北京大学

本科学位论文

题目 HERMES实验非极化氢靶上实光子遍举电产生过程中方位角不对称性的测量

姓　　名　卢显国

学　　号　00304124

院　　系　物理学院

指导教师　冒亚军 教授

二零零七年六月

Measurement of Azimuthal Asymmetries in Exclusive Electroproduction of Real Photons on an Unpolarized Hydrogen Target at HERMES

Xianguo Lu

School of Physics, Peking University

June, 2007

Submitted in total fulfilment of the requirements for the degree of B.S. in Physics

Abstract

This thesis presents the data analysis results for the azimuthal asymmetries in reaction [1] $ep_{\text{unp.}} \to e'p'\gamma$. Related to the generalized parton distributions (GPDs), these asymmetries can provide useful information on the nucleon structure. This analysis is based on the 2000 and 2005 measurements by the HERMES experiment, which used its hydrogen target and the positron/electron beams of the HERA storage ring. Asymmetries in kinematic bins are extracted, exhibiting a qualitative agreement with theoretical calculations.

Keywords: Exclusive Electroproduction of Real Photons, Azimuthal Asymmetries, Unpolarized Hydrogen Target, HERMES

摘要

本文分析了$ep_{\text{unp.}} \to e'p'\gamma$过程中方位角的不对称性[2]。该不对称性与一般化的部分子分布函数（Generalized Parton Distributions, GPDs）有关，因而能提供核子结构方面的有用信息。HERMES实验利用自身的氢靶和HERA存储环的正/负电子束流反应。本文基于该实验2000和2005年的数据，提取了不同动力学区间的方位角不对称度。分析结果与理论计算定性符合。

关键词： 实光子遍举电产生, 方位角不对称性, 非极化氢靶, HERMES实验

[1] unp. is short for *unpolarized*.

[2] unp. 是“非极化的（*unpolarized*）” 的缩写。

Contents

Chapter 1

Introduction

The elementary constituents of hadrons are known to be quarks and gluons, which interact via the strong interaction described by Quantum Chromodynamics (QCD). The QCD Lagrangian contains all phenomena of hadronic physics. However, in practice, it is difficult to derive strong interaction phenomena starting from first principles for lack of detailed knowledge of the hadron wave function in quark and gluon degrees of freedom. Alternatively, we rely on phenomenological functions such as form factors and parton distributions to obtain information about hadron structure. The new functions – generalized parton distributions (GPDs) – are discovered to be a new window for the exploration of nonperturbative QCD and three dimensional structure of hadrons [Ji97].

An essential tool to prob the hadron structure is hard reactions, such as inclusive [1] deep inelastic scattering (DIS), semi-inclusive DIS, Drell-Yan processes and exclusive reactions. The common important feature of hard reactions is the possibility to factorize the dynamics into (perturbatively calculable) short-distance part and (nonperturbative) long-distance part. For the inclusive process, information about the internal structure of the hadron is encoded into the so-called parton distributions, which can be interpreted as the densities of the longitudinal momentum carried by the partons. However, such inclusive properties are insufficient to provide a detailed description of the hadron wave function. For example, in parton distributions all transverse momenta are integrated out with some important features of the strong interaction thus lost. Contrary to inclusive DIS, exclusive reactions – so defined that the kinematical parameters of all initial- and final-state particles are specified, i.e. all outgoing particles are detected – undergo a higher sensitivity to the long-distance effects as a consequence of the coherent scattering off the hadron constituents. Such processes give a more direct access to the wave functions, shedding more light on the quark confinement and thus on the hadron structure, but on the other hand requiring more sophisticated theoretical analyses [Bel05].

GPDs are a generalization of ordinary parton distributions and elastic form factors. Besides the longitudinal parton momentum, GPDs also describe the spatial distribution of quarks and gluons in the plane transverse to the momentum of a fast moving hadron,

[1] With respect to the final states, requiring only the scattered lepton being detected.

hence forming a fully three-dimensional description of hadron structure [Die03]. Related to the nonperturbative structure of hadrons, GPDs enter the parameterization of exclusive processes, among which deeply virtual Compton scattering (DVCS) and hard exclusive meson production are the most favorable class in theoretical aspect [Ji98].

Experimental access to GPDs can be fulfilled by cross section or asymmetry measurements for exclusive reactions. However, since the exclusive rates drop much faster with momentum transfer than their inclusive counterparts, the experimental measurement for the exclusive process involving GPDs is a challenge, requiring high luminosity to compensate for small cross sections and detectors capable of ensuring the exclusivity of the final state. Pioneering experimental results have been obtained by HERMES [Air01], H1 [Adl01], ZEUS [Sau99], and CLAS [Ste01].

This thesis describes the first analysis using both the positron and electron data to simultaneously extract the beam-spin and beam-charge azimuthal asymmetries in HERMES acceptance. It is organized as follows. Chapter 2 gives the theoretical descriptions for DVCS. In Chapter 3, the HERMES experiment is introduced. Details of the data analysis are discussed in Chapter 4 and the results are presented in Chapter 5. Finally is the summary and outlook in Chapter 6.

Chapter 2

Deeply Virtual Compton Scattering

Deeply virtual Compton scattering (DVCS) is the hard photoproduction of a real photon, i.e. $\gamma^* N \to \gamma N'$, where $N(N')$ is a nucleon. The corresponding Feynman diagrams (Fig. 2.1) in leading twist [1] approximation have the handbag form [Die97]. It represents that a highly virtual photon scatters off a quark in a nucleon target, and then this quark emits a real photon and returns to form a recoil nucleon. The lower blobs correspond to the long-distant part, where the nonperturbative objects are parameterized by GPDs.

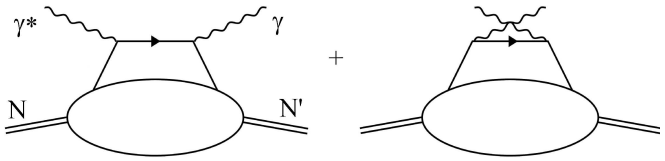

Figure 2.1: Handbag diagrams for DVCS.

DVCS can be accessed in the hard exclusive leptoproduction of a real photon off a nucleon,

$$l(k) + N(p) \to l'(k') + N'(p') + \gamma(q'), \tag{2.1}$$

where k, k', p, p' and q' denote the four-momenta of the initial- and final-state particles. In this process, besides DVCS, the real photon can also be emitted by the incoming or outgoing lepton via the Bethe-Heitler (BH) process. Leading Feynman diagrams for DVCS and BH process are shown in Fig. 2.2.

2.1 Kinematics

To describe the exclusive process (2.1), special variables like missing mass come into use in addition to the common ones for DIS, photon virtuality for example. In this section, kinematic variables needed in this analysis are defined (cf. Fig. 2.3), assuming a fixed target in the laboratory frame:

[1] Power-suppressed corrections to the probability for coherent scattering on a n-parton configuration go under the name of higher twists [Bel05].

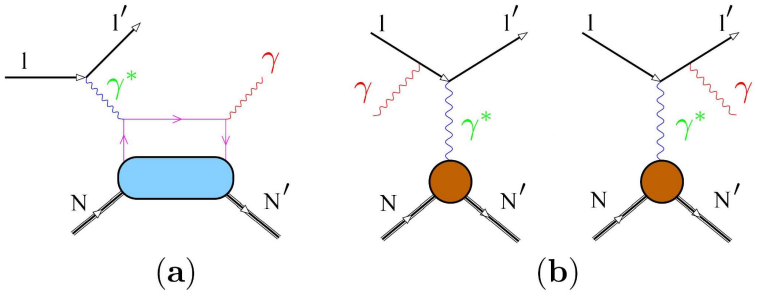

Figure 2.2: Leading Feynman diagrams for DVCS (a) and BH process (b).

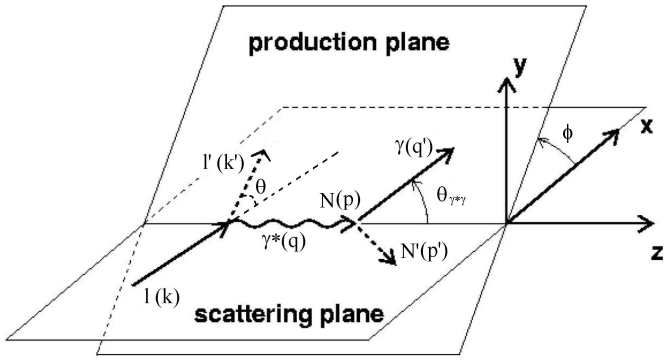

Figure 2.3: Kinematics of DVCS in the laboratory frame.

- The energy of the virtual photon [2],

$$\nu \equiv \frac{pq}{M_N} \overset{\text{lab}}{=} E - E', \tag{2.4}$$

where M_N is the nucleon mass and E (E') is the energy of the incoming (outgoing) lepton in the laboratory frame.

- The lepton energy fraction,

$$y \equiv \frac{pq}{pk} \overset{\text{lab}}{=} \frac{\nu}{E}. \tag{2.5}$$

- Photon virtuality,

$$Q^2 \equiv -q^2 \overset{\text{lab}}{\approx} 4EE' \sin^2 \frac{\theta}{2}. \tag{2.6}$$

[2] The expansion of the inner productut of 4-vectors a and b in an inertial frame S is defined as

$$ab \overset{\text{S}}{\equiv} a_0 b_0 - \left(\vec{a} \cdot \vec{b}\right), \tag{2.2}$$

$$= a_0 b_0 - \sum_{i=1}^{3} a_i b_i, \tag{2.3}$$

where $a_{0\ (1,2,3)}$, $\vec{a}$ are the time (space) component(s) and the 3-space vector in S.

The approximation holds if the lepton mass $m_l \ll E,\ E'$.[3]

- The squared invariant mass of the $\gamma^* - N$ system,

$$W^2 \equiv (q+p)^2 = M_N^2 + 2M_N\nu - Q^2. \tag{2.11}$$

- Bjorken variable,

$$x_B \equiv \frac{-q^2}{2pq} = \frac{Q^2}{2M_N\nu}. \tag{2.12}$$

- The polar angle between the virtual and real photons in the laboratory frame,

$$\theta_{\gamma^*\gamma} \overset{\text{lab}}{\equiv} \arccos\frac{\vec{q}\cdot\vec{q'}}{|\vec{q}|\,|\vec{q'}|}. \tag{2.13}$$

- The azimuthal angle, the angle between the production and scattering planes[4] in the laboratory frame,

$$\phi \overset{\text{lab}}{\equiv} \frac{\vec{q'}\cdot\vec{q}\times\vec{k'}}{\left|\vec{q'}\cdot\vec{q}\times\vec{k'}\right|}\cdot\arccos\frac{\vec{q}\times\vec{k'}\cdot\vec{q}\times\vec{q'}}{\left|\vec{q}\times\vec{k'}\right|\left|\vec{q}\times\vec{q'}\right|}. \tag{2.14}$$

- Missing mass squared, the squared invariant mass of the recoiling nucleon[5],

$$M_x^2 \equiv p'^2 \overset{\text{lab}}{=} W^2 - 2E_\gamma\left(\nu + M_N - \sqrt{\nu^2+Q^2}\cos\theta_{\gamma^*\gamma}\right), \tag{2.17}$$

where E_γ is the energy of the real photon in the laboratory frame.

- The squared 4-momentum transfer between the incoming and outgoing nucleon[6],

$$t \equiv (p'-p)^2 \overset{\text{lab}}{=} -Q^2 - 2E_\gamma\left(\nu - \sqrt{\nu^2+Q^2}\cos\theta_{\gamma^*\gamma}\right). \tag{2.19}$$

---

[3]

$$q^2 = (k-k')^2 \tag{2.7}$$
$$\overset{\text{lab}}{=} m_l^2 + m_l^2 - 2\left(EE' - \sqrt{E^2-m_l^2}\sqrt{E'^2-m_l^2}\cos\theta\right) \tag{2.8}$$
$$\approx -2EE'(1-\cos\theta),\ \text{since } m_l^2 \ll E, E', \tag{2.9}$$
$$= -4EE'\sin^2\frac{\theta}{2}. \tag{2.10}$$

[4] The production plane is spanned by the momenta of γ^* and γ, while the scattering plane by those of l and l'.

[5]

$$p'^2 = \left[(q+p)-q'\right]^2 \tag{2.15}$$
$$= W^2 - 2q'\,(q+p)\,. \tag{2.16}$$

And $q' \overset{\text{lab}}{=} \left(E_\gamma,\ E_\gamma\frac{\vec{q'}}{|\vec{q'}|}\right)$, $q+p \overset{\text{lab}}{=} \left(\nu+M_N,\ \sqrt{\nu^2+Q^2}\frac{\vec{q}}{|\vec{q}|}\right)$.

[6]

$$(p'-p)^2 \overset{\text{lab}}{=} M_x^2 + M_N^2 - 2M_N\,(\nu + M_N - E_\gamma)\,. \tag{2.18}$$

If the target nucleon stays intact, i.e. $M_x^2 = M_N^2$ [7], it can be expressed as the "constrained" momentum transfer,

$$t_c \equiv t(M_x^2 = M_N^2) \tag{2.21}$$

$$\stackrel{\text{lab}}{=} \frac{-Q^2 - 2\nu(\nu - \sqrt{\nu^2+Q^2}\cos\theta_{\gamma^*\gamma})}{1 + \frac{1}{M_N}(\nu - \sqrt{\nu^2+Q^2}\cos\theta_{\gamma^*\gamma})}. \tag{2.22}$$

2.2 Amplitudes

For an unpolarized fixed target in the laboratory frame, the four-fold cross section of Process (2.1) is given by [Bel02]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_B \mathrm{d}y \mathrm{d}\,|t|\,\mathrm{d}\phi} = \frac{\alpha_{em}^3 x_B y}{8\pi Q^2\sqrt{1+\epsilon^2}} \left|\frac{\mathcal{T}}{e^3}\right|^2,\ \epsilon \equiv 2x_B\frac{M_N}{Q}, \tag{2.23}$$

with the fine-structure constant α_{em} and the electron charge magnitude e.

