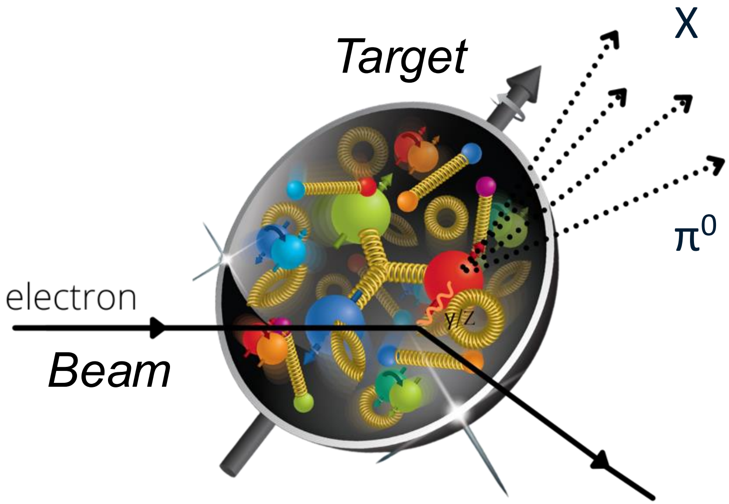
MARCH 11, 2026

NEUTRAL PION SIDIS MULTIPLICITY STUDIES AT CLAS12

VALERII KLIMENKO
Postdoctoral Appointee

MOTIVATION

- Neutral pion Semi-Inclusive Deep Inelastic Scattering Multiplicities measure the probability of producing a hadron (π^0) per DIS event
- Multiplicity measurements uniquely link hadronization with the proton's internal transverse motion, advancing both FF determinations and our 3D understanding of proton structure
- The **z** - dependence constrains **fragmentation functions (FFs):** how colored partons turn into color-neutral hadrons
- The $\boldsymbol{P_T^2}$ **-** dependence directly probes the unpolarized **TMD PDFs** revealing the proton's intrinsic transverse momentum dynamics
- It also allows to study isospin invariance as the neutral pion fragmentation function is thought to be dependent on the charged pion fragmentation functions

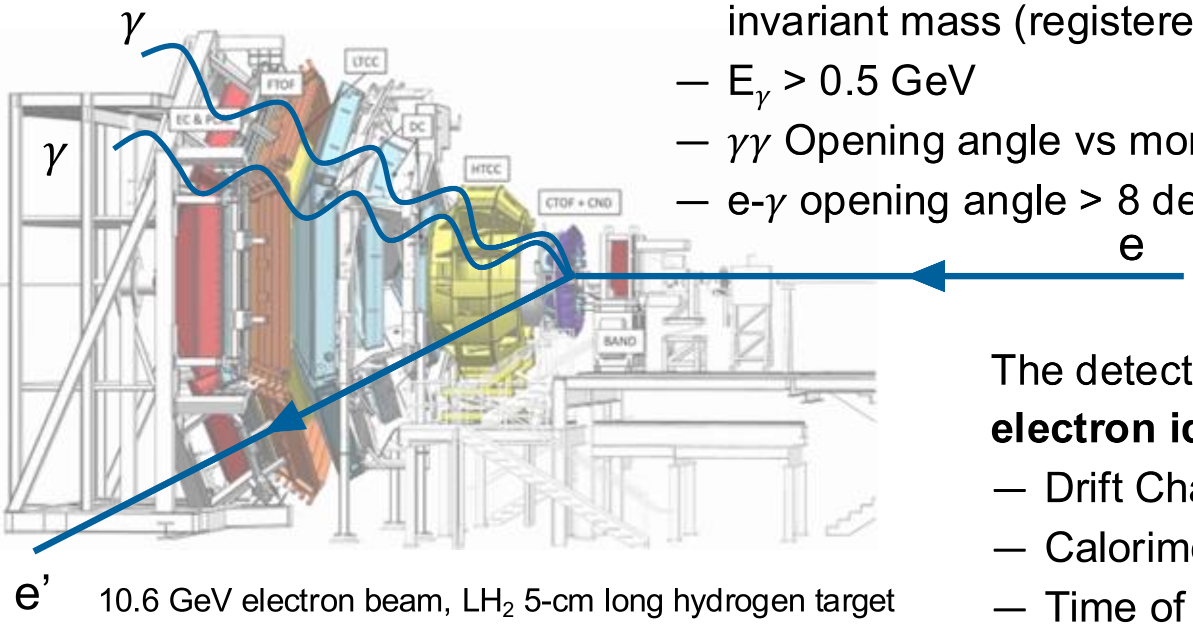


$$\frac{d^2 M_h}{dz\, dP_T^2}(x_B, Q^2, z, P_T^2) = \frac{\int_0^{2\pi} d\phi_h \frac{d\sigma^{SIDIS}}{dx\, dQ^2\, dz\, dP_T^2\, d\phi_h}}{\frac{d\sigma^{DIS}}{dx\, dQ^2}}$$

CLAS12

Selection of the ep → eπ^{0}X reaction with Forward Detector

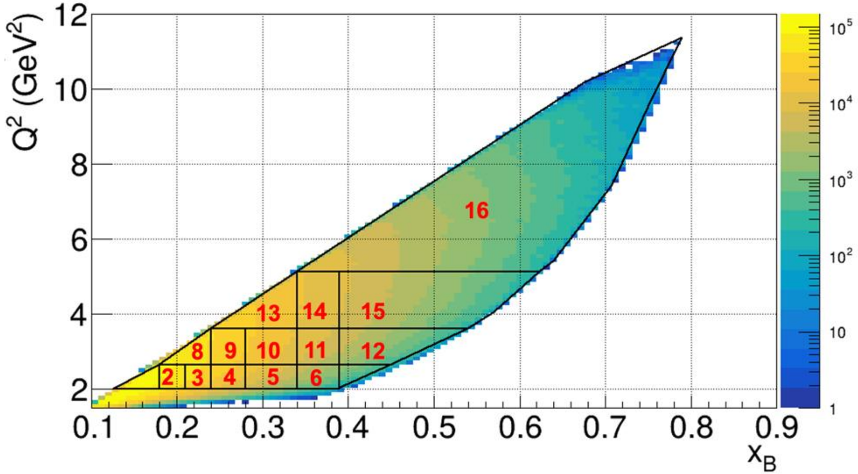
- **π^0 candidates** are reconstructed from photon pairs invariant mass (registered in Calorimeters)
- E_γ > 0.5 GeV
- $\gamma\gamma$ Opening angle vs momentum cut
- e-γ opening angle > 8 deg

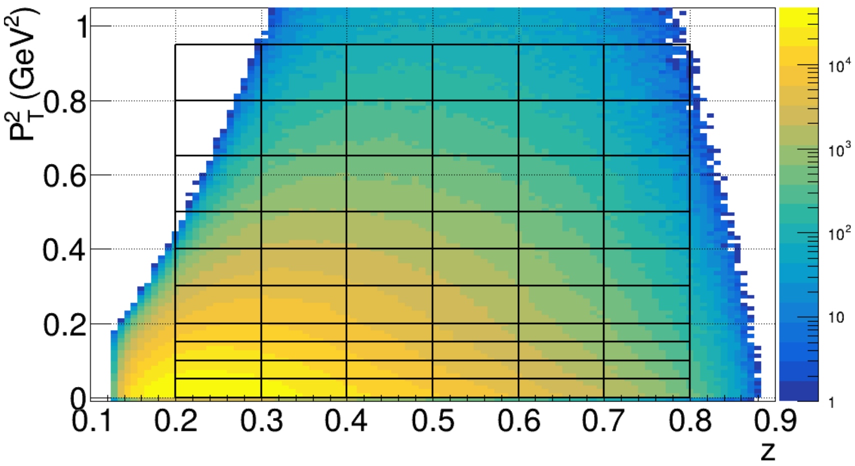


10.6 GeV electron beam, LH_2 5-cm long hydrogen target

The detector subsystems used for **electron identification**:
- Drift Chambers (track)
- Calorimeters (EM shower)
- Time of Flight (PID)
- Cherenkov counters (PID)

