Roland Kuhn

GLUON POLARIZATION IN THE NUCLEON

Roland Kuhn

Gluon Polarization in the Nucleon

A Measurement of Spin Asymmetries in Quasi-Real Photo-Production of Hadrons with High Transverse Momentum at COMPASS

HARLAND media

Bibliographische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Roland Kuhn
E-Mail: rk@rkuhn.info

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Die Dissertation wurde am 24.7.2007 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 25.9.2007 angenommen.

Gedruckt auf alterungsbeständigem Papier nach ISO 9760 (säure-, holz- und chlorfrei).

Printed in Germany

ISBN 978-3-938363-21-8

Abstract

During 2002–2004, the COMPASS experiment at the CERN SPS has recorded 1.5 fb^{-1} of deep inelastic scattering events with polarized muon beam and polarized deuterium target. The cross section for single-inclusive charged hadron production in dependence on the hadron's transverse momentum p_T is extracted. A PYTHIA and GEANT simulation is used to obtain the acceptance correction factors. The double spin asymmetry $A_{\mathrm{LL}}^{h,d}$ is measured in the region 1 GeV/$\mathbf{c}$ $< p_T <$ 3.5 GeV/$\mathbf{c}$, the p_T dependence of which is connected to the gluon polarization ΔG.

Kurzfassung

In den Strahlzeiten 2002–2004 hat das COMPASS-Experiment am CERN SPS eine Luminosität von etwa $1.5\,\mathrm{fb}^{-1}$ in tiefinelastischer polarisierter Myonstreuung mit polarisiertem Deuteron-Target aufgezeichnet. Der Wirkungsquerschnitt für die Produktion geladener Hadronen wurde in Abhängigkeit von deren Transversalimpuls p_T gemessen. Zur Bestimmung der Akzeptanzkorrektur wurde eine Monte-Carlo-Simulation verwendet, basierend auf PYTHIA und GEANT. Die Asymmetrie der Wirkungsquerschnitte mit Strahl- und Targetpolarisation parallel beziehungsweise antiparallel, $A_{\mathrm{LL}}^{h,d}$, wurde im Bereich $1\,\mathrm{GeV}/c < p_T < 3.5\,\mathrm{GeV}/c$ gemessen. Diese hängt von der Polarisation der Gluonen im Nukleon ab.

Contents

List of Figures

List of Tables

Acronyms and Abbreviations

ADC Analog to Digital Converter

ALICE heavy ion scattering experiment at the LHC

ANSI American National Standards Institute

APV Analog Pipeline readout chip originally designed for micro-strip gas chambers, used for silicon, GEM and RICH readout in COMPASS and other experiments

BMS Beam Momentum Station, a set of detectors arranged around a bending magnet in the beam line which is used for measuring the beam particle's momentum

CASTOR CERN Advanced STORage manager

CATCH COMPASS Accumulate, Transfer and Control Hardware

CDA Distance of Closest Approach

CERN formerly Conseil Européenne pour la Recherche Nucléaire, now European Laboratory for Particle Physics, situated in Geneva, Switzerland

CDR Central Data Recording

CHEOPS CHarm Experiment with OmniPurpose setup

CINDERELLA originally just a reference to the fairy tale, but retrofitted with an acronym by Jan Friedrich: *Command INterpreter for Data and Event REduction by Low Level Analysis* was the name of a Transputer-based experiment control and data analysis project at the Mainz Microtron MAMI.

COMGEANT COMpass adaption of the GEANT package

COMPASS COmmon Muon and Proton Apparatus for Structure and Spectroscopy

CORAL COmpass Reconstruction and Analysis Library

CPU Central Processing Unit of a computer

CRON Command Run ON, a scheduler for recurring system tasks on Linux (inherited from Unix)

DAQ Data AcQuisition system

DCS Detector Control System, in case of COMPASS realized using the PVSS framework

DESY Deutsches Elektronensynchrotron, Hamburg, Germany

DGLAP Dokshitzer, Gribov, Lipatov, Altarelli, Parisi [1, 2, 3, 4]

DMA Direct Memory Access, meaning that computer components other than the CPU access the main memory

DIM Distributed Information Management, a library developed at CERN, see [5]

DIS Deep Inelastic Scattering

DNS Domain Name Service, the internet service which resolves human readable computer names into internet protocol numeric addresses (and more)

DST Data Summary Tapes—also Data Summary Tables—stemming from the historical processing steps where the raw data are reconstructed to yield tracks, vertexes, and so on, the result of which takes much less space than the original data and thus is stored on tapes which "summarize" the raw data

ECAL Electromagnetic CALorimeter

EMC European Muon Collaboration

EPS Encapsulated PostScript, a special form of PS which follows a defined pattern to allow embedding into other documents

FPGA Field Programmable Gate Array, for more details see [6]

FLT First Level Trigger, usually implemented in hardware using specialized detector readout

GAMS GAMma Spectrometer, now used mainly as ECAL2 of the COMPASS experiment

GEANT detector description and simulation tool developed at CERN

GEM Gas Electron Multiplier, see [7]

GeSiCA GEM and Silicon Control and Acquisition module

GIF Graphics Interchange Format

HI deep inelastic scattering experiment at HERA

HCAL Hadronic CALorimeter

HERA Hadron Electron Ring Accelerator at the DESY laboratory in Hamburg, Germany

HotGeSiCA GeSiCA with new HotLink interface

HMC Hadron Muon Collaboration

IEC International Electrotechnical Commission

IEEE Institute of Electrical and Electronics Engineers

JPG Joint Picture Group graphics interchange file format

LEAR Low Energy Antiproton storage Ring

LED Light Emitting Diode

LEP Large Electron–Position collider

LEPTO a Monte Carlo program to produce deep inelastic lepton–nucleon scattering events, part of the LUND package

LHC Large Hadron Collider at CERN, Geneva, Switzerland

LO Leading Order in an expansion series

LUND package of Monte Carlo programs developed at Lund University, Sweden, which features a special model for jet fragmentation

mDST Mini Data Summary Tapes, a further compression step in the data analysis chain where basic selection criteria are already applied to the sample and original hit information is suppressed

MICROMEGAS Micromesh Gaseous Structure detectors

MW Muon Wall, a sandwich of tracking detectors with a hadron absorber in between

MWPC Multi-Wire Proportional Chamber

MYSQL database engine implementing the SQL standard

NLO Next-to-Leading Order in an expansion series

NMC New Muon Collaboration

OMEGA superconducting spectrometer magnet provided by CERN which was used in many experiments in the West Area, including WA89, see [8]

PAW Physics Analysis Workstation, predecessor of ROOT

PCI Peripheral Component Interconnect, the standard PC expansion slot technology since around 1994

PDF Portable Document Format *or* Parton Distribution Function

PERL Practical Extraction and Report Language, originally designed for automated text processing

PETRA Positron Electron Tandem Ring Accelerator, predecessor of HERA at DESY

PHAST COMPASS PHysics Analysis Software Tools, produces and processes MDST files

PNG Portable Network Graphics

PS PostScript, a stack-based programming language which is used by printers for page description, basis for the simplified PDF

PVSS Prozeßvisualisierungs- und Steuerungssystem (process visualization and control system)

PYTHIA a Monte Carlo program for high p_T physics in hadronic interactions, part of the LUND package

QCD Quantum ChromoDynamics

QED Quantum ElectroDynamics

RHIC Relativistic Heavy Ion Collider at the Brookhaven National Laboratory, NY, USA

RICH Ring Imaging CHerenkov detector, used in particle identification

RMS Root Mean Squared, the square root of the variance of a distribution

ROOT Rene's Object Oriented Toolkit, successor of PAW

RPC Remote Procedure Call, used in various protocols like the Network File System, DIM and others

SIDIS Semi-Inclusive Deep Inelastic Scattering

SLAC Stanford Linear Accelerator Center, Menlo Park, CA, USA

S-LINK Simple Link Interface developed at CERN for FIFO-like connection between components of detector readout systems [9]

SMC Spin Muon Collaboration

SPS Super Proton Synchrotron at CERN

SQL Structured Query Language, used in relational databases for manipulating and querying the stored data

SVG Scalable Vector Graphics

TCS Trigger Control System

TDC Time to Digital Converter

TIFF Targa Image File Format

TRAFFIC Track Finding and Fitting in CORAL

μDST Micro Data Summary Tapes, extracted from MDST by applying analysis specific event selection cuts and/or using a more compact storage format by omitting unused information

VITA incorporated, non-profit organization of vendors and users providing its members the ability to develop and promote open technology standards in real-time, modular embedded computing systems

VME VersaModule Eurocard

XML eXtended Markup Language, a subset of Standard Generalized Markup Language

XPM X-windows PixMap, an image file format

Chapter 1

Introduction

While the properties of the electron are very precisely known, its partners in forming the atom—neutron and proton—remain subject to investigations. This is owed to the fact that these particles are not point-like and structureless but consist of a complicated tangle of quarks and gluons, comparable to a ball of wool soaked with glue, albeit only about 10^{-15} m in size. During the 1960's the method of deep inelastic scattering was invented to probe the composition of the nucleon, and since then the understanding of the inner workings of them, driven by the so-called *strong force*, has been greatly enhanced. Experiments at DESY, SLAC and CERN together with neutrino experiments have shown us that quarks and gluons share the nucleon's momentum about equally and that the only quark flavors which show an abundance of quarks over their corresponding anti-quarks are the so-called *up* and *down* types. This is in accordance with a very figurative model called the constituent quark model, where nucleons are built from exactly three constituents which then consequently define all properties of the nucleon. While easy to explain and visualize, this model does not incorporate the gluons; this makes it even more wondrous how successful it has been, and still is for certain properties of the nucleon.

Staying in the picture of the glue soaked ball of wool, it is clear that it is not easy to pick out parts of the nucleon. There is always the question how much glue sticks to the parts which we can extract, and this quite literally is one of the main problems when dealing with the strong force. Our current best model of the strong force—quantum chromodynamics—is believed to describe all effects, but it is unfortunately impossible to exactly evaluate mathematically. Therefore approximations and simplifications are introduced, which allow the formulation of relations like the sum rules presented in chapter 6. Also different models like chiral perturbation are developed, which are tuned to be easier accessible mathematically in the low energy regime, where perturbative QCD calculations are most difficult. While often successful in predicting or describing experimental findings, the models up to now do not enable the calculation of the structure of the nucleon ab initio.

When talking about the structure of the nucleon, the question is often how the integral properties like mass or charge are split among the constituents. One of these properties is the so-called *spin*, angular momentum intrinsic to many particles, quantized in units of Planck's constant $\hbar$. It is known since the Stern-Gerlach experiment in 1922 that the spin of the proton is $\frac{1}{2}\hbar$, allowing the states $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$. This was later verified also for the neutron. Each quark carries a spin of $\frac{1}{2}\hbar$, while the gluons have $1\hbar$. It is obvious that the spins of all the quarks and gluons in the nucleon cannot simply be aligned and add up, but they also cannot be oriented completely at random: in the end the sum needs to equal $\frac{1}{2}\hbar$. To make it even more complicated, also a rotational motion of the nucleon's constituents around its axis would contribute to the sum. But since the mass and charge had been so conveniently attributed in the constituent quark model, it came as a surprise when the SMC experiment measured the contribution of the quark spins to the total nucleon spin to be only about one quarter. With the newest COMPASS results, this value has now climbed to 33%, but it is clear that the nucleon's spin cannot be explained in a simple quark model.

This insight inspired new experiment proposals like that of HMC at CERN, which forms COMPASS together with CHEOPS, an experiment aiming at the complementary field of hadron spectroscopy and using similar experimental techniques. The goal of HMC was to measure the contribution of the gluons' spin to the nucleon directly by selecting only those reactions where not a quark but a gluon is extracted from the nucleon using so-called *open charm production*. This requires a very advanced apparatus, because these reactions are not only rare but also hard to identify. The previous years of data analysis have shown that indeed the statistical accuracy reachable by such a measurement is not satisfactory. Instead it seems preferable to apply a less strict criterion for gluon selection and account for the accepted impurities afterward. The ansatz used already by the SMC experiment is to subtract the other spin dependent effects, scale this intermediate result by the inverse of the fraction of gluon-initiated events and multiply by the so-called *analyzing power*, a quantity which is calculable in the framework of quantum chromodynamics. The fraction of gluon-initiated events is obtained from a Monte-Carlo simulation and depends on the correct description of reaction type mixing in the Monte Carlo generator. These programs in turn are tuned to describe all data sufficiently well, but they are not based on a single theoretical model and have little interpretative power. Another ansatz has recently been published, which combines previous knowledge on the structure of the nucleon with speculations about the distribution of the gluons' spin, and utilizes quantum chromodynamics in a kinematic regime, where it can be reasonably well calculated. This results in predictions of experimentally measurable quantities, namely the asymmetry between the probability to produce hadrons off the nucleon when its spin is aligned with or against the direction of flight of the probe particle. These predictions are done for different assumed gluon spin distributions, which can then be compared to the experimental results. This is the topic of this thesis.

1.1 Outline

This thesis consists of three parts, of which the first comprises this general introduction and a more detailed presentation of the COMPASS experiment, its apparatus and physics goals. The second part gives insight into the work done by the author in the context of the experiment, namely the COMPASS configuration server, the online filter and the data acquisition system. The last part consists of an in-depth introduction to the underlying theory of deep inelastic scattering and the presentation of the actual analysis together with the results. The interested reader finds additional material on statistics relations, light-cone coordinates and a matching representation of the Dirac algebra, documentation on the various software packages developed in the context of the current thesis and various summary tables in the appendices.

1.2 A Few Words on Notation

Since our alphabet has only 26 characters some of them are used to denote different quantities in different contexts. To avoid ambiguities and allow faster reading, several font shapes distinguish the various meanings:

mathematical constants:	π, e, i
physical constants:	$c, e, \alpha, h, \hbar, \varepsilon_0$
physical units:	cm, GeV, μm
particles:	$p, \pi^-, \mu^+, \overline{\Lambda}^0, \overline{\Omega}_{cc}^-, \overline{\Xi}_{cc}^+$
three-vectors:	$\mathbf{p}$
Lorentz vectors:	q

Abbreviations and acronyms are set in small capitals to distinguish them from regular names. All such occurrences are listed on the pages preceding the introduction with their expansions.

The traditional units for data sizes are imprecise as their exact numerical meaning depends on the context, e. g. 1 MB of computer main memory means actually 1048576 bytes, while 1 GB of hard disk memory corresponds to 10^9 bytes—at least according to the hard disk manufacturer. Therefore in this thesis the following units are used:

1 kiB =	1024 B ≈	1.02 kB
1 MiB =	1048576 B ≈	1.05 MB
1 GiB =	1073741824 B ≈	1.07 GB
1 TiB =	1099511627776 B ≈	1.10 TB
1 PiB =	1125899906842624 B ≈	1.13 PB

Chapter 2

The COMPASS Experiment

COMPASS is an acronym for COmmon Muon and Proton Apparatus for Structure and Spectroscopy. Being a merger of the HMC [10] and CHEOPS [11] projects, COMPASS examines a variety of effects in high energy physics. The original proposal is available in [12], but not only the list of physics goals has grown but also new detector technology has been introduced. COMPASS is a fixed target experiment at the CERN SPS delivered with either secondary hadron beams (π^-, K^- and $\bar{p}$) of up to 290 GeV/c or a tertiary polarized muon beam of 160 GeV/c.

The following section introduces the physics program of both parts of the experiment together with a summary of recent measurements, while the second section concentrates on the apparatus itself. The analysis environment provided by the COMPASS collaboration is the basis of the current thesis and therefore presented in the third section. Parts of the experiment in which the current author has been involved are described in the following chapters in more detail. These are the detector configuration database, the data acquisition system and the online filter.

2.1 COMPASS Physics

COMPASS stands in the tradition of two lines of experiments: The deep inelastic scattering experiments EMC, NMC and SMC lead up to the HMC proposal, and the hadron spectroscopy experiments WA89, WA91, WA102 and several experiments at LEAR were to be followed by CHEOPS. From the two letters of intent it became clear that the demands on the spectrometer were similar for the next generation experiments of both lines, thus CERN encouraged the bundling of effort and the formation of a common collaboration. Therefore the physics program of COMPASS consists of two parts—the *hadron program* and the *muon program*—which will be introduced below. Further details concerning the physics program can be found in the COMPASS proposal [12].

2.1.1 Experiments with the Muon Beam

emc—the European Muon Collaboration—discovered that the structure functions of the nucleon depend on the choice of target nucleus, which is known as the *emc effect* [13]. nmc—the New Muon Collaboration—followed up on this as the systematic uncertainties in the emc measurement were large, owing to the fact that emc had not been designed to measure such an effect; the result [14] confirmed the effect, as did measurements at slac. As of today the emc effect is accepted albeit not completely understood. The second surprise discovered by emc is that only a part of the spin of the proton is carried by quarks [15]. This precipitated the so-called *spin crisis*,[1] which generated great activity among experimental and theoretical physicists, leading amongst others to the formation of smc—the Spin Muon Collaboration—which corroborated the emc observation that the polarization of the quarks inside the proton falls significantly short of the theoretical expectations [16, 17]. compass continues in the tradition of these experiments to further our knowledge on the structure of nucleons and hyperons.

2.1.1.1 Spin Structure of the Nucleon

One candidate for the explanation of the emc spin puzzle is that a sizable fraction of the nucleon's spin is carried by gluons. This would become possible if the spin of these ephemeral constituents was preferentially oriented in parallel to the nucleon's spin. One possible way to infer the magnitude of this effect is to fit all existing data points for nucleon structure functions using the dglap evolution equations, see e. g. [18] which concludes that the polarization of the gluons in the nucleon is about 25%. As this indirectly obtained quantity has a rather large uncertainty, a direct measurement is desirable. compass employs three ansatzes with the production of open charm, high p_T hadron pairs and single high p_T hadrons as is discussed in detail in section 6.6.

As shown by Jaffe and Ji [19, 20] the quark state inside the nucleon at the twist-two level is completely specified by the momentum distributions $q(x)$, the helicity distributions $\Delta q(x)$ and the transverse spin distributions $\Delta_T q(x)$. A measurement of the latter via the Collins effect [21] has been published in [22].

2.1.1.2 Lambda Polarization

Complementary information about the spin structure of the nucleon can be gained by examining the polarization of the strange sea.[2] Longitudinal polarization transferred from the virtual photon in deep inelastic scattering, or transverse polarization taken from transversely polarized nucleons, provide access to the spin density of the strange quarks.

[1] The term is not technically correct since the measurement simply found a very naïve assumption to be incorrect, but it stuck nevertheless.

[2] Since the nucleon consists only of u and d valence quarks the heavier flavors like s and c only are present as vacuum fluctuations. These fluctuations of $q\bar{q}$ pairs are called sea quarks.

Also the spontaneous polarization exhibited in the transverse polarization of Λ^0 baryons produced from unpolarized nucleons are being examined. First results from COMPASS can be found in [23, 24, 25, 26].

2.1.2 Experiments with the Hadron Beam

2.1.2.1 Hadronic Structure

The electromagnetic polarizabilities of the pion have been measured in $\gamma\gamma$ reactions [27, 28] as well as with a pion beam utilizing the Primakoff effect [29, 30, 31] and $\gamma p \to \gamma\pi^+ n$ [32]. The Primakoff measurement disagrees with the former, but was recently corroborated by the latter, thus making a new measurement desirable, which is done using the Primakoff method. Calculations in the framework of chiral perturbation theory [33] find values compatible with the $\gamma\gamma$ experiments. In addition, COMPASS will perform a similar measurement on kaons, for which presently no data exist.

Another field of interest is the chiral anomaly. It is described by an additional term in the chiral Lagrangian and enables parity violating transitions between an even and odd number of mesons in the initial and final states. Better experimental data is needed to test e. g. $O(p^6)$ corrections as computed in [34] to the process $\gamma \to 3\pi$.

2.1.2.2 Exotics

Mesons and baryons are constructed in the standard model from the quantum numbers of two or three quarks, respectively. As not all combinations of quantum numbers are possible in this scheme, particles carrying these combinations—if found—would indicate other types of hadrons, built of more constituents, or even including explicit gluonic degrees of freedom. In general, many color neutral objects are expected to exist, e. g. mesons of the type $q\bar{q}\ q\bar{q}$, $q\bar{q}g$ or $qq\ \bar{q}\bar{q}$, but it is not clear which constellations are bound states.

2.1.2.3 Charmed Hadrons

While charmed mesons are the object of intense studies, little is known about the properties of charmed baryons. For example, of the $J^P\ \frac{3}{2}^+$ 20-plet, the $\bar{\Sigma}_c^0$ and $\bar{\Omega}_c^0$ are still unconfirmed. In analogy to the $D^+ \to D^0\pi^+$ tagging mechanism, transitions like $\bar{\Omega}_c^{*0} \to \bar{\Omega}_c^0\pi^0$ can be used.

Even more interesting is the still untackled issue of doubly-charmed baryons. These lie within the 3.5–4 GeV/c^2 mass region and are challenging as they are hard to produce and difficult to identify.

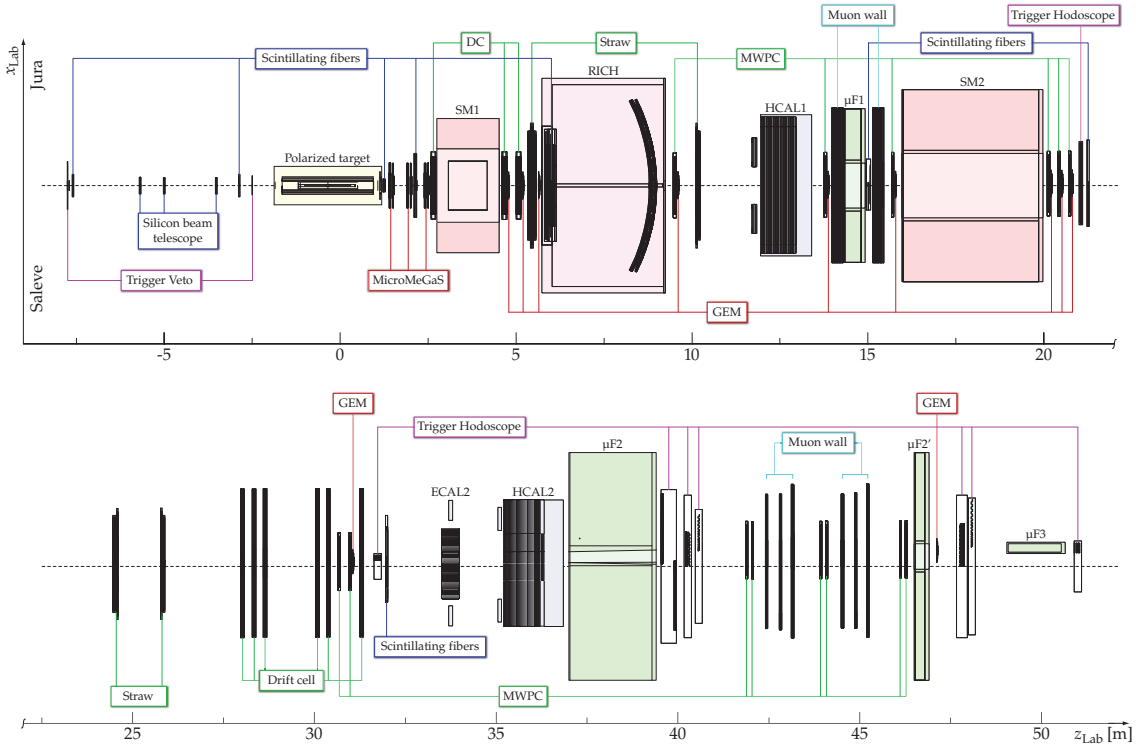


Figure 2.1: Side and top view of the COMPASS spectrometer with the polarized target for the muon program (taken from [23])

2.2 Experimental Setup

Corresponding to the two possible beam types there are two slightly different setups for the COMPASS experiment. However, as this thesis only investigates data taken with the muon beam, the hadron setup is not described here; please refer to [35] for more details.

Compared to the proposal [12] much has changed. The biggest modifications concern the first spectrometer magnet, the addition of a RICH detector and the type of tracking detectors where honeycomb trackers were replaced by gas electron multipliers; for more details see the COMPASS spectrometer paper [36].

2.2.1 Design Overview

COMPASS is a two-stage forward spectrometer with high precision tracking, electromagnetic and hadron calorimetry, particle identification, very low dead time trigger system and an exchangeable target setup. These properties are necessary to efficiently reconstruct particles at small and large angles with respect to the beam axis and interaction rates of up to 50 kHz in deep inelastic muon scattering as well as production reactions with hadron beam. The first spectrometer stage covers large angles of up to 180 mrad and momenta up to 20 GeV/c while the second stage accepts particles within 30 mrad with re-

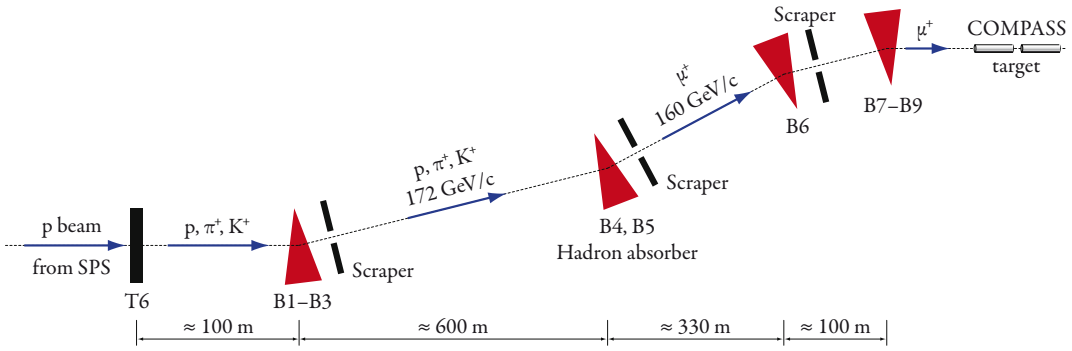


Figure 2.2: Schematic view of the M2 beam line. Bending magnets are grouped and focusing magnets omitted for clarity.

spect to the beam axis and momenta above 5 GeV/c. Each stage is equipped with a hadron calorimeter, the first stage also features a RICH counter for $\pi^{\pm}$–$K^{\pm}$ separation in the momentum region between 5–43 GeV/c and an electromagnetic calorimeter. The absorbing detectors of the first stage have a central hole that matches the acceptance of the second stage. In order to cover a large solid angle while achieving good spatial resolution and rate capability in the vicinity of the beam, the tracking stations are composed of different detector types, with silicon micro-strip or scintillating fiber detectors as the innermost component, GEM or MICROMEGAS covering an area of 30×30 cm^2 around the beam and straw tube trackers, MWPC, drift chambers and large area drift chambers extending up to 5×2.5 m^2 around the beam. Tracking detectors arranged around absorbers—60 cm iron in the first stage, 2.4 m concrete in the second stage of the spectrometer—perform muon identification in the kinematic region covered by the muon trigger system. The readout is designed to be effectively dead time free, utilizing pipelining, local data storage on the detector front-ends and asynchronous trigger signaling, in order to cope with high trigger rates limited only by the reachable data transfer speed. The resulting digitized raw data stream is digested by a large event building network including vast buffers for short term (a few SPS spills à 16.8 s in computer main memory) as well as medium term (1–2 days on disk arrays) temporary storage, from where the data are continuously transferred to permanent storage on the CERN tape libraries.

2.2.2 Beam

COMPASS uses secondary or tertiary beams, which are generated from the primary beam of approximately $1.2 \cdot 10^{13}$ protons per cycle delivered to the T6 production target at an energy of 400 GeV/c. The SPS follows an injection-acceleration-extraction cycle of currently 16.8 s. The protons are extracted after de-bunching during the so-called *on-spill time* of about 4.8 s.

After the T6 target, which constitutes the beginning of the M2 beam line depicted in

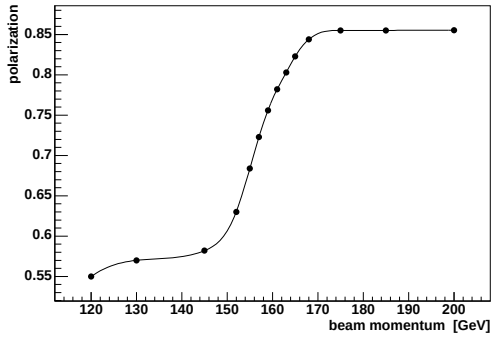


Figure 2.3: Beam polarization as a function of particle momentum in the laboratory frame

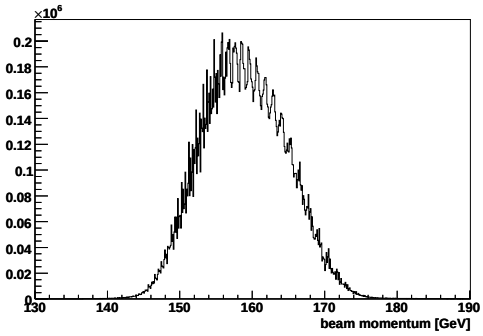


Figure 2.4: Reconstructed beam momentum. As the beam momentum station uses rather coarse-grained tracking, the spectrum shows a partly quantized structure.

figure 2.2, remaining beam protons are partly removed by momentum selection via magnets B1–B3 and collimators. The secondary beam of pions, protons, kaons and their anti-particles then traverses a decay line of $\approx$ 600 m, after which in case of the muon beam a hadron absorber (a beryllium block 9.9 m in length) filters out most of the other particles. The remaining particles then traverse another stage of momentum selection until finally reaching the experimental hall. This setup is flexible enough to allow experiments with polarized muons or charged hadrons of both polarities, the switch between them taking only a few hours.

The polarized muon beam is produced exploiting the parity violation in the weak decay of the pion. The anti-muons produced from the decaying positive pion beam are naturally right-handed due to the parity violation of the weak interaction. However, in the pion rest frame, the helicity of the muon has to be negative so that the spins of ν_μ and μ^+ couple to zero, since the ν_μ is produced with negative handedness, which together with its near-masslessness requires it to have negative helicity. This way the μ^+ which are produced in the forward direction are polarized anti-parallel to the direction of flight in the laboratory system. The polarization increases with decreasing polar angle of the direction of production of the muon, which translates into a rising beam polarization at higher momenta in the laboratory frame as shown in figure 2.3. By selecting positive hadrons of $\sim$ 177 GeV /c momentum at B1–B3 and transporting muons with a momentum around 160 GeV /c via B4–B6 to the hall, a balance between luminosity and polarization is achieved. $1.3 \cdot 10^{13}$ protons impinging on the production target per spill yield $2.3 \cdot 10^{8}$ muons in the momentum range of 150 GeV–170 GeV. Since this range is much larger than the required precision needed for the beam momentum in the reconstruction, a beam momentum station consisting of six fine-grained scintillator hodoscopes arranged around B6 provides a measurement of this quantity for each individual muon with a resolution of better than 1%. The resulting beam momentum spectrum is depicted in figure 2.4.

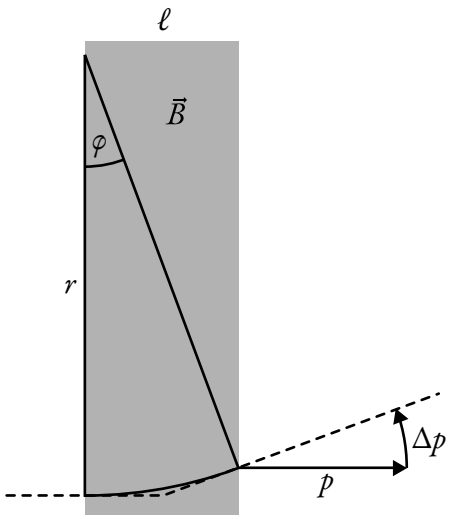


Figure 2.5: A particle traveling through a homogeneous magnetic field is bent on a circular path with radius r. The relation $\Delta p = \phi p$ holds exactly while the field integral being equal to the product of longitudinal dimension and field strength assumes a homogeneous field and small bending angles.

2.2.3 Magnets

A magnetic spectrometer makes use of the Lorentz force to determine the ratio of charge to momentum of a particle. If ℓ is the extent of the field of strength B, q the charge of a particle with momentum $p = \beta\gamma mc$, then the radius r of the circular trajectory of the particle inside the field is

$$\frac{\gamma m \beta^2 c^2}{r} = q\beta c B \qquad \Rightarrow \qquad r = \frac{p}{qB} \tag{2.1}$$

For small angles the deflection $\phi \approx \frac{\ell}{r}$ is anti-proportional to the momentum and proportional to the field integral.[3] The transverse momentum kick experienced by the particle after having traversed the magnetic field then is (see figure 2.5 for an illustration)

$$\Delta p = p\phi = q\ell B = 0.3\,\mathrm{GeV/c} \cdot \frac{\ell B}{1\,\mathrm{T\,m}} \tag{2.2}$$

The last step involved the conversion $1\frac{\mathrm{V}}{\mathrm{m}} = 1\frac{\mathrm{T\,m}}{\mathrm{s}}$, which can be derived by considering the expressions for the electric and magnetic forces acting upon a charge. Interpreting Δp not as a straight vector but the "turning around" of the particle's momentum makes the obtained relation valid also for large angles; if $\phi > \frac{\pi}{2}$ the particle is reflected.

The higher the particle's momentum, the smaller is its deflection in the magnetic field, so a better spatial resolution is needed for small angles to retain a certain momentum

[3] The integrated field strength, $\int B\,d\ell$, sometimes also referred to as *bending power*.

resolution. compass employs two spectrometer magnets whose main field component is directed downward, deflecting particles in the same direction and thereby enlarging the spatial separation of particles with different momenta. This property is used by the trigger system which selects events based mainly on geometrical criteria, including momentum selection and target pointing, see section 2.2.6.

The first spectrometer magnet is located 3.63 m downstream of the target. It is a conventional magnet with a bending power of 1.0 T m and has a gap of 172 cm in height that is opening downstream. This design provides a bigger acceptance and yields a better momentum resolution for particles with a big angle to the beam but also raises the problem of forces between the magnet and its surroundings.

The second spectrometer magnet is also a conventional magnet located 17.83 m downstream of the target. It can be operated with currents of 2000 A to 5000 A, with the standard setting of 4000 A providing a field integral of 4.4 T m. The gap of sm2 is $4 \times 2 \times 1\ \mathrm{m}^3$ (*depth* × *width* × *height*).

2.2.4 Detector Types

For beam definition two types of detectors are used: while $5 \times 7\ \mathrm{cm}^2$ silicon micro-strip detectors with a spatial resolution of around 10 µm define track position and angle, scintillating fiber detectors of $4 \times 4\ \mathrm{cm}^2$ measure the time of track traversal to 0.4 ns.

The tracking stations throughout the spectrometer are composed of three classes of detectors. The smallest, which are the only ones directly exposed to the unscattered beam particles, are scintillating fiber detectors ranging from $5 \times 5\ \mathrm{cm}^2$ to $12 \times 12\ \mathrm{cm}^2$ in size. The medium-sized gem or micromegas stations have an active area of $31 \times 31\ \mathrm{cm}^2$ and $40 \times 40\ \mathrm{cm}^2$ with a circular inactive area in the center so as not to be blinded by the beam. The spatial resolutions are 70 µm and 90 µm, respectively, while the time resolutions are 12 ns and 9 ns. For large area tracking only drift-based detectors are used which rely upon external track timing information to reconstruct the exact hit position. The sizes range from $178 \times 120\ \mathrm{cm}^2$ for the mwpc over $180 \times 127\ \mathrm{cm}^2$ for the drift chambers and $280 \times 323\ \mathrm{cm}^2$ for the straw tube trackers to $500 \times 250\ \mathrm{cm}^2$ for the large area drift chambers with spatial resolutions of 0.2 mm for straw tubes and drift chambers, 0.5 mm for the large area drift chambers and 1.6 mm for the mwpc.

The detectors mentioned up to now represent roughly half of the 200,000 readout channels used in compass. The biggest contributors are gem and mwpc, each around 30%, micromegas and straw tubes follow with ⅛ each while the rest is about evenly split among the remaining types. The data volume generated per event is dominated by rich and gem, which account for half the event size; a breakdown of the individual detector types is given in table 2.1.

The other half of the readout channels is used in particle identification. Here the rich with its 83,000 channels is responsible for most of the data volume. The vessel and mirror

detector type	relative data volume
RICH	24.1%
GEM	22.3%
MICROMEGAS	10.7%
Drift chambers	9.8%
Trigger (scalers, timing)	9.3%
Silicon micro-strip	8.9%
Scintillating fibers	3.7%
Muon walls	2.9%
MWPC	2.8%
Calorimetry	2.3%
Straws	1.9%
BMS	1.1%

Table 2.1: Fraction of data volume produced by the different types of detectors. The values have been extracted from three representative runs of the 2004 beam time.

system are adapted to the angular acceptance of the first spectrometer, 250 mrad in horizontal and 180 mrad in vertical direction. The RICH allows the identification of kaons in the energy range 5–43 GeV at a confidence level of 2.5σ which is crucial in reconstructing D^0 mesons. Each spectrometer stage is equipped with a hadron calorimeter in iron/steel–scintillator sandwich design with energy resolutions of $59\%/\sqrt{E/\,\mathrm{GeV}} \oplus 8\%$ and $66\%/\sqrt{E/\,\mathrm{GeV}} \oplus 5\%$, respectively. The 480 cells of HCAL1 cover $420 \times 300\,\mathrm{cm}^2$ with a hole in the center matching the opening of SM2, HCAL2 is built of 216 modules covering $440 \times 200\,\mathrm{cm}^2$. The second stage additionally features an electromagnetic calorimeter with nearly 3000 cells which covers with its $245 \times 184\,\mathrm{cm}^2$ the region relevant for Primakoff studies. Each of the two hadron calorimeters is followed by a muon identification system, commonly called *muon wall*. MW1 consists of 16 layers of micro drift tube detectors with quadratic cross section, eight before and eight after the 60 cm iron absorber while MW2 is built of six double layers of cylindrical stainless steel drift tubes behind the 2.4 m thick concrete absorber downstream of HCAL2.

2.2.5 Target

The COMPASS target for the muon program is polarized by dynamic nuclear polarization transfer. For the purpose of the explanation let us assume a single deuterium atom inside a large magnetic field. In this environment the spin-spin coupling between electron and nucleus is broken up, creating six energy levels from the three possible spin states of the nucleus and the two spin states of the electron, cf. figure 2.7. These six states divide naturally into a lower and an upper triplet because the magnetic moment of the electron is a factor of ≈1800 greater than the magnetic moment of the deuteron. Let the energy splitting between neighboring states inside each triplet be E_s and the energy needed for an electron spin flip be E_l. The latter is so large that the relaxation from the upper to the

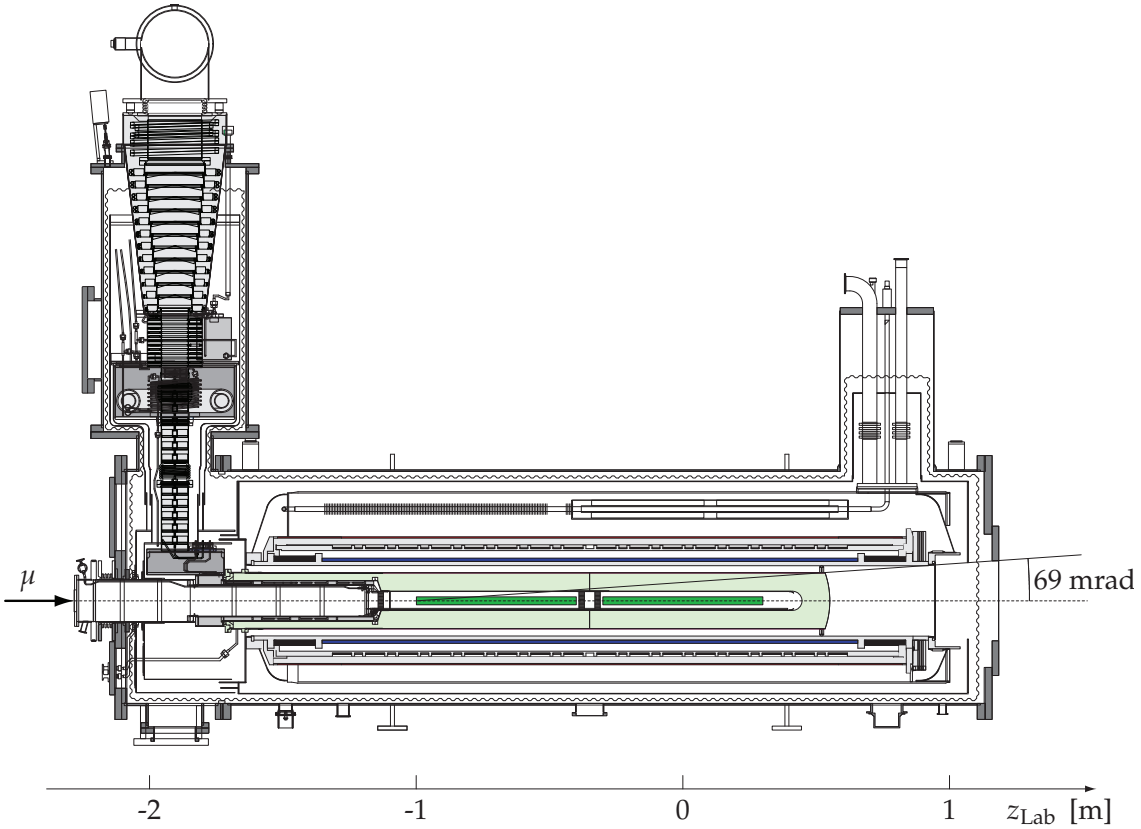


Figure 2.6: Side view of the compass polarized ^{6}LiD target. The components relevant for this analysis are colored: the two target cells in green, surrounded by the microwave cavity in light green and the solenoid in blue plus the dipole magnet in red. (from [37])

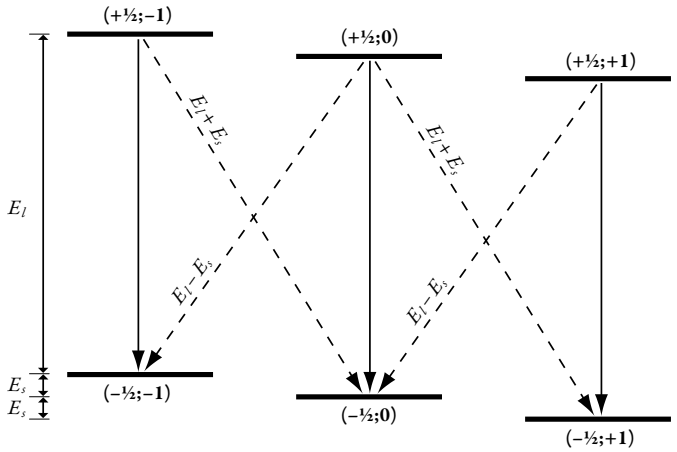


Figure 2.7: Energy level scheme employed in dynamic nuclear polarization. The solid lines represent electron spin flips while the dashed lines correspond to the microwave frequency with which the sample is irradiated, either $E_l + E_s$ for polarization anti-parallel to the magnetic field or $E_l - E_s$ for parallel polarization.

lower triplet is practically instantaneous. This means that an excitation of the atom with the energy $E_l + E_s$ followed by the electron's relaxation will change the deuteron's spin by one unit against the direction of the magnetic field. The opposite can be achieved by exciting the atom with the energy $E_l - E_s$. It is clear that the thermal energy of the system has to be smaller than E_s so that polarization can be built up.

At the magnetic field of 2.5 T created by the COMPASS target solenoid, this means cooling the target material using a ^{3}He/^{4}He dilution refrigerator to below 100 μK. The target setup is shown in figure 2.6, for details see [37]. The target polarization is measured by 5 coils mounted on each target cell and the value given has positive sign if the deuteron spins are polarized in parallel to the magnetic field lines. The polarization build up time is three to four days with a relaxation time > 2000 h at nominal magnetic field. However, in case of a loss of cooling power the polarization is destroyed within minutes.

From the analysis point of view the target consists of two cylindrical cells with a diameter of 3 cm and a length of 60 cm which contain oppositely polarized ^{6}LiD. This design was already chosen by the SMC experiment to minimize systematic errors by measuring both combinations of target and beam spin settings simultaneously. The spin orientations in the two target cells nevertheless need to be flipped frequently to avoid false asymmetries, as discussed in section 7.2.4. The easiest way is to use the so-called frozen spin mode, where the microwave irradiation is stopped and the magnetic field is—with the help of a 0.5 T transverse dipole magnet—adiabatically turned around. This procedure takes about half an hour and is done three times per day. The other method is to exchange the microwave frequencies transferred into the two target cells, thereby inverting the connection between solenoid field and target spins, referred to in the analysis as a change of *microwave setting*. This means building up new polarization in the opposite direction after completely destroying the old polarization, thus it can be done only every few weeks.

During the beam time of 2007, the target material has been replaced by NH_3 (so-called proton target), which in combination with the ^{6}LiD data will allow the extraction of information about the proton and neutron spin structure individually.

2.2.6 Trigger

The trigger system stands at the beginning of the readout chain, defining the point in time when an event occurred. COMPASS uses two trigger levels: the first level trigger (FLT) is implemented in hardware and if a physical reaction in the experiment fits a certain pattern, it initiates the readout of the front-end electronics of the detectors. The FLT discards uninteresting reactions and enriches certain wanted processes based on specialized detectors. The second level trigger is performed by the online filter CINDERELLA, a software which runs on the event building network, see chapter 5 for more details. Since it has the full information about the event at its disposal, the online filter is able to make more elaborate decisions than a hardware trigger would be capable of.

The compass muon trigger consists of four parts. Deep inelastic scattering at large Q^2 means that much more momentum is transferred than energy, leading to large scattering angles and high energy for the muon in the final state. The upper limit in Q^2 accessible by using a hodoscope based muon trigger is given by the acceptance of the second spectrometer magnet and amounts to 60 GeV$^2/c^2$, the lower limit is given by what is meant by *large* when it comes to Q^2, commonly understood as $Q^2 > 0.5$ GeV$^2/c^2$. The hodoscope system selecting muon scattering within these bounds consists of the *middle*, *ladder* and *outer* trigger modules, each having two planes, one of which is located downstream of an absorber, utilizing the great penetration power of muons to identify them while electrons and hadrons are absorbed. The individual hodoscope planes consist of 32 or 64 slabs whose signals are fed into a coincidence matrix together with those of the other matching plane, thereby achieving the selection of muons coming from the target at different angles.

For higher Q^2 the muon leaves the target at an angle that is too large to be covered by hodoscopes, but these reactions always include the production of several hadrons, some of which leave the target and may be registered in the hadron calorimeters. Therefore, in addition to requiring a certain amount of hadronic energy deposit in correlation with the hodoscope triggers, the signals from the hadron calorimeters are also used as a standalone trigger.

The third part of the muon trigger is designed to enrich photon–gluon fusion events. The cross section for this process is anti-proportional to Q^2, thus minimal scattering angles and high energy loss are the attributes of the most interesting events. The trigger system for these requirements is the so-called *inner* trigger, a set of two hodoscope planes situated very closely to the beam at the very end of the compass experiment. The spectrometer magnets with their added bending power of 5.4 T m turn around the momentum of the scattered muons by about 1.6 GeV/c (see section 2.2.3) which translates into a spatial shift at the position of the second plane of the inner trigger of 2 cm at an energy loss of 10 GeV or 17 cm at 60 GeV. This is the reason why this trigger element is the detector of compass which is farthest[4] from the target, yielding the largest possible lever arm.

The last part of the muon trigger is the veto system, which inhibits most triggers while so-called halo muons traverse the spectrometer. The great penetrating power of muons enables them to enter the hall at distances of tens of centimeters to the beam axis, creating signals in tracking detectors and hodoscopes which are not related to an interaction within the target. Therefore large hodoscope panels cover the end of the beam line and smaller scintillator panes are arranged around the beam, permitting only particles not farther than 2 cm from the beam axis. This veto system introduces dead time[5] but it is necessary because otherwise the purity of the hodoscope triggers would be too low and the event rates too high for the data acquisition system.

[4] Neglecting the beam momentum station, which is up to 130 m upstream of the target, while hodoscope HI05 is located 52 m downstream of the target.

[5] Dead time comprises all times at which interactions may occur but cannot be measured by the experiment. The dead time introduced by a veto system is also referred to as *trigger dead time* to distinguish it from the *daq dead time* caused by the finite speed of detector readout.

More information on the COMPASS muon trigger can be found in [38, 39].

2.3 Analysis Environment

Before physics analysis of the data can take place, the rather complex step of reconstruction needs to be done: the unprocessed detector information must be condensed into tracks and vertexes, which are the raw material for the test of any physics hypothesis against the measured data. This effort needs detailed knowledge of the workings of all parts of the apparatus, not only to produce a correct result, but also to come close to the best possible exploitation of the statistics recorded on tape. Because of the importance of this task, the design of the COMPASS Reconstruction and Analysis Library CORAL started shortly after the COMPASS proposal was finished. By that time object oriented programming techniques became widely available in the form of C++ and it was foreseeable that the software development tools would mature fast enough to allow the use of C++ in such mission critical software. The downside of the approach was that established physics analysis libraries like PAW [40] were not available for the new platform, leading to the implementation of e. g. histogramming classes inside CORAL and the export of the reconstructed information in the PAW Ntuple format. In parallel to COMPASS software development, the ROOT package [41] has been developed at CERN, which is the successor of PAW, implemented in C++. As this project reached a quality standard which allowed widespread use among physicists, Sergei Gerassimov conceived PHAST—the PHYSICS ANALYSIS SOFTWARE TOOLS [42]—which allows convenient access to the features of the reconstructed events, bundled with common algorithms to minimize the risk of repeating similar programming errors in many different analyses. PHAST allows the user to concentrate on the analysis of the events of the COMPASS data sample, taking away the tedious job of managing the data flow. At the same time the user interface is well suited for enabling physicists with former knowledge about Fortran to efficiently use PHAST without having to spend a long time learning new principles.

Traditionally, the product of the reconstruction program are so-called DST—data summary tables—from which selections are made for each analysis. Due to the popularity of PHAST, this role is played by MDST (mini data summary tables) in COMPASS, which are ROOT trees[6] containing the PHAST events. These events contain only the reconstructed information on tracks, vertexes, RICH, calorimeter and so on, and only those events are stored which have at least one reconstructed vertex. Combined with the compression employed by the ROOT library this leads to a reduction in data size by a factor $\approx$ 80 relative to the raw data.

The procedure of creating MDST files from the raw data is called production and is done at the CERN batch system, where COMPASS has a share of about 1000 CPU. First a tuned version of the reconstruction library has to be created and verified, then it is linked with

[6] The successor of the PAW Ntuple in ROOT is called a "tree" because of the hierarchical—or branch-like—structure.

Phast, which takes the reconstructed information and stores it event by event in root trees. The tuning of the reconstruction is an iterative process aiming at the best possible utilization of the detector data. It is necessary to adapt the reconstruction process to changes in the apparatus, checking the correctness of track reconstruction and vertexing and optimizing the efficiency in terms of reconstructed statistics. This process has first been exercised on period P2D of 2002, yielding several re-productions of the same raw data until the quality was satisfactory and it has been repeated for the following two beam times of 2003 and 2004, each of course based on a better starting point. The reproductions are indicated in the tables by giving the production slot—a number between one and three—and the Phast major version which has been used to write the mdst. Table E.2 on page 165 summarizes which versions have been used for the present analysis and how much disk space is necessary to store these data.

The smallest set of data typically handled during analysis is one run. Runs are defined during data taking to create aggregations of data for which the recording conditions have been stable. Therefore it was decided to group 200 sps spills—corresponding roughly to one hour of data taking—into one run. It should be mentioned that stopping the old run and starting a new one takes three or four spills which are wasted concerning physics. It turned out that the data quality observed by looking at various detector pseudo-efficiencies and event characteristics varies even during a run on a spill by spill basis. Spills which lie outside a target corridor for the benchmark quantities are collected in the so-called bad spill list. As spill and event numbers are retained on the mdst level, the user of Phast can decide whether to keep or discard the data contained in these spills. The bad spill list is maintained by the compass data stability team and hosted on the compass off-line pages [43].

The mdst are the input for physics analysis and are processed using Phast. For this purpose one or more so-called UserEvent functions are linked to the Phast core, which are then called in turn for each event which is processed. These functions have full access to all reconstructed properties of the event plus a collection of often used algorithms, e. g. for the calculation of x_F, etc. Once an event is identified as being interesting for the analysis at hand it can be saved again in a tree using the Phast format or the user can make use of the full features of root to create his own highly adapted subsample tree. The reason for not processing the mdst directly each time a histogram is added is the pure data size of multiple TiB which means that the disks themselves set a lower bound on the processing time.

The remainder of this section details the most prominent steps done by the reconstruction library. A thorough understanding of the tracking and vertexing is necessary when analyzing single-inclusive hadron production, as most of the event selection criteria rely entirely upon correct reconstruction of the raw data.

2.3.1 Tracking

Tracking means making the association between detector signals and the particles which created those signals. It involves several steps:

- preparation of hit clusters from the raw detector data,
- a pattern search to find track segments,
- bridging of the track segments through magnets and absorbers, and
- the final fit yielding optimal track parameters.

The first step is called *clustering* and it needs another preparatory step, the *decoding*, where the bits describing the data of each channel are extracted from the raw data stream, calibration constants are applied, features like timing and amplitude are extracted and finally a list of *hits* is created. Part of this step is the *mapping* of frontend IDs and channel numbers to detector planes and wire numbers, which is the basis for geometrical reconstruction. The position and orientation of all detector planes is taken from the geometry description file, making the connection between the plane and wire number and the three-dimensional position of a given hit in the global detector coordinate system (called Main Reference System). In this system the beam direction defines the z axis, y points upwards and x completes the right-handed coordinate system. A wire can be imagined as a usually very thin rectangle having the height of the active detector area and being placed at its position inside the detector volume. Hits of neighboring wires are grouped into *clusters* if their features match, i. e. if the confidence level for the hypothesis that the hits were created by the same particle is high enough; this procedure requires detailed knowledge of the function of each detector type and is thus provided by the detector experts.