The scattering amplitude $\mathcal{T}$ is given by the coherent sum of the DVCS amplitude $\mathcal{T}_{\text{DVCS}}$ and the BH amplitude $\mathcal{T}_{\text{BH}}$. Thus

$$|\mathcal{T}|^2 = |\mathcal{T}_{\text{BH}}|^2 + |\mathcal{T}_{\text{DVCS}}|^2 + \underbrace{\mathcal{T}_{\text{DVCS}}\mathcal{T}_{\text{BH}}^* + \mathcal{T}_{\text{DVCS}}^*\mathcal{T}_{\text{BH}}}_{\mathcal{I}}, \tag{2.24}$$

where $\mathcal{I}$ denotes the interference term. For a longitudinally polarized lepton beam of polarization λ, the three terms on the right hand side of Eq. (2.24) can be expanded in a Fourier series in ϕ to the twist-three approximation [Bel02]:

$$|\mathcal{T}_{\text{BH}}|^2 = \frac{e^6}{x_B^2 y^2 (1+\epsilon^2)^2 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} \sum_{n=0}^{2} c_n^{\text{BH}} \cos n\phi, \tag{2.25}$$

$$|\mathcal{T}_{\text{DVCS}}|^2 = \frac{e^6}{y^2 Q^2}\left(\sum_{n=0}^{2} c_n^{\text{DVCS}}\cos n\phi + \lambda s_1^{\text{DVCS}}\sin\phi\right), \tag{2.26}$$

$$\mathcal{I} = \eta\cdot\frac{-e^6}{x_B y^3 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}\left(\sum_{n=0}^{3} c_n^{\mathcal{I}}\cos n\phi + \lambda\sum_{n=1}^{2} s_n^{\mathcal{I}}\sin n\phi\right), \tag{2.27}$$

where $\eta = \pm 1$ corresponds to the sign of the lepton charge and $\mathcal{P}_{1,2}(\phi)$ are defined in the lepton BH propagators:

$$Q^2\mathcal{P}_1 \equiv (k - q')^2, \tag{2.28}$$

$$Q^2\mathcal{P}_2 \equiv [k - (p' - p)]^2. \tag{2.29}$$

[7]

$$M_x^2 = M_N^2 \stackrel{\text{lab}}{\Rightarrow} E_\gamma = \frac{2M_N\nu - Q^2}{2\left(\nu + M_N - \sqrt{\nu^2+Q^2}\cos\theta_{\gamma^*\gamma}\right)}. \tag{2.20}$$

The coefficients $c_{0,n}^{\mathrm{BH}}$ can be expressed solely in terms of the Dirac and Pauli form factors of the nucleon; $c_{0,n}^{\mathrm{DVCS},\mathcal{I}}$ and $s_n^{\mathrm{DVCS},\mathcal{I}}$ depend on the Compton form factors (CFFs) which can be given in terms of GPDs.

2.3 Experimental Observables

2.3.1 Beam-Spin Asymmetry

The helicity-independent part of σ_{LU} [8] – the left hand side of Eq. (2.23) – can be separated as

$$\sigma_{\mathrm{UU}}^{\eta} = \frac{\alpha_{em}^3 x_B}{8\pi Q^2\sqrt{1+\epsilon^2}}\left[\frac{\sum_{n=0}^{2} c_n^{\mathrm{BH}}\cos n\phi}{x_B^2 y(1+\epsilon^2)^2 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\mathrm{DVCS}}\cos n\phi}{yQ^2} - \eta\frac{\sum_{n=0}^{3} c_n^{\mathcal{I}}\cos n\phi}{x_B y^2 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}\right], \quad (2.30)$$

which depends on the beam charge. Hence

$$\sigma_{\mathrm{LU}} = \sigma_{\mathrm{UU}}^{\eta}\left[1+\lambda A_{\mathrm{LU}}^{\eta}\right], \quad (2.31)$$

where A_{LU}^{η} is the beam-spin asymmetry (BSA) dependent on the beam charge,

$$A_{\mathrm{LU}}^{\eta} = \frac{\frac{s_1^{\mathrm{DVCS}}\sin\phi}{Q^2} - \eta\frac{s_1^{\mathcal{I}}\sin\phi + s_2^{\mathcal{I}}\sin 2\phi}{x_B y t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}}{\frac{\sum_{n=0}^{2} c_n^{\mathrm{BH}}\cos n\phi}{x_B^2(1+\epsilon^2)^2 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\mathrm{DVCS}}\cos n\phi}{Q^2} - \eta\frac{\sum_{n=0}^{3} c_n^{\mathcal{I}}\cos n\phi}{x_B y t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}}. \quad (2.32)$$

It can be seen that the beam-charge dependence of A_{LU}^{η} arises from the interference parts in both the numerator and denominator. BSA has been measured by HERMES with positron beams [Air01] and by CLAS with electron beams [Ste01].

Because the BH part prevails in the denominator of A_{LU}^{η} with c_0^{BH} dominating over $c_{1,2}^{\mathrm{BH}}$, the interference part prevails in the numerator with $s_1^{\mathcal{I}}$ dominating over $s_2^{\mathcal{I}}$, and ϵ^2 is a small quantity, A_{LU}^{η} can be approximated as

$$A_{\mathrm{LU}}^{\eta} \approx -\eta\frac{x_B}{y}\frac{s_1^{\mathcal{I}}}{c_0^{\mathrm{BH}}}\sin\phi. \quad (2.33)$$

2.3.2 Beam-Charge Asymmetry and Beam-Spin Sub-Asymmetries

Similarly, the helicity- and charge-independent part of σ_{LU} can be separated as

$$\sigma_{\mathrm{UU}}^{0} = \frac{\alpha_{em}^3 x_B}{8\pi Q^2\sqrt{1+\epsilon^2}}\left[\frac{\sum_{n=0}^{2} c_n^{\mathrm{BH}}\cos n\phi}{x_B^2 y(1+\epsilon^2)^2 t\mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\mathrm{DVCS}}\cos n\phi}{yQ^2}\right]. \quad (2.34)$$

[8] LU stands for Longitudinally polarized beams and Unpolarized targets, while UU for Unpolarized beams and Unpolarized targets.

Thus

$$\sigma_{\rm LU} = \sigma^0_{\rm UU}\left[1 + \eta A_{\rm C} + \lambda A^{\rm DVCS}_{\rm LU} + \eta\lambda A^{\mathcal{I}}_{\rm LU}\right], \tag{2.35}$$

where the beam-charge asymmetry (BCA) $A_{\rm C}$ and the beam-spin sub-asymmetries (BSSA) $A^{\rm DVCS}_{\rm LU}$ and $A^{\mathcal{I}}_{\rm LU}$ read

$$A_{\rm C} = -\frac{\frac{\sum_{n=0}^{3} c_n^{\mathcal{I}} \cos n\phi}{x_B y t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}}{\frac{\sum_{n=0}^{2} c_n^{\rm BH} \cos n\phi}{x_B^2(1+\epsilon^2)^2 t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\rm DVCS} \cos n\phi}{Q^2}}, \tag{2.36}$$

$$A^{\rm DVCS}_{\rm LU} = \frac{\frac{s_1^{DVCS} \sin\phi}{Q^2}}{\frac{\sum_{n=0}^{2} c_n^{\rm BH} \cos n\phi}{x_B^2(1+\epsilon^2)^2 t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\rm DVCS} \cos n\phi}{Q^2}}, \tag{2.37}$$

$$A^{\mathcal{I}}_{\rm LU} = -\frac{\frac{\sum_{n=1}^{2} s_n^{\mathcal{I}} \sin n\phi}{x_B y t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}}{\frac{\sum_{n=0}^{2} c_n^{\rm BH} \cos n\phi}{x_B^2(1+\epsilon^2)^2 t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)} + \frac{\sum_{n=0}^{2} c_n^{\rm DVCS} \cos n\phi}{Q^2}}. \tag{2.38}$$

The BSA formulated by Eq. (2.32) is decomposed into $A^{\rm DVCS}_{\rm LU}$ and $A^{\mathcal{I}}_{\rm LU}$ which are independent of the beam charge.

Because the BH parts prevail in the denominators of the three asymmetries with $c_0^{\rm BH}$ dominating over $c_{1,2}^{\rm BH}$, and in the numerators $c_1^{\mathcal{I}}$ dominates over $c_{0,2,3}^{\mathcal{I}}$ and $s_1^{\mathcal{I}}$ over $s_2^{\mathcal{I}}$, the three asymmetries can be approximated as [9]

$$A_{\rm C} \approx -\frac{x_B}{y}\frac{c_1^{\mathcal{I}}}{c_0^{\rm BH}}\cos\phi, \tag{2.39}$$

$$A^{\rm DVCS}_{\rm LU} \approx \frac{x_B^2 t \mathcal{P}_1(\phi)\mathcal{P}_2(\phi)}{Q^2}\frac{s_1^{\rm DVCS}}{c_0^{\rm BH}}\sin\phi, \tag{2.40}$$

$$A^{\mathcal{I}}_{\rm LU} \approx -\frac{x_B}{y}\frac{s_1^{\mathcal{I}}}{c_0^{\rm BH}}\sin\phi. \tag{2.41}$$

Comparing Eqs. (2.33) and (2.41), we have

$$A^{\eta}_{\rm LU} \approx \eta A^{\mathcal{I}}_{\rm LU}. \tag{2.42}$$

[9] The parameterizations of $A_{\rm C}$, $A^{\rm DVCS}_{\rm LU}$, $A^{\mathcal{I}}_{\rm LU}$ – and also $A^{\eta}_{\rm LU}$ – in this analysis are not limited by such leading effect approximations; other order harmonics are taken into account in the asymmetry extraction (cf. Section 4.2.4).

Chapter 3

the HERMES Experiment

The HERMES experiment at the HERA storage ring at DESY is a fixed target experiment scattering longitudinally polarized positron/electron beams off polarized or unpolarized gas targets.

Because the HERMES detector (Fig. 3.1) is a forward spectrometer [Ack98] and the BH process dominates in the kinematic region covered by HERMES, it is difficult to extract the Fourier coefficients in $|\mathcal{T}_{DVCS}|^2$ and $\mathcal{I}$ by cross section measurements. However, the measurement of cross section asymmetries with respect to the beam charge and beam helicity provides an alternative way to access these coefficients and thus GPDs. Except for a short period in 1998, HERA was running with electron beams in 2005 and 2006 . Thus at HERMES, a precise measurement of these asymmetries with an electron beam became possible only after the 2005 data production had been released.

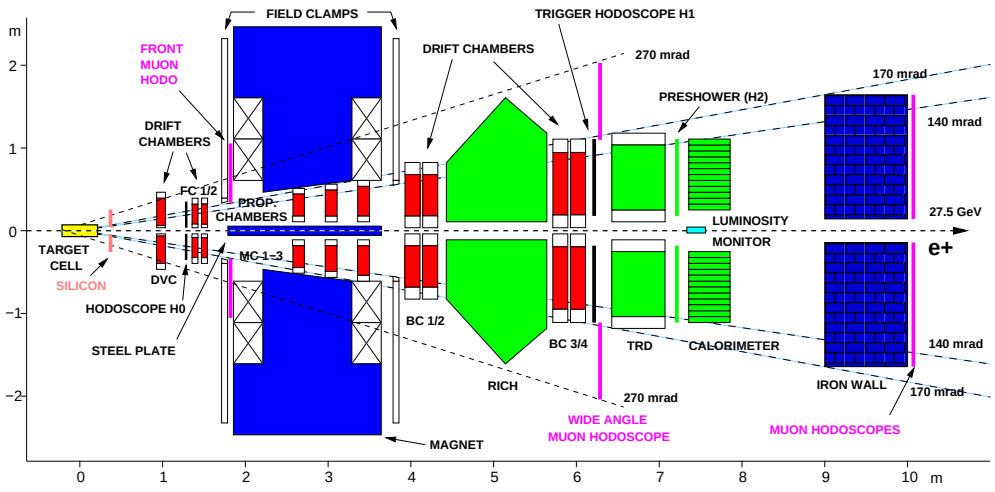

Figure 3.1: The HERMES spectrometer. Tracking system (red): drift vertex chamber (DVC), front chambers (FC1/2), magnet chambers (MC1-3) and back chambers (BC1/2,BC3/4). Particle identification system (green): ring-imaging Čerenkov-counter (RICH), transition radiation detector (TRD), pre-shower detector (H2), and calorimeter.

3.1 Beams and Targets

The positron/electron beams at HERA are transversely self-polarized. The Sokolov-Ternov effect builds up a polarization up to a maximum of 92% [Jac76], while the beam-beam interaction and non-vertical magnetic field components degrade it to about 60%, the observed level. The transverse beam polarization is converted to be longitudinal by the spin-rotator in front of HERMES and back by the one behind.

From 2002 to 2005, the target was transversely polarized. The transverse magnetic field in the target region introduces deflection to the charged particles. Thus certain transverse magnet correction (TMC) methods are required to calculate the vertices and scattering angles corresponding to the case without the transverse magnetic field.

3.2 Detection of Final-State Particles

At HERMES, the detection of the scattered lepton and real photon in Process (2.1) is mainly accomplished by the tracking and particle identification systems. Before 2006, the recoiling nucleon can not be detected.

The tracking system consists of the drift vertex chamber (DVC), front chambers (FCs), magnet chambers (MCs), and the back chambers (BCs).

- Event vertices and scattering angles are measured by DVC and FCs.
- The particle momentum is determined from the track deflection by the dipole magnets. The measurement is accomplished by BCs in combination with DVC and FCs.

The particle identification system consists of the ring-imaging Čerenkov-counter (RICH), transition radiation detector (TRD), pre-shower detector and the electromagnetic calorimeter. The last three are mainly for lepton-hadron separation, which is critical in this analysis (cf. Section 4.1).

RICH designed to separate pions, protons, and kaons, contributes to the lepton-hadron separation only for those particles with energies below about 4 GeV. Thus in this analysis it is not used.

TRD exploits the transition radiation of leptons to fulfill the lepton-hadron separation. When crossing boundaries between media of different dielectric constants, charged particles emit transition radiation proportional to its Lorentz factor. Thus with the same initial energy, a lepton deposits more energies than a hadron does, the latter through ionization mostly.

The pre-shower detector is based on the different formation mechanisms of leptonic and hadronic electromagnetic showers. Leptonic electromagnetic showers start usually much earlier than hadronic ones do, and hence generate more tracks in the detector and deposit more energies.

The calorimeter is of the size chosen in such a way that it contains most of the leptonic showers, but only a fraction of the hadronic ones. Thus for a lepton, the energy measured in the calorimeter approximately equals to its track momentum.

The PID system – $\mathrm{PID}_{2,3,5}$ – provides a criterion for the lepton-hadron separation. The probability of a particle being a lepton/hadron is a function of the detector response and the particle momentum. Define

$$PID_{\mathcal{D}} = \log_{10} \frac{P_l(p,R)}{P_h(p,R)}, \tag{3.1}$$

where $P_{l/h}(p,R)$ is the probability of a detected particle being a lepton/hadron if its momentum is p and the response of the $\mathcal{D}$-detector is R. The following combinations of $PID_{\mathcal{D}}$s make up the PID system:

$$PID_2 = PID_{preshower} + PID_{calorimeter}, \tag{3.2}$$

$$PID_3 = PID_{preshower} + PID_{calorimeter} + PID_{RICH}, \tag{3.3}$$

$$PID_5 = PID_{TRD}, \tag{3.4}$$

which reflect the relative probabilities of a detected particle being a lepton according to different combinations of detectors.

Chapter 4

Data Analysis

This analysis is based on the 00d0 e^+, unpolarized H production and the 05b1 e^-, unpolarized H and transversely polarized H productions. The net target polarization of the target-polarized data is negligible ($\langle$ g1Target.rPol $\rangle = 0.6\%$) and thus the whole data set is regarded as target-unpolarized.

4.1 Data Quality Cuts and Event Selection

To select events of Process (2.1), each candidate has to fulfill the requirements for the data quality, the DIS criteria (to select hard reactions) and exclusivity (to select exclusive processes).