KINEMATICS AND BINNING





— Large CLAS12 acceptance together with large statistics allow for multidimensional binning

— Semi-Inclusive DIS events selected with following kinematic cuts:

- $W > 2$ GeV
- $Q^2 > 2$ GeV^2
- $y < 0.75$

- $x_F > 0$ [$x_F = 2P_{h,L} / \sqrt{s}$] : current fragmentation region
- $M_x > 1.5$ GeV

Number of bins:

- X-Q^2 - 16
- Z – 6
- P_T^2 - 9
- Φ_h - 8

MULTIPLICITY

What do we extract?

Corrected SIDIS yield

$$N^{h,\,\mathrm{corr}}_{ijmn} = C^{h}_{\mathrm{rad}\,ijmn} \sum_{k} \frac{N^{h,\mathrm{raw}}_{ijmn} - N^{h,\mathrm{bkg}}_{ijmn}}{A^{h}_{ijmnk}}$$

Radiative Corrections

Raw Event Yield

Background

Acceptance and Efficiency Correction

k is φ bin

$$M^{h}_{ijmn} = \frac{1}{\Delta z_m \, \Delta P^2_{T,n}} \frac{N^{h,\,\mathrm{corr}}_{ijmn}}{N^{\mathrm{DIS,\,corr}}_{ij}}, \quad (i,j,m,n) \equiv (x_B\text{-bin},\, Q^2\text{-bin},\, z\text{-bin},\, P_T^2\text{-bin})$$

z-P_T^2 bin size

Corrected DIS yield

$$N^{\mathrm{DIS,\,corr}}_{ij} = C^{\mathrm{DIS}}_{\mathrm{rad}\;ij} \frac{N^{e,\mathrm{raw}}_{ij}}{A^{e}_{ij}}$$

Radiative Corrections

Raw Event Yield

Acceptance and Efficiency Correction

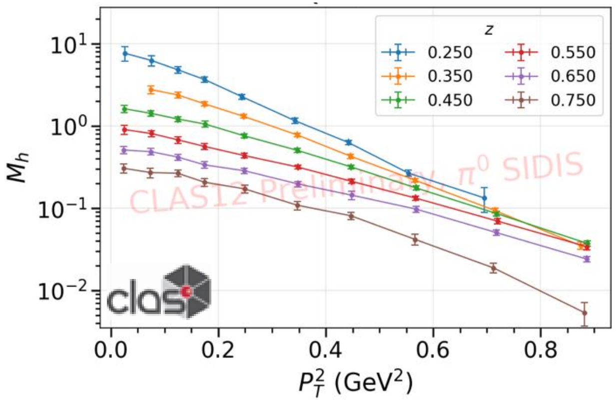
- Electron-side efficiencies partially cancel out in the ratio

RESULTS $M_H(Q^2, X_B, Z, P_T^2)$

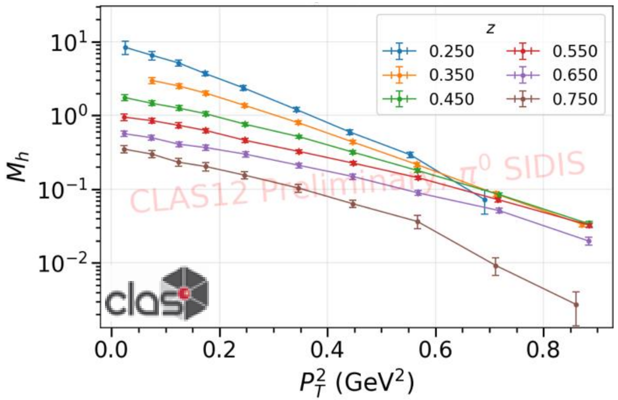
RESULTS

$M_h(p_T^2)$ for selected x-Q^2 bins

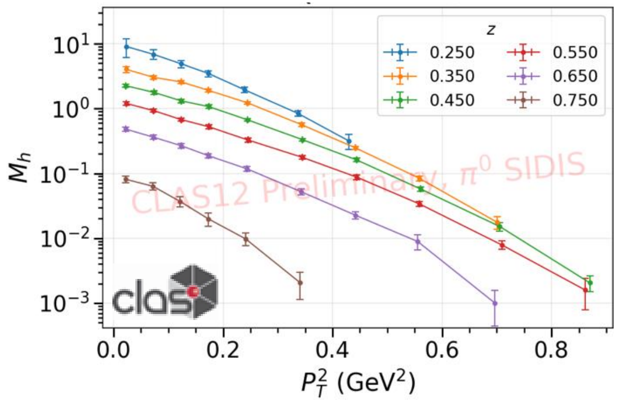
0.18 < x < 0.21, 2.0 < Q^2 < 2.66 GeV^2



0.24 < x < 0.28, 2.66 < Q^2 < 3.63 GeV^2



0.39 < x < 0.62, 3.63 < Q^2 < 5.12 GeV^2



- LO, gaussian approximation:

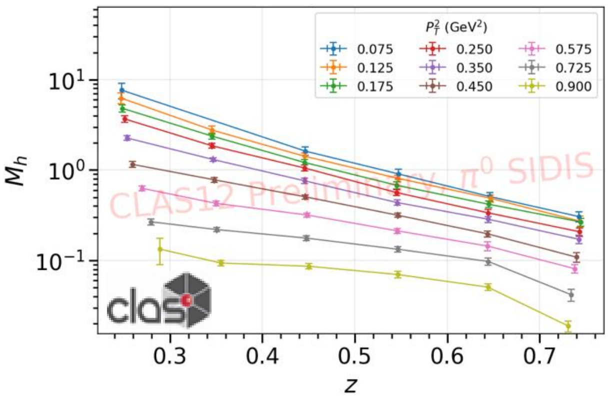
$$\frac{d^2M^h}{dz\,dP_T^2} = \left(\frac{d^4\sigma}{dx\,dQ^2\,dz\,dP_T^2}\right) \Big/ \left(\frac{d^2\sigma^{DIS}}{dx\,dQ^2}\right) = \frac{N}{\pi\langle P_T^2\rangle}\exp\left(-\frac{P_T^2}{\langle P_T^2\rangle}\right)$$

- The behavior is consistent with an LO Gaussian model, except for a high-P_T^2 downturn attributable to the phase-space restriction $M_x > 1.5$ GeV
- Only statistical uncertainties are shown

