


FACULTEIT WETENSCHAPPEN
VAKGROEP SUBATOMAIRE EN STRALINGSFYSICA
ACADEMIEJAAR 2006-2007

Exclusive π^0 production at HERMES

Detection – Simulation – Analysis

Exclusieve π^0 productie in HERMES

Detectie – Simulatie – Analyse

Arne Vandenbroucke
Promotor: Prof. Dr. Dirk Ryckbosch

Proefschrift ingediend tot het verkrijgen van de graad van
Doctor in de Wetenschappen: Natuurkunde

Most People are other People.
Their thoughts are someone else's opinions,
their lives a mimicry,
their passions a quotation.

0. Wilde

Contents

CHAPTER 1

Fundaments of Matter itself

It was about twenty years ago that I first asked my father what 'physics' actually was. He told me it was a science trying to explain various properties of nature, and gave me the example of why water was sticking to the surface of a cup when the water inside the cup was poured. A couple of years later it became clear that the word 'physics' itself was derived from the ancient Greek word 'φυσις', which means nature. Even later, I heard that the Greek word, 'φυσικη', means much more than just 'nature', it means 'the understanding of the concepts of nature'.

In the world of subatomic and particle physics, the fundamental blocks of matter are studied. Atoms are built out of a high density core surrounded by a cloud of electrons. The fundamental interaction responsible for keeping the positively charged nucleus and the negative electrons bound to each other is called the electromagnetic force. This force has the property that it repels particles with the same charge.

The atomic nucleus consists of positively charged protons and uncharged neutrons, altogether called nucleons. There has to exist a force which can overcome the electromagnetic repulsion between equally charged particles and which is responsible for binding these nucleons together. This force is called the 'Strong Force' and is studied in the field of nuclear and particle physics.

Nuclear physics studies the properties of atomic nuclei, whereas the particle physics community focuses on the building blocks of the neutron and proton itself, namely the quarks. According to the 'Standard Model', there exist 6 types of quarks. Each of the quarks has its corresponding antiquark. The quarks carry their own form of charge, called the 'color charge', which comes in three types: blue, green and red. The interaction between two quarks occurs via the exchange of a gluon, the gauge boson of the Strong Force. The gluon carries one unit of color and one unit of anti-color.

Particles built of a quark and an antiquark are called mesons, those built out of three (anti-)quarks are called (anti-)baryons. Examples of a baryons are the neutron and the proton. A common name for mesons and baryons is 'hadrons'. The production rate and properties of these hadrons are subject of study in particle physics.

Apart from their quark content, hadrons are ordered according to their spin. Spin is one of the most important properties of a particle. In the Standard Model of matter all fundamental particles have spin $\frac{1}{2}$. In contrast, particles which are the mediators of the fundamental forces have spin[1] 1. In other words, matter is described by spin $\frac{1}{2}$ particles, while forces are described by integer spin particles. Depending on its spin a particle is said to be a fermion (for half-integer spins) or a boson (for integer spins). Fermions and bosons behave fundamentally different. Nucleons have half-integer spin, thus they are fermions.

Naively one would expect that the spin of the proton is built out of the spin of the three constituent quarks. Since each quark has spin $\frac{1}{2}$, one could think that if two quark spins are aligned and one quark spin is anti-aligned the proton spin of $\frac{1}{2}$ is obtained. However, a measurement by the EMC collaboration in the 1980's showed that the constituent quarks are responsible for only about 30% of the nucleon spin. This observation lead to the so-called 'spin crisis'. At that point it became clear that the spin of the proton was not only built out of the spin of the three valence quarks. Quark orbital angular momenta and gluon spin are now believed to build the remaining $\sim 70\%$ of the spin.

The real challenge of a model describing the constituents of a nucleon is that it not only has to describe what the nucleon is made of, but that it needs to be able to predict the properties of the nucleon from the properties of its fundamental constituents. The comprehension of the nucleon spin in terms of the quarks and gluons is an indication of our understanding of the essential concepts of nature.

The HERMES experiment was built to solve the spin mystery and disentangle the various contributions to the nucleon spin. Via polarized electron-proton scattering information about the spin content of the nucleon can be obtained. A difference of counts between reactions where beam and target have the same polarization and those where beam and target have opposite polarization can be related to the so-called polarized quark distributions. Even though HERMES originally was set up to measure inclusive reactions, where only the scattered electron is detected, the spectrometer was 'enRICHed' after three years of running with a new component allowing to identify hadrons produced in the scattering reaction. The hadrons in these so-called semi-inclusive reactions contain information about the flavor of the quark involved in the scattering process. Data from HERMES have confirmed and detailed the observation from EMC about the nucleon spin not coming from the spin of the quarks alone.

In the hunt for other contributions to the nucleon spin, quark orbital angular momenta are expected to play an important role. Predictions for the quark orbital angular momenta are given in the framework of 'Generalized Parton Distribution' functions (GPDs). These GPDs are generalizations of the standard parton distribution functions and are accessible through hard exclusive reactions. In this type of reactions the final state is fully known.

Exclusive reactions have already been measured at HERMES. However, due to the limited acceptance of the spectrometer, the recoiling proton remained undetected. Exclusivity was achieved by imposing constraints on the missing mass. The latter is obtained from the kinematics of the scattered lepton and the produced meson.

In order to improve the measurement of exclusive reactions, the HERMES collaboration has decided in 2001 to design and construct a Recoil Detector surrounding the target cell so that the recoiling target can be directly measured. Together with the information from the forward spectrometer, all reaction products for the exclusive reactions can be identified and

[1] To be correct: this is not valid for the gravitational force; the graviton has spin 2

their kinematics determined. Hence, the background in the exclusive event sample will be strongly reduced compared to the present analysis.

In experiments trying to unravel the structure of nucleons, like the HERMES experiment, the nucleon is probed with a point-like particle like an electron or a muon. The reasons that high energy accelerators are needed for studies like these, is the resolution of the probe: when going to higher energies smaller distances can be accessed according to the particle-wave duality of matter.

A question often encountered by people not familiar with the field of particle physics is about the value of this type of research. There are different answers to this question. First of all, it needs to be stressed that understanding the fundaments of matter could potentially open the way to entirely new technologies. In this respect it is worth mentioning that when scientists at the end of the 19th century were experimenting with electromagnetic waves, objects like mobile phones, GPS or remote controls were completely unheard of. These objects, however, could only be developed with the knowledge acquired in understanding these electromagnetic waves. Nowadays not one physicist is able to predict any new technology based on the results of elementary particle research. However, all researchers will agree that not everything is known about the subject. Knowing that we do not know and investigating what we do not know opens doors to potential new technologies.

A second motivation for performing fundamental research is the spin-off of technologies developed in function of the experiments. One of the most striking examples of this is the development of medical imaging, a field which directly adopts technologies used in particle physics. The know-how in operating particle accelerators which are needed for the various particle physics experiments, has evolved to a much broader scientific field nowadays: solid state physicists and biomedical research, for instance, were boosted tremendously by the development of coherent monochromatic high intensity beams of synchrotron light. A last example illustrating this second point is the development of the 'World Wide Web', which started at CERN, a particle physics research center in Switzerland, from the effort to enhance the communication between various laboratories.

A last argument, finally, is related to the fact that mankind's world view itself is developing under the influence of knowledge. Our daily thinking is influenced by our interpretation of nature itself. We would not have the same conceptual thinking if we would not see the relation between an apple falling from a tree and the fundamental force of gravity. The development of modern art is hard to imagine if we saw ourselves in a geocentric frame.

After this introduction, the second chapter of this thesis introduces the theoretical framework of nucleon structure. The concept of structure functions and deep-inelastic scattering are presented as well as a discussion about the polarized and unpolarized structure functions $F_{1(2)}$ and $g_{1(2)}$. Subsequently, the formalism of Generalized Parton Distribution functions is introduced. The general concept of these functions, as well as some of their properties, are mentioned. A discussion on how we can experimentally access these functions ends this theory chapter.

Chapter 3 and 4 detail the experimental setup of the HERMES experiment. Chapter 3 focuses on the HERA electron beam, the polarized target and all detector components of the HERMES spectrometer, while Chapter 4 discusses how the information from various detector components is transformed into a format ready for physics analysis. Tracking and particle

identification needed to achieve this stage are explained.

The 'detection' part of this thesis is discussed in Chapter 5. That chapter is devoted to the design and construction of the silicon recoil detector for the HERMES experiment. A large fraction of the available PhD time was assigned to research and development studies for the readout of the new silicon detector. Most of the work related to this topic was performed in DESY Zeuthen, Germany.

Chapter 6 reports the 'simulation' section of the thesis. The performance of the HERMES electromagnetic calorimeter is extracted, based on an improved Monte Carlo description of the preshower and calorimeter.

The next chapter presents the 'analysis' branch of the thesis. The analysis of exclusive π^0 production on a hydrogen target is presented. Additionally, the π^0 and η production ratio on a hydrogen versus deuterium target as a function of $z(=\frac{E_h}{\nu})$ is investigated.

Chapter 8 finally gives a summary and concludes the work presented in this thesis.

CHAPTER 2

Structure of the Nucleon

Early scattering experiments in the 60's aimed at measuring the spatial extension of the nucleon. By elastically scattering electrons off nucleons information on the nucleon's shape could be probed. From these experiments it became apparent that the nucleon is not simply a point-like particle, but rather is an extended object.

Inelastic scattering experiments, started in the 70's, showed that the nucleon is built out of three quarks. These results yielded the so-called 'naive parton model'. The observation of scaling violation by second and third generation scattering experiments, however, was not expected by this model. The description of the building blocks of the nucleon required a more solid theoretical framework, which was found in QCD, the theory describing the strong interactions between quarks and gluons.

Experiments in the 80's and 90's have further explored the limits of QCD on the one hand, but on the other hand also have disproved various predictions from the early theory. The violation of the Gottfried rum rule and the Ellis-Jaffe sum rule are examples of the contrast between theoretical expectation and experimental measurements. The observation in the 80's by the EMC collaboration, that the spin of the nucleon is not simply generated by the quarks, dramatically showed our limited understanding of the constituents of nucleons.

When trying to determine the various aspects to the nucleon spin, orbital angular momenta and gluon spin are expected to play a non-negligible role. The power of being able to predict fundamental nucleon properties, such as spin, is very important in the understanding of the interactions between the most fundamental building blocks of matter.

A framework giving access to the total spin of quarks and gluons can be found in the 'Generalized Parton Distributions'. These combine information from elastic and inelastic scattering experiments and can also be accessed via exclusive reactions.

This section will give an overview of what has been learned from various scattering experiments, together with a conceptual overview of the theoretical framework explaining the observed spectra. An introduction into Generalized Parton Distributions as well as an overview of how the HERMES experiment could probe these functions will be given.

2.1 Nucleon form factors from elastic electron scattering

Since electrons are point-like particles, they form an ideal tool to probe the structure of nucleons. When electrons of energy E are scattered off a heavy object, the target recoil can be safely ignored. The cross section for this type of scattering can be written as:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = \underbrace{\left(\frac{2Z\alpha\hbar cE'}{|\vec{q}c|^2}\right)^2 \cdot \cos^2\frac{\theta}{2}}_{=\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mott}}} \cdot \left|F(\vec{q}^{\,2})\right|^2, \tag{2.1}$$

where Z is the charge of the target, α the fine structure constant, $\vec{q} = \vec{k} - \vec{k'}$, the three-momentum transfer to the target particle, $\vec{k}$ ($\vec{k'}$) being the three-momentum of the initial (scattered) electron, and θ the angle between initial and scattered electron. The function $F(\vec{q}^{\,2})$ is of particular importance here and describes the spatial extension of the target. Assuming spherical symmetry, it depends on $|\vec{q}|$ only. Obviously the cross section for scattering off compound objects should be smaller than when scattering off point-particles, for which the cross section is given by the Mott cross section as indicated in equation 2.1. The function $F(\vec{q}^{\,2})$ is referred to as a spatial form factor. For point-like particles we have $F(\vec{q}^{\,2}) \equiv 1$. Elastic electron scattering off the proton indicated that a dipole-like form factor described the data. Via a Fourier transformation from momentum space into position space, the charge distribution $f(r)$ can be calculated:

$$f(|\vec{x}|) = \frac{1}{(2\pi)^3}\int F(\vec{q}^{\,2})e^{-i\frac{\vec{q}\cdot\vec{x}}{\hbar}}d\vec{q}. \tag{2.2}$$

For dipole-like from factors the charge distribution follows an exponential form.

When the energy of the scattering particles is increased, the recoil of the target can no longer be neglected, and equation 2.1 needs to be modified. Instead of the three momentum transfer to the nucleon $\vec{q}$, the four-momentum transfer q has to be used. As a simplification, however, the electron mass can be safely ignored at higher energies. The resulting crosssection is described by the *Rosenbluth* formula:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = \left(\frac{2Z\alpha\hbar cE'}{|qc|^2}\right)^2 \cdot \cos^2\frac{\theta}{2} \cdot \frac{E'}{E} \cdot \left(\frac{G_E^2(Q^2) + \tau G_M^2(Q^2)}{1+\tau} + 2\tau G_M^2(Q^2)\tan^2\frac{\theta}{2}\right), \tag{2.3}$$

where $\tau = \frac{Q^2}{4M^2c^2}$, $Q^2 = -q^2$, M the mass of the target particle, Z the charge of the target particle, and G_E and G_M the electric and magnetic form factor of the nucleon respectively. In e.g. Reference [1] the above differential cross section is deduced relativistically. The form factors are defined for both neutron and proton. Given the well known charge of neutron and proton, the electric form factor has the following limiting values:

$$G_E^p(Q^2 = 0) = 1 \qquad G_E^n(Q^2 = 0) = 0.$$

The magnetic form factor is limited by the magnetic moments of proton μ_p and neutron μ_n, which can be expressed in terms of the nuclear magneton $\mu_N = \frac{e\hbar}{2M_p}$, as $\mu_p = 2.79\mu_N$ and

$\mu_n = -1.91\mu_N$, resulting in the following equations:

$$G_M^p(Q^2 = 0) = 2.79 \qquad G_M^n(Q^2 = 0) = -1.91,$$

Figure 2.1 shows these form factors as a function of Q^2. From the figure[1] it follows that both the neutron and proton magnetic form factors as well as the proton electric form factor fall off similarly with Q^2. The experimentally observed form factors can be described by a dipole fit [2]:

$$G^{\text{dipole}}(Q^2) = \left(1 + \frac{Q^2}{0.71\ \text{GeV}^2}\right)^{-2}. \tag{2.4}$$

As mentioned before, a dipole like form factor corresponds to a charge distribution which falls of exponentially:

$$\rho(r) = \rho(0)e^{-\alpha r} \qquad (\alpha = 4.27\ \text{fm}^{-1})$$

.

The neutron electric form factor is of particular importance. It is well known that the neutron itself is a neutral particle. However, it can be seen from Figure 2.1 that the neutron electric form factor is slightly negative, so it can be concluded that it is only appearing neutral from the outside, hence it must be built out of electrically charged constituents.

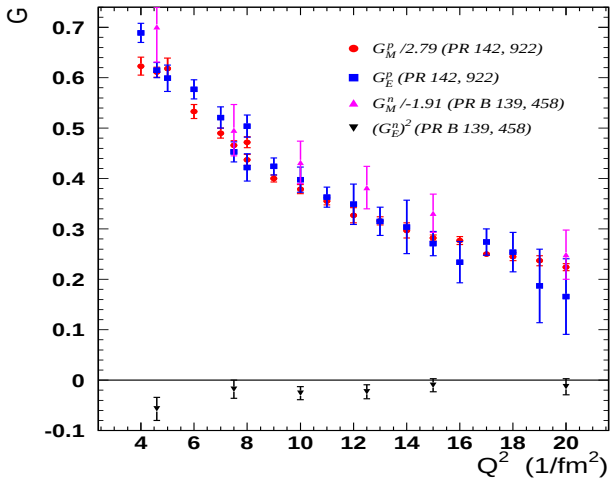

Figure 2.1: Proton and neutron electric and magnetic form factors as a function of Q^2. Neutron data are taken from Reference [3], proton data from Reference [4]. The values for the proton and neutron magnetic form factor are scaled by 2.79 and 1.91 respectively.

[1]The units of Q^2 in the figure are expressed in $1/\text{fm}^2$. To simplify notations, it is usual in high energy physics to use the conventions $c = 1$ and $\hbar = 1$. These conventions will be used throughout this thesis. Via the relationship $E = \frac{hc}{\lambda} (= h\nu)$ it can be seen that energies can be expressed as inversed distances.

2.2 Nucleon structure functions from Deep Inelastic Scattering

2.2.1 DIS kinematics

With increasing beam energies, the electron-proton scattering no longer occurs elastically, and an additional degree of freedom needs to be introduced in order to describe the process. In inelastic scattering, the energy E' of the scattered electron is kinematically not determined from the incident electron energy E and the scattering angle θ, unlike the case of elastic scattering. Instead, the produced particles in the reaction can take an arbitrary amount of the available energy. Therefore the resulting structure functions need an additional parameter. The cross section now is written as:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega \mathrm{d}E'} = \left(\frac{\alpha}{2E\sin^2\frac{\theta}{2}}\right)^2 \left[W_2(x,Q^2)\cos^2\frac{\theta}{2} + 2W_1(x,Q^2)\sin^2\frac{\theta}{2}\right], \tag{2.5}$$

with W_1 and W_2 two structure functions.

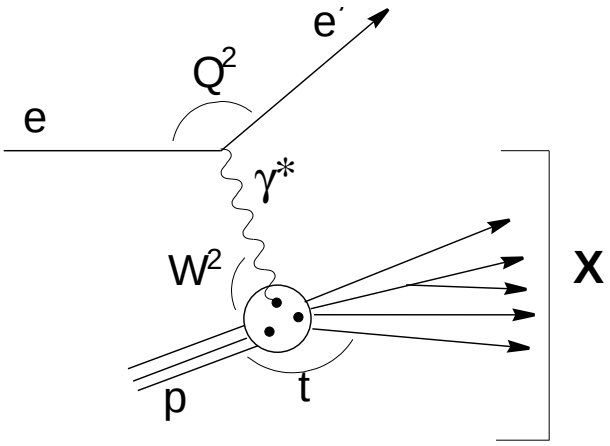

Figure 2.2: Schema of a DIS type reaction.

Figure 2.2 schematically shows the Deep Inelastic Scattering (DIS) reaction $e^- + p \rightarrow e^{-'} + X$. The electron exchanges a virtual photon γ^* with the target nucleon. If $k = (E, \vec{k})$ and $k' = (E', \vec{k'})$ are the four-momenta of initial and scattered electron respectively, Q^2 is determined by:

$$\begin{aligned} Q^2 &= -q^2 = -(k-k')^2 \\ &\overset{\text{lab}}{=} -\left(2m_e^2 - 2\left(EE' - \left|\vec{k}\right|\left|\vec{k'}\right|\cos(\theta)\right)\right) \\ &\approx -4EE'\sin^2(\frac{\theta}{2}) \qquad (m_e \approx 0), \end{aligned}$$

Higher Q^2 values correspond to higher virtual photon momenta. The higher a virtual photon's momentum, the shorter its wavelength. With increasing Q^2, smaller distances can accordingly be accessed.

The total energy available in the reaction is expressed by the center of mass energy s:

$$s = (k+P)^2 \overset{\text{lab}}{\approx} (2EM + M^2), \tag{2.6}$$

with P the four-momentum of the target nucleon and M its mass. Furthermore the variables ν, expressing the energy transfer from the lepton to the target, and W^2, indicating the invariant mass of all produced hadrons are introduced:

$$\nu = \frac{P \cdot q}{M} \overset{\text{lab}}{=} E - E' \tag{2.7}$$

$$W^2 = P'^2 = (P+q)^2 \overset{\text{lab}}{=} M^2 - Q^2 + 2\nu M, \tag{2.8}$$

where $P' = \sum_X p_X$ is the sum of the four-momenta of all produced particles. P is the four-momentum of the initial target with mass M, assumed to be at rest in the laboratory frame. By requiring W^2 to be sufficiently high, the region of nucleon resonances can be excluded. If $W^2 = M^2$ we are back into the regime of elastic scattering.

Finally, two scaling variables usually are introduced:

$$x = \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2M\nu} \overset{\text{lab}}{=} \frac{Q^2}{2\nu M_N} \tag{2.9}$$

$$y = \frac{p \cdot q}{P \cdot k} \overset{\text{lab}}{=} \frac{\nu}{E} \tag{2.10}$$

y indicates the fractional energy of the photon with respect to the initial energy. From equation 2.8 it follows that x will be equal to 1 for elastic reactions. For inelastic reactions we have $0 < x < 1$.

2.3 The DIS cross section

The cross section for deep inelastic scattering reactions can be written as:

$$\frac{d^2\sigma^{DIS}}{d\Omega dE'} = \left(\frac{\alpha\hbar}{Q^2 c}\right)^2 \frac{E'}{E} L^{\mu\nu} W_{\mu\nu}, \tag{2.11}$$

where $L^{\mu\nu}$ is the leptonic tensor, describing the emission of a virtual photon by an electron. The tensor is exactly calculable in QED and is given by:

$$L^{\mu\nu} = 2\left(k'^{\mu}k^{\nu} + k^{\mu}k'^{\nu} + (m_e^2 - k \cdot k')g^{\mu\nu} - 2i\epsilon^{\mu\nu\xi\tau} q_\xi s_\tau\right) \tag{2.12}$$

where $g^{\mu\nu}$ is the metric tensor, and $\epsilon^{\mu\nu\xi\tau}$ the totally anti-symmetric Levi-Civita tensor. The last part in expression 2.12 is anti-symmetric and describes the dependence on the initial lepton spin s_τ. The first part of the leptonic tensor is symmetric. It is easily seen that the spin-dependent part vanishes when summing over the spin states.

The absorption by the nucleon of a virtual photon is given in equation 2.11 by the hadronic tensor $W_{\mu\nu}$. This tensor is not exactly calculable, and therefore needs to be parametrized by

structure functions, which, as mentioned before, should be dependent on the two variables Q^2 and x (or ν)

$$\begin{aligned} W_{\mu\nu} &= \frac{F_1(Q^2,x)}{M}\left(-g_{\mu\nu}-\frac{q_\mu q_\nu}{Q^2}\right)+\left(P_\mu-\frac{P\cdot q}{q^2}q_\mu\right)\frac{F_2(Q^2,x)}{\nu M}\left(P_\nu-\frac{P\cdot q}{q^2}q_\nu\right) \\ &\quad +i\epsilon_{\mu\nu\xi\alpha}\frac{q^\xi}{\nu}\left(S^\alpha g_1(Q^2,x)+\left(S^\alpha-\frac{(S\cdot q)}{(P\cdot q)}P^\alpha\right)\frac{g_2(Q^2,x)}{\nu^2M^2}\right), \end{aligned} \tag{2.13}$$

where the spin independent part is described by the structure functions $F_1(Q^2,x)$ and $F_2(Q^2,x)$. The spin structure of the nucleon is also here described by the anti-symmetric part of the tensor, and parametrized via the structure functions $g_1(x,Q^2)$ and $g_2(x,Q^2)$. S^α is a pseudo-vector representing the spin four-vector of the nucleon. Since $W_{\mu\nu}$ is a tensor, S^α needs to couple to another pseudo-object, which is given by the pseudo-tensor $\epsilon_{\mu\nu\xi\alpha}$.

The structure functions F_1, F_2, g_1 and g_2 will be discussed in the next sections.

2.3.1 Unpolarized structure functions F_1 and F_2

When calculating the unpolarized cross section, one has to sum equation 2.11 over the initial lepton and hadron spins yielding the equation:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x\mathrm{d}Q^2}=\frac{4\pi\alpha^2}{Q^2}\left(F_1(x,Q^2)y^2\left(1-\frac{2m_e^2}{Q^2}\right)+\frac{F_2(x,Q^2)}{x}\left(1-y-\frac{Mxy}{2E}\right)\right) \tag{2.14}$$

The structure functions $F_1(x,Q^2)$ and $F_2(x,Q^2)$ have been measured in DIS experiments since the seventies. At the Stanford Linear Accelerator Center (SLAC) in San Francisco [5] it was observed that F_1 and F_2 depend only weakly on Q^2. According to the discussion of section 2.1 this means that electrons are scattered off point-like particles in the nucleon. Moreover, if the point-like particles have spin $\frac{1}{2}$, the structure functions F_1 and F_2 can be related to each other by comparing the general cross section expression for scattering of point like particles with spin $\frac{1}{2}$ with equation 2.14. Doing so, the following relation is obtained:

$$2xF_1(x)=F_2(x) \tag{2.15}$$

This relation is called the 'Callan-Gross' relation, and has been experimentally verified as can be seen from Figure 2.3. The relation gives dramatic evidence that the proton is built out of point-like particles with intrinsic spin $\frac{1}{2}$.

The absence of Q^2 dependence of the structure functions F_1 and F_2 was actually predicted before the experimental observation by Bjorken [6], in the regime where Q^2 is large (above about 1 GeV^2) and also $P\cdot q$ is large (above about 3.5 GeV), but their ratio, $2x$, is not. This absence of Q^2 dependence of $F_{1(2)}$ is referred to as 'Bjorken Scaling'.

In order to get an interpretation for the structure functions $F_1(x,Q^2)$ and $F_2(x,Q^2)$, the cross section for the scattering of an electron off a quark of flavor f can be considered:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}E'}=\left(\frac{\alpha}{2E\sin^2\frac{\theta}{2}}\right)^2\left[2\frac{e_f^2}{2m_f}\delta(x-1)\sin^2\frac{\theta}{2}+\frac{2m_f}{Q^2}\delta(x-1)\cos^2\frac{\theta}{2}\right], \tag{2.16}$$

where m_f is the quark mass. The δ-function is needed since we are considering elastic scattering off point-particles, hence one degree of freedom is lost. If the integration over E' is carried

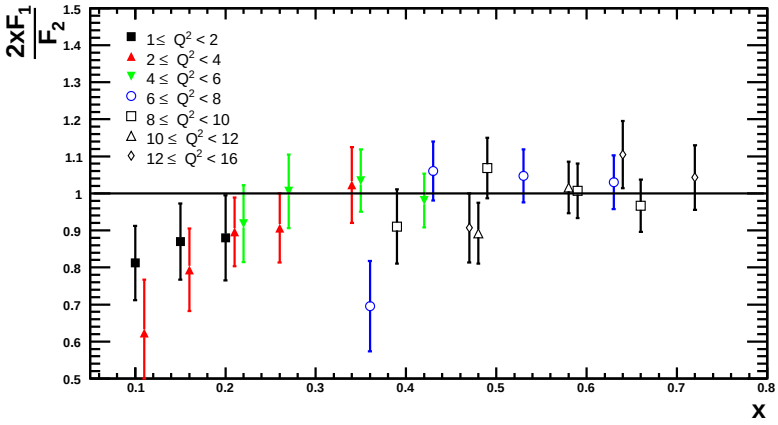

Figure 2.3: Experimental verification of the Callan-Gross relation. For visibility the data points have an offset in x. Data is averaged over the Q^2 bins. Data from Reference [5].

out, the Rosenbluth formula, equation 2.3, is obtained (under the assumption $G_E = G_M \equiv 1$). Comparing with equation 2.5, we can write:

$$\mathcal{W}_1^f = \frac{e_f^2}{2m_f}\delta(x_f - 1) \tag{2.17}$$

$$\mathcal{W}_2^f = \frac{2m_f e_f^2}{Q^2}\delta(x_f - 1) \tag{2.18}$$

where $x_f = -\frac{Q^2}{2q\cdot p_f}$, p_f the momentum of the quark, m_f its mass, and e_f its charge. The quark masses are related to the proton mass via $M = \Sigma_f m_f$. Individual quark momenta are not known, and therefore they need to be related to the proton momentum P: $p_f = z_f P$. This equation assumes that *each* component of p_f^μ gets the *same* fraction of the proton momentum P. Therefore it is only valid in the so-called 'infinite momentum frame', where the proton has an infinite longitudinal momentum and the quarks correspondingly also have only longitudinal momentum fractions. Accordingly we can write $x_f = \frac{x}{z_f}$, and $m_f = z_f M$.

Let $q_f(z_f)$ be the probability that a quark of flavor f carries momentum fraction z_f. Using equations 2.17 and 2.18 we can write:

$$\begin{aligned} W_1(x) &= \frac{1}{2}\sum_f \int_0^1 \frac{e_f^2}{2M}\delta(x - z_f)q_f(z_f)\mathrm{d}z_f = \frac{1}{2M}\sum_f e_f^2 q_f(x), \\ W_2(x) &= \sum_f \int_0^1 \frac{2M}{Q^2}e_f^2 z_f^2 \delta(x - z_f)q_f(z_f)\mathrm{d}z_f = \frac{2M}{Q^2}\sum_f e_f^2 x^2 q_f(x), \end{aligned} \tag{2.19}$$

where the property $\delta(\alpha x) = \frac{\delta(x)}{|\alpha|}$ has been used. The structure functions $W_{1(2)}$ are related to the dimensionless structure functions $F_{1(2)}$ via the relations:

$$\begin{aligned} F_1(x, Q^2) &= MW_1(x, Q^2) \\ F_2(x, Q^2) &= \frac{Q^2}{2Mx}W_2(x, Q^2). \end{aligned} \tag{2.20}$$

These calculations show that the structure functions $F_{1(2)}$ can be interpreted as describing the quark distributions within the nucleon.

The HERA facility made it experimentally possible to evaluate $F_2(x, Q^2)$ at high Q^2 and low x, as can be seen from Figure 2.4. From the figure it follows that at low x, F_2 strongly is dependent on Q^2, thus indicating that the quarks inside the nucleon are not really point-particles, but they are rather built out of other particles themselves. This behavior can be explained qualitatively: when going to lower momentum fractions x, more particles seem to share the momentum of the nucleon, while the resolving power of the electromagnetic probe increases at the same time with increasing Q^2, so all the particles sharing the momentum of the nucleon can be resolved. In other words, there are more particles at low x and they all can be seen with high Q^2. This implies larger parton distributions in the nucleon The Q^2 dependence of the structure function $F_{1(2)}$ is referred to as 'scaling violation', and forms experimental evidence of the existence of quark-antiquark pairs and gluons within the nucleon. These quark-antiquark pairs are referred to as the 'sea'-quarks. The gluon was experimentally confirmed from the existence of 3-jet events in e^+e^--scattering at PETRA.

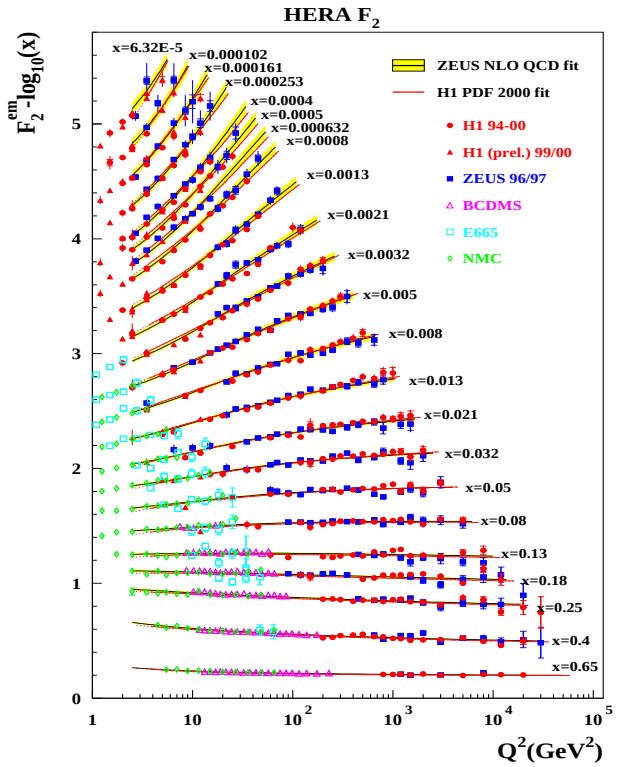

Figure 2.4: The unpolarized structure function F_2 as a function of Q^2 for different x bins. Figure taken from the ZEUS web page [7].

A qualitative picture of the valence and sea quark contributions in the nucleon can be seen in the right panel of figure 2.5. At higher x, the valence quark distribution $x \cdot v$ dominates, while at lower x the sea quarks contribute most to the nucleon momentum. The most probable value of the valence quark distribution is around 0.2, indicating that the nucleon momentum can not be made by summing the momenta of the 3 valence quarks alone. The left panel shows a calculation up to next to leading order (NLO, see later) at $Q^2 = 100$ GeV^2 for the

$u(x)$, $\bar{u}(x)$, $d(x)$, $\bar{d}(x)$ and $s(x)$-quark distributions as a function of x. The rise of all quark distributions at low x can be observed. At higher x it can be seen that there are about twice as many u as d quarks in the nucleon ($u(x) \approx 2d(x)$).

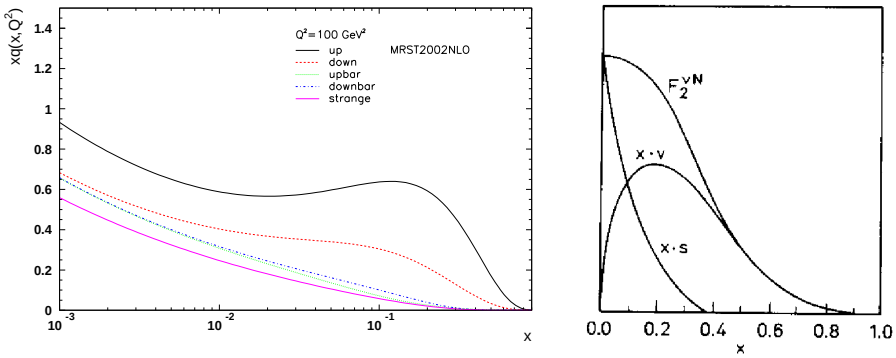

Figure 2.5: Left panel shows the u, $\bar{u}$, d, $\bar{d}$ and s-quark distributions as a function of x evaluated at $Q^2 = 100$ GeV^2 [8]. The right panel shows a qualitative picture of the sea $(x \cdot s)$ and valence $(x \cdot v)$ quark distributions as a function of x, the figure is taken from Reference [2].

From phenomenological principles, various sum rules can be postulated. One of these sum rules is the Gottfried sum rule, which gives a prediction for the integral of the difference of the proton and neutron F_2 structure functions:

$$S_G = \int_0^1 (F_2^p - F_2^n)\frac{\mathrm{d}x}{x}. \tag{2.21}$$

When we explicitly write equation 2.19 in terms of sea and valence u,s and d quark distributions, and taking equation 2.20 into account, we obtain:

$$\begin{aligned}
\frac{1}{x}F_2^p(x) &= \left(\frac{4}{9}\left(u_v^p(x) + u_s^p(x) + \bar{u}_s^p(x)\right) + \frac{1}{9}\left(d_v^p(x) + d_s^p(x) + \bar{d}_s^p(x)\right) + \frac{1}{9}\left(s_s^p(x) + \bar{s}_s^p(x)\right)\right) \\
\frac{1}{x}F_2^n(x) &= \left(\frac{4}{9}\left(u_v^n(x) + u_s^n(x) + \bar{u}_s^n(x)\right) + \frac{1}{9}\left(d_v^n(x) + d_s^n(x) + \bar{d}_s^n(x)\right) + \frac{1}{9}\left(s_s^n(x) + \bar{s}_s^n(x)\right)\right).
\end{aligned}$$

The subscripts v indicate valence quark distributions and the subscripts s the sea quark distributions for up u, down d and strange s quarks.

A very important principle in describing quarks is the isospin $SU(2)$ symmetry which states that the u-quark distribution in the neutron is the same as the d-quark distribution in the proton. Isospin symmetry demands that the neutron and proton can be transformed into each other by interchanging the u and the d quarks:

$$\begin{aligned}
u_v^n(x) &= d_v^p(x) \\
d_v^n(x) &= u_v^p(x) \\
u_s^p(x) = d_s^p(x) &= d_s^n(x) = u_s^n(x)
\end{aligned}$$

Sea quarks are always created as a quark-antiquark pair of the same flavor. Therefore we can write $u_s(x) = \bar{u}(x)$. As the proton is built out of two u and one d quark we can write:

$$\begin{aligned}\int_0^1 \left(u^p(x) - \bar{u}^p(x)\right) \mathrm{d}x = \int_0^1 u_v^p(x)\mathrm{d}x &= 2 \\ \int_0^1 \left(d^p(x) - \bar{d}^p(x)\right) \mathrm{d}x = \int_0^1 d_v^p(x)\mathrm{d}x &= 1 \\ \int_0^1 \left(s^p(x) - \bar{s}^p(x)\right) \mathrm{d}x &= 0\end{aligned}$$

Equation 2.21 now can be written as:

$$S_G = \frac{1}{3} + \frac{2}{3}\int_0^1 \left(\bar{u}(x) - \bar{d}(x)\right) \mathrm{d}x. \tag{2.22}$$

The NMC collaboration has measured the Gottfried sum rule explicitly: $S_G = 0.235 \pm 0.026$ [9]. The observation that $S_G \neq \frac{1}{2}$ forms experimental evidence of an asymmetric light quark sea. This could be explained by the Pauli principle: since the proton has more valence u quarks than d quarks it is easier to create a $d - \bar{d}$ pair than a $u - \bar{u}$ pair. This explanation, however, is not generally accepted.

There exists an alternative description of the unpolarized DIS cross section, equation 2.14. This description explicitly takes into account the photon-nucleon interaction. The DIS cross section accordingly is often represented as the absorption of a virtual photon by the proton. As the virtual photon is a massive particle it can be longitudinally or transversely polarized. The cross section therefore is written as:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \Gamma\left(\sigma_T^{\gamma^* p}(x, Q^2) + \epsilon\sigma_L^{\gamma^* p}(x, Q^2)\right), \tag{2.23}$$

where Γ represents a kinematic factor given by:

$$\Gamma = \frac{\alpha}{2\pi^2 Q^2}\frac{\nu^2(1-x)}{E^2 x}\frac{1}{1-\epsilon}, \tag{2.24}$$

and ϵ is the ratio of the longitudinal to transverse polarization of the virtual photon:

$$\epsilon = \frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1 - y + \frac{1}{4}y^2(\gamma^2 + 2)}, \tag{2.25}$$

with $\gamma = \sqrt{Q^2/\nu^2}$. The ratio R of the longitudinal and transverse absorption cross section is related to the structure functions F_1 and F_2:

$$R(x, Q^2) \equiv \frac{\sigma_L}{\sigma_T} = \frac{(1+\gamma^2)F_2(x, Q^2) - 2xF_1(x, Q^2)}{2xF_1(x, Q^2)}. \tag{2.26}$$

The structure function F_1 can now be related to the structure function F_2:

$$F_1(x, Q^2) = \frac{1+\gamma^2}{2x\left(1 + R(x, Q^2)\right)}F_2(x, Q^2) \tag{2.27}$$

In the Bjorken limit, $Q^2 \to \infty$, and $\nu \to \infty$, while keeping the ratio $\frac{Q^2}{2M\nu}$ finite, (i.e. $\gamma \to 0$), the longitudinal cross section σ_L vanishes, due to helicity conservation at the photon-parton vertex. It is easily seen that in this regime, the 'Callan-Gross' relation, equation 2.15, is obtained.

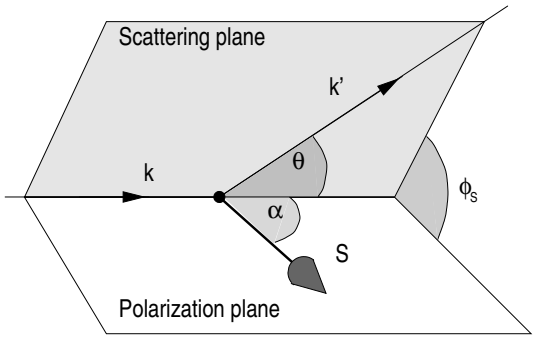

Figure 2.6: Kinematic variables in polarized lepton scattering from a polarized target.

2.3.2 Polarized structure functions g_1 and g_2

From formulas 2.12 and 2.13 it can be seen that no spin dependency will survive when the beam is unpolarized. Therefore, no generality is lost when defining the spin-dependent cross section as half the difference between right and left-handed incident electron cross sections [10]:

$$\begin{aligned}\frac{\mathrm{d}^3\left(\sigma(\alpha)-\sigma(\alpha+\pi)\right)}{\mathrm{d}x\mathrm{d}y\mathrm{d}\phi_S} &= \frac{e^4}{2\pi^2Q^2}\Bigg[\cos\alpha\left(\left(1-\frac{y}{2}-\frac{y^2\gamma^2}{4}\right)g_1(x,Q^2)-\frac{y}{2}\gamma^2 g_2(x,Q^2)\right)\\ &- \sin\alpha\cos\phi_S\gamma\sqrt{\left(1-y-\frac{y^2\gamma^2}{4}\right)}\left(\frac{y}{2}g_1(x,Q^2)+g_2(x,Q^2)\right)\Bigg],\end{aligned}$$

where the azimuthal angle ϕ_S and the angle α are angles between the target spin $\hat{S}$ and the scattered (incident) electron momentum $\vec{k}$ ($\vec{k}\,'$) respectively. These are indicated in Figure 2.6. Note that α is defined with respect to the incoming beam vector $\vec{k}$, and *not* with respect to the virtual photon direction $\vec{q}$. To access spin-dependent effects we can subtract cross sections either for different values of α or for right- and left-handed leptons, or we can measure the ϕ_S-dependence of the cross section.

The structure functions g_1 and g_2 can consequently be probed if both beam and target are polarized. Experimentally the spin structure functions $g_{1(2)}$ become available from measuring cross section differences:

$$\sigma_{LL} = \frac{1}{2}(\sigma^{\Leftarrow}_{\Rightarrow} - \sigma^{\Rightarrow}_{\Rightarrow}), \tag{2.28}$$

where longitudinally polarized leptons (L) scatter on a longitudinal polarized target (L), with polarization direction either parallel $\overset{\Rightarrow}{\rightarrow}$ or anti-parallel $\overset{\Leftarrow}{\rightarrow}$ to the spin direction of the beam. The cross section σ_{LL} can be written as:

$$\frac{\mathrm{d}^2\sigma_{LL}(x,Q^2)}{\mathrm{d}x\mathrm{d}Q^2} = \frac{8\pi\alpha^2 y}{Q^4}\cdot\left(\left(1-\frac{y}{2}-\frac{y^2}{4}\gamma^2\right)g_1(x,Q^2)-\frac{y}{2}\gamma^2 g_2(x,Q^2)\right), \tag{2.29}$$

where again $\gamma = \sqrt{Q^2/\nu^2}$. Experimentally, however, it is rather difficult to measure absolute cross sections, therefore one usually measures asymmetries or ratio's:

$$A_{\parallel} = \frac{\sigma_{LL}}{\sigma_{UU}}.$$

$\sigma_{UU} = \frac{1}{2}(\sigma^{\rightrightarrows} + \sigma^{\rightleftarrows})$ is the unpolarized cross section. Equation 2.29 can now be written as:

$$g_1(x,Q^2) = \frac{1}{1-\frac{y}{2}-\frac{y^2}{4}\gamma^2} \cdot \left(\frac{Q^4}{8\pi\alpha^2 y} \frac{\mathrm{d}^2\sigma_{UU}(x,Q^2)}{\mathrm{d}x\mathrm{d}Q^2} A_{\parallel}(x,Q^2) + \frac{y}{2}\gamma^2 g_2(x,Q^2) \right) \tag{2.30}$$

One of the primary goals of the HERMES experiment was to measure the structure function g_1. Experimentally one obtains $A_{\parallel}$ by counting the number of events when beam and target are parallel (anti-parallel) $N^{\rightrightarrows(\rightleftarrows)}$. These numbers need to be weighted by the dead time corrected luminosity in order to get a correct asymmetry:

$$A_{\parallel} = \frac{N^{\rightleftarrows}(x,Q^2)\mathcal{L}^{\rightrightarrows} - N^{\leftrightarrows}(x,Q^2)\mathcal{L}^{\rightleftarrows}}{N^{\rightleftarrows}(x,Q^2)\mathcal{L}_P^{\rightrightarrows} + N^{\leftrightarrows}(x,Q^2)\mathcal{L}_P^{\rightleftarrows}}, \tag{2.31}$$

where $\mathcal{L}_{(P)}^{\rightrightarrows(\rightleftarrows)}$ are given by:

$$\begin{aligned} \mathcal{L}^{\rightrightarrows(\vec{\leftleftarrows})} &= \int L^{\rightrightarrows(\rightleftarrows)}(t) \cdot \tau(t)\mathrm{d}t \\ \mathcal{L}_P^{\rightrightarrows} &= \int L^{\rightrightarrows}(t) \cdot \tau(t)|P_B(t)P_t(t)|\mathrm{d}t \end{aligned}$$

with $\tau(t)$ the detector dead time, and $P_B(t)$ and $P_t(t)$ the beam and target polarizations. From measurements at SLAC [11, 12] and from SMC [13] it followed that for longitudinal polarized targets the term proportional to g_2 in equation 2.30 has only a small effect and thus can be neglected. The polarized structure function g_1 can therefore be extracted from the observation of $A_{\parallel}$ and the knowledge of σ_{UU}. The latter is pretty well measured as seen from Figure 2.4. The world data on $g_1(x,Q^2)$ for the proton is plotted in Figure 2.7. A very good agreement between the various experiments can be seen even though these experiments mutually differ up to an order of magnitude in Q^2.

Due to the conservation of total angular momentum[2], a virtual photon with positive helicity can only couple to a quark with a spin orientation opposite to the photon spin, since the final state of the photon-quark interaction is a quark, and can not have a spin projection of $\frac{3}{2}$. This implies that when the photon helicity is parallel to the proton spin ($\overrightarrow{\rightarrow}$), the quark distributions with spin opposite to the proton spin $\Delta q_f^-(x,Q^2)$ are probed. Consequently, when photon and target spin are anti-parallel, the distribution $\Delta q_f^+(x,Q^2)$ is probed. This is indicated in Figure 2.8. The difference of these distributions is defined as the spin-dependent quark distributions:

$$\Delta q_f(x,Q^2) = q_f^+(x,Q^2) - q_f^-(x,Q^2) \tag{2.32}$$

In analogy to the unpolarized quark distributions $q_f(x,Q^2)$ these polarized quark distributions can in leading order be related to the spin-dependent structure function g_1:

$$g_1^p(x,Q^2) = \frac{1}{2}\sum_f e_q^2 \left(\Delta q(x,Q^2) + \Delta\bar{q}(x,Q^2)\right). \tag{2.33}$$

[2]i.e. $\hat{J}$, not to be confused with orbital angular momentum $\hat{L}$ ($\hat{J} = \hat{L} + \hat{S}$). In the example, however, orbital momenta are 0.

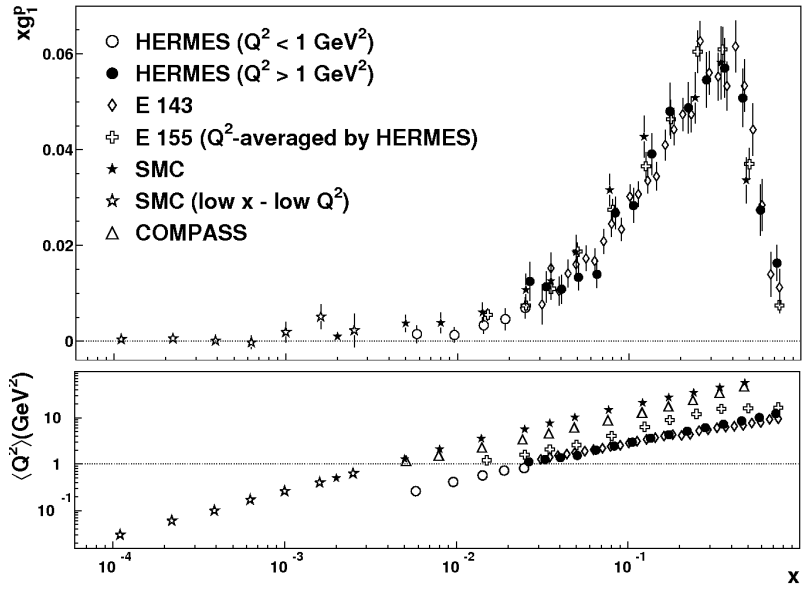

Figure 2.7: World data on $g_1^p(x, Q^2)$, together with the average (x, Q^2) for the different experiments. Figure taken from Reference [14].

Figure 2.8: Schematic of polarized DIS. The virtual photon helicity h_γ, quark helicity h_q and nucleon helicity h_N are indicated in the figure.

This equation can furthermore be decomposed in a flavor singlet S and a non-singlet NS component:

$$g_1^p(x, Q^2) = \frac{1}{2}\langle e^2 \rangle \left(\Delta q_f^S(x, Q^2) + \Delta \bar{q}_f{}^{NS}(x, Q^2) \right), \tag{2.34}$$

where $\langle e_q^2 \rangle = \frac{1}{N_f} \Sigma_q e_q^2$ is the average quark charge and

$$\begin{aligned} \Delta q^S(x, Q^2) &= \sum_f \left(\Delta q_f(x, Q^2) + \Delta \bar{q}_f(x, Q^2) \right) \\ &\equiv \Delta\Sigma(x, Q^2) \\ \Delta q^{NS}(x, Q^2) &= \frac{1}{\langle e^2 \rangle} \sum_f e_q^2 \left(\Delta q_f(x, Q^2) + \Delta \bar{q}_f(x, Q^2) - \Delta q_S(x, Q^2) \right) \end{aligned}$$

The singlet distribution $\Delta\Sigma(x, Q^2)$ is the sum of the polarized quark and anti quark distribu-

tions and is of particular importance since its first moment $\int \Delta\Sigma(x)\mathrm{d}x$ can be interpreted as the fraction of the nucleon spin originating from the quarks.

In the Bjorken scaling limit, the first moment of g_1 can be written as:

$$\Gamma_1 = \int_0^1 g_1(x)\mathrm{d}x = \sum_q e_q^2(\Delta q_f + \Delta\bar{q}_f), \tag{2.35}$$

where

$$\Delta q_f = \int_0^1 \Delta q_f(x)\mathrm{d}x,$$

the fraction of the nucleon spin originating from a quark of flavor f.

This moment can be decomposed into contributions from the axial charges a_0, a_3, and a_8, defined as the proton expectation values of the axial vector current:

$$A_j^\mu = \bar{\Psi}\gamma^\mu\gamma^5\left(\frac{\lambda_j}{2}\right)\Psi \qquad j = 0, \cdots, 8 \tag{2.36}$$

with λ_j the generators of the flavor $SU_f(3)$ in the standard Gell-Mann notation. We thus have:

$$\begin{aligned}\Gamma_1^p &= \frac{1}{2}\left(\frac{4}{9}(\Delta u(x) + \Delta\bar{u}(x)) + \frac{1}{9}(\Delta d(x) + \Delta\bar{d}(x)) + \frac{1}{9}(\Delta s(x) + \Delta\bar{s})\right) \\ &= \frac{1}{12}\left(\frac{4}{3}a_0 + a_3 + \frac{1}{3}a_8\right),\end{aligned}$$

where

$$\begin{aligned}a_0 &= \Delta u + \Delta\bar{u} + \Delta d + \Delta\bar{d} + \Delta s + \Delta\bar{s} \\ a_3 &= \Delta u + \Delta\bar{u} - \Delta d - \Delta\bar{d} \\ a_8 &= \Delta u + \Delta\bar{u} + \Delta d + \Delta\bar{d} - 2(\Delta s + \Delta\bar{s})\end{aligned} \tag{2.37}$$

The non-singlet axial charges a_3 and a_8 can be related to the two decay constants F and D, as obtained from the weak decay of hyperons [15]:

$$\begin{aligned}a_3 &= F + D = 1.2573 \pm 0.0028 \\ a_8 &= 3F - D = 0.579 \pm 0.025\end{aligned}$$

Measurements of the first moment of g_1^p provide the third necessary input to determine the spin carried by the quarks. The latter value can be estimated by placing $\Delta s + \Delta\bar{s} = 0$ in equations 2.37, implying $a_0 \equiv a_8$. Doing so, the important 'Ellis-Jaffe' sum rule is obtained:

$$\Gamma_1^p = \frac{1}{12}(F + D) + \frac{5}{36}(3F - D), \tag{2.38}$$

yielding a value of 0.189 ± 0.005, and consequently $a_0 = \Delta\Sigma \equiv 0.58$. The integral Γ_1^p was first measured by the EMC collaboration in 1987 to be $0.114 \pm 0.012 \pm 0.026$ [16]. This value is much smaller than the expected value. The observation by the EMC collaboration is commonly known as the 'spin crisis'. Recent measurements in DIS have confirmed the

violation of the Ellis-Jaffe sum rule and implicated that the spin of the nucleon is not simply built from the spin of the quarks. Instead, also the gluon polarization, $\Delta G(x)$, and orbital angular momenta, L^q need to be taken into account. Today it is common practice to write the nucleon spin as:

$$\frac{1}{2} = \frac{1}{2}\Delta\Sigma + L^q + J^G, \tag{2.39}$$

where J^G is the total angular momentum of the gluons.

2.4 A framework with quarks and gluons: QCD

From the previous sections it is clear that the naive parton model fails to explain the observed experimental data. Proton momentum and spin are not simply formed by the sum of the three valence quarks. The parton model, however, forms a nice approximation of the inner structure of a nucleon. A more elaborated theory of describing the interaction between quarks and gluons is QuantumChromoDynamics (QCD), derived from the old Greek word χρομος, which means color. Indeed, QCD forms the theoretical framework describing the quark-quark interactions in terms of the $SU_c(3)$ color symmetry. It has been experimentally verified that quarks come in three different colors. Moreover, the gauge bosons of QCD - the gluons - carry color charge themselves. Therefore, the gluons can couple to each other to form 'glueballs'. The interaction associated with color carriers is called the strong interaction. The fundamental interaction diagrams of this interaction are drawn in Figure 2.9.

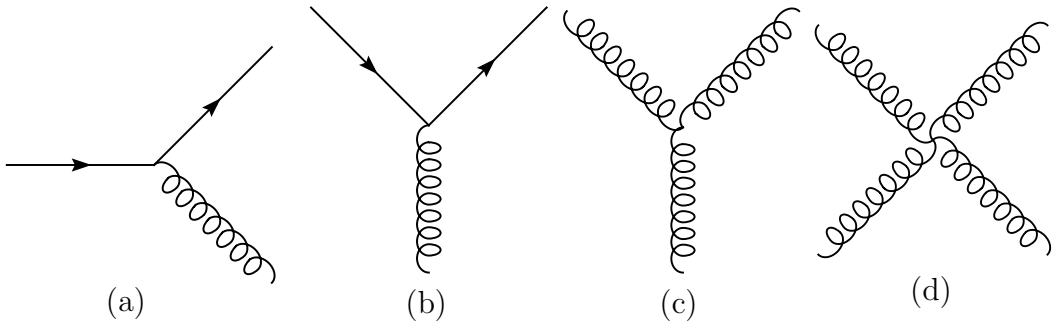

Figure 2.9: Fundamental vertices in QCD: Gluon emission by a quark (a), quark-antiquark pair creation by a gluon (b) and gluon self-coupling (c-d).

An important property of QCD is asymptotic freedom, which expresses the fact that at short distances the strong interaction coupling constant is small, and the particles move around essentially free. On the long range, however, the strong coupling constant α_s becomes large. The 'running' of the coupling constant is the reason of the success of the Bjorken parton model of freely moving quarks inside a nucleon. The asymptotic freedom effect in QCD can be explained by a screening effect, analogous as in the case of QED[3]. However, while in QED

[3]More generally, in any renormalizable theory, quantum corrections cause the coupling constant to be a function of the energy scale of the probe

the screening of the electromagnetic charges reduces the effective charge of an electron, gluons seem to have the opposite effect and are apreading the color charge of quarks.

Up to leading order (LO) in perturbation theory, the strong coupling constant α_s can be written as:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_0^2)}{1 + \left(\frac{\alpha_s(\mu_0^2)}{12\pi}\right)(11n - 2f)\ln\frac{Q^2}{\mu_0^2}} \tag{2.40}$$

n being the number of colors and f the number of flavors, μ_0^2 the scale at which we know α_s good enough to apply the evolution in Q^2. The scale μ_0 is introduced in the renormalization of the coupling. It is common practice to take $\mu_0 = M_Z$, the mass of the Z-boson. When describing the strong coupling constant, the variable Λ^2_{QCD} usually is introduced:

$$\ln\Lambda^2_{\mathrm{QCD}} = \ln\mu_0^2 - \frac{12\pi}{(11n - 2f)\alpha_s(\mu_0^2)},$$

so we can write α_s as a function of the parameter Λ_{QCD}:

$$\alpha_s(Q^2) = \frac{12\pi}{(11n - 2f)\ln\frac{Q^2}{\Lambda^2_{\mathrm{QCD}}}} \tag{2.41}$$

From the above equation it follows clearly that the strong force will decrease with increasing Q^2. This behavior has been experimentally observed as depicted in Figure 2.10. If Q^2 becomes of the order of Λ^2_{QCD}, the coupling becomes large and the partons involved will become strongly bound, and eventually form hadrons. In this regime perturbative QCD can no longer be applied. The value of Λ_{QCD} is not predicted by theory but has been measured to be $210 \pm 28 \pm 42$ MeV [17].

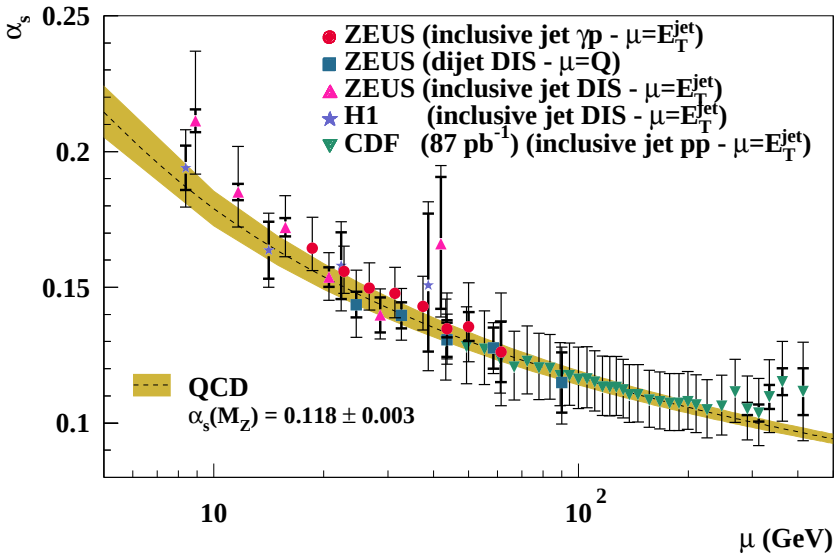

Figure 2.10: Running of the strong coupling constant as a function of μ. Figure taken from the ZEUS web page [7].

As the QCD coupling constant is Q^2 dependent, the parton distribution functions are Q^2 dependent as well. The leading order, i.e. Q^0, evolution of the structure functions $\Delta q^{NS}, \Delta\Sigma$ and ΔG can be described by the famous Dokshitzer-Gribov-Lipatov-Altarelli-Parisi equations (DGLAP, [18, 19, 20]):

$$\begin{aligned}
\frac{\partial}{\partial \ln Q^2}\Delta q^{NS}(x,Q^2) &= \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \mathcal{P}_{qq}(\frac{x}{\xi})\Delta q^{NS}(\xi)\frac{\mathrm{d}\xi}{\xi} \\
\frac{\partial}{\partial \ln Q^2}\Delta\Sigma(x,Q^20) &= \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \left(\mathcal{P}_{qq}(\frac{x}{\xi})\Delta\Sigma^{NS}(\xi) + 2n_f\mathcal{P}_{qg}\frac{x}{\xi}\Delta G(\xi)\right)\frac{\mathrm{d}\xi}{\xi} \\
\frac{\partial}{\partial \ln Q^2}\Delta G(x,Q^0) &= \frac{\alpha_q(Q^2)}{2\pi}\int_x^1 \left(\mathcal{P}_{gq}(\frac{x}{\xi})\Delta\Sigma^{NS}(\xi) + \mathcal{P}_{gg}(\frac{x}{\xi})\Delta G(\xi)\right)\frac{\mathrm{d}\xi}{\xi}
\end{aligned}$$

where n_f is the number of active quark flavors, and the functions $\mathcal{P}_{ij}(\frac{x}{\xi})$ represent splitting functions describing the probability that a parton i with momentum fraction x is radiated from a parton j with momentum fraction ξ. These functions are represented in leading order in Figure 2.11. x The Feynman graphs for the DIS interaction are shown for 0^{th} $(\mathcal{O}(\alpha))$ and

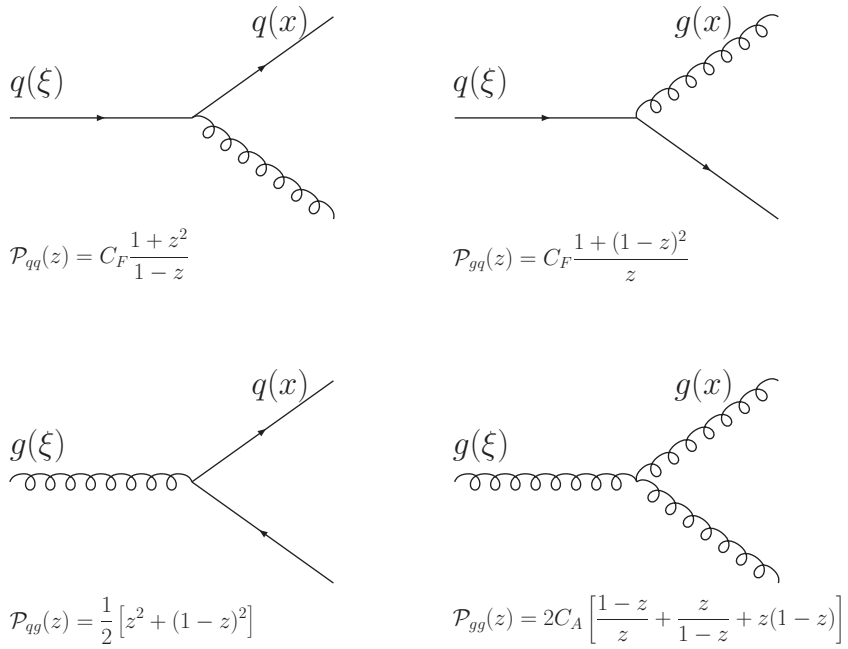

Figure 2.11: Leading order of the splitting functions in the DGLAP equations. $z = \frac{x}{\xi}$, C_A and C_F group theoretical constants, $C_A = n$, the number of colors, and $C_F = \frac{n^2-3}{2n}$. Equations taken from Reference [21].

1^{st} $(\mathcal{O}(\alpha\alpha_s))$ order in the strong coupling constant α_s in Figure 2.12. The process in panel (b) is called Photon-Gluon Fusion (PGF) and the one in panel (c) QCD-Compton Scattering (QCDC).

When moving to the next to leading order (NLO), the Feynman diagrams with two strong interaction vertexes need to be taken into account. The structure function g_1 then can be

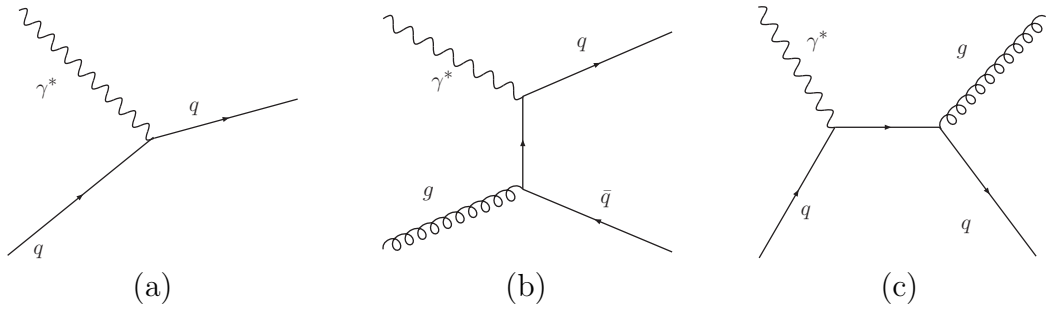

Figure 2.12: (a) represents the 0^{th} order Feynman graph in α_s for the reaction $\gamma^* + q \rightarrow q$. ($b$) and ($c$) represent the leading order diagrams: (b) is the Photon-Gluon Fusion process $\gamma^* + g \rightarrow q + \bar{q}$ and (c) represents gluon radiation $\gamma^* + q \rightarrow g + q$, the latter process is reffered to as QCD-Compton scattering.

written as [22]:

$$g_1(x.Q^2) = \frac{3}{2}\sum_{q,\bar{q}} e_q^2 \left[\Delta q + \frac{\alpha_s(Q^2)}{2\pi} C_a \otimes \Delta q(x.Q^2) + \frac{\alpha_s(Q^2)}{2\pi} 2C_G \otimes \Delta G(x, S^2)\right], \quad (2.42)$$

where C_q and C_G are coefficient functions. The product $\otimes$ denotes a convolution integral. One can see that the splitting functions $\mathcal{P}^{(1)}$ do not enter directly in the expression of g_1 at NLO. They are already taken into account implicitly in the expressions for ΔG and Δq.

Because of the dependence of g_1 on the polarized quark distributions Δq, it would be very interesting to invert the structure function and thus get access to Δq and ΔG. In References [23, 24, 25] a QCD analysis of the world data on polarized DIS is reported. In these analyses a parametrization of the polarized structure functions is extracted.

2.4.1 DIS in QCD

The hadronic tensor $W^{\mu\nu}$ in the DIS cross section, equation 2.11, can be expressed in a more general quantum field approach: it is the Fourier transform of the nucleon matrix element of the commutator of the electromagnetic current $J(q)$ [26]:

$$W^{\mu\nu} = \frac{1}{4\pi}\sum_X \langle PS|J^\mu|X\rangle\langle X|J^\nu|PS\rangle(2\pi)^4\delta(P+q-p_X) \quad (2.43)$$

$$= \int e^{iq\cdot\xi}\langle PS|[J^\mu(\xi), J^\nu(0)]|PS\rangle \mathrm{d}^4\xi, \quad (2.44)$$

where $p_X = \sum p_i$ is the sum over all hadronic final state momenta p_i, and $|PS\rangle$ a nucleon state with momentum P and spin S. The so-called *optical theorem* relates the hadronic tensor to the imaginary part of the forward virtual Compton scattering amplitude $\mathcal{T}$. This can be done for inclusive scattering, because of the summation over all hadronic final states in

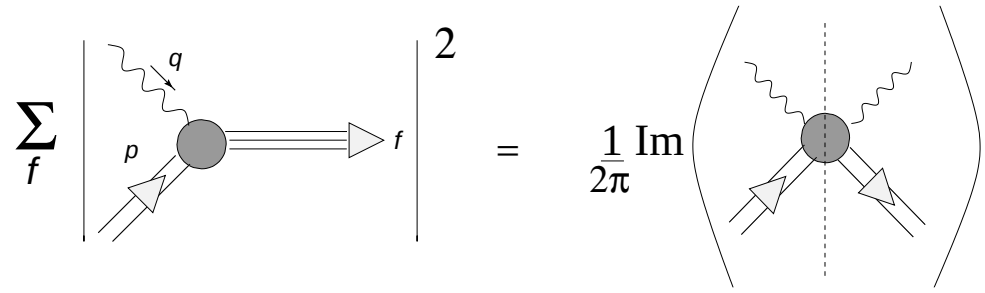

Figure 2.13: The optical theorem relating the hadronic tensor $W^{\mu\nu}$ to the virtual Compton scattering amplitude $\mathcal{T}$.

equation 2.43. The optical theorem is indicated in Figure 2.13. We can write:

$$\begin{aligned} 2\pi W^{\mu\nu} &= \mathcal{I}m\,(\mathcal{T}^{\mu\nu}) \\ \mathcal{T}^{\mu\nu} &= i\int e^{iq\cdot\xi} < PS|T(J^{\mu}(\xi)J_{\nu}(0))|PS > \mathrm{d}^4\xi, \end{aligned} \tag{2.45}$$

with $T()$ the time ordered product. The hadronic tensor can be seen as the product of a local, $J(0)$, and a non-local operator, $J(\xi)$.

The formalism to expand $\mathcal{T}$ as a product of local operators $\mathcal{O}$ is known as the *operator product expansion* (OPE). The OPE is a field theoretical expansion stating that the product of two non-local operators can be written as a sum of local operators:

$$\mathcal{O}_a(x)\mathcal{O}_b(0) = \lim_{x\to 0}\mathcal{O}_a(x)\mathcal{O}_b(0) = \sum_k c_k(x)\mathcal{O}_k, \tag{2.46}$$

the product of two operators is replaced by the sum, and the dependence on the position x is now in the coefficient functions $c_k(x)$. Applying the OPE to the commutator of the current operator J we get:

$$[J(\xi), J(0)] \sim \sum_k c_k(\xi^2)\xi^{\mu_1}\cdots\xi^{\mu_n}\theta_{\mu_1\cdots\mu_n}(0), \tag{2.47}$$

where $\theta_{\mu_1\cdots\mu_n}(0)$ are n local operators, depending on the spin n of the operator θ. The spin of an operator is related to the transformation properties under the Lorentz group. The operator θ could be of any canonical dimension d (in units of E). Since the tensor $T^{\mu\nu}$ has dimension 0, and the $|P\rangle$ state has dimension 1, the overall dimension of the entire matrix element $\langle PS|\theta|PS\rangle$ must be $d-2$. Every operator θ_μ acting on the $|P\rangle$ state introduces a dimension of $\frac{1}{M}$. M is used to indicate a mass scale which is related to Λ_{QCD}. The n operators of equation 2.47 sandwiched between the $\langle P|$ and $|P\rangle$ states accordingly imply the dimensionality of the matrix element to be:

$$\langle P|\theta_{\mu_1}\cdots\theta_{\mu_n}(0)|P\rangle \sim M^{d-2-n}. \tag{2.48}$$

The general form of the matrix element can be written as:

$$\langle P|\theta_{\mu_1}\cdots\theta_{\mu_n}(0)|P\rangle \sim P_{\mu_1}\cdots P_{\mu_n}M^{d-2-n}f_\theta, \tag{2.49}$$

with f_θ a scalar function. Since P is of the order Q, the dimension of the matrix element goes like $\langle P|\theta_{\mu_1}\cdots\theta_{\mu_n}(0)|P\rangle \sim Q^n M^{d-2-n}$. This implies that the dimension of the coefficient functions $c_k(\xi^2)$ need to be Q^{2-d}. The overall behavior of the tensor $\mathcal{T}^{\mu\nu}$ can be written as:

$$\mathcal{T}^{\mu\nu} \sim Q^{2-d} Q^n M^{d-2-n} = \left(\frac{Q}{M}\right)^{n+2-d} \sim Q^{2-t}, \tag{2.50}$$

with t the *twist* of the operator θ. Accordingly, twist is linked to the power with which M occurs in equation 2.47, it is the difference between the dimension d and the spin n of the operator: $t = d - n$. The dominating operators in the OPE are those with lowest twist. The concept of twist is commonly used in a less formal way as the order in $\frac{1}{Q^2}$ at which a particular effect is observed experimentally.

Up to twist three the objects of analysis in DIS are forward matrix elements of bi-local products of quark fields ψ:

$$\Gamma(x) = \frac{1}{2\pi}\int e^{i\lambda x}\langle PS|\bar{\psi}(0)\Gamma\psi(\lambda n)|PS\rangle \mathrm{d}\lambda, \tag{2.51}$$

with n a unit vector along the light cone (see appendix B):

$$n^\mu = \frac{1}{\sqrt{2}p}(1,0,0,-1) \tag{2.52}$$

When neglecting transverse momenta, only vector (γ^μ), axial-vector ($\gamma^\mu\gamma_5$) and tensor components ($\sigma^{\mu\nu}$) of the general expression for a 4×4 matrix (see Appendix A.2) will be non-zero. These components can be related to three structure functions [26]:

$$F_1(x) = \frac{1}{4\pi}\int e^{i\lambda x}\langle P|\psi(0)\not{n}\psi(\lambda n)|P\rangle \mathrm{d}\lambda \tag{2.53}$$

$$g_1(x) = \frac{1}{4\pi}\int e^{i\lambda x}\langle PS_\parallel|\psi(0)\not{n}\gamma_5\psi(\lambda n)|PS_\parallel\rangle \mathrm{d}\lambda \tag{2.54}$$

$$h_1(x) = \frac{1}{4\pi}\int e^{i\lambda x}\langle PS_\perp|\psi(0)[\not{S}_\perp,\not{n}]\psi(\lambda n)|PS_\perp\rangle \mathrm{d}\lambda \tag{2.55}$$

where $\not{n} = n_\mu\gamma^\mu$, and $S_{\parallel(\perp)}$ is the parallel (transverse) spin state. The functions F_1 and g_1 are already known, whereas the third function $h_1(x)$ can be written as the first moment of the transversity distribution δq:

$$h_1 = \int_0^1 \left(\delta q(x,Q^2) - \delta\bar{q}(x,Q^2)\right) \tag{2.56}$$

δq can be interpreted as the probability to find a quark with its spin aligned along the transverse spin of the nucleon minus the probability to find it reversely aligned:

$$\delta q(x) = q^{\Uparrow\uparrow}(x) - q^{\Uparrow\downarrow}(x). \tag{2.57}$$

It is important to note that the transverse spin distribution is not simply a rotation of the longitudinal distribution, as Lorenz boosts and rotations do not commute. An additional important feature of the transversity structure function is its odd chirality. Therefore it requires

a helicity flip of the interacting quark, which is forbidden in inclusive DIS, as indicated in figure 2.14. Transversity thus can not be accessed through inclusive DIS. Transversity distributions, however, can be accessed through semi-inclusive DIS which will be discussed in the next section. The first moment of the transversity function $h_1(x)$ can be related to the tensor charge, which also plays a role in the spin structure of the nucleon. More details about transversity at HERMES can be found in References [27, 28].

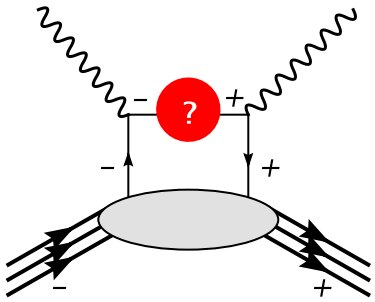

Figure 2.14: Handbag diagram indicating the required quark spin flip for probing the chiral odd transversity structure function h_1.

2.5 Polarized parton distribution functions from Semi-Inclusive DIS

Semi-inclusive deep inelastic scattering (SIDIS) is the DIS reaction where one or more produced hadrons are detected, apart from the scattered lepton. From the physical point of view this means that the cross section is not integrated over all hadronic final states, a particular final state rather is selected. The interest in SIDIS is to tag the struck quark from the produced hadron. This process is referred to as 'flavor tagging'. Schematically one can write:

$$e^- + N \rightarrow e^- + h + X, \tag{2.58}$$

where h denotes the produced hadron, and X represents all other reaction products. The process is schematically shown in Figure 2.15. The cross section for the process can be written as:

$$\sigma^h(x, Q^2, z) \propto \sum_f e_q^2 q_f(x, Q^2) D_f^h(z, Q^2), \tag{2.59}$$

where a new variable z given by

$$z = \frac{p \cdot p_h}{p \cdot q} \overset{\text{lab}}{=} \frac{E_h}{\nu} \tag{2.60}$$

is introduced. z expresses the fraction of the available energy carried by the produced hadron h. $D_f^h(z, Q^2)$ is a fragmentation function, which describes the fragmentation of a quark of type f into a quark of type h. Fragmentation functions describe the 'soft' processes involved in hadron formation. A commonly used model of fragmentation is the LUND model, based on string fragmentation [29]. The color field between a $q\bar{q}$ pair produced during the hard

interaction is represented by a massless relativistic string in this model. The string tension expresses the linear potential between the quark and the antiquark. Typical values for the string tension are about 1 GeV per fm. At sufficiently high energies, new $q\bar{q}$ pairs can be produced along the force field, corresponding to the breakup of the string. More information on fragmentation at HERMES can be found e.g. in Reference [30].

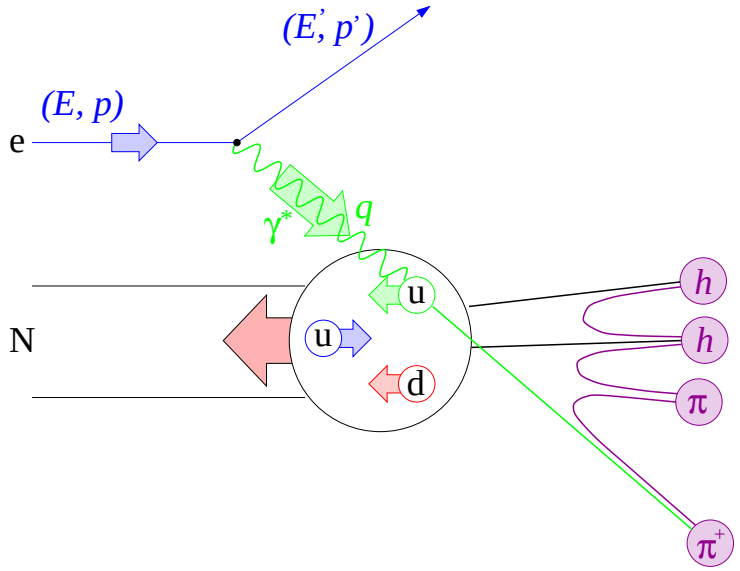

Figure 2.15: Schematic diagram of semi inclusive scattering.

A difficulty which immediately rises in SIDIS is how to make sure that the produced hadron is originating from the struck quark. To make a distinction, the 'current' and 'target' regions are introduced, where the former indicates that the hadron is coming from the interaction by the virtual photon, and the latter indicates that the hadron h is coming from target remnants.

Just like in DIS, in SIDIS a double spin asymmetry A_1^h can be defined. At LO QCD and assuming the structure function g_2 to be small, the asymmetry can be related to the quark distribution functions q and Δq:

$$A_1^h(x,Q^2) = C_R(x,Q^2)\frac{\Sigma_q e_q^2 \Delta q(x,Q^2)\int D_q^h(z,Q^2)\mathrm{d}z}{\Sigma_{q'} e_{q'}^2 q'(x,Q^2)\int D_{q'}^h(z,Q^2)\mathrm{d}z}, \tag{2.61}$$

where the integral over z is performed over the range where the produced hadron is coming from the current region. $C_R = (1+R(x,Q^2))/(1+\gamma^2)$ is a correction factor. From the observation of A_1^h one would want to extract the quark polarizations $\Delta q(x,Q^2)/q(x,Q^2)$. To perform the inversion, the concept of purities P_q^h is introduced:

$$P_q^h(x_i) = \frac{e_q^2 q(x_i,Q^2)\mathcal{D}_q^h(x_i,z_i,Q_i^2)}{\Sigma_{q'} e_{q'}^2 q'(x_i,Q_i^2)\mathcal{D}_{q'}^h(x_i,z_i,Q_i^2)} \tag{2.62}$$

$\mathcal{D}_q^h(x_i,z_i,Q_i^2)$ are experimentally measured fragmentation functions in a kinematic bin i, with average x_i, z_i and Q_i^2. These purities can be interpreted as the probability that, in a given bin

x_i, the photon struck a quark with flavor q when a hadron h is *detected* in the experiment. Equation 2.61 now can be written as:

$$A_1^h(x_i) = C_i \sum_q P_q^h(x_i) \frac{\Delta q(x_i, Q_i^2)}{q(x_i, Q_i^2)}. \tag{2.63}$$

C_i is a bin dependent correction factor. When considering all observed hadron types h, a system of linear equations is obtained:

$$\vec{A}(x_i) = C_i \cdot \mathbf{P}\vec{Q}(x_i), \tag{2.64}$$

with $\vec{A}$ the measured asymmetries, $\vec{Q}$ a vector containing the quark polarizations and $\mathbf{P}$ the purity matrix with elements P_q^h. The unknown vector $\vec{Q}$ can be obtained from minimizing:

$$\chi^2 = (\vec{A} - \mathbf{P} \cdot \vec{Q})^T \nu_a^{-1} (\vec{A} - \mathbf{P} \cdot \vec{Q}), \tag{2.65}$$

where ν_A^{-1} is the covariance matrix of the asymmetry vector $\vec{A}$. From the observed asymmetries for $\pi^{\pm}$ and $K^{\pm}$, the polarized quark distributions for the three lightest quarks have been extracted for the first time by the HERMES collaboration [31]. The distributions are shown in Figure 2.16. The assumption $\frac{\Delta s}{s} = \frac{\Delta \bar{s}}{\bar{s}}$ was made. One can see a positive (negative) polarization of the up-quark (down-quark), the up quark polarization is about twice as big as the one for the down quark ($\Delta u \approx -2\Delta d$). Moreover, all sea quark polarizations are consistent with zero. The observed polarizations are consistent with LO calculations from DIS data. From the evaluation of the flavor separated Δq functions, $\Delta\Sigma$ was measured to be $0.347 \pm 0.024 \pm 0.066$. This value is different from the one reported by EMC, $\Delta\Sigma = 0.114 \pm 0.012 \pm 0.026$ [16], due to the different Q^2 of the experiments. Like all QCD functions, also $\Delta\Sigma$ is subject to Q^2 evolution.

2.6 Generalized parton distributions via exclusive DIS

Guided by the observation that only a fraction of the nucleon spin is derived from the quark distributions, orbital angular momenta are expected to play an important role in the contribution to the nucleon spin. Obviously, orbital angular momenta can not be accessed from the 'standard' longitudinal spin structure functions $g_1^q(x, Q^2)$, as a two dimensional frame is needed to describe angular phenomena.

A more general description of the partons inside the nucleon is given by the generalized parton distribution functions, abbreviated to GPDs. In Reference [32] a sum rule for these GPDs which relates them to orbital angular momenta is proved. This sum rule will be accessed later. The framework of GPDs can be explained in a clean way by considering the Deeply Virtual Compton Scattering (DVCS) process, where a virtual photon is scattered off a quark inside a nucleon, thereby producing a real photon, while leaving the nucleon intact. The process is indicated in Figure 2.17. By definition, the final state of all particles in DVCS is known: a scattered lepton, a produced photon and a recoiling proton. Evidently, the initial state is known as well. Processes where the final and initial state are fully known are called exclusive processes. Elastic scattering is an example of an exclusive process.

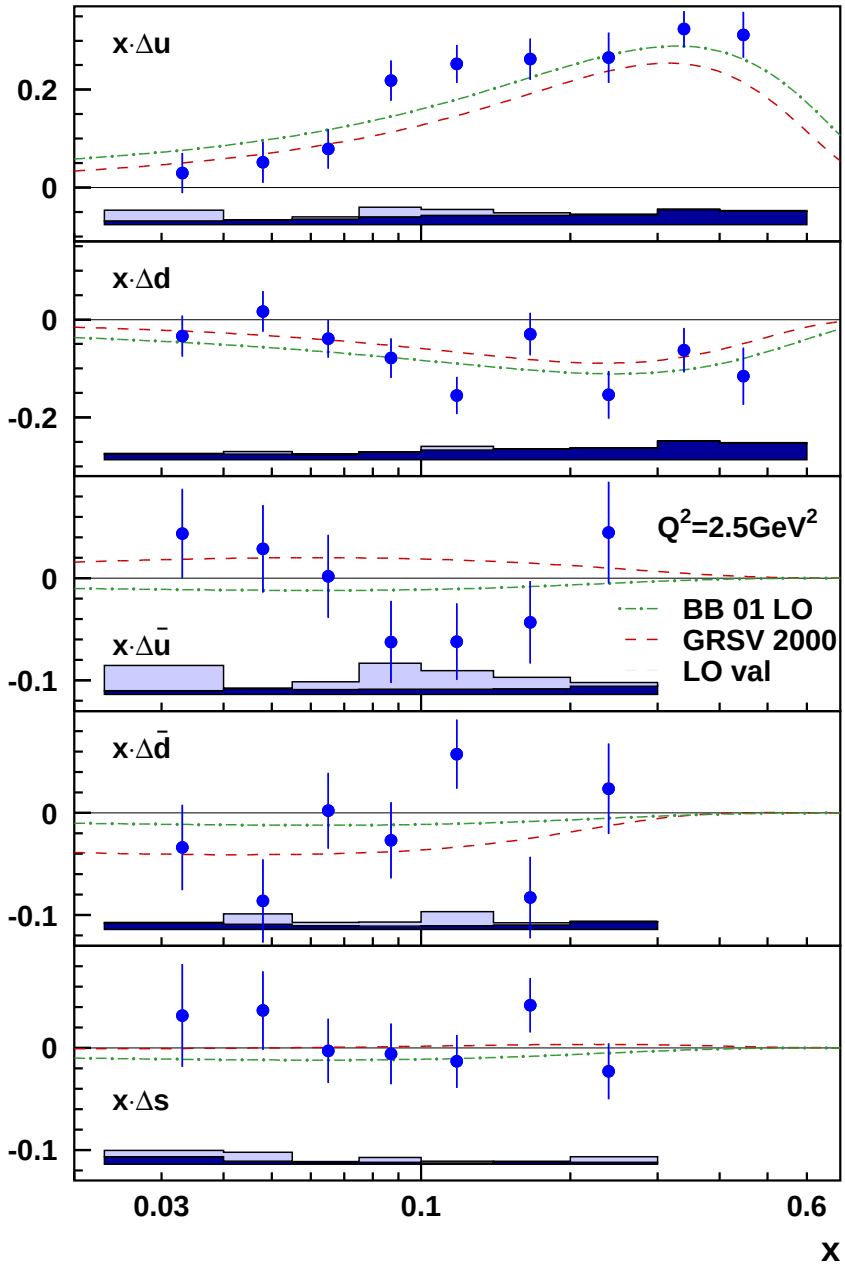

Figure 2.16: Polarized quark distributions for the u,d,$\bar{u}$,$\bar{d}$,s quarks. LO calculations from DIS data [24, 25] are indicated in the figure as well. The average Q^2 is 2.5 GeV^2. Figure from [31]. The light (dark) gray error band is the systematic error associated to uncertainties in the fragmentation model (contribution from Born asymmetries).

To describe exclusive processes, some new variables are introduced such as the average momentum $\mathcal{P} = \frac{P+P'}{2}$ and the momentum transfer to the nucleon $\Delta = (P' - P)$, which can be related to the Mandelstam t via the relation $t = \Delta^2$. Additionally the longitudinal momentum transfer to the real photon is defined as ξ and given by:

$$\xi = \frac{P^+ - P'^+}{P^+ + P'^+} \tag{2.66}$$

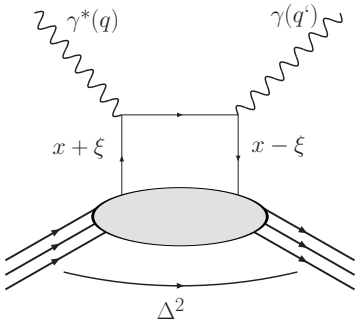

Figure 2.17: Symbolic representation of the variables involved in DVCS.

Since the P^+ momenta of physical states are non-negative, ξ is only physical in the interval $-1 < \xi < 1$. In the Bjorken limit one can write $\xi \sim \frac{x_B}{2-x_B}$. The variable x in Figure 2.17 is not to be confused with x_B. It is an internal variable, defined in the region $[-1, 1]$, and related to the quark momentum fraction. Depending on the interval in ξ where x is falling into, we can associate three different parton pictures: if $x \in [\xi, 1]([-1, -\xi])$, both momentum fractions $x+\xi$ and $x-\xi$ are positive (negative), which can be interpreted as the emission and re-absorption of a quark (antiquark) by the nucleon. If $x \in [-\xi, \xi]$, $x + \xi$ is positive, while $x - \xi$ is negative. The second momentum fraction can be interpreted as the emission of an antiquark with positive momentum $\xi - x$. The partonic interpretation of the range in x and ξ is indicated in Figure 2.18.

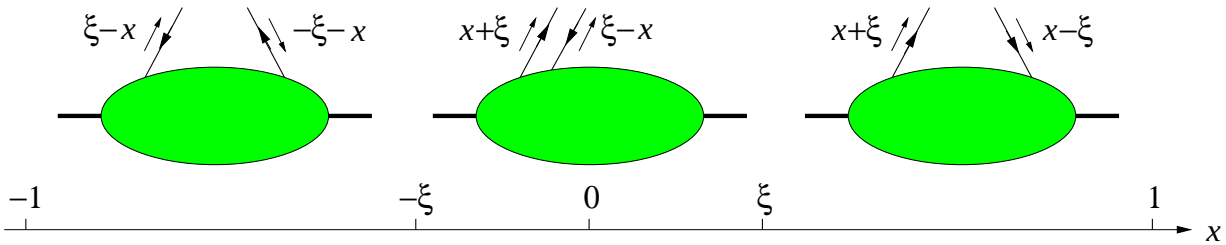

Figure 2.18: The parton interpretation of the coordinates x and ξ in the x-intervals $[-1, -\xi], [-\xi, \xi], [\xi, 1]$.

When comparing Figure 2.17 with the right panel of Figure 2.13, one can easily see that the diagram representing the forward virtual Compton scattering amplitude is a limiting case of the DVCS diagram. In the limit of zero momentum transfer, $\xi = 0$, the virtual Compton amplitude is obtained, therefore ξ is an indication of how far from the forward direction we are away, and is often referred to as the 'skewness' parameter.

Rewriting equations 2.53 to 2.55 for the DVCS process, or more general for the non-forward case, by sandwiching the three leading twist operators between two *different* states, we obtain

a set of 6 new functions for every quark flavor q:

$$\begin{aligned}\int \frac{d\lambda}{2\pi} e^{i\lambda x} \langle P'S'|\bar{\psi}_q(-\frac{1}{2}\lambda n)\not{n}\psi_q(\frac{1}{2}\lambda n)|PS\rangle &= H_q(x,\xi,t)\bar{u}(P'S')\not{n}u(PS) \\ &+ E_q(x,\xi,t)\bar{u}(P'S')\frac{i\sigma^{\mu\nu}\Delta_\nu n_\mu}{2M}u(PS)\end{aligned} \tag{2.67}$$

$$\begin{aligned}\int \frac{d\lambda}{2\pi} e^{i\lambda x} \langle P'S'|\bar{\psi}_q(-\frac{1}{2}\lambda n)\not{n}\gamma_5\psi_q(\frac{1}{2}\lambda n)|PS\rangle &= \tilde{H}_q(x,\xi,t)\bar{u}(P'S')\not{n}\gamma_5 u(PS) \\ &+ \tilde{E}_q(x,\xi,t)\bar{u}(P'S')\frac{\gamma_5(n\cdot\Delta)}{2M}u(PS)\ ,\end{aligned} \tag{2.68}$$

$$\begin{aligned}\int \frac{d\lambda}{2\pi} e^{i\lambda x} \langle P'S'|\bar{\psi}_q(-\frac{1}{2}\lambda n)n_\mu\sigma^{\mu\nu}\psi_q(\frac{1}{2}\lambda n)|PS\rangle &= H_{Tq}(x,\xi,t)\bar{u}(P'S')n_\mu\sigma^{\mu\nu}u(PS) \\ &+ E_{Tq}(x,\xi,t)\bar{u}(P'S')n_\mu\frac{\gamma^{[\mu}i\Delta^{\nu]}}{M}u(PS),\end{aligned} \tag{2.69}$$

following the notation from References [33], [34] and Reference [35][4], the symbol $[\mu\nu]$ indicates anti-symmetrization of the two indices. The functions $H, \tilde{H}, E, \tilde{E}, H_T$ and E_T are called 'Generalized Parton Distributions' (GPDs). Former nomenclature for these functions were also 'skewed' or 'off-forward' parton distribution functions because of their relation to the forward limit of the matrix elements. Equation 2.69 is again chiral-odd thus requiring a quark helicity flip. According to References [36, 37], two more transverse GPDs $\tilde{H}_{T_q}$ and $\tilde{E}_{T_q}$ enter[5] equation 2.69.

It is clear that in the limit $\Delta_\mu \to 0$ the functions $E, \tilde{E}$ and E_T vanish in the equations above. The appearance of the E functions is required to describe the situation where the nucleon spin flips ($S \neq S'$). Also for gluons there exist 8 GPDs, these are indicated by the superscript g (see e.g Reference [36]).

An important difference with the usual parton distributions is that the GPDs can not be regarded as squared amplitudes, since the proton and parton momenta on the left and on the right of the diagrams in Figure 2.17 are different. Therefore the GPD can not be interpreted as probability densities. These functions rather describe interferences of amplitudes.

