# Transverse Target Moments of Dihadron Production in Semi-inclusive Deep Inelastic Scattering at HERMES

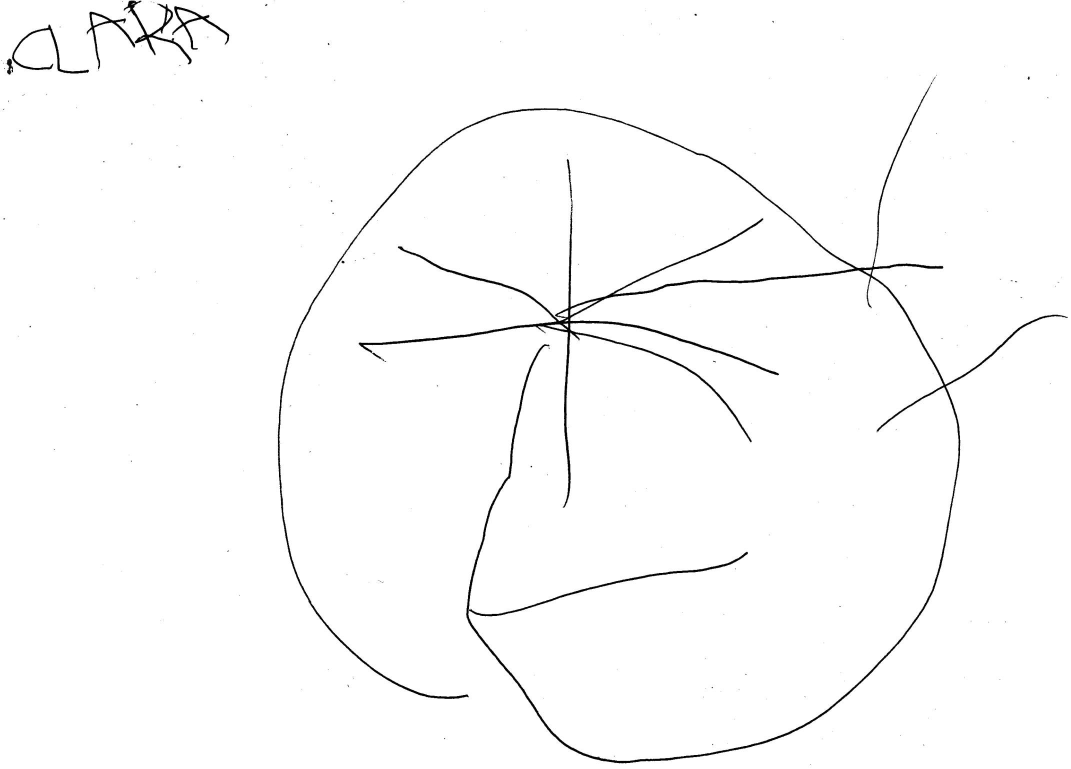
by

Stephen V. Gliske

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2011

Doctoral Committee:

Professor Wolfgang B. Lorenzon, Chair
Professor David W. Gerdes
Associate Professor Leopoldo A. Pando Zayas
Assistant Professor Clayton D. Scott



"*This* is what the Collin's function looks like." Clara Gliske, Spring 2010

For their own inspirational pursuit of knowledge and infallible determination,
I dedicate this dissertation to my dearest wife, Tamie,
and to our three daughters, Clara, Lily and Rebekah.

# ABSTRACT

Pseudo-scalar meson production in semi-inclusive deep inelastic scattering (SIDIS) at HERMES has provided essential information towards the understanding of the transverse momentum dependent structure of the proton. SIDIS dihadron (hadron pair) production also provides access to the structure of the proton and is complimentary to that provided by pseudo-scalars production, as the same parton distribution functions are involved. For example, while pion and kaon final states allow access to flavor combinations of the Sivers distribution function, SIDIS $\phi$ meson production (included in the $K^+K^-$ dihadron sample) allows direct access to the Sivers function for the strange quarks. The Sivers function for strange quarks is also related to the orbital angular momentum of the gluons. In the SIDIS cross section, the distribution functions are integrated with fragmentation functions for the respective final states. These fragmentation functions yield information regarding the quark hadronization process. Of particular interest, the Lund/Artru model of fragmentation makes specific predictions regarding the relation between results for dihadron and pseudo-scalar meson production for certain transverse momentum dependent moments. This dissertation presents the first transverse momentum dependent (non-collinear) analysis of the transverse target moments in SIDIS dihadron production, extracting results from the 2002-2005 HERMES data set for $\pi^+\pi^0$, $\pi^+\pi^-$, $\pi^-\pi^0$ and $K^+K^-$ dihadrons. A new transverse momentum dependent Monte Carlo generator, `TMDGen`, is also introduced. Additionally, several theoretical developments have been completed, including a new partial wave analysis of the fragmentation functions, computation of the next-to-leading twist dihadron cross section, and the first model calculation for transverse momentum dependent dihadron fragmentation functions.

# ACKNOWLEDGEMENTS

First and foremost, I would like to thank my adviser, Wolfgang Lorenzon. He allowed me to engage in research with HERMES when I, myself, still did not know what research I planned to eventually pursue. He has supported and guided me as I have chosen my path, and has given me the freedom to explore the aspects of most interest to me. Above all, I thank him for his trust in me, his candid guidance, and for his facilitating the foundation of my future career.

A special thanks must also be given to Alessandro Bacchetta. We became acquainted just a few months before I selected my dissertation topic and found it how closely it related to his. His patience and guidance as I have learned, reworked and reinterpreted a small portion of his research has been remarkable; ownership of ideas has never interfered with collective seeking of further understanding. Alessandro, thank you for your insight, your emails, and our many discussions.

I would also like to thank the many members and supporting staff of the HERMES collaboration. Many came and left before I joined, though what I have done would have been impossible without their earlier contributions. For those whose time at HERMES has intersected with mine, I must also express my thanks. Specifically, I would like to thank Wouter Deconinck and Avetik Airapetian, formerly fellow members of the U. of Michigan HERMES group.They introduced me to HERMES, to DESY, and to Hamburg, and have continued to answer questions and be supportive ever since. I also thank Brian Ball, who later joined the U. of Michigan group as an undergraduate, for our many discussions as we learned from each other. I would like to thank Naomi Makins, for not only introducing me to this topic but also for subsequent discussions. One such talk lead to the current connection between the Lund/Artru model and my partial wave analysis.

Marco Contalbrigo and Gunar Schnell must also be thanked, for their generous and continuous support in various analysis and at various conferences. Many times they have worked late so we could discuss my analyses over Skype. I express my gratitude to Klaus Rith, for his support of my work and his service as current spokesperson for HERMES. On one occasion, after I had presented a particularly “mathy” presentation to the collaboration, I tried to solicit comments from an otherwise unresponsive audience. At this point, Klaus made a comment that I have always remembered, “Let us just say, we trust you.” Thank you Klaus, and all HERMES members, for your trust and allowing me to seek my own, sometimes overly “mathy” way of doing things.

I also thank Richard Glischinski of Grosshansdorf, friend and $n$-th cousin a few times removed. You welcomed me to Germany, to your family and my family, introduced me to Kaszubia, and have been the light on so many otherwise gray and cloudy Hamburg days.

I acknowledge the many teachers and professors that have influenced me, at so many institutions in quite a few different states. Again, I cannot mention all by name. I must thank Janet Sutorius, for her significant influence in my mathematics education. I also express my thanks to William Klink at the University of Iowa. He laid the foundation for my perspective of quantum mechanics and group theory, and his teaching and guidance yet influence me today. I thank Warren Gooch of Truman State University and Larry Fritts of the University of Iowa, two of my music composition professors. Though my field of pursuit has changed, the significance of your influence remains.

Some of the greatest thanks and acknowledgment must be given to my parents. Thank you for your sacrifice, love and prayers. To my father, for giving me the gift of math, of science, of computers and problem solving: if my work goes beyond that which you have done, it is because I stand on the foundation you have built. To my mother, for giving me the gift of music, of literature, and of creativity: though the message I express may be unfamiliar to you, my ability to express it is due to you. And to my step-father, George, for inspiring me with your dedication and enthusiasm, and for allowing me to join your family and being willing to join ours: never have I known one who has followed their dreams with such passion and determination.

Lastly, I must thank those most close to me. I give my thanks to God. Though the nature of our relationship is not physically testable, He has certainly inspired me and encouraged me, and facilitates all good things I have been able to accomplish. I express my most heartfelt thanks my family: the love of my life, Tamie and our three children, Clara, Lily and Rebekah. You have ridden the roller coaster of failed dissertation topics and the uncertainty of graduation, and have supported and put up with me through it all. Clara, I thank you for your passion for learning, Lily, for your joy of life, and Rebekah, for your smiles. I dla Tamie, dziękuję ci za wszystko, bo wszysko z twojego powodu. I po tym czasie, gdy wszystko jest skonczone, jeszcze będziemy my. Ja bardzo kocham ciebie, na zawsza.

# TABLE OF CONTENTS

**CHAPTER**

# LIST OF FIGURES

**Figure**

# LIST OF TABLES

**Table**

# LIST OF APPENDICES

**Appendix**

# CHAPTER I

# Introduction

In the development and understanding of physical processes, one often considers a particular physical circumstance or setting. One of the most remarkable pairings of physical settings and physics concepts is the hydrogen atom and its relation to Quantum Mechanics and Quantum Electrodynamics (QED). Both historically and pedagogically, the hydrogen atom can be considered the key laboratory in the development and understanding of Quantum Mechanics, and hints towards the QED theory. In a very similar manner, the proton is the key laboratory in understanding certain aspects of Quantum Chromodynamics (QCD)—the theory of quarks and gluons and their interactions—and of the parton model—the description of hadrons as bound states of quarks and gluons, denoted partons. Trying to understand the internal structure of the proton in terms of its partonic constituents, including spin and angular orbital momentum degrees of freedom, has pushed our understanding of QCD and the constituent quark model, and may also hint towards physics beyond the standard model.

Experiments regarding the proton's internal structure, as well as experiments regarding the internal structure of the neutrons, deuterons, and other nucleons, raise important questions concerning the present interpretation of QCD. Such questions include how the spin and orbital angular momentum of each parton contribute to the total spin of the nucleon. Initially it was thought that the proton primarily constituted of three quarks, denoted valence quarks, each carrying one third of the total spin. Experimental evidence [1] later suggested that the valence quarks carry only a small fraction of the total spin, resulting in the situation known as the "Spin Crisis" and leading to the commissioning of several new experiments, including HERMES. The quarks were eventually determined to carry about 1/3 of the proton's spin [2, 3, 4]. A current review of the Spin Crisis can be found in Refs. [5, 6].

The HERMES experiment included a polarized gas target and utilized the electron/positron beam of the HERA accelerator. Many HERMES results related to the Spin Crisis involve semi-inclusive deep inelastic scattering, which can be understood as the process of scattering an electron or positron from a nucleus in the target gas, producing additional particles (thus inelastic), and detecting the scattered electron and some, but not all, of the produced particles (thus semi-inclusive). The virtual photon mediating the scattering must have high energy, allowing the individual partons to be probed (hence the deep quantifier) as opposed to interacting with the nucleus as a whole.

Proton structure experiments also address another open question in QCD: the nature of confinement. Confinement is the statement that quarks are never observed as free particles, but rather occur in bound states with other partons. Although confinement is experimentally confirmed, it is not currently known how to derive confinement from the theory of QCD, and it must be independently postulated. Any quark which might be observed as a free particle immediately undergoes a process resulting in the quark being in a bound state. Two analogous titles are given this process: hadronization, as bound states of quarks are denoted hadrons, and fragmentation, as the quark is said to "fragment" into the observed hadrons.

The hadronization process, though not specifically a feature of proton structure, is yet related to proton structure. Hadronization concerns how hadrons are initially formed, and thus addresses how the structure of a given hadron, such a proton, is created. More practically, proton structure experiments involve interacting with the partons, often removing quarks from the proton. Thus the measured results are the integration of features of both the proton's structure, i.e. the state of the quark before being struck, and the hadronization process, i.e. how the state of the measured particles relate to the state of the quark after being struck. Details related to proton structure are described by partonic distribution functions, which depend on the type of parton, and details regarding the hadronization process are described by fragmentation functions, which depend on both the type of parton struck and the actual observed particles.

In the earlier days of the parton model, it was assumed that one can neglect motion of the partons in any direction not parallel with the direction of the center of mass of the bound state. Effects dependent on momentum transverse to, rather than collinear with, the center of mass were considered to either "average out" or be prevented by symmetry considerations [7]. Neglecting the transverse motion of the quarks is denoted the "collinear assumption," as one considers only partonic motion collinear with the center of mass. Conversely, theories and functions which depend on the transverse momentum of the partons are denoted "transverse momentum dependent" (TMD).

Several results related the those leading to the Spin Crisis [1] could not be explained with collinear models. Two theories were put forth which depended on the transverse momentum of the partons. One theory, that of D. Sivers, placed the TMD effect within the structure of the nucleon [8], while the other theory, that of J. Collins, placed the TMD effect within the hadronization process [9]. The effects from each theory are thus called, respectively, the Sivers and Collins effects and are specifically encapsulated in the Sivers distribution function and the Collins fragmentation function. Theoretically, it was found that additional Wilson lines, related to additional initial and final state interactions [10] allow the presence of TMD effects. Early results from HERMES [11] demonstrated both processes are realized in nature, and the sub-field of TMD effects in nucleon structure and hadronization was established.

Additional TMD effects were developed, such as the Boer-Mulders effect [12], and reconsidered, such as the Cahn effect [13, 14]. An increasing number of theoretical papers developed the relation between the TMD distribution and fragmentation functions and various cross sections, e.g. Refs. [12, 15, 16]. These concepts were also detailed for SIDIS hadron pairs (denoted dihadrons) and vector mesons [17, 18].

Even with the development of TMD functions, the collinear distribution and fragmentation functions remain important, as each collinear function is an integral of a respective TMD function. Collinear functions are also easier to estimate in global fits, as the evolution with respect to energy scale is understood for collinear functions [19, 20, 21, 22, 23], while the evolution equations for TMD functions are not known. Knowledge of the evolution equations is required to compare results from experiments at different energy scales.

One of these distribution functions, for which both a collinear and TMD version exists, is the transversity distribution $h_1$ [15, 24, 25, 26]. This function occurs integrated with the Collins function in the SIDIS production of pseudo-scalar mesons. Unfortunately, since it is combined with a TMD function, certain difficulties arise in the comparison of results between different experiments. However, the SIDIS dihadron cross section contains two terms where $h_1$ occurs with collinear, yet Collins-like, fragmentation functions. This collinear access to transversity has historically been the primary motivation for SIDIS dihadron results, both experimentally and theoretically. For instance, the only published SIDIS dihadron results [27], as well as the published next-to-leading twist cross section [28], are both restricted to the collinear case.

However, much information is lost by not considering the TMD moments of the cross section. For instance, the number of unpolarized moments at leading twist reduces from 15 to three, and the number of transverse target moments, again at leading twist, reduces from 27 to two. As the Sivers function is a TMD function, none of the nine transverse target moments having contributions from the Sivers function occur in the collinear cross section. This is unfortunate, as the Sivers function for $s$-quarks, occurring in $\phi$ mesons and kaon-pair dihadron production, is of interest in understanding the process leading to gluon orbital angular momentum [29]. However, two of the nine transverse target moments related to Collins-like distribution functions exist in the collinear cross section. As the Collins function is TMD, one would not expect any of the moments to be present in the collinear case. This suggests that perhaps the two Collins-like functions occurring in the collinear moments are fundamentally different than the other seven Collins-like functions, though no strong statement can be made based on the published theory. The relation between the various Collins-like functions, and the possible distinction of two of the functions from the other seven, is clarified in Chapter II.

One other strong motivation for considering TMD dihadron production is to test the Lund/Artru string model of fragmentation [30], which predicts a sign change in the Collins function between pseudo-scalar meson production certain partial waves of vector meson production. In order to understand this prediction, it is necessary to fully quantify the relation between the Collins function in pseudo-scalar production and the nine Collins-like functions that occur in dihadron production. The previously developed theory is incomplete in defining this relation, and thus one of the important developments contained in this dissertation is the full quantification of the connection between Collins and Collins-like functions. Note, though, the Lund/Artru model can only be tested in the TMD case, not in the collinear case.

This dissertation presents the first TMD analysis of the transverse target moments in SIDIS dihadron production. Data was collected by the HERMES spectrometer during the years of 2002 to 2005. This document is organized as follows. Chapter II provides three essential ingredients. First, it includes the theoretical background, including definitions and conventions used in the remainder of this work. Second, Chapter II also presents a new partial wave analysis of the dihadron fragmentation functions, providing the explicit information needed to test the Lund/Artru model. A side result of the partial wave analysis is the computation of the next-to-leading twist cross section. Third, Chapter II contains the first model calculation of TMD dihadron fragmentation functions. The next chapter, Chapter III, outlines needed numerical methods. In particular, Chapter III includes a description of the acceptance correction method utilized in this analysis and also details a new TMD Monte Carlo generator, `TMDGen`. Prior to this work, no TMD Monte Carlo generator was available for SIDIS dihadron production, and only limited TMD generators were available for pseudoscalar production. Chapters IV and V detail the actual analysis of the dihadron data and the accompanying systematic studies. Finally, Chapter VI discusses the results and conclusions of this work.

# CHAPTER II

# Theory

This chapter includes four major sections. First, necessary definitions and conventions are presented. Second, items relevant to the cross section and its interpretation are given, the key items being an alternate partial wave expansion of the dihadron fragmentation functions (developed by the author) and the calculation of the next-to-leading twist dihadron cross section. The next-to-leading twist cross section has not been computed prior to this work, and the alternate partial wave expansion not only facilitates this computation but also aids in the interpretation of the moments. Next, the Lund/Artru fragmentation model is discussed. Testing a key prediction of the Lund/Artru model is one of the main motivations of the research contained in this dissertation. The final section of this chapter focuses on a new TMD spectator model calculation for dihadron fragmentation functions, a needed component for the Monte Carlo generator described in Section 3.2.

## 2.1 Definitions and Conventions

### 2.1.1 General Definitions

Semi-inclusive deep inelastic scattering (SIDIS) is the scattering of a lepton from a nucleon, such that the produced virtual photon has large invariant mass, the lepton and some additional specified particles are measured in the final state, and more than two additional final state particles remain unmeasured. SIDIS production of single hadrons with an electron or positron beam using a proton target is generally written as

$$e + p \to e' + h + X, \tag{2.1}$$

where $e$, $e'$ are the initial and scattered leptons, $p$ is the proton, $h$ is the measured hadron, and $X$ represents the unmeasured particles in the final state. See also Figure 2.1. SIDIS dihadron production is defined as

$$e + p \to e' + h_1 + h_2 + X, \tag{2.2}$$

where now there are two measured hadrons in the final state, $h_1$ and $h_2$. Note, SIDIS dihadron production involves multiple processes, including SIDIS vector meson production. For a more detailed treatment of possible processes included in $\pi^+\pi^-$ dihadron production, see Section III of Reference [31].

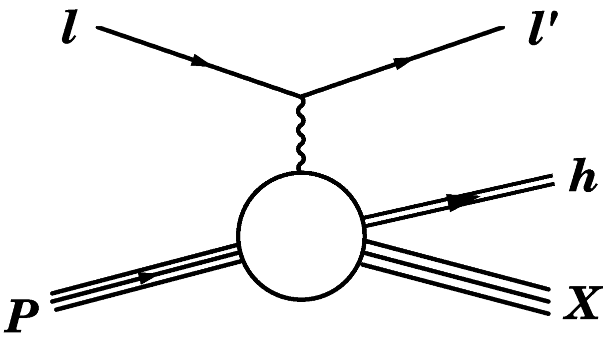


Figure 2.1: Diagrammatic depiction of SIDIS production of a single hadron. The incoming and scattered leptons are denoted $l$ and $l'$, the target is denoted $P$ and the outgoing hadron $h$. The additional unmeasured particles are denoted $X$. The non-perturbation QCD portion of the interaction is shown as an open circle.

According to the factorization theorem for SIDIS production [32], the process can be separated into three portions: a soft, non-perturbative part dealing with the distribution of quarks in the nucleon, the hard (perturbative) scattering of the virtual photon and the struck quark, and an additional soft, non-perturbative part dealing with the fragmentation of the struck quark into the measured particles, i.e. the hadronization process. This factorization theorem is independent of the specified final state particles, and thus applies to both hadron and dihadron production.

It will at times be important to distinguish with respect to which variables the cross section is differential. These variables will be denoted independent variables, as these are the statistically free variables. Additional variables will be denoted dependent variables and can be written as functions of the independent variables.

In defining angular and kinematic variables, it is necessary to adopt a convention for identifying which particle is $h_1$ and which is $h_2$. Following Ref. [33], let $p_1$ denote the positively charged hadron in the case of $\pi^+\pi^-$ and $K^+K^-$, and let $p_1$ denote the charged hadron in the case of $\pi^\pm\pi^0$.

The twist is rigorously defined as the difference between the dimension and spin of an operator in the operator product expansion of correlation functions [34]. In practice, a good working definition of twist is related to writing a Taylor series expansion of the discussed object (cross section, distribution function, fragmentation function, etc.) in terms of $1/Q^2$, with $Q^2$ the negative squared invariant mass of the virtual photon, defined in Equation 2.11. The leading term of the SIDIS cross section is twist-2, and thus next-to-leading twist is twist-3. Subtleties lie in the relation between twist and spin, and that the Taylor series expansion must be done in dimensionless quantities, i.e. ratios of other quantities over $Q^2$, though these subtleties will not be of concern within this dissertation.

### 2.1.2 Dirac Matrices

Gamma matrices will be in the chiral or Weyl representation, specifically,

$$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \quad \gamma^5 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \tag{2.3}$$

with $I_2$ the two by two identity matrix and the Pauli $\sigma$-matrices being

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.4}$$

Note that $\gamma^i$, $\gamma^5$, for $i = 1, 2, 3$ have opposite sign in this convention than in others. The commutator is defined as

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\mu]. \tag{2.5}$$

Projection operators $\mathcal{P}_\pm$ are also defined as

$$\mathcal{P}_+ = \frac{1}{2}\gamma^-\gamma^+ = \delta_{0,0} + \delta_{3,3}, \tag{2.6}$$

$$\mathcal{P}_- = \frac{1}{2}\gamma^+\gamma^- = \delta_{1,1} + \delta_{2,2}, \tag{2.7}$$

with

$$\gamma^+ = \frac{1}{\sqrt{2}}(\gamma^0 + \gamma^3) = \sqrt{2}\,(\delta_{2,0} + \delta_{1,3})\,, \tag{2.8}$$

$$\gamma^- = \frac{1}{\sqrt{2}}(\gamma^0 - \gamma^3) = \sqrt{2}\,(\delta_{0,2} + \delta_{3,1})\,. \tag{2.9}$$

### 2.1.3 Reference Frames

It will be useful to define three references frames, along with various coordinate systems in each frame. If one is only interested in defining the kinematic variables and angles in terms of measured momenta, such enumeration is somewhat pedantic. However, in Chapter III it will be necessary to invert all the relations, i.e. determine the measured momenta based on the cross section variables. In this case, such an enumeration is most useful.

The three different reference frames are defined according to the system at rest in the given frame: (I) the target, (II) the center of mass of the target, virtual photon system, and (III) the center of mass of the produced hadron system. For each frame, a number of coordinate systems are also relevant, denoted such that the Roman numeral indicates to which rest frame the coordinate system belongs. The full listing is given in Table 2.1.3, along with relations to other coordinate systems in use in the literature.

In some references, e.g. Ref. [17], a convention is used where any coordinate system with the $z$-axis coaxial with the virtual photon is a $\perp$-system, while any system with the $z$-axis coaxial with the produced hadron system's center-of-mass is a $T$-system. In $\perp$-systems the photon has no components of its momenta transverse to the $\boldsymbol{z}$-direction, while in $T$-systems

| Identifier | Description |
|---|---|
| $\mathbf{I_a}$ | The detector system: $\hat{\boldsymbol{z}}$ is along the lepton beam line and $\hat{\boldsymbol{y}}$ is in the physical up direction. |
| $\mathbf{I_b}$ | Similar to the detector system $\mathbf{I_a}$, except that $\hat{\boldsymbol{z}}$ is along the momentum of the actual beam lepton, which may differ from the ideal beam direction due to magnetic fields or radiative effects. |
| $\mathbf{I_c}$ | The prime frame of Ref. [35], Fig. 1, and Ref. [36], Fig. 1: $\hat{\boldsymbol{z}}$ is again in the direction of the beam lepton (as $\mathbf{I_b}$), but the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{z}}$ plane is now the lepton scattering plane. |
| $\mathbf{I_d}$ | The unprimed system of Ref. [35], Fig. 1, and Ref. [36], Fig. 1: system $\mathbf{I_b}$ is rotated about $\hat{\boldsymbol{y}}^{(\mathbf{I_b})}$ so that $\hat{\boldsymbol{z}}^{(\boldsymbol{I_c})}$ is in the direction of the virtual photon. Like frame $\mathbf{I_b}$, the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{z}}$ plane is the lepton scattering plane. |
| $\mathbf{II_a}$ | As with system $\mathbf{I_c}$, the $z$-axis is aligned with the virtual photon direction and the $x$- and $z$-axis lie in the lepton scattering plane. The difference in frames $\mathbf{I_c}$ and $\mathbf{II_a}$ is a boost opposite the direction of the virtual photon. |
| $\mathbf{II_b}$ | The $z$-axis is now in the direction of the produced meson system, and the $x$ axis remains in the lepton scattering plane. |
| $\mathbf{II_c}$ | As with system $\mathbf{II_c}$, the $z$-axis is aligned with the direction of the produced meson system. However, the $x$ axis of this system is in the hadron production plane. The primed system of Ref. [35], Fig. 2, is anti-aligned with coordinate system $\mathbf{II_c}$. |
| $\mathbf{III_a}$ | Frame $\mathbf{II_b}$ is boosted to the frame III. |
| $\mathbf{III_b}$ | Frame $\mathbf{II_c}$ is boosted to the frame III, corresponding to the unprimed system of Ref. [35], Fig. 2. |

Table 2.1: Description of coordinate systems.

the center of mass of the produced hadrons has no transverse components. Frames $\mathbf{I_d}$ and $\mathbf{II_a}$ are both $\perp$ frames, while frames $\mathbf{II_b}$, $\mathbf{II_c}$, $\mathbf{III_a}$, and $\mathbf{III_b}$ are all $T$ frames. The $T/\perp$ convention does not define which object is at rest in each frame. In fact, in some cases, a specific object at rest is implied, while in other cases, the notation allows for any boost along the $z$-axis. For the purpose of this dissertation, a more detailed notation is preferable.

### 2.1.4 Variable Definitions

#### 2.1.4.1 DIS Variables

Let the virtual photon 4-momentum be denoted $q^\mu$, and the three momentum and its magnitude be denoted $\boldsymbol{q}$ and $|\boldsymbol{q}|$, respectively. Let the target 4-momentum be denoted $P^\mu$, and the incoming and scattered lepton four-momenta as $k^\mu$ and $k'^\mu$.

| **Variable** | **Definition** |
|---|---|
| $E$ | lepton beam energy |
| $E'$ | scattered lepton energy |
| $\theta_{e2}$ | angle between lepton beam and scattered lepton<br>(polar angle of scattered lepton in frame $\mathbf{I_a}$) |
| $M$ | Target mass |

Table 2.2: Input quantities for DIS variables.

The DIS variables are all dependent on the quantities given in Table 2.2. As the target mass $M$ and beam energy $E$ are assumed fixed, only two of the DIS variables can be chosen as independent. The DIS variables can be computed from the quantities in Table 2.2, the most common variables being

$$\nu = \frac{q \cdot P}{M} = E - E', \tag{2.10}$$

$$Q^2 = -q^2 \approx 4EE' \sin^2(\theta_{e2}/2), \tag{2.11}$$

$$W^2 = (P+q)^2 = M^2 + 2M\nu - Q^2, \tag{2.12}$$

$$x = \frac{Q^2}{2M\nu}, \tag{2.13}$$

$$y = \frac{k \cdot P}{q \cdot P} = \frac{\nu}{E}, \tag{2.14}$$

$$s = (k+P)^2 = \frac{Q^2}{xy} + M^2 + m_l^2, \tag{2.15}$$

$$\gamma = \frac{2Mx}{Q} = \frac{Q}{yE_B} = \frac{Q}{\nu}, \tag{2.16}$$

$$\epsilon = \frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2 y^2}, \tag{2.17}$$

$$\sin\theta_\gamma = \gamma\sqrt{\frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1+\gamma^2}}. \tag{2.18}$$

The approximation in the definition of $Q^2$ is based on the assumption of effectively massless leptons. Note the lepton mass $m_l$ has been left in the equation for $s$. The quantity $Q^2$ is defined as the negative square of the mass of the virtual photon, while $W$ is the mass of the virtual photon, target system. The variables $x$ and $y$ are interpreted as the fraction of the target momentum carried by the struck quark, and the fraction of the beam energy transferred to the virtual photon, respectively. The variable $\nu$ is the virtual photon energy in the lab frame. The value $\epsilon$ is the ratio of longitudinal and transverse photon flux, and $\theta_\gamma$ is the angle between the lepton beam and the virtual photon. For a complete review of the definition of the various DIS variables, see Section 16 of Ref. [37].

#### 2.1.4.2 SIDIS Variables

The mass, energy, and momentum of the produced hadron system in the lab frame will be denoted, respectively, $M_h$, $E_h$, $\boldsymbol{P}_h$. The magnitude of the momentum will be written as $P_h := |\boldsymbol{P}_h|$. In the case of single hadron production, the momentum is measured and the particle identification yields the mass, thus specifying the energy. For dihadron production, the energy and momentum of the system is determined from the energy and momentum of the two measured hadrons, from which the dihadron mass can be determined. Thus for dihadron production, the invariant mass is considered one of the SIDIS variables, while for single hadron production the hadron mass is considered fixed.

The kinematic SIDIS variables generally used in the cross section (in addition to $M_h$) are $z$ and $P_{h\perp}$. They are defined, respectively, as the fraction of the virtual photon energy carried by the hadron system, and as the projections of $P_h$ perpendicular to the virtual photon direction. In the lab frame, they can be computed as

$$z = \frac{E_h}{\nu}, \tag{2.19}$$

$$P_{h\perp} = P_h \sin\theta_{\gamma h}, \tag{2.20}$$

where $\theta_{\gamma h}$ is the angle between the dihadron momenta and the virtual photon momenta, measured in the lab frame.

#### 2.1.4.3 Angles

Three planes are utilized in defining the relevant azimuthal angles. The lepton scattering plane includes the lepton beam, scattered lepton, and virtual photon. The hadron production plane includes the virtual photon and the center of mass of the produced hadron system. The decay plane includes the center of mass of the produced hadron system as well as the two measured hadrons.

Both SIDIS dihadron and single hadron production utilize the azimuthal angle $\phi_h$, the angle between the lepton scattering plane and the hadron production plane, and $\phi_S$, the angle between the lepton production plane and the transverse target polarization direction. Both of these angles are measured perpendicular to the virtual photon momenta direction, and thus can be measured in either reference frame **I** or **II**. A diagram showing these definitions is given in Figure 2.2. Note, a subtlety exists in regards to defining asymmetries about the virtual photon direction or about the lepton beam direction. For a full treatment, see Ref. [16] and the discussion in Section 3.2.1. The definitions for $\phi_h$, $\phi_S$ are given in agreement with the Trento convention [39], and can be explicitly computed according to

$$\phi_h = \mathrm{signum}\big[\,(\boldsymbol{k} \times \boldsymbol{P}_h) \cdot \boldsymbol{q}\big] \arccos \frac{(\boldsymbol{q} \times \boldsymbol{k}) \cdot (\boldsymbol{q} \times \boldsymbol{P}_h)}{|\boldsymbol{q} \times \boldsymbol{k}|\,|\boldsymbol{q} \times \boldsymbol{P}_h|}, \tag{2.21}$$

$$\phi_S = \mathrm{signum}\big[\,(\boldsymbol{k} \times \boldsymbol{S}) \cdot \boldsymbol{q}\big] \arccos \frac{(\boldsymbol{q} \times \boldsymbol{k}) \cdot (\boldsymbol{q} \times \boldsymbol{S})}{|\boldsymbol{q} \times \boldsymbol{k}|\,|\boldsymbol{q} \times \boldsymbol{S}|}, \tag{2.22}$$

with $\mathrm{signum}[a] = 1$ if $a > 0$, $-1$ otherwise, and where $\boldsymbol{S}$ is a vector indicating the target polarization in the lab frame. For HERMES, $\boldsymbol{S}$ is $(0, -1, 0)$.

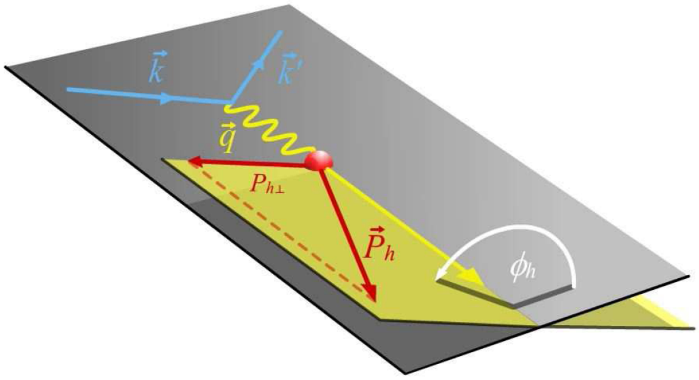


Figure 2.2: Diagram depicting the angles $\phi_h$ and $\phi_S$, from Ref. [38].

#### 2.1.4.4 Additional Dihadron Variables

As the mass of the additional particle in the final state is assumed to be known, the additional measured particle increases the number of independent variables by three. Likewise, all previous SIDIS variables are now defined to be with respect to the two hadron system. Letting $p_1^\mu$, $p_2^\mu$ be the 4-vectors of the two measured hadrons, one can define 4-vectors

$$P_h^\mu = p_1^\mu + p_2^\mu, \tag{2.23}$$

$$R^\mu = \frac{1}{2}\left(p_1^\mu - p_2^\mu\right). \tag{2.24}$$

The three additional variables are usually chosen to be the invariant mass of the dihadron system, $M_h = \sqrt{(P_h)_\mu (P_h)^\mu}$, and $\cos\vartheta$, $\phi_R$, the cosine of the polar angle as well as the azimuthal angle of $R^\mu$ in reference frame $\mathbf{III_a}$. Note, $R^\mu$ has azimuthal angle $\phi_R$ in both coordinate system $\mathbf{II_b}$ and $\mathbf{III_a}$, as this angle is unaffected by the boost in the $z$-direction. Other choices for independent variables include $z_1$, $z_2$, using Equation 2.19 for each of the individual hadron energies; $\zeta = (z_1 - z_2)/(z_1 + z_2)$, to the asymmetry between $z_1$, $z_2$; and/or $\varphi$, the azimuthal angle of $R^\mu$ in frame $\mathbf{III_b}$. Note that the polar angle $\vartheta$ is the same in both frame $\mathbf{III_a}$ and $\mathbf{III_b}$. The set of variables $(M_h, \cos\vartheta, \varphi)$, correspond with those chosen in exclusive meson production [35, 36].

To determine a closed form equation for $\phi_R$, one needs to compute the $x$-axis in coordinate system $\mathbf{II_b}$ from the available 3-vectors $\boldsymbol{k}$, $\boldsymbol{q}$, $\boldsymbol{P}_h$, $\boldsymbol{R}$, corresponding to the 3-momenta of input lepton and the virtual photon and the 3-vector portions of the 4-vectors $P_h^\mu$ and $R^\mu$. Let $\boldsymbol{n}$ be a vector parallel to the $x$-axis in coordinate system $\mathbf{II_b}$. Then $\boldsymbol{n}$ is coplanar with $\boldsymbol{k}$, $\boldsymbol{q}$ and thus can be written as a linear combination of these two vectors. Imposing the constraint that $\boldsymbol{n}$ is perpendicular to $\boldsymbol{P}_h$, and fixing the sign of $\boldsymbol{n}$ by requiring $(\boldsymbol{P}_h \times \boldsymbol{n}) \cdot (\boldsymbol{q} \times \boldsymbol{k}) > 0$, yields the expression

$$\boldsymbol{n} = (\boldsymbol{q} \cdot \boldsymbol{P}_h)\,\boldsymbol{k} - (\boldsymbol{k} \cdot \boldsymbol{P}_h)\,\boldsymbol{q}. \tag{2.25}$$

Let $\boldsymbol{R}_T$ denote the projection of $\boldsymbol{R}$ perpendicular to $\boldsymbol{P}_h$,

$$\boldsymbol{R}_T = \boldsymbol{R} - \frac{\boldsymbol{R}\cdot\boldsymbol{P}_h}{|\boldsymbol{P}_h|^2}\boldsymbol{P}_h. \tag{2.26}$$

One can then identify

$$\phi_R \;=\; \text{signum}\big[\,(\boldsymbol{n}\times\boldsymbol{R}_T)\cdot\boldsymbol{P}_h\big]\arccos\frac{\boldsymbol{n}\cdot\boldsymbol{R}_T}{|\boldsymbol{n}||\boldsymbol{R}_T|}, \tag{2.27}$$

which can also be equivalently written as

$$\phi_R \;=\; \text{signum}\big[\,(\boldsymbol{R}\times\boldsymbol{P}_h)\cdot\boldsymbol{n}\big]\arccos\frac{(\boldsymbol{q}\times\boldsymbol{k})\cdot(\boldsymbol{P}_h\times\boldsymbol{R}_T)}{|\boldsymbol{q}\times\boldsymbol{k}|\,|\boldsymbol{P}_h\times\boldsymbol{R}_T|}. \tag{2.28}$$

Note $\boldsymbol{P}_h\times\boldsymbol{R}_T = \boldsymbol{P}_h\times\boldsymbol{R}$. Although $\phi_R$ has been defined with words in many references, no closed form solution for $\phi_R$ has been presented prior to this work. Note, Equation 2.28 must be computed in reference frame **II**.

In some references [27], $\phi_{R\perp}$ is used instead of $\phi_R$.[1] The difference between $\phi_{R\perp}$ and $\phi_R$ is $Q^2$ suppressed, and thus ignored in leading twist analyses. The angle $\phi_{R\perp}$ is defined by taking the projection of $R_T$ perpendicular to $\boldsymbol{k}$, and then considering the azimuthal angle with respect to the lepton scattering plane. In analogy with the Trento convention for $\phi_h$, one can then write

$$\phi_{R\perp} \;=\; \text{signum}\big[\,(\boldsymbol{q}\times\boldsymbol{k})\cdot\boldsymbol{R}_T\big]\arccos\frac{(\boldsymbol{q}\times\boldsymbol{k})\cdot(\boldsymbol{q}\times\boldsymbol{R}_T)}{|\boldsymbol{q}\times\boldsymbol{k}|\,|\boldsymbol{q}\times\boldsymbol{R}_T|}, \tag{2.29}$$

consistent with the definition of $\phi_R$ in Ref. [18]. In the high $Q^2$ limit, $\boldsymbol{P}_h$ becomes parallel with $\boldsymbol{q}$, and Equation 2.28 and 2.29 become equal, as required.

An additional relationship between the quantities $\phi_R$ and $\varphi$ can be determined by comparing the rotation matrix that transforms frame $\mathbf{II_a}$ to $\mathbf{II_b}$ with the matrix that transforms frame $\mathbf{II_a}$ to $\mathbf{II_c}$. The difference in the azimuthal angles of these frames is simply $\varphi - \phi_R$. Also note that $\varphi = \phi_h + \phi_{R\perp}$. Comparing the rotation matrices yields the result

$$\tan(\varphi - \phi_R) = \tan(\phi_h)\cos\theta_{\gamma h}^{(\mathbf{II})}, \tag{2.30}$$

where $\cos\theta_{\gamma h}^{(\mathbf{II})}$ is the angle between the virtual photon and $\boldsymbol{P}_h$ in rest frame II. This expression then implies

$$\phi_R - \phi_{R\perp} = \phi_h - \tan^{-1}\left(\tan(\phi_h)\cos\theta_{\gamma h}^{(\mathbf{II})}\right). \tag{2.31}$$

Once again, this equation shows that in the limit of infinitely high $Q^2$, the difference between $\phi_R$ and $\phi_{R\perp}$ approaches zero, as in that limit the quantity $\cos\theta_{\gamma h}^{(\mathbf{II})}$ approaches unity.

To compute the quantity $\cos\theta_{\gamma h}^{(\mathbf{II})}$, one can boost $P_h^\mu$ to frame $\mathbf{II_a}$, where the $z$ component is $P_{h,z}^{(\mathbf{II_a})} = \left|\boldsymbol{P}_h^{(\mathbf{II_a})}\right|\cos\theta_{\gamma h}^{(\mathbf{II_a})}$. If one first computes the analogous quantity in the lab frame

---

[1] Some references, e.g. Refs. [31, 40], actually use the symbol $\phi_R$ but give it the definition of $\phi_{R\perp}$. Such references are generally focused on the high $Q^2$ limit (leading twist), where the angles become equivalent.

(rest frame I) via[2]

$$\cos\theta_{\gamma h} = \sqrt{1 - \left(\frac{P_{h\perp}}{P_h}\right)^2}, \tag{2.32}$$

one can then determine

$$\begin{aligned} P_{h,z}^{(\mathbf{II_a})} &= \left(\frac{\nu + M}{W}\right) P_h \cos\theta_{\gamma h} - \frac{|\boldsymbol{q}|}{W} E_h, & (2.33) \\ \left|\boldsymbol{P}_h^{(\mathbf{II_a})}\right| &= \sqrt{P_{h\perp}^2 + \left(P_{h,z}^{(\mathbf{II_a})}\right)^2}, & (2.34) \\ \cos\theta_{\gamma h}^{(\mathbf{II_a})} &= \frac{P_{h,z}^{(\mathbf{II_a})}}{\left|\boldsymbol{P}_h^{(\mathbf{II_a})}\right|}. & (2.35) \end{aligned}$$

Combining Equations 2.30 with 2.33 through 2.35 yields a precise relation relation between $\phi_R$ and $\phi_{R\perp}$. Note, that although the difference between $\phi_R$ and $\phi_{R\perp}$ is well known to be suppressed by $Q^2$, no exact expression such as Equation 2.31 was previously available.

#### 2.1.4.5 Intrinsic Variables

The transverse momentum of the struck quark (i.e. the quark to which the distribution function corresponds) will be denoted $\boldsymbol{p}_T$, with magnitude $p_T := |\boldsymbol{p}_T|$, and azimuthal angle $\phi_p$. Likewise, the transverse momentum of the fragmenting quark (i.e. the quark to which the fragmentation function corresponds) will be denoted $\boldsymbol{k}_T$, with magnitude $k_T := |\boldsymbol{k}_T|$ and azimuthal angle $\phi_k$. The quantities $p_T$ and $k_T$ are always positive, in contrast with some sources which define $k_T$ as the norm of the a four vector with only transverse components. The Minkowski metric would then make such a norm negative. The above identification of $p_T$, $k_T$ is consistent with the Amsterdam notation of the distribution and fragmentation functions, [15, 16, 28, 36], which is used throughout this dissertation. Note, some theorists do not follow the Amsterdam notation, and instead follow that of the Torino theory group, e.g. [41, 42, 43], where

$$\begin{aligned} p_T\Big|_{Torino} &= z k_T\Big|_{Amsterdam} = k'_T, & (2.36) \\ k_T\Big|_{Torino} &= p_T\Big|_{Amsterdam}. & (2.37) \end{aligned}$$

## 2.2 Partial Wave Expansion and Cross Section

### 2.2.1 Distribution Functions

Distribution functions will be given in the Amsterdam notation [12, 15, 44], as previously noted. The leading twist distribution functions can be interpreted as differences in helicity distributions or equivalently as helicity amplitudes [45]. Owing to the great deal of literature available concerning the distribution functions, a few references are provided rather than

---

[2] This is from the inversion of Equation 2.20, assuming $\theta_{\gamma,h} \leq \pi/2$.

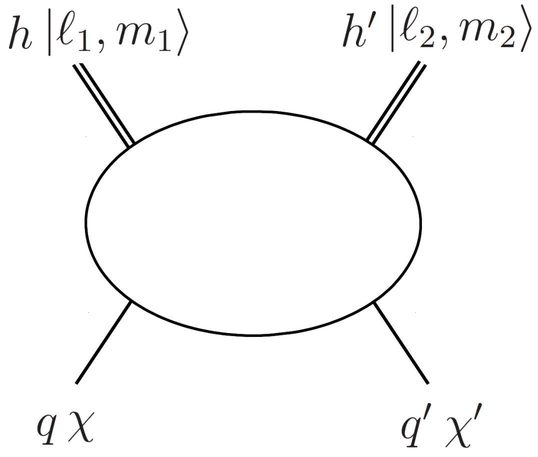


Figure 2.3: The generic diagram for the three leading order fragmentation functions. Via the optical theorem, the fragmentation functions are defined as the imaginary part of the above amplitude. The quarks are indicated as $q$, $q'$, with their spin states specified by $\chi$, $\chi'$, while the hadrons (dihadrons) $h$ and $h'$ are, respectively, in the partial waves $|\ell_1, m_1\rangle$ and $|\ell_2, m_2\rangle$.

attempting to summarize this large and complex field of study. Some of the more common leading twist fragmentation functions include the Sivers function $f_{1T}^{\perp}$ [8, 46], transversity $h_1$ [47, 48], the Boer-Mulders function $h_1^{\perp}$ [12, 49], and pretzelocity $h_{1T}^{\perp}$ [50, 51]. It should be remarked that the Sivers function for the $\phi$ meson, as well as for $K^+K^-$ hadron pairs, has been related to the orbital angular momentum of the gluons and can provide important tests for the relevant mechanisms [29].

### 2.2.2 Fragmentation Functions

The diagram used to define the leading twist fragmentation functions is given in Figure 2.3. A new convention is adopted in this work, where the name and symbol of the fragmentation are entirely associated with the quark spin states, i.e. $\chi$, $\chi'$ in Figure 2.3, while the various polarization states of the produced hadron(s), i.e. $|l_1, m_1\rangle$ and $|l_2, m_2\rangle$, are associated with partial waves of fragmentation functions. Such a convention requires a slight redefinition of the fragmentation functions and a new partial wave expansion. Note, previous definitions of the fragmentation functions either assume no polarization in the final state hadrons [16] (applicable for pseudo-scalar production), or define mixtures of certain partial waves as new fragmentation functions [15, 18, 52]. For example, one of the original sources for the SIDIS cross section, Ref. [15], introduces new fragmentation functions, including $G_{1L}$ and $G_{1T}$, based on the polarization of the final state. According to the convention proposed in this dissertation, there are only two fragmentation functions: the unpolarized fragmentation function $D_1$ which corresponds to the sum of non-quark-spin-flip diagrams, i.e those with $\chi = \chi'$, and the polarized fragmentation function or generalized Collins fragmentation function $H_1^\perp$, which corresponds to the sum of diagrams where the quark flips spin, i.e. $\chi \neq \chi'$.

One advantage to this new convention is that it places a clear distinction between the spin structure of fragmentation and the polarization structure of the produced system. However, the final result for the cross section with this convention is fully consistent with the cross section in the literature [18], given the relation between notations in Section 2.2.6.

To rigorously define the fragmentation functions, one must first consider the leading twist fragmentation correlation matrix $\Delta$, defined as the sum of all possible polarization states of Figure 2.3. This quantity is rigorously defined as a certain matrix element, e.g. Equation 24 of Ref. [15]. It is common to define a shorthand notation for an integral of traces of the fragmentation correlation matrix [15, 52]. For this document, let $\Delta^{[\Gamma]}$ be defined as short hand for

$$\Delta^{[\Gamma]}(z, M_h, |\boldsymbol{k}_T|, \cos\vartheta, \phi_R - \phi_k) = 4\pi \frac{z|\boldsymbol{R}|}{16M_h} \int dk^+ \ \mathrm{Tr}\left[\Gamma\Delta(k, P_h, R)\right]\Big|_{k^- = P_h^-/z}. \tag{2.38}$$

The fragmentation functions can be rigorously defined as traces of the fragmentation correlation matrix

$$D_1 = \Delta^{[\gamma^-(1+i\gamma^5)]}, \tag{2.39}$$

$$i\frac{|\boldsymbol{k}_T|}{M_h} e^{i\phi_k} H_1^\perp = \Delta^{[-i(\sigma^{1-}+i\sigma^{2-})\gamma^5]} = \Delta^{[(\gamma^2-i\gamma^1)\gamma^-\gamma^5]}. \tag{2.40}$$

These fragmentation functions are denoted non-expanded fragmentation functions when it is needed to distinguish them from fragmentation functions occurring in the partial wave expansion. In the case of pseudo-scalar meson production, some of the traces in Equations 2.39 and 2.40 are zero, thus reducing the definitions to the common expressions for pseudo-scalar productions [15].

### 2.2.3 Partial Wave Expansion

Note, for the rest of this dissertation, the $h$, $h'$ of Figure 2.3 are assumed to be dihadrons. However, the following equally applies to any hadron, dihadron, or higher multiplicity hadron Fock state, with the caveat that certain final states have limited partial waves. For example, only the pure $s$-wave state $|0,0\rangle$ is available for pseudo-scalar meson production.

The partial waves of the non-expanded fragmentation function can be defined in the direct product basis $|\ell_1, m_1\rangle\,|\ell_2, m_2\rangle$, defining the polarization states of both dihadrons occurring in Figure 2.3. However, nature generally prefers direct sum bases rather than direct product bases. For example, the four possible states of a quark, anti-quark pair in nature are a spin-0 pseudo-scalar meson and three polarizations of spin-1 vector mesons, not as two spin-aligned and two spin-anti-aligned states. Similarly, individual terms in the cross section are not related to partial waves in the direct product basis $|\ell_1, m_1\rangle\,|\ell_2, m_2\rangle$, but are related to partial waves in the direct sum basis $|\ell, m\rangle$, i.e. the overall spin-state of the two dihadron system.

Note, there are four quarks exiting the top of the diagram in Figure 2.3. Thus, in the usual shorthand, one writes

$$\begin{aligned}\frac{1}{2}\otimes\frac{1}{2}\otimes\frac{1}{2}\otimes\frac{1}{2} &= \left(\frac{1}{2}\otimes\frac{1}{2}\right)\otimes\left(\frac{1}{2}\otimes\frac{1}{2}\right), \\ &= (1\oplus 0)\otimes(1\oplus 0), \\ &= 2\oplus 1\oplus 1\oplus 1\oplus 0\oplus 0. \qquad (2.41)\end{aligned}$$

Thus, sixteen states are actually present: the five states of the spin-2 system, nine states arising from three sets of the three states of the spin-1 system, and the two spin-0 states. However, distinguishing between the three different spin-1 states, as well as between the two spin-0 states, is difficult. Theoretically, the different $\ell = 1$ and $\ell = 0$ states can be distinguished by using Generalized Casimir operators [53, 54]. These operators are related to the coupling scheme, i.e. the placement of parenthesis in Equation 2.41. Specifically, Equation 2.41 corresponds to coupling the first two and the last two quarks in the intermediate state. An alternate coupling scheme would be to couple the third quark to the first two in the intermediate step. Although the final line in Equation 2.41 does not depend on the coupling, different bases within the degenerate $1\oplus 1\oplus 1$ and $0\otimes 0$ space are implied by different couplings. Relations between choices of basis are then given by Racah Coefficients, and Clebsch-Gordan coefficients finally could be used to relate all the states [54, 55]. This would, unfortunately, require experimentally adjusting the coupling scheme, i.e. measuring the interference between a three quark state and a one quark state, which is not possible.

Thus, only nine combinations of the sixteen states are experimentally accessible, appearing as $2\oplus 1\oplus 0$. The experimentally observed spin-1 states are the sum of three distinct spin-1 systems: one arising from interference between the vector meson states (denoted *pp*-interference), and the other two arising from interference between vector meson states and pseudo-scalar states (denoted *sp*-interference). Likewise, the measurable spin-0 state contains the pseudo-scalar state as well as *pp*-interference between the two transverse polarization states.

The partial wave expansion into the $2\oplus 1\oplus 0$ states is accomplished by expanding the fragmentation functions of Equations 2.39 and 2.40 in terms of spherical harmonics. The polar angle is $\cos\vartheta$, while the azimuthal angle is $\phi_R - \phi_k$ as other constraints require these functions to only depend on this difference and not the angles $\phi_k$, $\phi_R$ individually [18].

Previously, partial wave analyses have only been preformed either at leading twist [18] or at next-to-leading twist but integrated over $P_{h\perp}$ [28]. In both cases, the partial wave analysis is done with respect to the direct product basis $|\ell_1 m_1\rangle\,|\ell_2 m_2\rangle$ of the two dihadrons occurring in Figure 2.3. Previous expansions are related to those listed here via Clebsch-Gordon coefficients, up to normalizations of the basis functions. In fact, the common Legendre polynomial expansion of the dihadron fragmentation functions [18, 28] corresponds with the $\cos\vartheta$ dependent factor in Equations 2.42 and 2.43. Thus, although the usual expansions in terms of Legendre polynomials occur with little motivation in the literature, the full justification is given by the partial wave expansion in the direct sum basis.

The partial waves of the fragmentation functions are then defined according to

$$D_1 = \sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell} P_{\ell,m}(\cos\vartheta)e^{im(\phi_R-\phi_k)}D_1^{|\ell,m\rangle}(z, M_h, |\boldsymbol{k}_T|), \tag{2.42}$$

$$H_1^{\perp} = \sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell} P_{\ell,m}(\cos\vartheta)e^{im(\phi_R-\phi_k)}H_1^{\perp|\ell,m\rangle}(z, M_h, |\boldsymbol{k}_T|), \tag{2.43}$$

and likewise for the higher twist fragmentation functions. All non-expanded fragmentation functions depend on the variables $z$, $M_h$, $|\boldsymbol{k}_T|$, $\cos\vartheta$, $\phi_R - \phi_k$, and possibly $Q^2$ (as $Q^2$ dependence is usually implicitly assumed)[3]. Alternately, the expansion could have been done with respect to $\phi_{R\perp} - \phi_k \equiv \varphi + \phi_h - \phi_k$. The final result for structure functions would then align with [35]. At high enough $Q^2$, the difference between the choice of expansion becomes negligible.

In the cross section, the real and imaginary portions of $D_1$ are separated, and thus it is useful to define the notation

$$D_1^{|\ell,m\rangle+} = D_1^{|\ell,m\rangle} + D_1^{|\ell,-m\rangle}, \tag{2.44}$$

$$D_1^{|\ell,m\rangle-} = D_1^{|\ell,m\rangle} - D_1^{|\ell,-m\rangle}, \tag{2.45}$$

assuming $m > 0$. The Legendre polynomials used in Equations 2.42 and 2.43 are

$$\begin{aligned} P_{0,0} &= 1, & P_{2,0} &= \frac{1}{2}\left(3\cos^2\vartheta - 1\right), \\ P_{1,0} &= \cos\vartheta, & P_{2,1} &= \sin 2\vartheta, \\ P_{1,1} &= \sin\vartheta, & P_{2,2} &= \sin^2\vartheta, \end{aligned} \tag{2.46}$$

with $P_{\ell,-m} := P_{\ell,m}$.

### 2.2.4 Explicit Cross Section

These new expanded fragmentation functions can be inserted into the hadronic tensor and the cross section can be written in terms of structure functions, following the same method as Ref. [16]. The leading twist cross section, written in terms of structure functions, is identical to that one would obtain from Ref. [18].[4] The interpretation of the structure functions is then the only difference in the TMD twist-2 cross section between this work and previously published papers. This work, however, also contains the TMD structure functions arising at twist-3, which has not been previously available.

As was noted before, the cross section with the non-expanded fragmentation functions is identical to that for single pseudo-scalar meson production. This allows one to compute the cross section for dihadron production at any twist level, given the pseudo-scalar cross section at the corresponding twist level. One just needs to apply the partial wave expansion

---

[3] The $Q^2$ evolution for TMD fragmentation functions is still an open question. Related references include [56, 57, 58, 59].

[4] Note, Equation C4 of Ref. [18] is missing the term proportional to $\cos\vartheta\cos 2\phi_h$, which should be accompanied by the fragmentation function $H_{1,OL}^{\perp}$ occurring in Equation 57 of the same reference.

of the fragmentation functions. Note this method is much less complicated than traditional methods, such as was employed for the collinear case [28].

A few details need description before the presentation of the cross section. Note, the terms of the cross section will be grouped into terms denoted $\sigma_{XY}$ based on the target and beam polarization. Specifically, subscripts $XY$ correspond to the beam ($X$) and target ($Y$) polarization, taking the values $U$ (unpolarized), $L$ (longitudinally polarized) and $T$ (transversely polarized). The total cross section is the sum of all terms $\sigma_{XY}$. The structure functions will likewise have subscripts $XY$, with the same meaning. In a few cases the structure functions are split into those for transverse and longitudinal virtual photon polarization. These are indicated with a subscript $XY,Z$, with $Z$ being either $L$ or $T$ to indicate the virtual photon polarization.

The depolarization factors occurring with leading twist moments are [16]

$$\begin{aligned} A(x,y) &= \frac{y^2}{2(1-\epsilon)} = \frac{1-y+\frac{1}{2}y^2+\frac{1}{4}\gamma^2y^2}{1+\gamma^2} \\ &\approx \left(1-y+\frac{1}{2}y^2\right), \end{aligned} \tag{2.47}$$

$$\begin{aligned} B(x,y) &= \frac{y^2}{2(1-\epsilon)}\epsilon = \frac{1-y-\frac{1}{4}\gamma^2y^2}{1+\gamma^2} \\ &\approx (1-y), \end{aligned} \tag{2.48}$$

$$\begin{aligned} C(x,y) &= \frac{y^2}{2(1-\epsilon)}\sqrt{1-\epsilon^2} = \frac{y\left(1-\frac{1}{2}y\right)}{1+\gamma^2} \\ &\approx y\left(1-\frac{1}{2}y\right), \end{aligned} \tag{2.49}$$

while the depolarization factors occurring with the next-to-leading twist moments are

$$\begin{aligned} V(x,y) &= \frac{y^2}{(1-\epsilon)}\sqrt{2\epsilon(1+\epsilon)} = \frac{2(2-y)}{1+\gamma^2}\sqrt{1-y-\frac{1}{4}\gamma^2y^2} \\ &\approx 2(2-y)\sqrt{1-y}, \end{aligned} \tag{2.50}$$

$$\begin{aligned} W(x,y) &= \frac{y^2}{(1-\epsilon)}\sqrt{2\epsilon(1-\epsilon)} = \frac{2y}{1+\gamma^2}\sqrt{1-y-\frac{1}{4}\gamma^2y^2} \\ &\approx 2y\sqrt{1-y}. \end{aligned} \tag{2.51}$$

The approximations hold in the high $Q^2$ limit when one neglects all but the next-to-leading twist. Note also, in the high $Q^2$ limit, the expressions in Equations 2.47 through 2.51 are independent of $x$ and only depend on $y$.

The cross section will be chosen to be differential with respect to $x$, $y$, $z$, $P_{h\perp}$, $\phi_h$, $\phi_S$, $M_h$, $\cos\vartheta$, and $\phi_R$. The phase space factor is taken from the single hadron phase space factor of Ref. [16], with an additional factor of $2M_hP_{h\perp}/4\pi$ to account for the phase space of $\cos\vartheta$, $\phi_R$ and $M_h$, and the fact that the cross section is chosen differential with respect to the polar angles $P_{h\perp}$, $\phi_h$ rather than the Cartesian 2-vector $\boldsymbol{P}_{h\perp}$.

Note, some discrepancy exists in the literature regarding the phase space factor. For example, Ref. [15] uses $s/Q^4$, while Ref. [16] uses $1/(xyQ^2)$. Note,

$$\frac{s}{Q^4} = \frac{1}{xyQ^2} + \frac{M^2}{Q^4}, \tag{2.52}$$

and thus the factors agree at high $Q^2$.

The magnitudes of the target polarization will be denoted $S_{\parallel}$ and $|\boldsymbol{S}_{\perp}|$, referring to the magnitudes of the longitudinal and transverse polarizations, as in Ref. [16]. These are the same quantities as $S_T$ and $S_L$ in Ref. [36]. The longitudinal polarization factor for the beam will be denoted $\lambda_e$.

