ATLAS Note

ANA-STDM-2020-07-INT1

21st February 2023

Draft version 0.4

Measurement of angular coefficients in $W \to \ell\nu$ events in low-μ pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Ludovica Aperio Bella[b], Aaron Armbruster[c], Alexander Bachiu[a], Stefano Camarda[c], Nicolo De Groot[d], Ruth Magdalena Jacobs[b], Daniil Ponomarenko[d], Manuella Vincter[a]

[a]*Carleton University,* [b]*DESY,* [c]*CERN,* [d]*Radboud University Nijmegen*

This note presents a measurement of angular coefficients and differential cross section as functions of transverse momentum and rapidity of W bosons. The measurement is performed in the $W \to e\nu$ and $W \to \mu\nu$ decay channels. The analysis uses the low pileup dataset collected with the ATLAS detector at 13 TeV in Run 2. The measurement is performed in the full solid angle and in the region of the W boson, defined as $|y_W| < 3.6$.

ChangeLog for version 0.3 → 0.4 in CDS (XX/02/2023)

- Answered to open CDS comments from Lorenzo.
- Added yields table to the Section 6.
- Added sliced 2D control plots to the Appendix F to show Data/Model agreement on the reco level.
- Updated for MJ estimation:
  - Updated MJ shape estimation formula (Eq. 49). Updated control plots and MJ estimation plots with new results obtained using new MJ shape estimation formula. Small changes in the MJ yields and unc. comes from the new *ROOT* version (6.18/04 → 6.28/00), in particular how the fit is performed for anti-isolated slices in the FR.
  - Implemented an improved acceptance correction method by usage of the coarse bins instead of inclusive SR (Section 5.4).
  - Added MJ acc. correction functions (Appendices E.5.2–E.5.5 and Appendices E.5.6–E.5.9) and MJ systematics breakdown plots (Appendix E.7).

ChangeLog for version 0.0 → 0.3 in CDS (30/09/2022)

- Answers to open CDS comments from Camilla and Frank.
- Polishing and restructuring text.

Contents

1 Introduction

In this note we present the measurement of the polar and azimuthal distributions of the $\ell\nu$ pairs produced from W decays in the centre-of-mass frame of the W boson using low-μ data collected from pp collisions at $\sqrt{s}$ = 13 TeV in the ATLAS detector at the Large Hadron Collider.

A detailed understanding of the Drell-Yan production mechanism for gauge bosons $p_1 + p_2 \rightarrow V + X$, where p_1 and p_2 are the incoming partons and $V = \{W, Z\}$, can be obtained by measuring the angular distributions of the leptonic decay products $\ell^{\pm} + \nu/\ell^{\mp}$ of a gauge boson. This is done by defining an event plane with beam and the gauge boson momentum direction as a reference frame, which provides detailed information of the lepton-hadronic correlation effects using the leptons produced by the decay of the gauge bosons. These correlations are described using nine hadronic "structure" functions and can be calculated using perturbative QCD for the parton model. A more detailed description of the theoretical formalism can be found in [1–4].

Without lepton selection requirements the angular distributions are determined by the vector boson polarization. However, the angular distributions become dominated by selection requirements when they are applied. These polarization effects can be recovered by comparing data to the ratio of MC distributions to those obtained using isotropic gauge boson decays. With enough data statistics a two-dimensional fit on the angular distributions can be used to extract the eight coefficients that describe them as discussed in Section 1.1.

The most recent measurement of these angular coefficients was performed by the LHCb Collaboration with $\sqrt{s} = 13$ TeV data for $Z \rightarrow \mu\mu$ events and extracted four of the eight coefficients as a function of p_{T}^{Z} [5]. All eight coefficients were measured as a function of p_{T}^{Z} and y^{Z} by the ATLAS Collaboration with $\sqrt{s} = 8$ TeV data for the Z boson using $Z \rightarrow \ell\ell$ where $\ell = \{e, \mu\}$ events [6] with the use one coefficient they were able to extract the weak mixing angle, $\sin^2\theta_W$ [7]. Other measurements of the Z boson coefficients have also been done previously by the CMS Collaboration who extracted five of the eight coefficients as a function of p_{T}^{Z} at $\sqrt{s} = 8$ TeV with just $Z \rightarrow \mu\mu$ events [8] and four of the eight coefficients as a function of p_{T}^{Z} by the CDF Collaboration at $\sqrt{s} = 1.96$ TeV with $Z \rightarrow ee$ events [9].

Only two of these angular coefficients for the W boson as a function of p_{T}^{W} have been measured previously by CDF Collaboration at $\sqrt{s} = 1.96$ TeV with $W^{\pm} \rightarrow \ell^{\pm}\nu$ events [10]. However, other measurements of the W boson polarization have measured previously by the ATLAS Collaboration with $\sqrt{s} = 7$ TeV data [11] and by the CMS Collaboration with $\sqrt{s} = 7, 13$ TeV data [12, 13]. These polarization fraction measurements are related to two coefficients and can be used to probe their values.

The angular coefficients for the W boson are important for modeling QCD production and have impact on other measurements like the W boson mass. Currently the modeling of these coefficients is done by extrapolating from the coefficients of the Z boson which increases the uncertainty on m^W measurements. The ATLAS Collaboration measurement of m^W at $\sqrt{s} = 7$ TeV reported an uncertainty from the modeling of the angular coefficients of $\delta p_{\mathrm{T}}^{\ell} = 5.8$ MeV and $\delta m_T = 5.3$ MeV on these distributions used to extract m^W [14]. The LHCb Collaboration reported a systematic uncertainty of $\delta m^W = 10$ MeV from the modeling on their measurement of m^W at $\sqrt{s} = 13$ TeV [15]. One of the criticisms of the most recent measurement of m^W by the CDF Collaboration at $\sqrt{s} = 1.96$ TeV is that they didn't include an uncertainty relating to the modeling of these angular coefficients [16].

1.1 Differential Cross-section

The differential cross-section, at all orders of QCD for non-zero transverse momentum, for the process $p_1 + p_2 \rightarrow W + X \rightarrow \ell\nu + X$ has a general form as seen in Equation 1. Where p_{T}^{W}, m^W, and y^W are the transverse momentum, mass, and rapidity in the laboratory frame and θ and ϕ are the polar and azimuthal angles of the leptons in the W boson rest frame. Rapidity is calculated using $y = \frac{1}{2}\ln\left[\left(E + p_z\right)/\left(E - p_z\right)\right]$ where E is the energy of the W boson and p_z is the momentum along the z-axis in the coordinate system.

Equation 1 is a result of the W boson having spin 1, and the lepton-neutrino pair having spin 1/2 and is further discussed in [1, 3].

$$\begin{aligned}\frac{d\sigma}{dp_{\mathrm{T}}^{W}\,dm^{W}\,dy^{W}\,d\cos\theta d\phi} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dp_{\mathrm{T}}^{W}\,dm^{W}\,dy^{W}}\Big\{&\left(1+\cos^2\theta\right)+\frac{1}{2}A_0\left(1-3\cos^2\theta\right)\\ &+A_1\sin 2\theta\cos\phi+\frac{1}{2}A_2\sin^2\theta\cos 2\phi\\ &+A_3\sin\theta\cos\phi+A_4\cos\theta+A_5\sin^2\theta\sin 2\phi\\ &+A_6\sin 2\theta\sin\phi+A_7\sin\theta\sin\phi\Big\}.\end{aligned} \tag{1}$$

The differential unpolarized cross section, denoted σ^{U+L}, has been factored out resulting in the coefficients $A_0 - A_7$ which are ratios of the helicity dependent cross-sections relative to the unpolarized cross-section. These coefficients are dependent on the W boson's kinematics and vanish when p_{T}^{W} is zero, except for A_4 which is present for all orders of QCD and is what generates the forward-backward asymmetry in $\cos\theta$. Integrating Equation 1 over the full lepton kinematic phase space,$\cos\theta$ and ϕ, results in the A_i coefficients vanishing which results in the triple differential unpolarized cross-section:

$$\int_{1}^{1} d\cos\theta\int_{0}^{2\pi} d\phi\frac{d\sigma}{dp_{\mathrm{T}}^{W}\,dm^{W}\,dy^{W}\,d\cos\theta d\phi} = \frac{d\sigma^{U+L}}{dp_{\mathrm{T}}^{W}\,dm^{W}\,y^{W}} \tag{2}$$

Integrating Equation 1 over ϕ results in:

$$\frac{d\sigma}{dp_{\mathrm{T}}^{W}\,dm^{W}\,dy^{W}\,d\cos\theta} = \frac{3}{8}\frac{d\sigma^{U+L}}{dp_{\mathrm{T}}^{W}\,dm^{W}\,y^{W}}\left\{\left(1+\cos^2\theta\right)+\frac{1}{2}A_0\left(1-3\cos^2\theta\right)+A_4\cos\theta\right\} \tag{3}$$

Integrating Equation 1 over $\cos\theta$ results in:

$$\frac{d\sigma}{dp_{\mathrm{T}}^{W}\,dm^{W}\,dy^{W}\,d\phi} = \frac{1}{2\pi}\frac{d\sigma^{U+L}}{dp_{\mathrm{T}}^{W}\,dm^{W}\,y^{W}}\left\{1+\frac{1}{4}A_2\cos 2\phi+\frac{3\pi}{16}A_3\cos\phi+\frac{1}{2}A_5\sin 2\phi+\frac{3\pi}{16}A_7\sin\phi\right\} \tag{4}$$

Integrating over either angle results in the loss of information of A_1 and A_6, which both require the full two-dimensional angular phase space.

1.2 Collins-Soper Frame

The Drell-Yan angular coefficients are reference frame dependent because they are dependent on the W boson kinematics. For this analysis the Collins-Soper (CS) frame [17] is chosen for direct comparison to the coefficients of the Z boson. The CS frame is defined using the rest frame of the W boson and the direction of the protons in the W boson rest frame. To define the z-axis the momentum of one proton and the inverse of the momentum of the second is used to create a bisecting angle, α as seen in Figure 1, and the sign is chosen to be in the same direction as the W boson's momentum in the laboratory frame. The y-axis is then chosen to be normal to the proton-proton plane and the x-axis is chosen to be orthogonal to complete a right-handed Cartesian coordinate system.

The direction of the leptons in the CS frame are defined by two angles, θ_{CS} and ϕ_{CS}. The angle θ_{CS} is defined as the angle between the momentum of the negatively charged lepton and the z-axis. In the case of W^+ there is no negatively charged lepton and instead the direction of the neutrino is used. The formula used to calculate $\cos\theta_{\text{CS}}$ can be expressed using the kinematics of the leptons in the laboratory frame [18]

$$\cos\theta_{\text{CS}} = \frac{p_z^{\ell\nu}}{|p_z^{\ell\nu}|} \frac{2}{m^W \sqrt{m^{W2} + p_{\text{T}}^{\ell\nu 2}}} \left(P_1^+ P_2^- - P_1^- P_2^+\right), \tag{5}$$

with

$$P_i^{\pm} = \frac{1}{\sqrt{2}} \left(E_i \pm p_{z,i}\right). \tag{6}$$

E_i and $p_{z,i}$ are the energy and longitudinal momentum of the lepton, respectively, such that $i = 1$ for the lepton and $i = 2$ for the anti-lepton. $p_z^{\ell\nu}$ is the longitudinal momentum of the dilepton system. The angle ϕ_{CS} is chosen to be the angle between the lepton plane and the xz-plane and must be calculated in the CS frame.

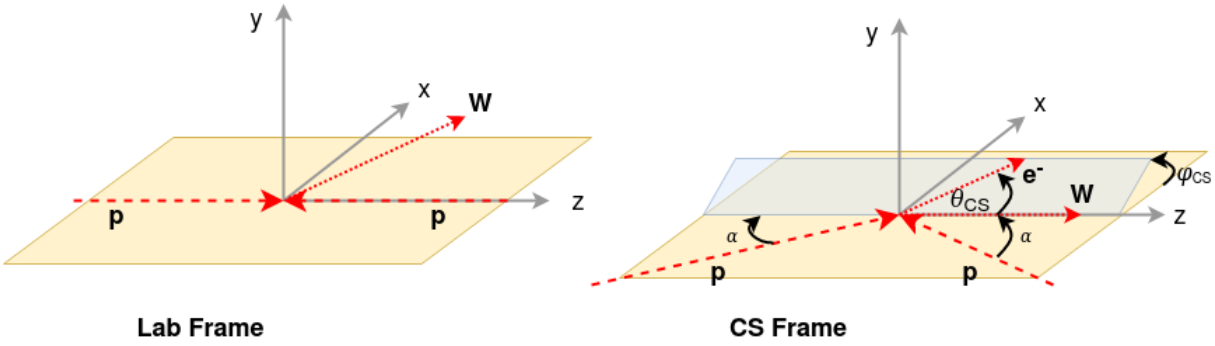


Figure 1: On the left is the lab frame with protons coming from opposite directions resulting in the W boson, in an arbitrary direction. On the right is the Collin-Soper frame.

1.3 Moments Method

A method to extract the coefficients from generated Monte Carlo events was proposed in [2] and was used in the analysis [6]. It is possible to isolate each coefficient while still remaining as functions of p_{T}^W and y^W. This is done by taking moments of the distributions with respect the appropriate product of trigonometric polynomials where the moment can be defined as

$$\langle P_i(\theta,\phi)\rangle = \frac{\int d\sigma(p_T, m, y, \theta, \phi) P_i(\theta,\phi) d\cos\theta d\phi}{\int d\sigma(p_T, m, y, \theta, \phi) d\cos\theta d\phi}. \tag{7}$$

For each coefficient this results in

$$
\begin{aligned}
\langle \frac{1}{2}\left(1-3\cos^2\theta\right)\rangle &= \frac{3}{20}\left(A_0-\frac{2}{3}\right); & \langle \sin 2\theta\cos\phi\rangle &= \frac{1}{5}A_1;\\
\langle \sin^2\theta\cos 2\phi\rangle &= \frac{1}{10}A_2; & \langle \sin\theta\cos\phi\rangle &= \frac{1}{4}A_3;\\
\langle \cos\theta\rangle &= \frac{1}{4}A_4; & \langle \sin^2\theta\sin 2\phi\rangle &= \frac{1}{5}A_5;\\
\langle \sin 2\theta\sin\phi\rangle &= \frac{1}{5}A_6; & \langle \sin\theta\sin\phi\rangle &= \frac{1}{4}A_7.
\end{aligned}
\tag{8}
$$

The underlying mechanic of taking the moment to isolate each coefficient is that it exploits the orthogonality of the spherical harmonics that the trigonometric polynomials are related to as in Table 1. The definition of the moment in Equation 7 requires the integration of the full lepton phase space for the orthogonality to hold so it can only be applied to MC samples.

$\mathbf{A_i}$	$\mathbf{P_i}$	$\mathbf{Y_l^m}\,(\theta,\phi)$	**Coupling**	**Non-Zero**
A_0	$1/2\left(1-3\cos^2\theta\right)$	Y_2^0		$O\left(\alpha_S^1\right)$
A_1	$\sin 2\theta\cos\phi$	$\left(Y_2^{-1}-Y_2^1\right)$	$\left(v_\ell^2+a_\ell^2\right)\left(v_q^2+a_q^2\right)$	$O\left(\alpha_S^1\right)$
A_2	$1/2\sin^2\theta\cos 2\phi$	$\left(Y_2^{-2}+Y_2^2\right)$		$O\left(\alpha_S^1\right)$
A_3	$\sin\theta\cos\phi$	$\left(Y_1^{-1}-Y_1^1\right)$	$v_\ell a_\ell v_q a_q$	$O\left(\alpha_S^1\right)$
A_4	$\cos\theta$	Y_1^0		$O\left(\alpha_S^0\right)$
A_5	$\sin^2\theta\sin 2\phi$	$\left(Y_2^{-2}-Y_2^2\right)$	$\left(v_\ell^2+a_\ell^2\right)\left(v_q a_q\right)$	$O\left(\alpha_S^2\right)$
A_6	$\sin 2\theta\sin\phi$	$\left(Y_2^{-1}+Y_2^1\right)$		$O\left(\alpha_S^2\right)$
A_7	$\sin\theta\sin\phi$	$\left(Y_1^{-1}+Y_1^1\right)$	$\left(v_\ell a_\ell\right)\left(v_q^2+a_q^2\right)$	$O\left(\alpha_S^2\right)$

Table 1: Summary of Drell-Yan angular coefficients, polynomials coupled to each coefficient, how the polynomials relate to the spherical harmonics, the vector and axial-vector couplings, and order at which they become non-zero.

1.4 Measurement Goals

The goal of this analysis is to measure four (A_0, A_2, A_3, and A_4) of the eight of the coefficients in Equation 1 as well as the differential cross-section. The coefficients A_1 and $A_5 - A_7$ are out of the scope of this measurement due the statistical sensitivity required as found in the sensitivity studies described in Appendix A . These coefficients will still be extracted at the same time but are expected to give inconclusive results. The coefficients are to be measured as a function of p_{T}^W and independently as function of —y^W — for W^- and W^+ in both an $e^-\overline{\nu}$ ($e^+\nu$) and $\mu^-\overline{\nu}$ ($\mu^+\nu$) channel.

2 Theoretical Predictions

This chapter will contain the various theoretical predictions to be compared against once they are compiled and the theoretical formalism. We plan to compare the measured differential distributions to fixed order predictions up to order α_S^3 in the perturbative expansion of the strong-coupling constant [19], which is next-to-next-to-leading order (NNLO) for the transverse-momentum distribution and next-to-next-to-next-to-leading order (N3LO) for the rapidity distribution. We also plan to include comparisons of the differential distributions with POWHEG+PYTHIA8 *and DYTurbo.*

2.1 Theoretical Formalism

The lepton angular distribution in the gauge boson rest frame is determined by the gauge boson polarization. The general structure of the angular distribution is given by nine helicity cross-sections corresponding to the nine helicity density matrix elements for the gauge boson [1]. In the LO parton subprocesses ($q\overline{q} \rightarrow V \rightarrow \ell\ell'$) only the transverse polarization gauge boson contributes to the helicity density matrix. At higher order, in general, all three polarizations of the gauge boson contribute. At $O\left(\alpha_{\mathrm{S}}\right)$ in perturbative QCD the angular distribution is described by six helicity cross-sections, which are functions of the transverse momentum and rapidity of the gauge boson. At the$O\left(\alpha_{\mathrm{S}}^2\right)$ contribution from additional three helicity cross-sections is non-zero. Following convention and notations of [2], let us briefly recall that the lepton-hadron correlations are described by the contraction of the lepton tensor $L_{\mu\nu}$ with the hadron tensor at the parton level $H^{\mu\nu}$, where $L_{\mu\nu}$ acts as an analyzer of the structure of $H^{\mu\nu}$ which carries the effective information on the polarization of the gauge boson produced in the interaction. The angular dependence can be extracted introducing helicity cross-sections corresponding to the non-zero combinations of the polarization density matrix elements

$$H_{mm'} = \epsilon_{\mu}^{*}\left(m\right) H^{\mu\nu} \epsilon_{\nu}\left(m'\right) \tag{9}$$

where $m, m' = +1, 0, -1$ and

$$\epsilon_{\mu}\left(\pm 1\right) = \frac{1}{\sqrt{2}}\left(0, \pm 1, -i, 0\right), \quad \epsilon_{\mu}\left(0\right) = \frac{1}{\sqrt{2}}\left(0, 0, 0, 1\right) \tag{10}$$

are the polarization vectors for the gauge boson defined with respect to the chosen gauge boson rest frame. The angular dependence of the differential cross-section can be written as

$$\frac{d\sigma}{dp_{\mathrm{T}}^{W} dm^{W} dy^{W} d\cos\theta d\phi} = \frac{3}{16\pi} \sum_{\alpha \in M} g_{\alpha}\left(\theta, \phi\right) \frac{d\sigma^{\alpha}}{dp_{\mathrm{T}}^{W} dm^{W} dy^{W}}, \quad M = \{U{+}L, L, T, I, P, A, 7, 8, 9\} \tag{11}$$

where the g_{α} represent the second-order harmonic polynomials multiplied by normalization constants, $\cos\theta$ and ϕ are the polar and azimuthal decay angles of the leptons in the gauge boson reference frame, and $d\sigma^{\alpha}$ denote helicity cross sections corresponding to nine helicity matrix elements.

$$
\begin{aligned}
\sigma^{U+L} &\sim H_{00} + H_{++} + H_{--} \\
\sigma^{L} &\sim H_{00} \\
\sigma^{T} &\sim 1/2\,(H_{+-} + H_{-+}) \\
\sigma^{I} &\sim 1/4\,(H_{+0} + H_{0+} - H_{-0} - H_{0-}) \\
\sigma^{P} &\sim H_{++} + H{--} \\
\sigma^{A} &\sim 1/4\,(H_{+0} + H_{0+} + H_{-0} + H_{0-}) \\
\sigma^{7} &\sim -i/2\,(H_{+-} - H_{-+}) \\
\sigma^{8} &\sim -i/4\,(H_{+0} - H_{0+} + H_{-0} - H_{0-}) \\
\sigma^{9} &\sim -i/4\,(H_{+0} - H_{0+} - H_{-0} + H_{0-})
\end{aligned}
\tag{12}
$$

The unpolarized differential cross section is denoted by σ^{U+L}, where $\sigma^{L,T,I,P,A}$ characterize polarization of the gauge boson, respectively cross-section for longitudinally polarized gauge boson, transverse interference cross-section, transverse-longitudinal interference cross-section, etc. [1]. The cross-sections are related to the vector v_x and axial-vector a_x coupling coefficients of $x = \{\ell, q\}$ to the W boson in the following way [1]:

$$
\begin{aligned}
\sigma^{U+L,L,T,I} &\sim \left(v_\ell^2 + a_\ell^2\right)\left(v_q^2 + a_q^2\right). \\
\sigma^{P,A} &\sim v_\ell a_\ell v_q a_q. \\
\sigma^{7,8} &\sim \left(v_\ell^2 + a_\ell^2\right)\left(v_q a_q\right). \\
\sigma^{9} &\sim \left(v_\ell a_\ell\right)\left(v_q^2 + a_q^2\right).
\end{aligned}
\tag{13}
$$

The $\sigma^{U+L,L,T,I,9}$ cross-sections receive contributions from the parity conserving (p.c.) part of the hadron tensors, the remaining four $sigma^{P,A,7,8}$ are proportional to the parity violating (p.v.) part of $H^{\mu\nu}$, i.e. they change sign under parity (P) transformation. Since the angular polynomials $g_{P,A,9}\,(\theta, \phi)$ change sign too, angular distributions involving $\sigma^{U+L,L,T,I,P,A}$ are parity conserving.

The standard polynomials, $g_\alpha\,(\theta, \phi)$, are:

$$
\begin{aligned}
g_{U+L} &= 1 + \cos^2\theta; & g_L &= 1 - 3\cos^2\theta; & g_T &= 2\sin^2\theta\cos 2\phi; \\
g_I &= 2\sqrt{2}\sin 2\theta\cos\phi; & g_P &= 2\cos\theta; & g_A &= 4\sqrt{2}\sin\theta\cos\phi; \\
g_7 &= 2\sin^2\theta\sin 2\phi; & g_8 &= 2\sqrt{2}\sin 2\theta\sin\phi; & g_9 &= 4\sqrt{2}\sin\theta\sin\phi.
\end{aligned}
\tag{14}
$$

The helicity cross sections $d\sigma^\alpha$ are obtained by convoluting the partonic cross sections of partons with momenta $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$

$$
\frac{d\sigma^\alpha}{dp_{\mathrm{T}}^W dm^W dy^W} = \sigma_{a,b} \int \mathrm{x}_1 dx_2 f_1^{h1}\left(x_1, \mu_F^2\right) f_2^{h2}\left(x_2, \mu_F^2\right) s \frac{d\hat{\sigma}_{ab}^\alpha}{dt\,du}\left(p_1, p_2, \alpha_{\mathrm{S}}\left(\mu_R^2\right)\right) \tag{15}
$$

where μ_F and μ_R denote respectively factorization and renormalization scale. The standard notation is then to factorize out unpolarized cross-section and present formula of Equation 11 as expansion into standard polynomials and dimensionless angular coefficients representing ratios of helicity cross-sections.

$$
\begin{aligned}
A_0 &= 2d\sigma^L/d\sigma^{U+L}; \quad & A_1 &= 2\sqrt{2}d\sigma^I/d\sigma^{U+L}; \quad & A_2 &= 4d\sigma^T/d\sigma^{U+L}; \\
A_3 &= 4\sqrt{2}d\sigma^A/d\sigma^{U+L}; \quad & A_4 &= 2d\sigma^P/d\sigma^{U+L}; \quad & A_5 &= 2d\sigma^7/d\sigma^{U+L}; \\
A_6 &= 2\sqrt{2}d\sigma^8/d\sigma^{U+L}; \quad & A_7 &= 4\sqrt{2}d\sigma^9/d\sigma^{U+L}.
\end{aligned}
\tag{16}
$$

This conveniently rewrites the equation into the form of Equation 1 or more compactly as

$$
\frac{d\sigma}{dp_{\mathrm{T}}^W dm^W dy^W d\cos\theta d\phi} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dp_{\mathrm{T}}^W dm^W dy^W}\left\{\left(1+\cos^2\theta\right)+\sum_{i=0}^{7}A_i\left(p_{\mathrm{T}}^W, y^W, m^W\right)P_i\left(\cos\theta,\phi\right)\right\}. \tag{17}
$$

In the limit of zero transverse momenta all coefficients except A_4 vanish. Another important thing to note is that because of the (p_{T}^W, y^W) dependence on the A_i coefficients depend strongly on the choice of the z-axis. The coefficients A_0, A_2, and A_3 are increasing functions of p_{T}^W and the deviations from lowest order expectation ($A_0 = A_2 = 0$) are quite large at modest value of p_{T}^W, i.e. p_{T}^W = 20 - 50 GeV. These coefficients are exactly equal at LO [20] but is not true at NLO where corrections produce deviations of less than 10% [2]. The deviation of A_1 from zero is much smaller even at large p_{T}^W. A_4 contributes the most at small p_{T}^W and decreases with as p_{T}^W increases. A_3 and A_4 contribute more for the W boson compared to the Z boson due to the fact that A_3 and A_4 are proportional to $v_\ell a_\ell v_q a_q$ which means that in the case of the Z boson they are proportional to the weak mixing angle and are suppressed.

The theoretical uncertainties due to the choice of the factorization and renormalization scales are very small for the cross section ratios A_i. Also, most of the uncertainties of the structure functions and of the choice of the factorization scheme cancels in the ratio.

3 Measurement Methodology

This chapter will cover the measurement methodology using a log-likelihood to extract the angular coefficients. A more detailed description can also be found in the internal note [21]. The methodology often expresses the measurement in terms of p_{T}^{W} but the same methodology is used for extracting the coefficients in y^{W}.

3.1 Likelihood Method

A likelihood is built from the nominal templates and the varied templates reflecting the systematic uncertainties. A set of nuisance parameters (NPs) $\theta = \{\beta, \gamma\}$ is used to interpolate between them which are constrained by auxiliary probability density functions.

The first category, β, are the NPs representing experimental and theoretical uncertainties. Each β^m in the set $\beta = \left\{\beta^1, \ldots, \beta^M\right\}$ are constrained by unit Gaussian probability density functions $G\left(0|\beta^m, 1\right)$ and interpolate between nominal and varied templates. They are defined by having a nominal value of zero and $\beta^m = \pm 1$ corresponding to $\pm 1\sigma$ for the considered systematic uncertainty.

The second category, γ, are the NPs that handle systematic uncertainties from the limited size of the MC samples. For each $\cos\theta_{\mathrm{CS}}$, ϕ_{CS} reco bin n there is a γ^n in the set $\gamma = \left\{\gamma^1, \ldots, \gamma^{N_{\mathrm{bins}}}\right\}$, where $N_{\mathrm{bins}} = 8 \times 11$ is the number of $\cos\theta_{\mathrm{CS}}$, ϕ_{CS} bins in each reco kinematic $(p_{\mathrm{T}}^{W}, y^{W})$ bin. γ^n has a nominal value of one and normalizes the expected events in bin n of the templates. They are constrained by Poisson probability density functions, $P\left(N_{\mathrm{eff}}^n | \gamma^n N_{\mathrm{eff}}^n\right)$, where N_{eff}^n is the effective number of MC events in bin n. Effective is used because of the non-uniform event weights used for to properly scale the MC to account for things like detector inefficiencies. The expected number of events in each bin n can be written as:

$$N_{\mathrm{exp}}^n\left(A, \sigma, \theta\right) = \left\{\sum_{j=0}^{N_{\mathrm{bins}}^{\mathrm{ana}}} \sigma_j \times L\left[T_{8,j}^n\left(\beta\right) + \sum_{i=0}^{7} A_{ij} \times T_{ij}^n\left(\beta\right)\right]\right\} \times \gamma^n + \sum_{B}^{\mathrm{bkgds}} T_B^n\left(\beta\right) \tag{18}$$

where:

- A_{ij}: Coefficient parameter A_i for observation bin j
- A: Set of all A_{ij}
- σ_j: Signal cross-section parameter
- σ: Set of all σ_j
- θ: Set of all NPs
- L: Integrated luminosity constant
- β: Set of all Gaussian-constrained NPs
- γ^n: Poisson-constrained NP
- T_{ij}: P_i template
- T_B: Background templates.

Summation over index j takes into account the contribution for all truth kinematic bins at generator levels in each reconstructed kinematic bin to account for migration between bins as explained in Section 7.1. Far away bins have very little contribution and are pruned if the expected number of events is smaller than 10^{-4} in order to improve CPU and memory requirements.

The likelihood is the product of Poisson probabilities across all N_{bins} and of auxiliary constraints for each NP β_m:

$$\mathcal{L}\left(A, \sigma, \theta | N_{\text{obs}}\right) = \prod_{n}^{N_{\text{bins}}} \left\{ P\left(N_{\text{obs}} | N_{\text{exp}}^{n}\left(A, \sigma, \theta\right)\right) P\left(N_{\text{eff}}^{n} | \gamma^{n} N_{\text{eff}}^{n}\right) \right\} \times \prod_{m}^{M} G\left(0 | \beta^{m}, 1\right) \tag{19}$$

3.2 Template Folding

We define a set of observables in truth space t with underlying function $f\left(t\right)$. Different detector effects like limited acceptance, reconstruction efficiency, and resolution will smear our observables t to have reconstructed values r with distribution $g\left(r\right)$. The probability to observe a reconstructed r given truth values t is expressed as $p\left(r|t\right)$. This probability can be convoluted with the truth distribution to get the reconstructed distribution:

$$g\left(r\right) = \int f\left(t\right) p\left(r|t\right) dt \tag{20}$$

Another way to express this is that the truth distribution $f\left(t\right)$ is folded through the convolution to produce $g\left(r\right)$.

Likewise, the polynomials in the differential cross-section decomposition can be folded in this manner to describe the reconstructed distributions. In this case the set of truth and reconstructed variables are $t = \left(\cos\theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}, p_{\text{T}}^{W,\text{Truth}}\right)$ and $r = \left(\cos\theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}}, p_{\text{T}}^{W,\text{Reco}}\right)$ respectively. The folded polynomial $\tilde{P}_{ij}\left(r | p_{\text{T}}^{W,\text{Truth}} \in \left(\Delta p_{\text{T}}^{W,\text{Truth}}\right)_j\right)$ is restricted to the finite bin width, $\left(\Delta p_{\text{T}}^{W,\text{Truth}}\right)_j$, and can be expressed as:

$$\tilde{P}_{ij}\left(r | p_{\text{T}}^{W,\text{Truth}} \in \left(\Delta p_{\text{T}}^{W,\text{Truth}}\right)_j\right) = \int_{\left(\Delta p_{\text{T}}^{W,\text{Truth}}\right)_j} \int_{\cos\theta_{\text{CS}}^{\text{Truth}}} \int_{\phi_{\text{CS}}^{\text{Truth}}} P_i\left(\cos\theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}\right) p\left(r|t\right) dt. \tag{21}$$

The folded polynomials acquire a $p_{\text{T}}^{W,\text{Reco}}$ dependence from the migrations in this observable and it is important to distinguish this from $p_{\text{T}}^{W,\text{Truth}}$. The $\tilde{P}_{ij}$ are a function of $p_{\text{T}}^{W,\text{Reco}}$ but are categorized in bins of $p_{\text{T}}^{W,\text{Truth}}$ by index j. It is also important to note that $\tilde{P}_{ij}$ are continuous versions of the templates T_{ij} in Equation 18 and not used in this analysis but to transition the concepts to the actual folded templates used.

There is no analytical expressions for $p\left(r|t\right)$ so MC is used to approximate them. This is done by calculating the reference coefficients, A_{ij}^{ref}, with the moments method described in Section 1.3. These reference coefficients fully describe the 2D angular PDF, $f_j\left(\cos\theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}\right)$ in the full phase space in each $p_{\text{T}}^{W,\text{Truth}}$ bin j:

$$f_j\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)=\sigma_j\left\{P_8\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)+\sum_{i=0}^{7}A_{ij}^{\mathrm{ref}}P_i\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)\right\}. \tag{22}$$

Where σ_j is the differential cross-section in the $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin j. For an MC event with weight w_{event} and corresponding truth and reconstructed variables t and r, respectively, we can estimate $p\left(r|t\right)$ with:

$$p^{MC}\left(r|t\right)=\frac{w_{\mathrm{event}}\left(r,t\right)}{f_j\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)}. \tag{23}$$

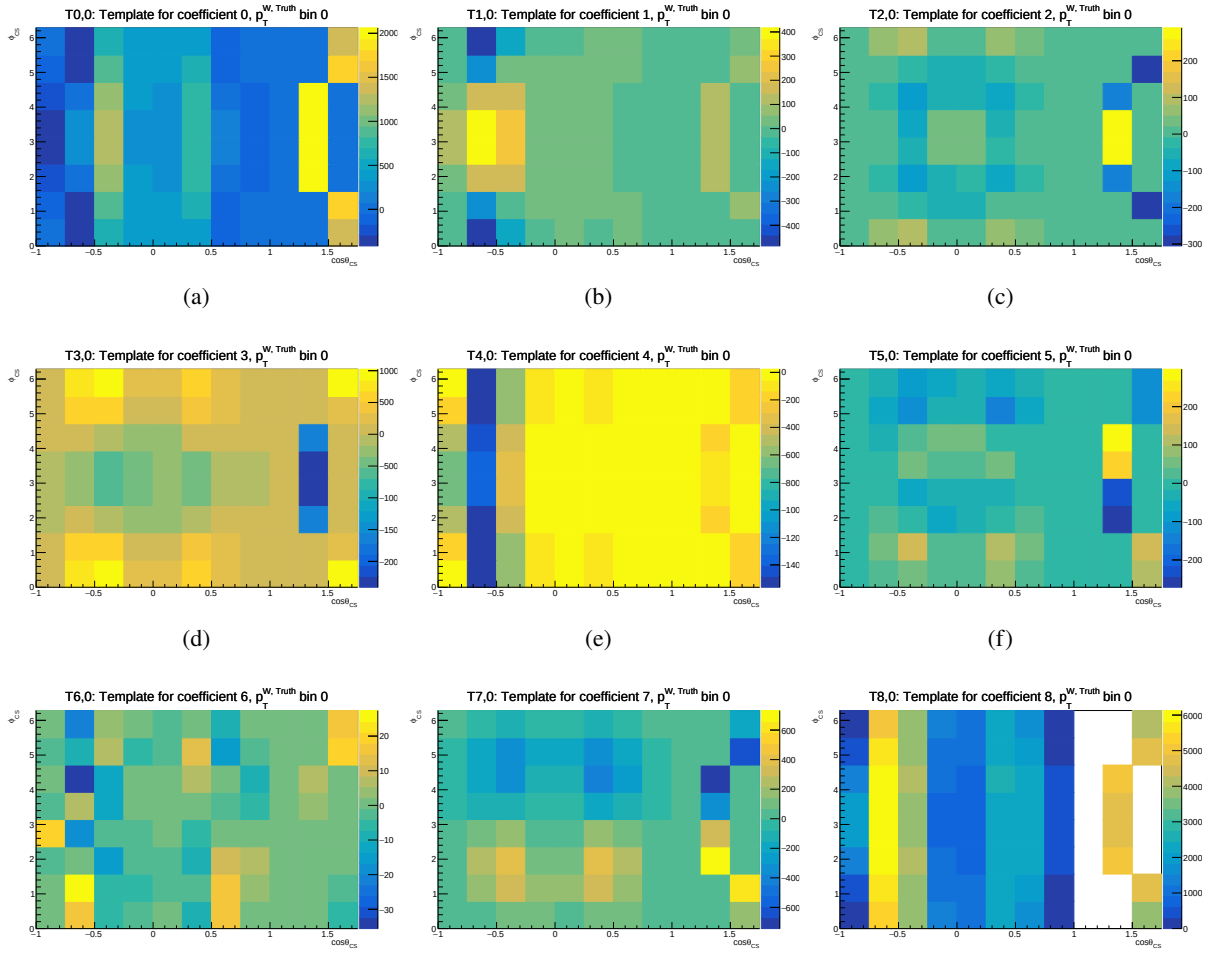
Here w_{event} is the produce of generator weights and scale factor weights used in the analysis. Division by f_j removes strong generator dependence on the coefficients used to build the templates assuming the MC follows the proper decomposition giving closure to unity. The final templates T_{ij} can then be expressed as the sum over all events within each bin j:

$$T_{ij}=\sum_{event\in\Delta_j}\frac{P_i\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}},\phi_{\mathrm{CS}}^{\mathrm{Reco}}\right)w_{\mathrm{event}}\left(r,t\right)}{\sigma_j\left\{P_8\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)+\sum_{i=0}^{7}A_{ij}^{\mathrm{ref}}P_i\left(\cos\theta_{\mathrm{CS}}^{\mathrm{Truth}},\phi_{\mathrm{CS}}^{\mathrm{Truth}}\right)\right\}}, \tag{24}$$

where Δ_j is the observable bins for the event and can be expressed as $\Delta_j=\left(\Delta p_{\mathrm{T}}^{W,\mathrm{Truth}}\right)_j$, $\left(\Delta p_{\mathrm{T}}^{W,\mathrm{Reco}}\right)_j$, $\left(\Delta\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}}\right)_j$, and $\left(\Delta\phi_{\mathrm{CS}}^{\mathrm{Reco}}\right)_j$. Examples of the templates for $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins can be seen in Figures 2, 3, and 4.

For the templates there are 11 bins in $\cos\theta_{\mathrm{CS}}$ in the range [-1,1.75] and 8 bins in ϕ_{CS} in the range [0, 2π]. The $\cos\theta_{\mathrm{CS}}$ range is extended past 8 bins in the range of [-1,1] with three additional $\cos\theta_{\mathrm{CS}}$ bins with bin boundaries [1, 1.25, 1.5, 1.75] in order to separate events where there is no real solution for p_{z}^{ν}. This is done to not distort the shape of $\cos\theta_{\mathrm{CS}}$ which is discussed in more detail in Section 4.4. The first bin is intentionally left empty in order to prevent bins from either solution region from merging if there are not enough events. The other two bins are for events that are either forward $\cos\theta_{\mathrm{CS}}>0$ or backward $\cos\theta_{\mathrm{CS}}<0$ and allow us to retain the ϕ_{CS} shape information for these events. The templates are weighted by the polynomials as seen in Equation 24 which means that bins will have negative values in some bins.

To understand how the shape of the angular distributions evolve with $p_{\mathrm{T}}^{W,\mathrm{Truth}}$, the polynomial templates are projected onto their relevant 1D axis. For A_0 and A_4 this is integrating over ϕ_{CS} as they are only dependent on $\cos\theta_{\mathrm{CS}}$ as seen in Figures 5 and 6, while for A_2 and A_3 this is integrating over $\cos\theta_{\mathrm{CS}}$ as seen in Figures 7 and 8. This doesn't work for A_1 as it depends on $\cos\theta_{\mathrm{CS}}$ and ϕ_{CS} as discussed in Section 1.1



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 2: Folded polynomial templates projected onto the $(\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}}, \phi_{\mathrm{CS}}^{\mathrm{Reco}})$ plane for $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin 0 for $W^- \to e^-\overline{\nu}$. The templates are weighted by the polynomials resulting in negative values in some bins.

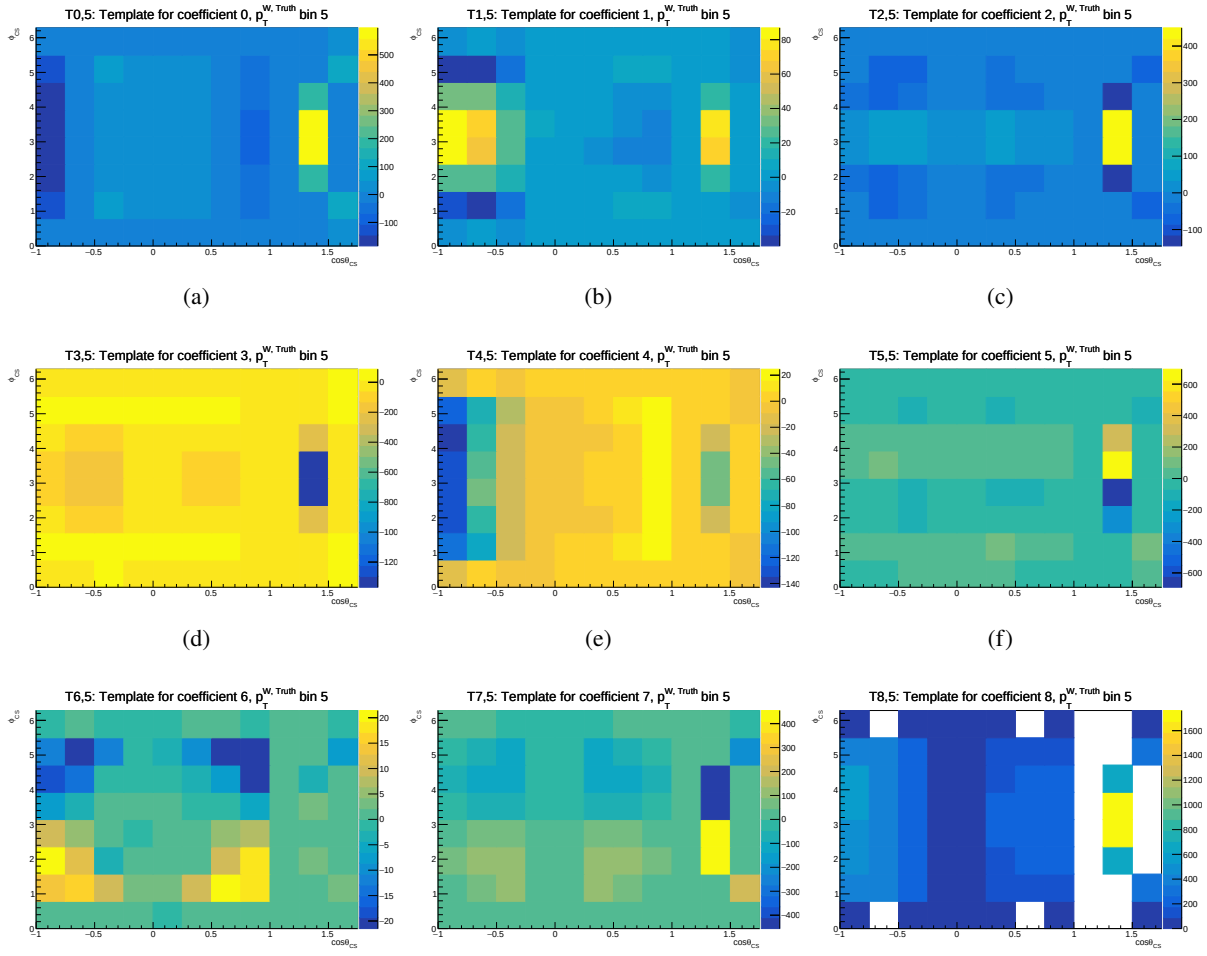


Figure 3: Folded polynomial templates projected onto the $(\cos\theta^{\text{Reco}}_{\text{CS}}, \phi^{\text{Reco}}_{\text{CS}})$ plane for $p_{\text{T}}^{W,\text{Truth}}$ bin 5 for $W^- \to e^-\overline{\nu}$. The templates are weighted by the polynomials resulting in negative values in some bins.

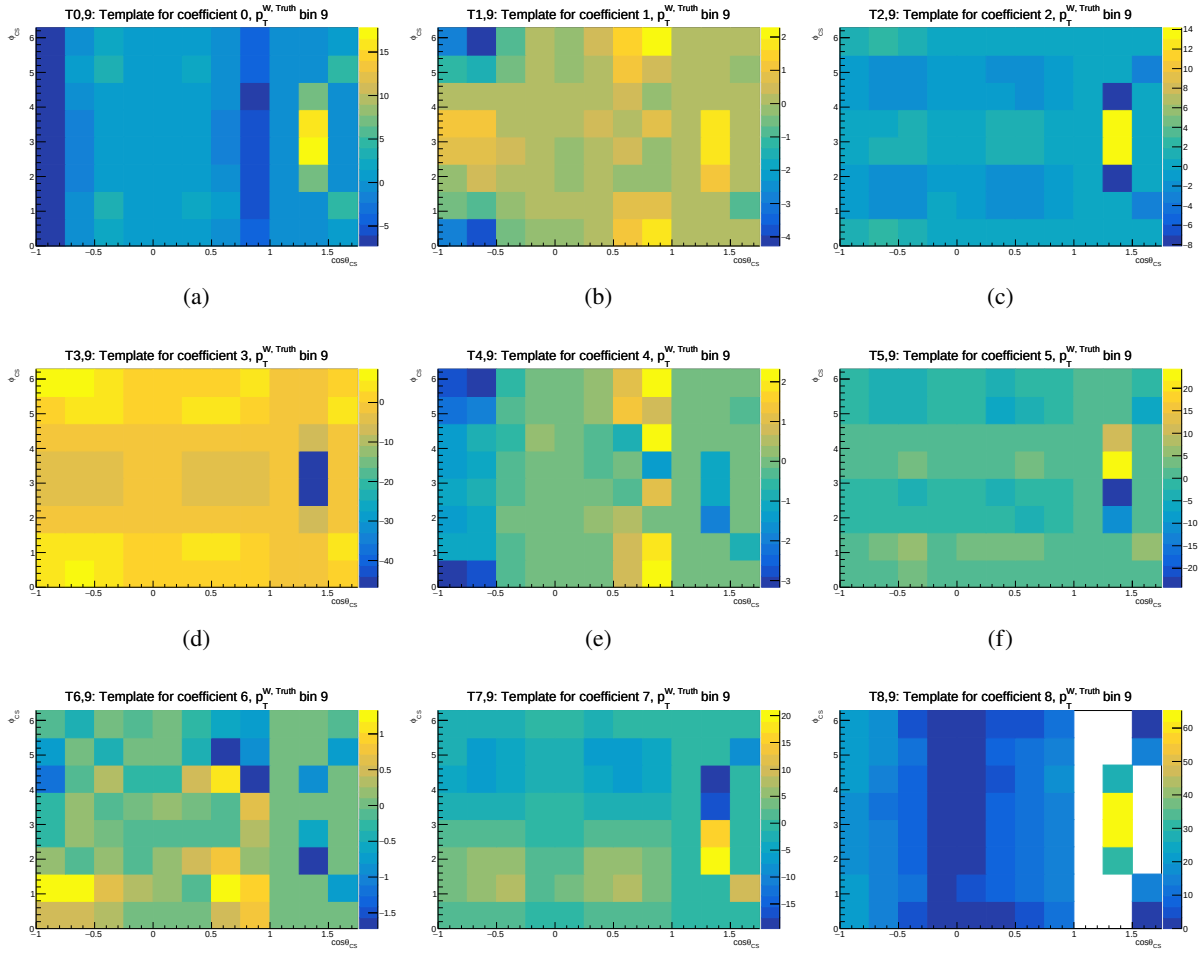


Figure 4: Folded polynomial templates projected onto the $(\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}}, \phi_{\mathrm{CS}}^{\mathrm{Reco}})$ plane for $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin 9 for $W^- \to e^-\overline{\nu}$. The templates are weighted by the polynomials resulting in negative values in some bins.

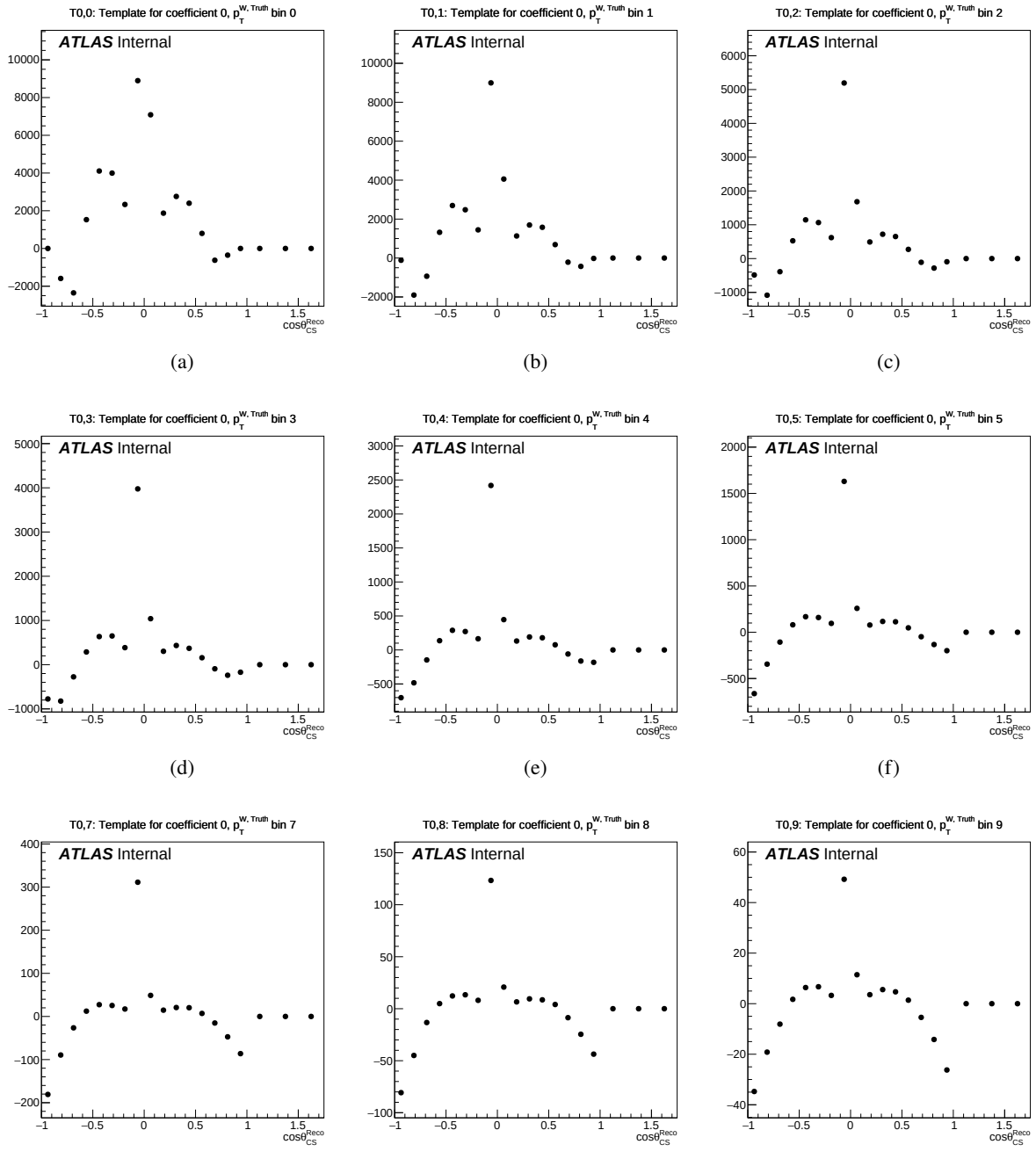


Figure 5: 1D Projections of folded polynomial templates onto $\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}}$ for $T_0 j$ for each $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin for $W^- \to e^-\overline{\nu}$.

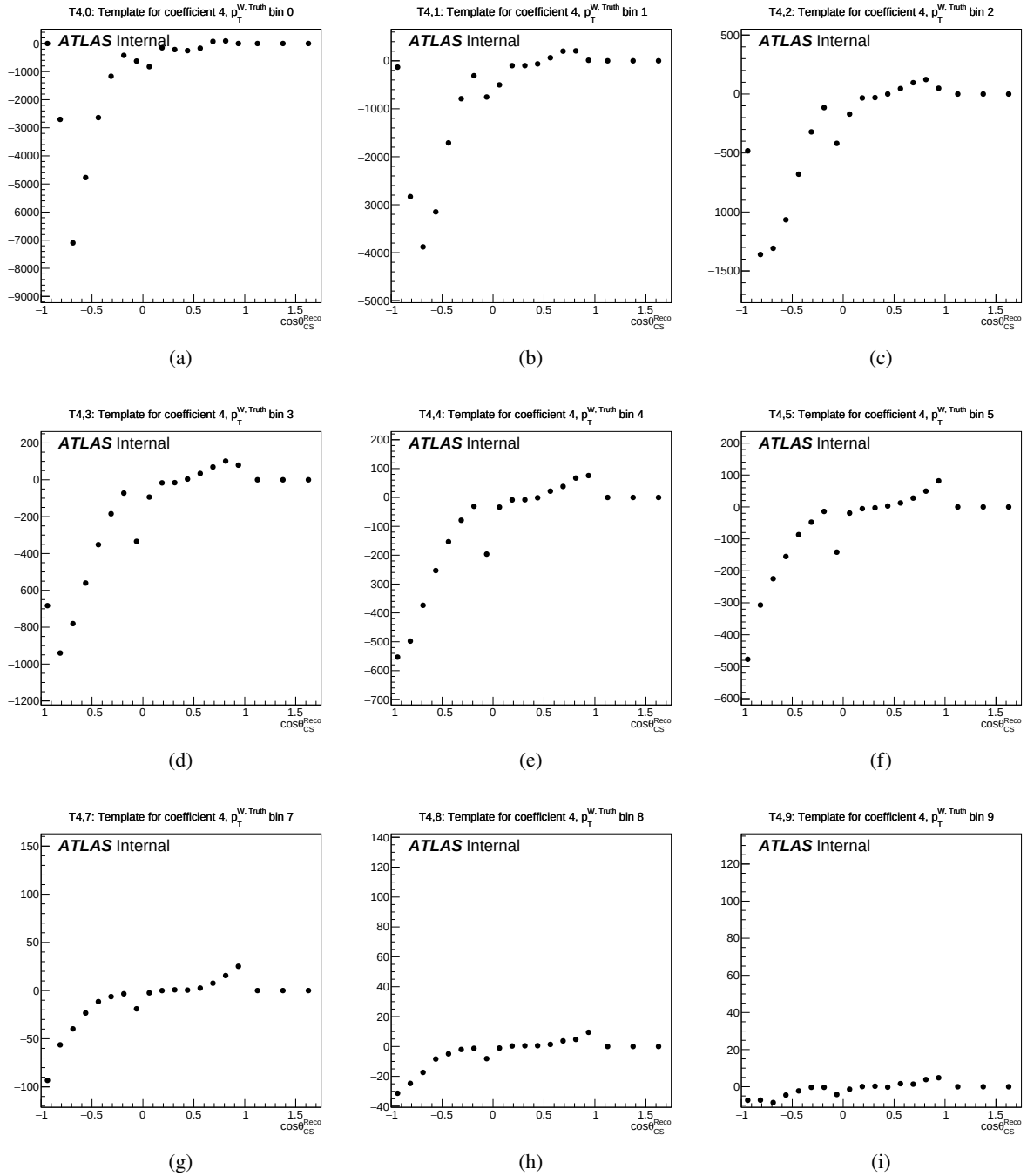


Figure 6: 1D Projections of folded polynomial templates onto $\cos\theta_{\mathrm{CS}}^{\mathrm{Reco}}$ for $T_4 j$ for each $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin for $W^- \to e^-\overline{\nu}$.

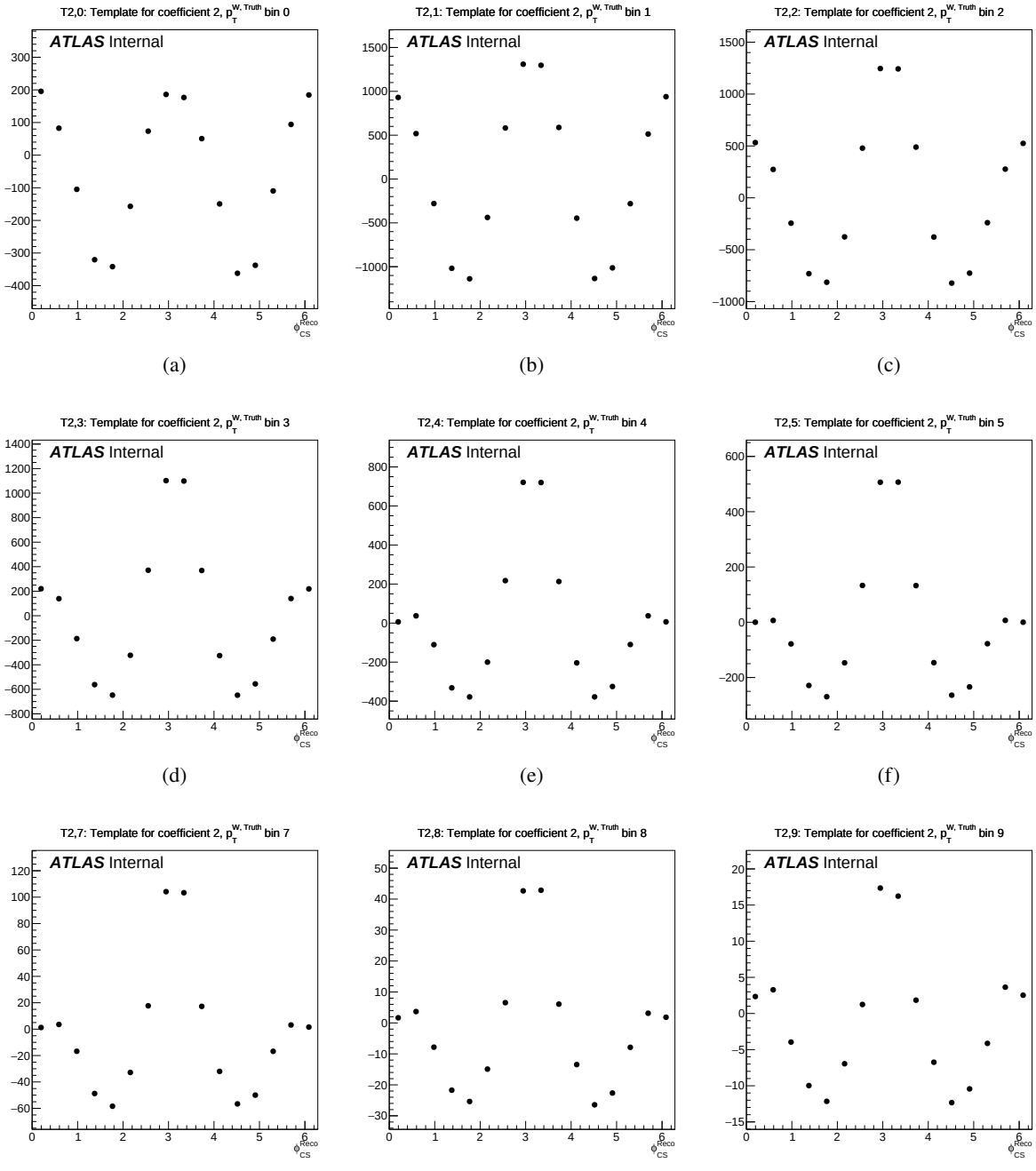


Figure 7: 1D Projections of folded polynomial templates onto $\phi_{\mathrm{CS}}^{\mathrm{Reco}}$ for $T_2 j$ for each $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin for $W^- \to e^- \overline{\nu}$.

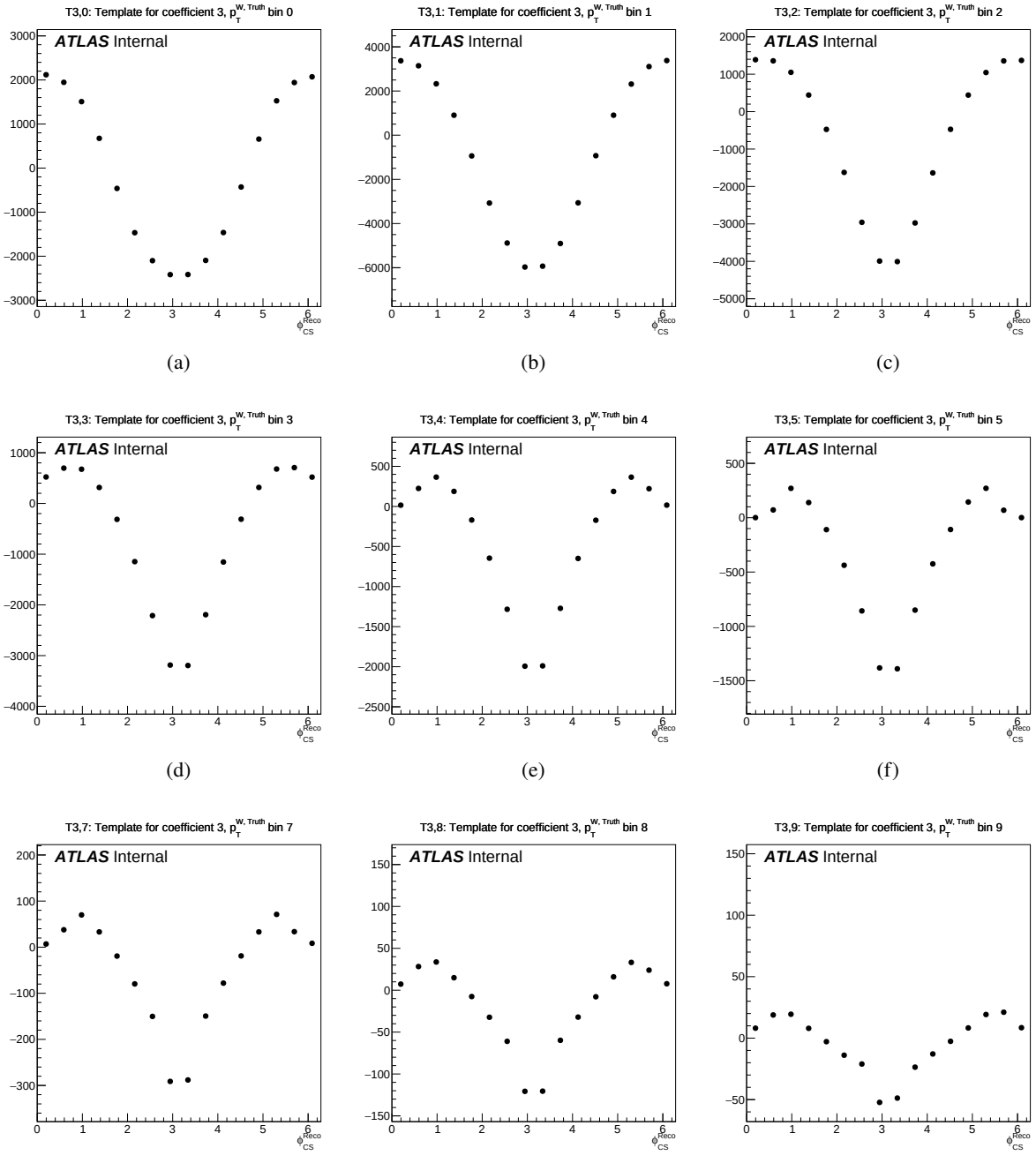


Figure 8: 1D Projections of folded polynomial templates onto $\phi_{\mathrm{CS}}^{\mathrm{Reco}}$ for $T_3 j$ for each $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin for $W^- \to e^- \overline{\nu}$.

3.3 Minimization Technique

To extract the coefficients the likelihood is minimized. To do this we can express Equation 18 as

$$N_{\text{exp}}\left(\alpha,\theta\right) = \Sigma_i\left[\Sigma_j\left(T_{ij}\Pi_k\alpha_k\right)\right]S_i\left(\theta\right), \tag{25}$$

where T_{ij} are bin values of templates, α_k multiply templates, $S_i\left(\theta\right)$ are systematic variations on $\Sigma_j\left(T_{ij}\right)$, and θ are NPs. For this analysis α are multipliers like A_{ij} and σ_j, or gamma parameters for MC statistics while θ are other NPs. The binned NLL is then

$$-\ln\mathcal{L} = \Sigma_n - \ln P\left(N_{\text{data}}|N_{\text{exp}}\right) + \Sigma_m P\left(N_{\text{MC}}|g_m N_{\text{MC}}\right) + \Sigma_1 G\left(0|\theta_1\right). \tag{26}$$

A layered cache is used to compute likelihood terms and derivatives in bottom-up approach compared to RooFit top-down approach as shown in Figure 9. Derivatives are implemented via chain rules for the NLL,

$$\frac{\delta\left(-\ln\mathcal{L}\right)}{\delta\beta} = \Sigma_{n+m}\frac{\delta\left(-\ln P\right)}{\delta\beta} + \Sigma_1\frac{\delta\left(-\ln G\right)}{\delta\beta}, \tag{27}$$

where the negative log Poisson is

$$\frac{\delta\left(-\ln P\right)}{\delta\beta} = \left(1 - \frac{N_{\text{obs}}}{N_{\text{exp}}}\right)\frac{\delta N_{\text{exp}}}{\delta\beta}. \tag{28}$$

Derivatives are computed within a layered cache similar to the 0^{th} order values where it is optimized for derivatives we know are zero and utilize efficient sparse matrix in Eigen [22] for using the full Hessian $\nabla^2\left(-\ln\mathcal{L}\right)$.

Full analytical Hessian information enables us to use a more efficient minimization algorithm based on Newton's method, where a more detailed description can be found in Appendix G which performs root finding on the gradient of the NLL to find the solution of $\nabla\left[-\ln\mathcal{L}\left(\vec{x}\right)\right] = 0$. This is an iterative method where we chose to iterate until $|-\ln\mathcal{L}\left(\vec{x}_{n+1}\right) - -\ln\mathcal{L}\left(\vec{x}_n\right)| < \epsilon$, where ϵ is the chosen precision which for this analysis a value of $\epsilon = 10^{-4}$ is used.

A sanity check is also performed to make sure that $\nabla\left[-\ln\mathcal{L}\left(\vec{x}\right)\right] = 0$ corresponds to a minima and not a maximum or saddle point. Checks are also done so that $\mathbf{H}$ is positive definite, which is that is is invertible and the covariance matrix, $\mathbf{H}^{-1}$ has all positive diagonals.

Full information of the Hessian allows us to compute the full uncertainty decomposition quickly and accurately because the total uncertainty for a given parameter is the square root of the diagonal of the covariance matrix. For a more detailed description see Appendix G.

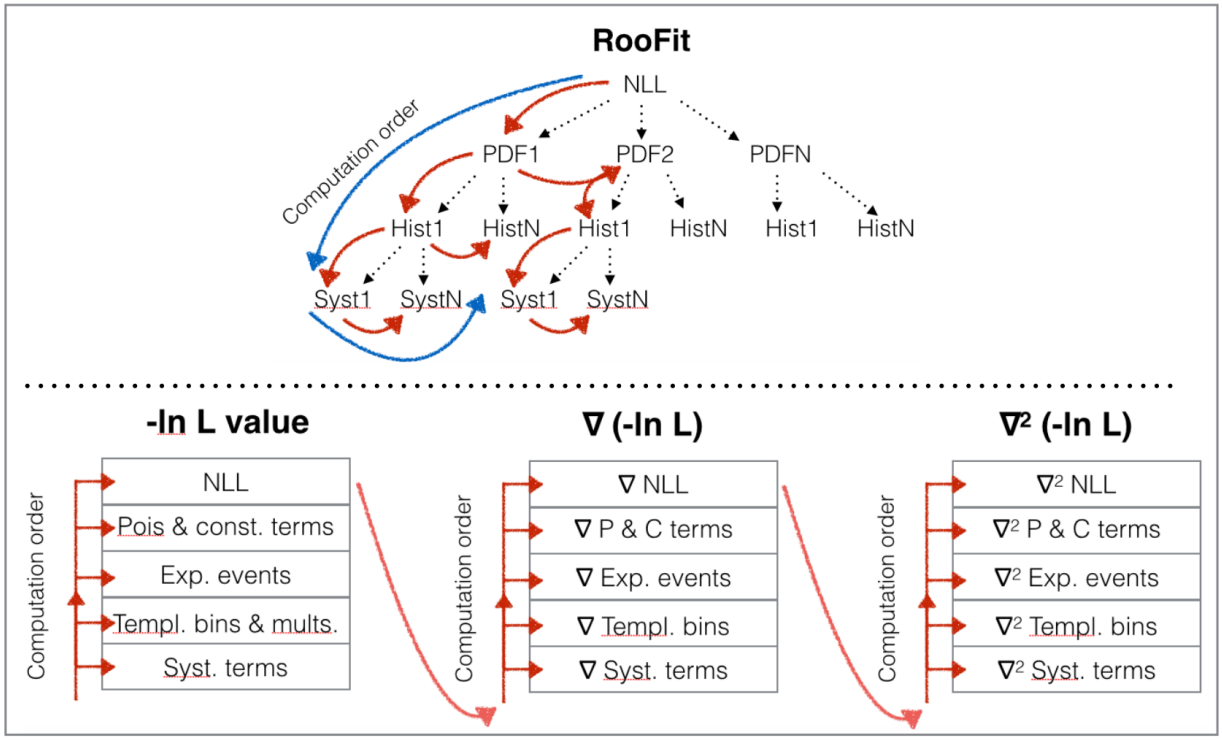


Figure 9: Comparison of computation order of Roofit which takes a top down approach to the bottom up approach taken for this analysis.

4 Analysis

This section describes the dataset consider in this analysis the event selection, the reconstruction of the unmeasured longitudinal momentum of the neutrino.

4.1 Data Sets

This analysis uses ATLAS special runs at 13TeV taken in Fall 2017 and July 2018 with low pile-up conditions ($\langle\mu\rangle \approx 2$) to optimise the measurement resolution of the hadronic system recoiling against the bosons that gives a direct measurement of the W-boson transverse momentum. The following runs are considered for analysis:

- $\sqrt{s}$ = 13TeV data taken in November 2017, ATLAS data period N, preliminary calibrated luminosity 144.931pb^{-1} with an uncertainty of ±2.1% (tag OflLumi-13TeV-010-lowmu),
  GRL data17_13TeV.periodN_DetStatus-v98-pro21-16_Unknown_PHYS_StandardGRL_All_Good_25ns_ignore_GLOBAL_LOWMU.xml:
  Runs 341294, 341312, 341419, 341534, 341615, 341649
- $\sqrt{s}$ = 13TeV data taken in June 2018, ATLAS data period G4+J, preliminary calibrated luminosity 190.249pb^{-1} with an uncertainty of ±1.5% (tag OflLumi-13TeV-010-lowmu),
  GRL data18_13TeV.periodG4J_MERGED_PHYS_StandardGRL_All_Good_25ns_ignore_GLOBAL_LOWMU.xml:
  Runs 354396, 355331, 355389, 355416, 355468

The total luminosity for the 13 TeV data amounts to 335.180pb^{-1} with an uncertainty of ±1.5%. The 13 TeV data was taken with a fixed levelled luminosity at $\langle\mu\rangle = 2$ (with minor exceptions).

