CHARGE SYMMETRY VIOLATION QUARK DISTRIBUTIONS VIA PRECISE MEASUREMENT OF π^+/π^- RATIOS IN SEMI-INCLUSIVE DEEP INELASTIC SCATTERING

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

---

by
Shuo Jia
December 2022

Examining Committee Members:

Dr. Nikolaos Sparveris, Advisory Chair,Temple University Department of Physics

Dr. Zein-Eddine Meziani, Temple University Department of Physics/Argonne National Lab Medium Energy Physics

Dr. Martha Constantinou, Temple University Department of Physics

Dr. Mark Jones, Jefferson Lab Department of Physics

©

by

Shuo Jia

December 2022

All Rights Reserved

ABSTRACT

CHARGE SYMMETRY VIOLATION QUARK DISTRIBUTIONS VIA PRECISE MEASUREMENT OF π^+/π^- RATIOS IN SEMI-INCLUSIVE DEEP INELASTIC SCATTERING

Shuo Jia

Temple University, December 2022

Dr. Nikolaos Sparveris, Chair

Charge symmetry in the nucleon parton distributions assumes the distribution of quarks in the proton is related to those in the neutron. Indirect experimental evidence constrains Charge Symmetry Violation (CSV) to be less than 9%. In Quantum Chromo-Dynamics (QCD), charge symmetry is broken by the mass difference between the up and down quarks. CSV in the valence region can be extracted from precision measurements of the cross section ratio of charged pion production in semi-inclusive deep inelastic scattering (SIDIS) from deuterium. Such measurements were proposed and carried out in experiment E12-09-002 at Jefferson Lab. The experiment was conducted in Hall C from fall 2018 to spring 2019 using the upgraded 10.6 GeV incident electron beam. In this SIDIS experiment, charged pions are detected in coincidence with scattered electrons covering the four-momentum transfer of the virtual photon Q^2 from 4 to 5.5 GeV^2, the Bjorken variable x for 0.35-0.65, and the fraction of energy transfer carried by the outgoing pion z from 0.4-0.7.

To my parents.

Contents

ABSTRACT . . . . . **iii**
DEDICATION . . . . . **iv**
LIST OF FIGURES . . . . . **vii**
LIST OF TABLES . . . . . **xii**
1 INTRODUCTION **1**
1.1 Charge Symmetry and Breaking 1
1.2 Experimental Limits 10
2 FORMALISM **18**
2.1 Deep Inelastic Scattering 18
2.2 Structure Function and Quark-Parton Model 20
2.3 Semi-Inclusive Deep Inelastic Scattering 23
2.4 Fragmentation 24
2.5 General Formulas for Semi-Inclusive Reactions 25
2.6 Factorization at 12 GeV JLab 27
2.7 Experimental Extraction of Charge Symmetry Violation 29
3 EXPERIMENTAL APPARATUS **37**
3.1 Overview 37
3.2 CEBAF Accelerator and Hall C Beamline 38
3.3 Target Chamber 42
3.4 Hall C Spectrometer 43
3.4.1 Hall C High Momentum Spectrometer (HMS) 44
3.4.2 Hall C Super High Momentum Spectrometer (SHMS) 48
3.5 Hall C Detectors 49
3.5.1 Hall C Detector Drift Chamber 52
3.5.2 Hall C Detector Hodoscope 53
3.5.3 Hall C Cherenkov Detector for Particle Identification 54
3.5.4 Hall C Lead-glass Calorimeter 58
3.6 Data Acquisition 61
3.7 SIMC 62
4 DATA ANALYSIS **65**
4.1 Analysis Overview 65
4.2 Kinematics 66
4.3 Spectrometer Momentum Correction 67
4.4 Event Selection 72
4.4.1 Coincidence Time Cut 72
4.4.2 Accidental Backgrounds Subtraction 74
4.4.3 Aluminum Dummy Backgrounds Subtraction 76

4.5 Particle Identification . . . 77
4.5.1 Electron Detection Efficiency . . . 77
4.5.2 HMS Cherenkov Detector . . . 77
4.5.3 Hadron Detection Efficiency . . . 81
4.5.4 SHMS Aerogel Cherenkov Detector . . . 83
4.5.5 SHMS Heavy Gas Cherenkov Detector . . . 89
4.5.6 RFtime Offset . . . 91
4.5.7 RFtime PID . . . 94
4.6 Tracking Efficiency . . . 104
4.7 DAQ Live Time . . . 105
4.8 Target Density Correction . . . 108
4.9 Beam Current and Charge Accumulation . . . 109
4.10 Radiative Backgrounds . . . 110
4.11 Radiative Correction . . . 111
4.12 Data Quality Check . . . 114
4.13 Bin Centering Correction . . . 117
4.14 Systematic Uncertainties . . . 120
4.14.1 Current Uncertainty . . . 120
4.14.2 Target Boiling uncertainty . . . 121
4.14.3 FADC Dead Time Correction Uncertainty . . . 122
4.14.4 Acceptance Uncertainty . . . 123
4.14.5 PID Systematic Uncertainty . . . 124
4.14.6 Radiative Backgrounds Uncertainty . . . 126
5 RESULTS . . . 127
5.1 Factorization . . . 127
5.2 CSV . . . 129
5.2.1 Yield Ratio . . . 129
5.2.2 Extraction of CSV . . . 131
5.2.3 Formalism Including CSV in the Fragmentation Functions . 136
5.2.4 Discussion of results . . . 142
5.3 Conclusion . . . 142
REFERENCES . . . 144

LIST OF FIGURES

1.1 A recent result of λ_s and λ_v from some groups[57][89][68] as defined in Eq.1.13, Q is 2 GeV which is around the CSV experiment setting. 7

1.2 The ratio $R_c(x, Q^2)$ vs. x using CCFR neutrino data combined with muon DIS data. Solid circles are from NMC $\mu - D$ data[5]; open triangles are from BCDMS group[87];solid triangles are from SLAC electron scattering data[6] 11

1.3 Theoretical valence CSV PDFs from Sather. Solid curve: $x\delta d_v$; dashed curve: $x\delta u_v$. The valence PDFs were preliminary fits to CCFR Tevatron structure functions evolved to $Q^2 = 12.6GeV^2$ . . . 13

1.4 MRST valence isospin symmetry breaking PDF corresponding to best fit value $\kappa = -0.2$ 14

1.5 MRST shallow χ^2 minimum 15

1.6 MRST valence CSV $\delta d - \delta u$ first momentum corresponding to best fit value $\kappa = -0.2$ and the 90% confidence value $\kappa = -0.8$ and $\kappa = 0.65$ 15

1.7 The charge symmetry violation $x\delta u_v$ and $x\delta d_v$ valence quark distribution for $Q^2 = 4GeV^2$. Dash-dotted, dashed and solid curves represent pure QED, pure QCD, and the total contributions, respectively. 17

2.1 Deep-Inelastic scattering of a lepton with four-momentum k off a proton with four-momentum P. The scattered lepton has four-momentum k'. If one of the final states is detected in coincidence with the lepton, it is semi-inclusive deep-inelastic scattering. 19

2.2 Momentum parton distributions (x times parton distribution function) for experiment $Q^2 = 4.6$ GeV^2 setting and high $Q^2 = 10000$ GeV^2 setting 23

2.3 the semi-inclusive pion production process, spectrometer HMS is for electrons detection, SHMS is for hadrons. 24

2.4 Fragmentation Functions for π production, $\pi-$ for left and $\pi+$ for right 25

2.5 Cross section for $H, D(e, e'\pi^{\pm}$ from Hall C experiment E-00-108. Curves denote cross section calculation using simple factorization using the models. Closed symbols have been corrected for exclusive ρ production backgrounds. 29

2.6 Ratios of proton to deuterium results for the sum (top) and difference (bottom) of the π^+ and π^- cross sections. Closed symbols have been corrected for exclusive ρ production backgrounds. 30

3.1 Jefferson Lab CEBAF 38
3.2 Electrons get accelerated in the RF cavity. The frequency is 1497 MHz, producing three beams at the third sub-harmonic, 500 MHz.The RF separator can split all beams and send them to the respective experimental halls at the end.[38] 40
3.3 Four Laser configuration for electron injector.[38] 40
3.4 Each sine wave represents the 1500 MHz accelerating frequency. The top three are before the 12 GeV upgrade; each experimental hall gets a 500 MHz beam. The center and bottom beams are modified to make room for the Hall D beam. One of the high pass Hall (A or C) would reduce the frequency to 250 MHz 41
3.5 Cryogenic and solid target ladders. Loop 2 is liquid hydrogen. Loop 3 is liquid deuterium. The aluminum dummy and carbon optics targets are below liquid targets. The cryo-targets are assembled above the solid target ladder. The beam enters from the left in the left picture.[88] 42
3.6 This is the target screen graphic user interface when the beam is not on any targets. Moving the target ladder can move the desired target to the beam position. Liquid cryo-target states are monitored and can be changed by cooling and heating. [88] 43
3.7 Hall C setup 44
3.8 Hall C top view and coordinate system 45
3.9 HMS side view 47
3.10 SHMS side view 47
3.11 HMS detectors 50
3.12 SHMS detectors 51
3.13 Drift Chamber 51
3.14 Hodoscope planes sketch. 53
3.15 Hodoscope pruning process. The drift chamber will record the position and direction information for each track. Each point at the first hodoscope plane will be predicted to arrive at a specific area on the second hodoscope plane. Tracks outside the reasonable region will be pruned.[35] 53
3.16 Cherenkov effect. The charged particle (red) going faster than the velocity of the light in the medium will cause a flash of light (blue). This is similar to the sonic boom but for the light. 56
3.17 Gas Pressure vs. Particle Momentum. The black curve represents the suggested gas pressure (which leads to the reflection index) of the HGC detector for good π/κ separation. The colored curves represent the Cherenkov threshold pressure for different particles. C_4F_8O performs equivalently to C_4F_{10} in this context. [88] 57
3.18 SHMS HGC Detector 58
3.19 SHMS aerogel Cherenkov detector number of photon electrons versus momentum. 59

3.20 Sketch of the electromagnetic shower. Reproduced from wiki sketch. 60
3.21 Hall C Calorimeters 60

4.1 Kinematic coverage Q^2 vs. x for all three Q^2 68
4.2 Kinematic coverage for three distinct Q^2 region 68
4.3 Hydrogen elastic scattering. The incoming electron k($E_b, \overrightarrow{p}$) scatter off a proton target at rest P(M_p,0), resulting in the scattered electron $k'(E', \overrightarrow{p'})$ and proton $P'(E_p, \overrightarrow{p_p})$ 70
4.4 W peak shift for data and SIMC. A Gaussian fit is done to determine the peak center. 70
4.5 HMS momentum offset as a function of center momentum 71
4.6 Coincidence time peak for one fall run and one spring run with similar central momentum setting and same cuts. Spring runs have a better coincidence time resolution. 73
4.7 Coincidence time for two runs with the same kinematic setting but opposite polarity. Blue peaks on both sides are accidental background peaks after some PID cut which will be discussed in Section.4.4.2 74
4.8 The different selections of accidental backgrounds 75
4.9 Compare different accidental backgrounds selection 76
4.10 The number of photoelectrons generated in the HMS Cherenkov detector depends on different momentum 78
4.11 HMS calorimeter 80
4.12 HMS calorimeter electron efficiency as a function of HMS momentum 80
4.13 Gain coefficients for the shower block at row 8 and column 11 . . . 82
4.14 SHMS E/p for one fall RunGroup without any PID cut before calibration 84
4.15 SHMS calorimeter calibration E/p. The green areas are selected for the minimization process, and after the calibration, the electron peak is calibrated to the correct position at 1. 84
4.16 SHMS calorimeter pion detection efficiency 85
4.17 SHMS calorimeter pion efficiency as a function of SHMS momentum 85
4.18 Aerogel Number of photoelectron vs. rftime pid 86
4.19 Aerogel Number of photoelectrons vs. rftime PID for low SHMS central momentum 87
4.20 SHMS aerogel Cherenkov detector efficiency as a function of SHMS momentum for D_2 runs 88
4.21 SHMS aerogel Cherenkov detector efficiency as a function of SHMS momentum 88
4.22 SHMS HGC inefficiency at the geometry center due to the mirror crack 90
4.23 SHMS heavy gas Cherenkov detector inefficiency due to mirror misalignment. 91
4.24 RFtime offset to put pion peak at 1 ns for one setting. The kaons and protons are slower 92
4.25 RFtime offset to move the maximum bin to 1 93
4.26 Difference of time of flight for different particles in SHMS compared to electrons versus different momentum on the x-axis. 93

4.27 RFtime vs. Momentum 95
4.28 The function form used for the peaks 96
4.29 Pion peak selected by a HGC pion cut on rftime kaon mass corrected spectrum 97
4.30 Pion and kaon peaks selected by an anti-HGC cut on rftime kaon mass corrected spectrum 98
4.31 Kaon to pion ratio plots for one RunGroup with different kinematic cut 99
4.32 Heavy gas efficiency from the fitting. The efficiency for the spring runs is higher than for the fall runs. The heavy gas detector mirror got realigned between the two run period. 99
4.33 Kaon to pion ratio for one setting as a function of z for both polarity 100
4.34 Kaon decay probability as a function of SHMS path position 100
4.35 Kaon decay probability for different momentum 101
4.36 Kaon decay probability for different momentum 101
4.37 The higher the momentum, the more kaons remain in rftime cut . . 102
4.38 P_K for two rftime cut cases, one is fixed from 0.2 to 1.54, the other one is varying from 0.2 to the middle of pion peak and kaon peak . 102
4.39 The positive pion purity as a function of momentum for one (x, Q^2) setting 103
4.40 N_{should}, the number of expected triggers that should have a track . . 105
4.41 N_{did},found triggers 105
4.42 SHMS tracking efficiency as a function of rate for Fall runs 106
4.43 SHMS tracking efficiency as a function of rate for Spring runs . . . 106
4.44 HMS tracking efficiency as a function of the SHMS 3/4 event rate . 106
4.45 The total live time during the experiment as a function of the SHMS 3/4 rate 107
4.46 Beam current for one fall run. There is one beam trip during the run 109
4.47 Radiative backgrounds 110
4.48 Diffractive ρ backgrounds 111
4.49 First order bremsstrahlung 112
4.50 High-order internal radiative correction 113
4.51 Electron straggling in the target 113
4.52 HMS quadruple Q1 current ratio for negative runs to positive runs. The average is plotted with standard derivation. The maximum, minimum, and first value ratios from each run and the mode are also tested. 115
4.53 Liquid D_2 target pressure and temperature check. Black circles are for negative runs, red circles are for positive runs. 115
4.54 Liquid H_2 target pressure and temperature check. Black circles are for negative runs, red circles are for positive runs. 116
4.55 Liquid D_2 and H_2 density 116
4.56 Yield consistency check between each runs in the same kinematic run groups 117
4.57 Yield ratio for one (x, Q^2) setting. Lines are yield ratios from SIMC simulation. 118
4.58 Sketch of bin center correction in the 1D case 119
4.59 Current ratio uncertainty vs. RunGroup 121

4.60 Target boiling correction ratio uncertainty in percentage vs. RunGroup . . . 121
4.61 FADC dead time correction . . . 122
4.62 FADC dead time correction ratio uncertainty vs. RunGroup . . . 123
4.63 Yield ratio from data (in points) compared with yield ratio from SIMC (in line) . . . 124
4.64 Ratio of RY ($N_D^{\pi^-}/N_D^{\pi^+}$) from data divided by yield ratio from SIMC for the SHMS and HMS . . . 124
4.65 Yield ratio RY compared with the two cases for one setting. The star is "with heavy gas cut", the circle is "no heavy gas cut, pion purity corrected" and the triangle is "no heavy gas cut, no purity corrected". . . . 125
4.66 The difference of R_Y between the two cases . . . 126

5.1 The difference ratio results of proton to deuterium pion scattering yield. . . . 128
5.2 The sum ratio results of proton to deuterium pion scattering yield. . 128
5.3 R^D_{meas} for the first Q^2 setting . . . 130
5.4 R^D_{meas} for the second Q^2 setting . . . 130
5.5 R^D_{meas} for the third Q^2 setting . . . 131
5.6 $A(x)$ term calculated from JAM20 PDFs[57], nnPDF[8] and CTEQ PDFs[89] . . . 132
5.7 $B(x, z)$ terms calculated from CTEQ6 PDFs[68] and JAM20 fragmentation functions[57] . . . 133
5.8 $B(x, z)$ calculated from JAM20 PDFs and fragmentation functions. 133
5.9 Result of unfavored over favored fragmentation function $\Delta(z)$ in comparison with some existing fragmentation functions . . . 135
5.10 The CSV results from fitting. The blue line and 1σ blue error band is from the fitting process. The points are the CSV results calculated from the fragmentation function ratio result for each Q^2 setting . . . 135
5.11 Residual for fitting, with the fitting parameters . . . 136
5.12 The calculated results of CSV from different PDF functions . . . 137
5.13 FFs from the DSS group . . . 137
5.14 CSV calculated from the yield ratio equation using JAM20 PDFs and FFs. The radiative ρ background is subtracted from data with coefficient γ . . . 140
5.15 CSV calculated from the yield ratio equation using JAM20 PDFs and DSS leading order FFs. The radiative ρ background is subtracted from data with coefficient γ . . . 141
5.16 The CSV results from all different cases . . . 141

LIST OF TABLES

3.1 Cryo-target states during normal operating conditions. [88] 43
3.2 HMS performance 46
3.3 SHMS performance 49
3.4 HMS calorimeter performance 61
3.5 SHMS calorimeter performance [55] 61

4.1 One fall RunGroup 67
4.2 One spring RunGroup 67
4.3 Hydrogen elastic runs are taken during the experiment 69
4.4 Target Entrance and Exit window wall thickness[49] 76
4.5 Systematic uncertainty 119
4.6 Systematic uncertainty 120
4.7 current uncertainty example 120

Chapter 1

INTRODUCTION

1.1 Charge Symmetry and Breaking

In a nuclear system, isospin (IS) and charge (CS) symmetries are introduced to describe the underlying physics of matter and forces. Meanwhile, Quantum chromodynamics (QCD) is the theory to describe the strong interaction between fundamental constituents of matter, quarks and gluons. Invariance under rotation in isospin space in all directions implies isospin symmetry, such that the QCD Hamiltonian commutes with the isospin operator(T), i.e.

$$[H, T] = [H, T^2] = 0$$

CS is defined as a specific isospin rotation. The invariance is with respect to a rotation of 180° about the y-axis, where the charge corresponds to the third axis[51][16].

$$[H, P_{CS}] = 0, P_{CS} = exp(i\pi T_2)$$

Consequently, isospin symmetry necessarily implies the validity of charge symme-

try; however, the converse is not necessarily true. As charge symmetry is a more restrictive symmetry than isospin symmetry, it is generally conserved in strong interactions to a greater degree than isospin symmetry. Thus, while in many nuclear reactions, isospin symmetry is violated at the few percent level, in most cases, charge symmetry is obeyed to better than one percent[37].

At nuclei level, the CS operator interchanges neutrons and protons. The proton and neutron masses are equal to about 0.1%. The binding energies of tritium ^3H and ^3He are equal to about 1%, after including the Coulomb corrections[65]. The scattering length of the interaction between neutron-neutron and proton-proton is similar to 1%[30]. Also, the mirror interaction in isospin space have the similar cross-section, e.g., $\sigma(n,^3He) \simeq \sigma(p,^3H)$[59].

At the quark level, CS turns a u quark into a d quark and protons into neutrons,

$$P_{CS}\left|d\right\rangle = \left|u\right\rangle, P_{CS}\left|u\right\rangle = -\left|d\right\rangle$$

CS indicates the invariance of a system under the interchange of up and down quarks while simultaneously interchanging protons and neutrons, i.e. Eq. (1.1) [52].

$$\begin{aligned} u^p(x,Q^2) &= d^n(x,Q^2) \\ d^p(x,Q^2) &= u^n(x,Q^2) \\ s^p(x,Q^2) &= s^n(x,Q^2) \equiv s(x,Q^2) \\ c^p(x,Q^2) &= c^n(x,Q^2) \equiv c(x,Q^2) \end{aligned} \tag{1.1}$$

At the current stage, the degree of charge symmetry violation (CSV) in valence quark distribution is unknown, however, as it is expected to be relatively small [37], CS has been universally assumed in the global analysis and extraction of PDFs. Nearly all extractions of PDFs from fits to data impose CS, which reduces

the number of free parameters. The validity of charge symmetry is required for sum rules and many relations between the structure functions[27][28][3].

However, CS is broken by a few percent at the quark level. QCD derives that CSV is related to the mass difference and the electromagnetic interactions. The latter plays a minor role at a high energy level, but the mass difference is an interesting part of CSV according to QCD. The CSV is estimated to be of the order of the up-down quark mass difference divided by some average mass expectation value of the strong Halmitonian,[53][30]

$$\frac{(m_d - m_u)}{\langle M \rangle}$$

where $\langle M \rangle$ has a value roughly 0.5-1.0 GeV. This would put the CSV effect at 1% or smaller [37].

Precision measurments of CSV in single-pion production in few-body systems [62][78] brought out the CSV at low energies[54][60][85][31]. The success in low energy CSV draw attention to the CSV at high energy range, the partonic part.

Charge-symmetry violation in the parton distributions is defined for the up and down quarks as [43]

$$\begin{aligned} \delta d(x) &\equiv d^p(x) - u^n(x), \\ \delta u(x) &\equiv u^p(x) - d^n(x), \\ CSV(x) &\equiv \delta d(x) - \delta u(x)\,. \end{aligned} \tag{1.2}$$

with similar definitions for anti-quark distributions. The strange quark distributions are the same in both the proton and neutron, as well as the antistrange distributions.It's beneficial to introduce the valence and sea quark distributions,

$$u_v(x) = u(x) - \overline{u}(x), d_v(x) = d(x) - \overline{d}(x) \tag{1.3}$$

For valence quark distributions, quark flavor number conseravations yields the quark normalization conditions:

$$\begin{aligned} \frac{1}{2}\int_0^1 u_v^p(x)dx = \frac{1}{2}\int_0^1 d_v^n(x)dx = \\ \int_0^1 d_v^p(x)dx = \int_0^1 u_v^n(x)dx = 1, \\ \int_0^1 \overline{s}(x)dx = \int_0^1 \overline{c}(x)dx = 0 \end{aligned} \tag{1.4}$$

The CSV in Eq1.2 can be decomposed into valence and sea pieces,

$$\begin{aligned} \delta u_v(x) &= u_v^n(x) - d_v^n(x), \\ \delta d_v(x) &= d_v^p(x) - u_v^n(x), \\ \delta \overline{u}(x) &= \overline{u}^n(x) - \overline{d}^n(x), \\ \delta \overline{d}(x) &= \overline{d}^p(x) - \overline{u}^n(x) \end{aligned} \tag{1.5}$$

From the constrain of valence quark normalization (Eq. 1.4) and the definition of CSV (Eq. 1.2), we arrive at the constraints

$$\int_0^1 \delta u_v(x)dx = \int_0^1 \delta d_v(x)dx = 0 \tag{1.6}$$

which leads to

$$\int_0^1 \delta d(x)dx = \int_0^1 \delta \overline{d}(x)dx, \int_0^1 \delta u(x)dx = \int_0^1 \delta \overline{u}(x)dx \tag{1.7}$$

and

$$\int_0^1 CSV(x)dx = 0\,. \tag{1.8}$$

Charge symmetry is widely assumed in parton distribution functions (PDFs), but it is so common that this assumption is not always mentioned. However, there are no direct measurements of the contributions from CSV. Nevertheless, the apparent violation of SU(2) flavor symmetry in the nucleon draws attention to CSV. The Gottfried sum rule[27] written in terms of the proton and neutron

structure functions F_2 is defined as $S_G = \int_0^1 (F_2^p - F_2^n)dx/x$. It can be expressed in the quark-parton model in terms of the quark momentum distributions $q_i(x)$:

$$S_G = \int_0^1 \sum_i e_i^2 [q_i^p(x) + \overline{q}_i^p - q_i^n(x) - \overline{q}_i^n(x)]dx \tag{1.9}$$

where e_i is the charge of a quark of flavor i. The integral represents the difference between the sum of the squares of the quark charges in the proton and neutron. Assuming charge symmetry and valence quark normalization condition1.4, the S_G becomes:

$$S_G = \frac{1}{3} + \int_0^1 \sum_i e_i^2 [2\overline{q}_i^p(x) - 2\overline{q}_i^n(x)]dx \tag{1.10}$$

The assumption of isospin symmetry between the proton and the neutron ($\overline{u} \equiv \overline{u}^p = \overline{d}^n$ and $\overline{d} \equiv \overline{d}^p = \overline{u}^n$) and the flavor-symmetric sea ($\overline{u} = \overline{d}$) the sea quark term in Eq.1.10 vanishes and the expected result is 1/3. The value of S_G is determined from $F_2^p - F_2^n$ which can be expressed as $F_2^p - F_2^n = 2F_2^d(1 - F_2^n/F_2^p)/(1 + F_2^n/F_2^p)$. The ratio $F_2^n/F_2^p \equiv 2F_2^d/F_2^p - 1$ is determined by the ratio of deuteron/proton measured cross-section. The New Muon Collaboration (NMC)[63] is 0.240 ± 0.016, significantly below the simple quark-parton model result of 1/3.

The violation of the Gottfried sum rule has an alternative explanation, that of CSV in the proton sea. Ellis and Stirling noticed [22] an elegant way of observing the CSV in proton sea by the Drell-Yan process. The proton-nucleon collisions, $pN \longrightarrow l^+l^- + X$ cross section asymmetry is defined as

$$A_{DY} = \frac{\sigma^{pp} - \sigma^{pn}}{\sigma^{pp} + \sigma^{pn}} \tag{1.11}$$

where σ^{pp} and σ^{pn} are the cross-sections for dileptons produced in p-p and p-n collisions. The asymmetry can be written in terms of valence and sea quark

structure functions in the Drell-Yan model

$$A_{DY} = \frac{(4u_v - d_v)(\overline{u} - \overline{d}) + (u_v - d_v)(4\overline{u} - \overline{d})}{(4u_v + d_v)(\overline{u} + \overline{d}) + (u_v + d_v)(4\overline{u} + \overline{d})} \tag{1.12}$$

By substitution of up quark to down quark ratios for the valence and sea parton distribution, $\lambda_v(x) = u_v(x)/d_v(x)$ and $\lambda_s(x) = \overline{u}(x)/\overline{d}(x)$, it becomes

$$A_{DY} = \frac{(4\lambda_v - 1)(\lambda_s - 1) + (\lambda_v - 1)(4\lambda_s - 1)}{(4\lambda_v + 1)(\lambda_s + 1) + (\lambda_v + 1)(4\lambda_s + 1)} \tag{1.13}$$

In the NA51 collaboration at the CERN [7] region, typical values for λ_v are slightly higher than 2. Sea isospin symmetry means $\lambda_s = 1$, for $\lambda_v = 2$, leads to $A_{DY} = 0.09$. The NA51 experiment measures cross-sections for proton beam scattered on proton and deuteron targets. The A_{DY} can be measured from the ratio of Drell-Yan cross-sections

$$A_{DY} = 2\frac{\sigma^{pp}}{\sigma^{pd}} - 1 \tag{1.14}$$

with the assumption $\sigma^{pd} = \sigma^{pp} + \sigma^{pn}$. The experimental result on A_{DY} is four standard deviations away from the values obtained with an isospin symmetric sea. Values of λ_v are taken within a range $[2, 2.7]$ and induce an additional systematic uncertainty on λ_s. The measured value of λ_s at $x = 0.18$ is

$$\lambda_s = 0.51 \pm 0.04(stat) \pm 0.05(syst) \tag{1.15}$$

This result suggests that isospin symmetry is violated in the light quark sea of the nucleon. Some recent results from several groups are shown in Fig.1.1

Fermilab Experiment 866 measured the Drell-Yan muon pair yield from 800 GeV/c proton bombardment of liquid deuterium and hydrogen targets. The $\overline{d}/\overline{u}$ and $\overline{d}-\overline{u}$ in the proton over the range $0.02 < x < 0.345$ are extracted. A significant difference between the $\overline{d}$ and $\overline{u}$ distribution is found. [29] The results from the

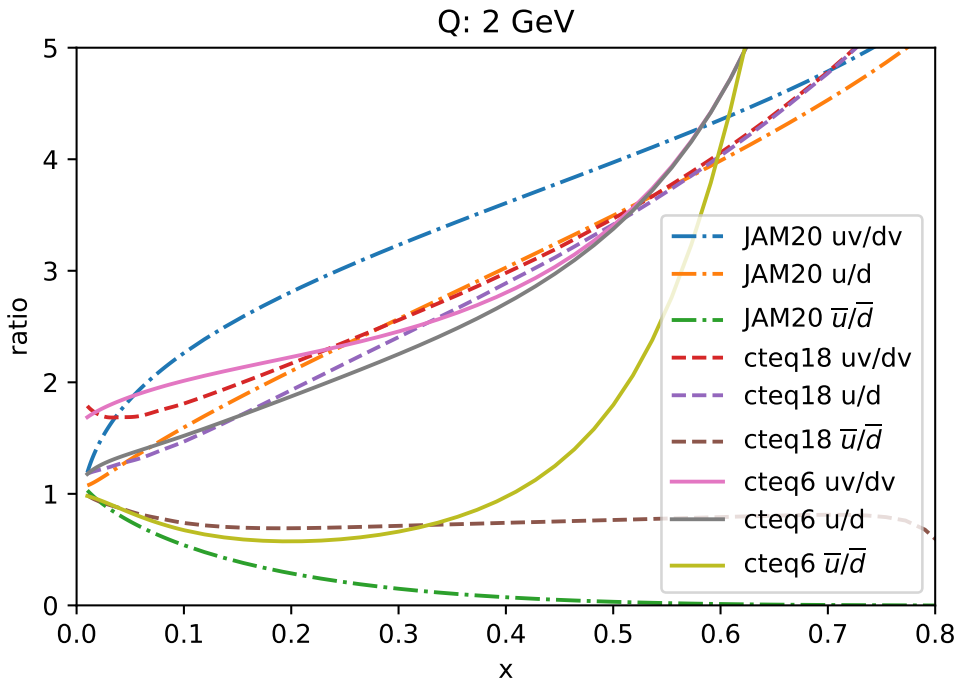


Figure 1.1: A recent result of λ_s and λ_v from some groups[57][89][68] as defined in Eq.1.13, Q is 2 GeV which is around the CSV experiment setting.

NA51 collaboration at CERN and the Drell-Yan experiment at FNAL do support the existence of flavor asymmetry in the sea quark region. However, this could also in principle be explained even if flavor symmetry were conserved if we assume CSV in the nucleon sea.

The valence quark CSV is also critical since it makes a more significant contribution than the sea quark CSV in the extraction of the Weinberg angle from neutrino Deep Inelastic Scattering (DIS). The "Nutev anomaly", which is from the NuTeV collaboration[91], can be explained by the assumption of CSV in the valence quark region. The NuTeV collaboration uses high purity neutral current (NC) ν and charged current (CC) $\overline{\nu}$ beam produced from the decay of pions and kaons provided by Fermilab. Neutrino interactions were observed in the NuTeV detector with the criterion that the visible energy in the calorimeter satisfied the relation $20 < E_{vis} < 180 GeV$, and the event vertex was contained in the calorime-

ter. After the selection, the results data sample of 1.62 million ν and 0.351 million $\overline{\nu}$ events were used to extract the independent value for the weak mixing angle, or Weinberg angle in terms of $\sin^2\theta_W$, using a procedure 1.16 suggested initially by Paschos and Wolfenstein[64].

$$R^{PW} \equiv \frac{\langle\sigma_{NC}^{\nu N_0}\rangle - \langle\sigma_{NC}^{\overline{\nu} N_0}\rangle}{\rho_0^2(\langle\sigma_{CC}^{\nu} N_0\rangle - \langle\sigma_{CC}^{\overline{\nu u} N_0}\rangle)} = \frac{1}{2} - sin^2\theta_W = \frac{R^{\nu} - R^{\overline{\nu}}}{1 - rR^{\nu}} \tag{1.16}$$

In Eq.1.16, $\langle\sigma_{NC}^{\nu N_0}\rangle$ is the total neutral current inclusive cross-section for the neutrinos on an isoscalar target. The quantity $\rho_0 \equiv M_W/(M_z cos\theta_W)$ is one in the Standard Model. The Paschos-Wolfenstein (PW) ratio can be written in terms of ratios of NC to CC ratio of the total cross-section on an isoscalar target induced by neutrinos R^{ν}, by antineutrinos $R^{\overline{\nu}}$ and the ratio of CC cross sections for antineutrinos to CC cross section induced by neutrinos r.

The result from NuTeV group is $\sin^2\theta_W = 0.2277 \pm 0.0013(Stat) \pm 0.0009(Syst)$ which is surprisingly three sigma away from the measured value for the Weinberg angle obtained from electroweak (EW) processes near the Z pole, $sin^2\theta_W = 0.2227 \pm 0.00037$, the NuTeV anomaly. There are no universally accepted complete explanations yet.

The derivation of PW ration requires lots of assumptions: an isoscalar target; neglects contributions from quark masses; assumes isospin symmetry in the parton distribution functions; neglects nuclear effects in PDFs; and neglects any contributions outside the Standard Model[41]. It is quite difficult to find new particles beyond Standard Model that contribute to the NuTeV neutrino measurement but do not change other measured quantities from their experimental limits. It's probably wise to consider Standard Model explanations. The correction to the Weinberg angle because iron is not an isoscalar target has already

been considered by the NuTeV group and it's well under control. The radiative correction for the charge current events used by the NuTeV group for the estimate of uncertainty might be too optimistic.

Another contribution comes from isospin violation, the charge symmetry violation in parton distributions. The correction to the PW ratios from parton charge symmetry violation ($\delta d_v(x) \equiv d_v^p(x) - u_v^n(x), \delta u_v \equiv u_v^p - d_v^n(x)$) has the form:

$$\begin{aligned} \Delta R_{CSV}^{PW} &\approx [1 - \frac{7}{3} sin^2\theta_W] \frac{\delta U_v - \delta D_v}{2(U_v + D_v)}, \\ Q_v &\equiv \int_0^1 x[q(x) - \overline{q}(x)]dx \end{aligned} \tag{1.17}$$

The small QCD radiative correction terms are simplified to be proportional to α_s, and only the valence PDFs contribution remains.

The phenomenological fits of PDF from the MRST group[46] include sea quark CSV and valence quark CSV. The valence quark CSV makes a substantially larger contribution to the extraction of the Weinberg angle from neutrino DIS. Using the sea-quark CSV and the best-fit value for valence quark CSV obtained by the MRST group would remove roughly 1/3 of the NuTeV anomaly. The CSV estimated value within the 90% confidence limit from MRST would completely remove the NuTeV anomaly. Compared with the phenomenological fits of MRST, the theoretical estimates of valence PDF CSV by Sather[70] and Rodionov *etal.*[69] both reduce the size of the NuTeV anomaly.

To extract the CSV effect in the valence quark distribution, we measured SIDIS with a charged pion in the Super High Momentum Spectrometer (SHMS) in coincidence with a scattered electron in the High Momentum Spectrometer (HMS). In order to determine the CSV effects, we evaluate the ratio of two charged pion polarity yields by this SIDIS measurement. Although the absolute yield is

not required, the precision of two spectrometers assured the accuracy of our ratio. This experiment has several Q^2 (4,4.7,5.5GeV^2) settings; for the range of x is $0.35 - 0.5$, $0.45 - 0.6$ and $0.5 - 0.65$ respectively, and z from 0.4 to 0.7. These kinematic settings were not accessible to be measured in the previous experiments; however, the precision and sufficient range make it possible to probe this small component of quark distribution functions. The experiment can not only constrain the limit of CSV on quark distribution but also enrich the data set for Thomas Jefferson National Lab (JLab) Hall C SIDIS program.

1.2 Experimental Limits

Without direct measurements that reveal the presence of CSV in parton distribution functions, there are only some upper limits on the magnitude of CSV.

The precise neutrino measurements of the CCFR collaboration[74] and muon interactions of NMC collaboration[37] give an upper limit of 9% at the region of $0.1 < x < 0.4$.The limits arise from comparing the structure function F_2^{ν} measured in neutrino induced charged current reactions and the structure function F_x^{γ} for charged lepton DIS, both measurements on isoscalar targets. In leading order, assuming parton charge symmetry, the structure functions have the form [37]

$$\begin{aligned} F_2^{\gamma N_0}(x) &\approx \frac{5Q(x)}{18} + \frac{x}{6}(c(x) + \overline{c}(x) - s(x) - \overline{s}(x)), \\ F_2^{W^{\pm} N_0}(x) &\approx Q(x) \pm x(s(x) - \overline{s}(x) + \overline{c}(x) - c(x)), \\ Q(x) &= \sum_{j=u,d,s,c} x(q_j(x) + \overline{q}_j(x)) \end{aligned} \tag{1.18}$$

where $F_2^{W^{\pm} N_0}$ is the structure function for charged current reaction produced by neutrinos. The relation between the charged-lepton and neutrino structure func-

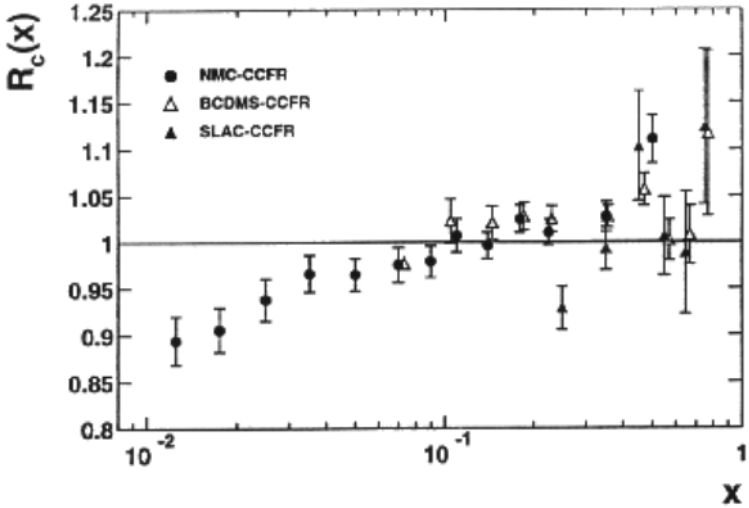


Figure 1.2: The ratio $R_c(x, Q^2)$ vs. x using CCFR neutrino data combined with muon DIS data. Solid circles are from NMC $\mu - D$ data[5]; open triangles are from BCDMS group[87];solid triangles are from SLAC electron scattering data[6]

tion is defined as $R_c(x, Q^2)$[13][82]:

$$\begin{aligned} R_c(x) &\equiv \frac{F_v^{\gamma N_0}(x) + x(s(x) + \overline{s}(x) - c(x) - \overline{c}(x))/6}{5\overline{F_2}^{WN_0}(x)/18}, \\ &\approx 1 + \frac{3x(\delta u(x) + \delta\overline{u}(x) - \delta d(x) + \delta\overline{d}(x))}{10Q(x)} \end{aligned} \tag{1.19}$$

The quality $\overline{F}_2^{WN_0}$ is the average of neutrino and antineutrino, F_2, charge changing structure functions. The strange and charm PDFs are known. It includes the CSV parton distribution $\delta u(x)$ and $\delta d(x)$. The neutrino measurements are from the CCFR group[74], which extracted the F_2 structure function for neutrino and antineutrino interactions using the Quadrupole Triplet Beam at FNAL. The F_2 structure functions from DIS reactions using high-energy muons or electrons measured by the NMC group[5], BCDMS muon scattering experiments,[87] and electron scattering results from SLAC [6]. The ratio R_c is plotted in Fig.1.2 vs. x. In the region $0.1 < x < 0.4$, the charge ratio R_c test is consistent with unity, with errors in the range $2 - 3\%$. Solving the CSV from R_c, one can set an upper limit

to the parton CSV effects in this x range at about $6-9\%$ level. For large values of x the upper limit on CSV effects is substantially greater. This is due to both the poorer statistics and the large Fermi motion corrections needed for the heavy target iron used for the CCFR group.

Theoretical estimates of parton CSV for valence quarks can be obtained by examining simple models for parton distribution functions, then observing how these change upon applying the operations of charge symmetry. Some theorists, Sather[25] and Rodionov et al[69], investigated valence quark distributions for the proton and neutron in a quark model and in the MIT bag model respectively.

Sather[70] assumed the CSV was from the neutron-proton mass difference $\delta M \equiv M_n - M_p$ and the differences in diquark masses. In a valence quark approximation, the nucleon is constructed by three valence quarks,

$$|p\rangle \sim [uud], |n\rangle \sim [udd] \tag{1.20}$$

The Adelaide group[75] developed a method for calculating twist-two valence parton distributions through the relation

$$q_v(x,\mu^2) = M \sum_X |\langle X| \frac{1+\alpha_3}{2} \phi(0) |N\rangle|^2 \times \delta(M(1-x) - p_X^+) \tag{1.21}$$

where $\alpha_3 = \gamma^0\gamma^3$. This equation denotes that one quark is removed from three valence quarks, leaving a residual diquark.

There are several potential sources of CSV in Eq. 1.21. The CSV comes from electromagnetic effects is at the order of electromagnetic coupling constant α. This is tiny compared with the other sources like the mass difference. One is the $n-p$ mass difference $\delta M \equiv M_n - M_p$, this is the result of all electromagnetic effects including the $u-d$ quark mass difference. A second effect is the difference in

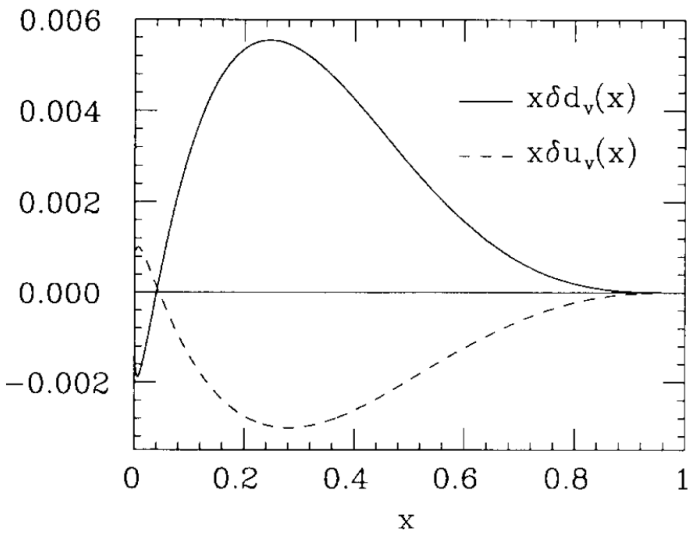


Figure 1.3: Theoretical valence CSV PDFs from Sather. Solid curve: $x\delta d_v$; dashed curve: $x\delta u_v$. The valence PDFs were preliminary fits to CCFR Tevatron structure functions evolved to $Q^2 = 12.6GeV^2$

diquark masses $\delta\tilde{m} = m_{dd} - m_{uu}$. With the variational method, the CSV can be written:

$$\delta q \approx \frac{\partial q_v}{\partial \delta\tilde{m}}\delta\tilde{m} + \frac{\partial q_v}{\partial M}\delta M \tag{1.22}$$

The valence quark CSV is related to the derivative of the valence PDFs,

$$\begin{aligned}
\delta d_v(x) &\equiv d_v^p(x) - u_v^n(x) \\
&= -\frac{\delta M}{M}\frac{d}{dx}[xd_v(x)] - \frac{\delta m}{M}\frac{d}{dx}d_v(x) \\
\delta u_v(x) &\equiv u_v^p(x) - d_v^n(x) \\
&= \frac{\delta M}{M}(-\frac{d}{dx}[xu_v(x)] + \frac{d}{dx}u_v(x))
\end{aligned} \tag{1.23}$$

where M is the n-p mass difference, $\delta M = 1.3MeV$, and $\delta m = m_{dd} - m_{uu} \sim 4MeV$ is the down-up quark mass difference. The numerical results showed $u^p(x)$ and $d^n(x)$ satisfied CS to within 1% while $d^p(x)$ and $u^n(x)$ were predicted to violate charge symmetry by 2-3% as in Fig.1.3.

According to the calculation of Rodionov[69], the CSV term for large x region can go up to 5%. He explicitly included the effect of quark transverse momentum

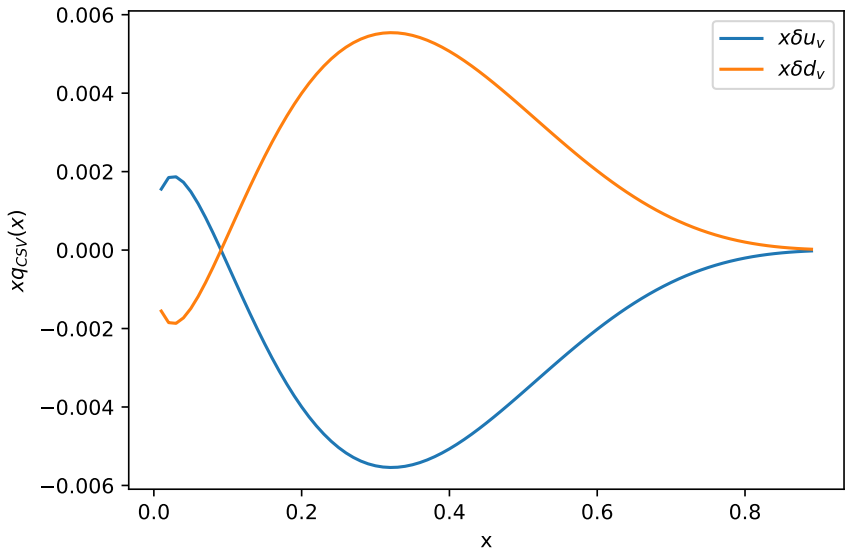


Figure 1.4: MRST valence isospin symmetry breaking PDF corresponding to best fit value $\kappa = -0.2$

within the MIT bag model.

Because CSV effects are typically very small at nuclear physics energy scales, all previous phenomenological parton distribution functions have assumed the validity of parton charge symmetry. However, the MRST group[46] has included isospin violation in a phenomenological evaluation of PDFs. They used a wide range of high-energy data to get a global fit of PDFs. The model for valence quark charge symmetry violating PDFs is chosen to be:

$$\begin{aligned} \delta u_v(x) = -\delta d_v(x) = \kappa f(x), \\ f(x) = (1-x)^4 x^{-0.5}(x - 0.0909) \end{aligned} \tag{1.24}$$

which is plotted in Fig.1.4. The integral over x of the valence CSV PDFs $\delta d_v \equiv d_v^p(x) - u_v^n(x)$ and $\delta u_v \equiv u_v^p(x) - d_v^n(x)$ must be zero, which is the quark normalization condition. At both small and large x, f(x) has the same form as the MRST valence quark distributions, and the first momentum of f(x) is zero. The coefficient κ was varied in a global fit to high energy data. The best fit is obtained at $\kappa = -0.2$ as in Fig.1.5. It corresponds to at maximum about a 3% violation of

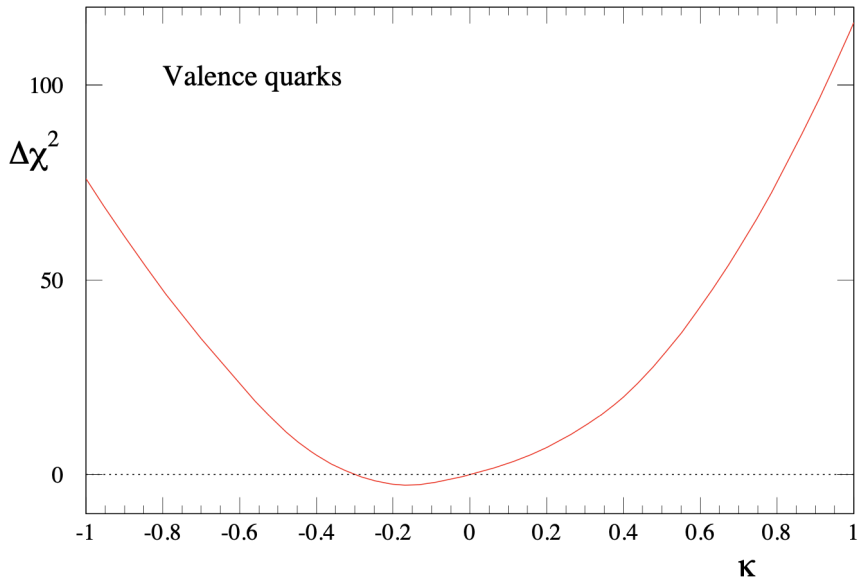
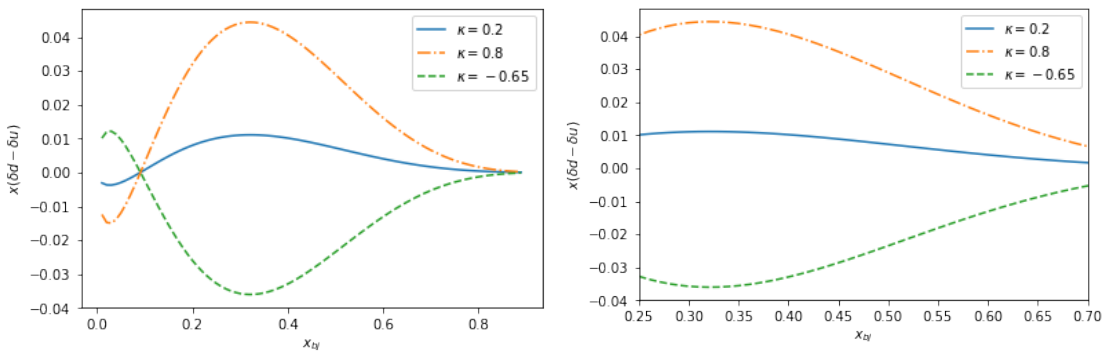


Figure 1.5: MRST shallow χ^2 minimum



(a) large x_{bj} range

(b) experiment x_{bj} range

Figure 1.6: MRST valence CSV $\delta d - \delta u$ first momentum corresponding to best fit value $\kappa = -0.2$ and the 90% confidence value $\kappa = -0.8$ and $\kappa = 0.65$

isospin for $u_v^n(x)$.