The cross section for unpolarized beam, unpolarized target, can be written as

$$\begin{aligned} d\sigma_{UU} \;=\; & \frac{\alpha^2 M_h P_{h\perp}}{2\pi xyQ^2}\left(1+\frac{\gamma^2}{2x}\right) \\ & \times \sum_{\ell=0}^{2} \Bigg\{ A(x,y) \sum_{m=0}^{\ell} \Bigg[ P_{\ell,m}\cos(m(\phi_h-\phi_R)) \\ & \qquad \times \left( F_{UU,T}^{P_{\ell,m}\cos(m(\phi_h-\phi_R))} + \epsilon F_{UU,L}^{P_{\ell,m}\cos(m(\phi_h-\phi_R))} \right) \Bigg] \\ & \quad + B(x,y) \sum_{m=-\ell}^{\ell} P_{\ell,m}\cos((2-m)\phi_h + m\phi_R) F_{UU}^{P_{\ell,m}\cos((2-m)\phi_h+m\phi_R)} \\ & \quad + V(x,y) \sum_{m=-\ell}^{\ell} P_{\ell,m}\cos((1-m)\phi_h + m\phi_R) F_{UU}^{P_{\ell,m}\cos((1-m)\phi_h+m\phi_R)} \Bigg\}. \end{aligned} \tag{2.53}$$

The longitudinally polarized beam, unpolarized target moments are

$$\begin{aligned} d\sigma_{LU} \;=\; & \frac{\alpha^2 M_h P_{h\perp}}{2\pi xyQ^2}\left(1+\frac{\gamma^2}{2x}\right)\lambda_e \\ & \times \sum_{\ell=0}^{2} \Bigg\{ A(x,y) \sum_{m=1}^{\ell} \Bigg[ P_{\ell,m}\sin(m(\phi_h-\phi_R)) \\ & \qquad \times \left( F_{LU,T}^{P_{\ell,m}\cos(m(\phi_h-\phi_R))} + \epsilon F_{LU,L}^{P_{\ell,m}\cos(m(\phi_h-\phi_R))} \right) \Bigg] \\ & \quad + W(x,y) \sum_{m=-\ell}^{\ell} P_{\ell,m}\sin((1-m)\phi_h + m\phi_R) F_{LU}^{P_{\ell,m}\sin((1-m)\phi_h+m\phi_R)} \Bigg\}. \end{aligned} \tag{2.54}$$

The unpolarized beam, longitudinally polarized target moments are

$$\begin{aligned} d\sigma_{UL} &= \frac{\alpha^2 M_h P_{h\perp}}{2\pi x y Q^2}\left(1+\frac{\gamma^2}{2x}\right) S_\parallel \\ &\times \Bigg\{ C(x,y)\sum_{\ell=1}^{2}\sum_{m=1}^{\ell} P_{\ell,m}\sin(-m\phi_h+m\phi_R) F_{UL}^{P_{\ell,m}\sin(-m\phi_h+m\phi_R)} \\ &\quad + B(x,y)\sum_{\ell=0}^{2}\sum_{m=-\ell}^{\ell} P_{\ell,m}\sin((2-m)\phi_h+m\phi_R) F_{UL}^{P_{\ell,m}\sin((2-m)\phi_h+m\phi_R)} \\ &\quad + V(x,y)\sum_{\ell=0}^{2}\sum_{m=-\ell}^{\ell} P_{\ell,m}\sin((1-m)\phi_h+m\phi_R) F_{UL}^{P_{\ell,m}\sin((1-m)\phi_h+m\phi_R)} \Bigg\}. \end{aligned} \tag{2.55}$$

The longitudinally polarized beam, longitudinally polarized target moments are

$$\begin{aligned} d\sigma_{LL} &= \frac{\alpha^2 M_h P_{h\perp}}{2\pi x y Q^2}\left(1+\frac{\gamma^2}{2x}\right) \lambda_e S_\parallel \\ &\times \sum_{\ell=0}^{2} \Bigg\{ C(x,y)\sum_{m=0}^{\ell} P_{\ell,m}\cos(m(\phi_h-\phi_R)) F_{LL}^{P_{\ell,m}\cos(m(\phi_h-\phi_R))} \\ &\quad + W(x,y)\sum_{m=-\ell}^{\ell} P_{\ell,m}\cos((1-m)\phi_h+m\phi_R) F_{LL}^{P_{\ell,m}\cos((1-m)\phi_h+m\phi_R)} \Bigg\}. \end{aligned} \tag{2.56}$$

The unpolarized beam, transversely polarized target moments are

$$\begin{aligned} d\sigma_{UT} &= \frac{\alpha^2 M_h P_{h\perp}}{2\pi x y Q^2}\left(1+\frac{\gamma^2}{2x}\right) |\boldsymbol{S}_\perp| \\ &\times \sum_{\ell=0}^{2}\sum_{m=-\ell}^{\ell} \Bigg\{ A(x,y)\Big[ P_{\ell,m}\sin((m+1)\phi_h-m\phi_R-\phi_S)) \\ &\qquad \times \left( F_{UT,T}^{P_{\ell,m}\sin((m+1)\phi_h-m\phi_R-\phi_S)} + \epsilon F_{UT,L}^{P_{\ell,m}\sin((m+1)\phi_h-m\phi_R-\phi_S)} \right) \Big] \\ &\quad + B(x,y)\Big[ P_{\ell,m}\sin((1-m)\phi_h+m\phi_R+\phi_S) F_{UT}^{P_{\ell,m}\sin((1-m)\phi_h+m\phi_R+\phi_S)} \\ &\qquad + P_{\ell,m}\sin((3-m)\phi_h+m\phi_R-\phi_S) F_{UT}^{P_{\ell,m}\sin((3-m)\phi_h+m\phi_R-\phi_S)} \Big] \\ &\quad + V(x,y)\Big[ P_{\ell,m}\sin(-m\phi_h+m\phi_R+\phi_S) F_{UT}^{P_{\ell,m}\sin(-m\phi_h+m\phi_R+\phi_S)} \\ &\qquad + P_{\ell,m}\sin((2-m)\phi_h+m\phi_R-\phi_S) F_{UT}^{P_{\ell,m}\sin((2-m)\phi_h+m\phi_R-\phi_S)} \Big] \Bigg\}. \end{aligned} \tag{2.57}$$

Lastly, the longitudinally polarized beam, transversely polarized target moments are

$$\begin{aligned} d\sigma_{LT} \;=\; & \frac{\alpha^2 M_h P_{h\perp}}{2\pi x y Q^2}\left(1+\frac{\gamma^2}{2x}\right)\lambda_e|\boldsymbol{S}_\perp|\sum_{\ell=0}^{2}\sum_{m=-\ell}^{\ell}\Bigg\{ \\ & C(x,y)P_{\ell,m}\cos((1-m)\phi_h+m\phi_R-\phi_S))F_{LT}^{P_{\ell,m}\cos((1-m)\phi_h+m\phi_R-\phi_S))} \\ & + W(x,y)\Bigg[P_{\ell,m}\cos(-m\phi_h+m\phi_R+\phi_S)F_{LT}^{P_{\ell,m}\cos(-m\phi_h+m\phi_R+\phi_S)} \\ & \quad + P_{\ell,m}\cos((2-m)\phi_h+m\phi_R-\phi_S)F_{LT}^{P_{\ell,m}\cos((2-m)\phi_h+m\phi_R-\phi_S)}\Bigg]\Bigg\}. \end{aligned} \tag{2.58}$$

In contrast with the 18 structure functions introduced in Ref. [16], dihadron production at twist-3 includes 162, exactly a factor of 9 more, as $2\oplus 1\oplus 0$ has dimension 9. However, the 18 structure functions involving a longitudinally polarized virtual photon are zero in dihadron production, just as the two single hadron structure functions involving a longitudinally polarized virtual photon are zero. Thus it is also common to write that unpolarized single hadron production has 16 moments, and dihadron production $16 \times 9 = 144$.

### 2.2.5 Structure Functions

Each structure function occurring in Equations 2.53 through 2.58 has the form of

$$F_{XY} = \mathcal{I}\left[wfD\right], \tag{2.59}$$

where $w$ is a $\boldsymbol{p}_T$, $\boldsymbol{k}_T$, $\phi_h$, $x$, $z$, $M$, $M_h$ dependent weight factor, $f$ is a distribution function and $D$ is a fragmentation function. The labeling scheme for $f$ and $D$ seems backwards, though it is given in this manner for historical reasons. Note that some structure functions include sums of several terms of this form. The operator $\mathcal{I}$ is shorthand for

$$\mathcal{I}\left[wfD\right] = \sum_q e_q^2 \int d^2\boldsymbol{p}_T d^2\boldsymbol{k}_T \delta^2\left(\boldsymbol{p}_T - \boldsymbol{k}_T - \frac{\boldsymbol{P}_{h\perp}}{z}\right) w f^q D^q. \tag{2.60}$$

Note that the distribution and fragmentation functions have flavor indices, but that the structure functions do not.

The weights $w$ will be written slightly different than in other sources. Typically the weights $w$ are written in terms of dot, cross or even wedge products of the vectors $\boldsymbol{p}_T$, $\boldsymbol{k}_T$, $\widehat{P}_{h\perp}$ [17, 16, 18]. However, this yields much more complicated expressions than is necessary. All weights can instead be written in terms of a factor involving possibly $|\boldsymbol{p}_T|/M$ and $|\boldsymbol{k}_T|/M_h$, multiplied by a single sine or cosine function of the involved angles. For example, the weight for the $\sin^2\theta\cos(4\phi_h - 2\phi_R)$ term in the unpolarized dihadron cross section is [18]

$$\begin{aligned} & \frac{\left[|\boldsymbol{k}_T|^2 - 4\left(\boldsymbol{k}_T\cdot\widehat{\boldsymbol{P}}_{h\perp}\right)\right]\left[(\boldsymbol{p}_T\cdot\boldsymbol{k}_T) - 4\left(\boldsymbol{p}_T\cdot\widehat{\boldsymbol{P}}_{h\perp}\right)\left(\boldsymbol{k}_T\cdot\widehat{\boldsymbol{P}}_{h\perp}\right)\right]}{2MM_h|\boldsymbol{k}_T|^2} \\ & - \frac{8\left(\boldsymbol{p}_T\cdot\widehat{\boldsymbol{P}}_{h\perp}\right)\left(\boldsymbol{k}_T\cdot\widehat{\boldsymbol{P}}_{h\perp}\right)^3}{2MM_h|\boldsymbol{k}_T|^2}, \end{aligned} \tag{2.61}$$

which can be written as

$$\frac{|\boldsymbol{p}_T||\boldsymbol{k}_T|}{MM_h}\cos(4\phi_h - \phi_p - 3\phi_k). \tag{2.62}$$

Furthermore, writing in terms of dot products also hides similarities between various moments. For instance, all the leading order, unpolarized terms involving the Boer-Mulder's function $h_1^\perp$ have a weight of the same form,

$$\frac{|\boldsymbol{p}_T||\boldsymbol{k}_T|}{MM_h}\cos((m-2)\phi_h + \phi_p + (1-m)\phi_k), \tag{2.63}$$

when written in terms of the cosine of the angles, rather than dot products. Written as dot products, the weights have very different form for each $m$. For example, see Appendix C of Ref. [18]. In particular, the $m$ in Equation 2.63 corresponds to the $m$ in the $|l, m\rangle$ angular momentum state of the dihadron. Thus writing in terms of sine or cosine functions not only makes the expressions simpler, this change also highlights deeper meanings and relationships between the structure functions.

The leading twist unpolarized beam, unpolarized target moments are

$$F_{UU,L}^{P_{\ell,m}\cos(m\phi_h - m\phi_R)} = 0, \tag{2.64}$$

$$F_{UU,T}^{P_{\ell,m}\cos(m\phi_h - m\phi_R)} = \begin{cases} \mathcal{I}\left[f_1 D_1^{|\ell,0\rangle}\right] & m = 0, \\ \mathcal{I}\left[2\cos(m\phi_h - m\phi_k) f_1 D_1^{|\ell,m\rangle +}\right] & m > 0, \end{cases} \tag{2.65}$$

$$\begin{aligned} F_{UU}^{P_{\ell,m}\cos((2-m)\phi_h + m\phi_R)} = & -\mathcal{I}\Bigg[\frac{|\boldsymbol{p}_T||\boldsymbol{k}_T|}{MM_h}\cos\big((m-2)\phi_h + \phi_p + (1-m)\phi_k\big) \\ & \times h_1^\perp H_1^{\perp|\ell,m\rangle}\Bigg], \end{aligned} \tag{2.66}$$

while the twist-3 structure functions are

$$\begin{aligned} F_{UU}^{P_{\ell,m}\cos((1-m)\phi_h + m\phi_R)} = & -\frac{2M}{Q}\mathcal{I}\Bigg[\frac{|\boldsymbol{k}_T|}{M_h}\cos((m-1)\phi_h + (1-m)\phi_k) \\ & \times \left(xhH_1^{\perp|\ell,m\rangle} + \frac{M_h}{M} f_1 \frac{\tilde{D}^{\perp|\ell,m\rangle}}{z}\right) \\ & + \frac{|\boldsymbol{p}_T|}{M}\cos((m-1)\phi_h + \phi_p - m\phi_k) \\ & \times \left(xf^\perp D_1^{|\ell,m\rangle} + \frac{M}{M_h} h_1^\perp \frac{\tilde{H}^{|\ell,m\rangle}}{z}\right)\Bigg]. \end{aligned} \tag{2.67}$$

The longitudinally polarized beam, unpolarized target structure functions are, at leading twist and for $\ell = 1, 2$ and $m > 0$,

$$F_{LU,T}^{P_{\ell,m}\sin(m\phi_h - m\phi_R)} = -\mathcal{I}\left[2\cos(m\phi_h - m\phi_k) f_1 D_1^{|\ell,m\rangle -}\right]. \tag{2.68}$$

The unpolarized beam, longitudinally polarized target structure functions are, at leading twist,

$$F_{UL}^{P_{\ell,m}\sin((2-m)\phi_h+m\phi_R)} = -\mathcal{I}\Bigg[\frac{|\boldsymbol{p}_T||\boldsymbol{k}_T|}{MM_h}\cos\big((m-2)\phi_h+\phi_p+(1-m)\phi_k\big) \times h_{1L}^{\perp}H_1^{\perp|\ell,m\rangle}\Bigg], \tag{2.69}$$

and additionally for $\ell = 1, 2$ and $m > 0$,

$$F_{UL}^{P_{\ell,m}\sin(m\phi_h-m\phi_R)} = -\mathcal{I}\left[2\cos(m\phi_h-m\phi_k)g_{1L}D_1^{|\ell,m\rangle-}\right]. \tag{2.70}$$

The longitudinally polarized beam, longitudinally polarized target moments are, at leading twist,

$$F_{LL}^{P_{\ell,m}\cos(m\phi_h-m\phi_R)} = \begin{cases}\mathcal{I}\left[g_{1L}D_1^{|\ell,0\rangle}\right] & m=0\\ \mathcal{I}\left[2\cos(m\phi_h-m\phi_k)g_{1L}D_1^{|\ell,m\rangle+}\right] & m>0.\end{cases} \tag{2.71}$$

The unpolarized beam, transversely polarized target moments are, at leading twist,

$$F_{UT,L}^{P_{\ell,m}\sin((m+1)\phi_h-m\phi_R-\phi_S)} = 0 \tag{2.72}$$

$$F_{UT,T}^{P_{\ell,m}\sin((m+1)\phi_h-m\phi_R-\phi_S)} = -\mathcal{I}\Bigg[\frac{|\boldsymbol{p}_T|}{M}\cos\big((m+1)\phi_h-\phi_p-m\phi_k\big) \times\Big(f_{1T}^{\perp}D_1^{|\ell,m\rangle+}+\mathrm{signum}[m]g_{1T}D_1^{|\ell,m\rangle-}\Big)\Bigg], \tag{2.73}$$

$$F_{UT}^{P_{\ell,m}\sin((1-m)\phi_h+m\phi_R+\phi_S)} = -\mathcal{I}\Bigg[\frac{|\boldsymbol{k}_T|}{M_h}\cos\big((m-1)\phi_h-\phi_p-m\phi_k\big)\ h_1H_1^{\perp|\ell,m\rangle}\Bigg], \tag{2.74}$$

$$F_{UT}^{P_{\ell,m}\sin((3-m)\phi_h+m\phi_R-\phi_S)} = \mathcal{I}\Bigg[\frac{|\boldsymbol{p}_T|^2|\boldsymbol{k}_T|}{M^2M_h}\cos\big((m-3)\phi_h+2\phi_p-(m-1)\phi_k\big) \times h_{1T}^{\perp}H_1^{\perp|\ell,m\rangle}\Bigg]. \tag{2.75}$$

The longitudinally polarized beam, transversely polarized target moments are, at leading twist,

$$F_{LT}^{P_{\ell,m}\cos((m+1)\phi_h-m\phi_R-\phi_S)} = -\mathcal{I}\Bigg[\frac{|\boldsymbol{p}_T|}{M}\cos\big((m+1)\phi_h-\phi_p-m\phi_k\big) \times\Big(g_{1T}D_1^{|\ell,m\rangle+}+\chi(m)f_{1T}^{\perp}D_1^{|\ell,m\rangle-}\Big)\Bigg]. \tag{2.76}$$

The twist-3 structure functions, except those for the unpolarized beam and unpolarized target, are not written out in terms of the distribution and fragmentation functions, as the specific formula are quite complex and results are not needed in the remainder of the dissertation.

It is worthy to note that the collinear fragmentation functions are simply integrals of specific partial waves of the TMD fragmentation functions. For example, the dihadron fragmentation function which has received the most interest [27, 40] occurs in the $F_{UT}^{\sin\vartheta\sin(\phi_R+\phi_S)}$ structure function, i.e. Equation 2.74 with $l = 1$ and $m = 2$. The specific fragmentation function, in both notations, is given later in Equation 2.90. Note, that the spin-flip inherent in the Collins functions requires a one unit change in the $z$-projections of angular momentum. Thus, the fragmentation function which survives in the collinear case is the one with $m = 1$ such that the $z$-projection of the angular momentum is zero. The $l = 2$, $m = 1$ partial wave in Equation 2.74 likewise survives in the collinear case for the identical reason. In a similar manner, the Boer-Mulders function also requires a spin-flip, and thus the partial wave of the Collins function that survives when paired with the Boer-Mulders function in Equation 2.66 are the $m = 2$ partial waves. A necessary, but not sufficient, power counting condition can be established. For a structure function that has a factor of $(p_T/M)^\alpha(k_T/M_h)^\beta$, one must have $|m| = |\alpha+\beta|$ if the structure function is to survive in the collinear case. Thus, the fragmentation functions in the structure functions surviving in the collinear case are not particularly special, but reflect the spin structure of a specific pair of distribution and fragmentation functions.

### 2.2.6 Relations with Previous Notation

The fragmentation functions occurring in the partial wave expansion of this dissertation represent a change of basis with respect to the common fragmentation functions defined in the literature. The relations between fragmentation functions in both notations can be determined by comparing the trace identities and expansion of this dissertation, specifically Equations 2.39, 2.40, 2.42, and 2.43, with the trace definitions and expansion common in the literature, specifically Equations 19-21 of Ref. [52] and Equation 57 of Ref. [18].[5] The results are

$$
\begin{aligned}
D_1^{|0,0\rangle} &= D_{1,OO} = \left(\frac{1}{4}D_{1,OO}^{s} + \frac{3}{4}D_{1,OO}^{p}\right), && (2.77)\\
D_1^{|1,0\rangle} &= D_{1,OL}, && (2.78)\\
D_1^{|1,\pm1\rangle} &= D_{1,OT} \mp \frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2}G_{1,OT}^{\perp}, && (2.79)\\
D_1^{|2,0\rangle} &= \frac{1}{2}D_{1,LL}, && (2.80)\\
D_1^{|2,\pm1\rangle} &= \frac{1}{2}\left(D_{1,LT} \mp \frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2}G_{1,LT}^{\perp}\right), && (2.81)\\
D_1^{|2,\pm2\rangle} &= D_{1,TT} \mp \frac{1}{2}\frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2}G_{1,TT}^{\perp}, && (2.82)
\end{aligned}
$$

---

[5] It should be noted two typographical errors exist in Equation 57 of Ref. [18]. On the line expanding $\bar{H}_1^{\sphericalangle}$, the factor of $\cos(2\phi_k - 2\phi_R)$ should in fact be $\cos(\phi_k - \phi_R)$, while on the line expanding $H_1^{\perp}$, the last symbol $H_{1,TT}^{\sphericalangle}$ should actually be $\bar{H}_{1,TT}^{\sphericalangle}$.

for the partial waves of the unpolarized distribution function. The relations for the combinations $D_1^{|\ell,m\rangle+}$, $D_1^{|\ell,m\rangle+}$ are

$$D_1^{|1,1\rangle+} = D_{1,OT}, \tag{2.83}$$

$$D_1^{|1,1\rangle-} = -\frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2} G_{1,OT}^{\perp}, \tag{2.84}$$

$$D_1^{|2,1\rangle+} = \frac{1}{2} D_{1,LT}, \tag{2.85}$$

$$D_1^{|2,1\rangle-} = -\frac{1}{2}\frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2} G_{1,LT}^{\perp}, \tag{2.86}$$

$$D_1^{|2,2\rangle+} = D_{1,TT}, \tag{2.87}$$

$$D_1^{|2,1\rangle-} = -\frac{1}{2}\frac{|\boldsymbol{k}_T|\,|\boldsymbol{R}|}{M_h^2} G_{1,TT}^{\perp}, \tag{2.88}$$

while for the partial waves of the Collins function,

$$H_1^{\perp|0,0\rangle} = H_{1,OO}^{\perp} = \frac{1}{4} H_{1,OO}^{\perp s} + \frac{3}{4} H_{1,OO}^{\perp p}, \tag{2.89}$$

$$H_1^{\perp|1,1\rangle} = H_{1,OT}^{\perp} + \frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} \bar{H}_{1,OT}^{\sphericalangle} = \frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} H_{1,OT}^{\sphericalangle} \tag{2.90}$$

$$H_1^{\perp|1,0\rangle} = H_{1,OL}^{\perp}, \tag{2.91}$$

$$H_1^{\perp|1,-1\rangle} = H_{1,OT}^{\perp}, \tag{2.92}$$

$$H_1^{\perp|2,2\rangle} = H_{1,TT}^{\perp} + \frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} \bar{H}_{1,TT}^{\sphericalangle} = \frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} H_{1,TT}^{\sphericalangle}, \tag{2.93}$$

$$H_1^{\perp|2,1\rangle} = \frac{1}{2} H_{1,LT}^{\perp} + \frac{1}{2}\frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} \bar{H}_{1,LT}^{\sphericalangle} = \frac{1}{2}\frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|} H_{1,LT}^{\sphericalangle}, \tag{2.94}$$

$$H_1^{\perp|2,0\rangle} = \frac{1}{2} H_{1,LL}^{\perp}, \tag{2.95}$$

$$H_1^{\perp|2,-1\rangle} = \frac{1}{2} H_{1,LT}^{\perp}, \tag{2.96}$$

$$H_1^{\perp|2,-2\rangle} = H_{1,TT}^{\perp}. \tag{2.97}$$

One point which is particularly subtle in the current literature is the difference between $H_{1,XT}^{\sphericalangle}$, $\bar{H}_{1,XT}^{\sphericalangle}$, and $H_{1,XT}^{\sphericalangle\prime}$, for $X{=}O,L,T$. Although the literature is not completely consistent, the notation of Ref. [18] is perhaps the most common, wherein the partial wave expansion and fragmentation correlator use the functions $\bar{H}_{1,XT}^{\sphericalangle}$, with the cross section written in terms of

$$H_{1,XT}^{\sphericalangle} = \bar{H}_{1,XT}^{\sphericalangle} + \frac{|\boldsymbol{k}_T|}{|\boldsymbol{R}|} H_{1,XT}^{\perp}. \tag{2.98}$$

The subtle difference in notation is difficult to see and already suggests that the choice of basis for the partial wave expansion in the literature is not optimal.

Note, the multiplicity of the $\ell = 0$ state discussed in 2.2.3 is reflected in the old notation in that the $|0,0\rangle$ states of both $D_1$ and $H_1^\perp$ have been written as a sum of $s$- and $p$-waves, as in Equation 2.77, 2.89, correspondingly Equations 39 and B5 of [18].

Using the above relations in notation, one can then compare the cross sections given in this dissertation and in published papers to ensure consistency. Such a process reveals several typographical errors in Equation C4 of Ref. [18], which can be shown to be related to inconsistencies within that document. Specifically, $-D_{1,OT}$, $-D_{1,LT}$, and $-D_{1,TT}$ need to be respectively replaced with $2D_{1,OT}$, $2D_{1,LT}$, and $2D_{1,TT}$.

One could likewise compare the cross section given in Ref. [28] with the above twist-3 structure functions to fully relate the higher twist notation. Such description is outside the scope of this dissertation, as herein it is only necessary understand the structure of the cross section at twist-3, not interpret any higher twist moments.

## 2.3 Model Predictions

### 2.3.1 The Lund/Artru Model

Models considered in Section 2.3 are relevant for SIDIS production of mesons, while previous portions of this chapter have considered SIDIS production of both mesons and baryons. It will be assumed that the parton struck by the virtual photon was a quark, rather than an anti-quark, but the conclusions follow identically for either case. It will be further assumed that the struck quark is transversely polarized, and thus the model predictions are related to the Collins fragmentation function. In cross sections, the Collins function is always paired with a distribution function involving a transversely polarized quark [16], e.g. transversity $h_1$, Boer-Mulders $h_1^\perp$, or pretzelocity $h_{1T}^\perp$. In at least the case of the transversity distribution function, one can further interpret the distribution function as the being related to the probability of the quark being the up state of transverse polarization, with the proton polarization defining the quantization axis.

The Lund/Artru string model of fragmentation [30] posits that the struck quark is initially connected with the target remnant via a gluon flux tube, or string. The Lund/Artru model further posits that, when the gluon flux tube breaks, the produced quark, anti-quark system has quantum numbers equal to that of the vacuum, $0^{++}$. This requires that the quark and anti-quark have their spins aligned, in order to have positive parity, and that the pair have one unit of orbital angular momentum in the opposite direction of the spin. One can quantize the system such that there are two cases: the anti-quark being produced with spin aligned or anti-aligned with that of the struck quark. From the point of view of the struck quark, i.e. facing in the direction of the quarks momentum with the transverse polarization in the up direction, the unit of angular momentum will cause the produced meson system to move preferentially to the left (right) in the case the quark spins are anti-aligned (aligned). This produces a left-right azimuthal asymmetry, with opposite signs depending on whether the anti-quark is aligned or anti-aligned with the fragmenting quark.

### 2.3.2 Relation to Amplitudes

Relative to the nucleon's transverse spin, the fragmenting quark can, in general, be in one of two spin states, though particular distribution functions may restrict the available spin states of the fragmenting quark. The two states are taken to be up or down, i.e. $\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$, and thus there are two possible diquark states with quarks anti-aligned, $\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \left|\frac{1}{2}, \mp\frac{1}{2}\right\rangle$ and two states with quarks aligned, $\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$. However, the direct product basis, where the spins of both quarks are specifically denoted, is not the basis in which measurable particles are diagonal. Instead, measurable hadrons are in the direct sum basis. Physicists often write the relation as $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$. The one spin-0 state is a pseudo-scalar meson, while three spin-1 states correspond to the three polarizations of vector mesons: $|1, 0\rangle$ being a longitudinal vector meson, and $|1, \pm 1\rangle$ being the two transverse polarizations.

The Clebsch-Gordan coefficients relating the two bases show that pseudo-scalar mesons are a symmetric combination of the two anti-aligned states, while longitudinal vector mesons correspond with the anti-symmetric combination of the two anti-aligned states. Transversely polarized vector mesons are uniquely the two aligned states, with no mixing. Note, as the model predictions are all relative to the struck quark polarization, there is no difference between the asymmetries for the two anti-aligned states $\left|\frac{1}{2}, \frac{1}{2}\right\rangle \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \left|\frac{1}{2}, \frac{1}{2}\right\rangle$. Thus, the anti-symmetric combination of the states should yield zero asymmetry. The final, amplitude level, prediction is then that the asymmetry for pseudo-scalar mesons has opposite sign to that for transversely polarized vector mesons (for each given flavor combination), while the asymmetry for longitudinal vector mesons is predicted to be zero. The model also predicts the sign, in that the pseudo-scalars should prefer left and the transverse vector mesons right, in regards to the left-right asymmetry from the point of view of the struck quark. Note, data related to the Collins asymmetry for pions [47, 60] is in agreement with the sign of this model.

### 2.3.3 Cross Section Level

To connect these amplitude level asymmetries with the Collins function, one needs to consider the cross section level, i.e. the sum of contributing amplitudes times the complex conjugate as shown in Figure 2.3. In other words, the Lund/Artru model makes predictions for the individual dihadrons, but the Collins function includes pairs of dihadrons.

Although not specifically developed by the Lund/Artru model, it will be assumed that the longitudinal $|1, 0\rangle |1, 0\rangle$ partial wave of the Collins function behaves as the model predicts the amplitude-level $|1, 0\rangle$ state to behave, i.e. that it is identically zero. Likewise, the Collins function for the two transverse partial waves $|1, \pm 1\rangle |1, \pm 1\rangle$ are assumed to have the oppose sign as the Collins function for pseudo-scalar meson production.

According to the Clebsch-Gordan coefficients, the two-dihadron (direct sum basis) states $|2, \pm 2\rangle$ are directly the squares of the transverse states, i.e.

$$|2, \pm 2\rangle \equiv |1, \pm 1\rangle |1, \pm 1\rangle . \tag{2.99}$$

Since each term in the cross section is related to a partial wave in the direct sum basis, and since the $\ell = 2$ state only occurs once in Equation 2.41, these partial waves have a clear access. In particular, neglecting possible twist-4 and higher effects, the structure functions

$F_{UT}^{P_{\ell,|m|}\sin((2-m)\phi_h+m\phi_R+\phi_S}$ in Equations 2.57 for $\ell = 2$, $m = \pm 2$, should have opposite sign as the corresponding pseudo-scalar Collins moments. Additionally, again neglecting twist-4 and higher effects, the unpolarized structure functions $F_{UU}^{P_{\ell,|m|}\sin((2-m)\phi_h+m\phi_R}$ in Equation 2.53 for $\ell = 2$, $m = \pm 2$ should have opposite sign as the Boer-Mulders moments of pseudo-scalar production.

For each of the above predictions regarding transverse vector mesons, the effect could be diluted by the presence of dihadrons, not arising from vector meson decay, yet in $|2, \pm 2\rangle$ partial waves. This dilution is assumed to be negligible, though no theoretical or experimental results specifically address this issue.

The longitudinal state $|1, 0\rangle\,|1, 0\rangle$ is a mixture of the $|2, 0\rangle$ state and the $|0, 0\rangle$ state arising from interference between vector meson polarization states. However, the $|0, 0\rangle$ partial wave of dihadrons not arising from vector mesons is known to be larger than the $|0, 0\rangle$ partial wave of vector meson production. Thus it is not possible to isolate the longitudinal $|1, 0\rangle\,|1, 0\rangle$ state and it is not possible to test this portion of the Lund/Artru model prediction.

### 2.3.4 The Gluon Radiation Model

It should be remembered that each fragmentation function (and each partial wave) depends on the flavor of the quark which is fragmenting. A common assumption is that the fragmentation functions for all flavors present in the observed hadron system are equal, which are collectively denoted the "favored" fragmentation function. For pion, $\rho$ meson, and pion-pair production, one generally also assumes that the contributions from $s$, $\bar{s}$, and heavier quarks are negligible. In this case, the fragmentation functions for flavors not present in the final state are also assumed equal and are denoted "disfavored" fragmentation functions.

It should be noted that the Lund/Artru model is only applicable when the struck quark is actually present in the observed hadron system, i.e. only for favored fragmentation functions. For disfavored fragmentation functions, another model is suggested by the author, denoted the "gluon radiation model." In this model, the struck quark emits a gluon in such a way that most of the momentum is transferred to the gluon. The struck quark becomes part of the remnant, and the gluon pair-produces to form a vector diquark. In the case of pseudo-scalar meson production, the diquark must interact further with the remnant in order to become a pseudo-scalar. In the case of vector meson production, this diquark can directly form the vector meson, and it will be assumed for vector meson production that further interactions with the remnant are higher order effects and negligible.

In cases where the diquark does interact further with the remnant, the gluon radiation model can be described within the paradigm of the Lund/Artru model, i.e. by considering fragmentation in terms of a struck quark and a gluon flux tube. In contrast to the the Lund/Artru model, the gluon radiation model considers the case where most of momentum of the struck quark is transferred to the pair-produced quark anti-quark system. Rather than requiring this quark anti-quark system to have the quantum numbers of the vacuum, the gluon radiation model has the system hadronize into the observed final state and thus have the quantum numbers of the final state. The portion of the gluon flux tube between the break point and the remnant is interpreted as an interaction between the pair produced system and the remnant, required for pseudo-scalar production but a higher order effect for vector meson production.

In both models, the anti-quark which is produced is present in the final state. However, the main difference between the models is which quark joins the produced anti-quark in final state: in one case, the stuck quark, and in the other case, the produced quark. In this manner, the Lund/Artru model describes favored fragmentation, and the gluon radiation model describes disfavored fragmentation.

One can then consider the relation, predicted by the gluon radiation model, between the disfavored fragmentation functions for vector meson production and the fragmentation functions for pseudo-scalar production. Comparing the Feynman diagrams for both models, as in Figure 2.4, one can determine that the perturbative portion of the gluon radiation model is present in the diagram for the Lund/Artru model, though the specific quarks which form the observed hadron is different in the two cases.

Assuming that the vector meson is produced in the $|1,1\rangle$ state, the emission of the gluon must flip the struck quark from the $\left|\frac{1}{2},\frac{1}{2}\right\rangle$ to $\left|\frac{1}{2},-\frac{1}{2}\right\rangle$ spin states. However, this diagram is exactly the initial part of the relevant diagram for the Lund/Artru model for the spin anti-aligned case, i.e. the one related to pseudo-scalar production. Thus the disfavored $|2,2\rangle$ Collins function is expected to have opposite sign as the favored pseudo-scalar Collins function. In a similar manner, the production of vector mesons in the $|1,0\rangle$ state is related to the antisymmetric combination of the spin aligned and the spin anti-aligned cases of the Lund/Artru diagram, and thus expected to be zero. However, it has already been discussed that predictions for the vector meson $|1,0\rangle$ state are not experimentally accessible.

### 2.3.5 Summary of Model Predictions

The following summary combines the results of both models and the discussion of amplitudes versus cross sections. The Lund/Artru model predicts that the $|2,2\rangle$ partial wave of the Collins function for SIDIS vector meson production, for quark flavors present in the produced vector meson, has the opposite sign as the respective pseudo-scalar Collins function. The gluon radiation model implicates that the $|2,2\rangle$ partial wave of the Collins function for SIDIS vector meson production, for quark flavors not present in the produced vector meson, also has the opposite sign as the respective pseudo-scalar Collins function. Note, while data [47] suggests that the pseudo-scalar favored and disfavored Collins functions are nearly equal and opposite, these models predict that in the vector meson case the favored and disfavored Collins functions have the same sign, for at least the $|2,\pm 2\rangle$ moments. Although the models yield expectations regarding the signs of certain partial waves, neither model addresses the relative size of the $|2,2\rangle$ versus $|2,-2\rangle$ partial waves, nor their relative size compared with the pseudo-scalar Collins function.

It should be noted that the sign change of the $|2,2\rangle$ moment for the Collins function (based on the Lund/Artru model) is truly a prediction. In contrast, the gluon radiation model was developed after preliminary results were first available, and thus the expectations of the gluon radiation model are more explanations of possible results rather than true predictions.

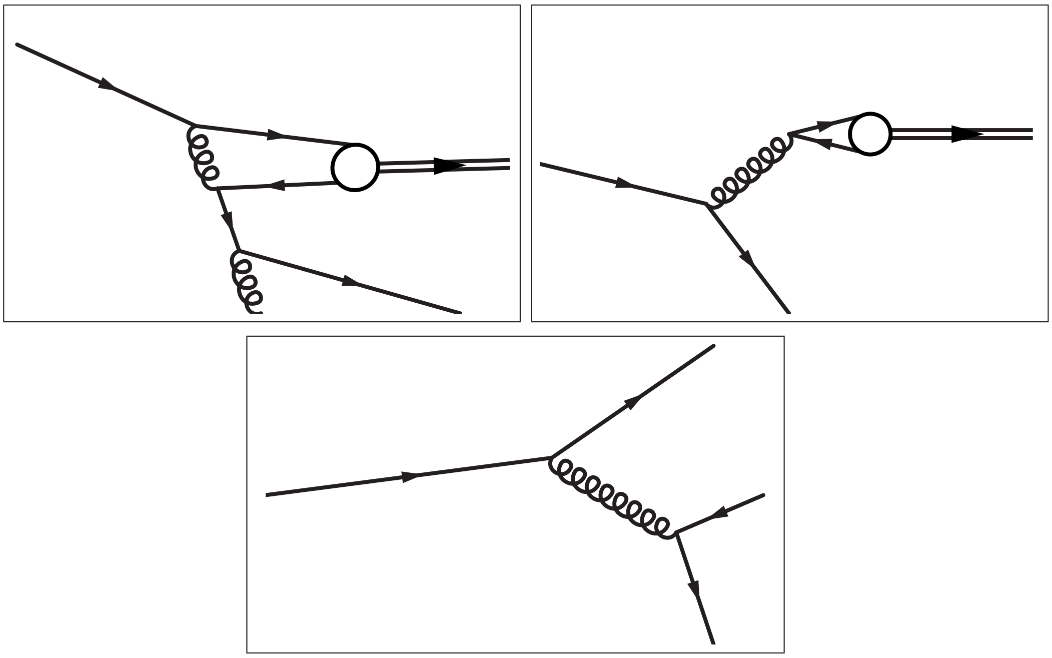


Figure 2.4: Feynman diagrams relevant for the Lund/Artru and Gluon Radiation models. In all panels, the quark entering from the right is the struck quark. For the upper panels, the particles connected to the bottom of the diagram are interactions with the target remnant. The open circle represents the non-perturbative hadronization process of the quark anti-quark pair forming a hadron. The upper left panel represents the Lund/Artru model and the upper right panel, the gluon radiation model. The lower panel is the perturbative sub-diagram, common to both diagrams. The difference between the Lund/Artru model and gluon radiation model is related to which quark returns to the remnant and which enters in the observed hadron. Note, the Artru model also requires an extra gluon emission, related to one half of the broken gluon flux tube.

## 2.4 Spectator Model of Dihadron Fragmentation

The only published models for dihadron fragmentation [31, 61] are spectator models. Reference [61] is designed for a pion, proton dihadron, including the Roper resonance, and is not directly applicable to the scope of this dissertation. Reference [31], however, is designed to be directly comparable with the previously published Hermes dihadron result [27]. This section, Section 2.4 is an extension of the work done in Ref. [31].

### 2.4.1 Fragmentation Correlation Matrix

Following Ref. [31], the $X$ of Equation 2.2 is replaced with a single, on-shell particle, "the spectator," with quantum numbers equal to that of the target. Unfortunately, Ref. [31] lacks three important aspects: 1) the results are all integrated over transverse momenta, 2) the model is only for $\pi^+\pi^-$ pairs, 3) the $l = 2$ states for the Collins function $H_1^{\perp|l,m\rangle}$ are all zero. While solving the third point, i.e. developing a model with non-zero $l = 2$ states for the

Collins function, is outside the scope of this dissertation, the first two deficiencies are solved in this section, Section 2.4. The TMD fragmentation functions can be computed by starting with the same fragmentation correlation function but without introducing the integration over transverse momentum. Additional final states can be included by modifying the vertex function and allowing the parameters to depend on the quark flavor.

A certain subtlety regarding the choice of using $M_h$ or $M_h^2$ as a dependent variable is worth a comment. When using $M_h$ instead of $M_h^2$, an extra factor of $2M_h$ must be included. The subtlety is whether this factor is included in the definition of the fragmentation functions, or is considered part of the overall phase space factor for the cross section. In this work, the fragmentation functions are defined to be differential with respect to $M_h$, and the extra factor of $M_h^2$ is included in the phase space factor of the cross section.

The specific expression for the correlator in this model is given in Equation 19 of Ref. [31]. Setting the quark mass to zero (as it done later in Ref. [31]), the correlator is

$$\begin{aligned}\Delta^q(k, P_h, R) \; = \; & \Big\{ |F^s|^2 \, e^{-2\frac{k^2}{\Lambda_s^2}} \not{k} \left( \not{k} - \not{P}_h + M_s \right) \not{k} \\ & + |F^p|^2 \, e^{-2\frac{k^2}{\Lambda_p^2}} \not{k}\not{R} \left( \not{k} - \not{P}_h + M_s \right) \not{R}\not{k} \\ & + F^{s*} F^p e^{-2\frac{k^2}{\Lambda_{sp}^2}} \not{k} \left( \not{k} - \not{P}_h + M_s \right) \not{R}\not{k} \\ & + F^s F^{p*} e^{-2\frac{k^2}{\Lambda_{sp}^2}} \not{k}\not{R} \left( \not{k} - \not{P}_h + M_s \right) \not{k} \Big\} \\ & \times \frac{1}{(2\pi)^3} \frac{1}{k^4} \delta\left( (k - P_h)^2 - M_s^2 \right) e^{-2\frac{k_T^2}{\Lambda_b^2}}, \end{aligned} \tag{2.100}$$

where $M_s$ is the mass of the spectator, $F^{s/p}$ are the vertex functions, and $\Lambda_{s,p,sp,b}$ are inverse slopes for the $k^2$ and $k_T^2$ cut-offs. Note, this correlator has an extra $k_T^2$-cutoff, with inverse slope $\Lambda_b$, than that in Ref. [31]. The extra factor was found to be necessary in numerical studies, and the inverse slope is given a subscript $b$ designating that it affects both $s$ and $p$ waves, and likewise their interference.

Implicit in Equation 2.100 is the assumption that the masses of both hadrons $h_1$ and $h_2$ are equal, i.e. $m_1 = m_2$. In case $m_1 \neq m_2$, the dihadron propagator needs to include additional terms. Effectively, every $\not{R}$ in Equation 2.100 would be replaced with a new quantity. Such computation is outside the scope of this dissertation, and thus the computations of this section, Section 2.4, are only be applicable to producing hadrons with approximately equal mass. Thus, it will be assumed $m_1 = m_2$, and the hadron mass is thus defined as $m := m_1$.

The inverse slopes in Equation 2.100 are parametrized as

$$\Lambda_{s,p,b} = \alpha_{s,p,b} z^{\beta_{s,p,b}} (1 - z)^{\gamma_{s,p,b}}, \tag{2.101}$$

$$\frac{2}{\Lambda_{sp}^2} = \frac{1}{\Lambda_s^2} + \frac{1}{\Lambda_p^2}. \tag{2.102}$$

The exponential form of the cutoff makes the inverses $\Lambda^{-1}$ analogous to the $b$-slope parameters in exclusive vector meson production [62], and are also related to the variance parameter in the so called "Gaussian Ansatz" for the $k_T$-dependence of the fragmentation functions.

The scalar vertex function is just a constant, $F^s = f_s$, one of the parameters of the model. The $p$-wave vertex function for pion-pion dihadrons is defined as

$$\begin{aligned}
F^p &= f_\rho \frac{\left(M_h^2 - M_\rho^2\right) - i\Gamma_\rho M_\rho}{\left(M_h^2 - M_\rho^2\right)^2 + \Gamma_\rho^2 M_\rho^2} + f_\omega \frac{\left(M_h^2 - M_\omega^2\right) - i\Gamma_\omega M_\omega}{\left(M_h^2 - M_\omega^2\right)^2 + \Gamma_\omega^2 M_\omega^2} \\
&\quad - i f'_\omega \frac{\sqrt{\lambda\left(M_\omega^2, M_h^2, m_\pi^2\right)}\Theta\left(M_\omega - m_\pi - M_h\right)}{4\pi\Gamma_\omega M_\omega^2 \sqrt[4]{4M_\omega^2 m_\pi^2 + \lambda\left(M_\omega^2, M_h^2, m_\pi^2\right)}}, \qquad (2.103) \\
\lambda\left(M_\omega^2, M_h^2, m_\pi^2\right) &= \left[M_\omega^2 - (M_h + m_\pi)\right]\left[M_\omega^2 - (M_h + m_\pi)\right], \qquad (2.104)
\end{aligned}$$

while for kaon-kaon dihadrons it is

$$F^p = f_\phi \frac{\left(M_h^2 - M_\phi^2\right) - i\Gamma_\phi M_\phi}{\left(M_h^2 - M_\phi^2\right)^2 + \Gamma_\phi^2 M_\phi^2}. \qquad (2.105)$$

The spectator mass $M_s$ is assumed proportional to $M_h$. There are a total of 14 free parameters thus far for pion-pion dihadrons, and two less for kaon-kaon dihadrons, since the quark mass is fixed at zero. The parameters are specifically the 9 parameters $\alpha_{s,p,b}$, $\beta_{s,p,b}$, $\gamma_{s,p,b}$, involved in the $z$ dependence of the $\Lambda$ slopes, the spectator mass factor, and the couplings $f_s$, $f_\rho$, $f_\omega$, $f_{\omega'}$, and $f_\phi$.

Using the properties of gamma matrices, the correlator can be rewritten as

$$\begin{aligned}
\Delta^q(k, P_h, R) &= \Bigg\{ |F^s|^2 e^{-2\frac{k^2}{\Lambda_s^2}} \Big[M_s k^2 + \Big(k^2 - 2(k \cdot P_h)\Big)\not{k} + k^2 \not{P}_h\Big] \\
&\quad + |F^p|^2 e^{-2\frac{k^2}{\Lambda_p^2}} \Big[M_s k^2 R^2 - 2k^2\Big((R \cdot k) - (R \cdot P_h)\Big)\not{R} - k^2 R^2 \not{P}_h \\
&\qquad + \Big(4(R \cdot k)\Big((R \cdot k) - (R \cdot P_h)\Big) + 2R^2(k \cdot P_h) - k^2 R^2\Big)\not{k}\Big] \\
&\quad + 2\mathrm{Re}\left[F^{s*} F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}} M_s \Big[2(k \cdot R)\not{k} - k^2 \not{R}\Big] \\
&\quad + 2\mathrm{Im}\left[F^s F^{p*}\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}} \Big[\Big(k^2 - 2(k \cdot P_h)\Big)\not{R}\not{k} + 2(R \cdot k)\not{P}_h\not{k} - k^2 \not{P}_h \not{R}\Big] \\
&\quad + \left(F^s F^{p*}\right) e^{-2\frac{k^2}{\Lambda_{sp}^2}} \Big[2\Big((R \cdot k) - (R \cdot P_h)\Big)k^2\Big]\Bigg\} \\
&\quad \times \frac{1}{(2\pi)^3} \frac{1}{k^4} \delta\left((k - P_h)^2 - M_s^2\right). \qquad (2.106)
\end{aligned}$$

Note, when applying the integral in Equation 2.100, the $\delta$-function in of Equation 2.106 forces the on-shell condition of $k^2$,

$$k^2 = \frac{z}{1-z}|\boldsymbol{k}_T|^2 + \frac{M_s^2}{1-z} + \frac{M_h^2}{z}, \qquad (2.107)$$

and introduces an extra face space factor of $(2zP_h^-)^{-1}$. Equation 20 of Ref. [31] can be derived from the above Equation 2.106 by integrating over $d^2\boldsymbol{k}_T$ and noting

$$d^2\boldsymbol{k}_T = |\boldsymbol{k}_T| d|\boldsymbol{k}_T| d\phi_k = \frac{d|\boldsymbol{k}_T^2|}{2} d\phi_k. \qquad (2.108)$$

Equation 20 of Ref. [31] additionally replaces $d\phi_k$ with $2\pi$, and has evaluated the $\delta$-function and thus includes the on-shell condition and extra phase space factor.

### 2.4.2 Model Prediction for Fragmentation Functions

The fragmentation functions are computed according to the trace identities given in Equations 2.39 and 2.40. Expressions for these fragmentation functions can then be expanded in partial waves via Equations 2.42 and 2.43, to yield the functions $D_1^{|l,m\rangle}$ and $H_1^{\perp|l,m\rangle}$.

Substituting Equation 2.106 into Equation 2.39 yields the model estimate

$$\begin{aligned} D_1 \;=\;& \frac{|\boldsymbol{R}|}{16\pi^2 M_h k^4}\Bigg\{ |F^s|^2\, e^{-2\frac{k^2}{\Lambda_s^2}} \Big[(1+z)k^2 - 2(k\cdot P_h)\Big] \\ &+ \mathrm{Re}\left[F^{s*}F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}} M_s \Big[4(k\cdot R) - z\zeta k^2\Big] \\ &+ |F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}} \Bigg[\Big(4(R\cdot k) - z\zeta k^2\Big)\Big((R\cdot k) - (R\cdot P_h)\Big) \\ &\qquad - R^2\Big((1+z)k^2 - 2(k\cdot P_h)\Big)\Bigg]\Bigg\}, \end{aligned} \tag{2.109}$$

dependent on $z$, $\boldsymbol{k}_T$, $M_h$, $\cos\vartheta$, $(\phi_R - \phi_k)$. Note that $D_1$ can be written as

$$D_1 \;=\; \tau_{ss}\,|F^s|^2\, e^{-2\frac{k^2}{\Lambda_s^2}} + \tau_{sp} M_s \mathrm{Re}\left[F^{s*}F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}} + \tau_{pp}\,|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}, \tag{2.110}$$

with

$$\tau_{ss} \;=\; \frac{|\boldsymbol{R}|}{16\pi^2 M_h k^4}\Big((1+z)k^2 - 2(k\cdot P_h)\Big), \tag{2.111}$$

$$\tau_{sp} \;=\; \frac{|\boldsymbol{R}|}{16\pi^2 M_h k^4}\Big(4(k\cdot R) - z\zeta k^2\Big), \tag{2.112}$$

$$\tau_{pp} \;=\; \Big((R\cdot k) - (R\cdot P_h)\Big)\tau_{sp} - R^2\tau_s. \tag{2.113}$$

It is useful to note the following products [18]

$$R\cdot P_h \;=\; \frac{m_1^2 - m_2^2}{2}, \tag{2.114}$$

$$k\cdot P_h \;=\; \frac{M_h^2}{2z} + z\frac{k^2 + |\boldsymbol{k}_T|^2}{2}, \tag{2.115}$$

$$k\cdot R \;=\; \frac{1}{z}\left(P_h\cdot R\right) + \frac{\zeta}{2}\left(P_h\cdot k\right) - |\boldsymbol{k}_T||\boldsymbol{R}_T|\cos(\phi_R - \phi_k), \tag{2.116}$$

$$P_h^2 \;=\; M_h^2, \tag{2.117}$$

$$R^2 \;=\; \frac{m_1^2 + m_2^2}{2} - \frac{M_h^2}{4}, \tag{2.118}$$

intermediate quantities,

$$|\boldsymbol{R}_T| \;=\; |\boldsymbol{R}|\sin\vartheta, \tag{2.119}$$

$$|\boldsymbol{R}|^2 \;=\; \frac{M_h^2}{4} - \frac{m_1^2 + m_2^2}{2} + \frac{\left(m_1^2 - m_2^2\right)^2}{4M_h^2}, \tag{2.120}$$

$$\zeta \;=\; \frac{1}{M_h}\left(\sqrt{m_1^2 + |\boldsymbol{R}|^2} - \sqrt{m_2^2 + |\boldsymbol{R}|^2 - 2|\boldsymbol{R}|\cos\vartheta}\right), \tag{2.121}$$

and light cone coordinates

$$
\begin{aligned}
zk^- &= \frac{2}{\zeta}R^- = P_h^-, \qquad &(2.122)\\
R^+ &= \frac{1}{P_h^-}\left(m_1^2 - m_2^2 - \frac{\zeta}{2}M_h^2\right), \qquad &(2.123)\\
k^+ &= \frac{z\left(|\boldsymbol{k}_T|^2 + k^2\right)}{2P_h^-}, \qquad &(2.124)\\
P_h^+ &= \frac{M_h}{2P_h^-}. \qquad &(2.125)
\end{aligned}
$$

Note that $R^2$ is the Lorentz invariant $R^2 := R^\mu R_\mu$, while $|\boldsymbol{R}|$ is the magnitude of the spatial components of $R^\mu$, and $|\boldsymbol{R}_T|$ is the magnitude of the transverse ($x$, $y$) components of $\boldsymbol{R}$.

Applying the partial wave expansion of Equation 2.42 to $D_1$ in Equation 2.109 yields the expressions

$$
\begin{aligned}
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|0,0\rangle} &= \left(\frac{z^2|\boldsymbol{k}_T|^2 + M_s^2}{1-z}\right)\left[|F^s|^2\, e^{-2\frac{k^2}{\Lambda_s^2}} - R^2\,|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right]\\
&\quad + \left(\Upsilon_{sp}^{D_1^{|0,0\rangle}}\right)\left[\mathrm{Re}\left[F^{s*}F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right]\\
&\quad + \left(\Upsilon_{pp}^{D_1^{|0,0\rangle}}\right)\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right], \qquad (2.126)\\
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|1,1\rangle} &= -2M_s|\boldsymbol{R}||\boldsymbol{k}_T|\left[\mathrm{Re}\left[F^{s*}F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right]\\
&\quad + \left(\Upsilon_{pp}^{D_1^{|1,1\rangle}}\right)\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right], \qquad (2.127)\\
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|1,0\rangle} &= -2\frac{M_s|\boldsymbol{R}|}{zM_h}\left(M_h^2 + z^2|\boldsymbol{k}_T|^2\right)\left[\mathrm{Re}\left[F^{s*}F^p\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right]\\
&\quad + \left(\Upsilon_{pp}^{D_1^{|1,0\rangle}}\right)\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right], \qquad (2.128)\\
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|2,2\rangle} &= \frac{1}{3}|\boldsymbol{k}_T|^2|\boldsymbol{R}|^2\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right], \qquad (2.129)\\
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|2,1\rangle} &= \frac{2}{3}\frac{|\boldsymbol{k}_T||\boldsymbol{R}|^2}{zM_h}\left(M_h^2 + z^2|\boldsymbol{k}_T|^2 + \frac{1}{2}z^2k^2\right)\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right], \\
&\qquad (2.130)\\
\frac{16\pi^2 M_h k^4}{|\boldsymbol{R}|}D_1^{|2,0\rangle} &= \left(\frac{2}{3}\frac{|\boldsymbol{R}|^2}{z^2M_h^2}\left(M_h^2 + z^2|\boldsymbol{k}_T|^2\right)\left(M_h^2 + z^2|\boldsymbol{k}_T|^2 + z^2k^2\right)\right.\\
&\quad \left. - \frac{4}{3}|\boldsymbol{k}_T|^2|\boldsymbol{R}|^2\right)\left[|F^p|^2\, e^{-2\frac{k^2}{\Lambda_p^2}}\right]. \qquad (2.131)
\end{aligned}
$$

with $D_1^{|l,-m\rangle} = D_1^{|l,m\rangle}$ in this model, and $\Upsilon$ functions

$$\begin{aligned}
\Upsilon_{sp}^{D_1^{|0,0\rangle}} &= \left(\frac{M_s}{z}\right)\Bigg[M_h\left(1+\frac{z^2|\boldsymbol{k}_T|^2}{M_h^2}\right)\left(\sqrt{m_1^2+|\boldsymbol{R}|^2}-\sqrt{m_2^2+|\boldsymbol{R}|^2}\right) \\
&\quad + 2\left(m_1^2-m_2^2\right)\Bigg], \qquad (2.132)
\end{aligned}$$

$$\begin{aligned}
\Upsilon_{pp}^{D_1^{|0,0\rangle}} &= \frac{1}{4z^2M_h^2}\Bigg(\left(\left(\sqrt{m_1^2+|\boldsymbol{R}|^2}-\sqrt{m_2^2+|\boldsymbol{R}|^2}\right)^2+\frac{4}{3}|\boldsymbol{R}|^2\right)\left(M_h^2+z^2|\boldsymbol{k}_T|^2\right) \\
&\quad \times\left(M_h^2+z^2|\boldsymbol{k}_T|^2+z^2k^2\right)+\frac{m_1^2-m_2^2}{2z^2M_h}\left(\sqrt{m_1^2+|\boldsymbol{R}|^2}-\sqrt{m_2^2+|\boldsymbol{R}|^2}\right) \\
&\quad \times\left(z^2k^2+(5-4z)\left(M_h^2+z^2|\boldsymbol{k}_T|^2\right)\right)+2\left(\frac{1-z}{z}\right)\left(m_1^2-m_2^2\right)^2 \\
&\quad + \frac{4}{3}|\boldsymbol{k}_T|^2|\boldsymbol{R}|^2, \qquad (2.133) \\
\Upsilon_{pp}^{D_1^{|1,1\rangle}} &= -\frac{1}{z}|\boldsymbol{k}_T||\boldsymbol{R}|\Bigg((5-4z)\left(m_1^2-m_2^2\right)+\frac{1}{M_h}\left(\sqrt{m_1^2+|\boldsymbol{R}|^2}-\sqrt{m_2^2+|\boldsymbol{R}|^2}\right) \\
&\quad \times\left(M_h^2+z^2|\boldsymbol{k}_T|^2+\frac{1}{2}z^2k^2\right)\Bigg), \qquad (2.134) \\
\Upsilon_{pp}^{D_1^{|1,0\rangle}} &= -\frac{|\boldsymbol{R}|}{2z^2M_h^2}\left(M_h^2+z^2|\boldsymbol{k}_T|^2+z^2k^2\right)\Bigg[\left(\sqrt{m_1^2+|\boldsymbol{R}|^2}-\sqrt{m_2^2+|\boldsymbol{R}|^2}\right) \\
&\quad \times\left(M_h^2+z^2|\boldsymbol{k}_T|^2\right)+2M_h\left(m_1^2-m_2^2\right)\Bigg]. \qquad (2.135)
\end{aligned}$$

The $\Upsilon$ functions depend on the mass difference between the two produced hadrons ($m_1-m_2$), and converge to zero when the difference is much smaller than the other mass scales ($M_h$, $|\boldsymbol{R}|$, $k$). Although, earlier it was assumed $m_1 = m_2$, the above $\Upsilon$ functions give a part of the needed correction for non-equal masses.

Proceeding in a similar manner for $H_1^{\perp}$, i.e. substituting Equation 2.100 into Equation 2.40, yields the model estimate

$$\begin{aligned}
H_1^{\perp} &= \frac{|\boldsymbol{R}|}{8\pi^2k^4}\Bigg(\left(k^2-2(k\cdot P_h)\right)R^+ + 2(R\cdot k)P_h^+ \\
&\quad -\left(k^2P_h^+ + \left(k^2-2(k\cdot P_h)\right)k^+\right)\frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|}\sin\vartheta e^{i(\phi_R-\phi_k)}\Bigg)\left[\mathrm{Im}\left[F^{s*}F^p\right]e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right], \qquad (2.136)
\end{aligned}$$

again dependent on $z$, $\boldsymbol{k}_T$, $M_h$, $\cos\vartheta$, $(\phi_R - \phi_k)$. The partial wave expansion is

$$
\begin{aligned}
\frac{8\pi^2 k^4}{|\boldsymbol{R}|} H_1^{\perp|0,0\rangle} &= \frac{1}{2z}\Bigg\{ 2\Big(z(1-z)k^2 - z^2|\boldsymbol{k}_T|^2\Big)\Big(m_1^2 - m_2^2\Big) \\
&\quad + M_h\left(2M_h^2 + z\,(2z-1)\,k^2 + 2z^2|\boldsymbol{k}_T|^2\right) \\
&\quad \times \left(\sqrt{m_1^2+|\boldsymbol{R}|^2} - \sqrt{m_2^2+|\boldsymbol{R}|^2}\right)\Bigg\} \\
&\quad \times \left[\mathrm{Im}\left[F^{s*}F^{p}\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right], \qquad (2.137) \\
\frac{8\pi^2 k^4}{|\boldsymbol{R}|} H_1^{\perp|1,1\rangle} &= -\frac{|\boldsymbol{R}|}{|\boldsymbol{k}_T|}\Big(k^2 + |\boldsymbol{k}_T|^2\Big)\Big(\left(1-z^2\right)k^2 - z^2|\boldsymbol{k}_T|^2\Big) \\
&\quad \times \left[\mathrm{Im}\left[F^{s*}F^{p}\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right], \qquad (2.138) \\
\frac{8\pi^2 k^4}{|\boldsymbol{R}|} H_1^{\perp|1,0\rangle} &= \frac{1}{z} M_h |\boldsymbol{R}| \Big( zk^2 - 2\left(M_h^2 + z^2(k^2 + |\boldsymbol{k}_T|^2)\right)\Big) \\
&\quad \times \left[\mathrm{Im}\left[F^{s*}F^{p}\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right], \qquad (2.139) \\
\frac{8\pi^2 k^4}{|\boldsymbol{R}|} H_1^{\perp|1,-1\rangle} &= -M_h^2 |\boldsymbol{R}||\boldsymbol{k}_T| \left[\mathrm{Im}\left[F^{s*}F^{p}\right] e^{-2\frac{k^2}{\Lambda_{sp}^2}}\right]. \qquad (2.140)
\end{aligned}
$$

Note that $H_1^{\perp|0,0\rangle}$ is zero in the case of $m_1 = m_2$.

### 2.4.3 Flavor Dependence

Note that the fragmentation functions computed in Section 2.4.2 are all based on the correlator in Equation 2.100 which has a flavor index. The flavor index is not carried through in the subsequent equations for sake of brevity, as the fragmentation functions for each flavor have identical form. The possible difference in the flavors is reflected in the ability to make different choices for the parameter per each flavor, as well as a possible change in sign of the 3-vector $\boldsymbol{R}$.

In applying the model to $\pi^+\pi^-$-dihadrons, which includes SIDIS $\rho^0$ production, it is sufficient to use two sets of parameters, as isospin relations give that $u \to \pi^+\pi^- X$ has the same correlator as $\bar{d} \to \pi^+\pi^- X$, $d \to \pi^-\pi^+ X$, $\bar{u} \to \pi^-\pi^+ X$. CP symmetry (replacing a quark with its anti-quark partner) also implies that the correlator of $s \to \pi^+\pi^- X$ is equal to that for $\bar{s} \to \pi^-\pi^+ X$. Thus one has a set of parameters for $u$ and $s$, and all others flavors are either equal or equivalent up to the sign of the $\boldsymbol{R}$.

For $\pi^+\pi^0$ production, one again has the flavors $u$ and $\bar{d}$ equivalent, and separately $d$ and $\bar{u}$ are equivalent, though isospin symmetry gives no relation between $u$ and $d$. One call also posit non-zero parameters for $s$ flavor, in which case $\bar{s}$ is again related to $s$ by changing the sign of $\boldsymbol{R}$. Given the parameters for $\pi^-\pi^0$ production, isospin symmetry (equivalently CP symmetry) specifies all the parameters for $\pi^+\pi^0$.

For $K^+K^-$ production, one again can use three sets of parameters, $u$, $d$, $s$, and the anti-quark flavors are each related to their conjugate via changing the sign of $\boldsymbol{R}$.

### 2.4.4 Numerical Results

Specific numerical results are computed using the `TMDGen` Monte Carlo generator, described later in Section 3.2. Thus, the numeric results for the above computed model, as well as the optimized parameter sets, are given in Section 3.2.5.

## CHAPTER III

# Numerical Methods

Two essential numerical tools are covered in this chapter. First, Section 3.1 describes a smearing and acceptance correction method, while the new Monte Carlo generator `TMDGen` is described in Section 3.2. A significant component of `TMDGen` is the inversion of the equations of Section 2.1.4, such that particle momenta can be determined from the cross section variables. It is hoped that the `TMDGen` will not only be useful for the analysis in this dissertation, but will also be useful for others analyses at various experiments and for theorists interpreting various SIDIS results.

## 3.1 Acceptance and Smearing Corrections

While theoretically one is interested in the true angular distribution of the produced particles, the HERMES detector does not have a full $4\pi$ angular acceptance. Additionally, higher order QED effects, *brehmsstrahlung* and detector resolution can smear the values of the measured variables. Estimating the true angular moments, given smeared data within acceptance, is the subject of this section. Note, for the actual dihadron analysis of this dissertation, smearing effects are negligible. However, the following method is applicable either with our without the presence of significant smearing. The methods described in this section, Section 3.1, are based on the authors work in Refs. [63, 64]. Furthermore, the method utilized in Ref. [65] can be shown to be a special case of the method described in Section 3.1.2.

### 3.1.1 Fredholm Integral Equation

To quantitatively discuss the effect of acceptance and smearing, it is necessary to introduce three quantities: $p\left(\boldsymbol{x}^{(T)}\right)$, the true distribution, i.e. the distribution of the true variables, as one would measure with a perfect, $4\pi$ detector; $p\left(\boldsymbol{x}^{(R)}\right)$, the reconstructed distribution or the distribution of the reconstructed variables; and $p\left(\boldsymbol{x}^{(R)}\middle|\,\boldsymbol{x}^{(T)}\right)$, the conditional probability of reconstructing certain values of the variables given their true value. The $D$-dimensional vectors $\boldsymbol{x}^{(T)}$, $\boldsymbol{x}^{(R)}$ include the $D$ parameters with which the cross section is differential. For simplicity, it is assumed in the following section that the domain of $\boldsymbol{x}^{(T)}$ and $\boldsymbol{x}^{(R)}$ are identical, though this is not necessary for the method.

From basic laws of probability, the relationship between the densities is

$$p\left(\boldsymbol{x}^{(R)}\right) = \eta \int d^D\boldsymbol{x}^{(T)}\; p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right) p\left(\boldsymbol{x}^{(T)}\right), \tag{3.1}$$

with

$$\frac{1}{\eta} = \int d^D\boldsymbol{x}^{(R)} d^D\boldsymbol{x}^{(T)}\; p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right) p\left(\boldsymbol{x}^{(T)}\right). \tag{3.2}$$

The proportionality constant $\eta$ is required to ensure that both $p\left(\boldsymbol{x}^{(R)}\right)$ and $p\left(\boldsymbol{x}^{(T)}\right)$ are normalized probability distributions, as it is possible that not all true data will be reconstructed. However, the absolute normalization is not relevant for the purpose of this dissertation, and we will focus on extracting $\eta p\left(\boldsymbol{x}^{(T)}\right)$, a non-normalized distribution. Once $\eta p\left(\boldsymbol{x}^{(T)}\right)$ is found, one can obtain $p\left(\boldsymbol{x}^{(T)}\right)$ by enforcing normalization.

One can interpret Equation 3.1 in terms of a smearing operator, mapping functions of the true domain to function of the reconstructed domain,

$$\tilde{g}\left(\boldsymbol{x}^{(R)}\right) = S\left[g(\boldsymbol{x}^{(T)})\right], \tag{3.3}$$

$$= \int d^D\boldsymbol{x}^{(T)}\; p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right) g\left(\boldsymbol{x}^{(T)}\right). \tag{3.4}$$

Thus Equation 3.1 can be simply written as

$$p\left(\boldsymbol{x}^{(R)}\right) = S\left[\eta p\left(\boldsymbol{x}^{(T)}\right)\right]. \tag{3.5}$$

Note, data from the detector can be used to estimate $p\left(\boldsymbol{x}^{(R)}\right)$, and data from a Monte Carlo, which includes the full simulation of the detector and all other effects for which one desires to correct, can be used to estimate $p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right)$. To obtain the desired $\eta p\left(\boldsymbol{x}^{(T)}\right)$, one needs to invert Equation 3.1. This equation, Equation 3.1, is a Fredholm integral equation of the first kind, and the problem of solving for $p\left(\boldsymbol{x}^{(T)}\right)$ belongs to the general class of inversion problems.