The next step is to search for track segments in regions of the spectrometer where charged particles are expected to travel approximately along straight lines. For this purpose the apparatus is divided into five zones—before the target, between target and SM1, between the two magnets, between SM2 and the second muon wall and after the second muon wall. In each zone the detector planes are grouped according to the angle between their strips and the y axis, where angles are artificially quantized to integer degrees; each such group is called a projection. Within each projection a simple search for two-dimensional track segments is done using a pivot planes algorithm. The name *pre-pattern* stems from the fact that certain tracks should leave a certain cluster pattern in the traversed detector planes. The algorithm successively selects all pairs of planes, computes all combinations of hits in the two planes and searches all other planes for clusters which would fit within their errors onto the straight line defined by the pair of pivot clusters. Additional freedom is given to clusters coming from regions where the magnetic field cannot entirely be neglected so that no bias is introduced towards higher momenta. The list of track segments found in each zone is cleaned up by three selection steps: A minimum number of

clusters and a track inclination smaller than a maximum angle is demanded. Then the remaining track segments are sorted according to the result of a quality function, which takes into account the number of participating clusters and their χ^2. In the last step the list is processed starting with the best rated track segments, discarding track segments which share too many hits with track segments of higher quality.

The two-dimensional track segment candidates from the different projections are combined to track pieces in space using a similar method as described above: From the set of projections pairs are successively chosen, combining consecutively all track segment pairs which have been found. This yields lines of intersection in three dimensional space around which clusters from all projections are searched with the same method as in the previous step. By adding the third dimension the clusters from the projections are assigned coordinates perpendicular to the measuring axis of their detector planes, which enables the exclusion of clusters which would then lie outside the active region of their detectors. Except for the limit on the the inclination angle the same procedure is used for cleaning up the set of three-dimensional track segment candidates as in the two-dimensional case.

The last step of the *pre-pattern* is a further cleaning of the set of three-dimensional track piece candidates, which amongst other things improves the definition of the track segments and removes falsely associated clusters. To decrease the computing time involved a dictionary of all possible tracks is used, which is obtained from Monte-Carlo simulation and translates between two different representations of the tracks: a track as a set of clusters and a track as a vector in the five-dimensional phase space. The dictionary is implemented as a look-up table indexed by the five phase space variables: horizontal and vertical position, horizontal and vertical angle of dip and inverse momentum times charge.[7] Each entry contains the corresponding set of cluster coordinates. This facilitates a fast calculation of χ^2 and its partial derivatives for a given phase space vector by interpolation, allowing efficient fitting of the track segment parameters for all candidates. If the resulting χ^2 exceeds a threshold the relevance of the individual clusters is evaluated. If their contribution to χ^2 is too large, they are removed from the track segment candidate and fitting starts over. Finally, a cut on the number of hits and on the resulting χ^2 cleans up the sample.

The third step of tracking—the so-called *bridging*—tries to build full tracks by sequentially connecting the up- and downstream track segments across all zone borders. In case of the bridging through the magnets, the up- and downstream parts are first extrapolated to the central plane of the magnet and then checked for compatibility of the track parameters. The preselected combinations are fitted globally and sorted by a quality function which mainly takes into account the χ^2 of the fit. The sorted list is processed in the direction of decreasing quality and combinations that share an up- or downstream track piece with a higher quality track combination are removed. The bridging through the absorbers uses a simplified version of the above method in which the track combinations

[7] This choice to represent the momentum yields a variable that fits positive, neutral and negative tracks within a contiguous region of phase space.

are not fitted, but the quality function acts directly on the sum of squared differences of the extrapolated track parameters.

In the last stage of tracking, the optimal track parameters and uncertainty matrices of the bridged tracks are calculated using a Kalman filter (see [44] and [45]). This iterative method successively adds the clusters to the fit and calculates the optimal track parameters and χ^2 for every step. Extrapolation through the apparatus in this stage takes into account multiple scattering in traversed material. Again, χ^2 cuts are applied to improve the quality of the track sample.

2.3.2 Vertexing

During vertexing, CORAL tries to find the primary and secondary interaction points and decay vertexes. In order to reconstruct the primary interaction point, first the beam and the scattered muon trajectories need to be identified since they constitute the anchors for the reconstruction. A track is taken for the scattered muon μ' if it has segments after the second muon filter, its charge is positive and its distance to the beam axis at the z position of the target center is below 2 cm. The corresponding beam muon is the beam track that approaches the μ' the closest, matches the timing of the μ' better than 1 ns and whose point of closest approach with the μ' is in the region around the actual target volume.

The vertexing is then performed in two steps. During the fast preselection, tracks are searched which are likely to originate in a common point in space using geometric and kinematic criteria. The mean value of the coordinates of the pair-wise points of closest approach is taken as a first estimate of the vertex position.

In the second step the optimum track and vertex parameters are calculated using the inverse Kalman filter method with the constraint that all tracks coming from one vertex share that one point in space. Based on the fit results, χ^2 is recalculated for each track and the track with the highest value is discarded from the vertex and a new fit calculated until the highest χ^2 value is below a certain threshold. For the reconstruction of the primary vertex the beam and the scattered muon track are not allowed to be removed. The optimized track and vertex parameters are saved in addition to the original track parameters.

Chapter 3

The COMPASS Configuration Server

At the collaboration meeting in Trieste February 2002 it became obvious that the initialization and configuration of the various parts of the COMPASS spectrometer were not in a state that allowed for a further growth in complexity and flexibility: the missing centralized infrastructure had born a plethora of more or less powerful but greatly different tools, written mostly by the detector experts for their own use. The incoherent user interfaces made it increasingly difficult for the shift crew to act upon detector failure or to change the mode of operation of the apparatus.

Especially during the setup and commissioning phase of COMPASS, the most commonly needed functionality was the reprogramming of the front-end and readout hardware. Even in stable running conditions it happens now and again that a detector front-end card loses its programming and needs to be initialized. Configuration files are used to describe this process, but with each detector group having their own conventions and naming scheme for these files the shift crew was forced to choose between non-obvious alternatives.

The solution was to introduce a database, recording the current state of the configuration done by the detector experts but also keeping the history of changes to this configuration. A simple program—tailored to the needs of the shift crew—controls the transfer of this information to the hardware, while a more complex daemon[1] called config_server handles all the details of programming the individual chips.[2]

The remainder of this chapter describes in detail the technology used and developed for this project, first covering the communication and database aspects and then discussing the user interface. In-depth information about the storage of information in the database and the inner workings of the config_server daemon can be found in section C.1 and is useful only to experts and developers.

[1] The term "daemon" refers to a program running in the background of a computer, waiting to be activated by some event and do its work.

[2] Most of the programmable ICs used in COMPASS are so-called FPGAs, field programmable gate arrays, see e. g. [6] for an introduction and a more specific view on the GEM/Silicon readout.

3.1 The Technology

The crucial point of the configuration database system is the communication with the hardware. Thankfully, all COMPASS detector systems are connected to data concentration modules which in turn use the VME standard for power supply and communication. Each VME crate—housing up to 23 readout modules—is equipped with a so-called VME-CPU, a general purpose PC mounted on a main board which features the standard VME connectors and the corresponding interface chip. These PCs run Linux with a kernel driver for the VME interface and are connected to the outside world via conventional ethernet networking. The configuration server daemon runs on these PCs and the means of communication with the shift crew are provided by the DIM library, while the data source is a MYSQL database server. The aforementioned components of this system are introduced in the following subsections.

3.1.1 VME

The VME standard was drafted in 1981 by Motorola, Mostek and Signetics and later adopted by ANSI, IEC, IEEE and VITA. As it can be implemented royalty free it is quite popular in high energy physics. The name stands for VERSAmoduleEurocard as it is the marriage of the electrical VERSAbus specification with the mechanical Eurocard standard. The specification describes an asynchronous non-multiplexed bus with 16 or 32 data lines and 16, 24 or 32 address lines with a maximum data rate of 10 Mbaud per line. In case of GeSiCA only 16 data lines are used, which allows a maximum transfer rate of slightly below 20 MB s^{-1}. For more details see [46].

The data concentrator modules are 9U (triple-height) Eurocard boards with extra connectors for plugging the TCS receiver (see section 4.1) and the S-LINK transceiver (see section 4.3) via feed-throughs to the back side of the back plane. The modules are powered and cooled by standard VME crates which are readily available with the extra connections on the back plane.

The software side of VME is represented by a Linux kernel driver which allows the mapping of VME address space into the virtual address space of the VME-CPU. This makes accessing GeSiCA registers transparently possible via simple memory accesses in any programming language. A C++ wrapper library around the setting up of the mapping and the actual memory access has been written to encapsulate the VME code in case a modification has to be made.

The Linux driver supports multiple mappings at once which is also necessary as the GeSiCA modules are accessed using 16-bit word size and the CATCH modules demand 32-bit access.

3.1.2 DIM

DIM stands for DISTRIBUTED INFORMATION MANAGEMENT and denominates a software library and utilities developed at CERN, see [5] for more information. The key feature is the DIM name service which associates keys with values across a computer network. Each process on a participating computer can publish so-called services, identified by a string like `"621/Configuration"`, its key. Other processes—called clients—query the name service—a process running on a specific PC—for the location of the services they need to send commands to or retrieve values from. The commands and values are of the simple variable types of the C language, including C-style strings. Reading values from a service can be done in three modes:

single-shot Retrieve the value synchronously once.

synchronous Subscribe to the service and get notifications upon each change of the value

asynchronous Subscribe to the service and cache all changes to the value locally.

For the purpose of the configuration server, each data concentration module is represented by a set of services. First there are seven values which can be inquired:

ModuleStatus Integer describing the momentary status of the module, whether it is currently being initialized, or the failure or success of a previously finished command.

ModuleType There are several types of data concentration modules (mainly CATCH and GESICA, but also the TCS controller and prescaler), which publish additional services according to their type.

Configuration A string representation of the configuration information as taken from the database.

ErrorString If the last action requested from the module resulted in errors or warnings, the ModuleStatus is set accordingly and this service gives detailed information about the reason of the failure.

PortFlags Status of the front-end connection ports of the CATCH.

RobSlot Each concentrator module delivers its data stream to a readout buffer card in a Readout Buffer Computer. This connection is given by an optical fiber link, which in case of an error needs to be checked by the DAQ expert. This information is also used for selecting specific detectors for readout in the data acquisition.

Programmed This flags the last known overall status of the module. In case of the CATCH modules, `ReadStatus` updates this information.

Second there are three commands which can be sent to the module:

Initialize Performs a (partial) initialization of the readout system. In case of the GeSiCA modules, a variety of flags controls which parts are reprogrammed and/or verified.

Reset In case of the CATCH modules, this issues a software reset

ReadStatus The configuration is verified by reading back as much of it as the hardware permits[3] and the status values are updated accordingly.

The typical usage pattern is to issue an Initialize command and monitor the ModuleStatus variable. As soon as a final success or failure status is reached, the user is informed.

3.1.3 MYSQL

As the configuration consists of the set of relations between concentrator modules, configuration files, readout buffers and FPGA programs, it is natural to use a relational database as storage back end. The history of changes is modeled using the modification time stamps as connecting elements between the database rows representing the state before and after the change. Relational databases usually implement the Structured Query Language, SQL. There are several implementations freely available and the author chose MYSQL by personal preference. Other components of the COMPASS DAQ, like the run log-book, also employ this database engine, allowing aggregation of those different uses on a single database server to minimize the maintenance effort needed.

3.2 User Interface

An intuitive and functional user interface is the key to user acceptance, which in turn is needed to accomplish the goal of the configuration server: a unified detector setup procedure for COMPASS. The typical end user of the system is the shift crew, who is only interested in keeping the experiment running. Therefore the programming of detector components is done with a simple, easy-to-use command line utility called LOAD. The workings of LOAD depend of course on the correct database contents, which is provided by the detector experts via a web front-end. Finally there is the system developer who needs to investigate closer if something goes wrong, which regarding the DIM component is done with the dimclient. These three pieces of software are installed on all COMPASS online computers and are portrayed in the following subsections.

[3] Certain hardware registers are write-only and thus cannot be read back, e. g. message buffers filled sequentially via one special register.

3.2.1 LOAD

The main job of LOAD is to handle the selection of modules which shall be programmed, send the appropriate commands to the config_server and present the results of the operation to the user. The module selection is intentionally versatile: giving the name of a Readout Buffer or Readout Buffer slot selects all connected modules, the name of a VME-CPU selects all modules in the same crate, allcatch selects all of type CATCH, allgesica selects all of type GESICA, all does everything and last but not least individual source IDs or ranges thereof may be directly given.

If only one module is selected, the status changes are displayed synchronously, indicating success or failure of the individual steps during the initialization. Otherwise only the final status together with the accumulated error and warning messages is printed synchronously and a summary is given after the last module signaled the ending of the programming process.

While the CATCH modules are always completely programmed upon initialization, the GESICA driver offers fine-grained control over which parts of the readout chain are to be programmed, see the output of LOAD -h for more details.

3.2.2 fedb.pl

This web front-end to the configuration database was written by Lars Schmitt. Its purpose is to provide an efficient and reliable interface to the underlying database, ensuring basic data consistency and that the history of changes is properly recorded. Programming the database with the appropriate values is the responsibility of the detector experts and forms the base of the function of the whole system, namely to allow non-expert shift crew personnel to re-initialize complicated front-end hardware in case of problems.

3.2.3 dimclient

dimclient is a low level debugging tool, which can list services, read values and send commands arbitrarily. The operation mode is given by the first argument:

dimclient list takes a glob pattern (like shell file name matching) and finds all services, commands and RPCs which match the pattern. The names of the found items are printed together with their data type (int, float, double or string) and in case of a service the current value is queried and printed.

dimclient lsfast does the same as the list command but does not query the values of services; this can result in a substantial time saving.

dimclient cmd takes the name (or a glob pattern) of a DIM command and the value to send, which can be either an integer or a string and sends the value to all matching commands. It does not wait for anything to happen in response as generic feedback in response to a command is not specified for DIM.

dimclient monitor subscribes to the named command and prints its value as it is updated. This is useful for certain services of the TCS modules like SPS spill status or spill summary information.

This tool is most useful for querying the configuration string of the modules registered in DIM, e. g. dimclient list 640/Configuration to get the current config_server image of the configuration for the first silicon station. Please note that querying a service does not trigger a reload of the configuration from the database. For this an Initialize command must be sent, either directly using dimclient cmd ⟨source ID⟩/Initialize or via LOAD.

Another use of this program is to execute the Restart command of the config_server daemon. The DIM name of this command is pccofeXX/Restart, where *XX* is to be substituted by the VME crate number. A restart is necessary only if new modules have been entered into the database or if the Version_tag is to be changed. The syntax is

dimclient cmd ⟨pccofeXX⟩/Restart [*new Version_tag*]

If no Version_tag is given the previous setting is kept. The config_server starts out with the Version_tag "latest".

3.3 Installation at COMPASS

The config_server daemon is started automatically by a CRON job via the restarter. This small program is part of the config_server distribution and responds to a termination of the config_server by starting it anew. This is a safeguard against rare crashes of the daemon, but is also exercised intentionally when a restart of the daemon is necessary in case of a large database configuration change or the deployment of a new config_server version. The aforementioned CRON job script, which runs every minute, checks for the config_server as well as the restarter processes and restarts both in case an anomaly is found.

All actions of restarter and config_server are logged to /tmp/config_server.log on the respective VME-CPU. In case of trouble, it is advisable to look at the recorded messages for a clue as to what is the underlying problem.

Chapter 4

The COMPASS Data Acquisition System

This section accompanies the experiment data from the detectors to their destination on tape. The hardware trigger system, which initiates the readout chain based on the conditions described in section 2.2.6, forms the input to the TRIGGER CONTROL SYSTEM (TCS), whose task it is to ensure the orderly transmission of event identifiers and time stamps to the concentrator modules. These in turn signal the connected detector front-ends to send out the data which they acquired during the time window corresponding to the trigger, and fill those data into the common COMPASS data format to send them on to the event building network. After assembling the events they are passed to the online filter software, which forms the second level of the COMPASS trigger system. Those which remain travel on to the central data recording facilities of CERN. Each of these steps will be highlighted in the following sections, followed by an overview of the experiment control at COMPASS as far as DAQ is concerned. The current author has been involved in the setup and maintenance of the event building network, the online filter and the experiment control.

4.1 Trigger Control System

The most important duty of the TCS is the distribution of the trigger signal to the detector front-ends. For this purpose a powerful laser system with passive optical splitters fans out the trigger decision to each of the 152 concentrator modules. The overall latency of this signal is below 1.5 µs and is mainly due to signal propagation times to and from the trigger barrack. The laser system is also used to distribute the COMPASS reference clock of 38.88 MHz which forms the time base for all TDC in the experiment. A special encoding scheme additionally allows the transmission of event labels, reset signals and trigger specific configuration information over the same physical medium. The information is decoded by the TCS receiver cards plugged to the back side of or built into the data concentrator modules (the latter is the case only for HOTGESICA).

Besides handling communications, the TCS watches over the DAQ dead time by generat-

ing the *DAQ busy* signal for the first level trigger, which is needed to ensure proper readout of all detectors. While most detector front-ends operate in a pipelined mode and thus do not require a fixed dead time after each trigger, experience shows that a dead time of 4 µs is needed after each event to avoid readout errors. More important at COMPASS is the so-called *variable dead time* which accounts for the depth of the front-ends' pipeline and their readout data rate by enforcing that no more than three triggers are issued within 75 µs.

Some detectors require constant monitoring of their performance. Therefore the TCS provides facilities to send so-called *calibration triggers* to each group of detector front-ends, including an advance signal for initiating specific calibration actions. In case of the HCALI this means that a LED pulser emits a well-defined flash of light which allows the online calibration of the photo multiplier amplification. The time delay between the advance signal and the trigger can be configured as well as the frequency of up to 300 Hz or whether the trigger is enabled during on-spill or off-spill time or both.

This system has been developed by Boris Grube and Igor Konorov, see [6, 23] for details.

4.2 Front-End Electronics

The COMPASS detector features more than 200,000 readout channels which need to be read out for every event (see [36] for a break down by detector types). Of course only a small fraction contain relevant information, namely those channels which witnessed the passing of a particle. The step of separating the signal from the noise is done by applying thresholds—either analog in form of a discriminator or digitally by processing the raw bits—at the earliest possible stage of data processing. Together with the formatting of the data for the transport towards the counting room, this is the main job of the front-end electronics. It is obvious that each detector type needs its own design, acting as unification layer between the different physical detectors and the next level of the COMPASS data acquisition.

For the operation of the front-end electronics, configuration of the readout channels (e. g. zero suppression modes, time gates, etc.) as well as calibration data (thresholds, noise parameters, etc.) are needed. As these data change frequently they are stored in volatile memory which needs to be programmed each time the electronics are powered up. For this task to be manageable by the shift crew a special configuration server software has been written, which is presented in chapter 3.

4.3 Data Concentration

A single optical fiber pair as used at the COMPASS experiment—using the S-LINK protocol—can transport up to 160 MB s^{-1}, which means that much bandwidth would be wasted if each detector front-end card had its own connection to the counting room.

Also the front-end cards need to be accessible from the commodity computer network for monitoring and programming. Thus, a layer of infrastructure is added, which consists of VME (see section 3.1.1) modules called GeSiCA or CATCH, depending on the detector type. Up to 23 modules live inside a VME crate together with a VME PC module (VME-CPU), which is mainly responsible for communications and runs the config_server daemon. The concentrator modules take in data coming from several front-end cards, merge the data belonging to the same event, format the event packets for transport, add the event identification received from the TCS and send them via S-LINK connections to the counting room. In case of the CATCH certain detector types have a low data rate per front-end port, which is taken into account by aggregating up to four modules into a single link using an S-LINK multiplexer.

4.4 Event Building

All steps leading up to this point are done in hardware to maximize the rate capacity of the system, but the next step is too complex for this approach. The experiment data arrive in the counting room via 92 optical S-LINK cables, over which the information pertaining to each event is spread out. These 92 streams of sub-detector information are multiplexed into 12 streams of complete event information by the ALICE DATE software package.

The Readout Buffer PCs receive the data streams using custom-made PCI cards which act as a buffer holding a few SPS spills worth of data between the S-LINK receiver cards and the PC main memory, hence their name *spill-buffer cards*. The readout part of the ALICE DATE software controls the flow of data from the spill-buffer cards to main memory with the help of a custom kernel driver using an interrupt-based DMA transfer scheme. The received S-LINK blocks are checked for errors, assembled from the up to four links into sub-events and forwarded to the recorder process, who sends each sub-event to the right Event Builder.

Each event is uniquely identified by its trigger number and built at one specific Event Builder PC, chosen without external communication in a round-robin fashion. Once all sub-events pertaining to one event are received at the Event Builder, the parts are assembled with a global data header and sent on to the online filter. Buffers at the sending and receiving sides as well as within the ALICE DATE processes even out most of the short-term (on the scale of seconds) congestions that appear in such a complex networking application, but it took great effort to adapt the system parameters to the COMPASS environment, where the event size is about two orders of magnitude smaller than at the ALICE experiment.

During 2004 there were 19 Readout Buffer and 12 Event Builder PCs, each featuring two CPU, 1 GiB of main memory and the latter including 640 GB disk arrays. For 2006 this system has been upgraded under the author's responsibility by including 10 new Readout Buffer PCs and replacing the Event Builder PCs with new ones having 4 GiB of main memory, faster processors and 1 TB disk arrays.

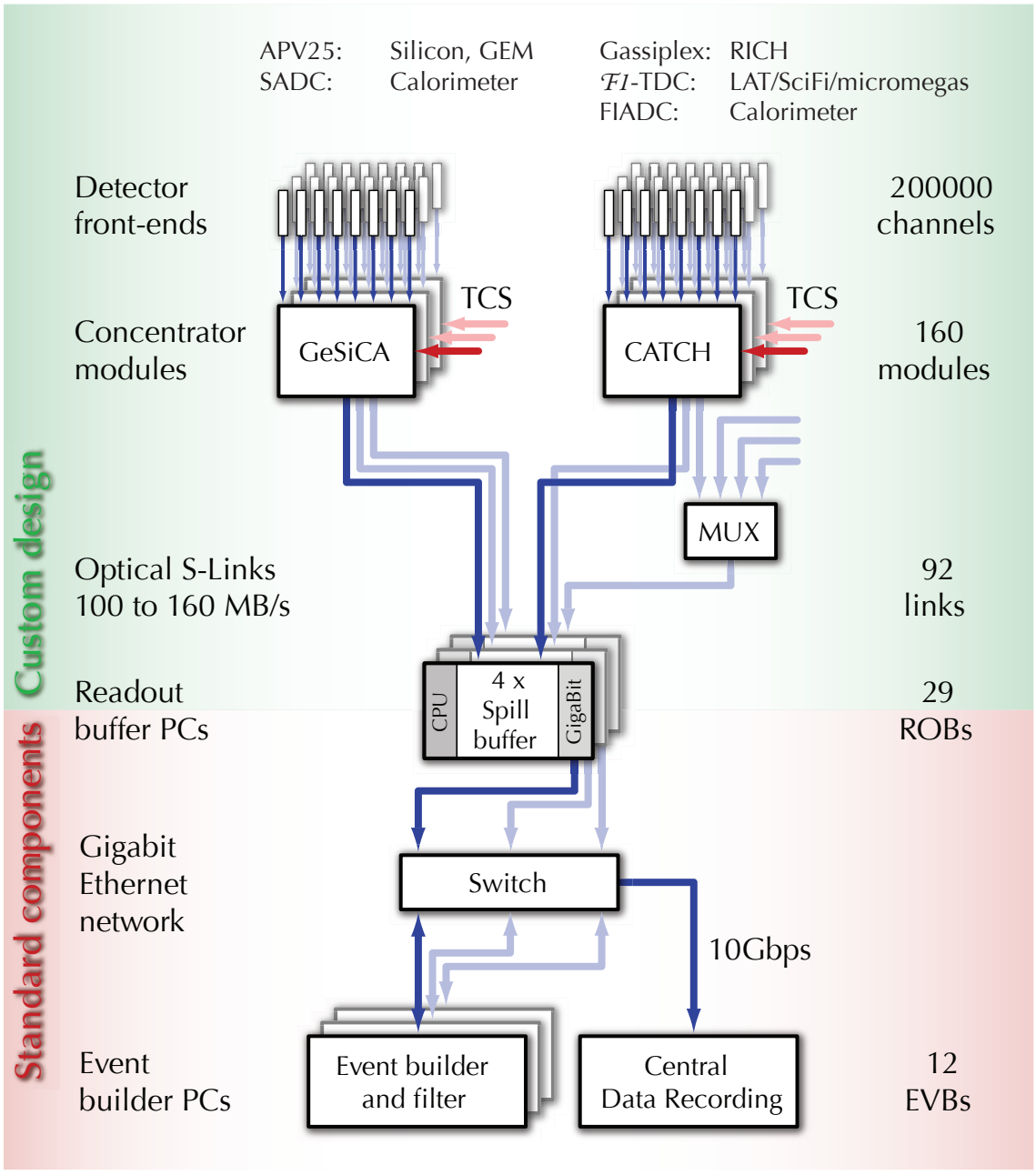


Figure 4.1: Schematic view of the COMPASS data acquisition system. The data flow is indicated by the arrows; the given numbers represent the setup as of 2006. [from [6]]

4.5 Online Filter

The original ALICE DATE Event Builder as used at COMPASS up to the beam time of 2002—actually developed as a test beam data acquisition for ALICE—could only write its output data stream directly to a file on disk. The CDR scripts described in the next section thus have an interface which is based solely on the creation of specific files and directories. As COMPASS finished its commissioning phase and grew efficient in producing data, it became obvious that the demands would soon exceed the capacities in volume as well as rate. Therefore after the beam time of 2002 work started on a program which makes a higher level trigger decision on the data stream before it gets written to disk. This software, which is described in detail in chapter 5, receives the event stream from a modified version of the Event Builder, partially decoding and reconstructing the events and either rejecting or accepting them for write-out to disk.

4.6 Central Data Recording

Central Data Recording means the hardware and software providing the permanent storage services of CERN as well as the software running on the Event Builder computers, which ensures the timely and orderly transfer of experiment data from the experiment hall to said services.

The permanent layer of the CERN storage concept is realized using tape libraries. This is advantageous from the point of view of infrastructure and maintenance cost, but it also implies large latencies for the retrieval and registration of data. Therefore a layer of disk servers grouped in stage pools acts as a cache for data on their way to the tapes or back to the user. The migration of data between tape and disk storage is managed by the CASTOR software, which employs algorithms to choose the least recently or least frequently used data to evict from the stage pool when new data are put into the system or requested from the tapes. All this is transparent to the user, who just links a special library to his own programs, enabling the pseudo–file system `/castor/cern.ch/`. The staging of data from tape to disk pool is done automatically and often used data typically are already available on dedicated experiment pools.

The connection between COMPASS and CASTOR is handled by a set of scripts running on the Event Builder machines. As soon as the data of one run are on disk, so-called bookmark files signal that these data are ready to be transferred. When the next slot to accept these data at the computing center is free, the transfer starts automatically. It is clear that on average the data rate produced by COMPASS must match the rate going to tape, but in case of temporary problems the Event Builder computers have large disk arrays which can buffer as much data as can be produced in one good day. This is the last and largest buffer in the COMPASS DAQ system.

Table 4.1 gives an overview of the amount of raw data which has been produced by COMPASS in its three physics beam times 2002–2004. Most of the data of 2003 and

2004 have passed through the filter, while in 2003 the filtering was really active (i. e. discarding events) only for a short period of time near the end of the beam time. Before, it was configured to only mark the data with the would-be decisions so that extensive testing and verification of the algorithms could take place until the collaboration was convinced of the stability and correctness of the program.

4.7 Experiment Control

Experiment control at compass comprises multiple systems, most notably the alice date runControl and the dcs for controlling operation parameters of the detectors and logging environment conditions, which is implemented using pvss. Additionally every expert group for compass components have their software tools for maintaining their equipment, which can be quite complex as in the case of the polarized target.

The original function of the alice date runControl is to orderly start and stop data taking. In view of the great number of PCs and other hardware systems involved, this task is challenging, especially taking into consideration that error conditions need to be handled correctly and efficiently in order not to waste beam time. In order to keep the maintenance effort at a reasonable level, the core logic is not written in a procedural programming language but in a cern-developed state machine description language.

After adapting the runControl to work with compass equipment, it has been expanded towards a more general experiment control system by including graphical configuration facilities for trigger selection, calibration trigger configuration, enabling and disabling of individual detectors, online filter configuration and several diagnostic tools concerning trigger and detectors, including a direct way to re-initialize detector front-ends in case of problems. Unification of all these services under a common user interface has significantly increased the effectiveness of the shift crew.

	size	
longitudinal	632.03 TiB	65.6%
transverse	169.99 TiB	17.6%
hadron	78.79 TiB	8.2%
commissioning	41.16 TiB	4.3%
testing	41.72 TiB	4.3%
total	963.69 TiB	

year	period	data size	comment
2002	P1A	20.56 TiB	commissioning
	P1B	20.60 TiB	commissioning
	P1C	11.54 TiB	
	P2A	24.78 TiB	
	P2B	16.76 TiB	transverse
	P2C	18.61 TiB	transverse
	P2D	17.98 TiB	
	P2E	21.80 TiB	
	P2F	9.19 TiB	
	P2G	19.78 TiB	
	P2H	14.09 TiB	transverse
total		195.69 TiB	

year	period	data size	comment
2003	P1A	26.97 TiB	
	P1B	12.46 TiB	
	P1C	15.14 TiB	
	P1D	23.59 TiB	
	P1E	28.54 TiB	
	P1F	18.09 TiB	
	P1G	25.13 TiB	transverse
	P1H	21.08 TiB	transverse
	P1I	22.54 TiB	
	P1J	34.13 TiB	
	P1K	2.69 TiB	transverse
	T01	11.59 TiB	testing
	T02	0.81 TiB	testing
total		242.76 TiB	

year	period	data size	comment
2004	T01	0.67 TiB	testing
	T19	15.15 TiB	testing
	T24	7.76 TiB	testing
	T41	0.45 TiB	testing
	T43	5.29 TiB	testing
	W21	3.51 TiB	
	W22	46.35 TiB	
	W23	29.62 TiB	
	W26	22.92 TiB	
	W27	13.16 TiB	
	W28	17.37 TiB	
	W29	16.17 TiB	
	W30	20.81 TiB	
	W31	22.35 TiB	
	W32	30.10 TiB	
	W33	23.48 TiB	transverse
	W34	18.88 TiB	transverse
	W35	28.71 TiB	transverse
	W36	17.34 TiB	transverse
	W37	32.28 TiB	
	W38	35.24 TiB	
	W39	19.20 TiB	
	W40	19.89 TiB	
	W44	32.91 TiB	hadron
	W45	12.59 TiB	hadron
	W46	33.29 TiB	hadron
total		525.49 TiB	

Table 4.1: Size of the raw data which are in CASTOR. Most of the data of 2003 and 2004 have been processed by CINDERELLA. This table does not include the data which have not been recorded and for which thus no reliable size estimate is available. The total data volume recorded in 2002–2004 is 963.69 TiB. Periods without a comment have been recorded with longitudinal target polarization.

Chapter 5

The COMPASS Online Event Filter

5.1 Motivation

The design goal of COMPASS was to build a dead time free high rate detector to make most efficient use of the beams provided by the SPS. Therefore the trigger system is optimized for minimum decision making time and not for trigger purity. To cope with the noise rates of several $10^6\ s^{-1}$, vetoes were added, which introduce about 7% trigger dead time at the usual readout rate of $10^4\ s^{-1}$, lessening the benefit of the nearly dead time free detector readout. At this point the online filtering of the data stream comes into play, which serves two main purposes:

- Reduce dead time introduced by vetoes by moving this functionality partly into the software domain. This means that the purity of the first level trigger is reduced intentionally to lower the trigger dead time and the added noise events are filtered out at a later stage, decoupled from the actual sampling of the detectors. However, this technique also has a limit, because the readout of the detector front-ends takes a certain amount of time, for example in case of the APV25 this amounts to 21 µs per event.[1] The depth of the readout pipelines limits the length of bursts during which the maximum sustainable event rate—the inverse of the readout time per event—can be exceeded, thus a dynamic DAQ dead time is necessary which outweighs the trigger dead time at rates greater than approximately $2 \cdot 10^4\ s^{-1}$.

- Allow for generally higher rates by rejecting events which are not interesting for physics analysis, thus increasing the purity of the trigger system without reducing its efficiency.[2] Storing uninteresting events on tape reduces the budget allocated for tape storage that is actually used as intended.

[1] This number is being halved for the 2007 run by doubling the readout frequency.

[2] Of course a cut will always reject a certain fraction of interesting events, but this is designed to be on the sub percent level, so that its effect on physics analysis is barely measurable

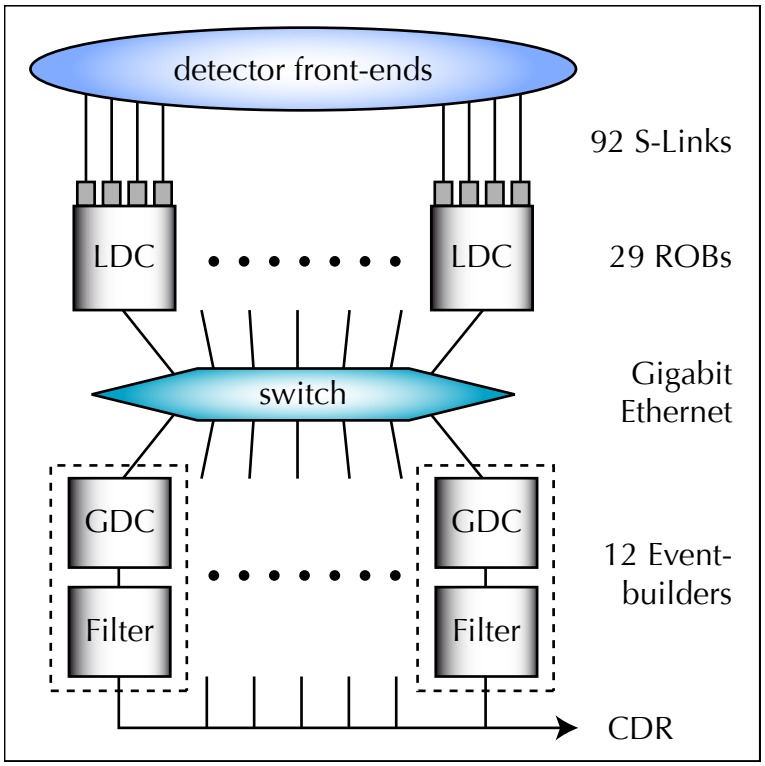


Figure 5.1: Schematic view of the integration of the CINDERELLA online filter in the COMPASS DAQ. The numbers given on the right are for the beam time of 2006.

Another positive effect of online decoding is that inconsistencies and missing information are detected during data taking, enabling the shift crew to react immediately, thereby improving the overall data quality.

While the current author has developed the general framework and integration concept, much work was put into this project by Thiemo Nagel. For all the gory details about the actual filter modules—doing the decoding of the detector data and making the decisions—see his diploma thesis [35]. The online filter project has been named CINDERELLA in reference to the fairy tale, in which one problem posed to the protagonist consists in separating the good peas from the spoiled ones.

5.2 Framework

In the absence of a dedicated filter farm, the online filtering was implemented on the Event Builder machines as shown in figure 5.1, where the data are first assembled from the fragments coming from the various sub-detectors. These dual processor machines have enough resources left after event building to allow for a simple analysis of the data before they are written to disk. Assuming a trigger rate of $10^4\,\mathrm{s}^{-1}$ and a spill structure of 4.8 s beam alternating with 12 s off-spill time, the online filter has to digest about 3000 events per second. Dividing the work between all 12 Event Builder PCs gives a rate of

$250\,s^{-1}$ per machine or an average decision making time of about 4 ms per event.

The data stream consists of an inhomogeneous mix of events, some of which take much more time to process than others, so to decouple the input stream from the filtering process a large buffer[3]—called *main buffer* for the remainder of this chapter—is used. This buffer is the center of synchronization for the individual components of the online filter:

- The *input thread* reads the ALICE DATE stream from a file or TCP/IP socket and fills the buffer, pausing if either the buffer runs full—thereby causing back pressure in the DAQ system—or the input stream has no more events ready at the moment. The latter should happen near the end of each spill, otherwise the data will not be digested upon the next spill start and the buffer will inevitably fill up over the course of the following spills until the DAQ system fails. It is a matter of tuning the online filter such that it does not consume more CPU cycles than are available per full spill cycle.

- The *filter threads* wait for events to appear in the buffer to process them and mark them with their decisions. The number of filter threads is tunable—e. g. to the number of CPU in the system—so that the computing capacity of the event builder machines can be fully utilized. The processing of each event is given by the execution of a filter chain, a preconfigured series of filter modules, each responsible for a self-contained task like decoding the data of one type of detector.

- The *output thread* in turn waits for processed and marked events to appear in the buffer to move them to the dedicated output buffer and append the online filter information, or to discard them. The information attached to the individual events contains the decisions of all filter modules which have processed the event so that an off-line analysis can directly examine the filter performance. At every end of spill a special event type is generated by the DAQ system to which detailed statistics about accept/reject ratios, the time used by the online filter and the complete online filter version information is attached. In this way, the statistic is embedded in the data stream which is written to permanent storage and can also be visualized online using the ALICE DATE monitoring library. The spill by spill stability of accept ratios for the different triggers is an important utility to detect problems within the whole apparatus during data taking as it immediately points out fluctuations in the general data quality.

- The *watchdog thread* ensures that none of the filter threads get stuck while processing an event for too long. Its configurable timeout is set such that infinite loops caused e. g. by incorrect treatment of corrupted data are broken up by killing the trapped thread, letting the event through unfiltered and starting a new filter thread to continue with the next available event. It should be noted that each occurrence of this action is recorded as an error visible to the shift crew and analyzed by the online filter team as it points to a possible weakness in the online filter code.

[3] 128 MiB during the beam time of 2004

5.2.1 Data Structures Used for Buffering

Efficient streaming demands the reading of contiguous and large chunks of data. Therefore the main buffer directly holds the unmodified ALICE DATE stream as a ring buffer and is filled by always trying to read the largest possible amount of data. During the wrap around of the ring buffer the last partial[4] event is copied to the beginning of the buffer so that events are always contiguous in memory, an assertion which greatly simplifies the data handling throughout the other parts of the online filter software. This stream is cataloged event-wise by the input thread, triggering another read operation when the data from the last read operation are exhausted. The catalog itself is another ring buffer of fixed size elements pointing to the raw event data and holding state information on the event, including a pointer to the output of the various decoding and decision making modules which will eventually be filled by the filter thread. This second ring buffer serves two purposes: to keep the meta information from interfering with the streaming of the raw data and to allow efficient access to the variably sized events.

The sizes of both buffers need to be tuned to the actual variations present in the data rate to ensure optimal data taking stability using the available memory. In practice, reasons for a degraded data taking stability are strongly dependent on the subjective perception of the people in charge, so we started out with a buffer size that seemed sufficient while not putting much memory pressure on the event builder machines, gained experience with it over a few weeks and doubled it to 128 MiB as rate induced problems suggested a bottle neck. With the new hardware available from 2006, the buffer has been further enlarged to 512 MiB to make use of the larger main memory. Afterward the remaining problems were attributed mostly to the front-end electronics or in some cases complete system failures of DAQ computers. The size of the meta-data ring buffer is adapted to the main buffer size by dividing by the average event size, adding a generous factor of 2 for the case that a sequence of small events is encountered and allocating the resulting number of catalog entries.[5]

As it was discovered that the main bottle neck of the original ALICE DATE event building process was the use of a system call that leads to the write-out of the events in many very small parts,[6] it became obvious that not only the input but also the output needs to be buffered. Discussions with the Linux kernel developer community have resulted in the addition of a 4 MiB output buffer, which is filled by the output thread and written in one go once the next event does not fit in any more. This measure alone increases the rate capability of the DAQ system substantially. A number cannot easily be given as the prob-

[4] Since the read operations cannot take into account the size of the obtained events, the probability for the last read before the buffer wrap around to happen exactly at an event boundary is very small.

[5] A catalog entry records its status—free, unfiltered data, filtered data—together with time stamps for all transitions between the different states plus pointers to the raw and decoded event data. Together with the ID of the filter thread processing the associated event this sums up to a size of 32 bytes per entry.

[6] ALICE DATE v3.7.1 uses the `writev` system call to write the individual data blocks from the readout buffer computers—17 parts for each event—which is particularly annoying considering the journaling file system used for data recording: the small packets fill up the transactions rather fast, leading to costly journal commits which stall the disk transfer.

lem depends on the complex interplay of the network and disk utilization while receiving experiment data, writing it to disk and simultaneously transferring recently recorded data to permanent storage. However, data taking instability as measured in the number of DAQ failures per day decreases significantly even if the online filter is only used for buffering.

5.2.2 Synchronization

The synchronization is the central part of the online filter framework as it is the only place where different threads interact. It consists of only about 200 lines of code which have been meticulously debugged and a score of nearly 1 PiB of data which passed through this code legitimates the attribute "bug free". The rationale behind the design is that each locking operation costs a non-negligible amount of CPU time,[7] which leads to the consequence that locks are only involved where they cannot be avoided, namely to synchronize the access of the concurrent filter threads to the buffer data structures. Given this locking, only three classes of threads need to be considered, for which the framework keeps three indexes pointing at the next event to be processed. The synchronization is done by a well defined sequence of transitions between the states of the entries of the meta-data buffer, called *blocks* in the following:

- The input thread waits for its next block to become *free* and for data to become available on the input data stream. After checking that the next event to be cataloged fits completely into the main buffer and ensuring that it has been transferred there, the block is filled with the meta-data for the event and is marked *new*. A small complication is given by the fact that there are events, which are not filtered at all[8]; those are marked *direct write-out*.
- The filter thread waits for its next block to become *new*, runs the filter chain and marks it either *accept* or *reject*. In case a *direct write-out* block is encountered, it is marked *direct write-out 2* and not filtered.
- The output thread waits for its next block to become marked either *accept*, *reject* or *direct write-out 2*, treats the event accordingly and marks the block *free*.

The correctness of this scheme is guaranteed by the atomicity of reads and writes of the machine word size. No locking is needed because blocks in a certain state are only altered by one specific thread, so e. g. waiting for the block to become *free* ensures that the input thread can safely modify it, because no other thread handles *free* blocks. The state *direct*

[7] The reason is that common locking implementations spin for a short time trying to acquire the lock, but if that fails they resort to using a system call. Because of the ever growing internal state of modern CPUs, system calls have a rather constant time overhead of about 1 µs in spite of the increasing clock frequency.

[8] Only the start of run record, the first event in run, calibration events and events with a size greater than that of the output buffer bypass the filtering step.

write-out 2 ensures that the output thread cannot overtake the filter threads but is blocked until they also acknowledged the *direct write-out* blocks by changing the flag.

In the context of thread synchronization within cinderella, waiting means to relinquish the cpu for at least 2 ms.[9] Given a single cpu system this leads to a batching of the processing: first the input thread catalogs some events, then some events are filtered until finally some events are moved to the output buffer and possibly written to disk. Considering the caching of code in modern cpu this sequence is beneficial to overall system performance when compared to the situation in which each event is processed as soon as it arrives. The event builder machines have two cpu each, so in principle the I/O and the filtering could run in parallel, but there are also the alice date and the cdr[10] processes to be taken into account.

5.2.3 Configuration

Many aspects of cinderella are configurable, details of the framework behavior as well as the composition of the filter chain and the inner workings of the filter modules themselves. This calls for an easily extensible and well structured configuration language. In order not to invent yet another configuration format, an existing markup language is used. xml merely specifies how configuration items are marked in the file, how white space and comments are handled, and so on. Parsing libraries for this format readily exist and we chose libexpat by personal preference. Editing xml documents with a normal editor can be painful, especially when creating them, thus the build procedure for cinderella produces a complete options file with all default values built in. At the same time, TEX code is generated which documents all the options—if the author of the code cared to give a description. The whole process depends on a set of xml files, one for each cinderella subsystem, which are transformed by create_config into their various output forms. create_config is part of the cinderella distribution and has been written so that all information about cinderella configuration comes from one place only. This approach makes it possible for the programmer to understand the structure of the configuration and enables easy extension of the whole system. Section C.2.2 gives an example of this system.

While it is clear that the configuration of the compass online filter is not an everyday job and hence does not need to be easy for the average physicist, it is desirable to have a flexible and powerful system at hand for debugging. This is realized by having two command line switches recognized by cinderella: -f loads the specified xml file, parsing its contents and overwriting the internal state of the configuration items contained in the file, while -F loads a file which lists the name of an xml file on each line, treating the lines in order as if the file names were specified using -f. This way many small xml files can

[9] On a Linux 2.4 kernel this means in fact a waiting period of at least 10 ms due to the timer interrupt resolution. Newer Linux kernels have typically a smaller minimum wait period, but the actual frequency of the timer interrupt is currently a topic of discussion among the developers.

[10] Central Data Recording, see section 4.6.

be created for various purposes, each changing only a small set of related options. These files can be grouped into often used functional units. The same mechanism is used when CINDERELLA is running in the ALICE DATE environment on the Event Builder computers, where the configuration file list is passed in via the ALICE DATE shared memory. *It is highly recommended to import the* default_config.xml *created by* create_config *for the specific version of CINDERELLA as the basis for all following configuration files.* We deliberately decided not to compile any default values into the executable to keep the configuration process consistent and understandable.

5.2.4 Setting Up the Filter Chain

During compilation, all possible filter module types are registered in the framework, but it depends on the configuration, as to which ones are placed in the filter chain. For this purpose the XML configuration tree as it is read from the configuration files may contain elements which create instances of filter modules to be inserted into the chain. This is done by annotating the corresponding XML elements with an attribute named instance, allowing multiple copies of the same module type to be created. These copies are preconfigured with the items found in the "default" instance for that module type before parsing the specific configuration. This modularization allows e. g. the reuse of decoding algorithms for different detectors of the same type. Each module must provide code to produce a list of resources which are needed by the module and a list of resources provided—in most cases these resources depend on information given by the configuration. This information is used by the framework to automatically determine the order in which the modules have to be executed in the filter chain.

5.2.5 Debugging Features

5.2.5.1 Messages

The first measure for debugging a non-trivial piece of software is debug output. In case of CINDERELLA a flexible and efficient reporting system was conceived and implemented from the beginning. An easy to use set of macros provides functions to communicate errors, warnings and status information as well as low level debugging information via a unified interface, where the programmer gives a severity level (see table E.1 on page 164) and specifies the logical part of CINDERELLA where the message originates from. There is no need to remove rather verbose debug statements—even from inside tight loops—after debugging because when preparing a version destined for production use all debug statements below a certain severity are deactivated during compilation.

There are two output channels for debug information, each of them fully configurable with minimum severities for all logical parts of CINDERELLA:

- The operator of the ALICE DATE system—commonly the shift crew running the experiment— needs to be informed of all problems detected by the online filter.

Here the rate of messages should be really low not to overload the messaging system or the operator; normal running conditions result in less than a handful of messages per run. Errors are prominently marked in the *message display* on a separate screen of the run control system.

During development or debugging this stream is written to the terminal where cinderella is run. Usually the minimum severity is then lowered for some parts of the program in order to get an immediate debug feedback.

- The other stream goes to a *log file* for later inspection and post mortem error analysis. The configuration can either specify a single file location or a directory to place the log files in. In the latter case the files are numbered up to a configurable maximum, so that e. g. the log files of the last 1000 runs are kept. The run number is also embedded in the file name to ease the finding of information should questions arise during the days following the recording of a specific run.

  Usually this message stream is kept more verbose to enable detailed failure analysis after an unforeseen behavior.

Both streams can be individually rate limited so that the logging itself does not worsen a bad situation further by placing extra load on the system in case of excessive errors. All messages are prefixed with the current time to microsecond accuracy, the severity level and logical part of Cinderella. The messages going to the log file are also tagged by the file name and line number of the code that generated them.

5.2.5.2 Event Dumps

If an event has structural errors which make automated decoding impossible, the raw event data are dumped to a file[11] for later inspection by a detector front-end expert. This aids the debugging of rare failures of the front-end electronics.

5.2.5.3 Dynamic Memory Allocation Debugging

Nearly all data structures in cinderella are allocated dynamically. Especially in case of the data generated by the filter modules on an event by event basis great care has to be taken to free the used resources after the processing of the event is finished. Otherwise the memory would fill up quickly, eventually leading to a failure of the daq system. Debugging errors related to dynamic memory allocation is one of the most difficult tasks a programmer has to face, because the actual failure typically shows up in a place which is not intuitively connected to the part of the program which is responsible for the error.

To aid debugging of these errors, the cinderella source code uses self-written wrapper functions around all library calls which deal with memory allocation. These wrappers either directly call the underlying library versions or they do some accounting and checking

[11] The files are created in a configurable directory in the same manner as the log files.

first, depending upon whether CINDERELLA has been configured for production use or debugging. By introducing the notion of memory allocation sections it is possible to check that all memory allocated within one section is also released. A second attempt to release a chunk of memory—which would normally simply crash the program—is detected and prevented. In case a problem is found, the offending code is indicated giving the source file name and line number together with the name of the variable in question.

5.2.5.4 Crash Handling

In case of a memory allocation failure or a bad memory access—commonly called a segmentation violation—the program is aborted. As this event is undoubtedly unintentional, it is important to facilitate the analysis of what went wrong. The problem with such an error condition is that potentially any data structure could have been damaged so that the possibilities for an automatic assessment are limited. CINDERELLA identifies the code location and the chain of function calls leading up to that point in a so-called back-trace which is printed to the log file before exiting the program. This functionality is also included in the production version because it does not cost performance and enables post mortem analysis in case of a failure in production use.

5.3 Event Selection Principles

In view of the short time of only 4 ms which are available per event for arriving at a decision, simple but effective algorithms are needed. Therefore tracking particles in a magnetic field is not possible, which further restricts the possibilities because all particle identification depends on reconstructed trajectories. Even for regions in which particles can be assumed to travel along straight lines a full-fledged track fit would be too time-consuming, so the algorithms need to be formulated in terms of correlations and hit patterns in specific detectors.

The requirements and possibilities differ between the muon and hadron programs of COMPASS and are mainly given by the target setup. In case of the muon program the main objective is to partially replace hardware veto systems with second level triggering to reduce the first level trigger dead time. The veto system is described in section 2.2.6, its purpose being to inhibit the trigger when so-called *halo muons* are entering the spectrometer off the beam axis, thus not scattering off the target nucleons but elsewhere and creating background hits in tracking detectors as well as hodoscopes. This negative identification is complemented in the online filter by requiring a minimum number of hits correlated in time in the scintillating fiber and silicon micro-strip detectors upstream of the polarized target and in the beam momentum station. To efficiently find these correlations along the time axis between hits of varying time resolution the conditional coincidence algorithm[12] has been developed which is described in [35].

[12] Having n hits and requiring m to be correlated in time has a complexity of the order of n^m when

For the hadron run an estimator of the charged particle multiplicity exiting from the target has been implemented by reusing the conditional coincidence code, but not for cutting at a certain threshold but for determining the number of hits per detector plane within the three silicon micro-strip station downstream of the target. After application of a truncated mean algorithm, the resulting distribution has sharp peaks for integer numbers of hits per plane, which are used as a good approximation for the number of particles leaving the target within the acceptance of the silicon detectors. This allowed the implementation of a trigger on diffractive events by requesting a multiplicity above two.

5.4 Calibration

For any reconstruction of detector data, the calibration information for the detector must be available. As these calibrations are mostly generated off-line a few days after the data have been taken, a fast re-calibration has to be implemented in the online filter, which can then correct for day/night shifts and calibration movements. The starting values for the calibration are read from the official calibration database and then refined during each run.

Because of the limited scope of the online filter, only detectors with TDC readout are used besides the silicon micro-strip counters. The latter require a rather complex timing reconstruction from consecutive signal amplitude samples which is difficult to tune by hand and up to now no automatic calibration algorithm has been conceived. Thankfully even the longer-term deviations are below the timing resolution so that this is not a problem.

For the beam momentum station and the scintillating fiber detectors, the offset in TDC counts between the signal registration at the front-end and the measured trigger time need to be calibrated regularly, as these detectors have a time resolution below 0.3 ns which is smaller than observed day–night effects. The offsets are measured by histogramming—for all hits recorded during a data taking period—the time difference between the hit time and the event trigger time, and determining by a fit the position of the peak created by in-time hits over the background of uncorrelated signals. Being limited in statistics, the full calibration procedure, which treats each wire independently, cannot be applied for the calibration refinement, thus detector planes are calibrated as a whole. The rationale behind this is that e. g. the timing movements introduced by temperature changes in the 130 m long cables of the beam momentum stations affect always all wires of a plane in the same way. What is not covered by this, is the case when e. g. a photo multiplier base has to be replaced for ⅙ of a scintillating fiber plane, but this should anyway only happen during a machine development break and thus a completely new calibration is needed for the whole detector.

comparing all hits with one another, while the conditional coincidence relies on sorting of time intervals and has a complexity of $n \log n$.

It has turned out that it is necessary to calibrate the timing offsets relative to each trigger. Also, the triggers have different time resolutions ranging from 650 ps for the inner trigger to 3.5 ns for the calorimetric trigger. To gather enough statistics for a calibration of all relevant detector planes relative to all triggers, a full run of 100–200 spills is needed. This dictates that the timing histograms gathered by the individual CINDERELLA instances on the Event Builder machines are transferred to a central auto-calibration daemon called calibrator, summed up and fitted. The results are stored in a special table of the official calibration database from where CINDERELLA reads them during startup. As the fitting of a few hundred histograms takes longer than the pause between two runs, CINDERELLA usually uses the results of the next-to-last run. The algorithm which determines the exact refinements to be used selects the latest data which satisfy certain quality criteria, which are given by the statistics in the respective histogram and the χ^2 of the fit. This is a safeguard against non-converging fits, should they have escaped detection in the calibrator. In case no conforming refinement data can be found, the shift crew is instructed to record a dedicated *refinement run*, for which a special configuration item named "filter-calibration" is available in the ALICE DATE runControl, see section 4.7.

