Studying Gluon Correlations and Nuclear Effects through Transverse Single-Spin Asymmetry Measurements at PHENIX and Promoting an Open Research Infrastructure in High Energy Physics

by

Dillon S. Fitzgerald

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2024

Doctoral Committee:

Professor Christine A. Aidala, Chair
Professor Wolfgang Lorenzon
Associate Professor Joshua Spitz
Professor Alejandro Uribe
Professor James Wells

Dillon Fitzgerald

dillfitz@umich.edu

ORCID: 0000-0001-6862-6876

To Veronica Verkest, whose love anchors and inspires me

ACKNOWLEDGEMENTS

So many interactions have shaped my trajectory during the pursuit of my PhD. While I will surely not have room to mention everyone who had an impact, I will do my best to list many that stuck out.

First and foremost, I would like to thank and show my appreciation for my partner, Veronica Verkest, whose love and support has inspired me, nurtured my growth, and pushed me to achieve my goals. She has stuck by me in the most difficult of times and taken care of me when I needed it most. She is a constant reminder that life is beautiful and should be enjoyed, helping me live mine to the fullest. Veronica is also a high energy nuclear physicist, facilitating my development from an intellectual standpoint with countless ideas and discussions.

I could not have envisioned a better match for a research advisor than Christine Aidala. She introduced me to the fascinating world of research in quantum chromodynamics, where my skills with computation and software were readily applied to study fundamental topics like proton structure and hadron formation. Christine was always happy to provide me with answers, countless resources, and access to a wealth of experts in the field. She provided so many outlets for me to learn and engage with the community during the course of my PhD, and I have grown tremendously as a scientist as a result. Not only was Christine an excellent mentor, she also went above an beyond to provide an environment where students felt comfortable and valued. I feel lucky to have participated in several jam sessions hosted by Christine, where members of her research group had the opportunity to bond beyond an interest in

physics, establishing a sense of community.

There have been numerous members of the Aidala research group, past and present, that have shaped my experience tremendously. The mentorship of Nicole Lewis stands out in particular, who taught me so much about QCD and transverse single-spin asymmetries both from an experimental and theoretical perspective. Joseph Osborn, Kara Mattioli, Sookhyun Lee, and Jordan Roth also taught me a great deal about QCD and jet physics, while Joseph Osborn taught me a lot about software practices and experimental infrastructure in addition. Desmond Shangase, Cynthia Nunez, and Chamindu Sangeeth Amarasinghe were always there for me when I needed a pick me up, a stimulating conversation, or a friend. Sharing a space to learn and work with these colleagues has been a privilege and a pleasure. The same goes for Ibrahim Chahrour, Devon Loomis, Manuel Ramirez Garcia, and Jose Marcos Arias, who joined the group in more recent years, but have already had a significant and positive impact on my journey.

I am grateful I had the opportunity to know Jordan Roth, and for the time we spent together. Jordan was incredibly intelligent and meticulous. He was also one of the most welcoming individuals I met when I was still new to the department and the research group. While Christine was away, Jordan stepped in to establish a sense of community, even organizing a trip down the Huron River followed by a game night for the whole group. Jordan shined bright, and that is how I will always remember him. He will forever be missed.

During my PhD, I have been a proud member of the PHENIX collaboration, comprising a wealth of experts from all over the globe. PHENIX promoted a welcoming and accessible research environment, hosting a number of "PHENIX Schools" — coordinated events where expertise was passed down to new members of the collaboration. I am grateful for the spin physics working group within PHENIX, a community of researchers working on common physics problems where issues could be discussed

collectively. I am also thankful for the conveners of this working group, Ralf Seidl and Sanghwa Park, who provided careful feedback and expertise on my analyses.

In more recent years, I have also been a proud member of the LHCb collaboration, working on tools related to open data. The Data Processing and Analysis project at LHCb, in particular, Work Package 6 focused on analysis preservation and open data, has been a key resource during this journey. I am very grateful for the knowledge and expertise passed on from Adam Morris, Sebastian Neubert, and Eduardo Rodrigues. I am also thankful for the very productive collaboration with Adam Morris, Piet Nogga, Tibor Simko, and Daan Eduard Rosendal, facilitating the development of a tool that goes far beyond the initially planned scope.

Much of my preparedness with research in high energy physics is owed to Robert Harr, a physics professor at Wayne State University who patiently worked with me for years as an undergraduate student and taught me about countless tools as well as research standards and practices. I still use many of the tips and tricks that he taught me to this day!

A true appreciation for physics comes with an appreciation of the underlying mathematics. Andrew Salch instilled this passion in me with his teachings at Wayne State University. He helped me see the beauty in mathematics and pushed me to further my understanding. I have never felt more inspired by a teacher.

My parents, Kevin Fitzgerald and Sandra Fitzgerald, have provided so much support and love that have shaped me into who I am today. They instilled in me the strength and courage to follow my passion, and nurtured those passions every step of the way, even when my focus changed from music to physics. I would not be who I am without them, and I am so grateful for everything they taught me and continue to do for me. I am also incredibly grateful for the love and support of my siblings, Ryan Fitzgerald and Shawn Fitzgerald, who are always there for me when I need them, no matter the circumstances.

While I did move around a few times in my youth, I have been in Michigan for over 15 years. I have been lucky enough to "grow roots" here, with a group of friends that I could rely on for anything. So much of my growth during high school, my undergraduate days, and beyond, was due to the support of such a fantastic group of friends. Jefferson Renusch, Joseph Kohler, Brent Cotter, Patrick Walsh, Rocco Deciechi, Kevin Eaglen, Barret Vermilion, Paige Hammerl, Maxwell Doll, and others, have had a tremendous impact on my life by being such close friends for so long.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Since the observation of large transverse single-spin asymmetries (TSSAs) in the 1970s, the subject has received much attention. Attempts to explain such asymmetries have provided a critical test of quantum chromodynamics (QCD), as a theoretical explanation requires additional correlations between quark and gluon fields or additional degrees of freedom in parton distribution functions and fragmentation functions that are not present in most perturbative QCD (pQCD) calculations. The Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory is the world's first and only polarized proton collider, with both proton-proton ($p^{\uparrow}p$) and proton-nucleus ($p^{\uparrow}A$) collisions recorded with polarized protons, allowing for an exploration of such observables at high energies where pQCD is valid and in a setting where gluons are accessible at leading order. This permits the study of twist-3 trigluon correlations within polarized protons, as well as modifications that arise to TSSAs in collision systems with additional nuclear matter. Measurements of TSSAs for midrapidity open heavy flavor electrons in $\sqrt{s} = 200$ GeV $p^{\uparrow}p$ collisions from the PHENIX experiment are reported in this dissertation to be consistent with zero within measured uncertainties. This production channel provides high sensitivity to gluons in polarized protons, as it proceeds predominantly through gluon-gluon fusion. This measurement provided the first explicit constraints on normalization parameters λ_f and λ_d of the antisymmetric and symmetric trigluon correlators, respectively, to the unpolarized gluon PDF. The 1σ confidence intervals determined by comparing the theoretical models to data were $\lambda_f = -0.01 \pm 0.03$ GeV and $\lambda_d = 0.11 \pm 0.09$ GeV, implying that trigluon correlations in transversely polarized protons are potentially sizable but also consistent with zero

based on the statistical precision of the measurement. Higher precision studies will need to be conducted in the future to pin down the strength of trigluon correlations in transversely polarized protons. First measurements of TSSAs for midrapidity π^0 and η mesons in $\sqrt{s_{NN}} = 200$ GeV $p^{\uparrow}$Au and $p^{\uparrow}$Al collisions from PHENIX are also reported in this dissertation, and are consistent with zero within measured uncertainties. This is compared with data from the same observable in $\sqrt{s} = 200$ GeV $p^{\uparrow}p$ collisions from which it was determined that no nuclear modification of the TSSA was observed. In addition to analyzing the unique data collected at RHIC, tools were developed to help promote an open research infrastructure in high energy physics by making public the unique data collected by the LHCb experiment at the Large Hadron Collider (LHC) at CERN. An application known as the LHCb Ntuple Wizard provides a means for external users (e.g. theorists and phenomenologists) to make queries of the collected LHCb data through an intuitive web interface, resulting in Ntuples delivered to the CERN Open Data Portal. The detection capabilities of LHCb make it ideal for studying the hadron formation process at the energy frontier, as well as a multitude of other measurements including the core flavor physics program. Providing public access to such datasets not only facilitates the preservation of data, but also has the potential to increase the overall impact and reach of the data to truly optimize the scientific output. The first public release of the LHCb Ntuple Wizard is expected in 2024.

CHAPTER I

Introduction

It has long been a goal of humans to understand the nature of the world around them in terms of its smallest and most fundamental constituents. In modern high energy physics experiments, we have access to processes that occur on the femtometer scale (10^{-15} m), five orders of magnitude smaller than the size of a typical atom. This enables the exploration of the interior structure of nucleons, such as protons and neutrons, once believed to be elementary. In reality, the interior structure of nucleons is incredibly complex. The goal of this dissertation is a continuation of this quest to understand nature in terms of its fundamental constituents. This is carried out in two ways. The first and more direct way is through measurements with the PHENIX detector to explore spin-momentum and spin-spin correlations within protons, analogous to the fine and hyperfine structure in atomic systems, and in the process of hadronic bound state formation. The resulting publications from these measurements can be found in Refs. [1, 2]. The second, more indirect approach, is through the development of tools that enable external researchers to access LHCb data, such that the full potential of the data can be realized. The focus of this approach is on the development of a tool called the LHCb Ntuple Wizard, described in Ref. [3].

1.1 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the quantum field theory describing the strong nuclear force, responsible for binding quarks and gluons into nucleons (protons and neutrons), and nucleons into nuclei. It is a remarkable quantum field theory that is non-abelian, with important features such as *asymptotic freedom* and *color confinement* that make its structure unlike any other known quantum field theory (QFT) in nature. As a consequence of these unique features, the degrees of freedom of the theory (quarks and gluons) are different than those we have access to experimentally (hadrons — bound states of quarks and gluons). This complication seeded a fascinating historical development in studying the structure of the nucleon, and eventually, quantum chromodynamics as a successful theory for describing such bound states. A brief historical overview will be presented in Section 1.2.

1.1.1 Lagrangian Density and Interactions

Quantum chromodynamics is an important component of the Standard Model of particle physics, and as such, the QCD Lagrangian density is encoded in the Standard Model Lagrangian density shown in Equation 1.1.

$$\begin{aligned}\mathcal{L}_{SM} = &-\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a + i\sum_f \bar{\psi}_f \not{D}\psi_f + D\phi^\dagger \cdot D\phi + M^2\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2 \\ &- \sum_{ij} h_{ij}\psi_{iR}\bar{\phi}\psi_{j,L} + \text{gauge-fixing terms, etc.,}\end{aligned} \tag{1.1}$$

where a is an index for the 12 generators in the adjoint representation of SU(3) $\otimes$ SU(2) $\otimes$ U(1), $G^a_{\mu\nu}$ is the gauge field strength tensor, and ψ_f are fermions (both quarks and leptons). $\not{D} = \gamma^\mu D_\mu$ where γ^μ are the Dirac matrices and D_μ is the covariant derivative, which encodes the space-time derivative and interaction terms with gauge boson fields. The Higgs field is symbolized by ϕ, where M represents the

Higgs boson mass, and λ is the Higgs self-coupling parameter. Finally, h_{ij} are Yukawa couplings between Higgs and fermion fields, which determine the mass of fermions. The QCD Lagrangian density is written in Equation 1.2. There are a few important distinctions between Eqs. 1.1 and 1.2, namely the presence of the Higgs field ϕ in Equation 1.1, and the span of a, which only covers the 8 generators in the adjoint representation of SU(3) in $\mathcal{L}_{QCD}$, written as follows,

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a + i \sum_f \bar{\psi}_f (\not{D} - m) \psi_f + \text{gauge-fixing terms. etc.} \tag{1.2}$$

In this case, ψ_f are quarks, with the index f representing the quark flavor. $D_\mu = \partial_\mu + igt^a A^a_\mu$ is the covariant derivative, which encodes the space-time derivative and interaction terms between the quark and gluon fields. The quark mass m comes from the interaction of the quark and Higgs fields, given by the $\sum_{ij} h_{ij} \psi_{iR} \bar{\phi} \psi_{j,L}$ term in Eq. 1.1, and $G^a_{\mu\nu}$ is the gluon field strength tensor in this case, defined as follows,

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f_{abc} A^b_\mu A^c_\nu. \tag{1.3}$$

The first two terms of Eq. 1.3 take the exact form of the gauge field strength tensor in quantum electrodynamics (QED), while the final term is due to the non-abelian nature of QCD and gives rise to gluon self interactions, specifically the 3- and 4-gluon vertices shown in Fig. 1.1. The quark-gluon interaction vertex, analogous to the charged fermion-photon interaction vertex in QED, can also be seen in Fig. 1.1.

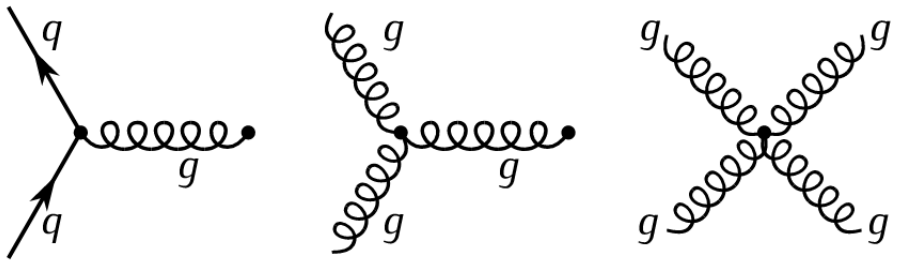

Figure 1.1: Interaction vertices of QCD, including the quark-gluon vertex and the gluon self-interaction vertices.

Interactions in QCD take place between objects that carry color charge (hence the name chromodynamics). Unlike QED, where the photon is electrically neutral, gluons themselves carry color charge, giving rise to many of the unique features of QCD. There are 3 types of charges in QCD, and it takes a combination of these 3 charges to create a neutral object. This inspired the name "chromodynamics", as one could envision combining 3 colors — red (R), blue (B), and green (G) — to make something color neutral (white).

1.1.2 Asymptotic Freedom

The coupling parameter between color charges g shown in Equations 1.2 and 1.3 is often expressed as $\alpha_s = g^2/4\pi$. One notable feature of this parameter is the asymptotic behavior, approaching zero in the limit of infinite momentum transfer. This was discovered independently in 1973 by David Gross and Frank Wilczek at Princeton [4], and David Politzer at Harvard [5]. The discovery was so fundamental to the development of QCD that all three physicists were awarded the Nobel Prize in 2004. Figure 1.2 shows α_s as a function of momentum transfer Q extracted from various measurements and lattice QCD calculations, where the precision of the perturbative expansions in α_s used in the extractions is indicated on the legend. It is common in the field to refer to a scattering process with a large momentum transfer as "hard",

and one with low momentum transfer as "soft". The terms hard and soft are also used to describe the scale of a process and the produced particles. A physical scale, known as $\Lambda_{QCD} \approx 200$ MeV, is typically taken as a threshold separating hard and soft physics, where perturbative techniques can be applied in calculating observables for hard processes.

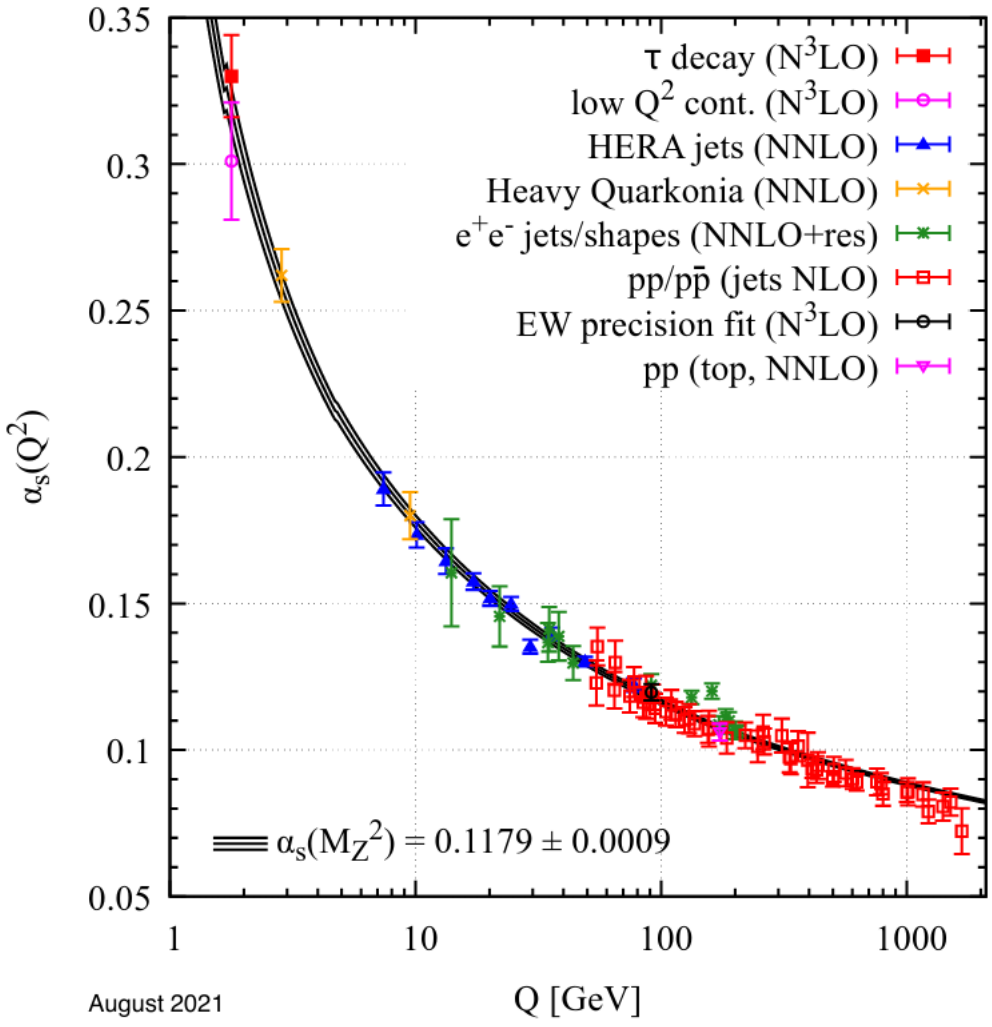

Figure 1.2: Plot of the strong coupling α_s as a function of momentum transfer or energy scale Q, showing the decrease of the coupling strength with increasing energy. This is known as asymptotic freedom. Figure taken from Ref. [6].

Following from the Callan-Symanzik equation, the relationship of the coupling

between quantum fields and the energy scale is encoded by a β function [7], as shown below,

$$\beta(g) = \frac{\partial g}{\partial \ln(\mu)}. \tag{1.4}$$

Here, μ is the so-called "renormalization scale" with units of energy, the value of which determines the value of the renormalized coupling, mass, and normalization of quark and gluon fields. A suitable choice of μ allows for increased accuracy of perturbative calculations in QCD [8]. The sign of the β function determines whether the coupling will increase or decrease with increasing energy. QCD has a negative β function, which implies a decrease in the coupling with increasing energy, as seen in Figure 1.2. In contrast, QED has a positive β function, which can be intuitively understood by the effect of vacuum polarization, in which the creation of virtual electric dipoles causes a screening effect leading to a decrease in the effective charge with increasing distance from the bare charge. The situation in QCD is more complicated, as the gluon self-coupling leads to a color anti-screening effect that is more prominent than the color screening from virtual color dipoles ($q\bar{q}$ pairs). This is summarized in Figure 1.3, in which the accumulation of virtual charges of the opposite (same) sign around the bare charge leads to a decrease (increase) in coupling strength for increasing distance in QED (QCD). The left hand side of Equation 1.4 can be determined from a Feynman diagram analysis at various levels in perturbation theory. This has been calculated at up to 5-loop precision [6] — a truly remarkable achievement of perturbative QCD (pQCD). The results from the 1-loop calculation are $\beta(g) = -\frac{g^3}{16\pi^2} b_0$ with $b_0 = (11N_c - 2n_f)/12\pi$ [6, 9]. $N_c = 3$ is the number of colors, and $n_f = 6$ is the number of flavors, resulting in a negative β function for QCD. Plugging this into Eq. 1.4 and solving reveals the asymptotic behavior of the coupling α_s, as shown in Eq. 1.5.

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + b_0 \alpha_s(\mu^2) \ln(Q^2/\mu^2)}, \tag{1.5}$$

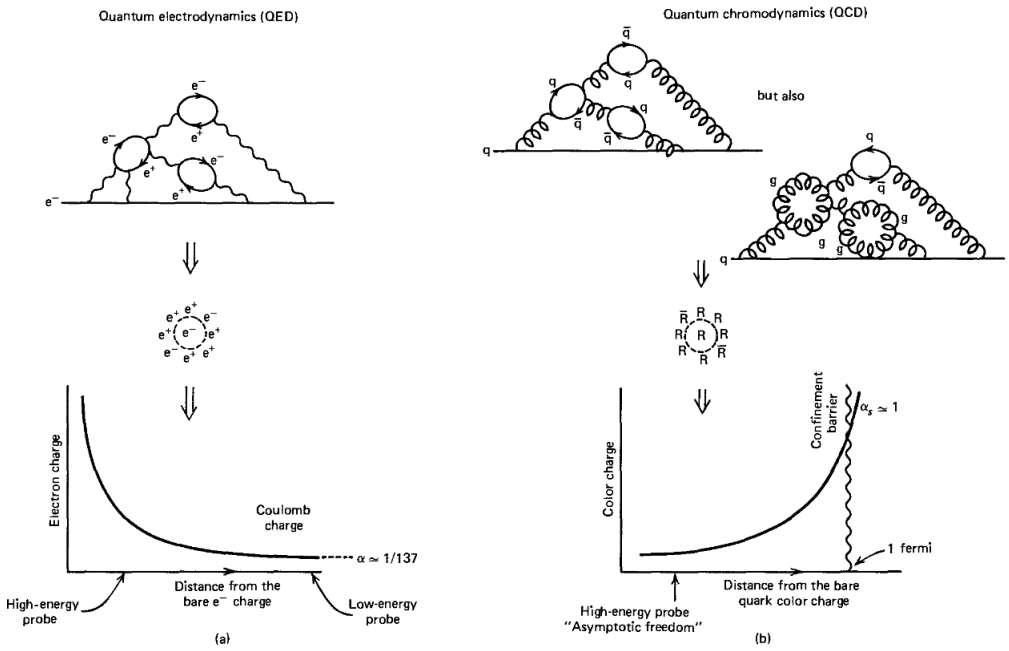

Figure 1.3: Diagram of the electromagnetic and strong couplings as a function of **distance**. This depicts the nature of the running coupling in terms of electric charges (QED) or color charges (QCD). See text for additional details. Figure taken from Ref. [10].

1.1.3 Color Confinement and Hadrons: Bound States of QCD

Another important consequence of quantum chromodynamics is color confinement, which states that color charges (quarks and gluons) cannot exist in isolation and must be confined to color-neutral bound states known as hadrons. Hadrons come in 3 different combinations — 3 quarks with colors RGB, known as baryons, 3 antiquarks with colors $\bar{R}\bar{G}\bar{B}$, known as antibaryons, and quark-antiquark pairs with colors $R\bar{R} + G\bar{G} + B\bar{B}$, known as mesons. The anticolor charges can be thought of as appropriate combinations of the opposing two color charges, i.e., $\bar{R}$ as a combination of B and

G, $\bar{B}$ as a combination of R and G, and $\bar{G}$ as a combination of R and B. This is summarized in Figure 1.4.

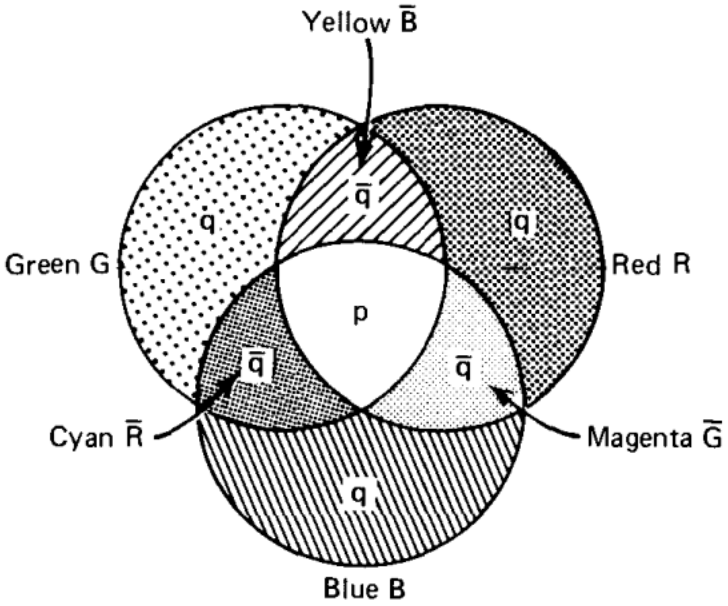

Figure 1.4: Diagram showing the color charges in QCD and how they are combined to create color neutral hadronic bound states. Figure taken from Ref. [10].

In the case of a $q\bar{q}$ color dipole, there is an attractive potential as shown in Fig. 1.5. The terminology "color flux tube" is often used to describe the gluon field between color charges in a bound state, given the field configuration. Separating the dipole would require energy to be put into the gluon field. As the separation between quarks increases it becomes energetically favorable to rupture the flux tube and create a new $q\bar{q}$ pair from the available energy in the gluon field, creating two mesons (and so on), as shown in Figure 1.6. The origin of color confinement is still poorly understood, but it is believed to be linked to the structure of the QCD vacuum [11].

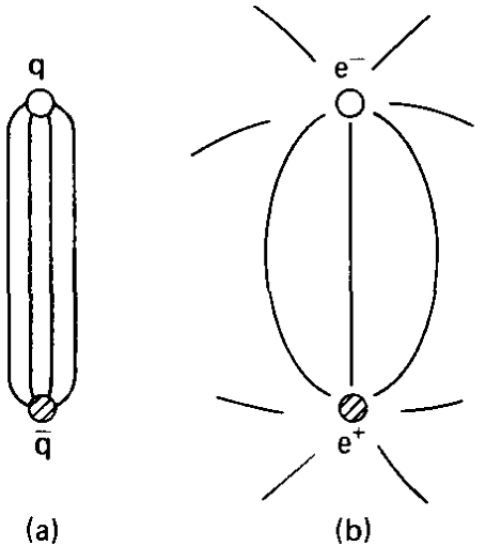

Figure 1.5: Field lines connecting a color dipole (left) in comparison with an electromagnetic dipole (right). Field lines in the color dipole are much more tightly compacted into a tube like shape, sometimes referred to as a color flux tube. Figure taken from Ref. [10].

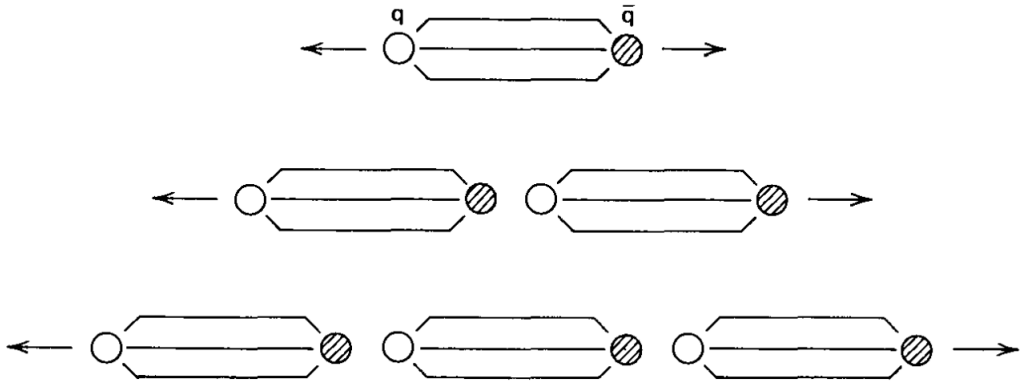

Figure 1.6: A schematic of the color confined nature of the $q\bar{q}$ dipoles (see text). Figure taken from Ref. [10].

1.2 Hadron Structure

The structure of hadrons, in particular the proton, has been the subject of study for decades. Lepton-proton scattering is a particularly useful tool for probing the interior structure of the proton in terms of its constituent quarks. To accomplish this, the energy of the interaction must be sufficiently high for the virtual photon to couple to a constituent quark rather than coherently to the entire proton. This is a regime known as Deep-Inelastic Scattering (DIS), in which the wavelength of the virtual photon is short enough to resolve nucleon structure (e.g. sub femtometer), resulting in fragmentation of the proton into a set of newly produced hadrons. In the DIS process, it is sufficient to measure only the final state lepton to extract information about the quark structure of the proton. A schematic of the DIS process is shown in Fig. 1.7, and the corresponding cross section is shown in Equation 1.6.

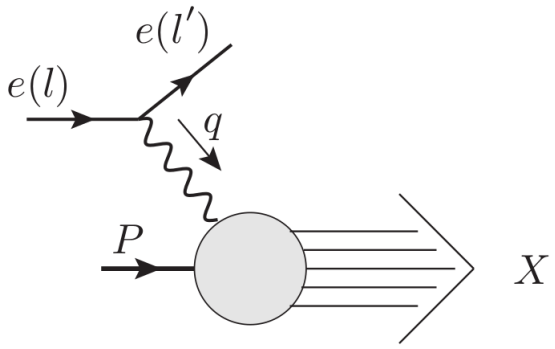

Figure 1.7: A schematic of the DIS process between an electron and a proton. l and l' are the initial and final state electron 4-momentum, while P is the proton 4-momentum and q is the virtual photon 4-momentum. Figure taken from Ref. [8].

$$E' \frac{d\sigma}{d^3\vec{l'}} \approx \frac{2\alpha^2}{sQ^4} L_{\mu\nu} W^{\mu\nu}, \tag{1.6}$$

where α is the electromagnetic coupling (also known as the fine structure constant), $\sqrt{s}$ is the center of mass energy of the collision, and Q is the momentum transfer,

defined as $Q = \sqrt{-q^2}$ where $q^\mu = l^\mu - l'^\mu$. $L_{\mu\nu}$ and $W^{\mu\nu}$ are referred to as the leptonic and hadronic tensors, defined in Eqs. 1.7 and 1.8, respectively. This expression for the deep inelastic scattering cross section is exact up to neglecting the electron and proton mass (i.e. the high energy limit).

$$L_{\mu\nu} = 2(l_\mu l'_\nu + l'_\mu l_\nu - g_{\mu\nu} l \cdot l') \tag{1.7}$$

$$W^{\mu\nu} = \frac{1}{4\pi} \int d^4z e^{iq\cdot z} \langle P, S | j^\mu(z) j^\nu(0) | P, S \rangle \tag{1.8}$$

Here, P and S are the momentum and spin vectors of the proton, while j^μ denotes the electromagnetic current for the quark fields. The hadronic tensor encodes a lot of information about the structure of the proton in terms of its constituent quarks. In particular, it can be decomposed into a general form where each term is the product of a tensor and a scalar function [8]. These scalar functions, known as structure functions, offer valuable information on the structure of both unpolarized hadrons ($F_1(x, Q^2)$ and $F_2(x, Q^2)$) and polarized hadrons ($g_1(x, Q^2)$ and $g_2(x, Q^2)$). This decomposition of the hadronic tensor is shown in Eq. 1.9,

$$\begin{aligned} W^{\mu\nu} = &\left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) F_1(x, Q^2) + \frac{(P^\mu - q^\mu P\cdot q/q^2)(P^\nu - q^\nu P\cdot q/q^2)}{P\cdot q} F_2(x, Q^2) \\ &+ i\epsilon^{\mu\nu\alpha\beta} \frac{q_\alpha S_\beta}{P\cdot q} g_1(x, Q^2) + i\epsilon^{\mu\nu\alpha\beta} \frac{q_\alpha \left(S_\beta - P_\beta \frac{S\cdot q}{P\cdot q}\right)}{P\cdot q} g_2(x, Q^2), \end{aligned} \tag{1.9}$$

where S_β is the spin vector, obeying the convention $S \cdot P = 0$. Considering first the case of deep inelastic scattering on an unpolarized proton target, terms proportional to the proton spin vector can be ignored. By further considering a reference frame known as the Breit frame in which the virtual photon momentum is $q^\mu = (0, 0, 0, -Q)$ and the proton momentum is $P^\mu = (0, 0, 0, P)$, Eq. 1.6 becomes Eq. 1.10.

$$\frac{d^2\sigma}{dxdy} \approx \frac{4\pi\alpha^2}{xyQ^2} \left[\left(1 - y - \frac{x^2y^2M^2}{Q^2}\right) F_2(x, Q^2) + y^2 x F_1(x, Q^2)\right]. \tag{1.10}$$

Eq. 1.10 is again exact up to the high energy limit approximation. The quantities x and y, typically referred to as "Bjorken-x" and "inelasticity", are defined in Eqs. 1.11 and 1.12, respectively.

$$x = \frac{Q^2}{2P \cdot q}, \tag{1.11}$$

$$y = \frac{q \cdot P}{l \cdot P}. \tag{1.12}$$

Further assumptions must be made to interpret the data from DIS experiments in a useful and intuitive way. For example, Figure 1.8 shows a compilation of global measurements of the structure function $F_2(x, Q^2)$ [12]. It can be seen from the data that at intermediate x values, the structure function is approximately independent of Q for a fixed x. This is a feature known as "Bjorken scaling". Since Q is related to the wavelength of the virtual photon, it can be thought of as the resolution scale of the process. The observation that the structure function F_2 is approximately independent of the resolution scale at high enough energies implies that in this limit, the proton can be thought of as a collection of point-like constituents. This observation verifies the applicability of the parton model (introduced in Section 1.2.1) in regimes where Bjorken scaling applies. In cases where the strong coupling is large and constituent quarks must be considered with their gluonic fields, the relevant correlation length can lead to violations of the scaling behavior and a breakdown of the notion of free partons.

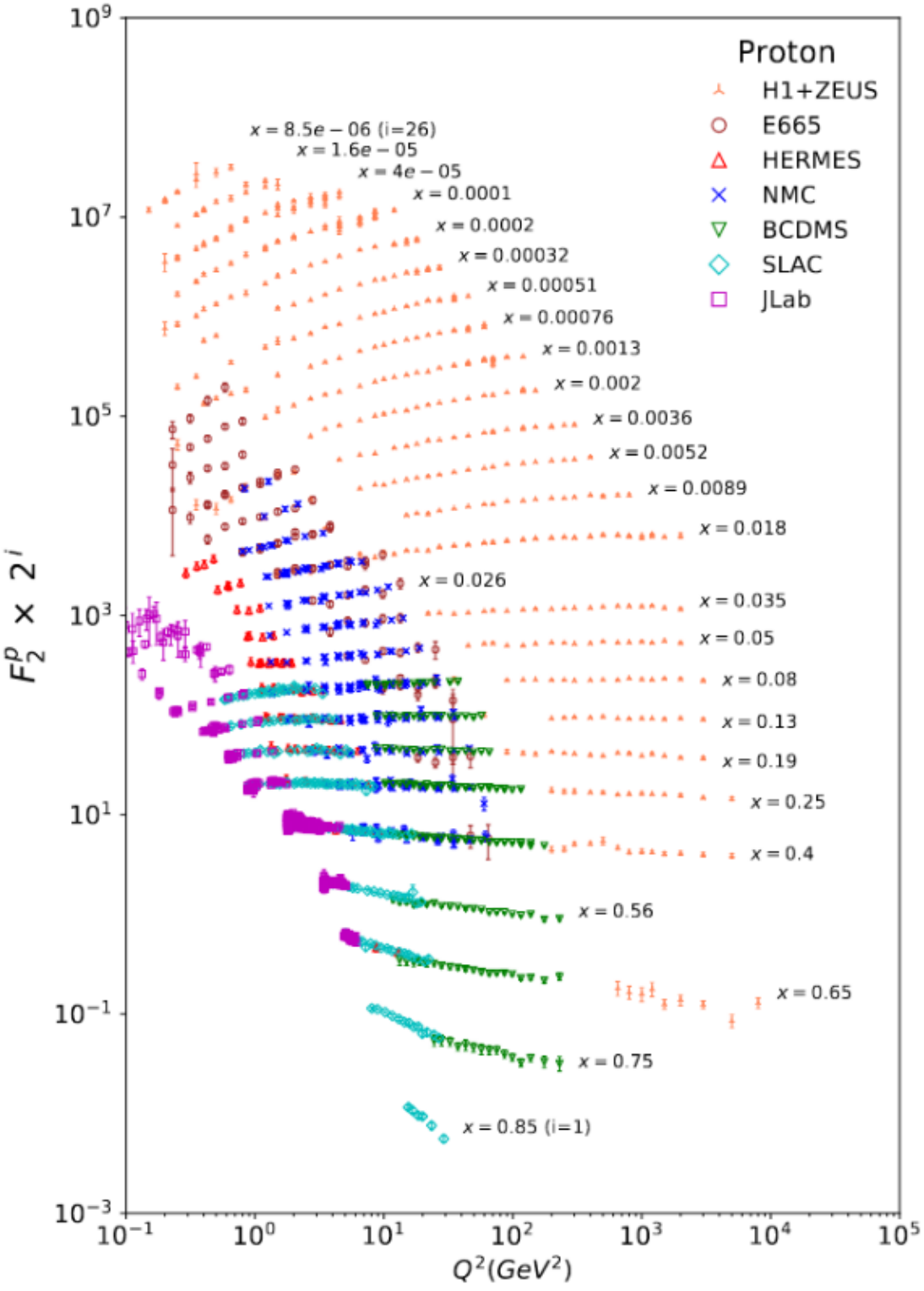

Figure 1.8: Unpolarized structure function F_2 for the proton compiled from global measurements. Figure taken from Ref. [12].

1.2.1 The Parton Model

The parton model was proposed by Richard Feynman in 1969 [13] and detailed by Bjorken and Paschos [14]. The salient feature of the model is that in high energy interactions, hadrons can be understood as collections of free point-like constituents (quarks and gluons) referred to as "partons". The argument intuitively relies on time dilation and length contraction in boosted systems. In the rest frame of the hadron, interactions between constituents occur on a time scale of 1 fm/c. In the center of mass frame of the DIS process, the hadron is boosted and this time frame is dilated depending on the energy of the collision. At high enough energies, interactions between constituents take sufficiently longer than the quark-lepton interaction, motivating the approximately free nature of partons at the moment of impact. This allows for a reformulation of the differential cross section for the DIS process as

$$d\sigma = \sum_a \int dx f_a(x) d\hat{\sigma}_{la}, \tag{1.13}$$

where the functions $f_j(x)$ are known as the parton distribution functions (PDFs), which describe the probability of sampling a parton with momentum fraction x from the parent hadron. PDFs are often denoted as $q_j(x), \bar{q}_j(x)$, and $g(x)$ for quarks, anti-quarks, and gluons, respectively, where j is the quark flavor. It is very important to note that in the parton model, Bjorken-x from Eq. 1.11 simplifies to the fraction of momentum the quark carries collinear to the proton momentum. This class of PDFs is known as collinear PDFs. There exists another class of PDFs called transverse-momentum-dependent PDFs (TMD PDFs), discussed in Section 1.4.2. The term $d\hat{\sigma}_{lj}$ represents the partonic cross section (also known as the partonic hard scattering coefficient), which can be calculated with perturbation theory. Calculating the partonic cross section to first order in the electromagnetic coupling α and comparing Eq. 1.10

to Eq. 1.13 yields the important result,

$$F_2 = \sum_j e_j^2 x q_j(x), \tag{1.14}$$
$$F_1 = \frac{1}{2x} F_2,$$

where the notation $f_j(x)$ has been replaced with $q_j(x)$ to denote that only quark PDFs contribute to the structure function F_2 in the parton model. Notice that Eq. 1.14 shows only x dependence for F_2, as expected from the observation of Bjorken scaling. The parton model and the observation of Bjorken scaling collectively led to the discovery of asymptotic freedom, discussed in Section 1.1.2. Figure 1.9 shows proton PDFs for quarks and gluons extracted from global fits to data at two different energy scales (note that μ^2 is often written in place of Q^2). u_v and d_v are the distributions for the valence quarks in the proton, which carry roughly 1/3 of the proton's total momentum.

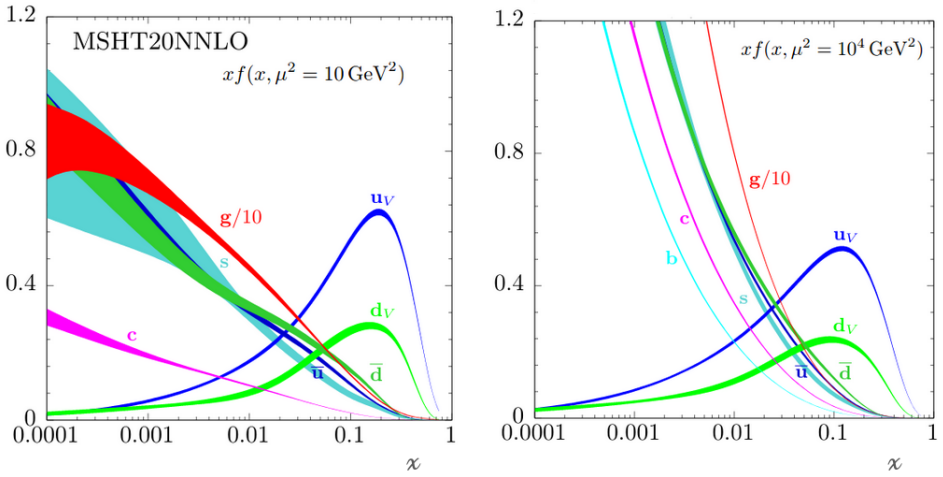

Figure 1.9: Parton distribution functions extracted from global fits at two different energy scales where the v subscript represents valence quarks. The rest of the curves are for sea quarks and gluons. Figure taken from Ref. [6]

It is interesting to note that at this point, no QCD was used in the analysis.

However, the lessons learned from analyzing DIS data led to the development of a powerful application of QCD in describing reactions involving hadrons, known as perturbative quantum chromodynamics (pQCD), described in the following.

1.2.2 Perturbative QCD (pQCD)

Calculations in QCD present a number of challenges; since quarks and gluons only exist in nature within hadrons, any interaction between quarks and gluons will inevitably involve hadrons. This implies components of corresponding cross sections that are not calculable from first principles using perturbative techniques due to the large coupling between quarks and gluons at hadronic length scales. However, asymptotic freedom allows for the application of perturbation theory in calculating scattering amplitudes for quark and gluon interactions at sufficiently high energies. Perturbative QCD (pQCD) provides a prescription for calculating cross sections in this case. The application of pQCD requires the use of an approximation known as *factorization* and sheds light on the principle of *universality*.

1.2.2.1 Factorization

Factorization states that cross sections can be separated (or factorized) into long and short distance components, where the short distance components (partonic cross sections) are calculated perturbatively with a Feynman diagram analysis. The long distance (nonperturbative) components describe the physics on length scales of typical hadronic bound states that cannot be calculated from perturbation theory. These functions describe the distribution of quarks and gluons within hadrons (initial state), or the process by which they form new hadrons (final state), typically in momentum space. Each hadron in the reaction has a corresponding nonperturbative function, where the initial state nonperturbative functions are generally referred to as parton distribution functions (PDFs) and the final state functions are generally referred to

as fragmentation functions (FFs). The assumptions that partons have only collinear momentum with their parent hadrons and that final state hadrons have only collinear momentum with the fragmenting parton, integrating over any transverse momentum, have been used to derive rigorous factorization theorems [15–18]. Examples of collinear factorization formulae for a number of reactions are shown in Section 1.2.2.3, where the variables x, x' represent collinear momentum fractions of partons with respect to their parent hadrons, and z is the collinear momentum fraction of the final state hadron with respect to the fragmenting parton. Factorization theorems also exist in which the PDFs and FFs explicitly depend additionally on transverse momenta k_T and j_T in the initial and final state, respectively. More details about this formalism can be found in Section 1.4.2.

Like the perturbative expansion in α_s used to calculate the partonic cross sections, the factorization of a hadronic cross section is also an expansion in a small parameter like $1/Q$ where Q is the hard scale. Typically, the leading term in this expansion is sufficient, and the corresponding factorization formula is said to be at "leading twist" or "twist-2". For certain observables, like transverse single-spin asymmetries (discussed in Section 1.4), a proper description using collinear factorization requires going beyond leading twist to the next term in the expansion, "twist-3". More information on the concept of twist, and how twist-3 collinear factorization can be used to describe TSSAs can be found in Section 1.4.1

1.2.2.2 Universality

The principle of universality implies process independence, or that PDFs and FFs are universal regardless of the process in consideration. This makes factorized pQCD calculations incredibly powerful, permitting nonperturbative functions measured in processes with the highest sensitivity to be used as inputs for other experiments. Experiments may have different kinematic coverage, requiring the evolution of the

PDFs and FFs to the appropriate energy regime. This is carried out with the DGLAP evolution equation introduced by Dokshitzer, Gribov, Lipatov, Altarelli, and Parisi in the 1970s [19–22], and formalized to the following [6]:

$$\frac{\partial f_a}{\partial \ln \mu^2} \propto \frac{\alpha_s(\mu^2)}{2\pi} \sum_b (P_{ab} \otimes f_b) \tag{1.15}$$

$$\frac{\partial}{\partial \ln \mu^2} D_i^h(x, \mu^2) = \sum_j \int_x^i \frac{dz}{z} P_{ji}(z, \alpha_s(\mu^2)) D_j^h \left(\frac{x}{z}, \mu^2\right), \tag{1.16}$$

where P_{ab} and $P_{ij}(z, \alpha_s(\mu^2))$ are the Altarelli-Parisi splitting functions [22], which are characterized by a perturbative expansion in α_s to describe successive parton radiation. In both Eqs. 1.15 and 1.16, μ^2 denotes the factorization scale, defined as a threshold scale to separate nonperturbative and perturbative physics. The factorization scale is typically taken to be the hard scale of the process Q, or some small integer fraction or multiple of Q, for example. It should also be noted that the factorization scale is distinct from the renormalization scale that is necessary in calculating the partonic cross section (perturbative part) in any pQCD factorization formula. For all factorization formulae written in this dissertation, the dependence on the factorization and renormalization scales is omitted for convenience.

1.2.2.3 Reactions Involving Hadrons

With pQCD in hand, the cross section for several important reactions with hadrons in the initial and/or final state can be calculated. Measuring cross sections for different types of reactions involving hadrons provides sensitivity to constrain nonperturbative functions for the various partons. A few important reactions are listed below.

Deep Inelastic Scattering: $lp \to lX$

This process is covered in detail in Section 1.2. There is only one hadron in

the initial state, and information about PDFs can be extracted from measuring the final state lepton. This makes it the simplest process for constraining quark PDFs, as charged leptons couple to quarks at leading order. The collinear factorization formula for calculating the differential cross section of the DIS process is shown in Eq. 1.17,

$$d\sigma = \sum_a f_a(x) \otimes d\hat{\sigma}_{la}. \tag{1.17}$$

Hadron production in electron positron annihilation: $e^+e^- \to hX$

There is only one hadron in the final state in this process, making it very useful in constraining final-state nonperturbative functions known as fragmentation functions (FFs), describing the probability of parton c fragmenting into hadron h with collinear momentum fraction z relative to the parton momentum. The collinear factorization formula used to calculate the differential cross section for this process is shown in Eq. 1.18,

$$d\sigma = \sum_c D_c^h(z) \otimes d\hat{\sigma}_{e^+e^- \to c}. \tag{1.18}$$

Drell-Yan: $pp \to l^+l^-X$

This process only involves hadrons in the initial state, where the leptoproduction necessarily proceeds through the $q\bar{q} \to l^+l^-$ channel. The principle of universality allows quark PDFs measured in DIS experiments to be appropriately evolved and used as an input for this process to extract information on the antiquark PDFs (and hence the sea quark PDFs). The collinear factorization formula for calculating the differential cross section for the Drell-Yan process is shown in Eq. 1.19,

$$d\sigma = \sum_{a,b} f_a(x) \otimes f_b(x') \otimes d\hat{\sigma}_{ab\to l^+l^-}. \tag{1.19}$$

Hadron production in proton collisions: $pp \to hX$

This process is necessarily complex as it includes 2 hadrons in the initial state and one in the final state, and therefore 3 nonperturbative functions in the factorization theorem. However, it offers leading order access to gluons, unlike any of the former listed processes. Therefore, hadron production in proton colliders provides sensitivity to constrain gluon PDFs. In this case, fragmentation functions can be measured in e^+e^- collisions, evolved, and used as an input in the extraction. The collinear factorization formula used to calculate the differential cross section for this process is shown in Eq. 1.20,

$$d\sigma = \sum_{a,b,c} f_a(x) \otimes f_b(x') \otimes d\hat{\sigma}_{ab\to c} \otimes D_c^h(z). \tag{1.20}$$

1.2.3 Polarized Structure

Our understanding of the spin structure of the nucleon has evolved greatly in the past several decades, from a naive assumption that the valence quarks of the proton carry all of its spin, to a much more complex understanding involving the spin and orbital angular momentum of quarks (valence and sea) as well as gluons.

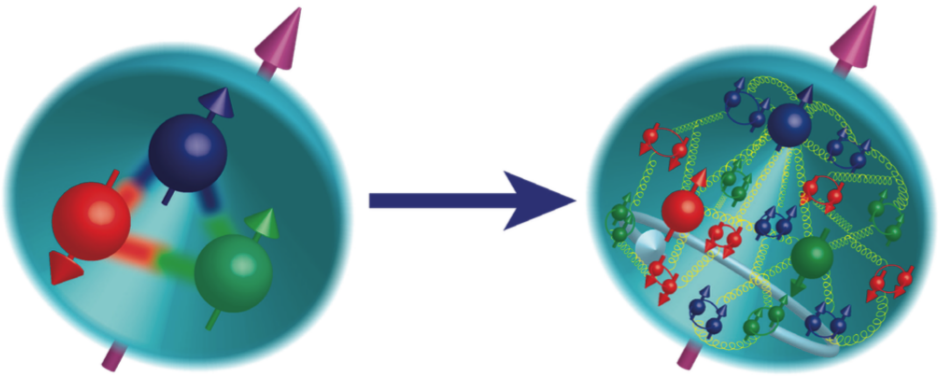

Figure 1.10: Cartoon depiction of our understanding of the proton spin structure in the 1980s (left) and now (right). The latter contains contributions from spin and orbital angular momentum of valence quarks, sea quarks, and gluons.

.

The structure functions related to polarized hadrons are introduced as $g_1(x, Q^2)$ and $g_2(x, Q^2)$ in Section 1.2. Polarized DIS experiments provide sensitivity to constrain these functions. Some useful measurements include the cross section asymmetry of a longitudinally polarized lepton on a longitudinally polarized proton with positive and negative helicity, or a transversely polarized proton with opposing spin orientations. This is summarized in Equations 1.21 and 1.22 [23].

$$\frac{d^2\sigma^{l^{\rightarrow}p^{\rightarrow}}}{d\sigma dE'} - \frac{d^2\sigma^{l^{\rightarrow}p^{\leftarrow}}}{d\sigma dE'} = -\frac{4\alpha^2 E'}{Q^2 EM\nu}[(E + E'\cos\theta)g_1 - 2xMg_2] \tag{1.21}$$

$$\frac{d^2\sigma^{l^{\rightarrow}p^{\uparrow}}}{d\sigma dE'} - \frac{d^2\sigma^{l^{\rightarrow}p^{\downarrow}}}{d\sigma dE'} = -\frac{8\alpha^2 (E')^2}{Q^2 M\nu^2}\left(\frac{\nu}{2E}g_1 + g_2\right)\sin\theta\cos\phi \tag{1.22}$$

Historically, the former process was used to constrain g_1, which can in turn be used with the latter process to constrain g_2. A similar analysis to that conducted in Section 1.2 can be carried out on Eq. 1.21, yielding an expression for the differential

cross section asymmetry in the parton model,

$$\Delta d\sigma = \sum_a \Delta f_a(x) \otimes \Delta d\hat{\sigma}_{la}. \tag{1.23}$$

Comparing Eq. 1.23 with Eq. 1.21 and considering the high energy limit, one obtains Eq. 1.24

$$g_1 = \frac{1}{2} \sum_j e_j^2 [\Delta q_j(x) + \Delta \bar{q}_j(x)], \tag{1.24}$$

with $\Delta q = q_+(x) - q_-(x)$, the difference in quark PDFs with positive and negative helicity. In words, g_1 can be intuitively understood as the sum of helicity distributions of quarks and antiquarks, typically considered over the light quark flavors u, d, s. Global measurements of g_1 for the proton are shown in Figure 1.11. The helicity distributions for partons are of great interest to the community, as they are key components in resolving the "proton spin puzzle" [24], previously known as the "proton spin crisis", the initially striking observation by the EMC experiment at CERN that the spin of the quarks carries a relatively modest fraction of the spin of the proton [25, 26]. We now understand it to be a combination of valence quark, sea quark, and gluon helicities, as well as orbital angular momenta.

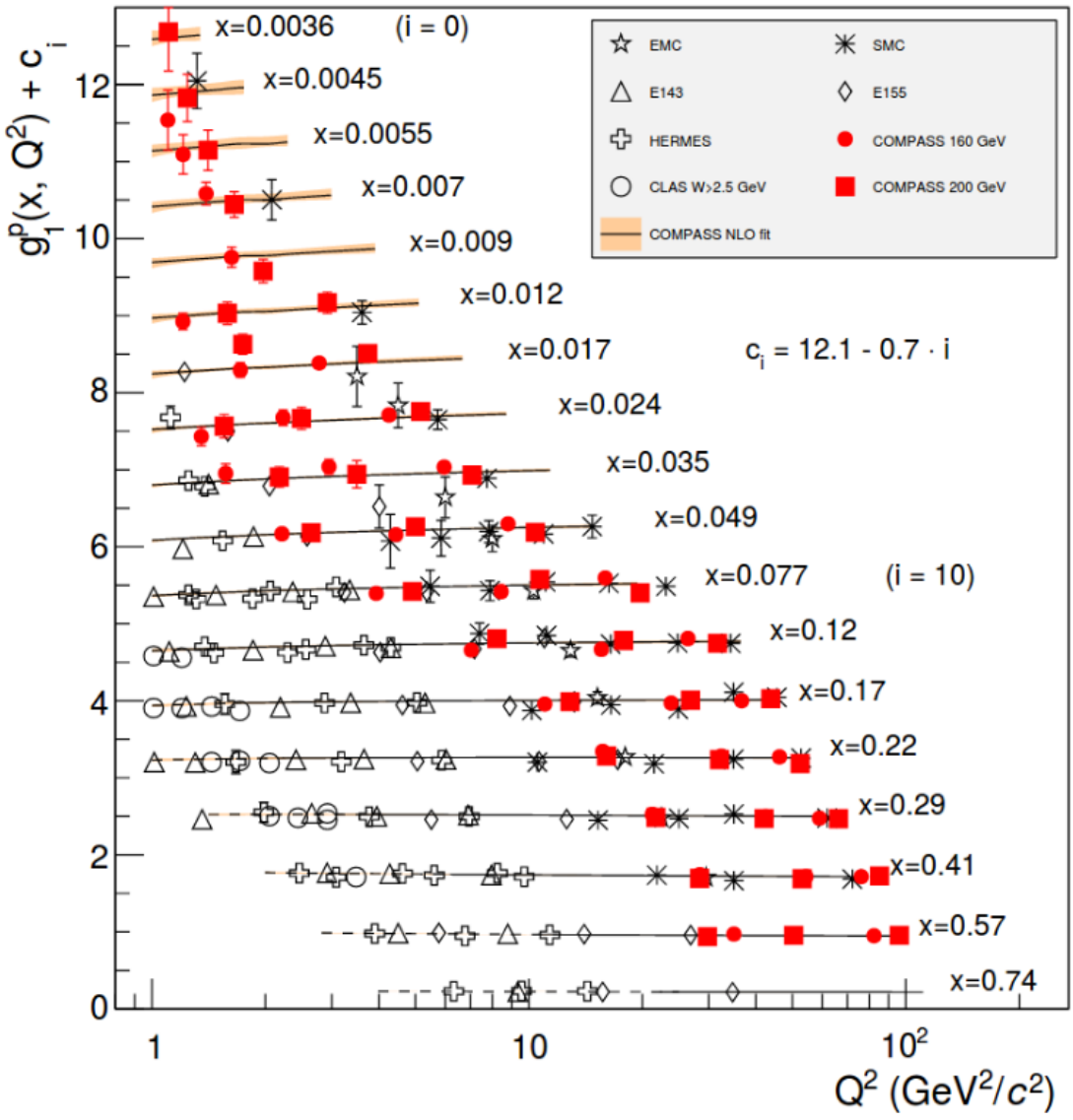

Figure 1.11: Polarized structure function g_1 for the proton, compiled from global measurements. Figure taken from Ref. [12].

Similar to the helicity distributions are the transversity distributions ($\delta q(x)$) which describe the difference in quark PDFs for a quark polarized parallel or antiparallel to the *transverse* proton polarization. The transversity distribution was first introduced by Ralston and Soper in 1978 [27], where the double-polarized Drell-Yan process was proposed to measure such a distribution. Unlike the helicity distribution, δq does not

have an interpretation in the parton model. Furthermore, transversity is a chiral odd distribution — in the helicity basis, it requires a helicity flip. Since scattering processes must conserve helicity, chiral-odd functions should appear in pairs for allowed processes in order to be overall chiral-even. The transversity distribution coupled with another chiral-odd fragmentation function provides a theoretical mechanism for generating large transverse single-spin asymmetries (see Section 1.4.2 for details). An important result relating the transversity, helicity, and unpolarized quark PDFs, known as the "Soffer bound" is shown in Eq. 1.25,

$$q(x) + \Delta q(x) \geq |2\delta q(x)|. \tag{1.25}$$

Unlike g_1, the situation is not so clear for g_2. If the quark mass is included in the partonic cross section calculation, one arrives at the following expression when considering free quarks in the final state, but off mass shell quarks in the initial state due to the binding potential,

$$d\hat{\sigma} \propto L_{\mu\nu}\hat{w}^{\mu\nu} \tag{1.26}$$

$$\hat{w}^{\mu\nu} \supset (2\epsilon^{\mu\nu\alpha\beta})(m_q s_\alpha)\left[\left(1 - \frac{m}{m_q}\right)p_\beta - \frac{m}{m_q}q_\beta\right]. \tag{1.27}$$

Equation 1.27 is the antisymmetric portion of the quark tensor $\hat{w}^{\mu\nu}$ (analogous to the hadronic tensor from Eq. 1.9), which contains all relevant terms for the calculation of Eqs. 1.21 and 1.22. It should be noted that this result does not satisfy the Ward-Takahashi identity [28, 29], and is therefore not gauge invariant unless $m = m_q$ or the quark and hadron are longitudinally polarized [23]. The parton model enforces the condition $m = m_q$ by requiring the initial state quark is also free. Following this, in comparing Eq. 1.27 to the antisymmetric portion of Eq. 1.9, it becomes evident that $g_2 = 0$ for quarks in the parton model, as there are no terms present analogous to the g_2 prefactors of the hadronic tensor. This illustrates the breakdown of the notion of free partons in processes involving transversely polarized hadrons, an important

topic that will be revisited in Section 1.4.

In a full QCD approach, the concept of non-interacting partons must be relaxed to consider the influence of the surrounding color fields within the proton on the struck quark. In the case of transversely polarized hadrons, contributions to the cross section are significantly smaller than the longitudinal case, requiring the inclusion of power-suppressed terms and those that break gauge invariance [23]. As shown by Efremov and Teryaev [30], gauge invariance can be restored through the inclusion of twist-3 contributions from a diagram involving an extra gluon exchange between the struck quark and the proton remnant, shown in Figure 1.12. This allows for small contributions to $g_2(x)$ arising from twist-3 quark-gluon correlations.

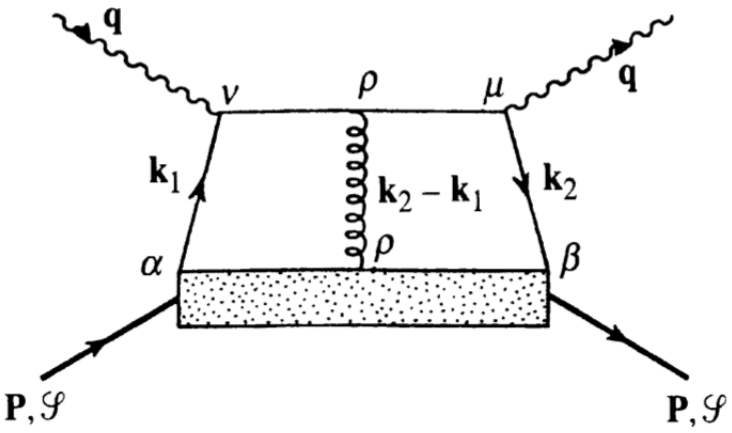

Figure 1.12: Depiction of the twist-3 contributions needed to interpret g_2 from Eq. 1.22. Figure taken from Ref. [23].

In a similar spirit, Wandzura and Wilczek proposed an approximate relationship between g_1 and g_2 at leading twist (twist-2) in the Bjorken scaling regime in 1977 [31],

$$g_2^{WW}(x) = -g_1(x) + \int_x^1 \frac{dx'}{x'} g_1(x'). \tag{1.28}$$

This has since been expanded into a more general form including higher order twist-

3 contributions (lacking $1/Q$ suppression relative to the twist-2 part), allowing for relatively clean extractions of twist-3 quark-gluon correlations [32],

$$g_2(x, Q^2) = g_2^{WW}(x, Q^2) + \bar{g}_2.(x, Q^2), \tag{1.29}$$

where twist-3 part is represented by $\bar{g}_2(x, Q^2)$.

1.3 Hadron Formation

In modern collision experiments, hadrons make up the majority of the final state particles, yet the mechanisms by which they are formed are still poorly understood. Two popular models to describe the hadronization process are known as the *Lund string fragmentation* model [33, 34] and the *cluster hadronization* model [35]. The former considers successive ruptures of color flux tubes, as shown in Fig. 1.6, while the latter considers the grouping of outgoing partons nearby in phase space. The two models yield considerably different predictions for a number of observables related to hadron and jet production (see for example, Ref. [36]).

In practice, the hadronization process is parametrized from data in nonperturbative functions known as fragmentation functions (FFs) introduced in Section 1.2.2. Collinear FFs describe the probability of a specific type of parton producing a hadron of a specific type with momentum fraction z collinear to the outgoing parton direction. Similar to PDFs, fragmentation functions can also depend on transverse momentum (in this case, of hadrons with respect to the parton momentum direction — see Section 1.4.2 for details). Such FFs are referred to as transverse-momentum-dependent FFs (TMD FFs). Fragmentation functions are the final-state analogue of parton distribution functions, with best constraints coming from e^+e^- experiments, as discussed in Section 1.2.2.3. A recent review of fragmentation functions for light quarks and gluons can be found in Ref. [37].

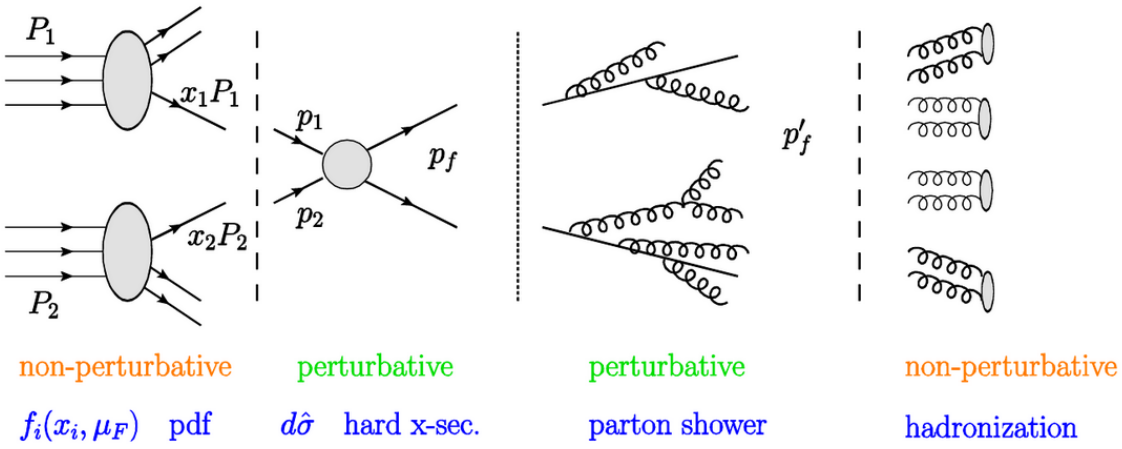

Figure 1.13: Schematic of a factorized scattering process with hadrons in the initial and final state, separately emphasizing the parton shower and hadronization steps. Figure taken from Ref. [38].

In a factorized picture, outgoing partons from hard scattering processes can be considered independently, and undergo radiative energy loss referred to as a *parton shower*. As the primary outgoing color charge continues to lose energy through radiation of gluons and successive splitting of gluons into quark-antiquark pairs, it evolves to an energy scale where the strong coupling is large, resulting in a transition of partonic degrees of freedom to hadronic degrees of freedom, a process known as *hadronization*. This is schematically depicted in Figure 1.13. The hadrons are ultimately measured in the tracking and calorimeter detectors of modern experiments in collimated sprays known as *jets*.

1.3.1 Parton Shower

Outgoing partons from hard scattering processes, like the struck quark in a DIS process, radiate a shower of color charges in accordance with the Alterelli-Parisi splitting functions [22]. The probability of a $1 \rightarrow 2$ splitting process for partons with separation angle θ is described by Eq. 1.30 [39].

$$dP_{i\rightarrow jk} = \frac{d\theta}{d} z P_{i\rightarrow jk}(z), \tag{1.30}$$

where z is the momentum fraction of the radiated parton with respect to the parent parton, and $P_{i\to jk}$ are the Altarelli-Parisi splitting functions. An important feature of this equation is the result for soft gluon radiation from a hard quark or gluon in the high-energy limit,

$$dP_{i\to ig} \approx \frac{2\alpha_s C_i}{\pi}\frac{d\theta}{\theta}\frac{dz}{z}. \tag{1.31}$$

The splitting functions are important in defining the radiation pattern of partons ejected from their parent hadrons, a subject of recently increasing interest to the QCD community. In 2017, a study was conducted using CMS Open Data to "expose" the QCD splitting functions [39], suggesting that similar meaningful studies can be conducted on the newly released LHCb Open Data (see Section V), particularly suited to explore the heavy quark sector. An important consequence of the quark mass for heavy quarks is a characteristic suppression of soft gluon radiation, coined as the "dead-cone effect" [40]. In the case of a heavy quark radiating soft gluons, the splitting probability from Eq. 1.31 is modified as follows,

$$dP_{i\to ig} \approx \frac{2\alpha_s C_i}{\pi}\frac{\theta^2 d\theta^2}{(\theta^2+\theta_{HQ}^2)^2}\frac{dz}{z}. \tag{1.32}$$

The dead-cone effect was first measured by ALICE [41] and is currently under investigation at LHCb.

1.3.2 Jets in QCD

Jets are collimated sprays of particles arising from outgoing partons in hard scattering events. They are defined by the algorithms used to cluster constituent particles into a composite object. The most notable algorithm is the "anti-k_T" algorithm, providing a method for experimental jet clustering that is free of infrared and collinear divergences [42]. Since the development of the anti-k_T algorithm, there has been an abundance of jet-related measurements, with much interplay between the experimen-

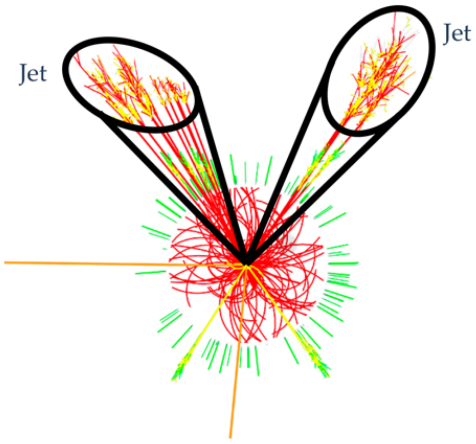

Figure 1.14: Cartoon of a detector cross sectional view with 2 measured jets, back to back with 2 high energy leptons. The jets are reconstructed from measured tracks and energy deposits in the calorimeters.

tal and theoretical communities. One particularly powerful aspect of jets is that they serve as a proxy for ejected partons in high-energy collisions, allowing for measurements that are approximately free from final-state fragmentation effects. This also implies that they serve as a map between an initially ejected parton and the set of corresponding collimated hadrons that make up the jet. In broader terms, they can be used as a composite object to cleanly extract information about hadron structure and initial-state effects, or they can be used as a laboratory to study final-state fragmentation and hadronization effects. Figure 1.14 shows a cross-sectional cartoon of a detector with jets reconstructed from measured tracks and calorimeter clusters.

1.3.3 Probing Hadronization

Many methods have been proposed to study the hadronization process by exploring the substructure of jets. One successful method involves measuring correlations between jets and their constituent particles. Some useful variables for such studies are the following:

$$z = \frac{\vec{p}_h \cdot \vec{p}_{jet}}{|\vec{p}_{jet}|^2} \tag{1.33}$$

$$j_T = \frac{|\vec{p}_h \times \vec{p}_{jet}|}{|\vec{p}_{jet}|} \tag{1.34}$$

$$r = \sqrt{(\eta_h - \eta_{jet})^2 + (\phi_h - \phi_{jet})^2}, \tag{1.35}$$

where the h subscript refers to hadron. Equations 1.33 and 1.34 are useful in constraining transverse momentum-dependent fragmentation functions (see Section 1.4.2 for details), while Eq. 1.35 is useful in studying how hadrons are spatially distributed within jets. Additionally, z from Equation 1.33 can be used to constrain collinear FFs. In the limit that the jet is a perfect proxy for the hard scattered parton, i.e. $\vec{p}_{jet} = \vec{p}_{parton}$, then z and j_T correspond exactly to variables appearing in collinear and TMD FFs. In practice, there are more subtleties in interpreting the experimental data (see Refs. [43, 44] for details). Hadrons are accessed via measurements of charged particles in the case of inclusive jets from ATLAS [45], Z-tagged jets from LHCb [46], and heavy flavor tagged jets from LHCb [47]. These data provide constraints on the hadronization process for gluon, light quark, and heavy flavor quark-initiated jets, respectively. For the Z-tagged jets at LHCb, there also exist measurements with identified charged hadrons, providing sensitivity to flavor-dependent TMD FFs [48].

