Albert-Ludwigs-Universität Freiburg

Fakultät für Mathematik und Physik

Dissertation

Christopher Ralph Regali

Exclusive event generation for the COMPASS-II experiment at CERN and improvements for the Monte-Carlo chain

Exclusive event generation for the COMPASS-II experiment at CERN and improvements for the Monte-Carlo chain

Dissertation

zur Erlangung des Doktorgrades
der
Mathematik und Physik
der
Albert-Ludwigs-Universität
Freiburg im Breisgau

vorgelegt von

Christopher Ralph Regali

Freiburg, Oktober 2016

Supervisor of the PhD thesis:	Prof. Dr. Horst Fischer
Second assessor:	Prof. Dr. Karl Jakobs
Dean:	Prof. Dr. Gregor Herten

Date of the oral exam:	20.12.2016

Contents

1. Introduction

One of the most important challenges of physics is the exploration of the matter surrounding us. By the end of the 19th century it was known that all visible matter consisted of the so-called atoms. Meyer and Medelejev succeeded in not only arranging them, but also in making precise predictions about further elements, which were still undiscovered back then [1]. This was a first indication that atoms could have a substructure.

By scattering α-particles on a thin gold foil, Rutherford was able to prove, that atoms consist of a small hard core, the nucleus, that contains the complete positive charge while the negative charge was distributed in a larger hull [2]. With the discovery of the neutron in 1932, all constituents of the atoms became known [3]: The electron, the proton and the neutron. In 1930 the neutrino was postulated from β-decay spectra and in 1935 a pion was introduced as the particle being responsible for the nuclear interactions [4].

As early as 1927, the heat capacity of hydrogen was measured at low temperatures. The degree of freedom of the nucleus resulting from its spin leads to an anomaly inexplicable by classic thermodynamics [5]. A precision measurement of the magnetic moment of hydrogen was also incompatible with the assumption of the proton being point-like [6]. Despite this, the set of known particles was thought to be complete by then. These two hydrogen measurements, however, were an indication that even the proton might have a substructure.

The discovery of the muon in cosmic radiation led to the discovery of more and more new particles in the following years. In 1964 Gell-Mann and Zweig suggested a model that was able to explain all known particles: the Quark model [7]. The breakthrough for this model came only later, when the proton's substructure was confirmed at SLAC[1] [8]. The measurements confirmed the proton to be made up of point-like constituents, called the quarks. In order to make the model complete, a set of gluons, which are the mediators of the strong interaction between these quarks, was postulated. These gluons were also searched for and first found at DESY[2] [9]. These findings brought the quark model wide acceptance and ultimately a Nobel prize [10]. Ever since then, the measurements

[1]**S**tanford **L**inear **A**ccelerator **C**enter
[2]**D**eutsches **E**lektronen **SY**ncrotron

of the nucleons' constituents have been conducted, becoming more and more precise over the years. This model has been an invaluable contribution to the field since one of the main goals of particle physics is to derive any physical quantities of the nucleons from knowledge about the quarks and gluons. The spin of the nucleon, which is a well-known quantum mechanical quantity, is of particular interest, since it can be measured with a very high precision. Surprisingly, however, it is still unknown how the constituents generate it. According to Jaffe and Manohar, the spin of the proton can be decomposed into [11]:

$$J_z = \frac{1}{2}\Delta\Sigma + \Delta g + \langle L_z \rangle, \tag{1.1}$$

where J_z is the total angular momentum, $\frac{1}{2}\Delta\Sigma$ is the direct contribution from quarks and anti-quarks, Δg is the contribution from the gluons and $\langle L_z \rangle$ is the contribution from the orbital angular momenta of the quarks, antiquarks and gluons. For protons, $J_z = \frac{1}{2}\hbar$ is a well-measured quantity [12]. With the introduction of polarized beams and targets the contribution of all quarks and anti-quarks in a nucleon, $\Delta\Sigma$, became accessible. The first experiment that was able to measure this quantity to a sufficient uncertainty was the EMC[3] experiment at CERN[4]. Its surprising result showed that all quarks and antiquarks together only accounted for about 38 % of the protons' spin [13]. The resulting problem was also called the "spin-puzzle". The COMPASS[5] experiment at CERN, a successor to SMC[6], carried out polarized muon proton scattering and constrained the gluon contribution to the proton spin to be smaller than $|\Delta g| < 0.2\,\hbar$ [14–16]. While being an important part of the solution to the "spin-puzzle", it still does not solve the puzzle completely. Another contribution comes from the orbital angular momenta, $\langle L_z \rangle$, of the quarks and gluons. Unfortunately, there is still no known direct experimental access to them.

A different approach of spin decomposition for the proton suggests an access to the total angular momenta via the generalized parton distributions [17]. These distributions link the elastic nucleon form factors to the parton distribution functions and may help to solve the "spin-puzzle". This indirect access is the main motivation for this thesis. In Chapter 2, the theory of particle scattering is introduced starting with Rutherford scattering up to the recent developments in the generalized parton distributions. The measurement of these is a very challenging task itself, since it is required that sums and differences of total cross sections of rare exclusive processes be measured. The most promising process is deeply virtual Compton scattering. A muon is scattered off a proton, which stays intact, while almost all of the transferred energy is converted into a single photon. Since the process is rare, a high luminosity is needed to find it. On the downside, however, the high rate creates a vast amount of background, which easily hides the single photon. The COMPASS-II experiment, standing in the tradition of EMC, SMC and COMPASS, is optimized to overcome these issues. The experimental setup is introduced in Chapter 3.

Technically, the measurement of sums of exclusive processes requires a very well-known acceptance of up to three percent uncertainty. To increase the accuracy of the experimental description in the Monte-Carlo, new features were introduced to the COMPASS-II simulation software chain. These improvements mainly focused on adding simulation details,

[3]**E**uropean **M**uon **C**ollaboration
[4]**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire
[5]**CO**mmon **M**uon **P**roton **A**pparatus for **S**tructure and **S**pectroscopy
[6]**S**pin **M**uon **C**ollaboration

which can only be accessed empirically, are introduced in Chap. 4. In order for spatial anisotropies to be taken into account in detector performance, a new set of tools was created allowing for a two dimensional extraction of efficiencies. These tools were used to create a full set of pseudo efficiencies for the 2012 pilot run data of the experiment. The tools, efficiency maps and the results are presented in Sec 4.1.

The exclusive photons, as well as the photons from neutral pions, need to be measured in the electromagnetic calorimeters, which is why their performance is of peculiar interest in the simulation. The data taking of COMPASS-II in 2016 and 2017 will produce vast amounts of data, that will need to be accompanied by even larger amounts of Monte-Carlo. To ease the generation of a simulation sample of the necessary size, the generation process needed an optimization. This system allows for the response of the spectrometer to the pile-up muons from the beam and the halo to be precomputed, and to save it to a new binary file format. The caching not only brings a drastic speed-up to the simulation as redundant muon tracking is omitted, but also allows to start the Monte-Carlo generation before the muon flux is known. This feature, that was already used in the 2012 pilot run Monte-Carlo generation, is described in Sec. 4.2.

To take the electronic noise generated by the photomultiplier and the front-end electronics into account, a new software module was created. This new module allows for extraction of the background profiles from real data on a cell-wise base as well as the re-injection of background according to these profiles into the reconstruction. A set of profiles was extracted from the 2012 pilot run data. This work is presented in Sec. 4.3.

Also, a new Monte-Carlo simulation software, called TGEANT[7], was created from scratch over the last years in order to ensure a good simulation of the experiment. The software simulates the interaction of particles and the detector response inside the experimental setup. Here, an appropriate event generator for exclusive reactions in the COMPASS-II kinematics comes into the game. A new C++ based event generator, HEPGen++, is presented in Chapter 5. It features different models for the exclusive muoproduction off protons. The generator also contains the production of single photons and different exclusively produced mesons, which are either experimentally valuable on their own or important background channels.

The exclusive production of neutral pions is of special interest since they decay into two hard photons, which makes them an important background, while, at the same time, the cross section can be used to constrain the generalized parton distributions. Currently the most sophisticated models for this exclusive neutral pion production are the model by S. Goloskokov and P. Kroll and the production model by S. Liuti and G. Goldstein. The C++ implementation of the model by Goloskokov and Kroll in HEPGen++ is presented in Chapter 6.

[7] **T**otal **GE**ometry **AN**d **T**racking

2. Theory

This chapter presents an introduction to the theoretical background which is essential for the rest of the thesis. In Sec. 2.1 elastic scattering is introduced from the Rutherford equation to the elastic form factors. Section 2.2 takes the step to deep inelastic scattering cross section, structure functions and parton distribution functions. This is generalized to the polarized case in Sec 2.3. With the polarized deep inelastic scattering theory, the spin puzzle naturally comes up. The issue of the integrated charge-square weighted sums of the helicity distributions not adding up to the nucleon spin is elaborated on in Sec. 2.4. As a possible solution, the generalized parton distributions are introduced in Sec. 2.5. Moreover, exclusive production processes need to be measured in order to get experimental access to these functions. They are classified as either deeply virtual compton scattering, introduced in Sec. 2.6, for the production of photons, or hard exclusive meson production in case of mesons (see Sec. 2.7). With the help of models, these cross section measurements can be used to constrain the generalized parton distributions. A current status is given in Sec. 2.8.

2.1 Elastic Scattering

2.1.1 Cross Sections

The elastic scattering of point-like, spin-less particles is well known and can be calculated exactly. When scattering leptons off nucleons the Rutherford cross section can be assumed to be valid [18]:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{Rutherford} = \frac{(e^2)^2}{(4\pi\epsilon_0)^2\,(4E_{kin})^2\,\sin^4\frac{\Theta}{2}}. \tag{2.1}$$

In this equation e is the elementary charge, ϵ_0 is the vacuum permittivity and E_{kin} is the kinetic energy of the beam particle. The angle Θ is the polar scattering angle, defined with respect to the unscattered "incoming" particle. At higher energies the relativistic effects and the spin become important. In order to consider helicity conservation an additional term can be multiplied to this cross section. The result is the so-called Mott cross section [18]:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)^*_{Mott} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{Rutherford} \cdot \left(1 - \beta^2 \sin^2\frac{\Theta}{2}\right). \tag{2.2}$$

2.1.2 Form Factors

When measuring experimental cross sections, it was noticed that only in the limit of $|\vec{q}| = |\vec{p} - \vec{p}'| \to 0$ the Mott cross section is in agreement with data. The $\vec{q}$ is defined as the momentum transfer, which is the difference between the incoming beam particles momentum $\vec{p}$ and the scattered beam particle $\vec{p'}$. Otherwise the data are systematically lower due to the geometric dimension of the target. Form factors can be introduced to account for this effect [18]:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)^*_{Exp} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)^*_{Mott} \cdot |F(\vec{q}^2)|^2. \tag{2.3}$$

These form factors are the Fourier transform of the charge distribution

$$f(r) = \frac{1}{(2\pi)^3}\int F(\vec{q}^2)e^{-i\vec{q}\vec{x}/\hbar}\mathrm{d}^3\vec{q}. \tag{2.4}$$

Here the intermediate step via $f(\vec{x})$ was omitted, as $\vec{x}$ in this case has an angular symmetry, so only the radial part is of interest. In the case of scattering off an electron $F(\vec{q}^2) = const$ hence $f(r) = \delta(r)/4\pi$ so the electron is point-like. Now considering the proton, a spin-$\frac{1}{2}$ particle with mass m_p, the cross section becomes [19]:

$$\begin{aligned}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = &\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\text{Mott}} \\ &\cdot\left[F_1^2 + \frac{Q^2}{4m_p}\left\{F_2^2 + 2(F_1^2 + F_2^2)^2\tan^2(\theta/2)\right\}\right].\end{aligned} \tag{2.5}$$

This is called the Rosenbluth cross section. It should be noted that the dependence of the form factors is written here in terms of $Q^2 = -q^2$ where q is the four-momentum transfer. The definition is given later in Tab. 2.1. It contains two elastic nucleon form factors,

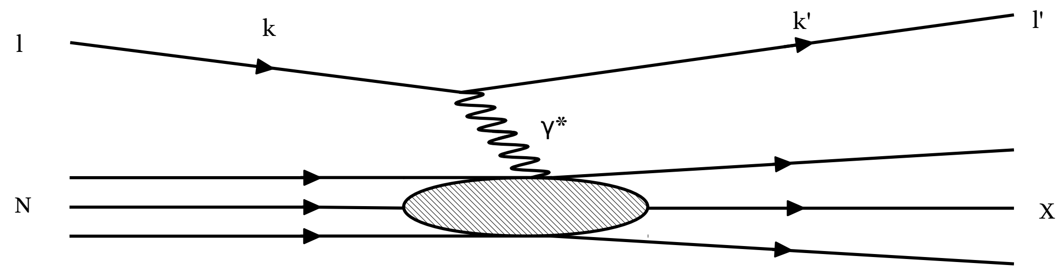


Figure 2.1: The graph of an arbitrary DIS process.

the so-called Dirac form factor $F_1(Q^2)$ and the Pauli form factor $F_2(Q^2)$. These can be recombined to the Sachs parametrization of the elastic form factors [20]:

$$G_M(Q^2) = F_1(Q^2) + F_2(Q^2), \tag{2.6}$$

$$G_e(Q^2) = F_1(Q^2) - \frac{Q^2}{4m_p} F_2(Q^2). \tag{2.7}$$

2.2 Deep Inelastic Scattering

2.2.1 Kinematic Variables

The process of leptons being scattered off nucleons can in general be written as

$$lN \to l'X$$

where l and l' is the lepton[1], N is the nucleon and X is the hadronic final state. This process is shown in Fig. 2.1 and can be used to derive the kinematic variables listed in Tab. 2.1. When the energy of the four-momentum transfer is high enough, $Q^2 > 1$ (GeV/c)^2, the virtual photon can interact with the constituents of the nucleon, the partons. These partons are subject to the strong interaction and consequently energy can be absorbed into excitation. The mass of the hadronic final state W can be used to destinguish between the region of elastic scattering ($W = m_p$), inelastic excitation ($W \lesssim 2\,\mathrm{GeV/c^2}$) and deep inelastic scattering where the proton can be fully fragmented (W large w.r.t. m_p).

2.2.2 Structure Functions

In inclusive DIS only the scattered lepton is measured while the final hadronic state is unknown, hence it is integrated over. In the case of semi-inclusive DIS, also abbreviated SIDIS[2], the final hadronic state is measured partially. Finally, in the exclusive DIS the complete final state is detected. In the Bjorken limit $Q^2 \to \infty, x_{bj} = const$ the cross section can be written as a product of the leptonic tensor and the hadronic tensor:

$$\mathrm{d}\sigma \sim L_{\mu\nu} W^{\mu\nu}. \tag{2.8}$$

The leptonic tensor $L_{\mu\nu}$ contains the hard part of the interaction and can be calculated with quantum electrodynamics theory. The hadronic tensor $W^{\mu\nu}$ is mainly unknown and

[1] Also contains anti-leptons

[2] **S**emi **i**nclusive **d**eep **i**nelastic **s**cattering

Table 2.1: Used kinematical variables for the DIS

Variable	Description
$k = (E, \vec{k})$	Momentum of the incident lepton, see Fig. 2.1.
$k' = (E', \vec{k}')$	Momentum of the scattered lepton, see Fig. 2.1.
$p = (E_p, \vec{p})$	Momentum of the target proton.
$q = k - k'$	Momentum of the virtual photon γ^*, see Fig. 2.1.
Θ	Polar angle of the scattered lepton with respect to the incident lepton.
$Q^2 := -q^2 = (k-k')^2 \stackrel{lab}{=} \frac{4EE'}{c^2}\sin^2\frac{\Theta}{2}$	Negative four-momentum transfer squared.
$x_{bj} := \frac{Q^2}{2pq} = \frac{Q^2}{2m_p\nu}$	Bjorken scaling variable.
$W^2c^2 := (p+q)^2 \stackrel{lab}{=} m_p^2 + 2m_p\nu - Q^2$	Invariant mass squared of the final hadronic state.
$\nu := \frac{p \cdot q}{m_p} \stackrel{lab}{=} E - E'$	The total energy of the γ^*, transferred from incident to scattered muon.
$y := \frac{\nu}{E}$	Fraction of the transferred energy between incident and scattered muon˙
$z := \frac{E_{had}}{\nu}$	The energy fraction of the γ^* of a final hadron with the energy E_{had}.
$\xi \approx \frac{x_{bj}}{2-x_{bj}}$	The skewness of an exclusive process.
$t := (p-p')^2 = -\Delta^2$	Squared four-momentrum transfer from target to the recoiled proton.
$t' := t - t_0 = t + 4\frac{m_p^2\xi^2}{1-\xi^2}$	t corrected for the minimal possible t_0.

therefore the interesting quantity for experiments. However, even though it is unknown, theory can be used to place some constraints on it. The hadronic tensor $W^{\mu\nu}$ contains six structure functions W_1 to W_6 in its most general form. To simplify, unpolarized scattering will be assumed first. In addition, since the invariant mass in the COMPASS-II experiment is limited to $\sqrt{s} = \sqrt{k+p} < 19\ \mathrm{GeV/c^2}$ which is well below the Z^0 mass, only virtual photons are considered. With the conservation of probability current at the vertex, the hadronic tensor can be simplified to [21]:

$$W^{\mu\nu} = W_1\left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + \frac{W_2}{m_p^2}\left(p^\mu - \frac{p\cdot q}{q^2}q^\mu\right)\left(p^\nu - \frac{p\cdot q}{q^2}q^\nu\right). \tag{2.9}$$

This expression can be used to write the cross section of the inclusive DIS as [22]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q^2\mathrm{d}\nu} = \frac{4\pi\alpha^2}{q^4}\left[W_2\cos^2\left(\frac{\Theta}{2}\right) + 2W_1\sin^2\left(\frac{\Theta}{2}\right)\right]. \tag{2.10}$$

The structure functions W_i that are used to express the hadronic tensor are generally functions of two independent variables, e.g. Q^2, ν. Experimental data has proven that, over a wide range of Q^2 and x_bj, the W_i can be substituted by two dimensionless functions of only x_{bj}:

$$MW_1(Q^2, \upsilon) \rightarrow F_1(x_{bj}) \tag{2.11}$$

$$W_2(Q^2, \upsilon) \rightarrow F_2(x_{bj}). \tag{2.12}$$

This phenomenon is called Bjorken scaling and can also be explained by the parton model. As the structure functions do not stronly depend on Q^2 and over a wide kinematic range the struck partons can be assumed to be point-like. This scaling behavior can be seen in Fig. 2.2.

By using equation 2.10 and replacing the W_1 and W_2 with F_1 and F_2 and then comparing coefficients with the cross section of scattering off point-like spin-$\frac{1}{2}$ particles, the so-called Callan-Gross relation follows:

$$2x_{bj}F_1(x_{bj}) = F_2(x_{bj}). \tag{2.13}$$

As it is experimentally well proven, the partons are spin-$\frac{1}{2}$ particles.

2.2.3 Parton Distribution Functions

The structure functions can be parametrized as a sum of charge-square weighted probability distributions to find a quark of the given flavour at a fixed x_{bj}. These probability functions are called parton distribution functions or PDFs[3] in short. The function $q_f(x_{bj})\mathrm{d}x_{bj}$ denotes the expectation value of the probability of hitting a quark of flavour f in the interval of $[x_{bj}, x_{bj} + \mathrm{d}x_{bj}]$. With these functions, F_2 can be written as:

$$F_2(x_{bj}) = x_{bj} \sum_f e_f^2(q_f(x_{bj}) + \bar{q}_f(x_{bj})). \tag{2.14}$$

Figure 2.3 shows these PDFs on a large range of x_{bj}. As expected, it can be observed that the valence quarks are dominant in the high x_{bj} regime, and that the sea quarks are gaining influence as x_{bj} is getting lower. In Fig. 2.2, a small violation of the Bjorken scaling is already visible. This deviation is coming from higher order corrections of the quantum chromodynamics (QCD) theory. The interpretation of the probabilities at the different energy scales is, that probability of finding a parton with a high momentum fraction is lower at higher energy scales. This is also the reason, why the F_2 structure function has a positive slope in Fig. 2.2 at small x_{bj} with rising Q^2. The PDFs can be evolved to any desired energy scale by the so-called DGLAP equations [24].

2.3 Polarized DIS

When it comes to polarized DIS, some new angles are introduced, see Fig. (2.4a, 2.4b). The azimuthal angle ϕ_S can be written as:

$$\phi_S = \frac{(\vec{q} \times \vec{l}) \cdot \vec{S}}{|(\vec{q} \times \vec{l}) \cdot \vec{S}|} \arccos\left(\frac{(\vec{q} \times \vec{l}) \cdot (\vec{q} \times \vec{S})}{|\vec{q} \times \vec{l}||\vec{q} \times \vec{S}|}\right), \tag{2.15}$$

where $\vec{S}$ is the nucleon spin.

[3]**P**arton **D**istribution **F**unctions

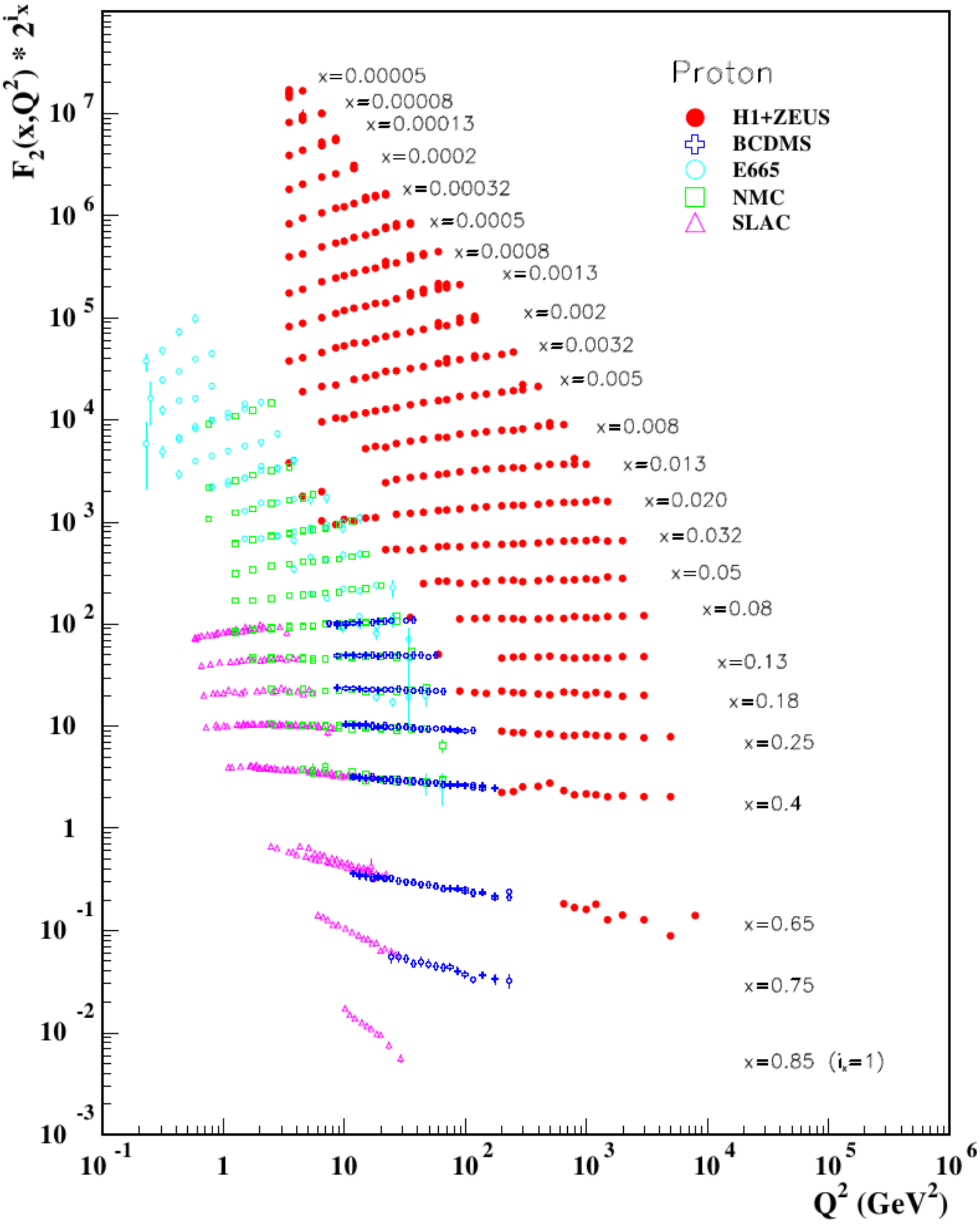


Figure 2.2: Dependence of the proton structure function F_2 on Q^2 for various values of fixed x_{bj} [12].

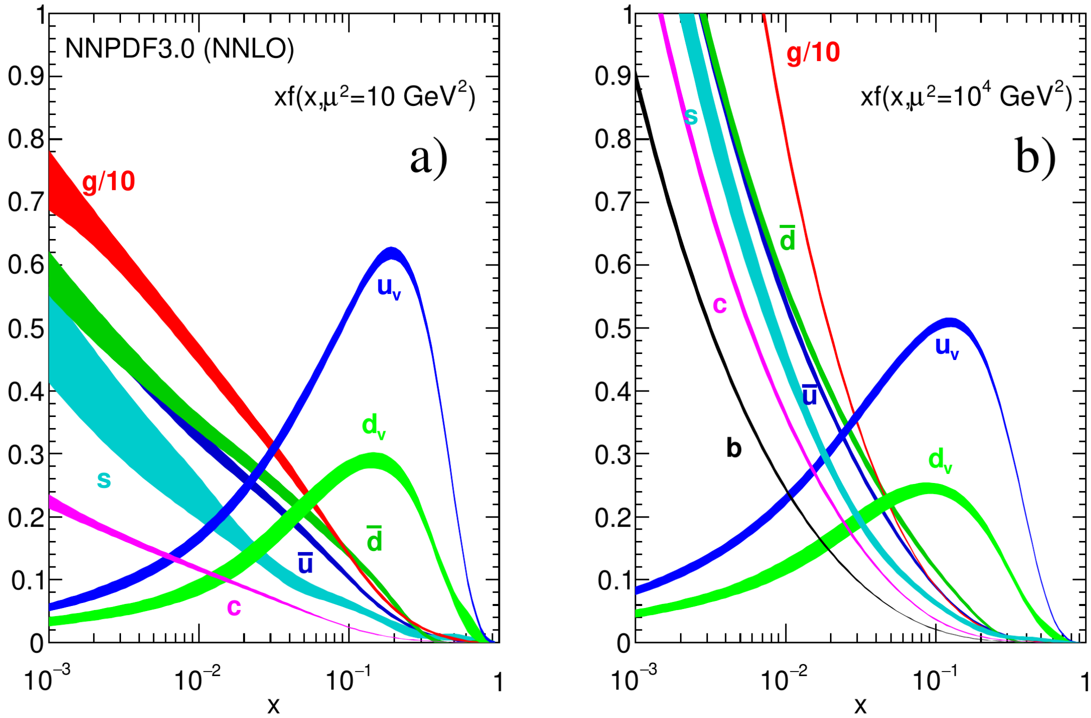
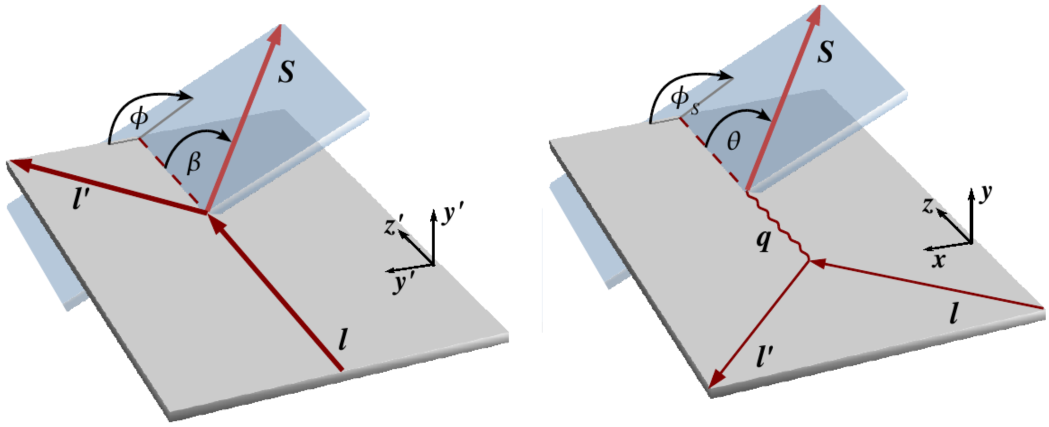


Figure 2.3: Dependence of the parton density functions for several quark types on x_{bj} at two different energy scales. In the lower energy regime, on the left, at $x_{bj} \approx 0.3$ the simple valence quark model holds quite well, as there is almost no sea quark contamination. At the higher energy scale, on the right, the sea quark and gluon PDF rise much faster. The simple valence quark model das absolutely not hold here [23].



(a) Definiton of the azimuthal angle ϕ as the angle between the ***l l'*** plane and the ***l S*** plane. Please note that the boldface letters in these figures denote vectors.

(b) Definition of the azimuthal angle ϕ_S as the angle between the leptonic ***l l'*** plane and the production plane γ^* ***S***.

Figure 2.4: Angle definitions in different reference systems [25].

2.3.1 Cross Section

The starting point for the polarized DIS is once more the equation 2.8. The tensors can be split into symmetric (S) and anti-symmetric (A) parts under interchange of ν and μ, resulting in the cross section equation [26]:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}E'} = \frac{\alpha^2}{2m_p q^4}\frac{E'}{E}\left[L^{(S)}_{\mu\nu}W^{\mu\nu(S)} + L'^{(S)}_{\mu\nu}W^{\mu\nu(S)} - L^{(A)}_{\mu\nu}W^{\mu\nu(A)} - L'^{(A)}_{\mu\nu}W^{\mu\nu(A)}\right]. \quad (2.16)$$

Here the term $L^{(S)}_{\mu\nu}W^{\mu\nu(S)}$ is proportional to the spin-averaged unpolarized cross section from equation 2.10. With s and s' as the covariant spin four-vectors of the lepton initially and finally, the parts of the leptonic tensor can be written as [26]:

$$L^{(S)}_{\mu\nu} = k_\mu k'_\nu + k'_\mu k_\nu - g_{\mu\nu}(k\cdot k' - m_p^2), \quad (2.17)$$

$$L^{(A)}_{\mu\nu} = m_p \varepsilon_{\mu\nu\alpha\beta} s^\alpha (k-k')^\beta, \quad (2.18)$$

$$\begin{aligned} L'^{(S)}_{\mu\nu} =& (k\cdot s')(k'_\mu s_\nu + s_\mu k'_\nu - g_{\mu\nu}k'\cdot s) - (k\cdot k' - m_p^2)(s_\mu s'_\nu + s'_\mu s_\nu - g_{\mu\nu} s\cdot s') \\ &+ (k'\cdot s)(s'_\mu k_\nu + k_\mu s'_\nu) - (s\cdot s')(k_\mu k'_\nu + k'_\mu k_\nu), \end{aligned} \quad (2.19)$$

$$L'^{(A)}_{\mu\nu} = m_p \varepsilon_{\mu\nu\alpha\beta} s'^\alpha (k-k')^\beta, \quad (2.20)$$

where $\varepsilon_{\mu\nu\alpha\beta}$ is the Levi-Civita symbol of fourth order in such way, that:

$$\varepsilon_{0123} = +1, \quad (2.21)$$

and $g_{\mu\nu}$ is the Minkowsky metric.

2.3.2 Structure Functions

The anti-symmetric terms of the hadronic tensor from equation 2.16 can also be expressed in terms of new structure functions G_1 and G_2:

$$\frac{1}{2m_p}W_{\mu\nu}^{(A)} = \tag{2.22}$$

$$\epsilon_{\mu\nu\alpha\beta}q^\alpha \left[m_p S^\beta G_1(P\cdot q, q^2) + [(P\cdot q)S^\beta - (S\cdot q)P^\beta]\frac{G_2(P\cdot q, q^2)}{m_p}\right]. \tag{2.23}$$

As G_1 and G_2 should also scale in the Bjorken limit, the Q^2 dependence can be dropped and dimensionless structure functions can be defined: $g_1(x_{bj})$ and $g_2(x_{bj})$.

Analogously to definition in equation 2.14, the polarized structure functions can be expressed in terms of corresponding polarized PDFs, called the helicity distributions Δq_f:

$$g_1(x_{bj}) = \frac{1}{2}\sum_f e_f^2 \Delta q_f(x, S), \tag{2.24}$$

$$g_2(x_{bj}) = 0. \tag{2.25}$$

The structure function g_2 cannot be interpreted and the simple parton model assumes it to be zero. Non-zero values can be obtained by allowing fermi motion of quarks inside the nucleon [26]. This is included in the QCD improved parton model. The function g_1 for the proton can be seen in Fig. 2.5. For a longitudinally polarized lepton and a transversely polarized nucleon a transversity distribution h_1^f can be defined. Here h_1^f is the difference in density of quarks with spin parallel to the nucleon spin and those which are anti-parallel.

2.4 Spin Puzzle

2.4.1 Quark Contribution

Out of all the results obtained from the polarized DIS, the following one was especially interesting. The helicity distribution Δq_f can be integrated to get the fraction of the total spin carried by the given quark flavour f. This can be formulated as the equation:

$$S_f^z = \frac{1}{2}\hbar \int_0^1 \mathrm{d}x_{bj} \Delta q(x_{bj}). \tag{2.26}$$

The sum of these S_f^z was expected to be equal to about 60% of the full proton spin of $J_z = \frac{1}{2}\hbar$, due to relativistic effects. Due to symmetry, the exchange relations for the parton density, helicity and transversity distribution are as follows:

$$\bar{q}_f(x_{bj}) = -q_f(-x_{bj}), \tag{2.27}$$

$$\Delta\bar{q}_f(x_{bj}) = -\Delta q_f(-x_{bj}), \tag{2.28}$$

$$\bar{h}_1^f(x_{bj}) = -h_1^f(-x_{bj}). \tag{2.29}$$

With these relations the spin sum can be conveniently written as:

$$\Delta\Sigma = \hbar \sum_f \int_{-1}^1 \mathrm{d}x_{bj} \Delta q_f(x_{bj}). \tag{2.30}$$

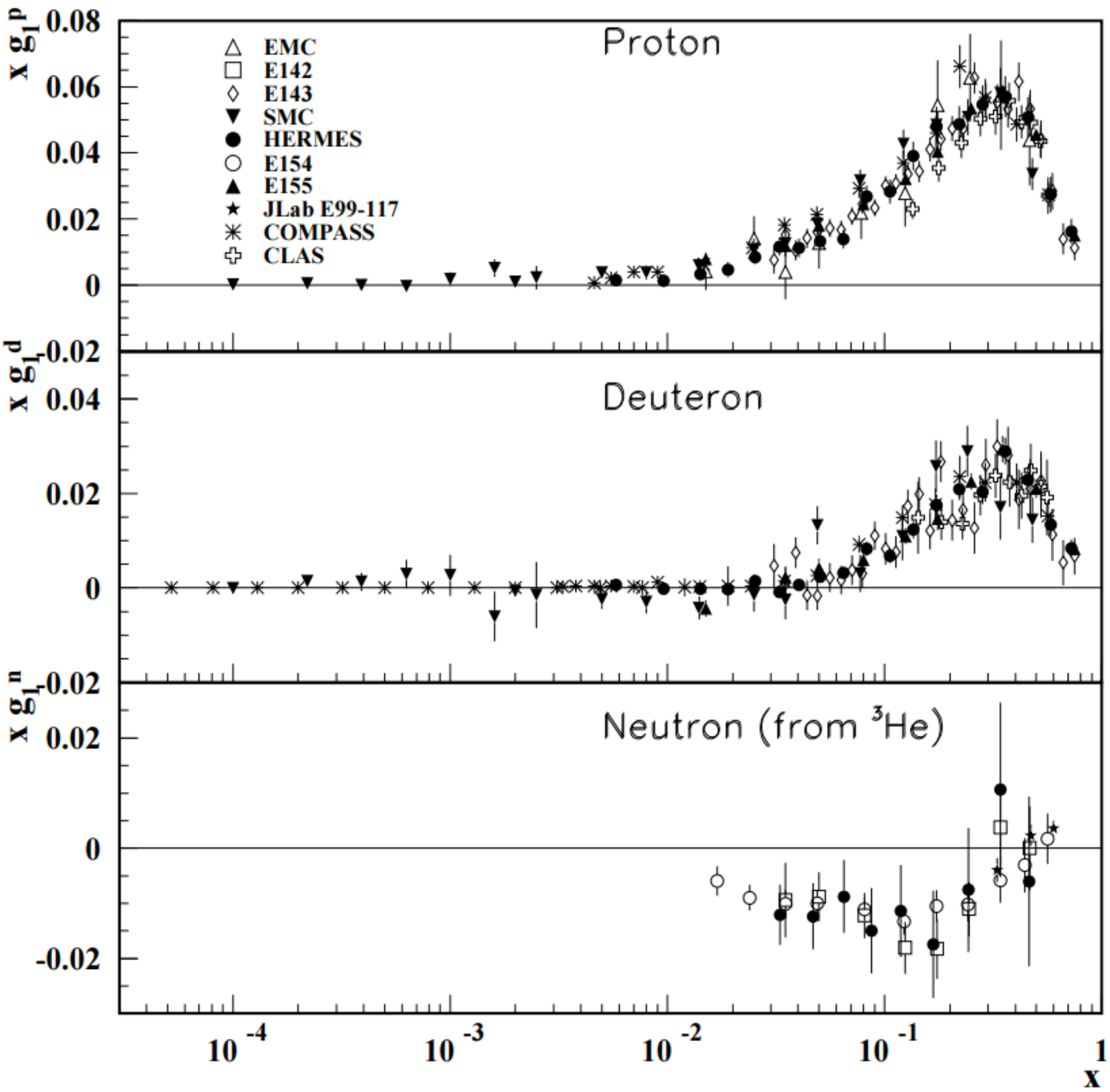


Figure 2.5: The polarized structure function g_1 against x_{bj} for the proton. [23]

The first experiment to measure this with sufficient accuracy, was the EMC experiment at CERN. The surprising result was that this sum is only measured to [27]:

$$\frac{1}{2}\Delta\Sigma = \frac{1}{2}(0.14 \pm 0.09 \pm 0.21)\hbar. \tag{2.31}$$

The measured range of Q^2 was $1.5\,(\mathrm{GeV/c})^2 < Q^2 < 70\,(\mathrm{GeV/c})^2$. So the sum of all quarks' and anti-quarks' spin only accounts for about 14% of the proton's total spin. Naturally, this raised the question of how the remaining spin of the proton is generated. The resulting problem has been called "spin puzzle" or "spin crisis".

2.4.2 Other Contributions

According to Jaffe and Manohar the total spin of the proton can be decomposed to [11, 28]:

$$J_z = \frac{1}{2}\Delta\Sigma + \Delta g + \langle L_z \rangle, \tag{2.32}$$

where Δg is the gluon polarization and $\langle L_z \rangle$ is the contribution of the angular momenta of quarks and gluons. This decomposition is valid in the infinite momentum frame and light-cone gauge. The measurement of the gluon spin contribution is a difficult task, as the experimental access is not straightforward. First assumed to be in the order of $\Delta g \approx 3\hbar - 4\hbar$ by several theories [29] the measurements conducted by the COMPASS experiment, as shown in Fig. 2.6, proved the value to be more than an order of magnitude smaller. The conclusion from all COMPASS physics channels combined can today be written as[14–16]:

$$|\Delta g| \lesssim 0.2\hbar, \tag{2.33}$$

with indications for Δg being positive [16]. The final contribution that is possible from the Jaffe-Manohar decomposition are the angular momenta of quarks and gluons, $\langle L_z \rangle$. These are even more difficult to constrain or measure because there is no direct experimental access to them.

2.5 Generalized Parton Distributions

2.5.1 Motivation

To tackle the obstacle of gaining access to the orbital angular momenta of quarks and gluons, Ji introduced a new spin decomposition for the nucleon [17]:

$$\frac{1}{2}\hbar = \sum_q J_q + J_g. \tag{2.34}$$

In this equation J_q is the total angular momentum of the quarks of flavour q and J_g is the total angular momentum of the gluons. This decomposition can be shown to be valid with longitudinally and transversely polarized moving nucleon, even though it is the natural decomposition in the nucleon rest frame [30]. The total angular momenta J_q and J_g

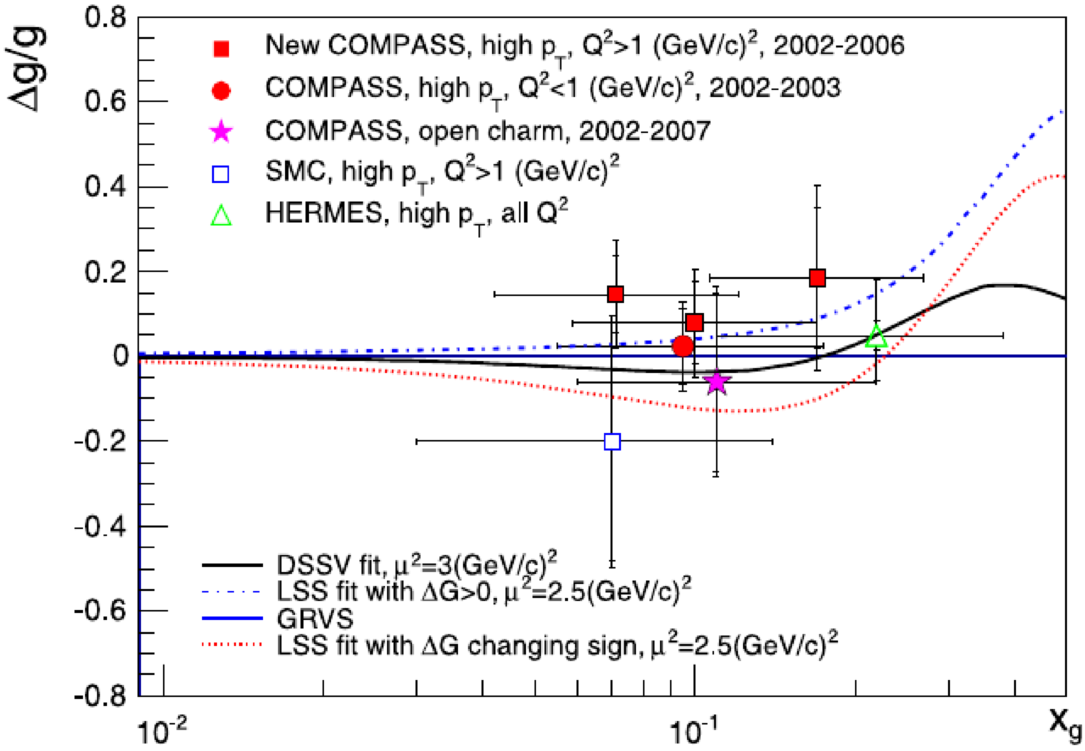


Figure 2.6: Measurements of the $\Delta g/g$ from COMPASS, HERMES and SMC. [15]

are accessible via moments of new distributions, called generalized parton distributions (GPD[4]). These GPDs are linked to the total angular momenta by Ji's sum rule [17, 31]:

$$J_q = \frac{1}{2}\hbar \int_{-1}^{+1} \mathrm{d}x x[H^q(x,\xi,t=0) + E^q(x,\xi,t=0)], \tag{2.35}$$

$$J_g = \frac{1}{2}\hbar \int_{0}^{+1} \mathrm{d}x x[H^g(x,\xi,t=0) + E^g(x,\xi,t=0)], \tag{2.36}$$

with the GPDs H and E. The definition for the skewness ξ was already given in Tab. 2.1 and is further explained in Sec. 2.5.2. Experimental access to these GPDs exists by measuring exclusive processes to a high precision.

2.5.2 Kinematics

In Sec. 2.2, the cross section was derived by using the factorization theorem to split the interaction into a hard and a soft part. In the case of large photon virtuality, this factorization holds true also for a finite momentum transfer to the target. Such exclusive processes can be drawn in so-called handbag diagrams as shown in Fig. 2.7. The name "handbag diagram" refers to the shape of the diagram. Here the GPDs encode the long-range part of the interaction while the short range is given by the partonic amplitude. This can be written as:

$$\mathrm{d}\sigma = [\text{partonic amplitude}] \otimes [GPD]. \tag{2.37}$$

[4] **G**eneralized **P**arton **D**istribtion

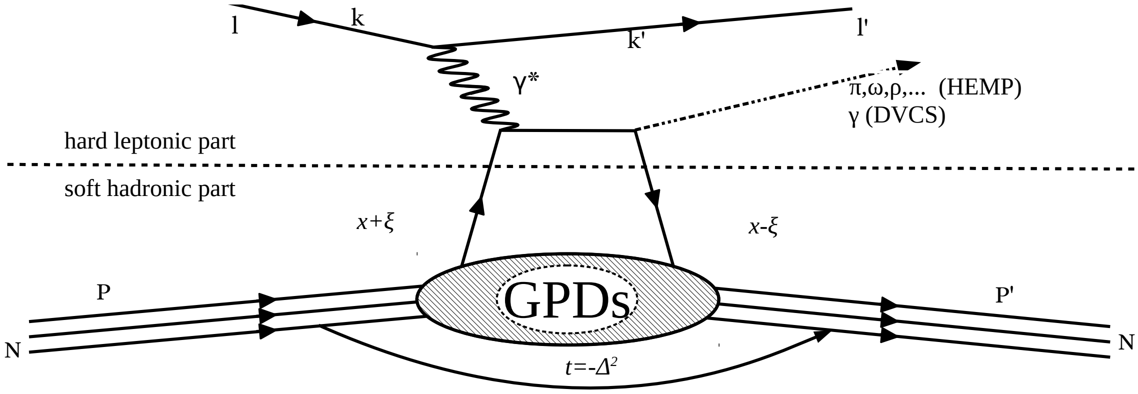


Figure 2.7: Handbag diagram of exclusive leptoproductions of photons or mesons.

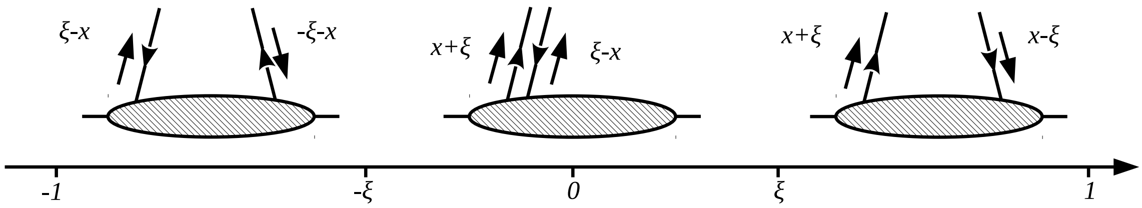


Figure 2.8: Parton model interpretation of the GPDs in different ξ regions. The arrows in the lines follow the convention for Feynman graphs, the stand-alone arrow denotes the space-like direction. In the first part ($x \in [-1, -\xi]$), an anti-quark is emitted and reabsorbed by the nucleon. The part in the middle ($x \in [-\xi, \xi]$) shows an emission of a quark followed by another emission by of an anti-quark. The last part, where $x \in [\xi, 1]$ shows the emission and reabsorption of a quark. Adapted from an original from [33].

The kinematic variables the GPDs depend on are the Mandelstam variable t, the so-called skewness ξ and the internal variable x. The Mandelstam variable t is the four-momentum transfer from the target nucleon to the recoiled target nucleon squared:

$$t = (p - p')^2 = -\Delta^2. \tag{2.38}$$

The skewness parameter ξ quantifies the difference of quark momentum between initial and final state from the mean value. In the COMPASS kinematics it can be approximated as a function of x_{bj} [32]:

$$\xi = \frac{p - p'}{p + p'} = x_{bj} \frac{1 + \frac{\Delta^2}{2Q^2}}{2 - x_{bj} + x_{bj}\frac{\Delta^2}{Q^2}} \approx \frac{x_{bj}}{2 - x_{bj}}. \tag{2.39}$$

With this skewness ξ, the longitudinal momentum fraction carried by the parton before(+) and after(-) the interaction can be written as $(x \pm \xi)$. The variable x is an internal variable and cannot be measured. Instead it needs to be removed by convoluting the GPDs with a scattering kernel and integrating over it. In the parton model the ξ variable can be interpreted in three different regions, as shown in Fig. 2.8. The three regions described are:

1. $x \in [\xi, 1]$: $x + \xi$ as well as $x - \xi$ are positive. This can be interpreted as an emission followed by the reabsorption of a quark.

Table 2.2: Classification of the GPDs in the scheme "*parton helicity conserving / parton helicity flip*".

	unpolarized	**polarized**
Nucleon helicity conservation	H^q / H^q_T	$\tilde{H}^q$ / $\tilde{H}^q_T$
Nucleon helicity flip	E^q / E^q_T	$\tilde{E}^q$ / $\tilde{E}^q_T$

2. $x \in [-\xi, \xi]$: As $x+\xi > 0$ can be interpreted as a quark being emitted by the proton, the $x - \xi < 0$ can be interpreted as the proton also emitting an anti-quark.

3. $x \in [-1, -\xi]$: This case is essentially the same as case 1, except that the emitted and reabsorbed quark is an anti-quark.

2.5.3 Classification

There are several definitions of the GPDs that differ by some factors, see the refs [17, 33, 34]. GPDs that conserve the nucleon helicity are labeled H (vector transition), whereas nucleon helicity flipping GPDs are named E (tensor transition). In case of the "polarized"[5], referring to the spin of the partons, rather than the target polarization, ones, typically a tilde is placed above. Therefore, for instance, the nucleon helicity conserving GPD in the polarized case is denoted as $\tilde{H}^q$(axial-vector transition), the nucleon helicity flip GPD is written $\tilde{E}^q$ (pseudoscalar transition). When it comes to quarks, there are four more GPDs that require a helicity flip in the parton. These are called "transversity GPDs" and they are typically indexed with a T as $\tilde{H}^q_T$. The transversity GPDs are usually used in the combination that is denoted with bar:

$$\bar{E}_T = 2\tilde{H}_T + E_T. \tag{2.40}$$

All of these GPD-types are listed in Tab. 2.2. The parton helicity flip GPDs were necessary to be introduced for hard exclusive meson production (HEMP) [35]. In the process of exclusive muoproduction of photons, also called DVCS[6], the only contribution is the leading twist interaction between the virtual photon and the quark which conserves helicity. Twist is a special quantum number introduced for DIS processes that depends both on the dimension and spin of an operator. It is a convenient tool to identify the order of an operator after the operator product expansion. It is defined as the dimension of an operator θ minus its spin [28]:

$$t_\theta = d_\theta - n_\theta. \tag{2.41}$$

In practice it is used to show the order of $\frac{1}{Q^2}$ of the effect seen in the experiment. It can be written as:

$$\mathrm{d}\sigma_\theta \sim \left(\frac{1}{Q^2}\right)^p, \tag{2.42}$$

where p is defined as:

$$p = 2(1 + t). \tag{2.43}$$

[5]See Sec. 2.5.4 for an explanation of the nomenclature
[6]**D**eeply **V**irtual **C**ompton **S**cattering

The leading twist for GPDs is twist-2. The higher-twist contributions can only occur in combination with a higher-twist meson wave function. This is explained in detail in Chapter 6.

2.5.4 Properties

The GPDs show some remarkable properties, which is the main reason today why they are such an extensively studied field of interest. These properties are introduced in this section.

2.5.4.1 Forward Limit

The forward limit is the case where $p' = p$ and helicities are equal in the initial and final state. In this special case the GPDs become the ordinary parton distribution and helicity distributions. This can be expressed in the equations [33]:

$$H^q(x,0,0) = q_q(x), \tag{2.44}$$
$$\tilde{H}^q(x,0,0) = \Delta q_q(x), \tag{2.45}$$

for $x > 0$.In the case of $x < 0$ the relations are [33]:

$$H^q(x,0,0) = -\bar{q}_q(-x), \tag{2.46}$$
$$\tilde{H}^q(x,0,0) = \Delta\bar{q}_q(-x). \tag{2.47}$$

When it comes to gluons, there is only the $x > 0$ case [33]:

$$H^g(x,0,0) = xg(x), \tag{2.48}$$
$$\tilde{H}^g(x,0,0) = x\Delta g(x). \tag{2.49}$$

No such forward-limits to known distribution exist for the GPDs E and $\tilde{E}$ for neither gluons nor quarks. A model calculation for the GPD H is shown in Fig. 2.9. The forward limit is marked by the red line.

2.5.4.2 Symmetry

The unpolarized GPDs for the gluons are even functions in x, while the polarized GPDs for the gluons are odd functions in x. This results from gluons being their own antiparticles. As the quark GPDs are neither even nor odd themselves often combinations of them are considered like the so-called “singlet” combinations [33]:

$$H^{q(+)} = H^q(x,\xi,t) - H^q(-x,\xi,t), \tag{2.50}$$
$$\tilde{H}^{q(+)} = \tilde{H}^q(x,\xi,t) + \tilde{H}^q(-x,\xi,t) \tag{2.51}$$

and the “nonsinglet” or “valence” combinations [33]:

$$H^{q(-)} = H^q(x,\xi,t) + H^q(-x,\xi,t), \tag{2.52}$$
$$\tilde{H}^{q(-)} = \tilde{H}^q(x,\xi,t) - \tilde{H}^q(-x,\xi,t). \tag{2.53}$$

The “singlet” combinations correspond to $C = +1$ exchange, while the “valence” combinations correspond to $C = -1$ exchange. Photon or vector meson production chooses

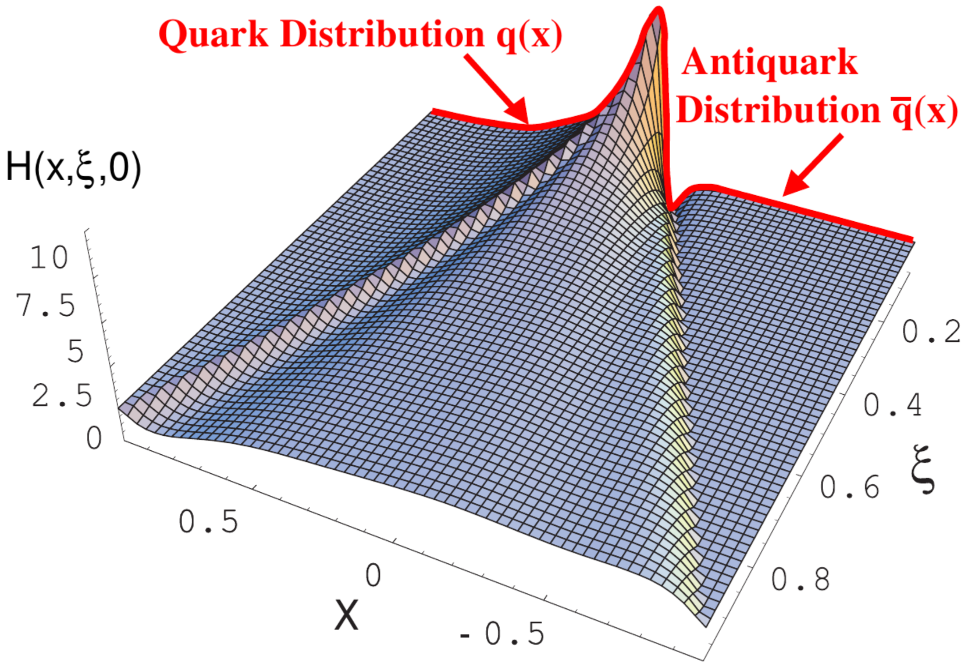


Figure 2.9: GPD H from a model calculation from [36]. Clearly visible and marked with the red line is the convergence towards the forward limit of the PDF $q(x)$.

charge-even combinations, whereas pseudoscalar meson production selects charge-odd combinations.