The data selection criteria are listed below:

- Each event was in a burst that passes the following cuts (for both spectrometer halves):

  badbits & 0x503E13DC ==0 (for 00d0 unpol.),
  badbits & 0x501E13DC ==0 (for 05b1 unpol.),
  badbits & 0x5D9FFFFD ==0 (for 05b1 pol.),

  which correspond to the following data quality cuts:

  - the dead time was reasonable:
    0.5 <= g1DAQ.rDearCorr21 <= 1.0;
  - the burst length was reasonable:
    0 < g1DAQ.rLength <= 11 sec;
  - the beam current was reasonable:
    5, 2 <= g1Beam.rMdmCurr <= 50 mA for 00d0, 05b1 respectively;
  - the burst was not the first one in a run;
  - the burst had no bad uDST records;
  - PID was available;

  - the burst was in a run marked as analyzable in the logrun.new file;
  - DQ info was available;
  - no dead blocks occurred in the calorimeter;
  - no dead blocks occurred in the preshower or lumi monitor;
  - TRD was functioning;
  - no HV trips occurred in the tracking chambers;
  - the latest beam polarization measurement was less than 5 minutes ago;
  - *for polarized runs*, the target was in a reasonable state;
  - *for polarized runs*, the fitted lumi was reasonable:
    1 <= g1Beam.rLumiFitBstGai <= 100.

- In addition the burst had to pass the following data quality cuts:
  - the raw luminosity count rate was between 5 and 10000 (3000) for unpol. (pol). runs:
    5 < g1Beam.rLumiRate < 10000(3000) for unpol.(pol.) runs respectively;
  - the dead time was larger than 0.8:
    0.8 < g1DAQ.rDearCorr21 <= 1.0;
  - TRD was checked without any problems:
    g1Quality.iTrdDQ==3;
  - the measured beam polarization was between 30 (20) and 80 for 00d0 (05b1):
    30(20) < | g1Beam.rPolFit | < 80 [1];
  - *for polarized runs*, the target was not in the flipping state:
    g1DAQ.iTargetBit & 0x2 != 2;
  - *for polarized runs*, the raw target polarization was between 0.5 and 1.5:
    0.5 < | g1Target.rPol | < 1.5;

- DIS criteria:
  - trigger 21 was fired:
    smTrack.bTrigMask & (1 << (21-1));
  - each event contained at least one track that satisfied:
    * having the same charge as the beam:
      g1Track.rEnergy*g1Beam.iBeamCharge>0;

[1] Because of incorrect records in the udst files, g1Beam.rPolFit in runs 26036-44748 of 05b1 should be checked to be negative (multiply -1 if positive).

  - identified as a lepton:
    g1Track.rPID2+g1Track.rPID5 > 2;
  - selected as a long track:
    !(g1Track.iSelect&0x0100)&&!(g1Track.iSelect&0x0200);
  - *for polarized runs*, Transverse Magnet Correction (TMC) was successful:
    method2 used for 05b1 pol., g1Track.bTMCStat&0x80000000 == 0;
  - lepton fiducial volume cut:
    - |g1Track.rVertZ|, |g1Track.rVertZCor2| < 18 cm for unpol., pol. runs,
    - |g1Track.rVertD| < 0.75 cm for unpol. runs,
    - $|x^e_{calo}| < 175$ cm, $30 < |y^e_{calo}| < 108$ cm;[2]
  - additional fiducial volume cuts:
    - |smTrack.rxOff + 172·tanθ_x| < 31 cm,
    - |smTrack.ryOff + 181·tanθ_y| > 7 cm,
    - |smTrack.ryOff + 383·tanθ_y| < 54 cm,[3]
    - |smTrack.rXpos+(383-275)·smTrack.rXslope| <= 100 cm,
    - |smTrack.rYpos+(383-275)·smTrack.rYslope| <= 54 cm;

- each event passes the following kinematics cuts:
  - $W^2 > 9$ GeV^2, $Q^2 > 1$ GeV^2, $\nu < 22$ GeV.

- Exclusivity:
  - each event contained exactly one track, and exactly one untracked cluster in the calorimeter;
  - photon fiducial volume cuts:
    - $|x^\gamma_{calo}| < 125$ cm, $33 < |y^\gamma_{calo}| < 105$ cm;[4]
  - the photon gives a signal in the preshower:
    smCluster.rPulsPre > 0.001 GeV;
  - each event passes the following kinematics cuts:
    - 5 GeV< E_γ < g1Beam.rHeraElEnergy, g1Beam.rHeraElEnergy > 27 GeV;[5]
    - $5 < \theta_{\gamma^*\gamma} < 45$ mrad;
    - $-1.5 < M_x < 1.7$ GeV, $-t_c < 0.7$ GeV^2, $0.03 < x_B < 0.35$, $Q^2 < 10$ GeV^2.

[2] x^e_{calo} = smTrack.rXpos+(738-275)·smTrack.rXslope, y^e_{calo} = smTrack.rYpos+(738-275)·smTrack.rYslope.
[3] $\tan\theta_x = \frac{k'_x}{k'_z}$, $\tan\theta_y = \frac{k'_y}{k'_z}$.
[4] x^γ_{calo} = smCluster.rXlEw, y^γ_{calo} = smCluster.rYlEw.
[5] E_γ=smCluster.rE.

The kinematic distributions of the selected DIS and hard exclusive events are shown in Fig. 4.1 and 4.2 respectively. It can be seen that the distributions of the DIS events from the e^+ and e^- data are not identical, which may be related to tracking problems or different calorimeter thresholds. And the discrepancy between the distributions of missing mass from the two data sets reveals possible problems in the photon energy measurement. Corrections and systematic errors will be investigated in future analyses.

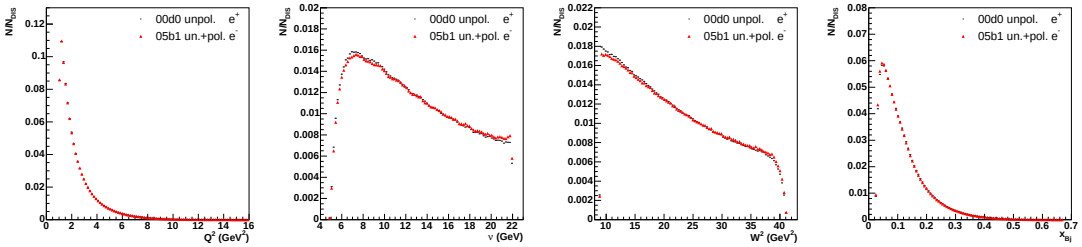

Figure 4.1: Normalized kinematic distributions of the DIS events.

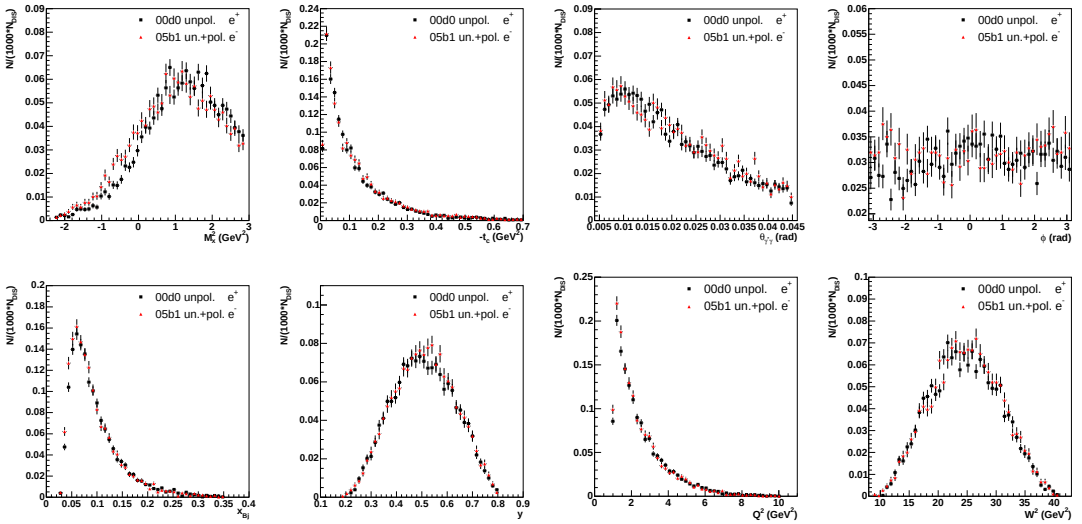

Figure 4.2: Kinematic distributions of the exclusive events, normalized per 1000 DIS events.

4.2 Asymmetry Extraction

In this analysis, the beam-spin asymmetry, beam-charge asymmetry and the beam-spin sub-asymmetries are extracted with the extended maximum likelihood (EML) method, which is an extension of the standard. This method takes into account the statistical property of the number of events and thus provides a more precise description for the experiment.

4.2.1 Extended Maximum Likelihood Method

Suppose that a population is characterized by M parameters, which are written in a vector form, $\underline{\theta} = (\theta_1,\ \theta_2, \ldots,\ \theta_M)^T$, where T denotes transpose. From the population, a sample of the observable x (which can be a vector with multi-components), $x_1,\ x_2, \ldots,\ x_N$, is obtained by N independent observations. The probability density function (pdf) $p(x;\underline{\theta})$ describes its normalized distribution [6]. In practice, $p(x;\underline{\theta})$ is usually evaluated by normalizing the unnormalized distribution function $P(x;\underline{\theta})$:

$$p(x;\underline{\theta}) = \frac{P(x;\underline{\theta})}{\mathbb{N}(\underline{\theta})}, \tag{4.2}$$

where $\mathbb{N}(\underline{\theta})$ is the normalization factor.

Under general conditions [7], especially in the large sample limit [Ken79], consistent estimators for $\underline{\theta}$ can be found by solving the following equations,

$$\frac{\partial L}{\partial \underline{\theta}} = 0, \tag{4.3}$$

or equivalently

$$\frac{\partial \ln L}{\partial \underline{\theta}} = 0, \tag{4.4}$$

in which $L(x_1,\ x_2, \ldots,\ x_N;\underline{\theta})$ is the likelihood function, giving the hypothetical probability for outcomes $x_1,\ x_2, \ldots,\ x_N$. This is the maximum likelihood (ML) estimation.

In the unweighted standard case, the number of events is fixable. L is then defined as the joint pdf for the data [Yao06],

$$L_{\mathrm{SML}} \equiv \prod_{i=1}^{N} \frac{P(x_i;\underline{\theta})}{\mathbb{N}(\underline{\theta})}. \tag{4.5}$$

[6]

$$\int_{\mathbf{X}} p(x;\underline{\theta})\mathrm{d}x = 1, \forall \underline{\theta} \in \underline{\mathbf{\Theta}}, \tag{4.1}$$

in which $\underline{\mathbf{\Theta}}$ are the permissible ranges for $\underline{\theta}$ and $\mathbf{X}$ the $\underline{\theta}$-independent domain of x.

[7] For example, $p(x;\underline{\theta})$ or $P(x;\underline{\theta})$ is properly parameterized; L is third-order differentiable in $\underline{\mathbf{\Theta}}$ without terminal maxima and possesses a single extreme in $\underline{\mathbf{\Theta}}$ (cf. Appendix A.1 and A.3).

Therefore

$$\ln L_{\mathrm{SML}} = \sum_{i=1}^{N} \ln P(x_i;\underline{\theta}) - N \ln \mathbb{N}(\underline{\theta}). \tag{4.6}$$

However, in nuclear and particle physics experiments, the observed number of events usually has a Poisson fluctuation about its expected value $\mathbb{N}(\underline{\theta})$. In this extended case the unweighted likelihood function is defined as [Bar90]

$$L_{\mathrm{EML}} \equiv \frac{e^{-\mathbb{N}(\underline{\theta})}[\mathbb{N}(\underline{\theta})]^N}{N!} \prod_{i=1}^{N} \frac{P(x_i;\underline{\theta})}{\mathbb{N}(\underline{\theta})}. \tag{4.7}$$

Therefore

$$\ln L_{\mathrm{EML}} = \sum_{i=1}^{N} \ln P(x_i;\underline{\theta}) - \mathbb{N}(\underline{\theta}). \tag{4.8}$$

The covariance matrixes of the SML and EML estimators are given asymptotically by

$$\left(-\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j} \right)^{-1} \tag{4.9}$$

evaluated at the roots of Eq. (4.4).

Both the SML and EML estimators are consistent (cf. Appendix A.3). However, their covariance matrixes are generally different. If the number of events does follow the Poisson distribution, as is the case in this analysis, the EML estimators properly incorporate this additional information.

The distinction between SML and EML can not be eliminated and the calculation of $\mathbb{N}(\underline{\theta})$ is a basic requirement for both the methods if $\mathbb{N}(\underline{\theta})$ depends on $\underline{\theta}$, which is generally the case unless the data are filtered or weighted. However, by filtering, the sample size is decreased; by weighting, the covariance matrix is not given by Eq. (4.9) [Sol64]. Furthermore, in the latter case, for the SML estimators, the covariance matrix is practically difficult to calculate; for the EML estimators, it requires another minimization procedure with the weights squared (cf. Appendix A.3) in which the square-weighted $\mathbb{N}(\underline{\theta})$ depends on $\underline{\theta}$. Therefore, by no means can the issues of $\mathbb{N}(\underline{\theta})$ be easily avoided.