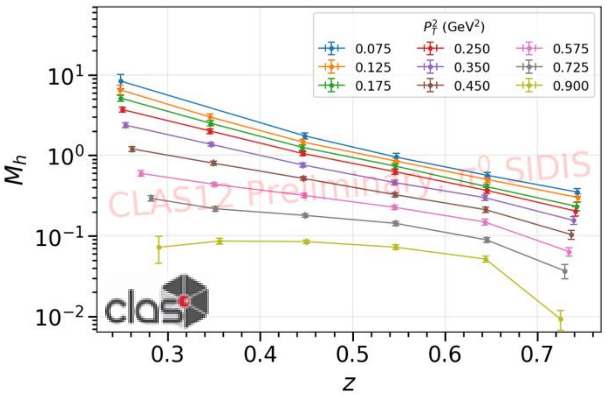
RESULTS

$M_h(z)$ for selected x-Q^2 bins

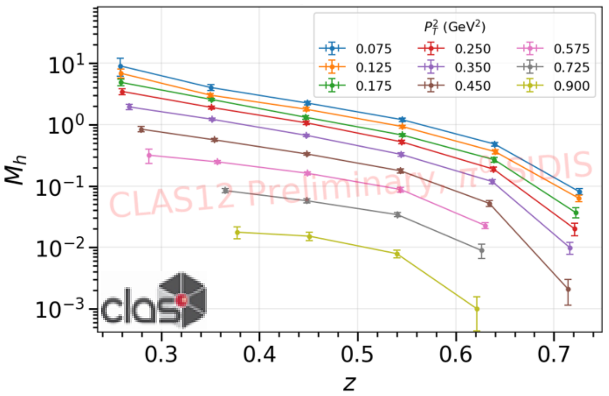
0.18 < x < 0.21, 2.0 < Q^2 < 2.66 GeV^2



0.24 < x < 0.28, 2.66 < Q^2 < 3.63 GeV^2



0.39 < x < 0.62, 3.63 < Q^2 < 5.12 GeV^2



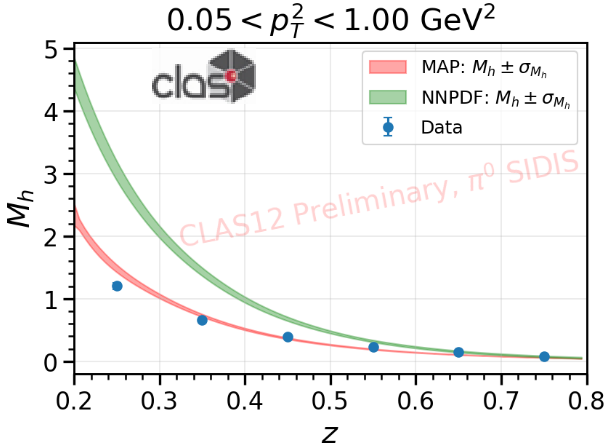
- ln($M_h(z)$) is approximately linear, but drops at large z due to the $M_x > 1.5$ phase-space cut
- Only statistical uncertainties are shown

RESULTS

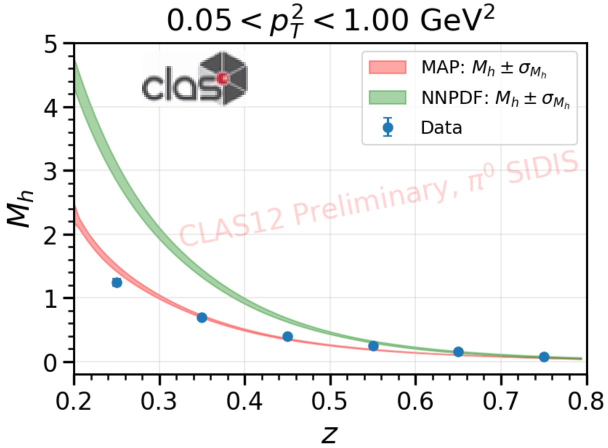
P_T^2 integrated multiplicity:

$$M_h = \frac{\sum_q e_q^2 f_q(x) D_q(z)}{\sum_q e_q^2 f_q(x)}$$

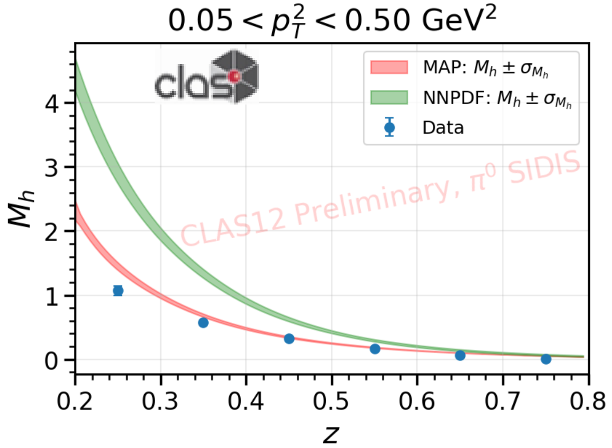
$M_h(z)$ integrated over P_T^2 for selected x-Q^2 bins



0.18 < x < 0.21, 2.0 < Q^2 < 2.66 GeV^2



0.24 < x < 0.28, 2.66 < Q^2 < 3.63 GeV^2



0.39 < x < 0.62, 3.63 < Q^2 < 5.12 GeV^2

- M_h was integrated over the P_T^2 range that preserves full z coverage
- The theoretical curves were computed with CT10nlo PDFs using two fragmentation-function sets:
  - **NNFF**, fitted only to single-inclusive hadron production in e^+e^- annihilation
  - **MAPFF**, which additionally includes charged-hadron multiplicity data from HERMES and COMPASS
- After integrating over P_T^2, **multiplicities** vs. z follow the trend of LO predictions using MAPFF FF and CT10nlo PDFs

MATRIX DECONVOLUTION

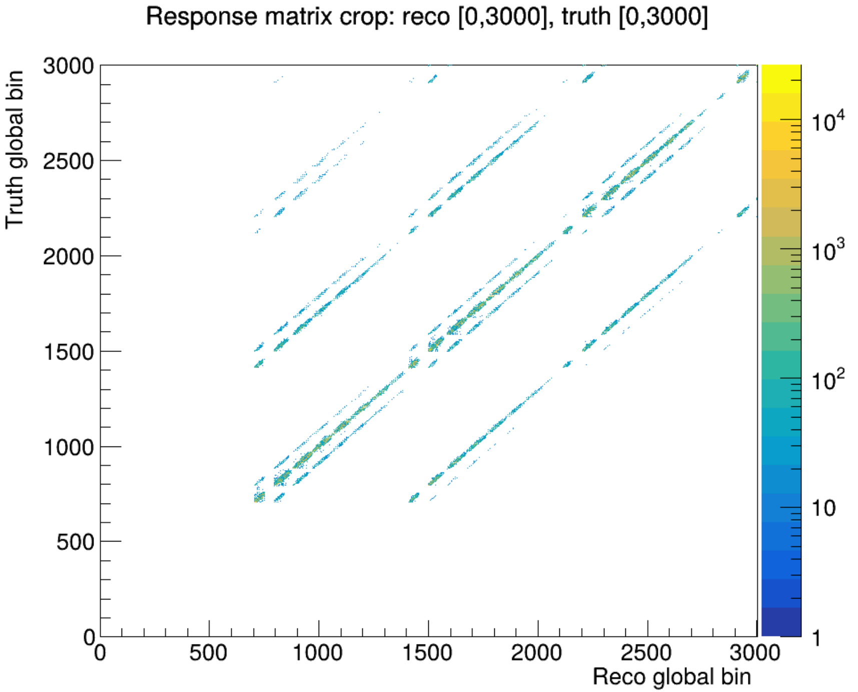
ACCEPTANCE AND EFFICIENCY CORRECTION

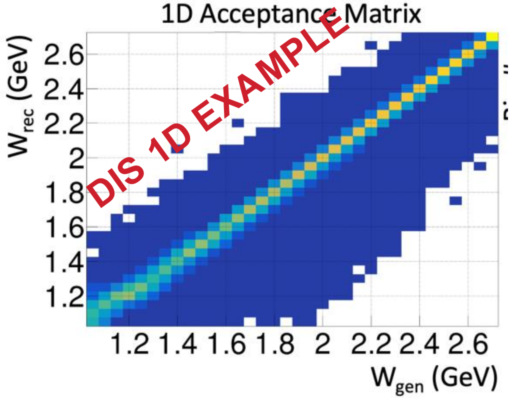
Bayesian Deconvolution

- Based on Response Matrix

$$\frac{P(\text{reconstructed in bin } i \mid \text{generated in bin } j)}{\text{total number of generated events}}$$