Another important symmetry comes from the time reversal relation which can be written as:

$$GPD(x, -\xi, t) = GPD(x, \xi, t). \tag{2.54}$$

It is valid for H, $\tilde{H}$, E, $\tilde{E}$ for gluons and quarks alike. Also, the complex conjugate allows for the relation:

$$[GPD(x, -\xi, t)]^* = GPD(x, \xi, t), \tag{2.55}$$

which is also valid for H, $\tilde{H}$, E, $\tilde{E}$.

2.5.4.3 Sum Rules

The GPDs show some interesting behavior in their first Mellin moments. The first moments deliver the local form factor currents as:

$$\int_{-1}^{1} \mathrm{d}x H^q(x, \xi, t) = F_1^q(t), \qquad \int_{-1}^{1} \mathrm{d}x E^q(x, \xi, t) = F_2^q(t), \tag{2.56}$$

$$\int_{-1}^{1} \mathrm{d}x \tilde{H}^q(x, \xi, t) = g_A^q(t), \qquad \int_{-1}^{1} \mathrm{d}x \tilde{E}^q(x, \xi, t) = g_p^q(t), \tag{2.57}$$

where F_1 and F_2 are the Dirac and the Pauli form factors, introduced in equation 2.5 and g_a and g_p are the axial and the pseudoscalar ones. An important constraint on the GPDs also relies on these moments by enforcing all moments of all GPDs to be writable as polynomials of the order n or $n+1$ in ξ for the nth moment. This feature is called "polynomiality of the GPDs", which is somewhat deceptive.

One of the most important sum rules for spin physics is Ji's sum rule, which was already mentioned in the motivation in equation 2.36.

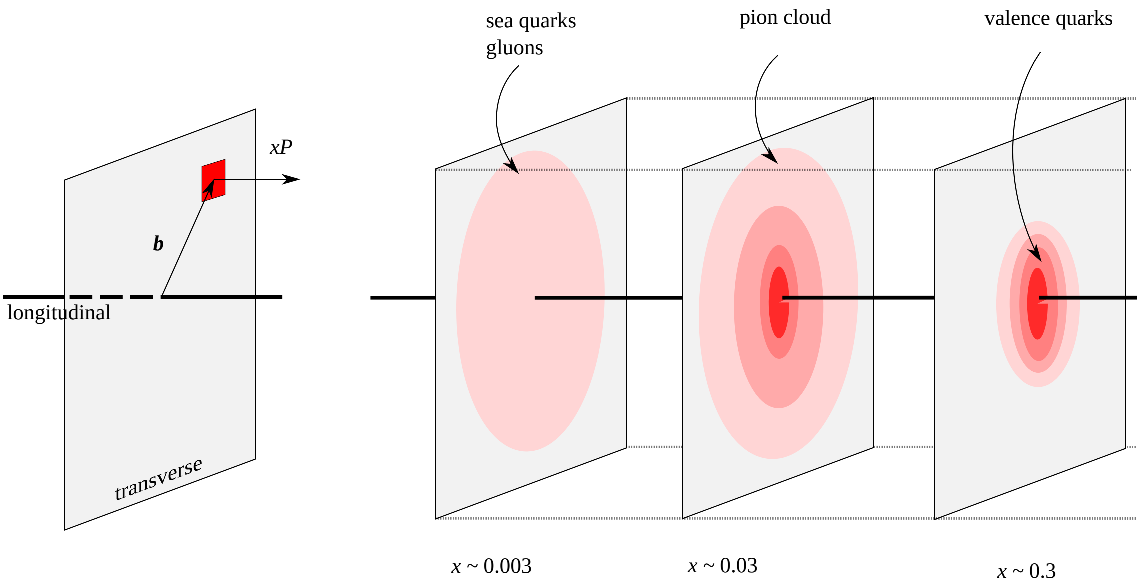


Figure 2.10: Schematic of nucleon tomography via impact parameter GPDs, note that $\vec{b}$ is shown as $\boldsymbol{b}$ here. Adapted from [38].

2.5.4.4 Impact Parameter Space

The GPDs can also be fourier transformed. If evaluated at $\xi = 0$ and in the case of the unpolarized H^q, the following relation can be shown [37]:

$$q(x, \vec{b}_\perp) = \int \frac{\mathrm{d}^2\Delta_\perp}{(2\pi)^2} H^q(x, 0, -\Delta_\perp^2) e^{-\mathrm{i}\vec{b}_\perp \cdot \Delta_\perp}. \tag{2.58}$$

In this equation $\vec{b}_\perp$ is the transverse impact parameter in space coordinates. As shown in [37], this form fulfills positivity constraints, and can be interpreted as a probability density distribution. This gives the impact parameter dependent parton distribution $q(x, \vec{b})$ great importance. On the one hand, it allows for a quasi three-dimensional nucleon tomography, as the impact parameter $\vec{b}_\perp$ can be described with two transverse space coordinates, indicated by $\perp$, while the momentum fraction dependence is longitudinally obtained. This combination is very well-chosen, so that the uncertainty principle is not a problem here. On the other hand, this relation allows to place constraints on the GPDs, because they need to be formulated in such a way that these properties stay valid. This can be used to limit a possible ansatz. Basic constraints are given below.

- Positivity and negativity:

$$\int \mathrm{d}^2\Delta_\perp \exp(-i\vec{b} \cdot \Delta_\perp) H_q(x, 0, -\Delta_\perp^2) \geq 0 \quad \text{for} \quad x > 0, \tag{2.59}$$

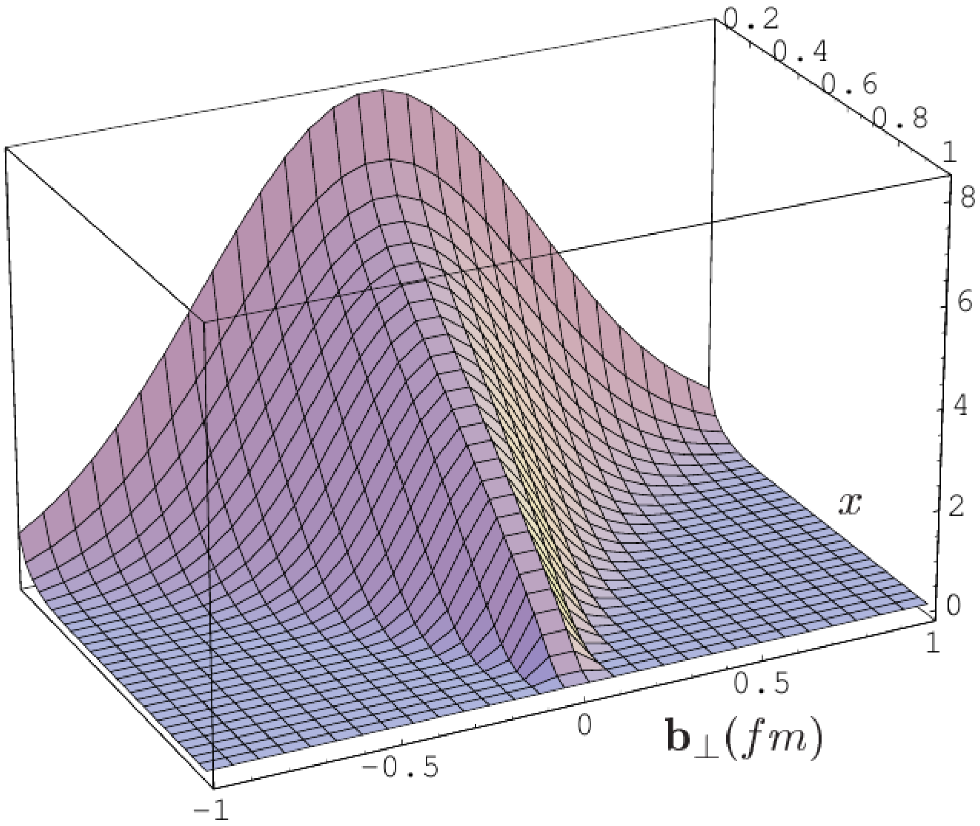


Figure 2.11: Impact parameter dependant parton distribution $u(x, \vec{b}_\perp)$. [37]

$$\int d^2\Delta_\perp \exp(-i\vec{b}\cdot\Delta_\perp) H_q(x, 0, -\Delta_\perp^2) \leq 0 \quad \text{for} \quad x < 0. \tag{2.60}$$

- Upper x limit:

$$\text{For } x \to 1 : \; H_q(x, 0, t) = H_q(x, 0). \tag{2.61}$$

- Lower x limit:

$$\text{For } x \to 0 : \; r_{\text{hadron}} \sim \alpha \ln \frac{1}{x}. \tag{2.62}$$

An ansatz that fulfills all of these constraints is shown in Fig. 2.11 and can be written as [37]:

$$q(x, \vec{b}_\perp) = q(x) \frac{1}{4\pi a(1-x)\ln\frac{1}{x}} \exp\left(-\frac{\vec{b}_\perp^2}{4a(1-x)\ln\frac{1}{x}}\right). \tag{2.63}$$

The parameter a is a model parameter introduced for the ansatz used in [37].

This principle has been researched to an even greater extent in [39].

2.6 Deeply Virtual Compton Scattering

As already mentioned in the classification section, Sec. 2.5.3, DVCS is the exclusive leptoproduction of real photons. At COMPASS, the production via muons is measured. This process was the first process that Ji suggested to measure the GPDs, in which he also formulated the amplitude for the DVCS process [17]. To make this cross section more readable the compton form factors can be introduced. They are essentially an integrated convolution of the hard scattering kernel or partonic amplitude with the GPD. This convolution was already mentioned in equation 2.37. A general convention for these so-called compton form factors is [40]:

$$\mathcal{F} = \int_{-1}^{1} \mathrm{d}x \; F(x, \xi, t) \cdot \left(\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon} \right) \; (F = H, E), \tag{2.64}$$

$$\tilde{\mathcal{F}} = \int_{-1}^{1} \mathrm{d}x \; \tilde{F}(x, \xi, t) \cdot \left(\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon} \right) \; (\tilde{F} = \tilde{H}, \tilde{E}). \tag{2.65}$$

The angle ϕ is here the angle between the leptonic and the hadronic plane. With these the full twist-3 amplitude for DVCS can be written as [32]:

$$|\mathcal{T}_{\mathrm{DVCS}}|^2 = \frac{e^6}{y^2 Q^2} \left\{ \sum_{n=0}^{2} c_n^{\mathrm{DVCS}} \cos(n\phi) + \sum_{n=1}^{2} s_n^{\mathrm{DVCS}} \sin(n\phi) \right\}. \tag{2.66}$$

Here a twist-hierarchy is visible [41]:

1. $c_0(\mathcal{F}, \tilde{\mathcal{F}})^{\mathrm{DVCS}}$ twist-2;
2. $(c, s)(\mathcal{F}, \tilde{\mathcal{F}})_1^{\mathrm{DVCS}}$ twist-3: Longitudinal transverse interference;
3. $(c, s)(\mathcal{F}, \tilde{\mathcal{F}})_2^{\mathrm{DVCS}}$ bilinear combinations of ordinary twist-2 and gluon transversity terms.

The c_n and s_n are functions of the above-mentioned compton form factors $\mathcal{F}$ and $\tilde{\mathcal{F}}$.

When it comes to exclusive leptoproduction of photons, there is another process called Bethe-Heitler process, which exists with the same final and initial state as the DVCS process. Both of the processes are compared in Fig. 2.12. Due to the nature of quantum mechanics, processes with identical initial and final states interfere on an amplitude level. Therefore, the amplitudes for Bethe-Heitler, DVCS and the interference need to be considered for the complete cross section:

$$|\mathcal{T}|^2 = |\mathcal{T}_{\mathrm{DVCS}}|^2 + |\mathcal{T}_{\mathrm{BH}}|^2 + \mathcal{T}_{\mathrm{Interference}}, \tag{2.67}$$

where the interference term contains:

$$\mathcal{T}_{\mathrm{Interference}} = \mathcal{T}^*_{\mathrm{DVCS}} \mathcal{T}_{\mathrm{BH}} + \mathcal{T}^*_{\mathrm{BH}} \mathcal{T}_{\mathrm{DVCS}}. \tag{2.68}$$

The Bethe-Heitler term is fully calculable in quantum electrodynamics and therefore allows for normalizing or benchmarking. Its involvement via the interference term is of

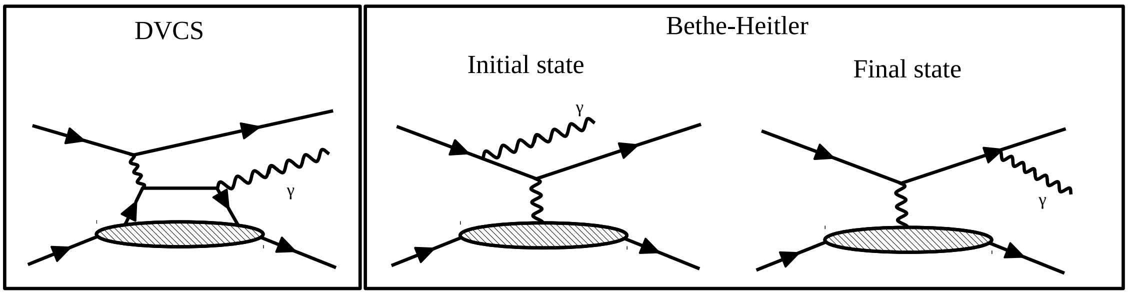


Figure 2.12: Processes with identical initial and final states: DVCS on the left, Bethe-Heitler on the right.

enormous importance, as this allows for further analysis of GPD quantities from the same data sample.

There are several interesting asymmetries that can be measured in order to put constraints on the GPDs involved. What makes them interesting is their dependency on the compton form factors and the GPDs. The $d\mathcal{M}$ is in the following [42]:

$$d\mathcal{M} = \frac{\alpha^3 x_{bj} y}{8\pi Q^2}\left(1 + \frac{4m_p^2 x_{bj}^2}{Q^2}\right)^{-1/2}. \tag{2.69}$$

The asymmetry for a polarized lepton and an unpolarized target in leading twist, also called single spin asymmetry or SSA in short, can be written as [42]:

$$\begin{aligned}\Delta_{SL} d\sigma &= d\sigma^{\uparrow} - d\sigma^{\downarrow}\\ &= -\frac{16(2-y)\sqrt{1-x_{bj}}}{\sqrt{1-y}x_{bj}\sqrt{-\Delta^2 Q^2}}\sin(\phi)\\ &\cdot \Im\left\{F_1\mathcal{H}_1 + \frac{x_{bj}}{2-x_{bj}}(F_1+F_2)\tilde{\mathcal{H}}_1 - \frac{\Delta^2}{4M^2}F_2\mathcal{E}_1\right\} d\mathcal{M}.\end{aligned} \tag{2.70}$$

The Compton form factors are defined in Eq. 2.64 and Eq. 2.65. The index notation SL stands for "Single Lepton" and is the usual abbreviation used in papers by D. Müller. There is no systematical nomenclature in this case, which is why [42] introduces these. The differential cross section $d\sigma^{\uparrow}$ denotes the result of leptons with their spin aligned in the momentum direction, whereas $d\sigma^{\downarrow}$ is the resulting cross section yielded by leptons with their spin aligned against their momentum direction. For the unpolarized target there is furthermore the charge asymmetry [42]:

$$\begin{aligned}\Delta_c^{\mathrm{unp}} d\sigma &= d^{+}\sigma^{\mathrm{unp}} - d^{-}\sigma^{\mathrm{unp}}\\ &= -\frac{16(2-2y+y^2)\sqrt{1-x_{bj}}}{\sqrt{1-y}yx\sqrt{-\Delta^2 Q^2}}\cos(\phi)\\ &\times \Re\left\{F_1\mathcal{H}_1 + \frac{x_{bj}}{2-x_{bj}}(F_1+F_2)\tilde{\mathcal{H}}_1 - \frac{\Delta^2}{4M^2}F_2\mathcal{E}_1\right\} d\mathcal{M}.\end{aligned} \tag{2.71}$$

For the COMPASS-II experiment the beam muons are produced via the pion decay. This production method generates the muons with a natural polarization according to their charge. Therefore, the spin and charge asymmetry for COMPASS-II need to be combined to a spin-charge asymmetry. The sum and the difference are both values of interest [43]:

$$\begin{aligned}\mathcal{D}_{CS,U} &= \mathrm{d}\overset{+}{\sigma}^{\leftarrow} - \mathrm{d}\overset{-}{\sigma}^{\rightarrow} = 2[P_\mu \mathrm{d}\sigma^{\mathrm{DVCS}}_{\mathrm{pol}} + e_\mu \Re(I)] \\ &\propto \left(\{s_1^{\mathrm{DVCS}} \sin(\phi)\}\right) + \left(c_0^I + c_1^I \cos(\phi) + \{c_2^I \cos(2\phi) + c_3^I \cos(3\phi)\}\right),\end{aligned} \tag{2.72}$$

$$\begin{aligned}\mathcal{S}_{CS,U} &= \mathrm{d}\overset{+}{\sigma}^{\leftarrow} + \mathrm{d}\overset{-}{\sigma}^{\rightarrow} = 2[\mathrm{d}\sigma^{\mathrm{BH}} + \mathrm{d}\sigma^{\mathrm{DVCS}}_{\mathrm{unpol}} + e_\mu P_\mu \Im(I)] \\ &\propto 2[\mathrm{d}\sigma^{\mathrm{BH}}] + \left(c_0^{\mathrm{DVCS}} + \{c_1^{\mathrm{DVCS}} \cos(\phi) + c_2^{\mathrm{DVCS}} \cos(2\phi)\}\right) \\ &\qquad + (s_1^I \sin(\phi) + \{s_2^I \sin(2\phi)\}).\end{aligned} \tag{2.73}$$

These can be used to extract several interesting quantities. From the sum $\mathcal{S}_{CS,U}$ the differential cross section in t and the so-called t-slope B can be extracted:

$$\mathrm{d}\sigma^{\mathrm{BH}} + \mathrm{d}\sigma^{\mathrm{DVCS}}_{\mathrm{unpol}} \rightarrow \frac{\mathrm{d}\sigma^{\mathrm{DVCS}}}{\mathrm{d}|t|} \propto \exp(-B|t|). \tag{2.74}$$

Using the definitions for the fourier coefficients in the cross sections in [32], their main dependence can easily be condensed to:

- $c_0^I \propto \Re\left\{C^I_{\mathrm{unp}} + \frac{\Delta^2}{Q^2}(C^I_{\mathrm{unp}} + \Delta C^I_{\mathrm{unp}})\right\}$;
- $c_1^I \propto \Re\left\{C^I_{\mathrm{unp}}\right\}$;
- $s_1^I \propto \Im\left\{C^I_{\mathrm{unp}}\right\}$.

The internal coefficients C are used for these fourier coefficients, as in [32]:

$$C^I_{\mathrm{unp}} = F_1 \mathcal{H} + \frac{x_{bj}}{2 - x_{bj}} \tilde{\mathcal{H}} - \frac{\Delta^2}{2m_p^2} F_2 \mathcal{E}. \tag{2.75}$$

The main contribution is clearly the compton form factor $\mathcal{H}$, of which the imaginary and the real part can be measured independently via the different fourier coefficients from the sum and difference. All further form factors are kinematically suppressed by at least $\frac{x_{bj}}{2-x_{bj}}$.

Coming back to the spin puzzle, all of this can be used to put constraints on the total angular momentum of quarks and gluons. Results from both experiments and lattice calculations are shown in the Fig. 2.13. The experimental measurements from JLAB Hall A and the HERMES collaboration are in good agreement with the results from lattice calculations.

2.7 Hard Exclusive Meson Production

The hard exclusive meson production, or HEMP in short, is another vital tool of constraining GPDs. The production of vector mesons such as ρ^0, $\rho^\pm$, ω, ϕ or J/ψ is possible as well as pseudoscalar mesons like the π^0, $\pi^\pm$, $K^\pm$ and η. The handbag scheme, that is used for factorization is shown in Fig. 2.14. In general, the meson production is interesting for GPD constraints because the different GPDs enter in different combinations for

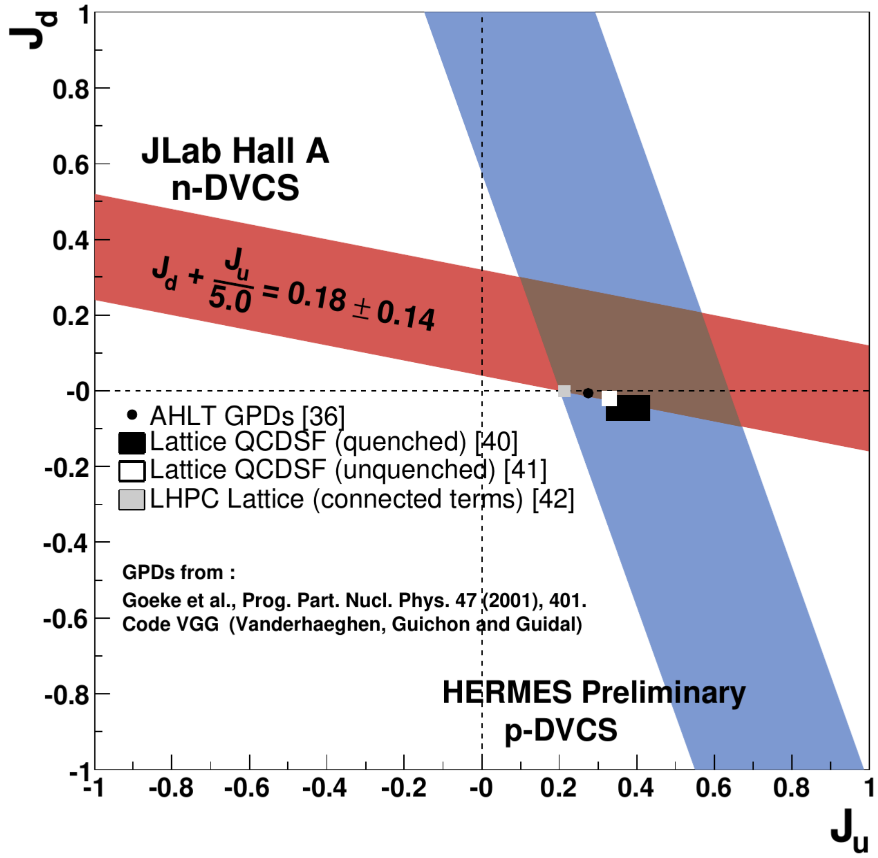


Figure 2.13: Constraints on the total angular momenta of valence u and d quarks from lattice calculations and experiments [44].

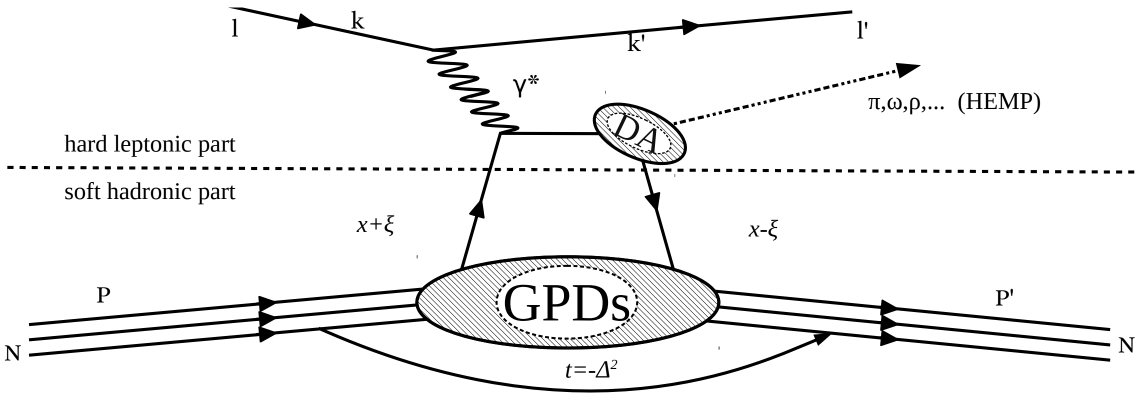


Figure 2.14: Hard exclusive meson production in handbag scheme. The meson distribution amplitude is written as DA.

each meson. On the theoretical side, the meson production is less clean than the DVCS process. This is due to the occurence of a second non-perturbative quantity: the meson distribution amplitude, or DA in short. This amplitude describes the coupling of the final state meson to the intermediate $q\bar{q}$ pairs or gluons. They also depend on the factorization constant μ_f and require a finite momentum transfer t to the nucleon.

In theoretical physics, the subprocess of virtual photoproduction is discussed regularly:

$$\gamma^* N \to N' V. \tag{2.76}$$

In the HEMP case the theory needs to be evolved to twist-3. For the pseudoscalar mesons there are only small contributions from $\tilde{H}$ and $\tilde{E}$, but the transversity GPDs $\bar{E}_T$ and H_T contribute to a great extent, since they are enhanced by the chiral condensate [45]. The transverse cross section for π^0 production is about an order of magnitude larger than the longitudinal one. For vector mesons measurements and predictions exist. Some of them are shown in Fig. 2.15.

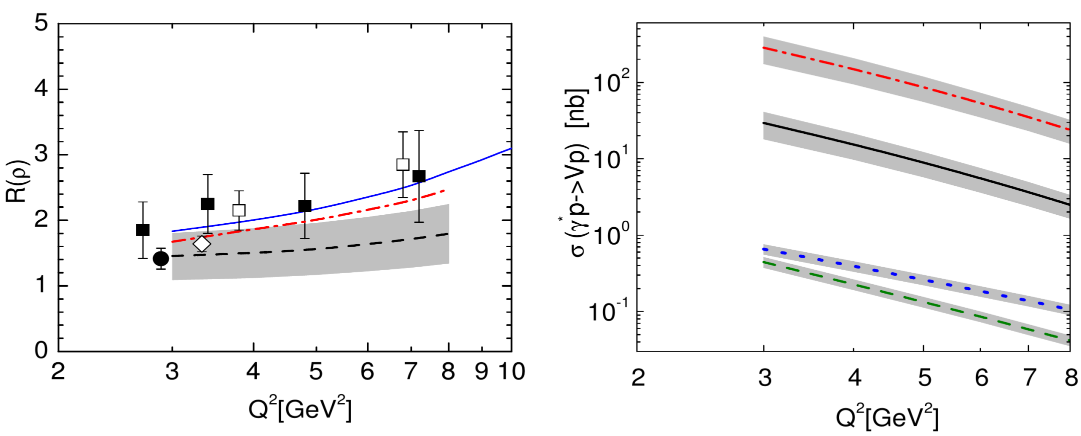


Figure 2.15: Left: The ratio $R(\rho)$ of longitudinal and transverse cross section for ρ^0 production. The full line is a prediction for HERA, dashed-dotted for COMPASS and dashed for HERMES.
Right: The integrated cross sections for the vector mesons: ρ^0, ω, ρ^+, K^{*0}[46].

2.7.1 Cross Section for a Transversely Polarized Target

Scattering muons off a polarized target is a typical measurement for COMPASS. There are ongoing studies with the aim of enabling COMPASS-II to measure with a polarized target and a recoil proton detector at the same time. Here the cross section and asymmetries are introduced following Diehl and Sapeta [47]. First the target spin vector $\vec{S}$ is defined as [47]:

$$\vec{S} = \begin{pmatrix} S_T \cos(\phi - \phi_S) \\ S_T \sin(\phi - \phi_S) \\ S_L \end{pmatrix}. \tag{2.77}$$

For a visualization of the angles see Fig. 2.16. The cross section can be written as a

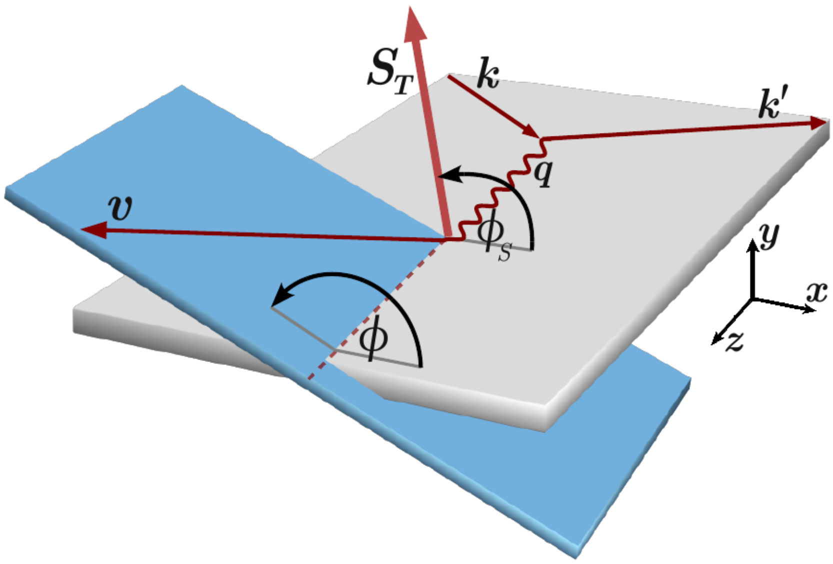


Figure 2.16: Angle definitions according to the Trento convention [48] [49].

contraction of a leptonic and a hadronic tensor, which is also done for the inclusive and semi-inclusive cross sections:

$$\mathrm{d}\sigma(lp \to lhX) \propto L^{\mu\nu} W_{\mu\nu} \frac{\mathrm{d}^3 l'}{2l'^0} \frac{\mathrm{d}^3 P_h}{2P_h^0}, \tag{2.78}$$

where the leptonic tensor can be written as:

$$L^{\mu\nu} = l'^\nu l^\mu + l^\nu l'^\mu - (l' \cdot l) g^{\nu\mu} + i P_l \epsilon^{\nu\mu\alpha\beta} q_\alpha l_\beta. \tag{2.79}$$

The convention is $\epsilon^{0123} = 1$ and the lepton beam polarization P_l is defined as $P_l = +1$ for a purely right handed beam and $P_l = -1$ for a purely left handed beam [47].

The spin density matrix is given by:

$$\rho_{ji} = \frac{1}{2}\left[\delta_{ji} + \vec{S} \cdot \sigma_{ji}\right] = \frac{1}{2}\begin{pmatrix} 1 + S_L & S_T \exp(-i(\phi - \phi_S)) \\ S_T \exp(i(\phi - \phi_S)) & 1 - S_L \end{pmatrix}, \tag{2.80}$$

with the hadronic tensor given by [47]:

$$W_{\mu\nu} = \sum_{i,j} \rho_{j,i} \sum_X \delta^4(P' + P_h - P - q) \sum_{\text{spins}} \langle p(i)|J_\mu(0)|hX\rangle\langle hX|J_\nu(0)|p(j)\rangle, \tag{2.81}$$

where J_μ is the electromagnetic current and $\sum_X$ denotes the integral over all hadronic final states. The final cross section can be calculated from these, but it is easier to express it in terms of the quantities:

$$\sigma_{mn} = \sum_{i,j} \rho_{ji} \sigma^{ij}_{mn} \propto \int \mathrm{d}t \mathrm{d}M_X^2 (\epsilon_m^{\mu *} W_{\mu\nu} \epsilon_n^{\nu}), \tag{2.82}$$

where the ϵ are the polarization vectors for the definite helicity m of the virtual photon that can be written as [47]:

$$\epsilon_0^\mu = \frac{1}{Q\sqrt{1+\gamma^2}} \left(q^\mu + \frac{Q^2}{P \cdot q} P^\mu \right), \tag{2.83}$$

$$\epsilon_{\pm 1} = \frac{1}{\sqrt{2}} (0, \mp 1, i, 0). \tag{2.84}$$

The σ^{ij}_{mn} in equation 2.82 are polarized photon interference terms that obey [47]:

$$\sigma^{ij}_{mn}(x_{bj}, Q^2) \propto \int \mathrm{d}t \mathrm{d}M_X^2 \sum_X \delta^{(4)}(P' + P_h - P - q) \sum_{\text{spins}} \left(\mathcal{A}^i_m\right)^* \mathcal{A}^j_n. \tag{2.85}$$

From hermiticity and parity invariance the following relations follow [47]:

$$\sigma_{nm} = \sigma^*_{mn}, \tag{2.86}$$

$$\sigma^{ji}_{nm} = (\sigma^{ij}_{mn})^*, \tag{2.87}$$

$$\sigma^{-i-j}_{-m-n} = (-1)^{m-n-i+j} \sigma^{ij}_{mn}. \tag{2.88}$$

With the above definitions, the master equation for polarized hard exclusive meson production can be derived [47]:

$$\begin{aligned}
&\left(\frac{\alpha}{8\pi^3} \frac{y^2}{1-\varepsilon} \frac{1-x_{bj}}{x_{bj}} \frac{1}{Q^2} \right)^{-1} \frac{\mathrm{d}\sigma}{\mathrm{d}x_{bj} \mathrm{d}Q^2 \mathrm{d}\phi \mathrm{d}\phi_s} \\
&= \frac{1}{2}\left(\sigma^{++}_{++} + \sigma^{--}_{++}\right) + \varepsilon \sigma^{++}_{00} - \varepsilon \cos(2\phi) \Re(\sigma^{++}_{+-}) - \sqrt{\varepsilon(1+\varepsilon)} \cos(\phi) \Re(\sigma^{++}_{+0} + \sigma^{--}_{+0}) \\
&- P_l \sqrt{\varepsilon(1-\varepsilon)} \sin(\phi) \Im(\sigma^{++}_{+0} + \sigma^{--}_{+0}) \\
&- S_L \left[\varepsilon \sin(2\phi) \Im(\sigma^{++}_{+-}) + \sqrt{\varepsilon(1+\varepsilon)} \sin\phi \Im(\sigma^{++}_{+0} - \sigma^{--}_{+0}) \right] \\
&+ S_L P_l \left[\sqrt{1-\varepsilon^2} \frac{1}{2} (\sigma^{++}_{++} - \sigma^{--}_{++}) - \sqrt{\varepsilon(1-\varepsilon)} \cos\phi \Re(\sigma^{++}_{+0} - \sigma^{--}_{+0}) \right] \\
&- S_T \Bigg[\sin(\phi - \phi_S) \Im(\sigma^{+-}_{++} + \varepsilon \sigma^{+-}_{00}) + \frac{\varepsilon}{2} \sin(\phi + \phi_S) \Im(\sigma^{+-}_{+-}) + \frac{\varepsilon}{2} \sin(3\phi - \phi_S) \Im(\sigma^{-+}_{+-}) \\
&\quad + \sqrt{\varepsilon(1+\varepsilon)} \sin(\phi_S) \Im(\sigma^{+-}_{+0}) + \sqrt{\varepsilon(1+\varepsilon)} \sin(2\phi - \phi_S) \Im(\sigma^{-+}_{+0}) \Bigg] \\
&+ S_T P_l \Bigg[\sqrt{1-\varepsilon^2} \cos(\phi - \phi_S) \Re(\sigma^{+-}_{++}) \\
&\quad - \sqrt{\varepsilon(1-\varepsilon)} \cos(\phi_S) \Re(\sigma^{+-}_{+0}) - \sqrt{\varepsilon(1-\varepsilon)} \cos(2\phi - \phi_S) \Re(\sigma^{-+}_{+0}) \Bigg],
\end{aligned} \tag{2.89}$$

where the target spins are written as $+$ and $-$ instead of $\pm\frac{1}{2}\hbar$ for readability and ε is the virtual photon polarization defined as:

$$\varepsilon = \frac{1 - y - \frac{1}{4}y^2\gamma^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}y^2\gamma^2}. \tag{2.90}$$

The γ is a kinematical function given to:

$$\gamma = \frac{2m_p x_{bj}}{Q}. \tag{2.91}$$

It is also common to re-write the cross section components in terms of their coupling to virtual photon polarizations:

$$\sigma_T = \frac{1}{2}(\sigma_{++}^{++} + \sigma_{++}^{--}), \tag{2.92}$$

$$\sigma_L = \sigma_{00}^{++}. \tag{2.93}$$

With these the total unpolarized cross section can be written as:

$$\sigma = \sigma_T + \varepsilon\sigma_L = \frac{1}{2}(\sigma_{++}^{++} + \sigma_{++}^{--}) + \varepsilon\sigma_{00}^{++}. \tag{2.94}$$

From equation 2.89 several measurable asymmetries can be derived. They are especially valuable since experimental access to asymmetries is easier than to total cross sections. In the case of a transversely polarized proton target, there is a total of eight meaningful asymmetries:

$$\begin{aligned}
A_{UT}^{\sin(\phi-\phi_S)} &= -\frac{\Im(\sigma_{++}^{+-} + \varepsilon\sigma_{00}^{+-})}{\sigma_0}, & & \\
A_{UT}^{\sin(\phi+\phi_S)} &= -\frac{\Im(\sigma_{+-}^{+-})}{\sigma_0}, & A_{LT}^{\cos(\phi-\phi_S)} &= \frac{\Re(\sigma_{++}^{+-})}{\sigma_0}, \\
A_{UT}^{\sin(3\phi-\phi_S)} &= -\frac{\Im(\sigma_{+-}^{-+})}{\sigma_0}, & A_{LT}^{\cos(\phi_S)} &= -\frac{\Re(\sigma_{+-}^{+0})}{\sigma_0}, \\
A_{UT}^{\sin(\phi_S)} &= -\frac{\Im(\sigma_{+0}^{+-})}{\sigma_0}, & A_{LT}^{\cos(2\phi-\phi_S)} &= \frac{\Re(\sigma_{+0}^{-+})}{\sigma_0}. \\
A_{UT}^{\sin(2\phi-\phi_S)} &= -\frac{\Im(\sigma_{+0}^{-+})}{\sigma_0}, & &
\end{aligned} \tag{2.95}$$

The longitudinal asymmetries can also be taken into account. They are introduced in [47].

2.8 GPD Constraints

Apart from the already mentioned mathematical constraints of the GPDs there are also some physical ones that can be placed on them. A full set of these constraints is given in Tab. 2.3. Since the GPDs have many parameters, a fully experimental measurement cannot be done in the near future. This is especially true for those GPDs that do not have a classical forward limit to relate them to, as those tend to be measured very well already.

Table 2.3: Contraints for GPDs, taken and adapted from [50]. For comparison the PDFs are assigned a five star rating. A_{UT} means the asymmetry of exclusive production with an unpolarized lepton beam and a transversely polarized nucleon. A_{LL} is the asymmetry of exclusive production with a longitudinally polarized lepton beam and longitudinally polarized nucleons.

GPD	**probed by**	**constraints**	**status**
H_{val}	ρ^0,ϕ cross section	PDFs, Dirac form factors	***
$H_{g,sea}$	ρ^0,ϕ cross section	PDFs	***
E_{val}	$A_{UT}^{sin(\phi-\phi_S)}(\rho^0,\phi)$	Pauli form factors	**
$E_{g,sea}$	-	sum rule for second moments	-
$\tilde{H}_{val}$	π^+ data	pol. PDFs, axial form factors	**
$\tilde{H}_{g,sea}$	$A_{LL}(\rho^0)$	pol. PDFs	*
$\tilde{E}_{val}$	π^+ data	pseudoscalar form factors	*
H_T, $\bar{E}_T(val)$	π^+ data, $A_{UT}^{sin(\phi_S}(\rho^0)$	transversity PDFs	*

The Tab. 2.3 also shows them as highly unreliable. It is, however, possible to model them in different ways allowing for fitting from the presently available data. There are two important approaches to modeling the GPDs. The first one is the so-called double distribution [51], or DD in short, ansatz. With this approach, the description of the GPD starts with a GPD in the forward limit and adds the t and ξ dependence via additional terms. This ansatz breaks QCD evolution and restricts the GPDs slightly too much, but still, the double distribution GPDs are widely used for their simplicity. Typically, the ansatz for the double distribution looks like:

$$F(\alpha,\beta) = h(\alpha,\beta)q(\beta), \tag{2.96}$$

where q is the quark distribution and h is the so-called profile function. A very important model based on the double distribution ansatz is the model bei Goloskokov and Kroll. A part of this thesis presents the complete implementation of the corresponding calculations for neutral pion production into an event generator. It will be introduced in Chapter 6. An alternative to this is the so-called dual representation[52], where the model for the GPD is based on a partial wave expansion of the t channel exchanges.

Measuring exclusive processes to a very high accuracy is the key tool to constraining the GPDs. Sums and differences of the cross sections give access to the interesting quantities of the theory. Many experimental issues have to be overcome to accomplish exclusive measurements within the needed accuracy. The COMPASS-II experiment, which is introduced in the next chapter, was designed to meet all the requirements needed for a successful measurement of these rare exclusive processes.

3. The COMPASS-II Experiment at CERN

The COMPASS-II experiment, which is a direct successor to the COMPASS experiment, is a fixed target experiment with the possibility for many different beams and targets. It is located in the north area at CERN at the end of the M2 beam line. The primary beam is provided by the SPS collider as a proton beam, which is converted into many other particles by hadronic interactions at the T6 target into. For an overview over the CERN accelerator complex, see Fig. 3.1.

The main goal of the COMPASS-II experiment are spin structure studies of hadrons and their spectroscopy. The program in 2012, which was a pilot run for the larger 2016 run, specifically focused on researching the spin structure of the proton. To gain access to this complicated quantity, exclusive processes need to be measured to a very high precision. In order to obtain these measurements, a polarized muon beam, that is described in Sec. 3.1, is scattered off of a liquid hydrogen. The complete target region is introduced in the Sec. 3.2. For an exclusive measurement of such production processes the reaction products need to be measured to a high precision to ensure real exclusivity. The COMPASS-II spectrometer, in detail explained in Sec. 3.3, is capable of doing so. It features a good particle identification, a large angular acceptance, calorimetry and large array of different tracking detectors. To save the produced data a sophisticated data acquisition system is needed. This system is described in Sec. 3.4 together with the reconstruction software and the data flow.

The description will mainly focus on the data taking in the years 2012 and 2016/2017 that are conducted using a polarized muon beam. A general description of the capabilities of the experimental setup is given in [53, 54].

3.1 The Polarized Beam

Beam Line

The SPS is delivering protons to the LHC ring and the various other experiments. With the M2 beam line, shown in Fig. 3.2, the primary proton beam can be converted to different

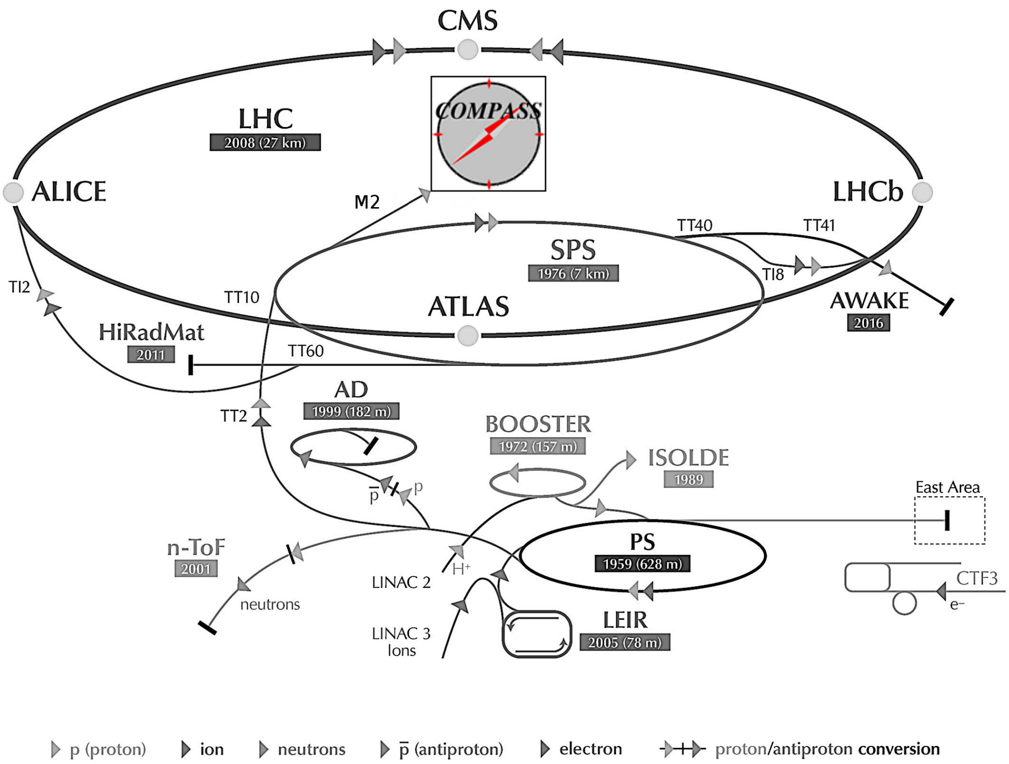


Figure 3.1: The CERN accelerator complex. COMPASS-II is located in the north area at the end of the M2 beam line [55].

secondary and tertiary beams. The protons, at a momentum of about 400 GeV/c, are scattered on the T6 target, a plate made of beryllium with a thickness of 50 cm. This block is interchangeable with shorter ones, resulting in lower fluxes at the COMPASS-II experiment for detector testing. The protons are typically delivered in spills of 9.6 s length containing an average of $1.2 \cdot 10^{13}$ protons. In an SPS supercycle of about 40 s length, one to three of these spills are emitted by the SPS. The number of spills and also the duration of the supercycle depend on the workload of other experiments and can change within minutes. The protons interact hadronically with the beryllium target creating different secondary particles, mainly pions and kaons. These pions and kaons can reach up to 280 GeV/c and are directly used for the COMPASS hadron programm. For the GPD programm, the mesons are guided through a decay tunnel, which is about 600 m in length. There the pions decay weakly to muons and (anti-) neutrinos. Due to the nature of the weak decay, being maximal parity violating, these muons are intrinsically polarized. To keep the beam well focused, the decay tunnel is equipped with alternating focusing and defocusing (FODO) quadrupole magnets. The polarization is not 100% though, since the mass of the muon leads to a part of the muons being polarized in the opposite direction. The polarization is momentum dependent and shown in Fig. 3.3. All remaining hadrons are absorbed by the hadron absorbers, which consists of seven to nine further beryllium blocks, each of 1.1 m length. After this, a set of deflecting magnets are installed. These are combined to the first red triangle in Fig. 3.2 and are used for momentum and charge separation of the beam particles. As a good compromise of momentum, polarization and intensity, a nominal beam momentum of 160 GeV/c is chosen for the muon program, as the intensity weakens with higher beam momenta.

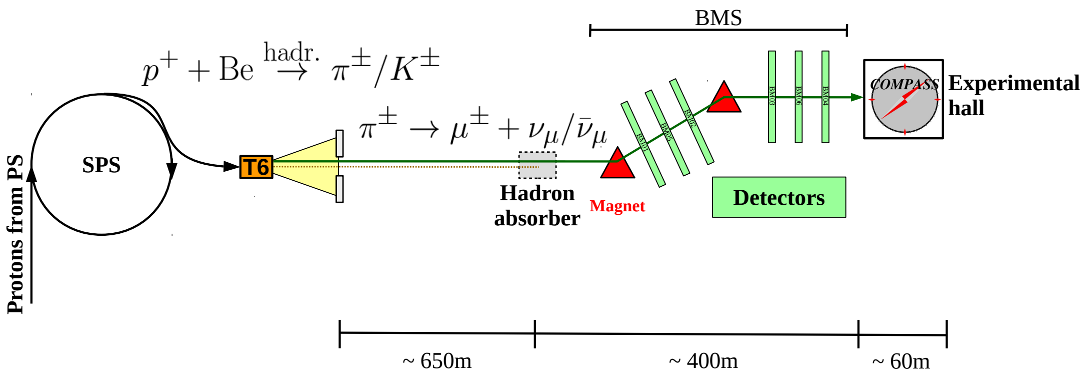


Figure 3.2: M2 beam line including the BMS, based on [56].

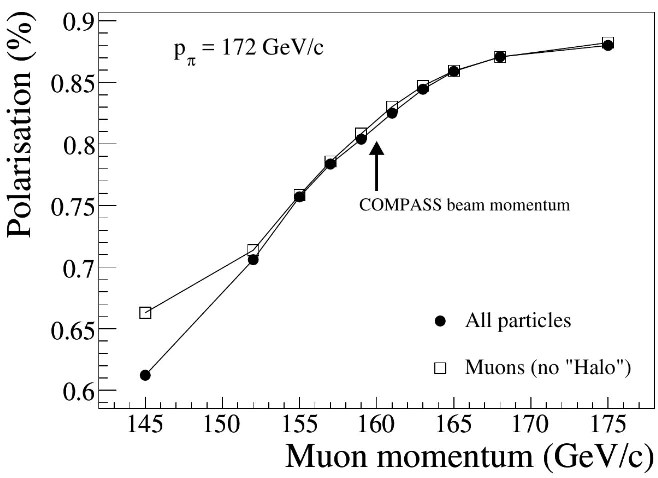


Figure 3.3: Beam polarization against muon momentum [53].

Beam Momentum Station

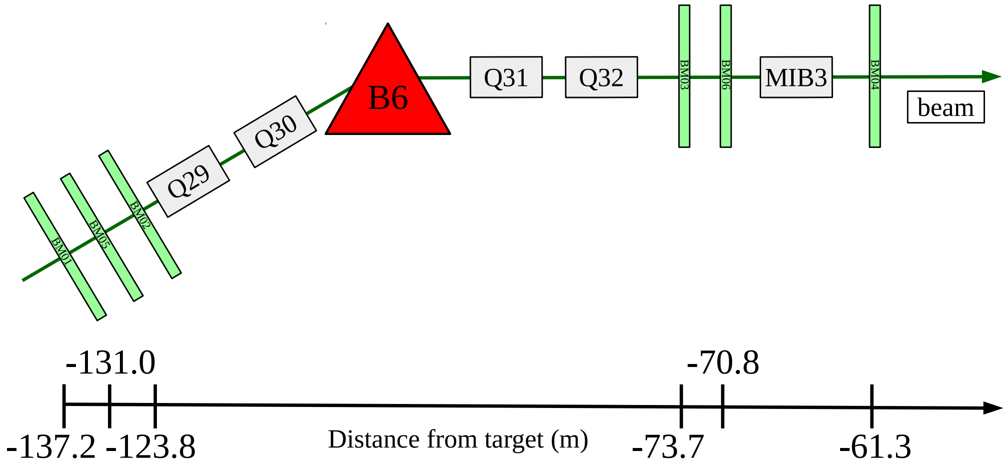


Figure 3.4: Schematic of the COMPASS beam momentum station, based on [53].

In order to conduct high-precision exclusive measurements, an accurate determination of the incoming muons' momentum is necessary for every incoming particle. This is achieved with the so-called beam momentum station, or BMS in short. The principle setup is shown in Fig. 3.4. The station consists of four scintillator hodoscopes (BM01 to BM04) with fast photomultiplier read-outs, which can deliver time resolutions up to 0.3 ns [53]. Two additional scintillating fibre hodoscopes (BM05, BM06) help to cope with the multi hit resolution in the BMS and to back-up the detection efficiency in the BMS. The hits from detectors in front of the target are also used for the final reconstruction, thus further increasing the reconstruction efficiency and accuracy.

3.2 Target Region

The target region is a very important part of the COMPASS-II spectrometer, as it is not only the main point of interaction but also the region where recoiled protons are detected. The target used for the exclusive measurements is a 2.5 m long tube, filled with liquid hydrogen. Its diameter is 40 mm. A very quite feature of the target is the usage of very little material. This is important due to the fact that the recoiled protons from exclusive processes have a very small momentum and measuring them is only possible if they leave the target. Therefore, the transverse material budget impacts the overall experiments' efficiency directly. The tube holding the hydrogen is made of 0.35 mm thick mylar foil, whereas the vacuum chamber is a carbon fiber tube of 1 mm thickness. The whole target scheme is shown in 3.5. A short review of its performance is given in [57].

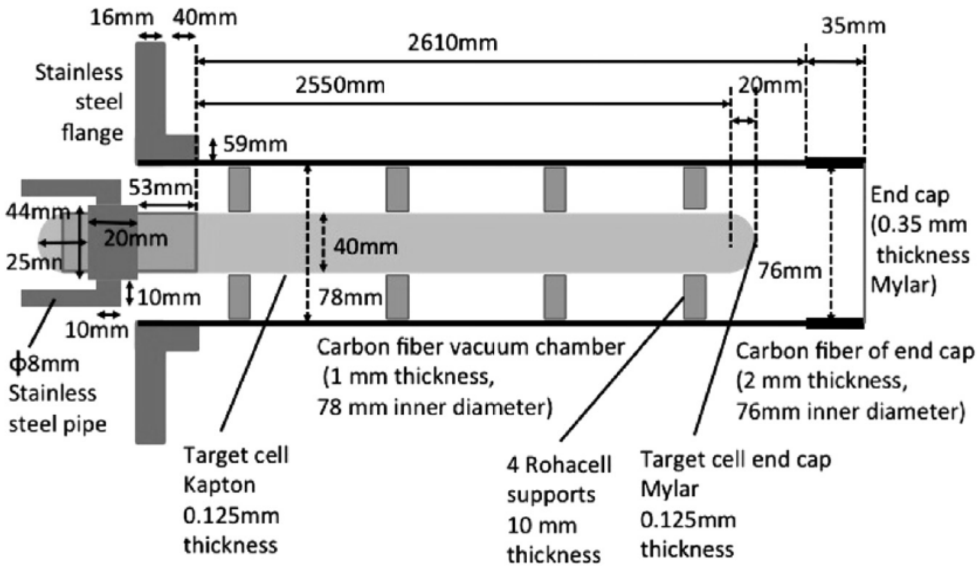


Figure 3.5: Schematic of the COMPASS-II target from [57].

