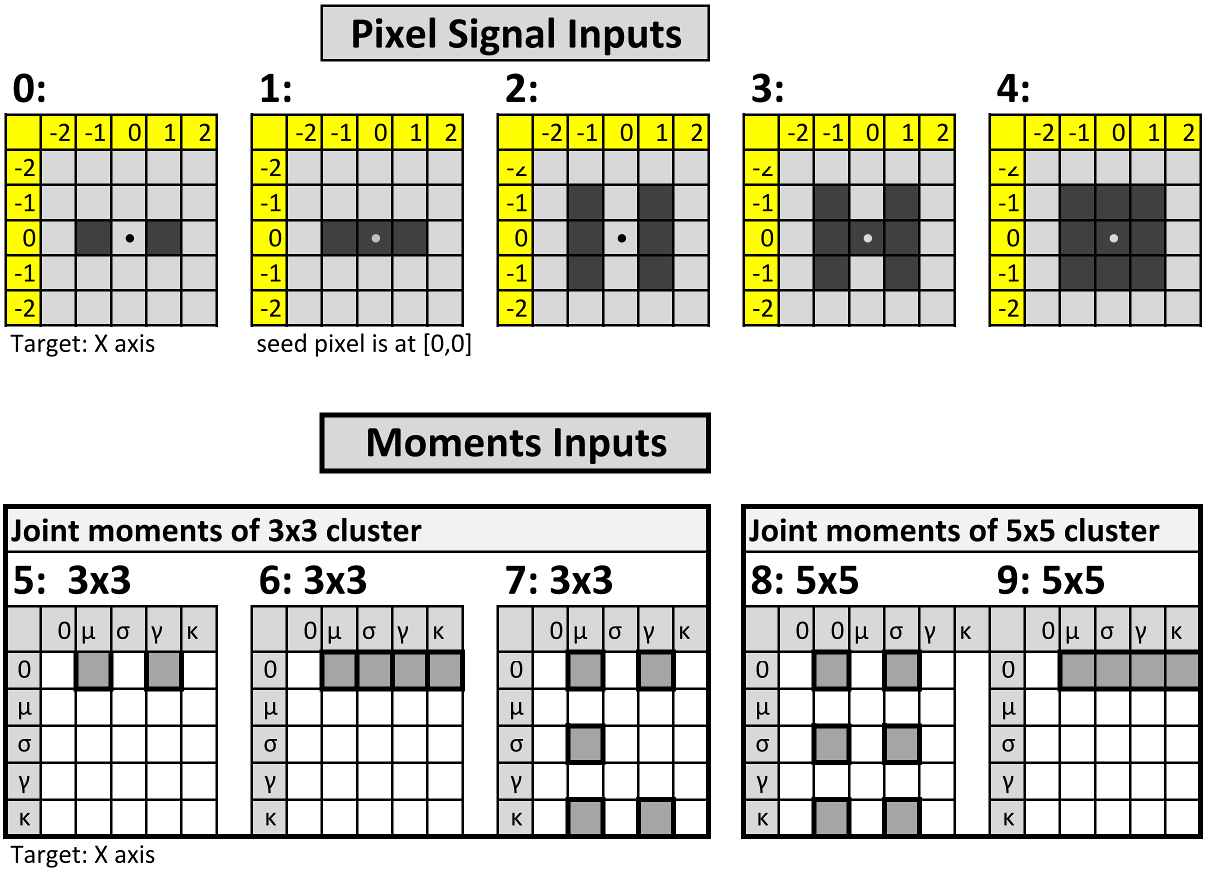