Just like all QCD functions, GPDs need to be scale dependent. The evolution equations can be split into two regimes, as mentioned before. One is where $|x| > \xi$, corresponding to the creation or annihilation of a quark or antiquark, while the other is the region where $|x| < \xi$, indicating the creation or annihilation of a quark-antiquark pair. The former region has the usual DGLAP evolution as a limiting case, while the latter has the Efremov-Radyuskin-Brodsky-Lepage (ERBL) [38, 39] kernels as limiting cases, which describe particle evolutions. More information on the evolution can be found in References [35, 40, 41].

[4] with a Taylor expansion of the path ordered integral $\mathcal{P}e^{-(\cdots)}$

[5] In Reference [33] a false asymmetry assumption is used. This false assumption causes two of the transverse GPDs to vanish.

2.6.1 Limits and sum rules

As already mentioned before, in the limit of zero momentum transfer ($p = p'$,$s = s'$), we should get the virtual Compton amplitude back. This imposes an interesting limit for the GPDs, stating that the standard parton distribution functions should be found for $\xi = \Delta^2 = 0$:

$$\begin{array}{lll} H^q(x,0,0) = q(x) & \tilde{H}^q(x,0,0) = \Delta q(x) & x > 0 \\ H^q(x,0,0) = -\bar{q}(-x) & \tilde{H}^q(x,0,0) = \Delta\bar{q}(-x) & x < 0 \\ H^g(x,0,0) = xg(x) & \tilde{H}^g(x,0,0) = x\Delta g(x) & x > 0 \end{array}$$

When we integrate equation 2.67 over x we get a non-forward matrix element of a local vector current:

$$\begin{aligned} & \int_{-1}^{1} \mathrm{d}x \int \frac{\mathrm{d}\lambda}{2\pi} e^{i\lambda x} \langle P'S'|\bar{\psi}_q(-\frac{1}{2}\lambda n)\not{n}\psi_q(\frac{1}{2}\lambda n)|PS\rangle \\ = & \int \mathrm{d}\lambda\delta(\lambda)\langle P'S'|\bar{\psi}_q(-\frac{1}{2}\lambda n)\not{n}\psi_q(\frac{1}{2}\lambda n)|PS\rangle = n^\mu\langle P'S'|\bar{\psi}_q(0)\gamma_\mu\psi_q(0)|PS\rangle \quad (2.70) \end{aligned}$$

In general, these non forward matrix elements define the Dirac and Pauli form factors, $F_1^q(t)$ and $F_2^q(t)$, as well as the axial and pseudoscalar ones, $G_A^q(t)$ and $G_P^q(t)$:

$$\begin{aligned} \langle P'S'|\bar{\psi}_q(0)\gamma^\mu\psi(0)|PS\rangle &= \bar{u}(P'S')\left(F_1^q(t)\gamma^\mu + F_2^q(t)\frac{i\sigma^{\mu\nu}\Delta_\nu}{2m}\right)u(PS) \\ \langle P'S'|\bar{\psi}_q(0)\gamma^\mu\gamma_5\psi(0)|PS\rangle &= \bar{u}(P'S')\left(G_A^q(t)\gamma^\mu\gamma_5 + G_P^q(t)\frac{i\gamma_5\Delta^\mu}{2m}\right)u(PS) \end{aligned}$$

Note that $F_1^q(t)$ and $F_2^q(t)$ are not the same as the structure functions $F_1(x,Q^2)$ and $F_2(x,Q^2)$ defined in DIS. According to the derivation above, the following sum rules for the GPDs accordingly are obtained:

$$\begin{array}{ll} \int_{-1}^{1}\mathrm{d}x\,(H^q(x,\xi,t)) = F_1^q(t) & \int_{-1}^{1}\mathrm{d}x\left(\tilde{H}^q(x,\xi,t)\right) = G_A^q(t) \\ \int_{-1}^{1}\mathrm{d}x\,(E^q(x,\xi,t)) = F_2^q(t) & \int_{-1}^{1}\mathrm{d}x\left(\tilde{E}^q(x,\xi,t)\right) = G_P^q(t) \end{array}$$

The Dirac and Pauli form factors can be related to the electromagnetic form factors defined in section 2.1:

$$\begin{aligned} G_E(Q^2) &= F_1(Q^2) - \frac{Q^2}{4M^2}F_2(Q^2) \\ G_M(Q^2) &= F_1(Q^2) + F_2(Q^2), \end{aligned}$$

with $F_1(t) = \sum_q e_q F_1^q(t)$ and $F_2(t) = \sum_q e_q F_2^q(t)$. Since the electromagnetic form factors are defined in elastic scattering, there is a unique relation between t and Q^2. In the sum rules defined above, the ξ dependence has vanished, while we only integrated x away. This is a consequence of Lorentz invariance: when integrating over x, also the reference to the light-cone direction x, with respect to which ξ is defined, vanishes.

More generally, an operator of spin n can be expressed in terms of n possible form factors for a quark of flavor q: A_{2i}^{qn}, B_{2i}^{qn} and C^{qn}. Following Reference [35] we can write:

$$\begin{aligned}\langle P'|O_q^{\mu_1\cdots\mu_n}|P\rangle &= \overline{u}(P')\gamma^{\mu_1}u(P)\sum_{i=0}^{[\frac{n-1}{2}]}A_{2i}^{qn}(t)\Delta^{\mu_2}\cdots\Delta^{\mu_{2i+1}}\overline{P}^{\mu_{2i+2}}\cdots\overline{P}^{\mu_n} \\ &+ \overline{u}(P')\frac{\sigma^{\mu_1\alpha}i\Delta_\alpha}{2M}u(P)\sum_{i=0}^{[\frac{n-1}{2}]}B_{2i}^{qn}(t)\Delta^{\mu_2}\cdots\Delta^{\mu_{2i+1}}\overline{P}^{\mu_{2i+2}}\cdots\overline{P}^{\mu_n} \\ &+ C^{qn}(t)\ \mathrm{Mod}(n+1,2)\ \frac{1}{M}\bar{u}(P')u(P)\Delta^{(\mu_1}\cdots\Delta^{\mu_n)}. \qquad (2.71)\end{aligned}$$

If we take a vector-like operator $\sim\gamma^\mu$ for the operator $O^{\mu_1\cdots\mu_n}$ in equation 2.71, we get a conceptual analog of the matrix element in the integrand on the left hand side of equation 2.67. After a Fourier transformation in x of the right hand side of equation 2.67 we see a relation between the GPDs H and E and the form factors A_i^{qn} and B_i^{qn}.

In general we can write the $(n-1)^{\text{th}}$ Mellin moment of the GPDs E and H as a sum of the general form factors $A^{qn}(t), B^{qn}(t)$ and $C^{qn}(t)$:

$$\begin{aligned}\int_{-1}^{1}dx x^{n-1}H^q(x,\xi,t) &= \sum_{i=0}^{[\frac{n-1}{2}]}A_{2i}^{qn}(t)(-2\xi)^{2i}+\mathrm{Mod}(n+1,2)\ C^{qn}(t)(-2\xi)^n \qquad (2.72)\\ \int_{-1}^{1}dx x^{n-1}E^q(x,\xi,t) &= \sum_{i=0}^{[\frac{n-1}{2}]}B_{2i}^{qn}(t)(-2\xi)^{2i}-\mathrm{Mod}(n+1,2)\ C^{qn}(t)(-2\xi)^n \qquad (2.73)\end{aligned}$$

The above relations indicate the important so-called polynomiality property of GPDs: the $n-$th moment in x of the GPDs E and H are polynomials in ξ of the order $(n+1)$ or of the order n depending on whether (n) is odd or even. Moreover, the sum of E and H is always a polynomial of order n. The contributions to H and E depending on C^{qn} are only present for even n. This property leads to the so-called D-term, which is important in the parametrization of GPDs.

When taking an axial-vector form for the operator in equation 2.71, we see that also the $(n-1)^{\text{th}}$ Mellin moments of GPDs $\tilde{H}$ and $\tilde{E}$ can be expressed as a polynomial of order $\xi^{(n-1)}$ with coefficients $\tilde{A}$ and $\tilde{B}$ respectively. There is no C^{qn} dependent term, however, in the polynomial expressions, because time reversal invariance constrains the analogue of expression 2.73 to come with odd powers of Δ^μ [40].

A last property is obtained from time reversal invariance. From this principle it follows that the GPDs need to be symmetric in ξ: interchanging momenta p and p' is reversing the sign of ξ. We thus have:

$$H^q(x,\xi,t) = H^q(x,-\xi,t) = \left[H(x,-\xi,t)\right]^*.$$

The framework of GPDs is of particular importance because of the relation to the total angular momenta of quarks and gluons as mentioned by Ji in Reference [32]. In general one

can write the angular momentum operator for quark, $\hat{J}_q$, and gluon, $\hat{J}_g$ as:

$$\begin{aligned}\hat{J}_q &= \int d^3x\ \hat{x} \times \vec{T}_q \\ &= \int d^3x\ \left[\psi^\dagger \frac{\vec{\Sigma}}{2}\psi + \psi^\dagger \hat{x} \times (-i\vec{D})\psi\right] \\ &= \frac{1}{2}\Delta\Sigma + L_q \\ \hat{J}_g &= \int d^3x\ \hat{x} \times (\vec{E} \times \vec{B}).\end{aligned} \tag{2.74}$$

The quark angular momentum is generated from the quark momentum density $\vec{T}_q$, while the gluon angular momentum from $\vec{E} \times \vec{B}$. $\vec{\Sigma}$ is the Dirac spin-matrix and $\vec{D} = \vec{\partial} + ig\vec{A}$ is the covariant derivative. If we now sandwich the operator $\hat{J}_{q,g}$ between two quark states we get:

$$J_{q,g}(\mu) = \left\langle P \left| \int d^3x (\vec{x} \times \vec{T}_{q,g})^z \right| P \right\rangle, \tag{2.75}$$

or, explicitly, knowing that the energy-momentum tensor T is an operator of spin $n = 2$ and evaluating the general expression of an operator of spin two in equation 2.71 we can write [35]:

$$\langle P'|T^{\mu\nu}_{q,g}|P\rangle = \bar{u}(P')\Big[A_{q,g}(t)\gamma^{(\mu}\bar{P}^{\nu)} + \frac{B_{q,g}(t)}{2M}\bar{P}^{(\mu}i\sigma^{\nu)\alpha}\Delta_\alpha + \frac{C_{q,g}(t)}{M}\Delta^{(\mu}\Delta^{\nu)}\Big]u(P) \tag{2.76}$$

Taking now equations 2.72 and 2.73 into account, we obtain:

$$\langle P'|T^{\mu\nu}_{q,g}|P\rangle = \int_{-1}^{1} E_q(x,\xi,t)x\mathrm{d}x + \int_{-1}^{1} H_q(x,\xi,t)x\mathrm{d}x \tag{2.77}$$

On the other hand, substituting expression 2.76 into equation 2.75 we get [32]:

$$J_{q,g} = \frac{1}{2}\left[A_{q,g}(0) + B_{q,g}(0)\right]\ . \tag{2.78}$$

Combining both results we can write down Ji's famous sum rule:

$$J^q = \frac{1}{2}\lim_{t\to 0}\int_{-1}^{1}\left(H^q(x,\xi,t) + E^q(x,\xi,t)\right)x\mathrm{d}x. \tag{2.79}$$

For gluons, a similar sum rule exists.

2.6.2 Impact parameter space

As discussed in the previous section, GPDs contain important information about the total angular momentum of quarks and gluons. At present, it forms the only mechanism giving

access to total angular momenta[6]. The GPDs, however, have no direct interpretation in terms of probability. Nevertheless, a very important analogy between form factors and GPDs can be observed. As mentioned before, a standard form-factor is defined as the non-forward matrix element of the local current operator. The *forward* limit of the electromagnetic form-factor describes how the *forward* matrix element, the charge, is distributed in space. This was shown in section 2.1, where the Fourier transform of the dipole fit yielded an exponential charge distribution. In full analogy, when taking the *forward* matrix elements of GPDs, the standard parton distribution functions (PDFs) are obtained. Taking now the Fourier transform describes how these PDFs are distributed in position space. This can be seen by moving to the 'impact parameter space', where we keep the momentum in the light-cone + direction, while Fourier transforming from transverse momentum to transverse position:

$$|p^+, \mathbf{b}, \lambda\rangle = \int \frac{d^2\mathbf{p}}{16\pi^3}\, e^{-i\mathbf{p}\cdot\mathbf{b}}\, |p^+, \mathbf{p}, \lambda\rangle, \tag{2.80}$$

where $|p^+, \mathbf{p}, \lambda\rangle$ is an infinite momentum helicity state, and $\mathbf{b}$ indicates a position in the transverse plane, defined as: $\mathbf{b} = \sum_i x_i \mathbf{b}_i$, where the sum goes over all partons i with momentum fraction x_i. It thus forms the center of transverse momentum. When now transforming the GPD H^q to impact parameter space, in the region where $\xi = 0$, we get the impact parameter dependent PDF $q(x, \mathbf{b})$:

$$q(x, \mathbf{b}) = \int \frac{\mathrm{d}\vec{\Delta}}{(2\pi)^2} e^{-i\mathbf{b}\cdot\vec{\Delta}} H^q(x, 0, t = -\Delta^2) \qquad (\Delta = (0, \vec{\Delta}, 0)) \tag{2.81}$$

This result is described in detail in Reference [42], where it is shown that $q(x, \mathbf{b})$ satisfies positivity constraints, thus allowing $q(x, \mathbf{b})$ to be interpreted as the probability to find a quark with momentum fraction x at transverse distance $\mathbf{b}$ from the center of transverse momentum. Recalling now that $\sum_q e_q \int \mathrm{d}x H^q = F_1(t)$, it follows from the observations above that the Fourier transform of the form factor $F_1(t)$ can be interpreted as the charge distribution as a function of transverse distance $\mathbf{b}$.

In the partonic picture of the infinite momentum frame, form factors describe transverse distributions, while PDFs describe longitudinal momenta distributions. From this picture it is obvious that GPDs contain both *position* and *momentum* information, two necessary inputs to get orbital angular momenta. This picture is presented in Figure 2.19.

Also for the GPD E, a Fourier transform to impact parameter space could be applied, yielding the function $e^q(x, \mathbf{b})$, which has no direct interpretation, since it refers to the nucleon helicity flip, and thus is non-diagonal. Therefore it is suggested to change from a longitudinal to transverse basis so that one can write [42]:

$$q_X(x, \mathbf{b}) = q(x, \mathbf{b}) - \frac{1}{2M}\frac{\partial}{\partial b_y} e^q(x, \mathbf{b}). \tag{2.82}$$

This equation describes the transverse distortion of the unpolarized impact parameter dependent PDF $q(x, \mathbf{b})$ if the state is polarized in a transverse direction.

The impact parameter space distribution of GPD $\tilde{H}^q$ is $\Delta q(x, \mathbf{b})$, which can be interpreted as the polarized analogue of $q(x, \mathbf{b})$.

[6]though, through 'transversity', L_q may be accessed [27]

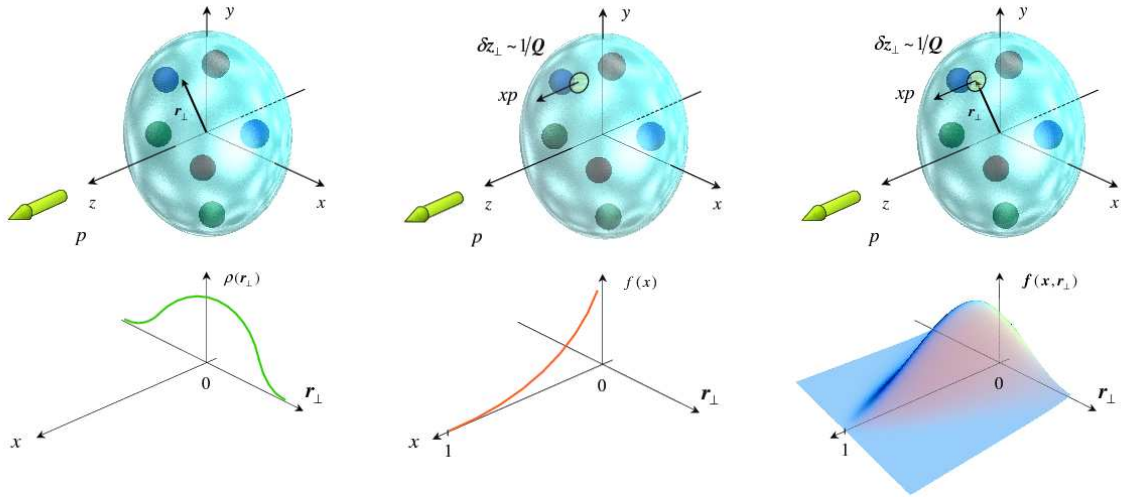

Figure 2.19: In the infinite momentum frame, form factors describe transverse distributions (left), and PDFs longitudinal momenta distributions (middle). GPDs generalize these distributions (right). Figure from Reference [41] (the authors use $\mathbf{r}_\perp$ rather than $\mathbf{b}$).

From the deviations above, it can be seen that the impact parameter space picture yields 'nucleon tomography', a process where the transverse electromagnetic distributions are evaluated in the $b_x - b_y$ plane for fixed values of x, in analogy with macroscopic tomography with X-rays. Reference [43] describes how GPDs give a three-dimensional picture of the nucleon.

2.6.3 Probing GPDs

Experimentally GPDs can be probed through exclusive processes such as DVCS, or exclusive meson production. If the recoiling target is not directly detected, two variables are introduced to indicate the exclusivity of the process. One is the missing mass M_X, which is just the square of the differences between the initial and detected four-momenta:

$$M_X = \sqrt{(P + q - p_h)^2}, \tag{2.83}$$

the other variable is the missing energy ΔE:

$$\begin{aligned} \Delta E &= \frac{M_X^2 - M_{\text{target}}^2}{2M_{\text{target}}} \\ &= \frac{t}{2M_{\text{target}}} + \nu - E_h \end{aligned} \tag{2.84}$$

A different combination of GPDs is obtained depending on both the observables and the final state of the process being considered. Different final states select different combinations of different GPDs. Cross sections give a quadratic combination of GPDs, while target or beam related asymmetries access a product of GPDs. The next two sections will briefly report on how GPDs can be measured via DVCS and meson production.

2.6.3.1 DVCS

The most direct process probing GPDs is DVCS. The DVCS amplitude gives access to the four GPDs $H, E, \tilde{H}$ and $\tilde{E}$. DVCS has the same final state as the Bethe-Heithler (BH) process, where the scattering lepton is emitting a Bremsstrahlung photon. Both processes are drawn in Figure 2.20. As a consequence of the indistinguishability of these processes, both amplitudes interfere, so that the photon production amplitude can be written as:

$$|\tau|^2 = |\tau_{\mathrm{BH}} + \tau_{\mathrm{DVCS}}|^2 = |\tau_{\mathrm{BH}}|^2 + |\tau_{\mathrm{DVCS}}|^2 + \underbrace{\tau_{\mathrm{DVCS}}\tau^*_{\mathrm{BH}} + \tau_{\mathrm{BH}}\tau^*_{\mathrm{DVCS}}}_{\mathcal{I}} \tag{2.85}$$

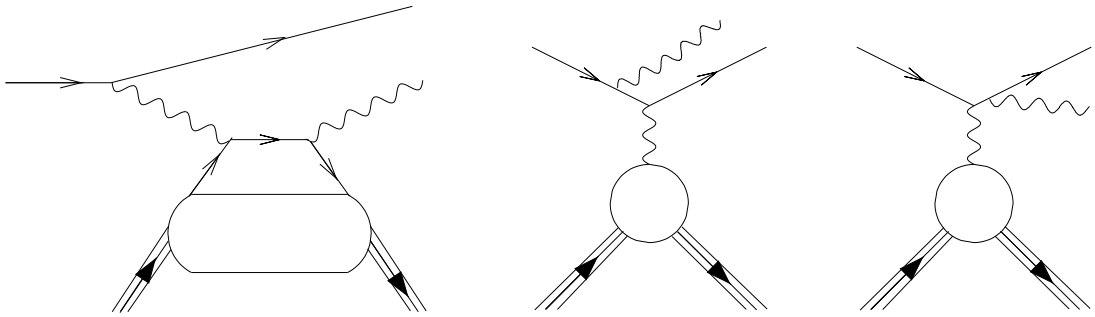

Figure 2.20: The final state of the DVCS (left) and the Bethe-Heithler (middle and right) are indistinguishable.

The cross section for DVCS has been measured at the HERA experiments H1 and ZEUS and is reported e.g. in Reference [44]. As mentioned before, the cross section measurements give access to a quadratic combination of GPDs.

At HERMES energies, the DVCS cross section is small, so that DVCS needs to be analyzed via asymmetry measurements. The asymmetries in general will depend on the azimuthal angle ϕ which is the angle between the lepton scattering plane (i.e. the plane defined by $\vec{k}$ and $\vec{k'}$) and the photon production plane (i.e. the plane defined by $\vec{q}$ and $\vec{p_\gamma}$). In Reference [45] it is shown that the amplitudes can be written as a finite sum of Fourier harmonics:

$$|\mathcal{T}_{\mathrm{BH}}|^2 = \frac{e^6}{x_B^2 y^2 (1+\epsilon^2)^2 \Delta^2 \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\mathrm{BH}} + \sum_{n=1}^{2} c_n^{\mathrm{BH}} \cos(n\phi) + s_1^{\mathrm{BH}} \sin(\phi) \right\} \tag{2.86}$$

$$|\mathcal{T}_{\mathrm{DVCS}}|^2 = \frac{e^6}{y^2 Q^2} \left\{ c_0^{\mathrm{DVCS}} + \sum_{n=1}^{2} \left[c_n^{\mathrm{DVCS}} \cos(n\phi) + s_n^{\mathrm{DVCS}} \sin(n\phi) \right] \right\} \tag{2.87}$$

$$\mathcal{I} = \frac{\pm e^6}{x_B y^3 \Delta^2 \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\mathcal{I}} + \sum_{n=1}^{3} \left[c_n^{\mathcal{I}} \cos(n\phi) + s_n^{\mathcal{I}} \sin(n\phi) \right] \right\}, \tag{2.88}$$

where $\mathcal{P}_{1(2)}$ are the BH propagators. They can be expanded in ϕ as $\mathcal{P}_i(\phi) = A_i + B_i \cos\phi$, where B_i is suppressed by at least $\mathcal{O}(1/Q)$. The +(-) sign in equation 2.88 stands for a negatively (positively) charged lepton beam. For unpolarized targets, only the $s_n^{\mathcal{I}}$ moments depend on the beam helicity λ. In the interference term, only the coefficients $c_0^{\mathcal{I}}, c_1^{\mathcal{I}}$ and $s_1^{\mathcal{I}}$ are leading twist coefficients, the other correspond to higher order corrections.

We can now consider asymmetries with respect to either the beam charge or to the beam spin for real photon production on an unpolarized target. We can additionally assume that the ϕ dependence in the BH amplitude cancels in numerator and denominator. Moreover, the $|\tau_{DVCS}|^2$ amplitude is small and can accordingly be ignored in the kinematic regime of fixed target experiments. At leading twist we can write:

$$A_{LU}(\phi) = \frac{\mathrm{d}\sigma(\overrightarrow{e^+}p) - \mathrm{d}\sigma(\overleftarrow{e^+}p)}{\mathrm{d}\sigma(\overrightarrow{e^+}p) + \mathrm{d}\sigma(\overleftarrow{e^+}p)} \sim \frac{x_B}{y}\frac{s_1^I}{c_0^{BH}}\sin(\phi). \tag{2.89}$$

$$A_C(\phi) = \frac{\mathrm{d}\sigma(e^+p) - \mathrm{d}\sigma(e^-p)}{\mathrm{d}\sigma(e^+p) + \mathrm{d}\sigma(e^-+p)} \sim \frac{x_B}{y}\frac{c_1^I}{c_0^{BH}}\cos(\phi). \tag{2.90}$$

In the above equations the notation LU stands for a longitudinally polarized beam and an unpolarized target. It was assumed that the denominator in the above equations is dominated by the $|\tau_{BH}|^2$ amplitude. For the beam charge asymmetry an unpolarized beam was assumed ($\lambda = 0$). The ϕ modulations of the asymmetries are shown in Figure 2.21. The additional $\sin\phi$ modulation in the beam charge asymmetry is due to a not entirely unpolarized beam, while the absence of the additional $\sin 2\phi$ moment shows the validity of the approximation: if higher order effects would play a role a $\sin 2\phi$ moment would be present.

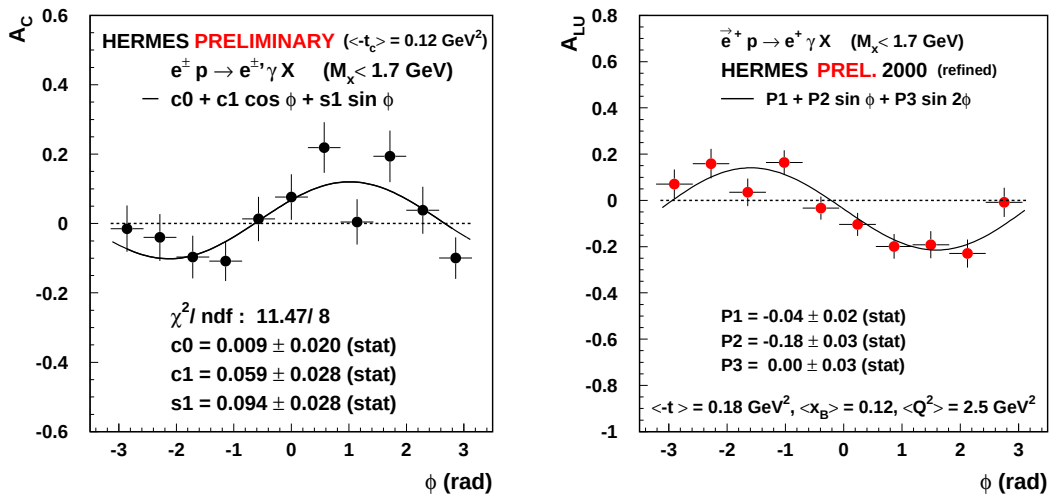

Figure 2.21: ϕ modulation of the beam charge A_C (left) and beam spin A_{LU} for DVCS on a proton target. Figure available from the HERMES web site [46].

Using the calculation of the $c_1^{\mathcal{I}}$ and $s_1^{\mathcal{I}}$ amplitudes from Reference [45], we see that these can be related to the real and imaginary part of the DVCS amplitude $M_{\mathrm{unp}}^{1,1}$ respectively. The matrix element $M_{\mathrm{unp}}^{1,1}$ can be written as a linear combination of the Compton Form Factors (CFFs) $\mathcal{H}, \mathcal{E}$ and $\tilde{\mathcal{H}}$:

$$M_{\mathrm{unp}}^{1,1} = F_1(t)\mathcal{H}_1(\xi,t) + \frac{x_B}{2-x_B}\left(F_1(t)+F_2(t)\right)\tilde{\mathcal{H}}_1(\xi,t) - \frac{t}{4M_p^2}F_2(t)\mathcal{E}_1(\xi,t), \tag{2.91}$$

where $F_1(t)$ and $F_2(t)$ are the Dirac and Pauli form factors. The CFFs are entirely twist-2 objects and can be related to the GPDs. At HERMES kinematics, the coefficients of the CFFs $\mathcal{E}$ and $\tilde{\mathcal{H}}$ in the previous equation are small.

According to Reference [47], the real and imaginary part of $\mathcal{H}$ can be related to the GPD H:

$$\begin{aligned} \mathrm{Im}\mathcal{H}_1 &= -\pi \sum_q e_q^2 \left(H^q(\xi,\xi,t) - H^q(-\xi,\xi,t)\right) \\ \mathrm{Re}\mathcal{H}_1 &= \sum_q e_q^2 \left(P \int_{-1}^{1} H^q(x,\xi,t) \left(\frac{1}{x-\xi} + \frac{1}{x+\xi}\right) dx\right) \end{aligned} \tag{2.92}$$

where P denotes Cauchy's principal value. This shows how GPDs can be accessed via beam spin and beam charge asymmetries.

Analogously, the transverse target related spin asymmetry A_{UT} can be expressed in terms of CFFs [45]:

$$A_{UT}(\phi,\phi_S) \propto \mathrm{Im}(F_2\mathcal{H} - F_1\mathcal{E}) \sin(\phi-\phi_S)\cos(\phi) + \mathrm{Im}(F_2\tilde{\mathcal{H}} - F_1\xi\tilde{\mathcal{E}}) \cos(\phi-\phi_S)\sin(\phi), \tag{2.93}$$

the angle ϕ_S is the angle between the lepton scattering plane and the spin of the proton. The angles ϕ and ϕ_S are shown in Figure 2.22. The $\sin(\phi-\phi_S)\cos(\phi)$ moment of the asymmetry gives access to the GPDs E and H. By combining the beam-spin and beam-charge related asymmetries with the transverse target asymmetry the GPDs H and E can be probed, yielding an estimate of J^q. One of these estimates can be found in Reference [48].

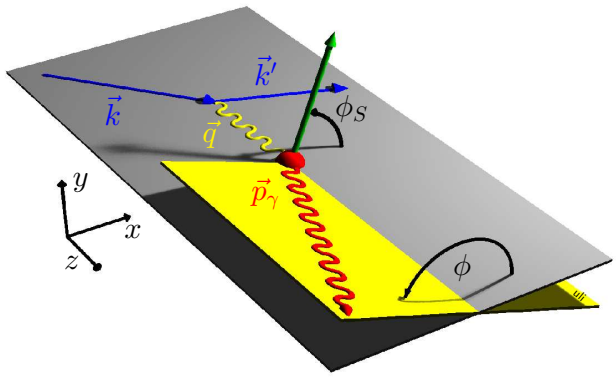

Figure 2.22: Definition of the angles ϕ and ϕ_S for the asymmetry A_{UT}.

2.6.3.2 Meson production

Another way of accessing GPDs is through exclusive meson production. In Reference [49] a general proof of factorization for the exclusive meson production amplitude is given. The authors prove that the amplitude for processes like:

$$\gamma^* + p \rightarrow M(q+\Delta) + p'(p-\Delta) \tag{2.94}$$

can be factorized in a hard part describing the interaction of the virtual photon with the quark lines H_{ij}, a meson light front wave function ϕ_j, and a GPD, f_i, for partons of type i:

$$\mathcal{M} = \sum_{i,j} \int_0^1 \mathrm{d}z \int \mathrm{d}\xi f_i(\xi, \xi - x, t) H_{ij}(Q^2\xi/x, Q^2, z)\phi_j(z) \tag{2.95}$$

The sum runs over the parton types i and j that connect the hard scattering part to the GPDs and to the meson. The proof is valid for longitudinally polarized virtual photons γ_L^* and applies to all types of mesons produced. In the same reference the authors proof that the amplitude for transversely polarized photons γ_T^* is suppressed by $\mathcal{O}(\frac{1}{Q^2})$. When working with transverse photons not only the hard scattering part needs to be changed but also more general non perturbative objects are required, yielding an additional complication.

The theorem is valid in the region where $Q^2 \gg \Lambda_{QCD}$ and $t \leq \Lambda_{QCD}$, and for leading power in Q^2. Figure 2.23 shows the factorization theorem graphically, with only the leading contribution to the hard part. Pseudoscalar meson production occurs via the quark exchange diagram drawn in Figure 2.23(b), while for vector meson production an additional interaction mechanism enters, namely through 2 gluon exchange, depicted in Figure 2.23(b).

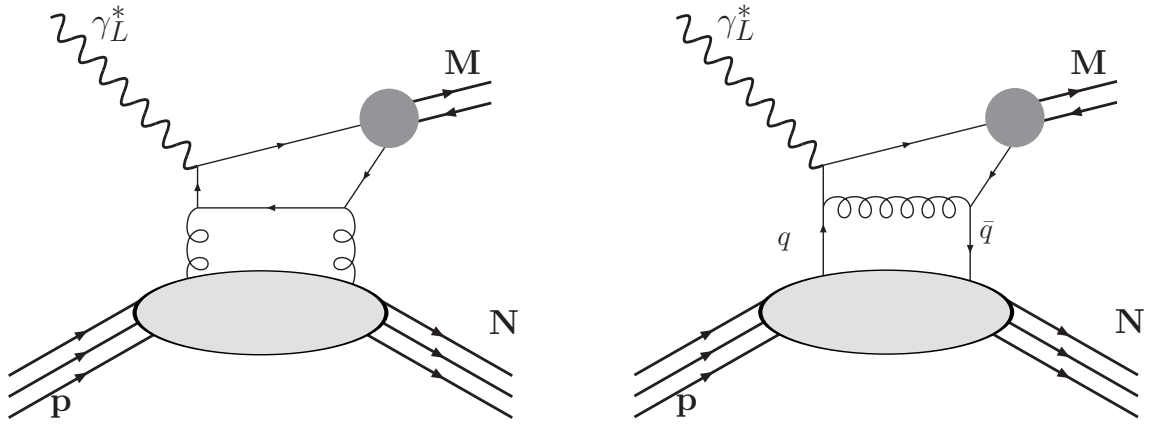

Figure 2.23: Scheme of the factorization theorem for the lowest order hard scattering part. The meson wave function is indicated by the dark circle, the soft part described by GPDs with the gray ellipse.

Depending on the quantum numbers of the produced meson, different GPDs can be probed, because parity constraints require the meson and the GPD to have the same parity. When the final meson is a longitudinally polarized vector meson, like a ρ_L, ϕ_L or J/ψ_L, the amplitudes are sensitive to the vector GPDs E and H. At low x, the gluon production diagram will dominate. When on the other hand the produced meson is a pseudoscalar, the axial-vector GPDs $\tilde{E}$ and $\tilde{H}$ enter the amplitude. In this case, the gluon coupling is prohibited by charge conjugation invariance: the hard scattering coefficient has to be zero when it couples a virtual photon and a pseudoscalar meson to a pair of gluons. If, finally, the vector mesons are transversely polarized, we get access to the transverse GPDs. Also here, the gluon coupling is forbidden due to helicity conservation in the hard scattering. Hard exclusive meson production thus acts as a helicity filter: when studying hard exclusive vector (pseudoscalar) meson production, the

helicity averaged GPDs H and E ($\tilde{H}, \tilde{E}$) are probed. Exclusive ρ, ϕ, ω vector meson production has already been studied at the HERMES experiment, see e.g. References [50, 51].

2.6.4 Pseudoscalar meson production

When studying pion production on a proton target the concept of the so called 'pion-pole' needs to be introduced first. Therefore we consider the π^+ cross section, which was measured in the $'70s$ by various experiments, albeit at low Q^2. More details can be found in e.g. Reference [52]. For longitudinal polarized virtual photons at small t, the cross section can be written as:

$$\sigma_L \sim \frac{-tQ^2}{(t-m_\pi^2)^2} g_{\pi NN}^2(t) F_\pi^2(Q^2,t), \tag{2.96}$$

where $F_\pi(Q^2,t)$ is the pion form factor, and the factor $g_{\pi NN}(t)$ a term which can be related to the probability to have a virtual π^+ meson inside the proton at a given t. The process can be seen as the scattering of a virtual photon from a virtual pion in the nucleon, where the pion becomes on-shell. One immediately sees a pole in the expression of σ_L if $t \to m_\pi^2$. Even though t is constrained to be negative, and the pole consequently can never be reached, the cross section for pion electroproduction is rising with decreasing $-t$. This observation is explained by the pole in equation 2.96. The cross section measured by these early experiments is indicated in Figure 2.24.

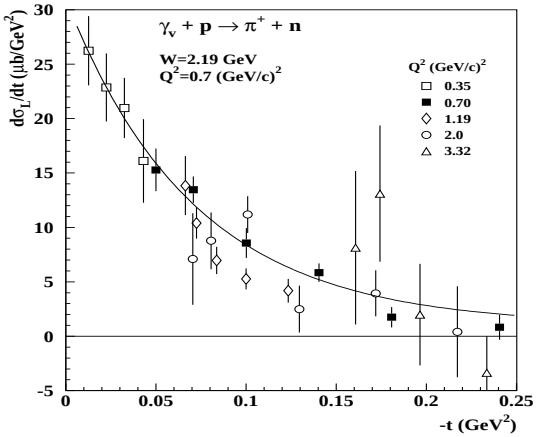

Figure 2.24: Longitudinal cross section for the reaction $\gamma_L^* + p \to n + \pi^+$ for various Q^2 and W, scaled towards $W = 2.19$ GeV and $Q^2 = 0.7$ GeV^2. Figure from Reference [53].

The pseudoscalar nucleon form factor $G_A^q(t)$ is dominated at low t by the contribution of the pion pole [54]:

$$\lim_{t \to m_\pi^2} G_A^q(t) = \frac{1}{2}\, \tau_{qq}^3\, \frac{4 g_A M_N^2}{m_\pi^2 - t}\,, \tag{2.97}$$

where $g_A \approx 1.267$ is the nucleon isovector axial charge and τ^3 is the Pauli matrix in the flavor space. From the sum rule of the GPD $\tilde{E}$ it now follows that also the latter GPD should have a 'pion pole' contribution [55].

The presence of the pion pole singularity in the GPD $\widetilde{E}$ leads to a strong dependence of the differential cross section of π^+ production on the *transverse* polarization of the target [56]. The analysis of the transverse target spin asymmetry for exclusive π^+ production is presently under study at HERMES [57].

In genera, the electroproduction for hard meson production can be written in the framework of GPDs as [58]:

$$\frac{d\sigma_L}{dt} = \frac{1}{16\pi\,(s-M_N^2)\,\Lambda(s,-Q^2,M_N^2)}\,\frac{1}{2}\sum_{h_N}\sum_{h_N^{'}}|\mathcal{M}^L(\lambda_M=0,h_N^{'};h_N)|^2\,, \qquad (2.98)$$

where $\mathcal{M}^L$ is the amplitude for the production of a meson with helicity $\lambda_M = 0$ by a longitudinal photon, and where h_N, $h_N^{'}$ are the initial and final nucleon helicities. Furthermore in equation (2.98), the standard kinematic function $\Lambda(x,y,z)$ is defined by

$$\Lambda(x,y,z) = \sqrt{x^2+y^2+z^2-2xy-2xz-2yz}\,. \qquad (2.99)$$

which gives $\Lambda(s,-Q^2,M_N^2) = 2M_N|\vec{q}_L|$, where $|\vec{q}_L|$ is the virtual photon momentum in the lab system. In the case of pion electroproduction, the amplitude $\mathcal{M}_\pi^L$ can be written as [54]:

$$\begin{aligned}\mathcal{M}_\pi^L &= -ie\,\frac{4}{9}\,\frac{1}{Q}\,\left[\int_0^1 dz\frac{\Phi_\pi(z)}{z}\right]\,\frac{1}{2}\,(4\pi\alpha_s)\\ &\times \left\{A_{\pi N}\,\bar{N}(p^{'})\not{n}\gamma_5\,N(p)\,+\,B_{\pi N}\,\bar{N}(p^{'})\,\gamma_5\frac{\Delta\cdot n}{2m_N}\,N(p)\right\}\,. \qquad (2.100)\end{aligned}$$

At low t (i.e. low Δ) it can be seen that the amplitudes A will dominate the cross section. For $\pi^0\,p$, the amplitudes A and B in equation (2.100) are given by [59, 60] :

$$A_{\pi^0 p} = \int_{-1}^1 dx\,\frac{1}{\sqrt{2}}\left(e_u\,\widetilde{H}^u\,-\,e_d\,\widetilde{H}^d\right)\,\left\{\frac{1}{x-\xi+i\epsilon}+\frac{1}{x+\xi-i\epsilon}\right\}, \qquad (2.101)$$

$$B_{\pi^0 p} = \int_{-1}^1 dx\,\frac{1}{\sqrt{2}}\left(e_u\,\widetilde{E}^u\,-\,e_d\,\widetilde{E}^d\right)\,\left\{\frac{1}{x-\xi+i\epsilon}+\frac{1}{x+\xi-i\epsilon}\right\}. \qquad (2.102)$$

while for $\pi^+\,n$ electroproduction on the proton, the amplitudes A and B are given by [59, 56] :

$$A_{\pi^+ n} = -\int_{-1}^1 dx\,\left(\widetilde{H}^u\,-\,\widetilde{H}^d\right)\,\left\{\frac{e_u}{x-\xi+i\epsilon}+\frac{e_d}{x+\xi-i\epsilon}\right\}, \qquad (2.103)$$

$$B_{\pi^+ n} = -\int_{-1}^1 dx\,\left(\widetilde{E}^u\,-\,\widetilde{E}^d\right)\,\left\{\frac{e_u}{x-\xi+i\epsilon}+\frac{e_d}{x+\xi-i\epsilon}\right\}. \qquad (2.104)$$

In the equation 2.100 the pion distribution amplitude Φ_π enters which in the asymptotic limit (i.e. $Q^2\to\infty$) can be written as [54]:

$$\Phi_\pi(z) = \sqrt{2}\,f_\pi\,6z(1-z)\,, \qquad (2.105)$$

with $f_\pi = 0.0924$ GeV the pion decay constant, known from the pion weak decay.

The function $\widetilde{E}^u - \widetilde{E}^d$ contains a strong pion pole singularity in the limit $t\to m_\pi^2$. The pion pole gives a dominant contribution to the π^+ electroproduction, and corresponds to the region

where $|x| < \xi$ in Figure 2.18. It can be seen that the amplitude $\mathcal{M}^L_{\pi^{+(0)}}$ has a $\frac{1}{Q}$ dependence. The reduced cross section, defined as $\sigma_{\rm red} = \sum_{h_N} \sum_{h'_N} |\mathcal{M}^L(\lambda_M = 0, h'_N; h_N)|^2$ thus should have a $\frac{1}{Q^2}$ dependence (and consequently the overall cross section a $\frac{1}{Q^6}$ dependence). The Q^2 dependence of the reduced cross section for exclusive π^+ production at HERMES is shown in Figure 2.25. A fit to the data of a function of the form $\frac{1}{Q^p}$ yielded values of $p = 1.9 \pm 0.5$ $(0.02 < x < 0.18), p = 1.7 \pm 0.6$ $(0.18 < x < 0.26)$ and $p = 1.5 \pm 1.0$ $(0.26 < x < 0.8)$ in the different x bins [61], in good agreement with the expectations. This observation forms evidence of the validity of expressions 2.98 and 2.103-2.104. Another property of the reduced cross section is that the pion-pole part is independent of x_B at a given Q^2, i.e. $\frac{\partial}{\partial x}\left(\mathcal{M}^L_{\pi^+,\rm pole}\right) = 0$ [58].

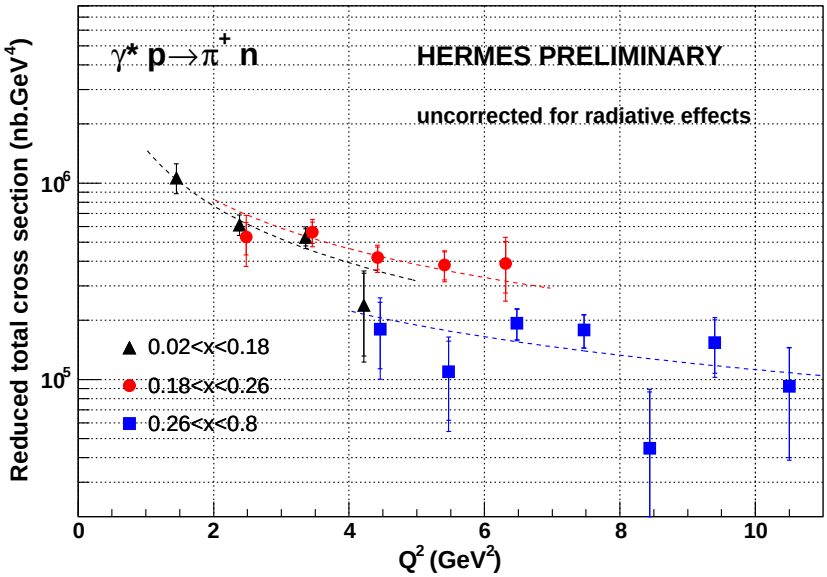

Figure 2.25: Q^2 dependence of the reduced cross section for the exclusive reaction $e + p \rightarrow e + \pi^+ + n$ obtained by HERMES in three different x bins. A fit of the form $\frac{1}{Q^p}$ is also shown in the figure. Fit values are mentioned in the text.

Comparing the amplitudes A and B for π^0 production, equations 2.101 and 2.102, with those for π^+ production, equations 2.103 and 2.104, it can be seen that in the $\pi^0 p$ channel the GPDs $\tilde{H}^u$ and $\tilde{H}^d$ enter with the same sign (since $\text{sign}(e_d) = -\text{sign}(e_u)$) whereas in the $\pi^+ p$ channel these enter with opposite signs. Taken the forward limit of the GPDs into account, $\tilde{H}^u \rightarrow \Delta u$ and $\tilde{H}^d \rightarrow \Delta d$, together with the fact that Δu and Δd have an opposite sign (see e.g. Figure 2.16), it can be seen that the production amplitude for π^0 mesons will be substantially smaller than the one for π^+ mesons. Moreover, as mentioned before, the function $\tilde{E}^u - \tilde{E}^d$ contains a strong pion pole contribution. This pole contribution is strongest at low t and absent in the case of π^0 prodcution. The relative importance of the latter term in the $\pi^+ \backslash \pi^0$ production ratio is indicated in Figure 2.26. The full line represents the total ratio, whereas the dotted line the pseudovector $(\tilde{H})$ contribution to the ratio only. In Reference [54] a ratio of $\pi^0 : \pi^+ \sim 1 : 5$ is predicted at $Q^2 = 10$ GeV^2 and $t = -0.3$ GeV^2.

Even though the predicted exclusive π^0 cross section is 20% smaller than the one for exclusive π^+ production, it is worth an effort to analyze π^0 cross section rates, especially

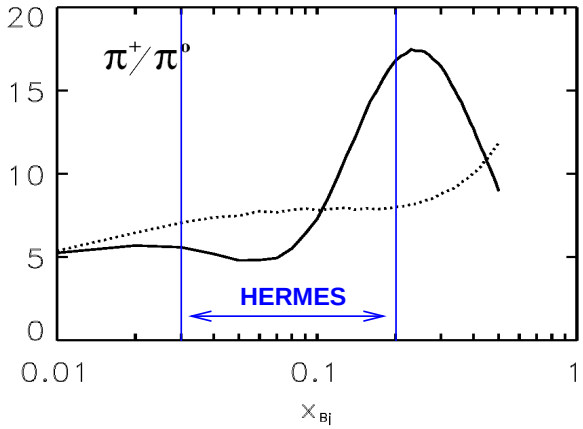

Figure 2.26: Ratio of the cross section for exclusive π^+ and π^0 production at $Q^2 = 10$ GeV^2 and $t = t_{\rm min}$. The dotted curve corresponds to the pseudovector contribution alone. Figure from Reference [59]. The x range accessible in the HERMES experiment is indicated as well.

in comparison with the π^+ cross section data. These data helps to disentangle the different GPDs: all 4 GPDs enter in the description of the DVCS process, while for pion production only the GPDs $\tilde{H}$ and $\tilde{E}$ are required. Data on exclusive π^0 production would strongly influence model calculation since it depends on the pseudovector contribution $\tilde{H}$ only (at low t). It consequently provides limits on the GPD $\tilde{E}$, in combination with the π^+ data. $\tilde{E}$ is not constrained by quark densities in the forward limit, and hence is more difficult to model.

Apart from the π^0 hard electroproduction, also the hard exclusive η production cross section does not contain a pion pole. The amplitudes A and B for the latter process are given by [62] :

$$A_{\eta p} = \int_{-1}^{1} dx \, \frac{1}{\sqrt{6}} \left(e_u \, \widetilde{H}^u + e_d \, \widetilde{H}^d - 2\, e_s \, \widetilde{H}^s \right) \left\{ \frac{1}{x - \xi + i\epsilon} + \frac{1}{x + \xi - i\epsilon} \right\} \quad (2.106)$$

$$B_{\eta p} = \int_{-1}^{1} dx \, \frac{1}{\sqrt{6}} \left(e_u \, \widetilde{E}^u + e_d \, \widetilde{E}^d - 2\, e_s \, \widetilde{E}^s \right) \left\{ \frac{1}{x - \xi + i\epsilon} + \frac{1}{x + \xi - i\epsilon} \right\} \quad (2.107)$$

The same reference predicts a ratio $\pi^0/\eta \sim 2/3$. The authors from Reference [58] on the other hand found a similar ratio of about $\pi^0/\eta \sim 4/5$.

Analyzing the ratio of pseudoscalar meson production offers the additional advantage that only the integral over the light cone fraction z of the meson distribution amplitude ($\phi_j(z)$ in equation 2.95) needs to be considered, because practically the same combinations of GPDs enter in the amplitude in both numerator and denominator. The integral of $\phi_j(z)$ is determined by the quantum numbers of the meson. The ratio hence will only be proportional to the ratio of GPDs. In Reference [63] the production ratio of hard exclusive π^0 and η mesons is predicted to be:

$$\pi^0 : \eta = \frac{1}{2} \left(\frac{2}{3}\Delta u + \frac{1}{3}\Delta d \right)^2 : \frac{1}{6} \left(\frac{2}{3} \cdot 1.27 \Delta u - \frac{1}{3} \cdot 1.27 \Delta d + \frac{2}{3} \cdot 1.14 \Delta s \right)^2 \quad (2.108)$$

These ratio's can also be used to predict the production ratio of π^0 and η mesons on a proton over a neutron target. In the case of exclusive π^0 production, the production ratio for

these targets can be written as:

$$(\pi^0)_n : (\pi^0)_p = \frac{1}{2}\left(\frac{2}{3}\Delta u_n + \frac{1}{3}\Delta d_n\right)^2 : \frac{1}{2}\left(\frac{2}{3}\Delta u_p + \frac{1}{3}\Delta d_p\right)^2 \tag{2.109}$$

Assuming SU(2) isospin symmetry we can write $\Delta d_n = \Delta u_p$ and $\Delta u_n = \Delta d_p$. The ratio now reads:

$$(\pi^0)_n : (\pi^0)_p = \frac{1}{2}\left(\frac{2}{3}\Delta d_p + \frac{1}{3}\Delta u_p\right)^2 : \frac{1}{2}\left(\frac{2}{3}\Delta u_p + \frac{1}{3}\Delta d_p\right)^2 \tag{2.110}$$

From Figure 2.16 it can be seen that $\Delta u \sim -2 \cdot \Delta d$. From the formula above it thus follows that π^0 production on the neutron will be suppressed with respect to the production rate on the proton. For the η a similar effect is predicted, although not that big:

$$(\eta)_n : (\eta)_p = \frac{1}{2}\left(\frac{2}{3}\Delta d_p - \frac{1}{3}\Delta u_p\right)^2 : \frac{1}{2}\left(\frac{2}{3}\Delta u_p - \frac{1}{3}\Delta d_p\right)^2, \tag{2.111}$$

where we have assumed $\Delta s \approx 0$. From the relation $\Delta u > \Delta d$, it follows that $2\Delta u - \Delta d > 2\Delta d - \Delta u$.

The authors of Reference [63] have qualified the effect using in their model calculation a representation of Δu_V and Δd_V from Reference [64] (and the assumption $\Delta s \approx 0$). The result of the calculation for $Q^2 = 20\text{GeV}^2$ is shown in Figure 2.27.

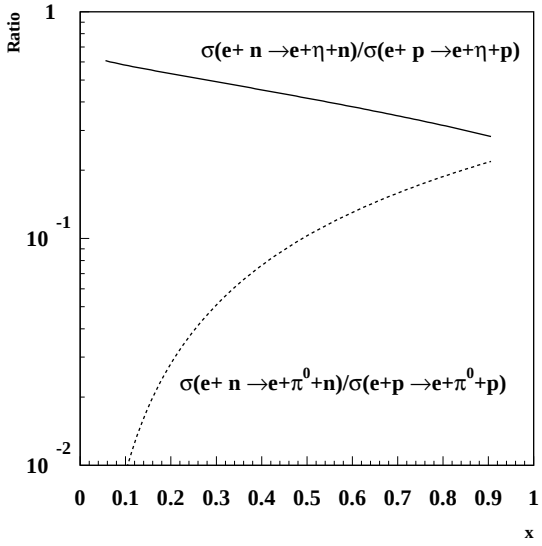

Figure 2.27: x dependence of the exclusive π^0/η production ratio on a neutron versus proton target for $Q^2 = 20$ GeV^2. Figure from Reference [63].

2.6.5 Parametrization of GPDs

This section will cover some models of the GPDs. First the general ansätze will be covered and then some specific models will be discussed. Theoreticians have already tried to make various

parametrization models of GPDs. On the one hand the limits as mentioned in section 2.6.1 are used as an input, while on the other hand, from impact parameter space transformation other constraints for the GPDs are obtained [42].

To comply with the polynomiality property of the GPDs the Double-Distribution (DD) functions often are introduced. A detailed discussion of the double distributions can be found in Reference [65]. These functions, $f(\beta,\alpha,t)$ are related to the GPDs by:

$$F^q(x,\xi,t) = \int_{-1}^{1} d\beta \int_{-1+|\beta|}^{1-|\beta|} d\alpha\ \delta(x-\beta-\alpha\xi)\ f(\beta,\alpha,t)\ , \tag{2.112}$$

where F^q is the related GPD. These DD functions always comply with the polynomiality condition but always have $h^{q(N)}_{N+1)} = e^{q(N)}_{N+1} = 0$. The parametrization thus is not complete. It can be completed by adding the so-called D-term [54]:

$$\begin{aligned} H^q(x,\xi,t) &= \int_{-1}^{1} d\beta \int_{-1+|\beta|}^{1-|\beta|} d\alpha\ \delta(x-\beta-\alpha\xi)\ F^q(\beta,\alpha,t) + \theta\left[1-\frac{x^2}{\xi^2}\right]\ D^q\left(\frac{x}{\xi},t\right)\ , \\ E^q(x,\xi) &= \int_{-1}^{1} d\beta \int_{-1+|\beta|}^{1-|\beta|} d\alpha\ \delta(x-\beta-\alpha\xi)\ K^q(\beta,\alpha,t) - \theta\left[1-\frac{x^2}{\xi^2}\right]\ D^q\left(\frac{x}{\xi},t\right)\ . \end{aligned}$$

The D-term generates the highest power of ξ for the n-th Mellin moment. Since for E^q and H^q the D terms are of equal magnitude, but have an opposite sign; they will vanish in the sum.

In a lot of models the Double Distributions are factorized as:

$$f(\beta,\alpha) = q(\beta)h(\beta,\alpha), \tag{2.113}$$

with $q(\beta)$ the ordinary quark distributions and $h(\beta,\alpha)$ a profile function.

A model of particular importance for the work presented in this thesis is the parametrization by Vanderhaeghen, Guidon and Guichal (VGG) [58]. This model is a direct phenomenological parametrization with only little underlying theory of the structure of the nucleon. This model assumes that the t dependence factorizes. This ansatz is only valid at small t. The VGG model consequently is only valid at small t. The model uses the following parametrizations:

$$\begin{aligned} H^{u[d]}(x,\xi,t) &= u(x)[d(x)]\, F_1^{u[d]}(t)/2[\cdot 2] \\ \tilde{H}^{u[d]}(x,\xi,t) &= \Delta u_V(x)[\Delta d_V(x)]\, G_A^{u[d]}(t)/G_A^{u[d]}(0) \\ E^{u[d]}(x,\xi,t) &= u(x)[d(x)]\, F_2^{u[d]}(t)/2[\cdot 2] \\ \tilde{E}^u = -\tilde{E}^d = \frac{1}{2}\tilde{E}_{\pi-pole} &= \Theta\left(-\xi \le x \le \xi\right)\ G_P(t)\ \frac{1}{\xi}\ \Phi_{as}\left(\frac{x}{\xi}\right)\ , \end{aligned} \tag{2.114}$$

where $G_{A(P)}$ are the axial (pseudoscalar) form factors, and $F_{1(2)}(t)$ the Dirac (Pauli) form factor. Φ_{as} is the meson distribution amplitude. The Θ function in the expression for $\tilde{E}$ ensures that it the latter function is only defined in the ERBL region. The model has been implemented in the HERMES Monte Carlo chain for generating exclusive pions. In this model the $\tilde{E}$ dependence for hard π^0 production is neglected, while $\tilde{E}$ is dominated by the pion pole in the case of π^+ production. Reference [58] also reports corrections to the leading order amplitude given before (equations 2.101 to 2.104). The correction due to the intrinsic

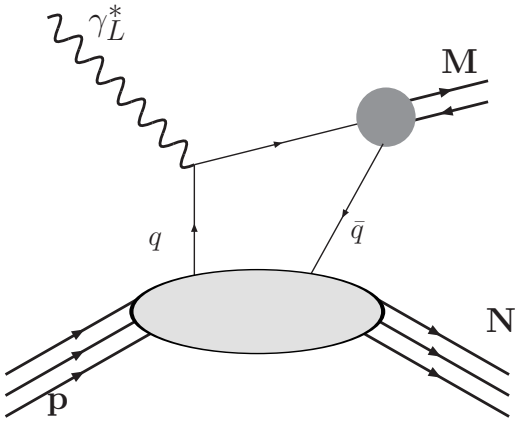

Figure 2.28: The 'soft overlap' diagram, concurring with the hard exclusive meson production at lower Q^2 values.

transverse momentum of quarks has been calculated as well as the soft overlap contribution. This mechanism can compete with the leading order amplitude at lower Q^2 values and is drawn in Figure 2.28.

Another GPD model, based on similar ansätze as the VGG model is the one described in Reference [59]. This model assumes the following factorization for the double distribution:

$$\Delta F(x, y; t) = h(x, y)\, \Delta q(x)\, f(t), \tag{2.115}$$

The t-dependence is governed by the form factor $f(t)$ which is taken to be equal to the nucleon pseudovector form factor:

$$f(t) = \frac{1}{(1 - t/\Lambda^2)^2}. \tag{2.116}$$

The $\tilde{E}$ dependent part is parametrized using the pion distribution amplitude:

$$\left(\Delta K^{du} + \Delta \bar{K}^{du}\right)(x, y, t \approx 0) \;=\; -\frac{2\sqrt{2} f_\pi M g_{\pi NN}}{m_\pi^2 - t}\, \delta(x)\, \Phi_\pi(y). \tag{2.117}$$

Figure 2.26 shows the π^0/π^+ production ratio based on this model for $Q^2 = 10$ GeV^2.

This ends the discussion on GPDs and at the same time the theoretical introduction into nucleon structure. Both unpolarized and polarized deep inclusive scattering and their relation to nucleon structure was presented. After a short description of semi-inclusive DIS, a description of the interesting framework of GPDs followed. The next chapters discuss the experimental tools required to measure GPDs. Chapter 7 discusses the analysis of exclusive π^0 production and the ratio of π^0 and η production on a deuterium and hydrogen target as a test for the predicted neutron suppression.

CHAPTER 3

The HERMES Experiment

The HERMES experiment is located at the DESY laboratory in Hamburg, Germany. The accelerator facility available there is a lepton-proton collider called HERA. The HERMES experiment makes use of the lepton beam only. The lepton beam consists out of either electrons or positrons. These two terms will be used alternately in this thesis when referring to the HERA lepton beam. In HERMES, leptons are scattered off an internally polarized gas target, and detected in a forward spectrometer. The accelerator facility, target and spectrometer will be covered in detail, followed by a discussion on luminosity, background and data acquisition.

3.1 The HERA accelerator

3.1.1 The HERA facility

The 'Hadron Electron Ring Accelerator' HERA is a storage ring of 6.3 km circumference. It is the only lepton-proton collider in the world. Moreover, it has the unique feature that both electrons and positrons can be accelerated and stored. Lepton energies up to 27.5 GeV and proton energies of 920 GeV are achieved. The proton beam makes use of superconducting magnets to focus and steer the beam. The most important beam properties of the HERA facility are given in Table 3.1. There are two interaction points (IP) where electrons and protons collide, namely in the north and in the south. At these two spots the experiments H1 and ZEUS can be found. HERA was built between May 1984 and November 1990. On October 19th 1991 the 1st electron-proton collision was observed. From 1995 on, HERMES has been taking data in the East Hall, making use of the electron beam only. Between 1999 and February 2003, Hera-B took data using the proton beam only. During 2001 and 2002 HERA was upgraded with an improved focusing at the collision points, and a better transfer line between the different pre-accelerators. The facility, as it was after 2002, is often referred to as HERA II.

	HERA			
Circumference	C_H	6335.83 m		
Revolution Frequency	f_0	47.317 kHz		
Number of buckets	N_b	220		
Number of colliding bunches	N_{cbun}	174		
	HERA-e		HERA-p	
Energy	E_e	27.5 GeV	E_p	920 GeV
Total Current	I_e	40 mA	I_p	90 mA
Number of bunches		189		180
Single Bunch Current	I_{be}	$211\mu A$	I_{bp}	$500\mu A$
Bunch Population	N_e	$2.79 \cdot 10^{10}$	N_p	$6.59 \cdot 10^{10}$
Bunch Length (RMS)	σ_{ez}	10.3 mm	σ_{pz}	191 mm
Bunch Time Length	$\sigma_{et} = \sigma_{ez}/c$	34.3 ps	$\sigma_{pt} = \sigma_{pz}/c$	637 ps
Peak Current	$\frac{N_e q}{\sqrt{2\pi}\sigma_t}$	52 A	$\frac{N_p q}{\sqrt{2\pi}\sigma_t}$	6.6 A

Table 3.1: Some properties of HERA taken from Reference [66].

Before particles enter the HERA accelerator, they pass through a chain of various pre-accelerators located on the DESY main site as can be seen in Figure 3.1. Lepton beams are produced and accelerated by Linac II, after which they are accumulated in the Positron Intensity Accumulator PIA, and finally injected into DESY II where an energy of maximum 7.0 GeV can be reached. From DESY II leptons can be extracted to DORIS, which is a source for synchrotron light, or to PETRA II, HERA's main pre-accelerator. In fact PETRA itself was the main accelerator at DESY before the construction of HERA. The major discovery of gluons at the DESY lab was achieved by using beams provided by PETRA. From summer 2007 on PETRA will be upgraded to PETRA III, serving as synchrotron light source for various X-ray experiments. From that moment HERA will no longer be operated. The proton beam on the other hand is produced by strip injection of H^- beams produced by LINAC 3 into DESY III, followed by the injection into PETRA.

A typical beam history can be seen in Figure 3.2. The lifetime of the electron beam is typically 12-16 hours, the proton beam has a much longer lifetime. In Figure 3.2 a steep decrease in electron current and lifetime can be observed around 18:00, which is due to an increased target gas density at the HERMES interaction point.

3.1.2 Polarization

HERA's lepton beam has the unique feature of being polarized. A polarized beam is important to perform double spin asymmetry measurements, which are useful to measure for example the structure function g_1. This section will detail the polarization of the lepton beam.

3.1.2.1 Spontaneous polarization

Particles in an externally applied magnetic field will have their spin either parallel or anti-parallel to that field. In a storage ring the particles will thus have their spin transversely polarized.

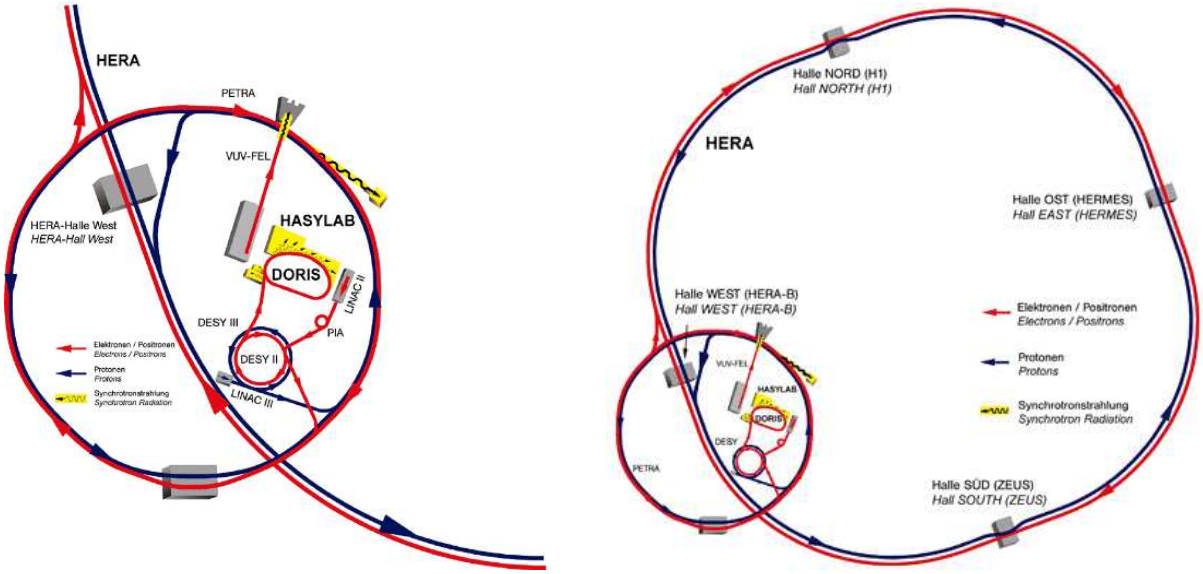

Figure 3.1: Schematic view of the HERA accelerator facility and its chain of pre-accelerators, indicated on the left half of the figure.

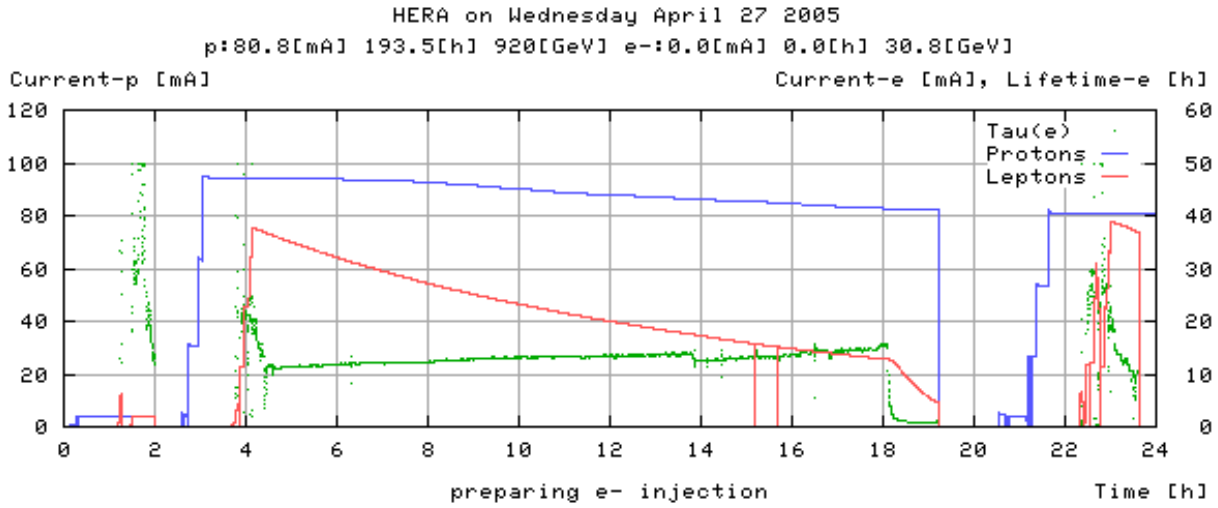

Figure 3.2: Typical beam history of the electron and proton beam in HERA. The upper curve corresponds to the proton beam current with the scale indicated on the left. The curve in the middle is the electron beam current, whose scale is absent in the figure. The lower curve shows the electron beam life time, the scale for the latter is depicted on the right. At 18:00 the lifetime dropped from about 15 hours to 2 hours due to the injection of high density gas in the HERMES target cell.

Initially, there will be as many particles with spin up $N^0(\uparrow)$ as with spin down $N^0(\downarrow)$, meaning we have an unpolarized beam. However, it is known that synchrotron radiation can introduce a spin flip. The flip probability $\mathcal{P}$ is dependent on the spin orientation:

$$\mathcal{P}(\uparrow\rightarrow\downarrow) \neq \mathcal{P}(\downarrow\rightarrow\uparrow)$$

After a time t the number of particles with spin up, $N^t(\uparrow)$, is different from the number of particles with spin down, $N^t(\downarrow)$:

$$N^t(\uparrow) = N^0(\uparrow) \times \mathcal{P}(\uparrow\rightarrow\downarrow) \neq N^0(\downarrow) \times \mathcal{P}(\downarrow\rightarrow\uparrow) = N^t(\downarrow)$$

The polarization P can now be defined as the difference of the number of particles with spin orientation $\uparrow$ from the number of particles with spin orientation $\downarrow$:

$$P = \frac{N(\uparrow) - N(\downarrow)}{N(\uparrow) + N(\downarrow)}$$

This spontaneous polarization effect is known as the Sokolov-Ternov effect [67]. The qualitative expression for the polarization built up in an initially unpolarized beam is given by [68]:

$$P(t) = \frac{8}{5\sqrt{3}}(1 - e^{-\frac{t}{\tau_{ST}}}), \tag{3.1}$$

with τ_{ST} for a circular ring given by:

$$\tau_{ST} = \frac{8}{5\sqrt{3}} \frac{2\pi m_e}{r_e h} \frac{\rho^3}{\gamma^5}, \tag{3.2}$$

where m_e (r_e) is the electron mass (Bohr radius), ρ the radius of the circular ring, and $\gamma = \frac{1}{\sqrt{1-(\frac{v}{c})^2}}$. The maximum theoretical polarization is about 92%. At HERA where $\rho = 603$ m and $E = 27.5$ GeV the built up time τ is about 25 minutes. For the LEP collider at CERN the build up time was about 10 hours [69].

Evidently, also depolarizing effects have to be taken into account when calculating beam polarizations. These can be caused by beam-beam interactions, depolarizing resonances or misalignment of the magnets. The depolarizing effects can be characterized by a single time constant τ_D such that the effective polarization P_{eff} can be written as:

$$P_{\text{eff}} = \frac{8}{5\sqrt{3}} \frac{\tau_D}{\tau_{ST} + \tau_D} \tag{3.3}$$

$$\tau_{\text{eff}} = \tau_{ST} \frac{\tau_D}{\tau_{ST} + \tau_D} \tag{3.4}$$

leaving a lower polarization with a faster rise time.

As explained before, the spontaneous polarization occurs in a transverse direction, whereas interesting physical observables require longitudinally polarized lepton beams. Therefore, the transversely polarized beam is rotated just in front of the experimental area. Evidently, the polarization needs to be turned back into the transverse direction right after the experimental area to keep profiting from the Sokolov-Ternov effect. Therefore, two sets of spin rotators have been implemented before and after the experimental north, east and west halls, as shown in Figure 3.3. The spin rotators close to H1 and ZEUS were only implemented after the upgrade to HERA II. The rotation is achieved by six dipole magnets, bending the beam three times in horizontal and three times in vertical direction. The principle of operation behind these dipole spin rotators is the Thomas-BMT equation [70], stating that for motion transverse to the magnetic field, the spin precesses around the field at a rate which is $\frac{g-2}{2}\gamma$ times faster

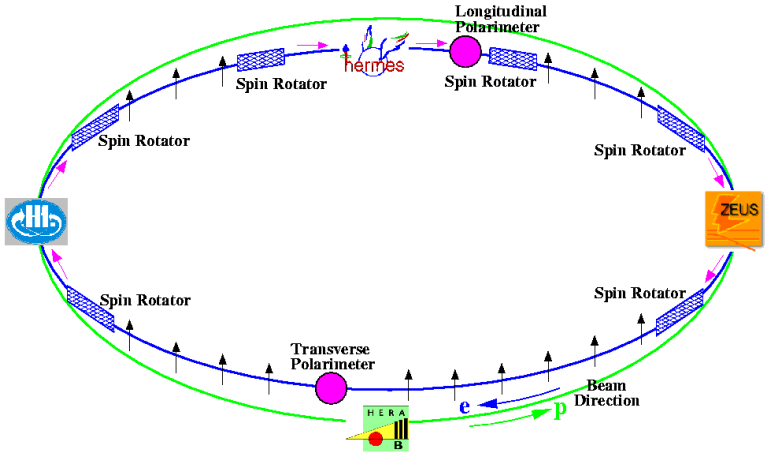

Figure 3.3: Schematic view of the HERA ring, with its 3 pairs of spin rotators, the LPOL and TPOL setup.

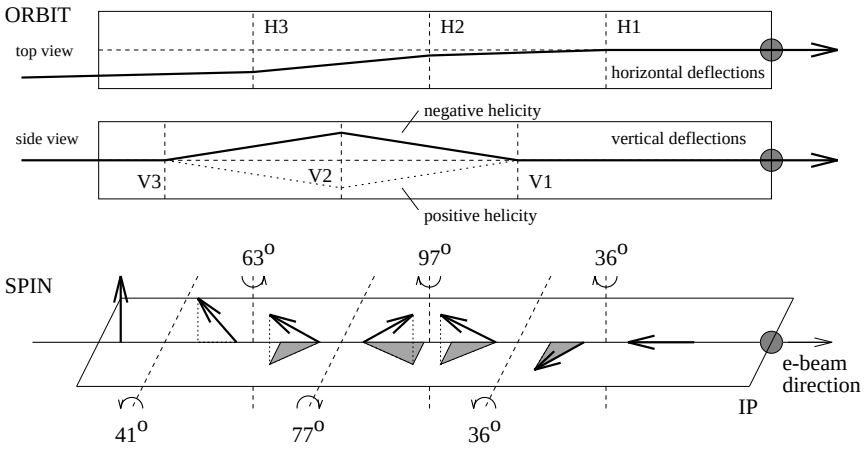

Figure 3.4: Schematic view of a spin rotator, consisting of 3 horizontal and 3 vertical bending dipole magnets. The length of the spin rotator is 56 m.

than the rate of rotation of the orbit direction. At HERA $\frac{g-2}{2}\gamma$ is between 60 and 80. The spin rotator is schematically depicted in Figure 3.4.

After the upgrade to HERA II, a larger beam-beam force was measured at both IPs resulting in lower polarization values. Moreover, due to stronger magnets in the interaction points in the north and south halls, the sensitivity to magnet misalignment has increased. The installation of two additional spin rotator pairs further decreased the maximum possible polarization. Values up to 50 % have been reached with the three spin rotators, compared to values up to 70 % before the upgrade to HERA II. In general the polarization is lower than 50 % as can be seen in Figure 3.5, where also the rise in polarization in the beginning of a fill can be seen.

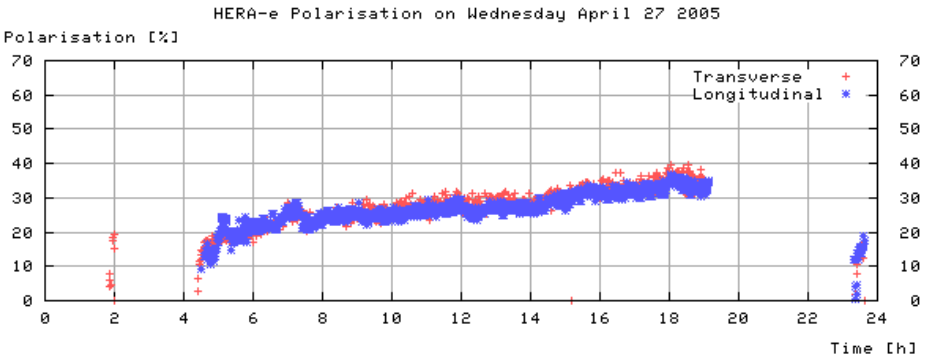

Figure 3.5: Polarization as measured by the TPOL and LPOL of the HERA e-beam.

3.1.2.2 Measuring the polarization

The polarization should be monitored continuously as it is subject to rapid change. At HERA, the polarization is measured by a transverse and a longitudinal polarimeter. The longitudinal polarimeter (LPOL) is situated close to the east hall, the transverse polarimeter (TPOL) is located in the west. Both polarimeters use Compton scattering as an underlying principle to measure the polarization. An asymmetry in the scattering of circularly polarized laser light off the beam is an indication of the beam polarization. In the case of the LPOL it is sufficient to measure the energy of the scattered photons, while the transverse polarization leads to an angular asymmetry only, so both the energy as well as the scattering angle need to be measured. The asymmetry can be measured by varying the polarization of the laser light from left to right. The ratio of LPOL/TPOL should obviously be one. More information about the polarimeters can be found in References [68], [71] and [72].

3.2 The internal polarized gas target

As HERMES is using a storage ring as beam source, the target density cannot be too high, otherwise the interaction rate would be so high that the beam current would drop substantially after a couple of revolutions. In order to comply with the primary physics goals of the HERMES experiment, namely measuring double spin asymmetries, both beam and target needed to be polarized. Since a gas is highly polarizable and has little dilution, the HERMES collaboration has decided to install an internal polarized gas target [73]. It has the additional advantage that the polarity of the gas can easily be flipped. Moreover, by choosing this target configuration, the measurement of recoiling particles is in principle possible. During 1996-97 longitudinally polarized hydrogen ($H_{||}$) was used as a target gas, between 1998 and 2000 the target cell was filled with longitudinally polarized deuterium ($D_{||}$), and between 2002 and 2005 transversely polarized hydrogen ($H_{\perp}$) were scattered off. In 1995 HERMES had a polarized ^{3}He target. This target setup will not be covered here, details can be found in Reference [74]. Additionally, also unpolarized gasses such as nitrogen, neon, krypton and xenon can be injected into the HERMES target cell.

The polarized target consists of an Atomic Beam Source (ABS), a target cell, a target gas analyzer TGA and a Breit-Rabbi Polarimeter (BRP). All these components are depicted in Figure 3.6. The most challenging requirement for the internal target is to determine the effective target polarization seen by the beam. More information about the HERMES polarized target in general can be found in Reference [75].

This section will detail the HERMES target region. First the polarized target will be discussed after which the unpolarized gas feeding system UGFS will be briefly mentioned.

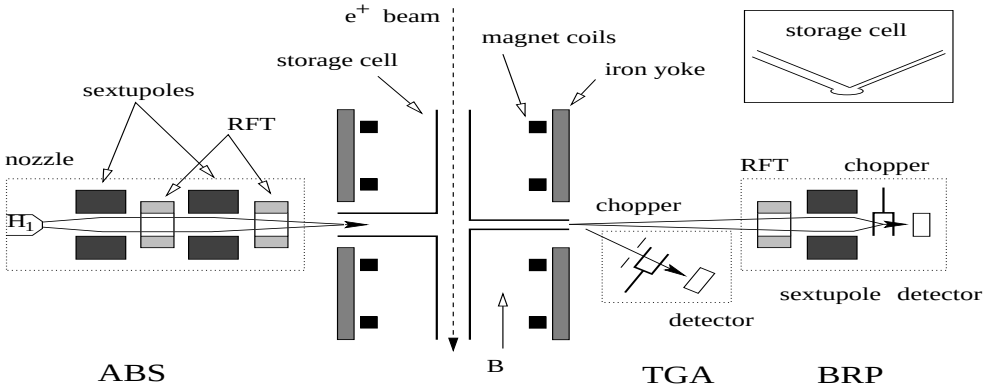

Figure 3.6: Schematic top view of the HERMES target. From left to right the Atomic Beam Source (ABS), the target cell inside the magnet, the Target Gas Analyzer (TGA) and the Breit-Rabi Polarimeter (BRP) can be seen. The locations of the radio-frequency transition units (RFT) are indicated as well. In the upper right a cross sectional view of the target cell is depicted.

3.2.1 Atomic beam source

The atomic beam source (ABS)[76] feeds polarized hydrogen or deuterium in the target cell. Molecular hydrogen (H_2) or deuterium (D_2) is dissociated by a radio frequency discharge of 13.56 MHz. To maximize the stability a few promille of oxygen is added to the molecular gas.

The atomic gas then flows further through a cooled nozzle at about 100 K, in order to freeze the water produced in the discharge. After the nozzle the gas expands in the dissociator chamber. A set of sextupole magnets focuses atoms with spin $+\frac{1}{2}$ according to the Stern-Gerlach principle, while atoms with opposite spin orientation are deflected. An additional set of sextupoles focuses the atomic beam into the entrance pipe of the target cell. Nuclear polarization is achieved by hyperfine transitions induced by RF-radiation. Polarization is achieved by three stages of high-frequency transitions: the strong (SFT), the medium (MFT) and weak (WFT) field transitions. In the case of hydrogen (deuterium) the MFT occurs at a frequency of 90 (25) MHz, the WFT at 14 (7) MHz and the SFT at 1430 (370) MHz. Polarizations of up to 97% (92%) for hydrogen (deuterium) have been reached. A powerful pumping system ensures low gas flow into the HERA beam vacuum.

3.2.2 Target chamber

The HERMES target chamber is shown in Figure 3.7. It consists of a storage cell, turbo molecular pumps to keep the pressure sufficiently low, a set of collimators and a wake field suppressor.

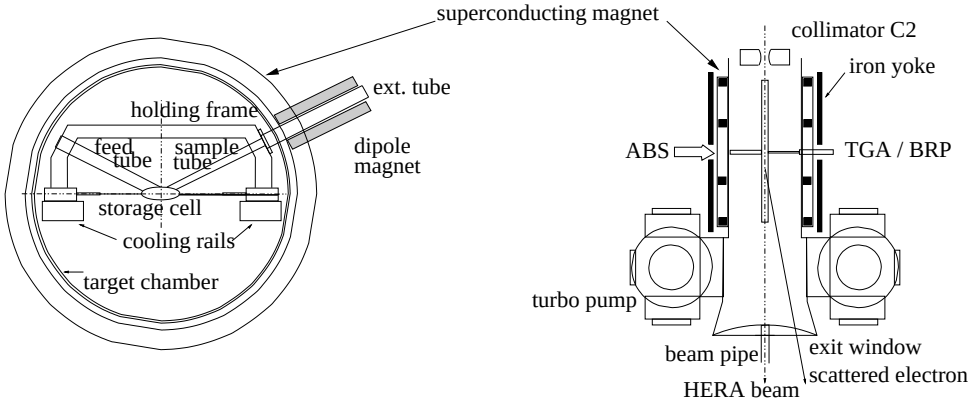

Figure 3.7: A cross sectional (left) and top (right) view of the HERMES target cell.

The HERMES storage cell [77] is made of $75\mu\mathrm{m}$ thin pure aluminum, is $400\ \mathrm{mm}$ long and has an elliptical shape. The cell is a bit longer in the downstream direction, where apertures are made for pumping. The aluminum is coated to minimize gas depolarization and recombination caused by wall collisions. In order to further decrease these effects the cell is cooled to a temperature of 100 K. Two side tubes are connected to the storage cell, one to inject polarized gas, the other to analyze the target gas. The latter tube is surrounded by a permanent dipole to prevent depolarization.

Heating of the target cell by the wake field of the beam should be prevented. Therefore, the target cell is connected to the beam vacuum pipe in the upstream as well as in the downstream direction via sixteen thin metal strips. These so called wake field suppressors ensure electrical connection between beam-pipe and target cell, while preventing thermal conduction between the two. Thermal contact should be avoided since the beam pipe is operated at room temperature. At the downstream end of the target cell, a $100\ \mu\mathrm{m}$ thick perforated titanium foil provides a smooth transition between the circular beam and elliptical target cell. The target cell together with its support structure and flange is shown in Figure 3.8.

The target chamber's exit window is $0.3\ \mathrm{mm}$ thick and made of stainless steel. It is kept as thin as possible in order to reduce multiple scattering as well as shower production from π^0 mesons (which decay in 2 photons) produced at small angles.

In order to reduce the amount of synchrotron light entering the target chamber and sequentially the spectrometer, a number of beam collimators were placed in front of the HERMES target. About a meter upstream of the target cell, an adjustable collimator C1 is located which is opened during injection and ramping of the beam. An additional fixed collimator C2 just upstream of the storage cell protects the cell from synchrotron light and showers originating from scattering from the C1 collimator.

For longitudinally polarized targets, the target chamber is put inside a superconducting magnet with four coils generating a uniform longitudinal field of up to 350 mT. This field

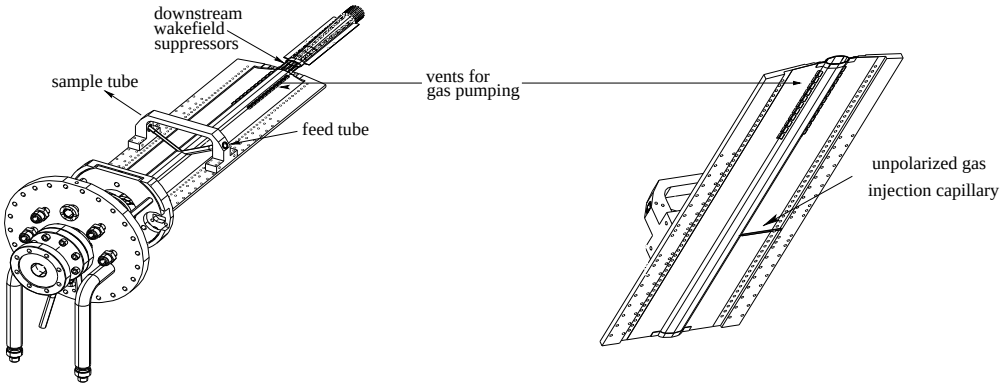

Figure 3.8: Schematic view of the HERMES target cell and its support structure. The downstream wake field suppressor can be seen as well as sample and feed tubes.

provides the quantization axis. The magnet and coils are indicated in Figure 3.7.

In the case of transversely polarized hydrogen, the target chamber was put in a transverse magnetic field, provided by a normal conducting magnet. The field was limited by the deflection of the beam due to the target magnet.

3.2.3 Target gas analyzer

In order to measure the atomic and molecular content of the gas extracted from the storage cell, a target gas analyzer (TGA) [78] was designed and developed. It is schematically shown in Figure 3.9. The TGA is mounted 7^o off-axis with respect to the Breit-Rabbi Polarimeter in order not to interfere with the latter. The TGA is built out of a pair of baffles, a chopper, and a Quadrupole Mass Spectrometer (QMS) with a beam ionizer and a Channel Electron Multiplier (CEM) for single ion detection. Particles in the TGA are ionized by 70 eV electrons, and subsequently mass filtered by the QMS before being detected by the CEM. The pair of baffles and the chopper precede the ionizing volume. The chopper rotates at a frequency of 5.5 Hz, blocking the sample beam periodically in order to allow subtraction of the residual gas signal. Both atomic and molecular gas fluxes can be measured. A typical signal of the TGA is shown in Figure 3.10.

3.2.4 Breit-Rabi polarimeter

The Breit-Rabi Polarimeter (BRP) [79] measures the polarization of the atomic target gas by precisely measuring the relative populations n_i of the hyperfine states of hydrogen or deuterium atoms. Nuclear polarizations can be directly related to the population density of the hyperfine states.

The BRP, schematically indicated in Figure 3.7 consists of two hyperfine transition units, a system of 2 sextupole magnets and a detector stage. The sextupoles (de)focus atoms with $m_s = (-) + \frac{1}{2}$. The hyperfine transition is established by a strong (SFT) and medium field transition unit (MFT) and allows for the exchange of hyperfine structure populations between various states. This permits selection of different combinations of initial hyperfine states reaching the detector. The detector stage is the same as for the TGA, constructed out of a

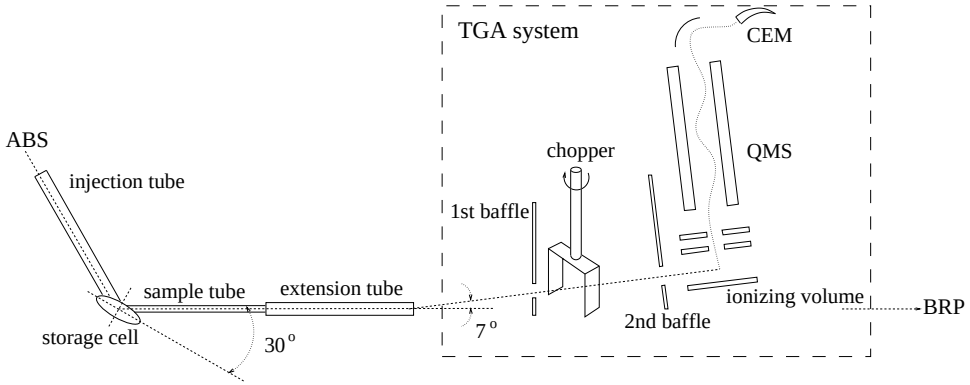

Figure 3.9: Schematic view of the TGA.

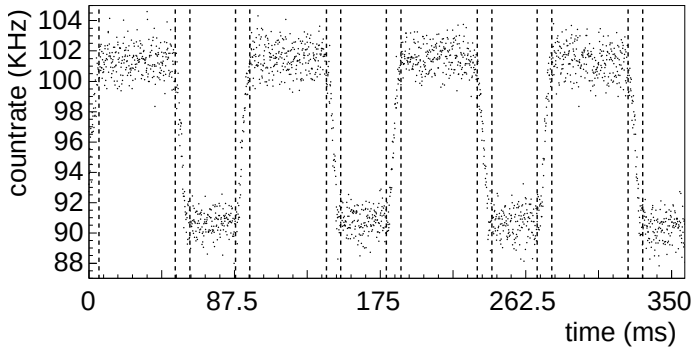

Figure 3.10: Time spectrum of the count-rate in the TGA. The vertical lines indicate undefined chopper positions. The gas flow determined by the TGA is the difference between count rates when the chopper is open and closed.

CEM and a QMS. Also here the QMS is preceded by a chopper for residual gas corrections. The only difference is that only hydrogen or deuterium atoms can be detected.

3.2.5 Unpolarized gas feeding system

Apart from a polarized target, also unpolarized gas can be injected into the HERMES target cell via the so-called unpolarized gas feeding system (UGFS). In Figure 3.8 the injection tube of the UGFS can be seen at the bottom side of the target cell. Various kinds of gasses have been used in HERMES : hydrogen, deuterium, krypton, xenon, neon and nitrogen. The gas densities can be chosen via a dosing valve. 'Low' densities mean a 45-hour lifetime contribution to the lepton beam lifetime, whereas 'high' densities mean a 2-hour lifetime contribution. The contribution to the beam lifetime τ_H can be calculated as: $\tau_H = \frac{\tau_0 \tau_G}{\tau_0 - \tau_G}$, where τ_G (τ_0) is the beam lifetime with (without) gas in the target cell. The consequence of 'high' density running to the beam lifetime and current is nicely illustrated in Figure 3.2, where at around 18:00 a change from low to high density target gas occurred.

3.3 The HERMES spectrometer

The HERMES spectrometer consists out of two identical parts, placed above and below the lepton beam pipe[1]. It is a forward spectrometer which is built out of tracking and particle identification detectors. The spectrometer contains a dipole magnet to determine the momentum of charged particles. Particles with scattering angles of ± 170 mrad in the horizontal direction and between 40 and 140 mrad in the vertical direction can be detected by the whole spectrometer. Since the original installation several upgrades were made mainly to increase the detector acceptance, such as the Lambda Wheels and the Wide Angle Hodoscopes. A detailed description of the spectrometer can be found in Reference [80]. This section will detail the building blocks of the spectrometer, which can be divided into three detector type groups: tracking chambers, particle identification detectors and triggering devices. References [81, 82, 83] could provide more insight in the physical principles behind certain detection techniques.