5.5 Technology

The Online Filter is implemented purely in C to get maximum performance. While it might be argued that C++ is in principle not slower than C, it is a fact that C++ requires more care when it comes to avoiding unnecessary copies of objects etc. The Linux kernel—being one of the biggest and most successful open source projects—has inspired key design features of CINDERELLA, most notably the abstraction of modular interfaces via structures containing function pointers. This technique allows the best part of C++—namely run-time polymorphism—to be implemented in C without adding the complexity overhead of fully object oriented design.

The filter is linked with the ALICE DATE monitoring library to be able to access the runControl shared memory segments for monitoring and controlling purposes. It receives the data stream via a TCP/IP socket, which is a very fast way of communication between two processes. This method was preferred over shared memory or Unix domain sockets because it can easily be used to move the filtering task to a dedicated filter farm.

5.6 Results and Outlook

During the beam time of 2004 CINDERELLA has been used for regular data taking at COMPASS. During the muon periods its goal was to enrich events which have a reconstructible beam track, which it did by discarding 23% of all events while keeping 99.6% of the desired ones. For the hadron pilot run at the end of the beam time, the track multiplicity cut was used to clean the diffractive trigger, discarding 45% of the events and

enabling the retirement of a prescaler with factor 2 for this trigger. The cut has been chosen in a very conservative way, but the loss ratio of wanted events has yet to be determined by physics analysis. While a filter configuration for the Primakoff trigger had also been developed, it was not activated because the trigger rates from the hardware trigger could just about be handled without further filtering.

For the future of compass the extension of this software is of highest priority, especially for the planned hadron run in 2008 which will bring a sizable increase in trigger and data rate. Sophisticated calorimetric triggers as well as tracking around the target may be needed to study rare processes for which no efficient hardware trigger can be constructed.

Chapter 6

Deep Inelastic Scattering

Since its beginnings in 1968 at the Stanford Linear Accelerator Center (SLAC) the field of deep inelastic scattering has received much attention from the experimental as well as from the theoretical side. There are several good text books on the subject, e. g. [47]. Therefore this chapter—after a short general introduction—only summarizes some of the more important aspects of deep inelastic scattering in conjunction with polarized structure functions, while a closer review is given of the recent calculations by B. Jäger et al [48] which initiated the analysis presented in the following chapter.

6.1 Introduction

It is well established that the spatial resolution achievable with conventional[1] wave optics is inversely proportional to the wave length of the light. The concept of quantum mechanics reveals a wave-like character within all possible forms of matter, thus transferring the wave concept to particle physics. De Broglie's law connects the wavelength of the quantum mechanical wave with the momentum of the particle by

$$\lambda = \frac{h}{|\mathbf{p}|} = \frac{\hbar c \cdot 2\pi}{pc} = \frac{197\,\text{MeV fm} \cdot 6.28}{pc} \tag{6.1}$$

Therefore, good resolution requires the probe to have high momentum. The further application of this knowledge leads to the technique of deep inelastic scattering, which means probing the internal structure of the nucleon with light of very short wave length: a momentum of 0.197 GeV/c corresponds to a wave length of 6.3 fm. The only source of such short wave length radiation which can be exploited experimentally are photons emitted by charged particles traveling nearly at the speed of light. It should be noted that eq. (6.1) is not invariant under changes of reference system and thus depends on whether the particle's momentum is evaluated e. g. in the laboratory frame or a frame in which

[1] Conventional in this context means that no phase information is measured.

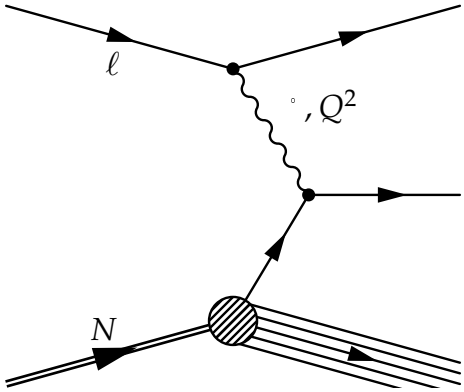


Figure 6.1: Leading order contribution to deep inelastic scattering in the Feynman picture. Higher order terms include multi-photon exchange, radiative corrections and external bremsstrahlung. The time axis extends to the right.

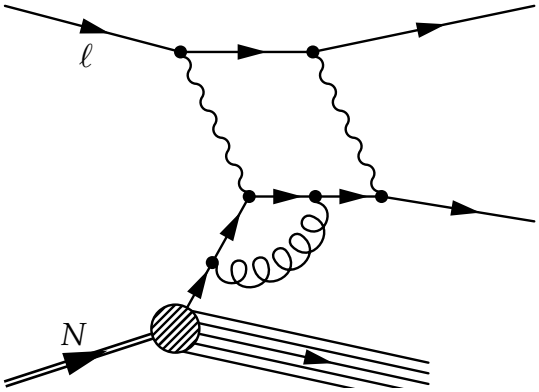


Figure 6.2: One possible non-leading order contribution to deep inelastic scattering with multi-photon exchange and a gluon loop. Each additional photon reduces the amplitude corresponding to the single graph roughly by $\frac{1}{137}$.

the particle is at rest. Thus, this direct interpretation is limited to a certain choice of reference frame, in case of deep inelastic scattering this is the so-called Breit frame, which is introduced in section 6.2. In calculations, the concept of resolution is realized by introducing so-called *scales*, like the factorization scale detailed in the following. Features below the scale are described by distribution functions, while features above the scale are resolved by the ab initio evaluation of the underlying quantum field theory, the so-called *hard scattering* matrix element.

Figure 6.1 shows the simplest picture of deep inelastic scattering. An incoming lepton emits a virtual photon of given energy ν and virtuality Q^2, which is absorbed by a quark in the nucleon. Assuming large momentum transfer, the struck quark can be treated as independent from the rest of the nucleon. In conjunction with the confinement exhibited by the strong interaction,[2] this leads to the production of additional hadrons, which is referred to by the term *inelastic*. The addition *deep* alludes to the fact that a considerable amount of energy is converted into mass while creating the additional hadrons from the remnants of the nucleon and the scattered quark. In other words, in deep inelastic scattering the lepton does not scatter off the nucleon as a whole, but interacts only with a so-called *parton* inside the nucleon.

This picture is of course simplified to enable the calculation. Nature always follows up on all possibilities, so in case we do not nail it down to one specific process we have to calculate all of them and sum them up correctly. Considering a microscopic process like deep inelastic scattering we can only fix the initial state by choosing target and beam, and we can preferentially select a class of final states by setting up a specific trigger in the experiment.[3] Internal complications like the ones shown in figure 6.2 contribute to the

[2] Confinement means that single quarks cannot exist as free particles; they always couple with other quarks and gluons to form hadrons.

[3] The issue of external bremsstrahlung will be ignored here; see section 7.4.3 for further discussion.

total cross section, but thankfully QED—which describes everything directly involving photons—inherently gives a rather strong ordering of contributions: each coupling of a photon to a charged particle corresponds to a factor $\frac{1}{137}$ in the probability for that diagram, thus diagrams with many such couplings can in most cases be neglected relative to the ones with less couplings. This ordering enables the formulation of the cross section as a perturbation series, which is the first prerequisite for calculating scattering cross sections. The single photon exchange picture used in deep inelastic scattering corresponds to the Born approximation or first order in this series.

The second useful simplification is the factorization of the process into two parts: the generation of a parton from the nucleon and the hard scattering between the parton and the virtual photon. While the latter can be calculated using the approximations of perturbative quantum field theory, the former is described by distribution functions, which are determined by experiment. Such a function contains information on the probability to find a parton with certain characteristics (e. g. a quark of flavour *up* with spin *down*) carrying a fraction x_B of the total nucleon momentum[4] when probing the nucleon at a scale Q^2. Factorization is based on the assumption of an energy scale—the so-called *hard scale*—which separates the hard scattering process from the description of the nucleon content, thus the features of the low-energy structure function and the high-energy scattering do not interfere and the structure functions are universal with respect to the scattering process; this scale is a measure for the resolution with which the nucleon is probed. It turned out as a big surprise that for $Q^2 \gtrsim 1\ \mathrm{GeV}^2/c^2$ the distribution functions are nearly independent of Q^2, a phenomenon called *scaling*. This means that, to leading order, deep inelastic scattering is dominated by the scattering off point-like charged particles, which can be identified with the quarks postulated in reference [49].

The necessity for experimental measurement does not mean that nothing can be calculated when it comes to distribution functions. The qualitative behavior near the kinematic edges—i. e. in case the parton carries only a very small or very large part of the total nucleon momentum—can be deduced by requiring that the theory is renormalizable and the wave functions possess certain mathematical properties like analyticity, see [50]. But most importantly, the nucleon is composed of strongly interacting particles, which leads to relations between the distribution functions for different partons and their evolution with Q^2, the famous DGLAP equations [1, 2, 3, 4]. This set of coupled integral-differential equations enables the comparison of measurements performed at different values of Q^2 and x_B and is a key ingredient for constructing an overall picture of the structure of the nucleon.

At present, the distribution functions for the light quark flavors u, d and s, as well as their anti-particles, are reasonably well established, as are the spin distributions of u and d quarks. In the simple parton model, all contributions would sum up to the total nucleon, but it turned out that only half of the nucleon momentum is carried by quarks. In case of the spin the problem is even more severe, as only about a third is explained

[4] This implies that we envisage the nucleon in a so-called *infinite momentum frame*, which is introduced in the following section.

by summing up the quark contributions [18]. While the missing momentum has been attributed to the gluons following their discovery at PETRA [51], the quantitative composition of the nucleon spin is still an open question. Possible explanations are either a substantial polarization of the gluons or orbital angular momentum—pictorial equivalent to imagining some part of the nucleon as a kind of spinning top. It is this question, which drives the main part of the COMPASS muon program, and also the present analysis.

6.2 Deep Inelastic Scattering

The kinematic variables commonly used when treating this topic are

$$\mathrm{p} = (E, \mathbf{p}) \qquad \text{and} \qquad \mathrm{p}' = (E', \mathbf{p}') \qquad \text{with} \qquad \mathrm{p}^2 = \mathrm{p}'^2 = m^2, \tag{6.2}$$

the four-momenta of the incoming and outgoing lepton of mass m. In a fixed target experiment, the initial nucleon momentum is $\mathrm{P} = (M, \mathbf{0})$ in the laboratory frame. Another reference frame, which is used frequently in theoretical calculations, is the Breit frame, where the lepton momentum after the scattering process is the reverse of its initial momentum. While this frame is beneficial for abstract considerations, the laboratory frame is more convenient when describing the experimental aspects and thus will be used throughout this thesis unless otherwise noted. Since only the single photon exchange is considered, we can immediately determine the virtual photon's momentum as

$$\mathrm{q} = \mathrm{p} - \mathrm{p}' = (\nu, \mathbf{q}) \tag{6.3}$$

while $Q^2 = -\mathrm{q}^2 > 0$ is the virtuality of the photon.[5] The momentum of the hadronic final state is given by $\mathrm{P}' = \mathrm{P} + \mathrm{q}$, with the invariant mass

$$W = \sqrt{\mathrm{P}'^2}/c \tag{6.4}$$

The criterion for deep inelastic scattering is $W > 2\,\mathrm{GeV}/c^2$, so that the region of elastic scattering and nucleon excitations is excluded. It is convenient to define the Bjorken scaling variable

$$x_B = \frac{Q^2}{2\mathrm{P}\cdot\mathrm{q}} \overset{\text{lab frame}}{=} \frac{Q^2}{2M\nu} \tag{6.5}$$

In case of elastic scattering $x_B = 1$, so it is a measure of the (in-)elasticity of the scattering process. In a reference frame where the nucleon travels with infinite momentum, this variable can be interpreted as the momentum fraction carried by the struck parton. This frame is chosen so that transverse momenta and intrinsic masses can be neglected. Other often used variables are the relative energy transfer from the lepton to the photon

[5] $\mathrm{q}^2 = 2m^2 - 2\mathrm{p}\cdot\mathrm{p}' \overset{\text{Breit frame}}{=} 2m^2 - 2(E^2 + p^2)$, which obviously is always negative.

$$y = \frac{\mathrm{P}\cdot\mathrm{q}}{\mathrm{P}\cdot\mathrm{p}} \overset{\text{lab frame}}{=} \frac{\nu}{E} \tag{6.6}$$

In case of semi-inclusive deep inelastic scattering, also the relative energy of the produced hadron is defined as

$$z = \frac{\mathrm{P}\cdot\mathrm{h}}{\mathrm{P}\cdot\mathrm{q}} \overset{\text{lab frame}}{=} \frac{E_{\text{hadron}}}{\nu} \tag{6.7}$$

where h is the four-momentum of the produced hadron. With these definitions we can approach the cross section of deep inelastic scattering. Assuming single photon exchange, the transition matrix element for the production of the hadronic final state X factorizes in two parts:

$$T_{fi} = \frac{e^2}{\mathrm{q}^2}\langle \ell(\mathrm{p}')|j^\mu(0)|\ell(\mathrm{p})\rangle\langle X(\mathrm{P}')|J_\mu(0)|N(\mathrm{P})\rangle \tag{6.8}$$

The factors e^2 and q^{-2} stem from the virtual photon coupling and propagation, respectively, while j^μ and J_μ are the electromagnetic currents of the lepton and nucleon to which the photon couples. X denotes the unobserved hadronic final state. The probability $\mathcal{P}$ to find a given final state in the detector is proportional to the squared absolute value of the matrix element, averaging over the spin states in the initial state and summing over the spin states in the final state. In inclusive deep inelastic scattering the hadronic final state is not detected, therefore we must also sum over all possible configurations of produced hadrons.

$$\mathcal{P} \propto \frac{1}{4}\sum_{\text{spins}}\sum_{X}|T_{fi}|^2 \propto L^S_{\mu\nu}W^{\mu\nu}_S \tag{6.9}$$

This representation makes it obvious how the problem is factorized into two subproblems: $L^S_{\mu\nu}$ is the part of the leptonic tensor that is symmetric with respect to an interchange of Lorentz indexes, which is calculable in QED:

$$L^S_{\mu\nu} = 2[p_\mu p'_\nu + p_\nu p'_\mu - g_{\mu\nu}\mathrm{p}\cdot\mathrm{p}'] \tag{6.10}$$

and $W^{\mu\nu}_S$ is the symmetric hadronic tensor. Assuming pure QED scattering,[6] Lorentz and gauge invariance together with parity conservation restrict it to the general form

$$W^{\mu\nu}_S = W_1(x_B, Q^2)\left(\frac{q^\mu q^\nu}{\mathrm{q}^2} - g^{\mu\nu}\right) + \frac{W_2(x_B, Q^2)}{M^2}\left(P^\mu - \frac{\mathrm{P}\cdot\mathrm{q}}{\mathrm{q}^2}q^\mu\right)\left(P^\nu - \frac{\mathrm{P}\cdot\mathrm{q}}{\mathrm{q}^2}q^\nu\right) \tag{6.11}$$

where W_1 and W_2 are structure functions which only depend on the invariants x_B and Q^2. Contracting the Lorentz indices in eq. (6.9) and computing the kinematic factors in the laboratory frame gives

[6] This excludes weak interaction contributions which arise from Z^0 exchange.

$$\frac{d^2\sigma}{d\Omega\, dE'} = \frac{\alpha^2}{4E^2 \sin^4\frac{\theta}{2}} \left(W_2(x_B, Q^2) \cos^2\frac{\theta}{2} + 2W_1(x_B, Q^2) \sin^2\frac{\theta}{2} \right) \tag{6.12}$$

with the solid angle element $d\Omega = d\cos\theta\, d\phi$, where the scattered lepton's direction is changed by the polar angle θ in azimuthal direction ϕ with respect to the incoming lepton direction. The dimensionless structure functions $F_1 = MW_1$ and $F_2 = \frac{P\cdot q}{M} W_2$ can then be measured varying x_B, Q^2 and the scattering angle θ in the experiment. While the acceptance in x_B and Q^2 of the experiment is often fixed by the design of the trigger system, in the early SLAC experiments the scattering angle has been selected by moving the detector around the target.

6.3 Polarized Deep Inelastic Scattering

The tensors given in equations (6.10) and (6.11) are symmetric, because the spins have been averaged over. If the spin projections are treated individually, an additional term is present in the lepton tensor:

$$L^A_{\mu\nu} = 2i\epsilon_{\mu\nu\alpha\beta} s^\alpha q^\beta \tag{6.13}$$

where ϵ is the totally anti-symmetric Levi-Civita tensor[7] and s is the lepton spin vector. To produce a non-zero effect in the contraction with the hadronic tensor, the latter has to be equipped with an anti-symmetric part, too. The symmetries already mentioned in the unpolarized case—Lorentz and gauge invariance, parity conservation—restrict this to

$$W^{\mu\nu}_A = i\epsilon^{\mu\nu\alpha\beta} q_\alpha \left[G_1(x_B, Q^2) S_\beta + G_2(x_B, Q^2) \frac{1}{M^2} (S_\beta \mathrm{P}\cdot\mathrm{q} - P_\beta \mathrm{S}\cdot\mathrm{q}) \right] \tag{6.14}$$

The two new structure functions $g_1 = (\mathrm{P}\cdot\mathrm{q}) G_1$ and $g_2 = \left(\frac{P\cdot q}{M}\right)^2 G_2$ characterize the spin distributions parallel and transverse to the nucleon spin direction:

$$\frac{d^2\sigma^{\rightleftarrows}}{d\Omega\, dE'} - \frac{d^2\sigma^{\rightrightarrows}}{d\Omega\, dE'} = \frac{4\alpha^2 E'}{M\nu Q^2 E} [(E + E'\cos\theta) g_1 - 2Mx_B g_2] \tag{6.15}$$

$$\frac{d^2\sigma^{\overrightarrow{\Uparrow}}}{d\Omega\, dE'} - \frac{d^2\sigma^{\overrightarrow{\Downarrow}}}{d\Omega\, dE'} = \frac{4\alpha^2 E'^2}{M\nu Q^2 E} \left[g_1 + \frac{2E}{\nu} g_2 \right] \sin\theta \tag{6.16}$$

It is obvious that g_2 can be neglected in the case of longitudinal spin settings since $E \gg M$ and the scattering angles are typically small. But g_2 is accessible if the target spin were aligned perpendicular to the beam in the scattering plane as can be seen from eq. (6.16).

[7] This tensor is defined as $\epsilon_{0123} = 1$ and each exchange of indices changes the sign.

Here $\sigma^{\overrightarrow{\Uparrow}}$ describes the case where the projection of the spin on the outgoing lepton direction is negative.

Since the absolute normalization of the cross section is difficult experimentally—we would have to know exactly the number of beam particles crossing the target volume and equally exactly the number of polarized nucleons in the target—it is convenient to use the knowledge about the unpolarized cross section and calculate the asymmetry

$$A_{\parallel} = \frac{\Delta\sigma}{\sigma} = \frac{\sigma^{\overrightarrow{\Leftarrow}} - \sigma^{\overrightarrow{\Rightarrow}}}{\sigma^{\overrightarrow{\Leftarrow}} + \sigma^{\overrightarrow{\Rightarrow}}} \propto \frac{g_1(x_B)}{F_1(x_B)} \qquad \text{(neglecting contribution of } g_2\text{)} \tag{6.17}$$

This separation of the spin states corresponds to a measurement with 100% polarization of beam as well as target particles. In a real experiment several factors dilute the measured asymmetry as discussed in section 7.2.3. Measuring the asymmetry means normalizing to the unpolarized cross section in the experiment, which substantially simplifies the extraction of g_1 compared to the direct subtraction of the absolute polarized cross sections.

Since the polarization transfer from the incoming lepton to the exchanged virtual photon depends on the kinematic variables, an asymmetry $A_{\parallel,\gamma^*}$ which assumes 100% photon polarization is also used, including the depolarization factor D which describes the polarization transfer from the incoming lepton to the virtual photon:

$$A_{\parallel} = D A_{\parallel,\gamma^*} \tag{6.18}$$

A crude approximation is $D \approx y$, while the full expression

$$D = y \cdot \frac{\left(1 + \frac{1}{2}\xi^2 y\right)(2 - y) - \frac{2y^2 m^2}{Q^2}}{y^2\left(1 - \frac{2m^2}{Q^2}\right)(1 + \xi^2) + 2(1 + R)\left(1 - y - \frac{1}{4}\xi^2 y^2\right)} \quad \text{with} \quad \xi = \frac{2Mx_B}{Q} \tag{6.19}$$

involves the ratio of the absorption cross sections for longitudinal and transverse photons on the nucleon, $R = \frac{\sigma_L}{\sigma_T}$. This quantity has been measured e. g. at SLAC [52].

6.4 Interpretation of the Distribution Functions

While the mathematical properties of the structure functions allow calculations to be done, which then interact with the progress of experimental measurement of the deep inelastic scattering cross sections, the meaning of the formal structure of the nucleon is not self-evident. Therefore, an appealing model was searched and proposed by Feynman: the *parton model* [53]. In that model, the nucleon consists of non-interacting partons, moving collinearly with infinite momentum; transverse momenta of the quarks inside the nucleon are neglected. The formal basis for this model is the operator product expansion developed by Wilson [54]. The idea is that the structure functions are related

structure function	Γ
F_1	γ^μ
g_1	$\gamma_5\gamma^\mu$
h_1	$\sigma^{\mu\nu}\gamma_5$

Table 6.1: Current structure which corresponds to each structure function in the Bjorken limit $Q^2 \to \infty$ with $x_B \to const.$

via the optical theorem to forward scattering amplitudes, which connect the nucleon wave function at different points along the light cone:

$$f(x_B;\, Q^2) = \int \frac{d\lambda}{2\pi} e^{ix_B\lambda} \langle N(\mathrm{PS})|\overline{\psi}(0)\Gamma\psi(\lambda n)|N(\mathrm{PS})\rangle \tag{6.20}$$

where λn is a vector along the light-cone (see section A.1), Q^2 stands for the scale up to which the nucleon structure is evaluated—see below—and Γ is the current connected with the respective structure function as shown in table 6.1. These bi-local transition operators can be treated using the operator product expansion formalism to expand them into a series of terms ordered in powers of $\frac{M}{Q}$, where the constant term is called leading twist—or twist-2—and the following are disregarded in the parton model because of the assumption $Q^2 \gg M^2$, i. e. only the terms are kept which exhibit *scaling* and thus do not vanish for $Q^2 \to \infty$. Using a representation of the Dirac matrices adapted to light-cone coordinates, which is given in section A.1, it can easily be seen that the leading twist contributions to the structure functions are simply number densities of partons while the higher twist terms describe parton correlations and have no simple interpretation as such. This leads to the usual parton model identification of sums of the parton distribution functions with the structure functions:

$$\begin{aligned}
F_1(x_B;\, Q^2) &= \frac{1}{2}\sum_i e_i^2 \left(q_i(x_B;\, Q^2) + \bar{q}_i(x_B;\, Q^2)\right) \\
G_1(x_B;\, Q^2) &= \frac{1}{2}\sum_i e_i^2 \left(\Delta q_i(x_B;\, Q^2) + \Delta\bar{q}_i(x_B;\, Q^2)\right) \\
H_1(x_B;\, Q^2) &= \frac{1}{2}\sum_i e_i^2 \left(\Delta_T q_i(x_B;\, Q^2) - \Delta_T\bar{q}_i(x_B;\, Q^2)\right)
\end{aligned} \tag{6.21}$$

where Δ and Δ_T mean the difference of quark densities with spin anti-parallel and parallel to the spin of the nucleon. The subscript T signifies that the nucleon spin is oriented at right angles to the beam, contrary to the situation, where the nucleon spin is aligned with or against the beam direction; the latter has been the typical experiment setup up to now. The presence of Q^2 marks the implicit inclusion of so-called *scaling violations*, which originate from the factorization of the cross section at a certain scale: the hard process can only resolve the structure of the nucleon up to that scale, therefore probing the nucleon at a higher scale gives access to additional structure and yields a different

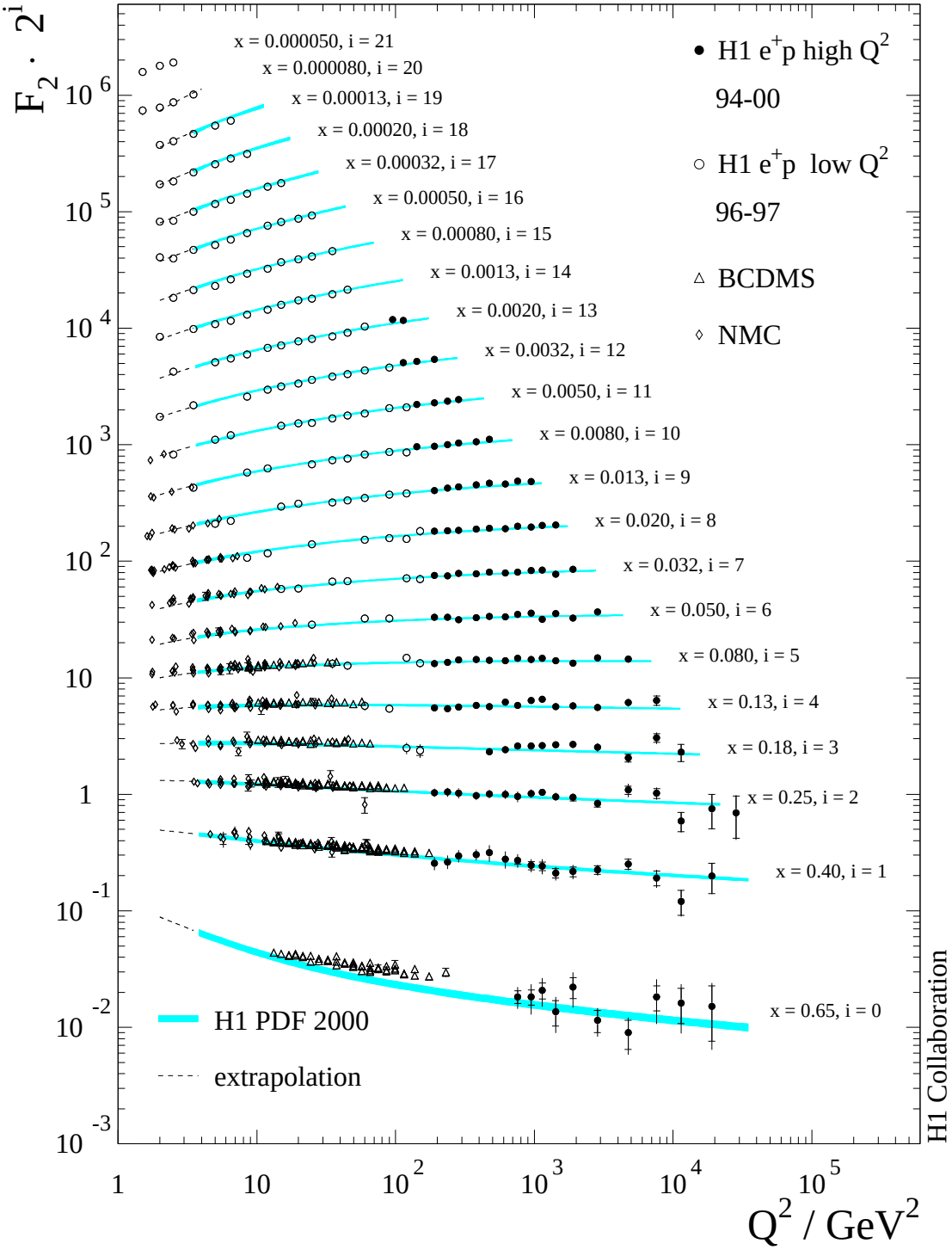


Figure 6.3: Scaling and its violation as measured at the H1 experiment at HERA. The curves and data points for the different values of x_B are scaled by 2^i with the index i given next to each value of x_B. Taken from [55].

result. These scaling violations are rather small, as is depicted in figure 6.3.

As for most of the COMPASS data Q^2 is very small, a word on this scale parameter is in order. The interpretation of Q^2 as the resolution of the virtual photon is only valid in the so-called Breit frame, where the energy transfer vanishes. In this frame the momentum of the photon is $q^* = (0, \mathbf{Q})$ so that the spatial resolution—an inherently non-Lorentz invariant quantity—would be $\frac{\hbar}{Q}$. However, for theoretical calculations Q^2 only denominates the scale of the process; this role can be played by any hard quantity which is invariant under the transformation between the laboratory frame and the Breit-Frame. Thus, in case of this high p_T analysis, p_T^2 plays the role of Q^2, while e. g. in the case of high p_T hadron pairs at $Q^2 > 1\,\mathrm{GeV}^2$ a combination of $\sum p_T^2$ and Q^2 could be used. The exact value of the scale should have no impact on the extracted physics results, since the choice of the scale is not fixed by a physical principle. But in practice a so-called *scale dependence* is observed, which usually reduces when including higher orders of the perturbation series in the calculation. It is customary to vary the scale by a factor of two in both directions and quote the resulting variance as scale uncertainty on the extracted quantity.

6.5 Sum Rules and the Spin Puzzle

The initial state of the scattering process in theory and experiment is—apart from the EMC effect mentioned in section 2.1.1—always the whole nucleon with its quantum numbers as they are known e. g. from the Stern-Gerlach experiment. These quantities cannot depend on the parameters of the scattering process used to probe the nucleon structure. On the other hand the constituents of the nucleon also have well-defined quantum numbers. In the quark parton model the quantum numbers of the nucleon are given by the first moments of parton structure functions.

Structure functions are number densities of quarks and anti-quarks with respect to certain properties, e. g. the number of quarks is given by

$$\begin{aligned}
n_u &= u^{(0)}(Q^2) - \bar{u}^{(0)}(Q^2) = \int_0^1 dx_B \left[u(x_B;\, Q^2) - \bar{u}(x_B;\, Q^2)\right] \overset{\text{proton}}{=} 2 \\
n_d &= d^{(0)}(Q^2) - \bar{d}^{(0)}(Q^2) = \int_0^1 dx_B \left[d(x_B;\, Q^2) - \bar{d}(x_B;\, Q^2)\right] \overset{\text{proton}}{=} 1 \qquad (6.22) \\
n_s &= s^{(0)}(Q^2) - \bar{s}^{(0)}(Q^2) = \int_0^1 dx_B \left[s(x_B;\, Q^2) - \bar{s}(x_B;\, Q^2)\right] \overset{\text{proton}}{=} 0
\end{aligned}$$

where $f^{(n)}$ is the n^{th} moment of f. The numbers given are the canonical quark numbers which satisfy the measured quantum numbers for the proton. While this sum rule is very

intuitive, its experimental test is very challenging, as instead of the measurable structure functions the quark parton densities enter directly. The first sum rule to give new insight in the structure of the nucleon was the Gottfried sum rule, named after [56],

$$\int_0^1 \frac{dx_B}{x_B}\left[F_2^p(x_B;\,Q^2) - F_2^n(x_B;\,Q^2)\right] = \frac{1}{3} \tag{6.23}$$

which is derived assuming that the $\bar{u}$ and $\bar{d}$ quarks in the nucleon are purely produced by gluon radiation, which is assumed to fluctuate into $u\bar{u}$ and $d\bar{d}$ pairs with equal probability. The NMC experiment at CERN has established in [57] that this sum rule is significantly violated, leading to the conclusion that the $\bar{u}$ and $\bar{d}$ distributions are not identical, the integral of $\bar{u} - \bar{d}$ being of the order of 0.15.

As for the spin structure of the nucleon, the most fundamental sum rule is that of Bjorken [58], which is based solely on current algebra and the universality of quarks with respect to the strong and weak interactions. Neglecting the contributions from the heavy quark flavors c, b and t, this sum rule is

$$\begin{aligned} S_B &= \int_0^1 dx_B\,\left[g_1^p(x_B;\,Q^2) - g_1^n(x_B;\,Q^2)\right] \\ &= \left[1 - \frac{\alpha_S(Q^2)}{\pi} - 3.583\left(\frac{\alpha_S(Q^2)}{\pi}\right)^2 - 20.215\left(\frac{\alpha_S(Q^2)}{\pi}\right)^3\right]\frac{g_A}{6} \end{aligned} \tag{6.24}$$

The connection with the weak axial-vector coupling constant g_A comes from the fact that the proton and neutron states are connected by the weak interaction as exhibited e. g. by the β-decay of the neutron. This is also visible when expanding the sum rule in the quark parton model as defined in eq. (6.21). The correction terms in eq. (6.24) are caused by the higher order contributions which are neglected in the quark parton model. This sum rule has been confirmed experimentally by the SLAC experiments E142/E143 and the CERN experiments EMC and SMC.

One drawback of the Bjorken sum rule is that it involves separate measurements on the proton and neutron, making it difficult to control the systematic uncertainties associated with comparing data taken with different targets. Decomposing the first moment of g_1^p in the quark parton model gives

$$\begin{aligned} g_1^{p\,(0)}(Q^2) &= \int_0^1 dx_B\, g_1^p(x_B;\, Q^2) \\ &= \int_0^1 dx_B \left[\frac{4}{18}\Delta u(x_B;\, Q^2) + \frac{1}{18}\Delta d(x_B;\, Q^2) + \frac{1}{18}\Delta s(x_B;\, Q^2)\right] \\ &\equiv \frac{4}{18}\Delta u + \frac{1}{18}\Delta d + \frac{1}{18}\Delta s \end{aligned} \tag{6.25}$$

using the notation

$$\Delta q(x_B;\, Q^2) \equiv q^{\uparrow}(x_B;\, Q^2) - q^{\downarrow}(x_B;\, Q^2) + \bar{q}^{\uparrow}(x_B;\, Q^2) - \bar{q}^{\downarrow}(x_B;\, Q^2) \tag{6.26}$$

We now express the first moment of g_1^p in the axial-vector currents of the proton

$$\begin{aligned} \Delta q_3 &= \Delta u - \Delta d = g_A && \propto \langle p\,|\bar{\psi}\gamma_\mu\gamma_5\lambda_3\psi|\,p\rangle \\ \Delta q_8 &= \Delta u + \Delta d - 2\Delta s && \propto \langle p\,|\bar{\psi}\gamma_\mu\gamma_5\lambda_8\psi|\,p\rangle \\ \Delta q_0 &= \Delta u + \Delta d + \Delta s && \propto \langle p\,|\bar{\psi}\gamma_\mu\gamma_5\psi|\,p\rangle \end{aligned} \tag{6.27}$$

using the Gell-Mann matrices λ_i (see section A.2). This leads to

$$g_1^{p\,(0)}(Q^2) = \frac{\Delta q_3}{12} + \frac{\Delta q_8}{36} + \frac{\Delta q_0}{9} \tag{6.28}$$

Ellis and Jaffe [59] evaluate this under the assumption that the contribution from the s quarks can be neglected, making $\Delta q_8 = \Delta q_0$. This yields the sum rule

$$g_1^{p\,(0)}(Q^2) = \frac{\Delta q_3}{12}\left[1 - \frac{\alpha_S(Q^2)}{\pi}\right] + \frac{5\Delta q_8}{36}\left[1 - \frac{7\alpha_S(Q^2)}{15\pi}\right] \tag{6.29}$$

The QCD corrections to this sum rule have been derived in [60] up to the order α_S^3. For the neutron only the isovector part Δq_3 changes sign, so taking the difference of the first moments of proton and neutron gives the Bjorken sum rule again. Using the current value $g_A = 1.270 \pm 0.003$ [61] and assuming that Δq_8 is known from hyperon decay measurements[8] to be 0.58 ± 0.02, the expected value is $g_1^{p\,(0)}(10\,\mathrm{GeV}^2) = 0.175 \pm 0.003$, however, the EMC experiment has measured a substantially lower value, which has been improved in accuracy by the E143/E155 experiments and SMC [16]. The value extracted by the latter is

$$g_1^{p\,(0)}(10\,\mathrm{GeV}^2) = 0.12 \pm 0.01 \tag{6.30}$$

[8] The assumption is that flavor $SU(3)$ can be used to describe the decays of strange baryons. It is known that this symmetry is not exact, so the uncertainty on Δq_8 is probably underestimated.

which is clearly in violation of eq. (6.29). This finding has been labeled *spin puzzle*, and it is not yet clear which of the assumptions made by Ellis and Jaffe are wrong. The most obvious possibility is to allow for a difference between Δq_8 and Δq_0, which in the quark parton model would mean that the strange quark content of the proton is polarized. Working up to order α_S the sum rule then is

$$g_1^{p\,(0)}(Q^2) = \left(\frac{\Delta q_3}{12} + \frac{\Delta q_8}{36}\right)\left[1 - \frac{\alpha_S(Q^2)}{\pi}\right] + \frac{\Delta q_0}{9}\left[1 - \frac{\alpha_S(Q^2)}{3\pi}\right] \tag{6.31}$$

Using the values given above we can extract $\Delta q_0 = 0.01 \pm 0.01$ as the contribution $\Delta\Sigma$ of the quarks to the spin of the nucleon at a scale of $10\,\mathrm{GeV}^2$. This value is surprisingly small compared to the naïve expectation of 1—in the style of the constituent quark model—or 0.58, which would be the parton model expectation for $\Delta s = 0$. However, the identification of Δq_0 and $\Delta\Sigma$ depends on the quark parton model, which does not allow for gluons to contribute to the quantum numbers of the proton. Altarelli and Ross [62] come to the conclusion that a sizable gluon polarization could compensate $\Delta\Sigma$ in the flavor-singlet term

$$\Delta q_0 = \Delta\Sigma - \frac{3\alpha_S(Q^2)}{2\pi}\Delta G(Q^2) \tag{6.32}$$

It should be noted that the suppression of the gluonic contribution by $\alpha_S(Q^2)$, for $Q^2 \to \infty$ is canceled by the evolution of same, so that this result is fundamentally different from the simple quark parton model. Current QCD fits in next-to-leading order by the COMPASS collaboration indicate that $\Delta\Sigma = 0.30 \pm 0.02$ at $Q^2 = 4\,\mathrm{GeV}^2$, with a contribution of $|\Delta G| \approx 0.2 \div 0.3$ from the gluons. This result makes an independent measurement of the gluon polarization even more desirable. Therefore the remainder of this chapter is devoted to the three ansatzes currently pursued by COMPASS to this end.

6.6 Parton Type Separation

The method of inclusive deep inelastic scattering as discussed so far implicitly sums over the contributions stemming from different types of partons, which within the standard model are the six quark flavors and the gluons. Of course, the question of which parton type contributes how much has soon been tackled with various complementary experiments. While deep inelastic neutrino reactions can distinguish between u- and d-type quark flavors and their anti-particles, semi-inclusive deep inelastic scattering has been used to tag valence, strangeness and charm content. The general semi-inclusive deep inelastic scattering graph is shown in figure 6.4, where the central blob hides the perturbative calculation of the hard scattering process of partons a and b into the parton c, which fragments into the observed final state hadron, plus the unobserved hadronic rest. Figure 6.5 depicts the ideal reaction for extracting information on the gluon polarization within the nucleon, often called *photon–gluon fusion*. Since the photon and gluon do

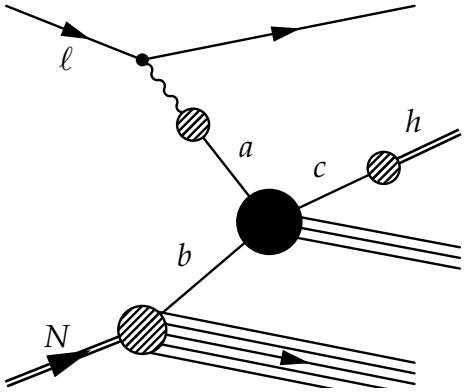


Figure 6.4: General Feynman graph representing single-inclusive deep inelastic scattering. The particles *a*, *b* and *c* represent different parton types taken from the photon (*a* is the photon in the non-resolved case) and nucleon or fragmenting into the observed final state hadron.

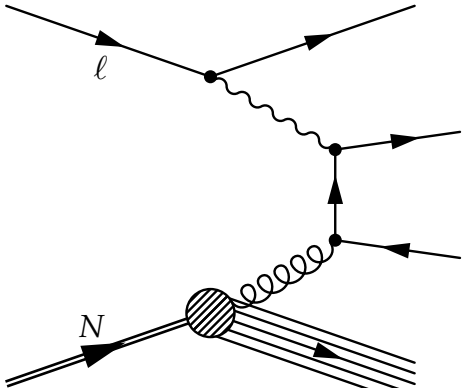


Figure 6.5: Leading order graph showing the most interesting sub-process for the extraction of the gluon polarization; the fragmentation of the charm quarks into hadrons is not shown.

not couple directly—the photon lacks strong charge in the same way as the gluon has no electromagnetic charge—a quark line is the only possible direct connection. If the available energy would be high enough, two hadron jets would be produced back-to-back from the two quarks, but as the center of mass energy at COMPASS amounts only to $\sqrt{s} = 17.3$ GeV, these two jets consist in most cases only of a handful of particles.

As stated earlier, nature tends to complicate things by superimposing all possible processes where we do not discriminate by measurement, leading to a dilution of the asymmetry measured in photon gluon fusion by the unrelated or much smaller asymmetries of all the other processes producing the same event signature. There are three ways to overcome this difficulty: the purity ansatz presented in the next section, which was successfully used by the SMC experiment, the open charm channel introduced in section 6.6.2, and the complete calculation of all contributing processes in next-to-leading order which was recently done by the group of A. Schäfer in Regensburg and is presented in section 6.6.3.

6.6.1 Purity Analysis of SIDIS Asymmetries

As the name implies, this method describes the measured experimental asymmetry as a sum of the diluted photon–gluon fusion asymmetry—see figure 6.5—and impurities created by other graphs. The ratio of the contribution from the different graphs is taken from a LUND [63] Monte Carlo simulation, while the analyzing power—the asymmetry created by inserting 100% gluon polarization into the photon gluon fusion graph—is calculated perturbatively. The most prominent background process is the LO DIS process depicted in figure 6.1 on page 50, the asymmetry of which has been measured by the COMPASS collaboration [18], among others. In case of low Q^2, additional asymmetric

background processes become important, most notably the resolved photon contributions which are discussed in more detail in section 6.6.3.

The perturbative calculation of the analyzing power requires a factorization scale to separate the non-perturbative low energy distribution function from the hard scattering process. In the absence of large Q^2, a high transverse momentum of the outgoing quarks can be used, since the average transverse momentum present in the nucleon is known to be of the order of only a few hundred MeV (the parton model does not incorporate the concept of intrinsic transverse momentum at all). Since the transverse momentum is invariant under the transformation from the Breit frame to the laboratory frame, the corresponding wavelength can also be interpreted as the resolution of the probe; a higher transverse momentum ensures higher locality of the scattering process.

The intrinsic transverse momentum of course attenuates gradually, leaving some freedom for optimization of the cuts to enrich the photon–gluon fusion in the sample. It has been found that the working point is given when hadron pairs are selected having $p_T > 0.7\,\mathrm{GeV}/c$ for each particle and $\sum p_T^2 > 2.5\,\mathrm{GeV}^2/c^2$. The result of this analysis has been published in [64].

6.6.2 Open Charm Production

COMPASS was especially designed to detect D^0 mesons in deep inelastic scattering, owing to the assumption that the intrinsic charm content of the nucleon can safely be neglected, so the detection of charmed mesons in the final state can, at leading order, only be attributed to the photon gluon fusion graph 6.5. Another picture is that charm loops inside the nucleon must be short-lived as prescribed by the uncertainty principle, thus the charm quarks are strongly correlated with their gluon parent; here the charm mass ensures locality and defines the resolution of the scattering process, so that no large Q^2 cut is required in the analysis. While being considered the golden channel for measuring the gluon polarization due to its lack of leading order background processes, it is experimentally extremely challenging, not least because of the small production cross section. This is reflected in the rather limited statistics obtained in the D^0 sample. Tagging the D^0 mesons coming from the D^{*-} and D^{*+} decays using the produced slow pions enhances the signal to background ratio dramatically and provides an addition to the golden channel. The estimation of the statistical uncertainty on $\Delta G/G$ achievable with the final COMPASS data sample using a deuteron target is 0.28.

6.6.3 Calculation of SIDIS Asymmetries in Next-to-Leading Order

Another ansatz is to calculate the asymmetries arising from all possible single-inclusive deep inelastic scattering processes for the actual experiment kinematics, in several scenarios for the gluon polarization, and to compare the obtained results with the measurement. Single-inclusive means that only exactly one hadron emerging from the hard

scattering process is considered, in contrast to semi-inclusive, which can mean anything between exclusive (complete hadronic final state) and inclusive (only scattered lepton). Calculations have shown that next-to-leading order corrections are sizable at intermediate values of p_T and are different for the polarized and the unpolarized cross section. Thus, a full next-to-leading order calculation of single-inclusive scattering in the COMPASS kinematics was done by B. Jäger, M. Stratmann and W. Vogelsang in [48]. In the following the key points of the paper are summarized.

The general single-inclusive deep inelastic scattering graph 6.4 contains four blobs: the hashed blobs denote the non-perturbative parts of the cross section, namely the structure functions of the photon $f_a^\gamma(x_a;\ \mu_f)$, the structure functions of the nucleon $f_b^N(x_b;\ \mu_f)$ and the fragmentation function $D_c^h(z_c;\ \mu_{f'})$ which describes the formation of the observed hadron h from the outgoing parton c, while the filled blob hides the sum of all possible hard scattering matrix elements $d\sigma_{ab\to cX}(S, x_a, x_b, P_h/z_c;\ \mu_f, \mu_{f'}, \mu_r)$; in the polarized case these four quantities are designated with a Δ. The μ_X are the scales introduced for calculative reasons as discussed in section 6.4. The hard scattering matrix element is fully calculated in perturbative QCD up to next-to-leading order in $\alpha_S(Q^2)$, the nucleon structure—apart from the gluon polarization—is sufficiently well known and the distribution functions have been measured at HERA, SLAC, NMC and SMC. The structure of the photon consists of a "direct" part, for which parton a simply is the photon, and a "resolved" part, where quantum fluctuations lead to quark and gluon content. At large momentum fractions x_a, the perturbatively calculable "point-like" contribution dominates the photon structure, but at small x_a the polarized photon structure is presently unknown. Therefore maximal and minimal saturation of the positivity constraint on the parton density are explored, leading to $|\Delta f^\gamma(x;\ \mu_f)| \leq f^\gamma(x_a;\ \mu_f)$. Effects of QED radiation in the initial or final states are not included, see section 7.4.3 for a discussion. The complete cross section then is

$$d\Delta\sigma = \sum_{a,b,c} \int dx_a\, dx_b\, dz_c\, \Delta f_a^\gamma(x_a;\ \mu_f) \Delta f_b^N(x_b;\ \mu_f) \\ d\Delta\sigma_{ab\to cX}(S, x_a, x_b, P_h/z_c;\ \mu_r, \mu_f, \mu_{f'}) D_c^h(z_c;\ \mu_{f'}) \quad (6.33)$$

The structure functions have been marked up with the relevant factorization scales μ_f for the parton content of photon and nucleon and $\mu_{f'}$ for the fragmentation into the hadron h. This means that momenta below this scale are treated non-perturbatively by attributing the effects to the respective structure functions, while momenta above this scale enter the perturbative calculation of the hard scattering process. As usual there is also a renormalization scale μ_r for the hard process. In the calculation, the three scales have been varied collectively in the range $p_T/2 \leq \mu_f = \mu_{f'} = \mu_r \leq 2p_T$. The resulting cross sections in LO and next-to-leading order are shown in figure 6.7. While in the polarized case—especially at higher values of p_T—the scale dependence is substantially reduced, the unpolarized cross section does not profit from going to the next order. Another indicator for estimating the applicability of a certain order in a perturbation series

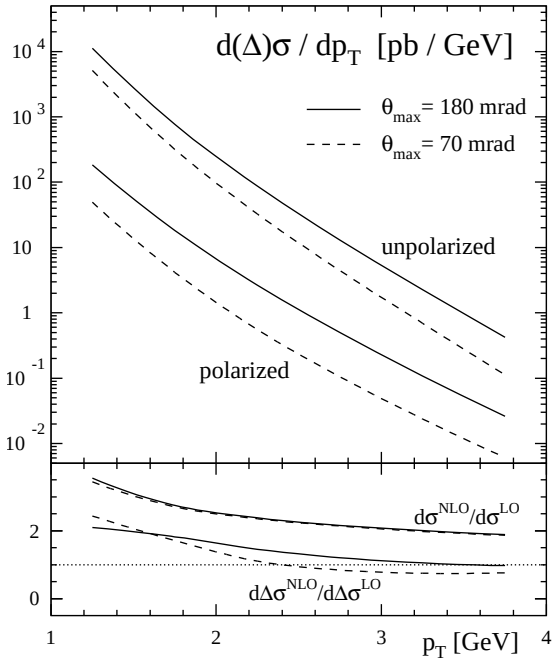


Figure 6.6: Unpolarized and polarized p_T-differential cross sections at next-to-leading order for the reaction $\mu d \to \mu' \pi^0 X$ for SMC ($\theta_{\max} = 70\,\text{mrad}$) and COMPASS ($\theta_{\max} = 180\,\text{mrad}$) target magnets. The lower panel shows ratios of next-to-leading order to LO results (K-factors). Taken from [48].

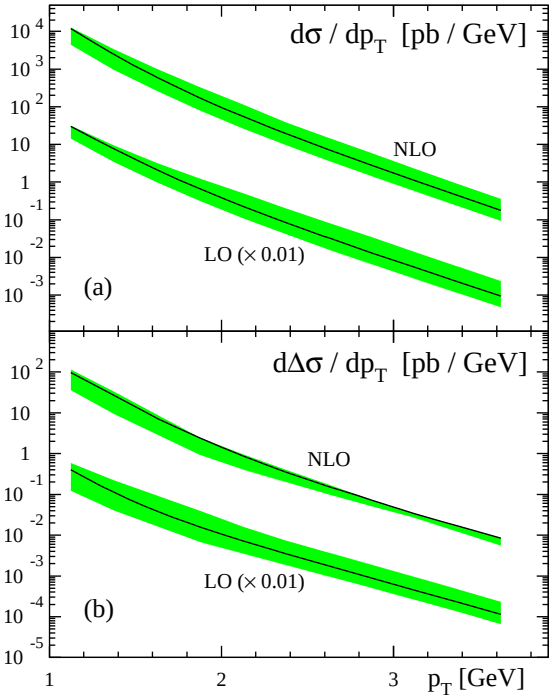


Figure 6.7: Polarized and unpolarized cross section in LO and NLO like in figure 6.6. The green bands reflect the scale dependence of the cross sections, while the solid lines correspond to the choice where all scales are set to p_T. All LO results are scaled by 0.01. Taken from [48].

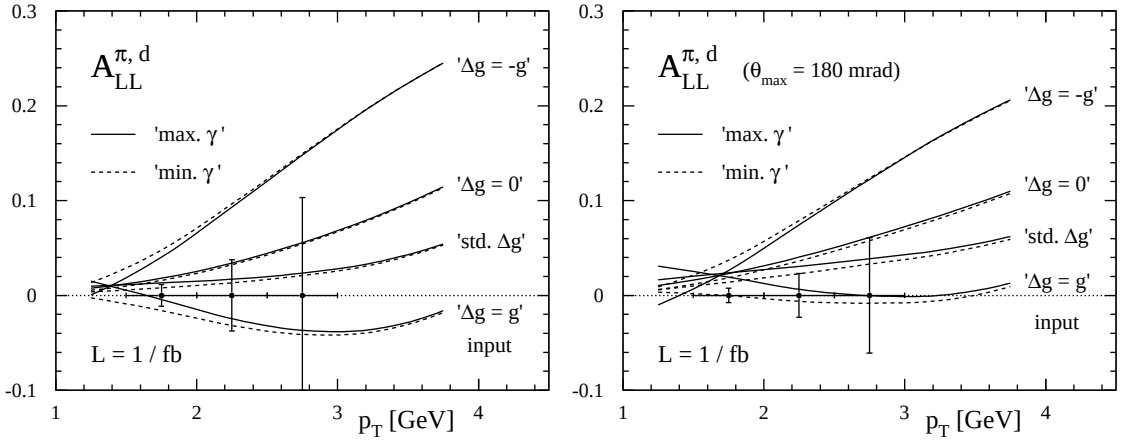


Figure 6.8: Double-spin asymmetry $A_{LL}^{\pi^0,d}$ at next-to-leading order for different gluon polarizations in the nucleon and minimal and maximal saturation of the polarized photon densities (see text). The error bars indicate the estimated statistical uncertainty for $\mathcal{L} = 1\,\mathrm{fb}^{-1}$ with COMPASS kinematics. The right plot assumes the new COMPASS target magnet, which has not been used for the data taking periods of 2002 to 2004. Taken from [48].

is the correction by the next order expressed as a factor relative to the lower order, the so-called K-factor shown in the lower panel of figure 6.6. Also here, the polarized series converges nicely for $p_T > 2\,\mathrm{GeV}/c$, while in the unpolarized case the inclusion of the next-to-leading order seems to double the cross section. The rise of the K-factors for $p_T < 2\,\mathrm{GeV}/c$ is interpreted as the break-down of the perturbation series.

Summarizing the interpretative power of the calculation, from all theoretical indicators the polarized cross section seems to be well calculable, while the unpolarized one might still be underestimated in NLO, as the authors remark in [48]. Therefore, the predicted asymmetries contain a scaling uncertainty, leading to large systematic uncertainties for a potential extraction of ΔG if the measured asymmetries are large.

All calculations have been done for the COMPASS kinematic boundaries, using the range $0.2 < y < 0.9$ and restricting the angle between the produced hadron momenta and the beam axis to $\theta < 70$ mrad as an approximation to the emittance of the target dipole magnet. Furthermore, the virtual photons are described by a polarized Weizsäcker-Williams equivalent photon spectrum

$$\Delta P_{\gamma\ell}(y) = \frac{\alpha}{2\pi}\left\{\left[\frac{1-(1-y)^2}{y}\right]\ln\frac{Q^2_{\max}}{m_\ell^2 y^2} + 2m_\ell^2 y^2\left(\frac{1}{Q^2_{\max}} - \frac{1-y}{m_\ell^2 y^2}\right)\right\} \tag{6.34}$$

with a cut at $Q^2_{\max} = 0.5\,\mathrm{GeV}^2/c^2$. The resulting double-spin asymmetry $A_{LL}^{\pi^0,d}$ for the production of neutral pions is shown in figure 6.8. As an outlook for the coming years it shows also how the asymmetries would change in view of the anticipated COMPASS target solenoid, which increases the angular acceptance from $\theta_{\max} = 70$ mrad to $\theta_{\max} = 180$ mrad. While the statistical precision increases as the cross section rises by a factor $4 \div 5$, the sensitivity to ΔG seems to be diluted slightly. This is attributed to the increasing importance of resolved photon processes at larger angles.