Several notable differences to the standard high μ data taking are present in addition to the lower pileup

- Lower topo-cluster thresholds are used in the reconstruction to optimise the response of topo-clusters for the hadronic recoil; this has also implications for the electron response
- Single e and μ triggers with significantly lower thresholds and looser identification criteria are run without prescale, most notably `HLT_e15_lhloose_nod0_L1EM12` and `HLT_mu14`

A more detailed description of the low-μ data condition can be found in [23].

4.2 MC Samples

Fully simulated and reconstructed ATLAS Monte Carlo samples ("MC") from the MC16 production campaigns are used to model the signal and background processes. All MC samples use settings specific to the special data run conditions, i.e. specifically the pileup overlay, topo-cluster noise thresholds and trigger menu are adapted. Given that the pileup distribution is already adjusted as to the dataset, no further μ-reweighting is performed. The main event samples generated and simulated ("HITS") for the main ATLAS MC16d campaign are reused and processed further with dedicated settings in digitisation and reconstruction. The simulation and reconstruction tags for these samples are s3126_r10244_r10210. These samples correspond to the 2017 conditions except for the bunch structure of the beam, which was "8b4e" in data and the more "standard" structure in the MC.

Process	Data set	Generator	σ·BR·ϵ_{filter} [nb] (th. unc.)	$N_{\text{evt}}^{\text{skim}}$ [10^6]	$N_{\text{evt}}^{\text{unskim}}$ [10^6]
$W^+ \to e^+\nu$	361100	Powheg+Pythia8	11.61 (5%)	40	40
$W^+ \to \mu^+\nu$	361101	Powheg+Pythia8	11.61 (5%)	40	40
$W^+ \to \tau^+\nu$	361102	Powheg+Pythia8	11.61 (5%)	0.28	5.0
$W^- \to e^-\overline{\nu}$	361103	Powheg+Pythia8	8.630 (5%)	30	30
$W^+ \to e^+\nu$	361104	Powheg+Pythia8	8.630 (5%)	29	29
$W^- \to \tau^-\overline{\nu}$	361105	Powheg+Pythia8	8.630 (5%)	0.24	4.0
$Z \to ee$	361106	Powheg+Pythia8	1.910×1.025 (5%)	10	10
$Z \to \mu\mu$	361107	Powheg+Pythia8	1.910×1.025 (5%)	10	10
$Z \to \tau\tau$	361108	Powheg+Pythia8	1.910×1.025 (5%)	0.12	1.0
$ZZ(q\overline{q}\ell\ell)$	363356	Sherpa 2.2.1	0.01556×0.141 (10%)	0.0064	0.010
$WZ(q\overline{q}\ell\ell)$	363358	Sherpa 2.2.1	0.003433 (10%)	0.0063	0.010
$WW(q\overline{q}\ell\nu)$	363359	Sherpa 2.2.1	0.02472 (10%)	0.0093	0.020
$WW(\ell\nu q\overline{q})$	363360	Sherpa 2.2.1	0.02472 (10%)	0.0093	0.020
$WZ(\ell\nu q\overline{q})$	363489	Sherpa 2.2.1	0.01142 (10%)	0.0047	0.010
$ZZ(4\ell)$	364250	Sherpa 2.2.2	0.001252 (10%)	0.0057	0.010
$WZ(3\ell\nu)$	364253	Sherpa 2.2.2	0.004583 (10%)	0.0062	0.010
$WW(2\ell 2\nu)$	364254	Sherpa 2.2.2	0.01250 (10%)	0.0073	0.010
$WZ(\ell 3\nu)$	364255	Sherpa 2.2.2	0.003235 (10%)	0.0050	0.010
Wt	410013	Powheg+Pythia8	0.03582 (10%)	0.0037	0.010
$W\bar{t}$	410014	Powheg+Pythia8	0.03399 (10%)	0.0037	0.010
$t\bar{t}$ (nominal)	410470	Powheg+Pythia8	0.8318×0.544 (7%)	1.2	2.0
$t(\text{t}-\text{chan.}t)$	410642	Powheg+Pythia8	0.03699 (10%)	0.016	0.030
$t(\text{t}-\text{chan.}\bar{t})$	410643	Powheg+Pythia8	0.02217 (10%)	0.011	0.020
$t(\text{s}-\text{chan.}t)$	410644	Powheg+Pythia8	0.002027 (10%)	0.0050	0.010
$t(\text{s}-\text{chan.}\bar{t})$	410645	Powheg+Pythia8	0.001268 (10%)	0.0052	0.010
$t\bar{t}$ (syst.)	410480	Powheg+Pythia8	0.8318×0.438 (7%)	0.85	1.5
$t\bar{t}$ (syst.)	410482	Powheg+Pythia8	0.8318×0.105 (7%)	0.40	0.50
$t\bar{t}$ (syst.)	410557	Powheg+Pythia8	0.8318×0.438 (7%)	0.85	1.5
$t\bar{t}$ (syst.)	410558	Powheg+Pythia8	0.8318×0.105 (7%)	0.40	0.50

Table 2: Monte Carlo samples at $\sqrt{s}$ = 13TeV. Given is a short description of the process, the ATLAS MC data set number (DSID), the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies (σ·BR·ϵ_{filter}) with the theoretical uncertainty in percent ("th. unc."), and finally the number of events analysed after skimming at derivation production ($N_{\text{evt}}^{\text{skim}}$) as well as the number of events originally processed and simulated ($N_{\text{evt}}^{\text{unskim}}$). In the case of $Z \to \ell\ell$ samples, the given $\epsilon_{\text{filter}} > 1$ is related to the fact, that the cross sections were calculated for $66 < m_{\ell\ell} < 116$GeV, but the generated mass range is larger. The last section of $t\bar{t}$ samples refers to variation samples for systematics studies. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}}/(\sigma\text{·BR·}\epsilon_{\text{filter}})$ is generally above 3fb^{-1} for signal and significant backgrounds, the exception are Powheg $W \to \tau\nu$ and $Z \to \tau\tau$ samples, that have about 0.45fb^{-1} only.

Table 2 summarizes the MC samples used in this analysis for signal templates as well as for EW+top background modeling. The cross-sections quoted in the table are used to normalize predicted event counts for the various samples and the respective uncertainties are listed as well.

The main signal event samples for W production are generated using the PowHEGevent generator [24–27] using CT10 [28] PDF, interfaced to Pythia8[29] using the AZNLO tune [30]. These Powheg+Pythia8samples are interfaced to Photos++ [31] to simulate the effect of final state QED radiation.

The effect of multiple interactions per bunch crossing ("pileup") is modelled by overlaying simulated minimum bias events over the original hard-scattering event. The minimum bias events were generated with Pythia8using the NNPDF2.3LO set of PDFs [32] and the A3 tune [33], which improves on the previous A2 tune [34] by using some Run2 results at $\sqrt{s}$ = 13 TeV and the Donnachie–Landshoff diffractive model that better describes the "visible" inelastic cross section compared to the default Schuler–Sjöstrand model.

For W processes, the samples are normalized to NNLO calculations performed using the DYTURBO program, an optimised version of DYNNLO [35, 36] using the MMHT2014nnlo PDF set [37]. The numerical values are taken from the corresponding ATLAS publications of the 2015 data at 13 TeV [38]. The uncertainties on those cross-sections arise from the choice of PDF set and PDF set internal uncertainties (~ 3 - 4%), from factorization and renormalisation scale dependence (¡ 1%), and the strong coupling constant α_s (~ 1 - 2%). A conservative total uncertainty of 5% is taken as an uncertainty on any event count predictions normalized using these cross-sections, e.g. for use of subtraction of electroweak backgrounds.

Backgrounds from top-quark pair-production $t\bar{t}$ as well as single-top production (Wt, t-channel, s-channel) are generated with Powheg+Pythia8. At $\sqrt{s}$ = 13 TeV samples that can be used to derive systematics on the $t\bar{t}$ background were produced: the "rad-low" variation can be obtained by internal weights, the "rad-up" variation is obtained by dedicated samples. Various combinations of di-bosons VV,V = W, Z are generated with Sherpa in all decay channels with at least one real lepton in the final state, for details see Table 2.

4.3 Event Selection and object definition

In this analysis, electrons or muons are considered as leptons. Electrons are reconstructed from energy clusters in the calorimeter associated to a track, while muons are reconstructed from a track that crosses both the inner detector and the muon chambers.

The object selection for this analysis follows the selection described in [39], however several studies were performed in order to maximise the sensitivity to the angular coefficients discussed in Appendix A.

$W \to \ell\nu$ candidates events are selected using the following criteria:

- Single electron (*HLT_e15_lhloose_nod0_L1EM12*) and muon (*HLT_mu14*) triggers were used for the electron and muon channels;
- Exactly one lepton in the event;
- Both electrons and muons are required to have $p_T > 25$ GeV;
- Electrons should have $|\eta| < 2.47$ and candidates in the crack region $1.37 < |\eta| < 1.52$ are excluded. Muons candidates are required to be within $|\eta| < 2.4$;

- The electron candidates must satisfy the *Tight LH* identification criteria. The muon candidates must satisfy the *Medium* identification criteria;
- All leptons are required to be isolated. We define lepton as isolated if it satisfies both isolation criteria: $ptvarcone20/p_T < 0.1$ and $topoetcone20/p_T < 0.05$;
- Leptons are required to originate from the primary vertex. The longitudinal impact parameter of each lepton track is required to be less than 0.5 mm ($\Delta z_0 \times \sin\theta < 0.5$ mm); Furthermore, the significance of the transverse impact parameter divided by its estimated uncertainty is required to satisfy for electrons $|d_0|/\sigma(d_0) < 5$ and for muons $|d_0|/\sigma(d_0) < 3$.

Dedicated lepton calibrations and efficiency corrections obtained specifically for the low-μ conditions are applied to the data and the MC and described in details in References [40, 41] for electrons and muons, respectively.

The electron reconstruction scale factors are obtained by extrapolation of the standard high-pileup SFs to the low-μ regime, and apply to both the 13 and 5 TeV datasets. The identification scale factors are measured in-situ separately for the 13 and 5 TeV data. Isolation and trigger efficiencies and SF are measured in-situ separately for the 13 and 5 TeV data sets. For the electron energy calibration, the global strategy follows the methodology applied in high-μ standard calibration and described in Ref. [42]. Electron energy scale and resolution corrections are measured in-situ using $Z \to ee$ events from the low pile-up dataset.

Muon reconstruction and *track-to-vertex association* requirements efficiencies measured insitu with $Z \to \mu\mu$ events are found to be compatible with the corresponding high-μ measurements and are measured to permille-level precision. The muon trigger and isolation efficiencies are measured insitu in the low-μ data, for each dataset, to typically better than $< 1\%$ precision, limited by the size of the $Z \to \mu\mu$ data samples. Finally, the muon calibration derived from the high-μ data has been cross-checked using both $Z \to \mu\mu$ and $J/\psi \to \mu^+\mu^-$ events in the low-μ datasets and has been found to be adequate at the current level of systematics. A dedicated correction for charge-dependent momentum biases, so-called sagitta correction, has been derived using a combination of different methods using the 13 TeV low-μ dataset.

Additional efficiency corrections were calculated specifically for this analysis as discussed in Appendix H for the *Tight LH* identification criteria applied to the electrons candidates and the additional isolation requirement ($topoetcone20/p_T < 0.05$) for both electrons and muons.

4.4 Neutrino Reconstruction

In the context of W boson production a high precision computation of the neutrino kinematics is mandatory, and this is done with the help of the *hadronic recoil*. In proton-proton collisions there is a non-zero transverse momentum for vector boson production that can be thought of to originate from initial state gluon and quark radiation:

$$\vec{p}_T(W/Z) = \vec{p}_T^{\,\text{lepton1}} + \vec{p}_T^{\,\text{lepton2}} = -\sum \vec{p}_T^{\,\text{ISRquark,gluons}} = -\vec{u}_T \,. \tag{29}$$

Here $p_T(W/Z)$ denotes the transverse momentum of the W or Z boson and p_T^{lepton} denote the transverse momenta decay leptons (charged lepton or neutrino). The *hadronic recoil*, $\sum \vec{p}_T^{\,\text{ISRquark,gluons}}$, is the quantity which accounts for the transverse momenta any partons recoiling from the colour-neutral boson, and is denoted as $\vec{u_T}$. In the contest of low pileup precision W analysis the use of *particle flow objects* (PFOs) as input constituents to the building of u_T, thus using the full available information on both topo-clusters

and tracks, represents a major improvement. The hadronic recoil in the transverse plane, $\vec{u_T}$, is computed as the vectorial sum of all charged and neutral PFOs. The modulus of this 2-vector is simply termed u_T. A critical quantity is ΣE_{T} that represents the total event activity and is related to the resolution of the u_T measurement. ΣE_{T} is defined as the scalar sum of the p_{T} of all PFOs. Naturally ΣE_{T} will be strongly dependent on the vector boson dynamics and grow with $p_{\mathrm{T}}^{V} = u_T$. To disentangle these effects, an even more important quantity is defined by $\Sigma \bar{E_T} = \Sigma E_{\mathrm{T}} - u_T$. $\Sigma \bar{E_T}$ may be thought of the event activity corrected for the "directed" recoil activity and thus represents the activity from underlying event, pileup, and emissions beyond first (hard) emission. [1]

The detailed description of the *hadronic recoil* definition and calibration is done in reference [43] and also summarised briefly here. The calibration is obtained as a function of $\Sigma \bar{E_T}$ and p_{T}^{V} of the boson, since the hadronic recoil is mainly sensitive to these. The correction is obtained in Z events and extrapolated to the W events.

The procedure consists of three steps :

- First, the $\Sigma \bar{E_T}$ distribution in the Monte-Carlo should be well modelled and match that of the data. More precisely, it is crucial to model correctly the correlation between $\Sigma \bar{E_T}$ and p_{T}^{V}, since we want to have a good description of the activity as a function of our measured physics observable. This is achieved thanks to a 2-dimensional reweighting, obtained in Z events. In the simulated W events, an additional reweighting of $\Sigma \bar{E_T}$ in bins of u_T is applied. A further 1-dimensional reweighting is obtained for each process (W^-, W^+ and Z) to recover the initial underlying $p_{\mathrm{T}}^{\mathrm{true}}$ spectrum.
- Second, the direction of the recoil is corrected, taking the projections on x and y axes of the recoil in Z events, and correcting for any data to MC differences.
- Finally, response and resolution corrections are, once again, obtained in-situ in Z data events, where the parallel and perpendicular components can be extracted in the data as a function of $\Sigma \bar{E_T}$ and $p_{\mathrm{T}}^{\ell\ell}$, and compared to the Monte-Carlo to extract corrective coefficients.

This Z boson based calibration is applied to W events ; uncertainties due to this extrapolation are included. A summary on these uncertainties, as well as their impact on the unfolded spectra, are discussed in section 7.

Once that the *hadronic recoil* is computed it is possible to determine indirectly the neutrino ν transverse momentum using the following expression:

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} := \vec{p}_{\mathrm{T}}{}^{\nu} = -(\vec{u_T} + \vec{p}_{\mathrm{T}}{}^{\ell}) \tag{30}$$

where $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is the missing transverse energy. The transverse momenta of additional muons with $p_{\mathrm{T}} > 5$ GeV and satisfying the medium identification is added at the definition of $E_{\mathrm{T}}^{\mathrm{miss}}$ while their calorimeter energy deposit is subtracted from the event.

Thanks to the clean environment guarantee by the low pileup dataset, in this analysis it was possible to maximise the sensitivity and enlarge the acceptance at high p_{T}^{W} without applying any requirements on m_T^W [2] and E_T^{miss}.

[1] E.g. a $V + 1$ jet event has $\Sigma \bar{E_T} \sim 0$ on parton level for any jet p_{T} and thus a finite $\Sigma \bar{E_T} > 0$ is due to additional activity.

[2] defined as:

$$m_T = \sqrt{2 p_T^{\nu} p_T^{l} (1 - \cos \Delta\phi^{\nu})} \tag{31}$$

However the unmeasured longitudinal momentum of the neutrino is inferred by assuming that the dilepton pair of the Drell-Yan process has invariant mass equal to m_W. A massless charged lepton with four-momentum p^μ and a massless neutrino with four-momentum q^μ are considered in the laboratory frame:

$$p^\mu = \begin{pmatrix} \vec{p} \\ |p| \end{pmatrix} \qquad q^\mu = \begin{pmatrix} \vec{q} \\ |q| \end{pmatrix} \tag{32}$$

$$\vec{p} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \qquad \vec{q} = \begin{pmatrix} q_x \\ q_y \\ q_z \end{pmatrix} \tag{33}$$

By imposing the m_W constrain:

$$(p+q)^\mu (p+q)_\mu = m_W^2 \tag{34}$$

and using the relations:

$$|p| = \sqrt{p_x^2 + p_y^2 + p_z^2} \qquad |q| = \sqrt{q_x^2 + q_y^2 + q_z^2} \tag{35}$$

and factorizing the expression in powers of q_z, leads to:

$$a q_z^2 + b q_z + c = 0 \tag{36}$$

with

$$a = (p_x^2 + p_y^2) \tag{37}$$

$$b = -p_z[2(p_x q_x + p_y q_y) + m_W^2] \tag{38}$$

$$c = |p|^2(q_x^2 + q_y^2) - [2(p_x q_x + p_y q_y) + \frac{m_W^2}{2}]^2 \tag{39}$$

Finally, it is possible to write the two solutions for longitudinal momentum of the neutrino as:

$$q_z(1,2) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{40}$$

The lepton momentum in the Collins-Soper framed can be expressed as [1]:

$$p_x^{CS} = \frac{1}{2}\frac{m_W}{m_W^2 + p_{T,W}^2}(2p_x - p_{T,W}^2) \tag{41}$$

$$p_y^{CS} = p_y \tag{42}$$

$$p_z^{CS} = \pm\frac{m_W}{2}\sqrt{1 - \frac{\left(p_x^{CS}\right)^2 + \left(p_y^{CS}\right)^2}{m_W^2/4}} \tag{43}$$

and as in Eq. (30), the neutrino p_{T}^{ν} and ϕ^{ν} are estimated by the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$ quantities.

The azimuthal and polar angles in the Collins-Soper frame are

$$\phi_{CS} = \arctan p_y^{CS}/p_x^{CS} \tag{44}$$

$$\theta_{CS} = \arctan \sqrt{\left(p_x^{CS}\right)^2 + \left(p_y^{CS}\right)^2}/p_z^{CS} \tag{45}$$

The transverse components p_x^{CS} and p_y^{CS} are uniquely determined from measurable laboratory frame quantities. On the other hand, the longitudinal component p_z^{CS} is determined only up to a sign. The twofold ambiguity in the reconstruction of the leptons momenta in the Collins-Soper frame translates in an ambiguity on the sign of $\cos\theta_{CS}$, while ϕ_{CS} is determined. In general, the twofold ambiguity related to the unmeasured longitudinal momentum of the neutrino results in two solutions for the rapidity of the W boson, y_W, and two solutions for $\cos\theta_{CS}$ which differ by a sign, while $p_{T,W}$ and ϕ_{CS} are uniquely determined.

The above formulas assume that the z axis is collinear to the direction of one of the collider beams. However in proton-proton collisions the sign of the z axis in the Collins-Soper frame is taken along the direction of the longitudinal momentum of W, i.e, we have $\cos\theta_{CS} \to \cos\theta_{CS}\mathrm{sign}(y_W)$ (see Eq. 5). With this definition, there are events for which the two solutions have opposite sign for y_W, hence $\cos\theta_{CS}$ is uniquely determined. These events, which are usually at central rapidities, allow to resolve the sign of $\cos\theta_{CS}$ in a fraction of the dataset, and open the possibility to measure the forward-backward asymmetry (the A4 coefficient) in W-boson production. More details on this subject are given in Appendix B.

5 Multi-Jet Estimate

Two main categories of backgrounds affect single W boson final states : the electroweak (single and diboson) and top backgrounds, obtained from the appropriate MC samples, and the the so-called multi-jet (MJ) background which can't be reliably simulated and has to be derived from data. Depending on the lepton flavour, the MJ background has significant contributions from real leptons produced in semi-leptonic decays of heavy quarks, pion decays, or photon conversions.

The data sample discussed in this analysis has a pile-up of $\mu \sim 2$, about a factor of 15–20 smaller than the standard 13 TeV running conditions. Lower pile-up will reduce the MJ background yield for at least three reasons:

- a reduced pile-up contribution to the isolation distributions, improving the discriminating power of the isolation cuts;
- better $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_T resolution, with correspondingly improved discriminating power of the event selection cuts;
- reduced contribution from pile-up jets themselves to the contamination of the sample.

The fake lepton background (called in the following multijet background or simply MJ) is concentrated at lower values of p_T^{ℓ}, E_T^{miss} and m_T. By relaxing the kinematic cuts in our signal region definition we introduce significant amount of multijet background. The background estimation method relies on the same methodology described in the *WpT* analysis (see [44]), however there are number of W-Ai analysis specific details that are listed in the following.

5.1 General procedure

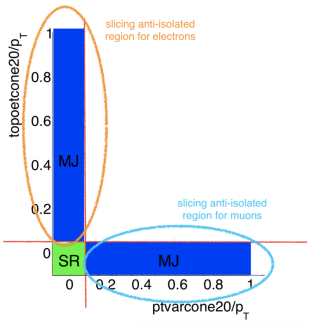
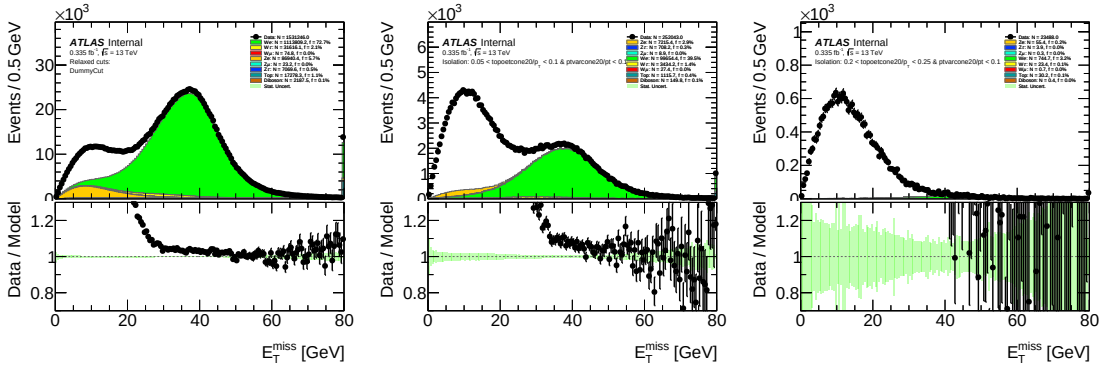


Figure 10: Sketch of the definition of the SR and CR1 for the multi jet estimate

The aim of the procedure is to estimate the multijet background yield in the signal region (SR). The multijet yields estimation relies on a fit on data using signal and MJ *templates* built from distributions as the p_T^{ℓ}, E_T^{miss} and m_T . As the W-Ai analysis is measuring angles, additionally to p_T^{ℓ}, E_T^{miss} and m_T variables, the $|\Delta\phi(\ell, E_T^{miss})|$ variable was used as well in the fit. For the signal, EWK and Top backgrounds those distributions are obtained using simulation, and include calibration corrections. The corresponding MJ background *templates*, are obtained from the data in a separate control region where we applying the same kinematic selections but inverting the isolation cuts. The choice of lepton isolation definition to define the SR and the CR1 for electron and muon channel is sketch in Figure 10 and explained in Section 4.3.

The definition of the event categories used for the MJ extraction are defined following naming convention in Ref. [44]:

- fit region (FR): which correspond to the analysis signal region (SR), with one isolated lepton ($ptvarcone20/p_T < 0.1$ and $topoetcone20/p_T < 0.05$), but all distributions includes overflow bins;
- control region 1 (CR1): Multijet enriched CR where the lepton is required to be anti-isolated ($topoetcone20/pT < 0.05$ and $ptvarcone20/p_T > 0.1$ for muons or $ptvarcone20/pT < 0.1$ and $topoetcone20/p_T > 0.05$ for electrons) all distributions includes overflow bins;



(a) (b) (c)

Figure 11: The E_T^{miss} distribution for $W^+ \rightarrow e^+\nu$ candidates in the fit region (a) and in the CR1 for first anti-isolated slice $0.05 < topoetcone20/p_T < 0.1$ (b) and anti-isolated slice $0.2 < topoetcone20/p_T < 0.25$ (c). The difference between data and simulation is attributed to the missing MJ contribution.

As example, the E_T^{miss} distribution for data in the FR is compared only with the signal and the electroweak and top backgrounds expectation form MC is illustrated in Figure 11(a). The difference between data and the MC simulation in the fit region is expected to be due to multijet background contribution. CR1 by construction is dominated by non isolated object which characterised multijet production, and serves to determine the multijet template distribution used in the FR. Appendix E.1 describe in details the correction applied on the definition of the recoil reconstruction in the CR1 as also described in Ref. [44]. As it is explained in the Appendix E.2 in this analysis the CR1 are built scanning only one isolation requirement at the time.[3]

The MJ template used in the FR is extracted from CR1 while subtracting from the data the contribution of signal, electroweak and top processes as simulated by MC. Figure 11(b) shown E_T^{miss} distribution for data and the MC in the CR1 where the lepton is required to to have $0.05 < topoetcone20/p_T < 0.1$ and Figure 11(c) shown E_T^{miss} distribution for data and the MC in the CR1 defined with a lepton to have $0.2 < topoetcone20/p_T < 0.25$. In Figure 11(b) the fraction of the electroweak and top related process contribution is 44.6%, and the fraction is drastically reduced in Figure 11(c), however in this slice also the data statistic is much lower.

To study the dependence of the multijet templates on the anti-isolation criterion used to define CR1, the CR1 is split into successive intervals (slice) of the isolation variable, instead of a single, integrated anti-isolated

[3] Only track-based isolation is relaxed and sliced for $\mu^{\pm}$ channels and for $e^{\pm}$ channels only calo-based isolation is relaxed and sliced (as displayed in Figure 10).

region. The isolation slices used to define $CR1(i)$ to extract the MJ templates are defined as the following intervals of lepton isolation :

- For $W \rightarrow \mu\nu$ channels $ptvarcone20/pT$ in $[0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5]$ and $topoetcone20/pT < 0.05$
- For $W \rightarrow e\nu$ channels $topoetcone20/pT$ in $[0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45]$ and $ptvarcone20/pT < 0.1$

Once the $CR1(i)$ slices are defined, the extraction of the number of multijet background events in the FR proceeds as follows, for a given isolation slice i:

- The multijet background template in each anti-isolation slices (also called bin), $N_{MJ}^{CR1(i)}$, is derived as follows:
$$N_{MJ}^{CR1(i)} = N_{data}^{CR1(i)} - N_{EW}^{CR1(i)} \tag{46}$$
where $N_{MJ}^{CR1(i)}$ is obtained from each one of the kinematic distributions used in the fit, p_T^{ℓ}, E_T^{miss}, m_T or $|\Delta\phi(\ell, E_T^{miss})|$.
- The fraction fit is performed for p_T^{ℓ}, E_T^{miss}, m_T or $|\Delta\phi(\ell, E_T^{miss})|$ distributions in the FR for each multijet template obtained in each anti-isolation slices. Figure 12 shows as en example the E_T^{miss} distribution in the FR, after fitting the multijet templates, $N_{MJ}^{CR1(i)}$, derived from different anti-isolation slices in CR1(i). It's written as:
$$N_{data}^{FR} = \alpha \cdot N_{MC}^{FR} + T(i) \cdot N_{MJ}^{CR1(i)}. \tag{47}$$
The fitting parameter $T(i)$ gives the estimated number of MJ events in SR: $N_{MJ}^{SR} \approx T(i) \cdot N_{MJ}^{CR1(i)}$. An overall normalisation factor for the signal and EW+top contribution (N_{MC}^{FR}), α, in this analysis is fixed to 1 during the fit.

In Eqs. 47 the multijet distributions in the fitting region are approximated by their anti-isolated counterparts from CR1(i). The residual dependence of the multijet background estimate on the anti-isolation criteria choose to built the template, introduces a systematic effect in the procedure. To reduce this dependence the MJ yields extracted in the FR using the different anti-isolation slices are used to extrapolate the final MJ yields in the signal region at very low isolation values. The final MJ background estimation is therefore extrapolated toward the SR for both the total yield and the template shape. Section 5.2 describe the final MJ yield extraction, and Section 5.3 the MJ template shape extrapolation.

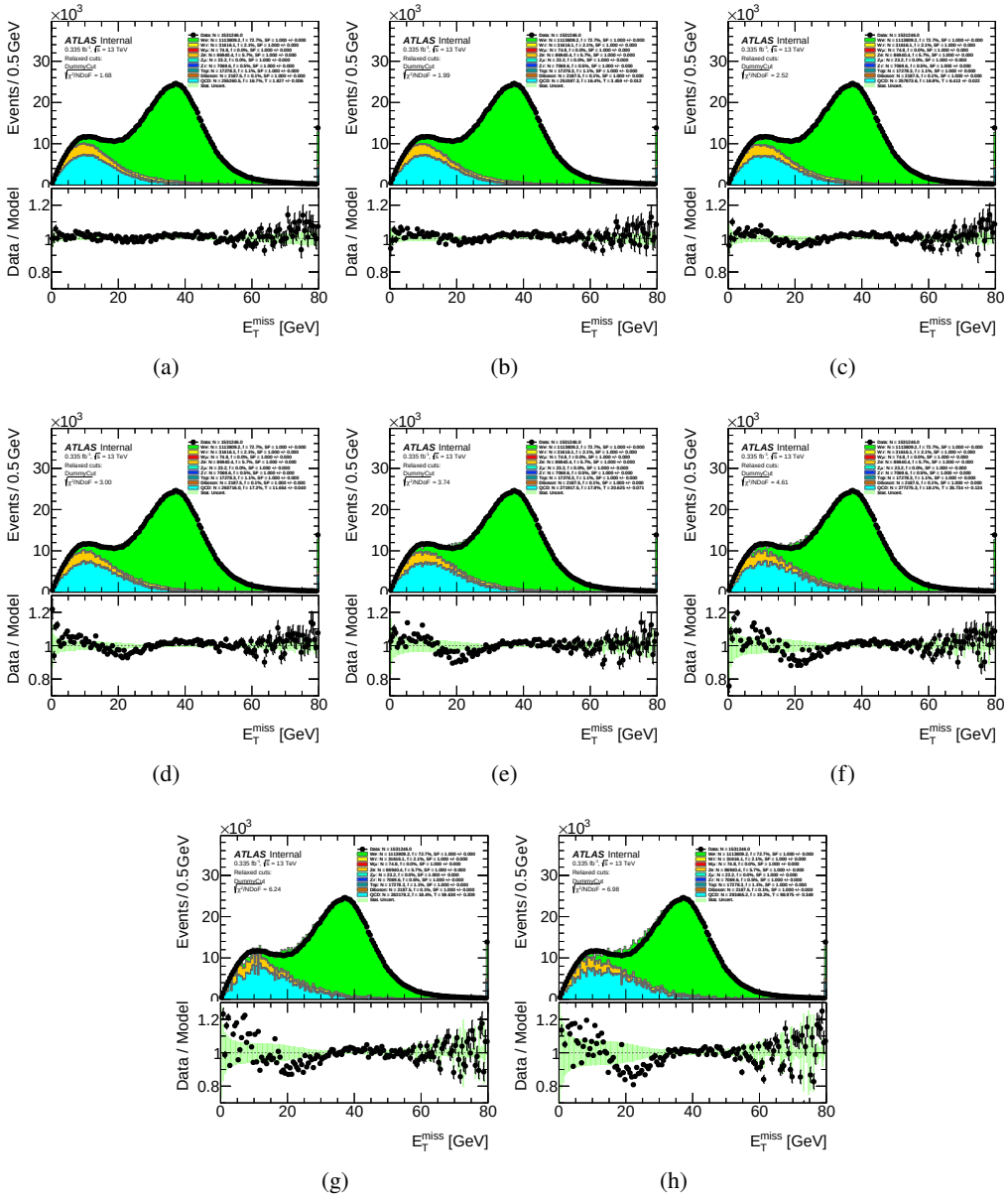


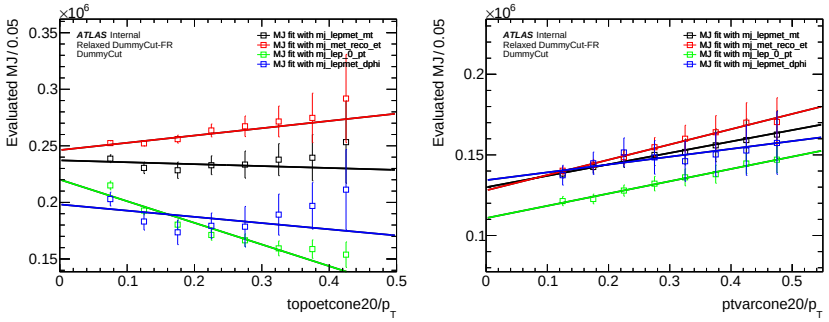
Figure 12: The multijet template distributions used in the fraction fit for $W^+ \to e^+\nu$ channel. The MJ distributions are derived in CR1 and fitted to the data in FR and indicated by light-blue color. 8 control region slices with different $topoetcone20/p_T$ cuts are constrained from (a) $0.05 < topoetcone20/p_T < 0.1$ to (h) $0.4 < topoetcone20/p_T < 0.45$.

5.2 MJ yield normalization

The total MJ yield in the analysis SR is evaluated by extrapolating the number of MJ events extracted from the fraction fit using different MJ template build from different anti-isolated bin towards the isolated regions. In the FR the MJ yield is extracted performing the fraction fit using Eq. 47 for each anti-isolation template $N_{MJ}^{CR1(i)}$ in different discriminating distributions. The MJ yields extracted using differnet anti-isolation slice for all the discriminating variable (p_T, m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$) are shown in Figure 13. The

error on each measured MJ yields is multiplied by $\sqrt{\chi^2/NDoF}$ resulting from the fraction fit to account for possible mismodelling in the template of the considered variables.

A linear fit is used to extrapolate the MJ yield extracted using different anti-isolation slices to the signal region (SR) separately for each discriminative variables p_T, m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$. The average isolation for the final extrapolation of the MJ background is taken as $X = 0.025$ from $b\bar{b}$ and $c\bar{c}$ samples studies [44]. The final central value of the MJ yield in SR is considered to be the average of the linear extrapolation at $X = 0.025$ of the four observables estimate. Figure 13 presents linear fit used for the the yield extrapolation to SR, for each channel, where p_T, m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$ ideally should converge at $X = 0.025$.



(a) $W^- \to e^-\overline{\nu}$ (b) $W^- \to \mu^-\overline{\nu}$

(c) $W^+ \to e^+\nu$ (d) $W^+ \to \mu^+\nu$

Figure 13: The estimated total number of the multijet background in SR scanned as a function of the isolation slice. Isolation scans in the $W^- \to e^-\overline{\nu}$, $W^- \to \mu^-\overline{\nu}$, $W^+ \to e^+\nu$ and $W^+ \to \mu^+\nu$ with recoil correction. Each point is plotted as the central of each isolation slice and correspond to the result of Eq. 47. Every set of points corresponds to the one of four multijet discriminative variables m_T, E_T^{miss}, p_T, and $|\Delta\phi(\ell, E_T^{miss})|$.

The results for the final yield estimation as well as the uncertainty breakdown in the $W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels is showed in Tab. 3. All the numbers and plots related to this estimation method are presented in the Appendix E.3.

The numbers of the MJ background estimated in the electron and positron channels are similar as expected due to the fake rates of a jet to fake an electron and to fake a positron should be the same. Good compatibility is observed between the two charges. Accounting for the systematics uncertainty the MJ yields in muon and anti-muon channels shows no significant discrepancy.

A closure test has been performed on this method using $b\bar{b} + c\bar{c}$ MC samples as a source of multijet templates in the CR1 passing all kinematic and anti-isolation selection. Results are in consistency with

data-driven method. Details on this are in the Appendix E.4.

5.3 Correction of the multijet background shapes in the signal region

The MJ templates from anti-isolated control regions themselves cannot correctly represent the MJ shape in the isolated regions. To solve this issue, a shape correction is applied to all MJ templates in CR1 according to the prescription below. The MJ shape is defined to be:

$$H_{MJ}^{[0.A,0.B]}[X] = H_{data}^{[0.A,0.B]}[X] - H_{EW+Top}^{[0.A,0.B]}[X] \tag{48}$$

where $H_{MJ}^{[0.A,0.B]}[X]$ is the normalized distribution of each kinematic variable X in $CR1$ after subtraction of the EW and top processes, and satisfying $A < isoCone/p_T < B$, where $isoCone$ is $ptvarcone20$ for $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ and $topoetcone20$ for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$. The shape correction representing the dependence on isolation variable (as example when $isoCone$ is $ptvarcone20$) is calculated via:

$$\begin{aligned}\Delta^{CR1}[X] = \frac{1}{4}\{&\frac{H_{MJ}^{0.1,0.15} - H_{MJ}^{0.3,0.35}}{4} + \frac{H_{MJ}^{0.15,0.2} - H_{MJ}^{0.35,0.4}}{4} + \\ &\frac{H_{MJ}^{0.2,0.25} - H_{MJ}^{0.4,0.45}}{4} + \frac{H_{MJ}^{0.25,0.3} - H_{MJ}^{0.45,0.5}}{4}\}\end{aligned} \tag{49}$$

which helps estimate the MJ shape in SR:

$$H_{MJ}^{SR}[X] = H_{MJ}^{0.1,0.15}[X] + 2 \cdot \Delta^{CR1}[X] \tag{50}$$

The correction term $\Delta[X]$ is purely built upon the MJ shape in 8 anti-isolated slices of CR1. With the above correction, the shapes of MJ templates each CR1(i) are extrapolated to an uniquely determined shape in SR as shown in Figure 14 to 17, which means that the fraction fit is effectively carried out only once for a given kinematic distribution.

Same strategy, but with different binning is used in case when $isoCone$ is $topoetcone20$ for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ channels.

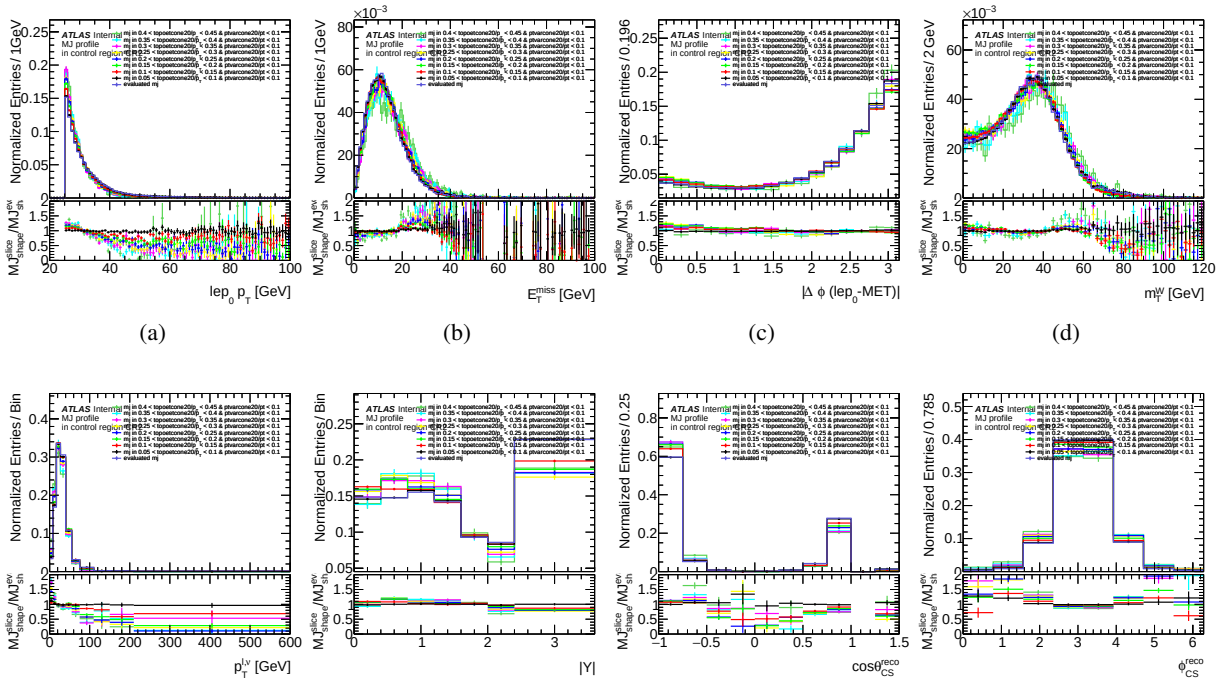


Figure 14: Normalized MJ shapes in anti-isolated slices of CR1 for $W^- \to e^-\overline{\nu}$ channel, along with the shape corrected to SR.

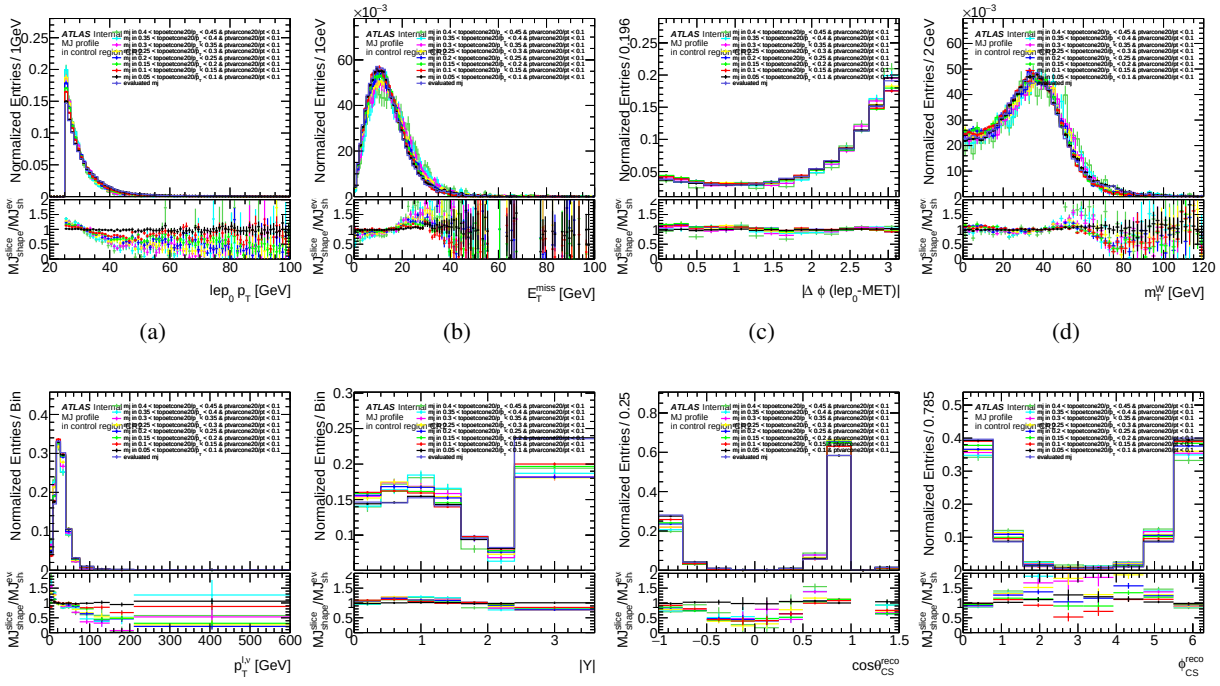


Figure 15: Normalized MJ shapes in anti-isolated slices of CR1 for $W^+ \to e^+\nu$ channel, along with the shape corrected to SR.

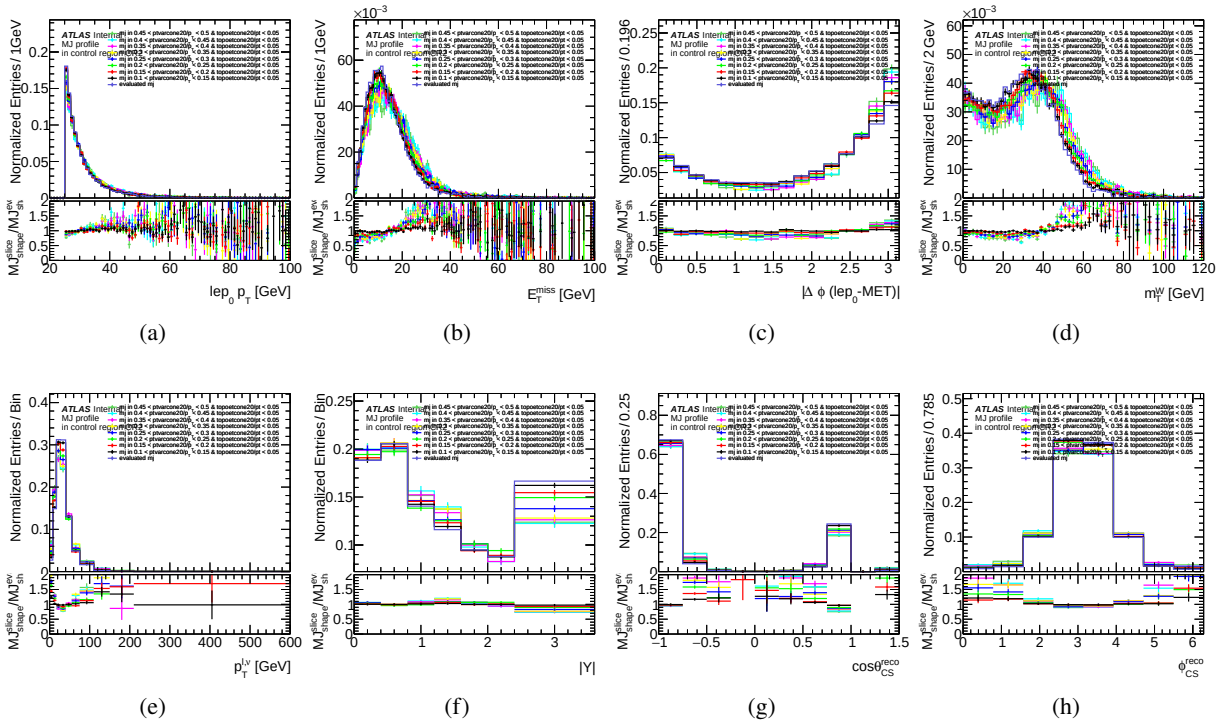


Figure 16: Normalized MJ shapes in anti-isolated slices of CR1 for $W^- \rightarrow \mu^- \overline{\nu}$ channel, along with the shape corrected to SR.

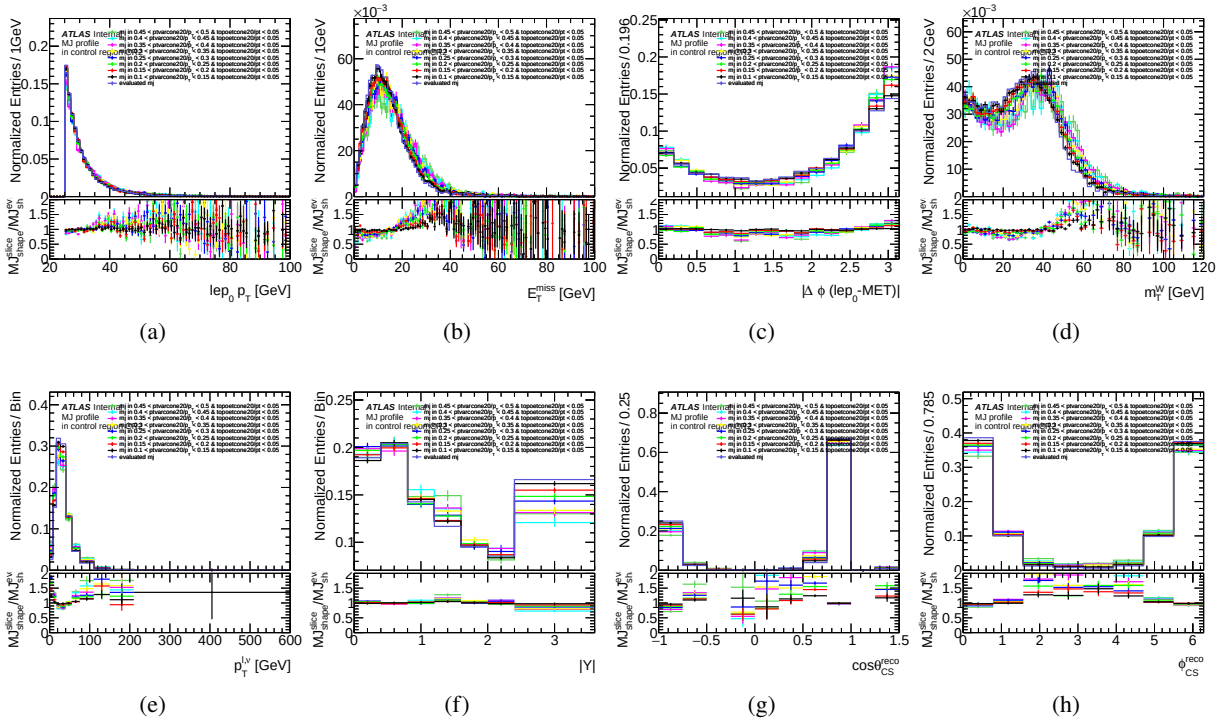
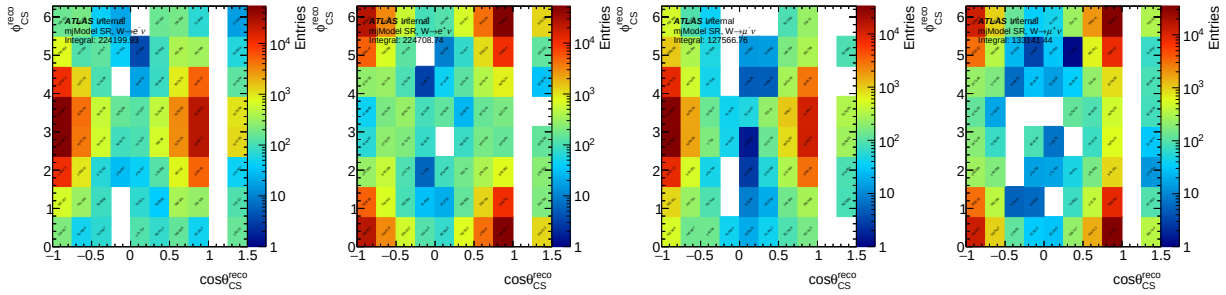


Figure 17: Normalized MJ shapes in anti-isolated slices of CR1 for $W^+ \rightarrow \mu^+ \nu$ channel, along with the shape corrected to SR.

5.4 MJ acceptance correction

Figure 18 shows the final 2D multijet templates as function of $\cos\theta_{CS}$ and ϕ_{CS} distributions normalised to the total multijet yield for the inclusive signal region.



(a) (b) (c) (d)

Figure 18: The multijet templates of the $\cos\theta_{CS}$ vs. ϕ_{CS} distributions for $W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels for the inclusive signal region.

In this analysis MJ templates need to be provided as function of $\cos\theta_{CS}$ and ϕ_{CS} for each $p_T^{\ell,\nu}$ and $|Y|$ measurement bin. As a reminder the extraction of the angular coefficient is done in $p_T^{\ell,\nu}$ and $|Y|$ bins as defined in the following:

$$\begin{aligned} p_T^{\ell,\nu} &: [0, 8, 17, 27, 40, 55, 75, 110, 150, 210, 600]; \\ |Y| &: [0, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 3.6] \end{aligned} \tag{51}$$

Running directly the data-driven method to provide multijet shape and yield estimation for each $p_T^{\ell,\nu}$ or $|Y|$ bin individually would not work due to significant statistics drop in the CR1 regions used to extract MJ templates. However using a single 2D multijet template ($\cos\theta_{CS}^{reco}$ vs. ϕ_{CS}^{reco}) derived from the inclusive SR for all $p_T^{\ell,\nu}$ and $|Y|$ bins wont work because there are strong dependence of the $\cos\theta_{CS}$ multijet shape as function $|Y|$ and ϕ_{CS} as function of $p_T^{\ell,\nu}$ as it is shown in Figure 19. Therefore the 2D $\cos\theta_{CS}$ vs ϕ_{CS} MJ distribution extracted in the inclusive signal region (Figure 18 shows this distributions for $W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels) could be extrapolated in each $p_T^{\ell,\nu}$ and $|Y|$ bins using so called *acceptance correction functions* to account the kinematic changes in the different $p_T^{\ell,\nu}$ and $|Y|$ bins : $[A_j = h_{MC}^{bin_j}/h_{MC}^{Inc.}.]$

To built those *acceptance correction* special background MC samples that are expected to have kinematics similar to MJ kinematics are used. The $b\bar{b}$+$c\bar{c}$ MC samples for $W^\pm \to \mu^\pm\nu$ channels and $JF17$ MC samples for $W^\pm \to e^\pm\nu$ channels are used to this scope (more details in Appendix E.5.1).

To validate this procedure the full MJ extraction was also extracted directly from data in 3 coarse $p_T^{\ell,\nu}$ and $|Y|$ regions define as :

$$\begin{aligned} p_T^{\ell,\nu} &: [0, 17, 55, 600]; \\ |Y| &: [0, 0.8, 2.0, 3.6] \end{aligned} \tag{52}$$

In this corse regions the full data-driven multijet estimation described in Section 5.1 was repeated to estimate the MJ yield and shape. The 2D MJ template extracted form data in corse $p_T^{\ell,\nu}$ and $|Y|$ regions

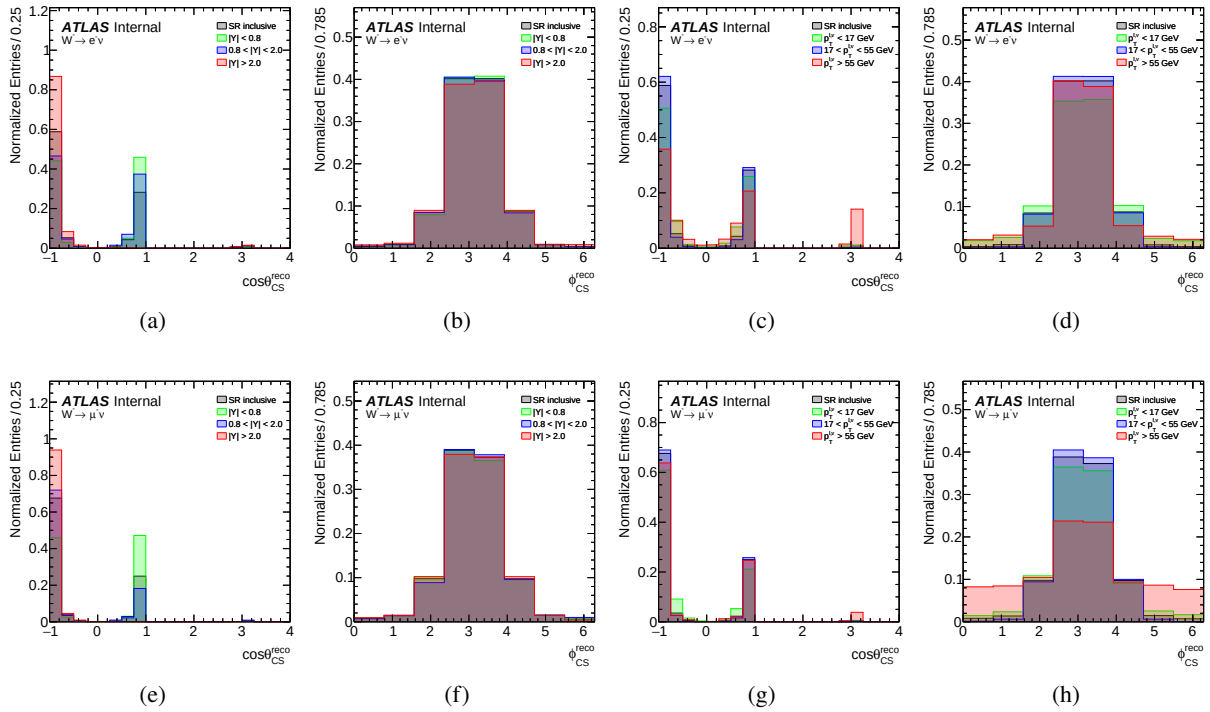


Figure 19: Projections for data-driven multijet templates derived for the inclusive signal region compared with 3 different $p_T^{\ell,\nu}$ or $|Y|$ bins in for $W^- \to e^-\overline{\nu}$ (top) and $W^- \to \mu^-\overline{\nu}$ (bottom) channels.

are compared with the 2D MJ template extracted from the data inclusive region and corrected with the corresponding *acceptance correction* [$B_j = h_{MC}^{bin_k}/h_{MC}^{Inc.}$ where the index k run on the corse region define in Formula 52]. Results of the closure test are shown on Figure 20 for $W^- \to e^-\overline{\nu}$ and $W^- \to \mu^-\overline{\nu}$ channels. As results form this closure test it was verified that in particular for the electron channel extrapolating the data driven inclusive MJ 2D template might results in suboptimal results.

Therefore the extraction of the final 2D multijet background template for a given $p_T^{\ell,\nu}$ or $|Y|$ bin j (define in Eq. 51) proceeds as follows:

- Derive the shape of multijet background events as function of $\cos\theta_{CS}^{reco}$ and ϕ_{CS}^{reco} distribution from data in 3 coarse bin, k, define in Eq. 52 ($h_{MJ,data}^{bin_k}$). The correction on the multijet background shapes described in the Section 5.3 is applied.
- With the dedicated MC samples a multijet template is built for the coarse $|Y|$ or $p_T^{\ell,\nu}$ bins k ($h_{MC}^{bin_k}$) and for the fine j-bin ($h_{MC}^{bin_j}$).
- *Acceptance correction* for each $p_T^{\ell,\nu}$ or $|Y|$ measurement bin are computed in MC as the difference in multijet templates between corse bin k to a given j bin shape : $A_j = h_{MC}^{bin_j}/h_{MC}^{bin_k}$.
- Apply acceptance correction to the data-driven template from the corse bin k to extrapolate to get template shape for a given j bin:

$$h_{MJ}^{bin_j} = h_{MJ,data}^{bin_k} \times A_j \tag{53}$$

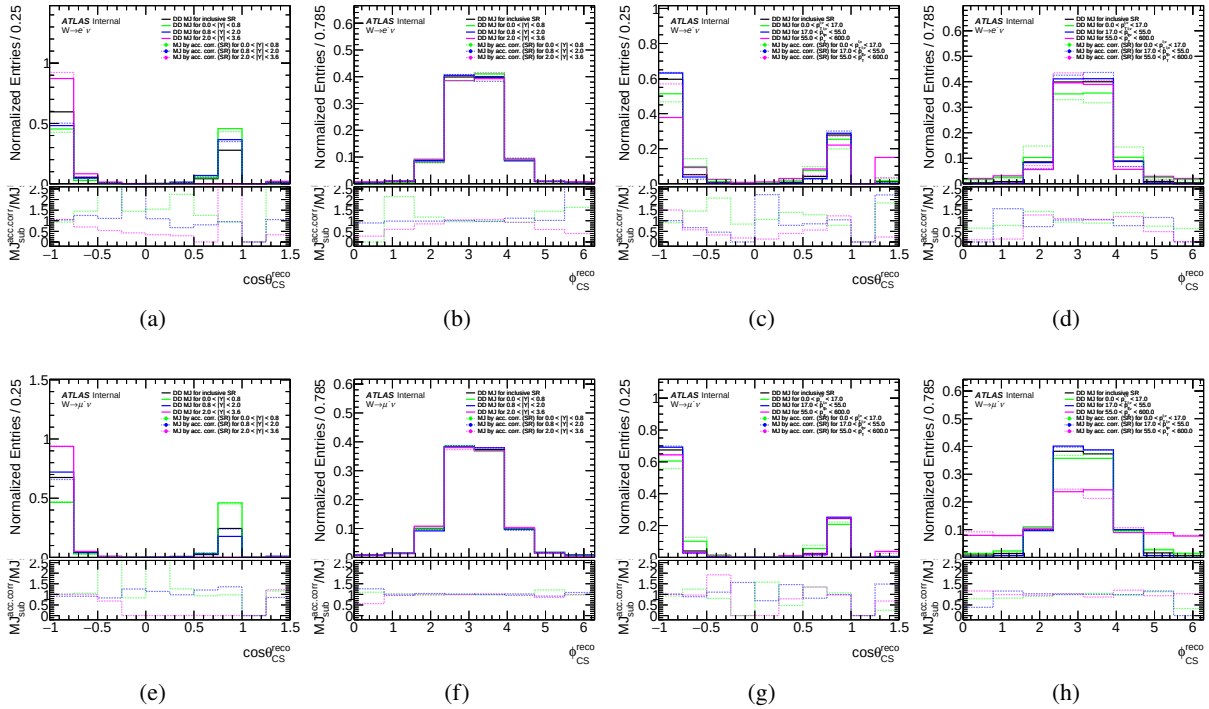
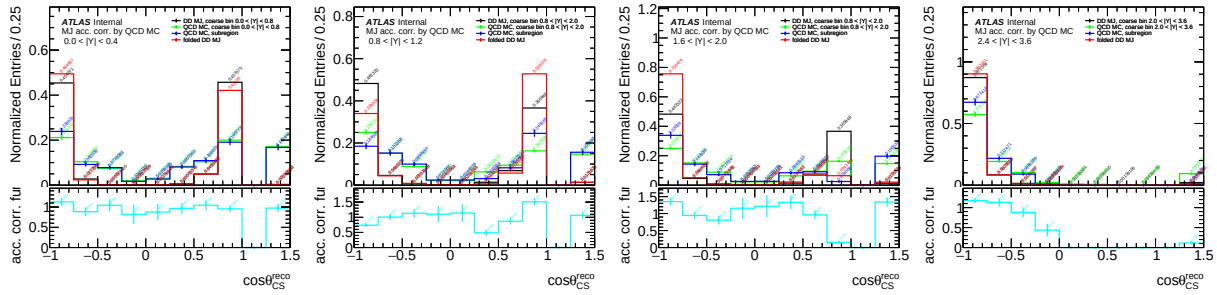


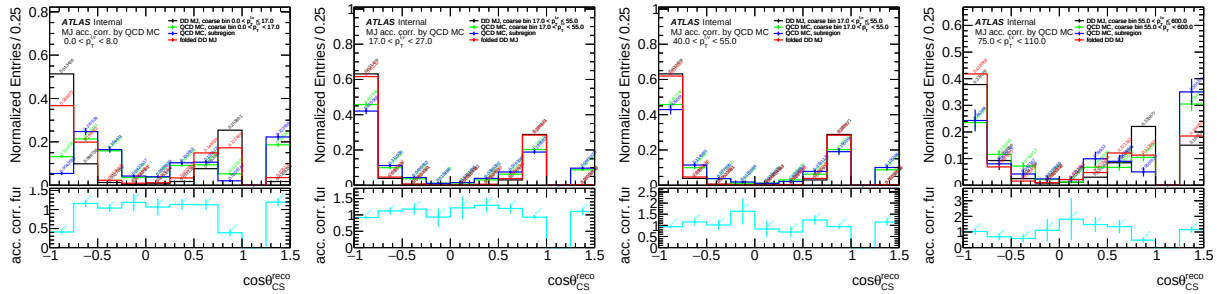
Figure 20: Projections for data-driven multijet templates (solid lines) and templates provided by acceptance correction method (dashed lines) derived for 3 individual subregions for $W^- \to e^-\overline{\nu}$ (top) and $W^- \to \mu^-\overline{\nu}$ (bottom) channels. The ratios between data-driven multijet templates and templates provided by acceptance correction method for each individual coarse subregion are presented on the ratio parts of the figures.

This procedure is shown for selected number of $|Y|$ and $p_T^{\ell,\nu}$ bins on Figures 21-22 for the 1D projection in $\cos\theta_{CS}^{reco}$ and ϕ_{CS}^{reco} distributions for $W^- \to e^-\overline{\nu}$ channel. Figures 23-24 shows same distributions for $W^- \to \mu^-\overline{\nu}$ channel. The 1D distributions are shown for simplicity: the acceptance corrections for the 2D distributions follows the same steps. Complite list of the acceptance corrections for 2D distributions is in the Appendices E.5.6–E.5.9. Tow sources of uncertainty are consider for this acceptance correction. One source of systematic is define as the relative difference between $h_{MC}^{bin_k}$ and $h_{MJ,data}^{bin_k}$ and is taken correlated within the coarse bin k; the other one is the statistical uncertainty of the MC samples template derived for j bins, $h_{MC}^{bin_j}$, which is included into analysis as fully uncorrelated uncertainty on acceptance correction functions.



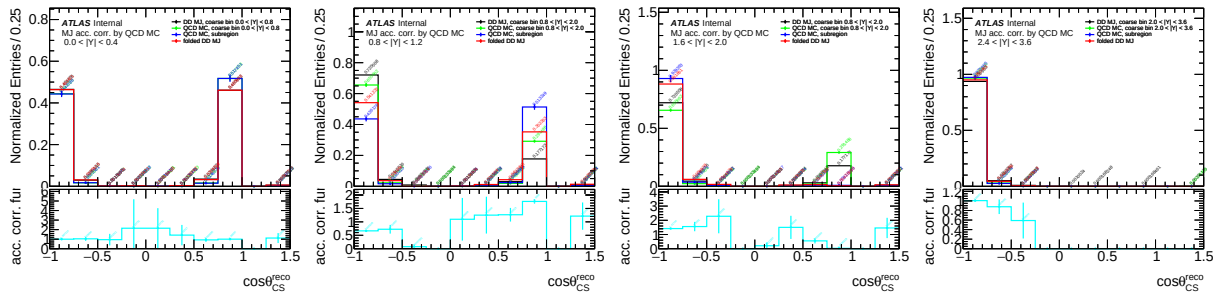
(a) (b) (c) (d)

Figure 21: Normalised multijet background templates for $\cos\phi_{CS}$ distributions (integrated over ϕ_{CS}) derived using acceptance correction functions (red) for some of $|Y|$ bins for $W^- \to e^-\overline{\nu}$ channel. Data-driven multijet templates (black) are derived using corresponting coarse bins (Eq. 52). Multijet acceptance correction functions are shown on the bottom parts of the figures.



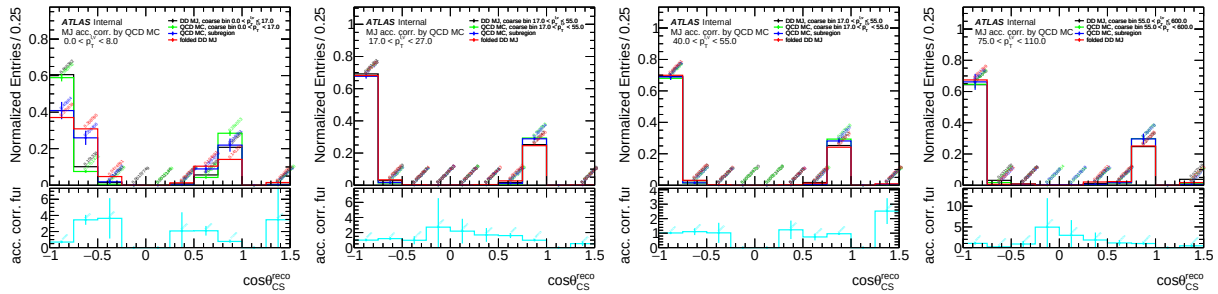
(a) (b) (c) (d)

Figure 22: Normalised multijet background templates for $\cos\phi_{CS}$ distributions (integrated over ϕ_{CS}) derived using acceptance correction functions (red) for some of $p_T^{\ell,\nu}$ bins for $W^- \to e^-\overline{\nu}$ channel. Data-driven multijet templates (black) are derived using corresponting coarse bins (Eq. 52). Multijet acceptance correction functions are shown on the bottom parts of the figures.



(a) (b) (c) (d)

Figure 23: Normalised multijet background templates for $\cos\phi_{CS}$ distributions (integrated over ϕ_{CS}) derived using acceptance correction functions (red) for some of $|Y|$ bins for $W^- \to \mu^-\overline{\nu}$ channel. Data-driven multijet templates (black) are derived using corresponting coarse bins (Eq. 52). Multijet acceptance correction functions are shown on the bottom parts of the figures.



(a) (b) (c) (d)

Figure 24: Normalised multijet background templates for cos ϕ_{CS} distributions (integrated over ϕ_{CS}) derived using acceptance correction functions (red) for some of $p_T^{\ell,\nu}$ bins for $W^- \to \mu^- \overline{\nu}$ channel. Data-driven multijet templates (black) are derived using corresponting coarse bins (Eq. 52). Multijet acceptance correction functions are shown on the bottom parts of the figures.

5.5 Sources of uncertainty in the MJ background estimation

The uncertainty of MJ background estimation arises from 3 major sources: total yield, the shape and statistics. These three components have different correlation patterns: the uncertainty of the total yield as well as the shape uncertainty is fully correlated across the measurement bin, while statistical uncertainty is uncorrelated due to its statistical nature.

The uncertainty breakdown for the final MJ estimation in the inclusive signal region for the $W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels is showed in Tab. 3. The complete brake down of the different sources of uncertainty investigated in this analysis are discussed below.

Channel	$W^- \to e^-\overline{\nu}$	$W^+ \to e^+\nu$	$W^- \to \mu^-\overline{\nu}$	$W^+ \to \mu^+\nu$
Total Number of MJ bkg	**224200**	**224709**	**127567**	**133141**
Luminosity and cross section	±6787 (3.03%)	±8622 (3.84%)	±555 (0.43%)	±702 (0.53%)
Intersection point	±27366 (12.21%)	±28321 (12.60%)	±14888 (11.67%)	±13419 (10.08%)
Extrapolation target	±1241 (0.55%)	±1143 (0.51%)	±1816 (1.42%)	±1601 (1.20%)
Choice of hists	±9122 (4.07%)	±9440 (4.20%)	±4963 (3.89%)	±4473 (3.36%)
Isolation correction	±596 (0.27%)	±813 (0.36%)	±709 (0.56%)	±546 (0.41%)
Correlated Uncertainty	±28197 (12.58%)	±29606 (13.18%)	±14915 (11.69%)	±13449 (10.10%)
Data Stat.	±475 (0.21%)	±448 (0.20%)	±799 (0.63%)	±810 (0.61%)
MC Stat.	±254 (0.11%)	±221 (0.10%)	±994 (0.78%)	±880 (0.66%)
Shape Correction	±1213 (0.54%)	±1527 (0.68%)	±769 (0.60%)	±1006 (0.76%)
Uncorrelated Uncertainty	±1327 (0.59%)	±1606 (0.71%)	±1489 (1.17%)	±1563 (1.17%)

Table 3: MJ background yield and the impact on the uncertainty on the yield due to shape uncertainty breakdown for inclusive signal region in the $W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels.

The following sources of uncertainty are evaluated for the MJ background yield extraction described in Section 5.2:

- Uncertainty due to the luminosity and the cross section of the simulated samples. In several ani-isolated slice in CR1 the contribution from the MC backgrounds is largest . This source of uncertainty is included in the uncertainty of MJ background.
- Intersection point: Uncertainty in the yield caused by the spread of the linear extrapolation as a function of isolation $k^i X + b^i$ using the different discriminating variable i (p_T, m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$). This uncertainty mainly affects the total number of the MJ background, which is estimated with the maximum and minimum $k^i X + b^i$ values. Here X is the isolation value for the central value, which is 0.025. These 4 lines used in the uncertainty sources which could lead to an bad intersection point. This is one of the reason that the statistical uncertainty is ignored from the uncertainty of the extrapolation to avoid double counting.
- Uncertainty from the choice of the extrapolation target, estimated by changing the isolation target from 0.025 to 0.
- Choice of discriminating variable: The uncertainty due to the choice of the distributions used to extract the MJ in the FR eg p_T, m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$, is estimated by removing form the linear extrapolation one of the results from the MJ yield extraction using a discriminative variable. The maximum difference is quoted as the uncertainty.