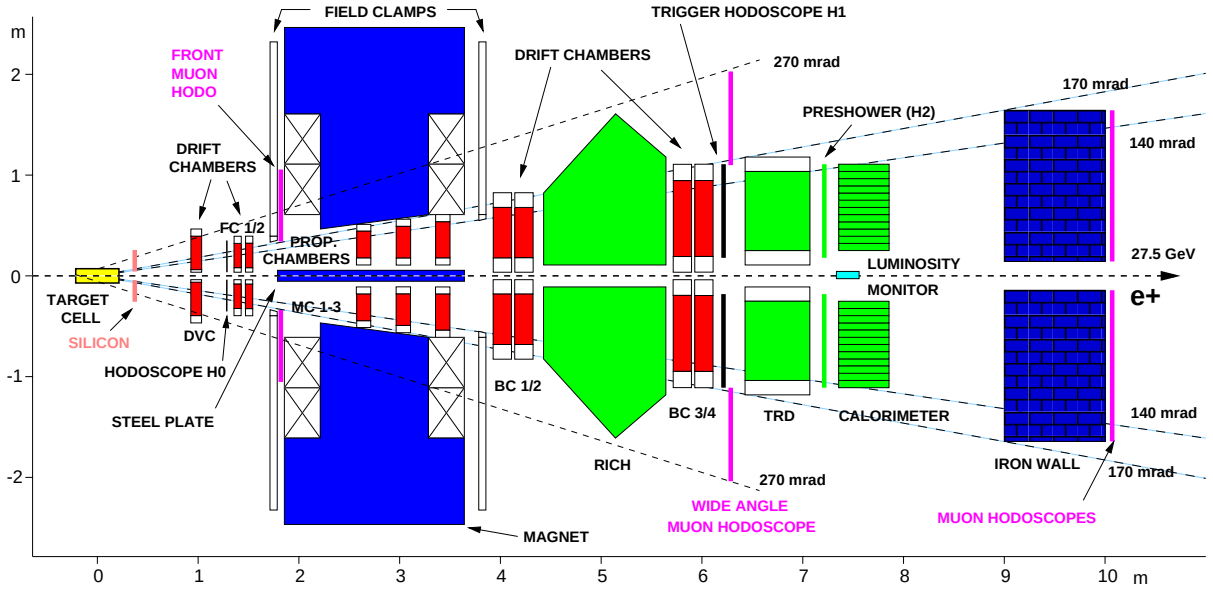

Figure 3.11: Schematic view of the HERMES spectrometer in the (y, z) plane. Tracking chambers are indicated in red, PID detectors in green, and triggering devices in black.

For the description of the HERMES spectrometer it is necessary to introduce the HERMES coordinate system. The origin of this reference system is located at the center of the target cell, with the z axis along the beam-pipe, pointing upstream. The y axis is chosen vertically, and the x axis set to have a right handed coordinate system.

The spectrometer magnet is operated at a deflecting power of 1.3 T·m. A massive iron plate located in the symmetry plane and called the septum plate shields both lepton and proton beam from the magnetic field. Furthermore, a correction coil with a deflecting power of 0.8 T·m was installed inside the shielding of the lepton beam to correct for fringe fields

[1]The proton beam pipe also passes in between both detector halves at a distance of 72 cm from the lepton beam.

and imperfections of the shielding. The stray field at the tracking chambers right in front and behind the magnet does not exceed 0.1 T.

3.3.1 Tracking

The main purpose of the tracking system within HERMES is twofold: first of all it needs to measure the track of a particle in 3-dimensional space, and secondly, it needs to be able to define the angle of deflection by the magnetic field in order to determine charged particles' momenta.

All HERMES tracking devices can be divided into two categories depending on the spatial acceptance of the devices: main tracking is coming from drift chambers, identified as the Drift Vertex Chambers (DVCs), Front Chambers (FCs) and Back Chambers (BCs), while extended tracking is achieved by wire chambers, called the Magnet Chambers (MCs) and a silicon tracker referred to as the Lambda Wheel (LW). Previously an additional tracking device was installed between the target and the DVC, namely the vertex chamber. This device has been removed and will not be covered here. All other tracking detectors will be covered in the next paragraphs.

3.3.1.1 Drift Chambers

The core of the HERMES tracking is made by the 2 FCs [84] and the 4 BCs [85] . They are conventional horizontal drift chambers. Each chamber has 6 layers of drift cells consisting out of a plane of alternating anode and cathode wires between a pair of cathode foils. The cathode foils and wires are at high voltage, the anode wires at ground. The cells are organized in 3 pairs, a vertical pair - the X plane, and U and V planes which are at an angle of $\pm 30^0$ with respect to the vertical plane. The 2 planes of one pair are staggered by half the cell size in order to help resolve ambiguities. A BC module is shown in Figure 3.12, next to a schematic of the field lines in a cell.

When charged particles traverse the chamber they ionize the gas. Electrons from the gas molecules will move to the anode wires, thereby further ionizing gas molecules resulting in an avalanche of positive ions which are attracted to the cathode wires. Due to the high mobility of electrons, their drift time across the cell can be used to determine where in the cell a particle passed.

The choice of gas mixture was made to cope with the strict safety rules concerning flammable gasses in the HERA tunnel. A mixture of 90% Argon, 5% CO_2 and 5% CF_4 was chosen.

Readout cards responsible for amplification, pulse shaping and discrimination are mounted on the drift chambers. The signals are transported to Fast-Bus TDCs in the electronic trailer, which is 30 m away.

The BCs determine the partial track behind the magnet. The chambers also provide the hit cell of the particle identification detectors located in between and behind the BCs. The size of these chambers is already quite large so the mechanical design needed special attention with respect to uniformity and stability. BC 1 is located in the fringe field of the spectrometer magnet. Simulations have shown that a vertical field of 0.1 T yields a deviation of less than 60 μm.

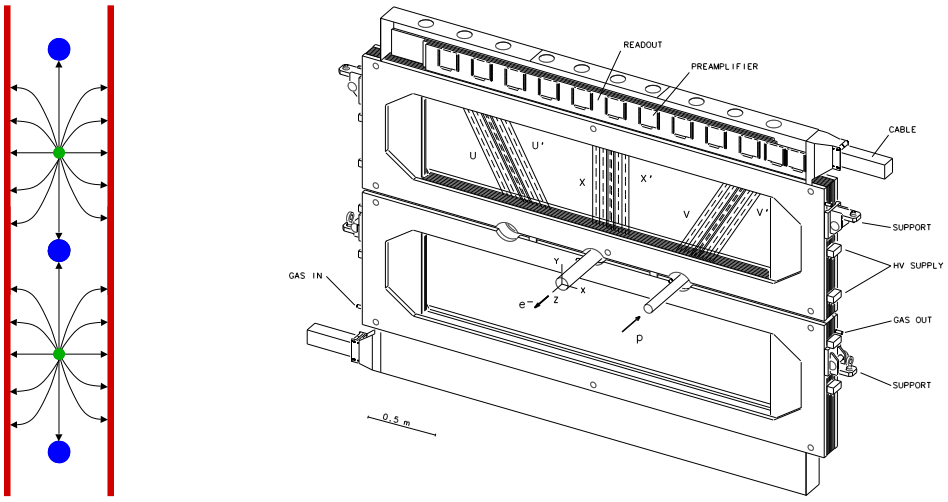

Figure 3.12: On the left: the field lines in a drift chamber cell. The thin anode wires, thick cathode wires and the cathode plane are visible, together with the field lines. Right: a schematic drawing of a BC module.

The average BC efficiency is above 99% for electron tracks, and around 97% when all particle tracks are taken into account, since hadrons deposit less energy in the chambers. A dependence of the efficiency on the atmospheric pressure has been observed. To avoid this, a system has been implemented which changes the high voltage to the BC according to the atmospheric pressure. This is illustrated in Figure 3.13, where also the resolution as a function of drift distance is shown. The 128 (192) wires of each BC1/2 (BC3/4) plane account for an integrated resolution of 275 (300) μm. Each FC plane contains 96 wires, yielding a resolution of 225 μm

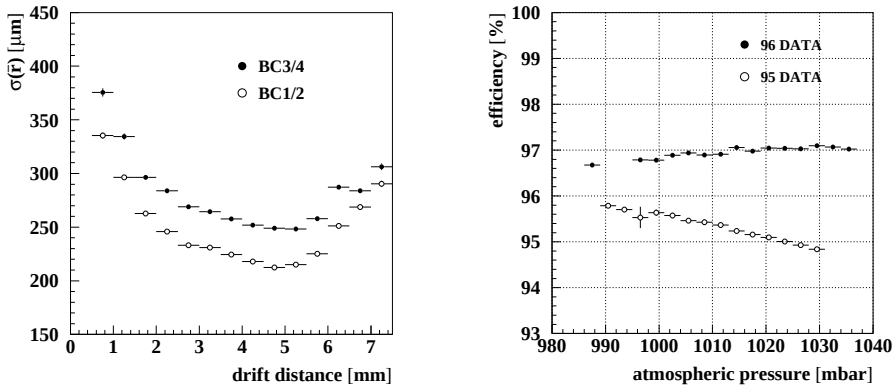

Figure 3.13: Left: the BC resolution as a function of drift distance. Right: the averaged plane efficiency as function of the atmospheric pressure.

The DVCs were installed later since redundancy in the front region was desired. Loss of a single FC plane could have had a large impact on HERMES tracking. The mechanical design as well as the readout is similar to the other drift chambers. The resolution is about 220 μm.

3.3.1.2 Magnet Chambers

The Magnet Chambers (MCs) [86] are Multi Wire Proportional Chambers (MWPCs), located inside the magnet. Initially they were build to help resolve multiple track events. However, low background rates have overcome this need. The MC's primary goal now is to identify tracks from low momentum particles, whose bending angle is too large to be detected by the BCs. These tracks are called 'short tracks' within the HERMES jargon. Pions from decaying Λs are an example of such particles. The MCs are, together with the FCs capable of fully identifying these particle tracks.

The operation inside a strong magnetic field of about 1 T automatically has led to the choice of MWPC for the MCs. These chambers have high efficiencies and are capable of dealing with high multiplicities. The positioning of the MCs is shown in Figure 3.14.

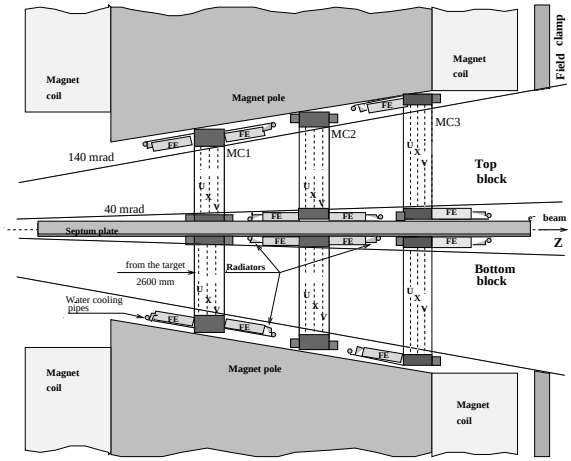

Figure 3.14: The positioning of the 6 MCs inside the magnet.

There are 6 MCs, each built out of 3 pairs of U, V, and X planes. The limited space as well as the requirement of preserving the vertical angular acceptance of the spectrometer imposed severe constraints on the MC design. To comply with these constraints, the frame width needed to be reduced inducing substantial forces in the modules, which complicates their design.

The chamber's gas mixture is the same as for the drift chambers, but the mixing ratio's have been optimized: 65% Argon, 30% CO_2 and 5% CF_4. The front end readout electronics are mounted on the chambers and are based on the LeCroy PCOS4 boards.

As there is poor air circulation inside the magnet gap, and the front end electronics's power consumption is substantial, a powerful cooling system was needed. An additional requirement to the cooling is that it also needs to keep the electronics above the dew point to avoid water condensation.

The right panel of Figure 3.15 shows the capabilities of the HERMES tracking for 'short' as well as for 'long' tracks. For short tracks the tracking is clearly better when also the DVCs are used.

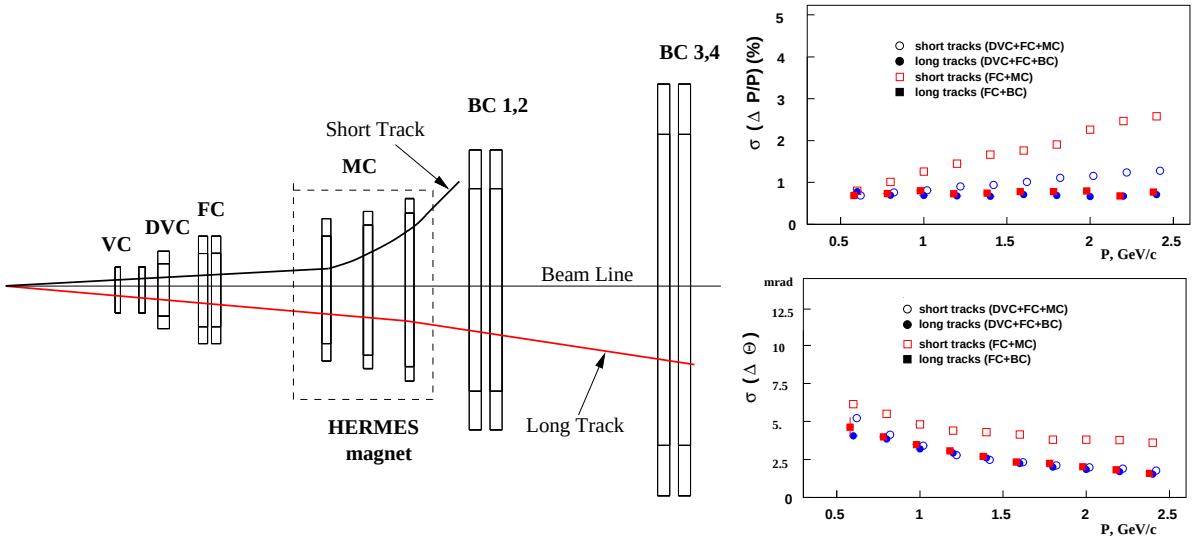

Figure 3.15: Left: a 'short' and a 'long' track in the HERMES spectrometer (top view). The Vertex Chambers (VC) are still in the picture. Top (bottom) right shows the momentum (angular) resolution of the HERMES tracking.

3.3.1.3 Lambda Wheel

After the 2000 shutdown an additional tracking device, the Lambda Wheel [87, 88], was installed to improve the detection of Λ and J/Ψ events. The main detector goal is to measure the momentum direction of a pion coming from a $\Lambda-$decay, and consequently allowing the reconstruction of a secondary vertex. Obviously, the detector needs to be positioned a couple of $\Lambda-$decay lengths away from the target cell if the pion is to be detected. As the $\Lambda-$decay length ($= \tau/c$) is about 7.9 cm, the position of the Lambda Wheel was chosen to be between $z = 45$ cm and $z = 50$ cm. Since pions from $\Lambda-$ decay have small momenta, the detector needed to be placed inside the beam vacuum. Silicon was chosen as active material, because of its fine granularity and vacuum properties.

The Lambda Wheel is build out of 12 modules, each consisting out of two trapezoidal double sided 300 μm thick silicon strip sensors. The strips are oriented parallel to the oblique side of the trapezoid, allowing all connections to be made from the top. The available radius is limited by the pumping cross. The radius of the Lambda Wheel was designed to be about 17 cm of which the inner 4 already are covered by the spectrometer acceptance. The modules are read out by on board HELIX chips, which were designed for use at the HERA experiment HERA-B. These chips will be extensively discussed in section 5.4.2. Sensors and electronics are cooled by ethanol.

As the Wheel is placed close to the accelerator, special care needed to be taken to screen the detector from the RF field generated by the beam. Therefore the pumping openings in the wake field suppressor were covered with a small mesh gauze. Additionally, an RF screen was installed between the Wheel and the target chamber, as depicted in Figure 3.16.

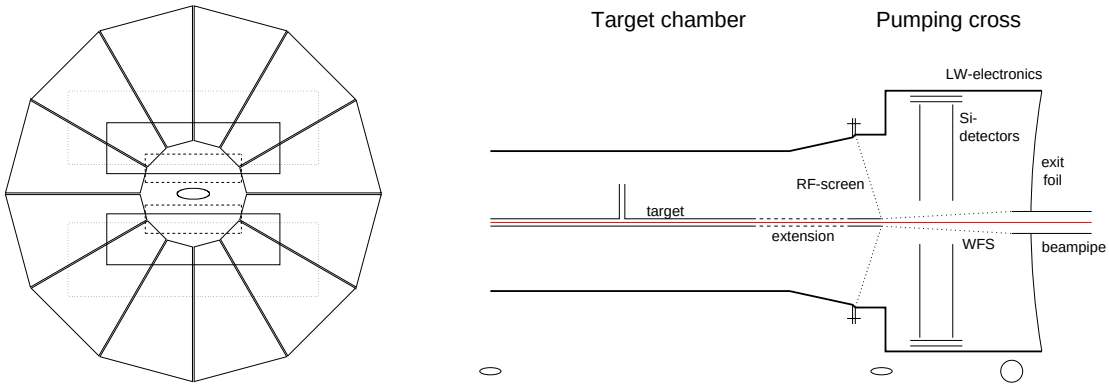

Figure 3.16: Left: Front view of the Lambda Wheel at $z = 45$ cm. The solid (dotted (dashed)) box indicates the HERMES acceptance at $z = 0$ cm ($z = 20$ ($z = -20$)). The right side of the figure shows a side view of the Lambda Wheel within the lepton beam vacuum. The two silicon layers of the Lambda Wheel are clearly visible, together with the readout electronics on top. The target cell and wake field suppressor (WFS) as well as the RF-screen are indicated in the figure.

3.3.2 Particle identification

Particle identification within HERMES is achieved in two steps: first electrons and hadrons are separated with a Transition Radiation Detector (TRD), and an electromagnetic calorimeter preceded by a preshower detector. Secondly, pions, kaons and protons are identified by the response of a Ring Imaging Cherenkov Detector (RICH). The RICH was installed in 1998. Before that, a conventional Čerenkov detector was used to identify pions and kaons. First the TRD and preshower-calorimeter combination are detailed, and then properties of the RICH are discussed.

3.3.2.1 Transition radiation detector

In order to achieve a sufficiently large electron-hadron separation, the installation of a detector with a pion rejection factor of about 100 was required. Because of the limited available space the technique of transition radiation was chosen: when a particle of sufficiently high energy traverses the boundary between two materials with different index of refraction it emits radiation. At HERMES energies only electrons emit such transition X- rays. The detection of an X-ray in coincidence with a charged particle thus can discriminate between electrons and hadrons. The detection of transition radiation is difficult, since it is emitted in the forward direction.

The HERMES TRD consists of a sequence of 6 modules. Each of these consist of radiator material and a proportional chamber. The radiator is a two-dimensional matrix of fibers, as the construction of a foil implied a lot of difficulties. The radiators are 6.35 cm thick. The proportional wire chamber is built out of 256 vertical wires in a gas mixture of 90% Xe and 10% CH_4. The mixture was optimized for efficient X-ray absorption. Signals from the wires are amplified and transmitted to Fast-Bus ADCs. The TRD is schematically drawn in Figure 3.17.

By ionizing of the radiator gas, electrons and pions will deposit a comparable amount of energy in the TRD, making the discrimination between both particle types hard. Figure 3.18

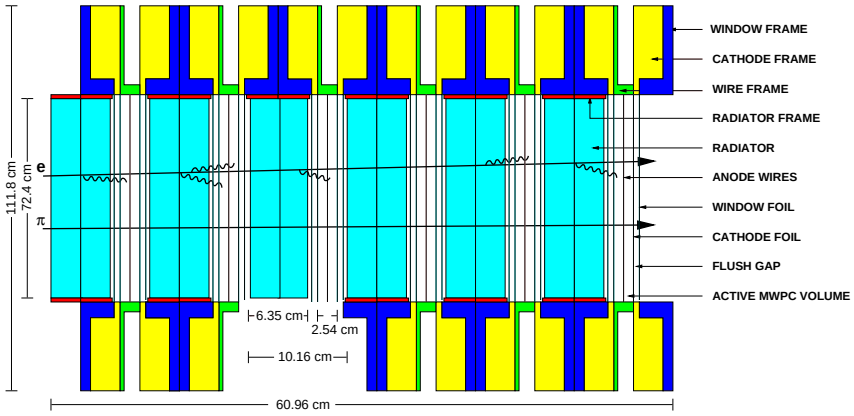

Figure 3.17: Schematic overview of the 6 modules of the HERMES TRD. Each module is built out of a radiator and a proportional chamber.

shows the energy loss in the TRD for pions and protons. Electrons deposit about 50% more energy than hadrons. Both distribution have large tails due to the production of knock-on electrons (δ-rays). When the transition radiation by the electrons is taken into account, a substantial difference in energy deposition can be seen between both particle types. In order to obtain a good hadron rejection factor, data from several modules need to be combined. A functional combination was found in the 'truncated mean' method: the largest signal from the 6 modules is abandoned while the average of the other five is taken. The method is illustrated in Figure 3.19. It yields a good pion rejection factor of above 100. By using a probability based method, this factor is further increased. Section 4.3 will detail this method.

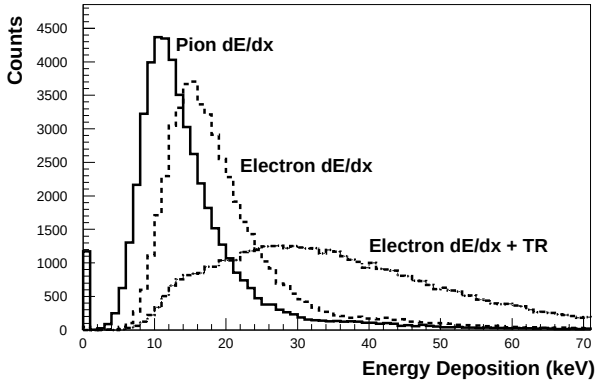

Figure 3.18: Energy deposition in one TRD module. The contributions of hadrons and leptons are shown.

3.3.2.2 Preshower-calorimeter

The HERMES calorimeter [89] was designed to perform an electron-pion separation of at least 10 at a first level trigger, and > 100 in the offline analysis. Moreover, the calorimeter needed to provide a first level trigger on scattered electrons itself, and a coarse position measurement

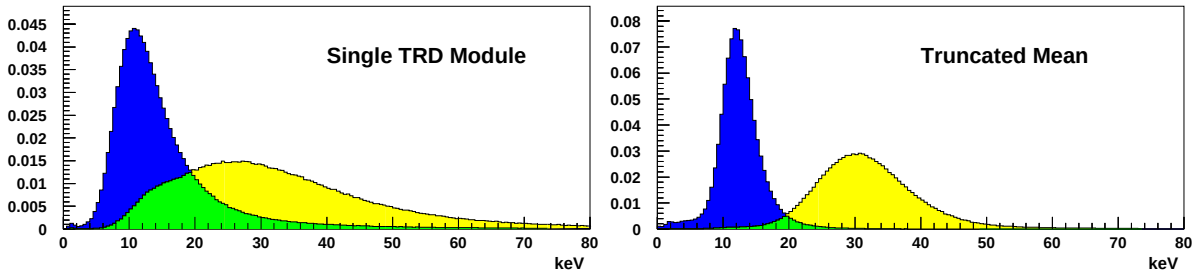

Figure 3.19: Illustration of the 'truncated mean' method. Left: the response of one TRD module. Right: the truncated mean of all 6 modules.

for these electrons. To cope with the design goals, a preshower has been installed in front of the calorimeter. These two detectors are reported on simultaneously in this section.

When discussing calorimetry an important variable is the radiation length X_0, indicating the length after which an electron still has $\frac{1}{e}(\sim 37\%)$ of its energy, it thus has lost $\sim 63\%$. This length can be approximated by [83]:

$$X_0 \sim 180 \cdot \frac{A}{Z^2}, \tag{3.5}$$

expressed in units [g/cm^2]. Another parameter is the critical energy E_c at which radiative processes start to dominate the energy loss.

In order to describe the energy loss the mean free path length λ is introduced. The mean free path for hadrons can be approximated as [83]:

$$\lambda \sim 35 \cdot A^{1/3}, \tag{3.6}$$

in units [g/cm^2].

An electron with an energy larger than the critical energy E_c will lose its energy primarily via the emission of a Bremsstrahlung photon. Bremsstrahlung photons in turn have a sufficiently high energy to convert into electron-positron pairs, which emit Bremsstrahlung themselves. A shower of electromagnetic radiation is consequently established until the energy of the electrons and positrons in the shower falls below the critical energy E_c.

The energy of the electrons and positrons in the shower is sufficiently high to create Čerenkov light. The Čerenkov photons can be detected by PhotoMultiplierTubes (PMTs). The electromagnetic shower is measured by the amount of Čerenkov photons reaching the PMTs.

Hadrons on the other hand, lose their energy along their path due to ionization radiation and hadronic interactions. The mean free path for hadrons is larger than for electrons.When a π^0 is created in a nuclear interaction, the neutral pion decays into two photons which in turn start an electromagnetic shower. Hadrons lose energy uniformly along their path, with the development of an electromagnetic component towards the end of their path.

The HERMES calorimeter is built out of 840 identical radiation hard F101 lead-glass blocks [90] with an area of 9×9 cm^2, and a length of 50 cm. Properties of F101 are given in Table 3.2. The block size was chosen in order to contain 99% of an electromagnetic shower inside a matrix of 3×3 blocks. Each block is mirror polished and wrapped in aluminized

Chemical Composition F101	weight %
PB_3O_4	51.23
SiO_2	41.53
K_2O	7.0
Ce	0.2
Radiation Length	2.78 cm
Critical Energy	17.97 MeV
Refraction index	1.65
Molière Radius	3.28 cm

Table 3.2: Chemical composition and calorimetric properties of F101 Lead Glass. Cerium is making the Lead Glass radiation hard, while also reducing its transparency.

Mylar foil, and covered with a Tedlar (= Polyvinyl Fluoride) foil to prevent light leakage into neighboring blocks. PMTs are glued to the end of the blocks with a silicon glue. The length of the blocks was optimized to improve the energy resolution. The energy deposition in the lead glass blocks is shown in Figure 3.20 for muons, electrons and pions. Muons act as minimum ionizing particles. Pions induce a minimum ionizing peak with a long tail due to hadronic showers. Electrons lose nearly all of their energy. The development of electromagnetic and hadronic showers is also schematically shown in the figure.

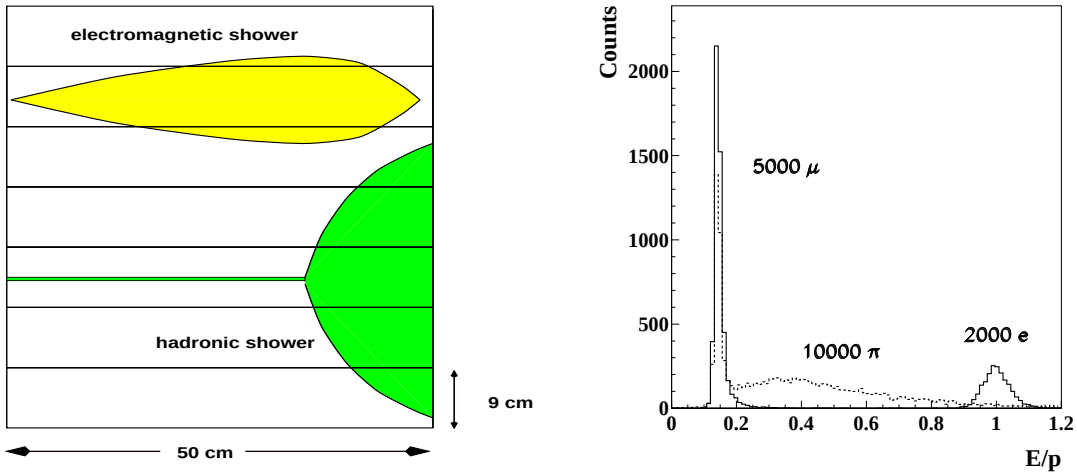

Figure 3.20: On the left the typical development of electromagnetic and hadronic showers in the HERMES calorimeter. Right: the $\frac{E}{P}$ ratio for 5 GeV muons, pions and electrons.

Electron-hadron separating capabilities of electromagnetic calorimeters can be substantially improved if the calorimeter is preceded by a material with a large mean free path length for hadrons [91]. By doing so, charge exchange relations such as $\pi^- p \rightarrow \pi^0 n$ are suppressed. Hadrons (electrons) have an interaction length of about 17 (0.56) cm in lead.

Therefore a $\sim 2X_0$ thick preshower built out of a 11 mm thick lead plate sandwiched between two 1.3 mm stainless steel sheets was constructed right in front of the calorimeter. The preshower is followed by a hodoscope H2 which will be detailed in section 3.3.3.1. A schematic

drawing of the preshower-calorimeter setup is shown in Figure 3.21. The segmentation of H2 and the calorimeter is clearly visible.

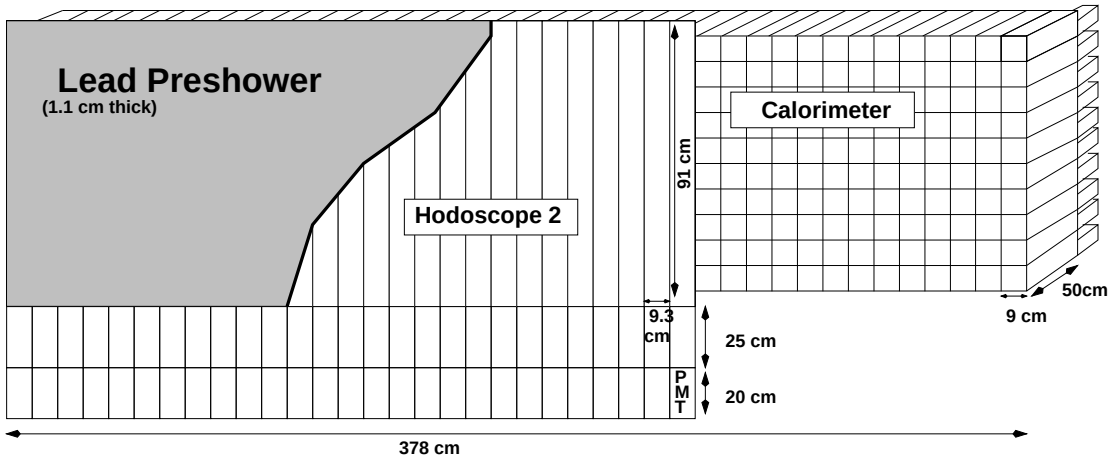

Figure 3.21: Schematic view of the preshower and calorimeter located in the lower detector half.

The linearity of the calorimeter has been verified in various test beams and is depicted in Figure 3.22. The data is well reproduced by a linear fit. However, for low energies, the data are better described by a cubic fit [90]:

$$E = 0.19 + 0.43 \cdot 10^{-2} \cdot \mathrm{ADC} - 0.10 \cdot 10^{-6} \cdot \mathrm{ADC}^2 + 0.10 \cdot 10^{-10} \cdot \mathrm{ADC}^3. \tag{3.7}$$

The cubic fit allows a correction of the electron energy deposited in the preshower. The block length of 50 cm ($\sim 18X_0$) was chosen to fully contain the electromagnetic shower induced by electrons and hence to improve the resolution. The energy resolution can be parametrized by [89]:

$$\frac{\sigma(E)}{E} = \frac{5.1 \pm 1.1}{\sqrt{E}} + (2.0 \pm 0.5) + \frac{10.0 \pm 2.0}{E}. \tag{3.8}$$

This is the resolution as measured in the HERMES experiment and is slightly worse compared to test beam data due to showering of the electrons in the material before the calorimeter. The energy resolution is plotted in Figure 3.23.

The hadron rejection factor of the HERMES calorimeter was measured to be between 10 and 100 depending on the positron energy and the calorimeter threshold. The rejection factor increases to values between 50 and 160 in the off-line analysis.

Even though F101LG is a radiation hard material, the calorimeter was built on a translating stage allowing the calorimeter to be moved 50 cm away from the beam pipe during injection and acceleration. Radiation doses are monitored by TF1 blocks placed behind the calorimeter.

It is important to realize that the preshower and calorimeter are the only detector stages within the HERMES spectrometer in which photons leave a signal. Consequently, the calorimeter is of the utmost importance to measure photons coming e.g. from π^0 decay. The calorimeter is required to determine photon energies and give a second space point. The latter is mandatory to determine photon tracks. The design of the HERMES calorimeter, however, was not optimized to detect photons.

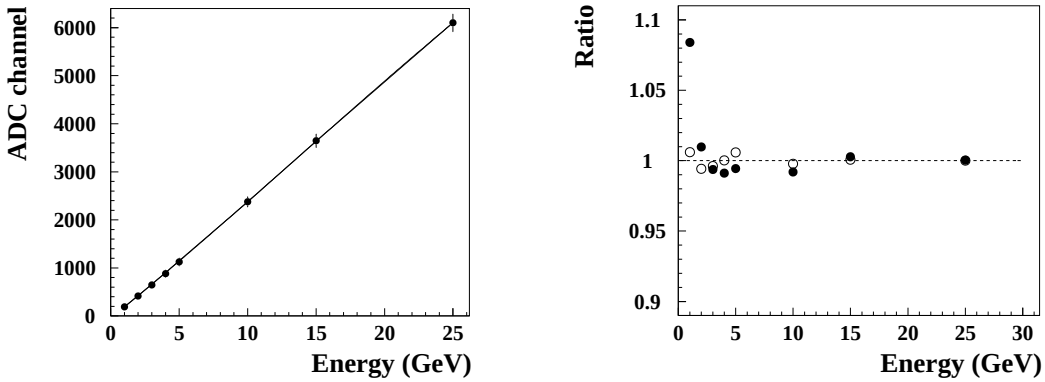

Figure 3.22: Left the linearity of the calorimeter response, on the right the closed (open) circles represent the ratios between the linear (cubic) fit and the data. Figure taken from Reference [90].

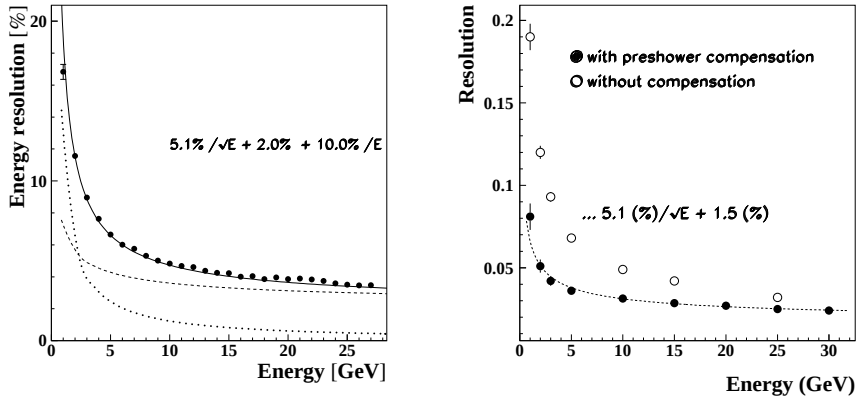

Figure 3.23: Calorimeter resolution as measured in HERMES (according to Reference [89], left) and in a test beam (according to Reference [90], right). The dashed (dotted) curve in the left panel represents the lead-glass (preshower) contribution.

3.3.2.3 Ring Imaging Cherenkov Detector

The HERMES Ring Imaging Cherenkov Detector (RICH) [92] was installed in 1998 as a replacement for the older threshold Čerenkov detector, in order to improve the discrimination between pions, kaons and protons. Nearly all hadrons in the HERMES kinematic and geometrical acceptance are either pions, kaons or protons with momenta between 2 and 15 GeV. Conventional threshold Čerenkov counters using heavy gases like C_4F_{10} are only useful at energies above 9 GeV, because of the high kaon threshold. In order to comply with the HERMES requirements for identification of pions, kaons and protons, a RICH with two radiators had to be built.

Particles of rest mass m_0 and momentum P with a velocity larger than the speed of light in a medium with refractive index n emit Čerenkov radiation at an angle θ defined by:

$$\theta = \arccos\left(\frac{1}{n}\sqrt{1+\left(\frac{m_0 c}{P}\right)^2}\right). \tag{3.9}$$

The produced Čerenkov light can be reflected by a spherical mirror and detected by an array of PMTs. The first radiator needs to have a low index of refraction in order to be able to detect the light from low momenta particles. Therefore aerogel with a refractive index $n = 1.03$ was decided on [93]. The second radiator material, C_4F_{10} gas slightly above atmospheric pressure, was chosen in order to be able to separate pions and kaons up to 15 GeV.

The expected Čerenkov angles as a function of momentum for electrons, pions, kaons and protons for aerogel and for C_4F_{10} are depicted in Figure 3.24. Electrons always have momenta above threshold of both radiators. In fact, two RICH detectors have been installed, one above and one below the HERMES symmetry plane. A schematic view of a HERMES RICH is shown in Figure 3.25, together with its principle of operation.

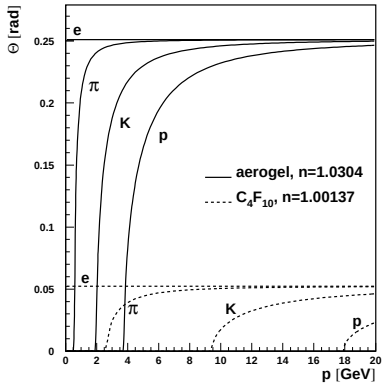

Figure 3.24: Čerenkov angle versus momentum for aerogel (solid lines) and C_4F_{10} (dashed lines) radiators. Pions, kaons and protons are clearly separated by combining the signal from both radiators.

The aerogel layer is built out of 5 layers of a 5×17 matrix of tiles. For the whole detector 850 tiles were needed. These were selected from a sample of 1680 tiles [93, 94]. The tiles are stacked according to their measured refractive index, thickness and surface quality. The total length of the aerogel layer is 5.5 cm. Dry nitrogen is constantly flushed through the aerogel layer to prevent degradation by the C_4F_{10} .

The mirror array exists out of 8 spherical mirror segments, with a radius of curvature of 2.20 m. All mirrors were aluminized to provide a reflectivity of 85%. The mirror alignment procedure by offline software was performed in three steps: first, the mirror as a whole was varied. By using a set of 24 parameters the response to single high energy lepton tracks was optimized . Next the individual mirrors were aligned, by comparing the response to a Monte Carlo simulation. Finally the single mirrors itself were divided in 3 matrices, which were aligned as if they were individual mirrors.

Most of the Čerenkov photons are emitted in the UV-light region. However, the reabsorption by the RICH-gas for photons with these wavelengths is high. Therefore most of the photons reaching the PMTs are in the visible light region, though some of the light coming from the C_4F_{10} radiator extends into the UV region. Conventional PMTs sensitive to the visible light region were chosen to detect the Čerenkov light. It was shown that the single photon resolution for both aerogel and C_4F_{10} are dominated by the pixel size of the photon detector. For cost compromises 3/4 inch PMTs were chosen, leading to a pixel size of 23.3 mm.

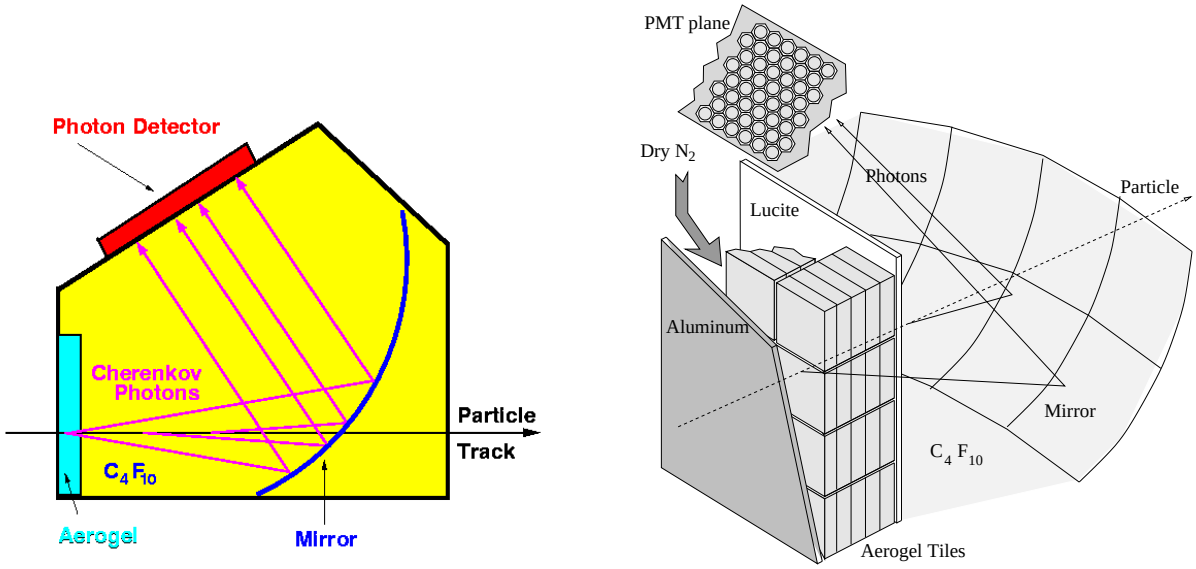

Figure 3.25: Schematic view of a RICH. Čerenkov light emitted by a particle is reflected by a mirror and detected by a PMT plane. Both radiators, the PMT plane and the mirror are shown in both figures.

All PMTs were stacked in a hexagonal packed matrix of 73 columns and alternately 26 and 27 rows. Only 38% of the area is covered by PMTs. In order to increase the light yield, aluminized funnels precede the PMTs. As the PMTs in the photon detector operate in the stray field of the spectrometer magnet, they needed to be shielded with $100\ \mu\mathrm{m}$ thick μ-metal. The readout is based on the LeCroy PCOS4 system.

The detected 'rings' are in fact not circular, bur rather distorted ellipses, as the photon detector is not exactly in the mirror's focal surface. The aerogel rings are rather difficult to detect, since only a few photons are produced in the aerogel layer. A typical event with a kaon in the upper and a pion and electron in the lower detector half is depicted in Figure 3.26. A 14.6 GeV electron leaves according to Figure 3.24 an aerogel and a gas ring. The 1.5 GeV π^- and the 5.5 GeV K^+ only leave an aerogel ring. The particle associtated with the 1.5 GeV track is clearly a pion since a 1.5 GeV kaon is below Čerenkov threshold. Moreover, a pion at 5.5 GeV would have had an additional gas ring, while the radius in case of a proton would be much smaller. The figure is an indication of the power of a RICH detector for particle identification.

From the above event it is clear that the RICH is a powerful PID instrument, though the reconstruction is not always that simple. In HERMES two reconstruction methods are used: Indirect Ray Tracing (IRT) and Direct Ray Tracing (DRT). The RICH PID Scheduler (RPS) combines the best features of both tracing methods by defining which method should be used. More information about the RICH PID reconstruction can e.g. be found in Reference [95].

3.3.3 Triggering

In order to separate useful physics hits from background contributions, a fast signal to trigger the spectrometer readout is mandatory. The HERMES trigger scheme is rather straightforward:

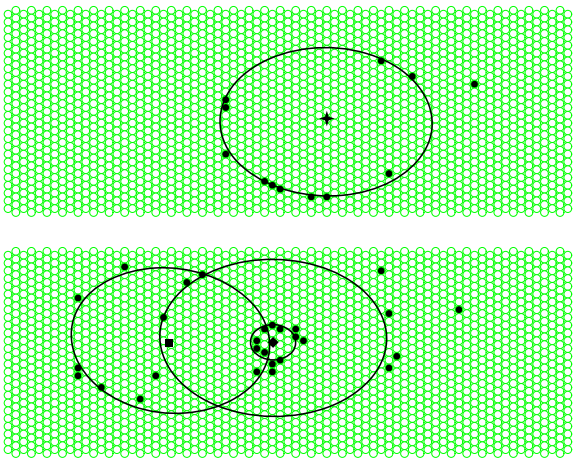

Figure 3.26: A typical RICH event with a 14.6 GeV electron in the lower half, yielding two 'rings', a 1.5 GeV π^- in the lower half and a 5.5 GeV K^+ in the upper half.

if a combination of signals of certain detectors are above a predefined threshold, all detector responses are read out. The entire HERMES trigger is first-level based only. For the purpose of triggering three hodoscopes, H0, H1 and H2 were built. They will be discussed first, later the various HERMES triggers will be covered.

3.3.3.1 Hodoscopes

In ancient Greek the word ὁδος means 'path'. Hodoscopes determine the path of charged particles and are built out of scintillating material in combination with PMTs. Sometimes a wavelength shifter or a light guide is needed to transfer scintillating photons to the PMT and eventually adjust the wavelength. H1 and H2 are located in the back part of the spectrometer (see Figure 3.11) and have a similar construction. They are built out of 42 paddles consisting of 9.3 $\times$ 91 $\times$ 1 cm^3 of scintillator material, a 25 cm long light-guide and a PMT at the outer end of the paddles. The paddles are staggered to obtain a maximum efficiency. The scintillator material is BC-412 from Bicron, which is a fast scintillator with a large attenuation length. H2 is often referred to as the preshower. The PMT signal is split and guided into ADCs and Constant Fraction Discriminators (CFDs). Discriminated signals go to TDCs and can be used for the time of flight technique for PID, which will be explained later. The logical OR of all CFD channels is used for trigger decision.

During 1995 running only the backward region was used for triggering, leading to a large background originating from the HERA proton beam in which protons are traveling in the -z direction. To improve trigger efficiencies, a hodoscope H0 was installed in the front region in 1996. H0 is constructed from a 3.2 mm thick plastic scintillator plate and is read out by two PMTs.

In 1998 HERMES was upgraded for the so-called charm program [96]. The goal was to detect the decay muons of J/Ψ and D mesons. To achieve these goals, the HERMES acceptance needed to be increased, and muons needed to be identified. The latter was achieved by placing a 'muon-hodoscope' behind a 1 meter thick iron absorber, which stops virtually all particles but muons. The acceptance increase was obtained by installing 'Front Muon-' and

'Wide Angle Muon' (WAM) hodoscopes. Their position is indicated in Figure 3.11. The steel of the HERMES magnet acts as a muon-filter for particles in the acceptance of the WAM.

3.3.3.2 HERMES Triggers

There are two main HERMES physics triggers: a DIS and a photo-production trigger. The DIS trigger selects scattered leptons via a signal in H0,H1,H2 and a calorimeter response above a programmable threshold. This threshold needs to correspond to a calorimeter signal above that of a minimum ionizing particle, which is 0.8 GeV. Initially the threshold was set to 3.5 GeV, but for polarized data taking it was lowered to 1.4 GeV. For triggering purposes, the responses of all calorimeter blocks in a column are added together to form a 'column sum'. The signals of 2 columns are combined pairwise into a 'segment sum'. These segment sums are the signals which are used for triggering.

The photo-production trigger is made to detect the production of hadrons. Consequently the calorimeter can not be used, and the BC is used instead to discriminate between useful events and background hits coming for example from collimator showers. The latter source of background can be further reduced by demanding the size of the signal in H0 not to exceed 8 times the size of a MIP signal.

Trigger 21 is the DIS trigger and trigger 28 the photo-production trigger. Apart from these two main physics triggers, there are other triggers for specific processes [97]. The overall trigger logic is established in a programmable logic unit which should be gated with a NOT BUSY signal. All triggers can be individually prescaled.

3.4 Luminosity

In order to do absolute measurements of e.g. cross sections, and to be able to compare data sets of different years or polarizations, it is very important that the luminosity is determined precisely. In HERMES this is done either by counting the number of DIS events in a data set, or by using the luminosity monitor [98]. A discussion of both methods is given in section 7.1.3. The design of the luminosity monitor will be described in this section.

Determining the luminosity can be done using a well known physics process, with a sufficiently high cross section to keep both systematic and statistic error low. An example of such process is the elastic scattering of beam particles off electrons from the target gas. In the case of an electron beam the process is called Møller scattering ($e^-e^- \rightarrow e^-e^-$), for a positron beam the process is referred to as Bhabha scattering ($e^+e^- \rightarrow e^+e^-$). An additional reaction which can be used in the case of a positron beam is the annihilation of an electron and positron into a pair of photons. All these reactions have small scattering angles and leave both particles with a similar amount of energy.

The luminosity monitor is built out of two symmetrical halves positioned on the left and right from the HERA lepton beam pipe. Each halve is built out of 12 NaBi(WO_4)$_2$ crystals, which is a highly radiation hard material. The blocks are read out by photomultipliers. Due to high radiation doses during beam injection and acceleration, the luminosity monitor was placed on a movable table allowing it to be moved in position right before data taking. When a physics run is about to be finished, the luminosity monitor is moved back into a position where it is safe from synchrotron radiation. A schematic view of the luminosity monitor can

be seen from Figure 3.27. The luminosity monitor has its own dedicated readout scheme. This scheme allows a dead time free operation of the luminosity monitor.

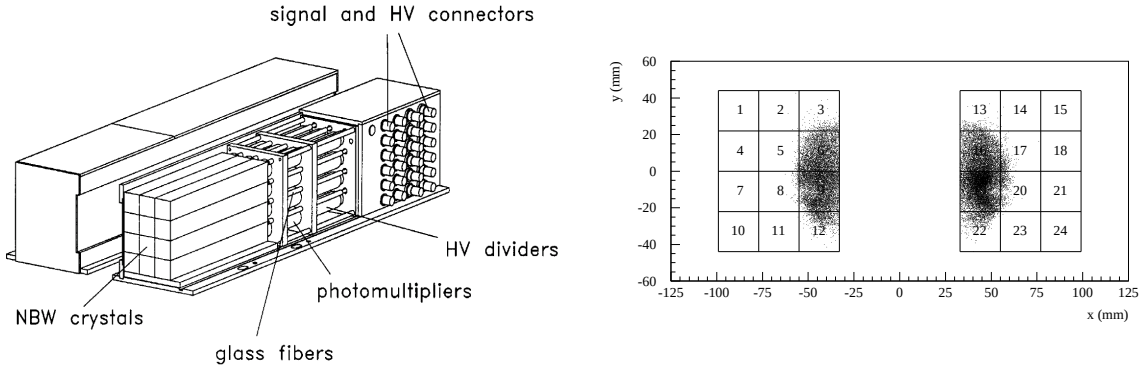

Figure 3.27: Schematic view of the luminosity monitor. On the right a front view with the reconstructed impact points from elastic scattering can be seen. The elliptical beam profile is well reflected.

Since the luminosity monitor is positioned outside of the standard HERMES acceptance, it requires its own dedicated trigger in order to select desired processes. Most background events will deposit energy in only one side of the luminosity monitor, while Bhabha or Møller events will have an equal energy deposition in both halves. Triggering occurs when signals in both luminosity detector halves exceed a 4.5 GeV threshold. The motivation for this choice can be seen in Figure 3.28.

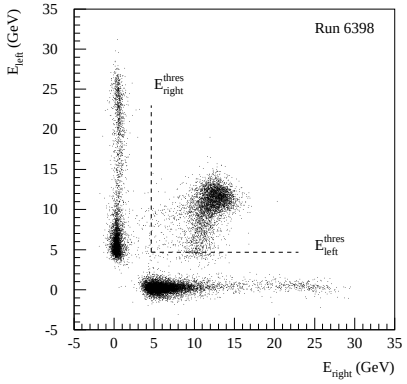

Figure 3.28: Energy deposited in the left luminosity monitor halve versus the right luminosity monitor halve. The trigger threshold is indicated in the figure. Energies of both halves add up to about 27.5 GeV, the HERA beam energy.

3.5 Background

Due to the nature of the HERMES experiment, all sensitive equipment is located close to the beam pipe. Additionally, due to the setup, the level of background radiation is high. The

main background is coming from synchrotron radiation and a beam halo of charged particles. However, also the proton beam induces a substantial background.

Beam tuning can be achieved by specially developed dipole magnets. A combination of these correction magnets can either move the beam position or the beam slopes. Several Beam Position Monitors (BPM) are installed in the HERA ring. They are used to monitor the beam position mainly during filling. This section will focus on background monitoring just before and during data taking.

3.5.1 Tuning scintillator telescope

The HERMES Tuning Scintillator Telescopes (TST) are built out of 2 layers of 4 scintillator plates around the beam pipe. The scintillators are read out by PMTs so they can give a fast indication of the beam halo. The second scintillator layer is preceded by 3 mm lead in order to induce signals from synchrotron radiation in the second plate. (Anti-)Coincidence rates between 2 layers thus are an indication of the amount of (synchrotron) charged particle background. The TSTs are located in front of the target, just in front of the FCs and right behind the luminosity detector. Evidently, all TSTs are positioned outside of the HERMES acceptance. They are mostly used to have an indication of the background before turning on the spectrometer.

3.5.2 Background monitoring

The response of various detectors can be used as an online indication of the instantaneous background. One of the most sensitive set of detectors are the Front Chambers, because of their location close to the target cell. Moreover, due to the absence of a magnetic field these chambers are hit by a lot of low momentum particles. The FCs are sensitive to both particle and synchrotron background.

By design the TRD is sensitive to synchrotron radiation. A small metal plate has been put right in front of the TRD to absorb some of the low energy photons. As an indication for the particle yield in the back region of the experiment, hodoscope H1 can be used.

It was also shown [98] that analyzing the luminosity detector data gives rather precise information about the beam position at the target. However, the luminosity detector is not a tuning device, as it is only moved into position after the background has reached an acceptable level.

3.5.3 Beam loss monitor

Because of the installation of the Lambda Wheels close to the target cell and close to the beam pipe a Beam Loss Monitor (BLM) was installed during summer 2002 [99]. This device can trigger a kicker magnet which is able to dump the HERA lepton beam if the radiation level becomes too high.

The BLM is located right behind the DVCs and consists of 2 pairs of three ionization chambers installed left and right of the beam. The chambers are filled with Argon gas which is ionized by halo particles and synchrotron radiation. Signals from the chambers are integrated and amplified before they are sent to an ADC, and a fast trigger circuit. The fast trigger is

formed by a discriminator and a certain time over threshold. The second trigger is formed by integrating the output over 1 s. These two triggers are needed because both instantaneous high doses as well as higher doses over a longer time period can severly damage detector equipment. A majority of two out of three triggers from one detector set dumps the beam. A veto, however, is needed to prevent dumping during beam injection, when radiation levels are always high.

3.5.4 Proton beam background

As the direction of the proton beam is opposite to that of the lepton beam, proton background can be monitored by paddles of the muon hodoscopes which are behind the lead wall. The installation of hodoscope H0 has also decreased the amount of proton background in the data.

3.6 Data acquisition

The backbone of the HERMES Data Acquisition (DAQ) is based on FastBus crates mastered by CERN Host Interfaces (CHI). To enhance performance most crates are equipped with Struck FastBus Readout Engines (FRE), featuring one or two Motorola 96002 Digital Signal Processing (DSP) units. After the acceptance of a certain trigger, all detector responses are read out by this DAQ chain according to their 'equipment number'.

Apart from event-based information, also monitoring and calibration data needs to be read out from the various detectors. These data are referred to as 'scalar' events and are read out every ten seconds. The time period between two scalar events is called a burst.

Information like pressure gauges, high voltage settings or temperature can be read out on a much longer timescale. This information is called 'slow control' and is read out by VME, CAMAC or RS232 and forms the so-called 'fill files'. Offline, these data are synchronized with the event data.

The data is organized in 450 MB blocks of data in 'Experimental Physics Input Output Package' (EPIO) [100] format. One such block of raw data is called a run. After a HERA fill is finished, all collected runs are backed up and copied over to the DESY main site.

In order to reduce the event size, hot wire suppression and pedestal subtraction is performed at the DSP level. Moreover, a double buffering mechanism is implemented at the DSP allowing event collection parallel with the next event readout. In order to monitor possible pedestal shifts, the first events of every run are read out unsparsified, i.e. without zero suppression[2].

In 2000 the DAQ was upgraded with new DSPs in order to allow the future installation of other detector components, like the Recoil Detector. The DAQ is capable to handle a trigger rate of 300 Hz with an event size of 2×13 Kb [101]. The overall HERMES dead-time has typically always been below 10 %[3].

[2]Usually only signals above a certain threshold (the zero line, i.e. pedestal + X-times the noise) are read out. This is called zero suppression.

[3]Achieved by prescaling triggers.

3.7 Performance monitoring

Even though all detectors can be calibrated after a certain data taking period, it is important to have an immediate indication about detector performance during data taking. Therefore a dedicated Gain Monitoring System (GMS) was installed and various online monitoring scripts have been written. These will be briefly covered here.

3.7.1 Gain monitoring system

The HERMES Gain Monitoring System (GMS) [102] has been designed in order to monitor the performance of all 952 PMTs of the calorimeter and the hodoscopes. The GMS is schematically shown in Figure 3.29. The system is based on a Nd:YAG laser with a wavelength of 532 nm, chosen to match the PMTs quantum efficiencies and a low attenuation in the lead-glass blocks. The laser light is split by a mirror to a trigger PMT and to the detectors and some reference detectors. A filter wheel was installed in front of the splitter in order to be able to vary the amount of light fed into the detectors. The reference detectors are photo-diodes with a very small temperature dependence.

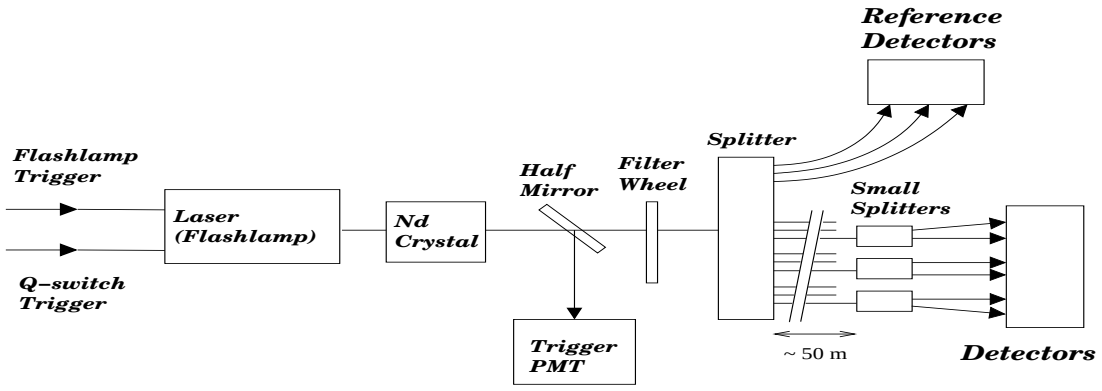

Figure 3.29: Schematic view of the Gain Monitoring system.

Based on the response of the reference detectors, the relative gain for all 952 PMTs can be determined. These measurements are compared with reference values. Differences are displayed on an online monitor. Every 20-30 seconds the display is updated.

3.7.2 Online monitoring

Important information like beam positions, leakage currents and detector temperatures need to be monitored continuously, as these directly affect the data taking. Online monitoring of these observables is based on a client-server network and a graphical interface.

The servers are based on the Distributed ADAMO Database (DAD) [103], which is an extension to ADAMO (Aleph DAta MOdel) [104]. They form the core of the slow control and online monitoring system. From these servers data can be read and displayed with simple scripts in a tcl/tk extension called PINK [105]. In the HERMES control room there are over 60 of these displays.

Apart from the above, the response of individual channels for every spectrometer component are monitored by an HBOOK [106] based program called the 'wire maps'. These indicate whether all ADC or TDC channels are read out correctly and give the expected responses.

This concludes the description of the HERMES experimental setup. The HERA beam, the HERMES target and the HERMES spectrometer were described. The next section describes how information from the various tracking and PID detectors are combined to form tracks and provide a high efficiency PID.

CHAPTER 4

Hermes Data Production

This chapter describes how the responses of the various hardware components in the HERMES experiment are combined to form useful observables. The ADC and TDC values coming from the hardware need to be translated into energy and time, tracks need to be found and fitted, particle identification has to be preformed. Moreover, data analysis requires information on the quality of the data taken is required. The data production chain forms the very important step in between data taking and data analysis.

The various steps of the path leading to a format usable for physics analysis are shown in Figure 4.1. Two branches can be seen, one corresponds to the 'Main Production', where the raw EPIO format is converted to the ADAMO table format by various programs. The other branch is referred to as the 'Slow Production' and combines all information about detector performance. The first two sections of this chapter explain the 'Main' and 'Slow' production. Next, the particle identification (PID) algorithm is detailed and finally the combination of main and slow production into the 'micro Data Summary Type' (uDST) format is explained.

4.1 The Main Production

The HERMES Main Production is indicated as the upper flow chart in Figure 4.1. In the first step the EPIO files coming from the DAQ are decoded by the HDC program. In the next step, the information from the various tracking devices is combined and tracks are being searched and reconstructed. This task is accomplished by HRC, XTC and TMC. Finally, alignment and efficiencies are calculated by the ACE program. This section discusses all these programs.

4.1.1 Decoding - HDC

The first step in the main production chain is the decoding of the raw data into hits in certain detector volumes. The HERMES decoder HDC is responsible for this first step and executes

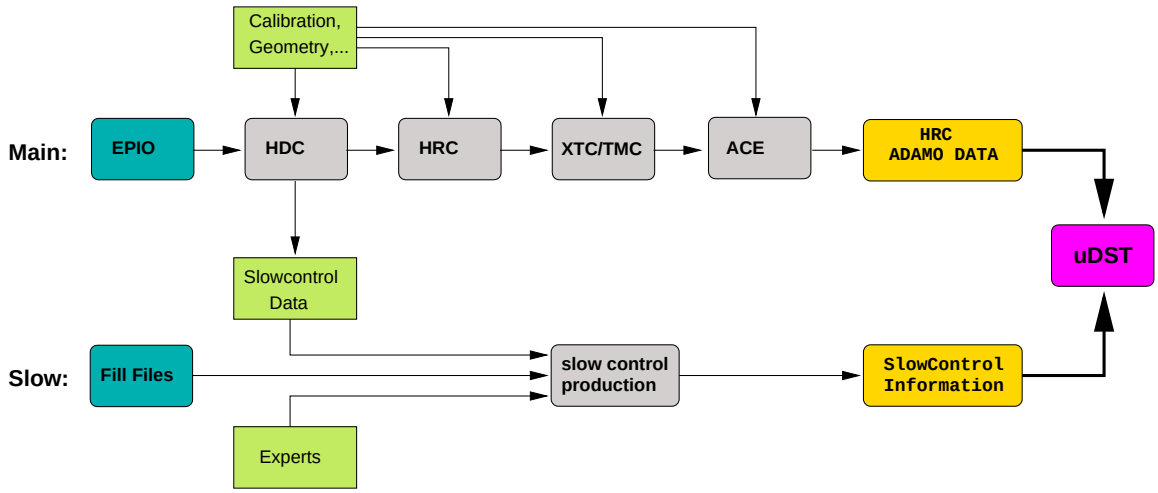

Figure 4.1: Schematic overview of the HERMES data production data chain. Starting from the online EPIO and Slow-Control information, the micro-DSTs (uDST) are produced. Alignment, calibration, and information from detector experts is necessary to come to the uDSTs.

three important tasks: mapping, calibration, and geometry.

First of all a hit in a certain 'equipment number' needs to be assigned to a certain detector volume. After this first mapping step, detectors are referred to with names like 'B2U7', meaning the seventh wire in the U plane of BC2.

The next step is to convert ADC numbers into physical quantities. From here on we can talk about the energy deposited into a detector volume. The exact conversion constants are coming from the various detector groups, who are responsible to calibrate their specific device. Finally detectors need to be located in the HERMES coordinate system, by making use of a geometry file. After passing through HDC the data are organized as localized, calibrated hits in the various detector volumes.

4.1.2 Reconstruction - HRC

The Hermes ReConstructing program HRC is responsible for assigning tracks to the hits in the different wire chambers, and to combine calorimeter hits into clusters of calorimeter blocks. HERMES Tracking is done in two steps: track finding and track fitting. More information about HRC can be found in Reference [107].

4.1.2.1 Track Finding

Track finding at HERMES is based on a 'tree-search' algorithm. The hits of the U, X, and V planes of all tracking chamber are combined separately to form one-dimensional track projections called 'tree-lines'. These are found by iteratively comparing the detector response with several predefined patterns. In each step the resolution is doubled. If the detector response matches a certain pattern, this pattern is referred to as the parent. Subsequently the response is compared to several possible child patterns.

For every parent there are a limited number of possible child patterns, reducing the number of pattern comparisons. A tree-search of depth 11 requires a comparison to 70 possible patterns on average. An example of the parent-children data bank is given in Figure 4.2.

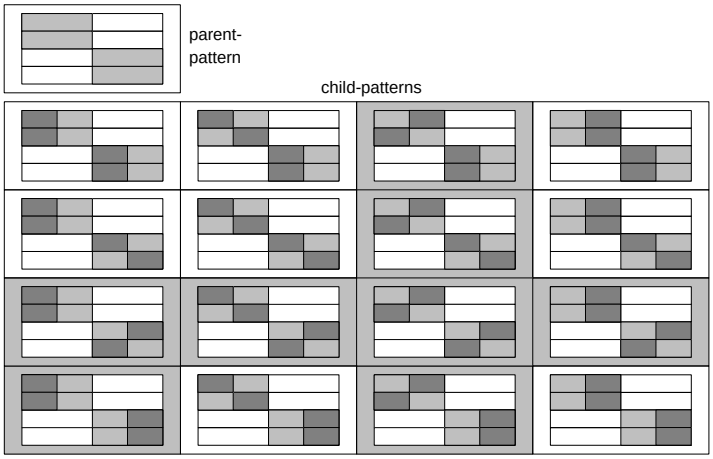

Figure 4.2: Example of the parent-children database. The example shows 4 detector planes; the parent track is indicated by the light gray areas in the upper left corner. The 16 (=2^4) possible child patterns obtained by splitting the parent hit in two are shown by the dark gray area superimposed on the light gray parent pattern. Only 8 child patterns correspond to a possible particle track. These are drawn on a white background.

The child pattern is either matched, in which case we go to the next step in the iteration or unmatched, meaning that it is impossible to form a tree-line. The branch consequently is abandoned. This iteration is continued for 11 or 10 steps, yielding a resolution of $2^{-11} \times 2890 \text{ mm} = 1.4 \text{ mm}$ for the back partial tree-lines, i.e. the part of the tree-line after the magnetic field, and $2^{10} \times 660 \text{ mm} = 0.6 \text{ mm}$ for the front partial tree-linek, since the width of the BC and FC modules respectively are 2890 mm and 660 mm. An example of the tree search algorithm for two tree-lines and 4 detector planes is depicted in Figure 4.3.

There are over 126 million possible patterns, which are reduced to 31,000 entries by using symmetry arguments. The algorithm was written to be able to deal with a certain amount of missing hits, so as to match detector inefficiencies.

4.1.2.2 Track Fitting

The tree search algorithm is used only to identify a tree-line, its parameters are calculated from a straight line fit to the detector hits along the trajectory, since the magnetic field before and after the magnet is negligible. The different resolutions of the tracking chambers are accounted for by weights in the χ^2 fit.

Once the tree-lines are found, they are combined to form a three dimensional track. The (x, y) position of the track is found from the rotation of U and V tree-lines, and a comparison with the X tree line. Once the (x, y) positions coming from the various tracking detectors are known, these can be combined to form partial tracks in the front region and in the back region, i.e. before and after the magnet.

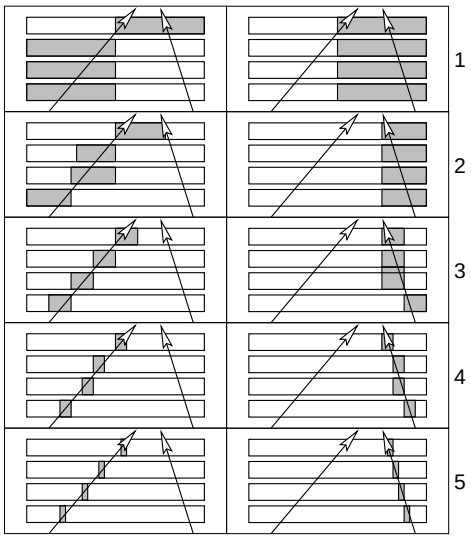

Figure 4.3: Tree-search principle for two tree-lines. In every step the detector resolution is doubled. In this example four detector planes are shown.

There are several ways to find the front partial track. It can be formed by either using only the FCs, or by using the FCs and DVCs (and earlier even by using the Vertex Chambers). Usually only the FCs are contributing to the front partial track, due to a non-trivial misalignment in the DVCs.

The combination of the front and back partial tracks can be done in various ways: One is to simply combine front and back partial track at the center of the magnet ($z = 275$ cm). This method is possible since the largest deviations between front and back partial track are only up to 1.0 cm for low energy particles, higher energy particle partial tracks are obviously lying even closer to each other. Another combination method uses the position of the back track as an additional point for the front track. This method is referred to as forced bridging, and manages to increase the momentum resolution significantly.

Particles with a momentum below 4.5 GeV never reach the back part of the spectrometer, and thus have only a front partial track. The latter partial track can, however, be combined with the information from the Magnet Chambers to form a so called 'short track', indicated in Figure 3.15.

The tracking algorithm requires a finer position resolution than the one provided in the geometry file. Additional alignment information is needed to improve the tracking capabilities of HRC. This information comes from dedicated alignment runs, in which the spectrometer magnet is turned off, yielding only straight tracks. In the offline alignment procedure, the positions of the various tracking devices are slightly varied until the best resolution is achieved.

4.1.2.3 Track properties: momentum and vertex

A very important parameter associated to a particle track is its deflection in the magnetic field which fixes the momentum of the particle. The momentum associated to a track is determined using a look-up table, segmented in 4 dimensions: the z-vertex of the track, two angles in the

x and y direction: θ_x and θ_y, and $\mathrm{d}\theta_x$, the change of θ_x in the magnetic field. The momentum then is found via an interpolation of the table. An accuracy of 0.1 % is achieved.

The table itself was generated using a Runge-Kutta algorithm to track randomly generated particles through the HERMES magnetic field. The latter was mapped for all three field components B_x,B_y and B_z as a function of x, y and z, with a position accuracy of less than 50 μm, and a field accuracy < 1 Gauss [108]. The field itself was interpolated by a linear combination of two quadratic interpolations.

If a track does not fit into the kinematics of the momentum lookup table, the 'shooting method' is used. The front and back partial track are iteratively compared until the right momentum is found. This method yields an accuracy of 0.2%.

From the combination of all tracks in an event, the event vertex can be determined. The routine calculating the event vertex is an iterative procedure. In the first step the minimum distance between all tracks of the event is calculated. If in the second iteration, the distance between a particular track in the event and the vertex calculated in the first iteration turns out to be too large, this track is omitted and the vertex recalculated.

The response of some detectors depends on the track parameters. The drift signal in the anode wires is e.g. only known after the vertical position of the track is known. Another example is the signal in the hodoscope which depends on the attenuation of the photons in the scintillating paddles. Therefore track parameters and momentum need to be recalculated after recalibration of some detector responses.

4.1.2.4 Clustering

Apart from tracking, HRC also performs the clustering of the calorimeter blocks related to the same electromagnetic shower induced by an incident particle. The main goal of the clustering is twofold: to know the total amount of energy deposited in the calorimeter by a particle and the coordinates of the impact point of a particle track with the calorimeter. Two clustering algorithms have been implemented.

The first algorithm searches for any hit calorimeter block, and then looks for hits in the neighboring calorimeter blocks. If the neighbors have hits, the search is continued to their neighbors. Blocks that have been tested are marked again. The cluster energy is then simply the sum of all blocks in the cluster. This algorithm is unable to separate overlapping clusters.

The second algorithm is based on the observation that over 98 % of a shower is contained in a cluster of 3 $\times$ 3 calorimeter blocks [90]. The cells with a local energy maximum, and the eight surrounding cells are added to form a cluster. This method is indicated in Figure 4.4. The overlap between two clusters C_1 and C_2 is accounted for by applying the following formula:

$$E^{(1)} = \sum_{i \in C_1 \setminus C_2} E_i \left(1 + \frac{\sum_{i \in (C_1 \cup C_2)} E_i}{\sum_{i \in (C_1 \cap C_2)} E_i} \right) \tag{4.1}$$

$$E^{(2)} = \sum_{i \in C_2 \setminus C_1} E_i \left(1 + \frac{\sum_{i \in (C_1 \cup C_2)} E_i}{\sum_{i \in (C_1 \cap C_2)} E_i} \right) \tag{4.2}$$

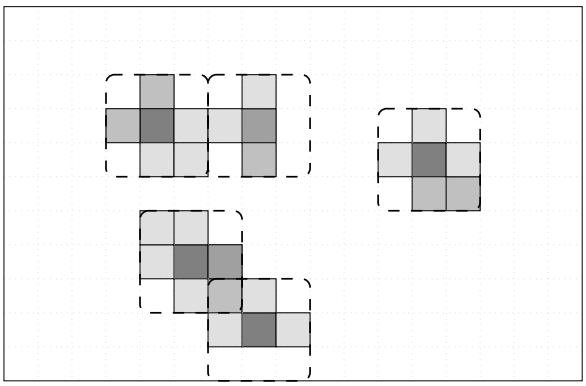

Figure 4.4: Example of the clustering algorithm. Local maxima in the calorimeter are indicated in dark gray, together with the eight neighboring cells a calorimeter cluster is formed. Figure from [109].

The cluster position was initially calculated using a square root weighting of the individual calorimeter blocks with energy deposition E_i in the cluster:

$$\begin{pmatrix} x_{\mathrm{cl}} \\ y_{\mathrm{cl}} \end{pmatrix} = \frac{1}{\sum_i \sqrt{E_i}} \begin{pmatrix} \sum_i \sqrt{E_i} x_i \\ \sum_i \sqrt{E_i} y_i \end{pmatrix} \tag{4.3}$$

This procedure, however, biases the cluster position towards the center of the central block, yielding a so called 'S-Curve' when plotting true versus reconstructed positions of a Monte Carlo sample, as indicated in Figure 4.5.

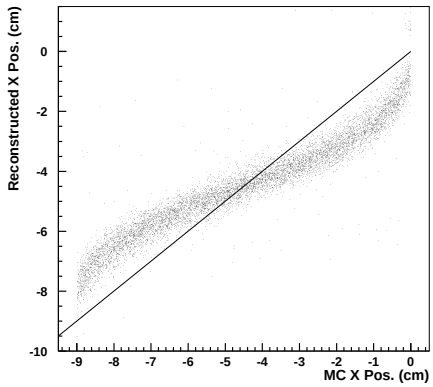

Figure 4.5: Generated versus reconstructed calorimeter position in a single calorimeter block. The reconstructed position is determined using square root weighting, equation 4.3. Figure from Reference [109].

This bias can be reduced by taking the exponential shower profile into account, using a logarithmic weighting method [110, 109]:

$$x_{\mathrm{cl}} = \frac{\sum_i w_i x_i}{\sum_i w_i}, \qquad w_i = \max(0, W_0 + \frac{\ln E_i}{\sum_i E_i}), \tag{4.4}$$

where W_0 is a threshold cutoff parameter which reduces background by excluding signals below a certain threshold. A typical value of W_0 is 4.8. The position reconstruction can additionally be improved by correcting for the incident angle of the particle.

4.1.3 External tracking - XTC

The HERMES eXternal Tracking Code XTC, was implemented to perform tracking outside of the standard HERMES acceptance. This code became part of the HERMES production chain when the acceptance was expanded with the Wide Angle Muon Hodoscope and the Lambda-Wheels. XTC will also be upgraded to perform tracking for the Recoil Detector.

During the 2002-2005 data taking period HERMES was equipped with a transverse magnet around the target cell, providing a transversely polarized target. This magnetic field, however, introduced an offset to the particle vertexes and track angles. Therefore a special correction code, the Transverse Magnet Corrections, TMC was introduced in between the HRC and ACE steps in the main production chain as indicated in Figure 4.1.

4.1.4 Efficiencies - ACE

Originally, ACE was planned to be responsible for Alignment, Calibration and Efficiency calculations of the HERMES experiment. In practice, however, ACE only calculates the efficiencies of the tracking chambers. Efficiencies, calibrations, and alignment calculations are done by experts in the respective detector groups, and are given as an input to the production chain.

4.2 Slow Production

The Slow Production is responsible for indicating the equipment state during data taking. Detector information comes from three sources and needs to be merged together. One source is the online slow control information, like leakage currents and temperatures. These are often referred to as 'fill files'. Another source comes from the expert files, made by the different detector groups. These contain offline calibration factors, or other important information. The third source comes from the main EPIO files after they have been passed through HDC. Information about for example the target spin state is included in these.

The EPIO files are marked with a VME time, while the fill files have UNIX time. Care must be taken when merging these files together. Therefore a dedicated HERMES time has been established indicating the number of seconds since January 1st, 1995, the date when HERMES started, and the number of microseconds since the last second.

4.3 Particle identification

The HERMES Particle IDentification has two distinct functions. One is to separate leptons from hadrons, and one to distinguish between various hadron types. The latter task is completed by the RICH, as already mentioned in section 3.3.2.3. The first task is the combined result of calorimeter, preshower, TRD, and, eventually the RICH detector. Additionally for low momentum particles, the time of flight (TOF) technique is used to identify particles. The

combined electron-hadron separation algorithm as well as the TOF technique are explained in this section.

4.3.1 Electron-hadron separation

The electron-hadron (or positron-hadron) separation within the HERMES experiment is based on a probability analysis. For every track with momentum p, polar angle θ, and deposited energy $E_{\rm dep}$ in detector volume $\mathcal{D}$, we can assign a probability $\mathcal{P}$ that the particle associated to the track is of type $\mathcal{T}$: $\mathcal{P}(\mathcal{T}|(E_{\rm dep},p,\theta))$. We can then apply Bayes' theorem and write:

$$\mathcal{P}(\mathcal{T}|(E_{\rm dep},p,\theta)) = \frac{\mathcal{P}_f(\mathcal{T}|(p,\theta))\cdot\mathcal{P}_p(E_{\rm dep}|(\mathcal{T},p,\theta))}{\sum_{\rm t=l,h}\mathcal{P}_f(\mathcal{T}_t|(p,\theta))\cdot\mathcal{P}_p(E_{\rm dep}|(\mathcal{T}_t,p,\theta))}, \tag{4.5}$$

where $\mathcal{P}_f(\mathcal{T}|(p,\theta))$ is the probability that a track with polar angle θ and momentum p is of type $\mathcal{T}$. This probability often is referred to as the 'particle flux'. $\mathcal{P}_p(E_{\rm dep}|(\mathcal{T},p,\theta))$ on the other hand, is the probability that a particle of type $\mathcal{T}$ with momentum p, and angle θ will deposit an energy $E_{\rm dep}$ in the detector. This distribution is referred to as the 'parent distribution'. Assuming the detector responses to be uniform within the acceptance, we can write the parent distribution as $\mathcal{P}_p(E_{\rm dep}|(\mathcal{T},p)$. The sum in equation 4.5 runs over leptons and hadrons only, as these are the two particle categories to be identified.

The probabilities $\mathcal{P}(\rm hadron|E_{dep},p,\theta)$ and $\mathcal{P}(\rm lepton|E_{dep},p,\theta)$ are combined into a logarithmic ratio PID:

$$\mathrm{PID} = \log\frac{\mathcal{P}(\rm lepton|E_{dep},p,\theta)}{\mathcal{P}(\rm hadron|E_{dep},p,\theta)} = \underbrace{\log\frac{\mathcal{P}_p(E_{\rm dep}|(\rm lepton,p))}{\mathcal{P}_p(E_{\rm dep}|(\rm hadron,p))}}_{\rm PID_{det}} - \log\underbrace{\frac{\mathcal{P}_f(\rm hadron|(p,\theta))}{\mathcal{P}_f(\rm lepton|(p,\theta))}}_{\Phi}. \tag{4.6}$$

$\rm PID_{det}$ is known from specific detector properties, and Φ can be estimated from the data. If a track has a PID $< 0 (> 2)$ it is more likely that the particle associated with the track is a hadron (lepton). The information from various detectors can now easily be combined defining several PID quantities:

$$\begin{aligned} \mathsf{PID2} &= \mathsf{PID_{cal}}+\mathsf{PID_{pre}} \\ \mathsf{PID3} &= \mathsf{PID_{cal}}+\mathsf{PID_{pre}}+\mathsf{PID_{cer/rich}} \\ \mathsf{PID5} &= \mathsf{PID_{TRD}} \end{aligned}$$

to form an overall PID = PID3 + PID5 - $\log\Phi$. The RICH is only capable of doing lepton-hadron separation for low momentum tracks [111]. The overall PID distribution is shown in Figure 4.6. The influence of the flux factor and the RICH to the PID distributions can be seen.

The parent distributions, which give the probability that a particle of type $\mathcal{T}$ deposits an energy $E_{\rm dep}$ in a detector volume, could in principle be derived from detailed detector measurements in test beams prior to installation. However, to take noise, smearing, and aging effects into account, it is better to use the response of the other PID detectors to determine the efficiency of a certain detector. Therefore, when parent distributions are calculated, some hard cuts on the response of the PID detectors are applied [112].

To calculate the particle flux ratio Φ exactly, the PID information would be required as an input, which obviously is not allowed. Therefore the determination is done iteratively,

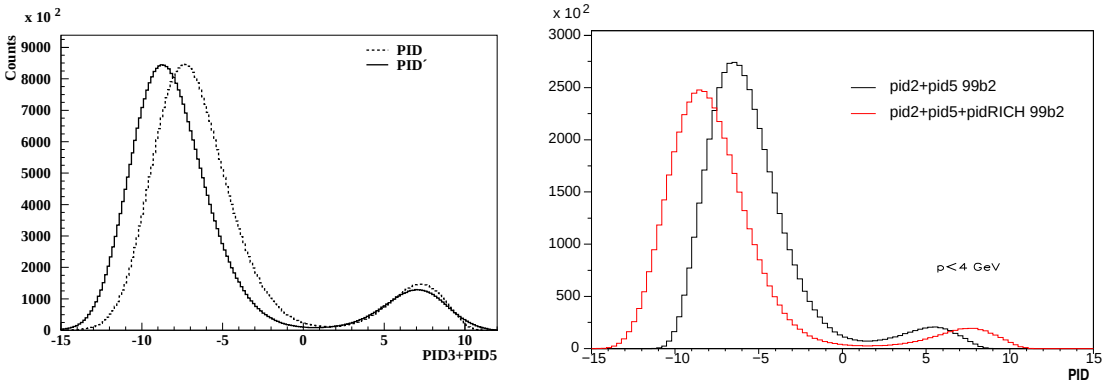

Figure 4.6: The left panel shows the influence of the flux correction to the PID distribution. The full line is flux corrected PID, while the dashed line is without flux correction. Right: the influence of the RICH to lepton/hadron separation for low momenta (P< 4 GeV).

whereby the flux ratio is uniform in the first step, $\Phi_0 = 0$. In the next step the PID from the previous step is taken: $\Phi_1 = \frac{\#\text{tracks with PID}<0}{\#\text{tracks with PID}>0}$. The iteration is continued until convergence is achieved. The flux ratio depends in principle on the target gas. The difference between polarized hydrogen and polarized deuterium targets, however, is small. Flux ratios for polarized deuterium are shown in Figure 4.7.

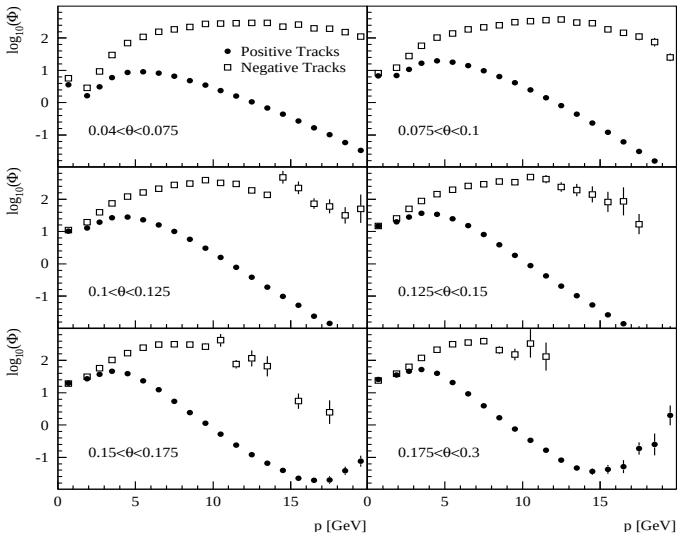

Figure 4.7: Logarithm of the flux ratio Φ for polarized deuterium during 2002 running. Open squares (filled circles) are the ratio for negative (positive) charged track. Figure from Reference [113].

.

The PID efficiency to determine leptons is dependent on the PID cut, and the kinematic variables. However, an overall efficiency of over 98% with a contamination less than a percent is generally reached. More details can be found in References [113, 114, 115].

4.3.2 Time of flight

Due to the bunched HERA electron beam, and the presence of the fast hodoscope arrays, the Time-Of-Flight technique can be used to separate low momentum hadrons. By measuring the time it takes for a particle to arrive at a hodoscope, and knowing its path from the information from the tracking algorithm one can determine its velocity and thus $\beta = \frac{v}{c}$. The mass of a particle is related to its momentum and to β via:

$$M^2 = P^2 \left(\frac{1}{\beta^2} - 1 \right).$$

Leptons with an energy above 10 MeV are essentially moving at the speed of light, they can consequently be used to calibrate the timing from the hodoscopes. A timing resolution of 0.49 ns could be achieved. The procedure is described in detail in Reference [116].

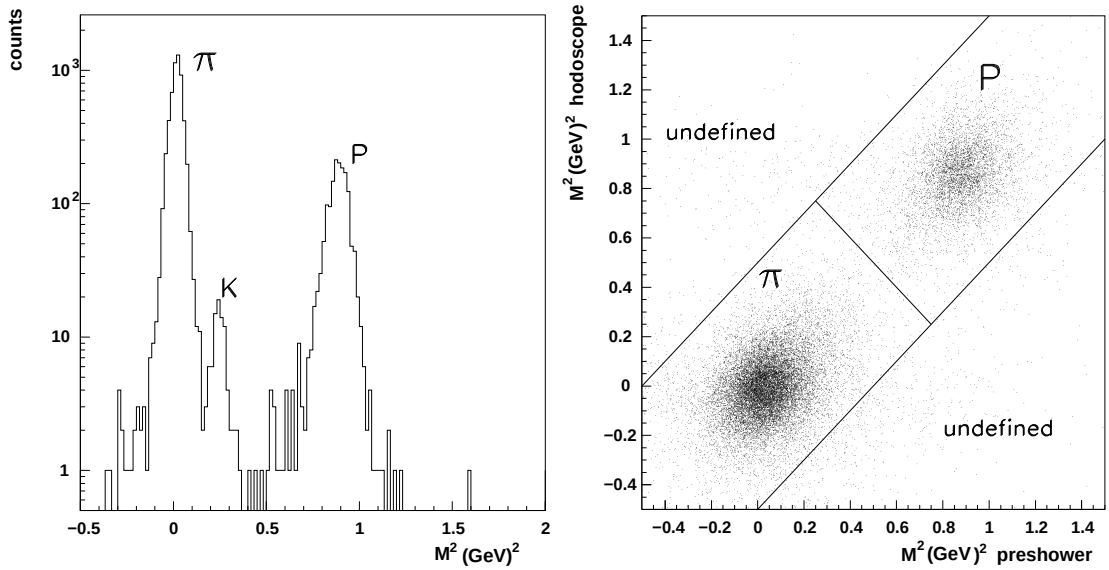

Figure 4.8: The left panel shows the M^2 distribution obtained by analyzing hodoscope H1 data for 0.6 GeV $< P <$ 1.1 GeV. The right panel shows the combined M^2 distributions obtained from hodoscopes H1 and H2.

For particle momenta below 1.5 GeV, pions, kaons and protons can be separated, as can be seen in Figure 4.8. Due to the low kaon flux, kaons are taken as a contamination in the pion sample for higher momenta.

Since there are two hodoscope arrays in the HERMES experiment, their information can be combined to improve efficiency, as indicated in Figure 4.8. Doing so, a pion efficiency above 98% with a proton contamination less than 4% could be reached. Proton efficiency is determined to be 85%, with less than 6% pion contamination.

4.4 Micro DST Production

The micro DST production (uDST) merges the slow and main production together to provide all necessary information for performing physics analysis, without being redundant and without

the need of unnecessarily large amounts of disk space. Some important PID calculations are done in the uDST production as well. The data are essentially organized in bursts and tracks. The output format has been chosen to be the ADAMO table format. The main event data are stored in the 'Track' table, which is filled once per event. Additional tables describe the state of the experiment. These are usually filled once per burst. Sometimes the state of the experiment changes within one burst, which for instance is the case when the target is flipped. A split of the quantity 'burst' is consequently required. The additional labeling introduced is the uDST-number, which is augmented whenever the state of the experiment changes within a burst. The data of a running period are ordered in run-number, burst-number, uDST-number, and event-number.

An 'a' production is the first production of a data taking period. It is run without expert files, and is not meant for data analysis. Its purpose rather is to serve as an online check for the different detector groups, who then can perform calibrations. These consequently serve as an input to the 'b' production.

4.5 Data quality

After a uDST production is completed, its quality is carefully checked by the data quality group. This group of people is looking for instance for human errors or high voltage trips affecting the data. After a sufficient amount of study, the group compiles a 'burst-list', which is a bit-pattern indicating the state of various detector components. Analyzers could check this bit-pattern against a mask. Doing so, a user requires the components essential to a particular analysis to be fully functional, while allowing components not used in the analysis to have certain malfunctions.

CHAPTER 5

A Silicon Recoil Detector for HERMES

A large part of the available time during this PhD project was spent on research and development for the new HERMES Silicon Recoil Detector (SRD). The Recoil Detector [117] was designed and built to measure recoiling target particles from exclusive reactions. The detector upgrade was decided on because of the relation of these exclusive reactions to GPDs. The HERMES Recoil Detector will be located around the target cell. Due to severe space constraint, the ABS and BRP needed to be removed. The HERMES experiment consequently had to stop its polarized target physics program. Recoil Physics is the main HERMES physics from 2006 up to the end in summer 2007. The Recoil Detector is built out of three detecting stages: a silicon, a scintillating fibre and a photon detector, all of which are located inside a superconducting magnet.

This chapter details the design and initial tests of the silicon detector. First, the overall detector requirements and corresponding design are mentioned. Later, the readout of the SRD is discussed in detail: both the readout chip and the DAQ chain are covered. The third section describes the prototype tests, both in a laser based test-stand, and in a MIP test beam at DESY.

5.1 Detector requirements

The Recoil Detector is built to constrain exclusivity at the event level. This is achieved by measuring all reaction products. Currently, the recoiling proton is outside of the HERMES acceptance. Exclusivity is achieved by putting constraints on the missing mass or missing energy of a data sample measured by the forward spectrometer. A detector built around the target cell increases the standard acceptance so that recoiling protons can be measured directly. Moreover, the Recoil Detector will remove inexclusive background from the exclusive data set. The detector design evidently needs to match the kinematics of exclusive reactions. This

section gives an overview of the detector requirements based on the kinematics of exclusive reactions.

5.1.1 Kinematics of exclusive reactions

The main processes to be studied with the Recoil Detector are Deeply Virtual Compton Scattering (DVCS) and Exclusive Meson Production. DVCS is the scattering of a virtual photon off a quark inside the proton with the creation of a real photon while the proton is staying intact:

$$\gamma^* + p \rightarrow \gamma + p. \tag{5.1}$$

This reaction has the advantage that the produced photon is directly related to the scattering quark, without the dilution of for example a meson distribution amplitude. The final state of DVCS is exactly the same as the one of the Bethe-Heitler (BH) process, where the electron is scattered by the proton as a whole, and a Bremsstrahlung photon is created. The latter process dominates at HERMES energies. The amplitudes of BH and DVCS add up coherently, yielding an interference term, which can be expressed in terms of GPDs as explained in section 2.6.3.1.

The benefit of exclusive meson production, on the other hand, is the possibility of extracting different combination of GPDs depending on whether pseudoscalar or vector mesons are produced. The additional information from hard exclusive meson production is mandatory to come to a good description of the nucleon in terms of GPDs.

Ideally, the Recoil Detector should be able to determine the full kinematics of the recoiling proton, and it should measure any other source of possible non-exclusive background. This background could e.g. come from the decay of Δ resonances into pions, neutrons and photons. An efficient detection of these particles would be beneficial to measure exclusive reactions with a low background.

The kinematics of a recoiling proton from the combination of BH and DVCS is shown in Figure 5.1. In the same figure the momenta of protons coming from intermediate Δ decay are given: $\gamma^* + p \rightarrow \gamma + \Delta^+(p \rightarrow +\pi^+)$. The latter reaction is called Associated DVCS (ADVCS). Recoiling protons are mainly going in a transverse direction up to 1.35 rad ($= 80^\circ$) with respect to the beam axis, and mostly have momenta below ~ 0.5 GeV. However, momenta up to 1.4 GeV and angles down to 0.1 rad are to be expected. Distributions for recoiling protons from exclusive ρ^0 production look similar.

5.1.2 Recoil Detector design

The Recoil Detector should be mounted around the target cell to cope with the θ distribution of the recoiling particles. Moreover, to deal with the different kinds of particles, it should be sensitive both to charged and neutral particles. Exclusive measurements require a large momentum acceptance as well. Therefore the Recoil Detector was designed to have three detectors: a silicon detector, a scintillating fibre and a photon detector. All components of the Recoil Detector are indicated in Figure 5.2.