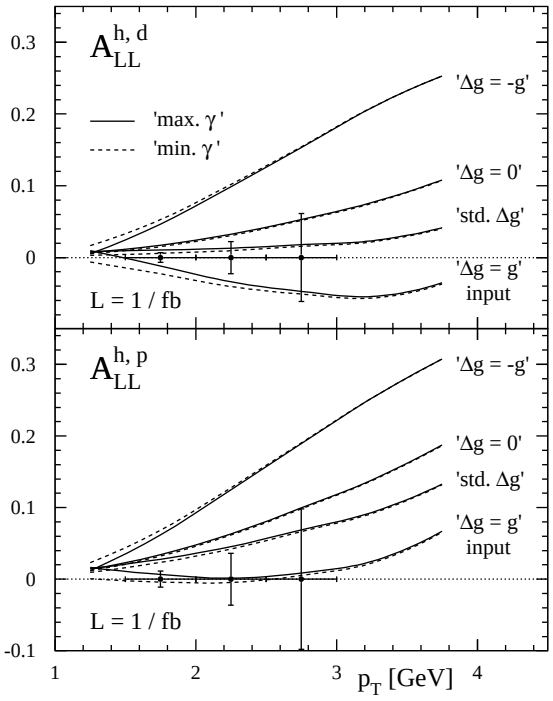


Figure 6.9: Double-spin asymmetry $A_{\mathrm{LL}}^{h,d}$—single-inclusive production of charged hadrons (p, $\pi^{\pm}$, $K^{\pm}$)—at next-to-leading order for different gluon polarizations in the nucleon, and minimal and maximal saturation of the polarized photon densities, dashed and solid lines, respectively (see text). The error bars indicate the estimated statistical uncertainty obtained with a luminosity of $\mathcal{L} = 1\,\mathrm{fb}^{-1}$ with COMPASS kinematics. Taken from [48].

p_T	$\delta A_{\mathrm{LL}}^{\pi^0,d}$	$\delta A_{\mathrm{LL}}^{h,d}$	ratio
1.75	0.01129 ± 0.00035	0.00653 ± 0.00047	1.728 ± 0.135
2.25	0.03739 ± 0.00035	0.02240 ± 0.00047	1.670 ± 0.038
2.75	0.10335 ± 0.00035	0.06112 ± 0.00047	1.691 ± 0.014
average			1.689 ± 0.013
ratio of cross sections			2.852 ± 0.075

Table 6.2: Ratio of estimated *statistical uncertainties* on single-inclusive spin asymmetries for production of neutral pions or charged hadrons, respectively. The uncertainties have been extracted from figures 6.8 and 6.9, using the EPS source code, in which the graphical representation of the error bars is stored with a precision of 0.25 pt, wherefore a quantization error is incurred as quoted above. Thus, the uncertainties given for the asymmetry errors are not Gaussian, but represent ranges. Nevertheless, the standard weight of $1/\delta^2$ has been used when averaging, and Gaussian error propagation has been employed to estimate the uncertainty on the ratio of cross sections, which is given by the square of the ratio of statistical uncertainties.

The most interesting plot concerning COMPASS analysis is shown in figure 6.9 and depicts the single-spin asymmetry $A_{\mathrm{LL}}^{h,d}$ for the single-inclusive production of charged hadrons. Neutral pion fragmentation functions are known to a higher precision, but before the installation of ECAL1 in 2006, COMPASS could not efficiently reconstruct these particles; instead it has excellent reconstruction features for charged particle tracks. The denomination "single-inclusive hadrons" is not to be confused with "leading hadrons", as the calculation via integration over the fragmentation function cannot incorporate the concept of a leading particle. This means that in the analysis one cannot select one specific hadron produced in a high p_T event but instead all hadrons connected to the production vertex must be counted. Therefore, each event may contribute multiple entries to all produced histograms.

6.6.3.1 Estimation of Unpolarized Cross Section for Charged Hadrons

While the published calculations only include an unpolarized cross section for the production of neutral pions,[9] statistical uncertainties on the experimental asymmetries are estimated also for charged hadrons using

$$\delta A_{\mathrm{LL}}^{X,N} = \frac{1}{P_{\mathrm{beam}} P_{\mathrm{beam}} f \sqrt{\sigma_{\mathrm{bin}} \mathcal{L}}} \tag{6.35}$$

Thus, the cross section for single-inclusive production of charged hadrons can be estimated by scaling the one for neutral pions with the squared ratio of estimated statistical uncertainties obtained for both cases. The procedure incurs an additional error caused by the finite resolution with which the error bars are represented in the article. The result is summarized in table 6.2. A p_T dependence of the ratio between charged hadrons and π^0 is not seen, therefore a factor of 2.85 ± 0.08 will be assumed independent of the transverse momentum.

The naïve expectation for this factor would be somewhat larger than two, since the difference between π^0 and charged pions should be small, thus charged pions contribute double the cross section, while the contribution of kaons and protons should be smaller because of their higher masses.

[9] The authors felt that especially the fragmentation functions for kaons and protons are not well constrained by the LEP data, which dominates their extraction. An improved global analysis, including pp data, is becoming available [65] and will be used to revisit single-inclusive hadron production at fixed-target experiments.

Chapter 7

High p_T Analysis

As described in the previous chapter, the goal of the current analysis is to extract hadron production asymmetries from the COMPASS data, and compare them to the calculations presented in section 6.6.3. However, it is not clear a priori that the calculations fit to the observations, unless they are successfully tested in a regime where they are expected to be correct. The biggest unknowns entering the asymmetries are the polarized structure functions of the nucleon and the photon, which thankfully do not have a significant influence on the unpolarized cross sections. Therefore it would be very surprising to find discrepancies between the prediction and measurement unless there is a deeper problem in the calculation. The analysis thus first turns to the unpolarized cross section before going on to extract the production asymmetry.

This chapter describes all details of the analysis performed and the obtained results, starting at the MDST level. One important part of the analysis process is data handling, which is detailed in the first section. The second section introduces the different experimental methods and algorithms used in asymmetry extraction, followed by the description of the data sample in section three. The unpolarized single-inclusive hadron production cross section is presented in section four, leading up to the measured production asymmetry in the second-to-last section, followed by a discussion of the result.

7.1 Analysis Tool-Chain

Due to the rare nature of the events from which the high p_T asymmetries are extracted, it is necessary to process all available data to get a meaningful value. During debugging of the procedures and algorithms this has to be repeated frequently, which would mean the processing of 8.1 TiB[1] of MDST files each time. Therefore the data are distilled in three steps, each time reducing the data size by orders of magnitude while retaining as much information as possible.

[1] The total size of the MDST for all longitudinal data of the data taking periods of 2002–2004, for further details see table E.2 on page 165.

7.1.1 μDST

The first step is done using the PHAST analysis package. A custom made UserEvent function selects events using very general criteria and writes a custom ROOT tree (μDST) containing hadron candidates. Since only the kinematic parameters are saved which might be interesting in later steps, and this only for the selected events, these μDSTs are about a factor 40 smaller than the mDSTs. For the exact criteria see section 7.3.2, however the purpose of the selection can be summarized by the following qualities of the chosen events:

- the event contains exactly one primary vertex with incoming and outgoing muon
- the beam muon's extrapolated trajectory crossed the whole length of both target cells
- there is at least one track associated to the vertex which has $p_T > 1$ GeV/c

The first condition ensures that all hadron candidates, which will be investigated later, must have originated in this vertex. This is necessary since COMPASS—being a fixed target experiment with very small scattering angles—sometimes attaches charged tracks to more than one vertex because the CDA allows this, see section 2.3.2 for details. The second condition means that both target cells are exposed to the same beam luminosity, which is essential for the asymmetry extraction methods described in section 7.2. The third is a cut which is required later in any case and reduces the statistics by more than one order of magnitude, thereby enabling much faster processing of the μDSTs.

7.1.2 Multidimensional Hadron Histogram

To get a clean hadron sample, cuts have to be imposed on more than a dozen different quantities. On the other hand it is necessary to compare different cut sets for tuning, cross checking and error estimation. Processing all μDSTs also takes a few hours, so a further step of data distillation is made: taking into account previous experience and adding some educated guesses, several potentially interesting cuts on the various quantities can be foreseen, which are interdependent in any imaginable way. This leads to the necessity to save this complete cut super-set as a multidimensional histogram. The axes represent expressions of the variables stored in the μDST on which cuts are imposed, and the binning gives the values on which cuts shall be possible when analyzing the histogram.

As an example, the angle θ_{beam} between the incoming muon direction and the outgoing hadron momentum is restricted in the theoretical calculation to 70 mrad, but considering the larger target magnet available from 2006 on, also a cut at 180 mrad has been studied. In addition it was interesting to look at the asymmetries without any such cut. This leaves the θ_{beam} axis with a binning $[0; 0.07; 0.18; \pi/2]$. A histogram axis can also contain logical conditions like $E_{\text{HCAL1}} > 0.6pc$ and arbitrarily complex combinations thereof. Variables of the μDST tree which are not part of the histogram axes are not subject to cuts.

Like in a normal ROOT histogram, the data are organized by bins, each being identified by its consecutively numbered position along the axes. Thus, each bin is represented by a tuple of integer coordinates.

To retain a high performance access to the data while making the histogram bigger in dimensions and total number of bins, the contents is stored in a relational database. The advantage of this approach is that such systems are readily available and have been extensively tested for correctness, in contrast to a private attempt at a better solution to the problem of fast searching in large data sets; MYSQL was chosen by personal preference. Relational databases store data in tables between which arbitrary $M : N$ relations can be formulated using the STRUCTURED QUERY LANGUAGE (SQL). The central part of the high p_T database is the table named data, which holds for each bin of the multi-dimensional histogram the number of entries and the sums of weights etc. needed for asymmetry and uncertainty calculation. This table is indexed[2] using a key which consists of

prod_id Every time the basic data (μDSTs) or the cut super-set changes, a new production ID is allocated to group together the histograms associated with that data set. The IDs are stored in the table named production together with a comment.

run_nb To allow for changes in the run grouping (e. g. to compare consecutive and global configurations) the histograms of the runs are kept separate instead of merging them right away. The table named run_groups contains the relation between different run groupings and run numbers.

bins_id All histograms of a specific production (as per prod_id) have the same binning, thus the coordinate tuples of the filled bins are stored in the table bins_id, which is indexed using a unique integer number for each tuple. This integer is used for accessing the data table instead of the coordinate tuple because an integer is easier to handle and faster to find than an object representing many different coordinates (e. g. concatenated numbers in a string).

For the final production a 9-dimensional histogram with 12960 bins has been used which resulted in a compression of data volume relative to the μDST size of about a factor 4,000.[3] During analysis, 13 dimensions with up to 2.5 million bins were necessary, reducing the compression ratio to about 20. Processing this data set takes about one minute,

[2] An index is a supplementary data structure automatically maintained by the database server, which allows efficient access to individual rows of the table, usually scaling logarithmically with the number of rows. There may be different keys for selecting rows, each of which needs a separate index. Accessing a table without an appropriate index causes a so-called *scan*, which scales linearly with the number of rows.

[3] On average 150 bins were populated per run, together with a storage size of 80 B resulting in 11 kiB per run which is to be compared with the average μDST size of 42.5 MiB per run

while the flexibility to change the algorithm for asymmetry extraction is retained as well as some freedom to choose different cuts from the same cut super-set. Debugging the algorithms has taken a considerable amount of time, for which this additional layer of data concentration has been very beneficial. The configuration data necessary to use and populate the database are also stored there. This way the administration software only has to integrate one library and the expressive power of SQL can be utilized also on these data.

The registration of the data with the database is performed by the custom program analyze, which reads μDSTs run by run, checks if there are already data from that run in the database, reads the histogram binning from the database, fills the histogram and writes the occupied bins into the database.

7.1.3 Hierarchical Result Tree

Once the data are in the database, they can be combined in any conceivable way, choosing different sub-samples, methods of asymmetry extraction, run groupings, and so on. This combination is done in the last step. The cut set (so-called configuration) is read from the database, all bins matching this cut set are selected and then for each run grouping (also read from the database) all bins for upstream and downstream cell are separately summed up. The resulting number of hadrons, sum of weights, etc. are then fed into the different algorithms described in section 7.2 and yield one measurement per run group and algorithm. These measurements are inserted into a tree of values, organized by cut set, asymmetry extraction method, p_T bin, microwave setting, period and run group. At each node in this tree the average and uncertainty of all measurements below this node can be retrieved, allowing a simple way to compare e. g. different settings or effects of microwave reversal.

Since this tree is relatively small it can be saved as a text file in human readable form. This is very convenient for cross checks with different people as text modification tools like Perl provide easy ways to accommodate any exchange format. Another advantage is that the text file can be copied easily to a notebook computer to enable the quick creation of different collections of results or the browsing of the results down to the run group level.

The management and the display of data from the hierarchical result tree is done by the Histo histogram manager class. Having started out simply as a histogram container and creation factory, this library has been extended to produce the asymmetry plots shown in section 7.5. For more details about Histo see section C.3.

7.1.4 Internal Cross Checks

The procedure described in this section allows a rather flexible and fast analysis of the data, but it comes for the price of increased complexity and more possibilities for error.

		polarization	
		+	−
solenoid	+	$d\sigma^{\overrightarrow{\Rightarrow}}$	$d\sigma^{\overrightarrow{\Leftarrow}}$
current	−	$d\sigma^{\overrightarrow{\Leftarrow}}$	$d\sigma^{\overrightarrow{\Rightarrow}}$

Table 7.1: Connection between polarization value signs and spin directions or rather polarized cross sections.

Therefore a short-cut has been implemented to compare the result to a much simpler one-step analysis for a given cut set. In practice this means that the cuts have been hard-coded into the PHAST user function and a printout has been added. The careful comparison of the output of this direct counting to the summation after going through the database shows no inconsistencies—all hadron counts are exactly equal.

7.2 Asymmetry Extraction

In the case of COMPASS, asymmetries are easier to measure than absolute cross sections, owing to the special design of the apparatus. This section details the methodology and the input quantities for asymmetry extraction.

7.2.1 Polarization Signs and Spin Orientation

As explained in section 2.2.5, for positive target polarization the deuteron spins are preferentially aligned with the magnetic field lines of the target solenoid. The field lines are parallel to the beam axis if the current in the coil is negative and anti-parallel in case of a positive current. This, together with the polarization of the beam in the upstream direction (see section 2.2.2), yields the connection between the various signs and the cross section shown in table 7.1. Keep in mind that the table has to be applied to each target cell separately and that the name of the microwave setting ('+' or '−') is derived from the polarization of the upstream target cell.

7.2.2 Effect of the Deuteron D-Wave Component

The wave function of the deuteron, containing a proton and a neutron and usually interchanged mesons to describe the interaction, has a total orbital angular momentum of $1\hbar$. This can be constructed from the aligned spins of the two nucleons using either a S-wave or a D-wave, while a P-wave is not possible due to symmetry considerations. As the decomposition of the two contributions is not directly observable, the probabilities for the two states are extracted using models. An evaluation of different models comes to the conclusion, that the probability to find the deuteron in a D wave is $P_D = 6.0 \pm 0.3\%$ [66].

The deuterons in the target material are aligned to the target solenoid magnetic field in three possible configurations:

$$|1,+1\rangle = \sqrt{1-P_D}|S,+1\rangle + \sqrt{P_D}\left(\sqrt{\tfrac{3}{5}}|D,-1\rangle - \sqrt{\tfrac{3}{10}}|D,0\rangle + \sqrt{\tfrac{1}{10}}|D,+1\rangle\right) \quad (7.1)$$

$$|1,0\rangle = \sqrt{1-P_D}|S,0\rangle + \sqrt{P_D}\left(\sqrt{\tfrac{3}{10}}|D,-1\rangle - \sqrt{\tfrac{2}{5}}|D,0\rangle + \sqrt{\tfrac{3}{10}}|D,+1\rangle\right) \quad (7.2)$$

$$|1,-1\rangle = \sqrt{1-P_D}|S,-1\rangle + \sqrt{P_D}\left(\sqrt{\tfrac{1}{10}}|D,-1\rangle - \sqrt{\tfrac{3}{10}}|D,0\rangle + \sqrt{\tfrac{3}{5}}|D,+1\rangle\right) \quad (7.3)$$

Here, on the left hand side the state vectors denote $|j,m_j\rangle$, while on the right hand side they contain $|\ell,m_s\rangle$. The measured target polarization depends only on the total angular momentum and its projection, where N^+ is the number of deuterons in the $|1,+1\rangle$ state, N^0 and N^- analog.

$$P_{\text{target,measured}} = \frac{N^+ - N^-}{N^+ + N^0 + N^-} \quad (7.4)$$

The polarization of the spins of the deuterons is given by the decomposition of this expression in terms of the wave functions given above, counting positive all contributions with deuteron spin projection $+1$ and negative those with projection -1.

$$P_{\text{target,spin}} = \frac{N^+\left[(1-P_D)+P_D(\tfrac{1}{10}-\tfrac{3}{5})\right] + N^-\left[-(1-P_D)+P_D(\tfrac{3}{5}-\tfrac{1}{10})\right]}{N^+ + N^0 + N^-} \quad (7.5)$$

$$= \frac{N^+(1-\tfrac{3}{2}P_D) - N^-(1-\tfrac{3}{2}P_D)}{N^+ + N^0 + N^-} \quad (7.6)$$

$$= P_{\text{target,measured}}(1-\tfrac{3}{2}P_D) \quad (7.7)$$

In the following, P_{target} stands for this corrected value of the target polarization.

7.2.3 Measured Asymmetries

The experiment counts the events (or hadrons) with certain characteristics, while the theoretical calculation yields cross sections—to be multiplied with the luminosity so that again numbers of events are obtained. But there is also another difference, namely that the calculations are done for specific spin configurations, corresponding to 100% polarization for beam and target. Thus, a number of factors enter the relation between the measured count rate asymmetry and the asymmetry related to the physical process:

$$A_{\text{exp}} = \underbrace{fDP_{\text{beam}}P_{\text{target}}}_{\beta} A^{\gamma^*}_{\text{phys}} \qquad \text{or} \qquad A_{\text{exp}} = \underbrace{fP_{\text{beam}}P_{\text{target}}}_{\beta} A_{\text{phys}} \quad (7.8)$$

where f is the dilution factor accounting for the presence of unpolarizable material in the target,[4] D is the depolarization factor which describes the polarization transfer from the incoming lepton to the virtual photon, and P_{beam} and P_{target} are the beam and target polarizations, respectively; as is explained later, the target polarization is not used in event weighting, therefore the two variants shown above are in that context often labeled fDP_{beam} or fP_{beam} methods, respectively. In the left case $A^{\gamma^*}_{\text{phys}}$ is the so-called virtual photon asymmetry, as the inclusion of the depolarization factor extrapolates to completely polarized photon flux. While comfortable for the calculation of the hard scattering process, this cannot be measured in deep inelastic scattering, where photons always receive only a part of the already incomplete polarization of the lepton beam. The next-to-leading order theoretical calculations presented in section 6.6.3 therefore include an appropriate description of the partial photon polarization, making the inclusion of D unnecessary.

7.2.4 Mechanisms to Avoid False Asymmetries

The event rates in the two target cells are

$$\text{upstream:} \qquad N_u = \phi_u a_u n_u \sigma_0 (1 + \beta A_{\text{phys}}) \tag{7.9}$$

$$\text{downstream:} \qquad N_d = \phi_d a_d n_d \sigma_0 (1 - \beta A_{\text{phys}}) \tag{7.10}$$

where β is the factor from eq. (7.8), ϕ_u, a_u and n_u the flux, acceptance and number of nucleons in the upstream target cell, and σ_0 is the unpolarized cross section. If we were to directly use the count rate asymmetry

$$\Delta = \frac{N_u - N_d}{N_u + N_d} \tag{7.11}$$

then any difference in the flux, acceptance or number of nucleons between the two target cells would generate an asymmetry which is not produced by the polarized cross sections but only by the experimental setup. Therefore several measures are taken to alleviate this effect.

First, a cut is done in the analysis, such that only events are used for which the extrapolated beam track crossed the whole target length inside the polarized volume. This ensures that $\phi_u \approx \phi_d$; the systematic effect of slightly less unscattered muons in the downstream cell because of the geometrical setup[5] is small enough to be neglected, but in any

[4] This factor describes which fraction of the target nucleons is polarizable. The naïve expectation—picturing ^{6}LiD as a bound state of an α particle and a deuteron—would be 0.5, but in fact this depends slightly on kinematics and thus $\langle f \rangle = 0.38$.

[5] Using the statistics for the next-to-last cut on table 7.2, subtracting one third for the hodoscope and transverse target cuts analog table 7.4, correcting for an acceptance/efficiency of about 60% [67], and applying a luminosity for period W28 of 33 pb^{-1} (see table 7.5), one arrives at a DIS cross section of 1 µb for

case it is cured together with the acceptance effects in the method of asymmetry extraction.

Second, the difference in acceptance and number of target nucleons is averaged out by inverting the target cell polarizations three times per day. The idea is that nothing changes but the association between $\sigma^{\rightleftarrows}$, $\sigma^{\rightleftarrows}$ and the two target cells so that the asymmetry induced by the experimental setup enters with different sign in the two sub-samples and hence cancels in the calculation of the physics asymmetry. It is obvious that the cancellation is only perfect if the statistics gathered with the two spin configurations are exactly equal and the false asymmetry also stays exactly the same throughout the time of the measurement. As the fast spin rotation can only be done by inverting the target solenoid field, this is not true. For one, the interaction between the target solenoid and the first spectrometer magnet is quite strong and leads to an up/down movement of the target by several 100 μm at the downstream end. Another problem is that the spectrometer performance depends slightly on external conditions like temperature and air pressure—think signal delays in cables and signal gain in detectors employing gas amplification—which is still not perfectly canceled by doing three reversals per day.

This leads to the third measure, the development of bad spill list and run groups. This task is done by the Data Stability Group of COMPASS, which publishes the official lists to be used in analysis. The procedure is to look at various characteristics of the data to determine certain benchmark performance estimators. The smallest practical unit for these tests is one SPS spill, containing on the order of 20.000 to 60.000 events.[6] Spills which lie outside the usual bands for one of the benchmarks are flagged "bad" and excluded from analysis. But the position of the bands also changes due to modifications to the spectrometer, so groups of adjacent runs which share the same benchmark values are analyzed together to extract an asymmetry, as within each group it is reasonable to assume that the acceptance for the upstream and downstream target cells did not change. For the beam times of 2002–2004, 185 run groups have been created, yielding the same amount of asymmetries. But this approach is still not perfect, as the inverted target solenoid field's influence on the charged particle trajectories between the target and the first spectrometer magnet does not cancel here.

Thus, a fourth measure is taken to overcome this obstacle, which is the inversion of the microwave setting, changing the connection between target solenoid current direction and target spin orientation. As this requires the destruction of the polarization as described in section 7.2.1 this is only done every few weeks. During the beam time, modifications to the spectrometer are kept to a minimum to make the asymmetries obtained within a few weeks comparable, so that the inverted microwave setting can be used as a check for systematic uncertainties remaining after the first three steps. If the statistics recorded with both settings are equal, these remaining false asymmetries should cancel

the COMPASS kinematics. The cross section for hard μ–e scattering should be of the same order, considering that the leading order matrix element is the same, apart from the charge. Therefore, the probability for a hard interaction inside a single target cell, given a target density of $1.7 \cdot 10^{25}$ cm^{-2}, is smaller than 10^{-4}.

[6] The luminosity delivered per spill by the SPS improved over the years, as well as the rate capability of the DAQ system, leading to the installation of additional triggers and the reduction of vetoes.

completely unless any modification was done to the apparatus which introduced a false asymmetry again.

7.2.5 Weighting

According to the recollection of J. Pretz [68], the method of individual event weighting was first used at the SMC experiment. The basic idea is that the measured asymmetry $A_{\text{exp}} = \frac{u-d}{u+d}$ is proportional to the physics asymmetry A_{phys} by some kinematic factor β as stated in eq. (7.8), so to extract the physics asymmetry one does

$$A_{\text{phys}} = \frac{1}{\langle\beta\rangle} A_{\text{exp}} \qquad \delta A_{\text{phys}} = \frac{1}{\langle\beta\rangle}\delta A_{\text{exp}} \approx \frac{1}{\langle\beta\rangle\sqrt{N}} \tag{7.12}$$

where $N = u+d$ is the number of events in the data sample, and u and d are the contrasting sub-samples, in our case events where the muon and nucleon spins were oriented parallel and anti-parallel, respectively. However, this method does not fully exploit the statistical possibilities as it gives the same weight to events for which β is big as for events where it is small, while it is clear that the latter ones contain less information on the physics asymmetry than the former ones.

Let us consider the extraction of asymmetries in very small bins of β where $\langle\beta\rangle_{\text{bin } i} = \beta_i$. Then each bin contributes an asymmetry $A_i = \frac{1}{\beta_i}\frac{u_i-d_i}{u_i+d_i}$ to the weighted mean

$$A^w_{\text{phys}} = \frac{\sum w_i A_i}{\sum w_i} = \frac{\sum \frac{A_i}{\delta^2 A_i}}{\sum \frac{1}{\delta^2 A_i}} = \frac{\sum A_i \beta_i^2 N_i}{\sum \beta_i^2 N_i} = \frac{\sum u_i\beta_i - \sum d_i\beta_i}{\sum u_i\beta_i^2 + \sum d_i\beta_i^2} \tag{7.13}$$

where we used the weight $w_i = \frac{1}{\delta^2 A_i}$ derived in appendix B, using eq. (7.12) and $A_i\beta_i N_i = u_i - d_i$. Going to infinitesimally small bins means that the sum $\sum u_i\beta_i$ over bins i with event counts u_i can be replaced by the sum of weights for all events j from the u target cell, $\sum_u \beta_j$. Essentially this means that from each event in the u sub-sample an asymmetry of $\frac{1}{\beta_i}$ is calculated with an uncertainty of the same magnitude, while the asymmetries calculated from the d sub-sample have the opposite sign. The uncertainty obtained using eq. (B.11) with $w_j = \beta_j^2$ for this weighted method is

$$\delta^2 A^w_{\text{phys}} = \frac{1}{\sum \beta_j^2} = \frac{1}{\langle\beta^2\rangle N} \tag{7.14}$$

Using the relation $\langle x^2\rangle = \langle x\rangle^2 + \sigma_x^2$ derived in appendix B, it can be seen that this uncertainty is smaller than the one in eq. (7.12) by the factor

$$\frac{\delta A_{\text{phys}}}{\delta A^w_{\text{phys}}} = \sqrt{\frac{\langle\beta^2\rangle}{\langle\beta\rangle^2}} = \sqrt{1 + \frac{\sigma_\beta^2}{\langle\beta\rangle^2}} \tag{7.15}$$

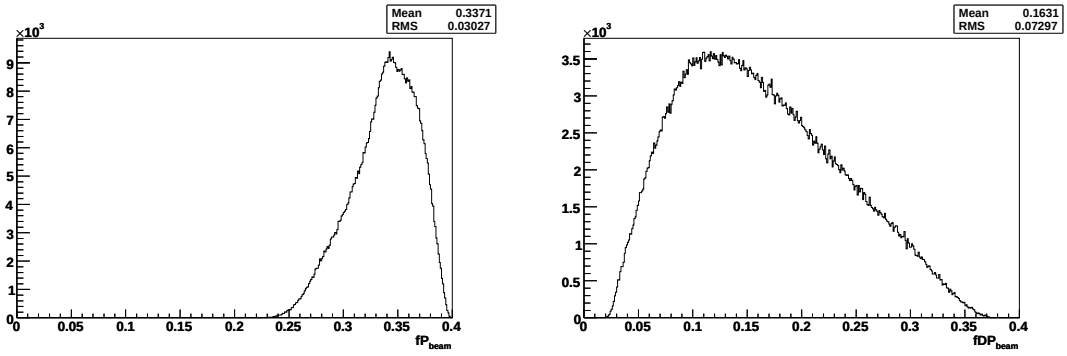


Figure 7.1: Distribution of possible weights for use in the asymmetry extraction algorithms.

It should be noted that weighting has so far only been considered using the optimal weight, which is given by the factor between the measured and the physics asymmetries. In COMPASS, either the muon–nucleon asymmetry or the virtual photon–nucleon asymmetry are measured, differing by the inclusion of the depolarization factor D. Figure 7.1 shows the distribution of the two corresponding choices for the weight together with their mean and standard deviation. Inserting these values into eq. (7.15) yields a possible gain in statistical uncertainty of 10% for the fDP_{beam} case—which is therefore used in the purity analysis introduced in section 6.6.1—but only 0.4% gain for the fP_{beam} case.

P_{target} is not included in the weight because in SMC substantial false asymmetries have been observed with the fDP_{target} weighting, which disappeared when using only fD. The reason is that if the weight contains quantities which show a long-term trend in time, the field reversal will not lead to a perfect cancellation of the apparatus asymmetry. Therefore the target polarization—measured every few minutes—is not evaluated event-by-event, but is averaged over the time interval covered by one asymmetry measurement and applied as a constant factor to the obtained asymmetry in the end.

Another consideration when using weighting is that the theoretical calculations so far have been done without employing this technique. In principle the weighting procedure should not introduce a bias in the extracted physics asymmetry,[7] but for a minuscule gain in statistical uncertainty the risk should not be taken that an assumption made in the derivation of the method is not well enough fulfilled. *Thus, the asymmetries presented in section 7.5 are extracted without utilizing event weighting.*

7.2.6 1st Order Method

This section only summarizes the keys points, for all the details see [69]. The application of weighting to the COMPASS case is based upon the assumption that

[7] Of course, the weight must not be correlated with the quantity which is being averaged. In our case, the A_i of eq. (7.13) are not correlated with the kinematic factors in question, because the effect of those has been divided out.

$$u = \phi a_u n_u \sigma_0 (1 + w P_u A_{\text{phys}}) \tag{7.16}$$

$$d = \phi a_d n_d \sigma_0 (1 - w P_d A_{\text{phys}}) \tag{7.17}$$

where ϕ is the incident muon flux, a_u is the acceptance function for the upstream target cell, n_u the number of target nucleons, σ_0 the unpolarized cross-section, $w = f P_{\text{beam}}$ with f being the dilution factor and P_{beam} the beam polarization, and P_u is the polarization of the material in the upstream target cell. ϕ and σ_0 already cancel in the counting rate asymmetry

$$\Delta = \frac{u - d}{u + d} = \frac{r - 1 + w A_{\text{phys}}(r P_u - P_d)}{r + 1 + w A_{\text{phys}}(r P_u + P_d)} \tag{7.18}$$

because we make sure to use only events for which the extrapolated beam track would have crossed both target cells. $r = \frac{a_u n_u}{a_d n_d}$ describes the asymmetry introduced by the apparatus, which neither has the same acceptance for events happening in the upstream or downstream cells, nor are the two cells identically filled. To cancel this asymmetry, the target polarization is inverted by adiabatically turning around the magnetic field as described in section 2.2.5. This gives a set of equations analog to eq. (7.16)–(7.18), but with primed quantities. Under the assumption that the spectrometer is not changed between the two settings, this only exchanges the association of u and d to the two target cells (this means that $a'_u = a_d$, analog for a'_d). Assuming that the difference in target polarization obtained in the two cells is small enough, it follows that the apparatus asymmetry cancels in the unweighted average of the asymmetries obtained using the two target spin configurations:

$$\frac{1}{2}(\Delta + \Delta') = (1 - \alpha^2) P_{\text{target}} \langle w \rangle A_{\text{phys}} \tag{7.19}$$

$$\sqrt{\frac{1}{N} + \frac{1}{N'}} = 2 P_{\text{target}} \langle w \rangle \sqrt{1 - \alpha^2}\, \delta A_{\text{phys}} \tag{7.20}$$

$$P_{\text{target}} = \frac{1}{4}(P_u + P_d + P'_u + P'_d) \tag{7.21}$$

$$\alpha = \frac{a_u n_u - a_d n_d}{a_u n_u + a_d n_d} = \frac{r - 1}{r + 1} \tag{7.22}$$

The uncertainty of the measurement of the target polarization—amounting to 3%—is not included in the expression for δA_{phys}, as its effect is accounted for in the systematic error.

α is the apparatus asymmetry, which of course is not known a priori. It could be obtained by Monte-Carlo simulation, but an estimate of it is also available by using the unweighted difference of the two extracted asymmetries

$$\alpha \approx \frac{1}{2}(\Delta - \Delta') \tag{7.23}$$

In case α is small, it is a small correction, which needs to be known only to 10% in order not to spoil the statistical accuracy. When applying weighting to this formalism, the assumption that $\langle w \rangle$ does not change between the two target spin configurations is needed, wherefore the target polarization—changing systematically between polarization attempts—is not included in the weight. With the assumption that the acceptance function can be taken as a constant across all bins in w the final extraction formula for the 1st order method is

$$A^w_{\text{phys}} = \frac{1}{2P_{\text{target}}(1-\alpha^2)} \left[\frac{\sum_u w_i - \sum_d w_i}{\sum_u w_i^2 + \sum_d w_i^2} - \frac{\sum'_u w_i - \sum'_d w_i}{\sum'_u w_i^2 + \sum'_d w_i^2} \right] \tag{7.24}$$

$$\delta A^w_{\text{phys}} = \frac{1}{2P_{\text{target}}\sqrt{1-\alpha^2}} \sqrt{\frac{1}{\sum_u w_i^2 + \sum_d w_i^2} + \frac{1}{\sum'_u w_i^2 + \sum'_d w_i^2}} \tag{7.25}$$

These equations will be referred to throughout this thesis as *1st order weighted method with global acceptance factor* (1st order global). In view of sizable apparatus asymmetries of the order of $\alpha = 0.1$ as are observed at COMPASS,[8] it is preferable to redefine the counting rate asymmetry to

$$\Delta = \frac{(1-\alpha)u - (1+\alpha)d}{(1-\alpha)u + (1+\alpha)d} = \frac{u - rd}{u + rd} \tag{7.26}$$

This changes the picture so that even if α is not small, it needs to be known only to 10%. The resulting formula for asymmetry extraction will be called *1st order weighted method with individual acceptance factors* (1st order individual):

$$A^w_{\text{phys}} = \frac{1}{2P_{\text{target}}} \left[\frac{\sum_u w_i - r\sum_d w_i}{\sum_u w_i^2 + r\sum_d w_i^2} - \frac{\sum'_u w_i - r\sum'_d w_i}{\sum'_u w_i^2 + r\sum'_d w_i^2} \right] \tag{7.27}$$

$$\delta A^w_{\text{phys}} = \frac{1}{2P_{\text{target}}\sqrt{1-\alpha^2}} \sqrt{\frac{1}{\sum_u w_i + r\sum_d w_i} + \frac{1}{\sum'_u w_i^2 + r\sum'_d w_i^2}} \tag{7.28}$$

Both sets of extraction formulas rely upon an approximately homogeneous acceptance function in the covered phase space. The biggest expected dependency of the acceptance function is on the p_T of the outgoing hadrons, as they are produced under the biggest angles and potentially suffer from the limited opening angle of the target magnet. However, as the data are binned in p_T, the acceptance should be reasonably constant inside each bin.

7.2.7 2nd Order Method

The assumptions about the apparatus asymmetry made in the previous section are not needed when employing the so-called 2nd order method described in [68], which owes

[8] Keep in mind that the statistical accuracy is of the order of $\delta A \approx 0.001$ for the lowest bins in p_T, which should not be spoiled by the uncertainty on the apparatus asymmetry.

its name to the fact that it expands eq. (7.18) for the two target spin configurations into a 2nd order equation for the physics asymmetry. The exact condition is that

$$\varkappa = \frac{\langle a_u \rangle \langle a'_d \rangle}{\langle a'_u \rangle \langle a_d \rangle} \approx 1 \tag{7.29}$$

which means that the apparatus—including the position of the illuminated target volume—does not change on average between the two microwave settings. This leads to a second order equation

$$a\left(A^w_{\text{phys}}\right)^2 + bA^w_{\text{phys}} + c = 0 \quad \Longrightarrow \quad A^w_{\text{phys}} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{7.30}$$

with

$$a = \gamma \langle \beta'_u \rangle_w \langle \beta_d \rangle_w - \langle \beta_u \rangle_w \langle \beta'_d \rangle_w \tag{7.31}$$

$$b = \gamma(\langle \beta'_u \rangle_w + \langle \beta_d \rangle_w) - (\langle \beta_u \rangle_w + \langle \beta'_d \rangle_w) \tag{7.32}$$

$$c = \gamma - 1 \tag{7.33}$$

where

$$\gamma = \frac{\sum_u w_i \sum'_d w_i}{\sum'_u w_i \sum_d w_i} \qquad \langle \beta_u \rangle_w = \frac{\sum_u w_i \beta_i}{\sum_u w_i} \tag{7.34}$$

and β is the factor from eq. (7.8). The uncertainty is given by Gaussian propagation from γ and $\langle \beta \rangle_w$ by

$$\delta^2 A^w_{\text{phys}} = \left(\frac{\partial A^w_{\text{phys}}}{\partial \gamma}\right)^2 \delta^2\gamma + \sum_{i=u,u',d,d'} \left(\frac{\partial A^w_{\text{phys}}}{\partial \langle \beta_i \rangle_w}\right)^2 \delta^2 \langle \beta_i \rangle_w \tag{7.35}$$

The contribution from the uncertainties on $\langle \beta \rangle_w$ is very small, amounting only to an increase of the total statistical uncertainty of 0.5% in the highest bin in p_T and being negligible in the bins with reasonable statistics.

7.3 The Data Sample

The basis of the sample is formed by all muon data taken with longitudinal target polarization during the beam times of 2002–2004. For the exact mDST production versions used see table E.2 on page 165. All kinematic plots and cut summary tables are obtained using the full statistics of 2004, unless noted otherwise.

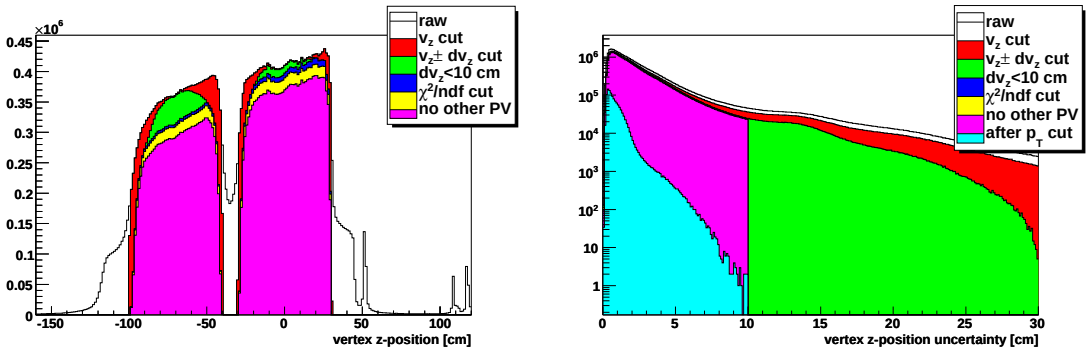


Figure 7.2: Cut series applied to the primary vertex. The colored areas show the entries left after applying the corresponding cut. The effect of the high p_T cut is not shown on the left plot due to the linear scale.

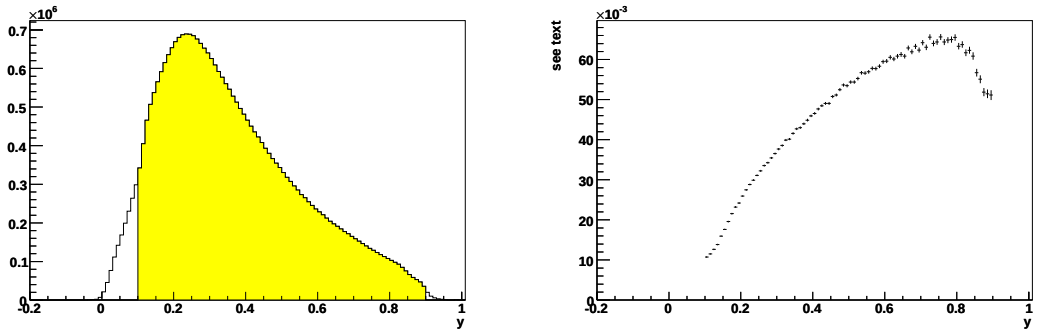


Figure 7.3: The left plot shows the distribution of the relative energy transfer y; the area selected by the cut $0.1 < y < 0.9$ is colored. The right plot depicts the fraction of the y distribution left after requiring at least one high p_T track in the event to the original distribution.

7.3.1 Raw Material

As explained in section 2.3 data taking is organized in periods, where within one period the general data characteristics should not change. Thus, there are periods of longitudinal mode and periods of transverse mode, denoting the arrangement of the target polarization relative to the beam direction. Table E.4 on page 167 gives an overview of the distribution of the statistics between these periods. Please note that the statistics for the four longitudinal spin configurations are equalized at the level of 10%.

7.3.2 mDST Cuts

The first step in the analysis chain is to select the interesting events from the mDST. In this case this means making sure that a muon has interacted inside the polarized target volume and created one or more additional particles. Furthermore there must not be any other such signature inside the same event, as then there would be a possibility for the additional particles to be associated with the wrong process. Apart from these vertex cuts and the $p_T > 1$ GeV cut only few cuts are done at the mDST level to retain maximal flexibility at the μDST level.

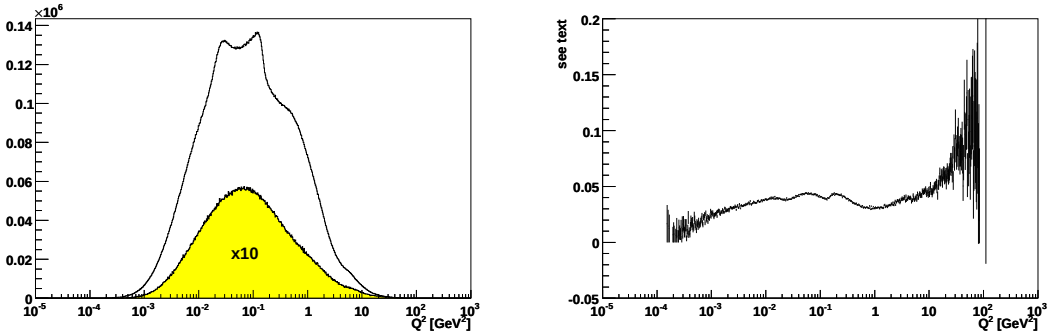


Figure 7.4: The left plot shows the distribution of the four-momentum transfer Q^2; the area selected by requiring at least one high p_T hadron attached to the primary vertex is colored. The right plot depicts the ratio of the two distributions. The structure seen in the uncut distribution is a characteristics of the (semi-) inclusive triggers, which vanishes when tagging the events with the strong hadron cuts as will be discussed in section 7.3.3.

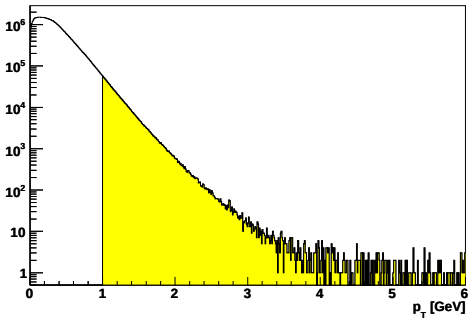


Figure 7.5: Distribution of p_T for all tracks associated to the vertex selected by the event cuts. The colored area is the sub-sample which is written to the μDST.

First, only events are considered, which have a *best primary vertex* as indicated by Phast. This vertex has to contain a beam particle and at least two outgoing particles, one of which has been flagged as scattered muon. After these cuts—given by the compass reconstruction as detailed in section 2.3—a series of cuts on the reconstructed vertex parameters is applied, which is illustrated in figure 7.2. These cuts do not take into account the transverse coordinates of the vertex as the exact angle between the target cells and the beam axis and the size and shape of the filled target volume are not known precisely a priori. Therefore, the transverse target cuts are a matter of tuning as will be discussed in section 7.6.2.1. The cuts applied at the mDST level are:

vertex z position The two target cells are arranged along the experimental z axis as the two intervals $[-100\,\text{cm}; -40\,\text{cm}]$ and $[-30\,\text{cm}; +30\,\text{cm}]$. Due to the small angles of the outgoing tracks caused by the fixed target nature of the compass experiment the vertex position resolution along the beam direction is limited. Therefore a possible misalignment of a few mm cannot be measured and the design values have to be trusted.

vertex z uncertainty range As already mentioned the uncertainty of the vertex z position measurement cannot be neglected, so the next cut demands that the whole interval $[v_z - \sigma_{v_z}; v_z + \sigma_{v_z}]$ lies within the same target cell. Figure 7.2 shows how this cut takes away vertices close to the ends of the target cells. While the cut on the z position alone does not have a big effect on the distribution of the z position uncertainty σ_{v_z} (right plot in figure 7.2), this cut effectively restricts $\sigma_{v_z} < 30\,\text{cm}$.

vertex z uncertainty The separation of the target cells is 10 cm, so a cut $\sigma_{v_z} < 10\,\text{cm}$ together with the previous cut guarantees that the vertex is associated to the correct target cell better than on the 2σ level. Another reason for this cut is to get rid of pseudo-vertices created by multiple scattering of pile-up muons in the target which happen to have picked up other tracks. As the outgoing tracks have to be extrapolated back into the target for vertexing the effect of this cut is much larger in the upstream cell.

χ^2 cut The vertex parameters are the result of a fit to a number of track parameter sets. We make a $\sim 3\sigma$ cut on the χ^2 of this fit to increase the probability that the reconstructed vertex is built up of the right tracks. The effect of this cut is small and evenly distributed as it should be.

no other PV If another primary vertex has been found in the event there is a chance that the hadrons which are later reconstructed are associated with wrong kinematic variables because they have been attached to the wrong vertex. Therefore only events with exactly one primary vertex are considered.

cut	number	ratio	reduction
all events	140626670	100.00%	0.00%
primary vertex (PV)	130506271	92.80%	7.20%
PV has beam	130506271	92.80%	0.00%
PV has scattered muon	81983213	58.30%	37.18%
PV has at least 2 particles	43543265	30.96%	46.89%
PV z position inside target	37122170	26.40%	14.75%
PV $z \pm \sigma_z$ inside target	33638007	23.92%	9.39%
$\sigma_z < 10\,\text{cm}$	32282878	22.96%	4.03%
χ^2 cut	31727476	22.56%	1.72%
not two PVs	30021701	21.35%	5.38%
y cut	28404851	20.20%	5.39%
beam momentum	28347404	20.16%	0.20%
at least one hadron with $p_T > 1$ GeV	1145249	0.81%	95.96%

Table 7.2: Impact of the event cuts demonstrated on the data of period W28 of 2004.

$0.1 < y < 0.9$ While strongly suppressed by the high muon mass, *radiative effects* can become sizable near the boundaries of the allowed kinematic region. Therefore the cuts shown in figure 7.3 are employed. Keep in mind that the y distribution changes in shape after the application of the high p_T cut described below.

beam momentum As discussed in section 2.2.2, the beam polarization is determined using a parameterization which was obtained by a simulation of the beam line. This parameterization is valid in the range $[140\,\text{GeV}; 180\,\text{GeV}]$, so the beam momentum is restricted to that range.

high p_T cut In the following only interested high p_T particles are of interest, which means that only those events are needed which have at least one track associated to the primary vertex with $p_T > 1$ GeV. p_T is the transverse momentum of the outgoing particle with respect to the virtual photon direction in the laboratory frame, which is the same as in the center of mass system of photon and nucleon. For the purpose of this cut a list is made of all tracks attached to the vertex, which is then ordered according to the p_T of the tracks. The event is selected if the first track has $p_T > 1$ GeV and all list members fulfilling this criterion are then individually written—in the same order—to the μDST. It is important to note that the μDST does not contain events but high p_T tracks as records. The distribution of transverse momenta and the effect on the Q^2 distribution of requiring at least one high p_T hadron in the event is shown in figures 7.4 and 7.5.

With this the list of cuts applied is complete at the level where the whole event information is available. This does not mean that no other event characteristics are subject to cuts, but those cuts are done at a later stage where the event structure has been broken up.

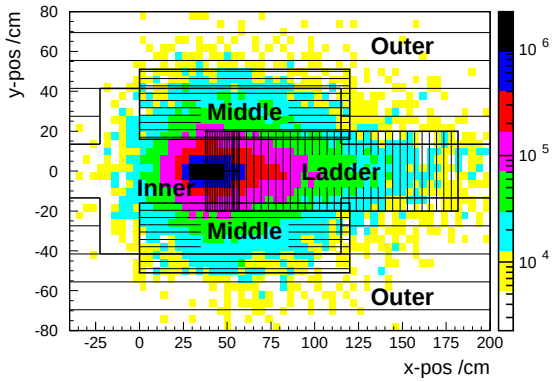


Figure 7.6: Position of the trigger hodoscopes in the plane perpendicular to the beam direction. The color scale indicates the muon rate per 16 cm^2 as expected from a Monte-Carlo simulation in a plane directly behind the second muon filter ($z = 40$ m). Taken from [38].

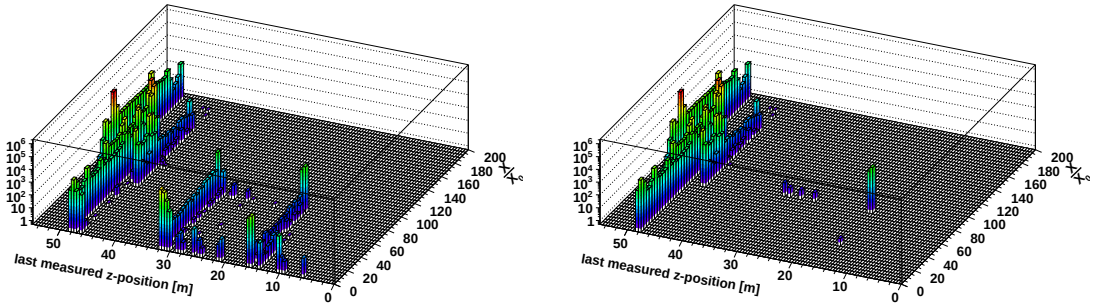


Figure 7.7: Last measured position of the scattered muon track versus number of radiation lengths crossed by that track. The left plot shows the distribution for all μDST entries while the right one depicts the situation after application of the hodoscope cut.

Table 7.2 lists the size of the μDSTs resulting from these cuts together with the number of entries. This is the source material for the hadron cuts described in the following section.

7.3.3 μDST Cuts

The mDST cuts deal with properties that are difficult to retain at the μDST level, or which would significantly increase their data volume; quantities like the event or hadron kinematics are represented compactly and of manifold use for systematic studies etc., so cutting on those is delayed till the μDST level. The reason for doing the cuts as late as possible is the gain in flexibility for the following analysis: the μDST for the complete period W28 of 2004 is 117 MiB in size and can easily be processed on a notebook computer.

The first cuts define the deep inelastic scattering event. The counting of single-inclusive hadrons is very sensitive to wrongly reconstructed event kinematics and there is no physics condition which could restrict the final state. Therefore all information given by the spectrometer is exploited to identify the scattered muon as such. First it is checked that the reconstructed muon track corresponds to the scattered muon that actually has

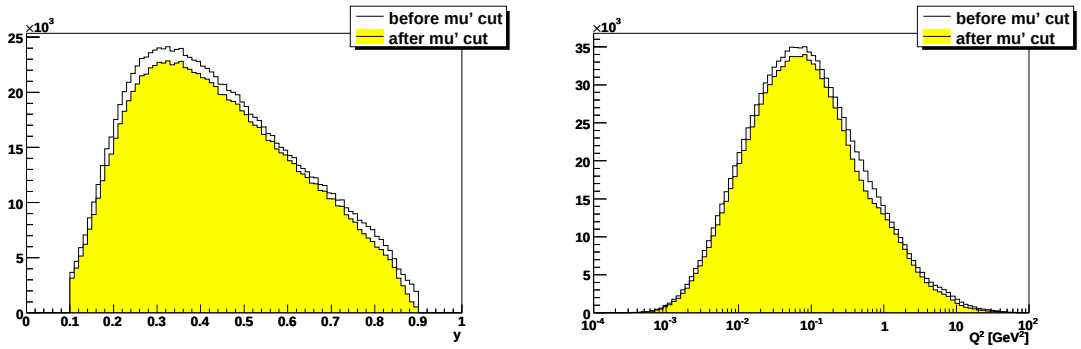


Figure 7.8: Effect of the combined hodoscope and radiation length cuts on the y and Q^2 distributions.

	upstream		downstream	
year	x[cm]	y[cm]	x[cm]	y[cm]
2002	−0.2	0.1	−0.3	−0.15
2003	0	0	−0.1	−0.25
2004	0	−0.1	0	−0.3

Table 7.3: Parameters of the position of the target volume (taken from [43] and private communication). The numbers given are offsets of the center of the upstream end ($z = -100$ cm) and downstream end ($z = 30$ cm) with respect to the beam axis. While the real target holder probably is tilted with respect to the beam axis, for the sake of simplicity it is treated as being sheared transverse to the beam axis in the analysis; the imprecision incurred is surely smaller than the uncertainty on the parameters.

created the event trigger. The *hodoscope cut*, illustrated in figure 7.7, requires the alleged muon track to have hits in both hodoscope stations belonging to the trigger subsystem which detected the event. In case more than one trigger fired—being quite common as the hodoscopes partly overlap—it is sufficient that the condition is fulfilled for at least one of the sub-triggers, whereas for events which were detected solely by the purely calorimetric trigger no cut is applied. The latter condition accounts for the presence of muon tracks which end around 16 m and 30 m downstream of the target. Figure 7.6 shows the spatial coverage of the different hodoscope trigger systems. Not shown is the beam hole of the muon filter of $40 \times 40\,\text{cm}^2$, which means that for hodoscopes partly covering this hole, the passing particles are not positively identified as being muons. The effect is visible in the right plot of figure 7.7, where around $z = 48$ m tracks are seen which have crossed down to zero radiation lengths. As there is no large excess of events in this region, these particles are probably muons, albeit unidentified; using the COMPASS standard cut for muon identification of 30 radiation lengths would remove 0.63% of the final sample. The effect of the scattered muon cuts is demonstrated in figure 7.8.