1.3.4 Transverse Spin in Hadronization

It was predicted by Collins in 1992 that the transverse polarization of a fragmenting quark can generate an asymmetry in the angular distribution of produced hadrons with respect to the parton momentum direction [49]. Such an effect is described theoretically with a chiral odd TMD FF that couples to the transversity PDF discussed in Section 1.2.3. Collins originally proposed to measure this in semi-inclusive DIS processes (SIDIS), where a final state hadron is measured in addition to the scattered electron, and pp collisions with hadrons that are produced approximately back-to-back in ϕ. More details about the Collins asymmetry can be found in Section 1.4.2.

On the other hand, it has been observed experimentally for decades in a number

of collision systems that Λ^0 baryons and other hyperons (baryons containing a strange quark) can be spontaneously polarized transverse to their momentum in the process of hadronization [50–58]. This has been observed in unpolarized collisions, implying an existing mechanism for unpolarized quarks to spontaneously obtain polarization during hadronization, or for polarized quarks in unpolarized hadrons to transfer their polarization to the final state Λ^0.

1.4 Transverse Single-Spin Asymmetries

Transverse single-spin asymmetry (TSSA) measurements quantify azimuthal modulations in particle production due to spin-momentum and spin-spin correlations in the polarized proton or in the process of hadronization. The initial state collision system can be $p^{\uparrow,\downarrow}(e,\ p,\ \mathrm{A})$, where the polarized protons have a non-negligible polarization fraction and the spin is transverse to their momentum direction in the center-of-mass frame. The key observable, A_N is defined in Eq. 1.36, with A denoting an asymmetry and the subscript N representing a spin that is normal to the momentum direction for the polarized proton.

$$\epsilon_N(\phi) = \frac{\sigma^{\uparrow}(\phi) - \sigma^{\downarrow}(\phi)}{\sigma^{\uparrow}(\phi) + \sigma^{\downarrow}(\phi)} = A_N P \cos\phi, \tag{1.36}$$

where $\sigma^{\uparrow,\downarrow}(\phi)$ are cross sections measured as a function of ϕ for spin up and down initial state protons respectively, and ϕ is the azimuthal angle. The polarization fraction of the initial state proton beam is expressed as P, and it dilutes the analyzing power. More generally, a cross section that is differential in ϕ can be parametrized in terms of the beam polarization angle ϕ_{pol}, P, and an asymmetry amplitude A_N as follows,

$$\frac{d\sigma}{d\phi}^{\uparrow,\downarrow} = \left(\frac{d\sigma}{d\phi}\right)_0 \left(1 + A_N P \sin(\phi_{pol} - \phi)\right), \tag{1.37}$$

where $\left(\frac{d\sigma}{d\phi}\right)_0$ represents the unpolarized differential cross section as a function of ϕ. A "spin up" proton bunch has a polarization vector aligned with the y axis, corresponding to $\phi_{pol} = \pi/2$, while a "spin down" proton bunch corresponds to $\phi_{pol} = -\pi/2$. Keeping this in mind, and plugging Eq. 1.37 into Eq. 1.36, one obtains the right-hand side of Eq. 1.36.

As mentioned in Section 1.2.3, observables involving hadrons with transverse spin provide a rigorous test of QCD, as their description in a collinear factorization scheme requires the relaxation of approximations such as assuming free partons and leading power factorization. Power-suppressed terms in the factorization expansion must be included, arising from interactions of the scattered parton with the proton remnant. The situation is no different here — numerous measurements of TSSAs in hadron-hadron scattering experiments show large magnitudes of asymmetries up to 40% at increasing $x_F = 2p_z/\sqrt{s}$ across a broad range of energies [59–62] (see Figure 1.15). This is contrary to the prediction that such asymmetries from the hard scattering of quarks should be very small ($< 1\%$) in Ref. [63]. Although recent calculations suggest small additional contributions from the perturbative expansion at two loops [64], these contributions can not explain the large magnitudes of A_N observed in Figure 1.15. The observation of large TSSAs can in fact not be explained with a leading twist collinear factorization theorem, which can be remedied by going to sub-leading twist in the factorization expansion or by modifying the definition of PDFs and FFs to include transverse momentum dependence to describe the internal motion of partons. These are referred to as the twist-3 factorization approach [30, 65] and the transverse-momentum-dependent (TMD) factorization approach [49, 66, 67] (see Ref. [68] for a recent review). Both frameworks are successful in modeling TSSAs in complementary regions of p_T [69–71], and are useful in constraining orbital angular momentum of quarks and gluons in protons [72–74]. Recent phenomenological arguments indicate that the separate approaches share a common origin in multiparton correlations [75].

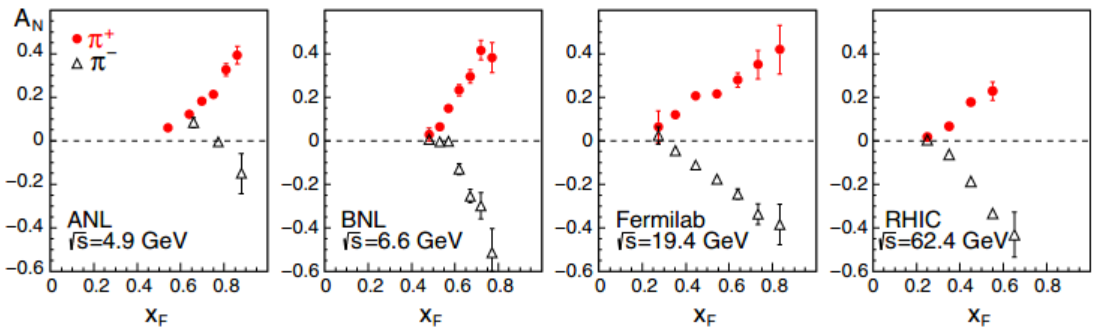

Figure 1.15: TSSAs in hadron-hadron collisions for charged pion production at various energies, showing a large effect (up to 40%) at increasing Feynman-x $x_F = 2p_z/\sqrt{s}$. Figure taken from Ref. [24].

1.4.1 Sub-leading Twist (twist-3)

The higher twist approach can be used to describe observed TSSAs with collinear factorization, meaning that the only relevant scale is the hard scale of the interaction Q (e.g. the momentum transfer in the DIS process). Twist is defined in two ways; calculating an observable to twist-n implies the inclusion of power-suppressed terms in the collinear factorization expansion to order $1/Q^{n-2}$. Leading-twist or twist-2 calculations therefore are not power suppressed by the hard scale, but do not generate observed TSSAs. More formally, the twist of an operator is defined as its canonical (mass) dimension minus its spin dimension. For example, consider the field-theoretic definition of quark and gluon PDFs [8, 76],

$$q(x) \propto \langle P|\bar{\psi}_q \frac{\gamma^+}{2} \psi_q|P\rangle \qquad g(x) \propto \langle P|G^{+\mu}\mathcal{W}G_{+\mu}|P\rangle \tag{1.38}$$

where $\mathcal{W}$ is a Wilson line (see Refs. [8, 77, 78] for details). It can be seen by analyzing Eq. 1.2 that the mass dimension of quark fields ψ and $\bar{\psi}$ is 3/2, while for the gluon field strength tensor it is 2. On the other hand, the spin of a quark is 1/2, while the spin of a gluon is 1. Combining the mass and spin dimensions for each field in the expression and taking the difference yields $3-1$ for $q(x)$ and $4-2$ for $g(x)$, both resulting in a twist of $n = 2$. Twist-3 correlators and beyond require the inclusion of

additional fields, implying a relationship to multiparton correlations within hadrons. For example, consider the following operators:

$$\mathcal{O}_{qgq} \propto \langle P|\bar{\psi}_q \gamma^+ G_{+\mu} \psi_q|P\rangle \tag{1.39}$$

$$\mathcal{O}_{ggg} \propto \langle P|G^{+\nu} G_{+\mu} G_{+\nu}|P\rangle. \tag{1.40}$$

A similar analysis yields $n = 5 - 2 = 3$ and $n = 6 - 3 = 3$ for Eqs. 1.39 and 1.40, respectively. Quark-gluon-quark (qgq) and trigluon (ggg) correlation functions are therefore twist-3 operators. More formally, these are known as the Efremov-Teryaev-Qiu-Sterman (qgq) correlator [30, 79], and the trigluon or three-gluon (ggg) correlators [80–84] respectively. The notion of trigluon correlation functions was introduced in Ref. [80] and they were subsequently clarified to be two independent functions [81–84]. The qgq correlator is much more constrained from data, especially given the fact that they are accessible in $ep^\uparrow$ collisions. Constraints from global fits are discussed in Ref. [75]. On the other hand, the ggg correlators have received less attention, with the best constraints coming from the Relativistic Heavy Ion Collider (RHIC) given that gluons are accessible at leading order in $p^\uparrow p$ collisions. Available data for constraining the trigluon correlation functions are discussed in Section 4.1.2, including work presented in this dissertation on the midrapidity open heavy flavor TSSA measurement from the $\sqrt{s} = 200$ GeV 2015 PHENIX $p^\uparrow p$ dataset.

The origin of twist-3 correlation functions can be interpreted as the quantum mechanical interference between a leading twist process and one with an extra gluon connecting the hard scattering region to one of the nonpeturbative regions in the square modulus of the scattering amplitude. This is illustrated in Fig. 1.16, where the left hand side shows the square modulus of the scattering amplitude (the dashed line in the middle separates $\mathcal{M}$ and $\mathcal{M}^*$), and the right hand side shows the factorization of such a process, including an illustration of the qgq correlator.

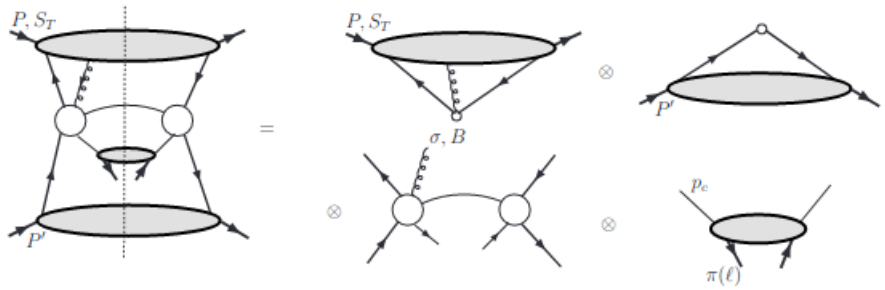

Figure 1.16: Schematic of the twist-3 factorization of a process where the qgq correlator is considered in the polarized proton (see text). Figure taken from Ref. [85].

Figure 1.17 shows a cartoon for the $p^{\uparrow}p \to hX$ process with labels A, B, and C for $p^{\uparrow}, p$, and h, respectively. The leading twist factorization for the cross section of this process with unpolarized hadrons is shown in Section 1.2.2.3. Considering the transverse polarization of one of the initial state protons complicates things. In order to generate non-negligible asymmetries within collinear factorization, one must consider twist-3 corrections. A factorization theorem for the TSSA (A_N) is shown in Eq. 1.41 [85, 86], with labels A, B, and C in the subscript corresponding to the hadrons in Fig. 1.17.

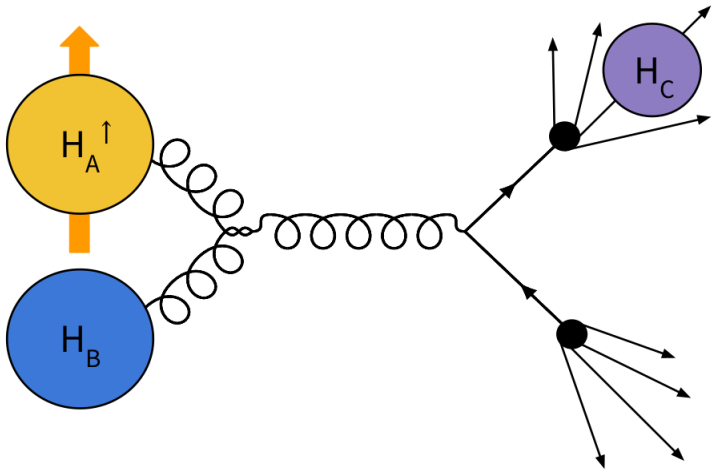

Figure 1.17: Cartoon depiction of a $p^{\uparrow}p \to hX$ scattering process in which the part of the partonic cross section shown is the $gg \to q\bar{q}$ channel.

,

$$
\begin{aligned}
A_N \propto & \sum_{a,b,c} \phi^{(3)}_{a/A}(x_1, x_2, \vec{s}_\perp) \otimes f_{b/B}(x') \otimes \hat{\sigma} \otimes D^C_c(z) \\
& + \sum_{a,b,c} \delta q_{a/A}(x, \vec{s}_\perp) \otimes \phi^{(3)}_{b/B}(x'_1, x'_2) \otimes \hat{\sigma}' \otimes D^C_c(z) \\
& + \sum_{a,b,c} \delta q_{a/A}(x, \vec{s}_\perp) \otimes f_{b/B}(x') \otimes \hat{\sigma}'' \otimes D^{C(3)}_c(z_1, z_2)
\end{aligned}
\tag{1.41}
$$

The (3) in the superscript corresponds to twist-3 correlators, while the remaining nonperturbative functions correspond to standard collinear twist-2 PDFs and FFs. The 3 independently factorized and summed terms in Eq. 1.41 arise from the 3 hadrons in the reaction, each of which has a contribution from a twist-3 correlator. Note that Fig. 1.16 shows this only for the first term, where the twist-3 correlator in the polarized proton ($\phi^{(3)}_{a/A}(x_1, x_2)$) is considered. This is known as the "Sivers like" twist-3 correlator, as it has a correspondence with the Sivers TMD PDF [66] introduced in Section 1.4.2. Similarly, the second and third factorized and summed terms contain the "Boer-Mulders like" and "Collins like" twist-3 correlators, related

to the Boer-Mulders TMD PDF [67] and the Collins TMD FF [49] (see Section 1.4.2). Both the Boer-Mulders and Collins twist-3 correlators couple to the transversity PDF ($\delta q_{x/X}(x, \vec{s}_\perp)$) introduced in Section 1.2.3.

The dependence of twist-3 correlation functions on two separate momentum fractions necessarily introduces complications, as it is difficult to extract multivariable functions from data. However, the contribution to TSSAs from twist-3 correlators arises from the so-called "soft gluon pole" (SGP) that occurs when $x_1 = x_2 = x$ [87], allowing for an extraction in terms of one variable, the collinear momentum fraction x. Furthermore, twist-3 correlators at a fixed x can be expressed in terms of k_T moments of their corresponding TMDs [69], allowing for information to be extracted about TMDs even in a case where collinear factorization applies and promoting synergy between the two frameworks. The twist-3 approach is often more suitable for measurements in pp collisions, as only one scale is typically measured. In this case, the p_T of measured particles is taken as a proxy for the hard scale Q. Given that it must be smaller than Q, a sufficiently high p_T should indicate a sufficiently hard interaction.

1.4.2 Transverse Momentum Dependent PDFs and FFs

The transverse momentum-dependent (TMD) formalism is often considered more intuitive as the observed asymmetries can be explained at leading twist. In this approach, internal dynamics related to interactions within the hadronic bound state are parameterized with the transverse momentum of partons (k_T) with respect to the proton momentum direction in the initial state, or the transverse momentum of hadrons (j_T) with respect to the parton momentum direction in the final state, rather than correlations between multiple partons within the hadrons. TMD factorization is applied by extending PDFs and FFs to depend on k_T and j_T, respectively, in the factorization formulae (illustrated in Figure 1.18). This explicit dependence means

that in order to constrain TMDs experimentally, two scales must be measured, a hard scale (Q) and a soft transverse momentum scale (k_T or j_T) with sufficient scale separation $Q \gg k_T$. This can be accomplished in e^+e^- annihilation, semi-inclusive deep inelastic scattering (SIDIS), the Drell-Yan process, or dijet production, for example. The small transverse momentum scale can be obtained from the transverse momentum imbalance of final state particles (q_T) or the transverse momentum of hadrons within final state jets (j_T). Figures 1.19 and 1.20 show tables of the 8 leading twist TMD PDFs and FFs depending on the combination of parton and hadron spin. For the TMD PDFs, the yellow shaded boxes on the diagonal correspond to PDFs that survive integration over k_T and are useful in collinear factorization as well. The Sivers TMD PDF in Figure 1.19 and the Collins TMD FF from Figure 1.20 are discussed in the following.

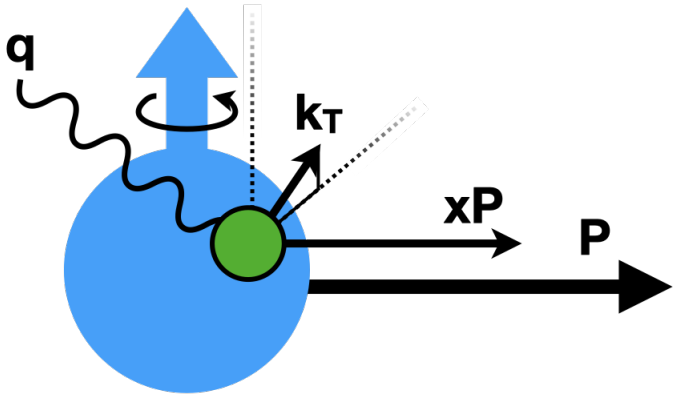

Figure 1.18: Cartoon representation of a parton within a proton that depends on collinear momentum fraction x and transverse momentum with respect to the proton momentum k_T.

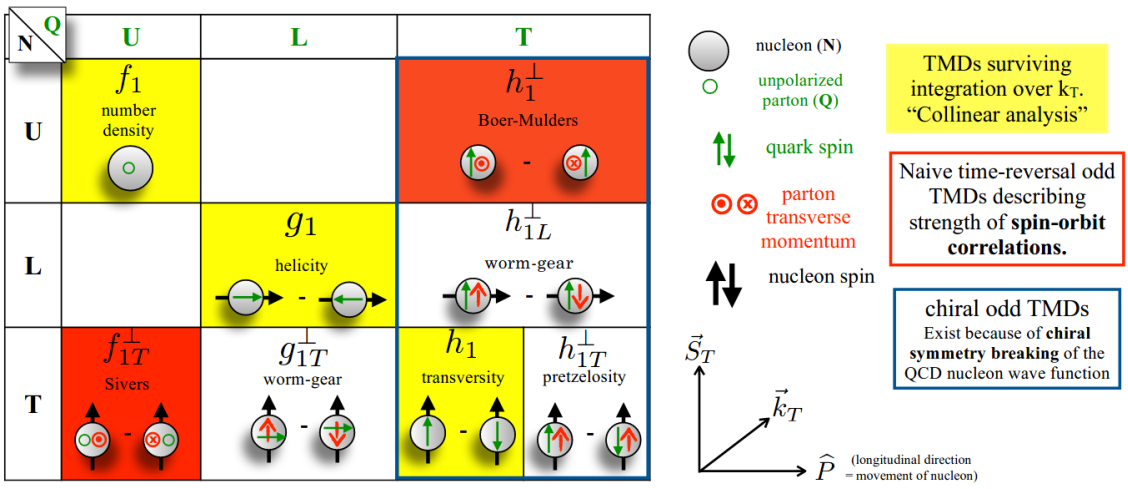

Figure 1.19: Table of 8 transverse momentum dependent parton distribution functions. Figure taken from Ref. [88].

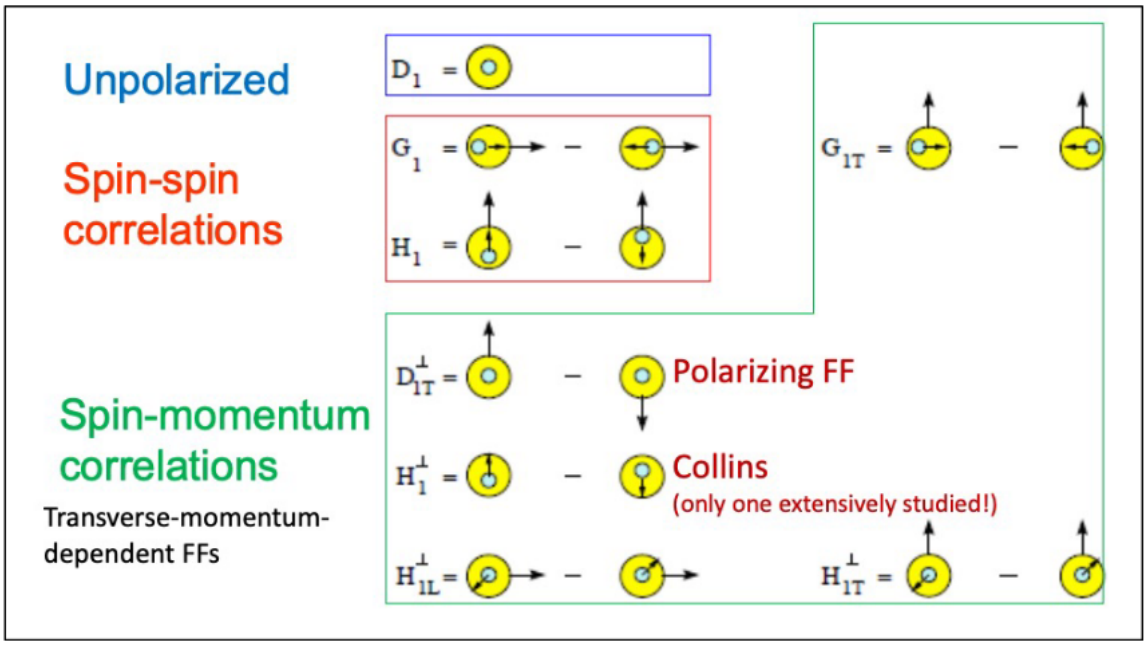

Figure 1.20: Table of 8 transverse momentum dependent fragmentation functions.

The Sivers TMD PDF was proposed by Dennis Sivers in 1990, kicking off the TMD formalism. This is a spin-momentum correlation between a transversely polarized hadron and a parton with transverse momentum k_T, sometimes thought of as a spin-

orbit correlation. It is defined by the triple product $\vec{S}_T \cdot (\hat{P} \times \vec{k}_T)$, where $\vec{S}_T$ is the transverse spin of the hadron, $\hat{P}$ is the hadron momentum direction, and k_T is the parton transverse momentum. The basic idea is that orbital angular momentum of the partons can play a role in the preferred scattering direction. For example, a spin aligned with the vertical can make it so partons in the hemispheres of the protons that are facing each other during the collision have a preferred scattering direction, leading to an azimuthal or left-right asymmetry.

In 1992, John Collins predicted a different mechanism for generating transverse single-spin asymmetries at leading twist, introducing the Collins TMD FF, a chiral odd FF in which a transversely polarized quark generates asymmetric fragmentation [49]. This implies a convolution of the Collins TMD FF with the transversity PDF as a transversely polarized quark is required and the process must be overall chiral even. The Collins effect is defined by the triple product $\vec{s}_T \cdot (\hat{k} \times \vec{P}_{hT})$, where $\vec{s}_T$ is the transverse quark spin, $\hat{k}$ is the quark momentum direction, and $\vec{P}_{hT}$ is the hadron transverse momentum. The azimuthal asymmetry in the distribution of produced hadrons can be intuitively understood through the conservation of angular momentum, as a transversely polarized quark hadronizes into an unpolarized hadron. In the same paper, he suggested that the Sivers asymmetry must be zero due to the time reversal invariant nature of QCD.

10 years later in 2002, Brodsky, Hwang, and Schmidt, revisited Collins' analysis of the Sivers asymmetry and concluded that corrections were needed to comply with gauge invariance. Collins subsequently found that the Sivers asymmetry can indeed be nonzero in the case of soft gluon exchange between the scattered parton and one of the initial state hadron remnants, prompting the prediction of a modified prescription of universality for the Sivers TMD PDF where a relative minus sign is included depending on if this soft gluon exchange facilitates an attractive or repulsive interaction [89]. The Sivers function in SIDIS vs Drell-Yan or similar color-annihilation

processes producing an electroweak boson, therefore, should have a relative minus sign as a consequence, a prediction that is so far favored by the (statistically limited) available data [90, 91].

The quark Sivers function has been confirmed with SIDIS data and extracted through several measurements from COMPASS [92–96] and HERMES [97]. Since gluons are not accessible at leading order in the SIDIS process, the gluon Sivers function remains poorly constrained. However, measurements presented in this dissertation, as well as other measurements from RHIC provide constraints on the trigluon correlation functions in polarized protons, which can be used to indirectly constrain the gluon Sivers function (see Section 4.1.2). The Collins asymmetry has additionally been confirmed in measurements from e^+e^- collisions at Belle [98, 99], BESIII [100], and BaBar [101, 102], SIDIS at HERMES [103] and COMPASS [104, 105], and pp collisions at STAR [106]. While the Sivers and Collins asymmetries have both been confirmed and are becoming better understood with more available data, further investigation is needed to understand the relative importance of the different mechanisms in different kinematic regions, and to explain the origin of the large TSSAs in Fig. 1.15.

1.4.3 TSSAs for Open Heavy Flavor Production

Open heavy flavor production at RHIC is predominantly through gluon-gluon fusion, with only a small contribution from quark-antiquark annihilation (see Section 4.1.2 for details). For gluon-gluon fusion, Eq. 1.41 simplifies tremendously for two reasons. The first is that in spin 1/2 nucleons like the proton, the gluon does not have a transversity distribution, so the second and third summed and factorized terms drop out. The second is that the sum over partons simplifies to only include gluons. For both of these reasons, this production channel provides a clean extraction of the Sivers-like twist-3 correlators for gluons relative to other production channels.

In contrast, light mesons such as π^0 and η are produced through a mixture of channels, and all terms in Eq. 1.41 must be considered, with terms containing the "Sivers like" and "Collins like" twist-3 correlators of particular importance. Both the Sivers and Collins mechanisms (or corresponding twist-3 correlation functions) must be considered, allowing for a determination of the relative importance of both mechanisms to the $p^{\uparrow}p \to hX$ asymmetries for light hadrons.

In Refs. [107] and [87] the $p^{\uparrow}p \to DX$ production channel was considered at $\sqrt{s} = 200$ GeV, and contributions to the TSSAs from twist-3 correlation functions were calculated. In Ref. [107], both $q\bar{q}$ annihilation and gg fusion were considered, and hence contributions from qgq and ggg correlators. It was determined that the contribution from qgq correlators was negligible to TSSAs for D^0 production at midrapidity, and the authors claimed that a non-zero TSSA measured in this channel at RHIC would be a clear indication of trigluon correlations in polarized protons. In this case, the trigluon correlation functions are defined as $T_G^{(f)}(x,x)$ (antisymmetric) and $T_G^{(d)}(x,x)$ (symmetric), where the superscript (f) and (d) s represent cases where the three gluon-field color indices contract with antisymmetric (f_{abc}) or symmetric (d_{abc}) structure tensors. Phenomenological models were introduced to describe the trigluon correlators following the lack of direct information on the functions. The models describe the ggg correlators with normalization parameters to the unpolarized gluon PDF. Following the argumentation from Ref. [86], parameters λ_f and λ_d are introduced as follows,

$$T_G^{(f)}(x,x) = \lambda_f g(x), \qquad T_G^{(d)}(x,x) = \lambda_d g(x). \tag{1.42}$$

In Ref. [87] only contributions from trigluon correlators were considered, as it was established that the qgq contribution should be negligible. For this analysis, trigluon correlation functions are instead defined as $N(x_1,x_2)$ (antisymmetric), and $O(x_1,x_2)$ (symmetric), yielding four independent contributions to the predicted

TSSAs, $\{N(x,x), N(x,0), O(x,x), O(x,0)\}$. Ref. [87] shows that at $\sqrt{s} = 200$ GeV the asymmetries depend on effective trigluon correlators $N(x,x) - N(x,0)$ and $O(x,x) + O(x,0)$, for which a direct mathematical relation to $T_G^{(f)}$ and $T_G^{(d)}$ exists in Ref. [83]. In a similar fashion, phenomenological models are employed to describe the ggg correlators in terms of normalization parameters K_G and K'_G, and the unpolarized gluon PDF $g(x)$. The parameters are introduced in Ref. [87] with the following assumptions:

$$O(x,x) = O(x,0) = N(x,x) = -N(x,0) \tag{1.43}$$

$$\text{[Model1]} \quad O(x,x) = K_G x g(x) \tag{1.44}$$

$$\text{[Model2]} \quad O(x,x) = K'_G \sqrt{x} g(x) \tag{1.45}$$

Assumptions introduced for the trigluon correlators in Eqs. 1.42, 1.44, and 1.45 (e.g., the functional dependence on x and the proportionality to the unpolarized gluon PDF) are oversimplified, but again are made due to the lack of direct information on the functions to give the community some footing. It is therefore useful to compare to different models with various x dependencies. An analysis to constrain these parameters ((λ_f, λ_d), K_G, and K'_G) from the midrapidity open heavy flavor electron A_N measurement presented in Ref. [1] and this dissertation is outlined in Chapter IV.

The TMD factorization approach has also been used to investigate TSSAs in open-charm production at RHIC (see Refs. [108–110]). However, open heavy flavor measurements in $p^{\uparrow}p$ collisions can only provide indirect constraints on the gluon Sivers PDF by first directly constraining the twist-3 trigluon correlators in polarized protons and then relying on the equivalence to k_T moments of the gluon Sivers PDF [69].

1.4.4 TSSAs in Proton-Nucleus Collisions

TSSA measurements in collisions involving a heavy nucleus have been proposed as a means to investigate the interplay between spin physics and low-x physics. It has been suggested that measuring the ratio of A_N in $p^\uparrow A$ collisions to that in $p^\uparrow p$ collisions for forward hadron production can provide sensitivity to gluon saturation in the nucleus [111]. This has been expanded on in several theoretical works [112–122]. The general idea is captured in Eq. 1.46:

$$\left.\frac{A_N^{pA\to h}}{A_N^{pp\to h}}\right|_{p_{h\perp}^2 \ll Q_s^2} \approx \frac{Q_{sp}^2}{Q_{sA}^2} e^{\frac{p_{h\perp}^2 \delta^2}{Q_{sp}^4}} \qquad \left.\frac{A_N^{pA\to h}}{A_N^{pp\to h}}\right|_{p_{h\perp}^2 \gg Q_s^2} \approx 1, \tag{1.46}$$

where Q_{sp} is the saturation scale in protons and Q_{sA} is the saturation scale in the heavy nucleus. It was assumed in the calculation that $Q_{sp} \ll Q_{sA}$, and both are much larger than δ, which comes from the Gaussian width assumed for the Collins FF.

Data from the 2015 RHIC dataset with $\sqrt{s} = 200$ GeV $p^\uparrow$Al and $p^\uparrow$Au collisions are very useful in investigating such questions, as it is the highest energy $p^\uparrow A$ data available, and the only collider dataset where such ideas can be investigated. The existing measurements from this data have so far yielded surprises and painted an inconsistent picture, making it difficult to draw conclusions (see Section 4.2.1 for details). An important point is that light hadron production in different rapidity regions shows inconsistent nuclear modification of A_N. With the recent measurement of midrapidity π^0 and η TSSAs in $\sqrt{s} = 200$ GeV $p^\uparrow p$ collisions [123], an opportunity was presented to measure the same observable in the $p^\uparrow A$ dataset to determine if any nuclear modifications were observed, in hopes of shedding light on this mystery. The results of this measurement are presented in Chapter IV.

CHAPTER II

Experimental Setup

The unique capabilities of the Relativistic Heavy Ion Collider (RHIC), including colliding polarized protons on polarized protons or a heavy nucleus, allow for the exploration of TSSAs in $p^{\uparrow}p$ and $p^{\uparrow}A$ collisions. This provides access to gluons at leading order in perturbation theory and allows for the exploration of parton-hadron spin-momentum correlations, even in collisions with a heavy nucleus. Furthermore, the PHENIX detector is ideal for measuring electromagnetic probes such as photons and electrons given the high granularity electromagnetic calorimeter, and the Ring Imaging Cherenkov Detector that allows for separation of electrons and charged pions. This is utilized to measure electrons from the decays of heavy flavor hadrons in $p^{\uparrow}p$ collisions, and the decays $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$ in $p^{\uparrow}A$ collisions. The corresponding TSSAs provide sensitivity to gluon correlations in polarized protons from the $p^{\uparrow}p$ data, and modifications from the more intense collision environment from the $p^{\uparrow}A$ data.

2.1 The Relativistic Heavy Ion Collider

The Relativistic Heavy Ion Collider (RHIC) is an accelerator complex located on the campus of Brookhaven National Lab in Long Island, New York. It contains two independent accelerator rings with a 3.83 km circumference and co-propagating

beams labeled "blue" (clockwise) and "yellow" (counterclockwise) that cross at six interaction points. RHIC came online in 2000 with experimental facilities PHENIX, STAR, BRAHMS, and PHOBOS at 4 of the interaction points [124]. Both BRAHMS and PHOBOS were designed to take data for the first several years of RHIC operation and finished running by 2006. PHENIX and STAR were designed as general-purpose detectors that would operate for a longer lifetime. PHENIX was decommissioned in 2016 and has been superseded by the sPHENIX experiment that saw its first collisions this year [125]. Both sPHENIX and STAR plan to take data until 2025, after which RHIC will be re-instrumented to prepare the Electron-Ion Collider facility [126, 127]. Figure 2.1 shows the arrangement of RHIC detectors along the rings during the start of its operation [124].

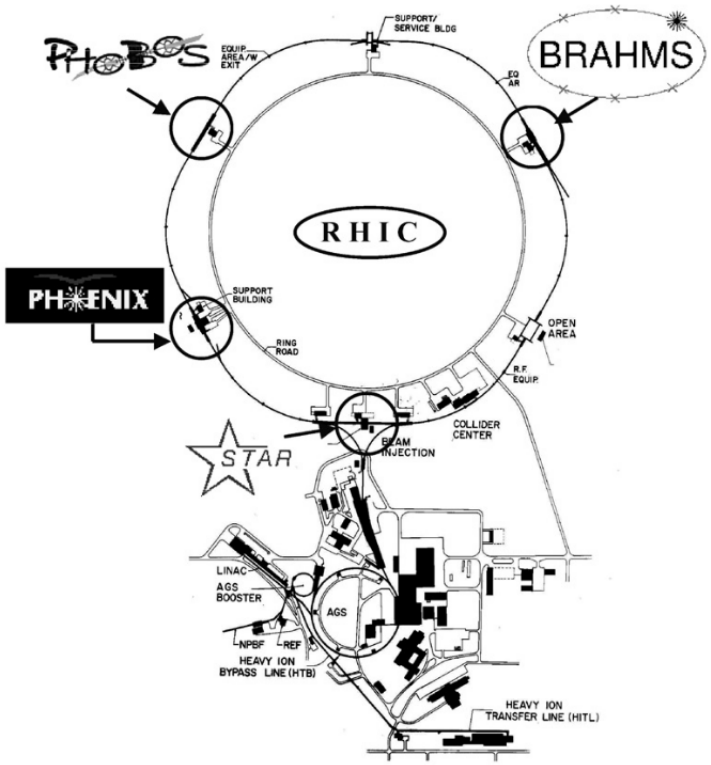

Figure 2.1: A map of the RHIC accelerator complex highlighting the experimental facilities PHENIX, STAR, PHOBOS, and BRAHMS that came online at the beginning of RHIC operation. This figure was taken from Ref. [124].

RHIC is an incredibly versatile accelerator/collider facility in that it has the ability to (*i*) collide polarized protons, (*ii*) collide nuclei with a wide variety of atomic numbers A, and (*iii*) host collisions across a wide range in energy. This is summarized nicely in Figure 2.2, with the red columns representing collisions with polarized protons, the blue columns representing collisions with heavy ions, and the striped columns representing asymmetric collision systems.

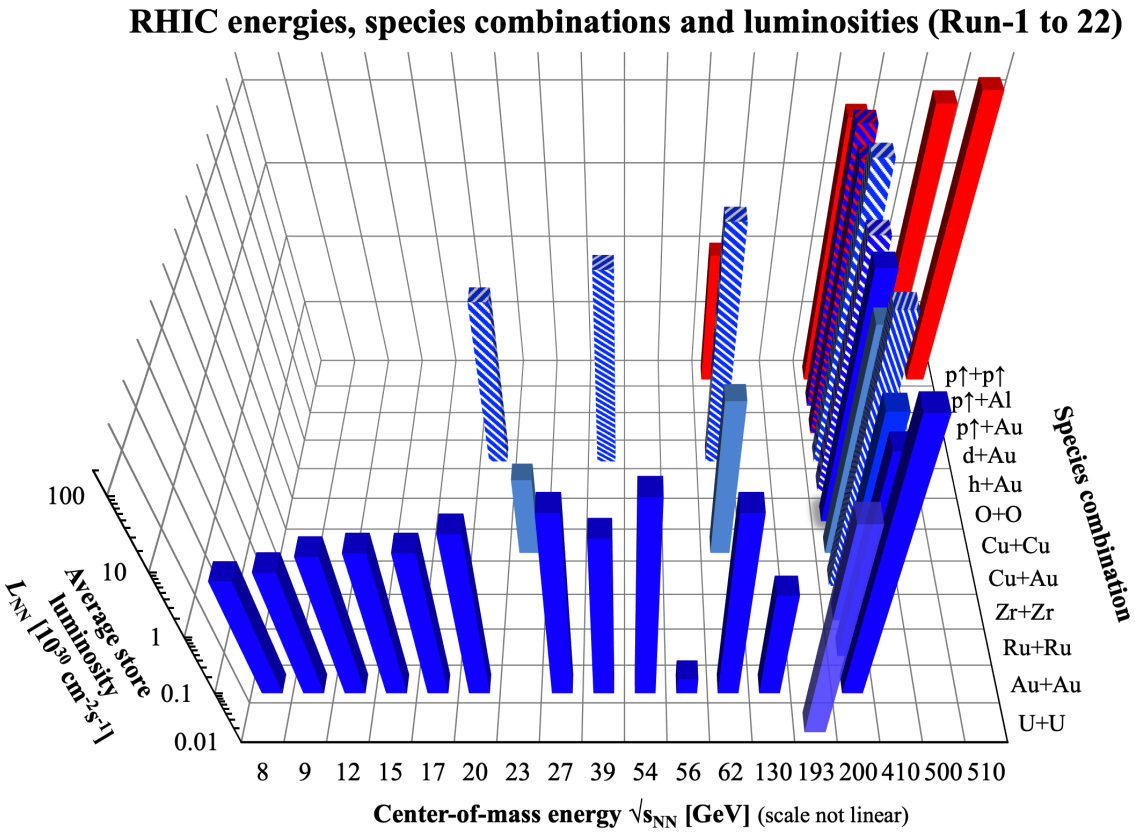

Figure 2.2: Column chart depicting the average store luminosity as a function of center of mass energy and collision system for RHIC Run 1 through 22.

Circular accelerators operate by steering the beam trajectory with magnetic fields (typically created with dipole magnets) while using radio frequency (RF) electric fields to accelerate charged particles. Charged particles experience a longitudinally oscillating voltage when passing through RF cavities at various locations along the

rings and form bunches due to the oscillating nature of the electric field. The number of bunches that can be stored in a circular accelerator is dependent on an important design parameter known as the RF harmonic number h in $f_{RF} = hf_{rev}$, where f_{RF} is the radio frequency, and f_{rev} is the revolution frequency. As the bunches are accelerated, the radio frequency and magnetic fields need to be adjusted accordingly to keep the beam bunched and in a stable orbit. The goal of accelerating the particles is to collide them at the desired energies. In order to optimize the number of collisions, it is ideal to ensure the beam is tightly focused along its trajectory. Transverse focusing of the beam can be realized with arrays of quadruple magnets at various locations along the ring. The transverse beam size is determined by two important parameters, (i) the emittance, formally defined as the area of the beam in phase space, and (ii) the beta function, related to the oscillation amplitude of the beam particles transverse to the beam momentum. The beta function is dependent on the position along the circumference of the ring based on the properties of the lattice of focusing magnets. An important quantity known as the betatron tune is defined as the ratio of the transverse oscillation frequency to the revolution frequency, or the number of transverse oscillations per revolution. The betatron tune must be kept under careful control by altering the strengths of the focusing and defocusing quadrupole magnets during run time, in particular, to avoid constructive behavior and resonances that lead to particle loss and growth of the transverse beam area [128].

In the RHIC rings, protons and gold ions are captured with an RF of 9 and 28 MHz, corresponding to a harmonic number h of 120 and 360 respectively. There are 111 bunches occupying each ring during operation, meaning for protons, every RF bucket is filled except for the last 9. The region of empty bunches is referred to as the abort gap. For gold ions every 3$^{\text{rd}}$ RF bucket is filled aside form the abort gap. The abort gap is necessary to allow the steering magnets enough time to power on when dumping the beam. The RF frequency is increased incrementally during acceleration,

eventually reaching a frequency of 197 MHz for all species, corresponding to $h = 2520$, with bunches occupying every 7^{th} RF bucket (aside from the abort gap). A storage cycle, also referred to as a fill, typically lasts around 8 hours. Each interaction region contains 18 quadrupole magnets, 9 on either side of the interaction point for focusing the beam before collision to increase the luminosity. The luminosity for top energy ($\sqrt{s} = 510$ GeV) pp collisions is $\mathcal{L} \approx 10^{32}$ cm^{-2} s^{-1}, while for top energy ($\sqrt{s_{NN}} = 200$ GeV) Au+Au collisions it is $\mathcal{L} \approx 10^{28}$ cm$^{-2}s^{-1}$. Currently, each interaction region (IR) has an important functionality, summarized in the following list [128].

- **IR12:** The location of the H-jet polarimeter [129] discussed in Section 2.1.2.
- **IR2:** The location of the Low Energy RHIC electron Cooling (LEReC) [130] and the Coherent electron Cooling (CeC) [131] projects.
  - High-intensity beams, in particular, heavy ions with a large charge suffer from intrabeam scattering (IBS). With so many protons or ions packed into a small bunch, repulsive Coulomb scattering between the bunch constituents, known as IBS, leads to emittance growth, debunching, and particle loss. Cooling the beam with a nearby beam of electrons has proven a useful method for counteracting these adverse effects.
- **IR4:** The location of the RF cavities as well as the *spin flipper* discussed in Section 2.1.1.1.
- **IR6:** The location of the STAR experimental hall [132] and the Alternating Gradient Synchrotron (AGS) to RHIC injection line.
- **IR8:** Previously the location of the PHENIX experimental hall [133]; now the location of the sPHENIX experimental hall [125].
- **IR10:** The location of the beam dump and electron lenses.

– Electron lenses are useful to counteract emittance growth, particle loss, and spread of the betatron tune from beam-beam effects in the interaction regions.

2.1.1 Polarized Protons at RHIC

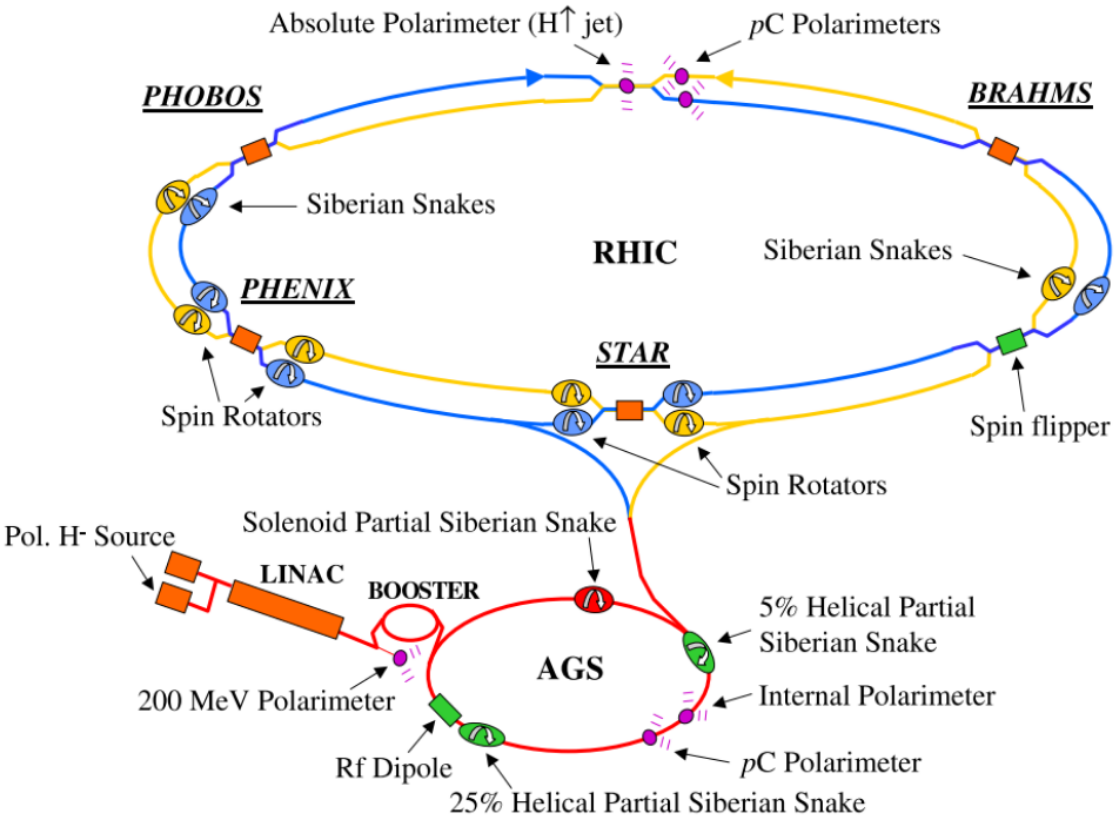

Figure 2.3: The layout of the RHIC accelerator complex showing devices responsible for maintaining, measuring, and controlling beam polarization.

Polarized protons begin their journey in the optically pumped polarized ion source (OPPIS) where 9×10^{11} H^- ions are sent into the linear accelerator (LINAC) in 500 μA 300 μs pulses [134]. The source of angular momentum in OPPIS is the TRIUMF laser system, which is used to optically pump a dense Rubidium vapor producing spin polarized Rubidium atoms at up to 2.5×10^{19} atoms/s. The design of OPPIS

and chain of charge exchange interactions that take place throughout the apparatus are described in detail in Ref. [135]. The general idea is that a low energy beam of unpolarized protons is sent through the Rb cell to undergo a charge-exchange process where the proton picks up a spin-polarized electron, resulting in a beam of electron spin-polarized Hydrogen atoms. The electron spin is then transferred to the proton in the nucleus via the Sona transition [136]. Finally, the proton spin-polarized Hydrogen atoms pass through a Sodium jet ionizer cell for the accumulation of another electron, resulting in a H^- ion that can more readily be propagated through the RHIC injection system. The pulse of H^- ions are sent through the radio frequency quadrupole for acceleration and focusing into the LINAC. The ions are subsequently accelerated to 200 MeV in the LINAC, then stripped of their electrons and injected in the Alternating Gradient Synchotron (AGS) Booster with 4×10^{11} polarized protons per bunch. In the Booster, the proton bunches are accelerated to 1.5 GeV and transferred to the AGS for additional acceleration to 25 GeV. The protons are finally transferred to RHIC resulting in 2×10^{11} protons per bunch filled in 111 bunches per ring with 106 ns bunch spacing. The spin of each bunch is independently controlled, with the bunch spin pattern being set in place by the initial pulses from OPPIS. Note that the intensities listed in this section come from the original design parameters. OPPIS was upgraded between 2009-2012 to yield 5-10 mA current pulses for higher luminosity running [135].

2.1.1.1 Maintaining Polarization

The steering and focusing of polarized proton beams require the use of external magnetic fields that alter the spin direction of polarized protons. The beam polarization for protons within the external magnetic fields present in circular accelerators

evolves as shown by the Thomas-BMT equation [137],

$$\frac{d\vec{P}}{dt} = -\left(\frac{e}{\gamma m}\right)[G\gamma\vec{B}_{\perp} + (1+G)\vec{B}_{\parallel}] \times \vec{P}, \tag{2.1}$$

where $\vec{P}$ is the polarization vector in the rest frame of the proton, $\gamma = E/m$ is the Lorentz factor, $G = 1.7928$ is the anomalous magnetic moment of the proton, and $\vec{B}_{\perp,\parallel}$ are the perpendicular and parallel components of the external magnetic field with respect to the accelerator plane. A few important features can be readily recognized from examining the Thomas-BMT equation: (i) for high beam energies, $G\gamma \gg (1+G)$ and the perpendicular component of the magnetic field has the dominant effect, (ii) $d\vec{P}$/dt is minimized when $\vec{B}$ and $\vec{P}$ are parallel, and (iii) the number of spin precessions per revolution in the case when $B_{\parallel} = 0$ is $G\gamma$. (iii) can be seen by comparing Equation 2.1 with the Lorentz Force Law,

$$\frac{d\vec{v}}{dt} = \left(\frac{e}{\gamma m}\right)[\vec{E} + \vec{v} \times \vec{B}], \tag{2.2}$$

where $\vec{E}$ is the external electric field and $\vec{v}$ is the velocity of the proton. Once the beam reaches the desired energy, there is no longer a net acceleration from the electric field. Furthermore, for $\vec{v}$ perfectly aligned in the accelerator plane, $\vec{B_{\perp}}$ is the only relevant component for the orbital motion in the accelerator. The extra factor of $G\gamma$ in front of $\vec{B}_{\perp}$ in Eq. 2.1 relative to Eq. 2.2 indicates that the spin procession is a factor of $G\gamma$ faster than the orbital motion in the accelerator when $B_{\parallel} = 0$. This important quantity can be understood as the number of full spin precessions per revolution and is referred to as the spin tune ν_{sp}. (ii) tells us that the ideal case to minimize depolarization is to have $\vec{P}$ perpendicular to the accelerator plane since $\vec{B}_{\perp}$ must be present to keep the protons in circular motion. Beams that propagate with polarization perpendicular to the accelerator plane will still experience depolarization effects from longitudinal magnetic fields due to magnetic imperfections or focusing

magnets. In particular, numerous depolarizing resonances occur in circular accelerators when the spin precession frequency is equal to the frequency that depolarizing magnetic fields are encountered. Since the steering and focusing magnets are fixed in place, the frequency that the beam encounters depolarizing fields depends on the beam energy. When the beam energy reaches the threshold of resonance, the stable polarization direction is driven away from the vertical by the depolarizing magnetic fields. One way to avoid this is by flipping the polarization direction with respect to the accelerator plane during propagation through the RHIC rings, as this would yield a relative negative sign in Eq. 2.1 such that incremental changes in $d\vec{P}/dt$ would cancel out.

Flipping the polarization direction while the beam is orbiting is accomplished with a technology known as *Siberian Snakes* invented by Derbenev and Kondratenko in 1976 [138], and expanded on by Ptitsin and Shatunov to a design consisting of helical dipole magnets that had the advantage of being more compact and modular while also minimizing shifts in the orbit trajectory [139]. A series of 4 helical dipole magnets each with a full helix period was the chosen technology for the Siberian Snakes at RHIC, with 2 snakes on each of the rings at opposing points of 3 o'clock and 9 o'clock, as shown in Figure 2.3. The performance of the snakes at $\gamma = 25$ with changes in the beam orbit, magnetic field, and spin orientation can be seen in Figure 2.4.

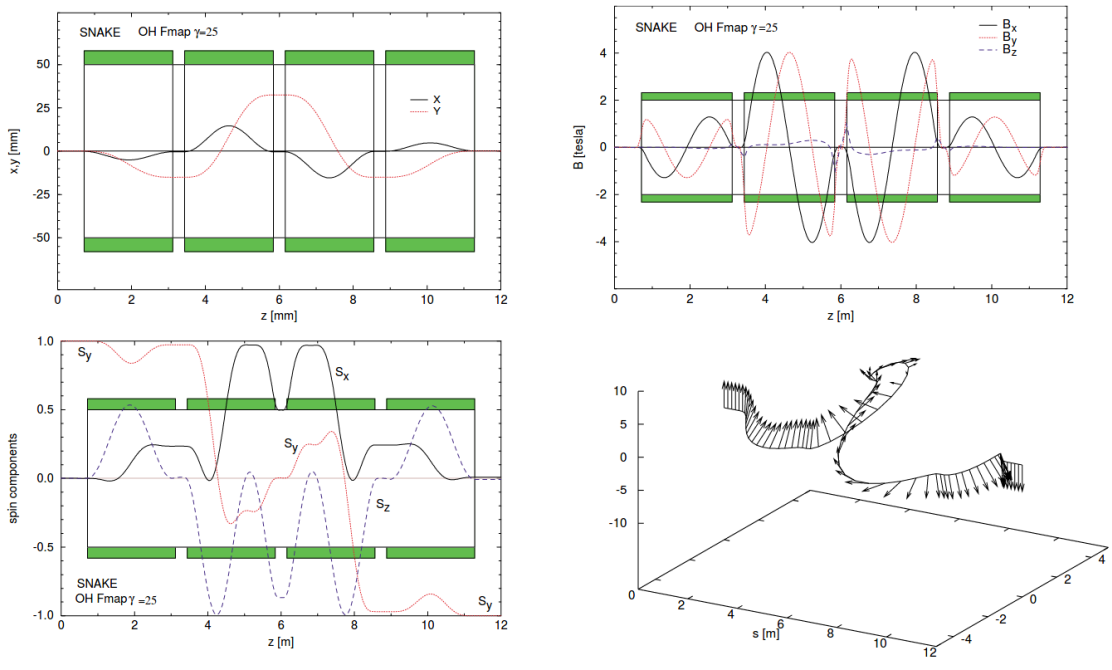

Figure 2.4: Siberian snake performance at $\gamma = 25$, with the beam orbit recorded in the top left panel, changes in the magnetic field recorded in the top right panel, changes in the beam spin orientation shown in the bottom left panel, and a trajectory plot with associated spin vectors shown in the bottom right panel. This figure was taken from Ref. [134].

The RHIC rings are also equipped with devices called *spin rotators*, similarly constructed from a series of 4 helical dipole magnets, but with different control parameters than the Siberian Snakes [134]. The spin rotators are located in each ring on both sides of the PHENIX and STAR interaction points at 8 o'clock and 6 o'clock respectively, as shown in Fig. 2.3. During some years of operation, data was taken with the beam polarization aligned parallel or antiparallel to the beam momentum. In this case the spin rotators function to change the orientation of the proton spins to be parallel or antiparallel with the momentum direction before the collision, and then back after the interaction point. This allows the accelerator and storage rings to operate under nominal conditions with the polarization direction perpendicular to

the accelerator plane to avoid depolarization effects.

To minimize systematic uncertainties arising from collisions of polarized protons, particularly from the spin pattern of the proton bunches, the installation of a *spin flipper* to flip the orientation of the proton bunches multiple times each fill has been a long-term goal of the RHIC facility. Spin flips can be achieved by adiabatically driving the beam through a spin resonance [134]. Standard implementations consisting of an AC dipole or solenoidal magnet have demonstrated an efficiency of 99% in various low-energy experiments (up to 2 GeV). They work by slowly sweeping the RF frequency of the magnet ν_{osc} across the spin tune ν_s. Single AC dipoles generate two spin resonances, one at $\nu_{osc} = \nu_s$, and another at $\nu_{osc} = 1 - \nu_s$, so the half-integer spin tune $\nu_s = 1/2$ at RHIC from the Siberian Snake configuration [134] would lead to overlapping spin resonances and interference preventing a spin flip [140, 141]. More complicated configurations therefore have to be considered. The first design consisted of 2 AC dipoles, and 4 DC dipole magnets, two on either side of the AC dipoles, and two in between with opposite polarity. While spin flips could be achieved with this configuration, it was revealed during commissioning that this induced global coherent betatron oscillations along the vertical axis and led to beam-beam interactions during collisions that would reduce the lifetime of the opposing beam [140]. It was decided to implement a new design consisting of 5 AC dipole magnets with 4 DC dipole magnets in between. A 97% efficiency was realized with this configuration at both 24 and 255 GeV, demonstrating efficient spin flipping at one of the lowest and the highest design energy of RHIC [141]. It should be noted that the spin flipper was not operational during the 2015 running year, and therefore does not affect the outcome of measurements presented in this dissertation.

The proton bunches are accelerated through two weak depolarizing resonances in the Booster that are dealt with via the so called "harmonic correction" method for closed orbits [142]. Several resonances related to the betatron tune are crossed

within the AGS, making the harmonic correction method much more tedious to apply. In addition, there is no straight section long enough within AGS to install a full Siberian snake which would enable a spin flip and protect from resonances related to the betatron tune. Instead, a series of partial Siberian snakes and an RF dipole are used to stabilize the polarization direction within AGS [134]. The result of all of this instrumentation is an average proton polarization of 59% for the blue beam and 60% for the yellow beam during the entirety of the 2015 data taking year [143].

2.1.2 Measuring Polarization

The polarization of the proton beam is regularly measured through an elastic scattering process of the proton beam on a Carbon target. The scattering is in the energy regime where Coulomb-Nuclear Interference produces a sizable asymmetry that can be used to measure the relative beam polarization, with measurements only taking seconds at peak design luminosity [134]. This way several measurements could be taken each fill for quality assurance. The pC polarimeters in each ring are complemented by a polarized hydrogen-jet polarimeter located at the 12 o'clock position of RHIC in one of the crossing regions, as seen in Figure 2.3. The hydrogen-jet polarimeter was used to calibrate the pC polarimeters to an accuracy of 3.4% by 2015, while also providing an absolute measurement of beam polarization for normalization. However, the H-jet polarimeter requires weeks of data taking (from multiple fills) due to the lower cross-section of elastic scattering on gaseous hydrogen than that of the carbon target. In both cases, the beam polarization can be extracted from the physical asymmetry,

$$A_N = \frac{1}{P_B}\frac{\sigma^{\uparrow} - \sigma^{\downarrow}}{\sigma^{\uparrow} + \sigma^{\downarrow}} \tag{2.3}$$

$$P_B = \frac{1}{A_N}\frac{\sigma^{\uparrow} - \sigma^{\downarrow}}{\sigma^{\uparrow} + \sigma^{\downarrow}}, \tag{2.4}$$

where P_B is the beam polarization, and A_N is the analyzing power. The analyzing power for this reaction is well understood, and is typically taken either from previous experimental measurements or measured in real time.

The forward scattered protons off of the carbon target in the pC polarimeters emerge at too shallow of an angle to reliably detect without hindering the quality of the beam [134]. The elastic scattering is instead identified by detecting the low-energy recoiling carbon nuclei, requiring a very thin target such that the recoiling nuclei are not absorbed. Carbon ribbon targets were developed at Indiana University Cyclotron Facility to satisfy this criterion. Accompanying the carbon target in both polarimeters were six silicon strip detectors placed to the left and right of the beam direction, and offset by $\pm 45°$ on both sides. Each silicon strip detector contains 12 10 mm $\times$ 2 mm strips. A schematic of the pC polarimeters is shown in Figure 2.5.

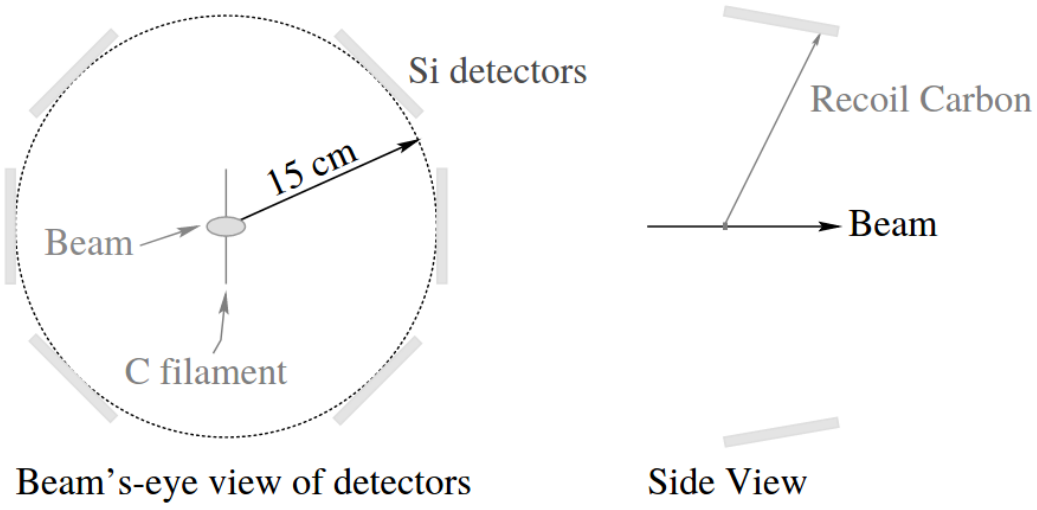

Figure 2.5: Schematic of the pC polarimeters located in each RHIC ring. Taken from Ref. [134].

An asymmetry is measured in each strip by using Equation 2.3 without taking into account the beam polarization — this quantity is referred to as the strip asymmetry

A_i, with each strip corresponding to a different position in azimuthal angle ϕ. The polarization is determined by fitting the results as a function of ϕ with a sinusoid, as shown in Equation 2.5 [144],

$$P_B \sin\phi = \frac{A_i}{\langle A_N \rangle}. \tag{2.5}$$

Here $\langle A_N \rangle$ is the cross-section weighted analyzing power (details of the calculation can be found in Ref. [145]). Equations 2.4 and 2.5 are equivalent when considering ϕ dependent cross-sections (see Eqs. 1.36 and 1.37 for details).

The H-jet polarimeter is a much larger apparatus, located at the 12 o'clock interaction region. It has 3 principal components, the polarized Atomic Beam Source (ABS), the scattering chamber, and the Breit-Rabi polarimeter (BRP) [129]. The H-jet target has a thickness of 10^{12} atoms/cm^2, which is small enough to allow for continuous operation without distorting the RHIC beams. The Breit-Rabi polarimeter measures the target proton polarization to $\approx 0.1\%$ accuracy, and the scattering chamber is equipped with silicon strip detectors to the left and right of the beam. The cross-section asymmetry $\Delta\sigma^{\uparrow,\downarrow}/\sigma$ is measured with the silicon strip detectors and used together with the target polarization to calculate the analyzing power A_N as shown in Eq. 2.3. A_N can, in turn, be used to calculate the beam polarization P_B due to the symmetry relation of the analyzing power for elastic pp scattering $A_N^{p^\uparrow p} = -A_N^{pp^\uparrow}$ [146]. The results obtained for P_B are used to normalize the fast polarization measurements from the pC polarimeters, and reduce the overall uncertainty of the beam polarization measurements at RHIC.

The transverse polarization direction at the interaction points of PHENIX and STAR is also measured and verified through local polarimetry. This is done by measuring far-forward neutron production in the Zero Degree Calorimeters (ZDCs) located at either experiment, originally installed to detect forward neutron multiplicities in heavy ion collisions in order to provide a universal event characterization amongst the different detectors [124]. Asymmetries from this production process were however

observed in the first polarized proton collisions at RHIC, and have since been used to monitor the polarization and bunch spin pattern in the interaction region [147].

2.1.3 Heavy Ions at RHIC

Heavy ions come from a different source than polarized protons and need to be fully stripped of their electrons before entering the RHIC rings. Figure 2.6 shows a schematic of the heavy ion injection system, with the Electron Ion Beam Source (EBIS) replacing the Tandem Van de Graaff as the source of ions in 2012. EBIS provides ions to both the RHIC rings as well as the NASA Space Radiation Laboratory. It consists of a 10 A electron gun, followed by a trapping region for singly charged ions, and an electron collector. Singly charged ions are injected into the trapping region from up to two separate sources with the ability to switch between species on a 1 s time scale [148, 149]. The theoretical principle behind an EBIS is known as electron impact ionization, where a focused electron beam (compressed by magnetic fields) bombards a gas of ions dispersed throughout the trapping region [150]. The trapping region contains cylindrical electrodes that create an axial electrostatic trap with adjustable trapping time, allowing for specific charge states to be achieved. This is tuned to be peaked at Au^{32+} for gold ions. The RHIC EBIS is designed to accept ions satisfying the criteria $q/m > 1/6$ in 10-40 μs pulses during extraction time.

In the case of gold ions, pulses of about 1.2×10^9 Au^{32+} ions with energy 17 keV per nucleon are accelerated by a radio frequency quadrupole and LINAC to respective energies of 300 keV and 2 MeV per nucleon [151]. The ions then arrive at the 30 m EBIS to Booster (ETB) transport line, where about 1.0×10^9 Au^{32+} ions are delivered to the Booster per EBIS pulse. This is about an 85% efficiency, which is notably better than the reported 56% efficiency of ions delivered to the Booster per pulse of the Tandem Van de Graaff. The EBIS pulse is delivered over the course of 1-4 Booster revolutions, and captured at a harmonic number of 4 to

create 4 bunches. The bunches are subsequently merged, accelerated to 105 MeV per nucleon, and output to the Booster to AGS (BTA) transport line, where they are stripped to Au^{77+} by passing through layers of aluminum and "glassy" carbon foil. The stripping process in the BTA transfer line has an efficiency of about 65%. During an AGS injection cycle, 8 bunches are delivered from the Booster and captured with a harmonic number of 16 in 2 groups of 4 adjacent RF buckets. Two subsequent merges take place before accelerating the final two bunches to 8.865 GeV per nucleon in the AGS. This procedure results in about 2.06×10^9 Au^{77+} ions at extraction from AGS, which are then sent to the AGS to RHIC (ATR) transport line where the remaining two electrons are stripped to produce Au^{79+} ions. This procedure is repeated until 111 bunches are filled in both the yellow and blue RHIC rings (for A+A collisions). In 2015, the first polarized proton on gold and aluminum ion collisions were recorded at RHIC, meaning that the injection system discussed here and in Section 2.1.1 can be operated in parallel. Once the ion bunches are injected into RHIC, they are accelerated to the desired energy for collision (the top design energy is 100 GeV per nucleon).

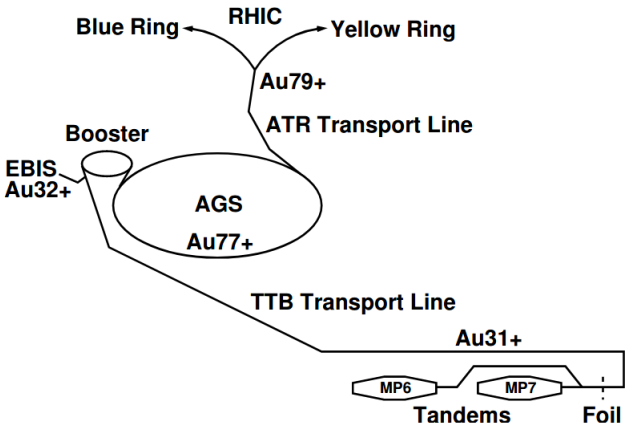

Figure 2.6: Schematic of the heavy ion injection system at RHIC, showing the intermittent stripping of electrons from a gold nucleus along its trajectory. This figure taken from Ref. [151].

2.2 The PHENIX Experiment

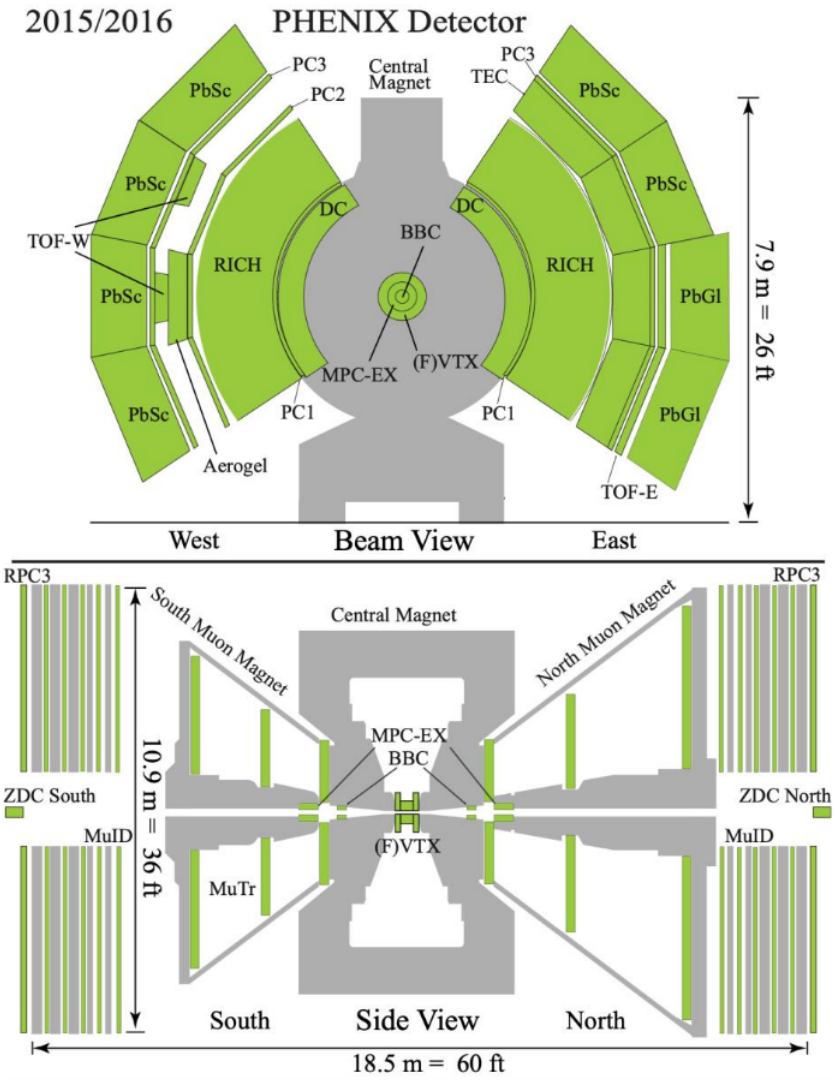

Figure 2.7: Schematic of the PHENIX detector as of its final years of running in 2015/2016. A cross sectional view of the central arm spectrometers is shown on top, with the various detector subsystems labeled, while a side view of the detector is shown in the bottom panel including the forward instrumentation.