A detector measuring low momentum recoiling protons essentially needs to be placed inside the beam vacuum, since most recoiling protons are too low in energy to penetrate the scattering chamber. Moreover, the detector has to be able to measure the energy of recoiling protons. Because of its good vacuum properties and its excellent energy resolution, silicon has

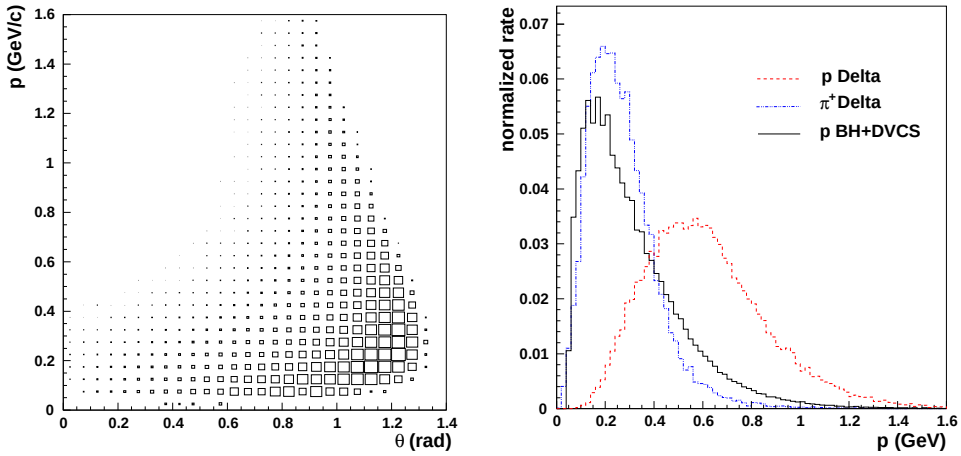

Figure 5.1: Left: kinematic distribution in P and θ for protons coming from BH/DVCS. Right: Momentum distribution for protons coming from BH/DVCS (black line) and for pions (dash-dotted line) and protons (dashed line) coming from intermediate Δ decay. Figure from Reference [117].

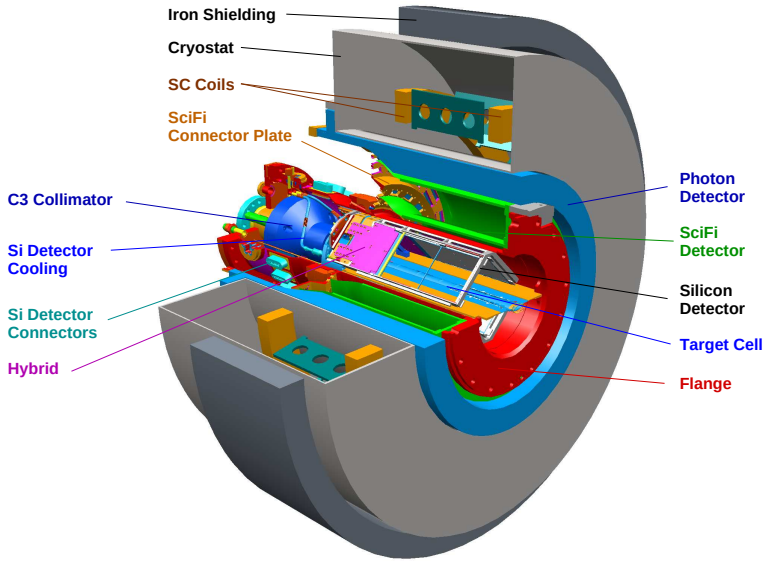

Figure 5.2: The HERMES Recoil Detector. The various detector components are indicated in the figure.

been chosen as the active material for the first detector stage. Momenta are measured from the energy deposited in the silicon. Figure 5.2 shows the shape of the silicon sensor's holding frame in gray.

Pions and higher energy protons will deposit less energy in the silicon. These particles have a high probability of escaping the scattering chamber. To measure these particles an additional detecting stage was required. The scintillating fibre technique was chosen to determine particle tracks. The use of scintillating fibres has the advantage that the fibres can easily be bent into

cylindrical shapes, which is needed to fully encircle the scattering chamber. Discrimination between pions and protons is done via the energy deposited into the fibres. A magnetic field provided by a 1 T superconducting magnet is required to determine particle momenta [117, 118]. The scintillating fibre tracker is indicated in green in Figure 5.2 and is built out of two co-axial cylindrical barrels. Each barrel is built out of 2 layers of fibers, rotated 30° with respect to each other.

Additional tracking points from the silicon detector significantly improve the overall tracking capabilities of the recoil detector. It therefore is beneficial if the silicon detector can also measure the charge deposited by minimum ionizing particles (MIP, see further). Even though most silicon detectors in high energy experiments are designed to measure MIPs, the detection of those particles is difficult in a system optimized to detect large energy depositions in the silicon.

To measure any additional photons either produced at the interaction point itself, or coming from the decay of an intermediate Δ resonance a photon detector was built. The photon detector consists of three layers, each containing converter material and an active stage and is depicted in blue in Figure 5.2.

In order to reduce background hits from the beam halo in the recoil detector, an additional collimator C3 was designed and positioned right in front of the target cell. The collimator can be seen in dark blue in Figure 5.2.

The option of a neutron detector was studied in the beginning of the project. However, both space and time limitations prevented a neutron detector to be designed and constructed.

5.2 Silicon strip detectors

Prior to the design of the Silicon Recoil Detector (SRD), the general response of silicon strip sensors to ionizing radiation is discussed here. First, the common construction of silicon sensors as used in particle physics experiments is reviewed. Then, the interaction of radiation with such sensors is presented. Excellent reviews of silicon sensors are given in References [119, 120].

5.2.1 Silicon sensors

Silicon is a semi-conducting material, meaning that in general it is not conducting due to the large band gap between valence and conduction band. However, a silicon crystal can be doped with donor and acceptor atoms, which take the place of silicon atoms in the lattice. An acceptor is an element of group III in the periodic system. It has one electron less than silicon. As such a hole is created, which can freely move through the lattice. Materials where holes can freely move is called p-type material. When on the other hand the silicon is doped with elements of group V, electrons are allowed to freely move. This type of material is called n-type material.

If p- and n-type silicon are brought into contact, the holes from the p-type migrate to the n-side. The electrons from the n- type on the other hand move to the p-side, so that a net positive (negative) charge is created on the n (p) side of the junction. The transition zone, being free of charge carriers, is called the depletion zone and has a potential difference across due to the net charge on the p and n-side of the junction. The combination of p- and n-type

material is called a $p-n$ junction. This junction is nothing else but a diode, through which current can flow from the side having a higher potential, i.e. the n-side, to the one with the lower potential, the p-side.

If ionizing radiation hits a silicon sensor, electrons from the valence band are excited to the conduction band where they can freely move. A hole in the valence band consequently is created. The hole can travel through the valence band albeit with a different mobility. The energy required to create such an electron-hole pair depends on the band gap, which is 1.1 eV for silicon. An energy higher than the band gap energy is required to create an electron-hole pair, as there are other competing processes through which incident energy can be dissipated like lattice vibrations or excitations or the creation of phonons. An average energy of 3.6 eV is required to create an electron-hole pair in silicon.

If electron-hole pairs are created in the neutral, non-depleted zone they recombine with the free carriers of the doped material. However, if electron-hole pairs would be created in the depleted region, they would move to the n- and p-side respectively, leaving a net charge on both sides of the junction.

Ideally, one would want to have a very large depletion zone to have as much charge as possible created at both sides of the junction. This can be achieved by applying a 'depletion' or 'bias' voltage over the diode by attaching a negative (positive) potential to the p- (n-) side. The diode consequently is reversely biased and the depletion zone is extended into the two regions. The extension is inversely proportional to the concentration of impurities. Usually, the concentration of acceptors on the p-side is much larger than the concentration of donors on the n-side. The depletion layer thus extends deeply into the n-side. This type of materials is referred to as p^+n junctions. Typical thicknesses are 300μm for the n-bulk, and sub-micron thickness for the p^+ layer, the latter only being necessary to build the active detector volume and collect the created charge.

To construct a usable sensor, several p^+ strips are positioned in parallel to each other on the n- bulk. The distance between two strips is referred to as the pitch. The incident position of ionizing radiation determines which strips contain most of the charge. Charge collected on the strips needs to be amplified and digitized. Amplification usually is performed by a charge sensitive preamplifier. A schematic drawing of a silicon sensor is given in Figure 5.3.

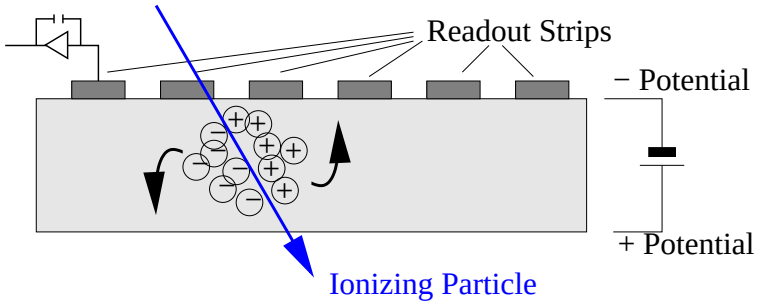

Figure 5.3: The principle of operation of a p^+n silicon strip detector. When a particle traverses a silicon sensor, electron-hole pairs are created along its path. The electron-hole pairs are separated under the influence of a bias voltage. For simplicity, coupling capacitors (see later) are not indicated in the figure.

The back side of a sensor is usually covered with a thin layer of highly doped n^+ type

silicon. This n^+ layer is implemented to provide a good ohmic contact to the external world, and to prevent the depletion area to reach the sensor surface. This side is often referred to as the 'ohmic' side, whereas the p^+n side is referred to as the 'junction' side.

In particle physics double sided silicon sensors are often used. These are constructed by segmenting the n^+ side into different strips rather than using a n^+ plane. One needs to be careful though, since the resistance between two n^+ strips is low. They act as if they were connected. Charge is thus spread over a wide area, and position information becomes coarse. The strips need to be blocked from each other, which is achieved by putting a p-stop implant in between two n^+ strips. As a consequence, the n^+ strips are electrically isolated from each other (for a fully depleted detector), yielding the possibility to have position information. The p^+ strips on the junction side, and the n^+ strips on the ohmic side can be oriented orthogonally to each other in order to give independent (x, y) coordinate information[1]. Together with the location of the sensor itself in space, a three dimensional space point can be reconstructed.

Due to e.g. thermal excitation a continuous leakage current is flowing between p- and n-side. To avoid this continuous current, the p^+ or n^+ strips are usually covered with an isolating SiO_2 layer. On top of the silicon-oxide, an aluminum readout strip is placed providing an ohmic contact to the external world and a capacitive coupling to the strips. This creates an effective isolation of the leakage current from the readout chain.

5.2.2 Energy loss in silicon

In general the average energy loss of an ionizing particle inside a material is given by the Bethe-Bloch formula. The formula does not account for the experimental fact that some high energy knock-on electrons (δ-rays) are escaping the material. The description of energy loss in thin detectors is more correct when using the truncated form of the Bethe-Bloch formula, which takes the escaping δ-rays into account. The latter depends on a cut-off energy T_{cut} depending on the size of the detector volume. The truncated Bethe-Bloch is given by [121]:

$$-\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x} = -\kappa z^2 \frac{Z}{A}\frac{1}{\beta^2}\left(\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 T_{\text{upper}}}{I^2} - \frac{\beta^2}{2}\left(1+\frac{T_{\text{upper}}}{T_{\max}}\right) - \frac{\delta}{2}\right), \tag{5.2}$$

where:

$$\begin{aligned}
\rho &= \text{density of the material } (= 2.330\,\tfrac{\text{g}}{\text{cm}^3})\\
\kappa &= 4\pi N_A r_e^2 m_e c^2\ (= 0.3071\,\tfrac{\text{MeVcm}^2}{\text{g}})\\
Z(A) &= \text{atomic number(mass) of the medium}\\
z &= \text{charge of incident particle (in units of elementary charge)}\\
T_{\text{upper}} &= \min\left(T_{\max}, T_{\text{cut}}\right)\\
T_{\max} &= \text{maximum kinetic energy transferred to an electron in a single collision}\\
&= 2m_e c^2\beta^2\gamma^2\ (\text{if } \tfrac{2\gamma m_e}{M} \ll 1 \text{ where M = mass incident particle})\\
I &= \text{Mean excitation energy } (= 173.0 \text{ eV for silicon})\\
\delta &= \text{'Density effect'} \text{ correction to energy loss by ionization}
\end{aligned}$$

A typical value of T_{cut} is on the order of 100 keV for electrons in silicon sensors [122]. Equation 5.2 is shown in Figure 5.4 for ionizing protons passing through a 300 μm thick silicon layer

[1]Note, however, that orthogonality is not a demand for independent coordinate information from both sides. The Lambda Wheels have e.g. strip oriented under an angle of 30° with respect to each other

in the momentum region relevant for the recoil detector. The density effect was neglected since only momenta up to 1.6 GeV are expected and this effect is only relevant at higher energies (and higher Z). Particles with an energy corresponding to the minimum plateau of the Bethe-Bloch are defined as 'minimum ionizing particles' (≡MIP). Protons with momenta larger than 1.1 GeV are minimum ionizing.

Formula 5.2 could be corrected with a radiative loss term, expressing the additional radiative energy losses. The effect of such correction is small and mostly only relevant for electrons.

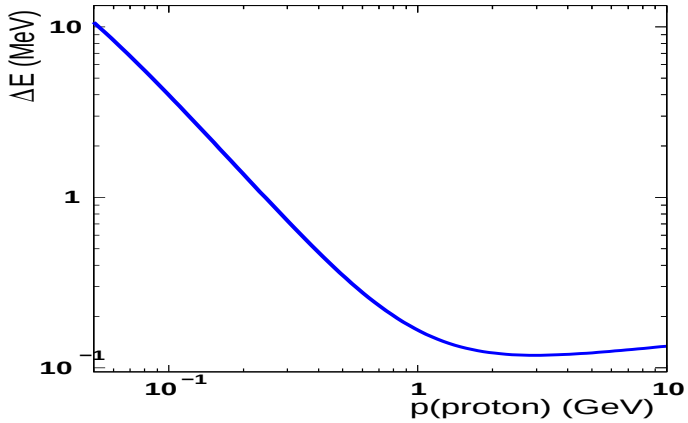

Figure 5.4: Average energy loss of a proton in 300 μm silicon, the loss was calculated by evaluating the Bethe-Bloch formula 5.2.

From Figure 5.4 it can be observed that the energy deposition in silicon for particles with less energy than a MIP follows a 'linear' behavior[2]. Indeed, for low momenta, only the first term in equation 5.2 will contribute. The energy deposition is proportional to $\frac{1}{\beta^2}$. The momentum of a low energy proton can thus be determined directly from its energy deposition. This feature is exploited in the design of the SRD: by measuring the amount of charge created in the silicon, proton momenta are reconstructed. At HERMES kinematics, protons are the only particles depositing these large amounts of energy in the silicon, allowing a trivial PID. In order to get the best energy resolution, the amount of material between the silicon sensors and the interaction point needs to be minimal. Therefore the SRD is put inside the ultra high HERA beam vacuum. The only material between the interaction point and the silicon modules is the 25 μm thick target cell[3].

Due to the possibility of large energy transfer in a single collision, the energy deposition inside thin detector layers is not Gaussian distributed. A difference between the 'most probable' and the 'average' energy deposition is the consequence. The energy deposition distribution can be approximated by a Landau distribution [82]:

$$\begin{aligned} L(\lambda) &= \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(\lambda+e^{-\lambda})} \\ \lambda &= \frac{\Delta E - \Delta E^W}{\xi} \end{aligned} \tag{5.3}$$

[2]Figure 5.4, however, has a log-log scale

[3]This was the original design thickness of the target cell for the recoil program. A target of 75 μm was finally used.

where $\xi = \frac{\kappa}{2}\rho\frac{Z}{A}(\frac{z}{\beta})^2$, ΔE the actual energy loss, and ΔE^W the most probable energy loss. ΔE^W can be determined by taking the mean of equation 5.3:

$$\Delta E^W = -\xi < \lambda > + < \Delta E > \tag{5.4}$$

The mean energy loss $< \Delta E >$ of a proton in a silicon layer is known from the Bethe-Bloch formula (equation 5.2). $< \lambda >$ can be calculated by taking the first moment of the Landau distribution:

$$< \lambda >= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \lambda e^{-\frac{1}{2}(\lambda + e^{-\lambda})} \mathrm{d}\lambda = \gamma + \ln 2$$

where γ is Euler's constant. Evaluating equation 5.4 for a MIP gives $\Delta E^W = 85.5$ keV. Knowing that 3.6 eV is needed to create one electron-hole pair in silicon it follows that a charge of $\sim$ 24000 electrons is 'most probable' to be created by a MIP. The Landau distribution for minimum ionizing particles in silicon is shown in Figure 5.5.

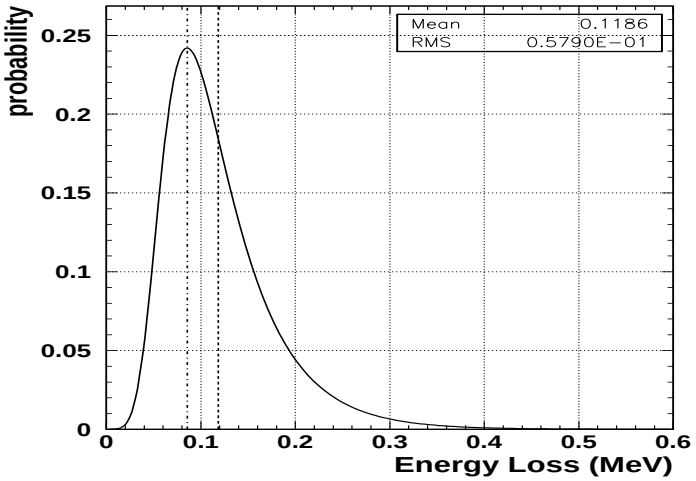

Figure 5.5: Energy loss distribution for minimum ionizing protons in silicon. The distribution is represented by a Landau curve. The dashed (dash-dotted) line indicates the mean (most probable) energy loss.

5.3 Silicon Recoil Detector conceptual design

A particle's energy can be measured by a sequence of thin silicon sensors where particles undergo fractional energy loss. The sequence should be followed by a thicker sensor where the particle is stopped in order to measure all its energy. This array of sensors is referred to as a a $\Delta E - \Delta E' - ... - E$ telescope. For the SRD a telescope consisting of two layers of silicon sensors was chosen. The second layer could not have an arbitrary thickness, since this would influence the measurement of particle momenta in the scintillating fibre tracker due to multiple scattering. Moreover, the production of double sided silicon sensors with arbitrary thickness is challenging, because a manufacturer's production process would have to be modified to

produce sensors of that thickness. Changing a production process is a time consuming and non-trivial work and was beyond the timescale and budget of the project. Due to the small scale of the project it was more beneficial to have an equal thickness for both layers, since in this case only one type of modules needed to be ordered and implemented. For these reasons, the thickness of both layers was chosen to be 300 μm.

According to a study performed in Reference [123] the required pitch of the silicon sensors should be on the order of 1 mm. The sensor chosen to be used for the SRD was the TIGRE sensor, built by Micron Semiconductors Ltd. The strip pitch of the TIGRE is 758 μm. The strip pitch of 758 μm yields a ϕ resolution of about 0.04 rad. The resolution in θ is less important, given the good momentum resolution of the forward spectrometer for the scattered lepton. The energy resolution will be mentioned further. More information on the TIGRE is given in section 5.5.1.

The available space for the construction of the SRD in the scattering chamber is tightly constrained. Therefore a diamond shape orientation of four silicon modules was chosen. The configuration covers 79% of the azimuthal angle ϕ. The acceptance in θ is between 23° and 80°. A drawing of the SRD, together with the target cell is shown in Figure 5.6.

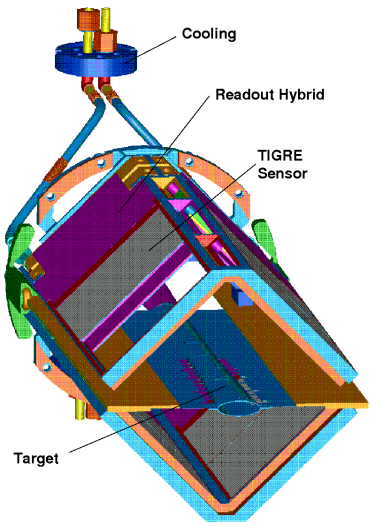

Figure 5.6: Drawing of the Silicon Recoil Detector. Two times four modules are arranged in a diamond shape around the target cell.

After having decided on the thickness and the amount of layers for the SRD, the energy deposition from low energy protons could be computed. According to a calculation detailed in Reference [124], the energy deposition in the first layer versus the energy deposition in the second layer for protons hitting the detector perpendicularly is plot in Figure 5.7. Protons with momenta lower than ~ 106 MeV get stuck in the first layer of silicon. Protons with momenta between ~ 106 and ~ 130 MeV punch through the first layer but get stuck in the second layer. If the momentum of a proton is higher than ~ 130 MeV it penetrates both silicon layers. The momentum resolution is best for particles with momenta below 130 MeV, and accounts about 3 %. For particles below 0.5 GeV the momentum resolution is less than 10 %. The resolution in t is less than 15 % for $-t < 0.02$ GeV^2 [125]

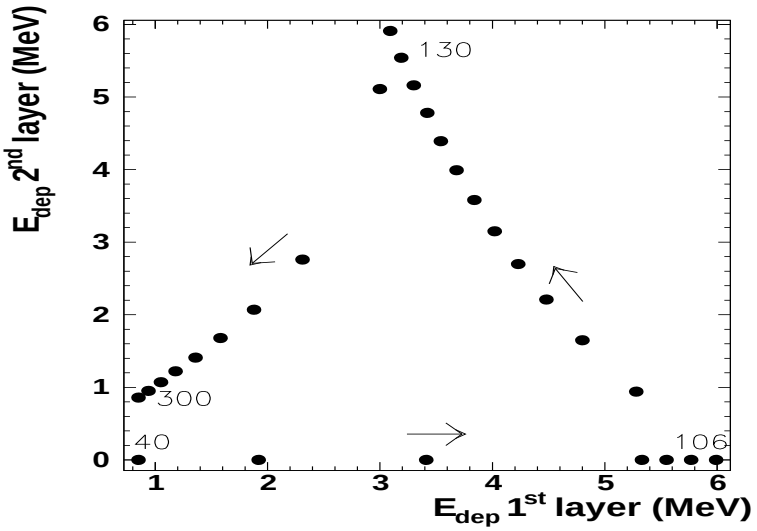

Figure 5.7: Energy deposition in two layers of 300 μm silicon for protons hitting the silicon perpendicularly. The arrows indicate increasing proton momenta. Protons with $P \leq 106$ MeV get stuck in the first layer, if $106 \leq P \leq 130$ MeV then the protons punch through the first layer but get stuck in the second. For $P \geq 130$ MeV the protons punch through both layers.

According to Figure 5.7, the largest energy deposition expected is about 6 MeV. The energy deposited and the corresponding charge created by such a particle is about 70 times the energy deposited by a MIP (~ 85.5 keV). Maximum charges of about 270 fC are expected. In order to be able to measure the whole energy deposition range from MIP to slow proton, an SRD module needs to respond linearly to input charges from 0 to about 270 fC, while having a sufficiently high signal to noise ratio for MIP particles. This demand of linearity is one of the most challenging of the entire recoil project.

Since the SRD modules are placed very close to the beam line, special care needs to be taken concerning the beam-induced RF-field. A sizable RF-field could prevent the SRD from operating. Therefore the target cell itself, and the connections to the HERA beam pipe were redesigned to reduce the RF field as much as possible. For these studies the beam was simulated by applying electric RF-signals through a dedicated aluminum rod mounted inside the target cell. Results of the tests and the according design changes are reported in Reference [126].

5.4 Readout chips

The choice of the appropriate readout chip is a very important decision in the design of any silicon detector. The chip's clock frequency should match the beam clock, the radiation tolerance should be adequate, and the gain and noise properties of the chip should comply with the expected signal. For the SRD two candidate chips were available, the APC and the HELIX chip. Both chips have a charge sensitive preamplifier suited for silicon readout and were already in use at HERMES for the readout of the Vertex Chambers (APC) and the Lambda

Wheels (HELIX). Moreover, both these chips are used in the other experiments located at the HERA ring, namely in H1 for the APC, and ZEUS (and HERA-B) in case of the HELIX.

As the SRD is designed to measure energies of recoiling protons, and since determining space-points for minimum ionizing particles would be beneficial, the dynamic range of the readout chip has to be large. This section will detail the properties of the readout chip, with first the selection of the best suited chip out of the two candidates, which was decided to be the HELIX chip. Therefore the HELIX is described in detail in the second subsection. A discussion on its programmable settings and the selection procedure of valid chips for the experiment follows.

5.4.1 Decision on the readout chip

Both the APC ([127],[128]) and the HELIX ([129], [130], [131]) readout chips have been considered for use with the silicon detector. The APC has the advantage that the gain of the input preamplifier can be adjusted in 2 steps. The dynamic range of the APC is about ± 8 MIP[4], i.e. 8 times the charge deposited by a MIP in silicon ($= 24,000$ electrons). The silicon test counter, a detector installed for a feasibility study of the SRD [132], demonstrated that using the low gain setting, signals corresponding to 6 MeV could be detected [133]. Detection of minimum ionizing particles was, however, not possible using the low gain settings due to the noise of that setup. The disadvantages of using the APC was that the control electronics would have to be designed and built, whereas for the HELIX the experience of the Lambda Wheels [134] could be used. The HELIX readout chip has the disadvantage that the gain of the preamp cannot be adjusted. The chip is linear over a range of about ± 10 MIP. The linear range of the HELIX is clearly not sufficient to comply with the goals of the SRD. However, if a design to reduce the effective gain could be realized, then the chip would be advantageous given the experience within the HERMES experiment. Tests on the APC chip were performed in the framework of this thesis in DESY Zeuthen using the H1 electronics and a charge injection setup. The HELIX chip was studied using the electronics and test stand from the ZEUS Micro Vertex Detector (MVD).

A possible way to extend the HELIX dynamic range is to distribute the charge coming from the silicon unevenly over two HELIX channels. This can be achieved by capacitive coupling: a capacitor placed between two HELIX input channels will divide the charge unequally between channels, resulting in a different effective gain of both channels. The method is shown in Figure 5.8: the signal charge from the detector is divided by a coupling capacitor into a high and low gain readout channel.

The construction and performance of the charge dividing setup was tested using a charge injection setup, analog to the one which will be discussed in section 5.4.3. The response of two channels, coupled through a 10 pF capacitor to an injected positive charge can be seen in Figure 5.9. It can be concluded that by using the charge division configuration, the dynamic range is extended sufficiently. The range can be optimised by varying the value of the coupling capacitor.

The investigation of the linear 'dynamic' range of the APC128 was performed using an H1 based setup [135]. The APC was controlled using a decoder chip and an OnSiRoc module,

[4]To avoid confusion: ± 8 MIP means linear from -8 to $+8$ MIP

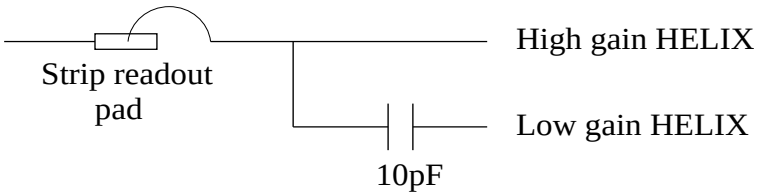

Figure 5.8: Circuit diagram for the readout with charge division. The capacitor value can be chosen to adjust the dynamic range.

which is a VME based module operating as a pattern generator and as an ADC. Charge was injected in an analogue way as for the HELIX tests.

As can been seen from Figure 5.10 the linear range of the APC in low gain mode is between -30 and 70 MIP. One also sees from the figure that if a look up table would be used to relate the ADC response to the deposited charge, the dynamic range could be extended down to roughly -60 MIPs. The data shown were averaged over 500 events. If switched to the high gain mode a range from -8 to $+8$ MIPs is observed. The charge injection tests showed a clear dependency of the linear region on the value of the reference voltage applied to the chip's preamplifier.

Because of the advantage that the HELIX is already implemented in the Lambda Wheel detector of the HERMES experiment and because of the reliable charge division method for the HELIX it has been decided to use the HELIX 128-3.0 as a readout chip for the Silicon Recoil Detector. More details about the chip selection can be found in Reference [124].

5.4.2 The HELIX chip

The HELIX chip is a typical front end readout chip. It was designed for the HERA-B experiment, which was also located at the HERA accelerator facility. The clock frequency was designed to match the HERA bunch crossing of 10 MHz. The HELIX is schematically depicted in Figure 5.11. The chip itself consists of 128 channels which all have a charge sensitive preamplifier, followed by a CR-RC shaper. The latter is needed to have the signal shaped within 1 clock cycle of 100 ns. After the shaper there is a condensator array, referred to as a pipeline, where the signal can be stored until readout. Every clock cycle a different pipe cell is addressed. The pipeline depth is $128 + 8$, so the maximal delay between signal and trigger is over 12 μs. The chip has a multi-event buffer of 8 events. Whenever a trigger occurs, the charge stored on one of the pipe cell condensators is read out via the pipe amplifier. The analog signal from all 128 channels then is serialized to form the 'AnalogOut' signal which is sent to the external world. The analog signal is followed by a trailer of eight bits giving information about the pipe-cell where the charge was stored.

There exists the possibility to daisy chain several chips together, to extend the amount of input channels from 128 to any multitude of that number. The analog out of all chips available in a daisy chain are serialized via a token communication between several chips. A possible broken chip can be excluded from the daisy chain via a fail-safe mechanism.

The HELIX 3.0 also features a comparator which can be used to trigger on the analog input. If the output of the preamplifier-shaper combination is above a certain reference voltage

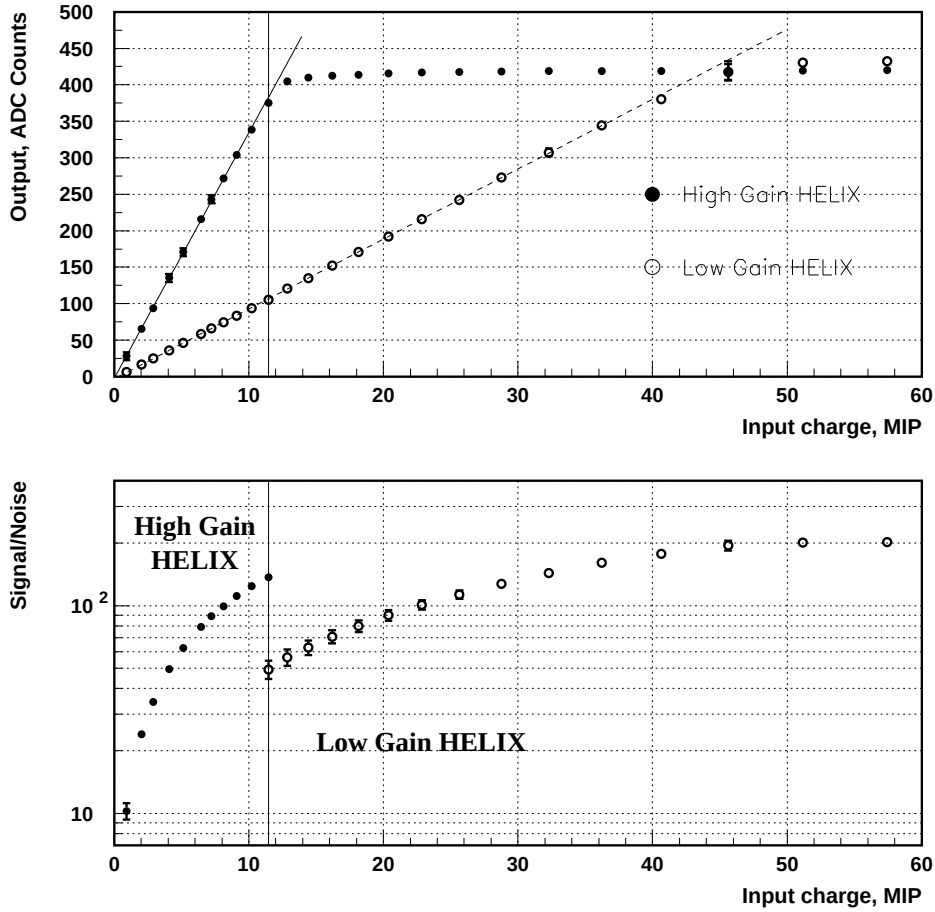

Figure 5.9: Responses of high gain and low gain HELIX channels coupled through a 10pF capacitor. Upper panel: HELIX dynamic range. Lower panel: signal to noise ratio. Once a high gain channel is saturated, the signal from a low gain channel is used.

a trigger is generated. Triggering on the signal from the silicon sensors never was a goal for the SRD. This feature consequently has never been used within this project.

The left panel of Figure 5.12 shows the difference of individual pipe-cell pedestal levels from a channel's average pedestal. The double peak structure is an indication that even and odd channels have a slightly different pedestal level, due to the construction of the chip. The right panel shows the same observable, in a 2 dimensional way. The measurements were done with the probe card setup which will be described in section 5.4.4. Since all pipe cells were addressed randomly during the measurement, no systematic effect can explain the observed phenomenon.

The double peak structure has also been observed during the initial tests with the ZEUS Micro Vertex Detector (MVD) hybrid, which also houses the HELIX 3.0 chip. However, it has never been observed in a measurement with the Silicon Recoil Detector hybrid, indicating that the noise on the SRD hybrid is higher than the intrinsic pedestal difference between two pipe cells of a particular channel.

During initial charge injection tests with a SRD prototype hybrid it became clear that the shaping time for the low gain channel was longer than the shaping time for the high gain signal.

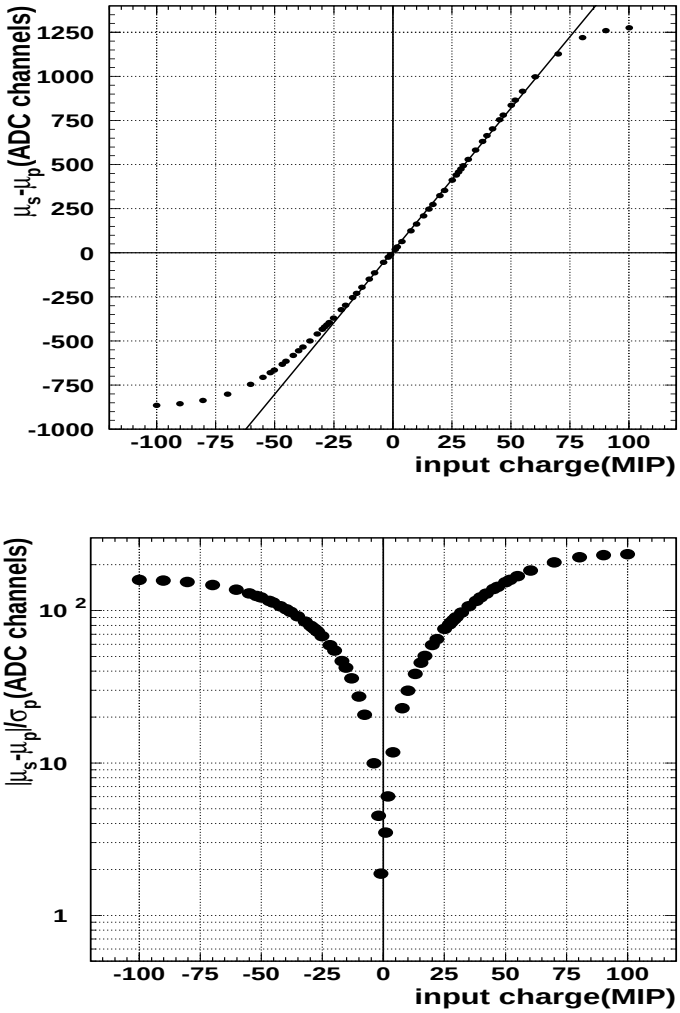

Figure 5.10: Upper panel: dynamic range for the APC128 in low gain mode with a reference voltage of 1.96 Volt to the preamplifier. $\mu_s(\mu_p)$ is the mean value of a Gaussian fit to the signal (pedestal) level of 500 (9500) events. A straight line was fitted to part of the data. The lower panel shows the signal to noise ratio measured as a function of the input charge.

This behavior can be seen from Figure 5.13 and is due to the different input capacitance of high and low gain channel due to the charge dividing capacitor.

5.4.3 Parameter selection

As seen from Figure 5.11 the HELIX chip requires various input voltages and currents to ensure its operation. These can all be varied to an optimal value specific to the project where the chip is to be used. The most important programmable parameters are those of the preamplifier and shaper. These can be adjusted by varying their bias current ($I_{\mathrm{pre(sha)}}$) and feedback transistor voltage ($V_{\mathrm{fp(fs)}}$). It was important to have a setup able to measure the shaper output to determine the optimal timing of the HELIX output. The timing is adjustable

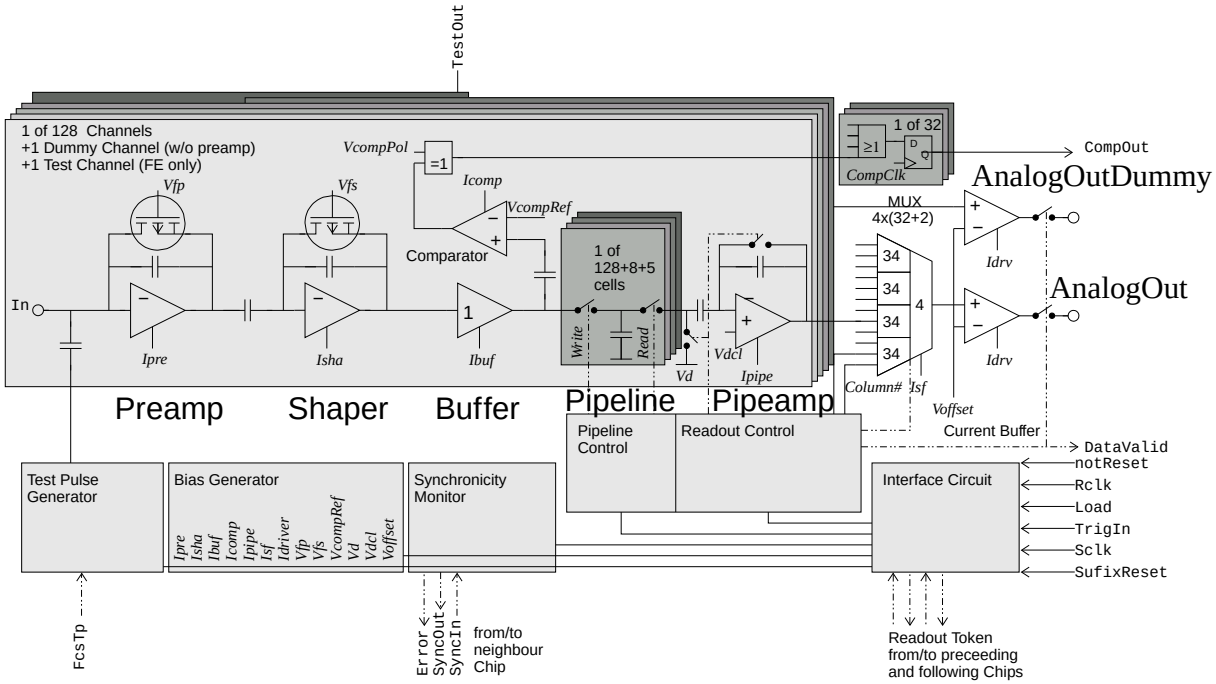

Figure 5.11: Schematic overview of the HELIX chip. Figure taken from Reference [131].

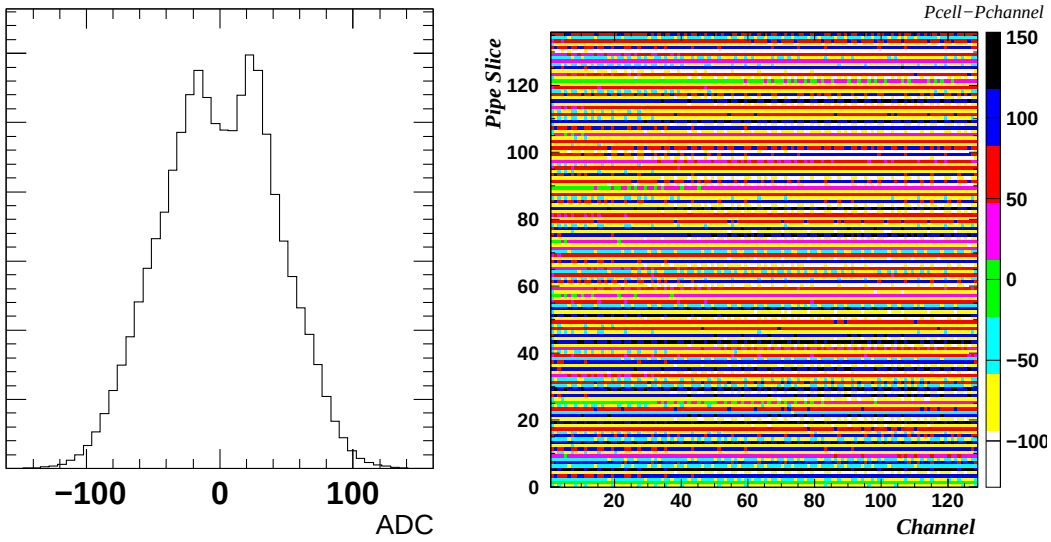

Figure 5.12: Difference of individual pipe-cell pedestal levels from a channel's average pedestal. Left is a one dimensional plot of the pedestal level of the entire chip. On the right side a two dimensional figure showing the pedestal level of every single capacitor in the HELIX pipeline. The figure shows the 'memory' of the chip. The double peak on the left and the red-yellow sequence on the right indicate a pedestal difference between even and odd pipe cell numbers. Plot taken from Reference [136].

via the aforementioned parameters.

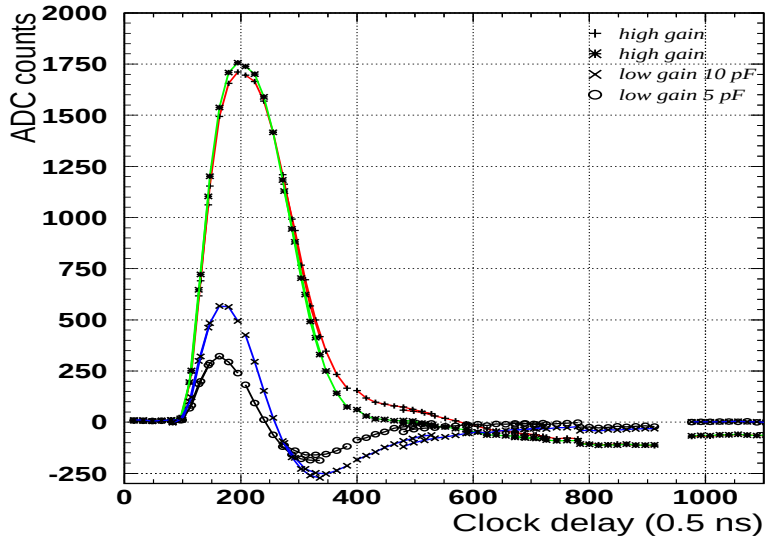

Figure 5.13: Output of the shaper for four channels. Two of them are high gain, and two low gain. + and × (∗ and ∘) have a charge division capacitor of 10 pF (5 pF). A difference in shaping time can be observed.

In the framework of this PhD work a test stand complying with the aforementioned goals was developed. The setup makes use of a charge injection hybrid, depicted schematically in Figure 5.14. A voltage step is connected over a 50 Ω resistor and a capacitor (C_{in}) to a hybrid input channel, leading to a high and low gain HELIX input channel. An additional load capacitor (C_{load}) was added to simulate the total capacitance of the TIGRE sensors. This charge injection setup is converting a voltage step into a charge. An attenuator was used to adjust the voltage step and by consequence the amount of charge injected into the system. The same charge injection setup was used for the chip selection tests mentioned before[5].

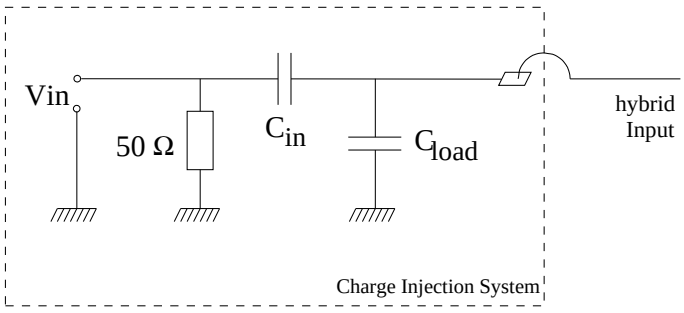

Figure 5.14: Schematic of the charge injection setup.

The controlling electronics for the shaper measurements is depicted in Figure 5.15. The setup's core is a 20 MHz clock, which is used in combination with a prescaler to delay the trigger to the charge injection system (indicated 'charge trigger' in the figure) in steps of 50 ns. This steps are further divided into steps of up to 0.5 ns when making use of the HLCU's

[5]Although C_{load} was absent in these tests

(see section 5.6.1) trigger delay. The trigger to the charge injection system needs to have a fixed timing with respect to the beam clock. Hence a coincidence with the beam clock was implemented.

The same 20 MHz clock is prescaled by a factor two to get a 10 MHz clock. A discriminator was used to get a 50 % duty cycle for the clock. This 10 MHz clock serves as an HLCU clock and - after being prescaled over 1000 times - as a trigger for the HLCU and the charge injection system. The trigger to the HLCU needs to be in coincidence with the output of the Input-Output (IO) register, the latter needed to trigger the readout.

The Input-Output register makes use of a flip-flop mechanism to ensure outputs synchronized with the 10 MHz clock. The flip-flop is set via the active phase of a prescaled clock, and reset via software. An output is generated after every active clock phase following a software reset. The flip-flop is indicated in Figure 5.16. Another output of the IO register, indicated in the top of the figure, serves as a veto to the charge injection trigger, in order to measure the pedestal. After a trigger is sent to the charge injection system, its output is vetoed during $22\ \mu$s to prevent multiple triggers.

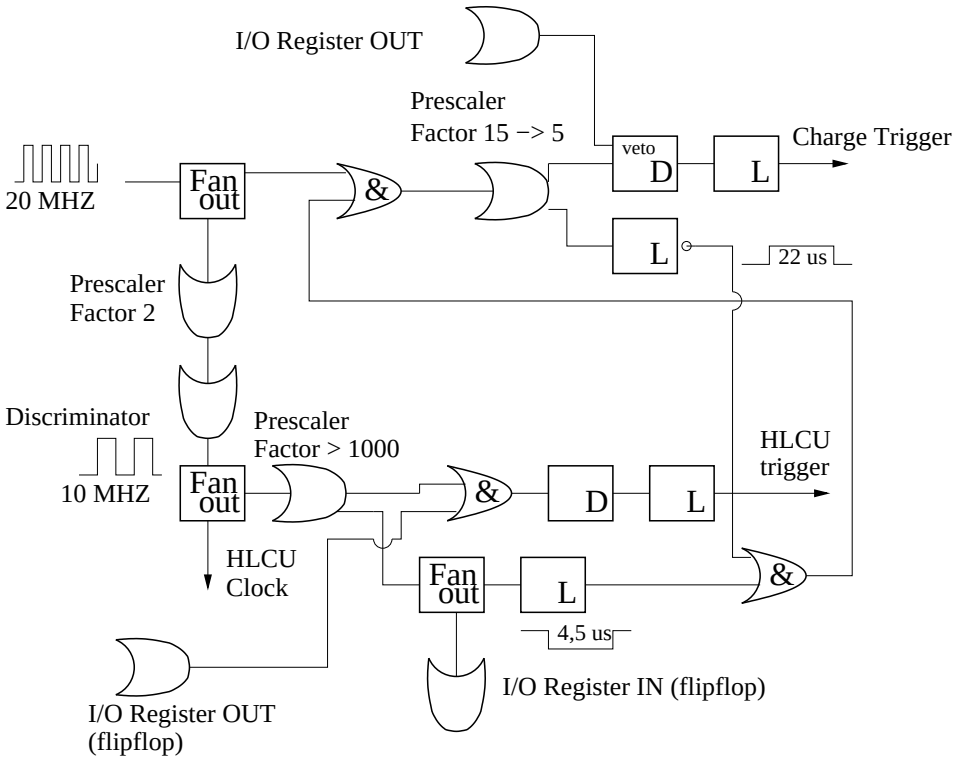

Figure 5.15: The NIM logic for the shaper measurement as explained in the text. **D** is a delay generating module, **L** adjusts the pulse length.

A user-friendly LabView program was written to control all steps of the measurement. The program also extracted the mean and sigma of signal and pedestal distributions, in order to provide all useful information directly after the test. A user only needed to specify a particular HELIX parameter and the number of events to be measured, all other issues were automatically taken care of.

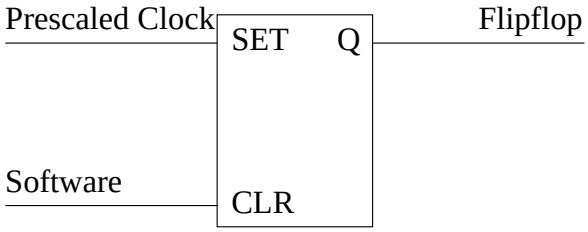

Figure 5.16: Schematic of the flip-flop mechanism for triggering: The flip-flop's output is set via a prescaled clock, the unit is reset via software.

Results of the tests for two different settings of V_{fs} in combination with a variation of I_{pre} are shown in Figure 5.17. Ideally one would like to have a similar rise time for the high gain and low gain shaper output since digitization happens at one particular moment in time. The proposed digitization time is indicated by the red line in the figure. Moreover a fast falloff is required, as no additional signal should leak into a different time slice. 288 (77) fC of charge was injected in the low (high) gain, in combination with a C_{load} of 33 pF. High V_{fs} in combination with high I_{pre} yield the best results.

In order to check the influence of any leftover charge from a possible previous event, a double voltage step was injected into the system in an additional measurement. This was implemented by using an arbitrary waveform generator. The time between two charge injections was two clock cycles (192 ns). As can be seen from Figure 5.18 both charges yield a similar amplitude for $I_{pre} \geq 160$ for the high as well as for the low gain channel, indicating that the shaper's tail is sufficiently small. More details about these tests can be found in References [137] and [138].

5.4.4 Selecting good chips

The HERMES SRD project needed 8 modules, 2 spares and one prototype. Therefore 104 HELIX chips were needed. To reach this number, a total of 6 HELIX chip wafers was ordered, resulting in 372 chips, all of which have been tested in a dedicated test stand.

All chips were probed via a probe card, accessing to all necessary HELIX pads. In order to make debugging of the setup easier, an additional fully functional reference chip housed by a dedicated separate driver card, was added. The availability of the reference chip also allowed testing of the various tokens of the daisy chain mechanism.

The chips were tested in three steps: first a selection was made based on the online chip test stand information, next, several other checks were made offline. All chips passing this second step were considered fully functional. From these 153 chips, the best 104 were selected via a dedicated analysis.

All programmable parameters were checked in the first test: the addressing was varied via an IO register, the different current and voltage registers were tested by programming them to various values and displaying the variation of the power consumption. The offline tests checked the pedestal and noise level of every individual pipe cell of the HELIX's memory. Moreover, the amplification was tested by analyzing the response to the HELIX internal test-pulse, which can inject ± 2 or $\pm 1 \times 24\,000$ electrons in a channel depending on whether the channel is even or odd.

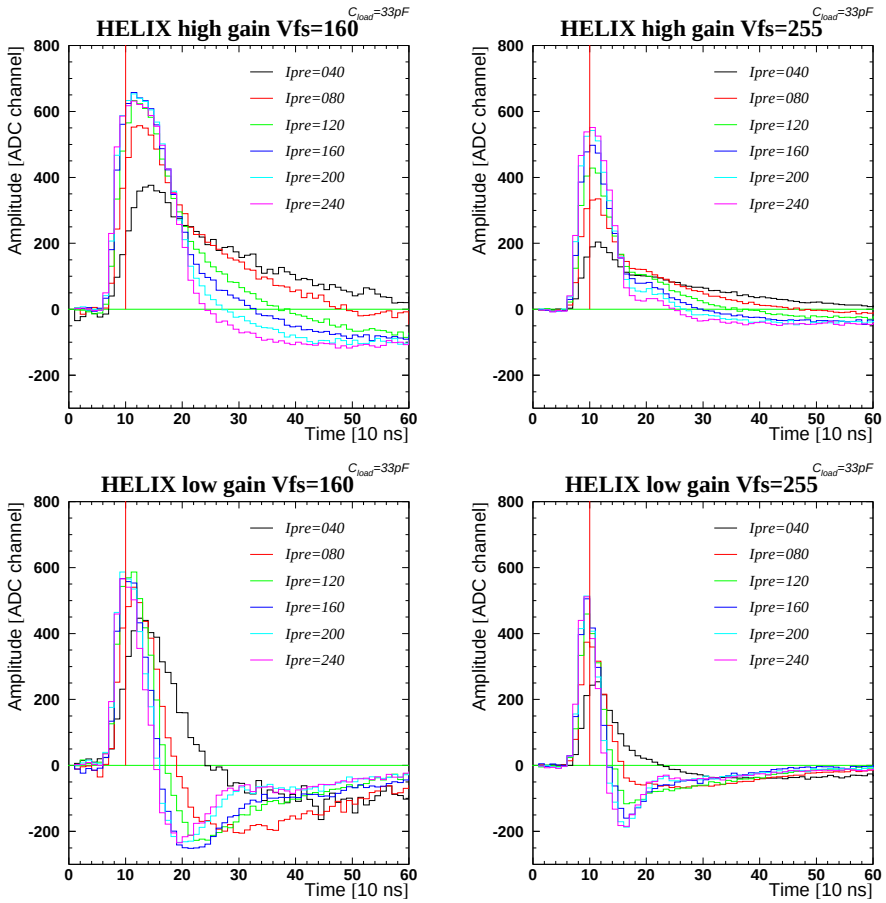

Figure 5.17: Shaper output for various V_{fs} and I_{pre}. The top (bottom) panel shows the high (low) gain channel's response to 77 (288) fC of charge. The red line at 10 ns indicates the proposed digitization time. High I_{pre} and high V_{fs} yield a maximal signal in both high and low gain chips.

To select the 104 best chips to be used in the project, a dedicated weighting method was developed: a chip got a certain score, based on its average chip noise (`acn`), its pedestal pipe cell spread (`ppcs`: the spread of the pedestal level of all pipe cells), and the amplification pipe cell spread (`apcs`). The `apcs` got a weight factor of 3, as the project requires small gain differences for every channel. The `acn` got a weight of 2, since noise is a setup dependent observable, and might be different in the HERMES experimental area. `ppcs`, finally, got a weight of 1, as this property could be calibrated out by performing dedicated pedestal runs and storing individual pipe cell pedestal information.

The best 52 chips were selected to serve as n-side chips, since due to the higher sensor capacitance on that side (see later), the collected charge is less than the charge collected on the p-strip. Apart from a chip's total score, hard cuts were imposed on the `acn`, `ppcs`, and `apcs` distributions. All details of the chip selection procedure can be found in Reference [139].

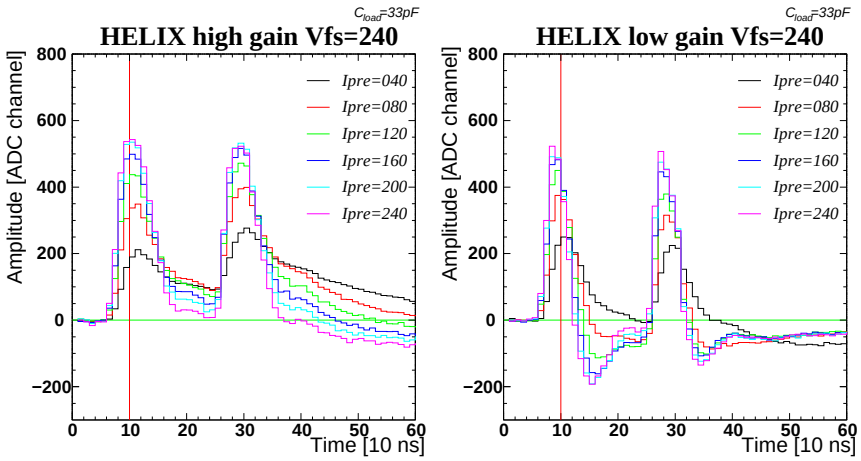

Figure 5.18: Shaper output for two consecutive charge injections with a time difference of 2 clock cycles for various values I_{pre}. The left (right) panel shows the high (low) gain channel's response to 77 (288) fC of charge. The red line at 10 ns indicates the proposed digitization time. The amplitude of both pulses is similar for both high and low gain chips.

5.5 Module design

After discussing the overall SRD design and the readout chip, the details of the construction of a single SRD module are discussed here. It must be stressed that all module components need to be vacuum compatible. First, the sensors itself are discussed, then the holding frame, the kapton flex-foils, and finally the readout hybrid. Figure 5.19 shows the p- and n- side of a SRD module.

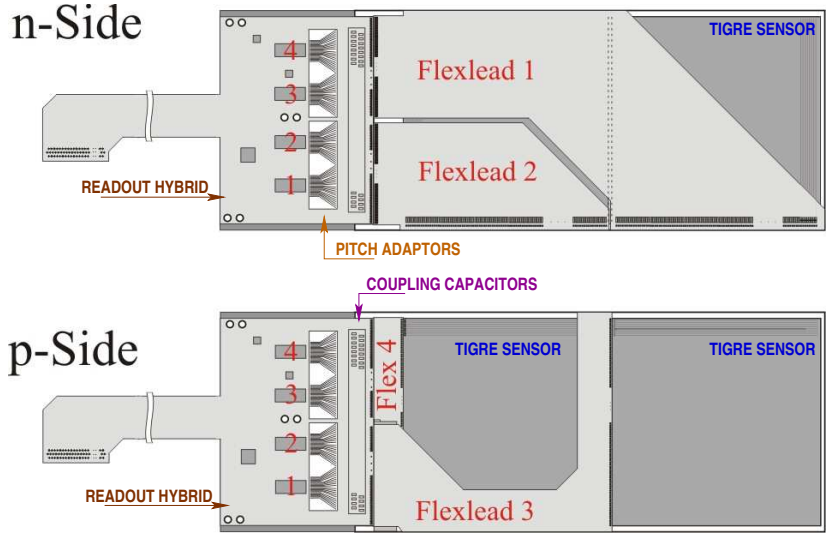

Figure 5.19: p- and n- side of a SRD module. Sensors, flex-foils (indicated as flexleads), coupling capacitors, pitch-adaptor, and readout hybrid are clearly visible. The HELIX chips are indicated by the numbers 1 to 4 on each sensor side.

5.5.1 TIGRE sensors

As mentioned before, the TIGRE sensor from Micron Semiconductors Ltd. was chosen as the silicon sensor for the SRD project. The TIGRE was designed for the Tracking and Imaging Gamma Ray Experiment. It is the largest existing double sided silicon sensor with an area of $9.9 \times 9.9\ \mathrm{cm}^2$, and a thickness of $300\ \mu\mathrm{m}$. The sensor has 128 strips, with a pitch of $758\ \mu\mathrm{m}$, and a width of $702\ \mu\mathrm{m}$. The TIGRE sensor's p- side strips are oriented perpendicular to the n- side strips. The voltage required to fully deplete a sensor is around 50 ± 10 V.

To ensure a uniform potential difference over the entire sensor, the depletion voltage is applied over a so called 'bias' and a 'guard' ring which surround all strips. This technique is common practice in the construction of silicon sensors. A poly-silicon resistor is added in between the silicon bulk and both the guard and bias bond pad. The bias voltage is applied to the sensor via the bias bond pad. The original bias resistor value of 50 $M\Omega$ was changed to 7 $M\Omega$. This change reduced the overall RC-time constant from 2.5 to 0.35 ms (given a strip capacitance of about 50 pF). With a time constant on the order of a ms, single strip hit rates up to a kHz can be processed without a charge build-up in the silicon.

The whole area of the sensor is covered with an SiO_2 layer to ensure AC-coupling of the signal. By doing so, the quasi constant leakage current is eliminated from the readout. An aluminum grid covering each strip individually transports the induced signal from the strip to the bond pad, where the connection to the external world is to be made.

All sensors were tested with a probe station, where the bias resistor was measured, and some overall I/V and C/V characteristics, like the depletion voltage, leakage current, and strip capacitance. An example of the latter measurement is given in Figure 5.20. The slope change in the figure corresponds to the situation where the sensor is fully depleted. The leakage current of all sensors was tested in a long term test, where the sensors were put under bias voltage for several weeks. More details about this test can be found in Reference [125].

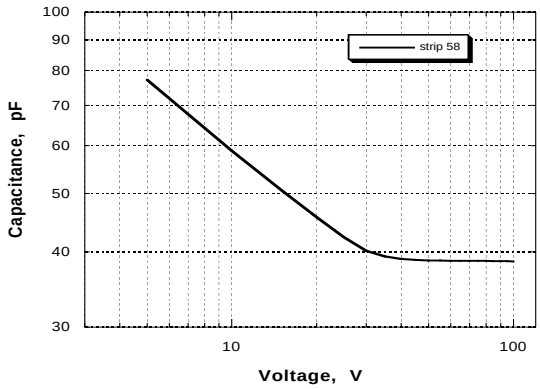

Figure 5.20: Strip Capacitance as a function of bias voltage. The slope change corresponds to the point where the sensor is fully depleted. Figure from Reference [140].

A silicon sensor can in general be seen as a combination of conductors and insulators giving rise to various internal capacitive couplings: coupling to the back plane, to neighboring strips, and evidently the coupling between the strips and the aluminum readout grid. Only a fraction of the original signal will be available on the aluminum grid for readout. Capacitance for p- and n-side are given in Table 5.1. These are different for both sides, yielding a higher signal

on the p-side than on the n-side. A schematic overview of the various internal capacitance in the TIGRE sensor is given in Figure 5.21.

	p-side	n-side
Strip Capacitance C_{STR}	34 pF	54 pF
Inter-strip Capacitance C_{int}	9 pF	7 pF
Coupling Capacitance C_c	1 nF	1nF

Table 5.1: Overview of various capacitances of a TIGRE sensor.

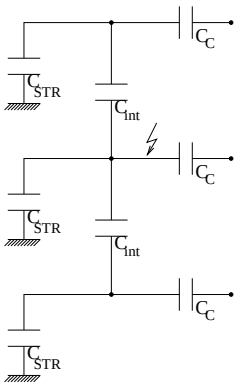

Figure 5.21: Schematic view of the capacitance inside the TIGRE sensor. Three strips are shown. When charge is created close to the middle strip, some of it is coupled to the back-plane, and to some neighboring strips.

Taking into account the various TIGRE capacitance, only about half of the signal is collected on the aluminum readout grid. This is advantageous to extend the dynamic range. For the detection of MIPs, however, these large capacitance imply that only about $12\,000$ electrons will reach the amplification stage, which severely complicates the detection of MIPs.

5.5.2 Holding frame

The material used for constructing the sensor's holding frame should be chosen carefully. Its modulus of elasticity, thermal conductivity, and thermal expansion coefficient should be close to those of silicon, since mechanical stress caused for example by a temperature rise should be avoided. Moreover, the material needs to be vacuum-compatible, and it preferably should be an insulating material. The need for vacuum compatibility reduces greatly the selection of possible materials.

Physical properties of aluminum, graphite and Shapal-M, an aluminum-nitride ceramic, are given in Table 5.2 together with those of silicon as a reference. From the table it can be concluded that Shapal-M's physical properties matches those of silicon best.

Two sensors are glued into a frame with a two component epoxy glue. In order to avoid the glue being spread over the flex-foils, the use of small containers in between foil and sensor was required. These 'preforms' force the glue to stay localised.

	silicon	aluminum	graphite	Shapal-M	Unit
Resistance	$10^{-4} - 10^{4}$	$5 \cdot 10^{-6}$	0.02	10^{12}	$\Omega \cdot \mathrm{cm}$
Thermal Expansivity	2.6	23	7.4	4.4	$\frac{10^{-6}}{\mathrm{K}}$
Modulus of Elasticity	170	70	15	160	GPa
Thermal Conductivity	150	130	65	100	$\frac{\mathrm{W}}{\mathrm{m \cdot K}}$
Outgassing Rate	n.a.	10^{-10}	$8 \cdot 10^{-11}$	$2.3 \cdot 10^{-11}$	$\frac{\mathrm{mbar \cdot l}}{\mathrm{s \cdot cm^2}}$
Costs per frame	n.a.	30	110	1200	€

Table 5.2: Physical properties for silicon, aluminum, graphite and Shapal-M.

To get the signals from the sensors to the readout hybrid, flex-foils containing conducting metallic traces were needed. The flex-foil itself is made out of kapton. Individual strips are bonded to metallic traces on the foils. The implementation of flex-foils in the design of the SRD leads to additional material between the silicon sensors and the HERMES target cell, worsening the momentum resolution. The design of the flex-foils has been optimised using a Monte Carlo simulation [141]. A design with $50\ \mu\mathrm{m}$ kapton, and traces built out of $5\ \mu\mathrm{m}$ copper, $5\ \mu\mathrm{m}$ nickel, and $0.1\ \mu\mathrm{m}$ gold was chosen. The flex-foils are indicated in Figure 5.19 as 'flexleads'. Conducting traces on flex-foil 1 and 2 bring the signals from the n-side to the readout hybrid. Strips on the n-side are oriented vertically. Strips on the p-side on the other hand are oriented horizontally. Traces on flex-foil 3 conduct the signal from the sensor furthest away from the hybrid. The flex-foil design was optimized to cover as few as possible of the TIGRE's surface.

5.5.3 Readout hybrid

A readout hybrid is required to provide signal amplification and communication to the external world. It houses the readout chips, mandatory for signal amplification as well as necessary pitch adaptors to match the different bond pad pitch of the readout chips and the kapton foils. The charge division capacitors are mounted on the hybrid as well. Additionally, the hybrid contains some receivers and line drivers to provide a solid connection to the outside world. Underneath the hybrid an aluminum heat sink is placed to dissipate the heating power of the active components on the hybrid. A picture of the SRD hybrid can be seen in Figure 5.22.

5.6 DAQ chain

The controlling and readout electronics for the SRD are based on the Lambda Wheel setup. The 'Hermes Lambda Control Unit' (HLCU) serves as a controller for the HELIX, and the 'Hermes ADC' (HADC) is the ADC for the SRD. Both these modules are VME based and were designed by NIKHEF, Amsterdam. As the HADC and HLCU are located in the HERMES electronics trailer, which is located about 30 meters from the interaction point, a dedicated repeater board was designed. This board is called 'Analog Clock Control' (ACC) which is located close to the SRD itself. An additional functionality of the ACC is the distribution of the required low voltage and bias voltage to the sensor. Moreover, it houses the interface

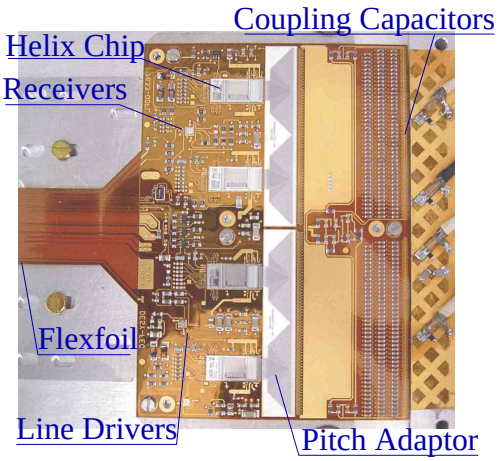

Figure 5.22: picture of the SRD readout hybrid, with the various components indicated in the figure.

to the radiation and temperature sensors on the hybrid. These modules are discussed in the next sections. The DAQ chain is depicted in Figure 5.23 where an SRD module is drawn as a diode with its two bias resistors. The AC coupling to the preamplifier for n- and p-side is also indicated in the figure.

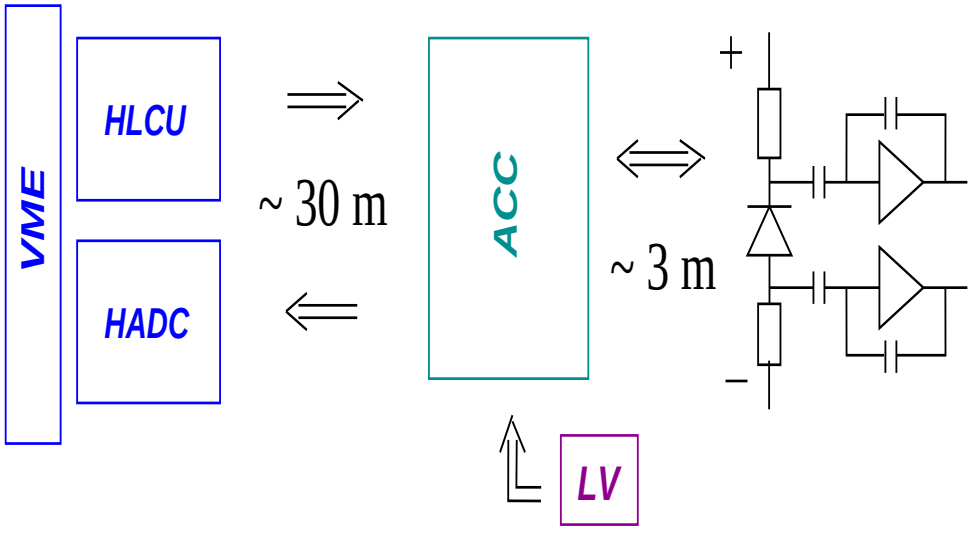

Figure 5.23: Overview of the DAQ chain for the Recoil Silicon Detector. The VME based modules HADC and HLCU are located in the HERMES Electronic Trailer (ET). The ACC is positioned close to the HERMES target platform and is sending and receiving the signals to and from the ET. It also provides the low voltage (LV) to the HELIX chips of the SRD. The HELIX's preamplifier and the silicon sensor diode itself, together with its bias resistors and the AC coupling, are indicated schematically in the figure.

5.6.1 Controlling and programming – HLCU

The HLCU provides the clock and trigger signals to the HELIX chip. Its outputs are organised in two groups of 8 signals: 'current' and 'TTL' signals. The 'current' group contains the LVDS ('Low Voltage Differential Signaling') signals. Clock and trigger for both n and p side are of this type. The 'TTL' group contains the TTL ('Transistor-Transistor Logic') signals 'reset', 'Sload', 'test-pulse n', and 'test-pulse p' with their respective grounds. Up to four SRD modules can be connected to the HLCU via its 4×2 RJ45 outputs. The HLCU has four inputs: a clock, trigger, reset, and test-pulse input.

Programming of the HELIX chip occurs over the trigger line. The chip's various programmable parameters, like I_{pre} and V_{fs} are assigned a register number. The first 6 bits of the pattern sent to the chip to program a particular register indicate the address of the chip to be programmed, the next 5 bits contain the register number, succeeded by 8 bits containing the value to which that register should be programmed. The maximum value of every bit thus is 255. An example of this sequence is shown in the left panel of Figure 5.24, where register 2 of chip with address 27 is programmed to a value of 83. An additional Sload signal is necessary to upload the values into the chip. During programming the reset signal should be switched on, in order not to confuse the bits on the trigger line with real triggers.

The HLCU also contains several programmable delays: the clock delay between the incoming and outgoing clock and the clock difference between p and n side, needed to cope with the different mobility of electrons and holes in silicon. Furthermore, the test-pulse and trigger can be delayed. All delays are programmable in steps of 0.5 ns. The delays are needed to get an optimal signal out of the sensors.

5.6.2 Module readout – HADC

The HADC module digitizes the analog signal coming from the HELIX . The serialized output of two HELIX chips is shown as an example in the right panel of Figure 5.24. The low and high gain output can be seen, as well as the trailer bits. On every active phase of the clock, one channel is digitized. Two hybrids can be connected to one HADC.

The HADC module also performs zero suppression and common mode noise calculation. A definition of common mode noise is given in appendix C. The common mode is calculated from the first 16 non-hit channels out of the first 32 channels of a chip. Moreover, the HADC converts the trailers to a bit pattern, and checks whether the bit pattern from all four chips of a SRD hybrid are the same.

5.6.3 Interfacing - ACC

The ACC serves as a repeater board for the signals coming from the HLCU, and has some line drivers to send the HELIX analog out to the HADC in the electronics trailer. The ACC also accepts the $\pm$ 2 V lines from the Low Voltage power supply as well as the bias voltage. Since the Low Voltage directly connects to the preamplifier, it is very important to make the Low Voltage as noise free as possible. All TTL signals needed for the Logic devices on the hybrid are being put between -2 and $+2$ V before being sent to the hybrid. As a consequence the entire hybrid is working between these voltages. Additionally, the ACC also contains the interface for

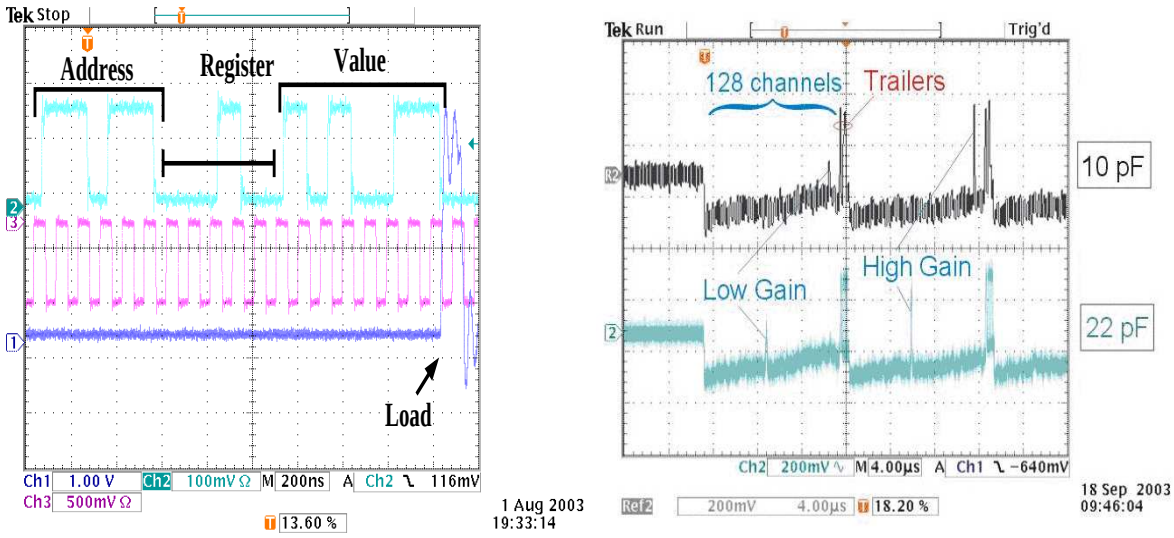

Figure 5.24: The upper line in the left panel shows the trigger-line when a certain chip parameter is programmed. Bits need to be set high or low during the active phase of the clock. The negative LVDS clock is also indicated in the figure. The observed bit pattern is 01101100101010011. The Chip address is 27 (011011), the register being programmed is register 2 (0010)(corresponding to the current to the HELIX shaper (isha)), and its value is 83 (1010011). The 'Sload' signal is also indicated on the figure. On the right hand side the serialized analog output is shown. The trailers are clearly visible.

the radiation and temperature monitors on the hybrid. One ACC module is needed for every SRD module.

5.7 Module performance testing

After having built the first prototype of the SRD, it needed to undergo several tests, in order to see whether the detector requirements were fulfilled. The first module also served as a test module to design and develop various test-stands which were meant to be used for testing and calibrating the final production modules prior to their installation in HERMES. In the frame of this work first extensive testing of the prototype was performed. For the sake of an easier reference, the prototype was baptized 'Hamlet'. A dedicated laser test stand was developed as well as a test beam setup at DESY. These test stands are the subject of the next two sections.

5.7.1 Laser test

A laser test stand has been proven to be a very useful tool for testing the properties of silicon detectors [142, 143, 144, 145]. The advantages of such a test stand are evident: systematic studies can be made without using an accelerator facility and the corresponding allocation of beam time. Moreover, large data samples can be achieved within a relatively short time, and the signal position is a priori known. The laser test stand was used for testing all production

modules. Additionally, the test stand can be used for rapid diagnostics of the quality of the SRD modules. Defects such as dead strips can quickly be searched for by a functional laser test stand.

This section describes the semi-automated setup made to determine the functionality of all SRD modules. The details about the laser test are reported in Reference [146]. Here, a description of the setup is given after a discussion on the decision on the laser wavelength. Next, a description of various measurements performed with Hamlet are mentioned. These measurements indicate on the one hand the functionality of the prototype, and on the other hand they are an indication of the tremendous potential of a laser test stand for simulating ionizing radiation.

5.7.1.1 Interaction of light with silicon

The absorption of photons in a semiconductor depends on the band-gap of the material. A photon with energy larger than the band-gap will create an electron-hole pair, whereas a photon with smaller energy won't be absorbed [142]. In silicon the band-gap energy is 1.1 eV, corresponding to a wavelength of about $1088\ \mathrm{nm}$. The difference between photon and band-gap energy is converted into heat.

A decision had to be made about the appropriate laser wavelength for the construction of a laser test stand for the SRD. Near-infrared light at a wavelength of about $1050\ \mathrm{nm}$ was being thought of as it has the advantage of being close to the silicon band-gap. On the other hand also red laser light at around $660\ \mathrm{nm}$ wavelength was considered, since it has the major advantage of being visible, allowing an easier construction of the setup since the laser spot could actually be seen with the unaided eye.

The reflectance and absorbance of a $200\ \mathrm{nm}\ \mathrm{SiO_2}$, $300\ \mu\mathrm{m}$ silicon crystal and $200\ \mathrm{nm}$ $\mathrm{SiO_2}$ sandwich was simulated as a function of wavelengths for normal incident photons. The result of the simulation is depicted in the right frame of Figure 5.25. This simulation was performed by the Solid State Department of Gent University [147]. Note, however, that if the light would hit one of the aluminum readout paths, it would be completely reflected, creating no charge in the sensor. Only a weak dependence on the wavelength of the incident photons can be seen up to 1250 nm. For wavelengths larger than about 1250 nm no light gets absorbed in the silicon. This wavelength is larger than the band gap of pure silicon due to the presence of the SiO_2 layers.

The penetration depth and mean depth of energy deposition as a function of wavelength are plotted in Figure 5.26, for a silicon thickness of $280\ \mu\mathrm{m}$. It can be seen that red light is completely absorbed, and electron-hole pairs are created very close to the detector surface. Infrared light on the other hand penetrates very deep into the silicon and electron-hole pairs are created along the path of the incident photons.

To summarize, the major difference between red and infrared light is the penetration depth in the silicon. Using infrared light, ionizing radiation is better simulated since electron-hole pairs are created uniformly along the path. However, the response of a silicon detector to a particular laser pulse always needs to be calibrated in terms of ADC channels in order to be able to interpret the results. The response to test beam data could for instance be used for such calibration. If only the functionality of the modules is to be verified, a red laser would be perfectly suitable. Because of these remarks, it has been decided to build the test stand using a red laser with a wavelength of $675\ \mathrm{nm}$ and an output power of 10 mW.

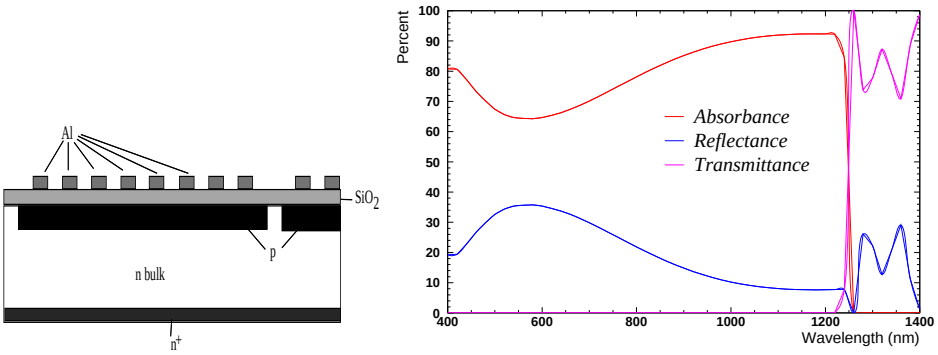

Figure 5.25: Left: a schematic cross section of a TIGRE sensor, on the right absorbance, transmittance and reflectance of a 200 nm SiO_2 – 300 μm Si – 200 nm SiO_2 sandwich, simulated by the Solid State Department of Gent University. At about 1250 nm no light gets absorbed by the silicon.

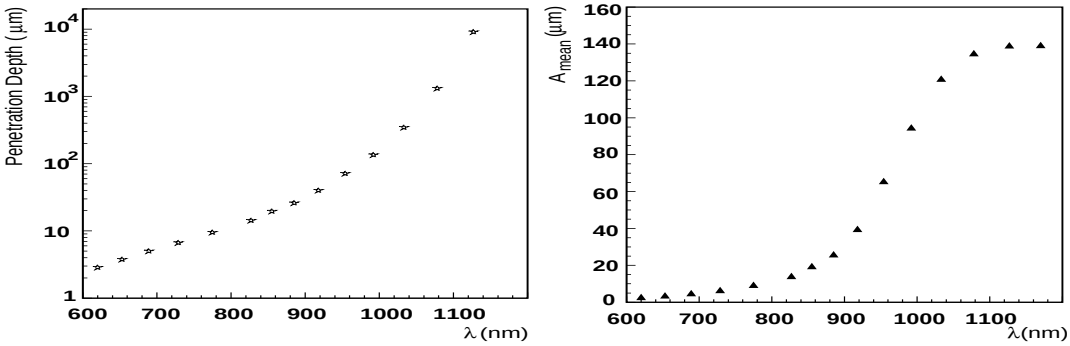

Figure 5.26: Penetration depth (left) and mean depth of energy loss A_{mean} (right) as a function of wavelength. Figure from [144].

5.7.1.2 Description of the setup

The silicon module is placed in a dedicated transport box which is bolted to a translating stage. The stage is controlled by a stepping motor, steerable via a PC's serial port. The whole setup is placed in a black box acting as a Faraday cage. A picture of the module in the black box is shown in Figure 5.27. By moving the table, laser light can be focused onto different areas of the detector. The range of the movable table is 10.5 by 10.5 cm, allowing a full scan of the entire TIGRE area at a time.

The laser diode is driven by a home made unit. By varying the input voltage to this driver, the laser output power can be adjusted. A NIM signal to the controller triggers the laser. The laser's output is transported by a light guide to a micro-focus lens, which focuses the laser light onto a spot of about a micrometer in diameter. The NIM signal required by the laser driver was provided by an IO register using the same flip-flop mechanism as mentioned before. The test-stand used the DAQ chain as mentioned in section 5.6, consisting of an HLCU, ACC and HADC.

A SRD module needs to be mounted on the translating stage with its p- side up, as this

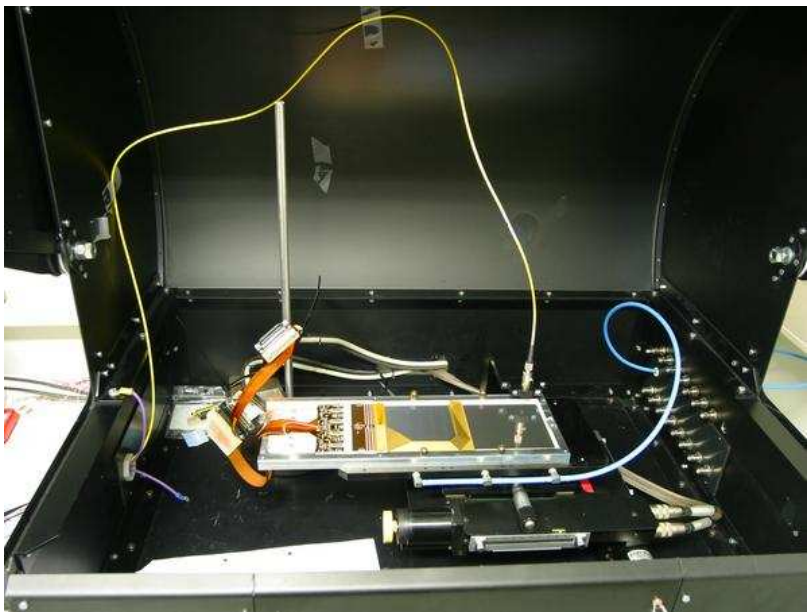

Figure 5.27: Photograph of the laser setup. The yellow light guide and the micro-focus lens are clearly visible, as well as the silicon module in its transport box. The blue hose is connected to a vacuum pump in order to fix the transport box to the table prior to bolting the box to the table.

clearly is the side with fewer flex-foils. Two distinct positions were foreseen to mount the transport box in the setup, each of them allowing one of both TIGRE sensors to be scanned. Since the modules were fully biased the n- side's functionality could be verified in principle in the same scan. This, however, complicated the alignment of the micro-focus with respect to the strips. It was found to be easier to scan p- and n- sides in two distinct measurements. In one measurement the table was moved horizontally, and in the other vertically. In the horizontal measurement the n-side strips were accessed while the light was focused on the same p-strip.

5.7.1.3 Response of Hamlet to the laser test stand

A selection of measurements for Hamlet will be reported here. All measurements described here are performed by injecting laser light to the p-side of the sensor furthest away from the readout hybrid. These measurements are not diluted by the flex-foils.

A full sensor scan was performed by moving the laser to every strip and triggering it 300 times. The mean of the response to the laser pulses for both high and low gain is plotted in the left panel of Figure 5.28. 11 bad strips can be observed, which all correspond to bad strips on the TIGRE sensor, as found during the TIGRE tests[6] [148]. During Hamlet's construction, it was decided not to bond bad TIGRE strips. These unbonded channels obviously do not respond to the injected light.

The right hand side of this figure shows the response of the left neighboring channel x when charge is injected in strip $x+1$ (top) and the response of the right neighboring channel

[6]A sensor with 11 broken strips is perfectly suitable for prototyping, not for a final module.

when strip $x-1$ is fired (bottom). An oscillation is seen in both figures. This behavior corresponds to the asymmetric cross talk behavior of the HELIX chip as is also reported by other groups [149]: if charge is injected into HELIX channel n, a larger fraction of this charge is measured on the neighboring channel $n+1$ $(n-1)$ than on the other neighbor $n-1$ $(n+1)$ depending on whether the injected channel n was even or odd[7]. This effect is due to an asymmetry in the construction of the pipe-cell layout.

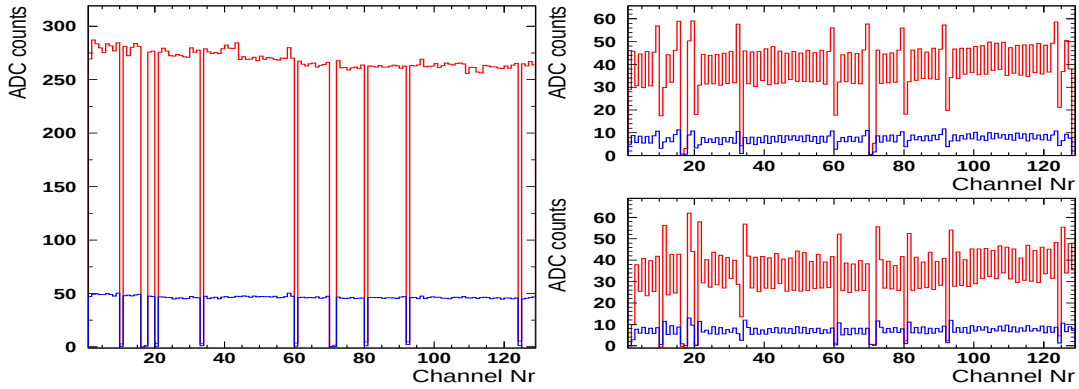

Figure 5.28: Left: The response of high (red) and low (blue) gain channel to the injected laser light for every channel. The upper (lower) right picture shows both for high and low gain the response of a channel when light is fed into the left (right) neighboring strip. 11 dead strips are visible in the figure.

In order to quantify the cross talk behavior, the response of the neighboring channels to the laser pulse was split into *even* and *odd* channels, resulting in two distributions, one with and one without asymmetric cross talk. The splitting into *even* and *odd* channels for the data presented in the upper right panel of Figure 5.28 results into the two upper plots of Figure 5.29. The left frame shows the symmetric+asymmetric cross talk, the right only shows the symmetric cross talk coming from strip-to-strip capacitance as well as stray capacitance of flex-leads and hybrid. Next, both arrays are subtracted from each other, giving an indication of the size of the asymmetric cross talk and is indicated in the lower left panel of the figure. The existence of dead strips leads to an additional difficulty in calculating the average cross talk: some hard cuts were needed in order not to bias the mean. These cuts are shown in blue in the lower right panel of Figure 5.29. The resulting mean of the distribution is depicted as a red line on the two lower plots.

A summary of the cross talk calculations for both left and right neighbor and for high and low gain HELIX is given in Table 5.3. An average value of 5.2% is found, which agrees well with the value of 4.9% as reported for the ZEUS Silicon tracker [149].

To estimate the symmetric component of the cross talk the distribution in the upper right of Figure 5.29 can be evaluated in a similar way as for the asymmetric cross talk. These resulting distributions only contain the cross talk's symmetric component. For the left as well as for the right neighbor these distributions are plot in Figure 5.30. Values for the cross talk obtained by calculating the mean of the distributions are indicated in the figure. The symmetric cross talk is of similar magnitude for both right and left neighbor. The cross talk

[7]Note that predicates 'even' and 'odd' depend on the numbering convention.

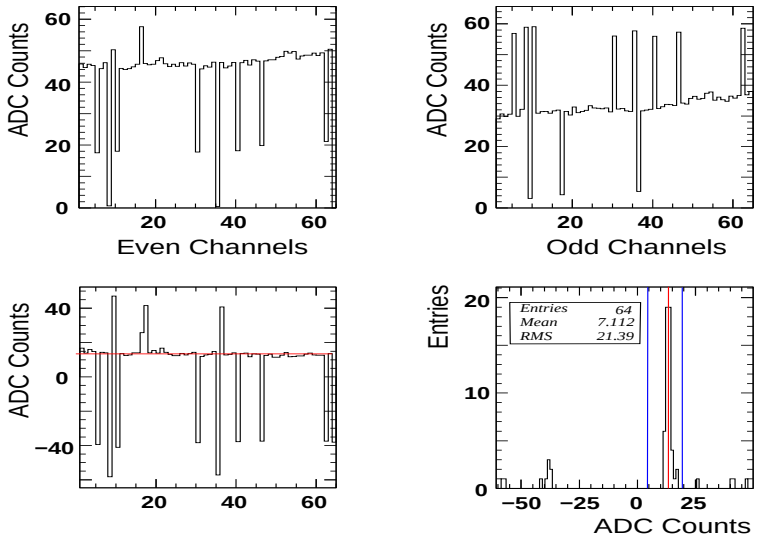

Figure 5.29: The two upper plots show the response of channel x when charge is injected in channel $x+1$; left (right) for even (odd) channels. The lower plots show the difference between top left and top right as an indication of the size of the asymmetric cross talk, left as an array and right as a histogram. The mean is indicated by a red line. Hard cuts needed to exclude broken channels are indicated by blue lines.

	Signal Mean (ADC Counts)	Asymmetric Cross talk (ADC Counts)	Ratio %
High Gain Right	269.0	14.3	5.3
High Gain Left	269.0	13.4	5.0
Low Gain Right	46.7	2.4	5.1
Low Gain Left	46.7	2.5	5.4

Table 5.3: Asymmetric cross talk values.

to the right, however, appears to be slightly higher, probably due to the laser spot position being not exactly in the middle of a strip. For the low gain channel the symmetric cross talk to the left (right) is $13.4 \pm 1.6\%$ ($12.9 \pm 1.2\%$), these values are *not* indicated in the figure. The numbers quoted here are similar to those observed in the analysis of the Erlangen test beam [150].

In Reference [151] a larger value of the cross talk is reported by using the same module and the same laser test stand. The values quoted there are most likely different due to a timing mismatch in the setup. The cross talk values quoted in this work were obtained after a dedicated investigation of the timing of the setup. During building of the laser test stand presented here, it became clear that if the setup has an incorrect timing between the various trigger signals, false cross talk values can be obtained. In general, a setup with silicon detectors is very sensitive to timing. Moreover, the active time of the laser was of the order of 10 ns, where a pulse of a few ps would be simulating a particle's signal in the SRD much better, thereby also simplifying the use of the setup. Nevertheless, if only the functionality of an

SRD module is to be tested the timing does not play such a crucial role, since in this case conclusions are made depending on whether a channel responds to the injected light or not.