The inversion of Fredholm integral equations is a well studied problem, though it has been shown to be ill-conditioned [66]. This means it is possible that slight variations in $p\left(\boldsymbol{x}^{(R)}\right)$ could cause drastically different solutions for $p\left(\boldsymbol{x}^{(T)}\right)$. The proof that the problem is ill-conditioned involves showing that the eigenfunctions are, in general, degenerate. However, this is a worst case scenario, and the cases of interest are not always so badly conditioned. Section 3.1.5 specifically considers the effect of the poor conditioning within the context of the specific numerical method used in this dissertation.

In some specific cases, smearing effects are found to be negligible, and one considers $p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right)$ to have the form

$$p\left(\boldsymbol{x}^{(R)} \middle|\, \boldsymbol{x}^{(T)}\right) = \delta^D\left(\boldsymbol{x}^{(T)} - \boldsymbol{x}^{(R)}\right) \epsilon\left(\boldsymbol{x}^{(R)}\right). \tag{3.6}$$

As the $\delta$-function removes the integral, one naively expects the solution to Equation 3.1 to be simply

$$\eta p\left(\boldsymbol{x}^{(T)}\right) = \left.\frac{p\left(\boldsymbol{x}^{(R)}\right)}{\epsilon\left(\boldsymbol{x}^{(R)}\right)}\right|_{\boldsymbol{x}^{(R)}=\boldsymbol{x}^{(T)}}. \tag{3.7}$$

However, one has neglected the fact that a change of variables would introduce extra factors and one has naively assumed that the original choice of variables is the choice leading to a factor of unity. Furthermore, in low density areas, $\epsilon$ may be zero (or near zero), causing Equation 3.7 to be singular (or near singular). Thus for numeric studies, it is generally best to represent acceptance as an an integral operator rather than a simple distribution.

The proposed method of this section corrects for all effects which are included in the data sample used to estimate $p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right)$. Thus, whether or not a smearing correction is applied depends on whether such an effect is included in the data set used for estimating $p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right)$. The mathematics of the correction method are, however, the same regardless of which effects are included.

### 3.1.2 Solution Using Basis Expansion

To solve Equation 3.1, a finite basis expansion is proposed. Let $\{f_i(\boldsymbol{x})\}_{i=0}^{N}$ be a set of $N$ basis functions, which can be assumed to be linearly independent with respect to the $L^2$ norm. In the context of this dissertation, the basis functions are products of Legendre polynomials in $\cos\vartheta$ multiplied with sine and cosine functions, the arguments of the sines and cosine functions being linear combinations of $\phi_h$, $\phi_R$, and $\phi_S$. Let the solution $\eta p\left(\boldsymbol{x}^{(T)}\right)$ and the conditional probability $p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right)$ be restricted to this finite basis,

$$\eta p\left(\boldsymbol{x}^{(T)}\right) = \sum_i \alpha_i f_i\left(\boldsymbol{x}^{(T)}\right), \tag{3.8}$$

$$p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right) = \sum_{i,j} \Gamma_{i,j} f_i\left(\boldsymbol{x}^{(R)}\right) f_j\left(\boldsymbol{x}^{(T)}\right). \tag{3.9}$$

The $\Gamma$ matrix, multiplied by a matrix taking into account the normalization of the basis, is a representation of the acceptance and smearing, while $\boldsymbol{\alpha}$ represents the final fit parameters which one desires to estimate. Thus, the basis expansion allows one to turn a difficult integral inversion problem into a parameter estimation problem—a much easier class of problems. Note, however, such a change does not remove the inherent poor conditioning of the problem.

#### 3.1.2.1 Fitting the Conditional Probability

The next step is to determine matrix $\Gamma$ by minimizing the integrated squared error (ISE),

$$\int d^D\boldsymbol{x}^{(R)} d^D\boldsymbol{x}^{(T)} \left[ p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right) - \sum_{i,j} \Gamma_{i,j} f_i(\boldsymbol{x}^{(R)}) f_j(\boldsymbol{x}^{(T)}) \right]^2. \tag{3.10}$$

It is helpful to define

$$B_{i,j} = \int d^D\boldsymbol{x}^{(\boldsymbol{R})} d^D\boldsymbol{x}^{(\boldsymbol{T})} \; p\left(\boldsymbol{x}^{(R)} \middle| \, \boldsymbol{x}^{(T)}\right) f_i\left(\boldsymbol{x}^{(\boldsymbol{R})}\right) f_j\left(\boldsymbol{x}^{(\boldsymbol{T})}\right), \tag{3.11}$$

$$F_{i,j} = \int d^D\boldsymbol{x}^{(\boldsymbol{T})} \; f_i\left(\boldsymbol{x}^{(\boldsymbol{T})}\right) f_j\left(\boldsymbol{x}^{(\boldsymbol{T})}\right). \tag{3.12}$$

Elements of the matrix $F$ are simply the $L^2$ overlaps between basis elements. This implies $F$ is symmetric, positive definite, and thus invertable. Elements of the matrix $B$ can be interpreted as the $L^2$ overlaps between basis elements $f_i$ and smeared basis elements $S\left[f_j\right]$. Note that $B$ is only symmetric in the case of symmetric smearing, including the case of no smearing. One can show that $B$ is positive semi-definite, and the fact that $B$ can have zero (or near zero) eigenvalues is exactly related to the inherent ill-conditioned nature of the Fredholm equation.

Using Equations 3.11 and 3.12, one can reduce Equation 3.10 to

$$\operatorname{Tr}\left[\Gamma F \Gamma^T F\right]-2 \operatorname{Tr}\left[\Gamma B^T\right]+\text{const.}, \tag{3.13}$$

where the constant is meant with respect to $\Gamma$. The ISE is minimized when $\Gamma=F^{-1} B F^{-1}$.

Note, $\Gamma F=F^{-1} B$ is the representation of the smearing operator restricted to the space spanned by the basis functions. Specifically, given a function $g(\boldsymbol{x}^{(T)})=\sum_i \beta_i f_i\left(\boldsymbol{x}^{(T)}\right)$, the smeared function can be found by acting $\Gamma F$ on the parameters $\boldsymbol{\beta}$, as

$$\begin{aligned}
S\left[g\left(\boldsymbol{x}^{(T)}\right)\right] & =\int d^D \boldsymbol{x}^{(T)} p\left(\boldsymbol{x}^{(R)} \mid \boldsymbol{x}^{(T)}\right) g\left(\boldsymbol{x}^{(T)}\right), & (3.14)\\
& =\int d^D \boldsymbol{x}^{(T)}\left[\sum_{i, j} \Gamma_{i, j} f_i\left(\boldsymbol{x}^{(R)}\right) f_j\left(\boldsymbol{x}^{(T)}\right)\right]\left[\sum_k \beta_k f_k\left(\boldsymbol{x}^{(T)}\right)\right], & (3.15)\\
& =\sum_{i, j, k} f_i\left(\boldsymbol{x}^{(R)}\right) \Gamma_{i, j} \beta_k \int d^D \boldsymbol{x}^{(T)} f_j\left(\boldsymbol{x}^{(T)}\right) f_k\left(\boldsymbol{x}^{(T)}\right), & (3.16)\\
& =\sum_i(\Gamma F \boldsymbol{\beta})_i f_i\left(\boldsymbol{x}^{(R)}\right), & (3.17)\\
& =\sum_i\left(F^{-1} B \boldsymbol{\beta}\right)_i f_i\left(\boldsymbol{x}^{(R)}\right), & (3.18)
\end{aligned}$$

assuming the integrals and sums are all convergent.

#### 3.1.2.2 Fitting the True Distribution

To compute $\boldsymbol{\alpha}$, one again forms the ISE, this time based on Equation 3.7,

$$\int d^D \boldsymbol{x}^{(R)}\left\{p\left(\boldsymbol{x}^{(R)}\right)-S\left[\eta p\left(\boldsymbol{x}^{(T)}\right)\right]\right\}^2 . \tag{3.19}$$

It is helpful to introduce the expectation value of the functions with respect to $p\left(\boldsymbol{x}^{(R)}\right)$,

$$b_i=\int d^D \boldsymbol{x}^{(R)} p\left(\boldsymbol{x}^{(R)}\right) f_i\left(\boldsymbol{x}^{(R)}\right) . \tag{3.20}$$

The ISE of Equation 3.19 then reduces to

$$\boldsymbol{\alpha}^T B^T F^{-1} B \boldsymbol{\alpha}-2 \boldsymbol{\alpha}^T B^T F^{-1} \boldsymbol{b}+\text{const.} \tag{3.21}$$

The minimum occurs when

$$\left(B^T F^{-1} B\right) \boldsymbol{\alpha}=B^T F^{-1} \boldsymbol{b}. \tag{3.22}$$

If $\left(B^T F^{-1} B\right)$ is invertable, the solution for $\boldsymbol{\alpha}$ is then

$$\boldsymbol{\alpha} = \left(B^T F^{-1} B\right)^{-1} B^T F^{-1} \boldsymbol{b}. \tag{3.23}$$

If $B$ is itself invertable, the solution reduces to

$$\boldsymbol{\alpha} = B^{-1} \boldsymbol{b}. \tag{3.24}$$

In solving either of Equations 3.23 or 3.24, one typically does not directly compute the inverse, but rather uses standard methods for solving matrix equations involving either a QR or LU decomposition. A $QR$ decomposition, using Householder transformations, is considered the most stable method for solving for $\boldsymbol{\alpha}$.

An alternate derivation of Equation 3.23 can be found by computing the ISE estimate of $p\left(\boldsymbol{x}^{(R)}\right)$, and then constructing the ISE. In this case, Equation 3.19 can be written

$$\begin{aligned} ISE &= \int d^D \boldsymbol{x}^{(R)} \left\{ p\left(\boldsymbol{x}^{(R)}\right) - S\left[\eta p\left(\boldsymbol{x}^{(T)}\right)\right] \right\}^2, & (3.25) \\ &= \int d^D \boldsymbol{x}^{(R)} \left\{ \sum_k \left(F^{-1}\boldsymbol{b} - F^{-1} B \boldsymbol{\alpha}\right)_k f_k\left(\boldsymbol{x}^{(R)}\right) \right\}, & (3.26) \\ &= \left(F^{-1}\boldsymbol{b} - F^{-1} B \boldsymbol{\alpha}\right)^T F \left(F^{-1}\boldsymbol{b} - F^{-1} B \boldsymbol{\alpha}\right), & (3.27) \end{aligned}$$

where the similarity with typical linear regression is more apparent. Minimizing Equation 3.27 also results in Equation 3.23. However, this derivation makes clear the need for the basis to be chosen such that both $p\left(\boldsymbol{x}^{(T)}\right)$ and $p\left(\boldsymbol{x}^{(R)}\right)$ are in the span, not just $p\left(\boldsymbol{x}^{(T)}\right)$. As $p\left(\boldsymbol{x}^{(R)}\right)$ is not explicitly fit to the basis in the first derivation, this subtlety is only apparent in the second derivation.

### 3.1.3 Uncertainty Calculation

For simplicity, in this section (Section 3.1.3) it will be assumed that both $B$ and $\left(B^T F^{-1} B\right)$ are invertable. The case they are not invertable is considered in Section 3.1.5, which details the computation of a pseudo-inverse, which can then be used in the formulas of this section, Section 3.1.3. The principle equation used in computing $\boldsymbol{\alpha}$ is then

$$B\boldsymbol{\alpha} = \boldsymbol{b}. \tag{3.28}$$

Three possible sources of uncertainty can be identified: two from propagating the uncertainty from $B$ and $\boldsymbol{b}$, and the third related to the inverse Hessian of the optimization function, in this case the ISE. Let $C^{(\boldsymbol{b})}$ be the covariance matrix of $\boldsymbol{b}$ and $C^{(B)}_{i,j;i',j'}$ be the covariance 4-form of $B$. Also define the matrix $C'^{(B)}$

$$C'^{(B)}_{i,i'} = \sum_{j,j'} C^{(B)}_{i,j;i',j'} \alpha_j \alpha_{j'}. \tag{3.29}$$

The covariance of $\alpha$, denoted $C^{(\boldsymbol{\alpha})}$ can then be written as

$$C^{(\boldsymbol{\alpha})} = B^{-1} C^{(\boldsymbol{b})} B^{-T} + B^{-1} C'^{(B)} B^{-T} + \left(B^T F^{-1} B\right)^{-1}. \tag{3.30}$$

The last term in Equation 3.30 is that arising from the inverse Hessian and does not scale with statistics. In numerical studies, it has been shown that the first two terms sufficiently quantify the relevant uncertainty, and that the final term is not meaningful. Equation 3.30 followed from the standard methods of propagating uncertainty (i.e. covariance matrices). Note, implicit differentiation of Equation 3.28 is used to find partial derivatives of elements of $\boldsymbol{\alpha}$ with respect to elements of $B$.

### 3.1.4 Numerical Calculation

The matrix $\Gamma$, introduced in fitting the conditional probability, never needs to be explicitly computed. Instead, only the matrices $F$ and $B$ and the vector $b$ need be estimated from available data. The matrix $F$ can be computed analytically, while $B$ and $b$ can be computed via Monte Carlo integration.

Given a set of data $\{\boldsymbol{x}^{(R,k)}\}_{k=1}^{N_R}$ drawn from $p\left(\boldsymbol{x}^{(R)}\right)$, i.e. the actual data from the HERMES detector, the expectation value and covariance matrix can be estimated via

$$b_i = \frac{V}{N_R}\sum_{k=1}^{N_R} f_i\left(\boldsymbol{x}^{(R,k)}\right), \tag{3.31}$$

$$\left(C^{\boldsymbol{b}}\right)_{j,j'} = \frac{\delta_{j,j'}}{N_R-1}\left[\frac{V^2}{N_R}\sum_{k=1}^{N_R} f_i^2\left(\boldsymbol{x}^{(R,k)}\right) - \left(b_i\right)^2\right]. \tag{3.32}$$

The quantity $V$ is the volume of the domain (assumed to be the same for both the $\boldsymbol{x}^{(T)}$ and $\boldsymbol{x}^{(R)}$ domains).

Computing $B$ and its covariance requires careful preparation of the Monte Carlo generator, as it is does not makes sense to discuss data drawn from a conditional probability. However, if data is generated uniform in $\boldsymbol{x}^{(T)}$, then $p_{MC}\left(\boldsymbol{x}^{(T)}\right) = V^{-1}$. One can then run this data through the simulation of the detector and the full data reconstruction scheme, yielding data distributed according to some $p_{MC}\left(\boldsymbol{x}^{(T)}, \boldsymbol{x}^{(R)}\right)$. The conditional probability is numerically related to the joint distribution as

$$p\left(\boldsymbol{x}^{(R)}\middle|\, \boldsymbol{x}^{(T)}\right) = \frac{p_{MC}\left(\boldsymbol{x}^{(T)}, \boldsymbol{x}^{(R)}\right)}{p_{MC}\left(\boldsymbol{x}^{(T)}\right)} = V\, p_{MC}\left(\boldsymbol{x}^{(T)}, \boldsymbol{x}^{(R)}\right). \tag{3.33}$$

Thus the matrix $B$ can be written as

$$B_{i,j} = V\int d^D\boldsymbol{x}^{\boldsymbol{(R)}} d^D\boldsymbol{x}^{\boldsymbol{(T)}}\; p_{MC}\left(\boldsymbol{x}^{(T)}, \boldsymbol{x}^{(R)}\right) f_i\left(\boldsymbol{x}^{\boldsymbol{(R)}}\right) f_j\left(\boldsymbol{x}^{\boldsymbol{(T)}}\right), \tag{3.34}$$

where elements of $B$ are simply expectation values of products of basis functions. Note, that the factor of $V$ in Equation 3.34 only affects the normalization of $\boldsymbol{\alpha}$, and thus can be absorbed into the yet unknown normalization factor.

Given data $\{\boldsymbol{x}^{(T,k)}, \boldsymbol{x}^{(R,k)}\}_{k=1}^{N_{MC}}$ drawn from a Monte Carlo prepared as described, the matrix $B$ and its covariance can then be computed via

$$\begin{aligned} B_{i,j} &= \frac{V^3}{N_{MC}} \sum_{k=1}^{N_{MC}} f_i\left(\boldsymbol{x}^{(R,k)}\right) f_j\left(\boldsymbol{x}^{(T,k)}\right), && (3.35) \\ \left(C^B\right)_{j,k;j',k'} &= \frac{\delta_{j,j'}\delta_{k,k'}}{N_\epsilon - 1}\left[\frac{V^4}{N_\epsilon}\sum_{k=1}^{N_\epsilon} f_j^2\left(\boldsymbol{x}^{(M,k)}\right) f_k^2\left(\boldsymbol{x}^{(T,k)}\right) - \left(B_{j,k}\right)^2\right]. && (3.36) \end{aligned}$$

Once the quantities $B$, $\boldsymbol{b}$ and their covariance are computed via Equations 3.31, 3.32, 3.35, and 3.36, the vector $\boldsymbol{\alpha}$ and its covariance can be found via Equations 3.28 and 3.30. The largest amount of the computation time is spent in summing over the two data samples to compute $B$ and $\boldsymbol{b}$, though in general the entire algorithm is quite fast.

### 3.1.5 Inverting the Matrix

While naively one might choose the set of basis functions to exactly match the modulations occurring in the cross section, one should also consider including extra terms to capture additional acceptance effects, i.e. one needs to represent well both $p\left(\boldsymbol{x}^{(T)}\right)$ and $p\left(\boldsymbol{x}^{(R)}\right)$ by the chosen basis. For example, one can consider including additional terms with the same azimuthal dependence as the terms from the cross section, but with a different Legendre polynomial of $\cos\vartheta$. The side effect of including additional basis functions, however, is that the matrix $B$ tends to be even more poorly conditioned. However, the matrix $B$ can be poorly conditioned even without increasing the number of basis functions.

To invert $B$, or to compute an estimate of a pseudo-inverse, a technique is used which is common in dimensional reduction problems. Let $N$ denote the dimension of $B$, and let the $m \times m$ identity matrix be denoted $I_m$, for any $m > 0$. Note that $B$ is positive semi-definite, and thus the eigenvalues of $B$ are non-negative and the eigenvectors are orthogonal. Consider the eigenvalue decomposition of $B$, $B = VDV^T$, where $D$ is diagonal and $V^TV = VV^T = I_N$.

Next, considers the $n$ largest eigenvalues, with $n$ chosen such that the sum of the $n$ largest eigenvalues equals a certain fraction of the sum of all eigenvalues. For the case of this dissertation, the fraction is chosen to be 95%. One then defines $V'$, an $N \times n$ matrix whose columns are the eigenvectors corresponding to the $n$ largest eigenvalues, as well as $D'$, an $n \times n$ matrix of the $n$ largest eigenvalues. One then has $B = VDV^T \approx V'D'V'^T$. Note that while $V'^TV' = I_n$, one has $V'V'^T \neq I_N$.

It can be shown that the matrix equation $B\boldsymbol{\alpha} = \boldsymbol{b}$ has solution

$$\boldsymbol{\alpha} = V'\left(D'\right)^{-1} V'^T \boldsymbol{b}, \tag{3.37}$$

and thus one can use $B^{-1} \approx V'\left(D'\right)^{-1} V'^T$ as a pseudo-inverse. In practice, $QR$ decomposition leads to more stable estimates of $\boldsymbol{\alpha}$ than using Equation 3.37. However, in propagating the uncertainty, the reverse is true. In this case, utilizing the pseudo-inverse of $B$ is more stable than using the inverse computed from $QR$ or any other matrix decomposition.

## 3.2 The `TMDGen` Generator

For systematic studies, it is necessary to either generate data according to the TMD SIDIS dihadron cross section, or to at least evaluate this cross section for given data points. Although `Pythia` [67] includes SIDIS dihadron production, no model is included for the angular dependence. In fact, no Monte Carlo generator has been previously available for the full angular dependence.

A limited generator, known as **`gmctrans`** [68], had been developed in FORTRAN by other members of the HERMES collaboration. This generator only included SIDIS single hadron production and was also restricted to the Gaussian ansatz for the $p_T$ and $k_T$ dependence for each distribution and fragmentation function. The generator also had several other disadvantages. Specifically, the program design made it difficult to extend to other cases, due to a large number of FORTRAN common blocks, exacerbated by the fact that the code was linked to a large number of other generators. The design did also not allow for the separation from the rest of the standard HERMES software, which is necessary for any broad use by other experiments or theorists.

For these reasons, a new generator, denoted **`TMDGen`**, has been developed. The generator is not simply a porting of the older FORTRAN code, but a complete redesign into an object oriented framework. Many advantages of the more modern language, such as class inheritance, namespaces, encapsulation, and longer names allow for cleaner, better organized code. The library is designed to be dependent on as few other libraries as possible and can both operate independent of the HERMES software suite and can connect to the HERMES processes chain. Additionally, the new **`TMDGen`** fully models the intrinsic transverse momenta, $\boldsymbol{p}_T$, $\boldsymbol{k}_T$, and places no model assumptions on their distribution.

The new generator includes SIDIS production of identified single hadrons, hadron pairs, and vector mesons, assuming a polarized electron or positron beam and fixed, polarized, proton target. The results of the generator can serve as input into a simulation of a detector, such as by using GEANT, followed by the usual data processing chain of an experiment.

The remainder of this chapter is organized as follows. Section 3.2.1 describes the basic methodology of the generator. Details regarding the computation of the track momenta from the cross section variables are given in Section 3.2.2. Code organization is discussed in Section 3.2.3. Section 3.2.4 describe models currently implemented in `TMDGen`. Finally, Section 3.2.5 compares numerical results from `TMDGen` with those from a `Pythia` production tuned to HERMES kinematics [69, 70, 71].

### 3.2.1 Methodology

The basic method of generation is the simple acceptance/rejection method. The method is as follows. Let $\boldsymbol{x}$ denote the independent variables, and let $f(\boldsymbol{x})$ be the distribution from which one desires to generate data. Let $g(\boldsymbol{x})$ be an additional distribution from which one can already generate data, generally chosen to be uniform. Also, let $M > 0$ be given such that $f(\boldsymbol{x}) \leq Mg(\boldsymbol{x})$ for all $\boldsymbol{x}$. Let $\mathcal{X}_g$ be a set of data generated from $g$. A data set following the distribution of $f$ can then be determined by rejecting data points $\boldsymbol{x}^* \in \mathcal{X}_g$ such that

$$\frac{f(\boldsymbol{x}^*)}{Mg(\boldsymbol{x}^*)} \leq r, \tag{3.38}$$

where $r$ is uniformly distributed in $[0, 1)$.

In the case of `TMDGen`, $f$ of Equation 3.38 is the cross section for the given process. The distribution $g(\boldsymbol{x})$ is set to unity, (though this can be overwritten by defining a new child class of the variable thrower class) and $M$ is set to $\max f(\boldsymbol{x})$. As $M$ is not known exactly, it is estimated by randomly sampling the distribution $f$, i.e. the cross section.

For single hadron production, the independent variables $\boldsymbol{x}$ are $x$, $y$, $z$, $P_{h\perp}$, $\phi_h$, $\psi$, $p_T$, and $\phi_{p_T}$, and for dihadron production $\boldsymbol{x}$ additionally includes $M_h$, $\cos\vartheta$ and $\phi_R$. These quantities are defined in Section 2.1.4, with the exception of $\psi$, which is merely the rotation angle from coordinate system $\mathbf{I_b}$ to $\mathbf{I_c}$, or equivalently, minus the azimuthal angle of the scattered lepton in coordinate system $\mathbf{I_b}$.

In the case that the lepton beam is purely in the $\hat{\boldsymbol{z}}$ direction in coordinate system $\mathbf{I_a}$, i.e. coordinate systems $\mathbf{I_a}$ and $\mathbf{I_b}$ are identical, a simple relation exists between $\psi$ and $\phi_{e_2}$, the azimuthal angle of the scattered lepton in coordinate system $\mathbf{I_a}$. Specifically $\psi = -\phi_{e_2}$. However, allowing the beam to have any direction in coordinate system $\mathbf{I_a}$ (i.e. a generic relation between systems $\mathbf{I_a}$ and $\mathbf{I_b}$) results in a more complicated expression. Note, again assuming coordinate systems $\mathbf{I_a}$ and $\mathbf{I_b}$ are identical, and assuming very small $\theta_\gamma$, the relation also holds that $\phi_S \approx \psi - \pi/2$. These relations have been useful in debugging `TMDGen` and may prove useful to others debugging code designed to link with `TMDGen`.

As the cross section is chosen differential with respect to the polar coordinates $p_T$, $\phi_p$ rather than the Cartesian $\boldsymbol{p}_T$, an additional factor of $p_T$ must be included in the phase space. Additionally, a factor of $\hbar^2/c^2 = 389.379$ $\mu$barns GeV$^2$ is included to convert the cross section to units of $\mu$barns.

Note that the target polarization is given with respect to the lepton beam, with the magnitude of transverse polarization being $P_\perp$ and of longitudinal polarization $P_\parallel$. The cross section is chosen differential with respect to $\psi$ rather than $\phi_S$, to avoid the introduction of additional phase space factor due to specifying $P_\parallel$, $P_\perp$ rather than $S_\parallel$, $S_\perp$ [36].

To determine $\phi_S$, $S_\parallel$, $S_\perp$ from the given $P_\parallel$, $P_\perp$, $\psi$ one just needs to transform the polarization vector from frame $\mathbf{I_a}$, where $P_\parallel$, $P_\perp$ are defined, to frame $\mathbf{I_d}$, where $S_\parallel$, $S_\perp$ are defined. The specific rotation matrices are

$$\begin{aligned}
\begin{pmatrix} S_\perp \cos\phi_S \\ S_\perp \sin\phi_S \\ -S_\parallel \end{pmatrix} &= \begin{pmatrix} \cos\theta_\gamma & 0 & -\sin\theta_\gamma \\ 0 & 1 & 0 \\ \sin\theta_\gamma & 0 & \cos\theta_\gamma \end{pmatrix} \\
&\times \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_{e_1} & 0 & -\sin\theta_{e_1} \\ 0 & 1 & 0 \\ \sin\theta_{e_1} & 0 & \cos\theta_{e_1} \end{pmatrix} \\
&\times \begin{pmatrix} \cos\phi_{e_1} & \sin\phi_{e_1} & 0 \\ -\sin\phi_{e_1} & \cos\phi_{e_1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -P_\perp \\ -P_\parallel \end{pmatrix}, \qquad (3.39)
\end{aligned}$$

with $\theta_\gamma$ being the angle between the beam lepton's momentum and the virtual photon's momentum. Equation 3.39 reduces to Equation 4 of Ref. [36] in the case that $\theta_{e_1} = 0$,

$\phi_{e_1} = 0$. Multiplying the matrices in Equation 3.39 results in

$$\begin{aligned} S_\perp \cos\phi_S &= P_\parallel \Big( \cos\theta_\gamma \cos\psi \sin\theta_{e_1} + \sin\theta_\gamma \cos\theta_{e_1} \Big) \\ &\quad + P_\perp \Big( \sin\theta_\gamma \sin\theta_{e_1} \sin\phi_{e_1} - \cos\theta_\gamma \cos\psi \cos\theta_{e_1} \sin\phi_{e_1} \\ &\qquad + \cos\theta_\gamma \sin\psi \cos\phi_{e_1} \Big), \qquad (3.40) \\ S_\perp \sin\phi_S &= P_\parallel \Big( \sin\psi \sin\theta_{e_1} \Big) + P_\perp \Big( \sin\psi \cos\theta_{e_1} \sin\phi_{e_1} - \cos\psi \cos\phi_{e_1} \Big), \qquad (3.41) \end{aligned}$$

from which $S_\perp$, and $\phi_S$ can be computed, and

$$\begin{aligned} -S_\parallel &= P_\parallel \Big( \sin\theta_\gamma \cos\psi \sin\theta_{e_1} - \cos\theta_\gamma \cos\theta_{e_1} \Big) \\ &\quad + P_\perp \Big( \sin\theta_\gamma \cos\psi \cos\theta_{e_1} \sin\phi_{e_1} + \cos\theta_\gamma \sin\theta_{e_1} \sin\phi_{e_1} \\ &\qquad - \sin\theta_\gamma \sin\psi \cos\phi_{e_1} \Big), \qquad (3.42) \end{aligned}$$

from which $S_\parallel$ can be computed.

Although neither $Q^2$, $P_{h\perp}$, or $W^2$ is generated directly, the domain of the generated variables can be restricted in order to keep these three variables within the physically accessible range. This is done by directly computing these variables according to

$$\begin{aligned} Q^2 &= 2xyME_B, & (3.43) \\ \nu &= yE_B, & (3.44) \\ W^2 &= M^2 + 2M\nu - Q^2, & (3.45) \end{aligned}$$

and rejecting events where these are out of a user specified range.

Note that the cross section for single hadron production can be written

$$\begin{aligned} \frac{d^6\sigma}{dx\,dy\,dz\,d\boldsymbol{P}_{h\perp}\,d\psi} &= \int d^2\boldsymbol{p_T} d^2\boldsymbol{k}_T \; \delta^2\left( \boldsymbol{p_T} - \boldsymbol{k_T} - \frac{\boldsymbol{P_{h\perp}}}{z} \right) \times \\ &\qquad \frac{d^{10}\sigma}{dx\,dy\,dz\,d\psi\,d\boldsymbol{P}_{h\perp}\,d\boldsymbol{p}_T\,d\boldsymbol{k}_T}, \qquad (3.46) \\ &= \int d^2\boldsymbol{p_T} \; \left. \frac{d^{10}\sigma}{dx\,dy\,dz\,d\psi\,d\boldsymbol{P_{h\perp}}\,d\boldsymbol{p_T}\,d\boldsymbol{k}_T} \right|_{\boldsymbol{k}_T = \boldsymbol{p}_T - \boldsymbol{P}_{h\perp}/z}, \qquad (3.47) \end{aligned}$$

and thus one can identify

$$\frac{d^8\sigma}{dx\,dy\,dz\,d\psi\,d\boldsymbol{P_{h\perp}}\,d\boldsymbol{p_T}} = \left. \frac{d^{10}\sigma}{dx\,dy\,dz\,d\psi\,d\boldsymbol{P_{h\perp}}\,d\boldsymbol{p_T}\,d\boldsymbol{k_T}} \right|_{\boldsymbol{k}_T = \boldsymbol{p}_T - \boldsymbol{P}_{h\perp}/z}. \qquad (3.48)$$

Similar expressions can be written for the dihadron cross section. To fully simulate the $p_T$ and $k_T$ dependence, one can effectively remove the integrals from the cross section in Section 2.2, include $\boldsymbol{p}_T$ in the list of independent variables, and set $\boldsymbol{k}_T$ to the value required by the $\delta$-function. To recover the physically realizable cross section, one needs to use just the physically realizable variables, effectively performing a Monte Carlo integration over $p_T$.

### 3.2.2 Computation of Track Momenta

Once the independent variables, those denoted as $\boldsymbol{x}$ in preceding paragraphs, are generated, it is necessary to invert the usual relations, Section 2.1.4, to determine the momentum of the scattered electron and produced mesons. First the electron and virtual photon momenta are determined, followed by the momentum of the center of mass of the produced hadron system, and finally, for dihadrons, the momenta of measured particles.

#### 3.2.2.1 Angles of the Scattered Electron

Using the definitions of the variables, Section 2.1.4, the energy and polar angle of the scattered electron in the frame/coordinate system $\mathbf{I_b}$, equivalently coordinate systems $\mathbf{I_c}$, is

$$\begin{aligned} E' &= (1-y)E_B, && (3.49)\\ \cos\theta_{e_2} &= 1-\frac{Q^2}{2(1-y)E_B} = 1-\frac{xy}{1-y}\frac{M}{E_B}. && (3.50)\end{aligned}$$

One can equivalently write

$$\begin{aligned} \cos\theta_{e_2} &= 1-\frac{\gamma^2 y^2}{2(1-y)}, && (3.51)\\ \sin\theta_{e_2} &= \pm\frac{y}{1-y}\gamma\sqrt{1-y-\frac{\gamma^2y^2}{4}}. && (3.52)\end{aligned}$$

Note, a relation between $\theta_{e_2}$ and $\theta_\gamma$ can be show, specifically

$$\begin{aligned} \sin\theta_\gamma &= \frac{(1-y)E_B}{\sqrt{\nu^2+Q^2}}\sin\theta_{e_2}, && (3.53)\\ &= \frac{E'}{|\boldsymbol{q}|}\sin\theta_{e_2}. && (3.54)\end{aligned}$$

Numerically, one computes $\theta_{e_2}$ from Equation 3.50 and $\theta_\gamma$ from Equation 3.53.

The momenta of the scattered lepton in coordinate system $\mathbf{I_c}$ has magnitude and direction given by Equations 3.49 and 3.50, with the azimuthal angle being zero. To determine the momenta in $\mathbf{I_a}$, one just needs to apply the necessary transformations between these coordinate systems.

#### 3.2.2.2 Angles of the Meson System

The energy of the produced hadron system can be computed by inversion of the kinematic variable definitions, and the momentum then directly follows from the mass, specifically

$$\begin{aligned} E_h &= yzE_B, \tag{3.55} \\ P_h &= \sqrt{E_h^2 - M_h^2}. \tag{3.56} \end{aligned}$$

In coordinate system $\mathbf{I_d}$, the momentum of the produced hadron system has polar and azimuthal angles $\theta_{\gamma h}$, $\phi_h$. Note, the angle $\theta_{\gamma h}$ was defined in Equation 2.32. A rotation about $\widehat{\boldsymbol{y}}$ of $\theta_\gamma$ followed by a rotation about $\widehat{\boldsymbol{z}}'$ of $-\phi_{e_2}$ moves a vector from this coordinate system to system $\mathbf{I_a}$. Using vector and matrix notation, one can then write

$$\widehat{\boldsymbol{P}}_h^{(\mathbf{I_a})}\left(\theta_h^{(\mathbf{I_a})}, \phi_h^{(\mathbf{I_a})}\right) = R_{\widehat{\boldsymbol{z}}'}(-\phi_{e_2}) R_{\widehat{\boldsymbol{y}}}(\theta_\gamma) \widehat{\boldsymbol{P}}_h^{(\mathbf{I_d})}\left(\theta_{\gamma h}^{(\mathbf{I})}, \phi_h\right), \tag{3.57}$$

from which the quantities $\theta_h^{(\mathbf{I_a})}$, $\phi_h^{(\mathbf{I_a})}$ can be determined.

It is also necessary to determine boost $P_h^{(\mathbf{I_d})}$ to frame $\mathbf{II_a}$ to compute other dihadron variables. In frame $\mathbf{I_d}$, the four momentum is

$$\left(P_h^{(\mathbf{I_d})}\right)^\mu = \left[E_h,\ P_h \sin\theta_{\gamma h}^{(\mathbf{I})} \cos\phi_h,\ P_h \sin\theta_{\gamma h}^{(\mathbf{I})} \sin\phi_h,\ P_h \cos\theta_{\gamma h}^{(\mathbf{I})}\right], \tag{3.58}$$

noting that $E_h$, $P_h$ is defined in frame $\mathbf{I}$, i.e $E_h \equiv E_h^{(\mathbf{I})}$, $P_h \equiv P_h^{(\mathbf{I})}$. The boost parameters from $\mathbf{I}$ to $\mathbf{II}$ are

$$\gamma = \frac{\nu + M}{W}, \qquad \gamma\beta_z = \frac{|\boldsymbol{q}|}{W}. \tag{3.59}$$

In frame $\mathbf{II_a}$ the $x$ and $y$ components are unchanged, while the energy and $z$ components are

$$\begin{aligned} E_h^{(\mathbf{II})} &= \frac{\nu + M}{W} E_h - \frac{|\boldsymbol{q}|}{W} P_h \cos\theta_{\gamma h}^{(\mathbf{I})}, \tag{3.60} \\ P_{h,z}^{(\mathbf{II_a})} &= \frac{\nu + M}{W} P_h \cos\theta_{\gamma h}^{(\mathbf{I})} - \frac{|\boldsymbol{q}|}{W} E_h. \tag{3.61} \end{aligned}$$

The magnitude of the three momentum in this frame is then

$$P_h^{(\mathbf{II})} = \sqrt{P_{h\perp}^2 + \left(P_{h,z}^{(\mathbf{II_a})}\right)^2}. \tag{3.62}$$

For other equations following, it is useful to note the light cone coordinates $P_h^+$ and $P_h^-$ in $\mathbf{II_b}$ and $\mathbf{II_c}$ are

$$\begin{aligned} P_h^\pm &= \frac{1}{\sqrt{2}}\left(E_h^{(\mathbf{II})} \pm P_h^{(\mathbf{II})}\right), \tag{3.63} \\ &= \frac{1}{\sqrt{2}}\left(\frac{\nu + M \mp |\boldsymbol{q}|}{W}\right)\left(z\nu \pm P_h \cos\theta_{\gamma h}^{(\mathbf{I})}\right), \tag{3.64} \\ & \tag{3.65} \end{aligned}$$

One also has the relation that $2P_h^+ P_h^- = M_h^2$.

#### 3.2.2.3 Angles of the Decay Products

For single pseudo-scalar meson production, the hadron system is purely the single meson, and the previous subsections are enough to determine all the measured momenta based on the cross section variables. For dihadron production, one needs to yet determine the momenta of the two measured hadrons. Computationally, one transforms $R^\mu$ from reference frame $\mathbf{III_a}$, where it can be written directly, to frame $\mathbf{I_a}$. Then knowing $R^\mu$ and $P_h^\mu$, one can compute $P_1^\mu$ and $P_2^\mu$. The four vector $R^\mu$ in frame $\mathbf{III_a}$ can be written as a function of angles $\vartheta$, $\phi_R$, and the relevant masses, as [18]

$$R^\mu = \left[\frac{m_1^2 - m_2^2}{2M_h},\ |\boldsymbol{R}|\cos\phi_R \sin\vartheta,\ |\boldsymbol{R}|\sin\phi_R\sin\vartheta,\ |\boldsymbol{R}|\cos\vartheta\right], \tag{3.66}$$

with

$$|\boldsymbol{R}|^2 = \frac{M_h^2}{4} - \frac{m_1^2 + m_2^2}{2} + \frac{\left(m_1^2 - m_2^2\right)^2}{4M_h^2}. \tag{3.67}$$

It should be noted that the Lorentz invariant $R^2 := R^\mu R_\mu$ has the value

$$R^2 = \frac{m_1^2 + m_2^2}{2} - \frac{M_h^2}{4}. \tag{3.68}$$

One can boost to frame $\mathbf{II_b}$ via boost parameters

$$\gamma = \frac{E_h^{(\mathbf{II})}}{M_h}, \quad \gamma\beta_z = -\frac{P_h^{(\mathbf{II})}}{M_h}. \tag{3.69}$$

The rotation matrix from frames $\mathbf{II_b}$ to $\mathbf{II_a}$ can best be expressed as functions of the Cartesian three momenta $\boldsymbol{P}_h^{(\mathbf{II_a})}$. The matrix represents a rotation about the $x$ axis, such that $P_h$ is in the $y$-$z$ plane, followed by a rotation about the $y$ axis. The matrix can be written as

$$R^{(\mathbf{II_b}\to\mathbf{II_a})} = \begin{pmatrix} \frac{P_{h,z}^{(\mathbf{II_a})}}{P_{h,xz}^{(\mathbf{II_a})}} & \frac{-P_{h,x}^{(\mathbf{II_a})}P_{h,y}^{(\mathbf{II_a})}}{P_{h,xz}^{(\mathbf{II_a})}P_h^{(\mathbf{II_a})}} & \frac{P_{h,x}^{(\mathbf{II_a})}}{P_h^{(\mathbf{II_a})}} \\ 0 & \frac{P_{h,xz}^{(\mathbf{II_a})}}{P_h^{(\mathbf{II_a})}} & \frac{P_{h,y}^{(\mathbf{II_a})}}{P_h^{(\mathbf{II_a})}} \\ \frac{-P_{h,x}^{(\mathbf{II_a})}}{P_{h,xz}^{(\mathbf{II_a})}} & \frac{-P_{h,y}^{(\mathbf{II_a})}P_{h,z}^{(\mathbf{II_a})}}{P_{h,xz}^{(\mathbf{II_a})}P_h^{(\mathbf{II_a})}} & \frac{P_{h,z}^{(\mathbf{II_a})}}{P_h^{(\mathbf{II_a})}} \end{pmatrix}, \tag{3.70}$$

with

$$P_{h,xz}^{(\mathbf{II_a})} = \sqrt{\left(P_{h,x}^{(\mathbf{II_a})}\right)^2 + \left(P_{h,z}^{(\mathbf{II_a})}\right)^2}. \tag{3.71}$$

To boost to $\mathbf{I_d}$, one simply changes the sign of $\beta_z$ in Equation 3.59. One can then transform $R^\mu$ from $\mathbf{I_d}$ to $\mathbf{I_a}$ with the same rotation matrices as those given in Equation 3.57. Once $R^\mu$ and $P_h^\mu$ are determined, one can then determine $P_1^\mu$, $P_2^\mu$ according to

$$P_1^\mu = \frac{1}{2}P_h + R^\mu, \tag{3.72}$$

$$P_2^\mu = \frac{1}{2}P_h - R^\mu. \tag{3.73}$$

| **Distribution Functions** | **Model Identifier** |
|---|---|
| $f_1$ | CTEQ [74] |
| $f_1$ | LHAPDF [75] |
| $f_1$ | BCR08 [76] |
| $f_1$ | GRV98 [77] |
| $g_1$ | GRSV2000 [78] |
| $f_{1T}$, $h_{1T}^{\perp}$, $h_1$ | Torino Group [79, 80, 81, 82, 83] |
| $f_1$, $g_1$, $g_{1L}$, $g_{1T}$, $f_{1T}$, $h_1$, $h_1^{\perp}$, $h_{1T}^{\perp}$ | Pavia Spectator Model [31] |

Table 3.1: Models of distribution function available in `TMDGen`.

### 3.2.3 Code Organization

One of the guiding principles in writing the generator, besides including the necessary cross sections, is to allow a very general framework that can be updated and improved with the least disturbance to the existing code. The generator is written in C++ and uses object orient design capabilities to assist in this goal. Each major task or task category is associated with a parent class, such as throwing the dependent variables, evaluating a distribution or fragmentation function, and evaluating a term of the cross section or the entire cross section. Specialization is accomplished by defining child classes. For example, there exists a child class of the distribution function class for each implemented model, while for the cross section class, there exists separate child classes for pseudo-scalar and dihadron production. Generalizing `TMDGen` to related TMD processes, such as in proton-proton collisions or in *e*-*p* collider, would only require defining additional child classes where needed.

In a few cases, specializations are not provided by children class, but rather by defining additional output functions. For example, the main controlling `TMDGen` class includes functions for each possible output type, current choices include `ROOT` [72] and `DAD/ADAMO` [73]. The use of preproccessor `#define` and `#ifdef` allow one to select which output options are compiled into the code, and users from additional experiments need only to define an extra function, detailing how to output to their format, and disable `ROOT` and/or `DAD/ADAMO` by undefining the necessary macros.

### 3.2.4 Implemented Models

A variety of models are available for both distribution and fragmentation functions, with full lists given in Tables 3.1 and 3.2. Note that some of the $f_1$ models are duplicated by allowing access to the LHAPDF library. However, use of LHAPDF requires the user to have installed this library, while the other $f_1$ functions are provided within the `TMDGen` package. To minimize dependencies on other libraries, access to (and thus dependence on) the LHAPDF libraries can be turned off via a compile time option. All other distribution and fragmentation models are fully packaged within `TMDGen`.

| **Frag. Functions** | **Final State** | **Model Identifier** |
|---|---|---|
| $D_1$ | pseudo-scalar | fDSS [84, 85] |
| $D_1$ | pseudo-scalar | Kretzer [86] |
| $D_1$, $H_1^\perp$ | dihadron | Spectator Model (Section 2.4) |
| $D_1$, $H_1^\perp$ | dihadron | Set given partial wave proportional to any other partial wave |

Table 3.2: Models of fragmentation function available in `TMDGen`.

### 3.2.5 Comparison with Other Generators

In this section, 1D distributions of each of the 5D kinematic variables from both `Pythia` and `TMDGen` are compared. The GRV98 model [77] is used for the unpolarized $f_1$ distribution function, and the spectator model is used for the unpolarized $D_1^{|0,0\rangle}$ fragmentation function. The spectator model fragmentation function parameter sets for the pion and kaon-pair dihadrons under consideration are chosen (by hand) to match the given `Pythia` distributions, while the `Pythia` distributions have been optimized to match HERMES kinematics. The actual values of the parameters for the spectator model are given in Table 3.3. As the specific choices were obtained by hand, it is possible that a more optimal set of parameters may exist. Note, the given models poorly constrain the $p_T$, $k_T$ distributions, meaning that it is possible to have the measurable variables match fairly well for various choices of parameters, yet with drastically different $\langle p_T^2\rangle$, $\langle z^2k_T^2\rangle$ values.

The comparisons for the dihadrons related to the $\rho$-triplet and $\phi$ mesons are given in Figures 3.2 through 3.4. The plots are made within the kinematic region

$$\begin{aligned} Q^2 &> 1\ \text{GeV}^4/c^2, & W^2 &> 10\ \text{GeV}^2/c^4, \\ 0.023 < x &< 0.4, & 0.2 < y &< 0.95, \\ 0.2 < z &< 0.8, & 0.05 < P_{h\perp} &< 1.6\ \text{GeV}/c, \end{aligned}$$

with $M_h < 1.6\ \text{GeV}/c^2$ for the $\pi\pi$ dihadrons and $M_h < 1.05\ \text{GeV}/c^2$ for the $K^+K^-$ dihadrons.

In general, the 1D comparisons are quite close, except perhaps the $x$ distribution. This discrepancy in the $x$ distribution is related to the flavor balance, and in particular, the flavor dependence of $f_1$, which is not a parameter being adjusted. Note, the full 5D distribution shows some deviations from `Pythia`, but `Pythia` also fails to match the full multivariate distribution seen in HERMES data. The multivariate differences between `TMDGen` and `Pythia` are more clearly observed when comparing within acceptance. Although the 1D plots match in $4\pi$ acceptance, they no longer match as well within acceptance.

For the `TMDGen` generator, one can also plot the magnitude of the intrinsic transverse momenta, $p_T = |\boldsymbol{p}_T|$ and $k_T = |\boldsymbol{k}_T|$. It is common to consider the distribution of $zk_T$ rather than that of $k_T$. The resulting `TMDGen` distributions are given in Figure 3.5. Note that, for all dihadrons, it is required that $p_T \approx zk_T$ in order to obtain a narrow $P_{h\perp}$ distribution. Likewise, the narrowness of the $P_{h\perp}$ distribution shows that all dihadron subprocesses have similar $p_T$ and $k_T$ dependence.

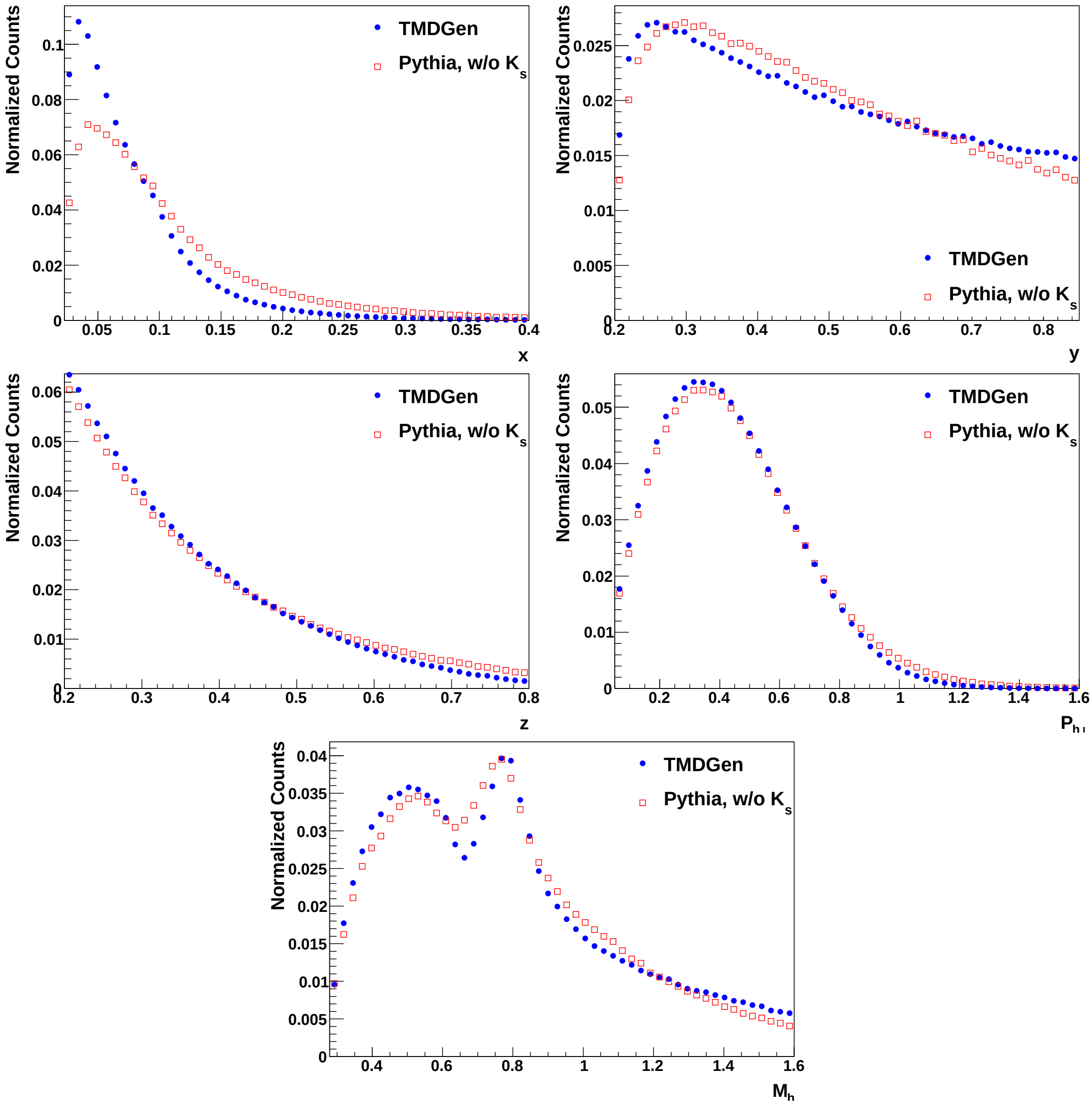


Figure 3.1: Comparison of 1D kinematic distributions from `TMDGen` and `Pythia`, in $4\pi$, for $\pi^+\pi^0$ dihadrons. Listing the rows from top to bottom, and within each row from left to right, the panels are respectively the $x$, $y$, $z$, $P_{h\perp}$, and $M_h$ distributions. `TMDGen` data is designated with blue circles, and `Pythia` data designated with red open squares.

| | $\pi\pi$ | $K^+K^-$ | | |
|---|---|---|---|---|
| **Parameter** | $u$, $\bar{d}$, $d$, $\bar{u}$ | $u$, $\bar{u}$ | $d$, $\bar{d}$ | $s$, $\bar{s}$ |
| $\alpha_s$ | 3.0 | 4.5 | 4.5 | 4.5 |
| $\beta_s$ | -0.751 | -0.125 | -0.125 | -0.125 |
| $\gamma_s$ | -0.193 | -0.6 | -0.6 | -0.6 |
| $\alpha_p$ | 7.0 | 0.0 | 0.0 | 4.5 |
| $\beta_p$ | -0.75 | 0.0 | 0.0 | -1.125 |
| $\gamma_p$ | -0.193 | 0.0 | 0.0 | -0.6 |
| $\alpha_k$ | 0.9 | 1.3 | 1.3 | 1.3 |
| $\beta_k$ | 0.125 | 0.6 | 0.6 | 0.6 |
| $\gamma_k$ | -0.6 | -0.8 | -0.8 | -0.8 |
| $M_s$ | 1.5 | 3.0 | 3.0 | 3.0 |
| $f_s$ | 500.0 | 500.0 | 200.0 | 500.0 |
| $f_\rho$, $f_\phi$ | 150.0 | 0.0 | 0.0 | 350.0 |
| $f_\omega$ | 0.63 | | | |
| $f_{\omega'}$ | 150.0 | | | |

Table 3.3: Parameter sets for spectator model of dihadron production. Parameters for the strange quark flavor in $\pi\pi$ production are set to zero. The quantities $f_\omega$ and $f_{\omega'}$ are not relevant for $K^+K^-$ production, and thus the corresponding entries in the above table are intentionally left blank.

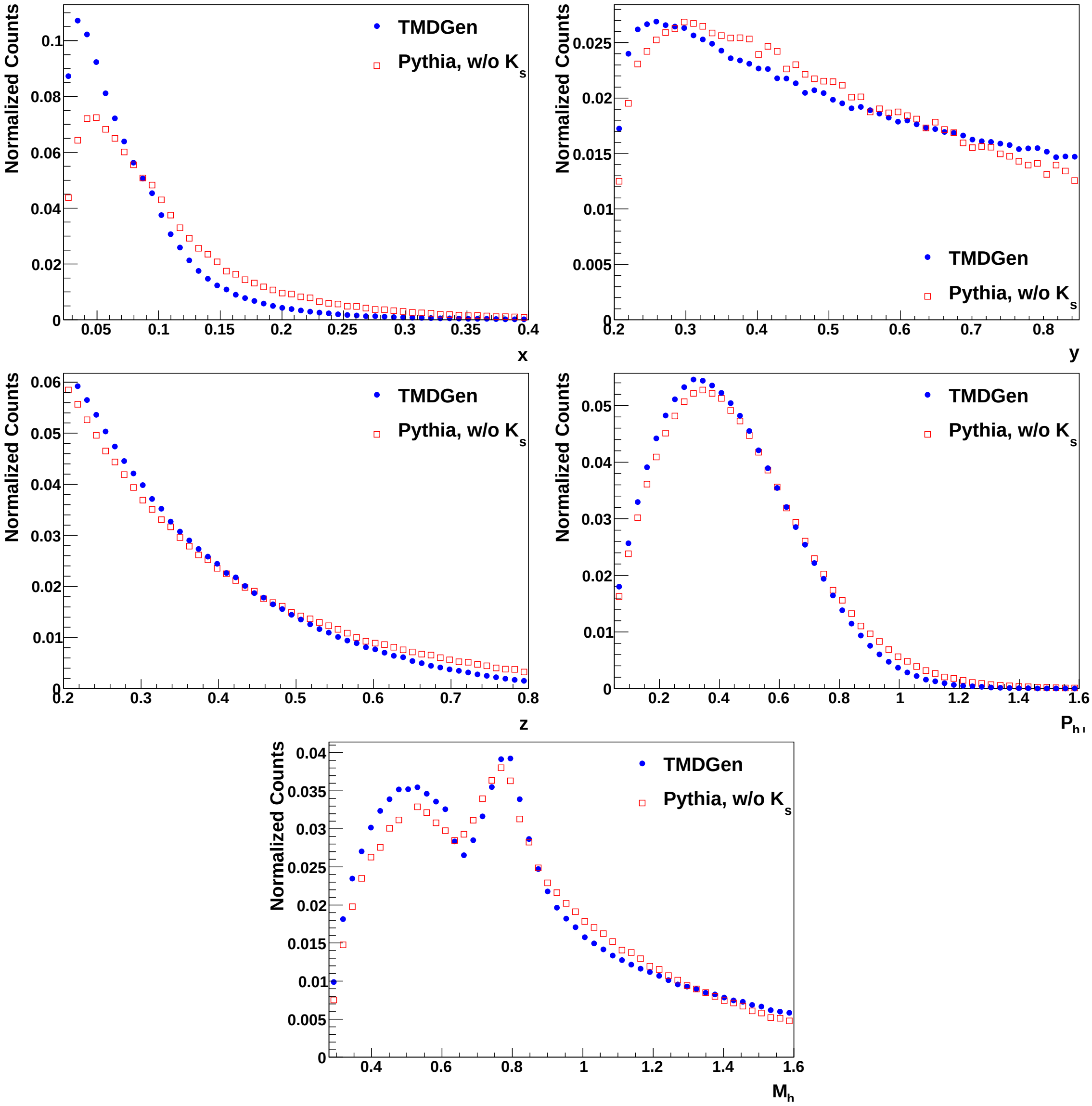


Figure 3.2: Comparison of 1D kinematic distributions from `TMDGen` and `Pythia`, in $4\pi$, for $\pi^+\pi^-$ dihadrons. Panels and markers are as in Figure 3.1.

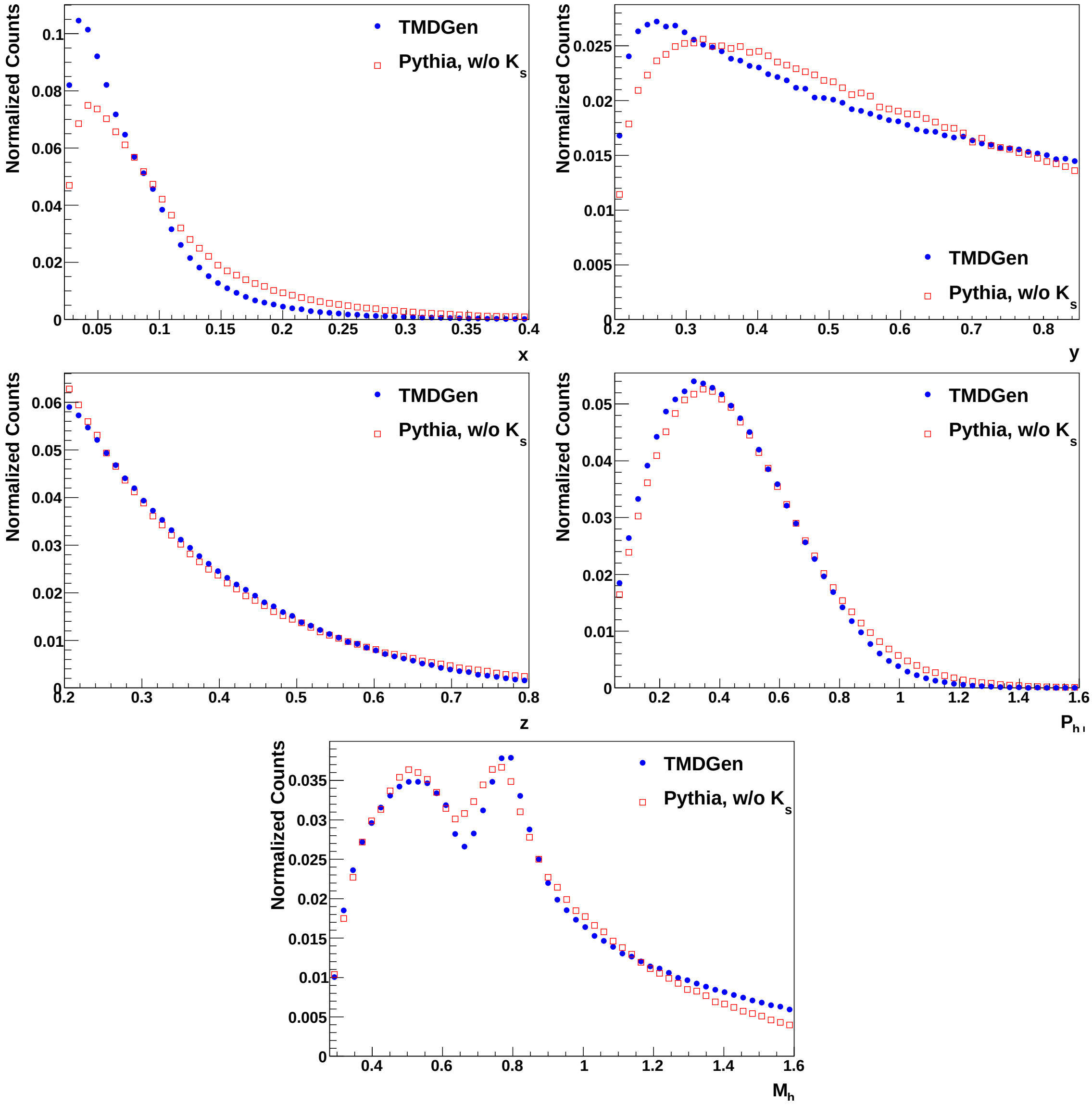


Figure 3.3: Comparison of 1D kinematic distributions from `TMDGen` and `Pythia`, in $4\pi$, for $\pi^-\pi^0$ dihadrons. Panels and markers are as in Figure 3.1.

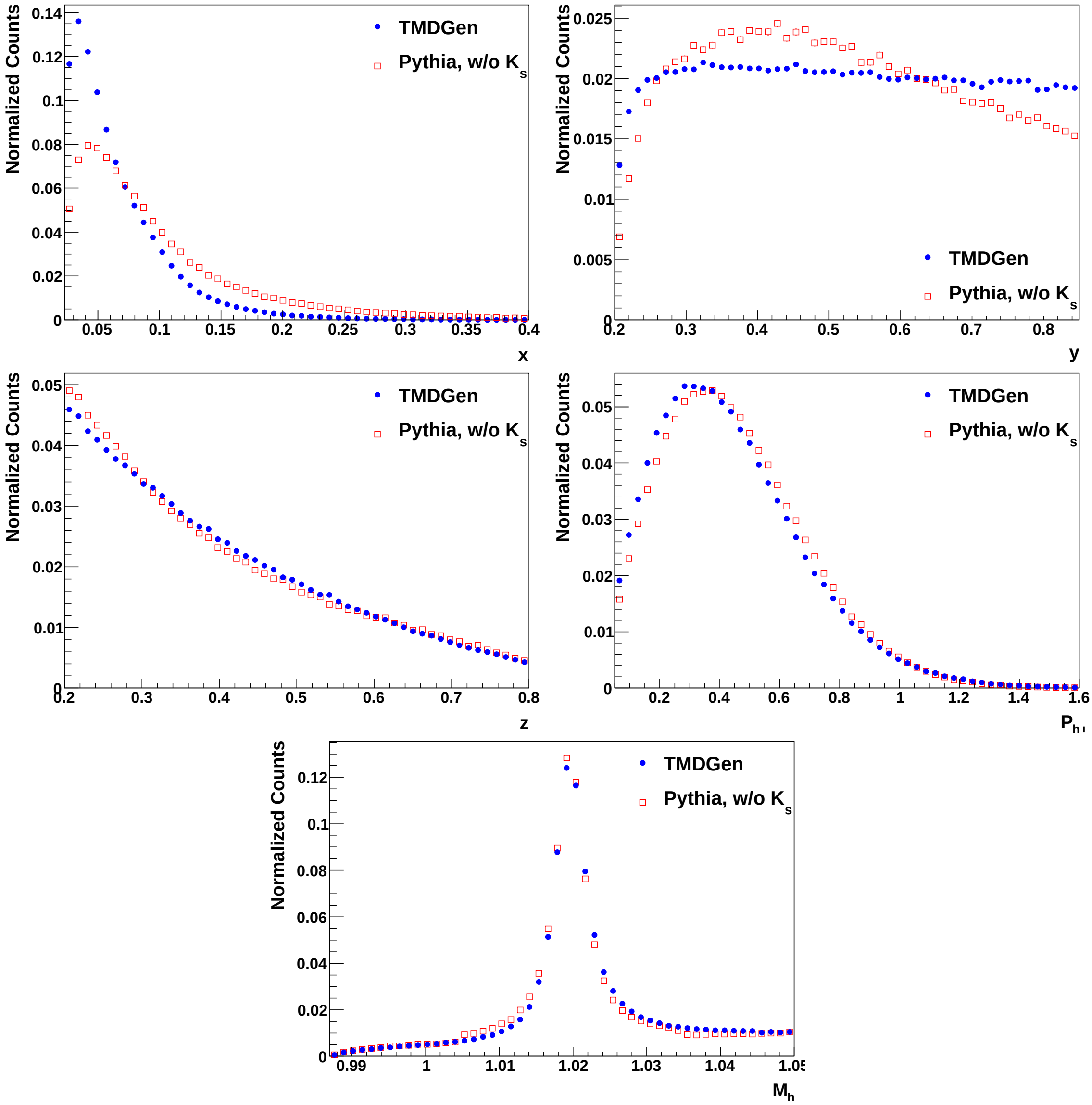


Figure 3.4: Comparison of 1D kinematic distributions from `TMDGen` and `Pythia`, in $4\pi$, for $K^+K^-$ dihadrons. Panels and markers are as in Figure 3.1.