The CSV defined in this experiment is $\delta d - \delta u$, plotted from the MRST models is in Fig.1.6a. The 90% confidence level is obtained for $\kappa = -0.8$ or $\kappa = 0.65$. The experiment x ranges from 0.3 to 0.7, the prediction of CSV from MRST models for the experiment range is in Fig.1.6b

The calculation of CSV by photon radiation and quark mass difference suggests that the CSV is 20% larger than the value based on quark mass differences alone at a scale of 5 GeV^2[86]. The calculations showed CSV violating effects as large

as 5% at large-x, while they were at the percent level in the momentum fractions:

$$\delta U = \int_0^1 dx x \delta u(x), \delta D = \int_0^1 dx x \delta d(x) \tag{1.25}$$

In the lattice calculation, the dynamical parton distributions generated radiatively from valence-like inputs at low scales are determined from the global fit by GRV[26], taking into account small-x data on deep inelastic and other hard scattering processes. The QED radiation on DIS processes is important, which is added to the valence quark charge asymmetries. The DGLAP evolution equations are modified by photon PDF. At the low scale Q_0^2, the major part of the input photon distribution " comes from the coherent emission of the photon from the elastic proton"[47]. Both QCD and QED radiation is treated as incoherent radiation from the quarks in the evolution.

The pure QED contributes to the charge symmetry violation in both δu_v and δu_v. At $Q^2 = 4GeV^2$, the QED contribution to the second momentum is -0.00099 for δU_v and -0.00009 for δD_v. The QCD contributions to the second momentums are derived by extrapolating the first lattice simulations to the physical point:

$$\delta U_v = 0.0023(7), \delta D_v = 0.0017(4) \tag{1.26}$$

where the number in brackets indicates the error in the last significant figure.

Combining the QED and QCD contributions, the total isospin violation distributions at $Q^2 = 4GeV^2$ are shown in Fig.1.7. The quantity is comparable with the MRST prediction.

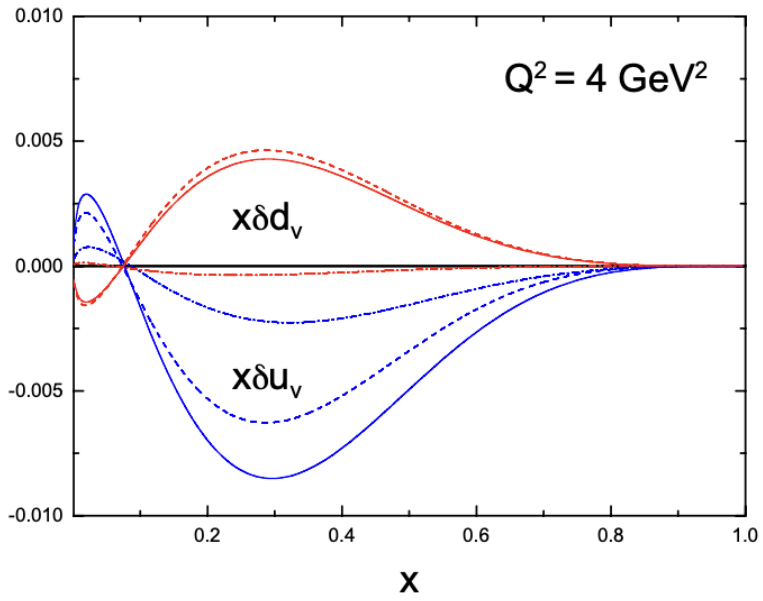


Figure 1.7: The charge symmetry violation $x\delta u_v$ and $x\delta d_v$ valence quark distribution for $Q^2 = 4GeV^2$. Dash-dotted, dashed and solid curves represent pure QED, pure QCD, and the total contributions, respectively.

Chapter 2

FORMALISM

2.1 Deep Inelastic Scattering

Deep Inelastic Scattering is a scattering process used to probe the insides of hadrons. In the lowest order deep inelastic electron-proton scattering, a single photon is exchanged between the electron and the proton. A lepton is scattered on a hadron and breaks the hadron into several pieces. The lepton in the final state is detected with four-momentum k'. If the final state hadron is detected in coincidence with the scattered lepton, it is semi-inclusive DIS (SIDIS).

$$l + N \hookrightarrow l' + X \tag{2.1}$$

A lepton l with four-momentum $k(\mathrm{E}, \overrightarrow{k})$ is scattered on a hadron N with four-momentum $P(E_T, \overrightarrow{p})$. The scattered lepton l' with four-momentum $k'(E', \overrightarrow{k'})$ is detected.

There are two independent variables that are Lorentz invariant to describe the kinematics of inclusive DIS. The momentum transfer to the nucleon, defined as $Q^2 = -q^s$, is directly related to the resolution of the lepton probe. The q is four

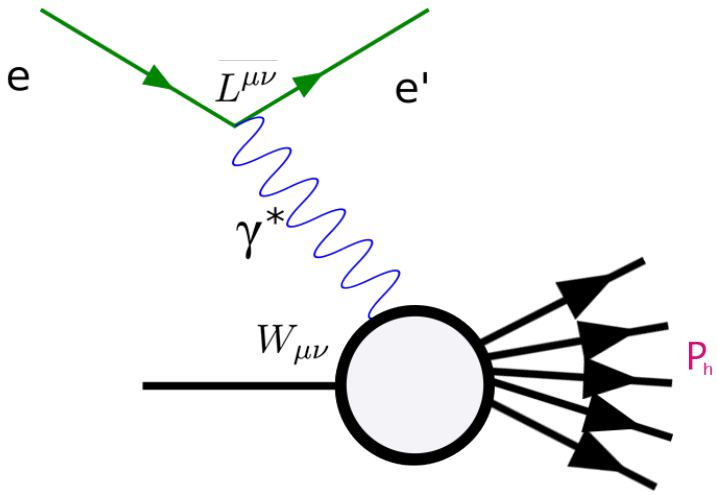


Figure 2.1: Deep-Inelastic scattering of a lepton with four-momentum k off a proton with four-momentum P. The scattered lepton has four-momentum k'. If one of the final states is detected in coincidence with the lepton, it is semi-inclusive deep-inelastic scattering.

momentum vector of the exchanged gauge boson $q = k - k' = (E - E', \overrightarrow{k} - \overrightarrow{k'})$. For $q^2 = 0$, the photon is real. For $q^2 < 0$, the exchanged photon is an off mass shell virtual photon. It carries some extra energy for a small period of time. The minus sign for Q^2 insures it is positive.

$$Q^2 = -q^2 = 2(EE' - \overrightarrow{k} \cdot \overrightarrow{k'}) \tag{2.2}$$

And the other independent variable is Bjorken scaling variable x, a measurement of the momentum fraction of the stuck parton that is hit compared to the whole neutron[12]:

$$x_{Bj} \equiv \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2M\nu} \tag{2.3}$$

where ν is the Energy loss of the lepton in the nucleon rest frame.

$$\nu \equiv \frac{P \cdot q}{M} = E - E' \tag{2.4}$$

A related invariant term, called the fractional energy loss is defined as:

$$y \equiv \frac{P \cdot q}{q \cdot k} = \frac{E - E'}{E} \tag{2.5}$$

Another important variable is the invariant mass squared of the hadronic final state:

$$W^2 = (P + k - k')^2 = (P + q)^2 = M^2 + 2M\nu + q^2 \tag{2.6}$$

For elastic scattering, $W^2 = M^2$, and Bjorken scaling variable $x_{bj} = 1$; for inelastic scattering, the x_{bj} is between $0 < x_{bj} < 1$, and the invariant mass squared is higher than M^2, since it get higher energy to form more final hadronic states,

$$W^2 = M^2 + \frac{1-x}{x}Q^2 > M^2 \tag{2.7}$$

2.2 Structure Function and Quark-Parton Model

The cross section for deep-inelastic scattering can be derived with Quantum Electromagnetic Dynamics:

$$\frac{d^2\sigma}{d\Omega dE} = \frac{\alpha^2 E'}{Q^4 E} L^{\mu\nu} W_{\mu\nu} \tag{2.8}$$

The $d\Omega$ is the solid angle of scattered lepton. The lepton tensor $L^{\mu\nu}$ describes the lepton-photon coupling in the electron vertex. The hadron tensor $W_{\mu\nu}$ describes the coupling of the virtual photon to the hadronic system in the hadron vertex. It contains all nucleon structure information, given by[39][21]:

$$W_{\mu\nu} = \frac{1}{4\pi} \int d^4x e^{iqx} \langle P\lambda | J_\mu(x) J_\nu(0) | P\lambda \rangle \tag{2.9}$$

For unpolarised DIS, due to gauge invariance:

$$q_\mu W_{\mu\nu} = q_\nu W_{\mu\nu} = 0 \tag{2.10}$$

The $W_{\mu\nu}$ can be expressed in terms of only two structure functions $W_1(Q^2, \nu)$ and $W_2(Q^2, \nu)$ which reveal essential information for its inner structure:

$$W_{\mu\nu} = W_1(Q^2,\nu)(-g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}) + W_2(Q^2,\nu)\frac{1}{M^2}(P_\mu - \frac{P\cdot q}{q^2}q_\mu)(P_\nu - \frac{P\cdot q}{q^2}q_\nu) \tag{2.11}$$

As a result, the cross-section formula in the rest frame of the proton where the electron scattering angle is θ:

$$\frac{d^2\sigma}{d\Omega dE'} = (\frac{d\sigma}{d\Omega})_{Mott}(2W_1(Q^2,\nu)tan^2\frac{\theta}{2} + W_2(Q^2,\nu)) \tag{2.12}$$

where the $(\frac{d\sigma}{d\Omega})_{Mott}$ is the Mott cross section for point like electron scattering:

$$\frac{d\sigma_{Mott}}{dQ^2} = \frac{4\alpha E'^2}{Q^4}cos^2\frac{\theta}{2} \tag{2.13}$$

In the DIS regime, in the limit $\nu \longrightarrow \infty$, $Q^2 \longrightarrow \infty$, νW_2 and W_1 become dependent on only one variable, x_{bj}[11][12]:

$$MW_2(Q^2,\nu) \longrightarrow F_1(x), \nu W_2(Q^2,\nu) \longrightarrow F_2(x) \tag{2.14}$$

The Elastic scattering falls rapidly with q^2 due to the proton form factors. However, the early DIS experiments performed at the Stanford Linear Accelerator Center (SLAC) show that the deep-inelastic scattering cross sections show a weak dependence on q^2. This indicates the scattering from point-like objects within the proton. The quark-parton model, proposed by Richard Feynman, is that the proton was made up of point-like constituents partons before the quarks and gluons were generally accepted. The inter partonic interactions can be neglected. The structure function becomes:

$$F_1(x) = \frac{1}{2}\sum_i e_i^2 q_i(x) \tag{2.15}$$

$$F_2(x) = x\sum_i e_i^2 q_i(x) \tag{2.16}$$

where q_i is the parton distribution functions (PDF), e_i is the parton charge. The relation $2xF_1(x) = F_2(x)$ is the Callan-Gross relation. This model well describes the results for limited Q^2 data. The quark-parton model is then improved by

Quantum ChromoDynamics (QCD) which evolved to other scales. Written in quark flavors:

$$\frac{1}{x}F_2^p(x) = (\frac{2}{3})^2[u^p(x) + \overline{u}^p(x)] + (\frac{1}{3})^2[d^p(x) + \overline{d}^p(x)] + (\frac{1}{3})^2[s^p(x) + \overline{s}^p(x)] \quad (2.17)$$

If we separate the valence and sea quark

$$u(x) = u_v(x) + u_s(x), d(x) = d_v(x) + d_s(x) \quad (2.18)$$

Extraction of individual parton distribution functions is done by global fits on all the experiments. Additional global constraints on net quark numbers are assumed to match the experiment. For up quarks in proton, the net quark number is 2:

$$\int_0^1 u_v(x)dx = \int_0^1 (u(x) - \overline{u}(x))dx = 2 \quad (2.19)$$

And the momentum fraction for up quark in proton $\int_0^1 xu(x)dx = 2/3$. For down quark in proton, the net quark number is 1:

$$\int_0^1 d_v(x)dx = \int_0^1 (d(x) - \overline{d}(x))dx = 1 \quad (2.20)$$

the momentum fraction for down quark in the proton is 1/3.

The current PDFs are precise with small estimated uncertainties, and the agreement between different groups is within a small error[57][89][8]. Fig. 2.2 shows some current PDFs from the JAM group, the CTEQ group and the NNPDF group. The parton distributions are plotted at experiment Q^2 setting in comparison with high Q^2 PDFs. This will be used in the equation to extract CSV.

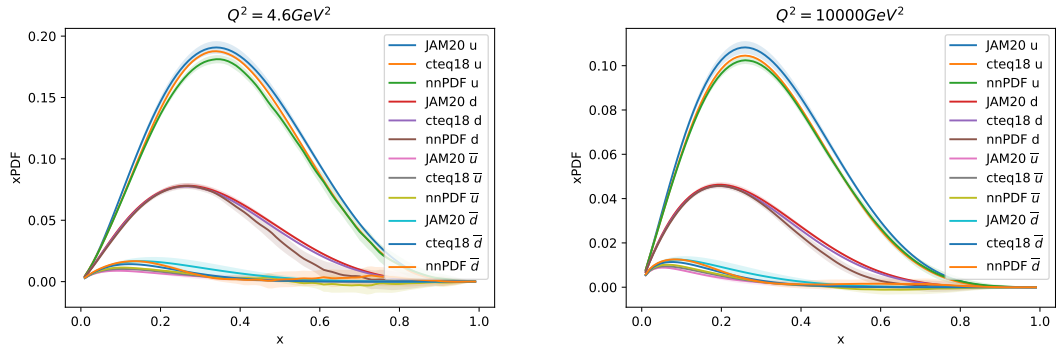


Figure 2.2: Momentum parton distributions (x times parton distribution function) for experiment $Q^2 = 4.6\ GeV^2$ setting and high $Q^2 = 10000\ GeV^2$ setting

2.3 Semi-Inclusive Deep Inelastic Scattering

Semi-Inclusive Deep Inelastic Scattering measures one or more final state hadrons in coincidence with scattered electrons.

$$l + N \hookrightarrow l' + X + h \tag{2.21}$$

In this experiment, the target is fixed, so the target four-momentum is reduced to $P(M, \overrightarrow{0})$. Here in conventional, M represents the mass of proton. In the final state of semi-inclusive scattering, the scattered lepton l' with four-momentum $k'(E', \overrightarrow{k'})$ is detected with a hadron h with four-momentum P_h as in Fig.2.3

For the SIDIS experiment, the final hadronic particle h is detected in coincidence with the scattered lepton. The energy fraction transferred to the hadron h in the laborator frame is defined as z, which ranges between $0 < z < 1$,

$$z \equiv \frac{P \cdot P_h}{P \cdot q} = \frac{E_h}{\nu} \tag{2.22}$$

The mass of the unobserved final state is W'^2,

$$\begin{aligned} W'^2 &= (P + k - k' - P_h)^2 = (M^2 + \nu - E_h)^2 - |\overrightarrow{q} - \overrightarrow{p_h}|^2 \\ &= M^2 + Q^2 \frac{1-x}{x} + M_h^2 - 2 \cdot (z+1) \frac{Q^2}{2 \cdot Mx} \end{aligned} \tag{2.23}$$

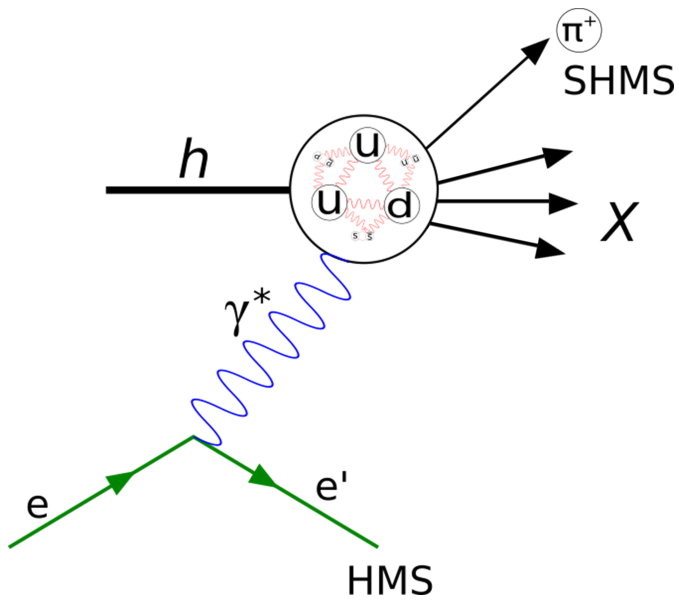


Figure 2.3: the semi-inclusive pion production process, spectrometer HMS is for electrons detection, SHMS is for hadrons.

The usual DIS kinematics such that $Q^2 > 1GeV^2$ and $W^2 > 4GeV^2$ are required. For the SIDIS region, the missing mass squared W'^2 is required to be greater than $2.5GeV^2$

2.4 Fragmentation

Isolated partons have never been observed in nature. Partons exist only inside the hadron and are bound by the interaction between each. This is known as confinement in QCD. Even if the quark in the bound state is ejected by a high-energy photon, it will immediately form a colorless bound state "jet" of hadrons[24]. The fragmentation process proceeds through the generation of additional quark-antiquark pairs, which eventually combine with the original quark or antiquark. The associated hadronization process is described by fragmentation function (FFs) $D_i^h(z, Q^2)(i = q, \overline{q}, g)$, which can be a measure of the probability density that an outgoing stuck parton produces a hadron h. The $D_q^h(z, Q^2)$ may be viewed as the

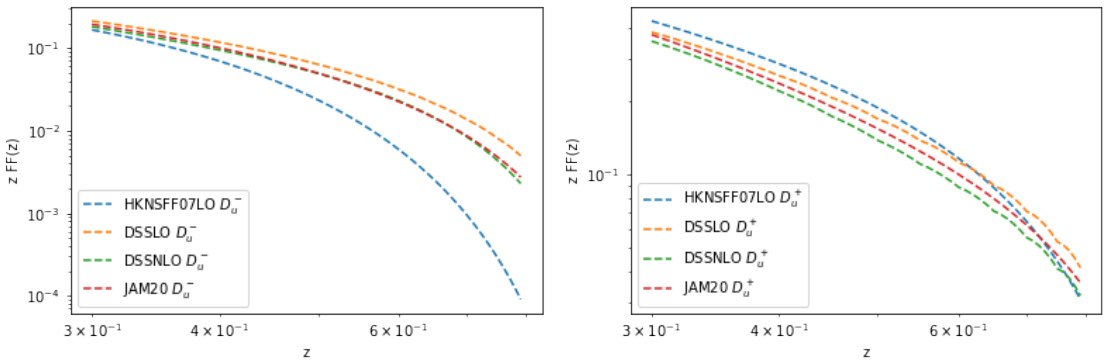


Figure 2.4: Fragmentation Functions for π production, $\pi-$ for left and $\pi+$ for right

final-state analogs of the initial-state parton distribution functions[92].

The fragmentation functions obey the momentum sum rule constraint

$$\sum_h \int_0^1 z D_i^h(z, Q^2) dz = 1 \tag{2.24}$$

separately for each flavor $i = q, \overline{q}, g$. The h represents the possible produced hadrons π, K, etc from each flavor stuck quark. The sum rule involves the sum over all possible hadrons.

The current knowledge of FFs is not as precise as PDFs. While evolving rapidly, the parameterizations for the FFs are not well constrained. The disagreement is shown in Fig.2.4[57][32][18][17].

The FFs are used in the sea quark terms in the formula.

2.5 General Formulas for Semi-Inclusive Reactions

The cross section of SIDIS given in the quark parton model is given by:

$$\frac{1}{\sigma_N(x)} \frac{d\sigma_N^h(x,z)}{dz} = \frac{\sum_i e_i^2 q_i^N(x) D_i^h(z)}{\sum_i e_i^2 q_i^N(x)}. \tag{2.25}$$

In Eq.2.25, $q_i^N(x)$ is the distribution function for quarks of flavor i and charge e_i, in the hadron N as a function of Bjorken x. $D_i^h(z)$ is the fragmentation function

for a quark of flavor i into hadron h depending on z. The quark longitudinal momentum fraction $z = E_h/\mu$ is energy ratio of the hadron and the virtual photon. The numerator in Eq.2.25, $N^{Nh} \equiv \sum_i e_i^2 q_i^N(x) D_i^h(z)$, is the yield of hadron h per scattering from nucleon N, as a function of z and x. In terms of these quantities, semi-inclusive production of a charged hadron from a proton can be described by

$$\begin{aligned} N^{p\pm}(x,z) =& \frac{4}{9}u^p(x)D_u^{\pm}(z) + \frac{4}{9}\overline{u}^p(x)D_{\overline{u}}^{\pm}(z) + \frac{1}{9}d^p(x)D_d^{\pm}(z) \\ &+ \frac{1}{9}\overline{d}^p(x)D_{\overline{d}}^{\pm}(z) + \frac{1}{9}s^p(x)D_s^{\pm}(z) + \frac{1}{9}\overline{s}^p(x)D_{\overline{s}}^{\pm}(z), \end{aligned} \tag{2.26}$$

$$\begin{aligned} N^{n\pm}(x,z) =& \frac{4}{9}u^n(x)D_u^{\pm}(z) + \frac{4}{9}\overline{u}^n(x)D_{\overline{u}}^{\pm}(z) + \frac{1}{9}d^n(x)D_d^{\pm}(z) \\ &+ \frac{1}{9}\overline{d}^n(x)D_{\overline{d}}^{\pm}(z) + \frac{1}{9}s^n(x)D_s^{\pm}(z) + \frac{1}{9}\overline{s}^n(x)D_{\overline{s}}^{\pm}(z). \end{aligned} \tag{2.27}$$

where $D_i^{\pm}$ are the fragmentation functions into positively or negatively charged hadrons. Assuming charge conjugation invariance, one gets

$$\begin{aligned} D_u^{\pi^+} &= D_{\overline{u}}^{\pi^-} \equiv D_1, D_d^{\pi^-} = D_{\overline{d}}^{\pi^+} \equiv D_2 \\ D_u^{\pi^-} &= D_{\overline{u}}^{\pi^+} \equiv D_3, D_d^{\pi^+} = D_{\overline{d}}^{\pi^-} \equiv D_4 \\ D_s^{\pi^+} &= D_{\overline{s}}^{\pi^-} \equiv D_{s,1}, D_s^{\pi^-} = D_{\overline{s}}^{\pi^+} \equiv D_{s,2} \end{aligned} \tag{2.28}$$

By making the additional assumption of charge symmetry, three independent fragmentation functions will obey the relations:

$$\begin{aligned} D_{fav} &= D_1 = D_2 \\ D_{unfav} &= D_3 = D_4 \\ D_s &= D_{s,1} = D_{s,2} \end{aligned} \tag{2.29}$$

and the parton distribution functions will satisfy the relations:

$$\begin{aligned} u^p(x) &= d^n(x) \\ d^p(x) &= u^n(x) \\ s^n(x) &= s^p(x) \end{aligned} \tag{2.30}$$

With these assumptions, we can rewrite $N^{n\pm}$ as:

$$\begin{aligned} N^{n\pm}(x,z) =& \frac{4}{9}d^p(x)D_u^{\pm}(z) + \frac{4}{9}\overline{d}^p(x)D_{\overline{u}}^{\pm}(z) + \frac{1}{9}u^p(x)D_d^{\pm}(z) \\ &+ \frac{1}{9}\overline{u}^p(x)D_{\overline{d}}^{\pm}(z) + \frac{1}{9}s^p(x)D_s^{\pm}(z) + \frac{1}{9}\overline{s}^p(x)D_{\overline{s}}^{\pm}(z). \end{aligned} \tag{2.31}$$

The difference of lepton production of positively and negatively charged hadrons is

$$\begin{aligned} N^{p+}(x,z) - N^{p-}(x,z) =& \frac{4}{9}u_v^p(x)(D_u^+(z) - D_{\overline{u}}^+(z)) + \frac{1}{9}d_v^p(x)(D_d^+(z) - D_{\overline{d}}^+(z)) \\ &+ \frac{1}{9}s_v^p(x)(D_s^+(z) - D_{\overline{s}}^+(z)), \end{aligned} \tag{2.32}$$

$$\begin{aligned} N^{n+}(x,z) - N^{n-}(x,z) =& \frac{4}{9}u_v^n(x)(D_u^+(z) - D_{\overline{u}}^+(z)) + \frac{1}{9}d_v^n(x)(D_d^+(z) - D_{\overline{d}}^+(z)) \\ &+ \frac{1}{9}s_v^n(x)(D_s^+(z) - D_{\overline{s}}^+(z)), \end{aligned} \tag{2.33}$$

These quantities depends on valence quark distributions $u_v(x) \equiv u(x) - \overline{u}(x)$, where $\overline{q}_i(sea) = q_i(sea)$. Also, the validity of the impulse approximation is assumed:

$$N^{D\pi^{\pm}}(x,z) = N^{p\pi^{\pm}}(x,z) + N^{n\pi^{\pm}}(x,z) \tag{2.34}$$

2.6 Factorization at 12 GeV JLab

To extract charge symmetry violating quark distributions, factorization is assumed. The semi-inclusive pion production can be described by two independent processes. First, in the scattering process, a virtual photon interacts with a quark in the hadronic system. Then the fragmentation process, that quark subsequently hadronizes forming a pion or other particle. The separation between the scattering and the fragmentation subprocess ensures the independence of the fragmentation

of the quarks. Assuming these two processes are independent, the cross section can be written as a product of the quark distribution function $q(x)$ and the fragmentation function, $D(z)$, weighted by the quark charge squared e_i^2:

$$\sigma \sim \sum_i e_i^2 q_i(x) D_i^h(z) \tag{2.35}$$

With this picture, the formalism for CSV extraction can be derived.

The semi-inclusive processes at a 12 GeV JLab should satisfy the factorization criteria for this experiment. The previous experiment done in the 6 GeV era has shown a sign that the factorization criteria is satisfied. The result in JLab 6 GeV HallC experiment E-00-108[58] with semi-inclusive charged pion production from hydrogen and deuterium targets verifies the factorization condition. In the experiment, the cross section is modeled following the factorization assumption by the CTEQ5 next-to-leading-order (NLO) parton distribution functions $q(x, Q^2)$[40], fragmentation function parametrization from Binnewise[10], and ratio of D^-/D^+ from HERMES Collaboration[66]. The cross section measured by the experiment is found to be consistent with the modeled cross section as in Fig.2.5

In addition, the simple ratios of the measured cross sections are formed insensitive to the factorization process:

$$\frac{\sigma_p^{\pi^+} + \sigma_p^{\pi^-}}{\sigma_d^{\pi^+} + \sigma_d^{\pi^-}} = \frac{4u + 4\overline{u} + d + \overline{d}}{5(u + \overline{u} + d + \overline{d})} \tag{2.36}$$

and

$$\frac{\sigma_p^{\pi^+} - \sigma_p^{\pi^-}}{\sigma_d^{\pi^+} - \sigma_d^{\pi^-}} = \frac{4u_v - d_v}{3(u_v + d_v)} \tag{2.37}$$

The two quantity should be independent of combinations of the two charge states and targets if the factorization holds. The results for the two ratios in the 6 GeV experiment are in Fig. 2.6. The ratios are independent of z up to $z \approx 0.7$, which

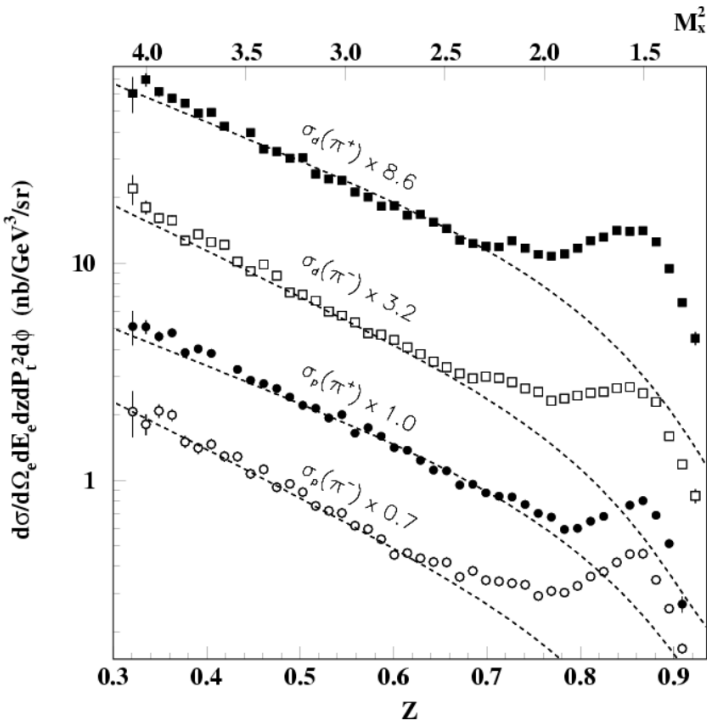


Figure 2.5: Cross section for $H, D(e, e'\pi^{\pm}$ from Hall C experiment E-00-108. Curves denote cross section calculation using simple factorization using the models. Closed symbols have been corrected for exclusive ρ production backgrounds.

indicates that the simple factorization assumption holds up to $z \approx 0.7$.

With results from the 6 GeV experiment promising, a similar test of factorization on this experiment is done. The H_2 runs are performed with the kinematics relevant to the experiment. The factorization assumption is tested to satisfy the kinematics coverage.

2.7 Experimental Extraction of Charge Symmetry Violation

Semi-inclusive pion production from lepton deep inelastic scattering on an isoscalar nuclear targets was suggested[37; 42] as a sensitive probe of CSV effects in nucleon valence distributions. For pion production, the semi-inclusive reactions

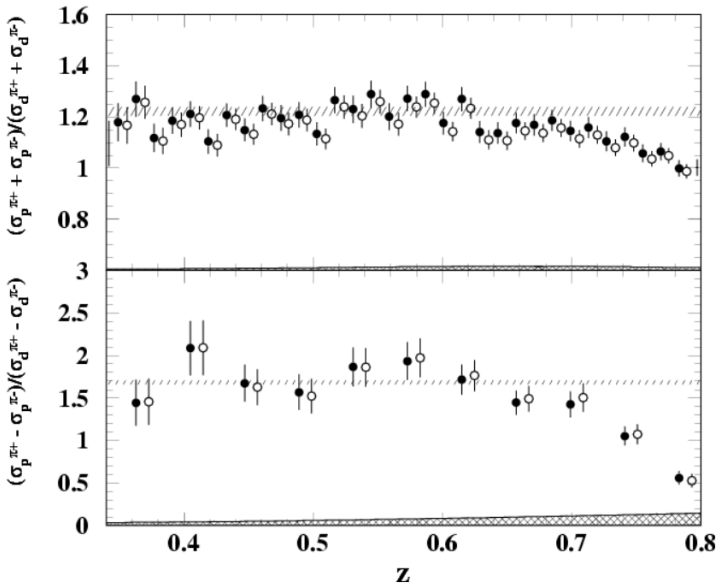


Figure 2.6: Ratios of proton to deuterium results for the sum (top) and difference (bottom) of the π^+ and π^- cross sections. Closed symbols have been corrected for exclusive ρ production backgrounds.

are proportional to the square of the quark charge, with the assumption of charge symmetry, the "favored" $\pi+$ production is four times more than the "unfavored" $\pi-$ production. The name "favored" comes from the quarks that are included in the lepton production. The authors proposed to measure the quantity $R^D_{meas}(x,z)$ defined by:

$$R^D_{meas}(x,z) = \frac{4N^{D\pi^-}(x,z) - N^{D\pi^+}(x,z)}{N^{D\pi^+}(x,z) - N^{D\pi^-}(x,z)} \tag{2.38}$$

where $N^{D\pi^+}(N^{D\pi^-})$ is the yield of $\pi^+(\pi^-)$ produced in coincidence with the scattered electron from deuterium. With the validity of the impulse approximation in Eq. 2.34, and Eq. 2.26, Eq. 2.27, the expression for R^D_{meas} is

$$R^D_{meas} = \frac{4N^{p\pi^-} + 4N^{n\pi^-} - N^{p\pi+} - N^{n\pi+}}{N^{p\pi^+} + N^{n\pi^+} - N^{p\pi^-} - N^{n\pi^-}} \tag{2.39}$$

The numerator part is

$$\begin{aligned}
&\frac{16}{9}u^pD_u^- + \frac{16}{9}\overline{u}^pD_{\overline{u}}^- + \frac{4}{9}d^pD_d^- + \frac{4}{9}\overline{d}^pD_{\overline{d}}^- + \frac{4}{9}s^pD_s^- + \frac{4}{9}\overline{s}^pD_{\overline{s}}^- \\
&+\frac{16}{9}u^nD_u^- + \frac{16}{9}\overline{u}^nD_{\overline{u}}^- + \frac{4}{9}d^nD_d^- + \frac{4}{9}\overline{d}^nD_{\overline{d}}^- + \frac{4}{9}s^nD_s^- + \frac{4}{9}\overline{s}^nD_{\overline{s}}^- \\
&-\frac{4}{9}u^pD_u^+ - \frac{4}{9}\overline{u}^pD_{\overline{u}}^+ - \frac{1}{9}d^pD_d^+ - \frac{1}{9}\overline{d}^pD_{\overline{d}}^+ - \frac{1}{9}s^pD_s^+ - \frac{1}{9}\overline{s}^pD_{\overline{s}}^+ \\
&-\frac{4}{9}u^nD_u^+ - \frac{4}{9}\overline{u}^nD_{\overline{u}}^+ - \frac{1}{9}d^nD_d^+ - \frac{1}{9}\overline{d}^nD_{\overline{d}}^+ - \frac{1}{9}s^nD_s^+ - \frac{1}{9}\overline{s}^nD_{\overline{s}}^+
\end{aligned} \tag{2.40}$$

The valence quark $u_v(x) = u(x) - \overline{u}(x)$ and $d_v(x) = d(x) - \overline{d}(x)$. With the charge conjugation invariance as in Eq.2.28 and the charge symmetry assumption as in Eq.2.29 and Eq.2.30, the denominator part is in Eq. 2.38

$$\begin{aligned}
N^{D\pi^+} - N^{D\pi^-} &= \frac{1}{3}(u_v^p + d_v^p)[D_u^+(z) - D_u^-(z)] \\
&= \frac{1}{3}(u_v^p + d_v^p)D_u^+(1 - \Delta(z))
\end{aligned} \tag{2.41}$$

where

$$\Delta(z) = \frac{D_u^-(z)}{D_u^+(z)} \tag{2.42}$$

The numerator Eq.2.40 becomes

$$\frac{15}{9}u^pD_u^- + \frac{15}{9}d^pD_u^- + \frac{15}{9}\overline{u}^pD_u^+ + \frac{15}{9}\overline{d}^pD_u^+ + \frac{3}{9}(s^p + s^n)(D_s^- + D_s^+) \tag{2.43}$$

With charge symmetry assumed in parton distribution function and fragmentation function, R_{meas}^D becomes

$$\begin{aligned}
R_{meas}^D(x,z) = &\frac{5\Delta(z)}{1-\Delta(z)} + 5\left[\frac{1+\Delta(z)}{1-\Delta(z)}\right] \times \frac{[\overline{u}^p(x) + \overline{d}^p(x)]}{[u_v^p(x) + d_v^p(x)]} \\
&+ \left[\frac{\Delta_s(z)}{1-\Delta(z)}\right]\frac{s(x) + \overline{s}(x)}{u_v^p(x) + d_v^p(x)}
\end{aligned} \tag{2.44}$$

For strange quark in sea region, the quark distributions obey $\overline{s}^p = s^p$ and $\overline{s}^n = s^n$. The charge conjugation implies that the fragmentation function to have the relation $D_{\overline{s}}^- = D_s^+$ and $D_{\overline{s}}^+ = D_s^-$. With all the equations, strange quark part

can be written

$$\begin{aligned} R^D_{strange} &= \frac{(s^p + s^n)(D^-_s + D^+_s)}{(u^p_v + d^p_v)D^+_u[1 - \Delta(z)]} \\ &= \left[\frac{\Delta_s(z)}{1 - \Delta(z)}\right] \frac{s(x) + \overline{s}(x)}{u^p_v(x) + d^p_v(x)} \end{aligned} \tag{2.45}$$

with the strange-favored ratio of the quark fragmentation functions

$$\Delta_s(z) = \frac{D^-_s(z) + D^+_s(z)}{D^+_u(z)} \tag{2.46}$$

Keeping charge-symmetry-violating functions, the quantity R^D_{meas} contains charge symmetry-violating quark distribution functions

$$\begin{aligned} \delta d(x) &= d^p(x) - u^n(x) \\ \delta u(x) &= u^p(x) - d^n(x) \end{aligned} \tag{2.47}$$

and the charge symmetry breaking fragmentation functions

$$\delta D(z) = \frac{D^+_u(z) - D^-_d(z)}{D^+_u(z)} \tag{2.48}$$

In the region of interest (moderately large z), the unfavored fragmentation function is considerably smaller than the favored term. We can neglect the CSV part of the "unfavored" fragmentation function, $D^+_d(z) = D^-_u(z)$. In consequence, the unfavored CSV term should be proportionately smaller than the favored CSV term. The sea quark charge symmetry breaking in this region can also be neglected ($\overline{u}^n = \overline{d}^p, \overline{d}^n = \overline{u}^p$). The strange quark part is the same with or without CSV assumption.

Expanding to the first order of the small quantities (δd,δu,δD), and ignoring the terms with both sea quark distribution and δD, the up and down quark part

in the numerator from Eq. 2.40 becomes

$$\begin{aligned}
&\frac{16}{9}u^pD_u^- + \frac{16}{9}\overline{u}^pD_u^+ + \frac{4}{9}d^pD_u^+ - \frac{4}{9}d^pD_u^+\delta D + \frac{4}{9}\overline{d}^pD_u^- \\
&+ \frac{16}{9}d^pD_u^- - \frac{16}{9}\delta dD_u^- + \frac{16}{9}\overline{d}^pD_u^+ + \frac{4}{9}u^pD_u^+ - \frac{4}{9}u^pD_u^+\delta D \\
&- \frac{4}{9}\delta uD_u^+ + \frac{4}{9}\overline{u}^pD_u^- \\
&- \frac{4}{9}u^pD_u^+ - \frac{4}{9}\overline{u}^pD_u^- - \frac{1}{9}d^pD_u^- - \frac{1}{9}\overline{d}^pD_u^+ \\
&- \frac{4}{9}d^pD_u^+ + \frac{4}{9}\delta dD_u^+ - \frac{4}{9}\overline{d}^pD_u^- - \frac{1}{9}u^pD_u^- + \frac{1}{9}\delta uD_u^- \\
&- \frac{1}{9}\overline{u}^pD_u^+ \\
&= \frac{15}{9}u^pD_u^- + \frac{15}{9}\overline{u}^pD_u^+ + \frac{15}{9}d^pD_u^- + \frac{15}{9}\overline{d}^pD_u^+ \\
&- \frac{4}{9}d^pD_u^+\delta D - \frac{4}{9}u^pD_u^+\delta D \\
&+ \frac{4}{9}\delta dD_u^+ - \frac{16}{9}\delta dD_u^- - \frac{4}{9}\delta uD_u^+ + \frac{1}{9}\delta uD_u^-
\end{aligned} \tag{2.49}$$

The denominator is

$$\frac{3}{9}(u_v^p + d_v^p)(D_u^+ - D_u^-) + (D_u^+ - D_u^-)(\frac{1}{9}\delta u - \frac{4}{9}\delta d) + \frac{1}{9}(u^p + d^p)D_u^+\delta D \tag{2.50}$$

which can be written:

$$\frac{3}{9}(u_v^p + d_v^p)(D_u^+ - D_u^-)(1 + \frac{\delta u - 4\delta d}{3(u_v^p + d_v^p)} + \frac{\delta D}{3(1-\Delta(z))}) \tag{2.51}$$

For the small quantity $(\delta d, \delta u, \delta D)$ in denominator, Taylor series is used:

$$\frac{1}{1-\delta} = 1 + \delta + \mathcal{O}(\delta^2) \tag{2.52}$$

The terms in R^D_{meas} numerator that includes just parton distribution function and fragmentation function terms are:

$$\begin{aligned}
&\frac{\frac{15}{9}u^pD_u^- + \frac{15}{9}d^pD_u^- + \frac{15}{9}\overline{u}^pD_u^+ + \frac{15}{9}\overline{d}^pD_u^+}{\frac{1}{3}(u_v^p + d_v^p)D_u^+(1-\Delta(z))}(1 - \frac{\delta u - 4\delta d}{3(u_v^p + d_v^p)} - \frac{\delta D}{3(1-\Delta(z))}) \\
&= (\frac{5\Delta(z)}{1-\Delta(z)} + \frac{5(\overline{u}^p + \overline{d}^p)(1+\Delta(z))}{(u_v^p + d_v^p)(1-\Delta(z))})(1 - \frac{\delta u - 4\delta d}{3(u_v^p + d_v^p)} - \frac{\delta D}{3(1-\Delta(z))}) \\
&= \frac{5\Delta(z)}{1-\Delta(z)} + \frac{15(\Delta(z)+1)(\overline{u}^p + \overline{d}^p)}{3(1-\Delta(z))(u_v^p + d_v^p)} + \frac{5\Delta(z)(4\delta d - \delta u)}{3(u_v^p + d_v^p)(1-\Delta(z))} - \frac{5\Delta(z)\delta D}{3(1-\Delta(z))^2}
\end{aligned} \tag{2.53}$$

The terms include charge symmetry breaking fragmentation functions δD (ignore the sea quark terms) are:

$$\begin{aligned}&\frac{-\frac{4}{9}d^pD_u^+\delta D-\frac{4}{9}u^pD_u^+\delta D+\frac{1}{9}\overline{d}^pD_u^+\delta D+\frac{1}{9}\overline{u}^pD_u^+\delta D}{\frac{1}{3}(u_v^p+d_v^p)D_u^+(1-\Delta(z))}\\&=-\frac{4\delta D}{3(1-\Delta(z))}\end{aligned}\tag{2.54}$$

Here the higher order terms from the Taylor expansion term of denominator are ignored. Same for the terms that include the CSV δd and δu:

$$\frac{-\frac{16}{9}\delta dD_u^-\frac{4}{9}\delta uD_u^+\frac{4}{9}\delta dD_u^+\frac{1}{9}\delta uD_u^-}{\frac{1}{3}(u_v^p+d_v^p)D_u^+(1-\Delta(z))}\tag{2.55}$$

The $R^D_{meas}(x,z)$ ratio becomes:

$$\begin{aligned}R^D_{meas}(x,)=&\frac{5\Delta(z)}{1-\Delta(z)}-\frac{[4+\Delta(z)]\delta D(z)}{3[1-\Delta(z)]^2}\\&+\left[\frac{1+\Delta(z)}{1-\Delta(z)}\right]\times\left[\frac{4[\delta d(x)-\delta u(x)]+15[\overline{u}^p(x)+\overline{d}^p(x)]}{3[u_v^p(x)+d_v^p(x)]}\right]\\&+\left[\frac{\Delta_s(z)}{1-\Delta(z)}\right]\frac{s(x)+\overline{s}(x)}{u_v^p(x)+d_v^p(x)}\end{aligned}\tag{2.56}$$

And the strange/favored ratio of quark fragmentation functions as in Eq. 2.46. To separate the x and z dependence more cleanly, one can multiply the ratio with a z dependent factor,

$$\begin{aligned}\tilde{R}^D(x,z)\equiv&\frac{1-\Delta(z)}{1+\Delta(z)}R^D_{meas}(x,z)\\=&\frac{5\Delta(z)}{1+\Delta(z)}-\frac{(4+\Delta(z))\delta D(z)}{3(1-\Delta^2(z))}+\frac{4[\delta d(x)-\delta u(x)]}{3[u_v^p+d_v^p]}\\&+\frac{5(\overline{u}^p(x)+\overline{d}^p(x))+\Delta_s(z)[s(x)+\overline{s}(x)]/(1+\Delta(z))}{[u_v^p(x)+d_v^p(x)]}\\\equiv&\tilde{R}^D_f(z)+\tilde{R}^D_{CSV}(x)+\tilde{R}^D_{sea}(x,z)\end{aligned}\tag{2.57}$$

where

$$\begin{aligned}\tilde{R}^D_f(z)&=\frac{5\Delta(z)}{1+\Delta(z)}-\frac{(4+\Delta(z))\delta D(z)}{3(1-\Delta^2(z))}\\\tilde{R}^D_{CSV}(x)&=\frac{4[\delta d(x)-\delta u(x)]}{3[u_v^p+d_v^p]}\\\tilde{R}^D_{sea}(x,z)&=\frac{5(\overline{u}^p(x)+\overline{d}^p(x))+\Delta_s(z)[s(x)+\overline{s}(x)]/(1+\Delta(z))}{[u_v^p(x)+d_v^p(x)]}\end{aligned}\tag{2.58}$$

The quantity $\tilde{R}^D(x,z)$ separates into three pieces. The first piece, $\tilde{R}^D_f$ in Eq. 2.58, depends only on z. The CSV contribution of the fragmentation function to the z-dependent term $\tilde{R}^D_f(z)$ was estimated to be 1% [42]. Therefore, it can be neglected in the expression of $\tilde{R}^D_f(z)$

$$\tilde{R}^D_f(z) \sim \frac{5\Delta(z)}{(1+\Delta(z))} \tag{2.59}$$

For most values of z, this term should decrease monotonically as z increases (not true for large values of z). More accurate and sophisticated fragmentation functions are carried out which will be used in this calculation. The second term, $\tilde{R}^D_{CSV}$ depends only on x. It is proportional to the nucleon CSV term. The last term $\tilde{R}^D_{sea}$ is proportional to the sea quark distributions and can be written as a sum of a strange sea term $R^D_{sea_S}$ and a non-strange sea term $R^D_{sea_{NS}}$, where

$$R^D_{sea_S} = \frac{\Delta_s(z)[s(x)+\overline{s}(x)]/(1+\Delta(z))}{[u^p_v(x)+d^p_v(x)]} \tag{2.60}$$

$$R^D_{sea_{NS}} = \frac{5(\overline{u}^p(x)+\overline{d}^p(x)}{[u^p_v(x)+d^p_v(x)]} \tag{2.61}$$

By rearranging Eq. 2.57, we can get the following formula:

$$D(z)R(x,z) + A(x)CSV(x) = B(x,z) \tag{2.62}$$

where

$$\begin{aligned} D(z) &= \frac{1-\Delta(z)}{1+\Delta(z)}, \Delta(z) = \frac{D^-_u(z)}{D^+_u(z)} \\ R(x,z) &= \frac{5}{2} + R^D_{meas}, \\ CSV(x) &= \delta d - \delta u, \\ A(x) &= \frac{-4}{3(u_v+d_v)}, \\ B(x,z) &= \frac{5}{2} + R^D_{sea_S}(x,z) + R^D_{sea_NS}(x). \end{aligned} \tag{2.63}$$

For a better estimate of the sea quark terms, PDFs from the CTEQ group [68] and the JAM group are used. Fragmentation functions are from De Florian and collaborators [17] and the JAM group [57]. In this experiment,$R^D_{meas}(x, z)$ (see Eq. (2.38)) is measured, and thus R(x,z) for 16 bins in x and z for 3 distinct Q^2 bins. At the end, there will be one equation similar to Eq. (2.62) for each of the kinematic settings. The CSV can then be extracted from all the experiment data.