Recoil Proton Detector: CAMERA

Since the resolution of the spectrometer is not sufficient to guarantee exclusivity, the recoiled proton needs to be measured indepedently. A new detector surrounding the target was installed in 2012. It is called CAMERA[1] and it consists of two rings, each made of

[1]**C**ompass **A**pparatus for the **M**easurement of **E**xclusive **R**e**A**ctions

24 scintillator slats of 275 cm and 360 cm length for ring A and B respectively. Every slat has two read-out photomultiplier tubes, one upstream and one downstream of the target. A time of flight measurement by this detector allows for an exact measurement of the momentum transfer t. The detector is shown in Fig. 3.6 with the LH2 target in place. To allow

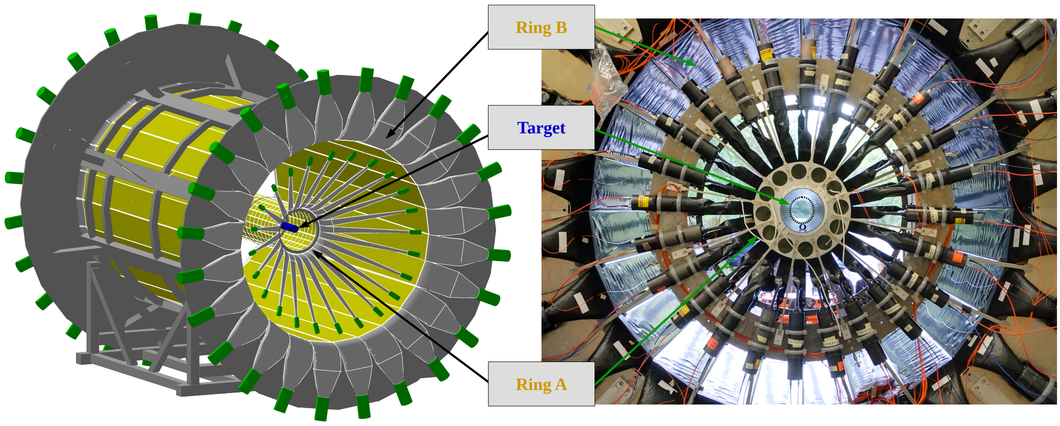


Figure 3.6: CAMERA as implemented in TGEANT (left), the new Geant4 based COMPASS-II Monte-Carlo with a target. A photograph of the real detector is shown in the right. The dotted line in the center indicates where the target is fitted in, as it is absent in this picture.

for a time of flight measurement, the protons also need to reach ring B. This sets a limit on the thickness of the scintillator material in ring A, which was chosen to be only 4 mm thin. This thickness allows most protons to pass, but at the same time limits the amount of light produced by the scintillation. The low amount of scintillation photons makes the detection more difficult. After the test run in 2012, in which the ring A scintillators could not reach the desired efficiency, all ring A scintillators were exchanged for new, high quality slats.

3.3 Spectrometer

The remaining part of the COMPASS-II experiment is the spectrometer. It consists of two stages, each equipped with a spectrometer magnet and surrounded by tracking detectors for charge and momentum measurements. The experiment also features a ring imaging Cherenkov detector(RICH-1) for particle identification, as well as a trigger system with target pointing which serves to show whether an event was of interest for analytical purposes. To measure the absolute energy of hard photons that are undetectable by tracking detectors, three electromagnetic calorimeters are used. For completeness of the energy measurement, and possibly a trigger decision, two hadronic calorimeters are placed behind the electromagnetic ones. At the very end of each stage for muon identification another round of trackers and hadron absorbers, called muon filters, are installed. These systems of detectors and absorbers are called muon walls. An image containing the whole spectrometer with all detectors, as used in 2016, is shown in Fig. 3.8. Both spectrometer parts are separately shown in Fig. 3.7, labeled with the particular detector names.

The two stage construction allows for a wide kinematic range to be covered while at the same time having a large angular acceptance.

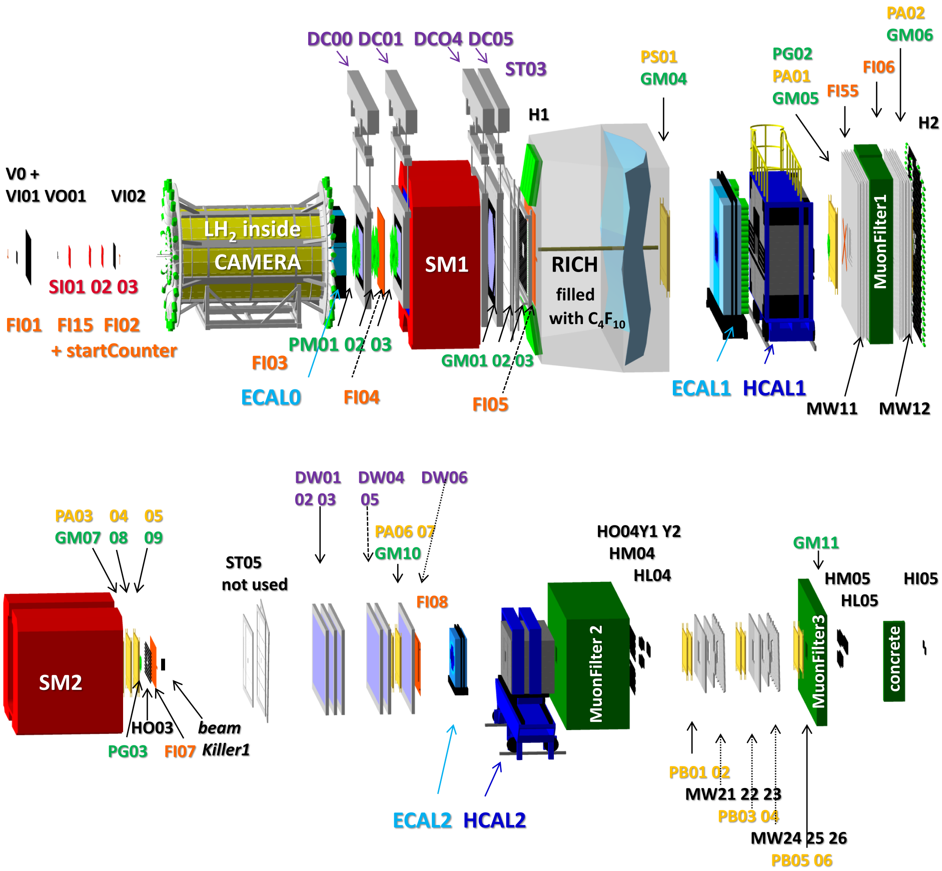


Figure 3.7: The spectrometer parts LAS and SAS with labels for each detector and part. The upper part is the target area with the LAS following, the lower part is the SAS. Please note that the scaling of both spectrometer parts is not the same for graphical reasons. For the correct scaling see Fig. 3.8. Both of these figures show the 2016 setup.

3.3.1 Tracking

For the COMPASS-II experiment, a variety of different tracking detectors are used. In the regions near the beam the particle flux is five orders of magnitude higher than in the outmost regions. Therefore, the different types of detectors can be grouped into different classes regarding their spatial coverage.

3.3.1.1 VSAT

The class of **VSAT** – very small area trackers – are detectors that can be placed innermost, in the beam. They need to be able to tolerate the highest rates, while at the same time offering the highest spatial or time resolution. At a particle rate of up to $10^5\,\mathrm{mm}^{-2}\,\mathrm{s}^{-1}$ [53] these detectors offer time resolutions of up to 3 ns or spatial resolutions of up to 100 μm. The detectors used are scintillating fibres (SciFi) and silicon micro strip detectors (SI). The SciFi stations are used in the BMS as well as upstream of the target to track the incoming muon. The SI stations are also placed upstream of the target for a precise measurement of the beam direction. In the latter years, pixelized micro-mesh gaseous detectors (PMM) and pixelized gaseous electron multipliers (PGEM) were added. They were fitted inside the **SAT** detectors of the respective kind to increase the resolution of tracking for the lesser

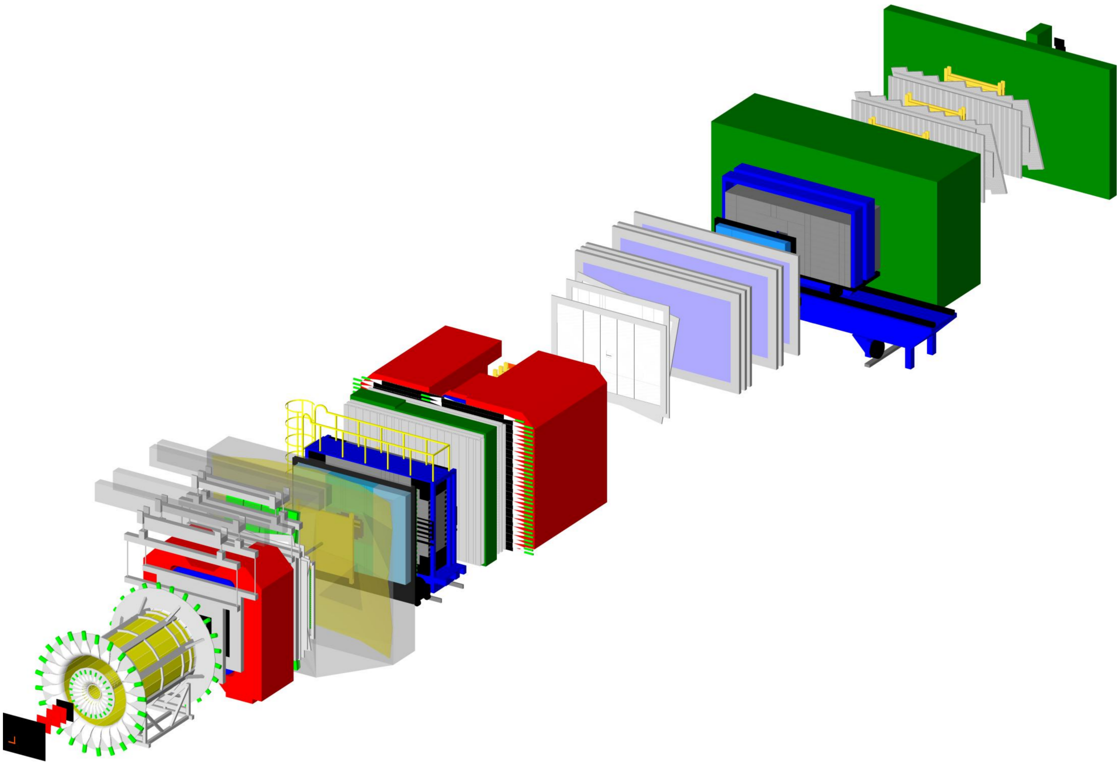


Figure 3.8: The complete COMPASS-II spectrometer for the 2016 run taken from TGEANT.

bent particles that enter the SAS afterwards. The pixel GEMs have a spatial resolution of 95 μm and a time resolution of 10 ns. The pixelized micromegas offer a time resolution of 9.1 ns and a spatial resolution of 79 μm.

3.3.1.2 SAT

The class of **SAT** – small area trackers – are covering the radii from 3 cm up to 40 cm away from the beam. They offer a good compromise between covered area, spatial resolution and time resolution. When it comes to this class of trackers, two types of detectors are used. The first are the micro-mesh gaseous detectors (Micromegas/MM). Their typical resolutions are 70 μm and 10 ns. The second detector type of this class are the gaseous electron multipliers (GEM). These detectors typically achieve resolutions of about 50 μm and 12 ns. Some of these detector planes are equipped with their respective pixelized counterparts in the center.

3.3.1.3 LAT

The last group of detectors is abbreviated with **LAT**, meaning large area trackers. These detectors offer active areas of several m^2. The detectors used for the large area coverage are the drift chambers (DC) and the multi-wire proportional chambers (MWPC). The last type of large area trackers are the so-called straw detectors. They are composed of many gas-filled conducting tubes with a wire in their center, which is differently charged from the outer tube. They are used for the tracking of particles that are heavily bent away from the beam axis.

3.3.1.4 Muon Identification

The muon identification is described in this section since it is done via tracking. The spectrometer has an absorber block at the end of each spectrometer stage. At the end of the LAS the muon filter 1 (MF1) is placed, which is made of a 60 cm thick iron plate. Together with tracking planes in front and behind of the material, the system is called the muon wall 1 (MW1). If a track can be extrapolated through the absorber, this track is usually identified as a muon, as they are the only particles that can reliably cross such absorption and radiation lengths. At the end of the SAS, the muon wall 2 (MW2) system is placed around the next absorber, the muon filter 2 (MF2). It is a concrete block of 2.4 m thickness. The last material block, called the muon filter 3 (MF3), is placed in front of the H5 hodoscope stations. Since there are no real tracking detectors following, it is not a muon wall. This filter, which is made of iron, makes sure, that the inner trigger, which is described in Sec. 3.3.6, is only activated by muons.

3.3.2 Calorimetry

The COMPASS-II experiment uses three electromagnetic calorimeters (ECAL) to measure the energy of photons and electrons, whereas for the measurement of the energy of hadrons two hadronic calorimeters are used. Some of the triggers, namely the inner trigger and the ladder trigger, used to require a signal in the HCAL, but this is not the case anymore in 2012 and 2016.

3.3.3 Electromagnetic Calorimeters

The first calorimeter in the spectrometer is the **ECAL0**, which is placed directly behind the CAMERA detector. It features solely the newly designed and constructed shashlik modules, which are built from alternating layers of scintillator and lead. The scintillators are connected by lightguides which are then read-out by a photomultiplier. The setup for ECAL0 changed from 2012 to 2016. It was decided to enlarge it even more, adding more shashlik modules. Both versions are shown in Fig. 3.9. The total module number was increased from 564 in 2012 to 1746 in 2016. The shashlik blocks used for building the ECAL0 feature 109 layers of lead and scintillator material and hence are shorter versions of the sampling modules used in ECAL1 and ECAL2, which feature 155 layers. A special energy dependant correction algorithm is used at reconstruction time to take into account lost energy. Another specialty of the ECAL0, is that the modules are read out in module blocks of 3×3 modules. This allows the employment of Micro-pixel Avalanche Photo Diodes (MAPD), which is a very new technique [58]. The total dimension of the ECAL0 was (132×110) cm in 2012 and is (204×206) cm in 2016, where the first number denotes the width and the second number is the height.

The **ECAL1** is located in the end of the LAS. It consists of several module types of lead glass, named GAMS, MAINZ and OLGA [54]. In addition, some parts of ECAL1 are equipped with shashlik modules. Their working principle and name are identical. The complete ECAL1 is shown in Fig. 3.10. The total modules of this calorimeter are 232 shashlik, 584 GAMS, 572 MAINZ and 320 OLGA, which sum up to a total of 1708 channels. The height of the ECAL1 is 286 cm, whereas the width is 395 cm.

The last electromagnetic calorimeter is **ECAL2**, which is located in the end of the SAS. It consists of of two types of lead glass modules called GAMS and GAMSRH, where the

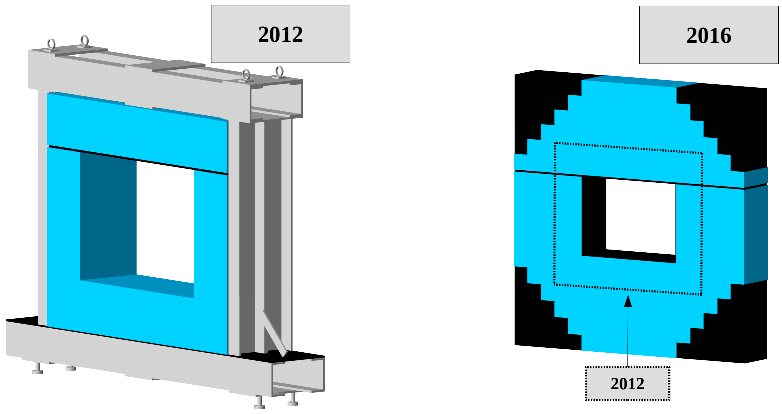


Figure 3.9: ECAL0 in the years 2012 and 2016 from TGEANT. All modules are shashlik ECAL0 type. The dashed line in the 2016 version of the ECAL shows the original 2012 modules which are now embedded.

Table 3.1: Modules used in the electromagnetic calorimeters for the COMPASS-II experiment.

Block type	**Front size** $[\text{cm}]^2$	**Material**
GAMS	3.82×3.82	TF1
GAMS RH	3.82×3.82	TF101
MAINZ	7.5×7.5	SF57
OLGA	14.3×14.3	SF5
Shashlik (ECAL0)	4.0×4.0	$109 \times$ (1.5 mm Sc. + 0.8 mm Pb)
Shashlik (ECAL1,2)	3.82×3.82	$155 \times$ (1.5 mm Sc. + 0.8 mm Pb)

RH stands for radiation hard. In the central region, likewise to ECAL1, shashlik modules are placed [54]. The shashlik and GAMS modules of ECAL2 are the same as the ones used in ECAL1. This calorimeter is shown in Fig. 3.11. In total 764 shashlik, 1440 GAMS and 768 GAMSRH blocks were used to construct this calorimeter which adds up to 2972 active channels. Because the ECAL2 is used to detect particles that are emitted under smaller angles than the other ECALs, it is mainly built from small modules. This ensures a good energy resolution for particles emitted under small angles with respect to the beam direction. The usage of these smaller modules results in the ECAL2 being smaller than ECAL1 despite having more channels. The total height of it 184 cm, the width is 245 cm.

All of the different modules used in the COMPASS electromagnetic calorimeters are listed in Tab. 3.1 with their respective size and material.

3.3.4 Hadronic Calorimeters

The hadronic calorimeters are only briefly covered, as they are not of much use for the DVCS and HEMP measurements. There are two hadronic calorimeters named **HCAL1** and **HCAL2**. Both of them are equipped with shashlik type modules that are different from the electromagnetic shashlik types in size and material. The absorber layers in these shashliks are made of iron, providing a better hadronic interaction length than lead. Each

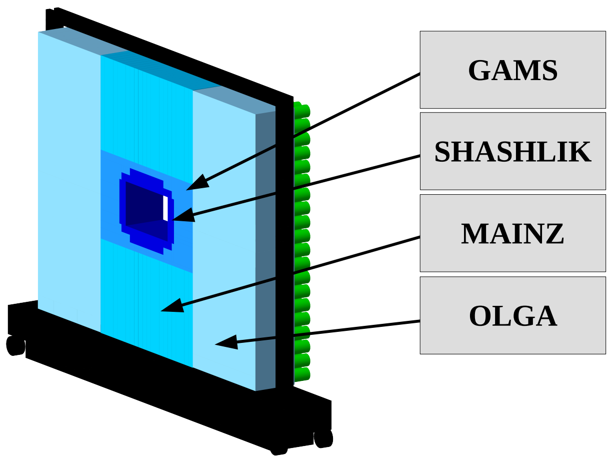


Figure 3.10: ECAL1 taken from TGEANT including mechanical support structure and the photomultiplier tubes.

module is read-out by photomultipliers tube. In earlier years the HCAL1 was used in coincidence with hodoscopes for trigger decisions.

3.3.5 Ring Imaging Cherenkov Detector

The ring imaging cherenkov detector (RICH1) is located in the LAS behind the first spectrometer magnet. The detector is filled with the heavy radiator gas C_4F_{10}. The purpose of the detector is particle identification. Particles passing through the volume with a velocity that is greater than the speed of light in the radiator gas will emit cherenkov radiation. The angle θ_C under which this light is emitted depends on the particles' velocity.

Through spheric mirrors at the end of the detector, these photons, which are arriving under the same angle, get reflected onto a circle of the same radius. So the radius of this circle is then correlated to the velocity of the particle. Since the momentum of the track can be independently measured from the track bending in the magnetic field, the rest mass becomes calculable and the particle identified. This can, in the context of the present work, be useful for hard exclusive meson production.

3.3.6 Trigger System

The spectrometer needs to deal with a high flux of about $2 \cdot 10^8$ muons per 4.8 s. However, it is impossible to write out all of the occuring muon proton interactions, because the data rate would be too high to handle by the data acquisition. To solve this issue, all the front-end cards buffer their data and wait for a decision on whether to save the data or not. The system which is supposed to make this decision is called a trigger system. The decision has to be made at a time scale of about 1 μs. In 2012 and 2016 only muon triggers were enabled. These triggers rely solely on the fast detection of the scattered muon to select interesting interactions. In principle, they all function alike. At least two sets of hodoscopes are checked with a coincidence unit.

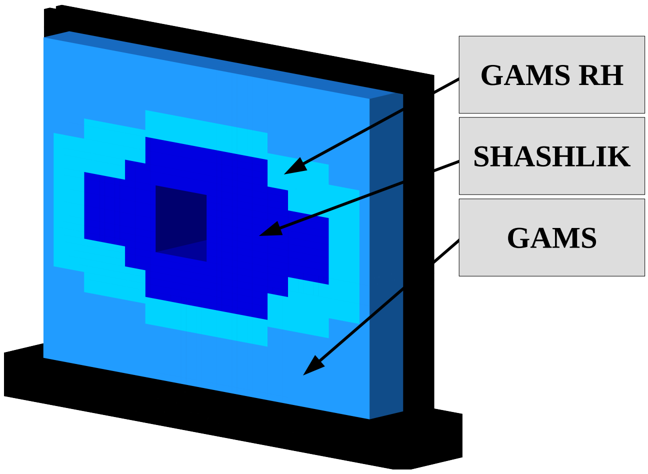


Figure 3.11: ECAL2 taken from TGEANT featuring the mechanical support structure.

The outer trigger (OT) and the large angle trigger (LAS) rely on the momentum component in the non-bending plane. Therefore, these triggers are built of horizontal scintillator slats stacked in top of each other. The correlation matrices are built in order for a trigger to be generated when the track would roughly hit the target, which is easy in this case, since the magnet bending does not contribute. This method is also called “target pointing trigger”. The matrices for target pointing are similar to diagonal ones.

To enlarge the kinematic range, that is accessable with the triggers, the y axis, where the magnet bending comes into play, needs to be considered. As the magnet bending makes a simple target pointing impossible, the matrices used here need to encode the minimal bending of the track, which in term corresponds to a minimal energy loss of the scattered muon. These matrices usually look like a triangle and the method is called “energy loss trigger”, which is visualized in Fig. 3.12. The ladder trigger (LT) is a trigger of this kind. The middle trigger (MT) combines both of the discussed concepts. A detailed description can be found in [59].

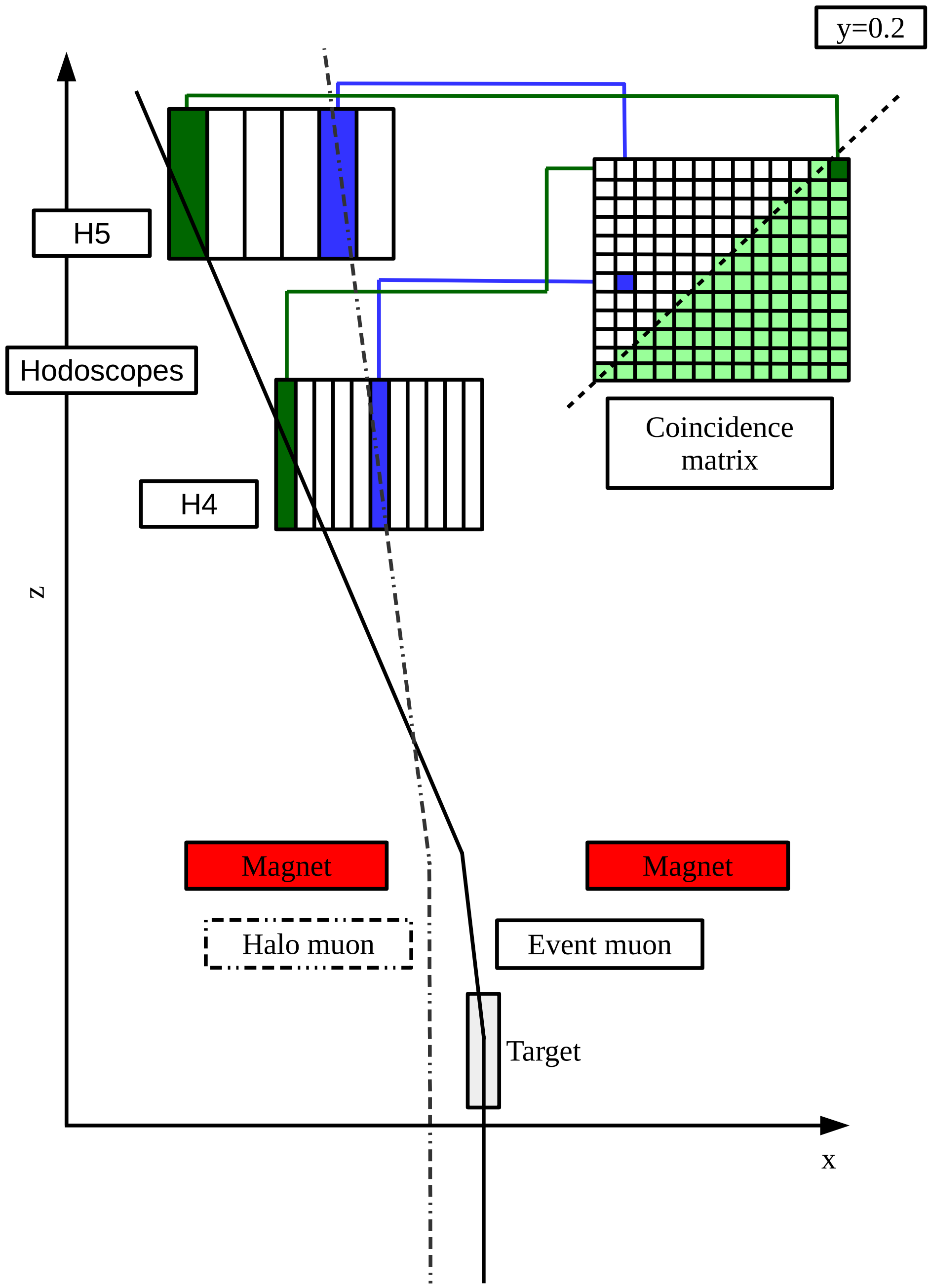


Figure 3.12: Trigger logic for the energy loss trigger, different bendings mean different energy losses. The figure was adapted from a picture in [59].

3.4 Data Acquisition and Reconstruction

3.4.1 Data Acquisition

The data rate at COMPASS is very high despite the triggering system. Trigger rates can be up to 100 kHz. This results in a data rate of about 1.5 GB/s. Finally, the resulting total data accumulated over a year of data taking can be around 600 TB. In the recent years, starting with 2014, a new DAQ was installed. This new system is based on FPGA[2] chips

[2]**F**ield **P**rogrammable **G**ate **A**rray

and therefore very efficient. A preliminary description is given in [60]. To achieve this, a multi-layer structure of data concentration is used, which is visualized in 3.13. The first layer are the frontend cards. They are converting the detector signal, time, amplitude or both, into a digital signal. Also, a small memory is embedded in these cards to provide a short buffering period, typically just enough to let the trigger system decide if this event is interesting. The front end cards cover about 250 000 detector channels. The next stage, where the sub event building is happening, are the read-out boards. These boards control the frontend cards themselves and merge their signals together. The most commonly used boards in COMPASS are the "CATCH"[3] modules for every detector, which is not otherwise noted, the "GANDALF"[4] and "TIGER"[5] combination for the recoil proton detector CAMERA and the drift chamber DC05 and the "HGeSiCA"[6] for the GEMs and silicons.

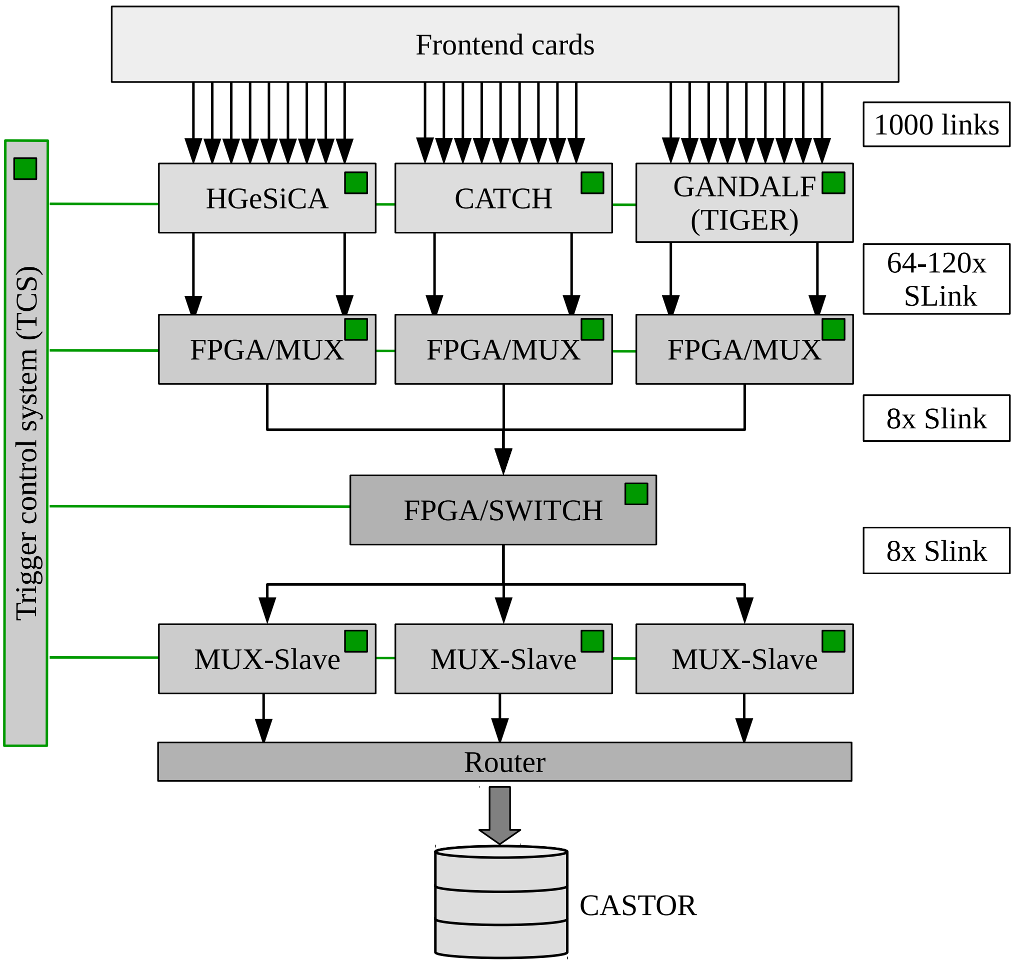


Figure 3.13: Simplified scheme of the new DAQ system from 2014 on. Green boxes mark the components that receive the trigger signal from the TCS. The TIGER module can also generate a trigger.

[3]**C**OMPASS **A**ccumulate, **T**ransfer and **C**ontrol **H**ardware
[4]**G**eneric **A**dvanced **N**umerical **D**evice for **A**nalog and **L**ogic **F**unctions
[5]**T**rigger **I**mplementation for **G**ANDALF **E**lectronic **R**eadout
[6]**Ge**m and **Si**licon **C**ontrol and **A**cquisition

The data from the read-out boards is then multiplexed by the first layer of FPGA modules to just eight S-Link[7] lines. The used multiplexer cards have enough memory to buffer one complete spill, which is about 16 GB. The switch card is the same FPGA card as the multiplexers and does the final event building. These events are sequentially sent to online PCs, which are also called MUX-Slaves, in a round-robin fashion. Finally, these online computers are then storing the concenated raw events on CASTOR via the switch system. From there, the raw data files are available for the reconstruction software.

3.4.2 Reconstruction

At COMPASS-II, the reconstruction of the raw data is done by a software called CORAL[8]. It consists of several parts that are executed one after another. The first step is done by a library called DAQ data decoding. This library takes care of reading the binary raw data and translating them into calibrated digits in the so-called decoding process. In the next phase, called clustering, these digits are then combined to hit clusters, each containing a full position and time and for some detectors also energy loss. In the special case of calorimeters in these phase, the amplitudes are extracted cellwise and the calibrations are applied.

With these hits and cell values the reconstruction is performed. For the tracking detectors a special kalman filter algorithm is applied. It starts with finding linear segments in regions without strong magnetic fields and then tries to bridge them together through the magnetic field using a fast lookup table for the bending in the field (dicofit). The resulting tracks contain information about the sign of the charge, the momentum and χ^2 from the fit. The reconstruction for the calorimeters is done by several algorithms for clustering the cell energy values together. For the ECAL2 a very sophisticated combined shower fitting algorithm based on the Lednev parametrization is available [61, 62]. For the other ECALs either the normal Lednev shower fitting algorithm or a simplified summation algorithm is available.

The reconstructed tracks are then extrapolated to the target. Points of intersection are then fitted and stored as the resulting vertices. When an incoming beam particle can be found, which intersects this vertex, it is also associated to this vertex, which in return is called a “primary” vertex. Track intersections without an incoming beam particle are called “secondary” vertices. Tracks intersecting calorimeter clusters will result in the clusters being marked as charged clusters. Clusters without associated tracks are called neutral clusters.

After this reconstruction is completed, only the momentum vectors and charges of all tracks and their first and last hits are saved to the resulting file, called a mDST[9]. However, there is one deviation from this principle: The recoil proton detectors such as the old RPD[10] and also the new CAMERA cannot be subjected to a simple tracking since their calibration can only be done after production. Therefore their digits are directly saved in the mDST file. Combining them to recoil tracks has to be done on analysis level. Furthermore, calorimeter clusters are added to the mDST as well as the geometry and magnetic fields

[7] **S**imple **Link** Interface
[8] **COMPASS** **R**econstruction and **A**lgorithm **L**ibrary
[9] **m**ini **D**ata **S**ummary **T**ape
[10] **R**ecoil **P**roton **D**etector

used for the reconstruction. This makes the mDST a small and efficient, but also self-sufficient data format for event storage.

In the case of Monte-Carlo, the decoding phase is not necessary and CORAL takes care of the clusterization process. Each detector group has their own detector class, equipped with a specialized and well-designed clusterization simulation. This includes the rejection of some hits for having too small energy deposit in the detector and also a smearing of the time and position according to each detector's resolution. This guarantees a good simulation of the real experimental setup. In this digitization process, "real hits" are created from the Monte-Carlo hits. All following modules have no information on whether the processed hits are real or were generated by Monte-Carlo in the first place. This ensures that any systematics coming from tracking or calorimeter shower fitting are visible with the help of the original Monte-Carlo information. The produced mDSTs from simulation usually contain the full Monte-Carlo input in order to allow for a comparison with the reconstructed tracks, vertices and calorimeter clusters. The complete COMPASS reconstruction scheme is shown in Fig. 3.14.

The analysis can then be performed using the PHAST[11] software. It allows users to write routines in C++ to select the data eventwise. In the software itself many often used functions like track extrapolation through magnetic fields and energy loss calculation are already implemented. With this tool it is easy to extract meaningful physical quantities from the mDST files.

[11]**PH**ysics **A**nalysis **S**oftware **T**ool

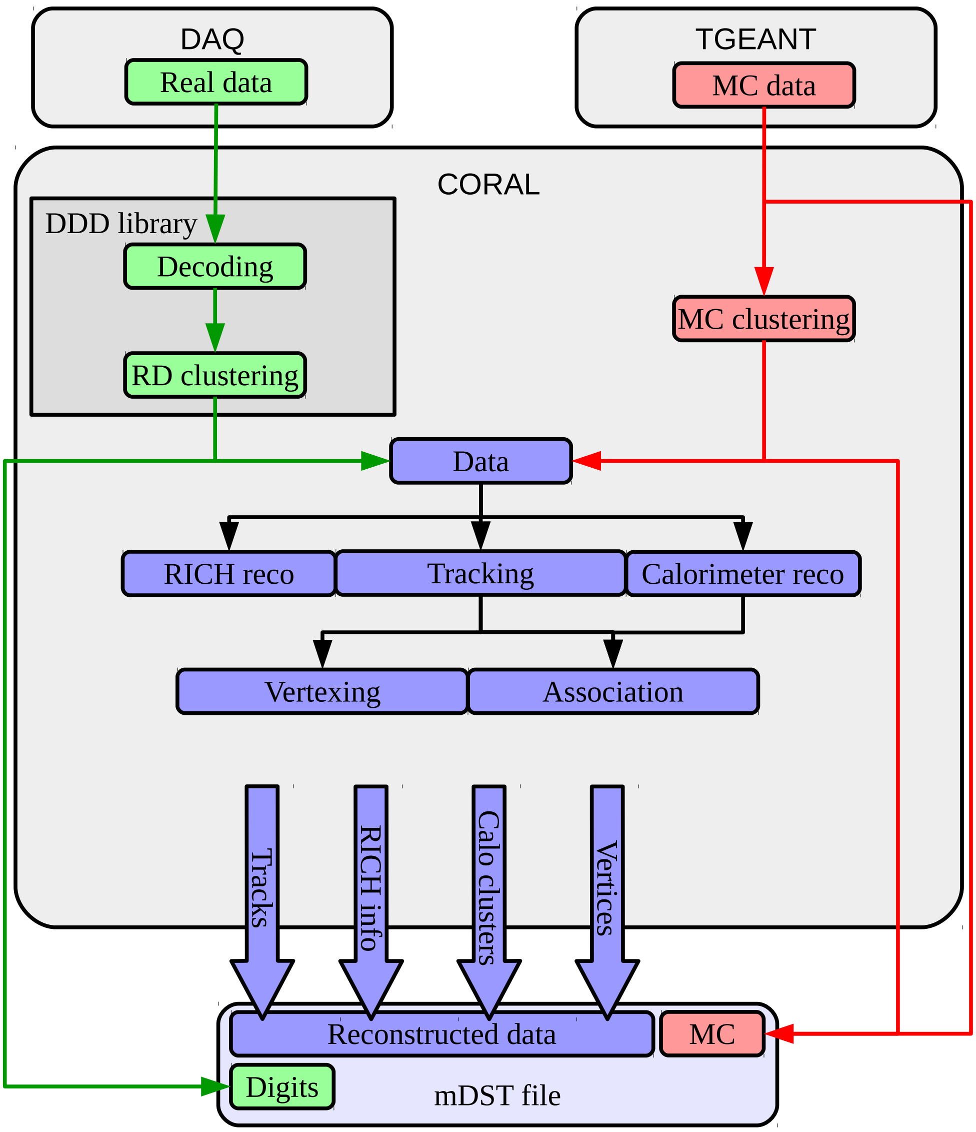


Figure 3.14: Data flow in the COMPASS reconstruction system. The saved real data digits are only for RPD and CAMERA.

4. Improvements on TGEANT

This chapter comprises three different new modules for improving the description of the experimental set-up in the COMPASS-II Monte-Carlo chain. As they are thematically linked by the the fact, that they all rely on data in some kind, they were concenated into a single chapter. The first section introduces the development and performance tests of a new two-dimensional efficiency module and its' advantages over the old system. A new pile-up muon module is shown in the second section. It allows for pre-computing the pile-up and disentangling the flux from the signal, which allows to start the simulation before the flux is known. The final module, which is presented in the third section of this chapter, is a toolbox to extract background profiles and rates for each cell in the ECALs and a method to mix random background into the signal according to these profiles. This module raises the accuracy of the ECAL simulation, which is crucial for the DVCS and the exclusive π^0 signal.

4.1 Pseudo Efficiency Maps in two Dimensions

In the scope of this thesis a new, two-dimensional efficiency map system was implemented into the reconstruction of the Monte-Carlo samples. In the old Monte-Carlo chain, only constant efficiencies per plane were available for COMPASS. This system was developed in order to improve the description of spatial anisotropies in the detectors. It includes sophisticated features such as improvement algorithms as well as a manual redrawing feature for the efficiency maps. Since these effects are only based on empirical detector performance, they cannot be described by theory and need to be extracted from the data. In Sec. 4.1.1, the extraction methods for real and pseudo efficiencies are introduced and their systematic differences are explained in detail. The next section, 4.1.2, covers the development of the new toolkit for two-dimensional efficiency management. In order to deal with the fact that very large statistics are necessary for each plane, improvement algorithms were developed for the histograms and they are introduced in Sec. 4.1.3. Moreover, the redrawing of the efficiency maps was implemented as well, so that heavy spatial anisotropies can be taken into account manually. The usage of the redrawing and an example are presented in Sec. 4.1.4. To estimate the difference and the systematics between pseudo and real efficiencies a comparison was conducted in Sec. 4.1.5. Finally, to show that the new system is working as expected, self-consistency test was conducted. The results of this test are given in Sec. 4.1.6. A summary is given in Sec. 4.1.7.

4.1.1 Extraction Method

The extraction starts for both methods, real and pseudo efficiency, by comparing illumination with found hits in a given detector plane. Illumination means that tracks that have been reconstructed with hits upstream and downstream of the given plane. These tracks are extrapolated directly on the plane, which results in a two-dimensional map of expected hits. If the detector measures a hit that is correlated in space and time to the track, the function for efficiency calculation gives a positive result. A negative result, on the other hand, is returned if there is either no hit at all, or the correlation is bad. The efficiency is then calculated by dividing the number of positive results by the sum:

$$e'(x, y) = \frac{N_{\text{pos}}(x, y)}{N_{\text{pos}}(x, y) + N_{\text{neg}}(x, y)}, \tag{4.1}$$

where $N_{\text{pos}}(x, y)$ is the number of positive efficiency checks in the bin (x, y) and N_{neg} is the number of negative checks. The bin width is detector-dependent, as rates differ greatly between the VSAT, SAT and LAT detectors, resulting in large differences in statistical power per active area. Real efficiency study means that the detector that is going to be checked, is excluded from the tracking. Naturally, in an experiment like COMPASS-II with over 300 planes, reconstructing a large amount of data for each plane is a huge task. To get an estimation, the so-called pseudo efficiency study can done, where the detector which is going to be checked, is included in the tracking. This can lead to a systematically higher efficiency of the detector, because of its ability to “pull” the track towards hits in its plane. Both methods and the systematic uncertainty introduced by only producing pseudo efficiencies is shown in Fig. 4.1.

The four reconstructed tracks are shown to introduce the concept of the pseudo efficiencies and to show the difference between pseudo and real efficiencies. The real efficiences

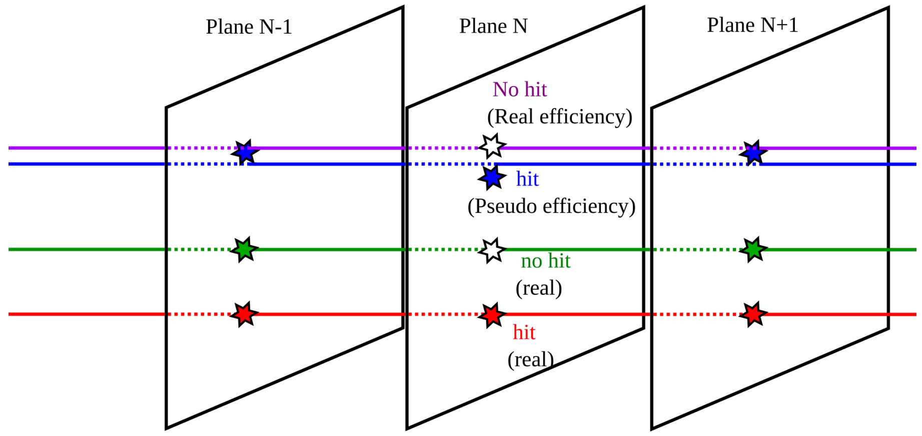


Figure 4.1: Schematic of efficiency production. In red and green: the negative and positive signals for real efficiency production. In blue and purple: a case of a reconstructed track that would give a positive result in pseudo-efficiency, because the hit in the N plane pulls the track down (blue case), while it would count as a negative in real efficiency, as depicted in purple.

are presented in the red and green lines. Green would count as negative, red as positive. The three upper hits can be reconstructed to the blue track if all three planes are enabled. In that case, the Nth plane would have a hit and therefore get a positive entry. Without the Nth plane, the tracking would most likely suggest the purple track, which is too far away from the hit in plane N and therefore results in a negative entry. As this effect cannot occur in the negative direction, pseudo efficiencies are systematically higher. The impact of this effect has been studied and the results are shown later in this section.

4.1.1.1 Old method

In the old Monte-Carlo chain, every detector had only a single efficiency assigned which was a single number calculated by convoluting the efficiency as a function of $e'(x, y)$ with the beam profile p_B dependent on the same parameters:

$$e = \int_x \int_y e'(x, y) \otimes p_B(x, y). \tag{4.2}$$

In the reconstruction, this number is used to decide whether the hit is accepted or not for every Monte-Carlo hit, regardless of its position. This is done by throwing a random number $r \in [0, 1]$ and accepting the hit if $r < e(\text{plane})$. While this assumption can hold in the limit of high statistics and symmetric inefficiencies, it will not represent any anisotropies caused by dead zones correctly. What is worse, is the fact, that local dead zones of the detector are not excluded from convolution and they lower the efficiency of the whole plane in completely unrelated regions.

4.1.1.2 New method

The new method relies on accurately describing the detector with a two-dimensional efficiency map. This method takes into account existing anisotropies in the detector's efficiency. Should a bin be empty due to no statistics beeing available in there, the bin is set to -1 to enable the system to destinguish between a zone with zero efficiency and one without any illumination. In the case of empty bins, the arithmetic mean of all filled bins is taken instead.

4.1.2 Toolkit

The new two-dimensional efficiency toolkit provides interfaces to load and modify sqlite databases containing the efficiency maps. They are saved as single rows in a table, which is described in the appendix F in more detail. During reconstruction time, all of the used planes are loaded into the memory to guarantee fast access. Every hit is then checked for its position in the respective detector plane. The bin is then calculated from the position of the hit in the so-called wire reference system (WRS). This bin contains the efficiency value that is used to decide whether the hit will be accepted or rejected. The decision is made by a simple uniform throw with $r \in [0, 1]$, a hit is accepted if $r < e(\text{plane}, x, y)$ holds true and is rejected else.

4.1.2.1 Initialization

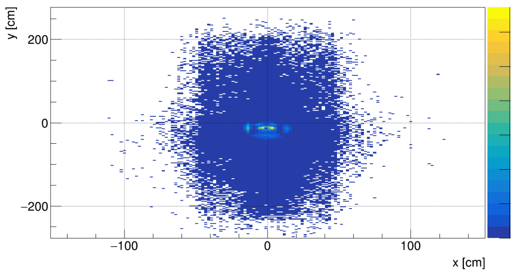
A python script is provided in order to create a new sqlite database. This script only creates the database structure without any further information. The first initialization process is then done by importing an alignment file from a Monte-Carlo software. This file includes all the mean efficiencies. Next, the initialization creates a database entry for each detector plane but it only featuring a single bin. This bin is set to the mean efficiency from the file. The sqlite database can now be enabled and the results after reconstruction are exactly the same as after reconstruction using the detectors.dat file. This is important, since creating efficiencies is a huge effort and usually not all detector planes are available at the same time. With this basic set of entries, the database can be modified plane by plane until it is completed, and at the same time at least the old flat efficiencies for any detectors are used, which are not yet studied in the new way.

4.1.2.2 Usage

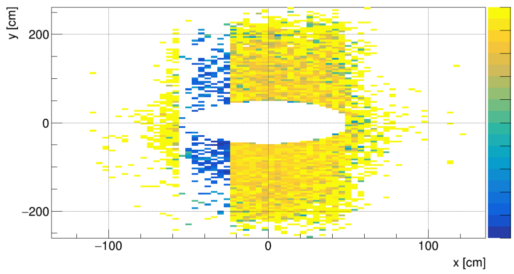
For the users creating the new efficiencies, the first steps are kept the same. The standard COMPASS-II tools already provides the necessary the raw maps. The only new things are a conversion tool, **efficConverter.py**, that compares the illumination map with the efficiency map and sets the empty bins to -1 and the **efficienciesDB**, which converts the ROOT histograms to the sqlite format and inserts them. An example of an illumination map, raw efficiency map and the combined efficiency map for the DVCS 2012 Monte-Carlo is shown in Fig. 4.2. It is possible to store several years, periods or beam charges in one database depending on the needs of the physics program. In the case of the muon program, it was decided to use one set of efficiencies per charge and year. It can be specified in the CORAL options file which set of efficiencies the system will use. The CORAL options flags, which are used to control the module, are listed in Tab. 4.2.

4.1.3 Improvement Algorithms

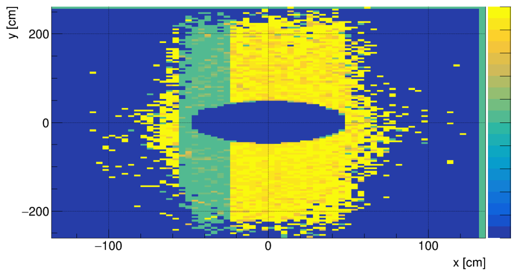
As already mentioned, in order to generate real efficiency maps from data, the used data sample needs to be reproduced for each plane. Therefore, the data samples' statistical power is limited by computational power. To improve the efficiency maps from the limited data, two new algorithms were developed for this toolkit. The first one is a geometric bridging of empty bins, while the second one is a floating mean. These two algorithms are described in the following section.



(a) The illumination of the plane.



(b) The calculated efficiency of the plane bin-wise. Unfilled bins have no hits in them or no illumination.



(c) The combination of the efficiency with the illumination. Unfilled bins in the illumination are set to -1 here. Therefore the z scale here ranges from -1 to 1.

Figure 4.2: Efficiency map combination with illumenation map. Mind the z scale in the last plot.

4.1.3.1 Geometric Bridging

The first improvement stage of the efficiency maps is a geometric bridging. The main idea is to fill unfilled bins, which have a homogeneous neighborhood with the mean of the neighbors value. Some of the inefficient detector features are straight lines, such as the borders of the active parts of the GEMs. These features should be taken into account and their borders must not be washed out or blurred. The idea is to calculate a mean of the neighboring values and at the same time keep track of how many neighbors contribute. The bridging matrix $\mathcal{G}$ is introduced so that the second generation of neighbors is also taken into account:

$$\mathcal{G} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{\sqrt{5}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \\ \frac{1}{2} & 1 & 0 & 1 & \frac{1}{2} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{\sqrt{5}} & \frac{1}{2\sqrt{2}} \end{bmatrix}. \tag{4.3}$$

The whole efficiency map is then handled as a big matrix with the elements e_{ij}. For all empty bins $e_{ij} = -1$ was already set. For all of these empty bins, the following value is calculated by first calculating the absolute value:

$$e'_{ij,\text{abs}} = \sum_{j'=j-2}^{j+2} \sum_{i'=i-2}^{i+2} \begin{cases} e_{i'j'}\, \mathcal{G}_{i'-i,j'-j} & \text{for} \quad e_{i'j'} > -1 \\ 0 & \text{else.} \end{cases} \tag{4.4}$$

Afterwards this value is normalized:

$$e'_{ij,\text{normalized}} = \frac{e'_{ij,\text{abs}}}{D}, \tag{4.5}$$

where the normalization factor D is the sum of all used elements of $\mathcal{G}$:

$$D = \sum_{j'=j-2}^{j+2} \sum_{i'=i-2}^{i+2} \begin{cases} \mathcal{G}_{i'-i,j'-j} & \text{for} \quad e_{i'j'} > -1 \\ 0 & \text{else.} \end{cases} \tag{4.6}$$

The sum of the elements of the $\mathcal{G}$ matrix can be interpreted as an area of a circle, therefore cutting on D means cutting on the number of available filled neighbor elements. In the scripts, the value for the cut-off was empirically chosen to:

$$D_{\text{min}} = 6.0, \tag{4.7}$$

meaning that about 55% of the neighboring bins have to be filled. This ensures that borders of different regions do not get washed out during the process.

4.1.3.2 Floating Mean

After the filling of the empty bins with the geometric bridging, a cutting weighted mean is applied to cut out extremely sharp fluctuations caused by low statistics bins. While the procedure is mathematically similar to the geometric bridging, this smoothing is used on all bins, not only on unset bins. Again, the new element is defined:

$$e'_{ij,\text{un}} = \sum_{j'=j-1}^{j+1} \sum_{i'=i-1}^{i+1} \begin{cases} e_{i'j'}\, \mathcal{G}'_{i'-i,j'-j} & \text{for} \quad e_{i'j'} > -1 \\ 0 & \text{else,} \end{cases} \tag{4.8}$$

with the normalization:

$$e'_{ij,\text{normalized}} = \frac{e'_{ij,\text{un}}}{D}, \tag{4.9}$$

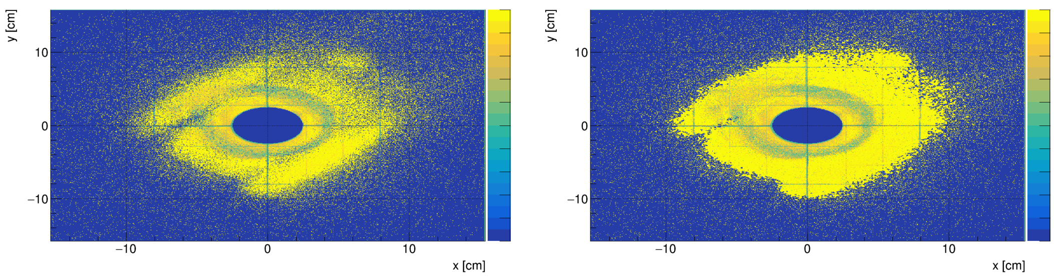
and:

$$D = \sum_{j'=j-1}^{j+1} \sum_{i'=i-1}^{i+1} \begin{cases} \mathcal{G}'_{i'-i,j'-j} & \text{for} \quad e_{i'j'} > -1 \\ 0 & \text{else.} \end{cases} \tag{4.10}$$

The mean kernel matrix is defined as:

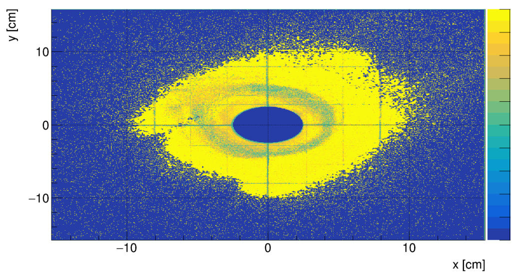
$$\mathcal{G}' = \begin{bmatrix} \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 1 & 2 & 1 \\ \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \end{bmatrix}. \tag{4.11}$$

The cuts applied this time are:



(a) The combined efficiency map without any modification.

(b) The map after application of the geometric bridging.



(c) The map after geometric bridging and the floating mean.

Figure 4.3: Efficiency maps of the plane GM04V1.

- At least four neighboring bins were used.
- The standard deviation of these four is smaller than 0.2.
- The difference of values after and before are higher than: $|e'_{ij} - e_{ij}| > 0.2$.

These cuts together only allow the smoothing of bins that are very different from an otherwise homogeneous neighborhood. Both methods' results are shown in Fig. 4.3. As visible from the picture, both algorithms should be applied one after another for the best results.

4.1.4 Editing Options

Some of the planes cannot be simply corrected by the improvement algorithms introduced earlier. Especially in the case of larger dead zones in the detectors, only correcting the efficiency maps manually will be fully correct. An example is shown in Fig. 4.4, in the straw plane ST05X1ub. A part of this plane is not functional, and because the straws only measure in one dimension, it has to be assumed that the unilluminated part of the plane is also dead in this region. This can hardly be done automatically and needs correction by hand. The new efficiency toolkit in TGEANT allows for a correction by hand in these cases by exporting the efficiency map directly to an image file. The targa image format is used for the export, and it can be read in by almost any graphics editing software. The information on the efficiency is stored as the grayscale value and the statistics in the alpha channel. It has to be mentioned, that the ST05 detector is not used in the 2016 run anymore and was just taken as an example for difficult to consider dead zones. This improves the overall description of the detector, since only non-zero bins are used for the arithmetic mean that is filled in all unilluminated bins. In the old method, the inefficiency of this part of the plane was lowering the efficiency of the whole plane. By marking zones dead in the new system, the mean efficiency of the plane is raised, which is most probably a better approximation for unilluminated areas.