4.2.2 Normalization Factors

From Eq. (2.35), not considering the acceptance effect and detection efficiency, the expected number density of exclusive events in the τ-ϕ-$\mathbf{x}$ space reads

$$n(\tau,\phi,\mathbf{x};\underline{\theta}) = \mathcal{L}(\tau)\sigma^0_{\mathrm{UU}}(\phi,\mathbf{x})\underline{b}^T(\tau)\underline{a}(\phi,\mathbf{x};\underline{\theta}), \tag{4.10}$$

with

$$\underline{b}(\tau) \equiv \begin{pmatrix} 1 \\ \eta(\tau) \\ \lambda(\tau) \\ \eta(\tau)\lambda(\tau) \end{pmatrix}, \qquad \underline{a}(\phi, \mathbf{x}; \underline{\theta}) \equiv \begin{pmatrix} 1 \\ A_{\mathrm{C}}(\phi, \mathbf{x}; \underline{\theta}) \\ A_{\mathrm{LU}}^{\mathrm{DVCS}}(\phi, \mathbf{x}; \underline{\theta}) \\ A_{\mathrm{LU}}^{\mathcal{I}}(\phi, \mathbf{x}; \underline{\theta}) \end{pmatrix}, \tag{4.11}$$

in which τ is time, $\mathbf{x}$ the set of kinematic variables $\{-t_c,\ x_B,\ Q^2,\ \dots\}$ not including ϕ. And $\mathcal{L}(\tau)$ is the luminosity which can be evaluated by the number density of the DIS events,

$$n_{\mathrm{DIS}}(\tau) = \mathcal{L}(\tau)\sigma_{\mathrm{DIS}}, \tag{4.12}$$

where σ_{DIS} is the total cross section of the DIS process. Thus

$$n(\tau, \phi, \mathbf{x}; \underline{\theta}) = n_{\mathrm{DIS}}(\tau) r(\phi, \mathbf{x}) \underline{b}^T(\tau) \underline{a}(\phi, \mathbf{x}; \underline{\theta}),\ r(\phi, \mathbf{x}) \equiv \frac{\sigma_{\mathrm{UU}}^0(\phi, \mathbf{x})}{\sigma_{\mathrm{DIS}}}. \tag{4.13}$$

Define summation operations with respect to different ranges of τ,

$$\underline{\sum}_{\tau} \equiv \begin{pmatrix} \sum_{\forall\tau, \eta(\tau)=+1, \lambda(\tau)>0} \\ \sum_{\forall\tau, \eta(\tau)=+1, \lambda(\tau)<0} \\ \sum_{\forall\tau, \eta(\tau)=-1, \lambda(\tau)>0} \\ \sum_{\forall\tau, \eta(\tau)=-1, \lambda(\tau)<0} \end{pmatrix}. \tag{4.14}$$

The corresponding number densities in ϕ and $\mathbf{x}$ can be generated,

$$\underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta}) \equiv \underline{\sum}_{\tau} n(\tau, \phi, \mathbf{x}; \underline{\theta}) = \begin{pmatrix} \overrightarrow{\mathcal{N}}^+(\phi, \mathbf{x}; \underline{\theta}) \\ \overleftarrow{\mathcal{N}}^+(\phi, \mathbf{x}; \underline{\theta}) \\ \overrightarrow{\mathcal{N}}^-(\phi, \mathbf{x}; \underline{\theta}) \\ \overleftarrow{\mathcal{N}}^-(\phi, \mathbf{x}; \underline{\theta}) \end{pmatrix}, \tag{4.15}$$

and also a constant matrix encoding the beam-charge and -helicity state,

$$\underline{\underline{\mathbb{B}}} \equiv \underline{\sum}_{\tau} n_{\mathrm{DIS}}(\tau) \underline{b}^T(\tau) = \begin{pmatrix} \overrightarrow{N}^+_{\mathrm{DIS}} & \overrightarrow{N}^+_{\mathrm{DIS}} & \overrightarrow{N}^+_{\mathrm{DIS}} \left\langle \overrightarrow{\lambda}^+ \right\rangle & \overrightarrow{N}^+_{\mathrm{DIS}} \left\langle \overrightarrow{\lambda}^+ \right\rangle \\ \overleftarrow{N}^+_{\mathrm{DIS}} & \overleftarrow{N}^+_{\mathrm{DIS}} & \overleftarrow{N}^+_{\mathrm{DIS}} \left\langle \overleftarrow{\lambda}^+ \right\rangle & \overleftarrow{N}^+_{\mathrm{DIS}} \left\langle \overleftarrow{\lambda}^+ \right\rangle \\ \overrightarrow{N}^-_{\mathrm{DIS}} & -\overrightarrow{N}^-_{\mathrm{DIS}} & \overrightarrow{N}^-_{\mathrm{DIS}} \left\langle \overrightarrow{\lambda}^- \right\rangle & -\overrightarrow{N}^-_{\mathrm{DIS}} \left\langle \overrightarrow{\lambda}^- \right\rangle \\ \overleftarrow{N}^-_{\mathrm{DIS}} & -\overleftarrow{N}^-_{\mathrm{DIS}} & \overleftarrow{N}^-_{\mathrm{DIS}} \left\langle \overleftarrow{\lambda}^- \right\rangle & -\overleftarrow{N}^-_{\mathrm{DIS}} \left\langle \overleftarrow{\lambda}^- \right\rangle \end{pmatrix}, \tag{4.16}$$

where $\rightarrow$ ($\leftarrow$) and $^+$ ($^-$) denote positive (negative) beam polarization and charge, N_{DIS} the number of the DIS events, and $\langle\lambda\rangle$ the average beam polarization over the DIS events. Generally $\underline{\underline{\mathbb{B}}}$ is non-singular.

From Eqs. (4.13), (4.15) and (4.16), we have

$$\underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta}) = \underline{\underline{\mathbb{B}}}\, r(\phi, \mathbf{x}) \underline{a}(\phi, \mathbf{x}; \underline{\theta}). \tag{4.17}$$

Since the first component of $\underline{a}$ equals 1, r can be solved as

$$\begin{aligned} r(\phi, \mathbf{x}) &= \left[r(\phi, \mathbf{x}) \underline{a}(\phi, \mathbf{x}; \underline{\theta}) \right]_1 && (4.18) \\ &= \left[\underline{\underline{\mathbb{B}}}^{-1} \underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta}) \right]_1 && (4.19) \\ &= \sum_{k=1}^{4} \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{1k} \left[\underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta}) \right]_k . && (4.20) \end{aligned}$$

In the large $\underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta})$ limit, the observed number densities $\underline{\mathcal{N}_{\rm obs}}(\phi, \mathbf{x})$ trend to their expected values $\underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta})$. Therefore, asymptotically

$$r(\phi, \mathbf{x}) = \sum_{k=1}^{4} \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{1k} \left[\underline{\mathcal{N}_{\rm obs}}(\phi, \mathbf{x}) \right]_k . \tag{4.21}$$

Thus the number density in ϕ and $\mathbf{x}$ over the whole range of τ reads

$$\begin{aligned} \mathcal{N}(\phi, \mathbf{x}; \underline{\theta}) &\equiv \underline{s}^T \, \underline{\mathcal{N}}(\phi, \mathbf{x}; \underline{\theta}) && (4.22) \\ &= \sum_{k=1}^{4} \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{1k} \underline{s}^T \, \underline{\underline{\mathbb{B}}} \, \underline{a}(\phi, \mathbf{x}; \underline{\theta}) \left[\underline{\mathcal{N}_{\rm obs}}(\phi, \mathbf{x}) \right]_k , \; \underline{s} \equiv (1,\ 1,\ 1,\ 1)^T . && (4.23) \end{aligned}$$

Finally we obtain the number density over the whole range of τ and ϕ,

$$\begin{aligned} \mathfrak{N}(\mathbf{x}; \underline{\theta}) &\equiv \sum_{\phi} \mathcal{N}(\phi, \mathbf{x}; \underline{\theta}) && (4.24) \\ &= \sum_{\phi} \sum_{k=1}^{4} \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{1k} \underline{s}^T \, \underline{\underline{\mathbb{B}}} \, \underline{a}(\phi, \mathbf{x}; \underline{\theta}) \left[\underline{\mathcal{N}_{\rm obs}}(\phi, \mathbf{x}) \right]_k && (4.25) \\ &= \underline{s}^T \, \underline{\underline{\mathbb{B}}} \sum_{i=1}^{\mathfrak{N}_{\rm obs}(\mathbf{x})} W_i \, \underline{a}(\phi_i, \mathbf{x}; \underline{\theta}), && (4.26) \end{aligned}$$

where $\mathfrak{N}_{\rm obs}(\mathbf{x})$ is the observed number density in $\mathbf{x}$ and

$$W_i = \begin{cases} \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{11} & (\eta_i = +1,\ \lambda_i > 0) \\ \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{12} & (\eta_i = +1,\ \lambda_i < 0) \\ \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{13} & (\eta_i = -1,\ \lambda_i > 0) \\ \left[\underline{\underline{\mathbb{B}}}^{-1} \right]_{14} & (\eta_i = -1,\ \lambda_i < 0) \end{cases} , \tag{4.27}$$

which is coupled to each data point $(\phi_i,\ \eta_i,\ \lambda_i)$.

Similarly, for a sample of a certain beam charge, the number density based on Eq. (2.31) reads [8]

$$\mathfrak{N}(\mathbf{x};\underline{\theta}) = \underline{s}^T \,\underline{\underline{\mathbb{B}}} \sum_{i=1}^{\mathfrak{N}_{\rm obs}(\mathbf{x})} W_i \,\underline{a}(\phi_i,\mathbf{x};\underline{\theta}), \tag{4.28}$$

with

$$\underline{a}(\phi,\mathbf{x};\underline{\theta}) \equiv \begin{pmatrix} 1 \\ A_{\rm LU}(\phi,\mathbf{x};\underline{\theta}) \end{pmatrix}, \qquad \underline{\underline{\mathbb{B}}} \equiv \begin{pmatrix} \overrightarrow{N}_{\rm DIS} & \overrightarrow{N}_{\rm DIS}\left\langle\overrightarrow{\lambda}\right\rangle \\ \overleftarrow{N}_{\rm DIS} & \overleftarrow{N}_{\rm DIS}\left\langle\overleftarrow{\lambda}\right\rangle \end{pmatrix}, \tag{4.29}$$

$$\underline{s} \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad W_i \equiv \begin{cases} \left[\underline{\underline{\mathbb{B}}}^{-1}\right]_{11} & (\lambda_i > 0) \\ \left[\underline{\underline{\mathbb{B}}}^{-1}\right]_{12} & (\lambda_i < 0) \end{cases}. \tag{4.30}$$

4.2.3 Likelihood Functions

If we focus on the average effect of the asymmetries in a certain kinematic region [$\mathbf{x}$], Eqs. (4.26) and (4.28) become

$$\boldsymbol{N}(\underline{\theta}) = \underline{s}^T \,\underline{\underline{\mathbb{B}}} \sum_{i=1}^{\boldsymbol{N}_{\rm obs}} W_i \,\underline{\mathfrak{a}}(\phi_i;\underline{\theta}), \tag{4.31}$$

in which $\boldsymbol{N}(\underline{\theta})/\boldsymbol{N}_{\rm obs}$ is the expected/observed number of the exclusive events in [$\mathbf{x}$] and $\underline{\mathfrak{a}}(\phi;\underline{\theta})$ the average asymmetries.

The negative log-EML functions to minimize are

$$-\ln L_{\rm EML} = -\sum_{i=1}^{\boldsymbol{N}_{\rm obs}} \ln\left[\underline{b}_i^T \underline{\mathfrak{a}}(\phi_i;\underline{\theta})\right] + \underline{s}^T \,\underline{\underline{\mathbb{B}}} \sum_{i=1}^{\boldsymbol{N}_{\rm obs}} W_i \,\underline{\mathfrak{a}}(\phi_i;\underline{\theta}), \tag{4.32}$$

where $\underline{b}_i \equiv (1,\ \eta_i,\ \lambda_i,\ \eta_i\lambda_i)^T$ and $(1,\ \lambda_i)^T$ respectively.

4.2.4 Parameterizations of Asymmetries

The asymmetries in Eq. (4.32) can be parameterized in several ways as shown below. Nevertheless, the results obtained are consistent (Fig. 4.3). For final results, the median parameterizations are chosen, i.e. 3 parameters for BSA and 9 parameters for BCA and BSSA.

- **BSA**

  1. *1 parameter*

$$A_{\rm LU}(\phi;s_1) = s_1 \sin\phi; \tag{4.33}$$

[8] The superscript $^\eta$ is omitted.

2. *3 parameters*

$$A_{\mathrm{LU}}(\phi; c_0, s_1, c_1) = c_0 + s_1 \sin\phi + c_1 \cos\phi; \tag{4.34}$$

3. *5 parameters*

$$A_{\mathrm{LU}}(\phi; c_0, s_1, c_1, s_2, c_2) = c_0 + s_1 \sin\phi + c_1 \cos\phi + s_2 \sin 2\phi + c_2 \cos 2\phi; \tag{4.35}$$

- **BCA and BSSA** [9]

1. *3 parameters*

$$A_{\mathrm{C}}(\phi; c_1) = c_1 \cos\phi; \tag{4.36}$$
$$A_{\mathrm{LU}}^{\mathrm{DVCS}}(\phi; s_1) = s_1 \sin\phi; \tag{4.37}$$
$$A_{\mathrm{LU}}^{\mathcal{I}}(\phi; s_1) = s_1 \sin\phi; \tag{4.38}$$

2. *9 parameters*

$$A_{\mathrm{C}}(\phi; c_0, s_1, c_1) = c_0 + s_1 \sin\phi + c_1 \cos\phi; \tag{4.39}$$
$$A_{\mathrm{LU}}^{\mathrm{DVCS}}(\phi; c_0, s_1, c_1) = c_0 + s_1 \sin\phi + c_1 \cos\phi; \tag{4.40}$$
$$A_{\mathrm{LU}}^{\mathcal{I}}(\phi; c_0, s_1, c_1) = c_0 + s_1 \sin\phi + c_1 \cos\phi; \tag{4.41}$$

3. *15 parameters*

$$A_{\mathrm{C}}(\phi; c_0, s_1, c_1, s_2, c_2) = c_0 + s_1 \sin\phi + c_1 \cos\phi + s_2 \sin\phi + c_2 \cos\phi; \tag{4.42}$$
$$A_{\mathrm{LU}}^{\mathrm{DVCS}}(\phi; c_0, s_1, c_1, s_2, c_2) = c_0 + s_1 \sin\phi + c_1 \cos\phi + s_2 \sin\phi + c_2 \cos\phi; \tag{4.43}$$
$$A_{\mathrm{LU}}^{\mathcal{I}}(\phi; c_0, s_1, c_1, s_2, c_2) = c_0 + s_1 \sin\phi + c_1 \cos\phi + s_2 \sin\phi + c_2 \cos\phi. \tag{4.44}$$

[9] The coefficients are local in the asymmetries, e.g. s_1s in $A_{\mathrm{LU}}^{\mathrm{DVCS}}$ and $A_{\mathrm{LU}}^{\mathcal{I}}$ are not related.