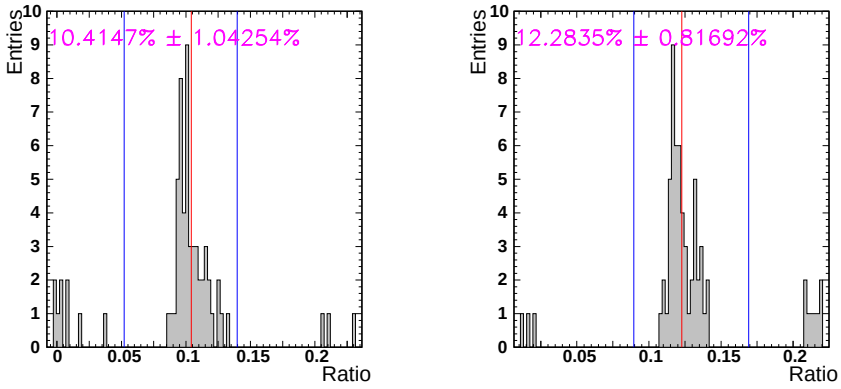

Figure 5.30: Symmetric cross talk for the left and right neighbors of the high gain chip. The red line indicates the mean of the distribution evaluated between the blue lines. The average cross talk is indicated by the number in the upper left of each panel. The error corresponds to the RMS of the distribution.

Another thing which could be checked by using the laser setup is the ratio of the high gain to low gain channel response. The observed ratio is depicted in Figure 5.31 and is about 5.75 ($\sim 85:15$), which is larger than the expected ratio of $78:22$ [124]. The difference is most likely due to the different sampling time in high and low gain channels as mentioned in section 5.4.2 resulting in a different fraction of the signal being digitized.

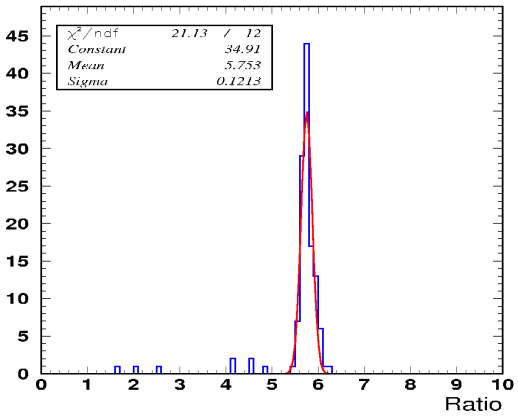

Figure 5.31: Ratio of the high gain to low gain channel response. A Gaussian is fitted to the spectrum. The fit parameters are indicated in the figure.

Due to the linear variation of the laser output power by changing the input voltage[8] to the laser driver, the linearity of a detector could be verified easily by scanning a sensor several

[8]The laser output power depends in fact on the current over the diode. The laser driver converts a voltage to a current.

times with varying laser power. In order to determine Hamlet's linearity, four scans of the same sensor were made. The response to each of these is shown on the left (right) hand side of figure 5.32 for the high (low) gain channel. An increasing response with increasing laser power can be observed. It can also be observed that the laser has sufficient power to saturate the high gain chip. The data is not pedestal subtracted nor common mode corrected; if we care about mean values, common mode correction is not needed. The pedestal only introduces a common offset, subtracting it is not needed when monitoring linearity.

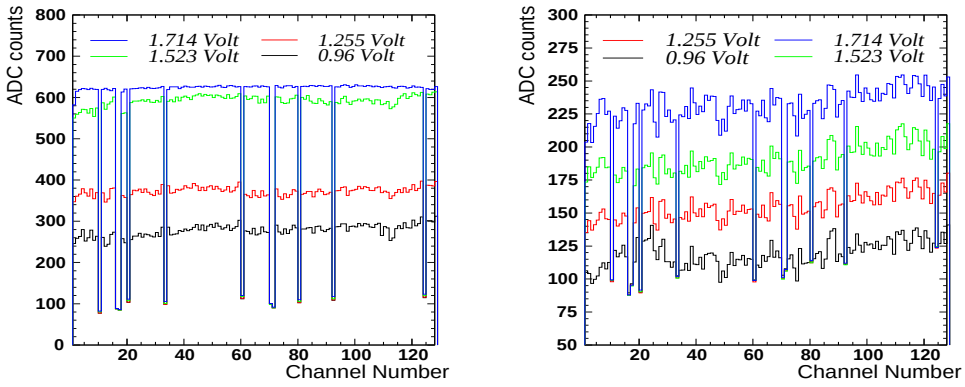

Figure 5.32: Response of the high (low) gain HELIX on the left (right) to various input voltages to the laser driver.

The linearity is tested by plotting the digitized data versus the laser input voltage and fitting a straight line to the data, as is done in Figure 5.33 for a single channel. For the high gain channel only the first three data points are used since the last data point corresponds to a saturated preamplifier. The fit is not very good for the high gain channel, whereas it is reasonably good for the low gain, which of course is less sensitive to small variations. One has to keep in mind that this measurement contains a large uncertainty due to the way the laser power is varied, namely by changing its input voltage. The light output of a laser diode does not depend linearly on the current through the diode for low currents. Moreover, a laser's light yield is temperature dependent. Therefore it was decided that a photo-diode should be implemented in the setup when these measurements are to be performed.

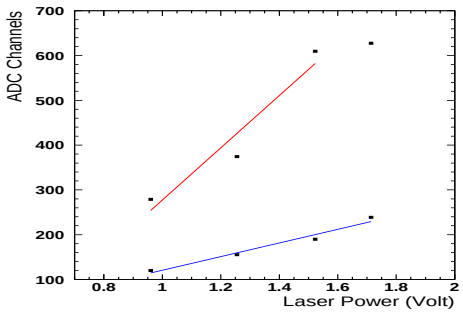

Figure 5.33: Fit of a straight line to the high (red) and low (blue) gain data. Here the response of channel 44 is shown.

The slope for all channels is indicated by black squares in Figure 5.34. On top (bottom) the high (low) gain chip's gain is shown.A uniform gain over the entire sensor can be observed. The red error bars indicate the error of the fit. A rather large error is seen, indicating that the method still needs some tuning, as mentioned before. The blue horizontal line is the mean of all slopes. The first 20 channels in the low gain chip seem to have a slightly varying slope. However, for the other channels the slope is constant within the error. A channel-to-channel slope variation can be explained by differences in stray capacitance and slightly different values of coupling capacitors.

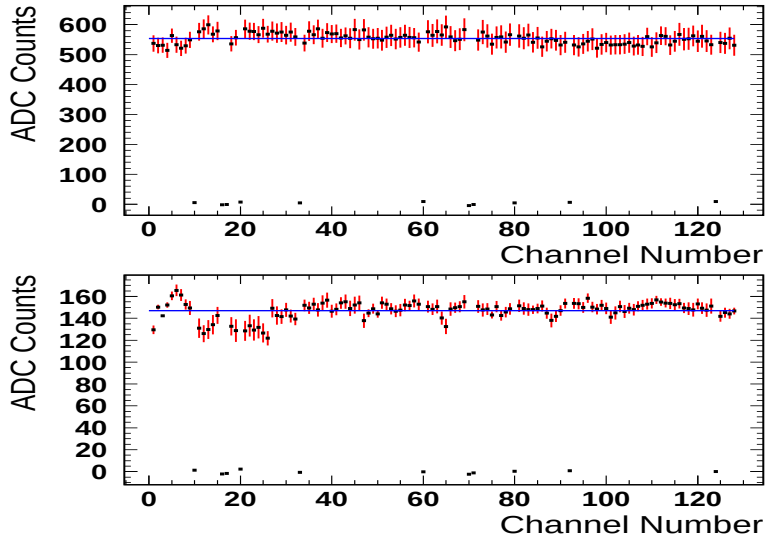

Figure 5.34: Slope values for high (top) and low (bottom) gain from a straight line fit to the data.

5.7.1.4 Concluding remarks

From all the measurements above it is clear that the first prototype of the SRD behaves as expected. Moreover, it was shown that the test as described above could realize a fast check of the functionality of an SRD module. From the measurements performed it can also be concluded that the laser test stand provides an excellent tool for systematic studies of the SRD. By generating large data samples, this setup is the only setup which is able to measure individual pipe cell variations in a reasonable time. Such a measurement has been reported in Reference [146]. By upgrading the setup with a photo-diode and a laser driver being able of producing shorter pulses, and eventually with an infrared laser, extensive tests could be performed, such as detailed cross talk measurements or position reconstructing algorithms.

5.7.2 Test beam

A test of a prototype in a test beam is always beneficial as the system is tested in realistic circumstances: a test beam area usually mimics the final experimental area much better than a laboratory. Not only because the signal of real physical particles is measured, but also because of the electrically noisy environment where the system has to be operated. Moreover, the full

setup including all electronics can be tested. Already very early in the project it was clear that the tandem accelerator facility of the University of Erlangen could be very useful [117]. This facility is able to accelerate protons to a kinetic energy between 2 and 12 MeV, which is exactly the kinematic region of the recoiling protons. A disadvantage is that the beam is continuous and not pulsed like the HERA beam, requiring additional modules in the setup. Tests at the Erlangen facility are reported in Reference [152].

Another excellent test beam facility was found at DESY II, where a beam of 1 to 6 GeV electrons is available. The response of the module to a MIP signal, being the lowest possible signal, is very interesting, especially in terms of signal over noise ratios. Moreover, the location and in house experience of the facility were advantageous. The most interesting fact of the DESY test beam is the availability of a beam telescope in experimental area 22 which was built by the ZEUS MVD group. The telescope consists of three times two high resolution, low noise silicon reference detectors with a very precise intrinsic position accuracy. The test bench is very useful for determining efficiencies and position resolution of a detector, see e.g. References [149, 153, 154, 155].

Apart from initial tests at the DESY II test beam, it was also decided to calibrate the MIP response of all SRD sensors using this facility. In order to have a reliable readout, the entire DAQ of the beam telescope was modified. This section will discuss the setup of the T22 test beam, as well as results of measurements with a preliminary prototype and with Hamlet.

5.7.2.1 The DESY II test beam facility

DESY II [156] serves as a preaccelerator for the PETRA and DORIS rings at DESY as can be seen from Figure 3.1. The creation of a test beam from the primary DESY II beam is schematically depicted in Figure 5.35. Inside the ring a carbon fibre target is positioned causing some electrons or positrons to emit Bremsstrahlung. The Bremsstrahlungsphotons have a direction tangent to the beam and consequently leave the ring. Photons are converted into electron-positron pairs via the interaction inside a metal plate. A dipole magnet and a collimator allow momentum and charge selection of the test beam particles.

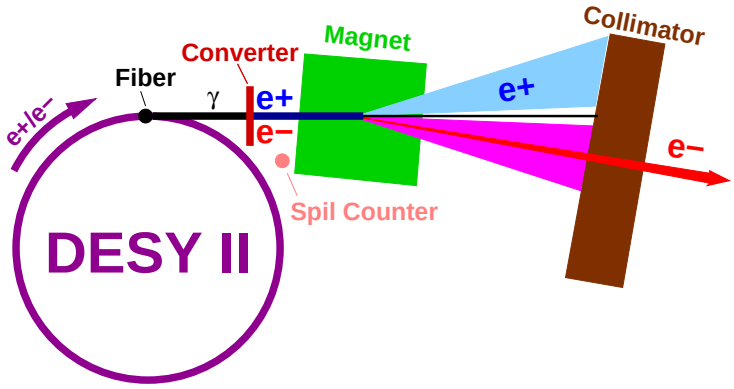

Figure 5.35: Schematic overview of beam generation at DESY II. DESY II itself can be filled either with electrons or positrons.

Individual bunches are accelerated and decelerated at a frequency of 12.5 Hz. Injection into the DESY II ring occurs at energies of 450 MeV, particles are ejected at 4.5 (7) GeV for

DORIS (PETRA). The main RF frequency of the synchrotron is 500 MHz, while the revolution frequency is 1 MHz. As can be seen from Figure 5.36, particles are injected every 2nd cycle. Ejection happens 40 ms after injection.

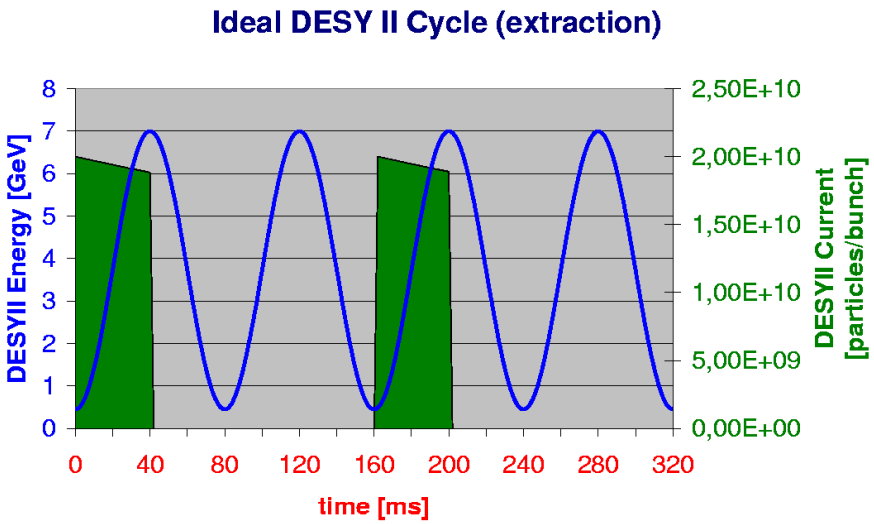

Figure 5.36: The blue line represents the DESY II clock frequency of 12.5 MHz. The filled area shows the beam current. Ejection happens 40 ms after injection.

Test beam particle rates depend on the thickness of the copper plate converting the photons to electron-positron pairs and on the selected energies. Rates of 1 kHz are expected for a 3 mm copper target and 3 GeV energy, while the rate for 6 GeV electrons after a 1 mm copper target is only 160 Hz [157].

The electronics for the SRD project is designed to run at a clock frequency of 10 MHz, the beam clock frequency of HERA. As mentioned before, the beam clock at T22 runs only at 1 MHz. Therefore a Phase Locked Loop (PLL) was needed to convert the 1 MHz clock to 10 MHz. Two additional problems with respect to the beam clock were observed: first of all it was seen that the 1 MHz clock disappears after about 78 ms after injection, and secondly, a beam clock modulation of 12.5 Hz was measured. The initial PLL used in the setup could not cope with these difficulties. A dedicated triggering scheme was accordingly developped accepting only triggers within the first 40 ms after particle injection into DESY II. Later, the PLL was replaced by an upgraded version.

5.7.2.2 Beam telescope at T22

The beam telescope at T22 is depicted in Figure 5.37 and consists of 3 pairs of single sided high resolution silicon sensors. Each pair consists of two 3.2×3.2 cm^2 and 300 μm thick silicon sensors with a strip pitch of 25 μm and a readout pitch of 50 μm, yielding a position resolution of 96 μm (15 μm) for a 1 GeV (6 GeV) beam. The improved resolution at higher energies is a consequence of a reduction in multiple scattering. Each sensor has 640 readout channels.

The Device Under Test (DUT) is placed on an $x - y - \phi$ translating stage, located in between the first and second reference detectors. In front of the first and behind the last reference detector two scintillating trigger fingers are positioned, accounting for a trigger surface of about 1 cm^2.

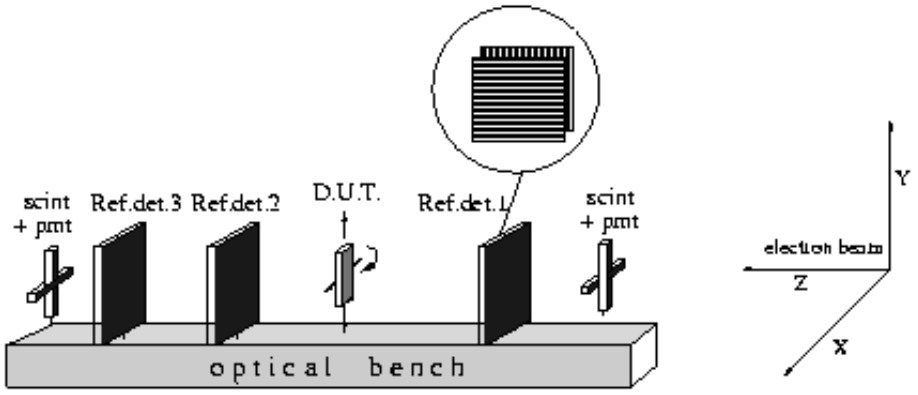

Figure 5.37: Schematic of the optical bench located at T22. The position of the 3×2 reference detectors can be seen as well as the position of the Device Under Test (DUT) and trigger scintillators.

The reference system is read out with the Viking Architecture 2 chip (VA2) made by IDE AS, Norway [158]. These are low noise frontend chips consisting of 128 input channels each having a preamplifier, a shaper, and a digital control unit. All amplified signals are fed into a multiplexer controlled by a 128 bit shift register. The equivalent noise charge (ENC) of the chip is $80 + 15$/pF. The VA is depicted schematically in Figure 5.38, together with its timing scheme. After the hold signal becomes active, a 'shift-in' bit is required by the shift register. A clock is needed to move the shift bit through all channels in order to provide the analog output. The necessary sequencing pattern was generated using the V551 module from CAEN [159]. The output of 10 chips was daisy chained, in order to have only one output for every reference detector. Digitization happened using the CAEN V550 module [160].

During setup of the test stand a 50 Hz noise was observed on the analog output of all 3 reference detectors. Therefore a dedicated RC filter was constructed. The pedestal and MIP signal distribution for reference detector 2 are depicted in Figure 5.39. The most probable value of the Landau (see equation 5.3) is given in the left panel as fit parameter $P2$ and corresponds to 287.5 ADC channels. The width of the pedestal is indicated by 'sigma' in the right panel and is about 5.4 ADC channels. The signal over noise (S/N) ratio consiquently amounts 53. The numbers mentioned correspond to the n-side sensor of reference detector 2. All other sensors had a S/N ratio around $70 - 80$, which is very high.

5.7.2.3 Measurements with an early prototype

The very first prototype build for the HERMES SRD was built intended to test the properties of the TIGRE sensor in combination with the charge division setup. The prototype was based on a ZEUS MVD hybrid, which houses four HELIX3.0 chips. Additionally, the prototype contained a charge dividing fanout and a TIGRE sensor. Only one side of a single TIGRE sensor was

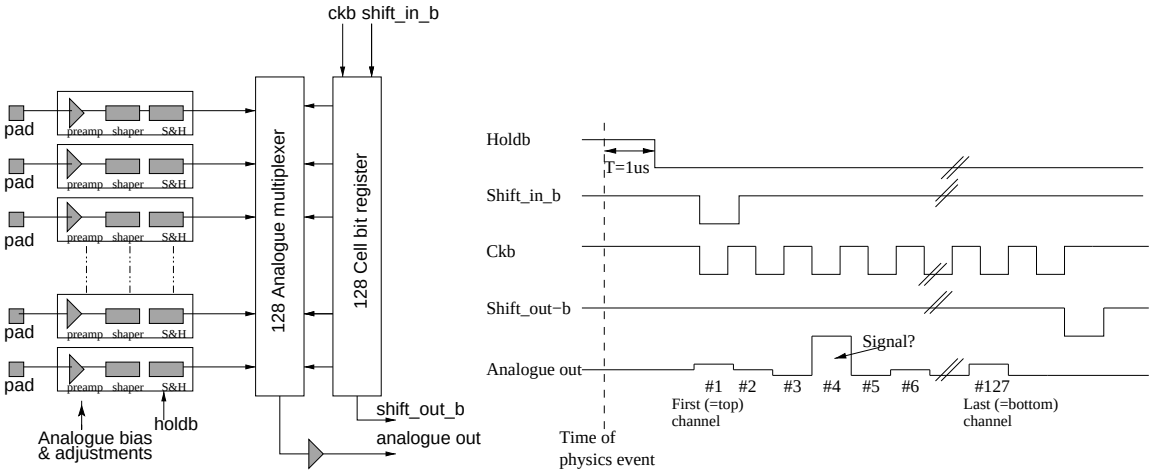

Figure 5.38: Schematic of the VA chip used to readout the reference detectors. The sequencing of the chips is shown on the right.

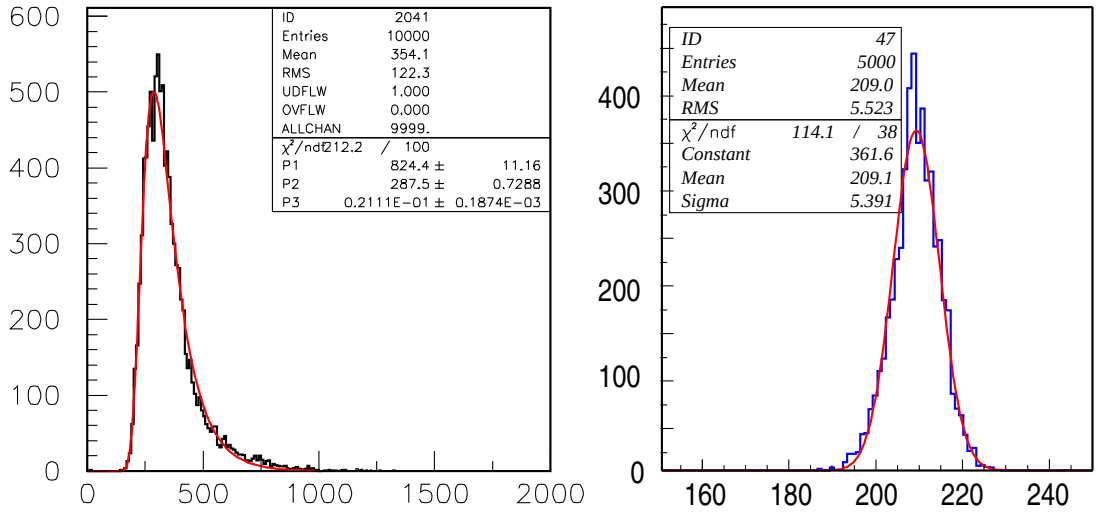

Figure 5.39: Response of the n-side sensor of reference detector 2. A Landau curve is fitted to the histogram. On the right the pedestal distribution for one channel of this sensor is shown. A Gaussian was fitted to the pedestal distribution.

bonded to the fanout. The latter had three different values of charge dividing capacitors of $5, 10$ and 22 pF, in order to test the behavior of the charge division. This prototype often is referred to as 'Henry' and was also used for initial setup of the laser test stand. With Henry also the double peak structure of Figure 5.12 was observed.

All the measurements with Henry were done using the ZEUS electronics setup. Digitization happened with a CAEN V550 while sequencing was performed by a SEQSI module [161]. In order to get the timing correct a delay curve was measured using a fine delay between trigger and clock signal. The delay curve is depicted in Figure 5.40.

Even though only one side of Henry was connected to the readout hybrid, the signal of both sides could be measured by unbonding the sensor, rotate it by 90 degrees and flipping

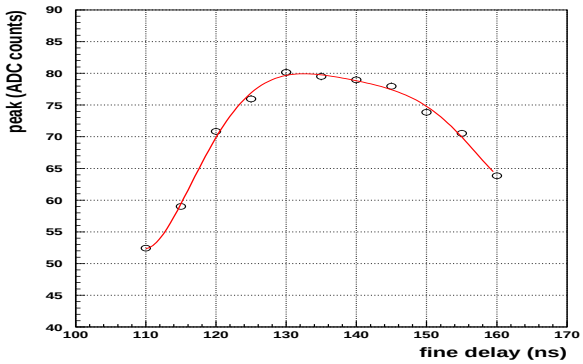

Figure 5.40: Landau peak values fitted to the spectrum as a function of the delay between trigger and clock delay.

it. A comparison of p- and n- side signals is shown in Figure 5.41. The 20% difference in signal height is due to the difference in capacitance between both sides. All signals of a side lie within a 5% band around the mean value. The different mean values could be due to different stray capacitances on the fanout.

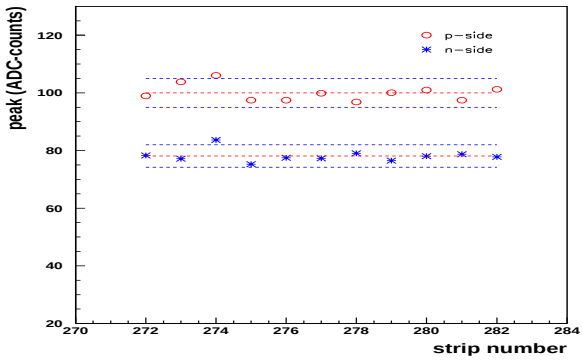

Figure 5.41: Most probable response of various p- and n- side strips of Henry to 1 GeV electrons. The dashed lines indicate a 5% band around the mean. Figure from Reference [162].

Using the information from the reference system interesting information about the response to minimum ionizing particles could be extracted. The position resolution of the sensor was measured to be $222\ \mu\mathrm{m}$ in agreement with the theoretical expected value of $\frac{\mathrm{pitch}}{\sqrt{12}}$. As the S/N ratio of the reference system is very high, the telescope can be used to determine the efficiency of the device under test. Henry's efficiency was measured to be $\epsilon = 99.73\%$ for the n- side when using a charge dividing capacitor of 10 pF. Due to the accurate position resolution of the reference system, both signal size and efficiency could be measured across a strip. These are shown in Figure 5.42. Efficiency and response are constant across a TIGRE's strip. All results of the measurements with Henry are detailed in [162] and [163].

From the initial measurements at T22 it could be concluded that the principle of charge division was functioning appropriately, and that MIP signals could be measured with a high

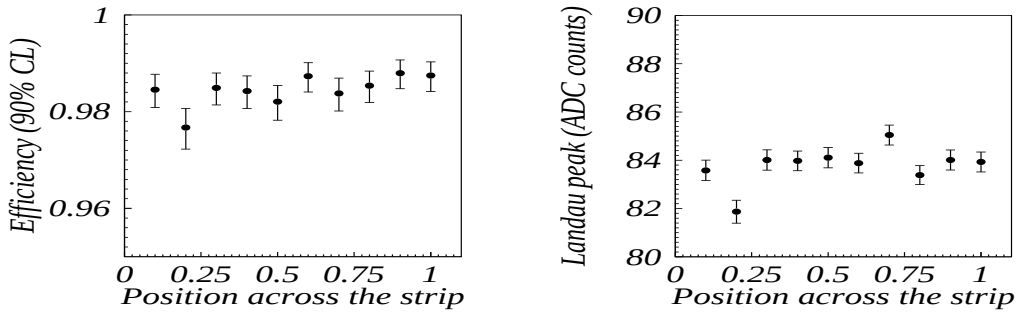

Figure 5.42: The left panel shows the efficiency across a strip while the right panel shows the Landau peak position across a strip. The units on the x axes are the fractional strip position. Figure from Reference [163].

efficiency by a TIGRE sensor in combination with a charge dividing setup. These observations were of key importance for the further development of the entire Recoil project.

5.7.2.4 Measurement with Hamlet at T22

For the calibration of all SRD production modules, the setup required some major changes. The mechanical setup had to be changed, and an HADC and HLCU needed to be installed. Also, the whole DAQ needed to be moved from a VME based CPU to Linux. Therefore the existing DAQ program was extensively modified. The full functionality of the HADC, including zero suppression and common mode correction was implemented. The software controlling the translating stage was ported to Linux as well.

Additionally, a dedicated decoder was written to convert the binary output of the DAQ program into the HBOOK ntuple format [106]. The output files of the decoder matched the input of the TESA package, which was specially developed by the ZEUS MVD collaboration in order to analyze the teststand data [164]. The routines of the decoder could easily be implemented in HDC. Special care was taken to optimise the speed of the decoder.

A requirement of the decoder is to group information from the high and low gain chip connected to the same channel together into one table entry. The serial output of a HERMES SRD module, however, has a sequence of low gain, high gain, low gain and high gain. There is a larger probability to have a signal in the high gain and not in the low gain, than vice-versa. Therefore it was decided to invert the entire output, thus looping first over the high gain entries followed by the low gain. The program was extensively tested and was able to create entries for low gain channels, even if there was no corresponding high gain entries, while still keeping all channels ordered in the readout chain.

The response of one strip to a 1 GeV electron beam, as measured with the first real prototype Hamlet with the modified setup, is depicted in Figure 5.43. On the left hand side the response of one strip is shown together with a fit of the sum of a Gaussian and a Landau function. The fit parameters are indicated in the figure. The Gaussian describes the pedestal of that particular strip. The mean of the Gaussian, $P5$ in the figure is around zero, as expected. The width of the Gaussian, $P4$, indicates the width of the pedestal distribution, i.e. the noise, and is about 3.5 ADC channels in this case. The peak of the Landau describes the

most probable energy deposition. This corresponds to $P2$ in the figure and amounts 15 ADC channels.

If for every event the strip with the highest ADC value is taken on the premise that that ADC value should be higher than three times the average noise, the histogram on the right hand side of Figure 5.43 is obtained. This histogram represents the signal distribution in the silicon sensors from a data sample. A Landau curve was fitted to the spectrum. The fit gave a most probable value of 15.64 ADC chanels. When taking the average of the most probable value of the Landau obtained from the comined Landau and Gaussian fit to every individual strip (like in the right of the figure), the value of 15.97 ADC channels is obtained. This value is close to the aforementioned 15.64 ADC channels, indicating a good stability of the method. Taking the noise of about 3.5 ADC channels into account, a S/N ratio of about 4.5 is obtained. This measurement was the first measurement of a MIP with the SRD prototype, and was done without having performed a delaycurve measurement. Due to time constraints efficiencies and signal distributions across strips could not be extracted. Results of the calibration run itself are detailed in Reference [165].

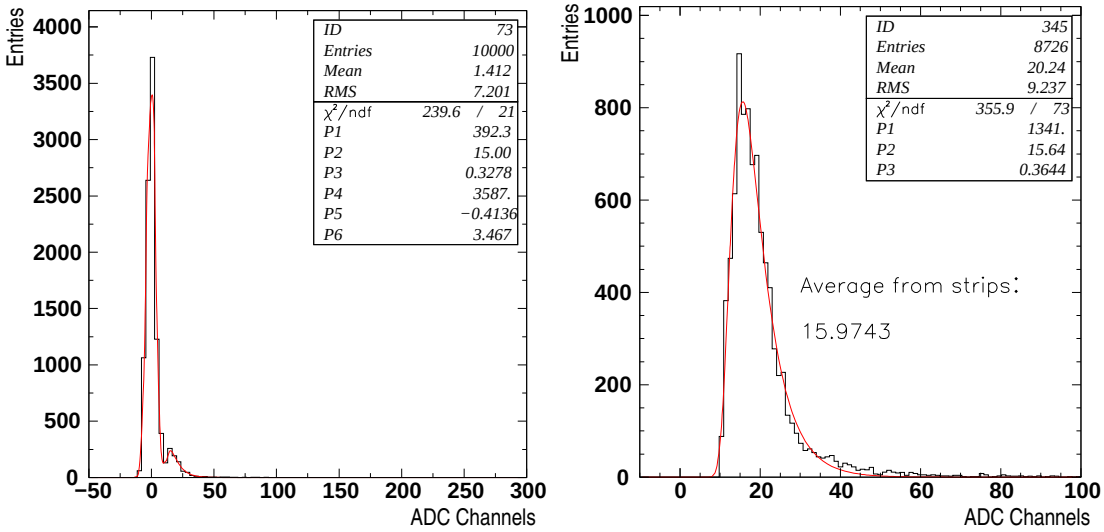

Figure 5.43: The left panel shows the response of one of Hamlet's strips in test beam 22. A Landau+Gaussian was fitted to the spectrum. Fit parameters are indicated in the figure. The right panel shows the distribution of the highest ADC value per event. A Landau was fitted to the spectrum.

This discussion ends the chapter on the Silicon Recoil Detector. The detector requirements and resulting design have been discussed, as well as the decision on the readout chip for the project. Next, the description of three test stands designed and build during the course of this thesis were discussed: A HELIX parameter test setup, able of measuring the output of the HELIX shaper with a very good precision; a laser test stand, which can be used for an extensive study of a SRD module for various input signal, and can serve as a diagnostic tool for fast module performance controlling. The third test stand was located at the test beam facility at DESY-Hamburg and allowed a performance check for the response of the SRD to minimum ionizing particles both in terms of position resolution and the signal to noise ratio. The laser

test and T22 setup have been used for testing and calibrating the production modules of the SRD project. The parameter test setup was used to decide on the HELIX parameter values.

CHAPTER 6

Monte Carlo Simulation of Preshower and Calorimeter

Detector simulation is a very important tool in experimental physics for several reasons. First of all, a Monte Carlo simulation can be used to verify the different steps in an analysis chain. To verify an analysis code, a user could generate a Monte Carlo sample with some large asymmetry. This generated asymmetry should also be extracted by any user analysis code. If not, the extraction method is doubtful at least. Secondly, the Monte Carlo can be used to investigate various effects in the experimental setup, such as misalignment between the detector halves, or irregularities in the beam slope. Another very powerful consequence of having a decent Monte Carlo description, is the ability to analyze background contributions in a particular data sample. Resolution and efficiencies can also be predicted by a reliable Monte Carlo simulation, the purity formalism explained in section 2.5.

Monte Carlo simulation is the only tool to study the response of the calorimeter to photons, since photons only interact with the calorimeter. In principle, also test beam data could determine the functioning of the calorimeter with respect to photons, however, no information about such test beam is available for the HERMES calorimeter.

For the measurements of DVCS and π^0 production, the calorimeter is the only device measuring the photon energies. The determination of energies is of vital importance when studying exclusive reactions. Indeed, when determining the missing mass (or energy) of a particular exclusive event candidate, the energy of the DVCS photon[1] should be measured precisely and combined with the momentum of the scattered lepton, measured by the deflection in the magnetic field. Especially in the absence of a Recoil Detector, these constraints on the missing energy or missing mass are mandatory. During the first years of HERMES running, however, the electromagnetic calorimeter was used mainly as a triggering and PID detector. Consequently, not a lot of effort was put in the Monte Carlo description of the calorimeter,

[1] in the case of π^0 production its 2 decay photons

resulting in a failed description of the experimentally observed spectra like, e.g. the π^0 invariant mass. As particle showers already start in the preshower, it is of the utmost importance to also verify the simulation of the preshower. This chapter covers the implementation of an improved description of both the preshower and the calorimeter, after a general introduction about the HERMES Monte Carlo.

The last section in this chapter shows the use of a good Monte Carlo description. This section determines the energy and position reconstruction capabilities of the calorimeter based on the new Monte Carlo description.

6.1 Organization of the HERMES Monte Carlo

The HERMES Monte Carlo can be split into two major parts: generation of the physics processes itself and simulation of the detector response. The first part is referred to as Generator Monte Carlo (GMC), whereas the second part is called HERMES Monte Carlo (HMC). An overview of the HERMES Monte Carlo chain can be found in Figure 6.1

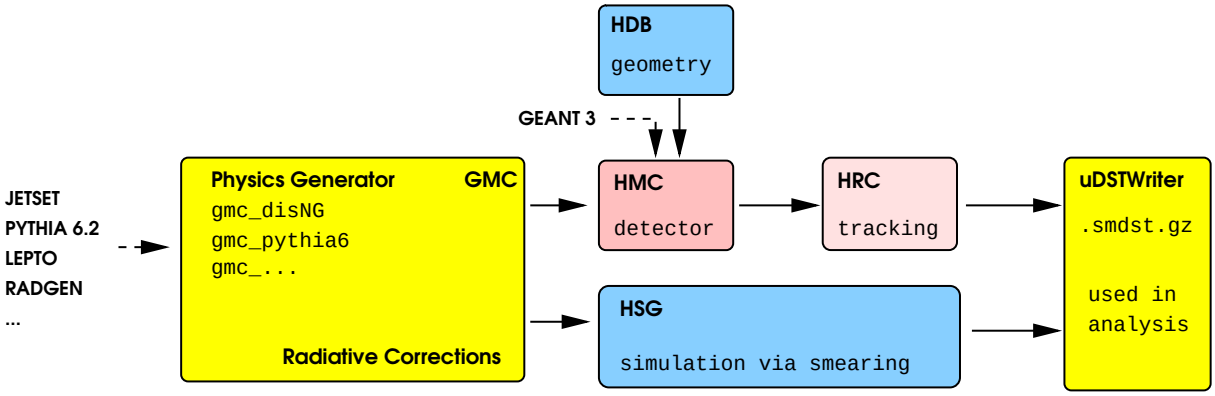

Figure 6.1: Schematic overview of the HERMES Monte Carlo chain as explained in the text.

In the case of GMC, there are various kinds of generators available dependent on the physics processes to be studied. The most popular generators are gmc_pythia and gmc_disNG. The former is based on the PYTHIA package [166], while the latter uses LEPTO [167] for generation of the DIS event itself, and the JETSET [168] package to calculate the fragmentation. Both gmc_pythia and gmc_disNG use RADGEN [169] to calculate QED radiative corrections. A variety of other generators exist to study specific topics such as DVCS, transversity or ρ^0-production.

The detector simulation itself, HMC, is based on the GEANT3 package, developed by the CERN computing group [170]. The detector geometry is described by the HERMES Database package HDB. The output of HMC is passed to HRC (see section 4.1.2) which is responsible for tracking. The output of HMC obviously needs to match the input of HRC. More details about HMC will be given in the next section.

HMC in combination with HRC is very time consuming: it takes a long time to generate a sufficiently large data sample even on modern CPUs. Moreover, some analyses need to iteratively tune Monte Carlo parameters, thus requiring a huge amount of CPU time. To

decrease the required processing time for such analyses, a dedicated smearing generator, HSG was developed [171, 172]. This generator smears the particles according to the resolution in p, θ_x and θ_y: $\frac{\Delta p}{p}(p)$, $\frac{\Delta \theta_x}{\theta_x}(p)$, and $\frac{\Delta \theta_y}{\theta_y}(p)$. It was shown that the smeared data reproduce the fully reconstructed Monte Carlo samples well for charged particle tracks.

Both HRC and HSG samples can be passed on to the uDSTWriter, to provide a handy format for the user analysis codes, so that essentially the same programs could be used to analyze both experimental data and Monte Carlo samples.

6.1.1 Setup of the HERMES Monte Carlo HMC

As already explained in the previous section the HERMES Monte Carlo is based on the GEANT3 package. GEANT is responsible for the event processing and determines how much energy is deposited in a certain detector. Obviously it needs to know the experimental setup before any simulation can be performed.

The setup in GEANT is described by volumes which can be filled with matter. On the one hand the geometrical shape of the volume needs to be given, and on the other hand the parameters of the material the volume is built of. Some volumes are declared as sensitive volumes, so that the energy deposited in that volume is assigned as a hit in that specific volume.

The geometry of the HERMES experiment is accounted for by a dedicated program called HDB. The program basically forms a human understandable interface to the geometry file needed by GEANT/HMC. It contains all material and volume definitions of the HERMES experimental setup.

After the geometry is read in, GEANT tracks particles through the detector. This is done in discrete steps. A particle is advanced by a certain amount which is determined by the step size. Interactions, energy losses and deflections are calculated over the step size. The step size itself consequently is a very important parameter. Ideally it should be as small as possible to have a high precision. Smaller step sizes, however, increase computing time. By default, the step size is calculated based on the minimum interaction length for a particle undergoing a certain process within a volume. These processes can be the decay of the particle, crossing of a volume boundary, Compton scattering, and many more. Furthermore, the user can set a maximum step length. If any secondary particle is generated within a step, its properties are added to a stack, and the parent particle is tracked further through the detector. At the end of every step it is evaluated whether the particle still is inside the same volume. If the particle has decayed, or if its energy is below a certain threshold, the track is abandoned.

As specific particle types undergo specific processes, there exist routines in GEANT for electrons, photons, charged and neutral hadrons, muons and neutrino's. As an example, the flow charts for electron and gamma tracking are shown in Figure 6.2. For electrons (or positrons) the step length is calculated as the minimum of the steps to the next Bremsstrahlung, $\delta-$ ray production, and in the case of positrons to the next positron annihilation process. Additionally, it is checked whether the particle stays within the volume. Once the step size is determined, multiple scattering and deflection in the magnetic field are calculated. After that, the specific process - Bremsstrahlung, δ-ray production, or positron annihilation - is applied to the particle. The routines `GBREM`, `GDRAY` and `GANNI` contain the actual cross sections of aforementioned

processes. Finally some variables like the change in particle momenta and the energy loss in the step are evaluated.

For photons a similar flow chart can be drawn, albeit easier, since they are not deflected in a magnetic field, and are not subject to multiple scattering. For these particles the step length is evaluated from the minimum distance to the next pair production, Compton scattering, photoelectric event, or to the next photo-induced fission. The latter process will only occur for heavy nuclei.

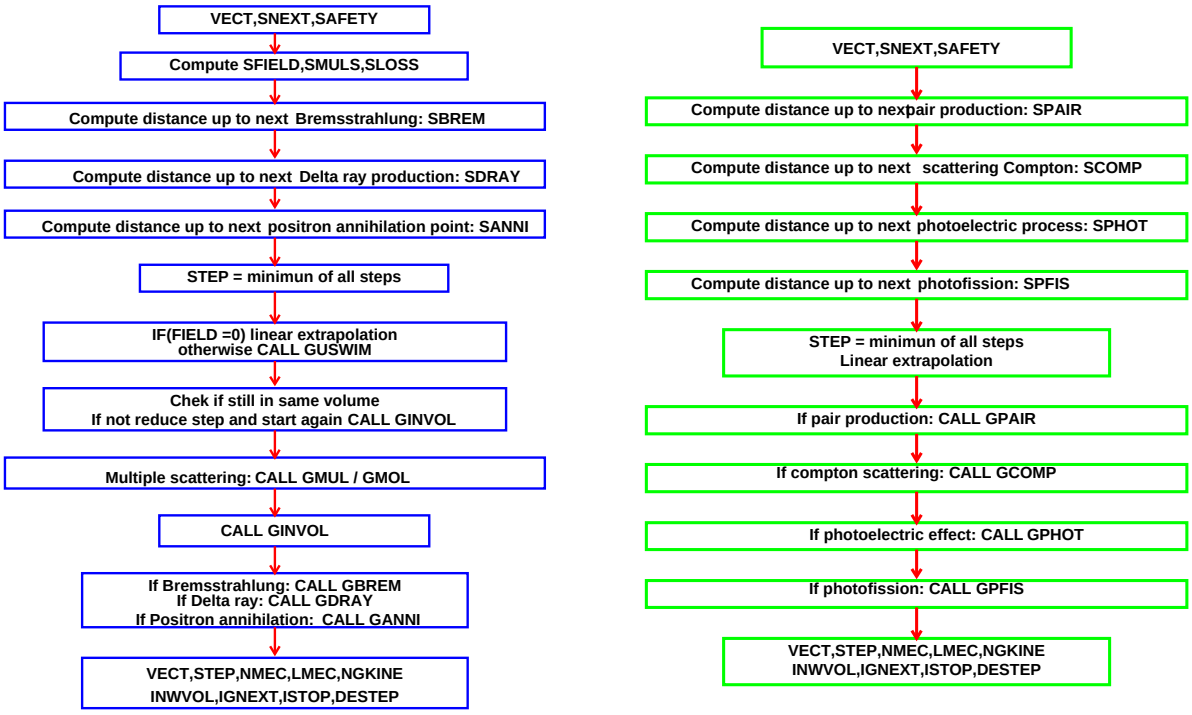

Figure 6.2: Flowcharts for the GEANT routine `GTELEC` (`GTGAMMA`) on the left (right). Figures taken from [170].

Figure 6.3 shows an event $e+p \rightarrow e'+p'+\pi^0$ tracked through the HERMES spectrometer by HMC. It is important to realize that the tracking of HMC is just a physical tracking of a particle through the spectrometer whereby various interactions are simulated. The red line crossing the entire spectrometer in the figure corresponds to the track of the scattered lepton. The particle track as it is used in analysis codes is reconstructed by HRC from the signal induced in the tracking chambers, indicated by the crosses in the figure. The scattered electron is located in the upper detector half, while the two photons from the π^0 decay are in the lower half. The scattered proton recoils in the transverse direction and hence is undetected. In fact, the proton is scattered in the region where the Recoil Detector is to be built. Only the charged electron leaves hits in the wire chambers, the RICH and the TRD. Both photons only interact in the calorimeter and preshower. Three electromagnetic showers can be seen in the calorimeter. In this particular event the energy of the scattered lepton is 21.5 GeV, while the photon energies are 1.3 GeV and 4.5 GeV respectively. Particles induce a larger shower depending on their energy. The lowest calorimeter cluster in Figure 6.3 can be identified as the one induced by the 4.5 GeV photon.

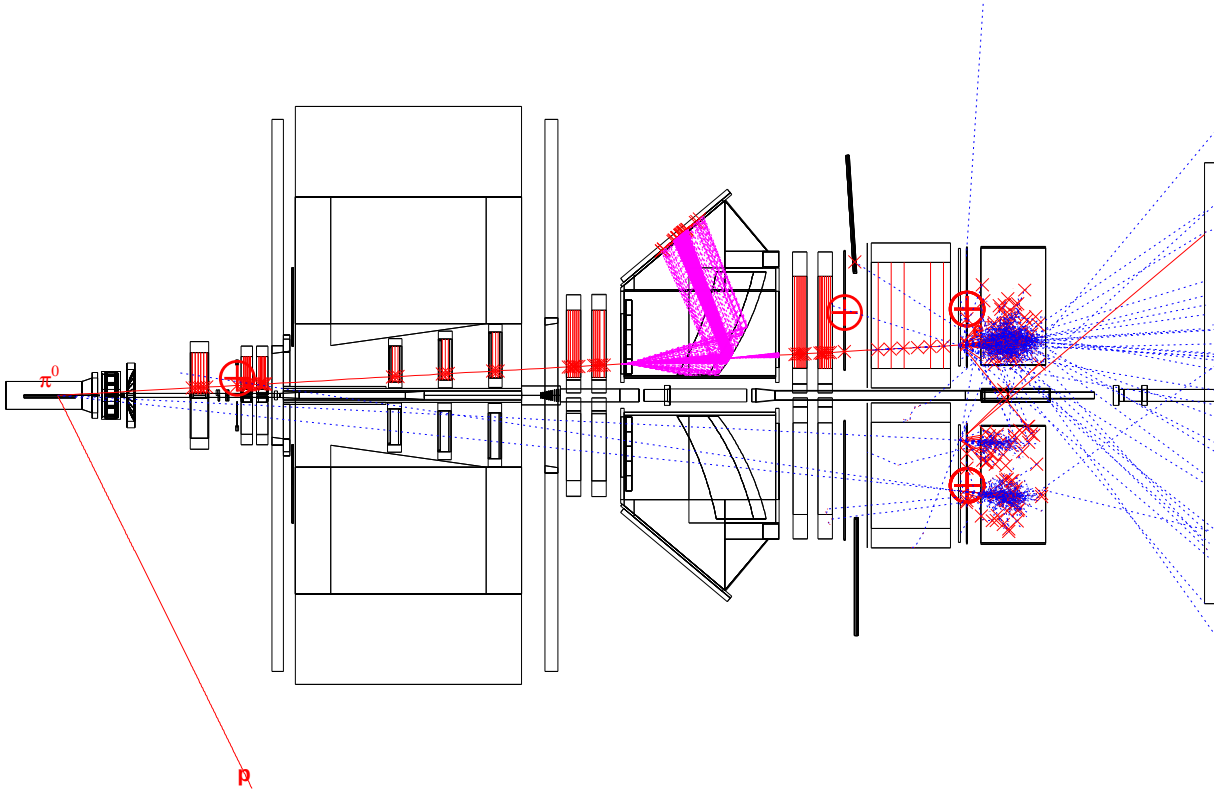

Figure 6.3: An $e + p \rightarrow e' + p' + \pi^0$ event tracked through HMC. Red lines correspond to charged particle tracks, blue dashed line correspond to photon tracks. The magenta lines visible in the RICH correspond with the path of optical photons. From top to bottom the showers induced in the calorimeter by the 21.5 GeV lepton, and the 1.3 GeV and 4.5 GeV photon can be seen. The target proton is scattered in a transverse direction.

After all tracks of an event are processed, including any secondaries generated, a certain amount of energy will be deposited in the sensitive volumes. This information needs to be further processed in order to match the response of the real physical detector volumes. Therefore efficiencies need to be taken into account, conversion to ADC and TDC values need to happen, instrumental limitations need to be calculated, and so on. The combination of all these effect is called 'digitization'. The digitization is ensuring that the HMC output format exactly matches the output of the HERMES decoder HDC.

6.2 Simulation of the preshower

It was already known since a long time that the HERMES Monte Carlo wasn't producing the right E/P distributions for leptons, nor did it describe the π^0 invariant mass spectrum correctly. As the electromagnetic shower is started in the preshower, the implementation of the preshower in HMC was checked first.

The preshower, depicted in combination with the calorimeter in Figure 3.21 on page 66, is a sandwich of a 0.44 inch $(= 11.176$ mm$)^2$ thick lead layer in between two layers of 0.05 inch $(= 1.27$ mm.$)^2$ of steel. The thickness was measured in the experimental area with an

[2]These values correspond to the design values

accuracy of about $100\ \mu\mathrm{m}$ and was found to be in perfect agreement with the design values [173]. However, when looking at the geometry file, the thickness of the preshower material was defined to be 11 mm. Moreover, the relative amount of lead and steel was found to be wrong, as well as the total density. The fraction of steel can simply be calculated as $2 \times 1.3/(11.3 + 2 \times 1.3) \approx 0.1871$, taking into account the measured values. The relative fraction of lead consequently is 0.8129. The density of the compound was calculated as:

$$\rho = 0.8129 \cdot \rho_{\mathrm{Pb}} + 0.1871 \cdot \rho_{\mathrm{Fe}} = 10.74\ \mathrm{g/cm^3} \qquad (\rho_{\mathrm{Pb}} = 11.4\ \mathrm{g/cm^3}, \rho_{\mathrm{Fe}} = 7.86\ \mathrm{g/cm^3}).$$

In the original geometry the density was specified as $9.73\ \mathrm{g/cm^3}$. Obviously, these changes affect the shower development in the Monte Carlo.

As can be seen from Figure 3.21 the preshower is followed by a hodoscope array H2. At typical HERMES energies, charged particles induce MIP signals in the hodoscope scintillators, i.e. corresponding to the minimum in the Bethe-Bloch formula, equation 5.2 on page 94). If, however, a particle shower is established by the interaction with the lead of the preshower, the signal in H2 will be large, since all charged particles involved in the shower will have a MIP energy deposition in the scintillator plastic. The signal of H2 therefore gives an indication about the interaction in the preshower itself. These two devices, the preshower and H2 are commonly called 'the preshower'. The words 'preshower pulse' are frequently used in the rest of this thesis to refer to the signal in H2 caused by the interaction of particles in the lead-iron sandwich and the scintillator array.

In the next three sections, the preshower response of data and Monte Carlo are compared for leptons, photons and hadrons respectively. These comparisons give an indication of the precision of the preshower description in HMC.

6.2.1 Response to leptons

Since the thickness of the preshower is about two radiation lengths, all electrons are expected to interact in the lead thus developing a particle shower. According to calculations detailed in Reference [174][3] about $20 - 30$ charged particles are expected for electron-induced showers after 2 radiation lengths. Given that a MIP deposits about 2 MeV in 1 cm of scintillator material, signals between $0.04 - 0.06$ GeV are expected.

The response of H2 to scattered positrons can be seen in Figure 6.4. Two distinct regions can be seen in the figure. The region above ~ 0.11 GeV of energy deposition corresponds to a regime where the hodoscope ADCs are saturated. In the region below ~ 0.11 GeV energy deposition, a fairly good agreement can be seen between experimental data represented by the dashed line, and the new Monte Carlo, represented by the full line. For comparison, the simulated response using the old description of the preshower is indicated in the figure as well. The simulation improvement is clearly visible.

For preshower pulses > 0.11 GeV, data and Monte Carlo are completely different. This region corresponds to the regime where the ADC is in overflow. The observed difference between data and Monte Carlo is explained below. ADC overflow can happen when too many particles are present. The response of the hodoscope is linear up to about 0.1 GeV of energy deposition. This value corresponds to the signal induced by about 50 MIP particles. Some particle showers will contain more than 50 particles after two radiation lengths.

[3] To be more specific: from Figure 5.14.1 in that reference and $\ln \frac{E}{E_c} \sim 5 - 7.5$ (E_c lead $= 9.51$ MeV)

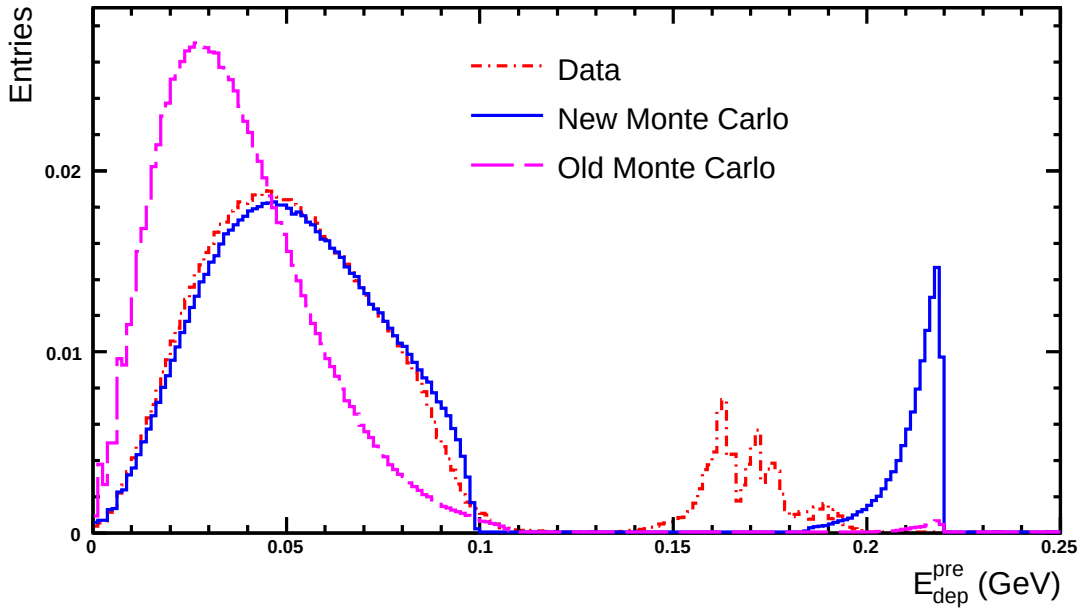

Figure 6.4: Preshower response for experimental data (red dashed-dotted line) and Monte Carlo with the new preshower implementation (blue full line). The pink dashed histogram presents the simulated preshower response before the applied changes. All histograms are normalized to unity.

Figure 6.5 compares the preshower response as a function of lepton energy from data and Monte Carlo, when the saturation region is omitted. The plotted value was obtained by the mean of a Gaussian fit to the response in each energy bin. A good agreement can be seen. As expected, the mean response of the preshower increases with energy: the number of particles in a shower is related to the energy of the initial electron.

As the preshower PMTs are located at the end of a scintillator paddle (see Figure 3.21), the scintillating light suffers an attenuation as a function of y. This attenuation is corrected for in HRC. As the HMC output also is passed through HRC, the attenuation needs to be included in the digitization procedure of the Monte Carlo. Figure 6.6 shows the preshower pulse as a function of positron momentum for different y-bins for both Monte Carlo and data. The preshower pulse clearly is independent of the y position for both samples, indicating a good implementation of the attenuation in both HRC and HMC of the attenuation.

The attenuation correction in HRC is independent of whether the ADC is saturated or not. As a consequence, the overflow bin becomes what it is in Figure 6.4 for data. In the case of Monte Carlo the overflow is what it is due to the handling of hodoscope ADC overflows in HMC. This difference between data and Monte Carlo in this overflow regime could easily be corrected in HRC.

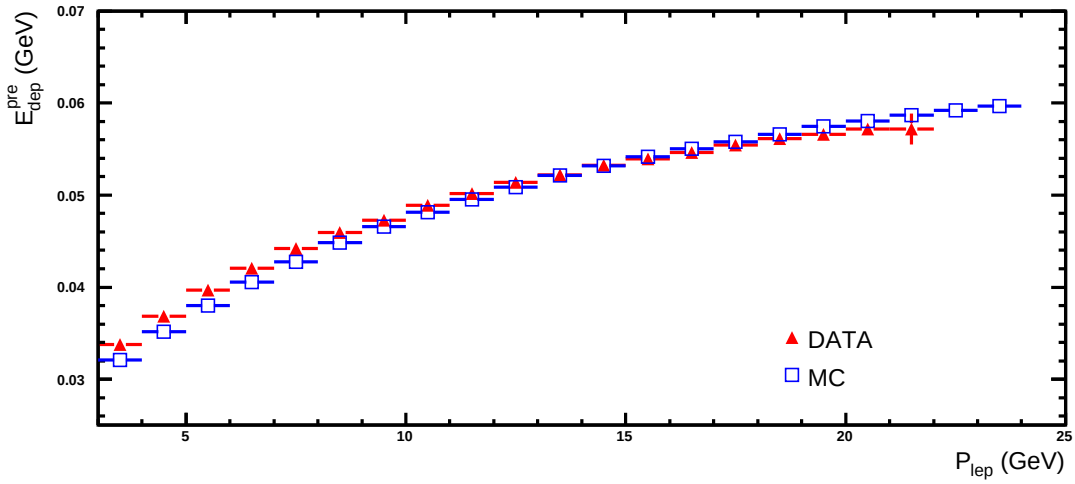

Figure 6.5: Open squares (closed triangles) represent the preshower response as a function of lepton energy for experimental Monte Carlo (data).

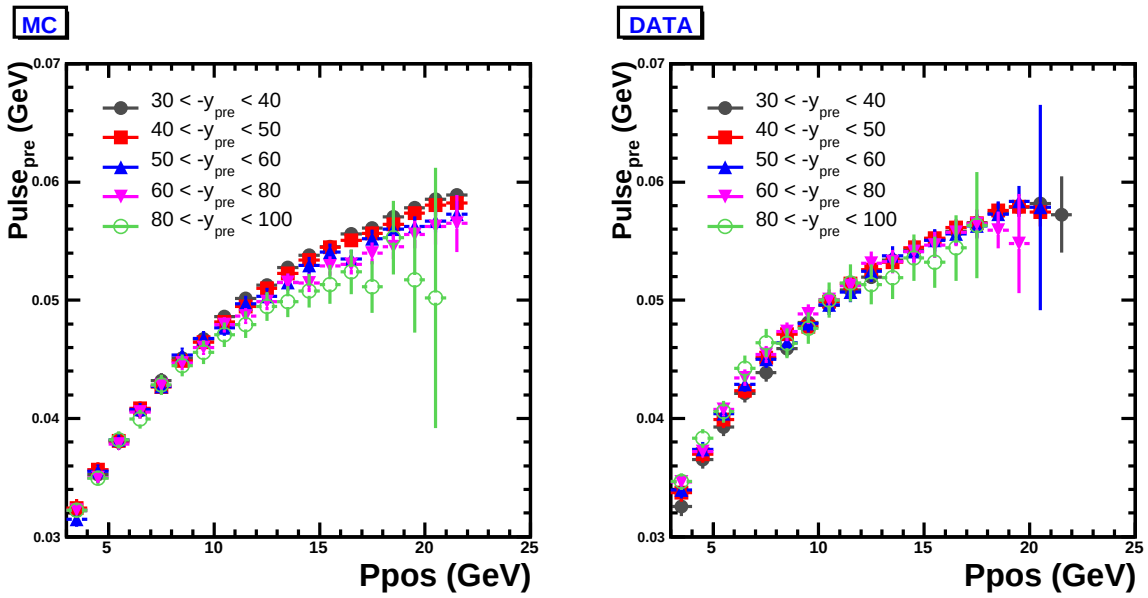

Figure 6.6: The left (right) panel shows the preshower pulse as a function of positron momentum for Monte Carlo (data) for various $y-$ bins as indicated in the figure.

6.2.2 Response to photons

Unlike charged particles, photons do not continuously lose energy due to collisional losses. They rather lose energy due to three distinct processes: the photo-electric effect (p.e.), Compton scattering and pair production. Additionally, a photon can lose energy by Rayleigh scattering, or photo-nuclear interactions, with most notably the giant dipole resonance (g.d.r.). The cross section for these processes is depicted in Figure 6.7. While at lower energies the photo-

electric effect dominates, pair production in the nuclear field, $\kappa_{\rm nuc}$, together with an additional contribution from pair production in the electron field, $\kappa_{\rm e}$, rules at higher energies.

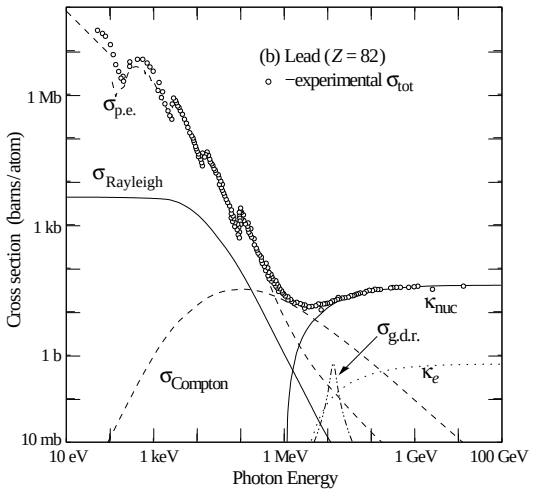

Figure 6.7: Total cross section for photons in lead, together with the partial contributions from various processes. Figure from Reference [121]. Abbreviations are explained in the text.

From the cross section, the mean free path, λ, expressed in cm, can easily be derived:

$$\frac{\sigma \times N_A \times \rho}{A} = \frac{1}{\lambda}, \tag{6.1}$$

where σ is expressed in cm^2/atom, N_A is Avogadro's number, representing the number of particles per mole, ρ the density of the material in g/cm^3, and A the mass number, in units $\frac{\rm g}{\rm mol}$. Once λ is known the interaction probability $\mathcal{P}_{\rm int}$ in a layer of thickness t can be calculated:

$$\mathcal{P}_{\rm int} = 1 - e^{-\frac{t}{\lambda}}. \tag{6.2}$$

When using tabulated values for the cross section in lead and iron [175], the interaction probability in the preshower can be calculated. As seen from Figure 6.8 the interaction probability in the lead-iron sandwich for photons with an energy range between 1 and 10 GeV is about $79 - 81$ %. About 80 % of the photons will be converted to a particle shower after passing the preshower.

It is expected that about 20 % of the photons does not interact in the lead-iron sandwich of the preshower. After traversing the preshower, these photons stay photons. The interaction probability of photons with an energy above 1 GeV in the plastic of the scintillator is negligible. Accordingly, about 20% of all photons are expected not to leave any signal in the preshower.

The data - Monte Carlo comparison for photons interacting in the HERMES preshower is depicted in Figure 6.9. A good agreement can be observed here as well. Furthermore, three distinct peaks can be seen: one corresponds with zero energy deposition, implying no interaction in the preshower, another one is the MIP peak, at about 2 MeV of energy deposition, and the last corresponds to the two-MIP peak. The one MIP peak is caused by interaction of either the positron or the electron in the scintillator paddle. The two MIP peak corresponds

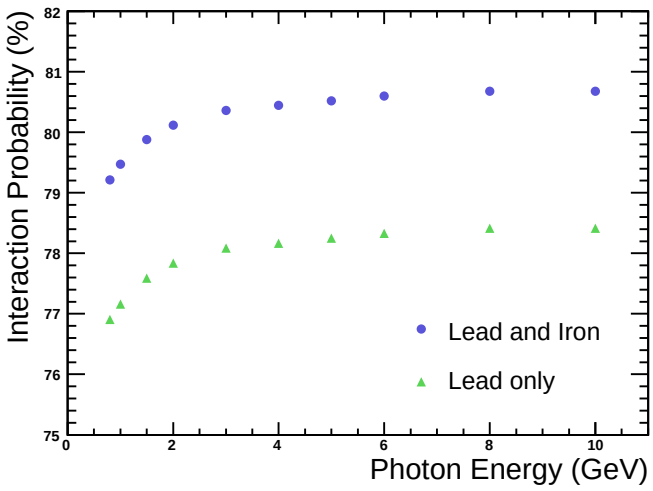

Figure 6.8: Interaction probability in the HERMES lead iron sandwich. Note the vertical scale.

to the situation where both electron and positron leave a signal in the same paddle. Apart from these three peaks a continuous tail can be seen in the figure. This tail corresponds to situations when a particle shower is established from the interaction of the photon with the lead of the preshower.

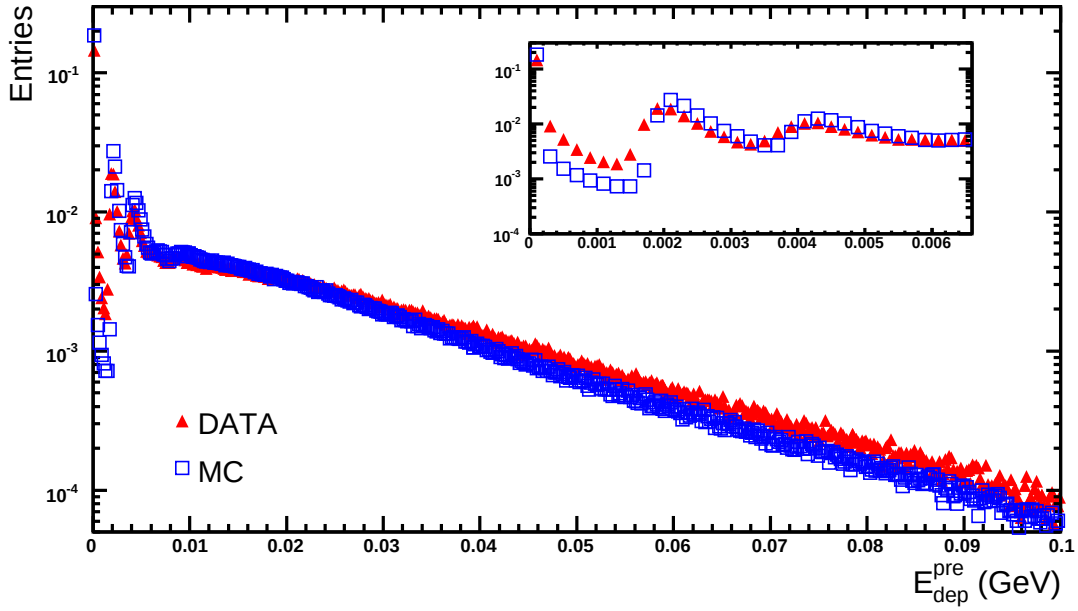

Figure 6.9: Preshower response to photons for data (red triangles) and Monte Carlo (blue squares). Both histograms are normalized to one. The inset is a magnification of the low energy deposition region ($E^{\mathrm{pre}}_{\mathrm{dep}} < 0.007$ GeV).

To estimate the fraction of non-interacting photons, the spectrum was normalized to 1, and integrated up to the first minimum around 0.014 GeV. For data (Monte Carlo) about 17.7 (19.3) % of the photons do not interact, in good agreement with the calculated values of Figure 6.8. The predicted values from Figure 6.8, however, are slightly higher than the observed ones. This can be explained by events where the photon-induced shower already starts in the RICH or TRD detectors which precede the preshower.

6.2.3 Response to hadrons

The interaction of hadrons with lead is fundamentally different than the one of leptons. In principle, hadrons can emit Bremsstrahlung radiation, however, the cross section of this process goes as $\sim \frac{1}{m^2}$. Even for the lightest hadron, the pion, the cross section for Bremsstrahlung is about $30\,000$ times smaller than for electrons. For energies below about 100 GeV the energy loss via Bremsstrahlung can be safely ignored for these particles. Instead, hadrons will lose energy by collisional losses, as described by the Bethe-Bloch formula (equation 5.2), or by strong interactions with atomic nuclei. The latter is responsible for the development of hadronic showers. The interaction length for these strong interactions, however, is typically much larger for high Z materials than the radiation length for electromagnetic interactions. In lead, for example, the nuclear interaction length is 17.1 cm. In the case of the HERMES preshower it is safe to assume hadrons not to interact in the lead of the preshower. After all, the purpose of installing the preshower was to improve electron-hadron separation. Most of the hadrons, however, leave a MIP signal in the H2, due to collisional losses in the plastic scintillator.

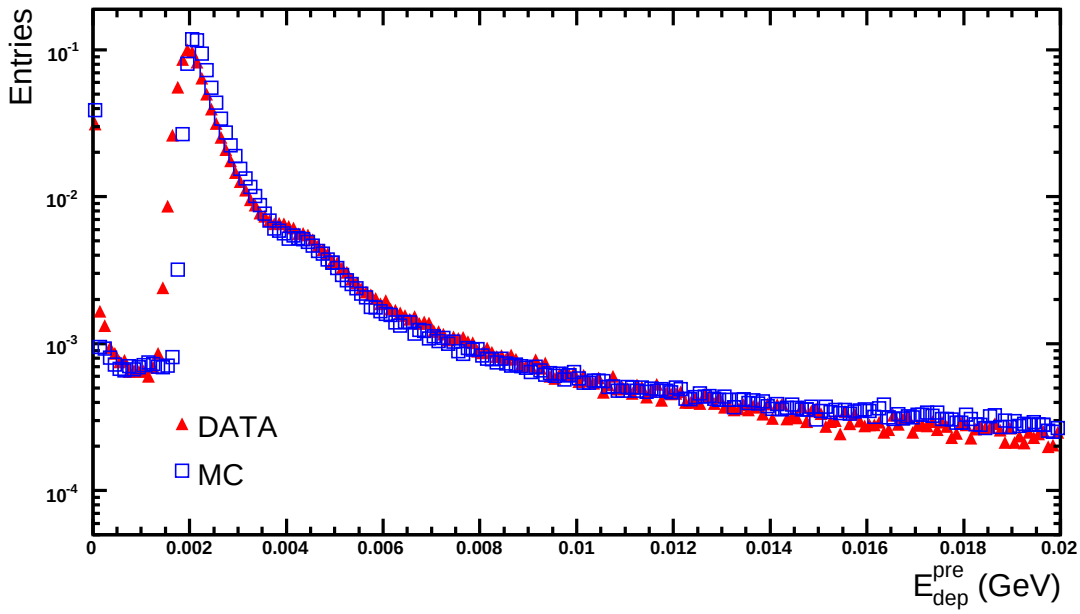

Figure 6.10: Monte Carlo (blue squares) and data (red triangles) comparison of the preshower response to hadrons. The histograms are normalized to one.

The preshower response to hadrons was compared between data and Monte Carlo. As seen from Figure 6.10 there is a good agreement. Also here both histograms were normalized to unity. A clear peak around 2 MeV corresponding to the MIP peak can be observed. The shoulder around about 4 MeV is partially caused by the situation where two hadrons are hitting the same paddle. The distribution in Figure 6.10 has a Landau shape due to δ-rays, i.e. electrons coming from ionization of target atoms caused by the incident radiation. The δ-rays are responsible for the tail of the distribution and part of the shoulder.

From the comparison between data and Monte Carlo it can be concluded that the improved description of the preshower simulates the behavior of leptons, photons and hadrons satisfactorily. Small deviations are still expected since the scintillator paddles are staggered in the experimental setup (see Figure 3.21) while in the geometry file, they are simply placed next to each other.

6.3 Simulation of the calorimeter

This section covers the simulation of the HERMES lead glass calorimeter. As explained in section 3.3.2.2 on page 63 the calorimeter is built out of 2×420 lead glass blocks. Each of the blocks measures $9 \times 9 \times 50$ cm. The blocks are wrapped with an aluminized Mylar foil which reflects the Čerenkov light created by the charged particles in the shower and with a Tedlar foil to make the blocks light-tight. A PMT is glued to the end of the blocks.

As mentioned before, it was known that the implementation of the calorimeter in HMC was imperfect. The simulation of an electromagnetic shower is difficult since a lot of particles are involved. Electron induced showers start when a Bremsstrahlung photon is emitted which then converts to an electron-positron pair. The secondary electron and positron pairs emit Bremsstrahlung themselves, which again converts to electron-positron pairs. A particle shower accordingly is created. The shower development continues until a critical energy E_c is reached, which is the energy at which the energy loss through Bremsstrahlung equals the energy loss due to ionization. This section will cover the description and implementation of an improved calorimeter simulation in HMC.

Figure 6.11 schematically shows the layout of the calorimeter simulation. The idea is to use a Look-Up Table (LUT) to represent the photo-electron yield per millimeter track length. The LUT takes the conversion from Čerenkov photons to photo-electrons into account, including various effects to the Čerenkov photons, which are detailed below. GEANT3 on the other hand is responsible for the generation of the electromagnetic shower. Since the output of HMC needs to match the input of HRC, the digitized response of a calorimeter block needs to be converted to an energy value. More details about this digitization process are given in section 6.3.4.

6.3.1 Čerenkov photons in a particle shower

The electric signal in the calorimeter comes from the Čerenkov light emitted by the charged particles in the shower. The Čerenkov photons propagate through the lead glass and are converted to photo-electrons at the surface of the PMTs.

All particles with a velocity larger than the speed of light in a medium ($\beta \cdot n > 1$) emit Čerenkov radiation. The energy loss of a particle via Čerenkov emission is very small, about

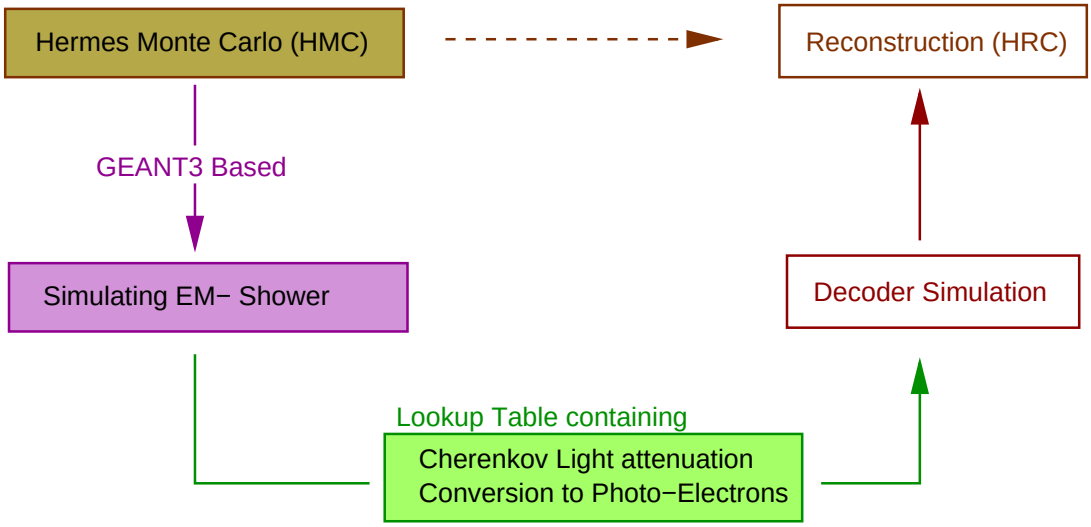

Figure 6.11: Schematic layout of the calorimeter simulation.

10^{-3} MeV cm^2 g $^{-1}$. Most energy is dissipated through Bremsstrahlung or ionization. The differential amount of Čerenkov photons emitted by a charged particle is given by:

$$\frac{\mathrm{d}^2 N^{\gamma}_{\mathrm{Cher}}}{\mathrm{d}\lambda \mathrm{d}x} = \frac{2\pi z^2 \alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right), \tag{6.3}$$

where α is the fine structure constant, z the particle's charge, and λ the wavelength of the Čerenkov photon. Typical PMTs are sensitive to light with wavelengths between 380 and 700 nm. Integrating equation 6.3 between these values, assuming that the Čerenkov angle $\theta_C (= \arccos \frac{1}{\beta n(\lambda)})$ depends only weakly on the wavelength, gives the following numerical value:

$$\frac{\mathrm{d}N^{\gamma}_{\mathrm{cher}}}{\mathrm{d}x} = \int_{380}^{700} \frac{2\pi z^2 \alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right)\mathrm{d}\lambda = 551.73 \cdot z^2 \cdot (1 - \cos^2\theta_C). \tag{6.4}$$

GEANT is in principle able to track the Čerenkov photons generated by the charged particles in the shower through the lead glass until they reach the PMTs. This approach, however, would be extremely time consuming, even on state-of-the-art CPUs, as a tremendous amount of photons would have to be tracked in each event. Therefore a different approach was needed.

The method chosen for the simulation of the HERMES calorimeter consists of calculating the amount of Čerenkov photons 'by hand' in a user routine in HMC. The routine is executed after every tracking step. If a particle is charged and if $\beta \cdot n > 1$, the amount of Čerenkov photons is calculated as:

$$N^{\gamma}_{\mathrm{Cher}} = 551.73(1 - \cos^2 \frac{1}{\beta \cdot n}) \cdot \Delta x, \tag{6.5}$$

where Δx is the step-length.

Before reaching the PMT, the Čerenkov photons are subject to various interactions. The photons undergo a strong attenuation in the lead glass, especially for short wavelengths (blue

light). This effect is enhanced by the chemical composition of the F101 blocks, which are composed out of 51.23% Pb_3O_4, 41.53% $Si0_2$, 7% K_2O and 0.2% Cerium. The presence of Cerium in the crystals makes the material radiation hard, while at the same time decreasing the transparency. The transmittance of the F101 blocks was measured as a function of the wavelength and is depicted in Figure 6.12(a). The transmittance T is defined as $T = I/I_0$, where I_0 is the intensity of the incident light and I the intensity of the light coming out of the sample. The two curves in the figure indicate the limits of the measurement.

The photons additionally undergo refraction at the boundaries between all surfaces. There are boundaries between the lead glass and the aluminized Mylar foils, between the lead glass and the silicon glue, and between the glue and the PMT surface. These refractions are indicated in panel (b) of Figure 6.12.

A third influence to the Čerenkov light is the reflection of the photons at the aluminized foils. The effect of aluminizing the outer side of the Mylar foil can be seen in panel (b) of the figure. Panel (c) shows a measurement of the reflection of the calorimeter wrapping. The absorption loss (= 1 - reflectivity) was measured at two different wavelengths. As a reference measurement, the reflectivity of a 100 % reflective foil produced by 3M was performed.

Finally, in order to estimate the amount of photo-electrons created from the Čerenkov light, the quantum efficiency of the PMTs has to be taken into account. The PMTs used for the HERMES experiment are Philips XP3461 PMTs, whose quantum efficiency is plotted in Figure 6.12(d).

6.3.2 The Look-Up Table for the calorimeter simulation

The idea of the improved calorimeter simulation is to include all effects which the Čerenkov photons are subject to in a Look-Up Table (LUT). This table incorporates the conversion of optical Čerenkov photons to photo-electrons in the PMT as well. The LUT contains the number of photo-electrons per millimeter track length. The functionality of the LUT is schematically shown in Figure 6.11.

Since electrons and positrons in the particle shower have momentum directions in the forward as well as in the backward direction, the corresponding Čerenkov photons will be emitted in either directions. Moreover, the Čerenkov production angle is velocity dependent. Therefore, the LUT needs to be dependent on the electron momentum at least.

An example of Čerenkov emission in the forward and backward direction is shown in Figure 6.13, where the projection into the (y, z) plane of a calorimeter block is shown. The upper (lower) panel gives the situation where a particle located at the axis of the block has a forward (backward) momentum direction, thus emitting Čerenkov radiation in the forward (backward) direction.

Apart from being momentum dependent, the LUT also needs to be dependent on the geometrical position where the Čerenkov radiation is emitted. Indeed, the collection efficiency depends on position, while the amount of attenuation clearly is dependent on the longitudinal position. For all of the above reasons the LUT depends on the $x, y, z, \theta_p\,(= \arccos \frac{p_z}{|\vec{p}|}), \phi_p\,(= \arctan \frac{p_x}{p_y})$ and $\beta(= \frac{v}{c})$ coordinates of the Čerenkov radiating particle.

For symmetry reasons the table was generated for one block octant only. A stand alone GEANT4 based simulation of a calorimeter block octant was developed [177] in order to generate the table. The propagation of optical photons was simulated in detail including the

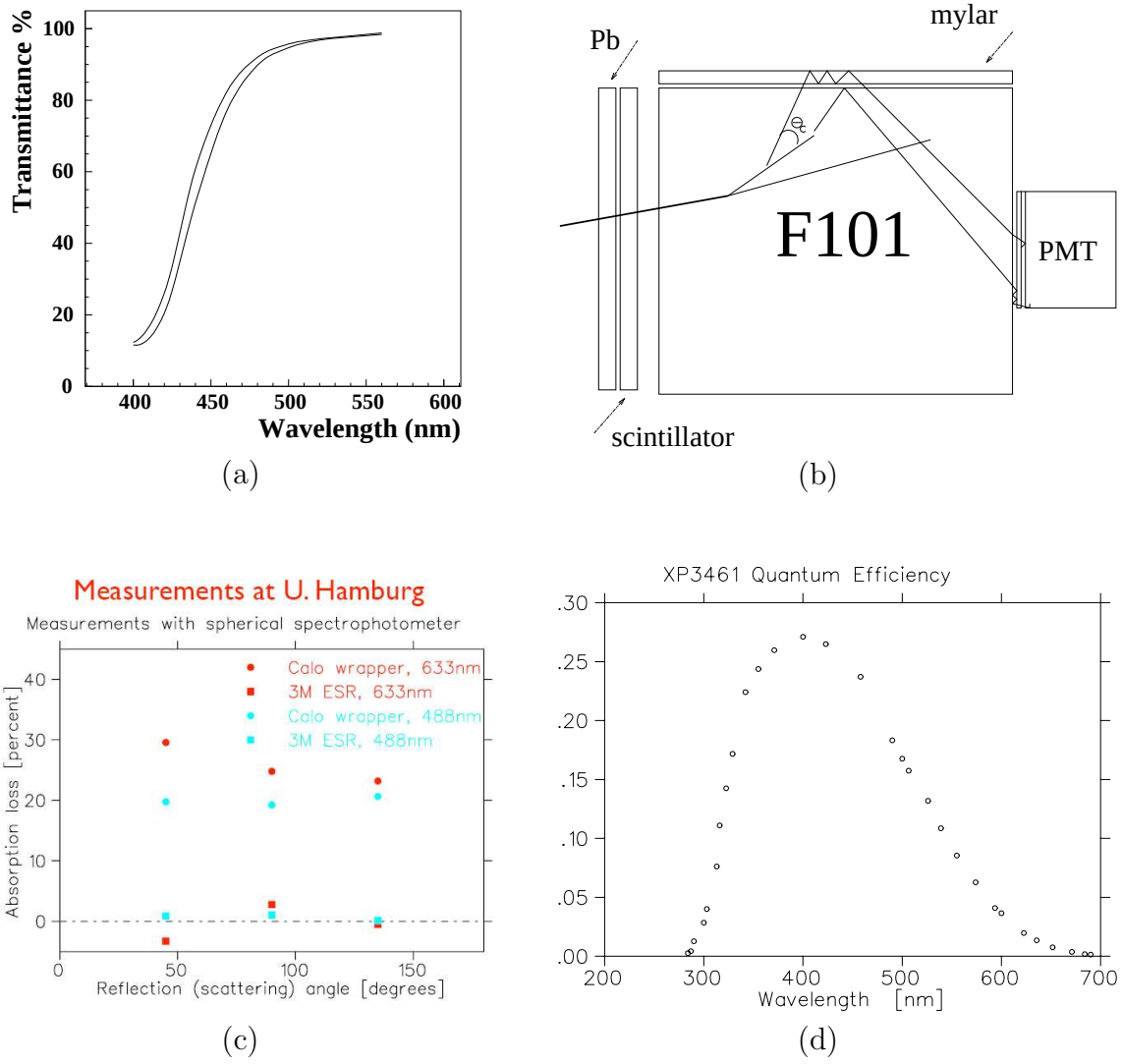

Figure 6.12: Various effects to the Čerenkov light: (a) shows the attenuation in the lead glass, (b) indicates the refraction and reflection at the different surfaces, (c) shows a measurement of the absorption loss of the aluminized Mylar surrounding the calorimeter blocks at two different wavelengths, and (d) shows the quantum efficiency of the PMT. (a) and (b) from [176], (c) from [177].

reflectivity of the foils, the transparency of the lead glass, as well as the refraction at all surfaces. Moreover, the quantum efficiency of the PMTs was included.

The most rapid variations of photo-electron yield were observed for the θ_p dependence of the yield. Consequently the segmentation of the table should be finest in θ_p. Rapid variations as a function of z and β were also observed. The table segmentation is indicated in Table 6.1 and is based on empirical observations of the photo-electron yield as a function of the coordinates. In an attempt to average over very rapid fluctuations the $(x, y, z, \theta_p, \phi_p, \beta)$ points are smeared around the specified bin values in the column labeled 'Generated Range' in

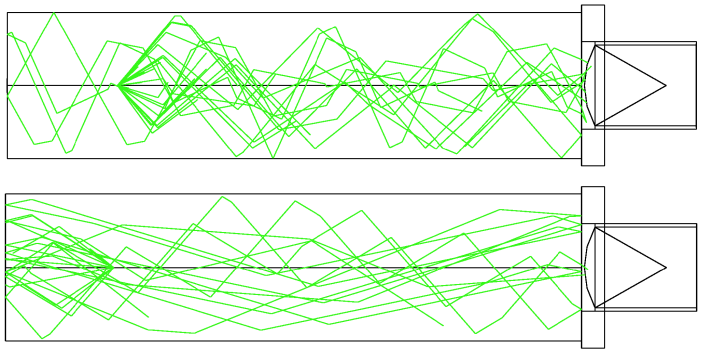

Figure 6.13: The upper (lower) panel shows the generation of Čerenkov light in the forward (backward) direction by a charged particle located at the axis (represented by the thick line) at about $z = 10$ cm in the block with a forward (backward) oriented momentum direction. The Čerenkov cone is projected in the (y, z) plane. Čerenkov photons are represented by the green lines. Multiple reflections can be seen inside the lead glass block. A different amount of photons reach the PMT surface in both cases. Figure from Reference [177].

Generated Range	Variable Range	Steps
$0 < x < y$	$0 < x < y$	10
$0 < y < 44$	$0 < y < 43.7$	10
$-249 < z < 249$	$-240 < z < 248$	16
$0 < \theta_p < 3.14159$	$0.014 < \theta_p < 3.126$	39
$0 < \phi_p < 6.2832$	$0.058 < \phi_p < 6.241$	10
$0.595 < \beta < 1.00$	$0.596 < \beta < 1.00$	16

Table 6.1: Segmentation of the calorimeter Look Up Table. The difference between 'Generated Range' and 'Variable Range' is explained in the text.

the table. The actual values are averaged over this smearing and given in the column labeled 'Variable Range'. The segmentation in x, y, θ_p and ϕ_p is done in equidistant steps over the generated range. For the z-range, however, the bin size is getting smaller towards the end of the block, since at these positions, more variations are expected. For a similar reason, the segmentation in β is finest for low β values.

For each bin in the 6 dimensional table, 1000 electrons were propagated along a one millimeter long track segment at the position and angle corresponding to that bin. The resulting optical photons were collected and recorded as the response for that bin. It was impractical to convert the entire calorimeter simulation to GEANT4 because of the large effort involved in redefining the geometry for the entire spectrometer. It took about 50 CPU days at WestGrid in Vancouver for the table to be calculated.

6.3.3 Coordinate transformation and table interpolation

For symmetry reasons, the LUT was generated for one block octant only. Hence, a coordinate transformation is required to transform from the HERMES coordinate system to the Look-Up

Table coordinate system. This is achieved by two subsequent transformations: a translation to a central position followed by a rotation to the desired quadrant. Depending on the position an additional mirroring is required to transform the coordinates to the correct octant. Not only (x, y) coordinates need to be translated, but also the momentum vector ϕ_p. An example of such a transformation is given in Figure 6.14. 8 points in the lower left corner are transformed to the reference octant in the upper right as an example. The LUT is only defined in the upper right octant, colored in gray in the figure. In this specific example, the 8 points are rotated over 180 degrees and mirrored around the line $x = y$. The ϕ_p momenta associated with these points are indicated by the arrow directions. These are also transformed by the algorithm. It can be verified from the figure that the implementation of the translation of position and momentum in a calorimeter block was carried out correctly.

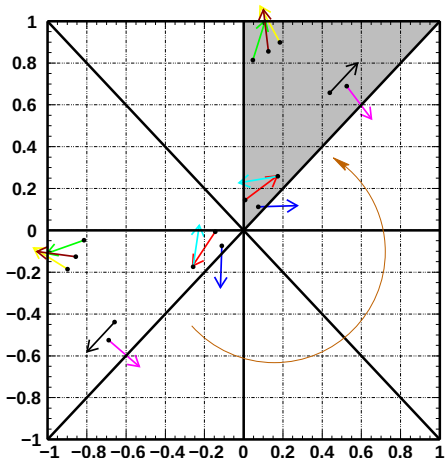

Figure 6.14: Cross sectional view of a calorimeter block. 8 points from the lower left octant are transferred to the upper right reference octant as indicated by the long arrow. The reference octant is colored gray in the figure. The short arrow directions indicate ϕ_p, which also need to be transformed when going to the reference octant.

In order to get the number of photo-electrons associated with a certain particle track position and direction, an interpolation in the Look-Up Table is required. There are various interpolation algorithms available for interpolating multidimensional tables. The algorithm to interpolate the table should be fast and precise. An interpolation is required for every mm charged track in the particle shower for which $\beta \cdot n > 1$. Therefore, care must be taken to select the appropriate interpolation method.

The easiest and very common algorithm is the linear interpolation method. The interpolated value is calculated by constructing a straight line through two points of the table and by evaluating this first order polynomial at the point of interest. For multidimensional tables, the algorithm becomes 'multi-linear': to interpolate in a table of n dimensions, $2^n - 1$ interpolations are required. An example of multidimensional linear interpolation in 3 dimensions is drawn in Figure 6.15. If instead of a linear interpolation between two points, polynomials of order m_i are used for dimension i, $\Pi_i(m_i + 1) - 1$ interpolations are required. It should be noted that polynomials of higher degrees do not necessarily improve the interpolated result. In particular, when the table needs to be extrapolated, polynomial functions tend to diverge faster than linear functions.

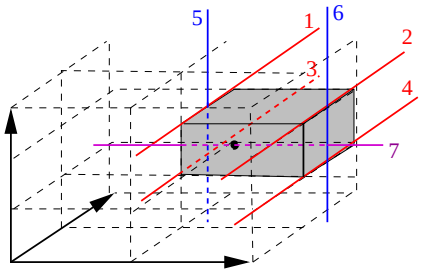

Figure 6.15: For a linear interpolation in 3 dimensions, 7 interpolations are required.

Another very common interpolation technique is the cubic spline method, where a cubic function is used to interpolate the table, with the requirement that the first and second derivative are continuous. This additional requirement results in longer calculation times. Therefore this algorithm was not considered for the implementation of the calorimeter Look-Up table interpolation.

A third interpolation algorithm investigated is the so-called 'simplex' method [178, 179, 180]. It is claimed to be the fastest interpolation technique possible [178]. The method is schematically drawn in Figure 6.16 in the case of three dimensions. A numerical example to illustrate the method will be given later. The simplex method consists of different steps: first, an $n-$ dimensional hypercube is to be found around the point $\mathbf{x} = (x_1 \cdots x_n)$. This hypercube is then transformed to the $n-$ dimensional unit cube. The unit cube can now be further divided into simplexes, which are the $n-$ dimensional analogues of triangles. In three dimensions, these volumes are tetrahedrons. The right simplex can be found by sorting the elements of the vector $\mathbf{x} = (x_1 \cdots x_n)$ in non-increasing order: $(x_{p(1)} \cdots x_{p(n)})$, where $p(1) \cdots p(n)$ is the permutation needed to sort the elements in that order. Once the right simplex is found, its $n+1$ vertices are used to construct a piecewise linear interpolation between the vertices of the simplex:

$$\hat{f}(\mathbf{x}) = f(X_{K+A_0})(1 - x^{p(1)}) + \left(\sum_{i=1}^{n-1} f\left(X_{K+A_i}\right) \left(x^{p(i)} - x^{p(i+1)}\right) \right) + f(X_{K+A_n}) x^{p(n)}, \quad (6.6)$$

where $\hat{f}(\mathbf{x})$ is the interpolated value, and $f(X_{K+A_i})$ the tabulated values at the $n+1$ vertexes $X_{K+A_0} \cdots X_{K+A_n}$. K is accounting for the required permutations to sort the elements of $\mathbf{X}$ in non-increasing order.

