

Measurements of the $\cos\phi$ and $\cos 2\phi$ Moments of the Unpolarized SIDIS π^+ Cross-section at CLAS12

Richard Capobianco,¹ M. Zurek,² Timothy B. Hayward,¹
Kyungseon Joo,¹ Stefan Diehl,¹ Andrey Kim,¹ and Marshall Scott²

¹*University of Connecticut, Storrs, Connecticut 06269*

²*Argonne National Laboratory, Argonne, Illinois 60439*

(Dated: November 13, 2023)

Semi-inclusive deep inelastic π^+ electroproduction has been studied with the CLAS12 detector at Jefferson Laboratory. Data were taken by Run Group A at Hall B of the laboratory using a polarized 10.6 GeV electron beam, interacting with an unpolarized liquid hydrogen target. The collected statistics enable a high-precision study of the $\cos\phi$ and $\cos 2\phi$ azimuthal moments of the unpolarized cross-sections. These azimuthal moments probe the Boer-Mulders function, which describes the net polarization of quarks inside an unpolarized proton, and the Cahn effect, which has a purely kinematic origin. The high statistics data will, for the first time, enable a multidimensional analysis of both moments over a large kinematic range of Q^2 , y , z , and P_T . We will present the status of this ongoing analysis, including the multidimensional unfolding procedures used for acceptance corrections.

Current Version of Analysis Note: v4
Released on November 13, 2023 for DNP-JPS 2023

Version v3 was released on February 22, 2023 for April APS 2023

Version v2 was released on October 1, 2022 for DNP 2022

Version v1 was released on August 2, 2022 for GRC 2022

CONTENTS

I. Introduction	3
II. Data Collection	6
A. Data Set	6
B. Resources	6
C. PID	6
D. Kinematic Cuts	7
E. Fiducial Cuts	9
F. Momentum Corrections	9
III. Distributions/Binning	14
A. Distributions	14
B. Kinematic Binning	16
IV. Monte Carlo Simulations	21
A. Description	21
B. Data vs MC Comparison	21
C. MC Event Matching	32
D. Momentum Corrections/Smearing	32
V. Analysis Procedure	38
A. Description	38
B. Acceptance Corrections	39
1. Acceptance Closure Tests	47
C. Other Corrections	51
D. Cosine Moment Extractions	51
VI. Plots Requested for Release	55
A. Images already shown in this Analysis Note	55
B. Additional Images to be Released	56
VII. Version Update Notes	67
VIII. Appendix	68
References	70

I. INTRODUCTION

The study of semi-inclusive deep inelastic scattering (SIDIS) experiments offers opportunities to address questions about the three-dimensional structure of nucleons via their relationship to transverse momentum-dependent parton distribution functions (TMDs). This analysis aims to study TMDs related to the unpolarized SIDIS cross-section for the positively charged pion, particularly the Boer-Mulders and Cahn effects, in order to gain a greater understanding of the transverse motion of quarks and gluons inside the proton.

SIDIS reactions are described by $l(\ell) + N(P) \rightarrow l'(\ell') + h(P_h) + X(P_X)$ where an initial lepton beam (l) scatters off of a target nucleon (N) to produce a scattered lepton (ℓ') and a single hadron (h). All other particles that may exist in the final state (X) are ignored. In this analysis, the initial lepton is a polarized electron while the nucleon is an unpolarized liquid hydrogen (i.e., a proton) target. In the final state, at least one scattered electron and π^+ pion are detected. Here, ℓ , P , ℓ' , P_h , and P_X are the 4-momenta of their respective particles.

From this reaction, several important kinematic variables used in this analysis include, but are not limited to:

$$Q^2 = -q^2 \quad (1)$$

$$\nu = \frac{q \cdot P}{M} = E - E' \quad (2)$$

$$y = \frac{P \cdot q}{P \cdot \ell} = \frac{\nu}{E} \quad (3)$$

$$x_B = \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2M\nu} = \frac{Q^2}{2MEy} \quad (4)$$

$$z = \frac{P \cdot P_h}{P \cdot q} = \frac{E_h}{\nu} \quad (5)$$

$$W^2 = (P + q)^2 \quad (6)$$

$$x_F = \frac{2P_h \cdot q}{|q|W} \quad (7)$$

$$\gamma = \frac{2Mx_B}{Q} \quad (8)$$

$$\epsilon = \frac{1 - y - \frac{1}{4}\gamma^2y^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2y^2} \quad (9)$$

The exchanged virtual photon is defined as $q = \ell - \ell'$ such that $Q^2 = -q^2$ is the hard scale of the process (the virtuality of the exchanged photon). If the electron beam has energy E and the scattered electron has energy E' then ν is defined as the difference between these two quantities. The variables y , x_B , and z are the fraction of beam energy transferred to the virtual photon, the fraction of target momentum carried by the struck quark, and the fraction of virtual photon energy carried by the hadron system. x_B here is therefore also known as “x-Bjorken”. W is the mass of the virtual photon target system while the “x-Feynman” variable (x_F) can be used to differentiate between current and target fragmentation and takes a positive value if the outgoing hadron moves in the same direction as the incoming electron in the struck quark center-of-mass frame. The quantity γ

describes the relationship between the energy transferred to the struck quark and the virtual photon energy while ϵ is the ratio of the longitudinal and transverse virtual photon flux.

Additionally, other important variables include P_T , which is the transverse momentum of the final state hadron, and ϕ_h , which is the angle between the lepton scattering plane and the hadron production plane, according to the “Trento-convention”.^[1] Both of these variables can be directly measured in the target rest frame where the q -vector is along the z-axis (shown in Figure 1).

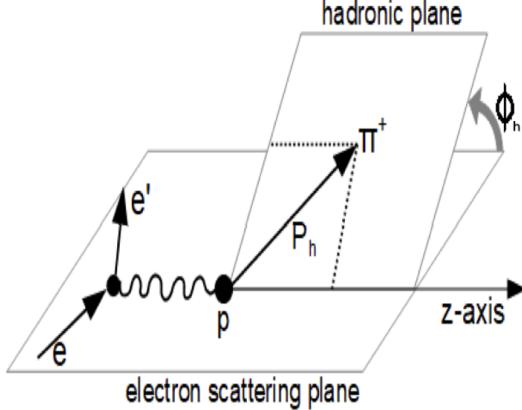


FIG. 1: Definition of the reaction kinematics of single pion SIDIS. Shows definition of ϕ_h . P_T is the perpendicular component of P_h with respect to the z-axis. This is the same center-of-mass frame in which x_F is also measured.

These variables help form the basis of this analysis, being in some way related to either a kinematic cut or the cross-section equation, though the variables Q^2 , x_B , y , z , P_T , and ϕ_h are the most relevant as they are the variables with which the differential cross-section (and therefore, the Boer-Mulders and Cahn effects) will be studied. The variable y and x_B can be used interchangeably in this respect, with the original intent of this analysis being to take measurements with respect to x_B instead of y , however, the analytical procedure has been slightly modified from the initial intended approach to substitute the kinematic binning of x_B with that of y . The reasons and implications of this substitution will be addressed in the Analysis Procedure section of this analysis note.

As this experiment involves an unpolarized target, by assuming single photon exchange and averaging over the beam polarization, the lepton-hadron SIDIS cross-section can be written as (see [2]):

$$\frac{d^5\sigma}{dydQ^2dzd\phi_h dP_{h\perp}^2} = A_0(1 + A_{UU}^{\cos\phi_h} \cos\phi_h + A_{UU}^{\cos 2\phi_h} \cos 2\phi_h) \quad (10)$$

where:

$$A_0 = \frac{\pi\alpha^2}{Q^2(1-\epsilon)}(1 + \frac{\gamma^2}{2x_B})(F_{UU,T} + \epsilon F_{UU,L}) \quad (11)$$

$$A_{UU}^{\cos\phi_h} = \frac{\sqrt{2\epsilon(1+\epsilon)}F_{UU}^{\cos\phi_h}}{(F_{UU,T} + \epsilon F_{UU,L})} \quad (12)$$

$$A_{UU}^{\cos 2\phi_h} = \frac{\epsilon F_{UU}^{\cos 2\phi_h}}{(F_{UU,T} + \epsilon F_{UU,L})} \quad (13)$$

Here the functions F are the structure functions with the first subscript representing the beam polarization, the second subscript representing the target polarization, and the third subscript (if

⁵⁴ there is one) representing the virtual photon polarization. Within these subscripts, the notations U ,
⁵⁵ L , and T are short for unpolarized, longitudinally polarized, and transversely polarized, respectively.

⁵⁶ At leading twist and next to leading twist, and by neglecting the higher twists (twists greater than
⁵⁷ or equal to three can be taken to be zero in the Wandzura-Wilezek approximation), these structure
⁵⁸ functions can be written as [2, 3]:

$$F_{UU}^{\cos 2\phi_h} = \mathcal{C} \left[-\frac{2(\hat{\mathbf{h}} \cdot \mathbf{p}_T)(\hat{\mathbf{h}} \cdot \mathbf{k}_T) - \mathbf{k}_T \cdot \mathbf{p}_T h_1^\perp H_1^\perp}{MM_h} + \dots \right] \quad (14)$$

$$F_{UU}^{\cos \phi_h} = \frac{2M}{Q} \mathcal{C} \left[-\frac{\hat{\mathbf{h}} \cdot \mathbf{k}_T}{M_h} x h H_1^\perp - \frac{\hat{\mathbf{h}} \cdot \mathbf{p}_T}{M} f_1 D_1 + \dots \right] \quad (15)$$

⁵⁹ where:

\mathbf{p}_T = the transverse momentum of the struck quark

\mathbf{k}_T = the transverse part of \mathbf{P}_h with respect to the quark direction

$\mathbf{P}_{h\perp}$ = the transverse part of \mathbf{P}_h with respect to the virtual photon momentum

$$\hat{\mathbf{h}} = \frac{\mathbf{P}_{h\perp}}{|\mathbf{P}_{h\perp}|}$$

M = Mass of Nucleon

M_h = Mass of Hadron

x = Momentum fraction

h = Unpolarized twist-3 TMD

h_1^\perp = Boer-Mulders function

H_1^\perp = Collins function

f_1 = Unpolarized TMD

D_1 = Unpolarized Fragmentation Function

⁶⁰ The primary goal of this analysis is to measure the azimuthal modulations in the unpolarized SIDIS
⁶¹ cross-section in order to study the Boer-Mulders and Cahn effects in terms of Q^2 , y , z , and P_T .¹ The
⁶² Boer-Mulders effect will be the main focus, as it is particularly interesting due to its sensitivity to the
⁶³ correlation between the quark's transverse momentum and intrinsic transverse spin in an unpolarized
⁶⁴ nucleon. The Cahn effect, meanwhile, is a kinematic effect that is also sensitive to the transverse
⁶⁵ motion of quarks and will be considered in this analysis as it is related to the $\frac{2M}{Q} \mathcal{C} \left[\frac{\hat{\mathbf{h}} \cdot \mathbf{p}_T}{M} f_1 D_1 \right]$ term
⁶⁶ in the structure function given above.² It is through the understanding of these effects that this
⁶⁷ analysis aims to provide a greater understanding the three-dimensional structure of the proton.

¹ y can be considered interchangeable here with the variable x_B .

² The Cahn effect is not itself a TMD function.

II. DATA COLLECTION

A. Data Set

For this analysis, data was collected using the CLAS12 detector at Jefferson Laboratory (JLab), whose upgrades from the CLAS detector have improved the facility's ability to conduct experiments with higher energies and statistics. The high luminosity and wide acceptance provided by the CLAS12 detector allow this study to be performed for the first time at JLab as a multidimensional analysis in terms of the kinematic variables detailed in the [Introduction](#).

The data under examination was collected during experiments conducted by Run Group A (RG-A) in the Fall of 2018. In this experiment, a 10.6 GeV polarized electron beam was collided with an unpolarized liquid hydrogen target. The currently analyzed data corresponds to the **Pass 1** calibrations of the **Inbending** runs from the Fall 2018 data set. Only events from the forward detector of CLAS12 (which contains a tracking subsystem that consists of drift chambers in a toroidal magnetic field and high and low threshold Cherenkov counters to identify the scattered electron and final state hadrons, respectively) are being included from these files for this analysis.

Preparations that are being made by the CLAS Collaboration to improve the event reconstruction calibrations of this data are expected to allow for the dropping of the forward detector requirements before this analysis is complete. The introduction of data collected from the central detector is planned to be investigated upon the release of the new Pass 2 calibration files, along with any elements of this analysis that may need to be revisited due to the changes in the data files.

B. Resources

The event selection procedures to identify $ep \rightarrow e'\pi^+(X)$ events for this analysis largely follow procedures developed and described by previously approved and published RG-A first experiment SIDIS analyses (inclusive π^+ and inclusive $\pi^+\pi^-$ papers) [4, 5]. In addition to these we make specific use of the following analysis notes:

- RG-A Common Analysis: https://clas12-docdb.jlab.org/DocDB/0009/000949/001/RGA_Analysis_Overview_and_Procedures-08172020.pdf
- “Back-to-back” (b2b) π^+ -proton analysis: <https://clas12-docdb.jlab.org/DocDB/0009/000935/007/v3%20to%20be%20submitted%20to%20the%20reviewers.pdf>

C. PID

The RG-A Analysis Overview and Procedures note, pg. 58 [6] goes into significant detail about the common particle identification scheme used for RG-A. The first layer of identification is the CLAS12 EventBuilder which is used to associate detector responses from various CLAS12 subsystems to particles via an identification protocol. The resulting information is output to dedicated data structures, called HIPO banks, for further physics analysis. This first level of particle assignment by

⁹⁴ the EventBuilder is relatively loose and significant refinement was added on top of it to improve the
⁹⁵ PID quality. The identification schema described in the following sections corresponds to the final
⁹⁶ selection criteria used in this analysis.

⁹⁷ A full description of the electron PID used here with numerous supporting plots can be found in
⁹⁸ the b2b analysis note starting on pg. 19. The criteria are as follows:

- Forward Detector
- > 2 photoelectrons detected in the HTCC (b2b note Fig. 10)
- > 0.07 GeV energy deposited in the PCAL (b2b note Fig. 11)
- sector dependent sampling fraction cut (b2b note Fig. 12)
- “diagonal cut” for electrons above 4.5 GeV (HTCC threshold) (b2b note Fig. 14)
- $y < 0.75$, not strictly an “electron cut”, but sets the minimum electron energy approximately greater than 2.4 GeV

⁹⁹ A full description of the π^+ PID used here with numerous supporting plots can be found in the
¹⁰⁰ b2b analysis note starting on pg. 25. The criteria are as follows:

- Forward Detector
- $p > 1.25$ GeV
- Refined chi2pid cuts. The corrected standard deviation σ of chipid is set to 0.88 for π^+ s.
 Lower bound of the cut is $\text{chi2pid} > -3\sigma$. Upper bound below 2.44 GeV is $\text{chi2pid} < 3\sigma$
 and above is set at 1/2 the distance in β between kaons and pions, by requiring $\text{chi2pid} < \sigma(0.00869 + 14.98587 \cdot e^{-p/1.18236} + 1.81751 \cdot e^{-p/4.86394})$.

D. Kinematic Cuts

¹⁰¹ Once an electron-pion pair is detected and passed through the PID cuts above, cuts to $W < 2$ GeV
¹⁰² and $Q^2 < 2$ GeV² are then made to select events within the SIDIS region of the data, as shown in
¹⁰³ Figure 2. The Q^2 cut is specifically designed to occur after the peak in its distribution to help avoid
¹⁰⁴ acceptance edge effects and facilitate optimal data to Monte Carlo matching.³

¹⁰⁵ To improve the selection of pion events and strengthen the forward detector requirements, further
¹⁰⁶ cuts are applied to the momentum of the π^+ pion and the polar angles of both particles, building
¹⁰⁷ upon the PID requirements outlined in the preceding section. These cuts are shown in Figure 3,
¹⁰⁸ with the cut to the pion momentum being $1.25 \text{ GeV} < p_{\pi^+} < 5 \text{ GeV}$ while the polar angle cut of
¹⁰⁹ $5^\circ < \theta < 35^\circ$ is applied to both particles identically.

¹¹⁰ Lastly, to mitigate contamination from the target fragmentation region and exclusive meson pro-
¹¹¹ duction, as well as to restrict radiative effects, cuts are implemented for $x_F > 0$, $M_X > 1.5$ GeV, and
¹¹² $y < 0.75$, respectively. These cuts are shown in Figure 4 below.

³ Prior Q^2 cuts made in some of the referenced notes tended to be smaller than the cut in this analysis.

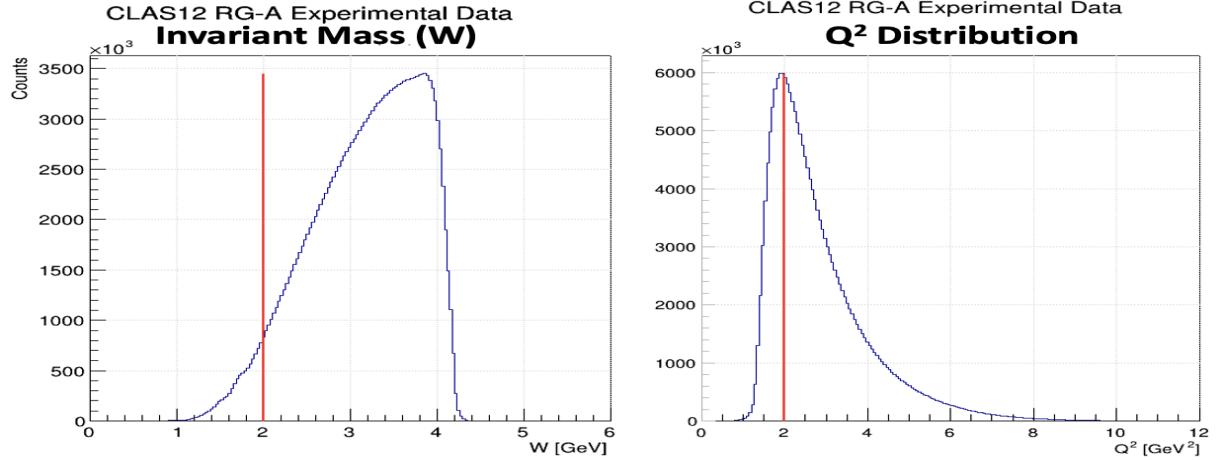


FIG. 2: W and Q^2 cuts shown as the red lines in the plots above.

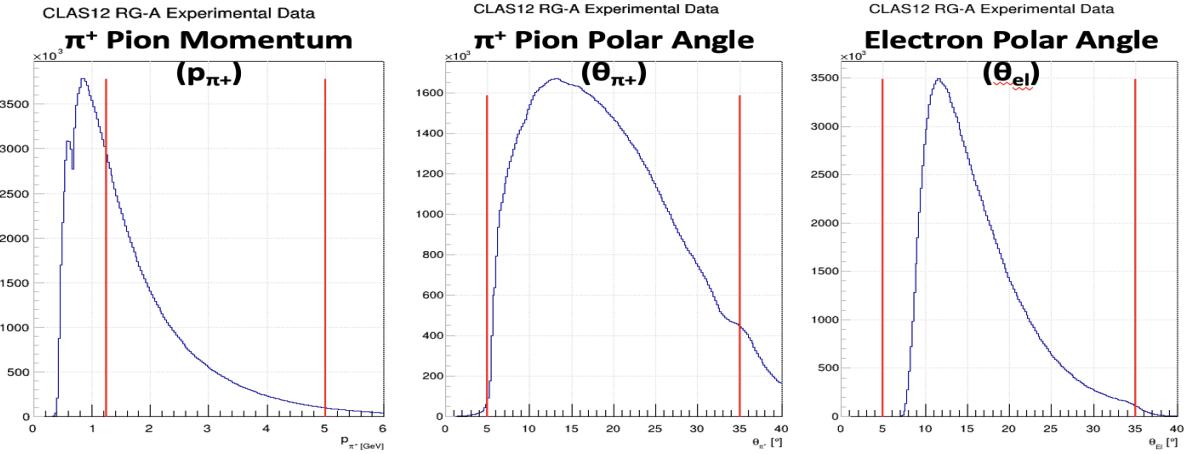


FIG. 3: p_{π^+} , θ_{π^+} , and θ_{El} cuts shown as the red lines in the plots above.

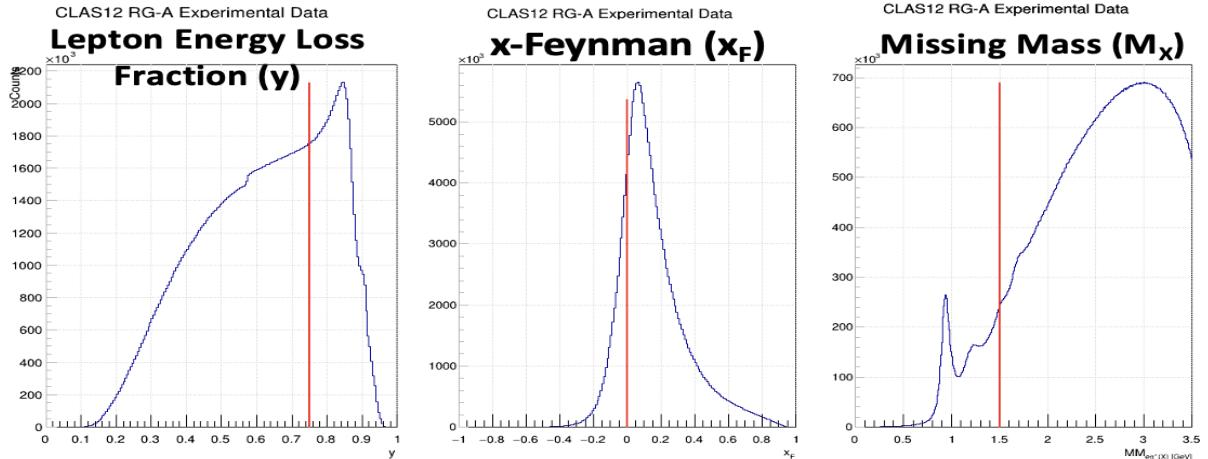


FIG. 4: y , x_F , and M_X cuts shown as the red lines in the plots above.

E. Fiducial Cuts

A detailed discussion and explanation of the fiducial cuts placed on the calorimeter and drift chambers can be found in the RG-A Common Analysis note, pg. 70 [6] or on pg. 29 of the b2b note. Additional fiducial cuts developed by another graduate student (Valerii Klimenko) were added to remove “bad detectors” from the experimental data. The effects of these cuts are illustrated in Figure 5 below.

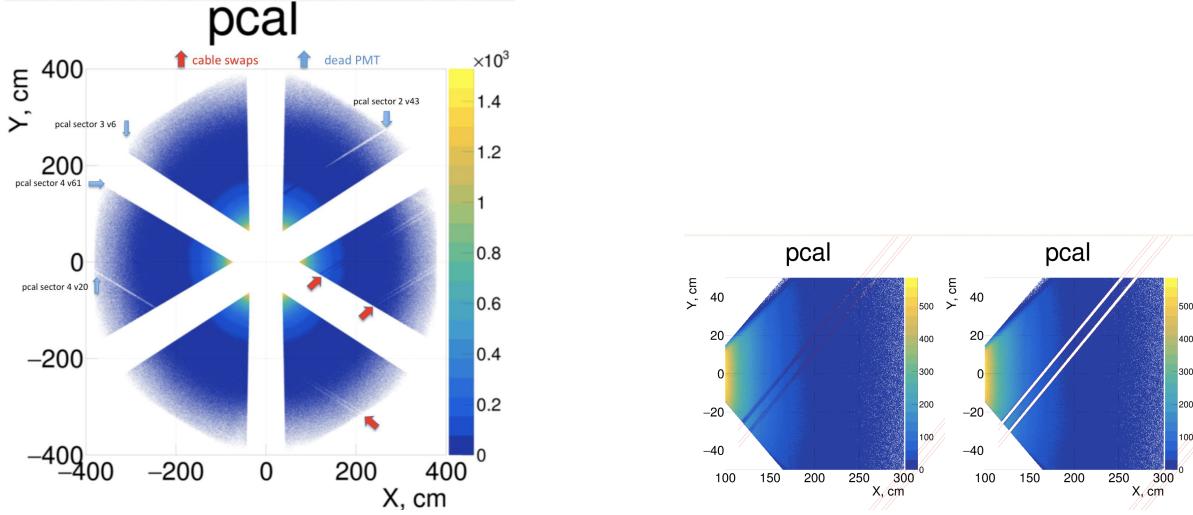


FIG. 5: PCAL stands for the Preshower Calorimeter. Code detailing exact cuts is included in the [Appendix](#) (Images and Code developed by Valerii Klimenko)

F. Momentum Corrections

Particle momentum corrections were also applied to the electron and π^+ pions in the selected events. These corrections were developed in a separate analysis which used exclusive events to correct the Missing Mass spectrum in the exclusive peak region of the $ep \rightarrow e\pi^+ N$ reaction. Using exclusive reactions with well-understood kinematics, it is possible to calculate the “correct” momentum for one particle based on its angle and the other particle’s kinematics using 4-vector momentum conservation. For the corrections applied in this analysis, both the electrons and π^+ pions were corrected with the exclusive $ep \rightarrow e\pi^+ N$ reactions, with the electron corrections also incorporating events from elastic scattering reactions to extend the kinematic range for which the corrections could be applied. The coverages of these corrections are shown below in Figures 6 to 8.

The event selection process to create these momentum corrections follows the standard RG-A procedures described in the earlier sections of this note. However, event selection procedures for the momentum correction did not include the additional fiducial cuts developed by Valerii Klimenko (see Figure 5). Events selected to create the momentum corrections also used a different set of Missing Mass/Invariant Mass cuts to plot events in the region of the exclusive/elastic peaks (i.e., the Neutron/Proton Mass peak). The other additional kinematic cuts described in the [Kinematic Cuts](#) section of this note were also not used when the momentum corrections described in this section were made.

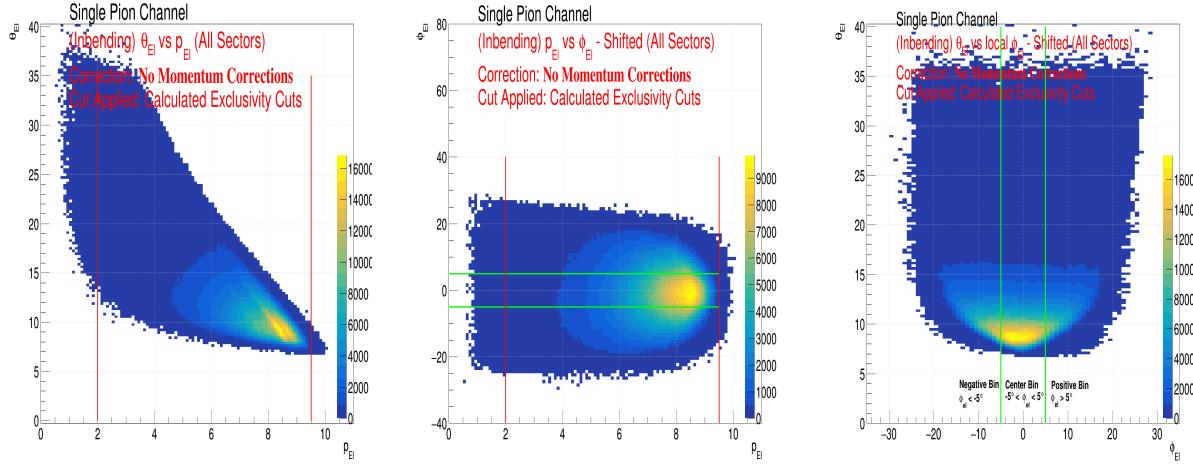


FIG. 6: Exclusive $ep \rightarrow e\pi^+N$ Reactions: Electron Momentum Range used for creating the corrections: $2.0 \text{ GeV} < p_{El} < 9.5 \text{ GeV}$ (see red lines). The green lines show the boarders for the ϕ_{El} bins ($\pm 5^\circ$)

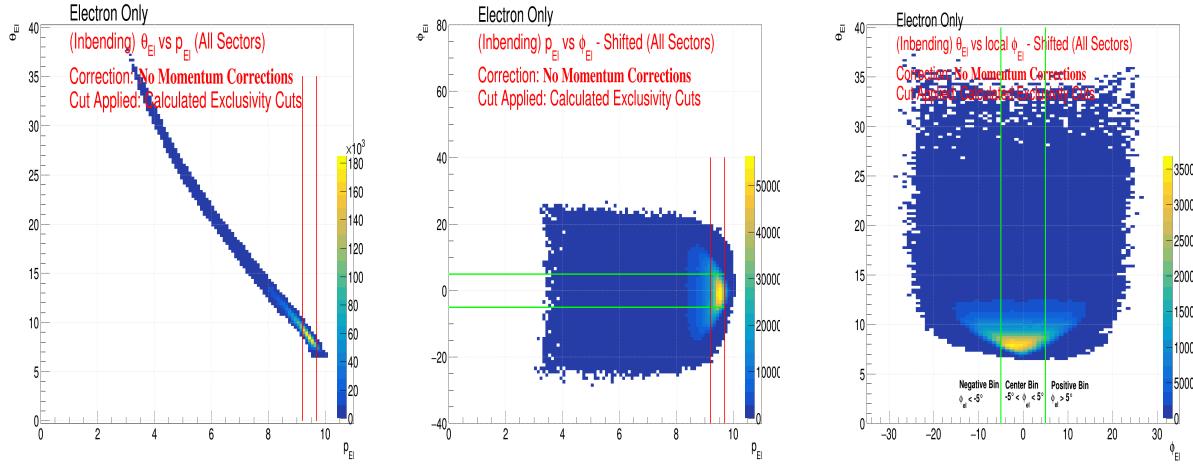


FIG. 7: Elastic Scattering Reactions: Electron Momentum Range used for creating the corrections: $9.2 \text{ GeV} < p_{El} < 9.7 \text{ GeV}$ (see red lines). The green lines show the boarders for the ϕ_{El} bins ($\pm 5^\circ$)

The corrections were obtained by comparing the calculated momentum with the experimentally measured value to plot $\Delta P_{\text{Particle}} = p_{\text{Calculated}} - p_{\text{Measured}}$ vs. p_{Measured} in different regions of the particle's azimuthal lab angle (ϕ). The final results of this procedure were independent momentum corrections for both particles that were sector-dependent functions of p_{Measured} and ϕ . These corrections were applied to the experimental data sample after the initial event selection cuts were applied. Figures 9 to 10 show the **Uncorrected** Missing Mass versus the particle momentum of both particles for all six of the detector's sectors and in the three ϕ regions used to develop the correction. Figures 11 to 12 meanwhile show the same distributions after the final corrections have been applied. For more details on these corrections, see **CLAS12 Momentum Corrections**. [7]

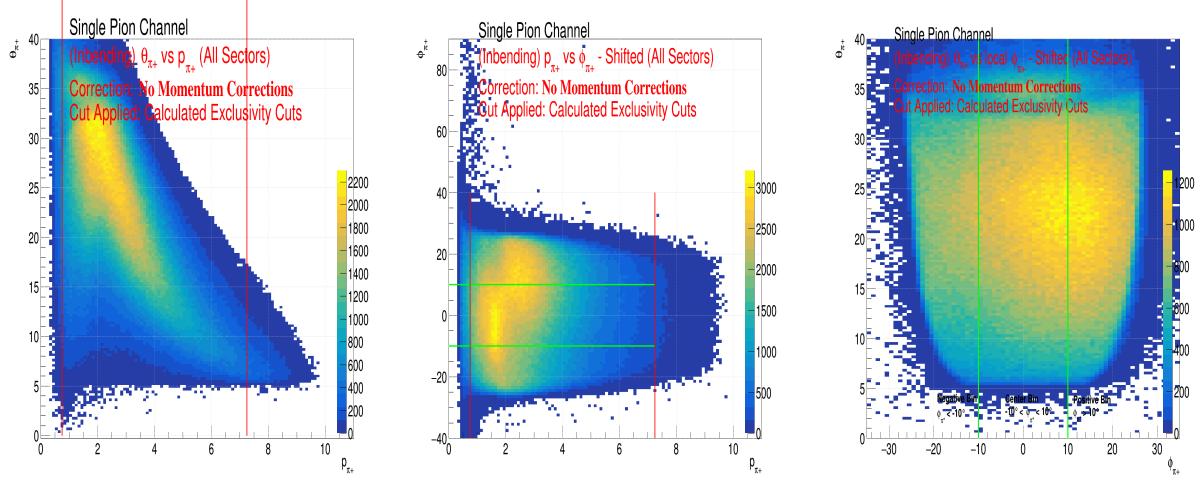


FIG. 8: Exclusive $ep \rightarrow e\pi^+N$ Reactions: π^+ Pion Momentum Range used for creating the corrections: $0.75 \text{ GeV} < p_{\pi^+} < 7.25 \text{ GeV}$ (see red lines). The green lines show the boarders for the ϕ_{π^+} bins ($\pm 10^\circ$)

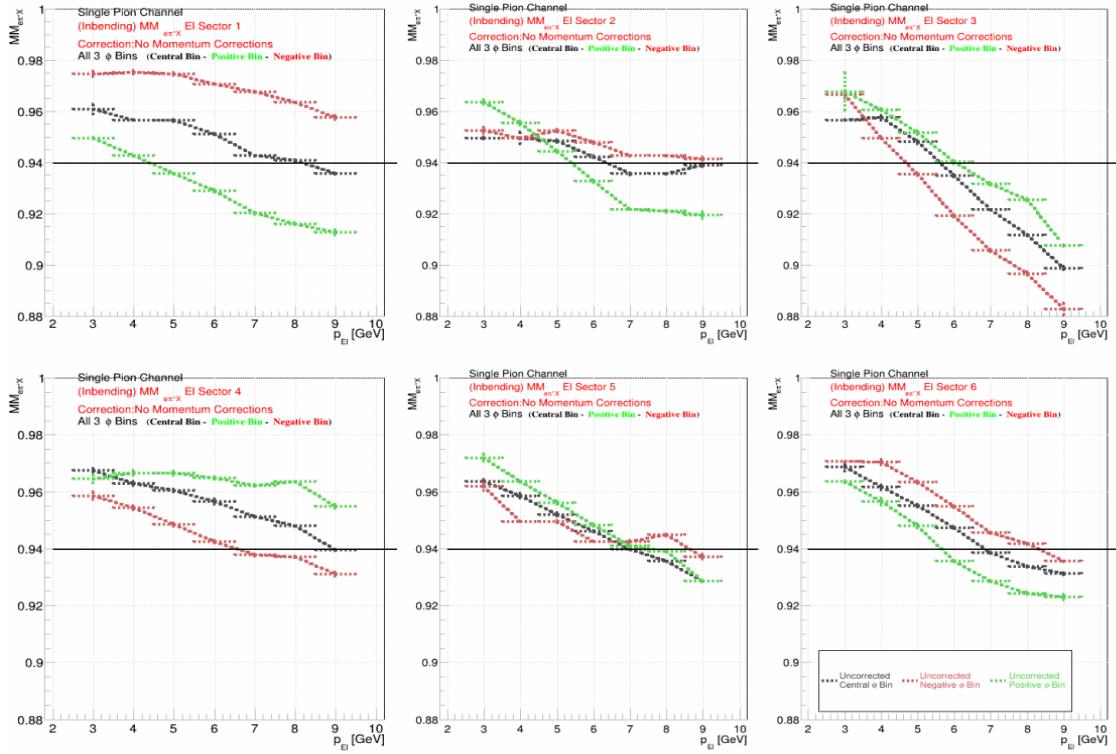


FIG. 9: Missing Mass vs Electron Momentum (Without Momentum Corrections)