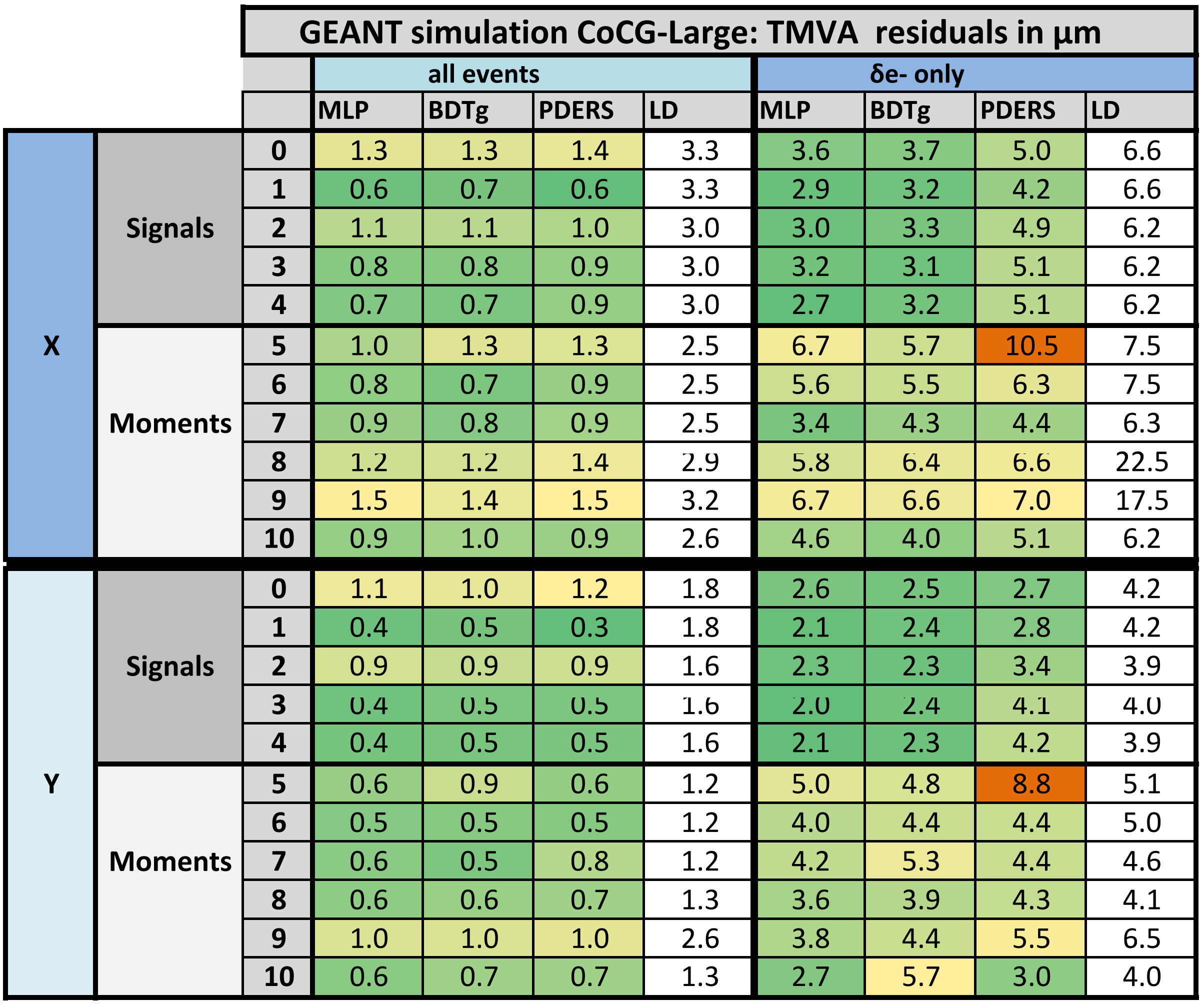