The next cut defines the *polarized target volume*. For this, the vertex must lie inside a cylindrical volume which is sheared from the experimental z-axis by ~ 1 mrad to be centered around the axis of the target holder. This volume is capped from the top to describe the filling level of target material inside the holder. The parameters of the volume are $r_{\max} = 14$ mm and $y_{\max} = 10$ mm while the details of the shearing depend on the beam time—see table 7.3 for details.

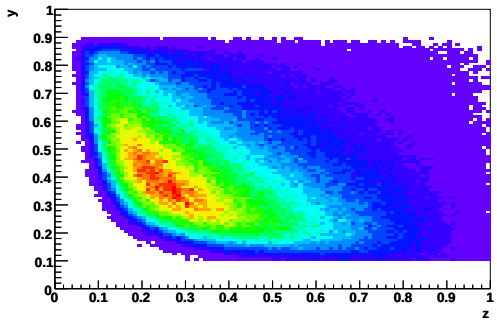


Figure 7.9: Correlation of the kinematic variables y versus z after muon identification cuts, target cut and photo-production cut. The hadron energy in the laboratory frame is given by yzE_{beam}.

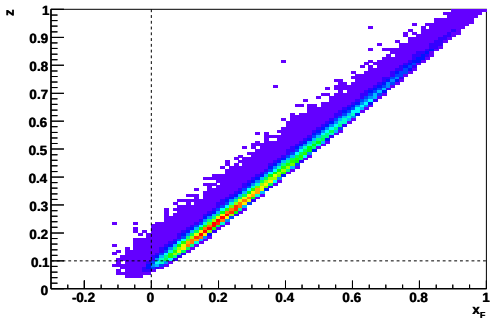


Figure 7.10: Correlation between z and x_F. The dashed lines illustrate the cuts, to which the sample is subjected.

The last event characteristics cuts select the *photo-production* regime $Q^2 < 0.5\,\text{GeV}^2$, restricts $y > 0.2$ and selects hadrons with production angle $\theta < 70$ mrad with respect to the beam axis, to be compatible with the theoretical ansatz presented in section 6.6.3.

The previous cuts ensure a proper selection of the deep inelastic scattering reaction, leaving hadron candidates in the data sample. The next task is to ensure that the hadrons are properly reconstructed. The first cut is to reject tracks which are identified as muons by having traveled through more than 30 radiation lengths of material. The next cut concerns the last measured point of the hadron candidate track, which must be situated downstream of SM1 to have properly defined momentum and direction of flight. This cut rejects the so-called fringe-field tracks, because their momentum is determined by the small bending of the trajectory in the outer regions of the magnetic field.

The next two cuts restrict the kinematic variables z and x_F to their physical regions and select the current fragmentation $x_F > 0$. The additional cut $z > 0.1$ restricts the sample to the region where the fragmentation functions can safely be applied, but it should be noted that the measured z would only match the fragmentation variable in the LO deep inelastic scattering case, where the struck quark's momentum can be assumed to be given by the momentum of the virtual photon. At higher order, the fragmenting quark has less energy than the virtual photon, as other partons are created, so the true z is higher than the calculated one.

The final cut uses the hadron calorimeter to positively identify two hadrons whose extrapolated tracks hit the active volume of HCAL1 or HCAL2. Unfortunately, the MDST of data taken before 2006 do not contain track parameters at the last measured point. Therefore, the uncertainty introduced by the extrapolation from shortly downstream of the target to the very end of the spectrometer is of the order of one centimeter. The size of the calorimeter cells has been enlarged in the analysis accordingly. Tracks which do not point into a hadron calorimeter are accepted unconditionally, while the others are required to have deposited at least 30% of the energy corresponding to their track momentum.

The summary of all cuts applied at the μDST level is given in table 7.4. The final distributions of the kinematic variables Q^2, y, z and p_T are shown in figure 7.11.

7.3.4 Elastic $\mu - e$ Scattering

The target obviously contains as many electrons as it contains protons, so a fraction of the recorded data are events where the μ^+ scattered elastically off an e^-. These events are easily identified, as the outgoing particle multiplicity is two and—to first order—the electron should have a momentum which is essentially given by the virtual photon momentum. Thus, the photon has to obey the dispersion relation

$$x = \frac{Q^2}{2m_e \nu} = 1 \tag{7.36}$$

In DIS, we calculate the very similar x_B, which only differs in the choice of mass m_e of the electron versus M of the nucleon. Thus, the μ^+e^- events should be found at $x_B = \frac{m_e}{M}$ or $\log_{10} x_B = -3.264$. Figure 7.12 gives an overview of the low p_T region for events which have only one additional track attached to the primary vertex. There clearly is a distinct peak for $p_T \to 0$ around $x_B = 0.0005$, which disappears when going to slightly higher transverse momenta as shown in figure 7.13. To further investigate this peak the histogram entries have to be re-weighted. Each bin in p_T corresponds to a circular ring area which is to a good approximation proportional to p_T; thus, assigning a weight of $\frac{1}{p_T}$ transforms the p_T distribution into the track density in dependence of p_T. It has been found that semi-inclusive deep inelastic scattering data can only be described by giving the partons in the nucleon a certain intrinsic transverse momentum $k_\perp$ which is of the order of $400 \div 1000$ MeV, so even for the LO DIS process we expect a smeared out track density at small p_T, and more so for higher order processes. This is demonstrated in figure 7.14. The right plot gives the track densities for the interesting region of x_B and different multiplicities and it shows that the distribution becomes increasingly flat for higher multiplicities. The left plot shows that the μ^+e^- peak also diminishes when going away from the elastic x_B region. The fact that it does not vanish completely is due to radiative effects which allow the hit electron to be off the mass shell before or after the interaction, thereby shifting the reconstructed inclusive kinematic variables.

7.4 The Unpolarized Cross Section

The value obtained in theoretical calculations often is a matrix element for the process under consideration. Together with kinematic factors and integrated over the relevant region of phase space it yields the cross section—usually denoted σ—which is the equivalent area the beam particle has to hit in order to initiate the reaction. The fact that many beam particles are given the chance for interaction and that also many target particles are available is expressed in the luminosity $\mathcal{L}$, which is the number of target particles per unit

cut	upstream cell number	ratio	red.	downstream cell number	ratio	red.
all tracks ($p_T > 2$)	132582	100.00%		218137	100.00%	
hodoscope cut	124156	93.64%	6.36%	204727	93.85%	6.15%
radial target cut	94213	71.06%	24.12%	156981	71.96%	23.32%
$Q^2 < 0.5\,\mathrm{GeV}^2$	70288	53.01%	25.39%	121132	55.53%	22.84%
$y > 0.2$	69348	52.31%	1.34%	118626	54.38%	2.07%
$\theta < 0.07$	46416	35.01%	33.07%	50170	23.00%	57.71%
PID is not muon	45649	34.43%	1.65%	49308	22.60%	1.72%
track after SM1	44579	33.62%	2.34%	46638	21.38%	5.41%
$0 < x_F < 1$	42901	32.36%	3.76%	45437	20.83%	2.58%
$0.1 < z < 1$	42761	32.25%	0.33%	45336	20.78%	0.22%
HCAL identification	37789	28.50%	11.63%	39419	18.07%	13.05%

Table 7.4: Impact of the hadron cuts on tracks from the upstream and downstream target cells demonstrated on the data of 2004. The "radial target cut" includes the vertical cut $y < 1\,\mathrm{cm}$ and is also applied to the extrapolated beam track throughout the longitudinal extent of the target volume.

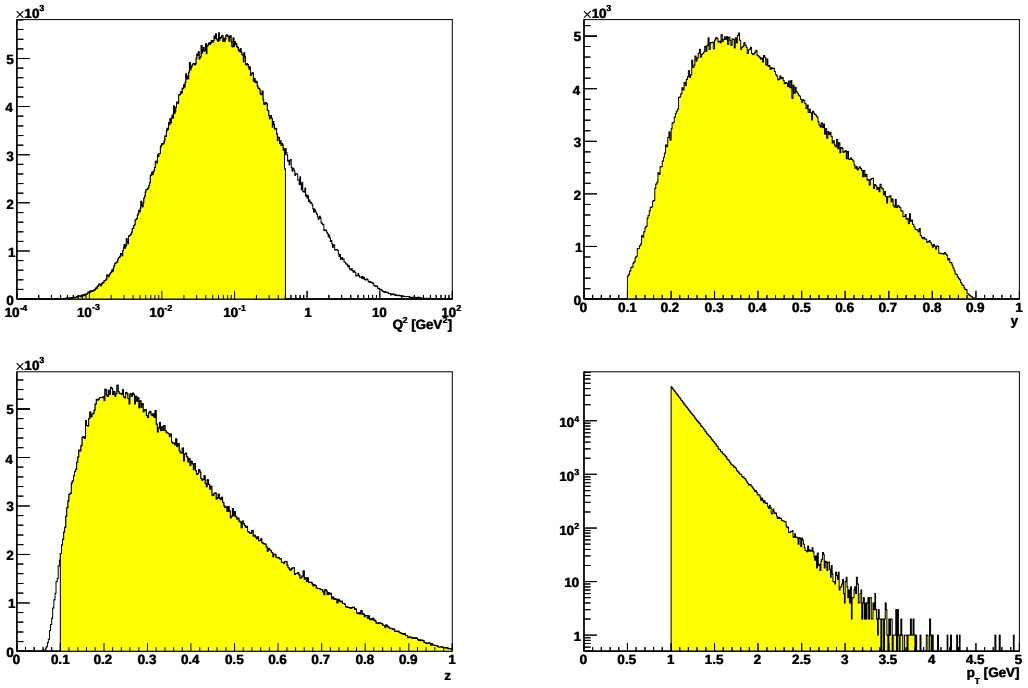


Figure 7.11: Final distributions of the four kinematic variables. In case of the Q^2 and z distributions the non-shaded area shows the part of the distribution which would be added by removing the cut on that specific variable only.

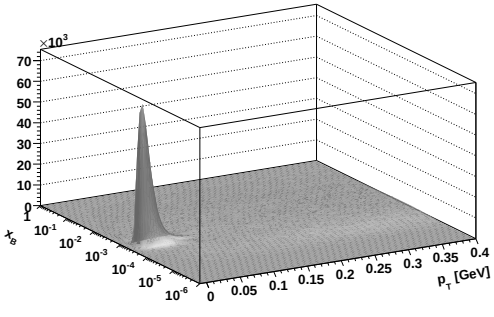


Figure 7.12: Events with only one additional track attached to the target show a distinct peak from elastic μ^+e^- scattering.

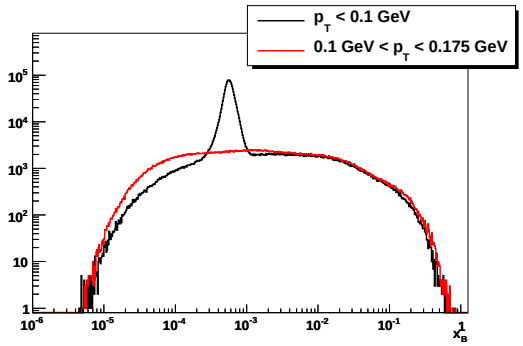


Figure 7.13: x_B distribution for events with multiplicity 1 at very low p_T.

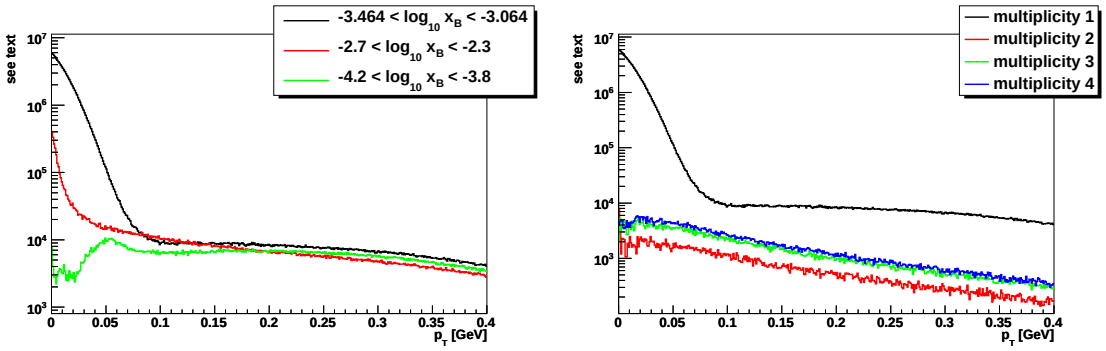


Figure 7.14: Low p_T region for multiplicity 1 in different x_B regions (left) and for different multiplicities around $x_B = \frac{m_e}{M}$ (right). Please note that the entries have been weighted with $\frac{1}{p_T}$ to obtain the density in the p_T plane. The μ–e peak is clearly seen at multiplicity 1 and x_B around $\frac{m_e}{M}$, and vanishes for higher multiplicities as well as when selecting different regions of x_B.

area multiplied by the number of incident beam particles. If the detection probability of the reaction in the detector is ϵ, then the expected number of registered events is

$$N_i = \mathcal{L}\sigma_i\epsilon_i \tag{7.37}$$

within each kinematic bin i. Thus, in addition to the event count, the luminosity and the detection probability need to be known for an extraction of the cross section from the experiment data. Following the presentation of each of these components, radiative effects will be discussed.

7.4.1 Luminosity

The COMPASS target during the 2004 run consisted of 337.2 mol of ^{6}LiD and 72.8 mol of other material, mostly ^{4}He [70]. The area density of nucleons has been evaluated to

$$\rho = 3.5 \cdot 10^{25} \frac{\text{nucleons}}{\text{cm}^2} \tag{7.38}$$

The number of muons delivered to the experiment is counted for each spill using a dedicated scintillator hodoscope in front of the target. Three effects have to be taken into account when using these data:

- The scintillator itself and its readout electronics are not completely dead time free. If two particles cross it less than $\sim$ 4 ns apart the second one will not be counted (see [36] for more information on the counting modules). The spread of beam rates measured during the period 04W28 has been $(95 \pm 4) \cdot 10^6$/ spill, yielding an average rate of 1 in 47 ns. The scintillator is read out in six segments out of which the central two each get about 30% of the flux or 1 in 156 ns. The Poisson distribution yields a probability of 0.32% for the occurrence of two or more events in 4 ns for these central hodoscope parts.

- The data acquisition of the COMPASS experiment is also not dead time free. If two triggers happen less than 5 µs apart or more than 3 (10) triggers occur within 75 µs (250 µs), triggers are lost. During 2002–2004 the DAQ dead time was $\sim$ 5%. Note that the trigger dead time introduced by the veto systems is taken into account by the Monte Carlo simulation in section 7.4.2 and thus not included here.

- The hodoscope counts muons traveling roughly in the direction of the target, it does not ensure that only those muons are counted which traverse the whole length of both target cells. Given the distribution of beam particle trajectories in space, this effect can be described using a geometric factor, which has been determined to be 0.58 (see [70] and references therein).

period	runs	luminosity	luminosity (incl. bad spills)
W22	179	$93.3\,\mathrm{pb}^{-1}$	$102.1\,\mathrm{pb}^{-1}$
W23	116	$55.3\,\mathrm{pb}^{-1}$	$67.2\,\mathrm{pb}^{-1}$
W26	105	$60.1\,\mathrm{pb}^{-1}$	$67.2\,\mathrm{pb}^{-1}$
W27	62	$34.2\,\mathrm{pb}^{-1}$	$38.5\,\mathrm{pb}^{-1}$
W28	160	$33.0\,\mathrm{pb}^{-1}$	$43.4\,\mathrm{pb}^{-1}$
W29	120	$35.3\,\mathrm{pb}^{-1}$	$39.7\,\mathrm{pb}^{-1}$
W30	123	$54.1\,\mathrm{pb}^{-1}$	$63.0\,\mathrm{pb}^{-1}$
W31	101	$55.1\,\mathrm{pb}^{-1}$	$60.3\,\mathrm{pb}^{-1}$
W32	169	$78.6\,\mathrm{pb}^{-1}$	$86.9\,\mathrm{pb}^{-1}$
W37	161	$88.5\,\mathrm{pb}^{-1}$	$97.7\,\mathrm{pb}^{-1}$
W38	168	$101.0\,\mathrm{pb}^{-1}$	$103.3\,\mathrm{pb}^{-1}$
W39	82	$53.2\,\mathrm{pb}^{-1}$	$57.2\,\mathrm{pb}^{-1}$
W40	68	$33.4\,\mathrm{pb}^{-1}$	$42.1\,\mathrm{pb}^{-1}$
total	1614	$775.1\,\mathrm{pb}^{-1}$	$868.6\,\mathrm{pb}^{-1}$

Table 7.5: Integrated luminosity for the data taking periods of 2004 which have been used for the extraction of the cross sections.

- The scintillator readout has a certain noise level, which is accounted for by applying thresholds to the signals using discriminator modules. In case of increased noise during certain periods of data taking, the obtained luminosity would be artificially increased, leading to a smaller reconstructed absolute cross section. While no systematic studies have been performed concerning this problem, it is generally believed that an effect of the order of 10% cannot be excluded.

Thus, a factor of $0.997 \cdot 0.95 \cdot 0.58 = 0.55$ needs to be applied to the raw counts produced by the hodoscopes. The extracted and corrected values are available in [70] for each run.

It is customary to express the luminosity as an inverse cross section, e. g. $1\,\mathrm{pb}^{-1}$ is $10^{36}\,\mathrm{cm}^{-2}$, giving an expectation value of 1 count for a process with a cross section of 1 pb. With this convention, the luminosity is summarized for the relevant data taking periods of the 2004 run in table 7.5, resulting in a total of $775.1\,\mathrm{pb}^{-1}$.

7.4.2 Acceptance Correction

The so-called acceptance correction describes the process needed in order to make experiment results comparable to theoretical calculations. It accounts for a number of effects which distort the measurement, the most important being the limited range of particle energies and trajectories *accepted* by the detector—hence the name. Even if a particle moves within the detector's acceptance it might not be detected due to *inefficiencies* in

individual detectors or the reconstruction process. A further effect is called *bin migration*, which means that due to inaccuracies and limited detector resolution some of the particle's reconstructed characteristics may fall into different kinematic bins than the true ones, leading to an inefficiency in the original bin and a pollution in the neighboring one.

It would be most convenient to extract the magnitude of these distortions by comparing a precisely known cross section to the measured count rates in all kinematic bins, since this would also provide an absolute normalization for the experiment luminosity. But the corrections are not independent of the process under consideration. The vertexing efficiency depends on the number of outgoing particles as well as their angles, the detector efficiencies depend on hit multiplicities, and so on. Especially the bin migration depends on the steepness of the slope of the cross section between bins. Thus, it is required to extract the correction factors using events of similar characteristics to the ones which are to be corrected. As this usually excludes the well-known benchmark cross sections, a different solution needs to be found.

The canonical approach for this problem is to use a Monte Carlo simulation to illuminate the detector with particle tracks similar in composition to the real experiment data, model detector response and reconstruction and obtain the losses and inaccuracies from the comparison between the simulated and reconstructed events. In the end the needed quantity is the cumulative effect of acceptance, efficiency and bin migration in each bin of the differential cross section, expressed in the ratio of simulated over reconstructed event counts. This factor then needs to be applied to the real reconstructed event counts, it is the inverse of ϵ_i in eq. (7.37).

The first part of the Monte Carlo software chain is the event generator. The program of choice for the COMPASS muon program is either LEPTO or PYTHIA, depending on the momentum transfer Q^2 in the deep inelastic scattering process. The domain of large Q^2 is covered by LEPTO while the region $Q^2 < 1\,\mathrm{GeV}^2/c^2$ is simulated with PYTHIA. This program has a very rich set of configuration parameters to allow the simulation of a wide range of kinematics and processes. We use it in a special mode where the beam particle—the virtual photon—is taken from an incoming muon and scatters off partons taken from the target nucleon. This process is well described by the code, owing to long years of fine tuning of the cross sections to experimental results from the deep inelastic scattering community. However, only unpolarized processes can be simulated, so a study of asymmetry extraction is currently not possible with this program. It should also be mentioned that the program does give information about which hard scattering process exactly was simulated in each event, be it photon gluon fusion or QCD compton scattering, but PYTHIA is based purely on empirical data of previous experiments, not on a unified model. Therefore care must be taken when interpreting this internal information.

In a second step the generated events are fed into a detector simulation package named COMGEANT, which is a COMPASS adaption of the successful GEANT package [71]. The main purpose of this program is to simulate the interaction of the produced particles with the material present in the detector and generate artificial detector responses based on the known characteristics of the sensitive areas of tracking detectors, calorimeters and

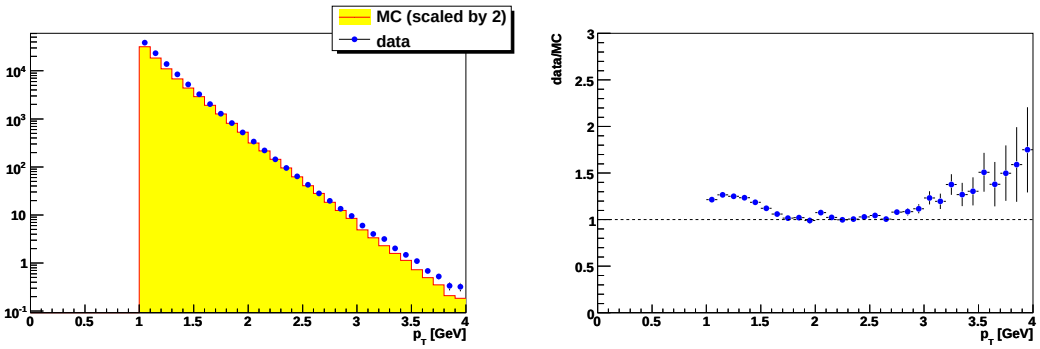


Figure 7.15: Distribution of p_T in real data and Monte Carlo simulation (left), and the ratio of the two (right). The Monte Carlo distribution has been scaled by a factor of 2, see text for details.

so on. Energy loss mechanisms like bremsstrahlung and random small angle deflection by multiple scattering are implemented as well as particle decay, but also detector occupancy by uncorrelated background events—so-called *pile-up*—and beam halo tracks are modeled. As a first use of the generated artificial detector response the data are used to simulate the decision process of the COMPASS trigger system.

It would be nice if the output of the second step would be artificial experiment data in exactly the same format as the real data, so that the unmodified reconstruction code could be used. But in addition to historical reasons and the question, where in the artificial data the underlying Monte Carlo truth should be stored, it is practical to use a different transport format. The simulation of the individual detector's response is closely related to the reconstruction code for that detector, since both are based on the detailed detector analysis performed by the respective experts. Thus, it makes sense to keep the code for simulation and reconstruction close together in the reconstruction package, where then the last step of the simulation—called digitization—happens.

Two Monte Carlo samples have been generated, each containing about 10 million events. The generator allows cuts on the hard scattering kinematics to be done at a very early stage of the simulation, speeding up the processing when only a part of the physically allowed phase space is to be populated. In this case, a cut has been placed on $Q^2 < 0.6\,\mathrm{GeV}^2/c^2$, allowing some room for event smearing which slightly changes the kinematics after the generation of the hard cross section. After the simulation of each event, a user-supplied selection routine is called for further cuts. Here, at least one hadron with a transverse momentum of $0.7\,\mathrm{GeV}/c$ has been required for the first sample. The simulated luminosity[9] of this sample is $30.7\,\mathrm{pb}^{-1}$, which is much smaller than the recorded experiment luminosity. Because of the extremely steep slope of the cross section in p_T, statistical errors from the Monte Carlo sample would become significant in the extracted cross section uncertainties for large p_T, therefore the second sample has the p_T cut raised to $1.7\,\mathrm{GeV}/c$, leading to a simulated luminosity of $3798\,\mathrm{pb}^{-1}$. Using this sample for $p_T > 2\,\mathrm{GeV}/c$

[9] PYTHIA writes the integral of the simulated cross section to the standard output, together with the number of events generated, disregarding the decision of the user-supplied selection routine. The ratio of these two numbers is the luminosity (see eq. (7.37) with $\epsilon = 1$).

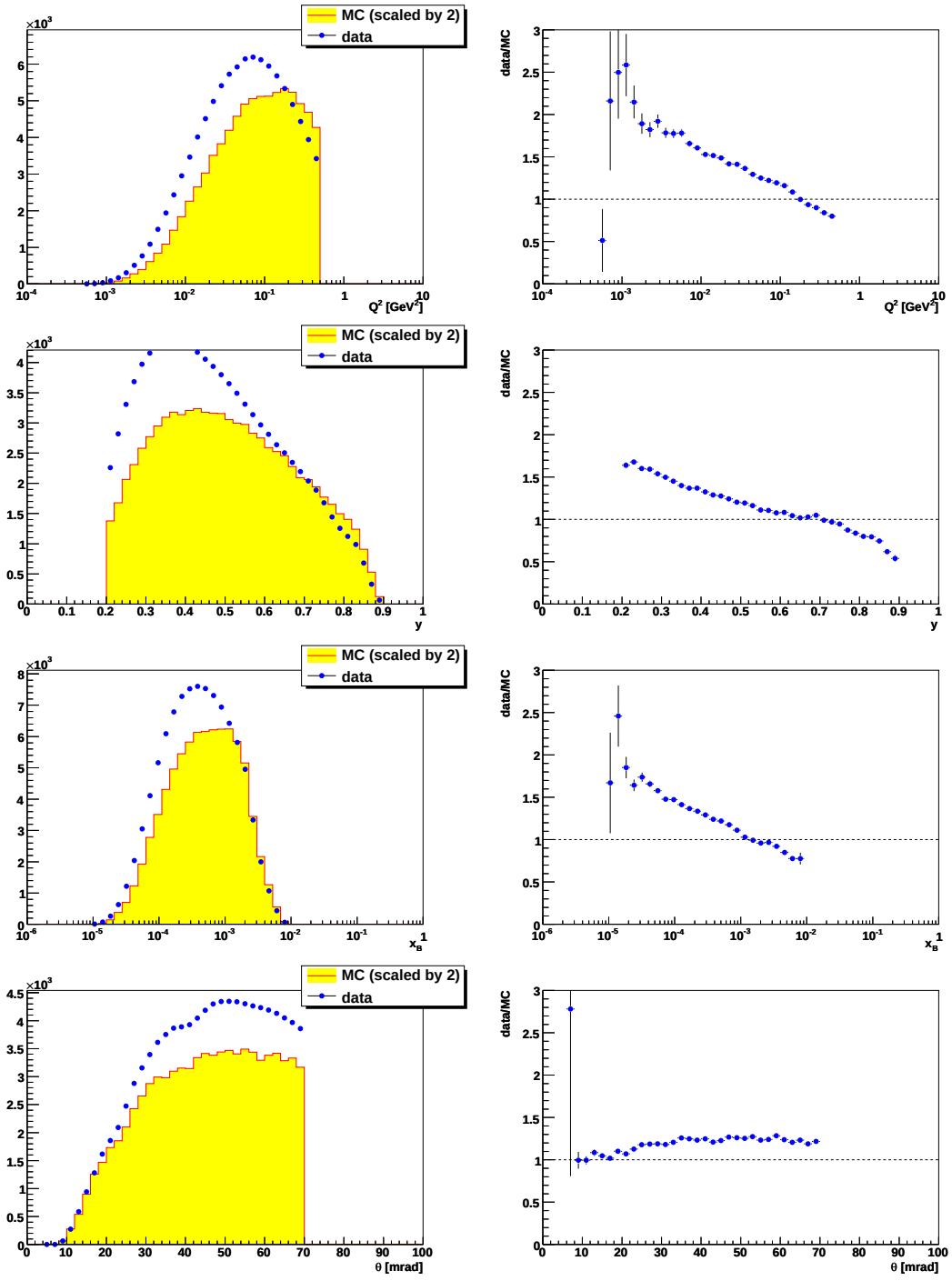


Figure 7.16: Distribution of kinematic variables in simulated and measured data for $p_T > 1$ GeV/c. The plots on the left side show the distributions themselves while the plots on the right hand side show the ratio of simulated over measured counts. The Monte Carlo distribution has been scaled by a factor of 2, see text for details.

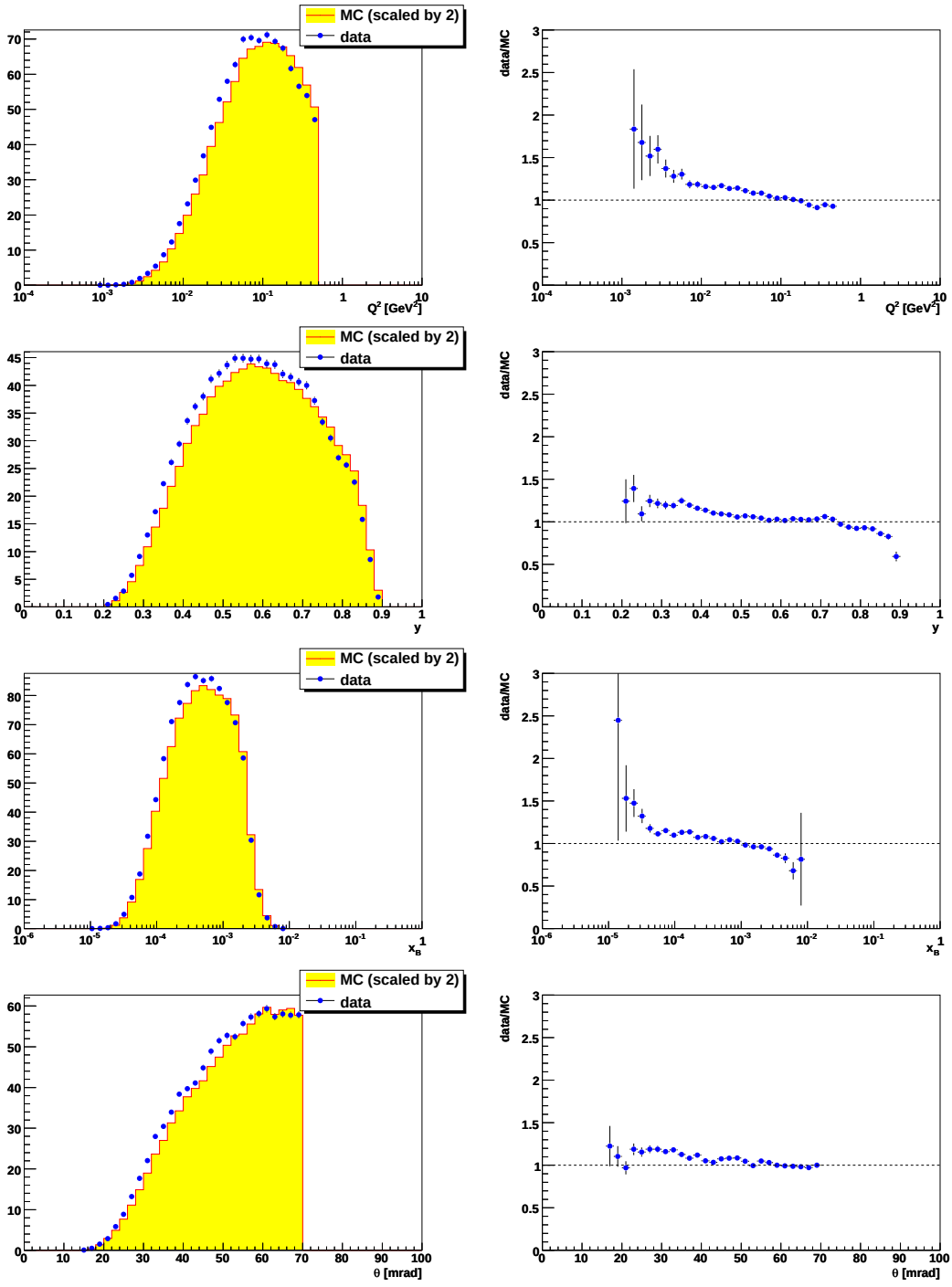


Figure 7.17: Distribution of kinematic variables in simulated and measured data for $p_T > 2\,\mathrm{GeV}/\mathbf{c}$. The plots on the left side show the distributions themselves while the plots on the right hand side show the ratio of simulated over measured counts. The Monte Carlo distribution has been scaled by a factor of 2, see text for details.

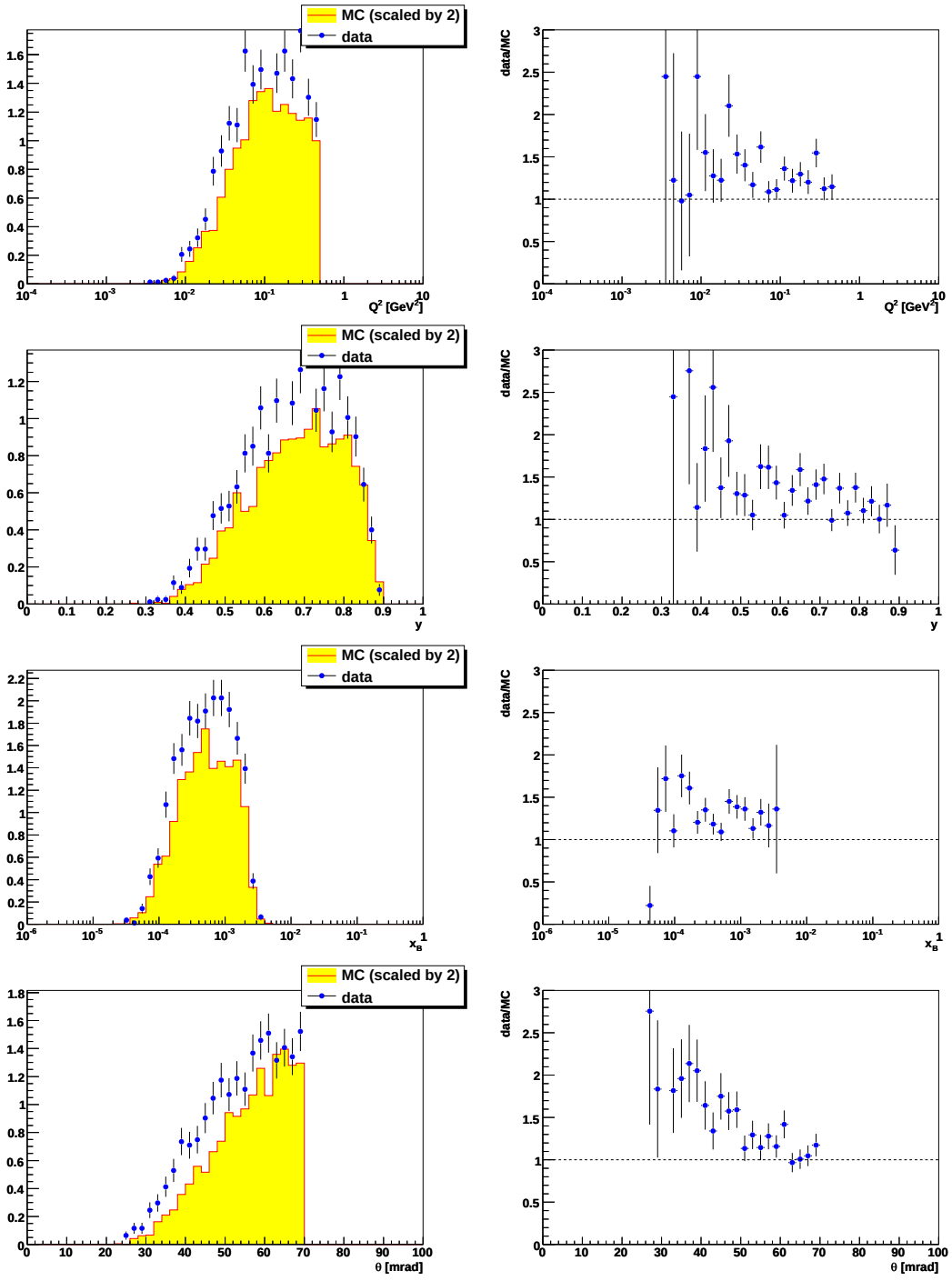


Figure 7.18: Distribution of kinematic variables in simulated and measured data for $p_T > 3\,\mathrm{GeV}/c$. The plots on the left side show the distributions themselves while the plots on the right hand side show the ratio of simulated over measured counts. The Monte Carlo distribution has been scaled by a factor of 2, see text for details.

ensures, that even at the highest p_T the statistics of the Monte Carlo sample are not dominating the uncertainty on the results. The safety distance of 0.3 GeV/c between the generator cuts and the used regions is again due to event smearing; however, from figure 7.15 it can be seen that there still is an effect seen in the first bin of the second sample ($2\,\text{GeV}/c < p_T < 2.1\,\text{GeV}/c$), where the generator cut reaches into the region used in the analysis, thereby artificially reducing the Monte Carlo events and increasing the value of the ratio to the measured cross section by about 5–10%.

The last step is to process the simulated data with the same reconstruction algorithms that are used for real data and apply the same analysis procedures for event selection, histogramming and so on. Figure 7.15 shows the comparison of the simulated and measured cross sections, differential in p_T. The absolute normalization of the Monte Carlo cross section has been scaled by a factor of 2 in this comparison for two reasons. First, the ratio of measured over simulated cross section is remarkably constant between 1.5 GeV/c and 3 GeV/c, only the normalization is off. Thus, for better visual comparison of the slopes, the scaling brings the two curves closer together. Second, PYTHIA employs only leading order matrix elements with partial incorporation of NLO effects via so-called *parton showers* (QCD bremsstrahlung). The difference in cross section observed in the theoretical calculations presented in section 6.6.3 between leading and next-to-leading order is very close to a factor of 2, as can be seen in the K-factors in figure 6.6. This could of course be coincidence, but it makes the comparison very interesting this way.

The structure in figure 7.15 below 1.5 GeV/c does not come as a surprise, considering that for photo-production the only hard scale in the process becomes rather small in this region. The NLO calculations are by their authors believed to break down if p_T is much below 2 GeV/c, and the leading order description in PYTHIA should not behave better in this respect. Figure 7.16 shows the distributions of the kinematic variables for $p_T > 1\,\text{GeV}/c$, which because of the steeply falling slope is dominated by hadrons with $p_T < 2\,\text{GeV}/c$. It can be seen that the mismatch in p_T goes together with an overall change in event kinematics, or put more simply, that the shape of the cross section is not well described. Therefore, the region of the first Monte Carlo sample—$p_T < 2\,\text{GeV}/c$—will not be further discussed concerning the extraction of the acceptance correction.

The deviation in figure 7.15 above 3 GeV/c could be caused by *radiative effects*, see section 7.4.3 for a discussion. The increasing discrepancy with higher p_T would in this scenario correspond to a radiative "background", diluting the true high p_T events since polarization effects are expected to vanish for $p_T \to 0$. Therefore, the asymmetries extracted later at $p_T > 3.5\,\text{GeV}/c$ need to be carefully interpreted.

The comparison of the single-inclusive deep inelastic scattering kinematics in figure 7.17 shows, that for $p_T > 2\,\text{GeV}/c$ the data are described within $10 \div 20\%$ by the Monte Carlo simulation, apart from the general factor of 2 in the absolute normalization. Of course, the plots are dominated by the smaller p_T, values below 2.5 GeV/c accounting for more than 90% of the sample. Thus, figure 7.18 contains the same comparison for $p_T > 3\,\text{GeV}/c$. Apart from the relative abundance of data discussed in the previous paragraph, the distribution of the event kinematics is still well described.

Given that the Monte Carlo sample is compatible with the high p_T sample in all kinematic respects, the last component needed for the acceptance correction is the comparison between simulated and reconstructed hadron counts in each bin. The cuts on the reconstructed side have been detailed in the previous sections and are derived from experimental necessities and the desire to match simulation and real data. On the simulation side the sample should contain the generated cross section within the cuts done also in the theoretical calculations. The latter are $0.2 < y < 1 \wedge \theta < 70\,\text{mrad} \wedge Q^2 < 0.5(\text{GeV}/c)^2$. The ratio obtained between reconstructed and simulated Monte Carlo events should then be a good approximation of the ratio needed in the extrapolation from the observed hadron counts to the theory cross section, irrespective of the differences in the underlying model for theory and Monte Carlo generator.

The result for the acceptance correction is shown in figure 7.20. The general trend towards better reconstruction efficiency for larger values of p_T is consistent with the observation that vertex finding and fitting works better the larger the angles between participating tracks are. As the virtual photon direction is mostly collinear with the beam muon's momentum, a larger transverse momentum of the hadron with respect to the virtual photon implies larger angles between the hadron and the incoming muon. The overall efficiency of around $30 \div 40\%$ is also consistent with the expectation for a spectrometer like COMPASS.

While the extraction of the acceptance correction is justified only in the region $p_T > 2\,\text{GeV}/c$, the smoothness of the curve obtained and the expected general trend suggest that even at lower p_T, the extracted values are not completely off. It seems unreasonable to assume that the acceptance correction for $p_T = 1 \div 2\,\text{GeV}/c$ is wrong by more than 10%, given that the discrepancy between the generated and measured cross sections is mostly an underestimation by 20% (cf. figure 7.15). Therefore, the region below $2\,\text{GeV}/c$ is included in the result plots, albeit separated by a dashed line.

7.4.3 Radiative Effects

As is known from classical electrodynamics, accelerated charge radiates electromagnetic waves. In the deep inelastic scattering process under consideration, the muon loses a substantial part of its very high energy in a very short time. Likewise, the produced hadron is formed from the nucleon at rest and quickly accelerated nearly to the speed of light in the laboratory frame. Therefore, momentum and energy loss by radiation need to be considered when evaluating the matrix element. The effects are twofold: the amplitude for the process is changed by allowing low energy radiation in addition to the non-radiating graph, and the kinematics of the radiating particles are changed for a fraction of the events, in which the emitted radiation has sizable energy.

The first effect can probably be expressed as a small correction, which should be independent of the lepton and quark helicities involved and therefore cancel in the asymmetries. However, a photon emitted by the muon before the scattering process would invert the effective beam polarization for this event, leading to a dilution of the asymmetry. It has

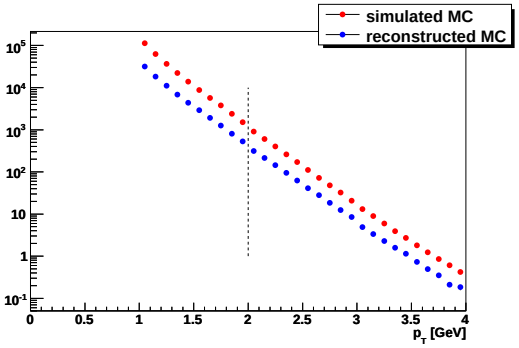


Figure 7.19: Distribution of p_T in the Monte Carlo sample for generated and reconstructed hadrons.

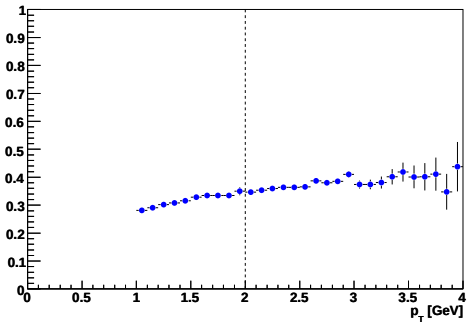
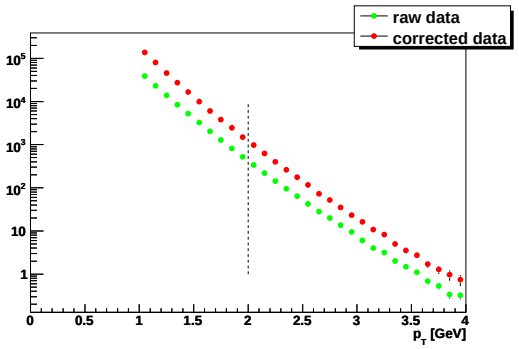


Figure 7.20: Acceptance correction factor obtained as ratio of the two distributions to the left.



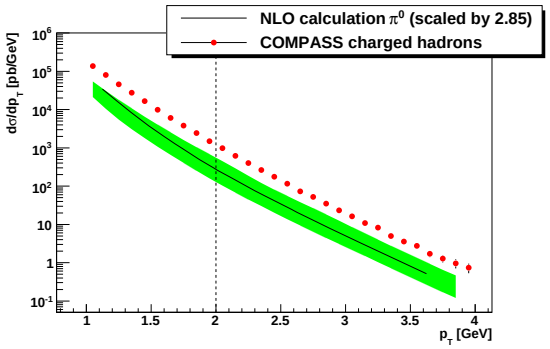


Figure 7.21: The left plot shows the comparison between the raw and corrected cross sections while the right plot includes the theoretical predictions for neutral pion production.

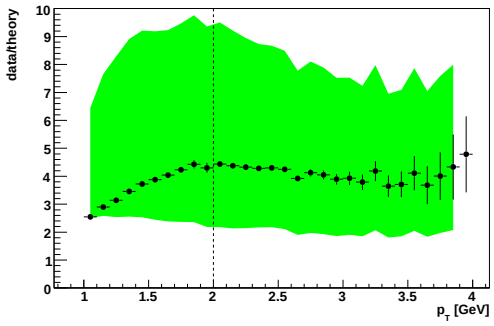


Figure 7.22: Ratio between measured cross section and NLO predictions. The gray band corresponds to the scale dependence shown in figure 6.7, the points to the choice of all scales equaling p_T. All theory curves have been scaled by the factor 2.85 derived in section 6.6.3.1.

been investigated which fraction of the DIS events are affected by radiative effects, without distinguishing pre- and post-reaction radiation, and the number was found to be of the order of 7% [72]. The model used was a PYTHIA Monte Carlo simulation coupled with a code which evaluates the relevant radiative graphs with the simulated kinematics to obtain a radiation probability.

The second effect has greater potential to distort the reconstructed cross section. Bremsstrahlung emitted by the muon before or after the scattering would change the reconstructed event kinematics from the real ones governing the hard scattering process. This could cause some of the abundant low p_T hadrons to be reconstructed with large p_T, owing to a wrongly reconstructed direction of the virtual photon. QED radiation of the produced hadrons would also change their direction, but since p_T is proportional to the longitudinal momentum, the energy loss by radiation would probably bias the transverse momentum change towards smaller values. It is unclear, however, how big coherent bremsstrahlung effects between the muon and the hadron are, as their relative phases depend on the details of hadron formation. Thus, it would be desirable to calculate the amplitude for both parts, $I_\mu = |\mathcal{M}_\mu|^2$ and $I_h = |\mathcal{M}_h|^2$, and estimate the range of the coherent part by $|I_\mu - I_h| < I_{\text{coherent}} < I_\mu + I_h$.

7.4.4 Result

Figure 7.21 shows the resulting measured cross section after acceptance correction in comparison with the uncorrected values and with the theoretical calculations. The scale on the ordinate is absolute as it includes the evaluated COMPASS luminosity. Using the factor between neutral pion and charged hadron cross sections obtained in section 6.6.3.1, the theory curve is scaled and compared to the measured cross section in figure 7.22, including the scale dependence as a colored band. In the region $p_T > 2\,\text{GeV}/c$, the ratio between experiment and theory is approximately constant for the lower edge of the scale band as well as the central curve, while the upper edge exhibits a slightly negative trend. Interestingly, there is no evidence for a relative excess of measured over calculated cross section at the highest p_T, analog to the one seen in the comparison between data and Monte Carlo in figure 7.15 on page 97. Whether this difference between the NLO calculations and the PYTHIA simulation is due to true higher order effects, or caused by the different description of the virtual photon spectrum, is unclear. But it should be noted, that none of the two explicitly include radiative effects.

In summary, the slope of the cross section is well described by the calculations, with rather large uncertainties and deviation in the absolute normalization. However, the authors state in [48]:

> "We note that all-order resummations of large logarithms in the perturbative series which appear when the initial partons have just enough energy to produce a high-p_T pion and a recoiling massless "jet" may lead to a considerable enhancement of the cross section at fixed-target energies as was recently demonstrated

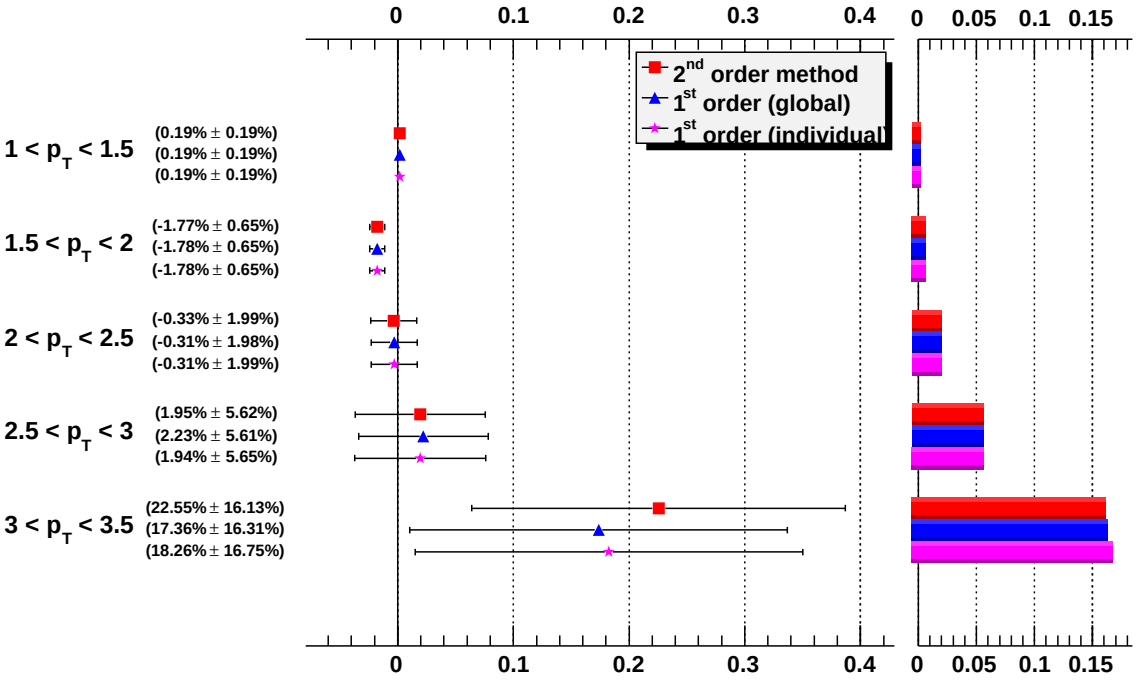


Figure 7.23: Muon–nucleon asymmetry $A_{\mathrm{LL}}^{h,d}$ in bins of p_T using different extraction methods. The error bars are only statistical. The pulls for the individual asymmetry measurements pertaining to each data point are shown in figures D.1–D.3 on pages 150f.

> for the process $pp \to \pi X$ [73]. Similar calculations for the case of photoproduction are not yet available but certainly desirable."

A tendency towards an underestimation of measured cross sections by NLO calculations has been observed in several other cases, generally increasing at lower center of mass energies, see [73] and references therein. The resummation has been shown to bridge the gap between theory and experiment to a large degree. Since it is basically an exponentiation of soft gluon radiation, it should not depend on the polarization of the quarks, and effects of changing the relative weights of the contributing sub-processes in the hard scattering are expected to be small [74]. Thus, despite a shortfall in predicted cross section, the calculated asymmetries should be comparable to measurement.

Therefore, I conclude that the measured single-inclusive cross section is well described by the NLO calculations, so that a comparison of the extracted asymmetries to the prediction for different ΔG should be tried in order to add to our knowledge about the gluon polarization.

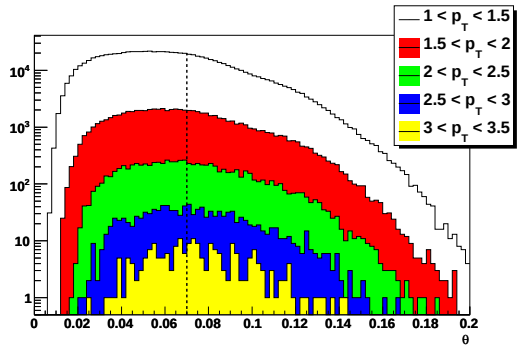


Figure 7.24: Distribution of the production angle of the hadron with respect to the incident muon momentum, in the bins in p_T as used for asymmetry extraction.

7.5 Asymmetries

Using the analysis techniques outlined in the previous sections, the muon–nucleon asymmetry $A_{\mathrm{LL}}^{h,d}$ shown in figure 7.23 is extracted. Because of the too small dispersion of fP_{beam} shown in figure 7.1 on page 80, no weighting has been employed, as discussed in section 7.2.5. The three algorithms for asymmetry extraction yield nearly identical results, which serves as an internal cross check, keeping in mind that the 1st and 2nd order methods differ substantially in the formulas used. At the highest p_T, the 2nd order method performs marginally better, owing to its slightly better use of the available statistics. The difference in asymmetry observed between the 1st and 2nd order methods in the highest p_T bin is significant, taking into account that the same data are used and therefore the slight difference in statistical uncertainty only allows fluctuations of $\delta A = 0.006$ according to section B.4. Other analyses have noticed a tendency towards numerical instability for the 2nd order method, when applied to data samples with very few events. This might also be the cause of the observed difference in this case, however, the pulls distribution (figure D.1 on page 150) does not indicate an anomaly.