- Isolation correction: The uncertainty due to the isolation correction is estimated by changing the isolation correction by 10%. This uncertainty is negligible comparing to the statistical uncertainty of the first slice of CR1. So only the contribution to the total multijet number is considered.

The following sources of uncertainty are estimated for the shape of the MJ background:

- Bin-by-bin uncorrelated data and MC statistics: the statistical uncertainty from the limited data and MC samples is taken from the results of shape extrapolation according to Eq. 50.
- Uncertainty due to the shape correction (described in details in 5.3), which mainly affects the shape of the MJ background. Shape uncertainty is defined to be the difference between the extrapolated shape in SR and the MJ shape of the first slice in CR1 (Eq. 49).

As the final output of MJ estimation three sources of uncertainty are propagated to the measurement (Table 5) as combination of uncertainties listed above. The total uncertainty related to the MJ yield extraction and the uncertainty related to the shape of the MJ background are taken as fully correlated across the measurement bin. The systematic uncertainty from the *acceptance correction* which raise from the $b\bar{b} + c\bar{c}$ or $JF17$ MC statistics in the given p_T^ℓ or $|Y|$ bin is taken as fully uncorrelated across the measurement bin and a correlated source across coarse bin k defined in Eq. 52 is also accounted to take care of possible data/MC differences .

6 Control Plots

In Table 4 $W^{\pm} \rightarrow \ell^{\pm}\nu$ signal selection event yields for the13 TeV dataset are shown. The yields for the signal and for the electroweak and top backgrounds are obtained from MC samples. The estimate of the multijet background passing due selection of leptons produced in semi-leptonic decays of heavy quarks, in-flight decays, or misidentification of photon conversions or hadrons, is done performed with data-driven techniques described in Section 5.

Channel	$W \rightarrow e^{-}\nu$	$W \rightarrow e^{+}\nu$	$W \rightarrow \mu^{-}\nu$	$W \rightarrow \mu^{+}\nu$
Wenu	863254.04 ± 267.09 (stat)	1113809.20 ± 304.28 (stat)	3.17 ± 0.52 (stat)	4.51 ± 0.63 (stat)
Wmunu	66.76 ± 2.56 (stat)	74.80 ± 2.88 (stat)	964361.45 ± 284.37 (stat)	1269812.83 ± 323.72 (stat)
Wtaunu	26349.67 ± 146.26 (stat)	31616.08 ± 166.55 (stat)	25398.98 ± 143.14 (stat)	31343.83 ± 165.80 (stat)
Zee	82869.92 ± 66.43 (stat)	86940.36 ± 67.86 (stat)	0.68 ± 0.22 (stat)	0.36 ± 0.19 (stat)
Zmumu	26.61 ± 1.59 (stat)	23.16 ± 1.30 (stat)	69322.58 ± 59.46 (stat)	74223.55 ± 61.39 (stat)
Ztautau	6934.17 ± 77.68 (stat)	7069.61 ± 78.17 (stat)	7610.59 ± 87.26 (stat)	7686.91 ± 87.63 (stat)
Top	16581.90 ± 55.33 (stat)	17278.30 ± 57.85 (stat)	16299.96 ± 54.65 (stat)	17028.52 ± 58.24 (stat)
Diboson	2170.16 ± 48.59 (stat)	2187.45 ± 51.63 (stat)	2202.69 ± 49.58 (stat)	2321.34 ± 52.10 (stat)
MultiJet	224199.93	224708.74	127566.76	133141.44
Data	1259751.00 ± 1122.39 (stat)	1531246.00 ± 1237.44 (stat)	1206859.00 ± 1098.57 (stat)	1533100.00 ± 1238.18 (stat)

Table 4: Number of data yield and the expected contribution from the MC process and the MJ. Errors quoted are just statistical evaluated as simple $\sqrt{(N)}$ for MC statistic and data statistic, for systematic error related to MJ estimate refer to Section 5.5

Control plots for the $W^{-} \rightarrow e^{-}\overline{\nu}$, $W^{+} \rightarrow e^{+}\nu$ (Figures 25-26) and $W^{-} \rightarrow \mu^{-}\overline{\nu}$, $W^{+} \rightarrow \mu^{+}\nu$ (Figures 27-28) channels at 13 TeV are provided here after applying the event selection and all the corrections described in Section 4.3. Only MJ systematics uncertainty is shown in the ratio plots. Above 100 GeV the data/MC agreement in the $p_T^{\ell,\nu}$ distributions degradate due to missing truth p_T^W reweighting correction above this range. This discrepancy doesn't affects W-Ai measurement as analysis is performed in the $p_T^{\ell,\nu}$ bins. However p_T^W reweighting above 100 GeV would be addressed in the next NTuple iterations.

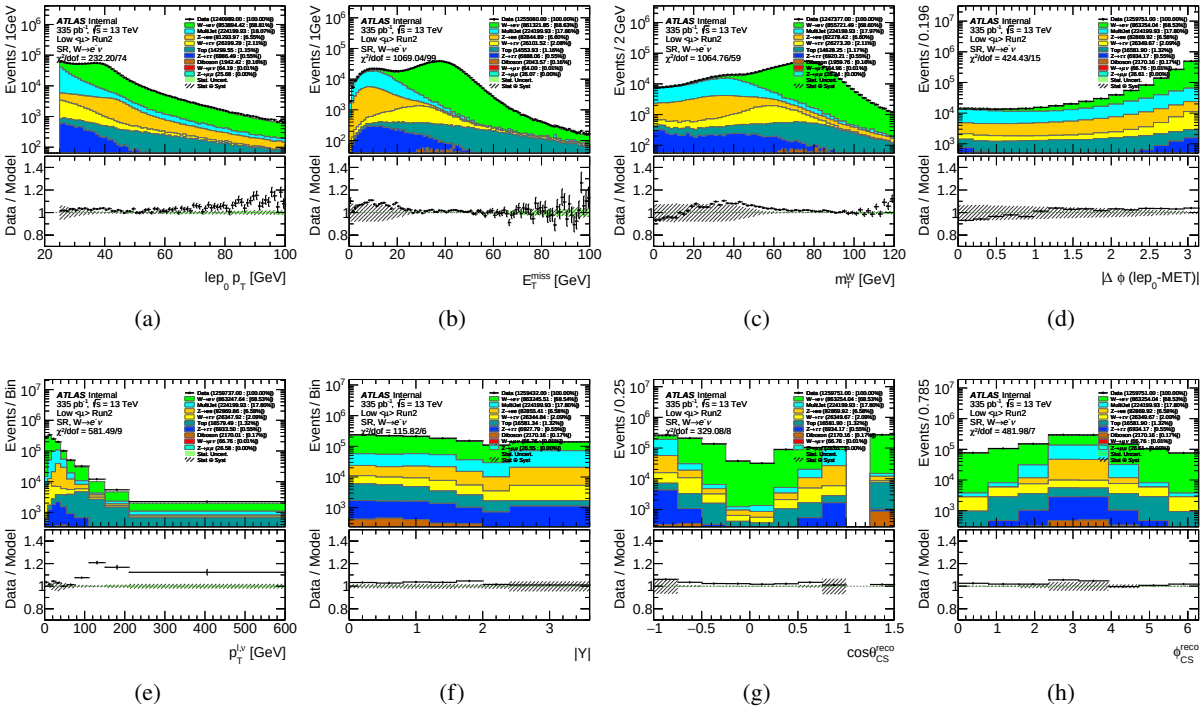


Figure 25: Control plots for inclusive signal region for $W^- \to e^-\overline{\nu}$ channel. Only overall multijet systematics uncertainty is shown.

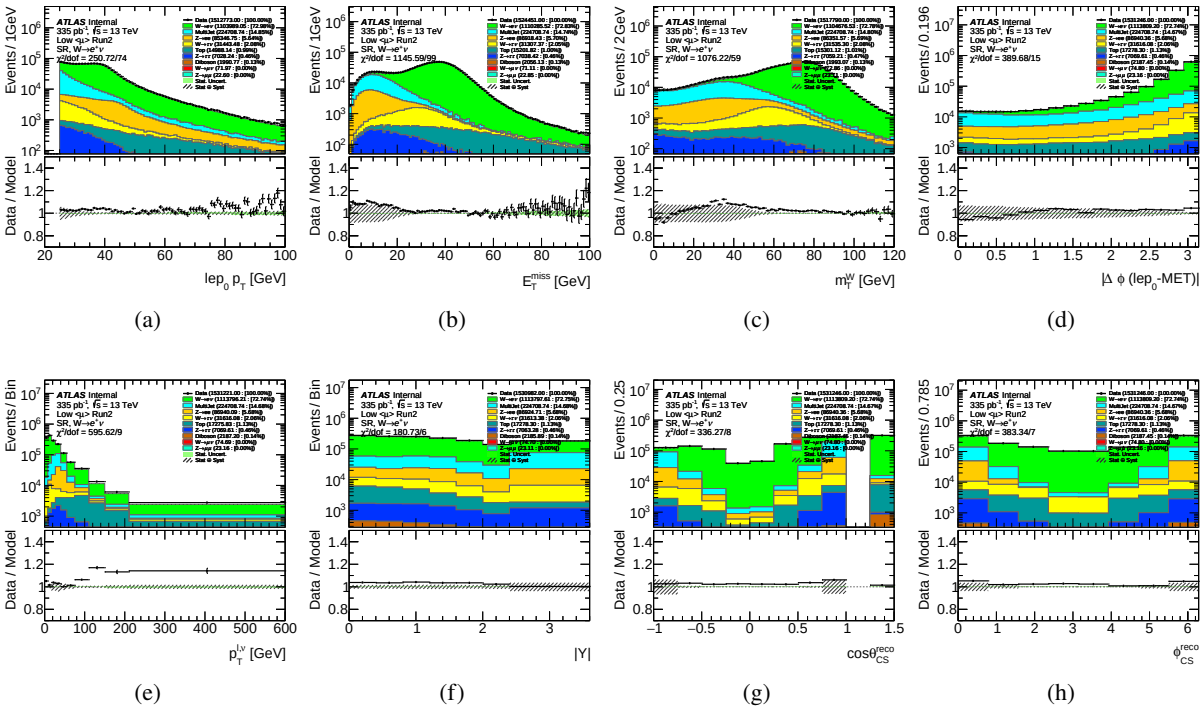
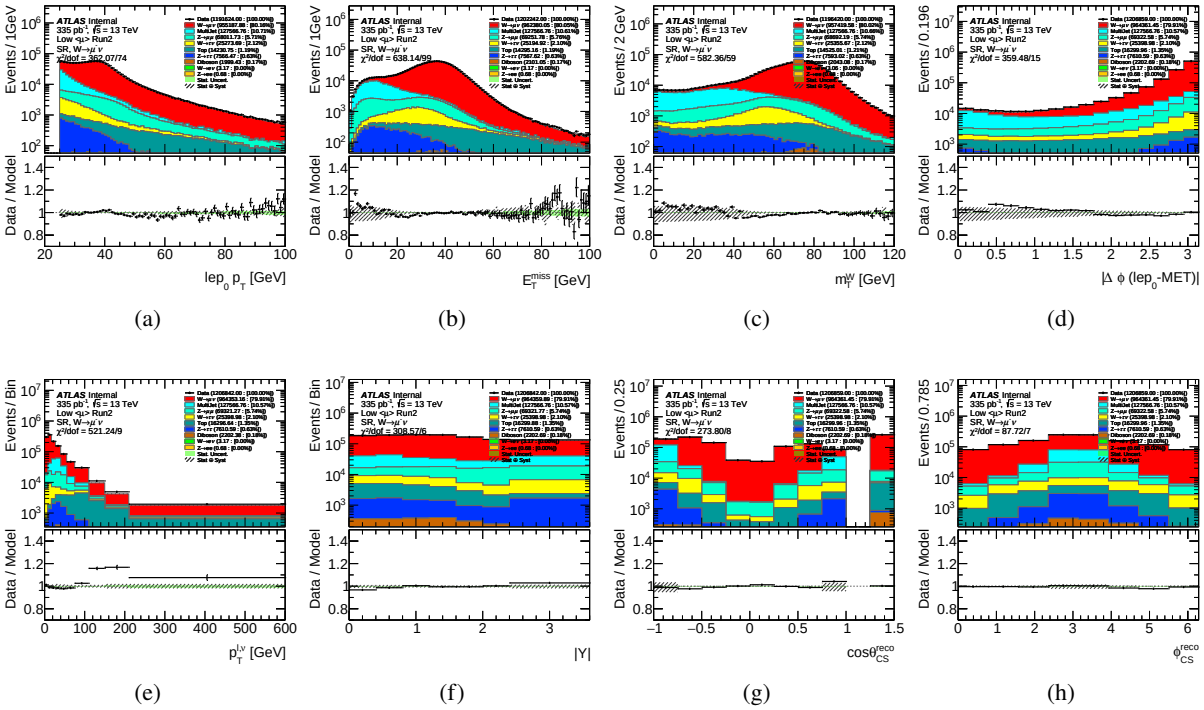


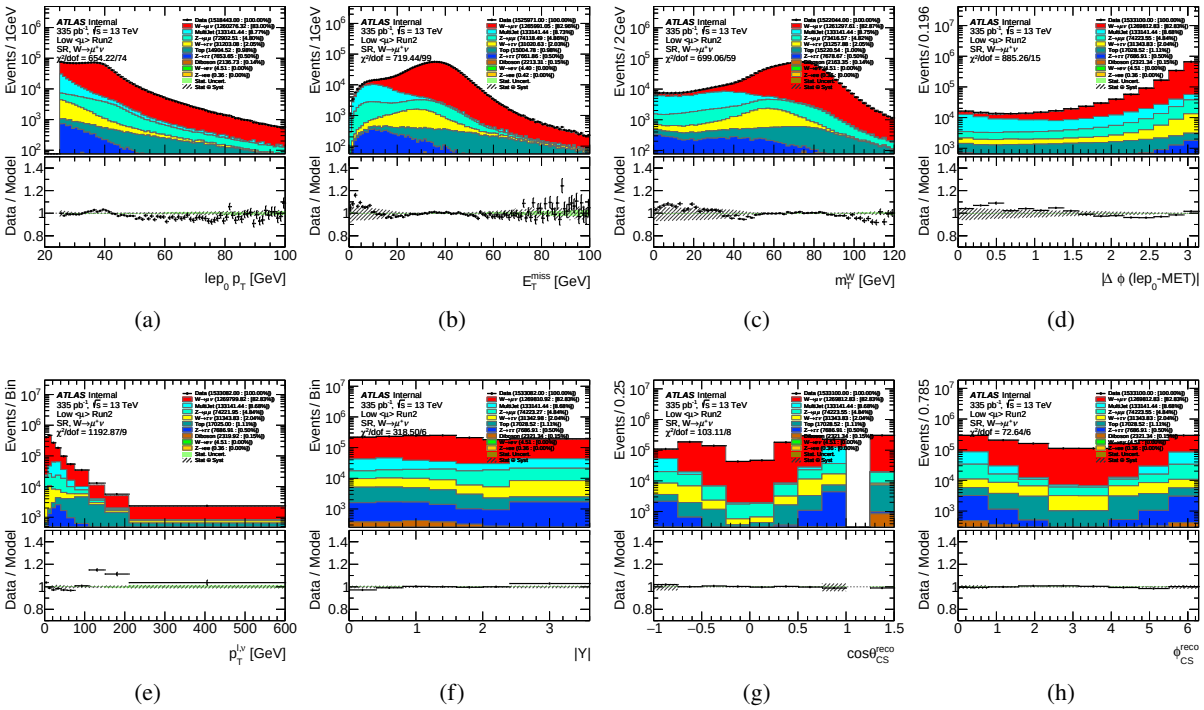
Figure 26: Control plots for inclusive signal region for $W^+ \to e^+\nu$ channel. Only overall multijet systematics uncertainty is shown.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 27: Control plots for inclusive signal region for $W^- \to \mu^- \overline{\nu}$ channel. Only overall multijet systematics uncertainty is shown.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 28: Control plots for inclusive signal region for $W^+ \to \mu^+ \nu$ channel. Only overall multijet systematics uncertainty is shown.

7 Uncertainties

The dominant source of uncertainty in this measurement is the statistical uncertainties in the data. These uncertainties are described in Section 7.1. The systematic uncertainties due to impact of experimental and theoretical uncertainties are described in Section 7.2 and 7.3.

7.1 Statistical Uncertainties

Uncertainties due to the finite data statistics can be divided into several categories as shown in Figure 29 for an example of A_0 in $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin 0. Uncertainties due to the parameter of interest are labeled as data-stat uncertainties where the available statistics of the data has the largest impact but resolution in the angular binning also impacts this value. The columns correspond to each coefficient while the rows correspond to the $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin. For example, the first column is for parameters in the same $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ as the parameter of interest and for rows 2-8 correspond to the other coefficients these are labeled as "shape" parameters which are non-migrational. While the first row corresponds to parameters for the same coefficient but in different $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins and are self-migrational parameters. The combination of different $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ and different coefficient number are labeled as "cross-migration". Lastly, this separation is also done for the normalization factors which are labeled "norm" and "norm-migration". The separated migration uncertainties for each coefficient can be found in Appendix D. There also is a non-migrational uncertainty which comes from the uncorrelated statistical uncertainty in the likelihood fit.

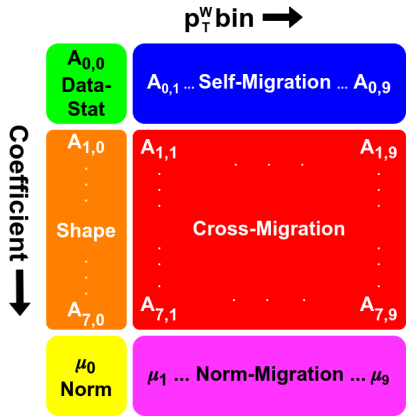


Figure 29: Categorization of parameters for the data-statistical uncertainty on the measured coefficients for an example of A_0 in $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin 0.

7.2 Experimental Uncertainties

The main experimental uncertainties affecting this measurement are expected to arise from hadronic recoil, resolution, and calibration. Lepton related systematic are expected to be subdominant. Multi-jet and EW+top background estimates contribute as well to the relevant experimental uncertainties which need to be evaluate. These experimental uncertainties only impact the numerator of the templates discussed in Section 3.2.

The systematic uncertainties connected to the hadronic recoil calibration (discussed in Section 4.4) are separated into 6 categories with a combined 62 nuisance parameters corresponding to the resolution, response, and the ΣE_{T} . Appendix C detailed some dedicated studies on the recoil systematics on the fit results.

The lepton systematic related to the leptons calibration and the leptons efficiency correction (discussed in Section 4.3) are expected to be subdominant for the but still need to be implemented. For the electron channels the SF efficiency SF are separated for the identification, isolation, reconstruction, and trigger. For the muon channels these are separated for the isolation, reconstruction, sagitta bias, trigger, and track-to-vertex-association.

The uncertainties for EW+top background estimates are taken as a flat uncertainty on their cross-sections. For single boson samples this is 5%, diboson samples 10%, and top backgrounds range from 7-10%; specifics for each sample can be found in Table 2 in Section 4.1. These variations are grouped into 3 categories for a total of 3 nuisance parameters based on the background type into boson, diboson, and top.

The uncertainties for the multi-jet background are discussed in Section 5 and Appendix E and are separated into 3 categories with a combined $(3 \times 10) = 30$ nuisance parameters corresponding to the acceptance, shape, and yield.

7.3 Theoretical Uncertainties

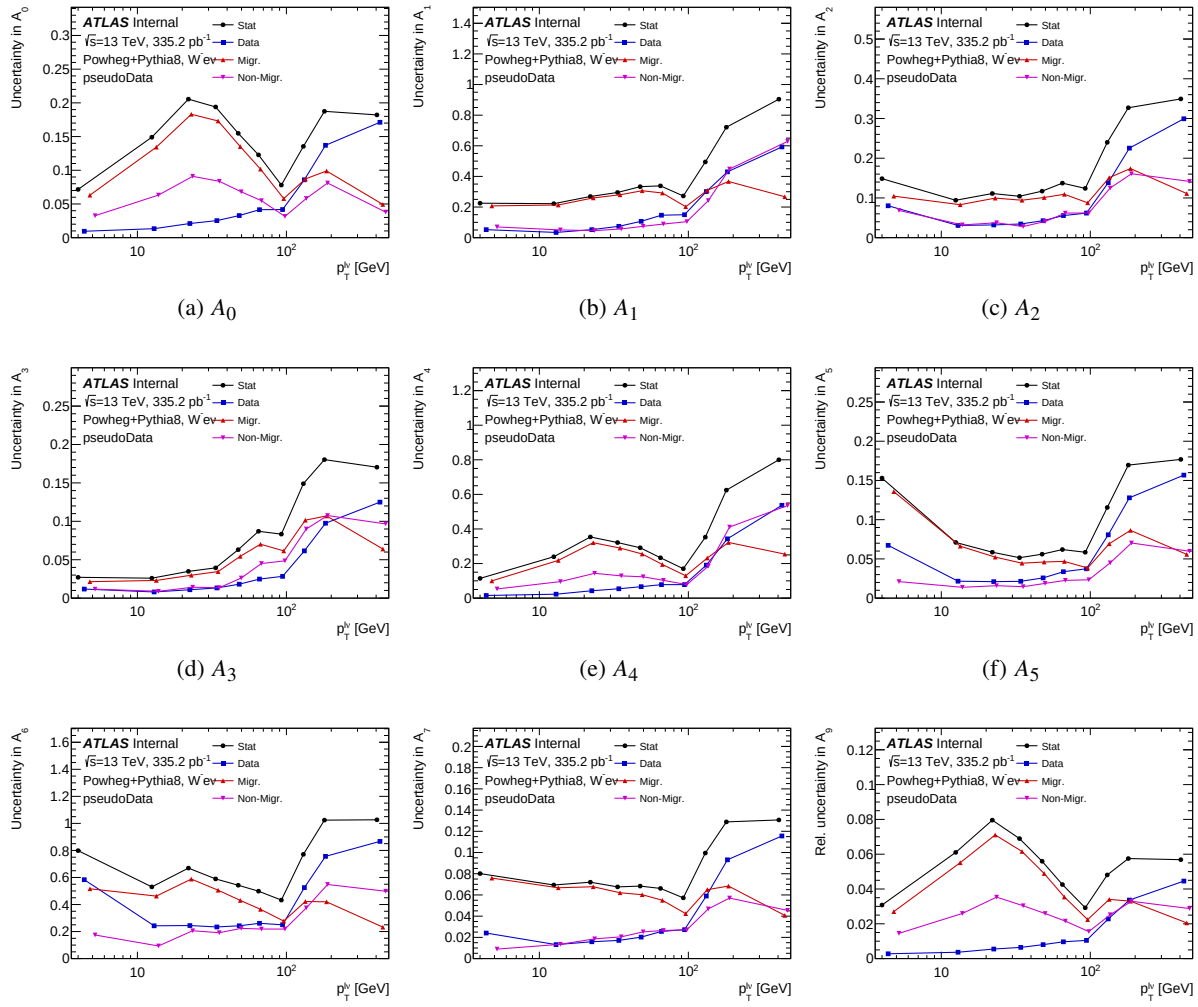
The only theoretical uncertainties considered are due to PDFs which have an impact on both the numerator and denominator of the templates. These are described by 26 NPs and are subdominant for $A_0 - A_7$ and approximately $\sim 3\%$ for the differential cross-section. Examples of the total expected PDF uncertainty can be found in Figures 38 and 47. Other theoretical uncertainties such as parton shower, generator modelling, and QED/electroweak corrections are not considered as they were found to be negligible for the Z angular coefficient measurement at $\sqrt{s} = 8$ TeV which had more precision than this measurement.

7.4 Expected Uncertainties

7.4.1 p_{T}^{W} Differential

Uncertainty	Description
Response_Sys	Systematic uncertainty on the response correction
Response_ExtSys	Systematic uncertainty on the response correction due to the extrapolation from Z to W
Response_Stat0_1-15	Statistical uncertainty on the response correction
Response_Stat1_1-15	Statistical uncertainty on the response correction
Resolution_ExtSys	Systematic uncertainty on the resolution correction due to the extrapolation from Z to W
Resolution_Stat0_1-14	Statistical uncertainty on the resolution correction
Resolution_Stat1_1-14	Statistical uncertainty on the resolution correction
SET_Sys	Systematic uncertainty on the ΣE_{T}
XSec_Boson	Flat 5% variation on the cross-section for all single boson samples
XSec_Diboson	Flat 10% variation on the cross-section for all diboson samples
XSec_Top	Flat variation between 7-10% on the cross-section for all top associated samples
MJ_Acc	Uncertainty on acceptance correction functions
MJ_Shape	Uncertainty on MJ template shape
MJ_Yield	Uncertainty on yield from intersection point
MC Stat	Poisson variation for the limited MC statistics used in template generation
PDF_1-26	Systematic uncertainty on the parton distribution function

Table 5: Description of the systematics nuisance parameters used in the analysis.

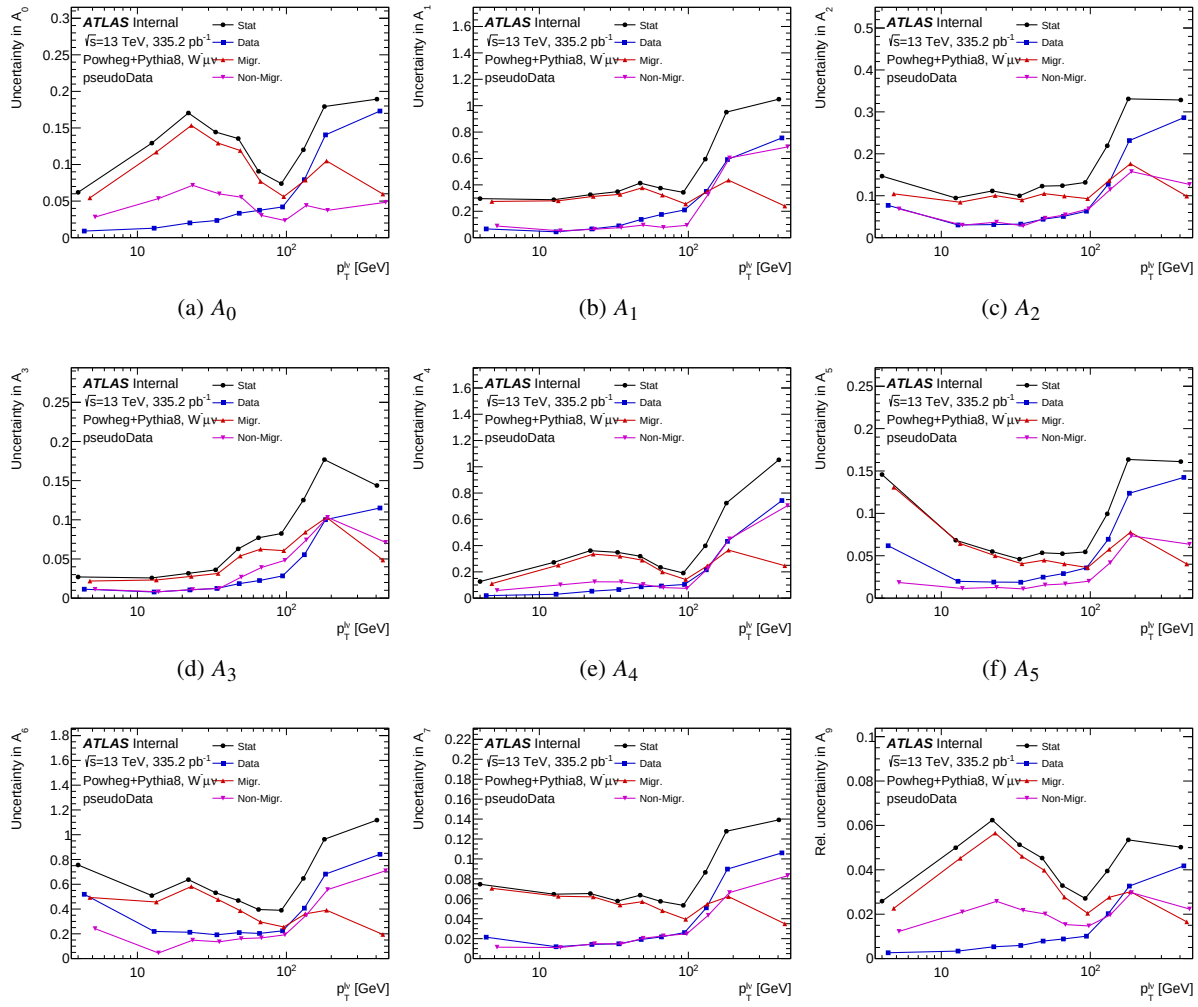


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 30: The uncertainty comparison for the statistical uncertainties in the $W^- \to e^-\overline{\nu}$ channel for measurement in p_{T}^{W}. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 104, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.

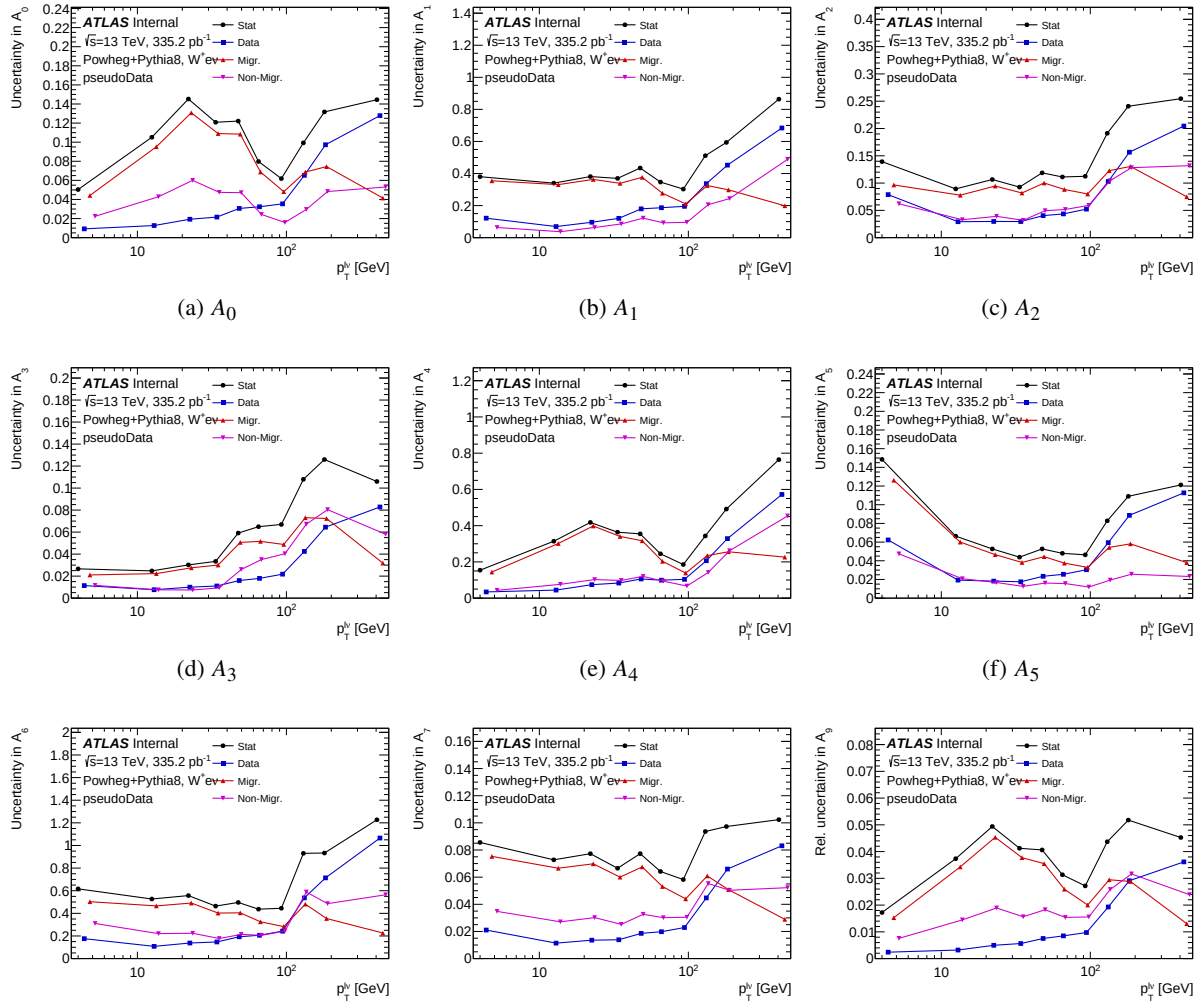


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 31: The uncertainty comparison for the statistical uncertainties in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in p_{T}^W. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 105, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.

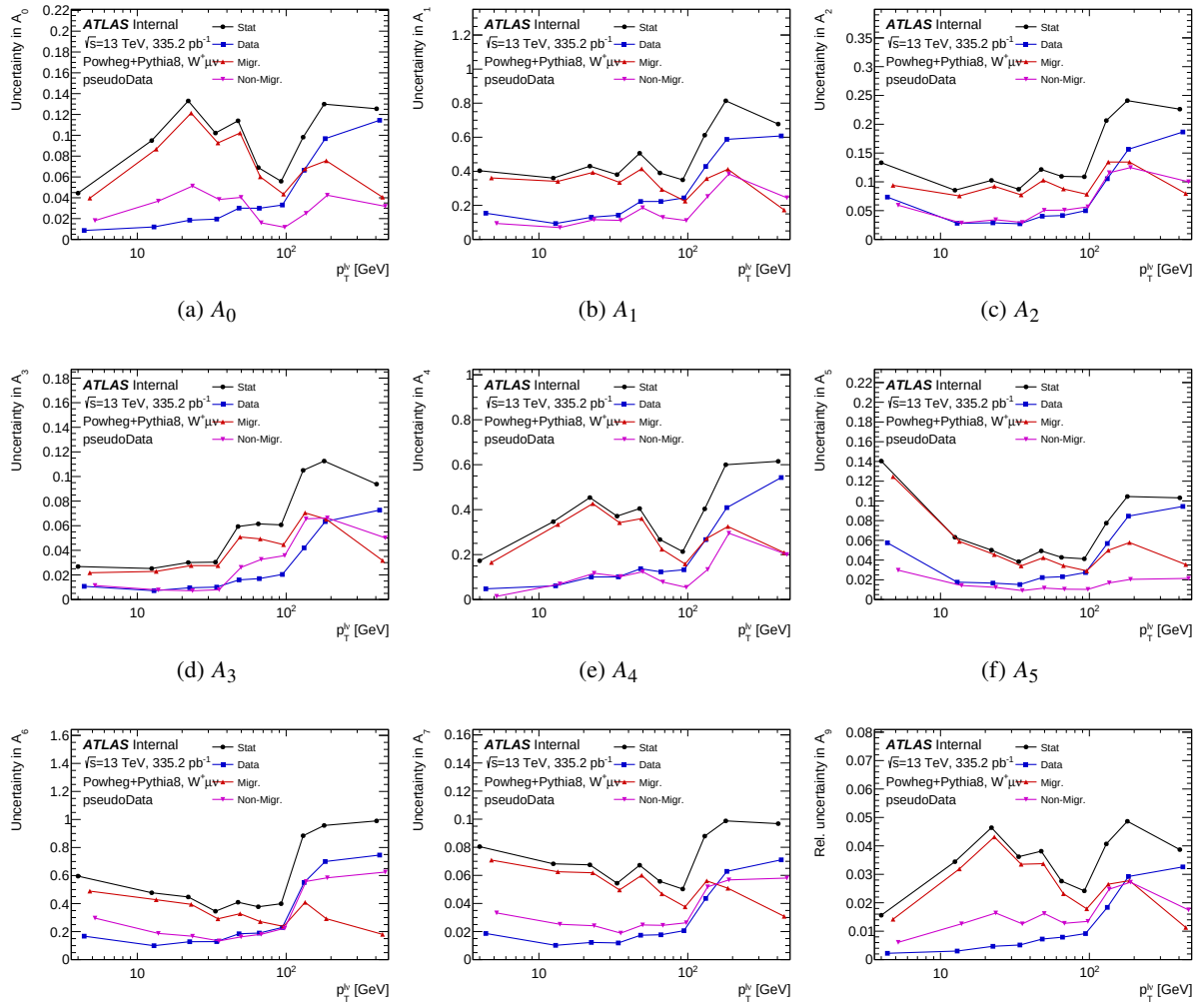


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 32: The uncertainty comparison for the statistical uncertainties in the $W^+ \to e^+\nu$ channel for measurement in p_{T}^{W}. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 106, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.

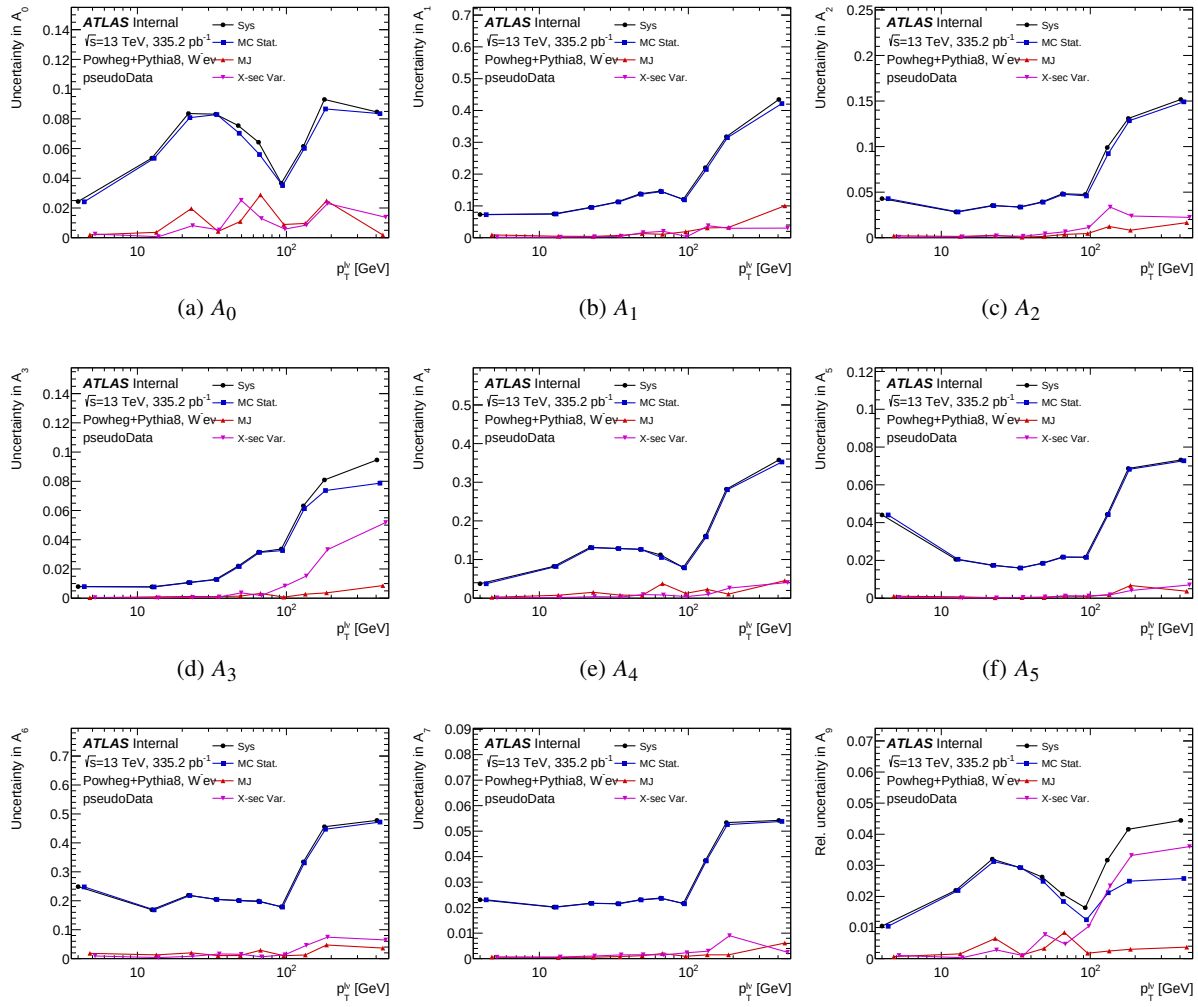


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 33: The uncertainty comparison for the statistical uncertainties in the $W^+ \to \mu^+\nu$ channel for measurement in p_{T}^{W}. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 107, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.

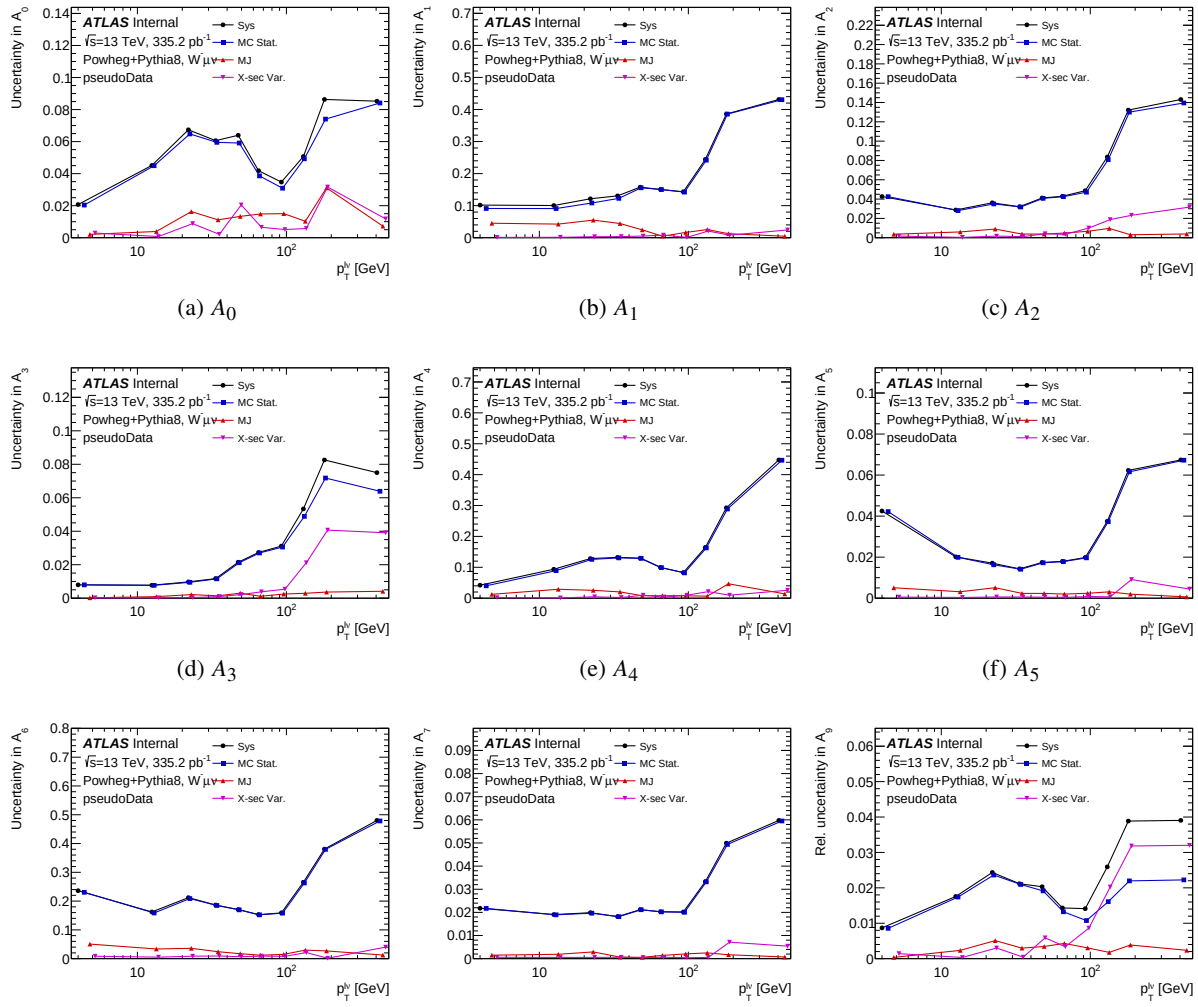


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 34: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^- \to e^- \overline{\nu}$ channel for measurement in p_{T}^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.

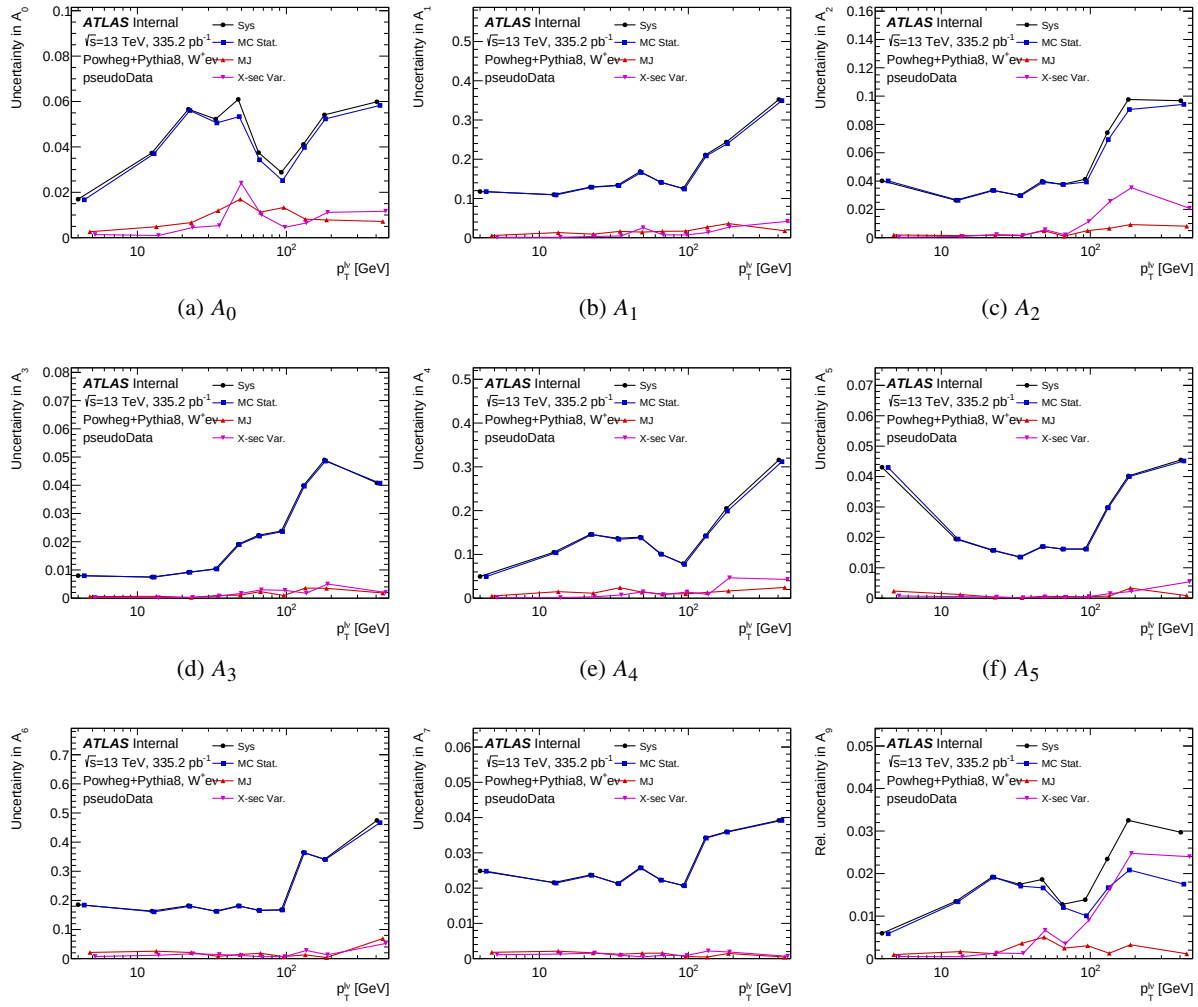


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 35: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in p_{T}^{W}. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.

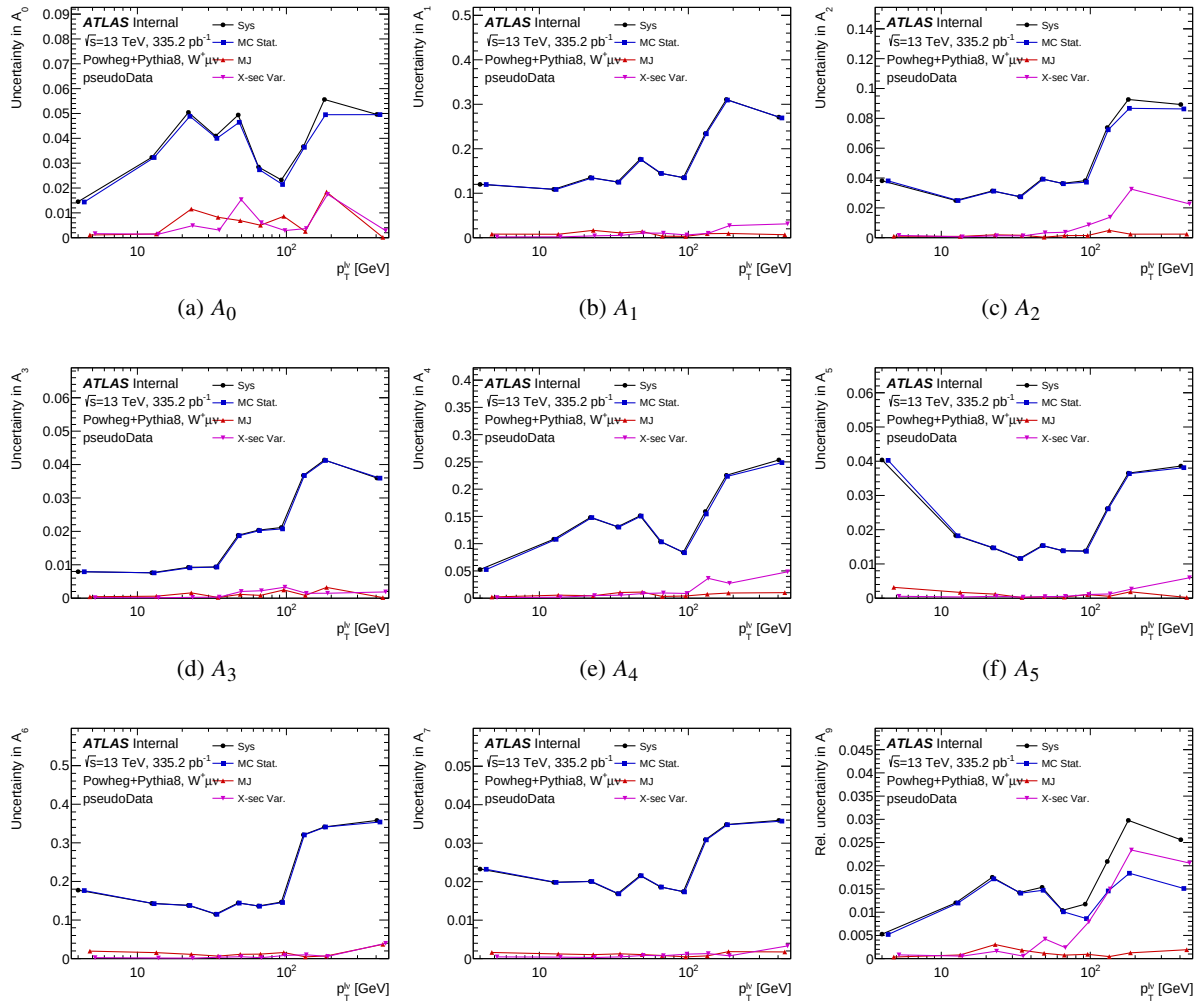


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 36: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^+ \to e^+\nu$ channel for measurement in p_{T}^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.

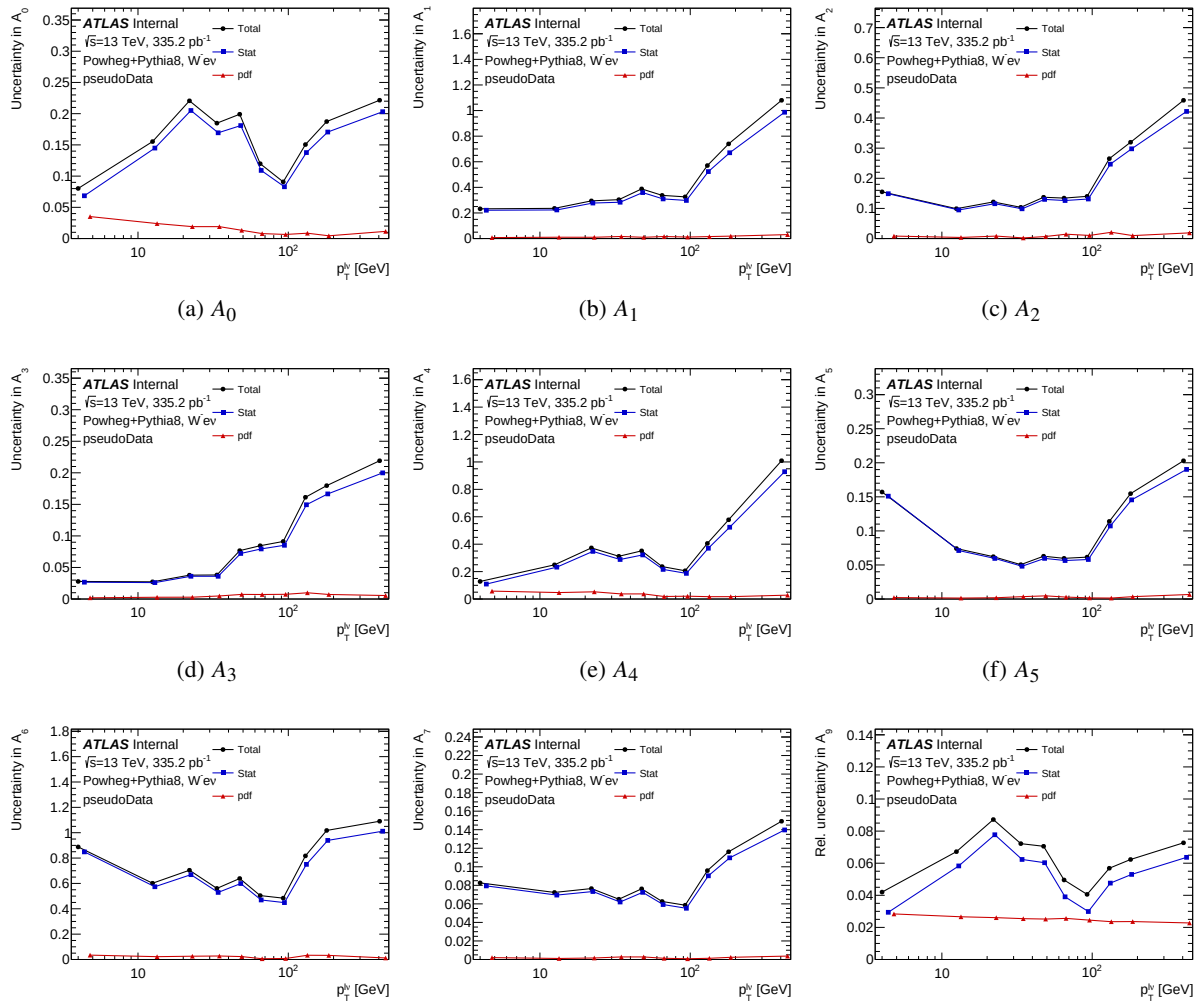


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 37: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^+ \to \mu^+\nu$ channel for measurement in p_{T}^{W}. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.



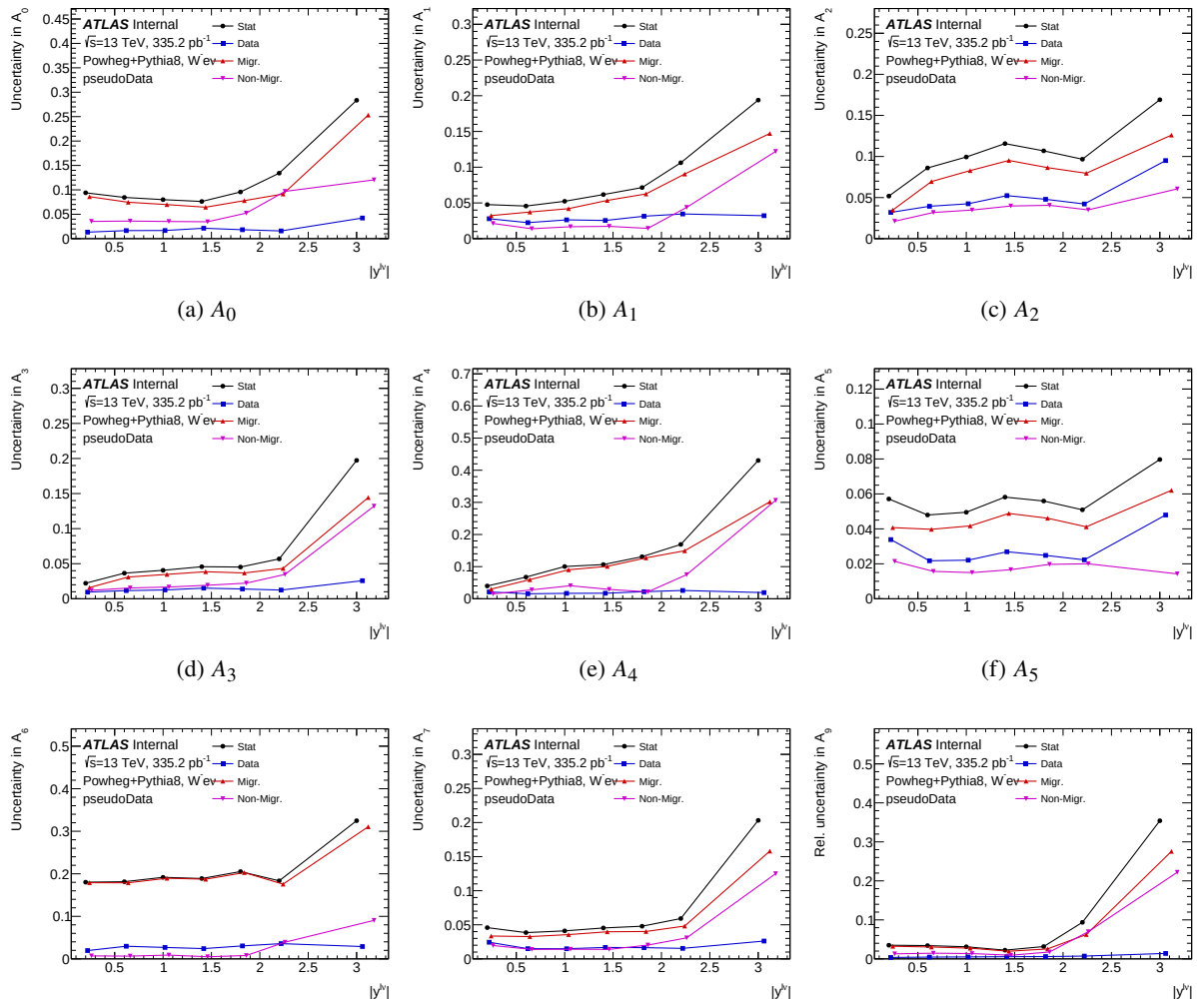
(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 38: Comparison of the PDF uncertainty in red to the statistical uncertainty in blue in the $W^- \to e^- \overline{\nu}$ channeling for measurement in p_{T}^W.

7.4.2 y^W Differential

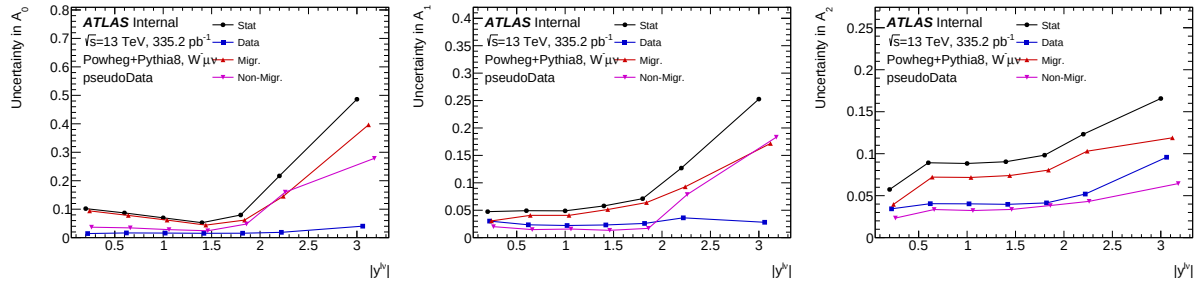


(a) A_0 (b) A_1 (c) A_2

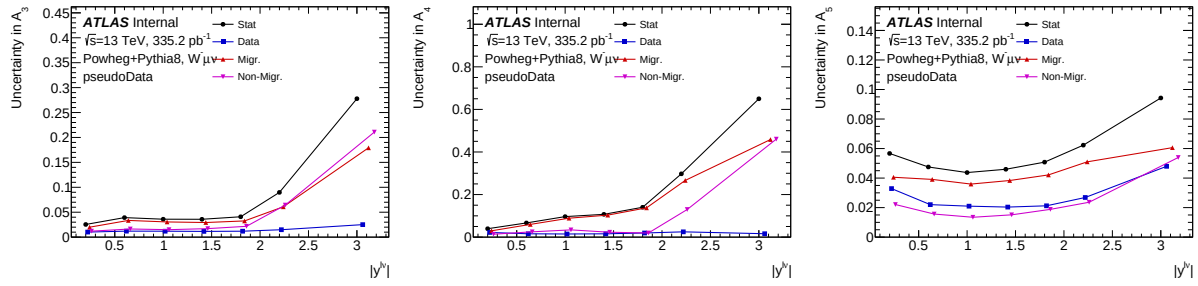
(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

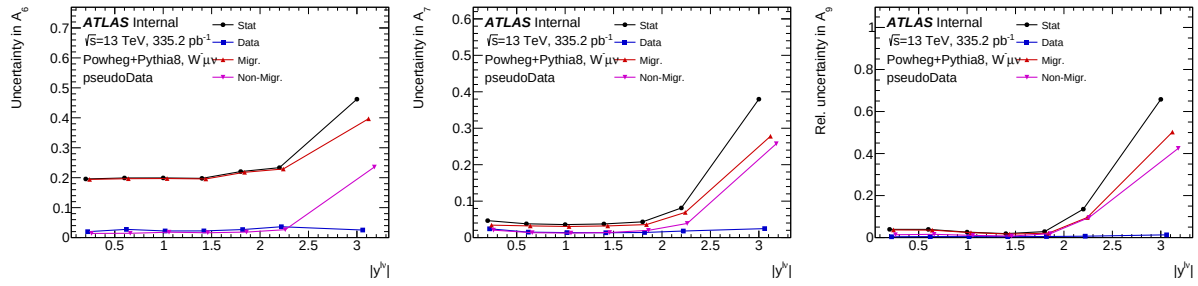
Figure 39: The uncertainty comparison for the statistical uncertainties in the $W^- \to e^-\overline{\nu}$ channel for measurement in y^W. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 108, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.



(a) A_0 (b) A_1 (c) A_2

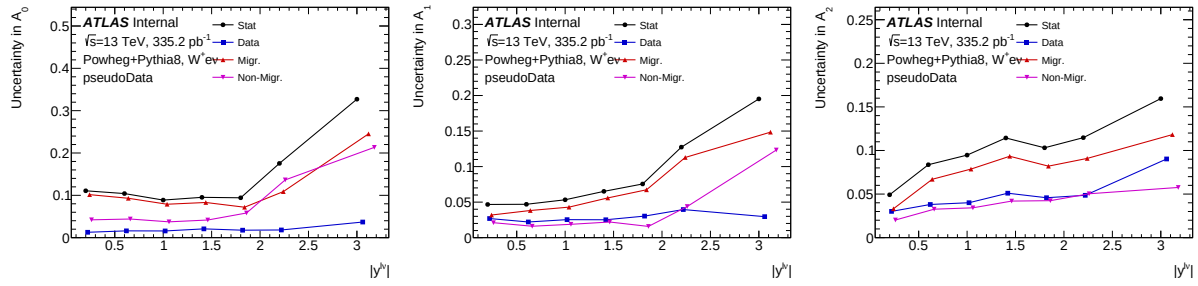


(d) A_3 (e) A_4 (f) A_5

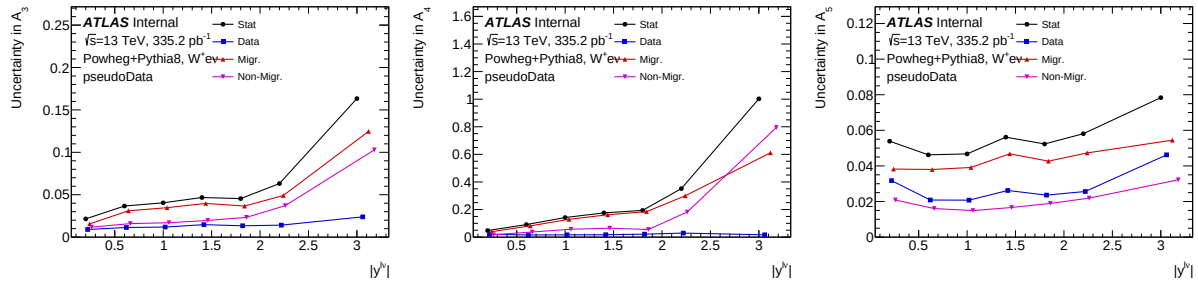


(g) A_6 (h) A_7 (i) A_9

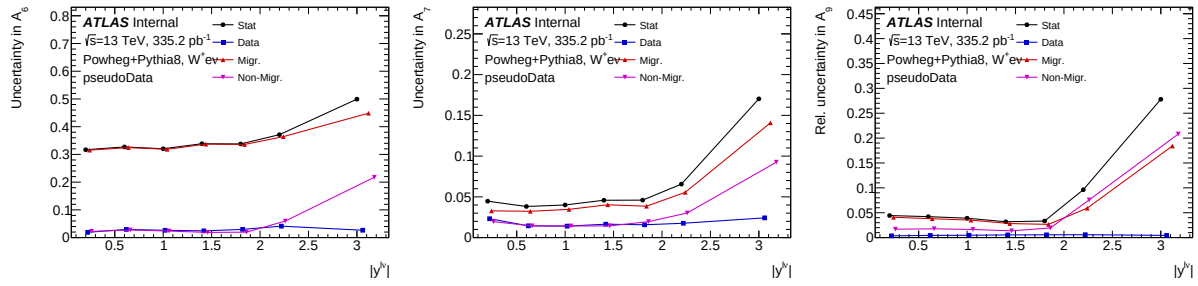
Figure 40: The uncertainty comparison for the statistical uncertainties in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in y^W. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 109, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.



(a) A_0 (b) A_1 (c) A_2

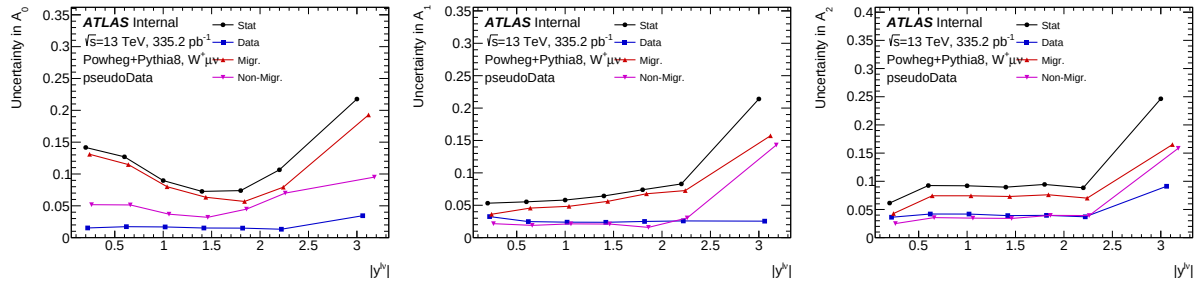


(d) A_3 (e) A_4 (f) A_5

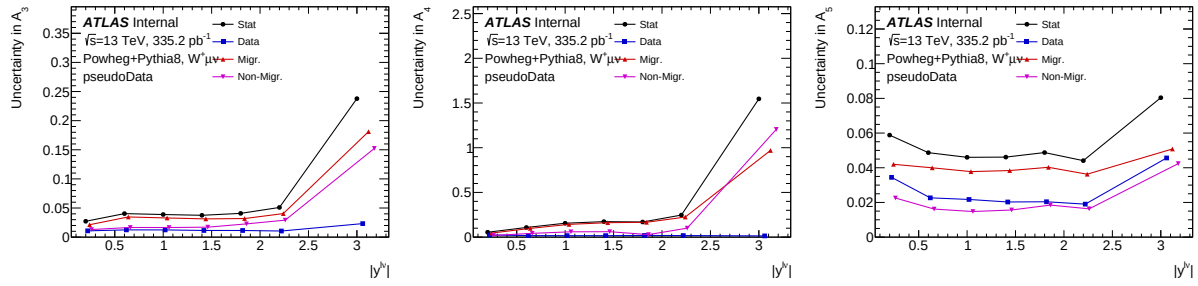


(g) A_6 (h) A_7 (i) A_9

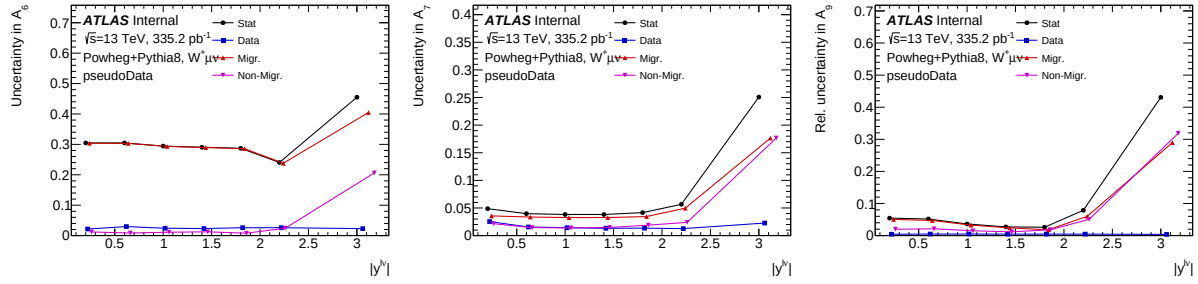
Figure 41: The uncertainty comparison for the statistical uncertainties in the $W^+ \to e^+\nu$ channel for measurement in y^W. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 110, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.