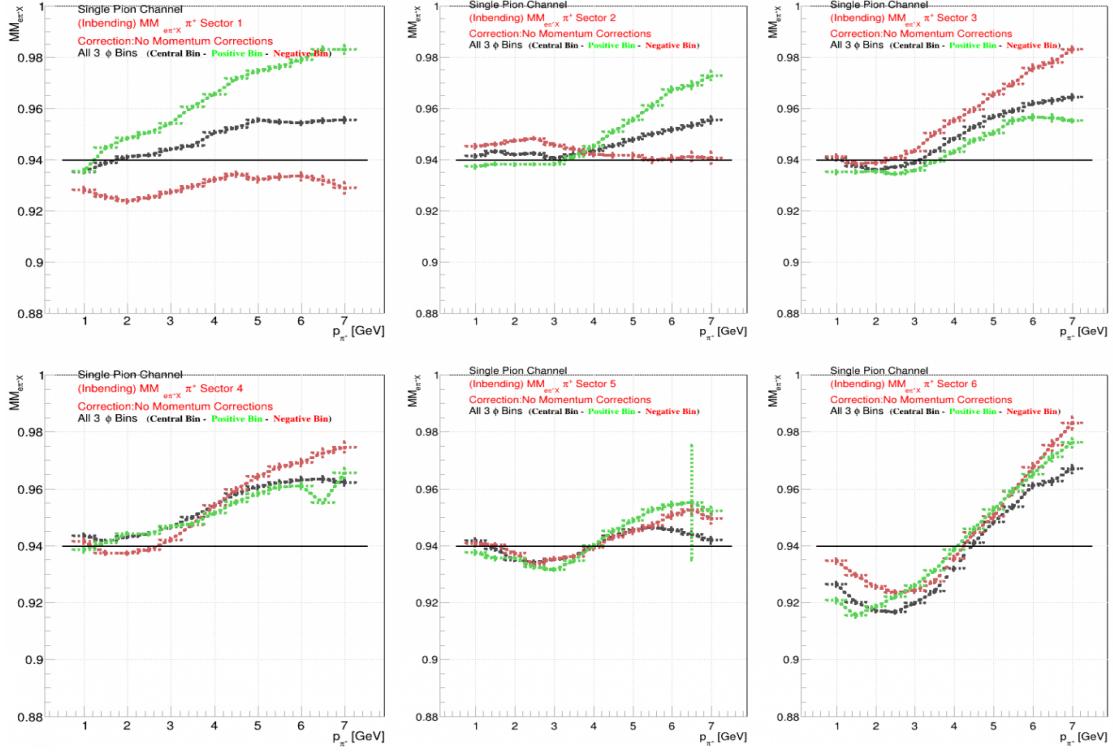
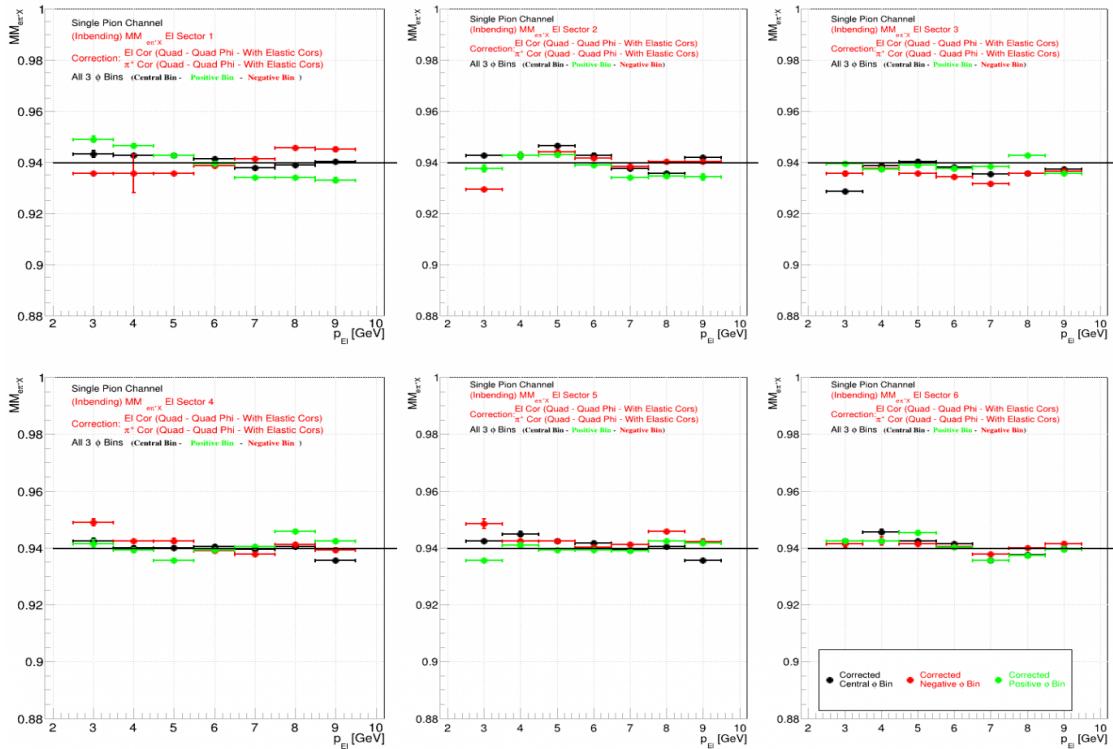
FIG. 10: Missing Mass vs π^+ Pion Momentum (Without Momentum Corrections)

FIG. 11: Missing Mass vs Electron Momentum (With Momentum Corrections)

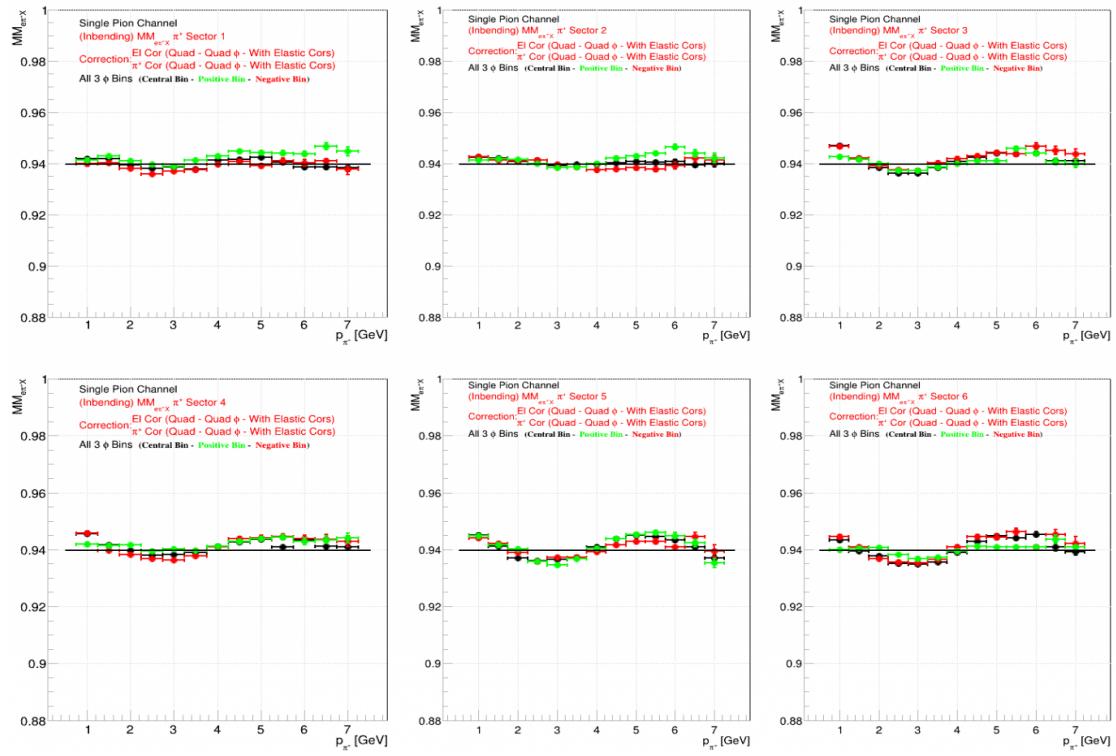


FIG. 12: Missing Mass vs π^+ Pion Momentum (With Momentum Corrections)

III. DISTRIBUTIONS/BINNING

A. Distributions

The following histograms in [Figure 13](#) show the 2D kinematic distributions of the distributions of Q^2 vs y , z vs P_T , and Q^2 vs x_B after all analysis cuts from the [Data Collection](#) section are applied. The figure shows the current Q^2 - y binning scheme as well as the old Q^2 - x_B scheme used in prior versions of this analysis note which are distinct from one another (See [Kinematic Binning](#) below for more details on this change).

[Figure 14](#) shows the distributions of both particle momenta in this same region. [Figure 15](#) shows the comparison between the particle kinematics in the experimental data and the Monte Carlo files with the cuts described throughout the previous sections as well as with the exclusive cuts used while testing the Monte Carlo smearing functions.⁴ See the [Monte Carlo Simulations](#) section for more details on the simulated data.

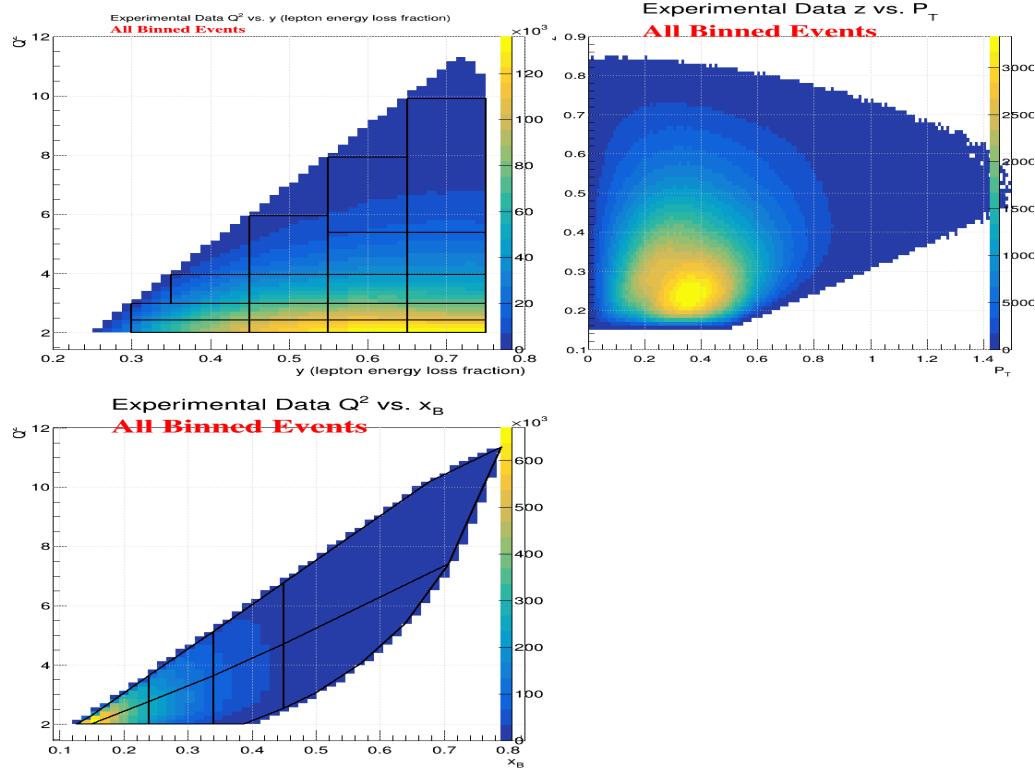


FIG. 13: 2D histograms of Q^2 vs. y (top left), z vs. P_T (top right), and Q^2 vs. x_B (bottom left) for all events that survive the cuts described in the [Data Collection](#) section of this analysis note

⁴ The cuts described here as “used while testing the Monte Carlo smearing functions” are the same as those described throughout the [Data Collection](#) section, but with the Missing Mass cuts being replaced with those used for the exclusive momentum corrections mentioned in the [Momentum Corrections](#) section.

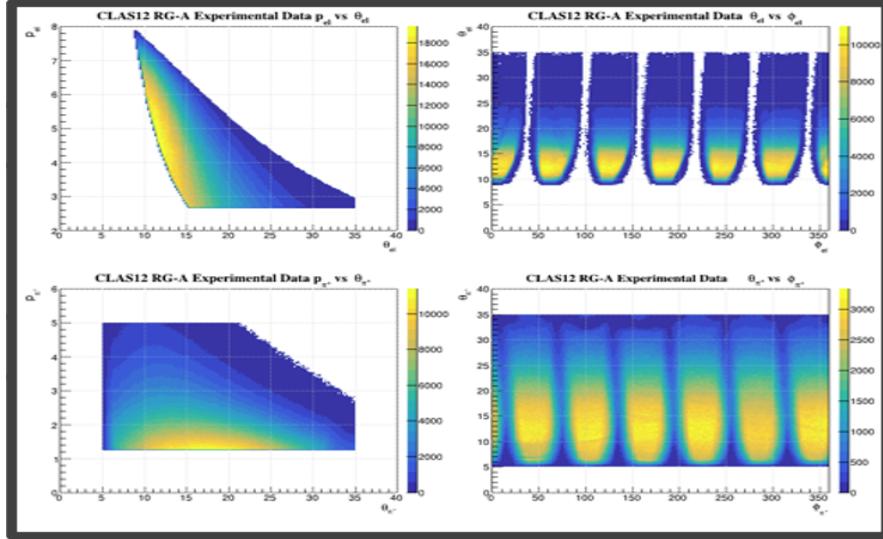
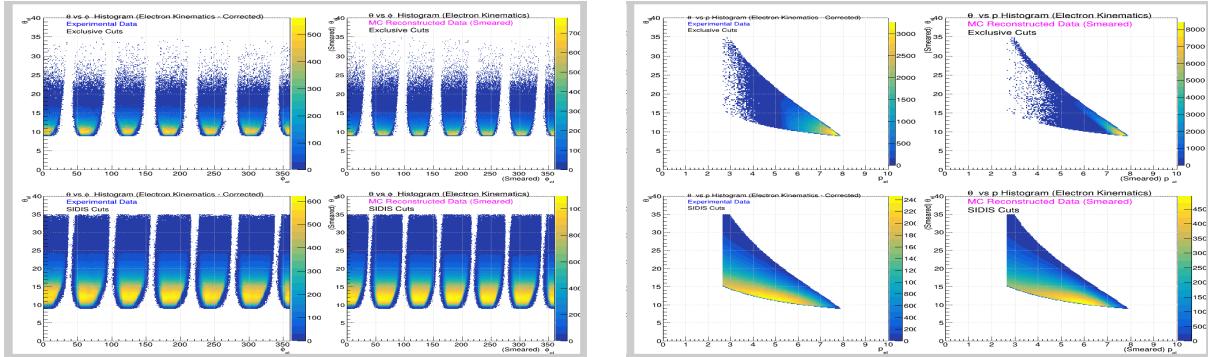
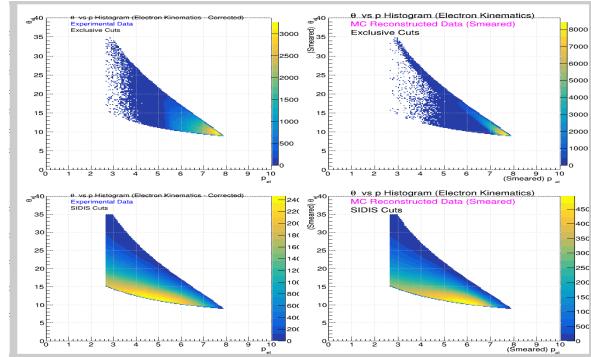


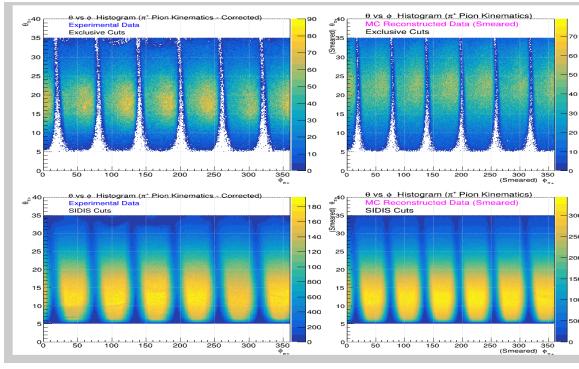
FIG. 14: 2D histograms of p vs. θ (left plots) and θ vs. ϕ (right plots) for the electron (top row) and π^+ pion (bottom row) kinematics for all events that survive the SIDIS cuts described in this analysis note



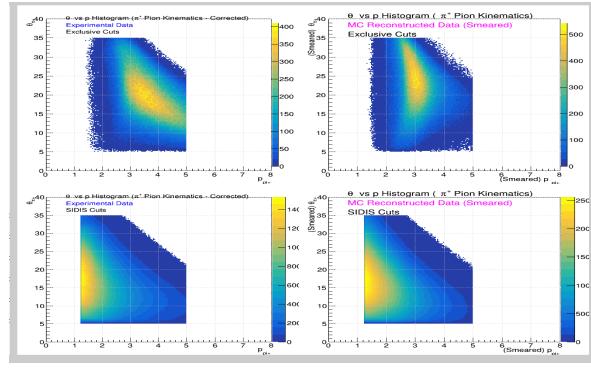
(a) ELECTRON Kinematics: θ_{el} vs. ϕ_{el} Plots



(b) ELECTRON Kinematics: θ_{el} vs. p_{el} Plots



(c) π^+ Pion Kinematics: θ_{π^+} vs. ϕ_{π^+} Plots



(d) π^+ Pion Kinematics: θ_{π^+} vs. p_{π^+} Plots

FIG. 15: Comparisons of the particle kinematics in the Data and Monte Carlo. Plots on the top rows use the **Exclusive cuts** used with the Momentum Correction/Smearing procedures. The bottom rows use the **SIDIS cuts** as described in this analysis note. The left plots show the **Experimental data** while the right plots show the smeared **Reconstructed Monte Carlo**

B. Kinematic Binning

The kinematic binning used by this analysis currently consists of 17 Q^2 - y kinematic bins for which there exists anywhere between 4 to 7 bins of z and of P_T . Combined, this means that there are anywhere from 20 to 42 z - P_T bins for any given Q^2 - y bins. As mentioned in the [Introduction](#), the use of the variable y instead of x_B was observed to have certain advantages in the analytical procedure, as it is proving to be easier to fit distributions of Q^2 vs. y with rectangular bins over the Q^2 vs. x_B distribution.

To illustrate the difference between the Q^2 vs. y and the Q^2 vs. x_B distribution, [Figure 13](#) shows a comparison of these two schemes with the rectangular Q^2 - y bins being better situated to fitting the data without the use of non-rectangular bins as shown in the Q^2 - x_B bins which had been used in a prior multidimensional analysis of π^+ Beam Spin Asymmetries (BSAs).⁵ [5] Regardless of which binning scheme is used, the z - P_T binning schemes are always defined as rectangular grids with unique designs based on the z vs. P_T distributions within a given Q^2 - y (or Q^2 - x_B) bin. The binning schemes used are shown in [Figures 16 to 20](#), with the schemes being designed to (ideally) give a relatively consistent amount of statistics per bin when possible.⁶

The ϕ_h distributions are plotted within each of the combined Q^2 - y - z - P_T bins as part of this analysis procedure (see the [Analysis Procedure](#) section below). This variable is currently being plotted with 24 equally sized bins from 0° to 360° (i.e., 15° per bin).

Further development of these binning schemes may still prove relevant as ensuring an optimized binning scheme will be necessary for taking the highest quality measurements that the experiment will allow.

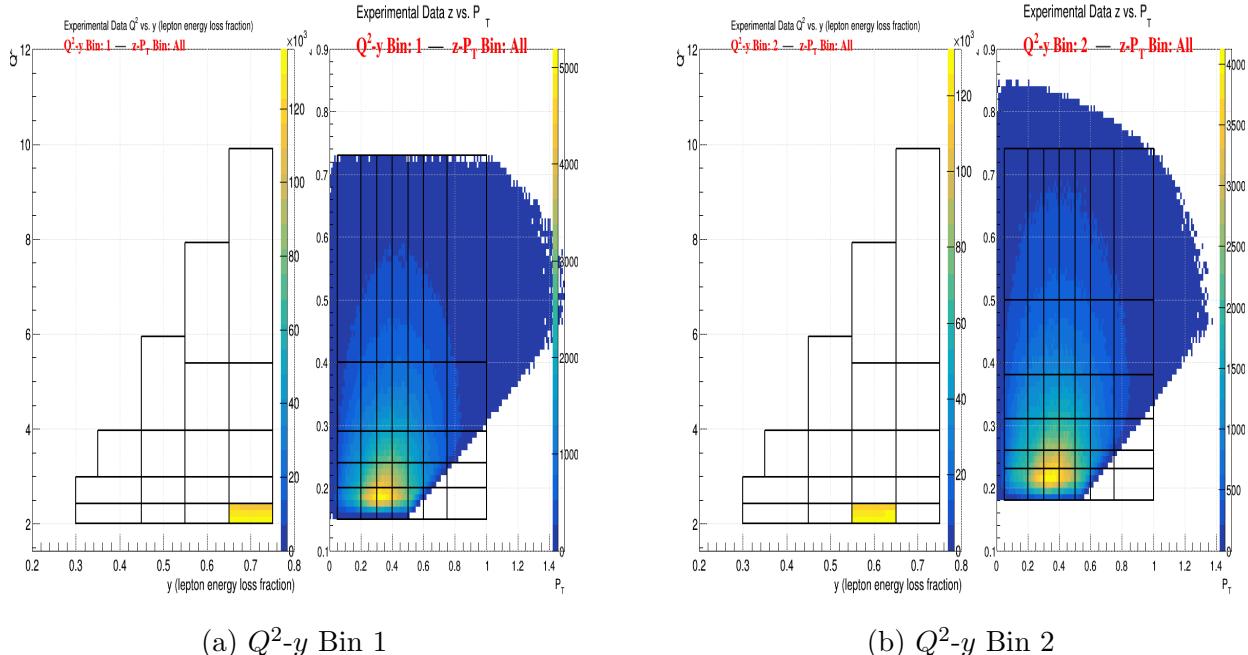


FIG. 16: 2D Kinematic Binning Schemes for Q^2 - y Bin 1 and 2

⁵ Modifications were made to the Q^2 - x_B binning scheme that was being used in this analysis as a result of a larger cut on Q^2 , so depictions of that scheme in this analysis note will differ from the one they were based on.

⁶ For a similar depiction of the old Q^2 - x_B binning scheme used in prior versions of this analysis note, see [Figure 77](#).

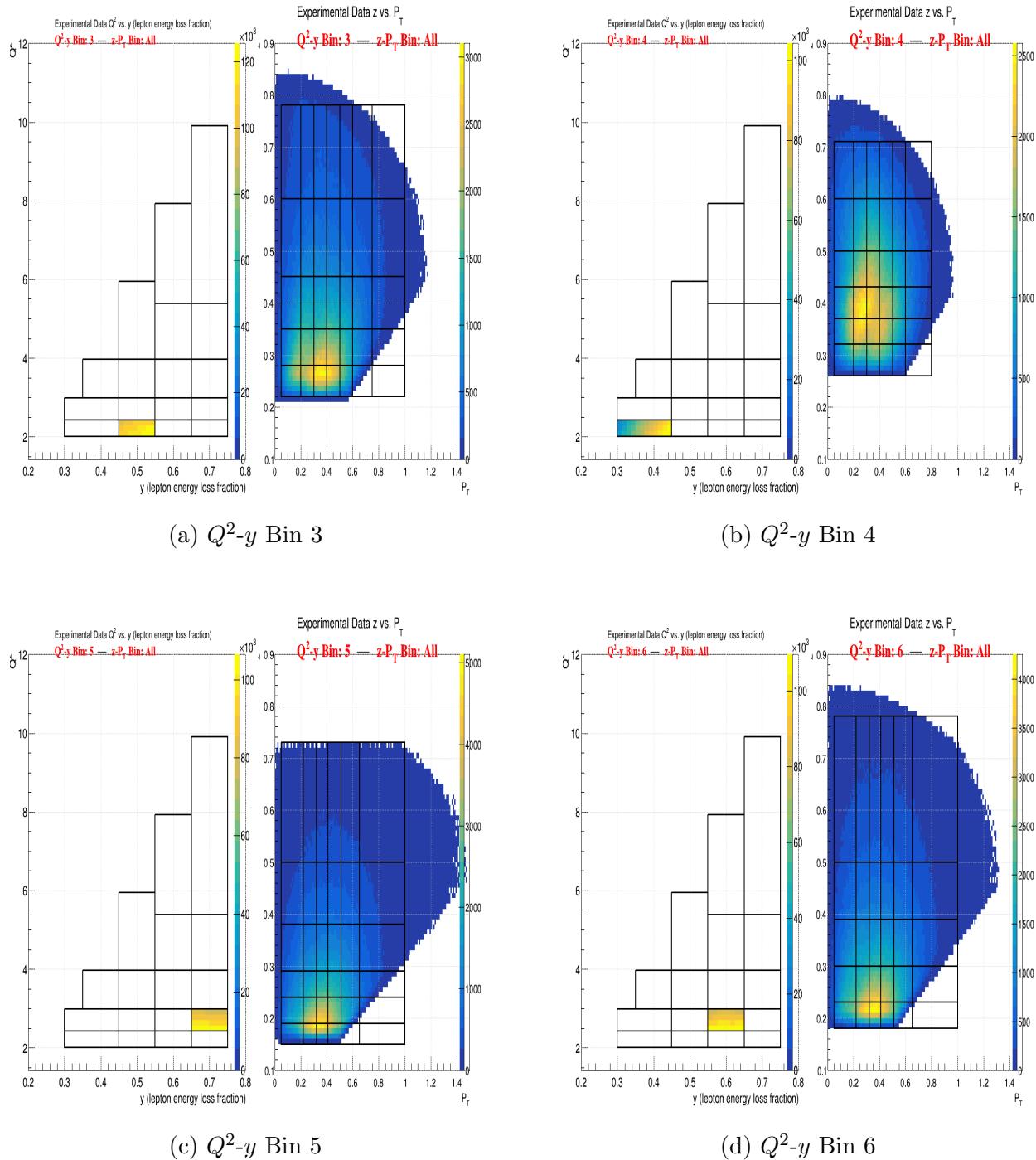
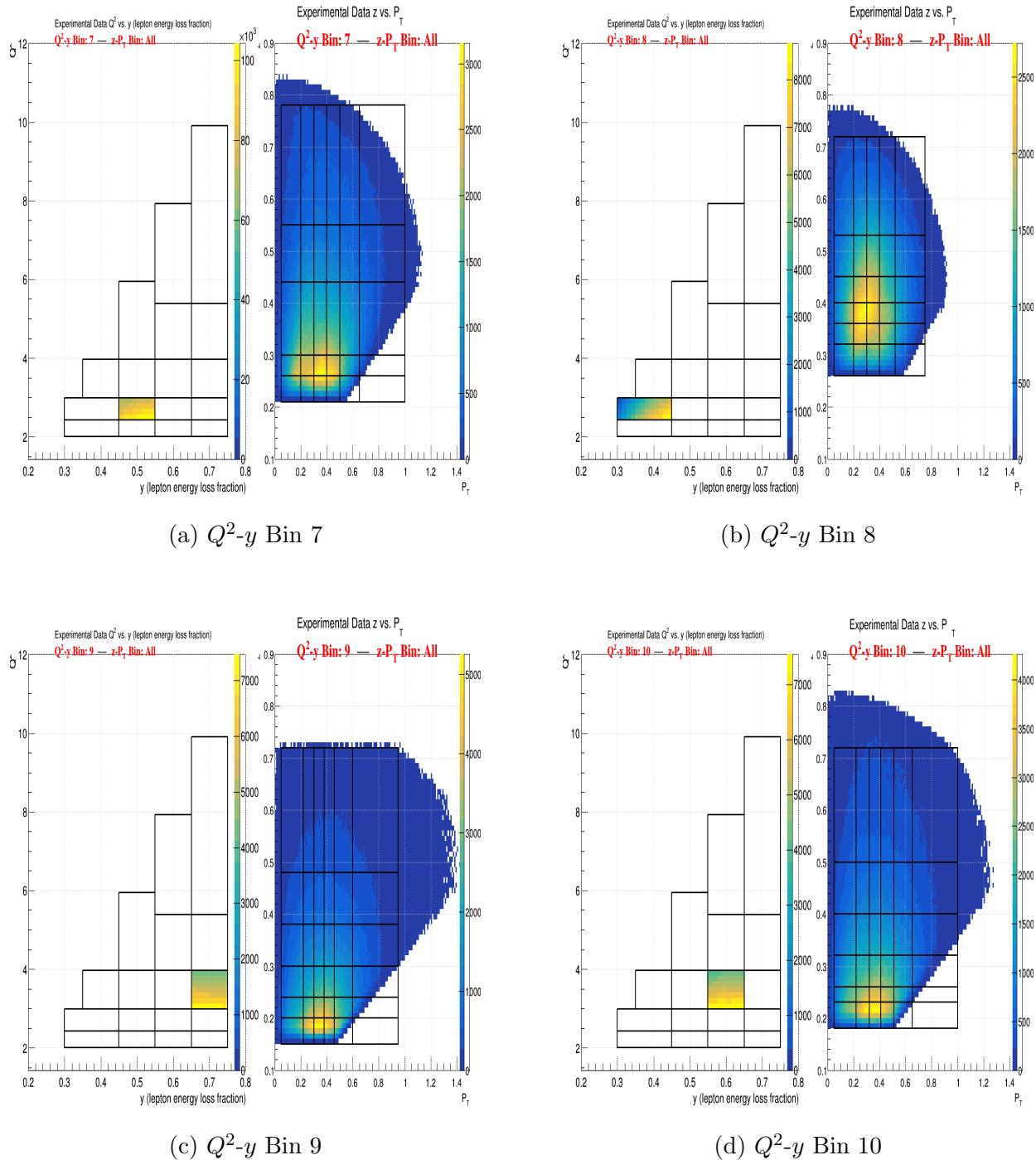


FIG. 17: 2D Kinematic Binning Schemes for Q^2 -y Bin 3, 4, 5, and 6

FIG. 18: 2D Kinematic Binning Schemes for Q^2 -y Bin 7, 8, 9, and 10

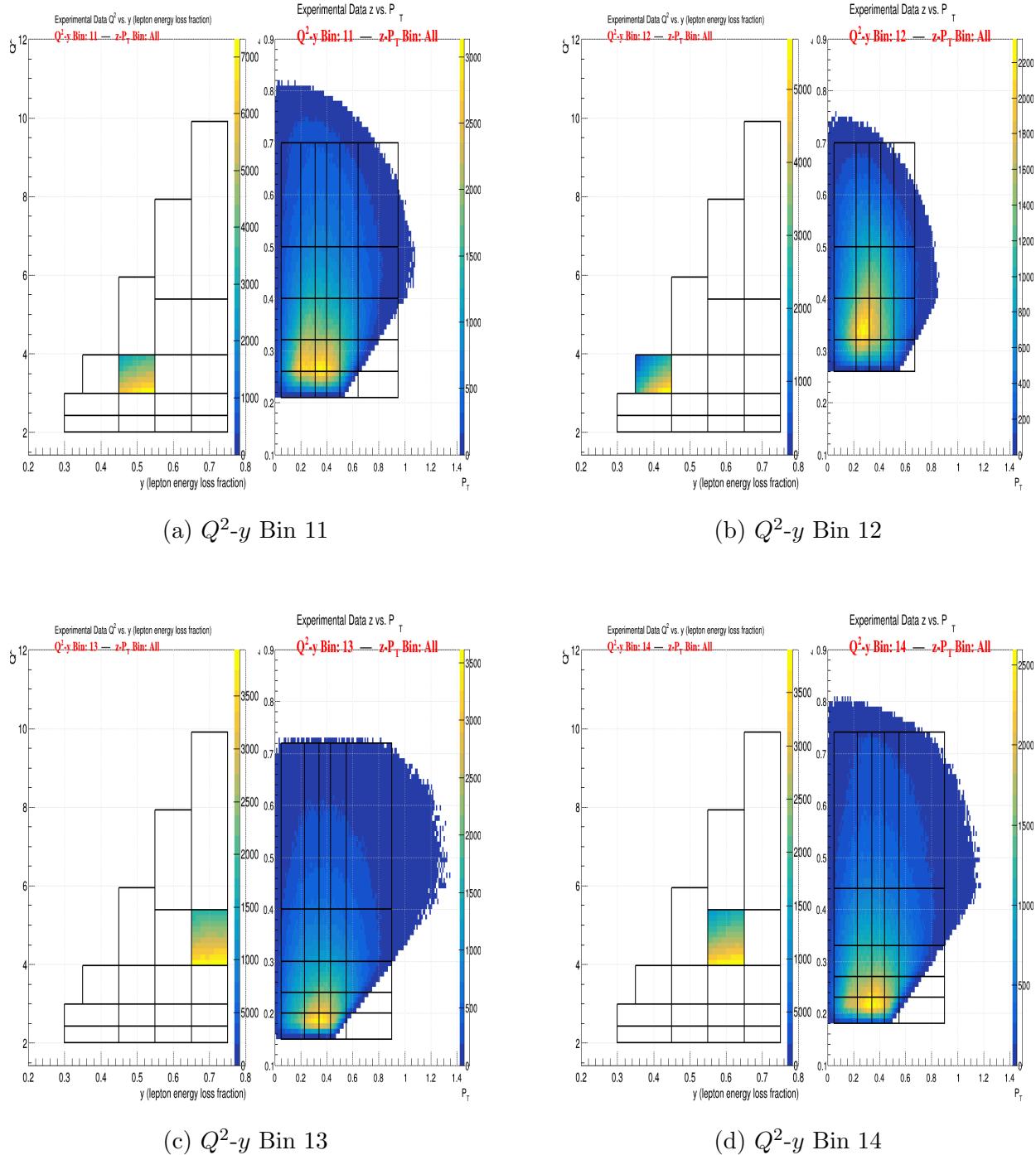
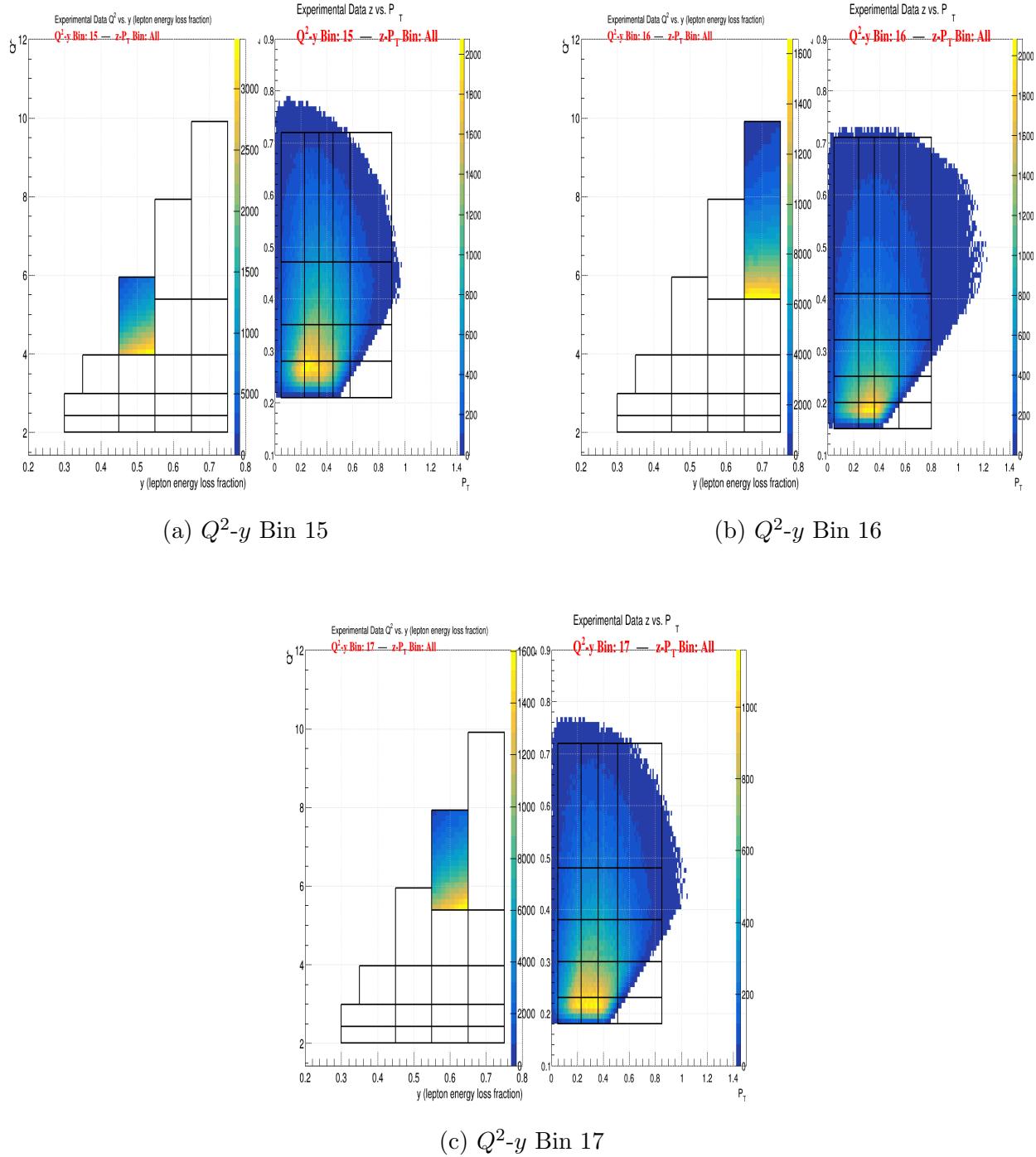


FIG. 19: 2D Kinematic Binning Schemes for Q^2 - y Bin 11, 12, 13, and 14

FIG. 20: 2D Kinematic Binning Schemes for Q^2 - y Bin 15, 16, and 17

IV. MONTE CARLO SIMULATIONS

A. Description

As several necessary corrections for this analysis require the use of simulated data sets to obtain, the use of a Monte Carlo generator was essential. The CLAS12 Fall 2018 RG-A experimental configuration has been described in detail in GEMC [8], a GEANT4-based simulation package that offers the possibility to easily implement detectors in a full GEANT simulation. The position of the detectors in Hall B have been matched to survey data and a realistic map of the magnetic field has been generated to accurately reproduce the experimental setup. LUND generators were used to produce realistic final states that were read by GEMC version 4.3.2 and passed through the detector system of CLAS12. The results of this process were cooked with COATJAVA version 6.5.3, and the reconstructed banks were added to the original generated banks for comparison. The generator used for SIDIS Monte Carlo analysis is clasdis [9] which is based on the PEPSI generator [10, 11], the polarized version of the well-known LEPTO generator. [12]

B. Data vs MC Comparison

Figure 21 shows the normalized comparisons between the Reconstructed (in red) Monte Carlo data and the Real Experimental data (in blue). The distributions shown are of Q^2 , y , z , P_T , x_B , and the momentum and lab θ and ϕ angles of both particles for all events described in **Figures 13 to 14**.