The cut on the production angle θ of the hadron with respect to the incident muon momentum nearly halves the statistics. The cut at $\theta < 70$ mrad had been introduced in the calculation as a rough model for the COMPASS spectrometer acceptance, given by the downstream opening of the target solenoid magnet, but figure 7.24 shows our data to reach much further in that parameter. The calculations for the new target magnet, which has an opening angle of 180 mrad and is in use since 2006, have revealed that this improved acceptance does not necessarily increase the significance of the obtained asymmetries as the change in covered phase space also changes the weight of the polarized gluon contribution—in this case in an unfavorable way. Nevertheless, the asymmetry obtained by omitting the θ cut is shown in figure 7.25. While the individual data points are statistically compatible with the ones from the standard cut set (in the bin $1.5\,\mathrm{GeV}/c < p_T < 2\,\mathrm{GeV}/c$ the difference amounts to 2.8σ), it does seem that there is a trend towards smaller absolute values, similar to a dilution of the asymmetry: the χ^2 of the data with respect to a constant asymmetry of zero is 10.5 with and 5.7 without the

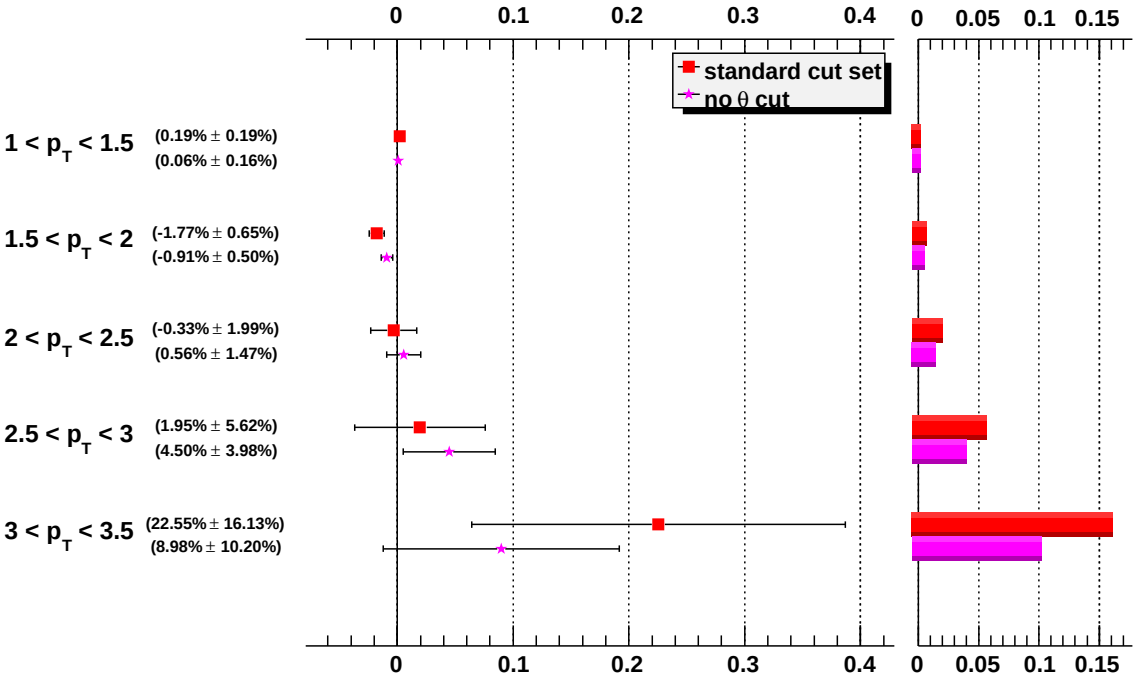


Figure 7.25: Muon–nucleon asymmetry $A_{\mathrm{LL}}^{h,d}$ without cut on the hadron production angle θ with respect to the incoming muon direction. The pulls for the individual asymmetries pertaining to each data point are shown in figure D.4 on page 153.

cut on θ.

7.6 Systematic Studies

The ability to invert the target polarization without changing the experimental setup and the magnetic fields provides a way to investigate false asymmetries which are not canceled by the extraction method. The most obvious source of such asymmetries is the change in magnetic fields in the target region when doing a fast polarization flip, leading to differences in charged particle tracking downstream of the target. Another effect is the movement of the target itself, which amounts to several hundred micrometers. In general every effect that is caused only by the change in target solenoid field direction causes a reproducible false asymmetry. By inverting the association between magnetic field and target polarization direction using the microwave reversal described in section 2.2.5, these false asymmetries cancel in the unweighted average of the asymmetries measured before and after the reversal.

However, not all systematic effects are only coupled to the change in solenoid sign, and also for reasons of optimum exploitation of the available statistics not the arithmetic but the weighted mean is used to combine the measured asymmetries, making the cancellation of the reproducible false asymmetries imperfect (cf. eq. (7.39)). The following sec-

tion discusses the quantitative consequences of these points and gives an estimate on the resulting systematic error. Section 7.6.2 attends to studies done using different cut sets, which contribute another part of the systematic error. All these investigations are not suited to quantify a possible universal offset in the asymmetry extraction, which is considered in section 7.6.3.

7.6.1 Statistical Estimate of the Systematic Error

The basis for this study are the individual run group asymmetry values and their statistical uncertainties grouped by beam time and microwave setting. Let A_+ be the overall asymmetry obtained with the positive microwave setting during one beam time and A_- its counterpart after microwave reversal. Considering only effects due to the solenoid field inversion, these two can be expressed in terms of the true asymmetry A_{phys} and the reproducible false asymmetry A_{rep}

$$A_+ = A_{\text{phys}} + A_{\text{rep}} \qquad A_- = A_{\text{phys}} - A_{\text{rep}} \quad \Longrightarrow \quad A_{\text{rep}} = \frac{1}{2}(A_+ - A_-) \tag{7.39}$$

The reproducible false asymmetry would cancel completely in the arithmetic average of A_+ and A_-, but because the statistics recorded within those two samples are not equal it is preferable to use the weighted mean. The difference between those two amounts to

$$\frac{w_+A_+ + w_-A_-}{w_+ + w_-} - \frac{1}{2}(A_+ + A_-) = A_{\text{rep}}\frac{w_+ - w_-}{w_+ + w_-} \tag{7.40}$$

and must be treated as a systematic error, albeit a very small one for the three beam times under consideration. With the reproducible false asymmetry measured, the individual data points in the A_+ sample can be corrected by subtracting A_{rep}, those pertaining to A_- by adding it. If no uncorrelated systematic effects remain, the ensemble of all modified individual asymmetry values should exhibit statistical fluctuations according to their uncertainties. But experience dictates that no experiment can be run at exactly stable conditions over many weeks and every time some characteristics of the spectrometer change non-uniformly during a run group, the extracted asymmetry will be affected. These small per–run group effects add to the fluctuations, an effect which can be quantified by calculating the χ^2 of a constant fit to the individual data points. Assuming a Gaussian distribution for these effects is not too far fetched—at least bigger false asymmetries should be more rare as we always strive for stability. Besides, a Gaussian distribution is the only model which is easy to calculate and interpret. The details of the algorithm are presented in section B.5, its essence is to find the Gaussian width ρ of the uncorrelated false asymmetries such that the χ^2 becomes compatible with the assumption of the constant fit on the 1σ level. Please note that this involves adding systematic and statistical errors in quadrature, not scaling the statistical uncertainties with a common factor.

Application of this method to the three beam times leads to the values shown in tables E.5–E.8 on pages 168ff. The tables show for each p_T bin and year the reproducible false

year	relative weight
2002	15%
2003	31%
2004	54%

Table 7.6: Statistical weight of the three beam times as extracted from the statistical uncertainty on the asymmetries extracted for the bin $1\,\mathrm{GeV}/c < p_T < 1.5\,\mathrm{GeV}/c$.

	2nd order		1st order (global)		1st order (ind.)		without θ cut	
	δA_{syst}	δA_{stat}	δA_{syst}	δA_{stat}	δA_{syst}	δA_{stat}	δA_{syst}	δA_{stat}
$1 < p_T < 1.5$	0.0014	0.0019	0.0014	0.0019	0.0014	0.0019	0.0013	0.0016
$1.5 < p_T < 2$	0.0038	0.0065	0.0038	0.0065	0.0038	0.0065	0.0030	0.0050
$2 < p_T < 2.5$	0.0122	0.0199	0.0122	0.0198	0.0121	0.0199	0.0046	0.0147
$2.5 < p_T < 3$	0.0275	0.0562	0.0265	0.0561	0.0249	0.0565	0.0199	0.0398
$3 < p_T < 3.5$	0.1115	0.1613	0.0858	0.1631	0.0799	0.1675	0.0657	0.1020

Table 7.7: Summary of systematic errors obtained by the statistical approach described in section 7.6.1. For more details refer to tables E.5–E.8 on pages 168ff. The statistical uncertainties are shown for comparison.

asymmetry A_{rep} with its contribution to the systematic error $(\delta A_{\mathrm{syst}})_{\mathrm{rep}}$, the 1σ interval for the size of random false asymmetries ρ together with the significance of the effect found, the contribution to the systematic error based on the 1σ upper bound on ρ, and finally the sum in quadrature for each year, both excluding and including the multiplication with the year's weight as given in table 7.6. Within each bin of p_T, the contributions from the three years are added in quadrature, resulting in the total systematic error.

Apart from the lowest p_T bin, the significance of random false asymmetries as assessed with this method is in all but one case below 1σ, and even the worst case only has 3.1σ. The calculated systematic errors are in all cases smaller than the statistical ones, typically around half the size. Due to the statistical nature of the approach, this is the expected outcome in case of no systematic effects. Therefore the conclusion is that the obtained estimates can be regarded as upper bounds on the systematic error.

This method is complemented by the so-called pulls method, where each individual value's distance from the central value is divided by its uncertainty and the obtained quantity is entered into a histogram. The expectation in case of purely stochastic fluctuations around the central value and correctly determined uncertainties is that the pulls should be distributed following a Gaussian function centered at zero with a width of one. The histograms for the pulls of all presented data points are shown in appendix D. The histograms are fitted with a normalized Gaussian function, finding in all cases mean, width and constant factor compatible with 0, 1 and 25, respectively. The latter is because the number of bins has been adjusted to 40% of the number of entries for each histogram (rounded down). It has been checked that the RMS of the distribution is compatible with the width of the Gaussian fit, and that the χ^2 probability does not indicate a bad fit. Interestingly, the distributions are closer to Gaussian than statistically expected, which manifests in that the χ^2 is in nearly all cases smaller than the number of degrees of freedom. As a conclusion, the pulls do not indicate any statistical irregularities in the

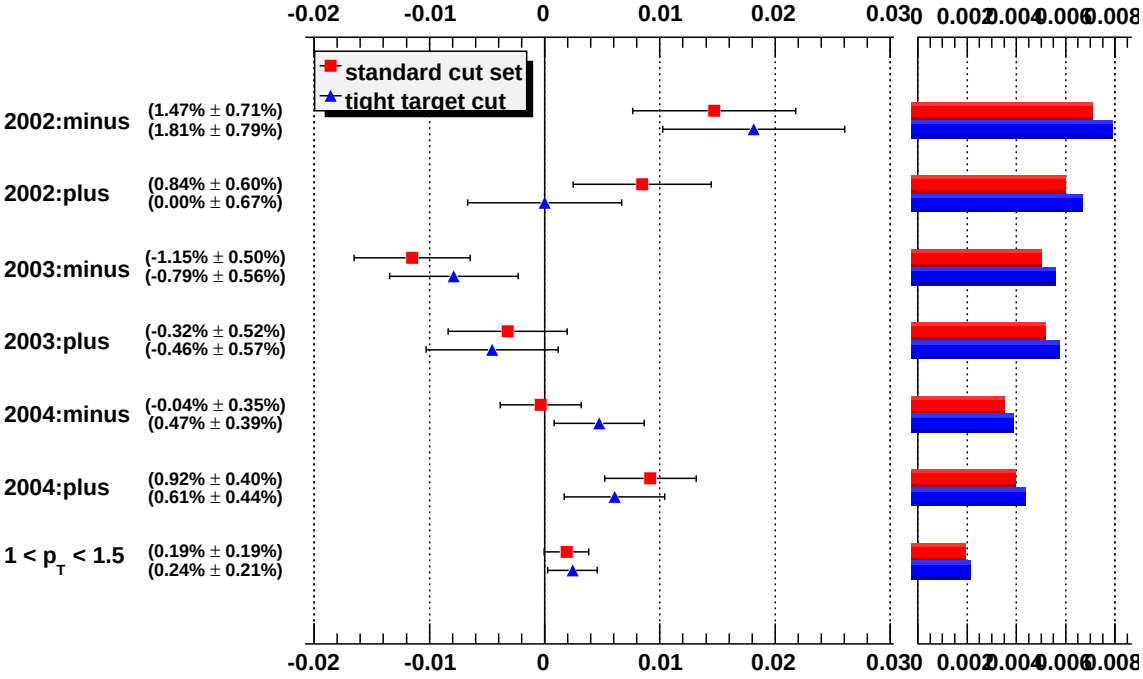


Figure 7.26: Study of different cut sets in the lowest p_T bin ($1\,\mathrm{GeV}/c < p_T < 1.5\,\mathrm{GeV}/c$) split by beam time and microwave setting; the lowest points represent the weighted average of the asymmetries displayed above. For an in depth discussion see section 7.6.2. The pulls for the individual asymmetries pertaining to each data point are shown in figure D.13 on page 162.

asymmetry extraction.

7.6.2 Different Cut Sets

False asymmetries, caused e. g. by irregular behavior of the apparatus under a solenoid field reversal, can manifest in specific sub-samples. Therefore, by varying cuts or splitting the sample in ideally equivalent parts, such effects can be detected. For this purpose, a tighter definition of the target volume has been studied as well as dividing the spectrometer in halves along the x or y direction. The first p_T bin has been chosen, because there the statistical resolution is best for detecting possibly small effects. The outcome of these studies is presented in the following sections.

7.6.2.1 Target Cuts

The definition of the exact target volume is challenging since there is no way to directly measure the position of the target cage inside the microwave cavity. Because of the necessity to minimize the amount of material used in holding the target, the mechanical construction is not stable enough to ensure a positioning accuracy on the sub-millimeter level, which would be desirable regarding the small transverse dimensions of the target

cells. Therefore, in an effort of the COMPASS off-line group, the density of reconstructed vertexes has been used to discriminate between the helium, the target cage frame and the target material inside the closed microwave cavity. During this investigation it became clear that not only the target has been mounted in different positions each year but also that the cage is not completely filled with ^{6}LiD. Hence after shearing the design target volume to match the measured positions, also 3 mm at the top of the target have to be cut off. However, the values of the geometrical correction parameters are not only limited in precision due to limited statistics, but also due to systematic effects e. g. in vertex reconstruction. This is why instead of a radius of 15 mm only 14 mm are used and the vertical cut for the target filling level is lowered from 12 mm above the axis to 10 mm.

It has been argued that even these conservative values are not far enough away from the real boundaries to ensure a homogeneous filling of the abstract analysis target volume with real ^{6}LiD in all conditions, so another set called *tight target* has been introduced with a radial cut at 12 mm and a vertical cut at 8 mm. The effect of this tightening is displayed in figure 7.26. The relative increase in uncertainty incurred by the cut is 11%, allowing for fluctuations of up to 43% of the systematic uncertainty for the smaller sample (1σ confidence level), according to section B.4. The observed changes in asymmetry amount to up to 3σ, but are quite well balanced, decreasing the reproducible false asymmetry—i. e. the difference between the asymmetries for the two microwave settings within each year—in 2003 and 2004 while increasing it in 2002. The net effect for the complete data set, however, amounts only to 0.6σ and thus is not significant.

7.6.2.2 Subdividing the Spectrometer

The COMPASS spectrometer is nearly symmetric around the beam axis for positive particles traveling at a momentum of 160 GeV/**c**, owing to the necessity of allowing the unscattered beam particles to travel through as little material as possible and avoiding to blind the detectors near the beam, thus introducing a "beam hole". Given the vertical magnetic field of the two spectrometer magnets, this symmetry is broken for slower particles in the left–right sense, but mostly retained in the up–down sense. For negative particles, the left–right symmetry is violated from the beginning, since their trajectories are bent in the opposite direction from the beam, making them leave the region of high-precision tracking earlier than their positive counterparts. This is reflected in a slightly lower reconstruction efficiency for negative particles, as can be seen from the slightly larger statistical uncertainties on the asymmetries shown in figure 7.27. The purpose of those two plots is to investigate false asymmetries by checking whether the acceptance canceling in the asymmetry extraction works even in the case of a significantly distorted apparatus symmetry.

Keeping in mind that the asymmetries from opposing halves of the spectrometer are uncorrelated, the only conspicuous difference in case of the positive hadrons (upper plot) is seen between upper and lower half in the year 2003 with microwave setting "minus". Without this 2.3σ effect, the difference between upper and lower half for the whole sam-

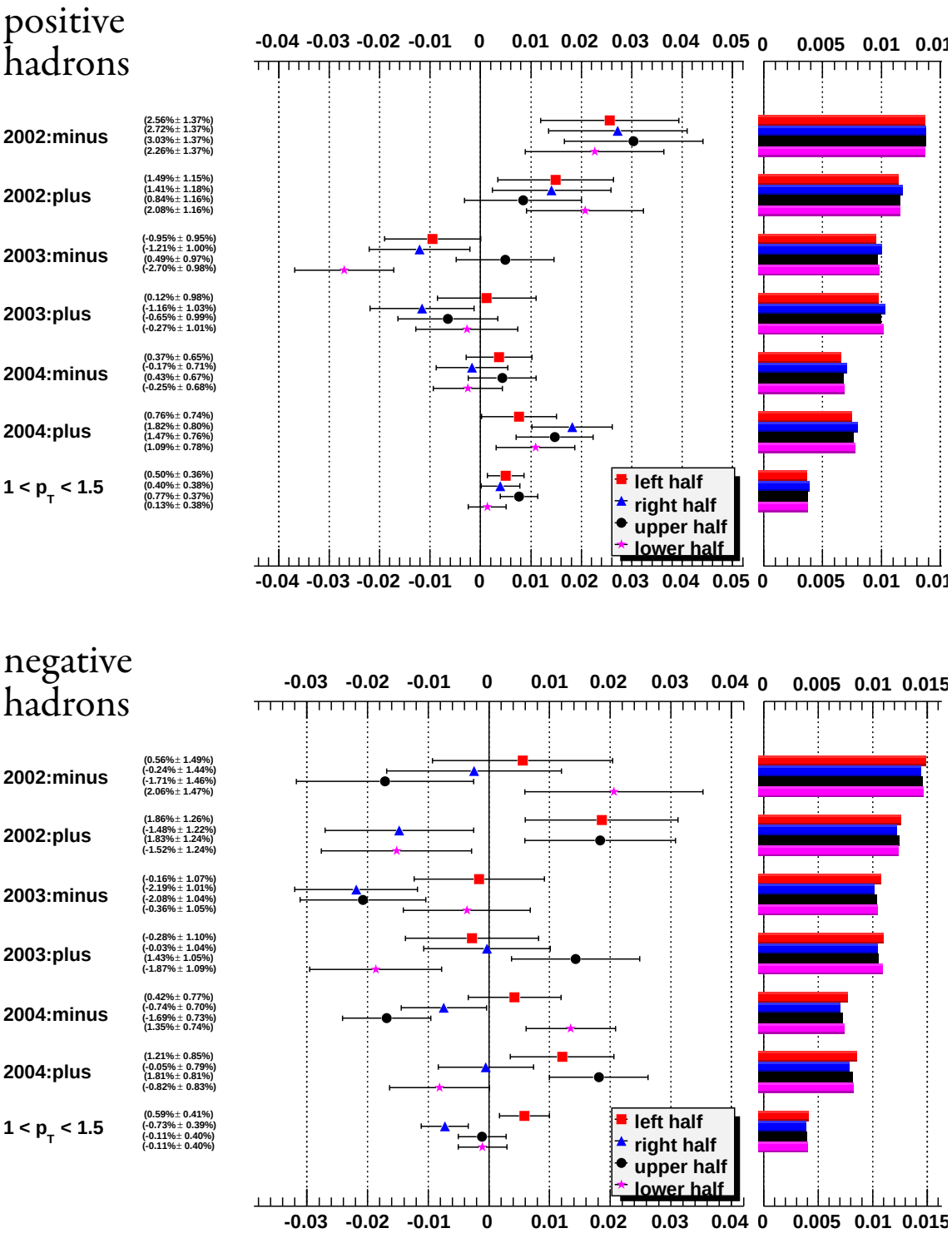


Figure 7.27: Studying spectrometer effects by dividing the acceptance in halves along the x or y axes. The produced hadrons are classified by their azimuthal angle around the beam direction. Left means positive x (Jura side), right means negative x (Salève side), up means positive y and down negative y. The upper plot shows the resulting asymmetries for positive hadrons, the lower plot for negative hadrons. The pulls for all data points are shown in figures D.5–D.12 on pages 154ff.

ple would be nearly zero, but even including it produces only a 1.2σ difference.

In the case of negative hadrons, the lower plot shows a richer structure. The magnetic interference of target solenoid and first spectrometer magnet produces an additional vertical magnetic field component, which either adds to or detracts from the field integral seen by the particles, depending on whether they travel downwards or upwards—along the bent field lines or with a small angle to them. Since the negative hadrons are already suffering from lower reconstruction efficiency in the outer spectrometer region on the Salève side, this difference leads to systematic acceptance changes which are correlated with the solenoid field direction. As discussed in section 7.2.4, this leads to a reproducible false asymmetry, which cancels in the average of both microwave setting. This is nicely demonstrated here, where the net difference between upper and lower half of the spectrometer is large for the individual sub-samples, but exactly zero for the complete sample.

Regarding the left–right division for negative hadrons, the individual differences are mostly of the order of 1σ, with the exception of the "plus" microwave setting during the year 2002, where an effect of 1.9σ is seen. However, since in almost all cases the asymmetry obtained from the right half of the spectrometer is smaller than that from the left half, the overall difference between left and right amounts to 2.3σ. This is the largest effect seen, but given its little significance it might still be a statistical fluctuation (the probability for 6 measurements lying on one side of the underlying value is still 3%).

7.6.3 A Possible General Offset

The methods used for quantifying possible systematic effects all rely upon the same algorithms for asymmetry extraction and provide only internal consistency checks. A truly general offset in measured asymmetries would go undetected. However, there is no reason to believe a priori that the asymmetry for $p_T < 1.5\,\mathrm{GeV}/c$ should be large, given e. g. the theory calculations which predict a very small asymmetry between -0.01 and 0.02, independent of the chosen ΔG. Therefore it can be concluded, that a general offset can at most be of the order of 0.02, given that the measured asymmetry in the lowest bin in p_T is 0.002.

7.6.4 Summary

Concerning systematic errors, different conventions are in use. Following [75], the conclusion would be that no systematic effects have been seen, so it does not make sense to quote a systematic uncertainty. On the other hand, the COMPASS collaboration prefers to publish upper bounds on the systematic uncertainty, estimated using statistical techniques. In case of no observed systematic effects, this leads to systematic errors which are roughly half the size of the statistical uncertainties, a connection that is quite unintuitive. Usually, the statistical uncertainty is decreased by taking more data, but the systematic uncertainty is given by the experiment and analysis method and should stay constant, unless the apparatus or the measurement technique can be further refined.

independent of A_{phys}	
statistical estimate	0.0014
target cut	0.0034
positive hadrons, up–down	0.0031
positive hadrons, left–right	0.0018
negative hadrons, up–down	0.0014
negative hadrons, left–right	0.0041
total	0.0067

proportional to A_{phys}	
dilution factor	5.0%
beam polarization	5.0%
target polarization	5.0%
total	8.7%

Table 7.8: Summary to the systematic uncertainty. The contributions independent of A_{phys} have been obtained from the lowest p_T bin and give upper bounds on the systematic error with a confidence level of 85%. The contributions proportional to A_{phys} have been determined by the COMPASS off-line group.

	A_{phys}	$(\delta A_{\text{syst}})_{\text{indep}}$	$(\delta A_{\text{syst}})_{\text{prop}}$	δA_{syst}
$1 < p_T < 1.5$	0.0019 ± 0.0019	0.0067	0.0002	0.0067
$1.5 < p_T < 2$	-0.0177 ± 0.0065	0.0067	0.0015	0.0069
$2 < p_T < 2.5$	-0.0033 ± 0.0199	0.0067	0.0003	0.0067
$2.5 < p_T < 3$	0.0195 ± 0.0562	0.0067	0.0017	0.0069
$3 < p_T < 3.5$	0.2255 ± 0.1613	0.0067	0.0195	0.0207

Table 7.9: Evaluation of the systematic uncertainties summarized in table 7.8 to the asymmetries extracted using the standard cut set and the 2nd order method.

Adhering to the experiment policy, it is preferable to use the most precise sub-sample for the determination of systematic uncertainties concerning the asymmetry extraction. Because of the steep drop in statistics for higher p_T, this effectively reflects that no systematic effects are seen where the statistical uncertainties are huge. The contributions to the statistically estimated part of the systematic uncertainty are given by the upper bounds calculated in section 7.6.1 and the absolute size of movements of the data points by the different cut sets examined in section 7.6.2. In the latter case, the difference ΔA obtained between data sets is calculated with error propagation, quoting $\Delta A + \delta(\Delta A)$ so that the values consistently correspond to a confidence level of 85%. The contributions from section 7.6.2.2 have been scaled by their relative weight in the measurement, as they affect only a sub-sample, e. g. negative hadrons.

Another source of systematic uncertainty comes from the dilution factor f and the calibration of polarization measurements for beam and target. These uncertainties amount to 5% each, as determined by the COMPASS off-line group, and translate directly into a 5% uncertainty in the asymmetry for each of the three parameters.

The contributions to the systematic error are summarized in table 7.8 and their effect is evaluated for the measured asymmetries in table 7.9.

7.7 Discussion

The comparison between the extracted asymmetries from figure 7.23 and the next-to-leading order predictions in figure 6.9 is shown in figure 7.28. Taking into account the systematic uncertainty displayed as a red band, the data are compatible with the standard Δg curve, as extracted using the GRSV fit [76]. However, especially the point at 1.65 GeV/c would favor a slightly higher gluon polarization, assuming that one can at least qualitatively interpolate between the four graphs. A substantially negative gluon polarization seems to be disfavored. This is corroborated by the left plot of figure 7.29, where the 68% confidence interval for $\Delta G/G$ is approximately $0.1 \div 0.7$ and $\Delta G/G < -0.2$ can be excluded on the 95% confidence level.

All of this discussion depends on whether the calculation of asymmetries can be trusted, while the unpolarized cross section comes out four times smaller than measured. If the "all-order resummations of large logarithms," hinted at in [48], are found to bring the unpolarized cross section to the measured magnitude, it is not completely clear whether the polarized cross section also is increased by the same factor. If not, that would change the prediction for the asymmetries and would probably lead to a different interpretation of the data.

The gluon polarization is not the only unknown in the NLO calculations, even less is known about the polarized structure of the photon. This is expressed in figure 7.28 in the dashed and solid lines, corresponding to two extreme scenarios for the photon PDF. While the difference obtained is very small at larger p_T, it dominates the gluon contribution at $p_T < 1.5$ GeV/c, where the statistical precision of the measurement is best.

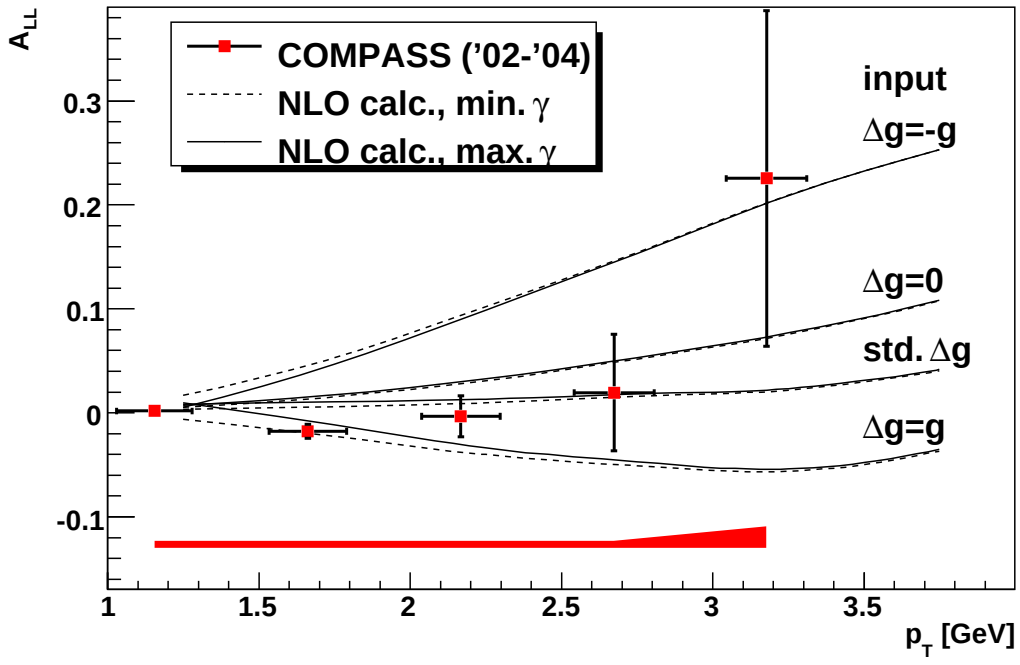


Figure 7.28: Measured double spin asymmetry in bins of p_T, compared to the predictions of three scenarios for the gluon polarization ΔG. The red band indicates the size of the systematic uncertainties of the measurement as detailed in table 7.9.

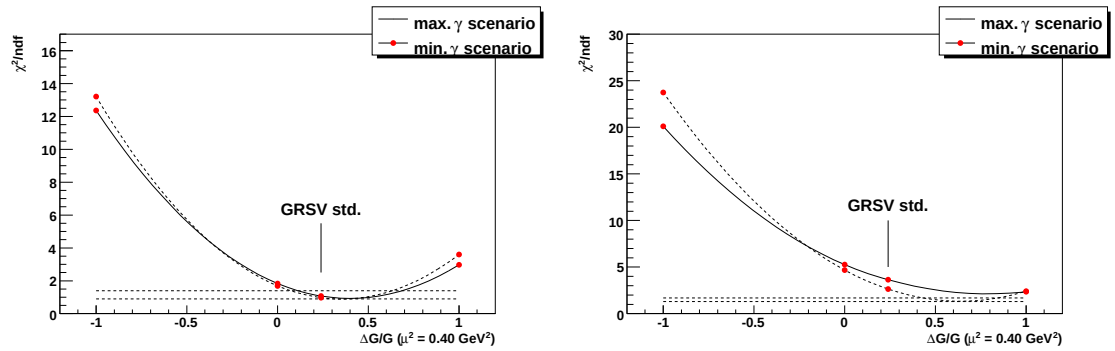


Figure 7.29: Reduced χ^2 calculated for the four theory curves in figure 7.28 with respect to the three (left) or four (right) largest-p_T data points. The number of degrees of freedom is two or three, respectively, accounting for ΔG as the only parameter. The distance between the horizontal dashed lines is one in units of χ^2.

Therefore, given reliable calculations for this kinematic range, COMPASS would be in the position to constrain the polarized photon PDF. However, the current theoretical results seem not yet to be mature enough, as indicated by the large scale dependence and K-factors shown in section 6.6.3.

If new calculations can reproduce the measured unpolarized cross section and improve their reliability at lower p_T, it will be most interesting to perform a real fit of the gluon polarization and of the polarized photon PDF. Unfortunately, the computational cost will probably be too high, so that discrete χ^2 distributions might have to suffice, as have been published in the RHIC case already.

Chapter 8

Conclusions

In the course of this thesis, three core software packages for the COMPASS experiment have been implemented and deployed. First was the config_server in the year 2002, responsible for programming the front-end electronics reading out most of the detectors and configuring the trigger control system TCS. Second was the online filter CINDERELLA, enabling COMPASS to record higher trigger rates by increasing the purity of the first level trigger decisions before the data are written to disk; this work has been done together with Thiemo Nagel, who took over responsibility for this project in 2004. The third was the data acquisition system, for which I took over responsibility in fall of 2005 from Lars Schmitt, who had just begun a major system upgrade with completely new hardware and software. After finishing the setup, the system has been maintained throughout the beam time of 2006 and—together with Damien Neyret—2007.

In the beam times of 2002–2004, COMPASS has recorded about 1.5 fb^{-1} of deep inelastic scattering with polarized deuteron target and polarized muon beam, allowing for unprecedented precision measurements of the spin structure of the nucleon. This thesis presents the first measurement of the polarized cross section for single-inclusive production of charged hadrons in deep inelastic lepton scattering, complementing earlier results from RHIC concerning the production of jets [77] or neutral pions [78] in pp collisions at $\sqrt{s} = 200\,\text{GeV}$. The measured unpolarized cross section matches the shape of a prediction [48]—obtained from calculations in next-to-leading order perturbative QCD—very well, but the absolute normalization is found to be four times as high as predicted. A possible explanation for this discrepancy is given by the authors, who believe that at relatively low center of mass energy, the contributions to the matrix element beyond the next-to-leading order contain sizable large logarithms arising from soft gluon bremsstrahlung. These should be summed up in all orders in order to obtain a physically meaningful result, a process which has been done for $p\,p$ scattering, but not yet for deep inelastic lepton scattering. It is expected, that the resummation is independent of spin polarizations, thus it is not unreasonable to assume that they will not severely affect the calculated double spin asymmetries.

The measured double spin asymmetry is found to be small, compatible with positive

gluon polarization within the model of the calculations. Evaluation of χ^2 for the four theory curves corresponding to different gluon polarizations yield a preference for a large positive gluon polarization. It should be noted that this statement depends on the GRSV model for the polarized gluon distribution functions and is based on the limited range in x_g covered by the COMPASS spectrometer. This range itself cannot easily be quantified, as the momentum fraction carried by the gluon cannot be directly measured and needs to be extracted using a model; it even depends on the polarized gluon distribution functions which we are trying to measure.

To further our knowledge on the composition of the nucleon spin, a joint effort of different experiments, operating in different kinematic regions, is needed, contributing the data for a global analysis. This encompasses fixed target muon scattering experiments with their high luminosity and well understood beam structure as well as polarized hadron colliders with their direct access to the gluon and to the parton kinematics, and will also require the extended x_B range offered by a polarized electron–proton collider like the proposed eRHIC project.

Appendix A

Formulas and Formalism

A.1 Light-Cone Coordinates

When dealing with fast particles it comes in handy to choose light-cone coordinates, in the following marked with a tilde:

$$\tilde{x}^{\pm} = \frac{1}{\sqrt{2}}(x^0 \pm x^3) \qquad \tilde{x}^{1,2} = x^{1,2} \qquad \tilde{g}^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{A.1}$$

$$x \cdot y = \tilde{x}^+\tilde{y}^- + \tilde{x}^-\tilde{y}^+ - \tilde{x}^1\tilde{y}^1 - \tilde{x}^2\tilde{y}^2 \tag{A.2}$$

Real photons moving in z direction then have $\tilde{p}^\mu = \Big(E, 0, 0, 0\Big)$. $\tilde{x}^+$ takes over the role of x^0 as evolution variable and quantization happens at equal $\tilde{x}^+$ instead of at equal time. It is convenient to choose a special representation of the Dirac matrices:

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & \sigma^3 \\ \sigma^3 & 0 \end{pmatrix} & \gamma^3 &= \begin{pmatrix} 0 & -\sigma^3 \\ \sigma^3 & 0 \end{pmatrix} \\ \gamma^1 &= \begin{pmatrix} \boldsymbol{i}\sigma^1 & 0 \\ 0 & \boldsymbol{i}\sigma^1 \end{pmatrix} & \gamma^2 &= \begin{pmatrix} i\sigma^2 & 0 \\ 0 & i\sigma^2 \end{pmatrix} & \gamma_5 &= \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \\ \gamma^+ &= \sqrt{2}\begin{pmatrix} 0 & 0 \\ \sigma^3 & 0 \end{pmatrix} & \gamma^- &= \sqrt{2}\begin{pmatrix} 0 & \sigma^3 \\ 0 & 0 \end{pmatrix} \\ \mathcal{P}_+ &= \tfrac{1}{2}\gamma^-\gamma^+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \mathcal{P}_- &= \tfrac{1}{2}\gamma^+\gamma^- = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{A.3}$$

The projection operators show that the upper components of the resulting spinors are

the "good", light-like components while the lower ones correspond to the violation of the light-likeness, which is suppressed like $\frac{M^2}{2p}$.

A.2 Gell-Mann Matrices

The Gell-Mann matrices are the usual choice of generators for the $SU(3)$ symmetry group. The particular representation makes it obvious that $SU(3)$ includes pair-wise $SU(2)$ of the components of the triplet:

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{aligned} \tag{A.4}$$

Appendix B

Statistics

This appendix is just a convenience collection of statistical relations which I have to derive for myself from time to time. So, not to have to do it each time again, I decided to devote some space to them here.

B.1 Basics

$$\langle x\rangle = \frac{1}{N}\sum x_i \qquad \sigma_x^2 = \frac{1}{N}\sum (x_i - \langle x\rangle)^2 \tag{B.1}$$

$$= \frac{1}{N}\left[\sum x_i^2 - 2\langle x\rangle \sum x_i + N\langle x\rangle^2\right] \tag{B.2}$$

$$= \langle x^2\rangle - \langle x\rangle^2 \tag{B.3}$$

The same result holds for the weighted case:

$$\langle x\rangle_w = \frac{\sum w_i x_i}{\sum w_i} \qquad \sigma_{x,w}^2 = \frac{\sum w_i (x_i - \langle x\rangle_w)^2}{\sum w_i} \tag{B.4}$$

$$= \frac{\sum w_i x_i^2 - 2\langle x\rangle_w \sum w_i x_i + \langle x\rangle_w^2 \sum w_i}{\sum w_i} \tag{B.5}$$

$$= \langle x^2\rangle_w - \langle x\rangle_w^2 \tag{B.6}$$

B.2 Error of an Asymmetry

Given an asymmetry $A = \frac{u-d}{u+d}$ extracted from $u+d = N$ events, the asymmetry according to Gaussian statistics is

$$\delta^2 A = \left(\frac{\partial A}{\partial u}\right)^2 \delta^2 u + \left(\frac{\partial A}{\partial d}\right)^2 \delta^2 d \tag{B.7}$$

$$= \frac{4d^2 u + 4u^2 d}{(u+d)^4} \tag{B.8}$$

$$= \frac{4ud}{N^3} \tag{B.9}$$

The maximum uncertainty for a given statistics is obtained for $u = d = \frac{1}{2}N$ and amounts to

$$\delta A = \frac{1}{\sqrt{N}} \tag{B.10}$$

B.3 Optimum Weight

Given a set of values x_i with corresponding uncertainties δx_i, which are to be averaged, what is the optimum weight? First, the uncertainty on the average can be computed using Gaussian uncertainty propagation:

$$\delta^2 \langle x \rangle_w = \sum_j \left(\frac{\partial \langle x \rangle_w}{\partial x_j}\right)^2 \delta^2 x_j = \sum_j \left(\frac{w_j}{\sum_k w_k}\right)^2 \delta^2 x_j \tag{B.11}$$

Since the uncertainty is positive, minimizing it yields the same result as minimizing the variance.

$$\frac{\partial \delta^2 \langle x \rangle_w}{\partial w_i} = 2 \sum_j \delta^2 x_j \frac{w_j}{\sum_k w_k} \cdot \frac{\delta_{ij} \sum_k w_k - w_j}{\left(\sum_k w_k\right)^2} \tag{B.12}$$

$$= \frac{2}{\left(\sum_k w_k\right)^3} \left[w_i \delta^2 x_i \sum_k w_k - \sum_j w_j^2 \delta^2 x_j \right] \stackrel{!}{=} 0 \tag{B.13}$$

This must be true for all δx_i, which implies that $w_i \delta^2 x_i$ is constant. As obviously the constant drops out in eq. (B.11), it is set to 1 and the result is

$$w_i = \frac{1}{\delta^2 x_i} \tag{B.14}$$

B.4 Allowed Fluctuation when Adding Data

When comparing data points which result from correlated samples it is useful to have a simple criterion at hand whether the variation between the results can be attributed to

decrease in uncertainty	increase in statistics	allowed movement
1%	2%	14%
2%	4%	20%
3%	6%	24%
4%	8%	27%
5%	10%	30%
6%	12%	33%
7%	14%	35%
8%	17%	38%
9%	19%	40%
10%	21%	42%

Table B.1: Allowed movement of data points at a confidence level of 68% in dependence of the decrease in statistical uncertainty or the growth in statistics. The allowed movement is given as a fraction of the original uncertainty.

statistical fluctuations or not. The most common case is that sample z_i is a superset of sample x_i, where the missing elements are denoted y_i and each of the individual measurements has its own Gaussian uncertainty. Let $w_x = \frac{1}{\delta x}$ be the weight associated to the expectation value $\langle x \rangle$, the latter of which will be called x in the following for brevity, and likewise for y and z. Using the relations derived above yields the following relations:

$$z w_z = x w_x + y w_y \tag{B.15}$$

$$w_z = w_x + w_y \tag{B.16}$$

The criterion for statistical compatibility between the two uncorrelated sub-samples x_i and y_i is that the difference between the expectation values x and y is smaller than a certain multiple of the standard deviation for that difference:

$$|x - y| < CL \cdot \sqrt{\frac{1}{w_x} + \frac{1}{w_y}} \tag{B.17}$$

The condition needs to be expressed in x and z and their uncertainties, which simplifies significantly by introducing the relative growth in statistics between x_i and z_i

$$w_z = w_x(1 + \varepsilon) \qquad w_y = \varepsilon w_x \tag{B.18}$$

Using this in eq. (B.17) yields

$$|x - z| \sqrt{1 + \frac{1}{\varepsilon}} < CL \cdot \delta x \tag{B.19}$$

Usually in a plot you see the change in uncertainty and the movement of the data point relative to the previous position. If the uncertainty gets smaller by 5%, then the relative

degrees of freedom	1σ		2σ		3σ		4σ	
	χ^2_{min}	χ^2_{max}	χ^2_{min}	χ^2_{max}	χ^2_{min}	χ^2_{max}	χ^2_{min}	χ^2_{max}
1	0.03938	2.00000	0.00326	4.00000	0.00032	6.00000	0.00003	8.00000
2	0.34229	3.69921	0.09314	6.18007	0.02882	8.49417	0.00938	10.7299
3	0.82824	5.20630	0.32893	8.02488	0.14672	10.5681	0.06856	12.9811
4	1.40859	6.62142	0.67397	9.71563	0.35900	12.4491	0.19998	15.0113
5	2.04631	7.98076	1.09552	11.3139	0.64812	14.2140	0.40027	16.9079
6	2.72319	9.30187	1.57299	12.8488	0.99784	15.8991	0.65958	18.7125
7	3.42887	10.5947	2.09324	14.3371	1.39598	17.5252	0.96840	20.4488
8	4.15686	11.8657	2.64746	15.7891	1.83364	19.1054	1.31892	22.1317
9	4.90276	13.1189	3.22943	17.2118	2.30420	20.6485	1.70489	23.7714
10	5.66346	14.3575	3.83461	18.6103	2.80266	22.1608	2.12139	25.3753
100	85.8335	114.182	77.3890	125.084	71.6064	133.439	67.0795	138.730
1000	955.050	1044.97	925.669	1076.81	904.653	1100.40	887.615	1120.05
10000	9862.50	10142.2	9763.57	10240.3	9693.56	10312.1	9636.60	10371.4
100000	99872.8	100450.	99355.6	100757.	99089.5	100981.	98887.1	101166.

Table B.2: χ^2 intervals corresponding to the central interval of the χ^2 probability distribution for a given confidence level.

growth in statistics is $\varepsilon \approx 10\%$, leading to $\sqrt{1+\frac{1}{\varepsilon}} = 3.3$, so that the point may move up to 30% of the original uncertainty for a confidence level of 68% ($CL = 1$). This is summarized for different settings in table B.1. In the case $\varepsilon \ll 1$ the formula simplifies to $\sqrt{\varepsilon}$ for the relative allowed movement at the 1σ confidence level.

B.5 Testing the Significance of Systematic Effects

When data are naturally grouped in units which should yield a bias-free measurement of a common quantity, the question arises whether the statistical errors alone account for the observed variations or if uncorrelated systematics add to them. If the measured quantity is not known a priori, as it would be e. g. in the case of deliberately canceling the physics asymmetry using inverted microwave settings, the weighted average of all data can be used as a best estimate. Like when fitting a constant value to the data points, a χ^2 can be calculated. If the statistical errors correctly model the fluctuations then the χ^2 probability $\mathcal{P}(\chi^2, n)$—with n being the number of degrees of freedom—should at 1σ confidence lie within the interval [15.7%, 84.3%]. This can be exploited to obtain statistical limits on the magnitude of uncorrelated systematic effects.

Let δ_i be the statistical errors of the measurements x_i and ρ be the assumed systematic one, then

$$\chi^2(\rho) = \sum_i \frac{(x_i - \langle x \rangle)^2}{\delta_i^2 + \rho^2} \tag{B.20}$$

is the modified χ^2. Using e. g. a binary search ρ can then be tuned so that $\mathcal{P}$ equals the

borders of the chosen confidence interval[1], leading to an equivalent confidence interval for the size of the uncorrelated systematic error. Table B.2 shows the χ^2 intervals resulting for certain numbers of degrees of freedom and confidence levels.

To estimate the influence of the uncorrelated systematic error on the weighted average, let $w_i = 1/\delta^2 x_i$ be the weights of the individual data points. The Gaussian uncertainty on the average caused by the systematic fluctuation ρ is then

$$\delta_{\text{syst}}\langle x\rangle = \rho\frac{\sqrt{\sum_i w_i^2}}{\sum_i w_i} \tag{B.21}$$

[1]Unless $\chi^2(0)$ already lies within the requested confidence interval, which would mean that the lower bound on ρ is zero.

Appendix C

Software

C.1 config_server Internals

Like config_server, LOAD and dimclient use the DIM name service to find the modules which are to be talked to. As these commands are frequently executed on control room computers and other software on these computers also employs DIM, but using a different name server, it is not the environment variable DIM_DNS_NODE which is considered but DIM_DNS_NODE_CS. The shell startup scripts of the common online account set things up correctly, but if a user wants to use these commands from his own account he will have to bear this in mind.

The heart of the COMPASS configuration server system is the config_server daemon, which runs on the VME-CPUs. First, the main program checks its location given by the *DB_CLIENT* and *DB_SERVER* environment variables, which define the VME_host_name to be looked up and the MYSQL server host to connect to. The DIM library also needs the environment variable *DIM_DNS_NODE* set. As outlined in section 3.1.2 DIM provides a special name service at which the config_server needs to register the Restart and Reexec commands. After doing so and creating the two VME mappings for 16-bit and 32-bit data access, the database is queried to find all modules attached to this VME-CPU and the corresponding module objects are instantiated. There are two ways to find out the correct module type: if the Type column is NULL, the source ID range of 640–767 is associated with normal GeSiCA modules, the rest with CATCH. Otherwise the Type information from the database determines which class to instantiate where all types but CATCH are handled by the GeSiCA module since they share the basic programming sequence.

Each module class is derived from the general Module, which defines general module behavior: it creates the seven standard DIM services listed in section 3.1.2 plus the three standard commands, it provides the database entry in a special class, which is refreshed from the MYSQL server on each call to Initialize and it defines convenient error reporting classes which can be thrown via the exception mechanism to be saved in the *ErrorString* DIM service. It also keeps a textual representation of the last read database contents in the

Configuration DIM service. The DIM services are registered in a service map which allows easy creation and update of services by simple C++ assignment.

The DIM commands registered by the Module class start a separate thread for each module being initialized so that different devices can be programmed in parallel. In case of the ADCs connected to GeSiCA modules this parallel execution saves a lot of time since every byte of the FPGA program is transferred via an I^2C transfer which takes thousands of CPU cycles to complete; these cycles would be wasted in a single-threaded environment. It is ensured that the same module initialization code does not run twice at the same time by utilizing a mutex; whether the command arriving later is queued or refused can be configured.

The modules only have to provide three methods: Initialize, ReadStatus and Reset, which are called when the corresponding DIM command is received. For LOAD to work correctly, all three methods must correctly set the final *ModuleStatus* DIM service to reflect the failure or success of the operation. These status codes are defined in cs_types.h, where the final status can be *MS_READY*, *MS_ERROR*, *MS_WARNING* or *MS_ERROR_WARNING* with the obvious meanings. The intermediate status codes *MS_INITIALIZE*, *MS_RESET*, *MS_READSTATUS*, *MS_DBXS*, *MS_ERROR_CONT* and *MS_WARNING_CONT* are purely informational. LOAD uses these to display a progress indication in case a single GeSiCA-type module is being programmed. This is possible since each write to a DIM service causes a value update—whether the value actually changed is immaterial—activating a call-back which had been registered by LOAD for this purpose. But beware, these notifications are not guaranteed to be delivered! If two updates happen too close in time the first one may be discarded.

In the following the GeSiCA module is described in more detail. The CATCH module is written and maintained by Fritz-Herbert Heinsius.

C.1.1 The GeSiCA Module

The GeSiCA module translates the Reset and ReadStatus commands to Initialize internally, since most of the executed code is the same. All hardware accesses are programmed in a write and a read form—for verification where it makes sense and is possible for the hardware—and it depends on the *init_mode* bit mask which are executed. If the command was ReadStatus, only verification is done, while in case of Initialize writing and subsequent verification is done. This uncovered many hardware problems while debugging the programming procedures. All three commands take an integer parameter which describes the subset of possible programming steps to be taken. Since it is usually only necessary to reset and program the APV chips when facing data errors, it would be inefficient to always reset the whole hardware chain. Therefore the Reset command only load the pedestals and thresholds into the ADCs and resets and programs the APVs. For the daring hardware developer also a fast mode is available which disables the verification

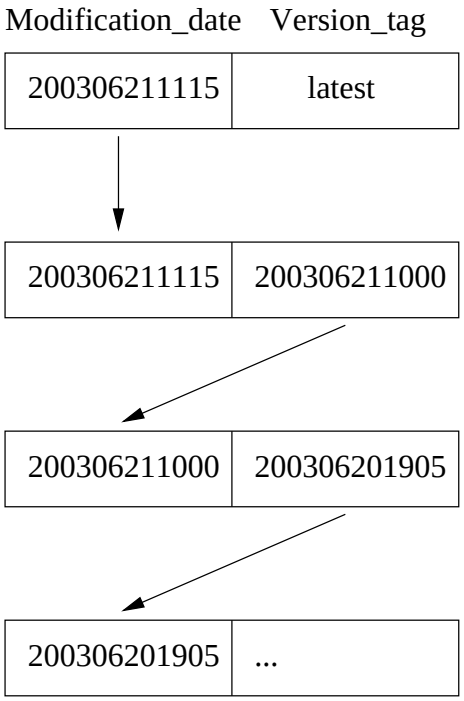


Figure C.1: Representation of the modification history of a table in the front-end database. The arrows show how to go backwards in history.

step. Further information on the available programming steps is available in the output of LOAD -h.

Each GeSiCA module keeps a *hardware_mutex* to guard against concurrent execution of the Reset, ReadStatus and Initialize commands. This is necessary since the generic code only protects each command from itself.

C.1.2 Database Layout

As detailed earlier one of the main motivations to introduce a front-end configuration database was to make modifications traceable. Thus, the database must be able to store the history of changes done to an equipment. For this purpose each table—with the exception of the web page related tables METATABLE, STARTTABLE and VERSION_TAGS—contains the two columns:

Version_tag stores the symbolic version tag if the entry is the current one in any of the defined versions. The default version selected by the config_server is latest, but there are also calibration, testing, inactive and disabled, which can be selected by giving it as the argument to a dimclient cmd ⟨pccofeXX⟩/Restart command. The references to Readout Buffer slots, module programs, TCS programs and front-end programs can be annotated with a different version by appending a colon and the desired version tag. This is used so that not

everything has to be duplicated when e. g. in fact only some ports shall be disabled in the testing configuration.

Modification_date is a column of type TIMESTAMP which records all changes to the column. In principle there should be no changes apart from retiring it when creating a new version. In this case, which is transparently handled for each time a user makes a change via the fedb.pl web interface, the old modification date is copied to the Version_tag column and Modification_date is set to the current time, which is also the modification date of the new entry.

Figure C.1 illustrates how the algorithm stores the modification history of all items in the database. Given below are the descriptions of the tables used in the database devdb in alphabetical order. As the CATCH module is written and maintained by Fritz-Herbert Heinsius alone, CATCH specific fields are not described here.

FRONTEND This is the biggest table in the database, where all front-end cards are registered. The meaning of some of its entries depends on the detector type to which they are associated. In case of the GeSiCA readout chain the rows of this table correspond to single APV chips or each of the two sides of an ADC card. In the following all entries which have a different meaning for CATCH and GeSiCA are marked with an asterisk:

Detector Descriptive name of the detector module the front-end card is connected to.

***Geographic_ID** Numeric identification of the front-end on the concentrator module. In case of GeSiCA this is the I^2C address of the APV chip or 0/1 for the lower/upper side of the ADC card.

TBname The canonical name of the detector plane to which this front-end is connected.

Format_ID Numeric distinction between different hardware versions of the front-end.

State Overall activation status of the front-end. It will not be accessed unless this flag is non-zero.

***Calibration_Info** In case of GeSiCA ADC entries this holds the name of the pedestal file which is to be loaded into the zero-suppression chip.

***Configuration** In case of GeSiCA a string describing either the APV or ADC register settings to be applied or tested.

FRONTEND_program The name of the corresponding entry of table FRONTEND_PROGRAM, which describes the software to be downloaded to the front-end card. This is NULL in case of the APV entries.

MODULE_source_ID This links the front-end card to the concentrator module. Its partner column lives in table MODULE.

Port The port on the concentrator module to which this front-end is connected.

***FRONTEND_serial_number** *unused in GeSiCA case*

***CableName** *unused in GeSiCA case*

***Number_of_TDC** *unused in GeSiCA case*

***Frontend_disabled** *unused in GeSiCA case*

***Gate_Window** *unused in GeSiCA case*

***Gate_latency** *unused in GeSiCA case*

Version_tag *see beginning of section*

Modification_date *see beginning of section*

FRONTEND_PROGRAM Most front-end cards have a reconfigurable logic in form of an FPGA chip. These chips lose their programming if the voltage drops—even for a short time—below a certain threshold, so they need to be reprogrammed frequently. These programs are available on the network file system in the experiment hall and this table is there to describe and locate them.

FRONTEND_program Name of the program as referred to from the FRONTEND table.

Description Short mention of features of this specific program file version.

FRONTEND_setup_file File name on the network.

Version_tag *see beginning of section*

Modification_date *see beginning of section*

MACROS All text fields in the other tables feature macro expansion of the form $macro_name. Macros may expand to contain other macros, but expansion is stopped at a nesting level of 10 to avoid endless loops.

Name Name by which the macros is to be invoked.

Expansion Expansion text, which may contain further macros.

Version_tag *see beginning of section*

Modification_date *see beginning of section*

METATABLE This table is only used for the fedb.pl web front-end. It contains a description of the editable items, possible values and sorting criteria.