Chapter 3

EXPERIMENTAL APPARATUS

3.1 Overview

This experiment was performed at the Thomas Jefferson National Accelerator Facility (TJNAF), a.k.a. Jefferson Lab, using the standard Hall C apparatus. Jefferson Lab is a National Laboratory with an accelerator known as the Continuous Electron beam Accelerator Facility (CEBAF). Its mission is "to provide forefront scientific facilities, opportunities, and leadership essential for discovering the fundamental structure of nuclear matter." The electron beam is generated from CEBAF with a frequency of 1497 MHz. The upgraded bending magnets at the arc ensure the electrons loop five times in the beamline and get accelerated enough in the linear accelerator sections. Then the electron beam is directed into three experimental halls, as in Figure 3.1 Since the high demand for accurate measurements, small angles, and large momenta spectrometers are required. HMS and SHMS in Hall C at Jefferson Lab are suitable for this experiment. They have identical acceptance for π^+ and π^-, which enables them to be the ideal spectrometers

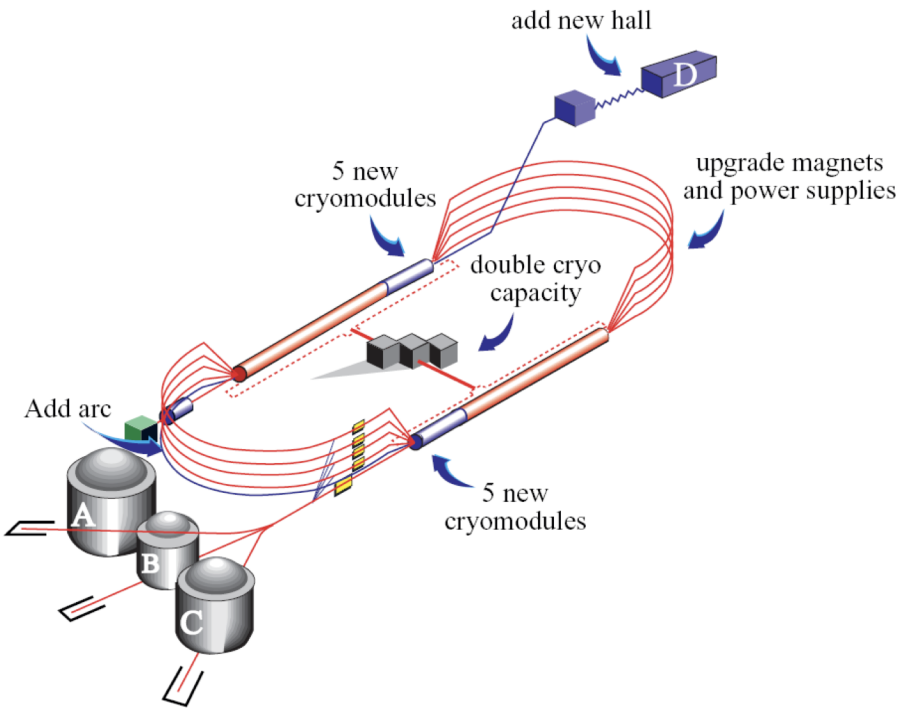


Figure 3.1: Jefferson Lab CEBAF

for measuring the charged pion ratio.

3.2 CEBAF Accelerator and Hall C Beamline

The Continuous Electron Beam Accelerator Facility (CEBAF) provides high-quality continuous electron beams incident in experiment halls in Thomas Jefferson National Accelerator Facility (TJNAF). The first beam was successfully delivered into experimental Hall C in 1994 and reached beam energy at 4 GeV in the following year. Implementing novel superconducting radiofrequency (SRF) accelerator technology for continuous wave (CW) operation allows beam energy up to 6 GeV with an electron beam delivered simultaneously into three experimental halls. Shortly after the first experiment, the upgrade of beam to the maximum energy of 12 GeV plan is put on the table. In 2012, the facility officially finished

all the 6 GeV experiments and started to be prepared for the 12 GeV upgrade. [2]

The principal modification of CEBAF needed to provide an increased beam energy is in the linac section of the beamline. Out of the existing 20 cryomodules accelerator, five more new cryomodules are added on both sides of the linac beamline. Each of the new cryomodules is upgraded to 8 7-cell superconducting RF cavities which are attached with an RF generator to create the flipping electric field for electrons to get accelerated as in Fig. 3.2. Each of the new cryomodule could accelerate electrons up to 100 MeV, which yielded an additional acceleration of 0.5 GeV per linac. Added to the existing cryomodules boost of 0.6 GeV per linac, the electrons can accumulate 1.1 GeV per linac now. The electrons get accelerated 2.2 GeV each pass with the upgraded bending arc. For Hall C, as well as Halls A and B, electrons get boosted five times and reach a maximum of 11 GeV energy. For the newly built Hall D, with the new arc, electrons were accelerated by one more linac, which yielded the JLab maximum beam energy, 12 GeV[48].

Simultaneous CW operation of three end stations was designed to send beams to all the halls. It is achieved by striking three independent laser lights to a single photocathode to generate three independent electron beams as in Fig. 3.3. Each has one-third of the main accelerator frequency.The RF cavity frequency is 1497 MHz, and the electrons delivered to hall C have a frequency of 249 MHz, as in Fig. 3.4, which leads to a 4 ns beam period. The RF separator at the end of the beamline will kick the corresponding beam bucket to each hall.

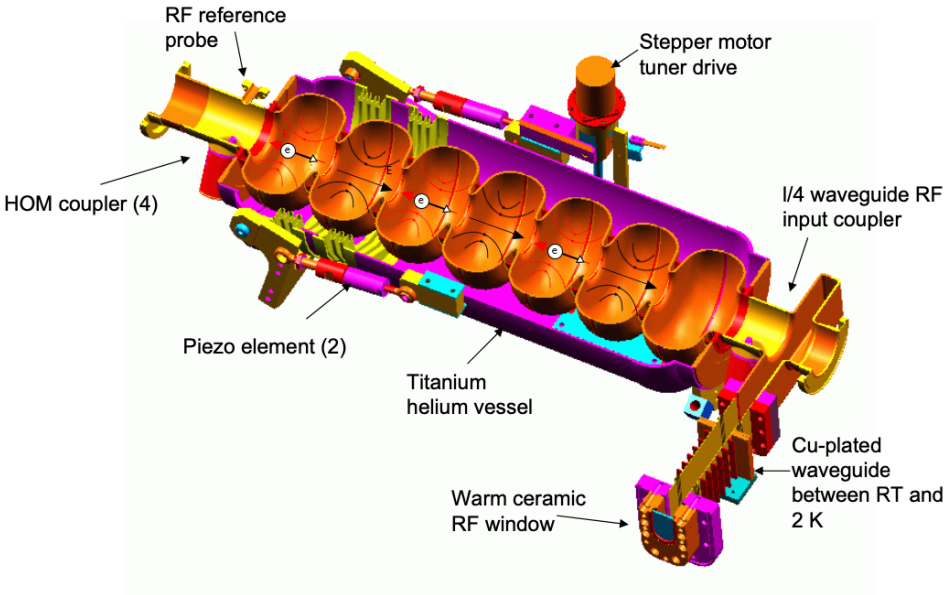


Figure 3.2: Electrons get accelerated in the RF cavity. The frequency is 1497 MHz, producing three beams at the third sub-harmonic, 500 MHz.The RF separator can split all beams and send them to the respective experimental halls at the end.[38]

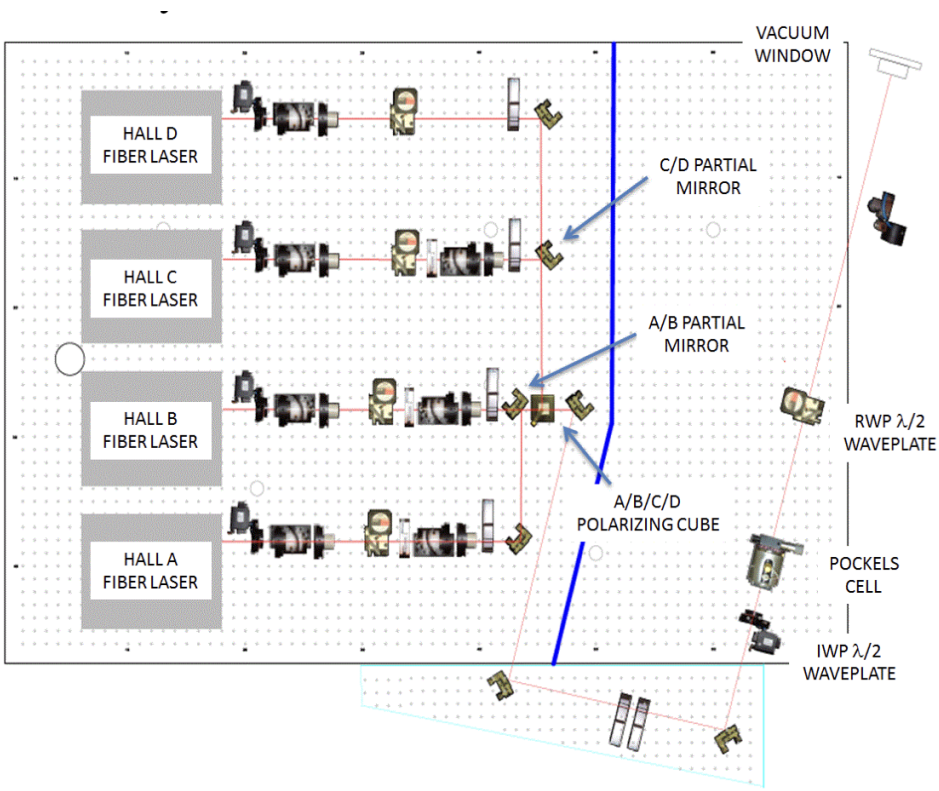


Figure 3.3: Four Laser configuration for electron injector.[38]

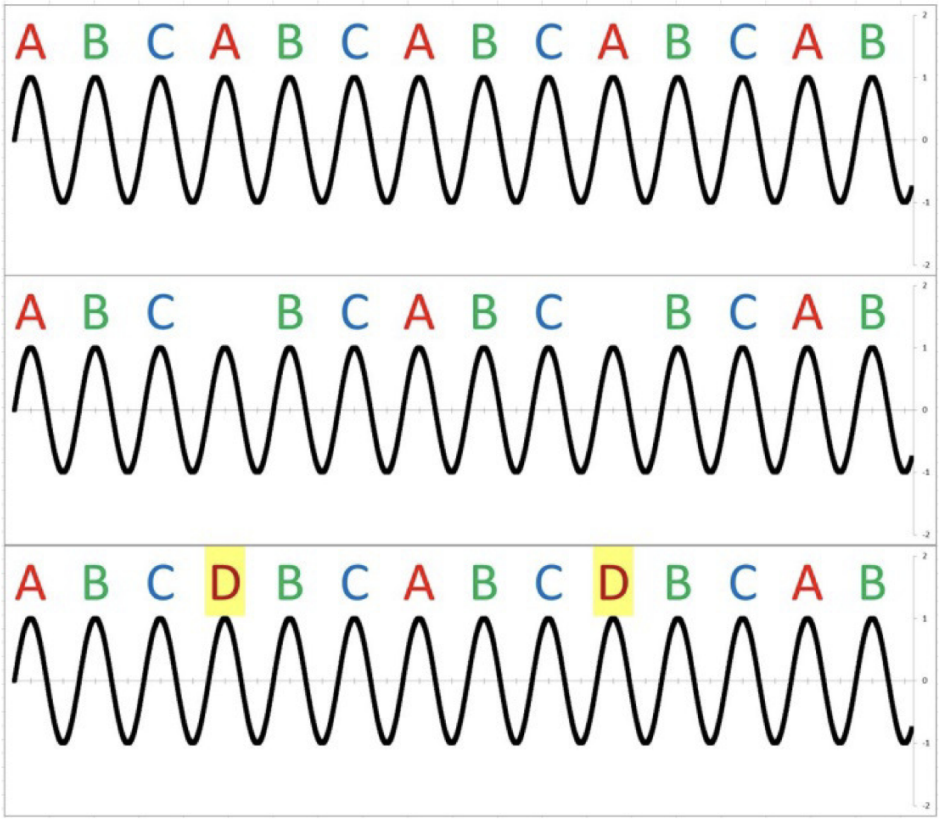


Figure 3.4: Each sine wave represents the 1500 MHz accelerating frequency. The top three are before the 12 GeV upgrade; each experimental hall gets a 500 MHz beam. The center and bottom beams are modified to make room for the Hall D beam. One of the high pass Hall (A or C) would reduce the frequency to 250 MHz

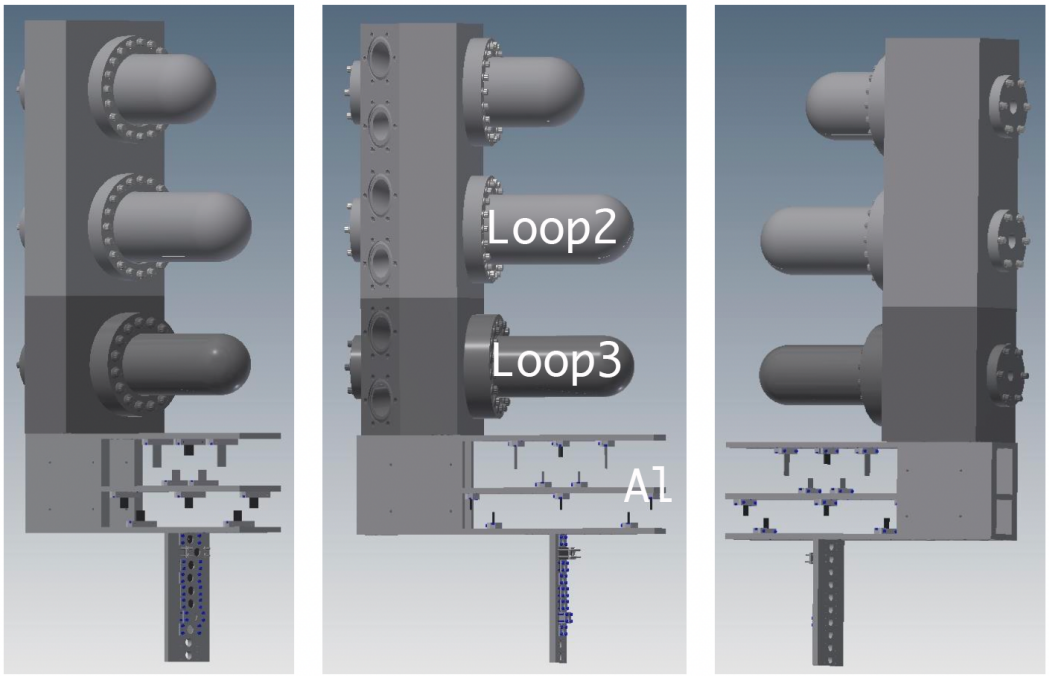


Figure 3.5: Cryogenic and solid target ladders. Loop 2 is liquid hydrogen. Loop 3 is liquid deuterium. The aluminum dummy and carbon optics targets are below liquid targets. The cryo-targets are assembled above the solid target ladder. The beam enters from the left in the left picture.[88]

3.3 Target Chamber

The experiment used a system of cryo-targets and solid targets in Hall C. Liquid hydrogen, liquid deuterium, and aluminum dummy target, including other targets, are mounted on a target ladder as shown in Fig. 3.5. The whole target ladder system is kept in a vacuum (few times 10^{-6} Torr) inside a scattering chamber with a single axis vertical motion system to move the desired target into the beamline axis position.

The liquid hydrogen and liquid deuterium target are used for data collection during the experiment. The aluminum dummy target is mounted on separate frames at the locations $Z = \pm 5$cm corresponding to the cryogenic target entrance and exit windows. The dummy target runs are used to subtract the actual cryo-target windows backgrounds. The hydrogen and deuterium targets are liquefied

Target	Temperature ($^{\circ}K$)	Pressure (PSIA)	Freezing T ($^{\circ}K$)	Boiling T ($^{\circ}K$)
H_2	19 ± 0.1	24 ± 2	13.86	22.24
D_2	22 ± 0.1	24 ± 2	18.83	25.13

Table 3.1: Cryo-target states during normal operating conditions. [88]

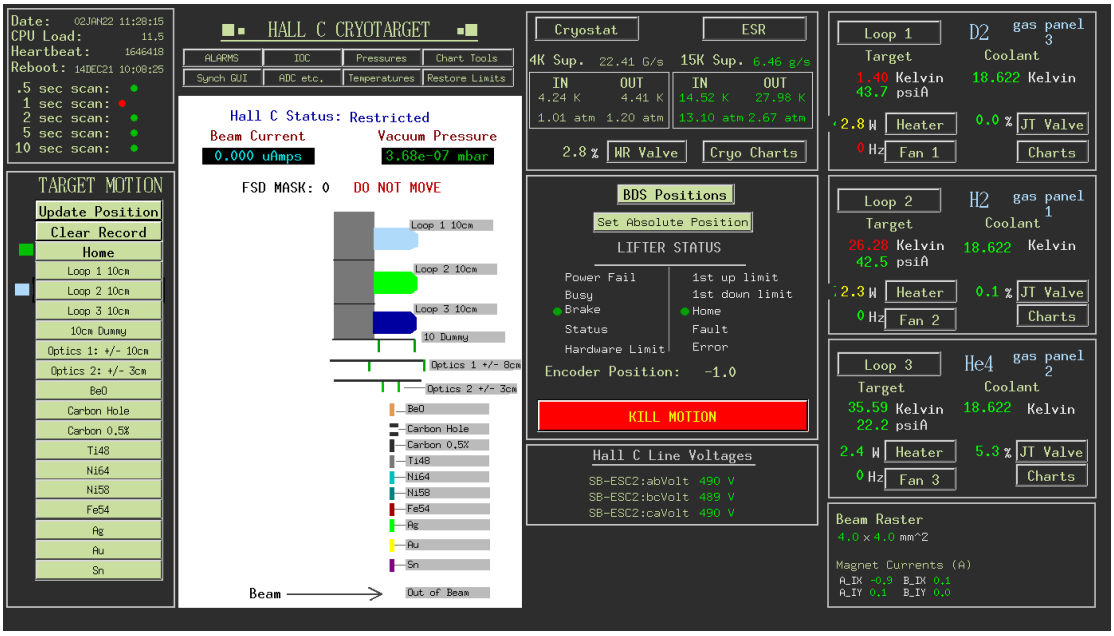


Figure 3.6: This is the target screen graphic user interface when the beam is not on any targets. Moving the target ladder can move the desired target to the beam position. Liquid cryo-target states are monitored and can be changed by cooling and heating. [88]

by the target group and controlled in a stable state, as in the Table 3.1. The temperature is controlled by the coolant loop and a high-power heater. The interface to control the coolant valve and heater is shown in Fig. 3.6 in the counting room.

3.4 Hall C Spectrometer

Hall C has two spectrometers, SHMS and HMS (Fig. 3.7), for particle measurement. The spectrometer can rotate around the pivot where the scattering chamber is located. There are two main coordinate systems in Hall C, as shown in

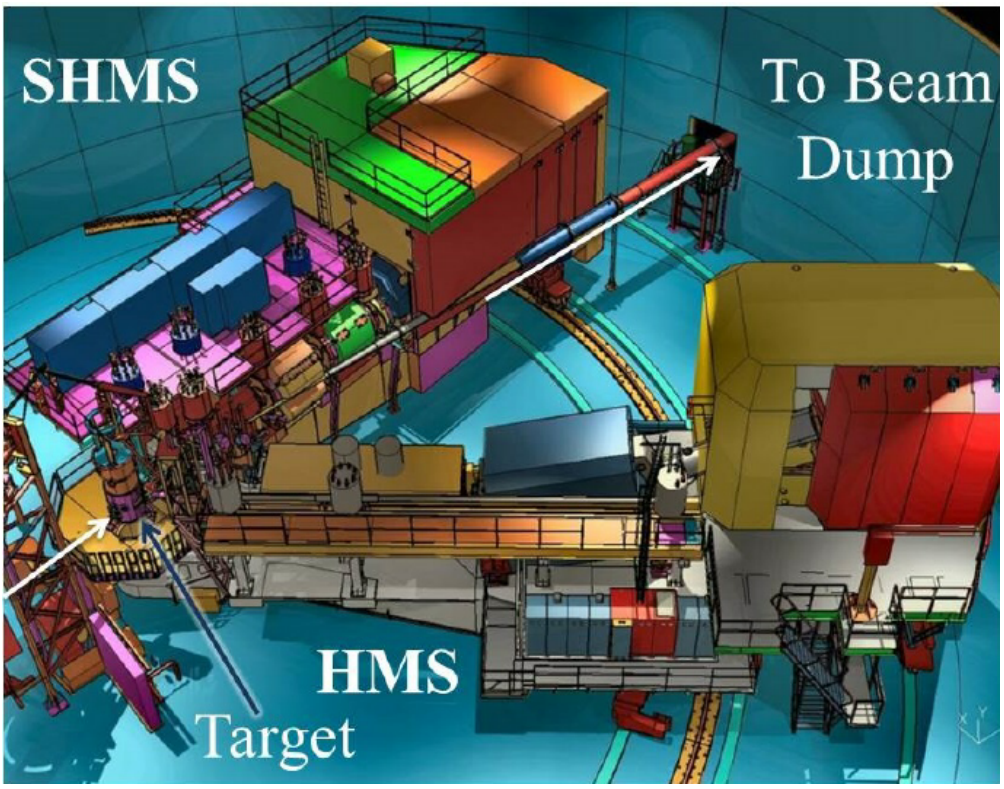


Figure 3.7: Hall C setup

Fig.3.8. The experiment beam coordinate system and the spectrometer coordinate system. For the beam coordinate system: z directs along the beamline; y points down to the ground; x is decided by the right-hand law to the right of the beam (look downstream). After the target chamber, the coordinates transfer to the spectrometer system. In the spectrometer coordinate system: z directs along the optical axis inside the spectrometer; x-direction is perpendicular to z-direction to the ground; y-direction is determined by right-hand law, which is left (look downstream) to both spectrometers. The central trajectory is the trajectory along the center of the entrance acceptance.

3.4.1 Hall C High Momentum Spectrometer (HMS)

The HMS is used as the electron arm in this experiment. It operates on the beam-right side of the beamline as in Fig.3.7. When beam electrons are scat-

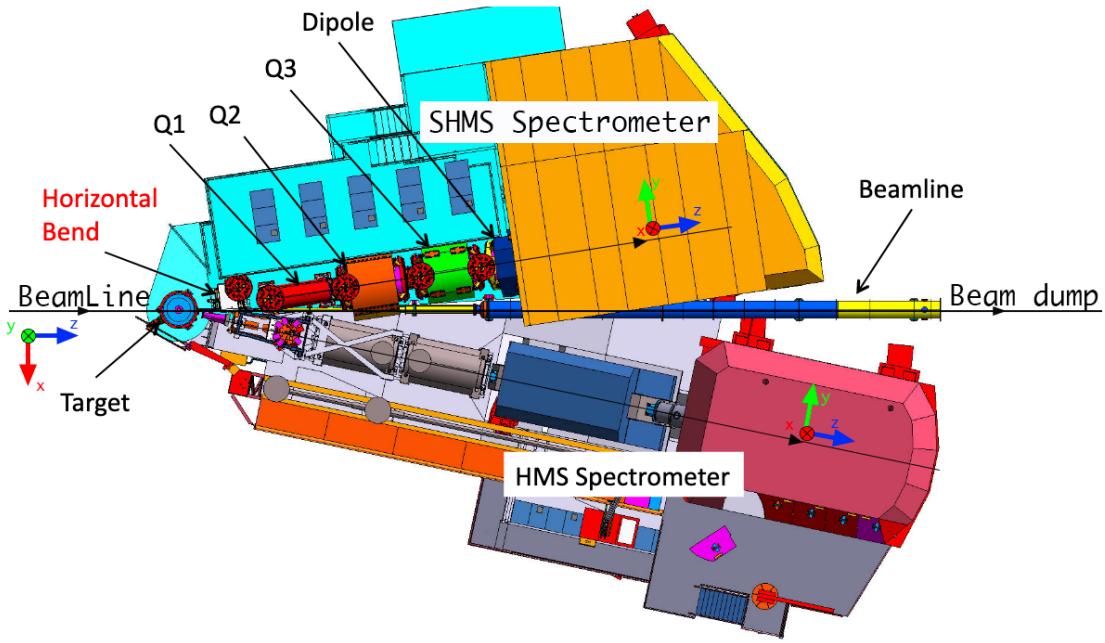


Figure 3.8: Hall C top view and coordinate system

tered in the target chamber, secondary particles are produced everywhere in all directions with all energies. Those produced particles at the configured spectrometer setting will be accepted. The accepted particles traverse three quadrupole magnets, Q1, Q2, and Q3, bend upwards by 25° in the dipole and produce a signal in the detector hut. The quadrupoles focus on the particle trajectories and determine the main acceptance. The quadrupoles are made by cold iron superconducting magnets, with soft iron around the superconducting coils. Followed by the QQQD magnets is the detector hut wrapped in a large concrete room as in Fig.3.9. The magnets and the detector hut are mounted on a big carriage with steel wheels mounted on rails that can rotate around the pivot. The rotation of the spectrometer can be controlled remotely by rotation GUI in the counting room. The NMR probe is placed at the center of the dipole and reads the magnetic field which determines the spectrometer's central momentum. The magnets

Range of central momentum	0.4-7.4GeV/c
Momentum deviation	$\pm 10\%$
Momentum resolution	$0.1\% - 0.15\%$
Scattering angle range	10.5° - 90°
Scattering angle plane angular acceptance	± 32 mrad
Out of plane angular acceptance	± 82 mrad
Scattering angle plane angular resolution	0.8 mrad
Out of plane angular resolution	1.0 mrad

Table 3.2: HMS performance

are precooled with liquid nitrogen and cooled by liquid helium to the operating temperature of 4.3 K.

The HMS can be moved to a scattering angle between 10.5° to 90°. Cameras down in the hall can monitor the rotation and monitor the angles. The momentum range can reach up to 7.4 GeV, which satisfies our momentum requirement (around 6 GeV). The dipole determines the momentum for the particle to travel along the central trajectory, the central momentum p_0. The deviation near the central momentum is δp which is defined as $\frac{p-p_0}{p_0}$ in percentage. The momentum acceptance for HMS is $\pm 10\%$ about the central momentum. The momentum resolution of the HMS is 0.1% to 0.15%. The angular acceptance is ± 32 mrad along the horizontal direction and ± 85 mrad along the vertical direction. The angular resolution is 0.8 mrad horizontal and 1.0 mrad vertical.

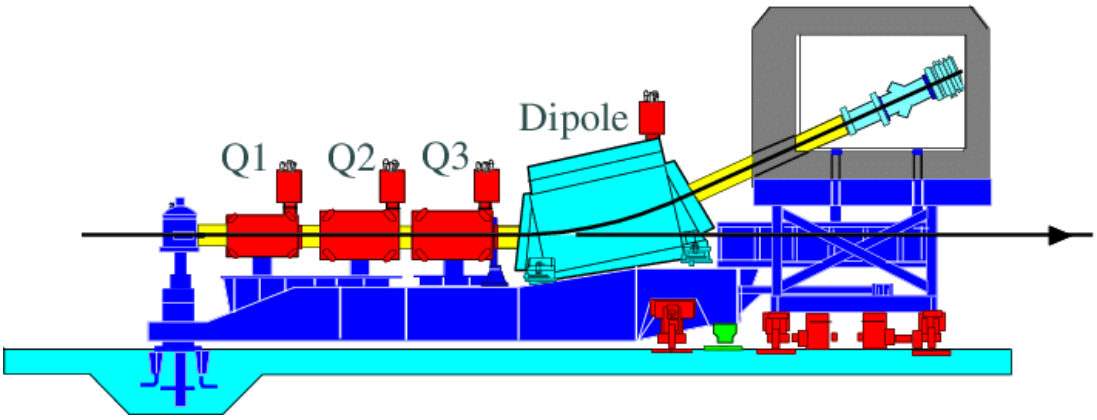


Figure 3.9: HMS side view

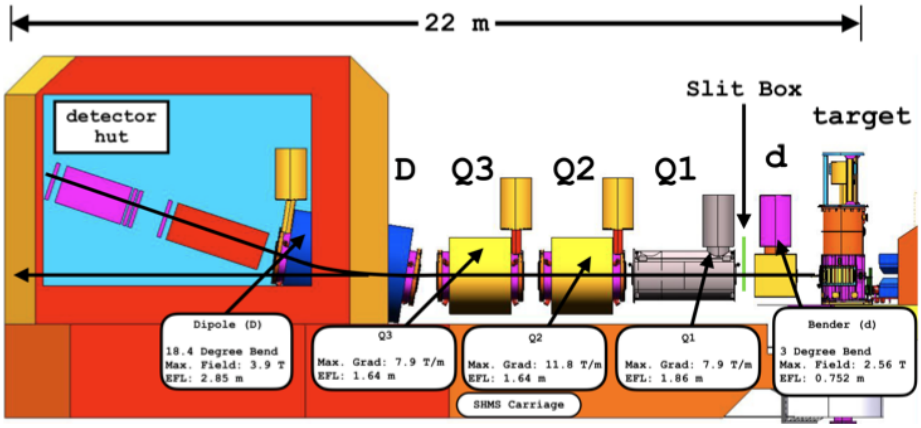


Figure 3.10: SHMS side view

3.4.2 Hall C Super High Momentum Spectrometer (SHMS)

The major upgrade of the Hall C 12 GeV era is the Super High Momentum Spectrometer (SHMS), as in Fig. 3.10. It replaces the old Short Orbit Spectrometer (SOS). The capability of the $11GeV/c$ momentum setting provides the ability to measure the secondary charged particles produced using the maximum available beam energy from the upgraded Jefferson Lab accelerator. It offers excellent kinematic control reproducibility, a thorough understanding of spectrometer acceptance, small-angle capability, high central momentum, and in-plane and out-of-plane acceptance well matched to the existing HMS to detect two particles in coincidence. The entire spectrometer is also mounted on a carrier that can rotate around the central pivot. The pions first go through a horizontal bender (HB) which decides the vertical acceptance and provides an initial 3° separation between scattered particles and the electron beam. It allows the spectrometer to measure small angle particles down to 5.5 °. It moves the subsequent pieces of the SHMS farther from the beamline, which allows more space for the following quadrupoles and shielding. The shielding, which is made of concrete, lead, boron, and plastic, reduces the background flux seen by the detectors. The first quadrupole focuses on the vertical direction while the second quadrupole defocused, and the third quadrupole focuses on maximizing the acceptance in the vertical scattering plane. The following dipole magnet bends the trajectories up by 18.4° into the detector hut. All of the magnets are set by the current. The whole trajectory for particles to travel in the SHMS spectrometer is 22 m. This length is critical to calculating different particles' travel times for separation. All the magnets are superconducting magnets that have a liquid cryogenic supply to

Range of central momentum	2-11GeV/c
Momentum deviation	-10% to 22%
Momentum resolution	0.03% to 0.08%
Scattering angle range	5.5° to 40°
Scattering angle plane angular acceptance	±18 mrad
Out of plane angular acceptance	±50 mrad
Scattering angle plane angular resolution	0.5-1.2 mrad
Out of plane angular resolution	0.3-1.1 mrad

Table 3.3: SHMS performance

keep them cool. The power supplies for magnets and the spectrometer rotation controls can be controlled remotely in the counting room. For our experiment, the scattered hadrons, the pions, are accepted into the spectrometer window after scattering inside the target chamber on the right side in Fig. 3.10.

3.5 Hall C Detectors

Two detector huts are attached at the end of each spectrometer magnet to record the particle information. The detectors are shielded to avoid the background from the hall. The configurations in the two spectrometers are similar, with differences in part to achieve particle triggering, tracking, and identification. The HMS detector hut (Fig. 3.11) starts with a pair of drift chamber planes for tracking purposes. The HMS aerogel detector is not installed during the experiment. Then the particles go through the first set of hodoscope planes, the Heavy Gas Cherenkov, and the second set of hodoscope planes. The hodoscope planes

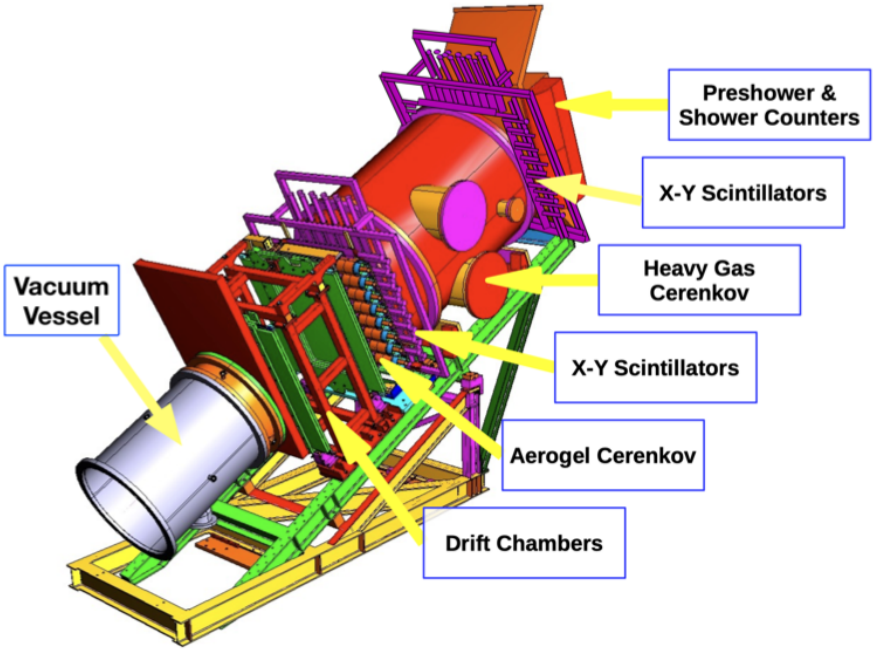


Figure 3.11: HMS detectors

provide trigger and timing information, and the Cherenkov detector can be used to identify different particles. Finally, the electrons deposit all their energy on the shower calorimeter. The SHMS detector hut (Fig. 3.12) starts with a noble gas Cherenkov detector, which is usually used for e/π separation; it was not installed during the fall run period but used by another experiment before the spring run period. Next are the drift chambers which provide tracking information. Then the two pairs of hodoscope planes are separated by heavy gas Cherenkov and aerogel Cherenkov detectors. The hodoscope planes provide triggering and timing, as in HMS. The heavy gas Cherenkov is designed for π/K separation, and the aerogel Cherenkov is designed for proton separation. Then the hadron deposit part of its energy on the calorimeter, while the electron deposit all the energy.

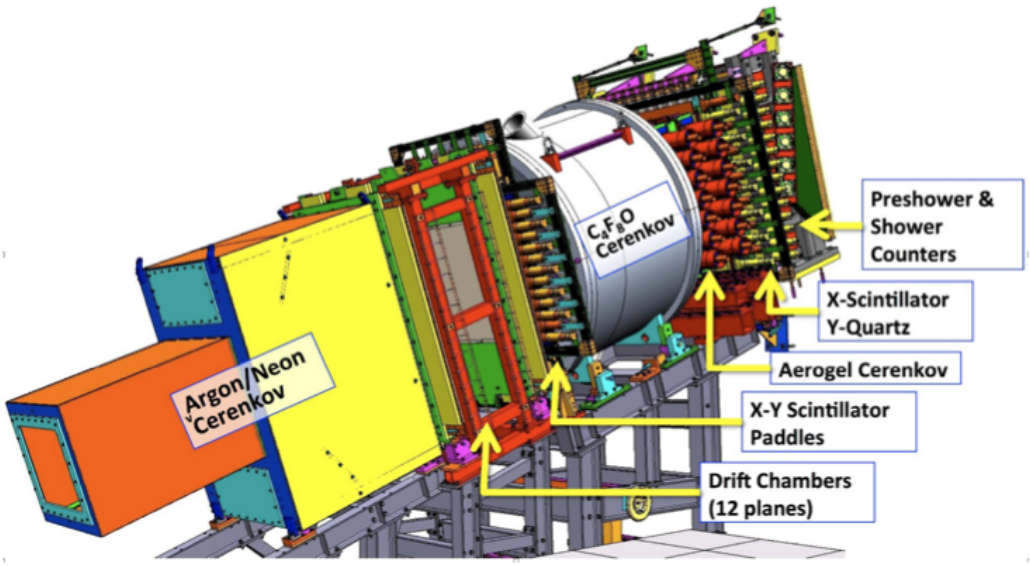
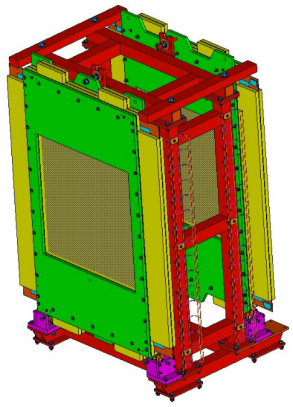
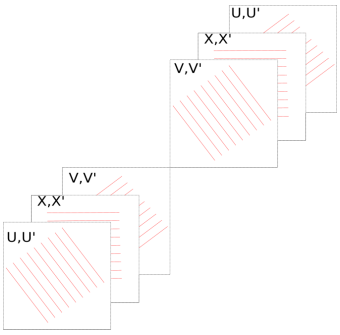


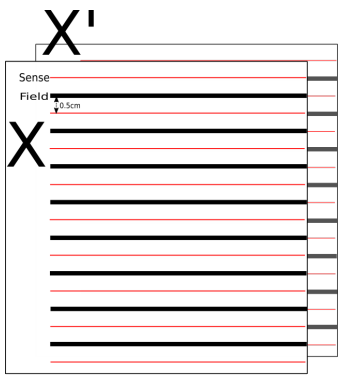
Figure 3.12: SHMS detectors



(a) SHMS Drift Chamber mounted in the detector hut frame



(b) Two chambers



(c) X and X' plane

Figure 3.13: Drift Chamber

3.5.1 Hall C Detector Drift Chamber

The drift chambers in both spectrometers are designed similarly, providing particles' position and angle information in the detector hut. The drift chambers are mounted on a frame, and the two-chamber packages are separated by about one meter (Fig. 3.13a). The center of the two planes is the "imaginary" focal plane. The two identical chambers have six planes of wires each (Fig.3.13b), between which are cathode planes of copper-coated mylar. Each wire plane consists of alternating anode wires and field wires. In planes labeled X and X', the wires are horizontal. In the U and U' planes, the wires are rotated 60°, and $-60°$ for V and V' planes. The primed planes are offset from the unprimed planes by 0.5 cm. The second chamber has an identical design but different direction.

The drift chambers are filled with mixed Ethane/Argon gas. When charged particles pass through the gas, it ionizes and creates more electrons. The induced particles can then ionize more and result in an avalanche. When the avalanche reaches the anode, it creates a measurable current. The drift time for the electron to the anode can be converted into the distance of the source particle from the anode. With the information from all the wires planes from two drift chambers, the particle position and direction are recorded.

The SHMS and HMS have similar drift chambers, which are 1 meter apart with an "imaginary" focal plane in the middle. The HMS drift chambers cover an active area of $100 \times 52cm^2$, and the SHMS drift chamber's operational area is $80 \times 80cm^2$. The active area is decided by the particles in the focal plane.

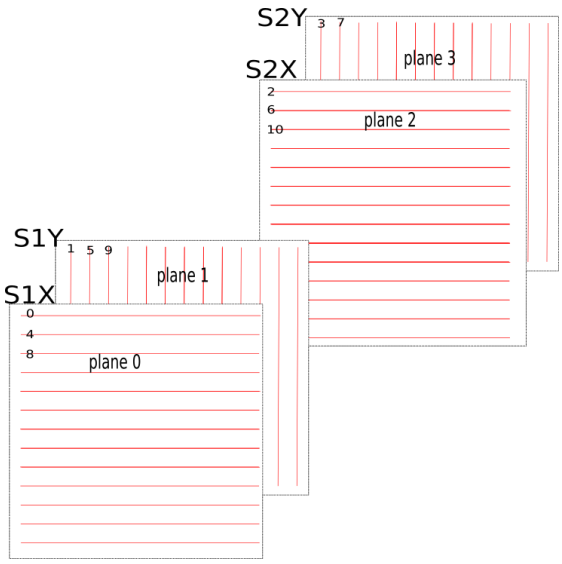


Figure 3.14: Hodoscope planes sketch.

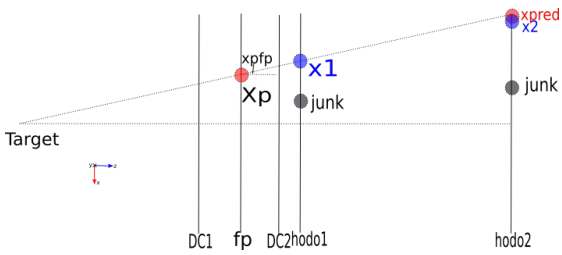


Figure 3.15: Hodoscope pruning process. The drift chamber will record the position and direction information for each track. Each point at the first hodoscope plane will be predicted to arrive at a specific area on the second hodoscope plane. Tracks outside the reasonable region will be pruned.[35]

3.5.2 Hall C Detector Hodoscope

The hodoscope system provides trigger and trigger time information. [72][36]

HMS Hodoscope

Two pairs of hodoscope planes are used in HMS for triggering information. The first set of hodoscope planes, S1X and S1Y, are placed after the drift chamber. The x planes are made of sixteen paddles, while the y planes are made of ten paddles. Each paddle has a photomultiplier tube (PMT) attached at each end. The second pair of hodoscope planes is 2.2 m away, with heavy gas Cherenkov detector in the middle. All four planes of HMS are made of plastic scintillating materials. When the charged particle propagates through the luminescence material, it emits light.

The emitted photons become measurable electrical pulses at the photomultiplier tubes attached to each paddle. The trigger that fires at least three out of four hodoscope planes is called HMS hodoscope 3/4 trigger. The 3/4 trigger rate would be the hodoscope 3/4 rate.

SHMS Detector Hodoscope

Triggering information comes from four planes of scintillator or quartz-bar hodoscopes. The hodoscope comprises four separate planes of detector paddles: S1X and S1Y, located immediately after the second drift chamber, and S2X and S2Y, approximately 2.2 m away along the z-direction. The first three are made of bars of scintillating plastic, and the fourth is synthetic quartz. The SHMS scintillator plane is comprised of 13 paddles in S1X, 13 paddles in S1Y, and 14 paddles in S2X. Each paddle has a PMT with negative high voltage attached to the two ends. The fourth plane, S2Y, is comprised of 21 bars of quartz, each with a length of 125 cm, a width of 5.5 cm, and a thickness of 2.5 cm. Each quartz bar is read out by a positive high voltage powered PMT at the two ends. The quartz bar only emits light from Cherenkov radiation which provides clean detection of charged particles. The trigger that fires at least three out of four hodoscope planes is called SHMS hodoscope 3/4 trigger. The 3/4 trigger rate would be the hodoscope 3/4 rate.

3.5.3 Hall C Cherenkov Detector for Particle Identification

Cherenkov Detector in spectrometers is used to separate different particles depending on the Cherenkov effect. It contains a radiator through which the charged particle passes and a photodetector to measure the radiated light. When

a charged particle travels faster than the light in the medium (group velocity $\frac{c}{n}$, c is the speed of light, and n is the refraction index of the medium), it will cause a flash of light, a Cherenkov light as in Fig. 3.16. Hall C spectrometers have several different Cherenkov detectors with different refraction indexes to separate different particles.

The threshold for the Cherenkov effect is

$$v_p \equiv \beta c > \frac{c}{n} \tag{3.1}$$

where n is the refraction index of the material in the Cherenkov tube. The velocity $\beta = p/\sqrt{m^2 + p^2}$ where m is the particle mass. The threshold reflection index for a given particle and momentum is :

$$n > \frac{\sqrt{m^2 + p^2}}{p} \tag{3.2}$$

The emitted light waves travel at light speed c/n, and the emission angle results in

$$cos(\theta_c) = \frac{1}{n\beta} \tag{3.3}$$

The number of emitted photons detected by the PMTs designed in the Cherenkov detector is related to the path length (L), PMT position, material transparency, etc. The number of photoelectrons can be written as

$$N_{p.e.} = L\frac{\alpha^2 z^2}{r_e m_e c^2}\int \epsilon(E) sin^2\theta_c(E) dE \tag{3.4}$$

where $\epsilon(E)$ is the efficiency for collecting the Cherenkov light and converting it into photoelectrons. The number of photoelectrons can be written as Eq.3.5 for the designed Cherenkov Detector in Hall C with known central momentum. The N_0 depends on the condition, such as the mirror alignment, PMT quantum efficiency,

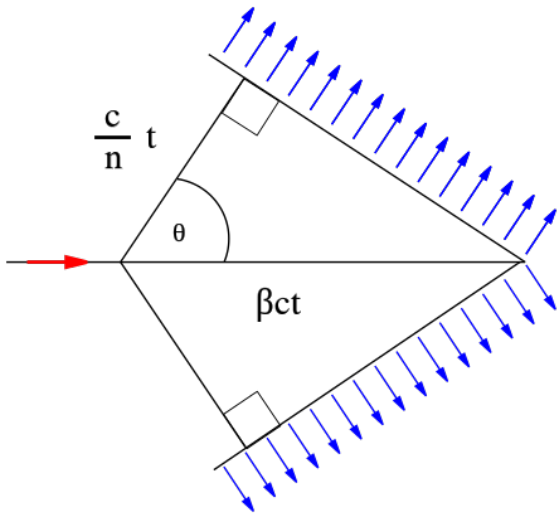


Figure 3.16: Cherenkov effect. The charged particle (red) going faster than the velocity of the light in the medium will cause a flash of light (blue). This is similar to the sonic boom but for the light.

reflection length, etc.

$$N_{p.e.} \approx LN_0 < sin^2\theta >= LN_0(1 - \frac{p^2 + m^2}{n^2p^2}) \tag{3.5}$$

HMS Gas Cherenkov Detector

The HMS gas Cherenkov detector consists of a 1.5 meter-long cylinder tank between the two hodoscope planes. The HMS gas Cherenkov detector is filled 0.45 atm C4F8O for e/π separation. The Cherenkov lights are reflected at the two mirrors and detected by the two PMTs.