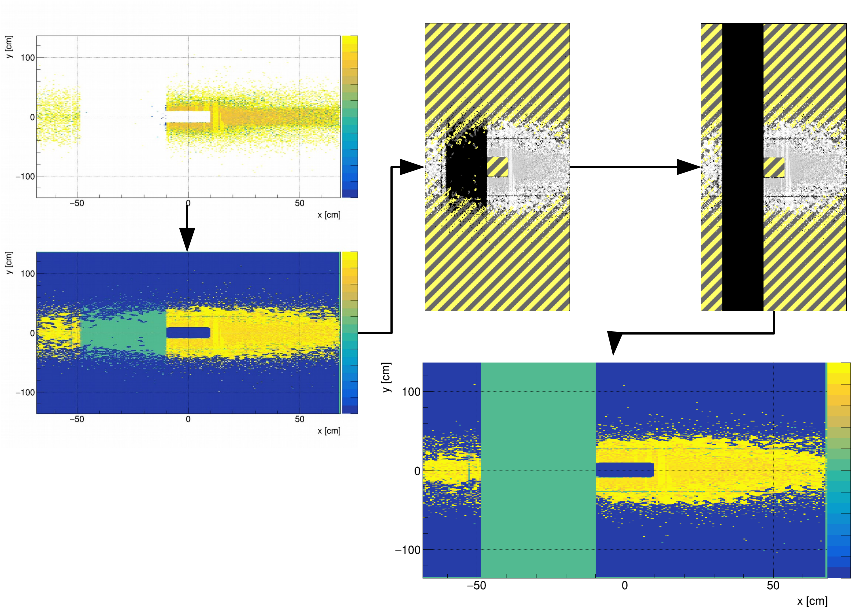
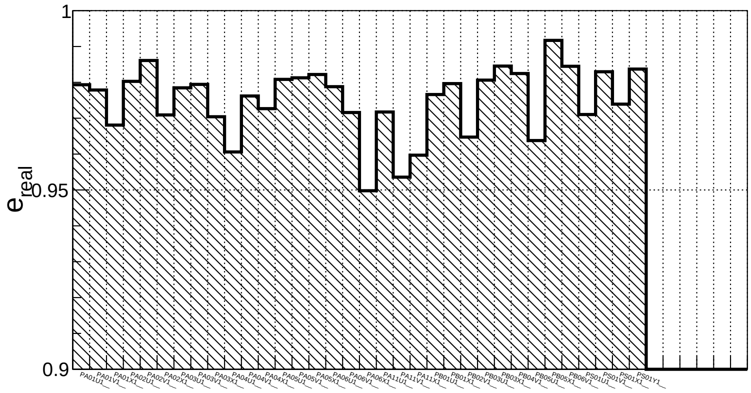


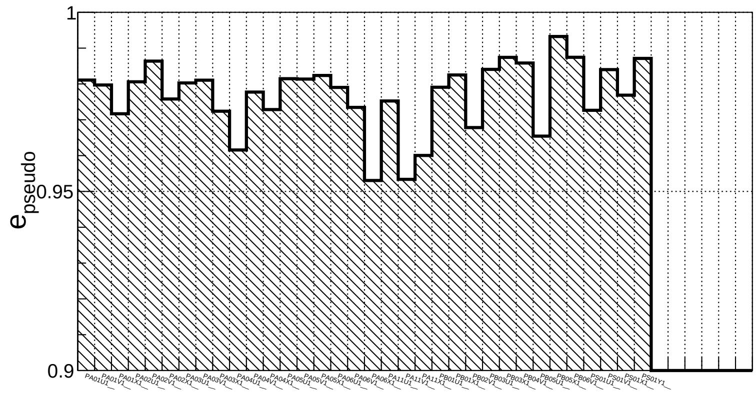
Figure 4.4: The flow of correcting an efficiency map manually at the example of ST05X1ub. It starts at the top left with the uncorrected map, then applying the bridging and the floating mean as well as illumination mapping at the bottom left. The map is then exported to a graphics file and loaded into a graphics editing software at the top right and edited there. The yellow and black hatching in these pictures indicates the alpha transparency of the image which represents bins without any illumination. Finally, the graphics file is reimported into the sqlite database resulting in the final efficiency map for the detector plane at the bottom right.

4.1.5 Comparison with Real Efficiencies

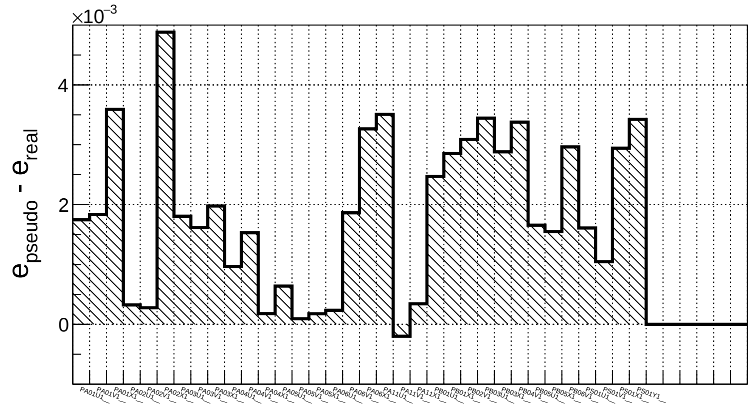
In the scope of this thesis it was decided to produce a full set of pseudo efficiencies separated by beam charge, which helps to cope with the systematics of different beam intensities. These pseudo efficiencies were then cross checked with the real efficiencies extracted by Artem Ivanov, who created a set of real efficiencies for the MWPC detectors for this test from the same real data sample. The results are listed in Tab. 4.1 and on the one hand, it is clearly visible that there is the systematic tendency to higher pseudo than real efficiencies. On the other hand, this systematics are one order of magnitude smaller, than the fluctuations between the detectors. The full table is given in the appendix E.1. As the systematic shift is always below 0.5 %, as is shown in Fig. 4.5, the usage of pseudo efficiency maps is well justified in first order approximation.



(a) The mean value of the real efficiency maps per MWPC plane.



(b) The mean value of the pseudo efficiency maps per MWPC plane.



(c) The difference of the mean values of the pseudo and real efficiency maps.

Figure 4.5: A comparison of the mean values of the efficiency maps per plane for the extracted real and pseudo efficiencies. These were extracted using the same parameters for the COMPASS-II efficiency extraction UserEvent and from the same real data sample from the 2012 COMPASS-II pilot run. A systematic shift to higher pseudo than real efficiencies is visible, which is also expected from theory. Furthermore, the shift is smaller than 0.5 % for all planes. (Real data run used: μ^{+}: 108511, μ^{-}: 108539)

Table 4.1: Comparison of real and pseudo efficiencies for the DVCS run in 2012. The real efficiency field denotes the arithmetic mean over all planes of the detectors of the respective tbname. In the case of the pseudo efficiency, this is done alike. The uncertainties for both these values comes from the standard deviation of the planes. The mean difference is the arithmetic mean of the differences per plane, which is why more digits are given for this value.

Detector	**real efficiency**	**pseudo efficiency**	**mean difference**
PA	0.98 ± 0.01	0.98 ± 0.01	-0.001529
PB	0.98 ± 0.01	0.98 ± 0.01	-0.003446
PS	0.978 ± 0.007	0.980 ± 0.007	-0.001994

4.1.6 Self-consistency Test

The self-consistency test was conducted on a test production of the production_16-02 DVCS mu+. The same extraction method was used as for the original efficiency extraction, which makes it a comparable test. Mind that in the regions where the statistics from real data were not sufficient, the arithmetic mean was used. This leads to the effect that in Fig. 4.6, the mechanical structure of the GEM plates is only visible in the center. While this is slightly disappointing, it is fixable by simply using more data to extract the pseudo efficiencies. Despite this, the overall agreement is almost perfect, making the self-consistency test a full success. The manually edited efficiency of the ST05X1db is shown in Fig. 4.7, this proves the system works as expected. The manual editing is important in this case, since straws are one-dimensional detectors made of straw tubes. If some of them get broken, they fully break on their whole length, which is the y axis in this case. Hence, they can safely be set to dead as done for this plane.

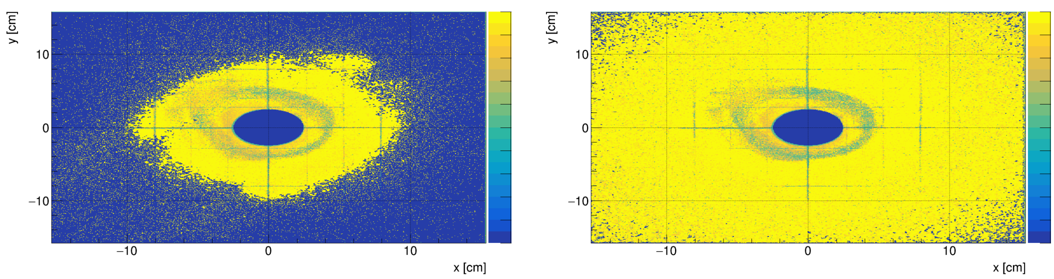


Figure 4.6: The final efficiency map for the GM04V1 detector plane is shown on the left. On the right: The reextracted efficiency for the self-consistency test. If the statistics in the input map are sufficient to resolve the inefficient borders of the GEM segments, they are also visible. In the outer regions, the arithmetic mean was taken, so there are no such structures visible.

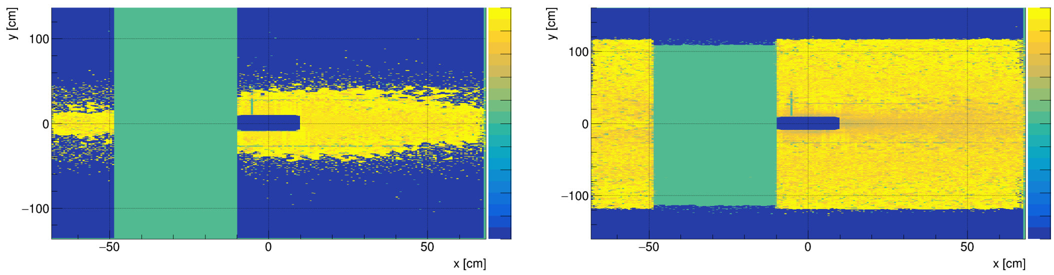


Figure 4.7: On the left: The final efficiency for the ST05X1db plane as extracted from real data. The plane was modified manually to mark the broken channels dead on the whole y range. On the right: The self-consistency test result from a Monte-Carlo sample. While the illumination is far better in Monte-Carlo, the dead zone introduced manually is clearly visible.

Table 4.2: Trafdic options for the two-dimensional efficiency maps

CsTGEANT option	**Value**	**Description**
2DEfficDB	[File]	The sqlite database to load.
2DEfficYear	2012-DVCS-muplus 2012-DVCS-muminus	An identifier flag showing which maps to use. Can also be used to identify different periods.
2DEffic AcceptNearbyYears	YES	If the detector has a better map in another flag option of 2DEfficYear, load it instead.

4.1.7 Summary

A complete toolkit for creating, refining and applying two-dimensional efficiency maps in the COMPASS-II Monte-Carlo chain was created. With the help of this toolkit two complete sets of pseudo efficiencies for the DVCS run in 2012 were extracted from real data, one set per beam charge to deal with the different intensities of the muon beams. As the efficiencies are intensity dependent, this is an important feature. The difference between the mean pseudo efficiencies for the different beam charges was found to be lower than 2 %. Each set was checked for dead zones in the detectors by visual examination and treated according to Sec. 4.1.4, to mark the dead zones for each plane correctly. While it could be shown for the MWPC detectors that there is the expected systematic offset between pseudo and real efficiencies, the deviation was found to be below 0.5 %. It has to be discussed, whether this systematic deviation is of enough influence to go through the computationally expensive process of extracting the real efficiencies. A self-consistency test was conducted, based on the latest Monte-Carlo production for the 2012 COMPASS-II pilot run, which shows very good results. The trafdic options to enable this feature are listed

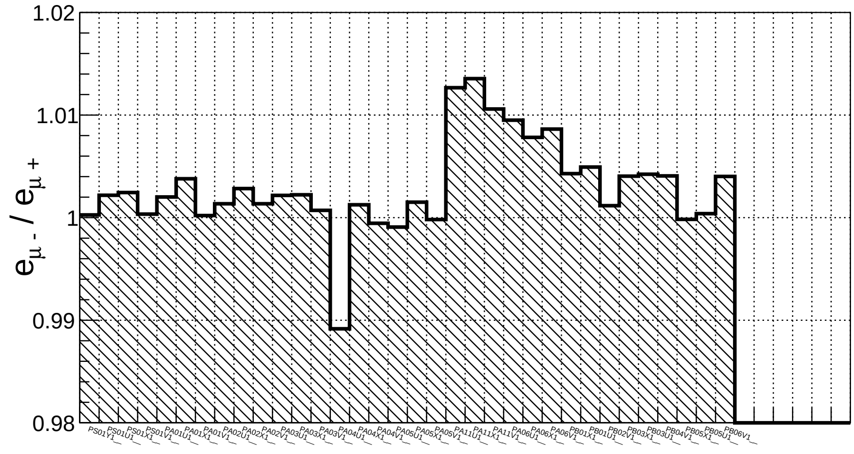


Figure 4.8: Ratio of mean pseudo efficiencies extracted from negative and positive muon data samples for different MWPC planes. The μ^- efficiencies are systematically higher than the μ^+ ones. This was expected since the μ^- flux of beam and halo is lower. (Real data run used: μ^+: 108511, μ^-: 108539)

in Tab. 4.2. The efficiency maps for the different beam charges differ by less than two percent for all properly illuminated planes, exemplary the MWPC detectors are shown in Fig. 4.8. The set of presented pseudo-efficiencies are currently the most accurate description of the detector performance in the 2012 pilot run. As they only differ marginally from the real efficiencies, they have been put in the database for the Monte-Carlo production as the reference for the 2012 run.

4.2 Pile-Up Enhancements

When it comes to reducing the uncertainty on the Monte-Carlo acceptance, statistics is the key. The COMPASS-II experiment, which uses muons as the beam particles, has a very high flux of them on-spill, resulting in a high flux of not-reacting muons. Not-reacting means, that muons are just crossing the spectrometer without engaging in a deep inelastic muon proton scattering. Nevertheless, these muons still cross many detectors and deposit energy in the calorimeters. To deal with these extra muons that are present in the time gate of an event, they need to enter the simulation too. This was already done before, but the statistics of the beam phase space was very limited and therefore unsuitable for a mass production aiming at high accuracy. Sec. 4.2.1 presents the creation of new beam files for the COMPASS-II test run in 2012 is shown. They greatly enhance the statistics of the beams phase space. In order to handle the upcoming DVCS run in 2016, which will generate unprecedented statistics for exclusive measurements, the Monte-Carlo statistics needs to be improved dramatically. This can be achieved by improving the performance of the simulation. In fact, a significant simulation speed-up has been achieved by caching the reactions of these pile-up muons. The new methods and formats needed for this approach are introduced in Sec. 4.2.2. Finally, a summary is given in Sec. 4.2.3.

4.2.1 New Beam Files

A new beam file, which parametrizes the beam particles, was created for each charge so that the accuracy of the beams' phase space representation can be improved. The new files feature millions of beam particles, whereas the old beam files only used 100 000 entries.

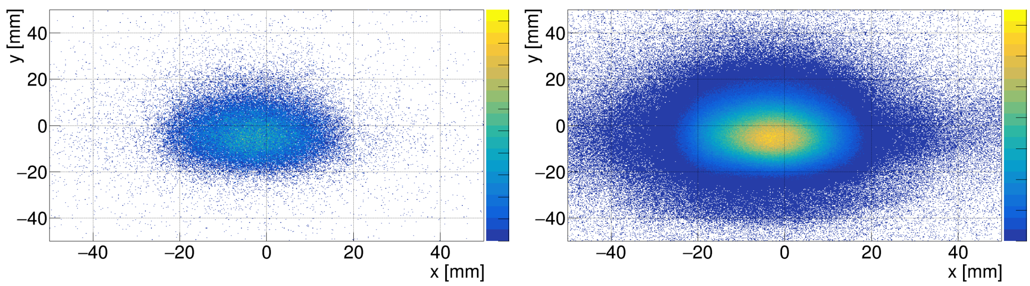


Figure 4.9: Spatial distribution of beam entries from the samples in the old (left) and new (right) beam files. The new beam files provide superior statistics and therefore sample the phase space much better.

The distributions for the kinetic energy and the spatial distribution are shown in Fig. 4.9 and Fig. 4.10. As the beam phase space is dependent on the charge, a second beam file was created for the μ^- case. For completeness, the spatial and momentum distributions are shown in Fig. 4.11.

4.2.2 External Pile-Up

A new function was introduced in TGEANT and CORAL in order to save computation time. When it comes to the upcoming run in 2016 with a very high statistics, the Monte-Carlo performance is very important. The muons rarely interact and most of them just fly

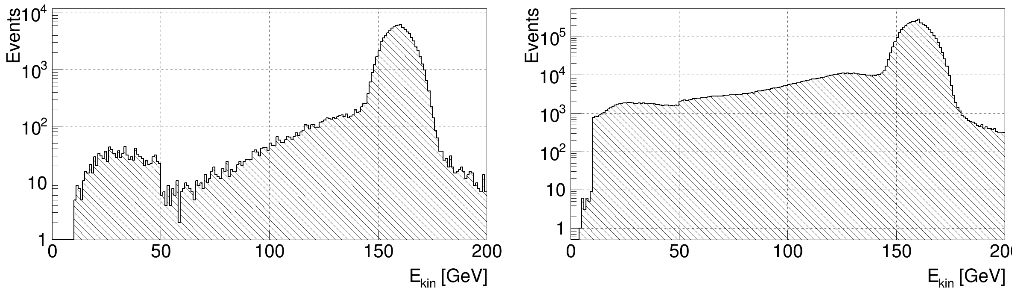


Figure 4.10: Energy distribution of beam entries from the samples in the old (left) and new (right) beam files. The new beam files provide superior statistics and therefore sample the phase space much better. Especially in the region below 100 GeV the old description was not sufficient, since it started fluctuating.

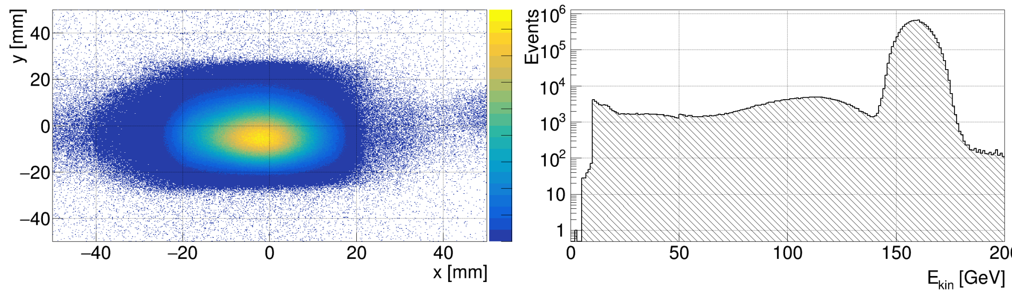


Figure 4.11: Energy and spatial distribution of the new beam file for the μ^- beam.

through the spectrometer and generate hits in the detectors. TGEANT needs to track a mean number of six of these additional muons per event, in a ± 100 ns timegate through the whole spectrometer. Since muons can hardly be destroyed, they need to be transported throughout the whole 60 m of the hall and deposit some energy in every detector they cross. Additionally, the time consuming shower simulation is triggered in each crossed calorimeter. This can be exploited to speed up the generation process by pre-computing these muons and their interactions in the spectrometer. This is technically not trivial since the amount of data to store is substantial. There are two million beam file entries and each one should be simulated at least ten times. This leads to roughly 20 000 000 events from which about six should be chosen randomly for each event. The usual TGEANT file format features a streamable gzip compressed ASCII file. This format is good for storing events for a long time or stream-piping the output directly to CORAL, but is has the problem of relatively slow reading speeds since the compression is taking a heavy load on the CPU. Also, there is no possibility to skip events or select random events from the file. This problem was overcome by introducing a new binary file format for TGEANT – the tbin format. As the events are not of the same size, a simple jumping to fixed position is not possible. Instead the format features a table of contents (ToC) section at the end of the file. The new binary format allows for fast random access operations while at the same time requiring almost no CPU time to do so.

Table 4.3: Trafdic options for the external pile-up module

CsTGEANT Option	Value	Description
ExternalPileUp	1	Enables the external pile up addition module.
ExternalPileUpFlux	0.03464 (μ^+) 0.01468 (μ^-)	The number of muons per ns.
ExternalPileUpTGate	100	The time range in which the pile-up muons are added. Given in ns.
ExternalPileUpFileList	[PATH]	The list of pile-up tbins to load.
ExternalPileUpCache	1	Copy the loaded tbin to a local folder before starting.
ExternalPileUpCacheFolder	[PATH]	The path to cache the tbin file to.

Options

At CORAL runtime, the TGEANT library loads a random file from the supplied tbin list. From there on the reading process is quite easy. The reading starts from the back. The first long integer is giving the size of the ToC. Then the ToC, which consists of long integers denoting the starting points of new events in the binary file, can be read in. These numbers are pushed in a list which is afterwards shuffled. The shuffle of the vector gives the function for loading the next event the interpretation of a random access iterator. A full visualization of the format is shown in appendix G.

The complete list of CORAL trafdic options is given in Tab. 4.3. To minimize the network load on batch systems, the system can copy the chosen file to a cache folder on a physical drive on the machine, if needed. This is enabled via the ExternalPileUpCache flag.

Procedure

In order to select the number of pile-up muons to be included, a Poissonian random number is drawn with the mean:

$$n_\mu = (2\, t_{\text{gate}} \times flux) - 1. \tag{4.12}$$

The reduction of the number by 1 is simple, because the main incident muon is already counting to the input flux. Afterwards, a time-offset is drawn uniformly as:

$$t_{\text{offset}} \in [-t_{\text{gate}}, t_{\text{gate}}]. \tag{4.13}$$

All timing information in the pile-up event is then shifted with this offset, and finally the event is merged into the signal event.

Event Merging

The merging of two events required new code inside of the TGEANT event library. A typical TGEANT event consists of a block of generator information and the spectrometers response to these generator tracks being tracked through the experimental hall. The generator information consists of a list of enumerated four-vectors with particle identification and status codes as well as their respective parent and daughter numbers. The response block contains a linked list of actual tracks including the particles that were created outside of the primary vertex. A hard photon producing an electron positron pair, or delta electron production can happen anywhere in the spectrometer. It is called a trajectory list. Both of these data blocks were originally designed to be fully self-consistent and therefore rely on the absolute number of each track for parent daughter relations. This needs to be taken account when merging two events. The new merging procedure takes place in the event class and features the full integration of such a pile-up event into the original event. First, all the parent-daughter relations are kept and offsetted to the next free particle-number of the event to merge into. Following this, the trajectory lists are concatenated. Calorimeter cell energy values are added up onto the already existing ones. This can be justified by the calorimeters featuring timing accuracy well below the capability of separation of the pile-up muons and the signal event. Finally, a new lujet is created with the information of the incident pile-up muon and added to the lujet list. The final combined event can then be processed by CORAL as usual. All in all, the read accesses to a local disk drive is still much faster than tracking the pile-up muons each time and takes well below one second per event for the full approximately five muons.

Flexibility enhancement

The flux of muons is a quantity difficult to estimate. There are several methods to extract the fluxes, including the direct use of beam scalers as done by E. Fuchey, or indirectly via the integrated DIS cross section, which is what S. Landgraf [63] did. The method of pile-up addition allows to generate the pile-up beforehand without knowing the flux, so the Monte-Carlo generation can start before this quantity is known. If any better flux value becomes known, it is possible to just rerun CORAL. This typically takes only two to three days on 200 cores for over 100 million events. Regenerating a whole sample of this size takes over a month on the same batch system. This new flexibility allows for a faster analysis of the data, since the Monte-Carlo becomes available way earlier.

4.2.3 Summary

New beam files for 2012 were created, enhancing the statistics from 100 000 to several millions of entries per charge. On the base of these new beam files, the pile-up was pre-computed in order to speed up the production of Monte-Carlo events. In order to make this work fast, a new indexed binary format was introduced to TGEANT – the tbin format. It allows for quick random access to the events. As a consequence, the tbin files are much larger than the usual tgeant files. In order to get a sufficiently accurate phase space description for the pile-up, the muons were tracked about ten times each. This results in about 20 million pile-up muon events or over 220 GB of generated tbin files. The LEPTO production of over 100 million DIS events requires about 500 pile-up muons that normally would have been tracked individually. So a speed-up of factor $f = \frac{500 \cdot 10^6}{20 \cdot 10^6} = 25$ was achieved in the pile-up generation time. This improvement will also greatly reduce CPU time needed to generate the 2016 Monte-Carlo sample vastly. In addition, the flux can be changed withing a matter of days and without regenerating everything from scratch, even on large Monte-Carlo samples. That is why the Monte-Carlo production can start even before the flux measurements are done. As a result, Monte-Carlo can become available earlier and the analysis work can start sooner.

4.3 Background Simulation in Electromagnetic Calorimeters

The background of the electromagnetic calorimeters at COMPASS-II is not simply derivable by theory. It is a combination of untracked and unvetoed halo muons making their way through the cells and electronic noise. The ladder is coming from photomultiplier tubes that are uniquely calibrated for each cell. Therefore, the background profile is highly cell-dependent. The halo muons' contribution is also cell-dependent, as their distribution is not isotropic. In this section, a set of tools are introduced to extract rates, energy profiles and time distributions cell-wise (Sec. 4.3.1). In the next section, Sec. 4.3.2, the method used to create background from the extracted profiles is explained. To check the validity of the module, an elefant test was conducted and the results are shown in Sec. 4.3.3. Finally, a summary is given in Sec. 4.3.6.

4.3.1 Profile Extraction

This section is about the extraction of the background profiles from the real data. It describes the process of extracting energy and time distribution of background clusters.

4.3.1.1 Data selection

The profile extraction starts with applying the following cuts to the full 2012 DVCS run data:

- only random trigger,
- no reconstructed beam in event.

The random trigger cut is applied to ensure that the resulting sample will contain time distribution of detector hits and clusters as uniform as possible. The remaining sample after this marks the largest amount of data that could already be used for the extraction of the background profiles. The next step is cutting out all events that have a reconstructed beam particle. Of course, the random trigger sample also contains events that have a reconstructed beam, which might lead to deep inelastic scattering with an arbitrary number of created neutral clusters. These events are simulated with the TGEANT and the LEPTO event generator, so this effect has to be cut out in order to avoid double counting of it in the final Monte-Carlo sample. The remaining event numbers after these cuts are given in Tab. 4.4.

Table 4.4: Event numbers after cuts for data selection.

Cut	**Events remaining**	**Fraction**
No cuts	4605262024	1.0
Random trigger	930964634	0.202152370
No beam particle	717510893	0.155802403

Next, all the clusters that were found are checked in each event. A cluster is a reconstructed energy spot in a calorimeter. It may contain several cells, depending on its energy and the type of the module. The clusters are reconstructed by CORAL from the energy values in each cell. For ECAL0 and ECAL1, the process starts at the most energetic cluster and then sums up all energies of the cell's neighbors. When no more neighboring cells have been struck, the algorithm searches the calorimeter for the next highest energy cell. The position of the cluster is then set to the mean of the struck cells weighted with their energy contribution. Only neutral clusters, which are defined as clusters that have no charged track crossing it, are saved to the final tree. The numbers split by each ECAL are listed in Tab. 4.5. These clusters, their energy and time distribution in each cell, are the base for all extracted profiles afterwards. To conveniently process a large amount of data on a batch system, all of the clusters are first written into a single tree. This tree contains the cell identification number, time and energy of the cluster. The ROOT trees from different computers can afterwards be easily merged. From these numbers the rate of background

Table 4.5: Neutral cluster numbers in final sample.

ECAL	**Clusters**	Fraction
ECAL 0	91402133	0.0514053
ECAL 1	470585785	0.26466128
ECAL 2	1216080205	0.68393342
Combined	1778068123	1.0

can be calculated as follows:

$$n_{\mathrm{bg}} = \frac{N_{\mathrm{ev}}}{N_{\mathrm{clus}}} = 2.47810610\,, \tag{4.14}$$

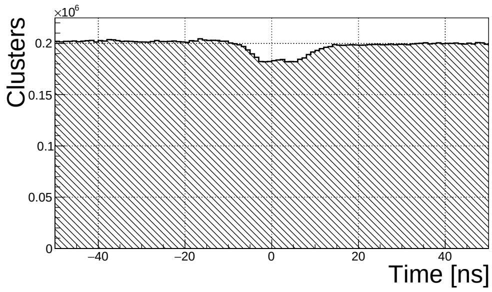
where N_{ev} is the total number of events from Tab. 4.4 and N_{clus} is the total number of neutral clusters from Tab. 4.5. This number is the base for the generated background rate later.

4.3.1.2 Resulting profiles

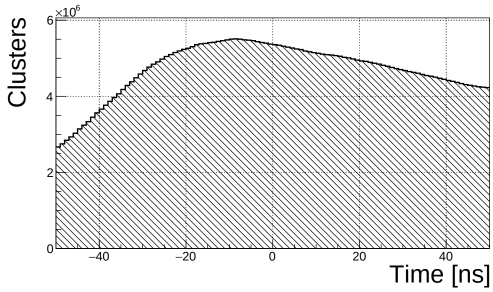
A set of distributions in time as well as energy was extracted for every calorimeter and every cell. The time distributions are shown in Fig. 4.12. In an ideal case, a flat distribution would be expected here. In Fig. 4.13 the energy distributions for each calorimeter are shown. For the electronic noise, a simple exponential distribution would be expected. It can be seen though, that especially ECAL1 and ECAL2 greatly differ from this expectation. In ECAL 2 a peak at around 2.2 GeV appears, and another at around 4 GeV. They are probably coming from undetected halo muons crossing the calorimeter. A leaking study was performed to see if these muons could be associated to another peak in the HCAL2, which is right behind ECAL 2. Unfortunately, the HCAL2 was too noisy in this energy region for any successful muon leaking veto.

4.3.2 Background Generation

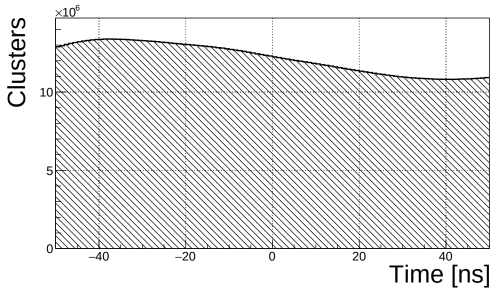
To generate the background from the extracted profiles, only three quantities need to be modeled. The absolute rate is already known from the first section. The distributions to model are:



(a) ECAL 0.



(b) ECAL 1.



(c) ECAL 2.

Figure 4.12: Time distributions for the neutral clusters as extracted from real data.

1. the time distribution,
2. energy distribution,
3. cell distribution.

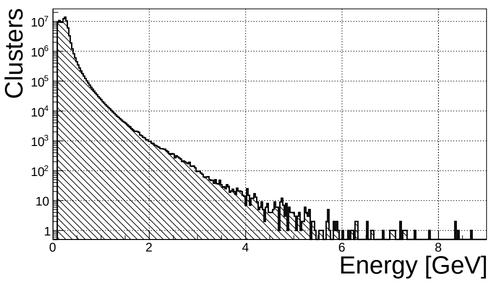
The time distribution is a feature of the calorimeter only, while the energy distribution is cell-dependent. This cell-dependence was already assumed, but it can be easily seen in the data too, by comparing energy distributions of several cells. Therefore the energy distributions need to be extracted and simulated cell-wise. And because of this, a cell distribution is necessary to take into account that some cells are more active than others.

4.3.2.1 Time distribution

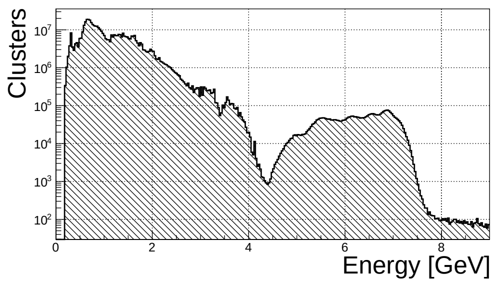
The time distribution, as visible in the Fig. 4.12 is not completely flat. Nevertheless, there are no very distinct features requiring any special modeling here. Systematics may be due to the extraction method. It was decided to use only a flat roll for the time with $t \in [-50, 50]$ ns, because this is the typical timegate of the calorimeters at COMPASS-II. The flat roll is the only roll that makes sense from the point of view of physics.

4.3.2.2 Energy distribution

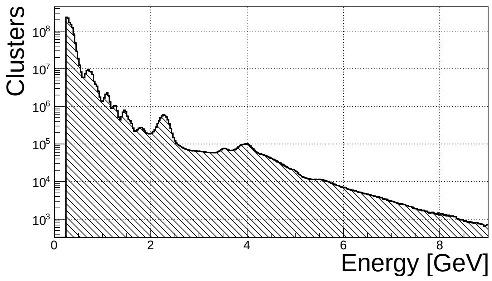
The energy distribution is highly cell-dependent, which is why it is only possible to use cell-wise profiles. This is clear considering that the background is a convolution of electronic noise and undetected particles. Undetected particles from the halo are more



(a) ECAL 0.



(b) ECAL 1.



(c) ECAL 2.

Figure 4.13: Energy distributions for the neutral clusters as extracted from real data.

common in the outer regions of the calorimeters, while the electronics for each cell is tuned individually and so inherently different. So because all contributions are cell-dependent in a non-trivial way, the only acceptable way is cell-wise extraction.

Discrete integral inversion method

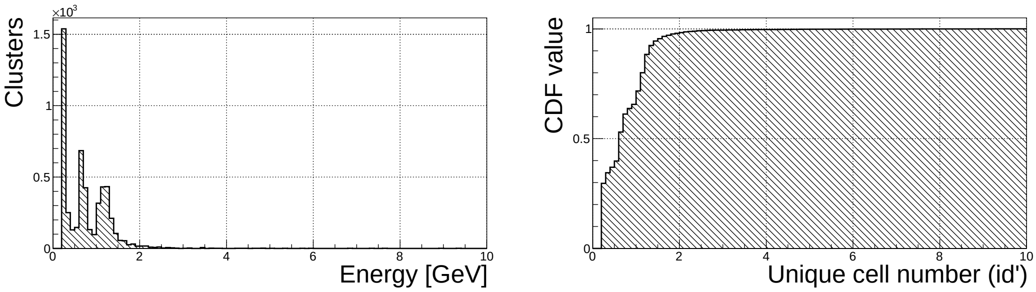
In order to cope with a finite number of bins, which is caused by finite statistics, and the profiles which are too complex to fit simple polynomials to them, the method of discrete cumulative density functions was employed. In the continuos case, this was already introduced in Sec. A. The new $f(n)$ is now the bin content of the energy distribution histogram in bin n. Then the CDF can then be defined as:

$$\mathrm{CDF}\,:\, g(n) = \frac{1}{\sum_0^{n,\max}} \sum_0^{\mathrm{n}} f(n). \tag{4.15}$$

With this CDF, which is a histogram itself, drawing values is simple. Instead of inverting the CDF, which is impossible, the trick is to simply find the first bin exceeding the drawn threshold from the left. In the case of energies, a continuous value is desired. To generate this, a second draw is added to the original bin center, with with r being a uniform draw in the bin width w_{bin}: $r \in [\frac{-w_{\mathrm{bin}}}{2}, \frac{+w_{\mathrm{bin}}}{2}]$:

$$e_{\mathrm{final}} = e_{\mathrm{center}} + r. \tag{4.16}$$

The uniform draw was chosen because there are thousands of cells, each with their own distribution. The uniform draw makes sure that no combination of surrounding bins lead



(a) Energy distribution.

(b) The resulting CDF.

Figure 4.14: The application of the discrete integral inversion method shown in the example of cell 1683 from the ECAL1.

to any complications, as there may be some distributions that are not well-behaved. These CDFs are calculated when creating the profiles with the tools **ecalNoiseAnalysis** and **ecalNoisePostProc** and then stored in a single ROOT file. An example of the application of this method is shown in Fig. 4.14.

4.3.2.3 Cell distribution

The distribution of the rate on the specific cells can be generated using the method already introduced for the energy distribution. A histogram is prepared in which the bins are filled with the following values:

$$f(id') = \sum_{0}^{n,\max} e(n, id), \tag{4.17}$$

where $e(n, id)$ is the energy value of bin n in the histogram of cell number id. The unified cell number id' was set to the new standard for this module in the electromagnetic calorimeter number i_{ECAL}:

$$id' = i_{\mathrm{ECAL}} \cdot 10\,000 + id. \tag{4.18}$$

This results in the histogram shown in Fig. 4.15a, from which another CDF can be built, which is shown in Fig. 4.15b. From this CDF the cell distribution can again be drawn.

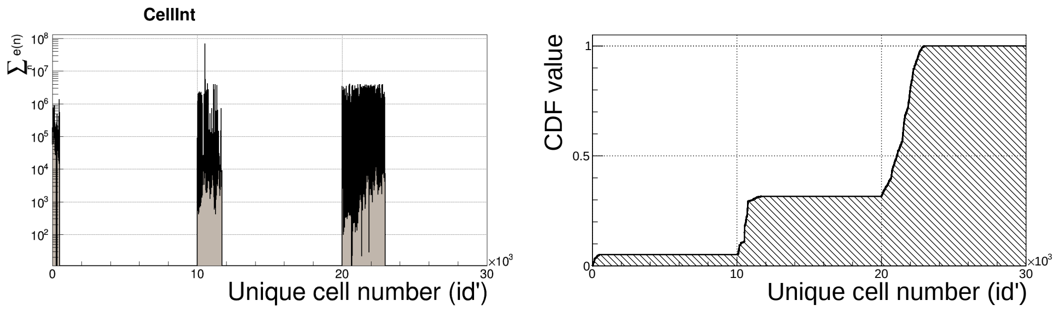
4.3.2.4 Rate

A Poissonian draw is used for each event in the Monte-Carlo in order to get the final cluster number from the rate of mean background clusters per event. The mean of the based poissonian distribution is set to the mean rate. The rate per cell and event can be interpreted as a baseline occupancy in the calorimeter modules. These are shown in Fig. 4.16.

4.3.3 Self-consistency Test

A whole set of pre cached CDFs was created for each cell in each electromagnetic calorimeter. To check the validity of the approach, the self-consistency test was conducted. This test was done using the **production_16-02** Monte Carlo sample with pure DVCS signal. The DVCS signal was cut out by a time cut:

$$|t_{\mathrm{cluster}}| \leq 2\,\mathrm{ns}. \tag{4.19}$$



(a) Energy integral distribution for the cells of each ECAL.

(b) The resulting CDF.

Figure 4.15: The final real data cell-wise energy distribution integrals plotted with their respective unique ids (id') and the resulting CDF after discrete inversion transform.

This also removes the background in this region, which is not perfectly correct, but in first order approximation seems fine, as it only introduces 2% uncertainty.

4.3.4 Energy Distribution

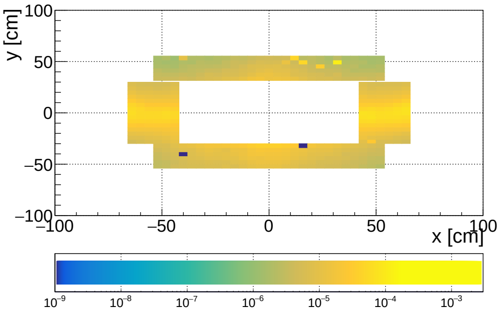
The energy distributions are shown in Fig. 4.17. These plots reproduce the data, which were already shown in 4.13, very well. The minor visible differences are coming from clusterization effects as well as the calorimeter smearing. All in all, a self-consistency test conducted for the energy distribution was successful.

4.3.5 Spatial Distribution

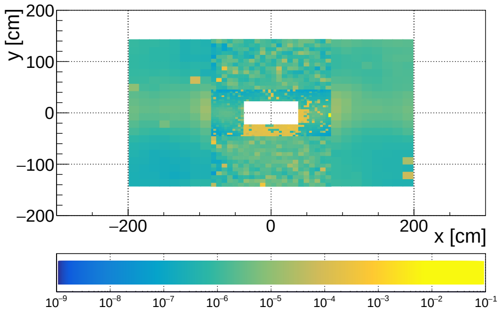
In the next step, the distribution of the rate along the cells needed to be checked and compared to the real data. The real data spatial distribution of the background rate was shown in Fig. 4.16. The reextracted rates are shown in Fig. 4.18. When comparing the two sets of rate distributions they agree very well. The slightly higher "clouds" are easily explainable by the fact, that the selected random triggers from the real data without beam are off spill, while the the Monte-Carlo sample features a heavy muon flux in each event. The clouds are unremoved halo muons that are already simulated with the TGEANT framework and cannot simply be removed, as the Monte-Carlo data contain no random trigger events.

Table 4.6: Trafdic options for the electromagnetic calorimeter background.

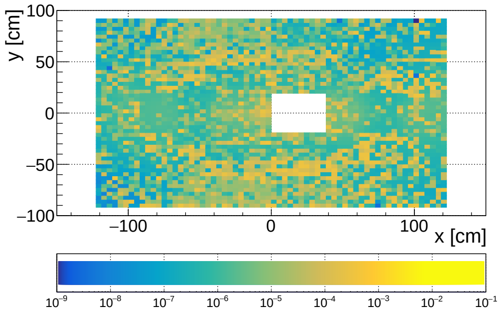
CsTGEANT option	**Value**	**Description**
EcalNoise	1	Enable the module.
EcalNoiseFile	[Path]	The CDF collection to use.
EcalNoiseRate	2.47810610	The number of clusters to add per event.



(a) ECAL 0.



(b) ECAL 1.

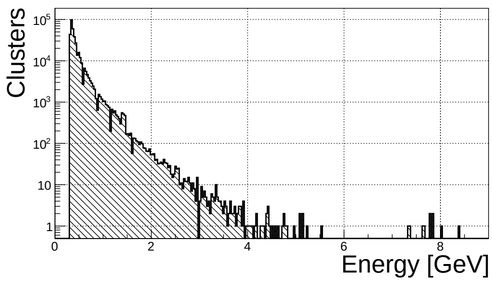


(c) ECAL 2.

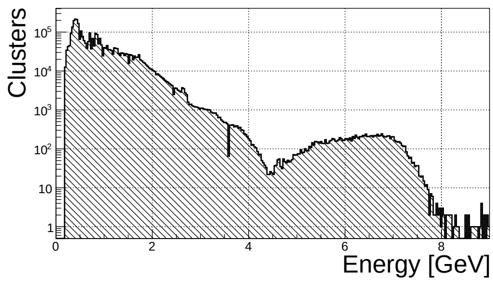
Figure 4.16: Background rates per cell and per event for each electromagnetic calorimeter. The respective palette is always drawn below the graphic.

4.3.6 Summary

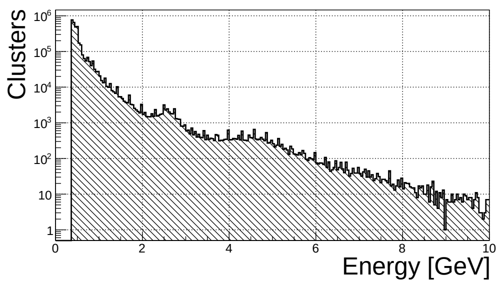
In this section, a completely new toolkit was presented, which allows cell-wise extraction of the energy dependent background of the electromagnetic calorimeters for the COMPASS-II experiment. The changes made to the reconstruction software, as well as the TGEANT software, are already upstream and made publicly available for all COMPASS-II physicists. It was shown that the extracted profiles are too complicated for the application of a simple polynomial fit. Furthermore, the method of discrete integral inversion sampling was introduced and applied to create discrete CDFs for each cell. A full set of CDFs were created from the complete data sample available at the time and documentation was written for future usage. Finally, a self-consistency test was conducted to verify the method and the CDFs used. It could be shown that the energy, as well as spatial distribution of the background clusters was in good agreement with the ones obtained from real data analysis. The results shown here, were already enabled in the creation of the Monte Carlo sample **production_16-02** that was widely used across different analysis works all across COMPASS-II. The trafdic options for enabling and controlling the module are listed in Tab. 4.6.



(a) ECAL 0.

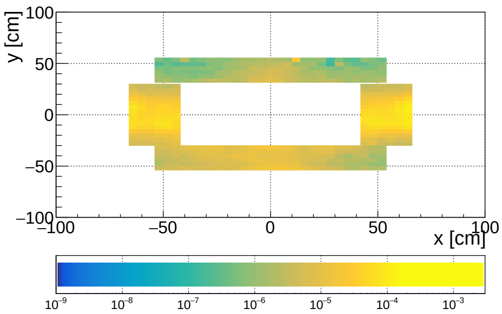


(b) ECAL 1.

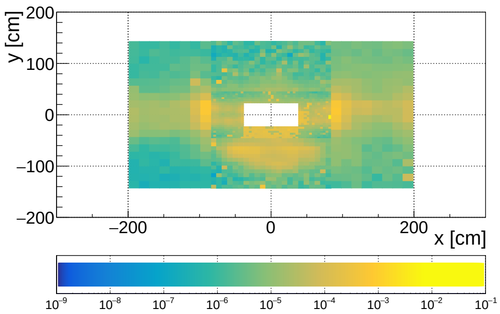


(c) ECAL 2.

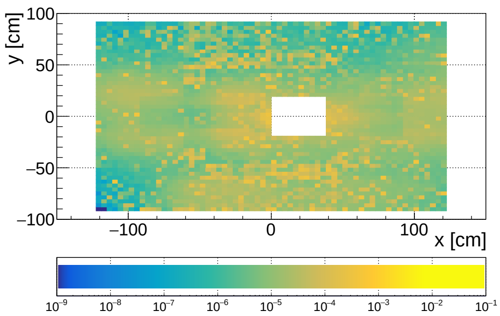
Figure 4.17: Energy distributions for the neutral clusters as extracted from Monte Carlo.



(a) ECAL 0.



(b) ECAL 1.



(c) ECAL 2.

Figure 4.18: Background rates per cell and per event for each electromagnetic calorimeter. The respective palette is always drawn below the graphic. These images were extracted from Monte Carlo.

5. HEPGen++ - an Exclusive Event Generator for COMPASS-II

This chapter presents the development of the event generator HEPGen++ for the exclusive production channels at the COMPASS-II experiment. It started as a port from the original HEPGen by A. Sandacz and P. Sznajder [64]. During its development, many issues were fixed and the design was improved. Now it is purely written in object-oriented C++. First, the general design of the generator is discussed and all kinematic equations are provided, see Sec. 5.1. In addition the data input and output options are described in Sec. 5.2 and it is shown how to control the generator. All of the used cross section models are discussed in the Sec. 5.3. HEPGen++ not only features many more cross section models and physics channels than the original HEPGen, but it also allows for evaluation of different models and parameter sets after reconstruction and analysis from the same data sample. This new approach can save an enormou amount of processing time by disentangling the theoretical model and parameters from the detector response and reconstruction. Whether an event gets successfully reconstructed and selected in an analysis is purely based on its kinematics. As the initial and final states of exclusive processes are always well defined, the kinematical variables, on which the reconstruction efficiency depends, can be used to compute to cross section according to different models. It is called "HEPGen++ in Phast" and it is covered in Sec. 5.4. For debugging and general utility, some helper libraries and tools were created. Their usage and interfaces are explained in Sec. 5.5. Finally, the results as well as recent applications of this generator are discussed in the summary, Sec. 5.6.

5.1 Design

HEPGen++ generates events for exclusive muoproduction of photons and mesons on nucleons. To generate an event, a valid kinematic needs to be randomly drawn. Valid means that neither the conservation of momentum, nor the conservation of energy can be violated. In general, two approaches for event generators are possible. The first one is a sampled-out non-weighting one. This approach is very intuitive, as kinematics, that are more likely to happen in the measurement just occur more often in the generated sample. While it is very easy to use afterwards, a complete scan of the full phase space can consume a lot of time and requires enormous statistics in the sample. Widely known examples for this generator class are PYTHIA [65] and LEPTO [66]. The second approach is the weighting event generation, were all the kinematics are generated uniformly. After the generation a differential cross section in all generated variables gets multiplied on top, as a weight. This leads also to the desired distributions in the final state. The advantage of this approach is that the sampling in areas with a low cross section is very accurate already at a lower statistical power. The downside, however, is that the regions of higher statistical power tend to get undersampled.

When it comes to HEPGen++, it was decided to take the advantages of both approaches. While the generator is a weighting one, the distributions of the drawn variables were modeled to be similar to the expected ones. This reduces undersampling in statistically strong regions, while, at the same time, keeping the flexibility of the weighted approach. In this way, the generation of the kinematics stays independent of the chosen generator, while the cross sections remain dependent on the drawn variables. This unified generation method is described in the following section, Sec. 5.1.1.

5.1.1 Generation of a Single Event

In this section, the generation of an exclusive muoproduction of a particle via virtual photon exchange is explained. The following method is almost completely independent of the produced particle(EPP[1]), which is quite useful since the same code can be used for all generators. As indicated in the theory chapter, a deep inelastic scattering reaction is always described by two independent variables. This can be (Q^2, x_{bj}) but all tuples of the variables ν, x_{bj}, Q^2, W are possible and equally valid. Choosing one set over another is typically motivated by technical reasons. The most important criterion is the resulting cross section that needs to be evaluated. If the cross sections that are going to be used are already in the parametrization of $\frac{d^2\sigma}{dQ^2 dx_{bj}}$, no additional work needs to be done. Otherwise Jacobian determinants need to be multiplied by them. For this generator, the set of (Q^2, ν) is chosen. The variable ν is a very nice observable for experimentalists as it is easy to interpret as energy transfer in the laboratory frame. Also, it has a rather intuitive scale which can be drawn flat, since all of the used cross section models predict a non-extreme behaviour in this variable. The second variable to draw is Q^2 as it is one of the most common variables in particle physics.

First of all, a ν according to user ranges is drawn. This limits the energy that is available for the rest of the processes:

$$\nu \in [\nu_{\mathrm{min,user}}, \nu_{\mathrm{max,user}}]. \tag{5.1}$$

[1]**E**xclusively **P**roduced **O**article

Variables with the index "user" are always taken from the datacard, which will be introduced later.

The next step is drawing a Q^2 from the desired user range. Naturally, the conservation of energy must not be violated, which is why the following constraint has to be kept: $Q^2 < \frac{2}{M_p}\nu x_{bj,\max}$. So finally Q^2 is drawn in:

$$Q^2 \in [Q^2_{\min}, Q^2_{\max}] \tag{5.2}$$

where minimal and maximal value of Q^2 are given to:

$$Q^2_{\min} = \max[Q^2_{\text{min,physical}} = -2m^2_\mu + 2(E\,E' - \vec{k}\cdot\vec{k}'), Q^2_{\text{min,user}}], \tag{5.3}$$

and

$$Q^2_{\max} = \min[Q^2_{\text{max,physical}} = \frac{2}{M_p}\nu x_{bj,\text{max,user}}, Q^2_{\text{max,user}}]. \tag{5.4}$$

The four-momenta of all particles in the events is first calculated in the centre of mass frame (CMS), because the equations stay shorter that way. Next, everything is transformed into the laboratory frame. By knowing the variables ν, Q^2 and the energy of the incoming beam muon E, the angle $\theta_{\mu'}$ of the scattered lepton with respect to the incoming beam lepton is already fixed:

$$\theta_{\mu'} = 2\arccos\left(\sqrt{1 - \frac{Q^2 - y^2 m^2_\mu c^2/(1-y)}{4|\vec{p}_\mu||\vec{p}_{\mu'}|}}\right), \tag{5.5}$$

where m_μ denotes the muon mass. The lepton momenta before and after the scattering are:

$$|\vec{p}_\mu| = \sqrt{E^2/c^2 - m^2_\mu c^2}, \tag{5.6}$$

$$|\vec{p}_{\mu'}| = \sqrt{(E-\nu)^2/c^2 - m^2_\mu c^2}. \tag{5.7}$$

To set the virtual photon γ^* from this, a ϕ angle needs to be drawn from the full range:

$$\phi_{\gamma*} \in [0, 2\pi]. \tag{5.8}$$

The θ_{γ^*} can be calculated from momentum and energy conservation to:

$$\theta_{\gamma^*} = \arccos\left(\sqrt{1 - \left(\sin(\theta_{\mu'})\frac{|\vec{p}_\mu|}{\sqrt{\nu^2/c^2 + Q^2}}\right)^2}\right). \tag{5.9}$$

The next step is deciding whether the reaction is "elastic", meaning the proton stays fully intact or if the target enters an excited intermediate state that decays afterwards. This excitation with following decay is called "diffractive dissociation" and is covered in Sec. 5.3.10.

Following this, there is the generation of the four momentum transfer t. It is easier to generate t' which starts at 0 though, as the functional dependence of most cross sections here is only a simple exponential function. Therefore this observable is drawn in the range of

$$t' \in [0, t'_{\mathrm{max}}]. \tag{5.10}$$

t' is related to t by correcting the so-called t_0 term that describes the minimal momentum transfer to make the given kinematics valid. For the equation see the Tab. 2.1 in the theory part.

The kinematics for the EPP are now almost fixed. Another ϕ angle needs to be drawn according to

$$\phi_{\mathrm{EPP}} \in [0, 2\pi], \tag{5.11}$$

for the azimutal distribution of the EPP. To fix the θ angle, again conservation of momentum and energy can be used:

$$\theta_{\mathrm{EPP}} = \arccos\left(\frac{t + Q^2 - m_{\mathrm{EPP}}^2 c^2 + 2E_{\gamma^*} E_{\mathrm{EPP}}/c^2}{2P_{\mathrm{in}} P_{\mathrm{out}}}\right), \tag{5.12}$$

where E_{EPP} is the energy of the exclusively produced particle and m_{EPP} is the mass of it. The momentum sums P_{in} and P_{out} can be written as:

$$P_{\mathrm{in}} = \frac{\sqrt{W^4c^4 - 2W^2c^2(m_p^2c^2 - Q^2) + Q^4}}{2Wc}, \tag{5.13}$$

$$P_{\mathrm{out}} = \frac{\sqrt{W^4c^4 - 2W^2c^2(m_a^2 + m_{\mathrm{EPP}}^2)c^2 + (m_a^2 - m_{\mathrm{EPP}}^2)^2c^4}}{2Wc}, \tag{5.14}$$

where m_a^2 is the mass of the excited target state squared. In case of an “elastic” event, it is simply the proton mass squared:

$$m_a^2 = m_p^2. \tag{5.15}$$

In the case of diffractive dissociation, this is not the case. The diffractive dissociation is covered in Sec. 5.3.10. The next step is the transformation into the laboratory frame, where the simulation will take run afterwards. For this the following matrix is constructed from the three-momentum of the $\vec{p}_{\gamma *}$ and the three-momentum of the incoming lepton $\vec{k}$:

$$M = \begin{pmatrix} \vec{h}_x & \vec{h}_y & \vec{h}_z \end{pmatrix}, \tag{5.16}$$

where the helper vectors are defined as:

$$\begin{aligned} \vec{h}_x &= \hat{\vec{p}}_{\gamma *}, \\ \vec{h}_z &= \frac{1}{|\vec{k}|}(\vec{h}_x \times \vec{k}), \\ \vec{h}_y &= \vec{h}_z \times \vec{h}_x. \end{aligned} \tag{5.17}$$

All of the particles get rotated using this rotation matrix. Afterwards, a standard Lorentz boost is applied to the resulting four-momentum vectors. In the laboratory system the recoiled particle four-momentum is easily writable as:

$$P_{p'} = (\nu - E_{\mathrm{EPP}} + m_p c^2,\ \vec{p}_{\gamma *} - \vec{p}_{\mathrm{EPP}}). \tag{5.18}$$

Finally, all of the particles get rotated once more, so the incoming beam gets put on the z axis.