- The algorithm starts from an initial prior (MC) and iteratively updates the unfolded spectrum using Bayes' theorem.
- The implementation provided by the RooUnfold package

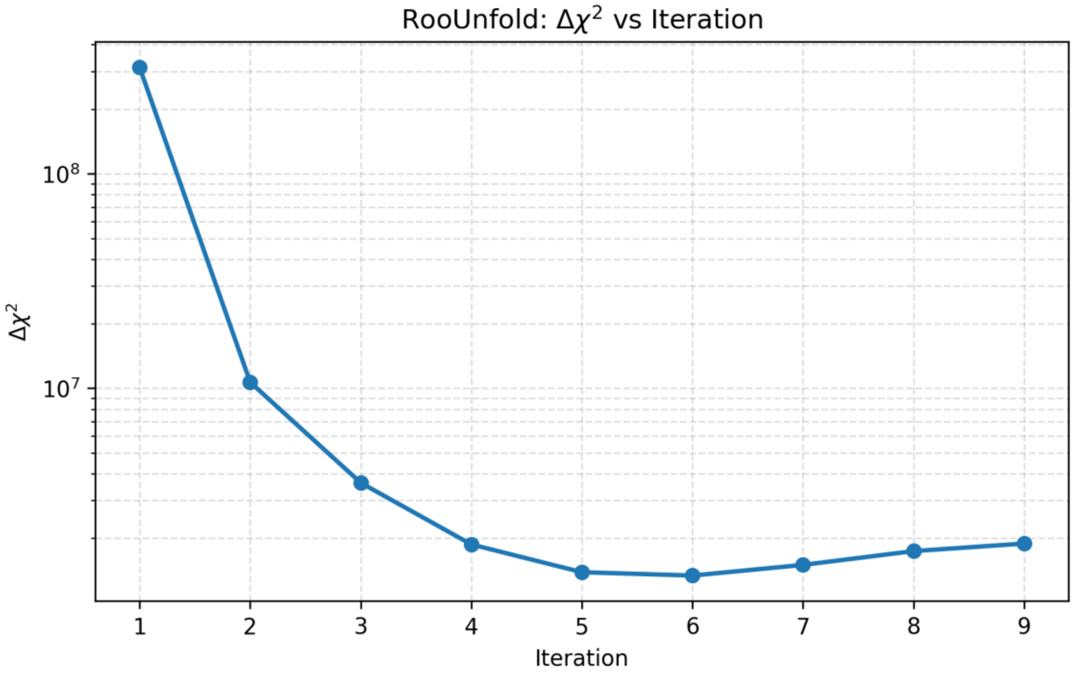




ACCEPTANCE AND EFFICIENCY CORRECTION

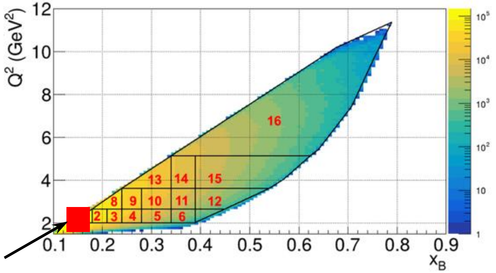
Bayesian Deconvolution

- The number of iterations acts as an effective regularization parameter
- Four Iterations was used in the nominal analysis
- Results with three and five iterations will be used to estimate the systematic uncertainty

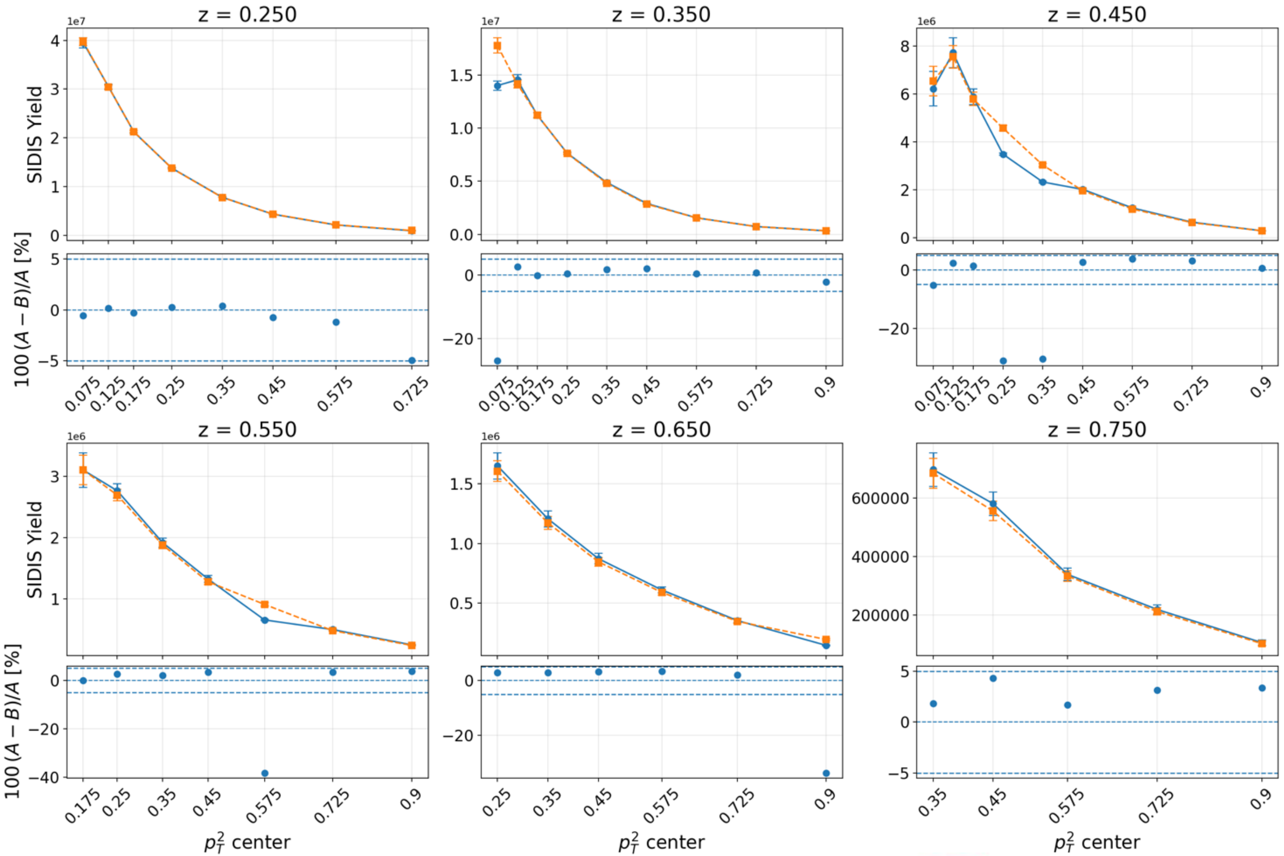


ACCEPTANCE AND EFFICIENCY CORRECTION

- A has been extracted from Φ fits
- SIDIS Yield **Bin-By-Bin** vs. **Bayesian** deconvolution
- Outliers are unstable fits
- ~5% effect
- Two approaches are consistent