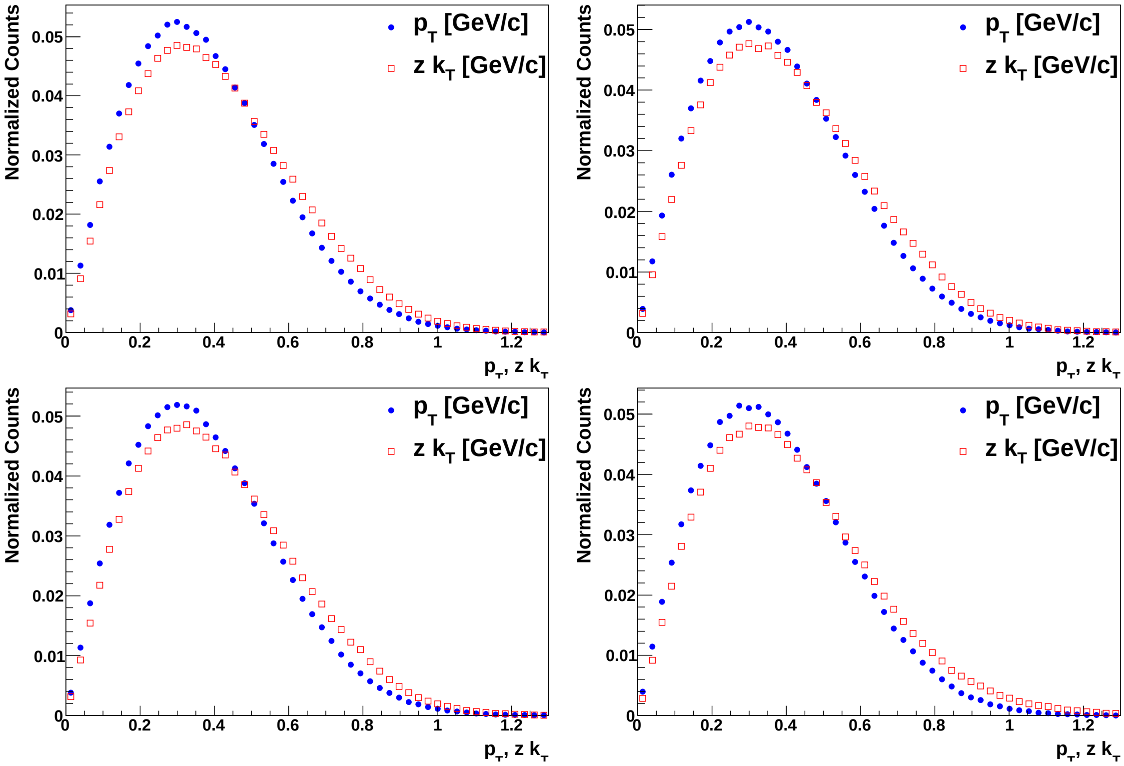


Figure 3.5: The $p_T$ and $zk_T$ distributions from `TMDGen` for all four dihadrons. Clock-wise from upper left, the panels are for $\pi^+\pi^0$, $\pi^-\pi^0$, $K^+K^-$, and $\pi^+\pi^-$. The distribution for $p_T$ is shown in blue circles, while $zk_T$ is shown in red open squares.

# CHAPTER IV

# Analysis

## 4.1 Experimental Setup

The HERMES spectrometer was located at the HERA storage ring from 1995 through 2007. HERMES utilized only the lepton storage ring, with the lepton beam incident on a fixed gas target, while the other experiments at the HERA ring used either just the proton storage ring (HERA-B) or both the proton and lepton rings (H1and ZEUS). HERA was located at the Deutsches Electronen Synchotron (DESY) laboratory in Hamburg, Germany. In June of 2007, the HERA ring and the three experiments still utilizing the ring (H1, ZEUS, and HERMES) were decommissioned. At this time PETRA, the injection ring used to fill HERA, began being used for other purposes.

### 4.1.1 The Spectrometer

The HERMES spectrometer consisted of a large number of different subsystems, with the upper portion being nearly perfectly symmetric with the lower portion. A schematic drawing, applicable for the 2002-2005 running period, is shown in Figure 4.1. Details of various components of the spectrometer can be found in many sources, with most complete technical reference being Ref. [87]. Many HERMES PhD dissertations also include detailed descriptions, e.g. [60, 88, 89, 38]. The components most relevant for this analysis are discussed in the following paragraphs.

**The Target** The target consists of a gas filled cell [90], the cell being an open ended tube within the lepton beam pipe. During 2002-2005, the cell was 40 cm long and had an elliptical cross section of 29 mm × 9 mm. The target cell could be filled with a variety of gases, including $H_2$, $D_2$, $^3$He, $N_2$, Xe and Kr, with the heavier gases being used near the end of the lepton fill. The cell was accompanied by a pump system to ensure the gas did not disperse into the lepton beam pipe, and the cell was also continuously fed by an atomic beam source (ABS). The ABS utilized radio frequency transitions between certain hyperfine states to produce atoms with polarized nuclei and unpolarized electrons. During 2002-2005, the nuclei were polarized transverse to the direction of the incoming lepton beam [91], while in other years the nuclei were either unpolarized or longitudinally polarized.

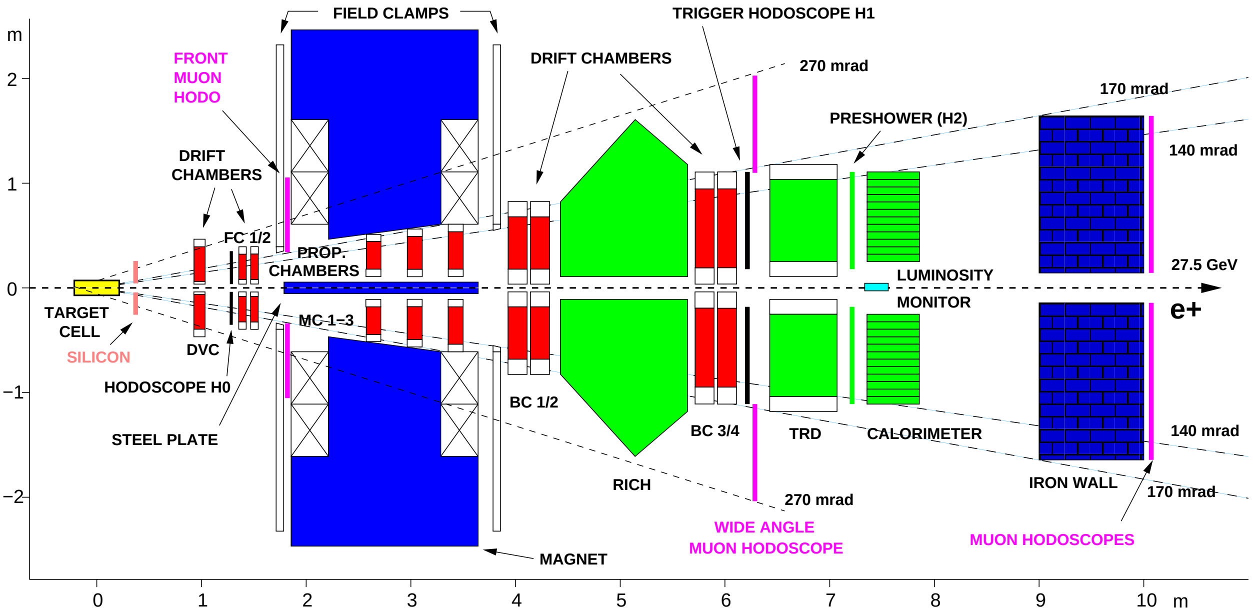


Figure 4.1: Schematic drawing of the 2002-2005 setup of the HERMES spectrometer. Taken from Ref. [87].

**Hodoscopes** Three scintillating hodoscopes are present [92], identified as H0, H1 and H2 in Figure 4.1. The three hodoscopes are made of paddles 1 cm thick and 9 cm wide and are connected to photomultiplier tubes. The paddles are also overlapped 2-3 mm to increase efficiency. H0 serves to veto particles originating from HERA's proton beam. The H1 and H2 hodoscopes are used for time of flight information. This, in turn, can be combined with the momentum determination from the tracking systems to compute the particle's mass, and therefore determine its identification. Furthermore, 11 mm of lead is installed in front of H2 to aid in lepton/hadron separation. The lead causes the lepton to shower, and thus leptons tend to deposit significantly more energy than hadrons. For this reason, H2 is also called the preshower.

**Transition Radiation Detector (TRD)** The TRD is designed to discriminate between hadrons and leptons and consists of a 6 cm thick 2D matrix of dielectric fibers. Electrons and positrons produce transition radiation which is measured in proportional wire chambers. Pions also interact with the TRD through ionization, though the amount energy deposited is much higher for leptons than for hadrons.

**Calorimeter** The electromagnetic calorimeter is designed to measure the energy of electrons, positrons and photons. It consists of 840 lead-glass blocks (420 per detector half) which are 9 cm square in cross section and 50 cm (18 radiation lengths) long. As hadrons do not deposit much energy in the calorimeter, the energy deposited, divided by the momentum determined by the tracking systems, can be used to distinguish leptons and hadrons. The calorimeter is also able to determine the position of photons to about 0.5 cm [93]. Combining information from the TRD, preshower and calorimeter allows a lepton-hadron separation efficiency of about 98%, with a contamination rate less than 2%.

**Ring Imaging Čerenkov Detector** The RICH is a dual radiator ring imaging Čerenkov Detector [94] and is designed to determine whether given hadrons are pions, kaons, or protons. The two (dual) radiators are a wall of silica aerogel tiles and a gas radiator volume of $C_4F_{10}$. The speed of light in the radiators is greater than the typical speed of the particles, and thus a cone of Čerenkov light is produced. This light is reflected by mirrors and detected by photomultiplier tubes. The opening angle of the cone can be detected with a resolution of about 7.2 mrad. The distributions of opening angle versus momenta for each particle type to not overlap over most of the momentum range, which allows determination of the particle identification. Unfortunately, at low momentum the distributions begin to overlap, thus impeding proper identification. Kaons and protons cannot be distinguished from each other if the momentum is below 2 GeV/$c$, and particle identification is not reliable for tracks with momenta below 1 GeV/$c$. An event level identification procedure (EVT) was also developed in 2007 [38], which superseded the next best algorithm, which is based on direct ray tracing (DRT) [95]. The EVT algorithm is particularly advantageous when the Čerenkov rings from more than one particle overlap.

**DIS Trigger** Although not a specific hardware component, an efficient trigger system is utilized to decide which events to store. While a given event is being stored, the detector cannot take new data, thus causing some dead time. One trigger, trigger 21 or the DIS trigger, requires coincident signals from the two forward hodoscopes (H1 and H2) and an energy deposition above a certain threshold in two adjacent columns of the calorimeter. This trigger serves to identify candidate deep inelastic scattering events through the detection of a high energy scattered lepton.

### 4.1.2 Further Considerations

During portions of the 2002-2005 running period, the HERMES target chamber was filled with transversely polarized hydrogen with an average polarization of around 75%. For this reason, these years are known as the transverse target running period. The lepton beam was longitudinally polarized, with an average polarization of about 85%. The beam energy was 27.6 GeV/$c^2$. The lepton beam consisted of positrons during 2002-2004 and of electrons in 2005.

During the transverse target running period, a transverse target magnet was additionally installed. This magnet defines the quantization axis for the target polarization and also increased the relaxation time of the polarization. However, the transverse target magnet was not included in the previously existing reconstruction code, and two transverse magnet correction methods were implemented [96]. Later, careful measurements of the field map were taken, and a new tracking code, denoted HTC, incorporated all magnetic fields and an advanced event level vertex finding algorithm. The analysis presented in this document utilizes the HTC procedure, while Ref. [27] utilized the transverse magnet correction codes. The HTC algorithm determines a common vertex, given several tracks assumed to originate

| Year | R.F. Freq. | R.F. Offset |
|---|---|---|
| 2002 | 500 | 260 |
| 2003 | 500 | 260 |
| 2004 | 500 | 300 |
| 2005 | 500 | 360 |

Table 4.1: Values used for the beam energy correction.

from a common vertex. The algorithm also provides a measure of the probability that all tracks did indeed originate from the common vertex, as well as the individual probability for each given track having originated from the vertex.

The lepton beam energy has been found to vary slightly over time. However, a correction method has been determined, based on the frequency and offset of the radio frequency cavities of the HERA accelerator. The fractional shift in energy between the actual $E$ and nominal $E_0$ beam energy is

$$\frac{E - E_0}{E_0} = -\frac{1}{\alpha}\frac{\delta f}{f}, \tag{4.1}$$

with $\alpha = 4.7 \times 10^{-4}$, and with $f$ and $\delta f$ being the frequency and offset values for the radio frequency cavities. The specific values of $f$ and $\delta f$ are given in Table 4.1.

## 4.2 Data Selection Requirements

Analyzable, polarized hydrogen runs are selected using the standard HERMES burst lists. As a sanity check, the target type and polarization state are also checked for each event. All tracks are required to be “long,” i.e. to reach the back portion of the spectrometer, as well as to flag trigger 21, the DIS trigger. Leptons are identified according to the sum of PID variables 3 and 5 being positive, while hadrons are associated with negative values of this sum. The sum is related to a combined analysis of four hardware components, including the RICH, the TRD (transition radiation detector), preshower and calorimeter. Tracks with an absolute value of this sum being larger than 100 are also rejected, as these values are considered spurious. Note that photons are identified as clusters in the calorimeter without associated charged tracks and are distinct from neutrons due to the energy deposited.

It is also required that the event vertex, as determined by the HTC tracking algorithm, be within ±20 cm of the center of the target, corresponding roughly to the target length. A cleaner lepton sample is obtained by requiring the sum of PID variables 3 and 5 be greater than unity for leptons. For $\pi^+\pi^-$ and $K^+K^-$ dihadrons, it is required that the two hadron tracks have opposite charge. Hadronic identification is accomplished with the EVT method, or DRT if EVT fails. The lepton momentum is required to be less than the beam energy. The momentum of pions is required to be within 1 to 15 GeV/$c$, and for kaons 2 to 15 GeV/$c$. The upper hadron momentum cut is to avoid spurious data, while the lower hadron momentum cut is required for clean RICH identification. The individual HTC track probabilities are also required to be greater than 0.01, while the overall HTC vertex probability is required to be

| **Physical Object** | **Cut** |
|---|---|
| Front field clamp | $\lvert X_{off} + 172.0 \tan\theta_X \rvert < 31$ cm |
| Vertical lower limit (septum plate) | $\lvert Y_{off} + 181.0 \tan\theta_Y \rvert < -7$ cm |
| Rear field clamp | $\lvert Y_{off} + 383.0 \tan\theta_Y \rvert < 54$ cm |
| Rear clamp | $\lvert X_{pos} + 108.0 X_{sl} \rvert < 100$ cm |
| | $\lvert Y_{pos} + 108.0 Y_{sl} \rvert < 54$ cm |
| Calorimeter | $\lvert X_{pos} + 463.0 X_{sl} \rvert < 175$ cm |
| | $30 < \lvert Y_{pos} + 108.0 Y_{sl} \rvert < 108$ cm |

Table 4.2: Fiducial volume cuts for charged particle tracks.

| **Physical Object** | **Cut** |
|---|---|
| Calorimeter | $\lvert X \rvert < 125$ cm |
| | $33 < \lvert Y \rvert < 105$ cm |

Table 4.3: Fiducial volume cuts for trackless calorimeter clusters.

greater than $10^{-5}$. To reduce background, a minimum value of 0.8 GeV$/c^2$ for the corrected cluster energy is also enforced, with the cluster energy correction described in Section 4.2.1. A data quality bit mask of `0x427ffffd` is used. This incorporates a large number of common data quality parameters, with each bit being defined in detail in Ref. [97].

Each track is also required to be within the standard HERMES fiducial volume. For charged particle tracks, one defines $\tan\theta_X$ and $\tan\theta_Y$ as the tangent of the angles between the track momentum and the $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ axes in the HERMES coordinate system, system $\mathbf{I_a}$. The fiducial volume cuts depend on these angles, as well as the $x$ and $y$ offsets, slopes, and positions, ($X_{off}$, $Y_{off}$, $X_{sl}$, $Y_{sl}$, $X_{pos}$, $Y_{pos}$) as given in the HERMES production files. The specific cuts are given in Table 4.2.

For energy clusters in the calorimeter without charged tracks, such as those caused by photons, a different set of fiducial volume cuts are used. In this case, the cuts are based on the $X$ and $Y$ positions of the clusters, with the specific cuts given in Table 4.3. The cuts are designed to ensure the clusters are not too near the edges of the calorimeter.

In addition to the above selection requirements related to data quality and track reconstruction, restrictions are also placed on the relevant kinematic variables. Specifically, it is required that

$$\begin{aligned} &Q^2 > 1 \text{ GeV}^2/c^4, && 0.023 < x < 0.4, \\ &W^2 > 10 \text{ GeV}^2/c^4, && 0.2 < y < 0.95, \\ &0.05 < P_{h\perp} < 1.6 \text{ GeV}/c, && 0.2 < z < 0.8. \end{aligned} \tag{4.2}$$

As the theory is applicable in the high $Q^2$ limit, it is required that $Q^2 > 1$. The lower limit on $W^2$ is to ensure the event is diffractive. Radiative effects become much more prominent at higher $y$, and thus $y$ is limited below 0.95. Many HERMES analyses use a more conservative

| **Var.** | **Bin Edges** | | | | |
|---|---|---|---|---|---|
| $x$ | 0.023 | 0.04 | 0.055 | 0.085 | 0.40 |
| $y$ | 0.20 | 0.60 | 0.70 | 0.80 | 0.95 |
| $z$ | 0.20 | 0.34 | 0.44 | 0.56 | 0.80 |
| $P_{h\perp}\left[\mathrm{GeV}/c\right]$ | 0.05 | 0.30 | 0.50 | 0.75 | 1.6 |

Table 4.4: Bin edges for the kinematic variables.

| **Dihadron** | $\boldsymbol{M_h}$ **Bin Edges [GeV/$\boldsymbol{c^2}$]** | | | | |
|---|---|---|---|---|---|
| $\pi\pi$ | 0.279 | 0.450 | 0.640 | 0.900 | 1.600 |
| $K^+K^-$ | 0.987 | 1.012 | 1.028 | 1.050 | |

Table 4.5: Bin edges for the invariant mass $M_h$.

upper $y$-cut of 0.85. Justification that this analysis can use a higher threshold without incurring significant radiative effects is based on the studies in Section 5.1. The lower limit on $P_{h\perp}$ is to ensure good resolution in $\phi_h$, as $\phi_h$ is ill-defined in the limit of $P_{h\perp}$ going to zero. The lower limit on $z$ is to ensure the produced hadron comes from the struck quark (i.e., the current fragmentation region) rather than being a target remnant (i.e., the target fragmentation region). The upper limit in $z$ is related to the exclusive vector meson background, and is further discussed in Section 4.5.4. The other requirements Equation 4.2 that are not specifically mentioned in this paragraph do not limit the data significantly, but instead are used to consistently define the domain of the cross section.

Fits are preformed in either 1D bins, in $M_h$, or in 2D binning, $M_h$ versus one of $x$, $y$, $z$, $P_{h\perp}$. The fitting functions are parametric with respect to only the angular variables, and the fit parameters represent integrals over the given kinematic bins. More details regarding fitting is given in Section 4.4. The specific kinematic bins are given in Table 4.4, while Table 4.5 lists the bins in $M_h$. The lower limits on $M_h$ are based on the production threshold for pion and kaon-pairs. The upper limit on $M_h$ for pion-pairs are to provide a consistent integration range, while for kaon-pairs, it is to identify an appropriate sideband region with invariant mass $M_h$ in the range just above the upper edge of the the $\phi$ meson peak. A significant amount of data for pure SIDIS kaon-pair production exists above the 1.05 GeV/$c^2$ threshold, but this data has yet been analyzed. For pion-pairs, one can identify in Table 4.5 two bins in the invariant mass region below the $\rho$ meson mass peak and one bin above the $\rho$ meson mass peak, while for kaon-pairs, bins are provided for the $\phi$ meson peak and one sideband region on either side.

### 4.2.1 Reconstruction of Neutral Pions

Neutral pions decay nearly immediately, and so the decay vertex, within detector resolution, is identical to the primary vertex. The neutral pions decay into two photons with a branching ration of greater than 98.8% [37]. Photons are identified in the detector as energy clusters in the calorimeter with no accompanying charged particle track. As the calorimeter is calibrated for leptonic clusters, the measured energy needs to be multiplied by a given factor for photonic clusters. For HERMES data, the factor is 0.97. For `Pythia` Monte Carlo data, the energy correction factor is 0.9255, while for `TMDGen` data the factor is 0.95. Note, the calibration of the calorimeter energy for photonic clusters is a known issue with the current HERMES productions, and is being investigated by the collaboration. Thus, some analyzers have used more complex energy correction factors, and the exact choice is still open. The given factors have been chosen to optimize the mean $\pi^0$ mass.

Assuming the two photons come from an single parent, the mass of the parent can be computed according to [88]

$$M_{\gamma\gamma} = 2\sqrt{E_{\gamma_1} E_{\gamma_2}} \sin \frac{\theta_{\gamma\gamma}}{2}, \tag{4.3}$$

with $E_{\gamma_1}$, $E_{\gamma_2}$ the energies of the two photons and $\theta_{\gamma\gamma}$ the angle between the photons' momenta, in the HERMES coordinate system ($\mathbf{I_a}$).

To determine the angle between the two photons, it is necessary to determine, for each cluster, the spatial distance between the location of the primary vertex and the center of the cluster in the calorimeter. The $x$ and $y$ position of the cluster are given in the standard HERMES data files, and the tracking methods provide the location of the vertex. Note, that the $x$ and $y$ position of the vertex is not used, as the magnitude of these positions is much smaller than other distances in this calculation. Also, historically, the $x$ and $y$ positions were not available in all tracking methods or were not considered reliable. It is not possible to determine exactly the $z$-position of the center of the cluster. Studies performed by other HERMES analyzers suggest using the value of 734 cm for photons which interact with the preshower (which is the case for the majority of photons in the kinematic region of interest), and a value of 747.5 cm for those that do not. Other studies have shown that the results for the invariant mass spectrum are not very sensitive to the choice of $z$-vertex.

Figure 4.2 shows the invariant mass of the two photon system $M_{\gamma\gamma}$ for the HERMES data set, within acceptance. No difference is seen in the $M_{\gamma\gamma}$ distribution in the positron and electron data samples, and so Figure 4.2 is the combined data set. Note, the peak position, and uncertainty from the fit, are $136.43 \pm 0.08$ MeV/$c^2$ and $136.42 \pm 0.09$ MeV/$c^2$, respectively, for the $\pi^+\gamma\gamma$ and $\pi^-\gamma\gamma$ data sets. This is quite close to the accepted pion mass of 139.6 MeV/$c^2$ [37]. The Gaussian $\sigma$ values are, respectively, $11.88 \pm 0.08$ MeV/$c^2$ and $12.06 \pm 0.09$ MeV/$c^2$, reflecting the detector resolution in determining the $\pi^0$ mass.

Comparisons of the $M_{\gamma\gamma}$ distribution between data and Monte Carlo are provided in Figure 4.3. The main difference between data and Monte Carlo is the is non-resonant $\gamma\gamma$ pair background, due to both combinatorics and other processes.

Based on Figure 4.2, $\pi^0$ events are identified by $0.115 < M_{\gamma\gamma} < 0.115$ GeV/$c^2$. The background fractions, based on the fits included in Figure 4.2, are 26.1% and 25.1%, respectively, among $\pi^+\pi^0$ and $\pi^+\pi^-$ candidate events. Discussion regarding correcting for this background is found in Section 4.5.1.

### 4.2.2 Vector Meson Reconstruction

The vector mesons also decay before moving any measurable distance and can be identified by a resonant peak in the spectrum of the invariant mass of the produced hadron system.

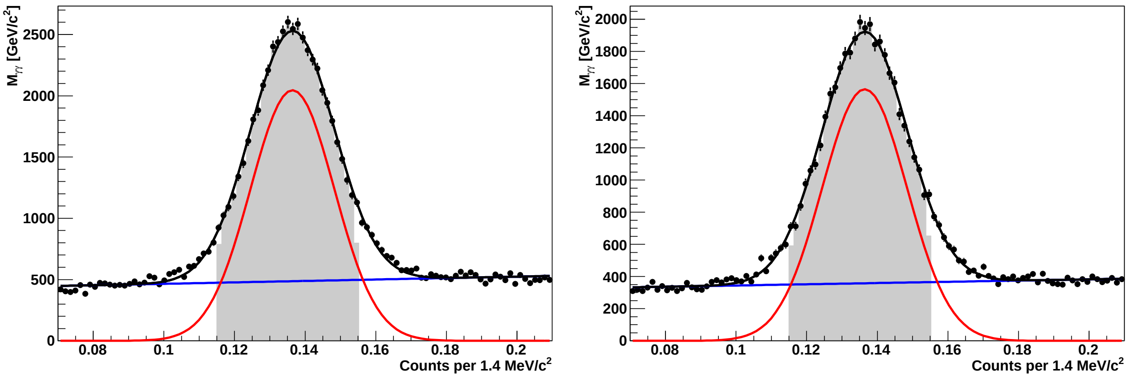


Figure 4.2: Invariant mass of the two photon system among $\pi^+\gamma\gamma$ (left panel) and $\pi^-\gamma\gamma$ (right panel) events from the full HERMES data set. A fit is provided, using a Gaussian distribution plus a linear background. The total fit line is drawn in black, while the individual Gaussian and background portions are drawn in red and blue, respectively. The gray shaded region, from $0.115 < M_{\gamma\gamma} < 0.155$ GeV$/c^2$, is defined as the $\pi^0$ mass window.

Non-SIDIS processes can also be present within the kinematic range, specifically exclusive vector meson and exclusive hadron pair production. Section 4.5.3 further discusses these non-SIDIS backgrounds.

Even within SIDIS production, a number of subprocesses contribute to the dihadron process. The `Pythia` prediction of the $M_h$ spectrum for several contributing processes is given in Figure 4.4. The processes contributing to the $\pi^\pm\pi^0$ spectrum are very similar to each other. These processes are also similar to those contributing to the $\pi^+\pi^-$ distribution, with the exception that the $\pi^+\pi^-$ distribution also includes contributions from $\eta$, $\eta'$ and $K_{0,S}$. The $K^+K^-$ spectrum includes purely resonant $\phi$ mesons and non-resonant kaons pairs, without the variety of additional subprocesses seen in the pion-pair distributions.

Figure 4.5 shows HERMES data fit to a Breit-Wigner distribution plus a linear background. The relative fraction of the vector meson signal versus other dihadron processes, along with possible extraction of the vector meson signal, is discussed in Section 4.5.4.

## 4.3 Kinematic Distributions in Acceptance

In Section 3.2.5, the `Pythia` and `TMDGen` Monte Carlo generators were compared within prefect $4\pi$ acceptance and with no radiative corrections (the Born level for the hard quark, virtual-photon vertex). One can also compare the Monte Carlo generators within acceptance. For the `Pythia` generator, radiative effects are also included using `RadGen` [98]. However, radiative effects have not yet been implemented in `TMDGen`.

The comparison is again given in 1D projections for each of the kinematic variables, $M_h$, $x$, $y$, $z$, $P_{h\perp}$, as well as the $\cos\vartheta$ distribution. These distributions are shown in Figures 4.6, 4.7, 4.8, and 4.9 for $\pi^+\pi^0$, $\pi^+\pi^-$, $\pi^-\pi^0$, and $K^+K^-$ dihadrons, respectively. For the dihadrons involving neutral pions, the two photon invariant mass $M_{\gamma\gamma}$ was already compared in Figure 4.3. The `TMDGen` generator was set to have all angular distributions turned off for

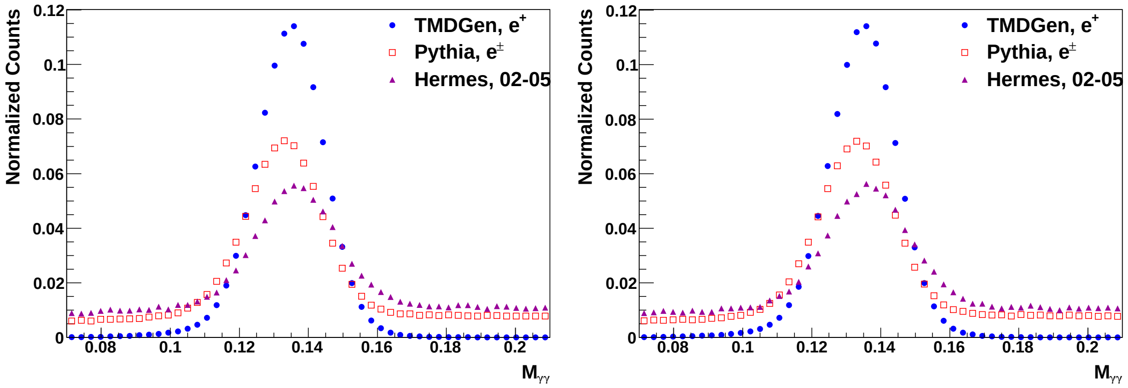


Figure 4.3: Comparison of the $M_{\gamma\gamma}$ distribution for $\pi^+\gamma\gamma$ (left panel) and $\pi^-\gamma\gamma$ (right panel) events for HERMES, `Pythia`, and `TMDGen` data within acceptance. Data from the `TMDGen` generator is shown in blue circles, `Pythia` with red open squares, and HERMES with purple triangles.

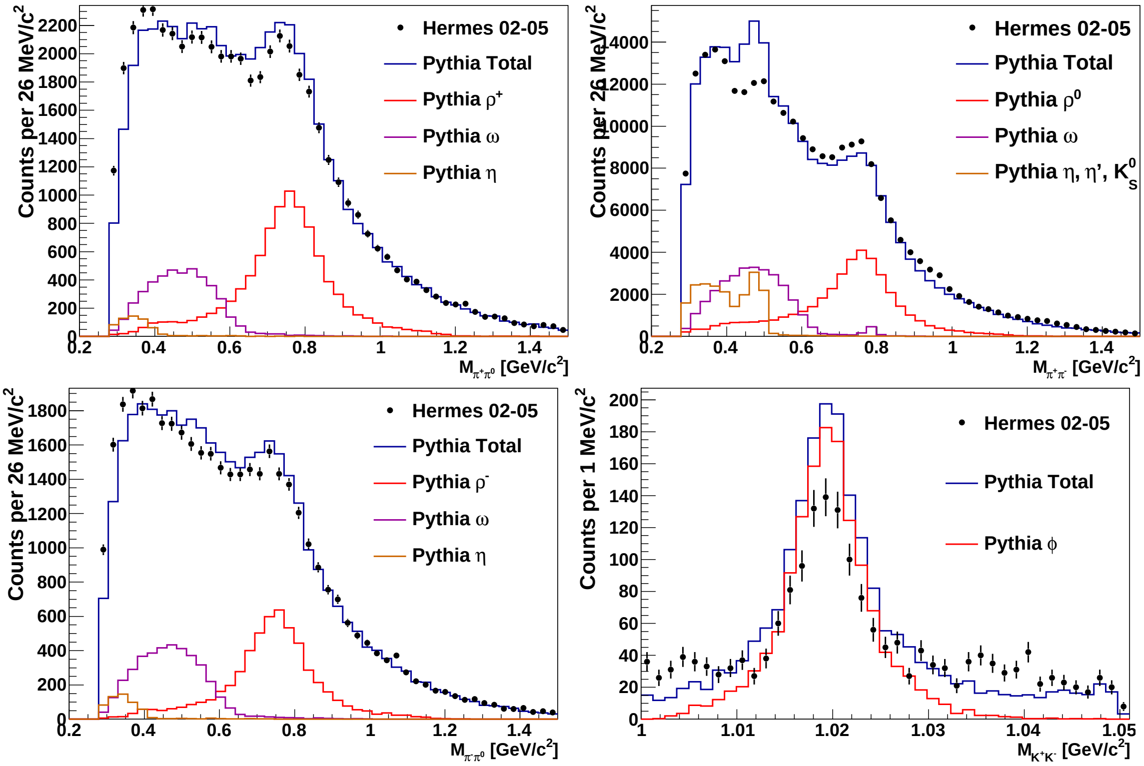


Figure 4.4: `Pythia` prediction of processes contributing to the dihadron invariant mass spectrum. Clock-wise from upper left, the panels are for $\pi^+\pi^0$, $\pi^+\pi^-$, $K^+K^-$, and $\pi^-\pi^0$ dihadrons. Black data points are HERMES data, 2002-2005, and a few select subprocesses are as indicated. In particular, the vector meson channel is indicated in red. Note the $\rho$ meson peak near 0.770 GeV$/c^2$ and the $\phi$ meson peak near 1.02 GeV$/c^2$.

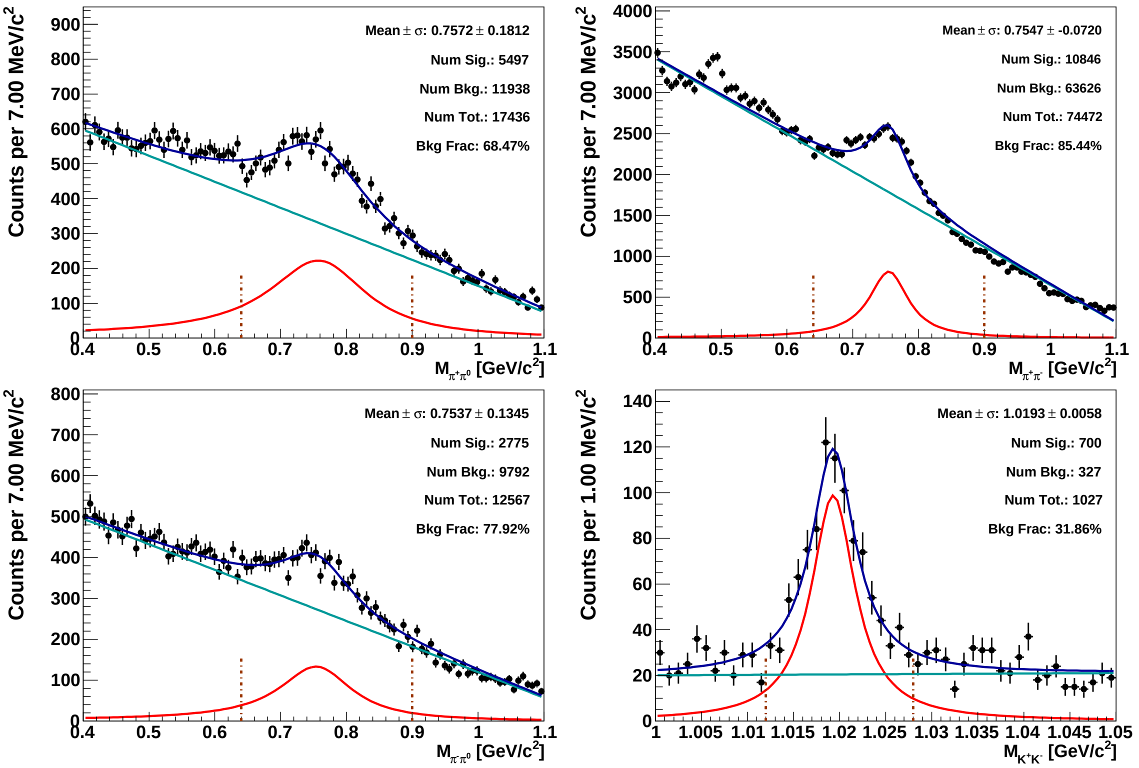


Figure 4.5: Dihadron invariant mass spectrum from HERMES data. Clock-wise from upper left, the panels are for $\pi^+\pi^0$, $\pi^+\pi^-$, $K^+K^-$, and $\pi^-\pi^0$ dihadrons. HERMES data is shown with black data points, with the fit shown in blue, and the Breit-Wigner and linear background functions separately drawn in red and cyan, respectively. The mass bin including the vector meson peak is indicated with vertical brown dash-dotted lines.

the kinematic plots in Figures 4.6 through 4.9. Thus, the $\cos\vartheta$ distributions show exactly the effect of acceptance on the distribution of this variable.

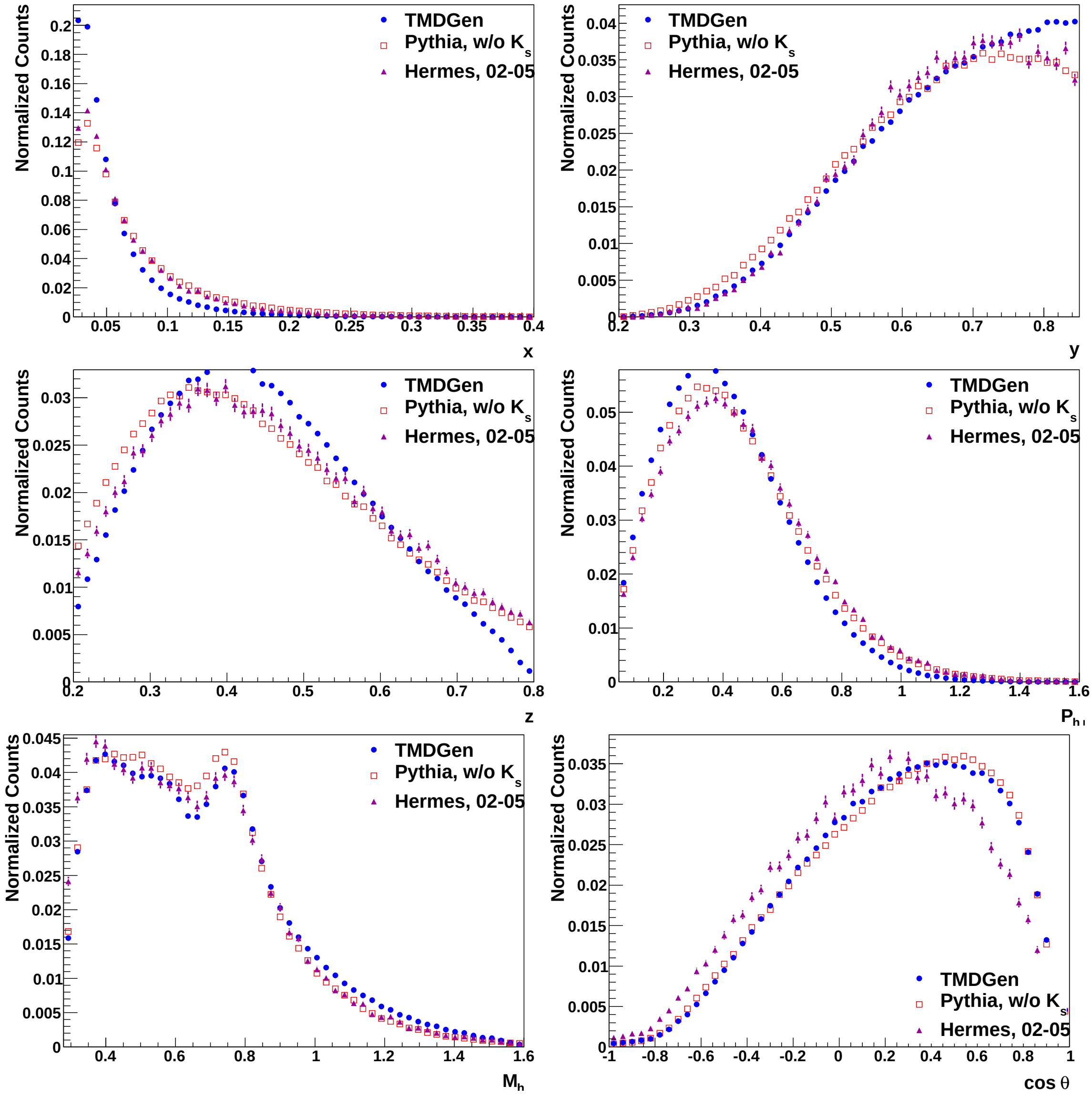


Figure 4.6: Comparison of kinematic variables for Hermes, `Pythia`, and `TMDGen` for $\pi^+\pi^0$ dihadron data within acceptance. The panels, clock-wise from upper left, are the $x$, $y$, $P_{h\perp}$, $\cos\vartheta$, $M_h$, and $z$ distributions. Data from the `TMDGen` generator is shown in blue circles, `Pythia` with red open squares, and Hermes with purple triangles.

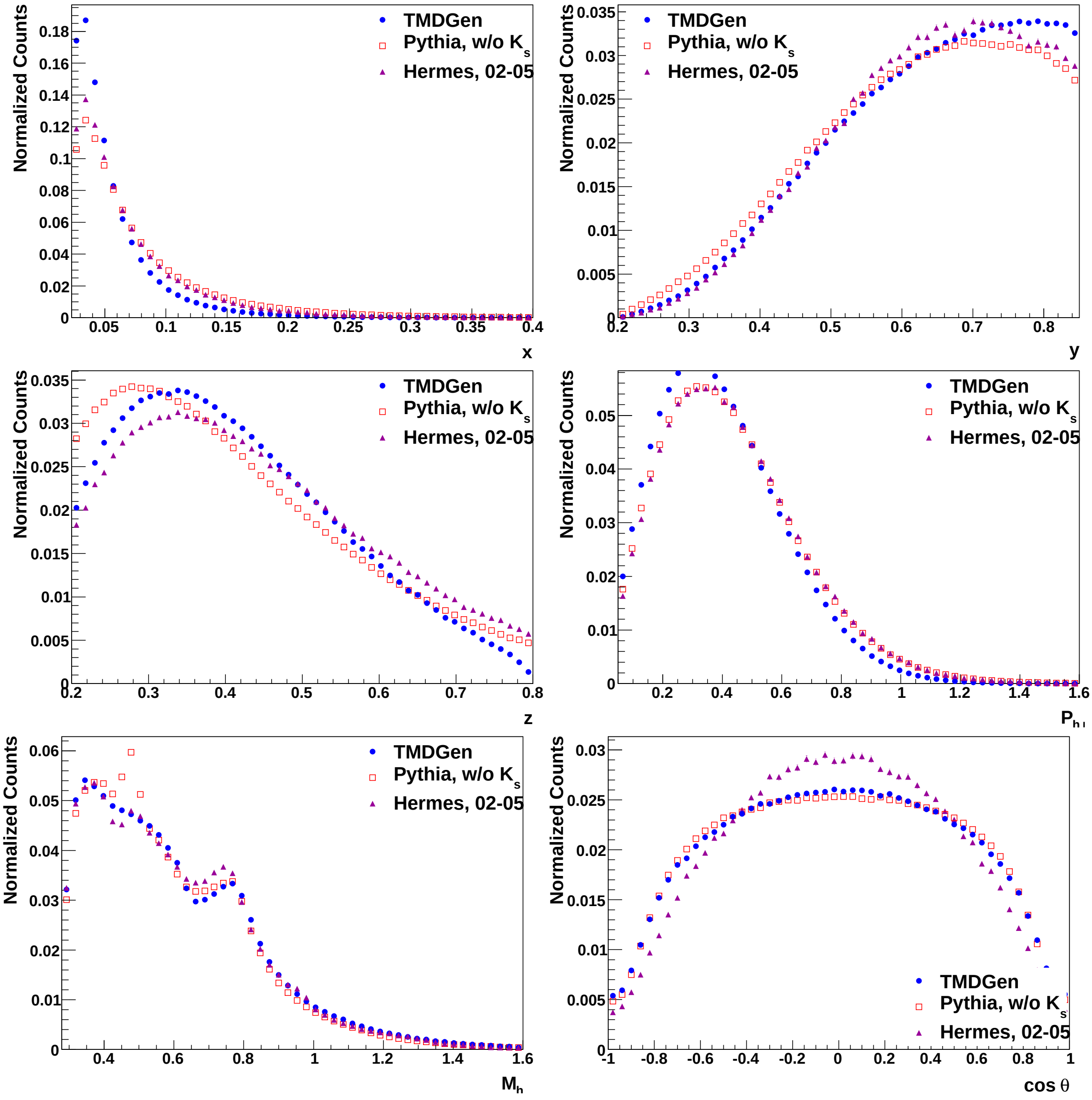


Figure 4.7: Comparison of kinematic variables for HERMES, `Pythia`, and `TMDGen` for $\pi^+\pi^-$ dihadron data within acceptance. Panels and markers are the same as in Figure 4.6.

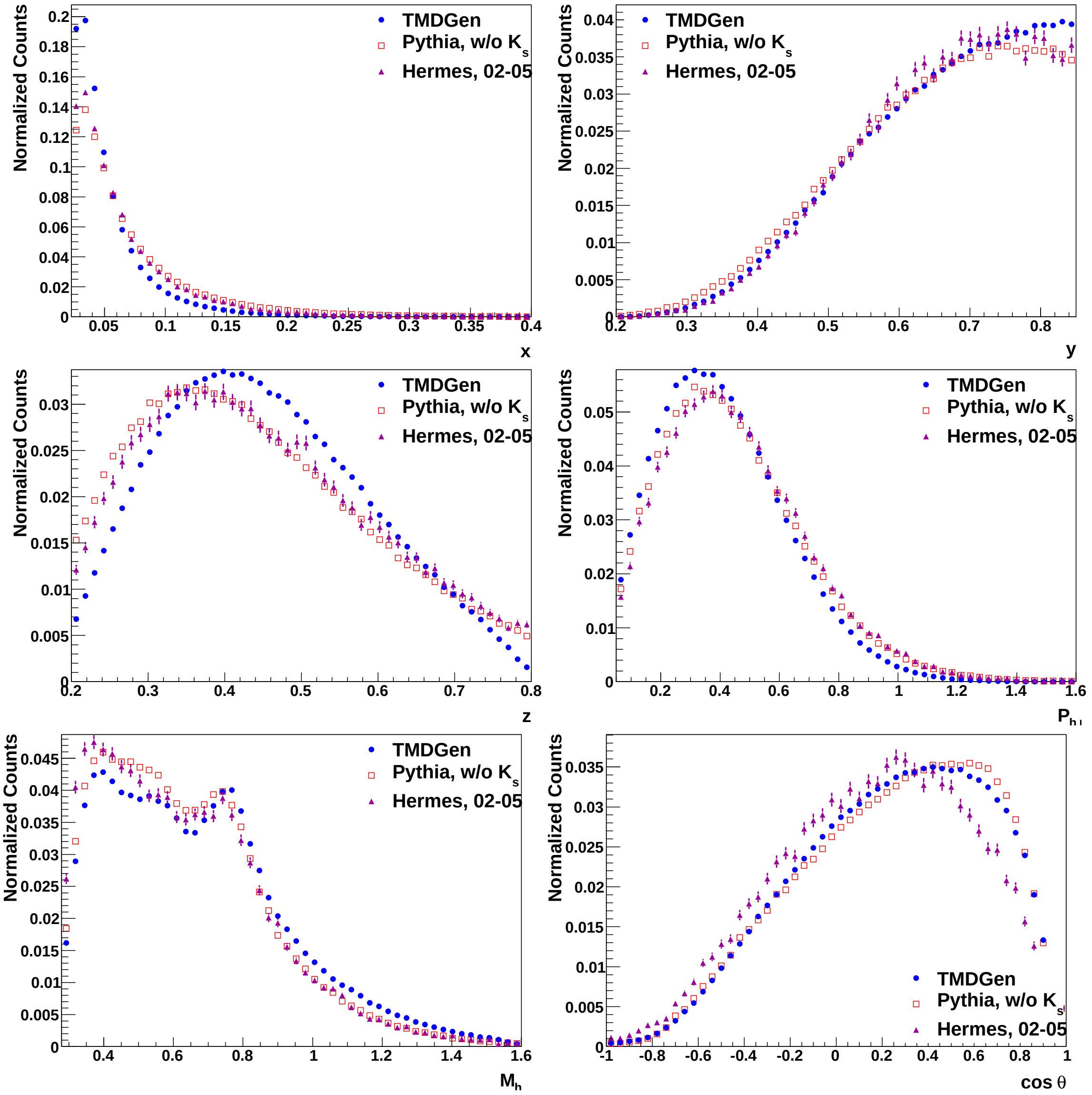


Figure 4.8: Comparison of kinematic variables for HERMES, `Pythia`, and `TMDGen` for $\pi^-\pi^0$ dihadron data within acceptance. Panels and markers are the same as in Figure 4.6.

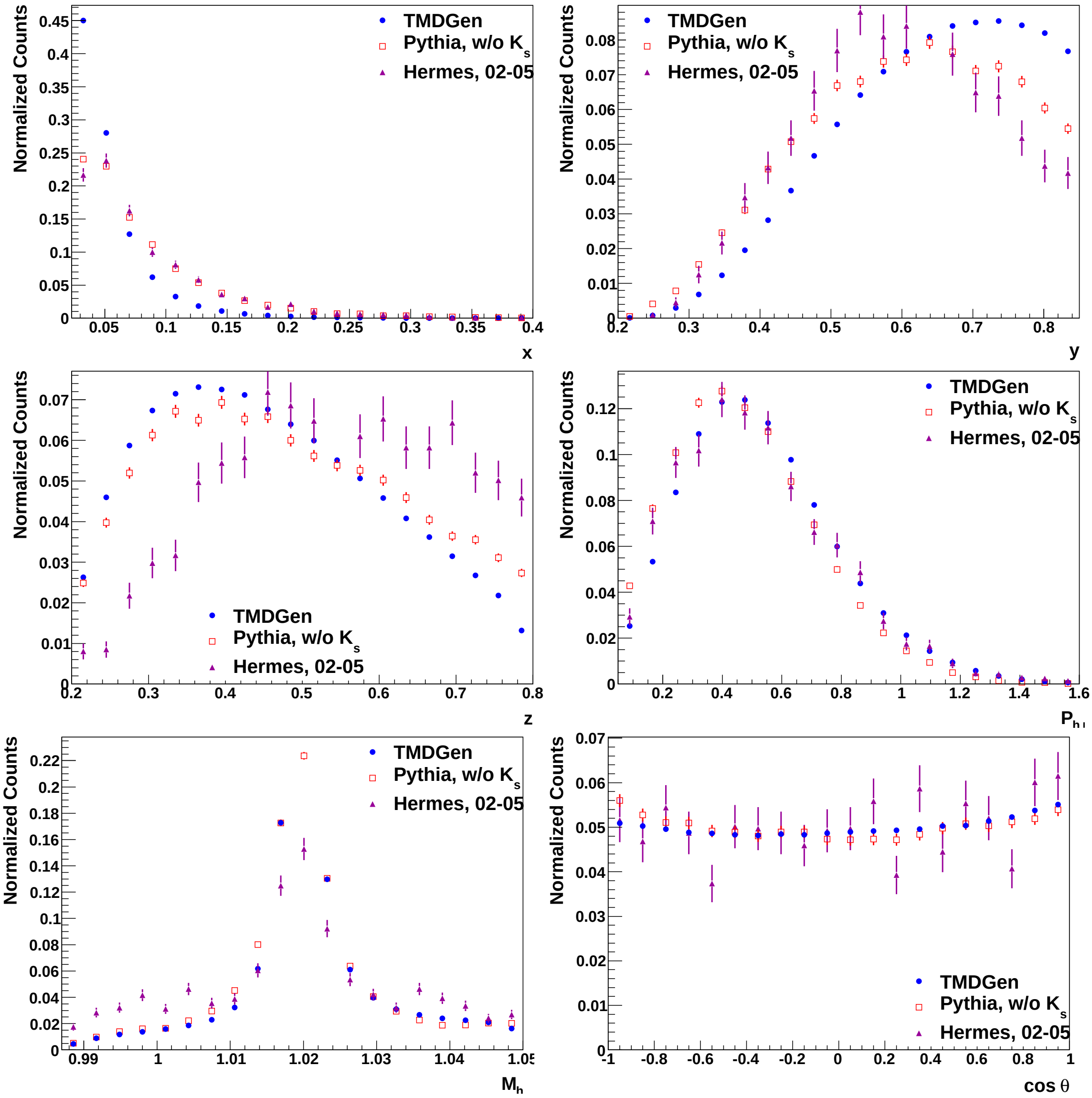


Figure 4.9: Comparison of kinematic variables for HERMES, `Pythia`, and `TMDGen` for $K^+K^-$ dihadron data within acceptance. Panels and markers are the same as in Figure 4.6.

## 4.4 Fitting Details

### 4.4.1 Fitting Functions

In cases where no acceptance correction is applied, a basic maximum likelihood estimate (MLE) is performed to determine the fit parameters. Specifically, the Hybrid S.J. routine [99, 100], as provided in the GNU General Scientific Library (GSL) [101] is used to find the roots of the first derivative of the MLE objective function. In cases where acceptance correction is applied, the ISE method of Section 3.1 is used.

In most cases, the fit function is chosen to include all of the unpolarized moments, including twist-3, as well as the Sivers and Collins transverse target moments. This equals 24 unpolarized moments (including the constant term) as well as 18 polarized moments, for a total of 42 moments. The fit function in this case can be specifically written as

$$\begin{aligned} f(\cos\vartheta, \phi_h, \phi_R, \phi_S) \quad = \quad & \sum_{\ell=0}^{2} \Bigg[ \sum_{m=0}^{\ell} a_1^{|\ell,m\rangle} P_{\ell,m} \cos(m\phi_h - m\phi_R) \\ & + \sum_{m=-\ell}^{\ell} \Bigg( a_2^{|\ell,m\rangle} P_{\ell,m} \cos((2-m)\phi_h + m\phi_R) \\ & \quad + a_3^{|\ell,m\rangle} P_{\ell,m} \cos((1-m)\phi_h + m\phi_R) \\ & \quad + b_1^{|\ell,m\rangle} P_{\ell,m} \sin((m+1)\phi_h - m\phi_R - \phi_S) \\ & \quad + b_2^{|\ell,m\rangle} P_{\ell,m} \sin((1-m)\phi_h + m\phi_R + \phi_S) \Bigg) \Bigg]. \end{aligned} \tag{4.4}$$

When applying the acceptance correction, one can consider including additional $\cos\vartheta$ dependence while including all the same azimuthal moments as in Equation 4.4. Among the Legendre polynomials up to $\ell = 2$, there are five linearly independent functions. Allowing each azimuthal moment of Equation 4.4 to occur with each of the possible Legendre functions results in 115 moments. Specifically, the fit function can be written as

$$\begin{aligned} f(\cos\vartheta, \phi_h, \phi_R, \phi_S) \quad = \quad & \sum_{m=0}^{2} \left( \boldsymbol{a}_1^{(m)} \cdot \boldsymbol{F}^{(m)} \right) \cos(m\phi_h - m\phi_R) \\ & + \sum_{m=-2}^{2} \Bigg( \left( \boldsymbol{a}_2^{(m)} \cdot \boldsymbol{F} \right) \cos((2-m)\phi_h + m\phi_R) \\ & \quad + \left( \boldsymbol{a}_3^{(m)} \cdot \boldsymbol{F} \right) \cos((1-m)\phi_h + m\phi_R) \\ & \quad + \left( \boldsymbol{b}_1^{(m)} \cdot \boldsymbol{F} \right) \sin((m+1)\phi_h - m\phi_R - \phi_S) \\ & \quad + \left( \boldsymbol{b}_2^{(m)} \cdot \boldsymbol{F} \right) \sin((1-m)\phi_h + m\phi_R + \phi_S) \Bigg), \end{aligned} \tag{4.5}$$

| **Year** | $\langle \boldsymbol{P_\perp} \rangle$ |
|---|---|
| 2002-2004 | $0.753 \pm 0.050$ |
| 2005 | $0.706 \pm 0.065$ |
| 2002-2005 | $0.723 \pm 0.059$ |

Table 4.6: Average target polarization per running period.

with $\boldsymbol{a}_i^{(m)} \cdot \boldsymbol{F}^{(m)}$ shorthand for

$$\boldsymbol{a}_i^{(m)} \cdot \boldsymbol{F}^{(m)} = \begin{cases} a_{i1}^{(m)} + a_{i2}^{(m)} \cos\vartheta + a_{i3}^{(m)} \frac{1}{2}\left(3\cos\vartheta^2 - 1\right) & \\ \quad + a_{i4}^{(m)} \sin\vartheta + a_{i5}^{(m)} \sin 2\vartheta & m = 0, 1 \\ a_{i1}^{(m)} + a_{i2}^{(m)} \cos\vartheta + a_{i3}^{(m)} \sin^2\vartheta & \\ \quad + a_{i4}^{(m)} \sin\vartheta + a_{i5}^{(m)} \sin 2\vartheta & m = 2 \end{cases}, \tag{4.6}$$

and equivalently for $\boldsymbol{b}_i^{(m)}$. Note the only difference between $\boldsymbol{F}^{(m)}$ for $m = 0, 1$ and $m = 2$ is in the third component, where $P_{2,0}$ is replaced with $P_{2,2}$, since among $P_{0,0}$, $P_{2,2}$ and $P_{2,0}$, there are only two linearly independent functions. The pair $P_{0,0}$, $P_{2,0}$ is chosen in most cases, except when the azimuthal moment occurs in the cross section with a factor of $P_{2,2}$.

The target polarization is not considered accurate on an event level, but the official average over certain running periods is considered accurate. The average values $\langle P_\perp \rangle$, per larger data sample, as provided by the target group within HERMES, are used. The fits are performed without explicitly using the factor of $\langle P_\perp \rangle$. One must then divide the results by $\langle P_\perp \rangle$. For plots in this chapter, the results have not been divided by $\langle P_\perp \rangle$, although results have been divided by this factor for all figures in Chapters V and VI as well as Appendices B and C. A table of the average target polarization per running period is given in Table 4.6. Note, the target group has also published polarization factors per certain sets of runs. These values have not been used, as the yearly and multi-yearly averages of Table 4.6 have been used instead. The uncertainty due to the target polarization results in a 7.3% scale uncertainty, which is indicated on each of the final result plots in Chapter VI and Appendix C.

Note, that using $P_\perp$ instead of $S_\perp$ introduces a small mixing between the unpolarized lepton and longitudinally polarized lepton terms in the cross section. However, Monte Carlo studies show that the maximum deviation between $P_\perp$ and $S_\perp$ is about $S_\perp = 0.97 P_\perp$. Given the size of the moments extracted in this dissertation, and considering that the data sample is approximately beam balanced, this mixing effect is considered negligible.

### 4.4.2 Verification of Acceptance Correction

Given the complexity of the fit functions and acceptance correction, it is necessary to verify with what accuracy the various moments can be extracted, and in particular, to select which moments are sufficiently accurate to be included in the final results. Details concerning the precision of the correction method is further considered in Section 5.1. Note, this study is denoted "Challenge A," to contrast to the similar study in Section 5.1, which is denoted "Challenge B."

| **Data Set** | $\boldsymbol{\pi^+\pi^0}$ | $\boldsymbol{\pi^+\pi^-}$ | $\boldsymbol{\pi^-\pi^0}$ | $\boldsymbol{K^+K^-}$ |
|---|---:|---:|---:|---:|
| Proto | 156k | 386k | 158k | 400k |
| $4\pi$ | 972k | 972k | 971k | 494k |
| Monte Carlo | 347k | 1.2M | 737k | 799k |
| HERMES | 53k | 259k | 40k | 2k |

Table 4.7: Table of relevant statistics for Challenge A. The various data sets are described in the text.

A large sample of `TMDGen` data was generated with no angular dependence, both in $4\pi$ and with running through the full detector simulation and data processing chain. Angular dependence can then be introduced by weighting the events. This method, that of generating with no angular dependence, allows for the greatest flexibility with the data set. The weights are computed for each event by evaluating the cross section using `TMDGen`, using the experimentally accessible cross section variables and integrating over $\boldsymbol{p}_T$. Table 4.7 includes a comparison of the amount of statistics in the actual HERMES data with the amount used in this Study.

There data sets are relevant in this study. The first, denoted proto-data, includes the simulation of the detector as well as the data processing chain, and also includes weights to induce angular dependence. The proto-data takes the place of the actual HERMES data for this study. The second data set, denoted $4\pi$ data, representing the true distribution. No detector simulation is used, though weights are used to induce the same angular distribution as is induced in the proto-data. The third data set is the Monte Carlo data, that which is used for estimating the joint distribution in the acceptance correction method. All of the data, proto, $4\pi$ and Monte Carlo are in fact generated with the `TMDGen`Monte Carlo generator for Challenge A. Note, the proto and Monte Carlo data both include a full simulation of the detector and data processing chain, and are identical except for the additional weights present in for the proto-data. To avoid introducing other effects into this study, all data for this study is generated with a positron beam.

The model for the unpolarized moments are based on distribution and fragmentation functions. Specifically, GRV98 [77] is used for $f_1$ and the Torino group's parametrization is used for $h_1^\perp$ [80]. The fragmentation functions used are those computed in Section 2.4 of this dissertation. Only the twist-2 unpolarized moments are included, as the twist-3 unpolarized $|0,0\rangle$ term breaks positivity, due to the $p_T$ and $Q^2$ range of the data sample. This is related to the fact that the $Q^2$ values for the data set are not as large as those for which one would ideally hope.

Using the available distribution and fragmentation functions for the transverse target moments results in moments that are extremely small. For the purpose of the studies in this section, Section 4.4.2, and for the later study in Section 5.1, some of the systematic effects could be masked by using too small of moments. Thus, instead of using distribution and fragmentation functions for the polarized moments, a simple model for the moments is chosen. The Sivers and Collins moments are set proportional to the $f_1 D_1^{|0,0\rangle}$ term, with the

| | **Sivers** | | **Collins** | |
|---|---|---|---|---|
| **Parameter** | $\boldsymbol{\ell \neq 2}$ | $\boldsymbol{\ell = 2}$ | $\boldsymbol{\ell \neq 2}$ | $\boldsymbol{\ell = 2}$ |
| $b_x$ | -0.75 | -0.75 | 0.75 | 0.75 |
| $\alpha$ | -1.0 | 2.0 | 2.0 | -1.0 |
| $\beta$ | -1.0 | 2.0 | 2.0 | -1.0 |
| $b_M$, $\pi\pi$ | 0.1 | 0.1 | 1.0 | 1.0 |
| $c_M$, $\pi\pi$ | -0.3 | -0.3 | 0.0 | 0.0 |
| $b_M$, $K^+K^-$ | 10.0 | -10.0 | 1.0 | 0.0 |
| $c_M$, $K^+K^-$ | -10.0 | 10.0 | 1.0 | 0.0 |

Table 4.8: Parameters used for the polarized moments, excluding the parameter $a$. Note, different values of $b_M$ and $c_M$ are used for pion-pair and kaon-pair dihadrons, due to the different mass ranges involved.

| **Partial Wave** | $\boldsymbol{a}$ |
|---|---|
| $\|0, 0\rangle$ | -0.05 |
| $\|0, 1\rangle$ | 0.05 |
| $\|2, 0\rangle$ | -0.06 |
| $\|2, \pm 1\rangle$ | -0.04 |
| $\|2, \pm 2\rangle$ | -0.02 |

Table 4.9: Values of the parameter $a$ used for the polarized moments. The same values are used for both Sivers and Collins moments.

proportionality constant having the form

$$a\big(1 - b_x \ln(x)\big) z^{\alpha}(1 - z)^{\beta}\big(1 + b_M M_h + c_M M_h^2\big). \tag{4.7}$$

The specific values for the parameters $a$, $b_x$, $\alpha$, $\beta$, $b_M$, $c_M$ are given in Tables 4.8 and 4.9. Note, the variation in the $a$ parameter over different partial waves was chosen so that when the data is fit, the fit parameters all have about the same magnitude. An alternate form to that in Equation 4.7 was also considered, which included $P_{h\perp}$ dependence. However, this other form was difficult to tune, as it often broke positivity at high values of $P_{h\perp}$. Thus, only the form in Equation 4.7 is used.