Comparisons of the ϕ_h distributions are also given in each of the subsequent Q^2 - y and z - P_T bins in **Figures 22 to 38** with these figures also including the Monte Carlo Generated distributions in green. The ϕ_h distributions are not expected to be identical as the Monte Carlo simulation was initialized with a flat distribution (as seen in the figures). The modulations that are seen in the Monte Carlo Reconstructed distributions of these plots, therefore, help to illustrate the acceptance effects' contributions to the ϕ_h modulations which will have to be corrected for before making the final measurements of the cross-section's azimuthal moments (See **Analysis Procedure** for more details).

In the z vs. P_T histogram of **Figures 22 to 38**, 4 Missing Mass cut lines have also been drawn to show the effects of this cut (see **Figure 4**) in different regions of the selected Q^2 - y bin. The green line represents the lowest range where the cut could be applied such that no event inside it should be affected by the analysis cut. The black line shows where the cut is applied based on the kinematics of events at the statistical center of the given Q^2 - y bin. The red line shows where all events begin to be removed by the cut (i.e., no events will appear beyond it when the analysis cuts are applied). The outer-most brown line is the same as the black line, but the mass of the neutron is used to plot it instead of the missing mass value used in the analysis cut (i.e., shows the region where the peak position of the exclusive events should be).

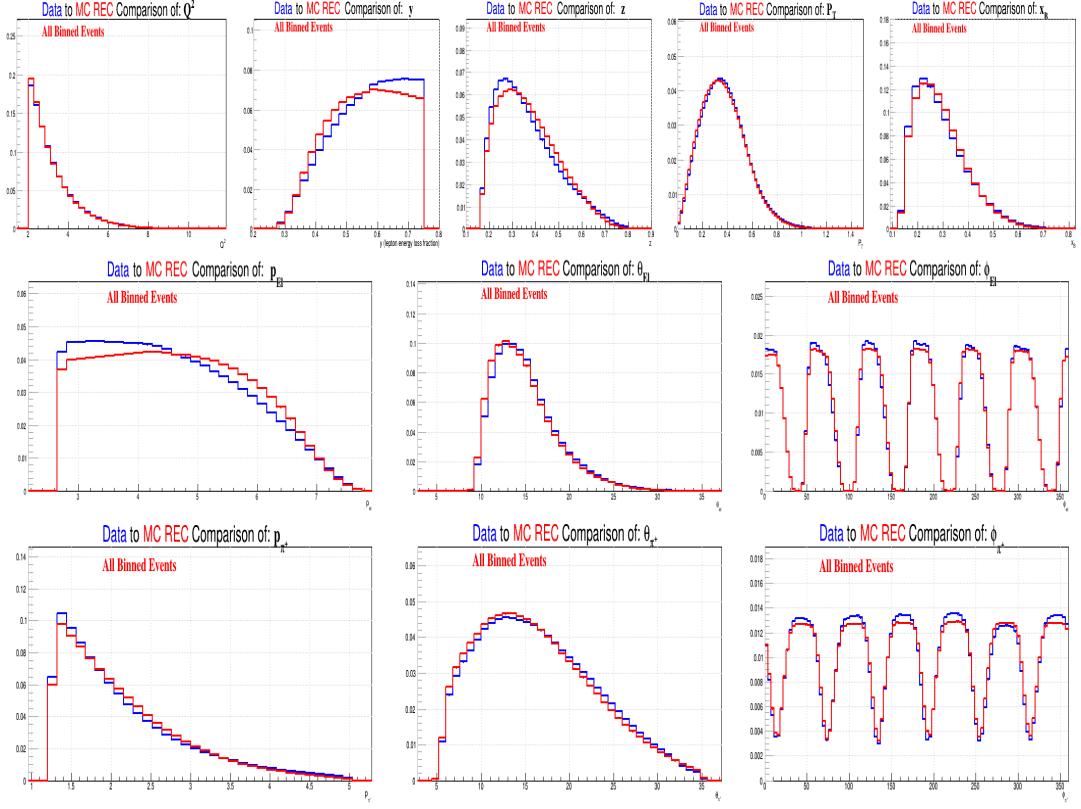


FIG. 21: Comparisons between Data and Monte Carlo for all events. Plots in row 1 are as follows: Q^2 , y , z , P_T , and x_B . Plots in the following two rows show the comparisons of the electron (row 2) and π^+ pion (row 3) kinematics for the momentum (p), polar (θ) and azimuthal (ϕ) lab angle

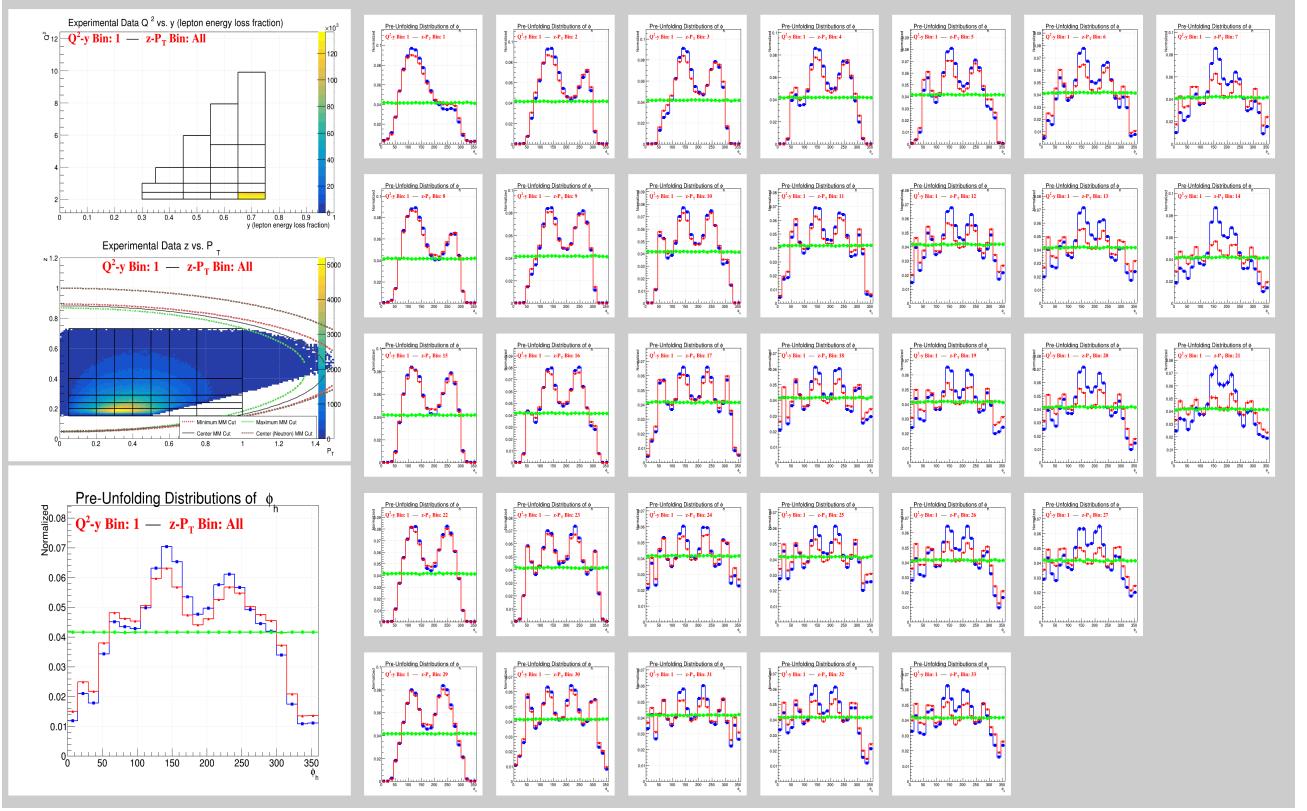


FIG. 22: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 1. The 2D plots show the Q^2 vs. y and z vs. P_T plots in this bin. The 1D histogram in the bottom left corner shows the integrated comparison of all events in the Q^2 - y bin. The rest of the histograms to the right show the distributions in each of the z - P_T bins following the same orientation as is shown in the z vs. P_T histogram

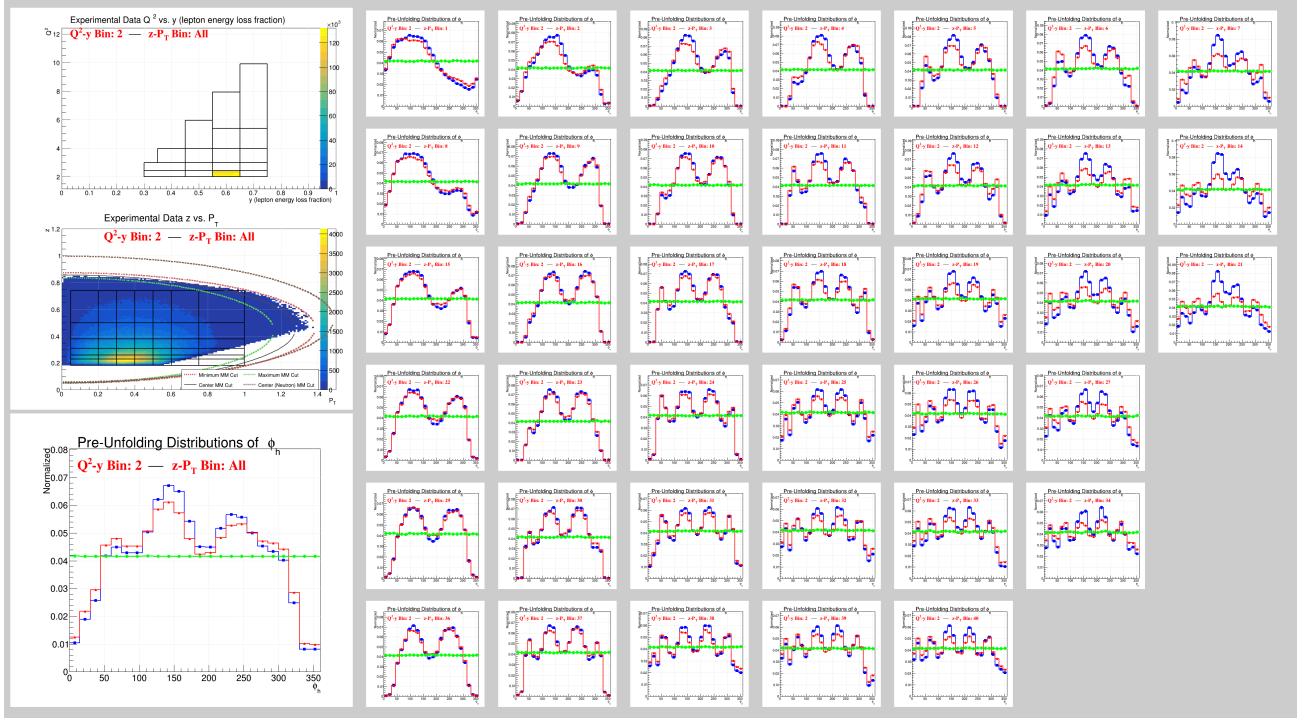


FIG. 23: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 2. See Figure 22 for more details

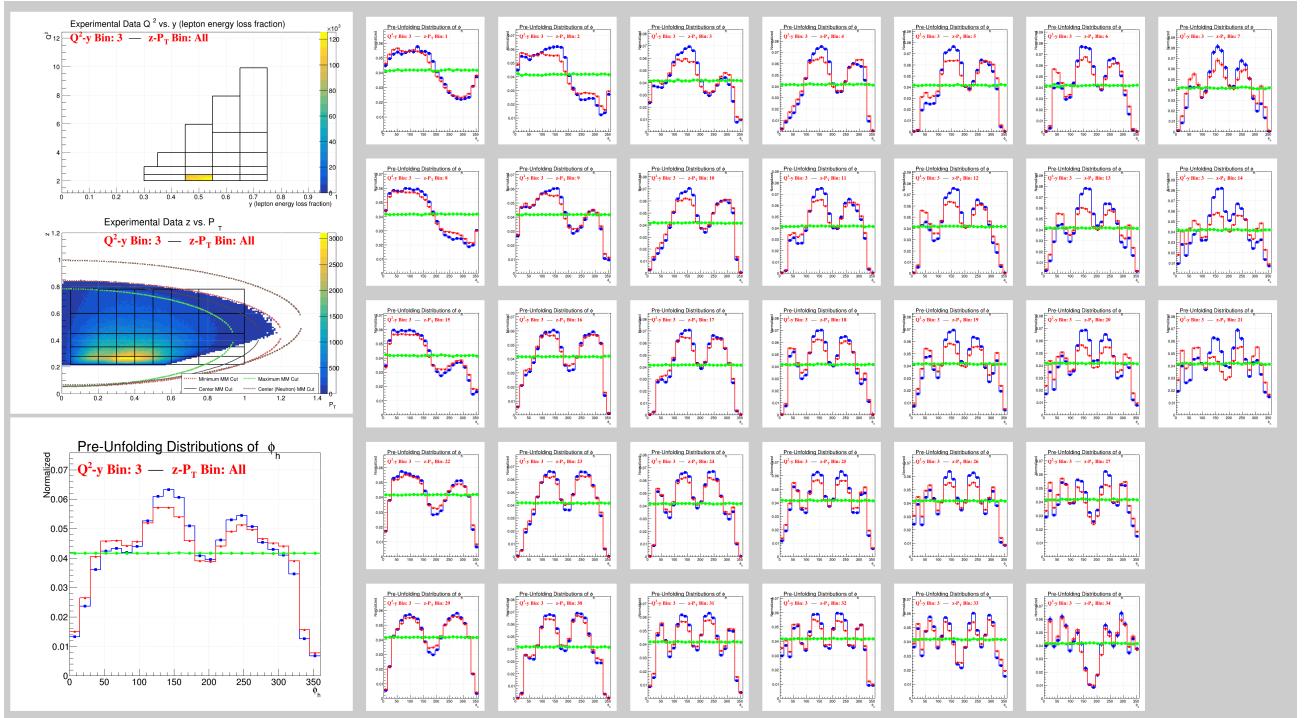


FIG. 24: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 3. See Figure 22 for more details

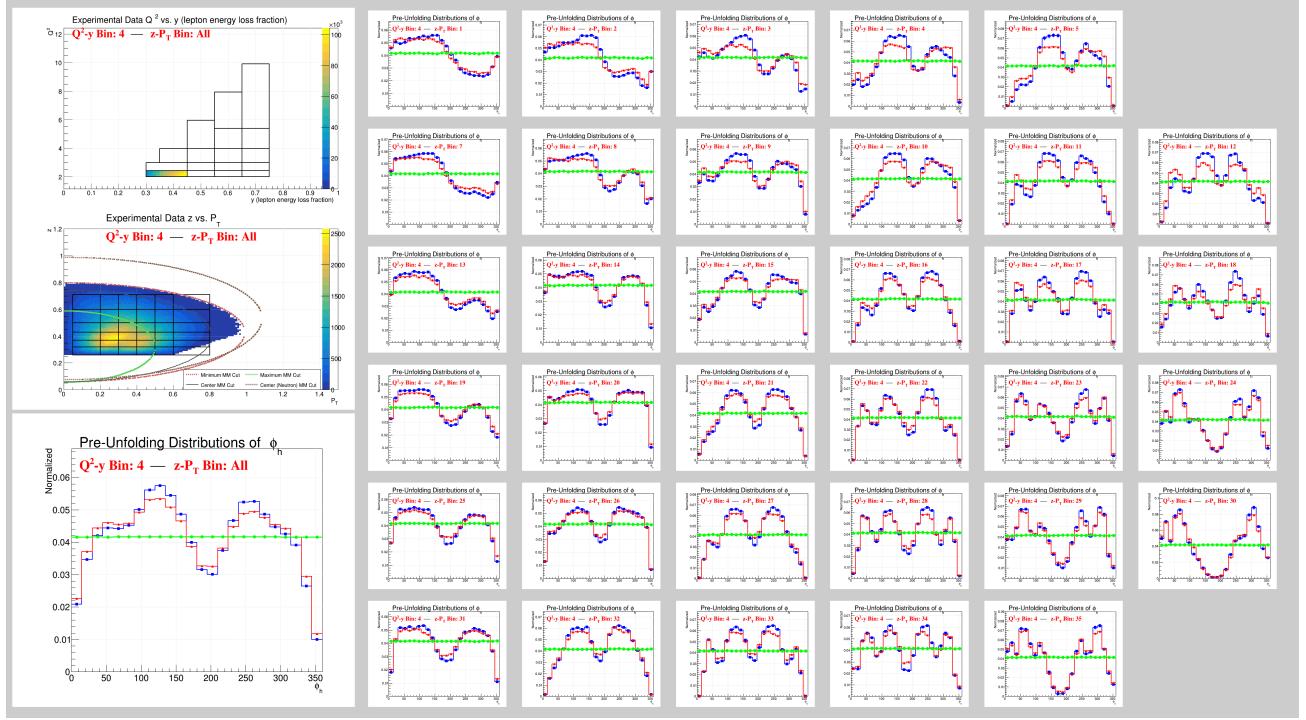


FIG. 25: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 4. See Figure 22 for more details

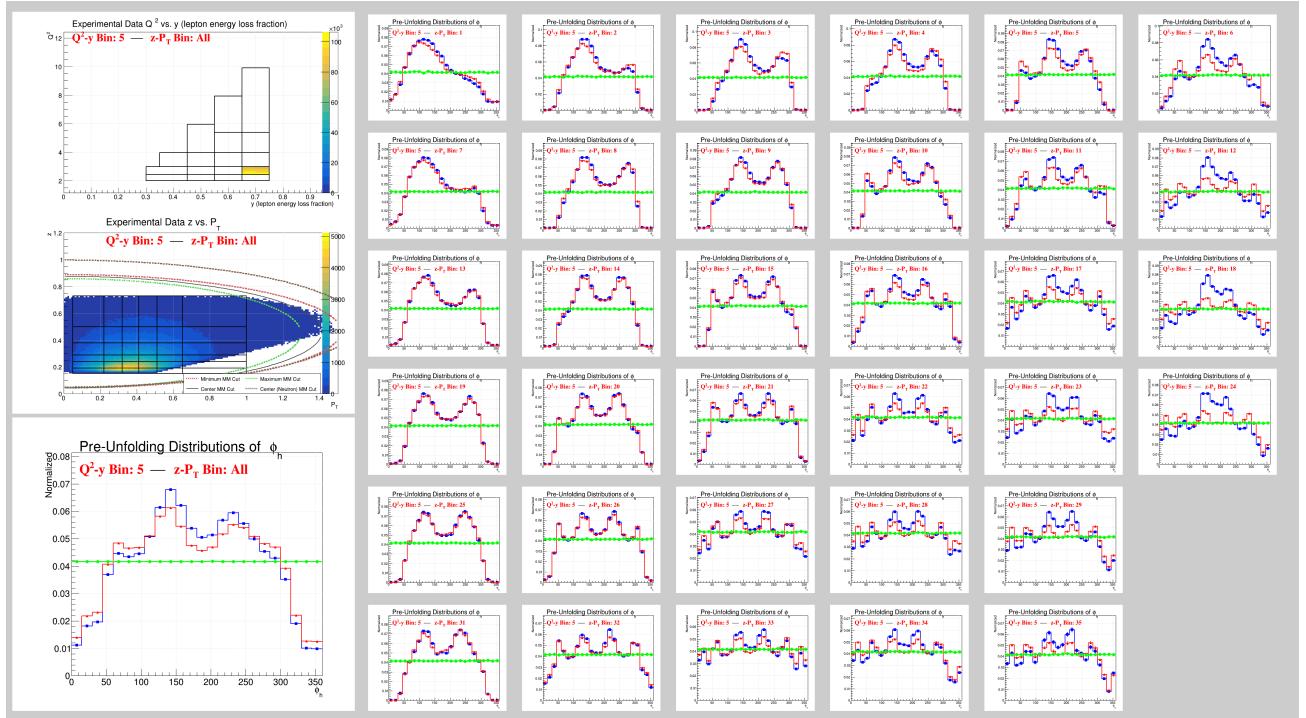


FIG. 26: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 5. See Figure 22 for more details

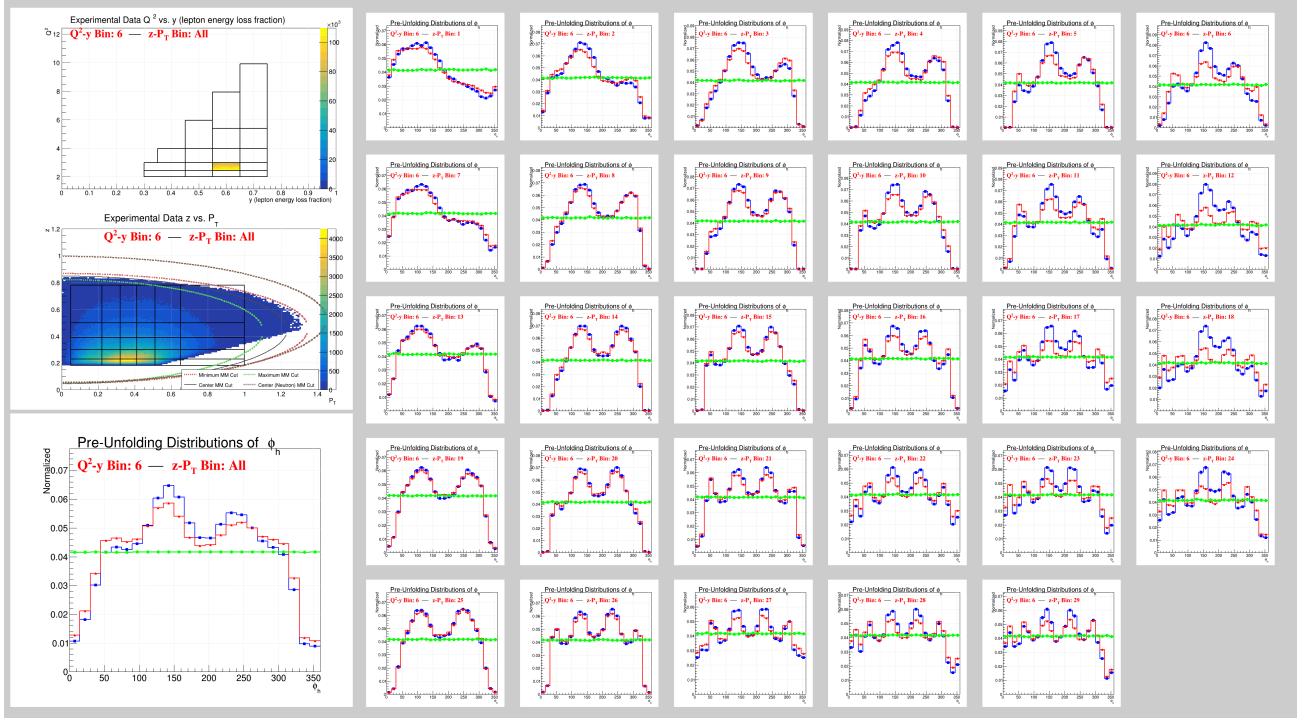


FIG. 27: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 6. See Figure 22 for more details

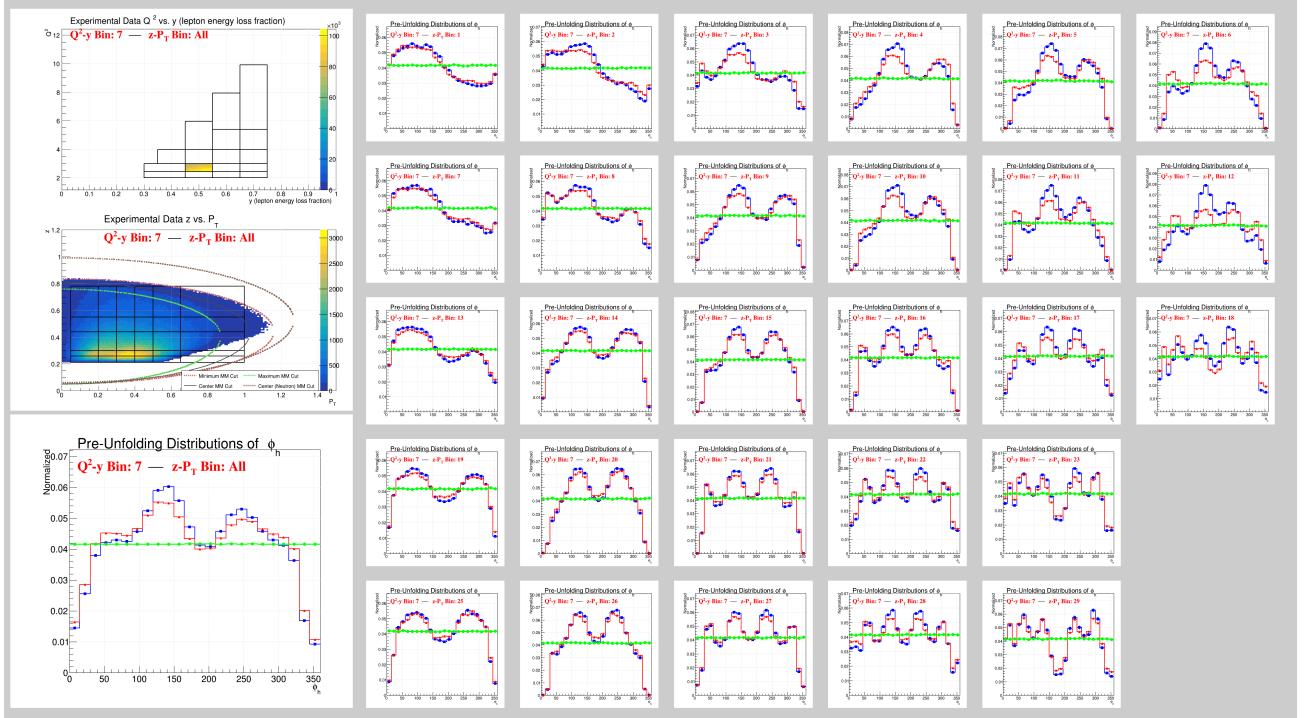


FIG. 28: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 7. See Figure 22 for more details

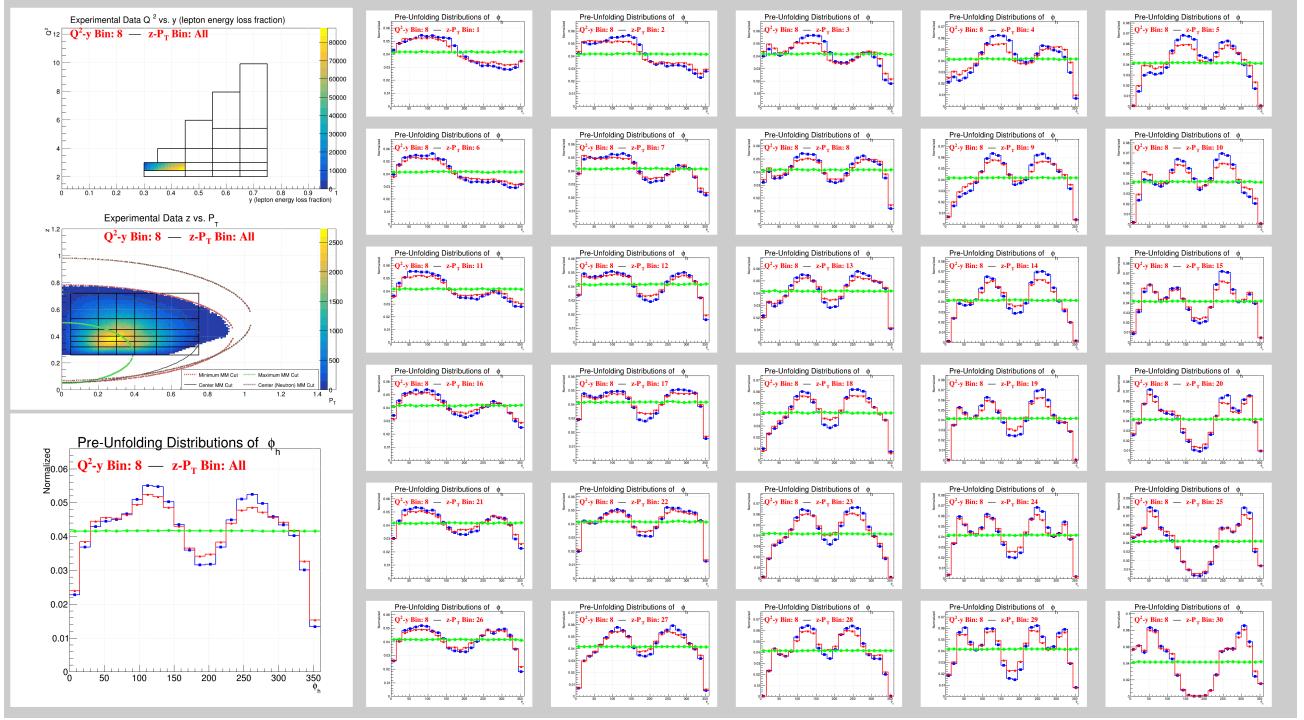


FIG. 29: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2-y Bin 8. See Figure 22 for more details