5.1.2 Sampling Choices

The distributions of Q^2 and t' were made with the integral inversion method, which is introduced and elaborated about in the appendix A. A $1/x$ distribution was chosen for Q^2 to have more statistical power where the real data has more statistical power. Higher powers such as $1/x^2$ or even $1/x^4$ would be applicable, which can be motivated by the cross section equations. As these show destinct dependences on the power of Q^2, the sampling could be optimized. The closer the chosen distribution is to the dependence of the cross section, the better, since the statistical power is concentrated, where it is needed most. Therefore, a representative sample of the complete phase space can be generated with less events, than with a non-optimized sampling.

In all of the following equations r shall be a uniform random number:

$$r \in [0, 1]. \tag{5.19}$$

For Q^2 the CDF is:

$$Q^2 = Q^2_{\text{min}} \left(\frac{Q^2_{\text{max}}}{Q^2_{\text{min}}} \right)^r, \tag{5.20}$$

and the phase factor is:

$$pf_{Q^2} = \frac{1}{Q^2} \frac{1}{\ln(Q^2_{\text{max}}/Q^2_{\text{min}})}. \tag{5.21}$$

In the case of t' distribution an exponential dependence is assumed with an arbitrary slope. The CDF for this distribution is:

$$t' = -\frac{1}{B} \ln[\exp(-Bt'_{\text{min}}) - \{\exp(-Bt'_{\text{min}}) - \exp(-Bt'_{\text{max}})\}r)], \tag{5.22}$$

and the phase factor becomes:

$$pf_{t'} = B \frac{\exp(-Bt')}{\exp(-Bt'_{\text{min}})} - \exp(-Bt'_{\text{max}}), \tag{5.23}$$

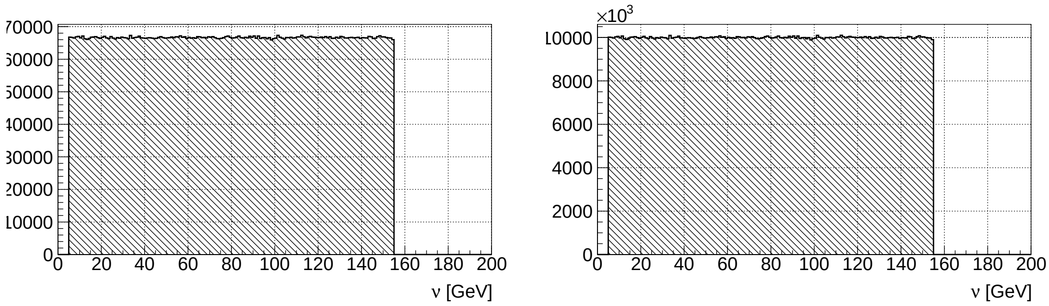
where B is the t' slope. When it comes to the ν draw, another method is used. This variable is to be drawn uniformly in a user specified interval:

$$\nu = \nu_{\text{min}} + (\nu_{\text{max}} - \nu_{\text{min}})r. \tag{5.24}$$

For normalization reasons it is very comfortable to also scale this integral back to one. Therefore, another phase factor is introduced for this roll:

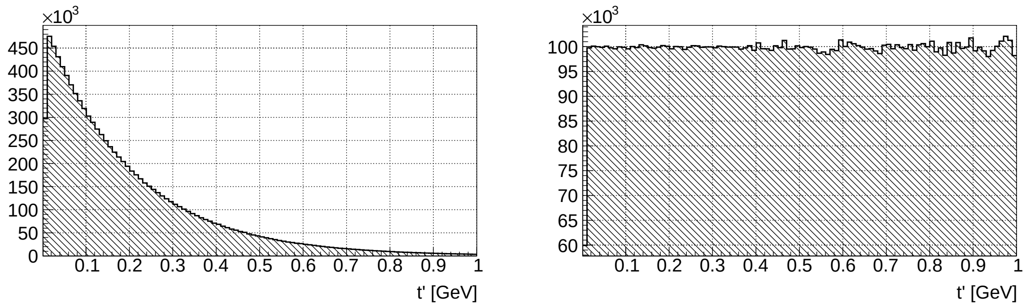
$$pf_\nu = \frac{1}{\nu_{\text{max}} - \nu_{\text{min}}}. \tag{5.25}$$

The distributions unweighted and with their respective phase factor weights are plotted in the figures from 5.1a to 5.3b.



(a) The ν distribution without weights. (b) The ν distribution with phase factor weights.

Figure 5.1: ν distributions on generator level. Here, only the normalization changes, as the distribution is already flat. The whole data sample contains 10 million events, so the normalization is that in each 1 GeV bin there are 10 million entries.



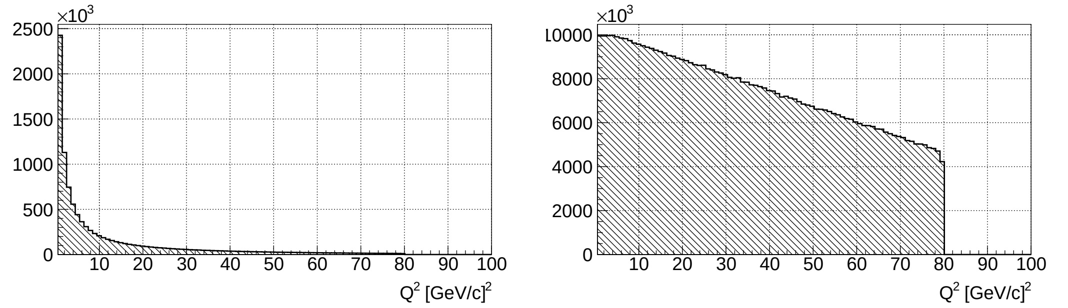
(a) The t' distribution without weights. (b) The t' distribution with phase factor weights.

Figure 5.2: t' distributions on generator level. Here the distribution is sampled very fine at small t' but then flattened out by the phase factor weights. As in the ν distribution the normalization puts the integral of an interval of 1 GeV/c to 10 million entries, which was the base statistics.

Overconstraints

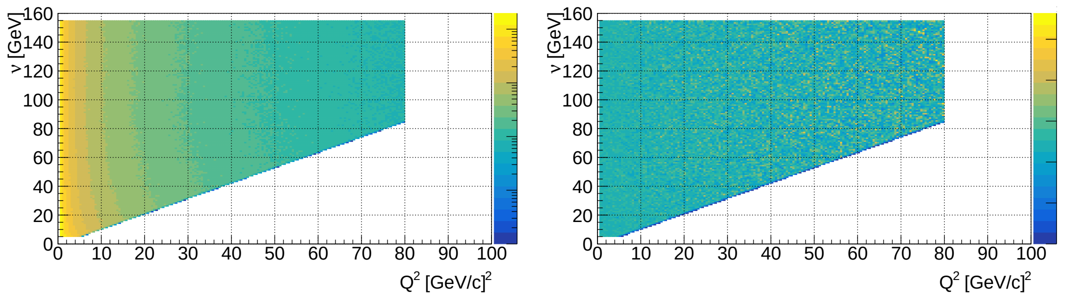
In the figures 5.3b and 5.4b a linear cut-off in Q^2 is visible. This is due to an overconstraint of the upperbound of Q^2. Remembering equation 5.4 it can be seen that x_{bj} limits can cut into the upper boundary more sharply than the user chosen Q^2. This rises two issues, the first of whom is the phase factor. This factor was introduced to remove the functional dependence $\mathcal{D}(x)$ from a non-uniform sampling in x. The non-uniform sampling works fine, as is visible in the Fig. 5.4a. The behaviour after phase-factor application is shown in 5.4b, where the whole region is homogenious.

The second problem is the normalization of the events with this flat distribution, since the number of filled bins in Q^2 is not the same for every ν. This problem is encountered when trying to calculate Monte-Carlo luminosity, where a proper integration over the cross section is required. An exemplary solution to this problem is given in the appendix B.



(a) The Q^2 distribution without weights. (b) The Q^2 distribution with phase factor weights.

Figure 5.3: Q^2 distributions on generator level. The slope is coming from energy conservation in combination with an invisible user cut on x_{bj}. See overconstraints section for a detailed explanation.



The Q^2 vs ν distribution. The sampling is more fine at small Q^2. Mind the logarithmic z scale in this picture.

The Q^2 vs ν distribution weighted with a correct dynamic phase factor. It flattens out the phase space in two dimensions.

Figure 5.4: Q^2 distributions on generator level. The slope is coming from energy conservation in combination with an invisible user cut on x_{bj}. See overconstraints section for a detailed explanation.

5.1.3 Software Design Principles

When designing HEPGen++ a few principles where set that were kept throughout the development of the whole software package. Since the ancestor version of HEPGen had problems running on modern operating systems and compiling on modern compilers the long time compatibility was the key issue. To ensure that in a few years this generator will still work the way it was intended, it was necessary to:

- not use any third-party libraries,
- make the code as standard compliant as possible,
- retain readability and understandability,
- prioritize maintainability over computational performance.

In the following sections these principles are shortly elaborated on and their influence on the final design is demonstrated.

Code Portability

Particle physics experiments at CERN have a history of having very inhomogeneous computer architectures as well as compiler and operating system versions. Therefore, the code of an event generator that will possibly be used for a long time has to be as portable as possible. In order to achieve that, no third-party library was used for the first part. All classes for vectors, four-vectors, as well as their mathematical methods were implemented in standard C++ 98. The old HEPGen became difficult to compile because of the usage of paw and cernlib, which was originally thought to be supported forever by CERN. To avoid the dependancy problem in HEPGen++, it was decided to keep it indepedent of any external libraries. A reliable random number generator, the James Random [67], was taken from CLHEP library [68] and embedded in the HEPGen++ repository.

The second design principle that should have lead to well-portable code, was keeping the standard strictly. The old C++ standard was chosen because it compiles on every C++ compiler even from the oldest GCC[2] installed on older machines. Even though it would be fast, risky unchecked direct memory operations like **memcpy** were mostly avoided. This greatly reduces the risk of memory access violations and problems with compiler optimizations.

Readability and Maintainance

Finally, the third and fourth point mentioned above, readability and maintainability, can be evaluated together. Strict object-oriented design is generally speaking not the fastest running code but it allows for a good maintainability by removing the need of several copies of the same code. An example of this approach is moving all of the kinematics related code to a base class called **HPhysicsGen** from which all specialized generators are derived. Removing a possible bug in this section of the code will automatically fix it for all implemented generators. Also, the readability of the code is greatly enhanced using by this method. A vector product written out in its components may be faster in FORTRAN than the method call of a vector-object, but it is much clearer to any new author who will take over the maintenance work. The example of **HPhysicsGen** is shown in the Fig. 5.5.

5.2 Data Input and Output

Most of the input and output formats were designed in an object-oriented inherited approach. This will be called "backend" design from now on, since the actual implementation of a format is done in derived backend classes, while all the logic and calls are handled through virtual base classes. The backend design allows seamless switching between the used backends and in some cases, for example the histograms, even enabling more than one at the same time.

5.2.1 Beamfile

The so-called beam file is a binary file that contains entries of beam particles measured in the experiment at one time. In its entirety, it should contain the complete beam phase space in momentum direction and energy as well as spatial coordinates. The format used

[2]**G**NU **C**ompiler **C**ollection

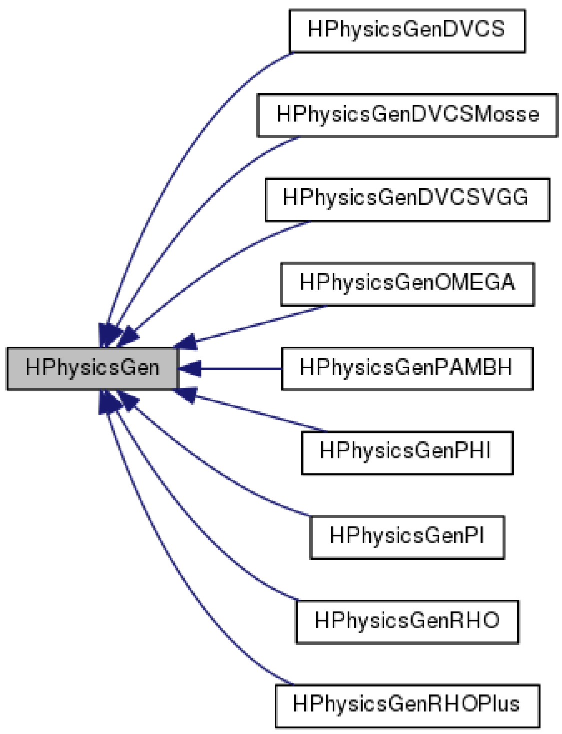


Figure 5.5: The inheritance diagram for **HPhysicsGen** generated directly from the C++ code with Doxygen. Most of the common code, which is the diffractive dissociation, muon kinematics and more are placed in the base class, while each generator only implements its cross section function as well as output particles.

at COMPASS was originally introduced as "xlaugato beam database format", though it is more a binary stream of entries than a real database. Just like all other binary formats that inherit the FORTRAN legacy, this one also suffers from the header length change that was introduced when the F77 compiler was dropped in favor of gfortran. HEPGen++ automatically looks for the right header size and reads the beam file regardless of the compiler of the xlaugato writer program that was used in the creation of the file. An illustration of the format is shown in the appendix C.

To calculate the momentum and position vector of a particle from this information, the following computations have to be made. All three components of the momentum need to be calculated from the slopes and the kinetic energy in order to get the full momentum vector. The equations for this are written in components of $\vec{k}$ as:

$$k_z = \frac{E_{\text{kin}}/c}{\sqrt{1 + \sin(\mathrm{d}x \cdot 10^{-3})^2 + \sin(\mathrm{d}y \cdot 10^{-3})^2}}, \tag{5.26}$$

$$k_x = k_z \sin(\mathrm{d}x \cdot 10^{-3}), \tag{5.27}$$

$$k_y = k_z \sin(\mathrm{d}y \cdot 10^{-3}), \tag{5.28}$$

where the beam is aligned along the z axis. To complete the four-momentum vector for the incident muon k, the mass has to be added:

$$k = (\sqrt{\vec{k}^2 c^2 + m_\mu^2 c^4}, \vec{p}). \tag{5.29}$$

Only the x and y component of the position of the particle is stored in the file. The x and y values are already in the coordinate system where the beam is aligned along the z axis, meaning that x and y read from the beam file give the transverse position. The longitudinal position z along the beam axis is set to 0 mm by convention for the muon program at COMPASS. For the hadron and Drell-Yan programs this number is implicitly assumed to -7.5 m.

The so-called "beam type" is a single number denoting whether this entry is a valid beam particle or a halo particle. In order to simulate signal events, only the beam (flag = 1) entries should be used, since they can intersect the target and their momentum resolution is good. The halo entries (flag = 2) have a much worse momentum resolution due to the nature of their measurement and will most probably not intersect the target material in the simulation.

5.2.2 LEPTO Output

The LEPTOv2 format was a common format in particle physics a while ago but is not commonly used anymore. However, this format was added in order to have a good backward compatibility. Like the beam file format, this format was originally introduced in FORTRAN, so it suffers from the same problems with the undefined length of the start word. HEPGen++ allows for different start and stop word lengths. They can be switched in the data card. In the appendix C the format is visualized. It consists of a header block that contains general information about the parameters that were in use for the production of the file. The event block that is written for each event contains the event-specific information. A complete description of each of the vectors and variables that are written here is given in the appendix to improve the legibility of this document.

5.2.3 Datacard

The datacards that control HEPGen++ are for the most standardized and even backwards compatible with the original HEPGen datacards. It is not directly possible to load HEPGen datacards into HEPGen++, since some editing is needed to remove all the line breaks for a given keyword. Not all of the variables in the card are actually used; some are used depending on the generators, while others are not used at all and their sole purpose is backward compatibility. The specialties of HEPGen++ are introduced in this section.

Generation Ranges

In principle, there are only a few relevant kinematic variables that constrain the generation. These are Q^2, ν, t' and $x_{bj,\max}$. To stay consistent with the generation of HEPGen $x_{bj,\max} = 1.0$ should also be kept. The values for Q^2 and ν are written into the **CUTL** line, while the t' limits have their own line called **TLIM**.

Beam Reading

The **BMRD** flag controls the beam file reading. Setting it to non-zero enables reading the beam file from the flag **BEAMFILE**. The flag itself has the following interpretation:

- 1: Read beam entries only,
- 2: Read halo entries only,
- 3: Read all entries available.

Generator Selection

The first number in the flag **VMES** controls the selection of the generator. The second number denotes the decay channel if the exclusively produced particle decays. As only fixed decay modes are implemented as of now, the second flag is not used yet. A list of implemented choosable generators is presented in list 5.1. The code by H. Moutarde is currently only restrictedly available in Freiburg and cannot be put in a public repository. Once his "PARTONS" framework [69] is released, a solution for integrating it with the generator can be found. For example a table containing amplitude of the different DVCS models could be loaded. As the DVCS amplitude is well-behaved in the current models, interpolating in this grid should not pose any problems. The calculation of the interference terms and the Bethe-Heitler terms can then be done in HEPGen++. A grid of final cross section values will be very hard to generate, since the Bethe-Heitler contribution has extreme slopes in some regions. This poses different problems already, and is discussed in the DVCS model sections (5.3.7 ff).

Further Flags

Further interesting flags are listed in Tab. 5.2.

Table 5.1: Flag choices for the first **VMES** flag.

VMES	Produced particle	Model
0	γ	Modified Strikman & Freund [70, 71]
1	$\pi^0 \to \gamma\gamma$	Goloskokov & Kroll 2006 and 2011 [72]
2	$\rho^0 \to \pi^+\pi^-$	Cross section measurements from HERMES [73]
3	$\phi \to K^+K^-$	Cross section measurements from HERMES [73]
6	$\omega \to \pi^+\pi^-\pi^0(\to \gamma\gamma)$	Cross section measurements from HERMES [73]
7	$\rho^+ \to \pi^+\pi^0(\to \gamma\gamma)$	Goloskokov & Kroll table [74]
8	$\omega \to \gamma\pi^0(\to \gamma\gamma)$	Cross section measurements from HERMES [73]
12	γ	VGG model implementation by L. Mosse. [75]
13	γ	Bethe Heitler by P.A.M. Guichon and A. Vidon. [76]

Table 5.2: Additional flags for HEPGen++

Flag	Value	Description
NGEV	long int	Number of events to generate.
DIFF	**0** / 1	Use of diffractive dissociation.
ENABLE_GFORTRAN	**0** / 1	Enables long startwords for LEPTOv2 format.
OUTFILE	[path]	Sets the name of the output file.
HISTOS_ASCII	**0** / 1	Enables kinematic histograms in ASCII format.
HISTOS_ROOT	**0** / 1	Enables kinematic histograms in ROOT format.
ENABLE_DEBUG	**0** / 1	Enables a lot of verbose debug output.
AUX1	-	Special settings for some generators.
AUX2	-	More special settings for some generators.

5.2.4 Histogramming

Adding histograms to the creation of a sample is quite simple. Several one-dimensional and two-dimensional histogram backends are available. If the ROOT framework is available and ROOT histogramming is enabled (Tab. 5.2), they will be put inside a ROOT file. The second option is to enable ASCII histogramming, in which case they will be written as plain text into a file. The system allows for weighted histogramming with double precision. The format is visualized in a graphical way in the appendix C. As mentioned there, for n bins the histogram has $n + 2$ data lines. The underflow bin is in the first line, the overflow bin in the $n + 1$th line. This allows for a good evaluation of the generated sample even without installing the ROOT framework.

5.3 Implemented Models

To implement a generator for HEPGen++ in the isotropic case, only a parametrization for the mass of the produced particle is needed as well as a parametrization for the cross section in dependence of any set of kinematical variables.

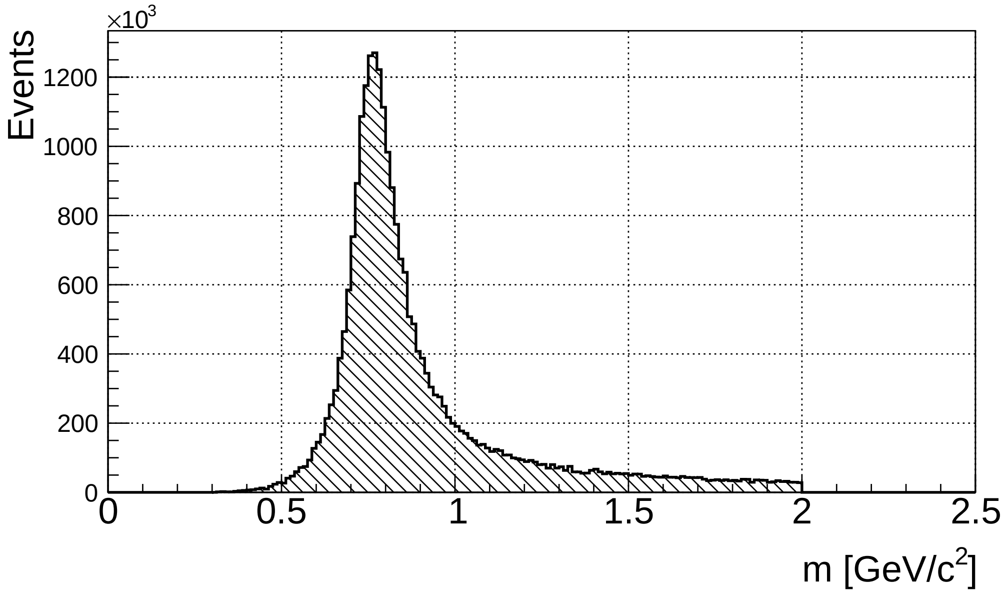


Figure 5.6: The mass distribution of the generated ρ^0 mesons.

5.3.1 Exclusive Rho: $\rho^0 \to \pi^+\pi^-$

The ρ^0 generator was the first generator to be implemented in the original HEPGen. For compatbility reasons, the goal was to implement it in C++ binary identically. Therefore, the mass distribution for the ρ^0 particle as well as the cross section needed to be computed in a similar way.

Mass Distribution

The mass distribution of the ρ^0 meson that almost always decays to two pions needs to be simulated according to the real data distributions. The $\frac{\mathrm{d}N}{\mathrm{d}M_X}$ distribution follows a relativistic Breit-Wigner distribution. The typical parametrization is introduced in [77]. The parameters for ρ^0 production were measured at HERA by the ZEUS collaboration [78]. The implemented mass distribution in HEPGen++ follows the distribution:

$$D(m) = \frac{m\, m_{\rho^0} g_{\rho^0} g^*_{\rho^0}}{\left((m)^2 - m^2_{\rho^0}\right)^2 + m^2_{\rho^0}(g^*_{\rho^0})^2}, \tag{5.30}$$

where

$$g^*_{\rho^0}(m) = g_0 \left(\frac{\sqrt{m^2 - m^2_{\pi^\pm}}}{q_0} \right)^3, \tag{5.31}$$

is the momentum dependant width with the constants $g_0 = 0.153\,\mathrm{GeV}$ being the mean width, $q_0 = 0.358\,\mathrm{GeV/c}$ being the momentum of the pion in the CMS and $m_{\rho^0} = 0.77\,\mathrm{GeV/c^2}$ is the mean mass of the ρ^0 meson. The mass of a charged pion is set to $m^2_{\pi^\pm} = 0.019488\,\mathrm{GeV^2/c^4}$. The resulting distribution is shown in Fig. 5.6.

Polarization

The polarization of the vector mesons like the ρ^0 can be measured by extracting angular distributions of their decay products. The ratio of the longitudinal cross section to transversly polarized one is typically denoted R with:

$$R = \frac{\mathrm{d}\sigma_{L,\gamma*p}}{\mathrm{d}\sigma_{T,\gamma*p}}. \tag{5.32}$$

By using generalized vector dominance and s-channel helicity conservation a parametrisation can be derived for this ratio. The complete derivation and fits to data was done in [79]. The parametrization can be fully written as:

$$\begin{aligned} R_V(Q^2) = & \frac{(Q^2 + m_{V,T}^2)^2}{m_{V,T}^4} \xi_V^2 \\ \times & \left[\frac{\pi}{2} \frac{m_{V,L}^2}{Q^2} - \frac{m_{V,L}^3}{\sqrt{Q^2}(Q^2 + m_{V,L}^2)} - \frac{m_{V,L}^2}{Q^2} \arctan\left(\frac{m_{V,L}}{\sqrt{Q^2}} \right) \right]^2, \end{aligned} \tag{5.33}$$

where ξ_V is the ratio of imaginary forward scattering amplitudes of longitudinally and transversely polarized vector mesons. It is theoretically a function of the kind of vector meson as well as W. The mass terms $m_{V,L}$ and $m_{V,T}$ are pole masses that are also parameters for the model. All of the parameter values from [79] are listed in Tab. 5.3. The fraction of longitudinally polarized mesons can then be written as:

$$f_L = \frac{(\varepsilon + \Delta) R_V}{1 + (\varepsilon + \Delta) R_V}, \tag{5.34}$$

where Δ is the difference in longitudinally to transversely polarized virtual photons given by:

$$\Delta = 2m_\mu^2 \frac{1 - \varepsilon}{Q^2}. \tag{5.35}$$

Table 5.3: Parameters for the Schildknecht parametrization for the polarization ratio.

Meson	ξ_V	m_T^2	m_L^2
ρ^0 - paper four parameters	1.06	$0.68\, m_\rho^2$	$0.71\, m_\rho^2$
ϕ - paper four parameters	0.9	$0.43\, m_\phi^2$	$0.60\, m_\phi^2$
ρ^0 - paper two parameters	1.0	$0.62\, m_\rho^2$	$0.93\, m_\rho^2$
ϕ - paper two parameters	1.0	$0.4\, m_\phi^2$	$0.6\, m_\phi^2$

Decay

The decay of the ρ^0 meson into the pions is, as it is a two-body problem, well calculable and straight-forward. With the ρ^0 being a vector meson though, the angular distribution of the decay products in θ depends on the polarization. For a longitudinally polarized vector meson, the following equations for the angle are used:

$$r \in [0, 1], \tag{5.36}$$

$$\cos(\theta) = \begin{cases} \sqrt[3]{1 - 2r} & \text{for} \quad r < 0.5 \\ \sqrt[3]{(2r - 1)} & \text{else.} \end{cases} \tag{5.37}$$

For transversely polarized vector mesons, the distributions for the angle θ look like:

$$r \in [0, 1], \tag{5.38}$$

$$\theta' = \arccos(|2r - 1|), \tag{5.39}$$

$$\cos(\theta) = \begin{cases} -2\cos\left(\frac{\theta'+\pi}{3}\right) & \text{for} \quad r < 0.5 \\ +2\cos\left(\frac{\theta'+\pi}{3}\right) & \text{else.} \end{cases} \tag{5.40}$$

Cross Section

The cross section of the exclusive muoproduction of ρ^0 was measured in dependence of Q^2 by the NMC collaboration [73]. It was parametrized according to the general function:

$$\frac{\mathrm{d}\sigma_{\gamma^* p}}{\mathrm{d}Q^2} = \sigma_0 \left(\frac{Q_0^2}{Q^2}\right)^\beta, \tag{5.41}$$

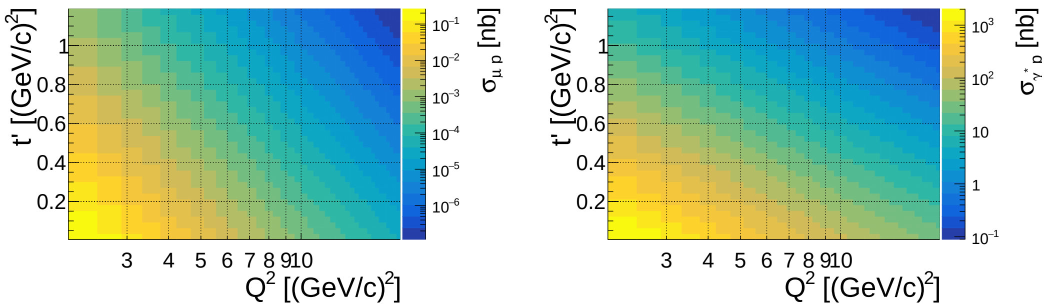
where the fit parameters are $Q_0^2 = 6.0\,(\mathrm{GeV/c})^2$, $\beta = 1.96$ and the base cross section is $\sigma_0 = 27.4\,\mathrm{nb}$. To account for a further t' dependence this cross section was modified with an exponential t' slope. According to [78] this is justified in first order approximation. The resulting term can now be written as

$$\frac{\mathrm{d}^2\sigma_{\gamma^* p}}{\mathrm{d}Q^2\mathrm{d}t'} = \sigma_0 \left(\frac{Q_0^2}{Q^2}\right)^\beta b_t \exp(-b_t t'), \tag{5.42}$$

where $b_t = 5.0$. This cross section needs to be multiplied with the virtual photon flux to get the muon proton cross section. The transverse virtual photon flux can be written as [73]:

$$\Gamma_T = \frac{\alpha(\nu - Q^2/2m_p)}{2\pi Q^2 E_\mu^2 (1 - \varepsilon)}, \tag{5.43}$$

where ε is the virtual photon polarization from equation 2.90. Finally, the resulting cross section equation is renormalized with a factor of $f_{GK} = 2.33$ to make the parametrization hit the prediction by Goloskokov and Kroll model. The total cross sections of NMC were known to be different from all other experiments in their absolute value [80], so renormalizing to the currently best model prediction while, at the same time, keeping the functional dependence measured by the experiment is justified. The cross section finally resulting from the measurements and calculations above is shown in Fig. 5.7a, the direct virtual photon proton cross section is shown in Fig. 5.7b.



(a) The full muon proton cross section corrected by the transverse photon flux.

(b) The virtual photon proton cross section only.

Figure 5.7: Cross sections for exclusive ρ^0 production as implemented in HEPGen++

5.3.2 Exclusive Phi: $\phi \to K^+K^-$

The ϕ generator was first introduced in HEPGen++ for the purpose of a master thesis by A. Gross where the analysis of COMPASS data for exclusive production of the phi mesons was the main focus.

Mass Distribution

As the ϕ distribution is very narrow, the width being at about 4 MeV, the relativistic Breit-Wigner distribution is omitted and a gaussian smearing is performed instead. A random mass difference is drawn from a gaussian distribution with mean $\mu = 0\,\mathrm{MeV/c^2}$ and width $\sigma = 4\,\mathrm{MeV/c^2}$. This is then added to the PDG mass for the final mass.

Polarization

The polarization is generated in the same way as for the rho generator, but with the parameters given for the phi mesons in Tab. 5.3.

Decay

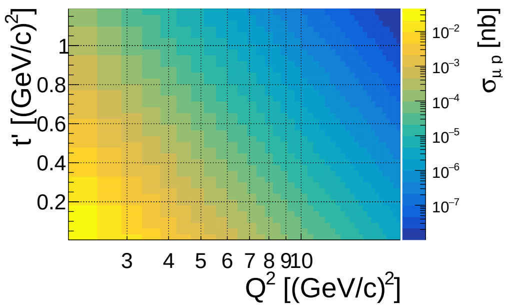
The decay is handled by the same calculations code like the ρ^0 meson generator, except that the decay particles are kaons with a different mass. The angular distribution is assumed to be identical.

Cross Section

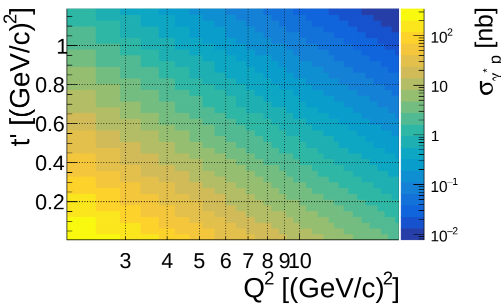
In the case of the cross section of the phi production, the same parametrization was used as the one for the rho. The parameters are slightly different and can be found in [73] to:

$$\beta = 2.27,$$
$$\sigma_0 = 3.425\,\mathrm{nb}.$$

The same normalization factor for these values to the GK model as for the rho production was used. The cross section looks very similar to the exclusive rho production, which is not surprising, since it follows the same parametrization. The cross sections are shown in the Fig. 5.8a and Fig. 5.8b. As visible, only the Q^2 slope and the absolute value of the cross sections differ from the ρ^0 generator.



(a) The full muon proton cross section corrected by the transverse photon flux.



(b) The virtual photon proton cross section only.

Figure 5.8: Cross sections for exclusive ϕ production as implemented in HEPGen++. The difference to the ρ^0 production is the Q^2 slope and the absolute scale of the cross section.

5.3.3 Exclusive Pions: $\pi^0 \to \gamma\gamma$

The field of exclusive π^0 production has recently gained a lot of momentum on the theoretical and experimental side alike. The pilot run of the COMPASS-II experiment, performed in 2012, delivered the first ever real exclusive neutral pion production data. This brought up the need to have a neutral pion event generator. The loading of the cross section

Table 5.4: AUX Flags for the π^0 generator

Flag	Value	Description
AUX1	[path]	Path to production cross section table.
AUX2	-1	Load self-consistent header containing cross section table. (-ng format)
AUX2	0	Load old HEPGen binning file.
AUX2	1	Load old HEPGen binning file with $\bar{E}_T$ contribution.

tables is controlled via the **AUX** flags in the datacard. The possible values for this generator are given in Tab. 5.4. In general, only the new format should be used. The standard datacards use these new tables (-ng-format, AUX2=-1) exclusively. Nevertheless, the option to load the old format persists. Note that in the old format the binning for the kinematics was hardcoded. Therefore, any modification in the binning of the table will result in a program crash.

Mass distribution

As with the phi meson, the pions have a very narrow width of only about $\Gamma = 8\,\mathrm{eV}$, so the mass is simply put to $m_{\pi^0} = 134.977\,\mathrm{MeV}$.

Cross section

The cross section for this generator offers two choices, both of which are based on tables for the Goloskokov and Kroll model for the exclusive neutral pion production. One is based on the older GK model, without transverse GPD influence(“GK09” [45]), while the second one is based on the latest official version of their model including transverse GPDs (“GK11” [72]). The latest version of the GK11 model is also completely implemented starting from the GPDs. As the implementation is very complex, a whole chapter (Chapter 6) is devoted specifically to it. In this section, only the tabulated implementation is presented.

HEPGen++ introduced a unified data table format using this parametrization, which is described in the appendix and which allows for changes to be made in a simple way. The parametrization is three dimensional in W, Q^2 and $-t'$. The cross section calculation from these tables is handled by first finding the bin, which is closest to the desired value. Each kinematic variable has to be checked from this bin, in order to determine whether the extrapolation has to be done forward or backward. Next, a linear extrapolation is done to get the final value that will be used for the calculation. Each bin contains two numbers, σ_L and σ_T which stand for the longitudinal and transverse cross section for $\gamma^* p \to \pi^0 p'$. In order to get the resulting cross section they are added up with the longitudinally polarized virtual photon cross section ε:

$$\sigma_{\gamma * p} = \sigma_T + \varepsilon \sigma_L. \tag{5.44}$$

Finally, to get the muon proton cross section again, the factor for the transverse virtual photon flux Γ_T is multiplied with it:

$$\sigma_{\mu p} = \sigma_{\gamma * p} \Gamma_T. \tag{5.45}$$

The transverse virtual photon flux can be parametrized in the COMPASS kinematical range as:

$$\Gamma_T = \frac{\alpha(\nu - \frac{Q^2}{2m_p})}{2\pi Q^2 E^2 y^2} \left[y^2(1 - 2m_{\frac{\mu^2}{Q^2}}) + (1 - y - 0.25\gamma^2 y^2) \frac{2}{1+\gamma^2} \right], \tag{5.46}$$

and the factor $\gamma^2 = \frac{Q^2}{\nu^2}$.

The tables each have different ranges, which are listed in Tab. 5.5. The cross sections

Table 5.5: Table options for table based generators.

table	Q^2 / [GeV/c]^2	W / [GeV/c]	t' / [GeV/c]^2
π^0 without $\bar{E}_T$	9 bins $\in [2, 10]$	11 bins $\in [5, 15]$	16 bins $\in [0.00, 0.75]$
π^0 with $\bar{E}_T$	15 bins $\in [2, 16]$	11 bins $\in [5, 15]$	16 bins $\in [0.00, 0.75]$
ρ^+	19 bins $\in [2, 20]$	12 bins $\in [5, 16]$	16 bins $\in [0.00, 0.75]$

resulting from these tables and extrapolation is shown in the Fig. 5.9 up to Fig. 5.12. It can also be observed from the figures that the difference is almost a factor of 10, which is a surprising result. The transversity GPDs were, in fact, introduced to cope with the problem of the experimental cross sections for pion production being much larger than expected. Another interesting feature which the plots show, is the t' dependence. While for the model without the GPD $\bar{E}_T$ a parametrization could have been found, the cross section shows a distinct non-trivial behavior for fixed Q^2 with its influence.

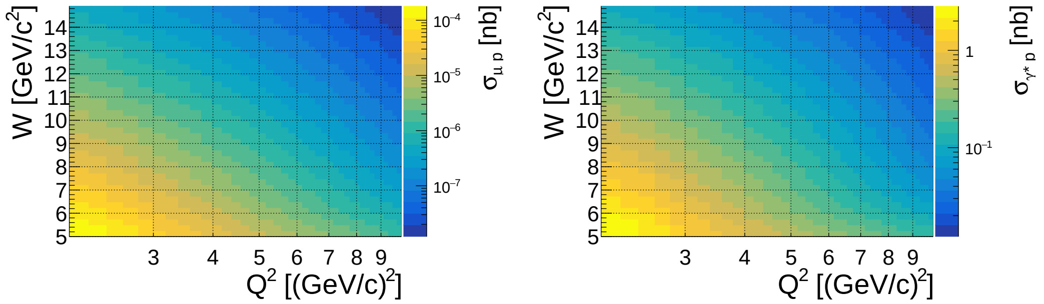


Figure 5.9: Cross sections for pion production without the GPD $\bar{E}_T$ contribution, at a fixed $t' = -0.5\,\mathrm{GeV/c}$.

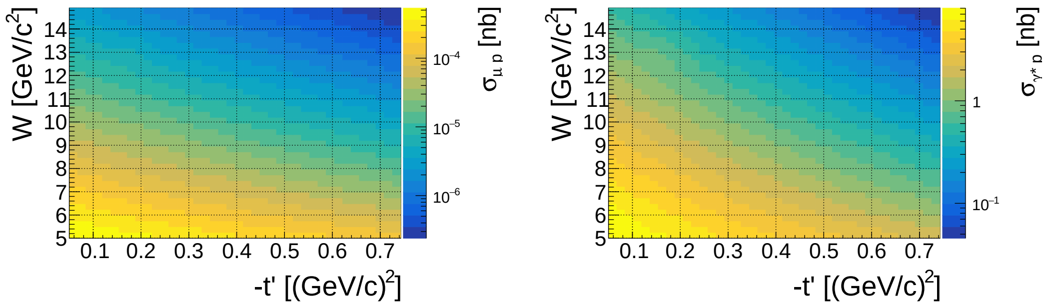


Figure 5.10: Cross sections for pion production without GPD $\bar{E}_T$ contribution at a fixed $Q^2 = 2.0\,\mathrm{GeV/c^2}$.

5.3.4 Exclusive Rho$^+$ Mesons: $\rho^+ \to \pi^+ + \pi^0(\to \gamma\gamma)$

This generator also uses the **AUX** flags for loading the cross section tables. The usage is generally identical to the neutral pion generator's **AUX** flags listed in Tab. 5.4.

Mass Distribution

The mass for the ρ^+ meson is generated the same way as the ρ^0. This is perfectly justified by the PDG giving the mass difference of both mesons to [23]:

$$\Delta m_{\rho^{0,\pm}} = (-0.7 \pm 0.8)\,\mathrm{MeV}. \tag{5.47}$$

Polarization

The polarization is generated in an identical manner as the rho generator.

Decay

The π^0 decays immediately into two photons in the final state. As this is a simple two-body decay, the calculations are simply done in the pions' centre of mass system. The resulting photons are then boosted along the pions' four-vector. The angular distribution is flat in the azimutal angle ϕ, for the polar angle θ a uniform $r \in [0, 1]$ is drawn. Then θ is calculated according to:

$$\theta = \arccos(2r - 1). \tag{5.48}$$

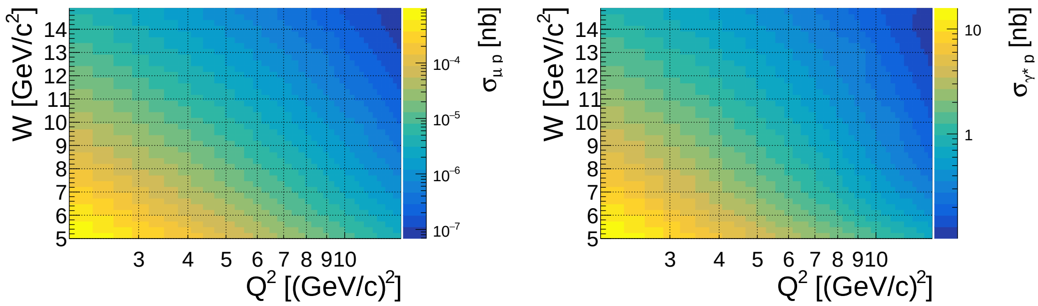


Figure 5.11: Cross sections for pion production with GPD $\bar{E}_T$ contribution at a fixed $t' = -0.5\,\mathrm{GeV/c}$.

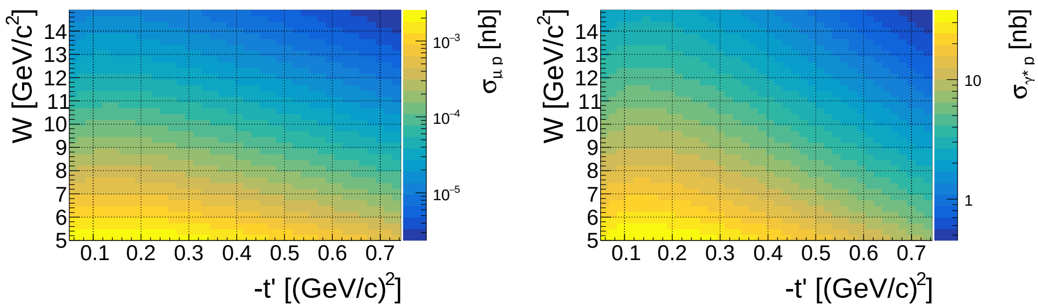


Figure 5.12: Cross sections for pion with GPD $\bar{E}_T$ contribution production at a fixed $Q^2 = 2.0\,\mathrm{GeV/c^2}$.

Cross Section

The cross section for this generator also comes from a data table by Goloskokov and Kroll. The ranges and binning of the table is listed in the Tab. 5.5. The cross sections are shown in the figures 5.13a to 5.13d.

5.3.5 Exclusive Omega Mesons: $\omega \to \pi^+\pi^-\pi^0(\to \gamma\gamma)$

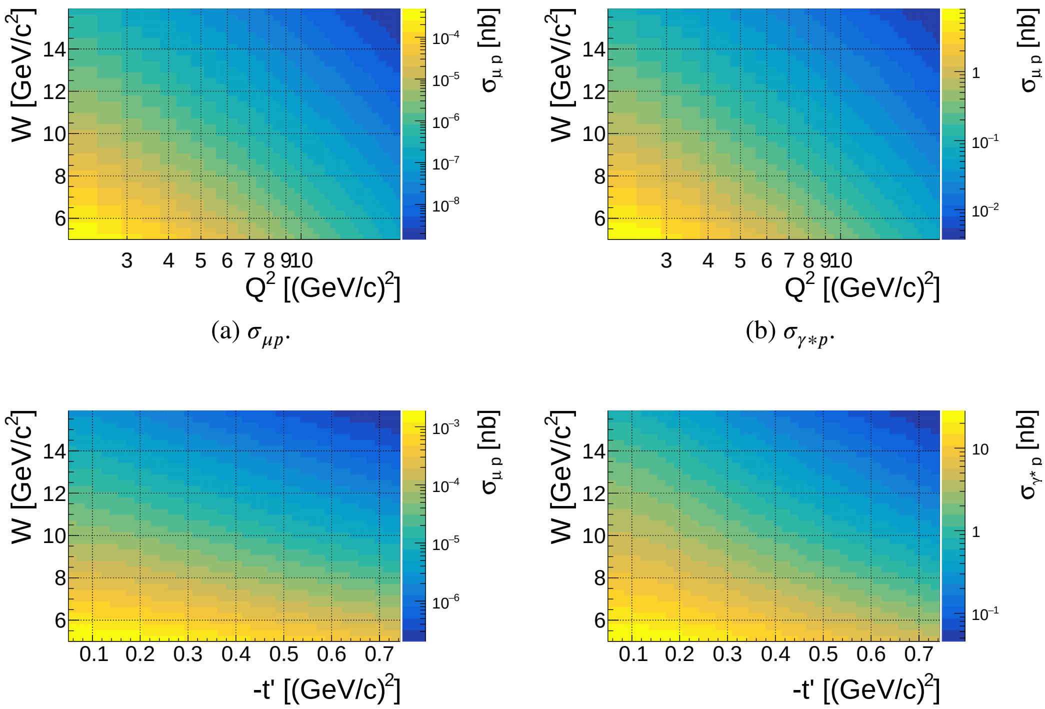
The omega generator with the decay into three pions is of importance for the acceptance estimation when looking for exactly this channel. This channel is the dominant decay mode of the meson, yet neither the exact cross section nor the angular distributions of the pions have ever been measured to a sufficient accuracy. According to theory, the pions show a special behaviour in the Dalitz plot, which this generator tries to model. The result is shown in Fig. 5.15.

Mass Distribution

The mass distribution follows a relativistic Breit-Wigner distribution like the ρ meson. It follows equation 5.30, but with the constants $g_0 = 0.00849\,\mathrm{GeV}$, $q_0 = 0.327\,\mathrm{GeV}$ and $m_\omega = 0.78265\,\mathrm{GeV/c^2}$ being the mean mass of the meson. The resulting distributions are shown in Fig. 5.14.

Polarization

Since no parameters are available for the polarization of the omega meson and the mass is near the rho mass, the parameters used for the rho meson are also used for the omega meson.



(a) $\sigma_{\mu p}$.

(b) $\sigma_{\gamma * p}$.

(c) $\sigma_{\mu p}$.

(d) $\sigma_{\gamma * p}$.

Figure 5.13: Cross sections for ρ^+ production with a fixed $t' = -0.5\,\mathrm{GeV/c}$ for the Q^2/W plots and $Q^2 = 2.0\,\mathrm{GeV/c^2}$ for the t'/W plots.

Decay

The decay in this channel works, by first randomly selecting one of the decay products and letting the omega decay into this product and a resulting "dipion". The highest possible energy and momentum is then given by:

$$E_{\max} = \frac{m_\omega^2 + m_1^2 - m_{\mathrm{dp}}^2}{2m_\omega}, \tag{5.49}$$

$$|\vec{p}|_{\max} = \sqrt{E_{\max}^2 - m_1^2}, \tag{5.50}$$

where m_1 is the mass of the selected "first" decay particle and $m_d p = m_2 + m_3$ is the mass of the dipion. Next, a random fraction of $|\vec{p}|$ is assigned to the decay particle 1:

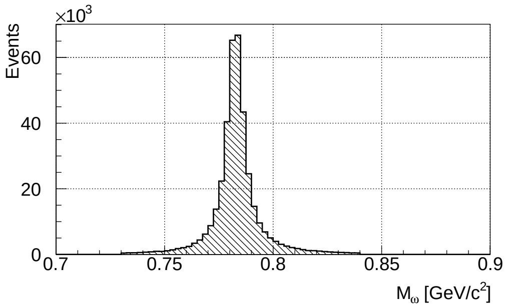
$$\vec{p}_1 = (r \cdot |\vec{p}|_{\max}, 0, 0). \tag{5.51}$$

Also, a polar angle between decay particle 1 and the dipion needs to be drawn:

$$\theta_{pq} = \arccos(2r' - 1). \tag{5.52}$$

Now, the kinematics for the dipion is fixed to:

$$E_{dp} = m_\omega - \sqrt{m_1^2 + |\vec{p}_1|}, \tag{5.53}$$



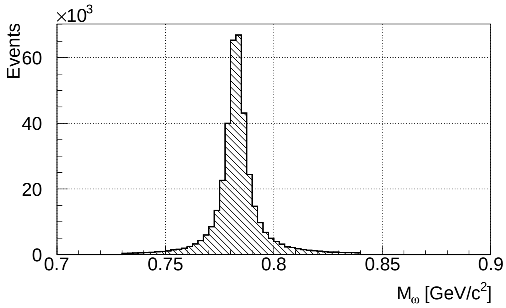


Figure 5.14: The mass distribution of the generated ω mesons. On the left: For the channel $\omega \rightarrow \pi^0 + \gamma$, on the right: For the channel $\omega \rightarrow \pi^+\pi^-\pi^0$.

$$m_{dp} = \sqrt{E_{dp}^2 - |\vec{p}_1|}, \tag{5.54}$$

$$q_{dp} = \frac{\sqrt{(m_{dp}^2 - m_{dp,\min}^2)[m_{dp}^2 - (m_2 - m_3)^2]}}{2m_{dp}} \tag{5.55}$$

With the accept-reject method, the distribution of q_{dp} is shaped according to:

$$p(q_{dp}) = 9|\vec{p}_1|^2 q_{dp}^2 (1 - \cos(\theta_{pq})^2). \tag{5.56}$$

After fulfilling this condition all of the decay particles' four momenta are set:

$$p_2 = (\sqrt{q_{dp}^2 - m_2^2}, q_{dp}\ \cos(\theta_{pq}), q_{dp} \sin(\theta_{pq}), 0), \tag{5.57}$$

$$p_3 = (\sqrt{q_{dp}^2 - m_3^2}, q_{dp}\ \cos(\theta_{pq}), q_{dp} \sin(\theta_{pq}), 0). \tag{5.58}$$

Now, a Lorentz boost is applied, to bring the second and third decay particle into the center of mass frame of the omega. The complete set of three decay particles are then rotated around the axis of the omega vector by a random angle before finally being boosted into the laboratory frame. This decay process produces the Dalitz plot shown in Fig. 5.15.

Cross Section

The shape of the cross section is assumed to be identical to that of the rho meson, with only the absolute value being smaller by a factor of 10. Therefore, the rho cross section is calculated and then divided by 10 to get the omega cross section.

New decay channel into $\pi^0 + \gamma$

The omega decay channel $\omega \rightarrow \gamma + \pi^0(\rightarrow \gamma\gamma)$ is a recent addition to the already existing omega generator. It features the decay channel into three photons, via a neutral pion. The cross section identical to the other omega decay channel is chosen. This generator was implemented on short notice to allow for background studies of this omega decay channel's contribution into the exclusive neutral pion production channel.

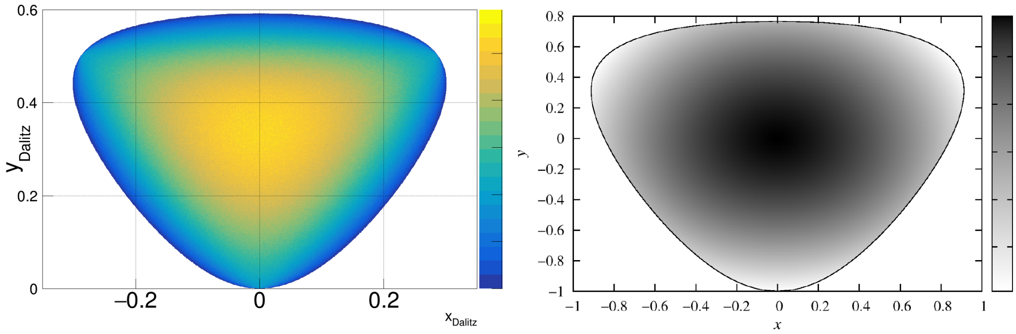


Figure 5.15: The Dalitz plot of the decay products from the ω generator (left) and a theoretical estimation from [81] (right).

Mass Distribution

The mass is drawn almost identically to the other decay channel. The only difference is the width: $q_0 = 0.380\,\mathrm{GeV}$, which was chosen for compatibility reasons. The resulting distribution is drawn next to the one representing the other decay channel for easier comparison. It can be seen in Fig. 5.14.

Polarization

From the 2012 data analysis of the COMPASS-II pilot run, Marianski determined the ratio between longitudinally and transversely polarized mesons to $R = 0.485$ [82].

Decay

This decay channel is easily calculable since the decay features two consecutive two-body decays which can both be calculated independently.

5.3.6 Exclusive J/ψ: $J/\psi \rightarrow e^+e^-$ or $\mu^+\mu^-$

For future experiments a generator for exclusive production of J/ψ is included in HEPGen++. The decay products are switchable via the **AUX1** flag. Setting it to "1" makes the generator produce only the electron-positron, while a setting of "2" only generates the decay into muons and anti-muons. Not setting the variables, however, will result in the generation of both decays with a branching ratio of 50%.

Cross Section

The cross section was taken from the ZEUS measurements [83]. The parametrization is written as:

$$\sigma_{\gamma * p} = \sigma_0 \left(\frac{W}{W_0}\right)^{\delta} \left(\frac{ams_{\psi}}{Q^2 + ams_{\psi}}\right)^{\beta} b_t \exp(-b_t t'), \tag{5.59}$$

where the parameters are taken from [83] to $\sigma_0 = 77\,\mathrm{nb}$, $\delta = 0.73$, $\beta = 2.44$, $ams_{\psi} = 9.59079\,(\mathrm{GeV/c})^2$ and $W_0 = 90\,\mathrm{GeV/c^2}$. The cross section is plotted in Fig. 5.16.