The PHENIX experiment ran from the beginning of RHIC operation in 2000 until its decommissioning in 2016. Despite the experiment no longer running, there are still a number of measurements being performed from the wealth of data captured during

its many years of operation. The focus of PHENIX was to measure a wide variety of probes in various collision systems, made possible by high granularity sub-detectors and high event rate capabilities. The layout of the PHENIX detector is shown in Figure 2.7, showing the central arm spectrometers (looking down the beam line) in the top panel, and the muon arms as well as other forward detector subsystems in the bottom panel. A detailed overview of the PHENIX detector is provided in Ref. [133].

2.2.1 Event Characterization Detectors

2.2.1.1 Beam-Beam Counter

The beam-beam counters (BBCs) provide crucial global event information such as the location of the collision vertex along the z-axis, the timing information for the event, and trigger information for the luminosity determination. They are located along the beamline at ±144 cm from the nominal collision point, just outside of the central magnet poles as shown in Fig. 2.7. The BBC has an acceptance of $3.1 < |\eta| < 3.9$ in pseudorapidity and full coverage in azimuth. They comprise arrays of 64 3 cm quartz radiators with photomultiplier tubes for readout. A detailed overview of the BBC detectors as well as other inner PHENIX detectors can be found in Ref. [152].

Charged particle detection in the BBC provides a criterion for identifying inelastic pp scattering events. The detection of at least one charged particle in both the north and south BBC detectors is a requirement to fire the least stringent trigger in PHENIX (the minimum bias or BBC local level 1 trigger). For luminosity determination, the minimum bias or BBCLL1 trigger had an additional requirement that the calculated collision vertex position along the z-axis should be within ±30 cm of the nominal interaction point [152]. If both criteria are satisfied, this prompts the data acquisition (DAQ) system to write out this event for further processing. The efficiency of such a trigger is about 50% for $\sqrt{s} = 200$ GeV pp collisions, and even higher for AuAu collisions. By recording the number of times this trigger is fired for each bunch

crossing, the luminosity can be determined. In particular, if this is done as a function of the bunch spin pattern, one can determine the ratio of counts for spin-up bunch crossings vs spin-down bunch crossings in order to calculate the relative luminosity.

2.2.1.2 Zero Degree Calorimeters

The zero-degree calorimeters (ZDCs) are very far-forward detectors designed to record the number of spectator neutrons produced in heavy ion collisions at each experimental hall of RHIC. As mentioned in Section 2.1.2, they have also proved to be useful detectors in $p^{\uparrow}p$ collisions for local polarimetry, as a sizable asymmetry of the produced neutrons was discovered at RHIC and can be used to determine the spin direction of either beam at the interaction region. Details about the ZDC detector design can be found in Ref. [153].

2.2.2 Central Arm Spectrometers

The central arm spectrometers have coverage of $|\eta| < 0.35$ in pseudorapidity and $\Delta\phi = \pi/2$ per arm in azimuthal angle, positioned slightly offset from back to back in order to maximize dilepton and diphoton pair acceptance [154], shown in the top panel of Figure 2.7. The central arm spectrometers contain the central magnet, consisting of an inner and outer set of electromagnetic coils that provide an axial magnetic field of up to 1.15 T $\cdot$ m [155].

2.2.2.1 Charged Particle Tracking

The primary function of the tracking detectors is to measure the trajectories and momenta of charged particles. Charged particle tracking in the central arm spectrometers consists of multiple detector subsystems, including Drift Chambers (DC), and Pad Chambers (PC1, PC2, PC3) [156]. The Drift Chambers perform precise track measurements in the $r - \phi$ plane, allowing for reconstruction of charged

particle p_T, while the Pad Chambers provide 3D tracking information in $r-\phi$ and z to reconstruct p_z and provide additional hits to correlate with signals in the DCs.

Drift Chambers The Drift Chambers are multiwire proportional chambers, consisting of a volume filled with gas and wires aligned in different orientations. Charged particles passing through the DCs ionize the gas, and the sense (anode) wires throughout the chamber are held at a high enough voltage to catch the resulting electrons. These signals are read out to determine which wire collected the particles and therefore the spatial location in a plane excluding the direction of the sense wire. The Drift Chambers are placed in the residual magnetic field of the central magnet, the maximum field strength within the volume of the DC is 0.06 T. This allows for the determination of momenta and charge of charged particles based on their curvature. Each DC is split into 20 equal sectors with coverage $\Delta\phi = 4.5°$, and contains six types of wire modules X1, X2, U1, U2, V1, and V2, as shown in Figure 2.8. The wire modules labeled 1 and 2 are stacked radially, with each X module containing 12 sense (anode) wires, and each U, V module containing 4. This results in 40 drift cells at different radii and azimuthal angles within the sector, providing single wire resolution and efficiency better than 150 μm in $r-\phi$ and 99% respectively, and single wire two track separation better than 1.5 mm [156]. The purpose of the U and V wire modules, which have stereo angles of about $\pm 6°$ with respect to the X wire modules, was to measure the z coordinate of the track with spatial resolution better than 2 mm. The z resolution is similar to that of the Pad Chamber layers, allowing hits to be correlated among the different tracking detectors and fit to charged particle trajectories that traverse many of the central arm sub detectors.

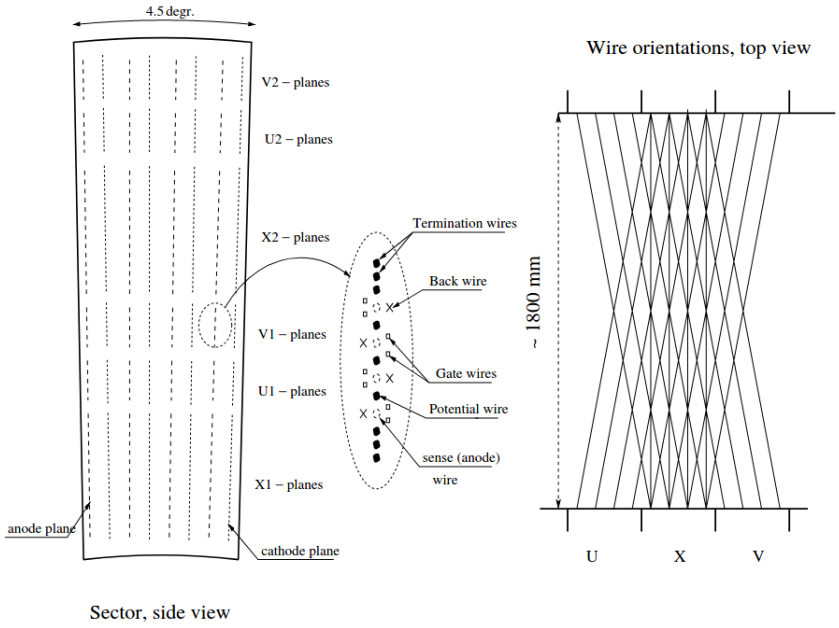

Figure 2.8: A schematic of a Drift Chamber sector showing a side view on the left along with a zoomed in view of the anode plane of one of the 40 drift cells, and an aerial view on the right showing the X, U, and V wire orientations. This figure was taken from Ref. [156].

Pad Chambers The Pad Chambers consist of 3 layers of multiwire proportional chambers (PC1, PC2, PC3) with a plane of field and anode wires suspended in a gas volume between a pixel-segmented cathode plane and a solid copper cathode plane [156]. The PCs are located outside of the central magnet and are purposed to measure points in space that correspond to straight-line particle trajectories. The segmentation of the pixelated cathodes is fine enough to ensure a resolution of a few millimeters in both the $r - \phi$ plane as well as the z direction. The resolution in the z direction is comparable to that of the stereo wire modules in the DCs such that hits in each detector subsystem can be correlated and fit to charge particle trajectories. The PC1 and PC3 layers are located in both spectrometer arms just outside of the DCs and just before the electromagnetic calorimeter, respectively. The

PC2 layer is only installed in the west arm and is positioned just outside the RICH as shown in Fig. 2.7. PC1 is particularly important for providing a p_z component of momentum as particles exit the DC, yielding a 3-dimensional momentum vector that determines the trajectory of particles through the RICH. In addition, the PC3 layer provides spatial information on charged particles just before they enter the EMCal. This makes identifying correlated hits in the DCs and PCs particularly important for electron identification [156].

2.2.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (EMCal) is located at the edge of either spectrometer arm (see Fig. 2.7). It consists of two types of calorimeters, six sectors of sampling lead-scintillator (PbSc) calorimeter and two sectors of Cherenkov lead-glass (PbGl) calorimeter with different granularity ($\Delta\phi \times \Delta\eta = 0.011 \times 0.011$ in PbSc and 0.008×0.008 in PbGl) [157]. Part of what made PHENIX such a unique detector was the high granularity of the EMCal. The 8 calorimeter sectors contained a total of 24,768 individual towers each read out with a photomultiplier tube.

The two calorimeter systems differ in important ways such as energy resolution, timing resolution, granularity, linearity, response to charge hadrons, and shower shape, providing important systematic cross-checks for measurements [157]. The PbSc calorimeter has a better timing resolution for electromagnetic showers of 200 ps, as opposed to the 300 ps timing resolution of the PbGl calorimeters. On the other hand, the PbGl has a finer granularity and nominal energy resolution of $5.9\%\sqrt{E}$ (GeV) $\oplus$ 0.8%, better than that of the PbSc calorimeter ($8.1\%\sqrt{E}$ (GeV) $\oplus$ 2.1%).

Figure 2.9 shows a PbSc calorimeter module containing 4 optically isolated towers, each with alternating layers of lead and scintillating material grouped into 66 sampling cells that are optically connected with 36 longitudinally penetrating wavelength shifting fibers [157]. The optical fibers deliver light to 30 mm phototubes at the back

of the tower.

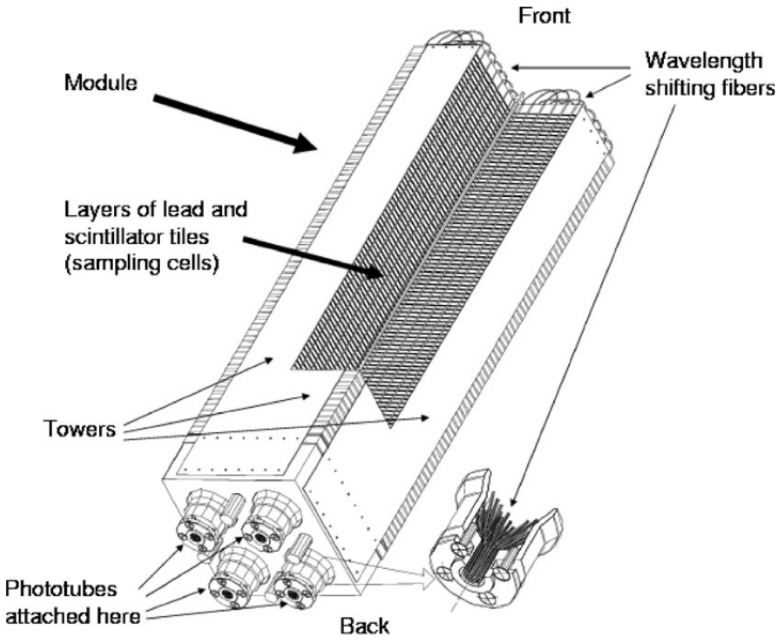

Figure 2.9: A lead-scintillator electromagnetic calorimeter module, consisting of 4 towers that are optically isolated and have independent phototube readouts. This figure was taken from Ref. [157].

For both the lead-scintillator and lead-glass type calorimeters, towers were grouped into rigid structures called supermodules, made up of 144 PbSc towers and 24 PbGl towers respectively. The sectors are then made up of 18 supermodules for the PbSc calorimeter and 192 supermodules for the Pb-glass, held in place by steel frames [157]. Figure 2.10 shows a supermodule of the lead-glass calorimeter, including a 6×4 matrix of PbGl towers, each with a photo multiplier attached to the far side of the tower.

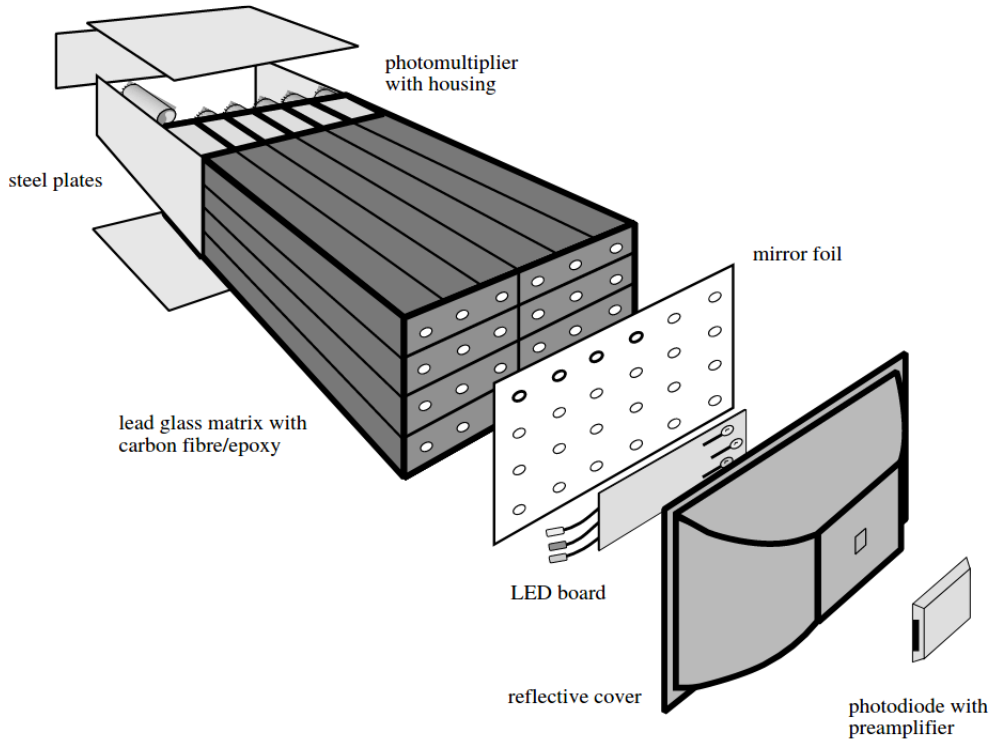

Figure 2.10: A supermodule of the lead-glass electromagnetic calorimeter, comprising 24 towers with independent phototubes for readout. This figure taken from Ref. [157].

2.2.2.3 Particle Identification

The PHENIX central arms have several particle identification (PID) detectors, in particular for the discrimination of electrons and charged pions. The main PID detector subsystem used in this dissertation work was the Ring Imaging Cherenkov Detector (RICH), consisting of a volume filled with CO_2 gas in each spectrometer arm with a Cherenkov threshold of $\gamma = 35$, corresponding to $p = 20$ MeV for electrons and $p = 4.9$ GeV for charged pions [158, 159]. RICH detectors operate by sensing photons in a ring shape from the characteristic conical Cherenkov radiation pattern emitted when a charged particle travels faster than the speed of light in a medium. The RICH detectors at PHENIX comprise two intersecting spherical mirrors made of 48 composite mirror panels that focus Cherenkov light onto two arrays of 1280

photomultiplier tubes. Figure 2.11 depicts a schematic of one of the RICH detectors and its various components.

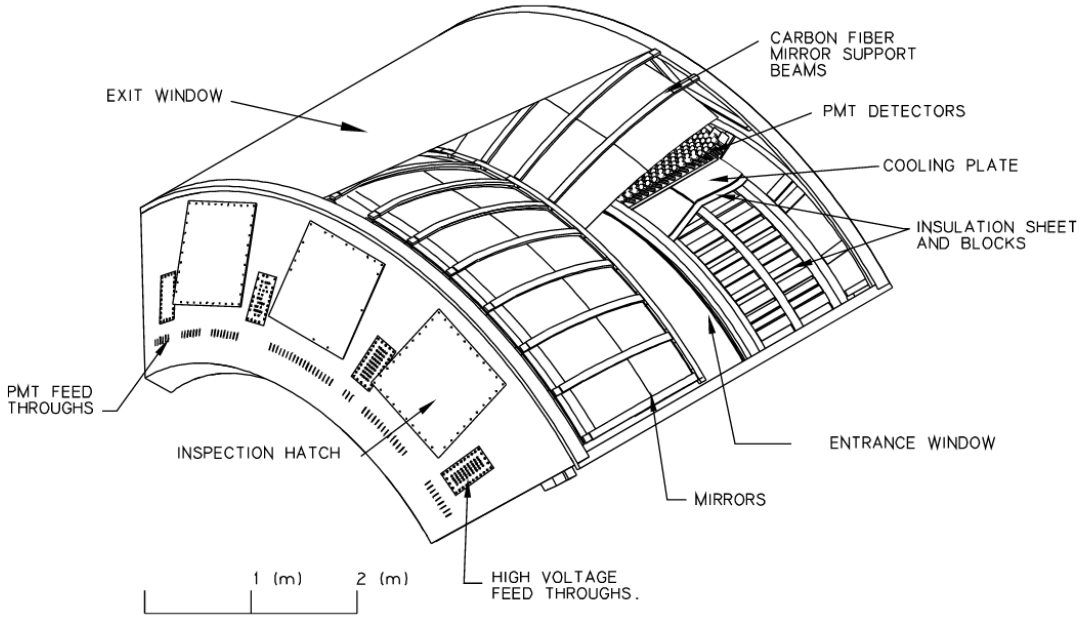

Figure 2.11: A schematic of the Ring Imaging Cherenkov Detector in one of the spectrometer arms, with two intersecting spherical mirrors to focus Cherenkov radiation onto arrays of photomultiplier tubes. This figure taken from Ref. [158].

The RICH was originally designed to contain ethane gas due to its high photon multiplicity in the Cherenkov shower but was ultimately operated with CO_2 which offered a trade-off of a higher Cherenkov threshold (and therefore a larger window of e/π separation) for lower e/π separation efficiency [158]. As mentioned in Section 2.2.2, correlated hits in the Drift Chamber and Pad Chamber 1 in a particular spectrometer arm provide a momentum direction to identify the trajectory of charged particles propagating through the RICH. Requiring some photomultiplier tubes fire in the RICH within a specific annulus corresponding to a threshold distance from the track projection in the RICH significantly improves the purity of detected electrons. In the case of CO_2 gas, the ring diameter of the Cherenkov light on the PMT array is about 11.8 cm, so a distance of at least 5 cm between the track projection and

photomultiplier tubes that fired is typically required for quality electron candidates.

2.2.2.4 The EMCal-RICH Trigger

The EMCal design permits the collection of rare events containing a high p_T photon or electron by triggering on events where the sum of energy deposited in various groupings of towers exceeds a predefined threshold. Towers are grouped with other adjacent towers into tiles of 2×2 and 4×4. The electron trigger consists of PMT signals in the RICH in addition to the energy sum deposited in a 2×2 tile of the EMCal being over a preset threshold, hence the name EMCal-RICH trigger (ERT). On the other hand, the high p_T photon triggers consist of 3 types labeled ERTA, ERTB, and ERTC each corresponding to 4×4 tiles with summed energy deposits above 3 respective energy thresholds. More information about the ERT trigger system can be found in a Ref. [160]. Figure 2.12 depicts a cartoon of the ERT inputs for various particles.

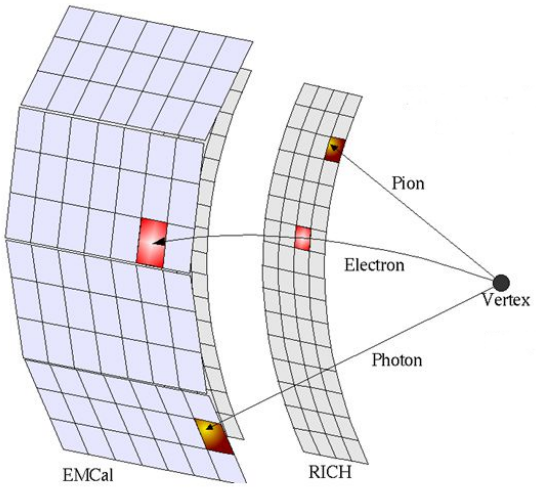

Figure 2.12: A schematic of the detector subsystems that make up the EMCal-RICH trigger, showing signal traces of various particles with electrons being registered in both the RICH and the EMCal, as required for the trigger condition.

2.2.3 Silicon Vertex Detector

The Silicon Vertex Detector (VTX) was a tracking detector installed close to the interaction region in 2011 that increased the resolution of the primary vertex reconstruction to 150 μm in pp collisions [161]. It also helped in identifying charged particles from the decay of heavy flavor hadrons that travel hundreds of microns before decaying. Figure 2.13 shows a schematic of the VTX, with finely segmented layers located at $r = 2.6, 5.1, 11.8$, and 16.7 cm, labeled B0, B1, B2, and B3, respectively. This came at the cost of a material budget per layer (expressed as a percentage of a radiation length) of $X_0 = 1.28\%, 1.28\%, 5.43\%$, and 5.43%, making a significant source of electrons from photonic conversions $\gamma \to e^+e^-$. However, the detector itself can be exploited to mitigate the electrons from photonic conversions by requiring hits in various layers of the VTX.

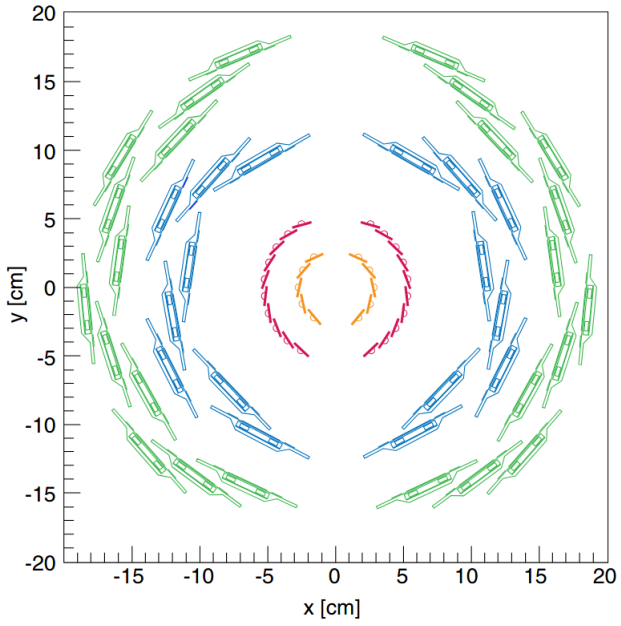

Figure 2.13: Layout of the VTX detector and its various layers B0, B1, B2, and B3 in (x, y) space. This figure taken from Ref. [161].

The Silicon Vertex Detector has acceptance $|\eta| < 1.2$ and $\Delta\phi \approx 0.8\pi$ per arm.

The layers B0 and B1 are made up of five and ten ladders of silicon pixel detectors respectively, with each ladder containing an array of 256×32 pixels of granularity $50\ \mu\text{m} \times 425\ \mu\text{m}$ per pixel [161]. Each pixel has its own readout channel, leading to a resolution of $109 \pm 37\ \mu\text{m}$ in the z direction and $13 \pm 5\ \mu\text{m}$ in the ϕ direction. The silicon pixel technology used to construct layers B0 and B1 was developed at CERN [162].

Layers B2 and B3 have 8 and 12 ladders per arm with 5 and 6 sensors, respectively, each segmented into stripixels with an effective size of $80\ \mu\text{m} \times 1000\ \mu\text{m}$ due to the stereoscopic readout of the sensors. More specifically, each pixel consists of two interleaved spiral implants that both register charge deposits from ionizing particles and carry the signal through two independent readout channels. The two implants are then connected to other pixels in the array at the edges of the spiral forming horizontal stripixels labeled X, and those with a 4.6° stereo angle labeled U (see Figure 2.14). This provides 2-dimensional (z, ϕ) position resolution with micrometer sensitivity, as the separate position or momentum components can be calculated from the stereo angle. This was a novel stripixel technology developed at BNL, with the advantage of having fewer readout channels than typical pixel detectors [163].

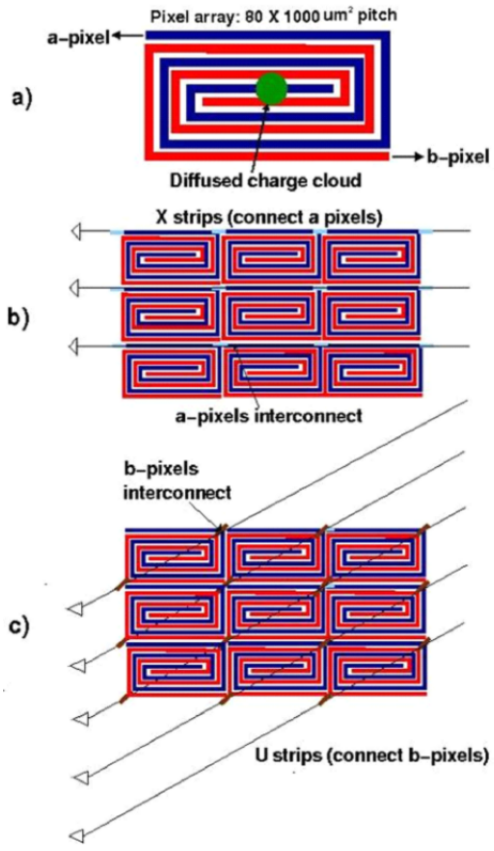

Figure 2.14: A diagram of the silicon strip pixel technology used in layers B2 and B3 of the VTX. This figure taken from Ref. [164].

CHAPTER III

Analysis Methods

Transverse single-spin asymmetry measurements at PHENIX have the potential to shed light on open questions in QCD, such as the gluonic structure of polarized protons, and effects from collisions with a heavier nucleus. The analysis procedures for the midrapidity open heavy flavor electron TSSA measurement on the 2015 $\sqrt{s} = 200$ GeV $p^{\uparrow}p$ dataset (see Ref. [1]), and the midrapidity π^0 and η meson TSSA measurement on the 2015 $\sqrt{s_{NN}} = 200$ GeV $p^{\uparrow}$Au and $p^{\uparrow}$Al datasets (see Ref. [2]), are presented in detail. The former places direct constraints on the twist-3 trigluon correlation functions in polarized protons, thereby indirectly constraining the gluon Sivers TMD PDF. The latter investigates modifications to TSSAs in collisions involving heavier nuclei. For both measurements, the data collected by the PHENIX experiment at RHIC are uniquely suited to answer the respective research questions.

3.1 Data Quality Assurance

The quality of the data used for measurement must be studied rigorously to ensure detection and polarization control capabilities of PHENIX and RHIC are working reliably. This is typically checked on a run-by-run basis for quality assurance of the PHENIX detector, and a fill-by-fill basis for that of the RHIC polarimetry. A run in this context refers to a collection of events and is specific to the operation of the

experiment (in this case PHENIX). Data taken during run time is indexed by run number, allowing collaborators to determine which detector subsystems were working reliably for which runs. A run is a subset of a fill, which refers to the period of running time after the beams are injected into the RHIC rings until they are dumped for the next fill (see Section 2.1 for details).

3.1.1 2015 Proton-Proton Collisions

Detailed data quality studies of the vertexing and tracking detector subsystems were conducted in the work corresponding to [161] and adopted for the open heavy flavor electron TSSA analysis. In particular, quality assurance variables were studied on a run-by-run basis at the event level for measured clusters in the VTX, and for tracks measured by the central arm spectrometers both with and without extrapolated fits to hits in the VTX detector. Runs containing quantities with a deviation from the mean value that is statistically significant are further inspected to determine if they should be removed from the good run list. Quality assurance variables for these studies are listed below.

- Event Level Variables
  - Number of narrow vertex minimum bias events
  - The fraction of events with a precise vertex measured by the VTX detector
- VTX Cluster Variables
  - Number of clusters measured in the VTX per event
- Track Level Variables
  - Tracks measured by the central arm spectrometer
    - Hadron tracks per event

    * Electron tracks per event
    * Ratio of electron to hadron tracks per event
  - Tracks measured by the central arm spectrometer with an extrapolated fit to corresponding hits in the VTX detector
    * Hadron tracks per event
    * Electron tracks per event
    * Ratio of electron to hadron tracks per event
    * Mean of the measured distance of closest approach (DCA) distribution for hadrons
    * Width of the measured distance of closest approach (DCA) distribution for hadrons

In addition, the work conducted in [123, 165] included studies of the data quality that went into the polarization and relative luminosity measurements. The outcome of all of these studies was a list of good runs suitable for the heavy flavor electron transverse single-spin asymmetry measurement.

3.1.2 2015 Proton-Nucleus Collisions

An internal study of π^0 mesons in pp, pAl, and pAu collisions was conducted on the 2015 PHENIX dataset in which the data quality related to the collision vertex (as measured by the beam-beam-counter), and π^0 rate (as measured by the EMCal) were investigated in detail. This work was not published, but it set the stage for Ref. [2] (discussed in this dissertation), as well as Ref. [123] (part of the dissertation of Nicole A. Lewis from the University of Michigan). The π^0 rate was studied in further detail as a function of p_T in Ref. [166] (part of Joseph D. Osborn's dissertation from the University of Michigan), leading to the removal of an additional run in the proton-Gold dataset. The quality assurance of measured quantities related to the beam

polarization also needed to be checked in detail to ensure polarization measurements were reliable, and that the spin patterns stored in the Spin Database reliably reflected the spin directions of the polarized proton bunches. For the 2015 datasets, much of these studies were conducted by a small team at PHENIX as a service to the collaboration, and several runs were flagged for removal. After removing the flagged runs, the following were studied to ensure the data was ready for measurement: (*i*) the counts (GL1P scaler sums) that went into the relative luminosity calculation and (*ii*) the trigger efficiency for η mesons (unlike the π^0 mesons, this was not previously studied by PHENIX in pA data). Figures 3.1 - 3.4 show GL1P scaler sums as a function of run number, where runs can be seen in both pAu and pAl with GL1P scaler sums that were significantly larger than others (by about 3 orders of magnitude). These runs were removed from the dataset. Finally, the outcome of study (*ii*) is shown in Figures 3.5 and 3.6 for pAu collisions, and Figures 3.7 and 3.8 for pAl collisions, with Figures 3.6 and 3.8 showing zoomed in versions for clarity. There were no obvious runs to remove based on this study.

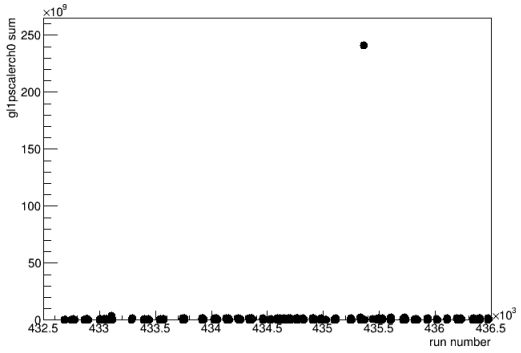

Figure 3.1: Trigger counts for each run in the pAu data, relevant for the relative luminosity calculation. It can be seen that run 435631 is several orders of magnitude larger than the rest of the runs.

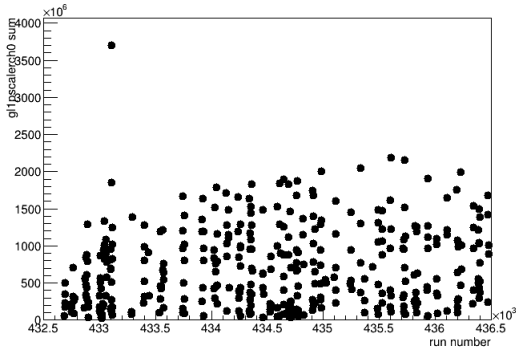

Figure 3.2: Trigger counts for each run in the pAu data after the removal of run 435631.

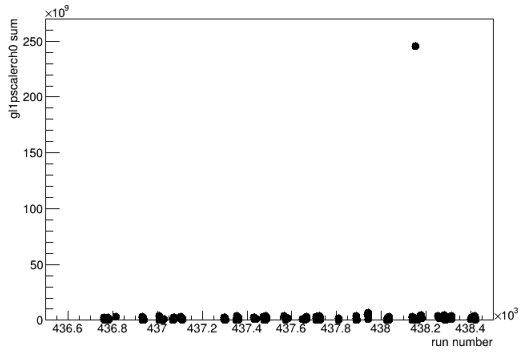

Figure 3.3: Trigger counts for each run in the pAl data, relevant for the relative luminosity calculation. It can be seen that run 438153 is several orders of magnitude larger than the rest of the runs.

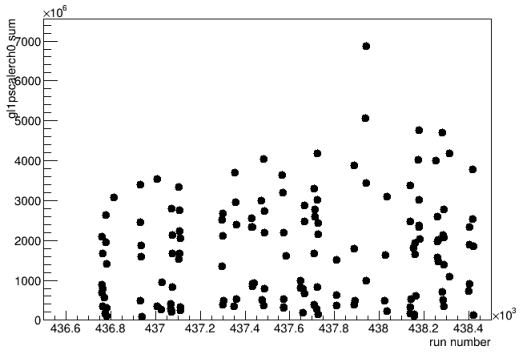

Figure 3.4: Trigger counts for each run in the pAl data after the removal of run 438153.

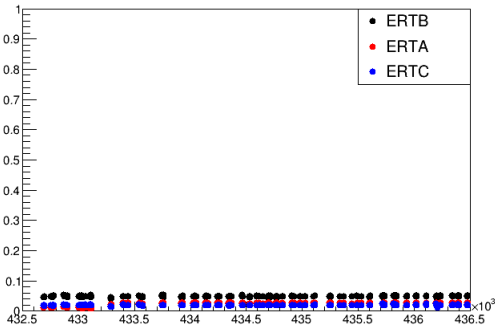

Figure 3.5: η trigger efficiency for the pAu data.

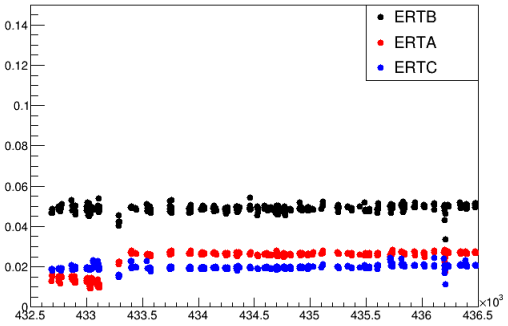

Figure 3.6: Zoomed in version of Figure 3.5.

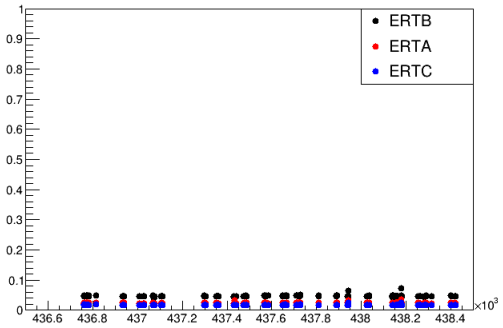

Figure 3.7: η trigger efficiency for the pAl data.

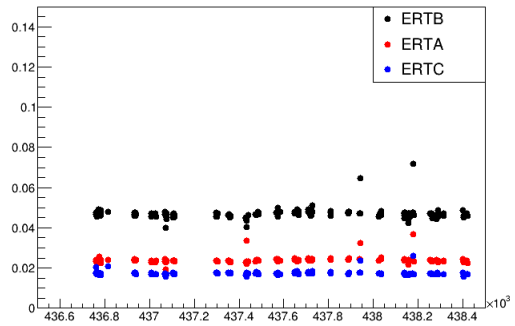

Figure 3.8: Zoomed in version of Figure 3.7.

3.1.2.1 Hot Towers

After curating a list of good runs, the performance of the electromagnetic towers and their corresponding electronics was investigated to determine if any were producing spurious signals (known as hot towers). The procedure for creating the hot tower lists follows exactly from the π^0 and η TSSA measurement with the 2015 pp dataset [123]. Maps of the calorimeter sectors are created by recording the number of electromagnetic clusters registered in each tower over the full running period as a function of tower position. Tower maps are created separately for low electromagnetic cluster energies in the range $0.5 < E_{cl} < 5$ GeV and high cluster energies in the range $5 < E_{cl} < 20$. A previous PHENIX QA study to determine a list of hot towers and dead (electrically inactive) towers was used to seed this study and determine the input tower list. EM clusters centered on or next to a previously flagged tower, or the edge of the calorimeter sector were not considered. Figures 3.9 and 3.11 show the tower maps for the low and high energy EM clusters respectively for a PbSc sector in the west spectrometer arm in the pAu dataset. The updated list of towers used for analysis is determined with an iterative procedure, flagging towers as hot if the number of clusters it recorded was larger than 6 times the RMS value of the total

number of clusters measured in the sector. Towers flagged as hot are removed from the tower map and the procedure is repeated with the newly calculated RMS value until no new hot towers are found. The final tower maps (corresponding to Figures 3.9 and 3.11 after removal of hot towers) are shown in Figures 3.10 and 3.12 respectively. While the figures only show this procedure for a particular sector in the pAu dataset, it is repeated for each sector in both the pAu and pAl datasets.

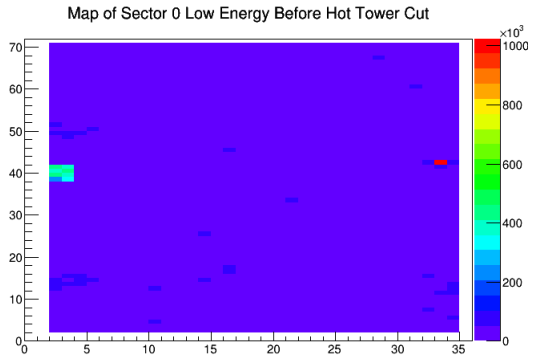

Figure 3.9: Sector 0, pAu, $0.5 < E_{cl} < 5$ GeV – before hot tower cuts

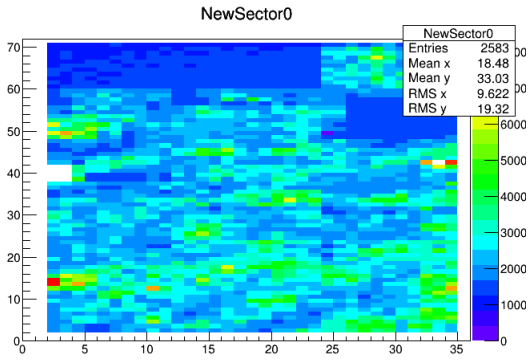

Figure 3.10: Sector 0, pAu, $0.5 < E_{cl} < 5$ GeV – after hot tower cuts

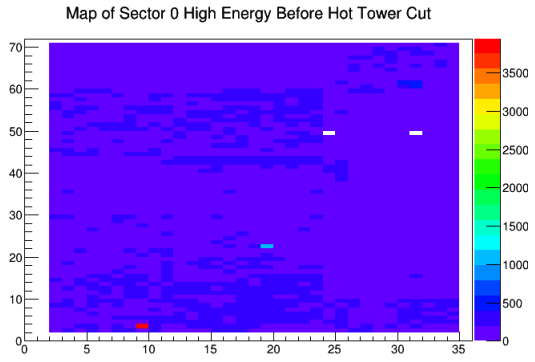

Figure 3.11: Sector 0, pAu, $5 < E_{cl} < 20$ GeV – before hot tower cuts

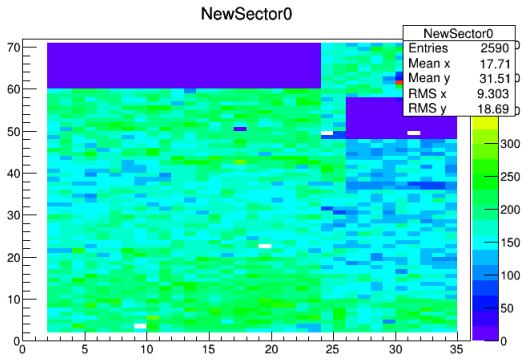

Figure 3.12: Sector 0, pAu, $5 < E_{cl} < 20$ GeV – after hot tower cuts

3.2 Signal Extraction

Quantities measured in different PHENIX detector subsystems can be used to create selection criteria for obtaining samples enriched in various produced particles. Candidates remaining after selection criteria are applied contain a mixture of both signal and background, and various methods are implemented to quantify the fraction of background in the candidate sample depending on the signal of interest.

3.2.1 Heavy Flavor Electrons

Heavy flavor electrons refer to electrons or positrons coming from the decay of a hadron containing a charm ($c/\bar{c}$) or beauty ($b/\bar{b}$) quark. This measurement was performed on the $\sqrt{s} = 200$ GeV polarized pp data set from the 2015 running year. Heavy flavor electrons were measured with the PHENIX central arm spectrometers ($|\eta| < 0.35$), utilizing the tracking (DC, PC1, PC2, PC3), vertexing (VTX), calorimetry (EMCal), and particle identification (RICH) detector subsystems. The analysis methods presented in this section follow very closely to that of [161], with the following modifications: (i) runs were removed that did not pass quality assurance tests related to polarimetry, (ii) electron candidates were separated by charge (into electron and positron samples), and (iii) a different p_T binning scheme was used, in order to allow for enough statistics to sort the candidate sample by the proton bunch polarization.

3.2.1.1 Selection Criteria

The p_T bins selected for this analysis are shown in Table 3.1. This choice was motivated by the measured p_T range in Ref. [161], where the $5\ GeV < p_T < 6\ GeV$ bin was dropped in this analysis due to its low signal to background for nonphotonic electrons, and coarser bins were used with the same bin edges in order to ensure there was enough statistics for spin sorting. Charged pions begin to fire the RICH with $p_T > 5\ GeV$, therefore different analysis cuts would have to be implemented in this

range, making it impractical to combine candidates with $p_T > 5\ GeV$ and candidates with $p_T < 5\ GeV$ in the same bin. It was found that above $p_T > 5\ GeV$, no set of analysis cuts yielded high enough statistics to spin sort without being completely dominated by background from charged pions or photonic electrons.

Table 3.1: p_T bins chosen for this analysis shown with average p_T values obtained in each bin for charge combined (+/-), and separated (+,-) samples after applying all analysis cuts. Due to the average p_T values being similar for the different samples, all results are plotted with the average p_T values from the charge combined sample (+/-).

p_T range (GeV)	$\langle p_T \rangle$ (+/−)	$\langle p_T \rangle$ (+)	$\langle p_T \rangle$ (−)
1.0 – 1.3	1.161	1.163	1.160
1.3 – 1.5	1.398	1.398	1.398
1.5 – 1.8	1.639	1.641	1.638
1.8 – 2.1	1.936	1.937	1.935
2.1 – 2.7	2.349	2.351	2.347
2.7 – 5.0	3.290	3.299	3.282

Electron Identification Selection Requirements: The following criteria are applied to quantities measured by the central arm spectrometers to identify electron candidates.

- $prob > 0.01$ — This variable measures the probability of a particle shower in the EMCal being electromagnetic. Low-probability candidates were removed from the analysis.

- $n0 > 1$ — This variable measures the number of phototubes fired within a nominally sized ring (annulus) in the RICH, used to identify electrons based on their characteristic Cherenkov ring.

- |dep| < 2 — The variable measures the quantity $(E/p - \langle E/p \rangle)/\sigma_{E/p}$, where $\langle E/p \rangle$ and $\sigma_{E/p}$ were calibrated with a measured electron candidate sample from the 2015 pp data (that did not rely on E/p cuts). It should be noted that E is measured in the electromagnetic calorimeter, p is measured with the tracking subsystems, and $\langle E/p \rangle \approx 1$ for electrons given that they deposit most of their energy in the electromagnetic calorimeter.

- $|emcsd\phi| < 3$ — This variable measures the difference between the track projection to the calorimeter and the measured energy deposit in the EMCal normalized to standard deviations along the ϕ direction. It is important to ensure the EMCal energy E and track momentum p are correctly matched to the same particle.

- $|emcsdz| < 3$ — This variable measures the difference between the track projection to the calorimeter and the measured energy deposit in the EMCal normalized to standard deviations along the z direction. It is important to ensure the EMCal energy E and track momentum p are correctly matched to the same particle.

- $disp < 5$ cm — This variable measures the displacement of the electromagnetic shower ring center with respect to the track projection position in the RICH photomultiplier tube array. PHENIX recommends this requirement to identify electrons.

- Conversion veto cut applied — This variable is outlined in Section 3.2.1.2. It is useful in removing the background from photonic conversions $\gamma \to e^+e^-$.

Track Selection Requirements: The following set of criteria are applied to quantities measured by the VTX detector and central arm spectrometers to ensure high-quality tracks.

- $\chi^2/ndf < 3$ — This variable measures the quality of the fit of tracks measured in the central arm spectrometers extrapolated to hits in the VTX, this cut was chosen such that only reliably fitted tracks are analyzed.

- $quality == 31 \;||\; quality == 63$ — This variable measures the quality of drift chamber tracks, PHENIX recommends the above cut for high-quality charged tracks in the DC.

- $(hitpattern\&3) == 3$ — This variable measures the pattern of hits in the VTX corresponding to a track, and is applied with a bitwise operator which requires at least one hit in each of the inner two layers of the VTX. This helps to remove background from photonic conversions generated by interactions with the beampipe or detector material.

- $nhit > 2$ — This variable measures the number of VTX layers with registered hits corresponding to a track and is applied with the criteria above to ensure there is also a hit in one of the outer two layers in the VTX. This helps to minimize photonic conversions arising from interactions with the inner layers of the VTX.

- $|zed| < 75$ — This variable measures the z coordinate where the track crosses the PC1 detector subsystem.

Quantity Distributions: Distributions for many of the selection variables are shown in the following figures, in particular with all other selection requirements in place aside from the requirement on the displayed variable, so the cut region for each variable can be seen explicitly. Figure 3.13 shows the *prob* distribution, while Fig. 3.14 shows the $n0$ distribution, and Fig. 3.15 shows the *dep* distribution. The χ^2 per number of degrees of freedom distribution for track projections fitted to VTX hits

is shown in Fig. 3.16, and the distribution of the number VTX layers with registered hits for all tracks is shown in Fig. 3.17.

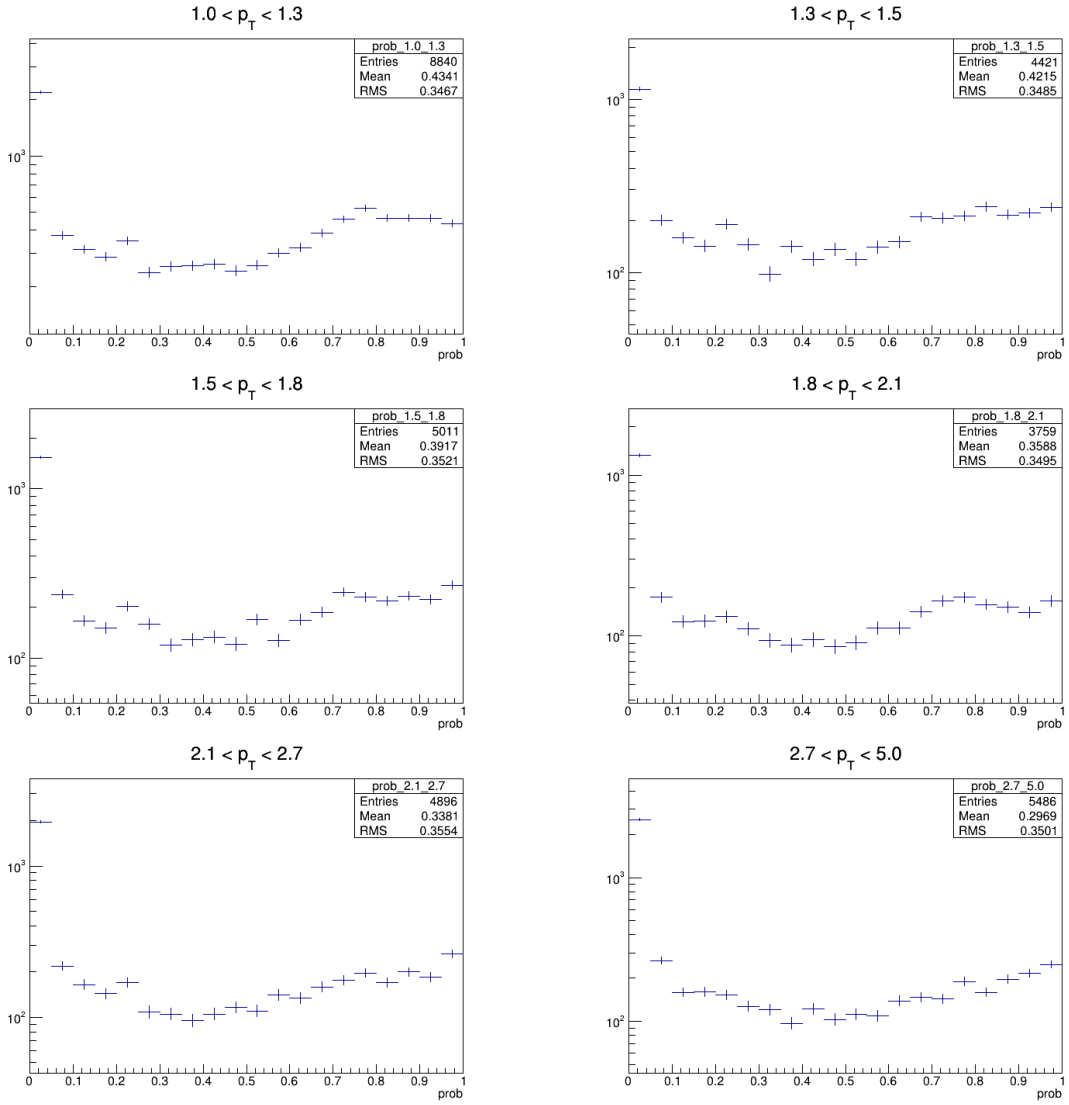

Figure 3.13: Distribution of the *prob* variable for candidates that pass all other electron identification and track selection requirements.

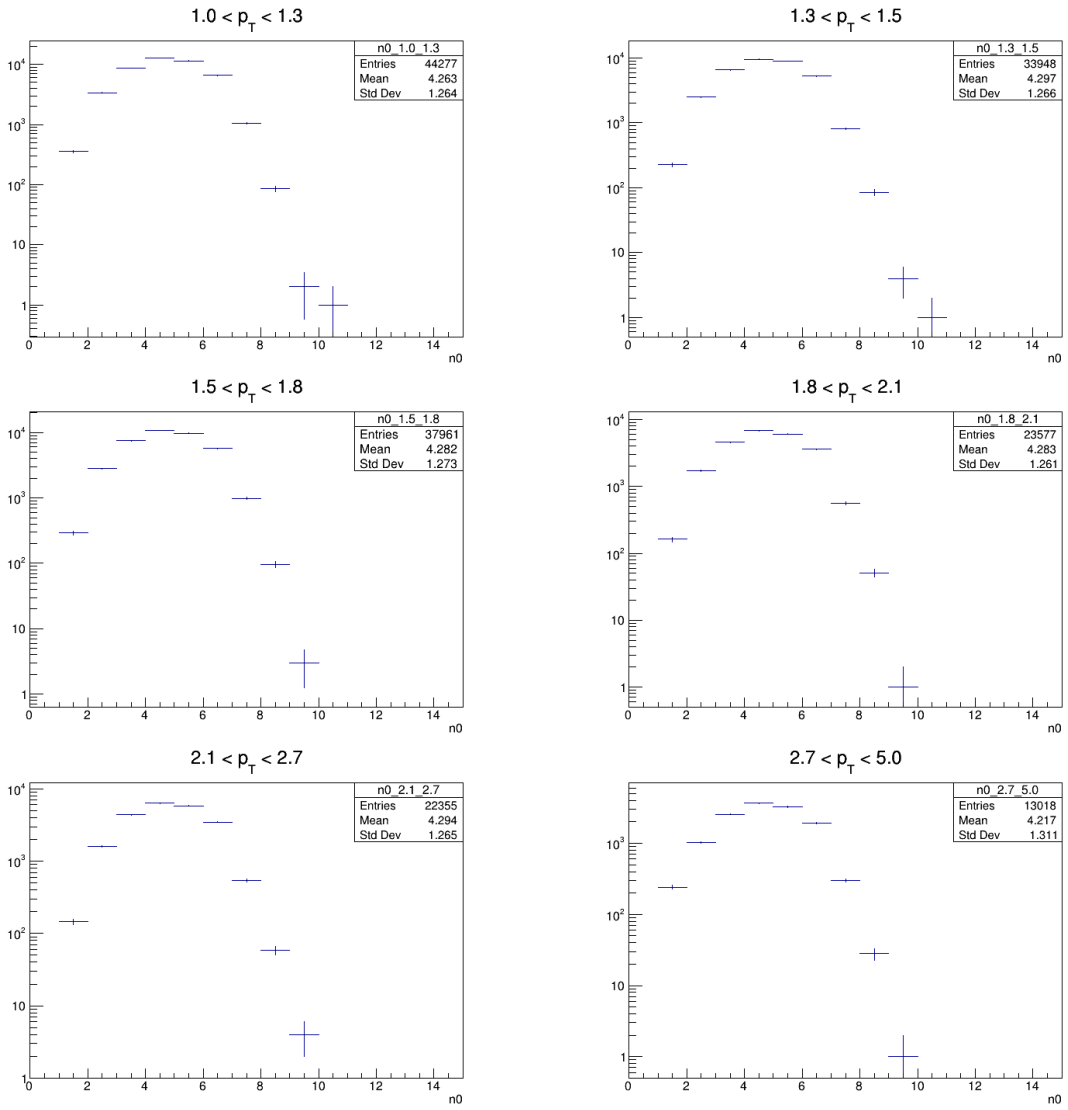

Figure 3.14: Distribution of the $n0$ variable for candidates that pass all other electron identification and track selection requirements, with an $n0 > 0$ cut to exclude the dominant hadron peak.

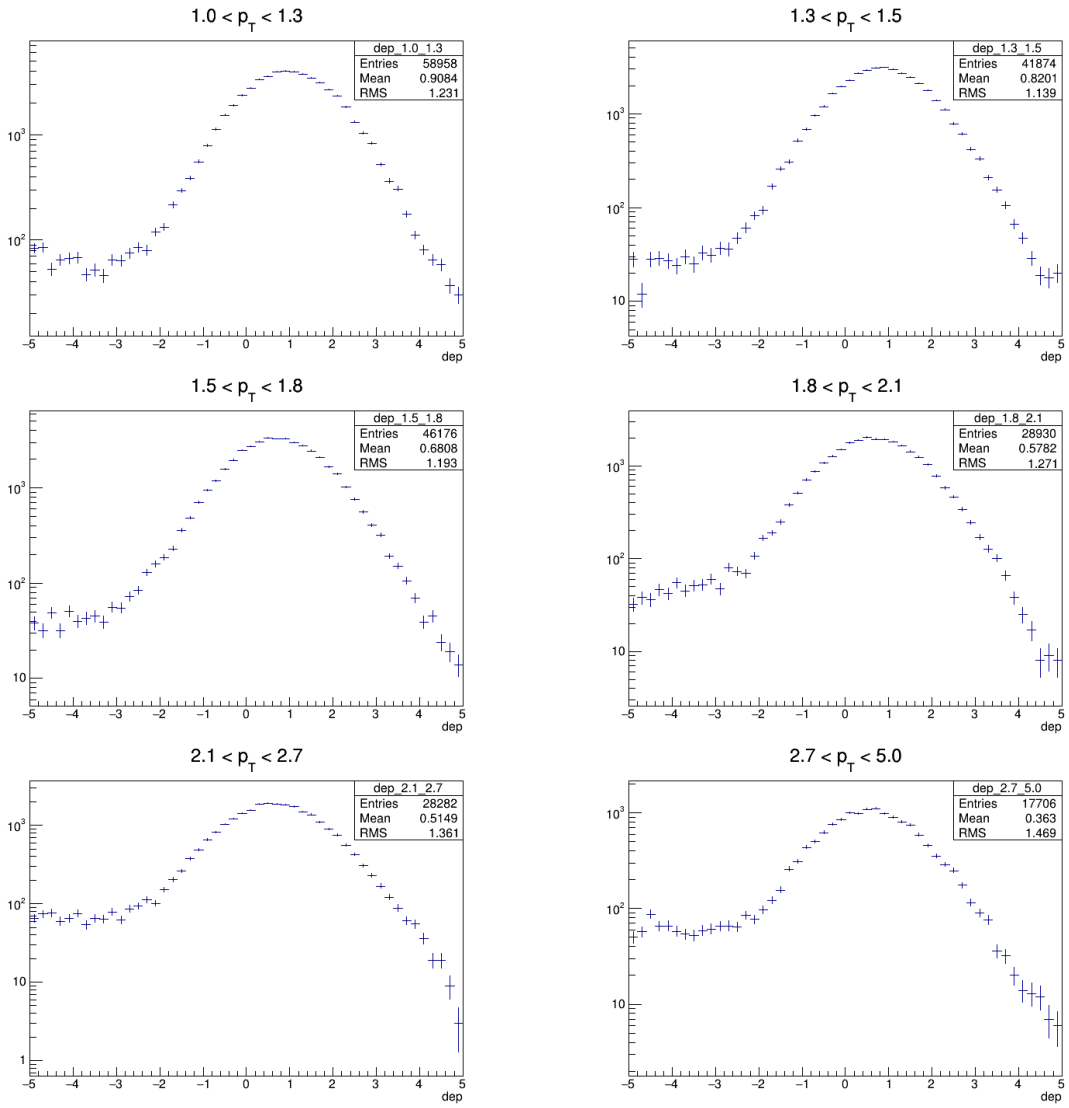

Figure 3.15: Distribution of the *dep* variable for candidates that pass all other electron identification and track selection requirements.

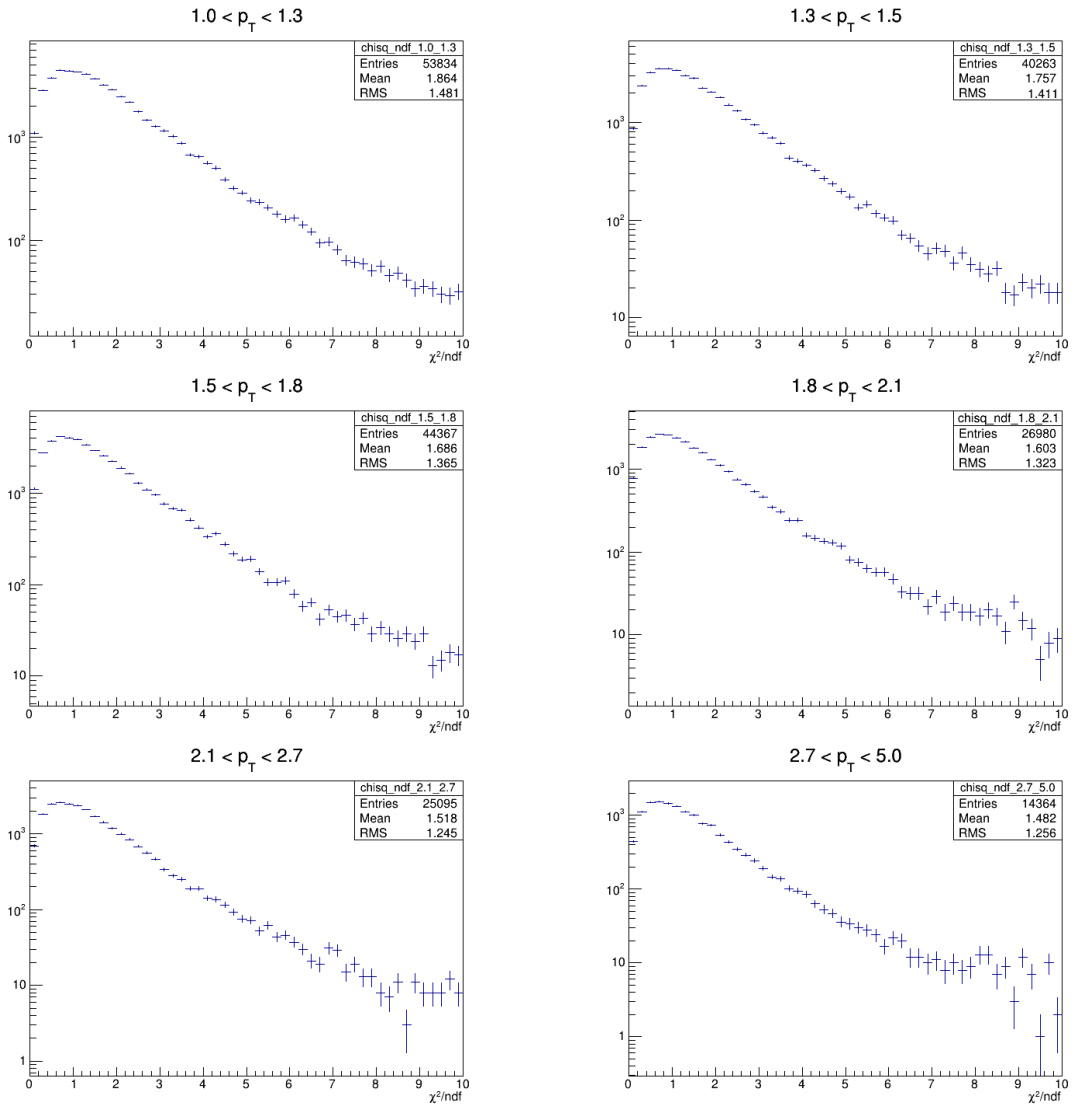

Figure 3.16: Distribution of the χ^2/ndf variable for candidates that pass all other electron identification and track selection requirements.

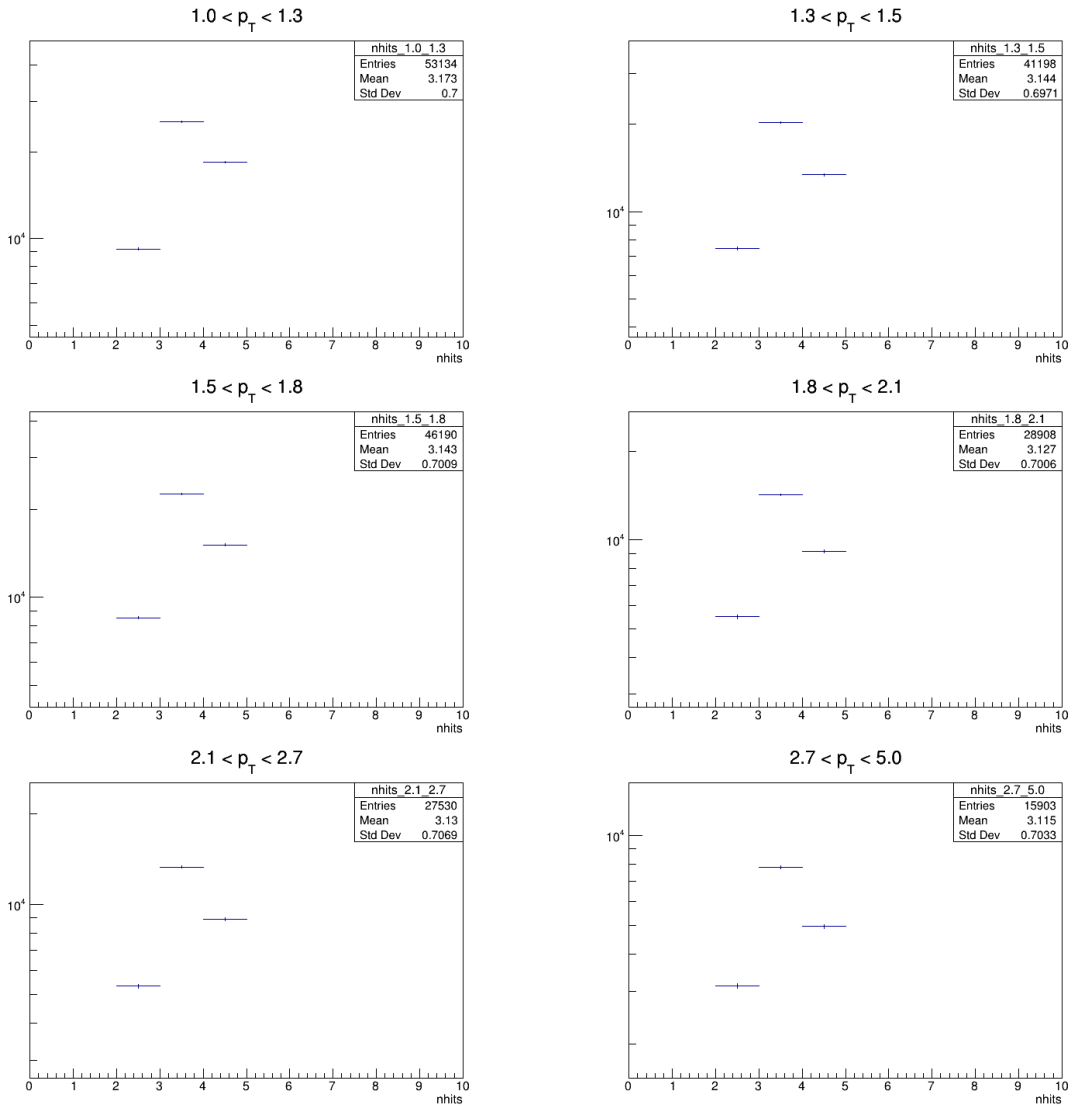

Figure 3.17: Distribution of the number of VTX layers with registered hits for candidates that pass all other electron identification and track selection requirements.

Spin Database Dependent Selection Requirement: The following selection requirement is to ensure only quality events are considered in calculating the relative luminosity.

- GL1P scaler counts per bunch crossing > 10000

3.2.1.2 The Conversion Veto Requirement

Electrons from photonic conversions $\gamma \to e^+e^-$ arising from interactions with the beam pipe or detector material comprise an important component of the background sources studied for this analysis. The VTX detector itself is a significant source of conversion electrons, with the inner and outer layers of the detector being 1.28% and 5.43% of a radiation length, respectively. However, the VTX detector can be used to reject conversion electrons based on the narrow opening angle of $\gamma \to e^+e^-$ conversions (much smaller than the opening angle of a typical hadronic decay). Tracks measured in the VTX detector are checked to see if any adjacent hits are observed within a narrow window size of $\Delta\phi \times \Delta z$, and are labeled as conversions and rejected if this criterion is met. It can be the case that uncorrelated hits from the event can lead to rejection via the conversion veto cut, ultimately decreasing the number of viable candidates to analyze. Therefore, the conversion veto window was studied in detail in Ref. [161] to determine the optimal size for the $\Delta\phi \times \Delta z$ window by achieving a low survival rate for photonic electrons from π^0 decays while maintaining a high survival rate for uncorrelated tracks (with charged hadrons used as a proxy since they are not produced via photonic conversions). Initial window sizes were determined in the heavy flavor electron separation analysis using the 2011 Au+Au dataset [167], but it was determined that the window size could be scaled by a factor of 2 to achieve optimal performance in the 2015 *pp* dataset, given the lower multiplicity of *pp* collisions leading to less uncorrelated hits in a given window of the VTX. Further information on the determination of the conversion veto window size and related studies can be found in Refs. [167] and [161].

3.2.1.3 Background Sources

All relevant background sources for the electron candidate sample in the transverse momentum range 1.5 GeV $< p_T <$ 5.0 GeV are listed in this section.

- **Misidentified hadrons $h^{\pm}$:** These consist of charged hadrons (primarily $\pi^{\pm}$) that pass our selection requirements and make it into the electron candidate sample.

- **Photonic electron sources:** These consist of Dalitz decays of neutral mesons $\pi^0 \to \gamma e^+ e^-$ and $\eta \to \gamma e^+ e^-$, as well as photonic conversions $\gamma \to e^+ e^-$, whether they are direct photons or photons from the decay of other hadrons.

- **Nonphotonic electron sources:** These consist of leptonic or semileptonic decays of hadrons (excluding Dalitz decays). Open heavy flavor decay electrons are also classified as nonphotonic electrons.

  - $J/\psi \to e^{\pm}$
  - $(K^{\pm}, K^0_S) \to e^{\pm}$ (Ke3)

Electrons from light vector meson decays, such as ρ, ω, and ϕ were considered in Ref. [161], but it was determined that they comprised a negligible source of background in the measured electron p_T range.

3.2.1.4 Hadron Contamination

Misidentified hadrons make up an important source of background in this analysis, which can be categorized in the following ways:

1. At low track p_T, misidentified hadrons arise primarily due to multiplicity effects — a hadron track can occasionally share a hit in the RICH with another charged particle track.

2. Approaching a track p_T of 5 GeV, charged pions will begin to fire the RICH. While this effect is minimized by limiting the p_T reach of this analysis to 1.5 GeV $< p_T <$ 5.0 GeV, it does cause an increase of the hadron contamination in the 2.7 GeV $< p_T <$ 5.0 GeV bin.

The hadron contamination is estimated with two separate methods that use independent detector subsystems, while the weighted average of the two methods is taken as the measured value, and the differences between the average and the input values are taken as systematic uncertainties. The two methods used in this analysis involve (i) the ratio of energy deposited in the electromagnetic calorimeter by the candidate track to the momentum of the track, and (ii) the survival rate of requiring hits in an annulus of the RICH corresponding to the candidate track. Method (i) is implemented with fits to the electron candidate *dep* distribution. Method (ii) is algebraic and based on the survival rate of the applied $n0$ cut, allowing for an estimate that is independent of *dep* shape.