SHMS Heavy Gas Cherenkov Detector

The SHMS Heavy Gas Cherenkov detector consists of a large cylindrical tank with an outer flange diameter of 1.88 m and a length of 1.3m. The tank contains four reflecting mirrors, coupled with four PMTs as in Fig3.18a. The tank can be filled with different gas with different pressure for different particle separations as in Fig.3.17 For this experiment, the HGC detector is supposed to do π/K separation. The tank is filled with C_4F_8O radiator gas at 1 atm, index of refraction 1.00137, which has a pion threshold $2.8GeV/c$ and kaon threshold $9.4GeV/c$. The

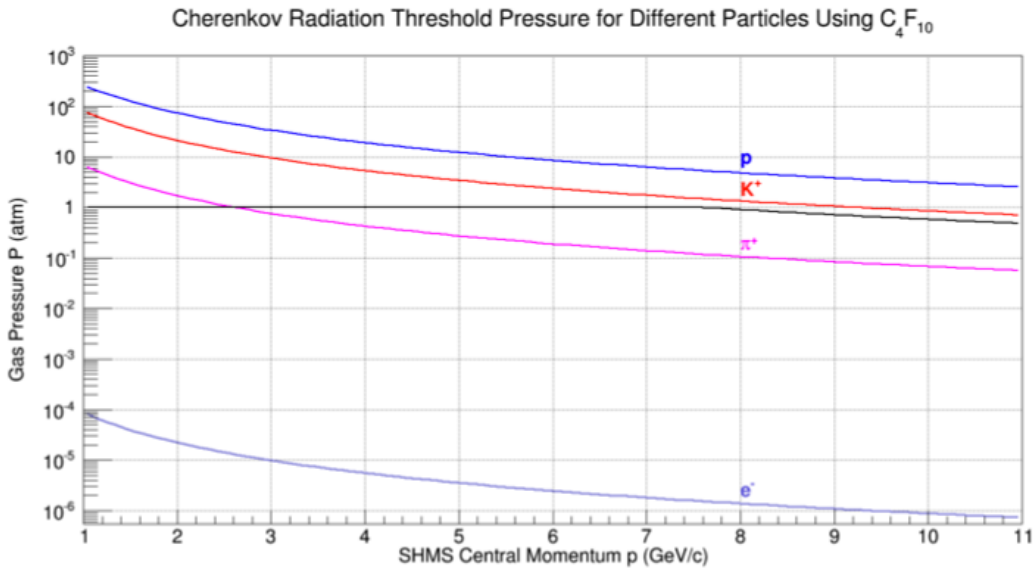
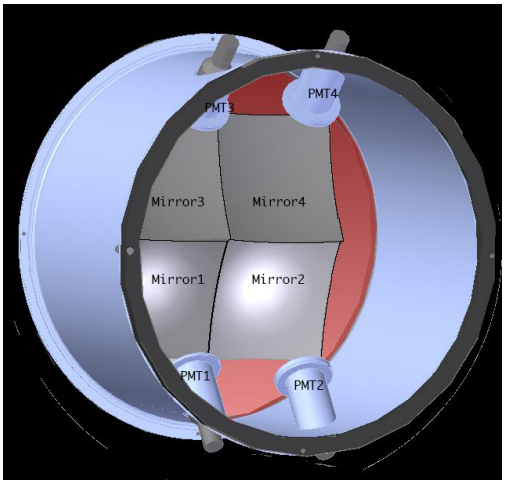


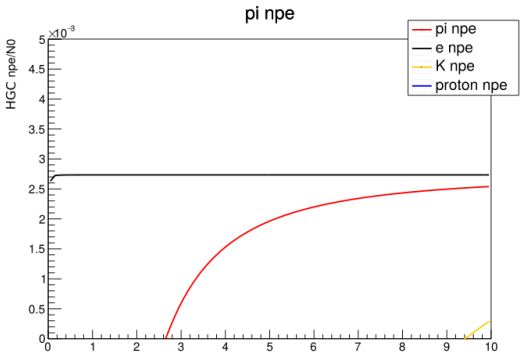
Figure 3.17: Gas Pressure vs. Particle Momentum. The black curve represents the suggested gas pressure (which leads to the reflection index) of the HGC detector for good π/κ separation. The colored curves represent the Cherenkov threshold pressure for different particles. C_4F_8O performs equivalently to C_4F_{10} in this context. [88]

number of photoelectrons as a function of momentum as calculated from the equation 3.5is shown in Fig. 3.18b, the y-axis unit is arbitrary and depends on tank conditions. The horizontal black line represents the number of photoelectrons produced by electrons. Since the mass of the electron can be ignored, the npe for the electron barely changes with momentum. The red line represents pions and the yellow line represents kaons. The proton threshold is way beyond the experiment's central momentum setting. The more energy a hadron contains, the larger momentum, and more photoelectrons are produced.

SHMS Aerogel Cherenkov Detector The Aerogel Cherenkov is placed after the Heavy Gas Cherenkov detector and before the hodoscope planes. The detector consists of a tray that holds the aerogel material and a light diffusion box with PMTs for light readout. The SHMS aerogel detector is filled with SP15 with a reflection index of 1.015. From the equation3.5, the pion threshold is 0.81 GeV/c, 2.84 GeV/c for kaon, and 5.4 GeV/c for protons. The proton threshold of



(a) SHMS HGC detector



(b) SHMS heavy gas Cherenkov detector number of photoelectrons versus momentum.

Figure 3.18: SHMS HGC Detector

5.4 GeV/c is higher than the momentum range of SHMS used in the experiment, so the protons are not supposed to radiate any light in the aerogel Cherenkov detector. However, knock-on electrons are produced by ionization or excitation of the constituent atoms when the proton passes through the aerogel. The knock-on electron may generate some light in the aerogel detector. To eliminate the protons, the aerogel detector cut needs to be set to be greater than 4 photoelectrons.

3.5.4 Hall C Lead-glass Calorimeter

Both spectrometers have a lead-glass electromagnetic calorimeter at the end of the detector hut since it provides a destructive measurement of energy. The calorimeter is a supplementary particle identification detector. The particle will deposit energy by the production of electromagnetic showers in the radiator. The total amount of light radiated in this case is proportional to the energy of the primary particle. The electrons will lose their energy primarily via ionization at low energies, and mainly by bremsstrahlung at high energies. The EM showers of electrons are produced by bremsstrahlung photons that decay to e^+e^- pairs via

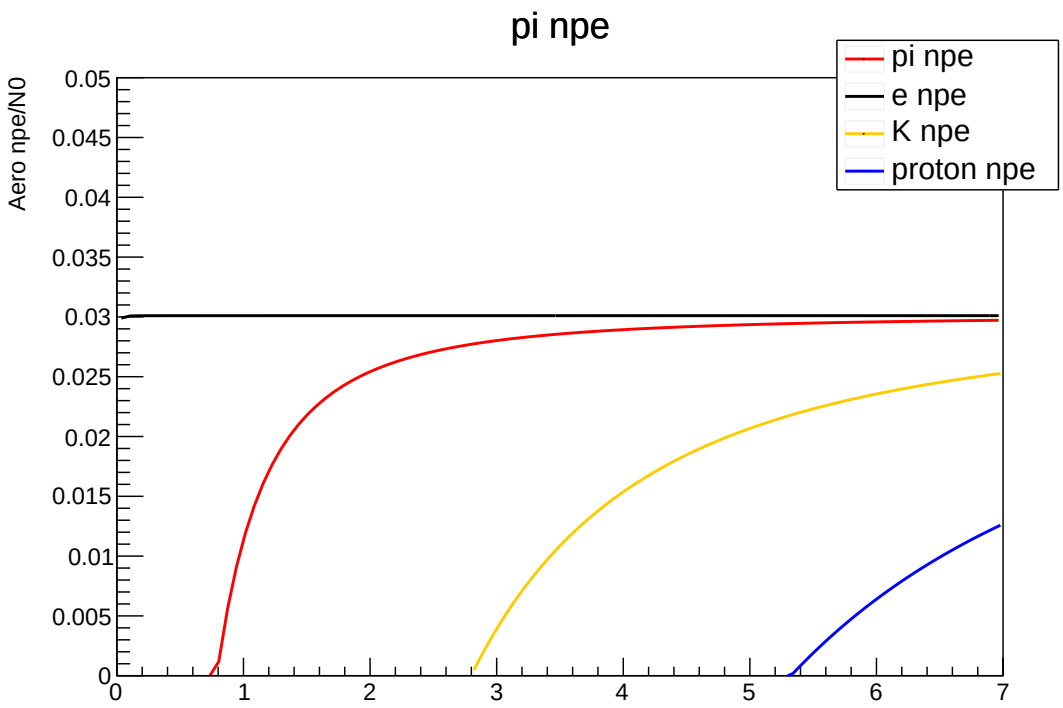


Figure 3.19: SHMS aerogel Cherenkov detector number of photon electrons versus momentum.

pair production. The pairs further radiate bremsstrahlung photons and create an EM shower cascade reaction until most of the energy is deposited on the calorimeter as in Fig3.20. The electron enters the radiator with initial energy E_0 and loses energy by producing electromagnetic cascades. The radiation length (X_0 g/cm^2), is the mean distance over which a high-energy electron loses all but 1/e of its energy by bremsstrahlung. Charged hadrons have a low probability to interact and produce a shower, and may pass through without interaction. In this case, they will deposit a constant amount of energy in the calorimeter [55][81].

HMS Calorimeter

The HMS calorimeter locates at the end of the detector hut, which is made of 4 layers of TF-1 lead glass modules. Each of the block is $10 \times 10 \times 70cm^3$ large. The refraction index is 1.65, with a density of 3.86 g/cm^3 which yields 3.65 radiation lengths per layer. First, two layers have two PMTs attached at each

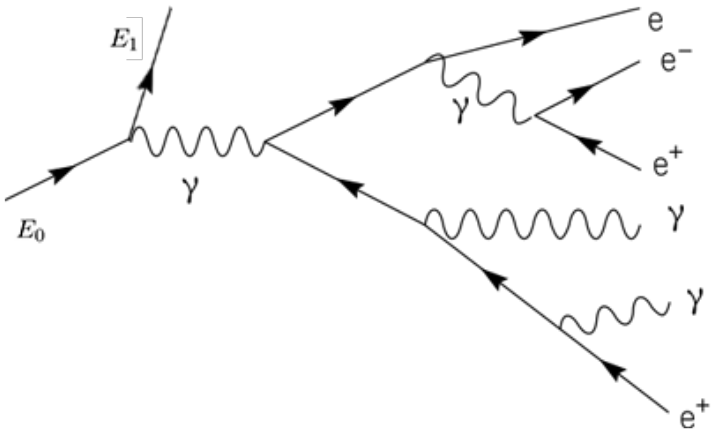
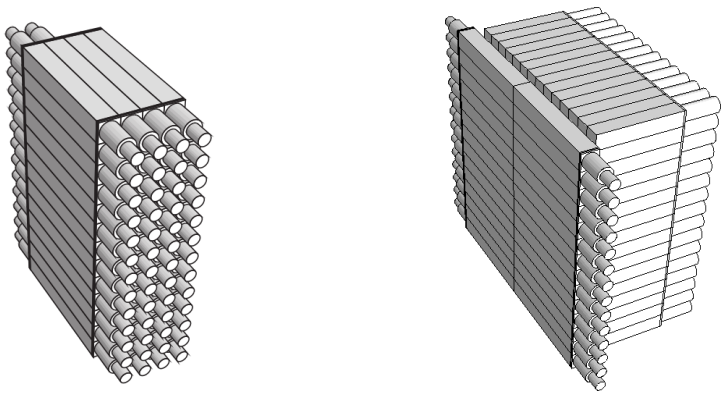


Figure 3.20: Sketch of the electromagnetic shower. Reproduced from wiki sketch.



(a) HMS Calorimeter (b) SHMS Calorimeter

Figure 3.21: Hall C Calorimeters

end while one PMT is attached at one end for the rest layers. To prevent shower leakage between the blocks, the HMS calorimeter is tilted by 5 degrees relative to the central trajectory.

SHMS Calorimeter

The SHMS lead glass shower calorimeter is the last detector which locates behind the S2Y hodoscope plane. It is made of two parts, pre-shower and shower counter as in Fig. 3.21b. The pre-shower layer is made of 2×14 TF-1 $10 \times 10 \times 70cm^3$ lead glass. The shower part is made by a 14×16 array of F-101 type

Effective area	$70 \times 130cm^2$
Total thickness	$\approx$ 14.6 rad. length
Dynamic range	0.5-7.3 GeV/c

Table 3.4: HMS calorimeter performance

Effective area	$120 \times 140cm^2$
Total thickness	$\approx$ 20 rad. length
Dynamic range	1.0-11.0 GeV/c
Energy resolution	$\approx 6\%/\sqrt{E}, EinGeV$

Table 3.5: SHMS calorimeter performance [55]

$9 \times 9 \times 50cm^3$ lead glass. Each lead glass is coupled with one PMT at the end. The combination of pre-shower and shower blocks is $\approx$ 20 radiation length, which will absorb $\approx 10GeV$ energy of the electrons. This is enough to measure all the energy of the electron under the full range of the SHMS setting. The general requirements for the SHMS calorimeter are in the table. 3.5.

3.6 Data Acquisition

The experiment will generate a large number of signals from every detector. This information is managed by the data acquisition (DAQ) system[84][33]. The data acquisition system converts the raw data to an appropriate format to be stored. Hall C uses the CEBAF Online Data Acquisition system (CODA) developed by the JLab data-acquisition group. The signals generated by the PMTs or wires are digitized by Analog to Digital Converters (ADCs), Time to Digital Converters (TDCs), or both. The digitized signals are read out by the corresponding

Read Out Controller (ROC) module. CODA interfaces with the Readout Controllers(ROCs) and processes signals from all the electronics.

Furthermore, there are also crates for the *scaler*. The counts, time, charge, current, *etc* are saved in the scaler tree in root files.

Also, the Experimental Physics and Industrial Control System (EPICS)[1] is a control system that controls devices and slow readout parameters. It is designed to interface remotely with detectors and devices. The EPICS system records many scaler parameters like the spectrometer momentum, angle, target pressure and temperature, magnets probes readout *etc*. The values read out from EPICS are saved in raw data files.

3.7 SIMC

SIMC is a standard Hall C Monte Carlo for coincidence reactions written in FORTRAN. The events are generated randomly and then weighted by the cross section and the spectrometer acceptance. It was developed from the Single arm Monte Carlos for the coincidence experiments. It includes models for several physics processes and also includes radiative effects[76]. Several reactions are implemented in SIMC:

- elastic and quasielastic $H(e, e'p), A(e, e'p)$.
- Exclusive pion production $H(e, e'\pi^+)n, A(e, e'\pi^\pm)$
- Kaon electroproduction
- Semi-Inclusive DIS pion production $H(e, e'\pi^\pm X, D(e, e'\pi^\pm)X$

- Semi-Inclusive DIS Kaon production $H(e, e'K^{\pm}X), D(e, e'K^{\pm})X$

- Diffractive ρ $H(e, e'\rho \hookrightarrow \pi^+\pi^-)p, D(e, e'\rho \hookrightarrow \pi^+\pi^-)p,$

The SIMC will read the input files. The input file indicates what reaction to use, what final states to expect, the beam, and target information *etc.* It can also enable or disable implementation or correction for raster, radiative correction *etc.* The SIMC process starts with generating events. The electron is generated with energy and position within the energy and spatial spread of the beam. Randomly generates the momenta and angles of the scattered electron and the hadron with flat distribution over a region larger than the actual spectrometer acceptance. This is the basic scattering vertex. Depending on the input file, the radiative process might be included, which allows all particles to lose some energy. The scattered electron and hadron are then transported through the spectrometer model. The spectrometer models contain detailed optics, apertures, and multiple scattering in the interfering material of the spectrometer. The particle is transported through the target chamber and spectrometer magnets and into the detector hut model. The particle events that passed all apertures successfully are recorded. Since it's imaginary spectrometers and detectors, there is no inefficiency applied. However, each successful event is assigned with a 'weight' related to the cross section, corrected by the luminosity factor and Jacobian transformation.

In this experiment, the elastic models are used to compare elastic scattering from data for momentum setting checks. The Semi-Inclusive DIS pion production is used with all the possible energy loss for radiative correction. The DIS cross section is from the latest model - F1F221 from Eric M. Christy (to be published). For the SIDIS cross section, the parameterization is from the latest fit of Peter

Bosted. The exclusive pion production, radiative delta resonance, and diffractive ρ are used for radiative background subtraction.

Chapter 4

DATA ANALYSIS

4.1 Analysis Overview

All electronics' raw signals are collected and stored in CODA files. Then the Hall C analyzer hcana processes and convert the raw data file to ROOT files. Hall C analyzer, hcana[15], is a C++-based event analysis tool to analyze data from the spectrometers in Hall C using the CERN ROOT framework. The analyzer has all information related to the detectors. It replays all the data for each run and generates root files. The parameter files for the different detectors are updated by calibration from some calibration runs. With the corresponding calibrated parameter files, the events are better analyzed. Then different cuts are applied to clean the particle samples. The efficiency study is related to each of the cuts to get accurate statistics. Some background particles that cuts can not simply eliminate are subtracted in other ways to get accurate counts per charge, which is the yield. The charge symmetry violating term can be extracted with the efficiency corrected yield.

4.2 Kinematics

During the CSV run period, twelve kinematics (x_{bj},Q^2) are done for the π^- to π^+ production yields on deuterium with four different center z settings each. This experiment measures semi-inclusive pion electroproduction from deuterium at large x to map out potential charge symmetry violation in the valence quark distributions. The kinematics were chosen with three different Q^2 values: 4.0, 4.75 and 5.5 GeV^2. At each Q^2 value, several x from 0.35 to 0.65, and for each different Q^2, x, four z from 0.4 to 0.7 are taken. For each different z, two polarities of hadron (π^+ and π^-) are taken. For each of the different (x, Q^2,z, polarity), several D_2 runs are used for data taking, and several aluminum(Dummy) runs are taken for background subtraction (as illustrated in Section. 4.4.3). The kinematic phase space is shown in Fig. 4.2, where the x-axis is Bjorken x and the y-axis is Q^2. Several center x settings are taken for each Q^2, which would overlap with each setting. With the coverage, one can extract $R^D_{measure}(x, z)$ at each kinematic setting. Also, at $Q^2 = 4.0$ and 4.7 GeV^2, additional data from liquid hydrogen is combined with deuterium data to perform simple factorization tests.

The experiment was completed in two parts. The first part was performed in November 2018, which finished the smaller part of Q^2 settings including all liquid hydrogen runs, and got enough statistics. The second part was conducted in March 2019, which finished the rest of the planned kinematic settings. For the experiment, run numbers are assigned by CODA at the beginning of each run. The run number ranges from 6009 to 6551 for the fall run period and from 7593 to 7830 for the spring run period. The experiment summary is attached here csv.

For convenience, runs with the same kinematic settings are sorted into one run

RunGroup	pos D_2	neg D_2	pos Al	neg Al	pos H_2	neg H_2
150	6194 6195	6245 6246 6247	6196	6242 6243 6244	6192 6193	6237 6240 6241

Table 4.1: One fall RunGroup

RunGroup	pos D_2	neg D_2	pos Al	neg Al	pos H_2	neg H_2
530	7808 7809	7796 7797	7811	7793 7794	nan	nan

Table 4.2: One spring RunGroup

group. RunGroup number increasing by 10 each is assigned for each run group in time order. The fall runs have RunGroup numbers from 10 to 420 in time order, and the spring runs have RunGroup numbers from 430 to 580. Hydrogen runs are only taken for some of the kinematic settings.

In addition, kinematic cuts are applied to ensure SIDIS events selection. The usual “DIS" kinematics such that $Q^2 > 1GeV^2$ and $W^2 > 4GeV^2$. Also, SIDIS requires the mass of the unobserved final state W'^2 (Equation 2.23) to be larger than 2.5 GeV^2. This last constraint is driven by the observation from E00-108 that factorization seems to hold up to z = 0.7 for their kinematics, $W'^2 > 2GeV^2$.

4.3 Spectrometer Momentum Correction

For this experiment, HMS is used for electron detection, and the newly installed SHMS is used for hadron detection. The HMS dipole has a nuclear magnetic

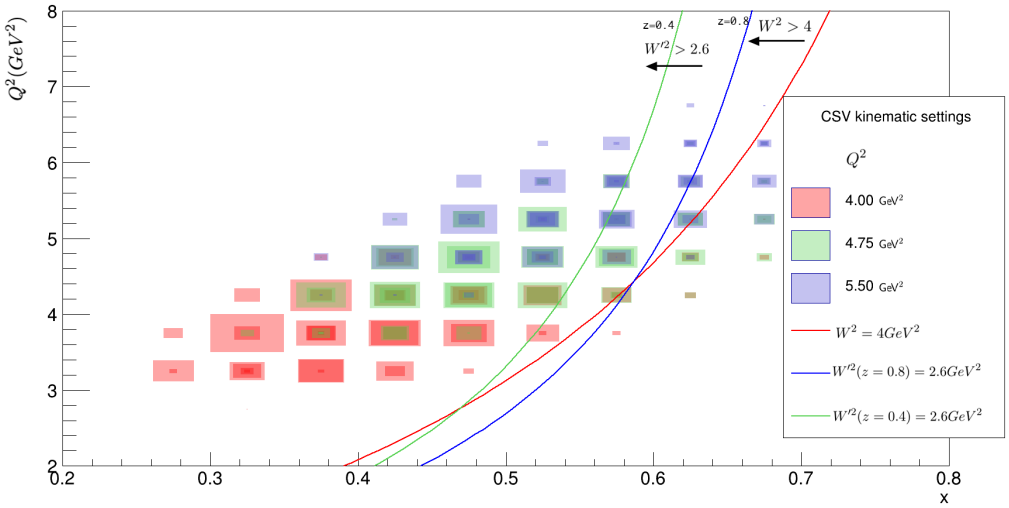
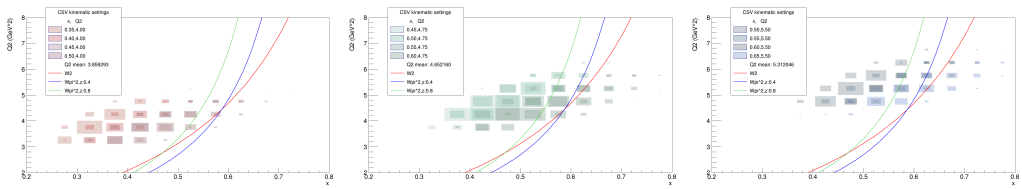


Figure 4.1: Kinematic coverage Q^2 vs. x for all three Q^2



(a) 4 GeV^2 (b) 4.75 GeV^2 (c) 5.5 GeV^2

Figure 4.2: Kinematic coverage for three distinct Q^2 region

Run Number	P_{HMS} (GeV)	θ_{HMS} (deg)	P_{SHMS} (GeV)	θ_{SHMS} (deg)
6009	6.358	21.06	5.414	23.63
6010	6.358	19.78	5.09	25.01
6012	6.358	18.56	4.768	26.445
6013	5.983	21.29	5.471	23.39
6016	5.714	22.43	5.744	22.29
6017	5.27	24.42	6.193	20.58
7822	5.714	23.22	5.65	22.25
7823	5.714	23.22	5.65	22.25
7826	5.714	21.93	5.36	23.45
7827	5.714	21.93	5.36	23.45
7830	5.714	20.71	5.07	24.71

Table 4.3: Hydrogen elastic runs are taken during the experiment

.

resonance (NMR) probe for field measurement. During the 12 GeV upgrade, the probe is removed and replaced. Additionally, the experiment runs at higher momentum settings than in the 6 GeV era. At a high central momentum setting above 5 GeV/c, the magnets start to saturate. Several hydrogen elastic runs are taken during the experiment to determine the offset for momentum as in Table 4.3.

The elastic scattering diagram is in Fig.4.3. In this process, the kinetic energy of a particle is conserved in the center-of-mass frame. The scattered electron energy can be calculated:

$$E' = \frac{E_b}{1 + 2\frac{E_b}{M_P} sin^2(\frac{\theta}{2})} \tag{4.1}$$

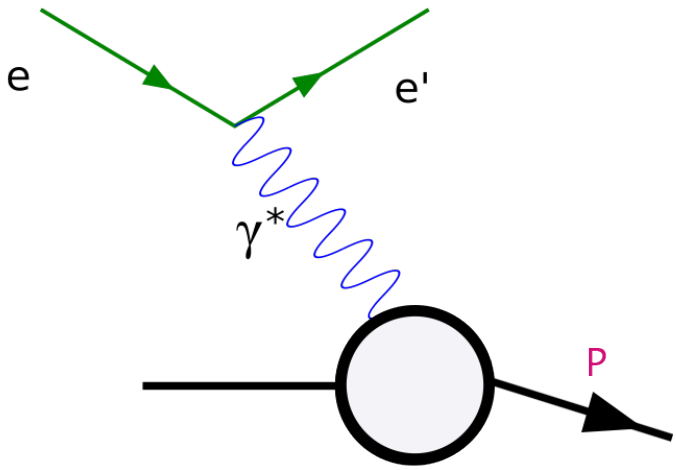


Figure 4.3: Hydrogen elastic scattering. The incoming electron k(E_b,$\vec{p}$) scatter off a proton target at rest P(M_p,0), resulting in the scattered electron $k'(E',\vec{p'})$ and proton $P'(E_p,\vec{p_p})$

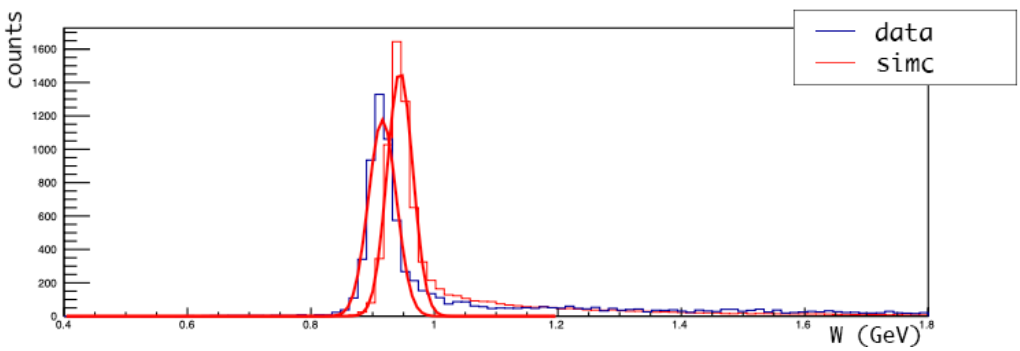


Figure 4.4: W peak shift for data and SIMC. A Gaussian fit is done to determine the peak center.

where θ is the scattering angle. The invariant mass W as in Eq.2.6 for elastic scattering equals the proton mass (M_p).

$$W^2 = (P + k - k')^2 = M^2 + 2M\nu - Q^2$$
$$\nu = E_b - E' \tag{4.2}$$
$$Q^2 = 4E_bE'sin^2(\frac{\theta}{2})$$

The invariant mass is a function of the electron beam energy, scatter angle and scatter momentum. A special beam energy measurement verifies electron beam energy. The scattering angle is verified by checking the photo of the rotation angle of the spectrometer for each run. The momentum offset is determined by the shift of the W peak between data and simulation. The simulation used here is

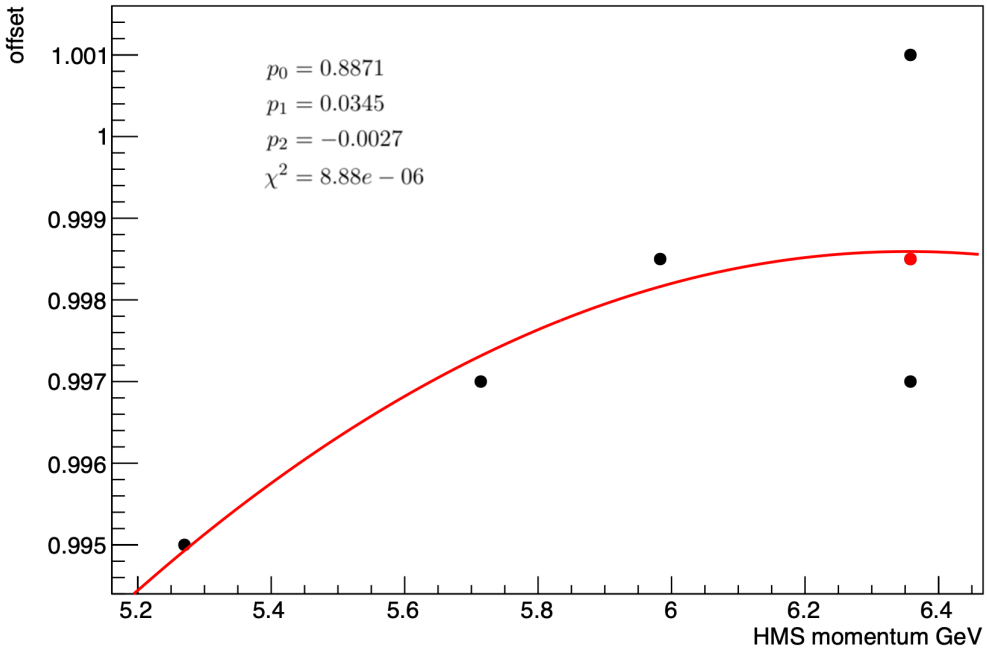


Figure 4.5: HMS momentum offset as a function of center momentum

from SIMC, a Monte Carlo simulation program, which takes the input kinematic setting and generates the corresponding events. The output W peak from SIMC is the same as M_P as in theory as in Fig.4.4. The momentum offset applied to the spectrometer could shift the W peak to where it should be.

Different runs with broad momentum settings are taken to cover the momentum range of the experiment as shown in Table. 4.3. The W peak shift is determined for each momentum and each spectrometer separately. The momentum offset calculated from the W peak shift for each momentum for HMS is in Fig.4.5. A simple polynomial fitting is applied to the offset plot.

More elastic runs are taken before and after the experiment. More detailed studies are done using all the elastic runs [79][80]. The momentum offset used in this experiment is from the detailed study. Momentum offset for HMS is $-0.000276P_{HMS}^3 + 0.002585P_{HMS}^2 - 0.008697P_{HMS} + 1.006440$, for SHMS is 0.998.

4.4 Event Selection

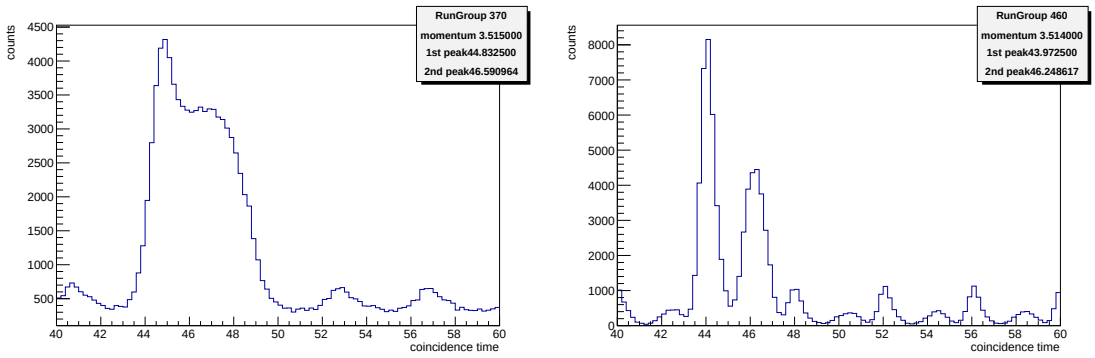
4.4.1 Coincidence Time Cut

The scattered electrons and produced pions should be measured simultaneously in two spectrometers for the SIDIS experiment. The time is calculated using the time difference between two spectrometer hodoscope planes. For a specific spectrometer momentum setting (P_{SHMS}), the total hadron drift length is the SHMS central length from the target to the hodoscope plane. The drifting time for the different particles is related to their mass:

$$\beta = \frac{P_{SHMS}}{\sqrt{P_{SHMS}^2 + M_{pion}^2}}, \tag{4.3}$$

$$t_{corr} = \frac{SHMS_Length}{c\beta} \tag{4.4}$$

where the β is the velocity for pions,and t_{corr} is time correction for pions. The pion time correction is added to the raw coincidence time. The kaons(K^+,K^-) and protons can also enter the SHMS acceptance and form triggers in the spectrometer. Kaons are slightly heavier than pions, so kaons' travel time is slightly slower than pions. As calculated from Eq.4.3 and Eq. 4.4, kaons are at most 2 ns slower than pions at the experiment momentum range. Kaons' peak is mixed with the coincidence time peak, which needs to be separated in another way. Protons are the heaviest(slowest) particle to enter the SHMS acceptance. It will form a proton peak on the right side of the pion peak. By cutting on coincidence time, the pion peak is selected. However, the proton coincidence time peak at higher momentum is closer to the mean pion peak. The protons might be included in the coincidence time cut, but other cuts will remove them.



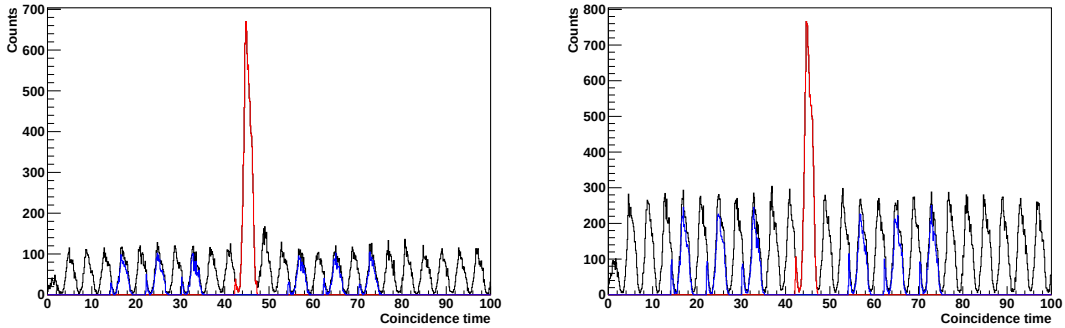
(a) One of the fall runs

(b) One of the spring runs

Figure 4.6: Coincidence time peak for one fall run and one spring run with similar central momentum setting and same cuts. Spring runs have a better coincidence time resolution.

In the summer of 2018, the reference time delay in the TDC was changed from using cables to using a gate generator. However, a mistake is made that the reference signals at the TDCs in the counting house are inverted. The polarity of the reference signal was the wrong polarity going into the TDC, so the trailing edge was setting the timing. The trailing edge has some jitter which causes the low resolution of the coincidence time peak[34]. This problem is researched and fixed after the fall run period[71]. The spring runs get better coincidence time resolution. The coincidence time cut used in this experiment is 5 ns with 2.5 ns on each side of the pion coincidence time center for fall runs and 3 ns with 1.5 ns on each side for spring runs. Due to the poor coincidence time calibration in fall runs, the pion peak and proton peak are broader than spring runs for the same SHMS central momentum as in Fig.4.6. Other cuts will separate the mixture of the main and proton peaks.

The coincidence spectrum for the same RunGroup but the different polarity is shown in Fig.4.7. For the positive polarity, the protons form a peak slower than the pion peak, while there is only one mean peak for the negative polarity spectrum. The proton peak is skipped for the accidental backgrounds discussed



(a) π^+ runs (b) π^- runs

Figure 4.7: Coincidence time for two runs with the same kinematic setting but opposite polarity. Blue peaks on both sides are accidental background peaks after some PID cut which will be discussed in Section.4.4.2

in Sec.4.4.2.

4.4.2 Accidental Backgrounds Subtraction

Besides the scattering process, random events happen anytime during the experiment. The random events can be removed by other selection cuts, but some can't. They are not predictable, so the way to remove them is by "blank controller". The electron beam is a continuous wave with a 4 ns period. The accidentals are triggering anytime during the SIDIS scattering or not. So the adjacent peaks can be used as accidental backgrounds. Several selection types of background peaks are tested to get a better estimate of accidental backgrounds.

Different selections are taken:

- 1. Five background peaks each on both sides of the mean peak. On the right side, the adjunct two background peaks are skipped to remove the proton coincidence time peak contamination.
- 2. Ten background peaks on the left side of the mean peak
- 3. Ten background peaks on the right side of the mean peak

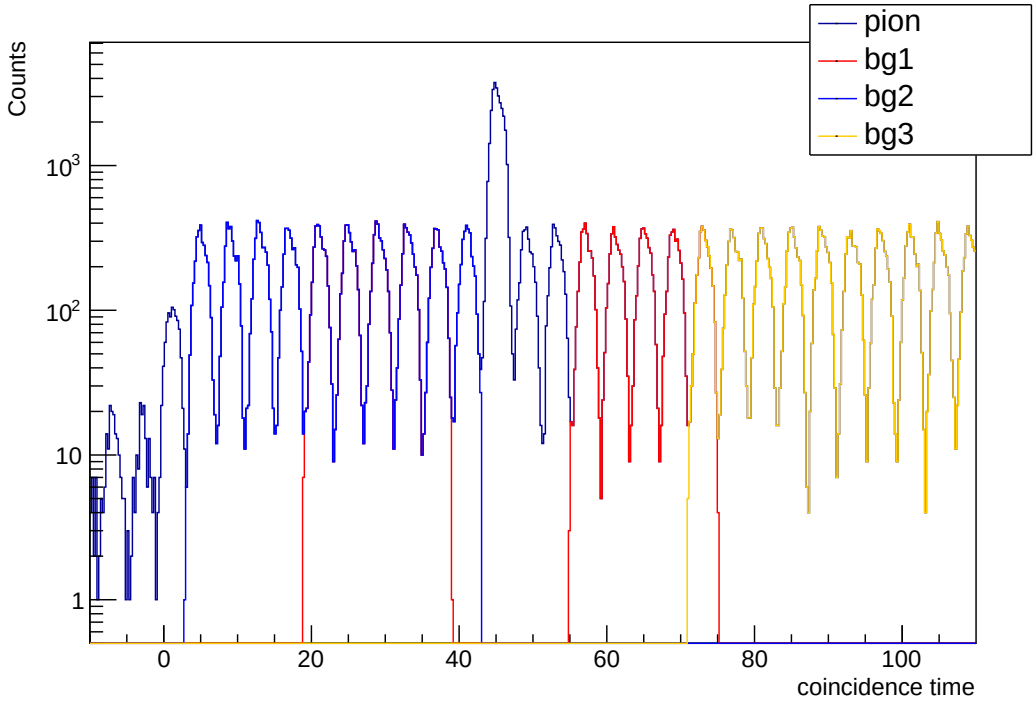


Figure 4.8: The different selections of accidental backgrounds

Several accidental background peaks are selected in the coincidence time spectrum. The beam period is 4 ns which applies to accidental peaks. All the chosen peaks are added together and then averaged to one period. There are no significant differences between different selections, as shown in Fig.4.9.

The accidental background peaks in this experiment are chosen to be three peaks on the left side and three on the right side of the pion peak. Since the fall runs have a broader coincidence peak than 4 ns, the accidental background peaks are taken to be the first, third, and fifth peaks on the left with the same width as the main pion peak. And on the right-hand side, the accidental peaks start from the third to reduce the background from the proton coincidence peak. The third, fifth, and seventh peaks are chosen to be the accidental peaks. All six peaks are normalized and subtracted from the mean pion peak.

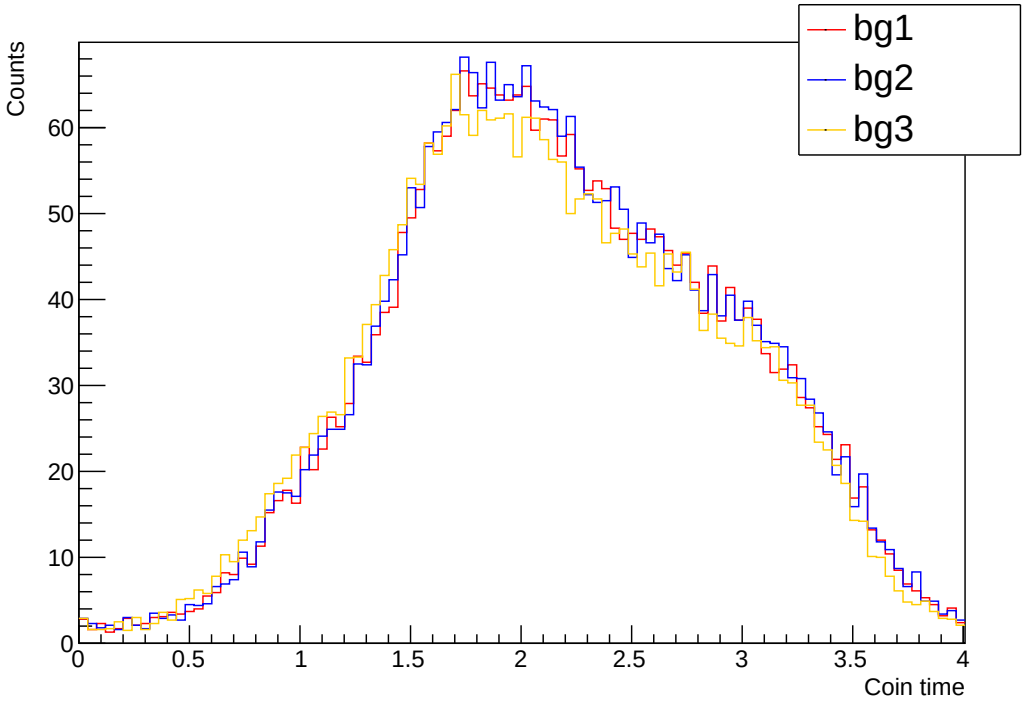


Figure 4.9: Compare different accidental backgrounds selection

.

Target	Entrance window(cm)	Exit window(cm)
LH2	0.015 ± 0.0011	0.0191 ± 0.0019
LD2	0.013 ± 0.0012	0.0188 ± 0.0013
Aluminum	0.06462 ± 0.0001	0.06459 ± 0.0001

Table 4.4: Target Entrance and Exit window wall thickness[49]

4.4.3 Aluminum Dummy Backgrounds Subtraction

The liquid Hydrogen and liquid deuterium are filled inside an aluminum(dummy) tube. When the electron beam arrives at the target tube, it will scatter on the liquid target and the target wall before and after. The target wall backgrounds are subtracted by measuring the scattering on an empty cell, the two aluminum walls at the same position as shown in Section3.3. Dummy runs are taken for every different kinematic setting and polarity. The H_2 and D_2 yield

is subtracted by the corresponding target wall yield.

$$Yield_{H_2/D_2,Dummysubtracted} = Yield_{H_2/D_2} - R_{H_2/D_2} Yield_{Dummy} \tag{4.5}$$

where R is the ratio of liquid target cell wall thickness divided by aluminum thickness:

$$R_{H_2 \text{ or } D_2} = \frac{H_2 \text{ or } D_2 \text{ wall thickness}}{\text{Dummy wall thickness}} \tag{4.6}$$

Calculated from Table 4.4, R_{H_2} is 0.262 and R_{D_2} is 0.245.

4.5 Particle Identification

The electrons and pions ($\pi+$ and $\pi-$) are supposed to be in coincidence and be as clean as possible. Different cuts are used to maintain the purity of the particle sample. However, the cuts from detectors are not perfect. There are still some unexpected particles mixed in the pure sample. Also, the cuts might accidentally prune some good events. Efficiency and purity are defined for different detectors to evaluate the performance.

4.5.1 Electron Detection Efficiency

The electron detecting efficiency in HMS can be described as the quality of the separation between electrons and other particles.

4.5.2 HMS Cherenkov Detector

For the HMS Cherenkov detector, both the electrons and pions fire the chamber. Due to the mass difference, the electron and pion have a different threshold and dependence on momentum as shown in eq.3.4. The number of photoelectrons

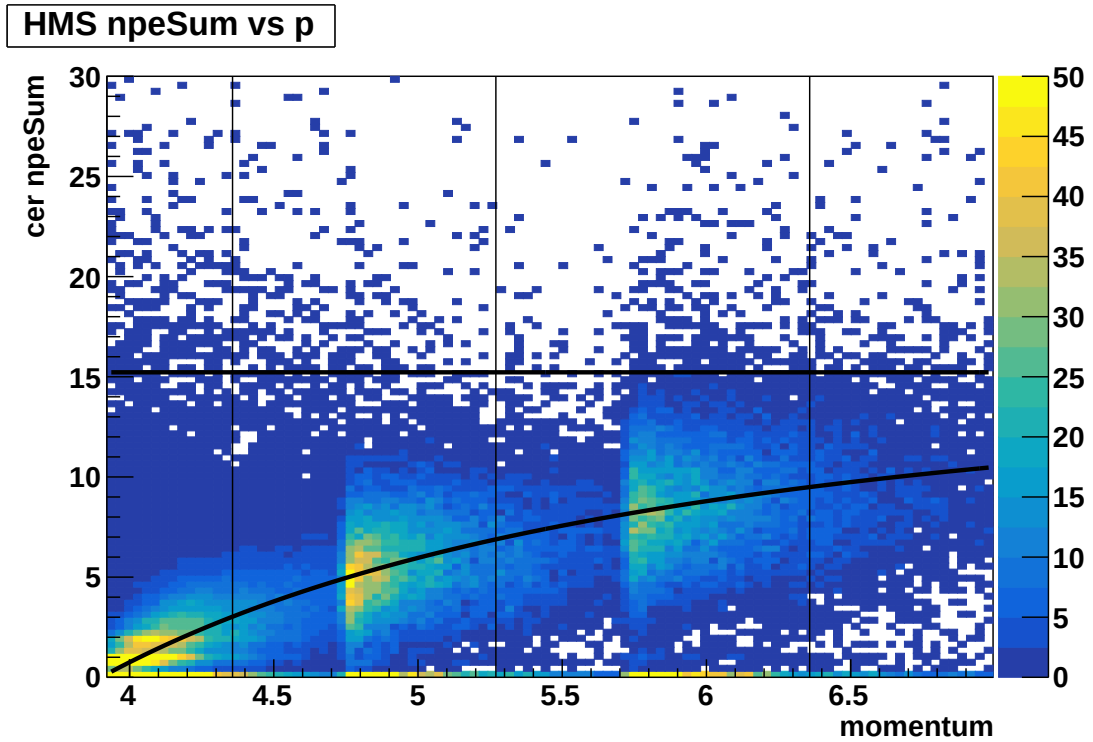


Figure 4.10: The number of photoelectrons generated in the HMS Cherenkov detector depends on different momentum

(Npe) measured for pion is momentum dependent as in Fig.4.10. The arbitrary coefficient N_0 is determined by the photoelectrons generated by electrons. A pion cut from the calorimeter is applied to remove the electrons. Three different runs with three center momentum setting at low, middle, and high are combined in Fig. 4.10. At high momentum, the pions produce more photoelectrons, similar to electron-produced Npe. And the spread of pion Npe makes it challenging to be separated cleanly from electron Npe. For this reason, HMS Cherenkov cut are only used for some efficiency calculation, but not in electron identification.

Calorimeter Calibration

The incident particle radiates photons and gets measured by the PMTs coupled at the end of the block. Ideally, all the calorimeter blocks with PMTs should have the same response for a given deposited energy. The calibration process matches

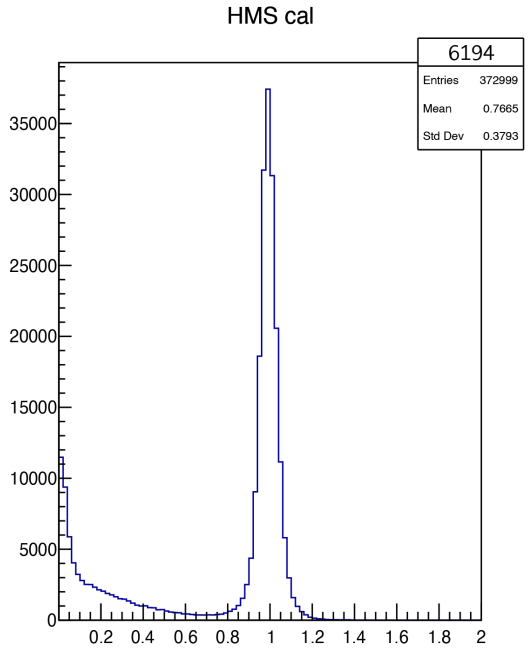
the gain constant for PMTs. For the HMS calorimeter, the electron energy should be the same as the center momentum setting. The existing HMS calorimeter calibration parameters are already good enough that the electron E/p peak is at 1. More details will be discussed in the SHMS calorimeter calibration.

HMS Lead Glass Calorimeter

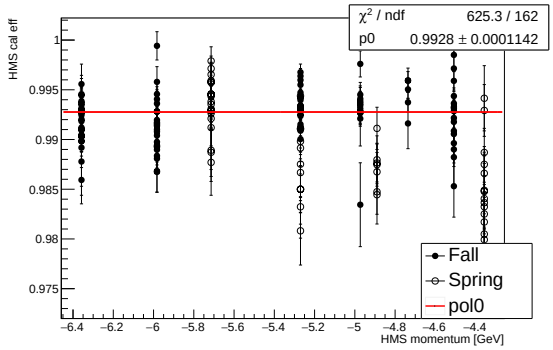
The electrons deposit all their energy on the HMS calorimeter as described in Section3.5.4. Since the electrons are massless, the E/p peak position for electrons should be 1 as shown in Fig.4.11a. A calorimeter cut E/p greater than 0.85 is used to select electrons. A clean electron sample without the calorimeter cut is required to calculate calorimeter electron detection efficiency. The calorimeter electron detection efficiency is defined as:

$$\varepsilon_{HMS,cal} = \frac{N_{did}}{N_{should}} \tag{4.7}$$

where N_{should} is the clean sample of electrons without a calorimeter cut, and N_{did} is the number of electrons in N_{should} that pass the calorimeter cut. To get a clean sample of electrons without the calorimeter, a strict SHMS pion cut and strict HMS Cherenkov Npe cut are used. Each lead glass in HMS is attached with at least one PMT. The gain constant for the PMT depends on the central momentum of the spectrometer. Because of that, the HMS calorimeter electron detecting efficiency is plotted as a function of central momentum setting as in Fig.4.7. A straight line is fitted on the efficiency. The result for D_2 runs is in Fig.4.11b, and for H_2 and Al runs are in Fig.4.12

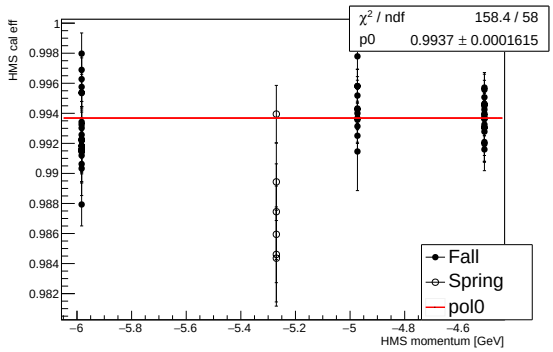


(a) HMS calorimeter E/p for one fall run without any PID cut

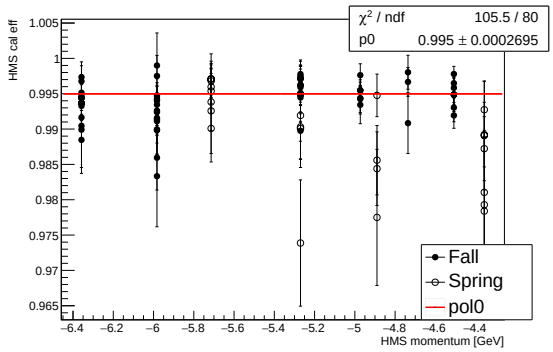


(b) HMS calorimeter electron detection efficiency

Figure 4.11: HMS calorimeter



(a) H2 runs



(b) Dummy runs

Figure 4.12: HMS calorimeter electron efficiency as a function of HMS momentum

4.5.3 Hadron Detection Efficiency

Hadrons are measured in SHMS in the CSV experiment. To get the expected pure pions, we must separate them from other particles. At our SHMS momentum and acceptance setting, we would simultaneously detect protons, kaons, and electrons as pions. Pion identification is usually done by a combination of Cherenkov detectors and calorimeters.