Figure 4.22: *The input variable configurations for pixel signals and moments with the in-pixel position along the X axis as target variable. The upper row shows the configurations based on the pure pixel signal. The lower row shows the moment based configurations. In the lower row the moments which were calculated for the spatial signal distribution of* 3×3 *and* 5×5 *pixel clusters, respectively, are separated. The configuration for the in-pixel position in Y as target variable is simply the transpose of the X axis configuration. The 10th configuration is not shown here and is the combination of first (mean) and third (skewness) moments of* 3×3 *and* 5×5 *pixel clusters.*

4.3.2 Performance of the multivariate analysis

Based on the previous results 11 different sets of input variables were selected for further investigation. Ten of them are shown in figure 4.22. The four above mentioned multivariate analysis techniques - multilayer preceptron neural network (MLP), boosted decision trees (BDT), probability density estimator (PDERS), and linear discriminant analysis (LD) - were applied to both the GEANT simulation as well as to the real test beam data using these input sets. Akin to the η and the charge cloud fit studies, the data was split up into the same energy ranges (fig. 4.2(a)) and additionally for the GEANT simulation into events with and without the secondary particles flag, i.e. δ-electrons.

GEANT specific cuts



			GEANT simulation CoCG-Large: TMVA residuals in µm							
			all events				δe- only			
			MLP	BDTg	PDERS	LD	MLP	BDTg	PDERS	LD
X	Signals	0	1.3	1.3	1.4	3.3	3.6	3.7	5.0	6.6
		1	0.6	0.7	0.6	3.3	2.9	3.2	4.2	6.6
		2	1.1	1.1	1.0	3.0	3.0	3.3	4.9	6.2
		3	0.8	0.8	0.9	3.0	3.2	3.1	5.1	6.2
		4	0.7	0.7	0.9	3.0	2.7	3.2	5.1	6.2
	Moments	5	1.0	1.3	1.3	2.5	6.7	5.7	10.5	7.5
		6	0.8	0.7	0.9	2.5	5.6	5.5	6.3	7.5
		7	0.9	0.8	0.9	2.5	3.4	4.3	4.4	6.3
		8	1.2	1.2	1.4	2.9	5.8	6.4	6.6	22.5
		9	1.5	1.4	1.5	3.2	6.7	6.6	7.0	17.5
		10	0.9	1.0	0.9	2.6	4.6	4.0	5.1	6.2
Y	Signals	0	1.1	1.0	1.2	1.8	2.6	2.5	2.7	4.2
		1	0.4	0.5	0.3	1.8	2.1	2.4	2.8	4.2
		2	0.9	0.9	0.9	1.6	2.3	2.3	3.4	3.9
		3	0.4	0.5	0.5	1.6	2.0	2.4	4.1	4.0
		4	0.4	0.5	0.5	1.6	2.1	2.3	4.2	3.9
	Moments	5	0.6	0.9	0.6	1.2	5.0	4.8	8.8	5.1
		6	0.5	0.5	0.5	1.2	4.0	4.4	4.4	5.0
		7	0.6	0.5	0.8	1.2	4.2	5.3	4.4	4.6
		8	0.6	0.6	0.7	1.3	3.6	3.9	4.3	4.1
		9	1.0	1.0	1.0	2.6	3.8	4.4	5.5	6.5
		10	0.6	0.7	0.7	1.3	2.7	5.7	3.0	4.0

Table 4.12: *The width σ (in μm) of a gaussian fit to the residuals for the four multivariate methods described in the text. The underlying data is a GEANT simulation of a CoCG-Large like pixel matrix. The table is split into four quadrants with the upper and the lower half showing the results for X and Y axis, respectively, and the left and right half showing the difference if no cut on the data is applied and if only events with δ-electrons are taken. The input variable configurations 0 to 10 are shown in figure 4.22.*

Table 4.12 shows the results for the GEANT data when no cut on the data is made (all events) and when only δ-electron events are used. The performance of the MVAs in the absence of a cut can be summarized as:

- The linear discriminant method (LD) is drastically outperformed by the other three MVA methods. It also performs much worse than the η method and the charge clouds fit.
- The best results are achieved with input configuration 1 and 6 and the results are always better for the smaller pixel sizes. This result is expected as the analysis of

the information content of possible input variables in the previous section already showed that the in-pixel position information is confined to the seed pixel and its two neighbors. Adding further information like the entire 3×3 pixel cluster (input configuration 4) does not improve the resolution but actually might worsen it, as the multi-dimensionality of the problem is increased by noisy dimensions/input variables.

- The inputs based on the statistical moments perform worse than the input based on the pure signal inputs. As the information is stored in just three pixels, a moment analysis of the spatial signal distribution is apparently a somewhat oversized method, especially the moments based on the 5×5 pixel cluster give worse results. However, the Y axis with the smaller pixel pitch shows generally a smaller gap between pixel signal and moment based input variables. For an even smaller pixel to charge cloud size ratio the moment approximation of the spatial signal distribution might become useful.
- The seed signal is essential. Input set 0 and 2 exclude the seed signal and the results are considerable worse than with an input set including the seed.
- When working with the moments the first four moments with respect to the target direction give the best results.
- Except for the linear discriminant the MVA methods show comparable results with PDE-RS outperforming the other when the input variables are restricted only to those carrying sufficient information.