Estimation Through *dep* Fitting: This method takes advantage of the different shapes of *dep* distributions for charged hadrons and electrons. The *dep* distribution of electrons should have a Gaussian shape with $\mu \approx 0$ and $\sigma \approx 1$. Hadrons in data are selected with a $n0 < 0$ cut — that is, charged tracks that did not fire the RICH photomultiplier tubes within a specific annulus. As shown in Figure 3.18, the *dep* distribution for hadrons in data is significantly different from that of electrons. The fits in Figure 3.18 serve as a template for fitting the electron candidate *dep* distribution, which is fit with a summation of the hadron *dep* template and a Gaussian in order to extract the fractional contribution from electrons and hadrons in each p_T bin. It should be noted that the only free parameter of the hadron *dep* fit template is the normalization. The resulting fits for electron candidates are shown in Figure 3.19. The hadron contamination fraction is measured as the integral of the hadron fit template within the *dep* cut region $|dep| < 2$ over the integral of the total fit in this region. This is shown in Figure 3.21 in the left panel, plotted alongside results obtained from the algebraic method.

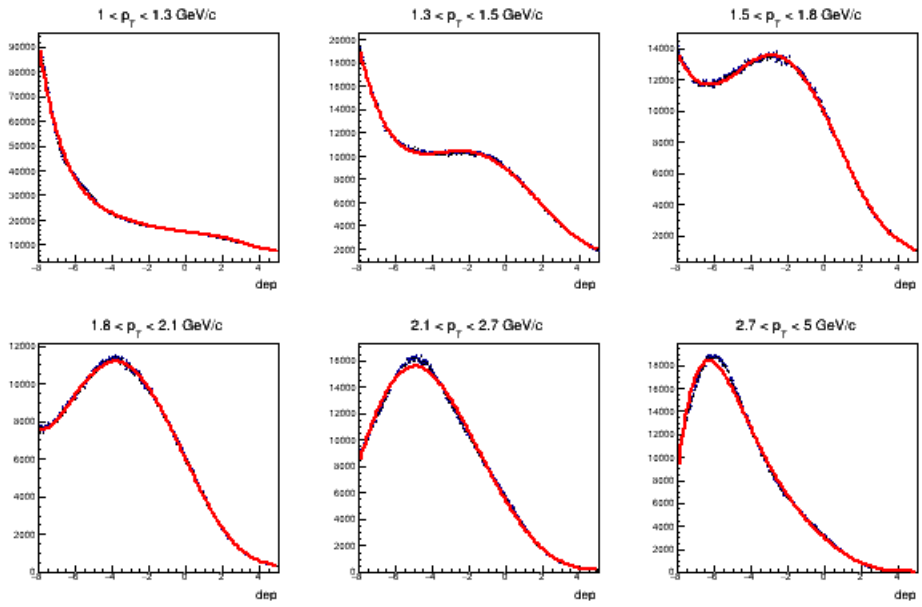

Figure 3.18: Distribution of the *dep* variable for hadron tracks in data in each p_T bin, fit in order to construct a template for the shape of the hadron contamination in the electron candidate *dep* spectrum.

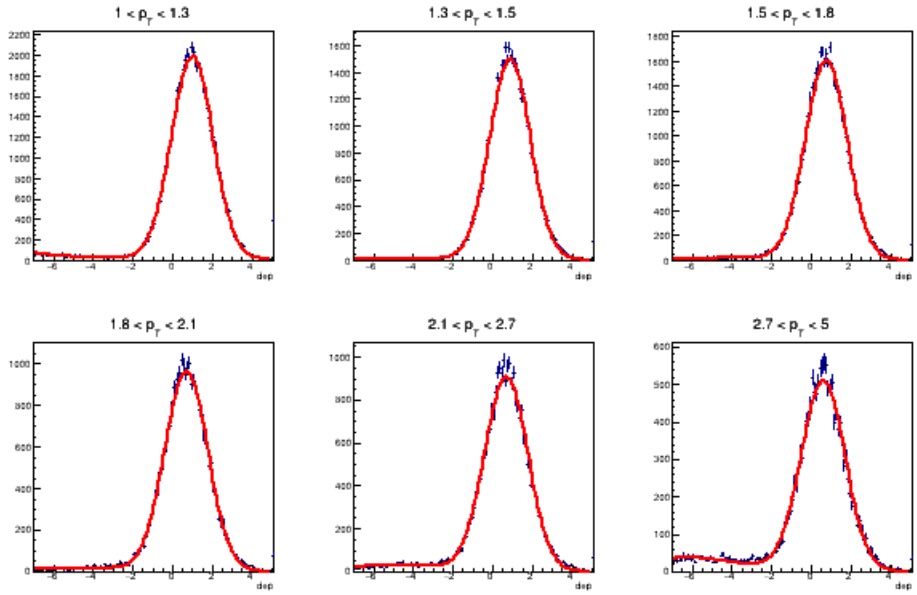

Figure 3.19: Distribution of the *dep* variable for electron candidates in each p_T bin, fit by a Gaussian + the hadron contamination template from Figure 3.18.

Algebraic Estimation: A separate method to estimate the hadron contamination that is independent of the electron candidate *dep* shape involves using the survival rate of the $n0 > 1$ cut. Electrons and charged hadrons (which are predominantly pions) fire the RICH differently due to their different mass and therefore have a different survival rate. The survival rate of the applied $n0$ cut for electrons ϵ_e was obtained from a simulated sample of electrons, while the survival rate for hadrons ϵ_h was estimated by isolating hadrons in data. The results of this study are shown in Figure 3.20. For this method, hadrons in data are accessed by applying a selection requirement of $dep < -6$ on the electron candidate sample, well into the region of the *dep* distribution dominated by hadron contamination (as seen in Figures 3.18 and 3.19). Equations 3.1 – 3.2 show the system of equations used to derive the number of hadrons in our electron candidate sample, which is given by Equation 3.3.

$$n_{non0} = n_e + n_h \tag{3.1}$$

$$n_{n0} = \epsilon_e n_e + \epsilon_h n_h \tag{3.2}$$

$$n_{h_{n0}} = \epsilon_h \frac{n_{n0} - \epsilon_e n_{non0}}{\epsilon_h - \epsilon_e} \tag{3.3}$$

Here n_{non0} is the number of electron candidates passing all analysis cuts described in Section 3.2.1.1 except for the $n0$ cut, n_{n0} is the number of electron candidates passing all analysis cuts, and $n_{e,h}$ are the number of electrons and hadrons in the n_{non0} sample respectively. The ratio of $n_{h_{n0}}/n_{n0}$ is the measured hadron contamination shown in the left panel of Figure 3.21, plotted along with the results obtained from the *dep* fitting method.

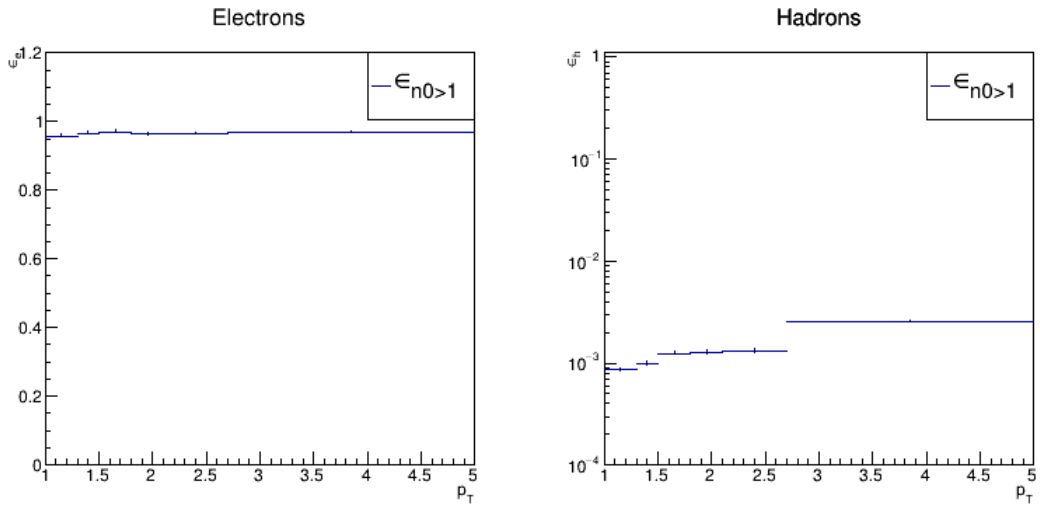

Figure 3.20: Survival rates of the $n0 > 1$ cut for both single electrons from simulation ϵ_e (left panel) and charged hadrons in data ϵ_h (right panel).

Measured Hadron Contamination: For both methods, the fraction of misidentified hadrons in our electron candidate sample is estimated and the results are shown in Figure 3.21 on the left panel. The weighted average of the two independent measurements was taken based on the statistical uncertainties in order to extract the measured hadron contamination. The difference between the value obtained from each method and the average value is taken as the systematic uncertainty for the upper and lower bounds of the measurement, as shown in Figure 3.21, where the right panel shows the measured hadron contamination fractions used in this analysis.

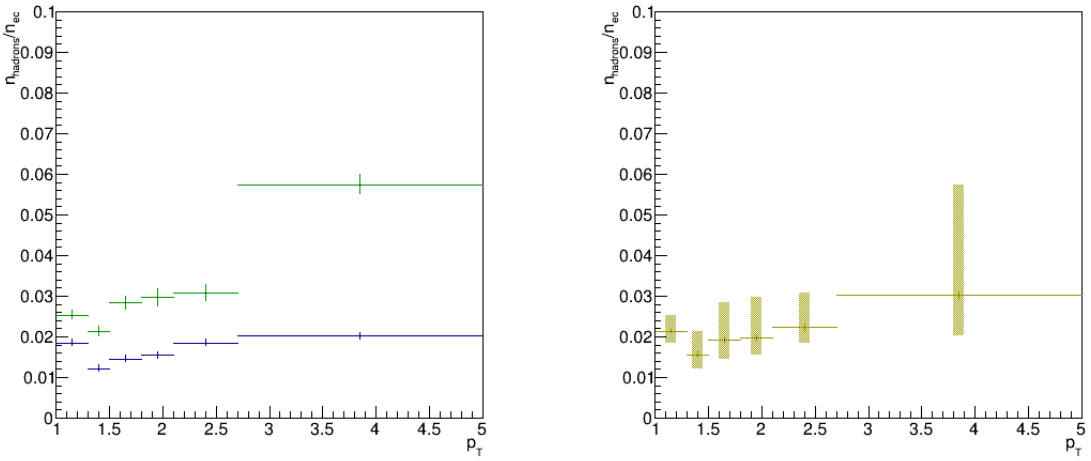

Figure 3.21: Hadron contamination extracted using both a *dep* fitting method and algebraic method. The hadron contamination fraction shown in the right panel is the weighted average of the two results shown in the left panel, and it is referred to as $f_{h^{\pm}} = \tilde{f}_{hc}$.

3.2.1.5 Electron Cocktail

In order to determine the background fractions of electrons from all other sources listed in Section 3.2.1.3, an electron "cocktail" with all relevant background sources was constructed in Ref. [161]. Information about the simulated electron charge was saved additionally for this analysis and used to construct the cocktail separately for electrons and positrons. The p_T distribution of the electron cocktail is shown in Figure 3.22, where the p_T weighting was done in accordance with past PHENIX $\sqrt{s} =$ 200 GeV cross section measurements at midrapidity. The ratio of yields of nonphotonic background sources in the electron cocktail to π^0 yields (the most prominent photonic background source) is shown in Figure 3.23.

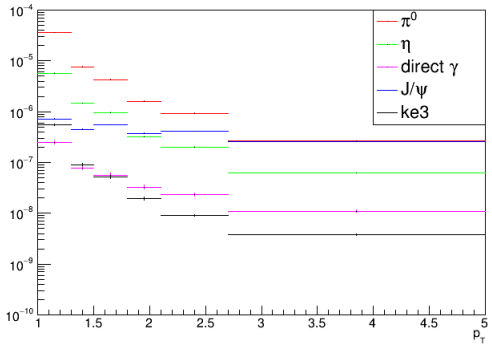

Figure 3.22: Electronic cocktail yields before normalizing to the electron candidate spectrum, referred to as $\tilde{n}_{i,j}$.

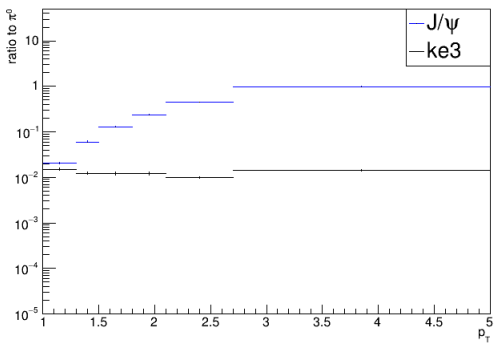

Figure 3.23: Ratio of nonphotonic electron background yields to π^0 yields, referred to as $\tilde{n}_j/\tilde{n}_{\pi^0}$.

The electron cocktail must still be normalized with respect to the measured electron candidate spectrum in order to extract the proper background fractions for each source. To achieve this, the fractions of photonic and nonphotonic electrons in our sample are measured with the use of the conversion veto cut described in Section 3.2.1.2. The open heavy flavor signal electrons belong to the nonphotonic electron sample, while the photonic electron sample is purely background, providing an opportunity to normalize the electron cocktail to the total number of measured electron candidates in data. In order to calculate the fraction of nonphotonic electrons in our electron candidate sample, F_{np}, we exploit the fact that the conversion veto cut affects the photonic and nonphotonic electrons differently. The survival rate of the conversion veto cut for nonphotonic electrons is due solely to uncorrelated tracks triggering the conversion veto cut, ϵ_{uc}, while the survival rate for photonic electrons is due to this as well as actual conversions being vetoed, $\epsilon_{uc}\epsilon_p$. Figure 3.24 shows ϵ_{uc} for different conversion veto window sizes as calculated from hadrons in data. It should be noted that the 2x conversion veto window described in Section 3.2.1.2 was chosen in this analysis, just as in Ref. [161], due to having a low survival rate of photonic electrons while the survival rate for nonphotonic electrons remains high and relatively flat

across p_T. The results for ϵ_{uc} for the 2x window are shown in Figure 3.25. Figure 3.26 shows the survival rate for simulated photonic electron sources, with the weighted average of the survival rate from different sources taken as ϵ_p, which is plotted again in Figure 3.27 for clarity.

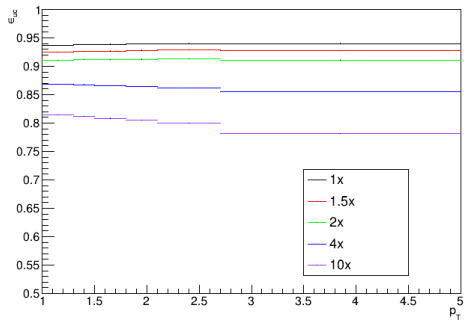

Figure 3.24: Survival rates for hadrons in data for various conversion veto cuts, with 2x used for this analysis

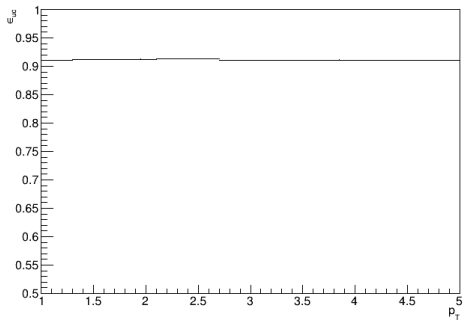

Figure 3.25: Survival rate of 2x conversion veto cut for uncorrelated tracks, taken from Figure 3.24, ϵ_{uc}

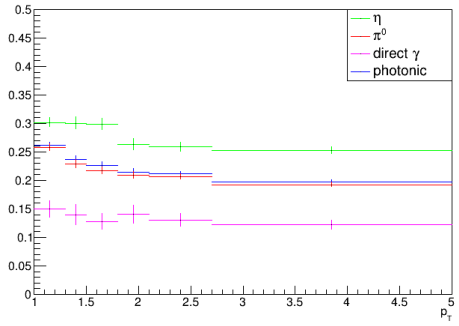

Figure 3.26: Survival rates for various photonic electron sources, the weighted average of which is ϵ_p.

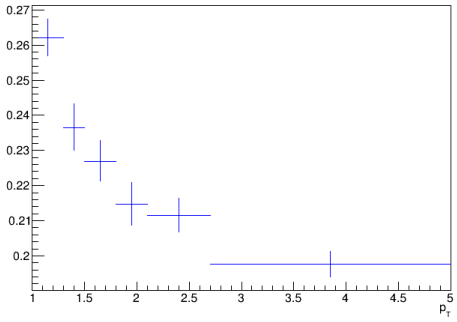

Figure 3.27: Survival rate of 2x conversion veto cut for photonic electrons from Figure 3.26 (ϵ_p).

Calculation of F_{np}: The conversion veto survival rates for nonphotonic (ϵ_{uc}) and photonic ($\epsilon_{uc}\epsilon_p$) electrons imply the validity of Equation 3.4 and 3.5:

$$n_e = n_{np} + n_p + n_{hc} \tag{3.4}$$

$$\tilde{n}_e = \tilde{n}_{np} + \tilde{n}_p + \tilde{n}_{hc} = \epsilon_{uc} n_{np} + \epsilon_{uc}\epsilon_p n_p + \tilde{n}_{hc} \tag{3.5}$$

where n_e represents the number of electron candidates, n_{np} is the number of nonphotonic electrons, n_p the number of photonic electrons, and $n_{hc} = f_{hc} n_e$ ($\tilde{n}_{hc} = \tilde{f}_{hc}\tilde{n}_e$) the number of misidentified hadrons. The quantities with a tilde overhead represent yields with the conversion veto applied, where the survival rates ϵ_{uc} and ϵ_p come from Figures 3.25 and 3.27 respectively. It follows from Equation 3.5 that the only unknowns in this system of equations are n_{np} and n_p, since the hadron contamination procedure described in Section 3.2.1.4 can be carried out with and without the conversion veto cut applied. Therefore, this system of equations can be solved and used to extract the nonphotonic electron fraction F_{np}, as shown in Equation 3.6.

$$F_{np} = \frac{\tilde{n}_{np}}{\tilde{n}_{np} + \tilde{n}_p} = \frac{n_{np}}{n_{np} + \epsilon_p n_p} = \frac{\epsilon_{uc}\epsilon_p n_e - \tilde{n}_e - \epsilon_{uc}\epsilon_p n_{hc} + \tilde{n}_{hc}}{(\epsilon_p - 1)(\tilde{n}_e - \tilde{n}_{hc})} \tag{3.6}$$

Figures 3.28 and 3.29 show the electron candidate yields with and without the conversion veto cut applied for the charge combined sample, i.e. $\tilde{n}_e$ and n_e respectively. It should be noted that the hadron contamination was recalculated without the conversion veto cut to obtain f_{hc} and therefore n_{hc}, giving us all necessary inputs for the F_{np} calculation shown in Equation 3.6. Figure 3.30 displays the results of the F_{np} calculation, which are used as input to the background fraction normalization procedure.

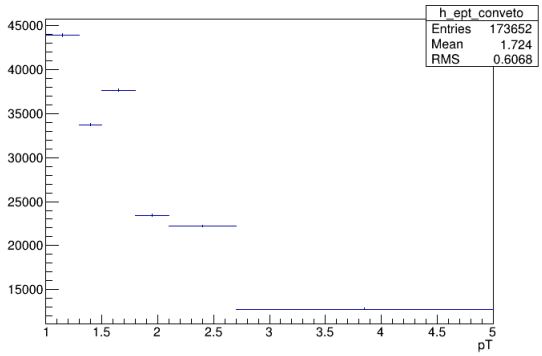

Figure 3.28: Electron candidate p_T spectrum with the conversion veto cut applied ($\tilde{n}_e$).

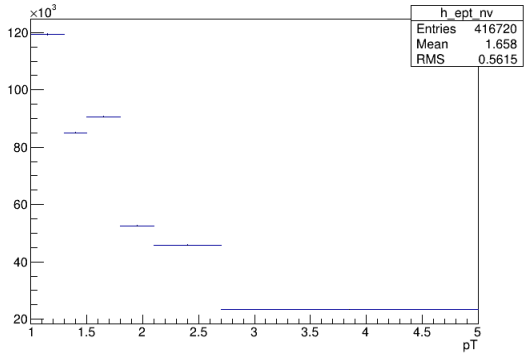

Figure 3.29: Electron candidate p_T spectrum without the conversion veto cut applied (n_e).

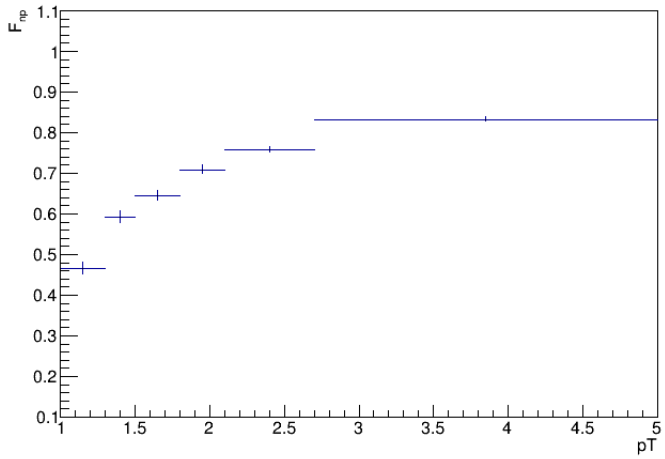

Figure 3.30: Fraction of nonphotonic electrons used to normalize the background fractions to our electron candidate spectrum (F_{np}).

Normalizing to the Electron Candidate Sample: As mentioned previously, F_{np} provides a means of normalizing the electron cocktail yields to the total electron candidates in the data. Since the photonic electrons in our sample are purely background, we can use F_{np} as shown in Equation 3.7 to normalize the photonic background fractions $i = \pi^0, \eta, \gamma$:

$$f_i = (1 - \tilde{f}_{hc})(1 - F_{np})\frac{\tilde{n}_i}{\tilde{n}_{\pi^0} + \tilde{n}_\eta + \tilde{n}_\gamma} \tag{3.7}$$

where $\tilde{f}_{hc}$ is the hadron contamination fraction with the conversion veto cut applied (shown in Figure 3.21), and the values $\tilde{n}_i$ are the photonic electron yields from the electron cocktail shown in Figure 3.22. We can't use Equation 3.7 for the nonphotonic background sources since our signal constitutes a part of the nonphotonic electron sample. We can however normalize the nonphotonic background fractions using the calculated π^0 background fraction f_{π^0}, and the electron cocktail yield fractions $\tilde{n}_j/\tilde{n}_{\pi^0}$ shown in Figure 3.23. Equation 3.8 is therefore used to calculate the background fraction for nonphotonic sources $j = J/\psi, Ke3$.

$$f_j = f_{\pi^0}\frac{\tilde{n}_j}{\tilde{n}_{\pi^0}} \tag{3.8}$$

3.2.1.6 Calculated Background Fractions

The resulting background fractions for photonic electron sources f_i, nonphotonic electron sources f_j, and hadrons misidentified as electrons, $f_{h^\pm} = \tilde{f}_{hc}$ are shown in Figure 3.31. The final results of the analysis are split by charge, therefore the background fractions are also calculated for the separate charges, shown in Figures 3.32 and 3.33 for positrons and electrons respectively. The systematic uncertainties on the background fractions were calculated by modifying the functions used to weight the electron cocktail p_T spectrum with uncertainty factors extracted from the fits. The electron cocktail normalization procedure described in Section 3.2.1.5 was repeated for 1500 samplings of the weighting functions, recalculating ϵ_p, F_{np}, and $f_{i,j}$ for each iteration. A resulting distribution of $f_{i,j}$ was created from which the RMS value was assigned as a systematic uncertainty on the reported background fractions. The calculation of the hadron contamination fraction $f_{h^\pm}$ and corresponding systematic uncertainties are discussed in Section 3.2.1.4.

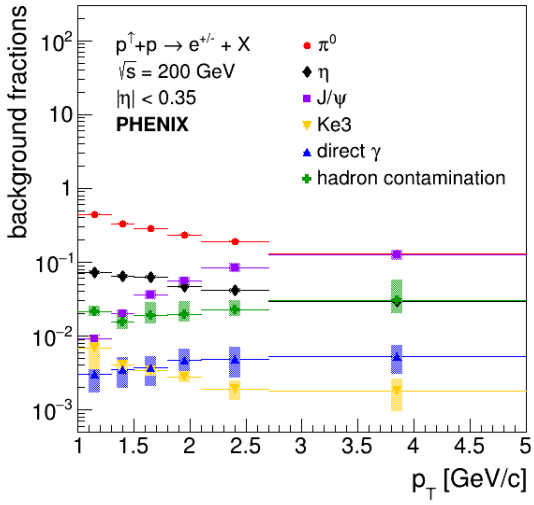

Figure 3.31: Fraction of measured electron and positron candidates attributed to each background source, used as an input to Equation 3.21 ($f_{i,j,h^{\pm}}$). This figure was taken from Ref. [1].

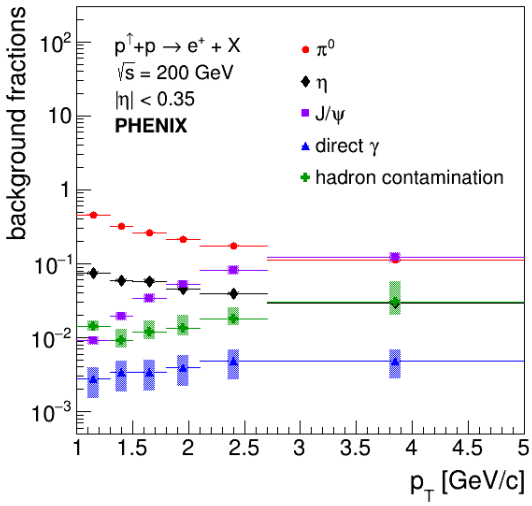

Figure 3.32: Fraction of measured positron candidates attributed to each background source, used as an input to Equation 3.21 ($f_{i,j,h^{+}}$). This figure was taken from Ref. [1].

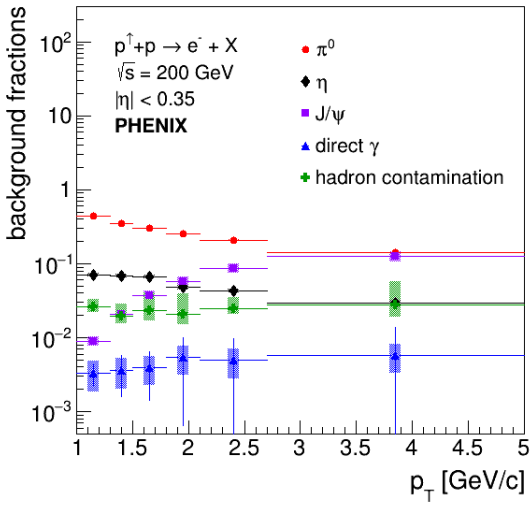

Figure 3.33: Fraction of measured electron candidates attributed to each background source, used as an input to Equation 3.21 ($f_{i,j,h^{-}}$). This figure was taken from Ref. [1].

Table 3.2: Fractions of background $f_{i,j,h^{\pm}}$ present in each p_T bin for the open heavy flavor positrons and electrons, used as inputs to the background correction procedure, and shown in Figs. 3.32 and 3.33 respectively. This table was taken from [1].

$e^{\pm}$	p_T range (GeV/c)	$\langle p_T \rangle$ (GeV/c)	$f_{\pi^0 \to e^{\pm}}$	$f_{\eta \to e^{\pm}}$	$f_{\gamma \to e^{\pm}}$	$f_{J/\psi \to e^{\pm}}$	$f_{h^{\pm}}$
e^+	1.0 – 1.3	1.161	0.458	0.0738	0.00274	0.00916	0.0140
	1.3 – 1.5	1.398	0.318	0.0592	0.00336	0.0195	0.00924
	1.5 – 1.8	1.639	0.264	0.0582	0.00339	0.0344	0.0120
	1.8 – 2.1	1.936	0.215	0.0458	0.00399	0.0520	0.0134
	2.1 – 2.7	2.349	0.173	0.0394	0.00481	0.0823	0.0179
	2.7 – 5.0	3.290	0.111	0.0297	0.00480	0.122	0.0300
e^-	1.0 – 1.3	1.161	0.439	0.0704	0.00335	0.00900	0.0261
	1.3 – 1.5	1.398	0.347	0.0692	0.00364	0.0206	0.0198
	1.5 – 1.8	1.639	0.299	0.0665	0.00394	0.0375	0.0230
	1.8 – 2.1	1.936	0.252	0.0478	0.00535	0.0577	0.0205
	2.1 – 2.7	2.349	0.208	0.0429	0.00490	0.0872	0.0245
	2.7 – 5.0	3.290	0.143	0.0296	0.00572	0.127	0.0279

3.2.2 Neutral Mesons

Both π^0 and η particles are pseudoscalar mesons with respective valence quark content $\frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d})$ and $\approx \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s})$. Transverse single-spin asymmetry measurements of π^0 and η mesons provide access to initial and final state spin-momentum correlations in the proton and/or process of hadronization, while comparing measurements in both systems can potentially yield insight on effects coming from strange quarks. This measurement was carried out on the $\sqrt{s_{NN}} = 200$ GeV $p^{\uparrow}$Au and $p^{\uparrow}$Al data sets from the 2015 running year. Both π^0 and η mesons are measured via the 2γ decay channel, with branching ratios $BR(\pi^0 \to 2\gamma) \approx 99\%$ and $BR(\eta \to 2\gamma) \approx 39\%$ [6], and an invariant mass peak of $m_{\pi^0} \approx 135$ MeV and $m_{\eta} \approx 548$ MeV respectively. Photons are measured with the electromagnetic calorimeters (EMCal), while the tracking subsystems (DC, PC1, PC2, PC3) are used to veto electromagnetic clusters coming from charged tracks, and the collision vertex is measured by the Beam Beam Counter (BBC). The analysis methods follow very closely from that of Ref. [123], where the same observable was measured for $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$

with the 2015 PHENIX pp dataset. There is an important distinction that only one beam (the proton beam) is transversely polarized in the 2015 pA collisions, while both beams were transversely polarized in the 2015 pp collisions.

3.2.2.1 Selection Criteria

The p_T bins selected for this analysis are shown in Tables 3.3 and 3.4 for the π^0 and η measurements, respectively. This binning is the same as that of Ref. [123].

Table 3.3: p_T bins chosen for the π^0 analysis shown with average p_T values obtained in each bin for pAu and pAl collision systems after applying all analysis cuts.

p_T range (GeV)	$\langle p_T \rangle$ (pAu)	$\langle p_T \rangle$ (pAl)
2 – 3	2.711	2.676
3 – 4	3.731	3.468
4 – 5	4.311	4.412
5 – 6	5.400	5.410
6 – 7	6.414	6.417
7 – 8	7.423	7.424
8 – 9	8.431	8.433
9 – 10	9.438	9.439
10 – 12	10.780	10.790
12 – 20	13.520	13.560

Table 3.4: p_T bins chosen for the η analysis shown with average p_T values obtained in each bin for pAu and pAl collision systems after applying all analysis cuts.

p_T range (GeV)	$\langle p_T \rangle$ (pAu)	$\langle p_T \rangle$ (pAl)
2 – 3	2.640	2.642
3 – 4	3.443	3.461
4 – 5	4.406	4.416
5 – 6	5.403	5.408
6 – 7	6.413	6.417
7 – 8	7.422	7.425
8 – 10	8.740	8.742
10 – 20	11.650	11.700

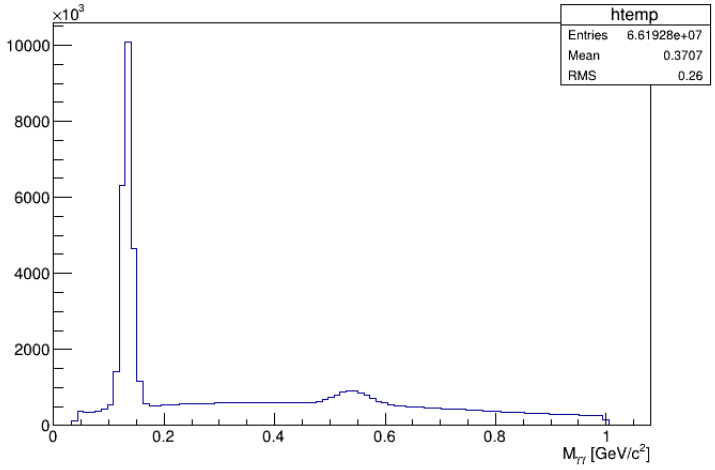

Figure 3.34: Invariant mass of all reconstructed photon pairs with $M_{\gamma\gamma} > 1$ GeV/c^2 – pAu.

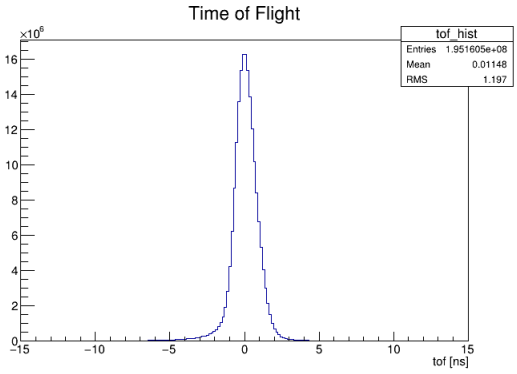

Figure 3.35: Time of flight distributions for ERT trigger photons – pAu.

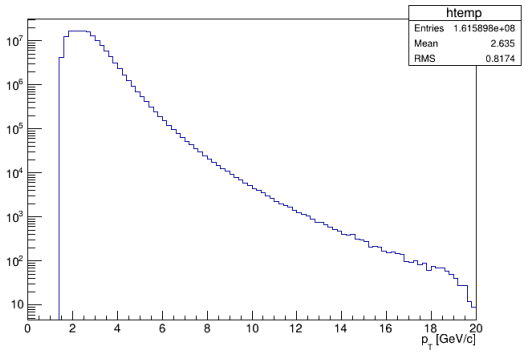

Figure 3.36: p_T distributions for ERT trigger photons – pAu.

The selection requirements used in this analysis follow from Ref. [123], where the same observable was measured for π^0 and η mesons in 2015 proton-proton PHENIX dataset. It should be noted that all cuts outlined below for the two particles are the same aside from that on the invariant mass distribution.

Event Selection: The selection requirements that each event in the analysis had to pass are listed below.

- One of the EMCal-RICH triggers (ERTB, ERTA or ERTC) had to fire.
- The z vertex position as measured by the Beam Beam Counter was required to satisfy $|z_{vtx}| \leq 30$ cm.

Photon Selection: The following selection requirements were placed on electromagnetic clusters in order to obtain a list of photon candidates for this analysis.

- $0.5 \leq E_{core} \leq 20.0$ GeV — This variable measures the energy deposited in the central tower of a calorimeter cluster from an electromagnetic shower [168]. This is motivated by the majority of the energy in an electromagnetic shower being concentrated in the center.

- Shower shape — A number of cuts were placed on the electromagnetic shower shape to increase the purity of photons in our dataset by rejecting charged hadrons and electrons.
  - $\chi^2 < 3$ (PbSc Sectors) — This variable determines the goodness of fit of the measured cluster to a typical electromagnetic cluster profile, with lower χ^2 values corresponding to an increased likelihood that the cluster came from an electromagnetic shower. This variable is not accessible in PbGl clusters, so a dispersion cut is used in place.
  - $disp < p_1 + p_2\theta + p_3\theta^2$ (PbGl sectors) — This variable is related to the second moment of the calorimeter tower position distribution corresponding to a given EM cluster. Here θ is the incident angle of the cluster on the calorimeter surface, and the parameters p_1, p_2 and p_3 were taken from previous PHENIX analyses.
  - $prob_\gamma > 0.02$ — The variable measures the probability that an electromagnetic shower came from a photon, based on the shower shape. This helps eliminate background from charged hadrons, which tend to have a wider shower in the EMCal.
- Tower quality — Calorimeter towers had to pass quality assurance cuts, including (i) exclusion of hot and dead towers (described in Section 3.1.2.1), and (ii) exclusion of towers on the edge of a calorimeter sector. Once a list of bad towers was generated, electromagnetic clusters centered in the list of bad towers or directly adjacent towers were rejected.
- $|tof| < 5$ ns — This variable measures the time of flight of the photon candidate. Figure 3.35 shows a histogram of the time of flight for all trigger photons in the pAu dataset (the pAl results look similar). They have a Gaussian shape centered around zero with a tail for $t < 0$ which comes from pile-up events that

are eliminated by the tof cut.

- Charged track veto — Implemented to eliminate clusters that are geometrically associated with a track in order to suppress the background from electrons and charged hadrons. More specifically, clusters were checked to see that they did not match with the angle of a track as measured by the PC3 tracking detector with $\Delta z < 12$ cm and $\Delta\phi < 8$ cm.
  - The matching track was required to have quality ≥ 7 and $p_T > 0.5$ GeV. If no charged track was found the photon passed this cut.

Photon Pair Selection: The candidates passing all of the photon selection criteria within the same event were then matched into pairs. Figure 3.34 shows the invariant mass spectrum for all photon pairs with $M_{\gamma\gamma} > 1$ GeV in pAu collisions (the distribution for pAl collisions is similar). The photon pairs are required to pass the following requirements.

- Both clusters were required to be in the same arm.
- $\Delta R > 8$ cm — This variable measures the distance between the centers of EM clusters in the photon pair.
- $\alpha = |E_1 - E_2|/(E_1 + E_2) < 0.8$ — This variable measures the energy asymmetry of the two photons in the pair.
- $112 < M_{\gamma\gamma} < 162$ MeV (π^0 mesons) — This is the signal range in the diphoton invariant mass spectrum corresponding to π^0 candidates.
  - The sidebands of the invariant mass distribution, $47 < M_{\gamma\gamma} < 97$ MeV and $177 < M_{\gamma\gamma} < 227$ MeV, are used to study the combinatorial background under the signal peak.

- $480 < M_{\gamma\gamma} < 620$ MeV (η mesons)— This is the signal range in the diphoton invariant mass spectrum corresponding to η candidates.
  - The sidebands of the invariant mass distribution, $300 < M_{\gamma\gamma} < 400$ MeV and $700 < M_{\gamma\gamma} < 800$ MeV, are used to study the combinatorial background under the signal peak.
- The supermodule of the trigger photon was required to match with the supermodule that fired one of the EMCal-RICH triggers.
- $p_T^{trig} > 1.5$ GeV — This variable corresponds to the transverse momentum of the photon that fired the ERT. Figure 3.36 shows the p_T distribution for trigger photons with $p > 1.0$ GeV from the pAu data (the distribution looks similar for the pAl data).
- Finally, lower energy photons can pass the ERT check when they are detected in the supermodule that fired the ERT, even if they are not the photon that fired the ERT. The p_T cut described above eliminates the majority of these photons, but the highest energy photon in the photon pair is required to be the highest energy photon in the event to further reduce these occurrences.

3.2.2.2 Quantifying Backgrounds

The background fraction is an estimate of the ratio of photon pairs within the π^0 (η) meson peak of $0.112 < M_{\gamma\gamma} < 0.162$ GeV ($0.480 < M_{\gamma\gamma} < 0.620$ GeV) that do not come from $\pi^0 \to \gamma\gamma$ ($\eta \to \gamma\gamma$) decays. The mathematical expression is shown in Equation 3.9,

$$r = N_B/(N_S + N_B) = B/(S + B). \tag{3.9}$$

The quantities S and B denote the number of candidates under the signal peak corresponding to the signal and combinatorial background respectively. The behavior

of the combinatorial background depends on both the p_T of the photon pair and the arm of the EMCal that they are measured in, especially given that the lead glass sectors are only in the east arm. Thus, invariant mass histograms for each p_T bin and arm are constructed and then fit with a Gaussian + third-order polynomial with a fit range 0.07 to 0.25 GeV (0.4 to 0.7 GeV) for the π^0 (η). The third-order polynomial fit can be seen to describe the combinatorial background under the signal peak well for π^0 mesons in pAu (panel a) and pAl collisions (panel b) and η mesons in pAu (panel c) and pAl collisions (panel d) of Figure 3.37. The third-order polynomial curves are then integrated from $0.112 < M_{\gamma\gamma} < 0.162$ GeV ($0.480 < M_{\gamma\gamma} < 0.620$ GeV) for π^0 (η) candidates to calculate B in Eq. 3.9, while $S + B$ is calculated by integrating histogram counts within the signal regions. Tables 3.5 and 3.6 show the results of this calculation for π^0 in pAu and pAl collisions, and η in pAu and pAl collisions, respectively. The "East Arm" and "West Arm" values are used for the background correction of the relative luminosity formula results and the "Both Arm" quantities are used for the square root formula. The "Both Arm" values are the average of the west and east arm background fractions, weighted by the histogram counts for each arm in the invariant mass ranges $0.112 < M_{\gamma\gamma} < 0.162$ GeV and $0.480 < M_{\gamma\gamma} < 0.620$ GeV for π^0 and η candidates, respectively.

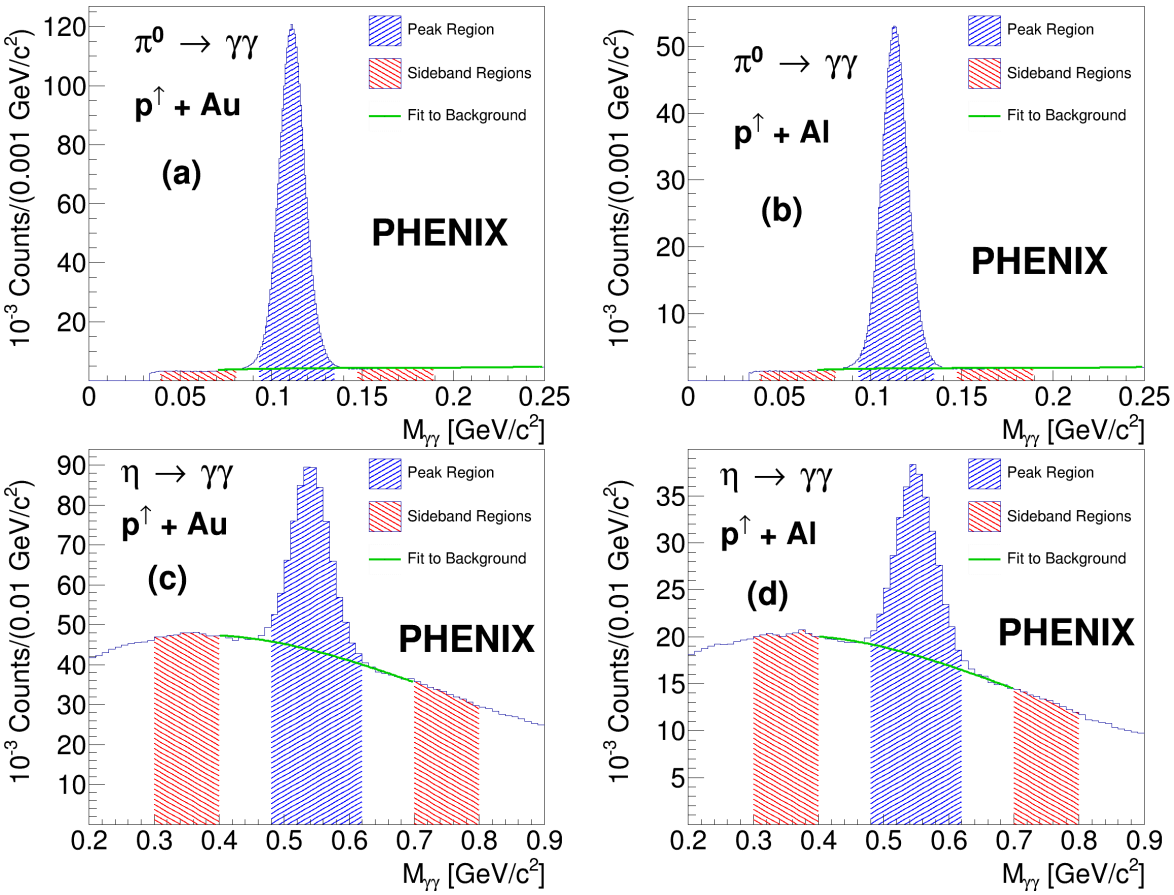

Figure 3.37: Invariant mass distributions around the $\pi^0 \to \gamma\gamma$ peak in (a) $p^{\uparrow}$Au collisions and (b) $p^{\uparrow}$Al collisions and around the $\eta \to \gamma\gamma$ peak in (c) $p^{\uparrow}$Au collisions and (d) $p^{\uparrow}$Al collisions for photon pairs within $4 < p_T$ [GeV/c] < 5 in the west central-arm spectrometer. The [blue] leftward-hatched regions are the signal peaks, used for quantifying yields for the A_N calculations, the [red] rightward-hatched regions are the sidebands, used to quantify yields for the A_N^{BG} calculations, and the [green] solid curves correspond to fits to the combinatorial background, used in calculating the background fractions. This figure was taken from Ref. [2].

Table 3.5: Background fractions in the π^0 invariant mass signal region in pA collisions for the west arm, east arm, and both combined.

pA	p_T [GeV]	r (West Arm)	r (East Arm)	r (Both Arms)
pAu	2 - 3	0.142	0.136	0.139
	3 - 4	0.102	0.1	0.101
	4 - 5	0.0848	0.0842	0.0845
	5 - 6	0.0773	0.0762	0.0768
	6 - 7	0.0764	0.074	0.0752
	7 - 8	0.0802	0.0743	0.0773
	8 - 9	0.0868	0.0766	0.0817
	9 - 10	0.078	0.0716	0.0748
	10 - 12	0.0628	0.0688	0.0659
	12 - 20	0.0667	0.0563	0.0611
pAl	2 - 3	0.128	0.121	0.125
	3 - 4	0.0951	0.0933	0.0943
	4 - 5	0.0818	0.0823	0.082
	5 - 6	0.0762	0.0774	0.0768
	6 - 7	0.0751	0.0742	0.0747
	7 - 8	0.0755	0.0759	0.0757
	8 - 9	0.0818	0.0776	0.0797
	9 - 10	0.0741	0.0701	0.0722
	10 - 12	0.0656	0.0698	0.0677
	12 - 20	0.066	0.0555	0.0607

Table 3.6: Background fractions in the η invariant mass signal region in pA collisions for the west arm, east arm, and both combined.

pA	p_T [GeV]	r (West Arm)	r (East Arm)	r (Both Arms)
pAu	2 - 3	0.794	0.79	0.793
	3 - 4	0.697	0.689	0.694
	4 - 5	0.61	0.614	0.612
	5 - 6	0.541	0.546	0.543
	6 - 7	0.507	0.511	0.509
	7 - 8	0.495	0.48	0.488
	8 - 10	0.492	0.485	0.488
	10 - 20	0.462	0.504	0.483
pAl	2 - 3	0.785	0.755	0.771
	3 - 4	0.674	0.654	0.665
	4 - 5	0.598	0.588	0.594
	5 - 6	0.543	0.541	0.542
	6 - 7	0.517	0.492	0.506
	7 - 8	0.483	0.498	0.49
	8 - 10	0.479	0.494	0.486
	10 - 20	0.432	0.426	0.429

3.3 Transverse Single-Spin Asymmetries

Transverse single-spin asymmetry measurements presented in this dissertation come from initial state collision systems $p^{\uparrow,\downarrow}$(p, Au, Al). The observable is introduced in Section 1.4 (see Eq. 1.36), where ϕ in this case is the azimuthal angle as defined in the PHENIX coordinate system. For midrapidity PHENIX measurements, there is limited detector coverage in ϕ (as discussed in Section 3.3.2.2) and mea-

surements are integrated over the azimuthal angle for each spectrometer arm. The average value of $|cos\phi|$ for all detected candidates is therefore applied as a correction factor. The relevant formulae for A_N and σ_{A_N} in this analysis are then,

$$A_N = \frac{1}{\langle|cos\phi|\rangle}\frac{1}{P}A_N^{raw} \tag{3.10}$$

$$\sigma_{A_N} = |A_N|\sqrt{\left(\frac{\sigma_{A_N^{raw}}}{A_N^{raw}}\right)^2 + \left(\frac{\sigma_P}{P}\right)^2} \tag{3.11}$$

where A_N^{raw} is calculated in terms of detected particle candidates that are sorted in terms of spin direction and spectrometer arm. Two separate methods are used to calculate A_N^{raw} — *the relative luminosity formula*, and *the square root formula*. The results from the relative luminosity formula are reported in the end, with the differences obtained using the two formulae considered for systematic uncertainties.

3.3.1 Raw Asymmetries

The relative luminosity formula is an exact formula for the transverse single-spin asymmetry, in agreement with Equation 1.36. Due to A_N being a ratio measurement, and further that the relative luminosity formula considers candidates measured in one spectrometer arm at a time — effects from acceptance and detector efficiency cancel out. However, the relative luminosity of different spin configurations must be measured, as discussed in Section 3.3.1.1. Equations 3.12 and 3.13 are used to calculate the raw asymmetry and statistical uncertainty using the relative luminosity formula:

$$A_N^{raw} = \frac{N_L^{\uparrow} - RN_L^{\downarrow}}{N_L^{\uparrow} + RN_L^{\downarrow}} \tag{3.12}$$

$$\sigma_{A_N^{raw}} = \frac{2RN_L^{\uparrow}N_L^{\downarrow}}{\left(N_L^{\uparrow} + RN_L^{\downarrow}\right)^2}\sqrt{\frac{1}{N_L^{\uparrow}} + \frac{1}{N_L^{\downarrow}}}. \tag{3.13}$$

Here $N^{\uparrow}$ and $N^{\downarrow}$ refer to particle candidate yields corresponding to initial states $p^{\uparrow}p$ and $p^{\downarrow}p$ respectively, and N_L and N_R refer to those measured in the detector to the left or right of the polarized proton going direction. The relative luminosity R is measured from data and discussed in Section 3.3.1.1. While Equation 3.12 is exclusively for the left side counts, there is an equivalent expression for the right side counts, but the signs in the numerator are flipped to preserve the "left-right" asymmetry convention. The statistical uncertainty of the relative luminosity is negligible compared to the statistical uncertainties on the yields.

The square root formula is not an exact expression for the asymmetry, but rather a geometric mean. Due to A_N being a ratio measurement, effects from both detector acceptance and the relative luminosity cancel out to first order. Equations 3.14 and 3.15 are used to calculate the raw asymmetry and statistical uncertainty using the square root formula,

$$A_N^{raw} = \frac{\sqrt{N_L^{\uparrow}N_R^{\downarrow}} - \sqrt{N_L^{\downarrow}N_R^{\uparrow}}}{\sqrt{N_L^{\uparrow}N_R^{\downarrow}} + \sqrt{N_L^{\downarrow}N_R^{\uparrow}}} \tag{3.14}$$

$$\sigma_{A_N^{raw}} = \frac{\sqrt{N_L^{\uparrow}N_R^{\downarrow}N_L^{\downarrow}N_R^{\uparrow}}}{\left(\sqrt{N_L^{\uparrow}N_R^{\downarrow}} + \sqrt{N_L^{\downarrow}N_R^{\uparrow}}\right)^2}\sqrt{\frac{1}{N_L^{\uparrow}} + \frac{1}{N_L^{\downarrow}} + \frac{1}{N_R^{\uparrow}} + \frac{1}{N_R^{\downarrow}}}. \tag{3.15}$$

3.3.1.1 Relative Luminosity

The relative luminosity R in Equation 3.12 is defined as the ratio of luminosity for $p^{\uparrow}p$ collisions to that of $p^{\downarrow}p$ collisions: $R = L^{\uparrow}/L^{\downarrow}$. R is calculated by summing up the number of times each bunch crossing fires a particular Global Level-1 (GL1) trigger throughout a run (in this case the BBC Local Level-1 trigger with a 30 cm z vertex cut) while keeping track of the spin direction. These trigger counts are referred to as GL1P scalers, and R is calculated in practice by taking the ratio of GL1P scalers

summed over all of the runs in a particular fill for opposing spin orientations. The relative luminosity was calculated separately for each beam, as shown in Figure 3.38 for the pp dataset, Figure 3.39 for the pAu dataset, and Figure 3.40 for the pAl dataset. It should be noted that accelerator fills are combined into groups of two due to limited statistics, so the relative luminosity used in calculating the asymmetry for each fill group is the weighted average of R for each fill in that group.

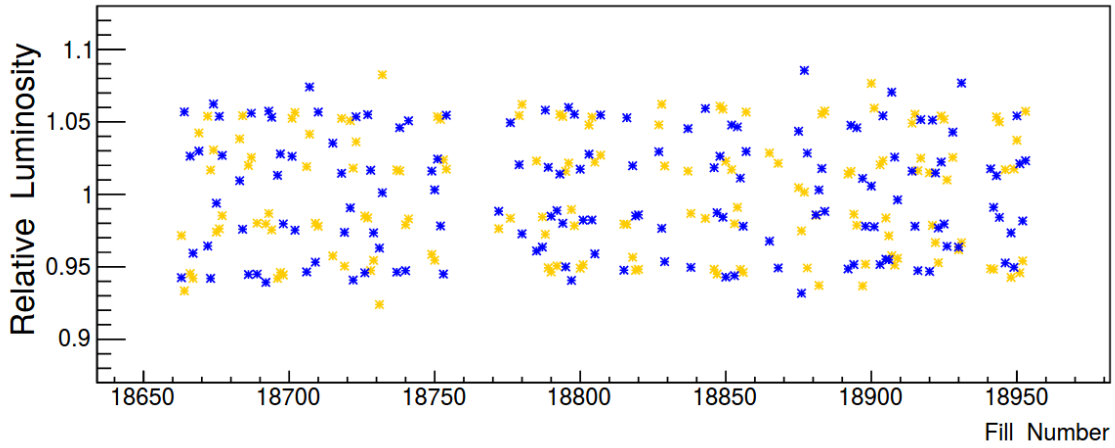

Figure 3.38: The relative luminosity calculated for each fill in pp collisions. The blue and yellow points show R for the blue and yellow beams respectively.

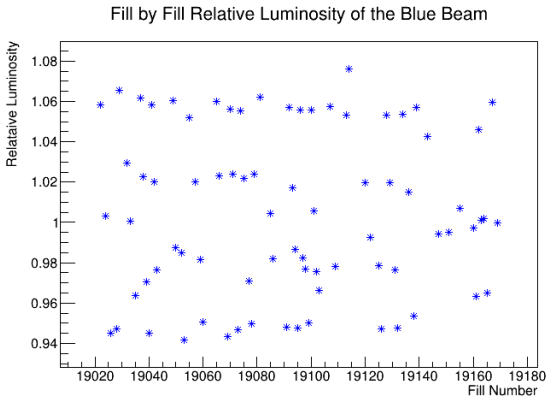

Figure 3.39: The relative luminosity calculated for each fill in pAu collisions.

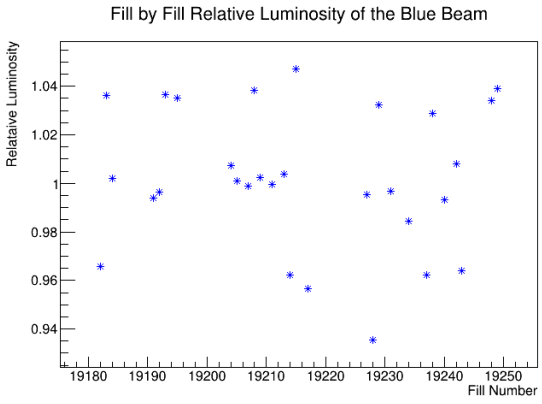

Figure 3.40: The relative luminosity calculated for each fill in pAl collisions.

3.3.2 Corrections to Raw Asymmetries

Equation 3.10 shows that the raw asymmetry must be corrected for the beam polarization fraction and the azimuthal detector acceptance in order to obtain a measured asymmetry for the signal candidates.

3.3.2.1 Polarization

The average polarization fraction of the proton beam must be taken into account in extracting A_N from data. As shown in Equation 3.10, a polarization fraction P below unity will dilute the asymmetry and must be corrected for. The CNI polarimetry group recorded the average beam polarization for each fill (documented in Ref. [143]). These values are used in this analysis and plotted in Fig. 3.41 for the pp data, and Figs. 3.42 and 3.43 for the pA data. Details on how the proton beam polarization is maintained and measured during acceleration and storage can be found in Section 2.1.1. Data from two successive fills are combined for analysis, where the polarization value for each of these fill groups is calculated as the average polarization of the two fills weighted by each fill's luminosity.[1]

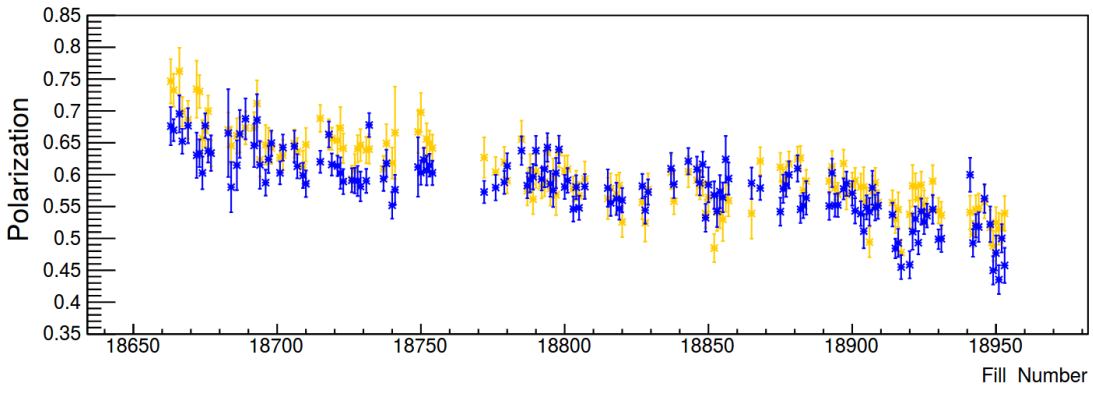

Figure 3.41: The average beam polarization of each fill in pp collisions. The blue and yellow points show P for the blue and yellow beams respectively.

[1]Fill luminosity is determined by summing the GL1P scalers for all runs in each fill.

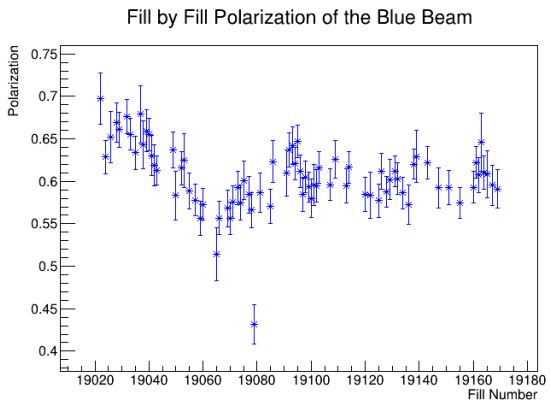

Figure 3.42: The average blue beam polarization of each fill in pAu collisions.

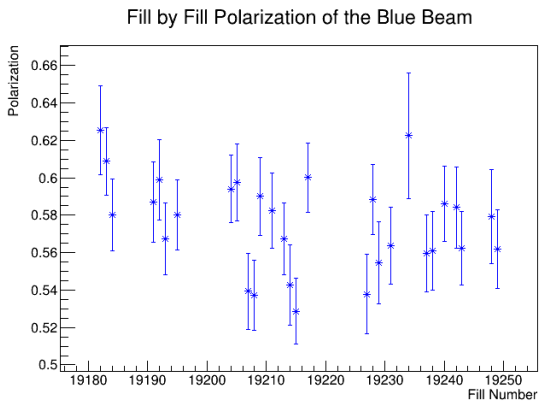

Figure 3.43: The average blue beam polarization of each fill in pAl collisions.

3.3.2.2 Azimuthal Acceptance

Transverse single-spin asymmetries are typically measured as a function of azimuthal angle ϕ. However, this method is not practical for midrapidity measurements at PHENIX, as the central arm spectrometer only covers $2 \times \pi/2$ in azimuth, and midrapidity TSSA measurements tend to be consistent with zero. For these reasons, this method is instead used as a cross check, as shown in Section 3.3.3.4. Instead, an azimuthal correction factor $\langle|cos\phi|\rangle$ is applied to account for the integration over the ϕ range of each spectrometer arm when calculating A_N. This correction factor is calculated as shown in Equation 3.16, the results of which can be seen in Table 3.7 for the electron candidates. The values labeled E and W correspond to the east and west spectrometer arms and are used in correcting A_N^{raw} calculated with the relative luminosity formula. The values labeled B are obtained by averaging over both arms and used to correct A_N^{raw} calculated with the square root formula. Figure 3.44 shows the ϕ distribution for the electron candidate sample.

$$\langle|cos\phi|\rangle = \frac{\int |cos\phi| d\phi}{\int d\phi} \approx \frac{\sum_{i=1}^{N} |cos\phi_i|}{N} \tag{3.16}$$

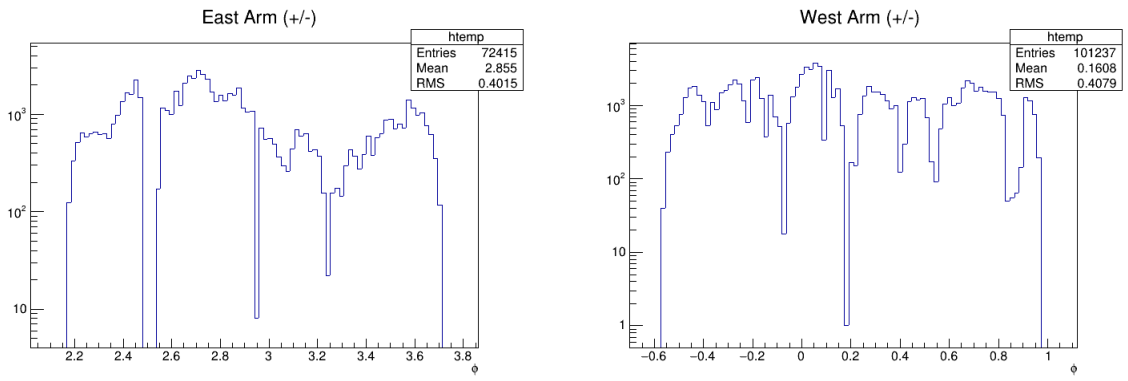

Figure 3.44: Electron candidate ϕ distribution in both spectrometer arms for 1.0 GeV $< p_T <$ 5.0 GeV.

Table 3.7: Azimuthal correction factors applied for different central arm spectrometers for the heavy flavor electron analysis in pp collisions. Note that the average values are typically closer to that of the west arm value, as there are more statistics in the west arm (seen in Figure 3.44).

p_T range (GeV)	$\langle p_T \rangle$ (GeV)	$\langle \lvert cos\phi \rvert \rangle$ (E)	$\langle \lvert cos\phi \rvert \rangle$ (W)	$\langle \lvert cos\phi \rvert \rangle$ (Both)
1.0 – 1.3	1.161	0.849	0.911	0.885
1.3 – 1.5	1.398	0.873	0.909	0.895
1.5 – 1.8	1.639	0.890	0.905	0.899
1.8 – 2.1	1.936	0.900	0.903	0.902
2.1 – 2.7	2.349	0.910	0.906	0.908
2.7 – 5.0	3.290	0.919	0.906	0.912

Figures 3.45 and 3.46 depict the ϕ distributions for the π^0 candidates in the $7 < p_T < 8$ GeV p_T bin from the pAu dataset in the west and east spectrometer arms, respectively. Figures 3.47 and 3.48 show this for the η candidate sample. The distributions in the pAl dataset are similar. The measured values of $\langle \lvert cos\phi \rvert \rangle$ are listed in Table 3.8 for the π^0 candidates in pAu and pAl collisions, and Table 3.9 for the η

candidates in pAu and pAl collisions. It should be noted that the latter two sectors in the east arm calorimeter (corresponding to the rightmost 2 peaks in the ϕ distribution for the east arms) are the PbGl sectors, which have a different efficiency for detecting photons (see Section 2.2.2.2).

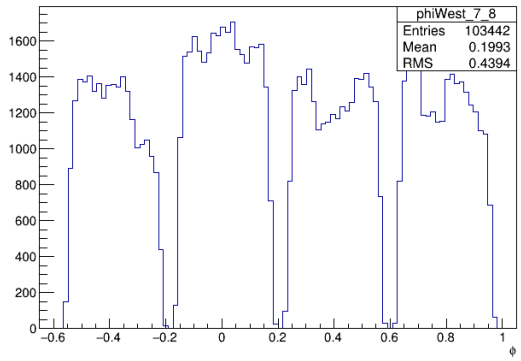

Figure 3.45: π^0 candidate ϕ distribution (7 GeV $< p_T <$ 8 GeV, pAu, West arm).

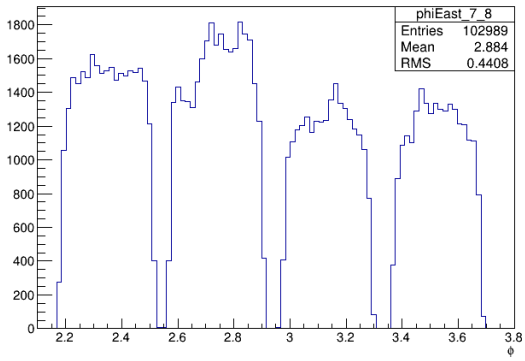

Figure 3.46: π^0 candidate ϕ distribution (7 GeV $< p_T <$ 8 GeV, pAu, East arm).

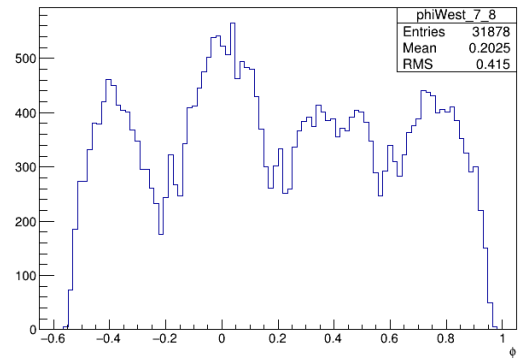

Figure 3.47: η candidate ϕ distribution (7 GeV $< p_T <$ 8 GeV, pAu, West arm).

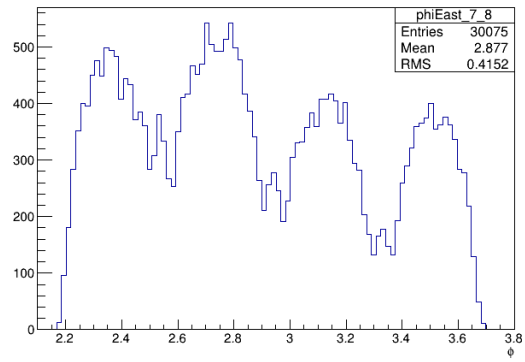

Figure 3.48: η candidate ϕ distribution (7 GeV $< p_T <$ 8 GeV, pAu, East arm.)

Table 3.8: Azimuthal correction factors $\langle|cos\phi|\rangle$ for the $A_N(p^{\uparrow}+\mathrm{A} \to \pi^0+X)$ analysis.

pA	p_T range (GeV)	$\langle\|cos\phi\|\rangle$ (W)	$\langle\|cos\phi\|\rangle$ (E)	$\langle\|cos\phi\|\rangle$ (Both)
pAu	2 – 3	0.909	0.769	0.844
	3 – 4	0.914	0.812	0.869
	4 – 5	0.899	0.838	0.872
	5 – 6	0.893	0.858	0.877
	6 – 7	0.890	0.870	0.880
	7 – 8	0.888	0.875	0.882
	8 – 9	0.889	0.878	0.883
	9 – 10	0.887	0.879	0.883
	10 – 12	0.888	0.881	0.885
	12 – 20	0.891	0.891	0.891
pAl	2 – 3	0.897	0.772	0.831
	3 – 4	0.897	0.801	0.853
	4 – 5	0.893	0.832	0.866
	5 – 6	0.889	0.851	0.872
	7 – 8	0.884	0.867	0.876
	8 – 9	0.884	0.871	0.878
	9 – 10	0.884	0.874	0.879
	10 – 12	0.885	0.876	0.881
	12 – 20	0.886	0.886	0.886

Table 3.9: Azimuthal correction factors $\langle|cos\phi|\rangle$ for the $A_N(p^{\uparrow}+\mathrm{A} \to \eta + X)$ analysis.

pA	p_T range (GeV)	$\langle\|cos\phi\|\rangle$ (W)	$\langle\|cos\phi\|\rangle$ (E)	$\langle\|cos\phi\|\rangle$ (Both)
pAu	2 – 3	0.926	0.811	0.879
	3 – 4	0.922	0.841	0.889
	4 – 5	0.912	0.865	0.891
	5 – 6	0.905	0.876	0.891
	6 – 7	0.900	0.882	0.891
	7 – 8	0.897	0.883	0.891
	8 – 10	0.894	0.885	0.890
	10 – 20	0.893	0.886	0.889
pAl	2 – 3	0.921	0.814	0.872
	3 – 4	0.917	0.839	0.884
	4 – 5	0.908	0.861	0.888
	5 – 6	0.900	0.870	0.887
	6 – 7	0.896	0.876	0.887
	7 – 8	0.893	0.877	0.886
	8 – 10	0.890	0.878	0.884
	10 – 20	0.886	0.882	0.884

3.3.3 Results Before Background Correction

With signal candidate yields sorted by spin and detector arm $N_{L,R}^{\uparrow,\downarrow}$, and average beam polarization P, relative luminosity R, and azimuthal correction factor $\langle|\cos\phi|\rangle$ all measured, one can apply Equation 3.10 to calculate the TSSA before background correction. It is useful to calculate the asymmetries as a function of fill number to check for possible time-dependent systematic effects. Poissonian statistics are as-

sumed for the statistical uncertainties on measured yields, so $N_{L,R}^{\uparrow,\downarrow}$ is required to be at least 10. This makes it practical to combine fills into groups of two chronologically, to allow for finer binning in p_T than when calculating the asymmetry for each fill individually. Figure 3.49 shows an example asymmetry plot as a function of fill group index from the heavy flavor electron analysis for electron and positron candidates within the p_T range $1.8 < p_T < 2.1$ GeV, in particular using the relative luminosity formula in the right spectrometer arm relative to the blue beam going direction, and only considering the polarization of the blue beam. The plot is fit to a constant in order to extract the average value of the asymmetry over all fill groups, weighted by the statistical error. This is how the asymmetry is calculated in each p_T bin for the square root and relative luminosity formulae, and for both spectrometer arms (left and right relative to the polarized beam going direction) when using the relative luminosity formula. For the pp dataset, both proton beams are transversely polarized with polarization directions controlled independently and changing from bunch to bunch. This means that the dataset can be split into two statistically independent subsets by keeping track of the polarization directions for only one beam at a time while effectively averaging over the polarization directions of the other beam. We denote these asymmetries as the yellow beam asymmetry and blue beam asymmetry depending on the beam polarization that we are keeping track of. The weighted average of the left and right relative luminosity asymmetries are taken as the asymmetry for each beam, and the weighted average of the asymmetry of each beam is taken as the final candidate asymmetry. In the pA datasets, only the proton (blue) beam is polarized, therefore the blue beam relative luminosity asymmetry is taken as the final candidate asymmetry.