Lead Glass Calorimeter Calibration

As shown in Section 3.5.4, the lead glass electromagnetic calorimeter for particle energy measurement is used in SHMS. The calorimeter works similarly to the HMS calorimeter. However, the hadrons only deposit part of their energy. The SHMS calorimeter has 2×13 PMTs for preshower and 14×16 PMTs for the shower part. The PMTs should have the same response to a given deposited energy. The energy that a PMT measure is:

$$E_i = G_i A_i \tag{4.8}$$

where G_i and A_i are the gain constant and raw ADC value for the ith block, respectively. The observed energy for one event (kth) in a calorimeter would be the sum of all the measured energy from all blocks:

$$E_k = \sum_i^n E_i^k, \tag{4.9}$$

The PMTs gain constant should be modified to make the same signal amplitudes for the same energy deposit to measure the correct energy deposit. Calibration for the SHMS calorimeter is done by electrons. Only negative momentum runs are used. The electrons' deposit energy is known from the momentum measurement

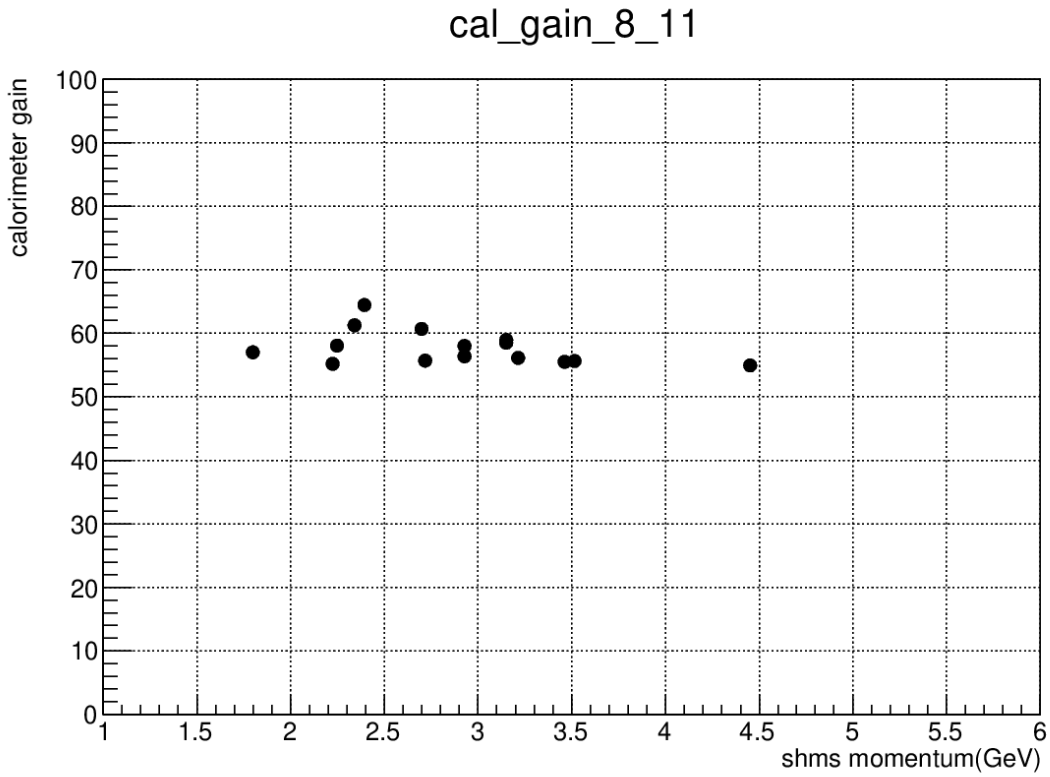


Figure 4.13: Gain coefficients for the shower block at row 8 and column 11

p. The E/p peak for electrons should be at 1 as in HMS. However, the existing parameters shift the peak a little as in 4.14. The gain constant for each block (n blocks) is then determined by minimizing the χ^2, which is defined as the squared difference between known energy (from track momentum p for the electron) to the measured energy in the calorimeter for all the selected electrons(N):

$$\chi^2 = \sum_k^N (p^{(k)} - \sum_i^n G_i A_i^{(k)})^2 \tag{4.10}$$

The gain-matching coefficients are generated from the minimization for each negative run. One of the gain constants from one block is plotted as a function of SHMS central momentum setting as in Fig.4.13. The average is taken for the final gain constant for the corresponding block. The calibrated gain-matching coefficients are then used in data replay. The electron energy peak position is calibrated

to the correct position as in Fig.4.15b.

Calorimeter Efficiency

The efficiency is defined similarly to HMS. The calorimeter E/p cut is chosen between 0.05 and 0.85. A clean pion sample without a calorimeter cut is required. For the negative polarity runs, electrons are also accepted in the SHMS spectrometer. The electron E/p will be at 1. To avoid electron contamination, positive runs are used for efficiency study.

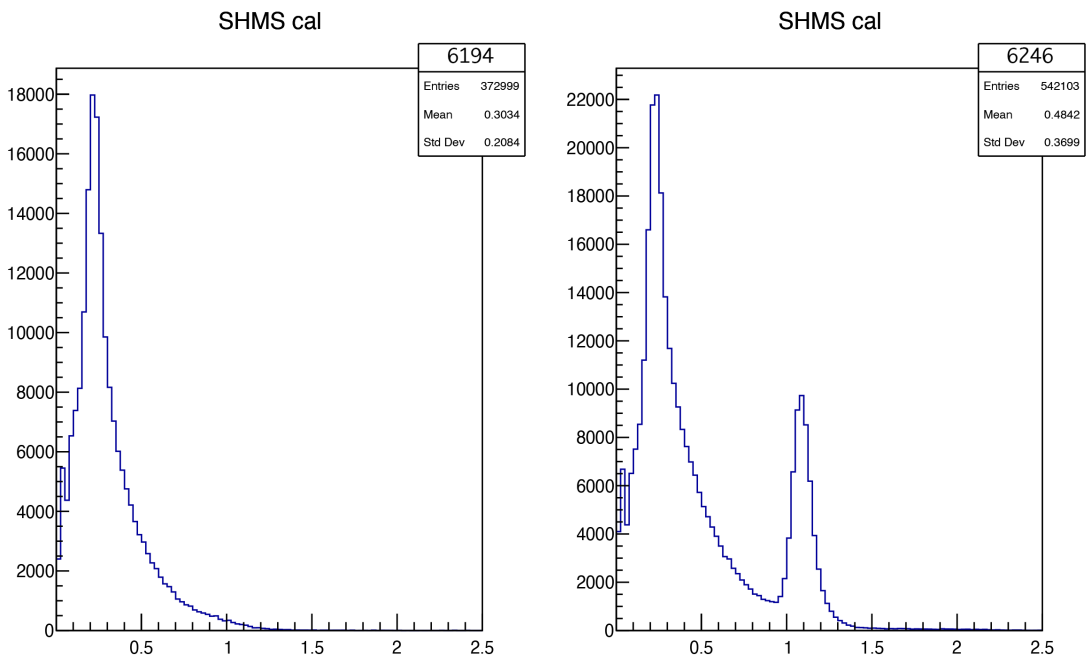
For the calorimeter pion detecting efficiency:

$$\varepsilon_{SHMS,cal} = \frac{N^{\pi}_{did}}{N^{\pi}_{should}} \tag{4.11}$$

The pure pion sample N_{should} is selected by strict HMS electron cut, high SHMS aerogel cut, heavy gas cut, and RFtime cut. The number of pions in N_{should} that passed the calorimeter cut would be N_{did}.

4.5.4 SHMS Aerogel Cherenkov Detector

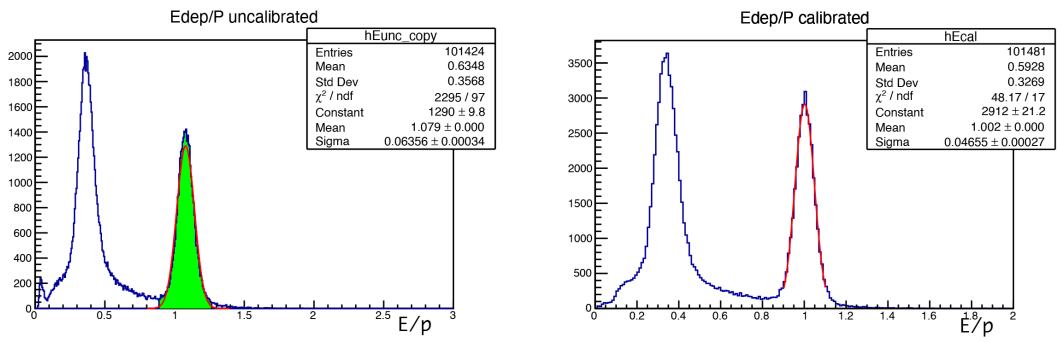
SHMS aerogel detector, as described in Section 3.5.3 is a good proton identification detector. The threshold for different particles and the relation between the number of photoelectrons as a function of momentum is in Fig. 3.19. For this experiment, the SHMS momentum setting range is below the proton threshold. However, there might be some knock-on electrons from the proton, also called delta rays. Electrons are emitted from atoms by the passage of protons through the aerogel. Proton traversing the aerogel transfers energy to the aerogel atoms via ionization or excitation of the constituent atoms. The electrons will fire the aerogel detector and produce some photoelectrons.



(a) π^+ runs

(b) π^- runs

Figure 4.14: SHMS E/p for one fall RunGroup without any PID cut before calibration



(a) Uncalibrated

(b) After calibration

Figure 4.15: SHMS calorimeter calibration E/p. The green areas are selected for the minimization process, and after the calibration, the electron peak is calibrated to the correct position at 1.

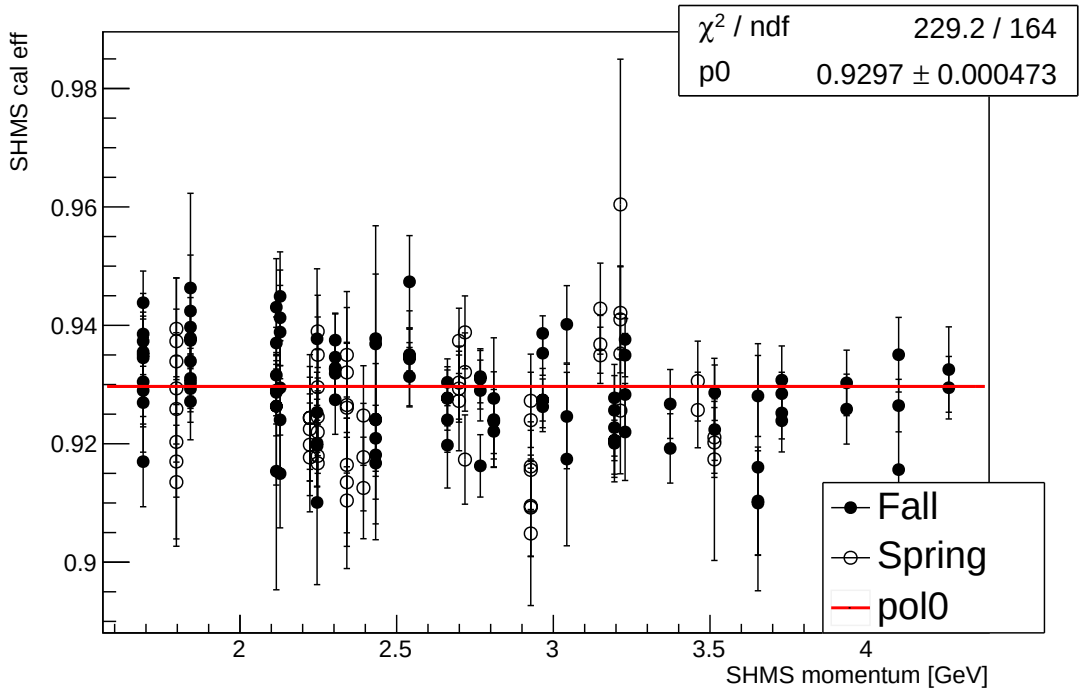
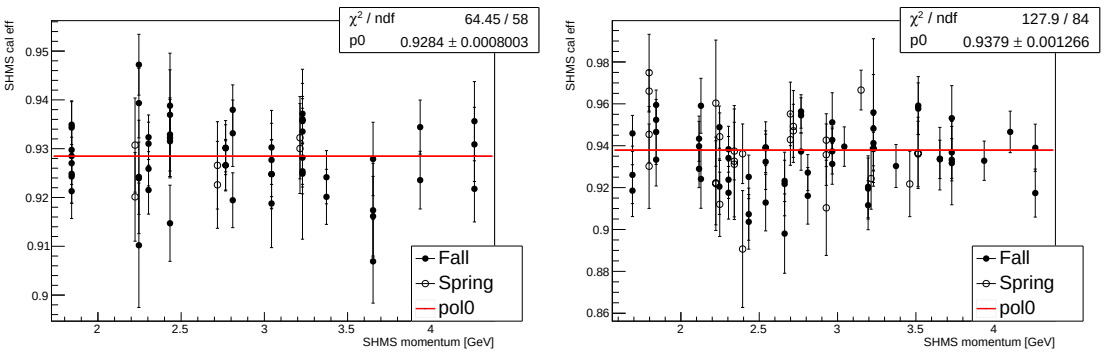


Figure 4.16: SHMS calorimeter pion detection efficiency



(a) H2 runs

(b) Dummy runs

Figure 4.17: SHMS calorimeter pion efficiency as a function of SHMS momentum

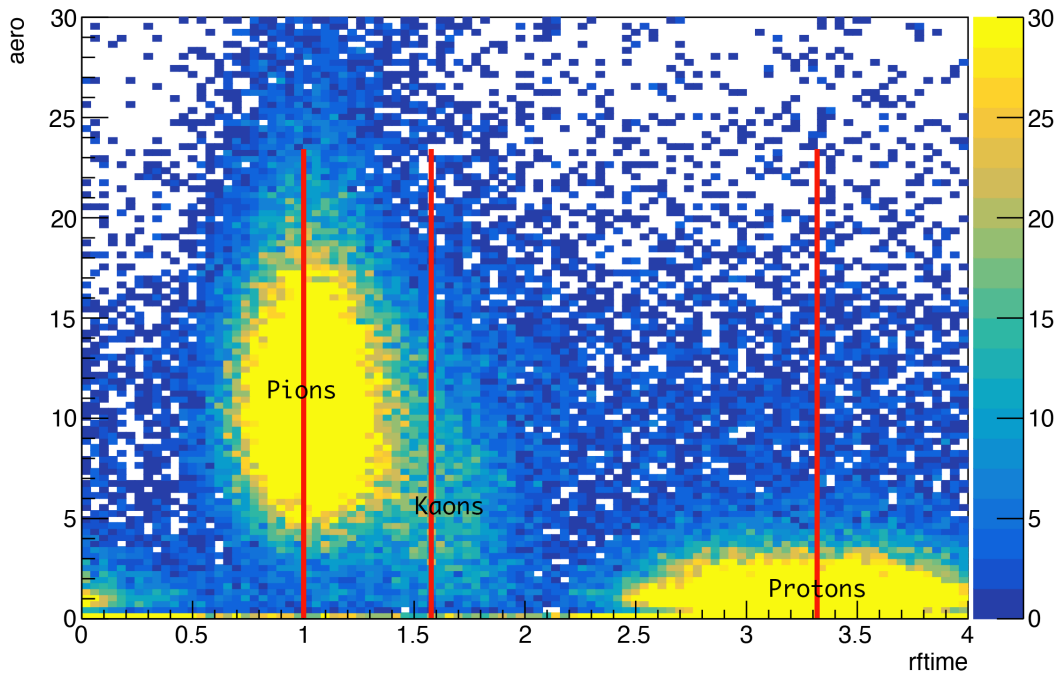


Figure 4.18: Aerogel Number of photoelectron vs. rftime pid

The Aerogel Npe vs rftime is shown in Fig. 4.18. The definition of rftime is discussed in section4.5.6. For that rungroup, the SHMS momentum is 3.5 GeV, from Eq.4.4, the proton coincidence peak is ∼2.5 ns slower than the mean peak. From the fall poor resolution coincidence spectrum, the proton peak cannot be separated from the mean peak. The protons are below the aerogel detector threshold. However, the knock-on electrons produced some aerogel Npe as shown on Fig.4.18 right-hand side. At this momentum, kaons are above the threshold. Kaons are estimated through RFtime cut.

For another RunGroup at SHMS central momentum 2.69 GeV as shown in Fig.4.19. The protons are ∼3.6 ns slower than pion, which can be eliminated by a coincidence time cut. In RFtime spectrum, the protons are ∼3.6 ns slower than the pion peak (at 1 ns), which will be larger than RFtime spectrum range of 4 ns. The proton RFtime would be folded back to ∼0.6 ns due to the Mod operation.

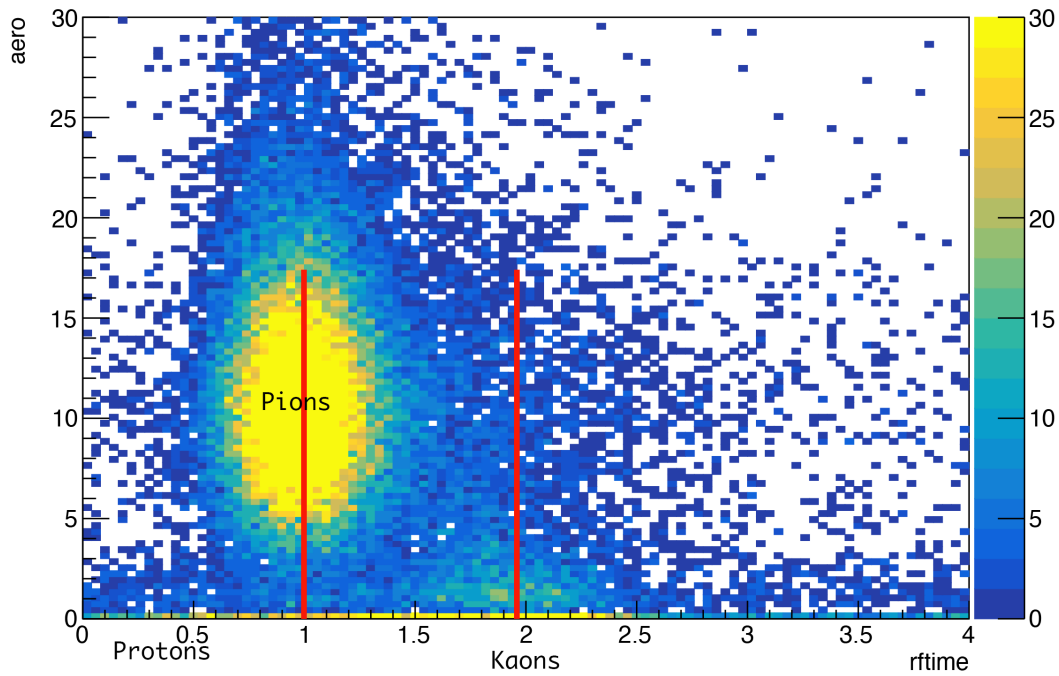


Figure 4.19: Aerogel Number of photoelectrons vs. rftime PID for low SHMS central momentum

There might still be some accidental protons from accidental backgrounds. Kaons are ∼1 ns slower than pions, and kaons are below the aerogel threshold. However, due to kaon decay ($K^+ \rightarrow \pi^+ + \pi^+ + \pi^-$ and $K^- \rightarrow \pi^- + \pi^+ + \pi^-$), kaons might also fire aerogel Cherenkov detector. This can be estimated by RF time cut.

To remove the effect of knock-on electrons, the aerogel cut is chosen to be greater than 4 Npe. The aerogel efficiency is calculated by:

$$\varepsilon_{aero} \equiv \frac{N^{\pi}_{did}}{N^{\pi}_{should}} \tag{4.12}$$

where the N^{π}_{should} is selecting the best pion sample without aerogel cut. Several strict electron cuts are used in the HMS spectrometer to get the best pion sample. Heavy gas Cherenkov detector Npe is set to be greater than 2. The inefficiency at the middle loses some pions, but the remaining particles should be pions. Rftime cut less than 1 ns is used to avoid the effect of kaons contamination. The numerator is the pion sample that did pass the aerogel Npe cut 4, which includes all cuts in

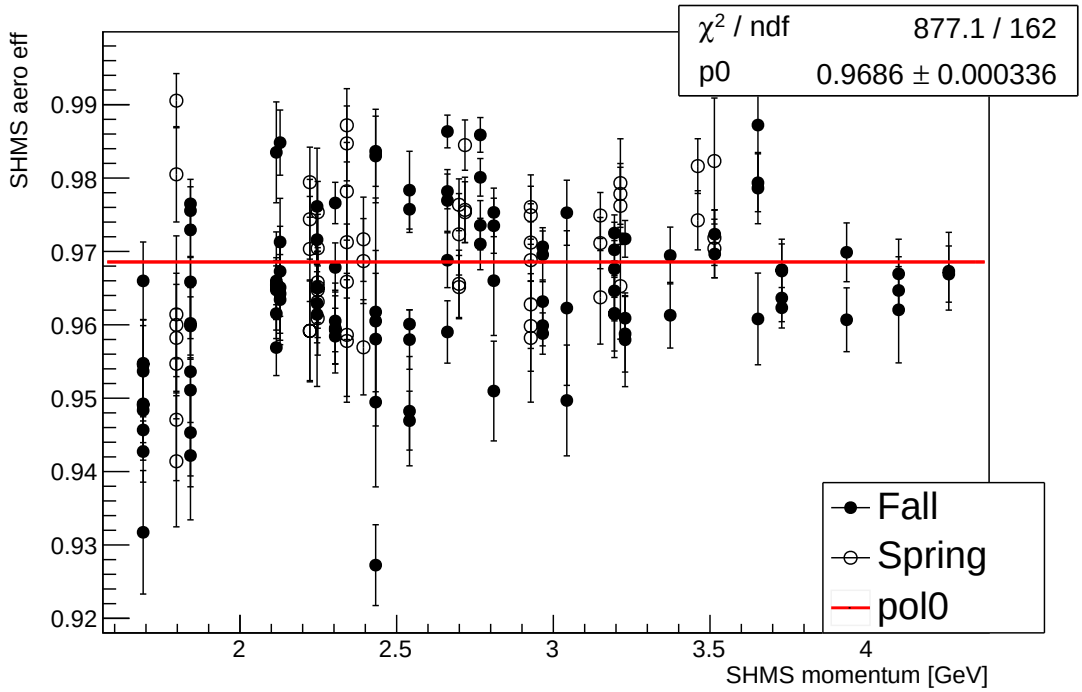
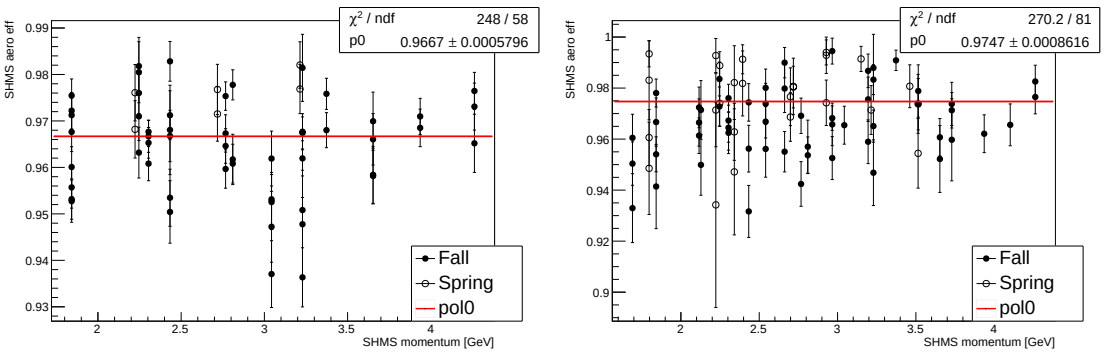


Figure 4.20: SHMS aerogel Cherenkov detector efficiency as a function of SHMS momentum for D_2 runs



(a) H2 runs

(b) Dummy runs

Figure 4.21: SHMS aerogel Cherenkov detector efficiency as a function of SHMS momentum

the denominator and the aerogel Npe cut greater than 4. The aerogel detecting efficiencies for D_2 runs are in Fig.4.20, for H_2 and Dummy runs are in Fig.4.21. The pion Npe in aerogel depends on the momentum as shown in Fig.3.19. However, the momentum settings are high enough for pions to pass the aerogel Cherenkov threshold and to radiate enough photoelectrons. The aerogel efficiency does not depend on momentum setting. A pol0 fitting is applied on the aerogel efficiency plot and used as the efficiency for all runs to avoid statistic fluctuation.

For low momentum less than kaon threshold, the aerogel cut could remove the kaons. For high momentum settings, the kaons will also fire the aerogel detector, but it will not fire the heavy gas Cherenkov detector.

4.5.5 SHMS Heavy Gas Cherenkov Detector

The SHMS heavy gas Cherenkov detector is designed for K/π separation. As illustrated in Section 3.5.3, the heavy gas Cherenkov detector has four reflecting mirrors corresponding to four PMTs for the Cherenkov light collection. However, a crack at the center is found in the mirror. This crack causes pion leaking at the center spectrometer setting.

For high momentum (greater than 2.8 GeV), pions start to fire light in heavy gas Cherenkov detector as calculated in Fig. 3.18b. The pion efficiency is defined as:

$$\varepsilon_{HGC} \equiv \frac{N^{\pi}_{did}}{N^{\pi}_{should}} \tag{4.13}$$

where the N_{should} is a clean pion sample without any heavy gas Cherenkov cut. And N_{did} is the number of pions in N_{should} that passed the HGC cut. The inefficiency at the geometry center is seen in Fig.4.22. The geometry center is at the center

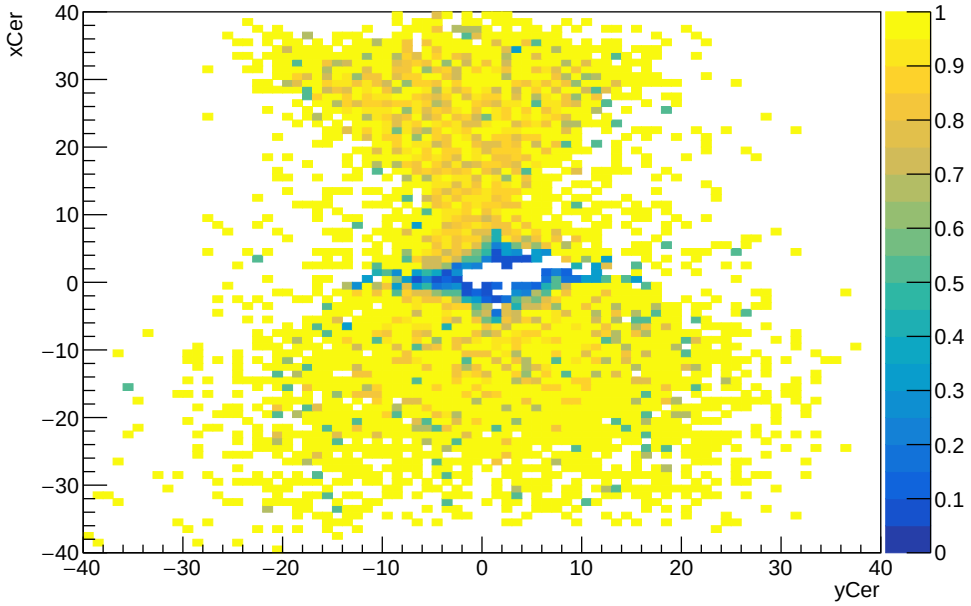


Figure 4.22: SHMS HGC inefficiency at the geometry center due to the mirror crack

momentum setting, which should have the densest desired pions. The mirror crack would lose the precious good pion statistics.

The number of photoelectrons can be calculated by equation Eq.3.4 and Eq.3.5. The N_0 is related to the actual conditions and is unknown. The tank is filled with gas, and the threshold is controlled by changing the pressure of the gas as in Fig.3.17. The tank is filled with 1 atm gas with a refraction index of 1.00137. The refraction index is sensitive to gas pressure. A fitting on the heavy gas Cherenkov number of photoelectrons is tested with two parameters.

$$N_{p.e.} = p_0(1 - \frac{p^2 + m^2}{(p_1 n)^2 p^2}) \tag{4.14}$$

The unknown p_0 is the scaling parameter, which will decide the total number of photoelectrons. In the case of the gas pressure change in the tank, a small offset p_1 on the refraction index is applied. The inefficiency near the center momentum region is removed.

The pions won't fire heavy gas Cherenkov detector at a low momentum setting

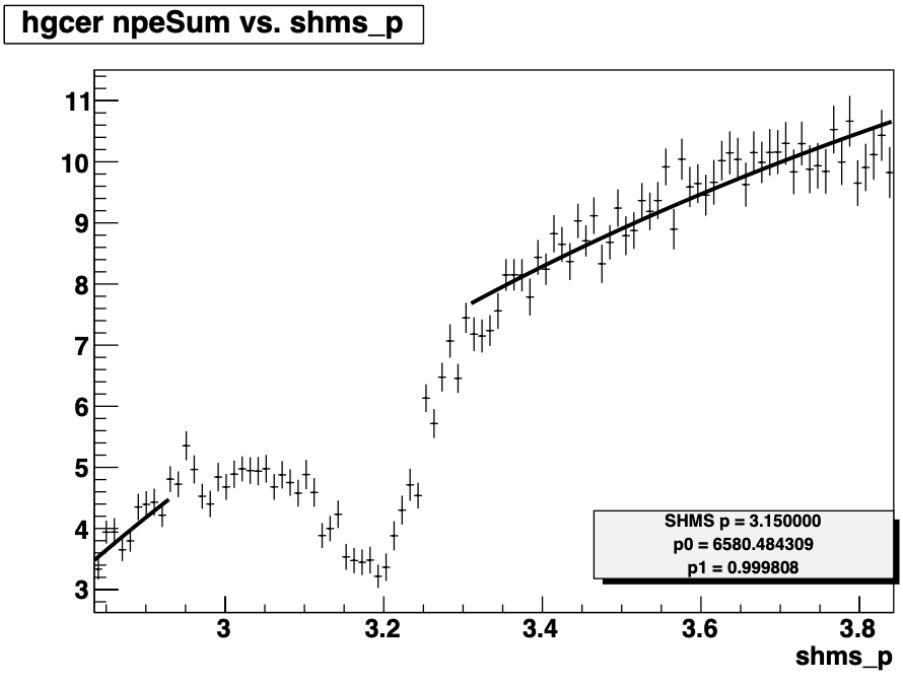


Figure 4.23: SHMS heavy gas Cherenkov detector inefficiency due to mirror misalignment.

(less than 2.8 GeV). It will fire the heavy gas Cherenkov detector at high momentum setting runs, while the kaons won't. the mirror crack at the middle will lose pions at the center. Besides that, the RFtime can be used to estimate pion kaon separation. So the HGC is combined with RFtime PID for pion selection. The RFtime PID is discussed later.

4.5.6 RFtime Offset

During the experiment, the electron beam accelerator frequency is 249 MHz for Hall C, and the beam period is 4.008 ns (4 ns for short) as in Sec.3.2. Hall C and the newly constructed Hall D share the same electron beam during the experiment. The beam chopper will separate the beam and deliver half of the beam to Hall C and the other half to Hall D. The resonance time is measured at the resonance frequency cavity at Fit.3.2. The difference between Resonance Frequency Time and Hodoscope time (RFtime) is the recorded particle time. This

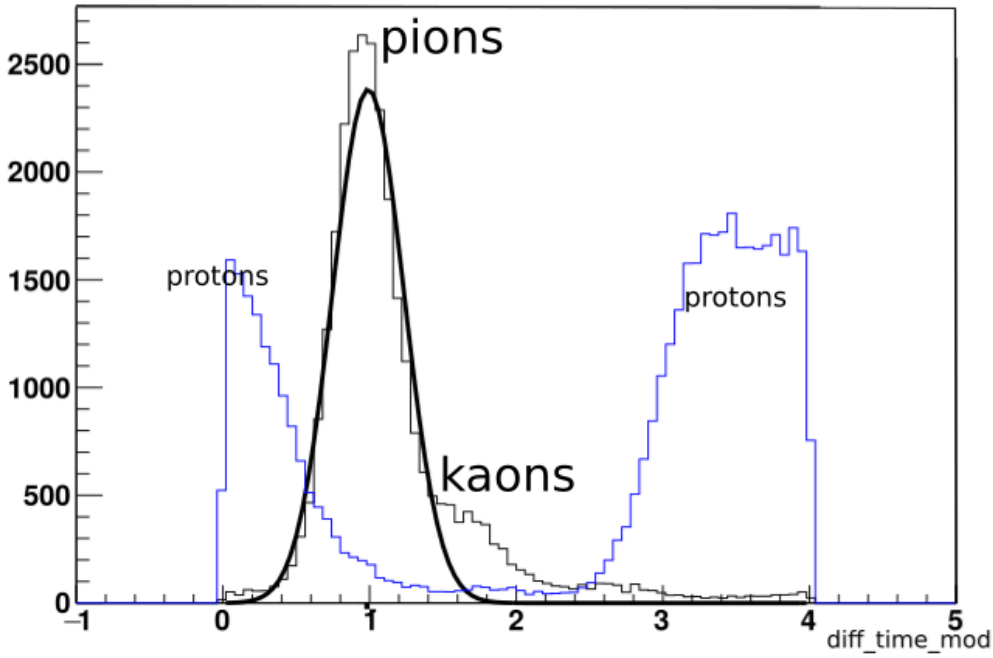
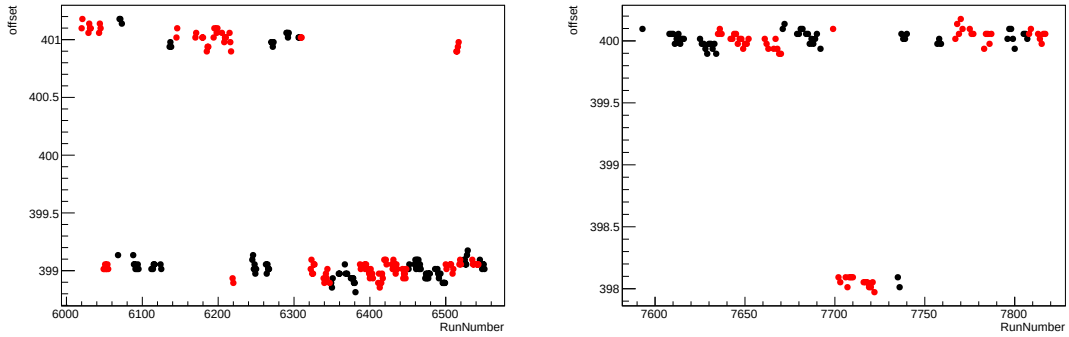


Figure 4.24: RFtime offset to put pion peak at 1 ns for one setting. The kaons and protons are slower

Resonance Frequency Time starts from arbitrary and ranges from negative to positive. A random offset ($\sim$ 400 ns) is added to shift all values to positive for all runs. Different particles travel the same path in each 4 ns period, so Mod(4ns) operation is done on the RFtime. The offset ensures that the Mod operation will not get confused at zero. Due to the beam chopper, it is not sure which half of the beam goes to Hall C. Sometimes, the pion RFtime peak from some runs is a half period (2 ns) different than other runs. To get a clear peak of pion at the RFtime spectrum at 1 ns, an additional offset of 2 ns is added to RFtime for some runs.

The offsets from all the runs are summarized together in fig.4.25

By fitting for all points from fall and spring, the RFtime offset is chosen to be 401.016 (399.012) for the fall runs and 400.027 (398.023) for the spring runs. Besides the offset from the beam, the hadron momentum setting will change the pion travel time. The time for different particles to travel from the target chamber



(a) Fall runs

(b) Spring runs

Figure 4.25: RFtime offset to move the maximum bin to 1

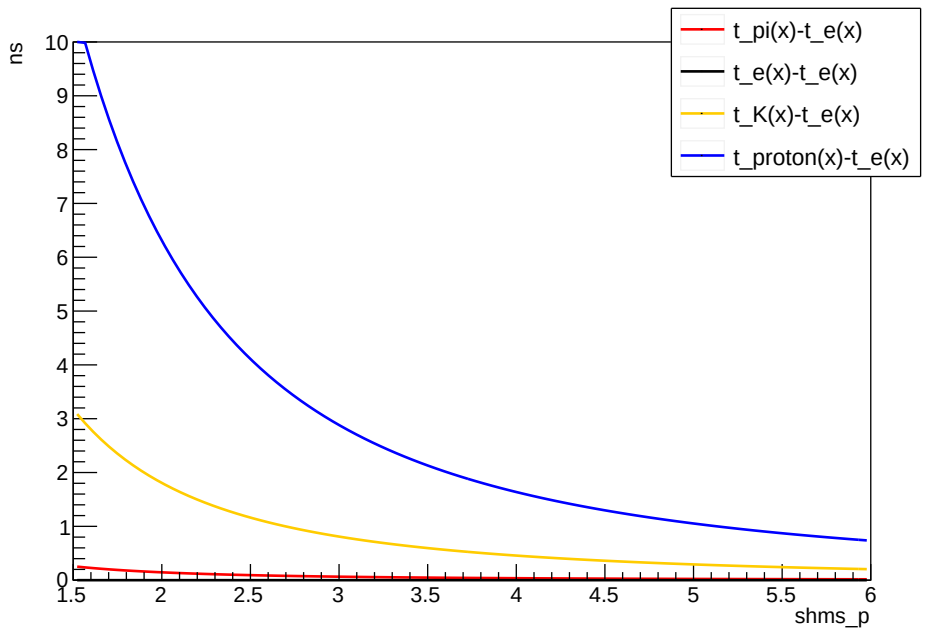


Figure 4.26: Difference of time of flight for different particles in SHMS compared to electrons versus different momentum on the x-axis.

to the hodoscope plane regarding their momentum and mass is calculated as in Figure 4.26 . The total length of the SHMS spectrometer is 20.1 m. The spectrometer central momentum setting p in GeV. The travel time for each particle can be calculated by its mass (GeV) using

$$t = \frac{\text{length} \cdot \sqrt{p^2 + m^2} \cdot 10^9}{c \cdot p} \tag{4.15}$$

where the 10^9 converts seconds to nanoseconds, c is the speed velocity in m/s. The electron mass is negligible, and the different center momentum won't change the electron flight time much. Depending on their different mass, the other particles would be a few nanoseconds slower than the electron. The RFtime offset is then modified by the spectrometer's central momentum setting for pions.

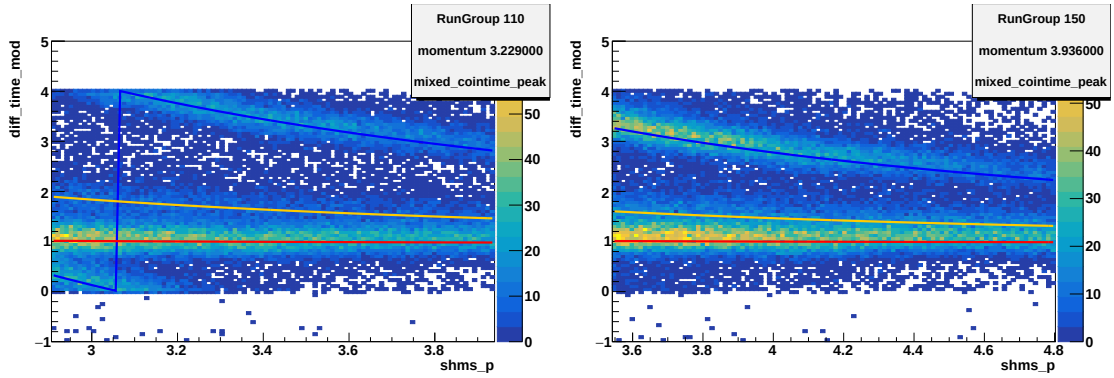
$$\text{rftime_offset}_\pi = \text{rftime_offset}_{raw} + t_\pi(\text{SHMS_P_central}) - t_\pi(\text{SHMS_P_average}) \tag{4.16}$$

After adding the offsets on all the settings, the pions are centered at 1 ns for all the runs at rftime spectrum.

4.5.7 RFtime PID

The rftime spectrum is used for particle identification purposes. The mass difference between different particles will locate at different masses on the rftime spectrum.

For one setting in Fig.4.27a, the SHMS center momentum setting is 3.23 GeV. The acceptance allows the momentum to vary from 2.9 to 3.9 GeV. The pions RFtime peak is already moved to 1 ns as the band under the red line shown in Fig. 4.27a. RFtime cut is chosen to be around 1ns to select the pion peak. The



(a) low momentum

(b) high momentum

Figure 4.27: RFtime vs. Momentum

kaon and proton travel time can be calculated as a function of momentum as in Fig. 4.26. At low momentum, the protons travel more than 3 ns slower than pions, which can be removed by the coincidence time cut. The proton RF time would be folded back to the beginning due to the Mod(4) operation. The protons can be removed by the combination of coincidence time cut and RFtime cut. The kaons have a similar RFtime as pions, which is difficult to be separated clearly from the pion band. Kaons need to be estimated in another way. For another center momentum, Fig.4.27b is showing a mixed kaon band and a closer proton band.

The mixed kaons are estimated by the RFtime spectrum. For one setting the momentum range between -10% to 20%, the rftime spectrum needs to be corrected by each particle. For the particles which have momentum different than the central momentum, a time correction is applied to the raw rftime which is calculated by the equation 4.15:

$$\begin{aligned} rf_{pi} &= rf_{raw} + t_{pi}(Preal) - t_{pi}(Pcentral), \\ rf_{K} &= rf_{raw} + t_{K}(Preal) - t_{K}(Pcentral), \\ rf_{proton} &= rf_{raw} + t_{P}(Preal) - t_{P}(Pcentral) \end{aligned} \tag{4.17}$$

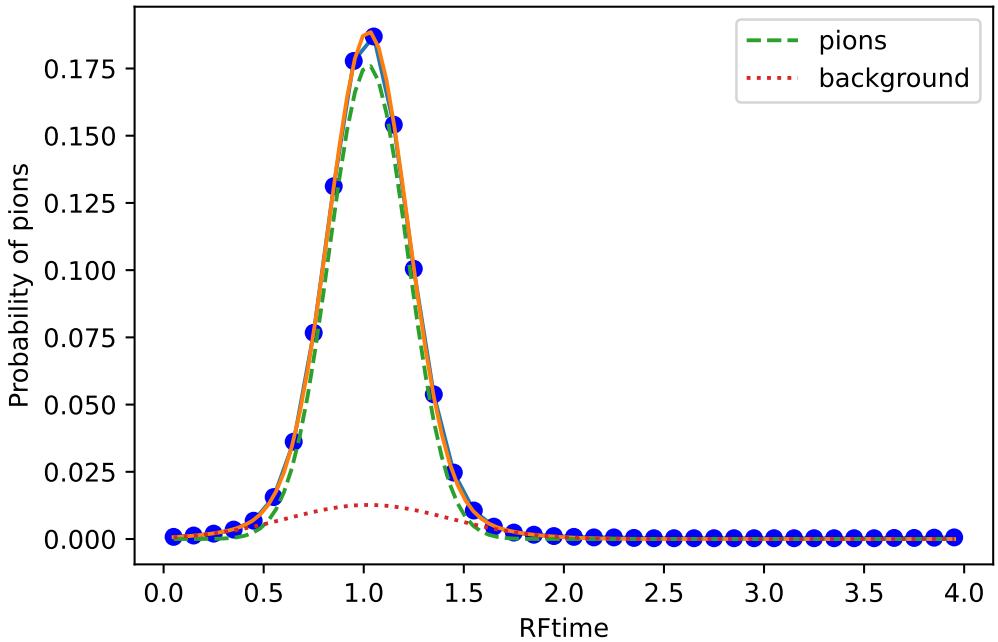


Figure 4.28: The function form used for the peaks

The kaon corrected rftime spectrum is used to estimate kaon contamination. There are more K^+ generated than K^-. Since the polarity won't change the rftime spectrum, the two polarities are considered together. Each RunGroup is separated into small samples by two kinematic variables, x_{bj} and z. One RunGroup is separated into two x_{bj} bins for less than the xbj center setting and greater than the center setting, and 20 bins in z from 0 to 1. The low and high edges of the z bins are empty.

After a HGC pion cut ($HGC_{npe} > 1$), pions are selected and form a peak as on Fig.4.29. The top one is for the π^+ and the bottom one is for the π^- for the same setting. The shape of the pion peak is roughly a gaussian function with some backgrounds due to the measurement inefficiency. This shape is estimated by selecting a good sample of electrons from elastic runs. The background peak is related to the main pion gaussian peak with lower amplitude and larger width.

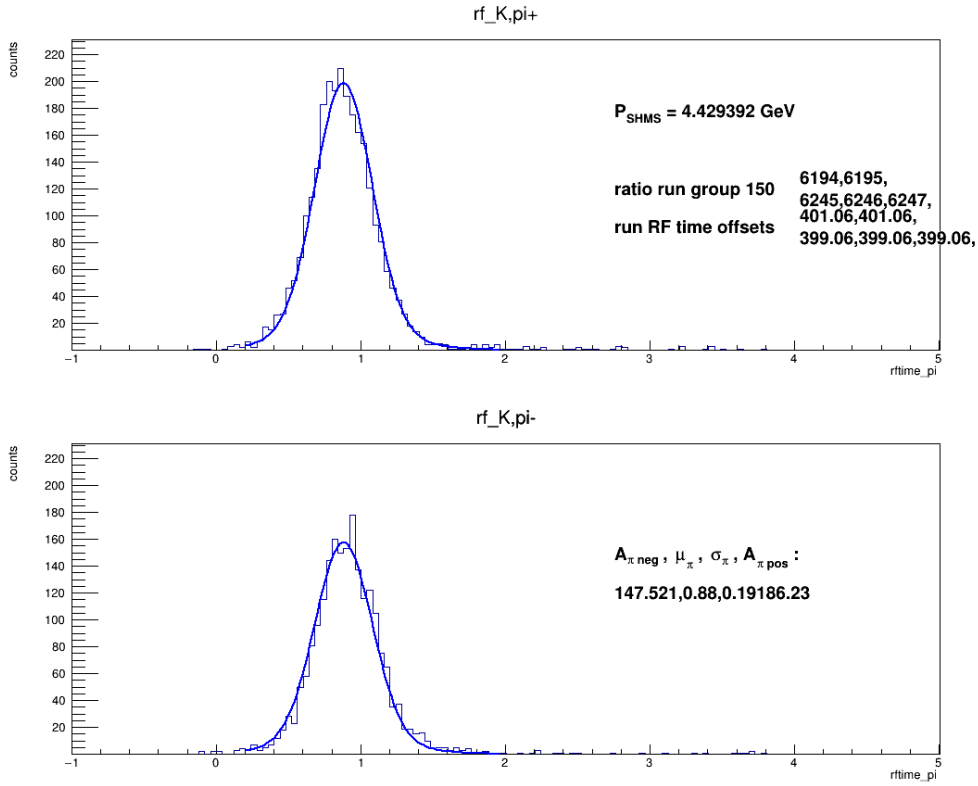


Figure 4.29: Pion peak selected by a HGC pion cut on rftime kaon mass corrected spectrum

Two functions are fitted on both positive and negative peaks simultaneously. The pion peak position μ and width σ are extracted. The integration of the two polarity peak could get the number of pions, $N^{\pi^+}_{\text{withHGC}}$ and $N^{\pi^-}_{\text{withHGC}}$.