Given $4\pi$ data and data reconstructed in acceptance, both with the given model induced, one can then compare the results of fitting each sample. The acceptance correction method of Section 3.1 is applied to the reconstructed data, and the results are denoted the "acceptance plus correction" results. For data in $4\pi$, an MLE fit is preformed, and the results from this fit are denoted the $4\pi$ results. Both fitting functions, Equation 4.4 and 4.5, were initially considered for fitting the data in acceptance, though the 42 parameter fit, Equation 4.4, was found to produce better results. Also, to solve $B\boldsymbol{\alpha} = \boldsymbol{b}$, both a $QR$ decomposition using

Householder transformations and the eigenvalue method of Section 3.1.5 were attempted, in both cases using the relevant algorithms from the GSL package [101]. In both cases, the eigenvalue method was used to compute $B^{-1}$, used in propagating uncertainties. The $QR$ decomposition yielded more accurate and precise results.

Plots showing the full results for all fit parameters are not shown, as this would require over 160 pages of plots. For the kinematic dependencies, the Collins $|2,2\rangle$ moment is chosen as representative. Plots comparing the $4\pi$ results with the acceptance plus correction results, 2D binning, are shown in Figures 4.10 through 4.13, for each dihadron type. In all but a few kinematic bins, the acceptance correction method is able to accurately reconstruct the moment. Note, this challenge was also repeated with the kinematic region reduced to avoid some of the problem areas, specifically the lowest $y$ and $z$ bins were removed and the maximum value of $M_h$ for pion-pairs was set to 1.2 $\mathrm{GeV}/c^2$. Reducing the kinematic domain did not improve the comparison but did significantly reduced the amount of statistics. For this reason, the full kinematic domain as defined in Section 4.2 is used.

Figure 4.14 includes the comparison of the $4\pi$ and acceptance plus correction results for pion-pair dihadrons, considering binning just in invariance mass. In this case, the Collins $|1,1\rangle$ moment, the same moment as that analyzed in Ref. [27], and the Collins $|2,\pm2\rangle$ moments are chosen as representative. Figure 4.15 similarly compares $4\pi$ and acceptance plus correction results for the 1D mass binning, except that the Sivers $|0,0\rangle$ and $|2,\pm2\rangle$ moments are chosen as representative. Again, the selected moments are generally reconstructed well. Note the particularly small scale on Figure 4.14.

To determine the overall effectiveness in reconstructing each of the moments from data within acceptance, a $\chi^2/ndf$ statistic is computed for each moment. The $\chi^2/ndf$ is computed by comparing the $4\pi$ versus acceptance plus correction results, varying over all bins within a specific binning choice. For example, the $\chi^2/ndf$ for the $M_h$-$x$ binning is computed by determining the individual $\chi^2$ values between the two results per each of the sixteen bins, and then taking the average value. The procedure is then repeated for each choice of 2D binning, as well as for the one choice of 1D binning considered. Results for each dihadron type are given in Appendix A, due to the four tables being somewhat lengthy. Specifically, the $\chi^2/ndf$ statistics are given in Tables A.1 through A.4.

Several trends in the data can be noticed among the $\chi^2/ndf$ statistic for all three of the pion-pair dihadrons. In particular, the $|2,\pm2\rangle$ moments tend to be reconstructed most accurately. Note, the the $|2,\pm2\rangle$ moments occur with a $\sin^2\cos\vartheta$ modulation, making them most sensitive to data with small values of $|\cos\vartheta|$. This is exactly the region where the acceptance is best. Although the $\cos\vartheta$ distributions for the $\pi^\pm\pi^0$ dihadrons in Figures 4.6 and 4.8 would suggest the optimal acceptance is near $\cos\vartheta = 0.5$, one must also consider the effect of symmetry. Since the $\sin^2\cos\vartheta$ is symmetric under the exchange of the sign of $\cos\vartheta$, one must consider the acceptance of the symmetrized $\cos\vartheta$ distribution, which does have the highest acceptance near small values of $|\cos\vartheta|$. For $\pi^+\pi^-$ dihadrons, the acceptance in $\cos\vartheta$ is symmetric, and Figure 4.7 shows that the acceptance is greatest near small values of $|\cos\vartheta|$.

The $|1,\pm1\rangle$ moments also tend to be reconstructed well for the pion-pair dihadrons, except in the 1D binning for a few cases. As the acceptance can depend on all variables, one generally expects 1D binning to yield worse results than higher dimensional binning.

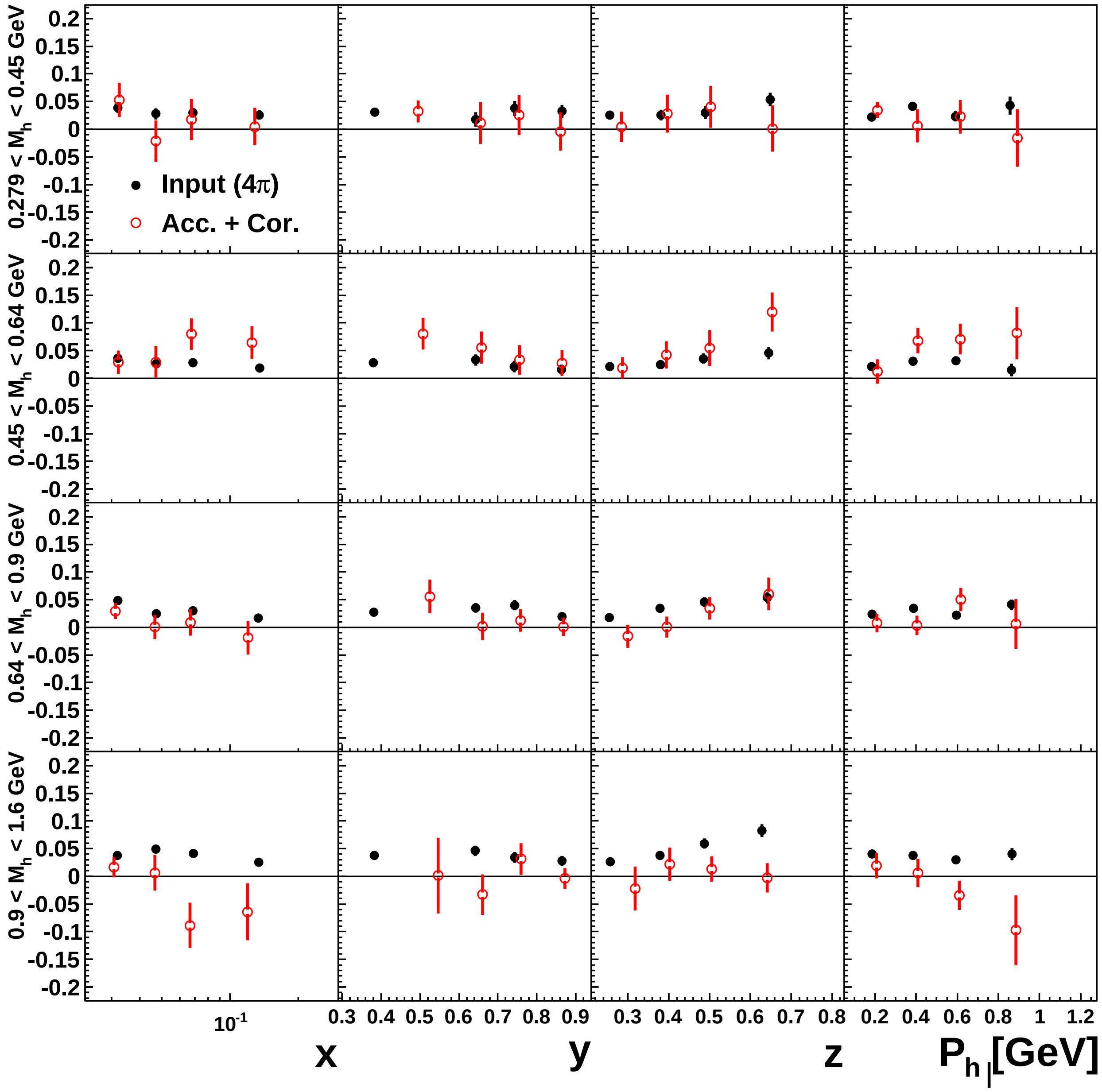


Figure 4.10: Comparison of $4\pi$ versus acceptance plus correction moments using `TMDGen` data for $\pi^+\pi^0$ dihadrons, with kinematic dependencies, for the Collins $|2,2\rangle$ moment. The results for the MLE fit to data in $4\pi$ are shown with black filled circles, while the moments from the acceptance correction fit to data within acceptance are shown with red open circles. The upper row of panels is for the lowest $M_h$ bin, with each row of panels being for the next higher $M_h$ bin. From left to right, the columns of panels represent binning with respect $x$, $y$, $z$, and $P_{h\perp}$. Note, each row represents an independent data sample, while each column is a different binning of the same sample.

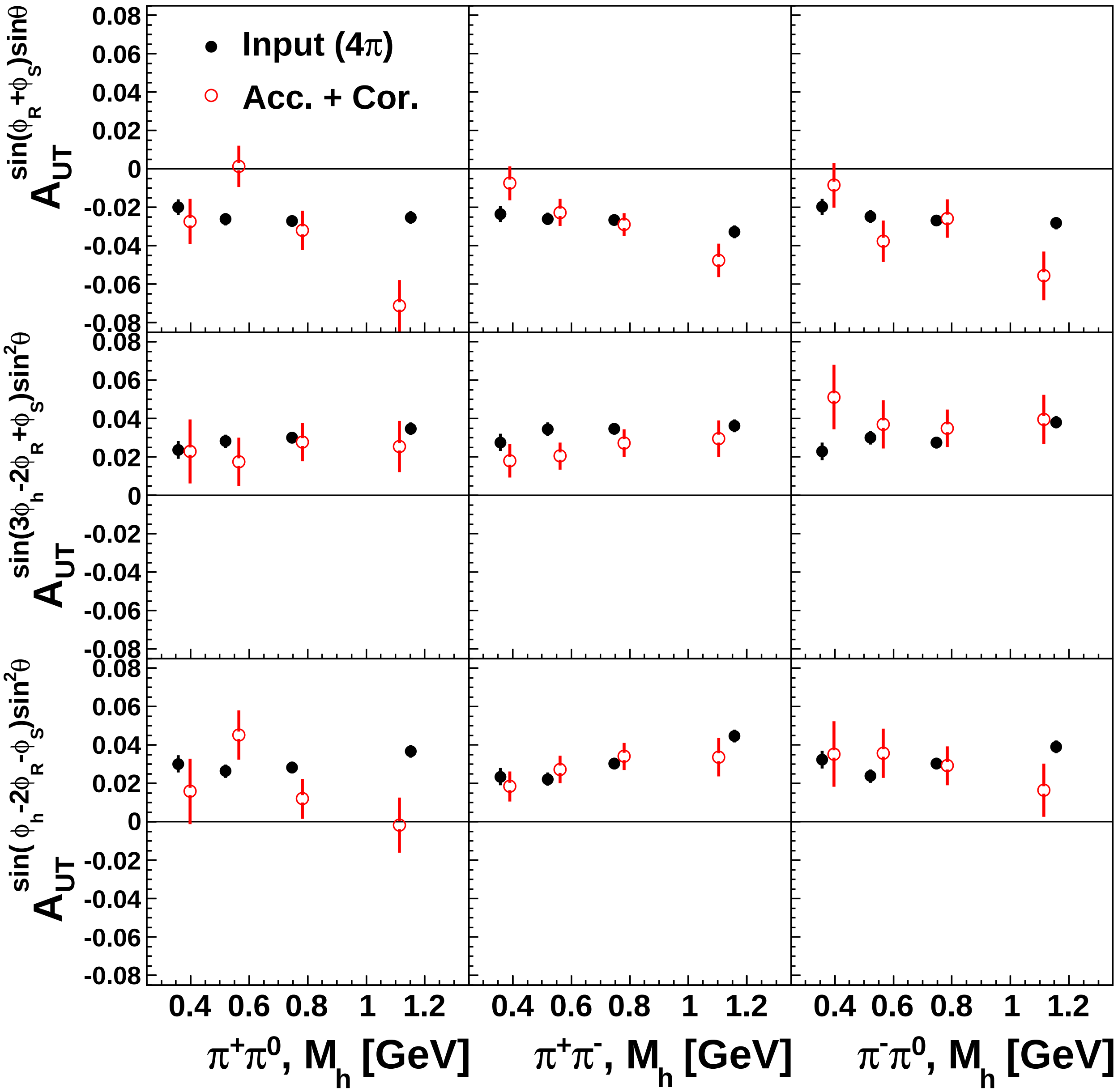


Figure 4.11: Comparison of $4\pi$ versus acceptance plus correction moments using `TMDGen` data for $\pi^+\pi^-$ dihadrons, with kinematic dependencies, for the Collins $|2,2\rangle$ moment. Panels and markers are the same as in Figure 4.10.

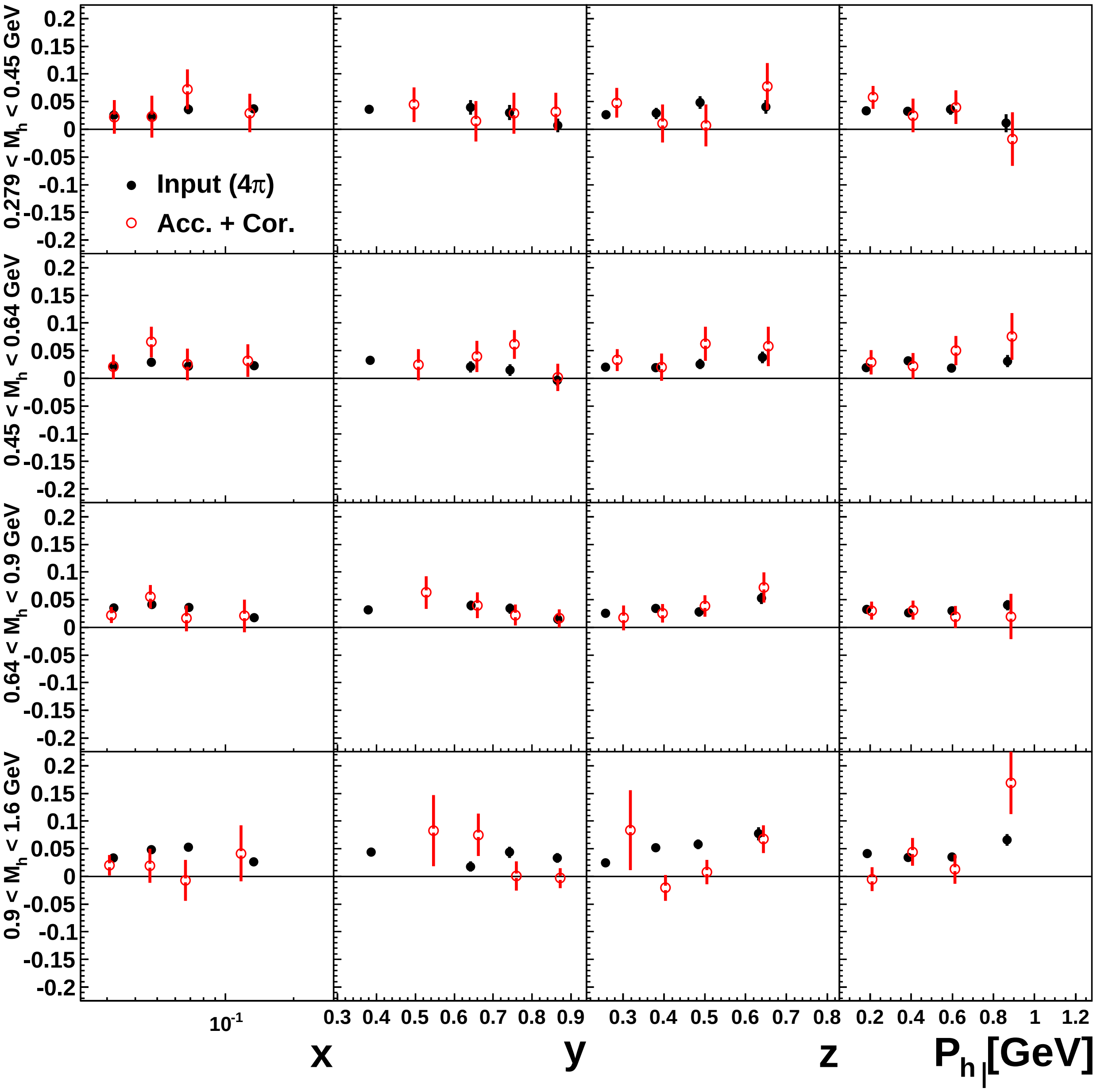


Figure 4.12: Comparison of $4\pi$ versus acceptance plus correction moments using `TMDGen` data for $\pi^-\pi^0$ dihadrons, with kinematic dependencies, for the Collins $|2,2\rangle$ moment. Panels and markers are the same as in Figure 4.10.

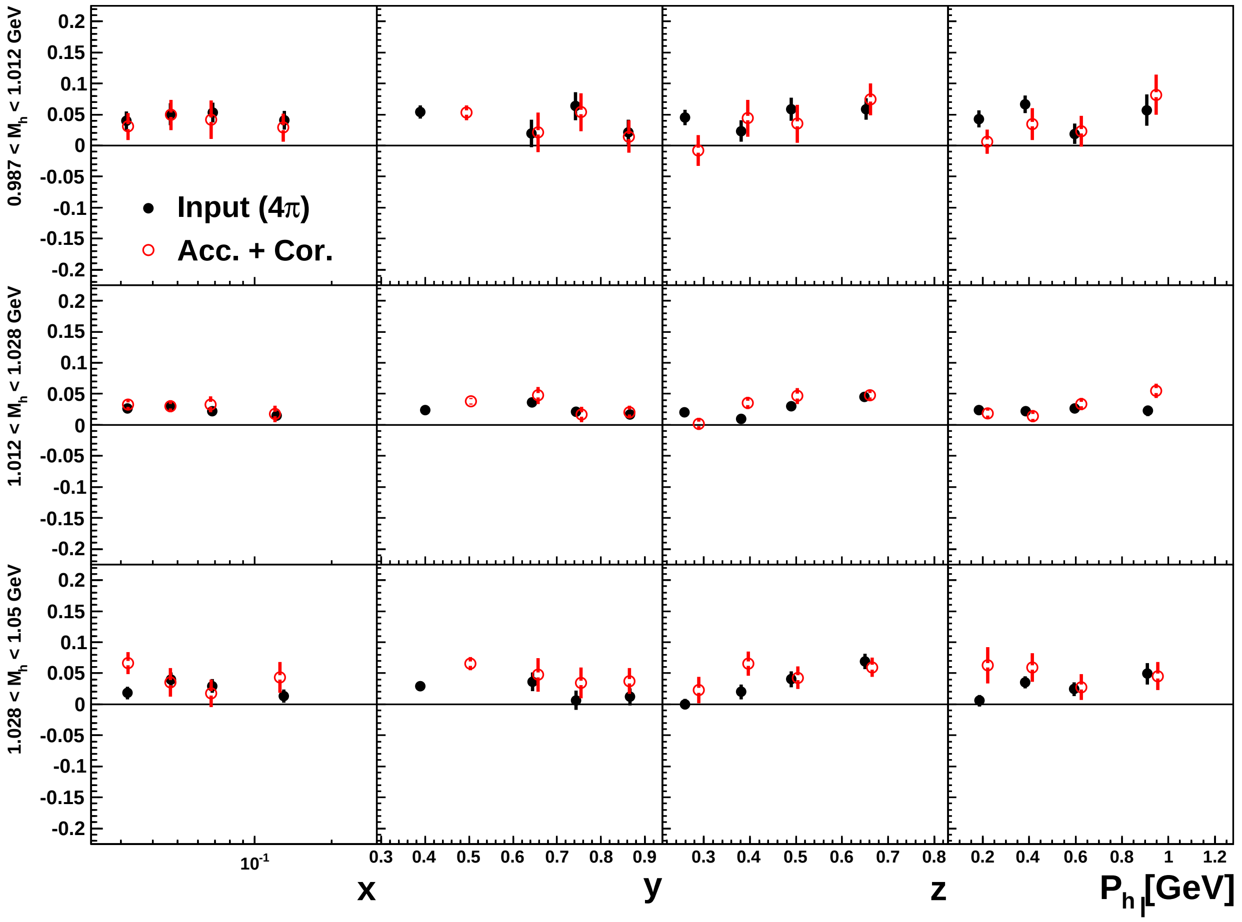


Figure 4.13: Comparison of $4\pi$ versus acceptance plus correction moments using `TMDGen` data for $K^+K^-$ dihadrons, with kinematic dependencies, for the Collins $|2,2\rangle$ moment. Panels and markers are the same as in Figure 4.10.

Note, these $|1,\pm 1\rangle$ moments occur with a $\sin\vartheta$ modulation, and the general shape of $\sin\vartheta$ is quite similar to that for $\sin^2\vartheta$. Thus the $|1,\pm 1\rangle$ moments are reconstructed well for the same reasons as that for the $|2,2\rangle$ moments.

The $|\ell,0\rangle$ moments are most sensitive to data with large values of $|\cos\vartheta|$. This is precisely where the acceptance most severely reduces the data for pion-pair dihadrons, with the acceptance approaching zero as $|\cos\vartheta|$ goes to one. Thus, the $|\ell,0\rangle$ moments are all reconstructed poorly for pion-pair dihadrons. Additionally, the $\pi^\pm\pi^0$ dihadrons have a strong asymmetry in their acceptance, with much worse acceptance for negative values of $\cos\vartheta$. This is a reflection that the acceptance for low momentum neutral pions is worse than the acceptance for low momentum charged pions, or specifically that the minimum photon cluster energy requirement removes more low momentum neutral pions than the minimum RICH momentum requirement removes charged pions. For $\pi^\pm\pi^0$ dihadrons, one would expect the strong asymmetry in the acceptance versus $\cos\vartheta$ would make reconstruction of the odd moments $|1,0\rangle$, $|2,\pm 1\rangle$ difficult. While the $|1,0\rangle$ moments, which have no azimuthal dependence, are

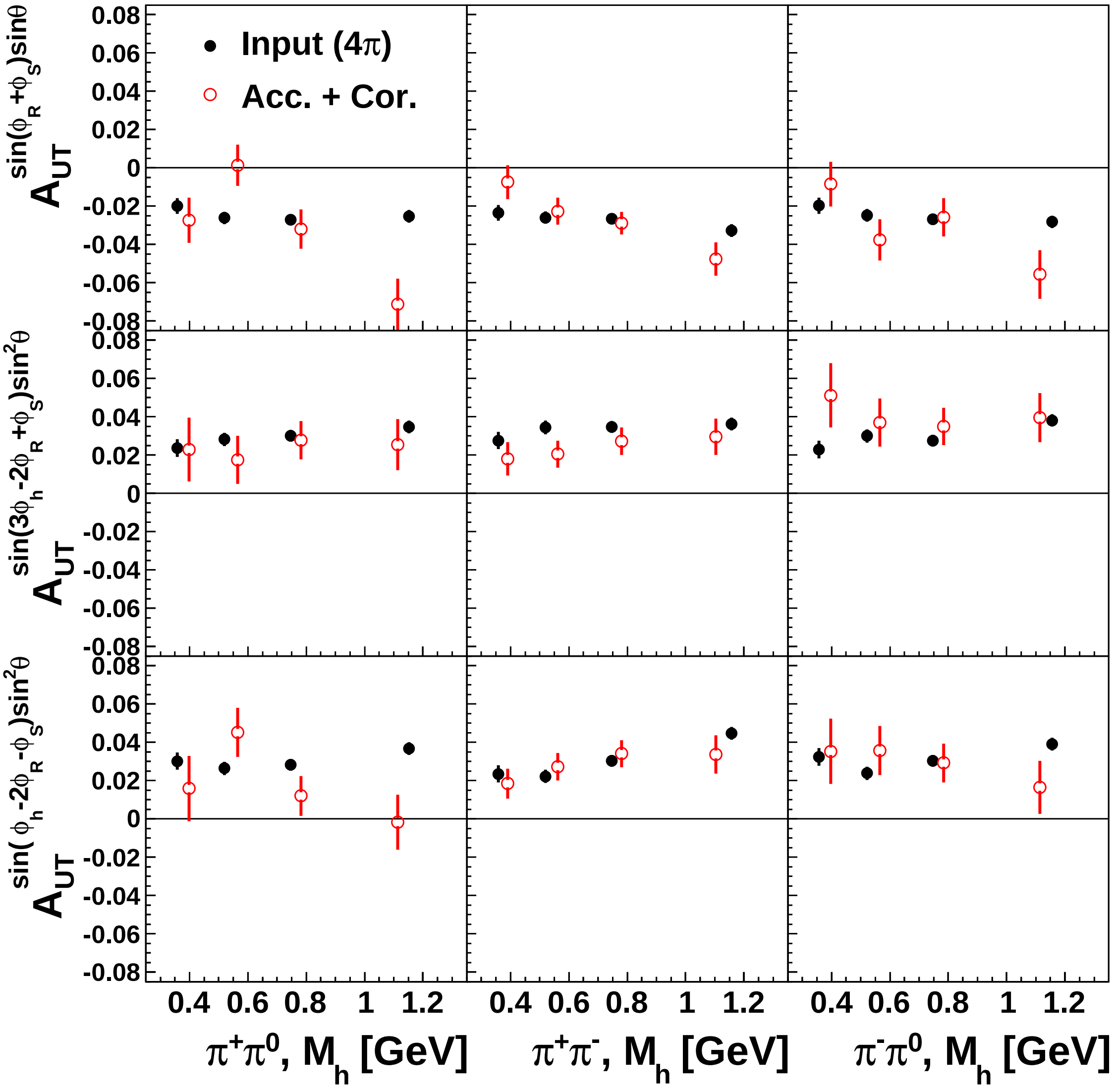


Figure 4.14: Comparison of $4\pi$ versus corrected moments using `TMDGen` data for pion-pair dihadrons, without kinematic dependencies. Each column represents a different dihadron, specifically, from left to right, the $\pi^+\pi^0$, $\pi^+\pi^-$, and $\pi^-\pi^0$ dihadrons. The rows of panels, represent different moments, and are, from top to bottom, the $|1,1\rangle$, $|2,-2\rangle$ and $|2,2\rangle$ Collins moments. The distribution in $4\pi$ is shown with black filled circles, while the moments in acceptance plus correction are shown with red open circles.

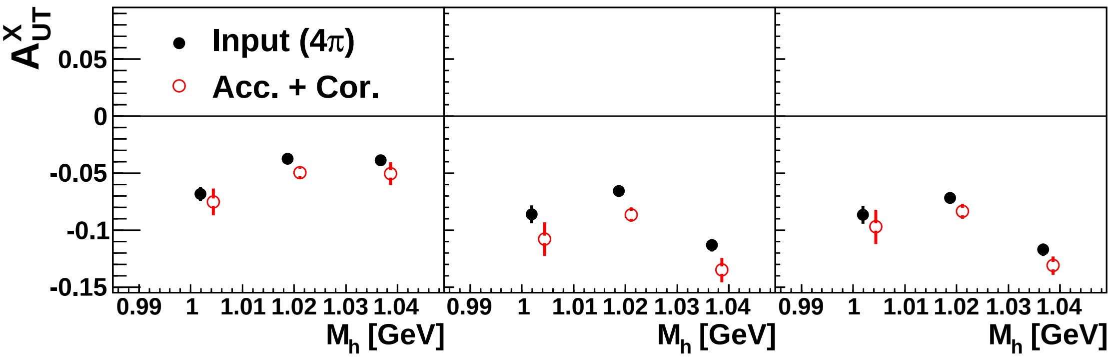


Figure 4.15: Comparison of $4\pi$ versus corrected moments using `TMDGen` data for $K^+K^-$ dihadrons, without kinematic dependencies. The panels, from left to right, are the $|1,-1\rangle$, $|2,2\rangle$ and $|2,-2\rangle$ Sivers moments. The distribution in $4\pi$ is shown with black filled circles, while the moments in acceptance plus correction are shown with red open circles.

reconstructed quite poorly, the other odd moments are reconstructed fairly well—most likely due to their additional azimuthal dependence.

For the $K^+K^-$ dihadrons, there are no clear trends in the $\chi^2/ndf$ statistics, Table A.4. This is mainly due to the significantly different acceptance in $\cos\vartheta$, as can be seen from the $\cos\vartheta$ distribution in Figure 4.9. The kaon mass being larger than the pion mass causes the RICH momentum cuts have a much less significant effect. For this reason, the $\chi^2/ndf$. statistics for all moments are much better for $K^+K^-$ dihadrons than for any of the pion-pair dihadrons. Note, though, the worst moments are still some of the $|\ell,0\rangle$ moments, as with pion-pair dihadrons.

For the main motivation of this dissertation, the moments of most interest are the $|1,1\rangle$ and $|2,\pm2\rangle$ Collins moments for pion-pair dihadrons and the $|0,0\rangle$ and $|2,\pm2\rangle$ Sivers moments for $K^+K^-$ dihadrons. Fortunately, the data is sensitive to these moments, and all of these moments can be reconstructed quite well using the given acceptance correction method. For the pion-pair dihadrons, all of the $|2,\pm2\rangle$ and $|1,\pm1\rangle$ Sivers and Collins moments can be reconstructed well, and thus these will be the moments for which results are given in Chapter VI. For the $K^+K^-$ dihadrons, at this point, all moments will be considered, though this will be reconsidered once the full systematic uncertainties are determined in Chapter V.

## 4.5 Processes and Backgrounds

### 4.5.1 Non-resonant Photon Pairs

As noted in the discussion regarding Figure 4.2 in Section 4.2.1, there exists a sizable amount of non-resonant photon pairs within the $\pi^0$ mass window. As no models exist for any asymmetry present in the non-resonant sample, the only possible method is to fit data from the sidebands and interpolate an effective asymmetry in the peak region. However, the statistics are somewhat limited, as shown in Table 4.10.

|  | $\pi^+\gamma\gamma$ |  |  | $\pi^-\gamma\gamma$ |  |  |
|---|---|---|---|---|---|---|
| **Year** | **L** | **P** | **H** | **L** | **P** | **H** |
| 2002 | 516 | 4214 | 592 | 384 | 3250 | 445 |
| 2003 | 299 | 2330 | 351 | 253 | 1865 | 250 |
| 2004 | 1660 | 14338 | 1903 | 1287 | 11209 | 1450 |
| 2005 | 4009 | 32274 | 4634 | 2884 | 24256 | 3314 |
| Total | 6484 | 53156 | 7480 | 4808 | 40580 | 5459 |

Table 4.10: Table of statistics within different two-photon invariant mass regions. L, P, and H indicate low-sideband, peak, and high-sideband regions, defined as $0.075 < M_{\gamma\gamma} < 0.095$ GeV/$c^2$, $0.115 < M_{\gamma\gamma} < 0.155$ GeV/$c^2$, and $0.175 < M_{\gamma\gamma} < 0.195$ GeV/$c^2$, respectively.

Data from the $M_{\gamma\gamma}$ sideband regions, defined in the caption to Table 4.10, for both $\pi^+\gamma\gamma$ and $\pi^-\gamma\gamma$ events, were fit using the acceptance correction method of Section 3.1 for both. The same Monte Carlo data is used for correcting data in the sidebands as for the data in the peak, since the `TMDGen` generator does not include any non-resonant photon pairs, and thus includes no sideband regions.

As a representative moment, the results for the $|2,2\rangle$ Collins moments versus $M_h$, for $\pi^\pm\gamma\gamma$ events, are shown in Figure 4.16. The kinematic dependencies for the $|2,2\rangle$ moment, just for $\pi^+\gamma\gamma$ events, is shown in Figure 4.17. The effect of the correction is generally small: the central values are only slightly shifted, and the uncertainties are slightly increased. This correction is assumed effective, and no further systematic uncertainty is assigned for non-resonant photon pairs.

### 4.5.2 Charge Symmetric Background

A number of processes produce electron-positron pairs, with one of the produced leptons mimicking the scattered electron from the SIDIS process. These processes are denoted the charge symmetric background, since they produce electrons and positrons symmetrically. The dominant processes are quasi-real photo-production of $\pi^0$ mesons in the target gas and the interaction of particles with the collimator. A number of HERMES analyses found it necessary to correct for this background, the biggest effect being seen in the low $x$ and low $Q^2$ data used in the $F_2$ structure function analysis [102]. The correction has been also used in the exclusive $\rho$ SDME and transverse target moment analysis [103].

In this analysis, the charge symmetric background fraction is quite low, as shown in Table 4.11. As the acceptance-correction fitting method is linear, it is sufficient to separately fit the like and unlike sign data, and then make the appropriate linear combination post-fitting. This is in contrast to using MLE, where fitting with negative weights is not numerically identical to weighting the results of two separate fits.

Data from the 2002-2005 running period, for the three pion-pair dihadrons and with the apparent scattered lepton having opposite charge as the beam (the unlike-sign data),

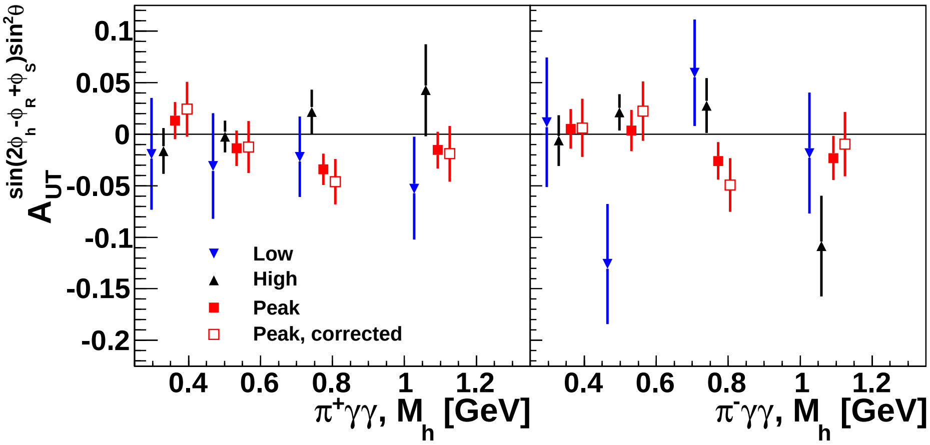


Figure 4.16: Comparison of the $|2,2\rangle$ Collins moment versus $M_h$ within three $M_{\gamma\gamma}$ regions. Left panel is for $\pi^+\gamma\gamma$ events; right panel is for $\pi^-\gamma\gamma$ events. Results from the lower sideband region are given with blue inverted triangles, from the higher sideband region with black upright triangles, and with red filled squares for data from the $\pi^0$ peak region. The corrected peak values are given with red, open squares.

have been fit using the acceptance correction method of Section 3.1. The statistics for unlike-charge lepton and $K^+K^-$ dihadrons are too low to allow fitting. Due to the low background fraction (1.2%) and low statistics, this background is considered negligible for $K^+K^-$ dihadrons. The results of fitting the unlike-sign data are mostly consistent with zero and have very high uncertainties, especially for the kinematic dependencies. As a representative moment, the $|2,2\rangle$ Collins moment for all three pion are given in Figure 4.18, for 1D binning in $M_h$. As a representative figure for the kinematic dependencies, Figure 4.19 shows the results with kinematic dependence, again for the $|2,2\rangle$ moment, but only for $\pi^+\pi^0$ dihadron data. Based on the results of the fits for the pion-pair dihadrons, it can be seen that there is not enough data to accurately estimate the background signal. Furthermore, due to the small background fraction, this background most likely has negligible effect on the final results. Thus, no correction and no systematic uncertainty is assigned for this background.

| **Year** | $\pi^+\pi^0$ | | $\pi^+\pi^-$ | | $\pi^-\pi^0$ | | $K^+K^-$ | |
|---|---|---|---|---|---|---|---|---|
| 2002 | 222 | 5.0% | 827 | 3.8% | 145 | 4.3% | 2 | 1.1% |
| 2003 | 120 | 4.9% | 477 | 3.9% | 74 | 3.8% | 1 | 1.0% |
| 2004 | 762 | 5.0% | 2849 | 3.9% | 487 | 4.2% | 4 | 0.7% |
| 2005 | 1608 | 4.7% | 7346 | 4.5% | 1667 | 6.4% | 18 | 1.4% |
| Total | 2712 | 4.9% | 11499 | 4.3% | 2373 | 5.5% | 25 | 1.2% |

Table 4.11: Data statistics for lepton charge unlike beam lepton charge, separated by year and dihadron type. For each year and dihadron, the total amount of statistics is given, in addition to the background fraction those statistics represent.

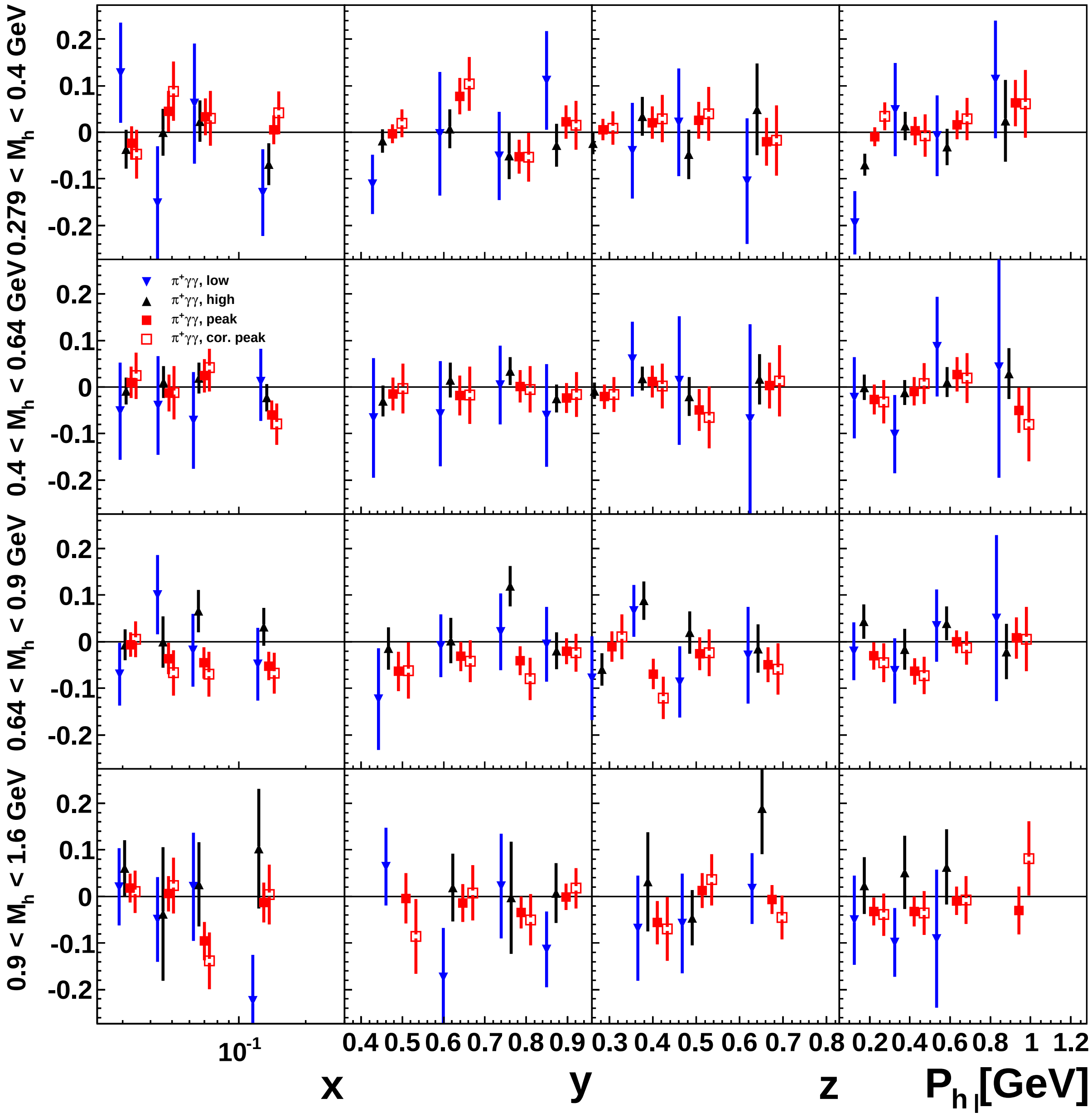


Figure 4.17: Comparison of the kinematic dependence of the $|2,2\rangle$ Collins moment within three $M_{\gamma\gamma}$ regions. Data is for $\pi^-\gamma\gamma$ events. The markers are the same as in Figure 4.16, and the panels are the same as for Figure 4.10.

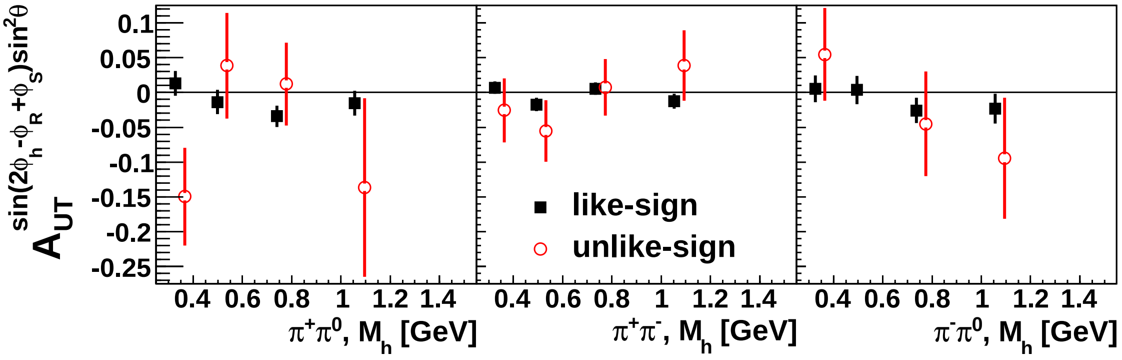


Figure 4.18: Comparison of the $|2, 2\rangle$ Collins moment versus $M_h$ for like versus unlike lepton sign. The panels, from left to right, are for $\pi^+\pi^0$, $\pi^+\pi^-$, and $\pi^-\pi^0$ dihadron events. Black squares are for lepton with like sign as the lepton beam, red open circles for unlike sign.

### 4.5.3 Exclusive Background

For interpretation of the results in terms of distribution and fragmentation functions, it is essential that the dihadrons are produced semi-inclusively and not exclusively. Specifically, this means that it is required that other particles are present in the final state. When other particles are produced, they take some of the available energy, thus reducing the fractional energy $z$ of the produced hadron system.

One can remove much of the exclusive vector meson process from the data sample by restricting the range of $z$ and/or the missing mass $M_X$. The missing mass is computed by considering the missing momenta and energy, i.e. the difference between known input and the considered outgoing particles, and computing the effective mass. In a previous publication [27], both $M_X$ and $z$ cuts were employed, though the following study shows that both cuts are not needed.

Figure 4.20 shows the missing mass $M_X$ versus $z$ distribution for each of the four dihadrons considered. The data sample is `Pythia` with a positron beam. No significant difference is seen when using `Pythia` data with an electron beam. Lines are drawn on the plots at several choices of cuts, corresponding to those cuts listed in Table 4.12. The actual background fractions are also included in Table 4.12. The extra $M_X$ cut is shown not to improve the background fraction, and thus a cut of $z < 0.8$ is sufficient.

Unlike the charge symmetric background, there exists no model independent correction method. There also exists no results for exclusive vector meson production in the correct kinematic domain. Thus, one cannot correct for the background using negative weights nor is it possible to estimate a systematic uncertainty. However, the background fraction is quite small. Most exclusive $\rho^0$ SDMEs for a transversely polarized target and unpolarized beam tend to be quite small, with the largest (in absolute magnitude) being 0.11 and all but one SDME less than than 0.07 in magnitude [103]. Thus one would expect the exclusive $\rho^0$ mesons can contribute no more than 0.004 to the polarized moments studied in this dissertation. The charged $\rho$ SDMEs have not been so well studied, but one can expect the

| **Cut** | $\pi^+\pi^0$ | | $\pi^+\pi^-$ | | $\pi^-\pi^0$ | | $K^+K^-$ | |
|---|---|---|---|---|---|---|---|---|
| $z < 0.7$ | 1903 | 0.5% | 14204 | 3.0% | 1316 | 1.0% | 3363 | 2.5% |
| $z < 0.8$ | 2128 | 0.5% | 15391 | 3.6% | 1421 | 1.4% | 3852 | 3.3% |
| $z < 0.9$ | 2247 | 0.6% | 16315 | 4.7% | 1502 | 1.7% | 4200 | 4.5% |
| $M_X < 1.5$ | 2253 | 0.7% | 16470 | 5.4% | 1510 | 2.0% | 4222 | 4.7% |
| $M_X < 2.0$ | 2163 | 0.6% | 15620 | 4.2% | 1443 | 1.5% | 3896 | 3.8% |
| $M_X < 1.5$ and $z < 0.8$ | 2128 | 0.5% | 15386 | 3.6% | 1420 | 1.4% | 3850 | 3.4% |
| $M_X < 2.0$ and $z < 0.8$ | 2109 | 0.5% | 15229 | 3.7% | 1410 | 1.3% | 3778 | 3.4% |

Table 4.12: Statistics and background fraction for exclusive production verses select $M_X$ and $z$ cuts and dihadron type, from `Pythia` positron data.

| **Dihadron** | **Est. VM Stats.** | **Bkg. Frac.** |
|---|---|---|
| $\pi^+\pi^0$ | 5497 | 68.5% |
| $\pi^+\pi^-$ | 10846 | 85.4% |
| $\pi^-\pi^0$ | 2774 | 77.9% |
| $K^+K^-$ | 700 | 31.9% |

Table 4.13: Table of the vector meson fraction within the resonant mass peak. Data is computed from the Breit-Wigner plus background fit of HERMES data, as shown in Figure 4.5.

relative size of the SDMEs to be comparable. Thus, the exclusive background fraction can be seen to have negligible effect on the final results, and no systematic uncertainty is needed.

### 4.5.4 Vector Meson Fraction

As discussed in Section 2.1.1, the SIDIS dihadron process includes many subprocesses. When one considers pure SIDIS vector meson production, these other subprocesses are considered background. In Figure 4.5, the invariant mass distributions are fit with a Breit-Wigner distribution plus a linear background, to estimate the amount of vector mesons. The estimated statistics and background fractions are given in Table 4.13.

Note though, in many cases it is not meaningful to discuss separating the vector meson signal from the other processes, especially when considering the interference between these processes. Additionally, the fraction of vector mesons and other dihadron subprocesses to each partial wave is unknown and cannot be estimated within the current theory. Thus, although the angular integrated background fractions is given in Table 4.13, it is unknown how much the other subprocesses contribute to the $\ell = 2$ sector nor how this contribution varies with $M_h$.

## 4.6 Comparison with Published Results

As a consistency check, one can perform a MLE fit of the $\pi^+\pi^-$ data to compare with Ref. [27]. The MLE fit is preformed with no corrections, using the older data productions, and using the cuts, binning, and extraction methods of Ref. [27]. Note, the fitting method of Ref. [27] involved a binned $\chi^2$-fit and anti-symmetrized in $\cos\vartheta$. The fitting method used therein also fixed the value of the unpolarized $P_{2,0}(\cos\vartheta)$ to some number $b$, and then scans over various possible values of $b$. For comparison, a MLE fit is used with the functional form of

$$f(\cos\vartheta, \phi_R, \phi_S) = 1 + a_1 \sin\vartheta \sin(\phi_R + \phi_S) + a_2 \sin 2\vartheta \sin(\phi_R + \phi_S), \tag{4.8}$$

where the unpolarized terms have been set to zero.

The idea for this comparison is a consistency check between the previous and current analyses, rather than a full repetition of the previous analysis. Note though, since so much of the analysis procedure is different between this analysis and the publication, this can serve only as a rough comparison.

The results of the MLE fit compared with the results of Ref. [27] are plotted in Figure 4.21. Only statistical uncertainty is shown. The central values are in good agreement. However, the uncertainties for the MLE fit tend to be smaller than those for the published results, by a factor of about 75%. Within the coarseness of this comparison, the agreement is quite good.

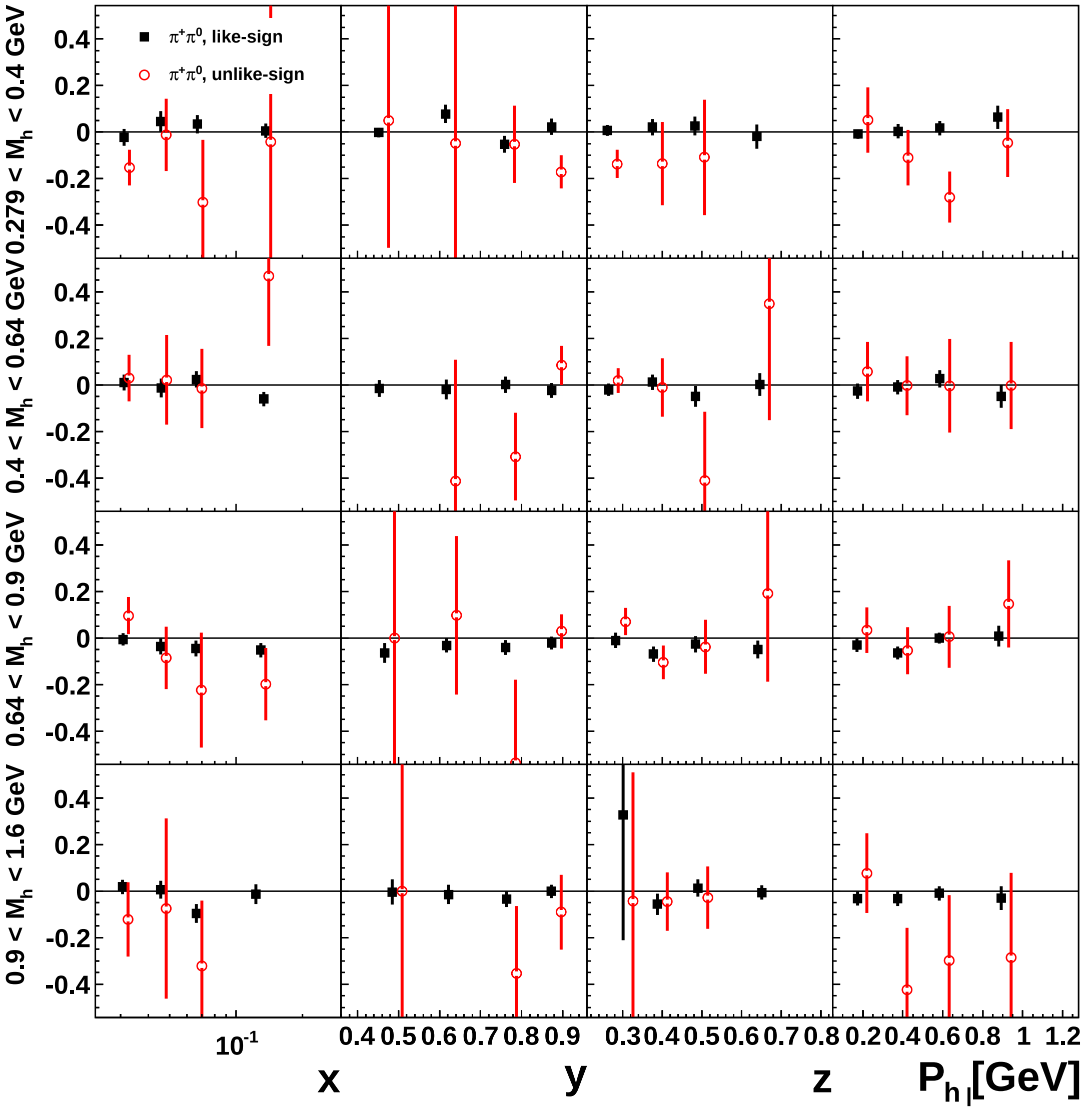


Figure 4.19: Comparison of the $\pi^+\pi^0$ dihadron $|2,2\rangle$ Collins moment, with kinematic dependencies, for like versus unlike lepton sign. Markers are as in 4.18, and the panels are the same as for Figure 4.10. Some data points, which brake positivity, are not shown.

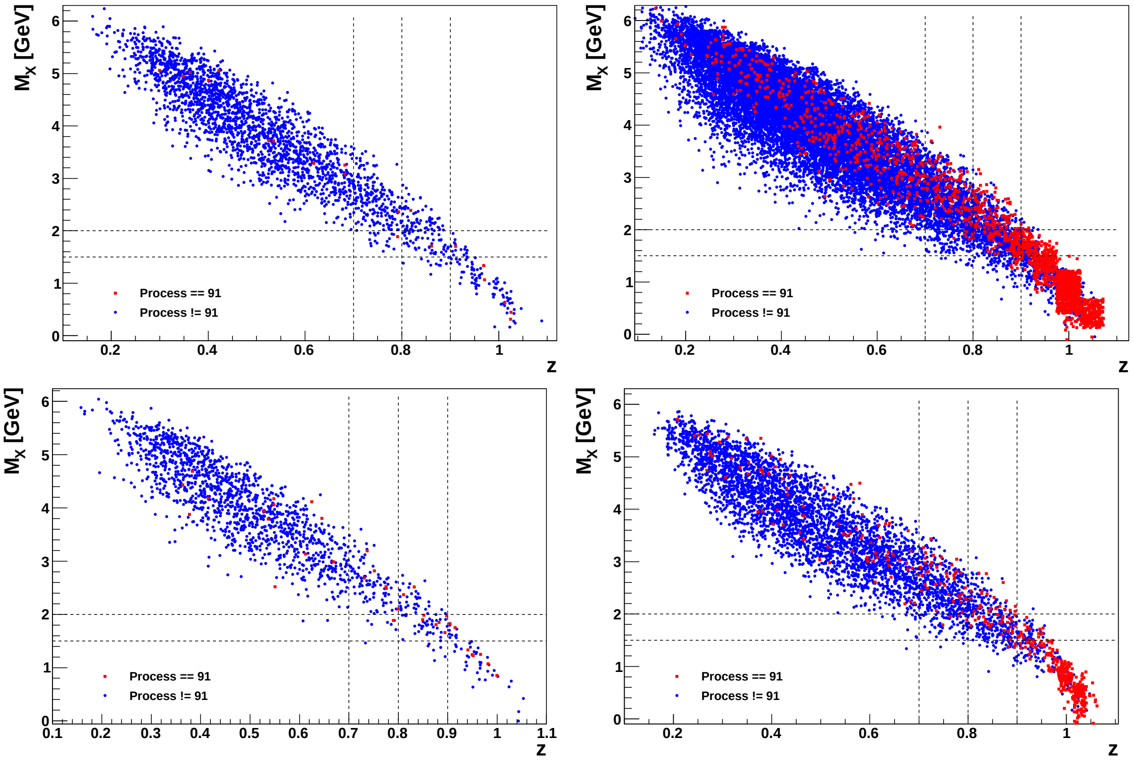


Figure 4.20: Distribution of missing mass $M_X$ versus $z$ for `Pythia` positron data. Panels represent data from each of the four dihadrons. Specifically, clock-wise from upper-left, $\pi^+\pi^0$, $\pi^+\pi^-$, $K^+K^-$, and $\pi^-\pi^0$. Lines are drawn at each of $z = 0.7$, $z = 0.8$, $z = 0.9$, $M_X = 1.5$ GeV/$c^2$, and $M_X = 2.0$ GeV/$c^2$.

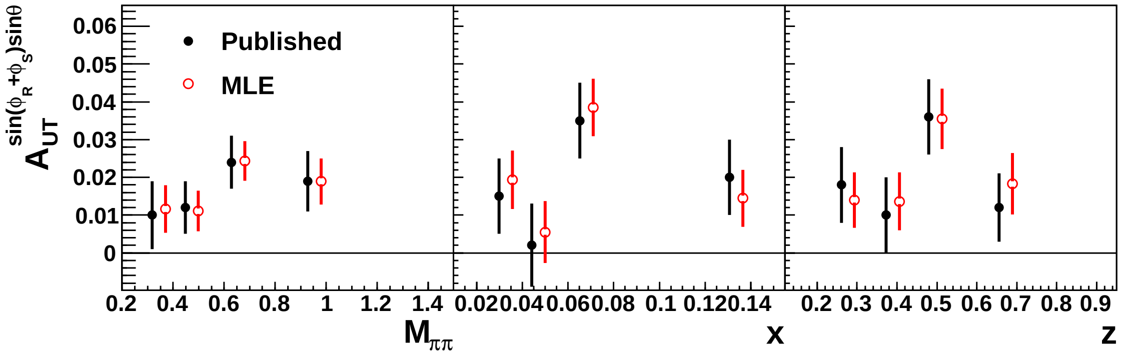


Figure 4.21: Comparison of uncorrected moments with previously published results. Data is for $\pi^+\pi^-$ pairs. For full description of the fitting method, see the text of this document. Published results are from Ref. [27].

# CHAPTER V

# Systematic Uncertainty Studies

This chapter details the estimation of the various sources of systematic uncertainty, with the exception of those already determined (in Chapter IV) to be negligible. Three sources of systematic effects are considered in this chapter: 1) residual smearing and acceptance effects after the acceptance correction, 2) variations between experimental setup between the 2002-2004 (positron beam) and 2005 (electron beam), and 3) uncertainty in the hadron identification procedure. The full results, comparing all sources of systematic uncertainty, are given in Appendix B, including all moments, dihadrons, and binning options for which final result are presented in Chapter VI and Appendix C.

Note, in all studies in this chapter using HERMES data, the target polarization has been removed from the fit parameters and the acceptance correction method of fitting from Section 3.1 has been applied. However, the non-resonant photon pair background is not taken into account in any figures or in the estimate of the uncertainty, as the effect upon the uncertainty estimation would be negligible.

## 5.1 Smearing and Acceptance

In Section 4.4.2, a model was introduced into `TMDGen` data, both in $4\pi$ and within acceptance. The comparison of the acceptance corrected and $4\pi$ moments is used to determine which moments can be reliably extracted within acceptance, with the study being denoted "Challenge A." In this section, the same model used in Section 4.4.2 is induced into `Pythia` data using the same procedure, i.e. by using `TMDGen` to evaluate the cross section and introducing weights. The reconstructed `Pythia` sample, i.e. the sample within acceptance, also has radiative effects included, via the subroutine `RadGen` [98]. Thus, this comparison shows which moments can be reliably extracted with both smearing and acceptance effects modifying the data sample, and is denoted "Challenge B." Table 5.1 compares the amount of statistics for the HERMES data set versus the data sets used in this study. Note, the reconstructed `Pythia` data includes data from both beam charges, while the `TMDGen` data used for the acceptance correction (for Challenge B) is only for positron beam. No significant difference has been shown by changing the beam charge for the Monte Carlo data used in the acceptance correction.

The moments versus invariant mass are shown in Figures 5.1 and 5.2, for pion-pair and $K^+K^-$ dihadrons, respectively. The comparison is generally quite good, with the worst

| **Data Set** | $\pi^+\pi^0$ | $\pi^+\pi^-$ | $\pi^-\pi^0$ | $K^+K^-$ |
|---|---|---|---|---|
| Reconstructed `Pythia` (proto) | 358k | 2.7M | 294k | 26k |
| $4\pi$ `Pythia` | 2.3M | 2.4M | 1.8M | 330k |
| `TMDGen` for Acc. Cor. (Monte Carlo) | 503k | 1.6M | 895k | 1.2M |
| HERMES | 53k | 259k | 40k | 2k |

Table 5.1: Table of relevant statistics for Challenge B. The various data sets are described in the text.

moments, among those plotted, being the $|1,1\rangle$ Collins moment for $\pi^\pm\pi^0$ dihadrons and the $|1,-1\rangle$ Sivers moment for $K^+K^-$ dihadrons. In this section, Section 5.1, the kinematic dependencies are not shown for sake of brevity. The systematic uncertainty estimated by this study, however, is shown versus the 2D binning in Appendix B.

To obtain an overall idea of how well the moments compare between the $4\pi$ results and acceptance plus correction results, the $\chi^2/ndf$ statistic per moment and set of binning is again computed, as it was in Section 4.4.2. The $\chi^2/ndf$ statistics results, for each dihadron type, are given in Appendix A, Tables A.5 through A.8. For pion-pair dihadrons, the general trends observed in the study of Section 4.4.2 ("Challenge A") are again observed in this study. The moments of most interest, the $|2,\pm2\rangle$ and $|1,1\rangle$ Collins moments, are still reconstructed fairly well.

It is worth noting that, for most moments, the 1D binning results tend to have $\chi^2/ndf$ statistics similar to the 2D binning results. This is an indication that the kinematic dependence of the smearing and acceptance effects is not overly strong. In other analysis where the kinematic dependence of the smearing and acceptance is significant, such as in Ref. [65], even 2D binning is insufficient and one must bin in all kinematic variables. The $|1,1\rangle$ Collins moments for $\pi^\pm\pi^0$ dihadrons, however, do have $\chi^2/ndf$ statistics that are significantly worse for the 1D binning than the 2D binning. The kinematic dependence of these moments (not shown) contain small systematic shifts, similar to the results for 1D binning shown in Figure 5.1. The difference in $\chi^2/ndf$ between binning options for the $\pi^\pm\pi^0$ dihadron $|1,1\rangle$ Collins moments is more a reflection of the higher uncertainty in the 2D binning, rather than a reflection of a difference in accuracy.

For $K^+K^-$ dihadrons, the $|1,0\rangle$ moments are shown to be quite poor, and the $|2,0\rangle$ moments are also much higher than the other moments. Since the statistics for the $K^+K^-$ study were much smaller, a lower threshold for the $\chi^2/ndf$ statistic should be chosen than for pion-pair dihadrons. This is due to the fact that higher statistical uncertainty will mask systematic differences, and the $\chi^2/ndf$ values will appear smaller due to the larger uncertainty on the moments being compared. Therefore, the results for all but the $|1,0\rangle$ and $|2,0\rangle$ partial waves will be considered in the final results, though it remains to be seen which moments will have small enough uncertainty to be meaningful.