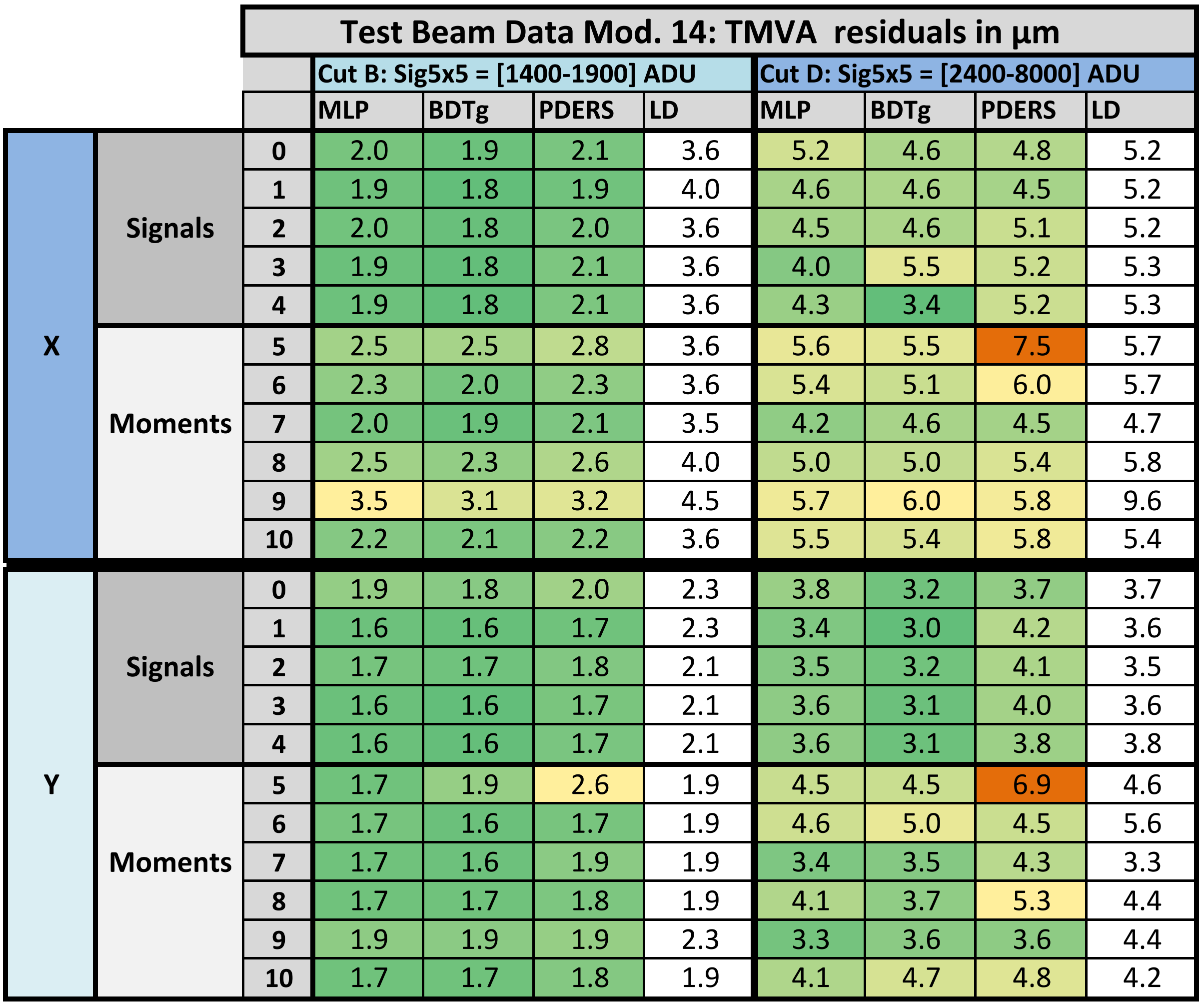
Where the simulated data set is restricted only to events with a δ-electron the results are:

- MLP and BDT show the best results, whereas the PDE-RS performs considerably worse.
- The pixel signal inputs outperform the moment based input configurations, although a restriction to only the seed and its two neighbors does not necessarily enhance the performance.
- For the moment based inputs a mixture of 3×3 and 5×5 cluster moments or even joint moments might possibly increase the performance.

Test beam data

Table 4.13 shows the results for the test beam data taken with module 14 (CoCG-Large type) for cut **B** and **D** as shown in fig. 4.2(a). The findings are representative for other test beam modules. For cut $B = 1400 - 1900\,\mathrm{ADU}$ - around the most probable value peak of the Landau distribution - they are:

- The linear discriminant is outperformed by the other three MVA methods.



Test Beam Data Mod. 14: TMVA residuals in µm										
			Cut B: Sig5x5 = [1400-1900] ADU				Cut D: Sig5x5 = [2400-8000] ADU			
			MLP	BDTg	PDERS	LD	MLP	BDTg	PDERS	LD
X	**Signals**	**0**	2.0	1.9	2.1	3.6	5.2	4.6	4.8	5.2
		1	1.9	1.8	1.9	4.0	4.6	4.6	4.5	5.2
		2	2.0	1.8	2.0	3.6	4.5	4.6	5.1	5.2
		3	1.9	1.8	2.1	3.6	4.0	5.5	5.2	5.3
		4	1.9	1.8	2.1	3.6	4.3	3.4	5.2	5.3
	Moments	**5**	2.5	2.5	2.8	3.6	5.6	5.5	7.5	5.7
		6	2.3	2.0	2.3	3.6	5.4	5.1	6.0	5.7
		7	2.0	1.9	2.1	3.5	4.2	4.6	4.5	4.7
		8	2.5	2.3	2.6	4.0	5.0	5.0	5.4	5.8
		9	3.5	3.1	3.2	4.5	5.7	6.0	5.8	9.6
		10	2.2	2.1	2.2	3.6	5.5	5.4	5.8	5.4
Y	**Signals**	**0**	1.9	1.8	2.0	2.3	3.8	3.2	3.7	3.7
		1	1.6	1.6	1.7	2.3	3.4	3.0	4.2	3.6
		2	1.7	1.7	1.8	2.1	3.5	3.2	4.1	3.5
		3	1.6	1.6	1.7	2.1	3.6	3.1	4.0	3.6
		4	1.6	1.6	1.7	2.1	3.6	3.1	3.8	3.8
	Moments	**5**	1.7	1.9	2.6	1.9	4.5	4.5	6.9	4.6
		6	1.7	1.6	1.7	1.9	4.6	5.0	4.5	5.6
		7	1.7	1.6	1.9	1.9	3.4	3.5	4.3	3.3
		8	1.7	1.7	1.8	1.9	4.1	3.7	5.3	4.4
		9	1.9	1.9	1.9	2.3	3.3	3.6	3.6	4.4
		10	1.7	1.7	1.8	1.9	4.1	4.7	4.8	4.2