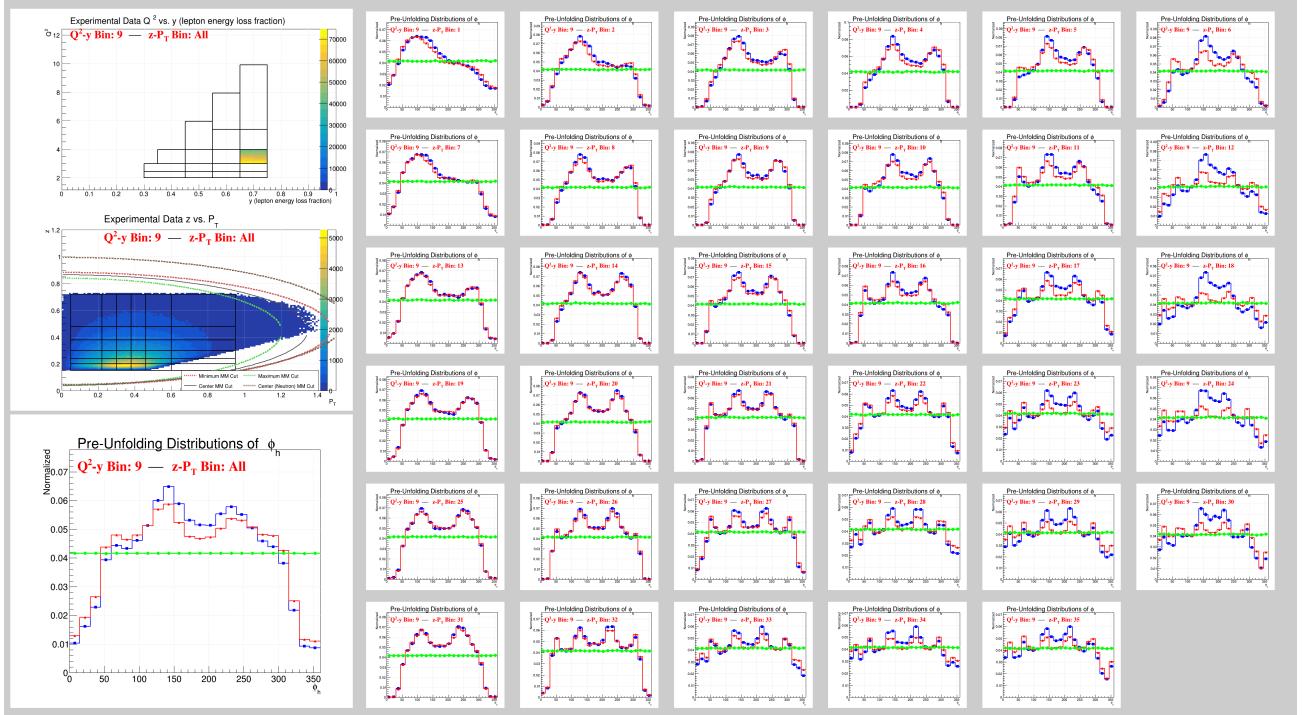


FIG. 30: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2-y Bin 9. See Figure 22 for more details

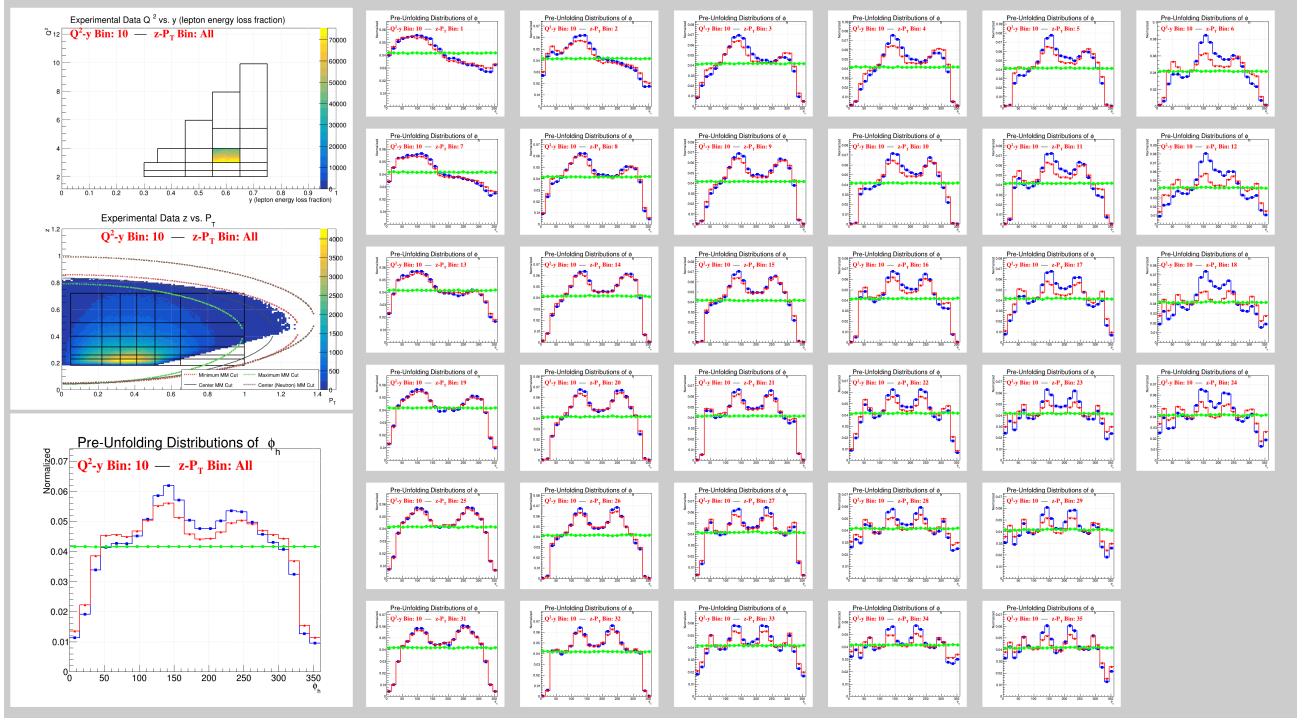


FIG. 31: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 10. See Figure 22 for more details

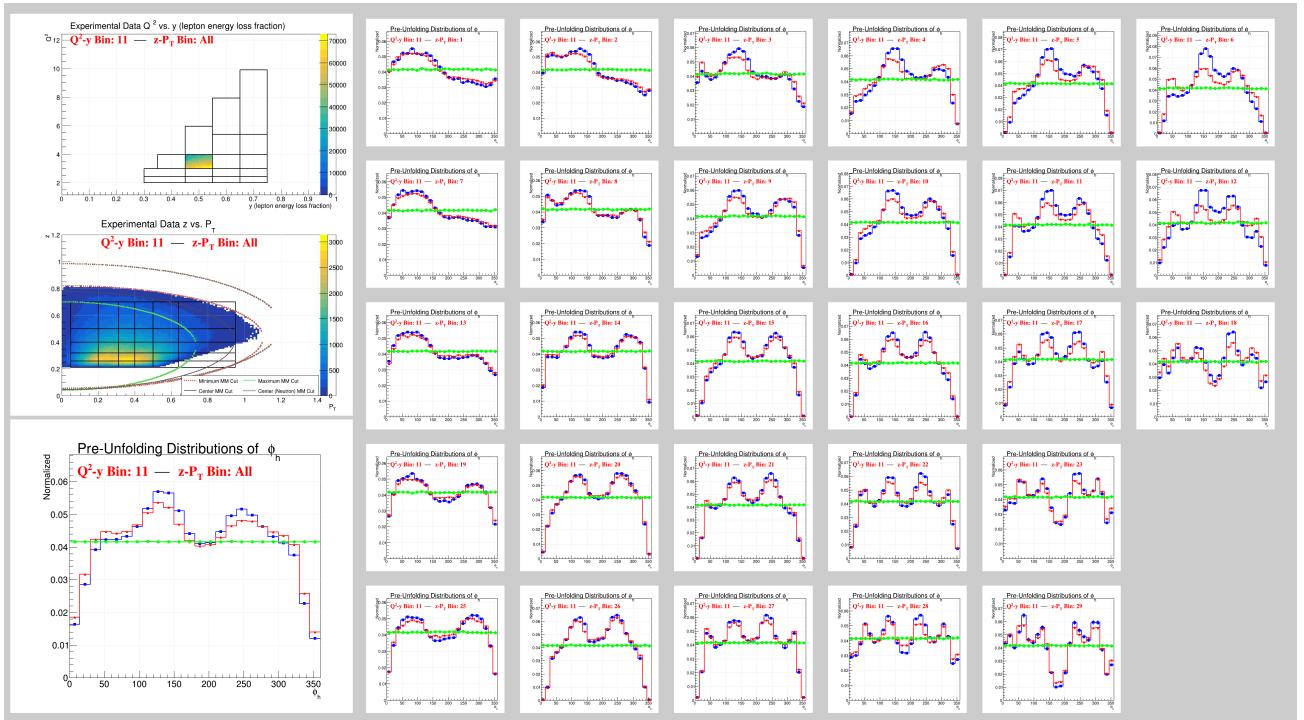


FIG. 32: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 11. See Figure 22 for more details

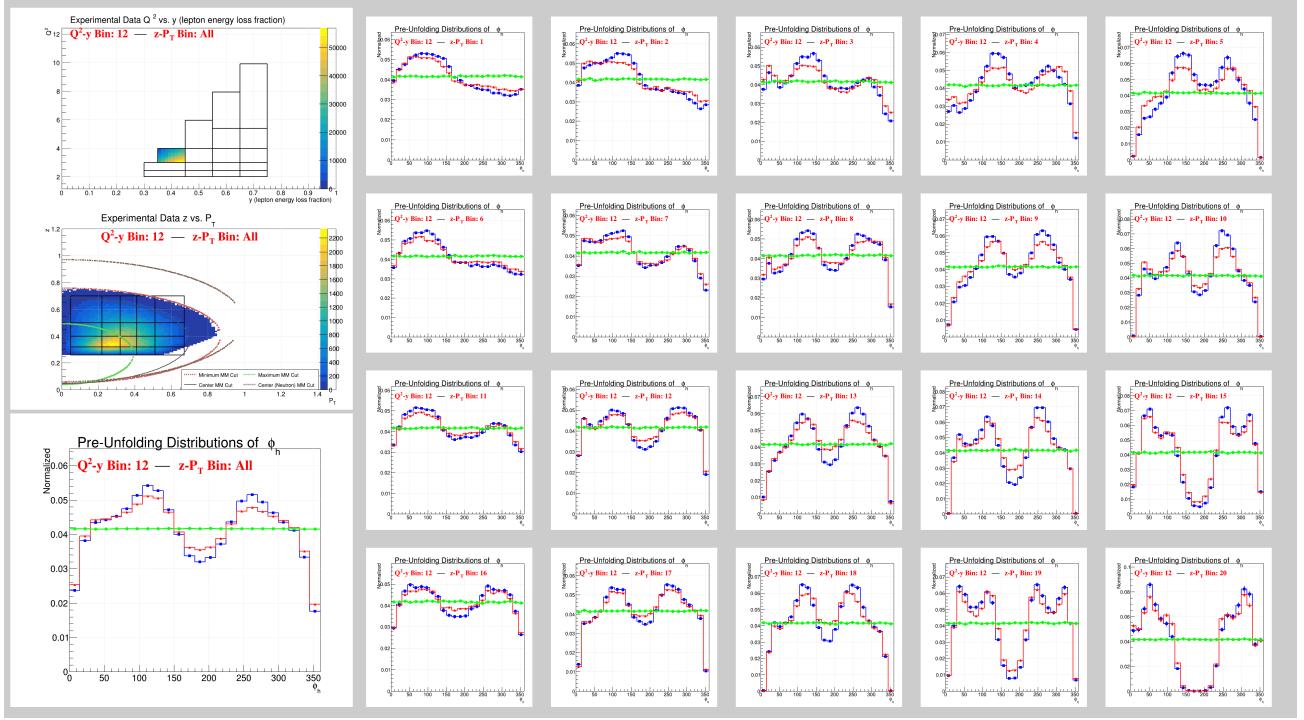


FIG. 33: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 12. See Figure 22 for more details

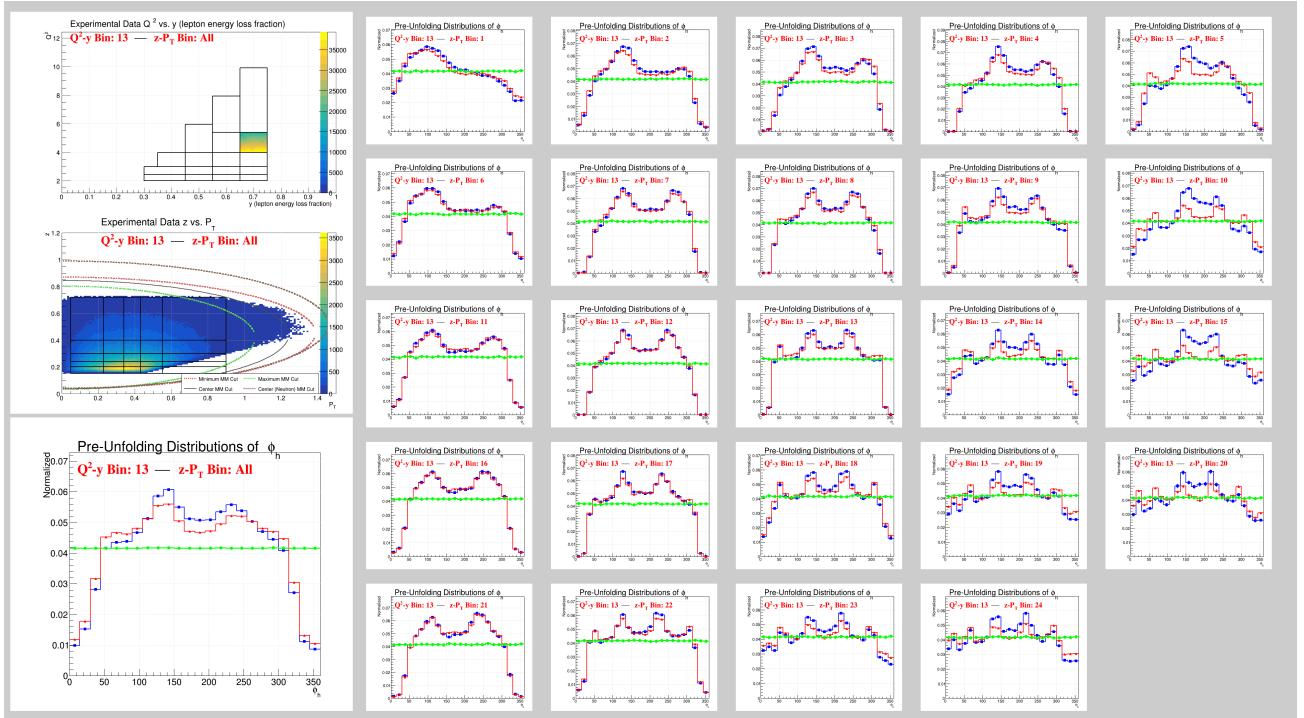


FIG. 34: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 13. See Figure 22 for more details

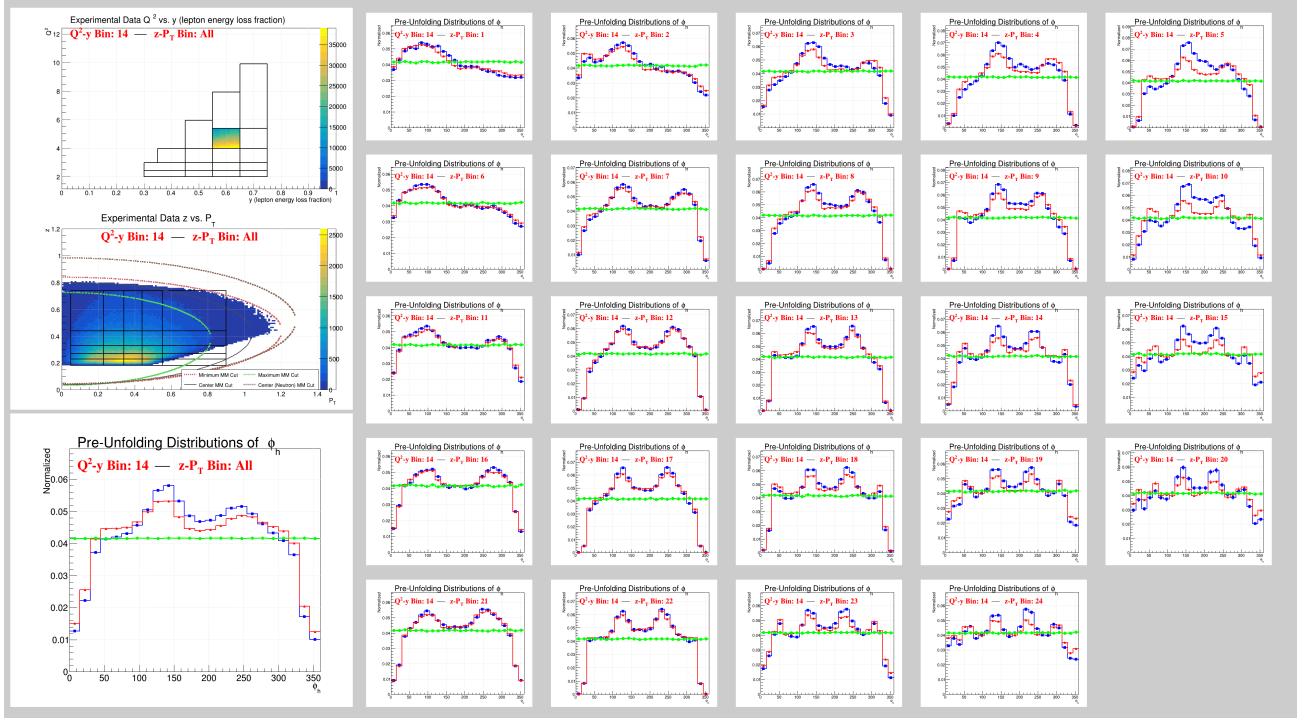


FIG. 35: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 14. See Figure 22 for more details

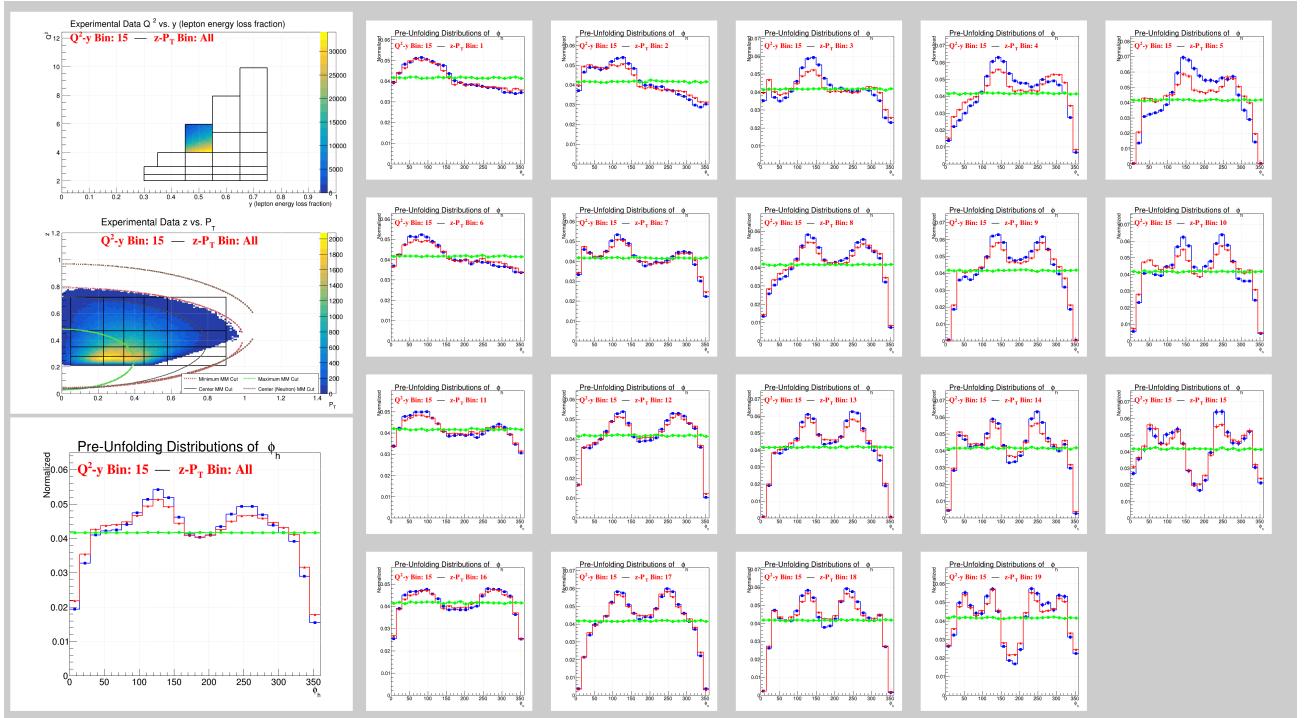


FIG. 36: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 15. See Figure 22 for more details

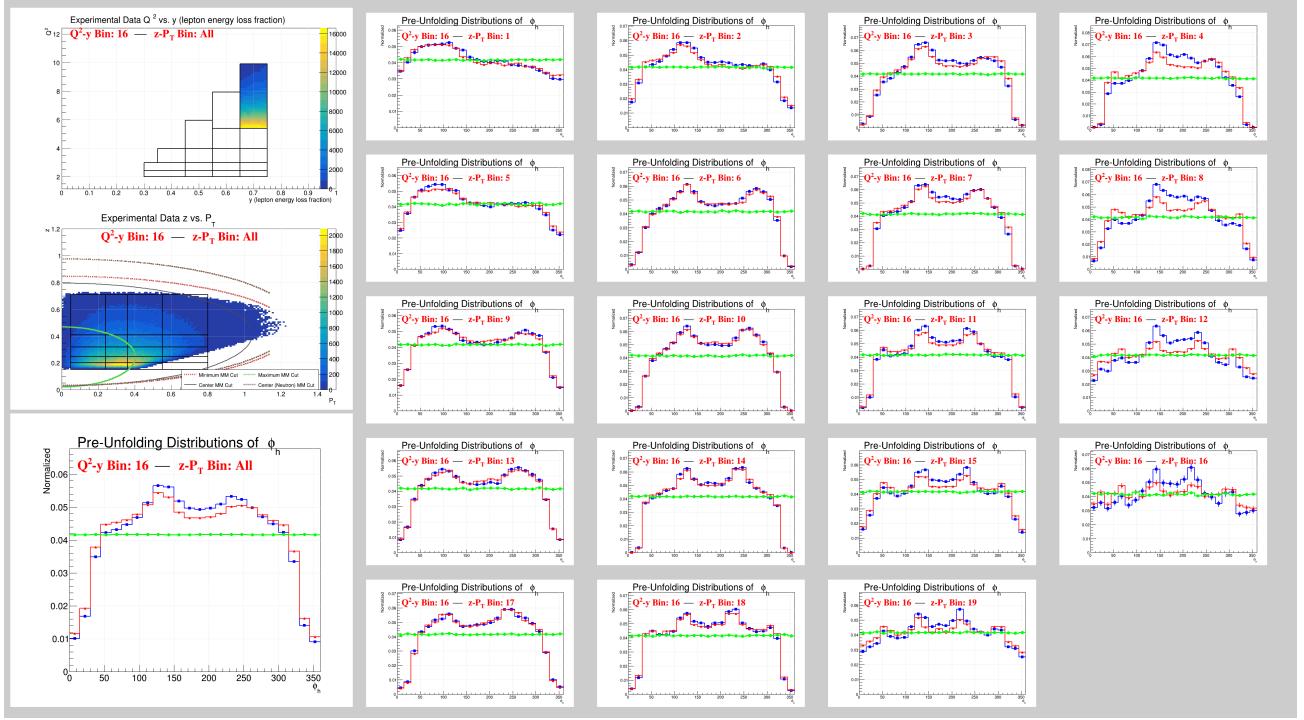


FIG. 37: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 16. See Figure 22 for more details

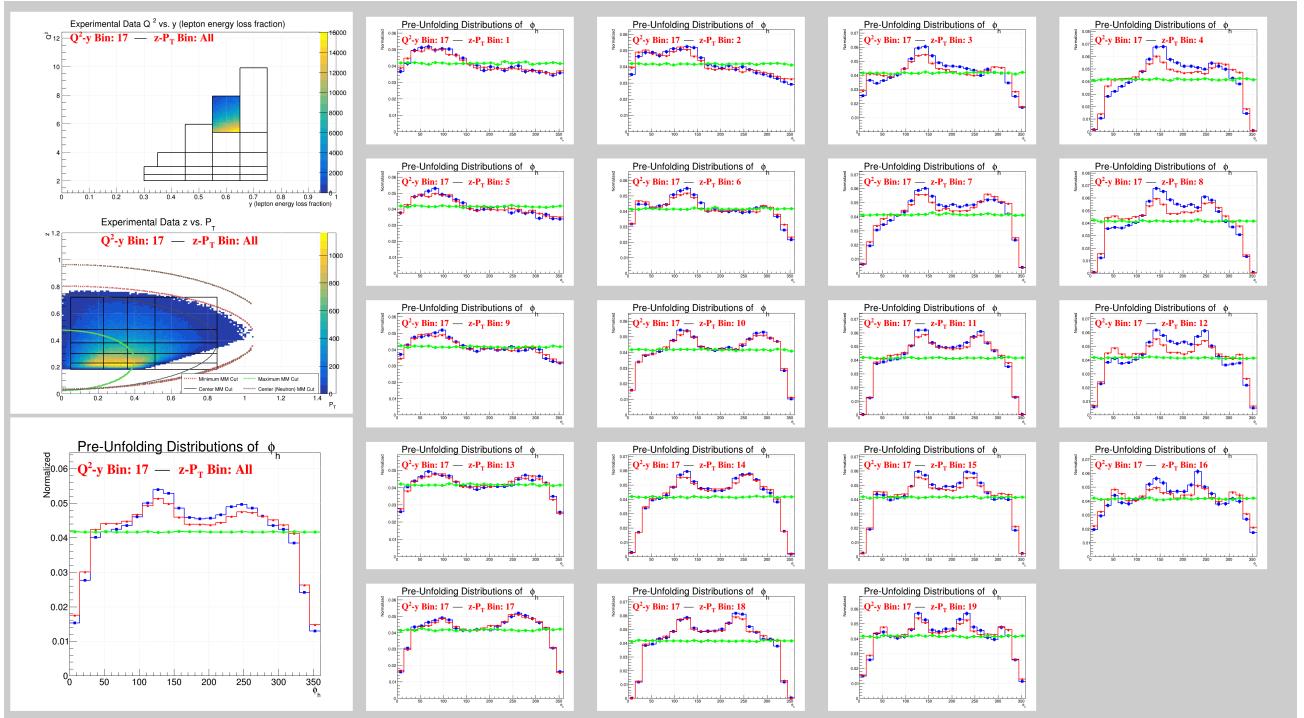


FIG. 38: ϕ_h Comparisons between Data and Monte Carlo for all events in Q^2 - y Bin 17. See Figure 22 for more details

C. MC Event Matching

In order to study the bin migration effects in the Monte Carlo simulation, reconstructed events had to be ‘matched’ to their generated counterparts. The Monte Carlo simulation used does not track information about the generated partner of a reconstructed particle, so this matching must be done event-by-event. This is done by matching the reconstructed particles to whichever generated particle has the most similar polar and azimuthal lab angles. In addition, the following requirements for a match are applied:

- Both the Reconstructed (REC) and Generated (GEN) particle must have the same electric charge.⁷
- $\Delta\theta_{(REC - GEN)} < 6^\circ$
- $\Delta\phi_{(REC - GEN)} < 10^\circ$
- Each particle can only have one match
- If more than one reconstructed particle fulfils the criteria above, the one with the smallest ‘ Δ ’ value defined below is chosen.

$$\Delta = \frac{|\Delta\theta_{(REC-GEN)}|}{|\theta_{REC}|} + \frac{|\Delta\phi_{(REC-GEN)}|}{|\phi_{REC}|} \quad (16)$$

With these criteria and after all other analysis cuts, the majority of reconstructed events (more than 98%) are able to be matched to a generated event. Unmatched events (defined as events in which either the electron or pion partners are not found within the given criteria) are excluded from the acceptance matrices.

D. Momentum Corrections/Smearing

As was done with the experimental data, the reconstructed Monte Carlo data will also be put through a similar momentum correction process using the same type of exclusive events used to correct the experimental data (See [Momentum Corrections](#)). Due to the nature of the Monte Carlo, it is expected that the momentum corrections can be applied in a simplified manner (i.e., it is likely that the Monte Carlo will not require sector or ϕ -dependent corrections and will be satisfied with a simple momentum-dependent correction). [Figure 39](#) shows the effects that a preliminary version of the Momentum Corrections that can have when applied to the Monte Carlo using this simplified procedure⁸. These figures show the Missing Mass vs. Particle Momentum before and after Momentum Corrections applied.

⁷ This requirement is done in place of having a PID requirement to allow for cases in which the event reconstruction mis-identifies a particle.

⁸ The correction factor for the electron in these examples is given as:

$$fe = \frac{(-6.9141 * 10^{-04}) * pp^2 + (5.5852 * 10^{-03}) * pp + (-5.2144 * 10^{-03})}{pp} \quad (17)$$

while the correction factor for the π^+ pion is given as:

$$fpip = \frac{(-1.8752 * 10^{-03}) * pp^2 + (1.0679 * 10^{-02}) * pp + (2.5653 * 10^{-03})}{pp} \quad (18)$$

where pp is the momentum of the particle. Application of the correction factors fe and $fpip$ follow the same methodology as is described in [CLAS12 Momentum Corrections](#).

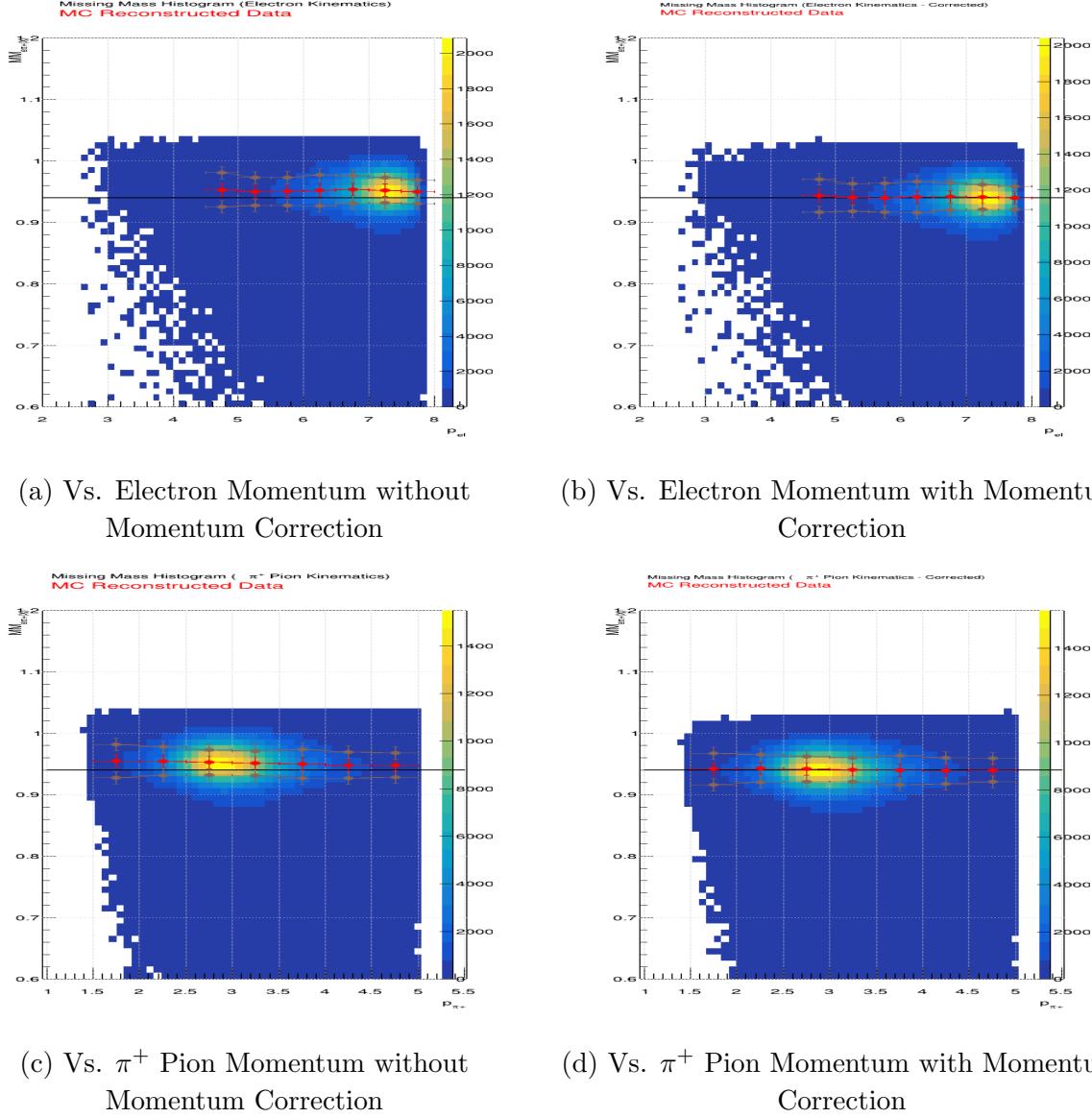


FIG. 39: Missing Mass vs. Momentum Plots for Exclusive Monte Carlo Corrections

Alongside these momentum corrections, the Monte Carlo will also require an additional smearing ‘correction’ whereby the resolution effects observed in the Monte Carlo data will be brought into better agreement with the experimental data.⁹ This momentum smearing can be done with the equation:

$$P_{\text{Smeared}} = P_{\text{Reconstructed}} + SF * (P_{\text{Reconstructed}} - P_{\text{Generated}}) \quad (19)$$

where SF is the “Smear Factor” parameter which controls how significant each simulated event is to be smeared. The exact value of that parameter is to be determined based on a study of similar plots as are used in the development and verification of the momentum corrections, namely by comparing the widths of the Missing Mass distributions from the experimental data and the Monte Carlo data, before and after smearing. Doing this ensures that the Monte Carlo provides the most reliable description of the event reconstruction performed by the detector for the purposes of

⁹ The Momentum Corrections are applied to the Monte Carlo before the additional smearing correction.

correcting the experimental data. Investigations of the most optimal choice for the definition of *SF* are still ongoing at the time of this analysis note's release, with the current value of *SF* being set to a constant value of 0.75 for the example plots shown in the figures below. Figures 40 to 43 show an example of how the effects of this smearing factor are evaluated as a function of the electron/ π^+ pion momentum. Figure 40 shows the Exclusive Missing Mass vs. Momentum plots for (from left to right) the experimental data, the unsmeared reconstructed Monte Carlo¹⁰, and the smeared reconstructed Monte Carlo using the given Smear Factor (*SF*). Figures 41 to 42 shows the sliced distributions of the Missing Mass whose fits provide the points for the peak Missing Mass positions and the widths of the distributions shown in Figure 40. Figure 43 shows the Monte Carlo to Data ratios for the widths of the fits for each of the sliced Missing Mass distributions (As shown in Figures 40 to 42). To ideally match the detector resolutions in the simulation, the ratios shown in Figure 43 should approach 1, providing the condition by which the effectiveness of each choice of *SF* can be judged.