(a) A_0 (b) A_1 (c) A_2

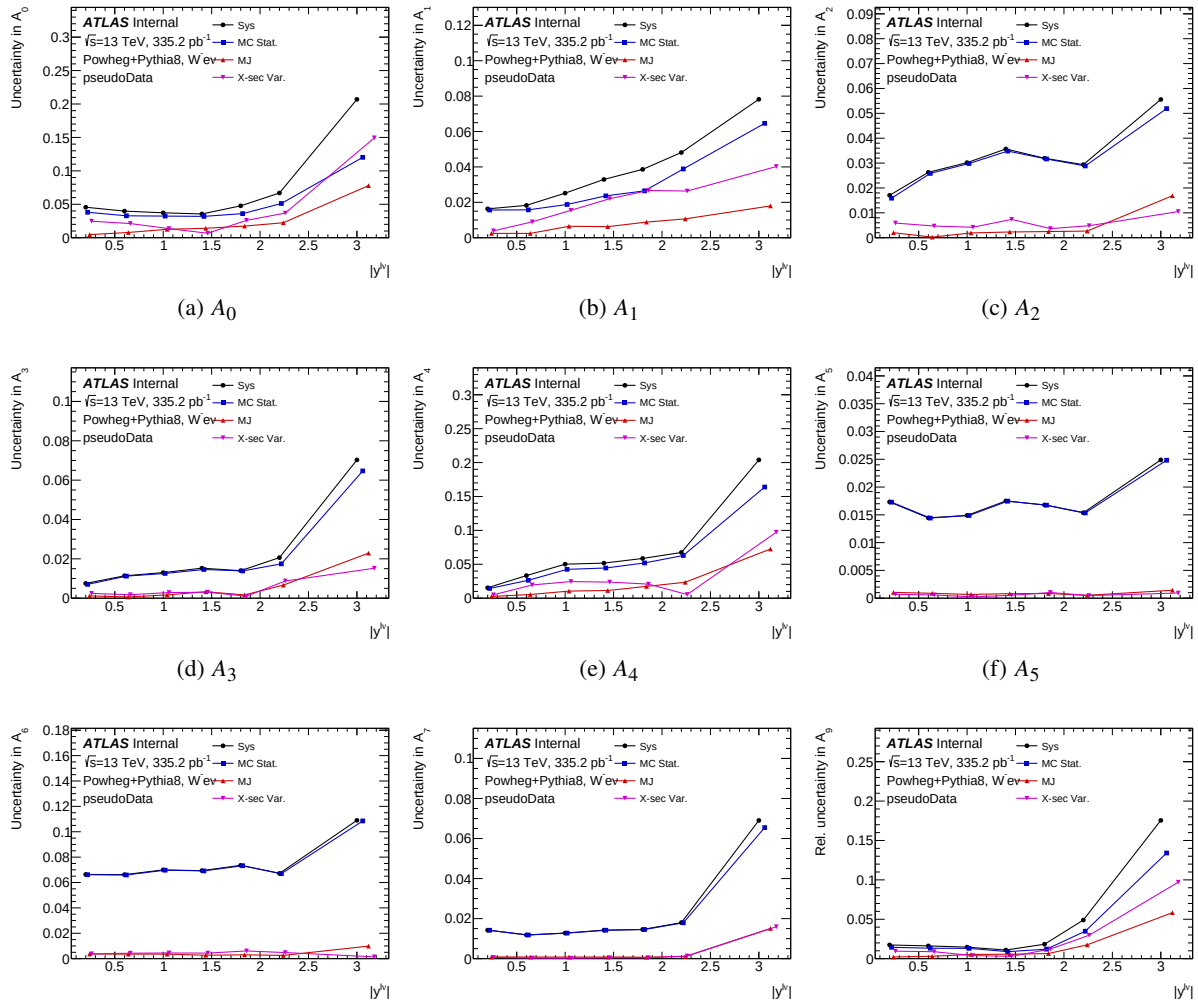


(d) A_3 (e) A_4 (f) A_5



(g) A_6 (h) A_7 (i) A_9

Figure 42: The uncertainty comparison for the statistical uncertainties in the $W^+ \to \mu^+\nu$ channel for measurement in y^W. Blue is the uncertainty due to the limited data statistics in each bin, red is the total migration uncertainty which is decomposed further in Figure 111, and pink are the uncorrelated statistical uncertainty from the likelihood fit. Black is the combination of all of these components to give the total statistical uncertainty.

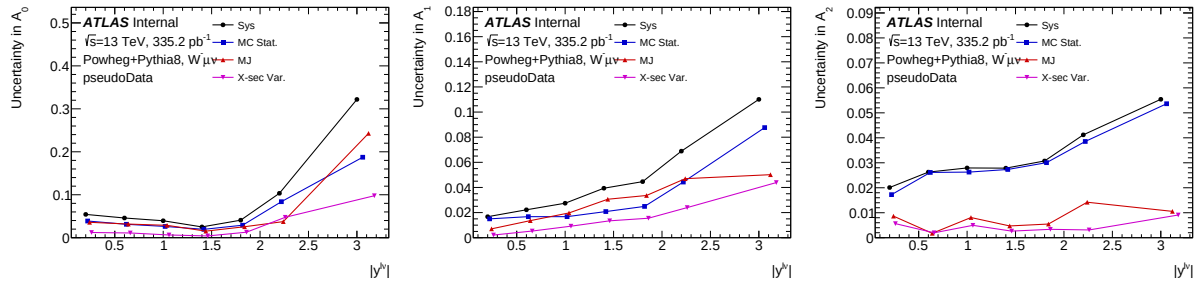


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

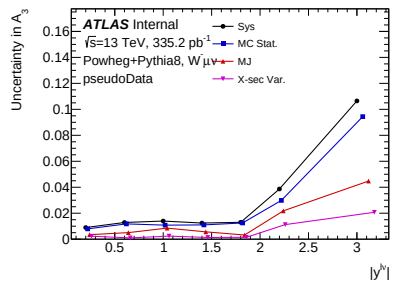
Figure 43: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^- \to e^-\overline{\nu}$ channel for measurement in y^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.



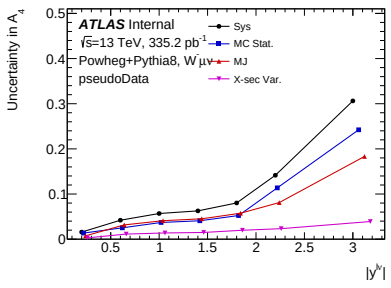
(a) A_0

(b) A_1

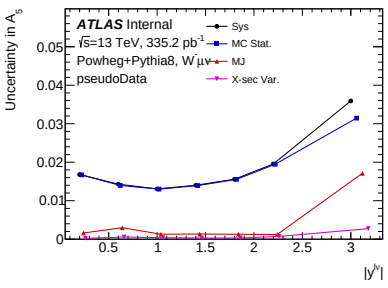
(c) A_2



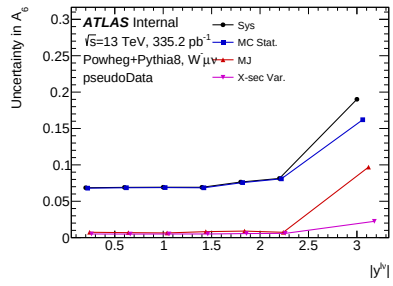
(d) A_3



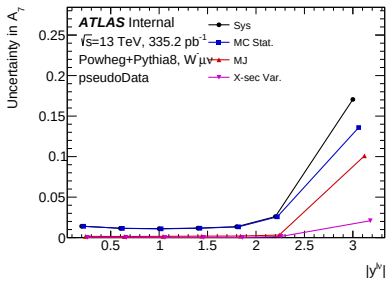
(e) A_4



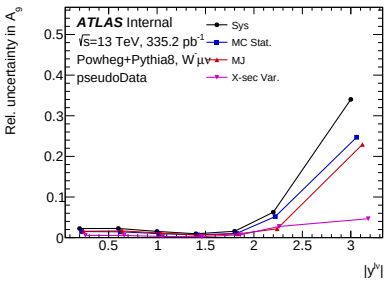
(f) A_5



(g) A_6

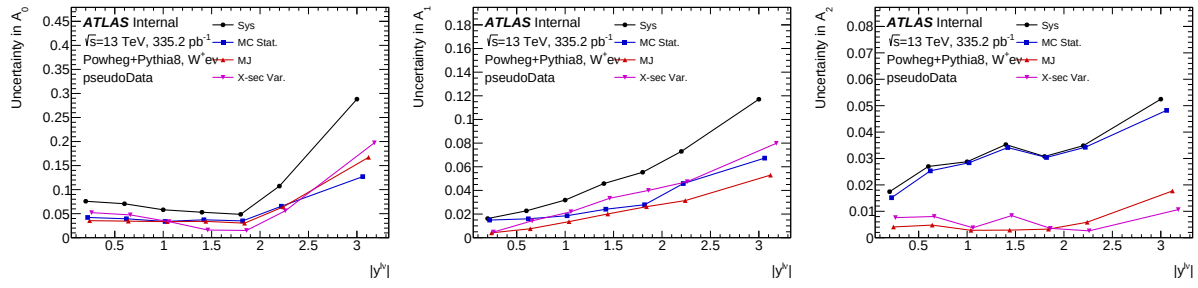


(h) A_7



(i) A_9

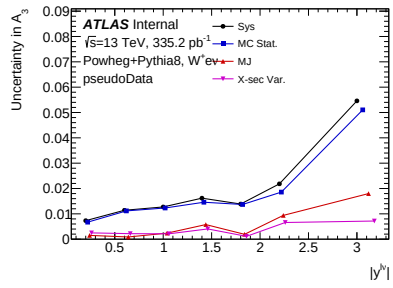
Figure 44: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in y^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.



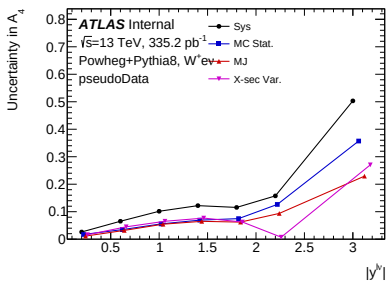
(a) A_0

(b) A_1

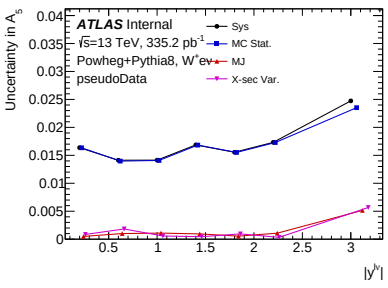
(c) A_2



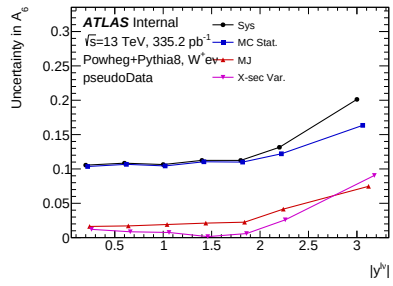
(d) A_3



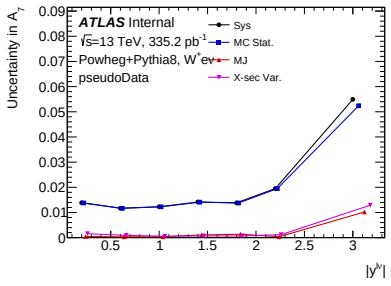
(e) A_4



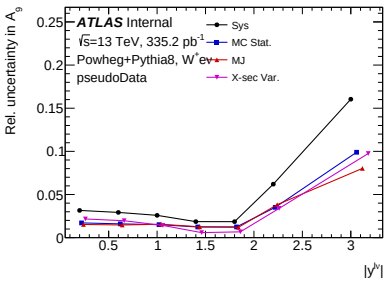
(f) A_5



(g) A_6

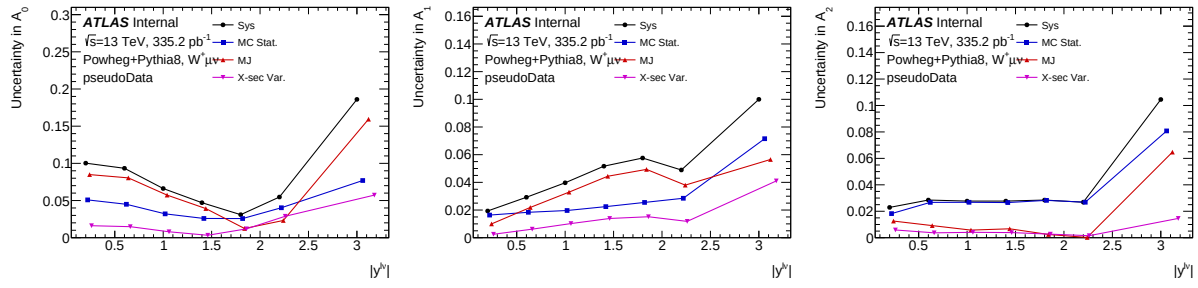


(h) A_7

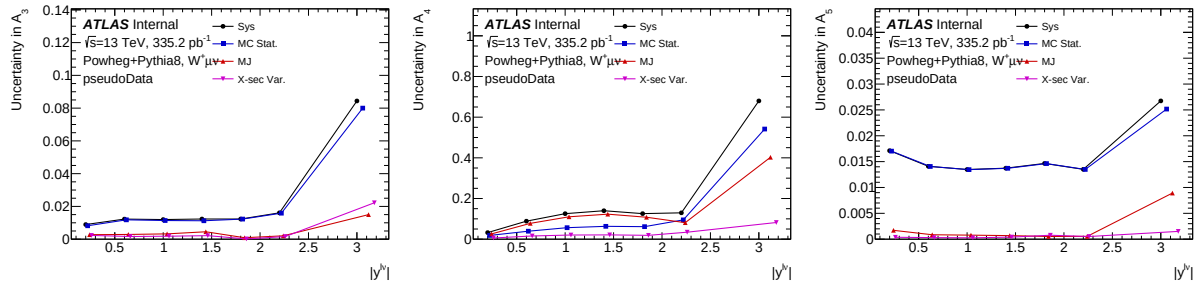


(i) A_9

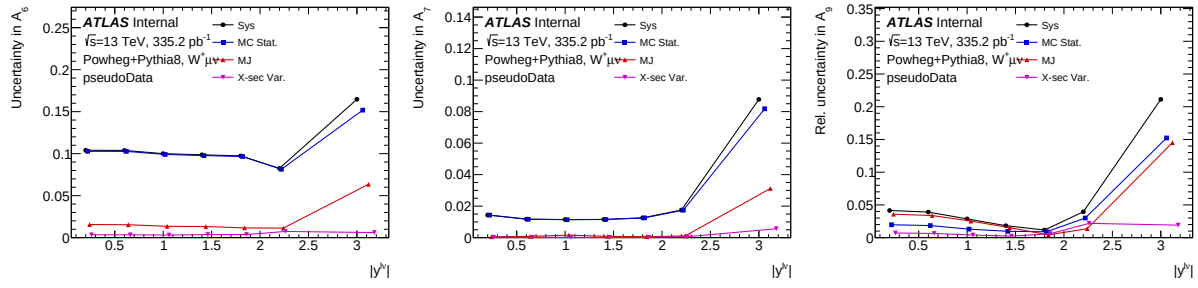
Figure 45: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^+ \to e^+\nu$ channel for measurement in y^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.



(a) A_0 (b) A_1 (c) A_2

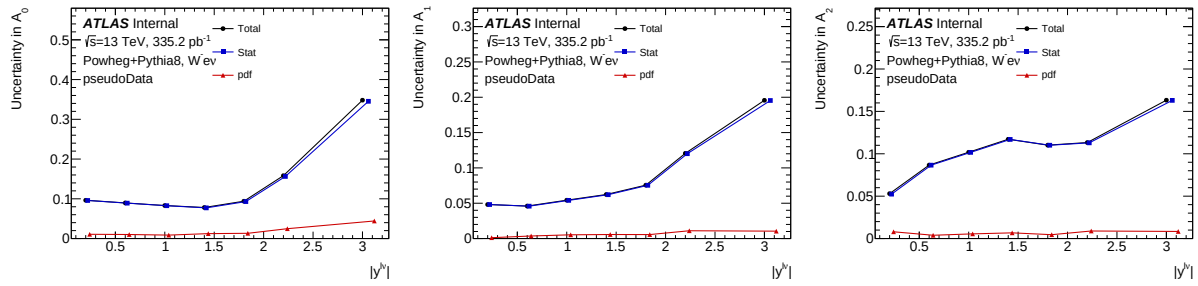


(d) A_3 (e) A_4 (f) A_5



(g) A_6 (h) A_7 (i) A_9

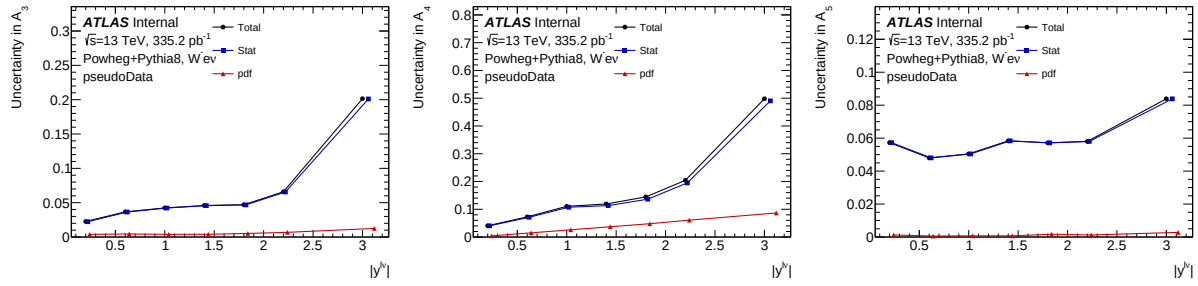
Figure 46: The uncertainty comparison for the background related systematic uncertainties and the dominant MC stat. systematic in the $W^+ \to \mu^+\nu$ channel for measurement in y^W. Blue is the uncertainty related to the limited MC statistics used to generate the templates, MJ is the uncertainty related to the MJ background estimate, and pink is the uncertainty from varying the background cross section. Black is the combination of these three systematic uncertainties.



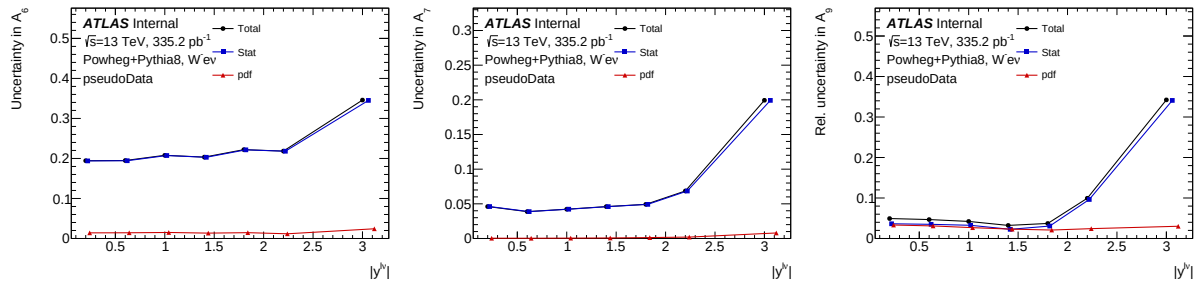
(a) A_0

(b) A_1

(c) A_2



(d) A_3 (e) A_4 (f) A_5



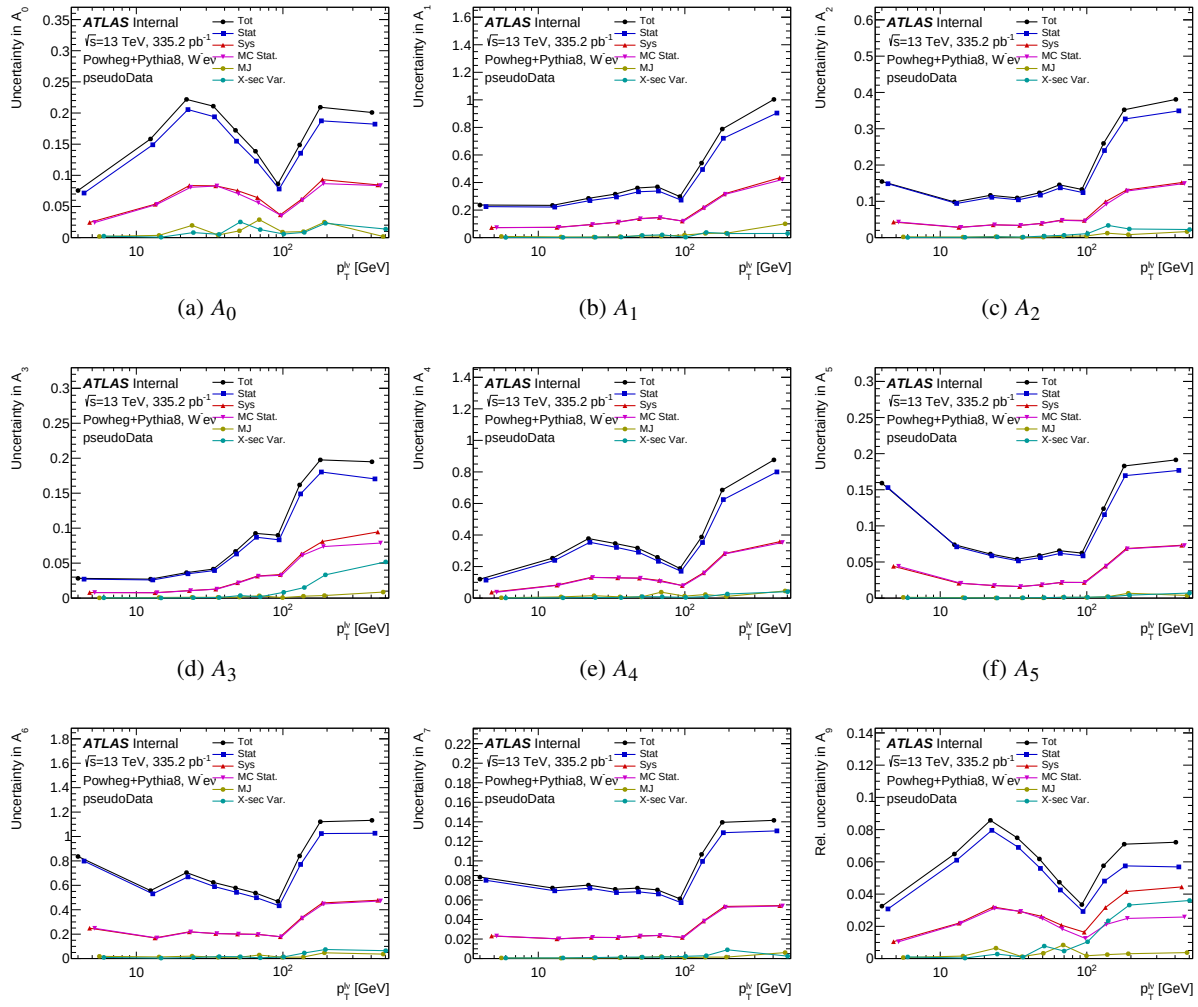
(g) A_6 (h) A_7 (i) A_9

Figure 47: Comparison of the PDF uncertainty in red to the statistical uncertainty in blue in the $W^- \to e^- \overline{\nu}$ channeling for measurement in y^W.

7.5 Summary of Uncertainties

A summary of the expected statistical and systematic uncertainties associated to the measurement of the A_i coefficients as a function of p_{T}^{W} is presented in Section 7.5.1 and as function of y^{W} is presented in Section 7.5.2 for each category of events ($W^- \to e^-\overline{\nu}$, $W^+ \to e^+\nu$, $W^- \to \mu^-\overline{\nu}$, and $W^+ \to \mu^+\nu$). The dominant uncertainty on the measurement of the A_i coefficients is the statistical uncertainty for all bins for both p_{T}^{W} and y^{W} differential measurements. The next dominant uncertainty is due to the signal MC statistical uncertainty, which is $\sim$ 1/3 of the data-statistical uncertainty due since the signal MC was produced with approximated ten times more events than the data sets used.

7.5.1 p_{T}^{W} Differential

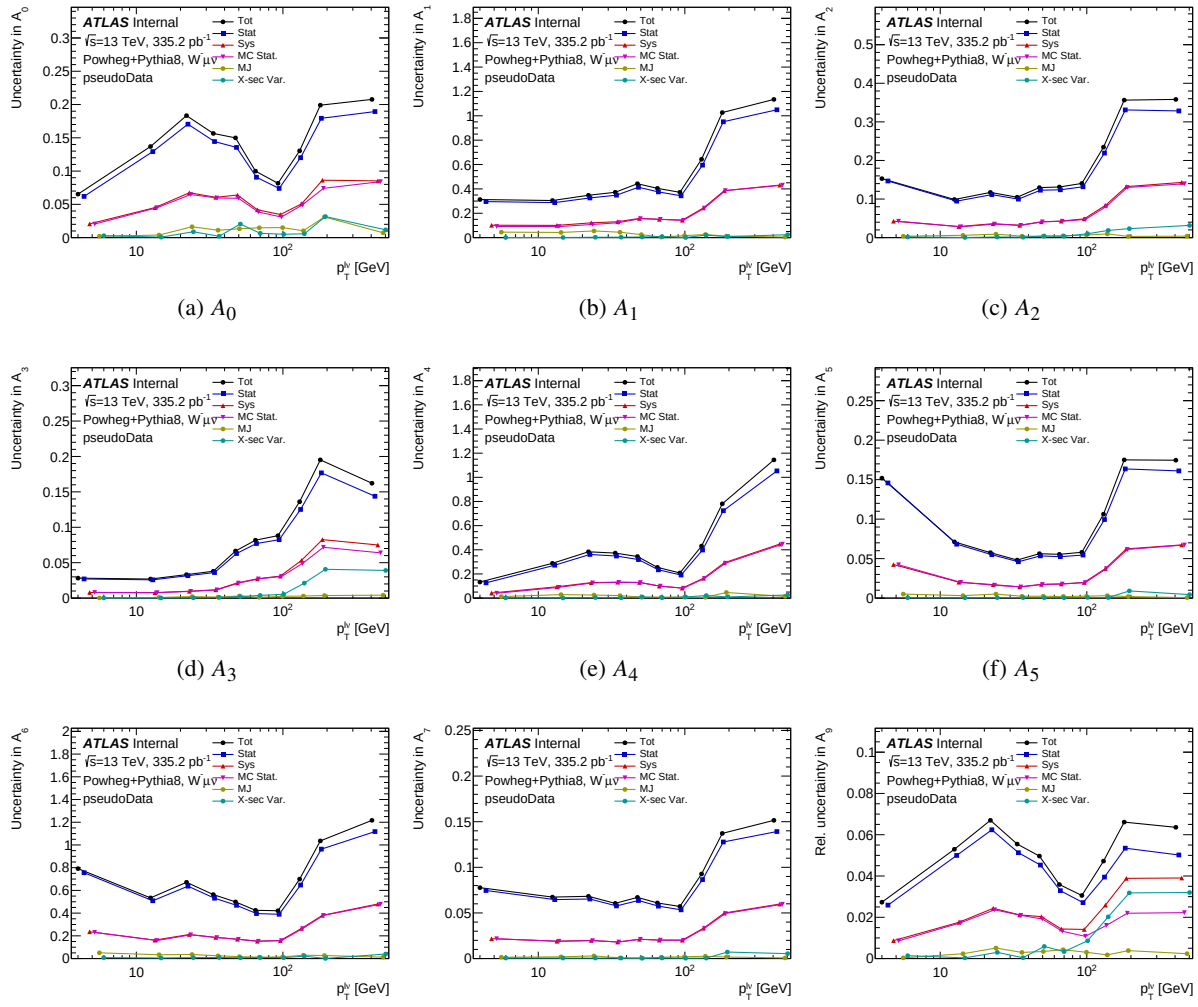


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 48: The uncertainty comparison for all contributions in the $W^- \to e^- \overline{\nu}$ channel for measurement in p_{T}^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.

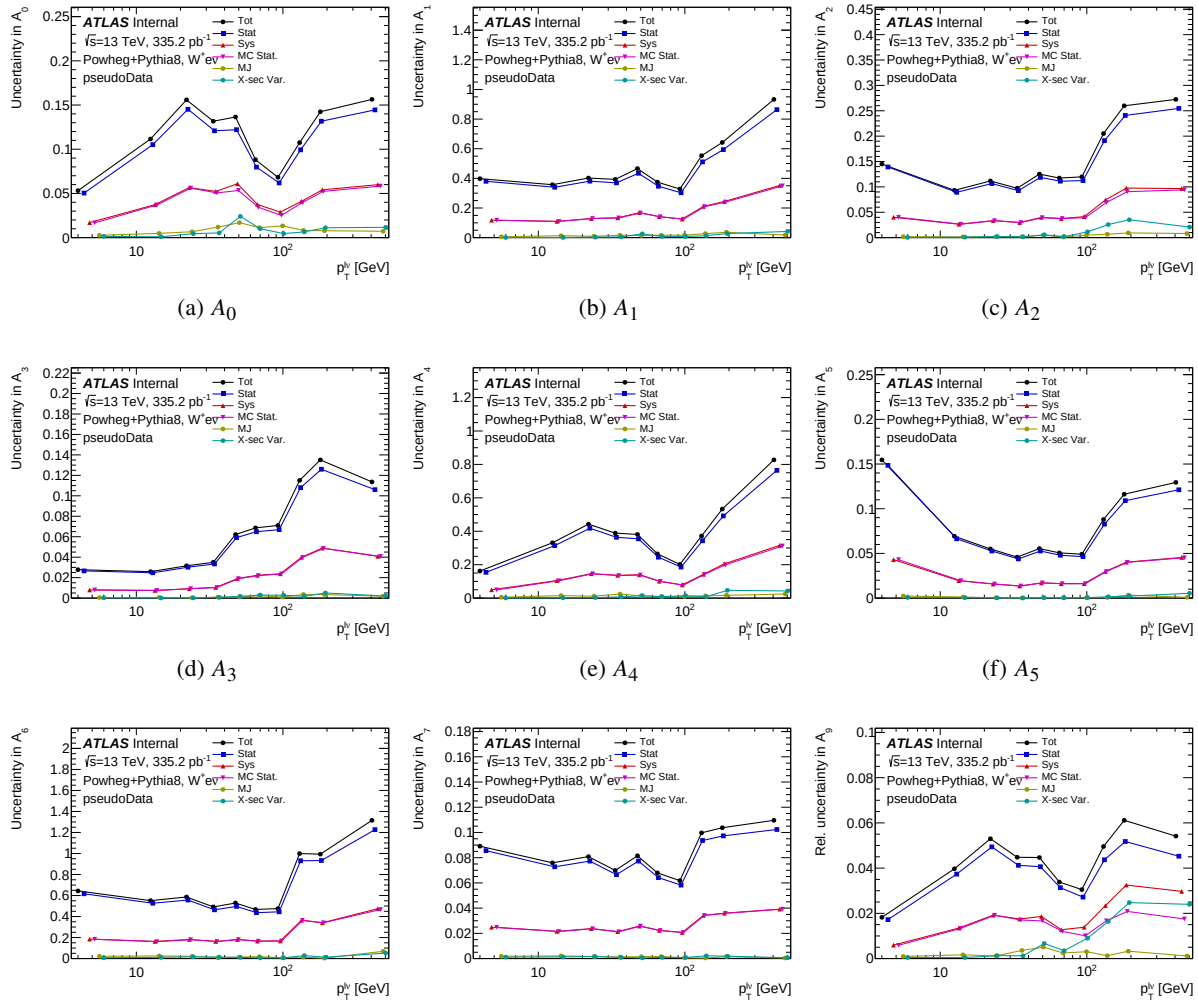


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 49: The uncertainty comparison for all contributions in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in p_{T}^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.

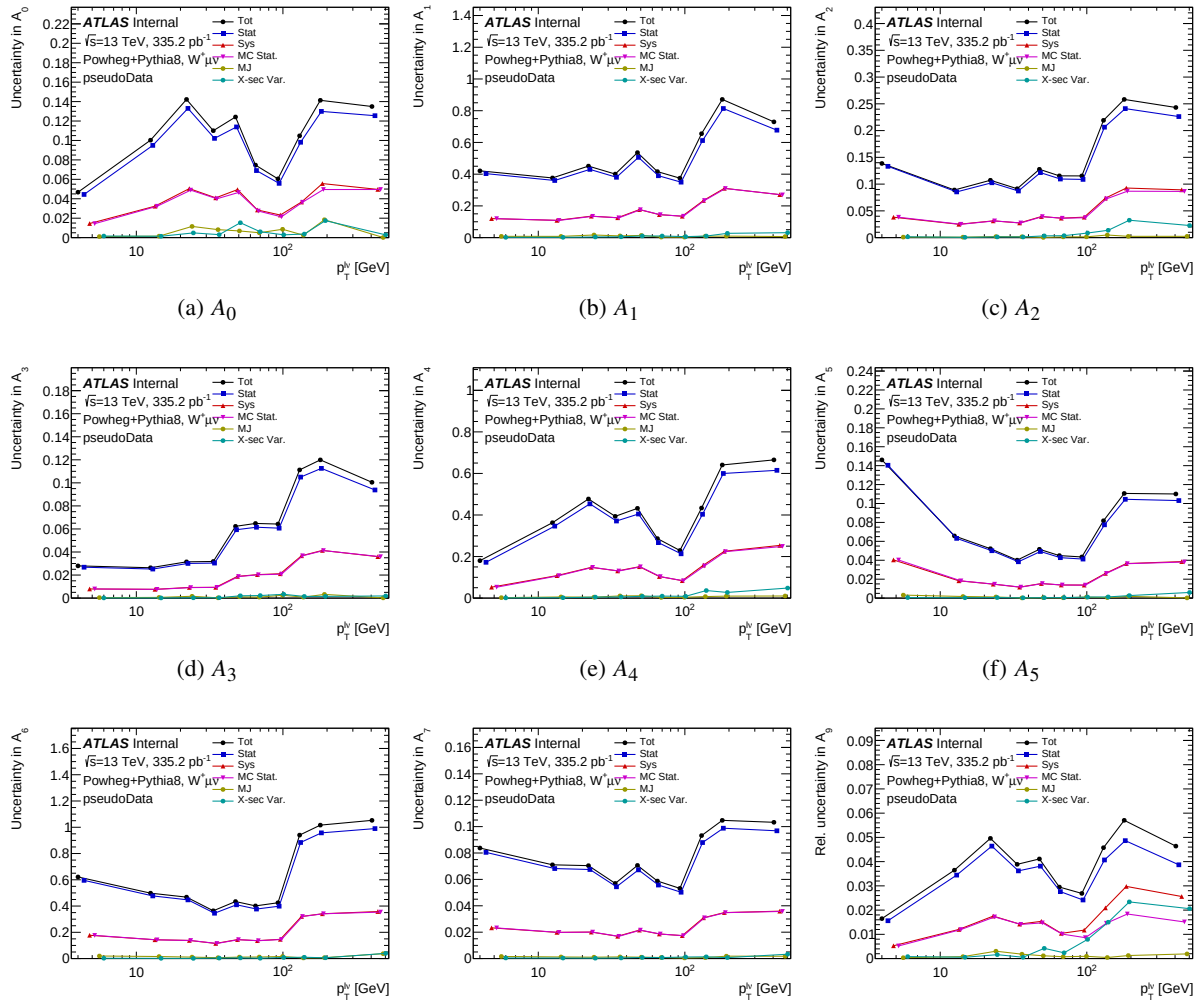


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 50: The uncertainty comparison for all contributions in the $W^+ \to e^+\nu$ channel for measurement in p_{T}^{W}. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.



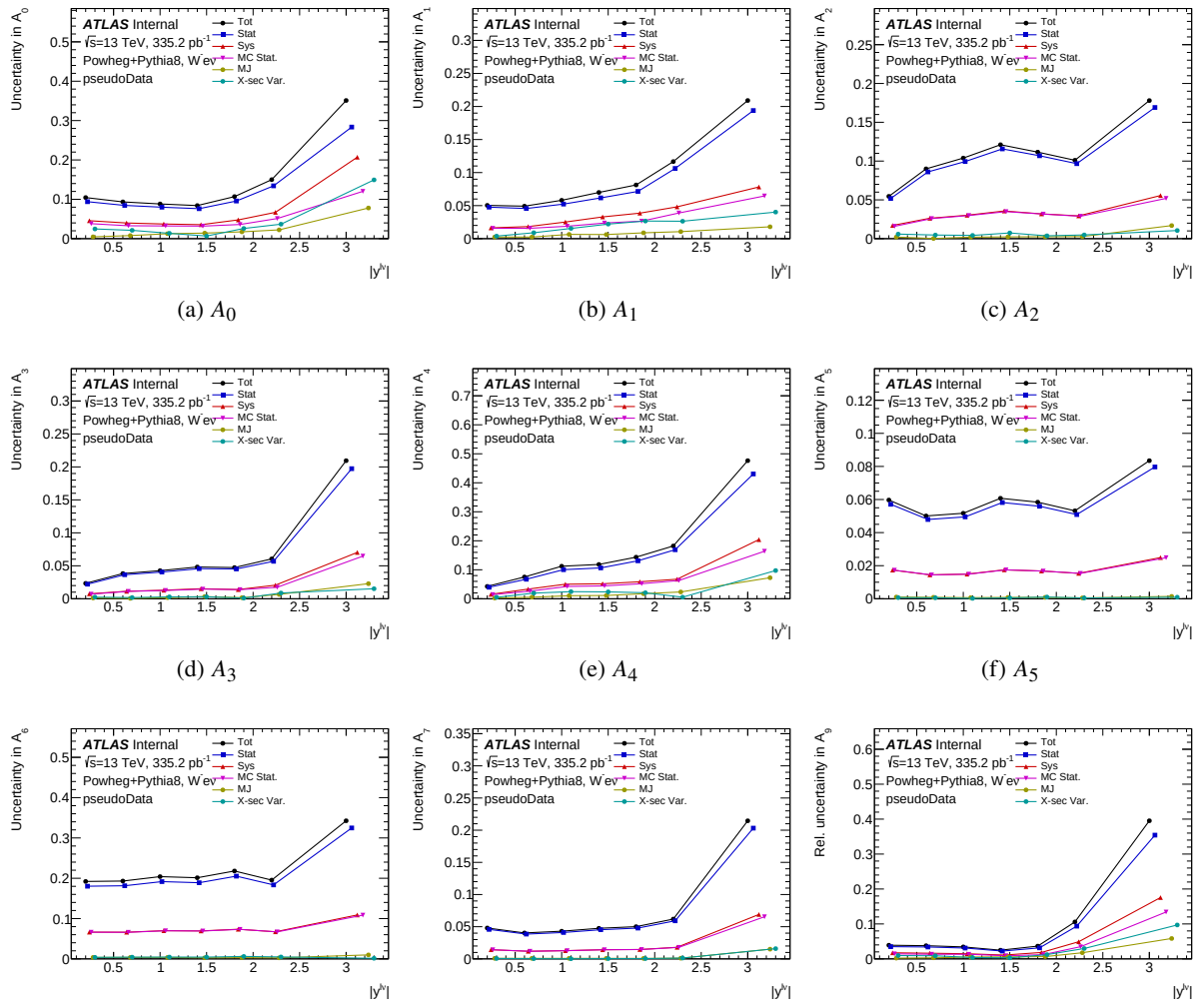
(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 51: The uncertainty comparison for all contributions in the $W^+ \to \mu^+\nu$ channel for measurement in p_{T}^{W}. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.

7.5.2 y^W Differential

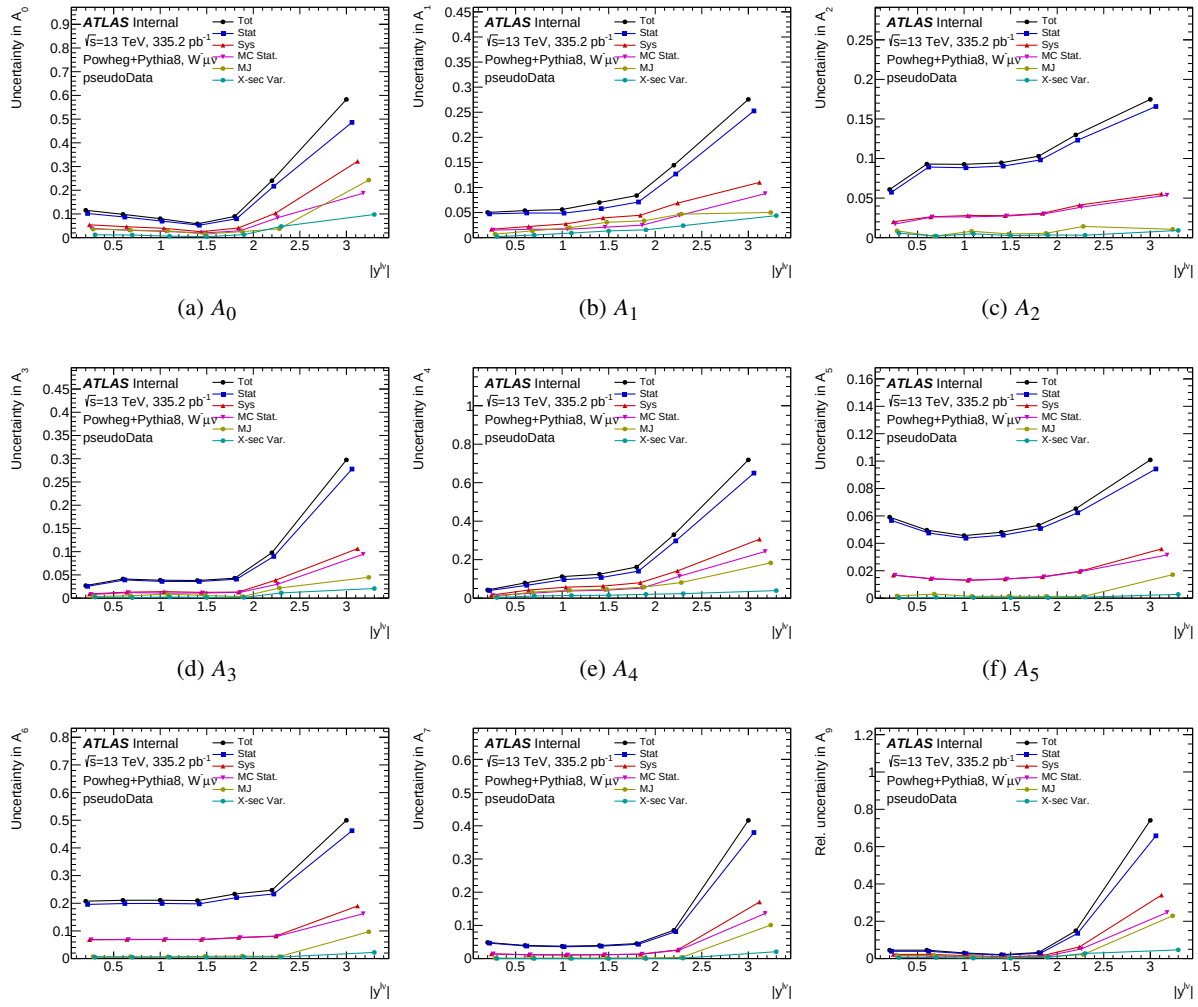


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 52: The uncertainty comparison for all contributions in the $W^- \to e^-\overline{\nu}$ channel for measurement in y^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.

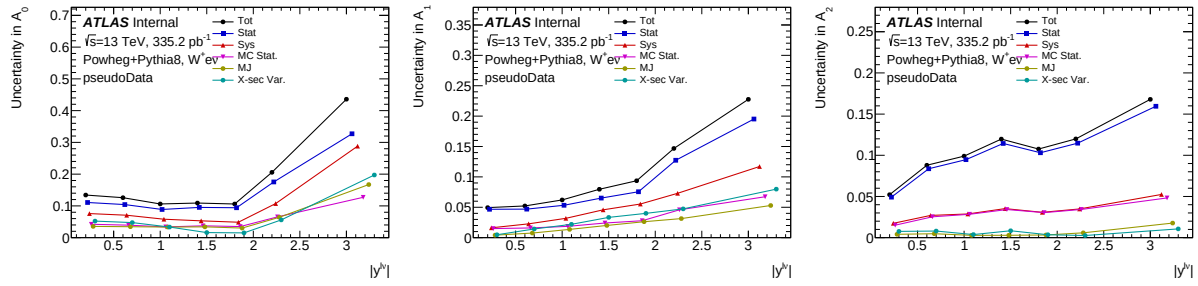


(a) A_0 (b) A_1 (c) A_2

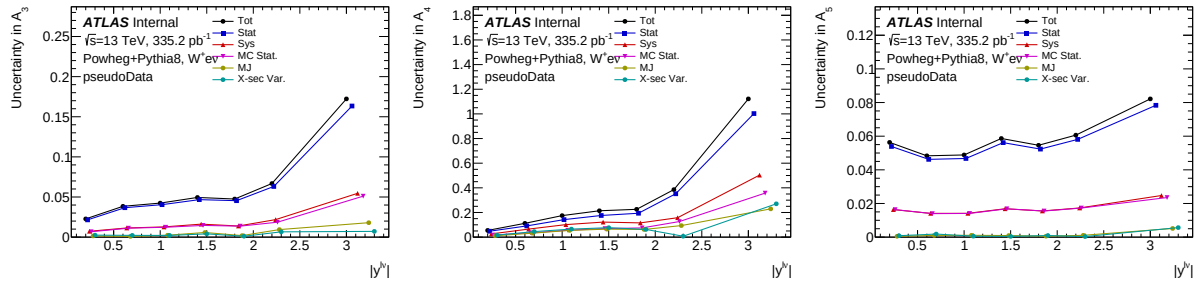
(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

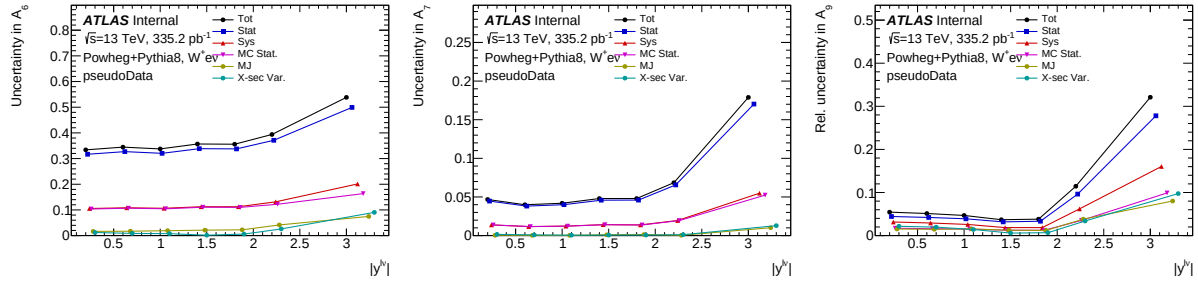
Figure 53: The uncertainty comparison for all contributions in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in y^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.



(a) A_0 (b) A_1 (c) A_2

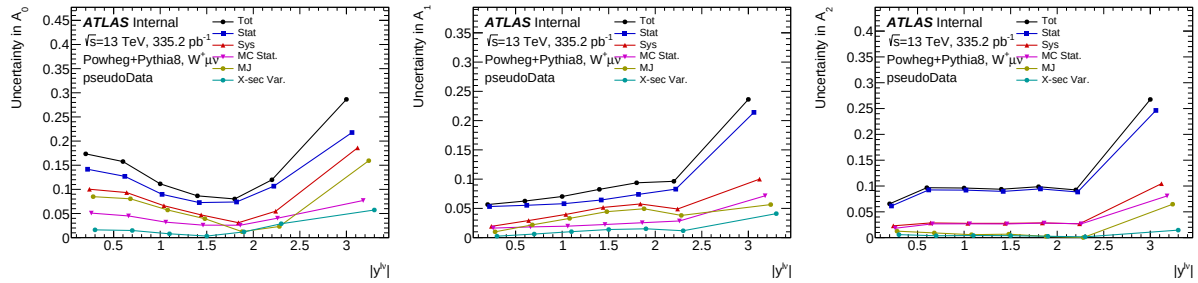


(d) A_3 (e) A_4 (f) A_5

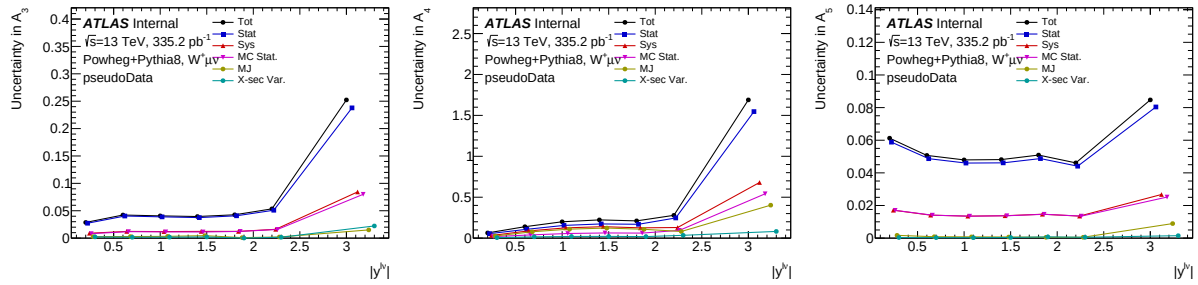


(g) A_6 (h) A_7 (i) A_9

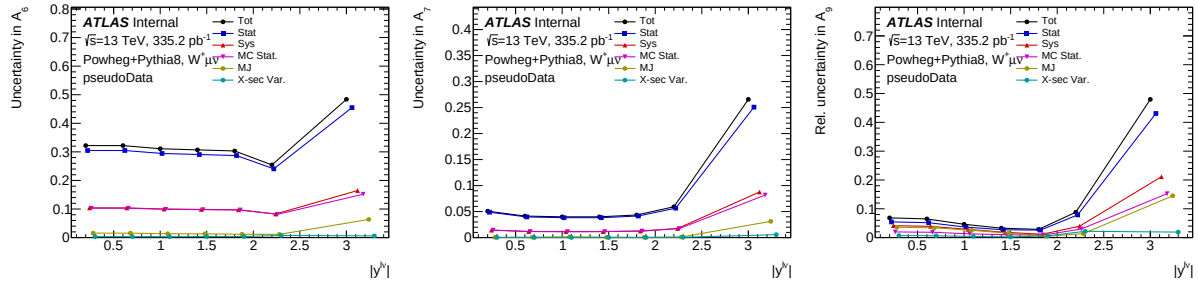
Figure 54: The uncertainty comparison for all contributions in the $W^+ \to e^+\nu$ channel for measurement in y^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.



(a) A_0 (b) A_1 (c) A_2



(d) A_3 (e) A_4 (f) A_5



(g) A_6 (h) A_7 (i) A_9

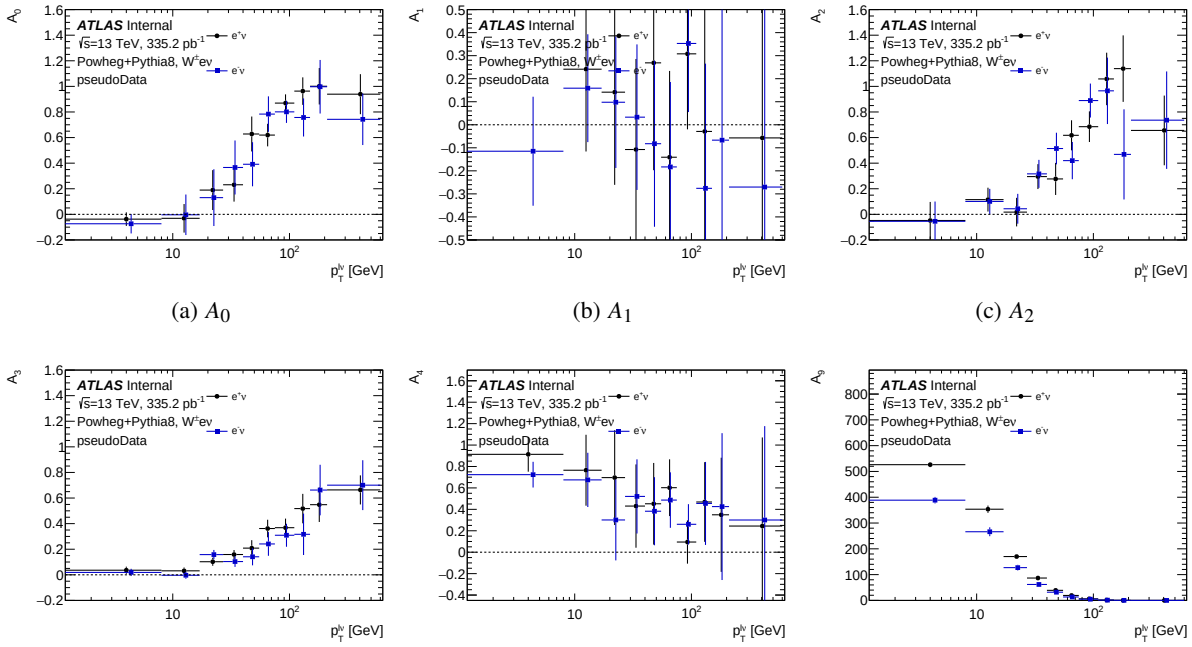
Figure 55: The uncertainty comparison for all contributions in the $W^+ \to \mu^+\nu$ channel for measurement in y^W. Blue is the total statistical uncertainty, red is the total systematic uncertainty, and black is the combination of the statistical and systematic uncertainties. The components of the systematic uncertainty are shown in pink, green, and teal to compare them to the statistical uncertainty directly.

8 Results

Results will be here once the coefficients are extracted using data. In the meantime results using pseudo data are currently used as placeholders.

8.1 Expected Results

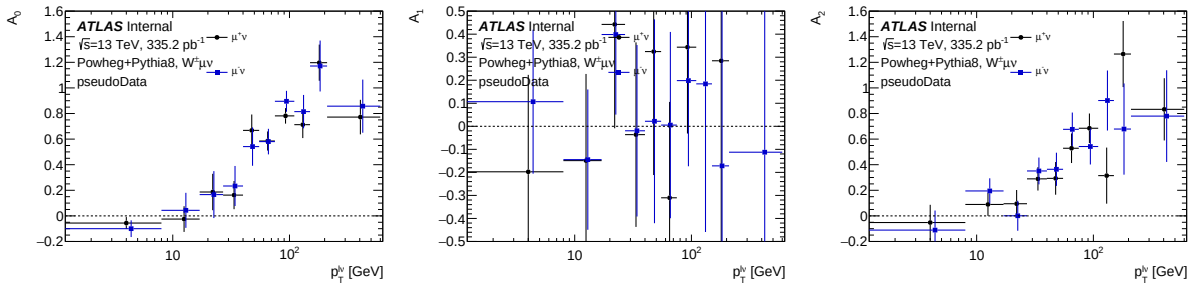
Here are the expected extracted results for the coefficients that are expected to be measured using pseudo data which is generated by randomly sampling the expected number events according to a Poisson distribution.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

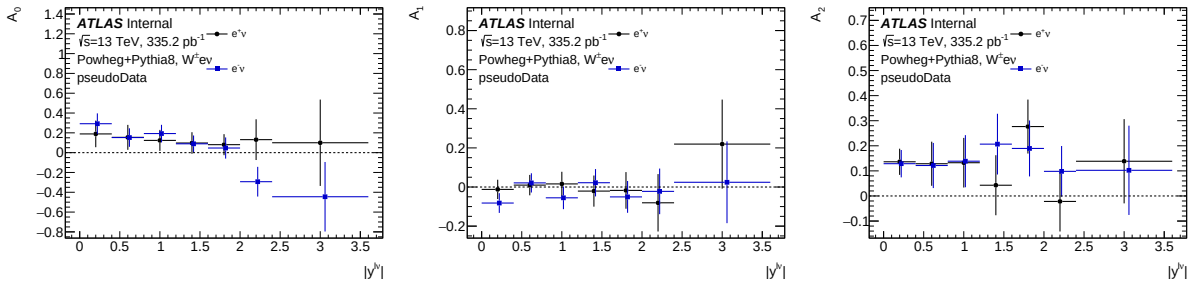
Figure 56: The coefficients extracted using pseudo data for channels $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ as a function of p_{T}^W. The errors represent the total uncertainty expected on the measurement.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

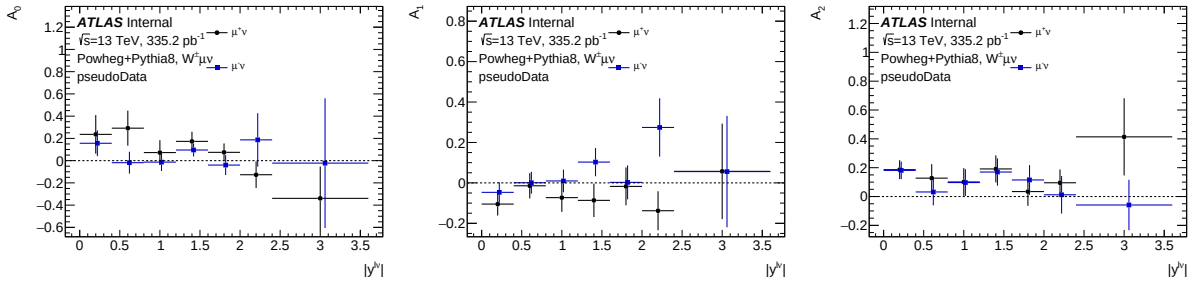
Figure 57: The coefficients extracted using pseudo data for channels $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ as a function of p_{T}^W. The errors represent the total uncertainty expected on the measurement.



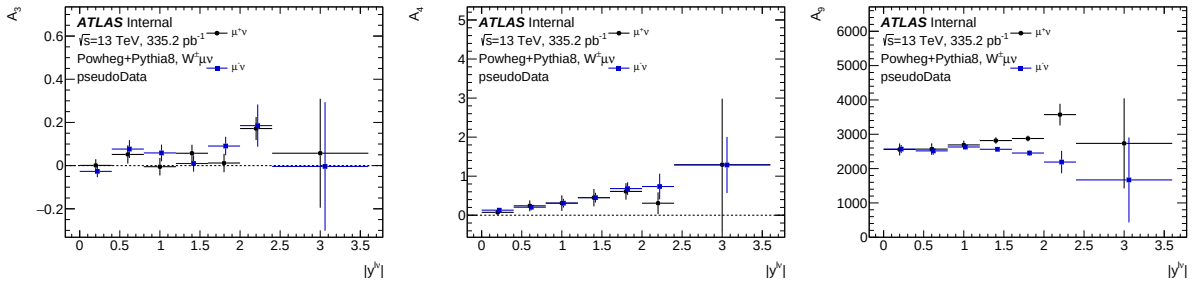
(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

Figure 58: The coefficients extracted using pseudo data for channels $W^- \to e^- \overline{\nu}$ and $W^+ \to e^+ \nu$ as a function of y^W. The errors represent the total uncertainty expected on the measurement.



(a) A_0 (b) A_1 (c) A_2



(d) A_3 (e) A_4 (f) A_9

Figure 59: The coefficients extracted using pseudo data for channels $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ as a function of y^W. The errors represent the total uncertainty expected on the measurement.

8.2 Channel Compatibility

Before extracting the measurements themselves for each channel, it is important to assess the compatibility between different channels in such a complex fitting procedure in terms of the reconstructed observables. The likelihood in Equation 19 was reparameterized in terms of new free parameters $\Delta A_{ij} = A_{ij}^{\mu\nu} - A_{ij}^{e\nu}$ and keeping the original $A_{ij}^{\mu\nu}$. The likelihood is maximized over this new set of parameters and uncertainties on ΔA_{ij} are extracted as described in Section 3.3. The expectation is that all the differences would be compatible with zero within the total uncertainty. This is performed in the p_{T}^{W} differential analysis in Section 8.2.1 for the differences $e^{+}\nu - \mu^{+}\nu$ and $e^{-}\nu - \mu^{-}\nu$ and in the y^{W} differential analysis in Section 8.2.2.

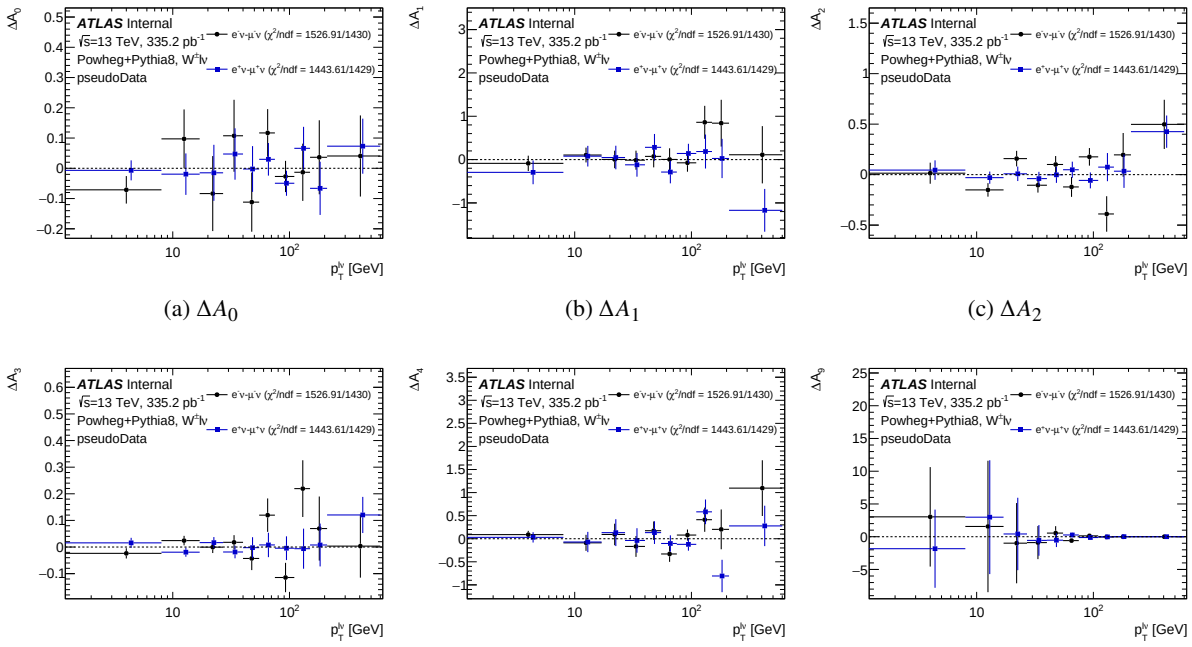
The observed compatibility between the channels is within the uncertainty of these measurements in most cases. It should be noted that the size of the statistical uncertainty on the differences between the channels is approximately twice the size of the expected statistical uncertainty on the combination between the two channels (one loses a factor $\sqrt{2}$ with respect to the individual channel statistical uncertainty when evaluating the difference between the two channels and one gains this same factor when combining them). The correlated systematic uncertainties are less important in these comparisons since these cancel, while the experimental ones are uncorrelated and add in quadrature between the two channels. The compatibility doesn't include the MC stat systematic but isn't needed as the channels are compatible just statistically.

To demonstrate the compatibility the goodness of fit is calculated which is summarized with the associated p-values in Tables 6.

Channel	χ^2**/ndf**	**p-value**
$e^{-}\nu - \mu^{-}\nu$, p_{T}^{W} differential	1526.91/1430	0.0371
$e^{+}\nu - \mu^{+}\nu$, p_{T}^{W} differential	1443.61/1429	0.3879
$e^{-}\nu - \mu^{-}\nu$, y^{W} differential	925.51/954	0.7401
$e^{+}\nu - \mu^{+}\nu$, y^{W} differential	931.64/952	0.6753

Table 6: Summary of goodness of fit and p-values for compatibility between channels.

8.2.1 p_{T}^{W} Differential $e^{+}\nu$ vs $\mu^{+}\nu$

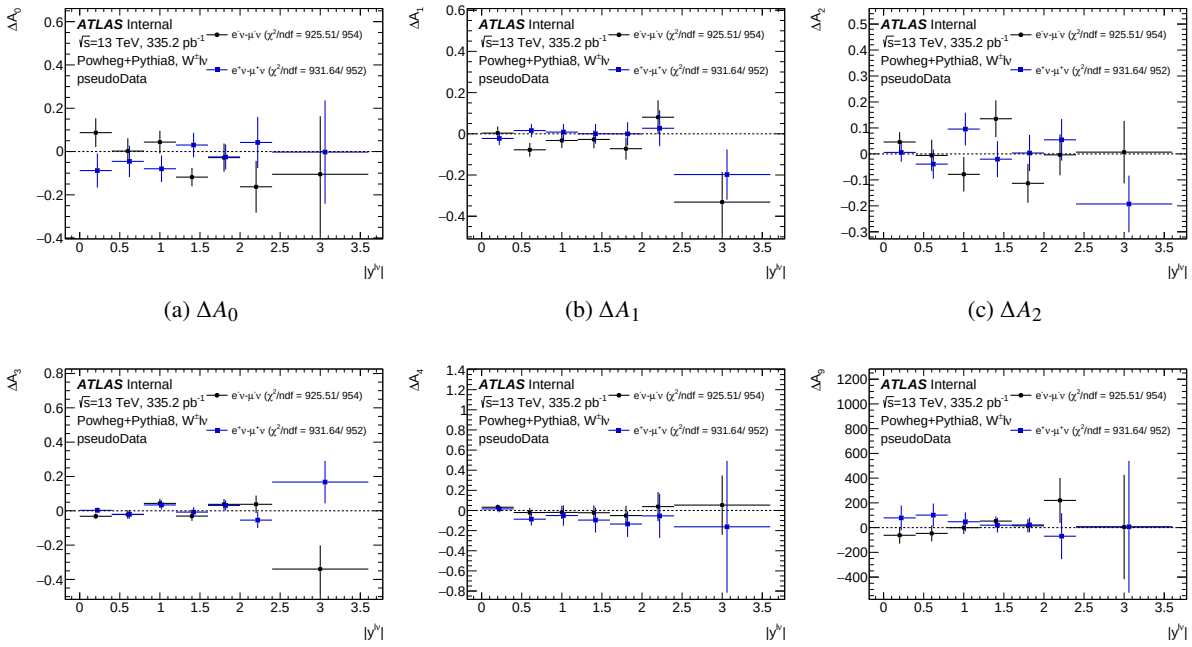


(a) ΔA_0 (b) ΔA_1 (c) ΔA_2

(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 60: Differences as a function of p_{T}^{W} between the $W^{-} \to e^{-}\overline{\nu}$ and $W^{-} \to \mu^{-}\overline{\nu}$ channels in black and $W^{+} \to e^{+}\nu$ and $W^{+} \to \mu^{+}\nu$ channels in blue for the measured coefficients. The points represent the measured difference using pseudo data in each p_{T}^{W} bin. The goodness of fit, χ^2/ndf, for all coefficients is also listed.

8.2.2 y^W Differential $e^+\nu$ vs $\mu^+\nu$



(a) ΔA_0 (b) ΔA_1 (c) ΔA_2

(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 61: Differences as a function of y^W between the $W^- \to e^-\overline{\nu}$ and $W^- \to \mu^-\overline{\nu}$ channels in black and $W^+ \to e^+\nu$ and $W^+ \to \mu^+\nu$ channels in blue for the measured coefficients. The points represent the measured difference using pseudo data in each y^W bin. The goodness of fit, χ^2/ndf, for all coefficients is also listed.

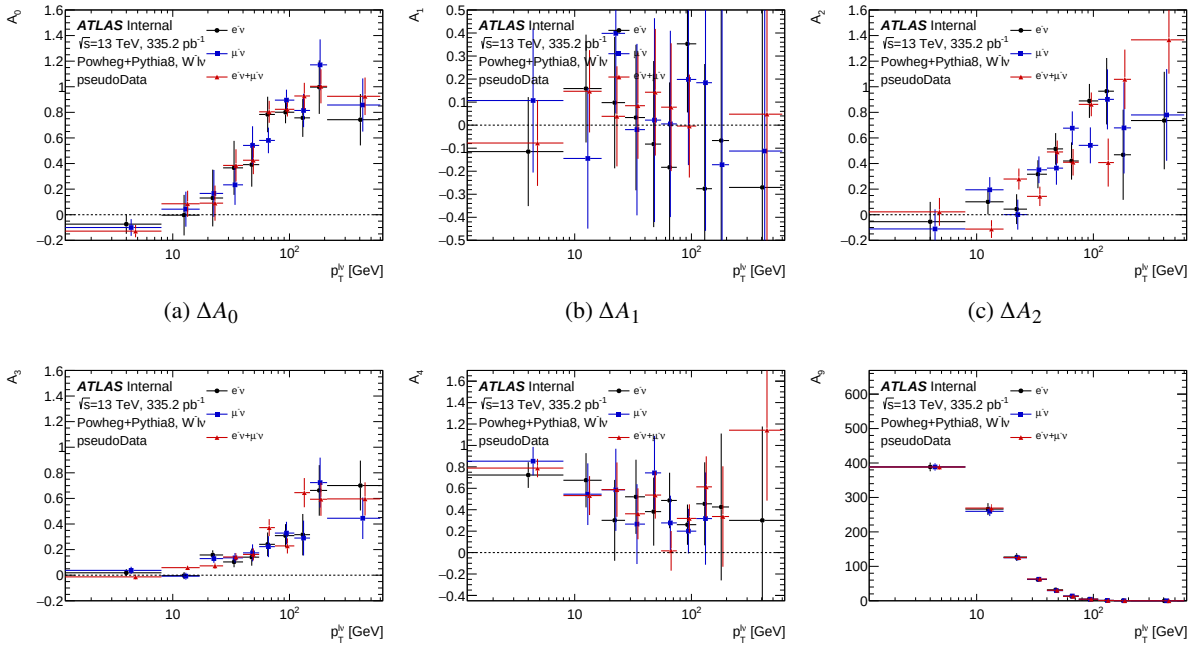
8.3 Channel Combination

With compatible channels we are able to combine them to gain a $\sim \sqrt{2}$ statistical sensitivity gain than either individual channel. The individual likelihoods are combined through a simple likelihood multiplication, keeping in mind that constraints on correlated systematics only keep one constraint term. Each likelihood can be decomposed into two types of terms: that which contains the observed data, denoted$\mathcal{L}_{\text{data}}$, and the auxiliary terms that constrain the nuisance parameters, θ, denoted $\mathcal{A}_i\left(\theta_i\right)$. The combined likelihood can be written as:

$$\mathcal{L}_{\text{cb}} = \left\{ \prod_{i}^{\text{channels}} \mathcal{L}_i\left(A, \theta | N_i\right) \right\} \left\{ \prod_{i}^{M} \mathcal{A}_i\left(\theta_i\right) \right\} \tag{54}$$

The combination results for p_{T}^{W} differential can be found in red in Figures 62 and 63 for $e^{-}\nu + \mu^{-}\nu$ and $e^{+}\nu + \mu^{+}\nu$ respectively as well as the results for the individual channels. Similarly the results for y^{W} differential can be found in Figures 64 and 65. The overall agreement between the combined and individual measurements can be more qualitatively grasped when presented in this way as opposed to the ΔA_{ij} measurement.

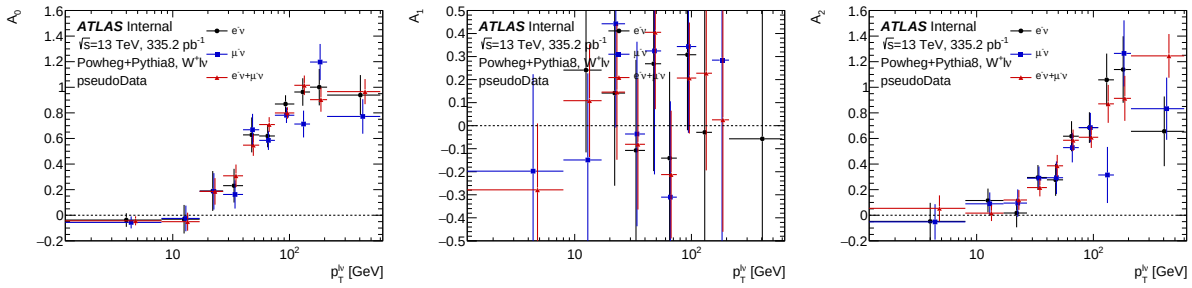
8.3.1 p_{T}^{W} Differential



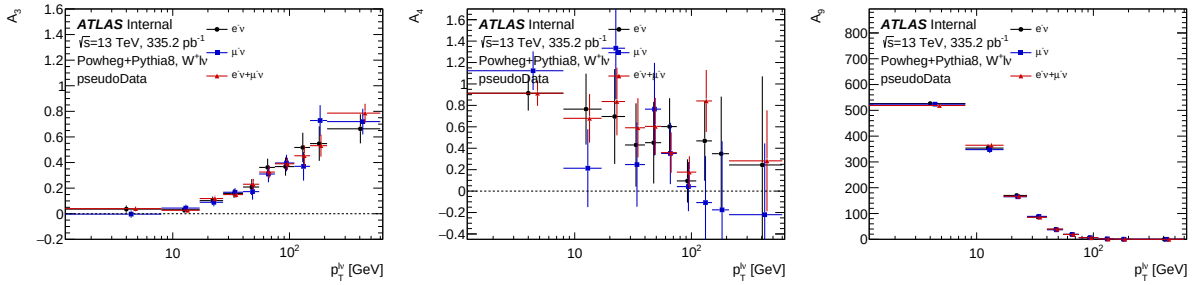
(a) ΔA_0 (b) ΔA_1 (c) ΔA_2

(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 62: Combined measurement of coefficients for W^{-} as a function of y^{W} differential in rapidity. The measured coefficients from the individual channels are overlayed in black and blue. The error bars represent the total uncertainty on the combined measurements.