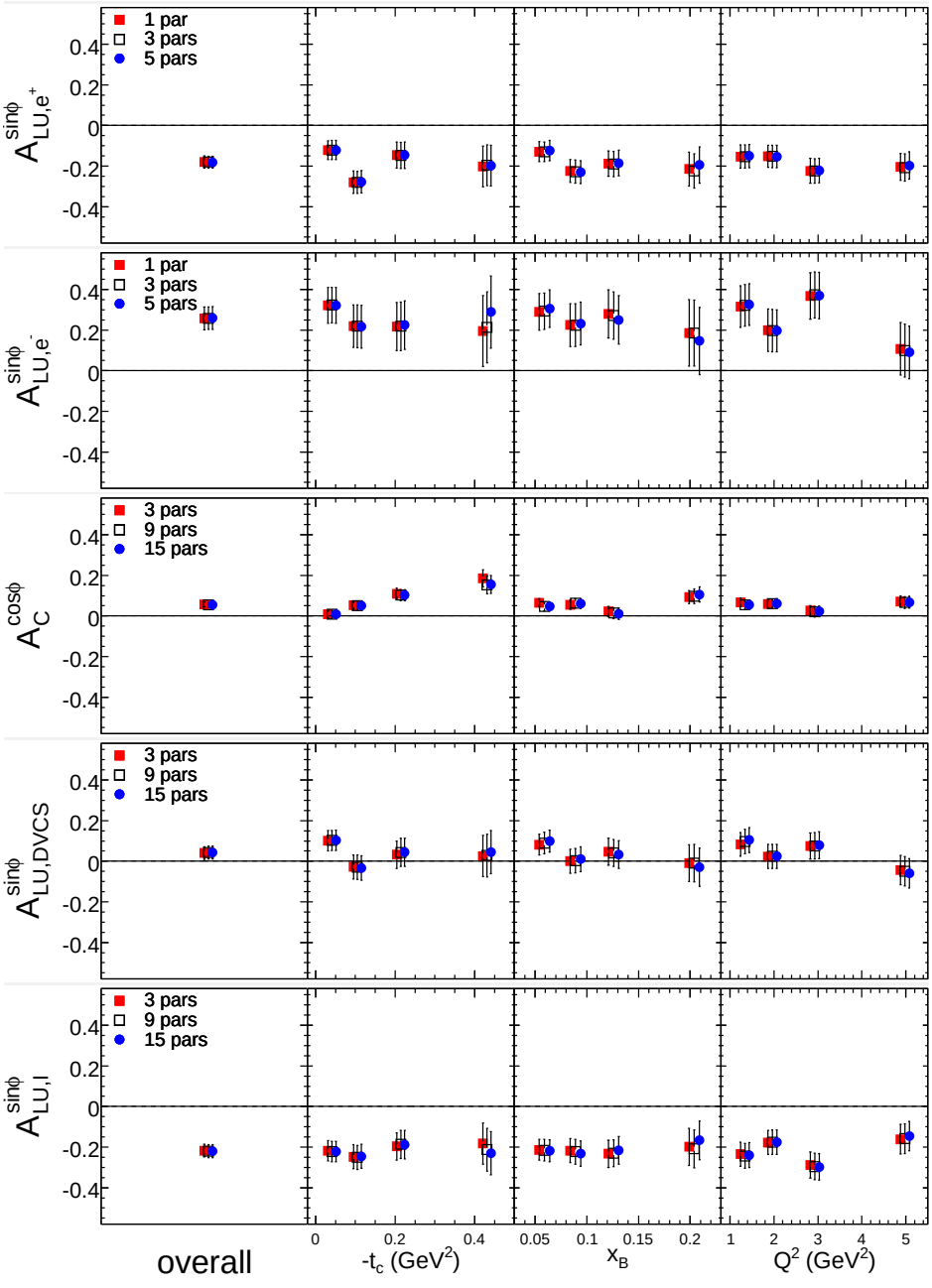

Figure 4.3: Fitting results for the leading moments with different parameterizations.

Chapter 5

Results

After the data quality cuts and event selection, DIS events and exclusive events are selected from the 00d0 e^+ and 05b1 e^- productions. The numbers of the DIS events and the average beam polarizations are shown in Table 5.1. These quantities are the building blocks of the "beam state matrix" $\underline{\underline{\mathbb{B}}}$ in Eqs. (4.16) and (4.29), and thus are important for the normalization factors of the exclusive events.

Because of the asymptotic properties of the ML estimators, the sample size influences the performance of the extraction method. The numbers of the exclusive events are shown in Table 5.2.

With the method described in Section 4.2, the asymmetries in kinematic bins of $-t_c$, x_B and Q^2 are extracted (Fig. 5.1 and Tables 5.3-5.5). In the overall kinematic region, $A_{\rm LU}^{\rm DVCS}$ is consistent with zero within 1.4 times the statistical uncertainty. Larger statistics would be needed to confirm a non-zero result. $A_{\rm LU}^{+}$, $-A_{\rm LU}^{-}$ and $A_{\rm LU}^{\mathcal{I}}$ are consistent within statistical uncertainties, as is expected from Eq. (2.42).

$A_{\rm LU}^{+}$ is compared to the result in [Ell03] based on a sample of 4014 positive- and 5373 negative-helicity-exclusive events from the 00c1 production. The two results are consistent (Fig. 5.2).

$A_{\rm C}$ is compared to the result in [Air07] based on a sample of about 9500 positron- and 700 electron-exclusive events from the 1998 and 2000 data. The two results are consistent and the statistical uncertainties of the current one are smaller by a factor of about 2 (Fig. 5.3) due to a large increase in the sample size of the electron data.

This work has been cross-checked by another two analyzers – T. Keri and D. Zeiler at HERMES. A good agreement is seen in the numbers of events (Tabel 5.6) and in the extracted asymmetries (Fig. 5.4).

$\overrightarrow{N}^{+}_{\mathrm{DIS}}$	$\overleftarrow{N}^{+}_{\mathrm{DIS}}$	$\overrightarrow{N}^{-}_{\mathrm{DIS}}$	$\overleftarrow{N}^{-}_{\mathrm{DIS}}$	$\left\langle \overrightarrow{\lambda}^{+} \right\rangle$	$\left\langle \overleftarrow{\lambda}^{+} \right\rangle$	$\left\langle \overrightarrow{\lambda}^{-} \right\rangle$	$\left\langle \overleftarrow{\lambda}^{-} \right\rangle$
1,858,977	3,184,492	1,512,975	2,158,798	0.544	-0.550	0.363	-0.304

Table 5.1: Numbers of the DIS events and the average beam polarizations over the DIS events.

kinematic bin		$\boldsymbol{N}_{\mathrm{obs}}$	$\overrightarrow{\boldsymbol{N}}^{+}_{\mathrm{obs}}$	$\overleftarrow{\boldsymbol{N}}^{+}_{\mathrm{obs}}$	$\overrightarrow{\boldsymbol{N}}^{-}_{\mathrm{obs}}$	$\overleftarrow{\boldsymbol{N}}^{-}_{\mathrm{obs}}$
overall		13528	2772	4937	2368	3451
$-t$ (GeV^2)	0.00–0.06	5506	1138	2030	934	1404
	0.06–0.14	3998	844	1478	698	978
	0.14–0.30	2867	554	1057	525	731
	0.30–0.70	1157	236	372	211	338
x_B	0.03–0.07	5074	1054	1769	925	1326
	0.07–0.10	3742	734	1401	662	945
	0.10–0.15	2979	638	1095	476	770
	0.15–0.35	1733	346	672	305	410
Q^2 (GeV^2)	1.0–1.5	3879	771	1370	705	1033
	1.5–2.3	3986	829	1447	706	1004
	2.3–3.5	3094	634	1149	530	781
	3.5–10.0	2569	538	971	427	633

Table 5.2: Numbers of the exclusive events in kinematic bins.

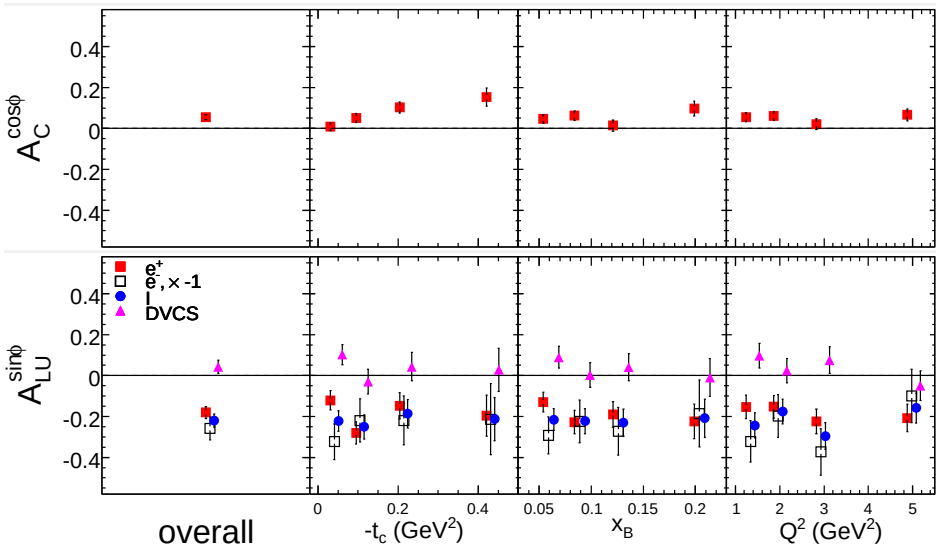

Figure 5.1: Final results for the leading moments. A_{LU}^{-} is scaled by -1.

kinematic bin		$\langle -t_c \rangle$ (GeV^2)	$\langle x_B \rangle$	$\langle Q^2 \rangle$ (GeV^2)	$A_C^{\cos\phi} \pm \delta_{\mathrm{stat}}$	$A_{\mathrm{LU,DVCS}}^{\sin\phi}$ $\pm\delta_{\mathrm{stat}}$	$A_{\mathrm{LU},\mathcal{I}}^{\sin\phi} \pm \delta_{\mathrm{stat}}$
overall		0.1201	0.0954	2.48	0.054 ± 0.012	0.044 ± 0.032	-0.221 ± 0.032
$-t_c$ (GeV^2)	0.00–0.06	0.0313	0.0776	1.97	0.009 ± 0.019	0.102 ± 0.050	-0.222 ± 0.050
	0.06–0.14	0.0953	0.1016	2.56	0.051 ± 0.022	-0.029 ± 0.059	-0.251 ± 0.060
	0.14–0.30	0.2039	0.1108	2.91	0.102 ± 0.027	0.043 ± 0.069	-0.187 ± 0.070
	0.30–0.70	0.4208	0.1213	3.57	0.153 ± 0.044	0.029 ± 0.105	-0.212 ± 0.105
x_B	0.03–0.07	0.1000	0.0539	1.45	0.046 ± 0.021	0.090 ± 0.052	-0.216 ± 0.053
	0.07–0.10	0.1035	0.0836	2.15	0.063 ± 0.023	0.002 ± 0.061	-0.223 ± 0.061
	0.10–0.15	0.1294	0.1208	3.12	0.015 ± 0.027	0.041 ± 0.066	-0.230 ± 0.066
	0.15–0.35	0.1989	0.1987	5.08	0.098 ± 0.036	-0.009 ± 0.092	-0.208 ± 0.093
Q^2 (GeV^2)	1.0–1.5	0.0834	0.0562	1.24	0.055 ± 0.023	0.097 ± 0.060	-0.243 ± 0.061
	1.5–2.3	0.1041	0.0773	1.86	0.062 ± 0.022	0.024 ± 0.060	-0.175 ± 0.060
	2.3–3.5	0.1274	0.1058	2.83	0.021 ± 0.026	0.076 ± 0.065	-0.296 ± 0.065
	3.5–10.0	0.1917	0.1702	4.88	0.066 ± 0.029	-0.050 ± 0.073	-0.158 ± 0.073

Table 5.3: Leading moments of BCA and BSSA extracted from the 00d0 and 05b1 combined data sets. The kinematic variables are averaged over the exclusive events.

kinematic bin		$\langle -t_c \rangle$ (GeV^2)	$\langle x_B \rangle$	$\langle Q^2 \rangle$ (GeV^2)	$A_{LU,e^+}^{\sin\phi} \pm \delta_{stat}$
overall		0.1171	0.0964	2.50	-0.181 ± 0.030
$-t_c$ (GeV^2)	0.00–0.06	0.0315	0.0782	1.98	-0.121 ± 0.047
	0.06–0.14	0.0946	0.1025	2.59	-0.280 ± 0.054
	0.14–0.30	0.2041	0.1128	2.96	-0.148 ± 0.064
	0.30–0.70	0.4192	0.1245	3.64	-0.196 ± 0.100
x_B	0.03–0.07	0.0951	0.0543	1.46	-0.129 ± 0.048
	0.07–0.10	0.1007	0.0837	2.15	-0.227 ± 0.057
	0.10–0.15	0.1283	0.1206	3.12	-0.189 ± 0.062
	0.15–0.35	0.1939	0.1983	5.07	-0.223 ± 0.084
Q^2 (GeV^2)	1.0–1.5	0.0797	0.0566	1.24	-0.153 ± 0.057
	1.5–2.3	0.1010	0.0775	1.87	-0.152 ± 0.055
	2.3–3.5	0.1239	0.1059	2.82	-0.225 ± 0.061
	3.5–10.0	0.1865	0.1699	4.87	-0.207 ± 0.066

Table 5.4: Leading moments of BSA with positrons from the 00d0 data set.

kinematic bin		$\langle -t_c \rangle$ (GeV^2)	$\langle x_B \rangle$	$\langle Q^2 \rangle$ (GeV^2)	$A_{LU,e^-}^{\sin\phi} \pm \delta_{stat}$
overall		0.1240	0.0942	2.45	0.258 ± 0.056
$-t_c$ (GeV^2)	0.00–0.06	0.0310	0.0768	1.94	0.323 ± 0.087
	0.06–0.14	0.0963	0.1003	2.51	0.219 ± 0.105
	0.14–0.30	0.2036	0.1082	2.85	0.219 ± 0.119
	0.30–0.70	0.4226	0.1178	3.49	0.213 ± 0.174
x_B	0.03–0.07	0.1062	0.0535	1.44	0.293 ± 0.090
	0.07–0.10	0.1072	0.0835	2.15	0.224 ± 0.105
	0.10–0.15	0.1309	0.1211	3.12	0.272 ± 0.117
	0.15–0.35	0.2060	0.1994	5.09	0.185 ± 0.163
Q^2 (GeV^2)	1.0–1.5	0.0878	0.0558	1.24	0.322 ± 0.102
	1.5–2.3	0.1081	0.0770	1.86	0.197 ± 0.104
	2.3–3.5	0.1320	0.1057	2.84	0.373 ± 0.113
	3.5–10.0	0.1991	0.1706	4.89	0.100 ± 0.130

Table 5.5: Leading moments of BSA with electrons from the 05b1 data set.

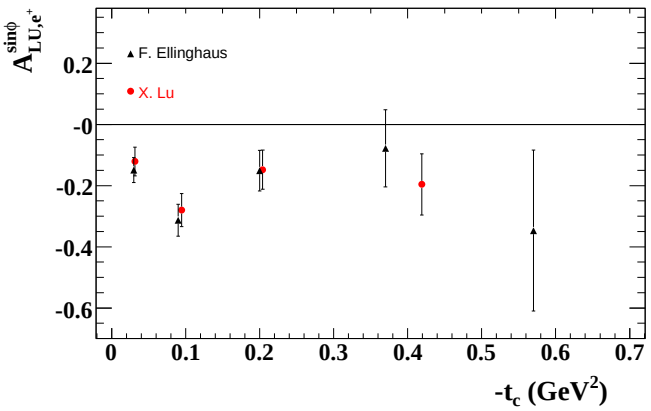

Figure 5.2: Comparison for $A_{\rm LU}^{+}$ between the current result and the one in [Ell03].

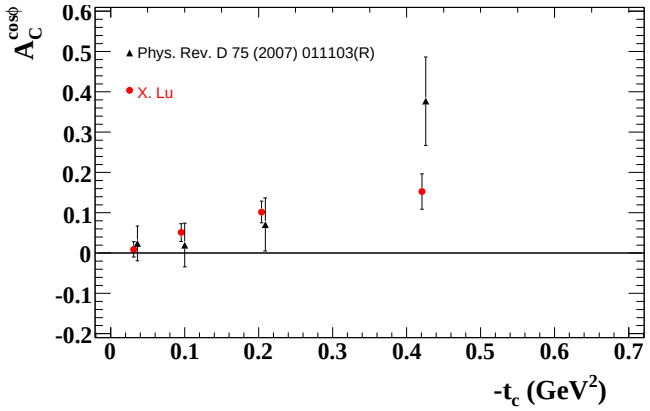

Figure 5.3: Comparison for $A_{\rm C}$ between the current result and the one in [Air07].