0.126 < x < 0.18, 2.0 < Q^2 < 2.66 GeV^2



UNFOLDING VALIDATION

CLOSURE TESTS

- Demonstrate that the unfolding procedure is unbiased, numerically stable, and robust against reasonable model variations

1. **Same-sample sanity closure** (Build and test the unfolding on the same MC sample)

   Goal: Implementation check only

2. **Independent-sample closure** (Split MC into two independent samples)

   Goal: Check of Intrinsic unfolding bias when response and pseudo-data are independent

3. **Stability check with repeated splits** (Repeat the independent A/B splits 10–20 times)

   Goal: Check of bias of unfolding in high-dimensional space

4. **Shape-mismatch closure** (Apply smooth weights to ALL events in B based on truth kinematics to match better unfolded data)

   Goal: Check of bias against realistic model mismatches

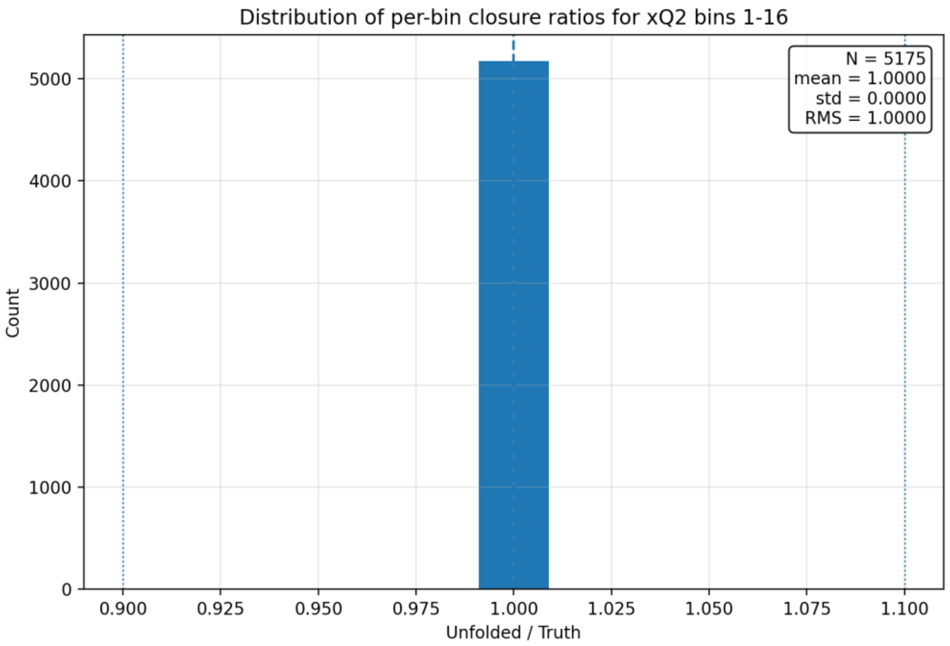
5. **Refolding consistency test** (Forward-Folding)

   Goal: Verify numerical self-consistency and stability

CLOSURE TEST #1

Purpose: This is a sanity / implementation check

- Synthetic pseudo-data are built from the same MC
- The pseudo-data are unfolded using the migration matrix from the same MC
- The unfolded result is then compared directly to the MC truth
- Closed exactly



FORWARD FOLDING

FORWARD FOLDING

Goal: Verify numerical self-consistency and stability

1. Unfold the data
2. Forward folding of the data with the same response matrix, Poisson fluctuated
3. Unfold the Forward folded data again and compare with (1)

Procedure

RG-A Data

Unfolded spectra

1. $X_1 R^{-1} \rightarrow Y_1$
2. $X_2 \leftarrow R\, Y_1$
3. $X_2 R^{-1} \rightarrow Y_2$

Forward Folded Data

Twice Unfolded spectra

The next slides show comparisons of

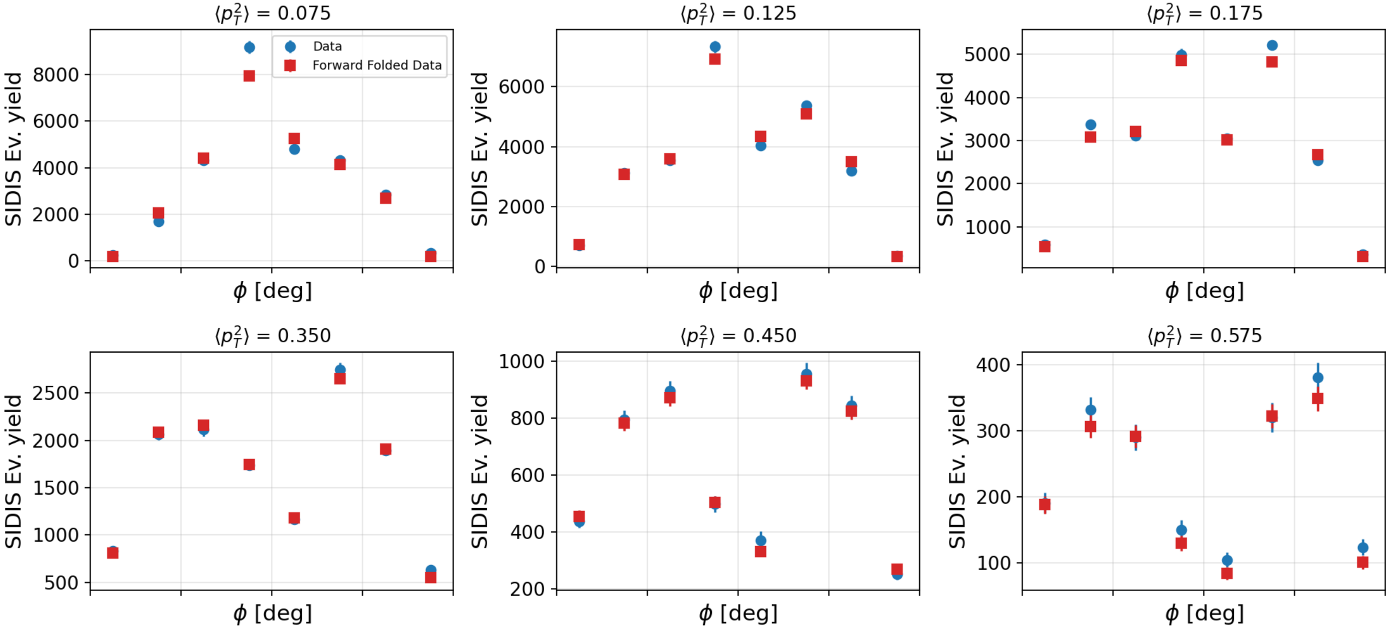
i. X_1 and X_2 (slide 16)

ii. Y_1 and Y_2 (slide 17)

iii. **Pull** of (ii) (slide 18)