To illustrate this method, a three dimensional example is considered in the following. Consider function values defined for every point (x, y, z) of the three dimensional space where x, y and z are integer. Let us additionally assume that we only have z values for odd numbers. Suppose now that the function value at the vector $\mathbf{x} = (4.2, 5.7, 6.2)$ needs to be evaluated. The following steps need to be considered:

1. Find a three dimensional cube around the vector: this is easy, the corner points of the cube are $(4,5,5)$,$(5,5,5)$,$(5,6,5)$,$(5,6,7)$,$(4,6,5)$, and $(4,6,7)$.

2. Transform the point into a unit-cube: this is done by applying the following formula:

$$x'_i = \frac{x_i - X_i^k}{X_i^{k+1} - X_i^k},$$

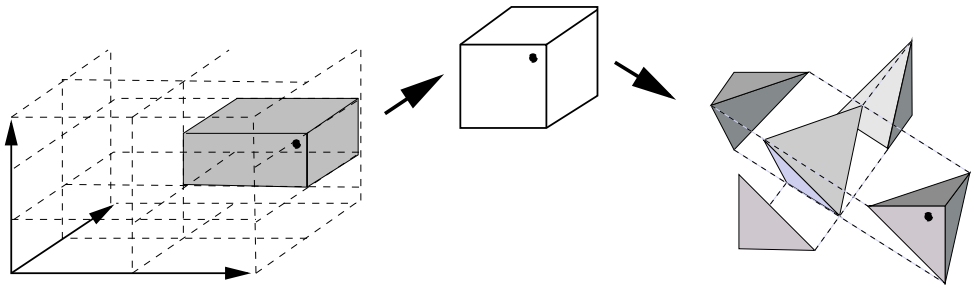

Figure 6.16: Schematic overview of the simplex interpolation method: a hypercube around the point to be interpolated, represented by the black circle, is transformed to the unit cube, which can then be further divided into simplexes.

with x'_i coordinate i of the vector in the unit cube, x_i the i^{th} coordinate of the original vector, and X_i^k the points of the hypercube surrounding the original vector. After this transformation we can write $\mathbf{x}' = (0.2, 0.7, 0.6)$.

3. Now we have to find the appropriate simplex by applying a permutation $\mathbf{p}$ which orders the elements of $\mathbf{x}'$ in non-increasing order. In this example we have $p(1) = 3$, $p(2) = 1$ and $p(3) = 2$. The permuted vector now reads: $\mathbf{x'^p} = (0.7, 0.6, 0.2)$. The corner points of the appropriate simplex correspondingly are: $(0,0,0)$,$(0,0,1)$,$(1,0,1)$ and $(1,1,1)$.

4. Transform the simplex back to the hypercube: the corners of the simplex surrounding the original vector $\mathbf{x} = (4.2, 5.7, 6.2)$ are: $(4,5,5)$, $(4,5,7)$,$(5,5,7)$ and $(5,6,7)$.

5. Perform the interpolation: Applying equation 6.6 we can write:

$$\hat{f}(\mathbf{x}) = f(4,5,5) \cdot (0.3) + f(4,5,7) \cdot 0.1 + f(5,5,7) \cdot 0.4 + f(5,6,7) \cdot 0.2$$

This concludes the example.

Two interpolation techniques were compared and tested for potential use in the calorimeter simulation. One is based on multidimensional linear interpolation, and the other one based on the simplex method. The multidimensional linear interpolation was implemented using the `FINT` routine from the `CERNLIB` package [181]. This algorithm, however, is only able to interpolate up to five dimensions. Therefore, the FINT routine was modified to be able to handle the required six dimension. The implementation of the simplex method is based on the description found in Reference [178].

A comparison between the simplex and the FINT method is given in Figure 6.17. The left panel shows the interpolation of 100 random vectors with both algorithms. One can conclude that all interpolations agree well for both methods. On the right hand side of the figure, the time needed for each interpolation is histogrammed. The calculations were performed on a 1 GHz Pentium III machine. The simplex method is more than twice as fast as the FINT method.

As seen from Table 6.1, the coordinate limits of the LUT do not correspond to the physical limits of the coordinates y, z, θ_p and ϕ_p[4]. The algorithm consequently also needs to be able to

[4]Given an $n = 1.6522$ for F101, particles will only emit Čerenkov radiation if $\beta > 0.6$.

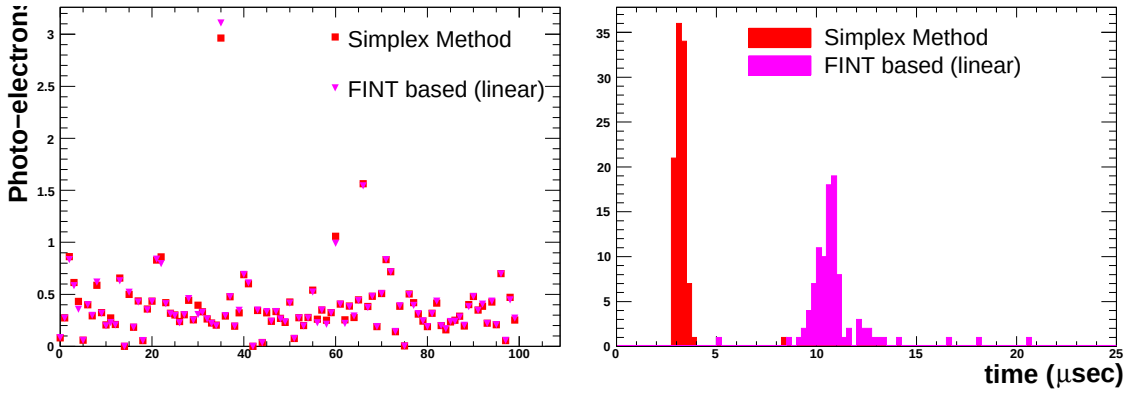

Figure 6.17: Comparison between the simplex and FINT interpolation techniques. The left panel shows the result of the interpolation for 100 random vectors. The right panel shows the time needed to interpolate the table.

Variable Range	Steps
$0 < x < y$	11
$0 < y < 45$	11
$-250 < z < 250$	18
$0 < \theta_p < 3.141593$	41
$0 < \phi_p < 6.29$	12
$0 < \beta < 1.00$	17

Table 6.2: Segmentation of the calorimeter Look Up Table for the simplex algorithm.

extrapolate the LUT. In the FINT routine, extrapolation is implemented. The simplex algorithm, on the other hand, is unable to handle extrapolations. Therefore, the original LUT was extended using the FINT method to extrapolate to the physical limits of all coordinates. The resulting table segmentation is given in Table 6.2. Depending on the interpolation algorithm chosen, a different Look-Up table has to be used.

6.3.4 Decoder simulation

Once the number of photo-electrons per mm track length is known from interpolating the table, the digitization in the ADCs needs to be simulated. This is done in various steps. First of all a Poisson fluctuation is added to the number of photo-electrons, followed by the conversion of the calorimeter block's response to an integer ADC value:

$$N_{\mathrm{ADC}} = \frac{\texttt{gPoiss}(rPulse) \cdot \mathrm{CaloNumberOfADCChan}}{\mathrm{CaloEnergyFullScale}},$$

where CaloNumberOfADCChan is 8192, and CaloEnergyFullScale 24000. The latter value was chosen to avoid ADC-saturation for incident electron energies below 30 GeV. Next thing to be included in the simulation is the electronic noise. The noise is added to a channel's pedestal

by applying the following formula:

$$ped = \mathrm{CaloPedChan} + \texttt{granor}() \cdot \mathrm{CaloPedWidth}/2.3,$$

where CaloPedChan is a channel's average pedestal, set to 300, and CaloPedWidth the full width at half maximum of the calorimeter ADC pedestal distribution, which is assumed to be equal to 4. The function `granor()` is an intrinsic GEANT function giving two random numbers according to a normal distribution. In the case presented here only one of the numbers is used. Once the pedestal is known, the number of ADC values is recalculated as:

$$N_{\mathrm{ADC}} = max(N_{\mathrm{ADC}} + ped, \mathrm{CaloNumberOfADCChan}).$$

Next, the average pedestal is taken out of the simulated ADC response again: $N_{\mathrm{ADC}} = N_{\mathrm{ADC}} - \mathrm{CaloPedChan}$. At this stage, we have fully simulated the HERMES calorimeter data acquisition. The output of HMC, however needs to mimic the output of the decoded HERMES data. Therefore, the number of ADC counts, N_{ADC}, needs to be converted to an energy value:

$$rPulse = \frac{N_{\mathrm{ADC}} * \mathrm{CaloEnergyFullScale}}{\mathrm{CaloNumberOfADCChan}}.$$

An additional scale factor is needed to convert the value of $rPulse$, which is expressed in units of photo-electrons, to a value in GeV. Therefore $rPulse$ needs to be multiplied by a scale factor: $rPulse = rPulse * \texttt{CaloPBGlassScaleFactor}$. The latter factor was determined by requiring $\frac{E_{\mathrm{gen}}}{E_{\mathrm{rec}}}$ to be equal to 1 for electrons. For the FINT method the value `CaloPBGlassScaleFactor` was found to be 0.001665, while for the simplex method it is 0.001647.

6.4 Results of the simulation

The simulation of the HERMES calorimeter has been tested extensively. The implementation of the LUT was tested, the longitudinal shower profile was analyzed and the generation of Čerenkov radiation verified. A comparison with the stand-alone GEANT4 based simulation, which was used to generate the Look-Up Table was done as well. This section reports on all of these tests. The comparison with experimental data is reported in the next section. data.

6.4.1 Testing the implementation of the Look-Up Table

One of the things to be verified was the implementation of the Look-Up Table and its interpolation in HMC. This was done by using high energy muons, perpendicularly incident on the calorimeter's surface. 100 GeV muons offer the simplification of eliminating the shower simulation; they keep on going straight without much deflection. The x,y, θ_p, ϕ_p and β thus stayed constant, so that the z-dependence only could be tested.

The z dependence of the photo-electron yield per mm track length is shown in Figure 6.18, for various (x, y) positions. The tabulated values are represented by the open circles. Both interpolation methods are able to reproduce the rapid variations in the photo electron yield given by the Look-Up Table, for all indicated positions, proving a good functionality of the algorithms. The figure at the same time an indicates the rapid variations in the Look-Up Table, and the segmentation of the table in z. The dependence of the tabulated values on the other variables shows similar large variations.

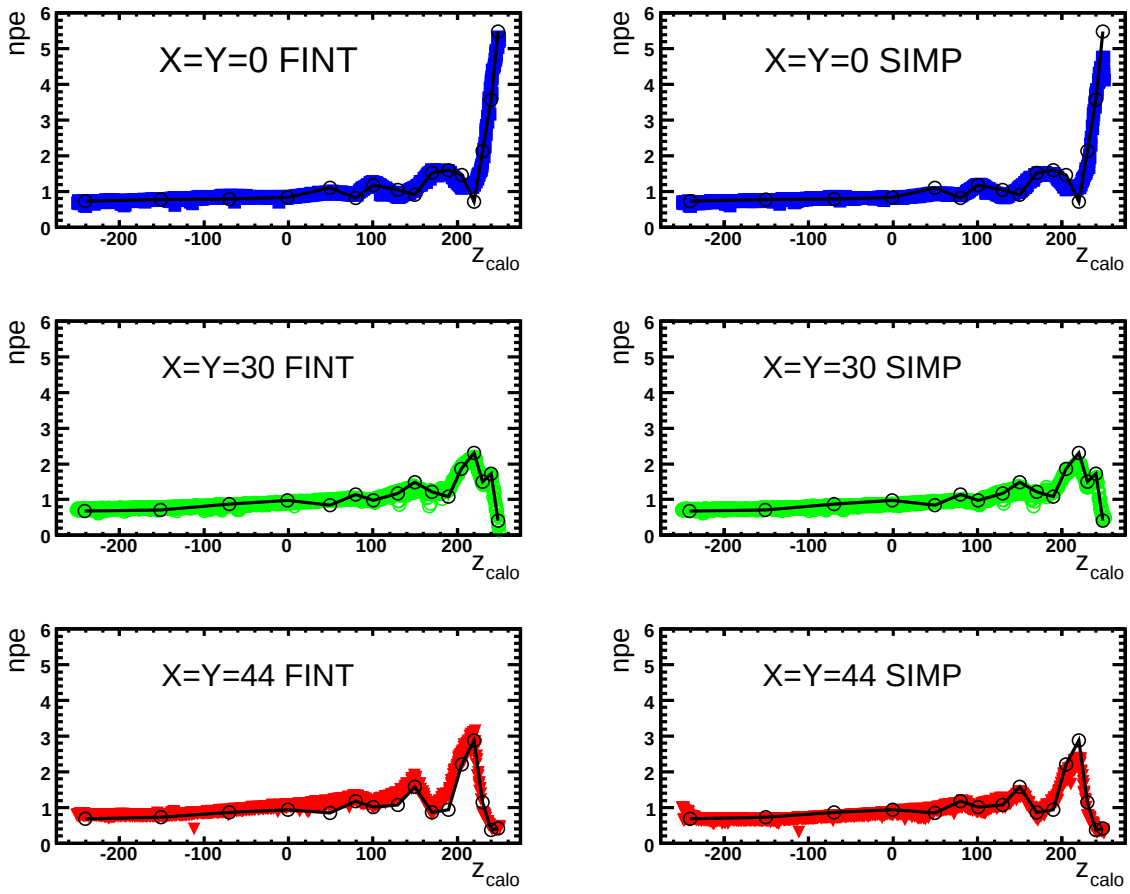

Figure 6.18: Number of photo-electrons (npe) per millimeter track length as a function of longitudinal position z. The tabulated values are represented by the open circles. It is seen that the segmentation in z is finest near the end of a block. The figures on the left (right) are made using the FINT (simplex) interpolation technique for muons traversing the lead-glass blocks. Muon positions in the lead glass from top to bottom are $(x, y) = (0, 0)$, $(30, 30)$ and $(44, 44)$. The upper (lower) panel thus corresponds to the center (edge) of a block.

6.4.2 Longitudinal shower profile in GEANT

The analysis of electromagnetic showers generated by HMC/GEANT3 was an important test of the simulation. It does not make sense to continue optimizing the simulation if already the shower itself is not described very well. For various incident energies the longitudinal profile was fitted by a function of the form:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}, \tag{6.7}$$

with E_0 the incident energy, t the depth in radiation lengths, and a and b fit parameters [182]. This was done for electrons as well as for photons, both perpendicularly incident on the calorimeter's surface.

Results of the simulation and a fit to the shower profile can be seen in Figure 6.19. The left panel shows the results for electrons, the right panel these for photon induced electromagnetic shower. The error bars in the right panel are the RMS of the distribution, they indicate the large spread in longitudinal shower profile. The error bars on the left are the error on the mean value, these form an indication of the statistical precision of the simulation. As expected, photon induced showers penetrate deeper in the lead-glass, since photons have a larger mean free path than electrons. The simulated profiles are well described by equation 6.7, indicating a good simulation of the longitudinal shower profile.

The option `IABAN=2` was selected for GEANT3, so that particles are tracked down to a user controlled cut off energy, rather than the default cut off of 10 MeV. The latter value is the default option and is set to increase computational efficiency. For electromagnetic showers, however, this value is too large. The cut off energy was chosen to be 200 keV.

From the right panel of Figure 6.19 it can be concluded that the shower profile has large fluctuations. The resolution of lead-glass calorimeters is limited by these fluctuations. A strong attenuation of optical photons in the lead glass enhances the effect of shower fluctuations. As can be seen from Figure 6.12 the attenuation is strongest for wavelengths between 350 and 475 nm. A typical high PMT quantum efficiency for these wavelengths, makes this effect even worse[5].

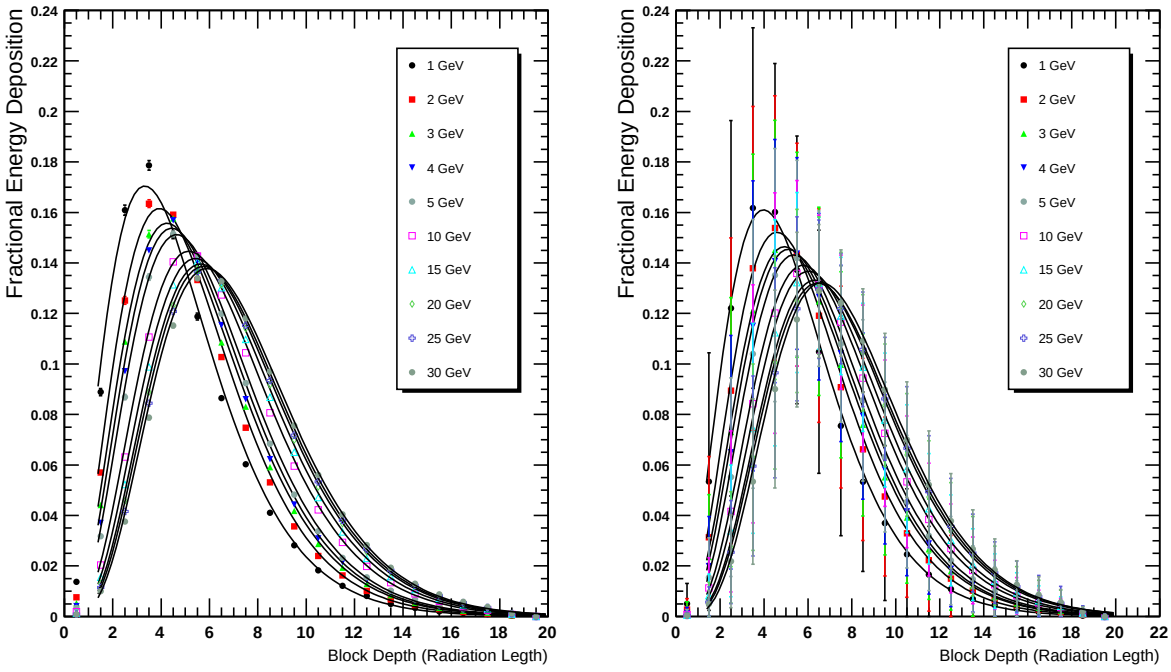

Figure 6.19: The left (right) panel shows the longitudinal shower profile for electron (photon) induced electromagnetic showers. The error bars in the right panel are the RMS of the distribution, those in the left panel the error on the mean. Longitudinal shower profiles have large RMS values. The right panel shows the statistical accuracy of the simulation.

[5]In Reference [183] an improved resolution is reported by making use of a blue light filter in front of the PMTs. With less light, these authors obtained a better resolution.

function	Leptons		Photons	
	p_1	p_2	p_1	p_2
Amaldi	0.017	0.4	0.017	1.2
Fit to mean	0.0189	0.344	0.021	1.02
Fit to median	0.0274	−1.21	0.022	−0.57

Table 6.3: Fit parameters p_1 and p_2 of the function $\ln\frac{E}{p_1} + p_2$ fitted to the mean and median energy deposition for both electron- and photon-initiated showers in lead glass. The functions are drawn in Figure 6.20.

From the longitudinal shower profile the mean depth of energy deposition can be easily determined:

$$< t >= \int_0^\infty t \cdot b\frac{(bt)^{a-1}e^{-bt}}{\Gamma(a)}\mathrm{d}t = \frac{a+1}{b}$$

The median of the profile, i.e. the value at which half of the energy is deposited, can be estimated by solving the following equation:

$$\int_0^{t_{\mathrm{median}}} b\frac{(bt)^{a-1}e^{-bt}}{\Gamma(a)}\mathrm{d}t = 0.5$$

In Reference [184], the median depth of energy deposition in an electromagnetic shower is given as:

$$t_{\mathrm{median}} = \ln\frac{E}{E_c} + C, \tag{6.8}$$

where C is 0.4 for electrons and 1.2 for photons, E_c is the critical energy, corresponding to 17.97 MeV for F101. Figure 6.20 shows the mean and median energy deposition in the calorimeter as a function of incident electron (left) or photon (right) energy. A function of the form $\ln\frac{E}{p_1} + p_2$ was fitted to the data points. Fit parameters are indicated in Table 6.3 Equation 6.8 is also shown in the figure. The function given by equation 6.8 describes the variation with incident energy reasonably, though not perfectly. This is an indication of an imperfection in the shower simulation. Surprisingly, however, the fit parameters to the *mean* energy deposition, are close to the predicted ones for the *median* of the distribution given by equation 6.8, while the fit to the median distribution gives a significantly different result. This indicates that the energy deposition represented by equation 6.8 corresponds to the mean energy deposition and not to the median, as mentioned in Reference [184].

Apart from verifications with empirical expectations, some electromagnetic shower observables were compared with results form the standalone GEANT4 based Monte Carlo, which was written to generate the Look-Up table. The escaping energy out of a nine block cluster as well as the total shower track length within the nine block cluster were compared, and are depicted in Figure 6.21. The error bars in the figure represent the RMS of the distributions. The dependence on energy of both observables obtained by the GEANT3 based HMC simulation is negligible in the case of GEANT3 , while the GEANT4 based Monte Carlo indicates a decrease in normalized shower track length, together with an increase in escaping energy. Moreover, larger fluctuations can be seen for the GEANT4 simulation. At the lowest energy points both simulations agree well. The GEANT4 behavior seems more natural from first principles: for

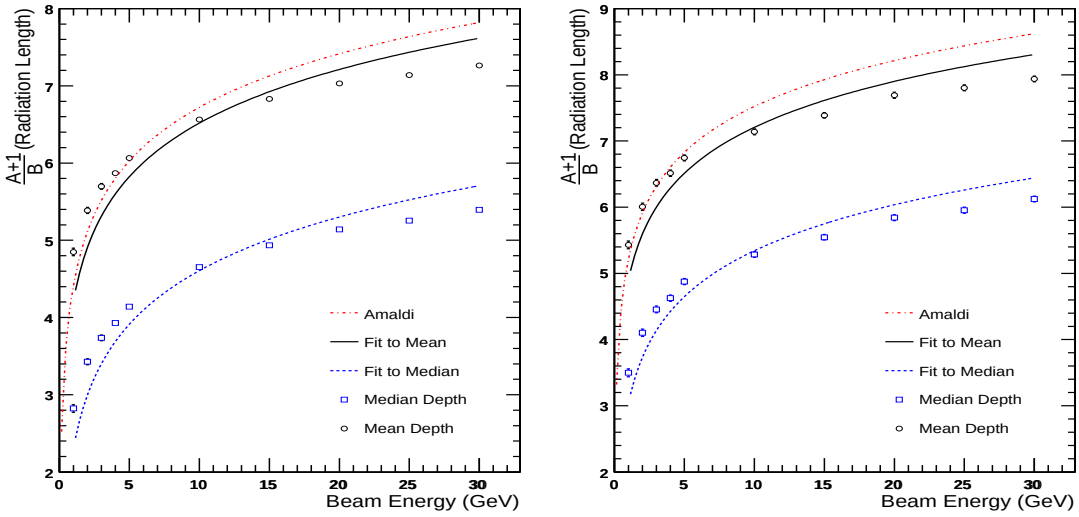

Figure 6.20: Mean (circles) and median (squares) energy deposition in the lead glass calorimeter as a function of incident electron (left) and photon (right) energy. The dash-dotted line represent the empirical expected value for the median (!) from Reference [184]. The full (dashed) line is a fit to the mean (median) energy deposition. Note the different y-scales in both panels.

higher energies, the transverse shower profile is larger, resulting in a larger fraction of the shower escaping from the nine block cluster. The observed differences are assumed to be an intrinsic property of both packages. Also in e.g. Reference [185] a difference in electromagnetic shower profile between GEANT3 and GEANT4 was reported. GEANT4 uses a different model for the description of multiple scattering [186]. Moreover, there are differences in the range of particles in detector material between both versions. Additionally, a different precision in machine numbers between GEANT3 and GEANT4 may result in non-negligible differences between the results of both simulation packages. GEANT3 is `FORTRAN` based, while GEANT4 is based on the `C++` framework.

From all of the above observations it can be concluded that the shower simulation by HMC is reasonably good. No evidence was found for large deviations compared to what is expected from empirical considerations, or compared to an improved description by GEANT4.

6.4.3 Testing Čerenkov light generation

The next test of the simulation was to check the Čerenkov light generation by HMC, and compare it to the amount of Čerenkov photons generated by GEANT4. Since muons with an energy of a couple GeV only lose their energy via collisional losses, the generation of Čerenkov light can be compared without the complication of particle shower development. The amount of light was determined by evaluating equation 6.4 after every step. The result for 1 and 10 GeV muons are shown in Figure 6.22. A very good agreement between the results of both packages can be seen.

The amount of Čerenkov light generated by electrons was also compared for various incident energies. The comparison of mean and width of the corresponding distributions, normalized

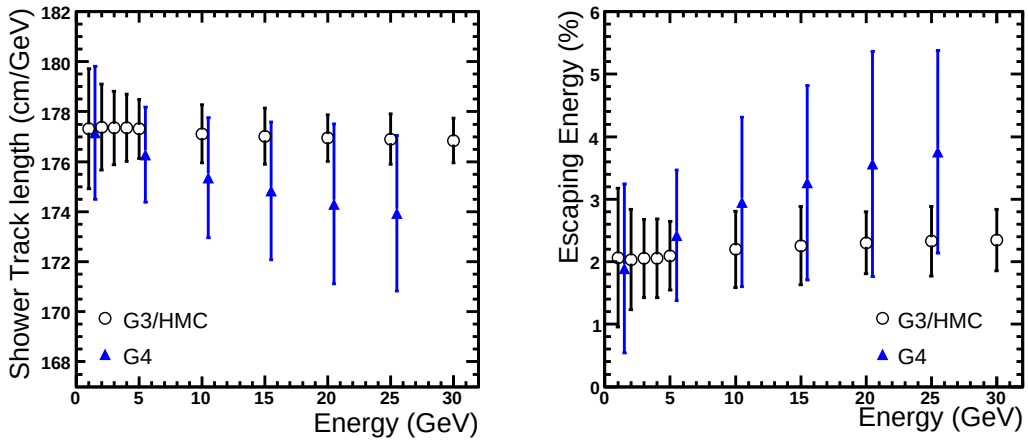

Figure 6.21: Left panel shows the normalized total shower track length in cm/GeV, the right panel the escaping energy from a 9 block cluster for both GEANT4 (triangles) and HMC/GEANT3 (circles), as a function of incident energy. The error bars indicate the RMS of the distributions. For visibility the GEANT4 points have an offset of +0.5 GeV.

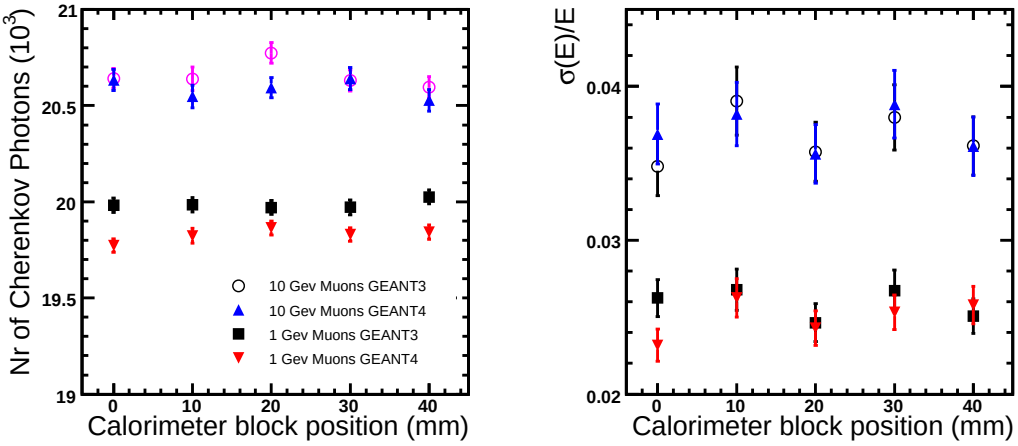

Figure 6.22: Number of Čerenkov photons for a 1 and 10 GeV muon beam as simulated with HMC and GEANT4. The standard deviation of the distribution is drawn in the right panel.

by the incident energy, is shown in Figure 6.23. GEANT4 finds about 56 photons per incident MeV, whereas HMC gets about 60 photons. This is a difference of about 7%, which is probably due to the difference in shower generation between both distributions. The GEANT4 based simulation produces also here larger fluctuations than the GEANT3 based package.

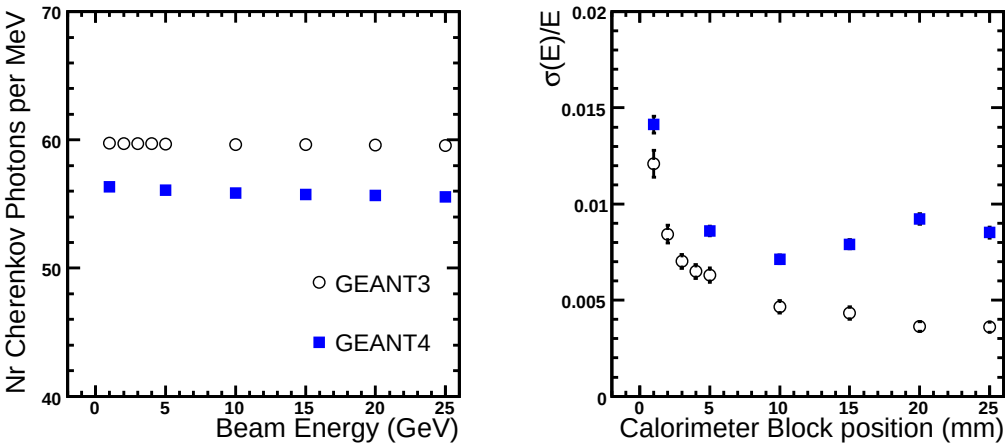

Figure 6.23: Number of Čerenkov photons per incident MeV as a function of incident energy. The standard deviation of the distribution is shown in the right panel.

6.4.4 Results for photo-electrons

Testing the number of photo-electrons was done with the same input particles as mentioned in section 6.4.3: the photo-electron yield for 1 and 10 GeV muons was compared as a function of calorimeter block position. Figure 6.24 shows the result of the simulation. The 1 GeV muon photo-electron yield distribution was fitted with a landau distribution, the 10 GeV photo-electron distribution with a Gaussian. A good agreement can be seen between the simplex and FINT interpolation algorithm. For both energies, the number of photo-electrons from HMC is roughly 15% larger than the number from the standalone GEANT4 code. Also here, the standard deviation of the GEANT4 simulation is larger than the one from HMC for both energies.

For the comparison with electron initiated showers, a scale factor was applied to get the $\frac{E_{\text{rec}}}{E_{\text{gen}}}$ ratio close to one. The factor applied to the GEANT4 base Monte-Carlo was 697, while for HMC it was 783, a difference of about 11%.

These observations suggest that an additional efficiency factor is required for HMC: if about 15% of the photo-electrons would be lost, the GEANT3 and GEANT4 distributions for muons would agree. This factor would yield more consistent results for electrons as well. The implementation of such an efficiency factor has not been investigated and is an option for further improvement of the simulation of the calorimeter.

From the comparison mentioned above, it can also be concluded that the FINT and simplex method produce similar results. The FINT method is expected to be a little more precise, due to a more efficient extrapolation. The simplex method however is faster. It was decided to keep the implementation of both interpolation methods in HMC. The user can decide which method is more advantageous for a particular analysis.

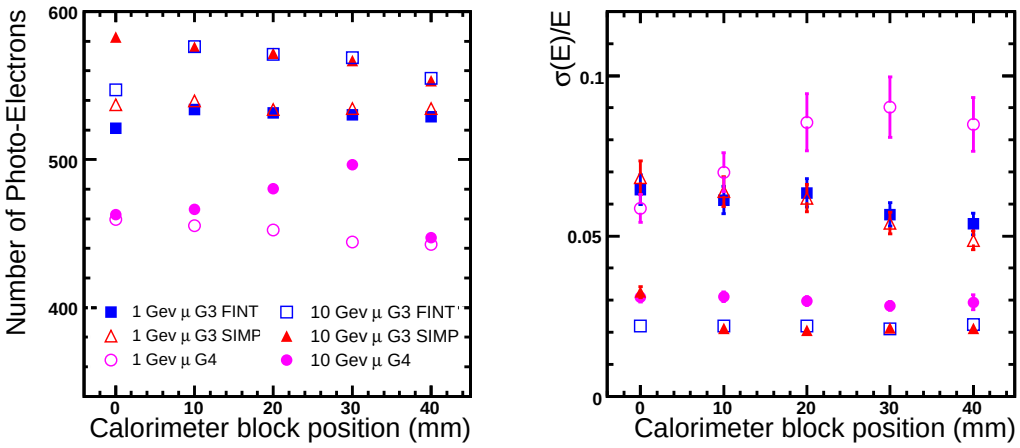

Figure 6.24: The left panel shows the most probable value (MPV) of the photo-electron yield for 1 and 10 GeV muons. The right panel displays the standard deviation of the distribution divided by the most probable value. Results obtained with GEANT4 (circles), and with GEANT3 in combination with the LUT and FINT interpolation (squares) and SIMP interpolation (triangles) respectively.

6.5 Comparison with experimental data

The real goal of any detector simulation is that it reproduces the detector's response. Unlike the previous comparisons, the Monte Carlo needs to be generated from the center of the target cell in order to be able to compare the Monte Carlo sample with the experimental data. Moreover, particles need to be generated over a specific momentum range, unlike the monochromatic beams used in the previous sections.

The only data we can comfortably compare with are positron or electron data. The tracking system in combination with the magnet offers the possibility to determine particle momenta, which can be compared to the energy deposition in the calorimeter. The energy over momentum ratio should be close to 1. Additionally, the incident position at the calorimeter surface can be predicted from tracking parameters and compared to the position calculated from the calorimeter response. These two observables will be subject of the first two sections of this paragraph. The last section will compare the reconstruction of photons, via the two-photon invariant mass spectrum and the resultant π^0 peak.

6.5.1 Energy over momentum for positrons

The ratio of energy determined by the calorimeter over the momentum measured by the tracking system, is not only used for particle identification (see section 3.3.2.2), it also serves the calibration by the calorimeter for all data productions. Since particle momenta and tracks are in general well reproduced by the HERMES Monte Carlo, the E/P ratio forms an excellent tool to verify the simulation of the energy reconstruction of the calorimeter.

As mentioned in section 3.3.2.2 a cubic correction is implemented in HRC to improve the linearity of the E/P ratio. This correction is given by equation 3.7 on page 66. In order to

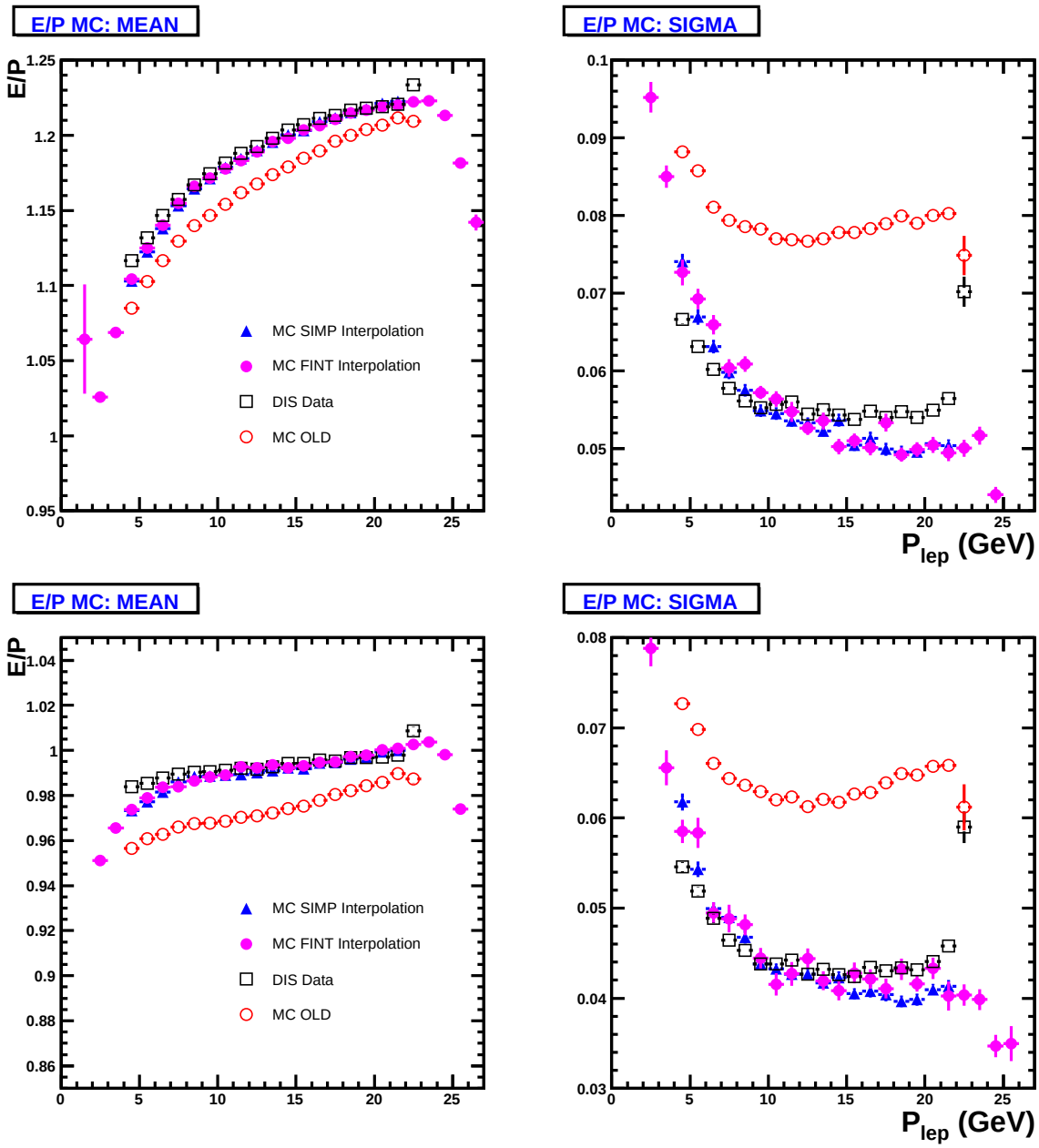

Figure 6.25: Comparison of the E/P ratio of experimental data with Monte Carlo. The upper panel shows the ratio without the cubic correction in HRC. The two plots on the left give the mean of the distributions, the two on the right show the standard deviation. The result of the old Monte Carlo is given for comparison.

test the Monte Carlo simulation, direct comparisons are preferably made. The complication of any other program should be avoided where possible. The comparison therefore should happen in the intermediate stage between HMC and HRC. Experimental data, however, are only available after HRC. In order to compare the E/P ratio without the cubic correction formula, the experimental data needs to be 'anti-cubic corrected'.

The comparison of the E/P ratio between data and Monte Carlo is seen in Figure 6.25.

The upper plots compare anti-cubic corrected data with the output of HMC. The lower plots compare experimental data as given in the micro-DSTs and where accordingly the cubic correction was applied to the Monte Carlo events. The data are binned in 1 GeV bins. A Gaussian was fitted to the data in every bin. The mean and width are depicted in the figure. A good agreement can be seen between the experimental data and the Monte Carlo for both interpolation techniques. For comparison, also the old Monte Carlo implementation of the calorimeter is given in the figure. The improvement in simulation is clearly visible.

The anti-cubic corrected data gives slightly better results. A maximal deviation of less than 2% between data and Monte Carlo for the mean of the E/P distribution can be seen in the lowest energy bins. The Monte Carlo also describes the width of the E/P distribution well. For the Monte Carlo sample only one particle per event was generated, whereas for data any number of particles could be present per event[6]. This could possibly deteriorate the resolution, and could explain the deviation for momenta above 15 GeV. The difference between experimental data and Monte Carlo can partly be attributed to a less than perfect simulation of the tracking devices and the magnetic field.

6.5.2 Position reconstruction

The position reconstruction performance of the calorimeter mentioned in section 4.1.2.4 could be used as an additional comparison between data and Monte Carlo. Assuming a perfect tracking system, a particle's back partial track, reconstructed by HRC, can be used to predict the particle's incident position at the calorimeter surface. This position is referred to as the 'generated' position in the following and is calculated as:

$$X_{\text{gen}} = \text{rcPartTrack.X} + 463 \cdot \text{rcPartTrack.XSlope}$$
$$Y_{\text{gen}} = \text{rcPartTrack.Y} + 463 \cdot \text{rcPartTrack.YSlope}$$

rcPartTrack.X(Y) is the position of the partial track at the center of the magnet, located at $z = 275$ cm, while X(Y)Slope is the track's slope. The formula above assumes the mean depth of energy deposition to be at $z = 738$ cm. The latter value is the standard value used in nearly all HERMES analyses and is the estimate of the mean position of energy deposition for leptons [187]. When comparing experimental data versus Monte Carlo the exact position of the mean energy deposition is not important. A potential bias by taking the wrong z-position should be reflected in both data and Monte Carlo. For these studies we only want that data and Monte Carlo agree.

The partial tracks are only available in the HRC production. Since the PID calculations are performed in the micro DST production, some hard cuts were required on the responses of the various particle identification detectors in order to select scattered electrons or positrons. A particle track was assigned an electron or positron track if the following condition was fulfilled:

$$\text{PulsTRD} > 29 \;\&\&\; \text{PulsPre} > 0.025 \;\&\&\; E/|P| > 0.92 \;\&\&\; E/|P| < 1.05.$$

No distinction was made between electrons and positrons.

Figure 6.26 shows the position reconstruction capabilities of the calorimeter for 2000 data in combination with the results of a Monte Carlo simulation. During the 2000 data taking period, positrons were scattered off the HERMES target.

[6]However, the 'old' Monte Carlo indicated in the figure was generated with PYTHIA .

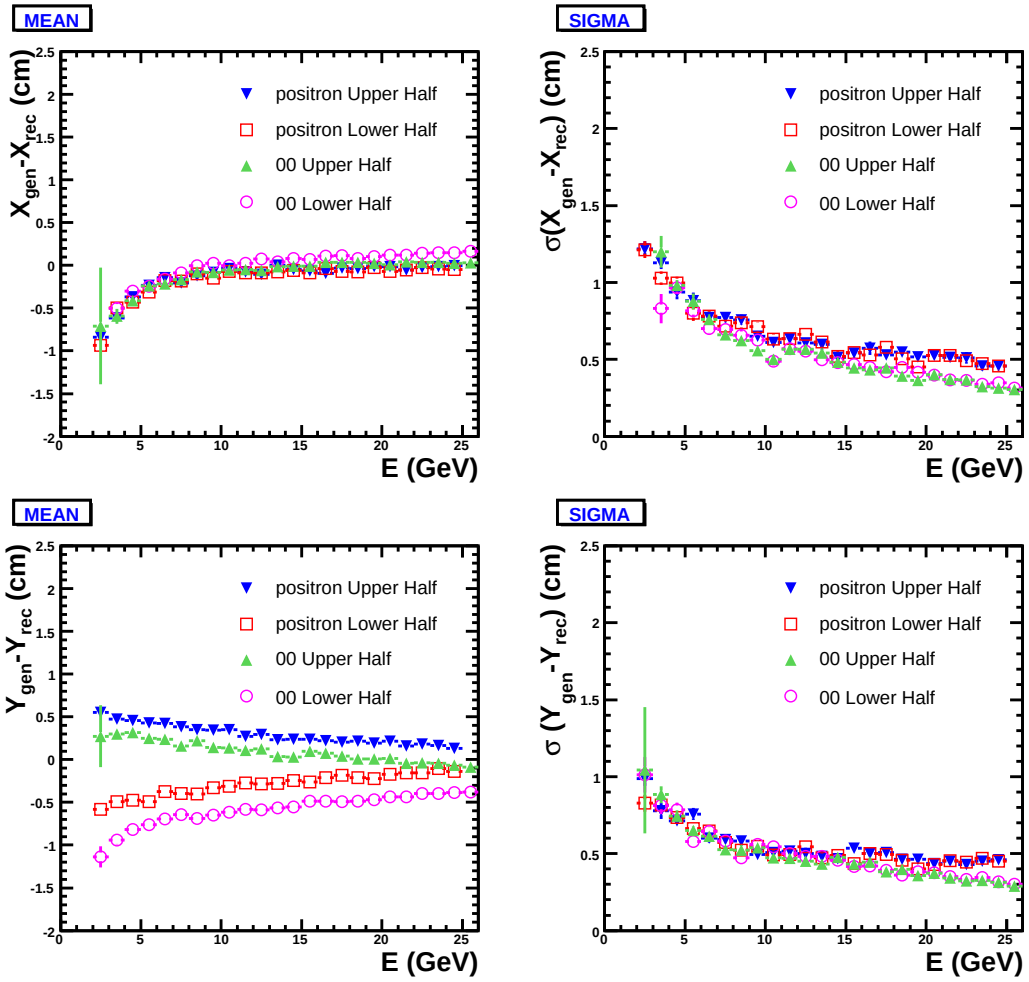

Figure 6.26: Comparison of the position reconstruction capabilities of the calorimeter between data and Monte Carlo. The various observables are indicated in the figure. Data are from 2000 (indicated as 00 in the legend), Monte Carlo was generated with positrons. The LUT was interpolated using the FINT technique. The left (right) plots show the mean (sigma) of the distribution.

For the Monte Carlo sample, positrons were generated between 2 and 25 GeV. The data and Monte Carlo samples were binned in 1 GeV bins, and a Gaussian was fitted to the $X_{\text{gen}} - X_{\text{rec}}$ and $Y_{\text{gen}} - Y_{\text{rec}}$ distributions. A good agreement for the x-position can be seen for both upper and lower calorimeter halves in terms of mean and standard deviation of the distribution. An offset between data and Monte Carlo can be seen in the lower left panel of the figure, corresponding to the reconstruction of the y position. While the Monte Carlo sample has the same positioning bias magnitude in the upper and lower calorimeter half, the bias of the experimental data in the lower half is larger than in the upper half, indicating a possible misalignment effect between both detector halves. The sign of the bias is inverted between both halves, since both Y_{rec} and Y_{gen} change sign when going to the other calorimeter half.

The variation of the standard deviation with energy can be explained from first principles: as the energy resolution is better for higher energies, also the position resolution should be

more precise for these energies. The resolution in x and y are of similar magnitude.

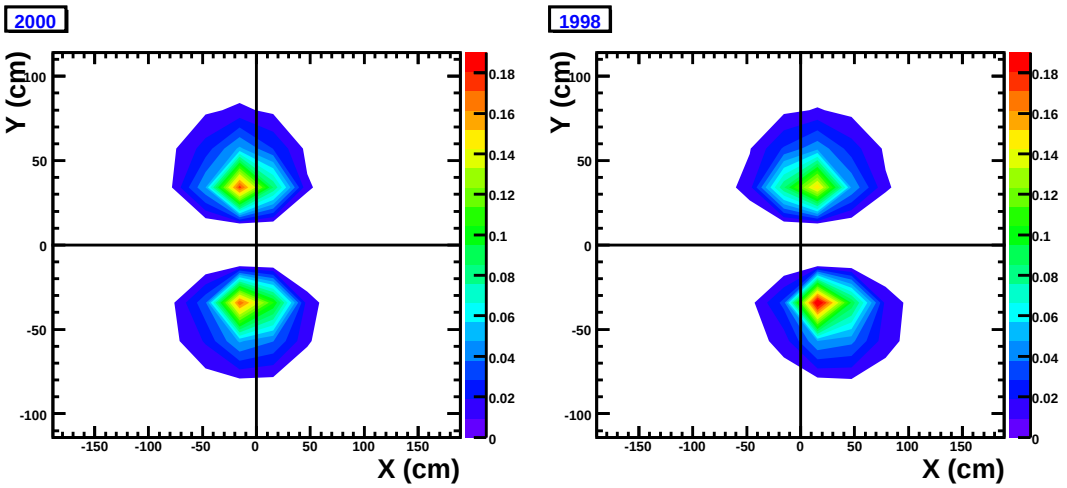

Figure 6.27: The left (right) panel shows the hit distributions for the 2000 (1998) across the calorimeter surface as an intensity scatter plot. Both samples are normalized to unity. In 2000 (1998) the beam consisted of positrons (electrons). The difference in X-position is due to the field induced by the spectrometer magnet. The difference in Y position is presumably due to a different beam slope at the HERMES IP between both data taking periods. Note that negative X values correspond to the right hand side when looking upstream.

The left panel of Figure 6.27 shows the hit distribution across the calorimeter surface for electrons and positrons for the 2000 sample. The favored calorimeter quadrant for positrons is the one where $x < 0$ and $y > 0$. A substantial amount of hits can be seen in the quadrant $x < 0$ and $y < 0$ as well. The right panel of the figure shows the hit distribution for the 1998 sample, where the primary beam particles were electrons. It can be seen that electrons are preferentially scattered in the $x > 0, y < 0$ quadrant with a substantial fraction in the $x > 0, y > 0$ quadrant. The difference in x position at the calorimeter surface is due to the spectrometer magnet. The difference in y position is presumably due to a change in beam slope.

Figure 6.28 shows the position reconstruction distributions for 1998 data in comparison with the results from a Monte Carlo simulation generated with electrons. Also here a good agreement can be seen between data and Monte Carlo for the mean and standard deviation of both X and Y distributions. Compared to Figure 6.26, the bias in X is inverted, while in Y it is similar in sign and in magnitude. The opposite bias with respect to Figure 6.26 is due to the fact that both X_{gen} and X_{rec} change sign when changing from electrons to positrons. An indication of misalignment between top and bottom detector halves can also here be seen from the reconstructed Y-position.

6.5.3 Two-photon invariant mass spectrum

The two-photon invariant mass spectrum, with the clearly visible π^0 peak forms an excellent tool for comparing data and Monte Carlo for photon induced showers. This observable has the advantage that no information from the tracking system is required, thus simplifying the

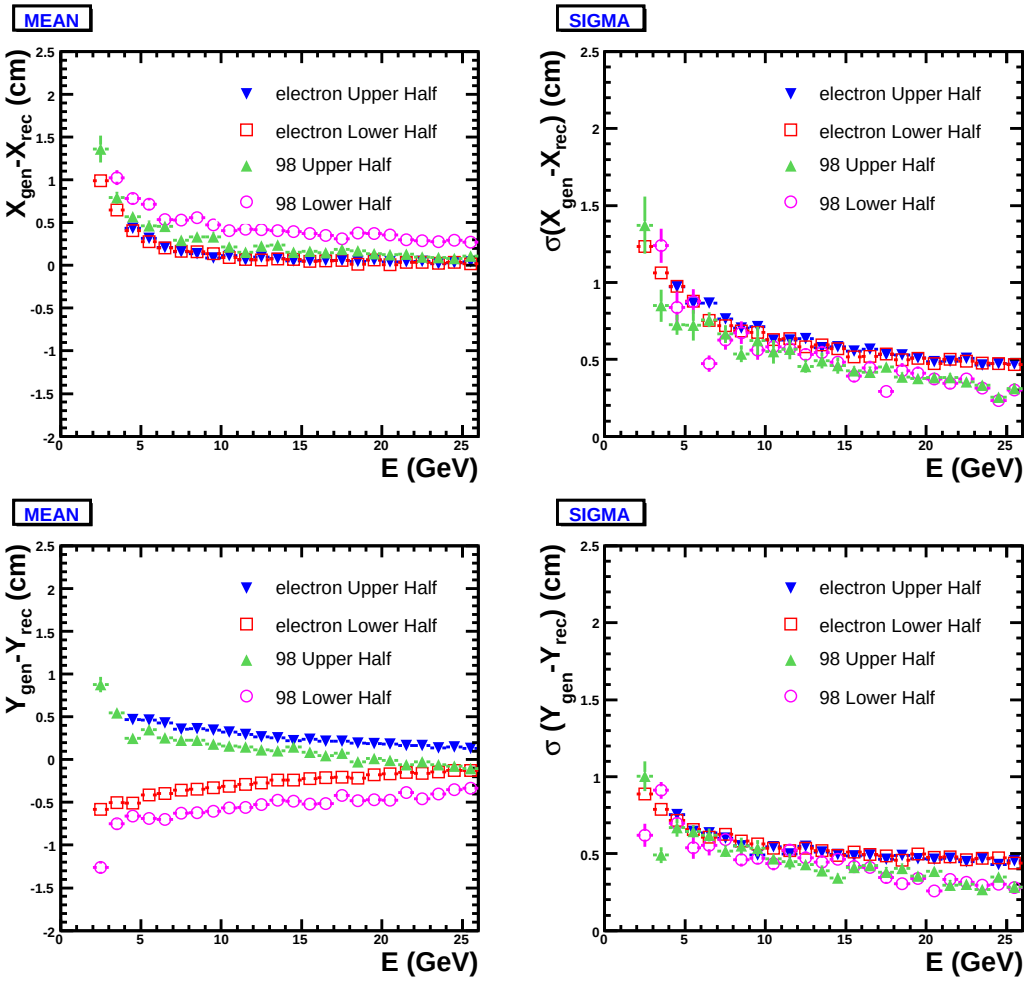

Figure 6.28: Comparison of the position reconstruction capabilities of the calorimeter between data and Monte-Carlo. The various observables are indicated in the figure. Data are from 1998 (indicated as 98 in the legend), Monte Carlo was generated with electrons, the LUT was interpolated using the FINT technique. The left (right) plots show the mean (sigma) of the distribution.

comparison. A drawback, however, is that the simultaneous response of two particles needs to be compared, whereas in the previous comparisons only one particle was used.

Figure 6.29 compares the two-photon invariant mass spectrum for various π^0 energies. Details about the construction of the invariant mass spectrum are given elsewhere in this thesis. In order to simulate the combinatorial background in the data sample, a physics generator was required, rather than the background generator which was used in all of the above comparisons. The Monte Carlo therefore was generated with PYTHIA, and tracked through HRC. A good agreement between experimental data and Monte Carlo can be seen over the entire energy range. A Gaussian was fitted to the central peak. Fit values are given in Table 6.4. For comparison also the fit values of a Gaussian fit to the two-photon invariant mass spectrum obtained with the old Monte Carlo are given. The fit values clearly indicate an improvement of the Monte Carlo description of the calorimeter.

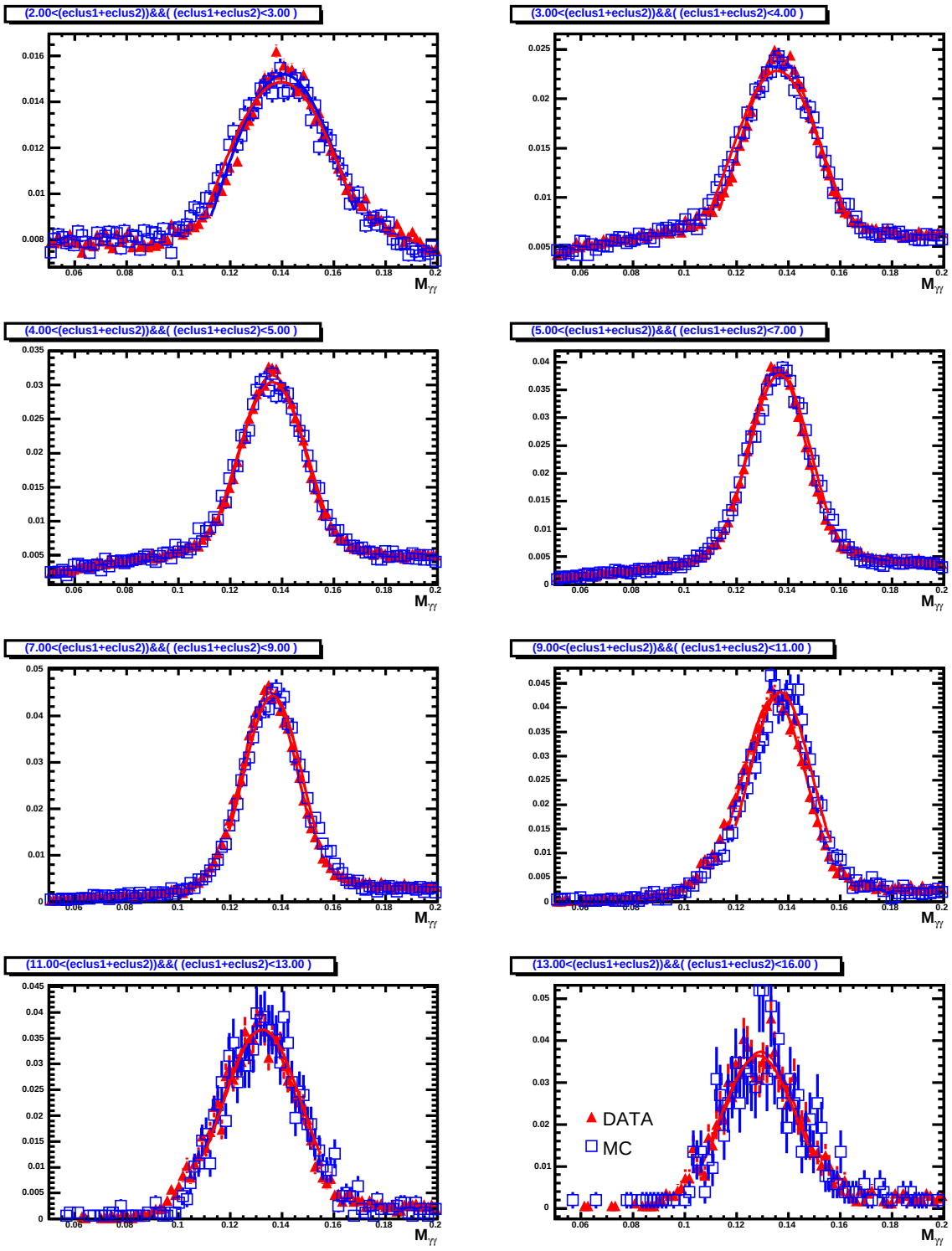

Figure 6.29: Two-photon invariant mass spectrum for various π^0 energies for data (triangles) and Monte Carlo (circles). A Gaussian was fitted to the central peak. Fit parameters are given in Table 6.4.

$E_{\gamma_1} + E_{\gamma_2}$ (GeV)	Mean Data (MeV)	Mean MC (MeV)	Mean Old (MeV)
$2-3$	139.6 ± 0.3	140.6 ± 0.2	138.0 ± 0.1
$3-4$	135.7 ± 0.2	136.7 ± 0.1	133.0 ± 0.1
$4-5$	136.3 ± 0.2	136.3 ± 0.1	133.4 ± 0.1
$5-7$	136.6 ± 0.1	135.8 ± 0.1	133.5 ± 0.1
$7-9$	136.4 ± 0.2	135.2 ± 0.1	133.9 ± 0.1
$9-11$	137.0 ± 0.3	134.5 ± 0.2	133.2 ± 0.1
$11-13$	132.4 ± 0.6	131.4 ± 0.3	128.6 ± 0.2
$13-16$	128.4 ± 1.2	129.4 ± 0.5	127.6 ± 0.5
$E_{\gamma_1} + E_{\gamma_2}$ (GeV)	Width Data (MeV)	Width MC (MeV)	Width Old (MeV)
$2-3$	29.65 ± 0.57	27.45 ± 0.31	32.38 ± 0.30
$3-4$	18.68 ± 0.24	16.44 ± 0.13	19.51 ± 0.11
$4-5$	14.48 ± 0.20	13.51 ± 0.12	15.74 ± 0.10
$5-7$	12.71 ± 0.17	11.88 ± 0.10	13.68 ± 0.07
$7-9$	11.52 ± 0.22	11.09 ± 0.13	12.68 ± 0.11
$9-11$	12.19 ± 0.34	12.49 ± 0.22	14.06 ± 0.16
$11-13$	15.30 ± 0.76	15.00 ± 0.44	17.35 ± 0.28
$13-16$	14.48 ± 1.94	14.33 ± 0.57	17.95 ± 0.59

Table 6.4: Mean and width of a Gaussian fit to the two-photon invariant mass spectrum depicted in Figure 6.29. For comparison also the fit values of for the old Monte Carlo are given. The latter is not drawn in Figure 6.29. A good agreement between data and new Monte Carlo can be seen over the entire interval.

The minimal difference between the width of the experimental data and the new Monte Carlo distributions could probably be improved by increasing the simulated electronic noise in the digitization routine of the Monte Carlo, or by including an additional efficiency in the simulation according to the discussion of section 6.4.4.

6.6 Performance of the HERMES calorimeter

The previous comparisons showed a good agreement between experimental data and Monte Carlo, indicating a solid and functioning description of the calorimeter within HMC. A good functioning calorimeter simulation offers the possibility of testing the performance of the HERMES electromagnetic calorimeter in terms of energy measurement and position reconstruction for photons. The performance is estimated from comparing generated and reconstructed distributions. This section covers the energy measurement as well as the position reconstruction performance of the HERMES calorimeter. The reported results indicate the behavior of the combination of the calorimeter and HRC, the reconstructing program. The latter is responsible for clustering and for position reconstruction, as already mentioned in section 4.1.2.4.

6.6.1 Energy measurement for photons

When analyzing the energy measurement for photons, it is mandatory to distinguish between cases when the photon-initiated electromagnetic shower starts in the preshower, and the case when the photon does not interact in the preshower. In the latter case, the penetration depth into the lead-glass will be deeper due to the larger mean free path for photons. The two cases can be distinguished depending on the energy deposition in the preshower.

Figure 6.30 shows the ratio of generated and reconstructed energy as a function of reconstructed energy. A distinction was made on the energy deposition in the preshower. The data were binned in 1 GeV bins, and a Gaussian was fit to the spectra. It can be observed that in the case of photons the reconstructed energy $E_{\rm rec}$ is larger than the generated energy, especially in the case where the photon does not interact in the preshower. This can be explained qualitatively: if a photon does not shower in the preshower, it penetrates deeper into the lead-glass. The attenuation of the Čerenkov photons will accordingly be less than in the case of electron initiated showers, since Čerenkov photons are generated closer to the PMT. The calorimeter, however, is calibrated via the E/P ratio of electrons. The various effects to the Čerenkov photons are accounted for by the calibration procedure. When applying these effects to photon induced showers, the attenuation of the Čerenkov photons is overestimated resulting in $\frac{E_{\rm gen}}{E_{\rm rec}} < 1$. When the photon interacts in the preshower, we have a similar although smaller effect. Also here the showering starts later than in the case of electrons. The shower, however, develops earlier than in the case of photons not interacting in the preshower.

In the same figure, a decrease of the ratio can be observed for energies above about 20 GeV. This corresponds to the regime where the reconstructed energy distribution becomes broader, due to the effect of the cubic correction applied in HRC (see later).

Within the HERMES collaboration, a standard way of correcting for the deeper penetration of photon-induced showers, is to take 97% of the measured energy, i.e. $E_\gamma = 0.97 \cdot E_{\rm rec}$. The dashed line in the figure shows that this value is a reasonable approximation in the case that the electromagnetic shower started in the preshower. The correction, however, is not optimal, and completely incorrect for photons not showering in the preshower. In an attempt to quantify the difference between generated and reconstructed distributions, the following function was fit to the spectrum:

$$\frac{E_{\rm gen}}{E_{\rm rec}} = p_0 \cdot (E_{\rm rec})^{p_1} - p_2 \cdot (E_{\rm rec})^{p_3}\,, \tag{6.9}$$

with $p_0 \cdots p_3$ fit parameters. One power-law describes the rise at lower energies, while the other power-law fits the decrease at higher energies. Fit values are indicated in the figure.

Obviously, the width of the $\frac{E_{\rm gen}}{E_{\rm rec}}$ distribution (indicated in Figure 6.30) is smaller when the photon does not interact in the preshower, since all energy is deposited in the calorimeter in this case. For a photon induced shower starting in the preshower the resolution was determined by a fit of a function of the form $p_1 + \frac{p_2}{\sqrt{E_{\rm rec}}}$. The fit gave following result:

$$\sigma(\frac{E_{\rm gen}}{E_{\rm rec}}) = (1.4 \pm 0.0) + \frac{(8.4 \pm 0.1)}{\sqrt{E_{\rm rec}}}\ \%. \tag{6.10}$$

The behavior of $E_{\rm gen}/E_{\rm rec}$ as indicated in Figure 6.30 is a little odd; a decrease of the ratio as a function of energy is expected since at higher energies, the shower penetrates deeper into the lead glass. Instead, an opposite behavior is seen. The observed behavior is expected to

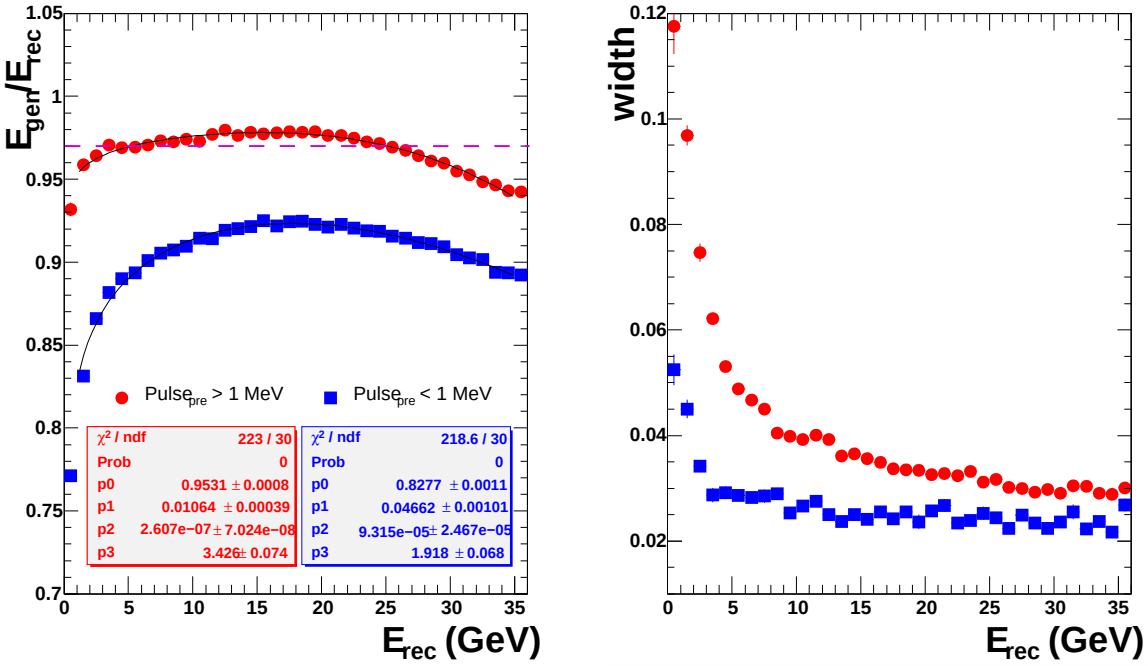

Figure 6.30: The left panel shows the mean of the $E_{\rm gen}/E_{\rm rec}$ ratio for photons and the right panel its width. Circles (squares) show the ratio when (no) energy is deposited in the preshower. The dashed line in the left panel represents the standard HERMES correction of 0.97 to the reconstructed energy.

be a consequence of the cubic correction in HRC. To test this hypothesis, the reconstructed energies were anti-cubic corrected. The $E_{\rm gen}/E_{\rm rec}$ ratio with this additional correction is plotted in Figure 6.31. The variation of the ratio as function of $E_{\rm rec}$ behaves as expected. Also here a distinction was made depending on the energy deposition in the preshower. The spectrum in Figure 6.31 was attempted to be parametrized by a function of the form:

$$\frac{E_{\rm gen}}{E_{\rm rec}} = p_0 + \frac{p_1}{\sqrt{E_{\rm rec}}}. \tag{6.11}$$

Fit parameters are indicated in the figure. The width of the distribution is similar to the one presented in Figure 6.30. The resolution in the case of energy deposition in the preshower was determined as:

$$\sigma(\frac{E_{\rm gen}}{E_{\rm rec}}) = (1.0 \pm 0.0) + \frac{(7.8 \pm 0.1)}{\sqrt{E_{\rm rec}}}\ \%, \tag{6.12}$$

which indicates a slightly better resolution than in the case presented by equation 6.10.

The $E_{\rm gen}/E_{\rm rec}$ ratio in Figure 6.31 does not show a fall off for $E_{\rm rec} > 20$ GeV as seen in Figure 6.30. This fall off is a consequence of the cubic correction. The effect is detailed in Figure 6.32, where the generated and reconstructed energy distributions for both methods can be seen. The generated energy follows a flat distribution and is indicated by the dashed histograms. The left (right) panel shows the anti-cubic corrected (standard) Monte Carlo. Due to the cubic correction, the reconstructed energy distributions become broader. Saturation, moreover, occurs at lower energies if the cubic correction is applied.

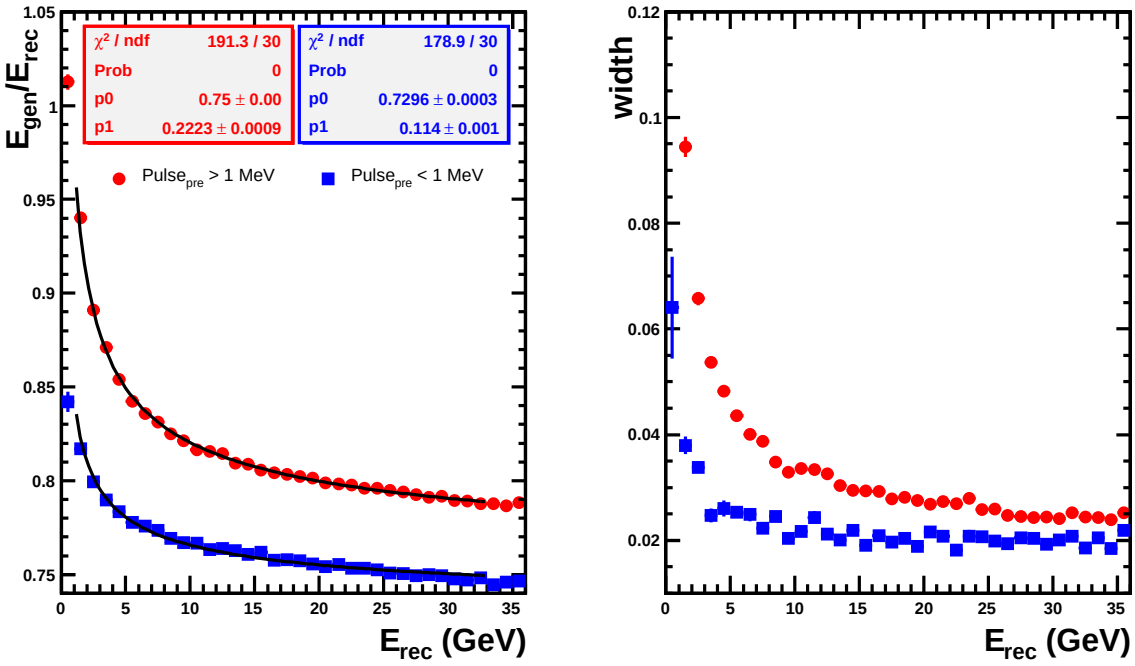

Figure 6.31: The left panel shows the mean of the $E_{\text{gen}}/E_{\text{rec}}$ ratio for photons and the right panel the width of that distribution for anti-cubic corrected Monte Carlo events. Circles (squares) show the ratio when (no) energy is deposited in the preshower.

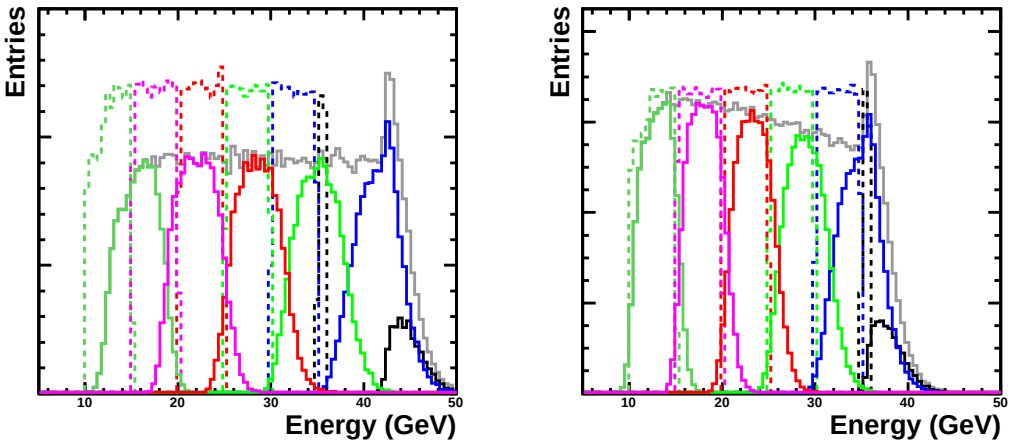

Figure 6.32: Comparison of the generated energy (dashed histograms) and the reconstructed energy (full histogram) by the calorimeter. The left panel shows the 'anti-cubic' corrected data. The right panel the standard HRC output. The data are ordered in various bins indicated by the colors. The gray histogram shows the histogram for the entire reconstructed data set. The effect of the cubic correction is a broadening of the reconstructed energy. Note the slope in the gray histogram in the right panel.

6.6.2 Position measurement for photons

Apart from the energy measurement, also position reconstruction is a mandatory task of the calorimeter for the reconstruction of photon tracks. The position reconstruction performance can be verified by comparing generated Monte Carlo positions with reconstructed ones. When using the background generator of HMC, particles are generated according to a flat distribution in three variables: the momentum magnitude p, and its angular components θ_x and θ_y. Since photons are not deflected by the magnetic field, the generated Monte Carlo can easily be used to calculate the position of the photon track in the calorimeter:

$$x_{\text{gen}} = (v_z + \alpha) \cdot \cos\varphi \tan\theta$$
$$y_{\text{gen}} = (v_z + \alpha) \cdot \sin\varphi \tan\theta$$

φ and θ are already converted from θ_x and θ_y internally in HMC. α is a parameter describing the energy loss depth in the calorimeter, and v_z the z vertex of the event. Various parametrization for α were assumed:

1. $\alpha = 720$, as an initial check; the value corresponds to the position of the calorimeter surface.
2. $\alpha = 747.5$, since this is the recommended value for all HERMES analyses. The factor is assumed to describe the mean distance of energy deposition for photons, and is taken from the HERMES analysis web site [187].
3. α according to a parametrization of the mean depth of energy deposition $E_{\text{dep}}^{\text{mean}}$, as determined from Figure 6.20
4. α according to a parametrization from Reference [184], namely equation 6.8.
5. α according to a fit to the median depth of energy deposition $E_{\text{dep}}^{\text{med}}$, determined from Figure 6.20.

Depending on the energy deposition in the preshower the energy dependence for either leptons or photons was taken for assumptions $\alpha = 3, 4, 5$.

Figure 6.33 shows the mean and the width of the generated minus reconstructed positions as a function of energy for photons in the calorimeter quadrant corresponding to $y > 0$ and $x < 0$, i.e. the upper right corner when looking downstream. For every energy bin a Gaussian was fitted to the distribution in that bin. Depending on α, a bias is introduced in the reconstructed positions. The best results are achieved for the case when α is parametrized according to the fit to E_{dep}^{mean}. This is not too surprising since this was a fit to Monte Carlo generated data itself. Also the parametrization according to E_{dep}^{med}, or according to Reference [184] introduce little bias in the sample. The main message from the figure is that the recommended value of $z = 747.5$ is wrong.

From Figure 6.33 can also be concluded that the position resolution is worse in y than in x. This could be a consequence of the position reconstruction algorithm described in section 4.1.2.4. As mentioned before, the data plotted in Figure 6.33 are for generated position in the calorimeter quarter corresponding to $x < 0$ and $y > 0$. Figure 6.34 shows the position reconstruction performance for the quarter corresponding to $x > 0$ and $y < 0$. As expected, exactly the opposite bias can be seen for the x and y distributions, indicating the stability of the method.

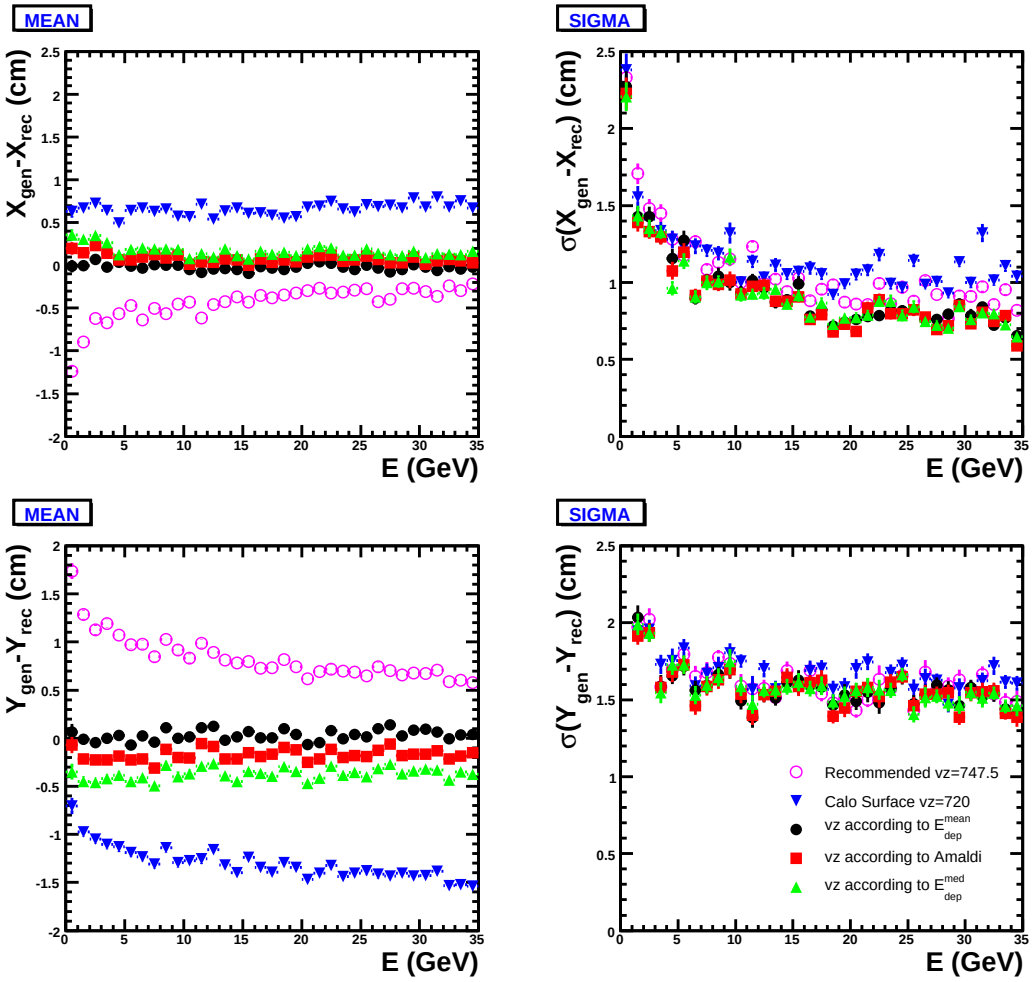

Figure 6.33: The upper (lower) half shows the mean of the generated minus reconstructed position distributions for x (y) positions. Left is the mean of the distribution, right its standard deviation. The plots represent the region where $x < 0$ and $y > 0$. The generated values were calculated for various longitudinal energy depositions as described in the text and in the legend. The (x, y) positions correspond to the upper right quarter of the calorimeter.

In conclusion, this chapter described the implementation of an improved Monte Carlo simulation for the preshower and calorimeter of the HERMES spectrometer. The new Monte Carlo is a combination of GEANT3 to simulate the electromagnetic shower and a Look-Up Table containing the photo-electron yield per mm track length. The various attenuation and refraction effects the Čerenkov photons are subject to are accounted for in the Look-Up Table. A good agreement with experimental electron and positron data was observed. The last section showed how the new Monte Carlo can be used to relate the signal in the calorimeter to the incident photon energy. The calorimeter resolution, e.g. equation 6.10 can be used as an input for HSG in order to obtain a fast calorimeter simulation based on smearing. Care, however, must be taken regarding the cubic correction in HRC.

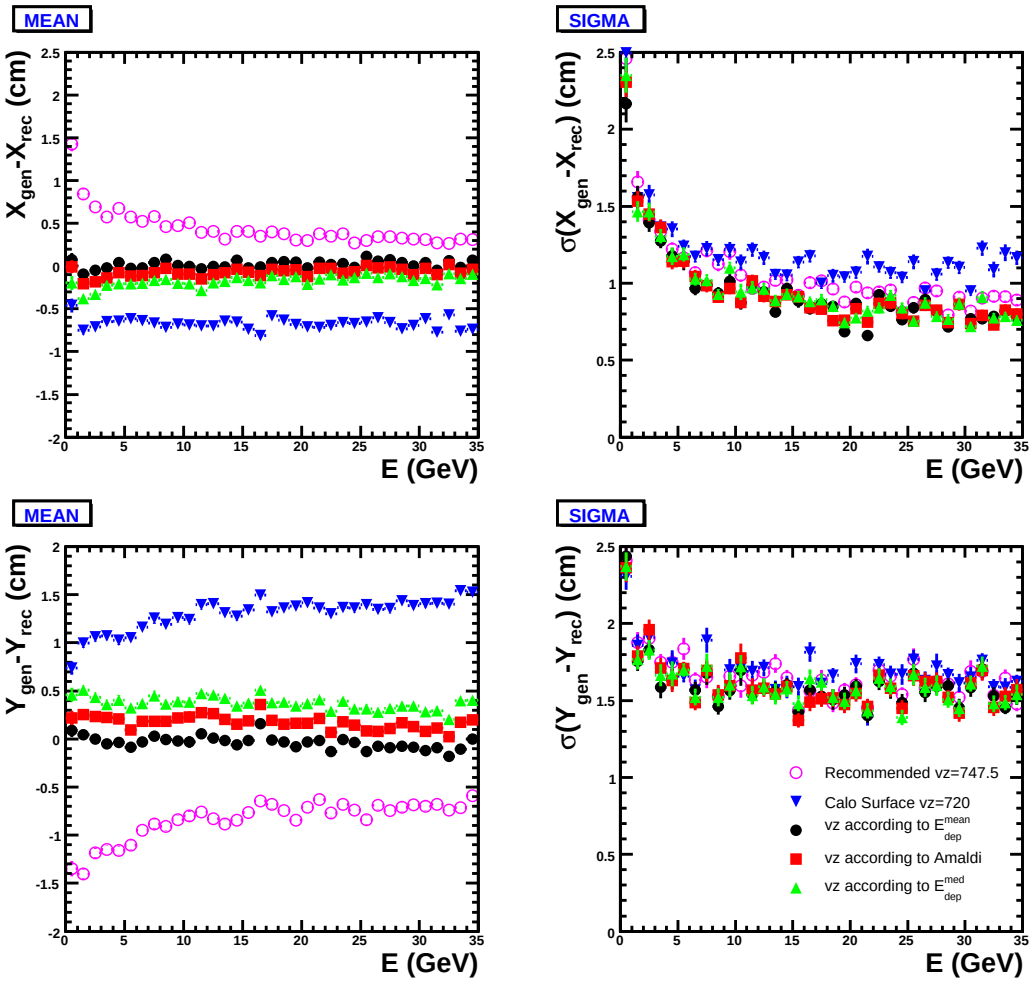

Figure 6.34: The upper (lower) half shows the mean of the generated minus reconstructed position distributions for x (y) positions. Left is the mean of the distribution, right its standard deviation. The plots represent the region where $x > 0$ and $y < 0$. The generated values were calculated for various longitudinal energy depositions as described in the text and in the legend. The (x, y) positions correspond to the lower left quarter of the calorimeter.

CHAPTER 7

Analysis of Exclusive π^0 production

This chapter describes the analysis of a potential exclusive π^0 signal in the HERMES data. First the various data selection criteria will be mentioned, followed by a discussion of the π^0 invariant mass reconstruction. Next, an attempt to improve the calorimeter cluster energy reconstruction will be given. A discussion of the sensitivity of the HERMES spectrometer with respect to exclusive π^0 signals based on the results of an exclusive π^0 event generator follows. Later, an attempt will be made to correct the exclusive π^0 event sample for the non-exclusive background. Finally, the predicted neutron suppression for exclusive π^0 and η pseudoscalar mesons is verified by analyzing the multiplicity ratio of these particles on hydrogen and deuterium targets.

7.1 Data selection

A HERMES data production of a particular year includes all data taken during that period. Accordingly, also runs where the experimental conditions were not optimal are stored to disk. Therefore it is important to examine the quality of the data. A year's production contains data taken on various gas targets and with all triggers. To select an appropriate subsample, various event selection cuts are additionally needed. This section covers the data quality cuts, the selection of a DIS subsample, and a discussion of the obtained luminosities.

7.1.1 Selecting good bursts by data quality constraints

Since hardware information is read out on the burst level[1], all data quality criteria should be applied at the burst level. The first step in the analysis code is to select valid bursts out of the several gigabytes of data of a particular production. For this purpose the burst-list from the

[1]for a definition of the quantity 'burst' see sections 3.6 and 4.4

data quality group was used. More information about the burst-list format in general can be found e.g. in Reference [113]. The burst-list is a bit pattern defined by the data quality group. A user can select particular data quality cuts by comparing the bit pattern from the burst-list against a user defined mask. The particular bits used to select valid data are discussed in the following. Similar selections were applied to the top and bottom detector halves. The applied data quality selections are:

Dead time: The dead time τ is calculated as the ratio of generated triggers over accepted triggers. The dead time obviously is always smaller than one. Large dead times correspond to large backgrounds, therefore these data sets should be rejected. Typically, two dead time ratio's are calculated: one overall ratio for all triggers (`rDeadCorr`) and one for the DIS trigger 21 (`rDeadCorr21`). An additional quality factor `rDeadCorrArt` indicates whether there are additional corrections needed for the ratio. When this value is 1, no additional corrections are required.

Reasonable deadtime	$0.5 \leq$ `g1DAQ.rDeadCorr` ≤ 1	**Bit 2**
	$0.5 \leq$ `g1DAQ.rDeadCorr21` ≤ 1	
	`g1DAQ.rDeadCorrArt == 1.0`	**Bit 30**

Burst length: As the burst is the time unit between two consecutive scalar events, the burst length should not be too long in order to have trustworthy scalar data.

Reasonable burst length	0.0 s $\leq$ `g1DAQ.rLength` ≤ 11 s	**Bit 3**

Beam current and luminosity: The luminosity is determined offline by the luminosity group from a fit to the ratio of the luminosity monitor counts to the beam current, the latter being measured by a dedicated device. The beam current is provided by the HERA Machine Group. For unpolarized runs the selection on the luminosity rate should be discarded, since most of the unpolarized data is taken during high density running at the end of a fill. During these high density runs, high luminosity count rates are expected, due to the high target density. For these data samples high luminosity count rates are not an indication of untrustable experimental conditions.

Reasonable beam current	5 mA $\leq$ `g1Beam.rMdmCurr` ≤ 50 mA	**Bit 4**
Reasonable luminosity	5 Hz $\leq$ `g1Beam.rLumiFitBstGai` ≤ 80 Hz	**Bit 5**

Bad bursts: Due to e.g. clock synchronization problems, some particular bursts need to be discarded. The first burst in a run is always omitted, since the initial phase of a run includes various initializations of detectors. Another example of bad bursts are those during which the PID detectors had problems or those which contain bad records.

first burst in a run	bitmask `g1Quality.iuDSTbad`	**Bit 6**
bad uDST records	bitmask `g1Quality.iuDSTbad`	**Bit 7**
no PID available	bitmask `g1Quality.iuDSTbad2`	**Bit 8**

Dead blocks in H2, Lumi or Calo; bad TRD: During some runs or bursts it happened that the calorimeter, H2 or the luminosity monitor contained some bad blocks. The determination of bad blocks is done via the GMS system. This system however is not completely error-free so special care needs to be taken when making selections based on dead blocks. For charged particle analysis the calorimeter is allowed to have one bad block. Therefore a dedicated bit 31 was introduced. In the presented analysis all calorimeter blocks are required to be working.

bad calorimeter blocks	bitmask `g1Quality.bCaloDead`	**Bit 17**
bad lumi or H2	bitmask `g1Quality.bH2LumiDead`	**Bit 18**
bad TRD	bitmask `g1Quality.iTrdDQ`	**Bit 19**

High Voltage trips in FC or BC: The controlling software of the High Voltage system records eventual trips in a special table. All entries of the table need to be 'looped' over. This is done in dedicated code provided by the data quality group.

HV trips in FC/BC	selection `g1HVtrip.iDetID` and `g1HVtrip.iChannel`	**Bit 20**

Hard cuts on run numbers: During some particular runs certain spectrometer devices were not working properly, therefore hard cuts on the run numbers are needed. The list of those runs is provided by the data quality group.

particular bad runs	cut on `g1DAQ.iRun`	**Bit 21**

Target state: Depending on polarized or unpolarized running different bits need to be selected. Unpolarized data is simply selected by the `g1Unpol` table, while for polarized data additional requirements for ABS, BRP and TGA are needed. Moreover, the target polarization needs to be in a well defined state. For the analysis presented in this thesis, only unpolarized data was needed, therefore no requirements were made on the target state, even for those runs where the target was polarized.

Polarized data	`g1Unpol.iGasType` $>$ `0`	**Bit 15**

Bit pattern: The resulting bit mask slightly varies by data production. Bits 17 and 18 are e.g. empty for the 04b0 production. The applied bit masks for every data production are given in Table 7.1. Details on the respective formats for every data production can be found from Reference [188]. The data quality bit mask used in this thesis discards the online logbook entries by the shift crew. Moreover, the beam polarization criteria were omitted since in the analysis presented here only unpolarized data is needed. The RICH data quality constraints were discarded as well, because the RICH only identifies charged hadrons. Here only neutral hadrons are considered.

7.1.2 Selecting DIS events

Having selected valid bursts, the search for DIS events within a data sample is made. The selection of DIS events is done in two steps: first an electron or positron track needs to be found in the HERMES acceptance, and secondly the kinematic variables Q^2, W^2 and y are calculated and constrained to match DIS kinematics.

7.1.2.1 Fiducial Volume Cuts

The first condition applying to every track in a particular event regards the tracking method. For 1997 data the method without Vertex Chambers (NOVC) is required. Next, it is checked that a track should not be a magnet track by requiring `(g1Track.iSelect & 0x0300) != 0`. A magnet track (or 'short' track) is a track for which only the FCs and the MCs fired, see e.g. Figure 3.15 on page 61. Subsequently, various constraints are imposed on the track parameters.

To calculate track locations at specific positions in the HERMES spectrometer, information from the `smTrack` table is needed. This leads to an additional constraint, namely the demand that every track in `g1Track` corresponds to exactly one track in the `smTrack` table. The various conditions are explained below:

Vertex cut: The z-vertex of a track should be constrained to avoid scattering events from calorimeter C2 which enter the data sample. Moreover, a cut on the transverse vertex position is applied. However, for polarized data taken during 2004 and 2005, the cut on the transverse vertex should be omitted due to the influence of the transverse magnetic field on the track.

Calorimeter position cut: Projected track positions at the outer rows of the calorimeter need to be excluded as at these positions the clustering of 9 calorimeter blocks can not be performed, since less than 8 calorimeter blocks surround the incident track position. Hard cuts on the x and y positions at the calorimeter accordingly need to be considered. The position at the calorimeter is calculated from the back partial track parameters.

Field Clamp Positions: As can be seen from Figure 3.11 there are field clamps in front and behind the magnet, requiring additional cuts to the track parameters. Particles hitting these field clamps could cause particle showers. These showers obviously need to be avoided in the data sample.

Septum Plate: From Figure 3.11 also the position of the septum plate in the middle of the magnet can be seen. The presence of this plate requires additional fiducial cuts.

Following the discussion above, a track was considered to be within the HERMES acceptance when following conditions were fulfilled:

vertex z position:	$\lvert z_{\text{vert}}\rvert \leq 18\text{cm}$
transverse vertex position:	$d_{\text{vert}} = \sqrt{x_{\text{vert}}^2 + y_{\text{vert}}^2} \leq 0.75\text{cm}$
calorimeter x position:	$\lvert$ `smTrack.rXpos` + 445· `smTrack.rXslope`$\rvert \leq 175$
calorimeter y position:	$30 \leq \lvert$`smTrack.rYpos` + 445· `smTrack.rYslope`$\rvert \leq 108$
front field clamp:	$\lvert$ `smTrack.rxOff` + $172 \cdot \tan\theta_x\rvert < 31$
septum plate:	$\lvert$ `smTrack.ryOff` + $181 \cdot \tan\theta_y\rvert > 7$
rear field clamp:	$\lvert$ `smTrack.ryOff` + $383 \cdot \tan\theta_y\rvert < 54$
rear field clamp x position:	$\lvert$ `smTrack.rXpos` + 108· `smTrack.rXslope` $\rvert \geq 100$
rear field clamp y position:	$\lvert$ `smTrack.rYpos` + 108· `smTrack.rYslope` $\rvert \geq 54$

`rX(Y)pos` and `rX(Y)slope` are defined for the back partial tracks which are originating at the center of the magnet at $z = 275$ cm. The `rx(y)Off` positions are related to the front partial track and are originating from the center of the target cell. It should be noted that due to the transverse magnetic field during polarized hydrogen running in 2004 and 2005, corrections should be applied to ϕ, θ and z_{vert} of each track in those productions. Additionally the cut on d_{vert} should be discarded for these data productions. More information about these corrections can be found e.g. in Reference [189].