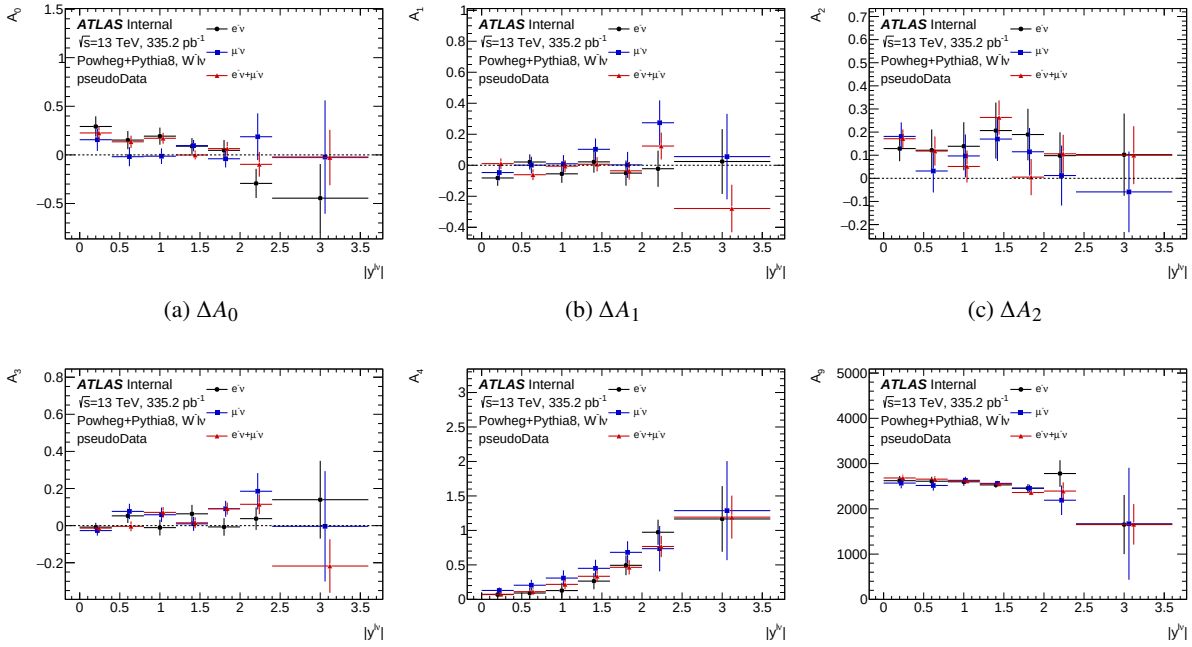
(a) ΔA_0 (b) ΔA_1 (c) ΔA_2



(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 63: Combined measurement of coefficients for W^+ as a function of y^W differential in rapidity. The measured coefficients from the individual channels are overlayed in black and blue. The error bars represent the total uncertainty on the combined measurements.

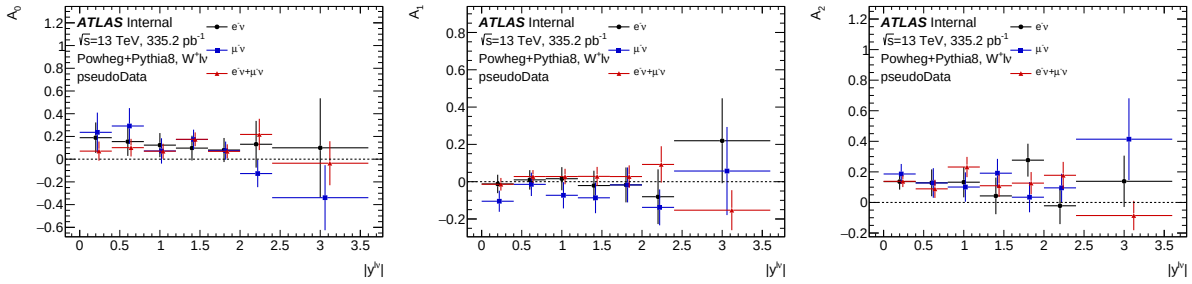
8.3.2 y^W Differential



(a) ΔA_0 (b) ΔA_1 (c) ΔA_2

(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 64: Combined measurement of coefficients for W^- as a function of p_{T}^W differential in transverse momentum. The measured coefficients from the individual channels are overlayed in black and blue. The error bars represent the total uncertainty on the combined measurements.



(a) ΔA_0 (b) ΔA_1 (c) ΔA_2

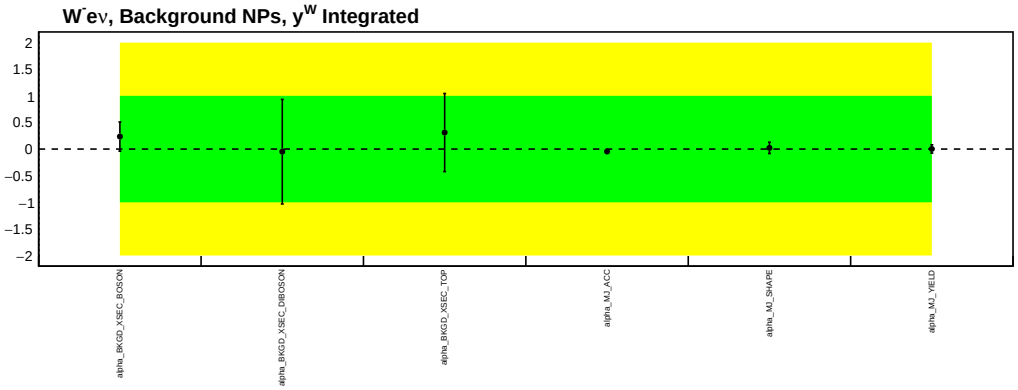
(d) ΔA_3 (e) ΔA_4 (f) ΔA_9

Figure 65: Combined measurement of coefficients for W^- as a function of p_{T}^{W} differential in transverse momentum. The measured coefficients from the individual channels are overlayed in black and blue. The error bars represent the total uncertainty on the combined measurements.

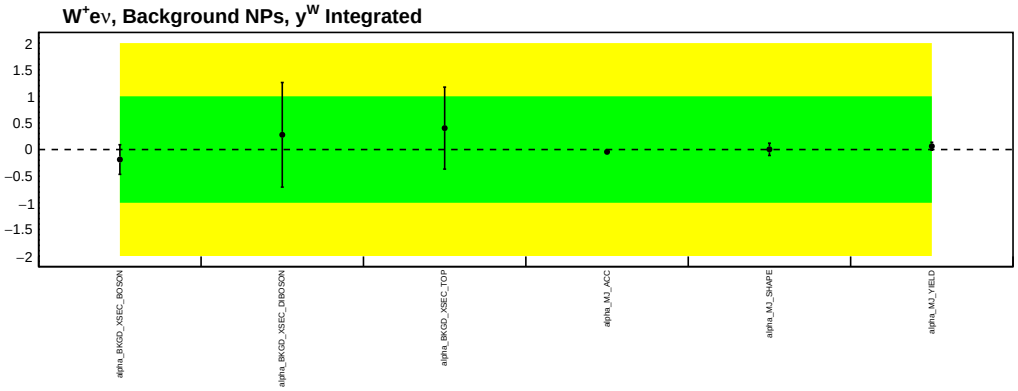
8.4 Nuisance Parameter Pulls

The best fit values of each nuisance parameter along with their post-fit constraints are presented here. Most parameters should have a fit value close to zero with a constraint close to unity. If a systematic variation does not mimic the PDF of the polynomials in the model, the floating coefficients won't absorb the variation and the result will be that the likelihood value changes in the variation to compensate. That is, the data will become incompatible with the PDF under the variation and the constraint on the nuisance parameter will be less than unity. In most cases this only happens by a small amount, but systematics that are over-conservative to begin with or which could even be measured using angular distributions can become significantly constrained. The pulls for each channel are shown for in the p_{T}^{W} differential measurement in Figures 66 - 69 and for the y^{W} differential measurement in Figures 70 - 73.

8.4.1 p_{T}^{W} Differential

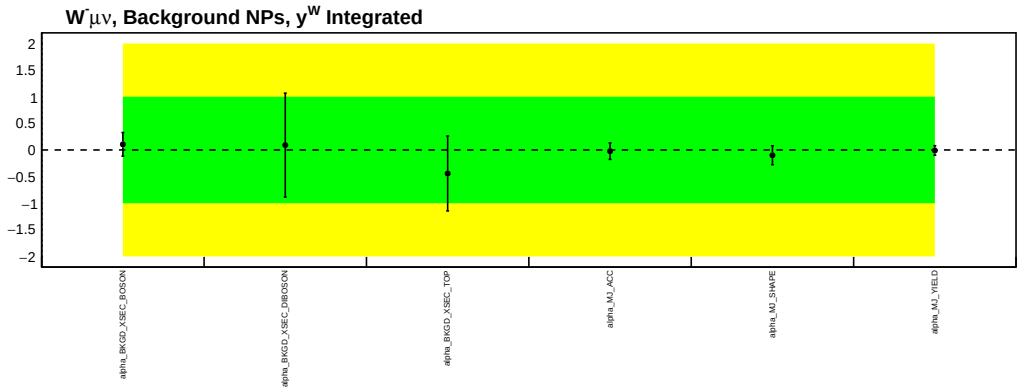


(a) $W^{-} \rightarrow e^{-}\overline{\nu}$ background NPs

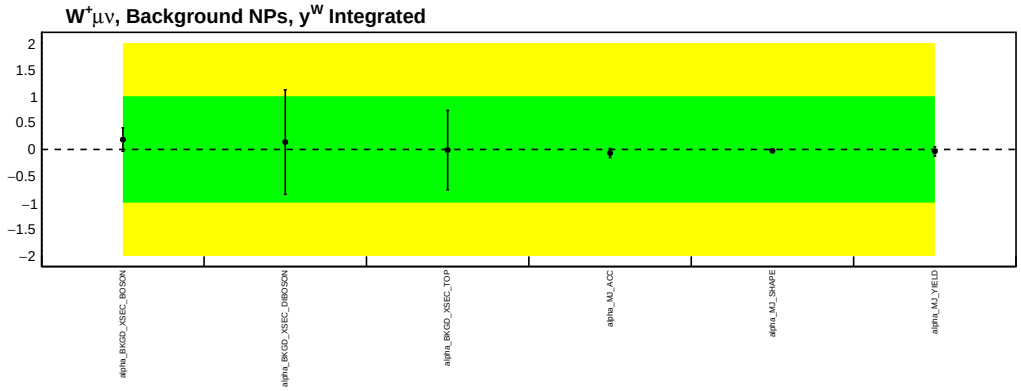


(b) $W^{+} \rightarrow e^{+}\nu$ background NPs

Figure 66: Best fit and post-fit constraint of the background nuisance parameters in the p_{T}^{W} differential configuration for $W^{-} \rightarrow e^{-}\overline{\nu}$ and $W^{+} \rightarrow e^{+}\nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

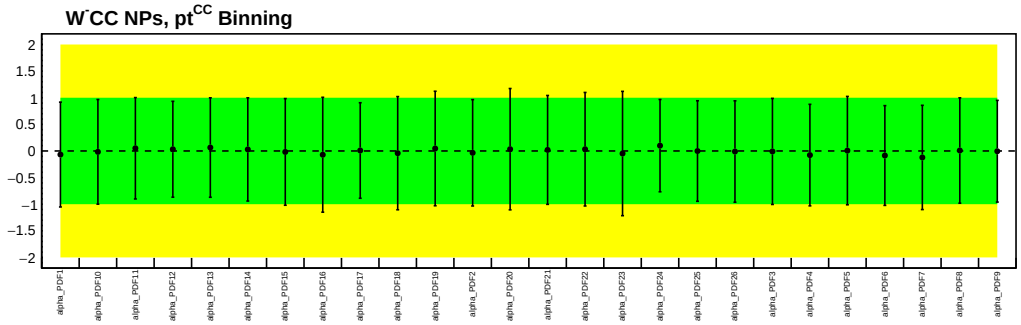


(a) $W^- \to \mu^- \overline{\nu}$ background NPs

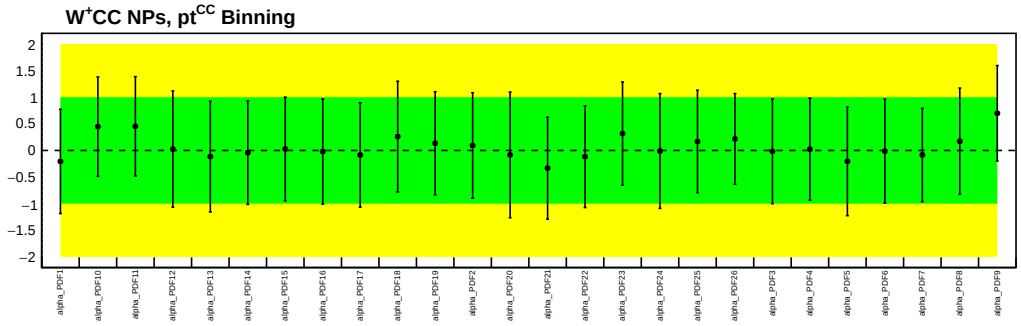


(b) $W^+ \to \mu^+ \nu$ background NPs

Figure 67: Best fit and post-fit constraint of the background nuisance parameters in the p_{T}^{W} differential configuration for $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

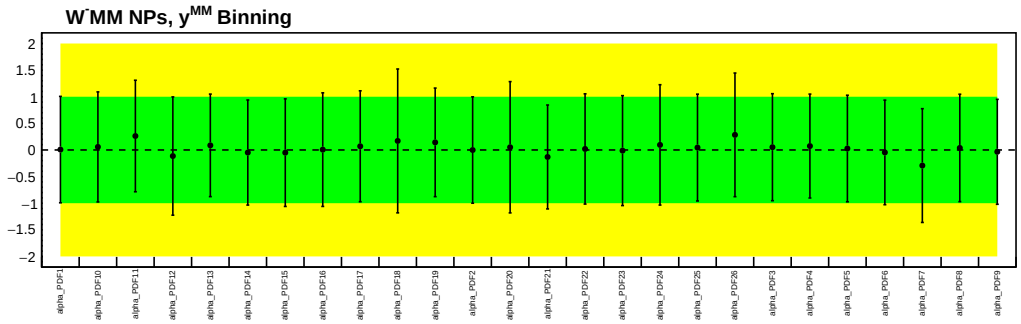


(a) $W^- \to e^-\overline{\nu}$ pdf NPs

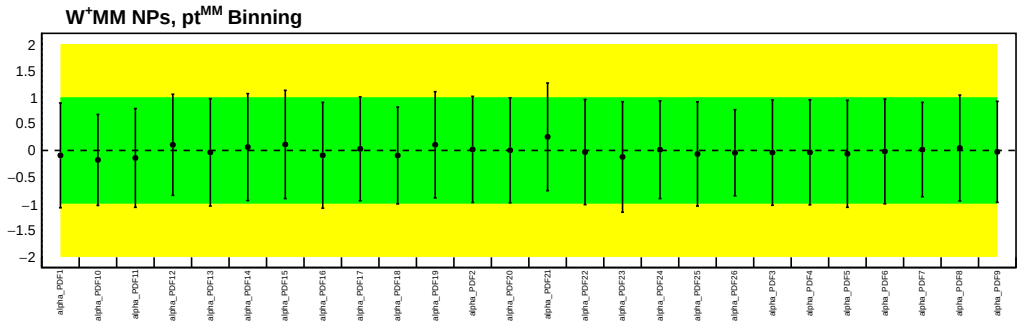


(b) $W^+ \to e^+\nu$ pdf NPs

Figure 68: Best fit and post-fit constraint of the pdf nuisance parameters in the p_{T}^{W} differential configuration for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.



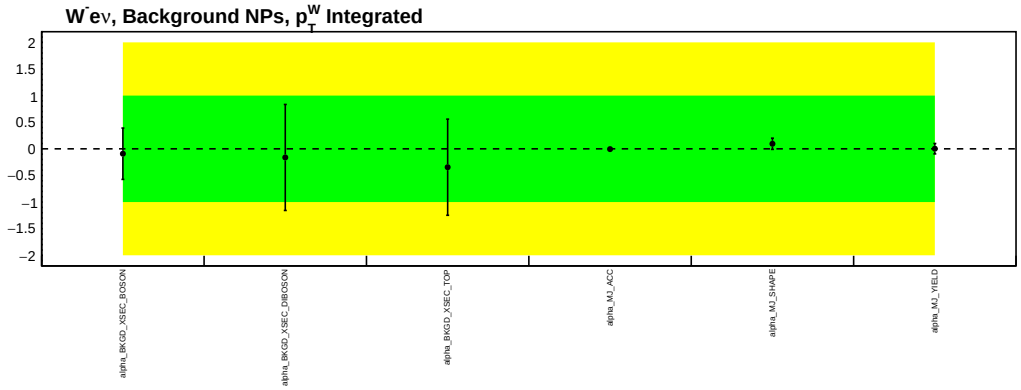
(a) $W^- \to \mu^- \overline{\nu}$ pdf NPs



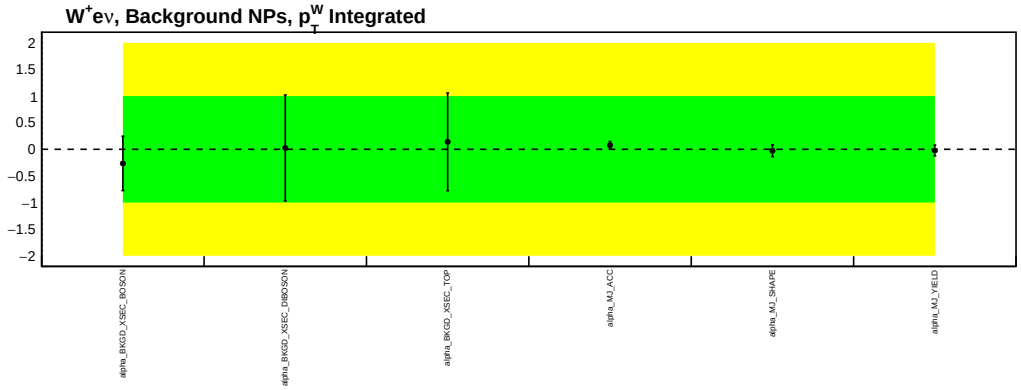
(b) $W^+ \to \mu^+ \nu$ pdf NPs

Figure 69: Best fit and post-fit constraint of the pdf nuisance parameters in the p_{T}^{W} differential configuration for $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

8.4.2 y^W Differential

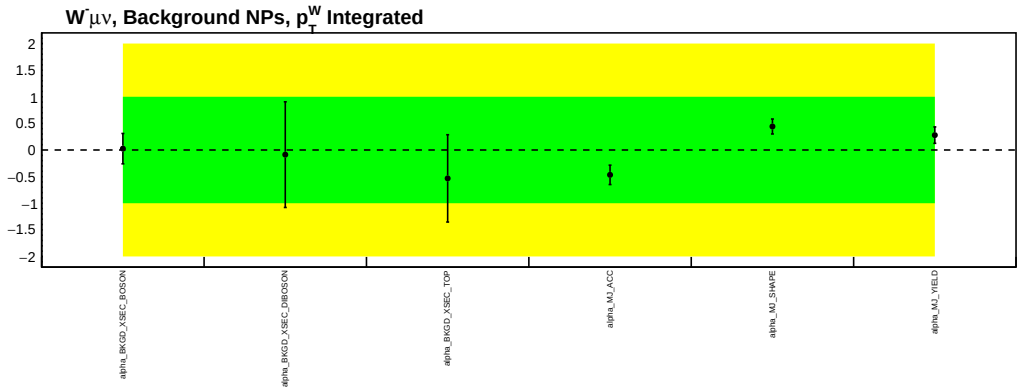


(a) $W^- \to e^-\overline{\nu}$ background NPs

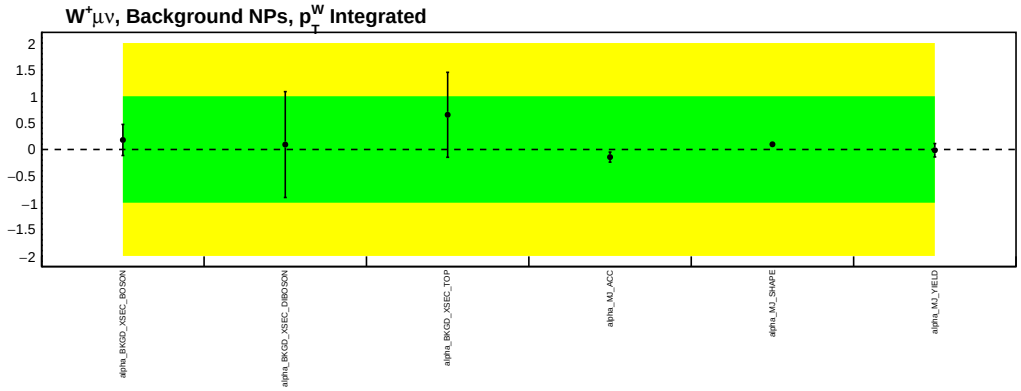


(b) $W^+ \to e^+\nu$ background NPs

Figure 70: Best fit and post-fit constraint of the background nuisance parameters in the y^W differential configuration for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

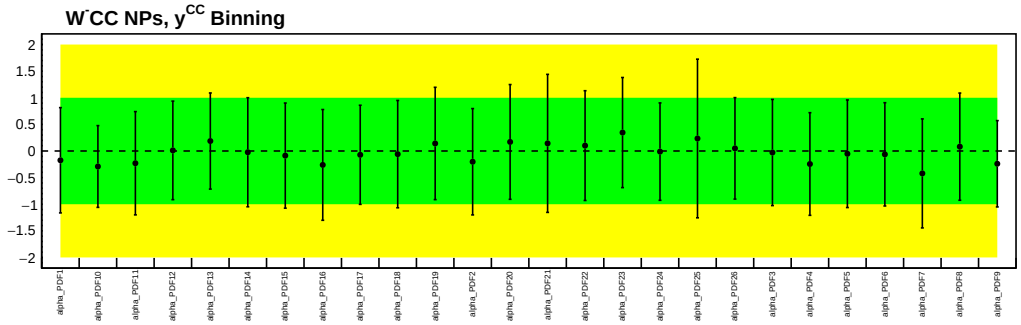


(a) $W^- \to \mu^- \overline{\nu}$ background NPs

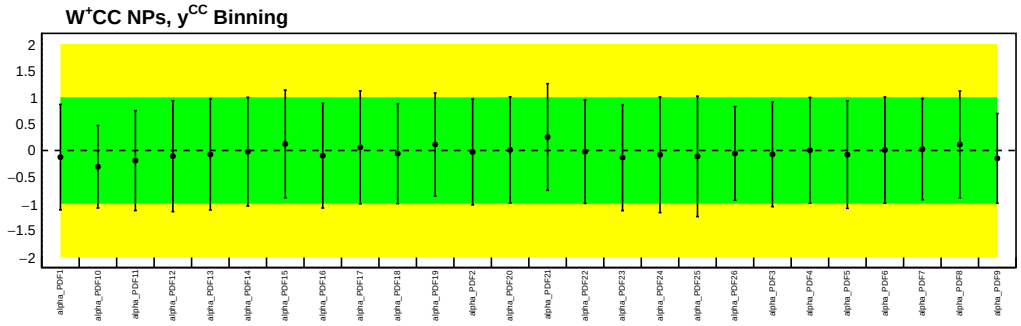


(b) $W^+ \to \mu^+ \nu$ background NPs

Figure 71: Best fit and post-fit constraint of the background nuisance parameters in the y^W differential configuration for $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

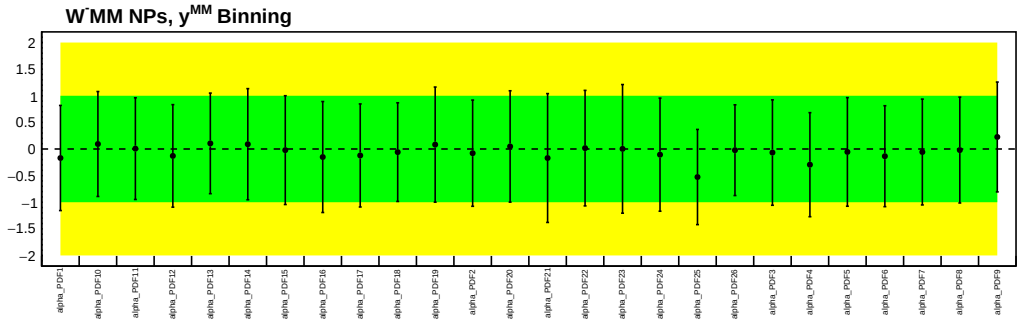


(a) $W^- \to e^-\overline{\nu}$ pdf NPs

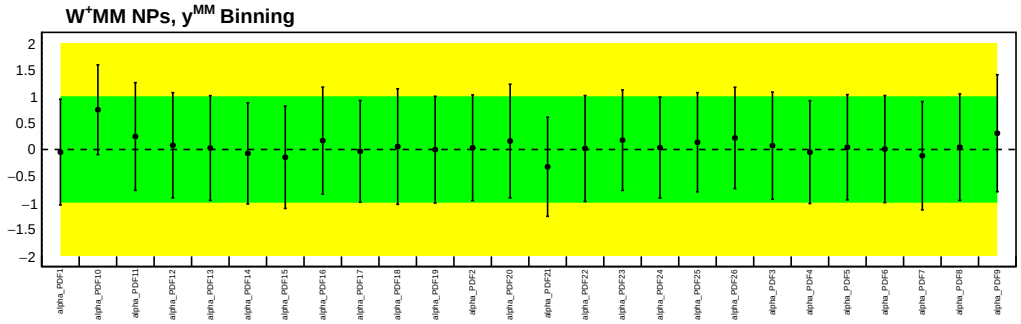


(b) $W^+ \to e^+\nu$ pdf NPs

Figure 72: Best fit and post-fit constraint of the pdf nuisance parameters in the y^W differential configuration for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.



(a) $W^- \to \mu^- \overline{\nu}$ pdf NPs



(b) $W^+ \to \mu^+ \nu$ pdf NPs

Figure 73: Best fit and post-fit constraint of the pdf nuisance parameters in the y^W differential configuration for $W^- \to \mu^- \overline{\nu}$ and $W^+ \to \mu^+ \nu$ channels. The green and yellow bands represent $\pm 1\sigma$ and $\pm 2\sigma$ from zero.

9 Comparison with Theoretical Predictions

This chapter will contain the comparison of results to theoretical predictions from Section 2 when those are compiled.

References

[1] E. Mirkes, *Angular decay distribution of leptons from W-bosons at NLO in hadronic collisions*, Nuclear Physics B **387** (1992) 3, ISSN: 0550-3213, URL: http://www.sciencedirect.com/science/article/pii/055032139290046E (cit. on pp. 5, 6, 9, 10, 29).

[2] E. Mirkes and J. Ohnemus, *W and Z polarization effects in hadronic collisions*, Phys. Rev. D **50** (9 1994) 5692, URL: http://link.aps.org/doi/10.1103/PhysRevD.50.5692 (cit. on pp. 5, 7, 9, 11).

[3] E. Mirkes and J. Ohnemus, *Angular distributions of Drell-Yan lepton pairs at the Tevatron: Order $\alpha - s^2$ corrections and Monte Carlo studies*, Phys. Rev. D **51** (1995) 4891, arXiv: hep-ph/9412289 (cit. on pp. 5, 6).

[4] E. Mirkes and J. Ohnemus, 'Polarization effects in Drell-Yan type processes h1 + h2 —> (W, Z, gamma*, J / psi) + x', *1994 Meeting of the American Physical Society, Division of Particles and Fields (DPF 94)*, 1994 1721, arXiv: hep-ph/9408402 (cit. on p. 5).

[5] R. Aaij et al., *First measurement of the $Z \to \mu^+\mu^-$ angular coefficients in the forward region of pp collisions at $\sqrt{s} = 13$ TeV*, (2022), arXiv: 2203.01602 [hep-ex] (cit. on p. 5).

[6] ATLAS Collaboration, *Measurement of the angular coefficients in Z-boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **08** (2016) 159, arXiv: 1606.00689 [hep-ex] (cit. on pp. 5, 7).

[7] ATLAS Collaboration, *Measurement of the effective leptonic weak mixing angle using electron and muon pairs from Z-boson decay in the ATLAS experiment at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2018-037, 2018, URL: https://cds.cern.ch/record/2630340 (cit. on p. 5).

[8] CMS Collaboration, *Angular coefficients of Z bosons produced in pp collisions at $\sqrt{s} = 8$ TeV and decaying to $\mu^+\mu^-$ as a function of transverse momentum and rapidity*, Phys. Lett. B **750** (2015) 154, arXiv: 1504.03512 [hep-ex] (cit. on p. 5).

[9] T. Aaltonen and others", *First Measurement of the Angular Coefficients of Drell-Yan e^+e^- Pairs in the Z Mass Region from $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **106** (24 2011) 241801, URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.241801 (cit. on p. 5).

[10] D. Acosta et al., *Measurement of the azimuthal angle distribution of leptons from W boson decays as a function of the W transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **73** (2006) 052002, arXiv: hep-ex/0504020 (cit. on p. 5).

[11] ATLAS Collaboration, *Measurement of the polarisation of W bosons produced with large transverse momentum in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment*, Eur. Phys. J. C **72** (2012) 2001, arXiv: 1203.2165 [hep-ex] (cit. on p. 5).

[12] CMS Collaboration, *Measurement of the Polarization of W Bosons with Large Transverse Momenta in W+Jets Events at the LHC*, Phys. Rev. Lett. **107** (2011) 021802, arXiv: 1104.3829 [hep-ex] (cit. on p. 5).

[13] *Measurement of the W boson rapidity, helicity, and differential cross sections in pp collisions at* $\sqrt{s}$ = 13 TeV, (2020) (cit. on p. 5).

[14] ATLAS Collaboration, *Measurement of the W-boson mass in pp collisions at* $\sqrt{s}$ = 7 *TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 110, arXiv: 1701.07240 [hep-ex] (cit. on p. 5), Erratum: Eur. Phys. J. C **78** (2018) 898.

[15] R. Aaij et al., *Measurement of the W boson mass*, JHEP **01** (2022) 036, arXiv: 2109.01113 [hep-ex] (cit. on p. 5).

[16] T. Aaltonen et al., *High-precision measurement of the ¡i¿W¡/i¿ boson mass with the CDF II detector*, Science **376** (2022) 170, eprint: https://www.science.org/doi/pdf/10.1126/science.abk1781, URL: https://www.science.org/doi/abs/10.1126/science.abk1781 (cit. on p. 5).

[17] J. Collins and D. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, Phys. Rev. D **16** (1977) 2219, URL: http://link.aps.org/doi/10.1103/PhysRevD.16.2219 (cit. on p. 6).

[18] C. M. Carloni Calame, G. Montagna, O. Nicrosini and A. Vicini, *Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders*, JHEP **10** (2007) 109, arXiv: 0710.1722 [hep-ph] (cit. on p. 7).

[19] M. Pellen, R. Poncelet, A. Popescu and T. Vitos, *Angular coefficients in W+j production at the LHC with high precision*, (2022), arXiv: 2204.12394 [hep-ph] (cit. on p. 9).

[20] C. Lam and W. Tung, *A Systematic Approach to Inclusive Lepton Pair Production in Hadronic Collisions*, Phys. Rev. D **18** (1978) 2447 (cit. on p. 11).

[21] A. J. Armbruster et al., *Methodology of angular coefficient and full solid angle cross-section analysis*, tech. rep., CERN, 2018, URL: https://cds.cern.ch/record/2314233 (cit. on p. 12).

[22] G. Guennebaud, B. Jacob et al., *Eigen v3*, http://eigen.tuxfamily.org, 2010 (cit. on p. 22).

[23] J. Kretzschmar, F. Balli and R. Stroehmer, *Samples and Physics modelling for low pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657141 (cit. on p. 24).

[24] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, Journal of High Energy Physics **2004** (2004) 040, URL: https://doi.org/10.1088/1126-6708/2004/11/040 (cit. on p. 26).

[25] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, Journal of High Energy Physics **2007** (2007) 070, URL: https://doi.org/10.1088/1126-6708/2007/11/070 (cit. on p. 26).

[26] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO vector-boson production matched with shower in POWHEG*, Journal of High Energy Physics **2008** (2008) 060, URL: https://doi.org/10.1088/1126-6708/2008/07/060 (cit. on p. 26).

[27] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010** (2010) 43, URL: https://doi.org/10.1007/JHEP06(2010)043 (cit. on p. 26).

[28] M. Guzzi et al., *CT10 parton distributions and other developments in the global QCD analysis*, (2011), arXiv: 1101.0561 [hep-ph] (cit. on p. 26).

[29] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Computer Physics Communications **178** (2008) 852, ISSN: 0010-4655, URL: https://www.sciencedirect.com/science/article/pii/S0010465508000441 (cit. on p. 26).

[30] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv: 1406.3660 [hep-ex] (cit. on p. 26).

[31] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006) 97, arXiv: hep-ph/0506026 (cit. on p. 26).

[32] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 26).

[33] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965 (cit. on p. 26).

[34] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107 (cit. on p. 26).

[35] S. Catani and M. Grazzini, *Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-boson Production at the Large Hadron Collider*, Phys. Rev. Lett. **98** (2007) 222002, arXiv: hep-ph/0703012 [hep-ph] (cit. on p. 26).

[36] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order*, Phys. Rev. Lett. **103** (8 2009) 082001, URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.082001 (cit. on p. 26).

[37] L. Harland-Lang, A. Martin, P. Motylinski and R. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: 1412.3989 [hep-ph] (cit. on p. 26).

[38] ATLAS Collaboration, *Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. B **759** (2016) 601, arXiv: 1603.09222 [hep-ex] (cit. on p. 26).

[39] H. Atmani et al., *Measurment of the p_T spectrum of W- and Z-bosons produced in pp collisions at $\sqrt{s}$ = 5 TeV and 13 TeV in low-pileup runs*, tech. rep., CERN, 2018, URL: https://cds.cern.ch/record/2632159 (cit. on p. 26).

[40] T. Xu, H. Atmani and L. Aperio Bella, *Electron corrections for low pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657152 (cit. on p. 27).

[41] A. Sydorenko, J. A. Kremer and T. Xu, *Muon corrections for low pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657116 (cit. on p. 27).

[42] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on p. 27).

[43] M. Li et al., *Hadronic recoil reconstruction and calibration for low pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657182 (cit. on pp. 28, 137).

[44] T. Xu et al., *Multi-jet background in low-pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657146 (cit. on pp. 31, 32, 35, 144, 150).

[45] J. Benitez et al., *Search for new resonances decaying to a W or Z boson and a Higgs boson in the $\ell\ell b\bar{b}$, $\ell\nu b\bar{b}$, and $\nu\nu b\bar{b}$ channels in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, tech. rep., CERN, 2015, URL: https://cds.cern.ch/record/2054042 (cit. on p. 107).

[46] M. Li et al., *Hadronic recoil reconstruction and calibration for low pile-up runs taken in 2017 and 2018*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2657182 (cit. on pp. 118, 119).

[47] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 166).

List of contributions

Ludovica Aperio Bella	Analysis design, initial sensitivity studies, Ntuple generation, hadronic recoil studies
Aaron Armbruster	Overall fitting methodology
Alexander Bachiu	Initial sensitivity studies, identification and isolation scale factors, coefficient extraction, systematic uncertainties
Stefano Camarda	Analysis design, neutrino reconstruction, initial sensitivity studies, theoretical predictions
Nicolo De Groot	Advising
Ruth Magdalena Jacobs	Initial Multi-jet studies, neutrino reconstruction studies
Daniil Ponomarenko	Multi-jet studies and generation. Implementation of the multijet analysis (fraction fits, anti-isolation scan; extrapolation to signal region and different W-Ai measurement bins; closure tests; systematics).
Manuella Vincter	Advising
Grigorii Tolkachev	PDF systematic uncertainties

Appendices

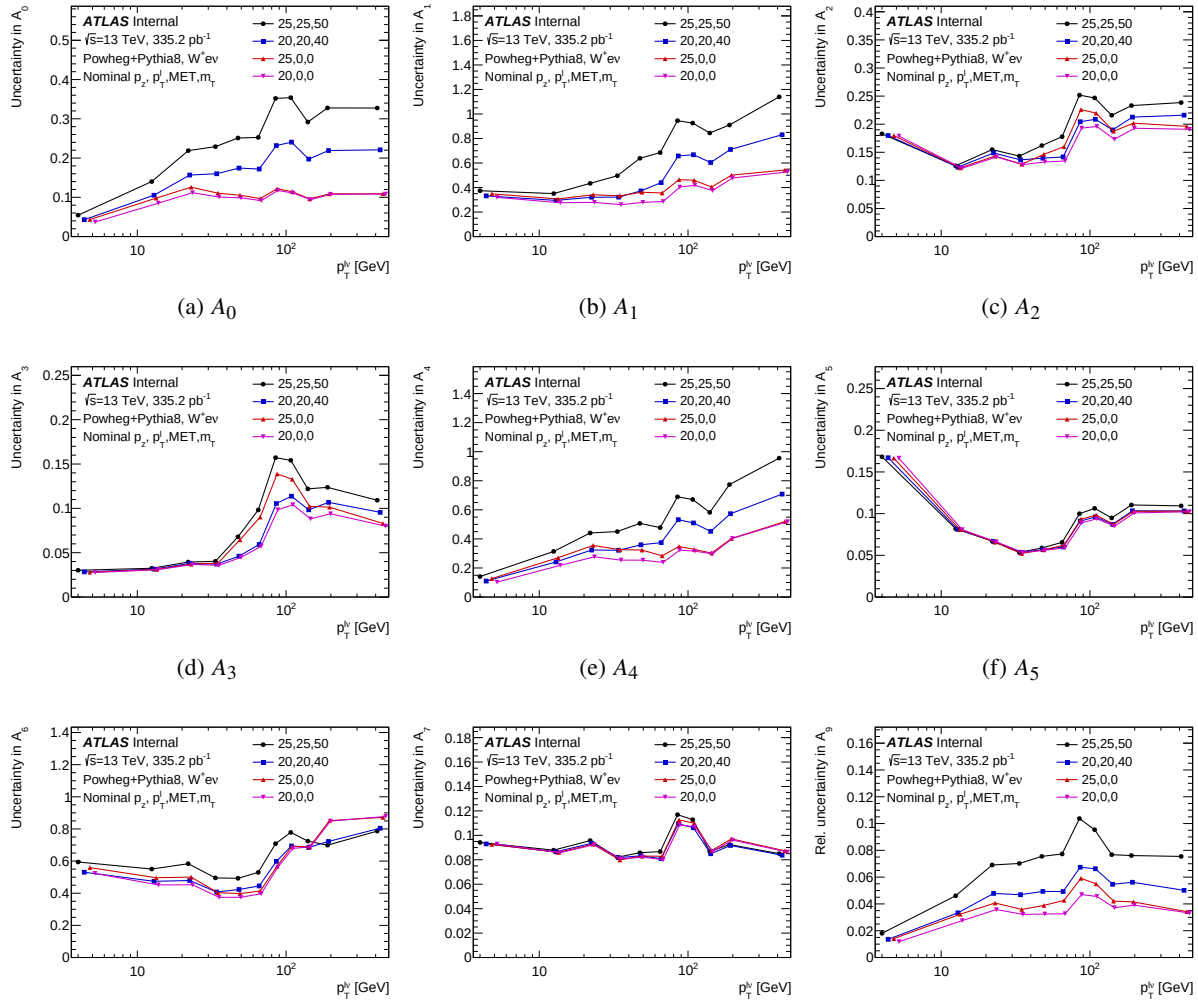
A Sensitivity studies

Given the limited luminosity recorded by ATLAS with low pile-up condition during the Run-2 data-taking (335.180pb^{-1} at 13 TeV), in order to increase the overall analysis sensitivity, different sensitivity studies were performed to maximise that statistical significance of the measurement. In the following the optimisation of the choice is kinematic cuts, identification and isolation criterion selection on the lepton are shown together of detailed studies on how to handle events where there is no solution for the neutrino longitudinal momentum.

A.1 Kinematic cuts choice

One method to gain statistical sensitivity is by relaxing the cuts on kinematics to get a larger acceptance of events. However, these cuts are in place to reduce the amount of background which will need to be addressed in another way. This is discussed in the Section A.2 where tighter constraints are applied on the identification and isolation. This study was to see whether there is a meaningful gain with relaxed sensitivity and where in the distributions the gain is. It used the nominal p_z^ν solution method as discussed in Section A.3 but was repeated using the rotation p_z^ν methodology with no relative difference in results. The study did not include background and was done to see the impact on just the signal statistical uncertainty.

There are three main kinematic cuts that were varied p_{T}^{ℓ}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and m_T which is their order in Figures 74 and 75. The most dominant for this analysis is the MJ and the kinematic cuts are one way of reducing it as it mostly is at low $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_T. As seen in Figure 74, reducing both of these cuts increases the statistical sensitivity by a large amount especially at high $p_{\mathrm{T}}^{\ell\nu}$ where the analysis is limited. The MJ is predominantly at low $p_{\mathrm{T}}^{\ell\nu}$ so it does not interfere with this statistical gain. The last thing to note was there was not a significant enough gain by adding events $20 \leq p_{\mathrm{T}}^{\ell} \leq 25$GeV to warrant the gain in MJ.

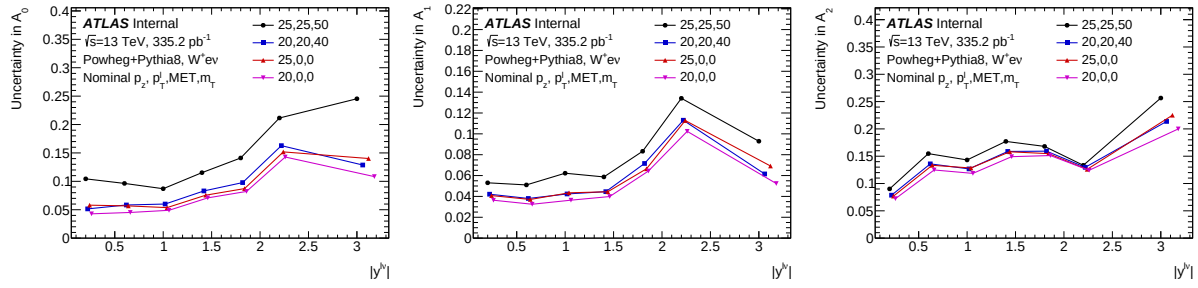


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

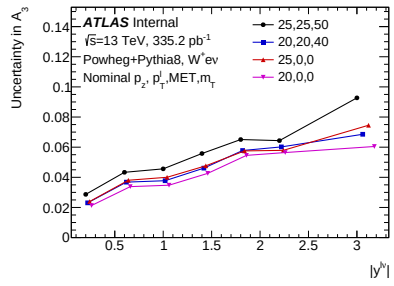
Figure 74: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for different selection criteria in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^+ \to e^+\nu$. The legend lists the various selection criteria in GeVfor p_{T}^{ℓ}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and m_T respectively; for example in black is when $p_{\mathrm{T}}^{\ell} > 25\mathrm{GeV}$, $E_{\mathrm{T}}^{\mathrm{miss}} > 25\mathrm{GeV}$, and m_T ¿ 50 GeV. This study was done using the nominal p_{z}^{ν} choice from the study in Section A.3.



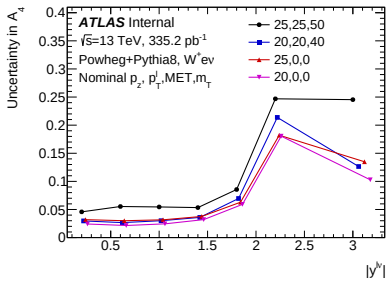
(a) A_0

(b) A_1

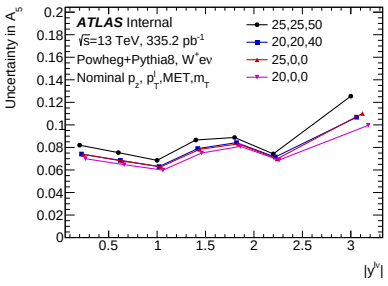
(c) A_2



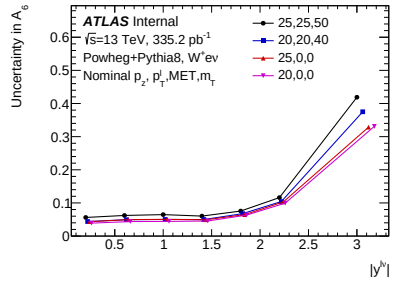
(d) A_3



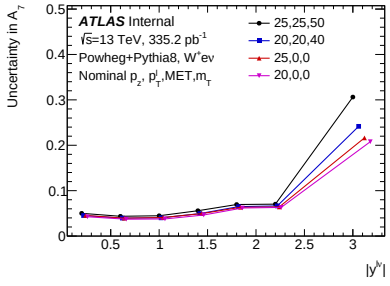
(e) A_4



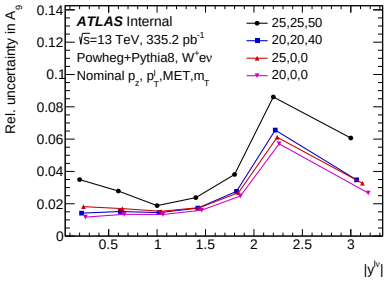
(f) A_5



(g) A_6



(h) A_7

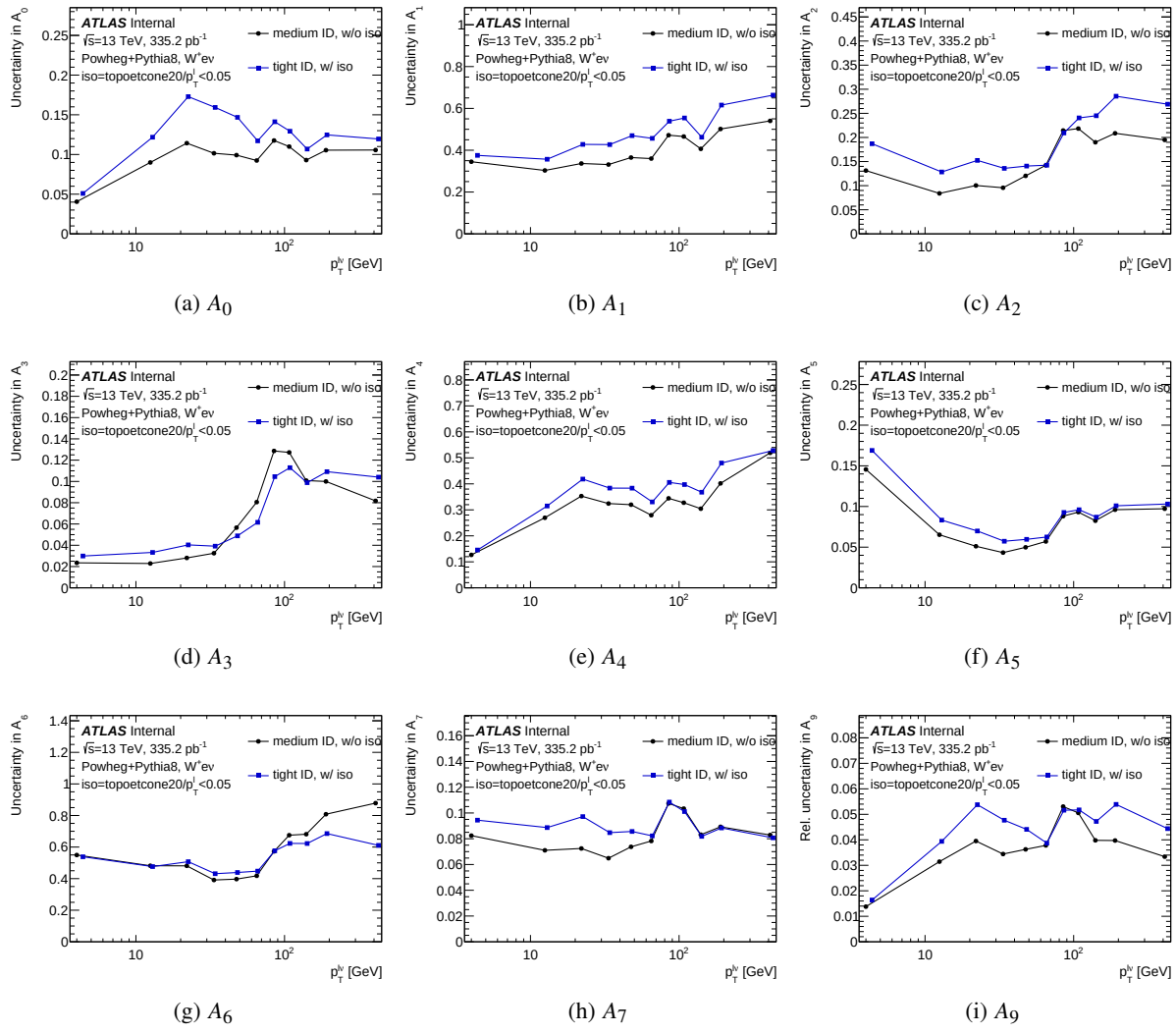


(i) A_9

Figure 75: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for different selection criteria in $y^{\ell\nu}$ binning for $W^+ \rightarrow e^+\nu$. The legend lists the various selection criteria in GeVfor p_{T}^{ℓ}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and m_T respectively; for example in black is when $p_{\mathrm{T}}^{\ell} > 25\mathrm{GeV}$, $E_{\mathrm{T}}^{\mathrm{miss}} > 25\mathrm{GeV}$, and m_T ¿ 50 GeV. This study was done using the nominal p_{z}^{ν} choice from the study in Section A.3.

A.2 Lepton selection optimisation

Having relaxed kinematic cuts allows more MJ background but can be reduced by having stricter identification and/or isolation requirements. The electron channel has more MJ so it needs both a tighter identification requirement while both electron and muon channels use the isolation requirement. Figures 76 - 79 show that these requirements do not remove too much statistical sensitivity with the inclusion of the identification and isolation cuts. There is net statistical significance gain with the relaxed kinematic cuts discussed in Section A.1 and these additional identification and isolation requirements.

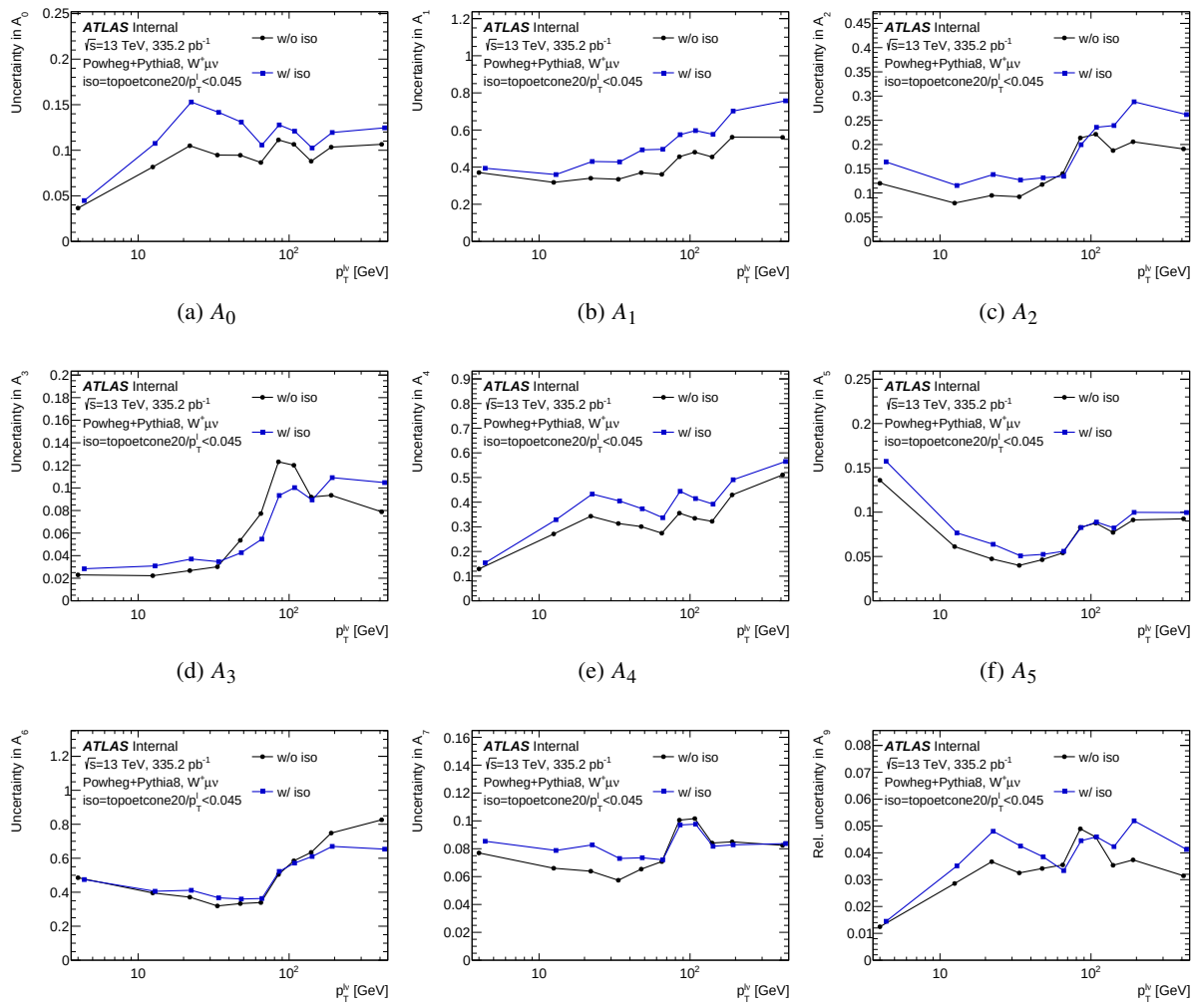


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 76: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the inclusion of adding isolation and a tight ID selection in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^+ \to e^+\nu$.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

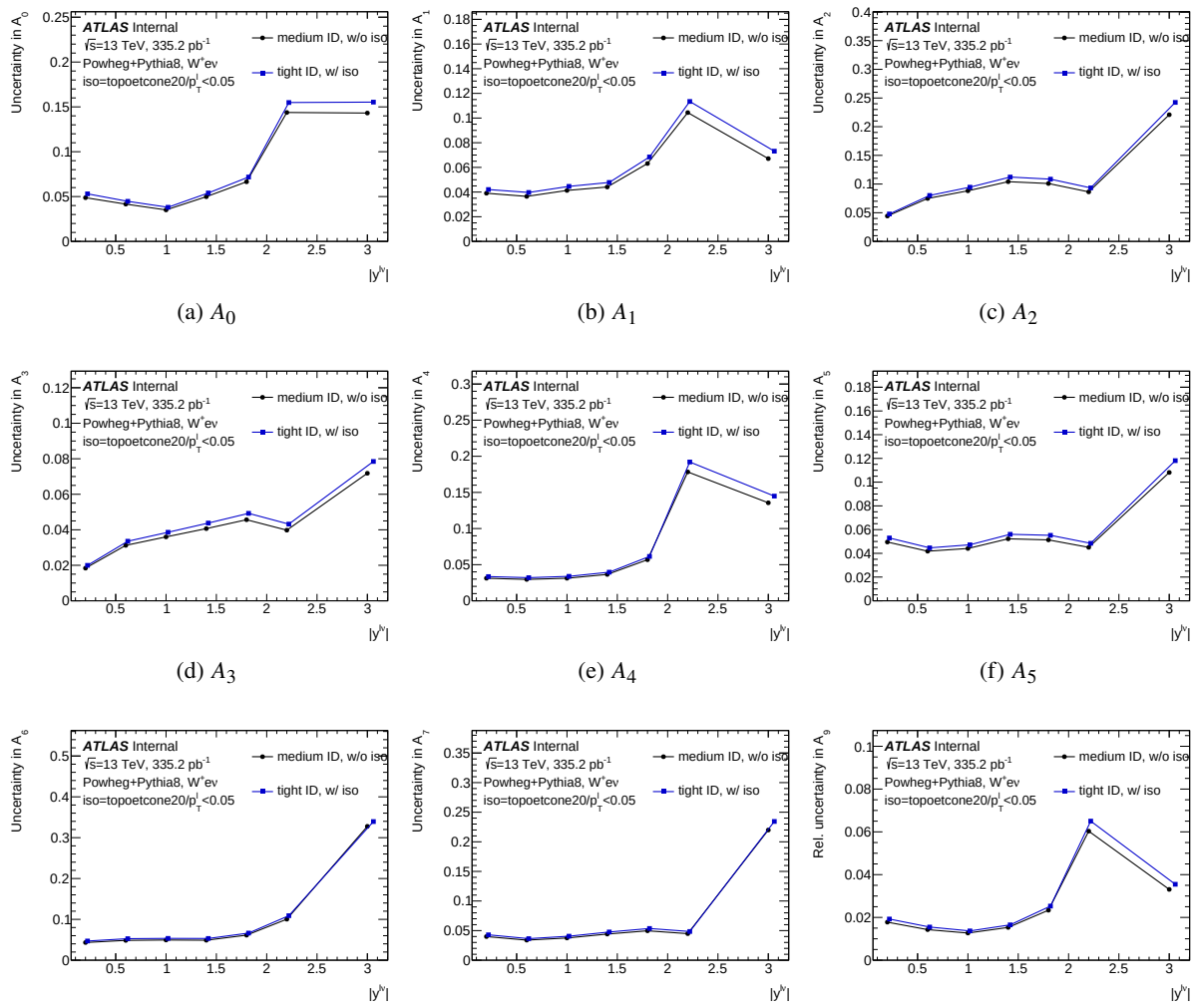
Figure 77: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the inclusion of adding isolation selection in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^+ \to \mu^+\nu$.

A.3 Neutrino p_z^ν no solution

The mass constraint results in an imaginary solution when $\Delta < 0$, which occurs when $m_T > m^W$. When this happens, events are not being able to be fully reconstructed, however, the ϕ_{CS} value can still be calculated as normal for these events. To use these events the templates were extended to have an additional 8×8 phase space in order to not loose ϕ_{CS} shape information and not interfere with the real solutions. However, what to use for the $\cos\theta_{\mathrm{CS}}$ value had to be decided, four ideas were tested:

Nominal

The nominal choice is to not assign any $\cos\theta_{\mathrm{CS}}$ information about these events by storing them within the same bin and just use the ϕ_{CS} information. This is done by setting $\cos\theta_{\mathrm{CS}} = 0$ and not solving for p_z^ν when



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

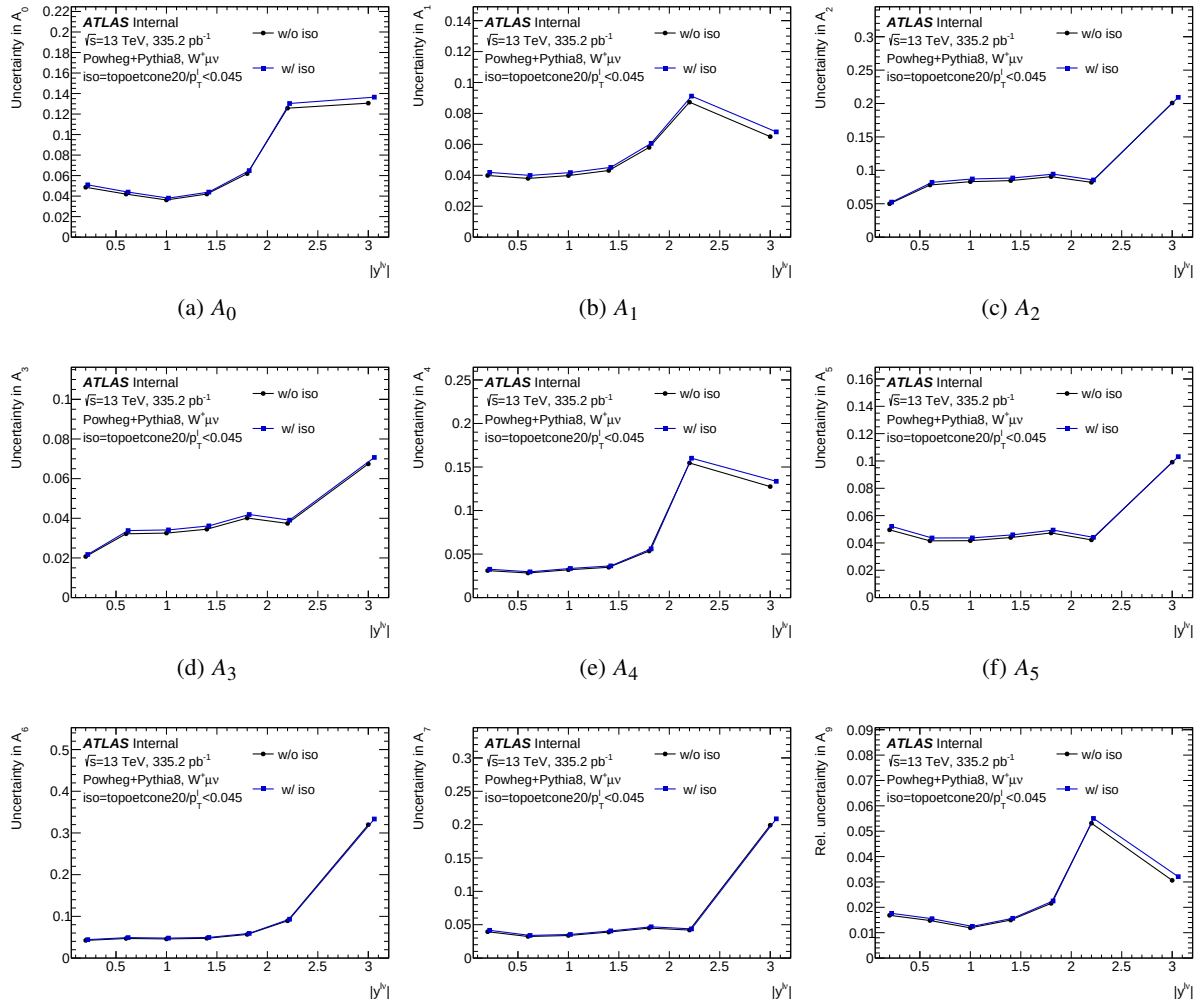
(g) A_6 (h) A_7 (i) A_9

Figure 78: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the inclusion of adding isolation and a tight ID selection in $y^{\ell\nu}$ binning for $W^+ \to e^+\nu$.

$\Delta < 0$. A drawback to this approach is that no attempt on the sign of $\cos\theta_{CS}$ is made so the W is always assumed to be traveling perpendicular to the negatively charged lepton (or neutrino in W^+ case) in the CS frame. The rest of the methods used were done in order to try and gain some of this $\cos\theta_{CS}$ information back.

VH Style

The second method tried, labeled VH style, took inspiration from a VH measurement just removes the imaginary part of the solution to $p_z^\nu = -b/2a$ [45]. This method allowed some solutions to have a negatively signed $\cos\theta_{CS}$ value but overall only increased the sensitivity to A_1 compared to the other methods.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 79: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the inclusion of adding isolation selection in $y^{\ell\nu}$ binning for $W^+ \to \mu^+\nu$.

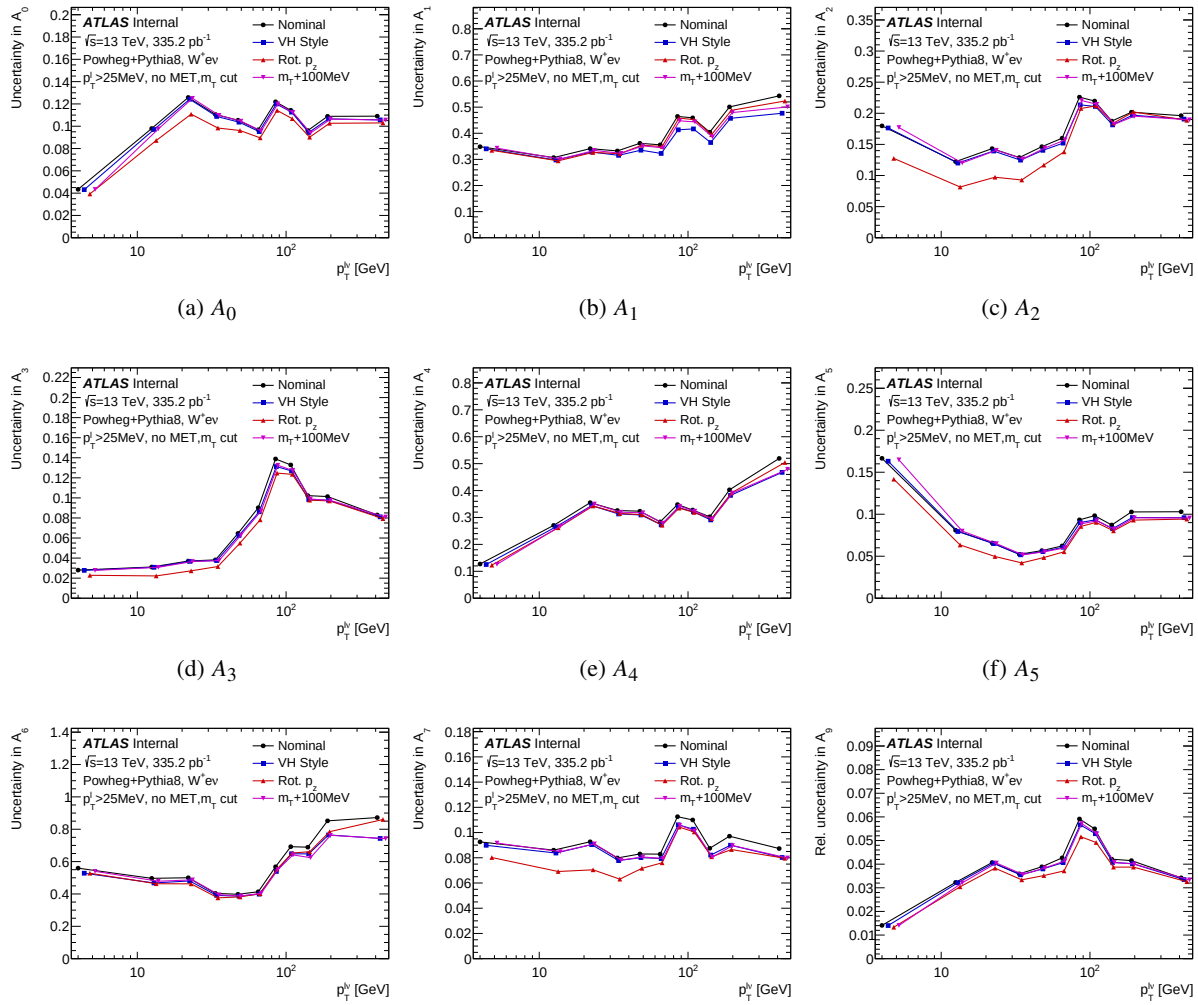
Rotation p_z^ν

The rotation p_z^ν method gets its name from rotating the $E_{\mathrm{T}}^{\mathrm{miss}}$ by the minimum amount needed for a solution. This rotation changes the values of the neutrino p_{T} and ϕ resulting in a different ϕ_{CS} value for these events. It was determined to give the best overall increase in sensitivity as there is a noticeable difference for low $p_{\mathrm{T}}^{\ell\nu}$ bins for A_0, A_2, A_3, A_5, and A_7 as well as a slight increase in sensitivity over the nominal.

Shifted m_T

The last method tested utilizes why we do not get a solution, which is when $m_T > m^W$. For the events that do not get an initial solution using the PDG m^W the solution is recalculated using $m_T + 100\mathrm{MeV}$ for the mass constraint. The 100MeV shift ensures that $\Delta \neq 0$ for all events like the VH style and this applied shift

is well within the resolution of m_T. One downside to this method is that even using this method there are still events with $\Delta < 0$ depending on the rest of the parameters while calculating it.

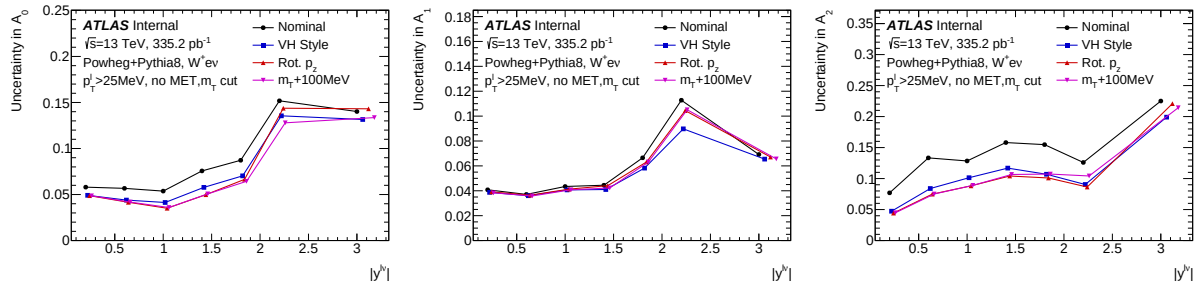


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

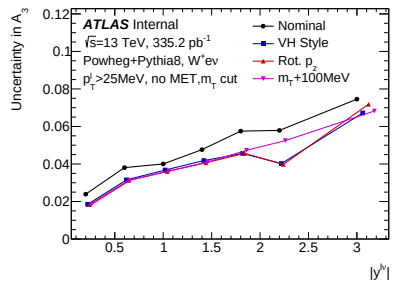
Figure 80: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the different neutrino solutions in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^+ \to e^+\nu$.