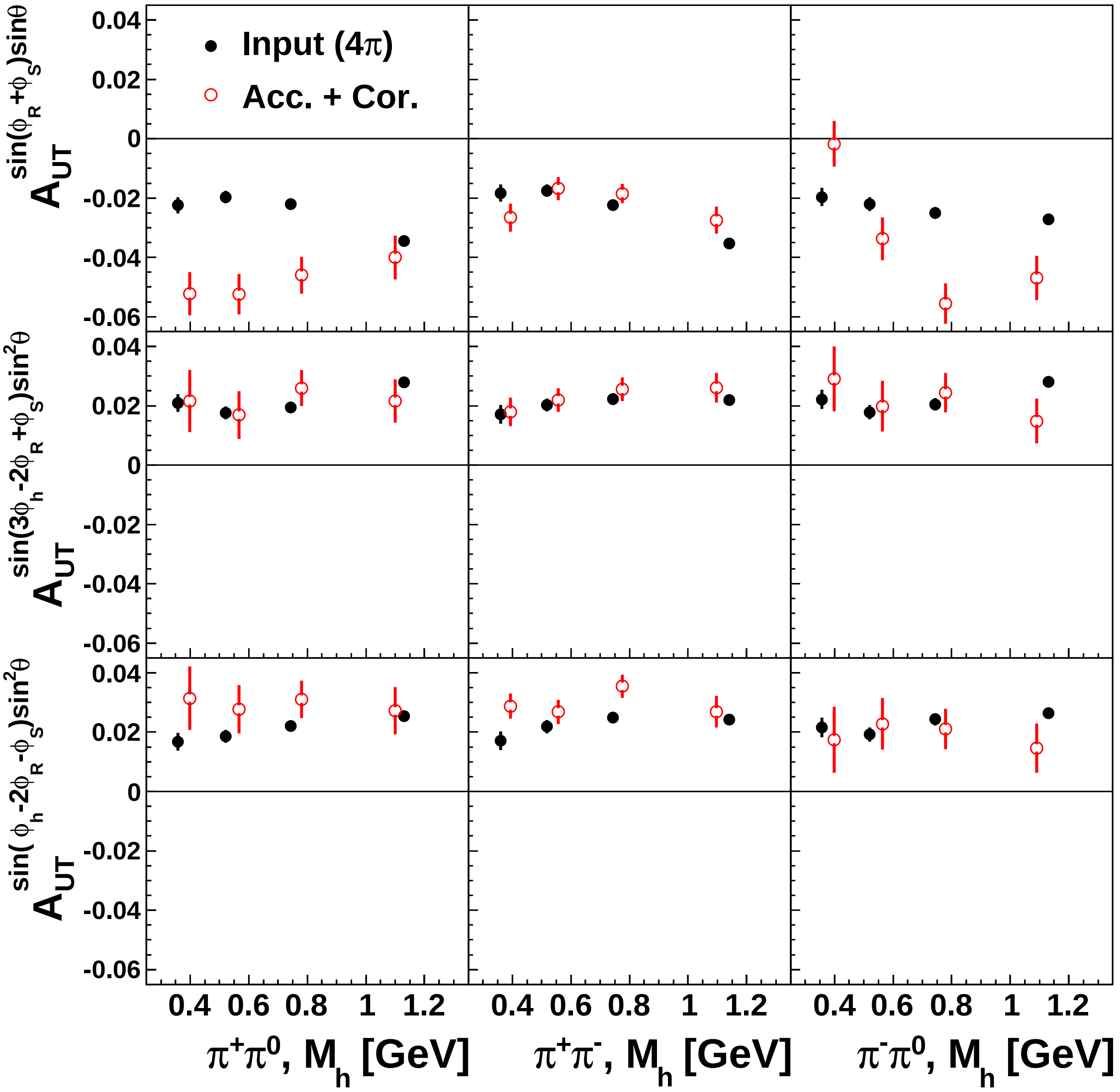


Figure 5.1: Comparison of $4\pi$ versus corrected moments using `Pythia` data for pion-pair dihadrons, without kinematic dependencies. Each column represents a different dihadron type, specifically, from left to right, $\pi^+\pi^0$, $\pi^+\pi^-$, and $\pi^-\pi^0$ dihadrons. The rows of panels represent different moments, and are, from top to bottom, the $|1,1\rangle$, $|2,-2\rangle$ and $|2,2\rangle$ Collins moments. The distribution in $4\pi$ is shown with black filled circles, while the moments in acceptance plus correction are shown with red open circles.

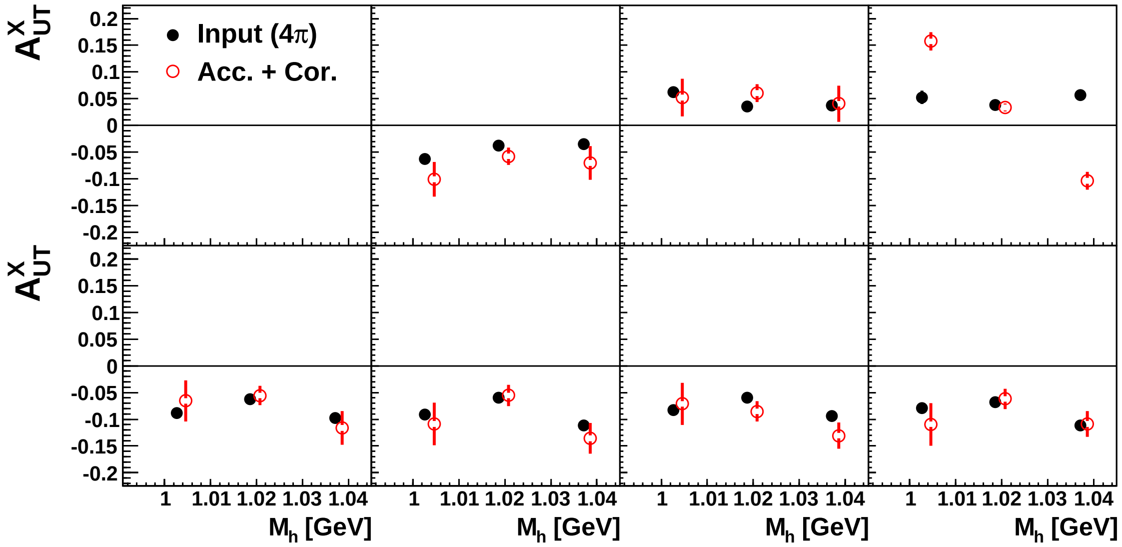


Figure 5.2: Comparison of $4\pi$ versus corrected moments using `Pythia` data for $K^+K^-$ dihadrons, without kinematic dependencies. The upper left panel is intentionally left blank. The other panels on the top row, from left to right, are the $|0,0\rangle$, $|1,1\rangle$ and $|1,-1\rangle$ Sivers moments. The panels on the second row are, from left to right, are the $|2,2\rangle$, $|2,1\rangle$, $|2,-1\rangle$ and $|2,2\rangle$ Sivers moments. As with Figure 5.1, the distribution in $4\pi$ is shown with black filled circles, while the moments in acceptance plus correction are shown with red open circles.

## 5.2 Year dependence

The HERMES data sample, as well as the `TMDGen` Monte Carlo data used in the acceptance correction, includes data from both beam charges, in roughly equal amounts. For instance, roughly 60% of the HERMES dihadron data was collected with an electron beam, and the remainder with a positron beam. Among the HERMES collaboration, uncertainties due to the differences between the two running periods (positron, 2002-2004, and electron, 2005) are historically called year dependent systematic uncertainties. However, "beam-charge" dependent rather than "year dependent" might have been a better choice, as the data samples in 2002 and 2003 are too small to consider individually and are instead included with the 2004 data. Although the SIDIS process is invariant with respect to beam charge, many systematic effects are not, most notably the beam position, offset, and slope, as well as the curvature of the beam through the target region.

Two systematic issues must be addressed. First, one must determine if there is any systematic effect due to fitting the data samples simultaneously rather than fitting separately and then combining the results. Determining whether a systematic effect is present, and the size of the resulting systematic uncertainty, shall be denoted "Study 1." In order to determine an effect, the source of such an effect would have to be included in simulations of the detector for the two running conditions.

The second study, denoted "Study 2," is then to determine whether there is a residual difference between the electron and positron sub-samples which is not accounted for in the detector simulation or data processing chain. Study 2 is accomplished by checking the consistency between the results for fitting electron and positron data separately, each corrected with the `TMDGen` data of the respective beam charge.

### 5.2.1 Year Dependent Study 1

In this study, the results for a combined fit of the electron and positron HERMES sub-samples, using combined electron and positron `TMDGen` data for correcting the acceptance, is compared with fitting and correcting the sub-samples separately and forming the appropriate linear combination. In all cases, the acceptance correction fit of Section 3.1 is used for fitting and correcting the data. Tables of $\chi^2/ndf$ statistics are computed, similar to those in Challenge A and B, and are listed in Appendix A.

In general, the results for both methods are very consistent, except for certain moments already shown to be unstable in previous studies. These unstable moments tend to be the $|\ell, 0\rangle$ moments for all dihadrons and the $|2, -1\rangle$ Sivers moments for $\pi^\pm\pi^0$ dihadrons. Based on these results, it appears that all known systematic differences between the running periods (included in the detector simulation) have either been corrected in the data processing chain or are negligible. Thus, one can use an electron-positron combined fit, and no systematic uncertainty is assigned thus far.

### 5.2.2 Year Dependent Study 2

In this study, the results for the positron beam (2002-2004) and electron beam (2005) data samples, are checked for consistency. A comparison of the results per $M_h$ bin, for both lepton charge samples as well as the combined sample, is given in Figures 5.3 and 5.4. For pion-pair dihadrons, there are only a few bins where there are significant differences between the sub-samples. For $K^+K^-$ dihadrons, there appear to be some systematic trends, though with only three bins, it is difficult to say whether the inconsistency is due to a larger trend or localized to certain individual bins.

Again, $\chi^2/ndf$ statistics are computed for each binning set and are provided in Appendix A. For Study 2, the $\chi^2/ndf$ values are slightly larger than one would hope. While some of the moments of interest have $\chi^2/ndf$ values near unity, several moments also have $\chi^2/ndf$ values in the range of 2.0 to 4.0, implying about a variation comparable to about 1.5 standard deviations of a one dimensional Gaussian variable. This is close enough to be purely statistical fluctuations, though this could also be indicative of some systematic effect. Note, the possible systematic effect would have to be one not previously observed at HERMES, as a large number of other HERMES analyses have also investigated possible systematic differences between these running periods, and known effects for these years have been incorporated in the the simulation and data processing chain. All other observables considered in this analysis, such as the 1D projects of the distribution of relevant variables, positions of mass peaks, etc., show no significant difference between the two running periods. Note also, the overall statistic provided in Tables A.13 through A.16 is quite good, though the most extreme moments are not considered in the overall. As no other indication exists of a yet unknown systematic difference between the positron and electron data samples, except

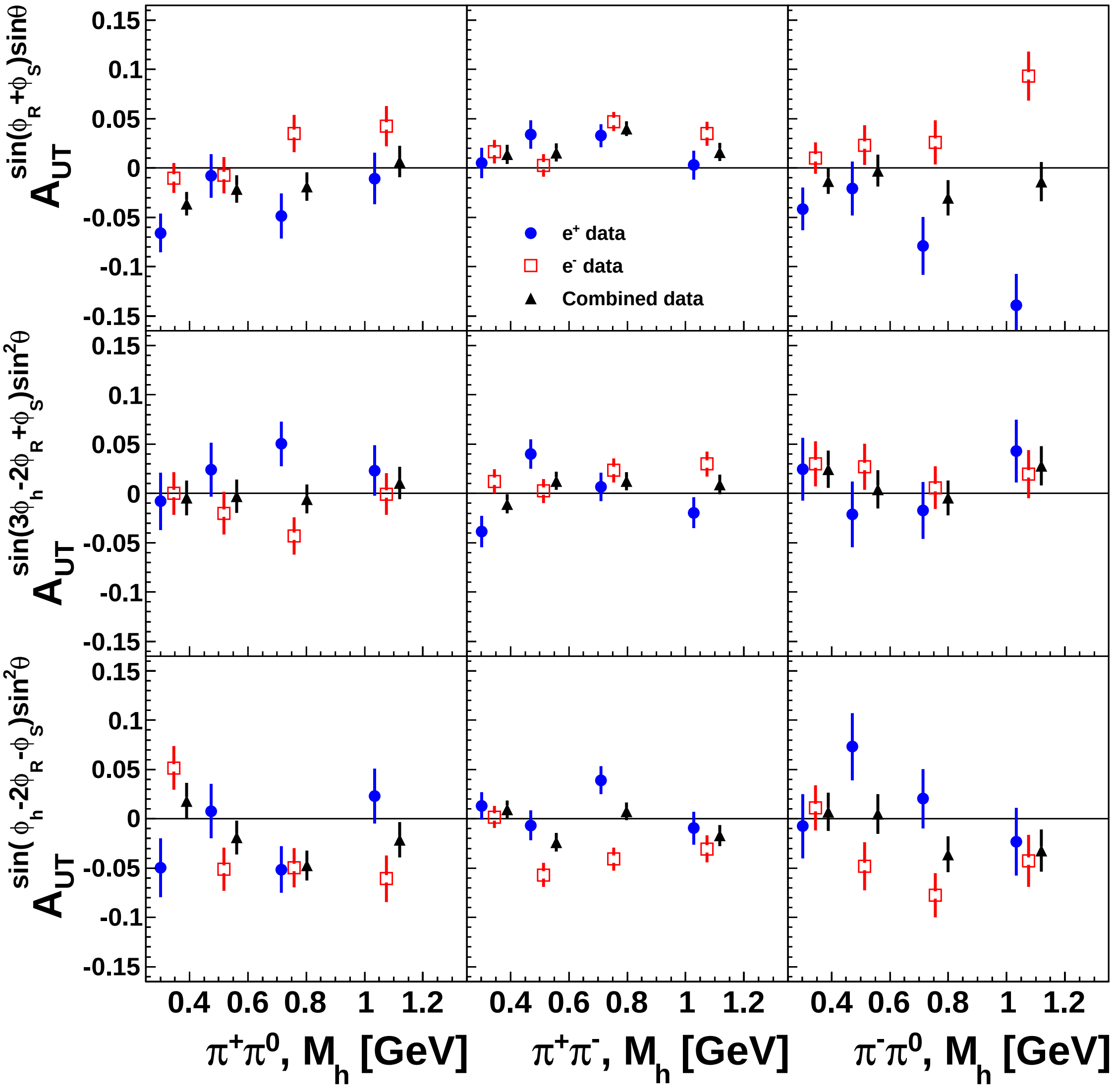


Figure 5.3: Comparison of select HERMES results versus lepton beam charge for pion-pair dihadrons, 1D binning. The rows, from top to bottom, represent the results for the $|1,1\rangle$, $|2,-2\rangle$, and $|2,2\rangle$ Collins moments. The columns, from left to right, are for dihadron types $\pi^+\pi^0$, $\pi^+\pi^-$, and $\pi^-\pi^0$, respectively. The blue circles are for positron beam, the red open squares are for electron beam, and the black triangles are results for the combined sample.

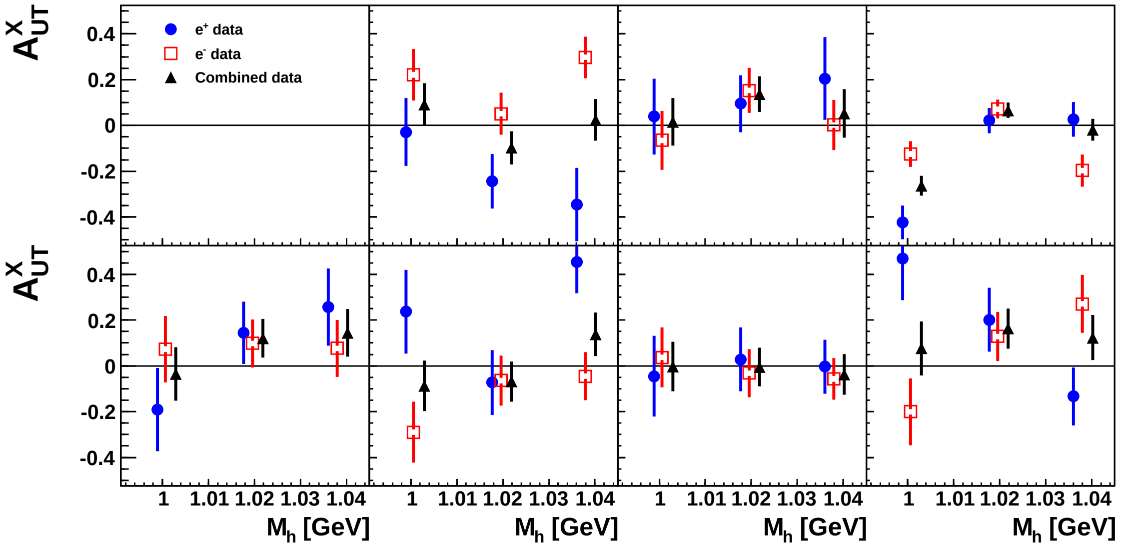


Figure 5.4: Comparison of select HERMES results versus lepton beam charge for $K^+K^-$ dihadrons, 1D binning. The panels represent various partial waves of the Sivers moments, and are arranged identical to Figure 5.2. As with Figure 5.3, the blue circles are for positron beam, the red open squares are for electron beam, and the black triangles are results for the combined sample.

that of the $\chi^2/ndf$ values for a few moments are in the range of 2.0 to 4.0, the conclusion is that the discrepancy is mostly statistical.

An uncertainty is assigned, per each moment and per each bin, by first determining the value of the assigned uncertainty that would reduce the $\chi^2$ value between the two sub-samples to unity. Since the discrepancy is assumed to be mostly statistical, this uncertainty is then divided by a factor of two. In the case that the $\chi^2$ value is less than one, no systematic uncertainty is assigned. Given that the results for the two sub-samples, electron and positron, are respectively $A_e \pm \delta A_e$ and $A_p \pm \delta A_p$, the uncertainty $\delta A_{year}$ is computed according to

$$\delta A_{year} = \frac{1}{4}\sqrt{\left(A_e - A_p\right)^2 - \delta^2 A_e - \delta^2 A_p}. \tag{5.1}$$

Note the factor of one quarter is the product of the extra factor of one half times the factor of one half arising from the fact that the uncertainty is being added to both samples. In the case that the squared difference is much larger than the sum of the variances, the resulting uncertainty is approximately one quarter of the magnitude of the difference, i.e. $|A_e - A_p|/4$. The results for this systematic uncertainty are shown in comparison with the other systematic uncertainties in Appendix B.

## 5.3 Hadronic Identification Procedure

Two methods exist for using the EVT algorithm for determining the particle identification of each hadron. One method is to simply assign each hadron the most likely identification. The other method is to consider the conditional probability of assigning a certain particle identification given the true identification. This conditional probability has been traditionally called the $P$-matrix among the HERMES collaboration. The second method is then to use the inverse of the $P$-matrix to assign a weight, interpreted as the probability that the given track is a certain type. Theoretical justification is given in Ref. [38], with the governing equation being a Fredholm integral equation, similar to that in Section 3.1. This second proposed method is denoted as RICH unfolding, as inverting a conditional probability statement has traditionally been denoted unfolding.

Neither method, a priori, is necessarily more accurate. Thus, both methods are considered, and half the magnitude of the difference in the results for each method is taken as the systematic uncertainty due to the hadronic identification procedure. Results from both methods, for select moments versus $M_h$, are shown in Figures 5.5 and 5.6. No significant difference between the methods is apparent, though the RICH unfolding method has slightly larger uncertainty. Note, the RICH unfolding results include the requirement that the absolute value of the weight is less than 5, to remove tracks with unreasonably high weights. The simpler method, without RICH unfolding, is the one used through this dissertation, and is, specifically, the method used for the final results.

This estimates of the uncertainty due to hadronic identification is shown in comparison with the other sources of systematic uncertainty in Appendix B. In general, this uncertainty is negligible in comparison with the other sources considered, though this was not known before conducting this study.

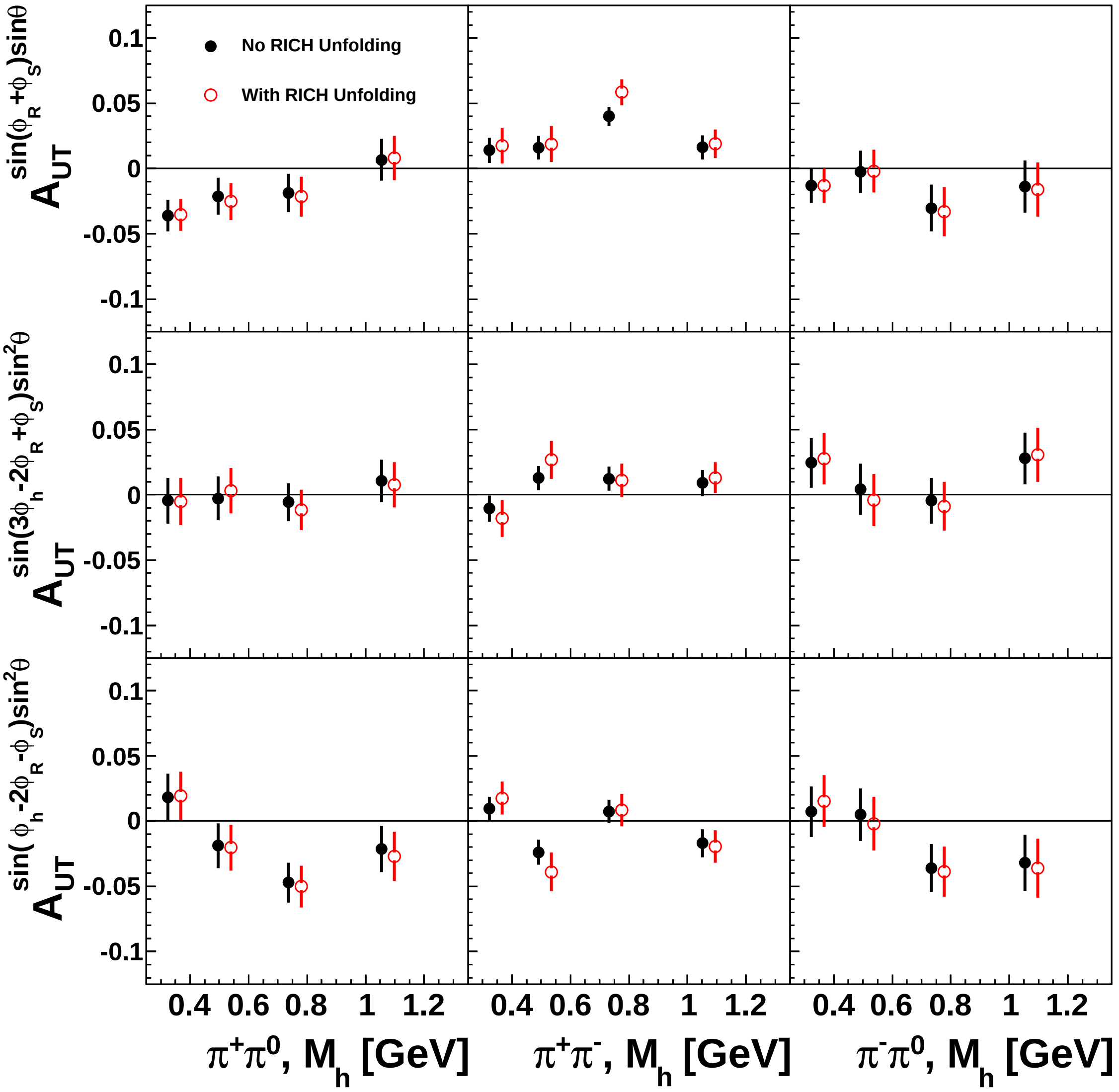


Figure 5.5: Comparison of select HERMES results versus hadron identification method for pion-pair dihadrons, 1D binning. The panels are arranged as in Figure 5.3. The black circles use the method of assigning the most likely particle type, while the red squares use the RICH unfolding method.

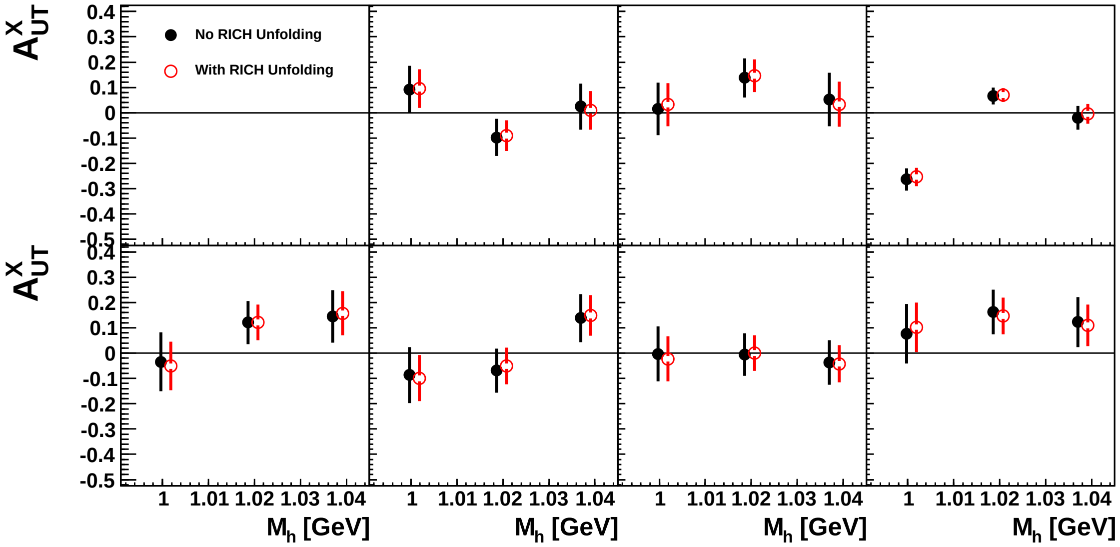


Figure 5.6: Comparison of select HERMES results versus hadron identification method for $K^+K^-$ dihadrons, 1D binning. The panels are arranged as in Figure 5.2. As with Figure 5.5, the black circles use the method of assigning the most likely particle type, while the red squares use the RICH unfolding method.

# CHAPTER VI

# Results and Conclusions

## 6.1 Results

Results for the Collins $|1,1\rangle$, $|2,-2\rangle$, and $|2,2\rangle$ moments of pion-pair dihadron production are presented and discussed in Sections 6.1.1 and 6.1.2. Results for the Sivers moments of $K^+K^-$ production, for all partial waves except $|1,0\rangle$ and $|2,0\rangle$, are presented and discussed in Section 6.1.3. Results for several additional moments, for all dihadron types considered, are given in Appendix C. The results in the appendix are considered stable, but their interpretation is more relevant in the context of a global fit rather than in the discussion of individual plots. Note, all result plots have a 7.3% scale uncertainty, indicated on each plot, due to uncertainty in the measurement of the target polarization.

### 6.1.1 Collins $|2,\pm 2\rangle$ Moments for Pion-Pair Dihadrons

The Collins $|1,1\rangle$, $|2,-2\rangle$, and $|2,2\rangle$ moments are shown versus invariant mass in Figure 6.1. One generally assumes that no hadron pairs are in the $|2,2\rangle$ partial wave except those resulting from a vector meson decay. Thus, one would expect non-zero moments only in the $M_h$ bin which includes the vector meson peak. Even within the mass bin including the vector meson peak, the moments are expected to be small, as the cross section for the SIDIS vector meson subprocess represents a fairly small portion of the total dihadron cross section. The results in Figure 6.1 support these expectations. Furthermore, in agreement with both the Lund/Artru and gluon radiation models, the $|2,2\rangle$ moments are small but non-zero (within the $M_h$ bin containing the $\rho$-mass peak), while the $|2,-2\rangle$ moments are consistent with zero.

The kinematic dependencies for the $|2,-2\rangle$ moment are shown in Figure 6.2. It is important to note that the moments are consistent with zero across all kinematic bins. Thus, the prior observation—that results for the $|2,-2\rangle$ moments in Figure 6.2 are consistent with zero—is not due to a cancellation of significantly positive and negative regions, but is generally true in all kinematic bins. This is again consistent with the Lund/Artru and gluon radiation models.

The kinematic dependencies for the $|2,2\rangle$ moment are shown in Figure 6.3. Within the uncertainties, no strong kinematic dependencies can be observed. The variation with each kinematic variable rarely changes the sign of the moment with statistical significance. A

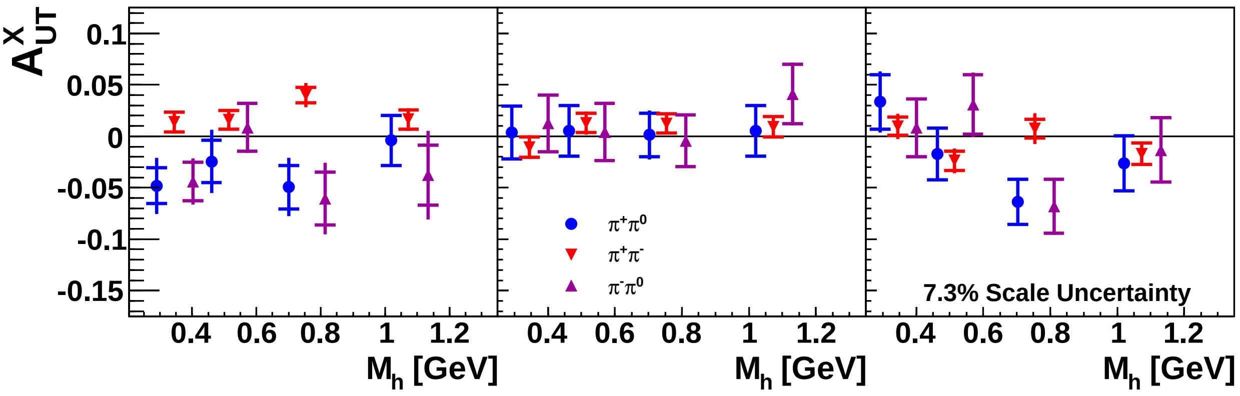


Figure 6.1: Final Results versus $M_h$ for the Collins $|1,1\rangle$, $|2,-2\rangle$, and $|2,2\rangle$ moments for pion-pair dihadrons. Statistical uncertainties are demarcated by the horizontal error bar, and the combined statistical and systematic uncertainties are indicated by the full error bar. The panels, from left to right, are for the $|1,1\rangle$, $|2,-2\rangle$, and $|2,2\rangle$ partial waves.

slight increase with $P_{h\perp}$ is generally observable, though, given the uncertainties, the results are also consistent with other possibilities.

In order to connect the given results with implications of the Lund/Artru and gluon radiation models, it is necessary to relate the disfavored and favored fragmentation functions with the observed dihadrons. To accomplish this, one generally assumes $u$ quark dominance, i.e. one assumes that the cross section is dominated by scattering off $u$ quarks, with the total contributions from other flavors being much smaller. The `TMDGen` models used in Chapters IV and V predict that $u$ quark scattering accounts for just under 70% of the cross section for all three pion-pair dihadrons. The $u$ quark dominance assumption implies that the $\pi^+$ and $\rho^+$ results are mainly due to favored fragmentation functions, while $\pi^-$ and $\rho^-$ results are primarily due to disfavored fragmentation functions. Results for $\pi^0$ are generally considered (and consistent with) the average of the results for $\pi^+$ and $\pi^-$. The Collins $|2,\pm2\rangle$ moments for $\rho^0$ production is more complex, possibly involving both the favored and disfavored fragmentation functions of $\rho^\pm$ production. Note, the given models and results both indicate that the moments for $\rho^0$ are not equal with the average of $\rho^+$ and $\rho^-$.

In order to test the Lund/Artru and gluon radiation models, a naive method to extract the vector meson signal has has been implemented. One defines the $M_h$ bin containing the vector meson peak as the peak $M_h$ bin. A simple linear interpolation, using the two bins on either side of the peak $M_h$ bin, is used to determine the non-vector meson signal in the peak $M_h$ bin. It is also assumed that the moments represent the value at the average $M_h$ of the given bin. Using the background fractions of Table 4.13, one can then estimate the vector meson signal. This is a particularly naive method, as it neglects shifts in the average kinematics between $M_h$ bins. However, the size of the uncertainties on the final results, especially when considering the kinematic dependencies, are high enough not to merit a more complicated procedure. Kinematic dependencies are also found to be not large, thus this naive method is sufficient.

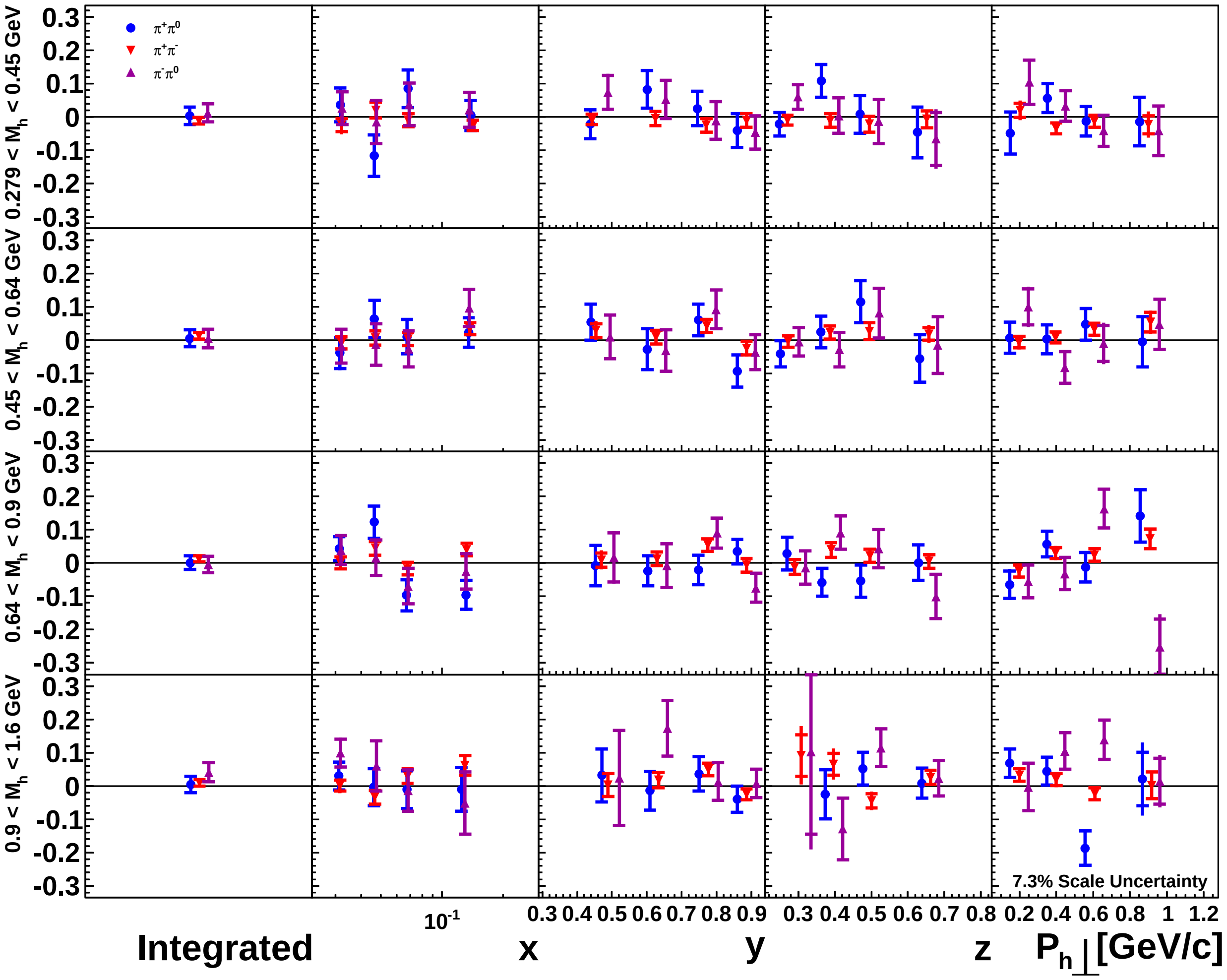


Figure 6.2: Final Results for the Collins $|2,-2\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. The panels are arranged as for other kinematic plots, with the addition of a column on the left for the 1D $M_h$ results. Additional columns, from left to right, represent binning with respect $x$, $y$, $z$, and $P_{h\perp}$, per each $M_h$ bin. The upper row of panels is for the lowest $M_h$ bin, with each row of panels being for the next higher $M_h$ bin. Blue circles indicate $\pi^+\pi^0$ dihadrons, red inverted triangles indicate $\pi^+\pi^-$ dihadrons, and upright purple triangles indicate $\pi^-\pi^0$ dihadrons.

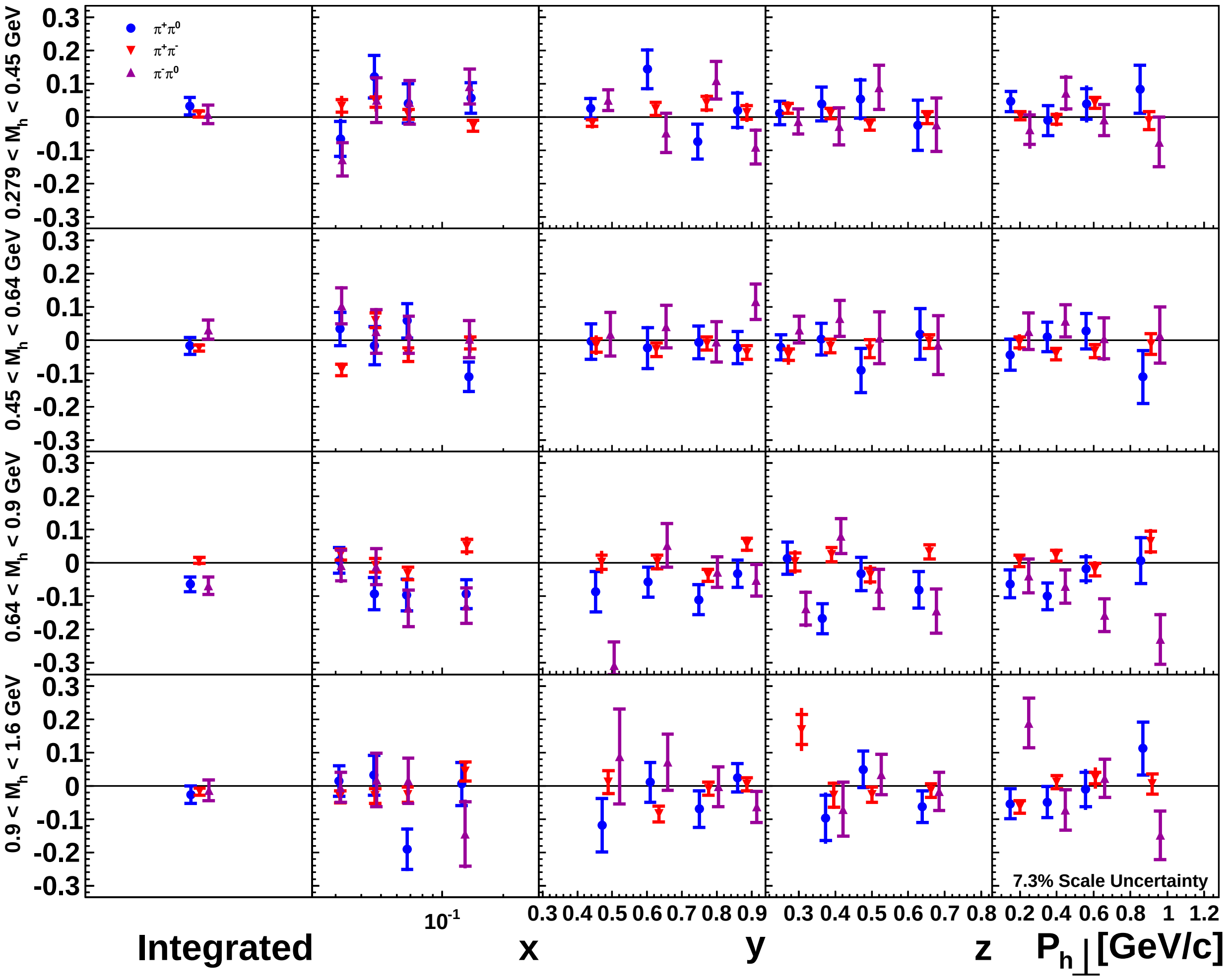


Figure 6.3: Final Results for the Collins $|2,2\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

The results of isolating the vector meson signal is shown in Figure 6.4. For their interpretation, let us assume $u$ quark dominance and that the pion results are as in Ref. [47]. The Lund/Artru model predicts the results for $\rho^+$ to be negative and the results for $\rho^0$ to be consistent with zero. The gluon radiation model implies the results for $\rho^-$ to also be negative and of comparable size to those for $\rho^+$. The magnitude of the moments are also expected to increase with $x$ and $P_{h\perp}$, based on the pion results, though the uncertainties on the kinematic dependencies are too large to confirm or invalidate these expectations. Furthermore, the Lund/Artru and gluon radiation models focus on the overall sign of the moments, allowing for different kinematic dependencies between the pseudo-scalar and dihadron results.

In general, the uncertainties are slightly higher than desirable, given the apparent size of the moments. However, the basic conclusions of the Lund/Artru model and the gluon radiation model regarding the sign of the SIDIS $|2,2\rangle$ moments are observed in the data.

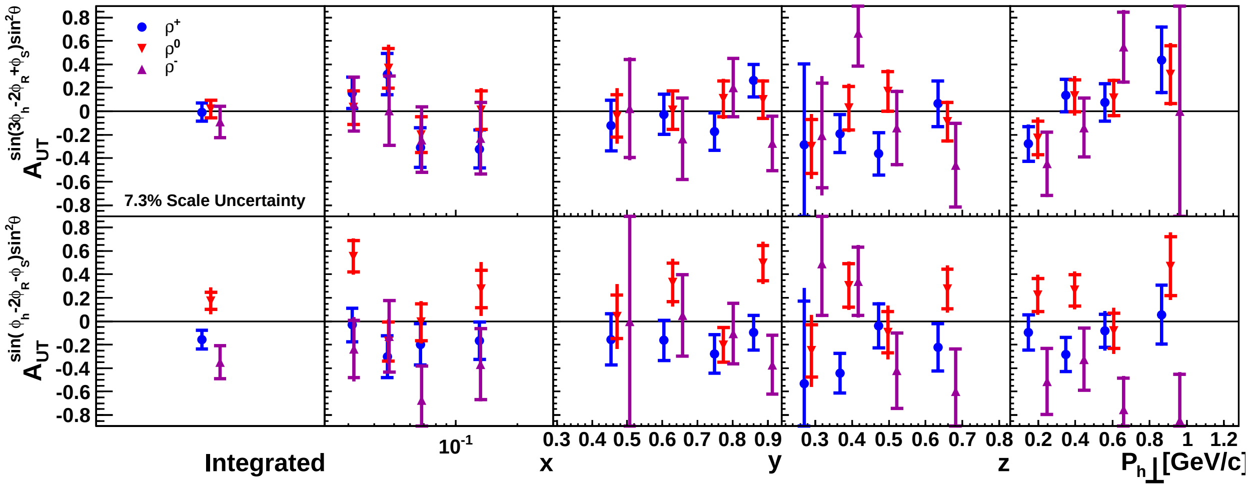


Figure 6.4: Final Results for the Collins $|2, \pm 2\rangle$ moments for $\rho$ mesons, including kinematic dependencies. Columns of and markers are arranged as in Figure 6.2. The top row indicates the $|2, -2\rangle$ moments, and the bottom row the $|2, 2\rangle$ moments.

### 6.1.2 Collins $|1, 1\rangle$ Moments for Pion-Pair Dihadrons

The Collins $|1, 1\rangle$ moment for pion-pair dihadrons is of theoretical interest as it allows collinear access to the transversity distribution function $h_1$. Results versus invariant mass are in Figure 6.1, left panel, while the kinematic dependencies are shown in Figure 6.6.

The results for $\pi^+\pi^-$ dihadrons are an updated version of the results in Ref. [27]. The results of Figures 6.1 and 6.6 include an acceptance correction, use of the angle $\phi_R$ rather than $\phi_{R\perp}$, and involve a different fitting procedure and function. Although the binning is slightly different, the comparison is quite close, as shown in Figure 6.5. When comparing the results versus $M_h$, the bins outside the mass peak are very consistent. The bin containing the $\rho^0$ mass peak has results differing by about one to two standard deviations, though part of the effect can be related to the narrower bin in this analysis. This discrepancy is also seen when comparing the $x$ and $z$ dependencies within this $M_h$ region near the $\rho$ meson mass peak. Some hoped that the improved results would yield moments two or three times larger, thus reducing the discrepancy between previous HERMES results [27] and COMPASS results [40]. Although the $|1, 1\rangle$ moment in the peak is larger than in Ref. [27], it is questionable whether this increase is not enough to recover the difference.

The kinematic dependence of the moments between this analysis and Ref. [27] are somewhat different. While both support larger values of the moments for medium values of $x$, the results herein indicate a decrease of the moment with increasing $z$, a feature not present in Ref. [27]. The results of Ref. [27] also tend to be much more noisy than the results presented here.

As with the $|2, 2\rangle$ moments, the $\pi^+\pi^-$ results are not the average of the $\pi^\pm\pi^0$ results, as one might expect from naive isospin invariance. The $\pi^+\pi^0$ and $\pi^-\pi^0$ moments are, in general, quite similar and opposite in sign to the $\pi^+\pi^-$ results. There is no indication among any of

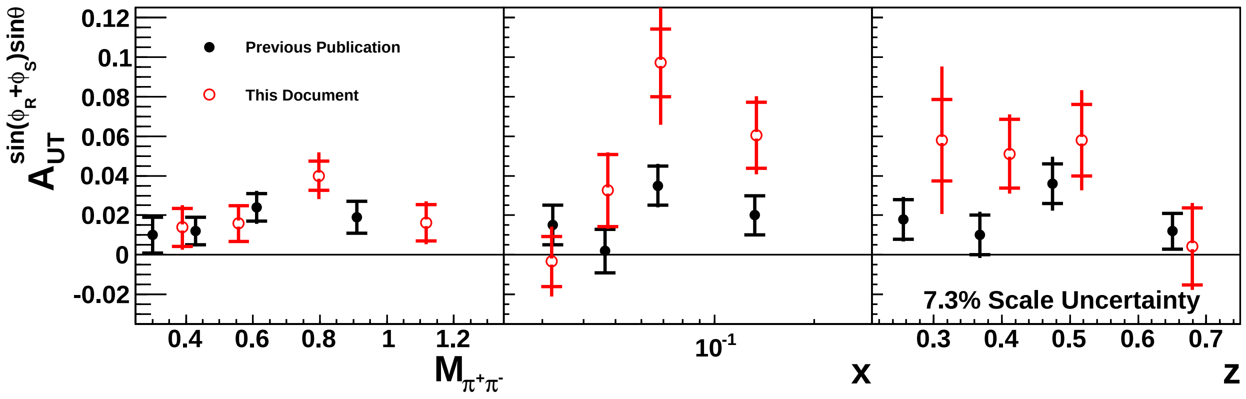


Figure 6.5: Comparison of the Collins $|1,1\rangle$ moment with previous results. Previous results from Ref. [27] are shown with filled black circles, while the results from this dissertation are shown with open red circles. The left panels show results with respect to $M_{\pi^+\pi^-}$. The middle and right panels show, respectively, the $x$ and $z$ dependence within the $M_h$ region including $\rho$ meson mass peak. Note the $M_{\pi^+\pi^-}$ region and the $x$ and $z$ bins are not consistent between Ref. [27] and this document. The full error bar indicates combined statistical and systematic uncertainty, while the horizontal bar indicates the statistical uncertainty.

the pion-pair dihadrons of a sign change of the fragmentation function across the $\rho$ meson peak, as suggested in Refs. [104, 105].

### 6.1.3 Sivers Moments for $K^+K^-$ Dihadrons

The currently theory is quite vague regarding predictions for $K^+K^-$ dihadrons and possible implications of the results. Unfortunately, the data is statistically limited, not allowing a determination of the $x$ dependence. Without the $x$ dependence, little can be said regarding the Sivers function.

It is worth noting, however, that no clear signal is observed within the middle $M_h$ bin, which contains the $\phi$ meson peak. The background fraction is fairly low, on the order of one third, and so a signal will not be masked nearly as strongly as in the pion-pair dihadron case. The only indication of a difference between the central $M_h$ bin versus the exterior $M_h$ bins is for the $|0,0\rangle$ partial wave, which includes contributions from both longitudinal $\phi$ mesons and non-resonant kaon pairs. Unfortunately, this data set can neither confirm nor exclude a non-zero Sivers function for strange quarks.

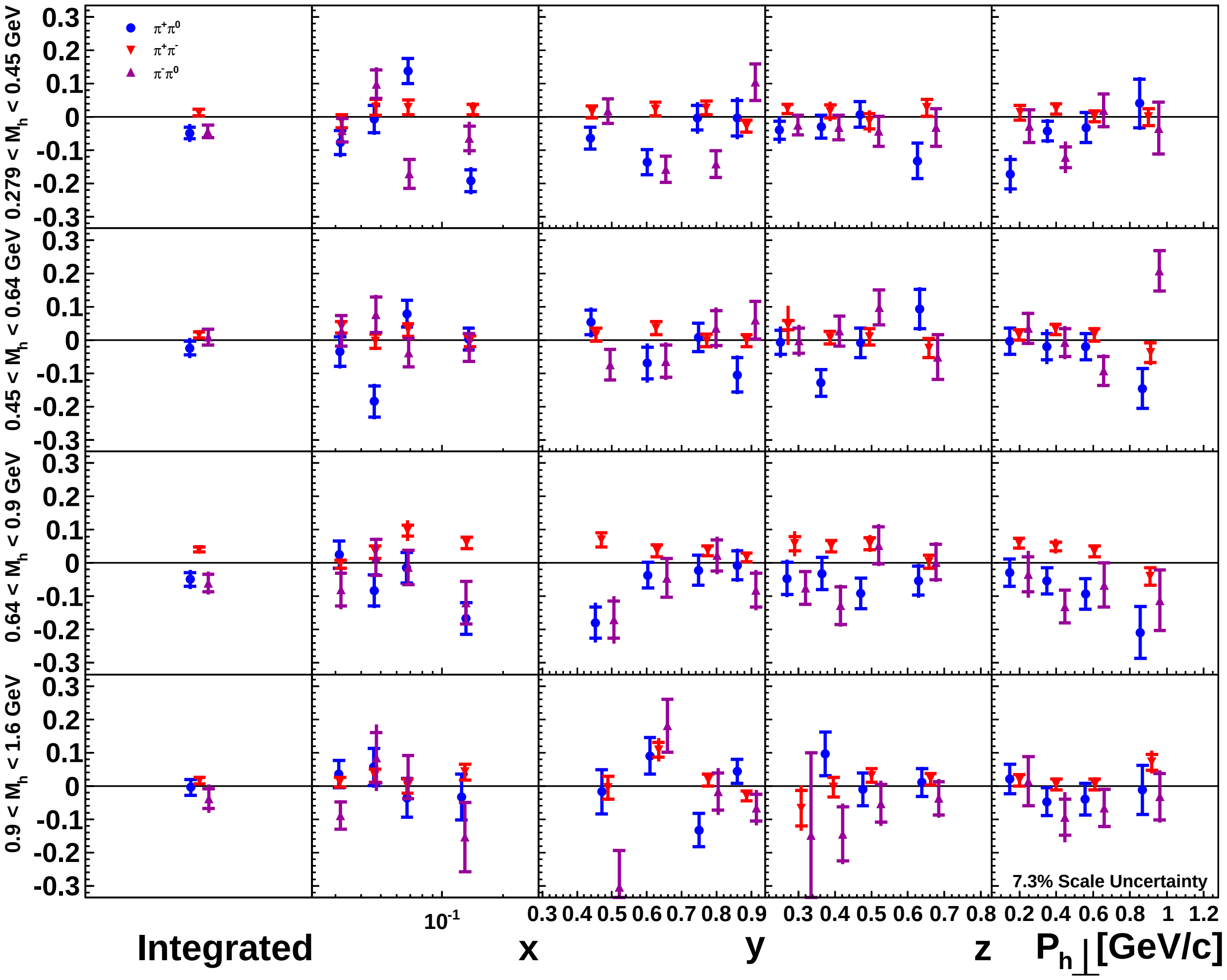


Figure 6.6: Final Results for the Collins $|1, 1\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

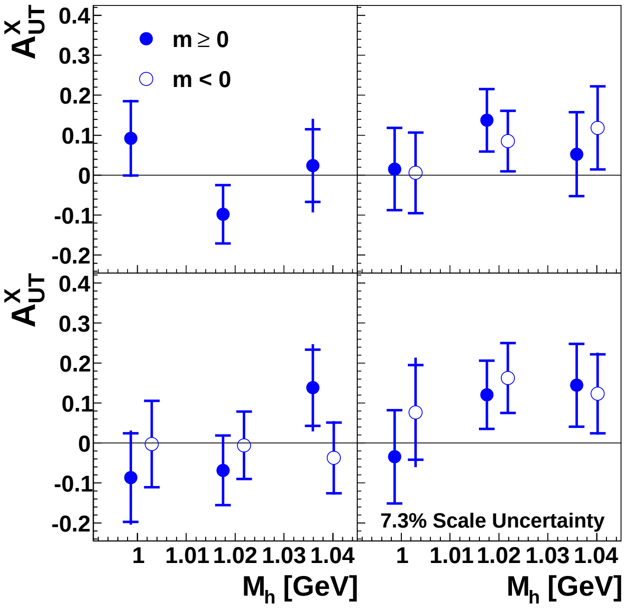


Figure 6.7: Final Results versus $M_h$ for select Sivers moments for $K^+K^-$ dihadrons. The panels are, top row, left to right, for $|0,0\rangle$, and $|1,\pm1\rangle$ partial waves, and the bottom panels are, left to right, for $|2,\pm1\rangle$ and $|2,\pm2\rangle$ partial waves. Filled circles indicate $m \geq 0$, and open circles $m < 0$.

## 6.2 Conclusions

This dissertation has focused on the TMD analysis of select transverse target moments of SIDIS dihadron production. The main motivation has been to test the Lund/Artru model of fragmentation, investigate the gluon shower model, and provide measurements related to the strange quark flavor of the Sivers function, with the side benefit of improving the HERMES results for collinear access to transversity.

Several theoretical advancements have been described, including a new partial wave expansion of the cross section and the computation of the next-to-leading twist dihadron cross section. This has allowed the quantification of the connection between the Collins moment of pseudo-scalar meson production and the several Collins-like moments in dihadron production, and organizes the complexity of the many moments occurring in the dihadron cross section. Specific predictions of the Lund/Artru model, related to moments of the cross section, have been determined, and a new gluon radiation fragmentation model has been put forth. Additionally, a new spectator model calculation for dihadron TMD fragmentation functions has also been completed.

A TMD Monte Carlo generator, `TMDGen`, has been written, which includes polarized SIDIS pseudo-scalar and dihadron production. Although `TMDGen` was written for this analysis, the generator was designed with the expansion to additional analysis and experiments in mind. `TMDGen` also includes a full simulation of the intrinsic transverse momenta, $\boldsymbol{p}_T$ and $\boldsymbol{k}_T$.

The analysis of the transverse target moments for SIDIS dihadron production from HERMES data is also presented, specifically for $\pi^+\pi^0$, $\pi^+\pi^-$, $\pi^-\pi^0$, and $K^+K^-$ dihadrons. A new acceptance correction method has been proposed and tested, and relevant systematic uncertainties have been estimated.

The results for the Collins $|2, \pm 2\rangle$ moments for pion-pair dihadrons are in agreement with expectations. The Collins $|2, -2\rangle$ moment seems everywhere consistent with zero, as expected from the struck quark being in the positive transverse polarization state. A signal is seen for the $|2, 2\rangle$ moment in the $\rho$ peak mass bin, while the results in other mass bins are consistent with zero.

Additionally, the signs of the $|2, 2\rangle$ partial waves for $\rho^+$ and $\rho^-$ production are consistent with the Lund/Artru and gluon radiation models, given the previous $\pi^\pm$ results. The uncertainties are too large to allow a good comparison regarding the kinematic dependencies, though the models are mainly concerned with the sign of the moments, not the full kinematic dependence.

The results for the Collins $|1, 1\rangle$ moment for $\pi^+\pi^-$ dihadrons is generally in agreement with Ref. [27], though the results of this document are much less noisy. Some kinematic dependence is now observable, which was masked by noise in the previous result. Although predictions are not available for the results for the $\pi^\pm\pi^0$ dihadrons, the results for $\pi^+\pi^0$ are similar to those for $\pi^-\pi^0$, and both tend to be of the opposite sign as the $\pi^+\pi^-$ results. These trends were also observed in the $|2, 2\rangle$ moments: that the $|1, 1\rangle$ is similar is not surprising.

Note, neither the fragmentation functions computed in Chapter II, nor those in Ref. [31], can be used as model predictions for $\pi^\pm\pi^0$ dihadrons, since they involve some parameters that need to be fit with data. However, the presented results allow such a fit for $\pi^\pm\pi^0$ dihadrons, as has been previously done for $\pi^+\pi^-$ dihadrons [40]. The fit of Ref. [40] for

$\pi^+\pi^-$ dihadron results should also be repeated with the newer data set, though it seems the newer results of in this dissertation will not reconcile the differences between the HERMES and COMPASS results.

With regard to the Sivers moments for $K^+K^-$ dihadrons, no clear indication exists that the signal for the $\phi$ meson production subprocess is different than that for non-resonant kaon-pair dihadron production. These results are consistent with the gluons having either relatively small, or possibly zero, orbital angular momentum.

The main goals and motivation of this dissertation have been accomplished, with the results generally in agreement with expectations. This dissertation represents the first transverse momentum dependent analysis of polarized SIDIS dihadron production. It is hoped that future analyses and experiments will utilize the theoretical developments provided herein and improve upon these results. Already, an analysis is in progress at Jefferson Laboratory, regarding collinear SIDIS dihadron production with a longitudinally polarized beam and unpolarized target [106], though the possibilities for a TMD analysis and for other polarization states are existent. Collinear analyses are also in progress, or have been completed, at PHOENIX and COMPASS, and transverse momentum dependent analyses could be conducted at these experiments in the future. The BELLE collaboration has released results for collinear dihadron fragmentation, occurring in electron positron annihilation, and it is hoped they will also consider transverse momentum dependent dihadron fragmentation. The most information might yet come from future experiments, such as the proposed electron ion collider. However, each incremental step will continue to further our understanding of hadronic structure and the fragmentation process, today and in the future.

# APPENDICES

# APPENDIX A

# Additional Tables

Two studies have been discussed in this dissertation wherein a model is induced in a data set with a perfect $4\pi$ detector as well as in a data set within acceptance, within acceptance meaning the data has been through a full simulation of the HERMES detector and has been reconstructed by the usual data processing chain. In Section 4.4.2, the data is generated by `TMDGen`, while in Section 5.1, the data is generated by `Pythia`, with the reconstructed data also including radiative effects.

For both studies, a $\chi^2/ndf$ per transverse target moment is computed, averaged over all bins within each choice of binning, as described in Section 4.4.2. As these tables are somewhat lengthy, they have been placed in this appendix rather than within the respective chapters. Results relevant for Section 4.4.2 are given in Table A.1 through Table A.4 and are denoted "Challenge A Results." Results relevant for Section 5.1 are given in Table A.5 through Table A.8 and are denoted "Challenge B Results."