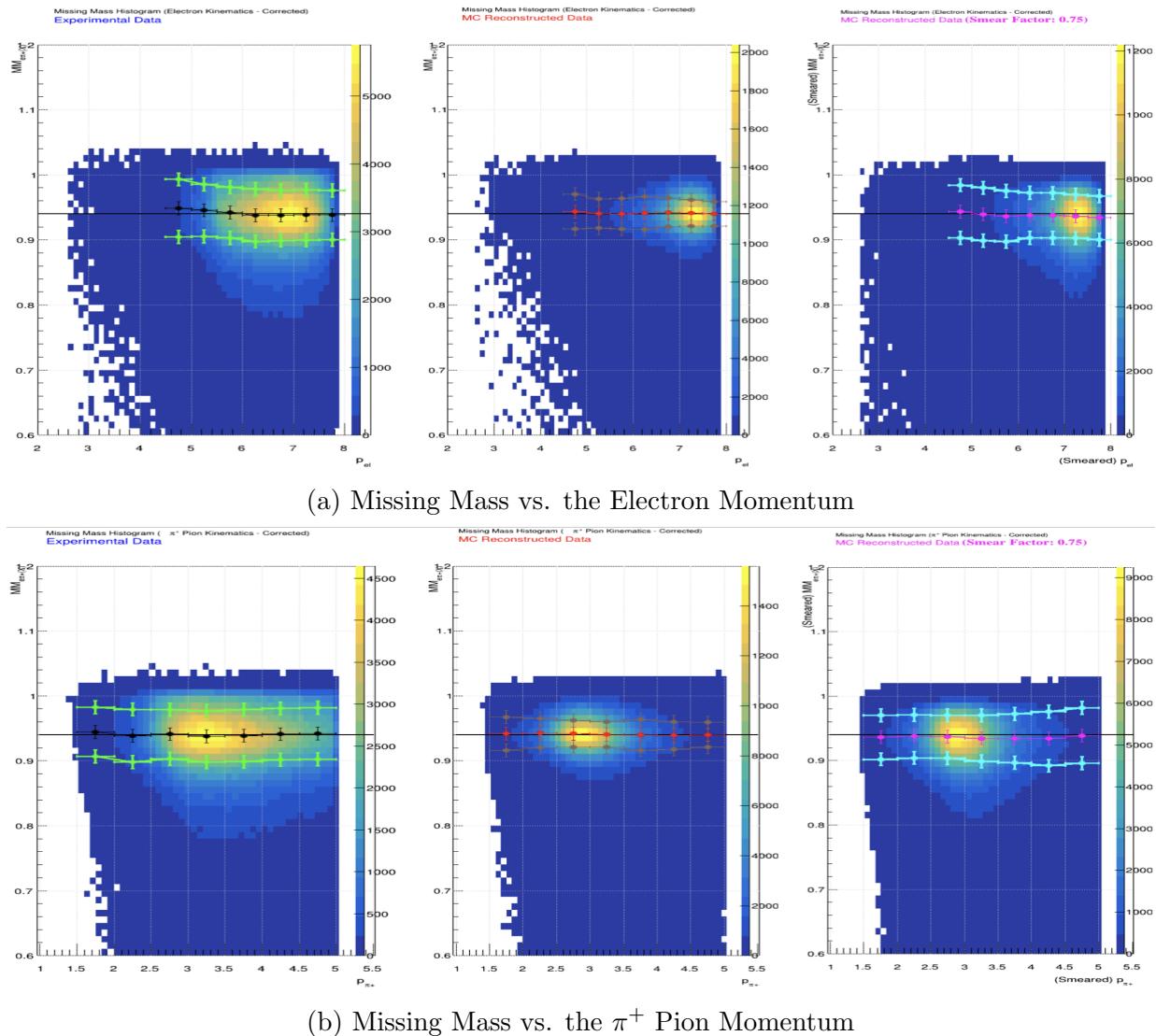


FIG. 40: Missing Mass vs. Momentum Plots with peak positions and plotted widths

¹⁰ ‘Unsmeared’ in this context means that no additional smearing function is applied to the reconstructed Monte Carlo.

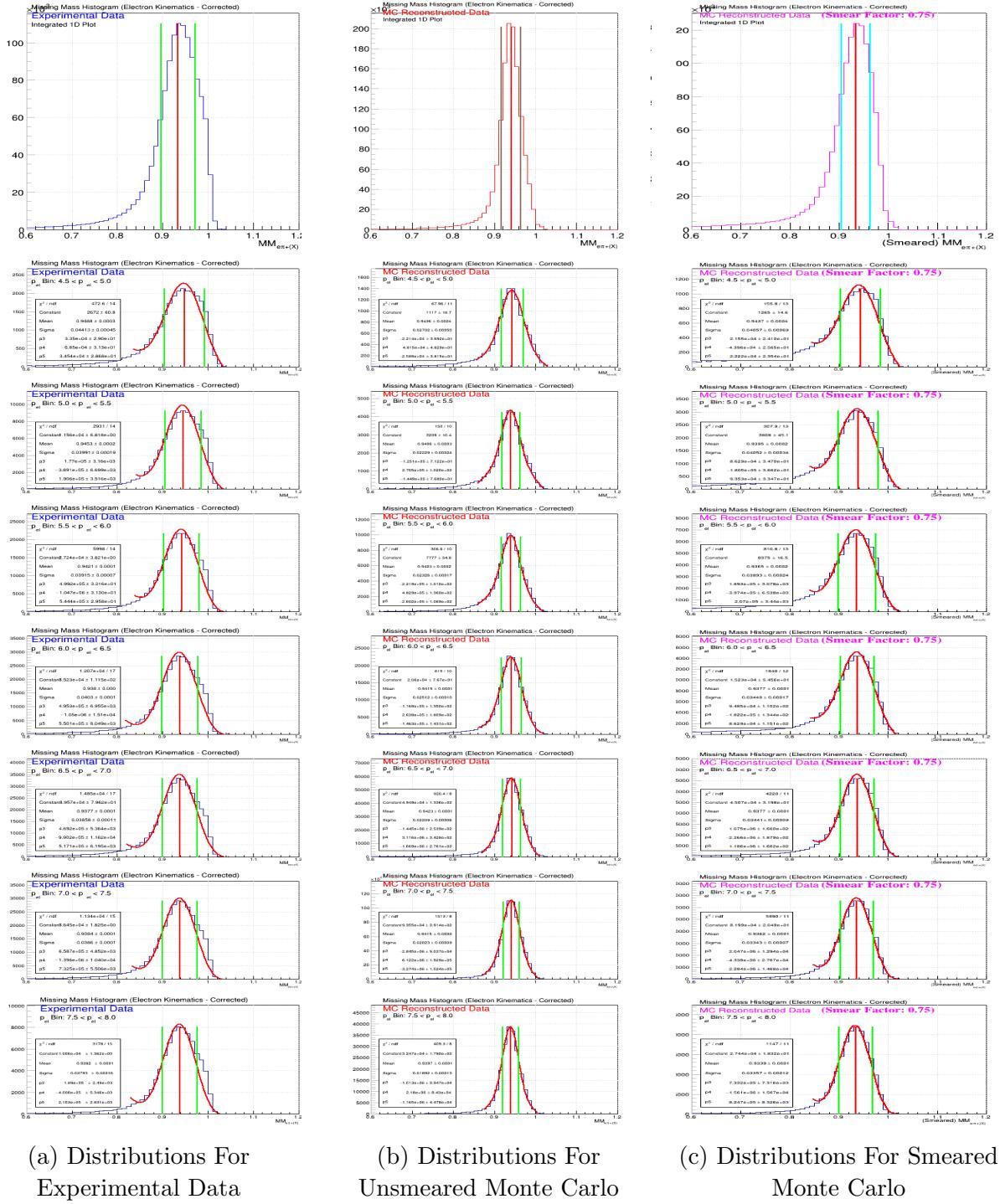


FIG. 41: Sliced Distributions of the Missing Mass in different regions of the Electron Kinematics. The vertical green lines show what is defined as the distributions' widths while the vertical red lines show where the fit estimates the peak positions to be. Plots in the top rows are the integrated distributions, with each of the plots in the subsequent rows being in different regions of Momentum (lowest momentum corresponds to the second row while the last row is for the highest momentum points)

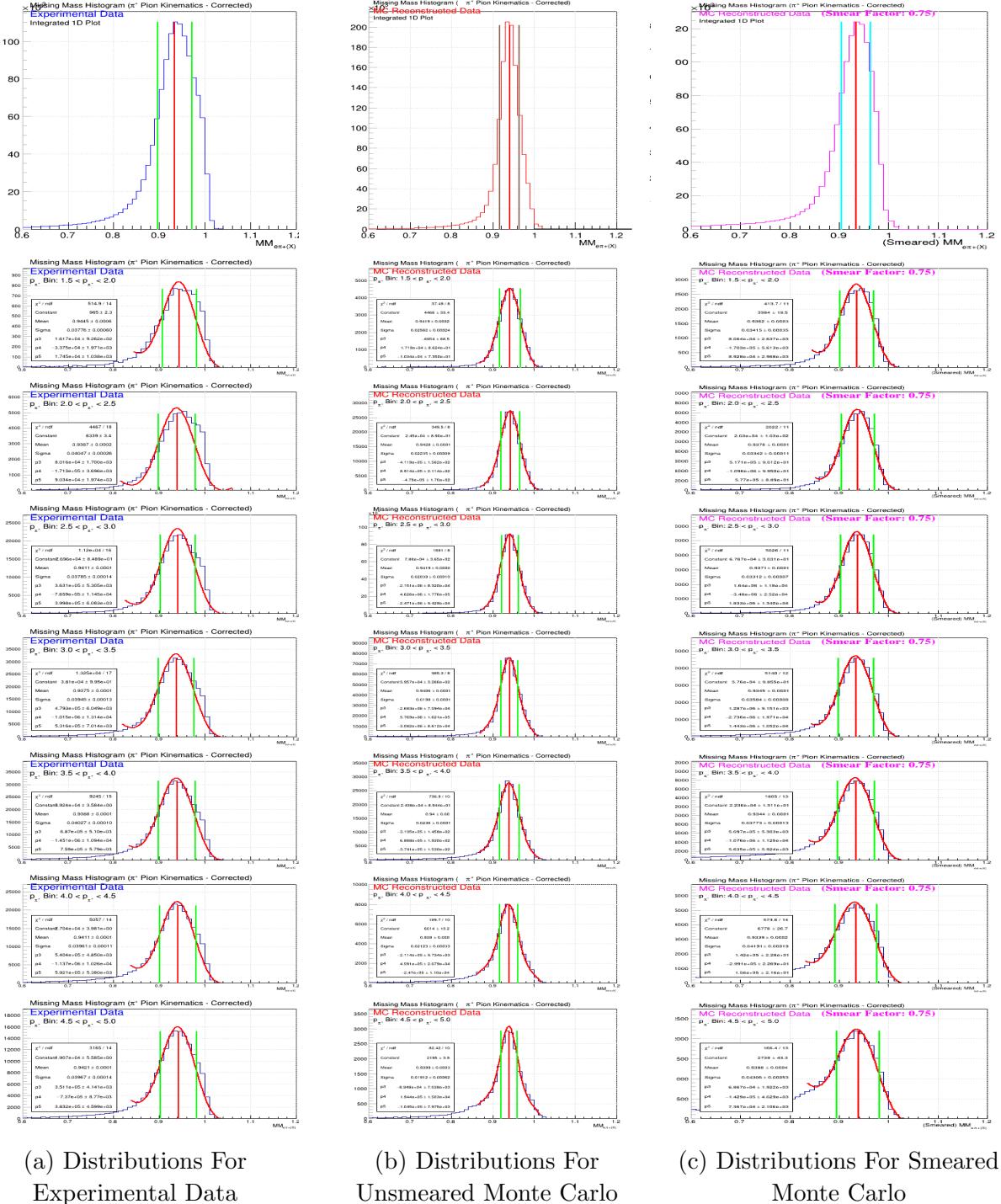


FIG. 42: Sliced Distributions of the Missing Mass in different regions of the π^+ Pion Momentum. See Figure 41 for more details

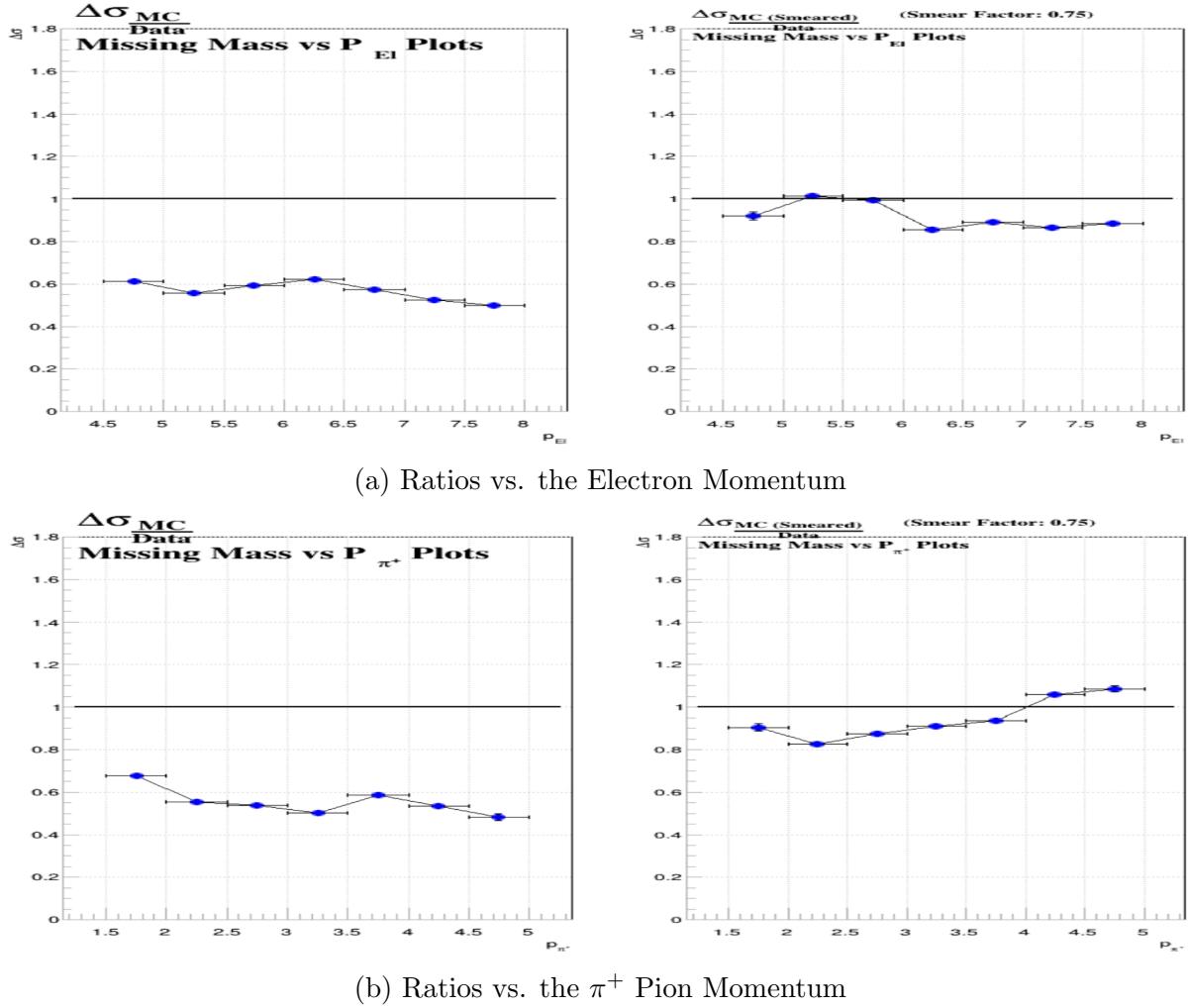


FIG. 43: Ratios of the widths of the Missing Mass distributions as a function of Momentum

V. ANALYSIS PROCEDURE

A. Description

To perform the multidimensional measurement of the $\cos \phi$ and $\cos 2\phi$ azimuthal moments of the unpolarized SIDIS π^+ cross-section, the cross-section moments (A_0 , $A_{UU}^{\cos \phi_h}$, and $A_{UU}^{\cos 2\phi_h}$) will be extracted from the fits of the ϕ_h distribution in different Q^2 , y , z , and P_T bins, with the function:

$$A(1 + B \cos \phi_h + C \cos 2\phi_h) \quad (20)$$

where the moments can then be expressly given as $A_0 = A$, $A_{UU}^{\cos \phi_h} = B$, and $A_{UU}^{\cos 2\phi_h} = C$.

To take the experimental measurements that are necessary, the SIDIS cross-section can be extracted from the following formula:

$$\frac{d^5\sigma}{dQ^2 dy dP_T dz d\phi_h} = \frac{1}{\Gamma_\nu} \frac{1}{\Delta Q^2 \Delta y \Delta P_T \Delta z \Delta \phi_h} \frac{N}{R \cdot BC \cdot \eta \cdot N_0} \frac{1}{(N_A \cdot \rho \cdot t / A_w)} \quad (21)$$

where:

Γ_ν = Virtual Photon Flux

R = Radiative Correction Factor

N = Bin Yields

BC = a factor which evolves bin-averaged differential cross-section to a specific kinematic point within the bin

η = Acceptance Correction

N_0 = Life-time corrected incident electron flux

N_A = Avogadro's Number

ρ = target density

t = target length

A_w = Atomic weight of the target

Also, the “ Δ ” terms on the right side of the equation refer to the bin volume of the particular bin in which the cross-section is being measured, while the last term in the equation refers to the target number density. Specifically, the whole term is one over the target number density.

By plotting the ϕ_h distributions in each of the combined Q^2 - y - z - P_T bins and applying the correction factors listed in Equation (21), it will be possible to fit the distribution of ϕ_h with Equation (20) and extract the multidimensional measurements of the azimuthal moments from those fits. As these corrections are crucially important to the accurate extraction of the A_0 , $A_{UU}^{\cos \phi_h}$, and $A_{UU}^{\cos 2\phi_h}$ moments from the differential cross-section, details of the correction factors within Equation (21) will be discussed in the following sections of this analysis note.¹¹

¹¹ Note that these corrections are in addition to the momentum correction/smearing corrections discussed in prior sections of this analysis note, which ought to be done for the experimental and simulated data. Further discussion of these corrections will not be included in this section.

B. Acceptance Corrections

The first of the correction factors from [Equation \(21\)](#) that is being considered is the Acceptance corrections factor η . In its simplest form, η can be calculated by using a bin-by-bin comparison between the Monte Carlo Reconstructed and Generated distributions whereby:

$$\eta = \frac{\text{Number of Events Reconstructed Per Bin}}{\text{Number of Events Generated Per Bin}} \quad (22)$$

This method results in a single correction factor per kinematic bin. However, a more appropriate approach for this analysis will be to use an Acceptance matrix to perform the acceptance corrections. An Acceptance matrix represents a more sophisticated way of describing geometric acceptances, detector efficiencies, and bin migration effects by providing more information about what happens to a generated event as it undergoes the reconstruction process to be detected by the detector. Therefore, to correct for acceptance while considering the effects of bin migration, an Acceptance matrix $A_{(i, j)}$ should be used where:

$$A_{(i, j)} = \frac{\text{Number of events generated in bin } j \text{ but reconstructed in bin } i}{\text{Total number of events generated in the } j\text{-th bin}} \quad (23)$$

This definition can also be used to define what is referred to as a Response matrix, which is defined in the same way as the matrix $A_{(i, j)}$, but without being divided by the total number of generated events in the j -th bin. The Response matrix becomes relevant when discussing the available tools and algorithms that perform the corrections with these matrices.

With the Acceptance matrix, it is possible to go from a set of experimental data (represented by a vector Y_i where this vector is equal to the number of events experimentally measured in the i -th bin) to the acceptance unfolded data (represented by a vector X_j where this vector is equal to the number of acceptance-corrected events in the j -th bin). This can be done by unfolding the relation:¹²

$$Y_i = A_{(i, j)} X_j \iff X_j = A_{(i, j)}^{-1} Y_i \quad (24)$$

where the right-hand side of the equation gives the acceptance corrected form of a given data set. To perform this unfolding correction, unfolding methods such as the Singular Value Decomposition (SVD) or Bayesian method must be used. Both of these methods are built into the ROOT library called “RooUnfold” making it a much simpler task to compare and test these methods to see which would work best for this analysis.

To perform the unfolding corrections with the matrices described in [Equation \(23\)](#), the SVD method uses an algorithm that calculates the inverted response matrix of the data, while the Bayesian method takes an iterative approach to predict the unfolded response matrix using an algorithm based on elements of Bayesian statistics. A full description of the SVD method can be seen in “SVD Approach to Data Unfolding” by Andreas Höcker and Vakhtang Kartvelishvili [13] while the Bayesian method was introduced by Giulio D’Agostini in “A Multidimensional unfolding method based on Bayes’ theorem” [14].

¹² If $A_{(i, j)}$ does not have any non-diagonal terms (i.e., if for $i \neq j$, $A_{(i, j)} = 0$) then these terms can be applied in the same way as the correction factor η was defined in [Equation \(22\)](#).

Implementations of both the SVD method (for single variable matrices with the same number of bins for the measured and the unfolded spectrum) and the Bayesian method are already available in the ROOT library called ‘RooUnfold’. In the case of the SVD method, the ‘RooUnfold’ library uses another existing library known as ‘TSVDUnfold’. These libraries take as inputs a set of 1D histograms for the generated and reconstructed simulations and a Response matrix, which are then all used to unfold the experimental data. With the implementation of the ‘RooUnfold’ library, these inputs are stored in a custom Response matrix object (i.e., `RooUnfoldResponse`) which allows for the easy application of several different unfolding methods such as those used in this analysis. The application of the ‘TSVDUnfold’ library’s basic example was used to help set the regularization parameter of `kreg = 13` (used to regularize the final unfolded distributions) for this analysis.

The regularization parameter used in the unfolding procedure presented in this analysis note was chosen as a preliminary test of the SVD method. This choice of the regularization parameter was shown to work reasonably well for the current systems to which it was applied. However, future developments in this analysis, especially regarding plans for the multidimensional unfolding, would likely generate the need to revisit this parameter’s value at a later time if the SVD method were to continue to be used (for reasons also discussed in this section, this method may not be used in the future for these multidimensional corrections). For the purposes of demonstrating the unfolded procedure in this Note, this parameter is a reasonable choice. The example used to model the application of the SVD method is available within the ROOT Reference Guide for `TSVDUnfoldExample.C`.¹³ The number of iterations used for the Bayesian method was increased from the default value of 4 used by ‘RooUnfold’ to 10 iterations to help ensure the convergence of the χ^2 value used to evaluate how well each iteration performed. Figures 44 to 49 shows the results of these 1D corrections in some of the Q^2-y bins described in this analysis note.

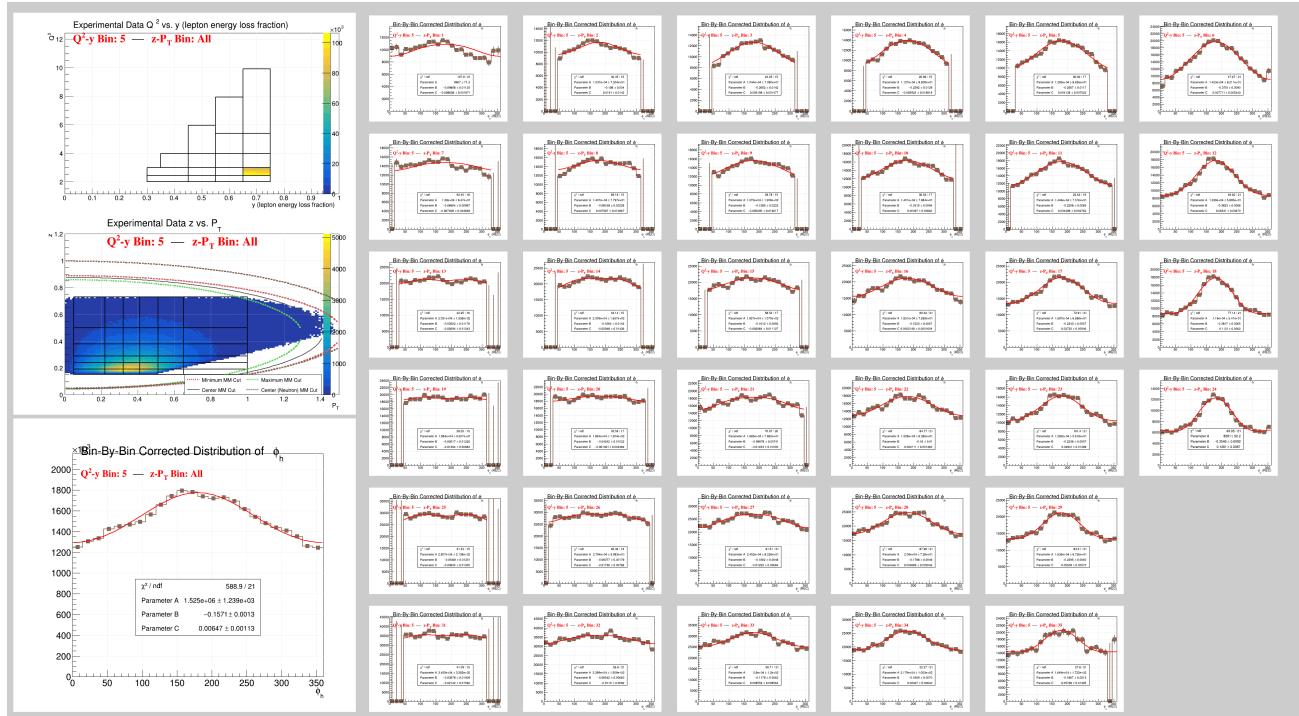


FIG. 44: 1D ϕ_h Bin-by-Bin Acceptance Corrections for all events in Q^2-y Bin 5. See Figure 22 for more (basic) details regarding the layout of this type of image

¹³ This example is available here: https://root.cern/doc/master/TSVDUnfoldExample_8C.html

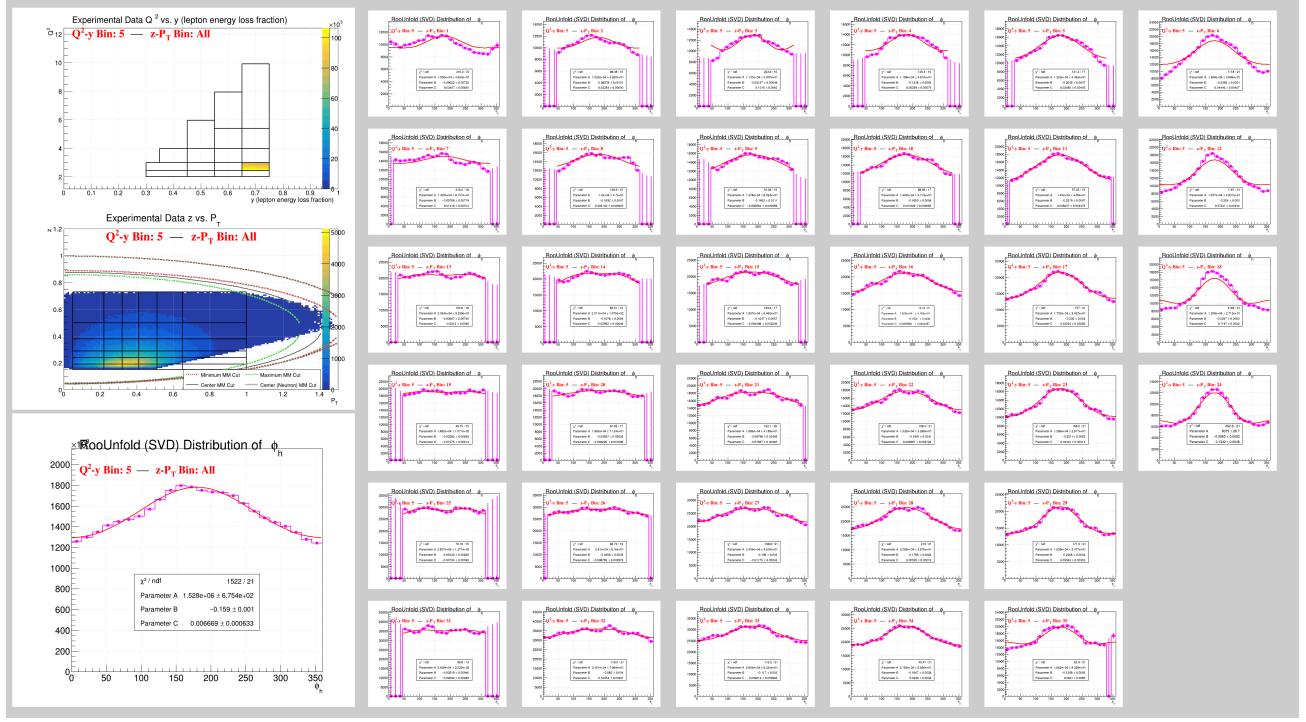


FIG. 45: 1D ϕ_h SVD Unfolding Corrections for all events in Q^2 - y Bin 5. See Figure 22 for more (basic) details regarding the layout of this type of image

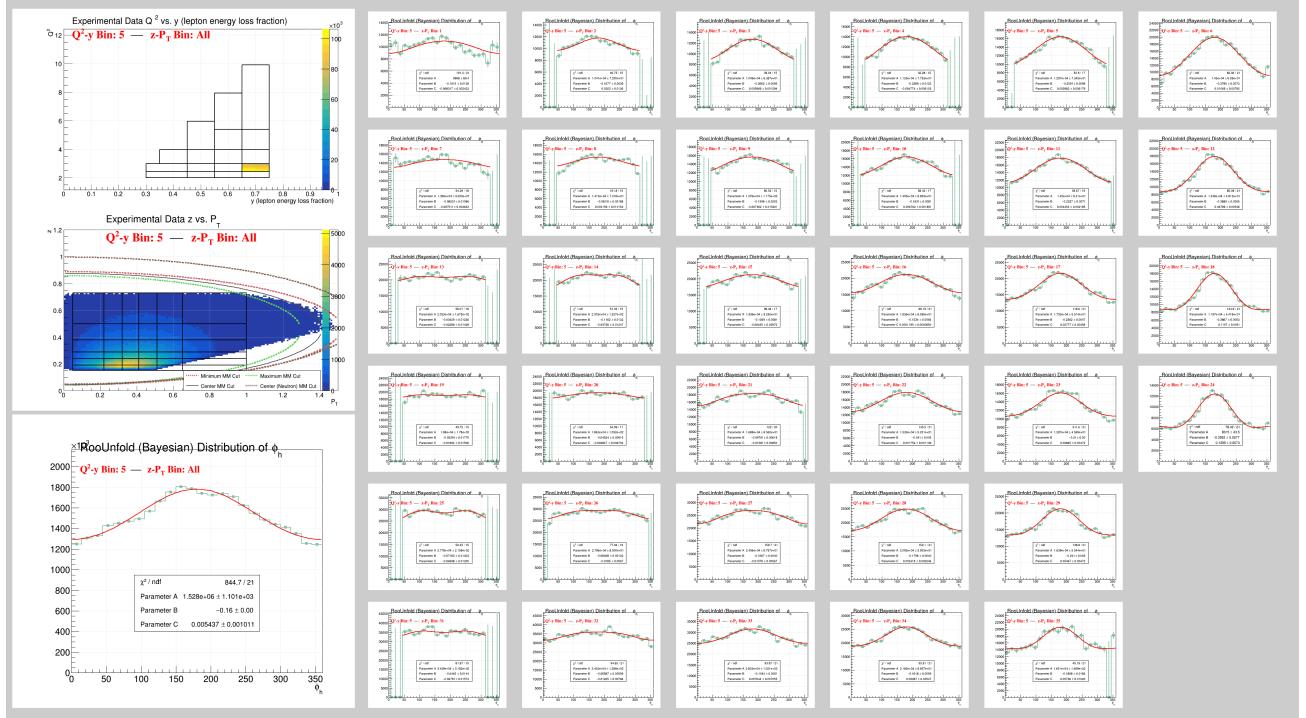


FIG. 46: 1D ϕ_h Bayesian Unfolding Corrections for all events in Q^2 - y Bin 5. See Figure 22 for more (basic) details regarding the layout of this type of image

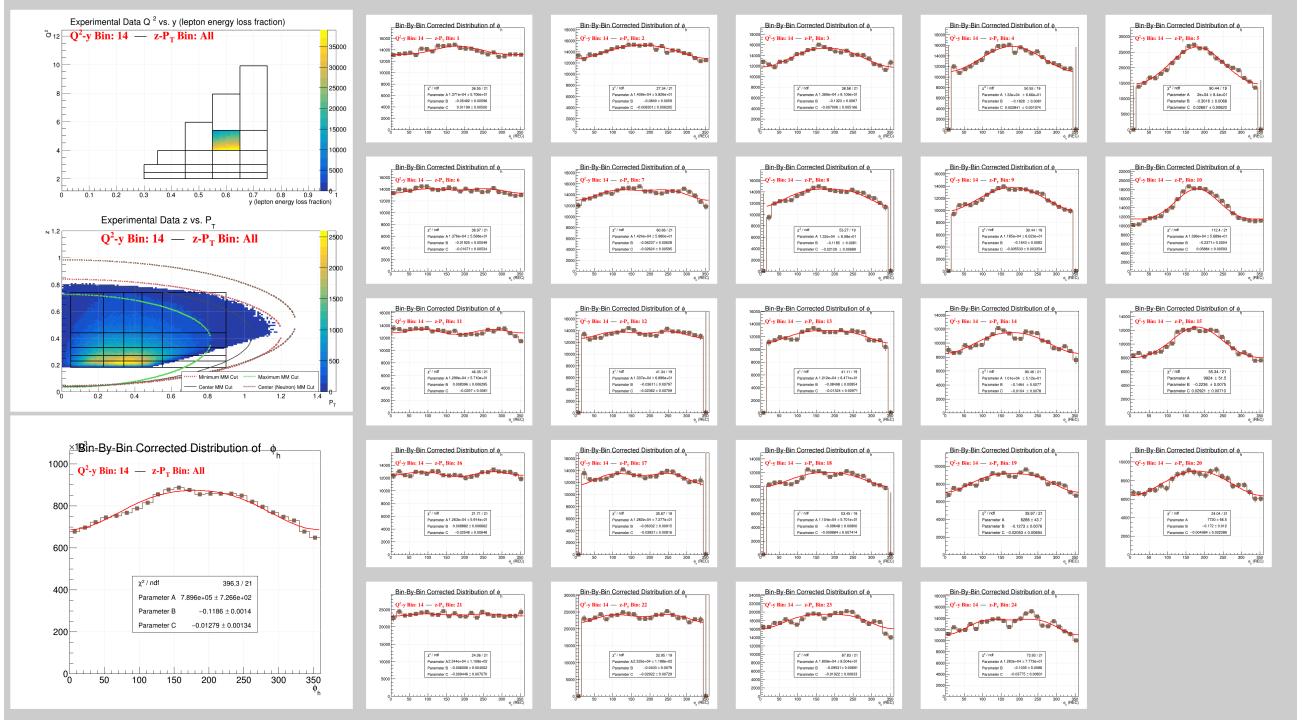


FIG. 47: 1D ϕ_h Bin-by-Bin Acceptance Corrections for all events in $Q^2\text{-}y$ Bin 14. See Figure 22 for more (basic) details regarding the layout of this type of image