To get the number of kaons, an anti-HGC cut ($HGC_{npe} < 1$) is applied to the sample. Because of the center's bad region, there will be some pions leaking into the anti-HGC cut. The size of the pion peak will depend on the kinematics which is related to the position at the heavy gas Cherenkov detector. This spectrum is then fitted with fixed pion peak μ and σ from the HGC cut pion peaks. Depending on the momentum, the kaon time is calculated and added to the pion time. The free parameter for the anti-HGC cut spectrum would be the amplitude of the two pion peaks, two kaon peaks, and the kaon σ. From this result, the four numbers are extracted, $N^{\pi^+}_{\text{antiHGC}}$, $N^{\pi^-}_{\text{antiHGC}}$, $N^{K^+}_{\text{antiHGC}}$ and $N^{K^-}_{\text{antiHGC}}$.

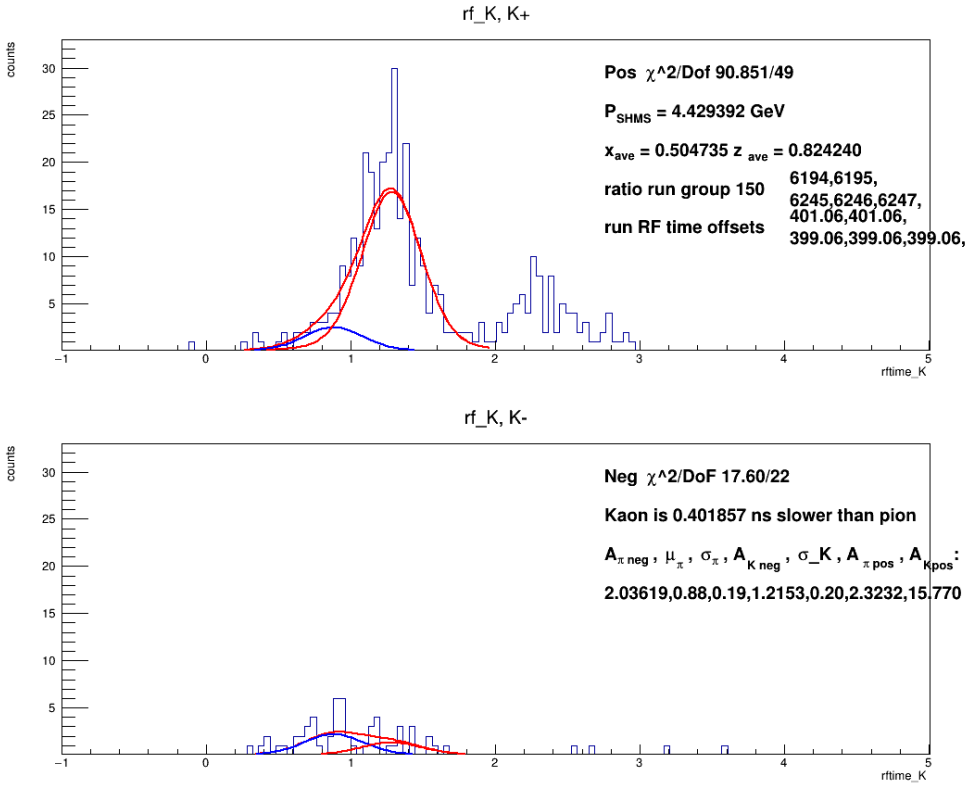


Figure 4.30: Pion and kaon peaks selected by an anti-HGC cut on rftime kaon mass corrected spectrum

Each RunGroup is separated into smaller sample depending on x_{bj} and z, the kaon to pion ratios are checked as a function of the variable z. More results are shown in Fig.4.31

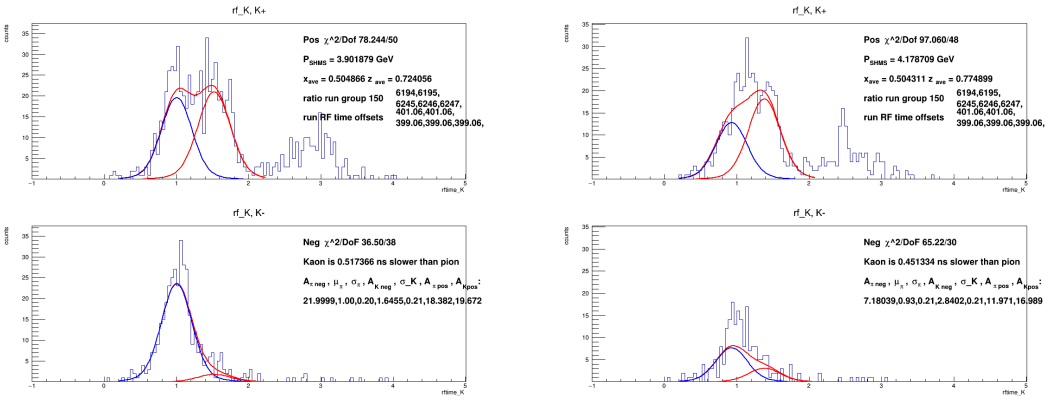
From the two fittings and the six number, the heavy gas efficiency and kaon to pion ratio can be formed:

$$\mathrm{HGC}_{eff} = \frac{N^{\pi}_{\mathrm{withHGC}}}{N^{\pi}_{\mathrm{withHGC}} + N^{\pi}_{\mathrm{antiHGC}}} \tag{4.18}$$

$$R_{K/pi} = \frac{N^{K}_{\mathrm{antiHGC}}}{N^{\pi}_{\mathrm{withHGC}} + N^{\pi}_{\mathrm{antiHGC}}} \tag{4.19}$$

For the heavy gas efficiency, one point from one group is taken, and the efficiency as a function of momentum is plotted in Fig.4.32

The kaon to pion ratio is plotted as a function of z for each (x, Q^2) setting. One of them is shown in Fig.4.33. A linear fit is applied on the ratio, the kaon to pion ratio for that setting is then calculated from the fit.



(a) $x > 0.45$ and $0.7 < z < 0.75$ (b) $x > 0.45$ and $0.75 < z < 0.8$

Figure 4.31: Kaon to pion ratio plots for one RunGroup with different kinematic cut

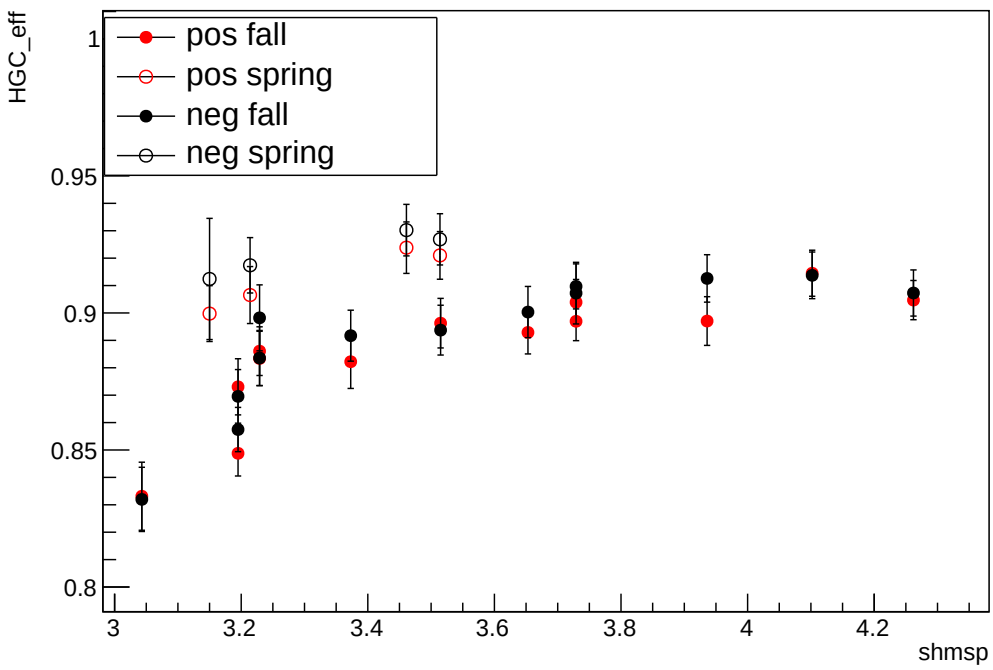
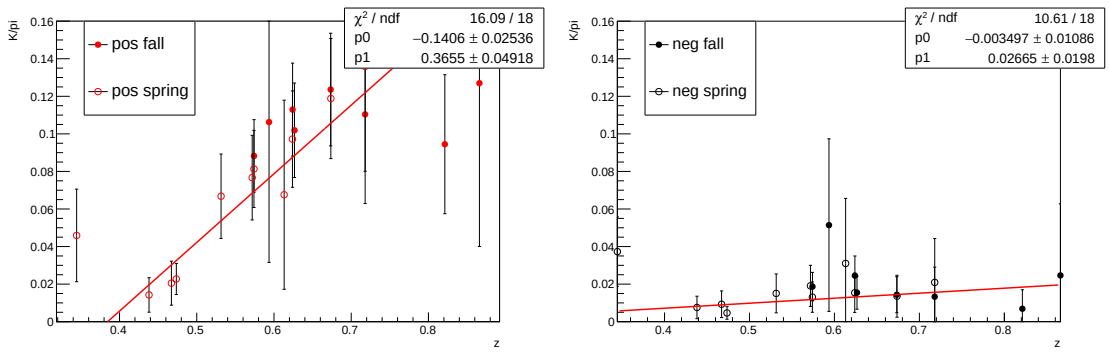


Figure 4.32: Heavy gas efficiency from the fitting. The efficiency for the spring runs is higher than for the fall runs. The heavy gas detector mirror got realigned between the two run period.



(a) Positive (b) Negative

Figure 4.33: Kaon to pion ratio for one setting as a function of z for both polarity

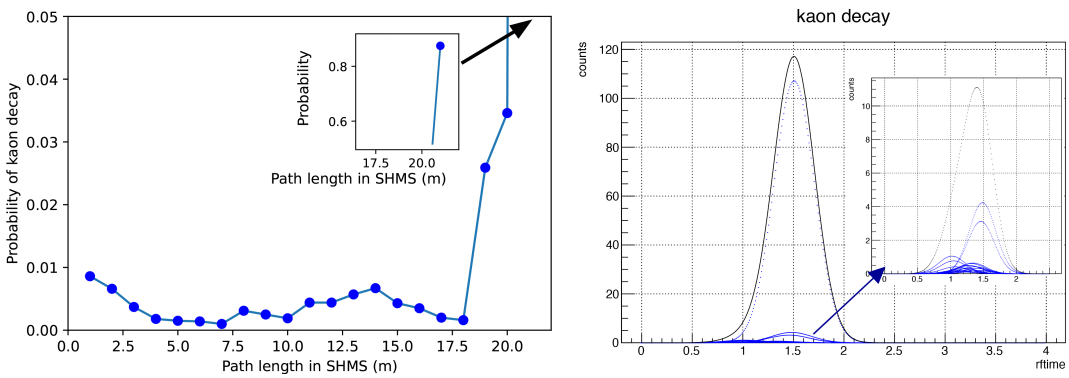
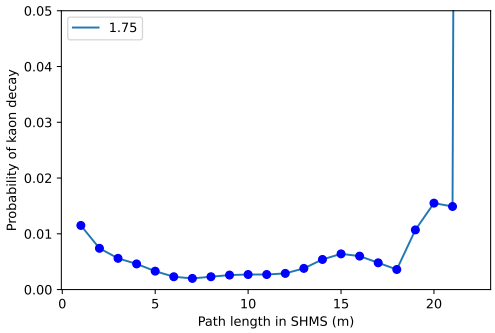


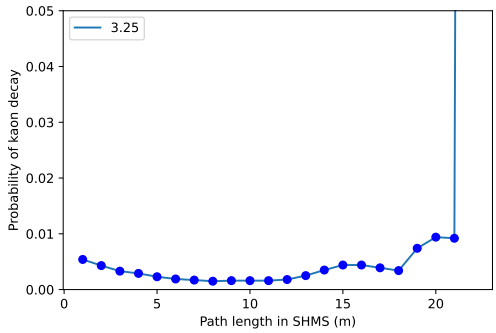
Figure 4.34: Kaon decay probability as a function of SHMS path position

The kaon decays into muons and pions anytime during the SHMS path. Among all the kaons that are generated, some of their decays are at the beginning in the target chamber. Some of their decay is in the middle of the SHMS path. Some of them decay at the end of the track in the detector hut. And most of them (usually over 80%) won't decay. Each of the decayed kaons, a small gaussian function is assumed. The peak position for the small gaussian function is calculated by the momentum of the setting and the position of the decay. The σ is assumed to be the same as pion peak σ. The amplitude is calculated by the probability divided by the not decayed kaons. The plots are shown in Fig.4.34

The probability is simulated by SIMC for ten momenta. Some probability are shown in Fig.4.35 and Fig.4.36.

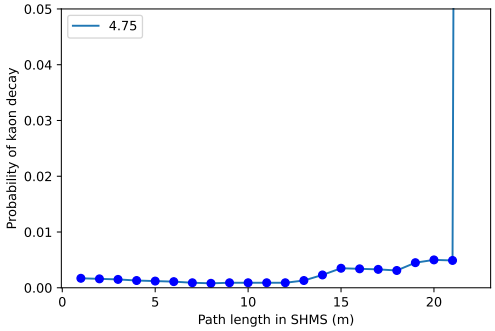


(a) $P = 1.75GeV$

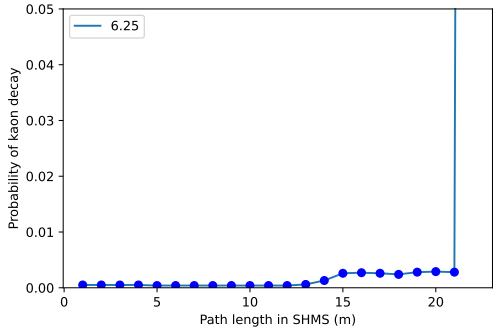


(b) $P = 3.25GeV$

Figure 4.35: Kaon decay probability for different momentum

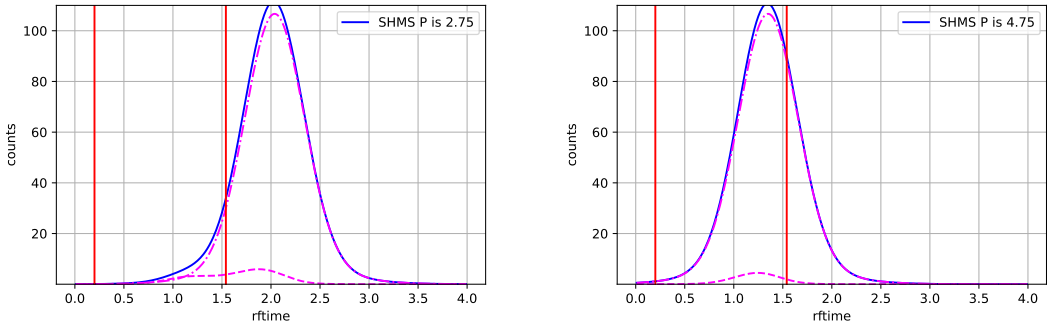


(a) $P = 4.75GeV$



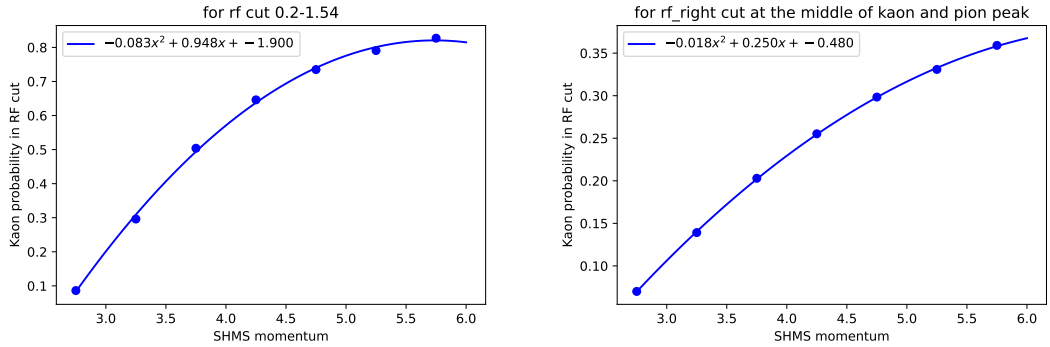
(b) $P = 6.25GeV$

Figure 4.36: Kaon decay probability for different momentum



(a) $P = 2.75GeV$ (b) $P = 4.75GeV$

Figure 4.37: The higher the momentum, the more kaons remain in rftime cut



(a) Rftime cut fixed (b) Rftime cut vary

Figure 4.38: P_K for two rftime cut cases, one is fixed from 0.2 to 1.54, the other one is varying from 0.2 to the middle of pion peak and kaon peak

After the anti-HGC cut, the not decayed kaons are selected. So the kaons estimated from the kaon to pion ratio $R_{K/pi}$ is only the not decayed kaons part. The decayed kaons are then added to the sample calculated by the probability.

At higher momentum, the kaons travel faster and the time difference between kaons and pions will be smaller. The kaon peak will be closer to the pion peak on the rftime plot as in Fig.4.37

Different rftime cuts would change the number of kaons. The portion of kaons that get inside the rftime cut P_K as a function of momentum for two cases is plotted in Fig.4.38

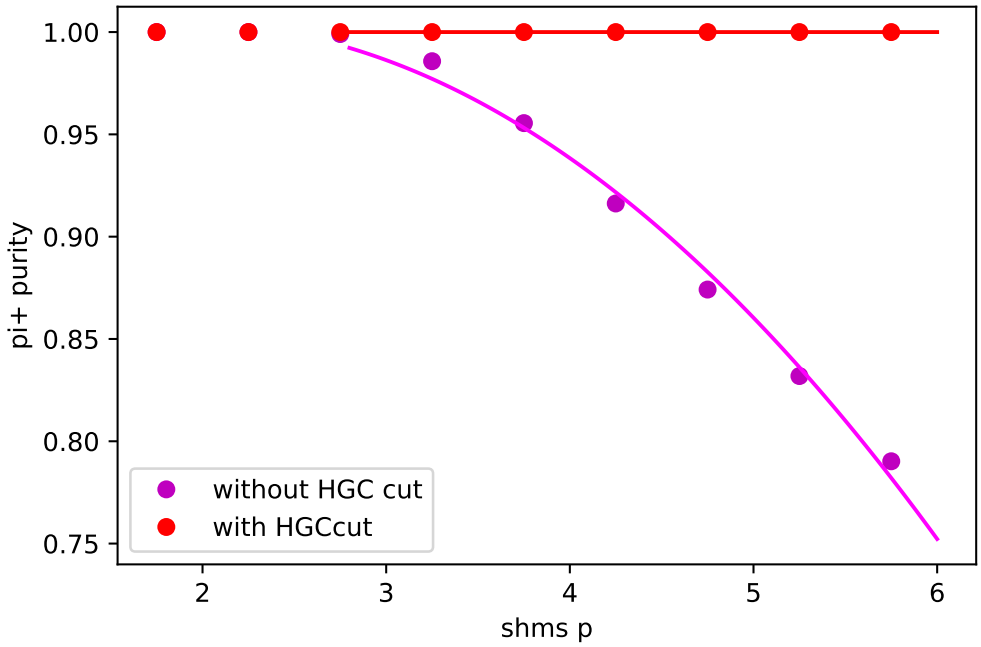


Figure 4.39: The positive pion purity as a function of momentum for one (x, Q^2) setting

The pion purity π_{purity} is then defined as:

$$\pi_{\text{purity}} = \left(1 - R_{K/\pi} P_K\right)\Big|\text{RFtimecut} \tag{4.20}$$

The pion purity is then checked for one setting. With the heavy gas cut, the not decayed kaons will be removed, the only remaining contamination would be the decayed kaons. The number of decayed kaons is small, which gives more than 99% pion purity. Without the heavy gas cut, more counts would be selected with kaon contamination. The pion sample is then corrected by the pion purity for each setting. The pion purity result for one setting is shown in Fig.4.39. The difference between the two cases is then compared for the systematic check of particle identification.

$$N^{\pi}_{true} = N^{\pi}_{raw} \cdot \pi_{\text{purity}} \tag{4.21}$$

4.6 Tracking Efficiency

Charged particles traverse in the spectrometer and make signals on detectors that will form a trigger. The drift chambers form tracking information for a particle. However, if there are too many hits in the drift chamber, it will be difficult to determine which hits to use and which is the proper track corresponding to the trigger. And if there are not enough hits in the drift chamber, then a track cannot not be determined. The tracking efficiency estimates the inefficiency of the drift chamber. The tracking efficiency is defined as:

$$\varepsilon_{tracking} = \frac{N_{did}}{N_{should}} \tag{4.22}$$

The N_{should} is a clean particle sample selected by the hodoscope, and the N_{did} is the subset of N_{should} that has been tracked in drift chambers. The hodoscope receives signals from the passing particles and forms a trigger. The hodoscope pruning algorithm will check if the event is a good event and set a flag (goodScinHit flag) for each of the events to be good or bad. In a simple sketch, as shown in Fig.4.40, the geometry for the SHMS spectrometer constrains the $\frac{xp_fp}{x_fp}$ to be 0.0018, and 0.0011 for HMS. The first hodoscope plane will measure the position x1. The predicted vertical position on the second hodoscope plane will be a function of the initial position x1 and the distance between the two hodoscope planes:

$$x_{project} = x1(1 + \frac{xp_fp}{x_fp}\text{distance_z}) \tag{4.23}$$

The predicted horizontal position y2 would be the same as y1. If the trigger position on the second plane meets the criteria (the difference between the predicted position and actual position is within 20 cm), this trigger is set to be a good event

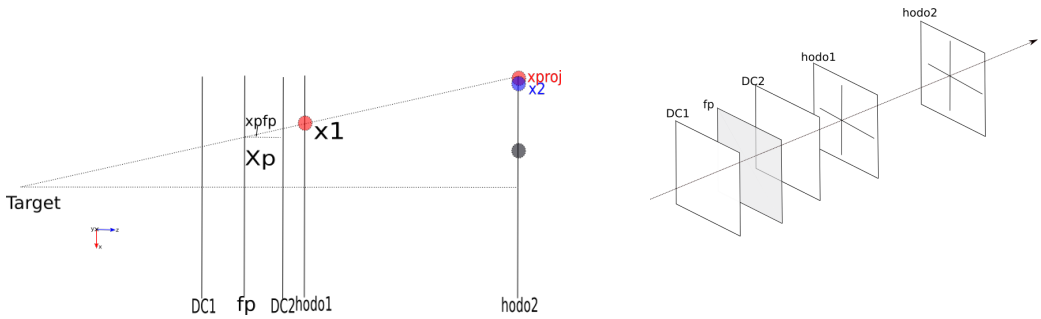


Figure 4.40: N_{should}, the number of expected triggers that should have a track

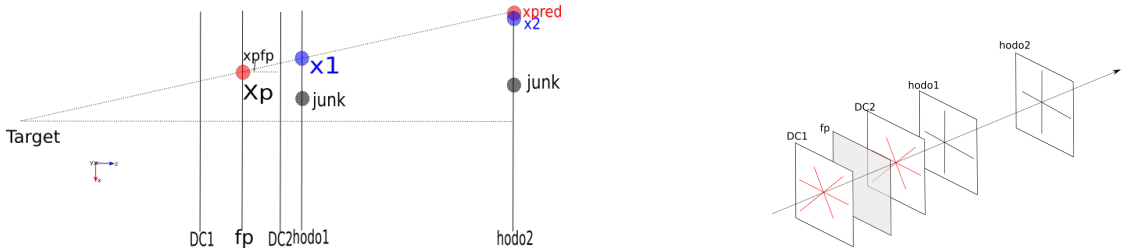
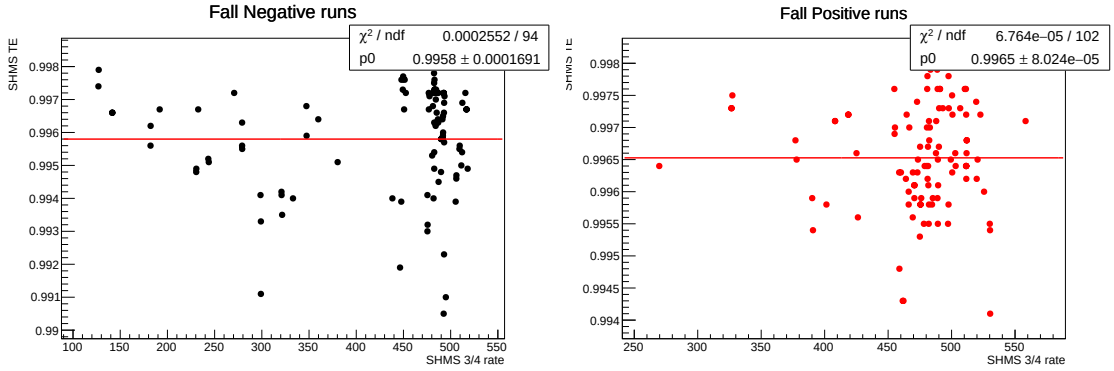


Figure 4.41: N_{did},found triggers

(goodScinHit is true). If the distance is much different than the grey point in Fig.4.40, the flag is set to be false. The N_{should} is selected by good event flag and electron (HMS tracking) or pion (SHMS tracking) PID cut without any tracking information. The N_{did} is the subset of N_{should}. For all the triggers from the expected sample, the number of triggers that did form good drift chamber tracks is N_{did}. This efficiency depends on the rate. If the event rate is too high, the number of hits in the drift chamber within the 200ns drift time might be considered too high to reliably find tracks in the drift chambers. The tracking efficiency as a function of rate is plotted as in Fig.4.42, Fig.4.43 and Fig.4.44. The tracking efficiency is applied to each run as it is.

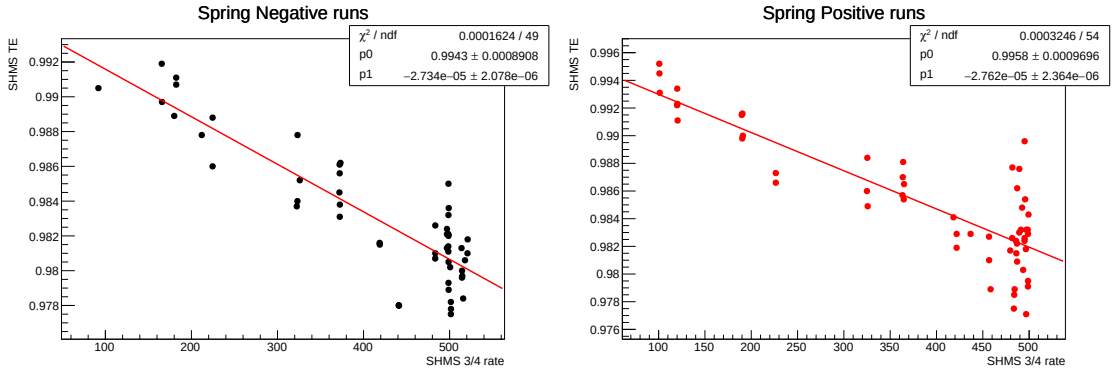
4.7 DAQ Live Time

Another efficiency comes from the Data Acquisition system (DAQ). The electronics get busy when there is a signal. If multiple signals piled up at a similar



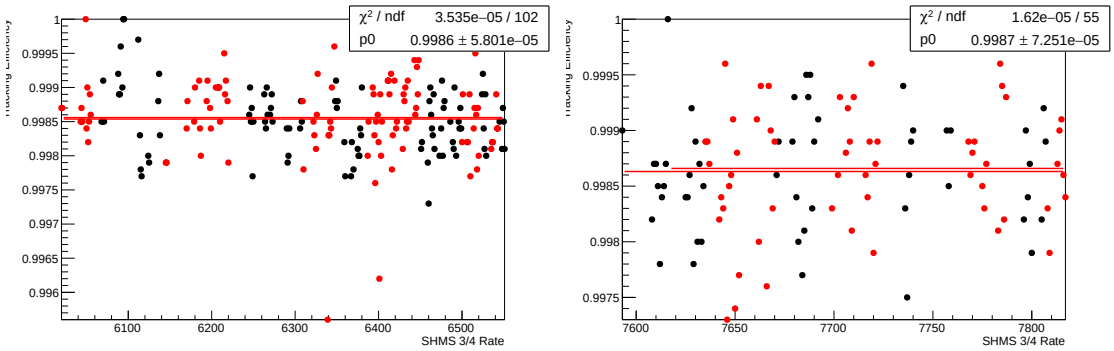
(a) Negative runs (b) Positive runs

Figure 4.42: SHMS tracking efficiency as a function of rate for Fall runs



(a) Negative runs (b) Positive runs

Figure 4.43: SHMS tracking efficiency as a function of rate for Spring runs



(a) Fall runs (b) Spring runs

Figure 4.44: HMS tracking efficiency as a function of the SHMS 3/4 event rate

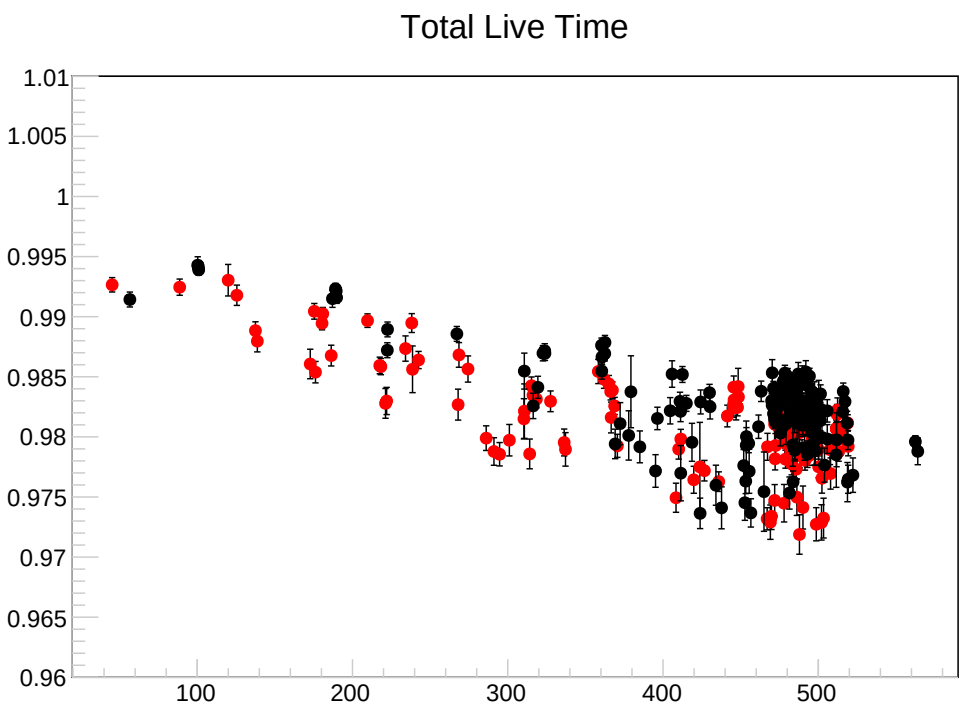


Figure 4.45: The total live time during the experiment as a function of the SHMS 3/4 rate

time, the electronics might not able to record all the events. This is electronic dead time. Another type of dead time is computer dead time, which is derived from the busy state of the data acquisition system. The computer dead time comes from the delay due to the data transfer process, the cable length, and the communication between ADC and TDC modules. The combined dead time inefficiency, which is total live time, is estimated by Electronics Dead Time Monitoring (EDTM) system[67][72]. The EDTM system introduces a known fixed-frequency signal into the data acquisition system. Since the signals are picked up by the data acquisition system, to avoid blocking of actual signal from the experiment, this known fixed frequency is set to be very low, at 10 Hz in the experiment. The ratio of accepted EDTM events and the known total events would be the total live time. The busier the event rate in the spectrometer, the more possible it might miss a good event like in Fig.4.45.

4.8 Target Density Correction

The experiment used liquid D_2 target and liquid H_2 target at really low temperatures as described in section 3.3. The high energy electron beam scattering on the target might cause some temperature fluctuations. The electron beam is already been spread to a $2 \times 2\ mm^2$ to avoid local burning by the raster system [90]. The raster system reduced temperature fluctuation, but the different currents might still make some difference in data yield.

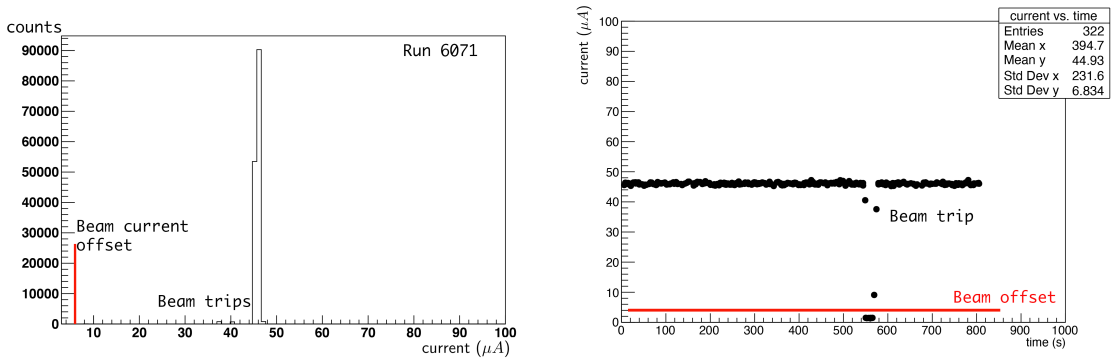
For the experiment, positive and negative polarity runs are taken to get the yield ratio. Since the rate for π^+ production is twice smaller than π^- production. To maximize the usage of beam time, the electron current is modified to get a similar positive event rate with a negative event rate. The difference of electron current for opposite polarity needs to be accounted for.

A series of runs were taken before this experiment with the same kinematic setting but a different current setting. The yield is then corrected by all the efficiency like rate-dependent tracking efficiency, live time, etc. The remaining effect of inefficiency should be only the target boiling. By comparing the yield from different current runs, the loss of yield due to target boiling can be studied. The target boiling correction used in this experiment is that the yield will lose $2.3 \pm 0.3\%$ per $100\mu A$ for liquid H_2 runs, $2.7 \pm 0.32\%$ per $100\mu A$ for liquid D_2 runs [45]. The equation for target boiling correction in this experiment is:

$$\varepsilon_{H_2} = 1 - \text{current}(2.3/10000) \tag{4.24}$$

$$\varepsilon_{D_2} = 1 - \text{current}(2.7/10000) \tag{4.25}$$

where the current is the current for this run in μA



(a) Beam current set at 46 μA

(b) Beam current vs. time

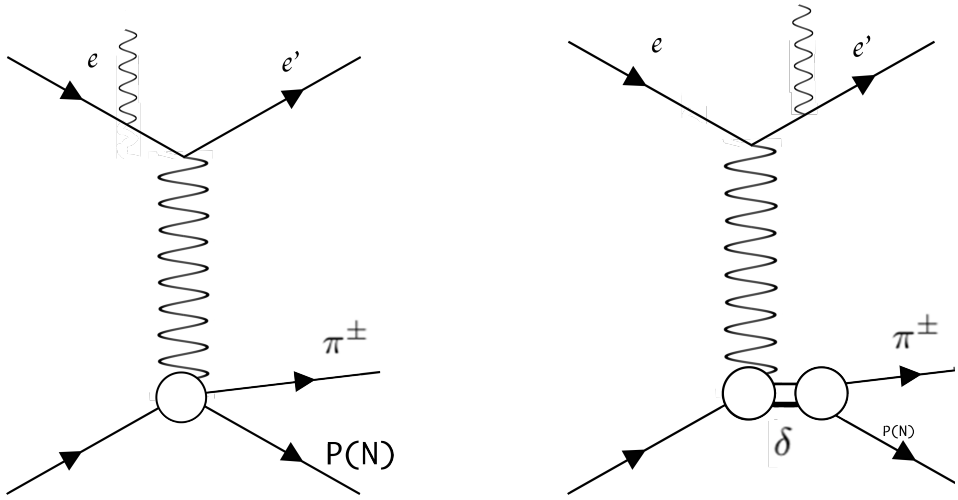
Figure 4.46: Beam current for one fall run. There is one beam trip during the run

4.9 Beam Current and Charge Accumulation

The current is measured by Beam Current Monitors (BCMs)[44]. The charge accumulation is needed for each run to calculate the pion production yield. Yield is defined as the number of counts in unit time (s) and unit charge (C). The yield can be compared between different runs and compared with simulations or other experiments. The total charge is calculated by integrating the beam current and time:

$$Q = \int I dt \tag{4.26}$$

During each run, there might be some beam trips due to any reason. Sometimes when the beam is not stable, the run might be a junk run. For example, one fall run in Fig.4.46, it is a 15 mins (900 s) run with the current setting at 46 μA, and there is one beam trip during the run. Beam current is recorded in the scaler tree together with time and event number *etc*. The event number corresponds to the event tree. So the beam current can be assigned to a data tree event. Then the beam trip is cut by a beam offset in both the scaler tree and event tree. For charge accumulation calculation, the charge where the beam current is below the offset is not counted in the total charge.



(a) Exclusive pion production (b) Delta pion production

Figure 4.47: Radiative backgrounds

4.10 Radiative Backgrounds

When an electron strikes a proton, it might rebound elastically, which is elastic scattering; it might disintegrate the proton completely, which is DIS as expected; it can kick the proton into a high energy state, which is the Δ resonance state. The yield of events from the radiative tail from exclusive pion production ($e + p \rightarrow e\prime + \pi^{+} + n$,$e + n \rightarrow e\prime + \pi^{-} + p$) was estimated using radiative effects implemented in the energy peaking approximation in SIMC combined with a fit to exclusive pion electroproduction for W above the resonance region [9][14] modified to reproduce the general features of the MAID model[20] at lower W. Generally, the contributions from those events are quite small, no more than 5% at z = 0.5, while growing to 7% at the largest z.

In addition, the pions from the decay of diffractively produced ρ particles will radiate into the spectrometer acceptance. The yield from these backgrounds should be subtracted from the semi-inclusive pion yield. At high energy, a photon emitted by the electron recoil might evolve into a vector meson before it actually

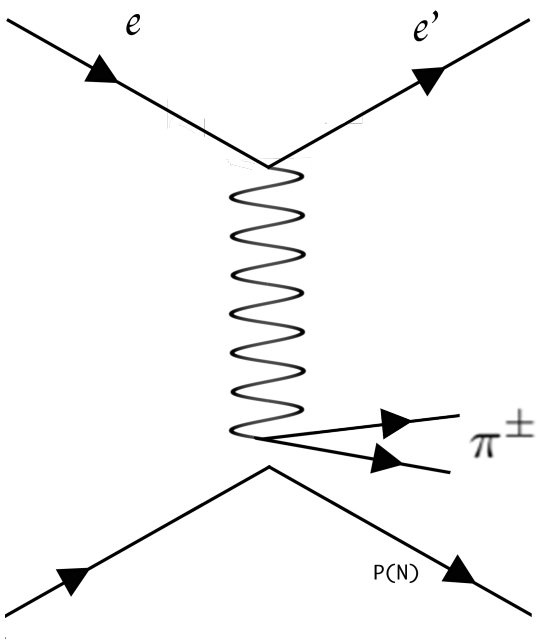


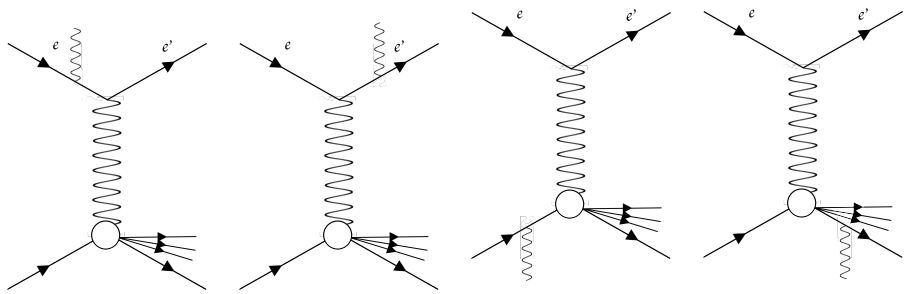
Figure 4.48: Diffractive ρ backgrounds

glace off the proton. The vector meson then disintegrates into pions as in Fig.4.48

The contribution from diffractive ρ production was estimated using the rho cross section from PYTHIA [77] modified to agree with HERMES data [4] with additional modifications to agree with the recent CLAS data at lower energies.

4.11 Radiative Correction

The charged particle moving in the experimental apparatus can lose its energy at any time during the experiment. The electromagnetic radiation produced by the change of its energy influenced by the Coulomb field of another charged particle is bremsstrahlung. These kinds of energy loss will change the real incident and scattered energy. The incoming and scattered electrons that interact with the nucleus involved in the scattering process are called internal bremsstrahlung. The electron can also lose energy anytime during the interaction with charged particles other



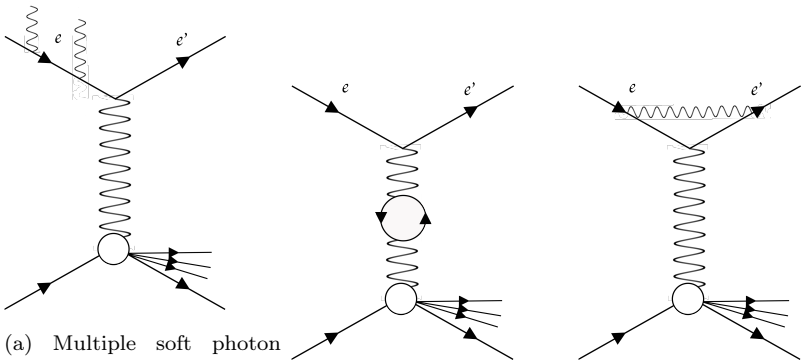
(a) electron current (b) electron current (c) hadron current (d) hadron current

Figure 4.49: First order bremsstrahlung

than the involved nucleus. Such processes are called external bremsstrahlung. Full first-order radiative correction derived by J.Schwinger[73], the Born cross section should be modified by a factor to get the measured cross section. Later Yennie, Frautschi, and Suura generalized the factor to higher orders for soft photon emission. Radiative correction formulas for coincidence electrons off protons reaction were then derived by Tsai [83] and Meister and Yennie [50]. The formulas are well documented in Mo and Tsai [56].

For the first order, internal bremsstrahlung of radiating a single photon is represented as photon emission by electron current (Fig.4.49a, Fig.4.49b) and photon emission by hadron current (Fig.4.49c, Fig.4.49d). Photon emission by hadrons is negligible. Higher-order processes such as multiple soft photon emission processes, interaction at the vertex, and vacuum polarization are shown in Fig.4.50. The electron-photon vertex correction is known exactly from QED. The vacuum polarization correction contains contributions from both lepton and hadronic loops, the former is calculated from QED and the latter is calculated in a similar way[23].

The electron straggling in the target is very similar to the radiative correction. It has to be taken into account since it's comparable with radiative corrections. The internal bremsstrahlung of electron straggling can be seen as two external radiators with one before and one after the scattering as in Fig.4.51. The effect



(a) Multiple soft photon emission (b) vacuum (c) vertex

Figure 4.50: High-order internal radiative correction

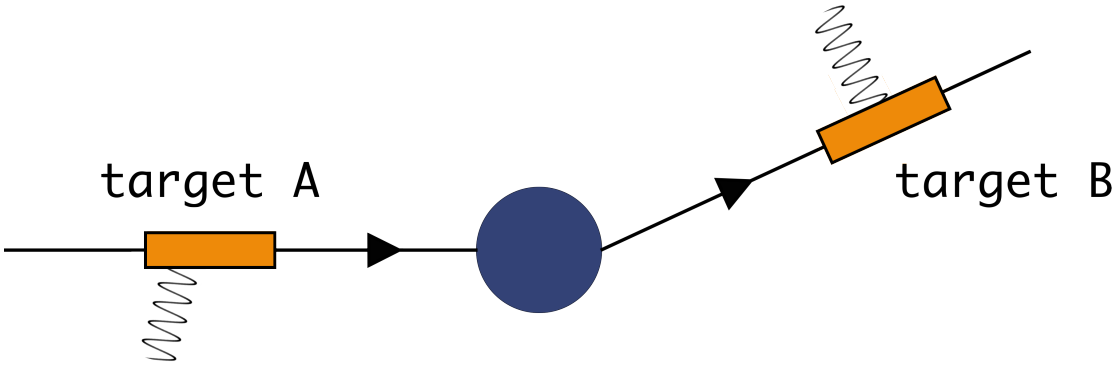


Figure 4.51: Electron straggling in the target

of electron straggling on cross section derived by Mo & Tsai [56]:

$$\sigma_{rad}(E_s, E_p) = \int_0^T \frac{dt}{T} \int_{E_{smin(Ep)}}^{E_s} dE'_s \int_{E_p}^{E_{pmax(E'_s)}} dE'_p I(E_s, E'_s, t)\sigma_r(E'_s, E'_p) I_e(E'_p, E_p, T-t) \tag{4.27}$$

where $I(E, E', t)$ represents the probability of finding an electron with initial energy E losing energy to E' after traveling a distance t (in units of radiation length) in the target with length T. The cross section $\sigma_r(E'_s, E'_p)$ is the Born cross section after internal radiative correction.

The radiative correction algorithm took care of all the above radiative processes. Based on the formalism originally applied to the inclusive electron scattering from Mo and Tsai[56], it is extended to the coincidence reactions and implemented in SIMC.

The yield of events from the radiative tail from exclusive pion production ($e+p(n) \rightarrow e' + \pi^{\pm} n(p) + \gamma$) is estimated using radiative effects implemented in the energy peaking

approximation in SIMC combined with a fit to exclusive pion electro-production for W above the resonance region. And the events from the radiative tail from the delta resonance region δ are also adapted in the SIMC.

The Born cross section of the process is modified due to radiation effects and the electron kinematics are also changed. These physical processes are experiment specific and are usually not included in the theoretical calculations. The process of removing these effects is the radiative correction. With a valid model Born cross section(σ_{born}^{model}) which covers all of the measured kinematic regions, the radiative cross section (σ_{rad}^{model}) can be calculated from it. The radiative correction factor becomes the ratio:

$$\text{RadCor} = \frac{\sigma_{born}^{model}}{\sigma_{rad}^{model}} \tag{4.28}$$

The data yield would be obtained by applying the radiative correction factor:

$$Y_{norad}^{data} = \text{RadCor} \cdot Y_{rad}^{data} \tag{4.29}$$

And the σ obtained from our data becomes:

$$\sigma_{born}^{data} = \text{RadCor} \cdot \sigma_{rad}^{data} \tag{4.30}$$

4.12 Data Quality Check

Data quality is checked in several different ways. The EPICS slowly read back the parameters of the devices every 30 seconds. It reads out magnet current values, the magnetic field strength, the pressure, and the temperature of the target. This experiment required a yield ratio of negative runs to positive runs to extract CSV values. It's important to make sure that the two polarity runs have the same values but different signs. The ratio between two polarities is plotted as the HMS Q1 current in Fig. 4.52. The current and magnetic field from all magnets is checked.