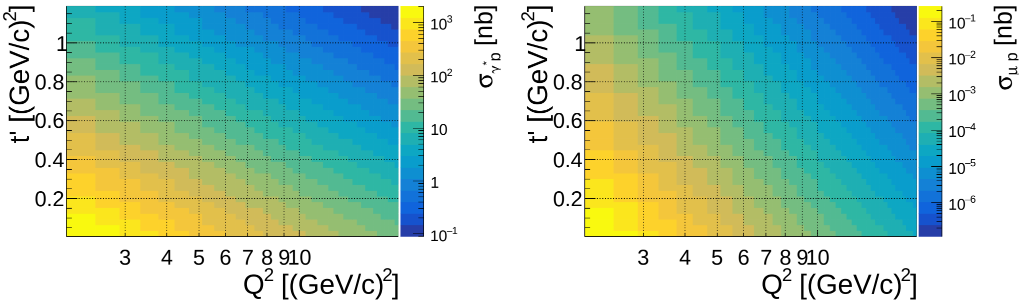


Figure 5.16: The cross sections for exclusive J/ψ production in HEPGen++.

Mass

The mass is sharply set to the resonance mass of $m_{J/\psi} = 3.0969\,\mathrm{GeV/c^2}$.

Polarization and Decay

The ratio of the longitudinally to the transversly polarized cross section is a simple Q^2 sloped distribution:

$$R_V = 0.07\,Q^2. \tag{5.60}$$

The θ distribution, the polar angle of the decay particles in the center of mass frame of the J/ψ, is dependant on the polarization. In the case of longitudinally polarized mesons the distributions for the angle θ looks like:

$$r \in [0, 1], \tag{5.61}$$

$$\theta' = \arccos(|2r - 1|), \tag{5.62}$$

$$\cos(\theta) = \begin{cases} -2\cos\left(\frac{\theta'+\pi}{3}\right) & \text{for} \quad r < 0.5 \\ +2\cos\left(\frac{\theta'+\pi}{3}\right) & \text{else.} \end{cases} \tag{5.63}$$

On the other hand, when it comes to the transversely polarized mesons, the transversely polarized ones, the distribution is given by:

$$\begin{aligned} & r \in [0, 1], \\ \cos(\theta) = & \frac{-2.0 + 4.0\,r + \sqrt{16.0\,r^2 - 16.0\,r + 5.0}}{\left(\left|-2.0 + 4.0\,r + \sqrt{16.0\,r^2 - 16.0\,r + 5.0}\right|\right)^{2/3}} \\ & + \frac{-2.0 + 4.0\,r - \sqrt{16.0\,r^2 - 16.0\,r + 5.0}}{\left(\left|-2.0 + 4.0\,r - \sqrt{16.0\,r^2 - 16.0\,r + 5.0}\right|\right)^{2/3}} \end{aligned} \tag{5.64}$$

5.3.7 Exclusive Photon Generation: FFS

The generator for exclusive photons combines the DVCS model of FFS [71, 84] that has been modified for COMPASS-II by A. Sandacz [70] and is combined with the Bethe-Heitler calculations from the BMK paper [32] where the propagators were recalculated by P.A.M. Guichon to include the lepton mass.

DVCS Cross Section

The implementation of the DVCS cross sections follows roughly the FFS model described in [84] with some modifications to the GPDs. Here B_0, x_0, α_p are the Regge parameters for the t-slope in the parametrization form:

$$B(x_{bj}) = B_0 + 2\alpha_p \ln\left(\frac{x_0}{x_{bj}}\right). \tag{5.65}$$

The imaginary part of the GPD H, which in this model is the only contributing GPD for DVCS, is parametrized following [70]:

$$\Im(H) = \frac{\pi}{x_{bj}} \frac{F_2(x_{bj}, Q^2)}{R} \exp\left(-\frac{B(x_{bj})}{2}|t|\right) \tag{5.66}$$

This assumes the so-called D-term to be 0 which breaks the polynomiality of the GPDs and is problematic. But, as this simplified model only serves to get an estimation of the DVCS content, it is well justified to do so. The skewness ratio R is assumed to be:

$$R = \frac{\Im(A(\gamma^* p \to \gamma^* p)_{t=0})}{\Im(A(\gamma^* p \to \gamma p)_{t=0})} = 0.5. \tag{5.67}$$

The ratio between the imaginary and the real part of the amplitude η is calculated. In the paper it is given to:

$$\eta = \frac{\pi}{2} \frac{\mathrm{d}\ln(F_2(x_{bj}, Q^2))}{\mathrm{d}\ln(1/x_{bj})}. \tag{5.68}$$

When using the derivative dispersive relation instead, it can be written as [70]:

$$\eta = \frac{\Re(H)}{\Im(H)} \approx \frac{\pi}{2} \frac{\ln[\Delta F_2(x_{bj}, Q^2)] + \Delta B \cdot \frac{t}{2}}{\ln(1/(\Delta x_{bj}))}, \tag{5.69}$$

where the discrete differences are [70]:

$$\Delta x_{bj} = \frac{.99\, x_{bj}}{1.01\, x_{bj}}, \tag{5.70}$$

$$\Delta F_2 = \frac{F_2(1.01\, x_{bj}, Q^2)}{F_2(0.99\, x_{bj}, Q^2)}, \tag{5.71}$$

$$\Delta B = 2\alpha_p \left[\ln\left(\frac{x_0}{1.01\, x_{bj}}\right) - \ln\left(\frac{x_0}{.99\, x_{bj}}\right)\right]. \tag{5.72}$$

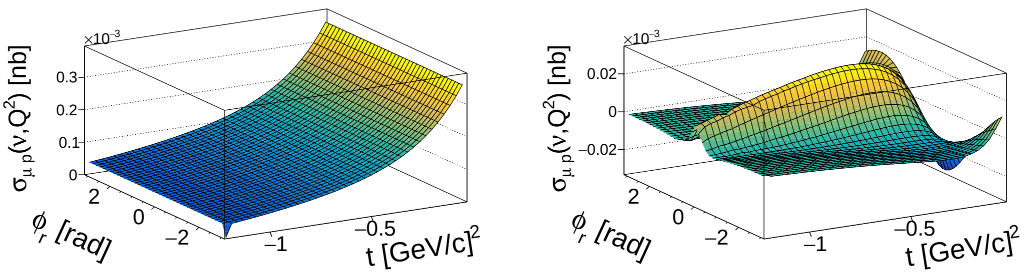
With η, the real part of H becomes simply:

$$\Re(H) = \eta \Im(H). \tag{5.73}$$

The DVCS amplitude squared can then be written as:

$$A^2 = \frac{2(2 - 2y + y^2)}{y^2(2 - x_{bj})^2 Q^2} 4(1 - x_{bj})(\Re(H)^2 + \Im(H)^2). \tag{5.74}$$

Afterwards, the cross section is calculated from this amplitude by multiplying it with a phase factor and transforming the resulting cross section from $\frac{\mathrm{d}^2\sigma}{\mathrm{d}x_{bj}\mathrm{Q}^2}$ to $\frac{\mathrm{d}^2\sigma}{\mathrm{d}\nu\mathrm{Q}^2}$. The DVCS cross section is shown in Fig. 5.17a.



(a) FFS-DVCS cross section.

(b) FFS-Interference cross section.

Figure 5.17: FFS model cross section calculations for the kinematics: $Q^2 = 3\,(\mathrm{GeV/c})^2$ and $\nu = 10\,\mathrm{GeV}$.

Interference Term

The formalism to calculate the interference amplitude is taken from [32]. The $\phi_r = \phi$ angle is the angle between the leptonic plane, spanned by k and k', and the photonic plane spanned by γ^* and p_γ. Taking the interference term directly from the paper and evaluating only to twist-2 leads to the amplitude:

$$\mathcal{I} = \frac{\pm e^6}{x_{bj} y^3 \Delta^2 \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\mathcal{I}} + c_1^{\mathcal{I}} \cos(\phi) + s_1^{\mathcal{I}} \sin(\phi) \right\}, \tag{5.75}$$

where the Fourier coefficients are:

$$c_0^{\mathcal{I}} = -8(2-y)\Re\left\{ \frac{(2-y)^2}{1-y} K^2 C_{\mathrm{unp}}^{\mathcal{I}}(\mathcal{F}) + \frac{\Delta^2}{Q^2}(1-y)(2-x_{bj})(C_{\mathrm{unp}}^{\mathcal{I}} + \Delta C_{\mathrm{unp}}^{\mathcal{I}})(\mathcal{F}) \right\}, \tag{5.76}$$

$$c_1^{\mathcal{I}} = -8K(2-2y+y^2)\Re(C_{\mathrm{unp}}^{\mathcal{I}}(\mathcal{F})), \tag{5.77}$$

$$s_1^{\mathcal{I}} = 8K\lambda y(2-y)\Im(C_{\mathrm{unp}}^{\mathcal{I}}(\mathcal{F})), \tag{5.78}$$

and the kinematical factor K is:

$$K^2 = -\frac{\Delta^2}{Q^2}(1-x_{bj})\left(1-y-\frac{y^2\varepsilon^2}{4}\right)\left(1-\frac{\Delta^2_{\mathrm{min}}}{\Delta^2}\right) \left\{ \sqrt{1+\varepsilon^2} + \frac{4x_{bj}(1-x_{bj})+\varepsilon^2}{4(1-x_{bj})}\frac{\Delta^2-\Delta^2_{\mathrm{min}}}{Q^2} \right\}. \tag{5.79}$$

Like in the DVCS case, the amplitude is converted to a differential cross section in ν and Q^2. The interference cross section is plotted in Fig. 5.17b.

Bethe-Heitler Cross Section

The Bethe-Heitler amplitude is also taken from the BMK formalism [32] to:

$$\mathcal{T}^2_{\mathrm{BH}} = \frac{e^6}{x_{bj}^2 y^2 (1+\varepsilon^2)^2 \Delta^2 \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \left\{ c_0^{\mathrm{BH}} + c_1^{\mathrm{BH}} \cos(\phi) + c_2^{\mathrm{BH}} \cos(2\phi) \right\}, \tag{5.80}$$

with the Fourier coefficients:

$$
\begin{aligned}
c_{0,\text{unp}}^{\text{BH}} = 8K^2 \Bigg\{ & (2+3\varepsilon^2)\frac{Q^2}{\Delta^2}\left(F_1^2 - \frac{\Delta^2}{4m_p^2}F_2^2\right) + 2x_{bj}^2(F_1+F_2)^2 \Bigg\} \\
& + (2-y)^2 \Bigg\{ (2+\varepsilon^2)\left[\frac{4x_{bj}^2 m_p^2}{\Delta^2}\left(1+\frac{\Delta^2}{Q^2}\right)^2 + 4(1-x_{bj})\left(1+x_{bj}\frac{\Delta^2}{Q^2}\right)\right] \\
& \qquad \cdot \left(F_1^2 - \frac{\Delta^2}{4m_p^2}F_2^2\right) \\
& + 4x_{bj}^2\left[x_{bj} + \left(1 - x_{bj} + \frac{\varepsilon^2}{2}\right)\left(1-\frac{\Delta^2}{Q^2}\right)^2 - x_{bj}(1-2x_{bj})\frac{\Delta^4}{Q^4}\right](F_1+F_2)^2 \Bigg\} \\
& + 8(1+\varepsilon^2)\left(1 - y - \frac{\varepsilon^2 y^2}{4}\right)\Bigg\{ 2\varepsilon^2\left(1-\frac{\Delta^2}{4m_p^2}\right)\left(F_1^2 - \frac{\Delta^2}{4m_p^2}F_2^2\right) \\
& \qquad - x_{bj}^2\left(1-\frac{\Delta^2}{Q^2}\right)^2 (F_1+F_2)^2 \Bigg\},
\end{aligned}
\tag{5.81}
$$

$$
\begin{aligned}
c_{1,\text{unp}}^{\text{BH}} = 8K(2-y)\Bigg\{ & \left(\frac{4x_{bj}^2 m_p^2}{\Delta^2} - 2x_{bj} - \varepsilon^2\right)\left(F_1^2 - \frac{\Delta^2}{4m_p^2}F_2\right) \\
& + 2x_{bj}^2\left(1-(1-2x_{bj})\frac{\Delta^2}{Q^2}\right)(F_1+F_2)^2 \Bigg\},
\end{aligned}
\tag{5.82}
$$

$$
c_{2,\text{unp}}^{\text{BH}} = 8x_{bj}^2 K^2 \left\{ \frac{4m_p^2}{\Delta^2}\left(F_1^2 - \frac{\Delta^2}{4m_p^2}F_2^2\right) + 2(F_1+F_2)^2 \right\}. \tag{5.83}
$$

Muon Propagators

The muon propagator terms $\mathcal{P}_1(\phi)$ and $\mathcal{P}_2(\phi)$ are set differently in HEPGen++ than in the BMK paper. There, the following parametrization is used:

$$
J = \left(1 - y - \frac{y\varepsilon^2}{2}\right)\left(1+\frac{\Delta^2}{Q^2}\right) - (1-x_{bj})(2-y)\frac{\Delta^2}{Q^2}, \tag{5.84}
$$

$$
\mathcal{P}_1(\phi) = -\frac{1}{y(1+\varepsilon^2)}\{J + 2K\cos(\phi)\}, \tag{5.85}
$$

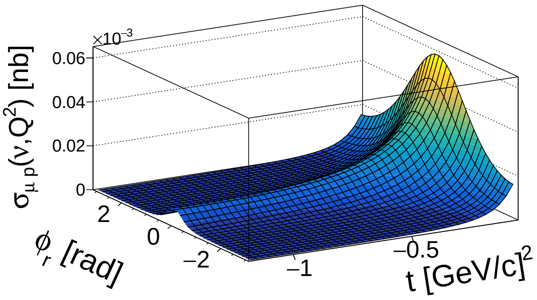
$$
\mathcal{P}_2(\phi) = 1 + \frac{\Delta^2}{Q^2} + \frac{1}{y(1+\varepsilon^2)}\{J + 2K\cos(\phi)\}, \tag{5.86}
$$

while in HEPGen++ the propagators are derived from the momentum four-vectors in the laboratory system:

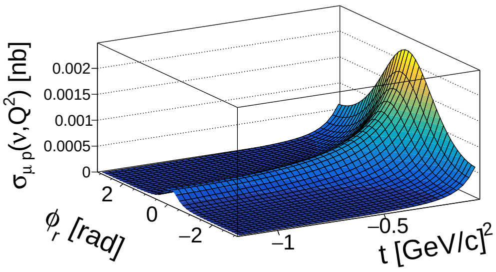
$$
\mathcal{P}_1(\phi) = \frac{(P_{\mu'} + P_\gamma)^2 - m_\mu^2}{Q^2}, \tag{5.87}
$$

$$
\mathcal{P}_2(\phi) = \frac{(P_\mu - P_\gamma)^2 - m_\mu^2}{Q^2}. \tag{5.88}
$$

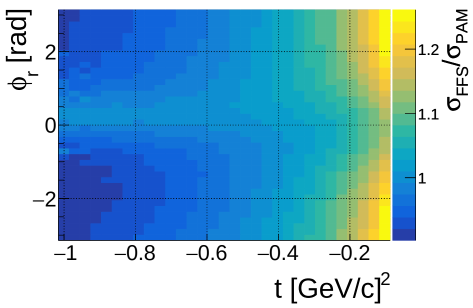
The propagator terms from BMK are approximations for massless leptons, like electrons. For COMPASS, which is using muons, it is better to use the four-vector notation. This can be shown by comparing different Bethe-Heitler cross section equations in the Fig. 5.18c to Fig. 5.18a. In the first figure, Fig. 5.18c, the four-vector propagators are used while the rest of the calculations remains the one for electrons. One can see that this approximation is very good over a wide kinematic range, as it fits the reference by P.A.M. Guichon in Fig. 5.18a reasonably well. Using the BMK approximation with the parametrization for the propagator terms, shown in Fig. 5.18b, results in a rather large discrepancy at smaller t. The reason for using the modified BMK model for different propagators instead of the an-



(a) Analytically correct version by P.A.M. Guichon.



(b) BMK using parametrization for propagators.



(c) BMK using four-vector propagators divided by the P.A.M. Guichon code.

Figure 5.18: Different Bethe-Heitler calculations at the kinematics: $Q^2 = 3\,(\mathrm{GeV/c})^2$ and $\nu = 10\,\mathrm{GeV}$. The modified Bethe-Heitler code fits the analytically correct Bethe-Heitler well.

alytically correct P.A.M. Guichon calculation is that with the BMK terms the interference term can be calculated, which cannot be done with the analytical formulation.

5.3.8 Exclusive Photons: VGG Model

The generator for exclusive photon production can also be used with the implementation of the VGG model [85, 86] by Laurent Mosse [75]. This model was used for the estimation of the accuracy of the measurement that COMPASS-II could achieve at the time of the proposal writing. It is much more sophisticated than the simple FFS model. However, it should be noted, that the Bethe-Heitler implementation is also not perfect in this implementation. For the purposes of this thesis, the Mosse code was refurbished, packaged into a single library and an interface to the HEPGen++ framework.

5.3.8.1 GPDs

The GPDs used for the VGG model, are double distribution types in the formalism of Radyushkin [87].

GPD H

The GPD H is modeled to [75, 88]:

$$H^q_{DD}(x,\xi,t) = \int_{-1}^{1} \mathrm{d}\beta \int_{-1+|\beta|}^{1-|\beta|} \mathrm{d}\alpha\ \delta(x-\beta-\alpha\xi) F^q(\beta,\alpha,t), \tag{5.89}$$

where the t dependence is factorized out in the shape of [75, 88]:

$$F^q(\beta,\alpha,t) = h(\beta,\alpha) q(\beta) f^q(\beta,t). \tag{5.90}$$

The α and β are typical notations for integration variables in double distribution ansatz models [88]. The PDFs $q(\beta)$ in the equation above were taken from MRST98 at $Q^2 = 1\,(\mathrm{GeV/c})^2$ [89] and are evolved using a parametrized evolution approach, with a Reggeized pole [75]. The second part of the t dependence term is the so-called profile function, which was modeled as:

$$h(\beta,\alpha) = \frac{\Gamma(2b+1)}{2^{2b+1}\Gamma^2(b+1)} \frac{[(1-|\beta|)^2)^2 - \alpha^2]^b}{(1-|\beta|)^{2b+1}}. \tag{5.91}$$

The profile is dependent on parameter b which is typically in the order of 1.

GPD E

The GPD E is also modeled as a double distribution type ansatz. Contrary to the GPD H, this distribution has to fulfill the forward limit of F_2. It can be written as [75]:

$$E^q(\beta,\alpha,t) = h(\beta,\alpha) e_q(\beta) f_e(t), \tag{5.92}$$

Where h is the same profile function, which was already used for GPD H, $f_e(t)$ is the t dependence of the nucleon form factor $f_t(t) = (1 - t/(0.71\,\mathrm{GeV}^2))^{-2}$ [75]. To make E fulfill the forward limit and Ji's sum rule, the e_q term is chosen to:

$$e_b(\beta) = A^q q_v(\beta) + B^q \delta(\beta), \tag{5.93}$$

with the parameters given to [75]:

$$A^q = \frac{2J^q - M^q}{M^{q_v}}, \tag{5.94}$$

$$B^q = \kappa^q - N_q A^q. \tag{5.95}$$

In these equations, κ^q is the anomalous magnetic moment of the quarks with flavour q, N_q the number of valence quarks with flavour q in the nucleon. J^q is the total angular momentum fraction carried by quark q, it is a free parameter. M^q and M^{q_v} are the momentum fractions of all (valence) quarks of flavour q defined by [75]:

$$\begin{aligned} M^q &= \int_0^1 \mathrm{d}x\ x\ ((q(x) + \bar{q}(x)), \\ M^{q_v} &= \int_0^1 \mathrm{d}x\ x\ q_v(x). \end{aligned} \tag{5.96}$$

GPD $\tilde{H}$

The GPD $\tilde{H}$ is modeled according to the double distribution ansatz [75]:

$$\tilde{H}(\beta,\alpha,t) = h(\beta,\alpha)\Delta q(\beta) g_A^q(t)/g_A^q(0), \tag{5.97}$$

where g_A^q is the axial form factor of flavour q. The helicity distribution $\Delta q(x)$ is taken from [90].

GPD $\tilde{E}$

Modeling the GPD $\tilde{E}$ is a bit more complicated, since it contains a pion pole. The pion pole means, that the parton from the target proton forms a quark-antiquark-triangle loop with any pion inside. This behaviour is especially strong at $\Delta \approx m_\pi$ but contributes to this GPD in general [91]. The t dependence of the pion pole can be written as [75]:

$$h_A^\pi(t) = -g_A \frac{4m_p^2}{t - m_\pi^2 c^2}. \tag{5.98}$$

With this, the GPD $\tilde{E}$ can be written in double distribution ansatz form as [75]:

$$\tilde{E}^\pi(x,\xi,t) = \theta(\xi - |x|) h_A^\pi(t) \frac{1}{\xi} \phi_{as}\left(\frac{x}{\xi}\right), \tag{5.99}$$

where $\phi_a s(z) = 3/4(1 - z^2)$ is the pion distribution function with the longitudinal momentum fraction z of the quark. With $z = x/\xi$ the factor $1/\xi$ needs to be introduced for normalization reasons.

5.3.8.2 Convolutions and DVCS Amplitude

The GPDs then need to be convoluted with hard scattering kernels and integrated over x to get access to the Compton form factors, which are needed for observable calculation. These scattering kernels are of the form:

$$C^\pm(x,\xi) = \left[\frac{1}{x - \xi + i\varepsilon} \pm \frac{1}{x + \xi - i\varepsilon}\right]. \tag{5.100}$$

Therefore, they must be integrated using Cauchy principal value [75]:

$$\lim_{\varepsilon\to 0} \int_0^1 \mathrm{d}x \frac{f(x,\xi)}{x - \xi + i\varepsilon} = P\int_0^1 \mathrm{d}x \frac{f(x,\xi)}{x-\xi} - i\pi f(\xi,\xi), \tag{5.101}$$

where [75]:

$$P\int_0^1 \mathrm{d}x \frac{f(x,\xi)}{x-\xi} = \int_0^1 \mathrm{d}x \frac{f(x,\xi) - f(\xi,\xi)}{x-\xi} + f(\xi,\xi)\ln\left(\frac{\xi}{1-\xi}\right). \tag{5.102}$$

The combinations of the CFFs lead to the helicity amplitudes, which can then be used to extract any meaningful quantity, like the DVCS amplitude:

$$T^{DVCS} = -\frac{e_l}{Q^2} \sum_{h=-1}^{1} C^h_{h_l',h_l}(\varepsilon,\theta,\phi) M_{h',h_p',h,h_p}, \tag{5.103}$$

where the $C^h_{h_l',h_l}$ are the decomposition coefficients given in [75], h is the virtual photon helicity, h' is the real photon helicity, h_p' is the nucleon helicity after scattering, h_p is the nucleon helicity in the initial state and h_l, h_l' are the initial and final state lepton helicities. The angle ϕ is defined as the azimutal angle between the leptonic and the production plane, θ is defined as the angle between the initial state proton and the real photon, ε is the virtual photon polarization vector. The full DVCS cross section is shown in Fig. 5.19c.

5.3.8.3 Bethe-Heitler Amplitude

The Bethe-Heitler implementation in this model is calculated like [75]:

$$T^{BH} = |e|e_l^2\varepsilon'^{*}_{\mu}(k')\bar{u}(k')\left[\gamma^{\mu}\frac{\gamma(k'+q')+m_l}{(k'+q')^2-m_l^2}\gamma^{\nu}+\gamma^{\nu}\frac{\gamma(k-q')+m_l}{(k-q')^2-m_l^2}\right]u(k)\times\cdots$$
$$\frac{1}{(p'-p)^2}\bar{u}(p')\Gamma^{em}_{\nu}(p',p)u(p). \tag{5.104}$$

In this equation the γ are the Dirac gamma matrices, $u, \bar{u}$ are Dirac Spinors, $e_l = \pm|e_l| = \sqrt{4\pi/137}$ is the lepton charge, q' is the four-momentum of the real photon and the electromagnetic current:

$$\langle p'|J^{em}_{\mu}(0)|p\rangle = |e|\bar{u}(p')\Gamma^{em}_{\mu}(p',p)u(p), \tag{5.105}$$

where the hadronic tensor is [75]:

$$\Gamma^{em}_{\mu}(p,p') = F_1^{p,n}(t)\gamma_{\mu} + iF_2^{p,n}(t)\frac{\sigma_{\mu\nu}(p'-p)^{\nu}}{2m_p}. \tag{5.106}$$

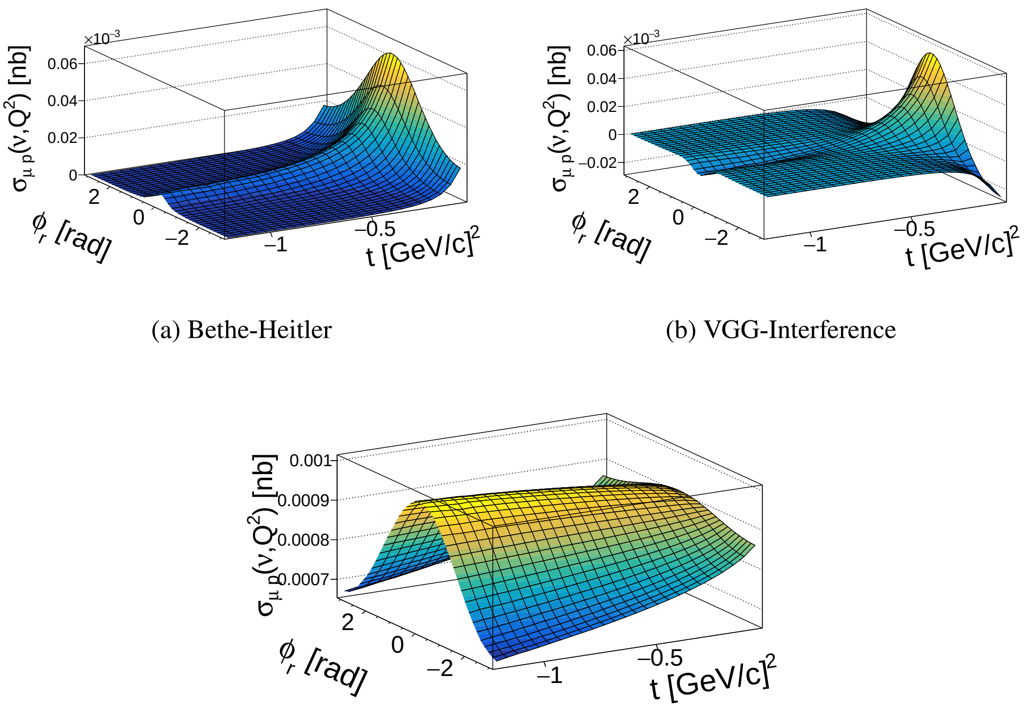
In this notation, the poles of the Bethe-Heitler cross section are clearly visible. The so-called s- and p-peak resulting from the real photon being emitted in the direction of the incoming lepton ($\sim \frac{1}{(k-q')^2-m_l^2}$) or being emitted in the opposite direction of the scattered lepton ($\sim \frac{1}{(k'+q')^2-m_l^2}$). The interference amplitude is obtained in the same way as already described for the FFS model. The Bethe-Heitler cross section in shown in Fig. 5.19a, the interference cross section in 5.19b.

5.3.9 Exclusive Photons: GK11 Model

The GK11 model was numerically implemented in C++ by H. Moutarde [69]. While the complete framework is not available to public yet, an early version was used in order to compare it with the other models. The GK11 model, is introduced in Chapter 6, since it is implemented for π^0 production into HEPGen++. The code of H. Moutarde was equipped with an interface to be used with HEPGen++. Plots of the cross sections for comparison are shown in the Fig. 5.20. The model can be used in leading order approximation or next to leading order calculation. The latter, however, takes much more time to calculate. For the usage in the generator, it is very important to keep in mind that the Bethe-Heitler contribution is done for electrons, neglecting the lepton mass completely. Therefore, the calculation becomes wrong in the kinematics coming close to the so-called s- and p-peaks, the regions of singularities in the Bethe-Heitler propagator.

5.3.10 Diffractive Dissociation

The model for the diffractive dissociation of the target nucleon is taken from [92]. With the appropriate flag set in the data card this becomes enabled in the generation, tainting the exclusive sample with non-exclusive background. This feature becomes especially useful when analysing data without a recoil detector for proper proton identification to account for possible events that are wrongly identified as exclusive. With a recoil detector in place, it allows for an estimation of it's proton identification capabilities under better simulation conditions.



(a) Bethe-Heitler

(b) VGG-Interference

(c) VGG-DVCS

Figure 5.19: Calculations for the cross sections resulting from the VGG model following the implementation by Laurent Mosse. The kinematics are $Q^2 = 3.0\,(\mathrm{GeV/c})^2$ and $\nu = 10\,\mathrm{GeV}$.

Probability of Diffractive Dissociation

The probability distribution of the diffractive dissociation, or "DD" in short, follows the equation:

$$\sigma_{\mathrm{DD}}(W^2) = 0.68\left(1 + \frac{36\,\mathrm{GeV}^2/\mathrm{c}^4}{W^2}\right)\ln\left(0.6 + 0.1\,\frac{W^2}{\mathrm{GeV}^2/\mathrm{c}^4}\right)\,\mathrm{mb}. \tag{5.107}$$

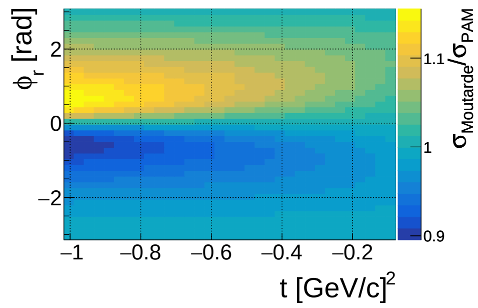
For the ratio between DD and elastic events, the elastic cross section which is measured in the paper for different nuclei, is also required. The protons are measured to [92]:

$$\sigma_{\mathrm{elastic,p}} = 7\,\mathrm{mb}. \tag{5.108}$$

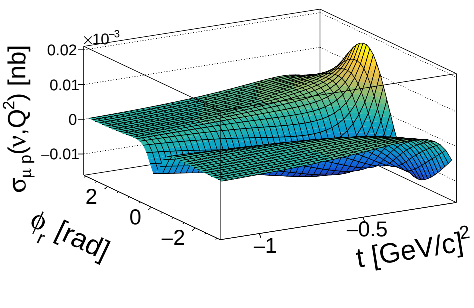
From these two the ratio of DD to exclusive events can be written as:

$$R_{\mathrm{DD}} = \frac{\sigma_{\mathrm{DD}}}{\sigma_{\mathrm{DD}} + \sigma_{\mathrm{elastic,p}}}. \tag{5.109}$$

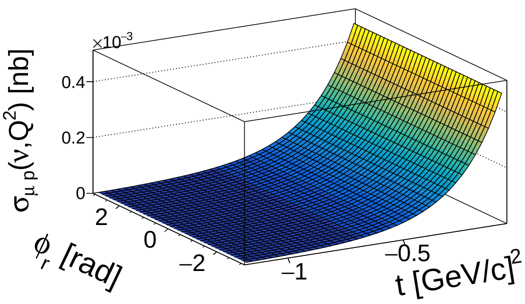
Finally, this ratio can be used with an accept-reject method to generate the desired distribution of DD events. Afterwards, it has to be decided whether the excited nucleon changes or stays the same. The probability of nucleon conversion under charged pion emission is 66 %, which corresponds to the valence quark composition in the struck nucleon. In the case of nuclei that also have neutrons inside, a number draw decides whether the struck nucleon is a proton or a neutron.



(a) Bethe-Heitler



(b) Interference



(c) DVCS

Figure 5.20: Leading order calculations for the cross sections resulting from the GK11 model implemented by H. Moutarde. The kinematics are $Q^2 = 3.0\,(\mathrm{GeV}/c)^2$ and $\nu = 10\,\mathrm{GeV}$. The Bethe-Heitler is not shown absolutely, but divided by the perfect cross section calculation by P.A.M. Guichon.

Mass Distribution

The mass distribution from [92] is shown in Fig. 5.21a. In order to model this distribution a minimal and a maximal mass square first needs to be determined. These are set to:

$$M^2_{X,\mathrm{min}} = m_p^2 + m_{\pi^0}^2, \tag{5.110}$$
$$M^2_{X,\mathrm{max}} = m_p^2 + 0.15\,W^2. \tag{5.111}$$

The peak mass can be estimated from the plot to:

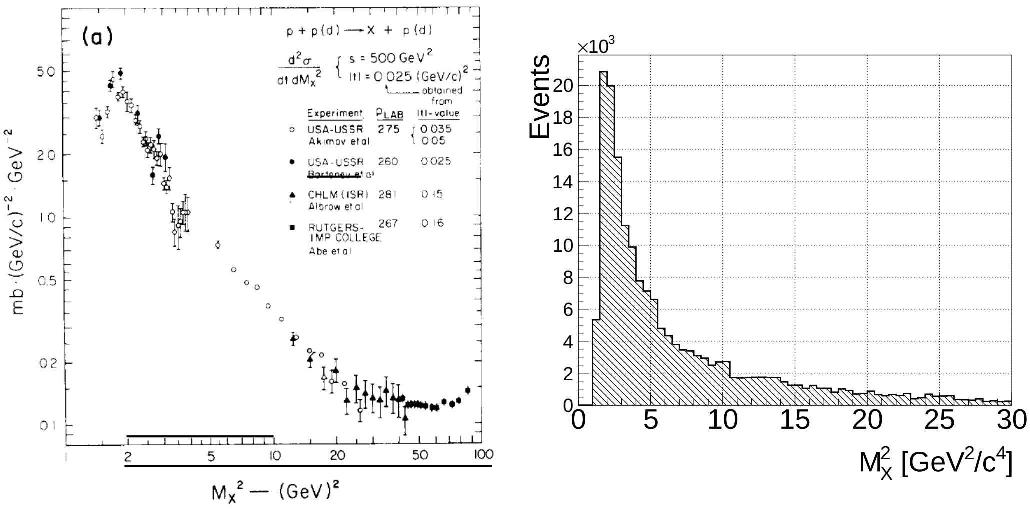
$$M^2_{X,\mathrm{peak}} \approx 2\,\mathrm{GeV/c^2}. \tag{5.112}$$

The distribution over the mass peak is relatively simply to approximate by:

$$p(M_X^2) = \frac{M^2_{X,\mathrm{peak}}}{M_X^2}, \tag{5.113}$$

while below the mass peak the shape of the distribution has to be considered:

$$p(M_X^2) = \frac{M^2_{X,\mathrm{peak}}}{M_X^2}\,\frac{M_X^2 - M^2_{X,\mathrm{min}}}{M^2_{X,\mathrm{peak}} - M^2_{X,\mathrm{min}}}. \tag{5.114}$$



(a) Experimental data from [92]. (b) HEPGen++ distribution.

Figure 5.21: Distributions for M_X^2 of the diffractive dissociation.

The resulting distribution is shown in Fig. 5.21b which fits to the data distribution in the given errors. The multiplicity of the charged and neutral particles can be drawn after the mass is known. It has been empirically confirmed, that they follow gaussian distributions [92] where the standard deviation σ for charged and neutral pions are in good agreement with:

$$\sigma_{\pm} \approx \mu_{\pm}/2, \tag{5.115}$$

$$\sigma_n \approx \mu_0/2, \tag{5.116}$$

with the gaussian means:

$$\mu_{\pm} \approx 2\sqrt{M_X - m_p}, \tag{5.117}$$

$$\mu_0 \approx \frac{\mu_{\pm} - 0.5}{2}. \tag{5.118}$$

Technically, the mean $\mu_{\pm}$ needs to be raised by 1 if a proton was hit. This is due to the fact, than one additional charged diffractive product is needed to fulfill charge conservation.

Two-Body Decay

For the case of a two-body-decay, the kinematics can be calculated exactly. The momentum of the particles in the centre of mass system of the decaying particle is:

$$p/c = \frac{\sqrt{M_X^4 - 2M_X^2(m_0^2 + m_1^2) + (m_0^2 - m_1^2)^2}}{2\sqrt{M_X^2}}. \tag{5.119}$$

As in the previous kinematical cases the ϕ angle is rolled flat while the θ angle is converted as arcus cosine:

$$\phi = 2\pi r, \tag{5.120}$$

$$\theta = \arccos(2r' - 1), \tag{5.121}$$

$$\tag{5.122}$$

where $r, r' \in [0, 1]$ are distributed uniformly. Finally, a boost to the laboratory system gives the right final momentum four-vectors for the diffractive dissociation products.

Higher Multiplicity

In case of higher multiplicity, the momentum balance cannot be simply calculated anymore. In the lower M_X^2 region, the phase space for the momentum distribution is isotropic. The validity for the model of the isotropic phase space can be roughly estimated to $M_X^2 < 4\,\mathrm{GeV}^2/\mathrm{c}^4$. In this region, the momentum distribution has the maximum:

$$p_{\max}/c = \frac{\sqrt{M_X^4 - 2M_X^2(m_p^2 + m_\pi^2) + (m_p^2 - m_\pi^2)^2}}{2\sqrt{M_X^2}}, \tag{5.123}$$

which follows from the kinematic limitations. The momenta of the produced particles are distributed like:

$$p(p) = r^2\, p_{\max}, \tag{5.124}$$

with $r \in [0, 1]$ uniformly.

For the region $M_X^2 > 4\,\mathrm{GeV}^2/\mathrm{c}^4$, the naive model does not hold anymore. In this region, transverse momentum is distributed following the probability function:

$$p(p_\perp) = \frac{2}{\langle p_\perp \rangle} \exp(1)\, p_\perp \exp\left(-2\frac{p_\perp}{\langle p_\perp \rangle}\right), \tag{5.125}$$

where the mean is different for protons or pions:

$$\langle p_\perp \rangle = \begin{cases} 1.5\,\mathrm{GeV/c} & \text{for protons,} \\ 1\,\mathrm{GeV/c} & \text{for pions.} \end{cases} \tag{5.126}$$

The transverse momentum is then set by drawing a flat ϕ angle for the particle and by setting the x component to 0 first. Next, the experimental pseudo-rapidity y is taken from data, where the limit of y is given to:

$$y_{\max} = \ln\left(\sqrt{\frac{M_X^2}{m_p}}\right). \tag{5.127}$$

The y can then be drawn in the range $y \in [-y_{\max}, y_{\max}]$ uniformly. With the final pseudo rapidity, the momentum x component can be calculated in the following way:

$$p_x = \frac{\sqrt{p_\perp^2 + m^2}}{2}(\exp(y) - \exp(-y)), \tag{5.128}$$

where m denotes the actual rest mass of the considered particle.

By default both models of phase space do not conserve the total momentum. There are many particles, each with randomly chosen momentum. A momentum balancing algorithm needs to be introduced to ensure the ensemble of particles fullfills the conservation of momentum.

First, the momentum balance is calculated from all the involved momentum four-vectors:

$$P_{\text{sum}} = \sum_i P_i = (E_{\text{sum}}, \vec{p}_{\text{sum}}). \tag{5.129}$$

Next, a power corrected three-momentum sum is built:

$$\vec{p}_p = \sum_i \sum_{c=1}^{3} \vec{e}_c p_{i,c}^3, \tag{5.130}$$

where $\vec{e}_c$ with $c \in [1,3]$ are the normalized unit vectors in $\mathbb{R}^3$ and p_i denotes the i'th component of $\vec{p}$. Each particle is then corrected by a fraction of these momentum sums in the momentum three vectors:

$$\vec{p}_{\text{new}} = \sum_{c=1}^{3} \vec{e}_c \; p_{\text{old,c}}^3 p_{\text{sum,c}} / p_{p,c}. \tag{5.131}$$

Then, the momentum four-vector is regenerated for each particle:

$$P_{\text{new}} = (\sqrt{m^2 + \vec{p}_{\text{new}}^2}, \vec{p}_{\text{new}}). \tag{5.132}$$

Finally, an iterative process is started and repeated until the energy balance is correct with an uncertainty of $\Delta E < 0.01$ GeV. For the rest of the process, the phase space has to be distinguished again. In the case of longitudinal phase space, the correction is simpler. The momentum for each particle is simply scaled with a scale-factor of:

$$f_p = 1 + \frac{M_X^2 - P_{\text{sum},0}}{\frac{1}{2}\sum_i \left[\frac{|\vec{p}_i|^2}{m_i^2} + |\vec{p}_i|^2\right]}. \tag{5.133}$$

In the case of longitudinal phase space, each hemisphere, which in this case can be identified with the x component of the momentum, $P_{i,1}$, which is greater or lesser than 0, needs to be viewed separately. The momentum sum gets split:

$$p_- = \sum_i P_{i,1}^3 \quad \forall\, P_{i,1} < 0\,\text{GeV/c}, \tag{5.134}$$

$$p_+ = \sum_i P_{i,1}^3 \quad \forall\, P_{i,1} > 0\,\text{GeV/c}. \tag{5.135}$$

The momentum for each particle is then corrected accordingly and depending on the sign of the x component of the momentum:

$$P_{i,1} = P_{i,1,\text{old}} \mp \left(|P_{i,1,\text{old}}|^3 \frac{M_X^2 - P_{\text{sum},0}}{2p_{\mp}} \right). \tag{5.136}$$

5.4 HEPGen++ in PHAST

On the level of analysis, the whole generator information is available to the user. Before HEPGen++ was introduced, this information was used solely to study reconstruction uncertainties and acceptance until now. With the introduction of the HEPGen++ library design and the features of separated phase space factors the user has completely new possibilities at his disposal when it comes to data analysis.

5.4.1 Motivation

The schematic of the COMPASS-II Monte-Carlo chain is shown in Fig. 5.22. The typical scale for time consumption for generating Monte-Carlo events are given in Tab. 5.6.

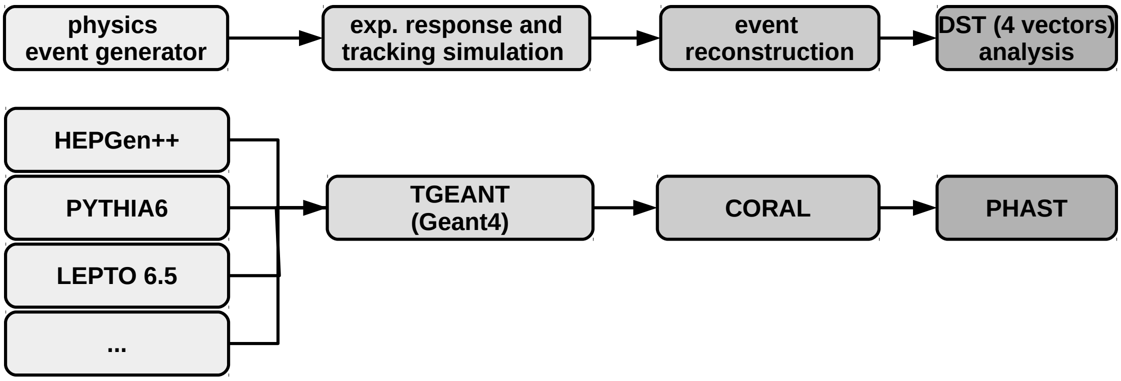


Figure 5.22: Schematic of the compass Monte-Carlo chain.

Table 5.6: Typical timescales for Monte-Carlo event production at COMPASS II.

Step	**Consumed time / event**
Event generation	$\mathcal{O}$(ms)
Detector response	$\sim 20\,$s
Reconstruction	$\mathcal{O}$(s)
Analysis	$\mathcal{O}$(ms)

It is clearly visible that the event generation takes a negligable amount of time compared to the reconstruction, which itself is considerably faster than the detector response simulation. For physics processes where no large amounts of data are available or that are extremely model dependant like the DVCS or exclusive π^0 production, different models may be possible. Some of these models have free parameters for better fitting to data in unknown kinematics. Changing any model parameters or complete models usually requires a full reproduction of the Monte-Carlo, which is time consuming. HEPGen++ in Phast allows to overcome this issue, by allowing to change model parameters or complete models at analysis time.

5.4.2 Kinematic Regeneration

At COMPASS-II, the whole generator information, including the generated momentum four-vectors of all involved particles, real and virtual ones, are available to the user in the analysis toolkit software PHAST. In the case of exclusive event generation the initial and final state per physics channel are then always well defined. The DVCS process is always as follows:

$$\mu\, p \to \mu'\, p'\, \gamma. \tag{5.137}$$

Therefore, knowing all the momentum four-vectors allows the user to recalculate all the necessary variables that the cross section can ever depend on. While it is impractical for each user to do so, a set of supporting tools were created for the purposes of this thesis.

With this simple toolkit a user can regenerate the necessary kinematics from the embedded generator information in the events from PHAST. From there all the implemented models in HEPGen++ are directly applicable with arbitrary parameters. It is possible for a DVCS event to get the weights, which represents the likelihood of the process, from VGG, GK11 and FFS with any number of different parameters. This is schematically visualized in Fig. 5.23. As this takes only milliseconds per event and per parametrization, a user can check which model fits the data best in which parametrization, without generating a new Monte-Carlo sample each time. This saves the major time consuming steps on the way and even may allow for fitting parameters directly to existing data.

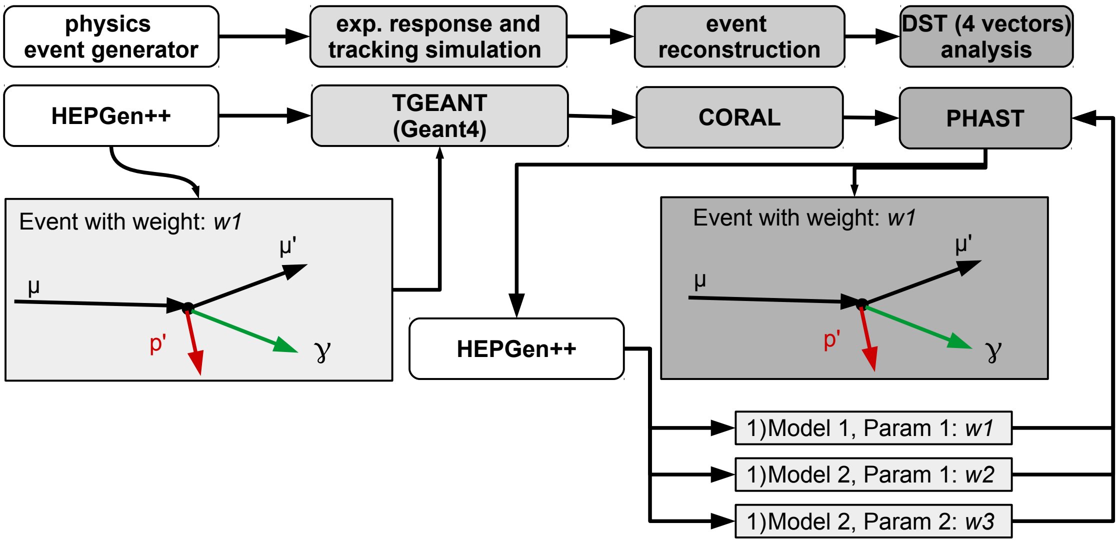


Figure 5.23: Schematic of the HEPGen++ in PHAST tool.

HEPGen++ also offers python bindings for these functions, so the weight-swapping can be done at an even later stage than PHAST. This offers great flexibility for physicists working on the final analysis.

A comparison of two different sets of Regge-parameters for the FFS model, as well as, VGG and the GK11 model are shown in the figures 5.24 as an example.

5.4.3 Lujet Double Precision Splits

Lujets are generator information lines describing particles. They contain a four-momentum, particle identification number and a status code. In a single lujet, only single precision floating point accuracy is forseen by CORAL. The Bethe-Heitler cross section has a very large slope near s- and p-peaks. To determine the cross section in this region accurately, a higher precision is needed than the one provided by the standard lujets, otherwise one faces relative errors of up to 20% in the Bethe-Heitler cross section values. Therefore, a new feature was introduced in order to overcome this problem in HEPGen++: The lujet-double-precision splits. The split that lujets have to be put together are in the same order as the first lujets and have negative k codes to separate them. When a normal analysis is done, which does not require such precision, it is sufficient to just ignore all negative k codes. In the case of high-precision analyses, the lujets can simply be added in C++ into a double precision variable, which restores the full 14 digits from the generator without

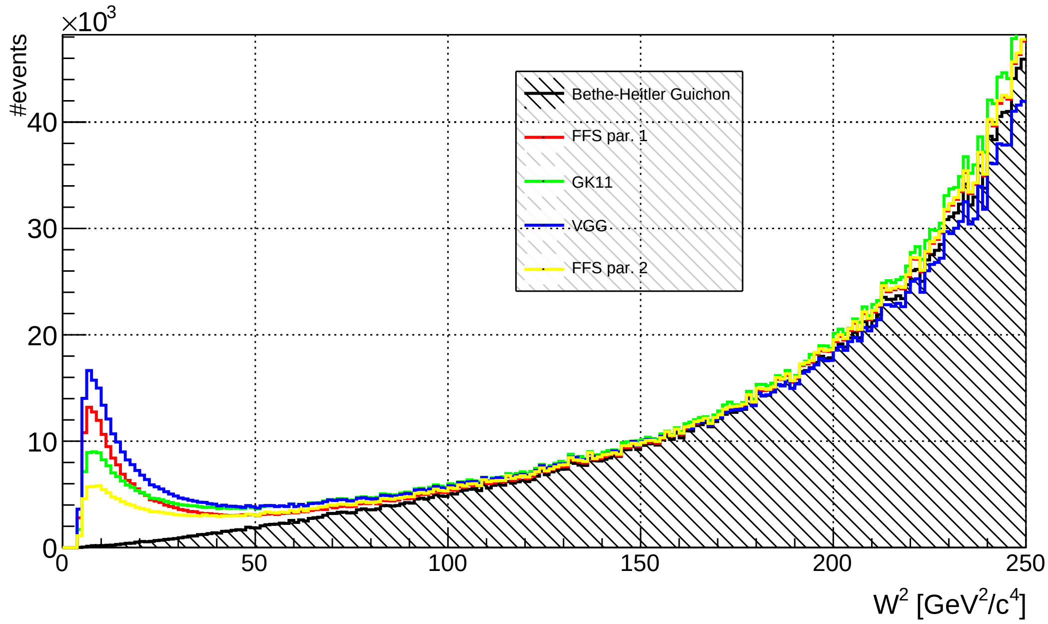


Figure 5.24: The distribution of W^2 from the latest Monte-Carlo sample for the 2012 DVCS run. The differently colored lines stand for different models and parametrizations. The correct pure Bethe Heitler contribution by P.A.M. Guichon is drawn in black.

breaking backwards compatibility at all. A more detailed description is given in appendix D.

5.5 Toolbox

Originally, most of the tools presented in this section were developed for debugging purposes. Some of them became very useful for generating and handling Monte-Carlo files and productions.

5.5.1 LEPTOv2 Management

For a more easy handling of LEPTOv2 files, a library was written to take care of all the low level file interactions. This library allows reading and writing of LEPTOv2 files in both common header lengths. The library is called “libslread” and built by default in any HEPGen++ installation. A full documentation including Doxygen is available online at the TGEANT homepage [93, 94]. The binary LEPTOv2 format is documented in appendix C.

As a proof of concept, a LEPTOv2 file splitter was written in well under 100 lines of code. It is called “leptoFileSplitter” and also available by default. This tool is of high importance when generating LEPTO Monte-Carlo with TGEANT. The LEPTOv2 file that is generated must have the same length as the number of beam particles in the beam file. TGEANT can only read entries from this file consecutively and typically a single TGEANT instance is kept running for 1000 to 5000 events. To cover the full phase space of the beam file, the LEPTO input file needs to be split in parts that TGEANT can completely simulate.

The "leptoFileAnalyzer" can be used for a direct evaluation of the LEPTOv2 event content. This tool makes all physical observables available for each event and interfaces with the histogramming backend of libhepgen. In this way, it is possible to generate distributions directly from the LEPTOv2 files.

5.5.2 OpenGL visualization

To get an overview of generated events, it is often helpful to view a visualization of an event. This can be done without running any simulation software since the HEPGen++ package contains an OpenGL based event viewer. This viewer can display events either from a LEPTOv2 file, with the help of the libslread mentioned earlier, or it can call the libhepgen directly and start a generator of its own. The visualization of an exclusive omega

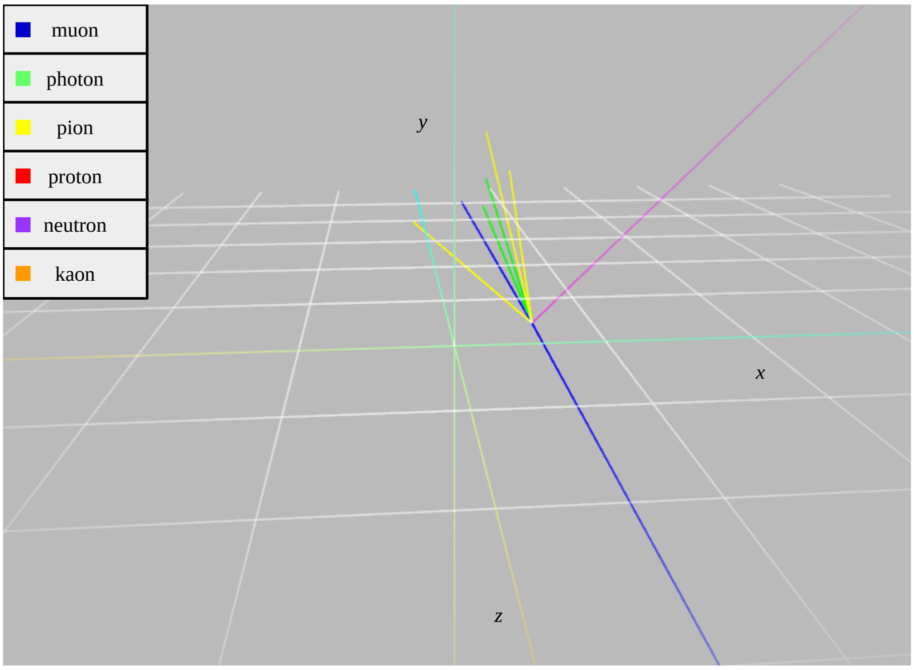


Figure 5.25: The OpenGL event viewer of HEPGen++. It also allows for displaying arbitrary LEPTOv2 files. The offset to the axis is due to the fact that the beam is taken from the beam file. Particle color codes are given in the legend. The event is pseudo-exclusive omega production with a decay into three pions. The target proton is dissociated into a neutron and a pion. The colours of the lines have been washed out close to the edges of the screen with fragment shader effects so that the visualization is more eye-friendly.

event is shown in Fig. 5.25 following the scheme:

$$\begin{aligned} \mu p^{+} &\rightarrow \mu p^{*+} \omega \mu' \rightarrow n\, 2\pi^{+}\, \pi^{-}\, 2\gamma, \\ p^{*+} &\rightarrow n + \pi^{+}, \\ \omega &\rightarrow \pi^{+}\pi^{-}\pi^{0}, \\ \pi^{0} &\rightarrow \gamma\gamma. \end{aligned}$$

The event display program also allows for manual event preselection. With the space bar, the user can display the next event, from either the LEPTOv2 file or the internal generator. If the event is interesting, pressing the "s" key will write this event to the output file. This allows the user to patch together special samples for performance tests in the Monte-Carlo. The output file can then directly by analyzed by the leptoFileAnalyzer tool mentioned above, or it can be processed by TGEANT to generate a detector response for any COMPASS-II setup. The full listing of the hotkeys can be found in Tab. 5.7

Table 5.7: Key bindings for the OpenGL event display.