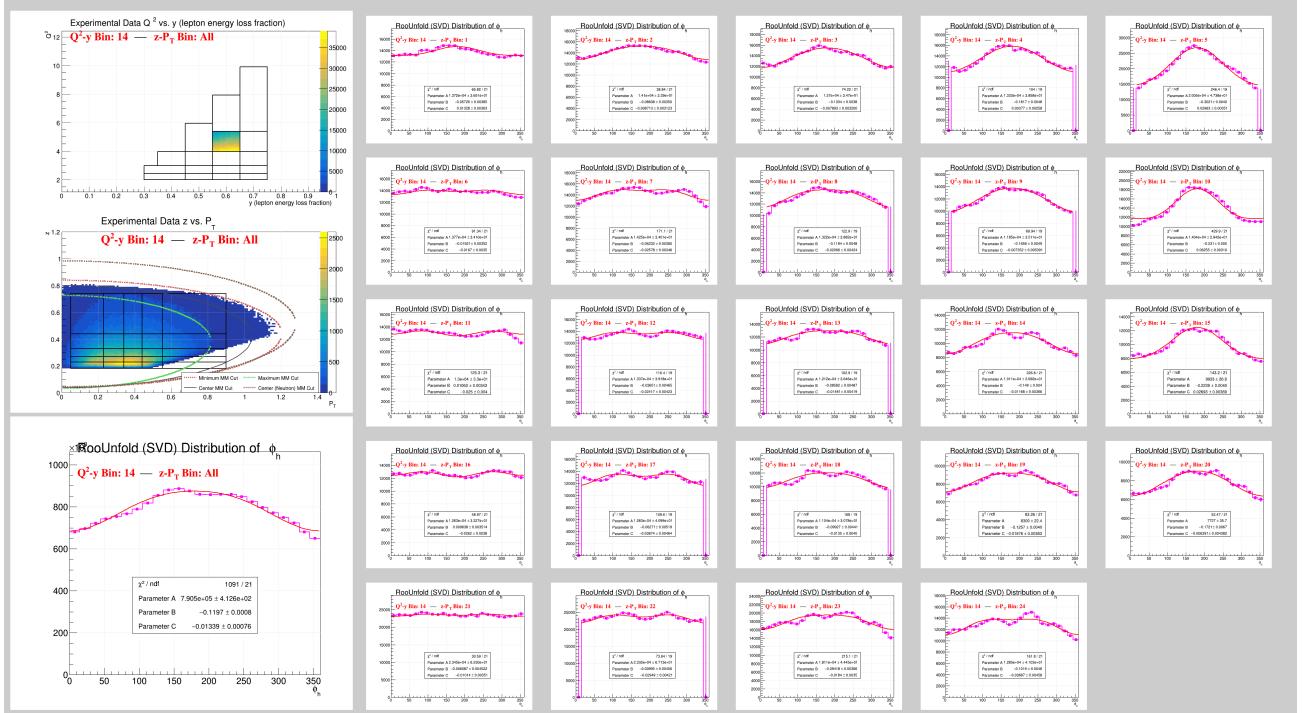


FIG. 48: 1D ϕ_h SVD Unfolding Corrections for all events in $Q^2\text{-}y$ Bin 14. See Figure 22 for more (basic) details regarding the layout of this type of image

method, which utilizes an iterative approach to predict the unfolded Response matrix based on

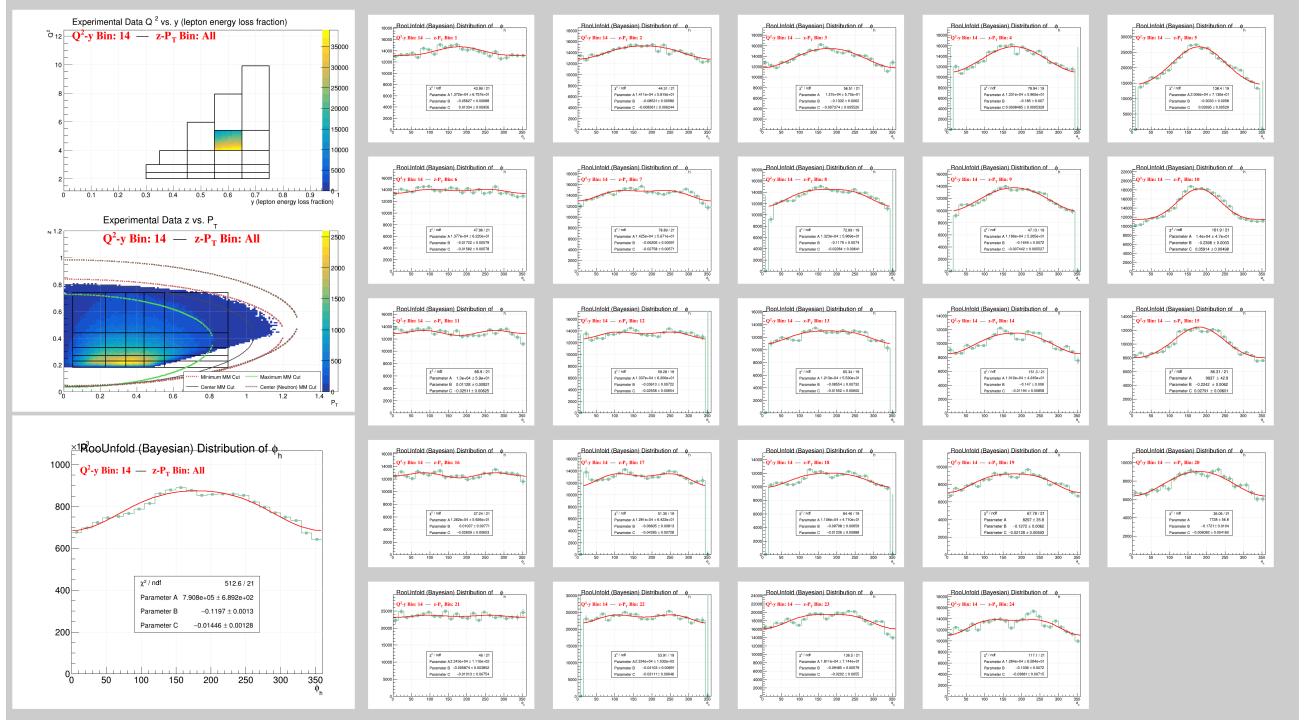


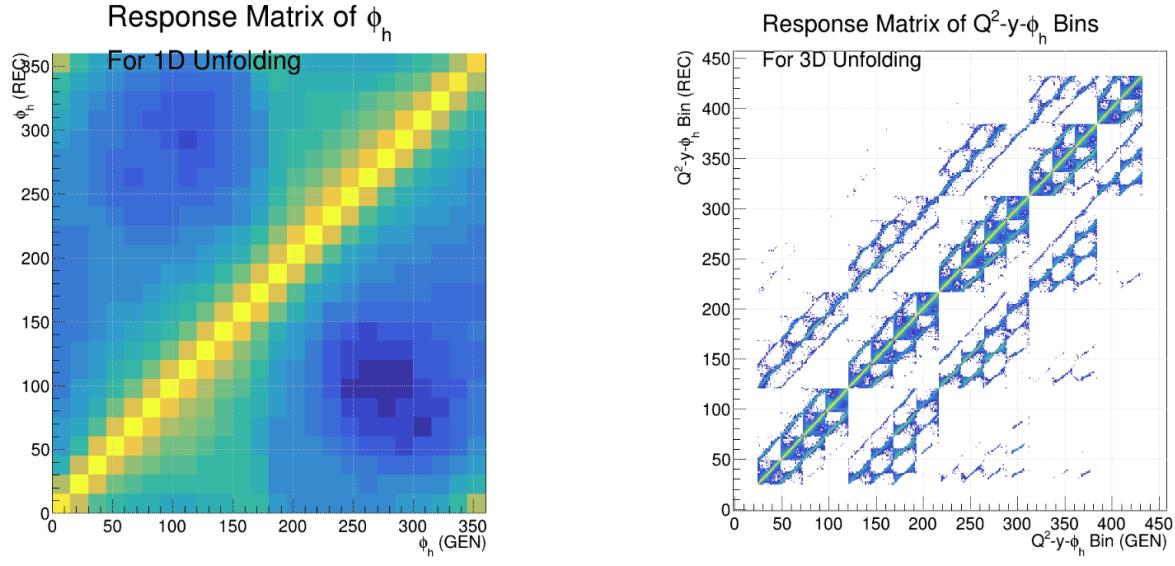
FIG. 49: 1D ϕ_h Bayesian Unfolding Corrections for all events in Q^2 - y Bin 14. See Figure 22 for more (basic) details regarding the layout of this type of image

elements of Bayesian statistics, is the most practical choice for this analysis. This is evident as the Bayesian method's iterative procedure does not rely on continuous variables, making it more suitable for multidimensional problems than other available unfolding methods.[14] Moreover, during the initial rounds of testing, the Bayesian method has demonstrated exceptional performance relative to the other unfolding methods tested, aligning well with the expectations set by its inherent design. Given that the ultimate goals of this unfolding procedure will involve correcting for acceptance and bin migration in five dimensions (i.e., in Q^2 , y , z , P_T , and ϕ_h), the Bayesian method emerges as the optimal choice.

Performing the ultimate five-dimensional unfolding corrections will require a flattened multidimensional binning scheme whereby all five variables can be unfolded simultaneously by unfolding the Response matrix of the flattened bins. This flattening of the kinematic bins is achieved by assigning all of the existing kinematic bins (defined in [Kinematic Binning](#)) with an integer value that can be combined to form the definition of the flattened multidimensional bins, which can then be treated as a single variable that can be unfolded with one-dimensional methods. All events that exist outside of the defined binning schemes are assigned a bin value of 0 for the construction of these flattened binning schemes (i.e., if an event is not found within any of the Q^2 - y bins, both the Q^2 - y and z - P_T bin numbers will be 0, while if it is found in a defined Q^2 - y bin and outside the defined z - P_T bins, only the z - P_T bin number will be given as 0). These '0' bins aid in the handling of events from outside the kinematic regions of interest, with future developments being planned to expand these bins with sets of 'migration bins' to replace this single '0' bin.¹⁴ With the large set of kinematic bins for each of the five crucial variables, handling the appropriate Response matrix becomes a practical challenge that must be overcome to complete the analysis. Figure 50 shows current examples of the Response

¹⁴ These 'migration bins' would serve the same purpose as the '0' bins, but would split the region defined by these bins 0 into separate bins to better account for the directions that these external migrations are coming from.

338 matrices for the one-dimensional unfolding of ϕ_h as well as the multidimensional unfolding of Q^2 , y ,
 339 and ϕ_h based on a flattened binning scheme of those three variables. Figures 51 to 52 show similar
 340 3D response matrices constructed by flattening the z , P_T , and ϕ_h bins in some of the individual Q^2 - y
 341 bins (as well as their resulting corrected/unfolded distributions). No 5D response matrix combining
 342 all of these variables simultaneously has yet been created at the time of this analysis note's release.¹⁵



(a) Example Response matrix of the ϕ_h variable.
 Uses all events without additional cuts for kinematic Q^2 - y - ϕ_h bins.
 binning.

(b) Example Response matrix of the flattened
 Uses all events without additional cuts
 for kinematic binning.

FIG. 50: Plots are drawn with a logarithmic scale for the z-axis. Reconstructed bins are plotted on the y-axis, while the Generated bins are plotted on the x-axis.

¹⁵ The 5D matrix is currently a work-in-progress delayed by the computer resources needed for its creation (i.e., computer memory and runtime).

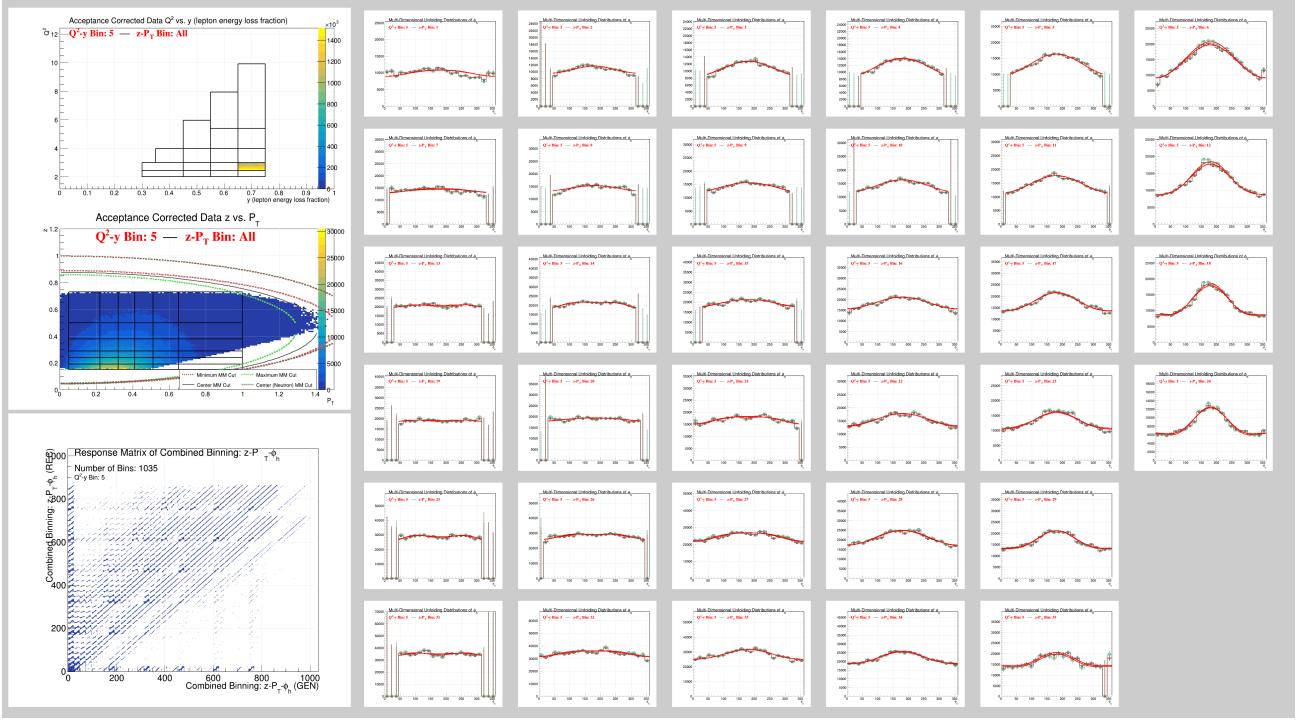


FIG. 51: ϕ_h Comparisons between the Multidimensional Acceptance Correction methods for all events in Q^2 - y Bin 5. The 2D plots have been corrected for acceptance (See [Cosine Moment Extractions](#) for details). The 2D response matrix in the bottom left corner shows flattened multidimensional bins used as the inputs for the unfolding algorithms. The Brown points in the histograms to the right show the distributions after applying a Bin-by-Bin Acceptance Correction while the dark green points show the distributions after applying the Bayesian Unfolding procedures

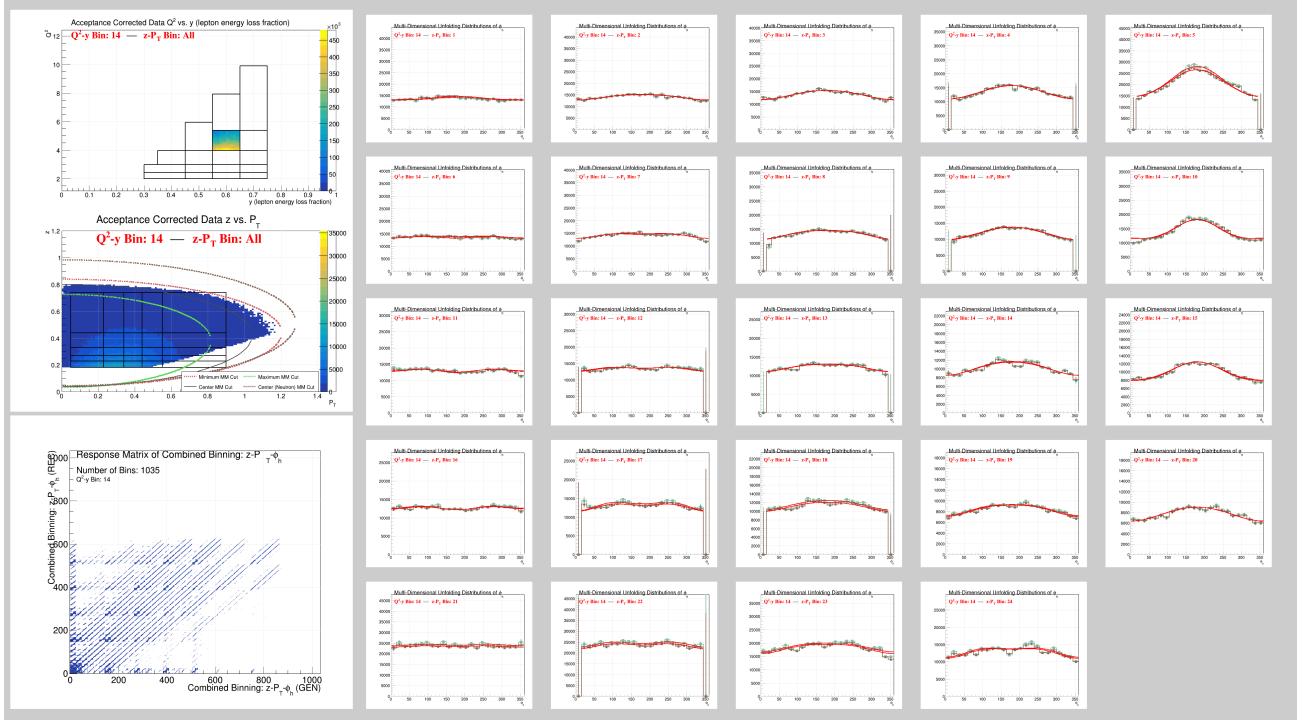


FIG. 52: ϕ_h Comparisons between the Multidimensional Acceptance Correction methods for all events in Q^2 - y Bin 14. See Figure 51 for more details

1. Acceptance Closure Tests

To ensure the acceptance corrections are applied appropriately, as described above, several closure tests can be performed to verify that the corrections are working as expected. A straightforward test is to provide the correction algorithms with the Monte Carlo simulated reconstructed data as if it were the experimental data that one wishes to correct/unfold. Comparing the resulting corrected distributions to the original generated Monte Carlo data (which, in the case of these closure tests, can be called ‘MC True’) would demonstrate whether the algorithm is working under the most ideal conditions possible (i.e., where the Monte Carlo perfectly describes all of the detector responses and effects in the collected data to be corrected). Figures 53 to 54 demonstrate the success of this test as all acceptance correction methods return the same distributions as the ‘MC True’ distributions in the images (see the light blue points in the plots).¹⁶

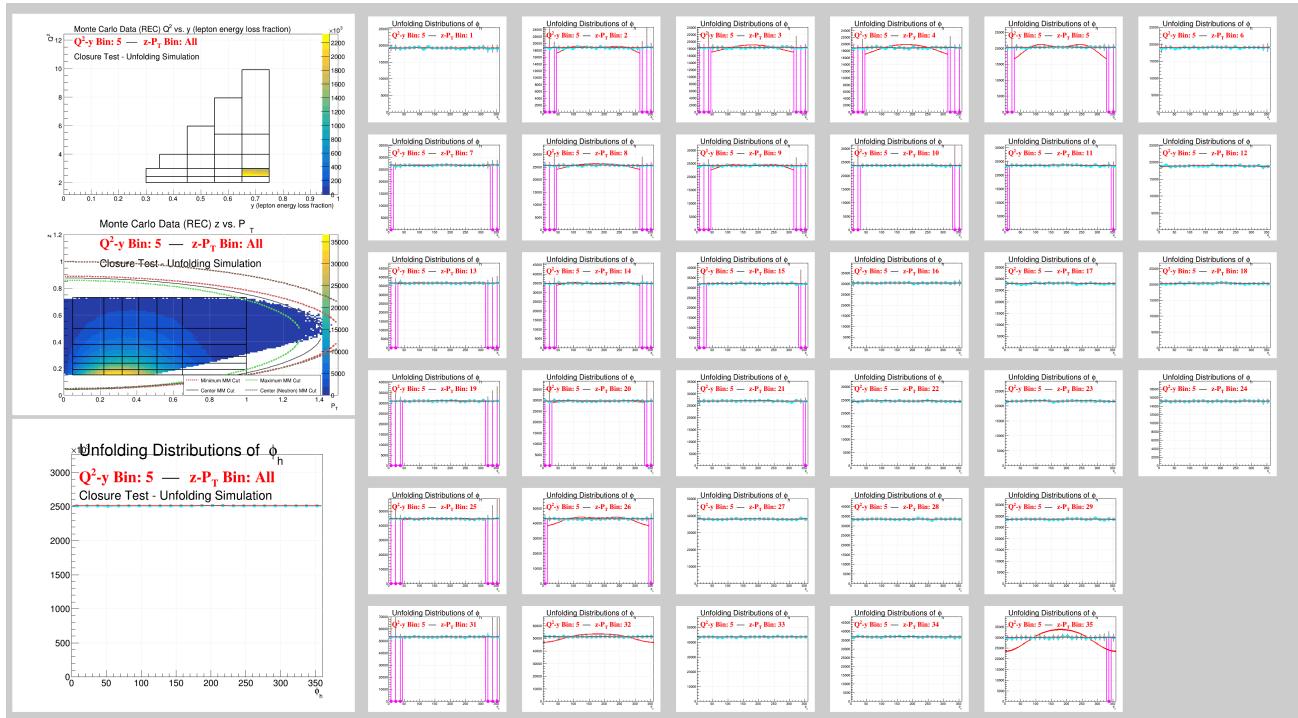


FIG. 53: ϕ_h Comparisons between the Simulated Acceptance Corrections for closure tests in Q^2 -y Bin 5. The distributions being corrected are the same as those used to construct the response matrix, so the corrections should all be flat like the light blue lines which shows the ‘MC True’ distributions that the corrections are trying to recreate.^a See Figure 44 and Figure 51 for more details on the layout of these images

^a See Footnote 16 for details on the quality of some of these fits.

For other specific examples, see either the [Plots Requested for Release](#) section of this note or the following web page:¹⁷ https://userweb.jlab.org/~richcap/Interactive_Webpage_SIDIS_richcap/Interactive_Unfolding_Page_Updated.html

¹⁶ Some plots shown in this closure test have poor fits due to the initial fitting parameters that may be biasing the results towards what is expected from the data. The unfolded/corrected results do appear better than these fits, which are technically a separate issue from this closure test. See [Plots Requested for Release](#) for a closer look on specific plots that demonstrate the conclusions put forward in this version of the analysis note.

¹⁷ This link is directed towards a web page where extra versions of these plots and others from this analysis note are kept and updated on a regular basis (includes additional descriptions of individual bins).

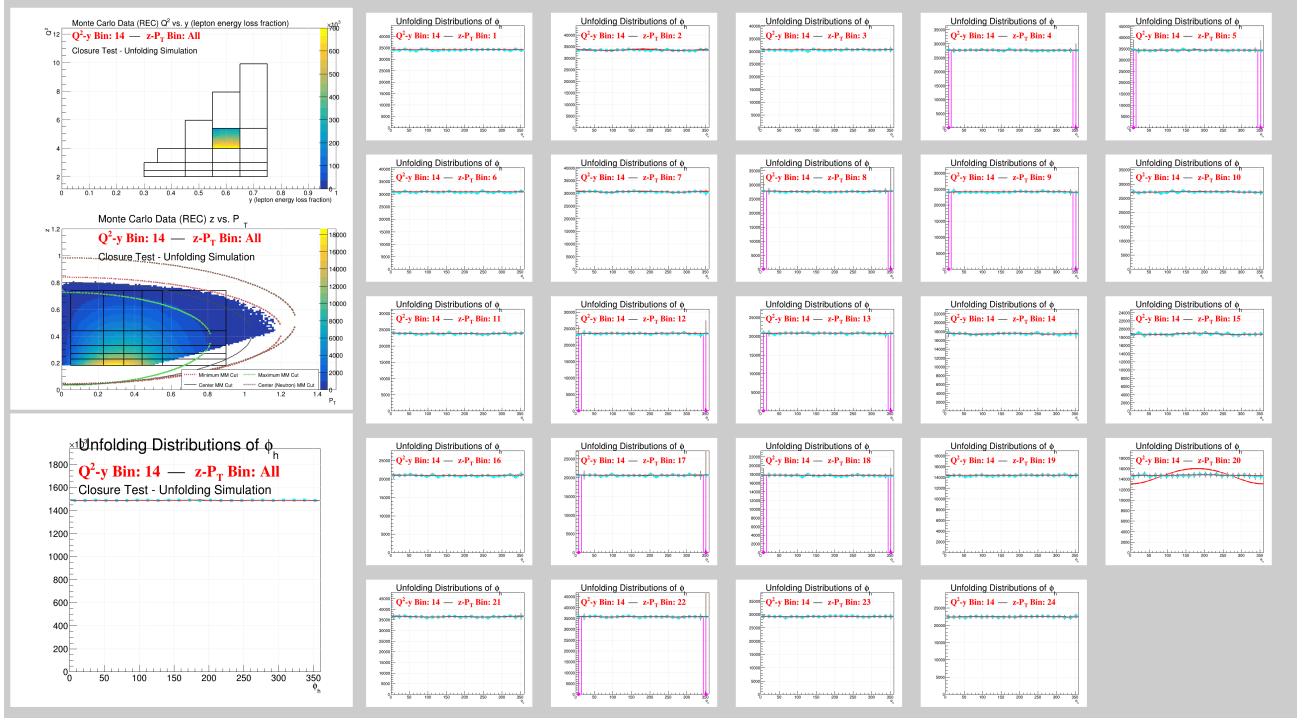


FIG. 54: ϕ_h Comparisons between the Simulated Acceptance Corrections for closure tests in Q^2 - y Bin 14. See Figure 53 for more details

In addition to using the same Monte Carlo to test that the corrections are applied correctly, one may also test whether they would work without first knowing the underlying physics this analysis seeks to measure. As the Monte Carlo generator did not include any ϕ_h modulations when it was initialized, a test of its ability to correct another simulation that does have some pre-defined modulations helps to verify the applicability of these corrections while using a simulation which does not accurately reflects all of the underlying physics. Using a weight function in the form of:

$$Weight = 1 + B \cos \phi_h + C \cos 2\phi_h \quad (25)$$

and applying it based on the generated ϕ_h angle when filling the generated and reconstructed histograms, one can modify the exciting simulation to create a separate data set that can be used to test how well the unmodulated simulation performs when trying to correct it. Figures 55 to 56 demonstrate the success of this test when the initial B and C parameters are set to a constant value (in the case shown, B and C were set equal to -0.5 and 0.025 respectively). Just as was the case with the previously discussed closure method, all acceptance correction methods return the same distributions as the ‘MC True’ distributions in the images, with each of the fits shown correctly predicting the fit parameters used to re-weight the distributions to within the margin of error specified by the fit functions.

Another simple test to ensure the multidimensional bins are flattened properly (i.e., so that the distributions never lose their intended meanings) is to check for any differences between the bin-by-bin corrections when the flattened bins are used versus the standard 1D histograms. The identical nature of each fit of the bin-by-bin corrections in Figure 57 demonstrates this consistency, which holds for all histograms produced for this analysis.

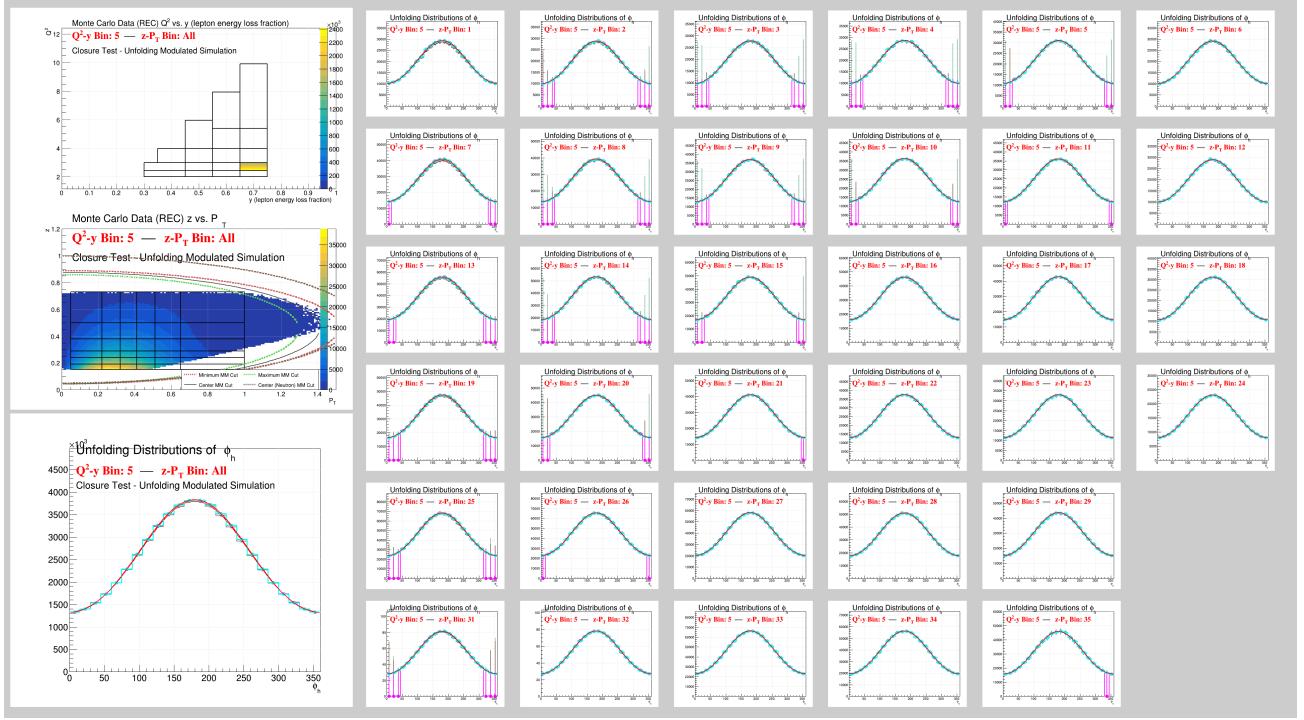


FIG. 55: ϕ_h Comparisons between the Simulated (**Modulated**) Acceptance Corrections for closure tests in Q^2-y Bin 5. The distributions being corrected have been modulated using Equation (25) with the initialized parameter values being $B = -0.5$ and $C = 0.025$. The light blue lines represent the ‘MC True’ distributions that the corrections are trying to recreate (which in this test are **not** flat)

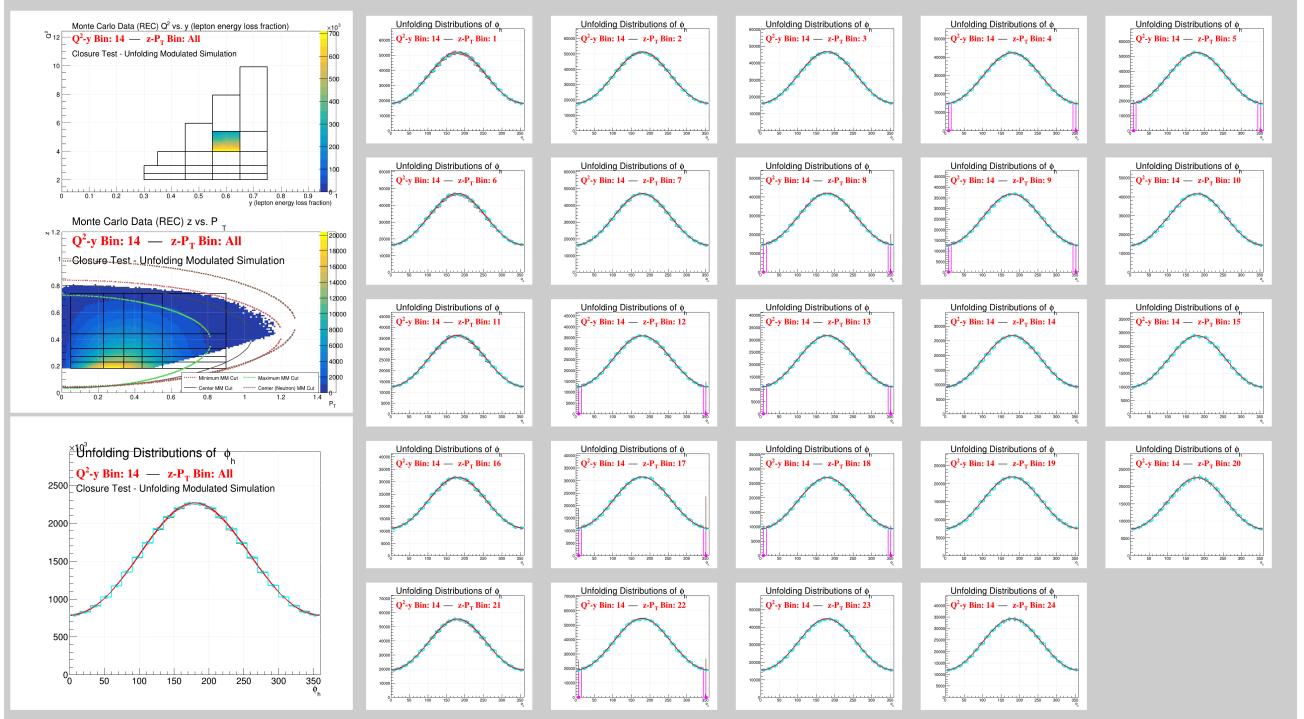


FIG. 56: ϕ_h Comparisons between the Simulated (**Modulated**) Acceptance Corrections for closure tests in Q^2-y Bin 14. See Figure 55 for more details

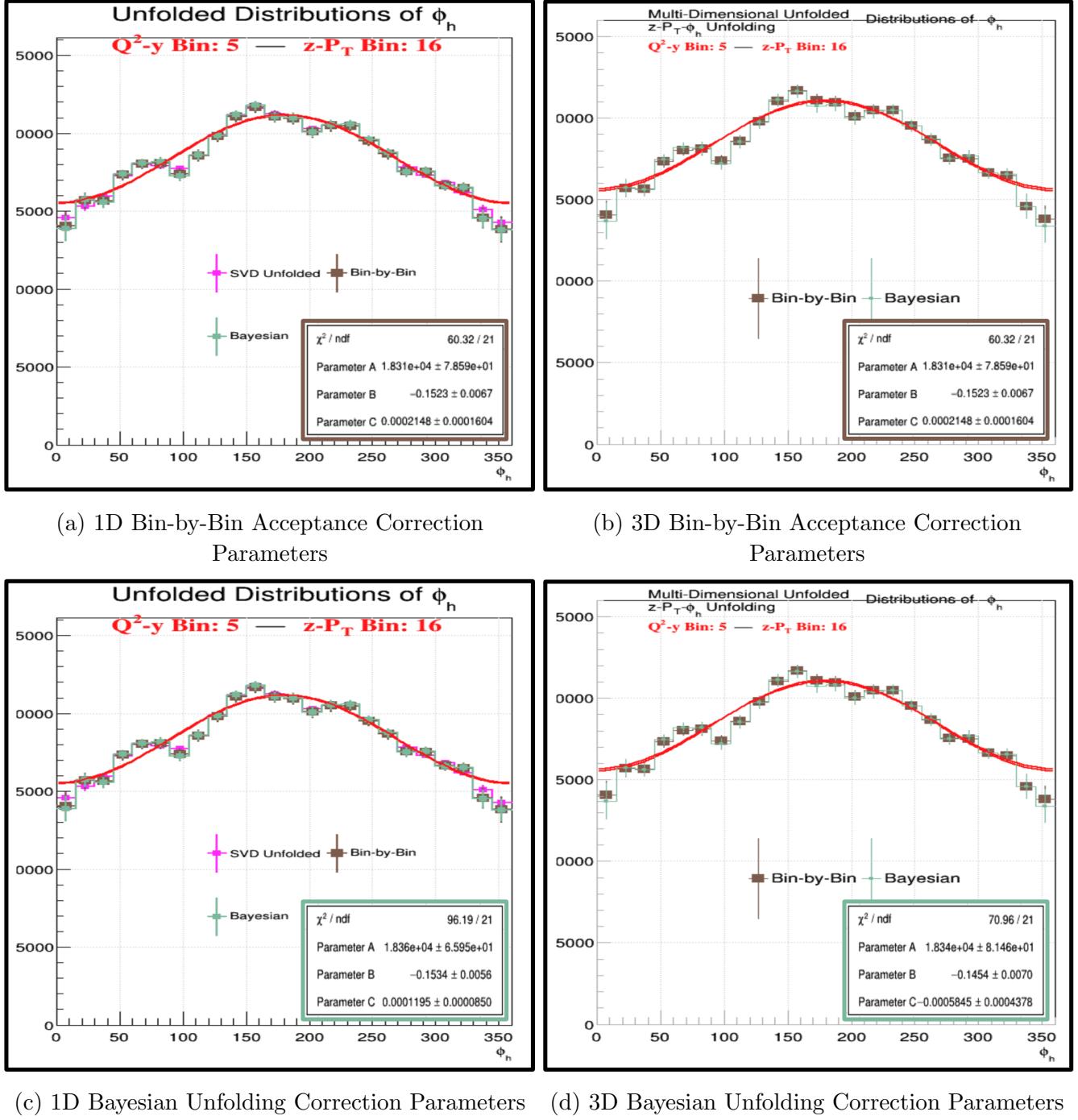


FIG. 57: Plots on the left show the corrected distributions using the 1D correction procedure for Q^2 -y Bin 5 - z - P_T Bin 16 (See Figures 44 to 49 for more information on these plots). Plots on the right show the same distributions using the 3D unfolding procedures for that same kinematic bin (See Figures 51 to 52 for more information on these plots). The top and bottom rows show the same plots, but display the fit parameters from separate correction methods. The y-axis gives the number of events in each ϕ_h after corrections (i.e., these plots are not normalized)

C. Other Corrections

Additional correction factors from [Equation \(21\)](#) that are to be considered include radiative corrections (R) and the bin averaging corrections (BC). BC is a conversion factor that converts the number of events within a given kinematic bin into the differential cross-section measurement corresponding to specific points within that bin. The radiative correction factor R can be determined in a similar method as the acceptance corrections, either by comparing simulated events with and without radiative effect taking the place of the number of reconstructed and generated events respectively so that:

$$R = \frac{\text{Number of Events Per Bin With Radiative Effects}}{\text{Number of Events Per Bin Without Radiative Effects}} \quad (26)$$

or by constructing a Response matrix with the radiative effects included in the simulation. Currently, the Monte Carlo simulation does not include these radiative effects, so the current plans call for another simulation with the radiative effects included.