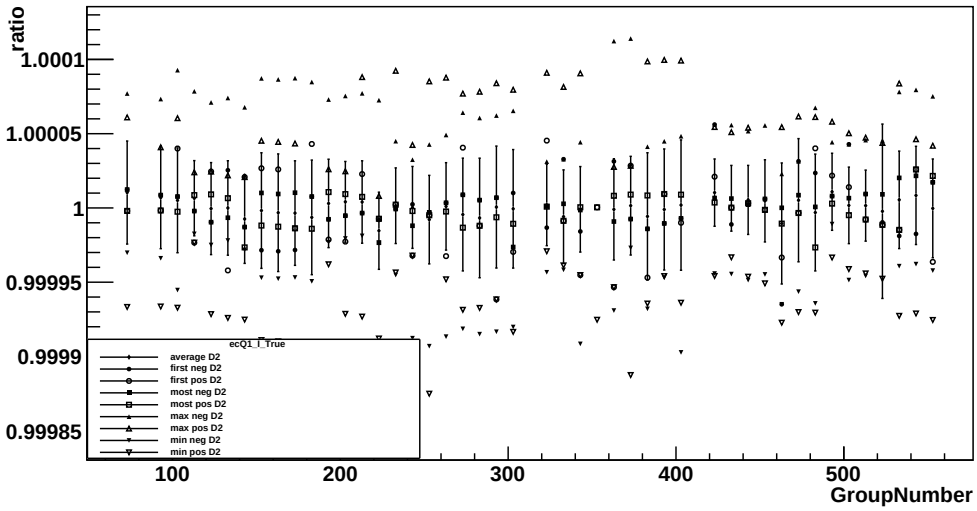
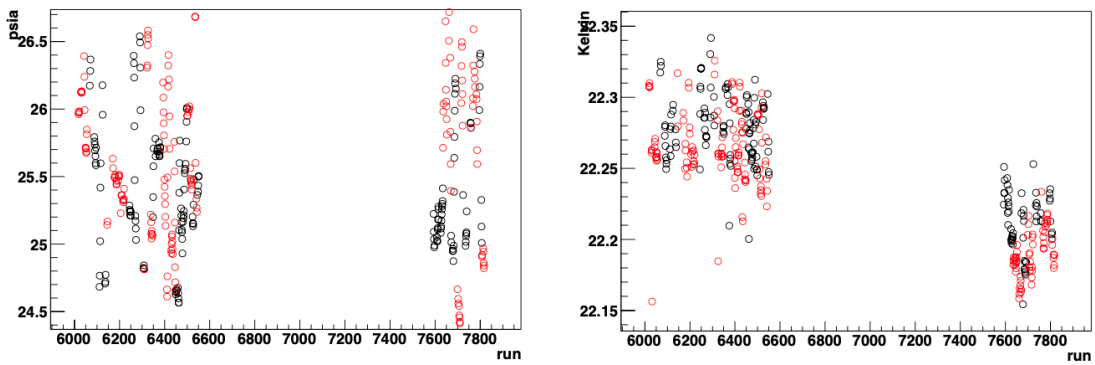


Figure 4.52: HMS quadruple Q1 current ratio for negative runs to positive runs. The average is plotted with standard derivation. The maximum, minimum, and first value ratios from each run and the mode are also tested.



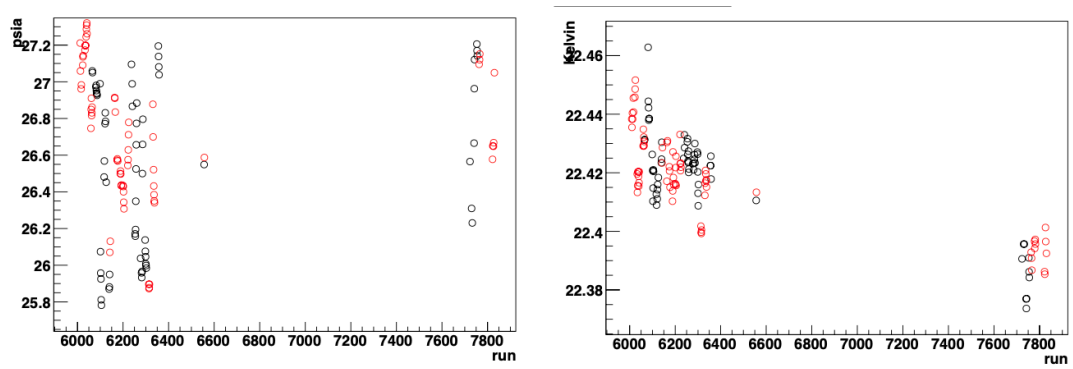
(a) Pressure in psi

(b) Temperature in Kelvin

Figure 4.53: Liquid D_2 target pressure and temperature check. Black circles are for negative runs, red circles are for positive runs.

The target variables are also checked. There are pressure and temperature probe at the target cells which return the values to the EPICS. The EPICS recorded the values slowly to the raw data file. The pressure and temperature are plotted for D_2 in Fig.4.53 and for H_2 in Fig.4.54. From the National Institute of Standards and Technology database [61], density can be calculated as shown in Fig. 4.55. This is checked during the experiment. The density fluctuation would cause a change in the final data yield. This is corrected by the target boiling correction.

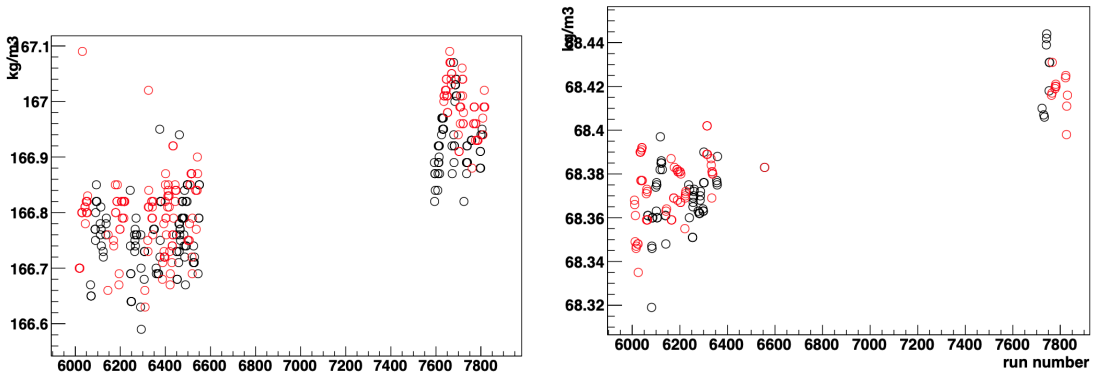
Another check is between different runs in the same run group. Each kinematic



(a) Pressure in psi

(b) Temperature in Kelvin

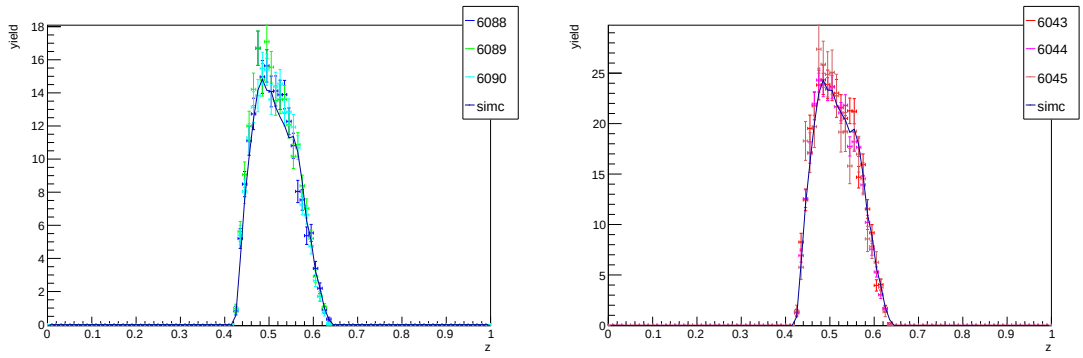
Figure 4.54: Liquid H_2 target pressure and temperature check. Black circles are for negative runs, red circles are for positive runs.



(a) The density of liquid D_2

(b) Density pf liquid H_2

Figure 4.55: Liquid D_2 and H_2 density



(a) Negative polarity runs

(b) Positive polarity runs

Figure 4.56: Yield consistency check between each runs in the same kinematic run groups

setting is taken by several small identical one-hour runs. And the simulation is done by SIMC for the semi-inclusive DIS process with radiative effect. The comparison between different runs is plotted for one run group as one example in Fig.4.56. All run groups are checked.

For one (x, Q^2) setting, four z settings are taken. Because of the acceptance of the spectrometer, there will be some overlap between different center z settings. The yield ratio $R_Y = \frac{Y_{neg}}{Y_{pos}}$ are checked between each center z setting. For the x centered at 0.45, Q^2 centered at 4.75 GeV^2, four center z scan are taken at $0.4, 0.5, 0.6, 0.7$. The yield ratios are plotted as a function of z as in Fig.4.57. The yield ratio should be a function of both x and z. But here in the figure, the yield is integrated over x and Q^2 for the kinematic setting.

4.13 Bin Centering Correction

To extract CSV from measure quantity R^D_{meas}, it is important to get the accurate kinematic value. The R^D_{meas} depends on both x_{bj} and z, and the calculation of all other terms relies on accurate kinematic variables. The x_bj and z bin width are defined as 0.05 in this experiment. For the large bin width, the problem arises as to which value inside the bin width should represent the kinematic value of this bin (zi, xj), as the

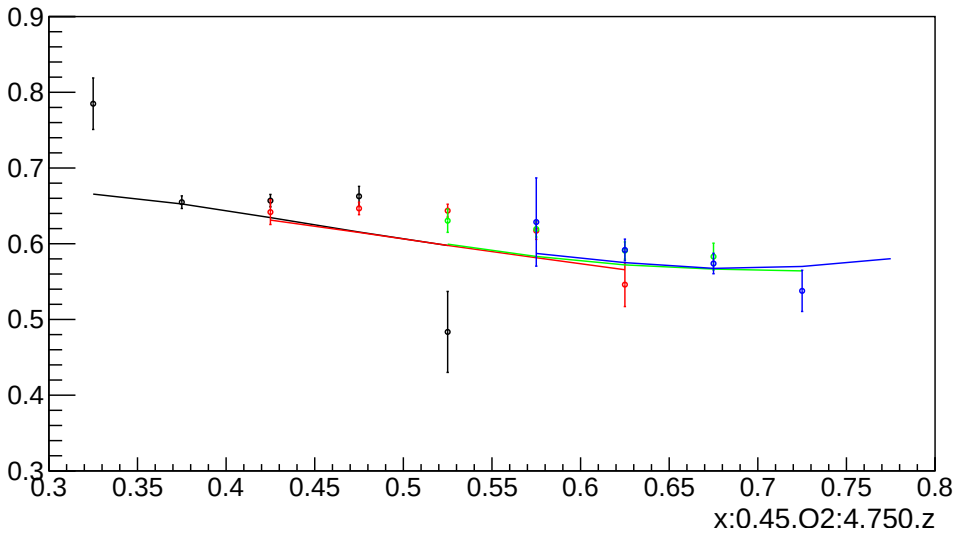


Figure 4.57: Yield ratio for one (x, Q^2) setting. Lines are yield ratios from SIMC simulation.

difference between the average and the center in Fig.4.58. The bin centering factor for each kinematic variable to get the average kinematic value for each kinematic bin is used here:

$$[f_{bc}^{x_bj}]_{(zi,xj)} = [\frac{x_{bj,ave}}{x_{bj,center}}]_{(zi,xj)} \tag{4.31}$$

The average for the kinematic quantity X in one kinematic bin (zi, xj) is calculated as

$$\overline{X}_{(zi,xj)} = [\frac{\sum_k^N X_k}{N}]_{(zi,xj)} \tag{4.32}$$

where the X_k is the value of the kinematic variable X for the event k in the N entries of the kinematic bin. And the variance of that bin is

$$Var_{(zi,xj)} = [\frac{\sum_k^N X_k^2}{N} - \overline{X}^2]_{(zi,xj)} \tag{4.33}$$

The weighted average is used for the bins where multiple runs overlap, including the yield ratio between negative and positive polarity.

$$X_{ave}^{pos}|_{(zi,xj)} = \frac{\sum \overline{X}_{runs}/Var_{runs}}{\sum 1/Var_{runs}}|_{(zi,xj)} \tag{4.34}$$

$$X_{ave}^{ratio}|_{(zi,xj)} = \frac{\frac{X_{ave}^{pos}}{Var^{pos}} + \frac{X_a^{neg}ve}{Var^{neg}}}{\frac{1}{Var^{pos}} + \frac{1}{Var^{neg}}}|_{(zi,xj)} \tag{4.35}$$

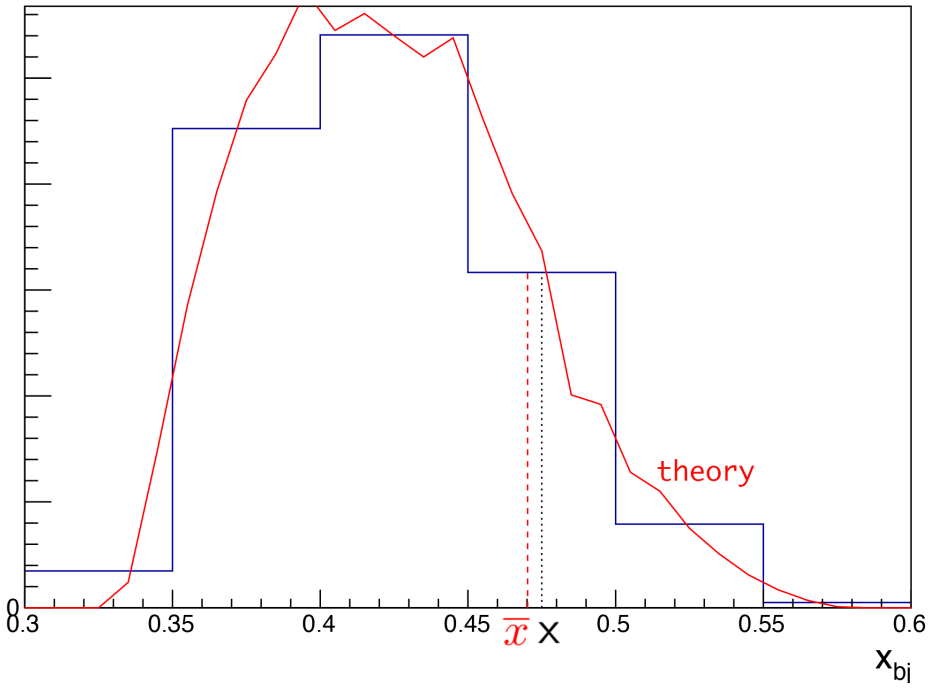


Figure 4.58: Sketch of bin center correction in the 1D case

For the measured value R^D_{meas} of each 2D kinematic bin (x, z), all the kinematic variables would have the average value. And the further calculation is using the average of all the kinematic variables for better precision.

Source	Fall Uncertainty	Spring Uncertainty
Current	0.45%	0.56%
Target Boiling	0.05%	0.07%
FADC correction	0.5%	0.1%
Tracking efficiency	$< 0.02\%$	$< 0.02\%$
Dead time	$< 0.02\%$	$< 0.02\%$

Table 4.5: Systematic uncertainty

Source	Uncertainty
Acceptance	1.5%
PID efficiency	related to kinematics
radiative backgrounds	related to kinematics

Table 4.6: Systematic uncertainty

polarity	current(μA)	uncertainty(μA)	min	max
neg	a	0.1	a-0.1	a+0.1
pos	b	0.1	b-0.1	b+0.1

Table 4.7: current uncertainty example

4.14 Systematic Uncertainties

4.14.1 Current Uncertainty

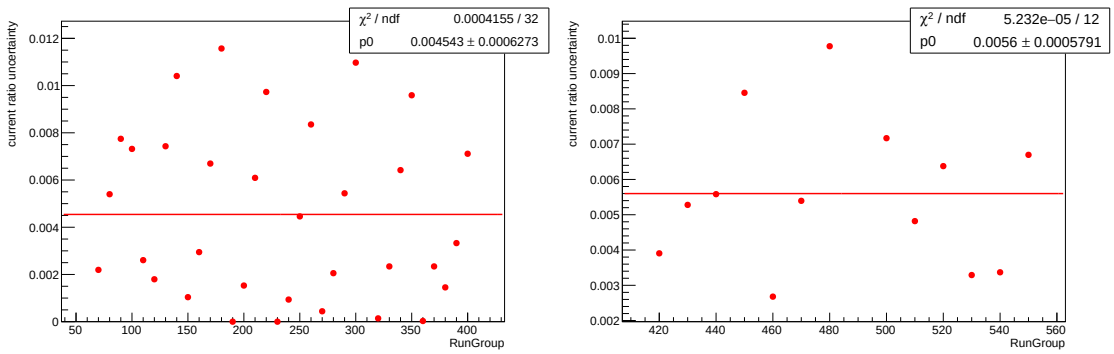
The current uncertainty is due to the uncertainty related to the Unser current monitor. The Unser monitor is a parametric current transformer used to measure the beam current. It provides non-destructive measurement and gives an absolute reference. The Unser monitor's output signal drifts and it's calibrated during the run period. The uncertainty from the beam current measurement is $\pm 100mA$ for the experiment.

For one setting, there are many runs taken with different charges. So the current for this RunGroup is normalized by the charge for each run.

$$Current = \frac{\sum_{runs} charge * current}{\sum_{runs} charge}$$

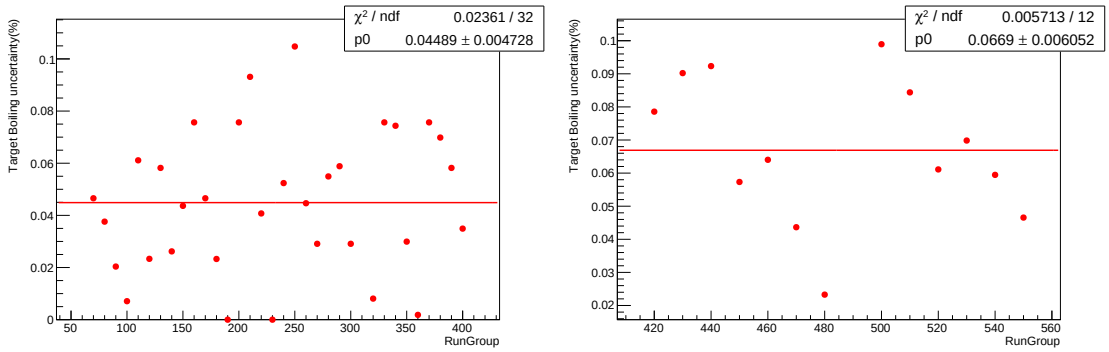
The current ratio uncertainty is calculated by the negative current divided by the positive current with current uncertainty ($100m\mu A$) as in Table.4.7.

$$\text{Currentuncertainty} = \frac{\frac{a-x}{b-x} - \frac{a+x}{b+x}}{a/b} \tag{4.36}$$



(a) Fall runs (b) Spring runs

Figure 4.59: Current ratio uncertainty vs. RunGroup



(a) Fall runs (b) Spring runs

Figure 4.60: Target boiling correction ratio uncertainty in percentage vs. RunGroup

The current ratio uncertainty as a function of each RunGroup is in Fig.4.59. The systematic uncertainty related to the current is around 0.45% for the fall runs and 0.56% for the spring runs.

4.14.2 Target Boiling uncertainty

The target boiling for the D2 is $2.3\% \pm 0.3\%$ per $100\mu A$, for H2 is $2.7\% \pm 0.32\%$ per $100\mu A$. The current for each RunGroup is still charge normalized. And the uncertainty due to the Target Boiling effect in the D_2 yield ratio is calculated by

$$\text{TB_uncertainty} = 0.3 \times (\text{neg_current} - \text{pos_current})/100 \tag{4.37}$$

The target boiling uncertainty as a function of RunGroup is in Fig.4.60 The uncertainty

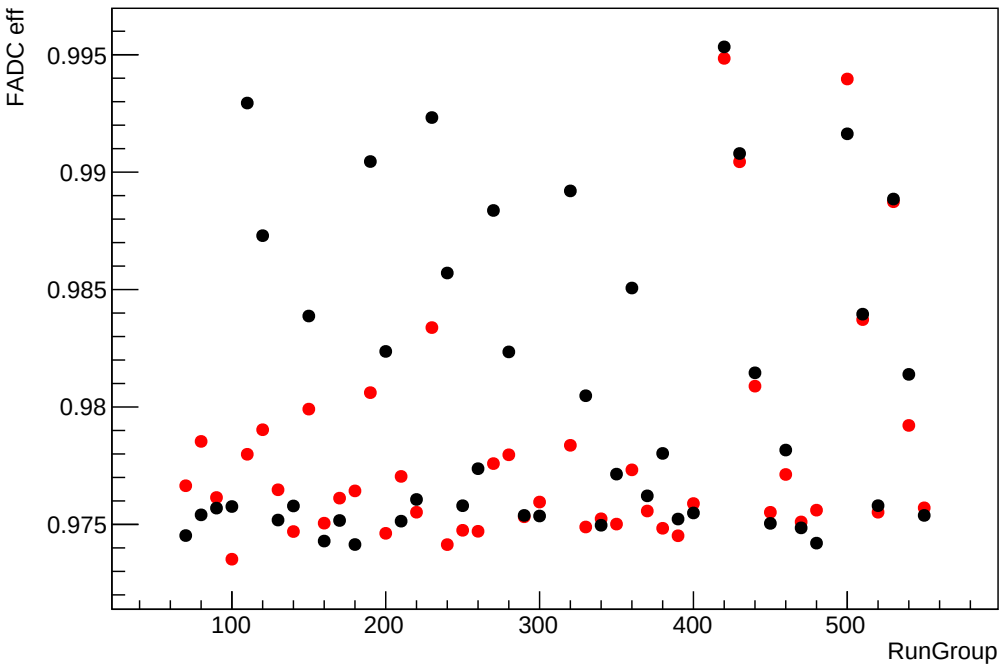


Figure 4.61: FADC dead time correction

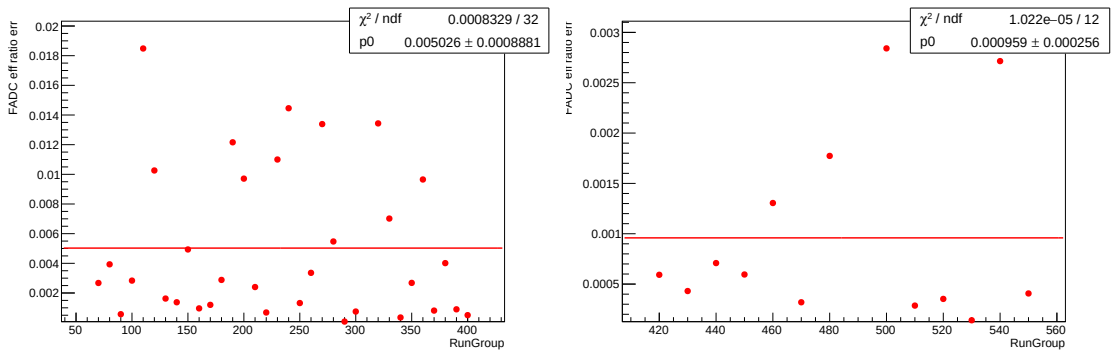
from the target boiling ratio is 0.045% for the fall runs and 0.067% for the spring runs.

4.14.3 FADC Dead Time Correction Uncertainty

The Flash Analog Digital Converter (FADC) dead time is due to the time window block of the previous particles. An electronic device receives the pulse of a signal and becomes busy. The signal from the next event might be ignored. The calorimeter has hundreds of PMTs. Some signals might be picked up by the other PMTs. But the Cherenkov detector has very limited PMTs. The main source of FADC dead time is the Cherenkov detector. The FADC dead time correction is related to the SHMS 3/4 trigger rate (in kHz) and HMS 3/4 trigger rate (in kHz):

$$\text{FADC_Efficiency} = 1 - 0.05 \times (\text{SHMS_rate}/1000 + \text{HMS_rate}/1000) \tag{4.38}$$

The efficiency decrease with the trigger detecting rate with a slope of 50 ± 30 ns. The



(a) Fall runs

(b) Spring runs

Figure 4.62: FADC dead time correction ratio uncertainty vs. RunGroup

FADC uncertainty is calculated by

$$\text{FADC_uncertainty} = \frac{(\text{FADC_neg/FADC_pos})|_{max} - \text{FADC_neg/FADC_pos})|_{min}}{\text{FADC_neg/FADC_pos})} \tag{4.39}$$

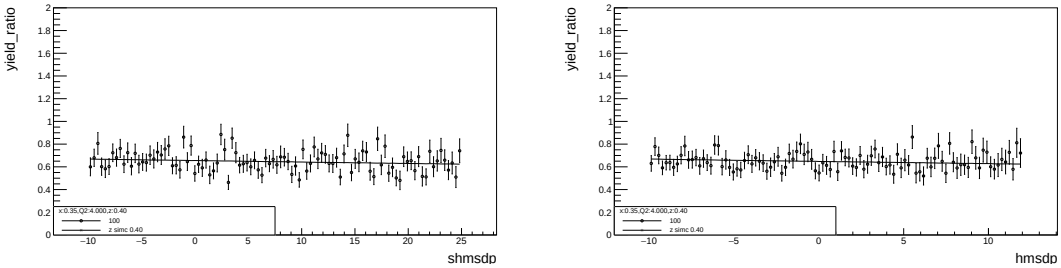
For the spring runs, the current for positive runs and negative runs are well controlled to let the event rate be the same for both polarities. So the FADC correction for the spring runs is similar to the negative and positive runs. The uncertainty of the FADC correction ratio is limited. For the fall runs, the FADC correction ratio uncertainty is 0.5% and for the spring runs, this number is 0.1%

luminosity is related to current and target density. The uncertainty from the luminosity would be propagated from current ratio uncertainty and target boiling uncertainty

4.14.4 Acceptance Uncertainty

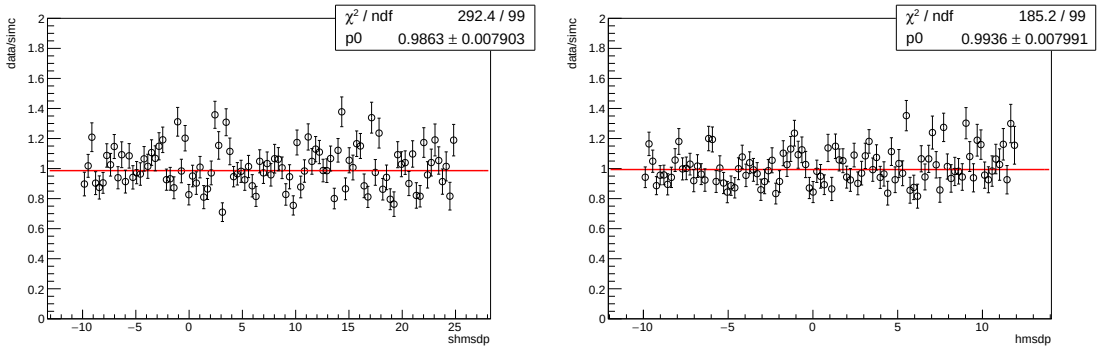
The SHMS and HMS acceptance uncertainty is studied by comparing the pion yield ratio $N_D^{\pi^-}/N_D^{\pi^+}$ from data from SIMC. For the high z region, the radiative backgrounds procedure like exclusive pion production, delta resonance, and diffractive ρ become large. So the lowest center $z = 0.4$ setting is examined.

The yield ratios (RY) as a function of δp which is the percentage of the momentum



(a) SHMS δp (b) HMS δp

Figure 4.63: Yield ratio from data (in points) compared with yield ratio from SIMC (in line)



(a) SHMS δp (b) HMS δp

Figure 4.64: Ratio of RY ($N_D^{\pi^-}/N_D^{\pi^+}$) from data divided by yield ratio from SIMC for the SHMS and HMS

from central momentum $\delta p = (p - p_{central})/p_{central}$ for both spectrometer is examined in Fig. 4.63. The data yield ratios are compared with the SIMC yield ratios.

The RY of data divided by SIMC is shown in Fig.4.64. A horizontal line is fitted to the ratio. The result is close to 1 and the uncertainty is around 1.5%.

4.14.5 PID Systematic Uncertainty

As discussed in Section 4.5.7, the pion selection at high momentum can be chosen in two ways: 1. Apply heavy gas number of photon electron cut, which will lose $10\% to 15\%$ good pion counts; 2. Don't apply heavy gas number of photon electron cut, apply pion purity to the sample yield. For the second case, the good pions are not removed but an extra correction is applied. The pion purity is applied to both positive and negative yield and then to the yield ratio.

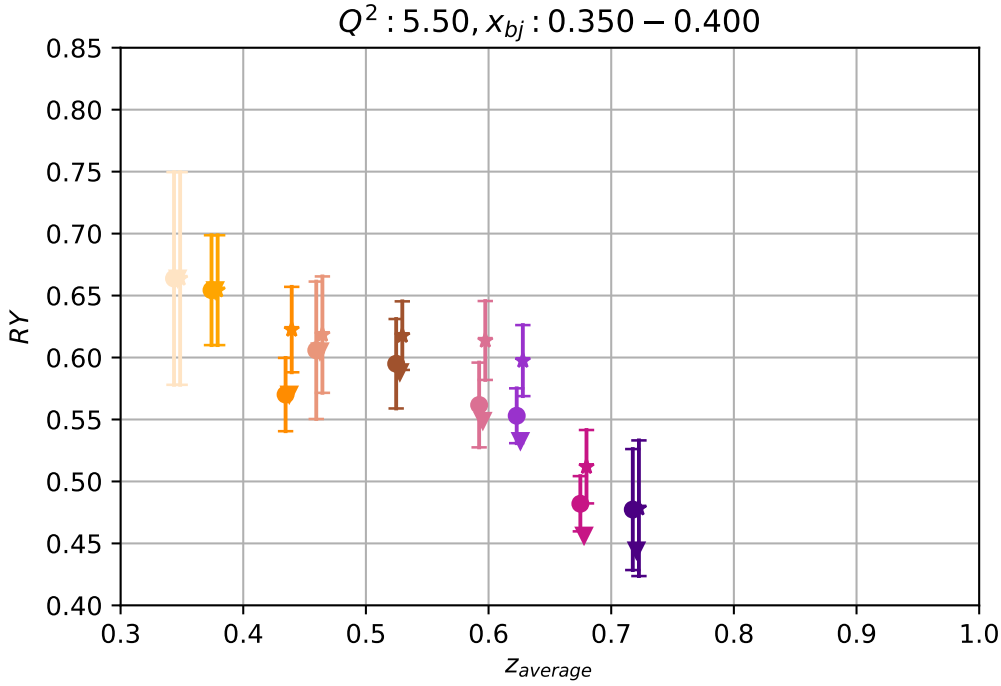


Figure 4.65: Yield ratio RY compared with the two cases for one setting. The star is "with heavy gas cut", the circle is "no heavy gas cut, pion purity corrected" and the triangle is "no heavy gas cut, no purity corrected".

The difference between the two cases is compared for each setting. One is shown in Fig. 4.65. The error bar from the yield ratio is not pion purity corrected. The heavy gas cut will remove some good pions, so the error bar from the "with heavy gas cut" case will be slightly larger than the "without heavy gas cut" case. For the pion purity, there are more K^+ produced than K^-, so the pion purity for the positive yield is lower than the purity in the negative yield. The yield ratio RY is defined as negative yield divided by positive yield so that the pion purity correction would increase the result (from the triangles to the circles). The difference between the yield ratio RY of the two cases is the systematic uncertainty from the particle identification as shown in Fig.4.66.

$$dRY = \frac{|RY_{withHGC} - RY_{noHGC,corr}|}{RY_{withHGC}} \tag{4.40}$$

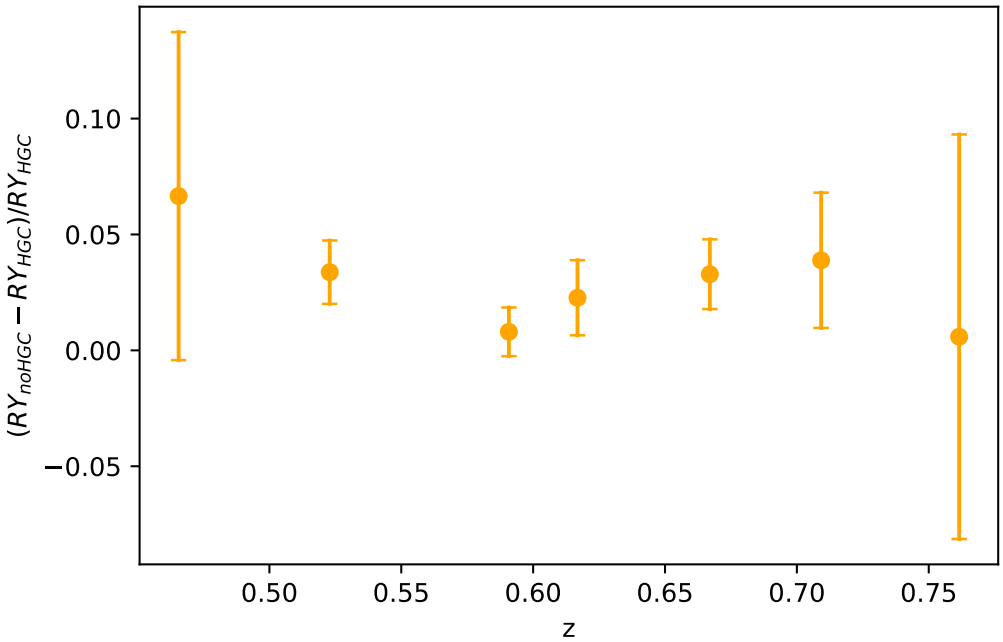


Figure 4.66: The difference of R_Y between the two cases

4.14.6 Radiative Backgrounds Uncertainty

The radiative backgrounds come from exclusive pion production, the delta resonance, and the diffractive ρ backgrounds. All the backgrounds are estimated by SIMC. Changing the background production model, the effect will change the CSV results. The uncertainty on CSV due to the radiative backgrounds is discussed in the results part.

Chapter 5

RESULTS

5.1 Factorization

As described in Section 2.6, the Hydrogen data are taken for some settings to verify the factorization assumption in JLab 12 GeV data. To verify the factorization assumption, two quantities are formed. One is the difference of π^+ scattered from the H_2 target minus the π^- yield divided by the same difference from the D_2 target. The other one is the summation of π scattered from H_2 divided by the summation from D_2 target.

$$\frac{\sigma_p^{\pi^+} + \sigma_p^{\pi^-}}{\sigma_d^{\pi^+} + \sigma_d^{\pi^-}} = \frac{4u + 4\overline{u} + d + \overline{d}}{5(u + \overline{u} + d + \overline{d})} \tag{5.1}$$

and

$$\frac{\sigma_p^{\pi^+} - \sigma_p^{\pi^-}}{\sigma_d^{\pi^+} - \sigma_d^{\pi^-}} = \frac{4u_v - d_v}{3(u_v + d_v)} \tag{5.2}$$

The two quantities should be independent of z. The H_2 data are taken for five (x, Q^2) settings, the quantity from all of the settings are shown in Fig.5.1 and Fig. 5.2.

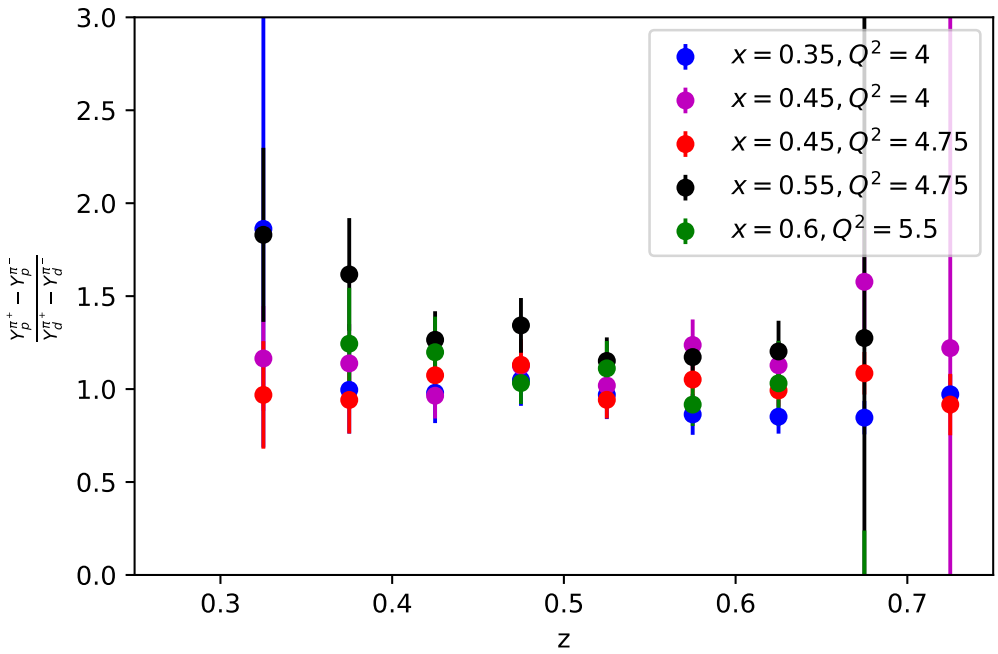


Figure 5.1: The difference ratio results of proton to deuterium pion scattering yield.

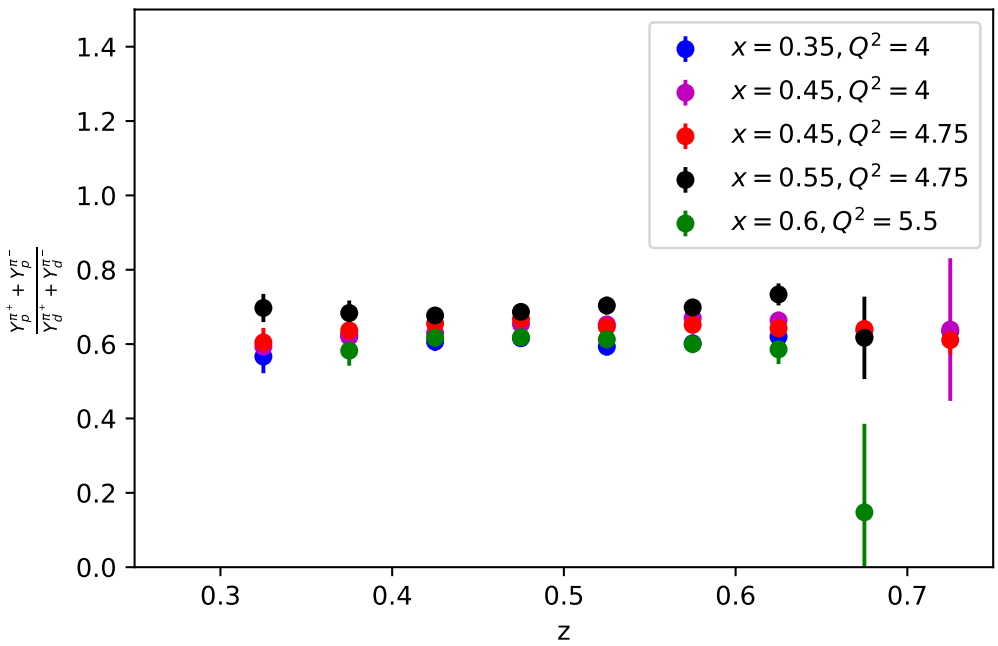


Figure 5.2: The sum ratio results of proton to deuterium pion scattering yield.

5.2 CSV

The CSV results are then extracted from the yield ratio from the pion production data of the deuterium target.

5.2.1 Yield Ratio

The yield of pions scattered from the deuterium target is first corrected by all the detecting efficiencies. Then all the radiative backgrounds are estimated and subtracted from the data yield. The ρ radiative background is subtracted from the data with a coefficient γ, which would change.

$$\begin{aligned} Y_{\text{target}}^{\text{pions}} &= \frac{N_{\text{pions}}}{Q \cdot \epsilon_{LT} \cdot \epsilon_{TB} \cdot \epsilon_{FADC} \cdot \epsilon_{trk} \cdot \epsilon_{PID}} \\ Y_{D_2}^{\text{data}} &= Y_{D_2}^{\text{pions}} - 0.245 \cdot Y_{\text{Dummy}}^{\text{pions}} \\ Y_{D_2}^{\pi^-/\pi^+} &= \text{raida_corr}(Y_{D_2}^{\text{data}} - Y_{\text{exc}} - Y_{\text{delta}} - \gamma Y_\rho) \end{aligned} \tag{5.3}$$

The yield ratio R_Y is calculated by the negative pion yield divided by the positive pion yield.

$$R_Y = Y^{D\pi^-}(x,z)/Y^{D\pi^+}(x,z) \tag{5.4}$$

The quantity R_{meas}^D is a function of the yield ratio:

$$R_{meas}^D(x,z) = \frac{4N^{D\pi^-}(x,z) - N^{D\pi^+}(x,z)}{N^{D\pi^+}(x,z) - N^{D\pi^-}(x,z)} \tag{5.5}$$

$$= \frac{4R_Y(x,z) - 1}{1 - R_Y(x,z)} \tag{5.6}$$

which is plotted as function of (x,z) for three distinct Q^2. The bins are selected to be 20 bins for x_{bj} from 0 to 1, and 20 bins for z from 0 to 1. The edge bins are empty. There are several central x_{bj} settings are taken for each Q^2, with several different central z scan settings. The overlap area for each kinematic point is merged by taking the weighted average among all the different runs. And all kinematic variables are bin-center corrected. The R_{meas}^D for all three Q^2 are shown in Fig. 5.3,5.4,5.5. Here the values are projected on the hadron fraction z axis.

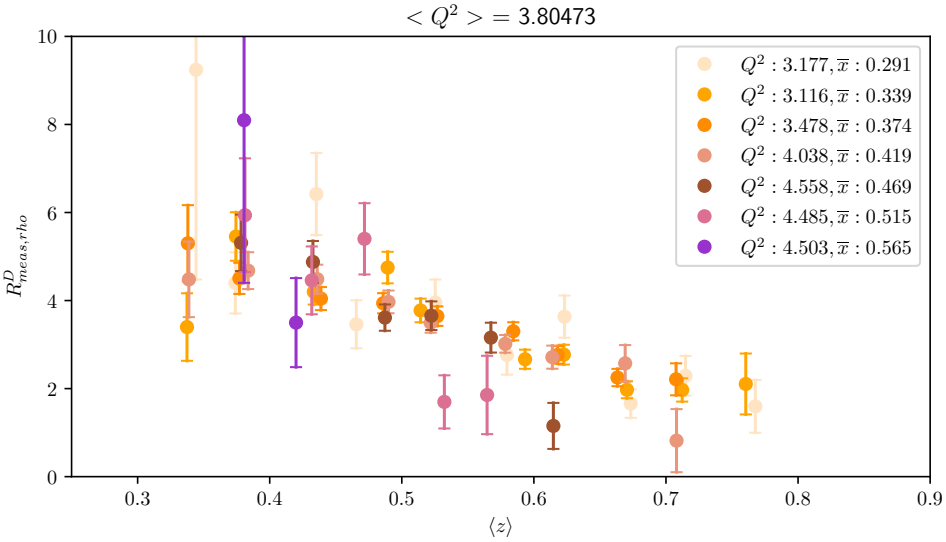


Figure 5.3: R^D_{meas} for the first Q^2 setting

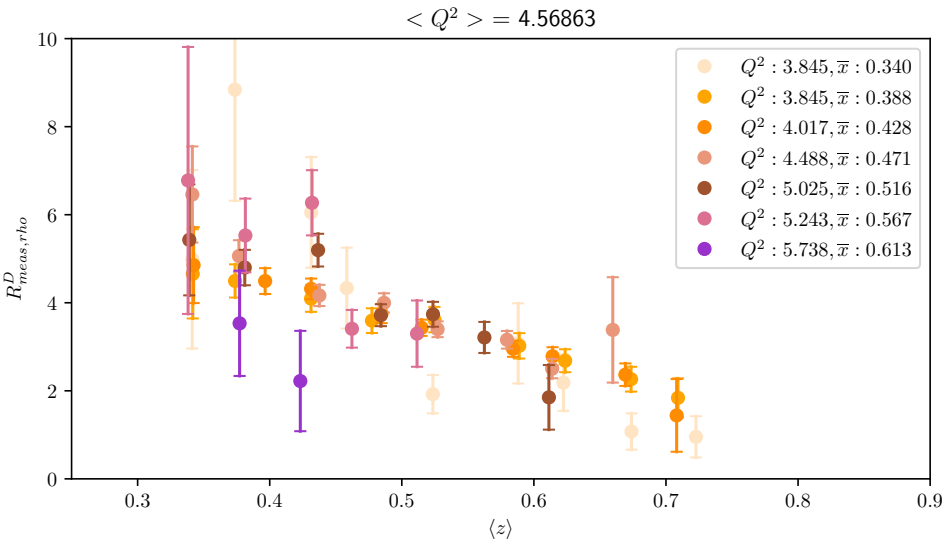


Figure 5.4: R^D_{meas} for the second Q^2 setting

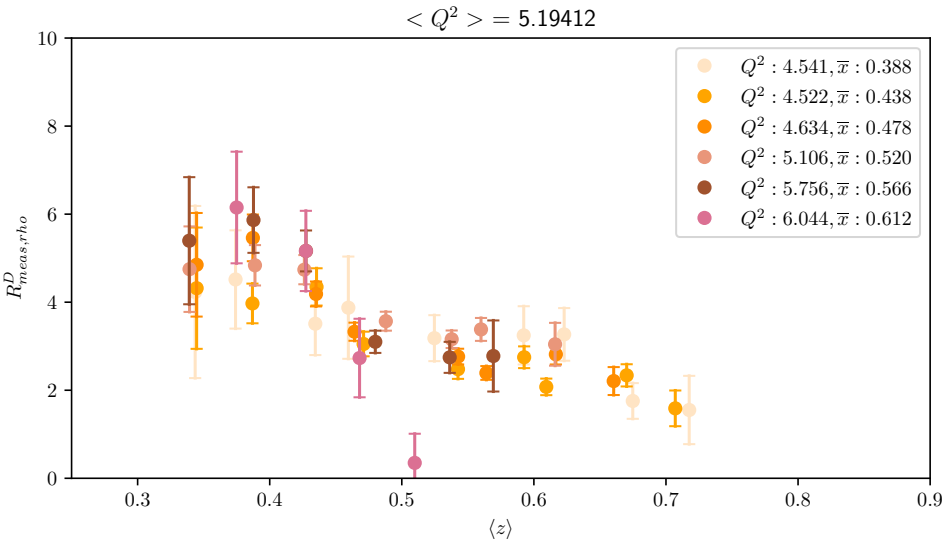


Figure 5.5: R^D_{meas} for the third Q^2 setting

5.2.2 Extraction of CSV

The CSV is then extracted from the R^D_{meas} by the equation:

$$D(z)R(x,z) + A(x)CSV(x) = B(x,z) \tag{5.7}$$

where

$$\begin{aligned} D(z) &= \frac{1-\Delta(z)}{1+\Delta(z)}, \Delta(z) = \frac{D_u^-(z)}{D_u^+(z)} \\ R(x,z) &= \frac{5}{2} + R^D_{meas}, \\ CSV(x) &= (\delta d - \delta u), \\ A(x) &= \frac{-4}{3(u_v + d_v)}, \\ B(x,z) &= \frac{5}{2} + R^D_{sea_S}(x,z) + R^D_{sea_NS}(x). \end{aligned} \tag{5.8}$$

The $A(x)$ term and $B(x,z)$ term are calculated from the existing PDFs and fragmentation functions as in Fig.5.6 and Fig. 5.8. The $B(x,z)$ term is related to sea strange quarks

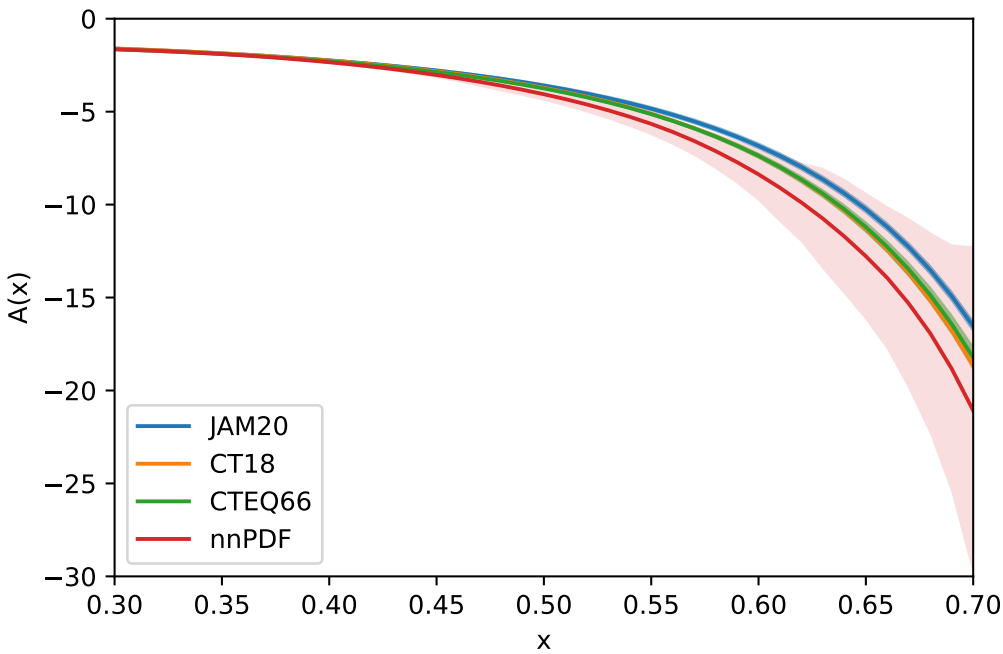


Figure 5.6: $A(x)$ term calculated from JAM20 PDFs[57], nnPDF[8] and CTEQ PDFs[89]

and sea non-strange quarks.

$$
\begin{aligned}
B(x,z) &= \frac{5}{2} + R^D_{sea_S}(x,z) + R^D_{sea_NS}(x) \\
R^D_{sea_S} &= \frac{\Delta_s(z)[s(x)+\overline{s}(x)]/(1+\Delta(z))}{[u^p_v(x)+d^p_v(x)]} \\
\Delta_s(z) &= \frac{D^-_s(z)+D^+_s(z)}{D^+_u(z)} \\
R^D_{sea_{NS}} &= \frac{5(\overline{u}^p(x)+\overline{d}^p(x)}{[u^p_v(x)+d^p_v(x)]}
\end{aligned}
\tag{5.9}
$$

The quantity from our measurement R^D_{meas} can then be written as:

$$
R^D_{meas}(x,z) = \frac{B(x,z) - A(x)CSV(x)}{D(z)} - \frac{5}{2} \tag{5.10}
$$

To extract CSV from the measured quantity R^D_{meas}, the unfavored to favored fragmentation ratio is assumed to be a function of z with two parameters:

$$
\Delta = z^\alpha (1-z)^\beta \tag{5.11}
$$

And the CSV term is assumed to be a function of x with three parameters:

$$
CSVx = x^a (1-x)^b (x-c) \tag{5.12}
$$

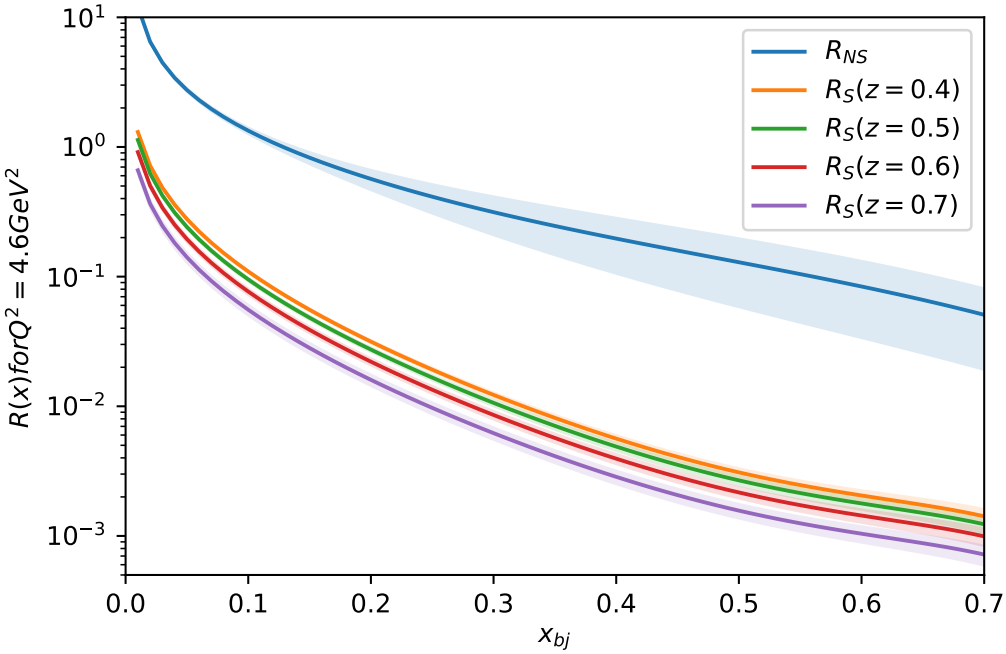


Figure 5.7: $B(x, z)$ terms calculated from CTEQ6 PDFs[68] and JAM20 fragmentation functions[57]

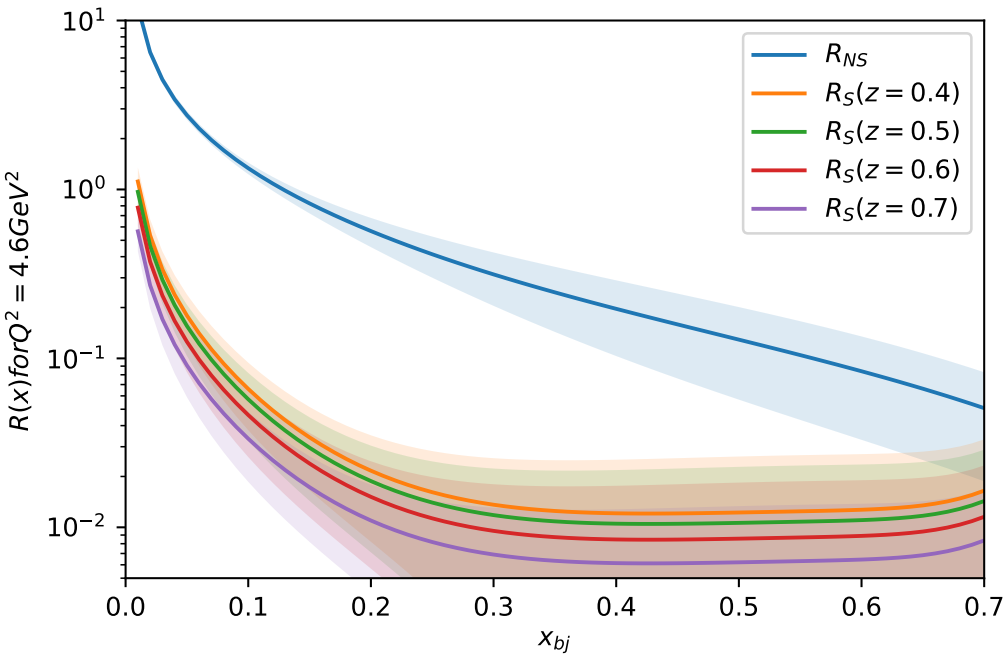


Figure 5.8: $B(x, z)$ calculated from JAM20 PDFs and fragmentation functions.