Table 4.13: *The residuals width σ (in μm) for test beam data of module 14 (CoCG-L) and the four multivariate methods described in the text. The table is split into four quadrants with the upper and the lower half showing the results for X and Y axis respectively. The left and right parts show the results for different cuts on the cluster energy as depicted in the table. The input variable configurations 0 to 10 are shown in figure 4.22.*

- With one exception the results for the shorter Y axis are almost independent from the input variable. The difference between pixel signal and statistical moment based inputs along the shorter Y axis are marginal unlike in the GEANT simulation. For the X axis the signal based inputs still perform better, again indicating that moment based input variables might be more useful if the charge cloud is spread over more pixels.

- The general performance of MLP, BDT, and PDERS is comparable to the η method and the charge cloud fit method and dominated by effects of the test beam experiment (multiple scattering, telescope resolution, etc.).

All these results have been obtained without fine tuning the TMVA settings and there is certainly room for improvement, however this is beyond the scope of this study.

4.4 Summary

This chapter explored the possibility for new position reconstruction methods with an emphasis on the δ-electron induced resolution worsening at higher energies. The results can be summarized as follows:

- For the given DEPFET sensor with $450\,\mu$m thickness and pixel dimensions of $32 \times 24\,\mu\text{m}^2$ the charge cloud has a width around $7 - 10\,\mu$m and is basically confined to the seed pixel and its two neighbors.

- For the majority of events with an energy close to the most probable value none of the studied method can significantly improve the traditional single η_{one} method. The reason for that is this - with the given charge cloud to pixel size - the charge and therefore the position information is mainly inside the seed pixel and its neighbor with the highest signal. However, for smaller pixel sizes the methods presented in this chapter are likely to improve over the η method.

- Although the traditional single η_{one} method fails to account for the changing charge sharing in events with more deposited energy, a simple approach with two η_{multi} distributions can drastically improve the spatial precision. One η_{MPV} distribution describes the charge sharing for events around the most probable value. The other $\eta_{\delta-e^-}$ covers events in the high energetic Landau tail.

- The charge cloud fit method has a position prediction power similar to the η_{multi} method but at a cost of somewhat higher algorithmic expenditure.

- The four multivariate methods, artificial neural networks (MLP), boosted decision trees (BDT), a probability density estimator (PDE), and linear discriminant (LD), did not improve the position reconstruction capabilities. However, as for the majority of events the information is stored in just two variables (the seed and its highest neighbor) the underlying problem is not really of a multidimensional nature. Unless a detector with rather small pixels with respect to charge cloud extension is employed, the additional technical effort of multivariate analysis is not offset by an advanced position reconstruction capability. The same is basically true for the high energetic part of the Landau distribution.

5

Summary and Conclusion

In particle collider experiments jet flavor tagging is an integral requirement to answer the current questions of particle physics and its related fields like cosmology. To achieve this, vertex detectors with excellent spatial resolution are needed. At the same time the detector should be as thin as possible. The DEPFET pixel sensor technology is a promising vertex detector candidate as its integrated first stage amplification allows for a thinned down sensor which still has a good signal to noise ratio. Therefore a DEPFET pixel based vertex detector concept for the International Linear Collider (ILC) was put forward. It envisions five layers of DEPFET matrices with 24x24 μm^2 sized pixels and row wise operation. The detector will be thinned down to well below 100 μm and have a total material budget contribution of $\sim$ 0.1% radiation length. With the inner layer placed at a radius of only 15 mm away from the interaction point the impact parameter resolution will be $\lesssim 4\,\mu$m. Since the occupancy should not exceed 1% the entire vertex detector needs to be read out in 50 μs which translates to a row rate of 20 MHz.
To advance this concept a prototype system was build with a 64×128 DEPFET pixel matrix with $32 \times 24\,\mu$m^2 sized pixel as the core component. The matrix is accompanied by two designated steering ASICs (SWITCHER) and a designated read out ASIC (CURO). The system has been successfully operated in the laboratory as well as in a beam test environment with an excellent signal to noise of over ratio 120 and a spatial resolution of better than 2 μm.
Several DEPFET matrices were used as a beam test telescope reference system in a large scale test beam campaign at the SPS facility at CERN with 120 GeV pions. This allowed for the first time to study the sensor homogeneity on an in-pixel scale. A near perfect homogeneity of the sensor is desired as any inhomogeneities would worsen the resolution and therefore the jet flavor tagging capabilities of a future vertex detector. A near perfect in-pixel homogeneity was observed. Although some sensors showed inhomogeneities of up to 8% r.m.s., this could be traced back to faulty biasing conditions.