Key	**Function**
Arrow up/down	Rotate view around x axis.
Arrow left/right	Rotate view around y axis.
Page up/down	Zoom in/out on z axis.
Left click	Start/stop automatic rotation.
Right click	Automatically show new evens all 750 ms.
Space	Show next event.
s	Save event to output file.
Escape	Close the viewer.

On the technical side, the event display uses vertex and fragment shaders in version 330. This allows for a great performance on modern graphics cards as well as great flexibility. The downside of this approach is that the creation and management of the modelview and projection matrices, which project the three-dimensional objects onto the screen need to be managed by the user. As a consequence, the amount of code that was necessary for the visualization increased, but this was the only way to keep the program supported for a longer time.

To enable this feature in the building process, Qt4 needs to be installed. Then, the passing of the flag "-DHEPGEN_EVDISGL=YES" in the cmake process enables the building of this tool.

5.5.3 Beam File Management

The tools available for the beam file management are the same as the LEPTOv2 tools. The libhepgen provides an API[3] for reading and writing beam files. Splitting and histogramming utilities are available an they are used in the same way as the already presented tools.. The beam file splitting comes in handy when generating large amounts of Monte-Carlo from high statistics beam files. These beam files need to be split into chunks of 100 000 entries each in order for LEPTO to read them correctly.

5.5.4 Cross Section Integrator

To evaluate the Monte-Carlo luminosity, the cross sections of the generators need to be integrated in the range of the generation. This can be done with the tool "lumiCalc". This

[3]**A**pplication **P**rogramming **I**nterface

is a simple command line based tool, that calls the cross section equations of the libhepgen, that are also used for "HEPGen-in-Phast", and uses the ROOT integrator on them. This tool is only built if ROOT was found on the target system. While it is not an essential for the usage of the generator, it comes in handy at the analysis time.

5.6 Summary

A working C++ implementation of a flexible event generator for hard exclusive meson production and deeply virtual compton scattering was presented. It features a modern modular design which can be easily extended in the future and which already has a multitude of different models implemented. Furthermore, the generator, due to the nature of the physics and the design, allows access to the models from the analysis level. This offers great flexibility to physicists working on improving model parameters or trying out different models while using only a small fraction of computing power. It is accompanied by a powerful toolkit that allows analysis, cutting or splitting of LEPTOv2 and beam files as well as visualizing events and selecting them by eye. HEPGen++ was used for all exclusive Monte-Carlo productions for the 2012 pilot run of COMPASS-II. The analysis of the DVCS and the neutral pion HEMP both used only HEPGen++ for the signal Monte-Carlo. There is an ongoing analysis of exclusive omega production, which aims for the extraction of spin density matrices and which also uses this generator. This makes HEPGen++ the prevailing generator for the COMPASS-II experiment.

6. Goloskokov and Kroll model for exclusive π^0 production

The exclusive neutral pion production is a very interesting channel for various reasons. The first reason is it being valuable for the GPD constraints directly. From the theoretical side, it is not as clean as the DVCS process, because of the formation of the meson in the final state. Nevertheless, it is a very valuable channel to constrain different GPDs. Another reason for the improvement work on this specific channel is, that the π^0 almost purely decays into two hard photons. Since the cross section for the production is estimated to be large with respect to the DVCS cross section, it is thought to be one of the main sources of background for the DVCS measurement. Until now, only pre-calculated tables from S. Goloskokov and P. Kroll were used for the weighting of the events. These tables did not contain a ϕ dependence for the interference terms and were limited in their Q^2 and W ranges. Furthermore, all of the GPD parameters were fixed too. The reasoning behind using the tables was that the generation of each point from the original Maple files took hours. This thesis features the first C++ implementation of the complete model by Goloskokov and Kroll for usage in the event generator as well as Python bindings, which allow for a very convenient and straight-forward access to the library. In the first section, Sec. 6.1, the GPDs and their parameters are introduced. Furthermore, this section also features the calculations of the subprocess amplitudes in twist-2 and twist-3, as well as the Sudakov form factors for attenuation. In the next Sec. 6.2, a method for interpolating in a precomputed grid of subprocess amplitudes is introduced. This method is already implemented in the library and allows a dramatic speed-up of the calculation of cross sections. The resulting cross section values can be plotted with the help of the interpolation method, examples of which are shown in Sec. 6.3. Finally, a summary is given in Sec. 6.4.

Table 6.1: GPD parameters from [72] and [95].

GPD	α_0	α' [GeV]$^{-2}$	b [GeV]$^{-2}$	N^u	N^d
$\tilde{H}$	0.48	0.9	0.59	-	-
$\tilde{E}_{n.p.}$	0.48	0.45	0.9	14	4
H_T	-0.17	0.45	0.3	1.1	-0.3
$\bar{E}_T$	0.3	0.45	0.5	6.83	5.05

6.1 Model

The GK model was chosen because it is the most sophisticated model available. While this model uses double distribution type GPDs, which are known to have flaws in polynomiality [33], it is the only model available to feature the transversity GPDs that are thought to contribute heavily to the π^0 production. The model can be used to compute the helicity amplitudes and therefore inherently allows for the calculation of any occuring interference terms.

6.1.1 GPDs

The necessary GPDs can be found in [72] with their respective parametrizations. The double distribution ansatz is chosen since it fulfills most needed properties of the GPDs, which is elaborated on in the theory section in more detail. The GPDs are then written as:

$$F_i^a(\bar{x}, \xi, t) = \int_{-1}^{1} \mathrm{d}\rho \int_{-1+|\rho|}^{1-|\rho|} \mathrm{d}\eta \delta(\rho + \xi\eta - \bar{x}) f_i^a(\rho, \eta, t), \tag{6.1}$$

with the double distribution ansatz:

$$f_i^a(\rho, \eta, t) = \exp[(b_i - \alpha_i' \ln \rho)t] \times F_i^a(\rho, \xi = t = 0) \frac{3}{4} \frac{[(1-\rho)^2 - \eta^2]}{(1-\rho)^3} \Theta(\rho), \tag{6.2}$$

where F_i is denoting the forward limit of the parametrized GPD, α_i' a Regge parameter, the slope of the residue function b_i. The typical constant of the Regge parametrization is absorbed into the forward limit F_i. The parameters used for the important GPDs for π^0 production are given in Tab. 6.1. Note that in the standard implementation only the GPD $\tilde{E}$ is used with parameters from this table. In the case of $\tilde{H}$, a special parametrization is used, which is introduced in the following section. Nevertheless, the old parametrization is available in the libGKPi0.

6.1.1.1 $\tilde{H}$

A new parametrization was chosen for the GPD $\tilde{H}$, which fulfills the forward limit correctly and is valid to large $-t$. The equation is [96]:

$$\tilde{H}(x, \xi, t) = \frac{3}{4\xi^3} \int_{\frac{x-\xi}{1-\xi}}^{\frac{x+\xi}{1+\xi}} p(y) \exp[t\, f(y)] \left(\xi^2(1-y)^2 - (x-y)^2\right) \mathrm{d}y, \tag{6.3}$$

where the components are:

$$p(y) = y^{-0.32}(1-y)^0 \left(c_1 + c_2\sqrt{y} + c_3\, y + c_4\, y^{3/2} + c_5\, y^2\right), \tag{6.4}$$

and:

$$f(y) = \left[-\alpha_i' \ln(y) + B\right](1-y)^3 + A\, y\,(1-y)^2. \tag{6.5}$$

Please note that y is only an integration variable here without the physical meaning of the y as energy transfer fraction, as introduced in the theory chapter. The used parameters are given in Tab. 6.2. The resulting GPDs are shown in Fig. 6.1.

Table 6.2: Parameters used for the special $\tilde{H}$ parametrization [96].

Parameter	**u**	**d**
A	1.264	4.198
B	0.545	0.206
α_i'	0.961	0.861
c_1	0.213	-0.204
c_2	0.929	-0.940
c_3	12.59	-0.314
c_4	-12.57	1.524
c_5	0.0	0.0

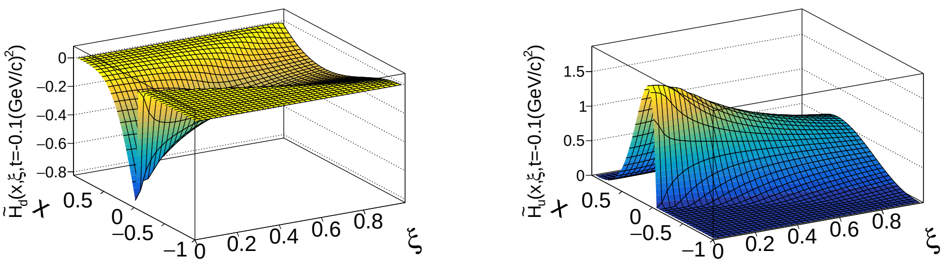


Figure 6.1: The GPD $\tilde{H}$ for d quarks on the left and u quarks on the right. The images are directly generated from the Python interface of HEPGen++.

6.1.1.2 $\tilde{E}$

The GPD $\tilde{E}$ is modeled after the general double distribution scheme introduced in equation 6.1. The final equation with the correct variable names as in the program source code can then be written as [96]:

$$\tilde{E}(x,\xi,t) = \exp(b\,t)\,(N\,\mathcal{D}(0,x,\xi) - 2\,N\,\mathcal{D}(1,x,\xi) + N\,\mathcal{D}(2,x,\xi)), \tag{6.6}$$

where the analytically integrated function part $\mathcal{D}$ is defined differently in three intervals.

For the range $(x+\xi) < 0$ the value is simply 0. If $(x-\xi) < 0$ is fulfilled, the function is given by:

$$\mathcal{D}(i,x,\xi) = \frac{3}{2\xi^3} \frac{\left(\frac{x+\xi}{1+\xi}\right)^{(2+i-k)} (\xi^2 - x + (2+i-k)\xi(1-x))}{(1+i-k)(2+i-k)(3+i-k))}. \tag{6.7}$$

At any other interval the equation can be written as [96]:

$$\begin{aligned}\mathcal{D}(i,x,\xi) =& \frac{3}{2\,\xi^3(1+i-k)(2+i-k)(3+i-k)} \Bigg\{ (\xi^2 - x) \\ &\cdot \left[\left(\frac{x+\xi}{1+\xi}\right)^{(2+i-k)} - \left(\frac{x-\xi}{1-\xi}\right)^{(2+i-k)} \right] \\ &+ \xi(1-x)(2+i-k) \left[\left(\frac{x+\xi}{1+\xi}\right)^{(2+i-k)} + \left(\frac{x-\xi}{1-\xi}\right)^{(2+i-k)} \right] \Bigg\}.\end{aligned} \tag{6.8}$$

The factor k is the Reggeized t dependence:

$$k = \alpha_0 + \alpha'\, t. \tag{6.9}$$

It should be kept in mind that the Regge parameters used are different for each GPD. The

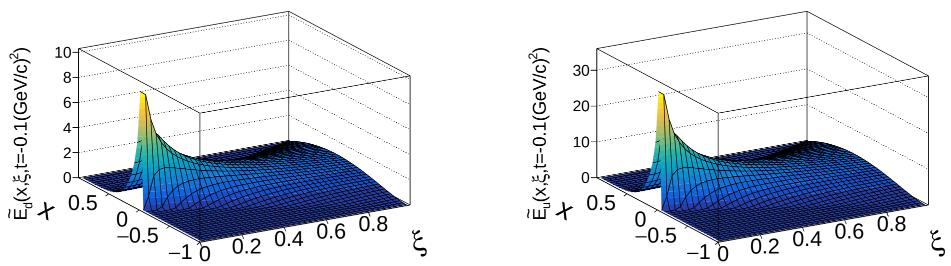


Figure 6.2: The GPD $\tilde{E}$ for d quarks on the left and u quarks on the right. Images directly generated from the python interface of HEPGen++.

resulting GPD $\tilde{E}$ for u and d quarks are shown in Fig. 6.2.

6.1.1.3 H_T

For twist-3 calculations, the transverse GPD H_T is of special importance. It can be calculated using the equation [95]:

$$H_T(x,\xi,t) = N\ \exp[b\,t] \sum_{j=0}^{5} c_j \cdot \mathcal{D}\left(\frac{j}{2}, x, \xi\right). \tag{6.10}$$

The used Regge parameters are given in Tab. 6.1. The special c_j parameters are listed in Tab. 6.3. A plot of the final GPDs can be seen in Fig. 6.3.

Table 6.3: Used parameters for H_T [95].

Parameter	**u**	**d**
c_0	3.653	1.924
c_1	-0.583	0.179
c_2	19.807	-7.775
c_3	-23.487	3.504
c_4	-23.46	5.851
c_5	24.07	-3.683

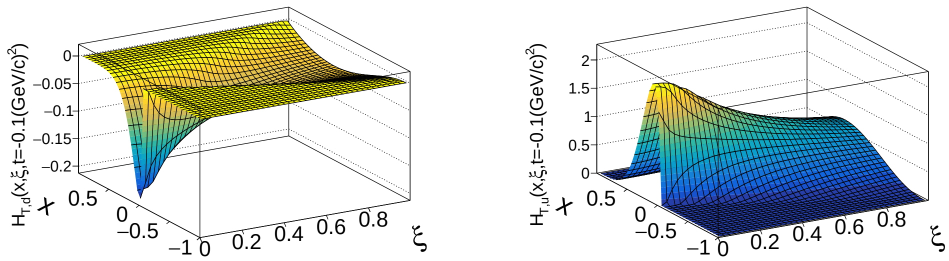


Figure 6.3: The transversity GPD H_T for d quarks on the left and u quarks on the right.

6.1.1.4 $\bar{E}_T$

The equations are different for the two quark types [95]:

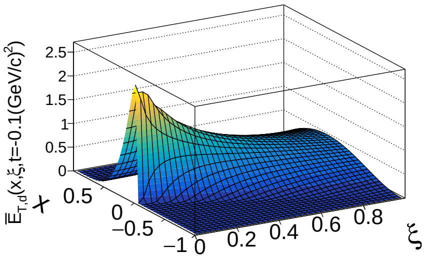
$$E_T^u(x,\xi,t) = N^u \exp[b\,t] \sum_{j=0}^{2} c_j^u \cdot \mathcal{D}\left(\frac{j}{2}, x, \xi\right), \tag{6.11}$$

$$E_T^d(x,\xi,t) = N^d \exp[b\,t] \sum_{j=0}^{4} c_j^d \cdot \mathcal{D}\left(\frac{j}{2}, x, \xi\right). \tag{6.12}$$

The Regge parameters for both of the equations are given in Tab. 6.1 and the coefficients for the power expansion are presented in Tab. 6.4. In Fig. 6.4, the resulting GPDs are drawn.

6.1.2 Sudakov Factor

The Sudakov form factor ensures, that the whole π^0 production is exclusive. In this context, exclusive means that there is no gluon radiation, which later hadronizes to any additional final state particle. This factor is applied by multiplying the convolution of the



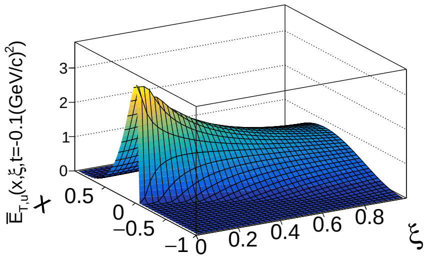


Figure 6.4: The transversity GPD combination $\bar{E}_T$ for d quarks on the left and u quarks on the right.

Table 6.4: Power expansion coefficients used for $\bar{E}_T$ [95].

Parameter	**u**	**d**
c_0	1	1
c_1	0	0
c_2	-1	-2
c_3	0	0
c_4	0	1

meson and the photon wave function with an additional factor e^{-S} where S is the so called Sudakov form factor. The factor itself can be written with the help of the function s [95]:

$$\begin{aligned} s(x,b,Q) =& 2\frac{c_f}{\beta_0}\left[\hat{q}\ln\left(\frac{\hat{q}}{\hat{b}}\right)-\hat{q}+\hat{b}\right] \\ &+c_f\frac{\beta_1}{\beta_0^3}\left[\hat{q}((\ln(2\hat{q})+1)/\hat{q}-(\ln(2\hat{b})+1)/\hat{b})\right. \\ &\left.+\frac{1}{2}(\ln(2\hat{q})^2-\ln(2\hat{b})^2)\right] \\ &+\frac{c_f}{\beta_0}\ln\left(\frac{\exp[2\gamma-1]}{2}\right)\ln\left(\frac{\hat{q}}{\hat{b}}\right) \\ &+4\frac{A_2}{\beta_0^2}\left[\frac{\hat{q}-\hat{b}}{\hat{b}}-\ln\left(\frac{\hat{q}}{\hat{b}}\right)\right], \end{aligned} \tag{6.13}$$

Table 6.5: Constants used in the GK model for exclusive pion production [95].

constant	value	description
n_f	3	Number of flavours
γ	0.57721	The Euler-Mascheroni constant
L_{QCD}	0.220 GeV	Lattice QCD constant
f_π	0.132 GeV	Neutral pion decay constant
N_c	3	Color factor
a_p	1.8 GeV	Twist-3 pion transverse size

where the terms used are:

$$\begin{aligned}
\beta_0 &= 11 - 2\frac{n_f}{3}, \\
\beta_1 &= 102 - 38\frac{n_f}{3}, \\
c_f &= \frac{4}{3}, \\
A_2 &= \frac{67}{9} - \frac{\pi^2}{3} - \frac{10}{27}n_f + 2\frac{\beta_0}{3}\ln\left(\frac{\exp(\gamma)}{2}\right), \\
\hat{q} &= \ln\left[\frac{xQ}{\sqrt{2}L_{\text{QCD}}}\right], \\
\hat{b} &= \ln\left[\frac{1}{bL_{\text{QCD}}}\right].
\end{aligned} \tag{6.14}$$

The constants in these terms are listed in Tab. 6.5 The final Sudakov form factor can now be written as [95]:

$$S(x,b,Q) = -C_1(x,b,Q) - C_2(b), \tag{6.15}$$

$$C_1(x,b,Q) = s(x,b,Q) + (\bar{x},b,Q), \tag{6.16}$$

$$C_2(b) = \frac{4}{\beta_0}\ln\left[\frac{\ln(b^{-2}L_{\text{QCD}}^{-2})}{\ln(\mu_R^2 L_{\text{QCD}}^{-2})}\right]. \tag{6.17}$$

The renormalization constant μ_R is a function $\mu_R(x,Q,b)$. For $x > \bar{x}$ the definition is:

$$\mu_R = \begin{cases} xQ & \text{for:} \quad xQ > b^{-1}, \\ b^{-1} & \text{else.} \end{cases} \tag{6.18}$$

In the other case, $x \le \bar{x}$, μ_R becomes:

$$\mu_R = \begin{cases} \bar{x}Q & \text{for:} \quad \bar{x}Q > b^{-1}, \\ b^{-1} & \text{else.} \end{cases} \tag{6.19}$$

6.1.3 Subprocess Amplitudes

After defining the GPDs, the next step is to calculate the subprocess amplitudes for the π^0 production. In principle, the exclusive production of pseudoscalar mesons can be described in the following way. The virtual photon interacts with a quasi on-shell quark from the target proton, which makes the struck quark virtual. By consecutively emitting a gluon it regains its status as a real quark and becomes one of the two quarks of the vector meson. The virtual gluon then engages in a quark-antiquark pair production, both of whom are on-shell. The produced antiquark becomes the second part of the produced meson, while the real quark is reabsorbed into the struck proton. Because the reabsorbed quark has a different momentum than the struck quark, a GPD is needed instead of a normal PDF. The quark going into the vector meson has a momentum fraction z of the final state meson, hence needs to be integrated over. Because the hard interaction part of the process contains transverse momentum, one also needs to integrate over this. In practice, it is easier to integrate over its' Fourier transformed variable, the impact parameter $\vec{b}$. The subprocess amplitude contains the integration of the wavefunction with the propagator terms and the Sudakov factor over z and $\vec{b}$.

6.1.3.1 Twist-2

In order to make the reading of the equations easier, a refactoring of the z variable is done in the following twist-2 case, where z denotes the momentum fraction of the quark in the final π^0. This is done as [95]:

$$z' = z(1-z). \tag{6.20}$$

Next, the wave function in impact parameter space can be written as [95]:

$$\Psi(z, \vec{b}) = z' \exp\left[-\frac{z'}{4}\frac{\vec{b}^2}{a_\pi^2}\right], \tag{6.21}$$

with the pion transverse size for the twist-2 case [95]:

$$a_\pi^2 = \frac{1}{8\pi^2}\frac{1}{f_\pi^2}. \tag{6.22}$$

The running of α_s is parametrized as:

$$\alpha_s = \frac{12\pi}{33-2n_f}\frac{1}{\ln\left(\frac{\mu_r^2}{L_{\mathrm{QCD}}^2}\right)}. \tag{6.23}$$

The propagators in the impact parameter space are given by:

$$\begin{aligned}\hat{T}_s = &-\frac{i}{4}H_0^{(1)}\left(\sqrt{\bar{z}(x-\xi)\frac{1}{2\xi}}|\vec{b}|\,Q\right)\Theta(x-\xi)\\ &-\frac{1}{2\pi}K_0\left(\sqrt{\bar{z}(\xi-x)\frac{1}{2\xi}}|\vec{b}|\,Q\right)\Theta(\xi-x),\end{aligned} \tag{6.24}$$

$$\hat{T}_u = -\frac{1}{2\pi}K_0\left(\sqrt{z(x+\xi)\frac{1}{2\xi}}|\vec{b}|\,Q\right). \tag{6.25}$$

Following these propagators, the wavefunction and the Sudakov factor, the subprocess amplitude can be constructed:

$$\mathcal{H}^{\pi^0}_{\text{twist}-2} = c_f \sqrt{\frac{2}{n_c}} \frac{Q^2}{\xi} \int_0^1 \mathrm{d}z \int \mathrm{d}^2\vec{b}\Psi(z,\vec{b})\alpha_s \exp[-S][\hat{T}_s - \hat{T}_u]. \tag{6.26}$$

To make this numerically more easily calculable, the radial symmetry of the impact parameter can be exploited in the following way:

$$\int \mathrm{d}^2\vec{b} = \int_0^{2\pi} \mathrm{d}\phi \int_0^{L^{-1}_{\text{QCD}}} \mathrm{d}|\vec{b}|\,|\vec{b}| = 2\pi \int_0^{L^{-1}_{\text{QCD}}} \mathrm{d}|\vec{b}|\,|\vec{b}|. \tag{6.27}$$

This removes the vector dependence and reduces the number of dimensions to integrate over by one. The final integral after further simplifications becomes:

$$\mathcal{H}^{\pi^0}_{\text{twist}-2} = 8\pi^2 f_\pi c_f \frac{Q^2}{\xi} \int_0^1 \mathrm{d}z \int_0^{L^{-1}_{\text{QCD}}} \mathrm{d}b\, b\, \Psi(z,b)\, \alpha_s\, \exp[-S]\,[\hat{T}_s - \hat{T}_u]. \tag{6.28}$$

This integral needs to be solved numerically. It is, however, noted that the resulting function behaves nicely. It is mostly flat with neither extremal values nor divergences. The appearing functions K_0 stand for the Bessel and $H^{(1)}_0$ present the first Hankel functions for the given order.

6.1.3.2 Twist-3

In the twist-3 case the amplitudes are more complicated, but the Sudakov factor and the coupling constant can be reused. The first component is once more the wave function [95]:

$$\Psi(z,\vec{b}) = \frac{4\pi}{\sqrt{2n_c}} f_\pi \mu_\pi a_p^2 \exp\left[-\frac{\vec{b}^2}{2a_p^2}\right] I_0(a_p, |\vec{b}|), \tag{6.29}$$

where I_0 is the cylindrical Bessel function of the zero'th order. With the Sudakov factor and the given wave function, the complete subprocess amplitude can be written as:

$$\begin{aligned}
\mathcal{H}^{\pi^0}_{\text{twist}-3} =& 8\frac{c_f}{\sqrt{2n_c}} \int \mathrm{d}z\mathrm{d}^2\vec{b}\, \Psi(z,\vec{b}\alpha_s \exp(-S) \\
&\times \left\{ \frac{-e_u}{x-\xi+i\epsilon}\delta^2(\vec{b}) + \frac{-e_u}{x+\xi-i\epsilon}\delta^2(\vec{b}) \right. \\
&- \frac{\bar{z}Q^2}{2\xi} e_u \left[\frac{i}{4} H^{(1)}_0\left(\sqrt{\bar{z}\frac{x-\xi}{2\xi}}bQ\right)\Theta(x-\xi) \right. \\
&\left. + \frac{1}{2\pi} K_0\left(\sqrt{\bar{z}\frac{\xi-x}{2\xi}}bQ\right)\Theta(\xi-x) \right] \\
&\left. - \frac{zQ^2}{2\xi} e_d \frac{1}{2\pi} K_0\left(\sqrt{z\frac{\xi+x}{2\xi}}bQ\right) \right\}.
\end{aligned} \tag{6.30}$$

This can be decomposed into five terms, the first two of which can be solved analytically. Therefore, splitting this equation into a part that has to be solved numerically and one that can be integrated out makes sense. This equation can be simplified by again using the cylindrical symmetry of the impact parameter $\vec{b}$. Also, the first two terms, including the two-dimensional delta distribution, can easily be integrated out and absorbed into the convolution directly. Applying all these simplifications the final subprocess amplitude parts 3-5 for twist-3 reads [95]:

$$\begin{aligned}\mathcal{H}^{\pi^0,(3-5)}_{\text{twist}-3} = -4\frac{c_f}{\sqrt{2n_c}}\frac{Q^2}{\xi}2\pi\int_0^1 \mathrm{d}z\int_0^{L^{-1}_{\text{QCD}}} \mathrm{d}b\, b\,\Psi(z,b)\,\alpha_s\,\exp(-S)\Bigg\{ \\ \bar{z}e_u\frac{i}{4}H_0^{(1)}\left(\sqrt{\bar{z}\frac{x-\xi}{2\xi}}bQ\right)\Theta(x-\xi) \\ +\bar{z}e_u\frac{1}{2\pi}K_0\left(\sqrt{\bar{z}\frac{\xi-x}{2\xi}}bQ\right)\Theta(\xi-x) \\ +ze_d\frac{1}{2\pi}K_0\left(\sqrt{z\frac{\xi+x}{2\xi}}bQ\right)\Bigg\}.\end{aligned} \tag{6.31}$$

These need to be solved numerically. The first two parts of the subprocess amplitude read:

$$\mathcal{H}^{\pi^0,(1,2)}_{\text{twist}-3} = 16\pi\frac{c_f}{n_c}\alpha_s f_\pi \mu_\pi a_p^2\left[\frac{-e_u}{x-\xi+i\epsilon}+\frac{e_d}{x+\xi-i\epsilon}\right]. \tag{6.32}$$

They do not contain any integrals anymore and can directly be used in the convolution. Therefore, in the program code they are absorbed into the convolution routine rather than into the subprocess integration routine.

6.1.4 Convolutions and Amplitudes

In the exclusive meson production process, one integral still needs to be performed. The struck quark from the target proton has a fraction of its momentum, called x. This integral is performed at the level of the convolution of the subprocess amplitudes with the GPDs. From the resulting convolutions, all helicity amplitudes can easily be constructed.

6.1.4.1 Twist-2

In the case of twist-2, the integration is straight forward:

$$H_{tc}(\xi,t,Q^2) = +\frac{1}{\sqrt{2}}\int_{-\xi}^1 \mathrm{d}x[e_u\tilde{H}_u(x,\xi,t)-e_d\tilde{H}_d(x,\xi,t)]\times\mathcal{H}^{\pi^0}_{\text{twist}-2}(x,Q^2,t), \tag{6.33}$$

$$E_{tc}(\xi,t,Q^2) = -\frac{1}{\sqrt{2}}\int_{-\xi}^1 \mathrm{d}x[e_u\tilde{E}_u(x,\xi,t)-e_d\tilde{E}_d(x,\xi,t)]\times\mathcal{H}^{\pi^0}_{\text{twist}-2}(x,Q^2,t). \tag{6.34}$$

The amplitudes resulting from this can be written as the combinations of the convolutions:

$$\mathcal{M}_{0+,0+} = \sqrt{1-\xi^2}\frac{e}{Q^2}\left[H_{tc}+E_{tc}\frac{\xi^2}{1-\xi^2}\right], \tag{6.35}$$

$$\mathcal{M}_{0-,0+} = \frac{e}{Q^2}\frac{\sqrt{-t'}}{2m_p}\xi\times E_{tc}. \tag{6.36}$$

6.1.4.2 Twist-3

In the case of twist-3, the integration is slightly more difficult, since the subprocess amplitude terms 1 and 2 contain poles. These poles need to be considered by applying Cauchy principal value to make these expressions solvable:

$$\int_{-\xi}^{1} \mathrm{d}x\, \mathcal{H}^{\pi^0,(1,2)}_{\text{twist}-3} \times H_T(x,\xi,t) \propto \int_{-\xi}^{1} \mathrm{d}x \left[\frac{-e_u H_{T,u}(x,\xi,t)}{x-\xi+i\epsilon} + \frac{-e_d H_{T,d}(x,\xi,t)}{x+\xi-i\epsilon}\right] \quad (6.37)$$

The partial integrals to solve this are:

$$H_{\text{int},+} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x \frac{e_u H_{T,u}(x,\xi,t) - e_d H_{T,d}(x,\xi,t)}{x+\xi}, \quad (6.38)$$

$$H_{\text{int},-} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x \frac{[e_u H_{T,u}(x,\xi,t) - e_d H_{T,d}(x,\xi,t)] - [e_u H_{T,u}(\xi,\xi,t) - e_d H_{T,d}(\xi,\xi,t)]}{x-\xi}. \quad (6.39)$$

The same formalism is applied to the convolutions of the GPD $\bar{E}_T(x,\xi,t)$. The resulting partial integrals are:

$$\bar{E}_{\text{int},+} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x \frac{e_u \bar{E}_{T,u}(x,\xi,t) - e_d \bar{E}_{T,d}(x,\xi,t)}{x+\xi}, \quad (6.40)$$

$$\bar{E}_{\text{int},-} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x \frac{[e_u \bar{E}_{T,u}(x,\xi,t) - e_d \bar{E}_{T,d}(x,\xi,t)] - [e_u \bar{E}_{T,u}(\xi,\xi,t) - e_d \bar{E}_{T,d}(\xi,\xi,t)]}{x-\xi}. \quad (6.41)$$

This solves the first two terms of the subprocess amplitude convolution. Naturally, the terms 3-5, $\mathcal{H}^{\pi^0,(3-5)}_{\text{twist}-3}$, need to be convoluted and integrated too:

$$\bar{E}_{\text{int}} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x\, \mathcal{H}^{\pi^0,(3-5)}_{\text{twist}-3} \times \left[e_u \bar{E}_{T,u}(x,\xi,t) - e_d \bar{E}_{T,d}(x,\xi,t)\right], \quad (6.42)$$

$$H_{\text{int}} = \frac{1}{\sqrt{2}} \int_{-1}^{1} \mathrm{d}x\, \mathcal{H}^{\pi^0,(3-5)}_{\text{twist}-3} \times \left[e_u H_{T,u}(x,\xi,t) - e_d H_{T,d}(x,\xi,t)\right]. \quad (6.43)$$

The amplitudes can be calculated from these terms in the following way:

$$\begin{aligned}\mathcal{M}_{0-,++} = \sqrt{1-\xi^2}\, e \Bigg[& H_{\text{int}} \\ & + \Gamma\alpha_s \left(H_{\text{int},-} + H_T(\xi,\xi,t) \left\{ i\pi - \ln\left(\frac{1-\xi}{2\xi}\right) \right\} \right) \\ & + \Gamma\alpha_s H_{\text{int},+} \Bigg],\end{aligned} \quad (6.44)$$

$$\begin{aligned}\mathcal{M}_{0+,-+} = \mathcal{M}_{0+,++} = -\frac{\sqrt{-t}}{4m} e \Bigg[& \bar{E}_{\text{int}} \\ & + \Gamma\alpha_s \left(\bar{E}_{\text{int},-} + \bar{E}(\xi,\xi,t) \left\{ i\pi - \ln\left(\frac{1-\xi}{2\xi}\right) \right\} \right) \\ & + \Gamma\alpha_s \bar{E}_{\text{int},+} \Bigg],\end{aligned} \quad (6.45)$$

$$\mathcal{M}_{0-,-+} = 0. \tag{6.46}$$

The factor Γ contains all previously absorbed constants. It is given by:

$$\Gamma = 16\pi \frac{c_f}{n_c} f_\pi \mu_\pi a_p^2. \tag{6.47}$$

6.1.5 Full Cross Section

The full cross section for exclusive pion production can be written as [45]:

$$\begin{aligned}\frac{\mathrm{d}^4\sigma_{\gamma^* p\to\pi^0}}{\mathrm{d}t\mathrm{d}x_{bj}\mathrm{d}Q^2\mathrm{d}\phi} =& \frac{1}{2\pi} \\ &\left[\frac{\mathrm{d}\sigma_T}{\mathrm{d}t} + \varepsilon\frac{\mathrm{d}\sigma_L}{\mathrm{d}t} + \varepsilon\cos(2\phi)\frac{\mathrm{d}\sigma_{TT}}{\mathrm{d}t} + \sqrt{2\varepsilon(1+\varepsilon)}\cos(\phi)\frac{\mathrm{d}\sigma_{LT}}{\mathrm{d}t}\right].\end{aligned} \tag{6.48}$$

The partial cross sections can be constructed from the amplitudes calculated earlier [45]:

$$\frac{\mathrm{d}\sigma_T}{\mathrm{d}t} = \frac{\kappa}{2}\left[\mathcal{M}^2_{0-,-+} + \mathcal{M}^2_{0-,++} + \mathcal{M}^2_{0+,-+} + \mathcal{M}^2_{0+,++}\right] \tag{6.49}$$

$$\frac{\mathrm{d}\sigma_L}{\mathrm{d}} = \kappa\left[\mathcal{M}^2_{0+,0+} + \mathcal{M}^2_{0-,0+}\right] \tag{6.50}$$

$$\frac{\mathrm{d}\sigma_{TT}}{\mathrm{d}t} = -\frac{\kappa}{2}\, 2\Re\left[\mathcal{M}^*_{0-,++}\cdot\mathcal{M}_{0-,-+} + \mathcal{M}^*_{0+,++}\cdot\mathcal{M}_{0+,-+}\right] \tag{6.51}$$

$$\frac{\mathrm{d}\sigma_{LT}}{\mathrm{d}t} = \frac{\kappa}{\sqrt{2}}\,\Re\left[\mathcal{M}^*_{0-,0+}\cdot\mathcal{M}_{0-,++} + \mathcal{M}^*_{0+,0+}(\mathcal{M}_{0+,++} - \mathcal{M}_{0+,-+})\right] \tag{6.52}$$

The κ function is a phase space function with an additional factor for changing the units to nb:

$$\kappa = \frac{1}{16\pi}\frac{1}{(W^2 - m_p^2)}\frac{1}{\Lambda}\, 0.3894\, 10^6\,\mathrm{nb}, \tag{6.53}$$

where the Λ function is the kinematic dependent function:

$$\Lambda = W^4 + Q^4 + m_p^4 + 2W^2Q^2 - 2\,W^2m_p^2 + 2\,Q^2m_p^2. \tag{6.54}$$

6.2 Technical Speed-Ups: SPA-interpolation

When trying to generate the full cross section on a per-event base, the technical limitations of the current computers become evident. Even though the convolution integration is done using a Gauss-Legendre integration method, which only requires 128 points, the code needs to numerically solve an integral over z and b for each of these points in twist-3 and twist-2. This adds up to 256 two-dimensional integrals over complex functions, or 512 double-precision floating-point ones, which takes about a minute on a newer computer. While it is acceptable to wait two minutes per point for creating graphs of the resulting cross sections, it is not feasible for an event generator. It should also be mentioned that, while one minute seems rather long for a complete evaluation of a single point of the cross section, it would take more than one hour to do the same evaluation in the original maple calculation. Clearly, the C++ implementation is already much faster.

The cross section calculation tools, which are coming with the package implement a very simple cache, which stores the integrals in the working directory after computation. Therefore, the calculations for any plots with t or ϕ dependence are fast. What is more, the cache will grow over time and more and more calculations will become faster.

A solution feasible for end users is implemented in the libGKPi0 library: The caching of precomputed subprocess amplitudes. Fortunately, the subprocess amplitudes are only dependent on x_{bj} and Q^2 and not t and ϕ so only a two dimensional grid is needed. For technical reasons, the two independent variables for the grid were chosen to (W, Q^2). The currently delivered grid ranges are given in Tab. 6.6. The subprocess amplitudes can be interpolated in between the cached datapoints in a linear way, since they are very smooth. Therefore, this method is called SPA-interpolation[1]. While it can, in principle, be extended to an even larger range, it has to be considered, that this model does not use evolution for the GPDs and therefore should neither be used at a too large Q^2, or a too small Q^2, which has some theoretical issues with the factorization. This cache is invariant of the GPDs used, making it the ideal tool for optimizing the GPDs according to data.

Table 6.6: Ranges and binnings of the provided subprocess integral grid.

Variable	**Number of bins**	**Range**
Q^2	100	(1-26.0) $(\mathrm{GeV}/c)^2$
W	100	(1-21.0) GeV/c^2

To test the accuracy of the SPA interpolation mode, the cross section was calculated in the in a narrow Q^2 and t range with a fixed W in SPA and full calculation mode. The two resulting plots were then combined to get a binwise relative difference between both. The result of the test was, that the points of discontinuity are well visible near the bin edges. The relative difference was always smaller than 1 %. The resulting plots are shown in Fig. 6.5. A combined forward-backward extrapolation was also tested, and is still available in the sourcecode. While it softens the discontinuities in near the bin borders, it also raises the relative error to about 1.4 %. Therefore, it was decided to keep the forward-only extrapolation mode as a default.

6.3 Results

With the help of the subprocess amplitude caching, the cross sections can be plotted in a reasonable time. The plots of the complete cross section can be found in Fig. 6.6. The ϕ dependence is shown in Fig. 6.7. The t evolution at different values of Q^2 for a fixed W is shown in Fig. 6.8.

[1]**S**ub**P**rocess**A**mplitude-interpolation

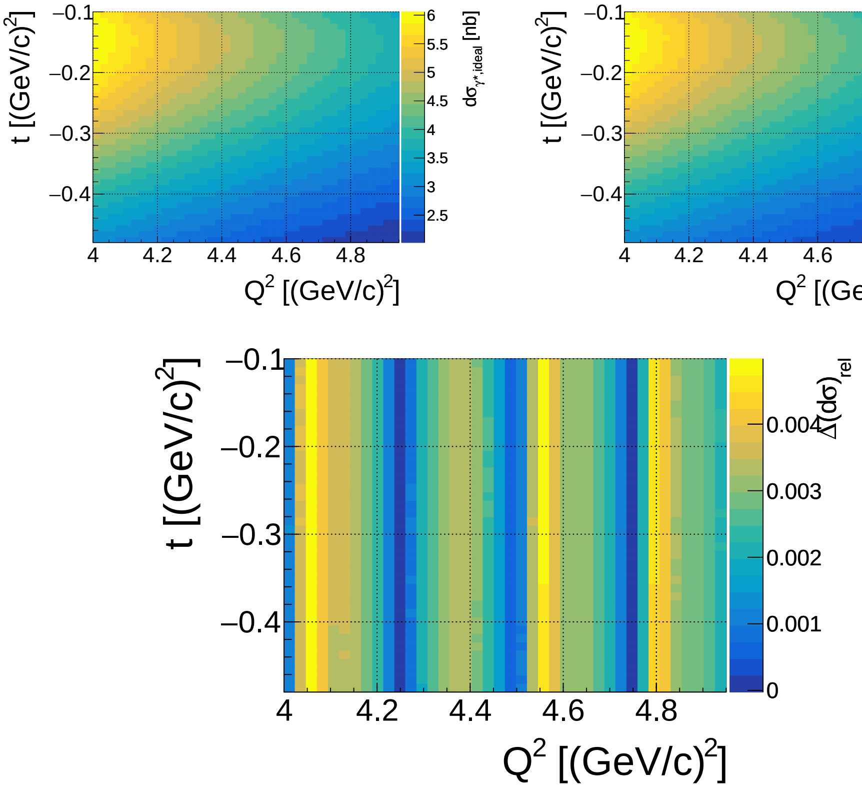


Figure 6.5: Cross section at $W = 6\,\mathrm{GeV}^2/\mathrm{c}^4$ for the full calculation mode (left) and the SPA-Interpolation mode (right) as well as the relative difference of both (bottom).

6.4 Summary

In the scope of this thesis, a full C++ implementation of the most recent Goloskokov and Kroll model was created. The values of the subprocess amplitudes and GPDs were cross checked to an accuracy of 10^{-4}, the cross sections deviate no more than 10^{-3}. Due to the subprocess amplitude caching and interpolation, the process of calculating and integrating the amplitudes and calculating the cross sections became fast. This implementation takes about 100 ms to do so in SPA-interpolation mode, and about 90 s in full calculation mode, the original maple files need an hour. The code is already in use at JLAB and at CERN for the COMPASS-II experiment and is currently the only working implementation in C++ of the model by Goloskokov and Kroll. A release of the π^0 analysis from the 2012 pilot run was finished in this year. It features HEPGen++ with the new libGKPi0 feature for the ϕ and t dependence of the cross section. This release will bring highly anticipated cross section data in the small x_{bj} range, where no data was available before. Any possible modification to the model that will arise during the inclusion of the COMPASS-II data, are easy to implement in the source code. If only the GPDs need to be changed, there is not even a need for regenerating the subprocess amplitude cache. As this regeneration takes the most computing time, different GPD parameter sets can be compared to existing data in a fast and simple way.

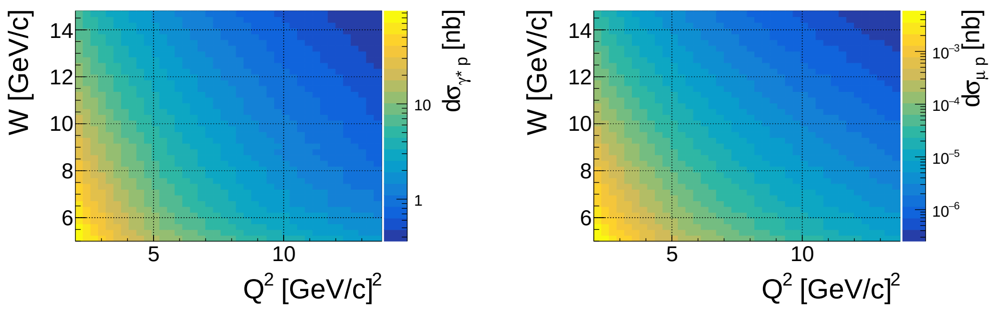


Figure 6.6: Cross section plots for $\frac{\mathrm{d}^3\sigma}{\mathrm{d}Q^2\mathrm{d}W\,\mathrm{d}t}$ at $t' = -0.5\,(\mathrm{GeV/c})^2$. They have been integrated over ϕ so that the interference terms cancel out.

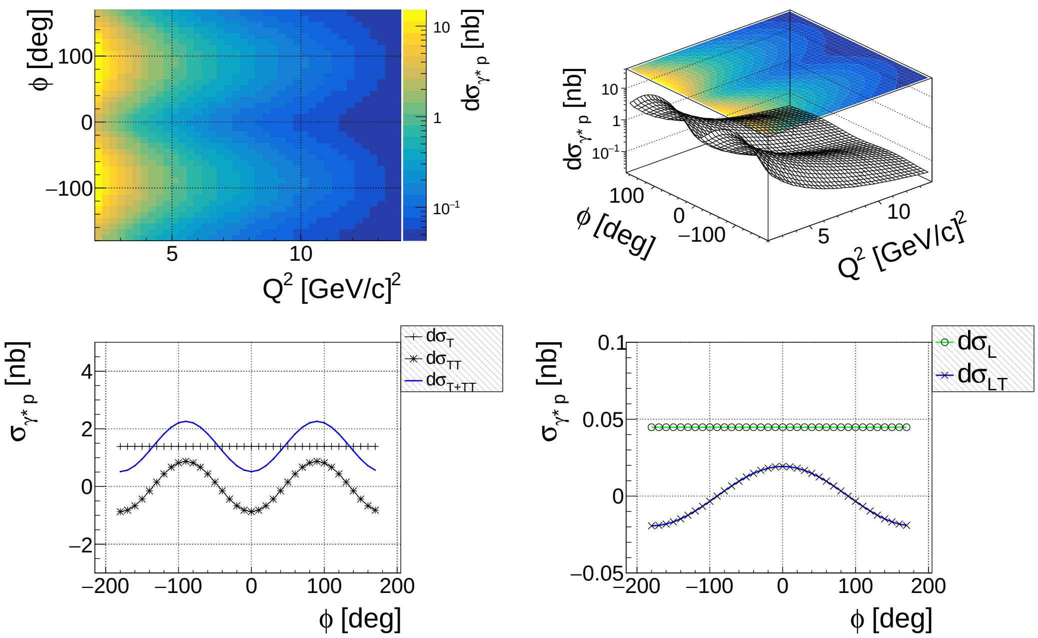


Figure 6.7: Cross section plots for $\frac{\mathrm{d}^3\sigma}{\mathrm{d}Q^2\mathrm{d}W\,\mathrm{d}t}$ at $t' = -0.5\,(\mathrm{GeV/c})^2$. The second row is evaluated at $Q^2 = 4\,(\mathrm{GeV/c})^2$. It shows transverse and longitudinal contribution. It is clearly visible, that at this kinematics the longitudinal contribution is negligable.

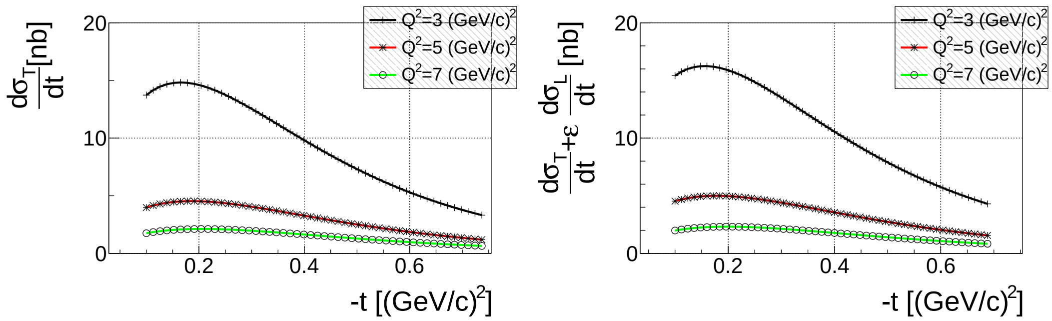


Figure 6.8: The cross sections for a fixed $W = 5\,\mathrm{GeV/c^2}$ at different Q^2 values against t. On the left, the twist-3 contribution, on the right twist-2 and twist-3 combined. The ratio of transverse to longitudinal virtual photons ε was calculated for each kinematic.

7. Conclusion and outlook

This thesis comprises the development of an event generator for exclusive photon and meson production in the COMPASS-II kinematic. In this process, the Goloskokov and Kroll model for exclusive neutral pion production has been implemented. Furthermore, several improvements for the Monte-Carlo software were introduced to the COMPASS-II reconstruction software. The main objective of this work was an increase of the overall accuracy and performance of the Monte-Carlo chain for experimental acceptance studies and model comparisons. The key to extracting the GPDs from the measured data is a very profound understanding of the experimental acceptance and the detector performances, which can only be achieved with a sophisticated simulation. To calculate the sums of cross sections, the acceptance of the experiment needs to be known to an uncertainty of about three percent to extract any significant results.

HEPGen++

The HEPGen++ generator takes care of the event generation of exclusive muon proton production of photons and different mesons. Furthermore, to study the diffractive contamination of the final exclusive sample, it allows for mixing in diffractive dissociation of the target proton mainly into pions. The extracted data will not suffice to directly access the CFFs, because their parameter space is large, hence the extraction will be model-dependent. To this day, several different models and parametrizations exist for the GPDs, which allow the calculation of DVCS cross sections. Finding out which model or parametrization fits the data best is easy when using HEPGen++ because of a revolutionary new feature called "HEPGen-in-Phast". This feature allows to disentangle cross section and detector response allowing to try different models or parameter sets on the analysis level. In most cases, this feature is about a factor of 1000 faster than reproducing the Monte-Carlo sample. The generator was designed in clean, object-oriented C++ making maintainance and expansion easy. By not using any external library, HEPGen++ is independent from external developments. The released data for DVCS and exclusive π^0 rely on the HEPGen++ event generator for acceptance corrections as well as comparison of the results to different theoretical predictions.

GK model for exclusive pion production

The exclusive neutral pion production is interesting for the experiment because of two main facts. The first one is that the GPDs can also be extracted by analysing the cross section of exclusive pion production. The second one is that they are a major background contribution to the DVCS channel. When it comes to the GPD extraction, it should be mentioned that the channel is theoretically not as clean as the exclusive photon channel, since it features a second QCD component for the formation of the pion in the final state. The transversity GPDs that are the main contribution to the pion production are not measurable in the exclusive photon production though. This makes the meson production an important source for completing the set of GPDs. The background for the DVCS measurement from this channel occurs due to the decay of the neutral pion into two hard photons. If one of the photons is lost, or the angle between the two photons is small, the resulting event can easily look like a single photon event in the spectrometer. The contamination of hidden π^0 needs to be known as accurately as possible, as it directly impacts the accuracy of the DVCS background estimation and therefore the cross section, too. The GPD model by S. Goloskokov and P. Kroll is used to describe the exclusive π^0 production process. It features assumptions for a full set of double distribution GPDs, including transversity GPDs that are very important for the pion production. HEPGen++ features the full calculation of all helicity amplitudes and cross section terms in twist-2 and twist-3. Also, the calculation of cross sections takes less than 0.1 s. The presented implementation, called libGKPi0, is a working C++ implementation of the model. It was also provided to other collaborations at JLAB.

Empirical improvements

The Geant4 based TGEANT software and HEPGen++ as the event generator became a quasi-standard in the the COMPASS-II Monte-Carlo chain and almost all recent analyses rely on this combination. A two-dimensional efficiency database was created to allow for local defficiencies. In this way, sets of two-dimensional efficiency maps can be selected for different beam charges, which is important for measuring absolute cross sections. Since the statistics for the extraction of efficiencies is always very limited due to computing time constraints, powerful improvement algorithms were developed. While it has been shown that the expected systematic difference between real and pseudo efficiencies exists, it was also found to be negligable compared to the spread of the detector planes. Because of this result, two full sets of pseudo efficiencies for the 2012 pilot run were extracted, one per beam charge. These sets were modified manually where necessary and Monte-Carlo was made to almost perfectly agree with the real data detector performance in the self-consistency check. In addition, a new module was introduced to improve the beam parametrization and performance of the pile-up muons. To enhance the overall Monte-Carlo production speed, a caching pile-up system was programmed. It allows for the reactions of the muons in the spectrometer to be precomputed and to be saved. A new binary file format was created to enable randomized access to these muon event data. Before reconstruction, a random number of randomly chosen muons are mixed into the signal data. With this approach of the pile-up treatment, the computation time could be speeded up by a factor of 25. Furthermore, this module enables the disentangling of the muon flux from the signal Monte-Carlo generation. This allows the Monte-Carlo generation to start even before the final flux calculation is ready.

Finally, the exclusive pion and photon measurements depend heavily on the performance of the electromagnetic calorimeters and their noise behaviour. The cells and their photomultipliers can show interference with other electronic equipment in the experiment. There are also undetected particles moving through the spectrometer. To cope with all of these effects at once, a new module for the ecal background generation was created. The module allows for extraction of cell-wise energy profiles from real data samples. These profiles can then be used to re-inject background into the calorimeters before the reconstruction, featuring a more sophisticated description of the electromagnetic calorimeter systems of COMPASS-II. Profiles were extracted from the 2012 run and were already used for the generation of Monte-Carlo samples already.

In conclusion, the Monte-Carlo chain is ready for the 2016 and 2017 run. A full and complete Monte-Carlo chain for the experiment now exists, which has been validated and tested on the data of the pilot run in 2012. The interchangeable models feature helps physicists in the analysis and the overall performance is very good. The detector performances, which are difficult to simulate, are now also included in the simulation via empirical insights, which delivers a comprehensive description of the full experiment. For any further requirements that may arise during the run and the future endeavours of COMPASS-II, HEPGen++ is well prepared due to the thoroughly planned object-oriented design. Any changes can be implemented easily and quickly making the reaction time short.

A. Sampling Techniques

Since HEPGen++ is a weighting event generator, all drawn distributions should be flat. If any distribution is a non-flat one, weighting with differential cross sections leads to a problem. When the cross section gets re-extracted from the Monte-Carlo sample, it will show a convolution of the non flat distribution with the functional dependence of the cross section. The generalized formulation of this is:

$$\frac{\mathrm{d}\sigma_{\mathrm{reextracted}}}{\mathrm{d}x} = \mathcal{D}(x) \otimes \frac{\mathrm{d}\sigma_{\mathrm{original}}}{\mathrm{d}x}, \tag{A.1}$$

where $\mathcal{D}(x)$ is the intrinsic functional dependence of the x distribution. Here, x can be any drawn variable. This can only be self-consistent if $\mathcal{D}(x)$ is normalized and uniform and must be fulfilled for every drawn observable, ν or Q^2 being examples.

Problems of strictly uniform distributions

While a uniform distribution is a sufficient approximation for sampling the ν phase space, since the cross sections are smooth in this variable, it is a bad choice for observables showing different behaviour like Q^2. In Fig.A.1 a SIDIS Q^2 distribution is plotted. For comparison two lines are drawn in. To reach an accurate representation of these data a Monte-Carlo sample according to the blue line would have to be generated. Clearly, this would sample the higher Q^2 way finer than necessary and therefore waste CPU time, which is expensive. The red line shows this in more detail. CPU time could be saved in the higher Q^2 regions, where the sampling is a bit too fine, with respect to the sampling in the lower Q^2 regions. In the lower Q^2 regions events would have to be weighted up, which is dangerous as with too low statistics the other phase space variables do not gain enough statistics for their own sampling to be accurate. The exclusive cross sections' dependence on Q^2 is shown in their respective sections, it is mostly $\mathrm{d}\sigma \sim \frac{1}{Q^2}$ or $\mathrm{d}\sigma \sim \frac{1}{Q^4}$.