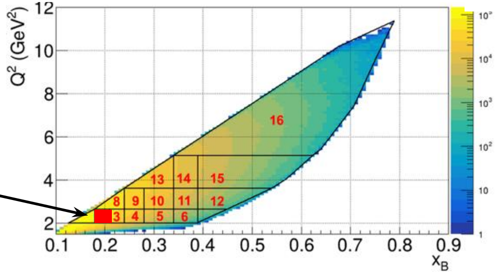
FORWARD FOLDING

i. X_1 and X_2



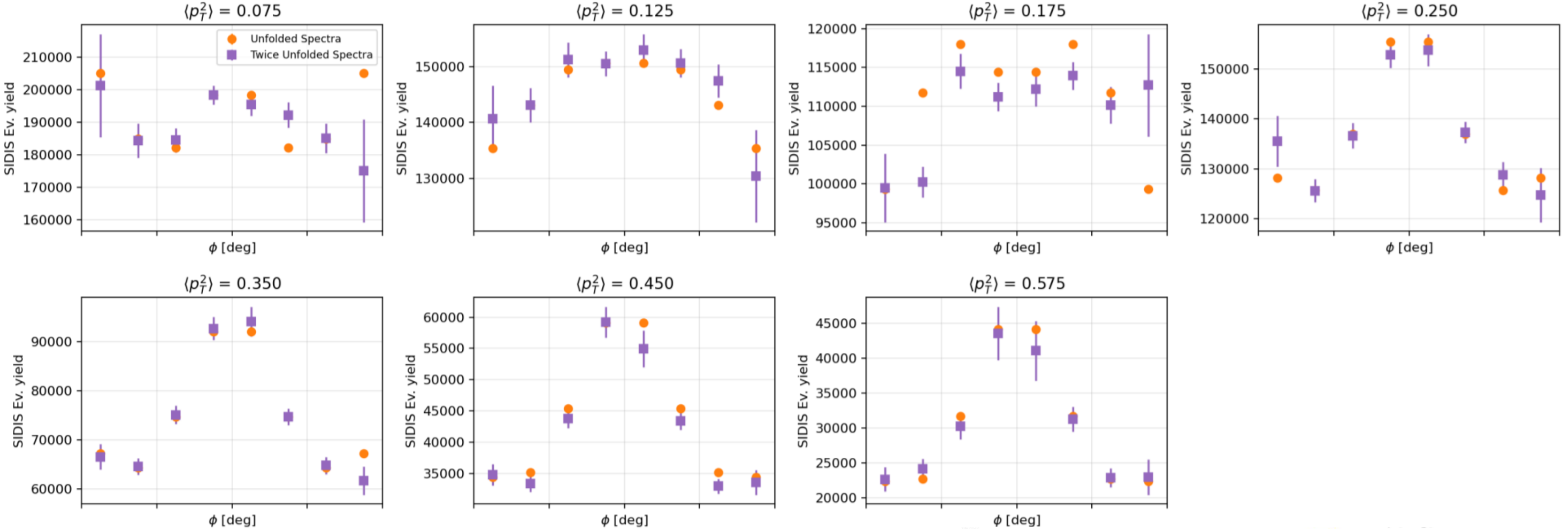
Data (X_1) vs. Forward folded data (X_2)

$0.18 < x < 0.21$, $2.0 < Q^2 < 2.66$ GeV^2



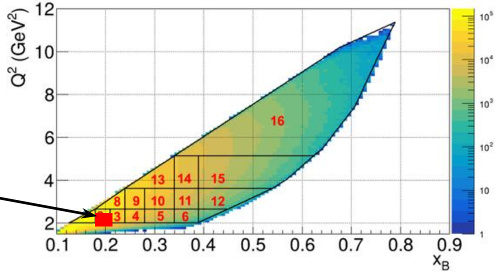
FORWARD FOLDING

ii. Y_1 and Y_2



Unfolded Data (X_1) vs. **Twice Unfolded data (X_2)**

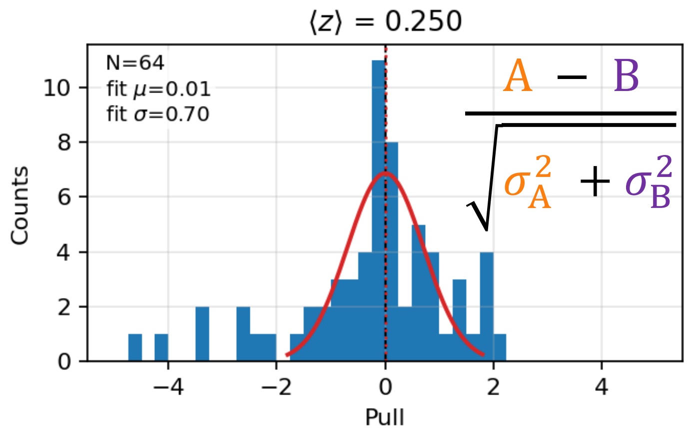
$0.18 < x < 0.21$, $2.0 < Q^2 < 2.66$ GeV^2

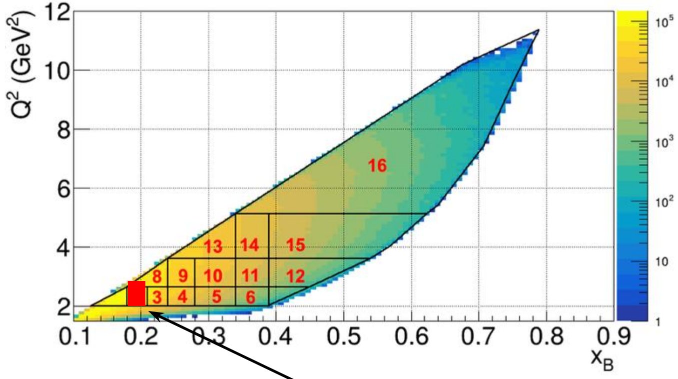


ii. Y_1 and Y_2

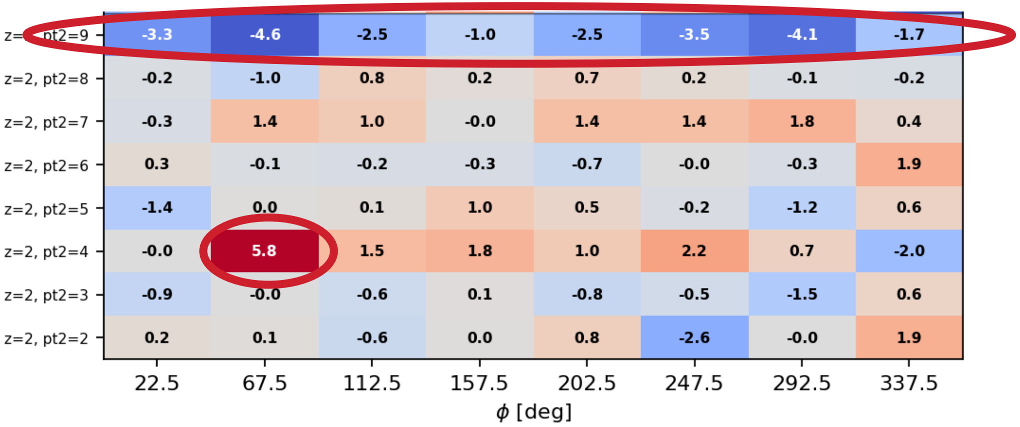
FORWARD FOLDING

- Pull: **Unfolded Data** vs. **Twice Unfolded Data**
- Integrated over ϕ and P_T^2
- Some deviations may arise from ϕ fits and bad bins
- Overall agreement for both samples is good
- $0.18 < x < 0.21$, $2.0 < Q^2 < 2.66$ GeV^2





$0.18 < x < 0.21$, $2.0 < Q^2 < 2.66$ GeV^2



CROSS CHECK WITH LEGACY CODE

LEGACY CODE CROSS CHECK

The Legacy analysis was updated as follows:

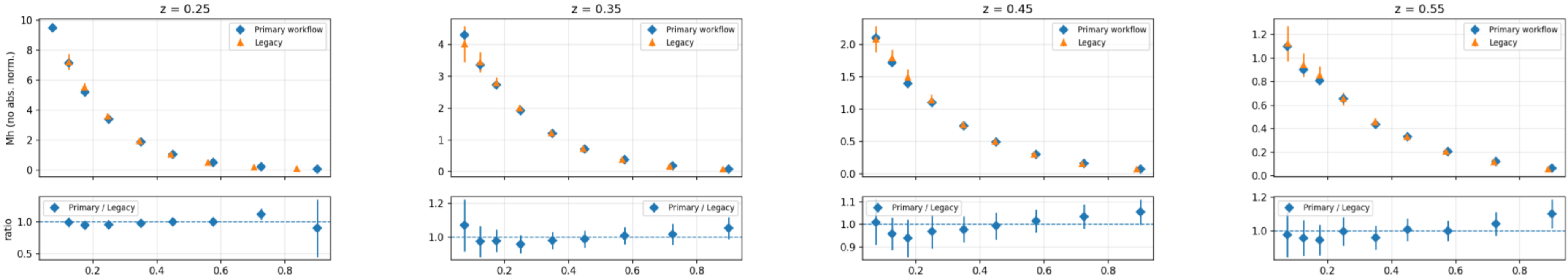
1. Switched to pass-2 data and the corresponding simulation sample
2. Refined and updated the event selection for pass-2
3. Binning scheme Update

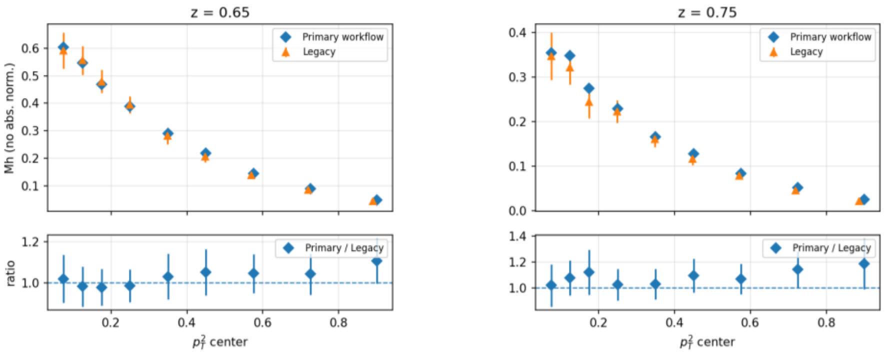
Updated analysis was simplified:

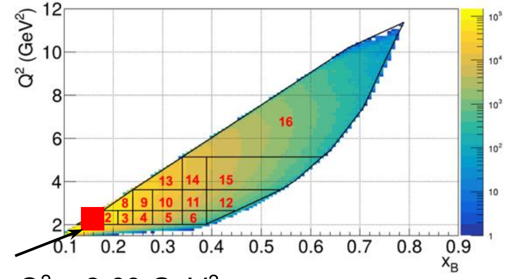
1. Matrix deconvolution -> bin-by-bin
2. Each ϕ bin was fitted and then summed to form a single bin
3. Acceptance handling was simplified. A well-reconstructed electron for both reconstructed and generated spectra was required

LEGACY CODE CROSS CHECK

- Mh (not normalized) **Legacy** vs. **Primary**
- Workflows are consistent within uncertainties







$0.126 < x < 0.18$, $2.0 < Q^2 < 2.66$ GeV^2

CONCLUSIONS AND OUTLOOK

Updated:

- Matrix deconvolution was incorporated
- Cross check with legacy code was performed
- Several closure test were conducted

In progress:

- Closure tests
- Estimate systematic uncertainties
- π^0 Efficiency studies
- VM contamination studies

Results:

- Preliminary Neutral pion SIDIS Multiplicities (x_B, Q^2, z, P_T^2) at CLAS12 are available
- After integrating over P_T^2 , multiplicities vs. z follow the trend of LO predictions using MAPFF fragmentation functions and CT10nlo PDFs
- The data will provide new constraints on unpolarized TMD PDFs and fragmentation functions (FFs) and enabling tests of isospin symmetry in FFs

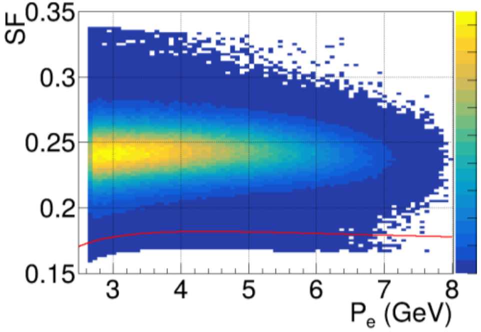
THANK YOU!

BACK UP SLIDES

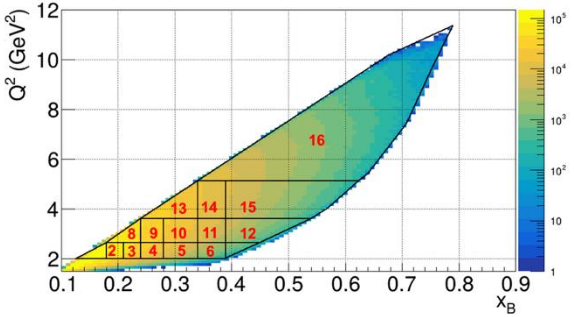
PID AND EVENT SELECTION

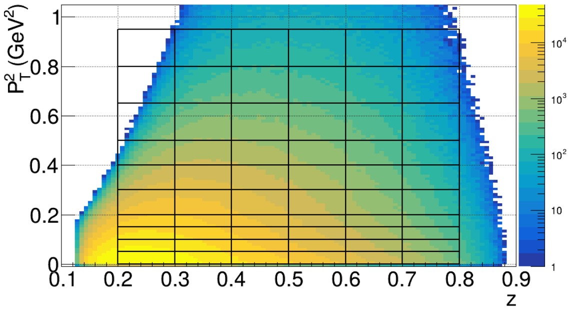
Starts with CLAS12 reconstruction algorithm

- Electron
  - 2 < p_e < 8 GeV
  - $y < 0.75$
  - Z-Vertex cut
  - DC and PCAL fiducial
  - Sampling Fraction



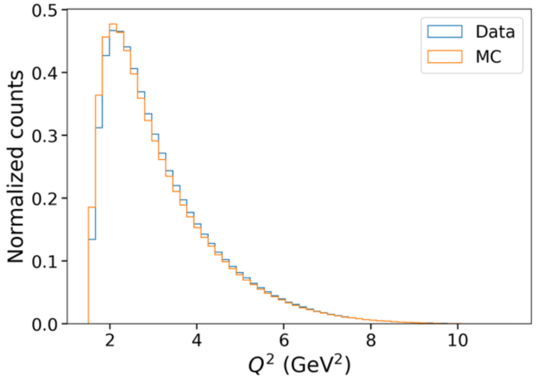
- Photon
  - $E_\gamma > 0.5$ GeV
  - e-γ opening angle > 8 deg
  - $0.9 < \beta < 1.1$
- π^0
  - Candidates are reconstructed from photon pairs
  - $x_F > 0$ [$x_F = 2P_{h,L}/\sqrt{s}$] : current fragmentation region
  - $M_x > 1.5$ GeV
  - $\alpha_{\gamma\gamma} > 6 \cdot \text{Exp}(1 - p_\pi) + 0.5$ deg