Two additional studies, denoted "Year Dependence Study 1" and "Year Dependence Study 2" are discussed in Section 5.2. Study 1 includes the comparison of fitting the positron (2002-2004) and electron (2005) data simultaneously versus fitting each data separately and forming the appropriate linear combination of the results. Tables of $\chi^2/ndf$ statistics, in the same format as those for Challenge A and B, are given in Table A.9 through Table A.12, Year Dependence Study 2 tests for consistency between the separate fits of the positron and electron data, and tables of $\chi^2/ndf$ statistics are given in Table A.13 through Table A.16.

| | $\boldsymbol{\chi^2/ndf}$ **per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{x}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{y}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{z}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{P_{h\perp}}$ |
| Sivers $\|0, 0\rangle$ | 24.262 | 7.933 | 7.442 | 26.445 | 11.041 |
| Sivers $\|1, -1\rangle$ | 4.130 | 1.798 | 1.501 | 1.423 | 1.368 |
| Sivers $\|1, 0\rangle$ | 32.172 | 40.630 | 17.661 | 50.027 | 11.113 |
| Sivers $\|1, 1\rangle$ | 2.105 | 1.335 | 1.259 | 0.842 | 1.038 |
| Sivers $\|2, -2\rangle$ | 2.267 | 1.055 | 1.516 | 1.505 | 1.073 |
| Sivers $\|2, -1\rangle$ | 1.558 | 3.525 | 3.379 | 7.401 | 4.048 |
| Sivers $\|2, 0\rangle$ | 206.813 | 63.409 | 59.620 | 201.947 | 65.524 |
| Sivers $\|2, 1\rangle$ | 3.309 | 2.847 | 3.466 | 30.484 | 4.804 |
| Sivers $\|2, 2\rangle$ | 1.851 | 1.047 | 1.668 | 0.785 | 1.426 |
| Collins $\|0, 0\rangle$ | 11.510 | 6.359 | 4.172 | 15.407 | 7.394 |
| Collins $\|1, -1\rangle$ | 1.947 | 1.501 | 1.109 | 0.610 | 1.487 |
| Collins $\|1, 0\rangle$ | 67.696 | 63.585 | 19.607 | 15.370 | 25.493 |
| Collins $\|1, 1\rangle$ | 5.863 | 1.851 | 1.835 | 1.423 | 2.025 |
| Collins $\|2, -2\rangle$ | 0.392 | 1.108 | 1.012 | 1.041 | 0.393 |
| Collins $\|2, -1\rangle$ | 2.708 | 2.208 | 3.038 | 21.793 | 1.687 |
| Collins $\|2, 0\rangle$ | 49.310 | 33.906 | 18.686 | 125.888 | 18.458 |
| Collins $\|2, 1\rangle$ | 5.632 | 2.865 | 1.766 | 8.277 | 3.025 |
| Collins $\|2, 2\rangle$ | 3.906 | 1.857 | 1.157 | 1.787 | 1.854 |

Table A.1: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^0$ dihadrons, Challenge A. The $\chi^2/ndf$ is computed over the various bins, for each different choice of binning. The number of degrees of freedom is 3 for the 1D $M_h$ binning and 15 for each of the 2D binning options.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 8.332 | 2.922 | 2.410 | 3.141 | 3.348 |
| Sivers $\|1,-1\rangle$ | 1.377 | 1.215 | 1.209 | 0.797 | 1.650 |
| Sivers $\|1,0\rangle$ | 3.672 | 13.696 | 6.074 | 24.869 | 4.027 |
| Sivers $\|1,1\rangle$ | 3.885 | 1.535 | 1.459 | 0.600 | 2.031 |
| Sivers $\|2,-2\rangle$ | 4.464 | 1.763 | 2.455 | 1.407 | 1.991 |
| Sivers $\|2,-1\rangle$ | 4.221 | 2.746 | 5.739 | 1.213 | 1.529 |
| Sivers $\|2,0\rangle$ | 17.283 | 13.656 | 16.336 | 10.363 | 8.876 |
| Sivers $\|2,1\rangle$ | 9.009 | 2.495 | 7.246 | 2.522 | 2.905 |
| Sivers $\|2,2\rangle$ | 5.555 | 2.251 | 1.558 | 1.069 | 1.922 |
| Collins $\|0,0\rangle$ | 0.597 | 1.297 | 2.691 | 2.121 | 1.671 |
| Collins $\|1,-1\rangle$ | 1.534 | 1.045 | 1.032 | 1.090 | 0.666 |
| Collins $\|1,0\rangle$ | 3.332 | 7.135 | 3.347 | 10.184 | 4.462 |
| Collins $\|1,1\rangle$ | 1.852 | 1.149 | 1.722 | 1.915 | 1.610 |
| Collins $\|2,-2\rangle$ | 1.796 | 0.709 | 0.547 | 1.265 | 1.020 |
| Collins $\|2,-1\rangle$ | 2.012 | 0.965 | 2.373 | 1.427 | 1.437 |
| Collins $\|2,0\rangle$ | 6.309 | 9.775 | 22.394 | 8.380 | 10.849 |
| Collins $\|2,1\rangle$ | 2.372 | 1.702 | 2.285 | 5.366 | 2.044 |
| Collins $\|2,2\rangle$ | 0.681 | 0.547 | 1.293 | 0.965 | 1.067 |

Table A.2: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^-$ dihadrons, Challenge A. See caption for Table A.1.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 18.911 | 7.036 | 8.475 | 15.532 | 4.144 |
| Sivers $\|1,-1\rangle$ | 1.200 | 1.152 | 1.677 | 1.247 | 1.819 |
| Sivers $\|1,0\rangle$ | 23.597 | 20.808 | 17.314 | 35.048 | 42.272 |
| Sivers $\|1,1\rangle$ | 2.678 | 1.058 | 1.423 | 1.194 | 2.728 |
| Sivers $\|2,-2\rangle$ | 2.819 | 1.554 | 1.405 | 0.920 | 1.849 |
| Sivers $\|2,-1\rangle$ | 18.069 | 6.053 | 6.073 | 6.626 | 5.915 |
| Sivers $\|2,0\rangle$ | 171.745 | 91.661 | 53.056 | 104.607 | 49.063 |
| Sivers $\|2,1\rangle$ | 9.204 | 3.040 | 3.794 | 3.380 | 2.215 |
| Sivers $\|2,2\rangle$ | 4.381 | 1.090 | 1.326 | 0.844 | 1.954 |
| Collins $\|0,0\rangle$ | 40.907 | 10.621 | 12.926 | 32.680 | 15.072 |
| Collins $\|1,-1\rangle$ | 1.305 | 1.411 | 0.893 | 1.178 | 0.537 |
| Collins $\|1,0\rangle$ | 116.182 | 71.356 | 45.478 | 67.321 | 32.063 |
| Collins $\|1,1\rangle$ | 2.167 | 1.806 | 0.983 | 0.592 | 1.051 |
| Collins $\|2,-2\rangle$ | 1.160 | 1.515 | 1.116 | 1.194 | 0.995 |
| Collins $\|2,-1\rangle$ | 8.307 | 4.052 | 3.649 | 10.863 | 2.406 |
| Collins $\|2,0\rangle$ | 164.146 | 38.817 | 56.540 | 143.452 | 98.096 |
| Collins $\|2,1\rangle$ | 3.287 | 2.363 | 2.969 | 3.225 | 2.572 |
| Collins $\|2,2\rangle$ | 1.107 | 0.558 | 0.909 | 1.289 | 0.888 |

Table A.3: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^-\pi^0$ dihadrons, Challenge A. See caption for Table A.1.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\vert 0,0\rangle$ | 2.893 | 0.879 | 0.753 | 0.773 | 0.773 |
| Sivers $\vert 1,-1\rangle$ | 1.426 | 0.638 | 1.887 | 0.384 | 0.956 |
| Sivers $\vert 1,0\rangle$ | 0.966 | 2.479 | 2.814 | 3.826 | 1.314 |
| Sivers $\vert 1,1\rangle$ | 0.424 | 0.766 | 0.739 | 0.484 | 0.585 |
| Sivers $\vert 2,-2\rangle$ | 2.557 | 1.629 | 4.248 | 2.167 | 1.543 |
| Sivers $\vert 2,-1\rangle$ | 6.033 | 1.530 | 3.755 | 0.862 | 1.931 |
| Sivers $\vert 2,0\rangle$ | 10.648 | 6.473 | 5.043 | 5.731 | 5.742 |
| Sivers $\vert 2,1\rangle$ | 4.680 | 1.571 | 2.500 | 0.526 | 1.000 |
| Sivers $\vert 2,2\rangle$ | 7.071 | 2.336 | 5.544 | 0.960 | 1.402 |
| Collins $\vert 0,0\rangle$ | 1.982 | 1.436 | 1.168 | 0.954 | 0.439 |
| Collins $\vert 1,-1\rangle$ | 2.345 | 0.806 | 0.955 | 0.743 | 0.757 |
| Collins $\vert 1,0\rangle$ | 2.110 | 4.455 | 2.847 | 1.814 | 1.585 |
| Collins $\vert 1,1\rangle$ | 0.282 | 0.686 | 0.596 | 0.338 | 0.481 |
| Collins $\vert 2,-2\rangle$ | 0.603 | 0.594 | 1.744 | 0.560 | 0.529 |
| Collins $\vert 2,-1\rangle$ | 2.109 | 1.656 | 1.686 | 0.803 | 1.311 |
| Collins $\vert 2,0\rangle$ | 0.785 | 1.816 | 2.229 | 1.304 | 2.237 |
| Collins $\vert 2,1\rangle$ | 1.681 | 0.406 | 1.861 | 1.357 | 1.419 |
| Collins $\vert 2,2\rangle$ | 2.914 | 0.789 | 1.397 | 1.742 | 1.341 |

Table A.4: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $K^+K^-$ dihadrons, Challenge A. See caption for Table A.1. Note, the number of degrees of freedom for $K^+K^-$ dihadrons is 2 for the 1D $M_h$ binning and 11 for each of the 2D binning options.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 10.257 | 6.154 | 5.898 | 6.199 | 6.886 |
| Sivers $\|1,-1\rangle$ | 8.649 | 2.064 | 1.872 | 2.681 | 3.024 |
| Sivers $\|1,0\rangle$ | 38.928 | 48.047 | 27.105 | 59.303 | 16.620 |
| Sivers $\|1,1\rangle$ | 1.072 | 1.729 | 2.029 | 1.393 | 1.549 |
| Sivers $\|2,-2\rangle$ | 8.710 | 1.312 | 2.256 | 1.948 | 2.242 |
| Sivers $\|2,-1\rangle$ | 14.156 | 7.346 | 5.586 | 11.712 | 5.233 |
| Sivers $\|2,0\rangle$ | 191.392 | 81.096 | 46.959 | 106.730 | 80.811 |
| Sivers $\|2,1\rangle$ | 9.984 | 1.987 | 6.877 | 4.140 | 4.155 |
| Sivers $\|2,2\rangle$ | 1.746 | 0.987 | 0.993 | 1.409 | 1.403 |
| Collins $\|0,0\rangle$ | 12.917 | 5.923 | 9.475 | 24.251 | 6.392 |
| Collins $\|1,-1\rangle$ | 0.806 | 1.851 | 1.135 | 2.099 | 2.088 |
| Collins $\|1,0\rangle$ | 47.455 | 31.840 | 37.332 | 45.703 | 20.431 |
| Collins $\|1,1\rangle$ | 16.554 | 2.497 | 3.843 | 4.319 | 3.131 |
| Collins $\|2,-2\rangle$ | 0.605 | 1.011 | 0.465 | 0.569 | 1.363 |
| Collins $\|2,-1\rangle$ | 12.480 | 2.694 | 2.772 | 14.441 | 3.673 |
| Collins $\|2,0\rangle$ | 33.781 | 32.088 | 28.132 | 174.760 | 16.624 |
| Collins $\|2,1\rangle$ | 3.693 | 2.127 | 2.664 | 10.043 | 1.161 |
| Collins $\|2,2\rangle$ | 1.596 | 0.740 | 1.227 | 1.364 | 1.048 |

Table A.5: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^0$ dihadrons, Challenge B. See caption for Table A.1.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 8.150 | 2.759 | 4.746 | 4.965 | 16.288 |
| Sivers $\|1,-1\rangle$ | 4.021 | 2.722 | 1.655 | 2.871 | 2.283 |
| Sivers $\|1,0\rangle$ | 14.778 | 8.770 | 11.368 | 27.578 | 7.217 |
| Sivers $\|1,1\rangle$ | 4.613 | 1.152 | 2.238 | 2.515 | 1.831 |
| Sivers $\|2,-2\rangle$ | 8.545 | 1.821 | 3.128 | 1.812 | 3.992 |
| Sivers $\|2,-1\rangle$ | 15.639 | 2.840 | 4.841 | 4.418 | 13.097 |
| Sivers $\|2,0\rangle$ | 21.056 | 8.615 | 21.184 | 17.354 | 20.621 |
| Sivers $\|2,1\rangle$ | 0.230 | 1.640 | 2.964 | 1.078 | 7.262 |
| Sivers $\|2,2\rangle$ | 0.496 | 0.797 | 1.442 | 1.794 | 1.469 |
| Collins $\|0,0\rangle$ | 6.811 | 3.382 | 4.183 | 2.828 | 11.761 |
| Collins $\|1,-1\rangle$ | 2.905 | 1.208 | 1.176 | 1.628 | 1.368 |
| Collins $\|1,0\rangle$ | 13.143 | 2.759 | 12.866 | 7.086 | 9.756 |
| Collins $\|1,1\rangle$ | 2.002 | 2.266 | 0.952 | 1.823 | 1.131 |
| Collins $\|2,-2\rangle$ | 0.471 | 1.314 | 1.254 | 0.717 | 0.876 |
| Collins $\|2,-1\rangle$ | 5.851 | 1.984 | 4.245 | 0.818 | 3.332 |
| Collins $\|2,0\rangle$ | 22.087 | 10.313 | 7.726 | 26.236 | 17.176 |
| Collins $\|2,1\rangle$ | 11.138 | 3.387 | 1.891 | 2.985 | 7.727 |
| Collins $\|2,2\rangle$ | 4.056 | 1.806 | 2.383 | 1.803 | 1.948 |

Table A.6: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^-$ dihadrons, Challenge B. See caption for Table A.1.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 25.319 | 8.626 | 8.475 | 20.388 | 7.444 |
| Sivers $\|1,-1\rangle$ | 6.011 | 2.119 | 2.794 | 1.246 | 1.993 |
| Sivers $\|1,0\rangle$ | 8.677 | 31.761 | 16.525 | 25.274 | 16.724 |
| Sivers $\|1,1\rangle$ | 1.572 | 0.666 | 1.667 | 1.881 | 0.814 |
| Sivers $\|2,-2\rangle$ | 4.522 | 0.572 | 1.364 | 1.402 | 1.925 |
| Sivers $\|2,-1\rangle$ | 9.338 | 5.064 | 6.190 | 13.393 | 6.561 |
| Sivers $\|2,0\rangle$ | 139.563 | 41.736 | 49.833 | 163.535 | 66.695 |
| Sivers $\|2,1\rangle$ | 4.105 | 2.379 | 2.771 | 11.088 | 5.461 |
| Sivers $\|2,2\rangle$ | 1.546 | 0.989 | 1.995 | 1.513 | 1.683 |
| Collins $\|0,0\rangle$ | 23.867 | 7.926 | 11.397 | 13.055 | 7.601 |
| Collins $\|1,-1\rangle$ | 0.793 | 1.354 | 2.126 | 1.734 | 1.947 |
| Collins $\|1,0\rangle$ | 7.822 | 11.987 | 19.515 | 43.712 | 20.314 |
| Collins $\|1,1\rangle$ | 10.736 | 2.290 | 4.205 | 2.273 | 4.068 |
| Collins $\|2,-2\rangle$ | 1.255 | 0.955 | 1.003 | 1.081 | 1.957 |
| Collins $\|2,-1\rangle$ | 1.238 | 1.324 | 1.183 | 2.410 | 2.824 |
| Collins $\|2,0\rangle$ | 51.900 | 30.394 | 96.140 | 75.190 | 41.508 |
| Collins $\|2,1\rangle$ | 4.916 | 2.660 | 3.965 | 5.619 | 2.545 |
| Collins $\|2,2\rangle$ | 0.834 | 1.055 | 0.822 | 1.023 | 1.211 |

Table A.7: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^-\pi^0$ dihadrons, Challenge B. See caption for Table A.1.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h\text{-}x}$ | $\boldsymbol{M_h\text{-}y}$ | $\boldsymbol{M_h\text{-}z}$ | $\boldsymbol{M_h\text{-}P_{h\perp}}$ |
| Sivers $\lvert 0,0\rangle$ | 1.979 | 0.813 | 0.662 | 1.038 | 0.991 |
| Sivers $\lvert 1,-1\rangle$ | 0.354 | 0.259 | 0.616 | 0.607 | 0.673 |
| Sivers $\lvert 1,0\rangle$ | 47.171 | 15.372 | 18.166 | 16.839 | 16.269 |
| Sivers $\lvert 1,1\rangle$ | 1.138 | 0.880 | 1.161 | 0.953 | 0.991 |
| Sivers $\lvert 2,-2\rangle$ | 0.328 | 1.353 | 0.941 | 1.578 | 0.852 |
| Sivers $\lvert 2,-1\rangle$ | 1.936 | 1.207 | 2.272 | 1.416 | 2.074 |
| Sivers $\lvert 2,0\rangle$ | 5.350 | 7.257 | 9.452 | 4.908 | 26.442 |
| Sivers $\lvert 2,1\rangle$ | 0.443 | 1.412 | 0.838 | 3.130 | 1.239 |
| Sivers $\lvert 2,2\rangle$ | 0.398 | 0.904 | 1.379 | 1.751 | 1.415 |
| Collins $\lvert 0,0\rangle$ | 1.032 | 0.719 | 0.944 | 1.465 | 0.652 |
| Collins $\lvert 1,-1\rangle$ | 1.364 | 1.242 | 1.042 | 1.342 | 1.327 |
| Collins $\lvert 1,0\rangle$ | 27.877 | 10.402 | 7.462 | 8.711 | 7.011 |
| Collins $\lvert 1,1\rangle$ | 3.638 | 1.385 | 0.774 | 1.678 | 1.330 |
| Collins $\lvert 2,-2\rangle$ | 1.015 | 0.929 | 1.948 | 1.399 | 0.786 |
| Collins $\lvert 2,-1\rangle$ | 0.385 | 1.355 | 0.655 | 1.401 | 1.694 |
| Collins $\lvert 2,0\rangle$ | 2.305 | 7.375 | 6.560 | 6.966 | 5.846 |
| Collins $\lvert 2,1\rangle$ | 0.429 | 1.592 | 1.212 | 1.053 | 1.550 |
| Collins $\lvert 2,2\rangle$ | 1.011 | 2.119 | 1.801 | 0.436 | 0.546 |

Table A.8: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $K^+K^-$ dihadrons, Challenge B. See caption for Table A.1.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | **$M_h$** | **$M_h$-$x$** | **$M_h$-$y$** | **$M_h$-$z$** | **$M_h$-$P_{h\perp}$** |
| Sivers $\|0,0\rangle$ | 0.377 | 0.188 | 0.104 | 0.995 | 0.206 |
| Sivers $\|1,-1\rangle$ | 0.292 | 0.061 | 0.062 | 0.301 | 0.067 |
| Sivers $\|1,0\rangle$ | 1.522 | 1.801 | 1.214 | 3.540 | 1.037 |
| Sivers $\|1,1\rangle$ | 0.131 | 0.028 | 0.036 | 0.148 | 0.035 |
| Sivers $\|2,-2\rangle$ | 0.083 | 0.014 | 0.019 | 0.149 | 0.065 |
| Sivers $\|2,-1\rangle$ | 2.475 | 0.598 | 0.498 | 2.323 | 0.686 |
| Sivers $\|2,0\rangle$ | 4.789 | 2.588 | 0.985 | 3.675 | 3.804 |
| Sivers $\|2,1\rangle$ | 0.142 | 0.048 | 0.040 | 0.543 | 0.098 |
| Sivers $\|2,2\rangle$ | 0.015 | 0.018 | 0.010 | 0.044 | 0.022 |
| Collins $\|0,0\rangle$ | 0.073 | 0.147 | 0.060 | 0.163 | 0.102 |
| Collins $\|1,-1\rangle$ | 0.042 | 0.026 | 0.012 | 0.020 | 0.018 |
| Collins $\|1,0\rangle$ | 5.567 | 3.441 | 2.799 | 2.629 | 5.550 |
| Collins $\|1,1\rangle$ | 0.052 | 0.021 | 0.026 | 0.059 | 0.017 |
| Collins $\|2,-2\rangle$ | 0.039 | 0.011 | 0.026 | 0.022 | 0.020 |
| Collins $\|2,-1\rangle$ | 0.021 | 0.027 | 0.023 | 0.192 | 0.037 |
| Collins $\|2,0\rangle$ | 0.595 | 1.113 | 0.574 | 3.365 | 1.958 |
| Collins $\|2,1\rangle$ | 0.068 | 0.058 | 0.029 | 0.198 | 0.036 |
| Collins $\|2,2\rangle$ | 0.041 | 0.013 | 0.019 | 0.059 | 0.016 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 0.690 | 0.533 | 0.342 | 0.963 | 0.720 |

Table A.9: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^0$ dihadrons, Study 1. See caption for Table A.1. An overall $\chi^2/ndf$ is provided for each binning option, though only moments with a $\chi^2/ndf < 9$ for the 1D $M_h$ binning is included in the overall statistic.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 0.319 | 0.062 | 0.096 | 0.094 | 0.108 |
| Sivers $\|1,-1\rangle$ | 13.728 | 2.750 | 2.433 | 4.484 | 3.031 |
| Sivers $\|1,0\rangle$ | 27.803 | 7.201 | 6.789 | 11.875 | 6.472 |
| Sivers $\|1,1\rangle$ | 0.214 | 0.025 | 0.039 | 0.108 | 0.057 |
| Sivers $\|2,-2\rangle$ | 0.031 | 0.017 | 0.023 | 0.011 | 0.028 |
| Sivers $\|2,-1\rangle$ | 0.118 | 0.024 | 0.073 | 0.031 | 0.066 |
| Sivers $\|2,0\rangle$ | 0.344 | 0.529 | 1.668 | 1.009 | 1.209 |
| Sivers $\|2,1\rangle$ | 0.062 | 0.032 | 0.029 | 0.037 | 0.050 |
| Sivers $\|2,2\rangle$ | 0.005 | 0.011 | 0.016 | 0.011 | 0.005 |
| Collins $\|0,0\rangle$ | 0.035 | 0.040 | 0.059 | 0.088 | 0.055 |
| Collins $\|1,-1\rangle$ | 0.022 | 0.017 | 0.024 | 0.014 | 0.025 |
| Collins $\|1,0\rangle$ | 0.242 | 0.132 | 0.958 | 0.084 | 1.382 |
| Collins $\|1,1\rangle$ | 0.049 | 0.020 | 0.038 | 0.013 | 0.019 |
| Collins $\|2,-2\rangle$ | 0.009 | 0.008 | 0.008 | 0.016 | 0.008 |
| Collins $\|2,-1\rangle$ | 0.010 | 0.025 | 0.017 | 0.037 | 0.017 |
| Collins $\|2,0\rangle$ | 1.125 | 0.812 | 1.237 | 0.468 | 1.497 |
| Collins $\|2,1\rangle$ | 0.045 | 0.031 | 0.059 | 0.254 | 0.056 |
| Collins $\|2,2\rangle$ | 0.037 | 0.013 | 0.012 | 0.017 | 0.018 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 0.127 | 0.106 | 0.256 | 0.135 | 0.271 |

Table A.10: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^-$ dihadrons, Study 1. See caption for Tables A.1 and A.9.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h}$-$\boldsymbol{x}$ | $\boldsymbol{M_h}$-$\boldsymbol{y}$ | $\boldsymbol{M_h}$-$\boldsymbol{z}$ | $\boldsymbol{M_h}$-$\boldsymbol{P_{h\perp}}$ |
| Sivers $\|0,0\rangle$ | 0.983 | 0.281 | 0.231 | 0.254 | 0.402 |
| Sivers $\|1,-1\rangle$ | 0.083 | 0.028 | 0.020 | 0.082 | 0.052 |
| Sivers $\|1,0\rangle$ | 5.886 | 2.682 | 2.082 | 3.432 | 1.528 |
| Sivers $\|1,1\rangle$ | 0.023 | 0.039 | 0.021 | 0.025 | 0.032 |
| Sivers $\|2,-2\rangle$ | 0.045 | 0.018 | 0.016 | 0.063 | 0.024 |
| Sivers $\|2,-1\rangle$ | 3.393 | 0.520 | 0.630 | 2.250 | 1.054 |
| Sivers $\|2,0\rangle$ | 7.528 | 3.138 | 2.946 | 2.354 | 1.849 |
| Sivers $\|2,1\rangle$ | 0.084 | 0.058 | 0.049 | 0.273 | 0.064 |
| Sivers $\|2,2\rangle$ | 0.021 | 0.019 | 0.014 | 0.020 | 0.023 |
| Collins $\|0,0\rangle$ | 0.151 | 0.093 | 0.076 | 0.178 | 0.154 |
| Collins $\|1,-1\rangle$ | 0.041 | 0.021 | 0.018 | 0.032 | 0.027 |
| Collins $\|1,0\rangle$ | 0.749 | 2.155 | 0.744 | 1.135 | 1.638 |
| Collins $\|1,1\rangle$ | 0.081 | 0.026 | 0.026 | 0.038 | 0.031 |
| Collins $\|2,-2\rangle$ | 0.039 | 0.017 | 0.017 | 0.015 | 0.020 |
| Collins $\|2,-1\rangle$ | 0.050 | 0.049 | 0.048 | 0.055 | 0.021 |
| Collins $\|2,0\rangle$ | 1.699 | 1.076 | 0.718 | 1.647 | 1.508 |
| Collins $\|2,1\rangle$ | 0.018 | 0.021 | 0.033 | 0.224 | 0.041 |
| Collins $\|2,2\rangle$ | 0.014 | 0.007 | 0.014 | 0.010 | 0.022 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 0.883 | 0.536 | 0.403 | 0.632 | 0.444 |

Table A.11: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^-\pi^0$ dihadrons, Study 1. See caption for Tables A.1 and A.9.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h}$-$\boldsymbol{x}$ | $\boldsymbol{M_h}$-$\boldsymbol{y}$ | $\boldsymbol{M_h}$-$\boldsymbol{z}$ | $\boldsymbol{M_h}$-$\boldsymbol{P_{h\perp}}$ |
| Sivers $\vert 0,0\rangle$ | 0.014 | 0.010 | 0.018 | 0.008 | 0.007 |
| Sivers $\vert 1,-1\rangle$ | 0.010 | 0.003 | 0.015 | 0.004 | 0.007 |
| Sivers $\vert 1,0\rangle$ | 0.360 | 0.154 | 0.238 | 0.142 | 0.067 |
| Sivers $\vert 1,1\rangle$ | 0.017 | 0.008 | 0.014 | 0.014 | 0.006 |
| Sivers $\vert 2,-2\rangle$ | 0.003 | 0.011 | 0.023 | 0.016 | 0.015 |
| Sivers $\vert 2,-1\rangle$ | 0.022 | 0.016 | 0.034 | 0.020 | 0.025 |
| Sivers $\vert 2,0\rangle$ | 0.005 | 0.104 | 0.927 | 0.462 | 0.146 |
| Sivers $\vert 2,1\rangle$ | 0.018 | 0.019 | 0.067 | 0.009 | 0.009 |
| Sivers $\vert 2,2\rangle$ | 0.027 | 0.028 | 0.023 | 0.015 | 0.010 |
| Collins $\vert 0,0\rangle$ | 0.015 | 0.018 | 0.036 | 0.010 | 0.013 |
| Collins $\vert 1,-1\rangle$ | 0.015 | 0.017 | 0.007 | 0.011 | 0.010 |
| Collins $\vert 1,0\rangle$ | 0.017 | 0.027 | 0.022 | 0.057 | 0.029 |
| Collins $\vert 1,1\rangle$ | 0.022 | 0.009 | 0.007 | 0.014 | 0.008 |
| Collins $\vert 2,-2\rangle$ | 0.013 | 0.009 | 0.013 | 0.013 | 0.012 |
| Collins $\vert 2,-1\rangle$ | 0.005 | 0.012 | 0.023 | 0.016 | 0.009 |
| Collins $\vert 2,0\rangle$ | 0.067 | 0.055 | 0.093 | 0.073 | 0.042 |
| Collins $\vert 2,1\rangle$ | 0.037 | 0.022 | 0.037 | 0.016 | 0.013 |
| Collins $\vert 2,2\rangle$ | 0.043 | 0.023 | 0.044 | 0.044 | 0.037 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 0.027 | 0.028 | 0.084 | 0.048 | 0.024 |

Table A.12: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $K^+K^-$ dihadrons, Study 1. See caption for Tables A.1 and A.9.

| | **$\chi^2/ndf$ per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | **$M_h$** | **$M_h$-$x$** | **$M_h$-$y$** | **$M_h$-$z$** | **$M_h$-$P_{h\perp}$** |
| Sivers $\|0,0\rangle$ | 17.591 | 7.096 | 7.334 | 2.969 | 10.694 |
| Sivers $\|1,-1\rangle$ | 0.051 | 0.404 | 0.597 | 0.288 | 0.466 |
| Sivers $\|1,0\rangle$ | 40.876 | 52.202 | 52.818 | 25.205 | 40.600 |
| Sivers $\|1,1\rangle$ | 0.495 | 0.637 | 1.323 | 0.865 | 1.033 |
| Sivers $\|2,-2\rangle$ | 2.103 | 0.613 | 0.541 | 0.546 | 1.103 |
| Sivers $\|2,-1\rangle$ | 15.068 | 3.275 | 5.699 | 4.385 | 4.765 |
| Sivers $\|2,0\rangle$ | 91.219 | 81.814 | 74.223 | 32.666 | 105.408 |
| Sivers $\|2,1\rangle$ | 2.320 | 2.824 | 2.776 | 1.372 | 3.607 |
| Sivers $\|2,2\rangle$ | 0.535 | 0.514 | 0.335 | 0.256 | 0.810 |
| Collins $\|0,0\rangle$ | 5.401 | 5.215 | 4.460 | 2.854 | 8.523 |
| Collins $\|1,-1\rangle$ | 0.263 | 0.776 | 0.653 | 0.863 | 0.431 |
| Collins $\|1,0\rangle$ | 13.891 | 45.293 | 51.312 | 21.691 | 23.983 |
| Collins $\|1,1\rangle$ | 1.291 | 0.770 | 0.803 | 0.551 | 0.669 |
| Collins $\|2,-2\rangle$ | 1.019 | 0.529 | 0.518 | 0.469 | 0.603 |
| Collins $\|2,-1\rangle$ | 4.734 | 2.183 | 1.942 | 2.593 | 2.850 |
| Collins $\|2,0\rangle$ | 29.137 | 45.527 | 39.443 | 25.770 | 60.356 |
| Collins $\|2,1\rangle$ | 1.594 | 1.324 | 2.896 | 1.296 | 3.115 |
| Collins $\|2,2\rangle$ | 1.291 | 0.601 | 0.656 | 0.712 | 1.126 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 1.347 | 1.287 | 1.374 | 0.995 | 1.911 |

Table A.13: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^0$ dihadrons, Study 2. See caption for Tables A.1 and A.9.

| | $\chi^2/ndf$ per Binning Option | | | | |
|---|---|---|---|---|---|
| **Moment** | $M_h$ | $M_h$-$x$ | $M_h$-$y$ | $M_h$-$z$ | $M_h$-$P_{h\perp}$ |
| Sivers $\|0,0\rangle$ | 41.586 | 9.109 | 7.231 | 6.785 | 18.924 |
| Sivers $\|1,-1\rangle$ | 0.963 | 0.514 | 0.761 | 0.516 | 0.840 |
| Sivers $\|1,0\rangle$ | 58.949 | 30.940 | 17.729 | 14.455 | 39.386 |
| Sivers $\|1,1\rangle$ | 1.622 | 0.846 | 0.644 | 0.363 | 0.873 |
| Sivers $\|2,-2\rangle$ | 2.927 | 0.895 | 0.678 | 0.804 | 1.505 |
| Sivers $\|2,-1\rangle$ | 7.955 | 3.291 | 4.217 | 1.250 | 6.070 |
| Sivers $\|2,0\rangle$ | 83.691 | 66.833 | 39.826 | 72.668 | 69.292 |
| Sivers $\|2,1\rangle$ | 1.996 | 1.100 | 1.042 | 1.560 | 2.063 |
| Sivers $\|2,2\rangle$ | 1.504 | 0.702 | 0.361 | 0.567 | 0.702 |
| Collins $\|0,0\rangle$ | 2.268 | 1.162 | 2.205 | 2.826 | 4.764 |
| Collins $\|1,-1\rangle$ | 0.906 | 0.528 | 0.627 | 0.284 | 0.532 |
| Collins $\|1,0\rangle$ | 18.955 | 15.265 | 6.044 | 7.677 | 31.230 |
| Collins $\|1,1\rangle$ | 0.582 | 0.326 | 0.354 | 0.374 | 0.455 |
| Collins $\|2,-2\rangle$ | 1.416 | 0.767 | 0.508 | 0.688 | 0.890 |
| Collins $\|2,-1\rangle$ | 1.111 | 1.147 | 1.216 | 0.891 | 1.609 |
| Collins $\|2,0\rangle$ | 43.253 | 9.824 | 48.356 | 20.935 | 69.267 |
| Collins $\|2,1\rangle$ | 2.829 | 1.676 | 2.668 | 3.247 | 2.377 |
| Collins $\|2,2\rangle$ | 2.260 | 0.526 | 0.618 | 0.747 | 0.895 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 1.667 | 0.977 | 1.152 | 1.023 | 1.708 |

Table A.14: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^+\pi^-$ dihadrons, Study 2. See caption for Tables A.1 and A.9.

| | $\boldsymbol{\chi^2/ndf}$ **per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{x}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{y}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{z}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{P_{h\perp}}$ |
| Sivers $\vert 0,0\rangle$ | 5.263 | 4.155 | 4.285 | 6.431 | 3.446 |
| Sivers $\vert 1,-1\rangle$ | 0.227 | 1.306 | 1.478 | 0.630 | 0.965 |
| Sivers $\vert 1,0\rangle$ | 22.580 | 37.623 | 51.008 | 45.214 | 38.451 |
| Sivers $\vert 1,1\rangle$ | 0.995 | 0.981 | 0.695 | 0.927 | 0.767 |
| Sivers $\vert 2,-2\rangle$ | 0.645 | 0.204 | 0.558 | 0.606 | 0.617 |
| Sivers $\vert 2,-1\rangle$ | 9.440 | 3.033 | 7.027 | 6.847 | 3.200 |
| Sivers $\vert 2,0\rangle$ | 61.082 | 59.477 | 68.170 | 55.123 | 67.931 |
| Sivers $\vert 2,1\rangle$ | 1.050 | 1.046 | 1.786 | 2.792 | 1.371 |
| Sivers $\vert 2,2\rangle$ | 0.424 | 0.539 | 0.520 | 0.321 | 0.675 |
| Collins $\vert 0,0\rangle$ | 3.232 | 2.697 | 3.634 | 3.501 | 4.339 |
| Collins $\vert 1,-1\rangle$ | 1.163 | 0.819 | 1.073 | 0.902 | 0.551 |
| Collins $\vert 1,0\rangle$ | 248.022 | 62.586 | 77.125 | 67.778 | 55.059 |
| Collins $\vert 1,1\rangle$ | 3.889 | 1.141 | 1.447 | 1.472 | 1.491 |
| Collins $\vert 2,-2\rangle$ | 0.181 | 0.353 | 0.257 | 0.736 | 0.469 |
| Collins $\vert 2,-1\rangle$ | 2.268 | 1.126 | 1.686 | 2.792 | 1.720 |
| Collins $\vert 2,0\rangle$ | 24.163 | 24.670 | 33.925 | 32.875 | 59.468 |
| Collins $\vert 2,1\rangle$ | 0.495 | 0.747 | 1.234 | 3.651 | 0.913 |
| Collins $\vert 2,2\rangle$ | 1.305 | 0.627 | 0.368 | 0.467 | 0.563 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 1.243 | 1.141 | 1.378 | 1.828 | 1.296 |

Table A.15: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $\pi^-\pi^0$ dihadrons, Study 2. See caption for Tables A.1 and A.9.

| | $\boldsymbol{\chi^2/ndf}$ **per Binning Option** | | | | |
|---|---|---|---|---|---|
| **Moment** | $\boldsymbol{M_h}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{x}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{y}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{z}$ | $\boldsymbol{M_h}$**-**$\boldsymbol{P_{h\perp}}$ |
| Sivers $\lvert 0,0\rangle$ | 2.973 | 0.565 | 0.468 | 0.744 | 0.653 |
| Sivers $\lvert 1,-1\rangle$ | 0.099 | 0.272 | 0.261 | 0.320 | 0.353 |
| Sivers $\lvert 1,0\rangle$ | 2.607 | 2.309 | 8.390 | 3.821 | 4.250 |
| Sivers $\lvert 1,1\rangle$ | 0.218 | 0.270 | 0.223 | 0.593 | 0.285 |
| Sivers $\lvert 2,-2\rangle$ | 2.244 | 0.562 | 1.019 | 0.661 | 0.661 |
| Sivers $\lvert 2,-1\rangle$ | 0.065 | 0.339 | 0.423 | 0.101 | 0.422 |
| Sivers $\lvert 2,0\rangle$ | 6.515 | 6.963 | 16.789 | 4.085 | 4.315 |
| Sivers $\lvert 2,1\rangle$ | 2.297 | 0.457 | 0.702 | 0.697 | 0.500 |
| Sivers $\lvert 2,2\rangle$ | 0.349 | 0.362 | 0.571 | 0.419 | 0.124 |
| Collins $\lvert 0,0\rangle$ | 0.473 | 0.252 | 0.161 | 0.280 | 0.256 |
| Collins $\lvert 1,-1\rangle$ | 0.340 | 0.267 | 0.546 | 0.619 | 0.303 |
| Collins $\lvert 1,0\rangle$ | 0.310 | 1.438 | 3.035 | 0.976 | 2.499 |
| Collins $\lvert 1,1\rangle$ | 0.471 | 0.125 | 0.500 | 0.375 | 0.193 |
| Collins $\lvert 2,-2\rangle$ | 0.240 | 0.191 | 0.182 | 0.412 | 0.166 |
| Collins $\lvert 2,-1\rangle$ | 1.893 | 0.442 | 0.375 | 0.633 | 0.453 |
| Collins $\lvert 2,0\rangle$ | 0.907 | 2.554 | 1.764 | 1.680 | 1.873 |
| Collins $\lvert 2,1\rangle$ | 0.559 | 0.249 | 0.463 | 0.263 | 0.324 |
| Collins $\lvert 2,2\rangle$ | 0.625 | 0.251 | 0.665 | 0.414 | 0.531 |
| Overall $\left(\chi^2_{M_h}/ndf < 9\right)$ | 0.875 | 0.914 | 1.869 | 0.875 | 0.929 |

Table A.16: Table of $\chi^2/ndf$ statistics per Sivers and Collins moments for $K^+K^-$ dihadrons, Study 2. See caption for Tables A.1 and A.9.

# APPENDIX B

# Kinematic Dependence of Systematic Uncertainties

Chapter V documents the estimation of three sources of systematic uncertainty, specifically uncertainty due to smearing and acceptance, year dependence (beam charge dependence), and the hadronic identification procedure. The full comparison of the contributions towards the total systematic uncertainty is contained in the figures in this appendix, at least for all moments for which final results are presented.

The results for $K^+K^-$ dihadrons are given in Figures B.1 and B.2. As the final results for $K^+K^-$ are only given with respect to the 1D $M_h$ binning, only the 1D binning systematics are shown in this appendix. The results for pion-pair dihadrons are given in Figures B.3 through B.26. Note, for pion-pair dihadrons, the results for the 1D and all 2D binning choices are shown on the same plot.

In general, the uncertainty related to the hadronic identification procedure is much smaller than the other systematic uncertainties in almost every case. The systematic uncertainty is generally dominated by the smearing/acceptance uncertainty, though there are many exceptions.

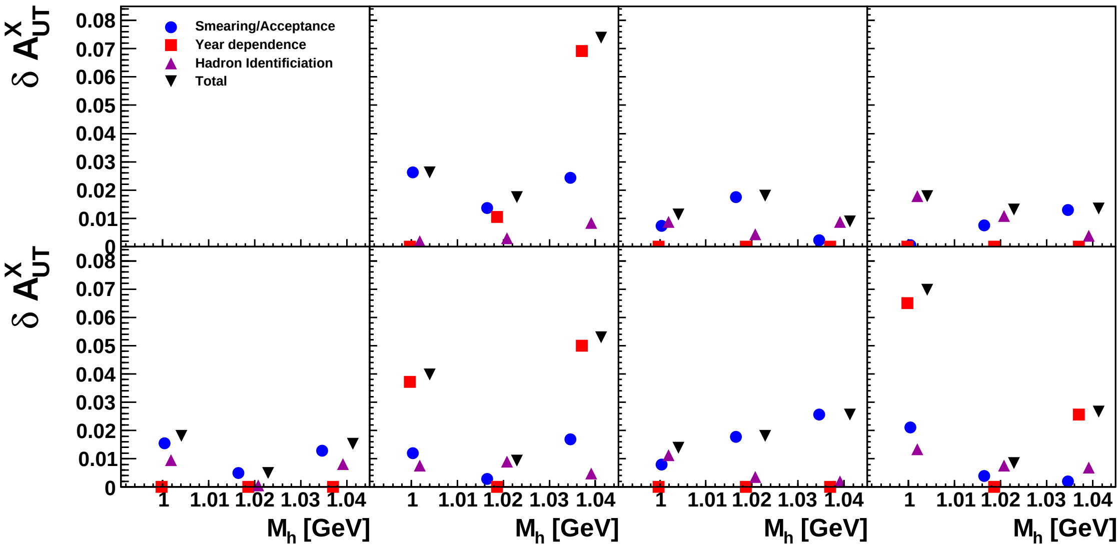


Figure B.1: Comparison of sources of systematic uncertainty for the Sivers moments for $K^+K^-$ dihadrons. The panels are arranged as in Figure 5.2. The upper left panel is intentionally left blank. The other panels on the top row are, from left to right, the $|0,0\rangle$, $|1,1\rangle$ and $|1,-1\rangle$ Sivers moments. The panels on the second row are, from left to right, the $|2,2\rangle$, $|2,1\rangle$, $|2,-1\rangle$ and $|2,2\rangle$ Sivers moments. The blue circles are for the uncertainty due to smearing and acceptance, the red squares for the year dependence, and the purple, upright triangles for the hadronic identification procedure. The total systematic uncertainty is shown with black, inverted triangles.

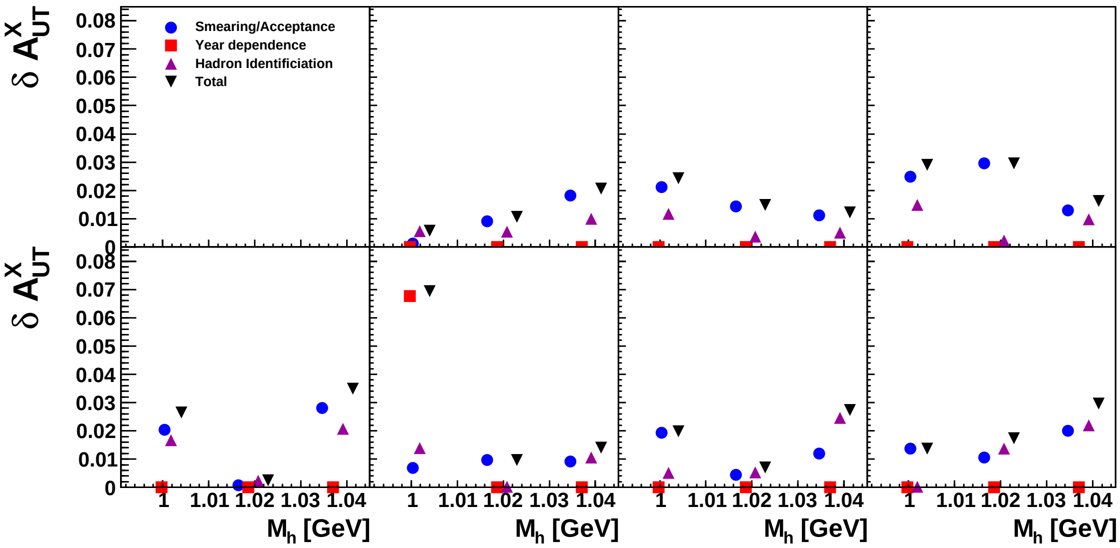


Figure B.2: Comparison of sources of systematic uncertainty for the Collins moments for $K^+K^-$ dihadrons. The markers are the same as in Figure B.1, as are the arrangement of the partial waves. Note, however, these are the Collins moments while Figure B.1 presents the Sivers moments.

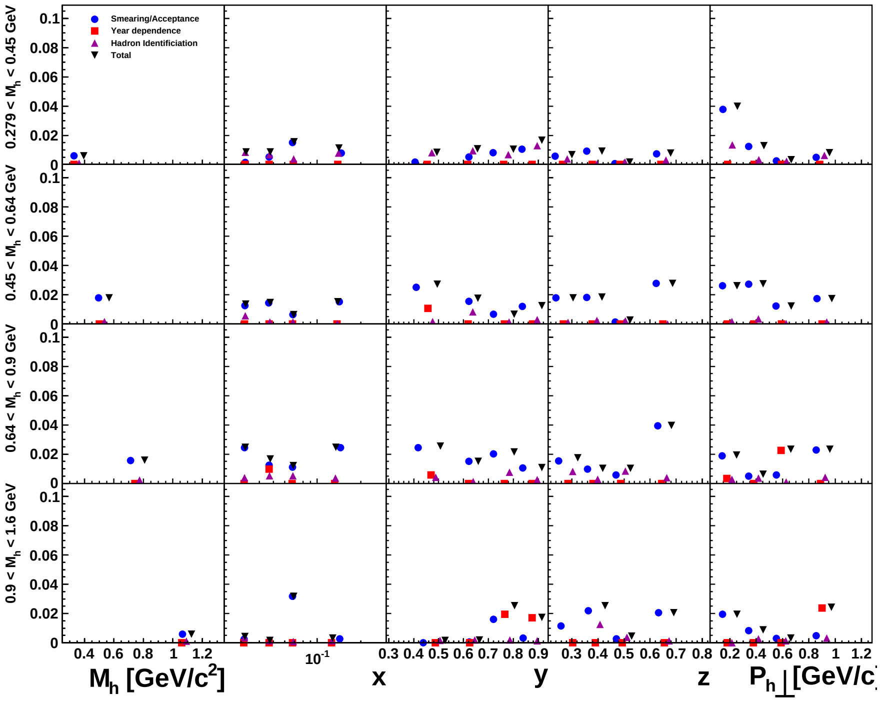


Figure B.3: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Sivers moments for $\pi^+\pi^0$ dihadrons. The panels are as in Figure 4.10, except the results versus $M_h$ are given in a new column on the far left. The markers are as in Figure B.1.

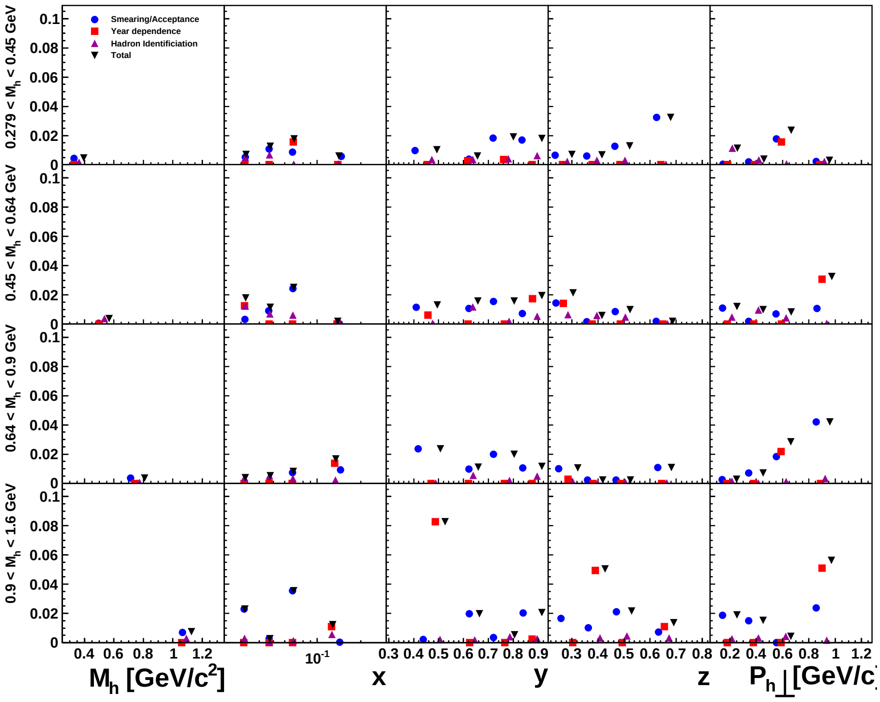


Figure B.4: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Sivers moments for $\pi^+\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

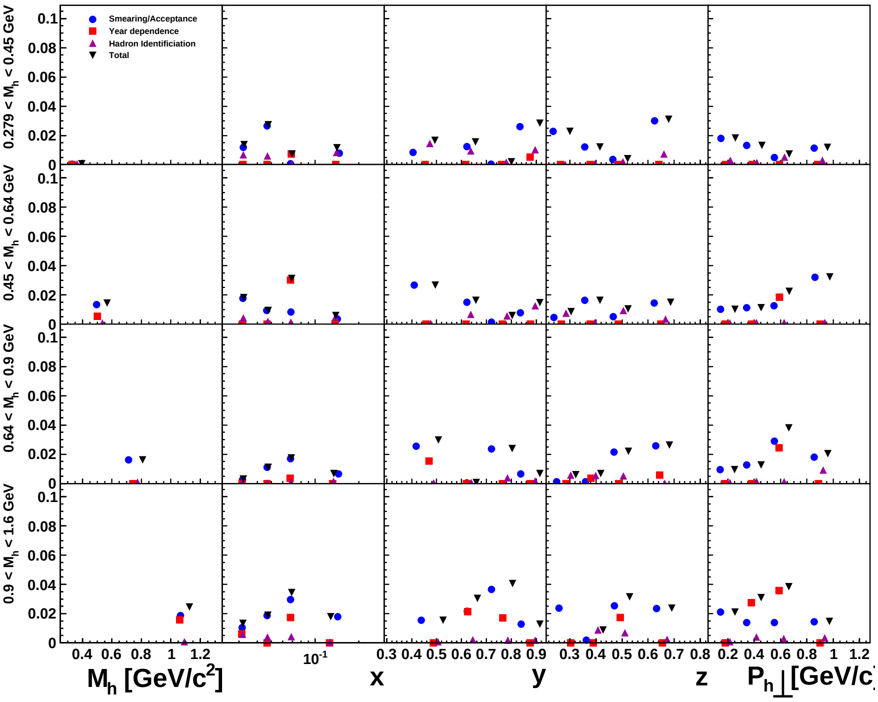


Figure B.5: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Sivers moments for $\pi^{+}\pi^{0}$ dihadrons. Markers and panels are as in Figure B.3.

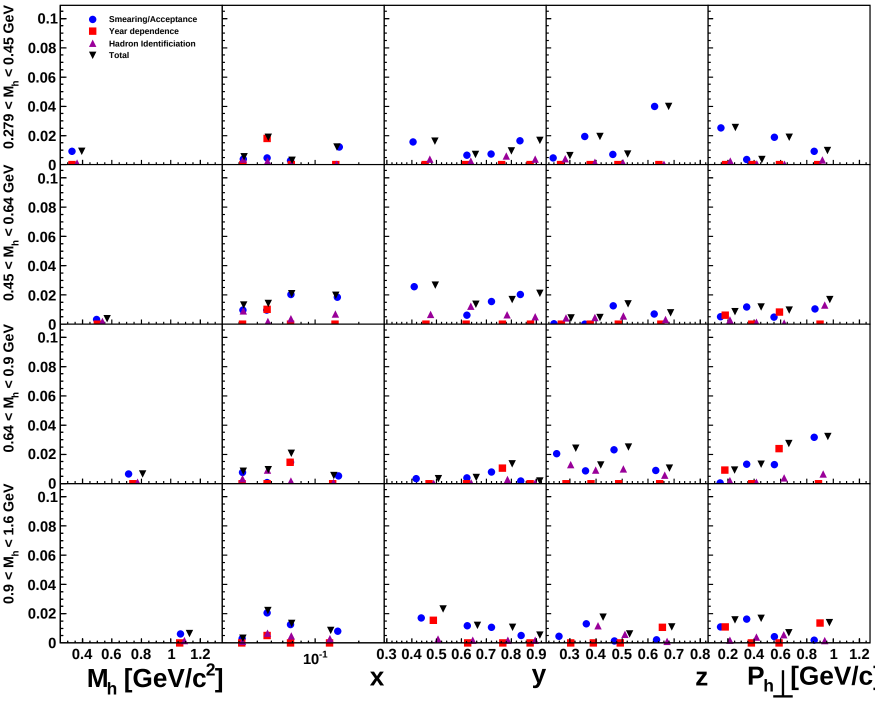


Figure B.6: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Sivers moments for $\pi^{+}\pi^{0}$ dihadrons. Markers and panels are as in Figure B.3.

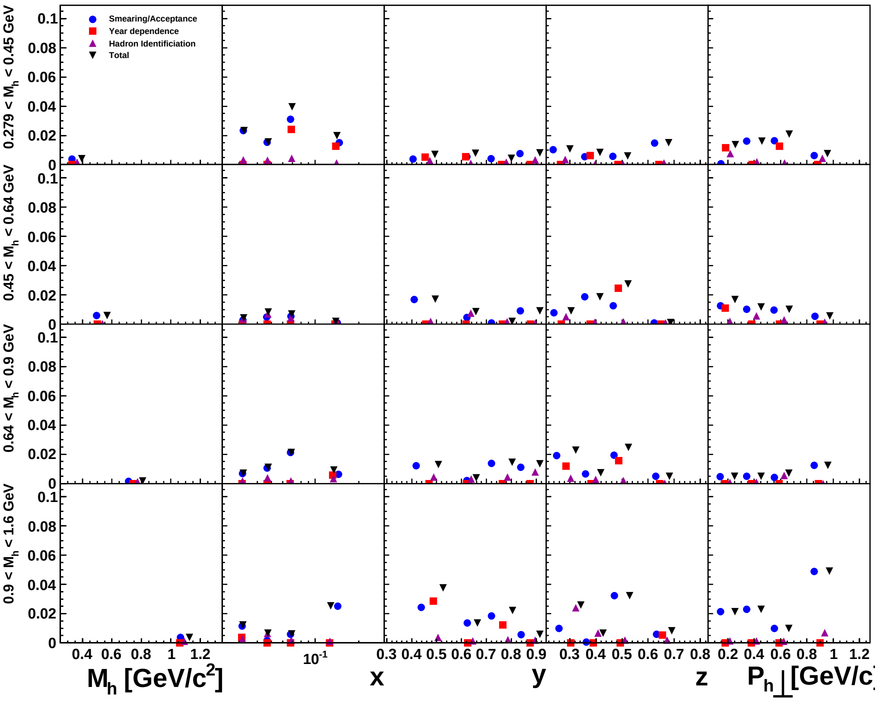


Figure B.7: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Collins moments for $\pi^+\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

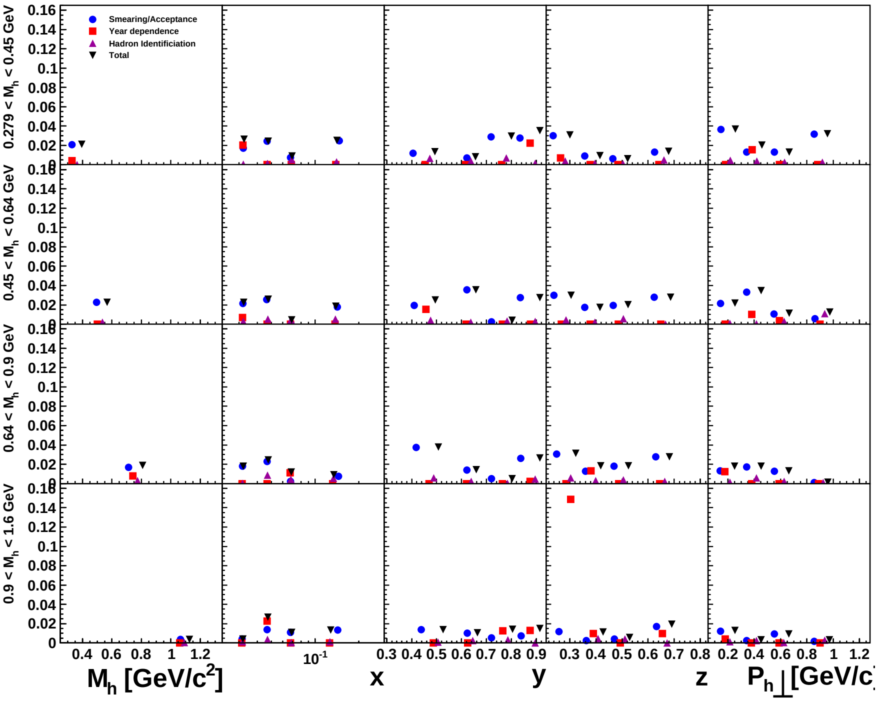


Figure B.8: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Collins moments for $\pi^+\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

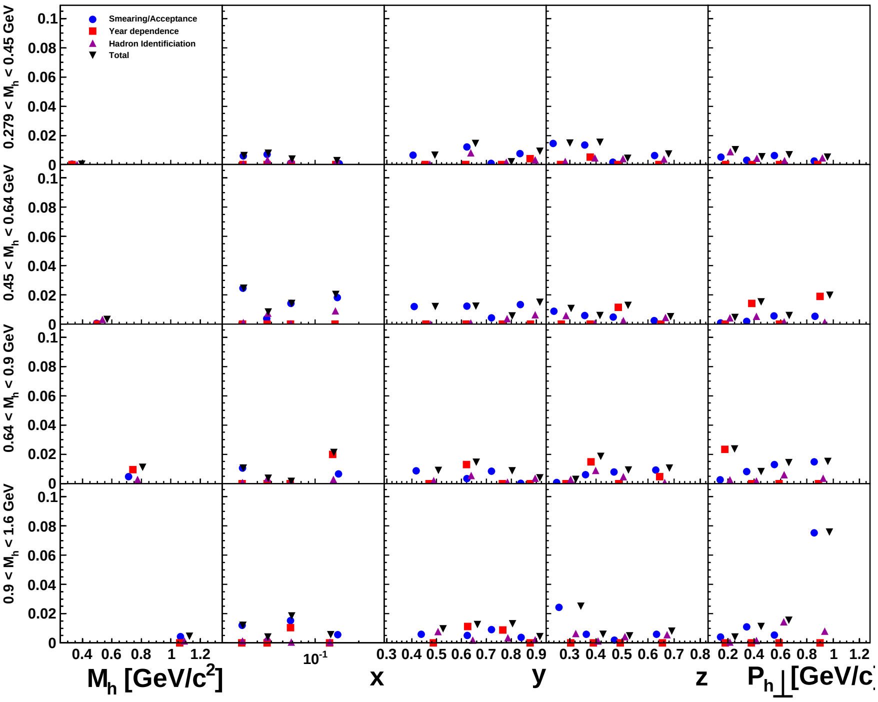


Figure B.9: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Collins moments for $\pi^+\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

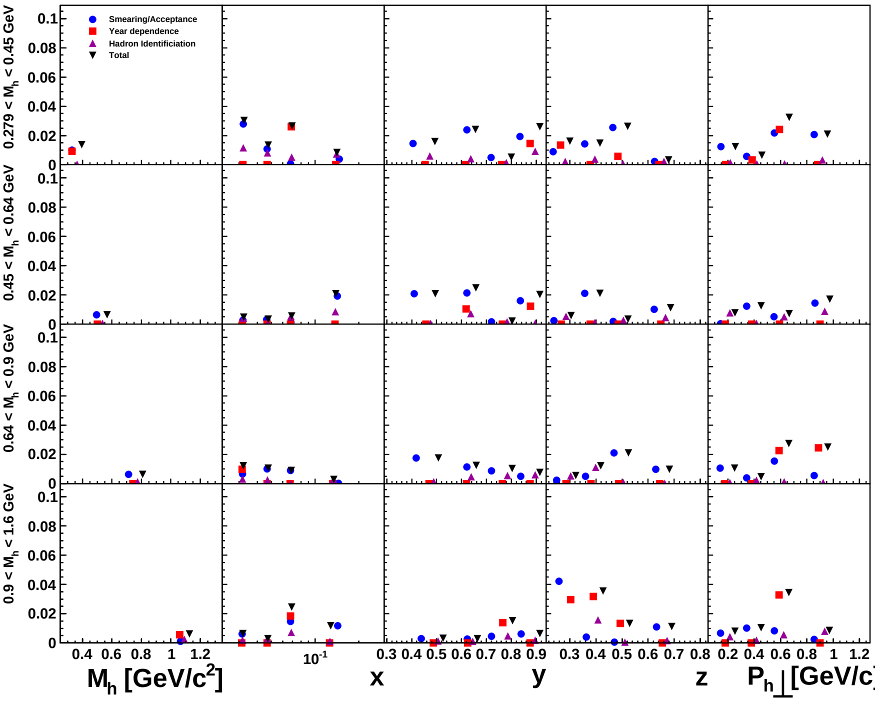


Figure B.10: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Collins moments for $\pi^+\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

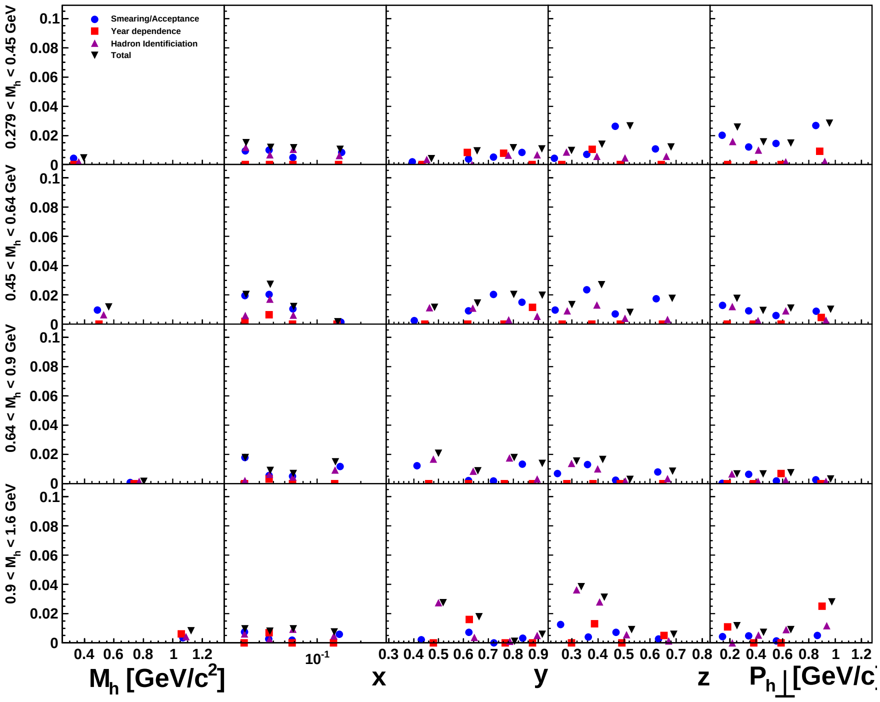


Figure B.11: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Sivers moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

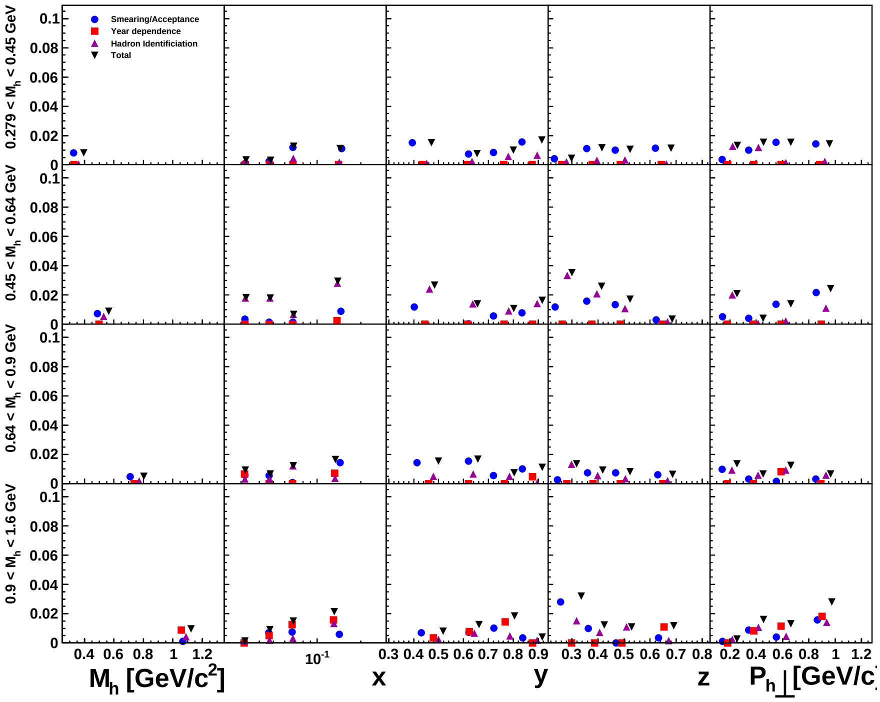


Figure B.12: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Sivers moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

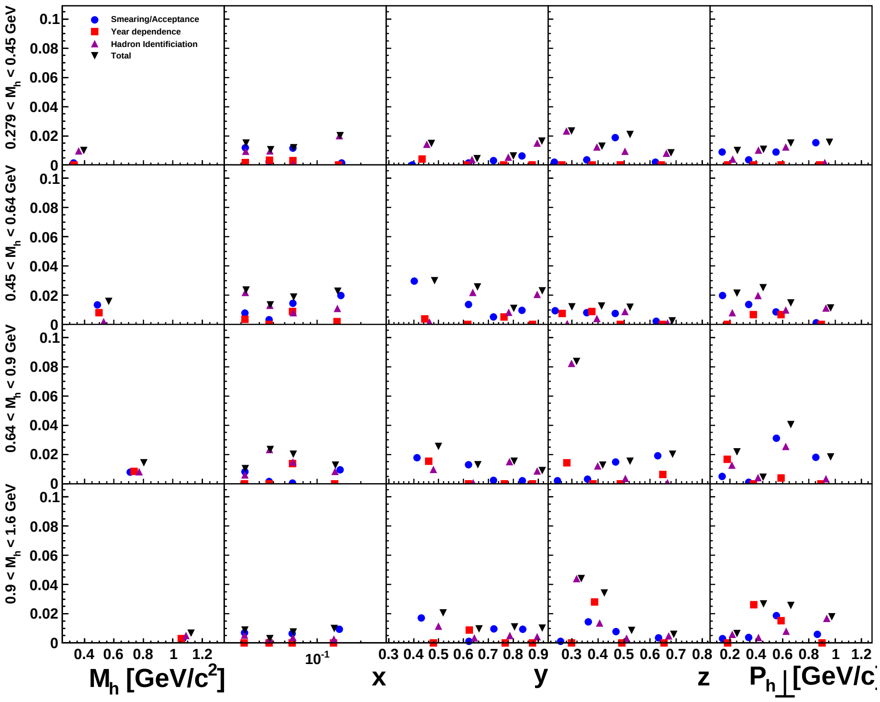


Figure B.13: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Sivers moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

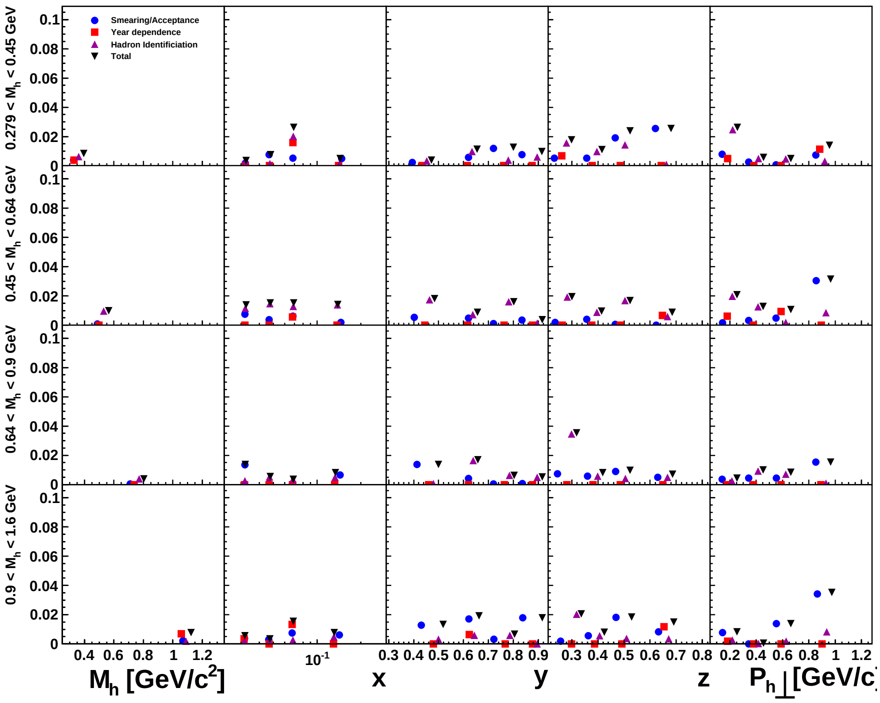


Figure B.14: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Sivers moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

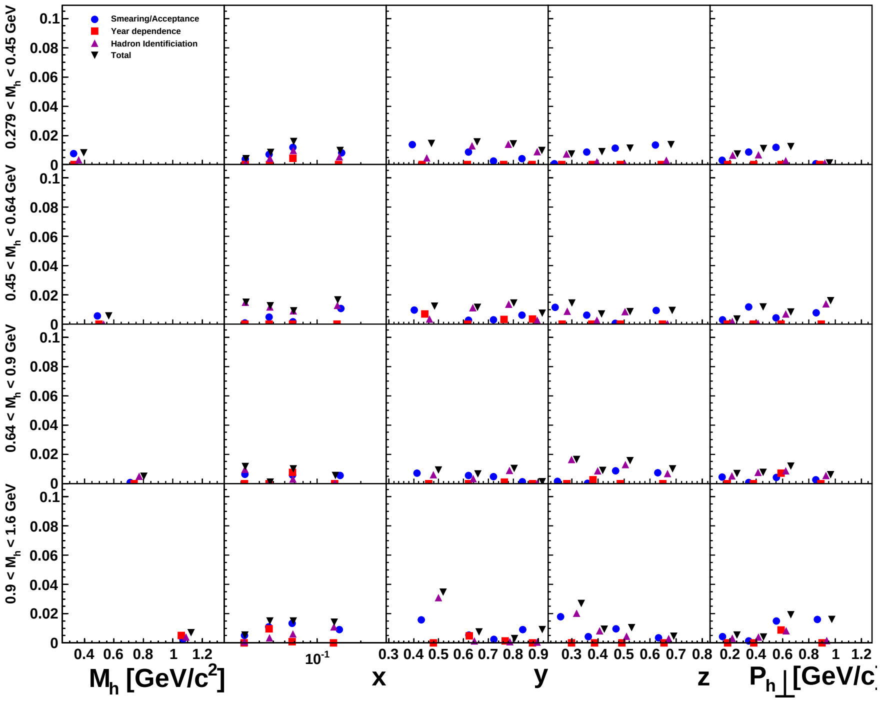


Figure B.15: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Collins moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

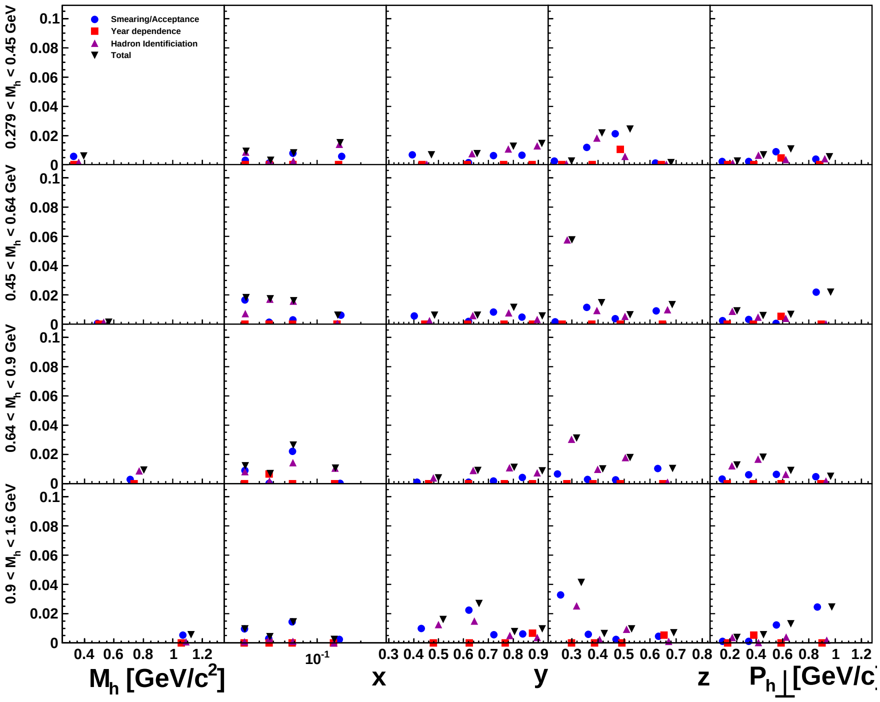


Figure B.16: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Collins moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

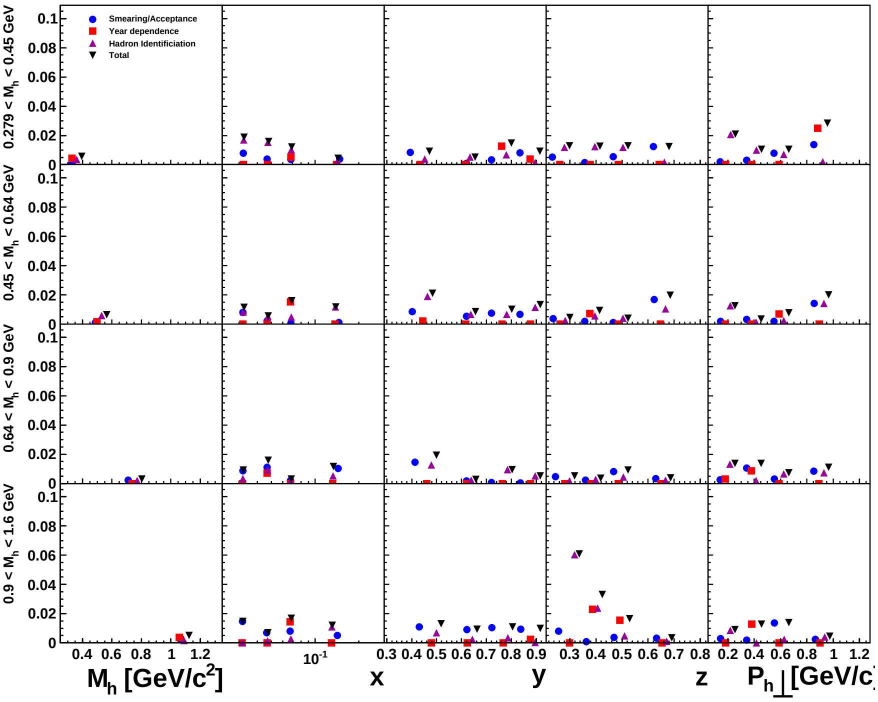


Figure B.17: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Collins moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

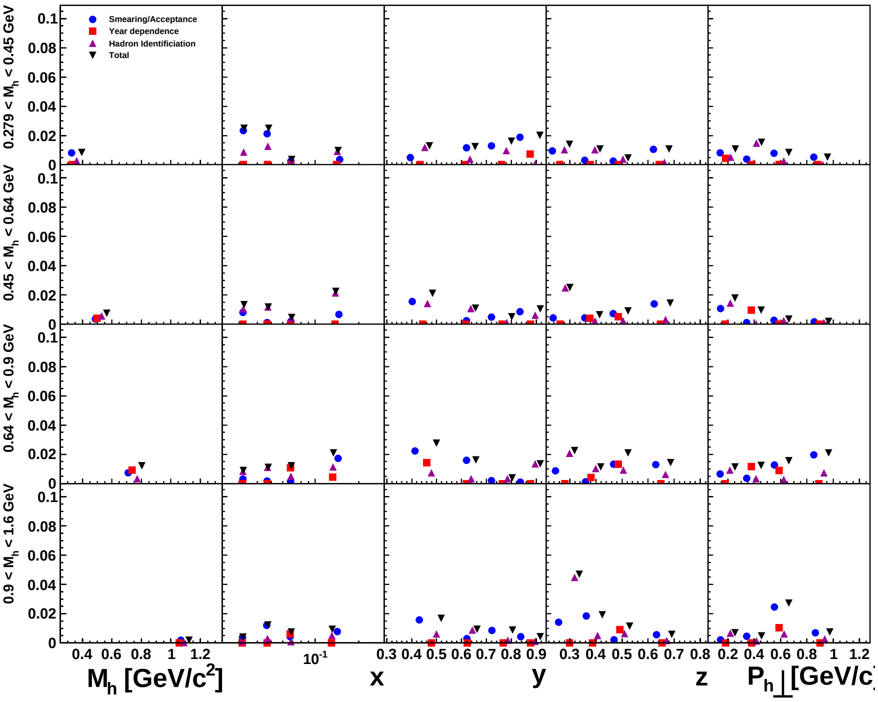


Figure B.18: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Collins moments for $\pi^+\pi^-$ dihadrons. Markers and panels are as in Figure B.3.