- **tablename** Name of the table this entry applies to.
- **colname** Name of the column in that table.
- **type** Data type, which can be c for strings, i for decimals and x for hexadecimals.
- **dropdown** If non-empty this is either a comma-separated list (starting with list), the word print which means read-only or a SQL statement yielding a list of values to choose from.
- **prio** The display of columns is ordered by this value.
- **orderby** The displayed rows are ordered by the columns which have a non-NULL value here. Higher numbers mean lower sorting priority.

MODULE Being the main entry point for the config_server, this table represents parameters of the concentrator modules as well as their connections to front-end cards, VME crates and Readout Buffers.

- **MODULE_source_ID** The unique characteristic of the module.
- **State** Global activation state of this module. If set to zero the module will not be programmed.
- **VME_host_name** DNS name of the VME-CPU, which also uniquely identifies the VME crate housing the module.
- **VME_base_address** Start address of the module's configuration space in on the VME bus.
- **VME_slot** Mechanical slot number inside the crate.
- **Detector** Name(s) of the detector(s) which are read out using this module.
- **MODULE_program** Reference to the programming entry in table MODULE_PROGRAM.
- **TCS_program** Reference to the programming entry for the connected TCS receiver card in table TCS_PROGRAM.
- **ROB_slot** The name of the Readout Buffercomputer and the S-LINK card number in the range 0–3.

Type Module type. Initially there were only CATCH and GESICA, which could be distinguished by their base address ranges, but then came the TCS controller and prescaler, which are also programmed using the config_server, and HOTGESICA modules, which have integrated TCS receivers.

MODULE_serial_number *unused in GESICA case*

Configuration *unused in GESICA case*

CableName *unused in GESICA case*

Slink_format_ID *unused in GESICA case*

Slink_mode *unused in GESICA case*

Trigger_mode *unused in GESICA case*

Version_tag *see beginning of section*

Modification_date *see beginning of section*

MODULE_TYPES This table only lists the different concentrator module types.

Type The number as it should be entered into the MODULE table.

Description Name of the module type.

SPILLBUFFER The data concentrator modules are connected to Readout Buffer slots, which are listed in this table.

ROB_slot The name of the port as it is entered into the MODULE table. This is the DNS name of the computer with the card number appended like pccorb12_2.

CableName *unused in GESICA case*

Type S-LINK comes in three variants which can transport 100 MB s^{-1}, 128 MB s^{-1} and 160 MB s^{-1} respectively.

Version_tag *see beginning of section*

Modification_date *see beginning of section*

STARTTABLE This table is used only by the fedb.pl web front-end to describe the possible item selections the user can make.

tablename Name of the table to select entries from.

dropdown SQL subexpression which is select'ed from the table to generate the list of items.

description	The descriptive type of items that are generated by this rule. This is printed next to the generated drop down menu on the web page.
query	The full SQL query to be used when generating the list of items. The string XXX will be substituted by the selected Version_tag.

TCS_PROGRAM The TCS receiver module which is plugged to the back of the data concentrator modules also needs to be programmed as it is also based on an FPGA. This table is very much comparable with the FRONTEND_PROGRAM table.

TCS_program	Name of the program as referred to from the MODULE table.
Description	Short mention of features of this specific program file version.
TCS_setup_file	File name on the network.
Version_tag	*see beginning of section*
Modification_date	*see beginning of section*

VERSION_TAGS This single-column table lists the known version tags and is used to create a drop down menu on the web page for version tag selection.

C.2 Cinderella Internals

C.2.1 Tools

Several small utilities have emerged along the way of creating CINDERELLA, the most important of which are listed here:

C.2.1.1 calibrator

The auto-calibration system outlined in section 5.4 is based on the central calibrator daemon to gather the statistics of one run and perform the fits necessary for auto-calibration. It runs on an online computer which usually is not working to full capacity so that the fitting of the several hundred histograms does not take too long and also does not disturb regular data taking.

C.2.1.2 cat_date

This is the Swiss Army Knife for ALICE DATE streams. It is linked with libshift so that it can read files transparently from the Event Builders or from CASTOR and it can write the output to a file—possibly also a libshift path—or to a network port, allowing the efficient long distance transfer of data. The most useful part, however, is that it can filter the data by trigger mask or spill number and that it can even accept only the data from specific detectors. The latter can be used to create excerpts for testing of CINDERELLA—leaving out all source IDs which are not decoded anyway—or for detector studies where many events are needed and the information from other detectors costs only transfer and processing time.

C.2.1.3 eob_monitor

The ALICE DATE libraries provide online monitoring facilities, to which clients can connect. Normally the Event Builder monitoring buffers are filled by the eventBuilder process, but in case CINDERELLA is running it takes over to allow the monitoring of the filtered data stream, especially the filter summary information attached to each end-of-burst event. Therefore the eob_monitor connects to all running CINDERELLA instances and uses the filtering capabilities of the monitoring library to receive all end-of-burst records. After adding up the statistics the obtained accept ratios are written to a MYSQL database which is in turn displayed on the shift crew's monitoring display. This way the spill-by-spill stability of data taking conditions for the individual triggers can be directly observed by the shift crew.

C.2.1.4 hist_draw

The histograms created for calibration are also saved to disk for later inspection should something have gone wrong. Since CINDERELLA is written in C it cannot directly link to ROOT to create these histograms, but that would also be exaggerated since the core functionality—counting in bins—can easily be implemented in CINDERELLA. hist_draw reads the CINDERELLA specific histogram files and displays them or saves them in the ROOT format for further processing.

C.2.2 Configuration System

This section is meant as a starting point for someone wanting to create a new CINDERELLA filter module. The existing prescaler module is used as an example as it is the most simple module available. First, let's look at the configuration, given in file prescaler_config.xml:

```
<Prescaler>
```

```
<Active type="CT_UINT" unit="boolean">
<default>1</default>
 The main activation flag for this filter. If it is set to zero,
 the filter will not be executed.
</Active>

<RunTriggers type="CT_ARRAY_UINT" unit="trigger number"
 location="yes_triggers">
<default>0</default>
 Triggers for which the filter should be executed.
</RunTriggers>

<DontRunTriggers type="CT_ARRAY_UINT" unit="trigger number"
 location="no_triggers">
 Triggers for which the filter may not be executed.
</DontRunTriggers>

<HotTriggers type="CT_ARRAY_UINT" unit="trigger number"
 location="hot_triggers">
<default>0</default>
 Triggers for which the filter module is run "hot" and filter decision are
 really enforced. For triggers that are not specified as "hot", the filter
 module is run in tagging-only mode.
</HotTriggers>

<Factor type="CT_UINT">
<default>10</default>

 Reduction factor of the prescaler (Integer). Precisely ever n'th event is
 accepted.
</Factor>

</Prescaler>
```

The CINDERELLA build system scans the source tree for all files ending in _config.xml, generating the various output files from them. First of all, the declaration of the struct prescaler_conf_s is generated, which is used by the module code to access the configuration items the user has given:

```
#ifndef _Prescaler_CONFIG_H_
#define _Prescaler_CONFIG_H_
struct prescaler_conf_s {
  /**
   * The main activation flag for this filter. If it is set to zero,
   * the filter will not be executed.
```

```
   **/
  unsigned int active;
  /**
   *  Triggers for which the filter should be executed.
   **/
  unsigned int ** yes_triggers;
  /**
   *  Triggers for which the filter may not be executed.
   **/
  unsigned int ** no_triggers;
  /**
   *  Triggers for which the filter module is run "hot" and filter decision are
   *  really enforced. For triggers that are not specified as "hot", the filter
   *  module is run in tagging-only mode.
   **/
  unsigned int ** hot_triggers;
  /**
   *  Reduction factor of the prescaler (Integer). Precisely ever n'th event is
   *  accepted.
   **/
  unsigned int factor;
};
typedef struct prescaler_conf_s prescaler_conf_t;
#endif
```

You can see how configuration items marked CT_UINT are translated to unsigned int and CT_ARRAY_UINT to and array of pointers to unsigned integer. It is necessary to use the double indirect representation to be able to signal the end of the array while still allowing all possible values for unsigned int to be given in the configuration file. The second generated file is the default configuration for the prescaler module. The values are taken from the <default> tags in the prescaler_config.xml file:

```
  <Prescaler instance="default">

    <!-- type: CT_UINT   unit: boolean -->
    <Active>1</Active>

    <!-- type: CT_ARRAY_UINT   unit: trigger number -->
    <RunTriggers>
      <Value>0</Value>
    </RunTriggers>

    <!-- type: CT_ARRAY_UINT   unit: trigger number -->
    <DontRunTriggers>
    </DontRunTriggers>
```

```
  <!-- type: CT_ARRAY_UINT   unit: trigger number -->
  <HotTriggers>
    <Value>0</Value>
  </HotTriggers>

  <!-- type: CT_UINT   unit: - -->
  <Factor>10</Factor>

</Prescaler>
```

Multiple <default> tags can be given to create multiple entries in array items. The third output is the TEX source which shows the configuration tree:

```
\begin{tree}{Prescaler}{-}{-}
Sorry, no description yet.
  \begin{tree}{Active}{CT_UINT}{boolean}
    The main activation flag for this filter. If it is set to zero,
    the filter will not be executed.
  \end{tree}
  \begin{tree}{RunTriggers}{CT_ARRAY_UINT}{trigger number}
    Triggers for which the filter should be executed.
  \end{tree}
  \begin{tree}{DontRunTriggers}{CT_ARRAY_UINT}{trigger number}
    Triggers for which the filter may not be executed.
  \end{tree}
  \begin{tree}{HotTriggers}{CT_ARRAY_UINT}{trigger number}
    Triggers for which the filter module is run "hot" and filter decision are
    really enforced. For triggers that are not specified as "hot", the filter
    module is run in tagging-only mode.
  \end{tree}
  \begin{tree}{Factor}{CT_UINT}{-}
    Reduction factor of the prescaler (Integer). Precisely ever n'th event is
    accepted.
  \end{tree}
\end{tree}
```

In addition to these user-visible output files, also internal glue is needed so that the configuration subsystem knows how to fill the values obtained from the XML files into their corresponding C structures. src/config/config_ylate.inc provides the offsets of the data destination relative to the beginning of the configuration structure in question:

```
/************************************************
 *  The main activation flag for this filter. If it is set to zero,
 *  the filter will not be executed.
```

```
 ************************************************/
{"CompassConfig.FilterConfig.Prescaler.Active", CT_UINT,
  (char*)&(((prescaler_conf_t*)ptr)->active)-(char*)ptr},

/************************************************
 *  Triggers for which the filter should be executed.
 ************************************************/
{"CompassConfig.FilterConfig.Prescaler.RunTriggers", CT_ARRAY_UINT,
  (char*)&(((prescaler_conf_t*)ptr)->yes_triggers)-(char*)ptr},

/************************************************
 *  Triggers for which the filter may not be executed.
 ************************************************/
{"CompassConfig.FilterConfig.Prescaler.DontRunTriggers", CT_ARRAY_UINT,
  (char*)&(((prescaler_conf_t*)ptr)->no_triggers)-(char*)ptr},

/************************************************
 *  Triggers for which the filter module is run "hot" and filter decision are
 *  really enforced. For triggers that are not specified as "hot", the filter
 *  module is run in tagging-only mode.
 ************************************************/
{"CompassConfig.FilterConfig.Prescaler.HotTriggers", CT_ARRAY_UINT,
  (char*)&(((prescaler_conf_t*)ptr)->hot_triggers)-(char*)ptr},

/************************************************
 *  Reduction factor of the prescaler (Integer). Precisely ever n'th event is
 *  accepted.
 ************************************************/
{"CompassConfig.FilterConfig.Prescaler.Factor", CT_UINT,
  (char*)&(((prescaler_conf_t*)ptr)->factor)-(char*)ptr},
```

For the module name to be associated with the size of the corresponding configuration structure, a list is currently maintained by hand near the beginning of src/config/configuration.c, into which a new module would have to be entered before it can be configured.

C.3 Histo

The origin of the ROOT package lies in histogramming large data sets, which implies that this functionality is made available with very low programming overhead for the user. There are two shortcomings in the approach:

- The use of C++ as programming language requires unique names and a declaration for every histogram.

- If a great number of histograms need to be customized (think axis titles, colors, and so on), this has to be explicitly stated for each single one.

While being part of the initial problem, C++ also entails the means to solve it. A custom container class has been written, which manages histograms and creates them as they are used, based on user-defined templates which are matched to the histogram title. It also manages the hierarchical result tree mentioned in section 7.1.3, including the routines to create pretty plots with labeled graphs for the comparison of asymmetries (which are discussed in their own section C.3.3).

C.3.1 User Interface

The user simply creates a Histo object with the default constructor. This object is already preconfigured so that simple histograms can be filled right away. The repertoire encompasses the ROOT histogram types TH1D, TH2D and TH3D—including their siblings for other numerical data types—and TProfile. Additionally the same filling methods can be used to gradually construct graphs (TGraphAsymmErrors). In the following a * denotes one of these five possible target types, which are then abbreviated as 1, 2, 3, P and G respectively.

C.3.1.1 Filling

Filling is done with the Fill* methods, which take a different number of arguments for the different objects to be filled, see table C.1. Histograms of other numerical data types are filled with the same methods, always passing the values as double.

FillG is special in several aspects. Let us first look at the *err** arguments:

- if none of them is present, all errors are set to 0
- if only *err1* and *err2* are present, they are taken to be symmetric x and y errors respectively
- if all four arguments are given, they are $e_x low$, $e_x high$, $e_y low$ and $e_y high$—in that order

The argument list documented in table C.1 is used to add a point at a given position. However, often it is necessary to simply advance the x position by a constant amount and add a point there. This is accomplished by substituting the x argument by a string *label* argument. When drawing the graph this point will be labeled with the given string.

Fill1	string *name*, double *x*
Fill2	string *name*, double *x*, double *y*
Fill3	string *name*, double *x*, double *y*, double *z*
FillP	string *name*, double *x*, double *y*
FillG	string *name*, double $x^{\ddagger}$, *y* double $err1^{\ddagger}$, $err2^{\ddagger}$ double $err3^{\ddagger}$, $err4^{\ddagger}$

SetDefault1	string *pattern*, int *bins_x*, double *x_low*, double *x_high*, Create1_t†, Finish1_t†, Draw_t†, char* *draw_option*†, Types *type*†
SetDefault2	string *pattern*, int *bins_x*, double *x_low*, double *x_high*, int *bins_y*, double *y_low*, double *y_high*, Create2_t†, Finish2_t†, Draw_t†, char* *draw_option*†, Types *type*†
SetDefault3	string *pattern*, int *bins_x*, double *x_low*, double *x_high*, int *bins_y*, double *y_low*, double *y_high*, int *bins_z*, double *z_low*, double *z_high*, Create3_t†, Finish3_t†, Draw_t†, char* *draw_option*†, Types *type*†
SetDefaultP	string *pattern*, int *bins_x*, double *x_low*, double *x_high*, double *y_low*, double *y_high*, CreateP_t†, FinishP_t†, Draw_t†, char* *draw_option*†
SetDefaultG	string *pattern*, CreateG_t†, FinishG_t†, Draw_t† char* *draw_option*†

Table C.1: Arguments of the Fill* and SetDefault* methods. Optional arguments are marked with †. For the arguments marked with ‡ see the detailed explanation in the text.

C.3.1.2 Defaults

Default settings can be added for 1-dimensional, 2-dimensional, 3-dimensional and profile histograms or graphs via the SetDefault* methods. These default settings are stored in a list for each histogram type, which is traversed each time a histogram is created. As soon as a match is found between the requested histogram title and the globbing pattern[1] these settings are applied to the newly created histogram. It is important to note that histograms need not be created beforehand, since they are automatically created when they are first filled with the Fill* methods. If special actions need to be taken for certain histograms, the creation can also be explicitly requested by using the GetHist* methods.

To allow greatest possible freedom in the customization done by the user, the SetDefault* methods take lists of functions which are called with a pointer to the object which is to be manipulated. More specifically, the functions need to fit the declaration

void func(<type>* *obj*);

where func is the name chosen arbitrarily by the user and <type> is the histogram type for CreateX_t and FinishX_t or TVirtualPad for Draw_t. If only a single function pointer (this includes NULL for "none") is needed, then it can be given directly as an argument to

[1] The function fnmatch of the C library is used, which usually is employed by command shells to perform filename matching, see man fnmatch

the SetDefault* method, otherwise a list has to be explicitly created using the constructor which takes a variable argument list like this:

Create1_t(*func1*, *func2*, ..., NULL)

It is important to terminate the list with a NULL *pointer.* The CreateX_t list is executed directly after histogram creation, the FinishX_t list execution is initiated by the user with the method Histo::FinishAll and the Draw_t functions run after the histogram has been drawn on the canvas[2].

For common tasks Histo provides readily usable static methods:

SumW2 Calls the SumW2 method of the ROOT histogram, which enables more statistics book-keeping and has other side effects, see [41] for details.

BinLog*M* Resets the binning of the Histogram along the M axis, so that the number of bins stays the same, but the bin edges are equidistant when the axis plotted in log scale. This has been used e. g. to produce the plots showing the Q^2 distribution.

SetLog*x* Sets the pad coordinates to be logarithmic along axis x before drawing the histogram. This function goes in the Draw_t lists.

AxisFromTitle Sets the axis title(s) from the histogram name. Multiple dimensions are separated using the string vs (notice the spaces surrounding it). An N-dimensional histogram type takes only the first N parts of the title, beginning by the last axis, which is the z-axis for three-dimensional histograms.

To keep the SetDefault* calls short all arguments after the necessary bin counts and histogram edges are optional. The default values for the arguments are empty for the lists and *draw_option*, except for the *draw_option* of a graph, which is "ap".

C.3.1.3 ROOT Object Ownership

If you feel that you are not completely comfortable with the ROOT concept of object ownership, please read the corresponding chapter in the ROOT User's Guide [41].

By default, all histograms and graphs create by the Histo object are disowned, which means that they don't belong to a specific ROOT directory. This has the advantage that the user has direct control over when and how the objects are destroyed, in contrast to the situation where the lifetime of the ROOT directory—e. g. a file—determines the lifetime of the objects it contains. There are two methods, DisownHistograms and DontDisownHistograms, which change this behavior, but they don't work retrospectively, so if you need to change the owner of a histogram after its creation, you have to do it by hand (calling SetDirectory(0) for the object in question).

[2] The type of the argument is TVirtualPad* because the ROOT macro gPad is used internally

C.3.1.4 Placing Histograms in ROOT Directories

Every time a histogram or graph is created, the current ROOT directory *gDirectory* is changed temporarily to the internally saved directory. Upon creation of the Histo object the internal directory is initialized with the current *gDirectory*, but the user can change this with the following methods:

- SetFile accepts either a TFile* or the name and optionally an open mode for a TFile which is to be created by the Histo object.
- SetDirectory simply sets the internal directory to the supplied TDirectory*.

If a file was previously set using SetFile, this file is closed when executing one of the above methods. An important usage of the internally saved directory is that the method WriteAll writes all objects to this directory unless a different directory is specified as the first parameter. The destructor of the Histo class calls WriteAll automatically if a writable TFile was registered using the SetFile methods.

C.3.1.5 Printing

After filling the histogram they usually are displayed and/or saved in a graphical format for later use. This can of course—via the GetHist* methods—manually and selectively be done by the user. An alternative is to use the general methods of Histo which operate on all registered objects at once. Histo::DrawAll takes a bit mask which decides whether to create a postscript output file, display the histograms on screen or both. If the first parameter is a string it is interpreted as the name of a directory into which the histograms file is to be saved.

C.3.2 Tools

Over time expose to ROOT leads to the creation of certain scripts like for the beautification of plots and the handling of histograms and trees from different files. Two of these scripts have been polished and extended into standalone programs, namely my CreateEPS and Boris Grube's MergeHist. Both are now part of the Histo distribution.

C.3.2.1 CreateEPS

The purpose of this program is the extraction of plots from ROOT files and subsequent storage of the resulting TCanvas in a format of the user's liking—as long as ROOT knows how to write this format. While I have not personally tried all these, the ROOT web page claims that all of the following are supported: PS, EPS, PDF, SVG, TIFF, JPG, PNG, GIF, XPM, C++ ROOT macro, ROOT binary, ROOT XML. Given this broad choice of output

formats, the inputs may seem restrictive: Either a the ROOT histogram is taken directly from a binary ROOT file or it is generated using the TTree::Draw function from a ROOT tree out of a binary ROOT file. The strong point is that histograms may be combined into one plot, either by drawing them onto the same canvas, or by dividing them by other histograms first, or multiplying them by an arbitrary function. In addition arbitrary functions can be drawn onto the canvas.

Besides the control over the design of the plots, which is necessary to keep up a certain style throughout a document, also tasks like labeling the axes or generating a legend are rather easy with this tool compared to having to write the same style control and position calculations every time a new plot is needed. The general syntax of the command is

CreateEPS -f ⟨input⟩ -o ⟨output⟩ -h ⟨spec⟩ [-g ⟨spec⟩] [-l ⟨spec⟩] [-q] [-v]

All of the options may be given more than once, but for -g and -l only the last occurrence is used. While -v as usual increases the verbosity level, -q lowers it, but as soon as the latter is encountered the ROOT message output concerning errors or the creation of new files is irrevocably suppressed. The ⟨spec⟩ is a comma-separated list of tokens which either have an equals sign and are interpreted as key-value pairs or they don't, in which case the value is taken to be "true". The comma may be doubled to produce a literal comma as part of a token. The ⟨spec⟩s given for the -h and -l options are processed sequentially, producing one histogram per -h or the (optional) legend with -l. While processing a ⟨spec⟩ missing values are searched in the global -g ⟨spec⟩. Histograms and trees are searched by their complete ROOT path inside the files given with -f, the search stopping at the first match found.

For the exact tokens within the ⟨spec⟩s please consult the source code, especially the main loop near the end of the file, as the knowledge and needs concerning beautiful plots always grow and new features are added steadily.

C.3.2.2 MergeHist

This handy tool replaces the hadd coming with the ROOT distribution. It offers fine-grained control over which items are merged, it consumes considerably less memory when merging large numbers of files and it merges histograms as it finds them and does not restrict itself to the set found incidentally in the first named input file. If trees are to be merged, they are written each into its own file, as ROOT has a built-in mechanism which creates new files should the tree grow larger than 1.9 GB[3]. To allow sensible management of the merged tree they are kept separate. The command line syntax is

MergeHist [-f] [-h] [-T [⟨trees⟩]] ⟨output⟩ ⟨input⟩…

[3] This limit is configurable using the static method TTree::SetMaxTreeSize.

where -f enables the overwriting of the output file in case of its previous existence, -h prevents the merging of histograms and -T enables the merging of trees, either unconditionally or only the comma-separated list of full tree paths inside the respective input files.

C.3.3 Hierarchical Result Tree

As mentioned in section 7.1.3 the analyze program generates a hierarchical result tree from the database contents, which can then be explored using the visualization part of that program. Actually the work is done by the Histo class, which provides not only the infrastructure for the storage of hierarchically organized measurement results, but also has a powerful plot generation engine to allow intuitive views of the results to be created.

C.3.3.1 Filling the Tree

First the tree has to be filled, which is done using the Histo::FillR method. The first argument is the path to the node in the tree which is to be filled, the path separator being the colon. Paths are represented using the Histo::path_t class, which can be constructed from ordinary strings and concatenated. An example path used in the high p_T analysis is

```
Histo::path_t("Joerg2w") + "1 < p_T < 1.5" + "minus:02P2A" + "12"
```

The second and third arguments to FillR are the value and the Gaussian error of same. The fourth and last argument is the Histo::ValueExtra, additional information attached to this data point. In case of the high p_T analysis this might be $\langle D \rangle$, but in principle it can be any list of doubles plus an optional name. An example construction could be

```
Histo::ValueExtra("group12") + 5 + 7
```

which would save the values 5 and 7 together with the annotation group12.

It is important to note that when filling the tree the values get added to all nodes which are traversed until finding the one which was specified, thus adding up—and correctly averaging—the values higher up in the tree in the nodes closer to the root. By taking the most elementary measurements, e. g. the asymmetry from a single run group, as the leaves of the tree, period-wise or microwave-wise summary becomes very easy. Just make sure to order the path the way you would like to average your values; if you need more than one way, don't panic: simply save the values twice and create distinct nodes for the two subtrees below the root.

C.3.3.2 Using the Tree

After filling the tree, Histo::MakeResultCanvas is used to extract information. It takes a Histo::ResultList to specify the nodes which are to be used from the tree plus a Histo::Options object which carries the style parameter and other options. The result list is constructed from triples of Histo::path_t, TStyle* and string, specifying the path to nodes in the tree (each path element being a globbing pattern), the style to be used for drawing the resulting graph and the label for this set of points, which will be put into the legend, should one be requested via the options. For the rich set of possible options please refer to the source code.

A second possibility is to produce a pulls plot, in which case the result list is replaced by a single Histo::path_t which again acts as a globbing pattern for each path component to select the set of measurements. The options determine whether the pulls are to be calculated against the mean or a fixed value.

Appendix D

Figures

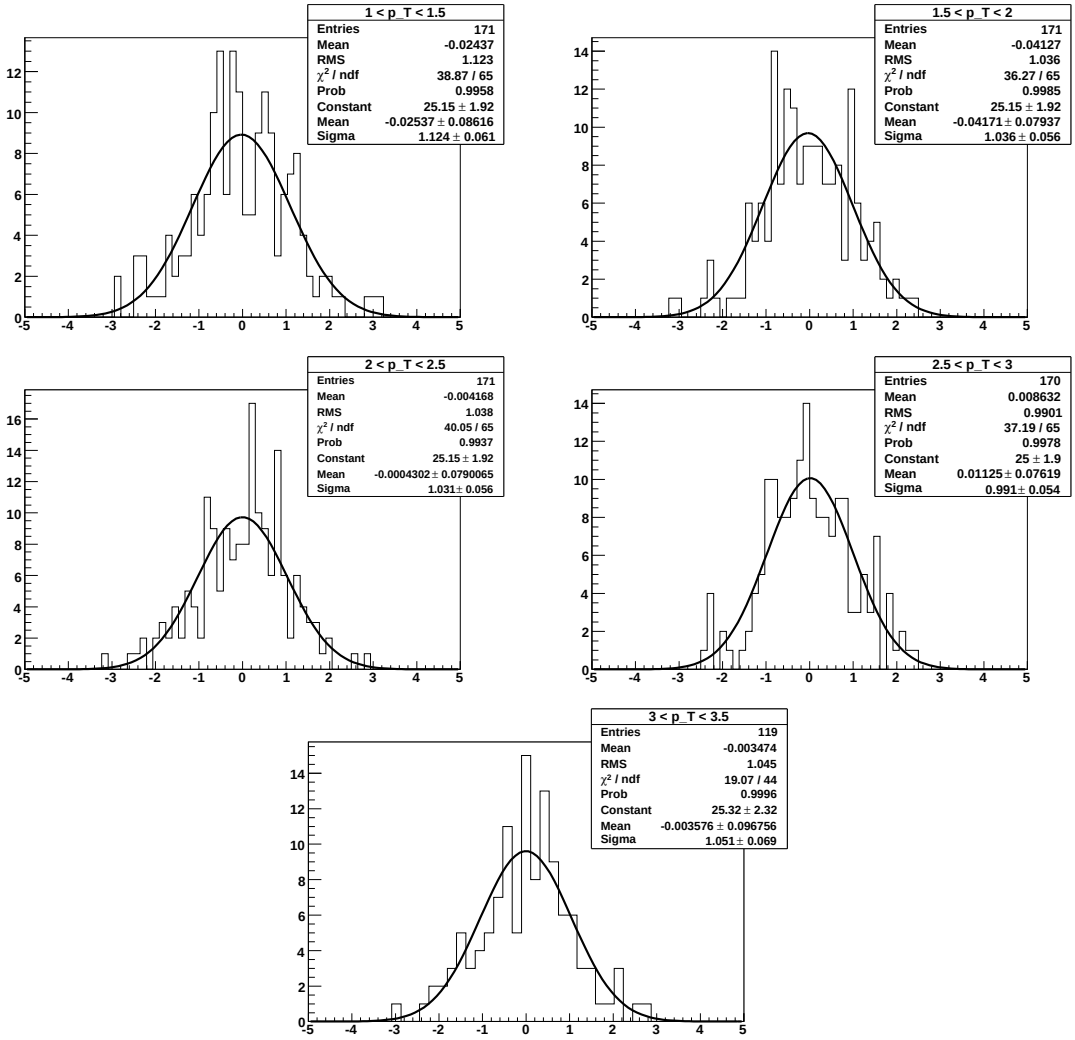


Figure D.1: Pulls for the asymmetry extracted using the 2nd order method. The corresponding data points are shown in figure 7.23 on page 105.

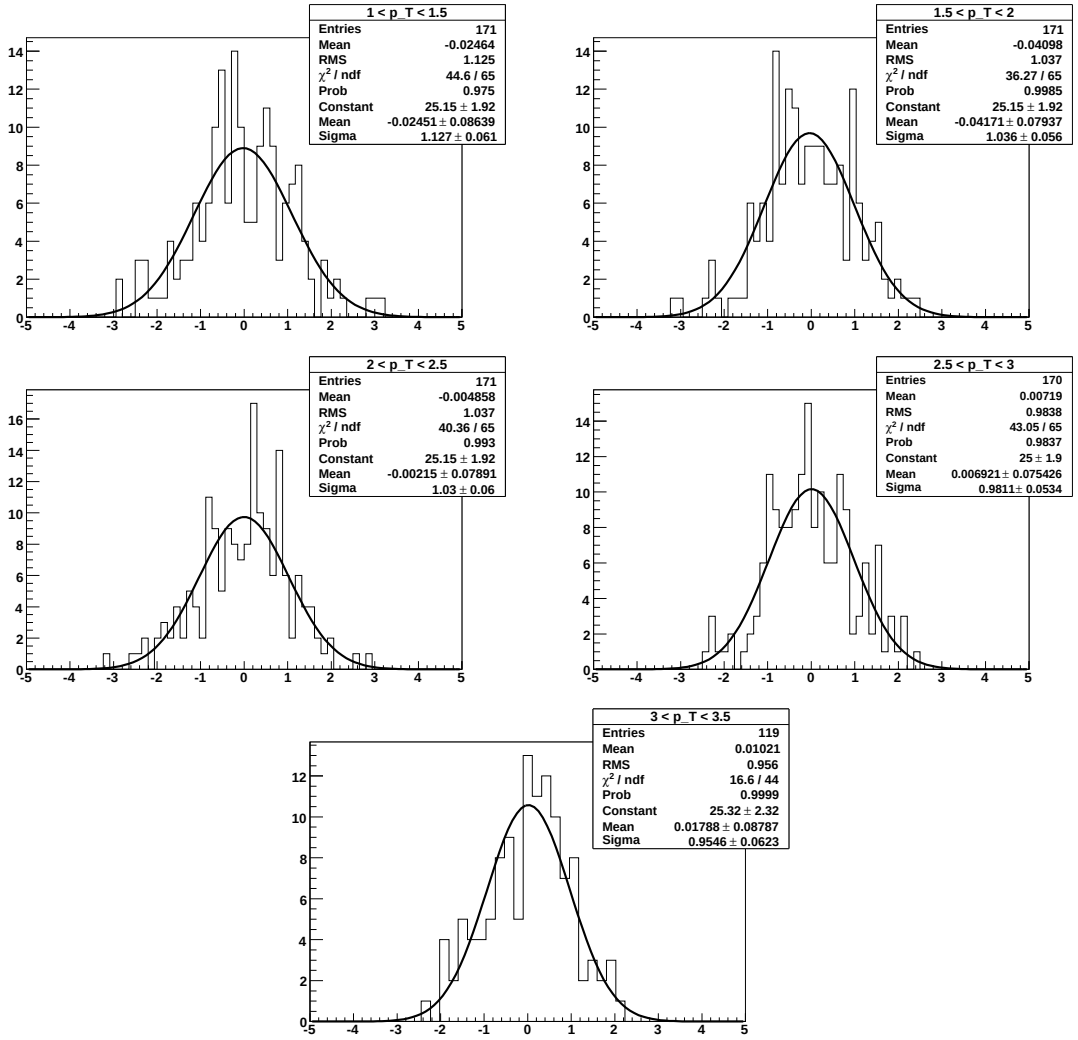


Figure D.2: Pulls for the asymmetry extracted using the 1st order method with global correction factor for the apparatus asymmetry. The corresponding data points are shown in figure 7.23 on page 105.

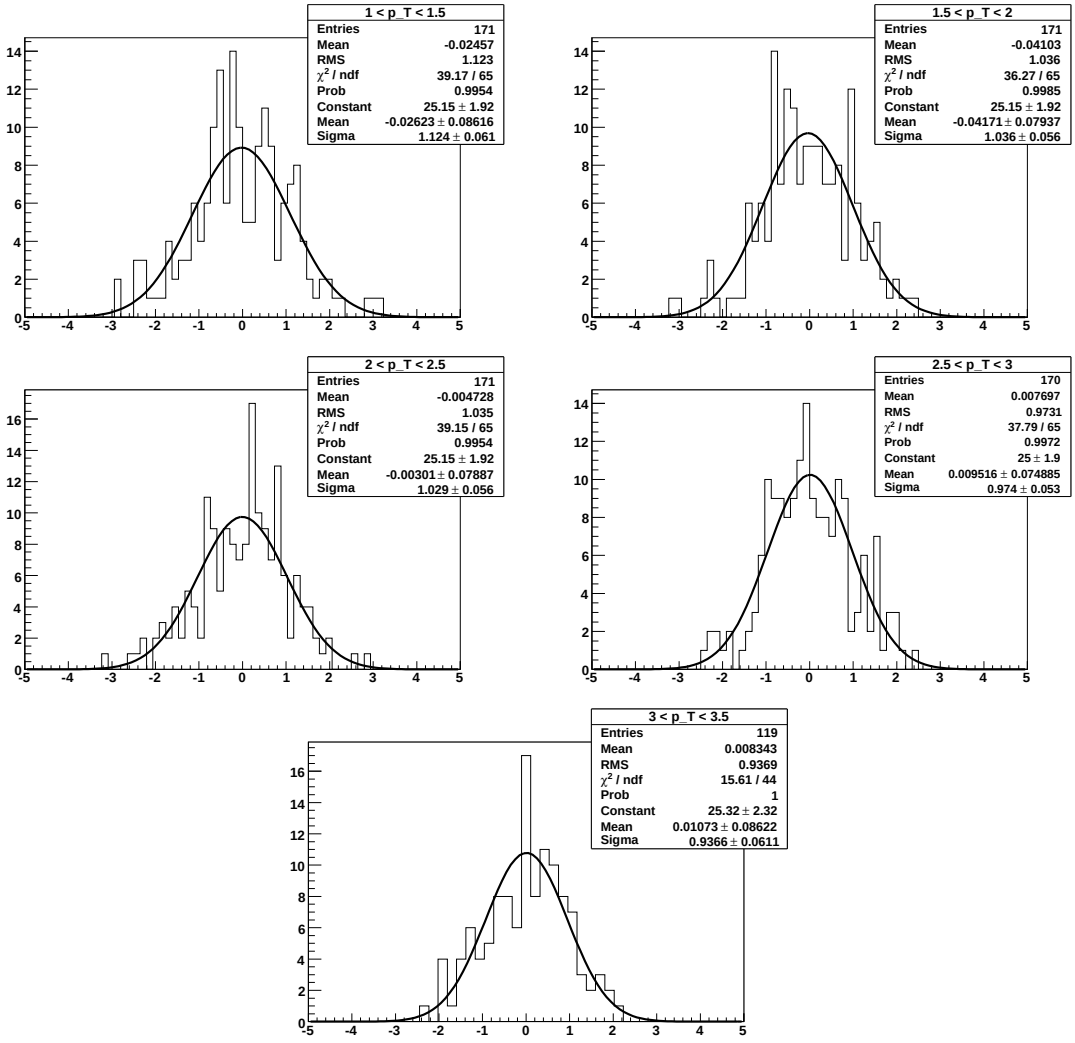


Figure D.3: Pulls for the asymmetry extracted using the 1st order method with individual sub-sample correction factors for the apparatus asymmetry. The corresponding data points are shown in figure 7.23 on page 105.

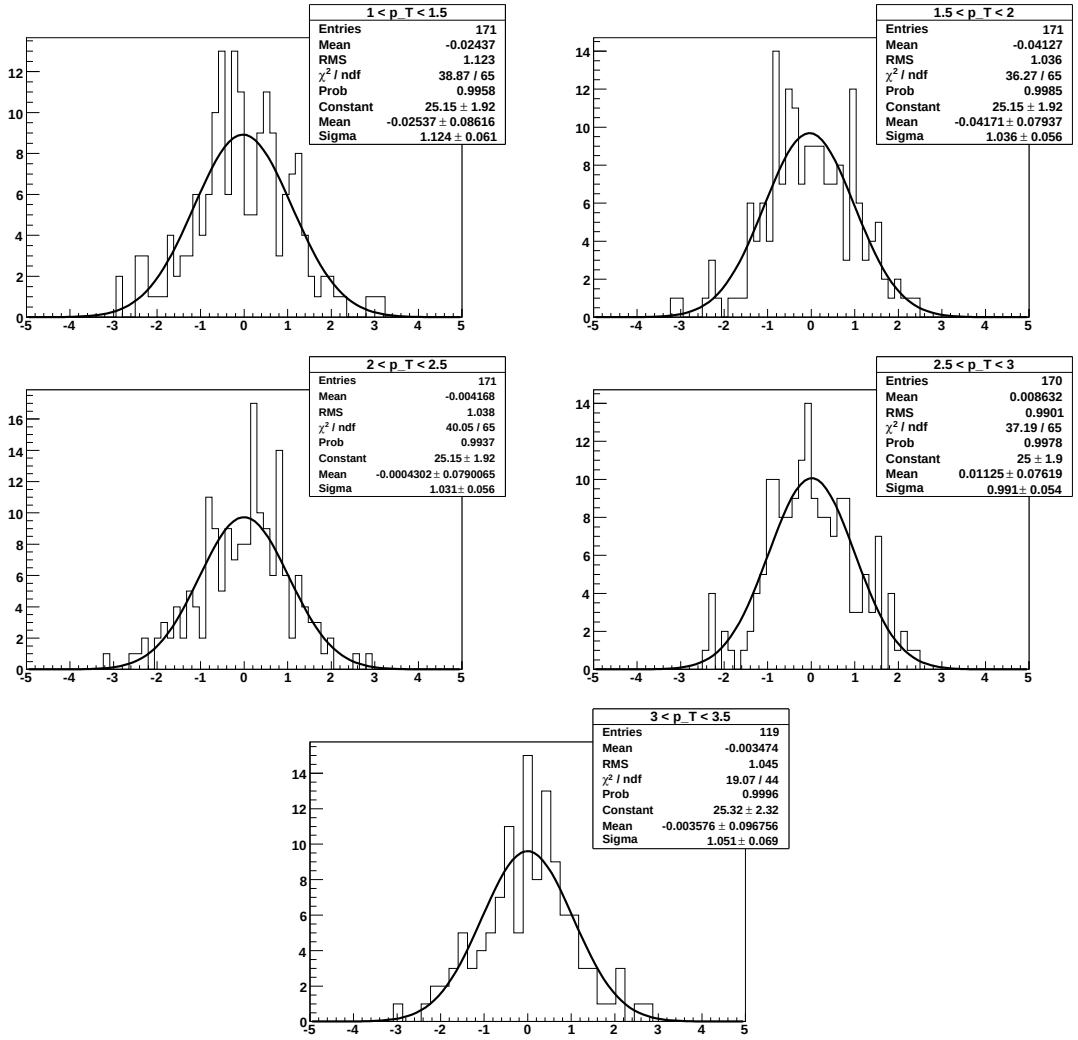


Figure D.4: Pulls for the asymmetry extracted without θ cut using the 2nd order method. The corresponding data points are shown in figure 7.25 on page 107.

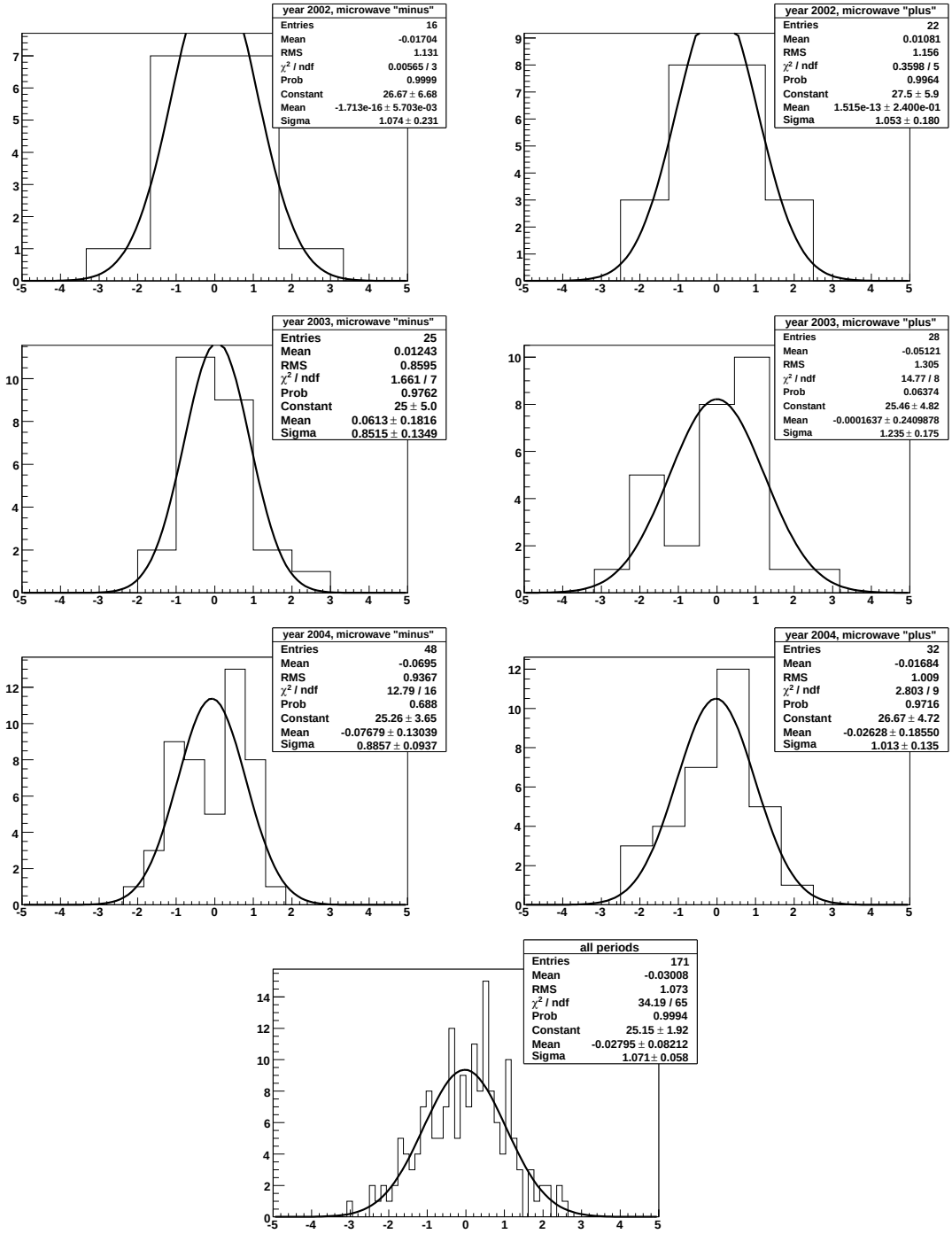


Figure D.5: Pulls for the asymmetry extracted in the upper half of the spectrometer using only positive hadrons and the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

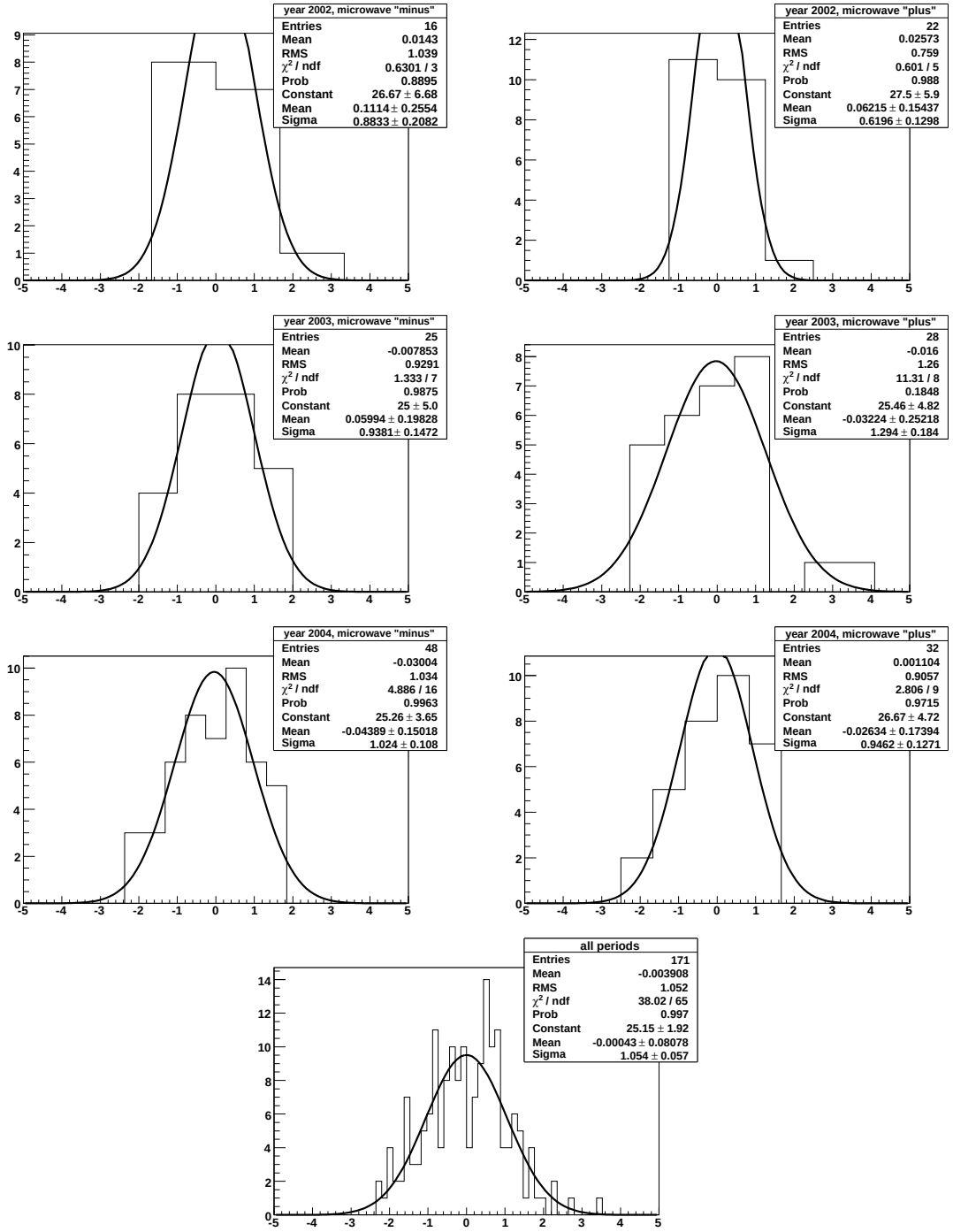


Figure D.6: Pulls for the asymmetry extracted in the lower half of the spectrometer using only positive hadrons and the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

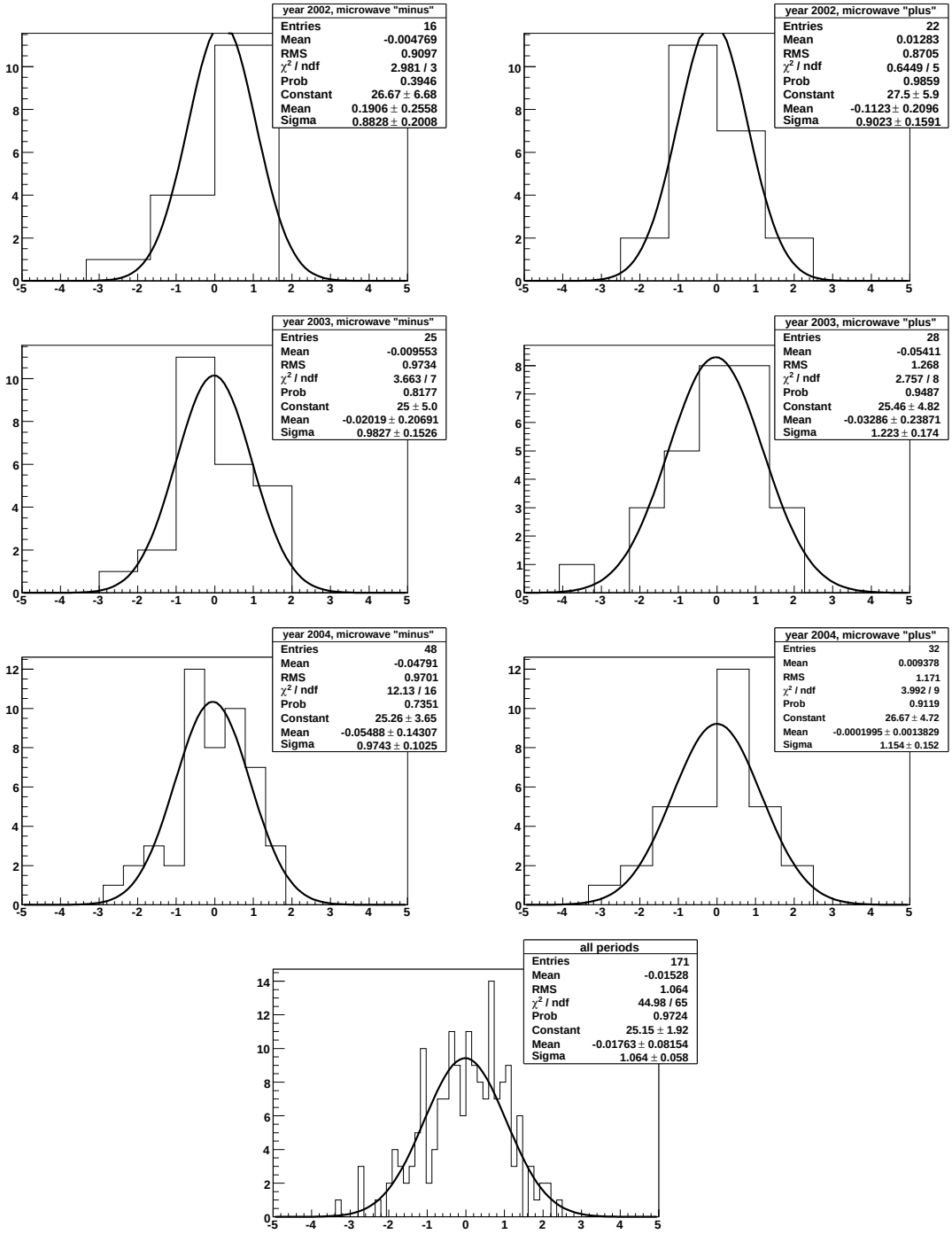


Figure D.7: Pulls for the asymmetry extracted in the left (Jura, positive x) half of the spectrometer using the 2^{nd} order method. The corresponding data points are shown in figure 7.27 on page 112.