		00d0 H unp.	05b1 H unp.	05b1 H pol.
N_{DIS}	X. Lu	5,043,469	307,236	3,364,537
	T. Keri	5,043,469	307,236	3,364,537
	D. Zeiler	5,048K	308K	3,364K
$N_{Excl.}$	X. Lu	7,709	524	5,295
	T. Keri	7,709	524	5,295
	D. Zeiler	7,709	524	5,295

Table 5.6: Cross check results for the numbers of the DIS and exclusive events.

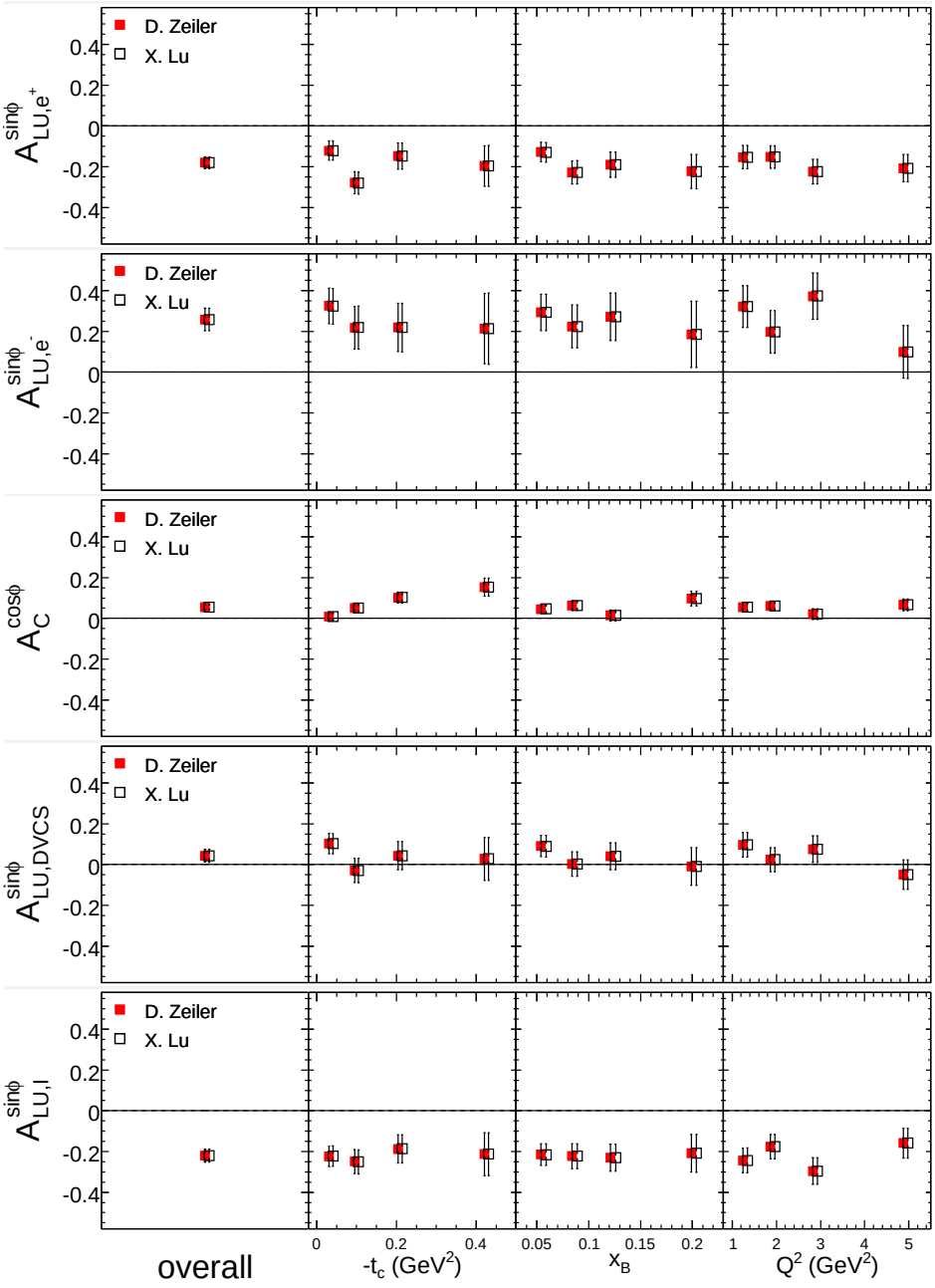

Figure 5.4: Cross check results for the leading moments.

Chapter 6

Summary and Outlook

In this thesis, the beam-charge and beam-spin asymmetries, A_{C} and A_{LU}^{η}, are extracted. With both the positron and electron data, the beam-spin sub-asymmetries $A_{\mathrm{LU}}^{\mathrm{DVCS}}$ and $A_{\mathrm{LU}}^{\mathcal{I}}$, which are induced by the DVCS and DVCS-BH interference contributions respectively, are further separated from the beam-spin asymmetry for the first time, and that $A_{\mathrm{LU}}^{\mathrm{DVCS}} \approx 0$, $A_{\mathrm{LU}}^{\eta} \approx \eta A_{\mathrm{LU}}^{\mathcal{I}}$ is observed.

These results are obtained in HERMES acceptance. In future systematic studies, influences from various sources will be quantified.

By comparing the asymmetries to theoretical predictions based on phenomenological models of GPDs, one can obtain model-dependent constraints on the total angular momentum of quarks in the nucleon [Ye06]. In future studies, all then available models will be used to compare to the data in order to help constraining GPDs within the models.

Appendix A

Basic Principles of Maximum Likelihood Estimation

A.1 Definitions

1. (a) $\underline{\theta} = (\theta_1,\ \theta_2, \ldots,\ \theta_M)^T$: M parameters in a vector form, characterizing a population. T denotes transpose.

   (b) $\underline{\mathbf{\Theta}}$: the permissible ranges for $\underline{\theta}$.

   (c) $\underline{\theta}^*$: the true values of $\underline{\theta}$, $\underline{\theta}^* \in \underline{\mathbf{\Theta}}$.

   (d) x: observable from the population, can be a vector with multi-components.

   (e) $\mathbf{X}$: the domain of x, independent of $\underline{\theta}$.

   (f) $x_1,\ x_2, \ldots,\ x_N$: a sample of x from N independent observations.

   (g) $\hat{\underline{\theta}}(x_1,\ x_2, \ldots,\ x_N)$: estimators for $\underline{\theta}$.

   (h) $p(x;\underline{\theta})$: probability density function (pdf) describing the normalized distribution of x,

   $$\int_{\mathbf{X}} p(x;\underline{\theta})\mathrm{d}x = 1, \forall \underline{\theta} \in \underline{\mathbf{\Theta}}. \tag{A.1}$$

   $\underline{\theta}$ describe the shape of the distribution.

   (i) $\langle f(x;\underline{\theta})\rangle_{\underline{\theta}'}$: the expectation value of any function of the observable, $f(x;\underline{\theta})$,

   $$\langle f(x;\underline{\theta})\rangle_{\underline{\theta}'} \equiv \int_{\mathbf{X}} p(x;\underline{\theta}')f(x;\underline{\theta})\mathrm{d}x, \tag{A.2}$$

   given a certain pdf with $\underline{\theta}'$.

   (j) $\underline{\underline{D}}(\underline{\theta};\underline{\theta}')$:

   $$\underline{\underline{D}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')\frac{\partial \ln p(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.3}$$

   $$= \left(\left\langle \frac{\partial \ln p(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.4}$$

   assumed to be positive definite, i.e., $p(x;\underline{\theta})$ is so parameterized that functions $\frac{\partial \ln p(x;\underline{\theta})}{\partial\theta_i}$, $i = 1,\ 2, \ldots,\ M$, are linearly independent in $\mathbf{X}$ [1].

[1] $\underline{\underline{D}}$ can be written in a more compact form,

$$\underline{\underline{D}} = \left\langle \underline{d}\ \underline{d}^T\right\rangle_{\underline{\theta}'},\ \underline{d}(x;\underline{\theta}) \equiv \frac{\partial}{\partial\underline{\theta}}\ln p. \tag{A.5}$$

(k) $\langle f(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})\rangle_{\underline{\theta}'}$: the expectation value of any function of the sample, $f(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})$,

$$\begin{aligned}&\langle f(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})\rangle_{\underline{\theta}'}\\ &\equiv\int\cdots\int_{\mathbf{X}^N} p(x_1;\underline{\theta}')p(x_2;\underline{\theta}')\cdots p(x_N;\underline{\theta}')f(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N,\end{aligned} \tag{A.13}$$

given a certain joint pdf with $\underline{\theta}'$.

(l) $L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})$: Likelihood function (LF), giving the hypothetical probability for outcomes $x_1,\ x_2,\ldots,\ x_N$, *assumed* to be third-order differentiable in $\underline{\boldsymbol{\Theta}}$ without terminal maxima.

(m) $l(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})$: log-likelihood function,

$$l(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})\equiv\ln L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta}), \tag{A.14}$$

omitting terms that are independent of $\underline{\theta}$.

(n) the Standard Likelihood Function (SLF):

$$L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})=\prod_{i=1}^{N}p(x_i;\underline{\theta}). \tag{A.15}$$

2. (a) $w(x)$: weight function, $w(x)>0,\ \forall x\in\mathbf{X}$.

For any non-zero number vector $\underline{v}$,

$$\underline{v}^T\ \underline{\underline{D}}\ \underline{v}=\underline{v}^T\left\langle\underline{d}\ \underline{d}^T\right\rangle_{\underline{\theta}'}\ \underline{v} \tag{A.6}$$

$$=\left\langle\underline{v}^T\ \underline{d}\ \underline{d}^T\ \underline{v}\right\rangle_{\underline{\theta}'} \tag{A.7}$$

$$=\left\langle(\underline{v}\cdot\underline{d})^2\right\rangle_{\underline{\theta}'} \tag{A.8}$$

$$\geq 0. \tag{A.9}$$

It follows that if $\underline{\underline{D}}$ is positive definite, " = " in Eq. (A.9) must not hold.

$$\underline{v}^T\ \underline{\underline{D}}\ \underline{v}=0\Leftrightarrow\underline{v}\cdot\underline{d}=0,\ \forall x\in\mathbf{X} \tag{A.10}$$

$$\Leftrightarrow\ \text{functions } d_i=\frac{\partial\ln p}{\partial\theta_i},\ i=1,\ 2,\ldots,\ M \text{ are linearly dependent in } \mathbf{X}. \tag{A.11}$$

Hence we come to the conclusion that $\underline{\underline{D}}$ being positive definite requires $\frac{\partial\ln p}{\partial\theta_i}$ being linearly independent in $\mathbf{X}$.

Here is a simple example. Suppose that $p(x;\theta_1,\ \theta_2)=p(x;\theta_1+k\theta_2)$ (k is a number independent of x and $\underline{\theta}$), which means the dependence of θ_1 and θ_2 degenerates into the dependence of $\theta_1+k\theta_2$ as a whole. We can see that $\frac{\partial\ln p}{\partial\theta_1}$ and $\frac{\partial\ln p}{\partial\theta_2}$ are linearly dependent,

$$\frac{\partial\ln p}{\partial\theta_2}=k\frac{\partial\ln p}{\partial\theta_1},\ \forall x\in\mathbf{X}. \tag{A.12}$$

(b) $p_w(x;\underline{\theta})$: weighted pdf,

$$p_w(x;\underline{\theta}) \equiv \frac{w(x)p(x;\underline{\theta})}{\int_{\mathbf{X}} w(x)p(x;\underline{\theta})\mathrm{d}x}, \tag{A.16}$$

which implies

$$\int_{\mathbf{X}} p_w(x;\underline{\theta})\mathrm{d}x = 1. \tag{A.17}$$

(c) $\underline{\underline{D_{w1}}}(\underline{\theta};\underline{\theta}')$, $\underline{\underline{D_{w2}}}(\underline{\theta};\underline{\theta}')$:

$$\underline{\underline{D_{w1}}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')w(x)\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.18}$$

$$= \left(\left\langle w(x)\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.19}$$

$$\underline{\underline{D_{w2}}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')w^2(x)\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.20}$$

$$= \left(\left\langle w^2(x)\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln p_w(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.21}$$

both positive definite from Definition 2a and Definition 1j.

(d) the Weighted Standard Likelihood Function (WSLF):

$$L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta}) = \prod_{i=1}^{N}\left[p_w(x_i;\underline{\theta})\right]^{w(x_i)}. \tag{A.22}$$

3. (a) $P(x;\underline{\theta})$, $\mathbb{N}(\underline{\theta})$: the unnormalized distribution function of x and its normalization factor,

$$\int_{\mathbf{X}} P(x;\underline{\theta})\mathrm{d}x = \mathbb{N}(\underline{\theta}). \tag{A.23}$$

$\underline{\theta}$ describe the size and shape of the distribution.

(b) $\underline{\underline{D_e}}(\underline{\theta};\underline{\theta}')$:

$$\underline{\underline{D_e}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.24}$$

$$= \left(\left\langle \frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.25}$$

assumed to be positive definite, i.e., $P(x;\underline{\theta})$ is so parameterized that functions $\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}$, $i = 1,\ 2,\ldots,\ M$, are linearly independent in $\mathbf{X}$.

(c) the Extended Likelihood Function (ELF):

$$L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta}) = \frac{e^{-\mathbb{N}(\underline{\theta})}[\mathbb{N}(\underline{\theta})]^N}{N!}\prod_{i=1}^{N}\frac{P(x_i;\underline{\theta})}{\mathbb{N}(\underline{\theta})}. \tag{A.26}$$

The Poisson pdf arises from the *assumption* that the expected number of events is $\mathbb{N}(\underline{\theta})$ and the observed one has a Poisson fluctuation about it.

4. (a) $\mathbb{N}_w(\underline{\theta})$: weighted normalization factor,

$$\mathbb{N}_w(\underline{\theta}) \equiv \int_{\mathbf{X}} w(x)P(x;\underline{\theta})\mathrm{d}x. \tag{A.27}$$

(b) $\underline{\underline{D_{ew1}}}(\underline{\theta};\underline{\theta}')$, $\underline{\underline{D_{ew2}}}(\underline{\theta};\underline{\theta}')$:

$$\underline{\underline{D_{ew1}}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')w(x)\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.28}$$

$$= \left(\left\langle w(x)\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.29}$$

$$\underline{\underline{D_{ew2}}}(\underline{\theta};\underline{\theta}') \equiv \left(\int_{\mathbf{X}} p(x;\underline{\theta}')w^2(x)\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\mathrm{d}x\right)_{M\times M} \tag{A.30}$$

$$= \left(\left\langle w^2(x)\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_i}\frac{\partial \ln P(x;\underline{\theta})}{\partial\theta_j}\right\rangle_{\underline{\theta}'}\right)_{M\times M}, \tag{A.31}$$

both positive definite from Definition 2a and Definition 3b.

(c) the Weighted Extended Likelihood Function (WELF):

$$L(x_1,\ x_2,\ldots,\ x_N;\underline{\theta}) = \frac{e^{-\mathbb{N}_w(\underline{\theta})}[\mathbb{N}_w(\underline{\theta})]^{N_w}}{N_w!}\prod_{i=1}^{N}\left[\frac{w(x_i)P(x_i;\underline{\theta})}{\mathbb{N}_w(\underline{\theta})}\right]^{w(x_i)}, \tag{A.32}$$

in which $N_w = \sum_{i=1}^{N} w(x_i)$.