The signal spectrum of a thin sensor follows the asymmetric Landau distribution. The long tail at higher energies is due to δ-electrons, which are knocked out of the atoms shell and given enough momentum to become a secondary particle inside the sensor. The tail of deposited energy left behind by this δ-electron will impair the reconstructed position of the original particle. The DEPFET sensor will be thinned down to reduce the material. This will inevitably change the signal spectrum of the silicon sensor to an even more asymmetric Landau distribution with a higher fraction of δ-electron affected events. Traditionally in a beam test, these events are thrown away, however, a position reconstruction technique that would be less susceptible to the influence of δ-electrons could increase the efficiency of the sensor. To tackle this problem and to generally find a better alternative to the standard position reconstruction method for silicon particle detectors, the so-called η method, a range of methods was studied. This study encompassed the test beam data gathered with the DEPFET telescope system and a simulated data set created with the GEANT 4 software. Three types of methods were involved in this study:

- An extension of the standard η method: The η method samples the charge sharing between the seed and its highest neighbor in the so-called η distribution and uses this distribution for position reconstruction. The novel approach is to have multiple η distributions for different energy regimes. This method reduced the position error caused by δ-electrons in simulated events from $1.8\,\mu$m to $0.8\,\mu$m (X axis, $32\,\mu$m wide) and from $1.3\,\mu$m to $0.5\,\mu$m (Y axis, $32\,\mu$m wide)

- A fit to the charge cloud shape: Using the high statistics and spatial resolution from the test beam data, the charge cloud was sampled and several analytical models were fitted. Using these functions a position reconstruction method based on a least square fit of the three central pixel signals was developed. This method showed in general similar results to the multi η method, although selecting the right function to model the charge cloud shape proved to be crucial.

- Multivariate analysis (MVA) methods: Four MVA techniques were applied for position reconstruction - linear regression, neural network, range-search probability density estimation, and boosted decision trees - using several preselected sets of input variables. The input configurations were either based on pixel signals or on statistical moments of the spatial signal distribution. Estimations of the information content of these input variables showed that the particle position is generally confined to the seed pixel and its two neighbors. These findings were confirmed by the MVA results and are also in accordance with theoretical predictions and measurements of the charge cloud size. Furthermore, the MVA results were at best equivalent to the η method, but often performed worse.

The general conclusion from this study is that for the given pixel to charge cloud size ratio of 2 to 3 the majority of the charge and therefore the information about the particles impact point is confined to the seed and its highest neighbor. This explains the excellent performance of the η method compared to much more advanced algorithms like neural networks. Even for the high energy part of the signal spectrum that is largely δ-electron influenced a simple approach with several η distributions shows the best effectiveness in terms of resolution improvement vs. algorithmic complexity.

Bibliography

[1] K. Nakamura and et. al., “Review of Particle Physics,” *J. Phys. G 37, 075021*, 2010.

[2] H. Bethe, “Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie,” *Annalen d. Phys., 397*, 1930.

[3] W. Leo, *Techniques for Nuclear and Particle Physics Experiments.* Springer, 1987.

[4] C. Leroy and P.-G. Rancoita, *Principles of radiation interaction in matter and detectors.* World Scientific Publishing, 2009.

[5] L. Landau, “On the energy loss os fast particles by ionization,” *J. Phys. (USSR) 8:201*, 1944.

[6] H. Bichsel, “Straggling in thin silicon detetcors,” *Rev. of Modern Phys., Vol 60, No 3*, 1988.

[7] C. Grupen, *Particle Detectors.* Cambridge University Press, 1996.

[8] I. Antcheva *et al.*, “Root – a c++ framework for petabyte data storage, statistical analysis and visualization,” *Computer Physics Communications*, vol. 180, pp. 2499 – 2512, 2009.