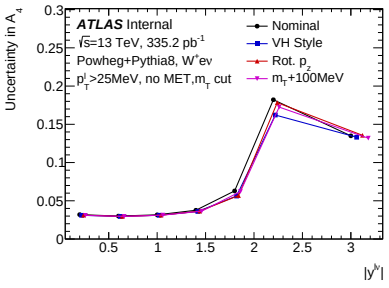
(a) A_0

(b) A_1

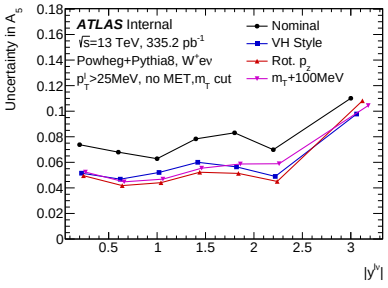
(c) A_2



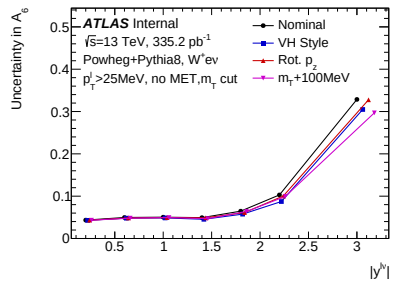
(d) A_3



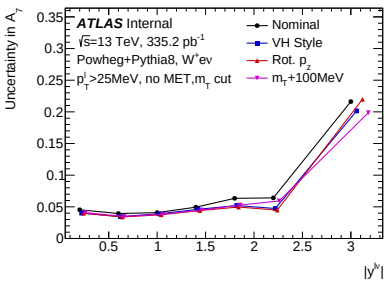
(e) A_4



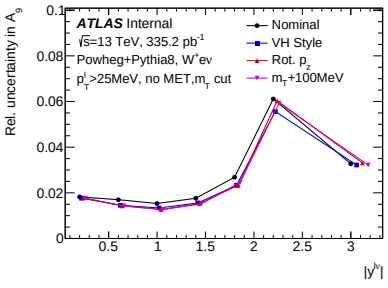
(f) A_5



(g) A_6



(h) A_7



(i) A_9

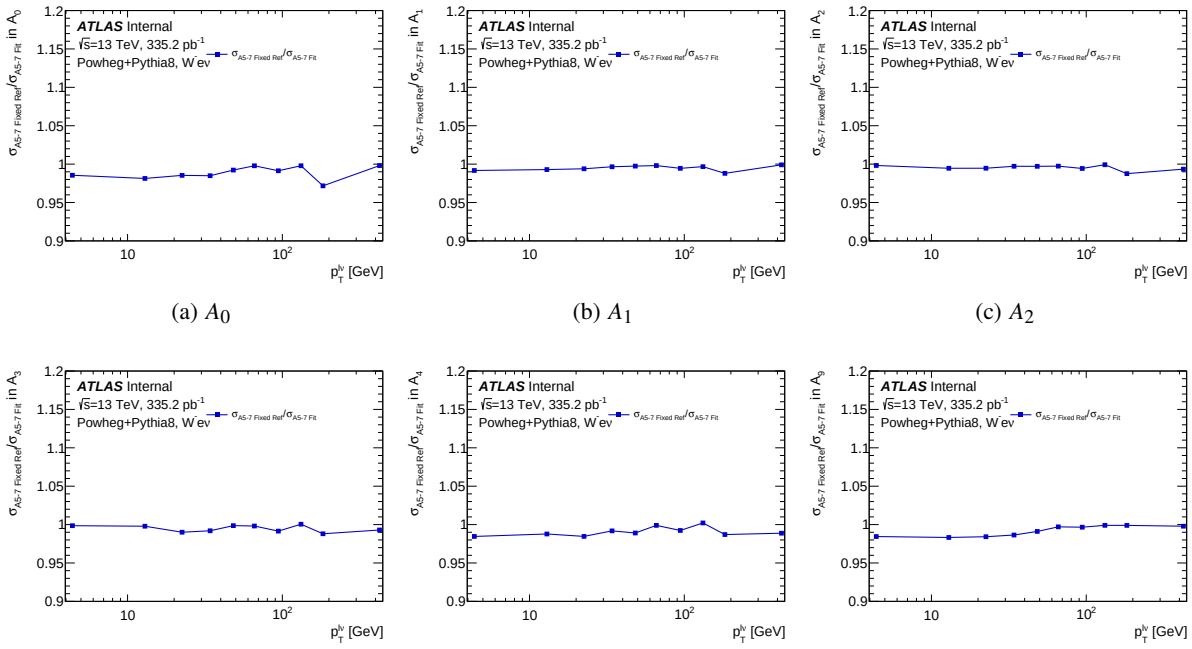
Figure 81: Comparison of the total statistical uncertainty for the coefficients $A_0 - A_7$ and relative statistical uncertainty for the total cross-section, A_9 for the different neutrino solutions in $y^{\ell\nu}$ binning.

A.4 Fixing $A_5 - A_7$

The coefficients $A_5 - A_7$ are expected to have small values since they don't start contributing until $O\left(\alpha_s^2\right)$. This means this analysis won't be sensitive to them but they need to be managed in the fit. We tested three different ways to see what the outcome on the other coefficients we are sensitive to:

- **Fit $A_5 - A_7$** - Extract the coefficients and allow them to float in the fit where no assumptions are made about them. This is considered to be the default option.
- **Fixed $A_5 - A_7$ reference** - Treat the coefficients as constants in the fit where they are set to and assumed as their reference values.
- **Fixed $A_5 - A_7$ zero** - Treat the coefficients as constants in the fit but set them to 0 and assume they are irrelevant.

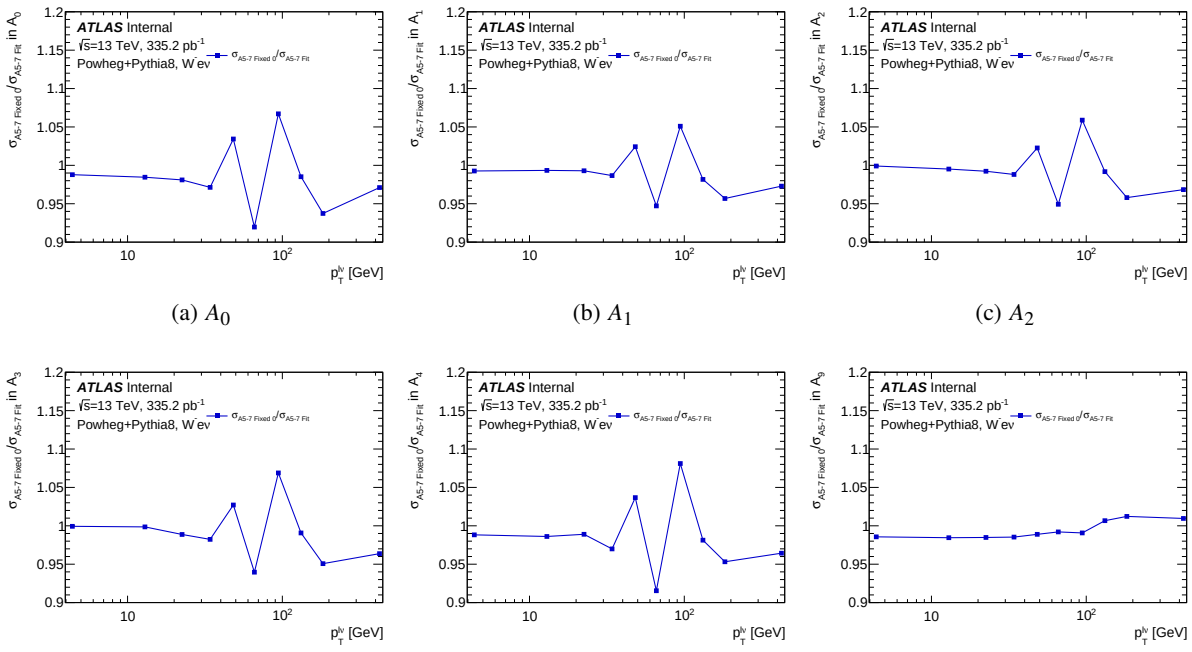
Fit $A_5 - A_7$ is the default and the other two methods were compared to it as seen in Figure 82 and 83 which take the ratio of the uncertainty on the coefficient divided by the uncertainty of the nominal. We can see the fixing $A_5 - A_7$ to their reference value reduces the uncertainty on the order of 1%. In contrast we see that fixing $A_5 - A_7$ to zero cause a reduction of uncertainty in some bins ¿5% but a gain in other bins, this seesaw shape is seen in all coefficients except the cross-section A_9. This region is known to be the least stable part of the fit because there is a transition of shape in the signal templates as well as this region is where the background makes a more meaningful contribution.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

Figure 82: Ratio of the uncertainty on each coefficient if $A_5 - A_7$ is fixed to their reference value as a constant in the fit over the uncertainty of letting $A_5 - A_7$ float in the fit. There is a consistent reduction in uncertainty on the order of 1% when fixing $A_5 - A_7$ to their reference values.

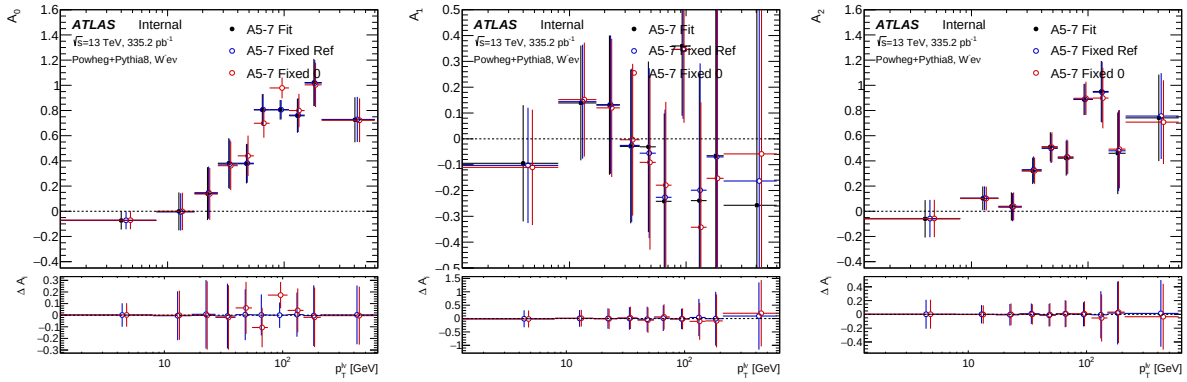


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

Figure 83: Ratio of the uncertainty on each coefficient if $A_5 - A_7$ is fixed to zero as a constant in the fit over the uncertainty of letting $A_5 - A_7$ float in the fit. There is a common seesaw shape which suggests an instability in the fit.

Comparing the values of the three methods as shown in Figure 84 we can see that fixing the coefficients to their reference values doesn't affect the value of the other coefficients in the fit while fixing them to zero does. From this study we concluded to stay with our default value. Fixing them to zero causes instability in the fit seen by the seesaw action of uncertainty and the change in values while fixing them at their reference values has very marginal gain but introduces an assumption where a systematic uncertainty might be needed to account for. The other thing to note is that χ^2/ndf only changed on the order of 10^{-3} so these other methods didn't suggest a better fit.



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_9

Figure 84: Comparison of the values for each treatment of $A_5 - A_7$. Fixing the coefficients to their reference values doesn't change the values for $A_0 - A_4$ and A_9 while fixing $A_5 - A_7$ to zero does cause a change to the values.

B Neutrino p_z^ν Sign Ambiguity

As discussed in Section 4.4, taking the mass constraint in order get p_z^ν results in a quadratic equation with a two-fold ambiguity. When $m_T < m^W$ there are two real solutions which results in a sign ambiguity in $\cos\theta_{CS}$. This appendix covers how choosing a solution at random we are able to statistically resolve the right distributions and have sensitivity to measure A_4 which depends on having the correct sign of $\cos\theta_{CS}$.

To understand this we look at Equation 5 which has two parts that contribute to the sign ambiguity. The first is $p_z^{\ell\nu}/|p_z^{\ell\nu}|$ which can be rewritten as the sign of $y^{\ell\nu}$ and the second is $\Delta P_\ell = \left(P_1^+ P_2^- - P_1^- P_2^+\right)$. The main source of solving the ambiguity in $\cos\theta_{CS}$, which happens 25% of the time, is when the two solutions have opposite signs in $y^{\ell\nu}$ which also causes a sign flip in ΔP_ℓ and both solutions of $\cos\theta_{CS}$ are equal. This happens predominately for central boson rapidity as seen in Figure 85 in red.

This sign flip can be understood by looking at the $W^- \to e^-\overline{\nu}$ case. In this case P_1 is fixed because we measure the lepton so ΔP_ℓ is simplified to

$$\Delta P_\ell = P_1^+ \left(E_\nu - p_z^\nu\right) - P_1^- \left(E_\nu + p_z^\nu\right). \tag{55}$$

If ΔP_ℓ flips sign then

$$\begin{aligned} E_\nu - p_z^\nu > E_\nu + p_z^\nu \Rightarrow p_z^\nu < 0 \quad &\text{for solution 1,} \\ E_\nu - p_z^\nu < E_\nu + p_z^\nu \Rightarrow p_z^\nu > 0 \quad &\text{for solution 2.} \end{aligned} \tag{56}$$

The sign flip of $y^{\ell\nu}$ can be understand by looking at the definition of y

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right), \tag{57}$$

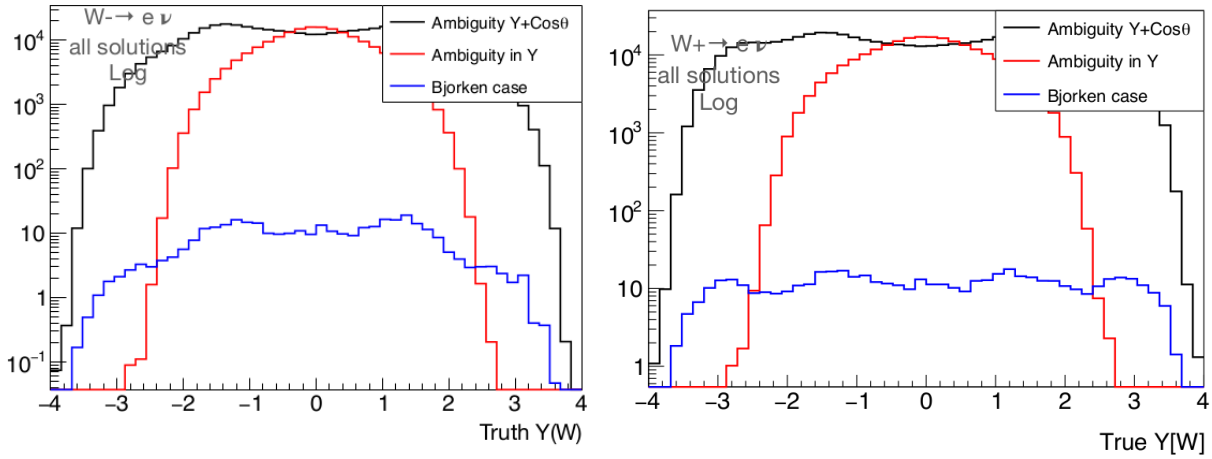
where if $y^{\ell\nu}$ flips sign then

$$\begin{aligned} E^{\ell\nu} + p_z^{\ell\nu} c > E^{\ell\nu} - p_z^{\ell\nu} c \text{ for solution 1,} \\ E^{\ell\nu} + p_z^{\ell\nu} c < E^{\ell\nu} - p_z^{\ell\nu} c \text{ for solution 2.} \end{aligned} \tag{58}$$

The sign flip happens mostly for central $y^{\ell\nu}$ due to fact the p_z^ν sign flip is more likely to occur when $p_z^\ell + p_z^\nu$ is close to zero. This can be seen by using $p_z^{\ell\nu} = p_z^\ell + p_z^\nu$ in Equation 58 which can be rearranged into

$$\begin{aligned} p_z^\ell + p_z^\nu > 0 \text{ for solution 1,} \\ p_z^\ell + p_z^\nu < 0 \text{ for solution 2.} \end{aligned} \tag{59}$$

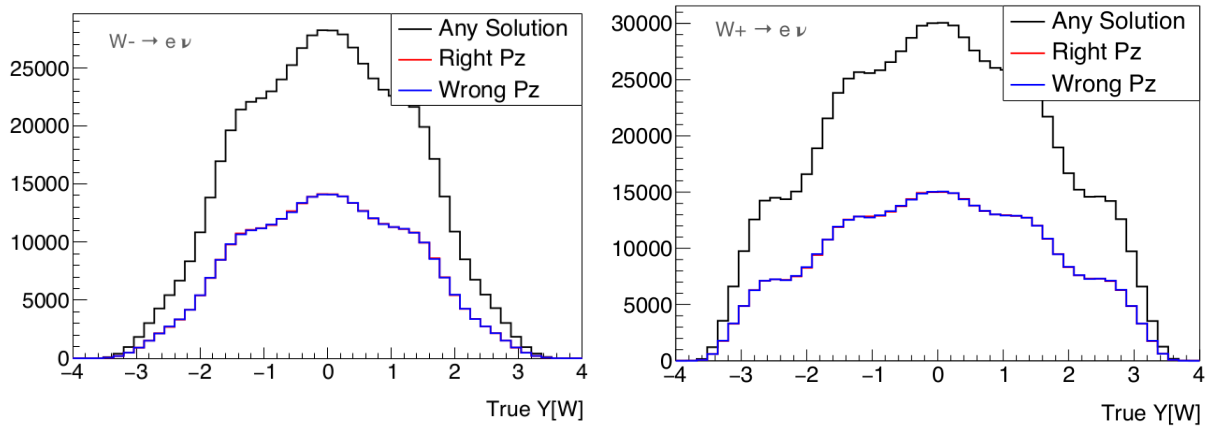
The other time that we can resolve this sign ambiguity is when the one of the solutions violates the Bjorken condition of $x < 1$ and is unphysical. This occurs when either $|p_z^\nu| > \sqrt{s}/2$ or $|p_z^\nu + p_z^\ell| > \sqrt{s}/2$ which is when the solution causes the longitudinal momenta of the neutrino or the W boson to exceed the beam energy. These events aren't frequent as seen in blue in Figure 85.



(a) $W^- \to e^-\overline{\nu}$ (b) $W^+ \to e^+\nu$

Figure 85: The distribution of events based on the ability to resolve the sign ambiguity. In black are events where there is both a sign ambiguity in $\cos\theta_{CS}$ and in $y^{\ell\nu}$, in red are events where the ambiguity in $\cos\theta_{CS}$ is solved by sign flipping cancelling but there is still an ambiguity in the sign of $y^{\ell\nu}$, and in blue are events where one solution violates Bjorken $x < 1$ and only one solution is valid.

To further understand the sensitivity it is important to look at what happens when we chose both the right and wrong p_z^ν solution, where we define the right p_z^ν as when the solution is closest to the truth p_z^ν. Both right and wrong solutions happen an equal amount as shown in Figure 86 where the wrong solution in blue is covering the right solution in red.

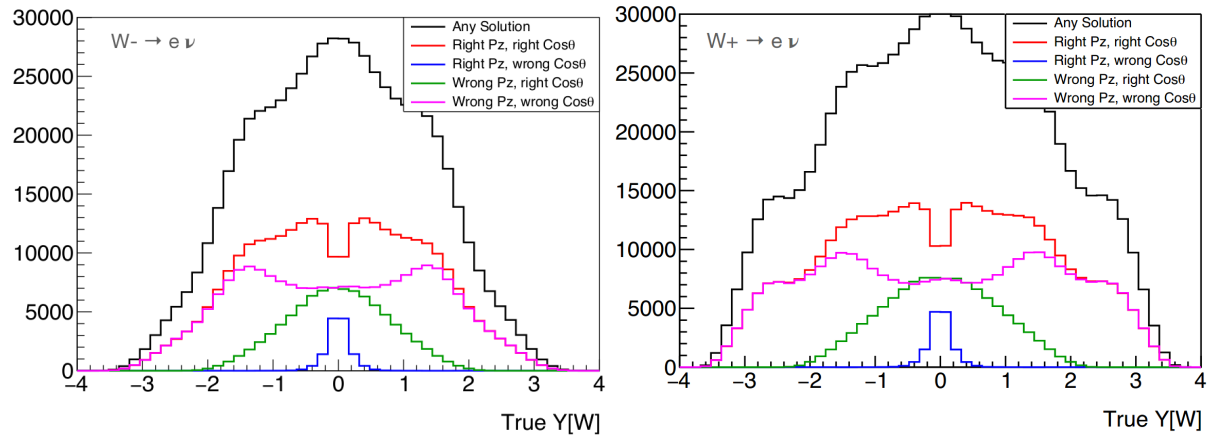


(a) $W^- \to e^-\overline{\nu}$ (b) $W^+ \to e^+\nu$

Figure 86: The distribution of the right (red) and wrong (blue) choice of p_z^ν which are equally distributed in $y^{W,\text{Truth}}$.

There are four possible combinations of choosing the correct solution of p_z^ν and getting the correct sign of $\cos\theta_{CS}$. The first is when we chose the correct p_z^ν and get the correct $\cos\theta_{CS}$ as shown in Figure 87 in red. These events are important for having sensitivity at large rapidity while the more central region we would get the correct sign of $\cos\theta_{CS}$ regardless. Second are events where we chose the correct p_z^ν but due the limited resolution the sign of $\cos\theta_{CS}$ flipped as shown in Figure 87 in blue. The events in where we do

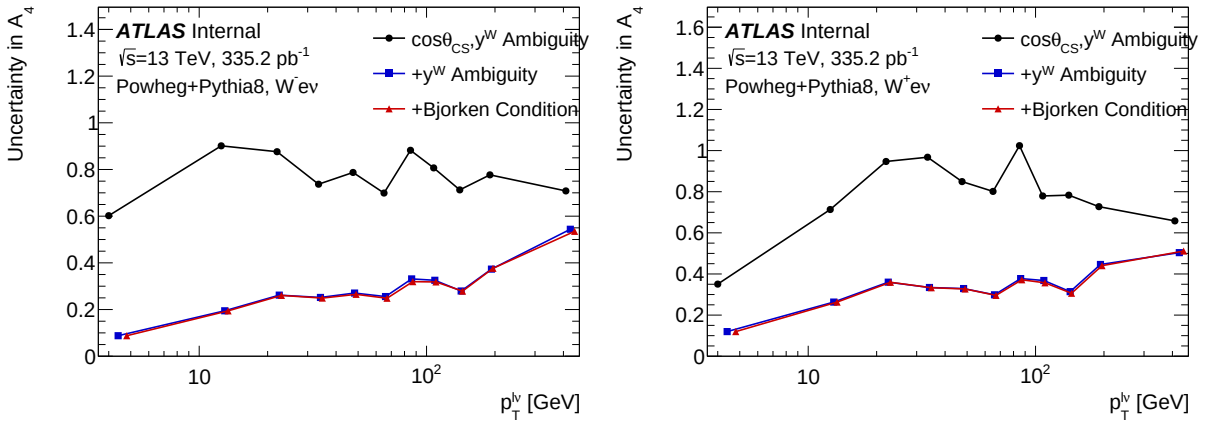
better than the random choice and gain the most sensitivity to A_4 in when we chose the wrong p_z^ν but get the right sign of $\cos\theta_{CS}$ because the sign of $y^{\ell\nu}$ also flipped. These events are shown in green in Figure 87 and are seen to be more central. Lastly, are events where we have the wrong p_z^ν and wrong sign of $\cos\theta_{CS}$ which dilutes our sensitivity mostly at high $y^{\ell\nu}$ as shown in pink Figure 87.



(a) $W^- \to e^-\overline{\nu}$ (b) $W^+ \to e^+\nu$

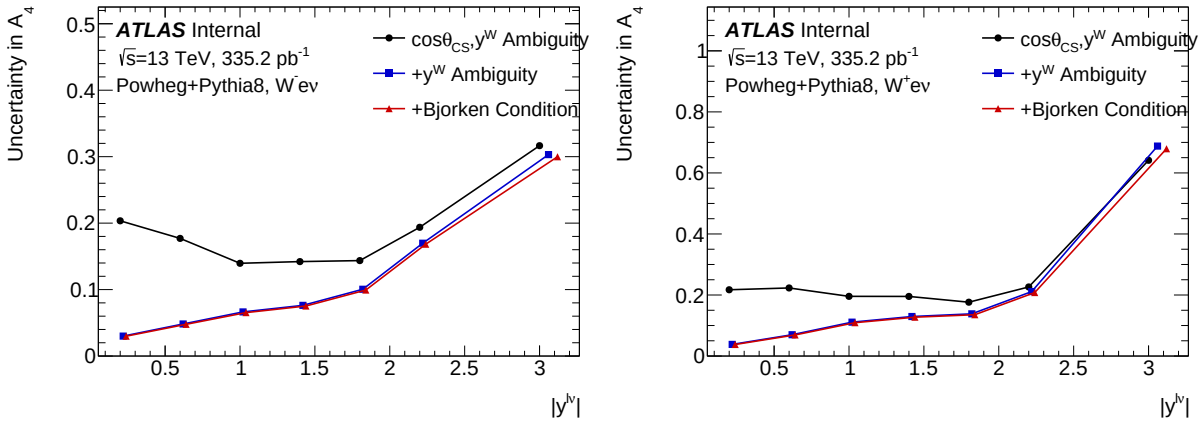
Figure 87: Distribution of events based for the four combinations of the choice of p_z^ν solution and the correct sign of $\cos\theta_{CS}$.

To show the sensitivity gained, A_4 was extracted dependent on the ambiguity level based on the three categories in Figure 85. Using only events that have both an ambiguity in the sign of $\cos\theta_{CS}$ and $y^{\ell\nu}$, where there is only a 50% chance to get the correct sign, we can see there isn't any sensitivity to A_4 as shown in Figures 88 & 89 in black. By combining these events with the events that only have a sign ambiguity in $y^{\ell\nu}$, where we get the right sign every time, we gain a large sensitivity as shown in blue in Figures 88 & 89. Lastly, for events that violate Bjorken $x < 1$ we gain a very small sensitivity increase as seen in red in Figures 88 & 89. In Figure 89 we also see that most of our sensitivity gain is in central $y^{\ell\nu}$ as expected.



(a) $W^- \to e^-\overline{\nu}$ (b) $W^+ \to e^+\nu$

Figure 88: Comparison of the total statistical uncertainty for A_4 for the inclusion of events dependent on the ambiguity in the sign of $\cos\theta_{CS}$ and y^W in $p_T^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$.



(a) $W^- \to e^-\overline{\nu}$ (b) $W^+ \to e^+\nu$

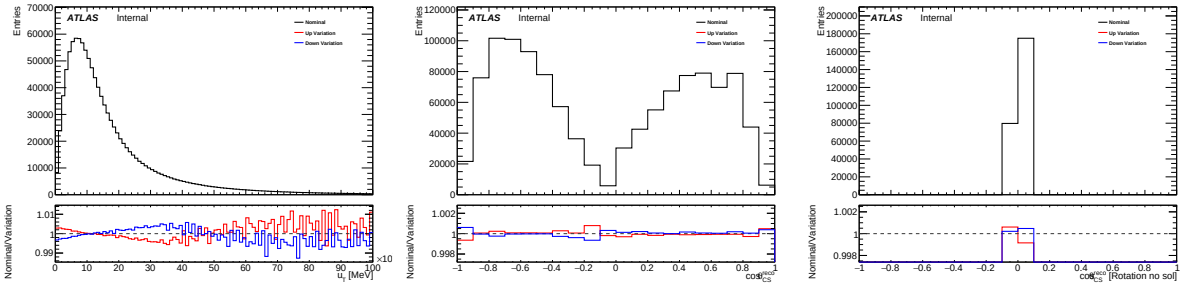
Figure 89: Comparison of the total statistical uncertainty for A_4 for the inclusion of events dependent on the ambiguity in the sign of $\cos\theta_{\mathrm{CS}}$ and y^W in $y^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$.

C Hadronic Recoil Systematics

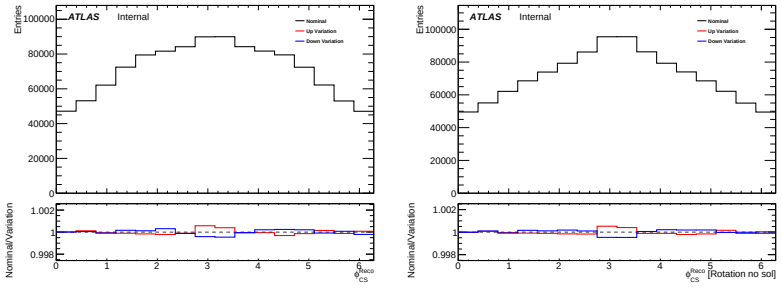
The hadronic recoil is expected to be the most dominant experimental systematic uncertainty for this measurement. There are 62 nuisance parameters associated to the statistical uncertainty on the resolution correction, statistical uncertainty of the response correction, systematic uncertainty on the response correction, systematic uncertainty on the ΣE_{T}, systematic uncertainty on the resolution correction due to the extrapolation from Z to W, and systematic uncertainty on the response correction due to the extrapolation from Z to W. More information on these corrections can be found in the support note [46].

C.1 Kinematic Distribution Impact

To understand the impact of the NPs variations on the hadronic recoil the angular distributions for $W^- \rightarrow e^- \overline{\nu}$ were plotted, Figures 90-93 show the variation for the four largest variations. The largest variation in the angular distributions is in Figure 91 where $\cos\theta_{\mathrm{CS}}$ varies at a 0.8% level.



(a) u_T (b) $\cos\theta_{\mathrm{CS}}$, solution (c) $\cos\theta_{\mathrm{CS}}$, no solution

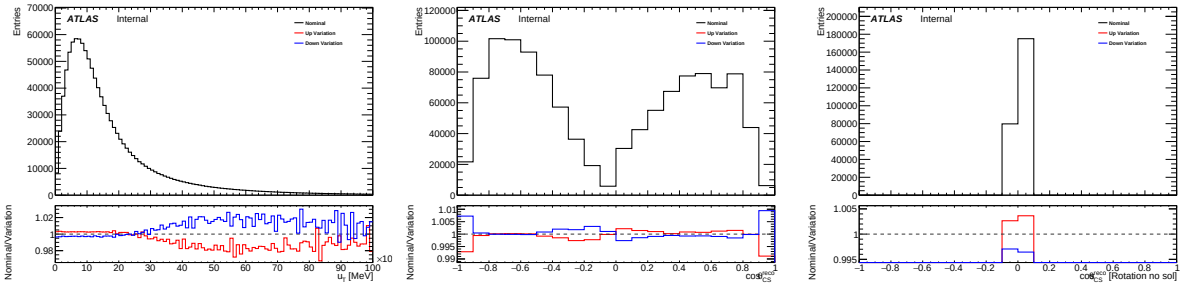


(d) ϕ_{CS}, solution (e) ϕ_{CS}, no solution

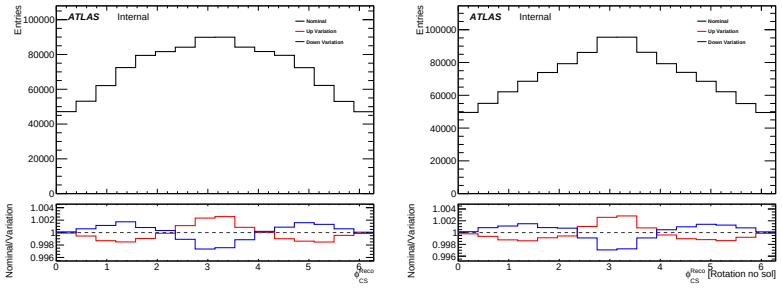
Figure 90: Systematic variation of resolution extrapolation systematic on the hadronic recoil and the impact on the angular distributions.

C.2 Template Variation

The variations in Section C.1 show the impact of some of the nuisance parameters on the integrated angular distribution level but this section will cover the impact the systematic variations have on the templates for $W^- \rightarrow e^- \overline{\nu}$. Examples of the impact on three NPs is shown for three $p_{\mathrm{T}}^{\ell\nu}$ bins in Figures 94 - 96 and



(a) u_T (b) $\cos\theta_{CS}$, solution (c) $\cos\theta_{CS}$, no solution



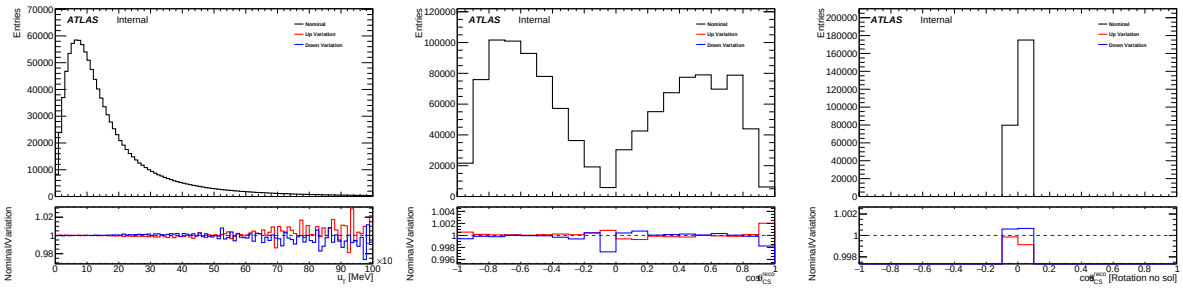
(d) ϕ_{CS}, solution (e) ϕ_{CS}, no solution

Figure 91: Systematic variation of response extrapolation systematic on the hadronic recoil and the impact on the angular distributions.

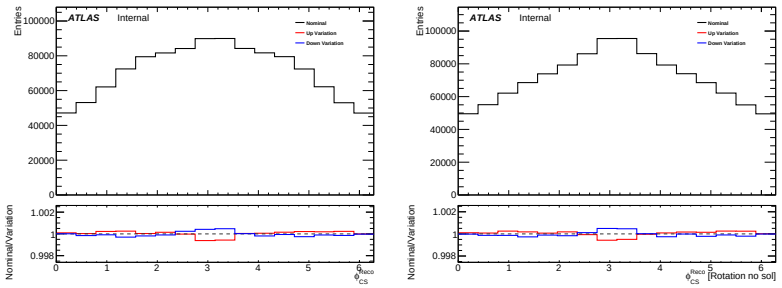
three $y^{\ell\nu}$ bins in Figures 97 - 99. They are plotted to show the percent difference between nominal and the variation.

C.3 Coefficient Impact

All NPs are propagated through the A_i framework in order to estimate their expected uncertainty on the fit. They are combined into three categories, Resolution, Response, and ΣE_{T} (labeled as SET) as shown in Figures 100 and 101. For the ΣE_{T} NP, the effects cancel for the coefficients $A_0 - A_7$, while the for the differential cross-section, A_9, they do not as the systematic is modeling the non-closure of the ΣE_{T} modeling procedure discussed in [46]. The total impact of the hadronic recoil systematics is not expected to be large as shown in Figures 102 and 103 where it is labeled as Sys and does not contribute to the total uncertainty on the same level as the statistical uncertainty.

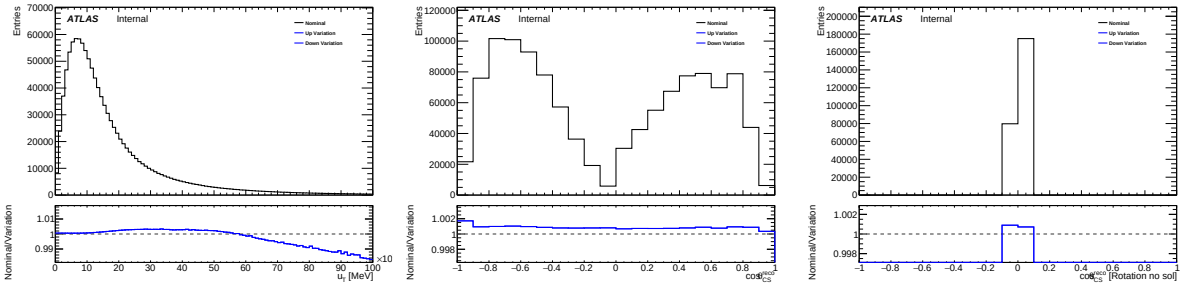


(a) u_T (b) $\cos\theta_{\mathrm{CS}}$, solution (c) $\cos\theta_{\mathrm{CS}}$, no solution

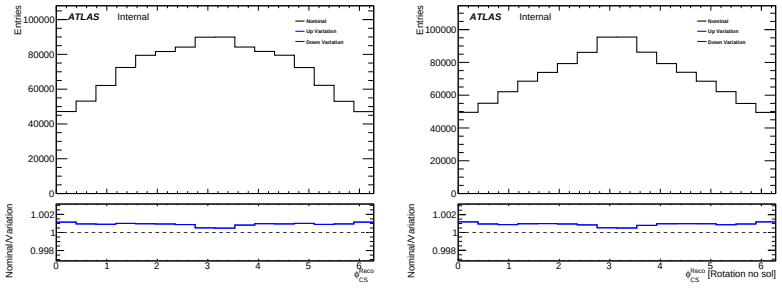


(d) ϕ_{CS}, solution (e) ϕ_{CS} no solution

Figure 92: Systematic variation of response systematic on the hadronic recoil and the impact on the angular distributions.

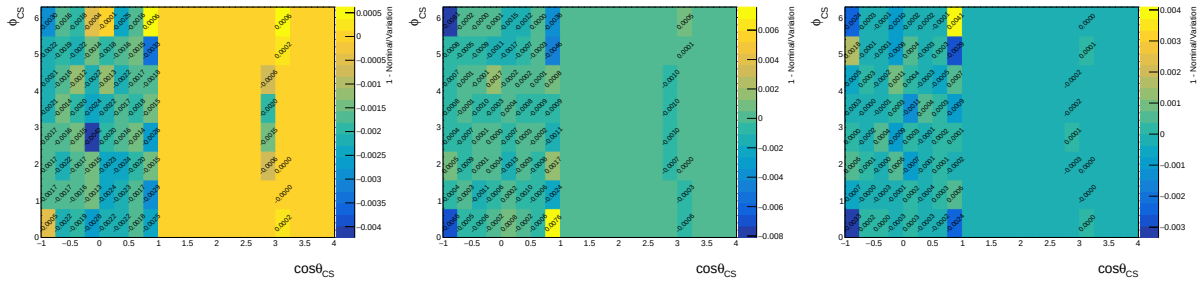


(a) u_T (b) $\cos\theta_{\mathrm{CS}}$, solution (c) $\cos\theta_{\mathrm{CS}}$, no solution

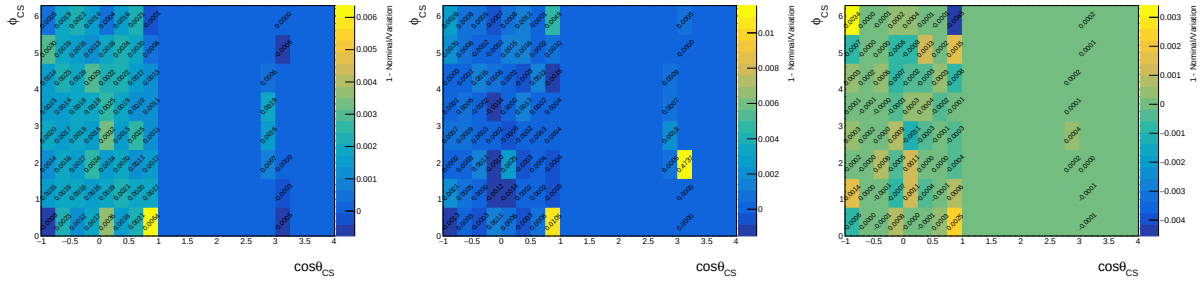


(d) ϕ_{CS}, solution (e) ϕ_{CS}, no solution

Figure 93: Systematic variation of ΣE_{T} systematic on the hadronic recoil and the impact on the angular distributions.

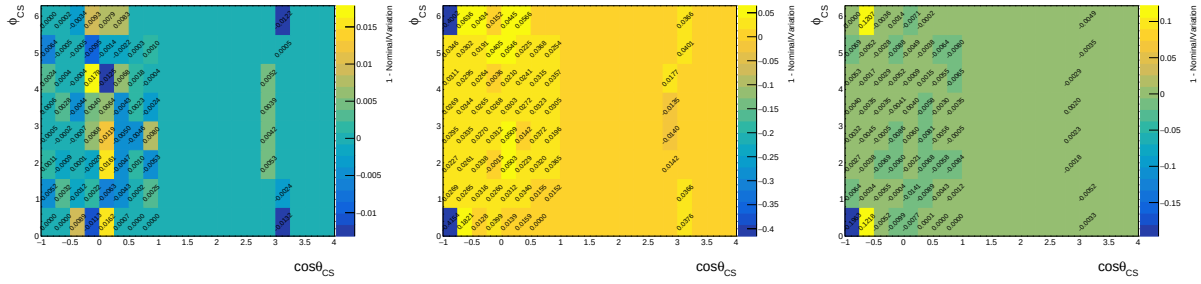


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

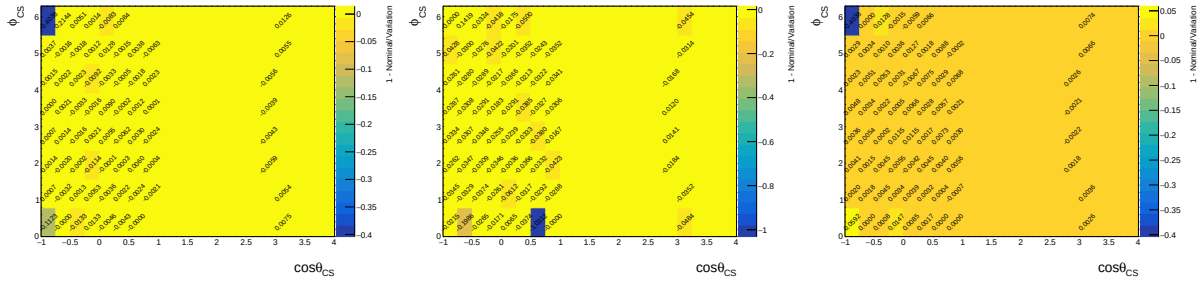


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

Figure 94: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $0 < p_{\mathrm{T}}^{\ell\nu} < 8$GeV.

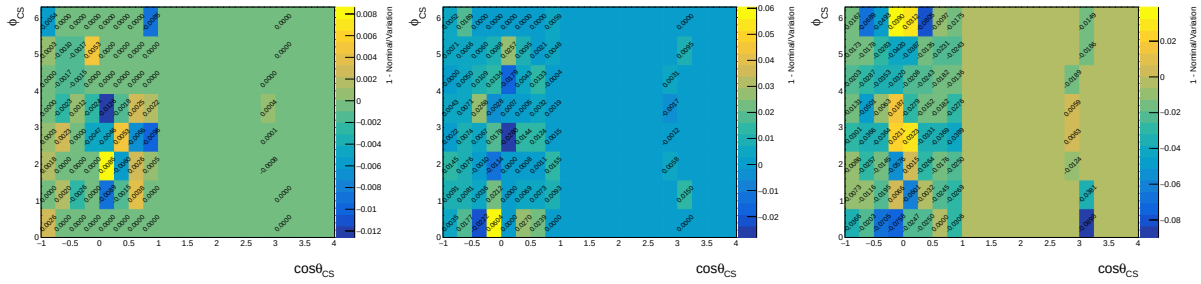


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

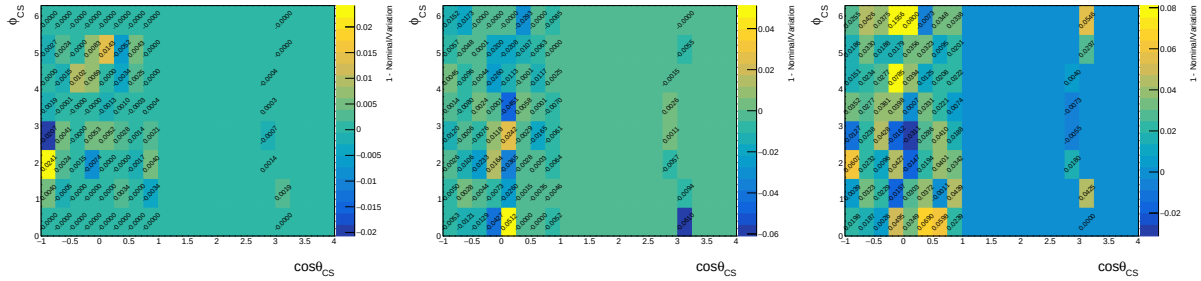


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

Figure 95: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $50 < p_{\mathrm{T}}^{\ell\nu} < 55$GeV.

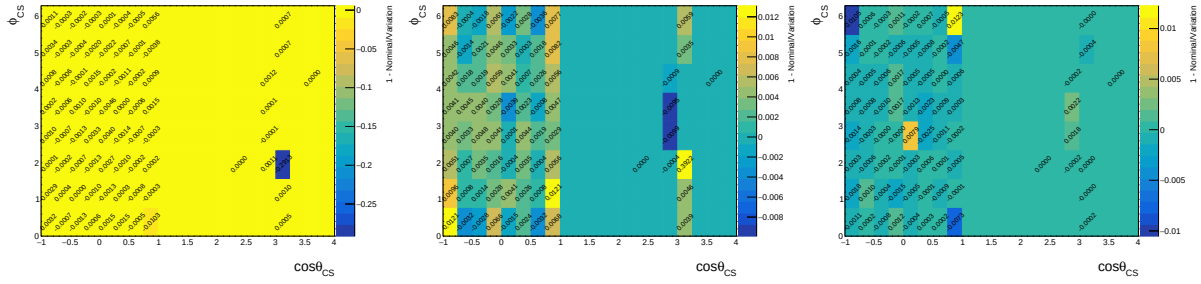


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

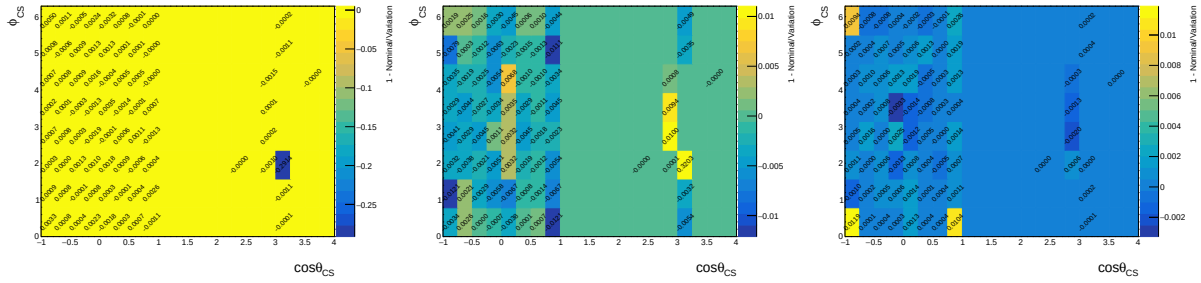


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

Figure 96: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $120 < p_{\mathrm{T}}^{\ell\nu} < 160$GeV.

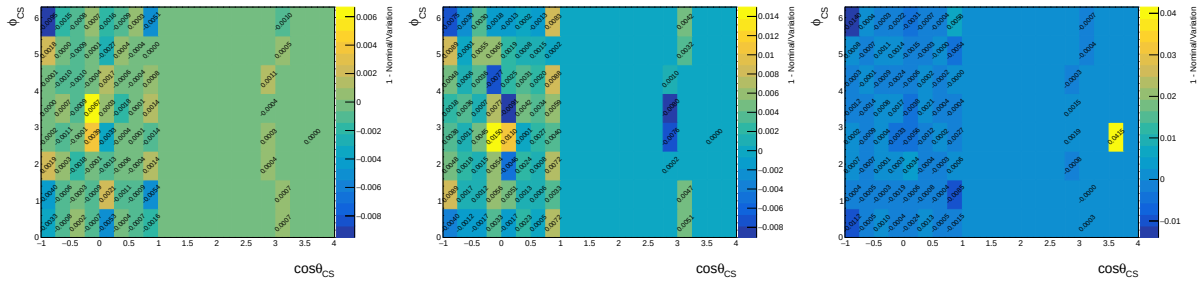


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

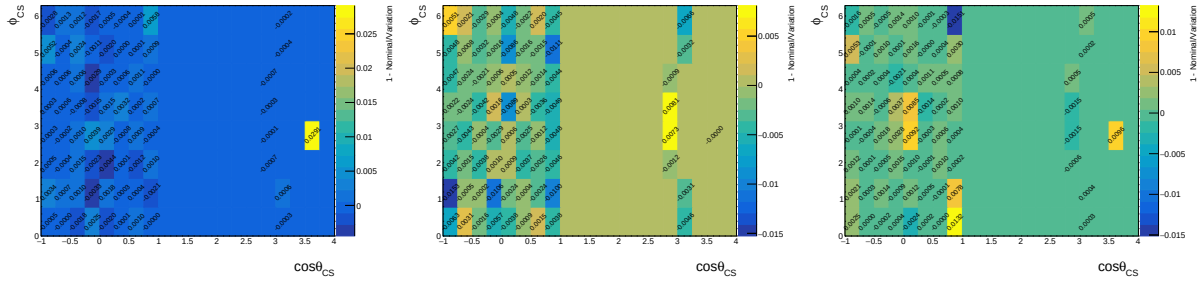


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

Figure 97: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $0 < y^{\ell\nu} < 0.4$.

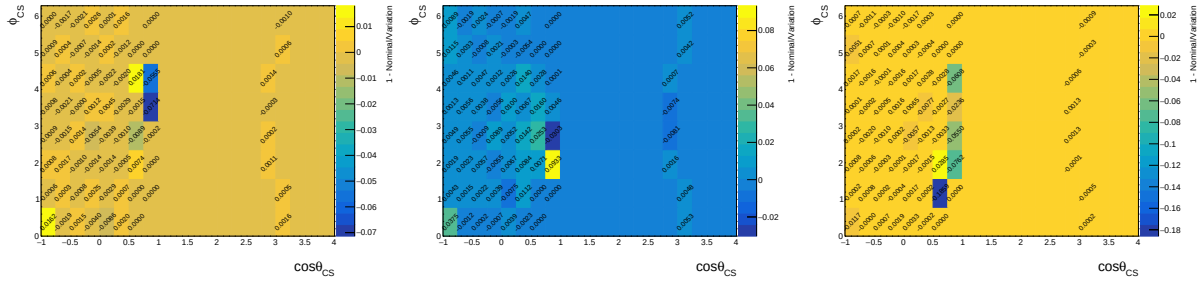


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

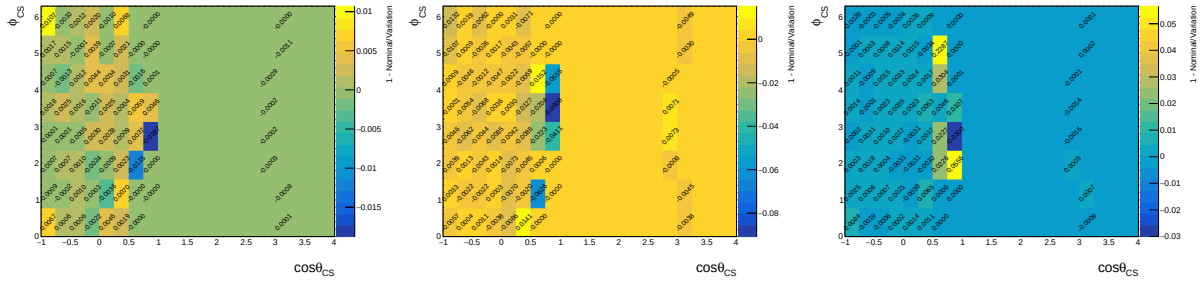


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

Figure 98: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $1.2 < y^{\ell\nu} < 1.6$.

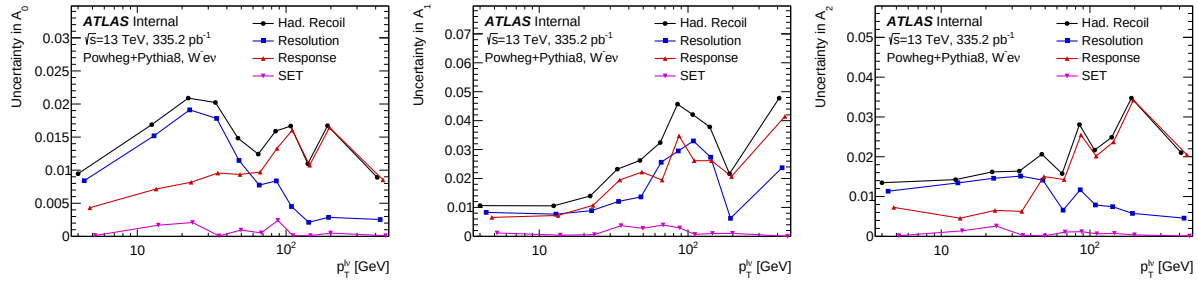


(a) Resolution Ext. Sys Up (b) Response Ext. Sys Up (c) Response Sys Up

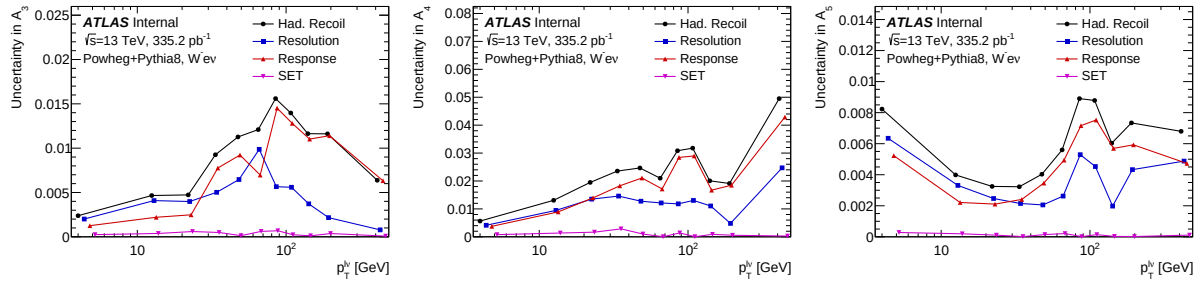


(d) Resolution Ext. Sys Down (e) Response Ext. Sys Down (f) Response Sys Down

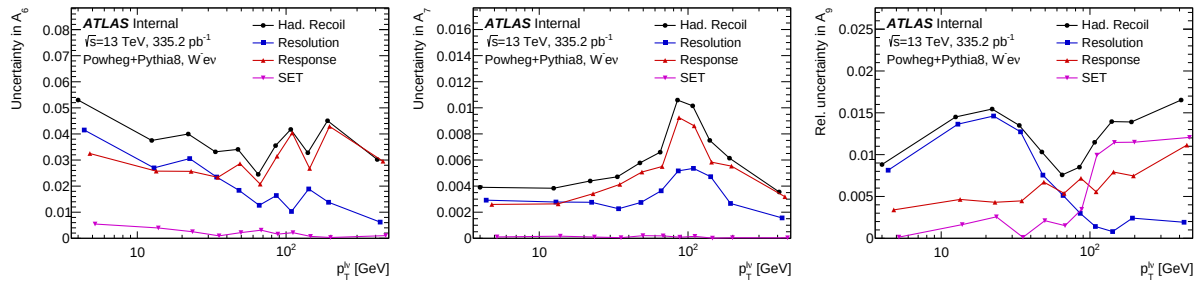
Figure 99: Up and down variation in the templates used to extract the coefficients for $W^- \to e^-\overline{\nu}$ for three different hadronic recoil NPs for the range $2.0 < y^{\ell\nu} < 2.4$.



(a) A_0 (b) A_1 (c) A_2

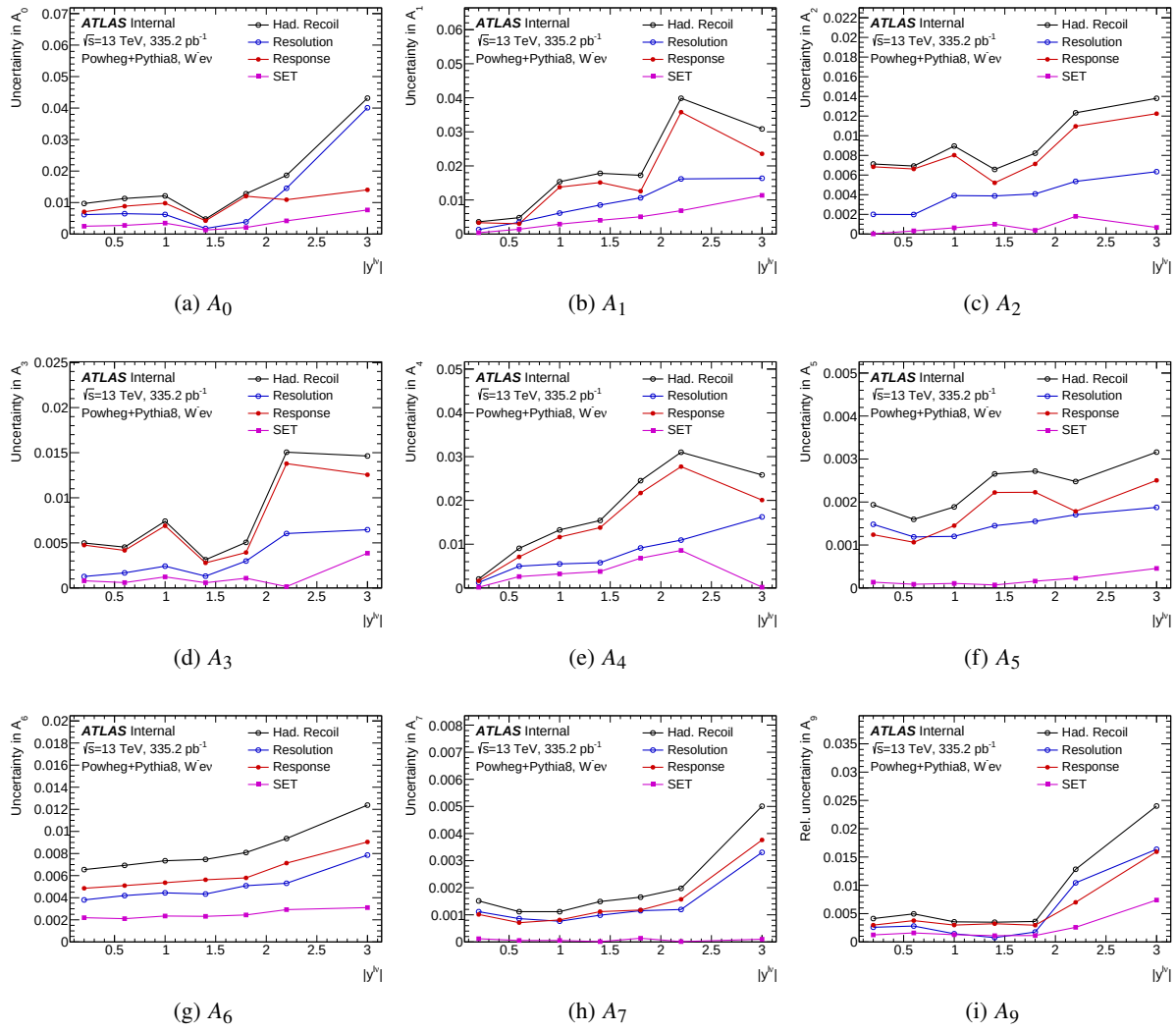


(d) A_3 (e) A_4 (f) A_5



(g) A_6 (h) A_7 (i) A_9

Figure 100: Hadronic recoil systematic uncertainty breakdown for the coefficients $A_0 - A_7$ in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$.

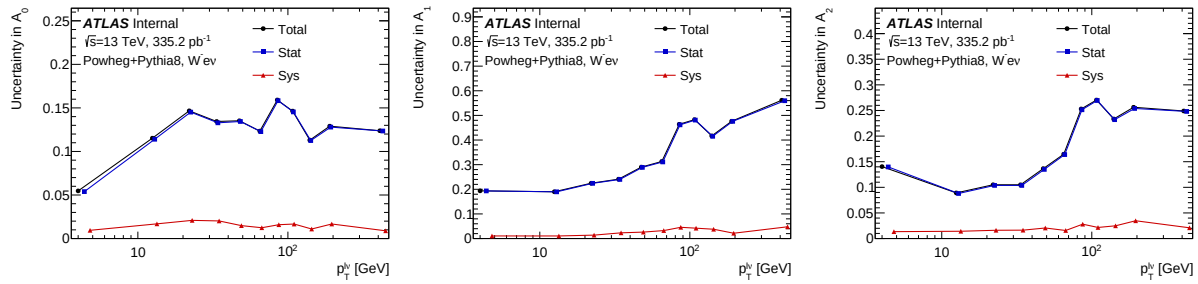


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

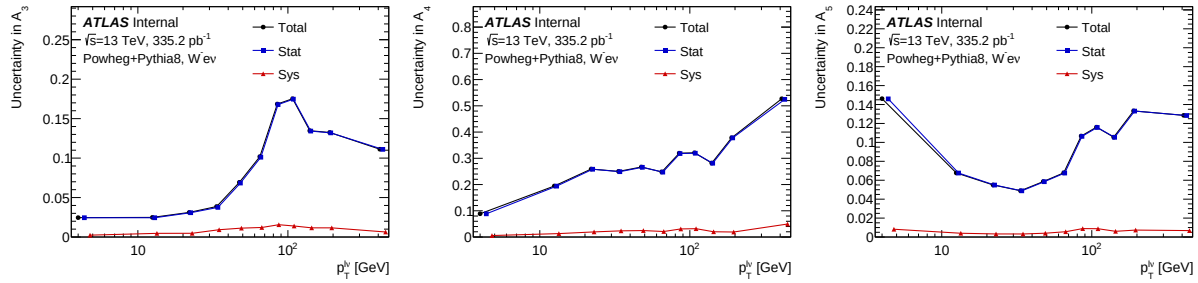
Figure 101: Hadronic recoil systematic uncertainty breakdown for the coefficients $A_0 - A_7$ in $y^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$.



(a) A_0

(b) A_1

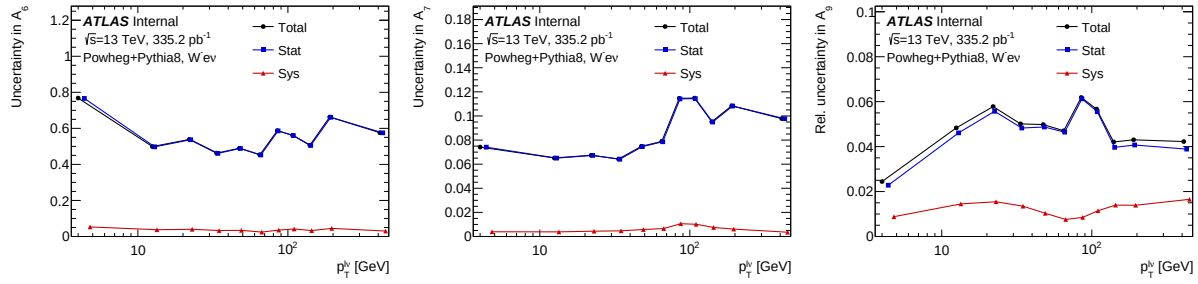
(c) A_2



(d) A_3

(e) A_4

(f) A_5

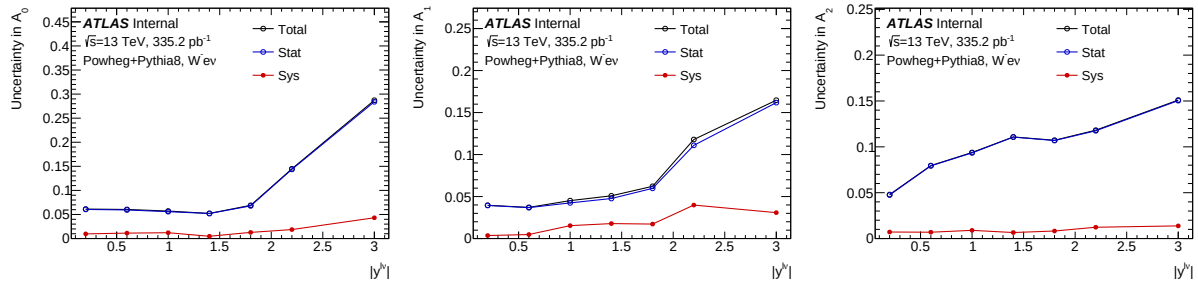


(g) A_6

(h) A_7

(i) A_9

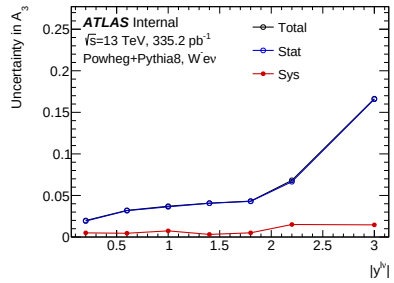
Figure 102: Total uncertainty breakdown for the two components statistical and hadronic recoil systematic for the coefficients $A_0 - A_7$ in $p_{\mathrm{T}}^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$.



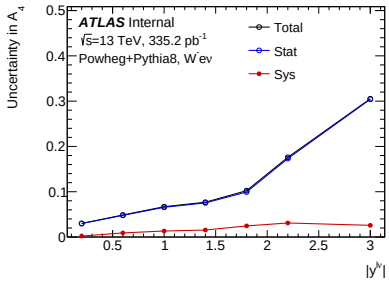
(a) A_0

(b) A_1

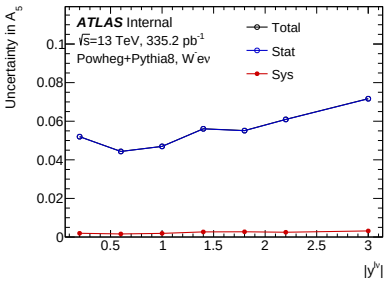
(c) A_2



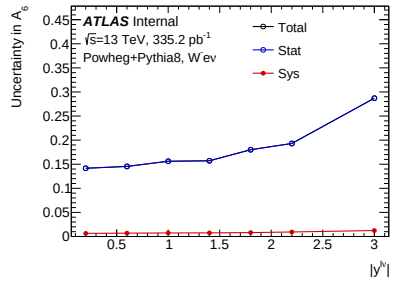
(d) A_3



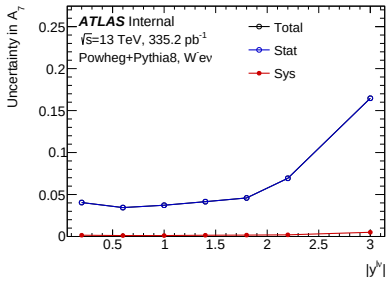
(e) A_4



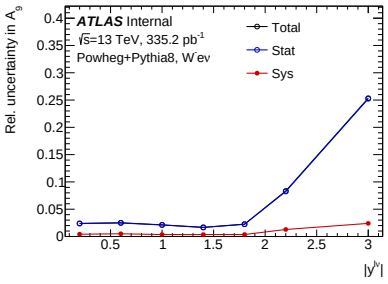
(f) A_5



(g) A_6



(h) A_7



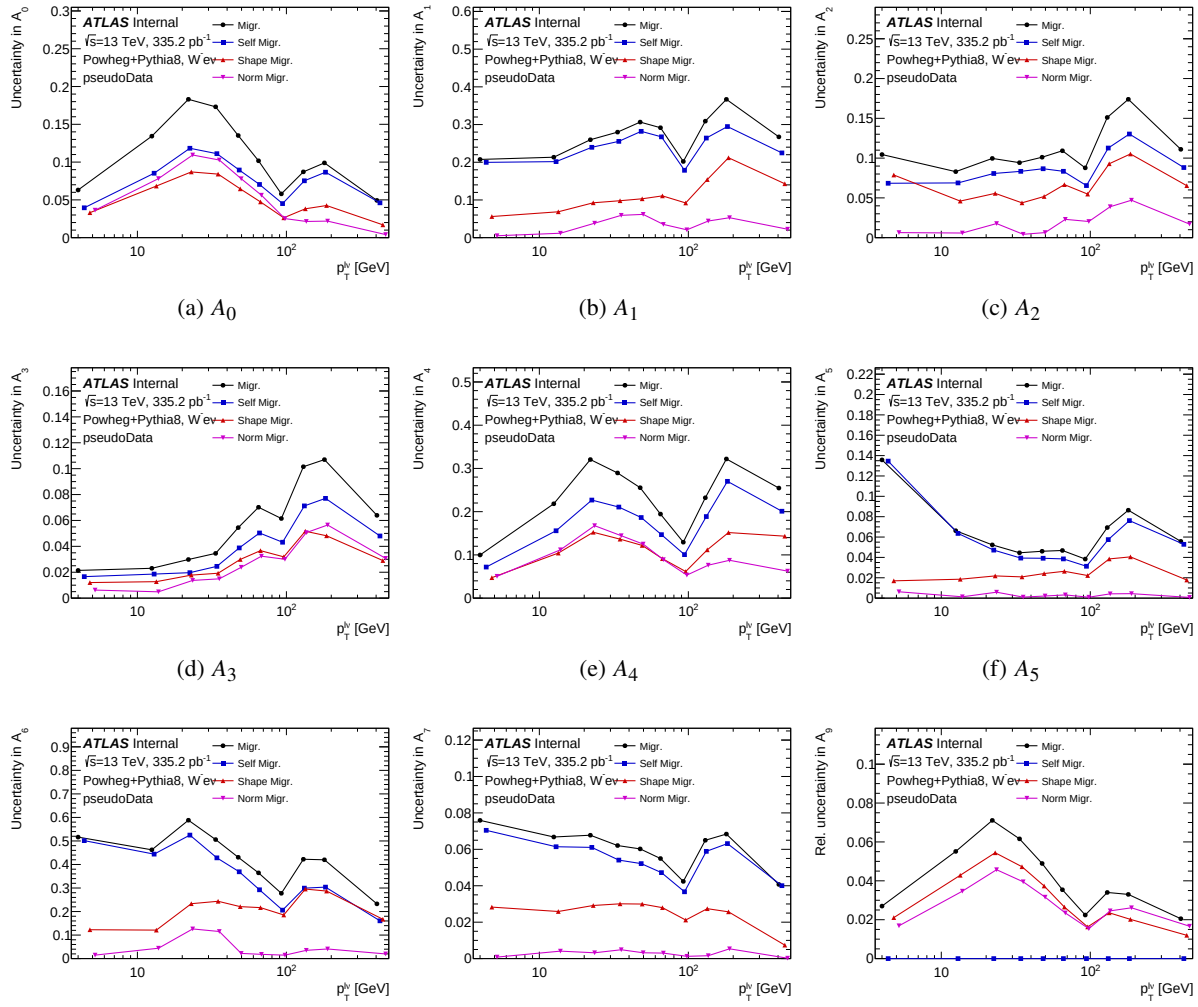
(i) A_9

Figure 103: Total uncertainty breakdown for the two components statistical and hadronic recoil systematic for the coefficients $A_0 - A_7$ in $y^{\ell\nu}$ binning for $W^- \to e^-\overline{\nu}$.

D Migration Uncertainties

This appendix covers the breakdown of each migration uncertainty category for each coefficient as described in Section 7.1. There are three categories of migration uncertainty: self which is the migration between $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ or $y^{W,\mathrm{Truth}}$ bins for the same coefficient, shape which is the migration between coefficients in the same $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bin, and norm which is the self migration for the normalization. Note that the self and norm migrations affect each coefficient but the differential cross-section, A_9, only is affected by the norm migration. The coefficients are ratios and are affected by their own measurement as well as the normalization, while A_9 is just a measurement of the normalization.