Also, we have the constrain that the integral of CSV of x from 0 to 1 to be zero:

$$\int_0^1 CSV x dx = 0 \tag{5.13}$$

From this constrain, the third parameter is constrained by the first two parameters:

$$c = \frac{\int_0^1 x^{(a+1)}(1-x)^b}{\int_0^1 x^a(1-x)^b} \tag{5.14}$$

where can be solved by the special function, beta function, which is a function of lambda function:

$$c = \frac{Beta(a+2, b+1)}{Beta(a+1, b+1)} \tag{5.15}$$

For the ρ radiative background subtraction, a coefficient is assumed to vary the size of ρ background correction. The ρ background is subtracted from the data yield:

$$\begin{aligned} Y_{corr}^{D\pi^-} &= Y^{D\pi^-} + \gamma Y_{\pi^-}^{\rho}, \\ Y_{corr}^{D\pi^+} &= Y^{D\pi^+} + \gamma Y_{\pi^+}^{\rho} \end{aligned} \tag{5.16}$$

which will slightly change the measured value R_{meas}^D by a small quantity:

$$\delta R_{\rho}^D(\gamma, Y_{\pi^+}^{\rho}, Y_{\pi^-}^{\rho}, Y^{D\pi^+}, Y^{D\pi^-}) \tag{5.17}$$

This small quantity is added in the fitting process to get the subtraction coefficient γ.

Fitted by Minuit[19], both $CSV(x)$ and $D(z)$ are extracted from our data. For fragmentation function ratio in Fig. 5.9, the blue line is the fitting result with 1σ error band. The CSV fitting results is the blue line in Fig. 5.10, also with 1σ error band. The CSV can be calculated by the fragmentation function ratio results, which are the points shown in Fig. 5.10.

The fragmentation function ratio extracted from the experiment data is close to the ratio calculated from several existing fragmentation functions. The CSV is then calculated by different existing PDFs and FFs by the equation derived from Eq.5.7:

$$CSV(x) = \frac{B(x,z) - D(z)R(x,z)}{A(x)} \tag{5.18}$$

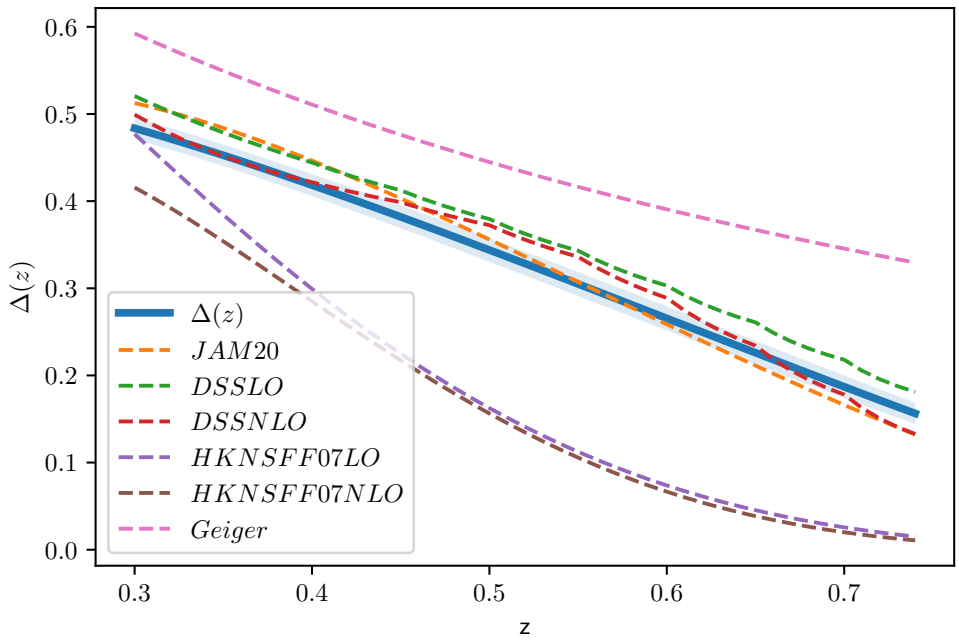


Figure 5.9: Result of unfavored over favored fragmentation function $\Delta(z)$ in comparison with some existing fragmentation functions

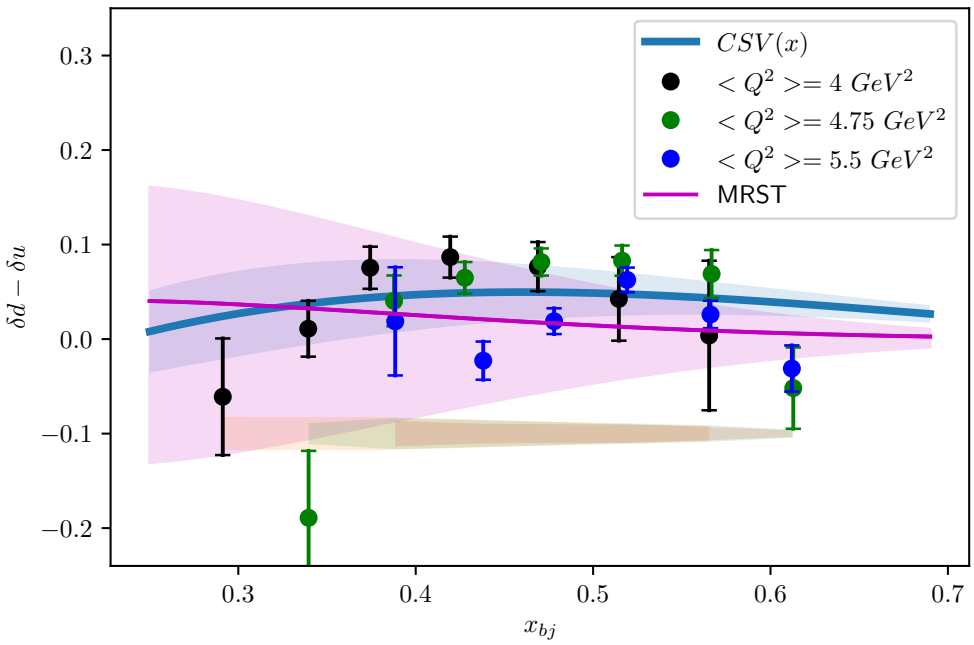


Figure 5.10: The CSV results from fitting. The blue line and 1σ blue error band is from the fitting process. The points are the CSV results calculated from the fragmentation function ratio result for each Q^2 setting

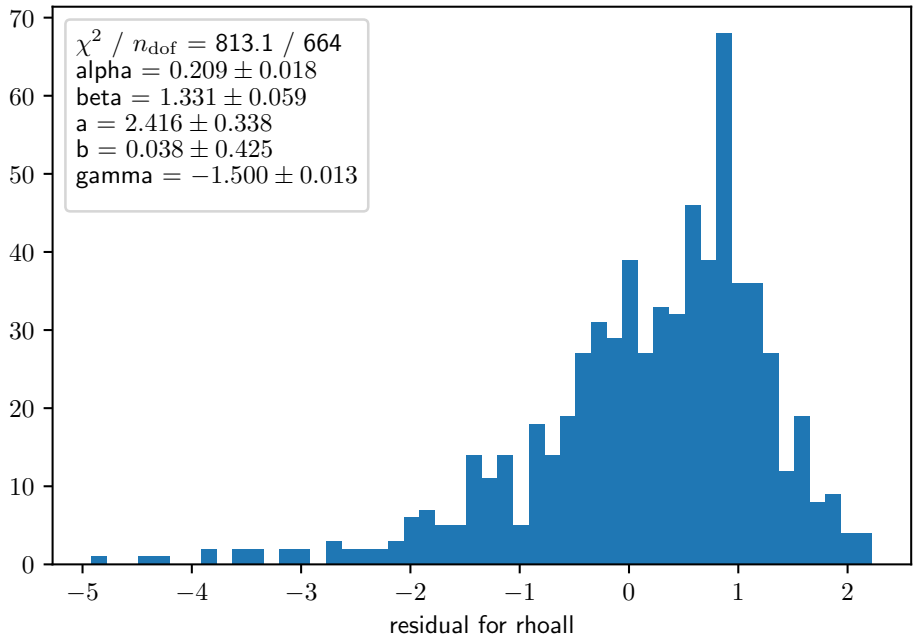


Figure 5.11: Residual for fitting, with the fitting parameters

Different PDFs will change both $A(x)$ and $B(x, z)$ term. The CSV calculated from JAM20PDFs, nnPDFs, and cteq18 PDFs are averaged as shown in Fig.5.12

The fragmentation functions are not well defined. For some groups, the charge symmetry in fragmentation functions is assumed (Eq.2.29), like the JAM group. But for the other group, it's different. For the DSS group results, the D_d^+ is assumed to be equal to D_u^-, while the other case doesn't. The equation 5.7 assumed charge symmetry in fragmentation function. This case is considered by another equation in section 5.2.3

5.2.3 Formalism Including CSV in the Fragmentation Functions

The Equation 5.7 ignored the quadratic term of CSV terms (δd^2, δu^2) during the process. A new equation is derived to get CSV from yield ratio R_Y.

From the yield of SIDIS hadron production Equation 2.26 and Equation 2.27, the

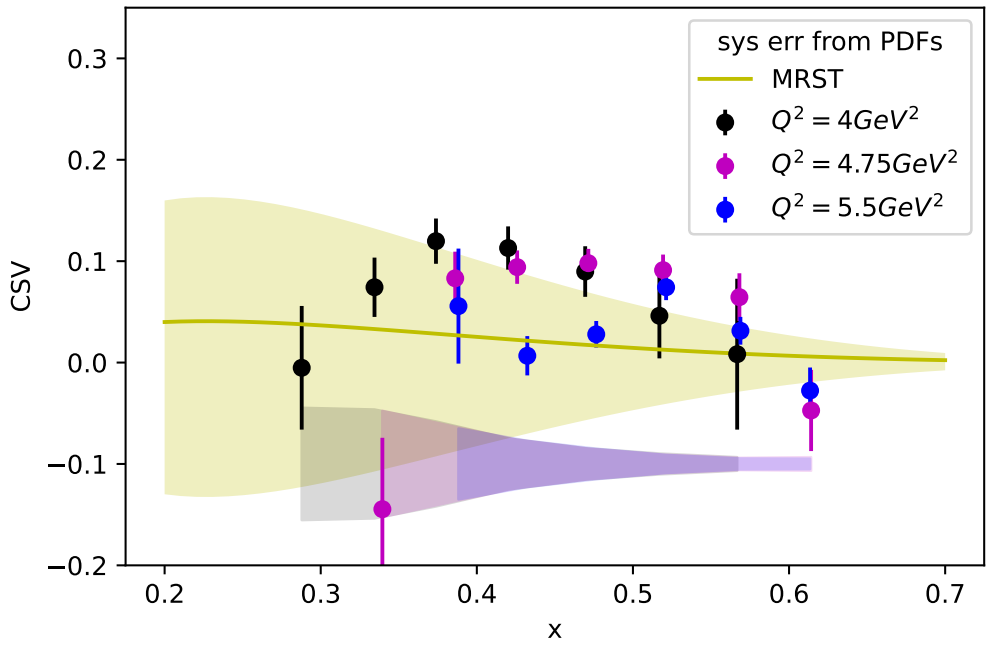
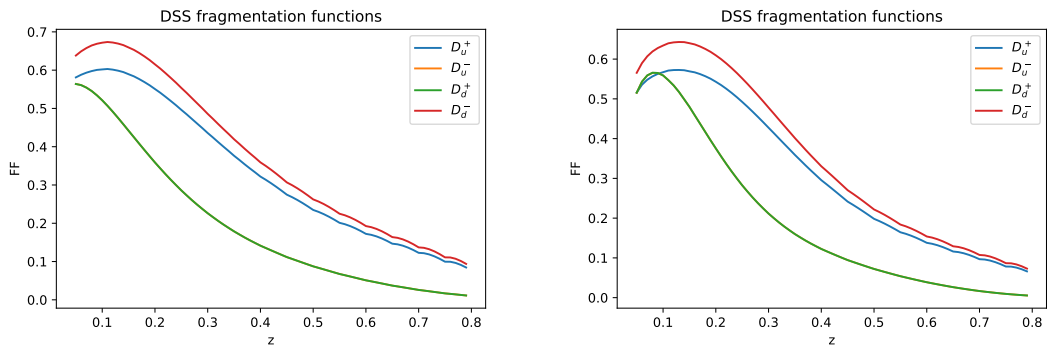


Figure 5.12: The calculated results of CSV from different PDF functions



(a) Leading-Order

(b) Next-to-Leading-Order

Figure 5.13: FFs from the DSS group

yield ratio, with impulse assumption, can be written as:

$$R_Y = \frac{N^{D-}}{N^{D+}} = \frac{N^{p-} + N^{n-}}{N^{p+} + N^{n+}} \tag{5.19}$$

With the assumption on CSV term $\delta d = -\delta u$, it leads to

$$\begin{aligned} \delta d &= -\delta u \\ u^n &= d^p + \delta u \\ d^n &= u^p - \delta u \end{aligned} \tag{5.20}$$

The sea quark charge symmetry is assumed, but the sea quark terms here are ignored

$$\begin{aligned} \overline{u}^p &= \overline{d}^n \\ \overline{d}^p &= \overline{u}^n \end{aligned} \tag{5.21}$$

With the charge conjugation invariance as in Eq.2.28 and the charge symmetry assumption as in Eq.2.29, The strange quark (s^p, s^n) terms and sea quark terms $(\overline{u}, \overline{d})$ are ignored. Define $\delta U = \frac{\delta u}{u_v^p + d_v^p}$, and the fragmentation function ratio $\Delta(z) = \frac{D_u^-}{D_u^+}$ the yield ratio R_Y becomes:

$$\begin{aligned} R_Y &= \frac{4(u_v^p + d_v^p)D_u^- + 4\delta u D_u^- + (d_v^p + u_v^p)D_u^+ - \delta u D_u^+}{4(u_v^p + d_v^p)D_u^+ + 4\delta u D_u^+ - \delta u D_u^- + (u_v^p + d_v^p)D_u^-} \\ &= \frac{4\Delta(z) + 4\delta U \Delta(z) + 1 - \delta U}{4 + 4\delta U - \delta U \Delta(z) + \Delta(z)} \\ &= \frac{(4\Delta(z) + 1) + \delta U(4\Delta(z) - 1)}{(4 + \Delta(z)) + \delta U(4 - \Delta(z))} \end{aligned} \tag{5.22}$$

solving for δU from yield ratio:

$$\delta U = \frac{(4\Delta(z) + 1) - Y(4 + \Delta(z))}{Y(4 - \Delta(z)) - (4\Delta - 1)} \tag{5.23}$$

The error for δU is:

$$\sigma_{\delta U} = \left(\frac{4\Delta + 1}{4 - \Delta} - \frac{(4\Delta - 1)(4 + \Delta)}{(4 - \Delta)^2}\right)\frac{\sigma_Y}{(Y - \frac{4\Delta - 1}{4 - \Delta})^2}$$

Here use the same equation R_Y as in Equation 5.19 but not ignoring the strange and $\overline{u}, \overline{d}$ terms. Use $u \equiv u^p, d \equiv d^p$, the numerator is:

$$\begin{aligned} &4(u + d)D_u^- - 4\delta d D_u^- + 4(\overline{u} + \overline{d})D_u^+ + (u + d)D_u^+ - \delta u D_u^+ \\ &+ (\overline{u} + \overline{d})D_u^- + 2sD_s^- + 2\overline{s}D_s^+ \end{aligned} \tag{5.24}$$

And the denominator is :

$$\begin{aligned} &4(u+d)D_u^+ - 4\delta d D_u^+ + 4(\overline{u}+\overline{d})D_u^- + (u+d)D_u^- - \delta u D_u^- \\ &+ (\overline{d}+\overline{u})D_u^+ + 2sD_s^+ + 2\overline{s}D_s^- \end{aligned} \tag{5.25}$$

The yield ratio is:

$$R_Y = \frac{(4u+4d+\overline{u}+\overline{d})D_u^- + (4\overline{u}+4\overline{d}+u+d)D_u^+ - 4\delta d D_u^- - \delta u D_u^+ + 2sD_s^- + 2\overline{s}D_s^+}{(4u+4d+\overline{u}+\overline{d})D_u^+ + (4\overline{u}+4\overline{d}+u+d)D_u^- - 4\delta d D_u^+ - \delta u D_u^- + 2sD_s^+ + 2\overline{s}D_s^-} \tag{5.26}$$

Assume $T = \frac{\overline{u}+\overline{d}}{u+d}$, $S = \frac{2(s+\overline{s})}{u+d}$ and $D_s^+ = D_s^- = D_u^-$, Y can be written:

$$R_Y = \frac{(4+T)\Delta(z) + (4T+1) - 4\delta d\Delta(z) - \delta u + S\Delta(z)}{(4T+1)\Delta(z) + (4+T) - 4\delta d - \delta u\Delta(z) + S\Delta(z)} \tag{5.27}$$

As defined before, $\delta U = (u^p - d^n)/(u+d)$, $\delta D = (d^p - u^n)/(u+d)$, assume

$$\delta D = -a_{CSV}\delta U \tag{5.28}$$

the yield ratio R_Y is

$$\begin{aligned} R_Y &= \frac{(4+T)\Delta(z) + (4T+1) + (4a_{CSV}\Delta(z) - 1)\delta U + S\Delta(z)}{(4T+1)\Delta(z) + (4+T) + (4a_{CSV} - \Delta(z))\delta U + S\Delta(z)} \\ &= \frac{(4+T+S)\Delta(z) + (4T+1) + (4a\Delta(z) - 1)\delta U}{(4T+1+S)\Delta(z) + (4+T) + (4a - \Delta(z))\delta U} \end{aligned} \tag{5.29}$$

From the yield ratio, δU is:

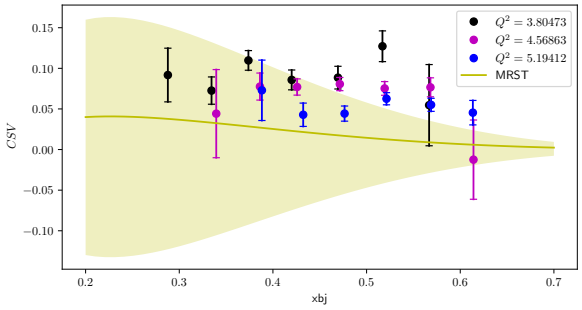
$$\delta U = \frac{(4+T+S)\Delta + (4T+1) - R_Y(4T\Delta + \Delta + S\Delta) - R_Y(4+T)}{(4a_{CSV} - \Delta)R_Y - (4a_{CSV}\Delta - 1)} \tag{5.30}$$

And CSV is $\delta d - \delta u$, which gives $CSV = -(a_{CSV}+1)(u+d)\delta U$

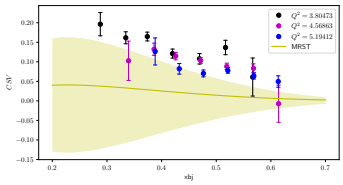
The CSV calculated by this equation is shown in Fig.5.14. The PDFs and FFs used here are from the JAM group. The radiative background ρ is subtracted from the data by two coefficients, 1 and 1.5.

If the charge symmetry is not assumed in fragmentation function (as in Equation 2.29), $D_d^- \neq D_u^+$, as in Fig. 5.13. The numerator from Equation 5.19 becomes

$$\begin{aligned} &4(u+d)D_u^- - 4\delta d D_u^- + 4(\overline{u}+\overline{d})D_u^+ + (u+d)D_d^- - \delta u D_d^- \\ &+ (\overline{u}+\overline{d})D_u^- + 2sD_s^- + 2\overline{s}D_s^+ \end{aligned} \tag{5.31}$$



(a) $\gamma = 1.5$



(b) $\gamma = 1$

Figure 5.14: CSV calculated from the yield ratio equation using JAM20 PDFs and FFs. The radiative ρ background is subtracted from data with coefficient γ

And the denominator becomes:

$$\begin{aligned} &4(u+d)D_u^+ - 4\delta d D_u^+ + 4(\overline{u}+\overline{d})D_u^- + (u+d)D_u^- - \delta u D_u^- \\ &+ (\overline{d}+\overline{u})D_d^- + 2sD_s^+ + 2\overline{s}D_s^- \end{aligned} \tag{5.32}$$

Assuming $\Gamma(z) = \frac{D_d^-(z)}{D_u^+(z)}$, the yield ratio R_Y becomes:

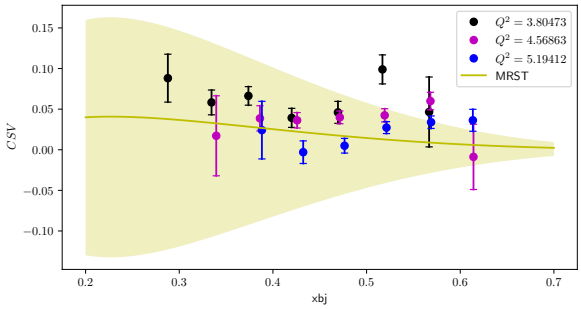
$$R_Y = \frac{(4+T+S)\Delta(z) + (4T+\Gamma(z)) + (4a_{CSV}\Delta(z) - \Gamma(z))\delta U}{(4T+1+S)\Delta(z) + (4+T\Gamma(z)) + (4a_{CSV} - \Delta(z))\delta U} \tag{5.33}$$

Therefore, the ΔU is:

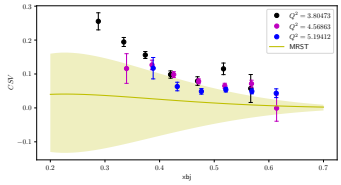
$$\delta U = \frac{(4+T+S)\Delta + (4T+\Gamma) - R_Y(4T\Delta + \Delta + S\Delta) - R_Y(4+T\Gamma)}{(4a_{CSV} - \Delta)R_Y - (4a_{CSV}\Delta - \Gamma)} \tag{5.34}$$

The CSV results calculated from the DSS fragmentation function is shown in Fig. 5.15

In summary, several different PDFs are tested, the PDFs from the JAM20 group, from the NNPDF group, and from the CTEQ group are used. Several different FFs are used, the fitting results from the CSV data set, JAM20 FFs, and DSS FFs. A new equation to get CSV directly from the yield ratio is used. Different radiative background subtractions (0,1,1.5) are included. The CSV results are averaged among all the results from all the different combinations shown in Fig. 5.16



(a) $\gamma = 1$



(b) $\gamma = 0$

Figure 5.15: CSV calculated from the yield ratio equation using JAM20 PDFs and DSS leading order FFs. The radiative ρ background is subtracted from data with coefficient γ

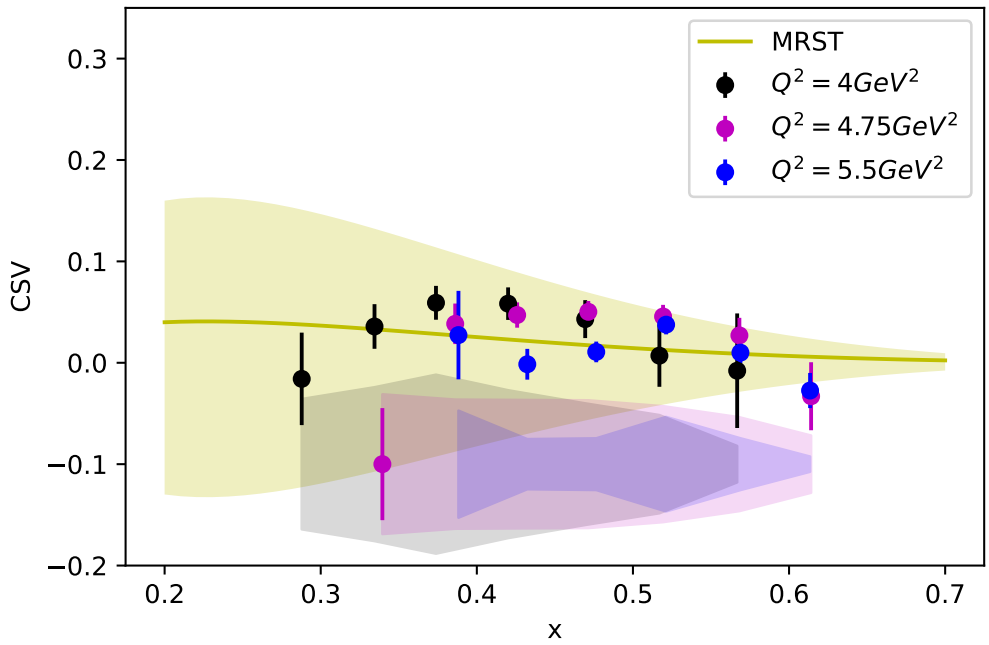


Figure 5.16: The CSV results from all different cases

5.2.4 Discussion of results

First, the Charge Symmetry Violation was extracted along with the fragmentation ratio $\Delta(z)$ from our data. The fragmentation ratio is comparable with the results of global analyses and fits, such as JAM20 and DSS, as shown in Fig. 5.9. The CSV results shown in Fig.5.10 are within the MRST prediction band with less uncertainty. We then compute CSV(x) from R^D_{meas} and use existing fits and models of the fragmentation functions as input. For some groups, like JAM20, charge symmetry is assumed in the fragmentation function parameterization. CSV(x) results using the JAM20 PDFs and FFs are shown in Fig. 5.14. These results are similar to the ones we obtain using the simultaneous extraction of $\Delta(x)$ from our data.

The DSS group did not assume CS in their model of the fragmentation functions. Their fit to world data found a better result allowing charge symmetry breaking in fragmentation functions, $D^-_d \neq D^+_u$. While the equation 2.62 assumes charge symmetry in the fragmentation function, we can instead use the equation 5.33 along with the DSS model to examine the impact of allowing CSV in the fragmentation functions. The results are shown in Fig.5.15. This CSV result is consistent with nominal ρ background subtraction, and also consistent with the MRST fit. This tends to lend weight to the notion that CSV in the fragmentation functions is an important ingredient in describing the data.

5.3 Conclusion

The precise measurement of the SIDIS charged pion production ratio was conducted on a deuterium target. The Charge Symmetry Violation parton distribution in the valence region and fragmentation ratio were extracted from the data covering a kinematic range of $0.3 < x < 0.6$, $0.3 < z < 0.7$, and $3GeV^2 < Q^2 < 6GeV^2$. The different

fragmentation function input suggests that a CSV contribution from the FF should be considered in a global analysis. The CSV results in parton distribution are consistent with MRST limits. Including CSV in the fragmentation improves the agreement with the MRST prediction.

These results should be included in an extensive global analysis which includes CSV in the parton distributions and fragmentation functions. Some H_2 data have been taken. They are used just for the factorization check in this analysis. Analysis on H_2 could help determine CSV in quark distribution or fragmentation functions.

REFERENCES

[1] Experimental physics and industrial control system.

[2] 12 gev upgrade proposal reaches critical stage, Feb 2021.

[3] Stephen L Adler. Sum rules giving tests of local current commutation relations in high-energy neutrino reactions. *Physical Review*, 143(4):1144, 1966.

[4] J. Agarwala and etc. Contribution of exclusive diffractive processes to the measured azimuthal asymmetries in sidis. *Nuclear Physics B*, 956:115039, Jul 2020.

[5] P. Amaudruz and etc. Gottfried sum from the ratio f_2^n/f_2^p. *Phys. Rev. Lett.*, 66:2712–2715, May 1991.

[6] P. Amaudruz and etc. Measurements of rd-rp and rca-rc in deep inelastic muon scattering. *Physics Letters B*, 294(1):120–126, 1992.

[7] A. Baldit and etc. Study of the isospin symmetry breaking in the light quark sea of the nucleon from the drell-yan process. *Physics Letters B*, 332(1):244 – 250, 1994.

[8] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10), oct 2017.

[9] C. J. Bebek, C. N. Brown, S. D. Holmes, R. V. Kline, F. M. Pipkin, S. Raither, L. K. Sisterson, A. Browman, K. M. Hanson, D. Larson, and A. Silverman. Electroproduction of single pions at low ϵ and a measurement of the pion form factor up to $Q^2 = 10$ gev^2. *Phys. Rev. D*, 17:1693–1705, Apr 1978.

[10] J. Binnewies, B. A. Kniehl, and G. Kramer. Pion and kaon production in e^+e^- and ep collisions at next-to-leading order. *Phys. Rev. D*, 52:4947–4960, Nov 1995.

[11] J. D. Bjorken. Asymptotic sum rules at infinite momentum. *Phys. Rev.*, 179:1547–1553, Mar 1969.

[12] J. D. Bjorken and E. A. Paschos. Inelastic electron-proton and γ-proton scattering and the structure of the nucleon. *Phys. Rev.*, 185:1975–1982, Sep 1969.

[13] C. Boros, J. T. Londergan, and A. W. Thomas. Evidence for substantial charge symmetry violation in parton distributions. *Physical Review Letters*, 81(19):4075–4078, nov 1998.

[14] P. Brauel, T. Canzler, D. Cords, R. Felst, G. Grindhammer, M. Helm, W. D. Kollmann, H. Krehbiel, and M. Schädlich. Electroproduction of +n, -p andk+,k+2211;0 final states above the resonance region. *Zeitschrift für Physik C Particles and Fields*, 3(2):101–123, 1979.

[15] Jefferson Lab Hall C. Root analyzer.

[16] C.Y. Cheung, E.M. Henley, and G.A. Miller. Charge symmetry breaking in the neutron-proton system. *Nuclear Physics A*, 348(4):365–387, 1980.

[17] D de Florian, R Sassot, and M Stratmann. Fragmentation functions for pions, kaons, protons and charged hadrons. *Journal of Physics: Conference Series*, 110(2):022045, may 2008.

[18] Daniel de Florian, R. Sassot, Manuel Epele, Roger J. Hernández-Pinto, and Marco Stratmann. Parton-to-pion fragmentation reloaded. *Physical Review D*, 91(1), Jan 2015.

[19] Hans Dembinski and Piti Ongmongkolkul et al. scikit-hep/iminuit, Dec 2020.

[20] D. Drechsel, S. S. Kamalov, and L. Tiator. Unitary isobar model –maid2007. *The European Physical Journal A*, 34(1):69, 2007.

[21] S.D Drell and J.D Walecka. Electrodynamic processes with nuclear targets. *Annals of Physics*, 28(1):18–33, 1964.

[22] S.D. Ellis and W.J. Stirling. Constraints on isospin breaking in the light quark sea from the drell-yan process. *Physics Letters B*, 256(2):258–264, 1991.

[23] R. Ent, B. W. Filippone, N. C. R. Makins, R. G. Milner, T. G. O'Neill, and D. A. Wasson. Radiative corrections for $(e, e^{'}p)$ reactions at gev energies. *Phys. Rev. C*, 64:054610, Oct 2001.

[24] R.D. Field and R.P. Feynman. A parametrization of the properties of quark jets. *Nuclear Physics B*, 136(1):1–76, 1978.

[25] Paolo Gambino. Precision electroweak measurements: a theorist point of view. *Nuclear Physics B - Proceedings Supplements*, 117:211 – 215, 2003.

[26] M. Glück, E. Reya, and A. Vogt. Dynamical parton distributions revisited. *The European Physical Journal C - Particles and Fields*, 5(3):461–470, 1998.

[27] Kurt Gottfried. Sum rule for high-energy electron-proton scattering. *Phys. Rev. Lett.*, 18:1174–1177, Jun 1967.

[28] David J. Gross and C.H.Llewellyn Smith. High-energy neutrino-nucleon scattering, current algebra and partons. *Nuclear Physics B*, 14(2):337–347, 1969.

[29] E. A. Hawker and etc. Measurement of the light antiquark flavor asymmetry in the nucleon sea. *Phys. Rev. Lett.*, 80:3715–3718, Apr 1998.

[30] EM Henley and GA Miller. Mesons in nuclei. *North Holland, Amsterdam*, 1:405, 1979.

[31] EM Henley and GA Miller. Excess of d over u in the proton sea quark distribution. *Physics Letters B*, 251(3):453–454, 1990.

[32] M. Hirai, S. Kumano, T.-H. Nagai, and K. Sudoh. Determination of fragmentation functions and their uncertainties. *Physical Review D*, 75(9), May 2007.

[33] DAQ group JLab. Introduction|data acquisition support.

[34] Mark Jones. Mark jones' email.

[35] Mark Jones. Track selection in hcana - hallcweb.jlab.org, Jun 2020.

[36] Mark Jones. Updates to hcana for hodoscope, Apr 2020.

[37] J.T.Londergan and A.W.Thomas. The validity of charge symmetry for parton distributions. *Prog.Part.Nucl.Phys.*, 1998.

[38] R. Kazimi et al. Four Beam Generation for Simultaneous Four-Hall Operation at CEBAF. In *Proc. of International Particle Accelerator Conference (IPAC'16), Busan, Korea, May 8-13, 2016*, number 7 in International Particle Accelerator Conference, pages 4240–4242, Geneva, Switzerland, June 2016. JACoW. doi:10.18429/JACoW-IPAC2016-THPOY060.

[39] Julius Kuti and Victor F. Weisskopf. Inelastic lepton-nucleon scattering and lepton pair production in the relativistic quark-parton model. *Phys. Rev. D*, 4:3418–3439, Dec 1971.

[40] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J. F. Owens, J. Pumplin, and W. K. Tung. Global qcd analysis of parton structure of the nucleon: Cteq5 parton distributions. *The European Physical Journal C - Particles and Fields*, 12(3):375–392, 2000.

[41] J. T. Londergan. Possible explanations for the NuTev Weinberg angle measurement. *Nucl. Phys. B Proc. Suppl.*, 141:68–73, 2005.

[42] J. T. Londergan, Alex Pang, and A. W. Thomas. Probing charge-symmetry-violating quark distributions in semi-inclusive leptoproduction of hadrons. *Phys. Rev. D*, 54:3154–3161, Sep 1996.

[43] J. T. Londergan, J. C. Peng, and A. W. Thomas. Charge symmetry at the partonic level. *Reviews of Modern Physics*, 82(3):2009–2052, jul 2010.

[44] Dave Mack. Bcmfinalcalibrationssummer2019, Oct 2019.

[45] Dave Mack. Final systematic error analysis of hms scaler-based fall '18 boiling slopes, Oct 20AD.

[46] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Uncertainties of predictions from parton distributions II: theoretical errors. *The European Physical Journal C*, 35(3):325–348, jun 2004.

[47] A. D. Martin and M. G. Ryskin. The photon pdf of the proton. *The European Physical Journal C*, 74(9):3040, 2014.

[48] R D McKeown. The jefferson lab 12 GeV upgrade. *Journal of Physics: Conference Series*, 312(3):032014, sep 2011.

[49] Dave Meekins. Hall c target configuration, Feb 2017.

[50] N. Meister and D. R. Yennie. Radiative corrections to high-energy scattering processes. *Phys. Rev.*, 130:1210–1229, May 1963.

[51] G. A. Miller, B. M. K. Nefkens, and I. Slaus. Charge symmetry, quarks and mesons. *Phys. Rept.*, 194:1–116, 1990.

[52] G.A. Miller, B.M.K. Nefkens, and I. Šlaus. Charge symmetry, quarks and mesons. *Physics Reports*, 194:1–116, Oct 1990.

[53] Gerald A. Miller. Charge symmetry breaking. *Nucl. Phys.*, A518:345–357, 1990.

[54] Gerald A. Miller, Allena K. Opper, and Edward J. Stephenson. Charge symmetry breaking and qcd. *Annual Review of Nuclear and Particle Science*, 56(1):253–292, 2006.

[55] H. Mkrtchyan, R. Carlini, V. Tadevosyan, J. Arrington, A. Asaturyan, M.E. Christy, D. Dutta, R. Ent, H.C. Fenker, D. Gaskell, and et al. The lead-glass electromagnetic calorimeters for the magnetic spectrometers in hall c at jefferson lab. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 719:85–100, Aug 2013.

[56] L. W. MO and Y. S. TSAI. Radiative corrections to elastic and inelastic ep and up scattering. *Rev. Mod. Phys.*, 41:205–235, Jan 1969.

[57] E. Moffat, W. Melnitchouk, T.C. Rogers, and N. Sato. Simultaneous monte carlo analysis of parton densities and fragmentation functions. *Physical Review D*, 104(1), Jul 2021.

[58] T. Navasardyan and etc. Onset of quark-hadron duality in pion electroproduction. *Phys. Rev. Lett.*, 98:022001, Jan 2007.

[59] B. M. K. Nefkens, W. J. Briscoe, A. D. Eichon, D. H. Fitzgerald, J. A. Holt, A. A. Mokhtari, J. A. Wightman, M. E. Sadler, R. L. Boudrie, and C. L. Morris. Test of charge symmetry in π^+ and π^- elastic scattering on tritium and ^{3}He. *Phys. Rev. Lett.*, 52:735–738, Feb 1984.

[60] J. A. Niskanen. Isospin breaking in the reaction at threshold. *Few-Body Systems*, 26(2):241–249, 1999.

[61] NIST. Nist chemistry webbook, srd 69.

[62] A K Opper, E Korkmaz, D A Hutcheon, R Abegg, C A Davis, R W Finlay, P W Green, L G Greeniaus, D V Jordan, J A Niskanen, G V O'Rielly, T A Porcelli, S D Reitzner, P L Walden, and S Yen. Charge symmetry breaking in np–>dpi(0). *Phys Rev Lett*, 91(21):212302, Nov 2003.

[63] etal P. Amaudruz. Gottfried sum from the ratio f_2^n/f_2^p. *Phys. Rev. Lett.*, 66:2712–2715, May 1991.

[64] E. A. Paschos and L. Wolfenstein. Tests for neutral currents in neutrino reactions. *Phys. Rev. D*, 7:91–95, Jan 1973.

[65] E. Pedroni, K. Gabathuler, J.J. Domingo, W. Hirt, P. Schwaller, J. Arvieux, C.H.Q. Ingram, P. Gretillat, J. Piffaretti, N.W. Tanner, and C. Wilkin. A study of charge independence and symmetry from + and total cross sections on hydrogen and deuterium near the 3, 3 resonance. *Nuclear Physics A*, 300(2):321–347, 1978.

[66] P.Geiger. *Ph.D. dissertation.* PhD thesis, 1998.

[67] Eric Pooser. Live time calculations, May 2018.

[68] Jonathan Pumplin, Daniel Robert Stump, Joey Huston, Hung-Liang Lai, Pavel Nadolsky, and Wu-Ki Tung. New generation of parton distributions with uncertainties from global qcd analysis. *Journal of High Energy Physics*, 2002(07):012–012, Jul 2002.

[69] E.N. RODIONOV, A.W. THOMAS, and J.T. LONDERGAN. Charge asymmetry of parton distributions. *Modern Physics Letters A*, 09(19):1799–1806, 1994.

[70] Eric Sather. Isospin-violating quark distributions in the nucleon. *Physics Letters B*, 274(3):433–438, 1992.

[71] Brad Sawatzky. Brad sawatzky's log.

[72] Brad Sawatzky. Hall c spectrometer optics - jlab, Oct 2018.

[73] Julian Schwinger. Quantum electrodynamics. iii. the electromagnetic properties of the electron—radiative corrections to scattering. *Phys. Rev.*, 76:790–817, Sep 1949.

[74] W. G. Seligman and etc. Improved determination of α_s from neutrino-nucleon scattering. *Phys. Rev. Lett.*, 79:1213–1216, Aug 1997.

[75] A. I. Signal and A. W. Thomas. Calculation of quark distribution functions using bag-model wave functions. *Phys. Rev. D*, 40:2832–2843, Nov 1989.

[76] SIMC. Simc overiew.

[77] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, May 2006.

[78] E. J. Stephenson, A. D. Bacher, C. E. Allgower, A. Gårdestig, C. M. Lavelle, G. A. Miller, H. Nann, J. Olmsted, P. V. Pancella, M. A. Pickar, J. Rapaport, T. Rinckel, A. Smith, H. M. Spinka, and U. van Kolck. Observation of the charge symmetry breaking $d + d \rightarrow^4 \mathrm{He} + \pi^0$ reaction near threshold. *Phys. Rev. Lett.*, 91:142302, Oct 2003.

[79] Holly Szumila-Vance. Commissioning the hms optics in the 2017-18 run period.

[80] Holly Szumila-Vance. Notes on the shms optics in the 2017-18 run period.

[81] V Tadevosyan, H Mkrtchyan, A Asaturyan, A Mkrtchyan, S Zhamkochyan, and For the Hall C Collaboration. The design and performance of the electromagnetic calorimeters in hall c at jefferson lab. *Journal of Physics: Conference Series*, 404(1):012052, dec 2012.

[82] A.W. Thomas, C. Boros, and J.T. Londergan. Evidence for charge symmetry violation in parton distributions. *Nuclear Physics A*, 654(1, Supplement 1):596c–601c, 1999.

[83] Yung-Su Tsai. Radiative corrections to electron-proton scattering. *Phys. Rev.*, 122:1898–1907, Jun 1961.

[84] Brigitte Vachon. Trigger and data acquisition (i), 2010.

[85] U. van Kolck, J.A. Niskanen, and G.A. Miller. Charge symmetry violation in pn→d0 and chiral effective field theory. *Physics Letters B*, 493(1-2):65–72, nov 2000.

[86] X.G. Wang, A.W. Thomas, and R.D. Young. Electromagnetic contribution to charge symmetry violation in parton distributions. *Physics Letters B*, 753:595–599, feb 2016.

[87] L. W. Whitlow, E. M. Riordan, S. Dasu, Stephen Rock, and A. Bodek. Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections. *Phys. Lett. B*, 282:475–482, 1992.

[88] S. A. Wood. 2019 version: Jefferson lab hall c standard equipment manual. *Journal of Physics: Conference Series*, Oct 2019.

[89] Keping Xie, T. J. Hobbs, Tie-Jiun Hou, Carl Schmidt, Mengshi Yan, and C. P. Yuan. The photon pdf within the ct18 global analysis. 2021.

[90] C. Yan, N. Sinkine, and R. Wojcik. Linear beam raster for cryogenic targets. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 539(1):1–15, 2005.

[91] G. P. Zeller and et al. Precise determination of electroweak parameters in neutrino-nucleon scattering. *Phys. Rev. Lett.*, 88:091802, Feb 2002.

[92] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.