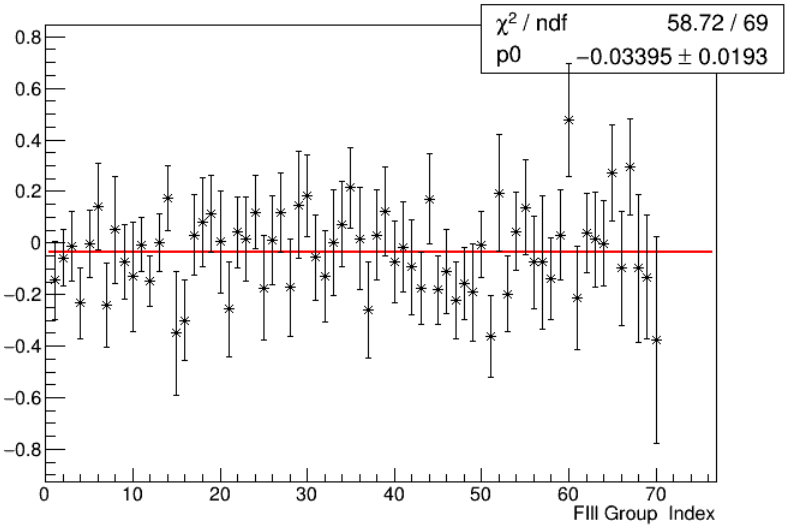

Figure 3.49: Relative luminosity asymmetry of the blue beam in the right spectrometer arm with respect to the blue beam going direction. The asymmetry is plotted as a function of fill group index, and in the p_T range from 1.8 GeV $< p_T <$ 2.1 GeV for electron and positron candidates.

3.3.3.1 Relative Luminosity Formula

Heavy Flavor Electrons: The two independent relative luminosity asymmetry formulae for each side of the detector allow for further dividing of the two statistically independent data samples corresponding to yellow and blue beam asymmetries into two more statistically independent data samples corresponding to each side of the detector, giving a total of four statically independent measurements to compare. Figures 3.50 and 3.52 show the left and right asymmetry results for both the yellow and blue beams respectively, calculated with Equations 3.10 and 3.12, while Figures 3.51 and 3.53 show the corresponding T tests evaluating the statistical significance of the differences in left and right asymmetries, calculated using Equation 3.17.

$$T(p_T) = \frac{A_N^{Left} - A_N^{Right}}{\sqrt{\sigma^2_{A_N^{Left}} + \sigma^2_{A_N^{Right}}}} \tag{3.17}$$

The uncertainties in the denominator in Eq. 3.17 are added in quadrature due to the left and right asymmetry measurements being calculated on completely uncorrelated datasets. We expect statistically consistent results to have T values to fall as if they were sampled from a T distribution. With only 6 p_T bins, this is difficult to see visually, but the T test output can still be used to verify that the results are statistically consistent. Figure 3.54 shows the weighted average of the left and right asymmetry results for the different beams, and Figure 3.55 shows the corresponding T values calculated using Equation 3.18.

$$T(p_T) = \frac{A_N^{Yellow} - A_N^{Blue}}{\sqrt{\sigma^2_{A_N^{Yellow}} + \sigma^2_{A_N^{Blue}}}} \tag{3.18}$$

In Fig. 3.54, the yellow points correspond to the yellow beam asymmetries and the blue points correspond to the blue beam asymmetries.

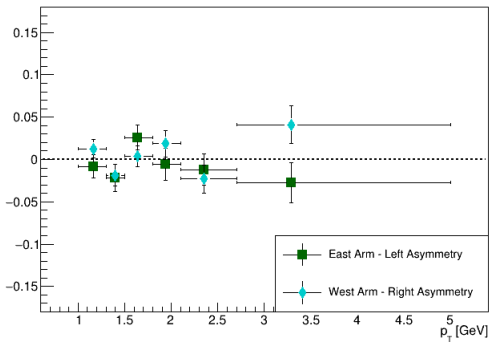

Figure 3.50: Yellow beam left and right $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity formula.

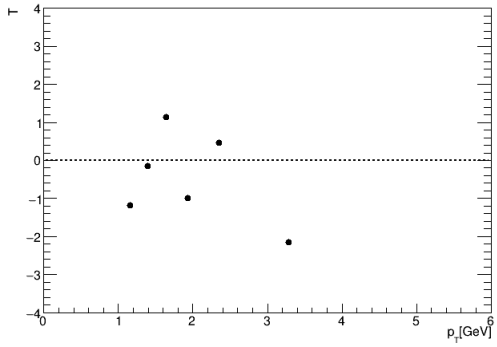

Figure 3.51: T test corresponding to the Fig. 3.50.

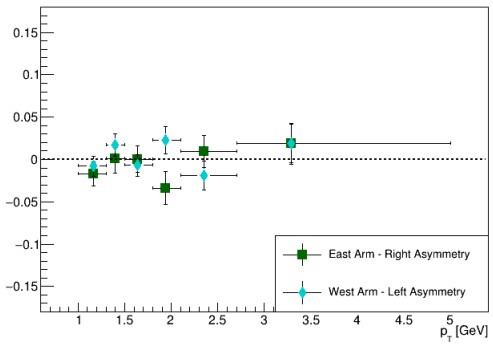

Figure 3.52: Blue beam left and right $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity formula.

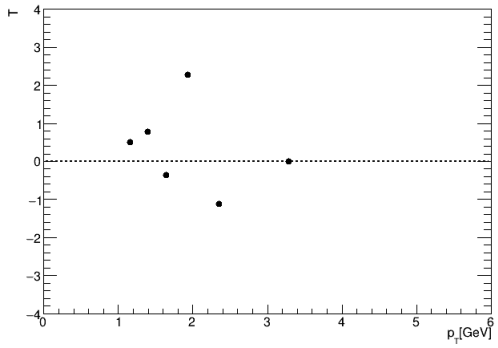

Figure 3.53: T test corresponding to the Fig. 3.52.

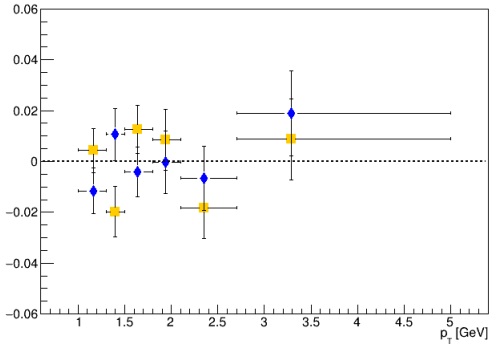

Figure 3.54: Yellow and blue beam $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity formula.

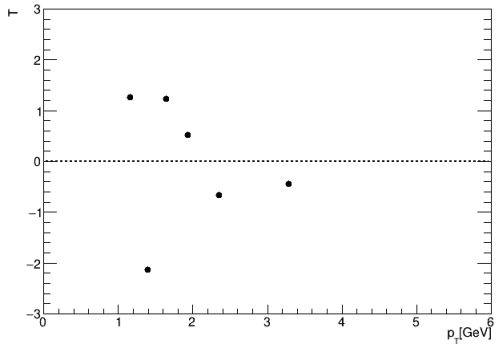

Figure 3.55: T test corresponding to Fig. 3.54.

Neutral Mesons: Only the protons are polarized in pA collisions, and they occupy the blue RHIC ring. Figures 3.56, 3.58, 3.60, and 3.62 show the left and right blue beam asymmetry results for π^0 in pAu and pAl collisions, and η in pAu and pAl collisions, respectively, calculated using Equations 3.10 and 3.12. Figures 3.57, 3.59,

3.61, and 3.63 show the corresponding T tests evaluating the statistical significance of the differences in left and right asymmetries, calculated using Equation 3.17.

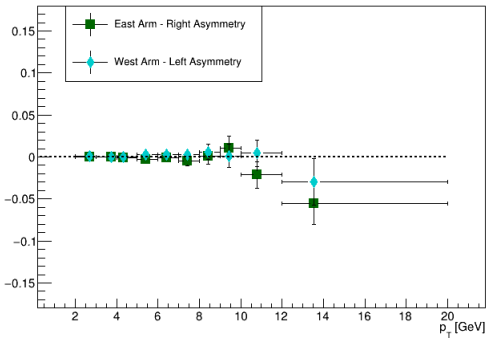

Figure 3.56: Blue beam left and right $A_N(p^{\uparrow}\text{Au} \rightarrow \pi^0 X)$ before background correction measured using the relative luminosity formula.

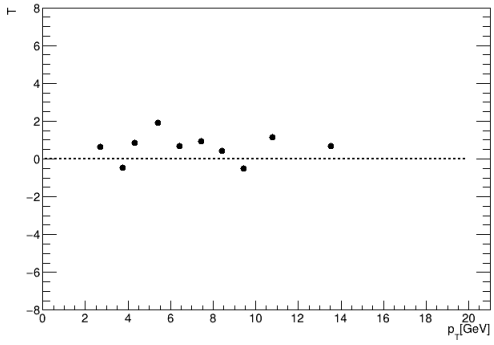

Figure 3.57: T test corresponding to Fig. 3.56.

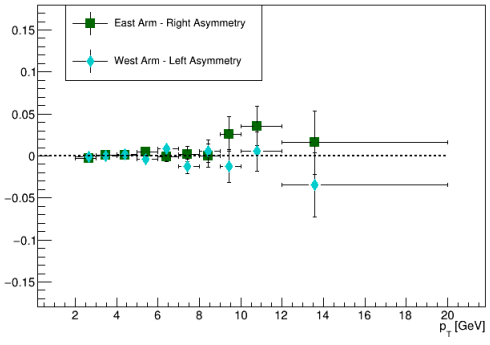

Figure 3.58: Blue beam left and right $A_N(p^{\uparrow}\text{Al} \rightarrow \pi^0 X)$ before background correction measured using the relative luminosity formula.

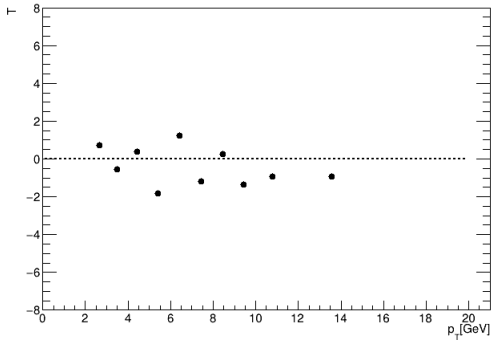

Figure 3.59: T test corresponding to Fig. 3.58.

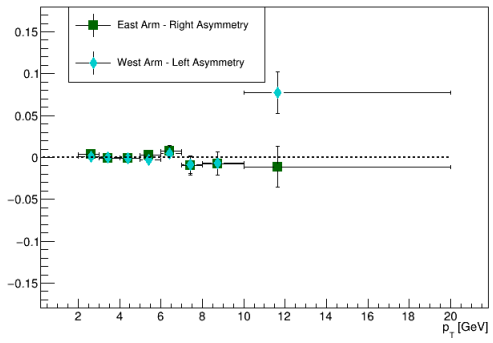

Figure 3.60: Blue beam left and right $A_N(p^{\uparrow}\text{Au} \to \eta X)$ before background correction measured using the relative luminosity formula.

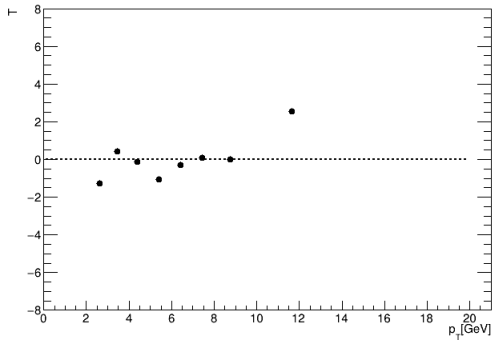

Figure 3.61: T test corresponding to Fig. 3.60.

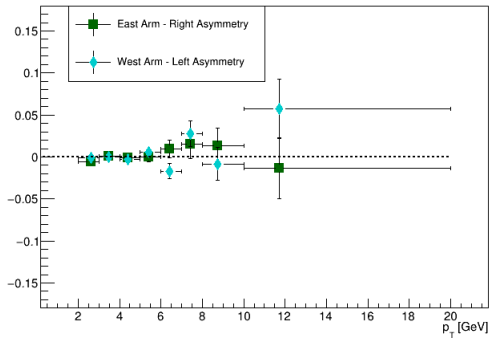

Figure 3.62: Blue beam left and right $A_N(p^{\uparrow}\text{Al} \to \eta X)$ before background correction measured using the relative luminosity formula.

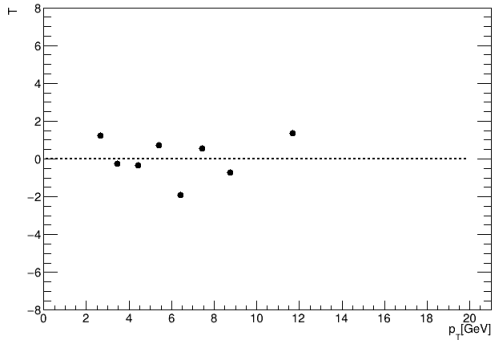

Figure 3.63: T test corresponding to Fig. 3.62.

3.3.3.2 Square Root Formula

Heavy Flavor Electrons: The results for A_N before background correction obtained using the square root formula (Equations 3.10 and 3.14) are shown in Figure 3.64, while the corresponding T values calculated using Equation 3.18 to demon-

strate statistical consistency between the results are shown in Figure 3.65. Note that for the asymmetry plots, the yellow points correspond to the yellow beam asymmetries while the blue points correspond to the blue beam asymmetries.

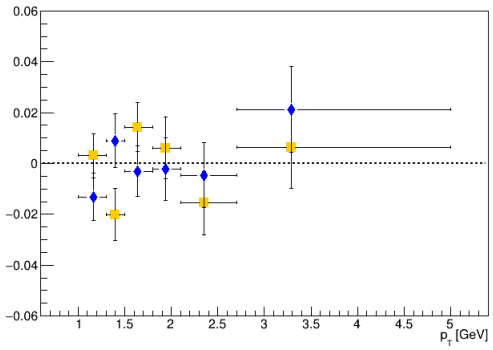

Figure 3.64: Yellow and blue beam $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the square root formula.

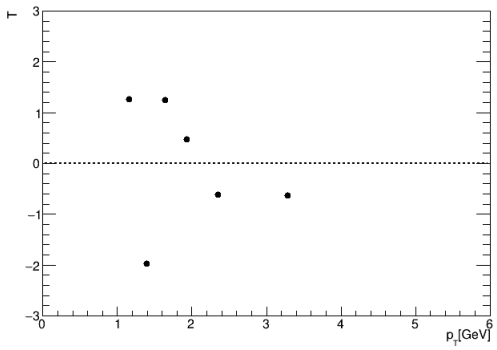

Figure 3.65: T test corresponding to Fig. 3.64.

Neutral Mesons: Similarly, the results for A_N before background correction obtained using the square root formula (Equations 3.10 and 3.14) are shown in Figures 3.66, 3.67, 3.68, and 3.69 for π^0 in pAu and pAl collisions, and η in pAu and pAl collisions, respectively.

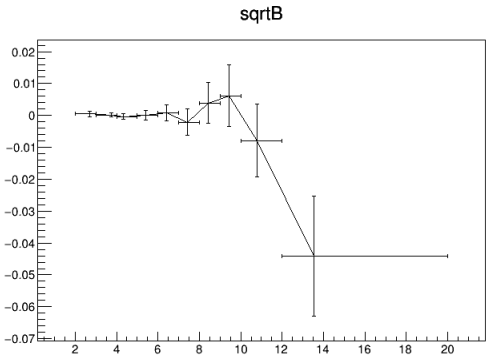

Figure 3.66: Blue beam $A_N(p^{\uparrow}\text{Au} \rightarrow \pi^0 X)$ before background correction measured using the square root formula.

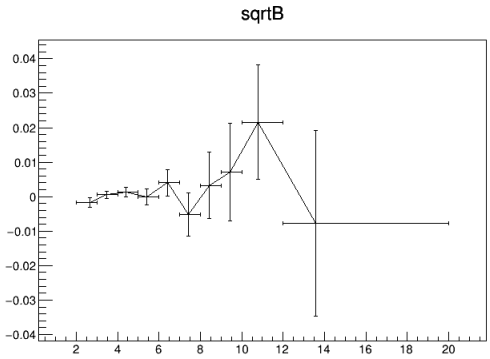

Figure 3.67: Blue beam $A_N(p^{\uparrow}\text{Al} \rightarrow \pi^0 X)$ before background correction measured using the square root formula.

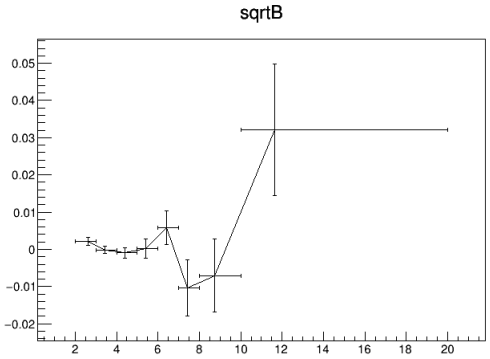

Figure 3.68: Blue beam $A_N(p^{\uparrow}\text{Au} \rightarrow \eta X)$ before background correction measured using the square root formula.

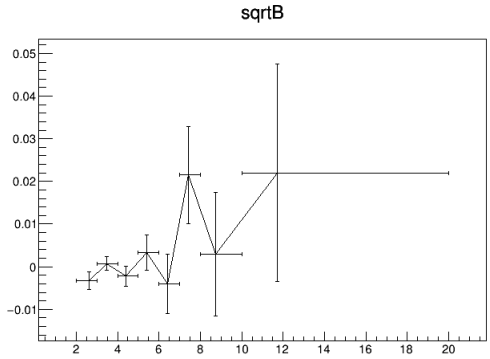

Figure 3.69: Blue beam $A_N(p^{\uparrow}\text{Al} \rightarrow \eta X)$ before background correction measured using the square root formula.

3.3.3.3 Asymmetry Formula Comparison

Heavy Flavor Electrons: The results using the two independent formulae for calculating A_N are compared and plotted together for the yellow beam asymmetries in Figure 3.70, the blue beam asymmetries in Figure 3.72, and beam averaged asymmetries in Figure 3.74. Equation 3.19 shows the T test procedure for comparing different calculations of mean values for the same quantity on the same dataset — the minus

sign in the denominator comes from the fact that the datasets are 100% correlated. Figures 3.71, 3.73 and, 3.75 show the T test results obtained using Equation 3.19 for the yellow beam, blue beam, and beam averaged asymmetries respectively.

$$T(p_T) = \frac{A_N^{Sqrt} - A_N^{Lumi}}{\sqrt{|\sigma^2_{A_N^{Sqrt}} - \sigma^2_{A_N^{Lumi}}|}} \tag{3.19}$$

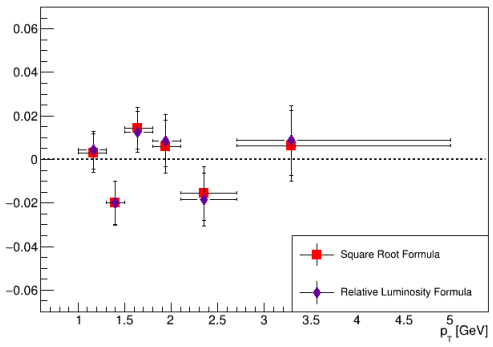

Figure 3.70: Comparison of yellow beam $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity and square root formulae.

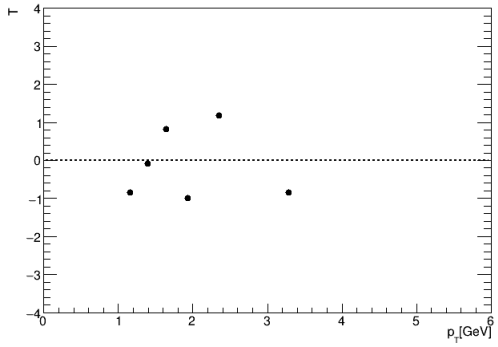

Figure 3.71: T test corresponding to Fig. 3.70.

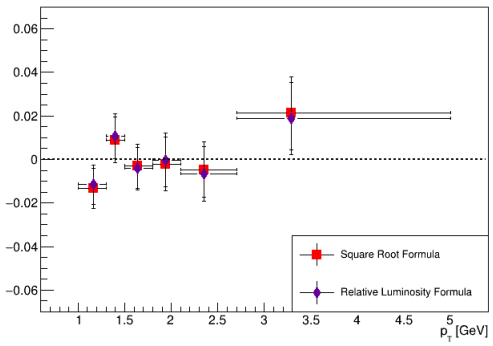

Figure 3.72: Comparison of blue beam $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity and square root formulae.

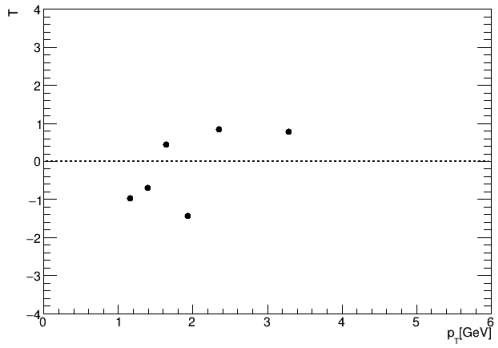

Figure 3.73: T test corresponding to Fig. 3.72.

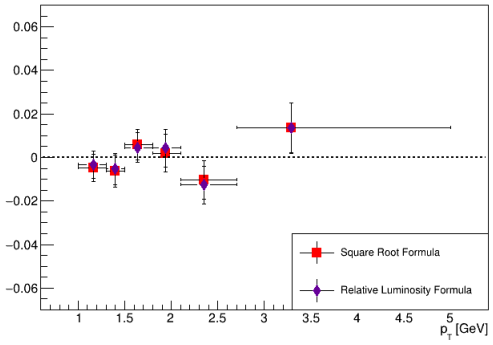

Figure 3.74: Comparison of beam averaged $A_N(p^{\uparrow}p \to e^{\pm}X)$ before background correction measured using the relative luminosity and square root formulae.

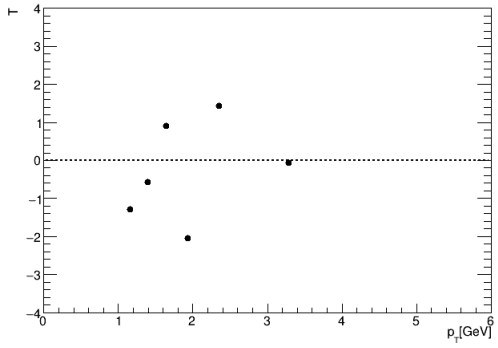

Figure 3.75: T test corresponding to Fig. 3.74.

Neutral Mesons: The results using the two independent formulae for calculating A_N are compared and plotted together in Figures 3.76, 3.78, 3.80, and 3.82 for π^0 in pAu and pAl collisions, and η in pAu and pAl collisions, respectively. Figures 3.77, 3.79, 3.81, and 3.83 show the corresponding T-test results obtained using

Equation 3.19.

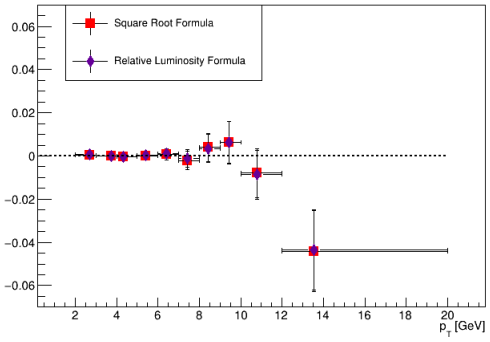

Figure 3.76: Comparison of $A_N(p^{\uparrow}\text{Au} \to \pi^0 X)$ before background correction measured using the relative luminosity and square root formulae.

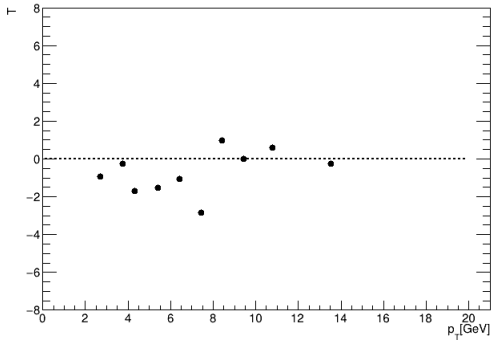

Figure 3.77: T test corresponding to Fig. 3.76.

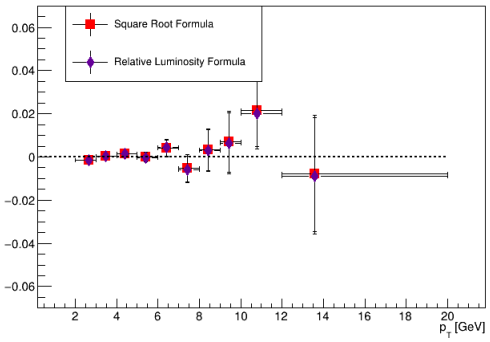

Figure 3.78: Comparison of $A_N(p^{\uparrow}\text{Al} \to \pi^0 X)$ before background correction measured using the relative luminosity and square root formulae.

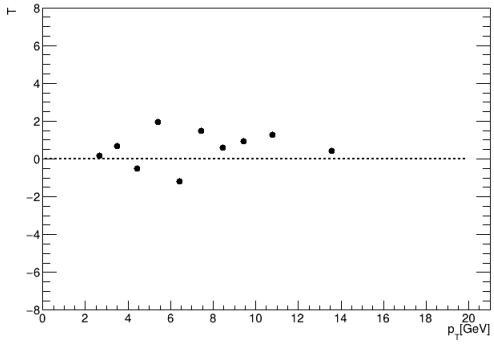

Figure 3.79: T test corresponding to Fig. 3.78.

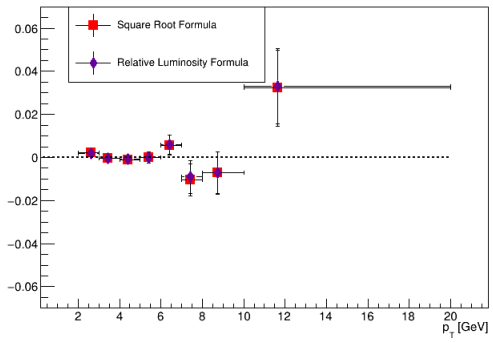

Figure 3.80: Comparison of $A_N(p^{\uparrow}\text{Au} \rightarrow \eta X)$ before background correction measured using the relative luminosity and square root formulae.

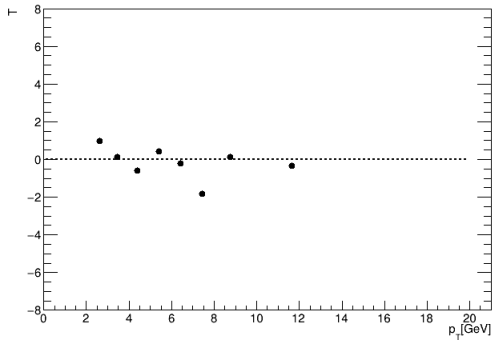

Figure 3.81: T test corresponding to Fig. 3.80.

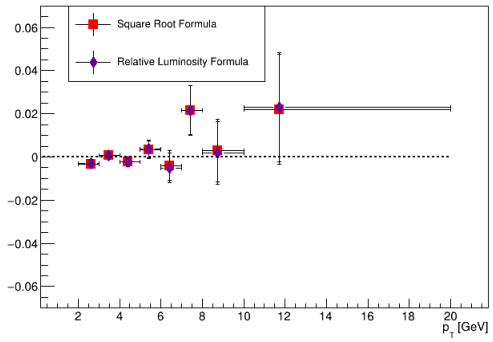

Figure 3.82: Comparison of $A_N(p^{\uparrow}\text{Al} \rightarrow \eta X)$ before background correction measured using the relative luminosity and square root formulae.

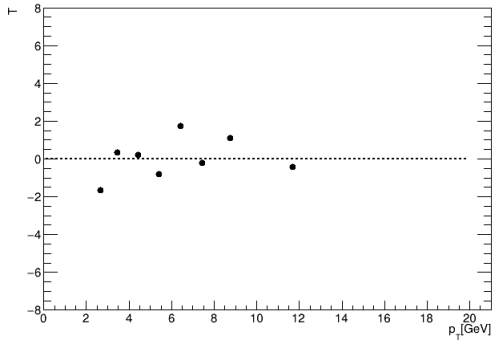

Figure 3.83: T test corresponding to Fig. 3.82.

3.3.3.4 Sinusoidal Modulation Cross Check

Equation 1.36 shows that A_N is typically extracted as the amplitude of a cosine modulation in ϕ. This is challenging to do for midrapidity PHENIX measurements due to the limited azimuthal coverage of the central arm spectrometers, and the

tendency for midrapidity TSSA measurements to be consistent with 0. It still serves as a useful cross check to results shown in previous sections. The relative luminosity formula was used to perform this cross check for practicality due to the ϕ dependent square root formula requiring detector coverage separated by π radians for each ϕ bin, while the PHENIX central arms are slightly offset from back to back. The specific formula used for this cross check is shown in Equation 3.20,

$$A_N \sin(\phi_s) = \frac{1}{P}\epsilon_N(\phi_s) = \frac{1}{P}\frac{N^{\uparrow}(\phi_s) - RN^{\downarrow}(\phi_s)}{N^{\uparrow}(\phi_s) + RN^{\downarrow}(\phi_s)}, \tag{3.20}$$

where ϕ_s denotes the azimuthal angle from the spin-up direction (y=0 in PHENIX coordinates), which increases to the left of the polarized beam going direction. It should be noted that no azimuthal correction is needed in Equation 3.20 since the calculated asymmetry is ϕ dependent.

Heavy Flavor Electrons: Electron candidates were split into 6 bins in ϕ_s (3 in each spectrometer arm) in order to maintain enough statistics in each p_T bin while still having enough points to fit to a sinusoid in ϕ_s. In every p_T bin, Equation 3.20 is applied to calculate the asymmetry as a function of ϕ_s, and the results are fit to the function "p0*sin(x)". Figure 3.84 and 3.85 show the fit results in the lowest p_T bin for the yellow and blue beam asymmetries respectively. The amplitude parameters $p0$ are extracted for both beams and plotted in Figure 3.86. The separate beam asymmetries shown in Figure 3.86 are averaged and compared to the relative luminosity results shown in Figure 3.54 obtained from integrating over the detector coverage in ϕ and applying the azimuthal correction factors discussed in Section 3.3.2.2. Results from the comparison can be seen in Figure 3.87.

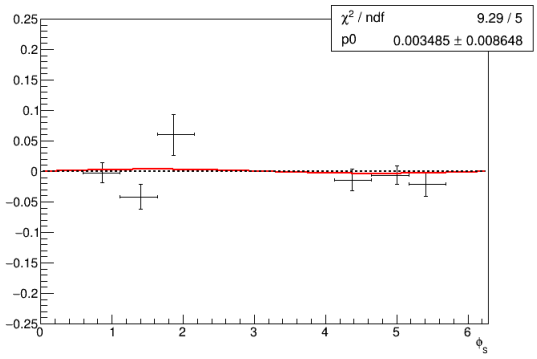

Figure 3.84: Yellow beam $A_N(p^{\uparrow}p \rightarrow e^{\pm}X)$ calculated as a function of ϕ_s within 1.0 GeV $< p_T <$ 1.3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

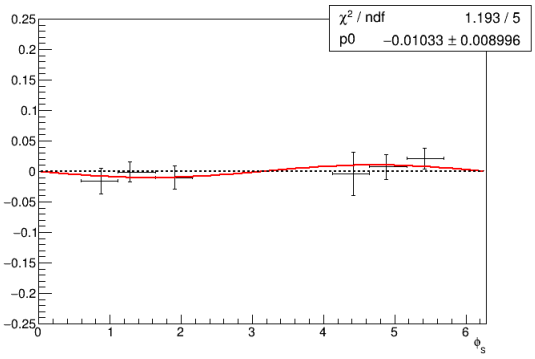

Figure 3.85: Blue beam $A_N(p^{\uparrow}p \rightarrow e^{\pm}X)$ calculated as a function of ϕ_s within 1.0 GeV $< p_T <$ 1.3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

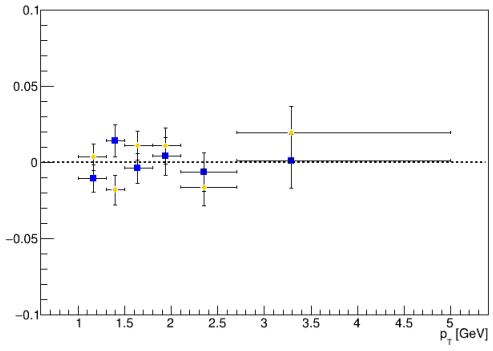

Figure 3.86: Yellow and blue beam $A_N(p^{\uparrow}p \rightarrow e^{\pm}X)$ extracted from amplitudes of sinusoidal fits in each p_T bin.

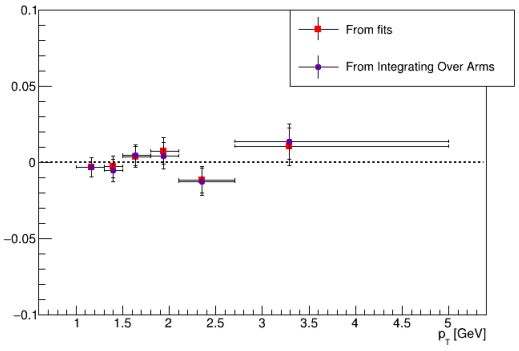

Figure 3.87: Comparison of $A_N(p^{\uparrow}p \rightarrow e^{\pm}X)$ from the weighted average of the fit amplitudes shown in Fig. 3.86, and the relative luminosity asymmetries shown in Fig. 3.74.

Neutral Mesons: Figures 3.88 and 3.89 show the fit results for π^0 candidates in the lowest p_T bin in pAu and pAl collisions, respectively. Similarly, Figures 3.92 and 3.93

show the fit results for η candidates in the lowest p_T bin in pAu and pAl collisions, respectively. The amplitude parameters $p0$ are extracted for π^0 and η candidates in pAu and pAl collisions and respectively plotted in Figures 3.90, 3.91, 3.94, and 3.95 along with the relative luminosity asymmetries calculated by integrating over ϕ.

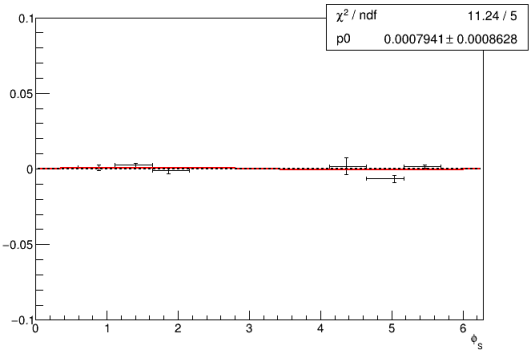

Figure 3.88: Blue beam $A_N(p^{\uparrow}\text{Au} \rightarrow \pi^0 X)$ calculated as a function of ϕ_s within 2 GeV $< p_T <$ 3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

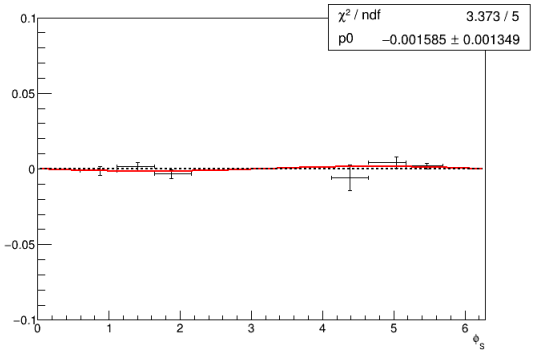

Figure 3.89: Blue beam $A_N(p^{\uparrow}\text{Al} \rightarrow \pi^0 X)$ calculated as a function of ϕ_s within 2 GeV $< p_T <$ 3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

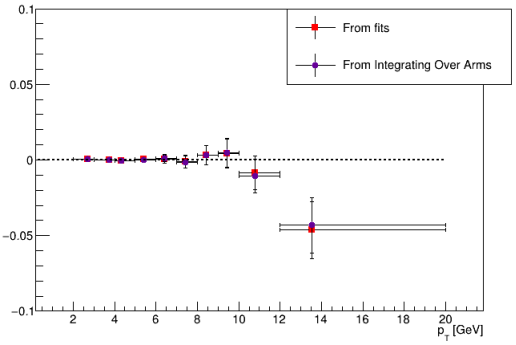

Figure 3.90: Comparison of $A_N(p^{\uparrow}\text{Au} \rightarrow \pi^0 X)$ from fit amplitudes in each p_T bin and relative luminosity results from Fig. 3.76.

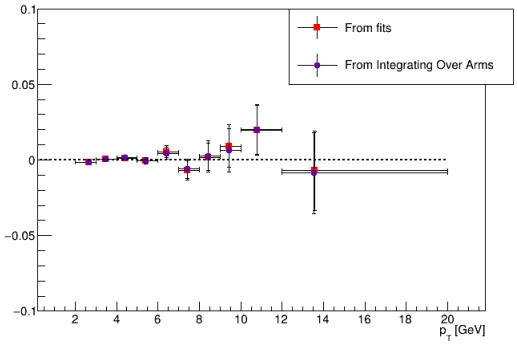

Figure 3.91: Comparison of $A_N(p^{\uparrow}\text{Al} \rightarrow \pi^0 X)$ from fit amplitudes in each p_T bin and relative luminosity results from Fig. 3.78.

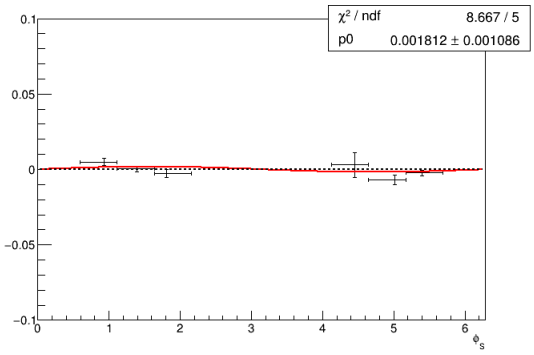

Figure 3.92: Blue beam $A_N(p^{\uparrow}\text{Au} \to \eta X)$ calculated as a function of ϕ_s within 2 GeV $< p_T <$ 3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

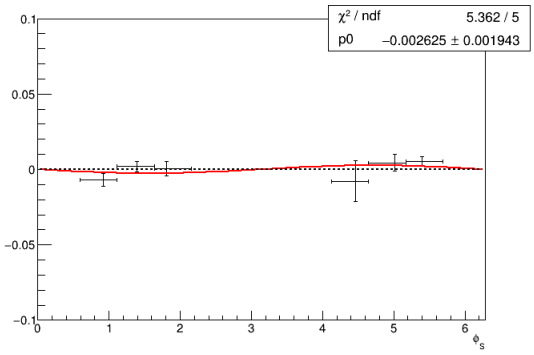

Figure 3.93: Blue beam $A_N(p^{\uparrow}\text{Al} \to \eta X)$ calculated as a function of ϕ_s within 2 GeV $< p_T <$ 3 GeV. The asymmetry is represented by the amplitude of the sinusoidal fit.

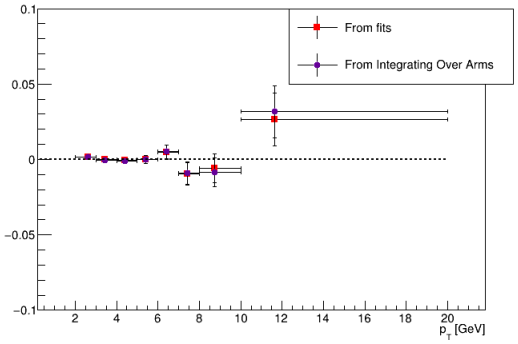

Figure 3.94: Comparison of $A_N(p^{\uparrow}\text{Au} \to \eta X)$ from fit amplitudes in each p_T bin and relative luminosity results from Fig. 3.80.

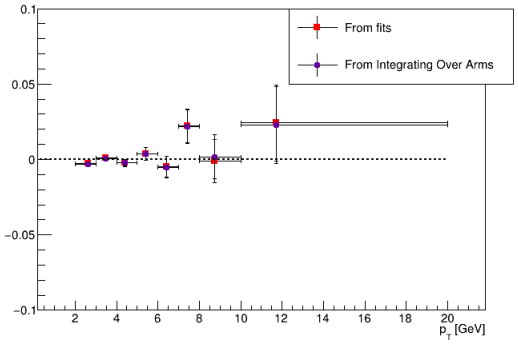

Figure 3.95: Comparison of $A_N(p^{\uparrow}\text{Al} \to \eta X)$ from fit amplitudes in each p_T bin and relative luminosity results from Fig. 3.82.

3.3.4 Background Correction

Not all candidates that pass the selection criteria are considered signal, as discussed in Sections 3.2.1.6 and 3.2.2.2. Therefore, one must correct the TSSA mea-

surement in each kinematic bin based on the measured background fractions using Equations 3.21 and 3.22,

$$A_N^S = \frac{A_N^{S+B} - \sum_i r_i A_N^{B_i}}{1 - \sum_i r_i} \tag{3.21}$$

$$\sigma_{A_N^S} = \frac{\sqrt{(\sigma_{A_N^{S+B}})^2 + \sum_i r_i^2 (\sigma_{A_N^{B_i}})^2}}{1 - \sum_i r_i} \tag{3.22}$$

where S stands for signal, B_i stands for the i^{th} background component, and $r_i = N_{B_i}/(N_S + N_{B_i}) = B_i/(S + B_i)$ is the background fraction. Quantification of background fractions r_i is discussed in detail in Section 3.2.1.6 for the heavy flavor electron measurement in pp collisions. In the case of neutral mesons where the particle is reconstructed and a resonance structure in the invariant mass spectrum can be used to verify the purity of the sample, only the fraction of combinatorial background under the signal peak r must be considered. The background fractions r are shown in Section 3.2.2.2 for the π^0 and η measurement in pA collisions.

Heavy Flavor Electrons: Plugging in the relevant background sources to Equations 3.21 and 3.22, one arrives at Equations 3.23 and 3.24 for calculating the background corrected asymmetries and statistical uncertainties for heavy flavor electron TSSAs. It was decided that the $Ke3$ contribution to the electron candidate sample was negligible and that this could be ignored in applying the background correction formulae. It should be noted that in previous analyses [123, 165], measurements for $A_N^{B_i}$ for each photonic background source are shown to be consistent with zero over a wide range in p_T. It was therefore decided to treat all photonic background sources purely as a dilution factor, as shown in Equations 3.23 and 3.24[2]. The values for $A_N^{h^\pm}$ and $\sigma_{A_N^{h^\pm}}$ are taken directly from published PHENIX results [169]. The procedure for calculating $A_N^{J/\psi \to e}$ and $\sigma_{A_N^{J/\psi \to e}}$ from $A_N^{J/\psi}$ and $\sigma_{A_N^{J/\psi}}$, which were taken from [170], is

[2]This implies that $A_N^{B_i} = 0$ and $\sigma_{A_N^{B_i}} = 0$ for all photonic background sources, $i = \pi^0, \eta, \gamma$.

outlined in Section 3.3.4.1.

$$A_N^{OHF\to e} = \frac{A_N^e - f_{h^\pm} A_N^{h^\pm} - f_{J/\psi\to e} A_N^{J/\psi\to e}}{1 - f_{h^\pm} - f_{J/\psi\to e} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}} \tag{3.23}$$

$$\sigma_{A_N^{OHF\to e}} = \frac{\sqrt{(\sigma_{A_N^e})^2 + (f_{h^\pm}\sigma_{A_N^{h^\pm}})^2 + (f_{J/\psi\to e}\sigma_{A_N^{J/\psi\to e}})^2}}{1 - f_{h^\pm} - f_{J/\psi\to e} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}} \tag{3.24}$$

Equations 3.25 and 3.26 are used to calculate the background corrected asymmetry for nonphotonic electrons, excluding the J/ψ from the background correction procedure, motivated by the large statistical uncertainties of the midrapidity $A_N^{J/\psi}$ values measured in [170]. Both open heavy flavor and nonphotonic electron TSSAs were reported in Ref. [1] (corresponding to work in this dissertation).

$$A_N^{NPe} = \frac{A_N^e - f_{h^\pm} A_N^{h^\pm}}{1 - f_{h^\pm} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}} \tag{3.25}$$

$$\sigma_{A_N^{NPe}} = \frac{\sqrt{(\sigma_{A_N^e})^2 + (f_{h^\pm}\sigma_{A_N^{h^\pm}})^2}}{1 - f_{h^\pm} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}} \tag{3.26}$$

Neutral Mesons: Photon pairs that have an invariant mass that matches the π^0 or η meson did not necessarily come from a $\pi^0 \to 2\gamma$ or $\eta \to 2\gamma$ decay. Even with all of the selection criteria in place to reduce the combinatorial background, there is no way to completely eliminate it. The TSSA needs to be corrected for this, as shown in Equations 3.27 and 3.28:

$$A_N^S = \frac{A_N^{S+B} - r A_N^B}{1 - r} \tag{3.27}$$

$$\sigma_{A_N^S} = \frac{\sqrt{(\sigma_{A_N^{S+B}})^2 + r^2(\sigma_{A_N^B})^2}}{1 - r}, \tag{3.28}$$

where S stands for signal, B stands for background, and $r = N_B/(N_S + N_B) = B/(S + B)$. This formula is applied to each asymmetry calculation individually, e.g. the square root result, and the left and right relative luminosity results. This is

because it was possible to calculate the background fractions in each spectrometer arm for the π^0 and η measurements, which was not the case for the heavy flavor electron measurement.

3.3.4.1 Background Asymmetries

The background asymmetries $A_N^{B_i}$ from Eq. 3.21 are obtained differently for the open heavy flavor electron and neutral meson measurements. In the latter, the resonance structure in the diphoton invariant mass spectrum provides a clean separation between signal and background regions, and the background asymmetry from Eq. 3.27 can be obtained from an analysis of the side bands (see Fig. 3.37). In the case of the heavy flavor electrons, there is no resonance structure, and the background asymmetries listed in Eq. 3.23 are obtained from previous measurements and mapped to decay lepton p_T as needed.

Heavy Flavor Electrons: Background asymmetries for charged hadrons and electrons from J/ψ decays must be known to apply Eq. 3.23. The midrapidity charged hadron A_N measurement in $\sqrt{s} = 200$ GeV from Ref. [169] could be input directly into the background correction procedure since charged hadrons misidentified as electrons are a relevant background source. On the other hand, the previously published $A_N^{J/\psi}$ measurement from PHENIX [170] had to be mapped to a function of decay electron p_T. $A_N^{J/\psi}$ was measured in 200 GeV collisions at PHENIX during data taken in 2006 [170], and while the measurement has large statistical uncertainties, it was determined that it would be difficult to improve with the 2015 data, mostly due to the addition of the VTX detector causing a large number of $\gamma \to e^+e^-$ conversions upon interacting with detector material. The old measurement was therefore used, and a toy Monte Carlo study was conducted to measure the dilution factor d of the asymmetry in $A_N^{J/\psi \to e} = dA_N^{J/\psi}$ arising from decay kinematics. A single particle sim-

ulation was performed to generate 10 million J/ψ particles within $|\eta| < 0.5$ and with a p_T spectrum that was weighted in accordance to a Kaplan function extracted from a previous PHENIX measurement [171]. The J/ψ particles were then forced to decay to e^+e^-, and candidates within nominal central arm acceptance of $|\eta| < 0.35$ were saved for further analysis. An artificial asymmetry was then applied to the J/ψ ϕ distribution per Eq. 1.37. In this simulation, $P = 1$ and $\phi_{pol} = 90°$ to simplify our equation to

$$N(\phi) = N_0(1 + A_N cos(\phi)) \tag{3.29}$$

where N_0 is a normalization factor that can be determined from simulation. It can be seen from Figure 3.96 that the majority of decay leptons within the $1.0\ GeV < p_T < 5.0\ GeV$ range are contained within $0\ GeV < p_T^{J/\psi} < 6.0\ GeV$, which corresponds to the measured range in [170]. It was determined from the simulations that above 1.5 GeV, we are typically observing the higher p_T decay lepton. Figures 3.97 and 3.98 show the fitted ϕ distribution of J/ψ particles in the left panel for the first decay lepton p_T bin, corresponding to an input asymmetry of $A_N^{J/\psi} = -0.06$ taken from [170], and of the decay lepton in the first p_T bin in the right panel. It should be noted that since the asymmetry is artificially inserted into the J/ψ ϕ distribution, there is no uncertainty on the fitted amplitude. The parameters $p0$ and $p1$ displayed in the fit panels correspond to N_0 and A_N respectively.

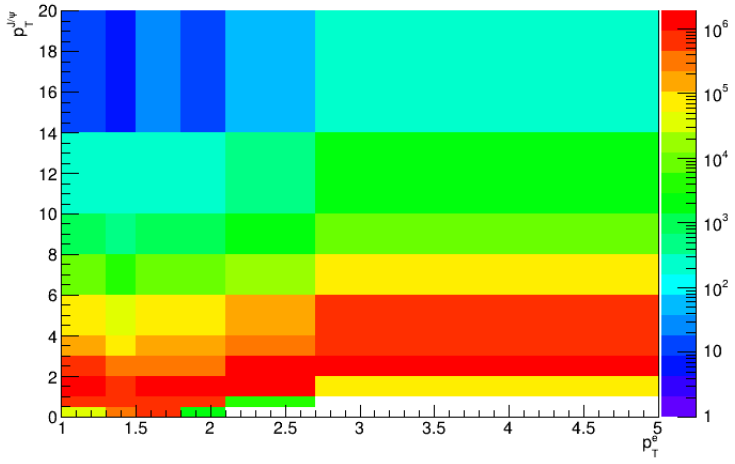

Figure 3.96: J/ψ p_T vs. decay lepton (e) p_T distribution as determined from 10 million simulated midrapidity J/ψ decays.

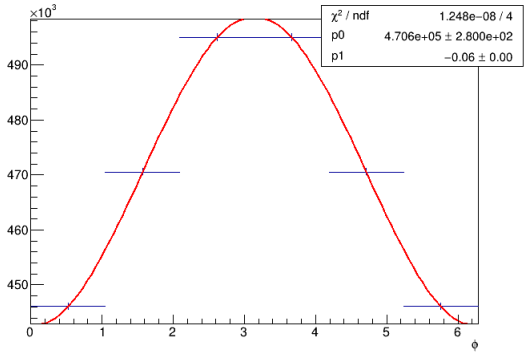

Figure 3.97: Simulated J/ψ asymmetry in accordance with Equation 3.29 within 1.0 GeV $< p_T^e <$ 1.3 GeV.

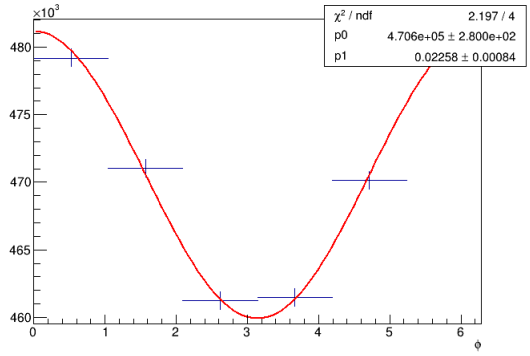

Figure 3.98: Simulated $J/\psi \to e$ asymmetry in accordance with Equation 3.29 within 1.0 GeV $< p_T^e <$ 1.3 GeV.

This process was repeated for 12 different input $A_N^{J/\psi}$ values to test the stability of the dilution factor calculation as a function of the input asymmetry. The results are shown Figure 3.99 for the p_T bin $1.0 - 1.3$ GeV. The gray band indicates the input changing by $\pm\sigma$ where σ is the statistical uncertainty taken from Ref. [170], while the vertical black line indicates the nominal input value of $A_N^{J/\psi} = -0.06$, and

corresponds to the ratios $A_N^{J/\psi\to e}/A_N^{J/\psi}$ shown in Table 3.10.

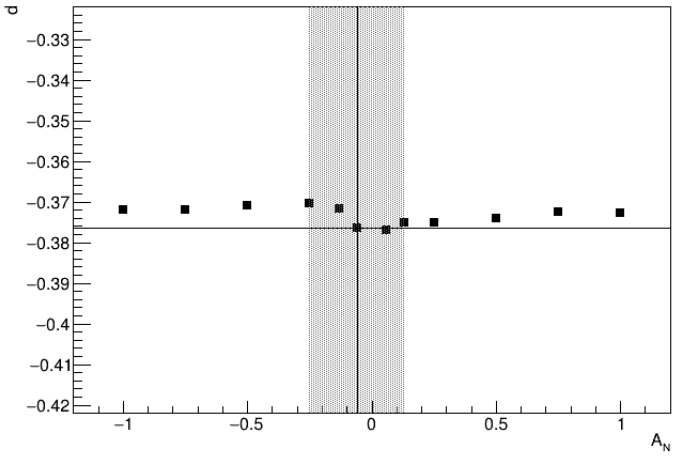

Figure 3.99: Ratio of $A_N^{J/\psi\to e}/A_N^{J/\psi}$ for various input $A_N^{J/\psi}$ values in the first p_T bin.

Table 3.10: Ratio of $A_N^{J/\psi\to e}/A_N^{J/\psi}$ in each p_T bin.

p_T range (GeV)	$\langle p_T \rangle$ (GeV)	d = $A_N^{J/\psi\to e}/A_N^{J/\psi}$
1.0 – 1.3	1.161	-0.376
1.3 – 1.5	1.398	-0.0804
1.5 – 1.8	1.639	0.489
1.8 – 2.1	1.936	0.714
2.1 – 2.7	2.349	0.808
2.7 – 5.0	3.290	0.886

Figures 3.100 and 3.101 show $\phi^{J/\psi}$ vs ϕ^e correlation plots for lowest and highest decay lepton p_T bin respectively, where it can be seen that the flip in sign of decay lepton asymmetry for $p_T^e < 1.5\ GeV$ and increase in magnitude as a function of p_T^e is due to the relative phase between $\phi^{J/\psi}$ and ϕ^e, which is close to $\pm\pi$ in the lowest p_T^e bins, and tends towards 0 as p_T^e increases. The calculated ratios in Table 3.10

are applied to the measured $A_N^{J/\psi}$ value from [170] in order to input $A_N^{J/\psi \to e}$ to the background correction formula for the open heavy flavor electron result.

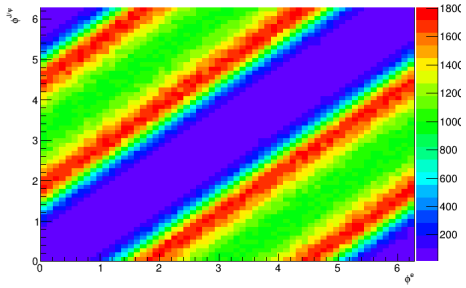

Figure 3.100: J/ψ ϕ vs. decay lepton (e) ϕ correlation plot within 1.0 GeV $< p_T^e <$ 1.3 GeV.

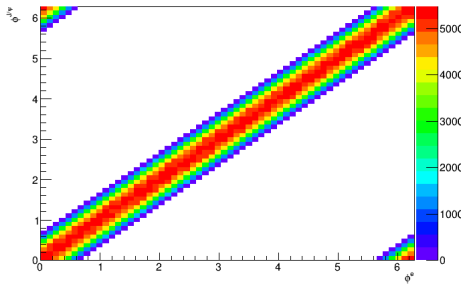

Figure 3.101: J/ψ ϕ vs. decay lepton (e) ϕ correlation plot within 2.7 GeV $< p_T^e <$ 5.0 GeV.

Neutral Mesons: The combinatorial background under the signal peak is studied with the sidebands of the invariant mass distributions, which are $47 < M_{\gamma\gamma} < 97$ MeV $\cup$ $177 < M_{\gamma\gamma} < 227$ MeV for the π^0 measurement and $300 < M_{\gamma\gamma} < 400$ MeV $\cup$ $700 < M_{\gamma\gamma} < 800$ MeV for the η measurement. Figure 3.102 shows a comparison of the relative luminosity and square root asymmetries for the π^0 background region in pAu collisions, while Figures 3.103 shows this for the π^0 background region in pAl collisions. Similarly, Figure 3.104 shows a comparison of the relative luminosity and square root asymmetries for the η background region in pAu collisions, and Figure 3.105 shows this for the η background region in pAl collisions.

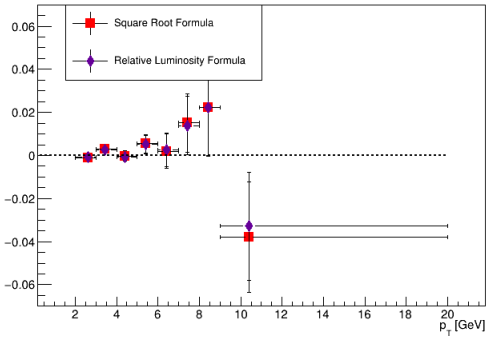

Figure 3.102: Relative luminosity and square root background $A_N(p^{\uparrow}\text{+Au} \rightarrow \pi^0 + X)$ comparison, measured in the π^0 invariant mass side band region.

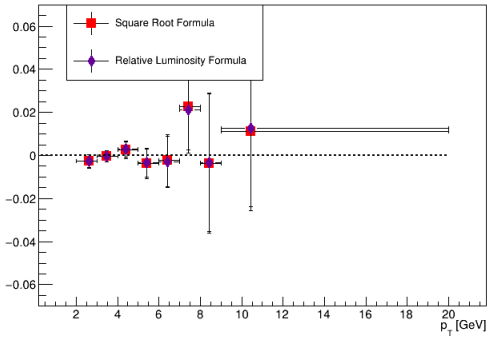

Figure 3.103: Relative luminosity and square root background $A_N(p^{\uparrow}\text{+Al} \rightarrow \pi^0 + X)$ comparison, measured in the π^0 invariant mass side band region.

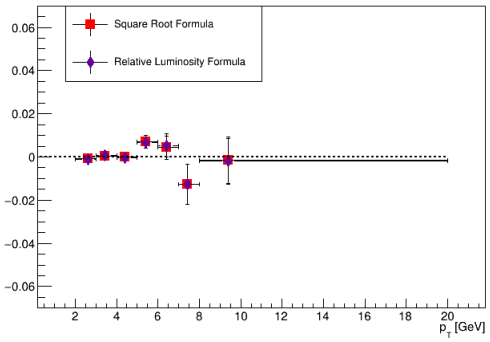

Figure 3.104: Relative luminosity and square root background $A_N(p^{\uparrow}\text{+Au} \rightarrow \eta + X)$ comparison, measured in the η invariant mass side band region.

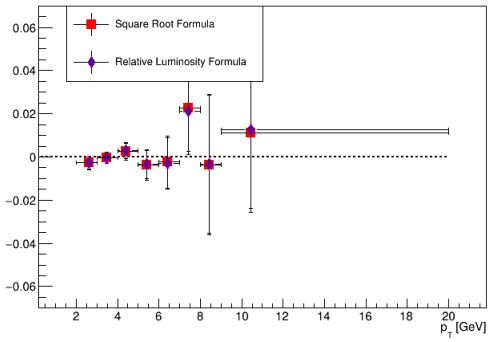

Figure 3.105: Relative luminosity and square root background $A_N(p^{\uparrow}\text{+Al} \rightarrow \eta + X)$ comparison, measured in the η invariant mass side band region.

3.3.4.2 Results After Background Correction

Heavy Flavor Electrons: With the background fractions calculated, and the background asymmetries measured previously at PHENIX [169, 170] and mapped to decay lepton p_T as described in Section 3.3.4.1, one can use Equations 3.23 and 3.24

to calculate the background corrected asymmetries and statistical uncertainties. The results of this procedure are shown in Figure 3.106 for open heavy flavor electron and nonphotonic electron results, in comparison with the results before background correction.

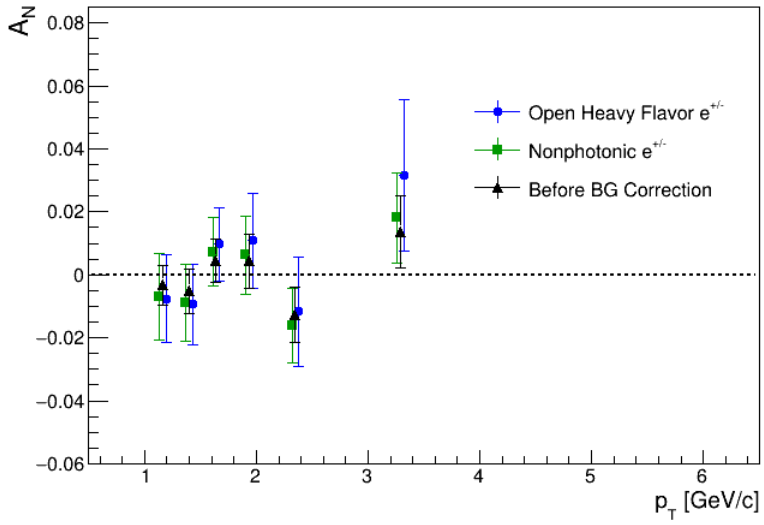

Figure 3.106: Comparison plot of $A_N(p^{\uparrow}p \rightarrow (OHF \rightarrow e^{\pm})X)$ and $A_N(p^{\uparrow}p \rightarrow NPe^{\pm}X)$ for open heavy flavor electron (blue circles) and nonphotonic electron (green squares) results, respectively, plotted alongside the results before background correction (black triangles).

Neutral Mesons: With background fractions and background asymmetries measured, one can use Equations 3.27 and 3.28 to calculate the background corrected asymmetries and statistical uncertainties. The resulting background corrected asymmetries for the relative luminosity formula and the square root formula are shown in Figures 3.107 - 3.110 for π^0 and η mesons in pAu and pAl collisions.

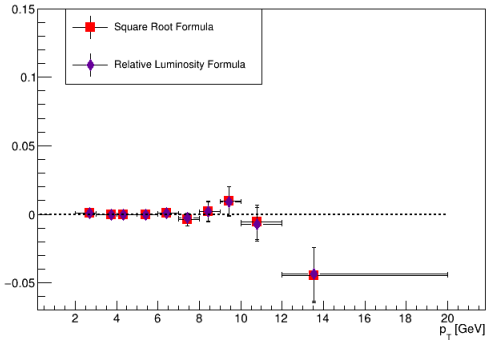

Figure 3.107: Background corrected relative luminosity and square root $A_N(p^{\uparrow}\text{Au}\rightarrow \pi^0 X)$ comparison.

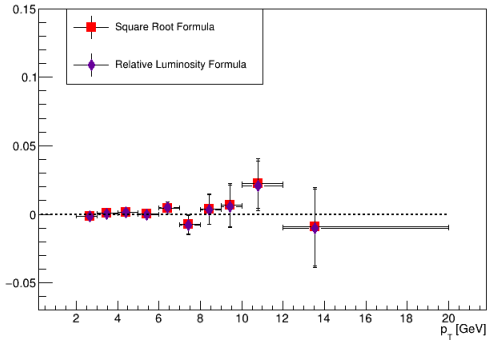

Figure 3.108: Background corrected relative luminosity and square root $A_N(p^{\uparrow}\text{Al}\rightarrow \pi^0 X)$ comparison.

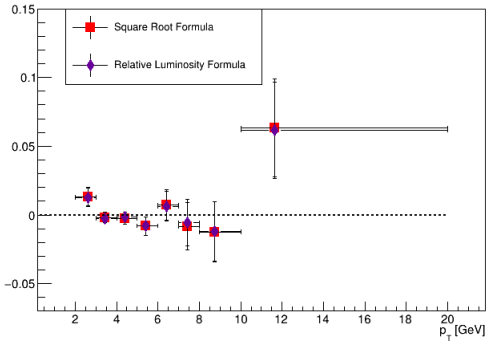

Figure 3.109: Background corrected relative luminosity and square root $A_N(p^{\uparrow}\text{Au}\rightarrow \eta X)$ comparison.

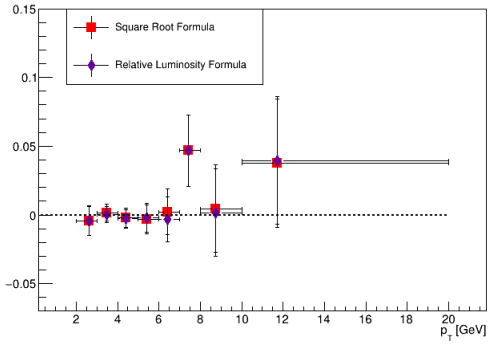

Figure 3.110: Background corrected relative luminosity and square root $A_N(p^{\uparrow}\text{Al}\rightarrow \eta X)$ comparison.

3.4 Systematic Studies

In this section, the procedures for the various systematic cross checks used in this analysis are outlined and results from the studies are summarized.

3.4.1 A_N Formula Difference

As discussed in Section 3.3.1, the asymmetry A_N is calculated by means of two different formulae — namely, the relative luminosity formula, and the square root formula. While the results obtained using the relative luminosity formula are reported as the measured asymmetry values, the difference between results obtained using the two methods is assigned as a systematic uncertainty using Equation 3.30,

$$\sigma_{diff}^{syst} = |A_N^{RelLumi} - A_N^{Sqrt}|. \tag{3.30}$$

The systematic uncertainties obtained from this procedure are summarized in Section 3.4.4 for both the open heavy flavor electrons and neutral mesons.

3.4.2 Background Correction

3.4.2.1 Heavy Flavor Electrons

Background Fractions: Systematic uncertainties on A_N are assigned based on results obtained by calculating background corrected asymmetries using the systematic uncertainty bounds of the background fractions rather than the nominal values, as shown in Equation 3.31

$$A_N^S \pm \sigma_{f^\pm}^{syst} = \frac{A_N^{S+B} - \sum_i (f_i \pm \sigma_{f_i}^\pm) A_N^{B_i}}{1 - (\sum_i f_i \pm \sigma_{f_i}^\pm)}, \tag{3.31}$$

where the S and B superscripts stand for signal and background respectively, $f_i \pm \sigma_{f_i}^\pm$ are the background fractions and corresponding systematic uncertainties for each background source, and $A_N^{B_i}$ are TSSAs for each background source.

In order to properly determine the systematic uncertainties on the background fractions used as inputs to Equation 3.31, additional systematic uncertainties on the background fractions in the $1.0 < p_T < 1.5$ GeV region were considered to those

shown in Figures 3.31– 3.33. This is due to the uncertainty of the nonphotonic electron fraction (F_{NP}). Studies conducted during a previous analysis on the heavy flavor electron sample in the same run [161] determined that the F_{NP} calculation was only reliable within uncertainties displayed on Figure 3.30 above 1.5 GeV, with concerns that it was underestimated in the $1.0 < p_T < 1.5$ GeV region. A conservative systematic uncertainty on F_{NP} was therefore applied, as shown in Figure 3.111, the upper and lower bounds of which were plugged into Eq. 3.7 to determine the additional systematic uncertainty on the background fractions below $p_T < 1.5$ GeV. The upper bound of the systematic uncertainties for F_{NP} in Figure 3.111 was taken as the central value of F_{NP} in the following p_T bin, and this was applied as a symmetric systematic uncertainty ($\sigma^{syst}_{F_{np}} = \mid F^{p_T^{i+1}}_{NP} - F^{p_T^i}_{NP} \mid$). It is expected that the fraction of nonphotonic electrons should decrease with decreasing p_T, proving this to be a conservative estimate of systematic uncertainty on the nonphotonic electron fraction. The systematic uncertainties on A_N from propagating uncertainties on the background fractions (shown in Eq. 3.31) include the additional systematic uncertainty from the F_{NP} calculation below 1.5 GeV presented in this section.

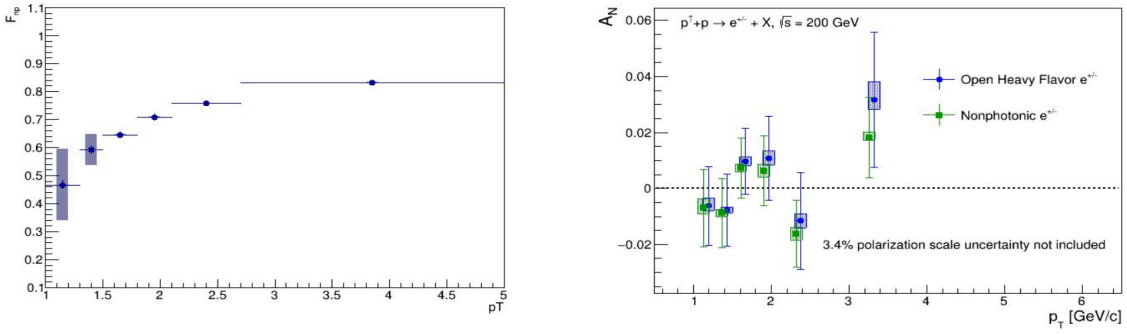

Figure 3.111: F_{NP} as a function of p_T with additional symmetric systematic uncertainties given by $\mid F^{p_T^{i+1}}_{NP} - F^{p_T^i}_{NP} \mid$ on the lowest two data points shown by the shaded blue rectangles (left panel) and the resulting final systematic uncertainties on A_N in the lowest two p_T bins (right panel) — this was the chosen method used in this analysis to assign additional uncertainties to F_{NP} and therefore A_N.