D.1 p_{T}^{W} Differential

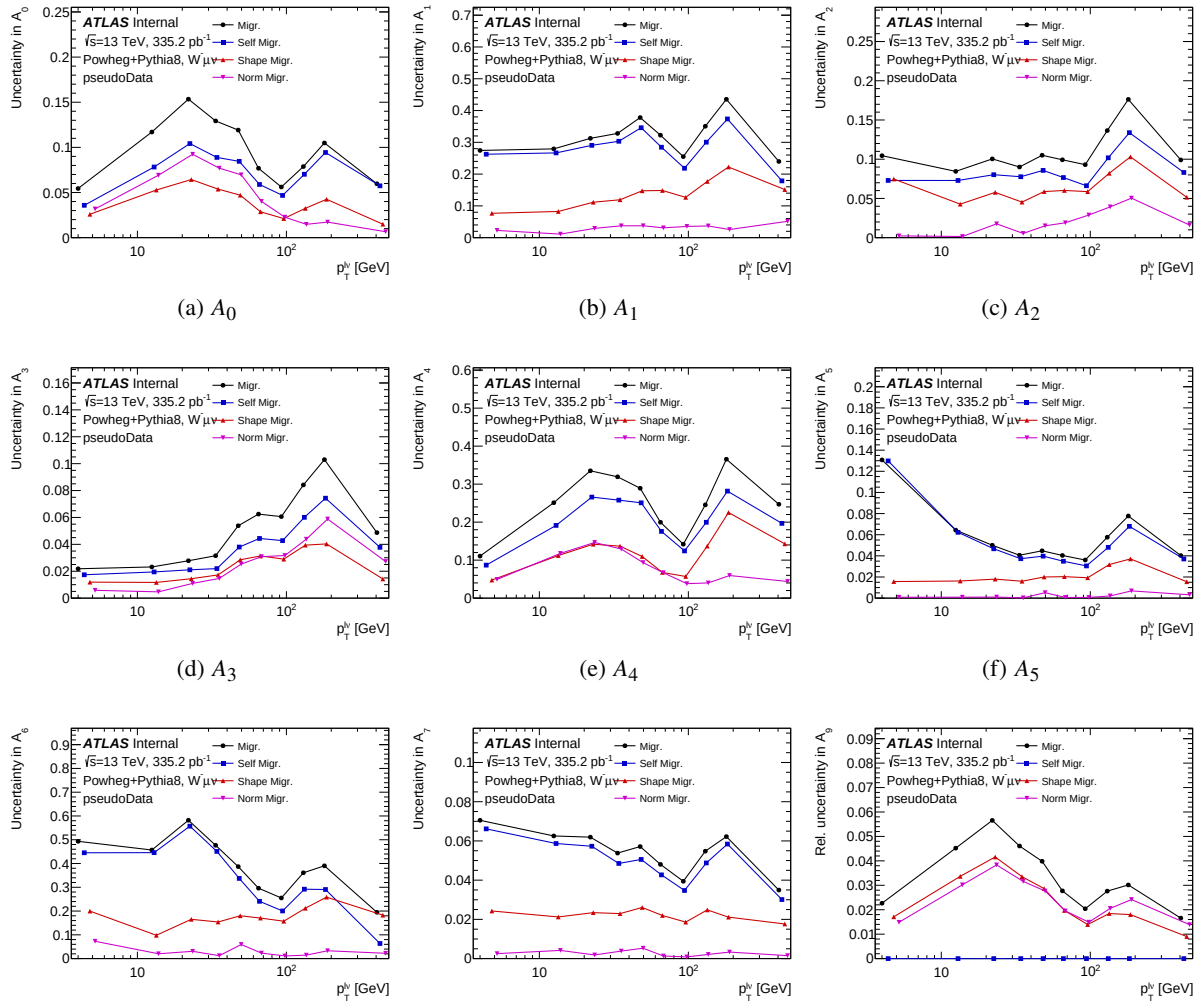


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 104: The uncertainty comparison for migration related uncertainties in the $W^- \to e^- \overline{\nu}$ channel for measurement in p_{T}^{W}. Blue is the self migration between $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

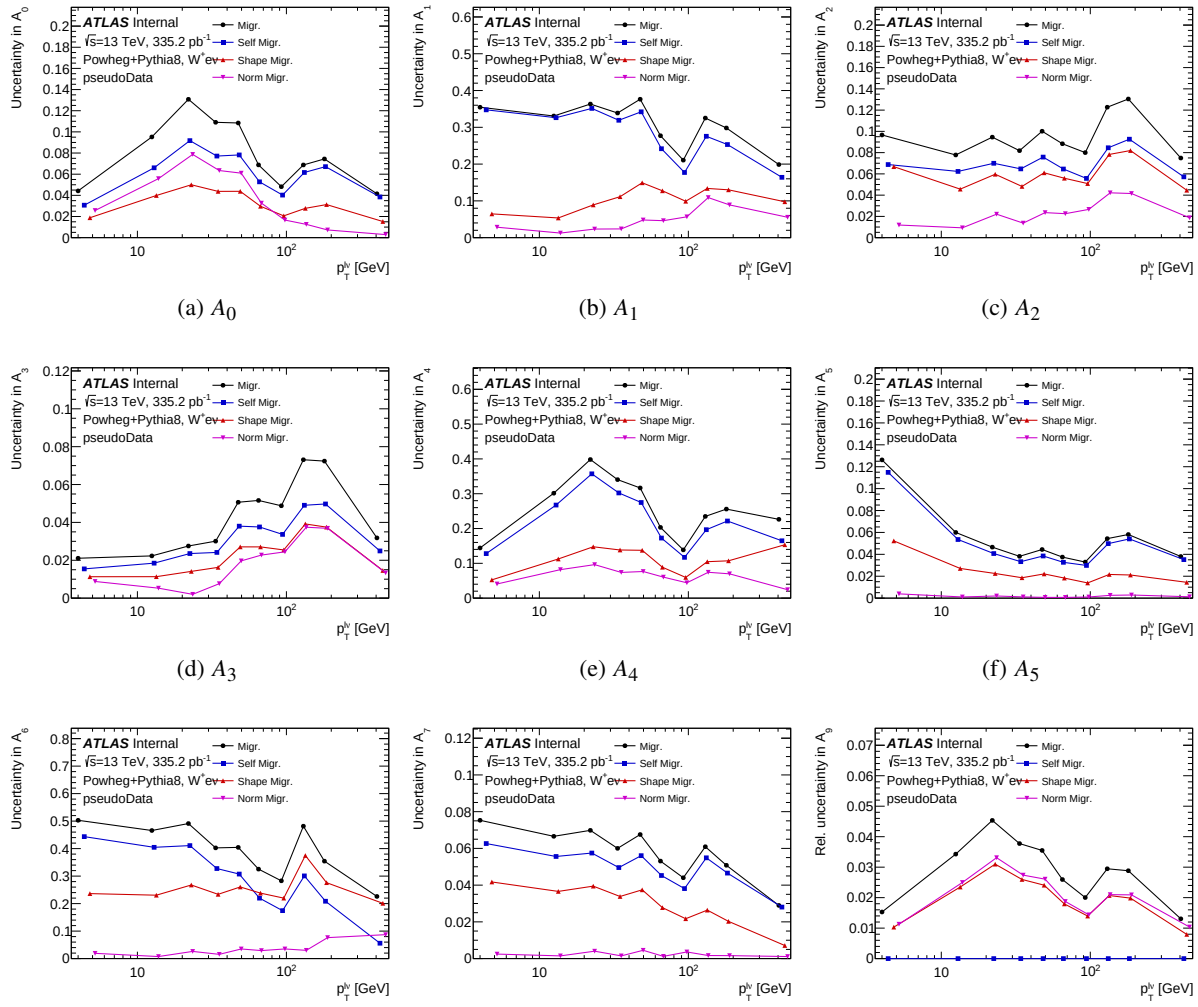


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 105: The uncertainty comparison for migration related uncertainties in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in p_{T}^{W}. Blue is the self migration between $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

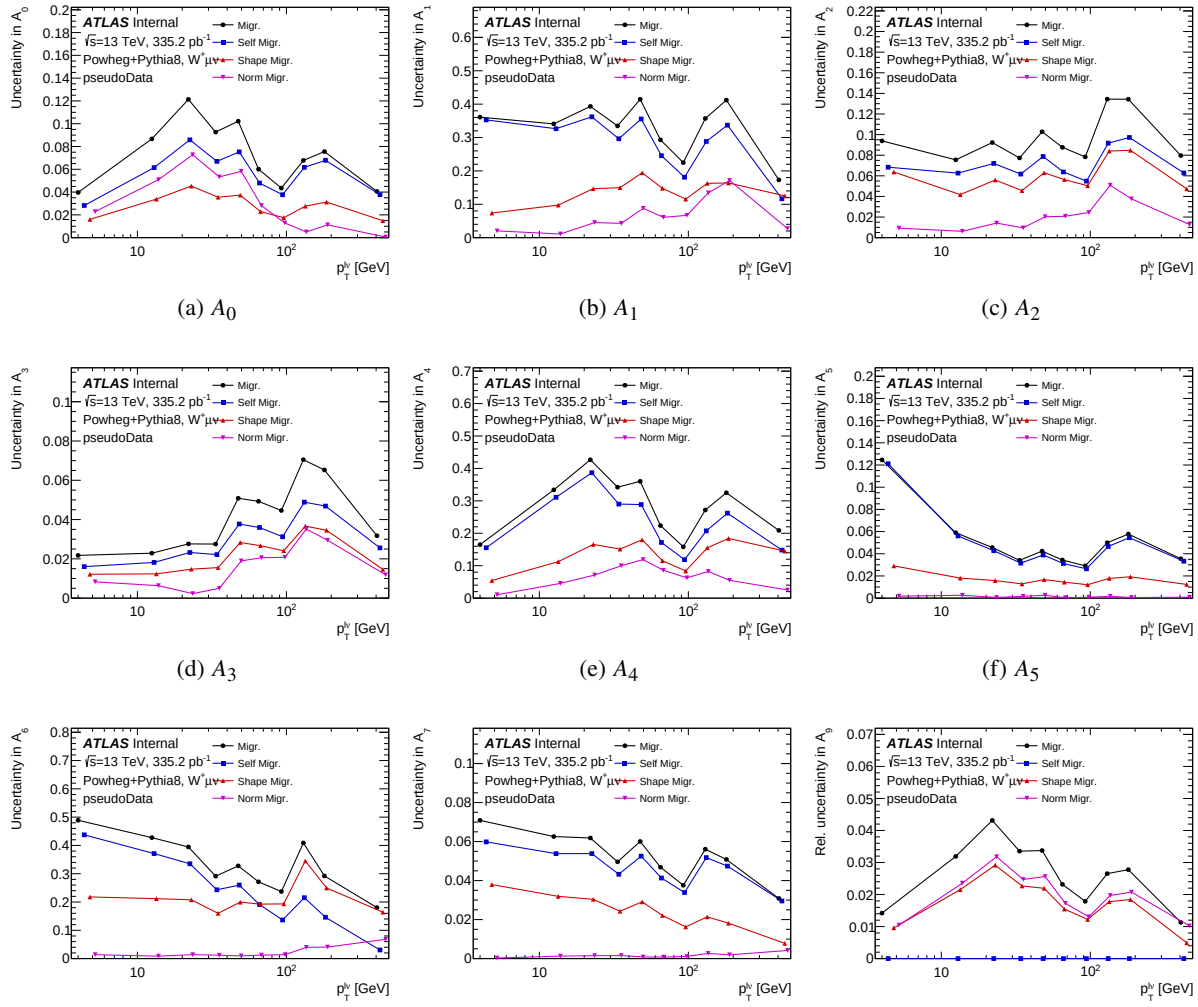


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 106: The uncertainty comparison for migration related uncertainties in the $W^+ \to e^+\nu$ channel for measurement in p_{T}^{W}. Blue is the self migration between $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.



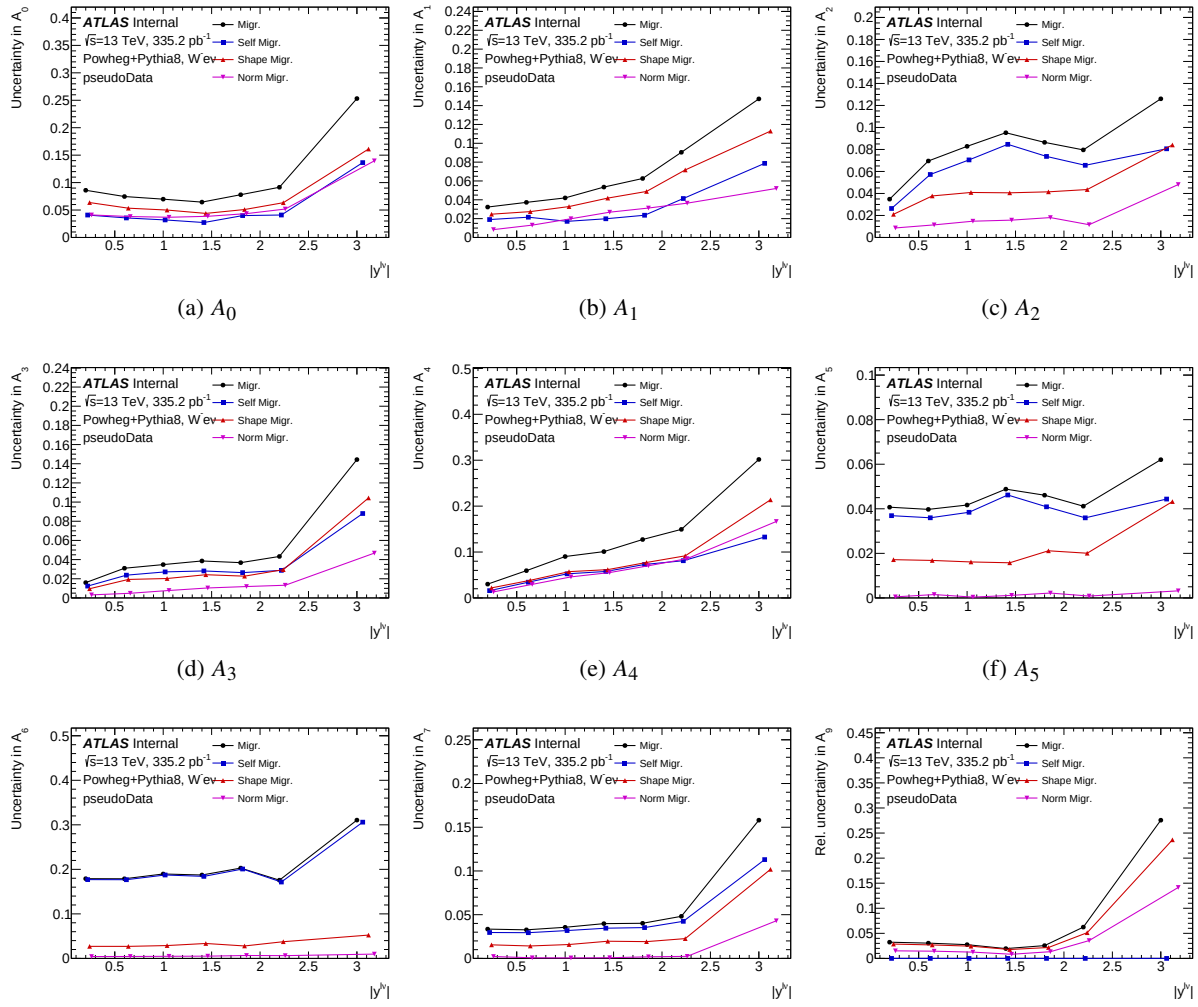
(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 107: The uncertainty comparison for migration related uncertainties in the $W^+ \to \mu^+\nu$ channel for measurement in p_{T}^{W}. Blue is the self migration between $p_{\mathrm{T}}^{W,\mathrm{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

D.2 y^W Differential

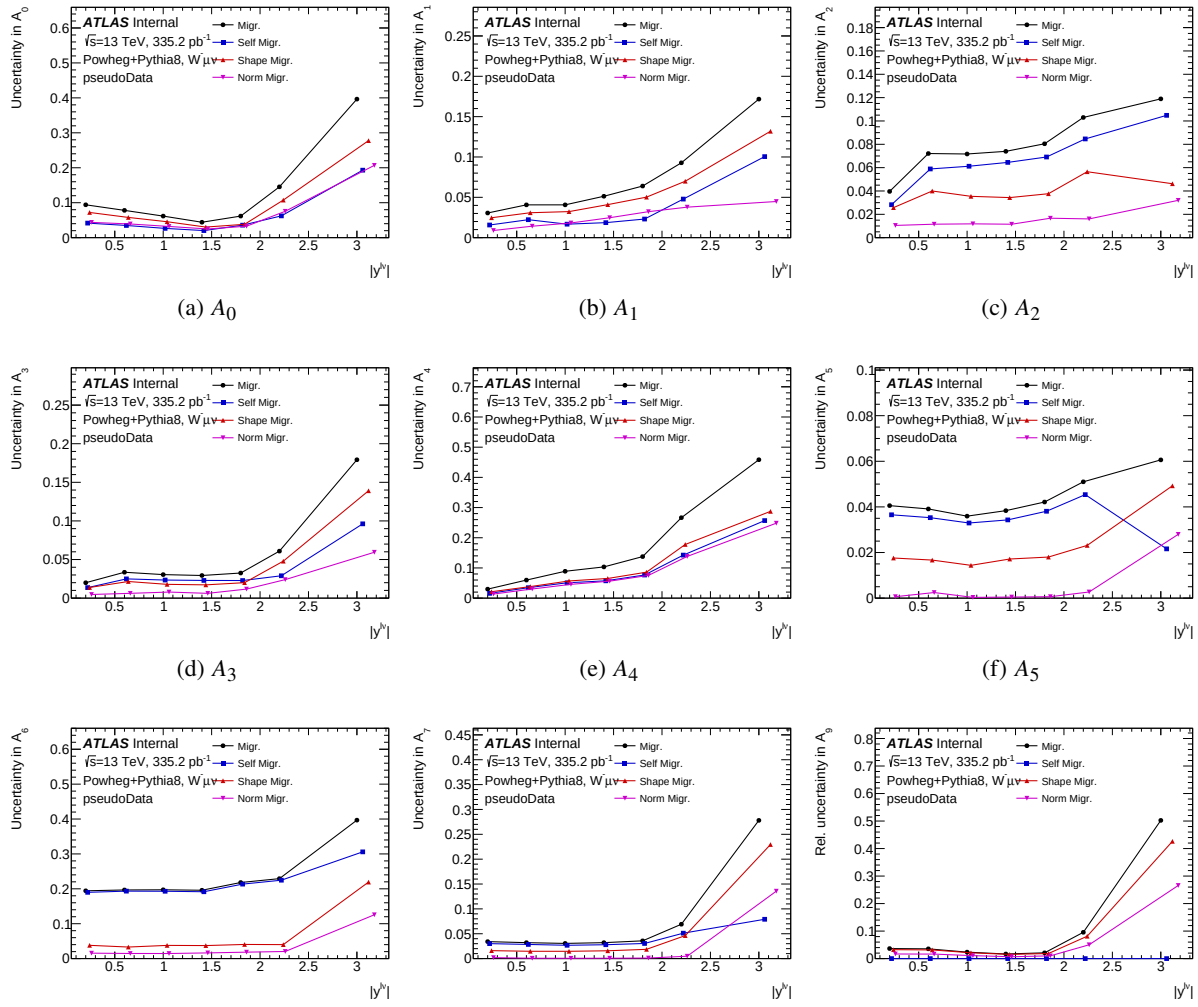


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 108: The uncertainty comparison for migration related uncertainties in the $W^- \to e^-\overline{\nu}$ channel for measurement in y^W. Blue is the self migration between $y^{W,\text{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

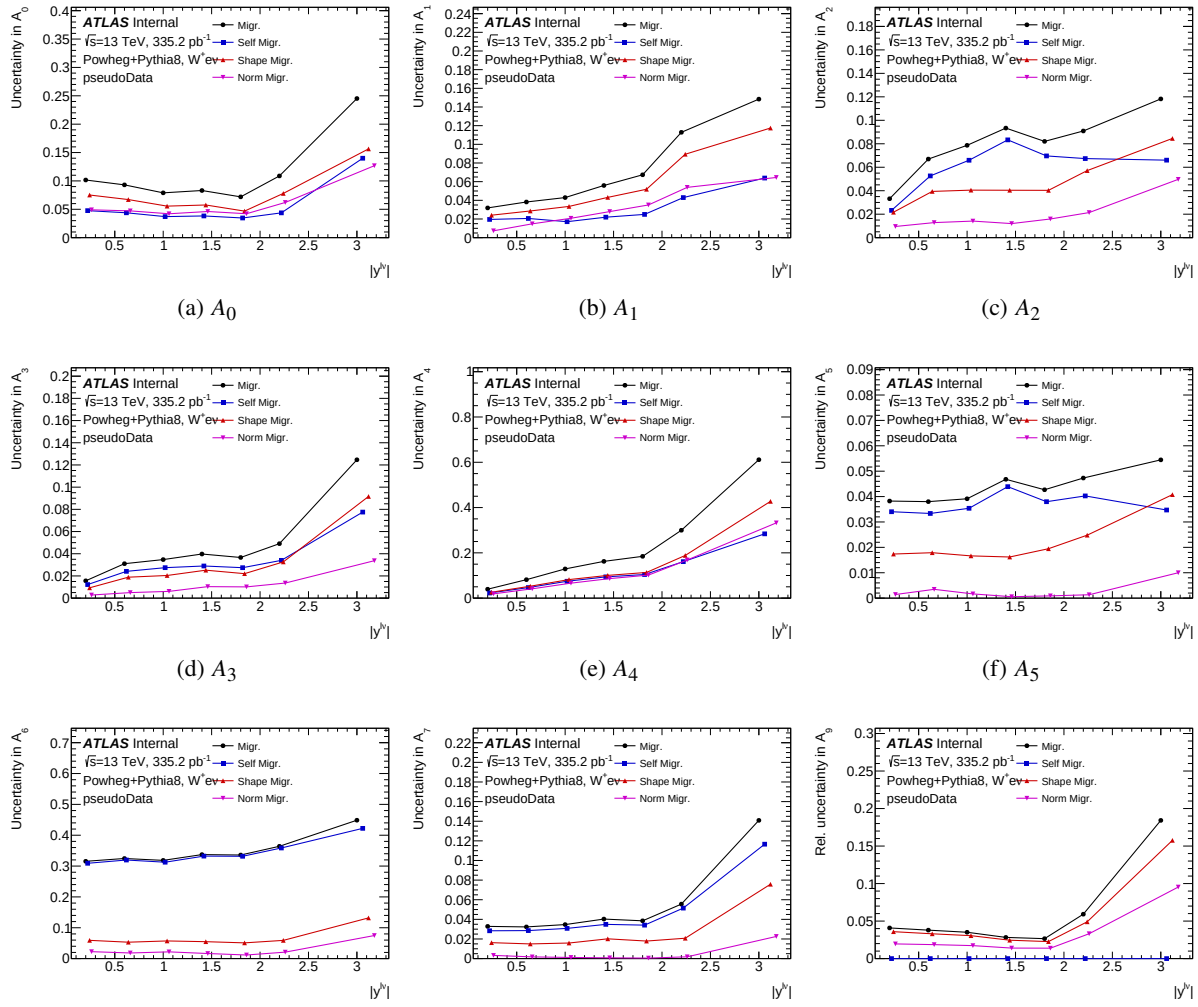


(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

Figure 109: The uncertainty comparison for migration related uncertainties in the $W^- \to \mu^- \overline{\nu}$ channel for measurement in y^W. Blue is the self migration between $y^{W,\text{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

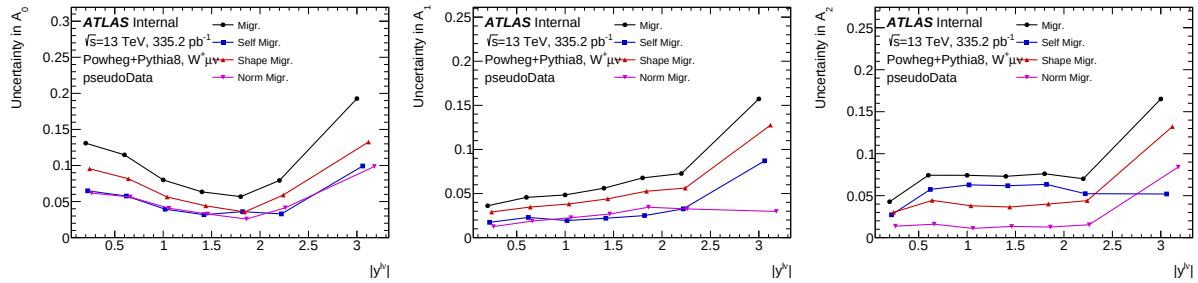


(a) A_0 (b) A_1 (c) A_2

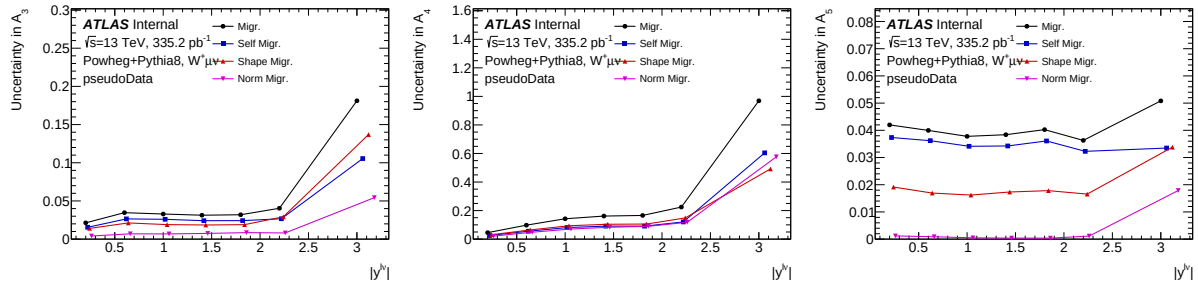
(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7 (i) A_9

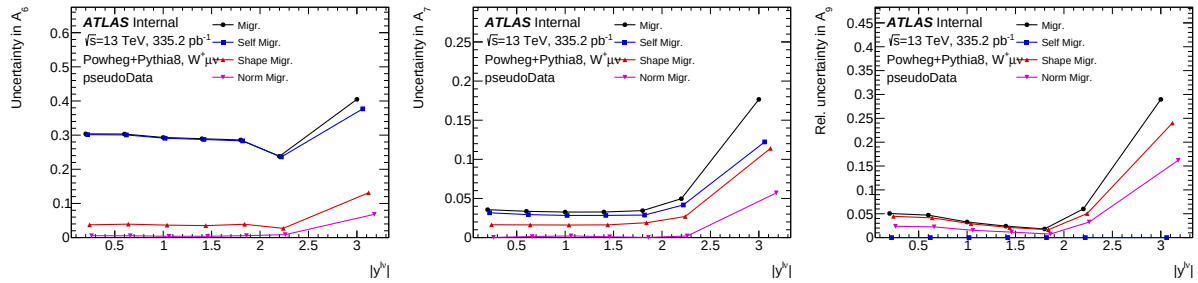
Figure 110: The uncertainty comparison for migration related uncertainties in the $W^+ \to e^+\nu$ channel for measurement in y^W. Blue is the self migration between $y^{W,\text{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.



(a) A_0 (b) A_1 (c) A_2



(d) A_3 (e) A_4 (f) A_5



(g) A_6 (h) A_7 (i) A_9

Figure 111: The uncertainty comparison for migration related uncertainties in the $W^+ \to \mu^+\nu$ channel for measurement in y^W. Blue is the self migration between $y^{W,\text{Truth}}$ bins for the same coefficient, red is the shape migration for migration between coefficients, and pink is the self migration for the normalization. Black is the combination of the all of these migration uncertainty components.

E Multi Jet Figures

Place your MJ studies here.

E.1 Recoil isolation correction

The recoil reconstruction, described in detail in Ref. [43], uses all PFOs in the event, excluding a cone of $\Delta R = 0.2$ around the selected leptons. To compensate for the underlying event and pile-up energy activity removed by this procedure, another same-size cone of activity in the event is selected, centred at the same η , but away in azimuth from any lepton and from hard activity; a vector, of magnitude given by the energy observed in this cone and oriented along the removed lepton, is added to the recoil. This procedure is successful in the signal region, where leptons are isolated and typically surrounded by soft activity.

However, for events in anti-isolated regions the selected lepton is on average close to a jet, and the cone replacement procedure fails. For example, for background events with a selected lepton of $p_T^\ell > 30$ GeV and $ptcone20/p_T > 0.1$, the total track energy around the lepton is at least $0.1 \times p_T^\ell \sim 3$ GeV. This energy is removed in the recoil reconstruction, and replaced by an inappropriately soft contribution. This effect biases the recoil, hence E_T^{miss}, m_T and $|\Delta\phi(\ell, E_T^{miss})|$, and worsens for increasing isolation values.

The projection of the recoil on the lepton momentum direction, $u_\parallel^\ell$, is most affected, as this is the direction affected by the cone replacement procedure. This projection is of particular importance for the reconstruction of E_T^{miss} and the W transverse mass.

This effect is corrected for in a dedicated hadronic recoil reconstruction, used for the present background analysis. For a given lepton, $isocone20$ ($ptcone20$ or $topoetcone20$), defined as the sum of the transverse momenta of tracks (calorimeter deposit) in a cone of $\Delta R = 0.2$ around it, is correlated to the removed energy around the lepton. This suggests the following correction:

$$\vec{u}^{corr} = \vec{u}^{baseline} + \vec{u}^{iso}, where \tag{60}$$

$$\vec{u}^{iso} \equiv k_{iso} \times isocone20 \cdot \vec{u}_\ell \tag{61}$$

where $\vec{u}^{baseline}$ is the standard recoil reconstruction, $\vec{u}^{iso}$ represents the present correction, and $\vec{u}_\ell$ is the unit vector aligned with the lepton direction and k_{iso} is a parameter accounting for a possible difference in scale between the $isocone20$ energy and the energy removed around the lepton. As the correlation coefficient between $isocone20$ and the energy of removed hard activities, k_{iso} has to be measured with data. The effect of this correction is by construction negligible when approaching signal region, hence not affecting the signal templates, used in the fraction fits, in a significant way.

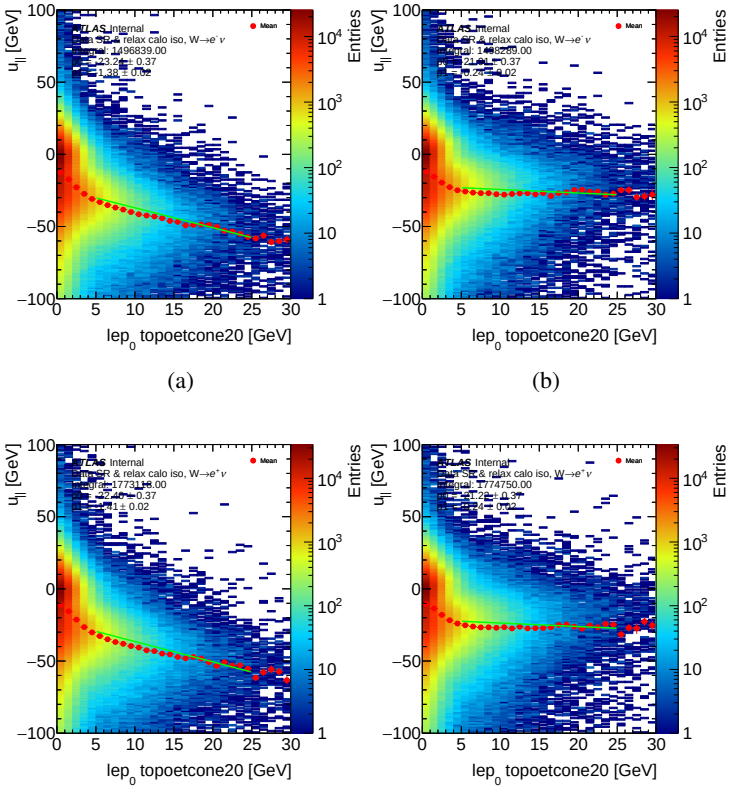
To determine k_{iso} it is assumed that the average value of $u_\parallel^\ell$ should be independent of the anti-isolation value. So $< u_\parallel^{\ell\ corr} >$ should be a constant after this correction. Denoting b_{iso} this constant:

$$\begin{aligned} < u_\parallel^{\ell\ corr} > &= < u_\parallel^{\ell\ baseline} > + < u_\parallel^{\ell\ iso} > \equiv b_{iso} \\ < u_\parallel^{\ell\ corr} > &= < u_\parallel^{\ell\ baseline} > + k_{iso} < isocone20 > \equiv b_{iso} \\ < u_\parallel^{\ell\ baseline} > &= -k_{iso} < isocone20 > + b_{iso} \end{aligned} \tag{62}$$

k_{iso} parameter	$W^- \to e^-\overline{\nu}$	$W^+ \to e^+\nu$	$W^- \to \mu^-\overline{\nu}$	$W^+ \to \mu^+\nu$
$ptcone20$	-	-	1.51 ± 0.02	1.50 ± 0.02
$topoetcone20$	1.38 ± 0.02	1.41 ± 0.02	-	-

Table 7: The fitted values of k_{iso}, the uncertainty are the statistical uncertainty from the fitting.

While this assumption may not hold to high accuracy, as the different isolation slices have different properties in terms of jet structure, potentially affecting $u_{\parallel}^{\ell}$, it is justified to compensate for the mis-match in the energy replacement procedure for anti-isolated leptons. So the correction factor k_{iso} can be fitted from the linear relationship between $< u_{\parallel}^{\ell\ baseline} >$ and $isocone20$. The effect of this correction is illustrated in Figure 112 and 113, which shows a large dependence of $< u_{\parallel}^{\ell} >$ on the value of $isocone20$ before correction and that this dependence is significantly reduced after correction. The k_{iso} value fitted for a given channel is the opposite of the slope obtained from the linear fit between $[5, 25]$ GeV, the results of which are listed in Table 7.



(a) (b)

(c) (d)

Figure 112: $< u_{\parallel}^{\ell} >$ as a function of $topoetcone20$, before (left) and after (right) correction in FR + CR1 for $W^- \to e^-\overline{\nu}$ (a, b) and $W^+ \to e^+\nu$ channel (c, d).

Example fraction fits before and after this correction are compared in Figure 114 and 115 for $W^- \to e^-\overline{\nu}$

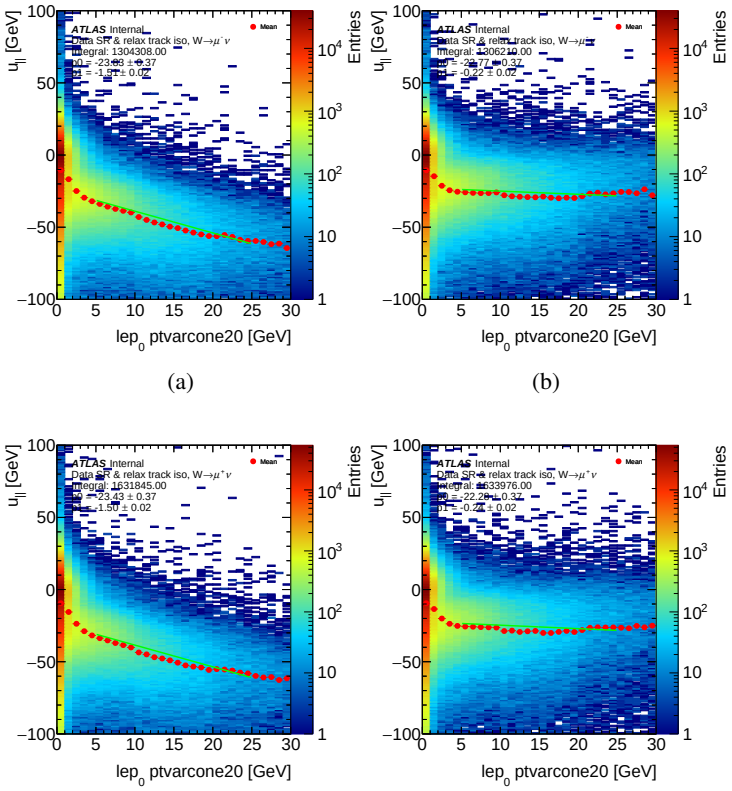


Figure 113: $< u_{\parallel}^{\ell} >$ as a function of $ptvarcone20$, before (left) and after (right) correction in FR + CR1 for $W^- \to \mu^- \overline{\nu}$ (a, b) and $W^+ \to \mu^+ \nu$ channel (c, d).

with the first and the third anti-isolation slices. The m_T, E_T^{miss} and $|\Delta\phi(\ell, E_T^{miss})|$ fits are visibly improved, as these variables directly benefit from the improved recoil definition; this is not the case for p_T^{ℓ}, which shows satisfactory agreement before the correction. The mis-modelling observed at high values for all distributions is driven by a mismodelling of the p_T^W distribution in the simulation.

Figures 116 and 117 show the effect of this correction on the $ptvarcone20/p_T$ (for $\mu^{\pm}$) and $topoetcone20/p_T$ (for $e^{\pm}$) scan in the electron and muon channels. The dependence of the fitted yield on the isolation range is affected.

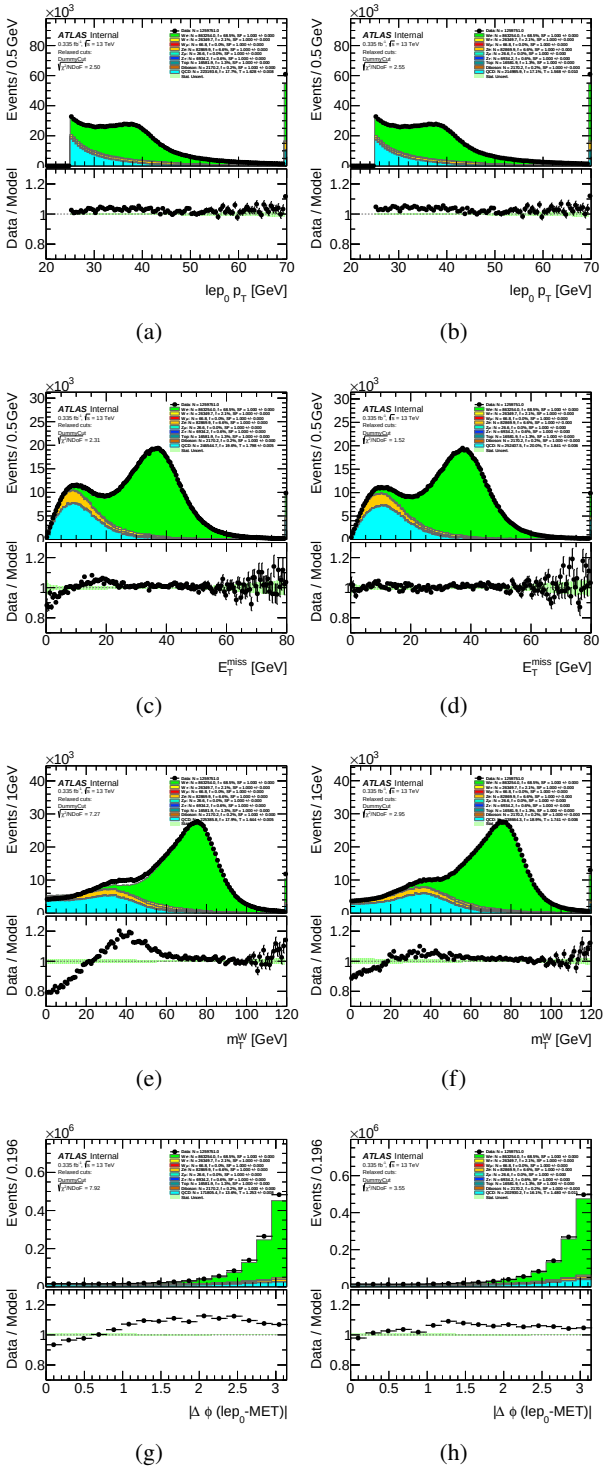


Figure 114: Multi-jet background template fits performed in distributions of $p_T^{e^-}$ (top), E_T^{miss} (middle) and m_T (bottom) with the MJ template obtained from isolation slice $0.05 < topoetcone20/p_T < 0.1$. The fits are presented in $W^- \to e^-\overline{\nu}$ channel before (left) and after (right) recoil correction.

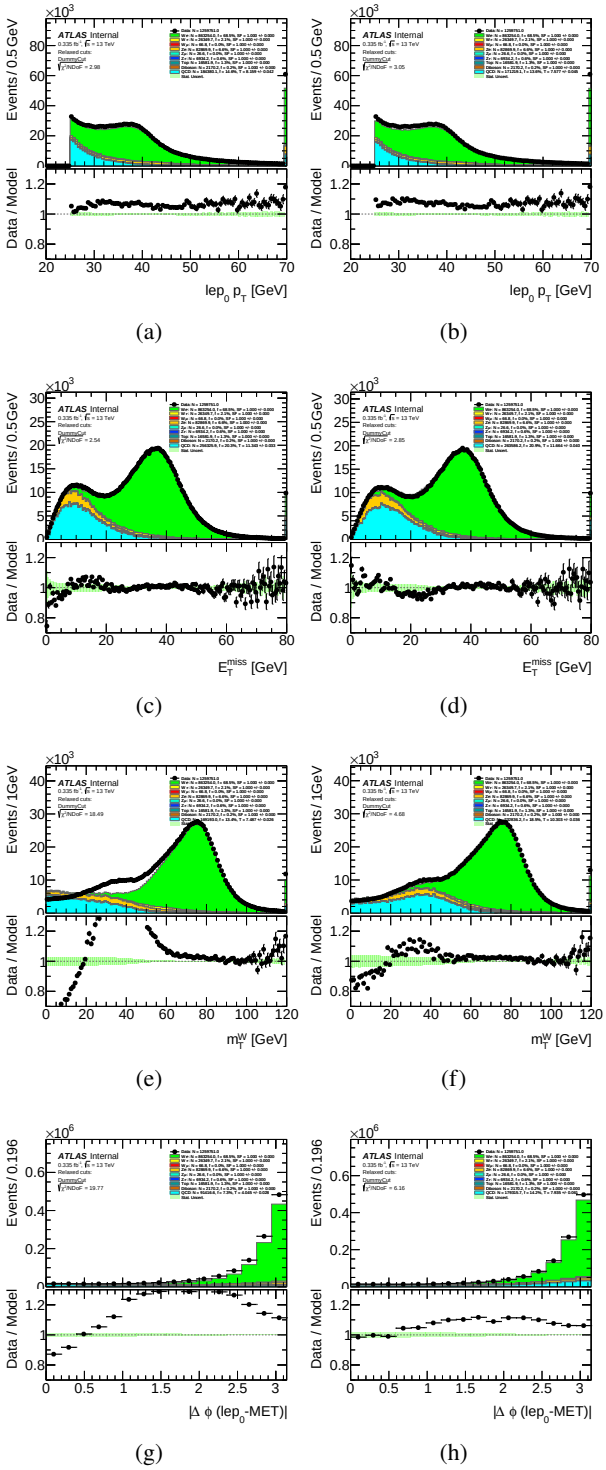
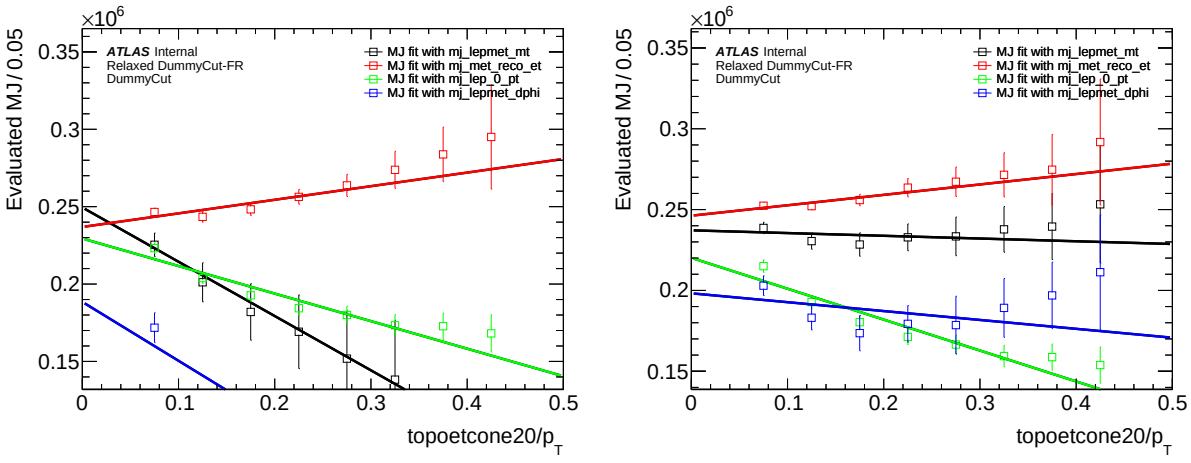
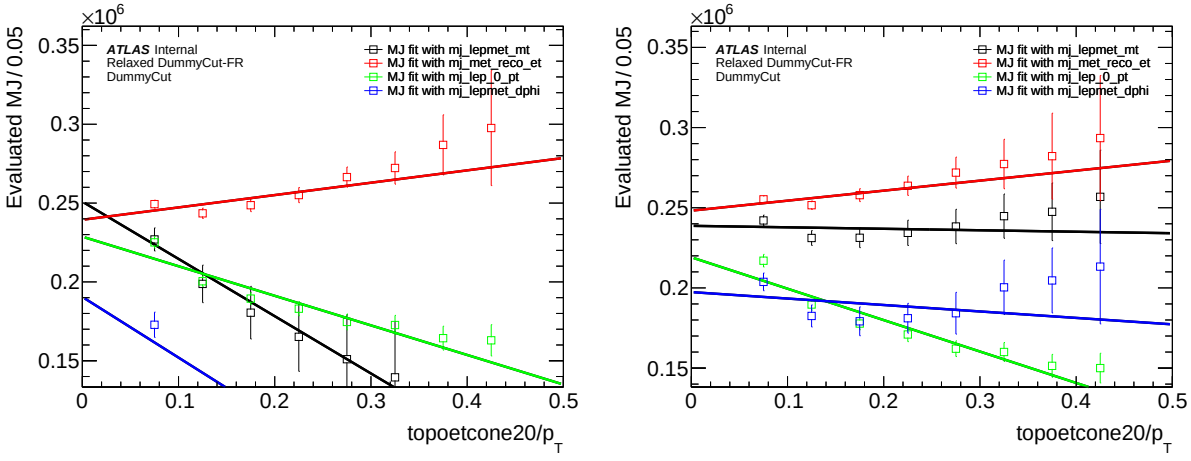


Figure 115: Multi-jet background template fits performed in distributions of $p_T^{e^-}$ (top), E_T^{miss} (middle) and m_T (bottom) with the MJ template obtained from isolation slice $0.2 < topoetcone20/p_T < 0.25$. The fits are presented in $W^- \to e^-\overline{\nu}$ channel before (left) and after (right) recoil correction.



(a) (b)



(c) (d)

Figure 116: The estimated total number of the multijet background in SR scanned as a function of the isolation slice. Isolation scan in $W^- \to e^-\overline{\nu}$ (top) and $W^+ \to e^+\nu$ (bottom) channels without (left) and with (right) recoil correction. Each point is plotted as average $topoetcone20/p_T$ of each dlice in CR2.

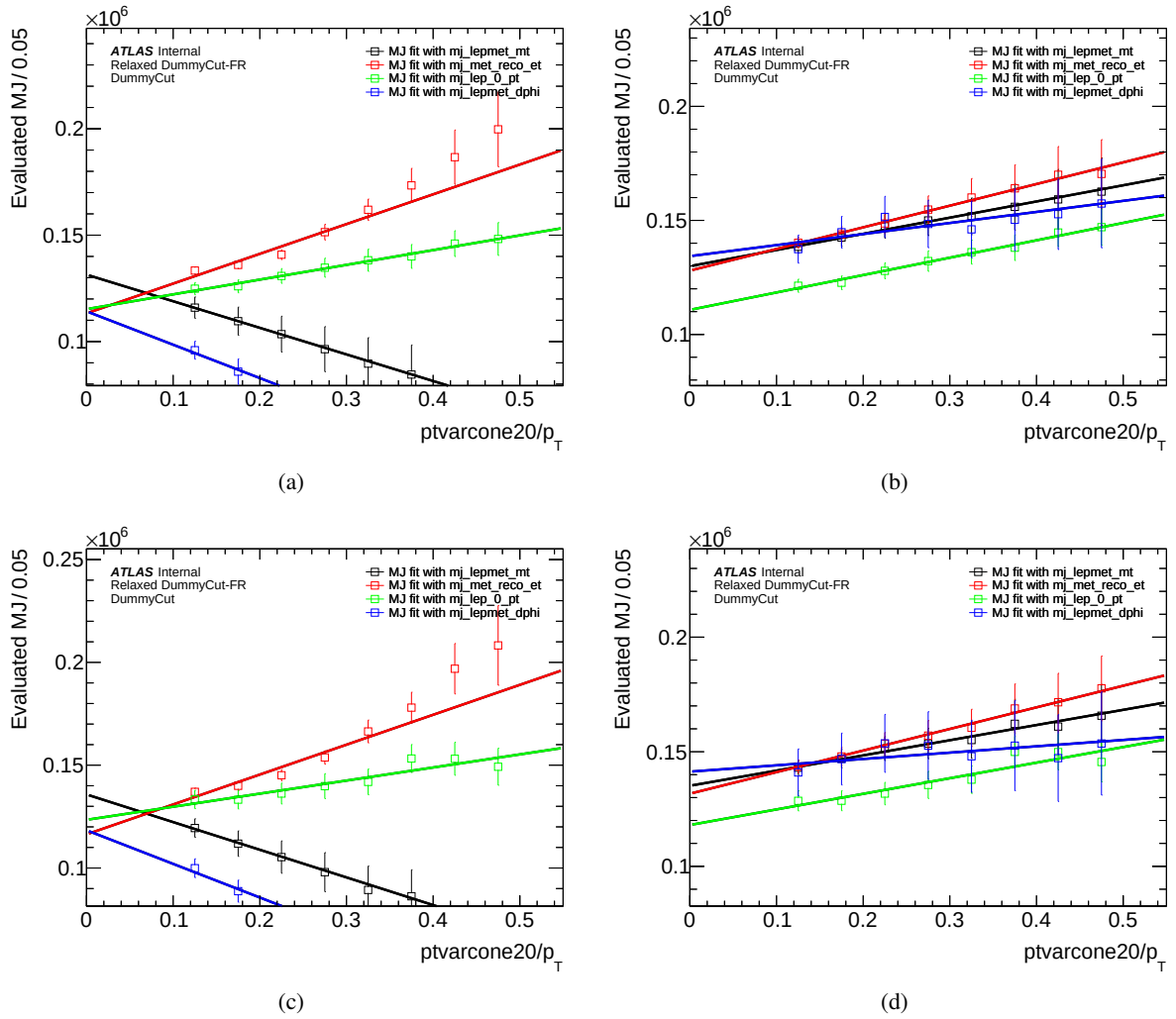


Figure 117: The estimated total number of the multijet background in SR scanned as a function of the isolation slice. Isolation scan in $W^- \to \mu^- \overline{\nu}$ (top) and $W^+ \to \mu^+ \nu$ (bottom) channels without (left) and with (right) recoil correction. Each point is plotted as average $ptvarcone20/p_T$ of each slice in CR2.

E.2 Choice of the track-based and calo-based isolation variables scans

This analysis uses different isolation variables for $e^{\pm}$ and $\mu^{\pm}$ channels:

- Slicing $ptvarcone20/pt$ for $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$: 8 slicing bins from 0.1 to 0.5
- Slicing in $topoetcone20/pt$ for $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$: 8 slicing bins from 0.05 to 0.45

In this chapter highlights why this decision has been made.

This analysis is based on multijet background estimation methodology same as used by WpT analysis (see [44]). Early multijet studies were focused on the track-based isolation variable. However, data/MC discrepancies for $\cos\theta_{CS}$ distributions in the $W^- \to e^-\overline{\nu}$ and $W^+ \to e^+\nu$ channels has been observed. In case of electrons it is shown on Figure 118(a). Effectively it had pointed to the problems in tails of the lepton η distributions. There are no such data/MC discrepancies observed for $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels. Example for muon channel is shown on Figure 118(b).

This issue comes from the multijet template shape estimation. Replacing track-based isolation variable ($ptvarcone20/p_T$) using same cone sized calo-based isolation variable ($topoetcone20/p_T$) fixes this. Results on this are shown on Figures 118(c) and 118(d) for $W^- \to e^-\overline{\nu}$ and $W^- \to \mu^-\overline{\nu}$ channels respectively.

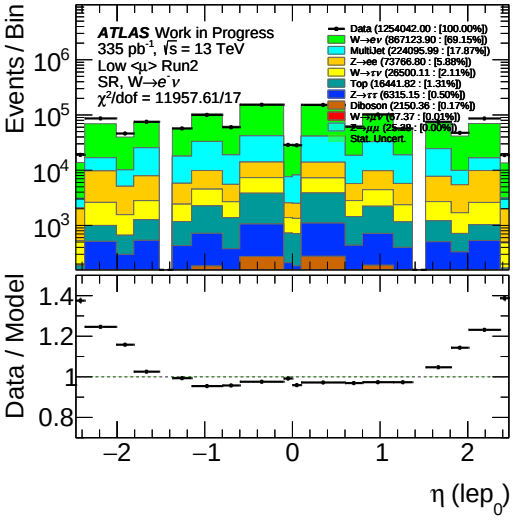
Impact on the evaluated normalized MJ shapes in anti-isolated slices of CR1 for η distribution for $W^- \to e^-\overline{\nu}$ as a function of track-based and calo-based isolation variables on Figure 119. This might happen due to tight cut on $topoetcone20/p_T$ during the $ptvarcone20/p_T$ anti-isolation slice scan.

Natural step would be to use same $topoetcone20/p_T$ isolation variable for all channels. However it was discovered that $topoetcone20$ is not properly modelled in the signal samples for $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels. Figure 120 shows $topoetcone20/p_T$ distributions for $Z \to ee$ and $Z \to \mu\mu$ selections. The 20% data/background discrepancy for central rapidity in the $Z \to \mu\mu$ is shown. There are no this feature for $Z \to ee$ channel.

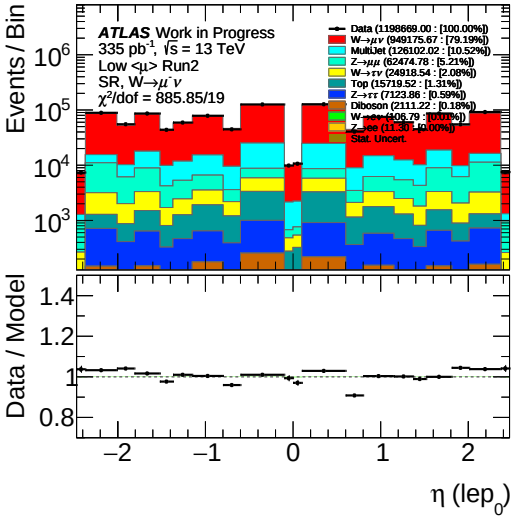
Same effect is observed for the first anti-isolated slice for the $W^- \to \mu^-\overline{\nu}$ and $W^+ \to \mu^+\nu$ channels. Figure 121 shows evaluation of the multi-jet background template fit in the $W^- \to \mu^-\overline{\nu}$ for the first two anti-isolated slices: $0.05 < topoetcone20/p_T < 0.1$ and $0.1 < topoetcone20/p_T < 0.15$. There are missing signal MC to subtract in the first anti-isolation slice that brings same 20% discrepancy. Impact on other anti-isolated slices is not that big as signal contribution to the CR1 drops. Linear extrapolations from anti-isolated slices for $W^- \to \mu^-\overline{\nu}$ channel are shown on Figure 122(a).

An overall 20% scaling of signal contamination helps to fix linear extrapolations as it is shown on Figure 122(b). However it is not best solution as we got negative shape in m_T^W as it is shown on Figure 123(b).

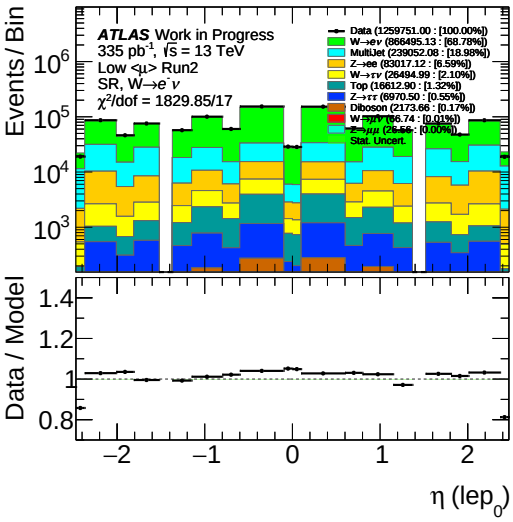
To get it right we should correct the MC mismodelling of $topoetcone20$, but in a meantime decided to use $ptvarcone20/p_T$ anti-isolation for muons and anti-muons channels.



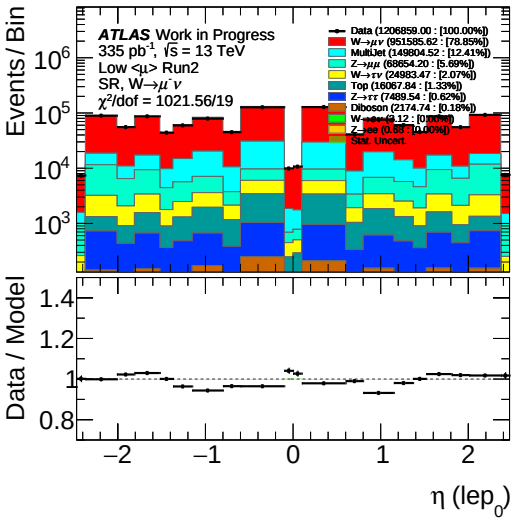
(a)



(b)

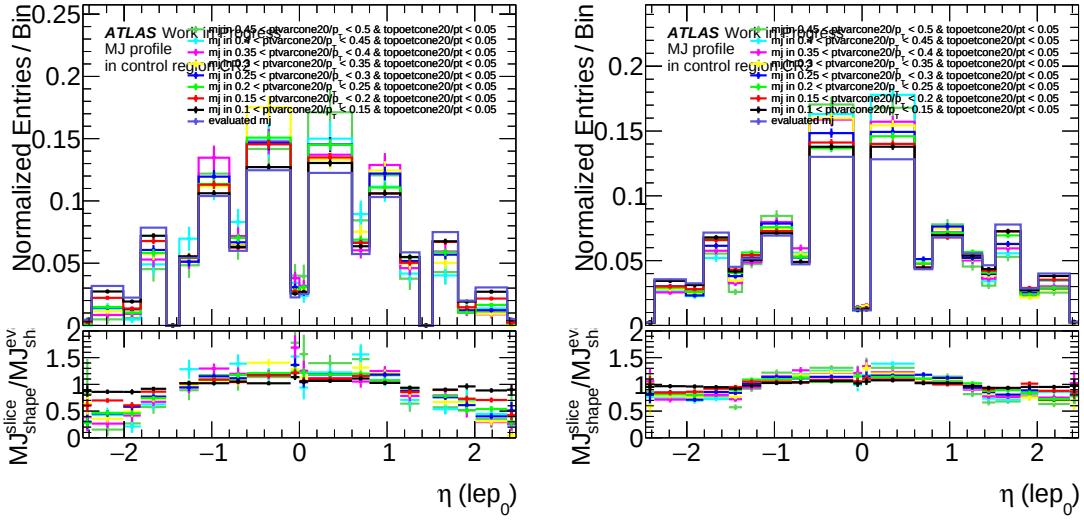


(c)

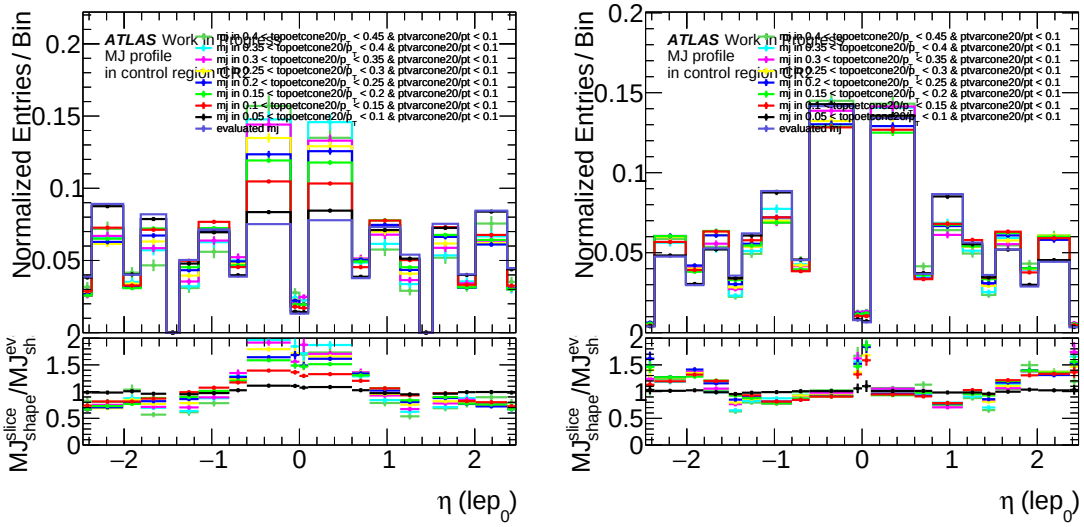


(d)

Figure 118: The η distributions for $W^- \to e^-\overline{\nu}$ (left) and $W^- \to \mu^-\overline{\nu}$ (right) channels using multijet estimation slicing over $ptvarcone20/p_T$ (top) and $topoetcone20/p_T$ (bottom).

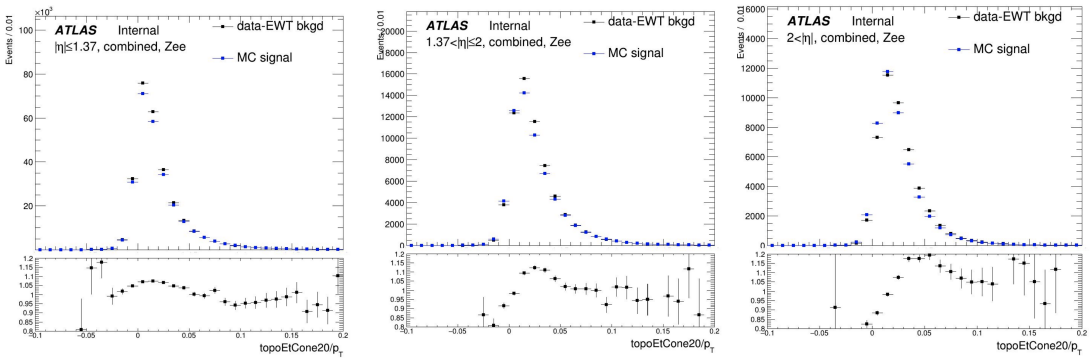


(a) (b)

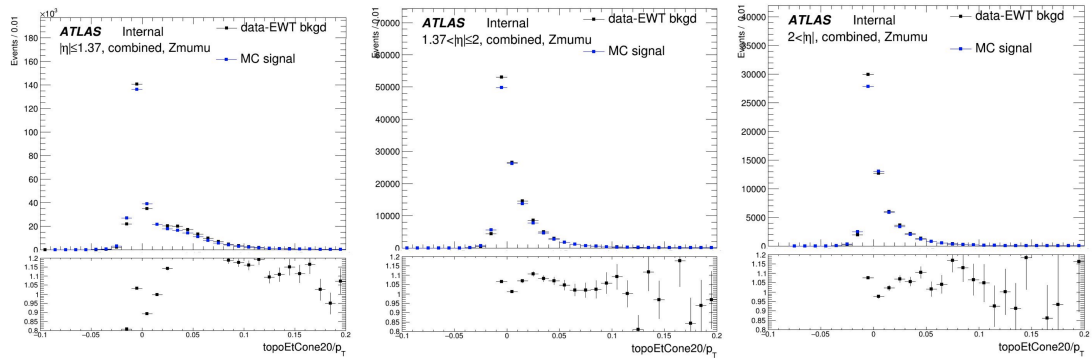


(c) (d)

Figure 119: Normalized MJ shapes in anti-isolated slices of CR1 for η distribution for $W^- \rightarrow e^-\overline{\nu}$ (left) and $W^- \rightarrow \mu^-\overline{\nu}$ (right) channels using multijet estimation slicing over $ptvarcone20/p_T$ (top) and $topoetcone20/p_T$ (bottom).

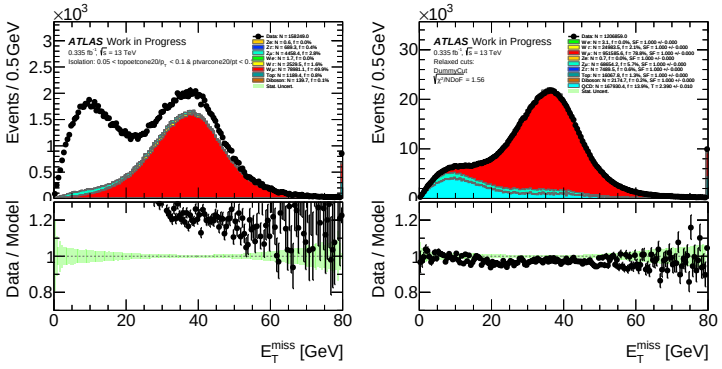


(a) (b) (c)

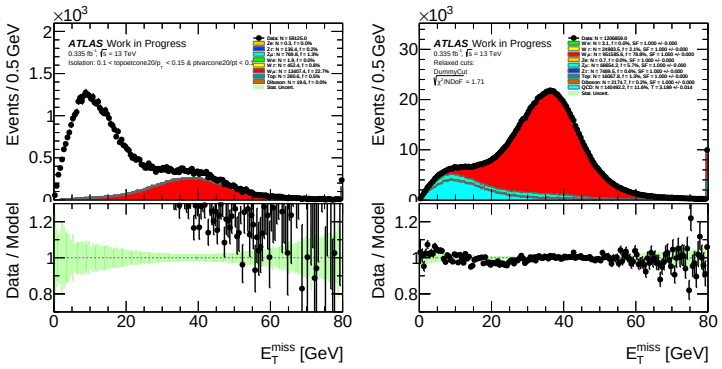


(d) (e) (f)

Figure 120: The relative $topoetcone20/p_T$ distributions for $Z \to ee$ and $Z \to \mu\mu$ selections.

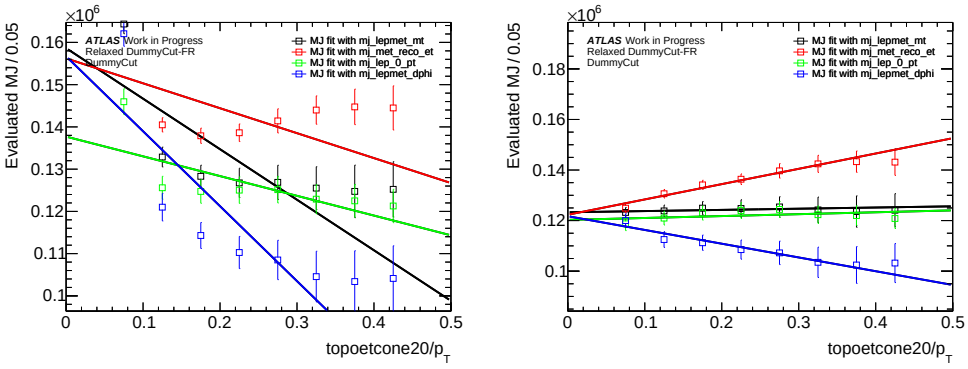


(a) (b)



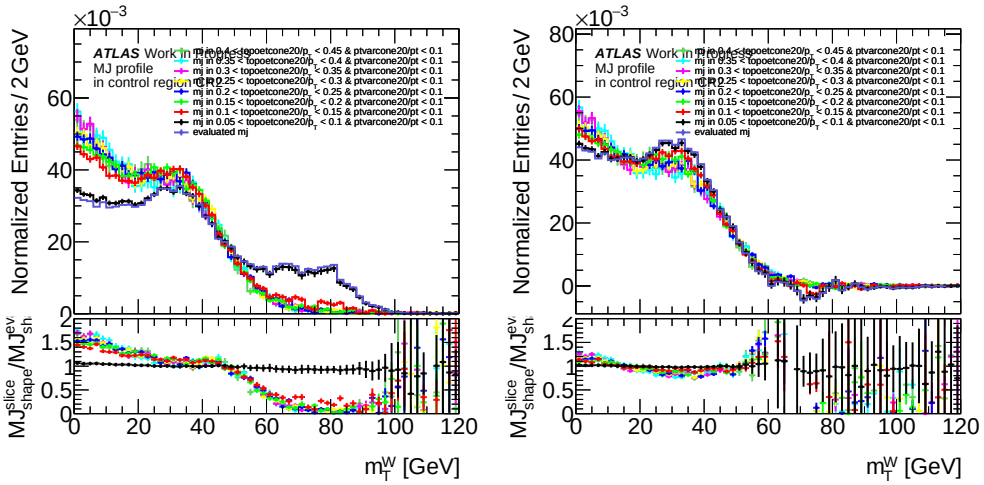
(c) (d)

Figure 121: The multi-jet background template fit in the $W^- \rightarrow \mu^- \overline{\nu}$ for the anti-isolated slices $0.05 < topoetcone20/p_T < 0.1$ (top) and $0.1 < topoetcone20/p_T < 0.15$ (bottom). Two regions are shown: CR1 (left) and fit region (right).



(a) No MC rescale (b) + 20% for MC

Figure 122: Linear extrapolations from anti-isolated slices for $W^- \rightarrow \mu^- \overline{\nu}$ channel before (left) and after (right) 20% rescale of the MC contribution to the CR1.



(a) No MC rescale

(b) + 20% for MC

Figure 123: Normalized MJ shapes in anti-isolated slices of CR1 for $W^- \to \mu^- \overline{\nu}$ channel before (left) and after (right) 20% rescale of the MC contribution to the CR1.

E.3 Multi-jet background template fits

Despite the fact in this analysis we use relaxed kinematic cuts for the signal region comparing to the WpT analysis [44], we keeping same naming convention for several samples are defined, corresponding to different regions in phase space and isolation. The cuts definitions for control regions are listed in Section 5.1.