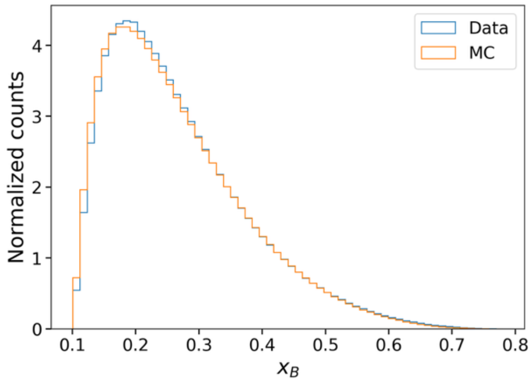


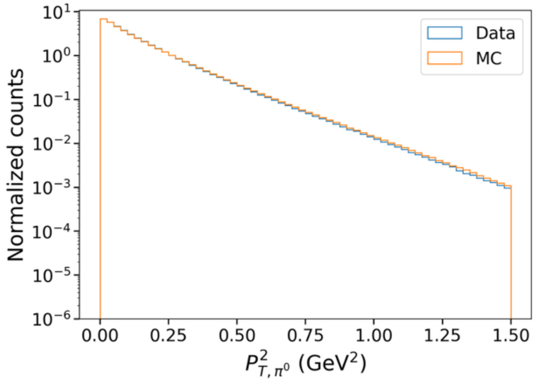
MONTE CARLO

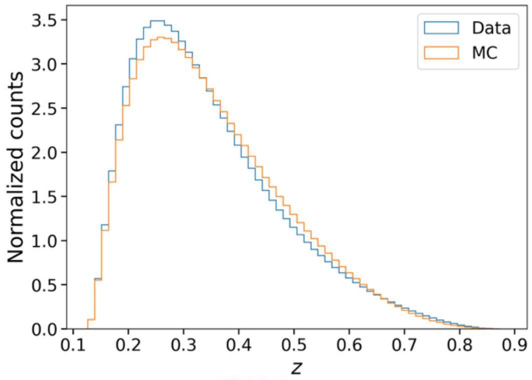
Event Generator used for the Acceptance



- CLASDIS-EG built on LEPTO
- Lund string model as in PYTHIA/JETSET
- LO electroweak DIS (arbitrary lepton polarization); O(α) matrix elements for boson–gluon fusion and gluon radiation; higher-order QCD via parton showers
- QED radiation is not included in the current MC sample
- Hadronization parameters tuned to CLAS12 semi-inclusive data (Q^2 > 1 GeV^2, W > 2 GeV)





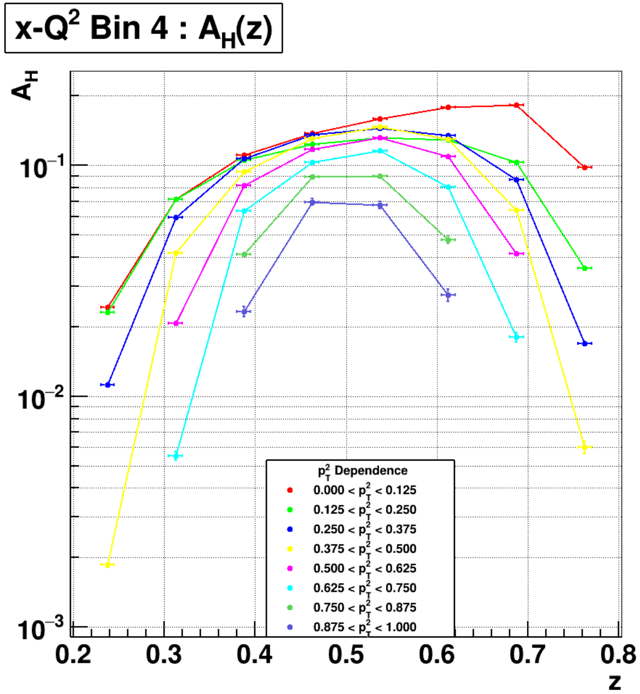


histograms normalized to unit area

ACCEPTANCE AND EFFICIENCY CORRECTION

Deconvolution

- Method: bin-by-bin (efficiency) 5D(x_B, Q^2, z, P_T^2, φ_h) correction
- Per-bin efficiency: $\varepsilon_i = \frac{N_{rec,i}}{N_{gen,i}}$ (reconstructed / generated)
- Deconvolved yield: $y_i = \varepsilon_i \cdot x_i$, where x_i is the measured yield
- Limitation: does not track bin migrations (i -> j); performance depends on generator and detector-simulation accuracy

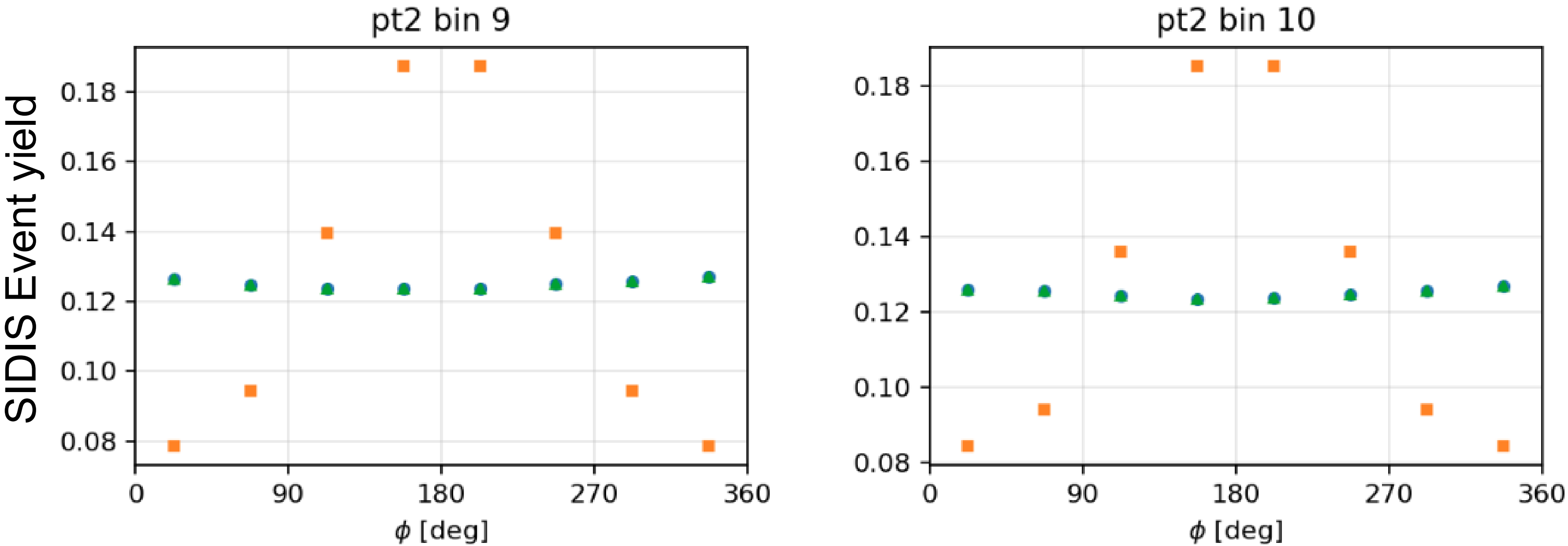


FORWARD FOLDING

i. Y_1 and Gen

- Generated phi spectra (Truth) **Gen clasdis**
- Unfolded result (Y_1) that is used to compute new prediction

1. $\mathbf{MC::}\boldsymbol{Gen} \rightarrow \mathrm{R} \rightarrow \boldsymbol{MC::Rec}$
2. $\mathbf{Y_1} \rightarrow \mathrm{R} \rightarrow \mathbf{X_2}$



Unit area normalization