The systematic uncertainties calculated with Eq. 3.31 ($\sigma_{f^{\pm}}^{syst}$) for open heavy flavor electrons are summarized in Table 3.13, while Table 3.14 shows this for nonphotonic electrons. The uncertainties $\sigma_{f^{\pm}}^{syst}$ are asymmetric due to the asymmetric systematic uncertainty of the hadron contamination background fraction $f_{h^{\pm}}$, calculated as described in Section 3.2.1.4.

Background Asymmetries: Systematic uncertainties on A_N are also assigned based on the assumption of setting $A_N^{\pi^0,\eta,\gamma} = 0$ in Equations 3.21 and 3.25 and $\sigma_{A_N^{\pi^0,\eta,\gamma}} = 0$ in Equations 3.22 and 3.26. They are obtained via the difference of calculated background corrected asymmetries using the measured uncertainty bounds on ($A_N^{\pi^0}$, A_N^{η}) and the nominal values, as shown in Equation 3.32 for the open heavy flavor results and in Equation 3.33 for the nonphotonic results. It was determined by calculation that the inclusion of the measured uncertainty of A_N^{γ} had a negligible effect, given the measured background fractions f_{γ} are on the order of 10^{-3} (see Figures 3.31 – 3.33). It was therefore not considered in the following,

$$A_N^{OHF\to e} \pm \sigma_{A_N^B}^{sys} = \frac{A_N^e - f_{h^{\pm}} A_N^{h^{\pm}} - f_{J/\psi\to e} A_N^{J/\psi\to e} - f_{\pi^0\to e}(0 \pm \sigma_{A_N^{\pi^0}}) - f_{\eta\to e}(0 \pm \sigma_{A_N^{\eta}})}{1 - f_{h^{\pm}} - f_{J/\psi\to e} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}} \tag{3.32}$$

$$A_N^{NP\to e} \pm \sigma_{A_N^B}^{sys} = \frac{A_N^e - f_{h^{\pm}} A_N^{h^{\pm}} - f_{\pi^0\to e}(0 \pm \sigma_{A_N^{\pi^0}}) - f_{\eta\to e}(0 \pm \sigma_{A_N^{\eta}})}{1 - f_{h^{\pm}} - f_{\pi^0\to e} - f_{\eta\to e} - f_{\gamma\to e}}. \tag{3.33}$$

Equations 3.32 and 3.33 provide a definition for the symmetric systematic uncertainties $\sigma_{A_N^B}^{syst}$. The measured uncertainties on $\sigma_{A_N^{\pi^0}}$ and $\sigma_{A_N^{\eta}}$ are taken from [123]. Figures 3.112 and 3.113 show the correlation matrices of $p_T^{\pi^0}$ vs. p_T^e and p_T^{η} vs. p_T^e from studies conducted for Ref. [161], which corresponds to the same heavy flavor electron sample aside from subtle differences discussed in Section 3.2.1.1. It can be seen in Figure 3.112 that the probability of observing an electron in our measured p_T^e range of $(1.0, 5.0)$ GeV from a π^0 decay with $p_T^{\pi^0} > 7$ GeV is negligible. Similarly, Figure 3.113 shows that the probability of observing an electron in our measured p_T^e

range of $(1.0, 5.0)$ GeV from an η decay with $p_T^{\eta} > 6$ GeV is negligible. Based on this analysis, the measured uncertainties input to Equations 3.32 and 3.33 for all electron p_T bins are $\sigma_{A_N^{\pi^0}} = 0.00187$ corresponding to the statistical uncertainty of the 6-7 $p_T^{\pi^0}$ bin and $\sigma_{A_N^{\eta}} = 0.00421$ corresponding to the statistical uncertainty of the 5-6 $p_T^{\pi^0}$ bin [123]. This is the most conservative uncertainty estimate based on statistical uncertainties measured in the relevant $p_T^{\pi^0,\eta}$ ranges corresponding to $1.0 < p_T^e < 5.0$ GeV. Furthermore, the inclusion of the systematic uncertainty in $\sigma_{A_N^{\pi^0}}$ and $\sigma_{A_N^{\eta}}$ affects only the fourth significant figure, and therefore is negligible. The systematic uncertainties calculated using Equation 3.32 are summarized in Tables 3.13 and 3.14 for heavy flavor and nonphotonic electrons respectively.

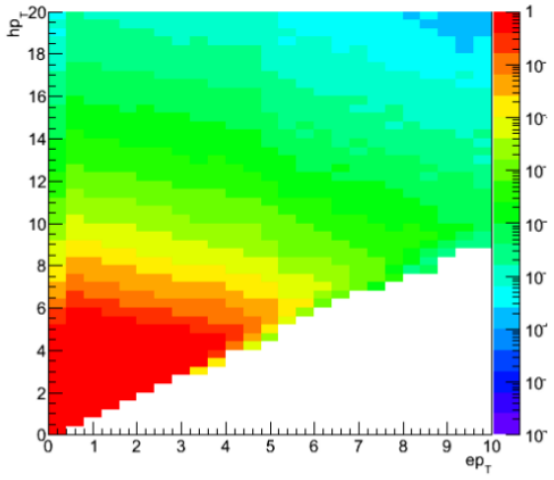

Figure 3.112: 2D distribution of π^0 p_T vs decay electron p_T.

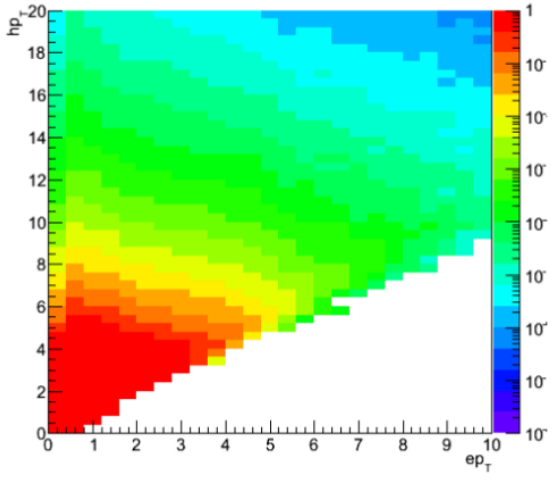

Figure 3.113: 2D distribution of η p_T vs decay electron p_T.

3.4.2.2 Neutral Mesons:

Systematic uncertainties on the background fractions are obtained by adjusting the fit ranges of the third-order polynomial describing the background under the signal peak and recalculating the background fractions. Two alternative fit ranges were used for the π^0 and η mesons respectively, and the maximum difference between

the nominally calculated background fractions and the newly calculated background fractions was taken as a systematic uncertainty on r. This procedure follows from Ref. [123], and the results are summarized in Tables 3.11 and 3.12. These systematic uncertainties are propagated to A_N by calculating background corrected asymmetries using the systematic uncertainty bounds of background fractions rather than the nominal values, as shown in Equation 3.34

$$A_N^S \pm \sigma_r^{syst} = \frac{A_N^{S+B} - (r \pm \sigma_r)A_N^B}{1 - (r \pm \sigma_r)}. \tag{3.34}$$

Here $r \pm \sigma_r$ are the background fractions and corresponding systematic uncertainties. The resulting systematic uncertainty on A_N for each p_T bin is summarized in Tables 3.15- 3.18.

Table 3.11: Systematic uncertainty assigned to r based on adjusting the invariant mass fit regions for π^0 mesons in pA collisions.

p_T [GeV]	σ_r (pAu)	σ_r (pAl)
2 - 3	0.0073	0.0071
3 - 4	0.0089	0.0088
4 - 5	0.0075	0.0071
5 - 6	0.0043	0.0049
6 - 7	0.0019	0.0017
7 - 8	0.0022	0.0023
8 - 9	0.012	0.011
9 - 10	0.0097	0.014
10 - 12	0.010	0.010
12 - 20	0.0030	0.0012

Table 3.12: Systematic uncertainty assigned to r based on adjusting the invariant mass fit regions for η mesons in pA collisions.

p_T [GeV]	σ_r (pAu)	σ_r (pAl)
2 - 3	0.0012	0.0038
3 - 4	0.0024	0.0021
4 - 5	0.0037	0.0042
5 - 6	0.0027	0.0035
6 - 7	0.0071	0.0063
7 - 8	0.0063	0.0077
8 - 10	0.0042	0.0063
10 - 20	0.017	0.010

3.4.3 Bunch Shuffling

Bunch shuffling is a procedure that involves randomizing the assigned spin direction of each bunch in the polarized proton beam such that any measurable physics asymmetry disappears, leaving only statistical fluctuations in the data unless additional sources of systematic uncertainty are present. The spin direction randomization is done on a fill-by-fill basis, and the fills are again grouped chronologically in pairs when calculating the asymmetry. The square root formula (Eq. 3.14) is used to avoid having to recalculate the relative luminosity for each iteration of the bunch shuffling. This procedure was repeated for 10,000 iterations, such that Gaussian statistics are applicable in analyzing the results. In this case, it is expected that fits of A_N/σ_{A_N} well replicate a $N(0,1)$ distribution — that is, a Gaussian distribution with a mean of 0 and a standard deviation of 1.

Heavy Flavor Electrons: Results for A_N/σ_{A_N} are shown in Figures 3.114 and 3.115 for the lowest 2 p_T bins in the heavy flavor electron analysis, where it can be seen in the fit panel located in the top right corner of each figure that the mean (entry labeled "Mean") is consistent with 0 and the standard deviation (entry labeled "Sigma") is consistent with 1. This behavior was observed for each p_T bin used in the analysis, implying that only statistical fluctuations are present in the data after the bunch shuffling procedure and that no additional systematic uncertainties need to be assigned based on these results.

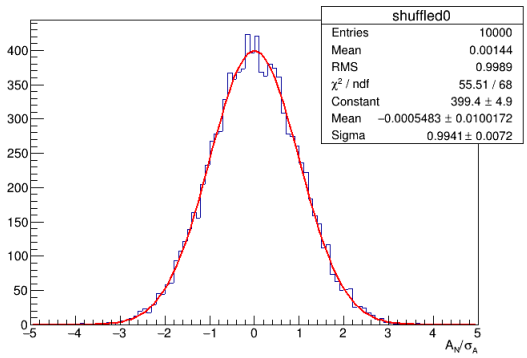

Figure 3.114: Distribution of A_N/σ_{A_N} for open heavy flavor $e^{\pm}$ obtained from the bunch shuffling procedure within $1.0 < p_T < 1.3$ GeV.

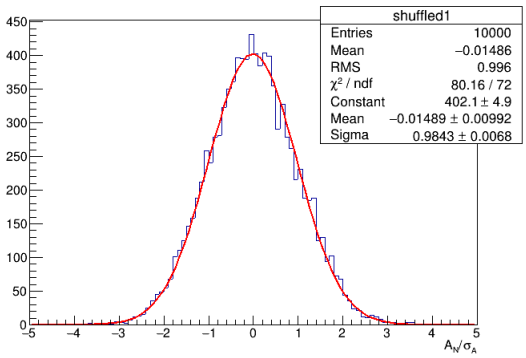

Figure 3.115: Distribution of A_N/σ_{A_N} for open heavy flavor $e^{\pm}$ obtained from the bunch shuffling procedure within $1.3 < p_T < 1.5$ GeV.

Neutral Mesons: Results for A_N/σ_{A_N} are shown for the lowest 2 p_T bins for π^0 candidates in pAu collisions in Figures 3.116 and 3.117, π^0 candidates in pAl collisions in Figures 3.120 and 3.121, η candidates in pAu collisions in Figures 3.124 and 3.125, and η candidates in pAl collisions in Figures 3.128 and 3.129. The fit panel located in the top right corner of each figure shows the mean (entry labeled "Mean") and the standard deviation (entry labeled "Sigma"). It can be seen from the summary plots that all of the means are consistent with zero, however, there are some p_T

bins, particularly at low p_T where the standard deviation of the fit is significantly different than 1. A systematic uncertainty was therefore added in each p_T bin where the standard deviation of the fit was more than 3 times the fit uncertainty greater than 1. The systematic uncertainties from this study are assigned such that the systematic and statistical uncertainty of A_N added in quadrature is equal to σ_{A_N} times the value obtained for the standard deviation from the fit of the A_N/σ_{A_N} distribution in the bunch shuffling procedure (σ_i), as shown in Eq. 3.35. The resulting systematic uncertainty on A_N for the relevant p_T bins is calculated as shown in Eq. 3.36, and summarized in Tables 3.15-3.18. Figures 3.118 and 3.119 show the mean and standard deviation values respectively obtained from the fits in each p_T bin in the π^0 pAu analysis. This is shown in Figures 3.122 and 3.123 for the π^0 pAl analysis, Figures 3.126 and 3.127 for the η pAu analysis, and Figures 3.130 and 3.131 for the η pAl analysis.

$$\sigma_i \sigma_{stat} = \sqrt{\sigma_{stat}^2 + \sigma_{syst,bs}^2} \tag{3.35}$$

$$\sigma_{syst,bs} = \sigma_{stat}\sqrt{\sigma_i^2 - 1} \tag{3.36}$$

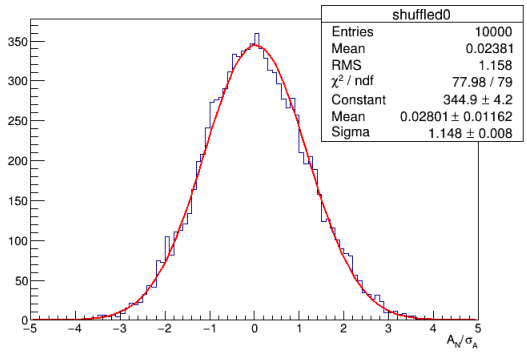

Figure 3.116: Distribution of A_N/σ_{A_N} for π^0 mesons within $2 < p_T < 3$ GeV in pAu collisions.

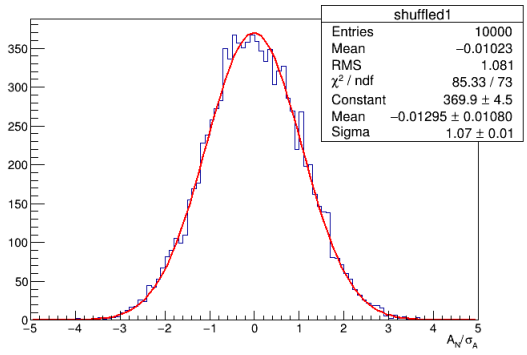

Figure 3.117: Distribution of A_N/σ_{A_N} for π^0 mesons within $3 < p_T < 4$ GeV in pAu collisions.

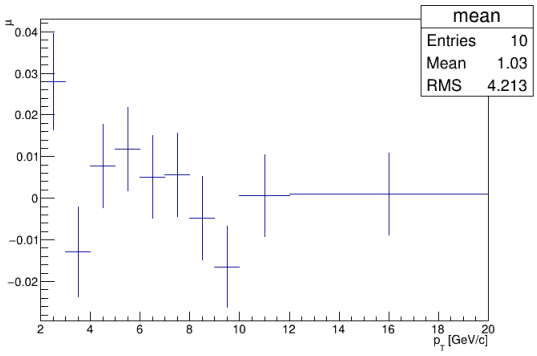

Figure 3.118: Summary plot of means and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for π^0 mesons in pAu collisions.

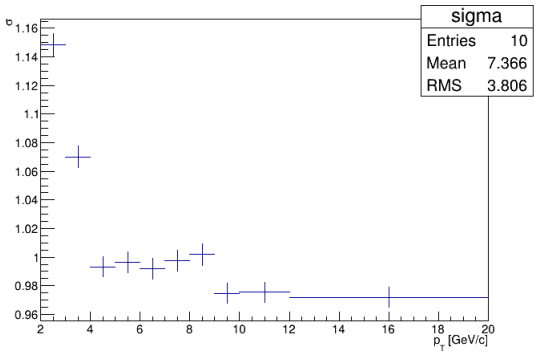

Figure 3.119: Summary plot of standard deviations and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for π^0 mesons in pAu collisions.

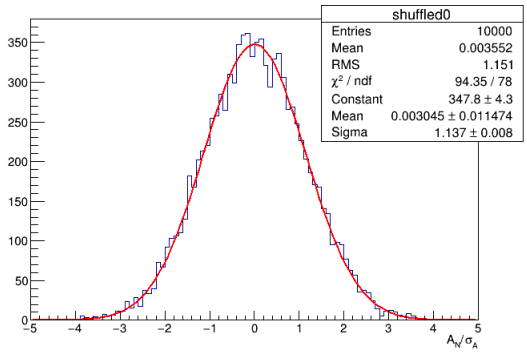

Figure 3.120: Distribution of A_N/σ_{A_N} for π^0 mesons within $2 < p_T < 3$ GeV in pAl collisions.

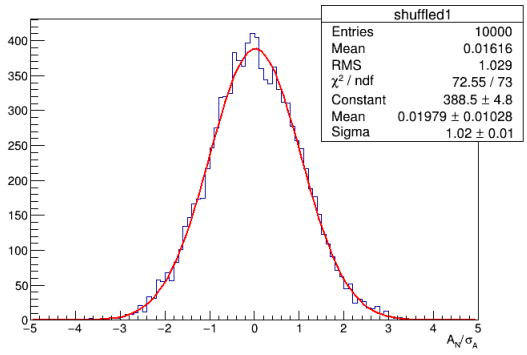

Figure 3.121: Distribution of A_N/σ_{A_N} for π^0 mesons within $3 < p_T < 4$ GeV in pAl collisions.

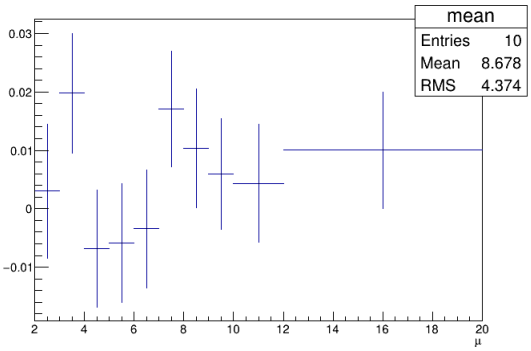

Figure 3.122: Summary plot of means and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for π^0 mesons in pAl collisions.

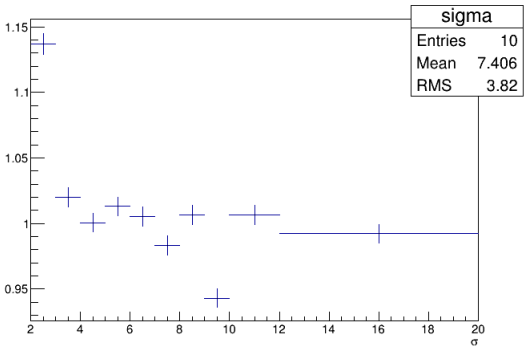

Figure 3.123: Summary plot of standard deviations and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for π^0 mesons in pAl collisions.

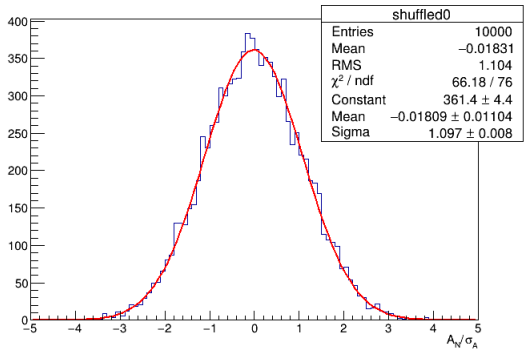

Figure 3.124: Distribution of A_N/σ_{A_N} for η mesons within $2 < p_T < 3$ GeV in pAu collisions.

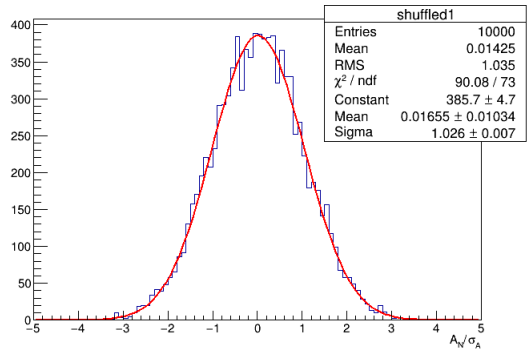

Figure 3.125: Distribution of A_N/σ_{A_N} for η mesons within $3 < p_T < 4$ GeV in pAu collisions.

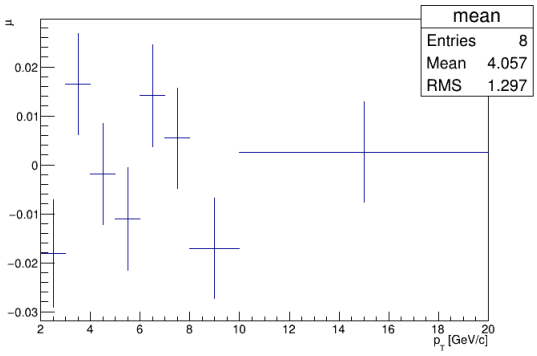

Figure 3.126: Summary plot of means and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for η mesons in pAu collisions.

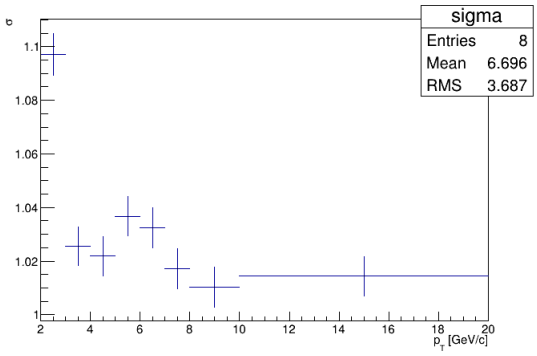

Figure 3.127: Summary plot of standard deviations and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for η mesons in pAu collisions.

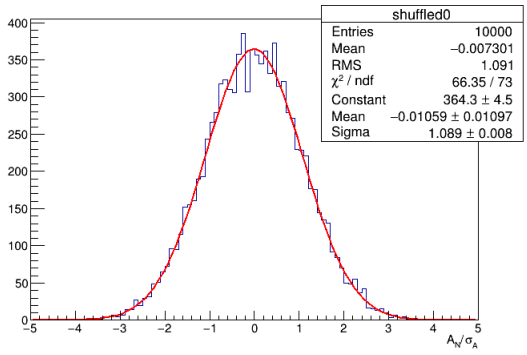

Figure 3.128: Distribution of A_N/σ_{A_N} for η mesons within $2 < p_T < 3$ GeV in pAl collisions.

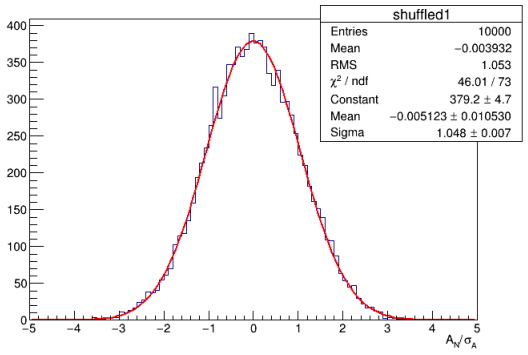

Figure 3.129: Distribution of A_N/σ_{A_N} for η mesons within $3 < p_T < 4$ GeV in pAl collisions.

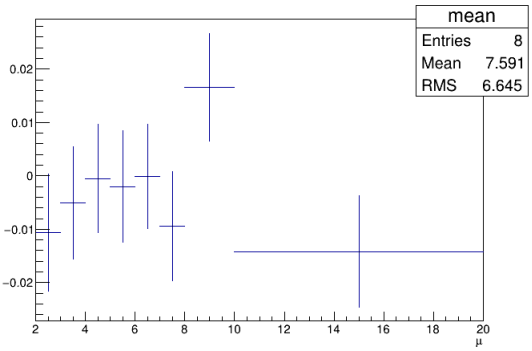

Figure 3.130: Summary plot of means and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for η mesons in pAl collisions.

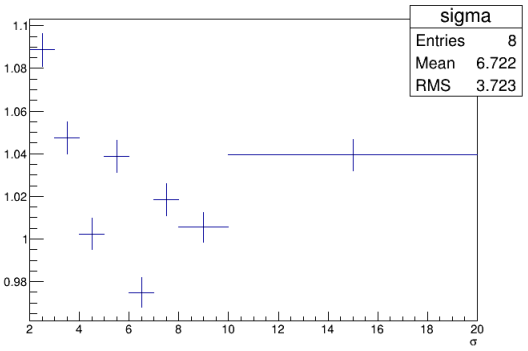

Figure 3.131: Summary plot of standard deviations and respective errors obtained from the Gaussian fits of the A_N/σ_{A_N} distributions for η mesons in pAl collisions.

3.4.4 Summary of Systematic Uncertainties

Heavy Flavor Electrons: All Systematic uncertainties described in Section 3.4 for the open heavy flavor electrons are summarized in Table 3.13 and for the nonphotonic electrons in Table 3.14.

Table 3.13: Summary of systematic uncertainty on $A_N^{OHF\to e^\pm}$ in $p^\uparrow p$ collisions.

p_T range (GeV)	$\langle p_T \rangle$ (GeV)	$\sigma^{syst}_{f^+}$	$\sigma^{syst}_{f^-}$	$\sigma^{syst}_{A_N^B}$	σ^{syst}_{diff}	$\sigma^{syst}_{tot^+}$	$\sigma^{syst}_{tot^-}$
1.0 – 1.3	1.161	0.0037	0.00193	0.00254	0.00146	0.00472	0.00351
1.3 – 1.5	1.398	0.00136	0.00104	0.00159	0.000907	0.00228	0.0021
1.5 – 1.8	1.639	0.00073	0.000581	0.00133	0.00139	0.00206	0.00201
1.8 – 2.1	1.936	0.00112	0.000915	0.000993	0.00235	0.00278	0.00271
2.1 – 2.7	2.349	0.000896	0.000694	0.000809	0.00237	0.00266	0.0026
2.7 – 5.0	3.290	0.00588	0.00362	0.000534	0.000133	0.0059	0.00366

Table 3.14: Summary of systematic uncertainty on $A_N^{NPe^\pm}$ in $p^\uparrow p$ collisions.

p_T range (GeV)	$\langle p_T \rangle$ (GeV)	$\sigma^{syst}_{f^+}$	$\sigma^{syst}_{f^-}$	$\sigma^{syst}_{A_N^B}$	σ^{syst}_{diff}	$\sigma^{syst}_{tot^+}$	$\sigma^{syst}_{tot^-}$
1.0 – 1.3	1.161	0.00295	0.00159	0.00249	0.00146	0.00413	0.00329
1.3 – 1.5	1.398	0.00108	0.000846	0.00153	0.000907	0.00208	0.00197
1.5 – 1.8	1.639	0.000267	0.000177	0.00125	0.00139	0.00189	0.00188
1.8 – 2.1	1.936	0.000222	0.000136	0.000914	0.00235	0.00253	0.00252
2.1 – 2.7	2.349	2.41e-05	0.000113	0.000717	0.00237	0.00248	0.00248
2.7 – 5.0	3.290	0.00192	0.000768	0.000451	0.000133	0.00198	0.000901

Neutral Mesons: All systematic uncertainties described in Section 3.4 are summarized in Table 3.15 for π^0 in pAu collisions, Table 3.17 for π^0 in pAl collisions, Table 3.16 for η in pAu collisions, and Table 3.15 for η in pAl collisions.

Table 3.15: Summary of systematic uncertainty on $A_N^{\pi^0}$ in $p^{\uparrow}$Au collisions.

p_T range (GeV/c)	σ_{diff}^{syst}	σ_r^{syst}	σ_{bs}^{syst}	σ_{tot}^{syst}
2 - 3	9.56e-05	1.4e-05	0.00056	0.000569
3 - 4	9.83e-05	2.21e-05	0.000267	0.000286
4 - 5	0.000257	5.24e-06	0	0.000257
5 - 6	0.000277	4.24e-05	0	0.00028
6 - 7	0.000245	1.13e-05	0	0.000245
7 - 8	0.00109	0.000133	0	0.00109
8 - 9	0.000526	0.000165	0	0.000551
9 - 10	0.000432	0.000334	0	0.000546
10 - 12	0.00139	0.000196	0	0.00141
12 - 20	0.000765	7.93e-05	0	0.000769

Table 3.16: Summary of systematic uncertainty on A_N^{η} in $p^{\uparrow}$Au collisions.

p_T range (GeV/c)	σ_{diff}^{syst}	σ_r^{syst}	σ_{bs}^{syst}	σ_{tot}^{syst}
2 - 3	0.000545	0.00025	0.003	0.00306
3 - 4	0.000759	4.13e-05	0.000858	0.00115
4 - 5	0.000696	1.39e-05	0.000942	0.00117
5 - 6	3.06e-07	0.000123	0.00182	0.00183
6 - 7	0.000904	8.49e-06	0.00278	0.00293
7 - 8	0.00258	5.29e-05	0	0.00258
8 - 10	0.000204	7.46e-05	0	0.000217
10 - 20	0.00209	0.00043	0	0.00213

Table 3.17: Summary of systematic uncertainty on $A_N^{\pi^0}$ in $p^{\uparrow}$Al collisions.

p_T range (GeV/c)	σ_{diff}^{syst}	σ_r^{syst}	σ_{bs}^{syst}	σ_{tot}^{syst}
2 - 3	6.72e-06	1.11e-05	0.000882	0.000882
3 - 4	6.35e-05	8.56e-06	0	6.41e-05
4 - 5	4.93e-05	1.2e-05	0	5.08e-05
5 - 6	0.000511	2.46e-05	0	0.000512
6 - 7	0.000417	6.12e-05	0	0.000422
7 - 8	0.000555	0.000229	0	0.0006
8 - 9	0.000368	5.32e-05	0	0.000371
9 - 10	0.000988	6.02e-05	0	0.00099
10 - 12	0.00163	7.1e-05	0	0.00163
12 - 20	0.00105	0.000169	0	0.00107

Table 3.18: Summary of systematic uncertainty on A_N^{η} in $p^{\uparrow}$Al collisions.

p_T range (GeV/c)	σ_{diff}^{syst}	σ_r^{syst}	σ_{bs}^{syst}	σ_{tot}^{syst}
2 - 3	0.000253	4.44e-05	0.00458	0.00459
3 - 4	0.00126	6.29e-06	0.0018	0.00219
4 - 5	0.0005	8.61e-06	0	0.0005
5 - 6	0.00096	9.93e-05	0.00293	0.00308
6 - 7	0.00549	2.48e-05	0	0.00549
7 - 8	6.62e-05	0.000428	0	0.000433
8 - 10	0.00274	2.15e-06	0	0.00274
10 - 20	0.00171	0.000286	0.0131	0.0133

CHAPTER IV

Results and Discussion

The final results for both the heavy flavor electron TSSA measurement in $\sqrt{s} =$ 200 GeV $p^{\uparrow}p$ collisions (see Ref. [1]) and the neutral pion and eta meson TSSA measurement in $\sqrt{s_{NN}} = 200o$ GeV $p^{\uparrow}A$ collisions (see Ref. [2]) outlined in Chapter III are discussed in the following. A discussion of the results in light of recent global measurements is also provided.

4.1 Heavy Flavor Electron TSSAs

The final transverse single-spin asymmetries for both open heavy flavor electrons and nonphotonic electrons are shown in Figure 4.1 (including positrons), while Figures 4.2 and 4.3 show the final results for the positrons and electrons, respectively. The nominal asymmetry values and statistical uncertainties are calculated with the relative luminosity formula. Both the electrons, positrons, and the charge combined sample are consistent with zero across the entire p_T range for both open heavy flavor and nonphotonic decay leptons. The statistical uncertainties are shown by the bars attached to each point, and the systematic uncertainties are represented by the shaded boxes. There is no dominant source of systematic uncertainty across all p_T bins (see Section 3.4 for details). The results for the asymmetries, statistical uncertainties, and systematic uncertainties are summarized in Table 4.1 for open heavy flavor electrons

and positrons, and Table 4.2 for non-photonic electrons and positrons.

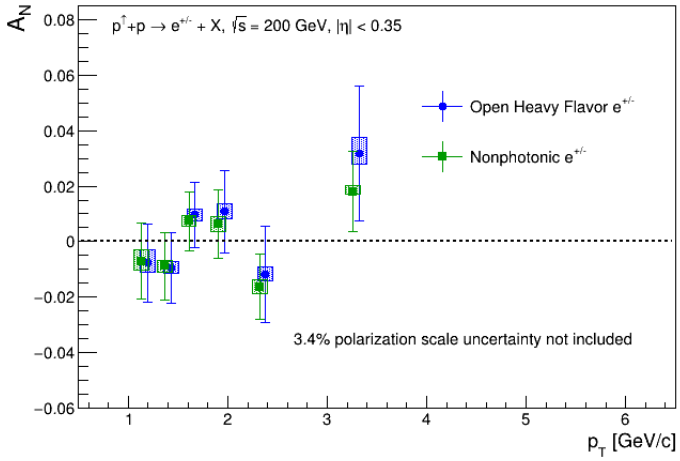

Figure 4.1: Final A_N for open heavy flavor $e^{\pm}$ and non-photonic $e^{\pm}$ results with full systematic uncertainties, indicated by the upper and lower bounds of the shaded rectangles on each point.

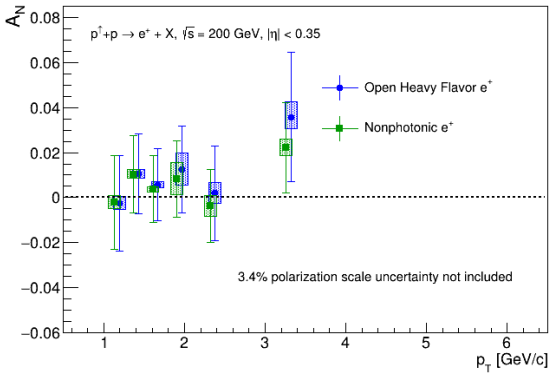

Figure 4.2: Final A_N for open heavy flavor e^+ and non-photonic e^+ results with full systematic uncertainties.

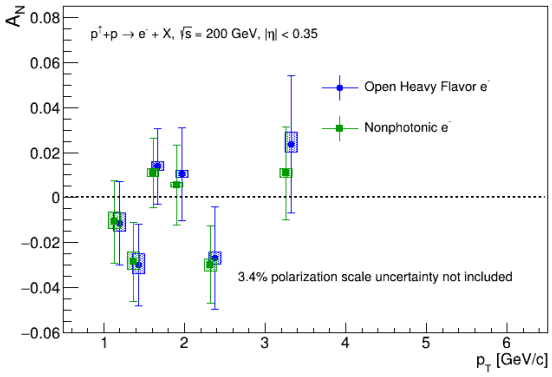

Figure 4.3: Final A_N for open heavy flavor e^- and non-photonic e^- results with full systematic uncertainties.

Table 4.1: Summary of final asymmetries $A_N^{\mathrm{OHF}\to\mathrm{e}^\pm}$ for open-heavy-flavor positrons and electrons with statistical $\sigma^{A_N^{\mathrm{OHF}\to\mathrm{e}^\pm}}$ and systematic uncertainties, shown in Fig. 4.27. This table was taken from Ref. [1].

$e^\pm$	p_T range (GeV/c)	$\langle p_T \rangle$ (GeV/c)	$A_N^{\mathrm{OHF}\to\mathrm{e}^\pm}$	$\sigma^{A_N^{\mathrm{OHF}\to\mathrm{e}^\pm}}$	$\sigma^{\mathrm{sys}}_{\mathrm{tot}^+}$	$\sigma^{\mathrm{sys}}_{\mathrm{tot}^-}$
e^+	1.0–1.3	1.161	-0.00256	0.0212	0.00330	0.00281
	1.3–1.5	1.398	0.0105	0.0178	0.00211	0.00189
	1.5–1.8	1.639	0.00571	0.0159	0.00134	0.00132
	1.8–2.1	1.936	0.0126	0.0192	0.00710	0.00708
	2.1–2.7	2.349	0.00208	0.0210	0.00473	0.00465
	2.7–5.0	3.290	0.0357	0.0287	0.00688	0.00501
e^-	1.0–1.3	1.161	-0.0113	0.0186	0.00474	0.00343
	1.3–1.5	1.398	-0.0297	0.0181	0.00502	0.00384
	1.5–1.8	1.639	0.0139	0.0167	0.00209	0.00191
	1.8–2.1	1.936	0.0105	0.0207	0.00176	0.00149
	2.1–2.7	2.349	-0.0267	0.0227	0.00269	0.00269
	2.7–5.0	3.290	0.0237	0.0305	0.00541	0.00363

Table 4.2: Summary of final asymmetries $A_N^{\mathrm{NP}e}$ for nonphotonic positrons and electrons with statistical $\sigma^{A_N^{\mathrm{NP}e}}$ and systematic uncertainties. This table was taken from Ref. [1].

$e^{\pm}$	p_T range (GeV/c)	$\langle p_T \rangle$ (GeV/c)	$A_N^{\mathrm{NP}}e$	$\sigma^{A_N^{\mathrm{NP}}e}$	$\sigma^{\mathrm{sys}}_{\mathrm{tot}^+}$	$\sigma^{\mathrm{sys}}_{\mathrm{tot}^-}$
e^+	1.0– 1.3	1.161	-0.00202	0.0207	0.00286	0.00268
	1.3– 1.5	1.398	0.0103	0.0172	0.00198	0.00181
	1.5– 1.8	1.639	0.00379	0.0148	0.00120	0.00120
	1.8– 2.1	1.936	0.00836	0.0170	0.00702	0.00702
	2.1– 2.7	2.349	-0.00371	0.0163	0.00452	0.00451
	2.7– 5.0	3.290	0.0220	0.0201	0.00401	0.00354
e^-	1.0–1.3	1.161	-0.0106	0.0182	0.00416	0.00316
	1.3–1.5	1.398	-0.0284	0.0174	0.00426	0.00337
	1.5–1.8	1.639	0.0111	0.0153	0.00174	0.00168
	1.8–2.1	1.936	0.00565	0.0178	0.00106	0.00103
	2.1–2.7	2.349	-0.0297	0.0171	0.00269	0.00268
	2.7–5.0	3.290	0.0108	0.0207	0.00225	0.00187

4.1.1 Comparison to Theoretical Models

Reference [107] discusses the calculation of A_N for D^0 and $\bar{D}^0$ mesons produced in $\sqrt{s} = 200$ GeV collisions via the twist-3 formalism (see Secton 1.4.3 for details). Figure 4.4 shows these results at midrapidity, with the 3 curves corresponding to varying strength and sign of contributions from the trigluon correlators. Per our request, the authors of Ref. [107] have provided these parametrizations with a granularity of 200 MeV in the range 0.4 - 4.4 GeV.

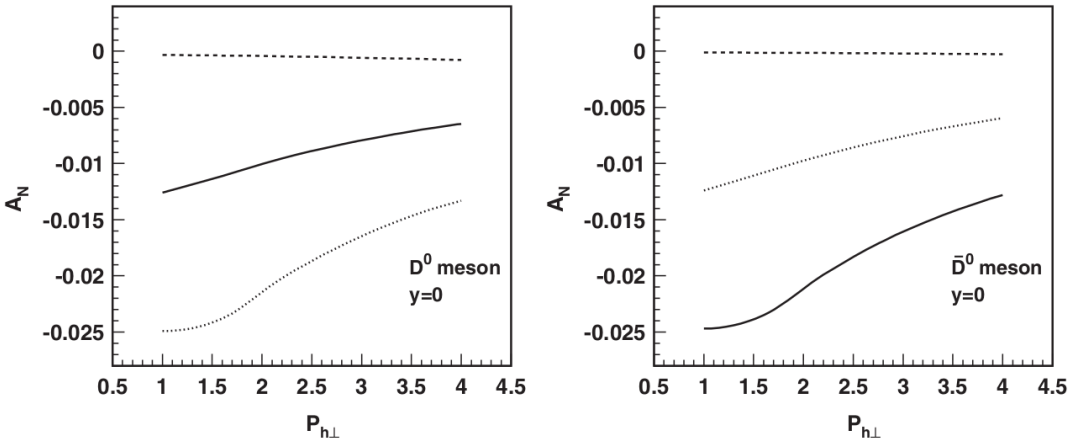

Figure 4.4: Twist-3 theoretical predictions for $A_N^{D^0}$ (left) and $A_N^{\bar{D}^0}$ (right) from Ref. [107]. The dashed curve corresponds to only qgq correlator contributions, showing it is negligible, while the solid and dotted curves correspond to varying strength and signs of the antisymmetric and symmetric ggg correlator contributions.

In order to extract the asymmetries as a function of decay lepton p_T, 10 billion semileptonic decays were generated for both D^0 and $\bar{D}^0$, where all decay channels including an e^+ or e^- were considered. The decaying particles were generated flat in pseudorapidity within $|\eta| < 0.5$, and with a p_T according to the cross section extrapolation from Ref. [161], which is fit with a modified Hagedorn function shown in Equation 4.1 and parameters shown in Table 4.3.

$$f^{D^0}(p_T) = \frac{A}{(exp(-Bp_T - Cp_T^2) + \frac{p_T}{D})^E} \tag{4.1}$$

Table 4.3: D^0 and $\bar{D}^0$ Hagedorn fit parameters, determined during the analysis of Ref. [161].

parameter	value
A	2.38357E-2
B	3.98803E-1
C	7.84526E-2
D	1.86030
E	7.66442

One must take into account an extra factor of $2\pi p_T$ when converting invariant cross sections to yields, therefore D^0 and $\bar{D}^0$ are assigned a p_T value by random sampling from a probability distribution function with the form of $2\pi p_T f^{D^0}(p_T)$. Finally, the decay e^+ or e^- is required to be within detector acceptance in pseudorapidity of $|\eta| < 0.35$. Figure 4.5 shows the $p_T^{D^0}$ vs $p_T^{e^+}$ correlation matrix for a 100 million event sample. Figure 4.6 shows this correlation matrix with additional requirements $0.3 < p_T^{D^0} < 4.5$ GeV and $1.5 < p_T^{e^+} < 5.0$ GeV, in which a significant reduction of statistics is observed. This was the motivation in generating 10 billion decays, to populate the relevant p_T range used in this transverse single-spin asymmetry analysis.

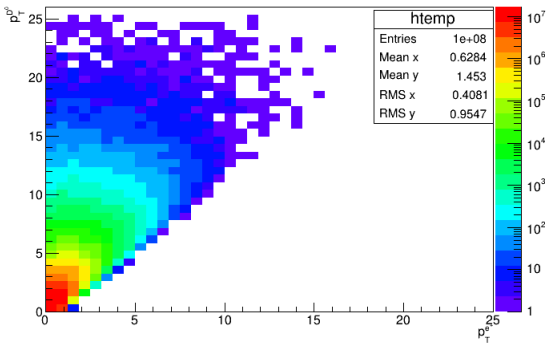

Figure 4.5: Correlation matrix of $p_T^{D^0}$ (vertical axis) vs $p_T^{e^+}$ (horizontal axis) for a 100 million event sample.

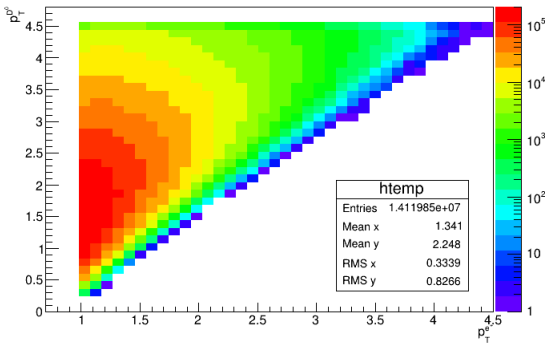

Figure 4.6: Correlation matrix of $p_T^{D^0}$ (vertical axis) vs $p_T^{e^+}$ (horizontal axis) for a 100 million event sample requiring $0.3 < p_T^{D^0} < 4.5$ GeV and $1.0 < p_T^{e^+} < 5.0$ GeV.

Figure 4.7 shows the reproduced $A_N^{D^0}$ and $A_N^{\bar{D}^0}$ curves in the full range provided by the authors of Ref. [107], which are used as inputs to the methods described in Section 3.3.4.1 for extracting $A_N^{D^0 \to e^+}$ and $A_N^{\bar{D}^0 \to e^-}$ from $A_N^{D^0}$ and $A_N^{\bar{D}^0}$, respectively. Figure 4.7 and 4.4 demonstrate good agreement in the range shown in the latter, with only small deviations arising from the provided granularity of the data points. Finally, the decay e^+ or e^- ϕ distribution is fit to extract $A_N^{D^0 \to e^+}$ and $A_N^{\bar{D}^0 \to e^-}$ respectively. This is done in 200 MeV bins within the range 0 - 5 GeV in p_T^e, except for the final bin which covers the range 2.8 - 5.0 GeV due to limited statistics. The results of this study are shown in Figure 4.8.

The $A_N^{D^0 \to e^+}$ and $A_N^{\bar{D}^0 \to e^-}$ curves from Figure 4.8 are plotted alongside the charge separated open heavy flavor $e^{\pm}$ results for positrons in Figure 4.9 and for electrons in Figure 4.10. Note that only the range of 1 - 5 GeV is considered for the theory curves (similarly to the data analysis), and that since the final p_T^e bin is 2.8 - 5.0 GeV, the theory curve only extends to the final bin center at 3.9 GeV.

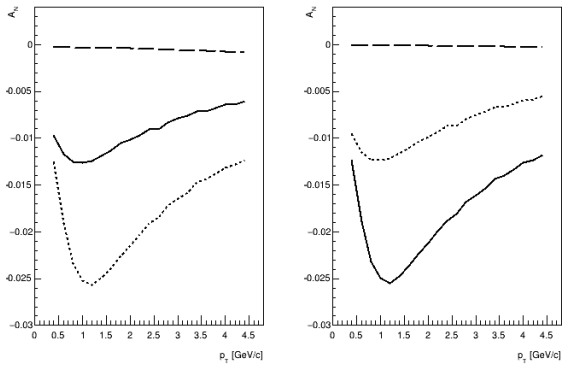

Figure 4.7: Reproduced $A_N^{D^0}$ (left) and $A_N^{\bar{D}^0}$ (right) – full p_T range of 0.4 GeV to 4.4 GeV provided by authors of Ref. [107].

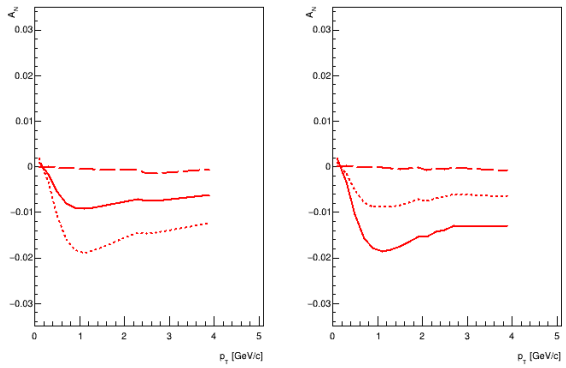

Figure 4.8: Resulting curves for $A_N^{D^0 \to e^+}$ (left) and $A_N^{\bar{D}^0 \to e^-}$ (right) in the range 0 - 5 GeV in p_T^e.

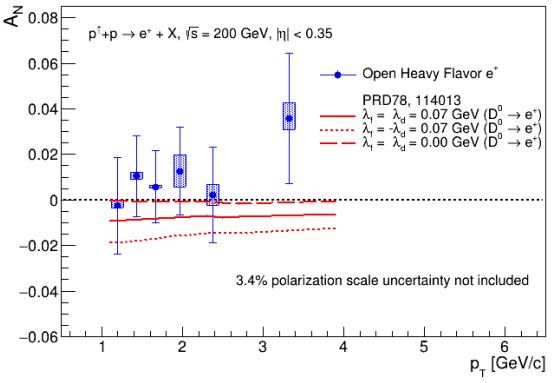

Figure 4.9: Final A_N for open heavy flavor positrons plotted alongside twist-3 theoretical predictions from [107].

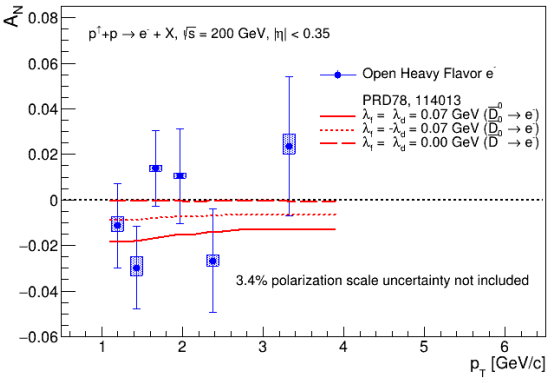

Figure 4.10: Final A_N for open heavy flavor electrons plotted alongside twist-3 theoretical predictions from [107].

A more detailed analysis of the theoretical parameters introduced in Ref. [107] and used to construct Figure 4.7 is discussed in Section 4.1.1.1. Reference [87] also discusses the calculation of A_N for D^0, $\bar{D}^0$, D^+, and D^- mesons produced in $\sqrt{s} = 200$ GeV collisions via the twist-3 formalism — comparisons to models introduced in this

reference are discussed in Section 4.1.1.2.

4.1.1.1 Constraining Theoretical Parameters (λ_f, λ_d)

The asymmetries plotted in Figure 4.4 from Section 4.1.1 receive contributions from 3 separate twist-3 correlation functions, the Efremov-Teryaev-Qiu-Sterman (qgq) correlator and both the antisymmetric and symmetric trigluon (ggg) correlators. The asymmetries for D^0 and $\bar{D}^0$ from Ref. [107] are shown in Equations 4.2 and 4.3.

$$A_N^{D^0}(p_T) = a_0(p_T) + \lambda_f a_1(p_T) + \lambda_d a_2(p_T) \tag{4.2}$$

$$A_N^{\bar{D}^0}(p_T) = b_0(p_T) + \lambda_f a_1(p_T) - \lambda_d a_2(p_T), \tag{4.3}$$

where a_0 and b_0 correspond to contributions from the qgq correlators for D^0 and $\bar{D}^0$, respectively, while a_1 and a_2 correspond to contributions from the antisymmetric (f-type) and symmetric (d-type) trigluon correlators, respectively (see Section 1.4.3 for details). λ_f and λ_d correspond to normalizations of the f and d type trigluon correlators to the unpolarized gluon PDF $g(x)$. The contribution of each separate term in Equation 4.2 is displayed in Figure 4.11, and those of Equation 4.3 in Figure 4.12. When $\lambda_f = \lambda_d = 0$, the only contribution to the asymmetry comes from the qgq correlators, i.e. terms a_0 and b_0. It can be seen from Figures 4.11 and 4.12 that these terms contribute negligibly to the asymmetry, allowing for the contribution of the trigluon correlators from a_1 and a_2 to be investigated by tuning λ_f and λ_d. Comparing Figure 4.11 and Equation 4.2 shows that a_1 and a_2 have opposite sign, and differ in magnitude.

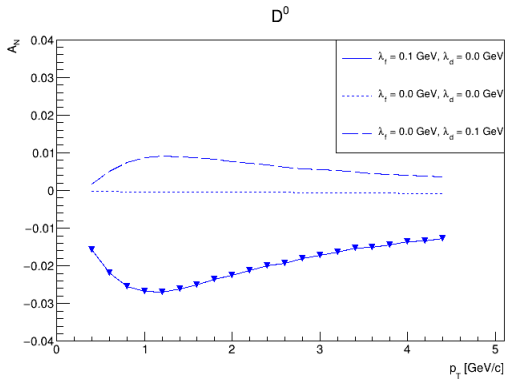

Figure 4.11: Plot of $A_N^{D^0}$ showing contributions from the qgq correlator (dotted), f-type ggg correlator (solid) and d-type ggg correlator (dashed). Based on the model from Ref. [107] shown in Eq. 4.2

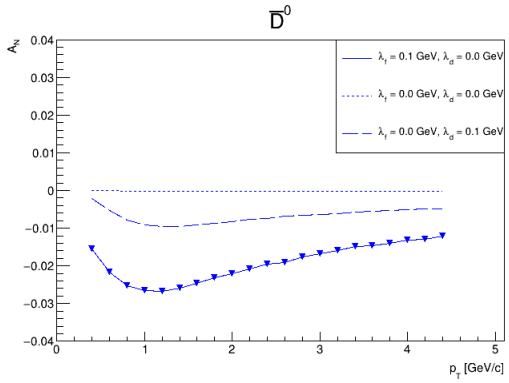

Figure 4.12: Plot of $A_N^{\bar{D}^0}$ showing contributions from the qgq correlator (dotted), f-type ggg correlator (solid) and d-type ggg correlator (dashed). Based on the model from Ref. [107] shown in Eq. 4.3

In order to find parameters that best fit the data, a scan in (λ_f, λ_d) parameter space was performed such that a set of theoretical asymmetries could be generated, corresponding to different combinations of λ parameter values, and compared to data. Initially, a matrix of (λ_f, λ_d) parameters from -1 GeV to 1 GeV in steps of 0.05 GeV was considered, with additional values of $\lambda_f = \pm 0.07$ GeV and $\lambda_d = \pm 0.07$ GeV taken from Ref. [107], resulting in 43×43 possible combinations. In order to accomplish this, a more efficient method for extracting decay lepton asymmetries from parent particle asymmetries was applied than that described in Sections 3.3.4.1 and 4.1.1. The same simulations of 10 billion semileptonic D^0 and $\bar{D}^0$ decays discussed in Section 4.1.1 were used to obtain correlations between p_T and ϕ of the decay lepton and D meson. The ϕ^e distribution was then weighted in accordance with Equation 4.4,

$$w(\phi^e) = 1 + A_N^{D^0}(p_T^{D^0}) \cos \phi^{D^0}. \tag{4.4}$$

The weighted distribution is then fit in the same p_T bins as used in the asymmetry

measurement, shown in Table 3.1. The fit function is shown in Equation 4.5.

$$f(\phi) = p_0(1 + p_1 \cos\phi), \tag{4.5}$$

where p_0 is just a normalization parameter, while p_1 corresponds to A_N^e. Some resulting fits used are shown in Figure 4.13 for positrons and Figure 4.14 for electrons.

These fits are used to extract $A_N^e(p_T)$ for a given parameter combination. Figures 4.15 and 4.16 show that the method outlined in this Section yields good agreement for $A_N^e(p_T)$ with that discussed in Sections 3.3.4.1 and 4.1.1 for the (λ_f, λ_d) parameter combinations plotted in Figures 4.9 and 4.10.

This procedure is repeated for each of the 43×43 parameter combinations such that the simulated asymmetries can be compared with the measured open heavy flavor electron asymmetries in each p_T bin. Figures 4.17 and 4.18 show a subset of parameter combinations corresponding to bounds of $\lambda_f = \pm 0.2$ GeV and $\lambda_d = \pm 0.2$ GeV in steps of 0.2 GeV for positrons and electrons respectively.

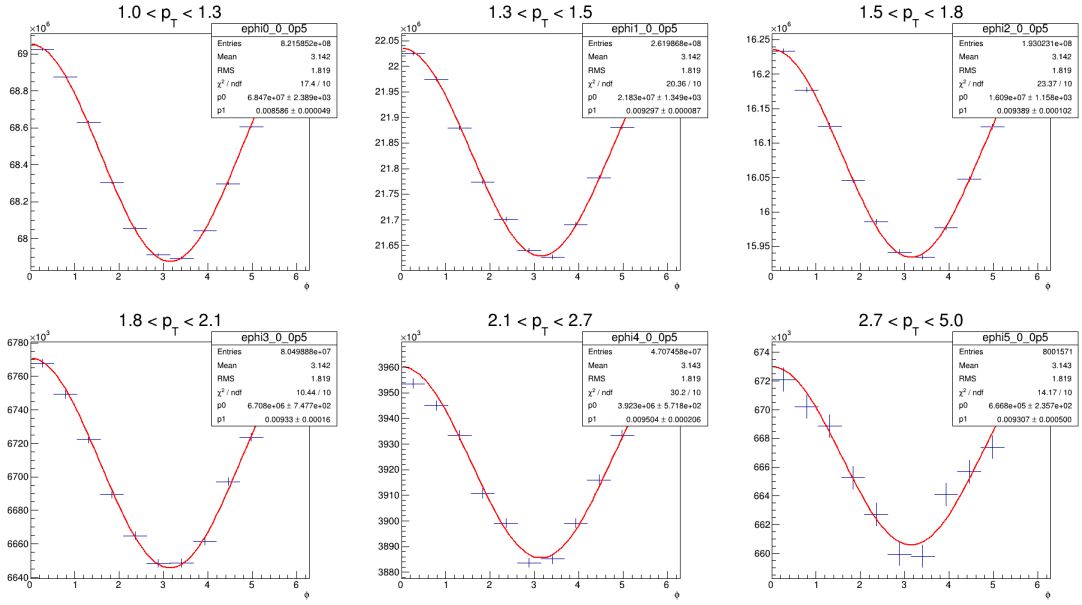

Figure 4.13: Fits of ϕ^{e^+} distributions using Equation 4.5, used to obtain $A_N(p_T)$ for positrons.

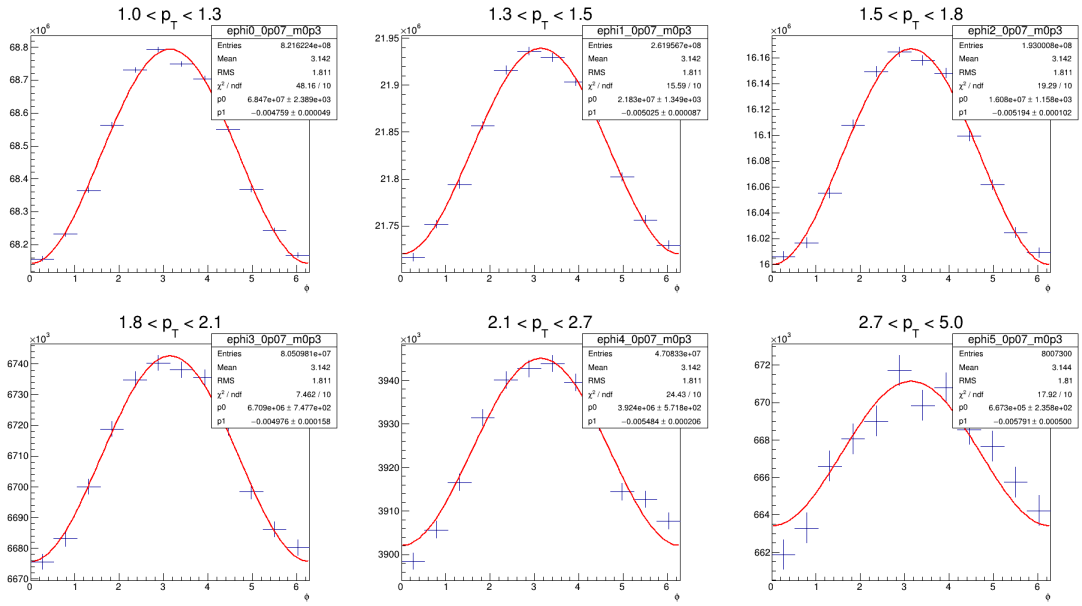

Figure 4.14: Fits of ϕ^{e^-} distributions using Equation 4.5, used to obtain $A_N(p_T)$ for electrons.

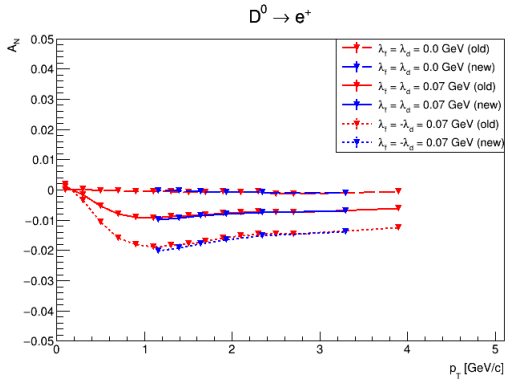

Figure 4.15: Demonstrated agreement between theory curves calculated using methods described in Sections 3.3.4.1 and 4.1.1 (red), and Section 4.1.1.1 (blue) for the parameter combinations chosen in Ref. [107] (positrons).

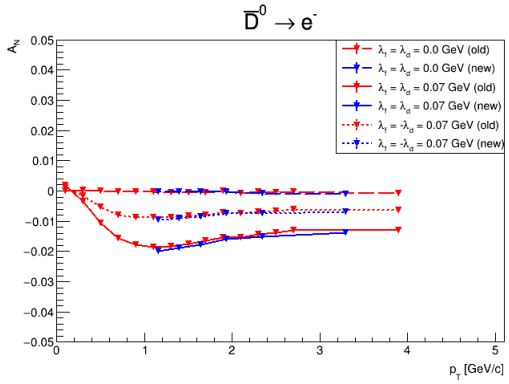

Figure 4.16: Demonstrated agreement between theory curves calculated using methods described in Sections 3.3.4.1 and 4.1.1 (red), and Section 4.1.1.1 (blue) for the parameter combinations chosen in Ref. [107] (electrons).

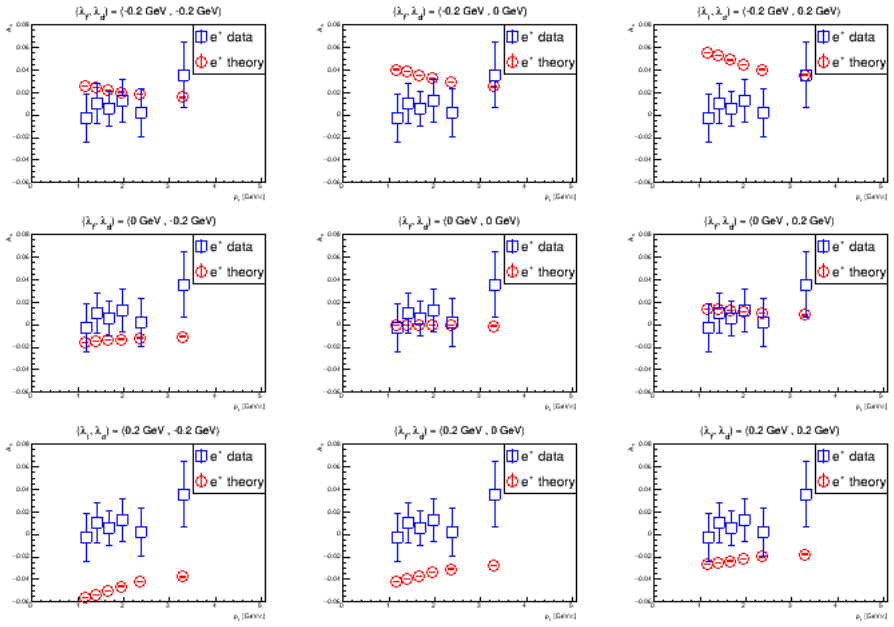

Figure 4.17: A subset of theory curves corresponding to various (λ_f, λ_d) combinations compared with the open heavy flavor positron A_N data points.

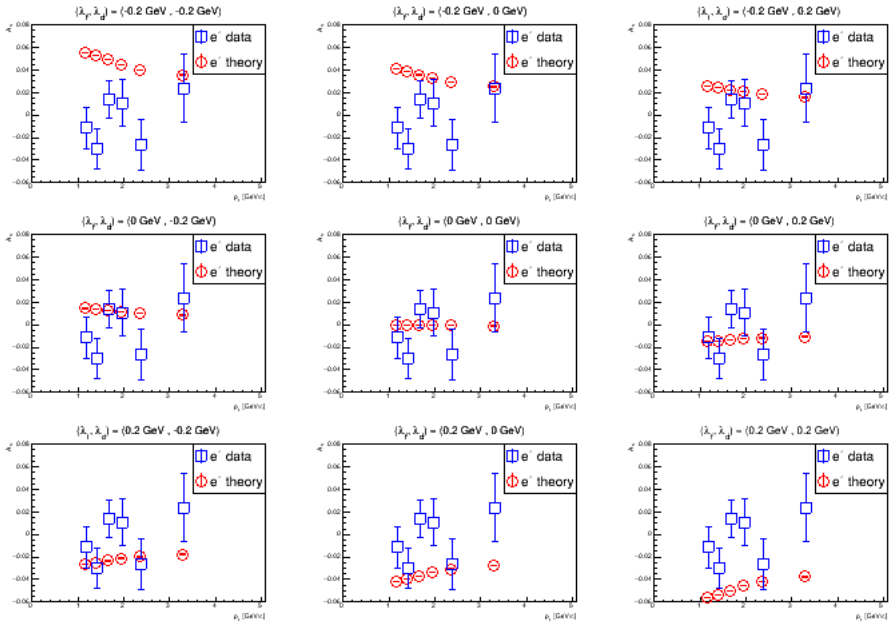

Figure 4.18: A subset of theory curves corresponding to various (λ_f, λ_d) combinations compared with the open heavy flavor electron A_N data points.

For each combination of (λ_f, λ_d), a χ^2 statistic is calculated, the minimum of

which is used to determine the best fit parameters with 1σ and 2σ confidence level (CL) bounds. $\chi^2(\lambda_f, \lambda_d)$ is calculated as shown in Equations 4.6 – 4.8.

$$\chi^2(\lambda_f, \lambda_d) = \chi^2_{(+)}(\lambda_f, \lambda_d) + \chi^2_{(-)}(\lambda_f, \lambda_d) \tag{4.6}$$

$$\chi^2_{(+)}(\lambda_f, \lambda_d) = \sum_i^{nbins} \frac{\left(A_N^{(+)data} - A_N^{(+)theory}(\lambda_f, \lambda_d)\right)^2}{\sigma^2_{(+)data}} \tag{4.7}$$

$$\chi^2_{(-)}(\lambda_f, \lambda_d) = \sum_i^{nbins} \frac{\left(A_N^{(-)data} - A_N^{(-)theory}(\lambda_f, \lambda_d)\right)^2}{\sigma^2_{(-)data}} \tag{4.8}$$

The corresponding $n\sigma$ CL region is determined with Equation 4.9,

$$\chi^2(\lambda_f, \lambda_d) - \chi^2_{min} < n^2. \tag{4.9}$$

$\sigma_{(+,-)data}$ in Equations 4.7 and 4.8 are taken as the measured statistical uncertainty. It was determined that adding the systematic uncertainty in quadrature had a negligible effect on the parameter constraints, as it typically only affected the third significant figure of the uncertainty, as shown in Tables 4.4 and 4.5. This validates the use of Equations 4.6 – 4.9, which are robustly defined in terms of statistical uncertainty.

Table 4.4: Comparison of σ_{stat} in each p_T bin to $\sqrt{\sigma^2_{stat} + \sigma^2_{sys}}$ for open heavy flavor positrons.

p_T range (GeV)	$\langle p_T \rangle$ (GeV)	σ_{stat}	$\sqrt{\sigma^2_{stat} + \sigma^2_{sys}}$
1.0 – 1.3	1.161	0.0212	0.0214
1.3 – 1.5	1.398	0.0178	0.0179
1.5 – 1.8	1.639	0.0159	0.0159
1.8 – 2.1	1.936	0.0192	0.0205
2.1 – 2.7	2.349	0.021	0.0215
2.7 – 5.0	3.290	0.0287	0.0295

Table 4.5: Comparison of σ_{stat} in each p_T bin to $\sqrt{\sigma_{stat}^2 + \sigma_{sys}^2}$ for open heavy flavor electrons.

p_T range (GeV)	$\langle p_T \rangle\,(GeV)$	σ_{stat}	$\sqrt{\sigma_{stat}^2 + \sigma_{sys}^2}$
1.0 – 1.3	1.161	0.0186	0.0192
1.3 – 1.5	1.398	0.0181	0.0188
1.5 – 1.8	1.639	0.0167	0.0168
1.8 – 2.1	1.936	0.0207	0.0208
2.1 – 2.7	2.349	0.0227	0.0229
2.7 – 5.0	3.290	0.0305	0.0309

Equation 4.6 and Equation 4.9 are ultimately used to determine the best-fit parameters and corresponding 1σ, 2σ and 3σ CL bounds, from which it was determined that the 0.05 GeV step size in λ parameter space was too coarse to reliably determine the confidence intervals. A higher resolution scan in (λ_f, λ_d) space was implemented to remedy this, zoomed in on the region of interest, with resolution of 0.01 GeV within -0.2 GeV $< \lambda_f <$ 0.2 GeV and -0.1 GeV $< \lambda_d <$ 0.3 GeV for a total of 41×41 parameter combinations. The same analysis was performed on the new theory curves to obtain best-fit parameters by determining the minimum $\chi^2(\lambda_f, \lambda_d)$ value from Eq. 4.6 in addition to 1σ, 2σ, and 3σ CL regions. The results of this procedure are shown in Figure 4.19. As seen in the middle row of Figures 4.17 and 4.18 (in particular the center most panel), the data seems to favor small contributions from the trigluon correlators. This can be achieved in two ways based on Equation 4.2 and 4.3 — with small normalization parameters (λ_f, λ_d), or with appropriately tuned combinations of (λ_f, λ_d) such that the contributions from the separate trigluon correlators cancel out. This is what is shown by the strip in parameter space corresponding to minimum $\Delta\chi^2_{e^{+/-}}(\lambda_f, \lambda_d)$ values as seen in panels (a) and (b) of Fig. 4.19, with the anti-correlated behavior for the electrons coming from the negative sign in front of

λ_d in Equation 4.3. These are the first data constraints of theoretical parameters λ_f and λ_d to date, with best fit values and 1σ confidence intervals of $\lambda_f = -0.01 \pm 0.03$ GeV and $\lambda_d = 0.11 \pm 0.09$ GeV, shown in panel (c) of Fig. 4.19.