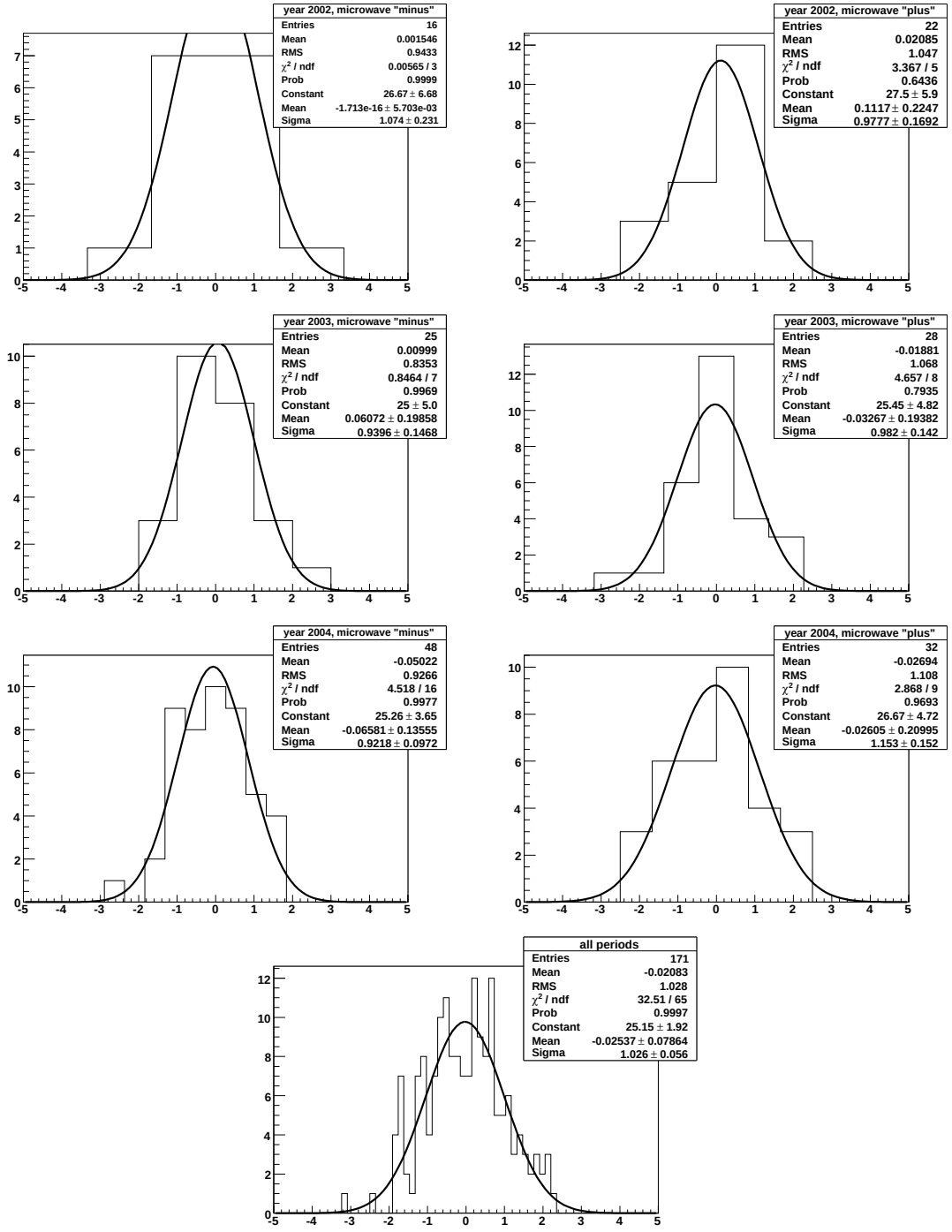


Figure D.8: Pulls for the asymmetry extracted in the right (Salève, negative x) half of the spectrometer using only positive hadrons and the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

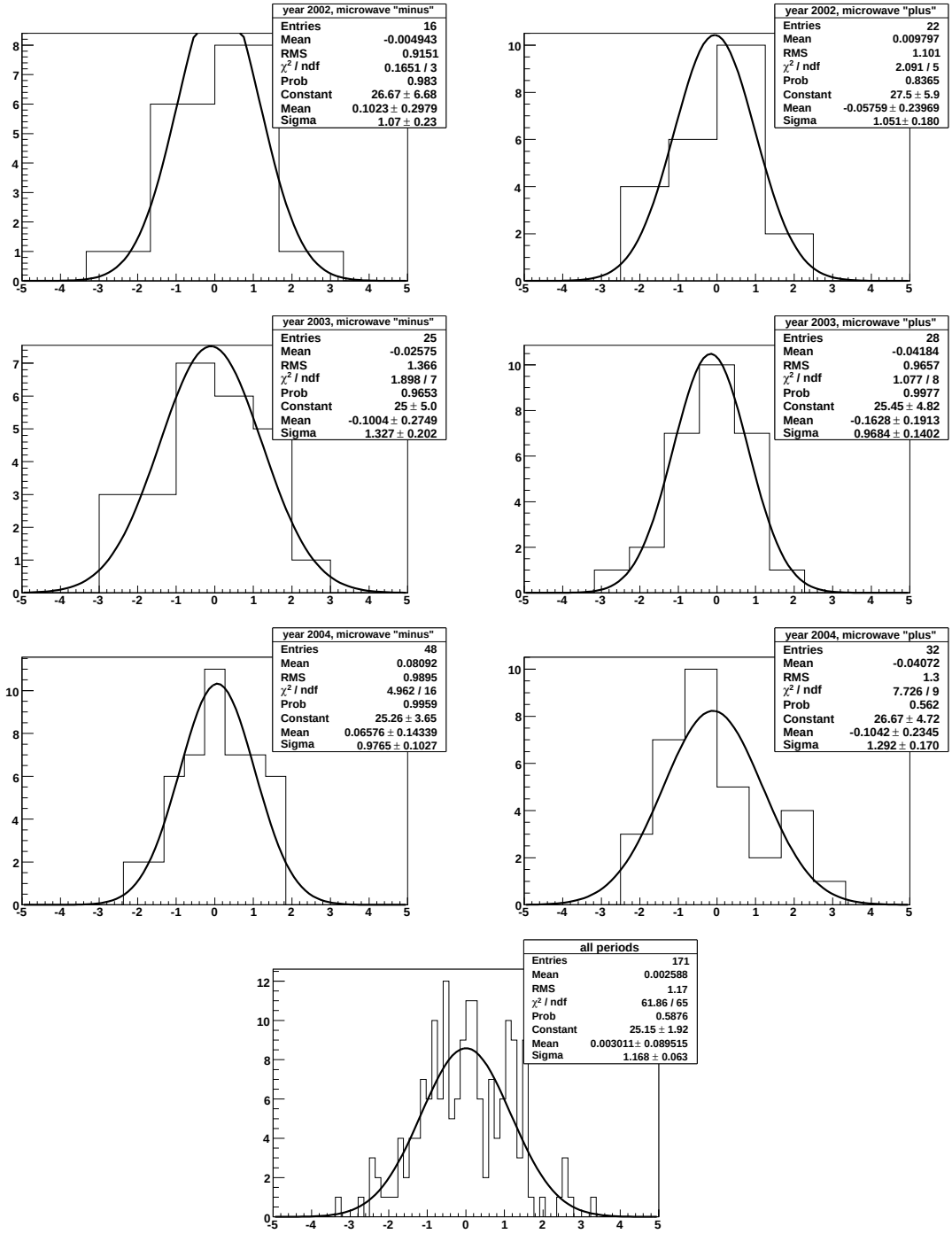


Figure D.9: Pulls for the asymmetry extracted in the upper half of the spectrometer using only negative hadrons and the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

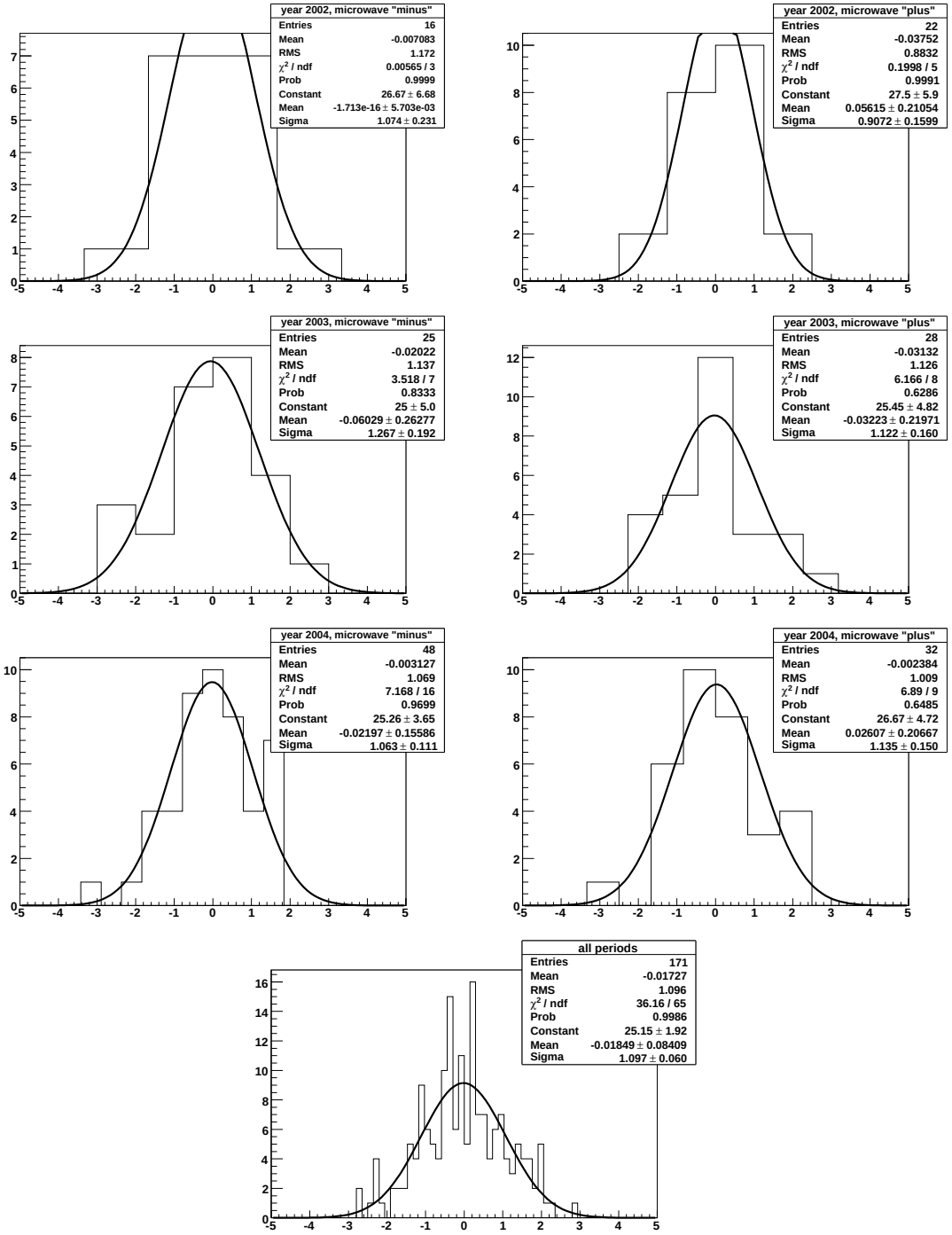


Figure D.10: Pulls for the asymmetry extracted in the lower half of the spectrometer using only negative hadrons and the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

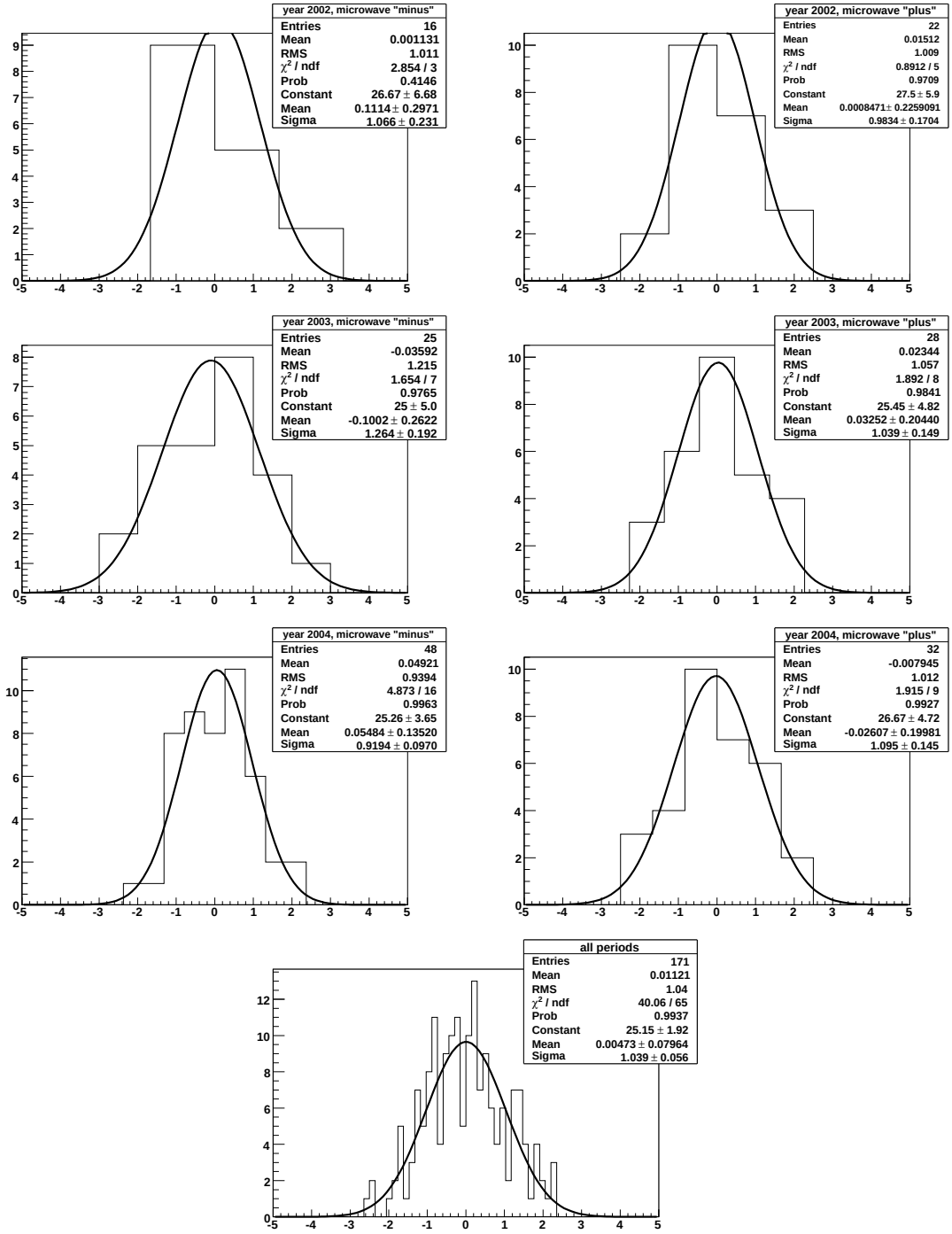


Figure D.11: Pulls for the asymmetry extracted in the left (Jura, positive x) half of the spectrometer using the 2nd order method. The corresponding data points are shown in figure 7.27 on page 112.

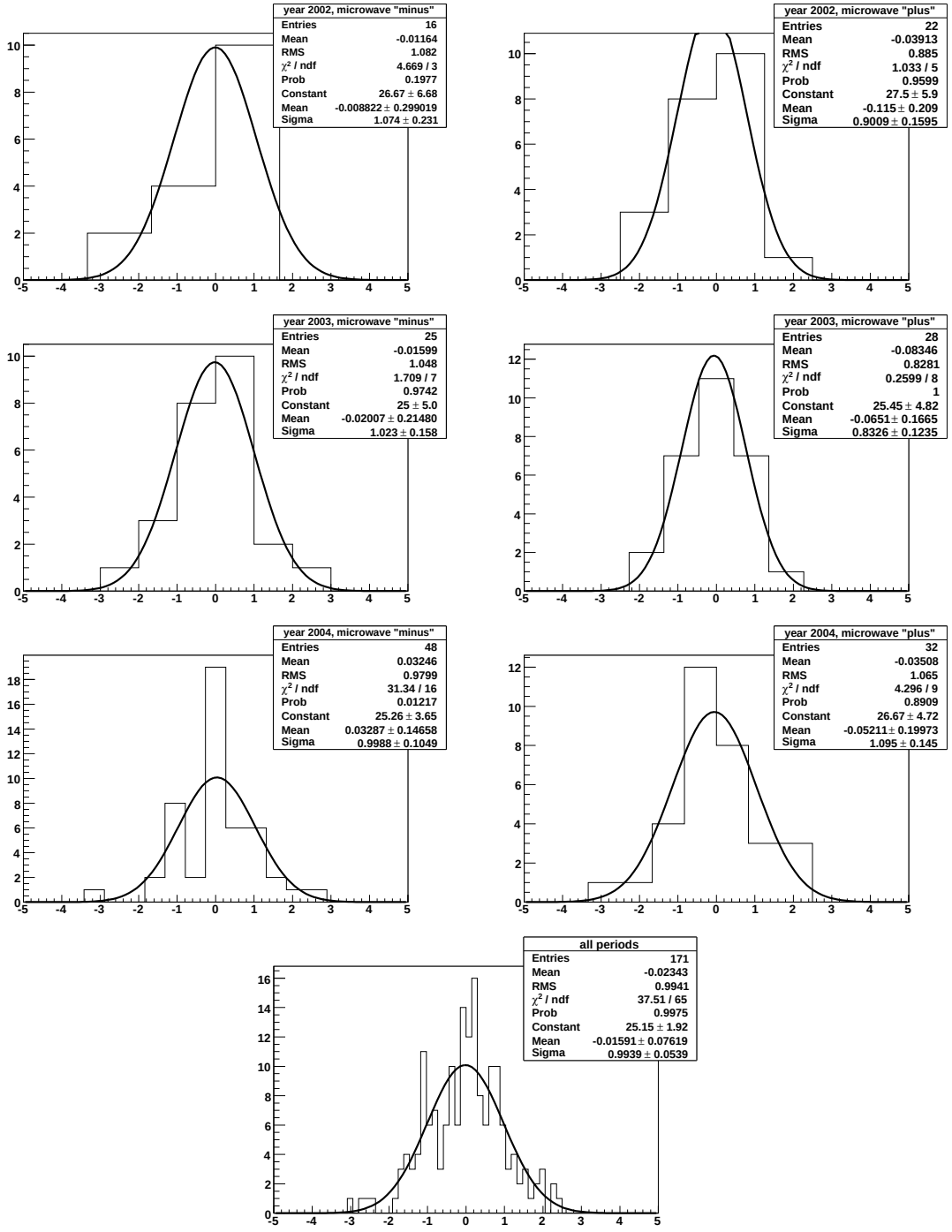


Figure D.12: Pulls for the asymmetry extracted in the right (Salève, negative x) half of the spectrometer using only negative hadrons and the 2$^{\text{nd}}$ order method. The corresponding data points are shown in figure 7.27 on page 112.

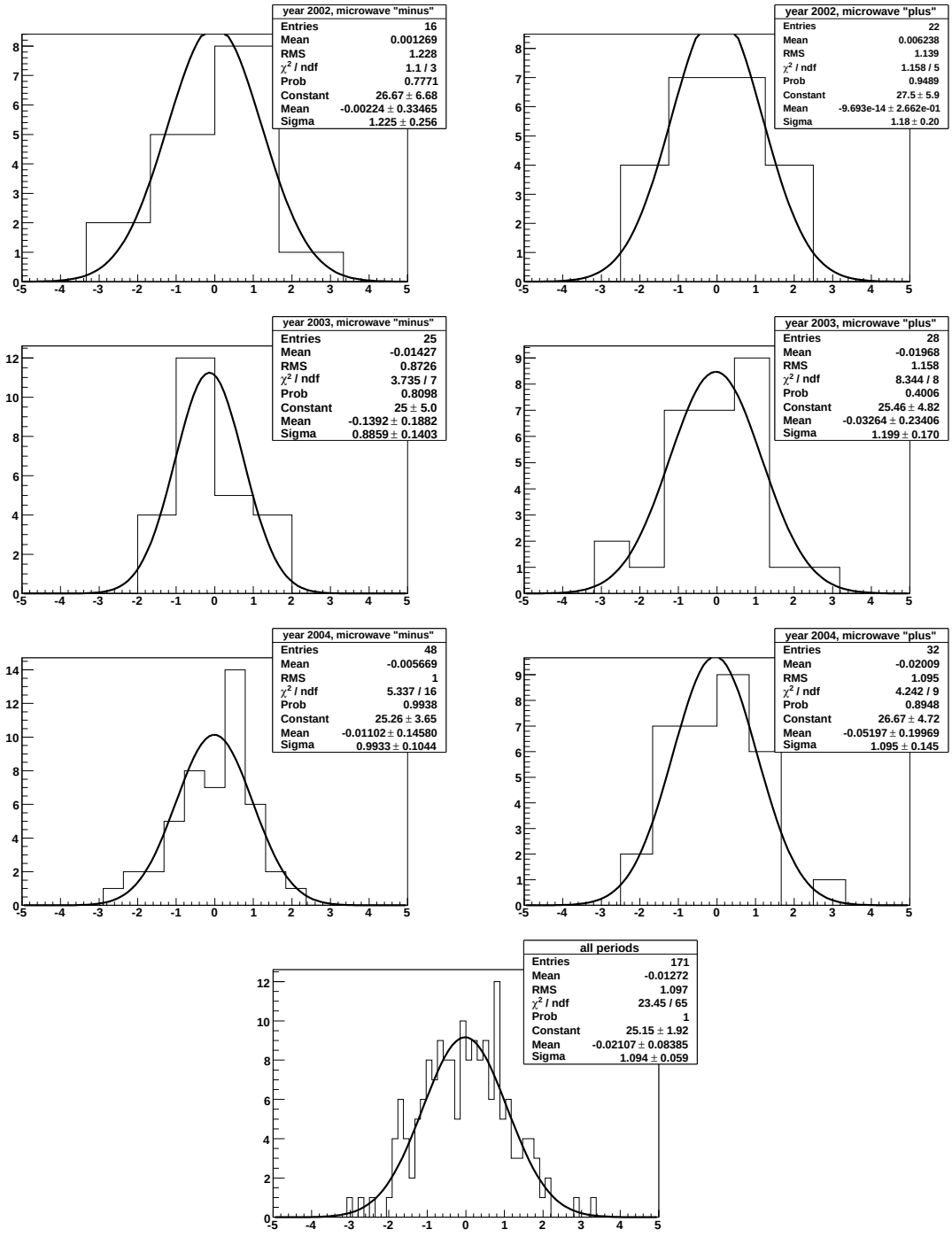


Figure D.13: Pulls for the asymmetry extracted with tighter target cuts using the 2nd order method. The corresponding data points are shown in figure 7.26 on page 110.

Appendix E

Tables

Name	Description
FATAL	A fatal condition which makes further execution of any code impossible. Not recoverable.
CRITICAL	Normal program execution is impossible, emergency mode is enabled. Call the expert and take immediate measures.
ERROR	A subsystem failed, but the program remains functional with reduced efficiency. Immediate intervention is recommended.
WARNING	A subsystem had to be restarted or an error condition was imminent but could be avoided using non-standard measures. An expert should be informed if this happens repeatedly.
NOTICE	Normal but significant condition. No action required.
INFO	Purely informational, do not care if you are not an expert.
DEBUG1	Debugging information for general debugging. Interesting mode changes are shown.
DEBUG2	Detailed debugging information which is used for debugging a specific subsystem.
DEBUG3	Very detailed information, including a significant amount of raw data. Only useful in very specific cases.

Table E.1: Description of the message priorities used in CINDERELLA. The level structure is inspired by the Apache project.

year	period	slot	size
2002	P1C	2-7	115.32 GiB
	P2A	2-7	186.67 GiB
	P2D	3-7	115.56 GiB
	P2E	2-7	166.04 GiB
	P2F	2-7	80.56 GiB
	P2G	2-7	134.40 GiB
	total		798.55 GiB
2003	P1A	2-7	260.02 GiB
	P1B	2-7	199.11 GiB
	P1C	2-7	253.06 GiB
	P1D	2-7	338.71 GiB
	P1E	3-7	428.91 GiB
	P1F	2-7	307.79 GiB
	P1I	2-7	350.13 GiB
	P1J	3-7	543.59 GiB
	total		2681.32 GiB
2004	W21	1-6	0.32 GiB
	W22	2-7	583.79 GiB
	W23	3-7	363.08 GiB
	W26	2-7	375.86 GiB
	W27	2-7	215.00 GiB
	W28	2-7	232.42 GiB
	W29	1-7	212.68 GiB
	W30	2-7	314.55 GiB
	W31	2-7	292.71 GiB
	W32	3-7	478.26 GiB
	W37	3-7	449.76 GiB
	W38	2-7	547.24 GiB
	W39	2-7	313.53 GiB
	W40	3-7	229.46 GiB
	total		4608.65 GiB
total			8088.52 GiB

Table E.2: Size of the mDST files for all periods with longitudinal target polarization. This includes runs and parts of runs which later are discarded by the bad spill rejection.

year	period	entries	size	compressed size
2002	P1C	1309764	201 MiB	130 MiB
	P2A	2627209	403 MiB	261 MiB
	P2D	1545115	236 MiB	153 MiB
	P2E	2106232	321 MiB	208 MiB
	P2F	1031766	158 MiB	102 MiB
	P2G	1758473	269 MiB	174 MiB
	total	10378559	1588 MiB	1028 MiB
2003	P1A	1222960	190 MiB	120 MiB
	P1B	1153789	179 MiB	115 MiB
	P1C	1414852	218 MiB	142 MiB
	P1D	1332594	208 MiB	135 MiB
	P1E	2308075	357 MiB	232 MiB
	P1F	1997841	308 MiB	200 MiB
	P1I	1946275	299 MiB	195 MiB
	P1J	2988395	459 MiB	299 MiB
	total	14364781	2216 MiB	1438 MiB
2004	W22	3267074	502 MiB	325 MiB
	W23	1833344	282 MiB	184 MiB
	W26	2084003	320 MiB	208 MiB
	W27	1209519	186 MiB	120 MiB
	W28	1183470	184 MiB	119 MiB
	W29	1244214	192 MiB	125 MiB
	W30	1768613	272 MiB	177 MiB
	W31	1811419	278 MiB	181 MiB
	W32	2802269	431 MiB	279 MiB
	W37	3044049	466 MiB	304 MiB
	W38	3346714	514 MiB	333 MiB
	W39	1756479	270 MiB	176 MiB
	W40	1131548	174 MiB	113 MiB
	total	26482715	4069 MiB	2645 MiB
total		51226055	7873 MiB	5111 MiB

Table E.3: Size of the μDST trees after the event selection cuts. The entries are hadron candidates and subject to the cuts described in section 7.3.3. Compressed size means the number of bytes actually used by the ROOT tree on disk.

					microwave +		microwave −	
year	period	runs	files	spills	sol. +	sol. −	sol. +	sol. −
2002	P1C	181	181	16772	—	—	279608	311153
	P2A	389	389	32696	—	—	516287	588358
	P2D	173	179	13905	316875	335946	—	—
	P2E	230	231	19897	464851	486976	—	—
	P2F	130	130	11005	238681	203540	—	—
	P2G	209	209	16417	130280	151569	183923	284106
	total	1312	1319	110692	1150687	1178031	979818	1183617
2003	P1A	163	271	10874	151153	188967	—	—
	P1B	139	235	10038	257551	240858	—	—
	P1C	153	273	11093	246627	327866	—	—
	P1D	180	328	12072	278701	260671	—	—
	P1E	260	451	19687	—	—	514968	459293
	P1F	180	335	15094	—	—	462662	391738
	P1I	131	301	15012	395341	464773	—	—
	P1J	135	442	22912	—	—	772807	543889
	total	1341	2636	116782	1329373	1483135	1750437	1394920
2004	W22	175	523	24365	736244	605146	—	—
	W23	105	313	14455	386034	400328	—	—
	W26	101	319	15141	—	—	459705	431744
	W27	59	185	8741	—	—	257710	247246
	W28	160	263	18983	—	—	234575	265406
	W29	120	218	18644	—	—	248185	274988
	W30	116	267	13930	—	—	371767	367757
	W31	101	249	14067	—	—	400769	351627
	W32	165	434	21002	—	—	568923	599191
	W37	161	398	23798	708980	569218	—	—
	W38	165	489	25405	686696	734952	—	—
	W39	79	272	12511	—	—	387182	340246
	W40	63	193	8674	—	—	239981	240944
	total	1570	4123	219716	2517954	2309644	3168797	3119149
total		4223	8078	447190	4998014	4970810	5899052	5697686

Table E.4: Overview of the statistics within the different data taking periods. mDST files are split into chunks of at most 1.5 GiB, which creates the difference between number of runs and number of files. The numbers given in the rightmost four columns count the number of high p_T hadrons as defined in section 7.3.3, not including the θ cut and with $p_T > 1\,\mathrm{GeV}/c$.

bin/year	A_{rep}	$(\delta A_{\text{syst}})_{\text{rep}}$	A_{ran}	$(\delta A_{\text{syst}})_{\text{ran}}$	δA_{syst}(year)	δA_{syst}
$1 < p_T < 1.5$						
2002	−0.00313	0.00052	$[0.011, 0.026], 2.0\sigma$	0.00437	0.00440	0.00078
2003	0.00414	0.00011	$[0.011, 0.024], 2.2\sigma$	0.00339	0.00339	0.00097
2004	0.00476	0.00057	$[0.000, 0.008], 0.1\sigma$	0.00100	0.00115	0.00062
						0.00139
$1.5 < p_T < 2$						
2002	0.02744	0.00451	$[0.000, 0.054], 0.3\sigma$	0.00898	0.01005	0.00175
2003	−0.00982	0.00032	$[0.000, 0.038], 0.1\sigma$	0.00537	0.00538	0.00153
2004	0.00588	0.00073	$[0.000, 0.046], 0.6\sigma$	0.00549	0.00553	0.00300
						0.00379
$2 < p_T < 2.5$						
2002	0.05385	0.00810	$[0.000, 0.178], 0.3\sigma$	0.02937	0.03047	0.00503
2003	0.00618	0.00023	$[0.000, 0.126], 0.2\sigma$	0.01810	0.01810	0.00514
2004	−0.01305	0.00166	$[0.023, 0.148], 1.1\sigma$	0.01780	0.01788	0.00985
						0.01220
$2.5 < p_T < 3$						
2002	−0.04827	0.00730	$[0.000, 0.521], 0.3\sigma$	0.08753	0.08783	0.01300
2003	0.00304	0.00004	$[0.000, 0.382], 0.3\sigma$	0.05566	0.05566	0.01597
2004	0.05101	0.00647	$[0.000, 0.260], 0.2\sigma$	0.03162	0.03227	0.01824
						0.02751
$3 < p_T < 3.5$						
2002	−0.23800	0.02576	$[0.000, 0.531], 0.0\sigma$	0.12876	0.13131	0.01337
2003	−0.16817	0.00265	$[0.000, 1.107], 0.5\sigma$	0.19721	0.19723	0.05596
2004	−0.22465	0.03141	$[0.270, 1.096], 1.2\sigma$	0.15220	0.15541	0.09548
						0.11148

Table E.5: Summary of the systematic errors for the standard cut set and 2nd order asymmetry extraction method without weighting.

bin/year	A_{rep}	$(\delta A_{\text{syst}})_{\text{rep}}$	A_{ran}	$(\delta A_{\text{syst}})_{\text{ran}}$	δA_{syst}(year)	δA_{syst}
$1 < p_T < 1.5$						
2002	−0.00314	0.00052	[0.011, 0.026], 2.0σ	0.00437	0.00440	0.00078
2003	0.00417	0.00011	[0.011, 0.024], 2.3σ	0.00341	0.00341	0.00097
2004	0.00476	0.00056	[0.000, 0.008], 0.1σ	0.00101	0.00115	0.00062
						0.00140
$1.5 < p_T < 2$						
2002	0.02750	0.00456	[0.000, 0.054], 0.3σ	0.00900	0.01010	0.00176
2003	−0.00971	0.00031	[0.000, 0.038], 0.1σ	0.00539	0.00540	0.00153
2004	0.00588	0.00073	[0.000, 0.046], 0.6σ	0.00549	0.00554	0.00300
						0.00380
$2 < p_T < 2.5$						
2002	0.05406	0.00815	[0.000, 0.179], 0.3σ	0.02953	0.03064	0.00508
2003	0.00579	0.00020	[0.000, 0.126], 0.2σ	0.01814	0.01814	0.00515
2004	−0.01329	0.00169	[0.016, 0.147], 1.0σ	0.01769	0.01777	0.00977
						0.01216
$2.5 < p_T < 3$						
2002	−0.05560	0.00844	[0.000, 0.523], 0.3σ	0.08768	0.08809	0.01310
2003	0.00321	0.00003	[0.000, 0.365], 0.3σ	0.05315	0.05315	0.01518
2004	0.05053	0.00641	[0.000, 0.247], 0.1σ	0.03001	0.03069	0.01736
						0.02652
$3 < p_T < 3.5$						
2002	−0.05519	0.00055	[0.000, 0.000], 0.0σ	0.00000	0.00055	0.00006
2003	−0.09314	0.00329	[0.000, 0.848], 0.1σ	0.14866	0.14870	0.04247
2004	−0.22545	0.02477	[0.000, 0.857], 0.3σ	0.11891	0.12147	0.07457
						0.08582

Table E.6: Summary of the systematic errors for the standard cut set and 1st order asymmetry extraction method with global correction factor.

bin/year	A_{rep}	$(\delta A_{\text{syst}})_{\text{rep}}$	A_{ran}	$(\delta A_{\text{syst}})_{\text{ran}}$	$\delta A_{\text{syst}}(\text{year})$	δA_{syst}
$1 < p_T < 1.5$						
2002	−0.00314	0.00052	$[0.011, 0.026], 2.0\sigma$	0.00436	0.00439	0.00078
2003	0.00418	0.00011	$[0.011, 0.024], 2.2\sigma$	0.00340	0.00340	0.00097
2004	0.00476	0.00056	$[0.000, 0.008], 0.1\sigma$	0.00100	0.00115	0.00062
						0.00139
$1.5 < p_T < 2$						
2002	0.02746	0.00456	$[0.000, 0.054], 0.3\sigma$	0.00896	0.01005	0.00175
2003	−0.00969	0.00030	$[0.000, 0.037], 0.1\sigma$	0.00537	0.00537	0.00153
2004	0.00589	0.00073	$[0.000, 0.046], 0.6\sigma$	0.00548	0.00553	0.00299
						0.00379
$2 < p_T < 2.5$						
2002	0.05350	0.00812	$[0.000, 0.176], 0.3\sigma$	0.02911	0.03022	0.00499
2003	0.00603	0.00021	$[0.000, 0.125], 0.2\sigma$	0.01792	0.01792	0.00508
2004	−0.01327	0.00168	$[0.014, 0.147], 1.0\sigma$	0.01765	0.01773	0.00977
						0.01209
$2.5 < p_T < 3$						
2002	−0.04552	0.00685	$[0.000, 0.485], 0.2\sigma$	0.08158	0.08187	0.01210
2003	0.00244	0.00003	$[0.000, 0.352], 0.2\sigma$	0.05127	0.05127	0.01467
2004	0.05095	0.00642	$[0.000, 0.228], 0.1\sigma$	0.02777	0.02850	0.01613
						0.02494
$3 < p_T < 3.5$						
2002	−0.08294	0.00330	$[0.000, 0.000], 0.0\sigma$	0.00000	0.00330	0.00033
2003	−0.10433	0.00212	$[0.000, 0.771], 0.1\sigma$	0.13643	0.13644	0.03879
2004	−0.20759	0.02629	$[0.000, 0.790], 0.2\sigma$	0.11027	0.11336	0.06987
						0.07992

Table E.7: Summary of the systematic errors for the standard cut set and 1st order asymmetry extraction method with individual correction factors.

bin/year	A_{rep}	$(\delta A_{\text{syst}})_{\text{rep}}$	A_{ran}	$(\delta A_{\text{syst}})_{\text{ran}}$	δA_{syst}(year)	δA_{syst}
$1 < p_T < 1.5$						
2002	−0.00730	0.00124	[0.007, 0.021], 1.6σ	0.00344	0.00365	0.00064
2003	0.00307	0.00009	[0.011, 0.021], 3.1σ	0.00297	0.00297	0.00084
2004	0.00506	0.00061	[0.000, 0.010], 0.4σ	0.00120	0.00135	0.00073
						0.00128
$1.5 < p_T < 2$						
2002	0.02281	0.00387	[0.000, 0.048], 0.4σ	0.00790	0.00879	0.00150
2003	−0.00672	0.00025	[0.000, 0.039], 0.4σ	0.00562	0.00562	0.00158
2004	0.00517	0.00065	[0.000, 0.030], 0.3σ	0.00360	0.00366	0.00200
						0.00296
$2 < p_T < 2.5$						
2002	0.02086	0.00340	[0.000, 0.123], 0.2σ	0.02035	0.02063	0.00339
2003	0.00922	0.00040	[0.000, 0.077], 0.1σ	0.01112	0.01113	0.00312
2004	0.00487	0.00061	[0.000, 0.000], 0.0σ	0.00000	0.00061	0.00034
						0.00462
$2.5 < p_T < 3$						
2002	−0.11598	0.01603	[0.000, 0.301], 0.1σ	0.05010	0.05260	0.00817
2003	0.05547	0.00043	[0.000, 0.338], 0.7σ	0.04903	0.04903	0.01366
2004	0.05815	0.00708	[0.000, 0.164], 0.1σ	0.01980	0.02103	0.01190
						0.01988
$3 < p_T < 3.5$						
2002	−0.33589	0.06022	[0.518, 1.299], 1.9σ	0.22865	0.23645	0.03717
2003	−0.06140	0.00073	[0.000, 0.851], 0.7σ	0.13386	0.13386	0.03757
2004	−0.11471	0.01685	[0.000, 0.523], 0.3σ	0.06726	0.06934	0.03898
						0.06567

Table E.8: Summary of the systematic errors for the cut set without θ cut and 2^{nd} order asymmetry extraction method without weighting.

Acknowledgements

First I would like to thank Stephan Paul for creating such a powerful working environment and letting me study with such knowledgeable and kind persons as will be mentioned below. There is no greater satisfaction than to rise to a difficult challenge and be given the tools to master it.

The single most important information source on all kinds of scientific and technical questions at our chair is Jan Friedrich, who also helped a lot in editing this thesis. His honest criticism and moral support are highly appreciated.

My work at CERN would have been much more difficult if not for Lars Schmitt who taught me some of the more obscure things there are to know about computers and gave me a solid base to work from when he handed over the data acquisition.

I would like to thank Igor Konorov for his support in all my projects with COMPASS and many good discussions, technical as well as otherwise. It is a pleasure working with someone as enthusiastic and competent.

Many thanks also to the whole group at E18. All of you made work a pleasure, with countless discussions as are only possible among physicists. Specifically I would like to thank Boris Grube for just being who he is, Thiemo Nagel for all his support, new perspectives and for taking over the online filter, Bernhard Ketzer and Joachim Hartmann for fruitful discussions, Annemarie Dinkelbach and Steffi Grabmüller for adding charm to the working group, Martin Simson for maintaining the coffee service, Alexander Mann for all his help with scanning and printing and Karin Frank for accomplishing top-notch secretary service despite being constantly — and at times extremely — overloaded.

Thanks also to Norbert Kaiser, who patiently read the theory chapter and provided me with valuable feedback.

Taking some time off and replenishing my spirits would not have been possible without Martti Sinisalo owning and letting a beautiful cottage in Karhujärvi near the polar circle in Finland.

Last and most importantly I would like to thank my wonderful wife for continuously and enduringly supporting my work. You enjoyed my successes, listened when I needed it and put me back on track when I started to run astray. And last but not least, you gave me two nice kids which I treasure above all.

Own Contributions

The concept of the COMPASS configuration server was developed by me in collaboration with Lars Schmitt, Igor Konorov and Fritz-Herbert Heinsius, while the implementation of the framework, the database part and the GeSiCA module has been done by me alone. The CATCH module has been written by Fritz-Herbert Heinsius.

The concept of the CINDERELLA online filter has been laid out by Lars Schmitt and myself with substantial extensions by Thiemo Nagel. The implementation of the core parts—thread interplay, buffering, initial version of the configuration infrastructure, message system, memory allocation debugging—has been done by me, the same is true for the ALICE DATE event format decoding, handling of ALICE DATE event types, trigger time and silicon raw data decoding. Thiemo Nagel and myself have conceived the conditional coincidence algorithm, which was subsequently implemented by him.

For the 2006 run the event building network of the DAQ has been setup anew, hardware as well as software. This work was started by Lars Schmitt but the main part was done by me.

Apart from the infrastructure supplied by COMPASS, which includes the badspill lists, everything in the analysis chapters has been done by myself.

Index

Bibliography

[1] Yu. L. Dokshitzer. Sov. Phys. JTEP **46** (1977) 641.
cited in frontmatter (p. XVIII), 6.1 (p. 51)

[2] V. N. Gribov and Lipatov L. N. Sov. J. Nucl. Phys. **15** (1972) 438.
cited in frontmatter (p. XVIII), 6.1 (p. 51)

[3] L. N. Lipatov. Sov. J. Nucl. Phys. **20** (1975) 95. *cited in frontmatter (p. XVIII), 6.1 (p. 51)*

[4] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. Nucl. Phys. **B126** (1977) 298. *cited in frontmatter (p. XVIII), 6.1 (p. 51)*

[5] Distributed Information Management System. http://dim.web.cern.ch/dim/.
cited in frontmatter (p. XVIII), 3.1.2 (p. 25)

[6] B. Grube. The Trigger Control System and the Common GEM and Silicon Readout for the COMPASS Experiment. Diploma thesis, TU München, 2001.
cited in frontmatter (p. XVIII), 3.0.2 (p. 23), 4.1 (p. 30), 4.4 (p. 32)

[7] F. Sauli. GEM: A new concept for electron amplification in gas detectors. Nucl. Instrum. Meth. **A386** (1997) 531–534. *cited in frontmatter (p. XIX)*

[8] W. Beusch et al. OMEGA PRIME: a project of improving the omega particle detector system. CERN SPSC 77-70 (1977). *cited in frontmatter (p. XX)*

[9] S-LINK homepage. http://hsi.web.cern.ch/hsi/s-link/. *cited in frontmatter (p. XXI)*

[10] E. Nappi et al. Semi-inclusive Muon Scattering from a Polarised Target. CERN SPSLC 95-27, SPSLC I204 (1995). *cited in chapter 2 (p. 5)*

[11] Yu. Alexandrow et al. CHEOPS: CHarm Experiment with Omnipurpose Setup. CERN SPSLC 95-22, SPSLC I202 (1995). *cited in chapter 2 (p. 5)*

[12] COMPASS collaboration. Common Muon and Proton Apparatus for Structure and Spectroscopy. CERN **SPSLC96-14** (1996). *cited in chapter 2 (p. 5), 2.1 (p. 5), 2.2 (p. 8)*

[13] J. J. Aubert et al. Measurements Of The Nucleon Structure Functions F2(N) In Deep Inelastic Muon Scattering From Deuterium And Comparison With Those From Hydrogen And Iron. Nucl. Phys. **B293** (1987) 740. *cited in 2.1.1 (p. 6)*

[14] P. Amaudruz et al. A Re-Evaluation of the Nuclear Structure Function Ratios for D, He, 6Li, C and Ca. Nucl. Phys. **B441** (1995) 3–11. hep-ph/9503291. *cited in 2.1.1 (p. 6)*

[15] J. Ashman et al. A measurement of the spin asymmetry and determination of the structure function g(1) in deep inelastic muon proton scattering. Phys. Lett. **B206** (1988) 364. *cited in 2.1.1 (p. 6)*

[16] B. Adeva et al. Spin Asymmetries A(1) And Structure Functions G1 Of The Proton And The Deuteron From Polarized High-Energy Muon Scattering. Phys. Rev. **D58** (1998) 112001. *cited in 2.1.1 (p. 6), 6.5 (p. 60)*

[17] B. Adeva et al. A Next-To-Leading Order Qcd Analysis Of The Spin Structure Function G(1). Phys. Rev. **D58** (1998) 112002. *cited in 2.1.1 (p. 6)*

[18] V. Yu. Alexakhin et al. The deuteron spin-dependent structure function g1(d) and its first moment. Phys. Lett. **B647** (2007) 8–17. hep-ex/0609038. *cited in 2.1.1 (p. 6), 6.1 (p. 52), 6.6.1 (p. 62)*

[19] R.L. Jaffe and Xiang-Dong Ji. Chiral odd parton distributions and polarized Drell-Yan. Phys. Rev. Lett. **67** (1991) 552–555. *cited in 2.1.1 (p. 6)*

[20] R.L. Jaffe and Xiang-Dong Ji. Chiral odd parton distributions and Drell-Yan processes. Nucl. Phys. **B375** (1992) 527–560. *cited in 2.1.1 (p. 6)*

[21] J. Collins. Fragmentation of transversely polarized quarks probed in transverse momentum distributions. Nucl. Phys. **B396** (1993) 161–182. *cited in 2.1.1 (p. 6)*

[22] E. S. Ageev et al. A new measurement of the Collins and Sivers asymmetries on a transversely polarised deuteron target. Nucl. Phys. **B765** (2007) 31–70. hep-ex/0610068. *cited in 2.1.1 (p. 6)*

[23] B. Grube. A Trigger Control System for COMPASS and a Measurement of the Transverse Polarization of Λ and Ξ Hyperons from Quasi-Real Photo-Production. PhD thesis, TU München, 2006. *cited in 2.1.1 (p. 7), 2.2 (p. 8), 4.1 (p. 30)*

[24] M. Wiesmann. A Silicon Microstrip Detector for COMPASS and A First Measurement of the Transverse Polarization of Λ^0-Hyperons from Quasi-Real Photo-Production. PhD thesis, Technical University Munich, 2004. *cited in 2.1.1 (p. 7)*

[25] J.M. Friedrich. Measurement of Transverse Λ Polarisation at COMPASS. In *SPIN2004: proceedings of the 16th International Spin Physics Symposium*, page 396, Trieste, Italy, October 2004. *cited in 2.1.1 (p. 7)*

[26] V.Y. Alexakhin. Longitudinal Polarization of Λ and $\overline{\Lambda}$ Hyperons in Deep-inelastic Scattering. In *SPIN2004: proceedings of the 16th International Spin Physics Symposium*, page 393, Trieste, Italy, October 2004. *cited in 2.1.1 (p. 7)*

[27] Mark II collaboration. Two-photon production of pion pairs. Phys. Rev. **D42** (1990) 1350–1367. *cited in 2.1.2 (p. 7)*

[28] The CELLO collaboration. An experimental study of the process $\gamma\gamma \to \pi^+\pi^-$. Z. Phys. **C56** (1992) 381–390. *cited in 2.1.2 (p. 7)*

[29] H. Primakoff. Photo-Production of Neutral Mesons in Nuclear Electric Fields and the Mean Life of the Neutral Meson. Phys. Rev. **81** (1951) 899. *cited in 2.1.2 (p. 7)*

[30] Y. M. Antipov et al. Experimental estimation of the sum of Pion Electrical and Magnetic Polarizabilities. Z. Phys. **C26** (1985) 495–497. *cited in 2.1.2 (p. 7)*

[31] M. Buénerd. Prospects for hadron electromagnetic polarizability measurement by radiative scattering on a nuclear target. Nucl. Instrum. Meth. **A361** (1995) 111–128. *cited in 2.1.2 (p. 7)*

[32] J. Ahrens et al. Measurement of the pi+ meson polarizabilities via the gamma p –> gamma pi+ n reaction. Eur. Phys. J. **A23** (2005) 113–127. nucl-ex/0407011. *cited in 2.1.2 (p. 7)*

[33] Juerg Gasser, Mikhail A. Ivanov, and Mikko E. Sainio. Revisiting gamma gamma –> pi+ pi- at low energies. Nucl. Phys. **B745** (2006) 84–108. hep-ph/0602234. *cited in 2.1.2 (p. 7)*

[34] J. Bijnens et al. THREE PSEUDOSCALAR PHOTON INTERACTIONS IN CHIRAL PERTURBATION THEORY. Phys. Lett. **B237** (1990) 488. *cited in 2.1.2 (p. 7)*

[35] T. Nagel. Cinderella: an Online Filter for the COMPASS Experiment. Diploma thesis, TU München, Physik-Department E18, 2005. *cited in 2.2 (p. 8), 5.1 (p. 38), 5.3 (p. 45)*

[36] P. Abbon et al. The COMPASS Experiment at CERN. Nucl. Instrum. Meth. **A577** (2007) 455–518. hep-ex/0703049. *cited in 2.2 (p. 8), 4.2 (p. 30), 7.4.1 (p. 94)*

[37] N. Takabayashi. Polarized target for the measurement of the gluon contribution to the nucleon spin in the COMPASS exeriment. PhD thesis, Nagoya University, 2002. *cited in 2.2.5 (p. 14), 2.2.5 (p. 15)*

[38] C. Bernet et al. The COMPASS trigger system for muon scattering. Nucl. Instrum. Meth. **A550** (2005) 217–240. *cited in 2.2.6 (p. 17), 7.3.3 (p. 88)*

[39] M. Leberig. Das COMPASS-Triggersystem zur Messung des Gluonbeitrags DG zum Protonspin. PhD thesis, Universität Mainz, 2002. *cited in 2.2.6 (p. 17)*

[40] Physics Analysis Workstation – PAW. http://wwwasd.web.cern.ch/wwwasd/paw/. *cited in 2.3 (p. 17)*

[41] ROOT web page. http://root.cern.ch. *cited in 2.3 (p. 17), C.3.1 (p. 144)*

[42] Phast web page. http://ges.home.cern.ch/ges/phast. *cited in 2.3 (p. 17)*

[43] COMPASS offline pages. http://wwwcompass.cern.ch/compass/software/offline/welcome.html. *cited in 2.3 (p. 18), 7.3.3 (p. 89)*

[44] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. Nucl. Instrum. Meth. **A262** (1987) 444–450. *cited in 2.3.1 (p. 21)*

[45] E.J. Wolin et al. Covariance matrices for track fitting with the Kalman filter. Nucl. Instrum. Meth. **A329** (1993) 493–500. *cited in 2.3.1 (p. 21)*

[46] VME FAQ. http://www.vita.com/vmefaq/. *cited in 3.1.1 (p. 24)*

[47] A.W. Thomas and W. Weise. The Structure of the Nucleon. WILEY VCH Verlag Berlin GmbH, 2001. *cited in chapter 6 (p. 49)*

[48] B. Jager, M. Stratmann, and W. Vogelsang. Longitudinally polarized photoproduction of inclusive hadrons at fixed-target experiments. Eur. Phys. J. **C44** (2005) 533–543. hep-ph/0505157. *cited in chapter 6 (p. 49), 6.6.3 (p. 64), 6.6.3 (p. 65), 6.6.3 (p. 66), 6.6.3 (p. 67), 7.4.4 (p. 104), 7.7 (p. 115), chapter 8 (p. 119)*

[49] M. Gell-Mann. Symmetries of Baryons and Mesons. Phys. Rev. **125** (1962) 1067–1084. *cited in 6.1 (p. 51)*

[50] S. D. Drell and Tung-Mow Yan. Connection Of Elastic Electromagnetic Nucleon Form-Factors At Large Q^2 And Deep Inelastic Structure Functions Near Threshold. Phys. Rev. Lett. **24** (1970) 181–185. *cited in 6.1 (p. 51)*

[51] D. P. Barber et al. Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA. Phys. Rev. Lett. **4312** (1979) 830–833. *cited in 6.1 (p. 52)*

[52] K. Abe et al. Measurements of R = sigma(L)/sigma(T) for 0.03 < x < 0.1 and fit to world data. Phys. Lett. **B452** (1999) 194–200. hep-ex/9808028. *cited in 6.3 (p. 55)*

[53] Richard P. Feynman. Very High-Energy Collisions of Hadrons. Phys. Rev. Lett. **23** (1969) 1415–1417. *cited in 6.4 (p. 55)*

[54] K. Wilson. Non-Lagrangian Models of Current Algebra. Phys. Rev. **179** (1969) 1499. *cited in 6.4 (p. 55)*

[55] H1 Collaboration, C. Adloff, et al. Measurement and QCD Analysis of Neutral and Charged Current Cross Sections at HERA. Eur.Phys.J. **C30** (2003) 1–32. hep-ex/0304003. *cited in 6.4 (p. 57)*

[56] K. Gottfried. Sum Rule For High-Energy Electron-Proton Scattering. Phys. Rev. Lett. **18** (1967) 1174. *cited in 6.5 (p. 59)*

[57] E.-M. Kabuß. Final results from the NMC. (1997). hep-ph/9706435. *cited in 6.5 (p. 59)*

[58] J. D. Bjorken. Applications Of The Chiral U(6) X (6) Algebra Of Current Densities. Phys. Rev. **148** (1966) 1467–1478. *cited in 6.5 (p. 59)*

[59] J. R. Ellis and R. L. Jaffe. A Sum Rule For Deep Inelastic Electroproduction From Polarized Protons. Phys. Rev. **D9** (1974) 1444. *cited in 6.5 (p. 60)*

[60] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren. The alpha(s)**3 approximation of quantum chromodynamics to the Ellis-Jaffe sum rule. Phys. Lett. **B404** (1997) 153–160. hep-ph/9702435. *cited in 6.5 (p. 60)*

[61] S. Eidelman et al. The Review of Particle Physics. Physics Letters B **592** (2004) 1. *cited in 6.5 (p. 60)*

[62] G. Altarelli and G. G. Ross. The Anomalous Gluon Contribution To Polarized Leptoproduction. Phys. Lett. **B212** (1988) 391. *cited in 6.5 (p. 61)*

[63] PYTHIA. http://www.thep.lu.se/torbjorn/Pythia.html. *cited in 6.6.1 (p. 62)*

[64] E. S. Ageev et al. Gluon polarization in the nucleon from quasi-real photoproduction of high-p(T) hadron pairs. Phys. Lett. **B633** (2006) 25–32. hep-ex/0511028. *cited in 6.6.1 (p. 63)*

[65] Daniel de Florian, Rodolfo Sassot, and Marco Stratmann. Global analysis of fragmentation functions for pions and kaons and their uncertainties. (2007). hep-ph/0703242. *cited in 6.6.3 (p. 69)*

[66] A.K.A. Azzam, M.A. Fawzy, E.M. Hassan, and A.M. Yasser. Electron-Deuteron Tensor Polarization and D-State Probability. Chinese Journal of Physics **434** (2005) 813–822. *cited in 7.2.2 (p. 75)*

[67] J.-F. Rajotte. presentation at Analysis Meeting, April 2007. rajotte@mail.cern.ch. *cited in 7.2.4 (p. 77)*

[68] J. Pretz. A New Method for Asymmetry Extraction. (2004). *COMPASS note 2004-11*. *cited in 7.2.5 (p. 79), 7.2.7 (p. 82)*

[69] J.-M. Le Goff. Asymmetry extraxtion. (2004). *COMPASS note 2004-3*. *cited in 7.2.6 (p. 80)*

[70] S. Koblitz F.-H. Heinsius. COMPASS luminosity for 2002-2004. (2006). *COMPASS note 2006-5*. *cited in 7.4.1 (p. 94), 7.4.1 (p. 95)*

[71] GEANT – Detector Description and Simulation Tool. http://wwwasd.web.cern.ch/wwwasd/geant/. *cited in 7.4.2 (p. 96)*

[72] B. Badelek. presentation at Analysis Meeting, Oct 2004. badelek@mail.cern.ch. *cited in 7.4.3 (p. 104)*

[73] Daniel de Florian and Werner Vogelsang. Threshold resummation for the inclusive hadron cross- section in p p collisions. Phys. Rev. **D71** (2005) 114004. hep-ph/0501258. *cited in 7.4.4 (p. 105)*

[74] W. Vogelsang. private communication. vogelsan@quark.phy.bnl.gov. *cited in 7.4.4 (p. 105)*

[75] Roger Barlow. Systematic Errors: Facts and Fiction. MAN/HEP/2002/1 (2002). hep-ex/0207026. *cited in 7.6.4 (p. 113)*

[76] M. Gluck, E. Reya, M. Stratmann, and W. Vogelsang. Models for the polarized parton distributions of the nucleon. Phys. Rev. **D63** (2001) 094005. hep-ph/0011215. *cited in 7.7 (p. 115)*

[77] B. I. Abelev et al. Longitudinal double-spin asymmetry and cross section for inclusive jet production in polarized proton collisions at s**(1/2) = 200-GeV. Phys. Rev. Lett. **97** (2006) 252001. hep-ex/0608030. *cited in chapter 8 (p. 119)*

[78] A. Adare. Inclusive cross section and double helicity asymmetry for π^0 production in p+p collisions at sqrt(s)=200 GeV: Implications for the polarized gluon distribution in the proton. (2007). arXiv:0704.3599. *submitted to Phys. Rev. D.* *cited in chapter 8 (p. 119)*