7.1.2.2 Kinematic Cuts

A particle track is assumed to be an electron or positron track if `PID3 + PID5 - Φ` > 1 (see section 4.3). For every event the electron (positron) track with the highest energy is granted to be the scattered beam electron (positron). An event is assumed to be a DIS event if trigger 21 fired and the associated Q^2 of the virtual photon is larger than 1 GeV^2. The resonance region is excluded by a cut on W^2. Furthermore a cut on the scaling variable y is applied, eliminating a regime where high QED radiative corrections are needed. The lowest kinematically allowed value of y is around 0.17 after applying the aforementioned constraints on W^2 and Q^2.

$$Q^2 > 1 \text{ GeV}^2$$
$$W^2 > 10 \text{ GeV}^2$$
$$y < 0.85$$

If an event track passes all fiducial and kinematic constraints, the event is presumed to be a DIS event. The resulting kinematic plane is depicted in Figure 7.1. From the figure it can also be seen that the maximal polar angle θ equals 220 mrad, as a consequence of the various fiducial volume cuts.

7.1.3 Calculating luminosities

If one wants to extract a cross section σ from an observed number of counts, it is mandatory to know the luminosity $\mathcal{L}$ delivered to the experiment. The latter is related to the observed number of counts by the relation: $\sigma = N/\mathcal{L}$.

In HERMES two alternative methods exist to calculate the total luminosity seen by the experiment: one takes into account the count rate in the luminosity monitor, and the other

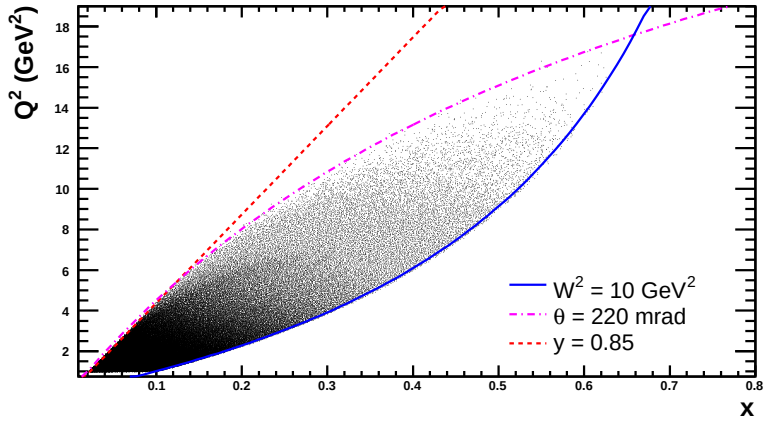

Figure 7.1: The (x, Q^2) kinematic distribution of DIS events in the HERMES acceptance. The dashed line corresponds to $y = 0.85$, the full line to $W^2 = 10$ GeV^2, and the dash-dotted line to $\theta = 220$ mrad.

one counts the number of DIS events. This offers the advantage that both calculations can be independently cross-checked. The difference between both methods forms an indication of the systematic error.

The method using the luminosity monitor is explained first. The count rate of this detector, $N_{\rm lumi}$, needs to be integrated over the burst length, $t_{\rm burst}$, and corrected for the dead time, $\tau^{\rm dead}_{\rm burst}$, in order to obtain the luminosity per burst[2]. Additionally, the luminosity monitor acceptance and efficiency need to be taken into account via a calibration factor $C^y_{\rm lumi}$. The constant is different from year to year and is available e.g. from References [190, 191]. For the years 2000 and 1999 the factor is equal to 0.417 ± 0.030 μb^{-1}. An additional constant, $C_{\rm targ}$, is required to take the number of electrons relative to the number of nucleons into account. This factor is equal to 1 for hydrogen targets, and 2 for deuterium targets. The final expression for the total integrated luminosity of a particular year y is :

$$\mathcal{L}_{\rm lumi} = C^y_{\rm lumi} \cdot C_{\rm targ} \cdot \Big(\sum_{\rm bursts} N_{\rm lumi} \cdot t_{\rm burst} \cdot \tau^{\rm dead}_{\rm burst} \Big). \tag{7.1}$$

The second method to determine the luminosity is based on the number of DIS events. This number reflects the luminosity indirectly. To calculate the associated luminosity, the numbers of electrons (or positrons) passing all above mentioned DIS cuts is counted, depending on the primary beam particles being electrons or positrons. A background correction for charge symmetric processes such as pair production is done by subtracting the number of positrons (electrons) from the aforementioned counts. The number of DIS events further needs to be scaled by the detection efficiency of DIS events. This efficiency, $\epsilon_{\rm DIS}$, is estimated from a fully tracked inclusive Monte Carlo. For the 1997 (1999) geometries the DIS efficiency is calculated to be 0.86 (0.81) [192]. Finally, to get a luminosity out of the number of DIS events, the number needs to be scaled by the DIS cross section on hydrogen targets, $\sigma_{\rm DIS}$, which can

[2]Another method would be to correct the observed event counts in the spectrometer for the dead time, and not the counts of the luminosity monitor.

be evaluated from world data on the F_2 structure function. For the kinematic range covered by HERMES, a value of $\sigma_{\rm DIS} = 60.9$ nb was calculated [192]. The luminosity based on the number of DIS events can thus be written as:

$$\mathcal{L}_{\rm DIS} = \frac{N^{e^+} - N^{e^-}}{\epsilon_{\rm DIS} \cdot \sigma_{\rm DIS}}. \tag{7.2}$$

The disadvantage of the latter method is that the efficiency needs to be recalculated from an inclusive Monte Carlo for every change in the detector setup or every change in the fiducial volume cuts. Moreover, a change in kinematic constraints to the DIS sample requires a recalculation of the inclusive cross section. A disadvantage of the method using the luminosity monitor count rate is the uncertainty involved in the luminosity calibration factor $C_{\rm lumi}$.

The number of DIS events for the years 1996, 1997, 2000, 2004 and 2005[3] are given in Table 7.1 for both hydrogen and deuterium data samples together with the values of the luminosity constant, $C_{\rm lumi}$. When calculating the ratio of number of DIS events on hydrogen over deuterium, the value 1.0371 ± 0.0004 is obtained, assuming Poisson statistics. To get the luminosity ratio between hydrogen and deuterium, the event ratio needs to be multiplied by $\sigma^D_{\rm DIS}/\sigma^H_{\rm DIS}$, as can be seen from equation 7.2. The cross section ratio is about 0.90 ± 0.00 [193] at average HERMES kinematics of $< x > \sim 0.11$ and $< Q^2 > \sim 2.5$ GeV^2. The luminosity ratio based on the number of DIS events accordingly is 0.93 ± 0.00. The same ratio determined by the luminosity monitor is 0.94 ± 0.08. The luminosity ratio between hydrogen and deuterium are in good agreement for both methods. The error on the ratio for the luminosity monitor based method is significantly higher. The difference in the two methods can serve as a measure of the systematic uncertainty in the luminosity determination.

7.2 π^0 invariant mass reconstruction

When reconstructing π^0 events, first valid photon-induced calorimeter clusters should be identified. Photons will only produce signals in the calorimeter (and eventually in the preshower) and therefore these particles appear as trackless calorimeter clusters in the data stream. Since photons are not deflected in a magnetic field, tighter fiducial cuts have to be applied in order to select photons originating from the event vertex. A lower limit on the cluster energy of 0.8 GeV was applied in order to discriminate real signals from electronic detector noise.

Cluster x position:	$\mid$ `smCluster.rXlEw` $\mid \leq 125$
Cluster y postion :	$30 \leq \mid$ `smCluster.rYlEw` $\mid \leq 180$
Cluster Energy:	`smCluster.rE` ≥ 0.8

The π^0 particle has a lifetime of $8.4 \pm 0.6 \times 10^{-17}$ s, and a branching ratio of $98.798 \pm 0.032\%$ for a decay into 2 photons [121]. To select π^0s out of a data sample, the invariant mass, $M_{\gamma\gamma}$, of any combination of two photons in an event is studied:

$$M_{\gamma\gamma} = \sqrt{(p_{\gamma_1} + p_{\gamma_2})^2} = \sqrt{2E_{\gamma_1}E_{\gamma_2}(1 - \cos\theta_{\gamma\gamma})} = 2\sin\frac{\theta_{\gamma\gamma}}{2}\sqrt{E_{\gamma_1}E_{\gamma_2}}, \tag{7.3}$$

[3]The number of DIS events obtained in 2005 with a deuterium target quoted here, and in the following, is obtained by taking only part of the entire 2005 production: runs $1390 - 20032$.

Year	Beam	Target	bit-mask	DIS	$\mathcal{L}_{\text{lumi}}$ (pb^{-1})	C_{lumi} (mbarn^{-1})
1996d0	e^+	polarized H	0x405E81FC	773 611	21.2 ± 1.6	422 ± 32
1996d0	e^+	unpolarized H	0x405E01CC	963 747	25.8 ± 2.0	422 ± 32
1997d1	e^+	polarized H	0x401E81FC	1 181 769	47.5 ± 4.0	426 ± 36
1997d1	e^+	unpolarized H	0x401E01CC	1 643 606	32.4 ± 2.7	426 ± 36
2000d0	e^+	unpolarized H	0xC03E01CC	5 194 973	135.3 ± 9.7	417 ± 30
2004b0	e^+	polarized H	0x401880FC	2 203 814	58.2 ± 2.4	969 ± 29
2005b1	e^-	polarized H	0x401E81FC	4 269 113	127.3 ± 5.9	302 ± 14
Total				16 230 633	447.7 ± 28.4	
1996d0	e^+	unpolarized D	0x405E01CC	1 614 312	49.3 ± 3.7	422 ± 32
1997d1	e^+	unpolarized D	0x401E01CC	1 868 853	62.5 ± 5.3	426 ± 36
2000d0	e^+	polarized D	0xC03E81CE	5 711 490	165.6 ± 11.9	417 ± 30
2004b0	e^+	unpolarized D	0x401801CC	3 100 138	94.0 ± 2.8	508 ± 15
2005b1	e^-	unpolarized D	0xC03E01CC	3 355 942	105.8 ± 4.3	348 ± 14
Total				15 650 735	477.2 ± 28.0	

Table 7.1: Luminosity collected during various data taking periods and for the targets used in this analysis. Beam charge and bad-bit pattern are indicated, as well as the particular data production for that year.

where $\theta_{\gamma\gamma}$ is the angle between the two photons as indicated in Figure 7.2. The angle $\theta_{\gamma\gamma}$ can be calculated as:

$$\cos\theta_{\gamma\gamma} = \frac{{\text{clus}_1}^2 + {\text{clus}_2}^2 - (x^{\gamma_1} - x^{\gamma_2}) \cdot (x^{\gamma_1} - x^{\gamma_2}) - (y^{\gamma_1} - y^{\gamma_2}) \cdot (y^{\gamma_1} - y^{\gamma_2})}{2 \cdot \text{clus}_1 \cdot \text{clus}_1}, \qquad (7.4)$$

with $\text{clus}_{1(2)} = \sqrt{|(x^{\gamma_{1(2)}}, y^{\gamma_{1(2)}}, z^{\gamma_{1(2)}} - z_{\text{vert}})|^2}$, the position of the cluster in the HERMES coordinate system, z_{vert} being the event vertex determined by the scattered lepton track. When plotting the two-photon invariant mass spectrum a peak can be observed around the π^0 mass of 135 MeV, as was already shown in Figure 6.29. The width of the peak is an indication of the resolution of the calorimeter for π^0 detections. As seen from equations 7.3 and 7.4 not only the energy and position of the calorimeter clusters enter in the invariant mass calculation, but also the longitudinal position of energy deposition, $z^{\gamma_1(2)}$.

Not all photons created at the event vertex will reach the calorimeter; some will interact in the material in front of the preshower or calorimeter. Photons can e.g. start showering in active detectors such as the RICH or the TRD, or they can induce particle showers by interacting with frame material. The efficiency of photon detection can be studied by using a Monte Carlo sample. The ratio of generated photons in the acceptance over the reconstructed ones needs to be calculated to do that. A condition was imposed on the reconstructed photons requiring that the photon originates from the target cell, and has a cluster energy above 0.8 GeV. Moreover, the distance between the projected calorimeter position of the generated photon and the reconstructed calorimeter position needed to be less than 15 cm. The resulting photon efficiency is shown in Figure 7.3 as a function of momentum. A constant efficiency of about $91.08 \pm 0.33\%$ can be seen. The Monte Carlo sample was generated using the

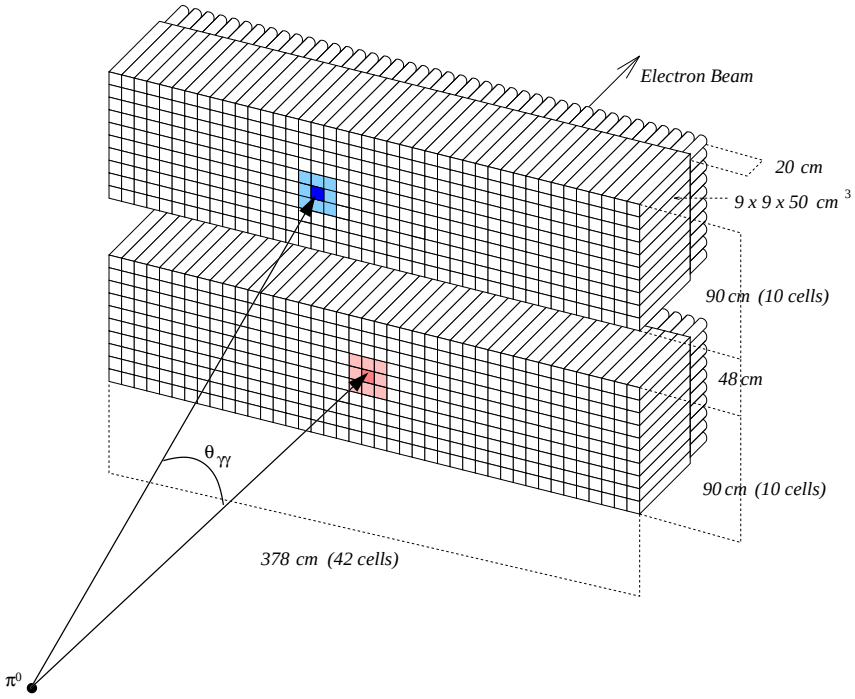

Figure 7.2: Schematic view of the decay of a π^0 into two photons and its detection by the HERMES calorimeter. The $z-$distance is not to scale in the figure.

background generator, i.e. generated uniform in energy and $\theta_x(= \arctan(\cos\phi \tan\theta))$ and $\theta_y(= \arctan(\sin\phi \tan\theta))$, and has only one photon per event.

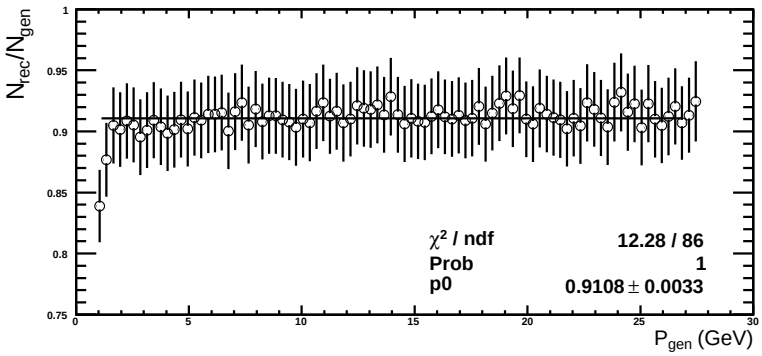

Figure 7.3: Photon reconstruction efficiency for photons generated in the HERMES acceptance with the background generator as a function of momentum. A constant was fitted to the spectrum. Fit values are indicated in the figure.

The efficiency for π^0 reconstruction was determined in the same way, with the same conditions with respect to the two reconstructed photons as in the single photon case and the additional constraint that the distance between 2 clusters is more than 15 cm. A reconstruction efficiency of $91.08 \pm 0.33\% \times 91.08 \pm 0.33\% = 83.0 \pm 0.6\%$ is expected from the efficiency of single photon events. The observed efficiency is shown in Figure 7.4. A non-constant behavior

can be seen. The decrease in efficiency for momenta above 11 GeV can be explained by the decrease in opening angle between both photons resulting in a calorimeter-cluster overlap: the two photons no longer produce two distinct clusters, they rather produce one overlapping cluster. This observation is important for DVCS analyses where only one calorimeter cluster is required in the analysis. Calorimeter clusters with energies above 11 GeV can originate from two photons from π^0 decay rather than from a single photon and thus form a potential background in the DVCS sample.

Also for momenta below 5 GeV a decrease in detection efficiency can be seen. This decrease is explained by the asymmetric energy distribution of the two photons from π^0 decay in combination with the low energy cutoff for the calorimeter clusters of 0.8 GeV. Between 5 and 11 GeV, a constant efficiency is observed. A fit yielded an efficiency of $81.2 \pm 0.5\%$, in good agreement with the expected $83.0 \pm 0.6\%$.

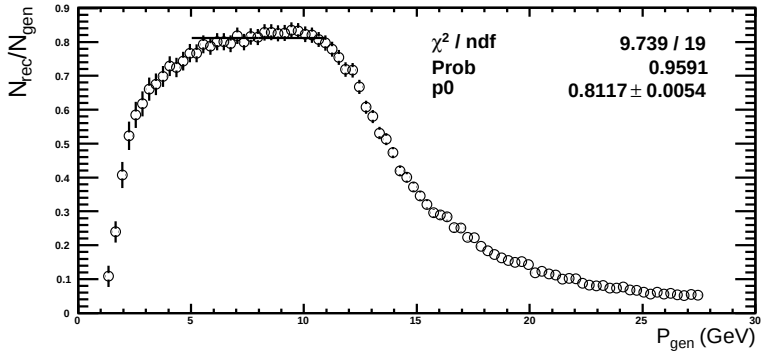

Figure 7.4: π^0 reconstruction efficiency for π^0 generated in the HERMES acceptance with the background generator as a function of π^0 momentum. A constant was fitted to the spectrum between 5 and 11 GeV. Fit values are indicated in the figure.

To test the performance of the calorimeter in more realistic circumstances, the π^0 efficiency was estimated using a physics generator like PYTHIA . The resulting efficiency is shown in the left panel of Figure 7.5. A similar behavior as for the data generated with the background generator can be observed, although the constant behavior between 5 and 11 GeV is less pronounced. The fit also yielded a slightly lower efficiency of about $77.3 \pm 0.6\%$. The observed behavior can be explained by the presence of at least one charged particle track per event, corresponding to the scattered DIS electron.

The same efficiency was estimated for the η particle as well, which has a branching ratio of $39.33 \pm 0.25\%$ in two photons. The mass of the η is $M_\eta = 547.30 \pm 0.12$ MeV [121]. A constant efficiency as a function of generated η energy can be seen in the right panel of Figure 7.5. This shows that for the same generated momenta, the opening angle between the two photons from η decay is larger than in the case of π^0, an effect of the larger η mass. The decrease in efficiency for low momentum η's is absent since both photons have larger energies than in the π^0 case. The constant fit yielded a result of $78.4 \pm 2.2\%$, which is comparable to the π^0 case, but again slightly lower than what was expected from the single photon efficiency.

The spectrum presented in Figure 7.5 contains only events with only one track and exactly two photons. When taking more tracks or photons into account a further drop in efficiency

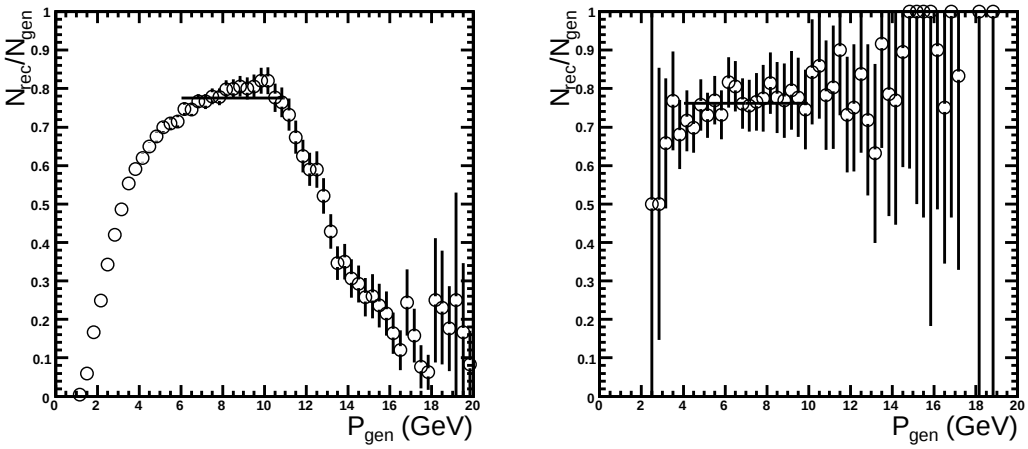

Figure 7.5: Left (Right): the π^0 (η) reconstruction efficiency as a function of generated π^0 (η) momentum for a PYTHIA sample with 1 charged track. A constant has been fitted to both spectra between 5 and 11 GeV. Fit values are indicated in the Table 7.2.

.

Track	$\epsilon_\pi^0(\%)$	$\epsilon_\eta(\%)$
0	81.2 ± 0.5	/
1	77.3 ± 0.6	78.4 ± 2.2
2	61.0 ± 0.6	59.5 ± 2.0
3	49.2 ± 0.8	45.8 ± 2.7
4	41.5 ± 1.5	32.5 ± 4.1
> 1	66.3 ± 0.4	63.7 ± 1.0

Table 7.2: Reconstruction efficiency for η (ϵ_η) and π^0 (ϵ_{π^0}) particles as a function of the number of tracks in an event for a Monte Carlo sample generated using the PYTHIA generator. The first row corresponds to the case presented in Figure 7.4, i.e. for a π^0 MC sample generated with the background generator.

was observed. Table 7.2 gives the result of the fit of the constant to both π^0 and η spectra for momenta between 6 and 11 GeV as a function of the number of tracks in a Monte Carlo event. If there are more tracks involved in an event the reconstruction efficiency clearly drops.

7.3 Energy resolution improvements

As mentioned before, the invariant mass reconstruction performance of the calorimeter for the π^0 is dependent on the energy and position reconstruction capabilities of the calorimeter. The previous chapter already showed a large difference in energy resolution between the cases when a photon was interacting in the preshower or not. Moreover, it is known that the resolution of lead-glass calorimeters is limited by fluctuations in the electromagnetic shower profile in combination with a strong attenuation of the Čerenkov photons in the lead glass

itself. The pulse in the preshower detector, however, may give an indication of the longitudinal shower development for a particular event: a large (small) energy deposition in the preshower corresponds to a shower maximum close to (far from) the calorimeter surface at $z = 720$ cm. Correspondingly, a large (small) attenuation of the optical Čerenkov photons in the lead glass is expected, yielding a smaller (bigger) signal in the PMT at the end of a calorimeter block. This behavior is indicated in Figure 7.6 (a), where the ratio $\mathrm{E_{gen}/E_{rec}}$ is plotted as a function of $\mathrm{E_{rec}}$ in different preshower pulse bins. Especially in the region $\mathrm{E_{rec}} < 10$ GeV a strong dependence of the ratio on the preshower pulse can be seen.

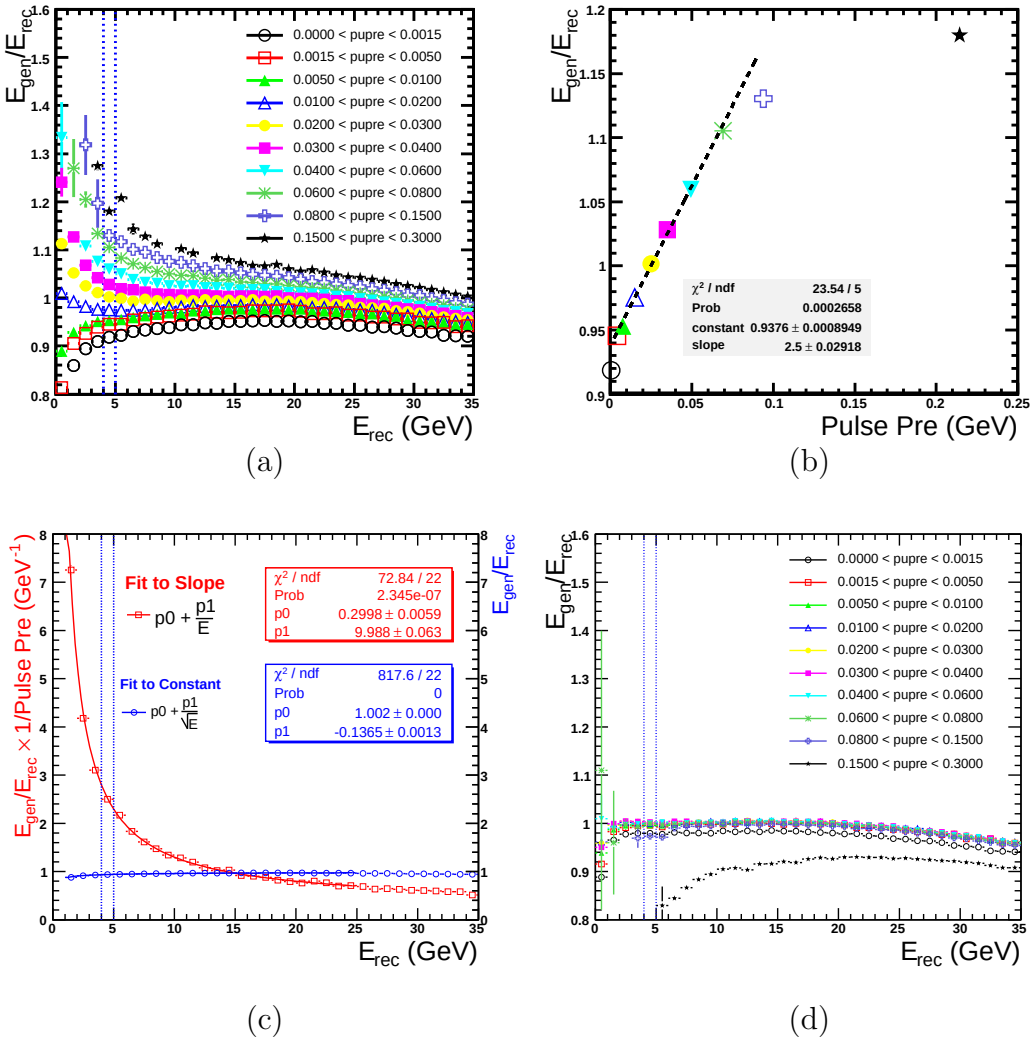

Figure 7.6: Panel (a) shows the dependence of the $\mathrm{E_{gen}/E_{rec}}$ ratio for a Monte Carlo sample on the preshower pulse as a function of $\mathrm{E_{rec}}$. Panel (b) shows the same dependence as a function of the preshower pulse, for the bin $4 < \mathrm{E_{rec}} < 5$ GeV. Panel (c) shows the fit to the slope and constant from the straight line fit to the spectrum in panel (b) for every $\mathrm{E_{rec}}$ bin. The units corresponding to the slope (constant) fit are indicated on the left (right) of the figure. Panel (d) gives the $\mathrm{E_{gen}/E_{rec}}$ ratio after applying the correction of equation 7.5.

In an attempt to correct the preshower pulse dependence of the $\mathrm{E_{gen}}/\mathrm{E_{rec}}$ ratio, the method proposed in Reference [194] was followed: for every bin in $\mathrm{E_{rec}}$ the ratio $\mathrm{E_{gen}}/\mathrm{E_{rec}}$ was plotted as a function of the preshower pulse. The mean of the ratio was determined by the mean of a Gaussian fit to the spectrum in that particular ($\mathrm{E_{rec}}$,Pulse Pre) bin. The method is presented in Figure 7.6 (a) and (b) for the bin 4 GeV $< \mathrm{E_{rec}} <$ 5 GeV, indicated by the dashed vertical lines in panel (a) of the figure. Panel (b) shows the $\mathrm{E_{gen}}/\mathrm{E_{rec}}$ ratio as a function of preshower pulse for that specific bin. The spectrum in panel (b) was fitted by a straight line. The first and last points were excluded from the fit, since the first data point corresponds to no energy deposition in the preshower, and the last one to a saturated preshower ADC.

Next, the constant and the slope of the fit to the spectrum in panel (b) were plotted as a function of $\mathrm{E_{rec}}$, shown in panel (c). The 'slope' ('constant') spectrum was fitted with a function of the form $p_0 + \frac{p_1}{E}(p_0 + \frac{p_1}{\sqrt{E}})$. From the fit parameters a correction function for $\mathrm{E_{rec}}$ could be obtained:

$$\begin{aligned} \mathrm{E_{gen}}/\mathrm{E_{rec}} = & \left((0.30 \pm 0.01) + \frac{(9.99 \pm 0.06}{\mathrm{E_{rec}}}\right) \cdot (\mathrm{PulsePre}) \\ + & \left((1.00 \pm 0.00) - \frac{(0.137 \pm 0.001)}{\sqrt{\mathrm{E_{rec}}}}\right) \end{aligned} \tag{7.5}$$

When applying this correction to the same data, panel (d) was obtained. It can be seen that the dependence on the preshower pulse disappears when applying equation 7.5. Only in the first and last preshower pulse bin a deviation from 1 can be observed. These two bins were indeed excluded in the fitting procedure.

Apart from the preshower pulse correction presented above, other corrections to the reconstructed photon energy were considered: the power-law correction given by equation 6.9, the HERMES 'standard' correction of $E_\gamma = 0.97 E_{\mathrm{clus}}$, and finally the 'anti-cubic' in combination with the square root correction of equation 6.11. Moreover, two calculations were used to estimate the cluster z-positions: one calulation uses the standard $z = 747.5$ cm, and the other the parametrization $z = 720 + (2.78 \cdot (\log(E_{\mathrm{clus}}/0.018) + 0.4(1.2)))$. Where 0.4 (1.2) is used in the case there is a (no) signal in the preshower. The latter calculation is referred to as 'Amaldi' in the following.

The mean and width of a Gaussian fit to the two-photon invariant mass spectrum in the π^0 region is indicated in Figure 7.7 for the data production years 2005 and 2000. From the figure it is clear that the π^0 resolution is best when applying the preshower correction of equation 7.5. When applying the power-law correction from equation 6.9 or the anti-cubic correction from equation 6.11, the resolution becomes worse, and when applying the standard 0.97 correction factor, it deteriorates further, independent of the z-reconstruction method. The improvement for the resolution in the lower energy bins is up to 50% when applying the preshower correction compared to the standard HERMES reconstruction method.

When looking at the mean of the two-photon invariant mass distribution it is hard to conclude which cluster energy correction gives the best results. Moreover, different HERMES data productions yielded different variations in the mean of the invariant mass distribution. One correction method could give good results for a particular production, while giving non-optimal results for another data production. This observation complicated the decision of the best method for reconstructing π^0 events. For the 2000 production the preshower correction method and the correction of 0.97 in combination with the z-position calculation from Amaldi

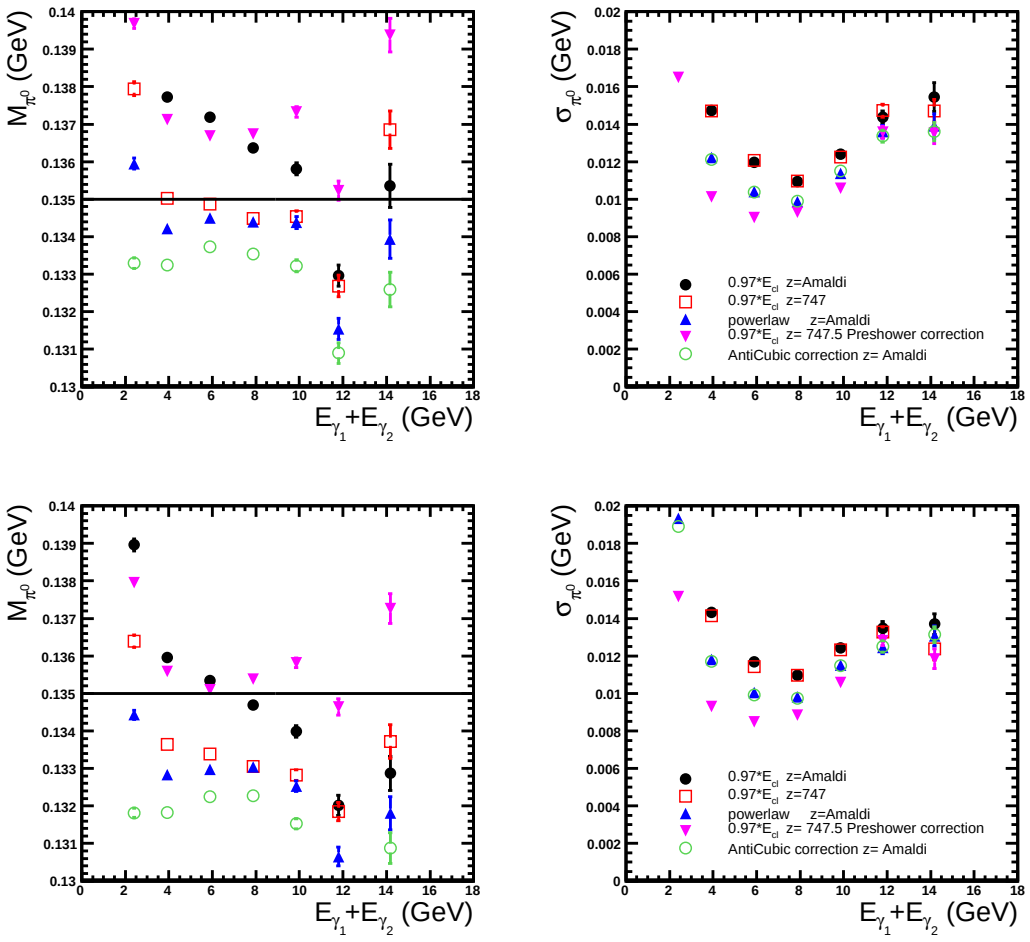

Figure 7.7: The left (right) two panels show the mean (width) of a Gaussian fit to the 2-photon invariant mass spectrum for various reconstruction methods as indicated in the figure legend. The upper (lower) two plots are for data productions of the year 2005 (2000).

yields the best results. For the 2005 production the standard HERMES reconstruction method and the power-law correction (equation 6.9) in combination with the z-position calculation from Amaldi give the best result. This observation indicates a varying calorimeter calibration over the different data production years.

When applying the preshower correction (equation 7.5), in combination with the z-position calculation from Amaldi, poorer results were obtained. This can be explained by the fact that the preshower correction itself already is a good indication of the longitudinal shower development. An investigation of the longitudinal shower profile as a function of preshower pulse could probably give better z- position estimates.

The preshower itself is only calibrated by using the MIP peak from the energy deposition of hadrons in the scintillator. The MIP peak itself corresponds to about 2 MeV of energy deposition. When extrapolating this value to larger energy depositions -up to 100 MeV as indicated in Figure 7.6- one can easily see that a small miscalibration of the MIP peak introduces large deviations when going to higher energy depositions. An improved preshower calibration

is likely to give more stable results for the various data productions. By using the one and two MIP signals from photons showering in the preshower in combination with the MIP signal from pions, the calibration probably could be improved.

An additional requirement for the calorimeter clusters to be separated by more than 15 cm lead to a decrease in energy dependence of the invariant mass peak for energies above 11 GeV. This condition avoids the two photon clusters to be overlapping.

7.4 Sensitivity to exclusive reactions

In order to check the sensitivity of the HERMES spectrometer to exclusive π^0 reactions, a Monte Carlo sample based on the Vanderhaeghen-Guidal-Guichon (VGG) model was generated and tracked through the spectrometer. More information about the VGG model can be found in section 2.6.5 and in Reference [58]. The spectra presented in this section are only used to show the sensitivity of the HERMES spectrometer with respect to exclusive π^0 production, therefore the Monte Carlo data sample was kept unweighted. Since the recoiling proton remains undetected in the absence of a recoil detector, the missing mass technique has to be applied in order to select exclusive processes.

The M_X^2 distribution for a Monte Carlo sample obtained with the VGG model is indicated in Figure 7.8. The spectrum in the figure is for $E_{\gamma_1} + E_{\gamma_2} > 7$ GeV. A Gaussian was fitted to the peak of the histogram, yielding a mean value of 0.743 ± 0.007 GeV^2, to be compared with the expected value of $0.938 \cdot 0.938 \approx 0.88$ GeV^2. The width of the distribution is about 1.28 GeV^2.

The tail on the left of the distribution is due to events where the recoiling proton has a substantial momentum; when applying a cut on the total recoiling proton momentum, $p_T^{\text{recoil}} < 2$ GeV, the tail disappears. The tail on the right of the peak vanishes when applying a cut on the z-component of the recoiling proton momentum, $p_z^{\text{recoil}} < 2.5$ GeV. This shows that the HERMES spectrometer only is capable of reconstructing exclusive events if the recoiling proton has a low momentum.

The dependence of the peak position and the width of the M_X^2 distribution on the sum of the photon energies can be seen in Figure 7.9. When applying the preshower correction from equation 7.5, an improvement in resolution of up to about 30% can be observed. However, the mean of the Gaussian fit to the M_X^2 distribution is biased towards lower values when applying this correction. The method using the anti-cubic correction in combination with equation 6.11 is slightly biased towards higher values, whereas the other three cluster energy corrections are coincident. A clear energy dependence can be seen for all methods. In the interval 7 GeV $< E_{\gamma_1} + E_{\gamma_2} <$ 13 GeV, the peak position is close to the expected value of 0.880 GeV^2.

The resolution in z, the energy fraction of the hadron with respect to the available energy ($z = E_h/\nu$), and in recoiling three-momentum p^{recoil} are indicated in Figures 7.10 and 7.11. No bias can be seen in either distributions. The resolution in z is about 0.05, as determined from a Gaussian fit. The distribution for the recoiling three-momentum p^{recoil} is asymmetric. It is very hard to reconstruct recoiling three-momenta with the information from the forward spectrometer. The Recoil Detector is expected to improve the resolution in p^{recoil} by a large amount [117].

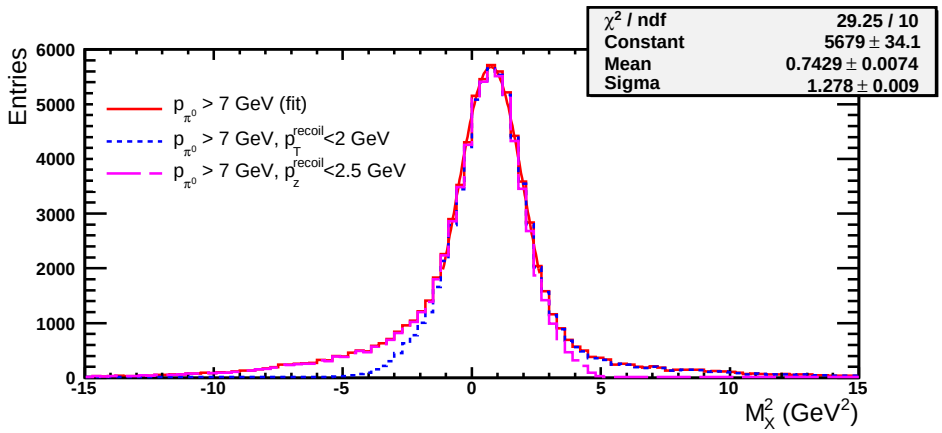

Figure 7.8: Missing mass (squared) distribution for the exclusive Monte Carlo events with $p_{\pi^0} > 7$ GeV. A Gaussian was fitted to the histogram. Fit parameters are indicated in the figure. The dashed (long-dashed) histogram shows the same distribution with the additional requirement that $p_T^{\text{recoil}} < 2$ GeV ($p_z^{\text{recoil}} < 2.5$ GeV).

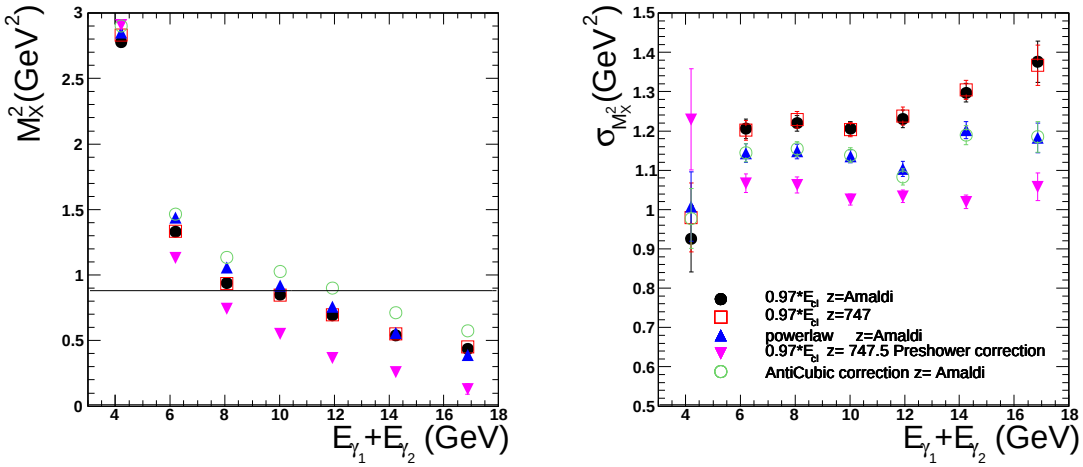

Figure 7.9: Mean (left) and width (right) of the missing mass distribution as a function of π^0 energy for various reconstruction methods.

The distributions presented in this section show that exclusive π^0 production can be measured at HERMES , even without Recoil Detector. The resolutions in M_X^2 and p^{recoil}, however, are not optimal. Nevertheless an attempt can be made to extract the π^0 cross section.

7.5 Background estimation in the exclusive π^0 sample

For the analysis of exclusive π^0 processes, additional cuts are required to the data sample: exactly one track and exactly two photons need to be present to have a valid event. Moreover, no other tracks should enter the event, like tracks not passing the fiducial cuts or magnet-tracks (see Figure 3.15):

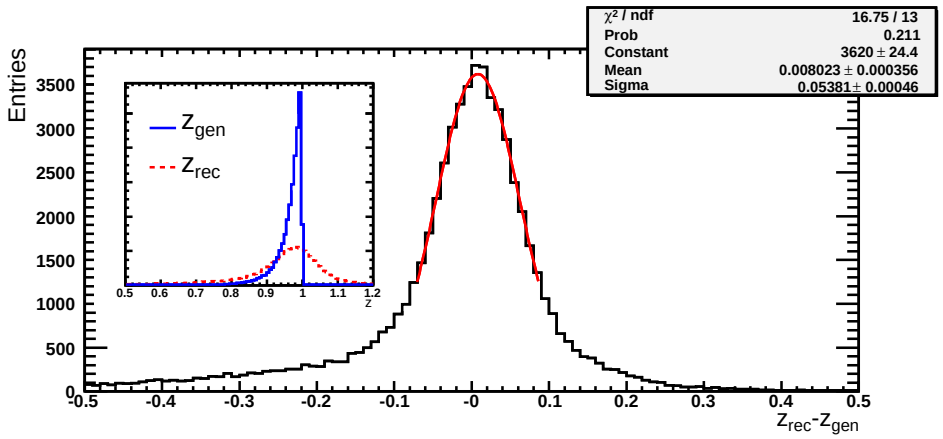

Figure 7.10: Difference between the generated and reconstructed fraction of energy z for the exclusive Monte Carlo. A Gaussian was fitted to the histogram, fit parameters are indicated in the figure. The inset shows the $z_{\rm gen}$ (full line) and $z_{\rm rec}$ (dashed line) distributions.

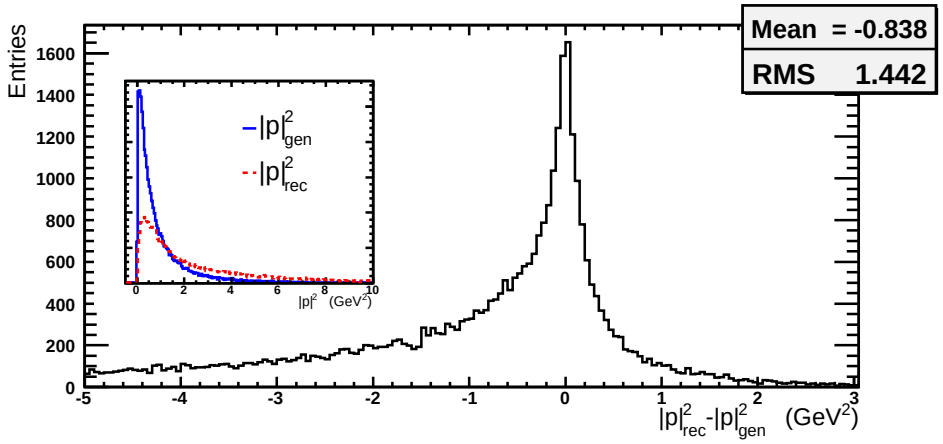

Figure 7.11: Difference between the generated and reconstructed recoiling momentum $p^{\rm recoil}$ for the exclusive Monte Carlo. The inset shows the $p^{\rm recoil}_{\rm gen}$ (full line) and $p^{\rm recoil}_{\rm rec}$ (dashed line) distributions. A large smearing can be observed.

.

ntrack $= 1$
nphot $= 2$
magnettrack/badtrack $= 0$

After applying these cuts, the M_X^2 spectrum was investigated in different energy bins. To estimate the amount of background in comparison with the π^0 yield expectations from the VGG model, the latter model was compared to the data. The comparison is shown in Figure 7.12. From the figure it is clear that the M_X^2 distributions are dominated by non-exclusive background, and that the VGG model certainly does not overestimate the experimental data.

A good estimation for the non-exclusive background in the exclusive event sample can be obtained by using Monte-Carlo techniques. Within HERMES there are two widely used Monte

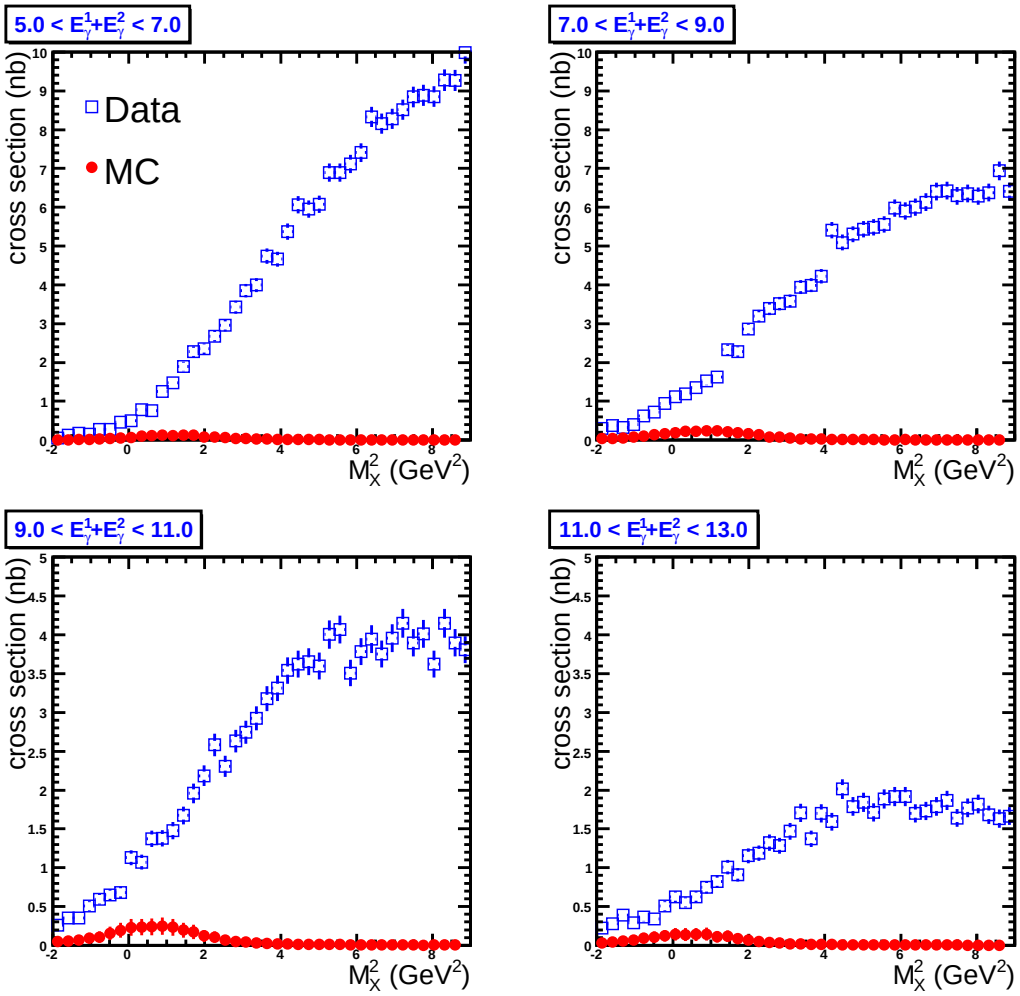

Figure 7.12: M_X^2 spectrum for hydrogen data (squares) in comparison with the spectrum from the VGG model for four different energy bins.

Carlo event generators, namely PYTHIA and DISNG.

Using PYTHIA , scattered lepton kinematics are generated according to the cross section of a particular process. PYTHIA also decides on the scattering process. The generated cross sections were adjusted to describe the HERMES data. These adjustments are described in e.g. Reference [195]. The resulting struck quark and remaining diquark from PYTHIA are fragmented into hadrons by the JETSET program, which is based on the LUND string model. The fragmentation described by JETSET was tuned to describe the experimentally observed semi-inclusive charged hadron multiplicity [172].

The DISNG package on the other hand generates the lepton kinematics uniformly in ν and Q^2, and not according to the cross section. Therefore, each event has to be assigned an event weight dependent on the cross section. The scattering process is decided on by the LEPTO program [167], whose output is further processed by JETSET.

The PYTHIA6 generator covers a variety of particle processes such as Deep Inelastic Scattering (DIS), Vector Meson Dominance (VMD) reactions, Photon-Gluon Fusion (PGF),$\cdots$.

Additionally, the PYTHIA Monte-Carlo contains cross sections for exclusive vector meson production. However, the PYTHIA generator does not contain any exclusive pseudoscalar meson generator. Therefore a Monte-Carlo generator like PYTHIA is a good tool to describe the non-exclusive background in an exclusive sample[4].

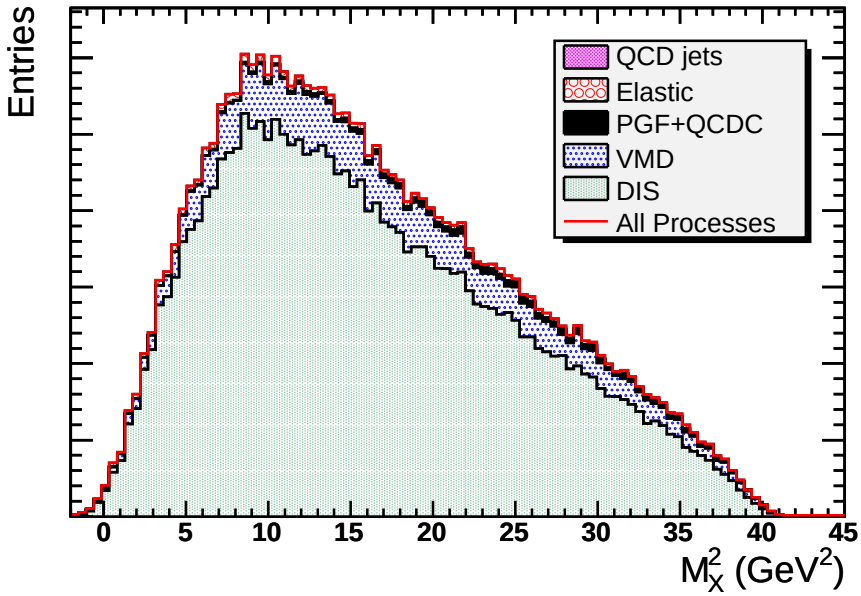

Figure 7.13: The contribution of different processes to the missing mass spectrum after applying exclusive π^0 cuts for a Monte Carlo sample generated by PYTHIA.

Figure 7.13 shows the different process contribution to the total M_X^2 spectrum after applying all exclusive cuts mentioned before. The main contribution is coming from the DIS process ($\gamma^* q \to q$), while a substantial amount comes from exclusive vector meson production (Vector Meson Dominance (VMD)). These entries in the sample could be originating from the decay of the ω meson, which has a lifetime of 7.8×10^{-23} s and decays into $\pi^+\pi^-\pi^0(88.8\%)$ and $\pi^0\gamma(8.5\%)$ [121]. Most of these events are having higher missing masses. A third contribution comes from Photon-Gluon Fusion (PGF, $g\gamma^* \to q\bar{q}$) and QCD-Compton processes (QCDC, $q\gamma^* \to qg$). PGF and QCDC are indicated in Figure 2.12(b) and (c). Smaller contributions come from elastic scattering, or QCD processes like quark-(anti)quark scattering (annihilation) and quark-gluon scattering.

The M_X^2 distributions for PYTHIA6 and data are compared in Figure 7.14. The π^0 was reconstructed using the standard HERMES reconstruction method, i.e. the cluster energy was corrected with the factor 0.97 and $z = 747.5$ was used. The aforementioned methods for π^0 reconstruction were all tested, but none of them yielded any improvement to the data - Monte Carlo comparison. For both data and Monte Carlo samples a normalization to the cross section level was performed. For doing so, the luminosity is required. In the case of experimental data,

[4]Even if PYTHIA would contain a description of hard exclusive pseudoscalar meson production, the process could be 'turned off' in the simulation

the luminosity was estimated by the count rate in the luminosity monitor. For the Monte Carlo, the luminosity can be calculated from information provided by the process generator.

The Monte Carlo overshoots the data in nearly all energy bins, in contrast to what is expected. The overshoot is even larger for the lower energy bins which are not shown in the figure. This result indicates that the PYTHIA Monte Carlo is not well tuned in the presented kinematic region. One would have to iteratively tune the Monte Carlo to the semi-inclusive π^0 sample in order to be able to use that Monte Carlo for subtraction of the non-exclusive background. This procedure is expected to take several months and was beyond the timescale foreseen for this work. The result somehow is surprising given the extensive tuning of the PYTHIA generator as mentioned e.g. in Reference [195]. For the analysis of exclusive π^+ production at HERMES the non-exclusive background is described with the PYTHIA event generator. In that analysis the non-exclusive background is well described by the PYTHIA Monte Carlo [196].

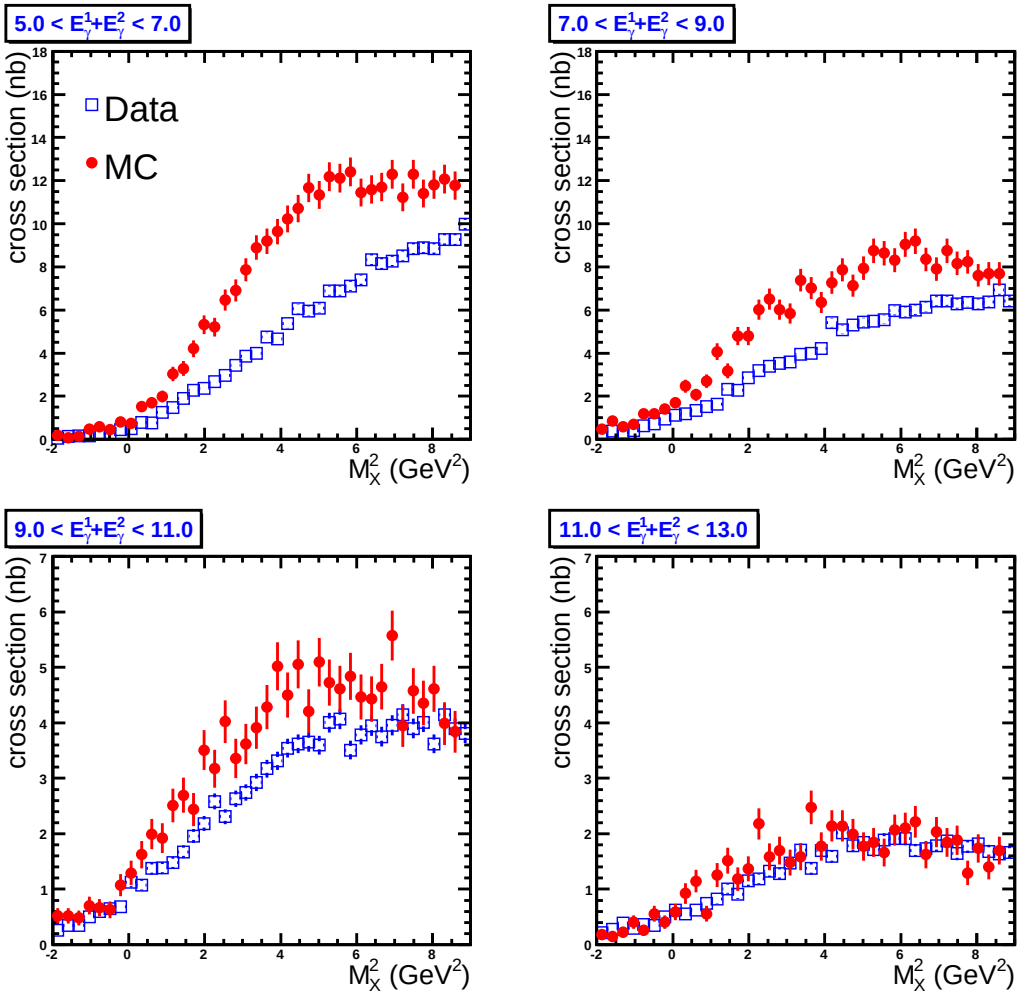

Figure 7.14: M_X^2 spectrum for all hydrogen data (squares) in comparison with the spectrum from PYTHIA6 for four different energy bins.

The comparison of the data to DISNG is shown in Figure 7.15. Also here an overshoot of the Monte Carlo can be seen in each energy bin. The description of the observed experimental spectrum by the Monte Carlo is bad. DISNG (i.e. LEPTO) was extensively tuned to describe the charged pion multiplicity rate over a large kinematic range and in different variables. It was expected that the same tune would describe the experimental π^0 spectra relatively well. This obviously is not the case, and more work is consequently required to tune the Monte Carlo in this particular region. The overshoot of the Monte Carlo sample was identified to be caused at the generator level, not by the updated Monte Carlo description of the calorimeter.

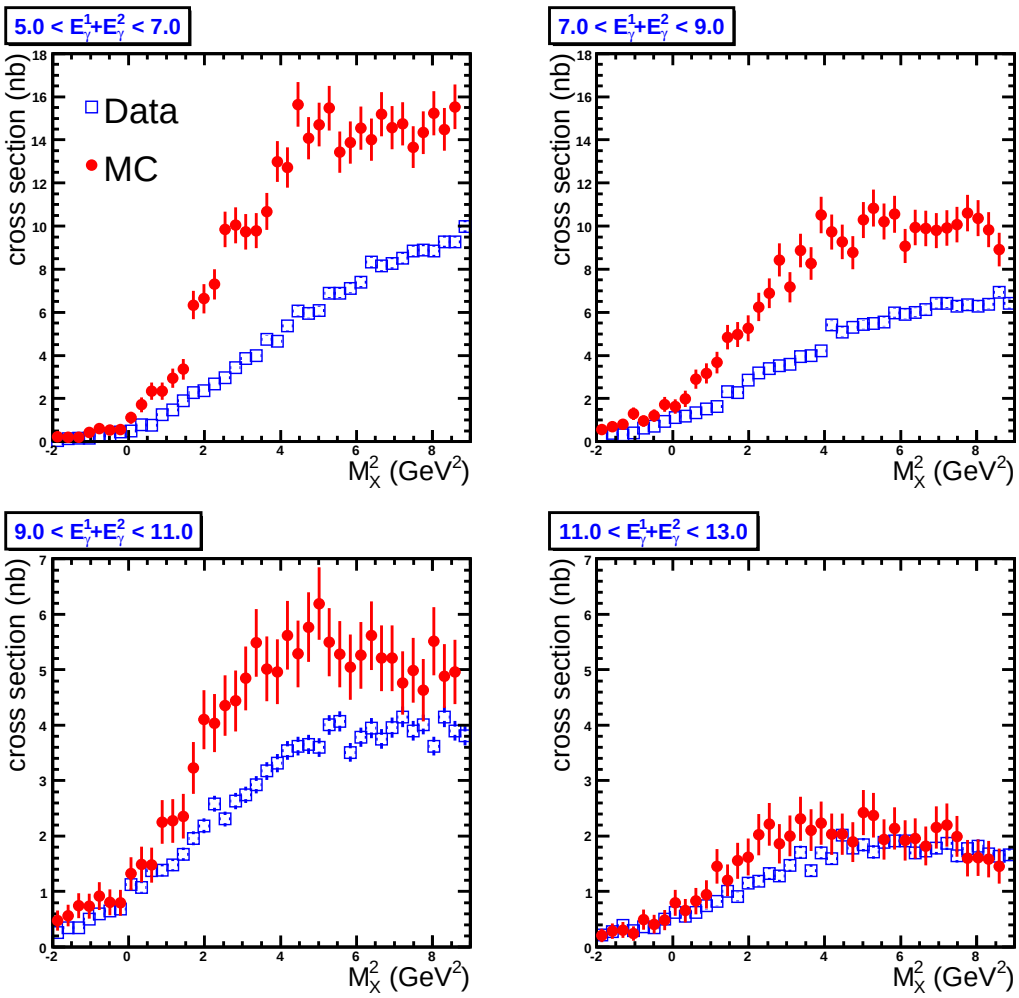

Figure 7.15: M_X^2 spectrum for hydrogen data (squares) in comparison with the spectrum from DISNG for four different energy bins.

After showing that the absolute normalization does not yield a good comparison between data and Monte Carlo, it was attempted to normalize both samples to unity. Such comparison compares the shape of both distributions. Also this comparison showed a non-negligible difference between data and Monte Carlo.

From the above comparison of the experimental data and the two Monte Carlo generators, DISNG and PYTHIA, it can be concluded that these do not describe the non-exclusive

background in the exclusive event sample adequatly. To improve the Monte Carlo description, a dedicated tuning of the Monte Carlo parameters is required. This tuning, however, is a non-trivial effort, given that most theory groups only give estimates for the fragmentation functions in the semi-inclusive region. By applying the aforementioned exclusive data selection criteria, a special kinematical domain is reached, where a semi-inclusive description may not be fully appropriate. A comparison between data and PYTHIA Monte Carlo over the entire M_X^2 range of the exclusive π^0 event sample is shown in Figure 7.16. It can be seen that at high M_X^2 the data and Monte Carlo agree relatively well with eachother. For $M_X^2 < 10$ GeV^2, the Monte Carlo overshoots the data.

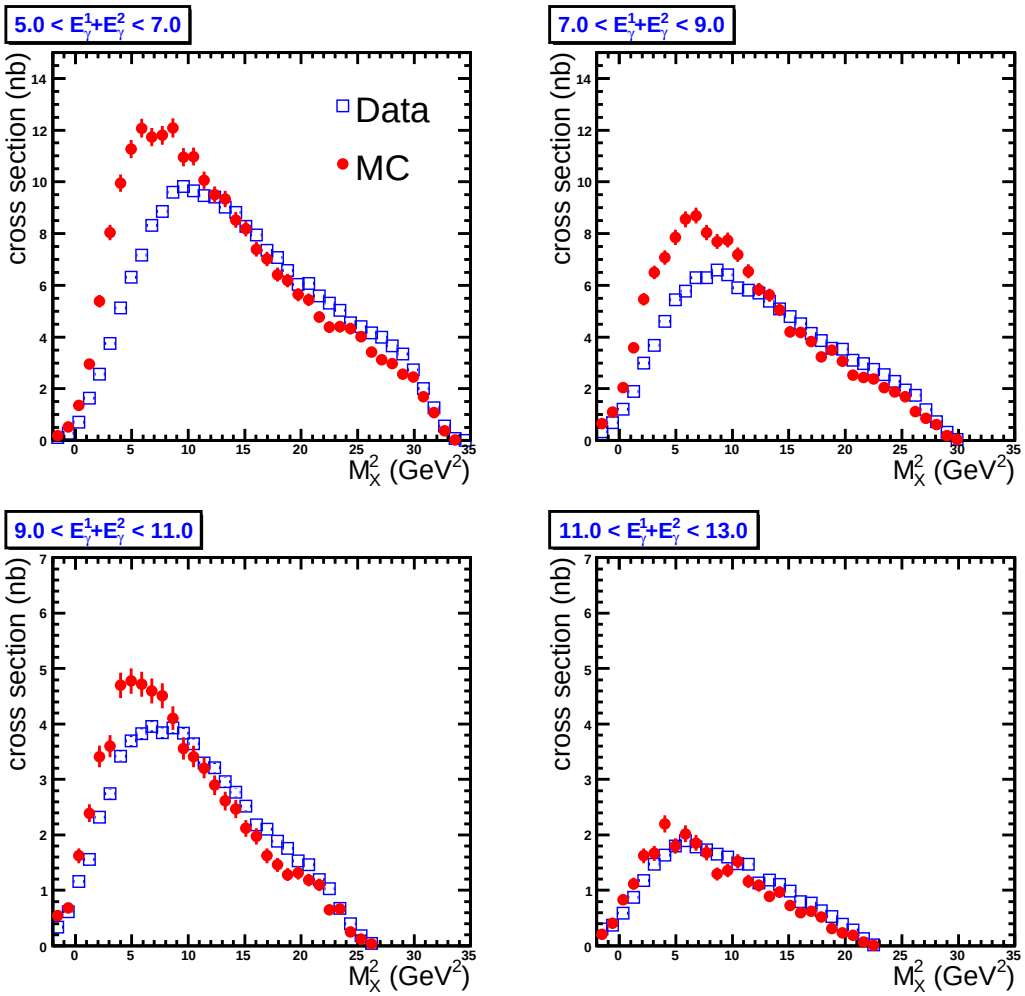

Figure 7.16: M_X^2 spectrum for all hydrogen data (squares) in comparison with the spectrum from PYTHIA6 for four different energy bins over the entire range in M_X^2.

7.6 π^0 and η multiplicity ratio's on H/D targets

In order to check the transition from the semi-inclusive to the exclusive regime, the production ratio of π^0 and η mesons was analyzed as a function of z, the energy fraction of the produced

meson ($= E_h/\nu$). As already mentioned in section 2.6.4, the production ratio for both of these mesons on a neutron target is expected to be suppressed. Figure 2.27 on page 44 shows the prediction from Reference [63].

In order to study the neutron suppression, π^0 and η multiplicities for both hydrogen and deuterium targets were counted as a function of z, and their ratio was examined. Since the binding energy of the deuterium nucleus is small, the multiplicity on a deuterium nucleus can be seen as the average of neutron and proton multiplicities. If neutron suppression for exclusive reactions is present, the multiplicity ratio, M_D/M_H, is expected to drop as a function of $z(= \frac{E_{\gamma_1}+E_{\gamma_2}}{\nu})$.

The data used for the multiplicity ratio determination are hydrogen and deuterium data from the data taking periods of 1996, 1997, 2000, 2004 and 2005. These data were binned as a function of z and two-photon invariant mass, $M_{\gamma\gamma}$. The π^0 (η) candidates were selected by requiring $0.06\ \mathrm{GeV} < M_{\gamma\gamma} < 0.2\ \mathrm{GeV}$ ($0.3\ \mathrm{GeV} < M_{\gamma\gamma} < 0.9\ \mathrm{GeV}$). In order to estimate the background in each $z-$bin, the spectra were fit with a Gaussian function to describe the peak of the spectrum, corresponding to the π^0 or η, plus a polynomial function to describe the combinatorial background in the sample. Both a first and second order polynomial were considered to describe the background. Both functions yielded similar background estimates. Next, the counts in each $z-$bin were corrected. To do so, the integral of the background polynomial function over the entire invariant mass range was subtracted from the number of entries in every $z-$bin. The fits of a Gaussian function and second order polynomial to the η-invariant mass window for a deuterium target are shown in Figure 7.17 as an example. The mean of the Gaussian is indicated in the figure by parameter p_1, and the width by parameter p_2. The cluster energies of both photons in the figure were corrected for the signal in the preshower using equation 7.5. The mean of the Gaussian fit is close to the η mass. Small differences of the mean of the Gaussian with respect to the η mass do not matter since the invariant mass window is integrated over. Only in the lowest z bin the fit procedure failed to describe the signal and background. Therefore, this bin is omitted in the further analysis.

Exclusive reactions occur at $z \sim 1$. Theoretically z can not be larger than 1. However, due to the experimental resolution, z values up to 1.2 were allowed in the analysis.

Figure 7.18 shows the result of the combined Gaussian and second order polynomial fit in the π^0-invariant mass window, for 4 z-bins. Also here is the mean (width) of the Gaussian fit indicated in the figure by the parameter p_1 (p_2). In the case of π^0 production it is seen that the signal to (combinatorial) background is larger than in the case of η production. The figure was made after correcting the cluster energies for the preshower pulse dependence.

Once all counts for both π^0 and η samples for the different targets are known, the multiplicity ratio's in each bin can be calculated. Both hydrogen and deuterium samples, however, need to be corrected for the luminosity. As mentioned before, the luminosity correction can be done in two ways: either by taking the number of DIS events into account, or by taking the luminosity from the luminosity monitor. For this analysis the normalization to DIS events was chosen, since possible acceptance and efficiency effects are expected to cancel when taking the ratio of deuterium to hydrogen multiplicities. When calculating the luminosity using the luminosity monitor count rate, an error coming from the luminosity constant (see Table 7.1) needs to be taken into account. This error would unnecessarily increase the error on the

obtained ratio's. The multiplicity ratio's were calculated as:

$$\frac{\mathcal{M}_D^{\pi^0}}{\mathcal{M}_H^{\pi^0}} = \frac{\frac{N_D^{\pi^0}}{\mathcal{L}_D}}{\frac{N_H^{\pi^0}}{\mathcal{L}_H}} = \frac{\frac{N_D^{\pi^0}}{N_{\rm DIS}^D}\epsilon_{\rm DIS}^D\sigma_{\rm DIS}^D}{\frac{N_H^{\pi^0}}{N_{\rm DIS}^H}\epsilon_{\rm DIS}^H\sigma_{\rm DIS}^H} = \frac{N_D^{\pi^0}}{N_H^{\pi^0}} \cdot \frac{N_{\rm DIS}^H}{N_{\rm DIS}^D} \cdot \underbrace{\frac{\sigma_{\rm DIS}^D}{\sigma_{\rm DIS}^H}}_{\mathcal{R}(x,Q^2)} = \frac{M_D^{\pi^0}}{M_H^{\pi^0}} \cdot \mathcal{R}(x,Q^2), \qquad (7.6)$$

where equation 7.2 has been used, and the fact that $\epsilon_{\rm DIS}^D = \epsilon_{\rm DIS}^H$. The cross section ratio $\mathcal{R}(x,Q^2)$ has been omitted in the following as the ratio clearly is (x,Q^2) dependent, and no initial averaging of the kinematic variables was made. Afterwards the observed ratio may be corrected for this cross section ratio $\mathcal{R}$.

Figure 7.19 shows the multiplicity ratio of π^0 and η mesons as a function of z. The average x, Q^2 and ϵ (according to equation 2.25 on page 14) are indicated in every z-bin as well. The predicted neutron suppression for exclusive reactions is clearly visible for both cases. This forms clear evidence that at HERMES kinematics exclusive η and exclusive π^0 production can be accessed. Moreover, the observation forms evidence that these exclusive regime is not readily described by the fragmentation function formalism, typical for SIDIS type reactions. In the case of π^0 production, there is a deviation of nearly 6σ from the multiplicity ratio in the semi-inclusive regime. For η production the significance of the deviation is about 3σ.

The effect was observed in case of π^0 production in Reference [197] as well, for a subsample of the presented data only. Results of the reference are in agreement with the ones presented here within the statistical errors.

Most of the systematic effects in the ratio should cancel since most unpolarized high density data was taken right after the polarized low density during the same fill. When the target was e.g. filled with polarized hydrogen during a particular fill, unpolarized deuterium usually was injected in the target cell at the end of that fill and vice-versa. Since hydrogen and deuterium data followed closely after one another, time dependent effects are expected to cancel in the ratio. Also the reconstruction efficiency for π^0 and η is independent of the target gas. Nevertheless a discussion on potential systematic effects in the ratio follows.

The data of the individual data taking periods were compared separately and no systematic deviations in the ratio could be observed. Only the 1996 data production showed some differences the origin of which could be pinned down to the large average number of photons per event in the 1996 sample. This is most likely due to a lower calorimeter threshold setting during that year. When applying an additional condition that the number of photons per event should be smaller than 5, all data sets agreed within the statistical error. For the 1997 sample the ratio between deuterium and polarized or unpolarized hydrogen showed no sizable differences.

The largest estimated systematic uncertainty in the analysis presented here is coming from the estimation of the combinatorial background in every $z-$bin. To estimate the size of this uncertainty, the multiplicity ratio was calculated using three different cluster energy correction methods. The method used for the data presented in Figure 7.19 is the standard HERMES reconstruction method. The two other use the preshower pulse for correcting the cluster energy in combination with either $z_{\rm calo} = 747$ or $z_{\rm calo}$ according to the parametrization from Amaldi. The latter introduces a systematical bias towards lower invariant mass values for the fitted Gaussian peak for both the π^0 and η sample. Figure 7.20 shows the invariant mass spectrum in the η region for all z-bins, but this time without applying the preshower correction. Compared with Figure 7.17 it can be seen that the width of the distribution is broader in all z-bins.

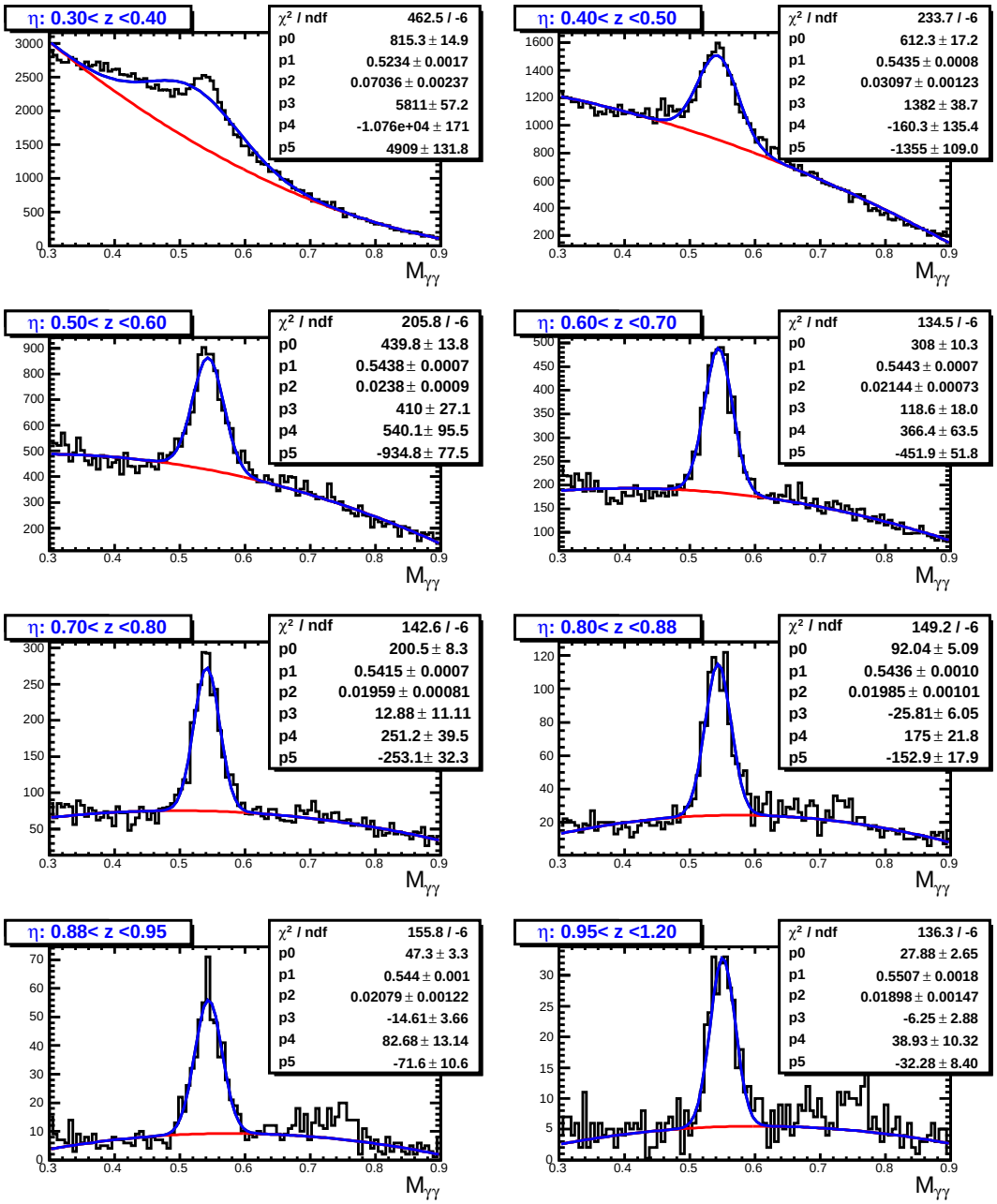

Figure 7.17: Two-photon invariant mass spectrum in various z-bins in the η mass region. The data presented is from 1996, 1997, 2000, 2004 and 2005 for a deuterium target. A Gaussian and a second order polynomial were fitted to the spectrum. The fit parameters are indicated in the figure. p_1 (p_2) is the mean (width) of the Gaussian function. The cluster energies were corrected using the preshower pulse. A calorimeter position of $z_{\text{calo}} = 747.5$ cm was assumed.

Only in the last three bins for both η and π^0 deviations in the resulting ratio between the various methods could be observed. A problem when estimating the size of the effect is that the determination of z itself depends on the calorimeter cluster energy. The shift in

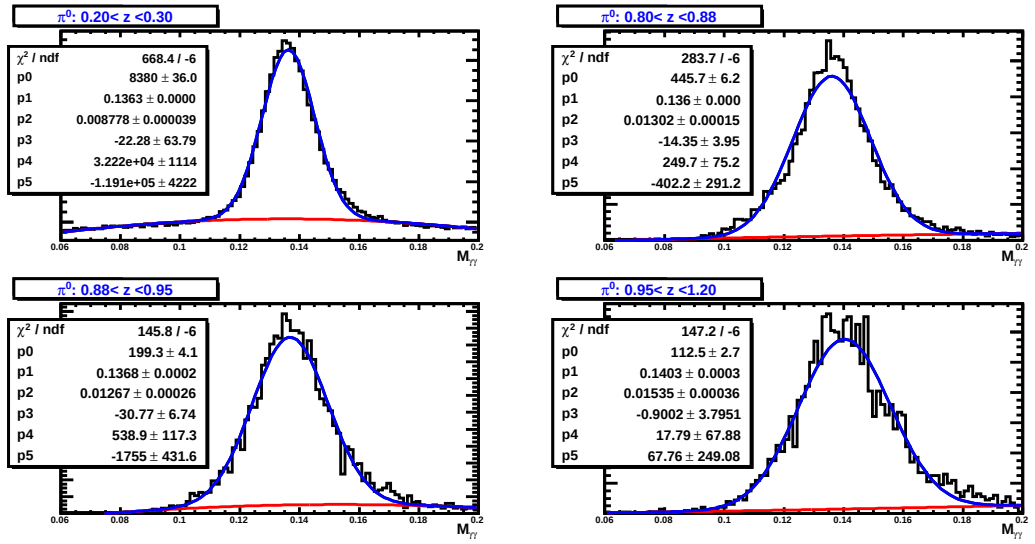

Figure 7.18: Two-photon invariant mass spectrum in various z-bins in the π^0 mass region. The data presented is from 1996, 1997, 2000, 2004 and 2005 for a deuterium target. A Gaussian and a second order polynomial were fitted to the spectrum. The fit parameters are indicated in the figure. p1 (p2) is the mean (width) of the Gaussian function. The cluster energies were corrected using the preshower pulse. A calorimeter position of $z_{\text{calo}} = 747.5$ cm was assumed. The y-axis is arbitrary.

the last two bins therefore is correlated: when in the penultimate bin the ratio increases, it decreases in the last bin and vice-versa. This correlation complicates quantifying the size of this error. Nevertheless, an upper limit for the error due to the reconstruction and background subtraction method can be estimated assuming both values to be uncorrelated. The average deviation for the η (π^0) particle was 6.7% (4.2%) in the last but one bin and 2.3% (1.5%) in the last bin.

The overall z-resolution is dominated by the energy resolution of the calorimeter. The scattered lepton energy which enters in the calculation via the variable ν is very well determined by the tracking system in combination with the magnet. The momentum resolution for charged particles worsened after the installation of the RICH in 1998, due to an increase in multiple scattering caused by the heavy RICH gas. This is not expected to have a significant effect on the multiplicity ratio, since the $z-$resolution is limited by the energy resolution of the calorimeter. Therefore the data sets of 1996-1997 can be combined with the 2000-2004-2005 data sets without any problem.

Another potential systematic effect is the trigger efficiency. This efficiency has been extensively studied for various other HERMES analysis (see e.g. Reference [22]). Only in 2000 a small efficiency drop was observed for a shorter time period. The statistics acquired during this period is negligible with respect to the entire data sample.

In order to estimate the uncertainty in the DIS yield, which is required for normalizing both hydrogen and deuterium samples, the numbers of DIS events corrected for dead time were evaluated for the different years. The results are shown in Table 7.3. The overall dead time for the hydrogen sample is 4.9%, and 8.5% for the deuterium sample. The ratio of dead time corrected DIS events for the deuterium over the hydrogen data sample is 0.997, the ratio of

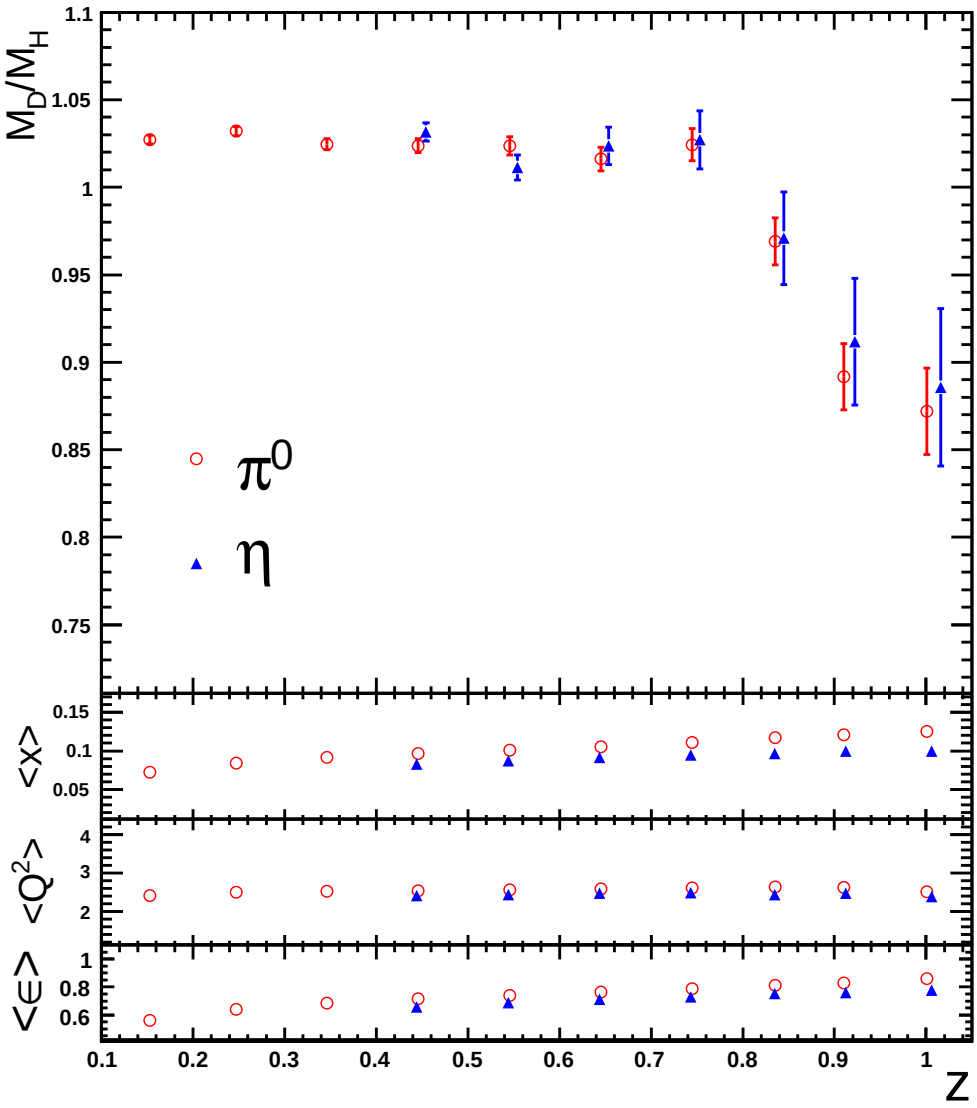

Figure 7.19: Multiplicity ratio of π^0 (circles) and η (triangles) mesons on a deuterium target and a hydrogen target as a function of $z(=\frac{E_h}{\nu})$. A clear drop in the ratio is seen for increasing z as an evidence for neutron suppression in exclusive reactions. The average x, Q^2, and ϵ for each z-bin are indicated in the figure. The error bars represent the statistical error only. The η-points in the upper panel have an offset of 0.01 for visibility.

identified DIS events for both samples is 0.964, a difference of about 3%. The dead time, however, should also be directly reflected in the π^0 and η multiplicities, and therefore is not expected to play a role in the ratio. The discussion on the dead time shows that all data samples were obtained in similar experimental conditions. The dead time varies only within 10% between the different data taking periods.

In order to quantify the expected multiplicity ratio in the semi-inclusive region, $z < 0.7$, the

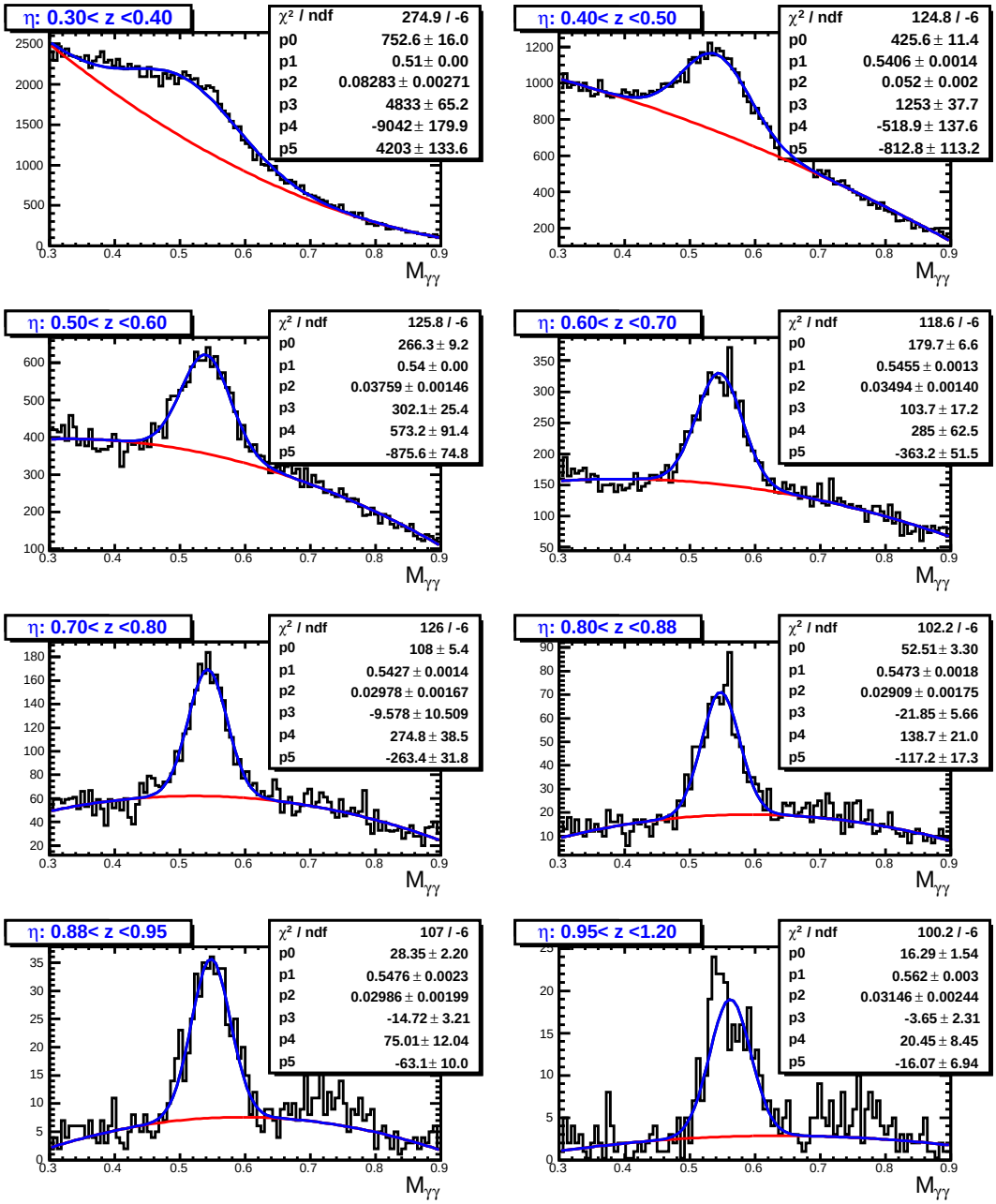

Figure 7.20: Two-photon invariant mass spectrum in various z-bins in the η mass region. The data presented is from 1996, 1997, 2000, 2004 and 2005 for a deuterium target. A Gaussian and a second order polynomial have been fitted to the spectrum. The fit parameters are indicated in the figure. p1 (p2) is the mean (width) of the Gaussian function. No preshower correction was applied. A calorimeter position of $z_{\text{calo}} = 747.5$ cm was assumed.

cross section for the semi-inclusive production of a hadron of type h needs to be considered:

$$\frac{d^3\sigma^h}{dxdQ^2dz} = \frac{\sum_f e_f^2 q_f(x,Q^2)\cdot D_f^h(z,Q^2)}{\sum_f e_f^2 q_f(x,Q^2)} \cdot \frac{d^2\sigma^{DIS}}{dxdQ^2}, \tag{7.7}$$

Year	$\mathcal{N}_{\rm DIS}(D)$	$\mathcal{N}/N(D)$	$\mathcal{N}_{\rm DIS}(H)$	$\mathcal{N}/N(H)$
1996	1 737 562	1.076	1 011 311	1.049
			803 145	1.038*
1997	2 058 629	1.102	1 697 564	1.033
			1 251 411	1.059*
2000	5 933 244	1.039*	5 536 272	1.066
2004	3 375 389	1.089	2 274 594	1.032*
2005	3 878 183	1.157	4 457 215	1.044*
	16 983 007	1.085	17 031 512	1.049

Table 7.3: Dead time corrected number of DIS events, $\mathcal{N}_{\rm DIS}$, for deuterium and hydrogen targets for different years. The asterisks denote polarized running. The ratio of dead time corrected numbers to measured numbers is also indicated in the table.