Conventions:

- Functions with usual variable lists, e.g. in the form of $f(x)$, $f(x_1,\ x_2,\ldots,\ x_N)$, $f(\underline{\theta})$, $f(x;\underline{\theta})$ or $f(x_1,\ x_2,\ldots,\ x_N;\underline{\theta})$, are denoted as f.
- Dependence of x_i is denoted by a subscript $_i$, e.g. $p_i \equiv p(x_i;\underline{\theta})$.
- Function at $\underline{\theta}^*$ are denoted by *, e.g. $p^* \equiv p(x;\underline{\theta}^*)$, $\langle\rangle_* \equiv \langle\rangle_{\underline{\theta}^*}$, $|_* \equiv |_{\underline{\theta}^*}$ but *no simplification for* $\underline{\underline{D}}(\underline{\theta};\underline{\theta}^*)$ *or* $\underline{\underline{D}}(\underline{\theta}^*;\underline{\theta})$ *except* $\underline{\underline{D}}^* \equiv \underline{\underline{D}}(\underline{\theta}^*;\underline{\theta}^*)$.
- $\frac{\partial}{\partial\underline{\theta}} \equiv (\frac{\partial}{\partial\theta_1},\ \frac{\partial}{\partial\theta_2},\ldots,\ \frac{\partial}{\partial\theta_M})^T$.

A.2 Properties of Likelihood Functions

Lemma A.2.0.1.

The law of large numbers (LLN),

$$\frac{1}{N}\sum_{i=1}^{N} f(x_i;\underline{\theta}) \xrightarrow{P} \langle f(x;\underline{\theta})\rangle_{\underline{\theta}^*}\ ,\ as\ N\to\infty, \tag{A.33}$$

where $\xrightarrow{P}$ *means converging in probability.*

Lemma A.2.0.2.

$\forall\underline{\theta}'\in\underline{\mathbf{\Theta}}$*, the following relations hold:*

$$\left\langle w\frac{\partial \ln p_w}{\partial\underline{\theta}}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'} = 0, \tag{A.34}$$

$$\left\langle w\frac{\partial^2 \ln p_w}{\partial\theta_i\partial\theta_j}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'} = -\left\langle w\frac{\partial \ln p_w}{\partial\theta_i}\frac{\partial \ln p_w}{\partial\theta_j}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'},\ \forall\ i,\ j, \tag{A.35}$$

$$N\left\langle w\frac{\partial \ln P}{\partial\underline{\theta}}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'} - \frac{\partial \mathbb{N}_w}{\partial\underline{\theta}}\bigg|_{\underline{\theta}'} \xrightarrow{P} 0,\ as\ \mathbb{N}(\underline{\theta}')\to\infty, \tag{A.36}$$

$$N\left\langle w\frac{\partial^2 \ln P}{\partial\theta_i\partial\theta_j}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'} - \frac{\partial^2 \mathbb{N}_w}{\partial\theta_i\partial\theta_j}\bigg|_{\underline{\theta}'} \xrightarrow{P} -N\left\langle w\frac{\partial \ln P}{\partial\theta_i}\frac{\partial \ln P}{\partial\theta_j}\bigg|_{\underline{\theta}'}\right\rangle_{\underline{\theta}'},\ \forall\ i,\ j,\ as\ \mathbb{N}(\underline{\theta}')\to\infty. \tag{A.37}$$

Proof. of Eq. (A.36)

$$l.h.s. = \left[\frac{N}{\mathbb{N}(\underline{\theta}')} - 1\right]\frac{\partial \mathbb{N}_w}{\partial\underline{\theta}}\bigg|_{\underline{\theta}'}. \tag{A.38}$$

Since N is Poisson-distributed with expectation value $\mathbb{N}$, in the large $\mathbb{N}$ limit, $\frac{N}{\mathbb{N}} - 1 \xrightarrow{P} 0$. [2] □

[2] In the large $\mathbb{N}$ limit, the Poisson pdf $\frac{e^{-\mathbb{N}}\mathbb{N}^N}{N!}$ approaches a Gaussian pdf with expectation value $\mathbb{N}$ and variance $\mathbb{N}$,

$$C\exp\left[-\frac{1}{2}\frac{(N-\mathbb{N})^2}{\mathbb{N}}\right], \tag{A.39}$$

where C is the normalization factor. Thus $\frac{N}{\mathbb{N}} - 1$ is normally distributed with expectation value 0 and variance $1/\mathbb{N}$,

$$C'\exp\left[-\frac{1}{2}\frac{(\frac{N}{\mathbb{N}}-1)^2}{\frac{1}{\mathbb{N}}}\right], \tag{A.40}$$

where C' is the normalization factor. Using Chebyshev inequality, we have

$$P\left(\left|\frac{N}{\mathbb{N}} - 1\right| < \epsilon\right) \geq 1 - \frac{\frac{1}{\mathbb{N}}}{\epsilon^2}, \tag{A.41}$$

Lemma A.2.0.3.
If $\hat{\underline{\theta}}$ satisfy the following conditions:

1. *consistency, i.e. $\hat{\underline{\theta}} \xrightarrow{P} \underline{\theta}^*$, as $N \to \infty$;*
2. $\left.\frac{\partial l}{\partial \underline{\theta}}\right|_{\hat{\underline{\theta}}} = 0$,

the following matrix relation for $\hat{\underline{\theta}}$ holds in the large N limit,

$$\left(-\left.\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right|_*\right)\left((\hat{\theta}_i-\theta_i^*)(\hat{\theta}_j-\theta_j^*)\right)\left(-\left.\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right|_*\right)=\left(\left.\frac{\partial l}{\partial\theta_i}\frac{\partial l}{\partial\theta_j}\right|_*\right). \tag{A.44}$$

Proof. Using Taylor's theorem, we have

$$l = l^* + \sum_{i=1}^{M}\left.\frac{\partial l}{\partial\theta_i}\right|_*(\theta_i-\theta_i^*) + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M}\left.\frac{\partial^2 l}{\partial\theta_j\partial\theta_k}\right|_{\underline{\theta}^\triangle}(\theta_j-\theta_j^*)(\theta_k-\theta_k^*), \tag{A.45}$$

where $\underline{\theta}^\triangle$ are some values between $\underline{\theta}$ and $\underline{\theta}^*$. And then from Condition 2,

$$\left.\frac{\partial l}{\partial\theta_i}\right|_{\hat{\underline{\theta}}} = \left.\frac{\partial l}{\partial\theta_i}\right|_* + \sum_{j=1}^{M}\left.\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right|_{\underline{\theta}^\triangle}(\hat{\theta}_j-\theta_j^*) = 0, \tag{A.46}$$

which can be written as

$$\underline{\underline{A_2}}^\triangle(\hat{\underline{\theta}}-\underline{\theta}^*) = \left.\frac{\partial l}{\partial\underline{\theta}}\right|_*, \tag{A.47}$$

where $\underline{\underline{A_2}}(x_1,\, x_2, \ldots,\, x_N;\underline{\theta}) \equiv \left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right)_{M\times M}$ and the dependence of $\underline{\theta}^\triangle$ is denoted by $^\triangle$. Note that $\hat{\underline{\theta}}-\underline{\theta}^*$ are the biases of $\hat{\underline{\theta}}$. Multiplying Eq. (A.47) by its transpose, we have

$$\underline{\underline{A_2}}^\triangle\,\underline{\underline{B}}\,\underline{\underline{A_2}}^\triangle = \underline{\underline{A_1}}^*, \tag{A.48}$$

where $\underline{\underline{B}}(\hat{\underline{\theta}}) \equiv \left((\hat{\theta}_i-\theta_i^*)(\hat{\theta}_j-\theta_j^*)\right)_{M\times M}$ and $\underline{\underline{A_1}}(x_1,\, x_2, \ldots,\, x_N;\underline{\theta}) \equiv \left(\frac{\partial l}{\partial\theta_i}\frac{\partial l}{\partial\theta_j}\right)_{M\times M}$. From Condition 1, $\underline{\theta}^\triangle \xrightarrow{P} \underline{\theta}^*$ as $N\to\infty$. Thus in the large N limit Eq. (A.48) becomes

$$\underline{\underline{A_2}}^*\,\underline{\underline{B}}\,\underline{\underline{A_2}}^* = \underline{\underline{A_1}}^*, \tag{A.49}$$

---

where ϵ is an arbitrary positive number. As $\mathbb{N}\to\infty$,

$$P\left(\left|\frac{N}{\mathbb{N}}-1\right|<\epsilon\right)=1, \tag{A.42}$$

i.e.

$$\frac{N}{\mathbb{N}}-1 \xrightarrow{P} 0. \tag{A.43}$$

with

$$\underline{\underline{B}} = \left((\hat{\theta}_i - \theta_i^*)(\hat{\theta}_j - \theta_j^*)\right)_{M\times M}, \tag{A.50}$$

$$\underline{\underline{A_1}} = \left(\frac{\partial l}{\partial \theta_i}\frac{\partial l}{\partial \theta_j}\right)_{M\times M}, \tag{A.51}$$

$$\underline{\underline{A_2}} = \left(-\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}\right)_{M\times M}. \tag{A.52}$$

□

Theorem A.2.0.1.

For WSLF, L^* is a local maximum in the large N limit.

Proof.

$$L = \prod_{i=1}^{N} p_{wi}^{w_i}, \tag{A.53}$$

$$\ln L = \sum_{i=1}^{N} w_i \ln p_{wi} \tag{A.54}$$

$$\xrightarrow{P} N \langle w \ln p_w \rangle_*, \text{ by LLN}, \tag{A.55}$$

$$\left.\frac{\partial \ln L}{\partial \theta_i}\right|_* \xrightarrow{P} N \left\langle w \left.\frac{\partial \ln p_w}{\partial \theta_i}\right|_* \right\rangle_* \tag{A.56}$$

$$= 0, \ \forall\, i, \text{ from Eq. (A.34)}. \tag{A.57}$$

$$\therefore \ln L^* \text{ is an extreme.} \tag{A.58}$$

$$\left.\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right|_* \xrightarrow{P} N \left\langle w \left.\frac{\partial^2 \ln p_w}{\partial \theta_i \partial \theta_j}\right|_* \right\rangle_* \tag{A.59}$$

$$= -N \left\langle w \frac{\partial \ln p_w}{\partial \theta_i}\frac{\partial \ln p_w}{\partial \theta_j}\right|_* \Bigg\rangle_*, \ \forall\, i,\ j, \text{ from Eq. (A.35)}. \tag{A.60}$$

$$\therefore \left(\left.\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right|_*\right)_{M\times M} \xrightarrow{P} -N \underline{\underline{D_{w1}}}^*, \text{ from Eq. (A.19)}. \tag{A.61}$$

Because $\underline{\underline{D_{w1}}}(\underline{\theta};\underline{\theta}')$ is positive definite (cf. Definition 2c), the Hessian matrix of $\ln L$, i.e. $\left(\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right)_{M\times M}$, is negative definite at $\underline{\theta}^*$. Thus $\ln L^*$ is a local maximum of $\ln L$ in the large N limit. Since $\ln L$ and L are both maximized at the same $\underline{\theta}$, the theorem is proved. □

Corollary A.2.0.1.

For SLF, L^* is a local maximum in the large N limit.

As a special case of WSLF with unity weight, i.e. $w(x) \equiv 1$, this statement holds as a corollary of Theorem A.2.0.1.

Theorem A.2.0.2.
For WELF, L^* is a local maximum in the large $\mathbb{N}$ limit.

Corollary A.2.0.2.
For ELF, L^* is a local maximum in the large $\mathbb{N}$ limit.

A.3 Maximum Likelihood Estimators

By derivative tests we only know about the local properties of L. For simplicity, here L is *assumed* to possess a single extreme in $\underline{\boldsymbol{\Theta}}$ regardless of its form. Then the extreme point can be taken as the general maximum likelihood (ML) estimators ($\underline{\hat{\theta}}_{\mathrm{ML}}$) for the four LFs, which means

$$\left.\frac{\partial L}{\partial \underline{\theta}}\right|_{\underline{\hat{\theta}}_{\mathrm{ML}}} = 0, \tag{A.62}$$

or equivalently,

$$\left.\frac{\partial \ln L}{\partial \underline{\theta}}\right|_{\underline{\hat{\theta}}_{\mathrm{ML}}} = 0. \tag{A.63}$$

In the Proof of Theorem A.2.0.1 and A.2.0.2, there is no constraint imposed on $w(x)$. Thus it is interesting to note that assigning arbitrary weights to the data still preserves the consistency of the ML estimators (cf. Section A.4). However, as is shown below, the covariance matrix depends on the weights.

Theorem A.3.0.3.
For WSLF, the covariance matrix of the ML estimators $\underline{\hat{\theta}}_{\mathrm{WSML}}$ reads

$$\underline{\underline{V}} = \frac{1}{N} \underline{\underline{D_{w1}}}^{*-1} \, \underline{\underline{D_{w2}}}^{*} \, \underline{\underline{D_{w1}}}^{*-1}, \tag{A.64}$$

which is given by the product of matrixes in the large N limit,

$$\left(-\frac{\partial^2 l}{\partial\theta_i \partial\theta_j}\bigg|_{\underline{\hat{\theta}}_{\mathrm{WSML}}}\right)^{-1} \left(\sum_{k=1}^{N} w_k^2 \frac{\partial \ln p_{wk}}{\partial\theta_i} \frac{\partial \ln p_{wk}}{\partial\theta_j}\bigg|_{\underline{\hat{\theta}}_{\mathrm{WSML}}}\right) \left(-\frac{\partial^2 l}{\partial\theta_i \partial\theta_j}\bigg|_{\underline{\hat{\theta}}_{\mathrm{WSML}}}\right)^{-1}, \tag{A.65}$$

where

$$l = \sum_{k=1}^{N} w_k \ln p_{wk}. \tag{A.66}$$

Proof. For WSLF, $l = \sum_{i=1}^{N} w_i \ln p_{wi}$. The ML estimators satisfy the conditions of Lemma A.2.0.3, so in Eq. (A.49)

$$\underline{\underline{A_1}}^* = \left(\sum_{k=1}^{N} \sum_{m=1}^{N} w_k w_m \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wm}}{\partial \theta_j} \bigg|_* \right) \tag{A.67}$$

$$= \left(\sum_{k=1}^{N} w_k^2 \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wk}}{\partial \theta_j} \bigg|_* \right) + \left(\sum_{k=1}^{N} \sum_{m \neq k}^{N} w_k w_m \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wm}}{\partial \theta_j} \bigg|_* \right), \tag{A.68}$$

with expectation value

$$\left\langle \underline{\underline{A_1}}^* \right\rangle_* = \left(\left\langle \sum_{k=1}^{N} w_k^2 \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wk}}{\partial \theta_j} \bigg|_* \right\rangle_* \right) + \left(\left\langle \sum_{k=1}^{N} \sum_{m \neq k}^{N} w_k w_m \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wm}}{\partial \theta_j} \bigg|_* \right\rangle_* \right) \tag{A.69}$$

$$= \left(N \left\langle w^2 \frac{\partial \ln p_w}{\partial \theta_i} \frac{\partial \ln p_w}{\partial \theta_j} \bigg|_* \right\rangle_* \right) + 0 \tag{A.70}$$

$$= N \underline{\underline{D_{w2}}}^*. \tag{A.71}$$

And

$$\underline{\underline{A_2}}^* = \left(- \sum_{k=1}^{N} w_k \frac{\partial^2 \ln p_{wk}}{\partial \theta_i \partial \theta_j} \bigg|_* \right) \tag{A.72}$$

$$\xrightarrow{P} \left(-N \left\langle w \frac{\partial^2 \ln p_w}{\partial \theta_i \partial \theta_j} \bigg|_* \right\rangle_* \right) \tag{A.73}$$

$$= \left(N \left\langle w \frac{\partial \ln p_w}{\partial \theta_i} \frac{\partial \ln p_w}{\partial \theta_j} \bigg|_* \right\rangle_* \right) \tag{A.74}$$

$$= N \underline{\underline{D_{w1}}}^*, \tag{A.75}$$

which is a non-singular number matrix. Thus the covariance matrix of the ML estimators reads

$$\underline{\underline{V}} \equiv \left(\left\langle (\hat{\theta}_i - \theta_i^*)(\hat{\theta}_j - \theta_j^*) \right\rangle_* \right) \tag{A.76}$$

$$= \left\langle \underline{\underline{B}} \right\rangle_* \tag{A.77}$$

$$= \underline{\underline{A_2}}^{*-1} \left\langle \underline{\underline{A_1}}^* \right\rangle_* \underline{\underline{A_2}}^{*-1} \tag{A.78}$$

$$= \frac{1}{N} \underline{\underline{D_{w1}}}^{*-1} \underline{\underline{D_{w2}}}^* \underline{\underline{D_{w1}}}^{*-1}. \tag{A.79}$$

□

Corollary A.3.0.3.