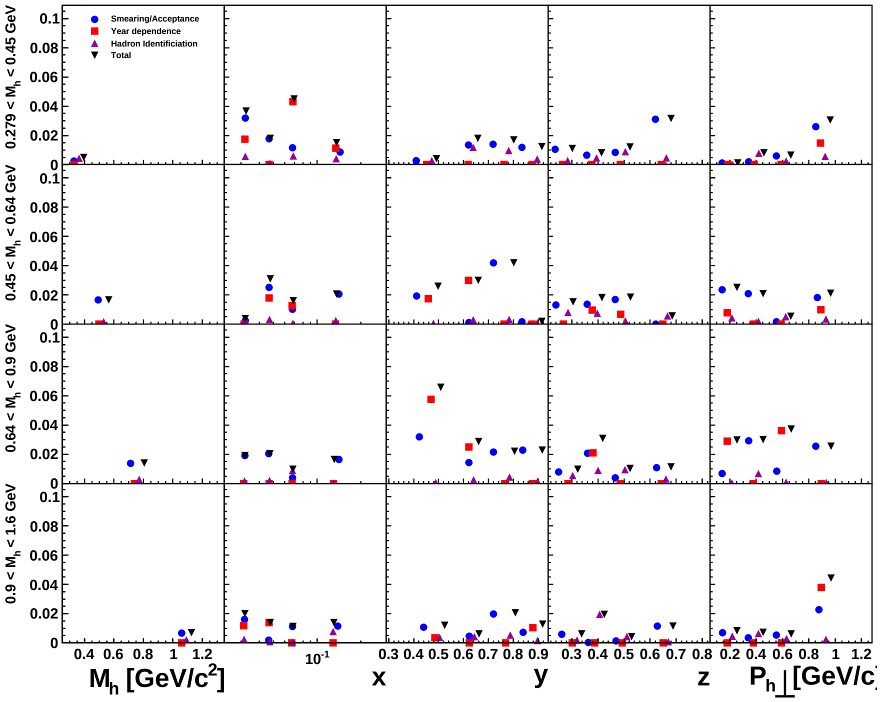


Figure B.19: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Sivers moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

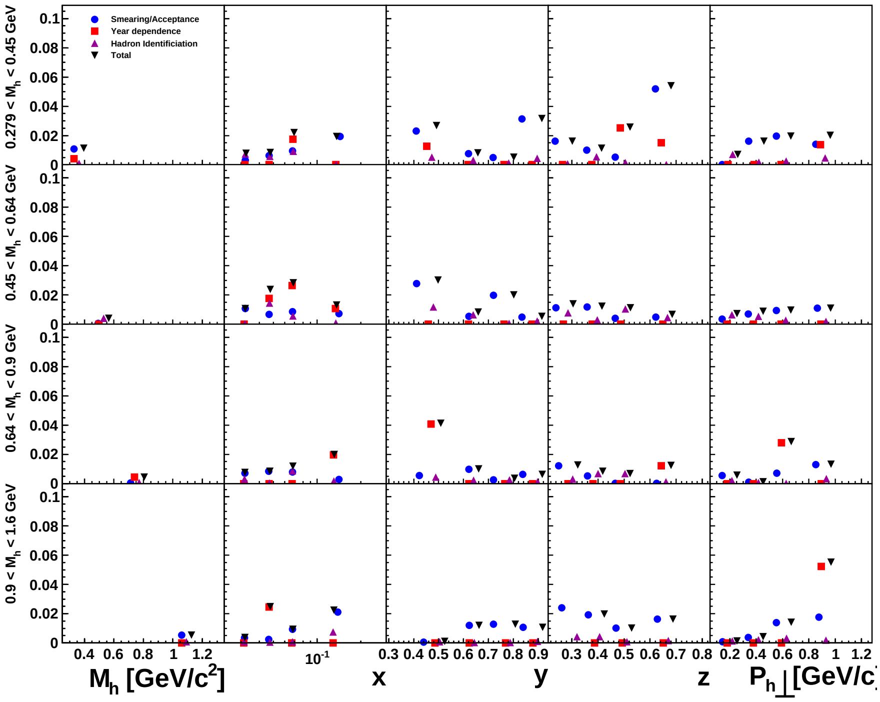


Figure B.20: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Sivers moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

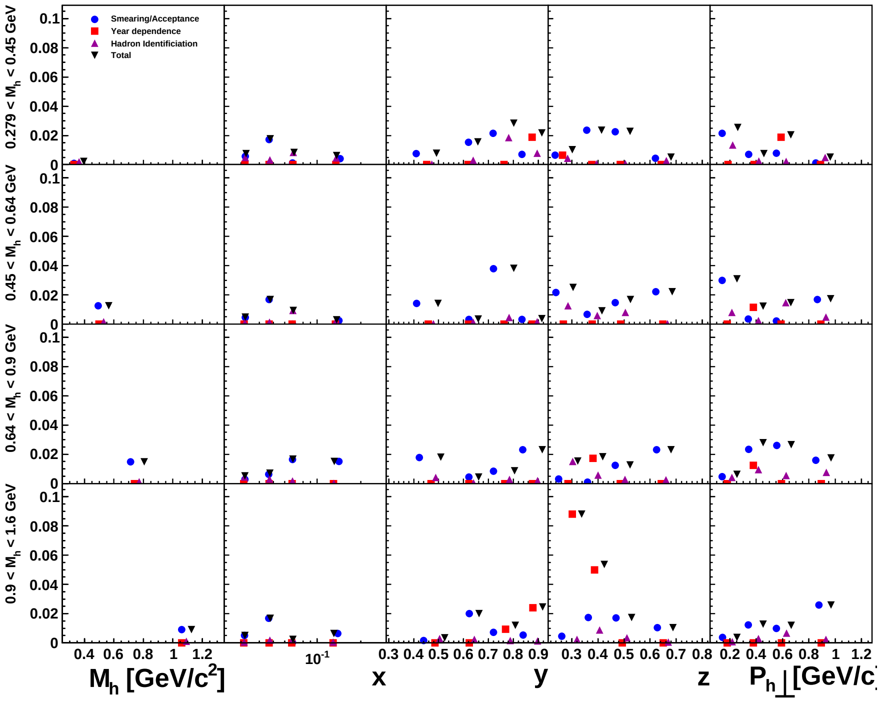


Figure B.21: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Sivers moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

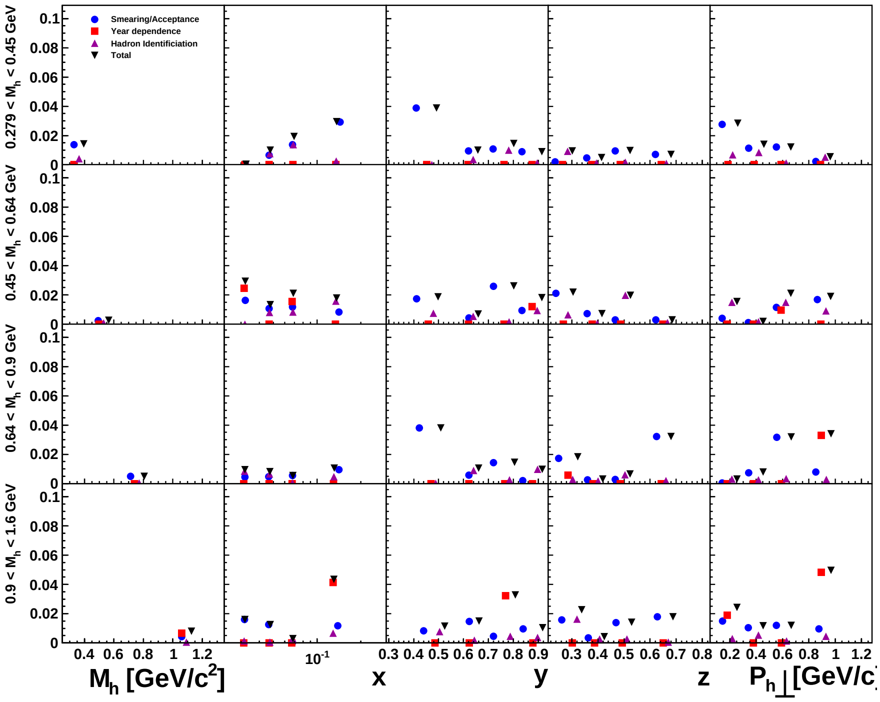


Figure B.22: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Sivers moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

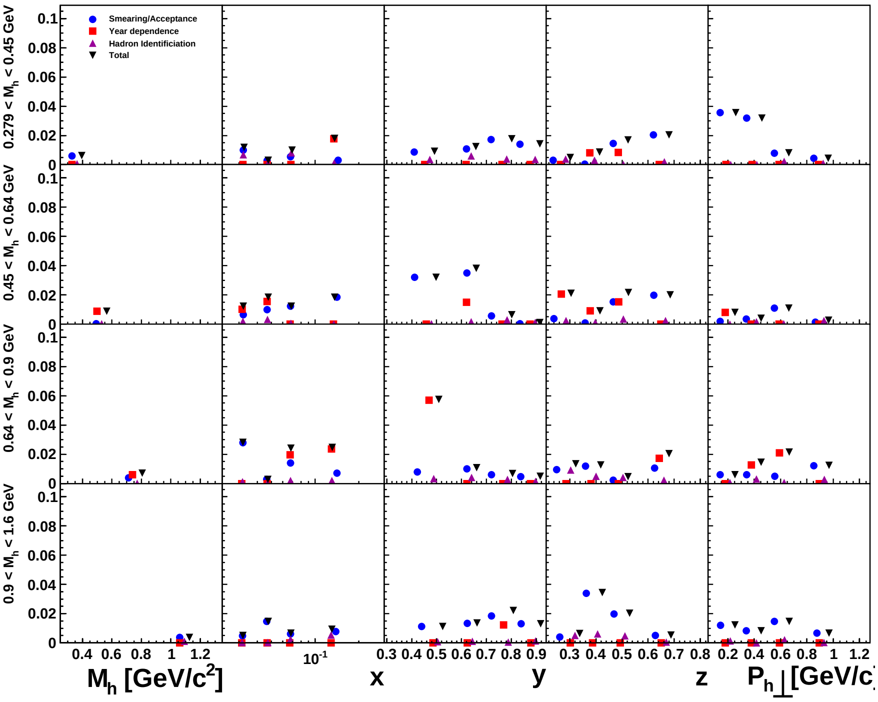


Figure B.23: Comparison of sources of systematic uncertainty for the $|1,-1\rangle$ Collins moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

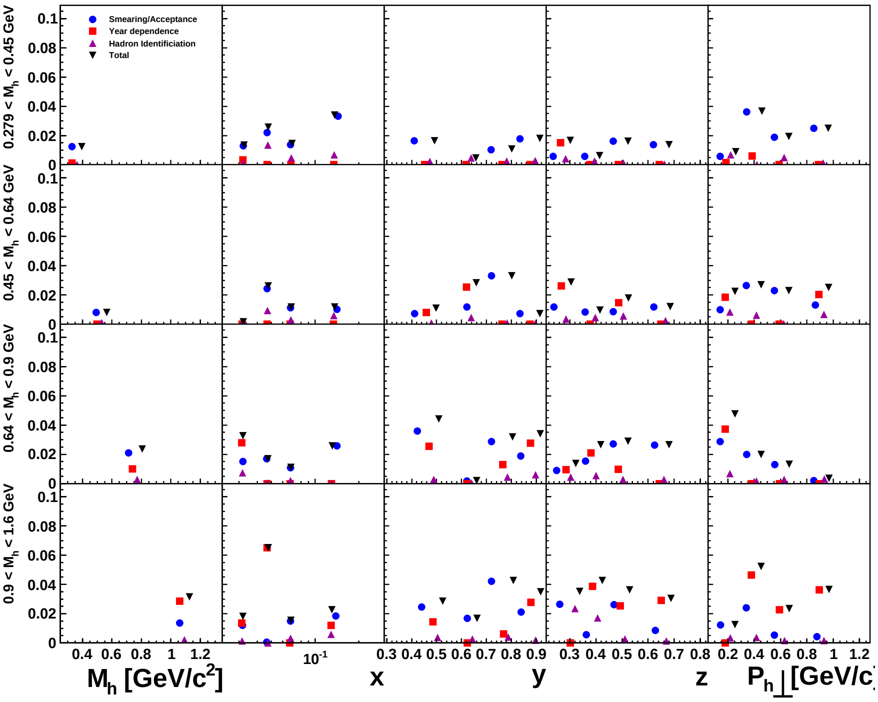


Figure B.24: Comparison of sources of systematic uncertainty for the $|1,1\rangle$ Collins moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

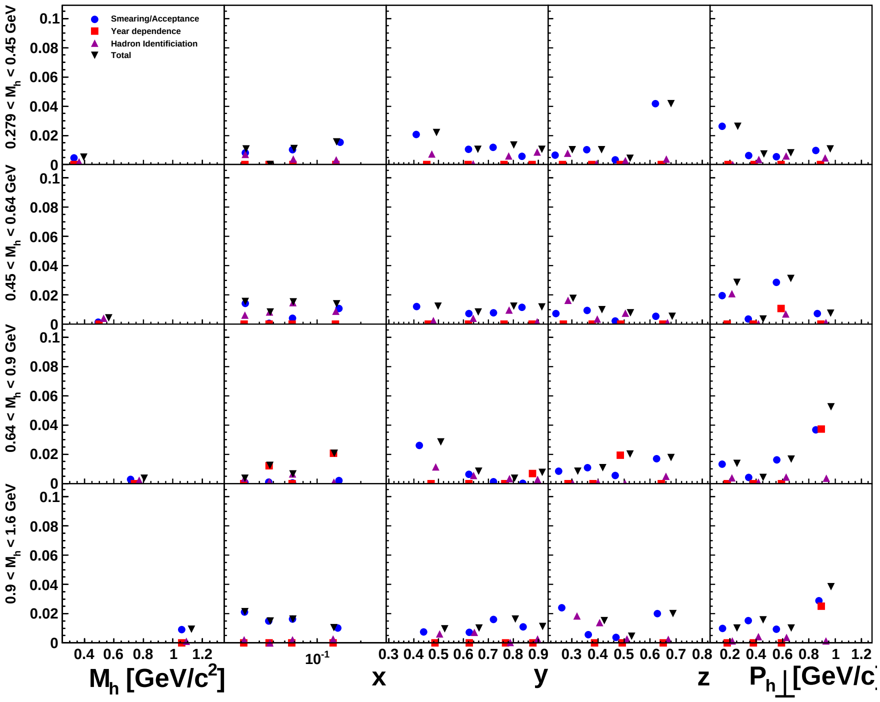


Figure B.25: Comparison of sources of systematic uncertainty for the $|2,-2\rangle$ Collins moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

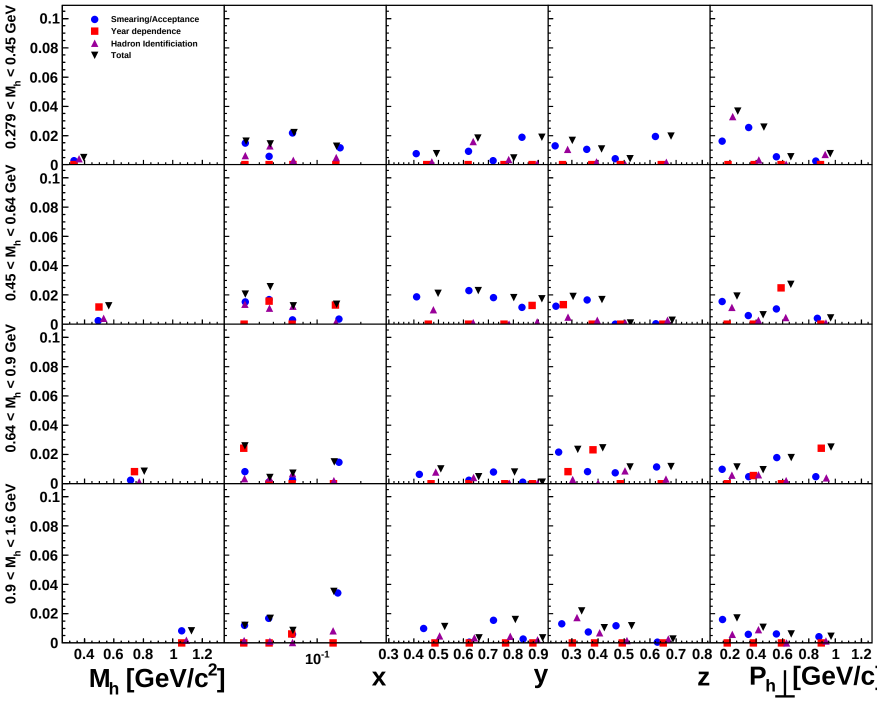


Figure B.26: Comparison of sources of systematic uncertainty for the $|2,2\rangle$ Collins moments for $\pi^-\pi^0$ dihadrons. Markers and panels are as in Figure B.3.

# APPENDIX C

# Additional Results

The moments of most theoretical interest are presented and discussed in Chapter VI. Additional moments, for all dihadron types considered, are provided in this appendix. Specifically, the Collins $|1,-1\rangle$ moment and Sivers $|1,\pm1\rangle$ and $|2,\pm2\rangle$ moments for pion-pair dihadrons are presented in Figures C.1 through C.5. Collins moments for $K^+K^-$ dihadron production are also provided in Figure C.6. As described in Chapter VI, these results are considered stable but their interpretation is more relevant in the context of a global fit rather than in the discussion of individual plots.

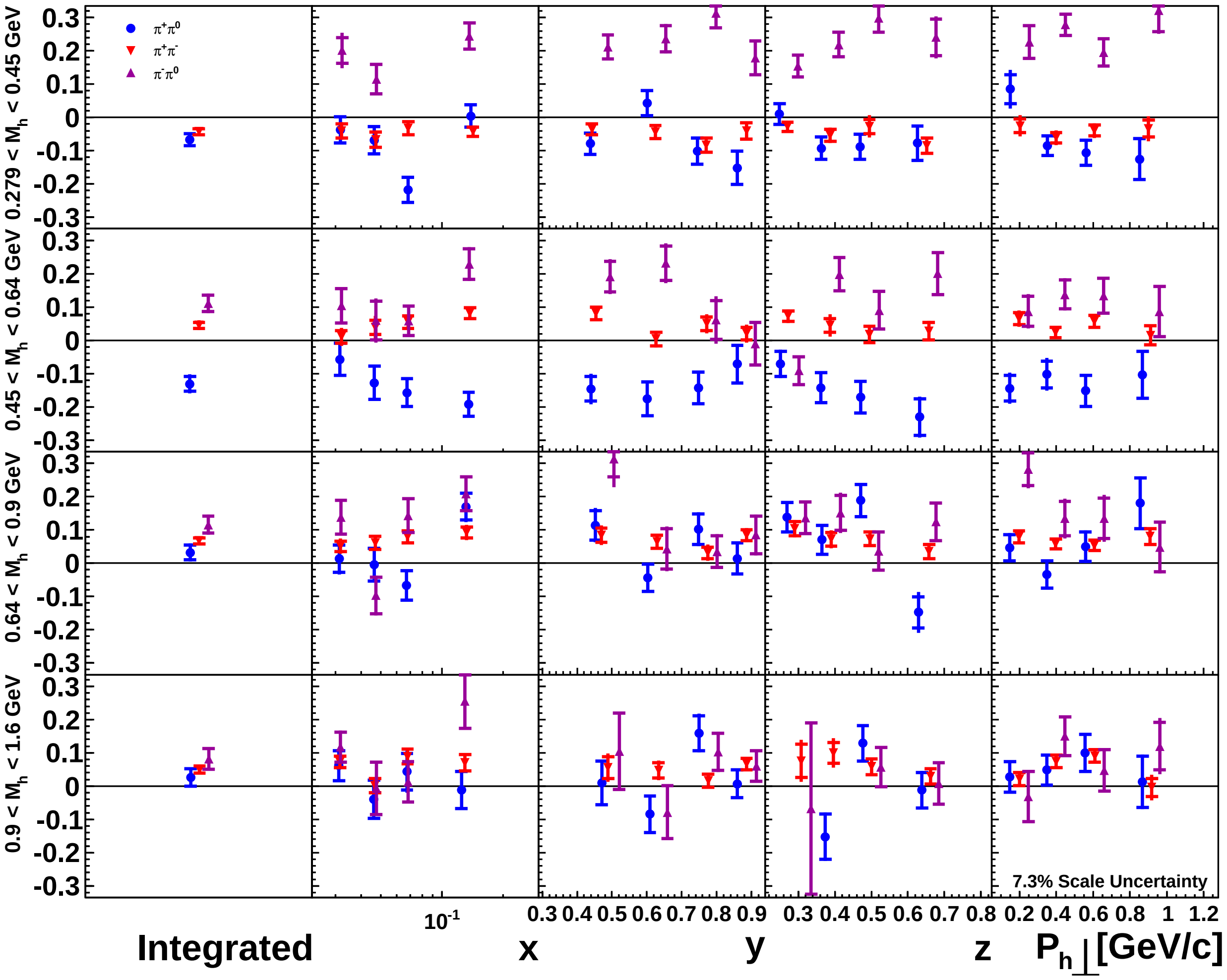


Figure C.1: Final Results for the Sivers $|1, -1\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

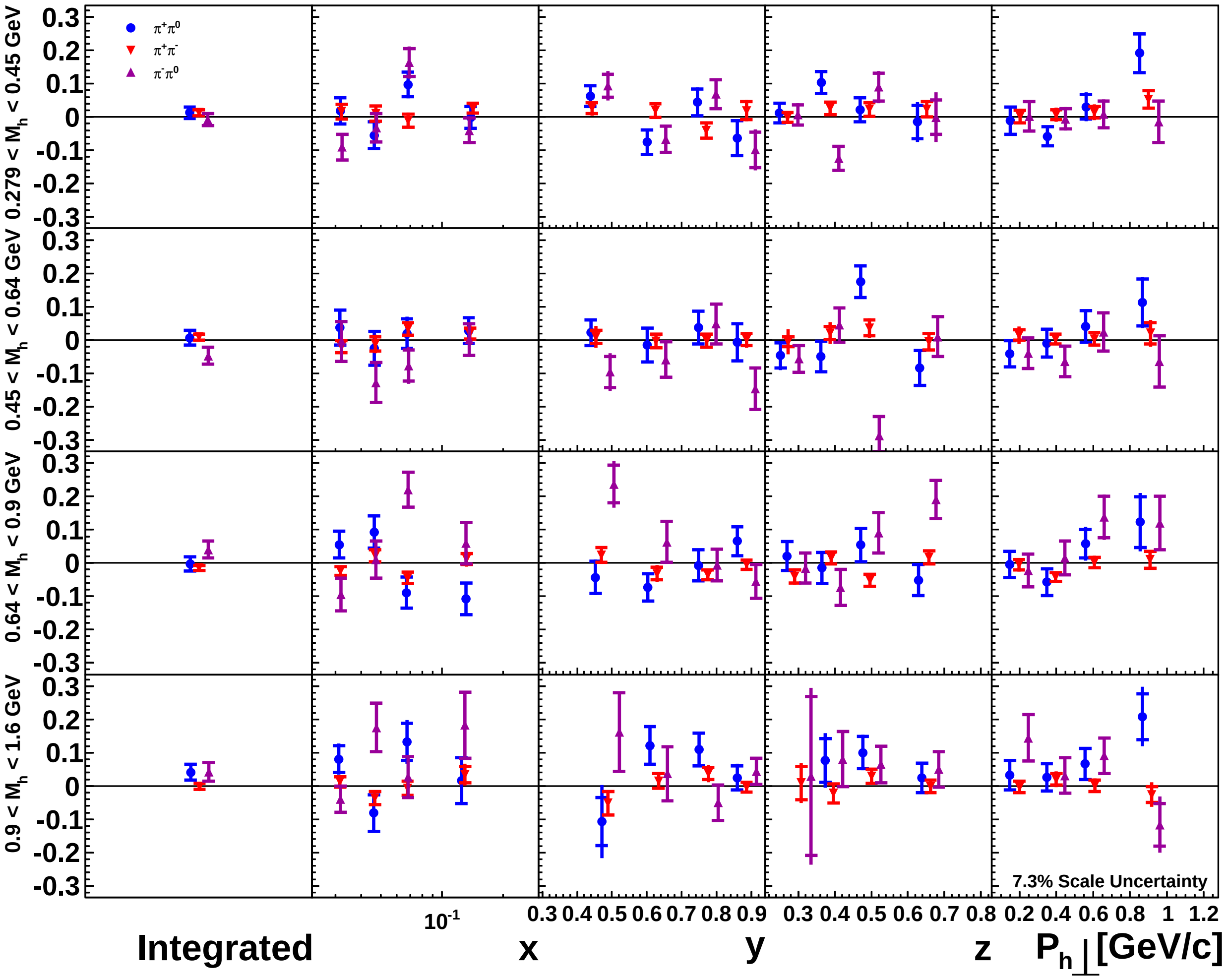


Figure C.2: Final Results for the Sivers $|1, 1\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

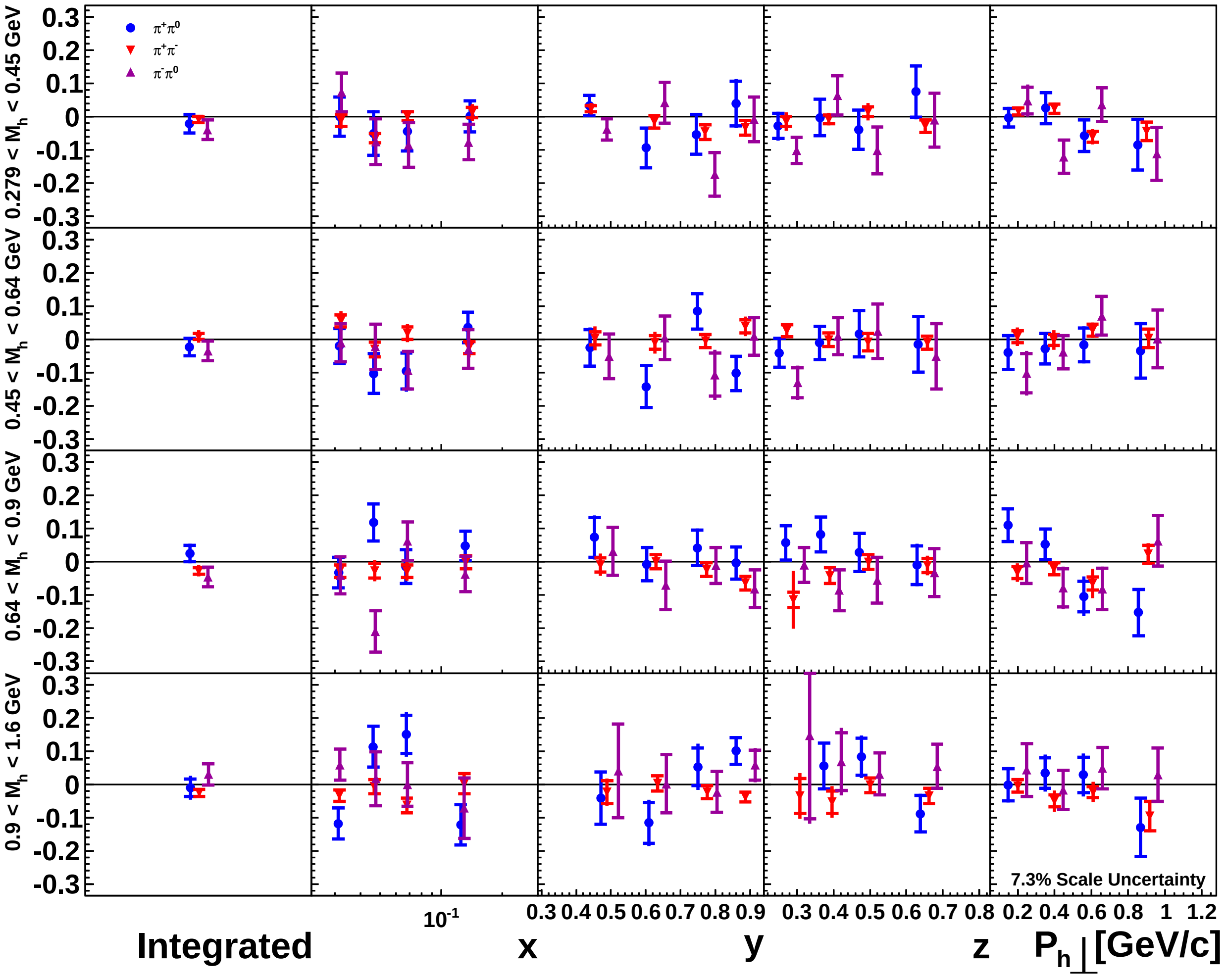


Figure C.3: Final Results for the Sivers $|2,-2\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

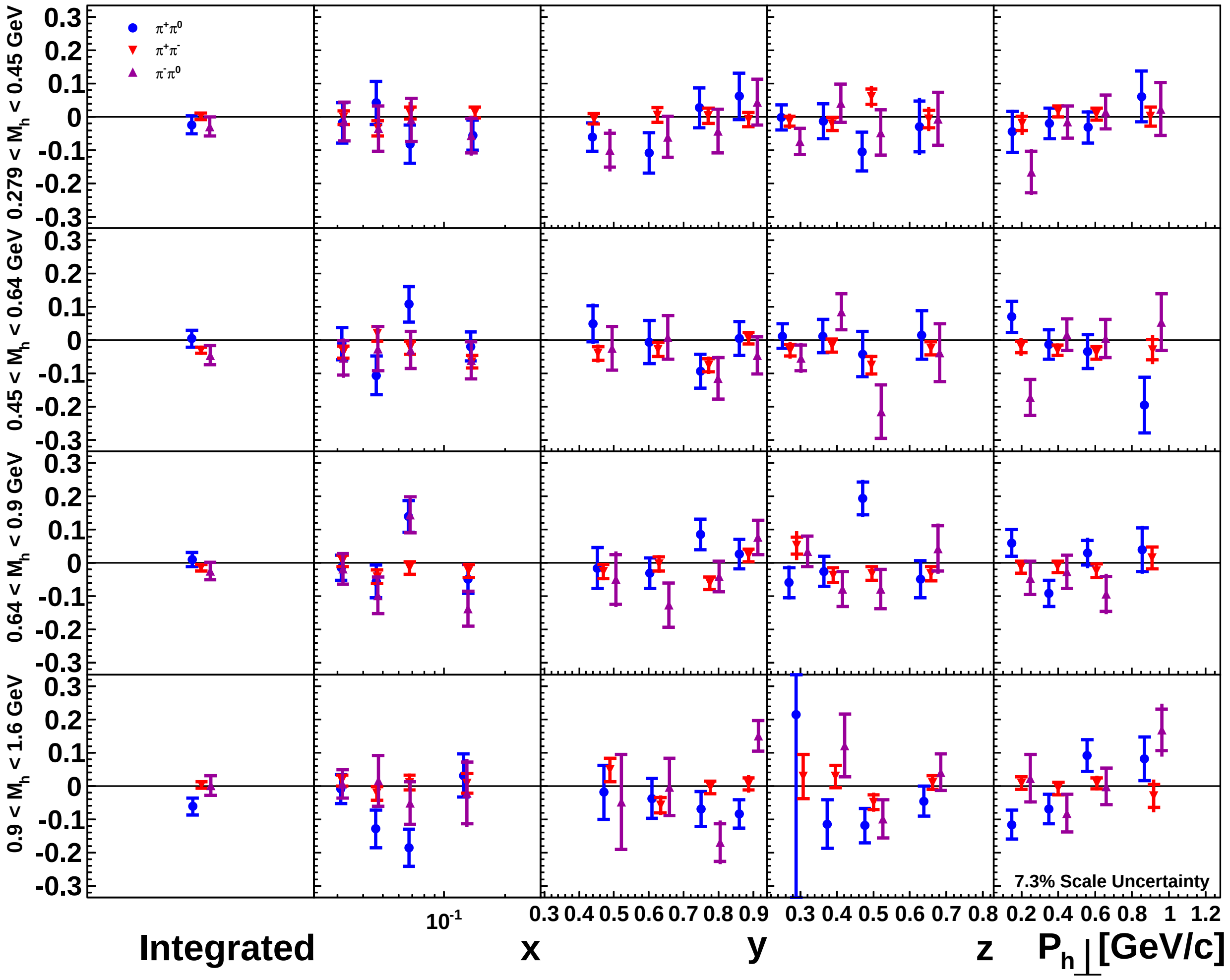


Figure C.4: Final Results for the Sivers $|2,2\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

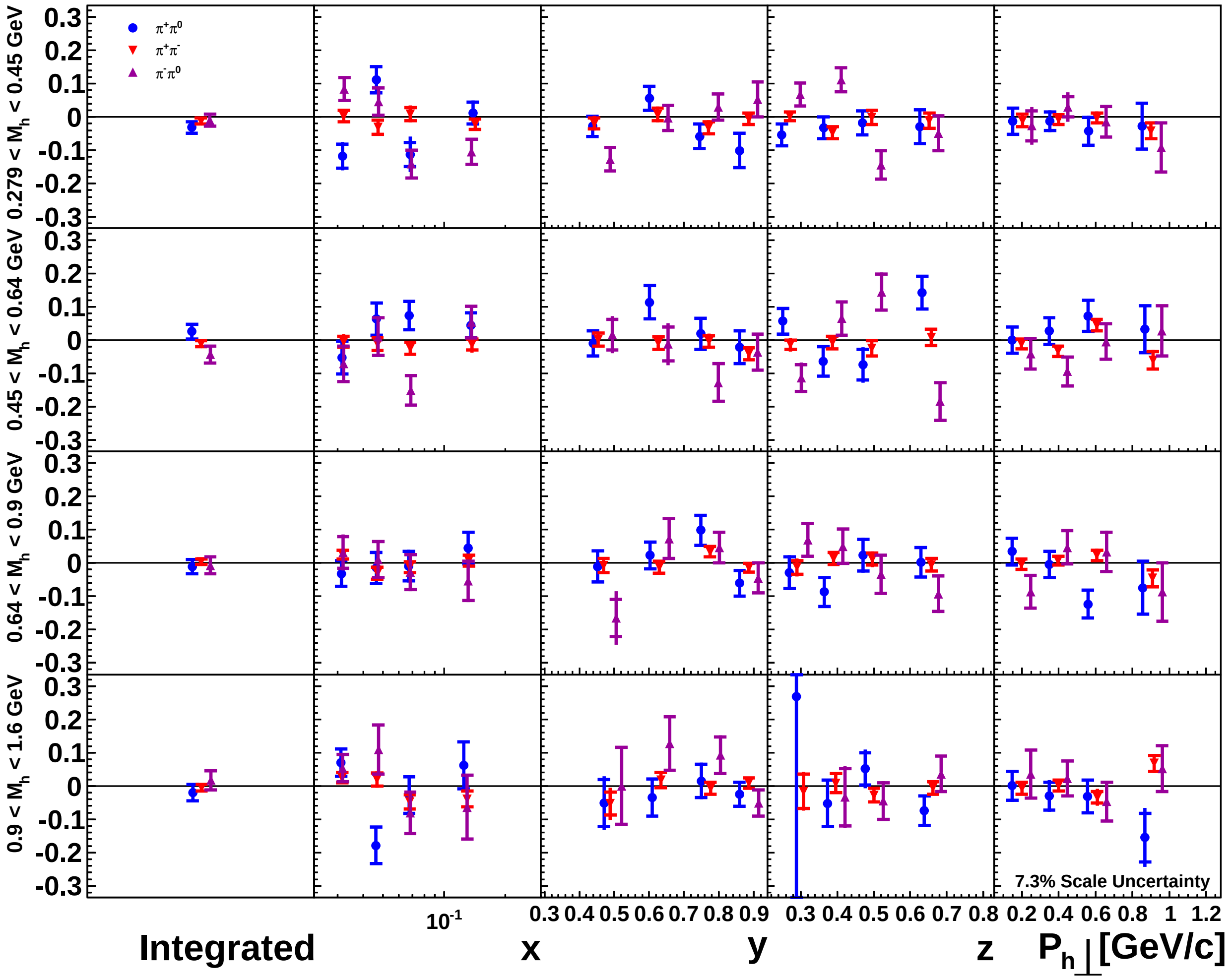


Figure C.5: Final Results for the Collins $|1, -1\rangle$ moments for pion-pair dihadrons, including kinematic dependencies. Panels and markers are arranged as in Figure 6.2.

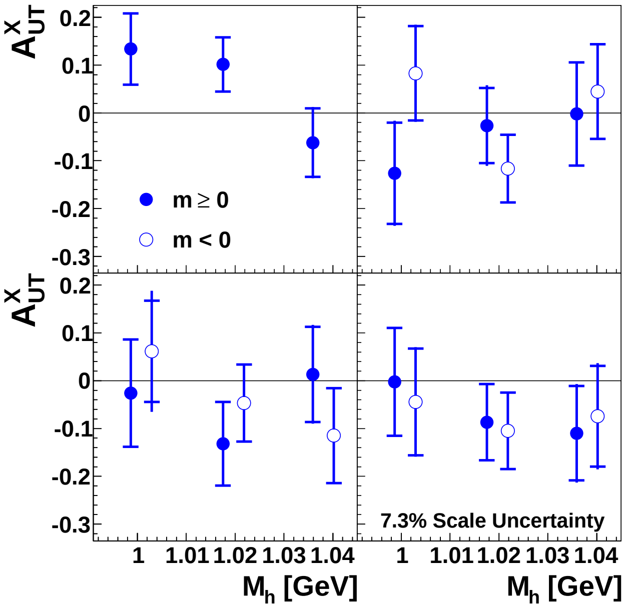


Figure C.6: Final Results versus $M_h$ for select Collins moments for $K^+K^-$ dihadrons. The panels are, top row, left to right, for $|0,0\rangle$, and $|1,\pm1\rangle$ partial waves, and the bottom panels are, left to right, for $|2,\pm1\rangle$ and $|2,\pm2\rangle$ partial waves. Filled circles indicate $m \geq 0$, and open circles $m < 0$.

# BIBLIOGRAPHY

# BIBLIOGRAPHY

[1] J. Ashman et al. (EMC Collaboration). A measurement of the spin asymmetry and determination of the structure function $g_1$ in deep inelastic muon-proton scattering. *Phys. Lett. B*, 206:364–370, 1988.

[2] A. Airapetian et al. (HERMES Collaboration). Precise determination of the spin structure function $g_1$ of the proton, deuteron and neutron. *Phys. Rev. D*, 75:012007, 2007.

[3] V. Y. Alexakhin et al. (COMPASS Collaboration). The deuteron spin-dependent structure function $g_1^d$ and its first moment. *Phys. Lett. B*, 647:8–17, 2007.

[4] D. de Florian et al. Global analysis of helicity parton densities and their uncertainties. *Phys. Rev. Lett.*, 101:072001, 2008.

[5] F. Myhrer and A.W. Thomas. A possible resolution of the proton spin problem. *Phys. Lett. B*, 663(4):302–305, 2008.

[6] A. W. Thomas. Interplay of spin and orbital angular momentum in the proton. *Phys. Rev. Lett.*, 101:102003, 2008.

[7] G. L. Kane, J. Pumplin, and W. Repko. Transverse quark polarization in large-$p_T$ reactions, $e^+e^-$ jets, and leptoproduction: A test of Quantum Chromodynamics. *Phys. Rev. Lett.*, 41(25):1689–1692, December 1978.

[8] D. Sivers. Single-spin production asymmetries from the hard scattering of pointlike constituents. *Phys. Rev. D*, 41(1):83–90, January 1990.

[9] J. C. Collins. Fragmentation of transversely polarized quarks probed in transverse momentum distributions. *Nucl. Phys. B*, 396:161–182, 1993.

[10] S. J. Brodsky, D. S. Hwang, and I. Schmidt. Final-state interactions and single-spin asymmetries in semi-inclusive deep inelastic scattering. *Phys. Lett. B*, 530:99–107, 2002.

[11] A. Airapetian et al. (HERMES Collaboration). Single-spin asymmetries in semi-inclusive deep-inelastic scattering on a transversely polarized hydrogen target. *Phys. Rev. Lett.*, 94:012002, 2005.

[12] D. Boer and P. J. Mulders. Time-reversal odd distribution functions in leptoproduction. *Phys. Rev. D*, 57(9):5780–5786, May 1998.

[13] R. N. Cahn. Azimuthal dependence in leptoproduction: A simple parton model calculation. *Phys. Lett. B*, 78:269, 1978.

[14] R. N. Cahn. Critique of parton-model calculations of azimuthal dependence in leptoproduction. *Phys. Rev. D*, 40:3107–3110, 1989.

[15] P. J. Mulders and R. D. Tangerman. The complete tree-level result up to order 1/Q for polarized deep-inelastic leptoproduction. *Nucl. Phys. B*, 461:197–237, 1996.

[16] A. Bacchetta et al. Semi-inclusive deep inelastic scattering at small transverse momentum. *J. High Energy Phys. (JHEP)*, 093, 2007.

[17] A. Bacchetta. *Probing the transverse spin of quarks in deep inelastic scattering.* PhD thesis, Vrije Universiteit, 2002.

[18] A. Bacchetta and M. Radici. Partial-wave analysis of two-hadron fragmentation functions. *Phys. Rev. D*, 67(9):094002, May 2003.

[19] V.N. Gribov and L.N. Lipatov. Deep inelastic *ep*-scattering in a perturbation theory. *Sov. J. Nucl. Phys.*, 15:438, 1972.

[20] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B*, 126(2):298–318, 1977.

[21] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and $e^+e^-$ annihilation by perturbation theory in Quantum Chromodynamics. *Sov.Phys. JETP*, 46:641, 1977.

[22] V.N. Gribov and L.N. Lipatov. $e^+e^-$ pair annihilation and deep inelastic *ep* scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:675–684, 1972.

[23] L.N. Lipatov. The parton model and perturbation theory. *Sov. J. Nucl. Phys.*, 20:94–102, 1975.

[24] J. P. Ralston and D. E. Soper. Production of dimuons from high-energy polarized proton-proton collisions. *Nucl. Phys. B*, 152:109–124, 1979.

[25] X. Artru and M. Mekhfi. Transversely polarized parton densities, their evolution and their measurement. *Z. Phys. C*, 45:669–679, 1990.

[26] R. L. Jaffe and X.-D. Ji. Chiral-odd parton distributions and polarized Drell–Yan process. *Phys. Rev. Lett.*, 67:552–555, 1991.

[27] A. Airapetian et al. (HERMES Collaboration). Evidence for a transverse single-spin asymmetry in leptoproduction of $\pi^+\pi^-$ pairs. *J. High Energy Phys. (JHEP)*, 2008(06):017, 2008.

[28] A. Bacchetta and M. Radici. Two-hadron semi-inclusive production including subleading twist contributions. *Phys. Rev. D*, 69(7):074026, April 2004.

[29] S. Brodsky and S. Gardner. Evidence for the absence of gluon orbital angular momentum in the nucleon. *Phys. Lett. B*, 643(1):22–28, 2006.

[30] X. Artru. Proposals for measuring transversity distributions in deep inelastic electron scattering and a model for E-704 asymmetries. In *Contribution to the $V^{th}$ Intl. Workshop on High Energy Spin Physics (SPIN-93).* Protvino, Russia, September 1993.

[31] A. Bacchetta and M. Radici. Modeling dihadron fragmentation functions. *Phys. Rev. D*, 74(11):114007, December 2006.

[32] X. Ji, J.-P. Ma, and F. Yuan. QCD factorization for semi-inclusive deep-inelastic scattering at low transverse momentum. *Phys. Rev. D*, 71(3):034005, February 2005.

[33] R. Seidl (PHENIX Collaboration). Interference fragmentation function measurements in PHENIX. In *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010)* [107].

[34] R. Jaffe, H. Meyer, and G. Piller. Spin, twist and hadron structure in deep inelastic processes. In Frieder Lenz, Harald Grießhammer, and Dieter Stoll, editors, *Lectures on QCD*, volume 496 of *Lecture Notes in Physics*, pages 178–249. Springer Berlin / Heidelberg, 1997.

[35] M. Diehl. Vector meson production from a polarized nucleon. *J. High Energy Phys. (JHEP)*, 09:064, 2007.

[36] M. Diehl and S. Sapeta. On the analysis of lepton scattering on longitudinally or transversely polarized protons. *Eur. Phys. C*, 41:515, 2005.

[37] C. Amsler et al. Review of particle physics. *Phys. Lett. B*, 667, 2008.

[38] R. Lamb. *The Boer-Mulders and Cahn effects: Azimuthal modulations in the spin-independent SIDIS cross section at* HERMES. Ph.D. dissertation, University of Illinois at Urbana-Champaign, August 2010.

[39] A. Bacchetta et al. Single-spin asymmetries: the Trento conventions. *Phys. Rev. D*, 70:117504, 2004.

[40] J. Wolny (COMPASS Collaboration). Transversity signal in two hadron pair production in COMPASS. In *Proc. of XVII Int. Workshop on Deep-Inelastic Scattering and Related Topics (DIS09).* Madrid, Spain, April 2009.

[41] M. Anselmino et al. Transversity and Collins functions from SIDIS and $e^+e^-$ data. *Phys. Rev. D*, 75(5):054032, March 2007.

[42] M. Anselmino et al. General partonic structure for hadronic spin asymmetries. *Phys. Rev. D*, 73(1):014020, January 2006.

[43] M. Anselmino et al. Single spin asymmetries in $lp \to hx$ processes: A test of factorization. *Phys. Rev. D*, 81(3):034007, February 2010.

[44] D. Boer, R. Jakob, and P. J. Mulders. Angular dependences in electroweak semi-inclusive leptoproduction. *Nucl. Phys. B*, 564(3):471–485, 2000.

[45] M. Boglione. Phenomenology of transverse momentum dependent distribution and fragmentation functions. In *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010)* [107].

[46] A. Airapetian et al. (HERMES Collaboration). Observation of the naive-$T$-odd Sivers effect in deep-inelastic scattering. *Phys. Rev. Lett.*, 103(15):152002, October 2009.

[47] A. Airapetian et al. (HERMES Collaboration). Effects of transversity in deep-inelastic scattering by polarized protons. *Phys. Lett. B*, 693(1):11–16, 2010.

[48] V. Barone, A. Drago, and P. G. Ratcliffe. Transverse polarisation of quarks in hadrons. *Phys. Rep.*, 359(1-2):1–168, 2002.

[49] F. Giordano (HERMES Collaboration). Cosine modulation of the unpolarized pion cross section at HERMES. In *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010)* [107].

[50] M. Diefenthaler (HERMES Collaboration). Signals for transverse-momentum dependent quark distributions studied at the HERMES experiment. In *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010)* [107].

[51] G. A. Miller. Non-spherical shapes of the proton: Existence, measurement and computation. *Nucl. Phys. News*, 2008.

[52] A. Bianconi et al. Two-hadron interference fragmentation functions: I. General framework. *Phys. Rev. D*, 62(3):034008, July 2000.

[53] S. Gliske, W. H. Klink, and T. Ton-That. Algorithms for computing generalized U($N$) Racah coefficients. *Acta Appl. Math.*, 88:229–249, 2005.

[54] W. Klink and T. Ton-That. Invariant theory of tensor product decomposition of U($N$) and generalized Casimir operators. *Not. of the AMS*, 56, 2009.

[55] S. Gliske, W. Klink, and T. Ton-That. Algorithms for computing $U(N)$ Clebsch Gordan coefficients. *Acta Appl. Math.*, 95:51–72, 2007.

[56] D. Boer. Sudakov suppression in azimuthal spin asymmetries. *Nucl. Phys. B*, 603(1-2):195–217, 2001.

[57] F. Ceccopieri, M. Radici, and A. Bacchetta. Evolution equations for extended dihadron fragmentation functions. *Phys. Lett. B*, 650(1):81–89, 2007.

[58] A. A. Henneman, D. Boer, and P. J. Mulders. Evolution of transverse momentum dependent distribution and fragmentation functions. *Nucl. Phys. B*, 620(1-2):331–350, 2002.

[59] A. Majumder and X.-N. Wang. Dihadron fragmentation function and its evolution. *Phys. Rev. D*, 70(1):014007, July 2004.

[60] U. Elschenbroich. *Transverse Spin Structure of the Proton Studied in Semi-Inclusive DIS.* PhD thesis, Universiteit Gent, 2006.

[61] A. Bianconi et al. Two-hadron interference fragmentation functions: II. A model calculation. *Phys. Rev. D*, 62(3):034009, July 2000.

[62] B. Ball. Diffractive slope extraction of exclusive $\rho^0_L$ and $\rho^0_T$ at HERMES. Contributed talk at 2008 APS Division of Nuclear Physics Meeting, Oakland, California, October 2008.

[63] S. Gliske. Analysis with kernel density estimation. In *Workshop on the Transverse Partonic Structure of Hadrons (TPSH).* Yerevan, Armenia, June 2009.

[64] S. Gliske. Kernel density estimation techniques for HERMES. HERMES Internal Report 08-015, December 2008.

[65] A. Airapetian et al. (HERMES Collaboration). Azimuthal modulations for charged hadrons, pions, and kaons produced in deep-inelastic scattering by unpolarized protons and deuterons. In preparation.

[66] A. D. Polyanin and A. V. Manzhirov. *Handbook of Integral Equations.* CRC Press, Boca Raton, 1998.

[67] T. Sjöstrand et al. High-energy-physics event generation with PYTHIA 6.1. *Comp. Phys. Comm.*, 135(2):238–259, 2001.

[68] N. Makins. GMC_TRANS manual. HERMES Internal Report 03-060.

[69] A. Airapetian et al. (HERMES Collaboration). Leading order determination of the gluon polarization from high-pT hadron electroproduction. *J. High Energy Phys. (JHEP)*, 08:130, 2010.

[70] A. Hillenbrand. *Measurement and Simulation of the Fragmentation Process at* HERMES. PhD thesis, Friedrich Alexander Universitaet, September 2005.

[71] P. Liebing. *Can the Gluon Polarization in the Nucleon be extracted from* HERMES *Data on single high-pT Hadrons?* PhD thesis, Universitaet Hamburg, 2004.

[72] R. Brun and F. Rademakers. `ROOT`—an object oriented data analysis framework. *Nucl. Instrum. and Meth. D*, 389:81–86, 1997. See also `http://root.cern.ch/`.

[73] W. Wander (HERMES Collaboration). `DAD`—Distributed Adamo Database system at HERMES. In *Proc. of Intl. Conf. on Computing in High Energy Physics (CHEP'95).* Rio de Janeiro, Brazil, September 1995.

[74] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys. (JHEP)*, 2002(07):012, 2002.

[75] M. R. Whalley, D. Bourilkov, and R. C. Group. The Les Houches accord PDFs (LHAPDF). `http://hepforge.cedar.ac.uk/lhapdf/`.

[76] A. Bacchetta, F. Conti, and M. Radici. Transverse-momentum distributions in a diquark spectator model. *Phys. Rev. D*, 78(7):074010, October 2008.

[77] M. Glück, E. Reya, and A. Vogt. Dynamical parton distributions revisited. *Eur. Phys. J*, C(5):461–470, 1998.

[78] M. Glück et al. Models for the polarized parton distributions of the nucleon. *Phys. Rev. D*, 63(9):094005, March 2001.

[79] M. Anselmino et al. Transversity and Collins functions from SIDIS and $e^+e^-$ data. *Phys. Rev. D*, 75(5):054032, March 2007.

[80] U. D'Alesio. Update on transversity and Collins functions from SIDIS and $e^+e^-$ data. Invited talk at the Ringberg Workshop "New Trends in HERA Physics 2008", Ringberg Castle, Tegernsee, Germany, October 2008.

[81] M. Anselmino et al. The Sivers function from SIDIS data. In *Proc. of XVI Int. Workshop on Deep-Inelastic Scattering and Related Topics (DIS08)*. London, U.K., April 2008.

[82] M. Anselmino et al. Sivers effect for pion and kaon production in semi-inclusive deep inelastic scattering. *Eur. Phys. A*, 39:89–100, 2009.

[83] V. Barone, S. Melis, and A. Prokudin. Boer-Mulders effect in unpolarized SIDIS: An analysis of the COMPASS and HERMES data on the $cos2\phi$ asymmetry. *Phys. Rev. D*, 81(11):114026, June 2010.

[84] D. de Florian, R. Sassot, and M. Stratmann. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev. D*, 75(11):114010, June 2007.

[85] D. de Florian, R. Sassot, and M. Stratmann. Global analysis of fragmentation functions for protons and charged hadrons. *Phys. Rev. D*, 76(7):074033, October 2007.

[86] S. Kretzer. Fragmentaion functions from flavour-inclusive and flavour-tagged $e^+e^-$ annihilations. *Phys. Rev. D*, 62:054001, 2000.

[87] K. Ackerstaff et al. The HERMES spectrometer. *Nucl. Instrum. and Meth. A*, 417:230–265, 1998.

[88] A. Vandenbroucke. *Exclusive $\pi^0$ Production at* HERMES*: Detection – Simulation – Analysis.* PhD thesis, Universiteit Gent, November 2006.

[89] W. Deconinck. *The Search for Exotic Baryons at the* HERMES *Experiment.* Ph.D. dissertation, University of Michigan, April 2008.

[90] C. Baumgarten et al. (HERMES Target Group Collaboration). A gas analyzer for the internal polarized target of the HERMES experiment. *Nucl. Instrum. and Meth. A*, 508:268, 2003.

[91] A. Airapetian et al. (HERMES Collaboration). The HERMES polarized hydrogen and deuterium gas target in the HERA electron storage ring. *Nucl. Instrum. and Meth. A*, 540(1):68–101, 2005.

[92] A. Airapetian et al. The time-of-flight technique for the HERMES experiment. *Nucl. Instrum. and Meth. A*, 540:305–310, 2005.

[93] J. Ely. HERMES calorimeter position reconstruction study. HERMES Internal Report 01-056, 2001.

[94] R. Lamb and A. Hillenbrand. Event level RICH algorithm: An explanation and user guide. HERMES Internal Report 07-017, July 2008.

[95] B. Hommez. *A Study of Fragmentation Processes in the* HERMES *experiment using a Ring Imaging Čerenkov Detector.* PhD thesis, Universiteit Gent, July 2003.

[96] W. Augustyniak. TMC—Vertex reconstruction in the presence of the HERMES transverse target magnet. HERMES Internal Note 07-008, 2008.

[97] J. Wendland. *Polarized parton distributions measured at the* HERMES *experiment.* Ph.D. dissertation, Simon Fraser University, 2003.

[98] I. Akushevich, H. Boettcher, and D. Ryckbosch (HERMES Collaboration). RADGEN 1.0. Monte Carlo generator for radiative events in DIS on polarized and unpolarized targets. In *Monte Carlo Generators for HERA Physics.* Hamburg, Germany, 1999.

[99] M.J.D. Powell. *Numerical Methods for Nonlinear Algebraic Equations*, chapter 6 and 7, pages 87–161. Gordon and Breach, 1970.

[100] J.J. Moré, B.S. Garbow, and K.E. Hillstrom. Testing unconstrained optimization software. *ACM Trans. on Math. Software*, 7(1):17–41, 1981.

[101] M. Galassi et al. *GNU Scientific Library Reference Manual (3rd Ed.).* Network Theory Ltd., 2009. `http://www.gnu.org/software/gsl/`.

[102] D. Gabbert. *Determination of the structure function* $F_2$ *at* HERMES. PhD thesis, Universitaet Hamburg, October 2008.

[103] A. Airapetian et al. (HERMES Collaboration). Exclusive $\rho^0$ electroproduction on transversely polarized protons. *Phys. Lett. B*, 679(2):100–105, 2009.

[104] R.L. Jaffe, X. Jin, and J. Tang. Interference fragmentation functions and the nucleon's transversity. *Phys. Rev. Lett.*, 80:1166, 1998.

[105] R.L. Jaffe, X. Jin, and J. Tang. Interference fragmentation functions and valence quark spin distributions in the nucleon. *Phys. Rev. D*, 57:5920, 1998.

[106] M. Aghasyan (CLAS Collaboration). Studies of spin-orbit correlations at JLab. In *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010)* [107].

[107] *Proc. of 19th Int. Spin Physics Symposium (SPIN 2010).* Forschungszentrum Jülich, Germany, September 2010.