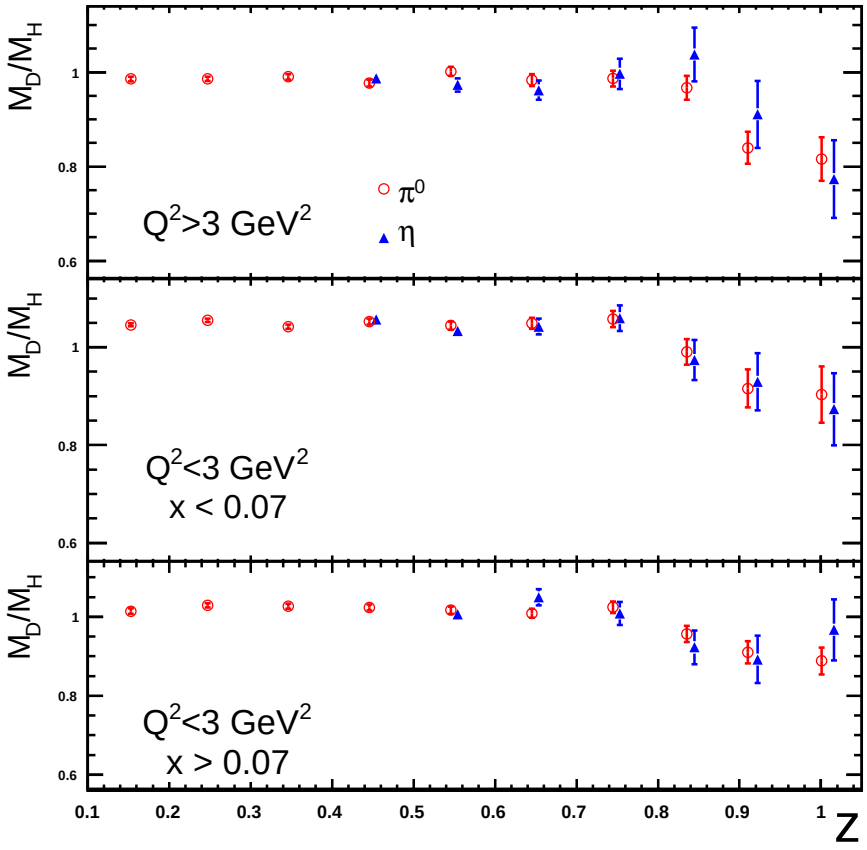

Figure 7.21: Multiplicity ratio for π^0 (circles) and η (triangles) for three different (Q^2, x) bins, the average kinematic values of the bins are mentioned in Table 7.5.

where, as before, $D_f^h(z, Q^2)$ is a fragmentation function, describing the production of a hadron of type h from a quark of flavor f, e_f^2 is the charge of the quark of flavor f and $q_f(x, Q^2)$ a quark distribution function. The sum runs over all quarks in the nucleon. For π^0 production

$< x >$	$< Q^2 >$	$\frac{M_D^{\pi^0}}{M_H^{\pi^0}}$	$\frac{\sigma_{\rm DIS}^D}{\sigma_{\rm DIS}^D}$	$\frac{M_D^{\pi^0}}{M_H^{\pi^0}} \cdot \frac{\sigma_{\rm DIS}^D}{\sigma_{\rm DIS}^D}$	$\frac{F_2^D}{F_2^H}$
0.16	4.64	0.99 ± 0.03	0.87 ± 0.01	0.86 ± 0.03	0.88 ± 0.01
0.048	1.51	1.04 ± 0.03	0.93 ± 0.01	0.97 ± 0.03	0.94 ± 0.01
0.095	2.23	1.02 ± 0.03	0.92 ± 0.01	0.94 ± 0.03	0.93 ± 0.01

Table 7.4: Measured multiplicity ratio for π^0 production, $\frac{M_D^{\pi^0}}{M_H^{\pi^0}}$, together with the $\frac{\sigma_{\rm DIS}^D}{\sigma_{\rm DIS}^H}$ ratio from Reference [193]. Column 5 is the product of column 3 and 4, and should be similar as the $\frac{F_2^D}{F_2^H}$ ratio. The latter ratio is given in the last column as determined from Reference [198]. The errors are statistical only.

we can assume that the following relations hold:

$$D_u^{\pi^0} = D_d^{\pi^0} = D_{\bar{u}}^{\pi^0} = D_{\bar{d}}^{\pi^0} \quad (7.8)$$

$$D_s^{\pi^0} = D_{\bar{s}}^{\pi^0}. \quad (7.9)$$

When we also assume SU(2) flavor symmetry for the parton distribution functions, we can write $u_p(x,Q^2) = d_n(x,Q^2)$, and $d_p(x,Q^2) = u_n(x,Q^2)$. If we additionally assume $s_p(x,Q^2) = s_n(x,Q^2)$, equation 7.7 can be written for a deuterium target as:

$$\frac{d^3\sigma^{\pi^0,D}}{dxdQ^2dz} = \frac{2\left(\left(\frac{4}{9}u_p(x,Q^2) + \frac{1}{9}d_p(x,Q^2)\right) D_u^{\pi^0}(z,Q^2) + \left(\frac{1}{9}s_p(x,Q^2)\right) \cdot D_s^{\pi^0}(z,Q^2)\right)}{2\left(\frac{4}{9}u_p(x,Q^2) + \frac{1}{9}d_p(x,Q^2) + \frac{1}{9}s_p(x,Q^2)\right)} \frac{d^2\sigma^{DIS,D}}{dxdQ^2} \quad (7.10)$$

The ratio of semi-inclusive π^0 production on a deuterium target over the one on a hydrogen target can now be written as:

$$\frac{\frac{d^2\sigma^{\pi^0,D}}{dxdQ^2}}{\frac{d^2\sigma^{\pi^0,H}}{dxdQ^2}} = \frac{\frac{d^2\sigma^{DIS,D}}{dxdQ^2}}{\frac{d^2\sigma^{DIS,H}}{dxdQ^2}} \sim \frac{F_2^D(x,Q^2)}{F_2^H(x,Q^2)}. \quad (7.11)$$

It can be seen that if the correct luminosity normalization is used, namely equation 7.6, the multiplicity ratio in the semi-inclusive region should be equal to one. The cross section ratio $\sigma_{\rm DIS}^D/\sigma_{\rm DIS}^D$ was calculated from the HERMES data in e.g. Reference [193]. The ratio $\frac{F_2^D(x,Q^2)}{F_2^H(x,Q^2)}$ was measured by the NMC collaboration at CERN [198] over a wide kinematic range.

The measured multiplicities in the semi-inclusive region, $0.1 < z < 0.4$, together with the DIS cross section ratio's from Reference [193] and the $\frac{F_2^D(x,Q^2)}{F_2^H(x,Q^2)}$ ratio from Reference [198], and the average x, and Q^2 values are compared in Table 7.4 for 3 bins in (Q^2, x). Within errors the results presented here correspond to the $\frac{F_2^D(x,Q^2)}{F_2^H(x,Q^2)}$ ratio from Reference [198]. This nicely shows the validity of the presented multiplicity ratio analysis and the cancellation of efficiency, acceptance and other systematic effects in the ratio.

Figure 7.21 shows the multiplicity ratio for both η and π^0 production for three kinematic regions, one corresponding to $Q^2 > 3$ GeV^2, one where $Q^2 < 3$ GeV^2 and $x < 0.07$, the third with where $Q^2 < 3$ GeV^2 and $x > 0.07$. Numerical values are given in Table 7.5. Neutron

	$< z >$	$< Q^2 >$	$< x >$	$< \epsilon >$	$M^D/M^H \pm \delta_{\text{stat}}$	$\frac{M^D}{M^H} \cdot \mathcal{R}(x,Q^2) \pm \delta_{\text{stat}}$
η	0.74	4.61	0.16	0.68	0.99 ± 0.03	0.86 ± 0.03
	0.83	4.64	0.16	0.70	1.03 ± 0.06	0.90 ± 0.05
	0.91	4.58	0.17	0.71	0.91 ± 0.07	0.79 ± 0.06
	1.00	4.44	0.16	0.71	0.77 ± 0.08	0.67 ± 0.07
	0.74	1.49	0.05	0.67	1.0 ± 0.09	0.93 ± 0.08
	0.83	1.47	0.05	0.70	0.95 ± 0.09	0.88 ± 0.08
	0.91	1.45	0.05	0.70	0.91 ± 0.06	0.85 ± 0.06
	1.00	1.46	0.05	0.72	0.87 ± 0.07	0.81 ± 0.07
	0.74	2.12	0.09	0.83	1.00 ± 0.03	0.91 ± 0.03
	0.83	2.09	0.10	0.84	0.92 ± 0.04	0.84 ± 0.04
	0.91	2.05	0.10	0.85	0.89 ± 0.06	0.81 ± 0.06
	1.01	1.98	0.10	0.86	0.96 ± 0.08	0.87 ± 0.07
π^0	0.74	4.70	0.18	0.74	0.98 ± 0.02	0.85 ± 0.02
	0.83	4.78	0.19	0.76	0.96 ± 0.03	0.84 ± 0.03
	0.91	4.75	0.20	0.78	0.83 ± 0.03	0.72 ± 0.03
	0.99	4.64	0.20	0.81	0.81 ± 0.05	0.70 ± 0.04
	0.74	1.42	0.05	0.73	1.05± 0.03	0.98 ± 0.02
	0.83	1.40	0.05	0.75	0.99 ± 0.03	0.92 ± 0.03
	0.90	1.39	0.05	0.77	0.91 ± 0.04	0.85 ± 0.04
	0.99	1.32	0.05	0.81	0.90 ± 0.06	0.84 ± 0.06
	0.74	2.06	0.10	0.86	1.02± 0.01	0.93 ± 0.01
	0.83	2.03	0.10	0.87	0.95 ± 0.02	0.86 ± 0.02
	0.91	1.98	0.10	0.88	0.91 ± 0.03	0.83 ± 0.03
	1.00	1.92	0.11	0.90	0.88 ± 0.03	0.80 ± 0.03

Table 7.5: Numerical values of the multiplicity ratio for the three highest z-bins for π^0 and η production on deuterium and hydrogen. Average x, Q^2 and ϵ in each bin are given. The error on the ratio is statistical only. The values in the last column are corrected for the ratio $\mathcal{R}(x,Q^2)$ from equation 7.6. Values for $\mathcal{R}$ were taken from Reference [193].

suppression seems to be strongest at high Q^2. For π^0 the ratio is the same within the statistical error in all 3 kinematic bins.

The neutron suppression is only predicted for the case of longitudinal polarized virtual photons. Longitudinal and transverse cross sections can not readily be separated at HERMES kinematics. Theoretically it is known that the σ_T/σ_L ratio follows a $1/Q^2$ behavior. At higher Q^2 the contamination from the transverse polarized virtual photons should thus decrease. The presented data agrees with this prediction, since the ratio seems to decrease at high Q^2.

The predicted ratio is smaller for π^0 production than for η production. From Figure 2.27 it can be seen that σ_n/σ_p is predicted to be 0.02 (0.6) at $\langle x \rangle = 0.20 (= 0.16)$, resulting in a M^D/M^H ratio of about 0.5 (0.8) respectively. The observed ratios presented here are of similar magnitude as the predicted values. However, the neutron suppression for π^0 and η production seems to be of similar magnitude, contrary to the calculations of Reference [63]. However, it should be noted that there may be a different amount of SIDIS background in the high z bin for both particles. A further analysis of π^0 to η production ratio would be very interesting, given the presented results in comparison with the predicted ones.

7.7 Conclusions

Various conclusions can be made from the analysis presented in this chapter. First of all, the comparison with the exclusive event generator based on the VGG model showed that at HERMES kinematics we are able to measure exclusive reactions, albeit with a broad M_X^2 distribution. The cross section predicted by the model is low and a comparison with experimental data showed a large non-exclusive background in the data sample. It was attempted to correct this background by using Monte Carlo generators which do not have cross sections for exclusive pseudoscaler processes implemented. However, data generated by these Monte Carlo generators overshoot the experimentally observed spectra. More work is needed to adapt those Monte Carlo generators to describe the non-exclusive background in the exclusive event sample. The Recoil Detector, however, is able to identify the recoiling proton, and consequently the background will be reduced by a large amount once data taken with that detector are ready for analysis. Moreover, the kinematics of the recoiling proton will be determined from the Silicon Detector directly, rather than being reconstructed from the forward spectrometer.

The analysis of the multiplicity ratio on hydrogen and deuterium of the pseudoscalar η and π^0 mesons showed clear evidence for the neutron suppression predicted in the exclusive regime. The systematic error in the determination of these ratio's was shown to be small. These observed effect shows that at HERMES kinematics exclusive production of π^0 (ánd η) mesons can be accessed.

CHAPTER 8

Summary, Conclusions and Outlook

This thesis discusses exclusive π^0 production at the HERMES experiment. After a theoretical introduction, the research and development for the hardware needed to measure exclusive reactions was detailed. Next, the propreties of the essential detector components of the HERMES spectrometer to measure exclusive π^0 production is described based on a Monte Carlo simulation. Finally, the analysis of exclusive π^0 production was discussed.

The research performed at the HERMES experiment, located at DESY in Hamburg, Germany, is aimed at the description and the understanding of the structure of nucleons. More precisely, it was built to measure the polarized structure functions $g_1(x, Q^2)$ and $g_2(x, Q^2)$ from polarized deep inelastic scattering. Form these polarized structure functions, the polarized quark distributions $\Delta q(x, Q^2)$ can be extracted. By doing so the contribution of the quarks to the spin of the nucleon can be estimated. These measurements detailedly confirmed results from other experiments that the quarks itself are responsible for only about $30\ \%$ of the nucleon spin. These observations showed how limited the understanding of the nucleon is. Rather than just the spin of the quarks, also the gluons and quark orbital angular momenta are expected to have a significant contribution to the nucleon spin. However, even if we would know the standard structure functions $F(x, Q^2)$ and $g(x, Q^2)$ with a very precise accuracy, the orbital angular momenta would still remain unknown, since these functions only contain one dimensional information. Only longitudinal information is embedded in these functions. Hence, another formalism is required in order to have a more complete image of the nucleon.

A promising formalism is found in the so-called 'Generalized Parton Distribution' functions. As described in Chapter 2 these are a generalization of the standard parton distribution functions in the sense that these functions are related to the non-forward matrix elements of the product of a local and non-local electromagnetic current operator, whereas the standard parton distribution functions correspond to the forward matrix element. Moreover, when integrated over x, it is seen that the GPDs are related to the conventional electromagnetic, axial and pseudovector structure functions. GPDs thus provide a unification mechanism for parton distribution functions and form factors.

By definition GPDs do not have a direct probabilistic interpretation. Therefore these functions may seem complex at first glance. However, a more intuitive picture can be obtained when moving to the impact parameter space. After the Fourier transformation to this space, it is seen that the GPDs contain both longitudinal momentum information and transverse position information. From this image the relation to orbital angular momenta is clear, since the latter is the product of a momentum and a position operator.

After introducing the GPDs, there may seem to be a lot of distribution functions required in order to obtain information about the nucleon. There are at least 4 GPDs for every quark flavor. However, when counting the polarized and unpolarized quark distribution functions, together with the form factors, it can be seen that the GPD formalism introduces less functions than before.

One method of accessing the GPDs is via the measurement of hard exclusive reactions. The factorization theorem proves that these processes can be written as a convolution of a hard scattering part and a soft part described by GPDs. In the case of meson production, an additional meson distribution amplitude enters in the factorization. The main physics goal of the HERMES experiment for the last two years of operation of the HERA accelerator facility is the verification of the predictions for Deeply Virtual Compton Scattering and hard exclusive meson production made by GPD models.

Hard exclusive vector meson production probes the GPDs $H(x,\xi,t)$ and $E(x,\xi,t)$, whereas pseudoscalar production accesses $\tilde{H}(x,\xi,t)$ and $\tilde{E}(x,\xi,t)$. The total spin of the meson produced in hard exclusive electroproduction thus acts as a GPD filter. The GPDs $\tilde{H}$ and $\tilde{E}$ are in the forward limit related to the spin dependent quark distributions Δq. In other words, via the formalism of GPDs polarized distribution functions can be accessed without the need of a polarized target by analyzing hard exclusive pseudoscalar meson production.

Unlike the GPD $\tilde{H}$, $\tilde{E}$ is not constrained in the forward limit and therefeore is more difficult to model. In modelling the GPD $\tilde{E}(x,\xi,t)$, accessible via pseusoscalar meson production, the pion pole plays an important role. The pole contribution is absent in the case of hard exclusive π^0 production, while it is not in the π^+ case. Hard exclusive π^+ electroproduction has already been observed at HERMES . The reduced cross section measured at HERMES follows the predicted $1/Q^2$ behavior. By comparing the production rate of π^+ mesons to the rate of π^0 mesons in similar experimental conditions, constraints on the relative importance of the pion pole can be imposed.

Because of the promising properties of GPDs for the understanding of nucleon structure, the HERMES collaboration decided to install a recoil detector surrounding the target cell. In Chapter 3 discusses the design requirements and the conceptual design of the recoil detector were discussed. Futhermore, a detailed description of the silicon recoil detector was presented.

The silicon recoil detector consists of 2×4 modules organised in a diamond shape around the target cell and is designed to measure the momentum of recoiling protons from their energy deposition in the silicon. Because recoiling protons have small momenta, large energy depositions are expected. This requires a readout chip for the detector which is capable of processing large input signals, which is a non-trivial demand since in most high-energy physics experiments silicon detectors are used as tracking detectors. Consequently, most of the readout chips are designed to amplify MIP signals. A too large input signal causes the preamplifier of these chips to be saturated.

The linearity and signal to noise ratio of two readout chips already being used at HERMES

and by the other experiments at the HERA facility, the APC and the HELIX chips, were determined. The tests were performed using a charge injection system. The APC chip showed a sufficiently large linearity for the project. In combination with a charge dividing capacitor, also the HELIX chip showed a linear behavior over a sufficiently large range. The charge division mechanism splits the output of a strip into a high gain and low gain HELIX input channel. Charge injection measurements showed that MIPs can be measured with a signal to noise ratio of about 10 by using the charge division setup. The APC only had a signal to noise ratio of about 3 for MIPs. Therefore the HELIX chip was chosen as the readout chip for the project.

The HELIX chip has a number of programmable parameters to adjust the behavior of the chip. The most important parameters are the currents sent to the preamplifier and the shaper, and the value of the feedback resistor for both preamplifier and shaper. These parameters determine the timing of the chip and consequently influence the signal to noise ratio. In order to be able to make a justified selection of these parameters, a dedicated charge injection setup was built with a very precise timing. In a fully automated setup, the output of the HELIX shaper could be digitized for every possible combination of the HELIX parameters. The setup was used to minimize the shaping time difference between high and low gain channel.

In order to test the full functionality of a silicon module a laser test stand was built. The setup makes use of a red light laser in combination with a microfocus lense. With the test stand it is possible to test the response of a particular strip to the injected light. Essential propreties as cross talk, linearity, and the amplification ratio between high and low gain channels can be tested as well.

The response to minimum ionizing particles of a Silicon Recoil Detector module was tested in a dedicated test beem setup at DESY Hamburg. The setup uses a beam telescope built out of 6 silicon sensors with a very high efficiency and an accurate position resolution. The readout of these 6 reference detectors was combined with the readout of the SRD module in one DAQ program. The setup has the advantage that a strip to strip efficiency and a signal to noise can be extracted. Moreover, the position reconstruction capability of the SRD could be tested by making use of the information of the beam telescope.

In the absence of a Recoil Detector, however, the preshower and calorimeter are the only detector components of the HERMES spectrometer sensitive to photons coming from π^0 decay. Consequently, these components are of vital importance in the analysis of π^0 and DVCS reactions. The calorimeter is responsible for the energy measurement of the photons coming from DVCS of from π^0 decay. Moreover, it gives a second space coordinate, apart from the event vertex, both being mandatory to determine the track of a photon. The performance of the calorimeter with respect to energy and position reconstruction can only be studied by using Monte Carlo techniques in lack of dedicated test beam results.

The calorimeter was originally built to provide a first level trigger and to perform electron-hadron separation. The energy of charged particles is measured by their deflection in the magnetic field. For the analysis of π^0 or hard γ production via the missing mass technique, an accurate description of the functionality of the calorimeter is indispensable. Therefore, the performance of the calorimeter was determined by an improved Monte Carlo simulation.

First of all, the simulation of the preshower was checked. It turned out that the thickness of the lead-iron sandwich was wrong in the geometry file required for the HERMES Monte Carlo. After implementing the correct thickness, a good agreement between data and Monte

Carlo was established for the signal of leptons, photons, and hadrons in the scintillator array behind the preshower.

The simulation of the response of the electromagnetic calorimeter of the HERMES experiment is set up in two steps: first, the electromagnetic shower is generated using the GEANT3 package. Secondly, the conversion from Čerenkov photons to photo-electrons is calculated by interpolating a dedicated Look-Up Table. The attenuation and reflections the Čerenkov photons in the shower are subject to before they reach the PMT, are accounted for in the Look-Up Table. The implementation of the GEANT3 package for simulating the electromagnetic shower was tested by comparing the generated shower profile with emperical predictions. Good concordance between the shower profile and the emperical predictions was obtained. A comparison between GEANT3 and a standalone GEANT4 based implementation of the HERMES electromagnetic calorimeter showed a good agreement. However, the GEANT4 results systematically showed larger shower fluctuations and a larger transverse shower profile than the GEANT3 based simulation.

Two algorithms were implemented to interpolate the 6-dimensional Look-Up Table. The interpolated values were compared with the tabulated numbers for high energetic muon data. Both algorithms gave a good interpolation of the look-up table.

The E/P ratio for electrons obtained with the improved Monte Carlo was compared to experimental data. A good agreement between Monte Carlo and experimental data was observed. Comparing to the results of the previous Monte Carlo implementation of the HERMES calorimeter, the new simulation showed an eminent performance improvement. The position reconstruction for electrons and the two-photon invariant mass spectrum with resulting π^0 peak were simulated and compared to experimental data as well. Also here good agreements were observed.

From a comparison of the reconstructed energy with the generated energy for single photon events, the performance of the HERMES calorimeter for photons could be determined. A distinction needs to be made between cases when the photon is showering in the preshower and when there is no showering in the preshower. Due to the calibration of the calorimeter, which is done using electrons, and due to the longer interaction length for photons, the attenuation of optical photons in the lead glass is overestimated. This effect yields larger reconstructed energies than the generated ones. A correction function was determined to correct the observed difference between generated and reconstructed energies.

Also the position reconstruction capabilities of the HERMES calorimeter were analyzed. The generated and reconstructed positions were compared at different z values. A bias in reconstructed x and y positions was obtained when comparing generated and reconstructed positions depending on the z position. The bias was removed for a z-position described by the empirical prediction of Reference [184]. It was also observed that the resolution in y is slightly worse than the resolution in x.

With the information obtained with the new calorimeter simulation, it is presumably also possible to use the observed resolution as an input to HSG, the smearing generator in the HERMES Monte Carlo, to provide a fast simulation of the response. Up to now, the implementation of the calorimeter in HSG strongly underestimates the experimental resolution of the calorimeter.

In the analysis chapter the reconstruction efficiency for π^0 and η mesons was determined. It was found that π^0 particles are reconstructed in HERMES with a high efficiency up to 12 GeV.

For η mesons, the efficiency remains constantly high over the entire momentum range.

A large dependence on the reconstructed energy by the calorimeter on the preshower pulse was observed. When correcting for this dependence, the energy reconstruction resolution is improved significantly. However, when applying the correction to the two-photon invariant mass spectrum, a bias in the π^0 peak position could be observed depending on the data production. This observation corresponds to a varying calorimeter or preshower calibration. An improved preshower calibration could be obtained by using the lepton and photon induced signal in the preshower. In fact, the preshower acts like a sampling calorimeter. Therefore also the z-position of a calorimeter cluster should be dependent on the preshower pulse. A parametrization of the z-position in terms of the prehower pulse would presumably improve the two-photon invariant mass spectrum by an additional factor.

When trying to analyze the exclusive π^0 production rate, a large non-exclusive background in the exclusive event sample was observed from comparisons with a Monte Carlo sample generated with an exclusive π^0 generator based on the VGG model. An attempt to correct the non-exclusive background in the event sample was made by using a semi-inclusive Monte Carlo sample. The comparison, however, showed that the Monte Carlo sample inadequatly describes the experimentally observed range. In order to analyze the exclusive π^0 cross section, either a better description of the observed spectrum by Monte Carlo is required or the non-exclusive background in the sample needs to be reduced. The first demand could be obtained by explicitely tuning the Monte Carlo in the specific kinematic range of the exclusive data-set, i.e. the high $z(=\frac{E_h}{\nu})$ range. This is non-trivial, since most semi-inclusive generators are only implemented to function in the range $0.2 < z < 0.8$. The second demand could be fulfilled by identifying the recoiling proton directly. This will be possible with a fully functional recoil detector.

As a test for the factorization theorem for exclusive meson production, the multiplicity ratio of both π^0 and η pseudoscalar mesons on hydrogen and deuterium targets was investigated. The data clearly shows a z dependence of this ratio. This observation indicates a suppression of π^0 and η production on the neutron at large z, since the deuterium nucleus is weakly bound. This behavior was predicted based on model calculations using the factorization theorem for exclusive reactions. reported on in this thesis. The ratio clearly drops as a function of z. The z-dependence shows a distinct transition from the semi-inclusive to the exclusive regime. For π^0 (η) production a deviation with a significance of 6σ (3.5σ) with respect to the semi-inclusive ratios was obtained. This observation forms evidence for the predicted neutron suppression in exclusive reactions.

The advantage of analyzing a production ratio is that the predictions are valid at lower Q^2 values. A dedicated future analysis of the ratio π^0/η as a function of z could also be a good test of model predictions. However, the difference in detection efficiency for both particles introduces an additional difficulty in the analysis, showing again the importance of the good Monte Carlo description of the calorimeter established in this work.

CHAPTER 9

Nederlandstalige Samenvatting

Deze thesis kadert in het onderzoek naar de bouwstenen van de materie. Algemeen geweten is dat alle materie opgebouwd is uit atomen die zelf bestaan uit een kern met daarrond een elektronenwolk. De atoomkern is opgebouwd uit neutronen en protonen herbergt bijna de volledige massa van een atoom. Typische afmetingen van de kern bedragen ongeveer $10^{-14} - 10^{-15}$ m. De straal van het ganse atoom daarentegen is ongeveer 10^{-10} m. Onnodig te zeggen dat de materiedichtheid in de kern zeer groot is. Als we daarbij in rekening brengen dat de protonen een positieve elektrische lading hebben, en wetende dat deeltjes met gelijke ladingen elkaar afstoten, kan het niet anders dan dat er een speciale kracht moet verantwoordelijk zijn om de atoomkern samen te houden. Die kracht wordt de 'sterke kracht' genoemd.

Een methode die vaak toegepast wordt om materie te onderzoeken is de te onderzoeken materie af te tasten met behulp van een probe. In de experimentele natuurkunde is die probe vaak een elektron, aangezien het een puntdeeltje is. Hoe meer energie het elektron heeft, hoe beter de resolutie, en hoe kleiner de afstanden die onderzocht kunnen worden. Verstrooiingsexperimenten eind de jaren '50 en begin jaren '60 toonden overduidelijk aan dat protonen en neutronen zelf geen puntdeeltjes zijn, maar bestaan uit nog andere deeltjes.

Gedurende de jaren '70 werd dit beeld uitgebreid. Betere experimentele technieken leidden tot hogere energieën en fijnere resoluties. Men kwam tot de constatatie dat protonen en neutronen zelf opgebouwd waren uit puntdeeltjes, quarks genoemd. Protonen bleken opgebouwd uit twee *up* en één *down* quark en neutronen uit twee *down* en één *up* quark.

Latere experimenten bij steeds hogere energie toonden aan dat het beeld van een proton opgebouwd uit drie quarks iets te eenvoudig is; als de schaal van de probe voldoende klein wordt kan men vaststellen dat er zich ook een groot aantal quark-antiquark paren bevinden in het proton. Deze worden voortdurend gecreëerd en geannihileerd in het veld van de sterke kracht. Bovendien is het nucleon niet enkel opgebouwd uit quarks en antiquarks, maar ook uit zogenaamde gluonen. Deze laatste zijn de overbrengers van de sterke kracht: tijdens een interactie tussen twee quarks wordt een gluon uitgewisseld. De gluonen interageren ook met andere gluonen.

Protonen en neutronen zijn niet de enige deeltjes die opgebouwd zijn uit quarks. Over het algemeen worden deeltjes opgebouwd uit quarks hadronen genoemd. Er bestaan deeltjes die gevormd worden door een quark-antiquark paar, de mesonen, en deeltjes die gevormd worden door drie quarks (of drie anti-quarks), de zogenaamde (anti-)baryonen. Mesonen en baryonen samen worden hadronen genoemd. Het bestuderen en beschrijven van hadronen aan de hand van hun bouwstenen is een onderdeel van de elementaire deeltjesfysica.

Eén van de meest fundamentele eigenschappen van deeltjes is hun spin. De fundamentele bouwstenen van de materie hebben halftallige spin: elektronen, muonen en quarks hebben allen spin $\frac{1}{2}$. Deeltjes met halftallige spin worden fermionen genoemd. Als twee deeltjes interageren met elkaar gebeurt dat door middel van uitwisseling van een zogenaamd ijkdeeltje. Bij de elektromagnetische kracht is dat ijkdeeltje een foton; bij de sterke kracht is het een gluon. Alle ijkdeeltjes hebben heeltallige spin; vandaar dat deze deeltjes ijkbosonen genoemd worden. Volgens het Standaard Model van de deeltjesfysica worden alle tot nog toe bekende krachten beschreven aan de hand van de uitwisseling van ijkbosonen tussen de fermionen.

Het proton zelf heeft ook halftallige spin. Eén van de grote uitdagingen van de theorie die de interactie tussen quarks door gluonen beschrijft is het verklaren van de spinstructuur van het proton aan de hand van de spin van de quarks. Spin is net zoals energie een grootheid die behouden wordt in een interactie: er kan geen spin verloren gaan of bijgecreëerd worden uit het vacuum. Het voorspellen en begrijpen van de spin van het proton aan de hand van de spin van zijn bouwstenen vormt een belangrijke test voor de theorie. Het toont hoe goed we in staat zijn om vanuit de beschikbare kennis bepaalde eigenschappen te voorspellen.

Het HERMES experiment, gesitueerd in DESY Hamburg, Duitsland is gebouwd om de spin structuur van protonen en neutronen te onderzoeken. Gepolariseerde elektronen worden verstrooid aan gepolariseerde trefkernen. Het verschil in het aantal waarnemingen tussen wanneer de elektronen en de trefkern dezelfde spin oriëntatie hebben, en de situatie waarbij de elektronenbundel en trefkern tegengestelde oriëntatie hebben verschaft ons informatie over de bijdrage van de quarks tot de spin van het proton. De metingen in HERMES bevestigden het resultaat van andere experimenten dat de spin van het proton niet eenvoudigweg op te bouwen is uit de spin van de quarks alleen. De quarks zelf dragen slechts ongeveer $30\ \%$ bij tot de spin van het proton. Men verwacht dat andere bijdragen tot de nucleon spin komen van het impulsmoment van de quarks en van de spin van de gluonen.

Het impulsmoment van quarks in het nucleon kan niet eenvoudig bepaald worden uit de standaard verstrooiingsexperimenten. In deze laatste wordt enkel het verstrooid elektron gemeten. Deze zogenaamde inclusieve reacties geven aanleiding tot gepolariseerde en ongepolariseerde parton distributie functies. Deze functies zijn bepaald voor elk quarktype en beschrijven de kans om bijvoorbeeld een quark van het type u te vinden in het nucleon. Het grote probleem echter is dat deze functies ééndimensionaal zijn: ze zijn enkel afhankelijk van de schaalvariabele x, die overeenkomt met de impulsfractie van het beschouwde quark. Om een impulsmoment te verkrijgen is tweedimensionale informatie noodzakelijk: deze grootheid is immers het vectorieel product van een positie- en een impulsoperator.

Een nieuw formalisme voor het beschrijven van quarks in het nucleon werd beschreven in de zogenaamde veralgemeende parton distributie functies, de GPD's. Deze functies zijn driedimensionaal. Er bestaan 4 GPD's voor elk type quark. De ééndimensionale partondistributiefuncties kunnen uit de GPD's verkregen worden door twee van de drie variabelen gelijk te stellen aan nul. Integreert men de x-afhankelijkheid van de GPD's weg, dan worden de

standaard vormfactoren bekomen. Die functies beschrijven de distributie van bijvoorbeeld elektrische en magnetische lading in het nucleon. Het GPD formalisme verenigt dus de traditionele ééndimensionale deeltjesdistributiefuncties met het formalisme der vormfactoren.

Experimenteel kunnen de GPD's gemeten worden aan de hand van exclusieve reacties. 'Gemeten' is een groot woord: experimenteel gaat men voorspellingen na die gebaseerd zijn op een model van de GPD's. Exclusieve reacties zijn reacties waarbij niet alleen het verstrooide elektron gedetecteerd wordt, maar ook alle andere reactieproducten. Via exclusieve productie van pseudoscalaire mesonen zoals de pionen en het η-meson kan informatie bekomen worden over de gepolariseerde quark distributies in het nucleon aan de hand van de relatie tussen de GPD's en de quark distributie functies.

Het type exclusieve reacties dat bestudeerd wordt in deze thesis is exclusieve π^0 productie: $e + p \rightarrow e' + p' + \pi^0$. Dit proces geeft toegang tot twee van de vier GPD's gedefinieerd voor elk type quark. De verhouding van exclusieve π^0 productie tot exclusieve π^+ productie bevat belangrijke informatie voor het modelleren van de GPD's. De werkzame doorsnede voor exclusieve π^+ productie werd reeds gemeten in HERMES en is in overeenstemming met modelberekeningen gebaseerd op het GPD formalisme. Het beschrijven van de exclusieve π^0 werkzame doorsnede vormt een serieuze uitdaging voor het model.

In het HERMES experiment wordt het terugstoot proton ontstaan in de exclusieve reactie tot nog toe echter niet gedetecteerd. Exclusieve reacties worden geselecteerd via de zogenaamde 'Missing Mass' techniek: als het verstrooid elektron en het geproduceerd meson gedetecteerd worden dan kunnen we exclusieve reacties selecteren door te eisen dat de niet-gedetecteerde massa de proton massa moet benaderen. Deze techniek werkt, maar zorgt voor een aanzienlijke achtergrond in de gegevens. Om die achtergrond te verminderen werd beslist om de HERMES spectrometer te voorzien van een recoil detector.

De recoil detector zelf bestaat uit verschillende sub-detectoren om enerzijds rechtstreeks het terugstoot proton te meten, en anderzijds om niet-exclusieve achtergrond in het event-sample te onderdrukken. Het impuls van het terugstootproton wordt bepaald door de energie depositie in een silicium detector. Wegens de lage impulsen van de terugstoot protonen moet de silicium detector in vacuum geplaatst worden.

Een deel van deze thesis beschrijft de initiële testen in de ontwikkelingsfase van de silicium detector. Een zeer belangrijke beslissing in het silicium project was het kiezen van type readout chip voor de detector. Via ladingsinjectie werd uit twee kandidaatchips, de APC en de HELIX, de meest geschikte chip geselecteerd. Wegens het beperkte gebied waarover de HELIX chip lineair het ingangssignaal versterkt, werd een ladingsdeling-mechanisme geïmplementeerd in de readout. Het analoge signaal afkomstig van de silicium sensor wordt hiertoe gesplitst in twee. Eén tak gaat naar de zogenaamde hoge versterkingschip, de andere tak wordt over een condensator naar de chip met lage versterking geleid. De zogenaamde hoge en lage versterkingschips zijn chips van hetzelfde type. De waarde van de condensator bepaalt de ladingsfractie die repectievelijk bij de chip met hoge en met lage versterking toekomt.

Omwille van het goed functioneren van de ladingsdeling tussen twee HELIX input kanalen werd geopteerd voor de HELIX chip. Hierdoor konden niet alleen grote signalen komende van terugstootprotonen verwerkt worden, maar ook signalen van zogenaamde MIPs, deeltjes met minimale energieafzetting in het silicium en dus een corresponderend klein signaal.

In het kader van het onderzoek verricht naar de optimale uitlezing van de siliciumdetector werd een speciale testopstelling gebouwd die het mogelijk maakte om het signaal van de

shaper in de HELIX chip te bekijken. Dat signaal kan gevarieerd worden aan de hand van programmeerbare parameters van de chip. De opstelling werkte volledig automatisch en had een heel preciese timing.

Om de functionaliteit van een prototype van de siliciumdetector te testen werd een opstelling gebouwd die gebruikt maakt van een laser met een golflengte van 675 nm. Het licht van de laser diode werd via een lens met microfocus op het siliciumoppervlak geprojecteerd, waar de fotonen een conversie ondergaan naar elektron-gat paren en zo een signaal induceren in het silicium. De laserbundel werd voldoende gefocust om licht op iedere strip afzonderlijk te projecteren. Zodoende kon elk uitleeskanaal van een module getest worden. De laser test was echter meer dan enkel een functionaliteitstest; ook de koppeling van het signaal naar naburige kanalen en de lineariteit van de respons kon getest worden met behulp van de opstelling.

Een derde testopstelling maakte gebruik van een testbundel in DESY Hamburg. Het gedrag van een module werd getest in een experimentele omgeving. De resulterende signaal-tot-ruis verhouding voor MIP deeltjes werd bepaald. Dankzij integratie van een bundeltelescoop in de opstelling kon de efficientie per strip eveneens vastgesteld worden.

Om exclusieve π^0 productie te bestuderen bleek het nodig te zijn het gedrag van de elektromagnetische calorimeter in de HERMES spectrometer goed te begrijpen. De enig mogelijke manier om de functionaliteit van de calorimeter na te gaan was met behulp van simulatietechnieken. Vooraleer de simulatie van de calorimeter zelf te beschouwen, werd eerst geverifieerd of de simulatie van de preshower adequaat was. Deze preshower is een sandwich van lood en staal die de calorimeter voorafgaat. Ongeveer 80 % van de fotonen interageert in de preshower. Als de simulatie van de preshower niet optimaal is, wordt de Monte Carlo beschrijving van de calorimeter wel heel moeilijk. Het bleek dat de dikte van de lood-staal sandwich verkeerd geïmplementeerd was in de simulatie. Na het corrigeren van deze fout werd een goede overeenstemming tussen experimentele data en het gesimuleerd signaal in de preshower gevonden.

De volgende stap was een verbetering van de bestaande simulatie van de calorimeter zelf. Reeds geruime tijd was duidelijk dat die simulatie het experimenteel spectrum niet correct beschreef. Daarom werd een nieuwe Monte Carlo beschrijving van de calorimeter geïmplementeerd. De beschrijving maakt enerzijds gebruik van het GEANT3 pakket om de elektromagnetische shower zelf te beschrijven. Anderszijds wordt gebruik gemaakt van een tabel die de conversie van Čerenkov fotonen naar foto-elektronen beschrijft. De tabel brengt de attenuatie en reflecties van de Čerenkov fotonen in rekening vooraleer deze de PMT bereiken. Twee algoritmes werden geïmplementeerd om de zesdimensionale tabel te interpoleren.

Er werd geverifieerd dat de nieuwe simulatie de energie afhankelijkheid van de E/P verhouding voor elektronen beschreef, alsook de positie reconstructie voor deze deeltjes. In beide gevallen werd een goede overeenkomst tussen simulatie en experimentele data waargenomen. Ook het experimentele twee-foton invariante massa spectrum, met resulterende π^0 piek, werd goed beschreven door de nieuwe simulatie.

Eenmaal de simulatie uitvoerig getest en voldoende goed geacht werd, kon de prestatie van de elektromagnetische calorimeter ten aanzien van de energiemeting van fotonen bepaald worden. Dit gebeurde door gegenereerde en gereconstrueerde data te vergelijken. Er bleek dat een onderscheid gemaakt moet worden tussen de situaties waar het foton al dan niet interageert in de lood-staal sandwich die de calorimeter voorafgaat. De gereconstrueerde energieën bleken telkens groter te zijn dan de gegenereede energieën. Dit effect kan verklaard

worden door het feit dat de calorimeter gecalibreerd wordt aan de hand van de E/P verhouding van elektronen, die gelijk moet zijn aan 1. Fotonen, echter, hebben een grotere vrije weglengte in de calorimeter, waardoor zij pas later zullen interageren. De attenuatie van de Čerenkov fotonen wordt dus overschat waardoor een relatief groter signaal ontstaat. Om dit effect te corrigeren werd een correctie functie gefit aan de gereconstrueerde energie.

Ook de positie reconstructie van de calorimeter kon getest worden met behulp van de nieuwe Monte Carlo. Er bleek duidelijk dat er een bias ontstaat in zowel de x- als y- positie reconstructie als een foutieve waarde voor z gekozen wordt. Als de parametrisatie van het longitudinaal shower profiel uit Referentie [184] gebruikt werd, bleek een bias-vrije reconstructie mogelijk voor zowel x- als y- coördinaat van de calorimeter.

Uit de nieuwe Monte Carlo simulatie bleek ook duidelijk dat de verhouding van de gegenereerde tot de gereconstrueerde energie afhankelijk was van de grootte van de puls in de scintillator-rij die tussen de preshower sandwich en de calorimeter gepositioneerd is. Dit is niet zo verwonderlijk. Immers, als er een groot signaal in de scintillator-rij geobserveerd wordt, betekent dit dat de shower al voor een groot stuk ontwikkeld is en dat diezelfde shower dus niet zo ver het lood-glas van de calorimeter zal binnendringen. Omgekeerd zal een klein signaal in de scintillator-rij betekenen dat de shower zich ver in het lood-glas van de calorimeter ontwikkelt. Een correctie-functie werd bepaald die de afhankelijkheid van de energie gemeten in de calorimeter van het signaal in de scintillator-rij uitdrukt. Als die functie echter toegepast werd op de experimentele data, werd een jaar tot jaar variatie geobserveerd in de positie van de π^0 piek in het twee-foton invariante massa spectrum. Dit betekent dat, ofwel de calorimeter, ofwel de scintillator-rij niet goed gecalibreerd zijn (ofwel, uiteraard, beide). De mogelijkheid tot een betere calibratie van de scintillator-rij werd geopperd door gebruik te maken van het signaal die elektronen en fotonen achterlaten in die scintillator rij. Het huidig calibratie algoritme maakt enkel gebruik van het MIP signaal van hadronen. Bovendien zou het in rekening brengen van de correlatie tussen het signaal in de preshower en de z-positie tot een betere energie resolutie kunnen leiden.

In het GPD formalisme speelt het factorizatie theorema een belangrijke rol. Dat theorema stelt dat voor exclusieve reacties de werkzame doorsnede geschreven kan worden als het product van een GPD met een meson golffunctie en een term die het harde verstrooiingsproces beschrijft. Dit is een ander factorizatie theorema dan dat voor semi-inclusieve productie. In dit type reacties wordt naast het verstrooide elektron ook een geproduceerd meson gedetecteerd. Het factorizatie theorema voor semi-inclusieve reacties stelt dat de werkzame doorsnede essentieel geschreven kan worden als het product van de werkzame doorsnede voor de inclusieve verstrooiingsreactie met een fragmentatie functie D_q^h. Deze laatste functie beschrijft de productie van een hadron van het type h uit een quark van het type q. De variabele bij uitstek om de overgang van het semi-inclusieve naar het exclusieve regime te beschrijven is de variable $z = E_h/\nu$. Deze variabele geeft de verhouding van de energie van het hadron h tot de beschikbare energie weer. Als $z \to 1$ komen we duidelijk in het exclusieve regime. Anderszijds wordt het semi-inclusieve regime gedefinieerd als $0.2 < z < 0.7$.

Als test voor het factorizatie theorema voor exclusieve meson productie werd de productie verhouding van de pseudoscalaire π^0 en η mesonen op waterstof en deuterium trefkern onderzocht. De HERMES data bewijzen duidelijk een z- afhankelijkheid van deze verhouding. Bij hoge z blijken minder η en π^0 mesonen geproduceerd te worden aan een deuterium trefkern dan aan een waterstof trefkern. De observatie is een indicatie voor de onderdrukking van π^0

en η productie uit het neutron eens het exclusieve regime bereikt wordt. Dit gedrag werd reeds voorspeld in Referentie [63] gebaseerd op modelberekeningen die gebruik maken van het factorizatie theorema voor exclusieve reacties. De z- afhankelijkheid van de productieverhouding toont dus niet alleen duidelijk de overgang van het semi-inclusieve regime naar het exclusieve regime, het bevestigt ook kwalitatief de voorspellingen van bovenvermelde referentie. Voor π^0 (η) mesonen werd een afwijking van 6σ (3σ) gevonden te opzichte van het semi-inclusieve gebied. De waarneming gepresenteerd in deze thesis kan dus geïnterpreteerd worden als een bevestiging voor het factorizatie theorema voor GPDs.

In deze thesis werd ook gepoogd de werkzame doorsnede voor exclusieve π^0 productie uit de bestaande HERMES dataset te extraheren. Het betreft een dataset in afwezigheid van de recoil detector. Een exclusief event sample werd verkregen door het opleggen van een aantal voorwaarden aan de HERMES data. Uit de vergelijking van het experimentele sample met de voorspellingen van een model voor exclusieve π^0 productie blijkt duidelijk dat het exclusieve event-sample een grote niet-exclusieve achtergrond bevat. Er werd dan ook gepoogd die niet-exclusieve achtergrond te beschrijven aan de hand van een semi-inclusief Monte Carlo sample. Jammer genoeg bleek dat het semi-inclusief Monte Carlo sample het geobserveerde spectrum niet adequaat beschrijft. Het reproduceren van de niet-exclusieve achtergrond vormt een uitdaging voor de semi-inclusieve Monte Carlo generatoren aangezien de meeste auteurs slechts een functionaliteit van hun Monte Carlo generator tot $z \sim 0.85$ garanderen. Het aanpassen van de Monte Carlo om het experimentele spectrum te beschrijven is dan ook een taak die buiten het tijdskader van deze thesis valt. Daarbij komt nog het feit dat met een recoil detector de niet-exclusieve achtergrond aanzienlijk gereduceerd wordt. De recoil detector is operationeel sinds juli 2006. Het wordt dan ook uitkijken naar analyses waarin data van de recoil detector geïncorporeerd worden.

APPENDIX A

Dirac Spinors and Parity

A.1 Dirac Spinors

When applying the relativistic quantization[1], $p_\mu \rightarrow i\hbar\partial_\mu$ to the relativistic energy-momentum relation, $p^\mu p_\mu - m^2c^2 = 0$ the Klein-Gordon equation for a wave function ψ is obtained:

$$\Big(-\hbar^2\partial^\mu\partial_\mu - m^2c^2\Big)\psi = 0 \tag{A.1}$$

Dirac tried to factor the energy-momentum relation, in order to get a field equation which was of first order in the derivatives ∂_μ. From the ansatz $p^\mu p_\mu - m^2c^2 = (\alpha^\kappa p_\kappa + mc)(\beta^\lambda p_\lambda - mc)$, the concept of γ-matrices, with the property $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ was introduced such that:

$$p^\mu p_\mu - m^2c^2 = (\gamma^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) = 0 \tag{A.2}$$

The Dirac equation is now defined as a solution of the equation above after substituting $p_\mu \rightarrow i\hbar\partial_\mu$:

$$\Big(i\hbar\gamma^\mu\partial_\mu - mc\Big)\psi = 0 \tag{A.3}$$

We can now consider plane wave solutions of the form:

$$\psi(x) \propto e^{-\frac{i}{\hbar}x\cdot p}u(p),$$

with $u(p)$ a bispinor, i.e. a quantity consisting of two spinors. In general a spinor can be viewed as a two-component object. The proposed solution for ψ can be inserted in equation A.3, yielding:

$$\big(\gamma^\mu p_\mu - mc\big)u = 0 \tag{A.4}$$

[1]Note that in this appendix, unlike elsewhere in this thesis c and $\hbar$ are not assumed to be equal to one

This is the so-called momentum-space Dirac equation which has 4 independent solutions: two of them correspond to positive energies, or *particle* states; the other two correspond to negative energies or *antiparticle* states. The physical interpretation of the two independent solutions for the (anti)particle states is given by the spin. The Dirac equation thus describes particles with spin $\frac{1}{2}$ together with their anti-particles.

A.2 Parity

The operator matrix element between two states $\langle\psi|$ and $|\psi\rangle$ yields expressions of the form $\psi_i^*\psi_j$. It is common practice to write the product $\psi_i^*\psi_j$ as a combination of quantities with particular transformation properties under the parity transformation.

In general, the Parity operator $\mathcal{P}$ denotes inversion of the coordinate system, i.e. it takes a point mirror image of an object. A vector $\vec{a}$'s image under such a transformation will point in the opposite direction: $\mathcal{P}(\vec{a}) = -\vec{a}$. Because of the $-$ sign the object is called a vector. Taking now the vector product of two vectors under parity tranformation: $\mathcal{P}(\vec{a} \times \vec{b}) = \vec{a} \times \vec{b}$ we see that it does not transform like a vector. The quantity $\vec{a} \times \vec{b}$ is said to be an axial- (or pseudo-) vector quantity. Also the scalar product of two vectors $\vec{a}$ and $\vec{b}$ does not change sign under parity transformation, whereas the triple product $\vec{c} \cdot (\vec{a} \times \vec{b})$ does. Therefore the latter quantity is referred to as a pseudoscalar quantity.

The Dirac spinors transform under parity like $\psi' \rightarrow \gamma^0\psi$. When defining $\bar{\psi} = \psi^\dagger\gamma^0$ it follows:

$$(\bar{\psi}\psi)' = \psi^\dagger\gamma^0\gamma^0\gamma^0\psi = \bar{\psi}\psi \tag{A.5}$$

$$(\bar{\psi}\gamma^5\psi)' = \psi^\dagger\gamma^0\gamma^0\gamma^5\gamma^0\psi = -\psi^\dagger\gamma^0\gamma^5\psi = -\bar{\psi}\gamma^5\psi \tag{A.6}$$

$$(\bar{\psi}\gamma^\mu\psi)' = \psi^\dagger\gamma^0\gamma^0\gamma^\mu\gamma^0\psi = -\psi^\dagger\gamma^0\gamma^\mu\psi = -\bar{\psi}\gamma^\mu\psi \tag{A.7}$$

$$(\bar{\psi}\gamma^\mu\gamma^5\psi)' = \psi^\dagger\gamma^0\gamma^0\gamma^\mu\gamma^5\gamma^0\psi = \psi^\dagger\gamma^0\gamma^\mu\gamma^5\psi = \bar{\psi}\gamma^\mu\gamma^5\psi \tag{A.8}$$

Equations A.5, A.6, A.7, A.8 respectively correspond to a scalar, a pseudoscalar, a vector and a pseudovector. Since a scalar and pseudoscalar contain one component each and a vector and pseudovector four components each of the product $\psi_i^*\psi_j$, an additional 6 are required to come to a general expression of $\psi_i^*\psi_j$ in terms of quantities with distinct parity properties. This is given by the antisymmetric tensor:

$$(\bar{\psi}\sigma^{\mu\nu}\psi)' = (\bar{\psi}\frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\psi)' = \psi^\dagger\frac{i}{2}\gamma^0\gamma^0(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\gamma^0\psi = \bar{\psi}\sigma^{\mu\nu}\psi \tag{A.9}$$

We have now constructed a full base for all 4×4 matrices in terms of $(1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5, \sigma^{\mu\nu})$.

APPENDIX B

Light Cone Coordinates

In the field theoretical description of quarks and gluons one often uses light cone coordinates. The use of these coordinates has several advantages as already pointed out by Dirac in 1949 [199]. He introduced the light front for the description of relativistic dynamics. The advantage of the use of light cone coordinates can be seen from the solutions of the general wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - v^2\frac{\partial^2}{\partial x^2}\right)\Psi(x,t) = 0. \tag{B.1}$$

If we introduce the coordinate transformation $\xi = t + \frac{x}{v}$ and $\eta = t - \frac{x}{v}$, the wave equation can be written as:

$$\frac{\partial^2}{\partial\xi\partial\eta}\Psi(\xi,\eta) = 0 \tag{B.2}$$

A general solution to this differential equation is: $\Psi(\xi,\eta) = f(t+\frac{x}{v}) + f(t-\frac{x}{v})$, which physically means that the wave front is traveling parallel to the line $t = \pm\frac{x}{v}$ in the $t-x$ space. The light front variables $x^{\pm}$ correspond to $v = c$. In the case of non-relativistic dynamics, v is much smaller than c, so the wave will travel along the $t-$axis, meaning that its front will travel parallel to the $x-$axis. If now, contrarily, the wave is travelling at the speed of light, the wave front will travel along the line $x^+ = t + x = C$ or along the line $x^- = t - x = C'$. In this case it is more convenient to describe the time-evolution of the system with the light front time, chosen to be x^+.

In four dimensions, the translation from traditional equal time (ET) coordinates (x^0, x^1, x^2, x^3) to light front (LF) coordinates (x^+, x^-, x^1, x^2) is done in the following way:

$$\begin{pmatrix} x^+ \\ x^- \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x^0 \\ x^3 \end{pmatrix}, \tag{B.3}$$

where the transverse components $\vec{x}_\perp = (x^1, x^2)$ are kept untouched. It is important to realize that time is expressed as $x^+ = \frac{x^0+x^3}{\sqrt{2}}$ rather than x^0 in ET coordinates. Moreover, the scalar

product needs to be written as:

$$v \cdot w = v^0 w^0 - \vec{v} \cdot \vec{w} = v^+ w^- + v^- w^+ - \vec{v}_\perp \cdot \vec{w}_\perp \tag{B.4}$$

The metric tensor thus becomes:

$$g^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{B.5}$$

Additionally, we can define the light cone γ matrices: $\gamma^\pm = \frac{1}{\sqrt{2}}(\gamma^0 \pm \gamma^3)$ and the projection operators $\Lambda^{+-} = \frac{1}{2}\gamma^{-(+)}\gamma^{+(-)}$

Dirac already pointed out that Lorenz boosts in the light front description are kinematic generators of a dynamical system, meaning that they leave the quantization surface constant ($t = 0$ in the traditional frame and $x^+ = 0$ in the light front system), whereas in the equal time, Lorenz boosts are dynamical generators, which move states off the quantization surface.

An interesting way to see the advantage of light-cone coordinates is coming from Reference [200] in the framework of DIS: the struck quark have a velocity close to the speed of light, i.e. close to the light cone. If an evolution function needs to be evaluated, the quark wave function will have to be evaluated at two different times, while in the light-cone formalism, these two states are evaluated at the *same* light cone time. Thus no evolution is needed. This picture is indicated in Figure B.1

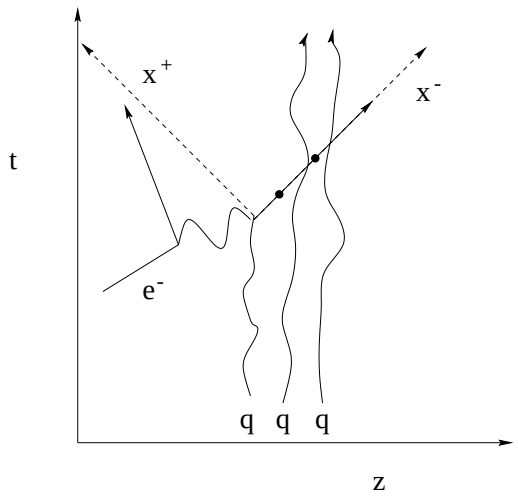

Figure B.1: Space-time graph of DIS. Figure from [200].

APPENDIX C

Common Mode Noise

The raw signal Q_i^k for each event i and strip k can be expressed as:

$$Q_i^k = S_i^k + P^k + N_i^k + C_i, \tag{C.1}$$

where P^k is the DC-offset of the channel, commonly known as pedestal level, S_i^k represents the channel's response to event i, N_i^k is the individual random noise contribution which is Gaussian distributed, and C_i is a random voltage offset, common for all channels of the chip, commonly called common mode noise.

In the absence of a signal S_i^k, the pedestal level of the device can be determined by averaging over all events:

$$P_k = \frac{1}{N_{events}} \sum_{i=1}^{N_{events}} Q_i^k. \tag{C.2}$$

The common mode noise needs to be calculated for each event individually:

$$C_i = \frac{1}{N_{channels}} \sum_{k=1}^{N_{channels}} (Q_i^k - P_k) \tag{C.3}$$

For a given event, the individual strip noise N_i^k can not be determined. Nevertheless, the average noise can be estimated evaluating the RMS of the pedestal distribution P_k:

$$N_{\text{rms}}^k = \sqrt{\frac{\sum_{i=1}^{N_{\text{events}}} \left(Q_i^k - C_i - P^k\right)^2}{N_{\text{events}}}} \tag{C.4}$$

Equation C.3 can be extended to the general case where $S_i^k \neq 0$:

$$C_i = \frac{1}{N_{channels}^*} \sum_{k=1}^{N_{channels}^*} (Q_i^k - P_k) \tag{C.5}$$

where the sum and corresponding averaging runs over the channels without hits ($N^*_{channels}$). In general a channel is said to have a hit when $S_i^k > P^k + 3 \cdot N^k_{\mathrm{rms}}$.

Acknowledgments

Having reached the end of this thesis I feel the urge to thank a number of people. First of all I am indebted to my promoter Prof. Dr. D. Ryckbosch for offering me a PhD position at the Gent University. I appreciated the trust he had during the entire PhD period, and especially enjoyed his confidence during my stay in Zeuthen. Whenever it was appropriate he always managed to pose interesting questions and give good suggestions related to the work. I also appraise the careful proof reading of this manuscript. His valuable comments, appropriate questions and relevant suggestions were valuable for improving the thesis.

During my stay at DESY Zeuthen I really enjoyed the cooperation with James Stewart who was the project leader of the silicon project. He gave enough responsibility to feel comfortable in the project and was always available for questions, discussions and comments. His way of working ensured a good motivation. I also want to thank Wolfgang Lange because he truly taught me the 'Art of Debugging'. I will never forget how he patiently looks for potential causes of malfunctioning electronics. I would moreover like to express gratitude to Helmut Böttcher and Wolf-Dieter Nowak for their help during my stay in Zeuthen whenever something was needed.

I very much enjoyed the company and discussions with my office and lab mates in Zeuthen, namely Misha Kopytin and Ivana Hristova. In addition, I enjoyed the cooperation with Matthias Reinecke during the development of the silicon project. Thanks also goes to Ingrid, Craig, Jos, Martin, Bernhard, Slava, Frank for their help and support.

Furthermore, I am indebted to Andy Miller for the collaboration in the development of the Monte Carlo simulation of the HERMES calorimeter, after all the 6 dimensional LUT was his idea. Even though due to the distance the communication was not always optimal, I enjoyed working together. His iron will to understand and solve problems, and his good suggestions and ideas were indispensable for the success of the calorimeter simulation. Besides his sense of humor made it pleasant to read his numerous emails.

Regarding the analysis, I owe a thank you very much to Delia Hasch and Cynthia Hadjidakis. They always were ready to answer questions and have constantly been very enthusiastic regarding exclusive physics processes and analysis.

Next, I would like to thank Elke Aschenauer for always having appropriate answers to questions I had. She usually manages to answer to the point and that is probably the reason that in nearly every HERMES thesis her help is acknowledged. Her devotion towards the HERMES experiment is respectable.

The (ex-)colleagues form the research group 'electromagnetic interactions' of Gent University: Peter Janssens, Arne Jr, Uli Elschenbroich, Freija Descamps, Gunar Schnell, Brecht Hommez, Yves Van Haarlem, Charlotte Van Hulse and Benidikt Zihlmann deserve a so long and thanks for all the fish. The first bunch of people made a good atmosphere at the lab, while Yves, Charlotte and Beni gave traveling to Hamburg an extra charm. Special thanks goes to my Gent office mates Michael Tytgat and Bino Maiheu. It was good fun sharing the office with those two. However, thanking Bino only for being a good office mate is an understatement: he was more like a 'Brother in Arms' during the PhD time. Thanks for sharing moments of frustration, joy, enthusiasm and pain both in Hamburg and Gent.

Thanks goes also to Jan, Luc and Klaas from our department for giving answers to questions that I raised upon them.

The social life in Berlin got boosted thanks to the foundation of the 'Stralauer Allee WG', therefore I greatly acknowledge Alejandro Lorca, Andrea Schindler and Lars Altwasser, my 'Mitbewohner' in '20b'. We had a great time there.

I also want to express gratitude to the numerous colleagues from the HERMES experiment for making a good atmosphere both on and offsite, answering many questions and for the nice evenings spent together. Special thanks goes to Josh for the most unplanned travel ever made.

These acknowledgments would not be complete without the mentioning of my friends from the 'G7': Bino, Sarah, Johanna, Lieven, Valery and Philippe. We all graduated in 2001, and ever since the friendship has not diminished in spite of the often large geographical separation.

Furthermore, I would like to thank the friends from Brugge: Jan, Nico, Poele, Koen, Pieter, Marie, Lieze, Eveline,.. it is great to have friends outside of the scientific community.

My parents deserve a huge thank you just for always being there.

I am beholden to Neptunus, the great god of sea, for always having me brought home safely. I guess we will meet at the equator.

Finally I want to thank Rudi Verspille for making the cover and binding this manuscript together. Thanks also goes to Alejandro and Freija for proof reading part of the thesis.

> There is one person who deserves a dedicated paragraph. For sharing precious moments and being such a lovely girl. For making me smile when she is around: Leentje !

Arne Vandenbroucke,
November 22, 2006

Bibliography

[1] D. Griffiths. *Introduction to Elementary Particles*. John Wiley and Sons, 1987.

[2] B. Povh et al. *Particles and Nuclei*. Springer-Verlag, 1999.

[3] E.B. Hughes et al. Neutron Form Factors from Inelastic Electron-Deuteron Scattering. *Physical Review*, 139B:458–471, 1965.

[4] T. Janssens et al. Proton Form Factors from Elastic Electron-Proton Scattering. *Physical Review*, 142:922–931, 1966.

[5] A. Bodek et al. Experimental Studies of the Neutron and Proton Electromagnetic Structure Functions. *Phys. Rev.*, D20:1471–1552, 1979.

[6] J. D. Bjorken. Asymptotic Sum Rules at Infinite Momentum. *Phys. Rev.*, 179:1547–1553, 1969.

[7] The Zeus Collaboration. Overview of Physics Results. *http://www-zeus.desy.de/publications.php3*.

[8] Durham HEP Databases. On-line Plotting and Calculation. *http://durpdg.dur.ac.uk/hepdata/pdf3.html*.

[9] M. Arneodo et al. A Re-Evaluation of the Gottfried Sum. *Phys. Rev.*, D50:1–3, 1994.

[10] K. Rith. Spin Asymmetries in Deep-Inelastic Electron-Nucleon Scattering: Selected HERMES Results. *Prog. Part. Nucl. Phys.*, 49:245–324, 2002.

[11] P. L. Anthony et al. Precision Measurement of the Proton and Deuteron Spin Structure Functions g_2 and Asymmetries A2. *Phys. Lett.*, B553:18–24, 2003.

[12] K. Abe et al. Measurement of the Proton and Deuteron Spin Structure Function g_1 in the Resonance Region. *Phys. Rev. Lett.*, 78:815–819, 1997.

[13] D. Adams et al. Spin Structure of the Proton from Polarized Inclusive Deep-Inelastic Muon-Proton Scattering. *Phys. Rev.*, D56:5330–5358, 1997.

[14] A. Airapetian et al. Precise Determination of the Spin Structure Function g_1 of the Proton, Deuteron and Neutron. in preparation, 2006.

[15] M. Anselmino et al. The Theory and Phenomenology of Polarized Deep Inelastic Scattering. *Phys. Rept.*, 261:1–124, 1995.

[16] J. Ashman et al. A Measurement of the Spin Asymmetry and Determination of the Structure Function g_1 in Deep Inelastic Muon-Proton Scattering. *Phys. Lett.*, B206:364, 1988.

[17] P. Z. Quintas et al. A Measurement of Lambda (MS) from Muon-Neutrino - Fe Non-singlet Structure Functions at the Fermilab Tevatron. *Phys. Rev. Lett.*, 71:1307–1310, 1993.

[18] Y. L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e^+-e^- Annihilation by Perturbation Theory In Quantum Chromodynamics. (in Russian). *Sov. Phys. JETP*, 46:641–653, 1977.

[19] V. N. Gribov and L. N. Lipatov. Deep Inelastic e-p Scattering in Perturbation Theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[20] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298, 1977.

[21] G. Altarelli. Partons in Quantum Chromodynamics. *Phys. Rept.*, 81:1, 1982.

[22] L. De Nardo. Measurement of the Structure Function g_1^d at HERMES and Extraction of Polarized Parton Distributions. *PhD Thesis, University of Alberta*, 2002. UMI-NQ-81178.

[23] Y. Goto et al. Polarized Parton Distribution Functions in the Nucleon. *Phys. Rev.*, D62:034017, 2000.

[24] M. Glück et al. Models for the Polarized Parton Distributions of the Nucleon. *Phys. Rev.*, D63:094005, 2001.

[25] J. Blumlein and H. Böttcher. QCD Analysis of Polarized Deep Inelastic Scattering Data and Parton Distributions. *Nucl. Phys.*, B636:225–263, 2002.

[26] R. L. Jaffe. Spin, Twist and Hadron Structure in Deep Inelastic Processes. *hep-ph/9602236*, 1996.

[27] U. Elschenbroich. Transverse Spin Structure of the Proton studied in Semi- Inclusive DIS. *PhD Thesis, Universiteit Gent*, 2006.

[28] A. Airapetian et al. Single-Spin Asymmetries in Semi-Inclusive Deep-Inelastic Scattering on a Transversely Polarized Hydrogen Target. *Phys. Rev. Lett.*, 94:012002, 2005.

[29] B. Andersson et al. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31, 1983.

[30] B. Maiheu. Hadronization in Electron-Proton Scattering at HERMES. *PhD Thesis, Universiteit Gent*, 2006.

[31] A. Airapetian et al. Quark Helicity Distributions in the Nucleon for Up, Down, and Strange Quarks from Semi-Inclusive Deep-Inelastic Scattering. *Phys. Rev.*, D71:012003, 2005.

[32] X.-D. Ji. Gauge Invariant Decomposition of Nucleon Spin. *Phys. Rev. Lett.*, 78:610–613, 1997.

[33] P. Hoodbhoy and X.-D. Ji. Helicity-Flip Off-Forward Parton Distributions of the Nucleon. *Phys. Rev.*, D58:054006, 1998.

[34] T. Feldmann. Generalized Parton Distributions. Talk at Internationale Universitätswochen für Theoretische Physik, Schladming, Austria, March 2006.

[35] X.-D. Ji. Off-Forward Parton Distributions. *J. Phys.*, G24:1181–1205, 1998.

[36] M. Diehl. Generalized Parton Distributions with Helicity Flip. *Eur. Phys. J.*, C19:485–492, 2001.

[37] M. Diehl and Ph. Hagler. Spin Densities in the Transverse Plane and Generalized Transversity Distributions. *Eur. Phys. J.*, C44:87–101, 2005.

[38] A. V. Efremov and A. V. Radyushkin. Asymptotical behavior of Pion Electromagnetic Form-Factor in QCD. *Theor. Math. Phys.*, 42:97–110, 1980.

[39] G. P. Lepage and S. J. Brodsky. Exclusive Processes in Quantum Chromodynamics: Evolution Equations for Hadronic Wave Functions and the Form-Factors of Mesons. *Phys. Lett.*, B87:359–365, 1979.

[40] M. Diehl. Generalized Parton Distributions. *Phys. Rept.*, 388:41–277, 2003.

[41] A. V. Belitsky and A. V. Radyushkin. Unraveling Hadron Structure with Generalized Parton Distributions. *Phys. Rept.*, 418:1–387, 2005.

[42] M. Burkardt. Impact Parameter Space Interpretation for Generalized Parton Distributions. *Int. J. Mod. Phys.*, A18:173–208, 2003.

[43] A. V. Belitsky and D. Mueller. Nucleon Hologram with Exclusive Leptoproduction. *Nucl. Phys.*, A711:118–126, 2002.

[44] S. Chekanov et al. Measurement of Deeply Virtual Compton Scattering at HERA. *Phys. Lett.*, B573:46–62, 2003.

[45] A. V. Belitsky et al. Theory of Deeply Virtual Compton Scattering on the Nucleon. *Nucl. Phys.*, B629:323–392, 2002.

[46] The Hermes Collaboration. Overview of Physics Results. *http://www-hermes.desy.de/notes/pub/trans-public-subject.html*.

[47] F. Ellinghaus. Beam-Charge and Beam-Spin Azimuthal Asymmetries in Deeply-Virtual Compton Scattering. *PhD Thesis, Humboldt Universität, Berlin*, 2004.

[48] F. Ellinghaus et al. Can the Total Angular Momentum of u-Quarks in the Nucleon be Accessed at HERMES? *Eur. Phys. J.*, C46:729–739, 2006.

[49] J. C. Collins et al. Factorization for Hard Exclusive Electroproduction of Mesons in QCD. *Phys. Rev.*, D56:2982–3006, 1997.

[50] M. Tytgat. Diffractive Production of ρ^0 and ω Vector Mesons at HERMES . *PhD Thesis, Universiteit Gent*, 2001.

[51] A. Airapetian et al. Exclusive Leptoproduction of ρ^0 Mesons from Hydrogen at Intermediate Virtual Photon Energies. *Eur. Phys. J.*, C17:389–398, 2000.

[52] J. Volmer. The Pion Charge Form Factor via Pion Electroproduction on the Proton. *PhD Thesis, Vrije Universiteit Amsterdam*, 2000.

[53] P. Brauel et al. Electroproduction of $\pi^+ n$, $\pi^- p$ and $K + \Lambda$, $K + \Sigma^0$ Final States above the Resonance Region. *Zeit. Phys.*, C3:101, 1979.

[54] K. Goeke et al. Hard Exclusive Reactions and the Structure of Hadrons. *Prog. Part. Nucl. Phys.*, 47:401–515, 2001.

[55] L. L. Frankfurt et al. N $\rightarrow$ Delta DVCS, Exclusive DIS Processes and Skewed Quark Distributions in Large N(c) Limit. *hep-ph/9808449*, 1998.

[56] L. L. Frankfurt et al. Hard Exclusive Pseudoscalar Meson Electroproduction and Spin Structure of a Nucleon. *Phys. Rev.*, D60:014010, 1999.

[57] I. Hristova. Background Correction for Exclusive π^+ Asymmetry. Talk at Collaboration Meeting, Hamburg, March 2006.

[58] M. Vanderhaeghen, P. A. M. Guichon, and M. Guidal. Deeply Virtual Electroproduction of Photons and Mesons on the Nucleon: Leading Order Amplitudes and Power Corrections. *Phys. Rev.*, D60:094017, 1999.

[59] L. Mankiewicz et al. Hard Exclusive Electroproduction of Pions. *Eur. Phys. J.*, C10:307–312, 1999.

[60] M. Vanderhaeghen et al. Hard Electroproduction of Photons and Mesons on the Nucleon. *Phys. Rev. Lett.*, 80:5064–5067, 1998.

[61] C. Hadjidakis. Exclusive Pion Production at HERMES . *Nucl. Phys.*, A755:557–560, 2005.

[62] L. Mankiewicz et al. Hard Exclusive Meson Production and Nonforward Parton Distributions. *Eur. Phys. J.*, C5:119–128, 1998.

[63] M. Eides et al. Hard Exclusive Electroproduction of Pseudoscalar Mesons and the QCD axial anomaly. *Phys. Rev.*, D59:114025, 1999.

[64] D. de Florian et al. Next-to-Leading Order Analysis of Inclusive and Semi- Inclusive Polarized Data. *Phys. Rev.*, D57:5803–5810, 1998.

[65] A. V. Radyushkin. Generalized Parton Distributions. *hep-ph/0101225*, 2000.

[66] M. Dohlus et al. Report from the HERA Taskforce on Luminosity Optimization: Theory and First Luminosity Scans. *DESY HERA 03-01*, 2003.

[67] A. A. Sokolov and I. M. Ternov. On Polarization and Spin Effects in the Theory of Synchrotron Radiation. *Phys. Dokl.*, 8:1203–1205, 1964.

[68] D. P. Barber et al. The HERA Polarimeter and the First Observation of Electron Spin Polarization at HERA. *Nucl. Instrum. Meth.*, A329:79–111, 1993.

[69] E. Gianfelice. Lepton Beam Polarisation in HERA with 3 Spin Rotators Pairs. Talk at Beschleuniger Betriebsseminar, Desy, Hamburg, 2004.

[70] L. H. Thomas. The Kinematics of an Electron with an Axis. *Phil. Mag.*, 3:1–21, 1927.

[71] M. Beckmann et al. The Longitudinal Polarimeter at HERA. *Nucl. Instrum. Meth.*, A479:334–348, 2002.

[72] J. Bohme. Polarisation Measurement at HERA II. *Acta Phys. Polon.*, B33:3949–3954, 2002.

[73] The Hermes Collaboration. The HERMES Technical Design Report. *Desy PRC*, 93-06, 1993.

[74] D. De Schepper et al. The HERMES Polarized ^{3}He Internal Gas Target. *Nucl. Instrum. Meth.*, A419:16–44, 1998.

[75] A. Airapetian et al. The HERMES Polarized Hydrogen and Deuterium Gas Target in the HERA Electron Storage Ring. *Nucl. Instrum. Meth.*, A540:68–101, 2005.

[76] A. Nass et al. The HERMES Polarized Atomic Beam Source. *Nucl. Instrum. Meth.*, A505:633–644, 2003.

[77] C. Baumgarten et al. The Storage Cell of the Polarized H/D Internal Gas Target of the HERMES Experiment at HERA. *Nucl. Instrum. Meth.*, A496:277–285, 2003.

[78] C. Baumgarten et al. A Gas Analyzer for the Internal Polarized Target of the HERMES Experiment. *Nucl. Instrum. Meth.*, A508:268–275, 2003.

[79] C. Baumgarten et al. An Atomic Beam Polarimeter to Measure the Nuclear Polarization in the HERMES Gaseous Polarized Hydrogen and Deuterium Target. *Nucl. Instrum. Meth.*, A482:606–618, 2002.

[80] K. Ackerstaff et al. The HERMES Spectrometer. *Nucl. Instrum. Meth.*, A417:230–265, 1998.

[81] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How To Approach*. Berlin, Germany: Springer, 1994.

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

[83] D. Green. The Physics of Particle Detectors. *Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 12:1–361, 2000.

[84] J. T. Brack et al. The HERMES Forward Tracking Chambers: Construction, Operation, and Aging Effects. *Nucl. Instrum. Meth.*, A469:47–54, 2001.

[85] S. Bernreuther et al. The HERMES Back Drift Chambers. *Nucl. Instrum. Meth.*, A416:45–58, 1998.

[86] A. Andreev et al. Multiwire Proportional Chambers in the HERMES Experiment. *Nucl. Instrum. Meth.*, A465:482–497, 2001.

[87] J. J. M. Steijger. The HERMES Silicon Project. *Nucl. Instrum. Meth.*, A447:55–60, 2000.

[88] M. G. van Beuzekom et al. First Experience with the HERMES Silicon Detector. *Nucl. Instrum. Meth.*, A461:247–250, 2001.

[89] H. Avakian et al. Performance of the Electromagnetic Calorimeter of the HERMES Experiment. *Nucl. Instrum. Meth.*, A417:69–78, 1998.

[90] H. Avakian et al. Performance of F101 Radiation Resistant Lead Glass Shower Counters. *Nucl. Instrum. Meth.*, A378:155–161, 1996.

[91] J. A. Appel et al. Performance of a Lead Glass Electromagnetic Shower Detector at FERMILAB. *Nucl. Instrum. Meth.*, 127:495, 1975.

[92] N. Akopov et al. The HERMES Dual-Radiator Ring Imaging Cerenkov Detector. *Nucl. Instrum. Meth.*, A479:511–530, 2002.

[93] E. Aschenauer et al. Optical Characterization of n = 1.03 Silica Aerogel used as Radiator in the RICH of HERMES. *Nucl. Instrum. Meth.*, A440:338–347, 2000.

[94] Y. Miyachi. The HERMES RICH Aerogel Radiator. *Nucl. Instrum. Meth.*, A502:222–226, 2003.

[95] B. Hommez. Hadron Identification with the HERMES RICH. *Nucl. Instrum. Meth.*, A502:294–299, 2003.

[96] M. Amarian et al. The HERMES Charm Upgrade Program: A Measurement of the Double Spin Asymmetry in Charm Leptoproduction. *Hermes Proposal*, 97-04, 1997.

[97] HERMES. HERMES First Level Trigger. *http://www-hermes.desy.de/groups/triggrp/*.

[98] T. Benisch et al. The Luminosity Monitor of the HERMES Experiment at DESY. *Nucl. Instrum. Meth.*, A471:314–324, 2001.

[99] M. G. van Beuzekom et al. The HERMES Silicon Project: The Radiation Protection System. *Nucl. Instrum. Meth.*, A512:44–51, 2003.

[100] CERN Computing and Networks Division. EPIO Manual. CERN *Program Library Long Writeup I101*, CERN , 1993.

[101] V. Shutov. HERMES DAQ Upgrade Status. Talk at Recoil Meeting, DESY, Hamburg, October 2004.

[102] H. Tanaka et al. A Gain Monitoring System with a Nd:YAG Laser for the Photomultipliers of the HERMES Experiment. *Nucl. Instrum. Meth.*, A515:725–732, 2003.

[103] W. Wander et al. DAD - Distributed ADAMO Database system at HERMES. Prepared for International Conference on Computing in High- energy Physics (CHEP 95), Rio de Janeiro, Brazil, 18-22 Sep 1995.

[104] CERN Programming Techniques Group. ADAMO Reference Manual. *CERN*, 1993.

[105] K. Ackerstaff et al. PinK: A TCL/TK Based Database Interface to ADAMO and DAD. Prepared for International Conference on Computing in High- energy Physics (CHEP 95), Rio de Janeiro, Brazil, 18-22 Sep 1995.

[106] R. Brun and D. Lienart. Hbook User Guide: CERN Computer Center Program Library Long Writeup: Version 4. 1998. CERN-Y250.

[107] W. Wander. Reconstruction of High Energy Scattering Events in the HERMES experiment. *PhD Thesis, University of Erlangen-Nürnberg*, 1996.

[108] G. Roeper. Messungen der Feldverteilung des HERMES Spektrometer Magneten. *Diplomarbeit, Universität Hamburg*, 1995.

[109] J. Ely. HERMES Calorimeter Position Reconstruction Study. *Hermes Internal*, 01-056, 2001.

[110] J. Ely. Measurement of the Single Spin Azimuthal Asymmetry in the Predominantly Exclusive Electroproduction of Photons from the Proton. *PhD Thesis, University of Colorado*, 2002.

[111] S. Schumann. Electron-Hadron Separation with the HERMES RICH Detector. *Hermes Summer Student Report*, 2001.

[112] J. Wendland. Particle Identification for HERMES Run I. *Hermes Internal*, 01-067, 2001.

[113] J. Wendland. Polarized Parton Distributions Measured at the HERMES Experiment. *PhD Thesis, Simon-Fraser University*, 2003.

[114] R. Kaiser. Particle Identification at HERMES. *Hermes Internal*, 97-25, 1997.

[115] J. Wendland. Improved Particle Identification at HERMES and Polarized Valence Quark Distributions in the Proton. *M.Sc. Thesis, Simon-Fraser University*, 1999.

[116] A. Airapetian et al. The Time-of-Flight Technique for the HERMES Experiment. *Nucl. Instrum. Meth.*, A540:305–310, 2005.

[117] The Hermes Collaboration. The HERMES Recoil Detector TDR. *Hermes Internal*, 02-003, 2002.

[118] M. Hoek. Design and Construction of a Scintillating Fibre Tracker for Measuring Hard Exclusive Reactions at HERMES. *PhD Thesis, Justus-Liebig Universität, Gießen*, 2006.

[119] A. Peisert. Silicon Microstrip Detectors. *Adv. Ser. Direct. High Energy Phys.*, 9:1–79, 1992.

[120] E. Gatti and P. F. Manfredi. Processing the Signals from Solid State Detectors in Elementary Particle Physics. *Riv. Nuovo Cim.*, 9N1:1–146, 1986.

[121] Particle Data Group. The Review of Particle Physics. *Eur. Phys. J.*, C, 2005.

[122] E. Switzer. Measurements of Electron Energy Deposition in the HERMES Silicon Recoil Detector. *Hermes Summer Student Report*, 2002.

[123] B. Kraus. Ein Silizium Detektor für das HERMES Experiment. *Diplom Thesis, Universität Erlangen-Nürnberg*, 2000.

[124] M. Kopytin et al. Decision on the Readout Chip for the New HERMES Silicon Recoil Detector. *Hermes Internal*, 02-020, 2002.

[125] B. Kraus. Deeply Virtual Compton Scattering and the HERMES Recoil-Detector. *PhD Thesis , Universität Erlangen-Nürnberg*, 2005.

[126] Z. Ye et al. Radio Frequency Test for the HERMES Silicon Recoil Detector. *Hermes Internal*, 05-014, 2005.

[127] R. Horisberger and D. Pitzl. A Novel Readout Chip for Silicon Strip Detectors with Analog Pipeline and Digitally Controlled Analog Signal Processing. *Nucl. Instrum. Meth.*, A326:92–99, 1993.

[128] M. Hilgers and R. Horisberger. Development of a Radiation Hard Version of the Analog Pipeline Chip APC128. *Nucl. Instrum. Meth.*, A481:556–565, 2002.

[129] M. Feuerstack-Raible et al. HELIX 128-x User's Manual Version 2.1. *HD-ASIC-33-0697, ASIC Labor Heidelberg*, 1999.

[130] M. Feuerstack-Raible. Overview of Microstrip Read-Out Chips. *Nucl. Instrum. Meth.*, A447:35–43, 2000.

[131] U. Trunk. Development and Characterisation of the Radiation Tolerant HELIX128-2 Readout Chip for the HERA-B Microstrip Detectors. *Ph.D. Thesis, Rupertus Carola University, Heidelberg*, 2000.

[132] J. Visser et al. Design and Performance of a Silicon Test Counter for HERMES. *Nucl. Instrum. Meth.*, A521:430–440, 2004.

[133] K. Fiedler. Identification of Spectator Nucleons at HERMES. *PhD Thesis,Universität Erlangen-Nürnberg*, 2001.

[134] J. J. M. Steijger. The Lambda Wheels, a Silicon Vertex Detector for HERMES. *Nucl. Instrum. Meth.*, A453:98–102, 2000.

[135] J. Burger et al. The Readout System of the new H1 Silicon Detectors. *Nucl. Instrum. Meth.*, A367:422–427, 1995.

[136] M. Kopytin. Longitudinal Target-Spin Azimuthal Asymmetry in Deeply-Virtual Compron Scattering. *PhD Thesis, Humboldt Universität, Berlin*, 2006.

[137] I.M. Gregor et al. HELIX128-3.0 Parameter Test for the HERMES Silicon Recoil Detector. *Hermes Internal*, in preparation, 2004 (draft).

[138] S. Balashov et al. Selection of the HELIX Parameters for the New HERMES Silicon Recoil Detector. *Hermes Internal*, in preparation, 2005 (draft).

[139] M. Kopytin et al. HELIX128-3.0 Production Tests. *Hermes Internal*, 05-002, 2005.

[140] B. Kraus. Silicon Sensor Tests. *internal information*, 2003.

[141] B. Kraus. Optimisation of the Flex-Foil Layering of the HERMES Recoil Detector. *Hermes Internal*, in preparation, 2003.

[142] Y. Tomita et al. Infrared Light Charge Injector as a Tool for the Study of Silicon Detectors. *Nucl. Instr. Meth.*, A270:403, 1988.

[143] J. Krizmanic et al. A 1060-nm Diode Laser System for Dynamically Probing Silicon Detectors. *Nucl. Instrum. Meth.*, A374:315–319, 1996.

[144] I. Abt et al. Characterization of Silicon Microstrip Detectors Using an Infrared Laser System. *Nucl. Instrum. Meth.*, A423:303–319, 1999.

[145] S. Shaheen et al. Characterization and Quality Control of Silicon Microstrip Detectors with an Infrared Diode Laser System. *Nucl. Instrum. Meth.*, A352:573–578, 1995.

[146] I.M. Gregor et al. A Laser Test-Stand for the New HERMES Recoil Silicon Detector. *Hermes Internal*, 04-016, 2004.

[147] D. Poelman and P.F. Smet. Methods for the Determination of the Optical Constants of Thin Films from Transmission Measurements: a Critical Review. *J. Phys.*, D36:1850–1857, 2003.

[148] B. Kraus. private communciation. 2004.

[149] M. Milite. The Internal Structure of Charmed Jets in Photoproduction at HERA and tests of the ZEUS Microvertex Detector Silicon Sensors. *Ph.D. Thesis, Universität Hamburg*, 2001.

[150] N. Pickert. What has Erlangen Been Up to Lately ? Talk at Recoil Meeting, DESY, Hamburg, September 2005.

[151] L Santamaria. Study on the Crosstalk Sources of the HERMES Silicon Recoil Detector at DESY. *Hermes Summer Student Report*, 2004.

[152] S. Balashov et al. Energy Calibration of the Silicon Modules for the Recoil Detector with the Tandem Facility. *Hermes Internal*, 05-020, 2005.

[153] L. A. T. Bauerdick et al. Beam Test of Silicon Strip Sensors for the ZEUS Micro Vertex Detector. *Nucl. Instrum. Meth.*, A501:340–358, 2003.

[154] M. Moritz. Measurement of the High Q^2 Neutral Current DIS Cross Section at HERA. *Ph.D. Thesis, Universität Hamburg*, 2002.

[155] I. Redondo Fernandez. Charm Electroproduction at HERA. *Ph. D. Thesis, Universidad Autonoma de Madrid*, 2001.

[156] M. Minty. General Overview of the DESY II and DESY III Synchrotrons. *DESY Internal*, 2003.

[157] N. Meyners. Test Beams at DESY . `http://adweb.desy.de/~testbeam`, 2004.

[158] IDE. The VA2. *Specification List*.

[159] CAEN. The V551. *Technical Information Manual*, 2000.

[160] CAEN. The V550. *Technical Information Manual*, 2000.

[161] M Morrisey. SEQSI. *Description and Manual*, 1996.

[162] I.M. Gregor et al. Test-beam Results of the First Prototype for the HERMES Silicon Recoil Detector. *Hermes Internal*, 02-053, 2002.

[163] I.M. Gregor et al. Study of the First Prototype for the HERMES Silicon Recoil Dtector with the ZEUS Beam Telescope. *Hermes Internal*, 04-019, 2004.

[164] C. Burgard. TESA. *A Test-Stand Analysis Framework*, 1999.

[165] I. Hristova et al. HERMES Recoil Silicon Detector Calibration to MIPs at T22 at DESY. *Hermes Internal*, 05-014, 2005.

[166] T. Sjostrand et al. PYTHIA 6.2: Physics and manual. 2001.

[167] G. Ingelman et al. LEPTO 6.5 - A Monte Carlo Generator for Deep Inelastic Lepton-Nucleon Scattering. *Comput. Phys. Commun.*, 101:108–134, 1997.

[168] T. Sjostrand. PYTHIA 5.7 and JETSET 7.4: Physics and manual. 1995.

[169] I. Akushevich et al. RADGEN 1.0: Monte Carlo Generator for Radiative Events in DIS on Polarized and Unpolarized Targets. 1998.

[170] CERN Application Software Group. GEANT , Detector Descriprion and Simulation Tool. *CERN Program Library Long Writeup*, W5013, 1994.

[171] F. M. Menden. Determination of the Gluon Polarization in the Nucleon. *Ph. D. Thesis, Universität Freiburg*, 2001.

[172] A. Hillenbrand. Measurement and Simulation of the Fragmentation Process at HERMES. *Ph. D. Thesis, Universität Erlangen-Nürnberg*, 2005.

[173] A. Kisselev. private communciation. 2005.

[174] B. Rossi. *High-Energy Particles, 4th Edition*. Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

[175] NIST. National Institute of Standards and Technology. *http://physics.nist.gov*.

[176] A. Fantoni. Costruzione del Calorimetro a Contatori di Vetro al Piombo dell'Esperimento HERMES per la Misura delle Funzioni di Struttura di Spin dei Nucleoni. *Ph. D. Thesis, Universita degli Studi della Calabria*, 1996.

[177] A. Miller. Calorimeter Simulation in HMC. Talk at Collaboration Meeting, Erlangen, October 2005.

[178] R. Ravotti et al. A Geometric Apporach to Maximum-Speed n-Dimensional Continuous Linear Interpolation in Rectangular Grids. *IEEE Trans. on Comp.*, 47:894–899, 1998.

[179] S Davies. Multidimensional Triangulation and Interpolation for Reinforcement Learning. *NIPS*, pages 1005–1011, 1996.

[180] T. Veldhuizen. Grid Filters for Local Nonlinear Image Restoration. *M. Sc Thesis, Dept. of Systems Design Engineering, University of Waterloo*, 1998.

[181] CERN Computing and Networks Division. Cernlib. CERN *Program Library Short Write-ups*, CERN, 1996.

[182] E. Longo and I. Sestili. Monte Carlo Calculation of Photon Initiated Electromagnetic Showers in Lead Glass. *Nucl. Instrum. Meth.*, 128:283, 1975.

[183] V.A. Hanin and E. Stern. Monte-Carlo Calculation of Energy Resolution of Lead Glass Detectors for High Energy Photons. *Nucl. Instrum. Meth.*, 157:455–460, 1978.

[184] U. Amaldi. Fluctuations in Calorimetry Measurements. *Phys. Scripta*, 23:409, 1981.

[185] J. Alvarez-Muniz et al. Comparative Study of Electromagnetic Shower Track Lengths in Water and Implications for Cerenkov Radio Emission. *Phys. Rev.*, D68:043001, 2003.

[186] S. Agostinelli et al. GEANT4: A Simulation Toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[187] HERMES. The uDST ddl Documentation Page. *http://www-hermes.desy.de/UDST_DOCU/ddl_main_page.html*.

[188] HERMES Data Quality group. Offline Data Quality - uDST Productions. *http://www-hermes.desy.de/groups/daqlgrp/OFFLINE_DQ/uDST/*.

[189] U. Elschenbroich et al. Single-Spin Azimuthal Asymmetries in SIDIS off a Transversely Polarized Proton Target - MC Studies. *Hermes Internal*, 03-007, 2003.

[190] HERMES Data Quality group. Luminosity constants of Different Years. *http://www-hermes.desy.de/groups/daqlgrp/OFFLINE_DQ/uDST/general/docu/DISstats.html*.

[191] U. Elschenbroich. Analysis of Luminosity Monitor Data for Different Years. *Hermes Internal*, 02-013, 2002.

[192] C. Hadjidakis et al. Total Cross Section for Hard Exclusive π^+ production. *Hermes Release Report*, 2004.

[193] J. Visser. Deep Inelastic Scattering off Hydrogen and Deuterium. *PhD Thesis, Rijksuniversiteit Groningen*, 2002.

[194] E. Avetysian. An attempt to Account the Preshower Pulse in the Calorimeter Energy measurement. Talk at Collaboration Meeting, Hamburg, December 2004.

[195] P. Liebing. Can the Gluon Polarization in the Nucleon be Extracted from HERMES Data on Single High-p_T Hadrons ? *Ph. D. Thesis, Universität Hamburg*, 2004.

[196] I. Hristova. Background Subtraction Studies for Exclusive π^+. Talk at Collaboration Meeting, Erlangen, October 2005.

[197] P. Di Nezza. Do You Believe in GPD's ? Talk at Collaboration Meeting, Hamburg, December 2001.

[198] M. Arneodo et al. Accurate measurement of F2(d)/F2(p) and R(d)-R(p). *Nucl. Phys.*, B487:3–26, 1997.

[199] Paul A. M. Dirac. Forms of Relativistic Dynamics. *Rev. Mod. Phys.*, 21:392–399, 1949.

[200] B. Klindworth. The $q-\bar{q}$ Potential in the Color-Dielectric Formulation of the Transverse Lattice. *PhD Thesis, New Mexico State University*, 1999.

Know a man
His face seems pulled and tense
Like he's riding on a motorbike in the
strongest winds
so I approach with tact
suggest that he should relax
but he's always moving much too fast
(· · ·)
He's yet to come back
But I've sees his picture
It doesn't look the same up on the rack
we go way back

I wonder bout his insides
Its like his thoughts are too big for his size
he's been taken, where I don't know
Off he goes
(· · ·)
Oh I rub my eyes
For he has returned
It seems my preconceptions are what should
have been burned
For he still smiles
And he's still strong
Nothing's changed
but the surrounding bullshit
that has grown

Oh now he's home
and we're laughing like we always did
my same old friend

Until a quarter to ten
I saw the strain creep in
He seems distracted
and I know just what is gonna happen next

Before his first step
he's off again

Pearl Jam