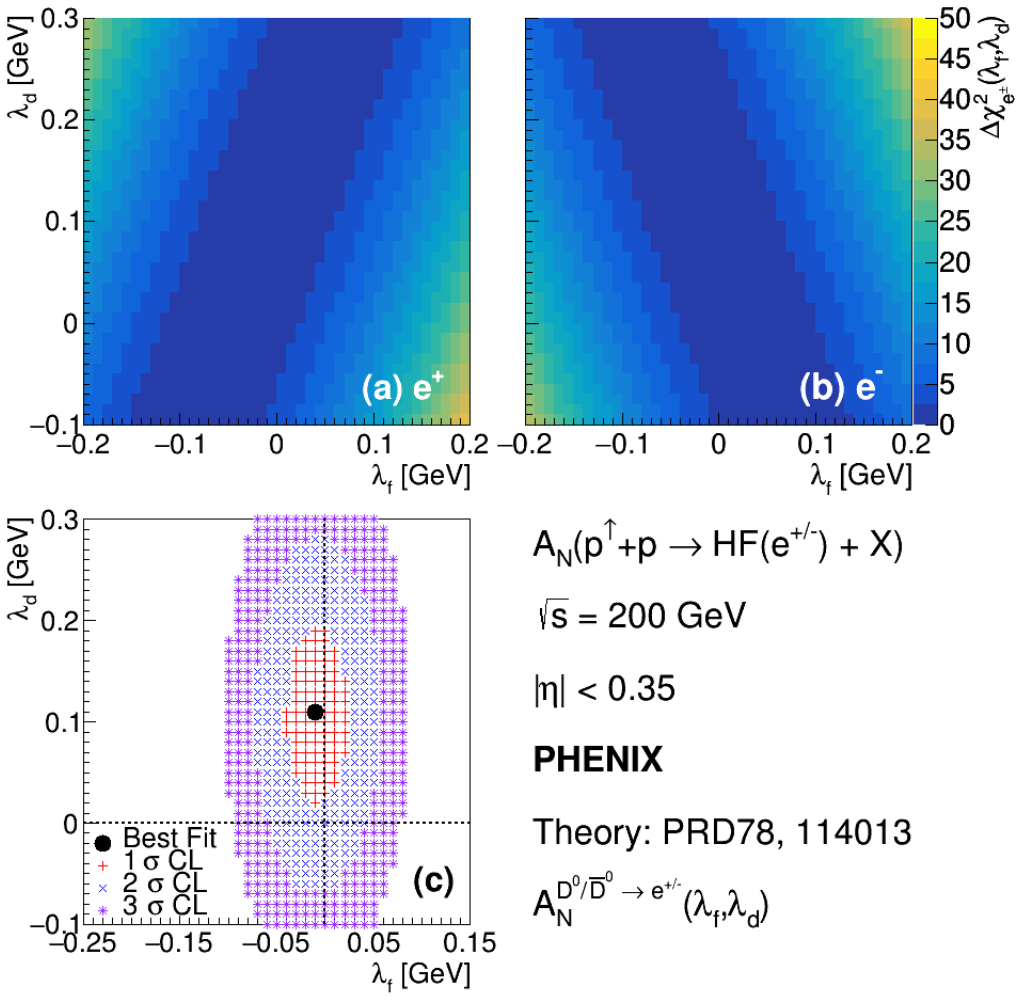

Figure 4.19: Results of the statistical analysis performed to extract best-fit parameters λ_f and λ_d by comparing data to theory [107]. $\chi^2(\lambda_f, \lambda_d) - \chi^2_{\text{min}}$ is shown for (a) e^+ and (b) e^-. Panel (c) shows the 1, 2, and 3σ confidence level regions, $\chi^2(\lambda_f, \lambda_d) - \chi^2_{\text{min}} < n^2$ $(n = 1, 2, 3)$. This figure was taken from Ref. [1].

4.1.1.2 Constraining Theoretical Parameters K_G and K'_G

In order to constrain theoretical parameters K_G and K'_G introduced in Ref. [87], the procedure from Section 4.1.1.1 is followed. In this case, the D meson asymmetries depend on only one parameter at a time for two separate models (model 1 with K_G, and model 2 with K'_G) describing how the trigluon correlation functions are normalized to the unpolarized gluon PDF. The theoretical asymmetries for a fixed value of K_G are shown in Figure 4.20, and for a fixed value of K'_G in Figure 4.21.

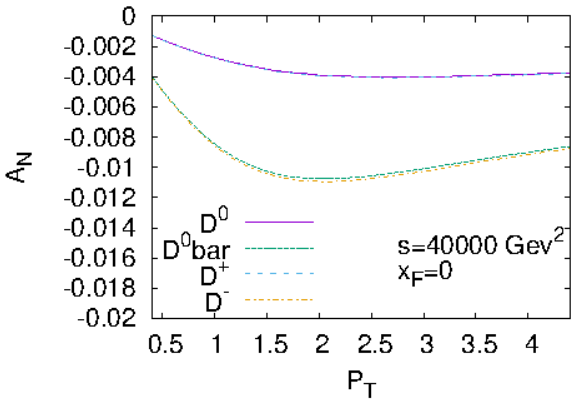

Figure 4.20: Theoretical predictions of A_N provided by the authors of Ref. [87] for $D^0, \bar{D}^0, D^+$, and D^- (model 1).

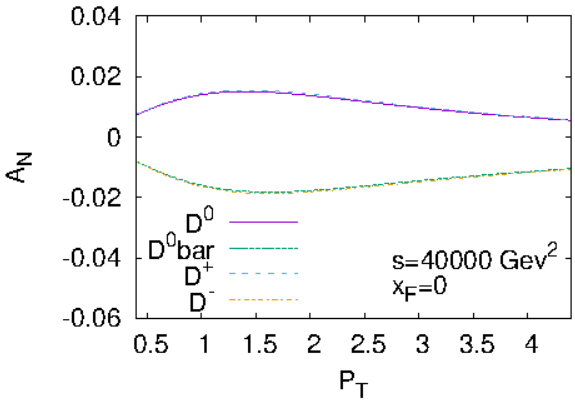

Figure 4.21: Theoretical predictions of A_N provided by the authors of Ref. [87] for $D^0, \bar{D}^0, D^+$, and D^- (model 2).

Given that asymmetries were provided for $D^0, \bar{D}^0, D^+$, and D^-, a procedure was implemented to combine appropriately scaled simulation samples of $D^0 \to e^+$ with $D^+ \to e^+$ and $\bar{D}^0 \to e^-$ with $D^- \to e^-$. The same $D^0 \to e^+$ and $\bar{D}^0 \to e^-$ samples discussed in Sections 4.1.1 and 4.1.1.1 were combined with newly generated $D^+ \to e^+$ and $D^- \to e^-$ samples to constrain K_G and K'_G. The fractions of f_{D^0} and f_{D^+} were determined in the simulated positron sample, and $f_{\bar{D}^0}$ and f_{D^-} in the simulated electron sample, where $f_{D^0} + f_{D^+} = f_{\bar{D}^0} + f_{D^-} = 1$. Figure 4.22 shows f_{D^0} for the positron sample (left panel) and $f_{\bar{D}^0}$ in the electron sample (right panel), where

$f_{D^+} = 1 - f_{D^0}$ and $f_{D^-} = 1 - f_{\bar{D}^0}$.

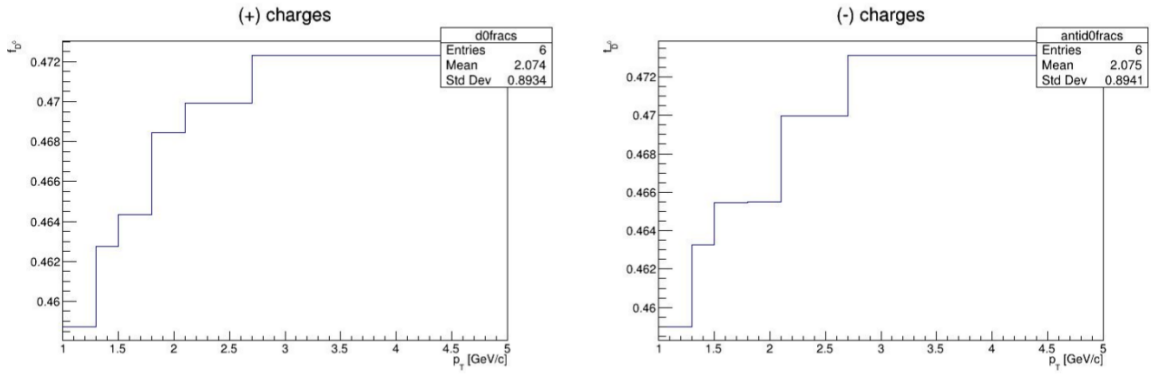

Figure 4.22: f_{D^0} in positron sample (left panel) and $f_{\bar{D}^0}$ in electron sample (right panel), where $f_{D^0} + f_{D^+} = f_{\bar{D}^0} + f_{D^-} = 1$.

The function and parameters used to weight the initially flat $D^{\pm}$ p_T spectra are shown in Equation 4.1 and Table 4.3, respectively. The parameters in Table 4.3 come from a fit to a $\sqrt{s} = 200$ GeV D^0 cross section measured in Ref. [161], used again here because there are no available cross section data for $\sqrt{s} = 200$ GeV $D^{\pm}$ production. The $D^{\pm} \to e^{\pm}$ p_T spectrum was scaled by a factor of 0.5 in accordance with the D^+/D^0 ratio measured by ALICE in $\sqrt{s} = 5.02$ TeV pp collisions [172]. Figure 4.23 depicts the D^+/D^0 production ratio measured in Ref. [172], showing consistency with 0.5 over a wide p_T range overlapping with that in Figures 4.20 and 4.21.

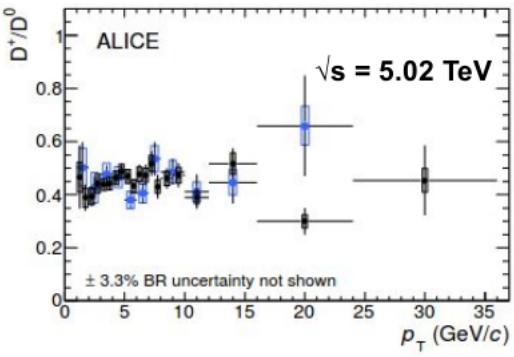

Figure 4.23: D^+/D^0 production ratio in $\sqrt{s} = 5.02$ TeV pp collisions measured in Ref. [172].

In order to calculate the fractions displayed in Figure 4.22, the $D^{\pm} \to e^{\pm}$ and the $D^0/\bar{D}^0 \to e^{\pm}$ p_T spectra are additionally scaled by their respective $e^{\pm} + X$ branching fractions (0.1607 for $D^{\pm}$ and 0.0649 for $D^0/\bar{D}^0$). The calculated fractions in Figure 4.22 are used to construct the theoretical input asymmetries to the χ^2 analysis via $A_N^{D\to e^+} = f_{D^0} A_N^{D^0\to e^+} + f_{D^+} A_N^{D^+\to e^+}$ (and charge conjugate), with $A_N^{D^0}$ and $A_N^{D^+}$ taken from Figures 4.20 and 4.21 for model 1 and model 2 respectively. Since the D meson asymmetries depend on only one parameter at a time, they can be written as follows,

$$A_N^{D,model1}(p_T) = K_G c_1(p_T) \tag{4.10}$$

$$A_N^{D,model2}(p_T) = K'_G c_2(p_T), \tag{4.11}$$

where c_1 and c_2 correspond to appropriate combinations of contributions from anti-symmetric and symmetric trigluon correlation functions. It is worth noting that both trigluon correlation functions depend on the same parameter in the two models (K_G for model 1 and K'_G for model 2), which was not the case in Section 4.1.1.1. Similarly to Section 4.1.1.1, independent scans were performed in K_G and K'_G parameter space with bounds $-0.005 < K_G < 0.005$ and step size of 0.0001 for K_G and bounds $-0.00025 < K'_G < 0.00075$ and step size of 0.00001 for K'_G (101 steps per parameter). Figures 4.25 and 4.24 show a subset of theory curves for various values of K_G (open black circles) and K'_G (black stars) compared with open heavy flavor positrons and electrons, respectively (blue squares).

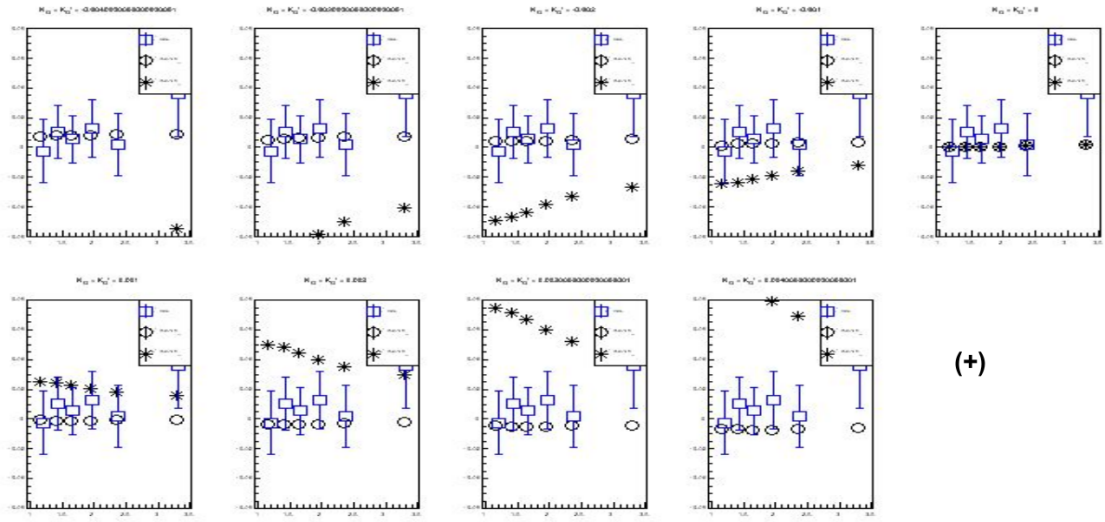

Figure 4.24: A subset of theory curves corresponding to various K_G values within $-0.005 < K_G < 0.005$ (black open circles) and K'_G values within $-0.00025 < K'_G < 0.00075$ (black stars) compared with the open heavy flavor positron A_N data points (blue squares) – e^+.

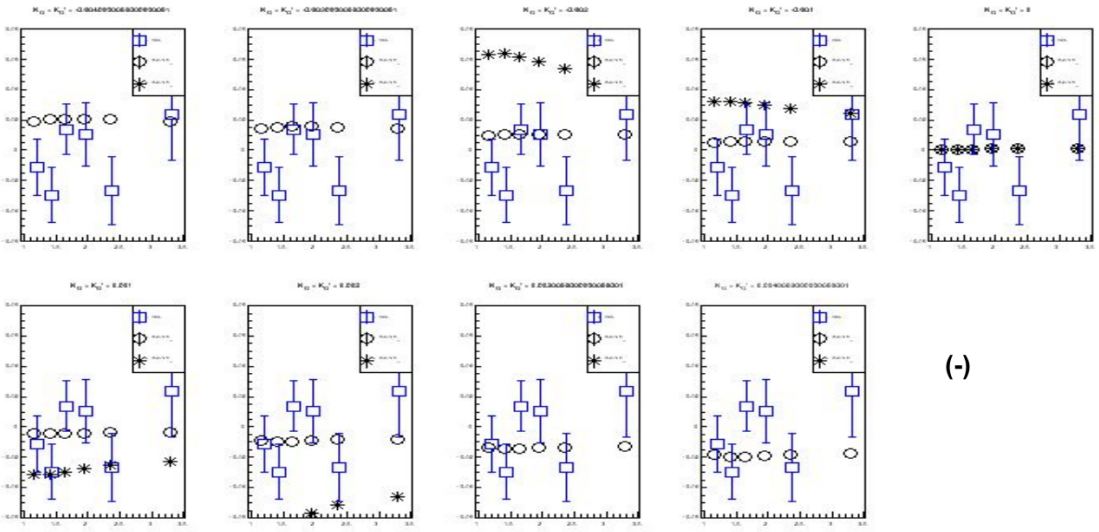

Figure 4.25: A subset of theory curves corresponding to various K_G values within $-0.005 < K_G < 0.005$ (black open circles) and K'_G values within $-0.00025 < K'_G < 0.00075$ (black stars) compared with the open heavy flavor electron A_N data points (blue squares) – e^-.

For each value of K_G and K'_G, a χ^2 statistic is calculated, the minimum of which is used to determine the best-fit parameters with 1σ and 2σ confidence intervals. $\chi^2(K_G)$ and $\chi^2(K'_G)$ are calculated as shown in Equations 4.12 – 4.14.

$$\chi^2(K_G||K'_G) = \chi^2_{(+)}(K_G||K'_G) + \chi^2_{(-)}(K_G||K'_G) \tag{4.12}$$

$$\chi^2_{(+)}(K_G||K'_G) = \sum_i^{nbins} \frac{\left(A_N^{(+)data} - A_N^{(+)theory}(K_G||K'_G)\right)^2}{\sigma^2_{(+)data}} \tag{4.13}$$

$$\chi^2_{(-)}(K_G||K'_G) = \sum_i^{nbins} \frac{\left(A_N^{(-)data} - A_N^{(-)theory}(K_G||K'_G)\right)^2}{\sigma^2_{(-)data}} \tag{4.14}$$

The corresponding $n\sigma$ confidence interval is calculated with Equation 4.15

$$\chi^2(K_G||K'_G) - \chi^2_{min} < n^2. \tag{4.15}$$

The || symbol in this context refers to an "or" operator, denoting that the statistic is calculated separately for each parameter. Just as in Section 4.1.1.1, $\sigma_{(+,-)data}$ in Equations 4.13 and 4.14 are taken as the measured statistical uncertainty of A_N for the separate charges.

Figure 4.26 shows the confidence intervals and best-fit parameters for K_G and K'_G calculated using Equation 4.15. The best-fit parameters are shown by the black dots, while the 1σ and 2σ confidence intervals are shown by blue xs and red crosses respectively. The resulting best fit parameters and 1σ confidence intervals are $K_G = 0.0006^{+0.0014}_{-0.0017}$ and $K'_G = 0.00025 \pm 0.00022$, consistent with the analysis conducted in Ref. [87].

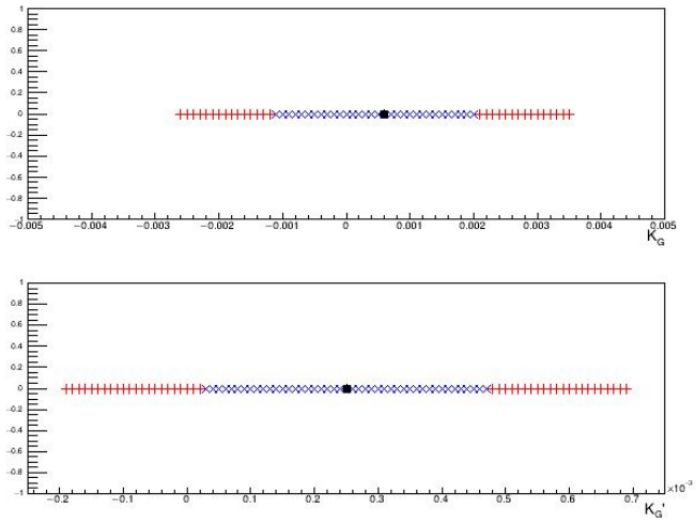

Figure 4.26: Best fit parameters $K_G = 0.0006^{+0.0014}_{-0.0017}$ and $K'_G = 0.00025 \pm 0.00022$ (black dots) corresponding to χ^2_{min} plotted alongside 1σ (blue x) and 2σ (red cross) confidence intervals calculated using Equation 4.15. A step size of 0.0001 was used for K_G, and 0.00001 for K'_G.

4.1.1.3 Final Results with Best Fit Parameters

The open heavy flavor electron and positron TSSAs are plotted again in Figure 4.27 along with the theoretical curves the best fit parameters $(\lambda_f, \lambda_d), K_G$, and K'_G obtained from the analysis discussed in Sections 4.1.1.1 and 4.1.1.2. The electron results are plotted as blue squares while the positron results are plotted as red circles. The bars show the statistical uncertainties while the shaded boxes depict the systematic uncertainties, which are roughly an order of magnitude smaller than the statistical uncertainties in most bins. The results are consistent with zero within uncertainties, as well as the theoretical curves with the best fit parameters. The legend shows the values of the best fit parameters, along with 1σ confidence intervals taken from Figures 4.19 and 4.26.

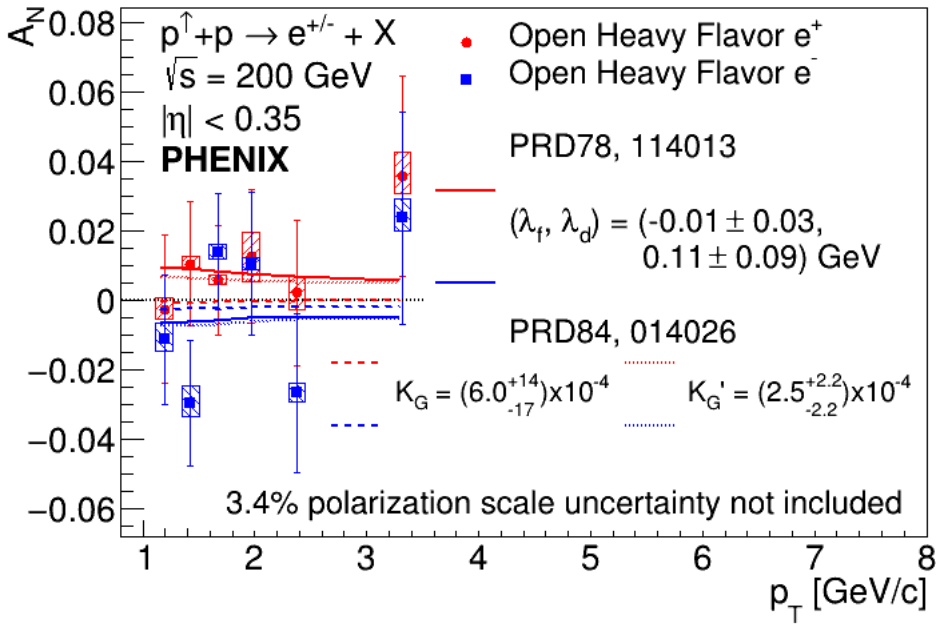

Figure 4.27: $A_N(OHF \rightarrow e^{\pm})$ (red) circles and (blue) squares for positrons and electrons, respectively. Also plotted are predictions of $A_N(D^0/\bar{D}^0 \rightarrow e^{\pm})$ from Ref. [107], and $A_N((D^0/\bar{D}^0 + D^{+/-}) \rightarrow e^{\pm})$ from Ref. [87] for best-fit trigluon-correlator-normalization parameters, with the red/blue solid, dashed, and dotted lines corresponding to central values of the 1σ confidence intervals shown in the legend. This figure was taken from Ref. [1].

4.1.2 Discussion

Before this measurement was published, the trigluon correlation functions were relatively unconstrained, with only few measurements capable of providing constraints [106, 123, 165, 173–177]. This includes a combination of TSSAs in jets [106, 173, 174, 176], direct photon [165], light hadron [123], and heavy flavor hadron [177] production channels. Jets, direct photons, and light hadrons, all involve mixtures of quark and gluon processes in the initial state, as shown in Figures 4.28, 4.29, and 4.30, respectively. On the other hand, heavy flavor production is dominantly produced via gluon-gluon fusion, as shown in Figure 4.31, where the red curves represent the

summed production channels and the yellow curve is for gg fusion. Direct photon and jet measurements offer cleaner constraints on initial state dynamics than light hadrons since effects from final state fragmentation are limited. In the former case, this is due to the photon being a color-neutral object that does not interact with color fields at leading order, while in the latter case, it is due to jets being a proxy for the hard scattered or fragmenting parton. Heavy flavor hadron production, however, is dominated by the gluon-gluon fusion production channel, receiving very little contribution from quark-antiquark annihilation. In addition, the mechanism by which spin-momentum correlations are generated in the final state is prohibited for gluons in spin 1/2 nucleons, as it requires the fragmentation of a transversely polarized parton. This makes heavy flavor production an even cleaner probe of initial-state dynamics given the dominant gluon-gluon fusion production mechanism.

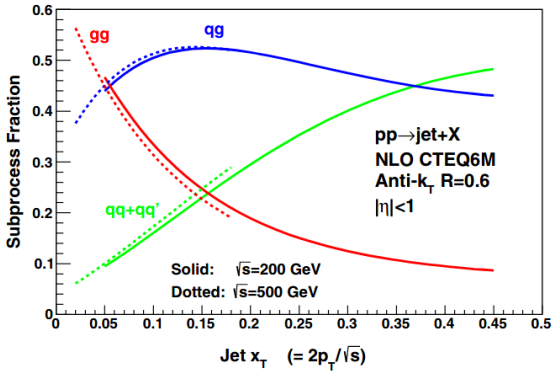

Figure 4.28: Fractional contribution of partonic processes contributing to midrapidity jet production in $\sqrt{s} = 200$ and 500 GeV pp collisions as a function of jet $x_T = 2p_T/\sqrt{s}$.

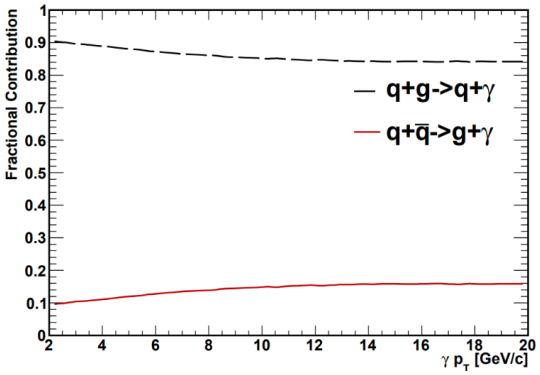

Figure 4.29: Fractional contribution of partonic processes contributing to direct photon production in $\sqrt{s} = 200$ GeV pp collisions as a function of p_T.

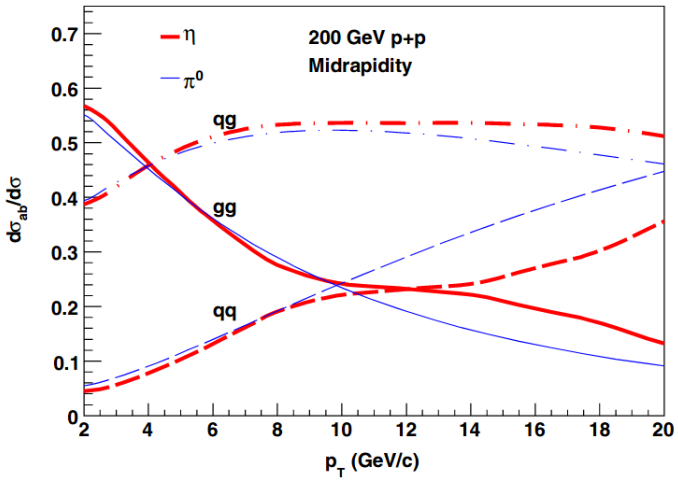

Figure 4.30: Fractional contribution of partonic processes contributing to midrapidity π^0 and η production in $\sqrt{s} = 200$ GeV pp collisions as a function of p_T.

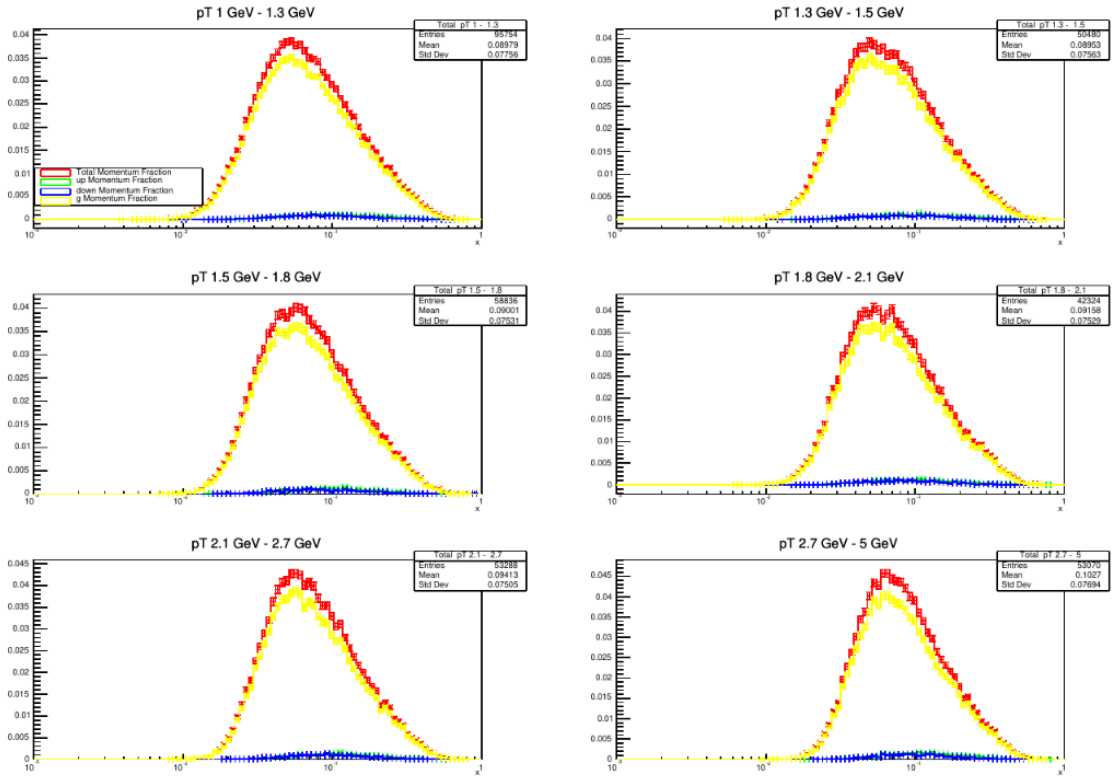

Figure 4.31: Fractional contribution of partonic processes contributing to midrapidity open heavy flavor electron production in $\sqrt{s} = 200$ GeV pp collisions as a function of collinear momentum fraction x in the various measured p_T bins.

The heavy flavor muon TSSA measured by PHENIX [177] has complementary kinematic coverage to this measurement but was performed on a lower luminosity data set (the 2012 running year). The heavy flavor electron measurement is the most up-to-date measurement of a TSSA in the heavy flavor production channel, and the first measurement to provide constraints on theoretical parameters λ_f and λ_d introduced in Ref. [107]. The consensus from the data so far is the trigluon correlations in polarized protons are small, which is corroborated by Figure 4.1.1.1, showing that $(\lambda_f, \lambda_d) = (0, 0)$ GeV lies just outside of the 1σ confidence level region. To this date, there are no global extractions of the trigluon correlation functions or the gluon Sivers PDF. The data presented in this dissertation is an important step towards such a goal, leading to a better understanding of the gluonic structure in transversely polarized protons.

4.2 Neutral Meson TSSAs

The final results for the π^0 TSSAs in pA collisions are shown in panel (a) of Figure 4.32, while the results for η TSSAs in pA collisions are shown in panel (b). In each panel, pAu results are shown as blue circles, pAl results are shown as green squares, and the pp results from Ref. [123] are shown as black diamonds. The statistical uncertainties are shown by the bars attached to each point while the systematic uncertainties are shown as shaded boxes. There are no dominant sources of systematic uncertainties across all p_T bins. The results are consistent with zero for both π^0 and η in both collision systems, as well as the corresponding measurements from the pp dataset, implying that no nuclear dependence was observed. Table 4.6 summarizes the results for the π^0 asymmetries while Table 4.7 summarizes the results for the η asymmetries.

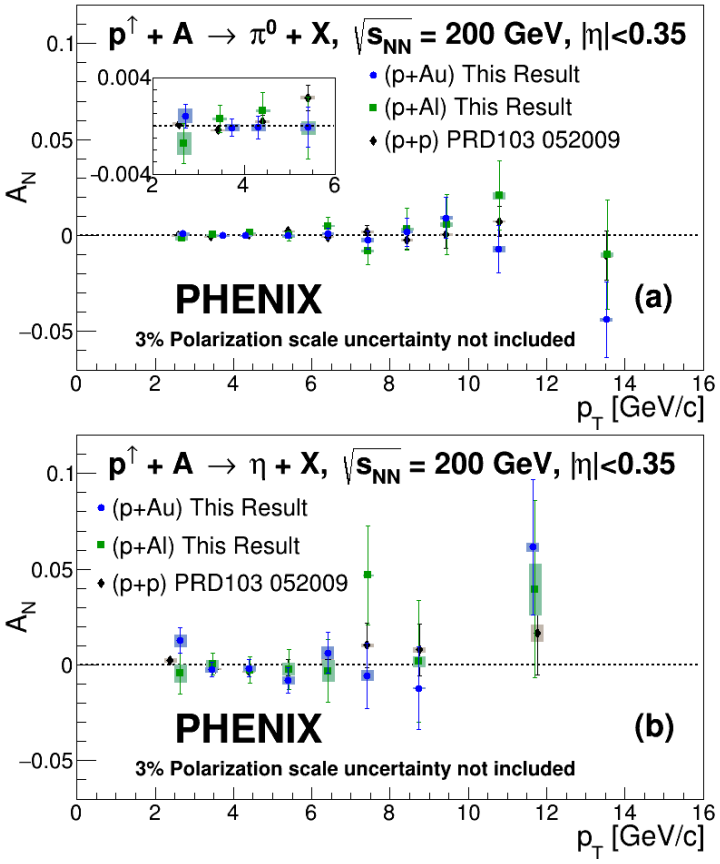

Figure 4.32: Transverse single-spin asymmetry for (a) π^0 and (b) η mesons in $p^{\uparrow}$Au collisions (blue circles), and $p^{\uparrow}$Al collisions (green squares) from this measurement, shown alongside the same measurement in polarized pp collisions from Ref. [123] (black diamonds). The error bars represent the statistical uncertainty (σ^{stat}) while the boxes represent the total systematic uncertainty (σ^{syst}). This figure was taken from Ref. [2].

Table 4.6: Summary of final asymmetries with statistical and systematic uncertainties for π^0 mesons in $p^{\uparrow}A$ collisions. σ^{syst} corresponds to the systematic uncertainties, displayed by the shaded boxes in Fig. 4.32. This table was taken from Ref. [2].

Collisions	p_T range [GeV/c]	$\langle p_T \rangle$ [GeV/c]	A_N	σ^{stat}	σ^{syst}
$p^{\uparrow}$Au	2–3	2.71	0.000818	0.000993	0.000569
	3–4	3.73	-0.000145	0.000701	0.000286
	4–5	4.31	-0.000135	0.000974	0.000257
	5–6	5.4	-0.00011	0.00164	0.00028
	6–7	6.41	0.00097	0.00281	0.00024
	7–8	7.42	-0.00243	0.00464	0.00109
	8–9	8.43	0.00179	0.00732	0.00055
	9–10	9.44	0.0093	0.0106	0.0005
	10–12	10.8	-0.0072	0.0122	0.0014
	12–20	13.5	-0.0438	0.0198	0.0008
$p^{\uparrow}$Al	2–3	2.67	-0.00147	0.00163	0.00088
	3–4	3.47	0.00056	0.00113	0.00006
	4–5	4.41	0.00126	0.00153	0.00005
	5–6	5.41	-0.00018	0.00254	0.00051
	6–7	6.42	0.00500	0.00429	0.00042
	7–8	7.42	-0.00809	0.00699	0.00060
	8–9	8.43	0.0035	0.0109	0.0004
	9–10	9.44	0.0058	0.0155	0.0010
	10–12	10.8	0.0208	0.0181	0.0016
	12–20	13.6	-0.0099	0.0286	0.0011

Table 4.7: Summary of final asymmetries with statistical and systematic uncertainties for η mesons in $p^{\uparrow}A$ collisions. σ^{syst} corresponds to the systematic uncertainties, displayed by the shaded boxes in Fig. 4.32. This table was taken from Ref. [2].

Collisions	p_T range [GeV/c]	$\langle p_T \rangle$ [GeV/c]	A_N	σ^{stat}	σ^{syst}
$p^{\uparrow}$Au	2–3	2.64	0.01279	0.00665	0.00306
	3–4	3.44	-0.00255	0.00377	0.00115
	4–5	4.41	-0.00168	0.00448	0.00117
	5–6	5.4	-0.00810	0.00667	0.00183
	6–7	6.41	0.0064	0.0108	0.0029
	7–8	7.42	-0.0056	0.0170	0.0026
	8–10	8.74	-0.0122	0.0216	0.0002
	10–20	11.7	0.0615	0.0351	0.0021
$p^{\uparrow}$Al	2–3	2.64	-0.0044	0.0107	0.0046
	3–4	3.46	0.00043	0.00575	0.00219
	4–5	4.42	-0.00278	0.00686	0.00050
	5–6	5.41	-0.0022	0.0104	0.0031
	6–7	6.42	-0.0032	0.0163	0.0055
	7–8	7.42	0.0468	0.0260	0.0004
	8–10	8.74	0.0017	0.0318	0.0027
	10–20	11.7	0.0395	0.0464	0.0133

4.2.1 Discussion

The 2015 RHIC dataset is the first and only collider dataset of $p^{\uparrow}A$ collisions, with only a handful of corresponding TSSA measurements to date [2, 178–183]. These data provide an opportunity to study how the partonic origins of the TSSA are

modified due to additional nuclear matter, stronger color fields, and higher particle multiplicities present in pA collisions. Currently, these measurements paint an inconsistent picture and warrant further investigation. Charged hadron TSSAs at intermediate rapidity reported by the PHENIX experiment show strong nuclear dependence [178, 179], while neutral pion TSSAs reported by the STAR experiment at farther forward rapidity indicate only moderate nuclear modification [181]. It should be noted that sizable asymmetries were observed in pp collisions in both cases. PHENIX also reported π^0 asymmetries in pA collisions (as well as η mesons) [2], but at midrapidity, in contrast to the forward rapidity measurement from STAR. Contrary to the observations at forward and intermediate rapidities, no nuclear modification was measured, and both the pp and pA measurements are consistent with zero. This measurement was a key component of this dissertation work and should prove useful in resolving the mystery of nuclear modification of TSSAs at different rapidity regions for light hadron production channels. Reference [179] specifically investigates the nuclear dependence of TSSAs for positively charged hadrons at intermediate rapidity from PHENIX. It is tempting to try to understand the observed A dependence of the TSSAs in terms of gluon saturation effects, but it should be noted that the average transverse momentum of the measured hadrons is $\langle p_T \rangle = 2.9$ GeV, well above the saturation scale in gold nuclei as predicted in Ref. [120] ($Q_s^{\mathrm{Au}} \approx 0.9$ GeV). The A dependence is apparent nonetheless but lacks theoretical understanding. The J/ψ TSSAs at intermediate rapidity reported by PHENIX seem to agree between pp and pAl collisions, but further investigation is needed for pAu results at low p_T [180]. There currently does not exist a theoretical framework that can explain all of these data. Factorization implies that the presence of a heavy nucleus in the collision system cannot affect the initial-state spin-momentum correlations in polarized protons, but the final-state spin-momentum correlations in the process of hadronization can be modified due to fragmentation in nuclear media. On the other

hand, considering factorization-breaking effects can lead to more complex modifications of the asymmetry from interactions with the nuclear remnant [184, 185]. A more complete theoretical understanding exists for the forward neutron asymmetries reported by PHENIX in Refs.[182, 183], where the nuclear dependencies of the TSSAs are understood to be due to the interplay of hadronic interactions such as one pion exchange and electromagnetic interactions in ultra peripheral collisions.

CHAPTER V

Open Data in High Energy Physics

Maximizing the scientific output from data recorded in high-energy physics experiments is crucial for the longevity of the field. This can be achieved in part by facilitating open access to and preservation of the recorded data. Since the establishment of the Data Preservation in High Energy Physics (DPHEP) collaboration, there have been substantial efforts to implement these practices in both currently and previously running HEP experiments [186]. For example, CERN has recently created an open data policy [187] followed by a more detailed open science policy [188] that outlines the commitment of CERN to make data collected at the LHC available at various levels of complexity, as established by the DPHEP collaboration [186]. These are defined below.

Level 1 Published results — this can include tables and figures but also preprocessed Ntuples or binned and unbinned fit likelihood functions.

Level 2 Outreach and education — usually in the form of highly preprocessed Ntuples.

Level 3 Reconstructed data — these data have been preprocessed to derive physics objects, such as charged particle candidates, photons, or particle jets. Reconstructed data may or may not be corrected for detector effects, such as efficiency and resolution.

Level 4 Raw data – the basic quantities recorded by the experimental instruments.

While each level of complexity is very interesting and important in its own right, providing open access to data for research purposes (i.e. level 3 data) is the focus of the work presented in this dissertation. The goal is to create an open data infrastructure that ensures the data is Findable, Accessible, Interoperable, and Reusable [189] for external researchers, in particular theorists and phenomenologists. Changes in policy are a crucial first step in accomplishing this, but they must be followed by organizational structures in the respective experiments to ensure the policy is adhered to. This takes a substantial amount of work, both from a bureaucratic and technical perspective, but there have been successful implementations of open access and data preservation in several experiments. Furthermore, this is a field of active development, with many interesting projects on the horizon, including the LHCb Ntuple Wizard [3], an important component of my dissertation work.

5.1 Successful Implementations

A number of HEP experiments have released a substantial amount of open data to the public domain, including currently running experiments such as CMS and LHCb at the LHC. In addition, previously running experiments such as H1, ZEUS, and HERMES at HERA have successfully implemented long-term data preservation to allow for continued analysis of the unique data collected years after HERA operation ceased.

A particularly successful example of open data releases is given by the CMS experiment [190]. CMS has been continually releasing open data since 2014, and they have an active community of users that have performed a number of studies on the data, a few of which can be found in Refs. [39, 191–198]. They have now released all of 2010-2012 *pp* data, as well as the 2010-2011 heavy ion data, and most of the 2015

pp data. Part of the reason for the success of CMS open data and the active user base is the yearly open data workshops held by the collaboration targeted toward scientists familiar with collider experiments [199]. An overview of CMS open data activities, including available tools, current challenges, and planned improvements can be found in Ref. [200].

In more recent years, the LHCb experiment [201] has made substantial efforts to release level 3 data to the public. This includes the first public release of LHCb Run 1 data (2011-2012) at the end of 2022, and efforts to develop a novel tool known as the LHCb Ntuple Wizard [3] to provide access to both Run 1 and Run 2 data, discussed in Section 5.2. The first Run 1 release amounted to about 20% of the Run 1 data, a roughly 200 TB volume now accessible via the CERN Open Data Portal [202]. The remaining 80% was initially blocked from release due to internal policy that has now been overcome, and there are active efforts ongoing to copy and host all of the Run 1 data (1 PB) to the open data portal. The open data team on LHCb is very small, making it difficult to provide the same level of resources in the way of user support as CMS. However, the LHCb Starterkit is publicly accessible and provides detailed explanations of the preliminary software tools and necessary steps to perform LHCb data analysis [203]. In addition, the LHCb software environment is accessible to the public, so the resources to analyze LHCb data are available aside from the computing infrastructure. The issues related to the large data volumes, the high learning curve for using LHCb software, and the lack of computing infrastructure are all addressed with the LHCb Ntuple Wizard discussed in Section 5.2, currently under active development.

The H1 [204], ZEUS [205], and HERMES [206] experiments at the formerly running HERA collider [207] (1992-2007) made a collaborative data preservation effort after the end of HERA operation [208–210]. In fact, members of the HERA experiments make up many of the core coauthors of the DPHEP study group document

produced in 2012 [186]. This collaborative effort resulted in an archive containing online data volumes amounting to 359 TiB from H1, 239 TiB from ZEUS, and 57 TiB from HERMES, with a roughly equal amount on tape, for the long-term preservation of HERA data. This, and ongoing efforts to preserve software environments in the respective experiments, facilitate continued analysis of this unique data many years after data taking stopped.

5.2 The LHCb Ntuple Wizard

The LHCb Ntuple Wizard is a novel application that allows external users to make specific queries of LHCb data via an intuitive web interface [3]. The queries are processed with the existing LHCb computing infrastructure resulting in Ntuples delivered to the CERN Open Data Portal [187]. Although LHCb has successfully released much of the collected Run 1 data (2011-2012) (discussed in Section 5.1), the procedure to do so is not scalable for Run 2 and 3 data due to the large volume of reconstructed (level 3) data collected by LHCb. This is illustrated in Figure 5.1, showing data volumes for Run 1+2 and projected Run 3 data volumes for the 4 main experiments at the LHC. The data volume depends on both the trigger rate and the average event size for the respective experiments. The large data volumes at LHCb can be understood mainly due to differences in trigger rate. The trigger rate is about 1 kHz for ATLAS [211] and CMS [212], while it is closer to 10 kHz for LHCb [213]. The large data volumes for Run 2 (and projected data volumes for Run 3) prompted the development of a new method to expose LHCb Run 2 data (and in the future Run 3 data) to external users. In addition, the barrier of entry for LHCb data analysis is high for external analysts, due to the requirement of learning and using the LHCb software environment to produce Ntuples from the reconstructed data. This step is effectively bypassed by the Ntuple Wizard, as the configuration of the Ntuples is handled through an intuitive web interface rather than a custom

software environment.

	ALICE	ATLAS	CMS	LHCb
Run-2	2 PB	0.5 PB	2 PB	10 PB (including Run-1)
Run-3	4 PB	1 PB	4 PB	45 PB
Total	**6 PB**	**1.5 PB**	**6 PB**	**55 PB**

Figure 5.1: Table of level 3 data volumes for the experiments at the LHC in Run 1 + 2, and projected level 3 data volumes for Run 3.

A flowchart of the Ntuple Wizard is shown in Figure 5.4, depicting the key steps of the application. These steps are outlined in detail in Sections 5.2.3 - 5.2.5. However, a proper understanding of the Ntuple Wizard architecture and functionality first necessitates a discussion of the LHCb data flow pipeline.

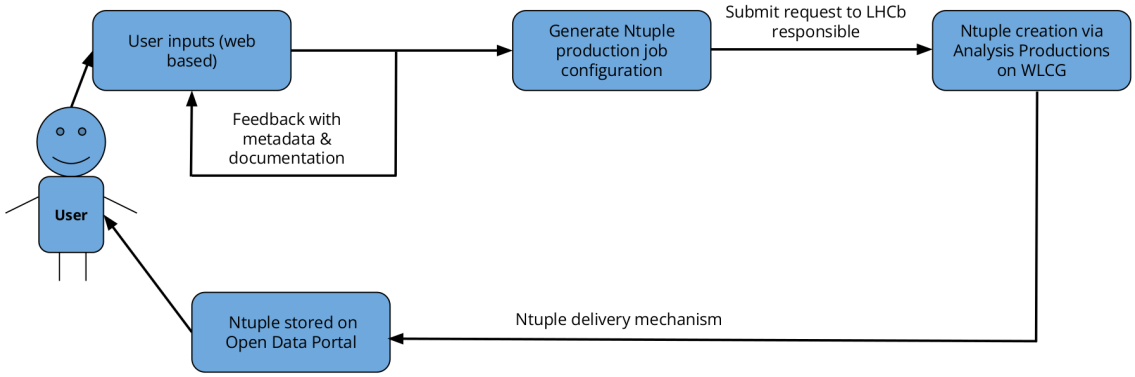

Figure 5.2: A logical flow chart depicting the various steps of the Ntuple Wizard.

5.2.1 The LHCb Data Flow Pipeline

A schematic of the LHCb data flow pipeline is depicted in Figure 5.3, showing a number of custom LHCb software tools for the various levels of data processing.

During the processing, raw data such as hits and energy deposits in various detector subsystems are converted into physics objects like charged particle trajectories, jets, and vertices during the *reconstruction* step, and filtered into subsamples enriched in particular physics signatures in the *stripping* step. Additional details about the LHCb data flow pipeline can be found in the LHCb computing technical design reports [214, 215].

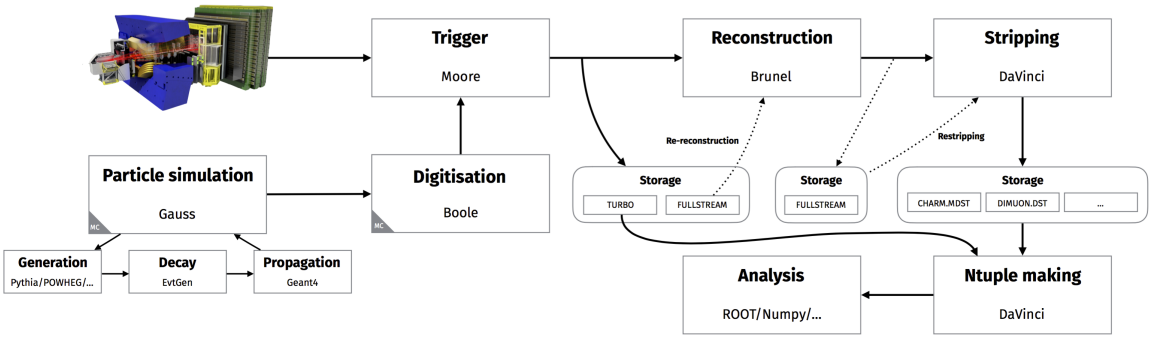

Figure 5.3: LHCb data flow in Runs 1 and 2. This figure was taken from Ref. [3].

The stripping step comprises collections of algorithms known as *stripping lines* that are applied to the reconstructed data to yield filtered subsets identified to have related physics signatures, known as *streams*. The motivation behind splitting the data into streams is to speed up the data processing pipeline for analyses by decreasing the number of files a user needs to process. The collection of events contained in the various streams is in general not disjoint.

Much of the interest of LHCb lies in analyzing the decays of heavy flavor hadrons. As such, the target physics signatures for many stripping lines are particle decays, for example, $B^+ \to \bar{D}^0\pi^+$ and subsequent decay $\bar{D}^0 \to K^+\pi^-$. The particles B, D, K, and π are the lightest hadrons containing b, c, s and u/d quarks respectively. A useful concept to aid in understanding and visualizing such cascading decays are graphical structures known as trees. A tree representing a particle decay is called a *decay tree*, with each particle in the decay corresponding to a node in the tree, and each edge

(or branch) corresponding to a parent-child relationship. Stripping line selections are designed to identify events containing sets of physics objects matching the desired decay tree, including any selection criteria placed on various nodes (i.e. particles). Some example criteria include selection requirements on kinematic variables, particle identification variables, and track or vertex reconstruction qualities. Collections of candidates meeting these criteria are saved in an LHCb-specific data format.

All steps up to stripping and storage in Fig. 5.3 are handled centrally by the experiment. The remaining steps *Ntuple making* and *Analysis* depend on user inputs. Typically, an LHCb user interfaces with the reconstructed data by processing files belonging to a particular stream, and selecting on a particular stripping line or other reconstructed physics objects. During the Ntuple making step, the data is converted into a software-framework-independent data structure known as an Ntuple, where each candidate is encoded by a tuple of measured or calculated variables. Ntuples contain only basic data types, or in some instances, arrays of basic data types, and are saved in ROOT files [216].

An LHCb application known as DaVinci [214, 215, 217] handles both the stripping and the Ntuple making step, as shown in Fig. 5.3. It is an event selection and data analysis software package built using the Gaudi framework [214, 215, 218] that provides access to various algorithms that can be combined in sequence to select, process, and analyze events. The particular algorithm used for Ntuple configuration is called *DecayTreeTuple*. Variables are registered to DecayTreeTuple objects through the use of LHCb specific tools and functions known as *TupleTools*. TupleTools correspond to a set of variables that fall under a particular category that can be added to the Ntuple, either by applying the tool to the entire decay tree, or just to particular nodes. Multiple DecayTreeTuple objects can be configured in a single DaVinci session, and the decay tree structured output of many of the stripping lines is readily converted into a DecayTreeTuple object. The DaVinci application is configured using Python [219]

scripts that specify the sequence of algorithms applied during data processing, as well as the output file content.

5.2.2 Architecture

The architecture of the Ntuple Wizard is depicted in Fig. 5.4 as a directed acyclic graph highlighting the key components and how they are interconnected. The three main components of the core application are (*i*) the metadata and documentation generated during deployment, displayed to the user as needed in the web interface, (*ii*) the web interface, and (*iii*) the configuration files generated by the web interface. Component (*i*) is discussed in Section 5.2.3, while (*ii*) and (*iii*) are discussed in Section 5.2.4.

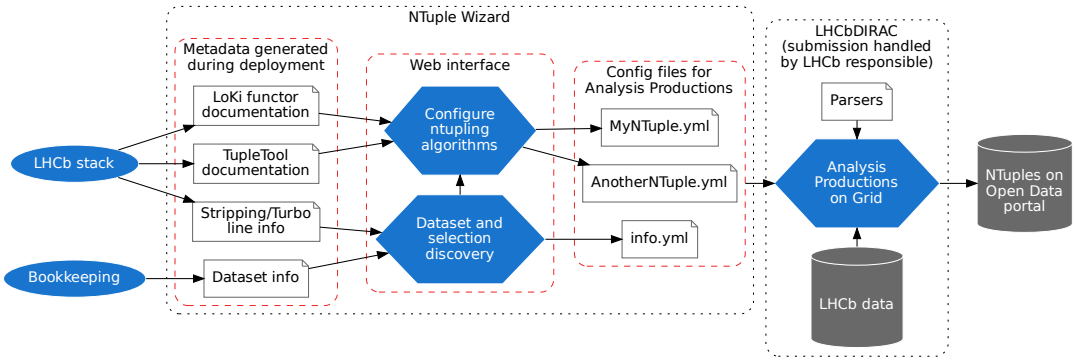

Figure 5.4: Architecture of the Ntuple Wizard. This figure was taken from Ref. [3].

5.2.2.1 Security Considerations

As discussed in Section 5.2.1, the DaVinci application is configured via Python files. However, accepting arbitrary Python code from external users presents obvious security risks. This motivates capturing the configuration submitted by users in a pure data-structure format. YAML files were chosen to satisfy this need, as shown in Fig. 5.4, which are parsed internally to produce the necessary Python scripts to steer

the DaVinci application.

5.2.3 Metadata and Documentation

Metadata and documentation of available tools and datasets are necessary to facilitate the core function of the Ntuple Wizard — generating configuration for the DaVinci application and Analysis Productions batch processing system [220, 221]. This requires information on available datasets, stripping lines, and TupleTools (introduced in Section 5.2.1). This information comes from several sources, grouped into two categories, (*i*) the LHCb software stack, and (*ii*) the LHCb bookkeeping system (i.e. database). Figure 5.4 outlines the information extracted from these sources and how it is propagated to the web interface.

Metadata from the LHCb software stack is acquired both directly from the DaVinci application and indirectly through the LHCb Doxygen pages. Doxygen is a standard tool for generating documentation from annotated source code [222]. Before the development of the Ntuple Wizard, the automatic pipeline for generating Doxygen documentation from the source code repositories of the various LHCb applications had failed. An important component of this work required another LHCb collaborator and I to repair the LHCb Doxygen pipeline, such that the pages were automatically generated during the release of new versions of DaVinci (and other LHCb applications), and the documentation for available TupleTools could be extracted for the Ntuple Wizard. On the other hand, the list of available stripping lines, and the interface for configuring all available TupleTools are extracted directly from running the DaVinci application.

The LHCb bookkeeping system, or LHCb database provides information about the file paths in the database system available for queries. An existing code base was leveraged for extracting metadata about the LHCb database, initially created for the recent release of LHCb Run 1 data on the CERN Open Data Portal.

An important feature of the Ntuple Wizard is the "static" backend of the application. The metadata is unchanging up to new legacy releases of the DaVinci application (which are few and far between) at which point updated versions of the Ntuple Wizard can be iterated. All necessary metadata and documentation are collected at deployment time and served as static files via HTTP. This means that no additional access to LHCb internal resources is necessary for the Ntuple Wizard once it is deployed.

5.2.3.1 External Documentation

While a lot of crucial information can be displayed to users of the Ntuple Wizard in real time directly through the application interface, there will undoubtedly be cases where additional documentation is needed to fully inform the user on how to create an optimal query. This external documentation is currently a work in progress and is being created using the Sphinx Python package [223] in a similar spirit as the LHCb Starterkit [203].

5.2.3.2 Source Code Management

Given that the metadata and documentation for the Ntuple Wizard are extracted mainly from LHCb internal sources, the code repository that handles this (referred to as the backend repository) is necessarily private. It is hosted on Gitlab (a platform to facilitate source code management with Git [224]) and written primarily in Python. The logic is steered by a workflow management tool called Snakemake [225]. The metadata and documentation are gathered categorically in JSON files, and then compressed and deployed to the web using a publishing service for static webpages called Gitlab Pages. The continuous integration (CI) feature of Gitlab is utilized to ensure the following:

1. Proper coding practices are followed

2. The codebase successfully builds

3. The output can be deployed to the web

The CI is triggered when any commits are pushed to any branch of the repository, but deployment to the web is only triggered when a branch is successfully merged into the master branch, or when a specific tag is created. Commits to the master branch are forbidden to avoid merge conflicts; changes can only be made to the master branch via merge requests (a merge request on Gitlab is effectively the same as a pull request on Github).

5.2.4 User Interface

The name "Ntuple Wizard" comes from the design principle that the user interface consists of a sequence of steps that the user is guided through to create the desired query, much like any other software wizard. It is a dynamic web page with no server-side component, where all necessary metadata are served as static files (see Section 5.2.3) and are rendered in the appropriate context to instruct the user. Formulation of a query consists of two main steps: dataset discovery and Ntuple configuration, outlined in the following subsections.

5.2.4.1 Dataset and Selection Discovery

The landing page of the Ntuple Wizard is the overarching **Production configuration**. It serves as the main hub where users have access to the list of available physics processes and corresponding stripping lines (i.e. selection algorithms), and once they make a selection, the ability to configure the corresponding Ntuple. This page is rendered as a table with each row corresponding to a selected physics process (e.g. decay) the user would like to analyze. The table is empty until the user makes a selection. The first step is therefore to choose from the list of available physics processes, handled by the **Decay search** step. This can be intuitively launched from the

landing page. Figure 5.5 shows an example of the Decay search web page, consisting of both a list of available physics processes, and several filtering functionalities. The list can be filtered by the following.

Decay head — the top level decaying particle

- Matching to the input particle
  - Including corresponding antiparticles
  - Excluding corresponding antiparticles
- Input category (multiple inputs accepted)
  - Matching to any input category
  - Matching to all input categories
  - Matching to none of the input categories

Contains — particles contained in the decay (multiple inputs accepted)

- Matching to any input particle
- Matching to all input particles
- Matching to none of the input particles
- Including or excluding the corresponding antiparticles from the matching procedure

Tags — categories or labels to describe physical processes (multiple inputs accepted)

- Matching to any input tags
- Matching to all input tags
- Matching to none of the input tags

Stripping line name — An LHCb specific selection algorithm, defined in Section 5.2.1

Multiple selections can be made from the list, at which point the selections are highlighted in blue, as shown in Fig. 5.5. It should be noted that the highlighted decay, in this case, corresponds to the example decay described in Section 5.2.1, which will be used to outline the various steps of the Ntuple Wizard. All user-made selections can be viewed simultaneously by clicking the "Show only selected" checkbox near the top right corner of the page. Once the user is satisfied with their selections, they can hit the green "Select" button at the bottom of the page, which will route the user back to the main **production configuration** step, but now with rows in the table corresponding to the selected physics processes. Each row has several corresponding columns. The first column lists the physics process and provides several buttons related to the Ntuple configuration. The second column contains a drop-down menu where a stripping line corresponding to the selected physics process must be chosen. This will specify a sequence of selection algorithms to be applied on the LHCb data to identify candidates from the desired process. Once a stripping line is selected, interaction with the adjacent button and final column is no longer blocked. A help button is located directly next to the stripping line drop-down menu, which contains links to the corresponding stripping versions exposing the sequence of algorithms used to identify candidates. This documentation is included to help inform the user on the optimal choice of stripping line. For a given stripping line, the corresponding stripping versions are rendered as blue capsule-shaped badges on each item in the drop-down menu. Tooltips are displayed to the user when they hover over these badges, specifying the data taking year corresponding to a given stripping version. In some cases, multiple stripping versions may be listed for the same data-taking year, this will have consequences for the third column of the table that will be addressed shortly. The final column of the table is another drop-down menu where the user can choose a data set. Data sets are broken up by stream, running year, magnet polarity (for systematic studies related to detector coverage or charge parity), and stripping

version. The stream is indicated by the name of the items in the drop-down menu, while the rest of the metadata is rendered in capsule-shaped badges with overlay tooltips, similar to the stripping line menu items. In the case of multiple stripping versions existing for a given running year, separate options will be rendered in the dataset selection drop-down menu. It is recommended to choose only one item per combination of running year and magnet polarity. In such cases the most recent stripping version should be selected. The production configuration step, containing a row with corresponding selections in the various columns, is shown in Figure 5.6.

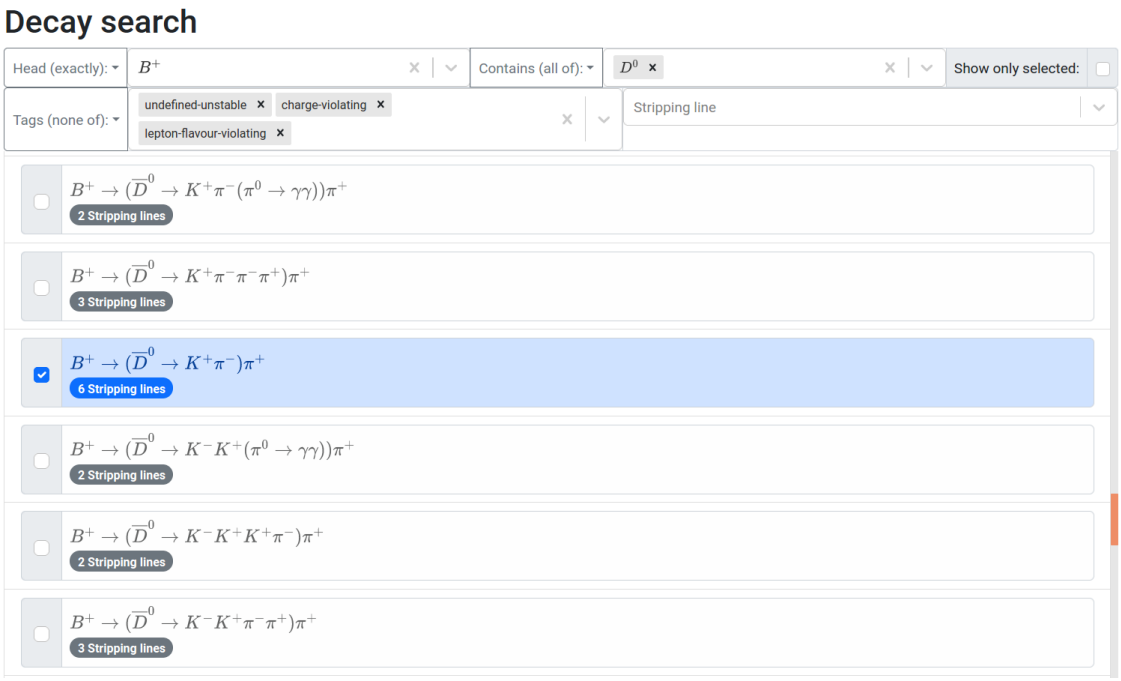

Figure 5.5: Example of the decay candidate search function of the Ntuple Wizard. This figure was taken from Ref. [3].

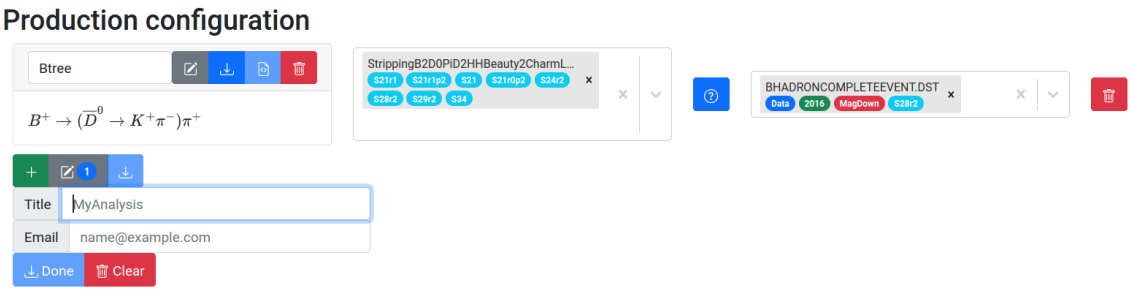

Figure 5.6: Example of the data set selection and production configuration step of the Ntuple Wizard. This figure was taken from Ref. [3].

5.2.4.2 Ntuple Configuration

After a physics process is selected, the configuration of the corresponding Ntuple can be initiated from the **production configuration** page. As mentioned previously, the first row in the column contains buttons related to Ntuple configuration. These buttons have mouse over tool tips to guide the user to the next steps. One such button creates an object for the Ntuple configuration, prompting the user to provide a name for the Ntuple and exposing the button to configure the Ntuple. This button will route the user to the **DecayTreeTuple configuration** step of the process, where an interactive decay tree is rendered with each particle in the decay corresponding to a node in the graph. The user will immediately notice 5 TupleTools corresponding to the DecayTreeTuple. These are the default TupleTools used in LHCb analyses, and are applied to the entire decay. At this stage, the user can choose to configure or remove any of the existing TupleTools, or add new TupleTools to the entire decay, particular nodes, or any grouping of nodes. An example of the Ntuple configuration page is shown in Figure 5.7 with the top panel showing default TupleTools corresponding to the entire DecayTreeTuple, and the bottom panel showing TupleTools added to selected nodes (the selected nodes are highlighted in blue and underlined in the "Current selection" field).

The TupleTools are rendered in a table under the decay tree (one per row), and the rows displayed depend on the currently selected nodes of the decay tree. If no nodes are selected, TupleTools applied to the entire decay are listed. The interfaces for the TupleTool configuration are extracted from DaVinci and served directly to the application from the static HTTP files containing the metadata, as discussed in Section 5.2.3. They can be accessed by clicking the edit pad button in the row corresponding to the desired TupleTool. A modal is rendered on screen where the user can input the desired configuration, with documentation of the TupleTool (including links to the Doxygen pages) listed at the bottom of the modal. Figure 5.8 shows an example of a TupleTool configuration modal for the particularly important tool that saves trigger information, *TupleToolTISTOS*, where TISTOS stands for Trigger Independent of Signal, Trigger On Signal.

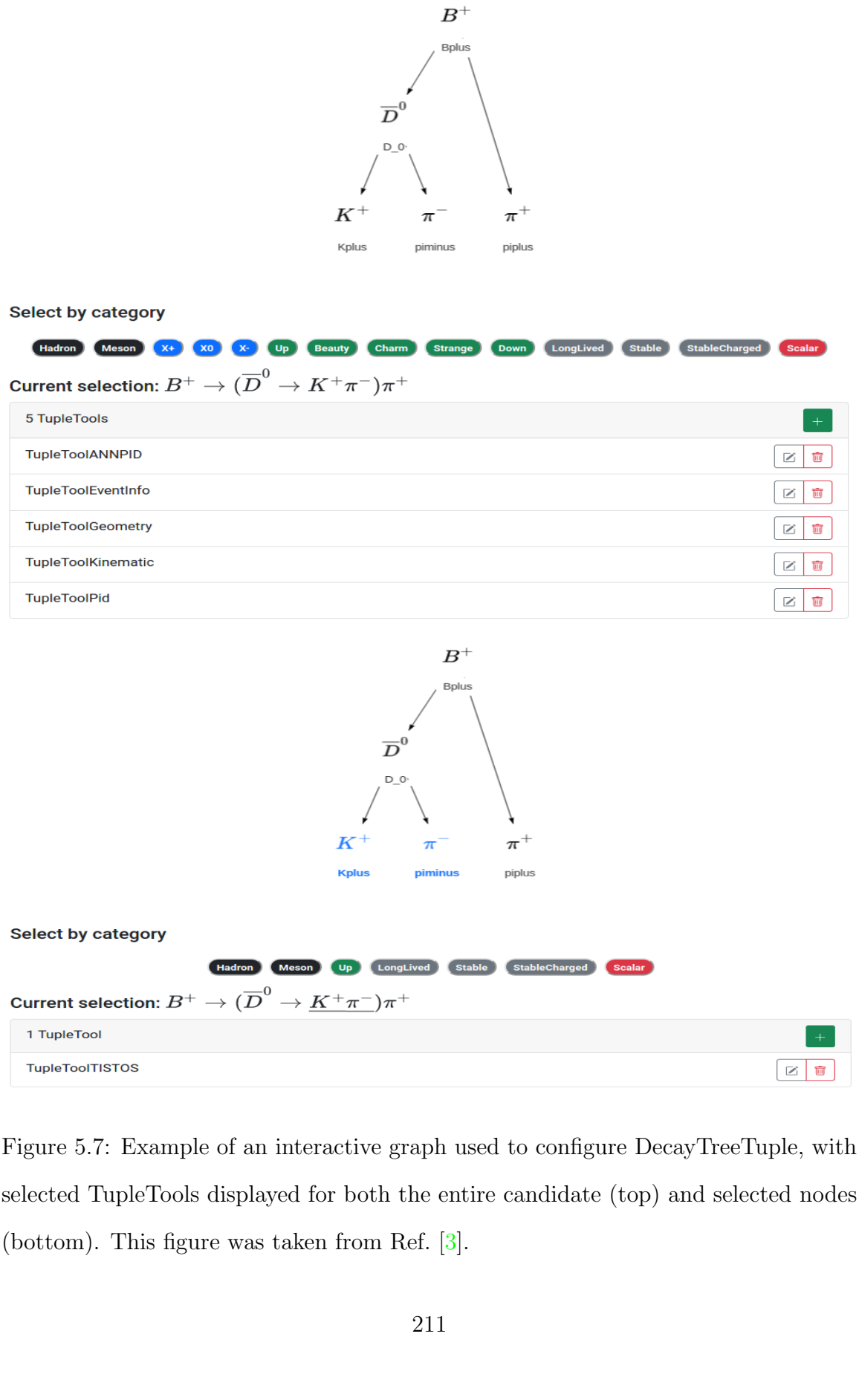

Figure 5.7: Example of an interactive graph used to configure DecayTreeTuple, with selected TupleTools displayed for both the entire candidate (top) and selected nodes (bottom). This figure was taken from Ref. [3].