D. Cosine Moment Extractions

The results shown in this analysis note only include distributions that have been corrected up to 3 dimensions in acceptance. As was described above, [Equation \(20\)](#) can be used to fit the corrected distributions of ϕ_h to extract the moments from the [cross-section equation](#). Typically, this will be done for the entire range of the ϕ_h variable, which is plotted from 0° to 360° across 24 bins. In cases where the edge bins have exceptionally low acceptance, however, the fit range will be reduced to avoid fitting bins with less reliable corrections. This applies in cases where very few or no events are measured in a particular reconstructed bin. As a particular ϕ_h bin's acceptance approaches 0, any corrections applied to that bin will become more prone to errors that could negatively impact the overall fits. The Bayesian unfolding method, in particular, requires attention to this matter, as the algorithm will still attempt to calculate a value for these bins, with their corrected statistics often being overestimated as compared to what would otherwise be expected. To mitigate these issues, empty bins in the reconstructed distributions and bins with particularly low acceptances (below 0.02) are automatically emptied (i.e., their bin content is set to 0) when plotting the corrected distributions. The fit range of the ϕ_h distribution is then adjusted to go from the first to the last filled bins in the given distribution being fitted. To avoid issues caused by low acceptances towards the center of the ϕ_h distributions, bins that are removed due to these acceptance issues also have their bin error artificially set to what their bin content would have been, plus their original bin error.¹⁸ Further improvements to how these edge bins are handled are still planned for future versions of this analysis note, with this current description reflecting how the plots shown in the [Plots Requested for Release](#) section of this analysis note were being fitted.

Once each of the individual histograms has been fit for each of the z - P_T bins, the fit parameters B and C are then plotted as functions of either z or P_T for each Q^2 - y bin. [Figures 58](#) to [59](#) present this information as functions of z , with the plots showing the P_T dependence via separate P_T bins. These bins are defined by the physical center of the z - P_T bins of a given column, while the value

¹⁸ Many of the kinematic bins that require this are likely to be redefined to avoid these types of acceptance issues entirely, so this aspect of the fitting procedure is expected to become less relevant in future iterations of the binning schemes.

⁴¹⁰ used to plot z is taken from the statistical center of the bin-by-bin corrected 2D plots of the z vs.
⁴¹¹ P_T distributions for each bin. To get these values, the standard bin-by-bin correction procedure is
⁴¹² applied to 2D histograms with finer binning relative to the defined schemes used in the analysis.¹⁹
⁴¹³ The statistical center of the bins is then taken from the average x or y value (depending on the
⁴¹⁴ variable being examined) of the events that lie in that particular bin. The standard deviation of
⁴¹⁵ these averages is then used as the x error bars for these plots.

⁴¹⁶ While the plots shown in this analysis note do not contain the full set of corrections required for
⁴¹⁷ this analysis²⁰, they do begin to show a clear dependence on the various kinematic variables, with
⁴¹⁸ the plots themselves indicating a larger dependence of the cosine moments on P_T as compared to z
⁴¹⁹ in most cases. Further work is still needed however to make any conclusive claims regarding these
⁴²⁰ observations.

¹⁹ The size of bins used is 0.01/bin for both variables.

²⁰ Only the 3D Acceptance Corrections have been applied. Still missing the other corrections described in [Other Corrections](#).

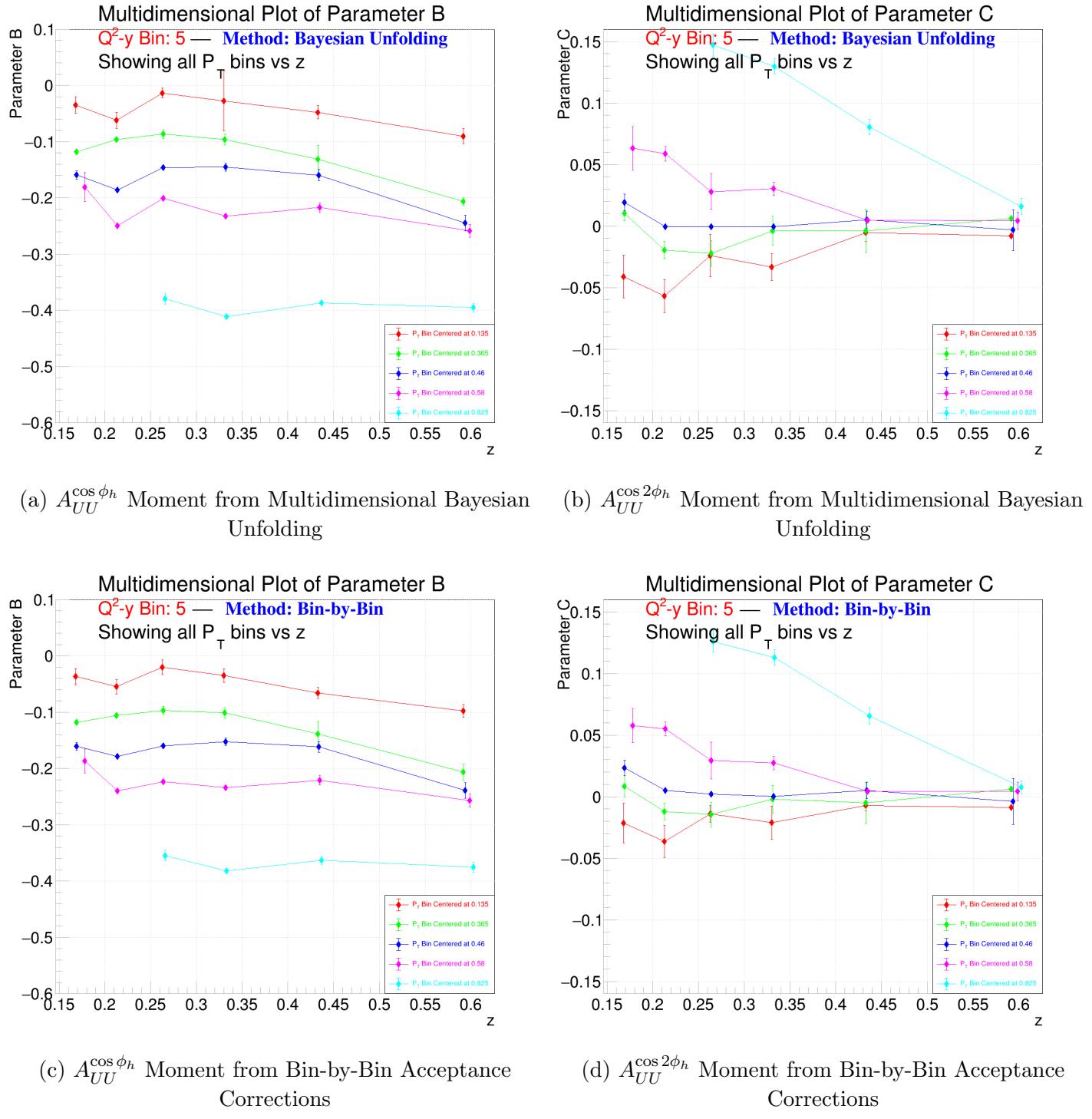


FIG. 58: Cosine Moments as functions of z in separate P_T bins for Q^2 -y Bin 5 (See Figure 51 for the individual distributions - points from the second P_T column were not included in these plots due to poor fitting)

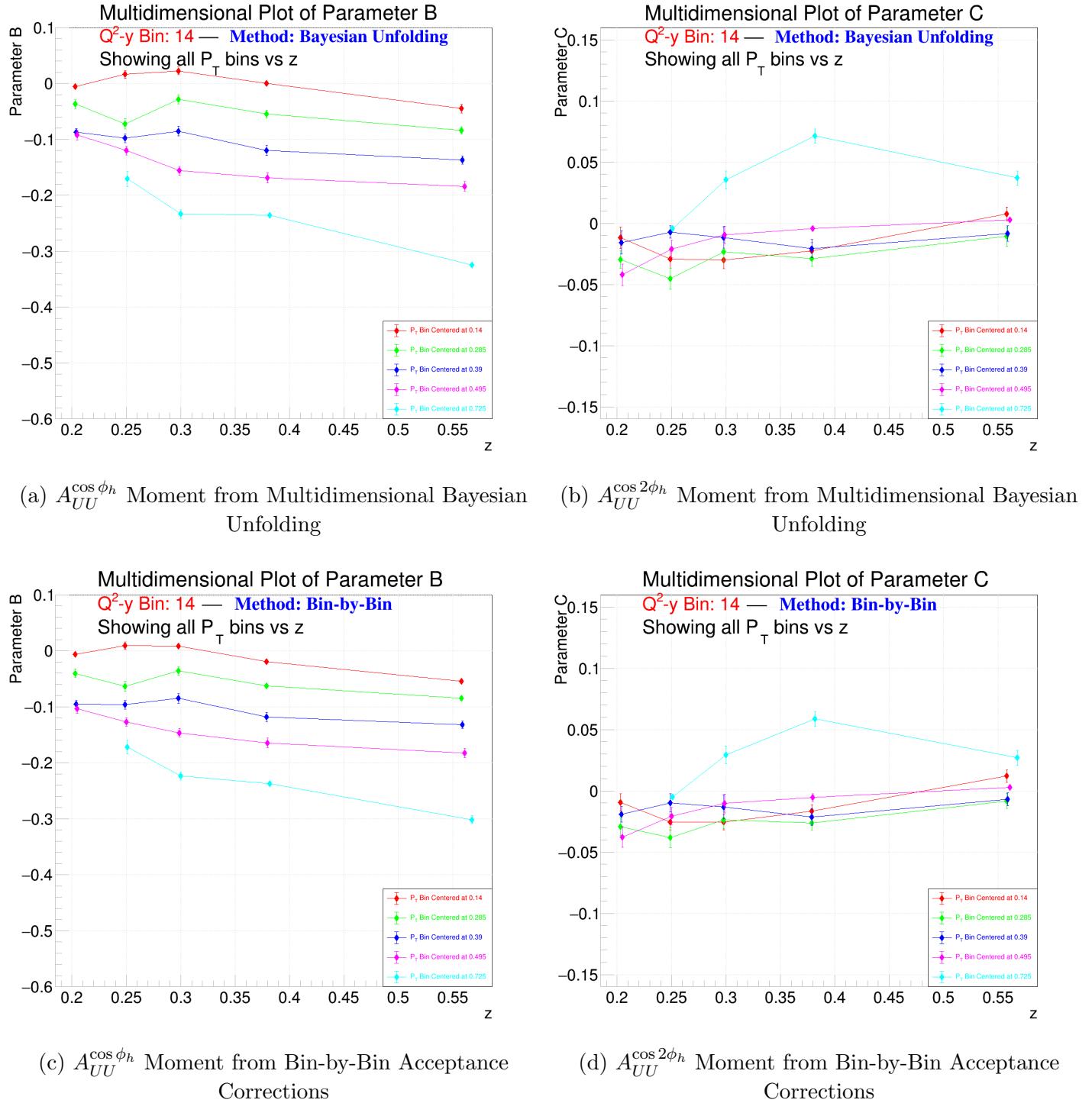


FIG. 59: Cosine Moments as functions of z in separate P_T bins for Q^2 - y Bin 14 (See Figure 52 for the individual distributions)

VI. PLOTS REQUESTED FOR RELEASE

⁴²¹ The images listed in this section are planned to be shown at the upcoming “2023 Fall Meeting of
⁴²² the Division of Nuclear Physics of the American Physical Society and the Physical Society of Japan”
⁴²³ (DNP-JPS Meeting) in Waikoloa, Hawaii.²¹ Images already shown elsewhere in this analysis note will
⁴²⁴ be listed here, while other images not included elsewhere will be included below with a description.

A. Images already shown in this Analysis Note

- **Figures 2 to 5** - Analysis Cuts
 - Images that show the cuts applied in this analysis.
- **Figure 40** - Missing Mass vs. Momentum
 - Images show the effects of applying an additional smearing correction to the Monte Carlo simulation.
- **Figure 43** - Ratios of Missing Mass Peak widths vs. Momentum
 - Images that demonstrate the effectiveness of applying the additional smearing correction to the Monte Carlo simulation.
- **Figure 13** - Q^2 - y Bin Definitions with all events
 - This image shows the kinematic distributions of Q^2 , y , z , P_T , and x_B for all events that survive the analysis cuts. The plot showing the Q^2 vs. y distributions includes the definitions of the Q^2 - y bins, so this image will be used to help depict the binning schemes used in this analysis.
- **Figure 50 (b)** - Response Matrix of Q^2 - y - ϕ_h bins
 - One of the possible examples of the response matrix that can be used for multidimensional acceptance unfolding corrections.
- **Figure 57** - Acceptance Corrected/Unfolded distributions of Q^2 - y Bin 5 - z - P_T Bin 16
 - Images that can be used to compare the 1D and 3D acceptance unfolding corrections in a particular kinematic bin.
 - One additional version of the plot for the 1D corrections will also be shown that simultaneously shows the fit parameters of the SVD unfolding method with the other two results shown in the images included in this note. See **Figure 66** for these values.
- **Figures 58 to 59** - Extraction of $A_{UU}^{\cos\phi_h}$ and $A_{UU}^{\cos 2\phi_h}$ Moments
 - Images which show the multidimensional extraction of the $A_{UU}^{\cos\phi_h}$ and $A_{UU}^{\cos 2\phi_h}$ moments from events in Q^2 - y Bin 5 and 14. Images show these parameters as functions of z in separate P_T bins.

²¹ Some images may only included as backup slides. Images listed in this section are not necessarily listed in a particular order.

B. Additional Images to be Released

425 Images shown here may have already appeared in one of the earlier images from this analysis note
 426 but will appear in this section in the manner that they are intended to be shown at the DNP-JPS
 427 meeting.

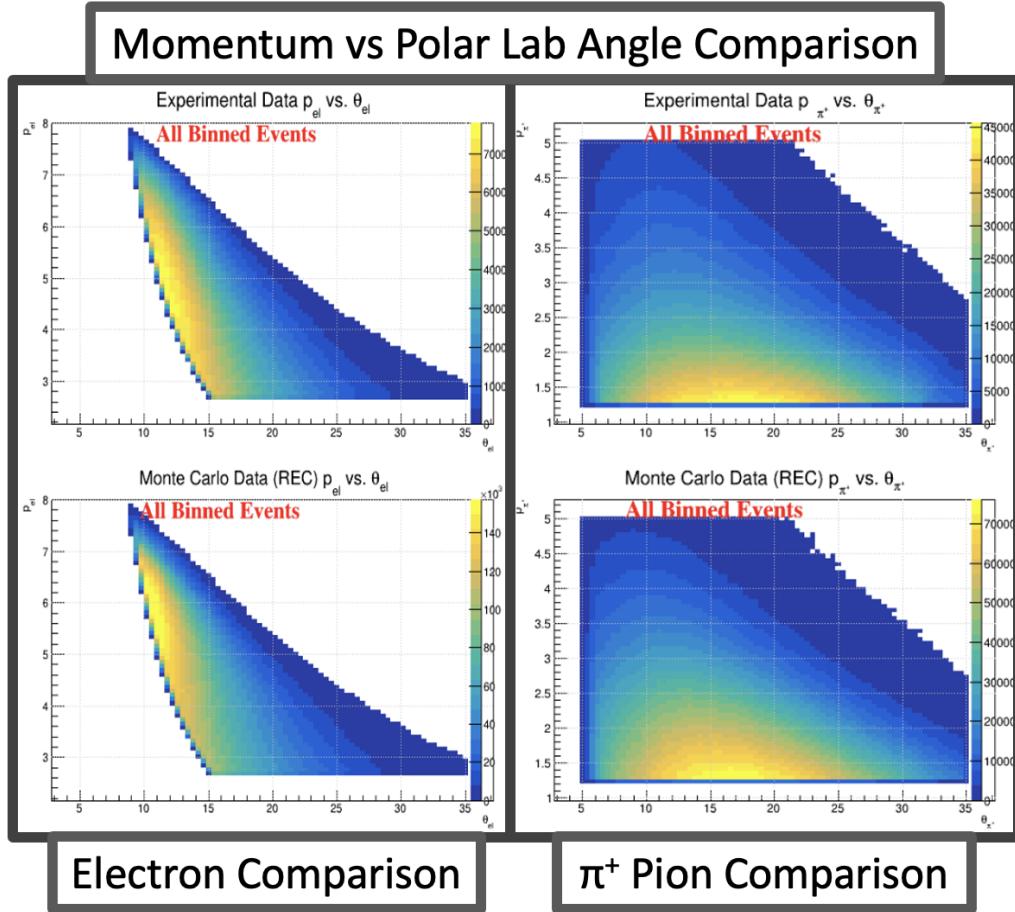


FIG. 60: Comparison of Data and Monte Carlo for Momentum vs. θ distributions for both the Electron and π^+ pion kinematics. Images in the top row come from the Experimental Data (after all analysis cuts), while the images in the bottom row come from the Reconstructed Monte Carlo data

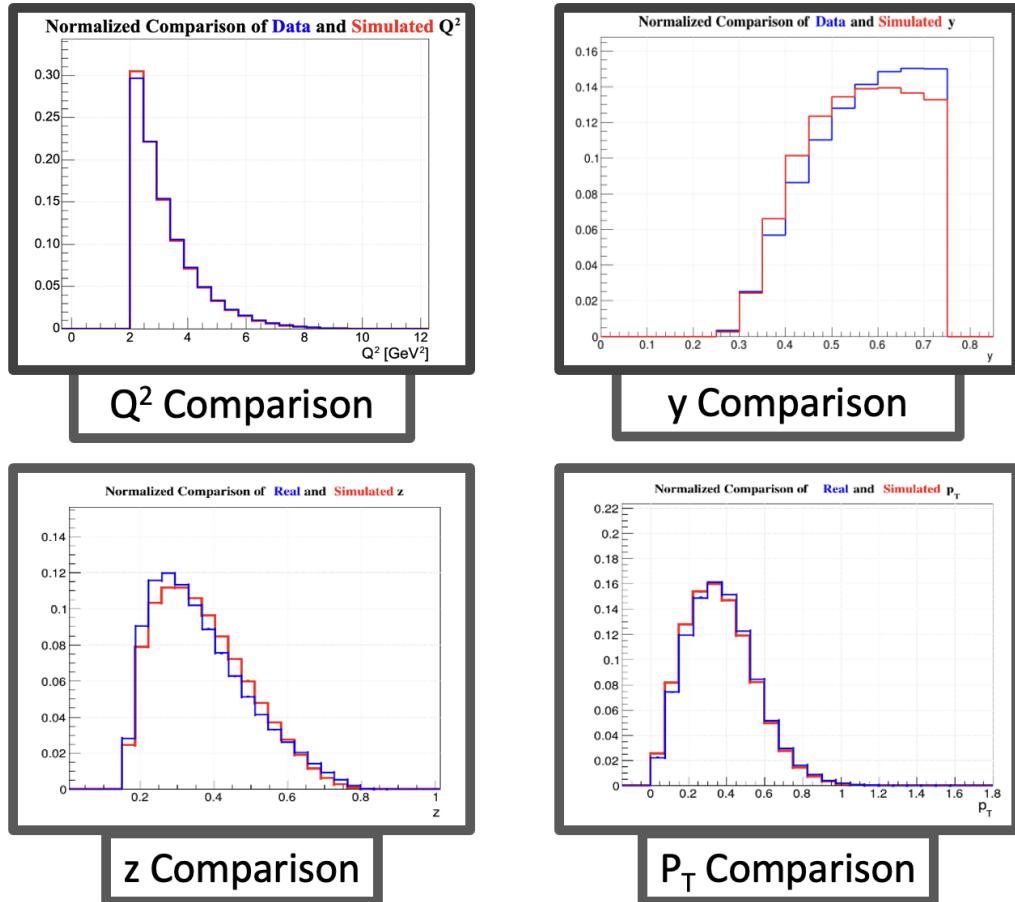


FIG. 61: Comparison of Data and Monte Carlo for the Kinematic Variables of Q^2 , y , z , P_T . Similar images are also included in [Figure 21](#)

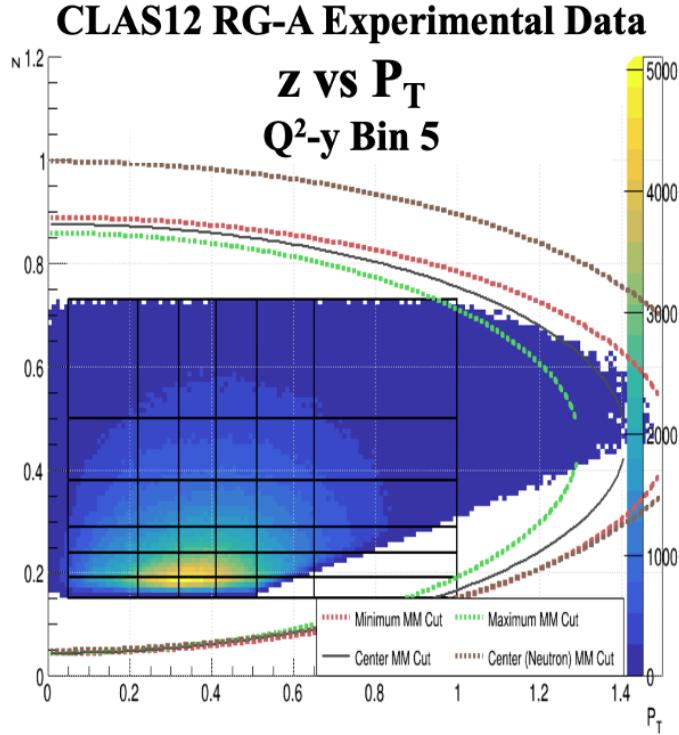


FIG. 62: z vs. P_T distribution for events in Q^2 - y Bin 5 (included as an example of the currently defined z - P_T bins). This image is also included in Figure 26

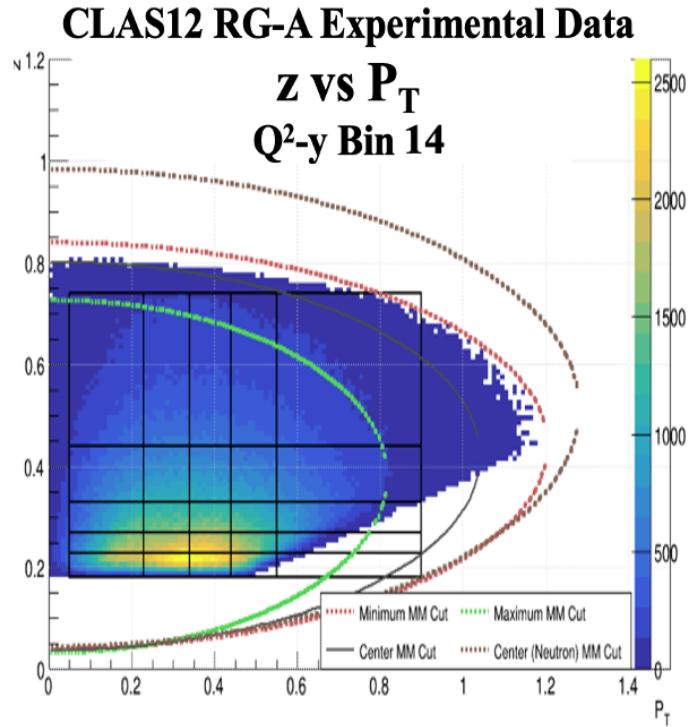


FIG. 63: z vs. P_T distribution for events in Q^2 - y Bin 14 (included as an example of the currently defined z - P_T bins). This image is also included in Figure 35

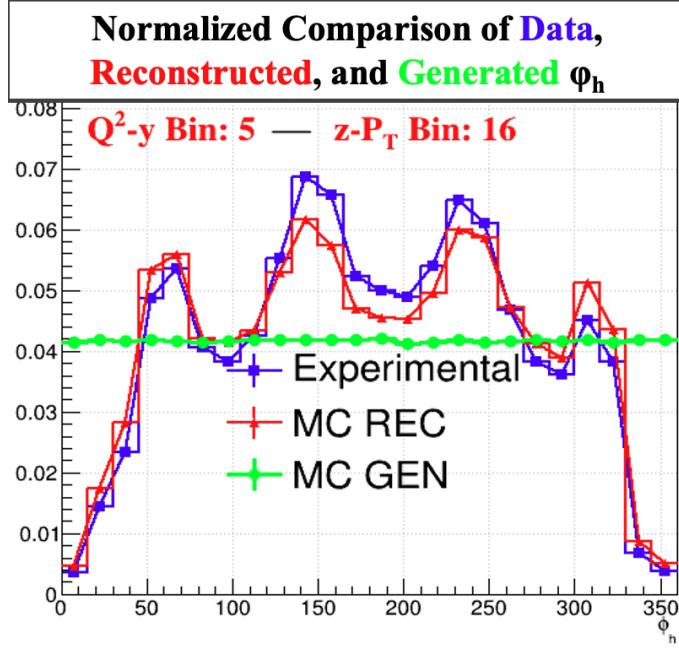


FIG. 64: ϕ_h distributions for events in Q^2 - y Bin 5 - z - P_T Bin 16. This image is also included in Figure 26

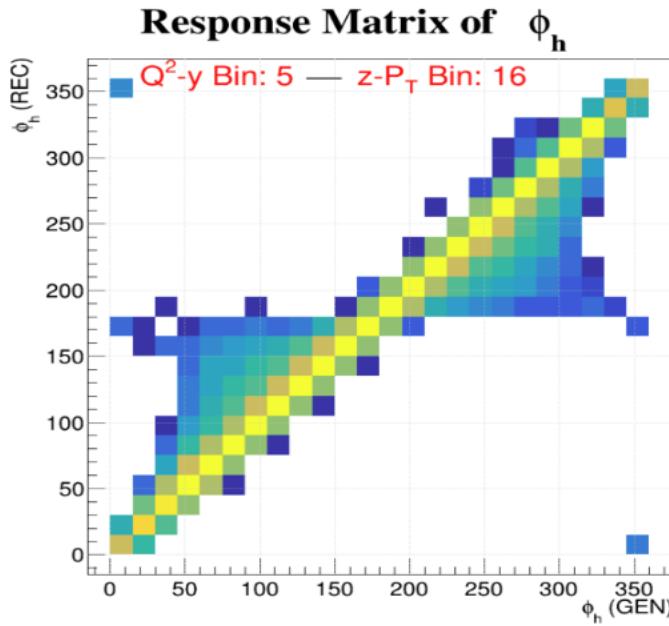


FIG. 65: Response Matrix for the 1D Unfolding process for events in Q^2 - y Bin 5 - z - P_T Bin 16. This matrix was used in the creation of some of the images in Figure 57 (i.e., the 1D unfolding examples)

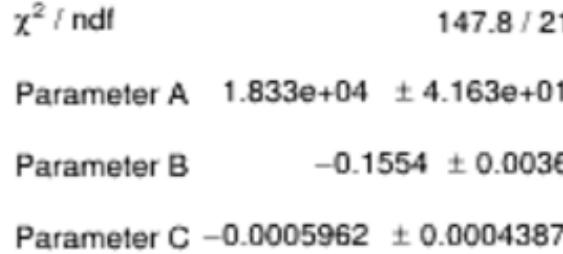


FIG. 66: SVD Unfolding fit parameters from the events in Q^2 -y Bin 5 - z - P_T Bin 16. These are the fit parameters of the SVD distributions shown in the 1D Unfolding Example in Figure 57

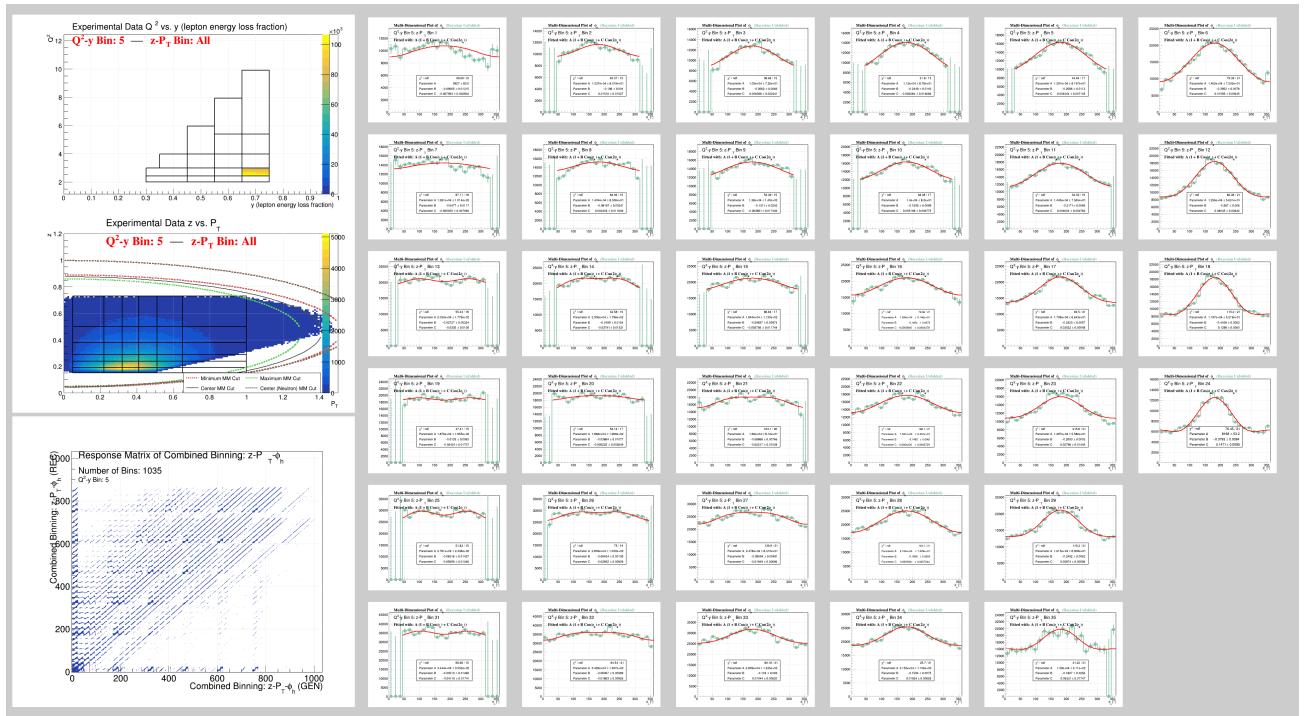


FIG. 67: 3D Bayesian Unfolding examples of each z - P_T bin in Q^2 -y Bin 5. These distributions are all also featured in Figure 51, but are separated here only to show the Bayesian Unfolded distributions