[9] M. Bergen and S. Seltzer, “Tables of energy losses and ranges of electrons and positrons”,” *NASA Report NASA-SP-3012*, 1964.

[10] L. Reimer and E. Lohrmann, *Scanning Electron Microscopy.* Springer, 1998.

[11] K. Kanaya and S. Okayama, “Penetration and energy-loss theory of eletrons in solid targets,” *J. Phys. D: Appl. Phys., Vol 5*, 1972.

[12] T. Everhart and P. Hoff, “Determination of kilovolt electron energy dissipation vs penetration distance in solid materials,” *J. o. Appl. Phys., Vol 42, Issue 13*, 1971.

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

[14] G. Lutz, *Semiconductor Radiation Detectors.* Springer, Berlin, 1999.

[15] E. Gatti, A. Longoni, P. Rehak, and M. Sempietro, “Dynamics of electrons in dirft detetcors,” *Nucl. Instr. and Meth. A253*, 1987.

[16] E. Gatti and P. Rehak, "Semiconductor drift chambers - an application of a novel charge transport scheme," *Nucl. Instr. and Meth. A225*, 1984.

[17] D. Collaboration, "DEPFET pixel vertex detector for the ilc," *ILC VDX Review, Oct 2007, FNAL*, 2007.

[18] C. Sandow, "Ladungsansammlung und Loeschvorgang von DEPFET-Pixel-Sensoren," Master's thesis, Bonn University, 2005.

[19] P. Hettkamp, "Messungen zur Charakterisierung der Ladungsmessungen von DEPFET-Matrizen durch Lasereinstrahlung," Master's thesis, Bonn University, 2007.

[20] K. Schmieden, "Charakterisierung einer neuen Generation von DEPFET-Sensoren mit Hilfe eines Lasermesssystems," Master's thesis, Bonn University, 2008.

[21] S. Esch, "Auslese neuer DEPFET-Sensoren mit verbesserter Ansteuerung und charakterisierende Messungen zu Antwortverhalten und Homogenitaet," Master's thesis, Bonn University, 2009.

[22] R. Kohrs, *Development and Characterization of a DEPFET Pixel Prototype for the ILC Vertex Detector.* PhD thesis, Bonn University, 2008.

[23] C. Damerell and D. Jackson, "Vertex detector technology and jet flavour identification at the future e^+ e^- linear collider," *DPB Summer Study (Snowmass 96)*, 1996.

[24] C. Damerell, "Vertex detectors and the linear collider," *LCWS 05, Standford, USA*, 2005.

[25] J. Abdallah and et al., "b-tagging in delphi at lep," *Eur.Phys.J. C, 32*, 2004.

[26] T. Behnke, C. Damerell, J. Jaros, and A. Myamoto, "ILC reference design report volume 4 - detectors," *ILC-REPORT-2007-001*, 2007.

[27] N. Wermes, "Pixel detectors for charged particles," *Nucl. Instr. and Meth. A 604*, 2009.

[28] E. Anderssen *et al.*, "Fluorocarbon evaporative cooling developments for the ATLAS pixel and semiconductor tracking detectors," Tech. Rep. CERN-OPEN-2000-093. ATL-INDET-99-016, CERN, Geneva, Oct 1999.

[29] G. Aad and et. al., "ATLAS pixel detector electronics and sensors," *JINST*, vol. 3, 2008.

[30] T. Behnke, S. Bertolucci, R. D. Heuer, and R. Settles, "TESLA: The superconducting electron positron linear collider with an integrated x-ray laser laboratory. technical design report. pt. 4: A detector for TESLA," *DESY-01-011.*

[31] M. Trimpl, *Design of a current based readout chip and development of a DEPFET pixel prototype system for the ILC vertex detector.* PhD thesis, Bonn University, 2005.

[32] P. Fischer, C. Kreidl, and I. Peric, "Steering and readout chips for DEPFET sensor matrics," *Proc. of the TWEPP 2007, Prague*, 2007.

[33] L. Andricek, G. Lutz, M. Reiche, and R. H. Richter, "Processing of ultra -thin silicon sensors for future e+ e- linear collider experiments," *IEEE Trans. o. Nucl. Sci. Vol. 51, No. 3*, 2004.

[34] N. Wermes and et al., "New results on DEPFET pixel detectors for radiation imaging and high energy particle detection," *IEEE Trans. Nucl. Sci. Vol. 51*, 2004.

[35] R. Kohrs, L. Andricek, P. Fischer, M. Harter, M. Karagounis, H. Krueger, G. Lutz, H. Moser, I. Peric, M. Porro, L. Reuen, R. Richter, C. Sandow, L. Strueder, M. Trimpl, and N. Wermes, "Development of a prototype module for a DEPFET pixel vertex detector for a linear collider," *IEEE Trans. Nucl. Sci. Vol. 52*, 2005.