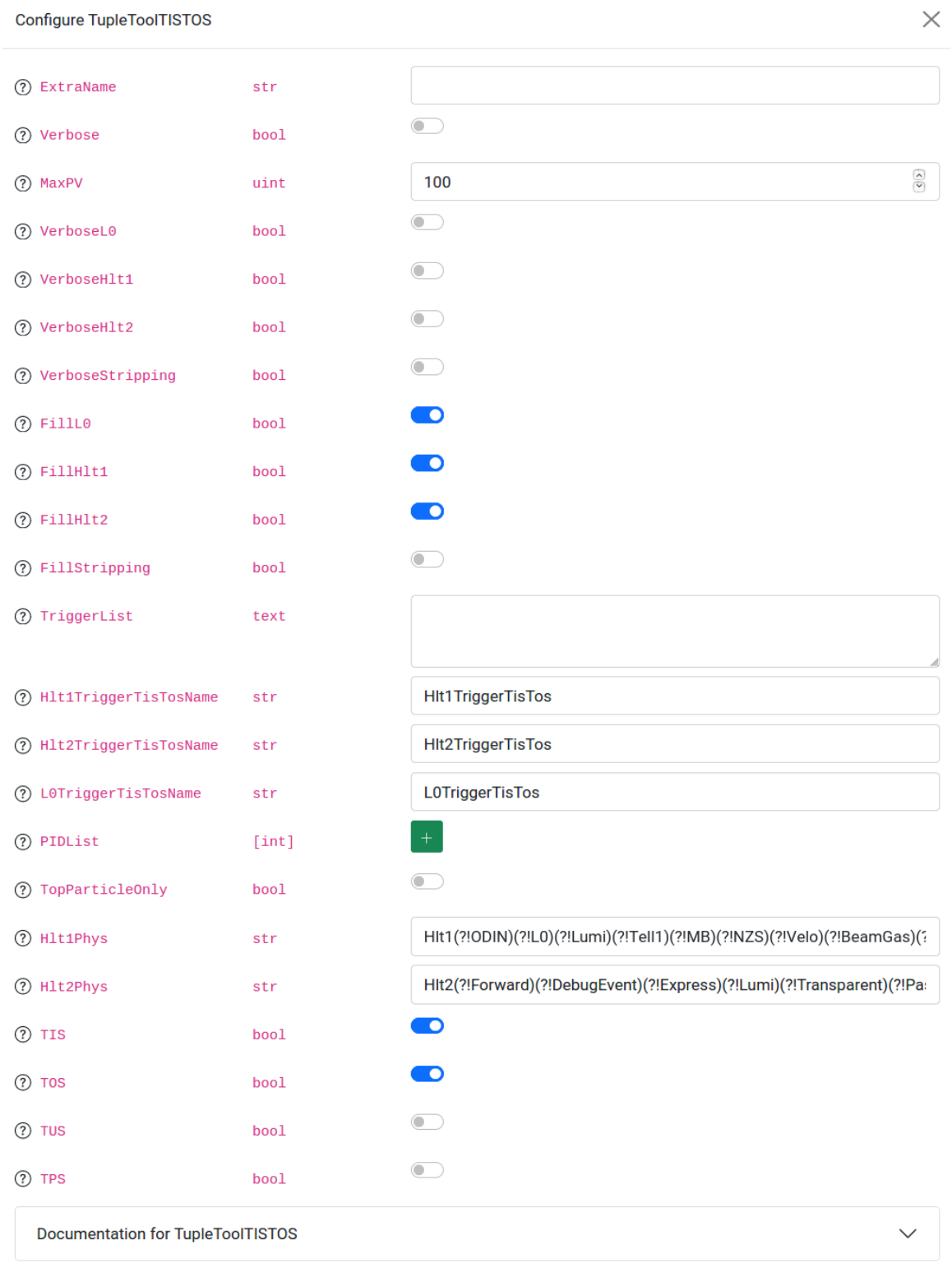

Figure 5.8: Example of the configuration interface of a TupleTool within the Ntuple Wizard, (in particular, *TupleToolTISTOS* for saving trigger information), including links to relevant documentation at the bottom of the modal. This figure was taken from Ref. [3].

5.2.4.3 Configuration files

As mentioned in Section 5.2.2, the configuration produced by the application is captured in YAML files. Namely, a YAML file to steer the entire job in the Analysis Production batch processing system, and a YAML file for each Ntuple to be created (e.g. each physics process selected) to configure the DaVinci application. Figure 5.9 shows the `info.yaml` file that steers the Analysis Productions system, created from the configurations input in Figures 5.5 - 5.8. A `job` is created and indexed for each data set added to a particular row configured in the **production configuration** step. The `options` field specifies the Python scripts used to configure the DaVinci application. The Python scripts are generated internally and the DecayTreeTuple configuration is done within them directly from the corresponding YAML files with the help of an LHCb specific parser named *AnalysisHelpers*. The YAML file corresponding the the Ntuple configuration produced in Figures 5.5 - 5.8 (`Btree.yaml`) is displayed in Figure 5.10. The structure of this configuration file is explained in detail in Ref. [3].

Figure 5.9: Output of info.yaml, the data file used to configure the AnalysisProductions batch processing system.

```
defaults:
  application: DaVinci/v46r7
  wg: DPA
  automatically_configure: true
  inform:
    - dfitzger@cern.ch
  output: DVNtuple.root
job0:
  input:
    bk_query: >-
      /LHCb/Collision16/Beam6500GeV-VeloClosed-MagDown/Real
      Data/Reco16/Stripping28r2/90000000/BHADRONCOMPLETEEVENT.
         DST
  options:
    - Btree.py
```

Figure 5.10: Output of Btree.yaml, the data file used to configure the DecayTreeTuple algorithm. This figure was taken from Ref. [3].

```
inputs:
  - /Event/BhadronCompleteEvent/Phys/
     B2D0PiD2HHBeauty2CharmLine/Particles
descriptorTemplate: ${Bplus}[B+ -> ${D_0}(D~0 -> ${Kplus}K+ ${
   piminus}pi-)${piplus}pi+]CC
tools:
  - TupleToolKinematic:
      ExtraName: ''
      Verbose: false
      MaxPV: 100
      Transporter: ParticleTransporter:PUBLIC
  - TupleToolPid:
      ExtraName: ''
      Verbose: false
      MaxPV: 100
  - TupleToolANNPID:
      ExtraName: ''
      Verbose: false
      MaxPV: 100
      ANNPIDTunes:
        - MC12TuneV2
        - MC12TuneV3
        - MC12TuneV4
        - MC15TuneV1
      PIDTypes:
        - Electron
        - Muon
        - Pion
        - Kaon
        - Proton
        - Ghost
  - TupleToolGeometry:
      ExtraName: ''
      Verbose: false
      MaxPV: 100
      RefitPVs: false
      PVReFitter: LoKi::PVReFitter:PUBLIC
      FillMultiPV: false
  - TupleToolEventInfo:
      ExtraName: ''
      Verbose: false
      MaxPV: 100
branches:
  Bplus:
    particle: B+
    tools: []
```

```
  D_0:
    particle: D~0
    tools: []
  Kplus:
    particle: K+
    tools: []
  piminus:
    particle: pi-
    tools: []
  piplus:
    particle: pi+
    tools: []
groups:
  Kplus,piminus:
    particles:
      - K+
      - pi-
    tools:
      - TupleToolTISTOS:
          ExtraName: ''
          Verbose: false
          MaxPV: 100
          VerboseL0: false
          VerboseHlt1: false
          VerboseHlt2: false
          VerboseStripping: false
          FillL0: true
          FillHlt1: true
          FillHlt2: true
          FillStripping: false
          TriggerList: []
          Hlt1TriggerTisTosName: Hlt1TriggerTisTos
          Hlt2TriggerTisTosName: Hlt2TriggerTisTos
          L0TriggerTisTosName: L0TriggerTisTos
          PIDList: []
          TopParticleOnly: false
          Hlt1Phys: >-
            Hlt1(?!ODIN)(?!L0)(?!Lumi)(?!Tell1)(?!MB)(?!NZS)
              (?!Velo)(?!BeamGas)(?!Incident).*Decision
          Hlt2Phys: >-
            Hlt2(?!Forward)(?!DebugEvent)(?!Express)(?!Lumi)
              (?!Transparent)(?!PassThrough).*Decision
          TIS: true
          TOS: true
          TUS: false
          TPS: false
name: DecayTreeTuple/Btree
```

5.2.4.4 Source Code Management

The source code for the user interface is managed using Gitlab (see Ref. [226]), and it is referred to as the frontend repository or frontend of the application. It is written in JavaSript [227] using the ReactJS frontend library [228] and a number of other packages specified in the `package.json` file in the root directory. Metadata is served via HTTP from the backend of the application (the static web page described in Section 5.2.3). Components rendered on the various web pages come from the `react-bootstrap` framework [229], including custom components built from standard components contextualized with LHCb metadata. The web pages are appropriately routed together using the `react-router-dom` package [230], such that the configuration is preserved when moving from one page to another. Similarly to the backend repository, the continuous integration (CI) feature of Gitlab was used to ensure the following:

1. Proper coding practices are followed
2. The codebase successfully builds
3. The output is deployed to an LHCb internal webpage for testing (only for tagged versions)

In general, the CI is executed whenever a commit is pushed to any branch, but item (3) specifically only runs when tagged versions of the repository are created. Commits to the master branch are forbidden — it is only updated through merge requests to prevent merge conflicts.

Once a version is tagged, a corresponding release is made on the Node Package Manager (npm) [231], found in Ref. [232]. This allows the components of the Ntuple Wizard application to be accessed and used publicly, which is crucial in integrating the application with the CERN Open Data Portal.

5.2.5 Integration with the Open Data Portal

At the conception of the Ntuple Wizard idea, we only envisioned interfacing with the CERN Open Data Portal (ODP) at the end of the chain when Ntuples are ready to be delivered to an output location, as shown in Figure 5.2. However, the collaboration with the CERN IT team has proven to be far more fruitful than we initially imagined. There are now ongoing efforts to fully integrate the Ntuple Wizard application into the CERN ODP, meaning that this is the location that users will go to access the application, as well as monitor their productions and retrieve the job output. This will help greatly in increasing the visibility and awareness of the application, given that the CERN ODP is already the central hub for open data in HEP. The CERN IT team has therefore created and maintained an additional layer, known as the *Ntupling Service*. This service has its own associated code repositories, one for the deployment of the application, with the Ntuple Wizard interface embedded, and one to store the user requests, which are submitted to the batch processing system upon approval.

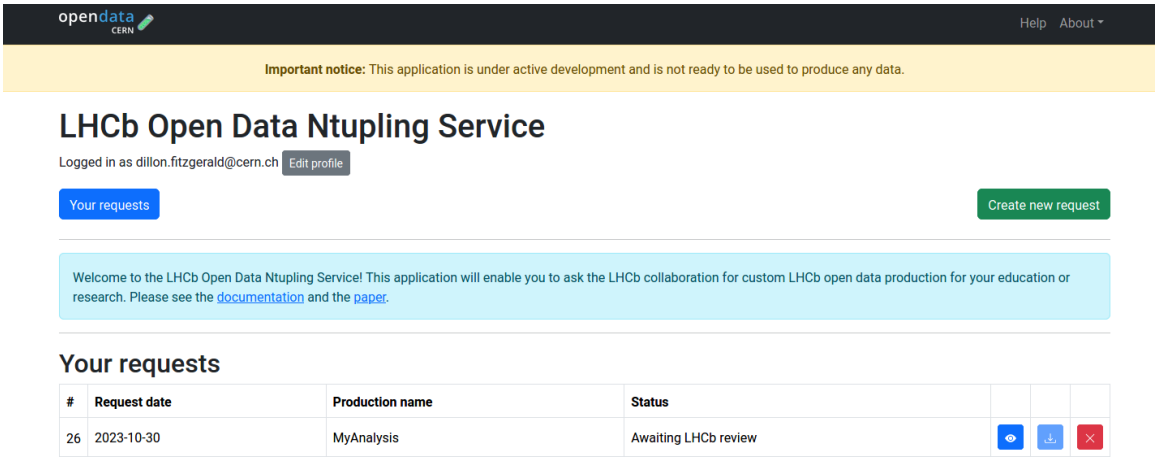

Figure 5.11: Landing page of the LHCb Ntupling Service on the CERN Open Data Portal, including information about previously made requests.

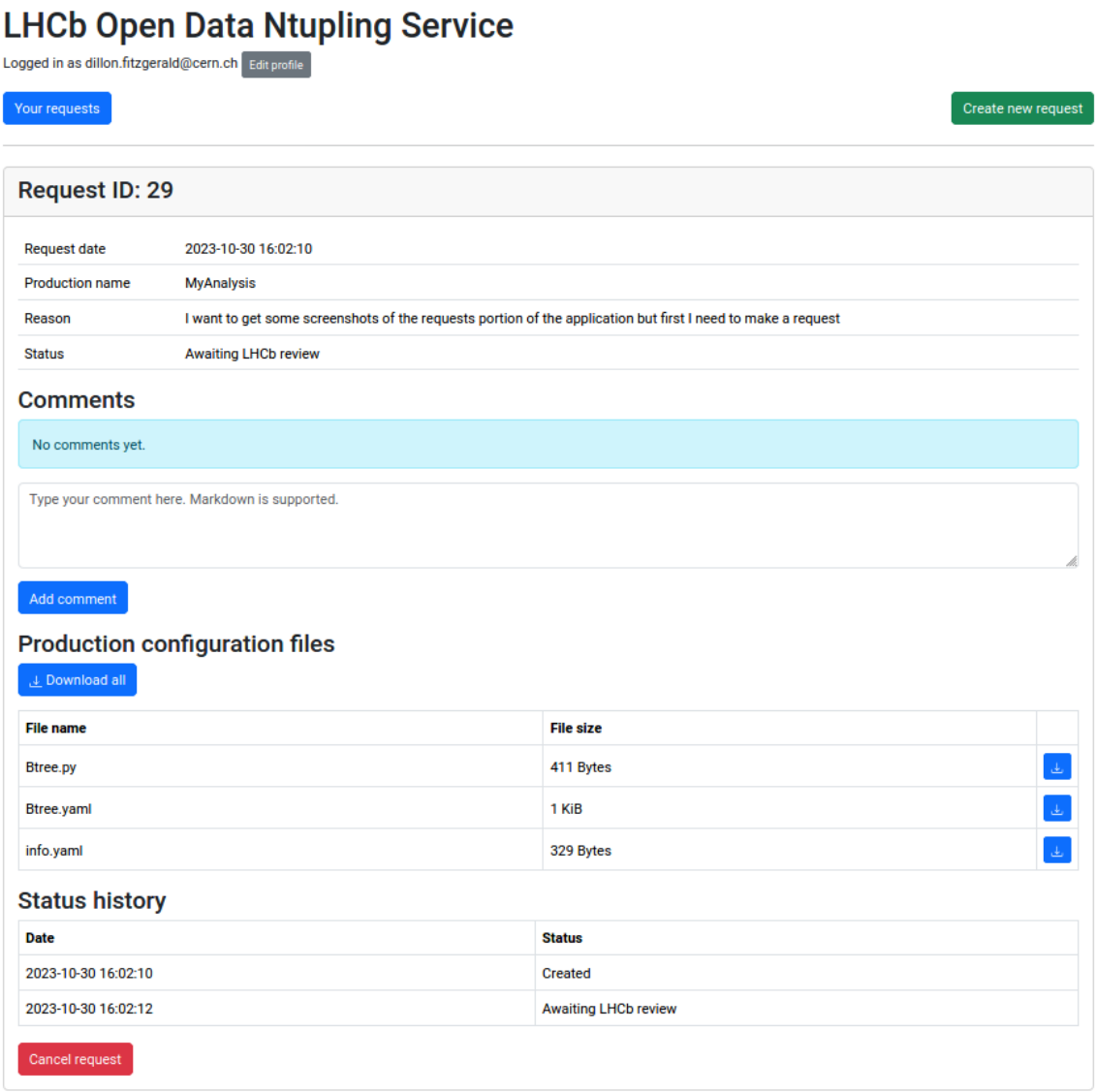

Figure 5.12: Expanded view of a previously made request on the LHCb Ntupling Service, accessed via the blue button with a picture of an eye on it in Fig. 5.11.

Figure 5.11 shows the Ntupling Service main page, where the Ntuple Wizard user interface discussed in Section 5.2.4 can be accessed by clicking the green button labeled "Create new request". Each request made by the user is displayed in a table listing the date of the request, the production name, the status (e.g. awaiting review), and buttons to view or cancel the request and download the output files when they are ready. Figure 5.12 shows the expanded view of a particular request,

including information about the request, the configuration files, and a mechanism for submitting comments to the LHCb open data team.

After submitting a request using the Ntupling Service, it is sent to the LHCb open data team for approval (see Section 5.2.6 for details). Upon approval, a test production will be initiated by the Analysis Productions batch system. Text-based output from the test production is routed back to the user with details about the projected data volume and the columns that will be present in the Ntuple. At this stage, the user can verify or modify the request. Once the request is verified, this will trigger the full production to be processed by the batch system (see Section 5.2.6 for details). The end result is Ntuples delivered back to the CERN ODP for retrieval. This procedure is outlined in Figure 5.13.

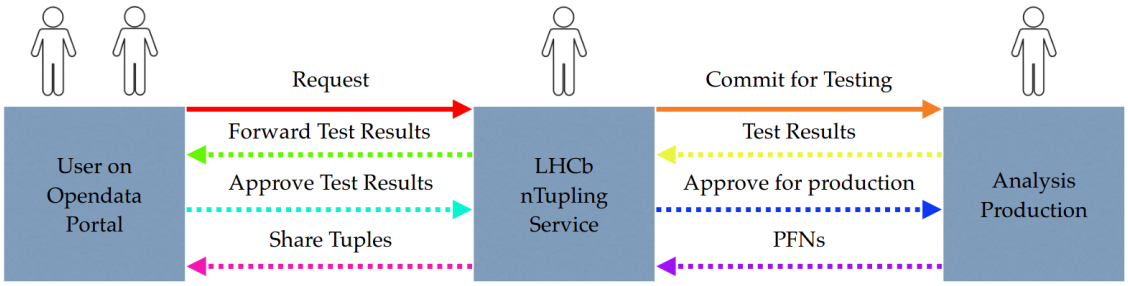

Figure 5.13: Schematic of the Ntuple Wizard workflow, including integration with the Ntupling Service hosted on the CERN Open Data Portal, and the Analysis Productions batch processing system at LHCb.

There are still some policy decisions to be made at this point, but the Ntuples will likely be stored for a time on the ODP and then deleted, with the expectation that the user will download them and transfer the data to their own private storage volume. There is discussion of promoting certain example Ntuples to be minted with a DOI, and to have the corresponding example configurations accessible to users of the service.

5.2.6 Batch Processing

Analysis Productions is a centralized batch processing system for creating Ntuples using the DaVinci application [220, 221]. It is hosted on Gitlab, where typically LHCb users create a new branch for their production while in the testing/development phase and then create a merge request for approval when they are ready to process the full set of data. The continuous integration feature of the Analysis Productions repository initiates a test production over a small subset of the data when a commit is pushed to the remote repository, and initiates the full production over the full dataset with one of LHCb's available computing grids when a merge request is approved and merged with the master branch.

Both test and full productions are handled by the LHCbDIRAC [214, 215] production system, where productions define how specified datasets will be processed. LHCbDIRAC will launch and manage computing jobs until the data set is fully processed, without the need for user intervention. Productions are configured in steps that specify the application to run, the scripts to configure the application, and the location of the desired dataset in the LHCb Bookkeeping system. The `info.yaml` files generated by the Ntuple Wizard (see Fig. 5.9) consist of a single production (i.e. job) per data set, each consisting of a single step with the DaVinci application.

The Ntupling Service layer developed in collaboration with CERN IT interfaces with Analysis Productions via the repository where user requests are stored. When a user submits a request through the Ntupling Service (see Section 5.2.5), it creates an issue in the service requests repository with the label "Awaiting LHCb review". Approval is handled by an LHCb open data team member by changing the label assigned to the issue to "Approved for test production". This will automatically trigger a commit in the Analysis Productions repository via the Analysis Productions Application Programming Interface (API), which will in turn launch a test production and return some feedback to the Ntupling Service for approval from the user. Once

approved, the label in the Ntupling Service repository will be updated accordingly to "Confirmed for real production", which will programmatically trigger a merge request in the Analysis Productions repository. Upon approval of the merge request by the LHCb Analysis Productions team, the full production will begin to run on an available computing grid. Once the production is complete, the Ntuples will be delivered to the CERN ODP for retrieval. The user will receive email notifications when the labels change in the Ntupling Service requests repository, keeping them up to date on the status of their request. The procedure discussed in this paragraph is shown in Figure 5.13. The mechanism for delivering the final Ntuples to the CERN Open Data Portal is still under active development, after which alpha testing can begin promptly within the collaboration.

5.2.7 First Public Release Plans

The Ntuple Wizard application, as well as the Ntupling Service that integrates the application with the CERN ODP and the LHCb Analysis Productions system, is well underway toward a first public release. There are only a few steps remaining before alpha testing can begin, as discussed in Section 5.2.6. The goal of the alpha testing will be to identify any remaining bugs and get feedback on the system from LHCb users. This will be followed by a period of beta testing among a community of theorists and phenomenologists associated with the LHC that will provide further suggestions for improvement. After the beta testing period is concluded, and subsequent versions of the applications are updated to implement user feedback, we can proceed with a first public release. The first public release is planned sometime next year following a press release given the high level of excitement in the community about the novel tool we are developing to provide open access to the data collected at LHCb. The time periods for the alpha and beta testing phases have yet to be determined. There are still some important policy discussions to be had internally within LHCb, such as

how much of the Run 2 data to initially provide access to, that will dictate the final release schedule.

5.2.8 Future Developments

Beyond the first public release, there are a number of future developments of the Ntuple Wizard that are underway by graduate students who have recently joined the team. This includes the configuration of custom jet and candidate reconstruction from standard collections of LHCb particles, as well as DecayTreeFitter, a tool that allows the user to provide mass constraints to resonant particles within a decay. Ideally, we will provide support to include the full set of algorithms for event selection and analysis available within DaVinci. In practice, we will address the algorithms we find most useful first, and attempt to get to the full set of algorithms as resources allow.

Additionally, the Ntuple Wizard will eventually need to be adapted to handle Run 3 data. Run 3 started in 2022, for which the majority of the data filtering is done in real time with the LHCb high-level trigger (HLT). This means that, for the most part, the data will be fully reconstructed online and the final preselection algorithms will run in the HLT, and the concept of stripping lines will not be as prevalent. However, the output will have a similar level of abstraction as that of the stripping procedure, allowing for a relatively simple adaptation of the application for Run 3 data when it is made public.

CHAPTER VI

Summary

Despite the proton being ubiquitous in nature and studied for over a century, there is still much to learn about its structure in terms of the constituent quarks and gluons. This goes for any hadrons as well as the mechanisms by which they are formed from quarks, or more specifically, the fragmentation of gluonic color flux tubes between quarks. Data sets available from both the former PHENIX experiment at RHIC and the LHCb experiment at the LHC permit unique approaches to studying these questions. RHIC is the world's first and only polarized proton collider, allowing for a wealth of spin-momentum and spin-spin correlations to be studied between hadrons and their constituents. LHCb on the other hand is a forward spectrometer equipped with instrumentation that makes it ideal to study the hadron formation process (hadronization) by measuring correlations between particles in jets. Such measurements at PHENIX and an effort to establish open access to LHCb data so that more can aid in the quest to understand the hadronization process (in addition to numerous other physics topics that can be addressed with LHCb data) comprise the core components of this dissertation work.

The quantum field theory that governs interactions between quarks and gluons, and ultimately the structure of hadronic bound states, is known as quantum chromodynamics (QCD). QCD is a unique QFT within the Standard Model due to properties

known as color confinement, and asymptotic freedom. The latter leads to weakly interacting, or approximately free quarks and gluons at high enough energies or short enough distance scales. This enables the use of perturbative QCD (pQCD) to mathematically describe many processes involving hadrons, empowered by the principle of universality but contingent on assumptions such as factorization. Factorization states that cross sections of processes involving hadrons can be mathematically factorized into short and long distance components (where a perturbative expansion can be applied for the short distance part). The long distance components are on the length scale of hadronic bound states, and are described with nonperturbative functions that parameterize the structure and formation of bound states. The initial state nonperturbative functions describing hadronic structure are generally referred to as parton distribution functions (PDFs), while the final state nonperturbative functions describing hadron formation are generally referred to as fragmentation functions (FFs). The nonperturbative functions are typically described in terms of collinear momentum fractions (x, z) as well as momentum transfer Q. Universality implies these functions are the same no matter how they are measured. In particular, a description of them in one energy regime can be evolved to obtain the appropriate description in another energy regime via the DGLAP integro-differential equation. In practice, these functions are extracted from data, and as a consequence of universality, nonperturbative functions constrained from data in one experiment can be evolved and used as an appropriate input to extract unknown information from another experiment.

Another consequence of asymptotic freedom is the successful application of the parton model in describing many phenomena involving hadrons, which treats hadrons as collections of approximately free point like constituents with collinear momentum. These constituents are referred to as partons (hence the name parton distribution functions). Together, factorization, universality, and the parton model, have allowed for the description of many spin-independent cross sections in processes involving

hadrons, and even the helicity distributions of quarks inside the proton. The assumptions begin to break down in processes involving hadrons with spin transverse to their momentum direction. In this case, a successful description requires either going to higher order in the factorization expansion or preserving the internal dynamics of partons within hadrons in the nonperturbative functions. More specifically, two viable approaches for describing such observables are the higher twist formalism (preserving the collinear factorization scheme), or the implementation a new factorization scheme that depends on the transverse momentum of partons within hadrons (replacing standard collinear nonpeturbative functions with corresponding TMD functions). Transverse spin observables therefore offer a rigorous test of QCD, and a path towards a better understanding of hadronic structure. Historically, much of the data used to constrain the spin structure of hadrons came from lepton-hadron scattering experiments, where processes involving gluons cannot be accessed at leading order. This makes the data at RHIC incredibly valuable, as it allows for the exploration of observables that depend on the proton spin in a collision system with leading order access to gluons. One particularly interesting set of functions that RHIC data can constrain are the twist-3 trigluon correlation functions, and the related gluon Sivers TMD PDF. Data from measurements presented in this dissertation are crucial in doing so.

The transverse single-spin asymmetries for midrapidity electrons and positrons from the decay of open heavy flavor hadrons were measured in $\sqrt{s} = 200$ GeV polarized pp collisions at PHENIX from 2015, and reported in Ref. [1]. This was a particularly important measurement for PHENIX, as heavy flavor production at $\sqrt{s} = 200$ GeV is dominated by gluon-gluon fusion, providing direct access to the trigluon correlations functions and indirect access to the gluon Sivers PDF. This measurement leveraged the excellent electron detection and identification capabilities of the PHENIX central arm spectrometers. The TSSAs for heavy flavor mesons D^0

and $\bar{D}^0$ were calculated for $\sqrt{s} = 200$ GeV pp collisions in Ref. [107], considering contributions from both qgq and ggg twist-3 correlation functions. The qgq contributions were shown to be negligible, implying that a non-zero A_N in this channel would clearly indicate trigluon correlations in polarized protons. In addition, measuring this for D^0 and $\bar{D}^0$ would provide constraints on contributions from the independent antisymmetric and symmetric trigluon correlation functions. A similar calculation was provided in Ref. [87] in the same channel but with a slightly different formalism (e.g. only considering contributions from the antisymmetric and symmetric trigluon correlators). Before this dissertation work, the trigluon correlation functions were relatively unconstrained. The data presented in Ref. [1] provide modest constraints, including the first extraction of theoretical parameters $\lambda_f = -0.01 \pm 0.03$ GeV and $\lambda_d = 0.11 \pm 0.09$ GeV, corresponding to normalizations of the trigluon correlation functions to the unpolarized gluon PDF [107]. Extractions of theoretical parameters K_G and K'_G from Ref. [87] (also corresponding to normalization parameters of the trigluon correlation functions to the unpolarized gluon PDF) were also performed, yielding consistent results with Ref. [87].

The 2015 datasets from RHIC are the only collider datasets in the world including polarized proton collisions on heavy nuclei, in particular gold and aluminum ions. This unique data provides several opportunities to study how TSSAs are modified in the presence of additional nuclear media, stronger color fields, and higher multiplicity collisions. One particularly interesting application lies in the comparison of TSSAs in pA collisions to that of the same process in pp collisions — a proposed measurement to access the gluon saturation scale. Regardless, the existing measurements from $p^{\uparrow}A$ data at RHIC have yielded many surprises and so far paint an inconsistent picture. This motivated the measurement of midrapidity TSSAs of π^0 and η mesons in $\sqrt{s_{NN}} = 200$ GeV pAu and pAl collisions from the 2015 PHENIX data, reported in Ref. [2]. This measurement takes advantage of the high granularity electromagnetic

calorimeter at PHENIX for the reconstruction of $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$ decays across a wide range of p_T. The data are plotted in comparison with that of Ref. [123], the same observable in polarized pp collisions, to see if any nuclear modification of the TSSA was observed. Both mesons (π^0 and η) in both collision systems (pAu and pAl) were consistent with zero and the corresponding pp measurements to high precision, showing no evidence for modification due to the presence of additional nuclear media in the collision system.

Another avenue for advancing the understanding of QCD is to enable access to quality datasets uniquely equipped to answer open questions in the field. LHCb is an ideal facility for studying the hadronization process, with high-precision tracking and vertexing, electromagnetic and hadronic calorimetry, and excellent particle identification capabilities for the separation of protons, kaons, and pions. Part of CERN's commitment to the public is the responsibility to provide access to the unique data collected at the LHC after a set amount of years for internal analysis. The LHCb data are relatively untapped in terms of potential QCD studies, given that most of the collaboration is focused on flavor physics, for which there are overlapping detector requirements. In 2022, LHCb released their first set of data to the public, consisting of about 200 TB (20%) of the Run 1 (2011-2012) collider data. At the time, the remaining Run 1 data was blocked from release due to internal LHCb policy that has since been amended. Work is currently underway for the release of the remaining Run 1 data to the CERN Open Data Portal. While this is a huge achievement for the LHCb collaboration, it is not a scalable strategy for the release of Run 2 (2015-2018) data and beyond. This is where another core component of my dissertation work comes in; the development of an application that enables users to query available LHCb data via a web interface resulting in Ntuples delivered to the CERN Open Data Portal. The application is known as the LHCb Ntuple Wizard, and it is described in Ref. [3]. An initial version of the application is already developed, and current work is ongoing

with the IT team at CERN to integrate the application directly within the CERN Open Data Portal. A public release of the entire service integrated within the portal is expected in 2024.

The measurements and applications presented in this dissertation are both important and timely for the advancement of QCD. TSSAs are a unique probe of hadron structure that aim to provide a more complete three-dimensional picture of the internal dynamics. Additionally, making the most out of existing datasets is crucial in preparation for successful analysis and interpretation of new data from the Run 3 LHC, sPHENIX, and the future EIC. The landscape of QCD is evolving. With a wealth of measurements from colliders around the world and recent advancements in theory and computation, the answers to questions that have long eluded us are within reach.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] N. J. Abdulameer et al. Improving constraints on gluon spin-momentum correlations in transversely polarized protons via midrapidity open-heavy-flavor electrons in p↑+p collisions at s=200 GeV. *Phys. Rev. D*, 107(5):052012, 2023.

[2] N. J. Abdulameer et al. Transverse single-spin asymmetry of midrapidity $\pi 0$ and η mesons in p+Au and p+Al collisions at sNN=200 GeV. *Phys. Rev. D*, 107(11):112004, 2023.

[3] Christine A. Aidala, Christopher Burr, Marco Cattaneo, Dillon S. Fitzgerald, Adam Morris, Sebastian Neubert, and Donijor Tropmann. Ntuple Wizard: An Application to Access Large-Scale Open Data from LHCb. *Comput. Softw. Big Sci.*, 7(1):6, 2023.

[4] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.

[5] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973.

[6] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[7] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[8] John Collins. *Foundations of perturbative QCD*, volume 32. Cambridge University Press, 11 2013.

[9] Stephen P. Martin and James D. Wells. *Elementary particles and their interactions*. Springer, 2022.

[10] F. Halzen and Alan D. Martin. *Quakrs and leptons: an introductory course in modern particle physics*. John Wiley & Sons, 1984.

[11] Dmitri E. Kharzeev and Eugene M. Levin. Color Confinement and Screening in the θ Vacuum of QCD. *Phys. Rev. Lett.*, 114(24):242001, 2015.

[12] P A Zyla et al. Review of Particle Physics. *PTEP*, 2020:083C01, 2020.

[13] R. P. Feynman. The behavior of hadron collisions at extreme energies. *Conf. Proc. C*, 690905:237–258, 1969.

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

[15] John C. Collins, Davison E. Soper, and George F. Sterman. Factorization for Short Distance Hadron - Hadron Scattering. *Nucl. Phys. B*, 261:104–142, 1985.

[16] John C. Collins, Davison E. Soper, and George F. Sterman. Soft Gluons and Factorization. *Nucl. Phys. B*, 308:833–856, 1988.

[17] Geoffrey T. Bodwin. Factorization of the Drell-Yan Cross-Section in Perturbation Theory. *Phys. Rev. D*, 31:2616, 1985. [Erratum: Phys.Rev.D 34, 3932 (1986)].

[18] A. H. Mueller, editor. *Perturbative quantum chromodynamics*. WSP, Singapore, 1989.

[19] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

[20] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[21] L. N. Lipatov. The parton model and perturbation theory. *Yad. Fiz.*, 20:181–198, 1974.

[22] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977.

[23] Elliot Leader. *Spin in particle physics*, volume 15. Cambridge University Press, 2 2011.

[24] Christine A. Aidala, Steven D. Bass, Delia Hasch, and Gerhard K. Mallot. The Spin Structure of the Nucleon. *Rev. Mod. Phys.*, 85:655–691, 2013.

[25] J. Ashman et al. A Measurement of the Spin Asymmetry and Determination of the Structure Function g(1) in Deep Inelastic Muon-Proton Scattering. *Phys. Lett. B*, 206:364, 1988.

[26] J. Ashman et al. An investigation of the spin structure of the proton in deep inelastic scattering of polarised muons on polarised protons. *Nuclear Physics B*, 328(1):1–35, 1989.

[27] J. P. Ralston and D. E. Soper. Production of dimuons from high-energy polarized proton-proton collisions. *Nucl. Phys. B*, 152(1):109, 1979.

[28] J. C. Ward. An identity in quantum electrodynamics. *Phys. Rev.*, 78:182–182, Apr 1950.

[29] Y. Takahashi. On the generalized Ward identity. *Nuovo Cim.*, 6:371, 1957.

[30] A. V. Efremov and O. V. Teryaev. QCD Asymmetry and Polarized Hadron Structure Functions. *Phys. Lett. B*, 150:383, 1985.

[31] S. Wandzura and F. Wilczek. Sum rules for spin-dependent electroproduction-test of relativistic constituent quarks. *Physics Letters B*, 72(2):195–198, 1977.

[32] J. P. Chen, A. Deur, S. Kuhn, and Z. E. Meziani. Spin structure functions. *J. Phys. Conf. Ser.*, 299:012005, 2011.

[33] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.

[34] Silvia Ferreres-Solé and Torbjörn Sjöstrand. The space–time structure of hadronization in the Lund model. *Eur. Phys. J. C*, 78(11):983, 2018.

[35] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. A Modified cluster hadronization model. *Eur. Phys. J. C*, 36:381–395, 2004.

[36] Yang-Ting Chien, Abhay Deshpande, Mriganka Mouli Mondal, and George Sterman. Probing hadronization with flavor correlations of leading particles in jets. *Phys. Rev. D*, 105(5):L051502, 2022.

[37] A. Metz and A. Vossen. Parton fragmentation functions. *Progress in Particle and Nuclear Physics*, 91:136–202, 2016.

[38] Adrian Signer. Low-energy Precision Physics and the High-energy Frontier. *Phys. Procedia*, 51:25–30, 2014.

[39] Andrew Larkoski, Simone Marzani, Jesse Thaler, Aashish Tripathee, and Wei Xue. Exposing the QCD Splitting Function with CMS Open Data. *Phys. Rev. Lett.*, 119(13):132003, 2017.

[40] Yuri L. Dokshitzer, Valery A. Khoze, and S. I. Troian. On specific QCD properties of heavy quark fragmentation ('dead cone'). *J. Phys. G*, 17:1602–1604, 1991.

[41] S. Acharya et al. Direct observation of the dead-cone effect in quantum chromodynamics. *Nature*, 605(7910):440–446, 2022. [Erratum: Nature 607, E22 (2022)].

[42] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[43] Yang-Ting Chien, Zhong-Bo Kang, Felix Ringer, Ivan Vitev, and Hongxi Xing. Jet fragmentation functions in proton-proton collisions using soft-collinear effective theory. *JHEP*, 05:125, 2016.

[44] Zhong-Bo Kang, Kyle Lee, John Terry, and Hongxi Xing. Jet fragmentation functions for Z-tagged jets. *Phys. Lett. B*, 798:134978, 2019.

[45] Georges Aad et al. Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector. *Eur. Phys. J. C*, 71:1795, 2011.

[46] Roel Aaij et al. Measurement of charged hadron production in Z-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, 123(23):232001, 2019.

[47] Kara Mattioli. *Measurement of Charged Hadron Distributions in Beauty and Charm Quark Tagged Jets at the Large Hadron Collider beauty Experiment.* PhD thesis, Michigan U., 2022.

[48] R. Aaij et al. Multidifferential study of identified charged hadron distributions in z-tagged jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 108:L031103, Aug 2023.

[49] J. Collins. Fragmentation of transversely polarized quarks probed in transverse momentum distributions. *Nucl. Phys. B*, 396(1):161, 1993.

[50] G. Bunce et al. Lambda0 Hyperon Polarization in Inclusive Production by 300-GeV Protons on Beryllium. *Phys. Rev. Lett.*, 36:1113–1116, 1976.

[51] K. Heller, O.E. Overseth, G. Bunce, F. Dydak, and H. Taureg. Λ^0 hyperon polarization in inclusive production by 24 GeV protons on platinum. *Physics Letters B*, 68(5):480–482, 1977.

[52] Kenneth J. Heller et al. Polarization of Lambdas and anti-Lambdas Produced by 400-GeV Protons. *Phys. Rev. Lett.*, 41:607, 1978. [Erratum: Phys.Rev.Lett. 45, 1043 (1980)].

[53] Edmond C. Dukes, Oliver E. Overseth, Gerry M. Bunce, Yousef Makdisi, Peter Yamin, Douglas A. Jensen, Michael N. Kreisler, Martin Marcin, Kamal K. Raychaudhuri, Marilyn W. Sullivan, Kenneth Heller, and Mike Shupe. Polarization of σ^0 hyperons in inclusive production from 28.5 gev/c protons on beryllium. *Physics Letters B*, 193(1):135–139, 1987.

[54] K. B. Luk et al. Polarization of Ω^- Hyperons Produced in 800 GeV Proton - Beryllium Collisions. *Phys. Rev. Lett.*, 70:900–903, 1993.

[55] Kenneth J. Heller. Spin and high energy hyperon production, results and prospects. *Conf. Proc.*, C960910:23–30, 1996.

[56] J. Felix et al. Λ^0 Polarization in 800 GeV/c $pp \to p_f(\Lambda^0 K^+)$. *Phys. Rev. Lett.*, 88:061801, 2002.

[57] Georges Aad et al. Measurement of the transverse polarization of Λ and $\bar{\Lambda}$ hyperons produced in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. *Phys. Rev. D*, 91(3):032004, 2015.

[58] Y. Guan et al. Observation of Transverse $\Lambda/\bar{\Lambda}$ Hyperon Polarization in e^+e^- Annihilation at Belle. *Phys. Rev. Lett.*, 122(4):042001, 2019.

[59] R. D. Klem, J. E. Bowers, H. W. Courant, H. Kagan, M. L. Marshak, E. A. Peterson, K. Ruddick, W. H. Dragoset Jr., and J. B. Roberts. Measurement of Asymmetries of Inclusive Pion Production in Proton Proton Interactions at 6 and 11.8 GeV/c. *Phys. Rev. Lett.*, 36:929, 1976.

[60] D. L. Adams et al. Analyzing power in inclusive π^+ and π^- production at high x_F with a 200-GeV polarized proton beam. *Phys. Lett. B*, 264:462, 1991.

[61] C. E. Allgower et al. Measurement of analyzing powers of π^+ and π^- produced on a hydrogen and a carbon target with a 22-GeV/c incident polarized proton beam. *Phys. Rev. D*, 65:092008, 2002.

[62] I. Arsene et al. Single Transverse Spin Asymmetries of Identified Charged Hadrons in Polarized $p + p$ Collisions at $\sqrt{s} = 62.4$ GeV. *Phys. Rev. Lett.*, 101:042001, 2008.

[63] 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.*, 41:1689, Dec 1978.

[64] S. Benic, Y. Hatta, H.-n. Li, and D.-J. Yang. Single-spin asymmetries at two loops. *Phys. Rev. D*, 100(9):094027, 2019.

[65] J.-W. Qiu and G. Sterman. Single transverse spin asymmetries. *Phys. Rev. Lett.*, 67:2264, Oct 1991.

[66] D. Sivers. Single-spin production asymmetries from the hard scattering of point-like constituents. *Phys. Rev. D*, 41:83, Jan 1990.

[67] D. Boer and P. J. Mulders. Time-reversal odd distribution functions in leptoproduction. *Phys. Rev. D*, 57:5780, May 1998.

[68] M. Anselmino, A. Mukherjee, and A. Vossen. Transverse spin effects in hard semi-inclusive collisions. *Prog. Part. Nucl. Phys.*, 114:103806, 2020.

[69] X. Ji, J.-W. Qiu, W. Vogelsang, and F. Yuan. A Unified picture for single transverse-spin asymmetries in hard processes. *Phys. Rev. Lett.*, 97:082002, 2006.

[70] Y. Koike, W. Vogelsang, and F. Yuan. On the relation between mechanisms for single-transverse-spin asymmetries. *Phys. Lett. B*, 659(5):878, 2008.

[71] F. Yuan and J. Zhou. Collins function and the single transverse spin asymmetry. *Phys. Rev. Lett.*, 103:052001, Jul 2009.

[72] X.-d. Ji, J.-P. Ma, and F. Yuan. Three quark light cone amplitudes of the proton and quark orbital motion dependent observables. *Nucl. Phys. B*, 652:383, 2003.

[73] Y. Hatta, K. Tanaka, and S. Yoshida. Twist-three relations of gluonic correlators for the transversely polarized nucleon. J. High Energy Phys. 02 **04 (2013)** 003.

[74] Y. Hatta and X. Yao. QCD evolution of the orbital angular momentum of quarks and gluons: Genuine twist-three part. *Phys. Lett. B*, 798:134941, 2019.

[75] J. Cammarota, L. Gamberg, Z.-B. Kang, J. A. Miller, D. Pitonyak, A. Prokudin, T. C. Rogers, and N. Sato. Origin of single transverse-spin asymmetries in high-energy collisions. *Phys. Rev. D*, 102(5):054002, 2020.

[76] Adam Freese, Ian C. Cloët, and Peter C. Tandy. Gluon PDF from quark dressing in the nucleon and pion. *Phys. Lett. B*, 823:136719, 2021.

[77] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, 1974.

[78] I. Balitsky. High-energy QCD and Wilson lines. 1 2001.

[79] J.-W. Qiu and G. F. Sterman. Single transverse spin asymmetries in direct photon production. *Nucl. Phys. B*, 378:52, 1992.

[80] Xiangdong Ji. Gluon correlations in the transversely polarized nucleon. *Phys. Lett. B*, 289(1):137, 1992.

[81] A. V. Belitsky, X. Ji, W. Lu, and J. Osborne. Singlet g_2 structure function in the next-to-leading order. *Phys. Rev. D*, 63:094012, Apr 2001.

[82] V. M. Braun, A. N. Manashov, and B. Pirnay. Scale dependence of twist-three contributions to single spin asymmetries. *Phys. Rev. D*, 80:114002, Dec 2009.

[83] H. Beppu, Y. Koike, K. Tanaka, and S. Yoshida. Contribution of twist-3 multigluon correlation functions to single spin asymmetry in semi-inclusive deep inelastic scattering. *Phys. Rev. D*, 82:054005, Sep 2010.

[84] Z.-B. Kang and J.-W. Qiu. Single transverse-spin asymmetry for D-meson production in semi-inclusive deep inelastic scattering. *Phys. Rev. D*, 78:034005, Aug 2008.

[85] C. Kouvaris, J.-W. Qiu, W. Vogelsang, and F. Yuan. Single transverse-spin asymmetry in high transverse momentum pion production in pp collisions. *Phys. Rev. D*, 74:114013, 2006.

[86] Jian-wei Qiu and George F. Sterman. Single transverse spin asymmetries in hadronic pion production. *Phys. Rev. D*, 59:014004, 1999.

[87] Y. Koike and S. Yoshida. Probing the three-gluon correlation functions by the single spin asymmetry in $p^{\uparrow}p \to DX$. *Phys. Rev. D*, 84:014026, 2011.

[88] Caroline Riedl. Recent Highlights from Spin-Physics Experiments. *SciPost Phys. Proc.*, 8:002, 2022.

[89] John C. Collins. Leading twist single transverse-spin asymmetries: Drell-Yan and deep inelastic scattering. *Phys. Lett. B*, 536:43–48, 2002.

[90] L. Adamczyk et al. Measurement of the transverse single-spin asymmetry in $p^{\uparrow} + p \to W^{\pm}/Z^0$ at RHIC. *Phys. Rev. Lett.*, 116(13):132301, 2016.

[91] Bakur Parsamyan. First measurement of transverse-spin-dependent azimuthal asymmetries in the Drell-Yan process. *PoS*, DIS2017:243, 2018.

[92] C. Adolph et al. Experimental investigation of transverse spin asymmetries in muon-p SIDIS processes: Collins asymmetries. *Phys. Lett. B*, 717:376–382, 2012.

[93] C. Adolph et al. II – Experimental investigation of transverse spin asymmetries in μ -p SIDIS processes: Sivers asymmetries. *Phys. Lett. B*, 717:383–389, 2012.

[94] C Adolph et al. Sivers asymmetry extracted in SIDIS at the hard scales of the Drell–Yan process at COMPASS. *Phys. Lett. B*, 770:138–145, 2017.

[95] C. Adolph et al. Collins and Sivers asymmetries in muonproduction of pions and kaons off transversely polarised protons. *Phys. Lett. B*, 744:250–259, 2015.

[96] M. G. Alexeev et al. Measurement of P_T-weighted Sivers asymmetries in leptoproduction of hadrons. *Nucl. Phys. B*, 940:34–53, 2019.

[97] A. Airapetian et al. Observation of the Naive-T-odd Sivers Effect in Deep-Inelastic Scattering. *Phys. Rev. Lett.*, 103:152002, 2009.

[98] Kazuo Abe et al. Measurement of azimuthal asymmetries in inclusive production of hadron pairs in e+ e- annihilation at Belle. *Phys. Rev. Lett.*, 96:232002, 2006.

[99] R. Seidl et al. Measurement of Azimuthal Asymmetries in Inclusive Production of Hadron Pairs in e+e- Annihilation at s**(1/2) = 10.58-GeV. *Phys. Rev. D*, 78:032011, 2008. [Erratum: Phys.Rev.D 86, 039905 (2012)].

[100] M. Ablikim et al. Measurement of azimuthal asymmetries in inclusive charged dipion production in e^+e^- annihilations at $\sqrt{s} = 3.65$ GeV. *Phys. Rev. Lett.*, 116(4):042001, 2016.

[101] J. P. Lees et al. Measurement of Collins asymmetries in inclusive production of charged pion pairs in e^+e^- annihilation at BABAR. *Phys. Rev. D*, 90(5):052003, 2014.

[102] J. P. Lees et al. Collins asymmetries in inclusive charged KK and $K\pi$ pairs produced in e^+e^- annihilation. *Phys. Rev. D*, 92(11):111101, 2015.

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

[104] C. Adolph et al. Collins and Sivers asymmetries in muonproduction of pions and kaons off transversely polarised protons. *Phys. Lett. B*, 744:250–259, 2015.

[105] C. Adolph et al. Experimental investigation of transverse spin asymmetries in muon-p SIDIS processes: Collins asymmetries. *Phys. Lett. B*, 717:376–382, 2012.

[106] Mohamed Abdallah et al. Azimuthal transverse single-spin asymmetries of inclusive jets and identified hadrons within jets from polarized pp collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 106(7):072010, 2022.

[107] Z.-B. Kang, J.-W. Qiu, W. Vogelsang, and F. Yuan. Accessing tri-gluon correlations in the nucleon via the single spin asymmetry in open charm production. *Phys. Rev. D*, 78:114013, 2008.

[108] M. Anselmino, M. Boglione, U. D'Alesio, E. Leader, and F. Murgia. Accessing Sivers gluon distribution via transverse single-spin asymmetries in $p^\uparrow p \to DX$ processes at BNL RHIC. *Phys. Rev. D*, 70:074025, Oct 2004.

[109] U. D'Alesio, F. Murgia, C. Pisano, and P. Taels. Probing the gluon Sivers function in $p^\uparrow p \to J/\psi\, X$ and $p^\uparrow p \to D\, X$. *Phys. Rev. D*, 96(3):036011, 2017.

[110] U. D'Alesio, C. Flore, F. Murgia, C. Pisano, and P. Taels. Unraveling the Gluon Sivers Function in Hadronic Collisions at RHIC. *Phys. Rev. D*, 99(3):036013, 2019.

[111] Z. B. Kang and F. Yuan. Single Spin Asymmetry Scaling in the Forward Rapidity Region at RHIC. *Phys. Rev. D*, 84:034019, 2011.

[112] Y. V. Kovchegov and M. D. Sievert. A New Mechanism for Generating a Single Transverse Spin Asymmetry. *Phys. Rev. D*, 86:034028, 2012. [Phys. Rev. D **86**, 079906(E) (2012)].

[113] A. Schäfer and J. Zhou. Color entanglement for γ-jet production in polarized pA collisions. *Phys. Rev. D*, 90(9):094012, 2014.

[114] A. Schäfer and J. Zhou. Transverse single spin asymmetry in direct photon production in polarized pA collisions. *Phys. Rev. D*, 90(3):034016, 2014.

[115] Y. V. Kovchegov and M. D. Sievert. Calculating TMDs of a Large Nucleus: Quasi-Classical Approximation and Quantum Evolution. *Nucl. Phys. B*, 903:164–203, 2016.

[116] J. Zhou. Transverse single spin asymmetry in Drell-Yan production in polarized pA collisions. *Phys. Rev. D*, 92(1):014034, 2015.

[117] Y. Hatta, B. W. Xiao, S. Yoshida, and F. Yuan. Single Spin Asymmetry in Forward pA Collisions. *Phys. Rev. D*, 94(5):054013, 2016.

[118] Y. Hatta, B. W. Xiao, S. Yoshida, and F. Yuan. Single spin asymmetry in forward pA collisions II: Fragmentation contribution. *Phys. Rev. D*, 95(1):014008, 2017.

[119] J. Zhou. Single spin asymmetries in forward $p - p/A$ collisions revisited: the role of color entanglement. *Phys. Rev. D*, 96(3):034027, 2017.

[120] S. Benić and Y. Hatta. Single spin asymmetry in forward pA collisions: Phenomenology at RHIC. *Phys. Rev. D*, 99(9):094012, 2019.

[121] Y. V. Kovchegov and M. G. Santiago. Lensing mechanism meets small- x physics: Single transverse spin asymmetry in $p^\uparrow + p$ and $p^\uparrow + A$ collisions. *Phys. Rev. D*, 102(1):014022, 2020.

[122] S. Benić, D. Horvatić, A. Kaushik, and E. A. Vivoda. Odderon mechanism for transverse single spin asymmetry in the Wandzura-Wilczek approximation. *Phys. Rev. D*, 106(11):114025, 2022.

[123] U. Acharya et al. Transverse single-spin asymmetries of midrapidity π^0 and η mesons in polarized $p + p$ collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 103(5):052009, 2021.

[124] M. Harrison, T. Ludlam, and S. Ozaki. RHIC project overview. *Nucl. Instrum. Meth. A*, 499:235–244, 2003.

[125] A. Adare et al. An Upgrade Proposal from the PHENIX Collaboration. 1 2015.

[126] A. Accardi et al. Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all. *Eur. Phys. J. A*, 52(9):268, 2016.

[127] R. Abdul Khalek et al. Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report. *Nucl. Phys. A*, 1026:122447, 2022.

[128] Vincent Schoefer. Acceleration and collisions at rhic, August 2021.

[129] A. Zelenski et al. Absolute polarized H-jet polarimeter development, for RHIC. *Nucl. Instrum. Meth. A*, 536:248–254, 2005.

[130] Alexei Fedotov et al. Operational Electron Cooling in the Relativistic Heavy Ion Collider. In *12th International Particle Accelerator Conference* , 8 2021.

[131] Vladimir Litvinenko et al. Coherent Electron Cooling (CeC) Experiment at RHIC: Status and Plans. In *10th International Particle Accelerator Conference*, page TUPTS078, 2019.

[132] K. H. Ackermann et al. STAR detector overview. *Nucl. Instrum. Meth. A*, 499:624–632, 2003.

[133] K. Adcox et al. PHENIX detector overview. *Nucl. Instrum. Methods Phys. Res., Sec. A*, 499:469, 2003.

[134] I Alekseev et al. Polarized proton collider at RHIC. *Nucl. Instrum. Meth. A*, 499:392–414, 2003.

[135] A. Zelenski et al. The RHIC polarized source upgrade. In *19th International Spin Physics Symposium*, page 295, 2011.

[136] P G Sona. New method proposed to increase polarization in polarized ion sources of h^- and d. *Energ. Nucl. (Milan), 14: 295-9.*, 14:295–299, May 1967.

[137] V. Bargmann, Louis Michel, and V. L. Telegdi. Precession of the polarization of particles moving in a homogeneous electromagnetic field. *Phys. Rev. Lett.*, 2:435–436, May 1959.

[138] Ya. S. Derbenev, A. M. Kondratenko, S. I. Serednyakov, A. N. Skrinsky, G. M. Tumaikin, and Yu. M. Shatunov. RADIATIVE POLARIZATION: OBTAINING, CONTROL, USING. *Part. Accel.*, 8:115–126, 1978.

[139] V. Ptitsin and Yu.M. Shatunov. Helical spin rotators and snakes. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 398(2):126–130, 1997.

[140] M Bai, C Dawson, Y Makdisi, W Meng, F Meot, P Oddo, C Pai, P Pile, and T Roser. Relativistic heavy ion collider spin flipper commissioning plan. *Journal of Physics: Conference Series*, 295(1):012162, may 2011.

[141] H. Huang, J. Kewisch, C. Liu, A. Marusic, W. Meng, F. Méot, P. Oddo, V. Ptitsyn, V. Ranjbar, and T. Roser. High Spin-Flip Efficiency at 255 GeV for Polarized Protons in a Ring With Two Full Siberian Snakes. *Phys. Rev. Lett.*, 120(26):264804, 2018.

[142] L.H. Yu, E. Bozoki, J. Galayda, S. Krinsky, and G. Vignola. Real time harmonic closed orbit correction. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 284(2):268–285, 1989.

[143] RHIC Spin Group. Rhic polarimetry results by fill. https://www.cnipol.bnl.gov/fills/?rp=15&fn=&ft=&be=&mode=1&sb=Select.

[144] I. Nakagawa et al. RHIC polarimetry. *Eur. Phys. J. ST*, 162:259–265, 2008.

[145] I. Nakagawa et al. p-carbon polarimetry at RHIC. *AIP Conf. Proc.*, 980(1):380–389, 2008.

[146] A. Bravar. The Absolute Polarimeter for RHIC. *AIP Conference Proceedings*, 675(1):830–835, 07 2003.

[147] Elke-Caroline Aschenauer et al. The RHIC Cold QCD Program. 2 2023.

[148] J. G. Alessi, D. Barton, E. Beebe, S. Bellavia, O. Gould, A. Kponou, R. Lambiase, R. Lockey, A. McNerney, M. Mapes, Y. Marneris, M. Okamura, D. Phillips, A. I. Pikin, D. Raparia, J. Ritter, L. Snydstrup, C. Theisen, and M. Wilinski. The Brookhaven National Laboratory electron beam ion source for RHIC. *Review of Scientific Instruments*, 81(2):02A509, 02 2010.

[149] E. Beebe, J. Alessi, S. Binello, T. Kanesue, D. McCafferty, J. Morris, M. Okamura, A. Pikin, J. Ritter, and R. Schoepfer. Reliable operation of the Brookhaven EBIS for highly charged ion production for RHIC and NSRL. *AIP Conference Proceedings*, 1640(1):5–11, 01 2015.

[150] G. Zschornacka, M. Schmidt, and A. Thorn. Electron Beam Ion Sources. In *CAS - CERN Accelerator School: Ion Sources*, pages 165–201, 2013.

[151] Chris Gardner et al. Operation of the RHIC Injector Chain with Ions from EBIS. In *6th International Particle Accelerator Conference*, page THPF046, 2015.

[152] M. Allen et al. PHENIX inner detectors. *Nucl. Instrum. Meth. A*, 499:549–559, 2003.

[153] Clemens Adler, Alexei Denisov, Edmundo Garcia, Michael J. Murray, Herbert Strobele, and Sebastian N. White. The RHIC zero degree calorimeter. *Nucl. Instrum. Meth. A*, 470:488–499, 2001.

[154] Gabor David. Overview of the phenix detector, August 2021.

[155] S. H. Aronson et al. PHENIX magnet system. *Nucl. Instrum. Meth. A*, 499:480–488, 2003.

[156] K. Adcox et al. PHENIX central arm tracking detectors. *Nucl. Instrum. Methods Phys. Res., Sec. A*, 499:489, 2003.

[157] L Aphecetche et al. PHENIX calorimeter. *Nucl. Instrum. Methods Phys. Res., Sec. A*, 499:521, 2003.

[158] M. Aizawa et al. PHENIX central arm particle ID detectors. *Nucl. Instrum. Methods Phys. Res., Sec. A*, 499:508, 2003.

[159] A. Adare et al. Double Spin Asymmetry of Electrons from Heavy Flavor Decays in $p+p$ Collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 87(1):012011, 2013.

[160] Emcal/rich trigger (ert). https://www.phenix.bnl.gov/experiment/trigger.html.

[161] C. Aidala et al. Measurement of charm and bottom production from semileptonic hadron decays in $p + p$ collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. D*, 99(9):092003, 2019.

[162] R Ichimiya, N Apadula, Y Akiba, E Atomssa, S Chollet, O Drapier, H En'yo, K Fujiwara, F Gastaldi, R Granier de Cassagnac, M Kasai, K Kurita, M Kurosawa, E J Mannel, H Ohnishi, Y Onuki, R Pak, C Pancake, M Sekimoto, E Shafto, W Sondheim, and A Taketani. Status and overview of development of the silicon pixel detector for the phenix experiment at the bnl rhic. *Journal of Instrumentation*, 4(05):P05001, may 2009.

[163] Z Li. Novel silicon stripixel detector: concept, simulation, design, and fabrication. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 518(3):738–753, 2004.

[164] R. Nouicer et al. Status and Performance of New Silicon Stripixel Detector for the PHENIX Experiment at RHIC: Beta Source, Cosmic-rays and Proton Beam at 120 GeV. *J. Instrum.*, 4(04):P04011, apr 2009.

[165] U. Acharya et al. Probing Gluon Spin-Momentum Correlations in Transversely Polarized Protons through Midrapidity Isolated Direct Photons in $p\uparrow+p$ Collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 127(16):162001, 2021.

[166] C. Aidala et al. Nonperturbative transverse momentum broadening in dihadron angular correlations in $\sqrt{s_{NN}} = 200$ GeV proton-nucleus collisions. *Phys. Rev. C*, 99(4):044912, 2019.

[167] A. Adare et al. Single electron yields from semileptonic charm and bottom hadron decays in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. C*, 93(3):034904, 2016.

[168] Electromagnetic calorimeter. https://www.phenix.bnl.gov/detectors/emcal.html.

[169] S. S. Adler et al. Measurement of transverse single-spin asymmetries for midrapidity production of neutral pions and charged hadrons in polarized p+p collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 95:202001, 2005.

[170] A. Adare et al. Measurement of Transverse Single-Spin Asymmetries for J/ψ Production in Polarized $p + p$ Collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 82:112008, 2010. [Phys. Rev. D **86**, 099904(E) (2012)].

[171] A. Adare et al. Transverse momentum dependence of J/psi polarization at midrapidity in p+p collisions at s**(1/2) = 200-GeV. *Phys. Rev. D*, 82:012001, 2010.

[172] Shreyasi Acharya et al. Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s}$ = 5.02 TeV with ALICE. *Eur. Phys. J. C*, 79(5):388, 2019.

[173] L. Adamczyk et al. Longitudinal and transverse spin asymmetries for inclusive jet production at mid-rapidity in polarized p+p collisions at $\sqrt{s}$ = 200 gev. *Phys. Rev. D*, 86:032006, Aug 2012.

[174] L. C. Bland et al. Cross Sections and Transverse Single-Spin Asymmetries in Forward Jet Production from Proton Collisions at $\sqrt{s} = 500$ GeV. *Phys. Lett. B*, 750:660, 2015.

[175] C. Adolph et al. First measurement of the Sivers asymmetry for gluons using SIDIS data. *Phys. Lett. B*, 772:854, 2017.

[176] L. Adamczyk et al. Azimuthal transverse single-spin asymmetries of inclusive jets and charged pions within jets from polarized-proton collisions at $\sqrt{s} = 500$ GeV. *Phys. Rev. D*, 97(3):032004, 2018.

[177] C. Aidala et al. Cross section and transverse single-spin asymmetry of muons from open heavy-flavor decays in polarized p+p collisions at $\sqrt{s}$ = 200 GeV. *Phys. Rev. D*, 95(11):112001, 2017.

[178] N. J. Abdulameer et al. Transverse single-spin asymmetry of charged hadrons at forward and backward rapidity in polarized p+p, p+Al, and p+Au collisions at sNN=200 GeV. *Phys. Rev. D*, 108(7):072016, 2023.

[179] C. Aidala et al. Nuclear Dependence of the Transverse Single-Spin Asymmetry in the Production of Charged Hadrons at Forward Rapidity in Polarized $p + p$, p+Al, and p+Au Collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. Lett.*, 123(12):122001, 2019.

[180] C. Aidala et al. Single-spin asymmetry of J/ψ production in $p + p$, p+Al, and p+Au collisions with transversely polarized proton beams at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. D*, 98(1):012006, 2018.

[181] J. Adam et al. Comparison of transverse single-spin asymmetries for forward π^0 production in polarized pp, pAl and pAu collisions at nucleon pair c.m. energy $\sqrt{s_{\rm NN}} = 200$ GeV. *Phys. Rev. D*, 103(7):072005, 2021.

[182] C. Aidala et al. Nuclear Dependence of the Transverse-Single-Spin Asymmetry for Forward Neutron Production in Polarized $p + A$ Collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. Lett.*, 120(2):022001, 2018.

[183] U. Acharya et al. Transverse single spin asymmetries of forward neutrons in p+p, p+Al and p+Au collisions at $\sqrt{s_{NN}} = 200$ GeV as a function of transverse and longitudinal momenta. *Phys. Rev. D*, 105(3):032004, 2022.

[184] T. C. Rogers and P. J. Mulders. No Generalized TMD-Factorization in Hadro-Production of High Transverse Momentum Hadrons. *Phys. Rev. D*, 81:094006, 2010.

[185] T. C. Rogers. Extra spin asymmetries from the breakdown of transverse-momentum-dependent factorization in hadron-hadron collisions. *Phys. Rev. D*, 88(1):014002, 2013.

[186] Z. Akopov et al. Status Report of the DPHEP Study Group: Towards a Global Effort for Sustainable Data Preservation in High Energy Physics, May 2012.

[187] CERN Open Data Policy for the LHC Experiments, Nov 2020. https://cds.cern.ch/record/2745133.

[188] CERN Open Science Policy, 2022. https://cds.cern.ch/record/2835057.

[189] M. D. Wilkinson et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, March 2016.

[190] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[191] Patrick T. Komiske, Ian Moult, Jesse Thaler, and Hua Xing Zhu. Analyzing N-Point Energy Correlators inside Jets with CMS Open Data. *Phys. Rev. Lett.*, 130(5):051901, 2023.

[192] Cari Cesarotti, Yotam Soreq, Matthew J. Strassler, Jesse Thaler, and Wei Xue. Searching in CMS Open Data for Dimuon Resonances with Substantial Transverse Momentum. *Phys. Rev. D*, 100(1):015021, 2019.

[193] M. Andrews, J. Alison, S. An, Patrick Bryant, B. Burkle, S. Gleyzer, M. Narain, M. Paulini, B. Poczos, and E. Usai. End-to-end jet classification of quarks and gluons with the CMS Open Data. *Nucl. Instrum. Meth. A*, 977:164304, 2020.

[194] Aram Apyan, William Cuozzo, Markus Klute, Yoshihiro Saito, Matthias Schott, and Bereket Sintayehu. Opportunities and challenges of Standard Model production cross section measurements in proton-proton collisions at $\sqrt{s}$ =8 TeV using CMS Open Data. *JINST*, 15(01):P01009, 2020.

[195] Patrick T. Komiske, Radha Mastandrea, Eric M. Metodiev, Preksha Naik, and Jesse Thaler. Exploring the Space of Jets with CMS Open Data. *Phys. Rev. D*, 101(3):034009, 2020.

[196] S. Elgammal, M. Louka, A. Y. Ellithi, and M. T. Hussein. Search for the production of dark matter candidates in association with heavy dimuon resonance using the CMS open data for proton–proton collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. Plus*, 138(6):548, 2023.

[197] Patrick T. Komiske, Serhii Kryhin, and Jesse Thaler. Disentangling quarks and gluons in CMS open data. *Phys. Rev. D*, 106(9):094021, 2022.

[198] Saksevul Arias, Eleazar Cuautle, and Hermes León Vargas. Jet fragmentation properties with CMS open-data. *Phys. Scripta*, 98(3):035305, 2023.

[199] Cms open data workshop 2023. https://cms-opendata-workshop.github.io/2023-07-11-cms-open-data-workshop/.

[200] Kati Lassila-Perini, Clemens Lange, Edgar Carrera Jarrin, and Matthew Bellis. Using CMS Open Data in research – challenges and directions. *EPJ Web Conf.*, 251:01004, 2021.

[201] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[202] CERN Open Data Portal. https://opendata.cern.ch/.

[203] LHCb Starterkit. https://lhcb.github.io/starterkit-lessons/index.html.

[204] I. Abt et al. The H1 detector at HERA. 7 1993.

[205] K. Edwards et al. ZEUS, A DETECTOR FOR HERA: LETTER OF INTENT. 6 1985.

[206] HERMES technical design report. 7 1993.

[207] G. A. Voss and B. H. Wiik. The Electron proton collider HERA. *Ann. Rev. Nucl. Part. Sci.*, 44:413–452, 1994.

[208] Dirk Krücker, Karsten Schwank, Patrick Fuhrmann, Birgit Lewendel, and David M. South. Data preservation for the HERA experiments at DESY using dCache technology. *J. Phys. Conf. Ser.*, 664(4):042029, 2015.

[209] David M. South and Michael Steder. The H1 Data Preservation Project. *J. Phys. Conf. Ser.*, 396:062019, 2012.

[210] Janusz Malka and Katarzyna Wichmann. The ZEUS data preservation project. *J. Phys. Conf. Ser.*, 396:022033, 2012.

[211] Georges Aad et al. Operation of the ATLAS trigger system in Run 2. *JINST*, 15(10):P10004, 2020.

[212] Mia Tosi. The CMS trigger in Run 2. *PoS*, EPS-HEP2017:523, 2017.

[213] Roel Aaij et al. Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. *JINST*, 14(04):P04013, 2019.

[214] The LHCb Collaboration. LHCb computing: Technical Design Report. (CERN-LHCC-2005-019; LHCb-TDR-11), 2005. http://cds.cern.ch/record/835156.

[215] The LHCb Collaboration. Upgrade Software and Computing. (CERN-LHCC-2018-007, LHCB-TDR-017), 2018. https://cds.cern.ch/record/2310827.

[216] I. Antcheva et al. ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization. *Comput. Phys. Commun.*, 180:2499–2512, 2009.

[217] DaVinci. https://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[218] Gaudi Project. https://gaudi.web.cern.ch/gaudi/.

[219] Python. https://www.python.org/.

[220] M. Ferrillo. New generation offline software for the LHCb upgrade I. 2022. https://cds.cern.ch/record/2806414.

[221] Chris Burr. Analysis Productions: A declarative approach to ntupling, May 2023. https://indico.jlab.org/event/459/contributions/11589/.

[222] Doxygen. https://www.doxygen.nl/.

[223] Sphinx. https://www.sphinx-doc.org/en/master/.

[224] GitLab. https://git-scm.com/.

[225] Snakemake. https://snakemake.readthedocs.io/en/stable/#.

[226] LHCb Ntuple Wizard frontend. https://gitlab.cern.ch/lhcb-dpa/wp6-analysis-preservation-and-open-data/lhcb-ntuple-wizard-frontend.

[227] https://www.javascript.com/.

[228] https://react.dev/.

[229] React bootstrap. https://react-bootstrap.netlify.app/.

[230] React router dom. https://www.npmjs.com/package/react-router-dom#react-router-dom.

[231] https://www.npmjs.com/.

[232] https://npm.io/package/lhcb-ntuple-wizard.