The solution to this problem is to sample with a distribution that is similar to the real data distribution for the event generation. To keep the phase space flat, the distribution must be invertable so that a $\mathcal{D}^{-1}(x)$ can be found, for x being any drawn variable. The

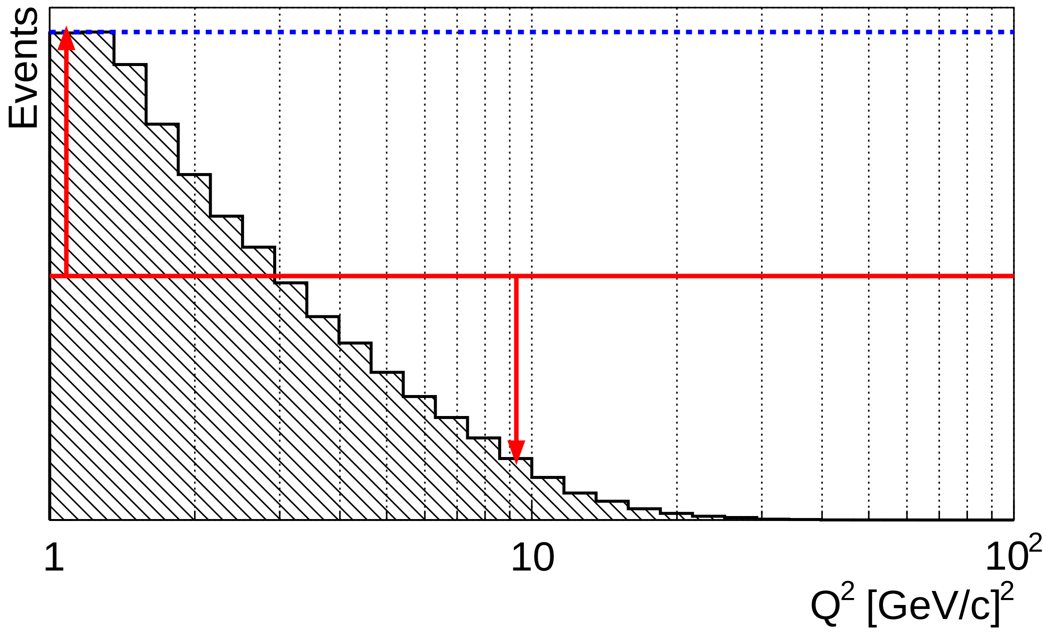


Figure A.1: SIDIS Q^2 distribution from the 2012 run showing the problem of uniform sampling in non-uniform natural distributions. Sampling uniform with statistics equivalent to the red line, will underrepresent the events in the low Q^2 region, while it will overrepresent the events in the higher Q^2 region.

normalized inversion can then be multiplied to the usual weights and in that way solve the problem:

$$\frac{\mathrm{d}\sigma_{\text{reextracted}}}{\mathrm{d}x} = \frac{\mathcal{D}^{-1}(x)\mathcal{D}(x)}{|\mathcal{D}^{-1}(x)\mathcal{D}(x)|} \cdot \frac{\mathrm{d}\sigma_{\text{original}}}{\mathrm{d}x}. \tag{A.2}$$

In principle, one could take all the Q^2 dependences for each of the cross sections and generate according to them. While this would allow the generator to be even more efficient at sampling correctly, it makes the usage of different cross section models for the same process more difficult and enlarges the code that needs to be maintained.

Arbitrarily distributed random numbers

In general, there are many different methods to generate random numbers in different distributions. The methods used in HEPGen++ are presented in the following section. Whenever possible, the efficient method of Integral inversion was used, whereas in all remaining cases it was receded to the accept-reject method was resorted to.

Integral inversion method

The integral inversion method, also named cumulative density function method (CDF) is an efficient method for sampling according to a desired distribution. This distribution must be integrable and the integral must be invertible for the method to work. The first step is to build the CDF in short according to:

$$\text{CDF} : g(x) = \frac{1}{\int_{x,\text{min}}^{x,\text{max}} f(x')\mathrm{d}x'} \int_{x,\text{min}}^{x} f(x')\mathrm{d}x', \tag{A.3}$$

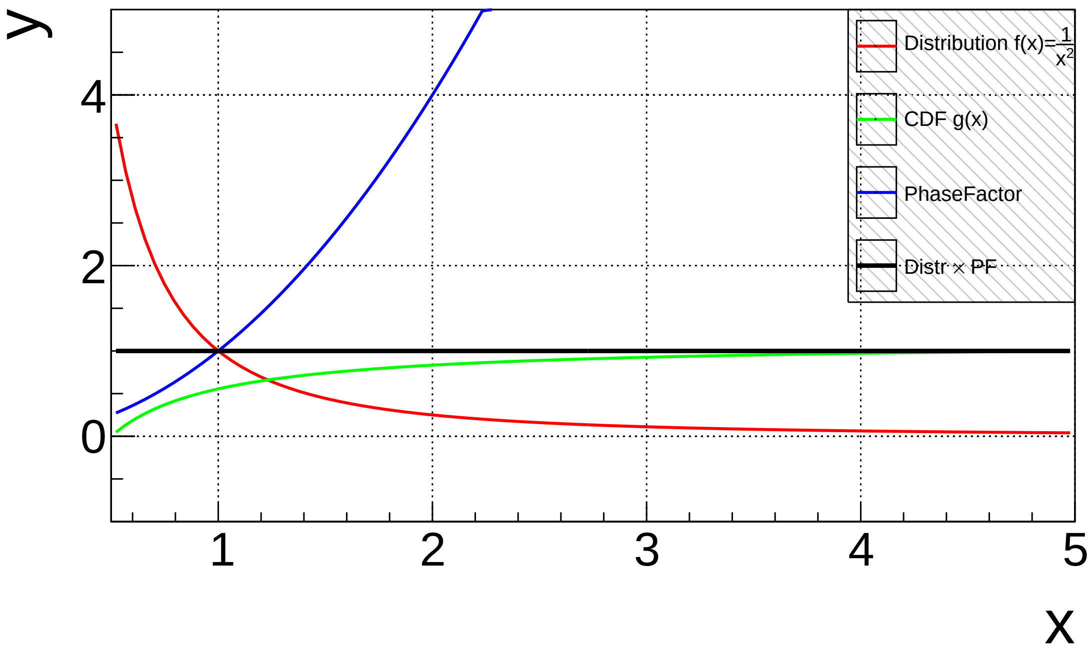


Figure A.2: Visualization of drawing random numbers from a distribution $\frac{1}{x^2}$ for the example in Sec. A .

where $f(x)$ is the desired distribution. The function $g(x)$ is normalized: $g(x) \in [0,1]$. This needs to be analytically solved for the next step, which is inverting the function. With the inversion, the function $g^{-1}(y)$ is obtained, where $y \in [0,1]$ can now be drawn uniformly. After applying the function $g^{-1}(y)$ the resulting random numbers will be distributed according to the original $f(x)$ distribution.

As an example, which is visualized in Fig. A.2, the calculation is shown here for the distribution $1/x^2$ with the range $x \in [0.5, 5]$. The range is needed because of the pole at 0. First the CDF needs to be calculated according to equation A.3:

$$h(x) = \int f(x)(d)x = -\frac{1}{x},$$
$$[h(x)]_{0.5}^{5.0} = 1.8,$$
$$g(x) = \frac{1}{1.8}\left[h(x) - h(0.5)\right].$$

After calculating the CDF, it needs to be inverted:

$$g^{-1}(y) = \frac{5}{9y - 10}, \tag{A.4}$$

where $y \in [0,1]$ uniform leads to a distribution of $g^{-1}(y)$ of $1/y^2$, which is exactly what was wanted. The inversion of the CDF can be interpreted and visualized as drawing uniform on the y axis. This is shown in Fig.A.3, where a random number $y = 0.5$ is drawn. Following a straight line parallel to the x axis until there is an intersection with the CDF then leads to $x \approx 0.9$.

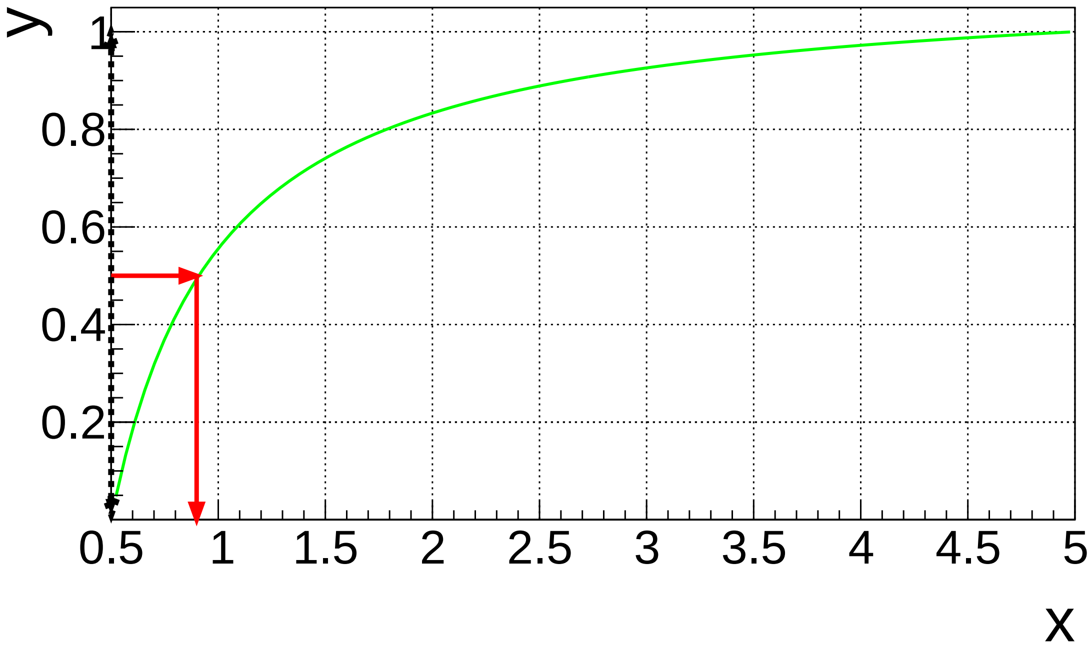


Figure A.3: Drawing a uniformly distributed random number $r \in [0, 1]$ can be visualized as selecting a value on the y-Axis and then intersecting with the CDF function to get the corresponding x value. In this example drawing $r = 0.5$ would result in my distributed random number as $r' \approx 0.9$.

The remaining issue here is that the resulting distribution has a non-uniform functional dependence, which is why this method was chosen in the first place. In equation A.2 the solution was to introduce a $\mathcal{D}^{-1}$ and normalize it. This normalized factor, that will be called "phase factor" from now on, can be built according to:

$$\hat{\mathcal{D}}^{-1}(x) = \frac{1}{f(x)} \cdot \frac{1}{\int_{x,\mathrm{min}}^{x,\mathrm{max}} f(x')\mathrm{d}x'}. \tag{A.5}$$

For the example case, $f(x) = \frac{1}{x^2}$, this factor was computed and it is shown in Fig.A.2. To demonstrate the validity of this approach the convolution of distribution and phase factor is drawn as the black dashed line. It is a straight line at 1, so the resulting distribution after sampling like this and weighting the numbers with the according phase factor would always result in a normalized uniform distribution, but with a non-uniform sampling. Such approach allows for higher statistics to be in regions where they are in the real data without double accounting for the events.

Accept-reject method

The accept-reject method is a very simple method for generating random numbers according to arbitrary distributions. No further requirements on the distribution are needed, except that the local maximum k needs to be known in the range in which the random numbers are generated. The first step is generating a uniform distribution in the full range allowed by the desired distribution $h(x)$:

$$x \in [x_{\mathrm{min}}, x_{\mathrm{max}}]. \tag{A.6}$$

The next step is evaluating the probability of this drawn random number according to:

$$p(x) = \frac{h(x)}{k}. \tag{A.7}$$

In the next step a second random number is drawn uniformly:

$$r \in [0, 1]. \tag{A.8}$$

For the case that $r < p(x)$ the random number x is accepted. Otherwise, the drawn x is rejected and the process gets restarted with a completely new x.

This method has the problem of being inefficient in any case and it even becomes worse the more the maximum deviates from the mean. Fortunately, in the case of HEPGen++, all distributions which cannot be generated using the CDF method are showing a good behaviour with this method.

B. Phase Space Integration

When integrating the total cross section the integral needs to be modified:

$$\sigma_{\text{tot}} = \int_{\nu,\min}^{\nu,\max} \mathrm{d}\nu \int_{Q^2,\min}^{Q^2,\max(\nu,x_{bj})} \mathrm{d}Q^2 \int_{t',\min}^{t',\max} \mathrm{d}t' \; \sigma_{\text{diff}}(\nu, Q^2, t'), \tag{B.1}$$

with $[Q^2, \max(\nu, x_{bj})]$ being taken from the equation 5.4. Practically, it is easier to introduce a new function:

$$\mathcal{Y}(\nu, Q^2) = \begin{cases} 1 & \text{for} \quad Q^2 < 2M_p \nu x_{bj,\max}, \\ 0 & \text{else.} \end{cases} \tag{B.2}$$

In the usual COMPASS-II kinematical range, the other boundaries do not come into play. Therefore, the cut on the upper bound of Q^2 is enough to ensure a well-behaved integration. With this new function, the cross section integral becomes:

$$\sigma_{\text{tot}} = \int \mathrm{d}\nu \int \mathrm{d}Q^2 \int \mathrm{d}t' \; \sigma_{\text{diff}}(\nu, Q^2, t') \cdot \mathcal{Y}(\nu, Q^2), \tag{B.3}$$

where all static integral ranges that are only taken from user input are dropped for legibility. With this helper function, the problem of normalization of the Monte-Carlo sample is solved.

C. File formats used in HEPGen++

The following appendix lists and visualizes all file formats that are used in HEPGen++.

C.1 Beamfile

The beam file format is visualized in the figure C.1.

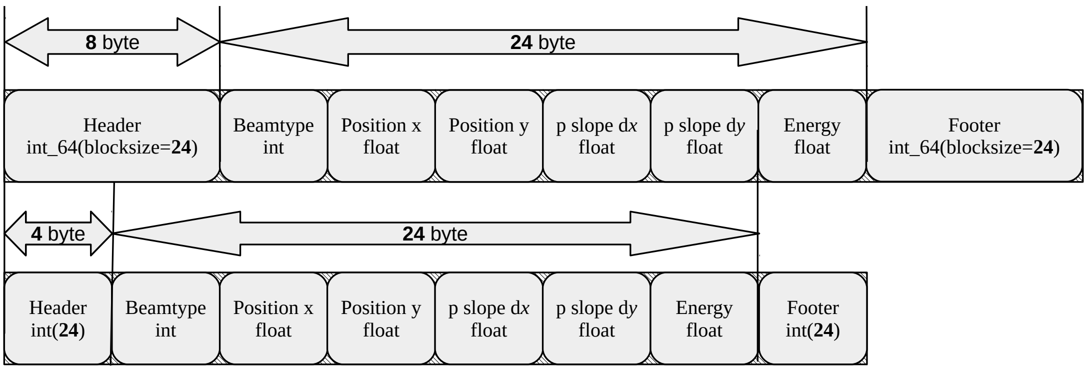


Figure C.1: The beam file format used at COMPASS and for HEPGen++. The upper part shows the gfortran compatible header sizes, the lower part the old, legacy and F77 compatible one. The energy in this format is just the kinetic energy, not containing the particle mass.

C.2 LEPTOv2

The format is visualized in the figure C.2.

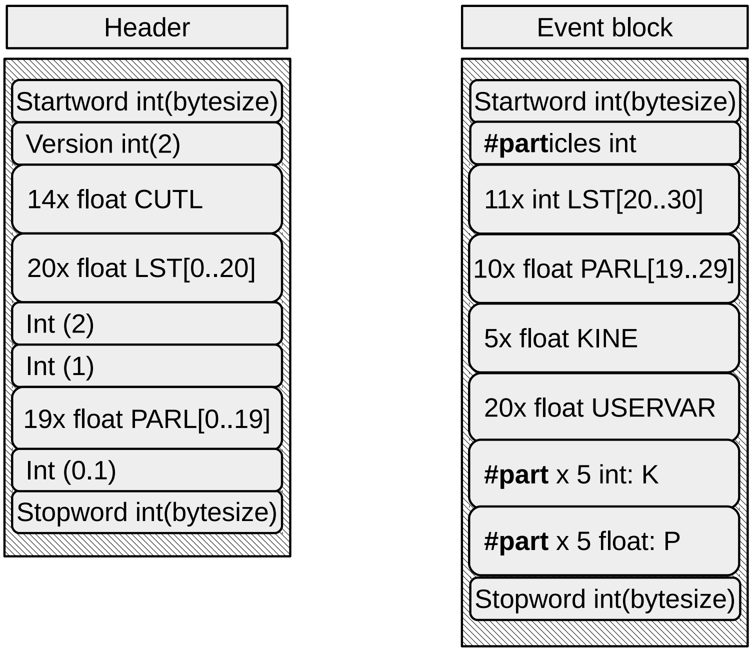


Figure C.2: The LEPTOv2 format as implemented in HEPGen++. The complete specification which variable contains what value is added in the appendix.

C.3 Generator variables

The different blocks of generator variables from the LEPTOv2 formats are filled according to the tables in this section.

Table C.1: **k codes** for status display.
In the case of HEPGen++ only the codes 1,11,21 can occur. The rest is given for completeness from the PYTHIA standard.

Number	Description
1	Undecayed final state
2,3	Unfragmented jets
4	Particle did not decay in given space
11	Decayed or fragmented particle or jets
12	A fragmented jet
13	A removed jet
14	A parton that branched into further partons
15	A forced decay
21	Documentation line
31	A line with information on sphericity, thurst or cluster

Table C.2: **USERVAR** vector contents

Number	Description
0	Beam position x (if beamfile was used)
1	Beam position y (if beamfile was used)
2	Sum of weights
4	Polarization of the vector meson
7	$Q^2_{\text{min}}(\nu)$
8	$Q^2_{\text{max}}(\nu)$
9	Total phase factor
10	Target mass
11	Probability for coherent scattering
12	Slope for coherent scattering
13	Slope for incoherent scattering
14	alf
15	Weight for DVCS only
16	Weight for Bethe-Heitler only
17	B_0 - Regge parameter for FFS-DVCS
18	$x_{bj,0}$ - Regge parameter for FFS-DVCS
19	α' - Regge parameter for DVCS

Table C.3: **PARL** vector contents

Number	Description
0	Target mass (A)
1	Target proton count(Z)
2	Beam energy
3	Struck hadron energy
4	Produced particle
5	Physics program (IRPOC)
6	Exclusively produced particle
7	$\Theta_{\max}$
10	Target mass
11	Probability for coherent scattering
12	Slope for coherent scattering
13	Slope for incoherent scattering
14	alf
15	Generator limit: $t'_{\min}$
16	Generator limit: $t'_{\max}$
17	Beam charge
18	Beam helicity
20	Beam energy
21	$\Theta_{\mu'}$
23	Active mass square M_X^2
24	t'
25	t
26	t_0
28	Θ_π
29	Θ_ϕ

Table C.4: **LUJETS** particle contents

Array	Number	Description
k	0	Status (see table C.1)
k	1	Particle identification number
k	2	Particle origin LUJET line number
k	3	LUJET line number where daughters begin
k	4	LUJET line number where daughters end
p	0	Momentum in x direction: p_x
p	1	Momentum in y direction: p_y
p	2	Momentum in z direction: p_z
p	3	Energy
p	4	Mass

C.4 ASCII histograms

The ASCII format for the ROOTless histogramming backend is visualized in the figure C.3.

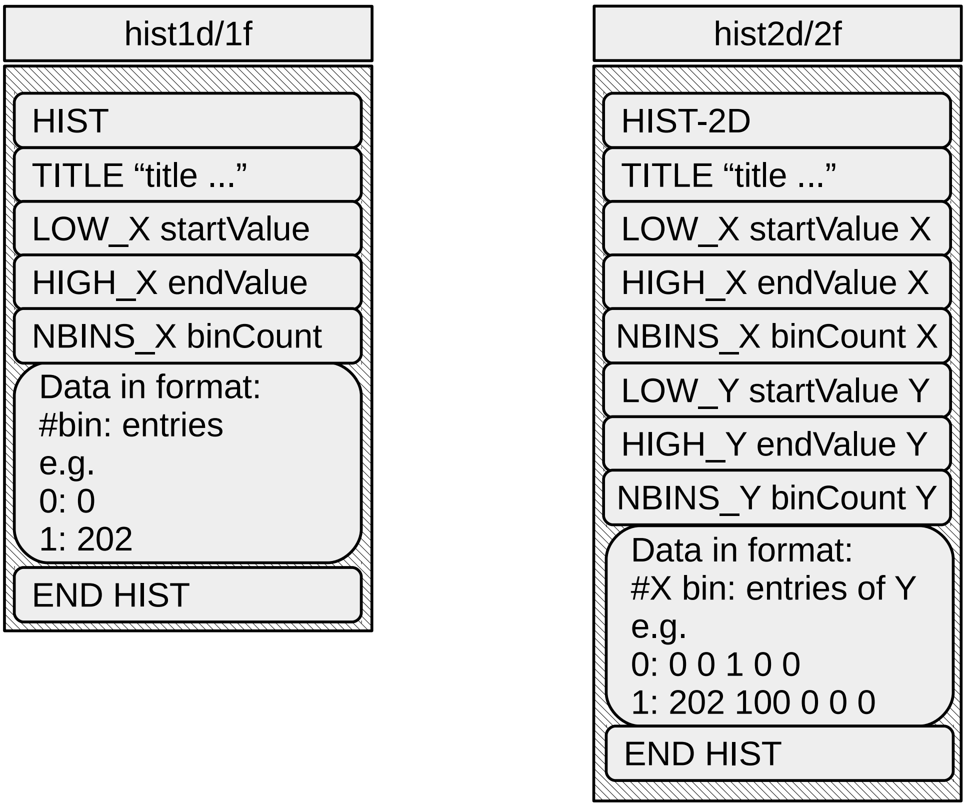


Figure C.3: The text file format for histograms made with HEPGen++. Mind that for a number n of bins in x or y there are always $n+2$ entries. The entry 0 is the underflow bin, the entry $n+1$ is the overflow bin.

D. Lujet splits

The Bethe-Heitler cross section has a very large slope near the so-called cat's ear region. To regenerate the cross section in this region accurately, a higher precision is needed than the one provided by the standard lujets. However, this lead to relative errors of up to 20% in the Bethe-Heitler cross section values.

Therefore, a new feature was introduced in order to overcome this problem in HEP-Gen++: The lujet-double-precision splits. In a single lujet, only a single precision floating point variable is reserved by CORAL.

A convenient feature of the memory representation of floating point variables was used in order to remove this limitation. In computers, floating point variables are stored in 4 byte or 32 bit, in a scientific representation [97]:

$$f = S \times M \times 2^E, \tag{D.1}$$

Where S is the sign bit, M is the mantissa, representing the significant and E is the exponent. The first bit gives the sign, the next eight bits represent the exponent and the remaining bits are for the mantissa. The exponent is stored with a special convention allowing it to range from -127 to 128 without a special sign bit, but this is not of special relevance for the method introduced here [97].

This representation allows the eight-byte double precision floating point variables to be cut into two single-precision ones in the following way:

$$f^d \approx f^s_{\text{high}} + f^s_{\text{low}}. \tag{D.2}$$

Or in memory formulation:

$$S \times 2^E \sum_{i=0}^{52} M \times 2^{-i} \approx \left[S_1 \times 2_1^E \sum_{i=0}^{23} M_1 \times 2^{-i} \right] + \left[S_2 \times 2_2^E \sum_{i=0}^{23} M_2 \times 2^{-i} \right]. \tag{D.3}$$

The original accuracy of double precision with its mantissa size is:

$$a_d = 53 \log_{10}(2) \approx 16, \tag{D.4}$$

whereas the single precision float has:

$$a_s = 23 \log_{10}(2) \approx 7 \tag{D.5}$$

and the combined doubled single precision results in:

$$a_{2s} = 2\, a_s \approx 14. \tag{D.6}$$

Even though two decimals are lost, this method doubles the precision of the lujets. The remaining two digits are technically not significant anymore, which is shown in the Fig. D.1.

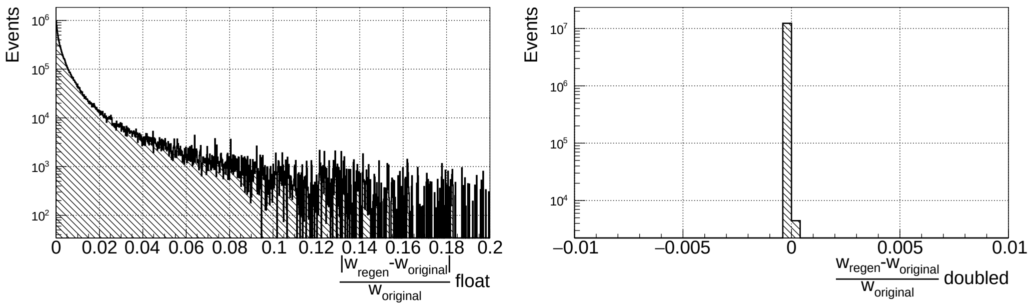


Figure D.1: Weights recalculated off of kinematics extracted from single precision floats (left) and doubled single precision floats (right).

E. Efficiencies

E.1 Comparison real and pseudo efficiency

The following tables feature the full efficiency comparison cross checked by Artem Ivanov.

Table E.1: Full comparison of the detector planes of PB of the MWPCs from 2012-DVCS runs with μ^+ beam.

Detector	**Real efficiency**	**Pseudo efficiency**	**Difference**
PB01U1__	0.9766	0.9791	-0.0025
PB01X1__	0.9797	0.9825	-0.0029
PB02V1__	0.9647	0.9678	-0.0031
PB03U1__	0.9806	0.9841	-0.0034
PB03X1__	0.9845	0.9874	-0.0029
PB04V1__	0.9825	0.9858	-0.0034
PB05U1__	0.9638	0.9654	-0.0017
PB05X1__	0.9917	0.9932	-0.0015
PB06V1__	0.9845	0.9874	-0.0030

Table E.2: Full comparison of the detector planes PS of the MWPCs from 2012-DVCS runs with μ^+ beam.

Detector	Real efficiency	Pseudo efficiency	Difference
PS01U1__	0.9710	0.9726	-0.0016
PS01V1__	0.9830	0.9840	-0.0010
PS01X1__	0.9739	0.9769	-0.0029
PS01Y1__	0.9837	0.9871	-0.0034

Table E.3: Full comparison of the detector planes of PA of the MWPCs from 2012-DVCS runs with μ^+ beam.

Detector	Real efficiency	Pseudo efficiency	Difference
PA01U1__	0.9793	0.9811	-0.0017
PA01V1__	0.9779	0.9797	-0.0018
PA01X1__	0.9681	0.9717	-0.0036
PA02U1__	0.9803	0.9806	-0.0003
PA02V1__	0.9861	0.9864	-0.0003
PA02X1__	0.9709	0.9758	-0.0049
PA03U1__	0.9785	0.9803	-0.0018
PA03V1__	0.9794	0.9811	-0.0016
PA03X1__	0.9704	0.9724	-0.0020
PA04U1__	0.9606	0.9616	-0.0010
PA04V1__	0.9762	0.9777	-0.0015
PA04X1__	0.9727	0.9729	-0.0002
PA05U1__	0.9808	0.9815	-0.0006
PA05V1__	0.9813	0.9814	-0.0001
PA05X1__	0.9822	0.9824	-0.0002
PA06U1__	0.9788	0.9790	-0.0002
PA06V1__	0.9716	0.9735	-0.0019
PA06X1__	0.9498	0.9531	-0.0033
PA11U1__	0.9717	0.9752	-0.0035
PA11V1__	0.9536	0.9534	0.0002
PA11X1__	0.9597	0.9600	-0.0003

F. Efficiencies database

The structure of the sqlite database is given in Tab. F.1. The efficValues flag is filled with an ASCII encapsulated histogram format. All of the x bins for a fixed y are concenated and separated with whitespace. After the last value, a “
” is added to indicate a line break.

Table F.1: SQLite database structure for efficiencies storage

Name	**Type**	**Description**
ID	INT	Autoincrementing indexing field.
TBNAME	TEXT	The TBNAME of the detector.
DETNAME	TEXT	The detector name.
UNIT	INT	The unit number.
YEAR	TEXT	An identifier flag for a year, a charge or whatever you want to group efficiencies with.
binsX	INT	Number of bins in x direction.
binsY	INT	Number of bins in y direction.
startX	REAL	The beginning of the detector in x.
startY	REAL	The beginning of the detector in y.
endX	REAL	The end of the detector in x.
endY	REAL	The end of the detector in y.
efficMean	REAL	The arithmetic mean of all filled bins.
efficValues	TEXT	ASCII encapsulated histogram.

G. Binary data format for TGEANT: tbin

The binary format for the external pre-computed pileup is shown in figure G.1. The header data words are given in table G.1.

Table G.1: Header and footer words for the tbin format.

Header	**Value**
Main	0xdeadbeef
SubBlock	0xcafebabe
DataBlock	0xdabbad00

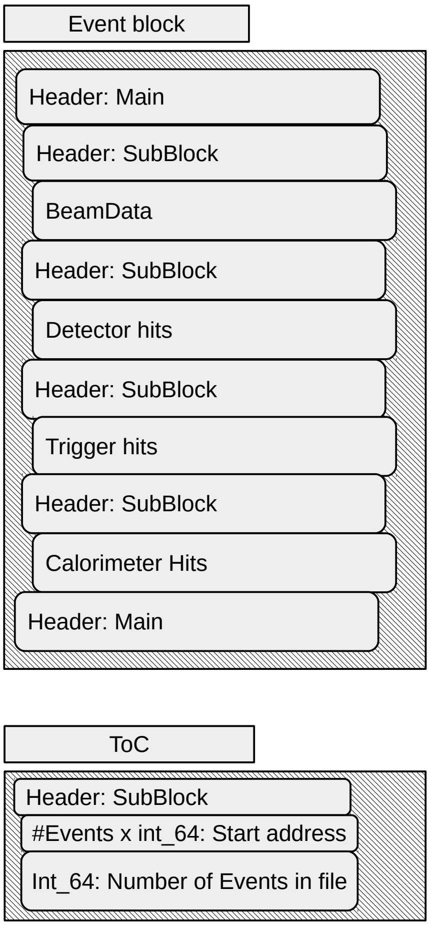


Figure G.1: The data layout in the tbin format. The used data blocks are explained in figure G.2. The header words are listed in table G.1.

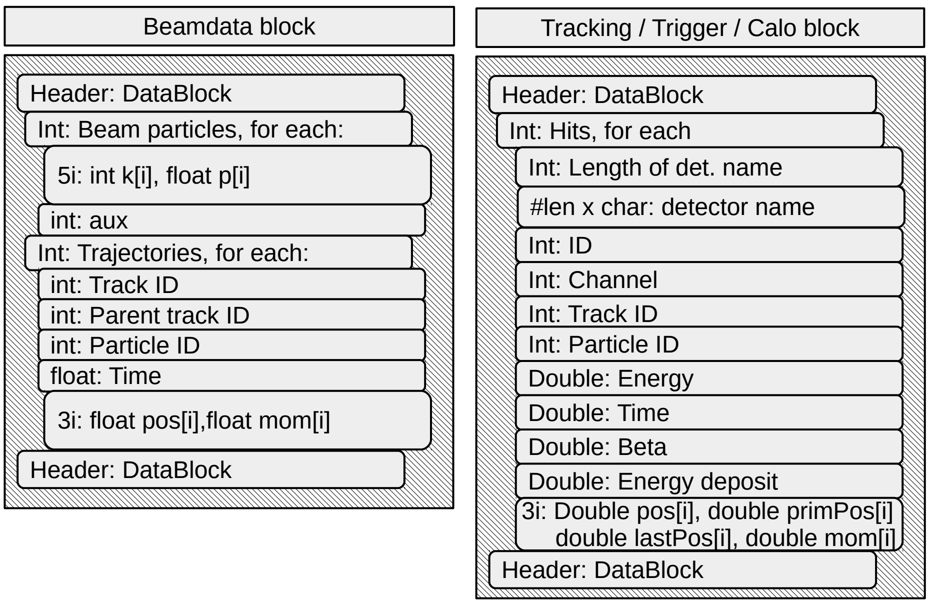


Figure G.2: The used data blocks for the tbin binary pile up format. The header words are listed in table G.1.

List of Figures

List of Tables

Bibliography

[1] D. Mendelejew, “Über die Beziehungen der Eigenschaften zu den Atomgewichten der Elemente”, *Zeitschrift für Chemie* (1869) 405–406.

[2] E. Rutherford, “The scattering of α and β particles by matter and the structure of the atom”, *Phil. Mag.* **21** (1911) 669–688.

[3] J. Chadwick, “The Existence of a Neutron”, *Proceedings of the Royal Society of London. Series A* **136** (1932) 692–708.

[4] H. Yukawa, “On the interaction of elementary particles”, *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57.

[5] D. M. Dennison, “A Note on the Specific Heat of the Hydrogen Molecule”, *Proceedings of the Royal Society of London. Series A* **115** (1927) no. 771, 483–486. `http://www.jstor.org/stable/94849`.

[6] L. S. Ornstein, W. Kapuscinski, and J. G. Eymers, “Intensity Measurements in the Secondary Spectrum of Hydrogen”, *Proceedings of the Royal Society of London. Series A* **119** (1928) 83–91.

[7] M. Gell-Mann, “A Schematic Model of Baryons and Mesons”, *Phys. Lett.* **8** (1964) 214–215.

[8] E. D. Bloom *et al.*, “High-Energy Inelastic $e - p$ Scattering at 6° and 10°”, *Phys. Rev. Lett.* **23** (1969) 930–934. `http://link.aps.org/doi/10.1103/PhysRevLett.23.930`.

[9] P. Söding, “On the discovery of the gluon”, *Eur. Phys. J. H* **35** (2010) no. 1, 3–28. `http://dx.doi.org/10.1140/epjh/e2010-00002-5`.

[10] Nobel Media AB, “The Nobel Prize in Physics 1969”, (2014). `https://www.nobelprize.org/nobel_prizes/physics/laureates/1969/`.

[11] R. Jaffe and A. Manohar, “The g_1 problem: Deep inelastic electron scattering and the spin of the proton”, *Nucl. Phys. B* **337** (1990) no. 3, 509 – 546. `http://www.sciencedirect.com/science/article/pii/0550321390905069`.

[12] K. A. Olive, “Review of Particle Physics”, *Chin. Phys. C* **40** (2016) .

[13] F. Close and R. Roberts, “Consistent analysis of the spin content of the nucleon”, *Phys. Lett. B* **316** (1993) no. 1, 165 – 171. `http://www.sciencedirect.com/science/article/pii/0370269393906736`.

[14] **COMPASS** Collaboration, C. Adolph *et al.*, “Leading and next-to-leading order gluon polarization in the nucleon and longitudinal double spin asymmetries from open charm muoproduction”, *Phys. Rev. D* **87** (2013) . `http://link.aps.org/doi/10.1103/PhysRevD.87.052018`.

[15] **COMPASS** Collaboration, C. Adolph *et al.*, “Leading order determination of the gluon polarisation from DIS events with high-p_T hadron pairs”, *Phys. Lett. B* **718** (2013) no. 3, 922 – 930. `http://www.sciencedirect.com/science/article/pii/S0370269312012221`.

[16] **COMPASS** Collaboration, C. Adolph *et al.*, “Leading-order determination of the gluon polarisation using a novel method”, `arXiv:1512.05053 [hep-ex]`.

[17] X. Ji, “Gauge-Invariant Decomposition of Nucleon Spin”, *Phys. Rev. Lett.* **78** (1997) no. 4, 610–613. `http://dx.doi.org/10.1103/PhysRevLett.78.610`.

[18] B. Povh, K. Rith, C. Scholz, and F. Zetsche, “Teilchen und Kerne. Eine Einführung in die physikalischen Konzepte”, Springer-Verlag GmbH, sixth ed., 2004.

[19] M. N. Rosenbluth, “High Energy Elastic Scattering of Electrons on Protons”, *Phys. Rev.* **79** (1950) 615–619. `http://link.aps.org/doi/10.1103/PhysRev.79.615`.

[20] L. N. Hand, D. G. Miller, and R. Wilson, “Electric and Magnetic Form Factors of the Nucleon”, *Rev. Mod. Phys.* **35** (1963) 335–349. `http://link.aps.org/doi/10.1103/RevModPhys.35.335`.

[21] R. Devenish and R. Amanda Cooper-Sarkar, “Deep Inelastic Scattering”, Oxford University Press, 2004.

[22] M. Kaku, “Quantum field theory: A modern introduction”, Oxford University Press, New York, 1993.

[23] **Particle Data Group** Collaboration, J. Beringer *et al.*, “Review of Particle Physics*”, *Phys. Rev. D* **86** (2012) . `http://link.aps.org/doi/10.1103/PhysRevD.86.010001`.

[24] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language”, *Nucl. Phys. B* **126** (1977) no. 2, 298 – 318. `http://www.sciencedirect.com/science/article/pii/0550321377903844`.

[25] H. Wollny, “Measuring azimuthal asymmetries in semi-inclusive deep-inelastic scattering off transversely polarized protons”, PhD thesis, Freiburg, (2010). `http://cds.cern.ch/record/1287595`.

[26] M. Anselmino, A. Efremov, and E. Leader, “The theory and phenomenology of polarized deep inelastic scattering”, *Physics Reports* **261** (1995) no. 1–2, 1 – 124. `http://www.sciencedirect.com/science/article/pii/0370157395000115`.

[27] **EMC** Collaboration, 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. B* **206** (1988) 340–346.

[28] R. L. Jaffe, H. Meyer, and G. Piller, “Spin, twist and hadron structure in deep inelastic processes”, pp. 178–249. Springer Berlin Heidelberg, Berlin, Heidelberg, (1997). `arXiv:hep-ph/9602236 [hep-ph]`. `http://dx.doi.org/10.1007/BFb0105860`.

[29] R. Carlitz, J. Collins, and A. Mueller, “The role of the axial anomaly in measuring spin-dependent parton distributions”, *Phys. Lett. B* **214** (1988) no. 2, 229 – 236. `http://www.sciencedirect.com/science/article/pii/0370269388914748`.

[30] X. Ji, “Proton Tomography Through Deeply Virtual Compton Scattering”, (2016). `https://inspirehep.net/record/1455791/files/arXiv:1605.01114.pdf`.

[31] X. Ji, “Deeply virtual Compton scattering”, *Phys. Rev. D* **55** (1997) 7114–7125. `http://link.aps.org/doi/10.1103/PhysRevD.55.7114`.

[32] A. V. Belitsky, D. Mueller, and A. Kirchner, “Theory of deeply virtual Compton scattering on the nucleon”, *Nucl. Phys. B* **629** (2002) 323–392, `arXiv:hep-ph/0112108 [hep-ph]`.

[33] M. Diehl, “Generalized parton distributions”, *Physics Reports* **388** (2003) no. 2–4, 41 – 277. `http://www.sciencedirect.com/science/article/pii/S0370157303003338`.

[34] K. Goeke, M. V. Polyakov, and M. Vanderhaeghen, “Hard exclusive reactions and the structure of hadrons”, *Prog. Part. Nucl. Phys.* **47** (2001) 401–515, `arXiv:hep-ph/0106012 [hep-ph]`.

[35] S. V. Goloskokov and P. Kroll, “Transversity in hard exclusive electroproduction of pseudoscalar mesons”, *Eur. Phys. J. A* **47** (2011) no. 9, 112. `http://dx.doi.org/10.1140/epja/i2011-11112-6`.

[36] M. Burkardt, C. A. Miller, and W. D. Nowak, “Spin-polarized high-energy scattering of charged leptons on nucleons”, *Rept. Prog. Phys.* **73** (2010) 016201, `arXiv:0812.2208 [hep-ph]`.

[37] M. Burkardt, “Impact parameter space interpretation for generalized parton distributions”, *Int. J. Mod. Phys. A* **18** (2003) 173–208, `arXiv:hep-ph/0207047 [hep-ph]`.

[38] **COMPASS** Collaboration, F. Gautheron *et al.*, “COMPASS-II Proposal”, *SPSC-P-340, CERN-SPSC-2010-014* (2010) . `http://wwwcompass.cern.ch/compass/proposal/compass-II_proposal/compass-II_proposal.pdf`.

[39] A. V. Belitsky, X. Ji, and F. Yuan, “Quark imaging in the proton via quantum phase space distributions”, *Phys. Rev. D* **69** (2004) , `arXiv:hep-ph/0307383 [hep-ph]`.

[40] H. Moutarde, “Extraction of the Compton Form Factor H from DVCS measurements at Jefferson Lab”, *Phys. Rev. D* **79** (2009) , `arXiv:0904.1648 [hep-ph]`.

[41] C. E. Hyde, M. Guidal, and A. V. Radyushkin, “Deeply Virtual Exclusive Processes and Generalized Parton Distributions”, *J. Phys. Conf. Ser.* **299** (2011) , `arXiv:1101.2482 [hep-ph]`.

[42] A. V. Belitsky, D. Mueller, L. Niedermeier, and A. Schaefer, “Leading twist asymmetries in deeply virtual Compton scattering”, *Nucl. Phys. B* **593** (2001) 289–310, `arXiv:hep-ph/0004059 [hep-ph]`.

[43] A. Radyushkin, “Exclusive Reactions at High Momentum Transfer IV: Proceedings of the 4th Workshop”, World Scientific Publishing Company, 2011.

[44] **Jefferson Lab Hall A** Collaboration, M. Mazouz *et al.*, “Deeply virtual compton scattering off the neutron”, *Phys. Rev. Lett.* **99** (2007) 242501, `arXiv:0709.0450 [nucl-ex]`.

[45] S. V. Goloskokov and P. Kroll, “An Attempt to understand exclusive π^+ electroproduction”, *Eur. Phys. J. C* **65** (2010) 137–151, `arXiv:0906.0460 [hep-ph]`.

[46] S. V. Goloskokov, “Cross sections and spin asymmetries in vector meson leptoproduction”, `arXiv:0910.4308 [hep-ph]`. `https://inspirehep.net/record/834883/files/arXiv:0910.4308.pdf`.

[47] M. Diehl and S. Sapeta, “On the analysis of lepton scattering on longitudinally or transversely polarized protons”, *Eur. Phys. J. C* **41** (2005) 515–533, `arXiv:hep-ph/0503023 [hep-ph]`.

[48] A. Bacchetta, U. D’Alesio, M. Diehl, and C. A. Miller, “Single-spin asymmetries: The Trento conventions”, *Phys. Rev. D* **70** (2004) , `arXiv:hep-ph/0410050 [hep-ph]`.

[49] K. Schmidt, “Transverse target spin asymmetries in exclusive vector meson muoproduction.”, PhD thesis, Freiburg, (2014). `https://freidok.uni-freiburg.de/data/9893`.

[50] P. Kroll, “A set of generalized parton distributions”, *Nuovo Cim. C* **036** (2013) no. 05, 167–172, `arXiv:1303.6433 [hep-ph]`.

[51] A. V. Radyushkin, “Double distributions and evolution equations”, *Phys. Rev. D* **59** (1999) , `arXiv:hep-ph/9805342 [hep-ph]`.

[52] M. V. Polyakov and A. G. Shuvaev, “On ’dual’ parametrizations of generalized parton distributions”, `arXiv:hep-ph/0207153 [hep-ph]`.

[53] **COMPASS** Collaboration, P. Abbon *et al.*, “The COMPASS experiment at CERN”, *Nucl. Instrum. Meth. A* **577** (2007) 455–518, `arXiv:hep-ex/0703049 [hep-ex]`.

[54] **COMPASS** Collaboration, P. Abbon *et al.*, "The COMPASS setup for physics with hadron beams", *Nucl. Instrum. Meth. A* **779** (2015) 69 – 115. http://www.sciencedirect.com/science/article/pii/S0168900215000662.

[55] F. Marcastel, "Schematics of the CERN accelerator complex.", http://cds.cern.ch/record/1621583/files/.

[56] S. Sirtl, "Azimuthal asymmetries in semi-inclusive deep-inelastic hadron muoproduction on longitudinally polarized protons.", PhD thesis, Freiburg, (2016). https://freidok.uni-freiburg.de/data/11139.

[57] E. Bielert *et al.*, "A 2.5 m long liquid hydrogen target for COMPASS", *Nucl. Instrum. Meth. A* **746** (2014) 20 – 25. http://www.sciencedirect.com/science/article/pii/S0168900214001363.

[58] N. Anfimov, V. Anosov, *et al.*, "Shashlyk EM calorimeter prototype readout by MAPD with superhigh pixel density for COMPASS-II ", *Nucl. Instrum. Meth. A* **718** (2013) 75 – 77. http://www.sciencedirect.com/science/article/pii/S0168900212014519.

[59] C. Bernet *et al.*, "The COMPASS trigger system for muon scattering", *Nucl. Instrum. Meth. A* **550** (2005) no. 1–2, 217 – 240. http://www.sciencedirect.com/science/article/pii/S0168900205012568.

[60] M. Bodlak *et al.*, "New data acquisition system for the COMPASS experiment", *Journal of Instrumentation* **8** (2013) no. 02, C02009. http://stacks.iop.org/1748-0221/8/i=02/a=C02009.

[61] **COMPASS** Collaboration, A. Lednev, "Shower separation program for ECAL2", *COMPASS Note* (2009) . https://wwwcompass.cern.ch/compass/notes/2009-7/2009-7.pdf.

[62] A. Lednev, "Electron shower transverse profile measurement", *Nucl. Instrum. Meth. A* **366** (1995) no. 2, 292 – 297. http://www.sciencedirect.com/science/article/pii/0168900295007172.

[63] S. Landgraf, "Luminosity Calculation for the COMPASS Experiment using the F_2 Structure Function.", Master's thesis, Freiburg, 2015. http://hpfr03.physik.uni-freiburg.de/arbeiten/diplomarbeiten/landgraf_master_2015.pdf.

[64] A. Sandacz and P. Sznajder, "HEPGEN - generator for hard exclusive leptoproduction", arXiv:1207.0333 [hep-ph].

[65] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[66] G. Ingelman, A. Edin, and J. Rathsman, "LEPTO 6.5: A Monte Carlo generator for deep inelastic lepton - nucleon scattering", *Comput.Phys.Commun.* **101** (1997) 108–134, arXiv:hep-ph/9605286 [hep-ph].

[67] F. James, "RANLUX: A FORTRAN implementation of the high quality pseudorandom number generator of Luscher", *Comput. Phys. Commun.* **79** (1994) 111–114. [Erratum: *Comput. Phys. Commun.* **97**, 357(1996)].

[68] L. Lönnblad, "CLHEP: a project for designing a C^{++} class library for high energy physics", *Comput. Phys. Commun.* **84** (1994) no. 1, 307–316. `https://cds.cern.ch/record/259545`.

[69] B. Berthou *et al.*, "PARTONS: PARtonic Tomography Of Nucleon Software. A computing platform for the phenomenology of Generalized Parton Distributions", `arXiv:1512.06174`.

[70] A. Sandacz, "Modifications to FFS model and predictions", `https://wwwcompass.cern.ch/compass/gpd/meetings/200904_april22/AS_gpd-Apr2009.ppt`.

[71] L. L. Frankfurt, A. Freund, and M. Strikman, "Diffractive exclusive photon production in DIS at DESY HERA", *Phys. Rev. D* **58** (1998) . `http://link.aps.org/doi/10.1103/PhysRevD.58.114001`. [Erratum: *Phys. Rev. D* **59**, no.11, 119901 (1999)].

[72] S. V. Goloskokov and P. Kroll, "Transversity in hard exclusive electroproduction of pseudoscalar mesons", *Eur. Phys. J. A* **47** (2011) 112, `arXiv:1106.4897 [hep-ph]`.

[73] **NMC** Collaboration, M. Arneodo *et al.*, "Exclusive ρ^0 and ϕ muoproduction at large Q^2", *Nucl. Phys. B* **429** (1994) no. 3, 503 – 529. `http://www.sciencedirect.com/science/article/pii/055032139490152X`.

[74] P. Kroll, "Cross section table for exclusive ρ^+ production", (2014). private communication.

[75] L. Mossé and N. d'Hose, "Etude de la diffusion Compton virtuelle dans le régime profondément inélastique pour le dispositif expérimental COMPASS", PhD thesis, Orsay, (2002). `https://cds.cern.ch/record/1488973`.

[76] P. Guichon and A. Vidon, "Bethe-Heitler calculation", (2015). private communication.

[77] J. D. Jackson, "Remarks on the phenomenological analysis of resonances", *Il Nuovo Cimento* **34** (1964) no. 6, 1644–1666. `http://dx.doi.org/10.1007/BF02750563`.

[78] **ZEUS** Collaboration, M. Derrick *et al.*, "Measurement of elastic ρ^0 photoproduction at HERA", *Z. Phys. C* **69** (1995) 39–54, `arXiv:hep-ex/9507011 [hep-ex]`.

[79] D. Schildknecht, G. A. Schuler, and B. Surrow, "Vector meson electroproduction from generalized vector dominance", *Phys. Lett. B* **449** (1999) 328–338, `arXiv:hep-ph/9810370 [hep-ph]`.

[80] **HERMES** Collaboration, A. Airapetian *et al.*, “Exclusive leptoproduction of ρ^0 mesons from hydrogen at intermediate virtual photon energies”, *Eur. Phys. J. C* **17** (2000) 389–398, `arXiv:hep-ex/0004023 [hep-ex]`.

[81] F. Niecknig, B. Kubis, and S. P. Schneider, “Dispersive analysis of $\omega \to \pi$ and $\phi \to 3\pi$ decays”, *Eur. Phys. J. C* **72** (2012) , `arXiv:1203.2501`.

[82] B. Marianski, “Extraction of R for exclusive ω production”, (2016). private communication.

[83] **ZEUS** Collaboration, S. Chekanov *et al.*, “Exclusive electroproduction of J/psi mesons at HERA”, *Nucl. Phys. B* **695** (2004) 3–37, `arXiv:hep-ex/0404008 [hep-ex]`.

[84] L. Frankfurt, A. Freund, and M. Strikman, “Deeply virtual compton scattering at HERA – A probe of asymptotia”, *Phys. Lett. B* **460** (1999) no. 3–4, 417 – 424. `http://www.sciencedirect.com/science/article/pii/S0370269399008035`.

[85] 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. D* **60** (1999) . `http://link.aps.org/doi/10.1103/PhysRevD.60.094017`.

[86] P. A. M. Guichon and M. Vanderhaeghen, “Virtual Compton scattering off the nucleon”, *Prog. Part. Nucl. Phys.* **41** (1998) 125–190, `arXiv:hep-ph/9806305 [hep-ph]`.

[87] A. V. Radyushkin, “Symmetries and structure of skewed and double distributions”, *Phys. Lett. B* **449** (1999) 81–88, `arXiv:hep-ph/9810466 [hep-ph]`.

[88] I. V. Musatov and A. V. Radyushkin, “Evolution and models for skewed parton distributions”, *Phys. Rev. D* **61** (2000) 074027, `arXiv:hep-ph/9905376 [hep-ph]`.

[89] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, “Parton distributions: A New global analysis”, *Eur. Phys. J. C* **4** (1998) 463–496, `arXiv:hep-ph/9803445 [hep-ph]`.

[90] E. Leader, A. V. Sidorov, and D. B. Stamenov, “Polarized parton densities in the nucleon”, *Phys. Rev. D* **58** (1998) 114028. `http://link.aps.org/doi/10.1103/PhysRevD.58.114028`.

[91] M. Penttinen, M. V. Polyakov, and K. Goeke, “Helicity skewed quark distributions of the nucleon and chiral symmetry”, *Phys. Rev. D* **62** (2000) 014024. `http://link.aps.org/doi/10.1103/PhysRevD.62.014024`.

[92] K. Goulianos, “Diffractive interactions of hadrons at high energies”, *Physics Reports* **101** (1983) no. 3, 169 – 219. `http://www.sciencedirect.com/science/article/pii/0370157383900108`.

[93] C. Regali and T. Szameitat, "TGEANT | A Geant4 based monte carlo simulation for the COMPASS-II experiment", *Resources and Forums for TGEANT* . `http://na58-project-tgeant.web.cern.ch/`.

[94] C. Regali, "HEPGen++ – An exclusive event generator for the COMPASS-II experiment", *Doxygen source code documentation* . `https://cregali.web.cern.ch/cregali/hepgen-plusplus/documentation/`.

[95] P. Kroll, "Twist-3", (2015). private communication.

[96] P. Kroll, "Equations for the GK11 model for π^0 production", (2016). private communication.

[97] IEEE Task P754, "ANSI/IEEE 754-1985, Standard for Binary Floating-Point Arithmetic", IEEE, New York, NY, USA, 1985. `http://ieeexplore.ieee.org/iel1/2355/1316/00030711.pdf`. Revised 1990. A preliminary draft was published in the January 1980 issue of IEEE Computer, together with several companion articles.

—

Acknowledgements

First and foremost I would like to express my sincere gratitude to my advisor Prof. Horst Fischer. He has been continually supporting me since I first started with my state exam in his group. His patience, knowledge and motivation have been priceless for my Ph.D. thesis, as well as the related studies. This great support over all these years made the thesis what it is today.

I also want to thank Prof. Königsmann dearly for the opportunity to join his working group and to become part of the COMPASS collaboration.

My sincere thanks also goes to Tobias Szameitat for being a great colleague and friend over the last years. Many stimulating discussions were held and a lot of innovations were made because of his input. He also proofread my work and was always ready to listen in difficult situations, which is why he is a true T-Bro.

I want to thank all of the people helping with the further proofreading of this thesis. Matthias Gorzellic, Riccardo Longo, Judy Simon, Tristan Schindler, Stefan Sirtl and especially Ana Martic molded my first versions into an enjoyable read.

All of the Freiburg group and the COMPASS collaboration have contributed immensely to my personal and professional growth. You were always a source of good advice and collaboration, which I appreciate a lot. I could not have hoped for a better working environment to pursue my scientific dreams.

I am very lucky to have great friends in my life, who have always supported me in difficult periods through various ways. Many thanks to Christian Dreher, David Kalb and Pirmin Stüble for answering calls at any time and cheering me up whenever I needed it most. My time in Freiburg was also very enjoyable because of the many friends that became part of my life here. I am not only grateful for the time spent with friends from the ExitAsia gym, but also for the lessons I learned: "Perseverance always prevails" (N. Loton).

Finally, a special thanks goes to my family. Words cannot express how grateful I am to my mother and my father for the sacrifices that you've made on my behalf. Your support sustained my endeavours and always will. I would also like to thank my sisters and grandmother for their support throughout this long time. At the end I would like to express my gratitude and appreciation to my girlfriend Verena Schindler, who not only helped with proofreading, but also spent many sleepless nights with me, supporting me wherever she could and fixing any occurring problems during these busy times.

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe.

Freiburg, den 25. Oktober 2016