[36] R. Kohrs, L. Reuen, L. Andricek, P. Fischer, G. Giesen, H. Krueger, G. Lutz, M. Mathes, H. Moser, I. Peric, R. Richter, C. Sandow, J. Treis, M. Trimpl, E. v. Toerne, J. J. Velthuis, and N. Wermes, "First test beam results on DEPFET pixels for the ILC," *Nucl. Instr. Meth. A 565*, 2006.

[37] M. Trimpl, L. Andricek, P. Fischer, R. Kohrs, H. Krueger, G. Lutz, H. Moser, I. Peric, M. Porro, L. Reuen, R. Richter, C. Sandow, L. Strueder, J. Treis, and N. Wermes, "A DEPFET pixel matrix system for the ilc vertex detector," *Nucl. Instr. Meth. A 560, 21-25*, 2006.

[38] L. Reuen and et al., "Performance of a DEPFET prototype module for the ilc vertex detector," *IEEE Trans. o. Nucl. Sc. Vol. 53, No. 3*, 2006.

[39] J. Velthuis, R. Kohrs, L. Reuen, H. Krueger, M. Mathes, E. von Toerne, M. Trimpl, and N. W. et al., "Status of DEPFET," *Nucl. Instr. Meth. A 569*, 2006.

[40] J. Velthuis, R. Kohrs, M. Mathes, L. Reuen, M. Koch, H.Krueger, E. von Toerne, and N. W. et al., "Depfet, a monolithic active pixel sensor for the ilc," *Nucl. Instr. Meth. A 579*, 2007.

[41] J. J. Velthuis, G. Nunes, Hanninger, R. Kohrs, M. Mathes, L. Reuen, M. Koch, H. Krueger, E. von Toerne, and N. W. et al., "A DEPFET based beam telescope with submicron precision capability," *IEEE Trans. Nucl. Sci. Vol. 55*, 2008.

[42] L. Reuen, "The DEPFET sensor as the first device ander test using the EUDET JRA1 telescope," *EUDET-Memo-2007-40*, 2007.

[43] J. Furletova, "The integration of DEPFET pixel detector into the EUDET telescope," *IEEE 2008, Nuclear Sience Symposium Dresden*, 2008.

[44] J. Furletova and L. Reuen, "JRA1 - The DEPFET sensor as the first fully integrated DUT in the EUDET pixel telescope: The PS test beam 2008," *Eudet-Memo-2008-34*, 2009.

[45] C. Sandow, L. Andricek, P. Fischer, R. Kohrs, H. Krueger, G. Lutz, H. G. Moser, L. Reuen, R. H. Richter, L. Strueder, J. Treis, M. Trimpl, and N. Wermes, “Clear performance of linear DEPFET devices,” *Nucl. Instr. Meth. A 568*, 2006.

[46] I. Peric, *Design and Realisation of Integrated Circuits for the Readout of Pixel Sensors in High-Energy Physics and Biomedical Imaging.* PhD thesis, Bonn University, 2004.

[47] I. Gregor and J. Mnich, “Eudet: Detector r&d towards the international linear collider,” *IEEE 2008, Nuclear Sience Symposium Dresden*, 2008.

[48] J. Treis, *Development and operation of a novel PC-based high speed beam telescope for particle tracking using double sided silicon microstrip detectors.* PhD thesis, Bonn University, 2002.

[49] M. Mathes, *Development nd Characterization of Diamond and 3D-Silicon Pixel Detectors with ATLAS-Pixel Readout Electronics.* PhD thesis, Bonn University, 2008.

[50] R. Turchetta, “Spatial resolution of silicon microstrip detectors,” *Nucl. Instr. and Meth. A*, 1993.

[51] S. Agostinelli *et al.*, “G4–a simulation toolkit,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003.

[52] R. Penrose, “A generalized inverse for matrices,” *Proc. Cambridge Phil. Soc.*, vol. 51, pp. 406–413, 1955.

[53] *TMVA 4 User Guide: Toolkit for Multivariate Data Analysis with ROOT.*

[54] J. Friedman, “Greedy function approximation: A gradient boosting machine,” *Annals of Statistics*, vol. 29, pp. 1189–1232, 2001.

[55] T. Carlia and B. Koblitz, “A multi-variate discrimination technique based on range-searching,” *Nucl. Instr. Meth. A*, vol. 501, pp. 576–588, 2003.

[56] A. G. Frodesen, O. Skjeggstad., and H. Tøfte, *probability and statistics in particle physics.* Universitetsforlaget Bergen - Oslo - Tromsø, 1978.

[57] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse.* Teubner, 1998.