For SLF, the covariance matrix of the ML estimators $\underline{\hat{\theta}}_{\mathrm{SML}}$ *reads*

$$\underline{\underline{V}} = \frac{1}{N} \underline{\underline{D}}^{*-1}, \tag{A.80}$$

which is given asymptotically by

$$\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{SML}}}\right)^{-1}, \tag{A.81}$$

where

$$l=\sum_{k=1}^{N}\ln p_k. \tag{A.82}$$

Theorem A.3.0.4.

For WELF, the covariance matrix of the ML estimators $\hat{\underline{\theta}}_{\mathrm{WEML}}$ *reads*

$$\underline{\underline{V}}=\frac{1}{N}\underline{\underline{D_{ew1}}}^{*-1}\,\underline{\underline{D_{ew2}}}^{*}\,\underline{\underline{D_{ew1}}}^{*-1}, \tag{A.83}$$

which is given by the product of matrixes in the large $\mathbb{N}$ *limit,*

$$\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)^{-1}\left(\sum_{k=1}^{N}w_k^2\frac{\partial\ln P_k}{\partial\theta_i}\frac{\partial\ln P_k}{\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)^{-1} \tag{A.84}$$

$$=\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)^{-1}\left(-\frac{\partial^2 l_2}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{WEML}}}\right)^{-1}, \tag{A.85}$$

where

$$l=\sum_{k=1}^{N}w_k\ln P_k-\mathbb{N}_w, \tag{A.86}$$

$$l_2\equiv\sum_{k=1}^{N}w_k^2\ln P_k-\mathbb{N}_{w^2},\ \mathbb{N}_{w^2}\equiv\int_{\mathbf{X}}w^2P\mathrm{d}x. \tag{A.87}$$

Corollary A.3.0.4.

For ELF, the covariance matrix of the ML estimators $\hat{\underline{\theta}}_{\mathrm{EML}}$ *reads*

$$\underline{\underline{V}}=\frac{1}{N}\underline{\underline{D_e}}^{*-1}, \tag{A.88}$$

which is given asymptotically by

$$\left(-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\bigg|_{\hat{\underline{\theta}}_{\mathrm{EML}}}\right)^{-1}, \tag{A.89}$$

where

$$l=\sum_{k=1}^{N}\ln P_k-\mathbb{N}. \tag{A.90}$$

The covariance matrix of WELF in Eq. (A.85) can be obtained by two minimization procedures, one for l and the other for l_2. However, since $\left(\sum_{k=1}^{N} w_k^2 \frac{\partial \ln p_{wk}}{\partial \theta_i} \frac{\partial \ln p_{wk}}{\partial \theta_j}\right)$ depends on w and w^2, the covariance matrix for WSLF in Eq. (A.65) is not so straight forward to obtain.

The estimators of (W)SLF and (W)ELF are consistent. However, the covariance matrixes of the extended estimators are generally different from those of the standard as a consequence of the former taking into account the *assumed* statistical property of the number of events.

A.4 Monte Carlo Test

A Monte Carlo (MC) test is performed to exemplify the statement

> *Assigning arbitrary weights to the data still preserves the consistency of the ML estimators.*

Two samples are generated from pdf

$$p(x; A^*) = \frac{1 + xA^*}{10},\ A^* = 0.1430,\ x \in (-5, 5), \tag{A.91}$$

with the acceptance-rejection method. The sample sizes are $N_1 = 10^3$ and $N_2 = 10^6$ respectively. Define two weight functions in $(-5, 5)$,

$$w_1(x) = 5 + x, \tag{A.92}$$

$$w_2(x) = (5 + x)^2. \tag{A.93}$$

Then we have two weighted pdfs,

$$p_{w1} = \frac{(5 + x)(1 + xA)}{\frac{50}{3}(3 + 5A)}, \tag{A.94}$$

$$p_{w2} = \frac{(5 + x)^2(1 + xA)}{\frac{500}{3}(2 + 5A)}. \tag{A.95}$$

The (weighted) distributions in 100 bins of the two samples are shown in (Fig. A.2) Fig. A.1, where the Least-Square(LS)-fitting results for A and $(N_w \equiv \sum_{i=1}^{N} w_i)$ N are displayed.

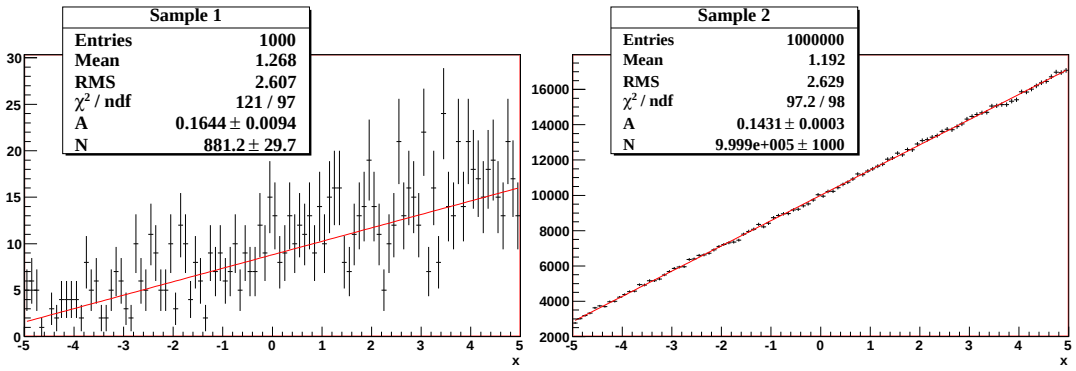

Figure A.1: Two samples from $p(x; A^*) = \frac{1+xA^*}{10}$, $A^* = 0.1430$, $x \in (-5, 5)$. The LS-fitting results for A and N are displayed.

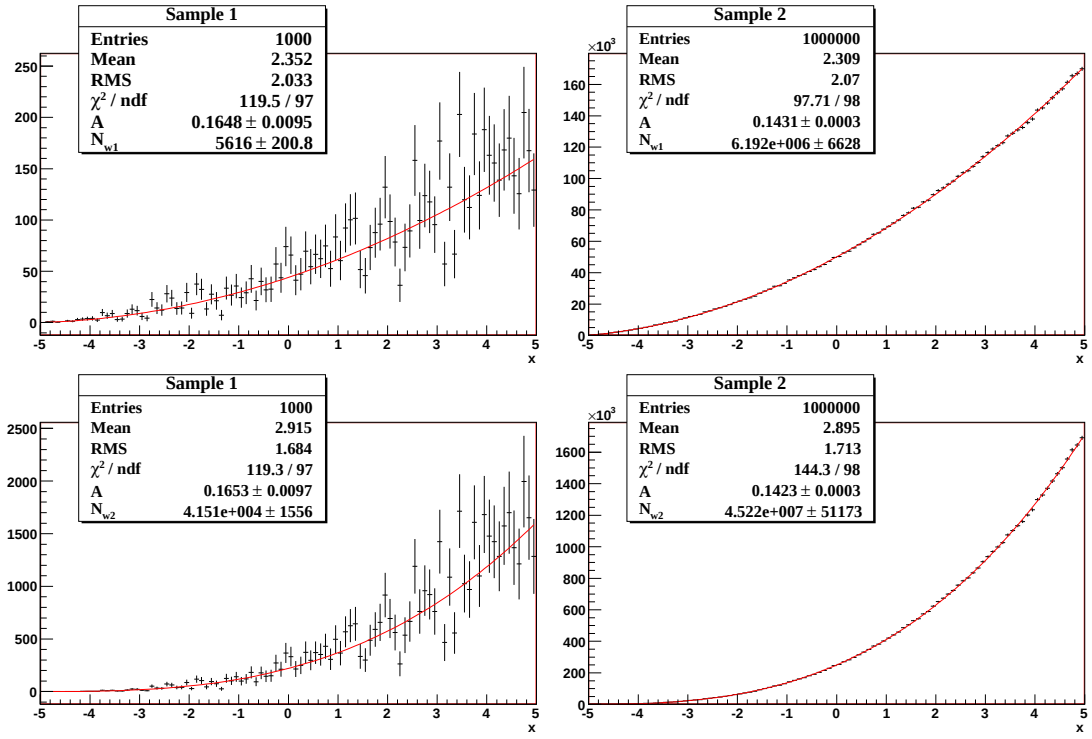

Figure A.2: Weighted (top panels by w_1 and bottom ones by w_2) distributions of the samples. The LS-fitting results for A and $N_{w1,w2}$ are displayed.

The corresponding negative log-SLF and -WSLFs read

$$-l = -\sum_{i=1}^{N} \ln(1 + x_i A), \tag{A.96}$$

$$-l_{w1} = -\sum_{i=1}^{N} (5 + x_i) \ln \frac{1 + x_i A}{3 + 5A}, \tag{A.97}$$

$$-l_{w2} = -\sum_{i=1}^{N} (5 + x_i)^2 \ln \frac{1 + x_i A}{2 + 5A}. \tag{A.98}$$

The minimization procedures of the above functions are performed with ROOT class TMinuit, which takes the minimum point as the central value of A and the inverse of the second derivative as the variance. Results are shown in Table A.1. It can be seen that the three estimators are consistent.

$N = 10^3$	A	σ_A	$N = 10^6$	A	σ_A
l	0.148	0.008	l	0.1431	0.0003
l_{w1}	0.156	0.007	l_{w1}	0.1431	0.0002
l_{w2}	0.155	0.004	l_{w2}	0.1431	0.0001

Table A.1: ML estimation results for the two samples. Note that the variances are given directly by TMinuit, and are incorrect in the case of weighting.

Bibliography

[Ack98] K. Ackerstaff *et al.* [HERMES Collaboration], Nucl. Instrum. Meth. A **417**, 230 (1998) [arXiv:hep-ex/9806008].

[Adl01] C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **517**, 47 (2001) [arXiv:hep-ex/0107005].

[Air01] A. Airapetian *et al.* [HERMES Collaboration], Phys. Rev. Lett. **87**, 182001 (2001) [arXiv:hep-ex/0106068].

[Air07] A. Airapetian *et al.* [HERMES Collaboration], Phys. Rev. D **75**, 011103 (2007) [arXiv:hep-ex/0605108].

[Bar90] R. J. Barlow, Nucl. Instrum. Meth. A **297**, 496 (1990).

[Bel02] A. V. Belitsky, D. Mueller and A. Kirchner, Nucl. Phys. B **629**, 323 (2002) [arXiv:hep-ph/0112108].

[Bel05] A. V. Belitsky and A. V. Radyushkin, Phys. Rept. **418**, 1 (2005) [arXiv:hep-ph/0504030].

[Die97] M. Diehl, T. Gousset, B. Pire and J. P. Ralston, Phys. Lett. B **411**, 193 (1997) [arXiv:hep-ph/9706344].

[Die03] M. Diehl, Phys. Rept. **388**, 41 (2003) [arXiv:hep-ph/0307382].

[Ell03] F. Ellinghaus, Ph.D thesis, Berlin Humboldt University (2003).

[Jac76] J. D. Jackson, Rev. Mod. Phys. **48**, 417 (1976).

[Ji97] X. D. Ji, Phys. Rev. Lett. **78**, 610 (1997) [arXiv:hep-ph/9603249].

[Ji98] X. D. Ji, J. Phys. G **24**, 1181 (1998) [arXiv:hep-ph/9807358].

[Ken79] M. Kendall *et al.*, The advanced theory of statistics, 4th ed. of Vol. 2 of the 3-volume ed. (1979) ISBN: 0 85264 255 5.

[Yao06] W. M. Yao *et al.* [Particle Data Group], J. Phys. G **33**, 1 (2006).

[Sau99] P. R. B. Saull [ZEUS Collaboration], arXiv:hep-ex/0003030.

[Sol64] F. T. Solmitz, Ann. Rev. Nucl. Part. Sci. **14**, 375 (1964).

[Ste01] S. Stepanyan *et al.* [CLAS Collaboration], Phys. Rev. Lett. **87**, 182002 (2001) [arXiv:hep-ex/0107043].

[Ye06] Z. Ye [HERMES Collaboration], PoS **HEP2005**, 120 (2006) [arXiv:hep-ex/0512010].

Acknowledgements

I would like to thank my supervisor Prof. Yajun Mao at Peking University. He provided me with an excellent working environment and important opportunities. With his strict requirements, I initially grasped the basic skills for data analysis. I would like to thank my supervisor Wolf-Dieter Nowak at HERMES for his kind support and guidance on my current work, and also for his thoughtful arrangement for my future study. Special thanks to my direct advisor Zhenyu Ye, from whom I gained valuable knowledge about DVCS analysis and received beneficial suggestions about life and research.

I would like to thank Hayg Guler, Tibor Keri, Andy Miller, Caroline Riedl, Dietmar Zeiler and all other HERMES members for helping me with the DVCS analysis. I would like to thank Xiaorui Lu, Hongxue Ye and Weilin Yu again for their help with my life in Germany, and also for those many interesting and helpful discussions.

I would like to thank the HEP Group members at Peking University. Special thanks to Ran Han, Yutie Liang, Ruizhe Yang and Peng Zhou for their valuable help since I joined this group.

I owe many thanks to my family for their understanding and support throughout these years.