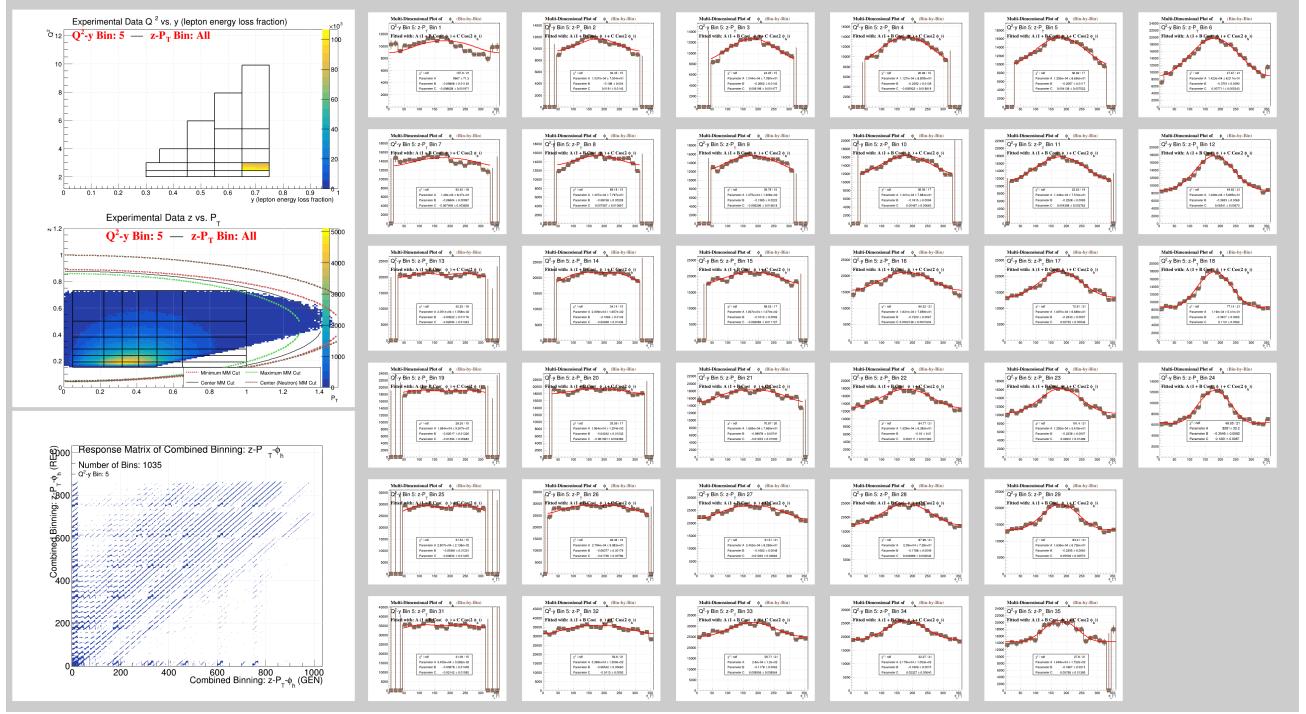


FIG. 68: Bin-by-Bin Acceptance Correction examples of each $z\text{-}P_T$ bin in $Q^2\text{-}y$ Bin 5. These distributions are all also featured in Figure 51, but are separated here only to show the Bin-by-Bin Corrected distributions

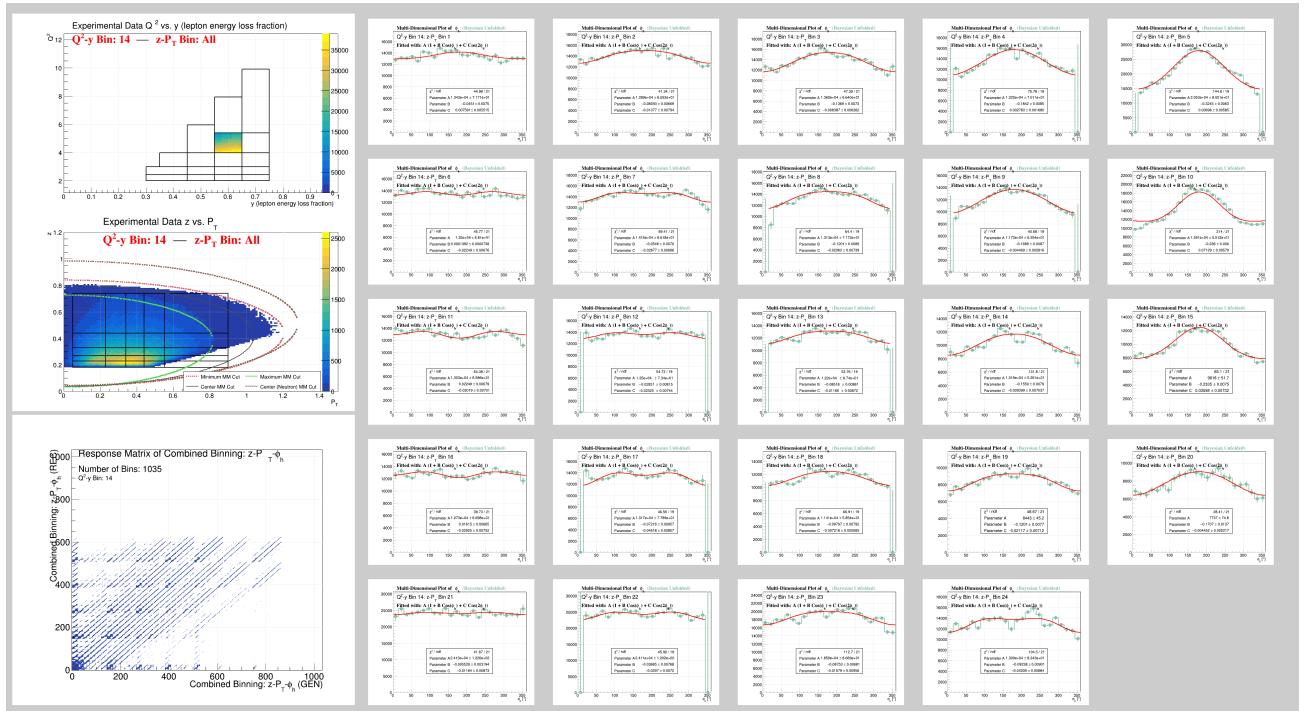


FIG. 69: 3D Bayesian Unfolding examples of each $z\text{-}P_T$ bin in $Q^2\text{-}y$ Bin 14. These distributions are all also featured in Figure 52, but are separated here only to show the Bayesian Unfolded distributions

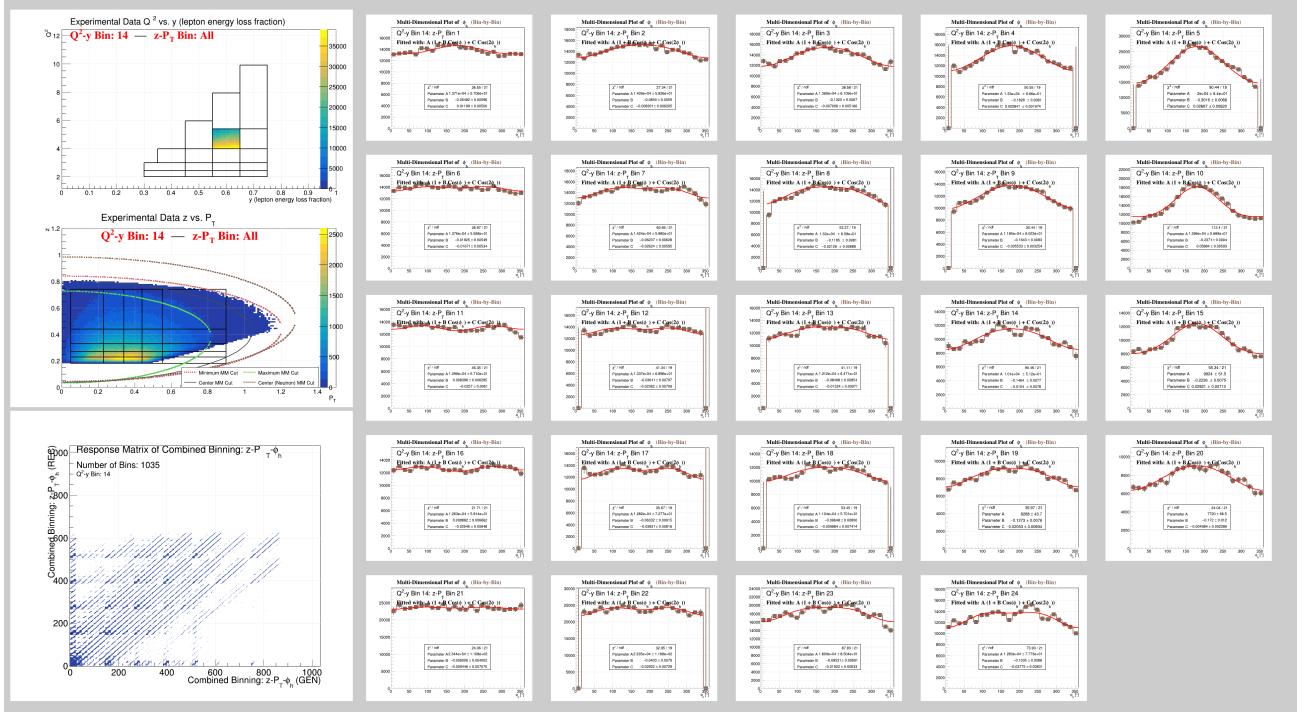


FIG. 70: Bin-by-Bin Acceptance Correction examples of each z - P_T bin in Q^2 - y Bin 14. These distributions are all also featured in Figure 52, but are separated here only to show the Bin-by-Bin Corrected distributions

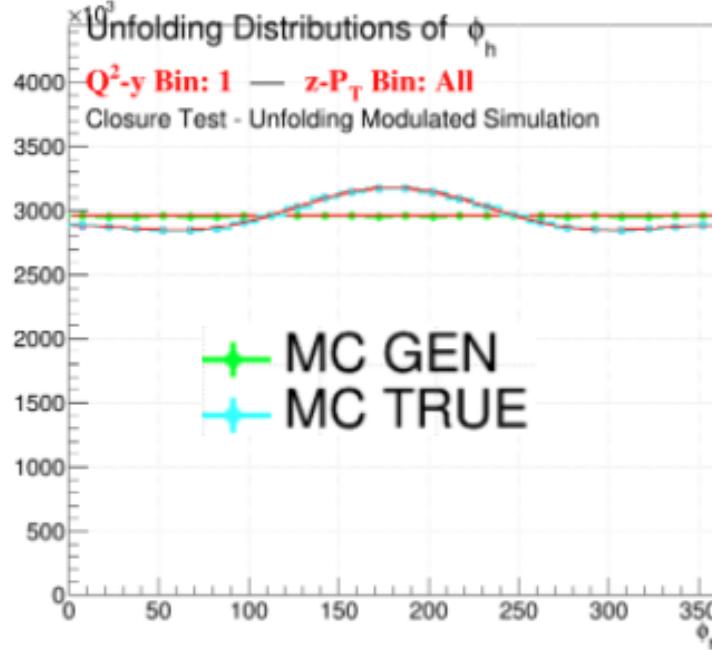


FIG. 71: Example of adding modulations to an existing generated Monte Carlo distribution of ϕ_h using Equation (25). Plot uses all events in Q^2 - y Bin 1 for the example with the weighing parameters set equal to $B = -0.05$ and $C = 0.025$

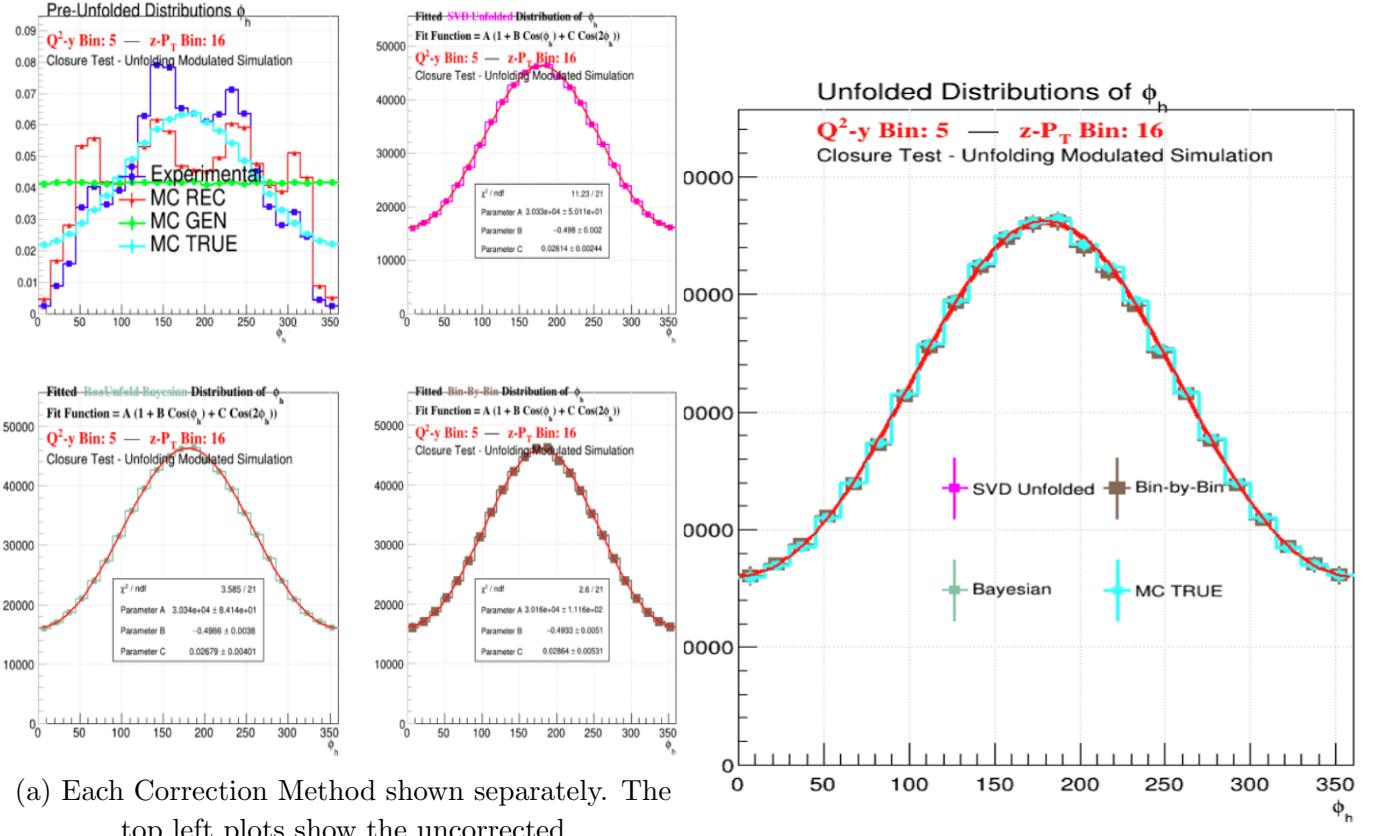


FIG. 72: Direct example of adding modulations to an existing generated Monte Carlo distribution of ϕ_h using Equation (25). The plots use events from Q^2 - y Bin 5 - z - P_T bin 16 with the weighing parameters set equal to $B = -0.5$ and $C = 0.025$. These plots were also included in Figure 55

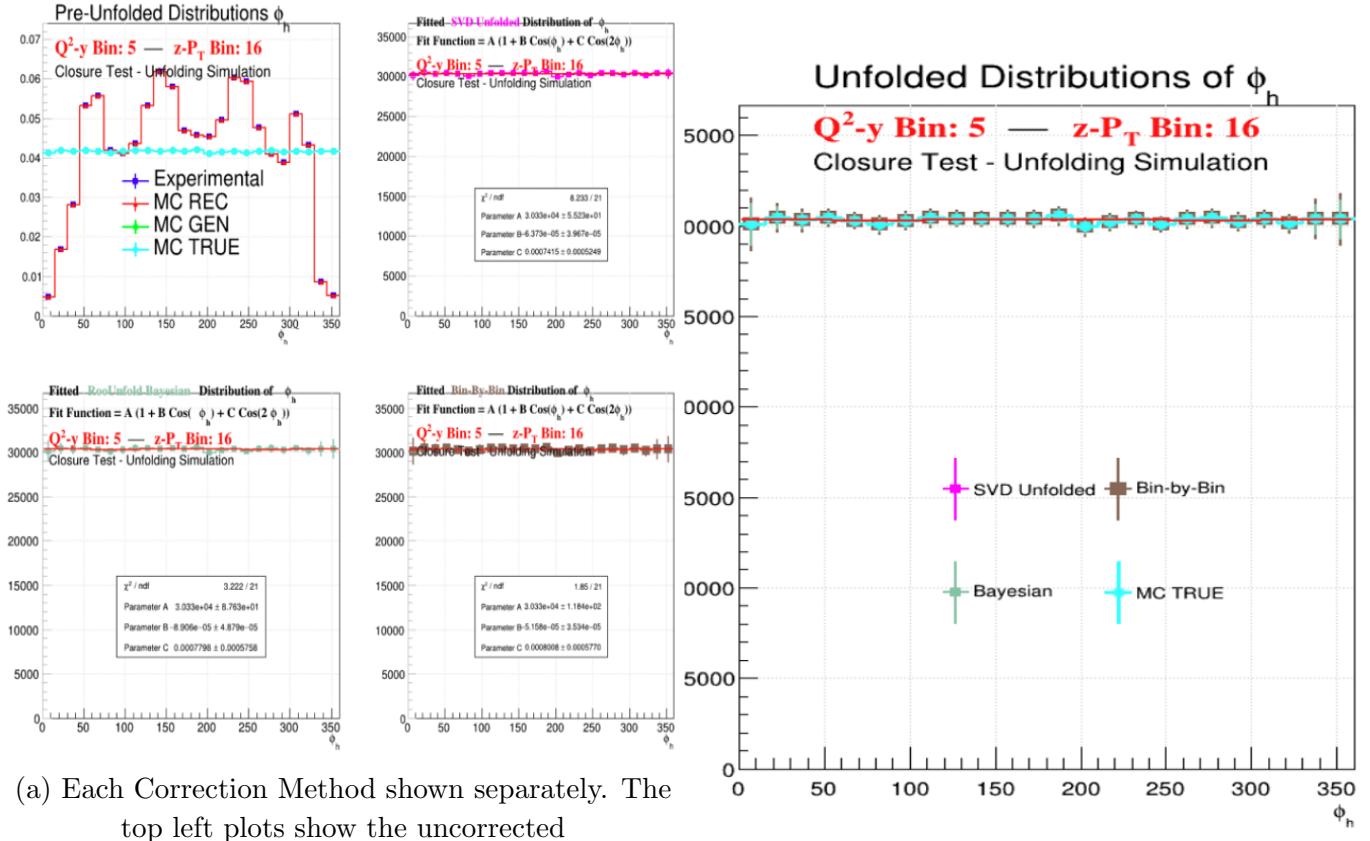


FIG. 73: Direct example of unfolding the unmodulated Monte Carlo distributions to return the initial generated distribution of ϕ_h . The plots use events from Q^2 - y Bin 5 - z - P_T bin 16 for the example. These plots were also included in Figure 53

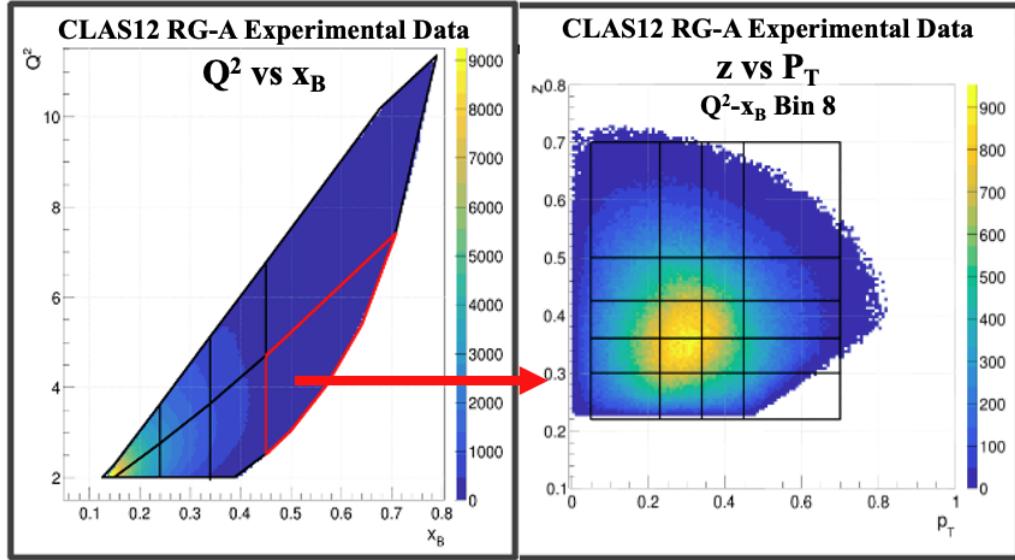


FIG. 74: Example of the original z - P_T bins using the old Q^2 - x_B binning scheme. Image shows the z - P_T bins for Q^2 - x_B Bin 8. This figure was already approved for release prior to Version v4 of this analysis note (now just included to demonstrate the changes made since that last update). See [Figure 77](#) for all other versions of this binning scheme as of the prior release version of this note

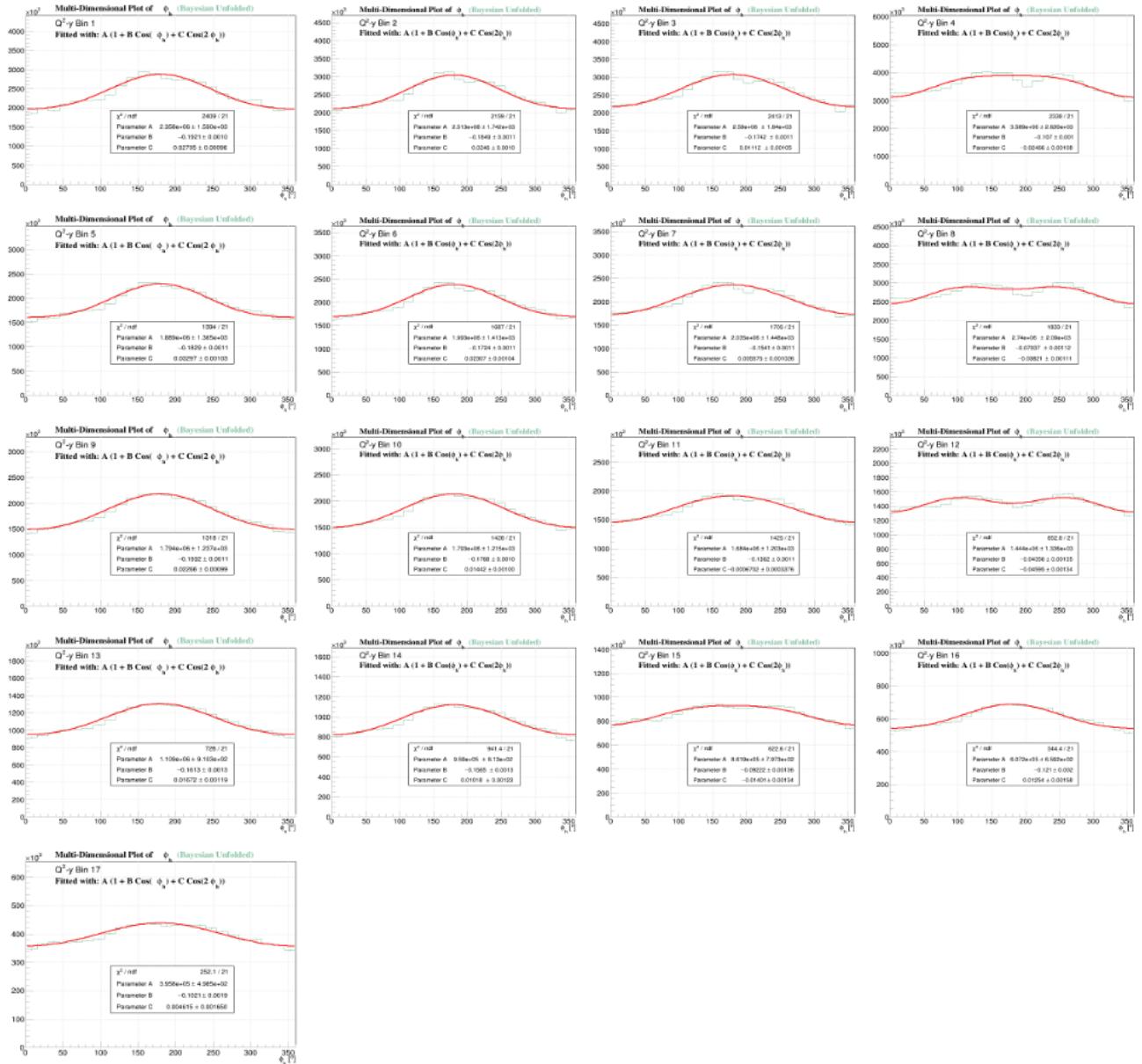


FIG. 75: Example of Multidimensional Bayesian Unfolding procedure with just the Q^2 , y , and ϕ_h variables. These fits are produced when unflattening the results of corrections made with the example Response matrix shown in Figure 50 (b)

VII. VERSION UPDATE NOTES

⁴²⁸ This section is meant to give a brief description of some of the most significant changes to this
⁴²⁹ analysis note as of the prior iteration. Changes made as of Version v4 of this note include:

- General revisions to the structure of this note (including a slight rewrite of several sections).
- Switched from using a Q^2 - x_B binning scheme to a new Q^2 - y scheme for reasons explained in this note.
 - See [Kinematic Binning](#) for details on the new binning scheme and [Figure 77](#) for previously released images of the older scheme.
- Drew cut lines for the missing mass cuts in some of the images showing the z - P_T distributions.
- Updated the description of the momentum corrections/smearing to reflect the current status of those efforts (momentum corrections for the experimental data have been finalized and released to the whole of the CLAS Collaboration).
 - Exact experimental corrections used are not included in this note (link is provided to where they have been made publicly available instead).
 - Corrections used in the Monte Carlo are included in the relevant section of this analysis note.
 - The additional Monte Carlo smearing corrections are now applied in a different, simplified way (as described in this note), with only each particle's momenta being affected by the smearing function (instead of the momentum **and** lab angles of the particle's being modified by the smearing function).
 - No longer referencing FX's smearing function as was previously done in prior versions of this analysis note.
 - * FX's smearing function had been used as a prior baseline for making improvements to the smearing correction developed for this analysis, but the new methods described in this note have now replaced the need to refer to this original smearing function.
- Added the description of various closure tests now being performed to check that the application of the unfolding corrections is being done properly.
- Added description of the multidimensional unfolding procedures (up to 3D unfolding).
- Added plots showing the $\cos \phi_h$ and $\cos 2\phi_h$ moments extracted from the current fits as functions of z in separate P_T bins.
- Link to additional versions of the images in this analysis note has been updated to reference the current location where the latest (regular) updates are posted. Link to current page is: https://userweb.jlab.org/~richcap/Interactive_Webpage_SIDIS_richcap/Interactive_Unfolding_Page_Updated.html

VIII. APPENDIX

⁴³⁰ Code not referenced in other analysis notes.

```

double func(double x, double k, double b){
    return k * x + b;
}

struct line{
    double k;
    double b;
};

double isOutOfLines(double x, double y, line topLine, line botLine){
    return y > func(x, topLine.k, topLine.b) || y < func(x, botLine.k, botLine.b);
}

bool BadElementKnockOut(double Hx, double Hy, int sector, int cutLevel);

bool BadElementKnockOut(double Hx, double Hy, int sector, int cutLevel){

    double widthChange = 0;
    if (cutLevel == 0) widthChange = -1;
    if (cutLevel == 2) widthChange = 1;

    if (sector == 5) return 1;

    if (sector == 1){
        double k = tan(29.5*3.1415/180);
        double b = -92;
        return (isOutOfLines(Hx, Hy, {k, b + widthChange} , {k, b - widthChange - 2.4}) &&
               isOutOfLines(Hx, Hy, {k, b + widthChange - 9.1} , {k, b - widthChange - 9.1 - 2.4}) &&
               isOutOfLines(Hx, Hy, {k, b + widthChange - 127} , {k, b - widthChange - 127 - 2.4}) &&
               isOutOfLines(Hx, Hy, {k, b + widthChange - 127 - 8} , {k, b - widthChange - 127 - 8 - 2.4}));
    }
    if (sector == 2){
        double k = tan(30.4*3.1415/180);
        double b = 120.5;
        return (isOutOfLines(Hx, Hy, {k, b + widthChange} , {k, b - widthChange - 4.4}));}
    }
    if (sector == 3){
        return ((Hx - widthChange) > - 303 || (Hx + widthChange) < -310);
    }

    if (sector == 4){
        double k = tan(-29.6*3.1415/180);
        double b = -232.8;
        return (isOutOfLines(Hx, Hy, {k, b + widthChange} , {k, b - widthChange - 3.5}));}
    }

    if (sector == 6){
        double k = tan(-30.6*3.1415/180);
        double b = -185;
        return (isOutOfLines(Hx, Hy, {k, b + widthChange} , {k, b - widthChange - 2}) &&
               isOutOfLines(Hx, Hy, {k, b + widthChange - 8.3} , {k, b - widthChange - 8.3 - 2.2}));
    }

    return 0;
}

```

FIG. 76: Fiducial Cuts Developed by Valerii Klimenko. Hx and Hy are the coordinates of matched hit in the PCAL.

431

Old binning scheme used in prior versions analysis notes.

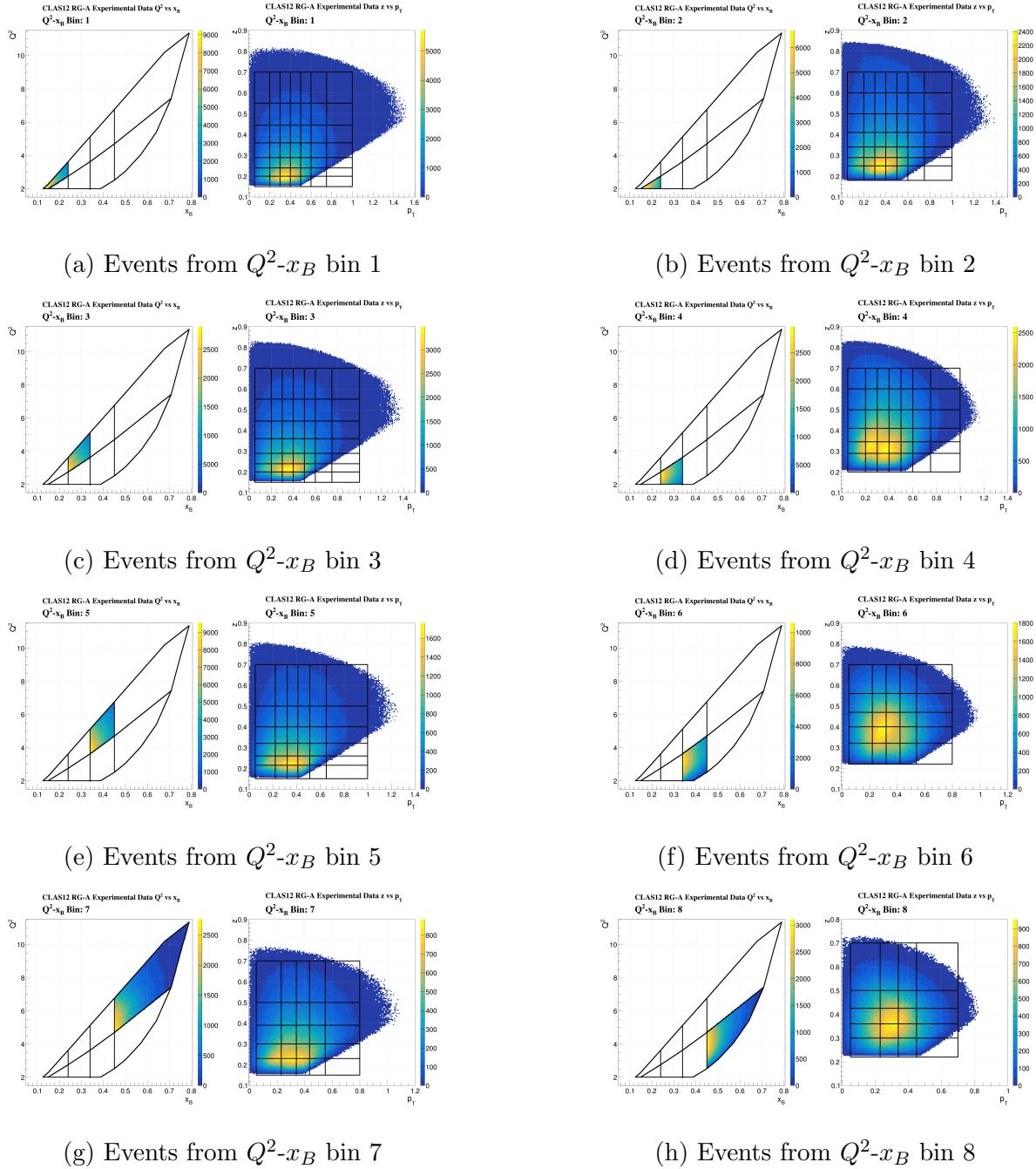


FIG. 77: 2D distributions of Q^2 , x_B , z , and P_T after all analysis cuts are applied to the RG-A data. Figure shows events in the old Q^2 - x_B binning scheme which was replaced as of Version v4 of this analysis note

-
- [1] A. Bacchetta, U. D'Alesio, M. Diehl, and C. A. Miller, "Single-spin asymmetries: The trento conventions," *Phys. Rev.*, vol. D70, p. 117504, 2004.
 - [2] N. Harrison, "Exploring the structure of the proton via semi-inclusive pion electroproduction." Thesis, University of Connecticut, available at https://www.jlab.org/Hall-B/general/thesis/Harrison_thesis.pdf, 2015.
 - [3] A. Bacchetta *et al.*, "Semi-inclusive deep inelastic scattering at small transverse momentum," *JHEP*, vol. 02, p. 093, 2007.
 - [4] T. B. Hayward, "Dihadron beam spin asymmetries on an unpolarized hydrogen target with CLAS12." Thesis, College of William & Mary, available at https://www.jlab.org/Hall-B/general/thesis/THayward_thesis.pdf, 2021.
 - [5] S. Diehl and K. Joo, "A multidimensional study of sidis π^+ beam spin asymmetry over a wide range of kinematics.." Internal Note., 2020.
 - [6] CLAS, "11 GeV polarized electrons on liquid hydrogen target to study proton structure, 3d imaging, and gluonic excitations, RG-A analysis overview and procedure." Internal Note, under review. Snapshot from August 2020: https://clas12-docdb.jlab.org/DocDB/0009/000949/001/RGA_Analysis_Overview_and_Procedures-08172020.pdf.
 - [7] R. A. Capobianco, "Clas12 momentum corrections." Available at https://clasweb.jlab.org/wiki/index.php/CLAS12_Momentum_Corrections.
 - [8] M. Ungaro *et al.*, "The CLAS12 Geant4 simulation," *Nucl. Instrum. Meth. A*, vol. 959, p. 163422, 2020.
 - [9] H. Avakian, "clasdis." <https://github.com/JeffersonLab/clasdis>, 2020.
 - [10] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.
 - [11] L. Mankiewicz, A. Schafer, and M. Veltri, "Pepsi: A monte carlo generator for polarized leptoproduction," *Comput. Phys. Commun.*, vol. 71, pp. 305–318, 1992.
 - [12] G. Ingelman, A. Edin, and J. Rathsman, "LEPTO 6.5: A Monte Carlo generator for deep inelastic lepton - nucleon scattering," *Comput.Phys.Commun.*, vol. 101, pp. 108–134, 1997.
 - [13] A. Höcker and V. Kartvelishvili, "SVD approach to data unfolding," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 372, pp. 469–481, apr 1996.
 - [14] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth. A*, vol. 362, pp. 487–498, 1995.