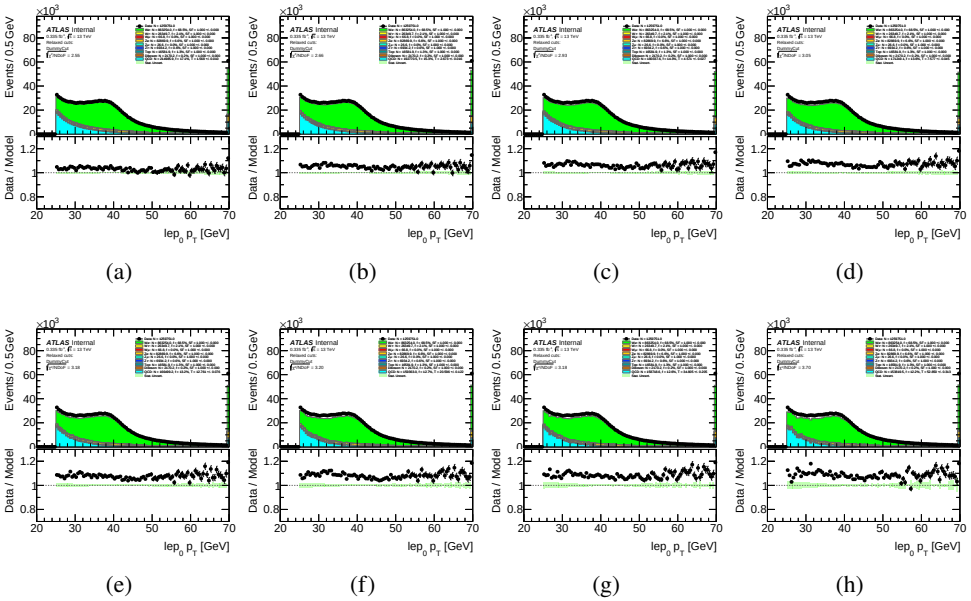
E.3.1 Electron channel multi-jet background template fits

u_T Cut	m_T^W fits b	m_T^W fits k	E_T^{miss} fits b	E_T^{miss} fits k	p_T fits b	p_T fits k	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits b	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits k
–	237173 ± 4899	-16912 ± 32212	246225 ± 3049	64434 ± 23147	220160 ± 3918	-191203 ± 18749	198207 ± 7539	-54934 ± 43369

Table 8: k and b are parameters for the electron extrapolations on 13 TeV. The uncertainty of these two parameters are provided by the extrapolation, from the statistical uncertainty of the parameter T and the number of events in CR of each slices and multiplied by $\sqrt{\chi^2/NDoF}$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1259751 ± 1122	998253 ± 330	222831 ± 472	85732 ± 97	137099 ± 482	1.568 ± 0.010	1.000 ± 0.000	582.97	90	2.55	222831 ± 472	85732 ± 97	137099 ± 482	214986 ± 3878
–	1	0.1	0.15	1259751 ± 1122	998253 ± 330	83148 ± 288	11023 ± 35	72125 ± 291	2.673 ± 0.016	1.000 ± 0.000	638.20	90	2.66	83148 ± 288	11023 ± 35	72125 ± 291	192770 ± 3731
–	2	0.15	0.2	1259751 ± 1122	998253 ± 330	41682 ± 204	2230 ± 16	39452 ± 205	4.571 ± 0.027	1.000 ± 0.000	772.86	90	2.93	41682 ± 204	2230 ± 16	39452 ± 205	180338 ± 4192
–	3	0.2	0.25	1259751 ± 1122	998253 ± 330	23298 ± 153	700 ± 9	22598 ± 153	7.577 ± 0.045	1.000 ± 0.000	838.65	90	3.05	23298 ± 153	700 ± 9	22598 ± 153	171219 ± 4713
–	4	0.25	0.3	1259751 ± 1122	998253 ± 330	13388 ± 116	340 ± 6	13048 ± 116	12.761 ± 0.076	1.000 ± 0.000	908.53	90	3.18	13388 ± 116	340 ± 6	13048 ± 116	166499 ± 5650
–	5	0.3	0.35	1259751 ± 1122	998253 ± 330	7932 ± 89	194 ± 5	7738 ± 89	20.596 ± 0.122	1.000 ± 0.000	918.94	90	3.20	7932 ± 89	194 ± 5	7738 ± 89	159363 ± 6601
–	6	0.35	0.4	1259751 ± 1122	998253 ± 330	4688 ± 68	128 ± 4	4560 ± 69	34.805 ± 0.205	1.000 ± 0.000	909.12	90	3.18	4688 ± 68	128 ± 4	4560 ± 69	158719 ± 8145
–	7	0.4	0.45	1259751 ± 1122	998253 ± 330	2998 ± 55	88 ± 3	2910 ± 55	52.850 ± 0.313	1.000 ± 0.000	1230.99	90	3.70	2998 ± 55	88 ± 3	2910 ± 55	153820 ± 11233

Table 9: Numbers for the electron p_T fits on 13 TeV. Numbers in FR are not affected by $topoetcone20/p_T$ Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



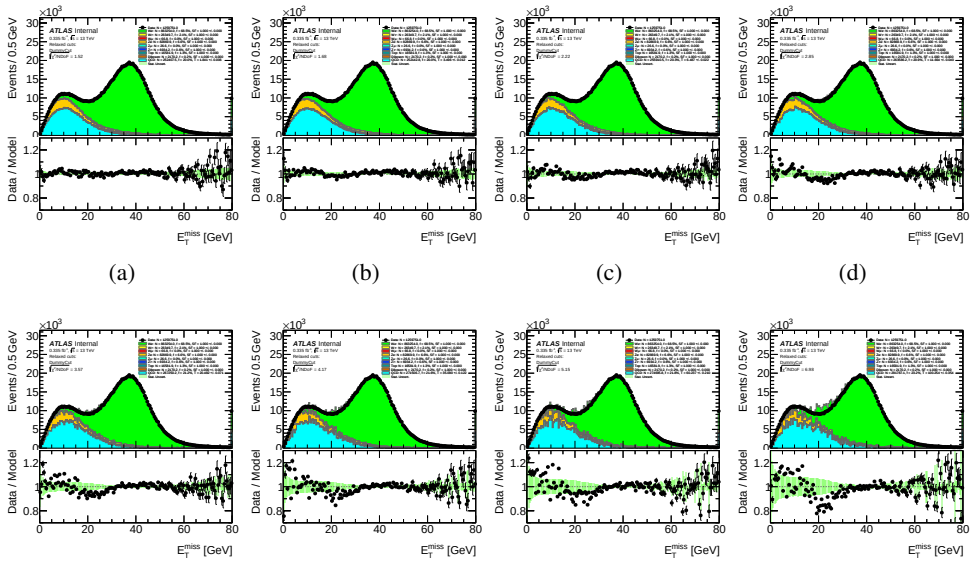
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 124: Multi-jet background template fits performed in distributions of p_T^ℓ of the electron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

m_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1259751 ± 1122	998253 ± 330	222831 ± 472	85732 ± 97	137099 ± 482	1.841 ± 0.006	1.000 ± 0.000	369.84	160	1.52	222831 ± 472	85732 ± 97	137099 ± 482	252408 ± 1889
–	1	0.1	0.15	1259751 ± 1122	998253 ± 330	83148 ± 288	11023 ± 35	72125 ± 291	3.496 ± 0.012	1.000 ± 0.000	453.36	160	1.68	83148 ± 288	11023 ± 35	72125 ± 291	252143 ± 2247
–	2	0.15	0.2	1259751 ± 1122	998253 ± 330	41682 ± 204	2230 ± 16	39452 ± 205	6.487 ± 0.022	1.000 ± 0.000	787.33	160	2.22	41682 ± 204	2230 ± 16	39452 ± 205	255910 ± 3535
–	3	0.2	0.25	1259751 ± 1122	998253 ± 330	23298 ± 153	700 ± 9	22598 ± 153	11.664 ± 0.040	1.000 ± 0.000	1302.68	160	2.85	23298 ± 153	700 ± 9	22598 ± 153	263586 ± 5712
–	4	0.25	0.3	1259751 ± 1122	998253 ± 330	13388 ± 116	340 ± 6	13048 ± 116	20.482 ± 0.071	1.000 ± 0.000	2035.08	160	3.57	13388 ± 116	340 ± 6	13048 ± 116	267250 ± 9085
–	5	0.3	0.35	1259751 ± 1122	998253 ± 330	7932 ± 89	194 ± 5	7738 ± 89	35.090 ± 0.122	1.000 ± 0.000	2788.86	160	4.17	7932 ± 89	194 ± 5	7738 ± 89	271507 ± 13650
–	6	0.35	0.4	1259751 ± 1122	998253 ± 330	4688 ± 68	128 ± 4	4560 ± 69	60.237 ± 0.210	1.000 ± 0.000	4245.53	160	5.15	4688 ± 68	128 ± 4	4560 ± 69	274696 ± 21840
–	7	0.4	0.45	1259751 ± 1122	998253 ± 330	2998 ± 55	88 ± 3	2910 ± 55	100.254 ± 0.354	1.000 ± 0.000	7784.22	160	6.98	2998 ± 55	88 ± 3	2910 ± 55	291787 ± 39007

Table 10: Numbers for the electron E_T^{miss} fits on 13 TeV. Numbers in FR are not affected by $topoetcone20/p_T$ Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

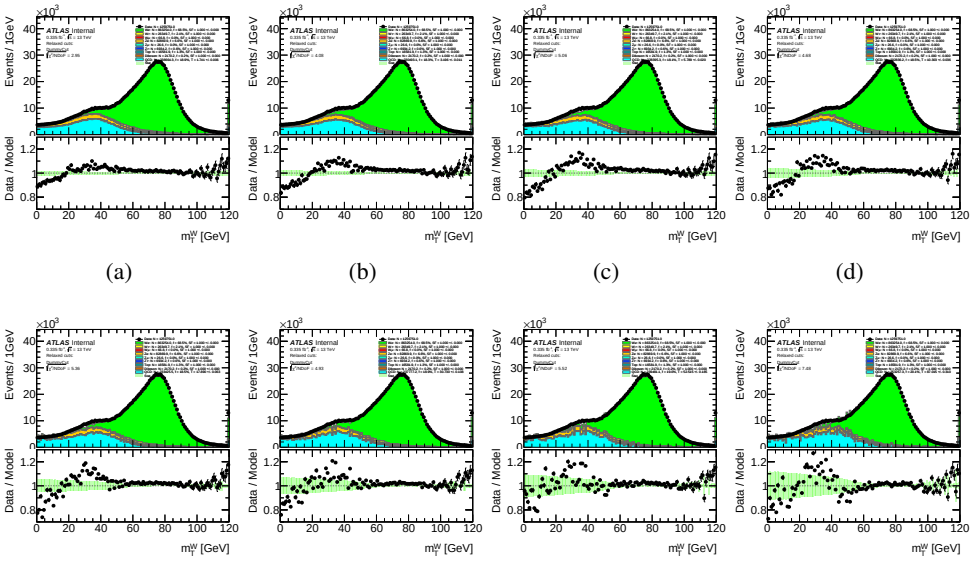
Figure 125: Multi-jet background template fits performed in distributions of E_T^{miss} of the electron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

m_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1259751 ± 1122	998253 ± 330	222831 ± 472	85732 ± 97	137099 ± 482	1.741 ± 0.006	1.000 ± 0.000	1047.64	120	2.95	222831 ± 472	85732 ± 97	137099 ± 482	238664 ± 3505
–	1	0.1	0.15	1259751 ± 1122	998253 ± 330	83148 ± 288	11023 ± 35	72125 ± 291	3.196 ± 0.011	1.000 ± 0.000	1995.10	120	4.08	83148 ± 288	11023 ± 35	72125 ± 291	230493 ± 5016
–	2	0.15	0.2	1259751 ± 1122	998253 ± 330	41682 ± 204	2230 ± 16	39452 ± 205	5.789 ± 0.020	1.000 ± 0.000	3076.53	120	5.06	41682 ± 204	2230 ± 16	39452 ± 205	228395 ± 7245
–	3	0.2	0.25	1259751 ± 1122	998253 ± 330	23298 ± 153	700 ± 9	22598 ± 153	10.303 ± 0.036	1.000 ± 0.000	2627.78	120	4.68	23298 ± 153	700 ± 9	22598 ± 153	232836 ± 8302
–	4	0.25	0.3	1259751 ± 1122	998253 ± 330	13388 ± 116	340 ± 6	13048 ± 116	17.890 ± 0.063	1.000 ± 0.000	3446.51	120	5.36	13388 ± 116	340 ± 6	13048 ± 116	233426 ± 11947
–	5	0.3	0.35	1259751 ± 1122	998253 ± 330	7932 ± 89	194 ± 5	7738 ± 89	30.730 ± 0.108	1.000 ± 0.000	2919.18	120	4.93	7932 ± 89	194 ± 5	7738 ± 89	237777 ± 14136
–	6	0.35	0.4	1259751 ± 1122	998253 ± 330	4688 ± 68	128 ± 4	4560 ± 69	52.515 ± 0.185	1.000 ± 0.000	3656.49	120	5.52	4688 ± 68	128 ± 4	4560 ± 69	239481 ± 20411
–	7	0.4	0.45	1259751 ± 1122	998253 ± 330	2998 ± 55	88 ± 3	2910 ± 55	87.015 ± 0.310	1.000 ± 0.000	6705.47	120	7.48	2998 ± 55	88 ± 3	2910 ± 55	253257 ± 36295

Table 11: Numbers for the electron m_T^W fits on 13 TeV. Numbers in FR are not affected by $topoetcone20/p_T$ Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

m_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1259751 ± 1122	998253 ± 330	222831 ± 472	85732 ± 97	137099 ± 482	1.480 ± 0.011	1.000 ± 0.000	201.62	16	3.55	222831 ± 472	85732 ± 97	137099 ± 482	202930 ± 6118
–	1	0.1	0.15	1259751 ± 1122	998253 ± 330	83148 ± 288	11023 ± 35	72125 ± 291	2.539 ± 0.019	1.000 ± 0.000	368.68	16	4.80	83148 ± 288	11023 ± 35	72125 ± 291	183102 ± 7545
–	2	0.15	0.2	1259751 ± 1122	998253 ± 330	41682 ± 204	2230 ± 16	39452 ± 205	4.399 ± 0.034	1.000 ± 0.000	726.67	16	6.74	41682 ± 204	2230 ± 16	39452 ± 205	173551 ± 10858
–	3	0.2	0.25	1259751 ± 1122	998253 ± 330	23298 ± 153	700 ± 9	22598 ± 153	7.935 ± 0.061	1.000 ± 0.000	606.69	16	6.16	23298 ± 153	700 ± 9	22598 ± 153	179316 ± 11329
–	4	0.25	0.3	1259751 ± 1122	998253 ± 330	13388 ± 116	340 ± 6	13048 ± 116	13.683 ± 0.107	1.000 ± 0.000	1144.88	16	8.46	13388 ± 116	340 ± 6	13048 ± 116	178533 ± 17874
–	5	0.3	0.35	1259751 ± 1122	998253 ± 330	7932 ± 89	194 ± 5	7738 ± 89	24.451 ± 0.191	1.000 ± 0.000	763.19	16	6.91	7932 ± 89	194 ± 5	7738 ± 89	189187 ± 18180
–	6	0.35	0.4	1259751 ± 1122	998253 ± 330	4688 ± 68	128 ± 4	4560 ± 69	43.177 ± 0.336	1.000 ± 0.000	609.65	16	6.17	4688 ± 68	128 ± 4	4560 ± 69	196898 ± 20573
–	7	0.4	0.45	1259751 ± 1122	998253 ± 330	2998 ± 55	88 ± 3	2910 ± 55	72.586 ± 0.580	1.000 ± 0.000	1076.57	16	8.20	2998 ± 55	88 ± 3	2910 ± 55	211259 ± 35461

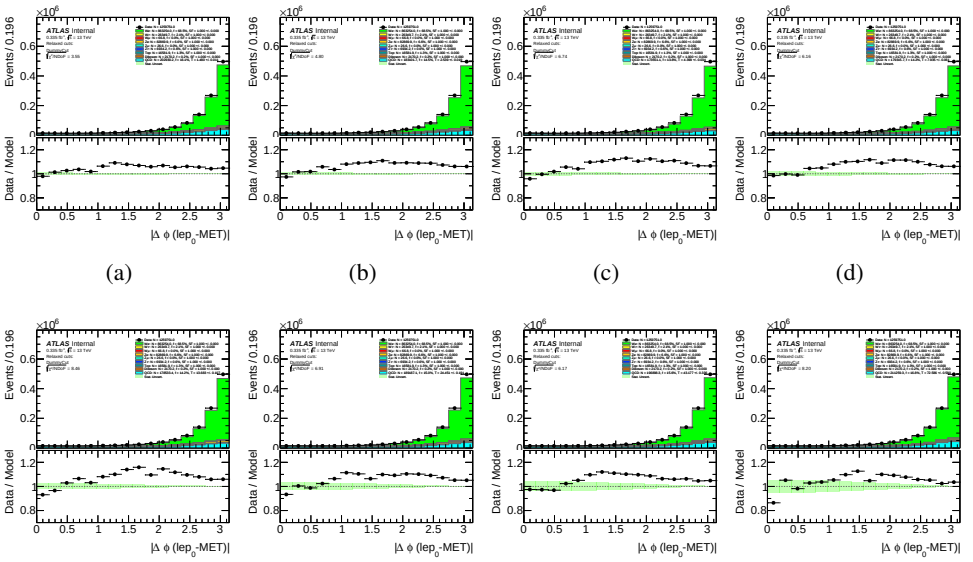
Table 12: Numbers for the electron $|\Delta\phi(lep_0 - MET)|$ fits on 13 TeV. Numbers in FR are not affected by $topoetcone20/p_T$ Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 126: Multi-jet background template fits performed in distributions of m_T^W of the electron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 127: Multi-jet background template fits performed in distributions of $|\Delta\phi(lep_0 - MET)|$ of the electron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

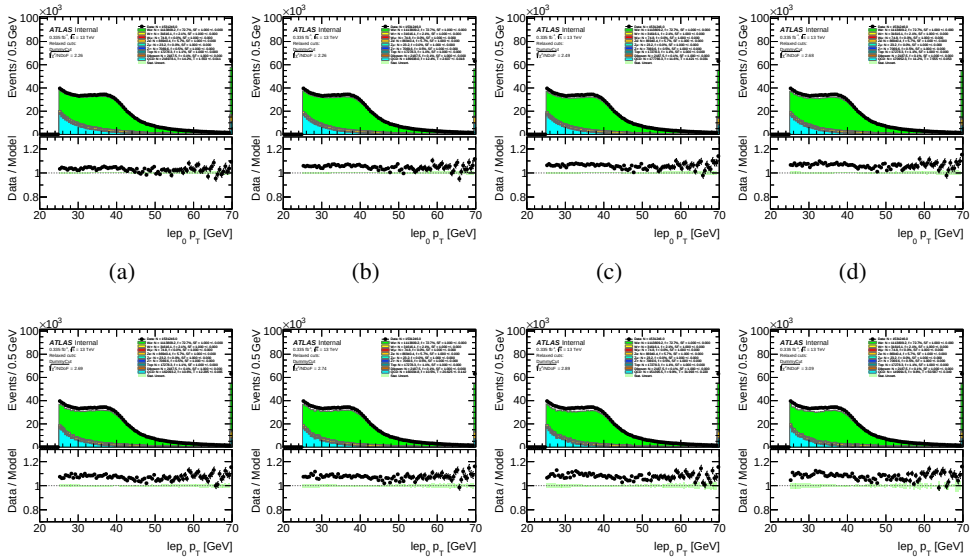
E.3.2 Positron channel multi-jet background template fits

u_T Cut	m_T^W fits b	m_T^W fits k	E_T^{miss} fits b	E_T^{miss} fits k	p_T fits b	p_T fits k	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits b	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits k
–	238724 ± 4593	-9203 ± 29650	248183 ± 3441	62448 ± 25705	219107 ± 3722	-195944 ± 17121	197394 ± 6877	-40250 ± 39215

Table 13: k and b are parameters for the electron extrapolations on 13 TeV. The uncertainty of these two parameters are provided by the extrapolation, from the statistical uncertainty of the parameter T and the number of events in CR of each slices and multiplied by $\sqrt{\chi^2/NDoF}$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1531246 ± 1237	1258999 ± 370	252043 ± 502	112314 ± 111	139729 ± 514	1.553 ± 0.011	1.000 ± 0.000	459.07	90	2.26	252043 ± 502	112314 ± 111	139729 ± 514	216978 ± 3936
–	1	0.1	0.15	1531246 ± 1237	1258999 ± 370	86953 ± 295	14207 ± 40	72746 ± 298	2.607 ± 0.018	1.000 ± 0.000	460.61	90	2.26	86953 ± 295	14207 ± 40	72746 ± 298	189638 ± 3483
–	2	0.15	0.2	1531246 ± 1237	1258999 ± 370	43013 ± 207	2802 ± 18	40211 ± 208	4.421 ± 0.031	1.000 ± 0.000	557.80	90	2.49	43013 ± 207	2802 ± 18	40211 ± 208	177766 ± 3843
–	3	0.2	0.25	1531246 ± 1237	1258999 ± 370	23488 ± 153	859 ± 10	22629 ± 154	7.555 ± 0.053	1.000 ± 0.000	648.62	90	2.68	23488 ± 153	859 ± 10	22629 ± 154	170952 ± 4460
–	4	0.25	0.3	1531246 ± 1237	1258999 ± 370	13584 ± 117	400 ± 6	13184 ± 117	12.295 ± 0.085	1.000 ± 0.000	651.74	90	2.69	13584 ± 117	400 ± 6	13184 ± 117	162093 ± 4903
–	5	0.3	0.35	1531246 ± 1237	1258999 ± 370	7999 ± 89	240 ± 5	7759 ± 90	20.625 ± 0.142	1.000 ± 0.000	675.05	90	2.74	7999 ± 89	240 ± 5	7759 ± 90	160035 ± 5895
–	6	0.35	0.4	1531246 ± 1237	1258999 ± 370	4895 ± 70	163 ± 4	4732 ± 70	31.993 ± 0.220	1.000 ± 0.000	750.66	90	2.89	4895 ± 70	163 ± 4	4732 ± 70	151396 ± 7142
–	7	0.4	0.45	1531246 ± 1237	1258999 ± 370	3073 ± 55	108 ± 4	2965 ± 56	50.587 ± 0.349	1.000 ± 0.000	858.89	90	3.09	3073 ± 55	108 ± 4	2965 ± 56	149991 ± 9252

Table 14: Numbers for the positron p_T fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



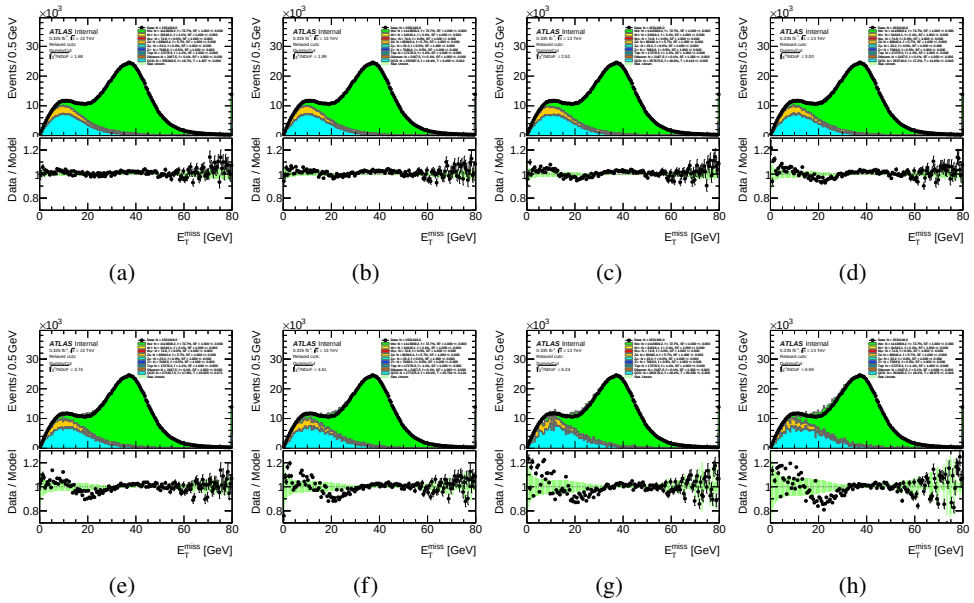
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 128: Multi-jet background template fits performed in distributions of p_T^ℓ of the positron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

w_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1531246 ± 1237	1258999 ± 370	252043 ± 502	112314 ± 111	139729 ± 514	1.827 ± 0.006	1.000 ± 0.000	450.59	160	1.68	252043 ± 502	112314 ± 111	139729 ± 514	255360 ± 2161
–	1	0.1	0.15	1531246 ± 1237	1258999 ± 370	86953 ± 295	14207 ± 40	72746 ± 298	3.459 ± 0.012	1.000 ± 0.000	635.87	160	1.99	86953 ± 295	14207 ± 40	72746 ± 298	251597 ± 2680
–	2	0.15	0.2	1531246 ± 1237	1258999 ± 370	43013 ± 207	2802 ± 18	40211 ± 208	6.413 ± 0.022	1.000 ± 0.000	1013.08	160	2.52	43013 ± 207	2802 ± 18	40211 ± 208	257874 ± 4034
–	3	0.2	0.25	1531246 ± 1237	1258999 ± 370	23488 ± 153	859 ± 10	22629 ± 154	11.654 ± 0.040	1.000 ± 0.000	1436.35	160	3.00	23488 ± 153	859 ± 10	22629 ± 154	263716 ± 6016
–	4	0.25	0.3	1531246 ± 1237	1258999 ± 370	13584 ± 117	400 ± 6	13184 ± 117	20.625 ± 0.071	1.000 ± 0.000	2233.17	160	3.74	13584 ± 117	400 ± 6	13184 ± 117	271917 ± 9657
–	5	0.3	0.35	1531246 ± 1237	1258999 ± 370	7999 ± 89	240 ± 5	7759 ± 90	35.734 ± 0.124	1.000 ± 0.000	3400.49	160	4.61	7999 ± 89	240 ± 5	7759 ± 90	277275 ± 15411
–	6	0.35	0.4	1531246 ± 1237	1258999 ± 370	4895 ± 70	163 ± 4	4732 ± 70	59.630 ± 0.209	1.000 ± 0.000	6224.33	160	6.24	4895 ± 70	163 ± 4	4732 ± 70	282178 ± 26792
–	7	0.4	0.45	1531246 ± 1237	1258999 ± 370	3073 ± 55	108 ± 4	2965 ± 56	98.975 ± 0.349	1.000 ± 0.000	7794.24	160	6.98	3073 ± 55	108 ± 4	2965 ± 56	293465 ± 39049

Table 15: Numbers for the positron E_T^{miss} fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

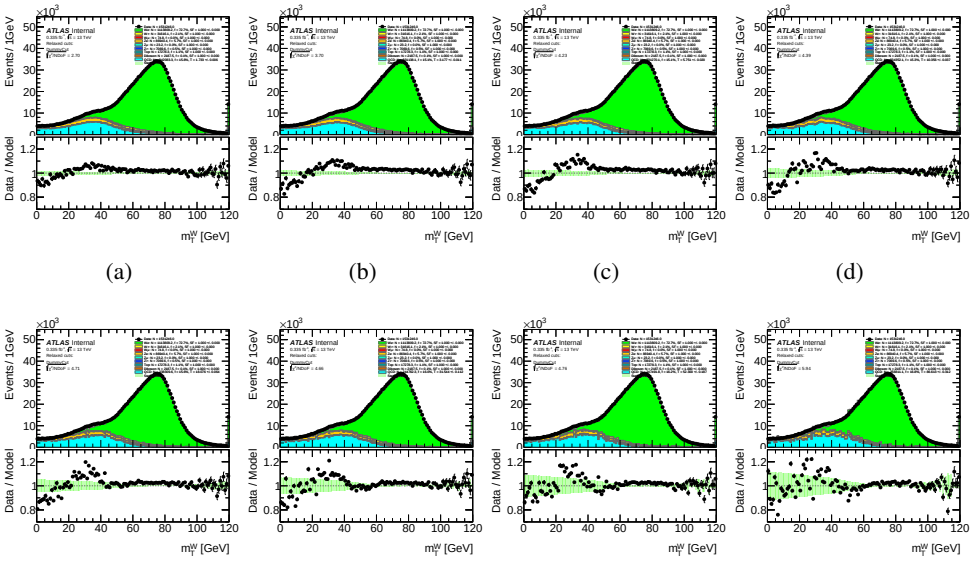
Figure 129: Multi-jet background template fits performed in distributions of E_T^{miss} of the positron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

w_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1531246 ± 1237	1258999 ± 370	252043 ± 502	112314 ± 111	139729 ± 514	1.733 ± 0.006	1.000 ± 0.000	873.20	120	2.70	252043 ± 502	112314 ± 111	139729 ± 514	242083 ± 3347
–	1	0.1	0.15	1531246 ± 1237	1258999 ± 370	86953 ± 295	14207 ± 40	72746 ± 298	3.177 ± 0.011	1.000 ± 0.000	1643.78	120	3.70	86953 ± 295	14207 ± 40	72746 ± 298	231138 ± 4636
–	2	0.15	0.2	1531246 ± 1237	1258999 ± 370	43013 ± 207	2802 ± 18	40211 ± 208	5.751 ± 0.020	1.000 ± 0.000	2144.13	120	4.23	43013 ± 207	2802 ± 18	40211 ± 208	231270 ± 6141
–	3	0.2	0.25	1531246 ± 1237	1258999 ± 370	23488 ± 153	859 ± 10	22629 ± 154	10.356 ± 0.037	1.000 ± 0.000	2312.03	120	4.39	23488 ± 153	859 ± 10	22629 ± 154	234352 ± 7884
–	4	0.25	0.3	1531246 ± 1237	1258999 ± 370	13584 ± 117	400 ± 6	13184 ± 117	18.076 ± 0.064	1.000 ± 0.000	2661.53	120	4.71	13584 ± 117	400 ± 6	13184 ± 117	238310 ± 10711
–	5	0.3	0.35	1531246 ± 1237	1258999 ± 370	7999 ± 89	240 ± 5	7759 ± 90	31.544 ± 0.112	1.000 ± 0.000	2607.77	120	4.66	7999 ± 89	240 ± 5	7759 ± 90	244762 ± 13785
–	6	0.35	0.4	1531246 ± 1237	1258999 ± 370	4895 ± 70	163 ± 4	4732 ± 70	52.300 ± 0.187	1.000 ± 0.000	2722.92	120	4.76	4895 ± 70	163 ± 4	4732 ± 70	247492 ± 17962
–	7	0.4	0.45	1531246 ± 1237	1258999 ± 370	3073 ± 55	108 ± 4	2965 ± 56	86.613 ± 0.312	1.000 ± 0.000	4234.75	120	5.94	3073 ± 55	108 ± 4	2965 ± 56	256811 ± 29106

Table 16: Numbers for the positron m_T^W fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

w_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.05	0.1	1531246 ± 1237	1258999 ± 370	252043 ± 502	112314 ± 111	139729 ± 514	1.458 ± 0.012	1.000 ± 0.000	154.79	16	3.11	252043 ± 502	112314 ± 111	139729 ± 514	203775 ± 5534
–	1	0.1	0.15	1531246 ± 1237	1258999 ± 370	86953 ± 295	14207 ± 40	72746 ± 298	2.509 ± 0.020	1.000 ± 0.000	288.61	16	4.25	86953 ± 295	14207 ± 40	72746 ± 298	182513 ± 6880
–	2	0.15	0.2	1531246 ± 1237	1258999 ± 370	43013 ± 207	2802 ± 18	40211 ± 208	4.456 ± 0.035	1.000 ± 0.000	446.29	16	5.28	43013 ± 207	2802 ± 18	40211 ± 208	179172 ± 8988
–	3	0.2	0.25	1531246 ± 1237	1258999 ± 370	23488 ± 153	859 ± 10	22629 ± 154	8.002 ± 0.063	1.000 ± 0.000	385.75	16	4.91	23488 ± 153	859 ± 10	22629 ± 154	181087 ± 9281
–	4	0.25	0.3	1531246 ± 1237	1258999 ± 370	13584 ± 117	400 ± 6	13184 ± 117	13.973 ± 0.111	1.000 ± 0.000	564.63	16	5.94	13584 ± 117	400 ± 6	13184 ± 117	184215 ± 13015
–	5	0.3	0.35	1531246 ± 1237	1258999 ± 370	7999 ± 89	240 ± 5	7759 ± 90	25.821 ± 0.207	1.000 ± 0.000	579.23	16	6.02	7999 ± 89	240 ± 5	7759 ± 90	200360 ± 16956
–	6	0.35	0.4	1531246 ± 1237	1258999 ± 370	4895 ± 70	163 ± 4	4732 ± 70	43.248 ± 0.348	1.000 ± 0.000	551.20	16	5.87	4895 ± 70	163 ± 4	4732 ± 70	204657 ± 20249
–	7	0.4	0.45	1531246 ± 1237	1258999 ± 370	3073 ± 55	108 ± 4	2965 ± 56	71.933 ± 0.589	1.000 ± 0.000	1073.74	16	8.19	3073 ± 55	108 ± 4	2965 ± 56	213284 ± 35728

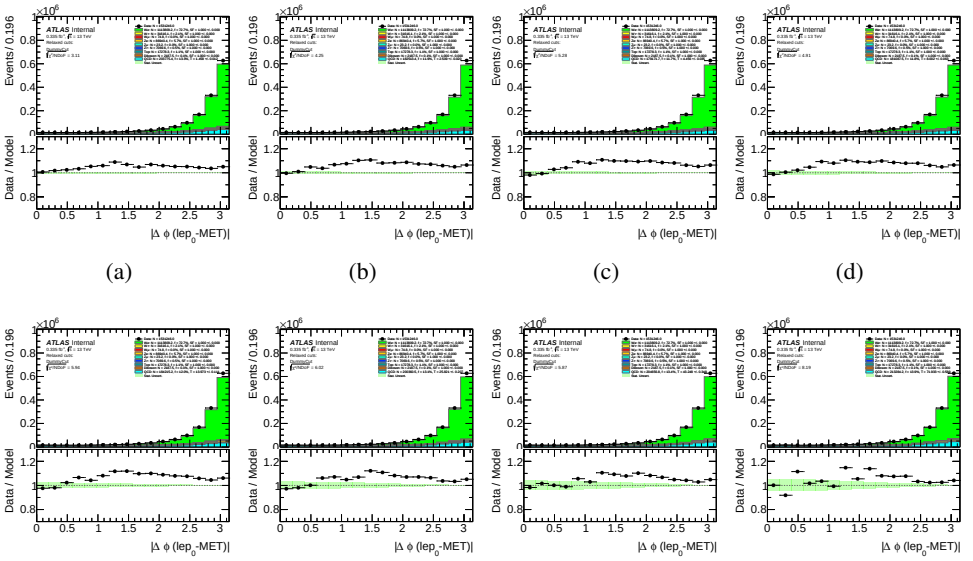
Table 17: Numbers for the positron $|\Delta\phi(lep_0 - MET)|$ fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 130: Multi-jet background template fits performed in distributions of m_T^W of the positron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 131: Multi-jet background template fits performed in distributions of $|\Delta\phi(lep_0 - MET)|$ of the positron channel, columns from (a) to (h) are the fits with the isolation slices from $0.05 < topoetcone20/p_T < 0.1$ to $0.4 < topoetcone20/p_T < 0.45$.

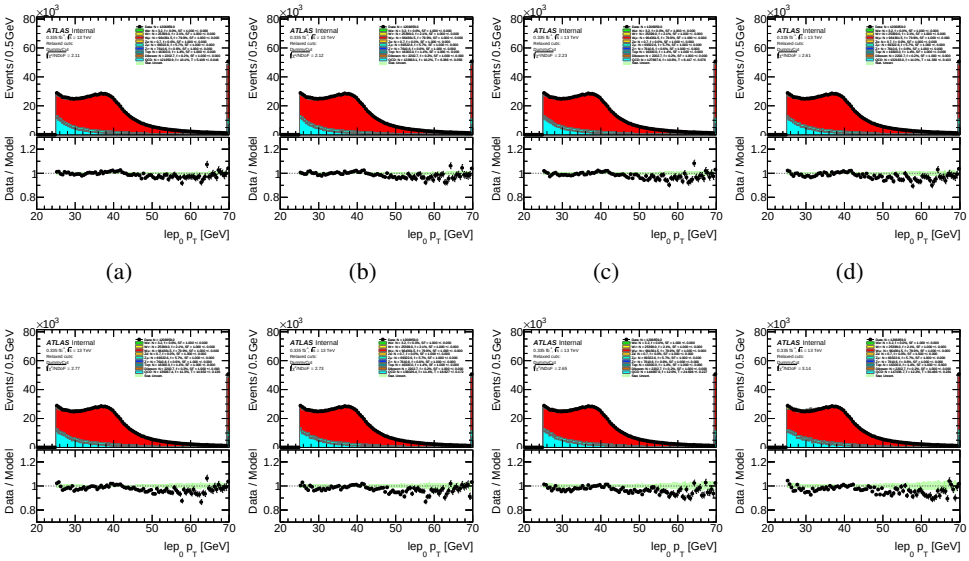
E.3.3 Muons channel multi-jet background template fits

u_T Cut	m_T^W fits b	m_T^W fits k	E_T^{miss} fits b	E_T^{miss} fits k	p_T fits b	p_T fits k	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits b	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits k
–	129929 ± 4588	70936 ± 21952	127983 ± 4236	94904 ± 20964	110773 ± 3654	76260 ± 14547	134318 ± 8297	48500 ± 34133

Table 18: k and b are parameters for the electron extrapolations on 13 TeV. The uncertainty of these two parameters are provided by the extrapolation, from the statistical uncertainty of the parameter T and the number of events in CR of each slices and multiplied by $\sqrt{\chi^2/NDoF}$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	23759 ± 160	5.108 ± 0.046	1.000 ± 0.000	400.50	90	2.11	25514 ± 160	1755 ± 14	23759 ± 160	121351 ± 2878
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	19209 ± 141	6.386 ± 0.058	1.000 ± 0.000	403.43	90	2.12	19741 ± 141	532 ± 7	19209 ± 141	122663 ± 3017
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	15194 ± 124	8.417 ± 0.076	1.000 ± 0.000	446.43	90	2.23	15432 ± 124	238 ± 5	15194 ± 124	127888 ± 3472
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	11665 ± 109	11.330 ± 0.103	1.000 ± 0.000	614.52	90	2.61	11810 ± 109	145 ± 5	11665 ± 109	132163 ± 4496
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9113 ± 96	14.932 ± 0.136	1.000 ± 0.000	688.79	90	2.77	9202 ± 96	89 ± 3	9113 ± 96	136067 ± 5246
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	7331 ± 86	18.827 ± 0.172	1.000 ± 0.000	672.26	90	2.73	7395 ± 86	64 ± 2	7331 ± 86	138026 ± 5612
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5859 ± 77	24.696 ± 0.227	1.000 ± 0.000	632.30	90	2.65	5904 ± 77	45 ± 2	5859 ± 77	144688 ± 6140
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4823 ± 70	30.486 ± 0.281	1.000 ± 0.000	889.94	90	3.14	4854 ± 70	31 ± 2	4823 ± 70	147037 ± 7924

Table 19: Numbers for the muon p_T fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 132: Multi-jet background template fits performed in distributions of p_T^ℓ of the muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	23759 ± 160	5.901 ± 0.028	1.000 ± 0.000	729.82	160	2.14	25514 ± 160	1755 ± 14	23759 ± 160	140202 ± 2468
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	19209 ± 141	7.490 ± 0.035	1.000 ± 0.000	1126.82	160	2.65	19741 ± 141	532 ± 7	19209 ± 141	143877 ± 3329
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	15194 ± 124	9.784 ± 0.046	1.000 ± 0.000	1618.14	160	3.18	15432 ± 124	238 ± 5	15194 ± 124	148658 ± 4470
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	11665 ± 109	13.266 ± 0.063	1.000 ± 0.000	2197.39	160	3.71	11810 ± 109	145 ± 5	11665 ± 109	154749 ± 6003
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9113 ± 96	17.562 ± 0.084	1.000 ± 0.000	3243.40	160	4.50	9202 ± 96	89 ± 3	9113 ± 96	160039 ± 8333
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	7331 ± 86	22.377 ± 0.107	1.000 ± 0.000	3929.58	160	4.96	7395 ± 86	64 ± 2	7331 ± 86	164057 ± 10306
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5859 ± 77	29.019 ± 0.140	1.000 ± 0.000	4332.88	160	5.20	5904 ± 77	45 ± 2	5859 ± 77	170017 ± 12364
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4823 ± 70	35.327 ± 0.171	1.000 ± 0.000	5368.57	160	5.79	4854 ± 70	31 ± 2	4823 ± 70	170386 ± 15037

Table 20: Numbers for the muon E_T^{miss} fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

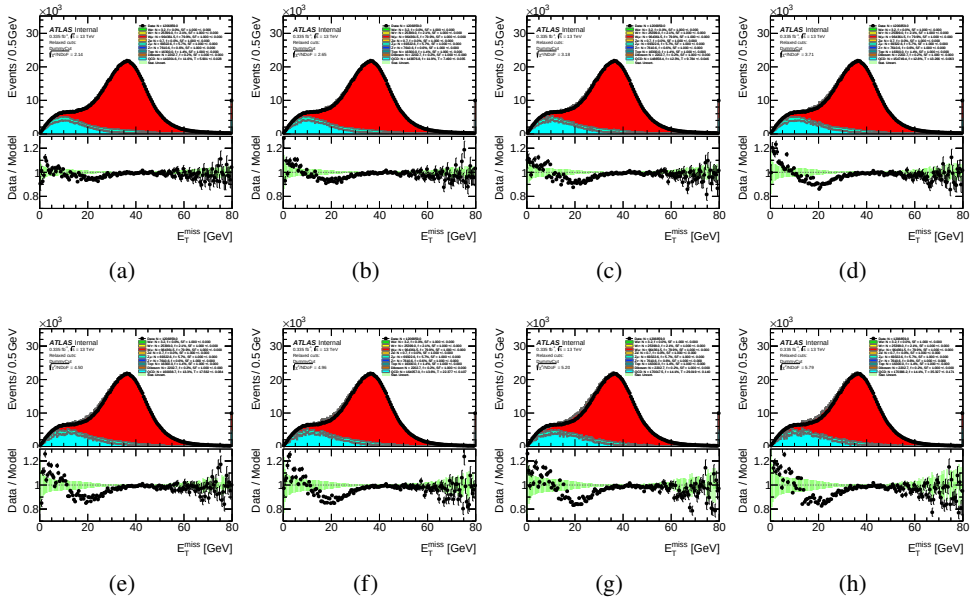


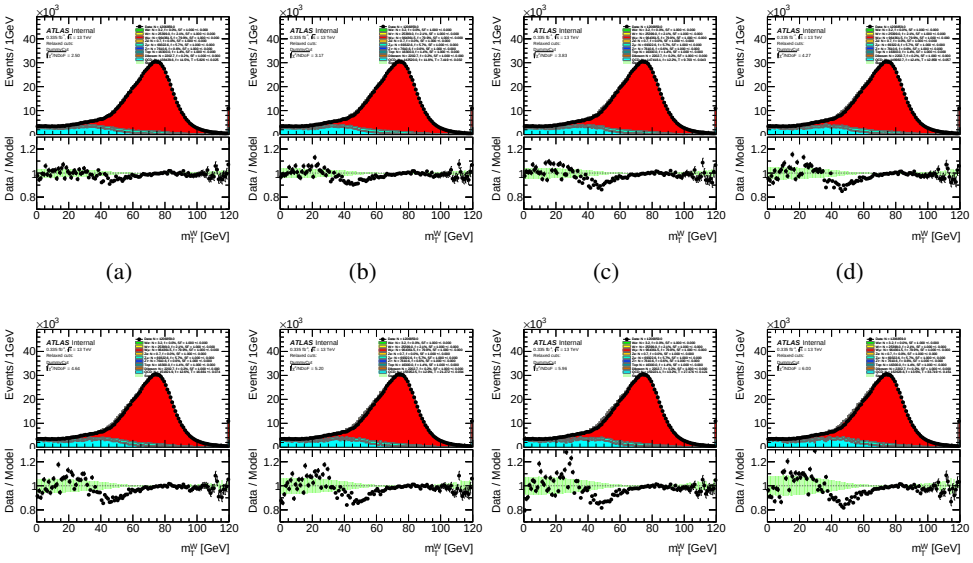
Figure 133: Multi-jet background template fits performed in distributions of E_T^{miss} of the muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	23759 ± 160	5.826 ± 0.025	1.000 ± 0.000	749.26	120	2.50	25514 ± 160	1755 ± 14	23759 ± 160	138430 ± 2782
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	19209 ± 141	7.419 ± 0.032	1.000 ± 0.000	1203.39	120	3.17	19741 ± 141	532 ± 7	19209 ± 141	142520 ± 3851
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	15194 ± 124	9.703 ± 0.043	1.000 ± 0.000	1756.70	120	3.83	15432 ± 124	238 ± 5	15194 ± 124	147418 ± 5236
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	11665 ± 109	12.858 ± 0.057	1.000 ± 0.000	2183.20	120	4.27	11810 ± 109	145 ± 5	11665 ± 109	149983 ± 6598
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9113 ± 96	16.661 ± 0.074	1.000 ± 0.000	2585.71	120	4.64	9202 ± 96	89 ± 3	9113 ± 96	151822 ± 8048
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	7331 ± 86	21.272 ± 0.094	1.000 ± 0.000	3250.81	120	5.20	7395 ± 86	64 ± 2	7331 ± 86	155952 ± 10182
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5859 ± 77	27.176 ± 0.121	1.000 ± 0.000	4258.66	120	5.96	5904 ± 77	45 ± 2	5859 ± 77	159221 ± 13145
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4823 ± 70	33.719 ± 0.151	1.000 ± 0.000	4321.83	120	6.00	4854 ± 70	31 ± 2	4823 ± 70	162629 ± 14761

Table 21: Numbers for the muon m_T^W fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	23759 ± 160	5.783 ± 0.042	1.000 ± 0.000	327.36	16	4.52	25514 ± 160	1755 ± 14	23759 ± 160	137404 ± 6145
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	19209 ± 141	7.538 ± 0.055	1.000 ± 0.000	343.68	16	4.63	19741 ± 141	532 ± 7	19209 ± 141	144801 ± 6950
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	15194 ± 124	9.972 ± 0.073	1.000 ± 0.000	468.49	16	5.41	15432 ± 124	238 ± 5	15194 ± 124	151509 ± 9026
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	11665 ± 109	12.727 ± 0.095	1.000 ± 0.000	552.69	16	5.88	11810 ± 109	145 ± 5	11665 ± 109	148454 ± 10410
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9113 ± 96	16.030 ± 0.120	1.000 ± 0.000	697.56	16	6.60	9202 ± 96	89 ± 3	9113 ± 96	146073 ± 12462
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	7331 ± 86	20.499 ± 0.154	1.000 ± 0.000	673.49	16	6.49	7395 ± 86	64 ± 2	7331 ± 86	150283 ± 13578
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5859 ± 77	26.069 ± 0.196	1.000 ± 0.000	715.69	16	6.69	5904 ± 77	45 ± 2	5859 ± 77	152737 ± 15451
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4823 ± 70	32.638 ± 0.248	1.000 ± 0.000	914.65	16	7.56	4854 ± 70	31 ± 2	4823 ± 70	157416 ± 19434

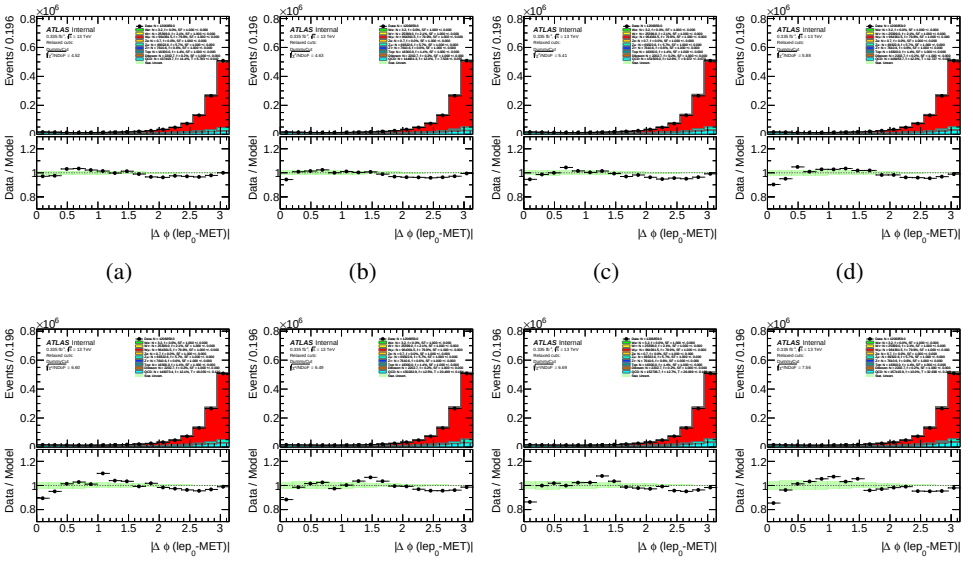
Table 22: Numbers for the muon $|\Delta\phi(lep_0 - MET)|$ fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 134: Multi-jet background template fits performed in distributions of m_T^W of the muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 135: Multi-jet background template fits performed in distributions of $|\Delta\phi(lep_0 - MET)|$ of the muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

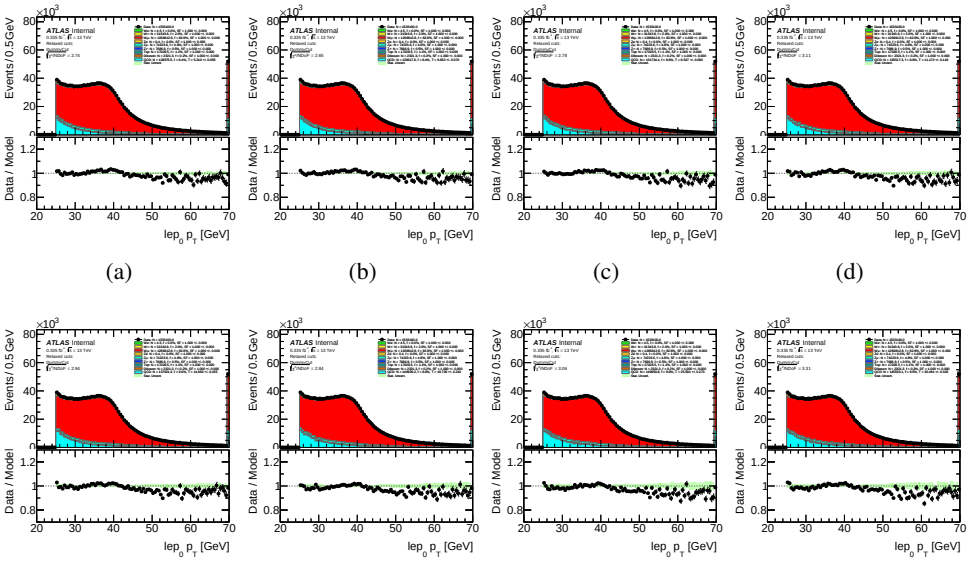
E.3.4 Antimuon channel multi-jet background template fits

u_T Cut	m_T^W fits b	m_T^W fits k	E_T^{miss} fits b	E_T^{miss} fits k	p_T fits b	p_T fits k	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits b	$\lvert\Delta\phi(\ell, E_T^{miss})\rvert$ fits k
–	135147 ± 4779	66271 ± 22086	131684 ± 4622	94214 ± 21894	118021 ± 5091	68043 ± 18764	141309 ± 12566	27667 ± 47305

Table 23: k and b are parameters for the electron extrapolations on 13 TeV. The uncertainty of these two parameters are provided by the extrapolation, from the statistical uncertainty of the parameter T and the number of events in CR of each slices and multiplied by $\sqrt{\chi^2/NDoF}$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1533100 ± 1238	1402422 ± 387	27011 ± 164	2334 ± 15	24677 ± 165	5.210 ± 0.055	1.000 ± 0.000	677.72	90	2.74	27011 ± 164	2334 ± 15	24677 ± 165	128575 ± 4387
–	1	0.15	0.2	1533100 ± 1238	1402422 ± 387	20028 ± 142	694 ± 9	19334 ± 142	6.652 ± 0.070	1.000 ± 0.000	631.34	90	2.65	20028 ± 142	694 ± 9	19334 ± 142	128618 ± 4356
–	2	0.2	0.25	1533100 ± 1238	1402422 ± 387	15718 ± 125	305 ± 7	15413 ± 126	8.547 ± 0.090	1.000 ± 0.000	697.99	90	2.78	15718 ± 125	305 ± 7	15413 ± 126	131734 ± 4883
–	3	0.25	0.3	1533100 ± 1238	1402422 ± 387	12307 ± 111	177 ± 4	12130 ± 111	11.172 ± 0.118	1.000 ± 0.000	870.88	90	3.11	12307 ± 111	177 ± 4	12130 ± 111	135517 ± 5904
–	4	0.3	0.35	1533100 ± 1238	1402422 ± 387	9514 ± 98	115 ± 4	9399 ± 98	14.663 ± 0.155	1.000 ± 0.000	777.06	90	2.94	9514 ± 98	115 ± 4	9399 ± 98	137821 ± 6006
–	5	0.35	0.4	1533100 ± 1238	1402422 ± 387	7668 ± 88	71 ± 3	7597 ± 88	19.736 ± 0.210	1.000 ± 0.000	723.91	90	2.84	7668 ± 88	71 ± 3	7597 ± 88	149936 ± 6671
–	6	0.4	0.45	1533100 ± 1238	1402422 ± 387	5912 ± 77	54 ± 3	5858 ± 77	25.581 ± 0.273	1.000 ± 0.000	840.40	90	3.06	5912 ± 77	54 ± 3	5858 ± 77	149851 ± 7747
–	7	0.45	0.5	1533100 ± 1238	1402422 ± 387	4811 ± 69	38 ± 2	4773 ± 69	30.491 ± 0.326	1.000 ± 0.000	986.12	90	3.31	4811 ± 69	38 ± 2	4773 ± 69	145533 ± 8691

Table 24: Numbers for the antimuon p_T fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 136: Multi-jet background template fits performed in distributions of p_T^ℓ of the anti-muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1533100 ± 1238	1402422 ± 387	27011 ± 164	2334 ± 15	24677 ± 165	5.817 ± 0.028	1.000 ± 0.000	927.36	160	2.41	27011 ± 164	2334 ± 15	24677 ± 165	143548 ± 2839
–	1	0.15	0.2	1533100 ± 1238	1402422 ± 387	20028 ± 142	694 ± 9	19334 ± 142	7.652 ± 0.036	1.000 ± 0.000	1452.56	160	3.01	20028 ± 142	694 ± 9	19334 ± 142	147944 ± 3900
–	2	0.2	0.25	1533100 ± 1238	1402422 ± 387	15718 ± 125	305 ± 7	15413 ± 126	9.956 ± 0.048	1.000 ± 0.000	1841.33	160	3.39	15718 ± 125	305 ± 7	15413 ± 126	153457 ± 4919
–	3	0.25	0.3	1533100 ± 1238	1402422 ± 387	12307 ± 111	177 ± 4	12130 ± 111	12.946 ± 0.062	1.000 ± 0.000	2611.75	160	4.04	12307 ± 111	177 ± 4	12130 ± 111	157037 ± 6555
–	4	0.3	0.35	1533100 ± 1238	1402422 ± 387	9514 ± 98	115 ± 4	9399 ± 98	17.071 ± 0.082	1.000 ± 0.000	3098.87	160	4.40	9514 ± 98	115 ± 4	9399 ± 98	160457 ± 8081
–	5	0.35	0.4	1533100 ± 1238	1402422 ± 387	7668 ± 88	71 ± 3	7597 ± 88	22.230 ± 0.107	1.000 ± 0.000	4114.85	160	5.07	7668 ± 88	71 ± 3	7597 ± 88	168885 ± 10708
–	6	0.4	0.45	1533100 ± 1238	1402422 ± 387	5912 ± 77	54 ± 3	5858 ± 77	29.292 ± 0.142	1.000 ± 0.000	4418.26	160	5.25	5912 ± 77	54 ± 3	5858 ± 77	171589 ± 12624
–	7	0.45	0.5	1533100 ± 1238	1402422 ± 387	4811 ± 69	38 ± 2	4773 ± 69	37.221 ± 0.181	1.000 ± 0.000	4297.88	160	5.18	4811 ± 69	38 ± 2	4773 ± 69	177656 ± 14117

Table 25: Numbers for the antimuon E_T^{miss} fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

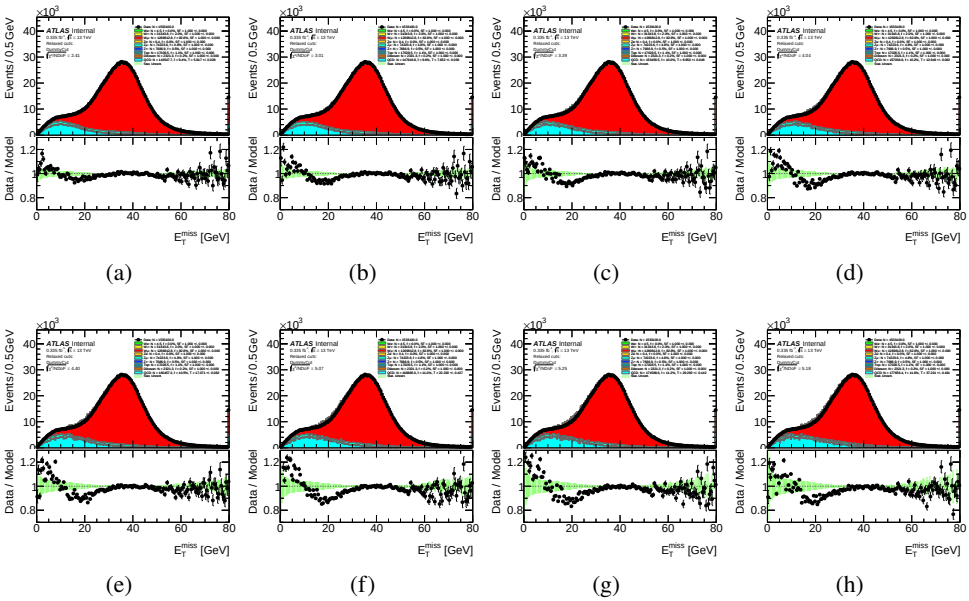


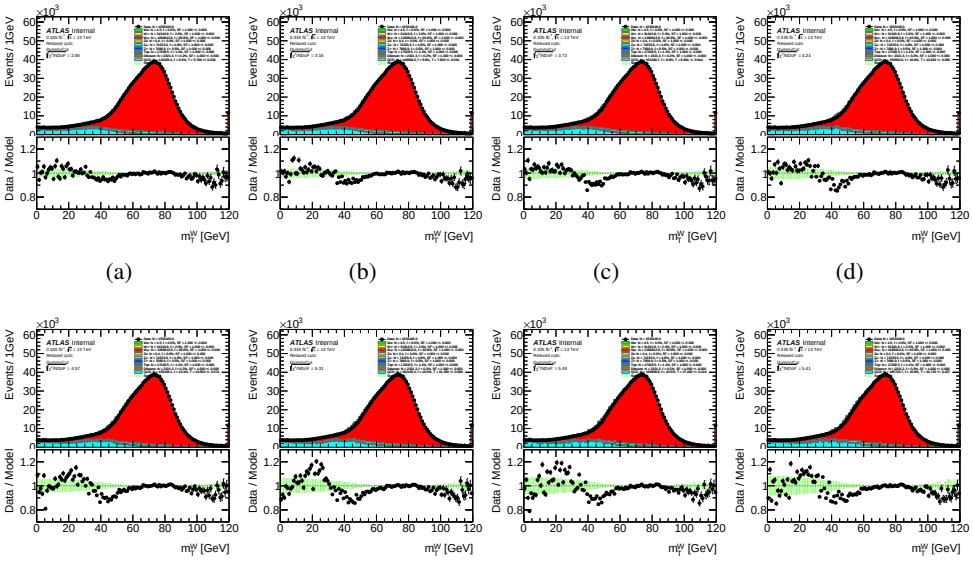
Figure 137: Multi-jet background template fits performed in distributions of E_T^{miss} of the anti-muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1533100 ± 1238	1402422 ± 387	27011 ± 164	2334 ± 15	24677 ± 165	5.796 ± 0.026	1.000 ± 0.000	843.98	120	2.65	27011 ± 164	2334 ± 15	24677 ± 165	143025 ± 3045
–	1	0.15	0.2	1533100 ± 1238	1402422 ± 387	20028 ± 142	694 ± 9	19334 ± 142	7.603 ± 0.034	1.000 ± 0.000	1201.45	120	3.16	20028 ± 142	694 ± 9	19334 ± 142	146996 ± 3987
–	2	0.2	0.25	1533100 ± 1238	1402422 ± 387	15718 ± 125	305 ± 7	15413 ± 126	9.821 ± 0.044	1.000 ± 0.000	1662.31	120	3.72	15718 ± 125	305 ± 7	15413 ± 126	151381 ± 5229
–	3	0.25	0.3	1533100 ± 1238	1402422 ± 387	12307 ± 111	177 ± 4	12130 ± 111	12.662 ± 0.056	1.000 ± 0.000	2153.46	120	4.24	12307 ± 111	177 ± 4	12130 ± 111	153591 ± 6623
–	4	0.3	0.35	1533100 ± 1238	1402422 ± 387	9514 ± 98	115 ± 4	9399 ± 98	16.502 ± 0.074	1.000 ± 0.000	2505.03	120	4.57	9514 ± 98	115 ± 4	9399 ± 98	155108 ± 8015
–	5	0.35	0.4	1533100 ± 1238	1402422 ± 387	7668 ± 88	71 ± 3	7597 ± 88	21.350 ± 0.096	1.000 ± 0.000	3379.07	120	5.31	7668 ± 88	71 ± 3	7597 ± 88	162195 ± 10656
–	6	0.4	0.45	1533100 ± 1238	1402422 ± 387	5912 ± 77	54 ± 3	5858 ± 77	27.482 ± 0.124	1.000 ± 0.000	3619.58	120	5.49	5912 ± 77	54 ± 3	5858 ± 77	160990 ± 12283
–	7	0.45	0.5	1533100 ± 1238	1402422 ± 387	4811 ± 69	38 ± 2	4773 ± 69	34.722 ± 0.157	1.000 ± 0.000	3514.50	120	5.41	4811 ± 69	38 ± 2	4773 ± 69	165729 ± 13656

Table 26: Numbers for the antimuon m_T^W fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1533100 ± 1238	1402422 ± 387	27011 ± 164	2334 ± 15	24677 ± 165	5.712 ± 0.042	1.000 ± 0.000	852.00	16	7.30	27011 ± 164	2334 ± 15	24677 ± 165	140967 ± 10245
–	1	0.15	0.2	1533100 ± 1238	1402422 ± 387	20028 ± 142	694 ± 9	19334 ± 142	7.594 ± 0.056	1.000 ± 0.000	870.33	16	7.38	20028 ± 142	694 ± 9	19334 ± 142	146815 ± 11306
–	2	0.2	0.25	1533100 ± 1238	1402422 ± 387	15718 ± 125	305 ± 7	15413 ± 126	9.969 ± 0.074	1.000 ± 0.000	885.25	16	7.44	15718 ± 125	305 ± 7	15413 ± 126	153658 ± 12633
–	3	0.25	0.3	1533100 ± 1238	1402422 ± 387	12307 ± 111	177 ± 4	12130 ± 111	12.589 ± 0.095	1.000 ± 0.000	1066.86	16	8.17	12307 ± 111	177 ± 4	12130 ± 111	152700 ± 14770
–	4	0.3	0.35	1533100 ± 1238	1402422 ± 387	9514 ± 98	115 ± 4	9399 ± 98	15.748 ± 0.119	1.000 ± 0.000	1091.83	16	8.26	9514 ± 98	115 ± 4	9399 ± 98	148021 ± 15713
–	5	0.35	0.4	1533100 ± 1238	1402422 ± 387	7668 ± 88	71 ± 3	7597 ± 88	20.094 ± 0.154	1.000 ± 0.000	1381.08	16	9.29	7668 ± 88	71 ± 3	7597 ± 88	152659 ± 19641
–	6	0.4	0.45	1533100 ± 1238	1402422 ± 387	5912 ± 77	54 ± 3	5858 ± 77	25.134 ± 0.192	1.000 ± 0.000	1141.79	16	8.45	5912 ± 77	54 ± 3	5858 ± 77	147234 ± 18907
–	7	0.45	0.5	1533100 ± 1238	1402422 ± 387	4811 ± 69	38 ± 2	4773 ± 69	32.181 ± 0.246	1.000 ± 0.000	1272.65	16	8.92	4811 ± 69	38 ± 2	4773 ± 69	153601 ± 22506

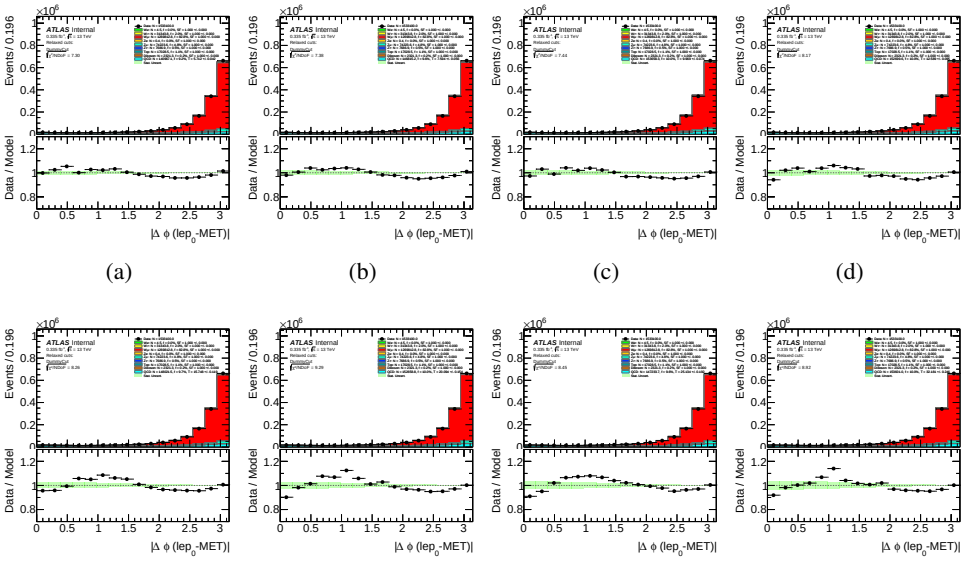
Table 27: Numbers for the antimuon $|\Delta\phi(lep_0 - MET)|$ fits on 13 TeV. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 138: Multi-jet background template fits performed in distributions of m_T^W of the anti-muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 139: Multi-jet background template fits performed in distributions of $|\Delta\phi(lep_0 - MET)|$ of the anti-muon channel, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.1$ to $0.4 < ptvarcone20/p_T < 0.5$.

E.4 Closure tests for data-driven MJ estimation algorithm

To make sure anti-isolated slices extrapolation method works properly number of closure tests has been performed. We used $b\bar{b} + c\bar{c}$ MC samples for $W^- \rightarrow \mu^-\overline{\nu}$ and $W^+ \rightarrow \mu^+\nu$ channels.

The properties of the $b\bar{b}$ and $c\bar{c}$ samples at $\sqrt{s}$ = 13 TeV are given in Table 28. The heavy-flavour production cross sections are higher than signal and other backgrounds, but their calculation carries significant uncertainties from the dijet production cross section (calculated at leading order), and the $b/c \rightarrow \mu$ decay fractions and fragmentation functions.

Process	Data set	Generator	σ·BR·ϵ_{filter} [nb]
$b\bar{b} \rightarrow \mu$	361250	Pythia8B+A14+NNPDF23LO	186
$c\bar{c} \rightarrow \mu$	361251	Pythia8B+A14+NNPDF23LO	58

Table 28: MC simulation of $b\bar{b}$ and $c\bar{c}$ decay to muon at $\sqrt{s}$ = 13 TeV.

This MC samples were used to derive MJ templates in the CR1 (CR2) anti-isolated slices. Linear extrapolations from anti-isolated slices are shown on Fig. 140 and listed in the Tables. 29. It is clear to see impact from limited MC samples statistics to the anti-isolated templates. Results are compatible with liner extrapolations shown at Fig. 13 that are derived by data-driven method.

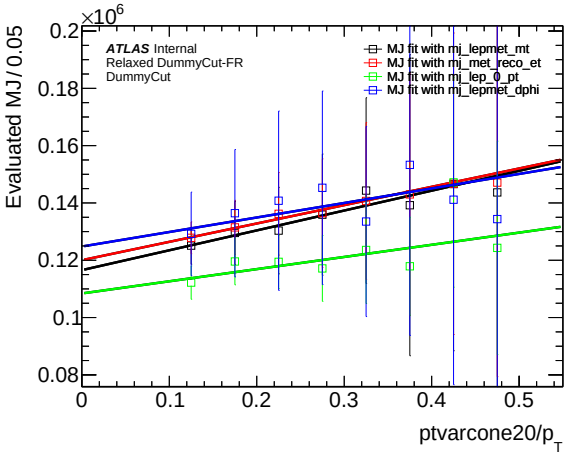


Figure 140: Linear extrapolations from anti-isolated slices for $W^- \rightarrow \mu^-\overline{\nu}$ channels using $b\bar{b} + c\bar{c}$ MC samples as a source of templates in anti-isolated region.

u_T Cut	m_T^W fits b	m_T^W fits k	E_T^{miss} fits b	E_T^{miss} fits k	p_T fits b	p_T fits k	$\lvert\Delta\phi(lep_0 - MET)\rvert$ fits b	$\lvert\Delta\phi(lep_0 - MET)\rvert$ fits k
–	116548 ± 13931	69343 ± 78742	120025 ± 12112	64022 ± 69345	108412 ± 9743	42472 ± 46358	124791 ± 24369	50671 ± 115205

Table 29: k and b are parameters for the muon extrapolations on 13 TeV using $b\bar{b} + c\bar{c}$ MC sample as a source of templates in anti-isolated region. The uncertainty of these two parameters are provided by the extrapolation, from the statistical uncertainty of the parameter T and the number of events in $CR2$ of each slices and multiplied by $\sqrt{\chi^2/NDoF}$.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	29180 ± 592	3.845 ± 0.031	1.000 ± 0.000	503.42	90	2.37	25514 ± 160	1755 ± 14	29180 ± 592	112200 ± 5807
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	23045 ± 524	5.189 ± 0.042	1.000 ± 0.000	714.02	90	2.82	19741 ± 141	532 ± 7	23045 ± 524	119573 ± 8146
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	17122 ± 451	6.976 ± 0.057	1.000 ± 0.000	723.30	90	2.83	15432 ± 124	238 ± 5	17122 ± 451	119447 ± 9343
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	13481 ± 400	8.690 ± 0.072	1.000 ± 0.000	910.74	90	3.18	11810 ± 109	145 ± 5	13481 ± 400	117147 ± 11489
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9429 ± 324	13.108 ± 0.110	1.000 ± 0.000	1644.71	90	4.27	9202 ± 96	89 ± 3	9429 ± 324	123587 ± 18688
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	6993 ± 277	16.859 ± 0.139	1.000 ± 0.000	1153.85	90	3.58	7395 ± 86	64 ± 2	6993 ± 277	117897 ± 17101
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5174 ± 235	28.443 ± 0.242	1.000 ± 0.000	2606.78	90	5.38	5904 ± 77	45 ± 2	5174 ± 235	147169 ± 36579
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4409 ± 212	28.196 ± 0.247	1.000 ± 0.000	3394.43	90	6.14	4854 ± 70	31 ± 2	4409 ± 212	124324 ± 37238

Table 30: Numbers for the muon p_T fits on 13 TeV using $b\bar{b} + c\bar{c}$ MC samples. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	29180 ± 592	4.381 ± 0.027	1.000 ± 0.000	672.90	160	2.05	25514 ± 160	1755 ± 14	29180 ± 592	127834 ± 5556
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	23045 ± 524	5.705 ± 0.030	1.000 ± 0.000	1510.58	160	3.07	19741 ± 141	532 ± 7	23045 ± 524	131469 ± 9434
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	17122 ± 451	7.955 ± 0.038	1.000 ± 0.000	2503.54	160	3.96	15432 ± 124	238 ± 5	17122 ± 451	136211 ± 14421
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	13481 ± 400	10.121 ± 0.046	1.000 ± 0.000	3417.58	160	4.62	11810 ± 109	145 ± 5	13481 ± 400	136444 ± 18950
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9429 ± 324	14.922 ± 0.067	1.000 ± 0.000	5045.71	160	5.62	9202 ± 96	89 ± 3	9429 ± 324	140692 ± 27390
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	6993 ± 277	20.446 ± 0.093	1.000 ± 0.000	8781.68	160	7.41	7395 ± 86	64 ± 2	6993 ± 277	142980 ± 42285
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5174 ± 235	28.353 ± 0.128	1.000 ± 0.000	9889.89	160	7.86	5904 ± 77	45 ± 2	5174 ± 235	146704 ± 52612
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4409 ± 212	33.362 ± 0.153	1.000 ± 0.000	14594.91	160	9.55	4854 ± 70	31 ± 2	4409 ± 212	147099 ± 67716

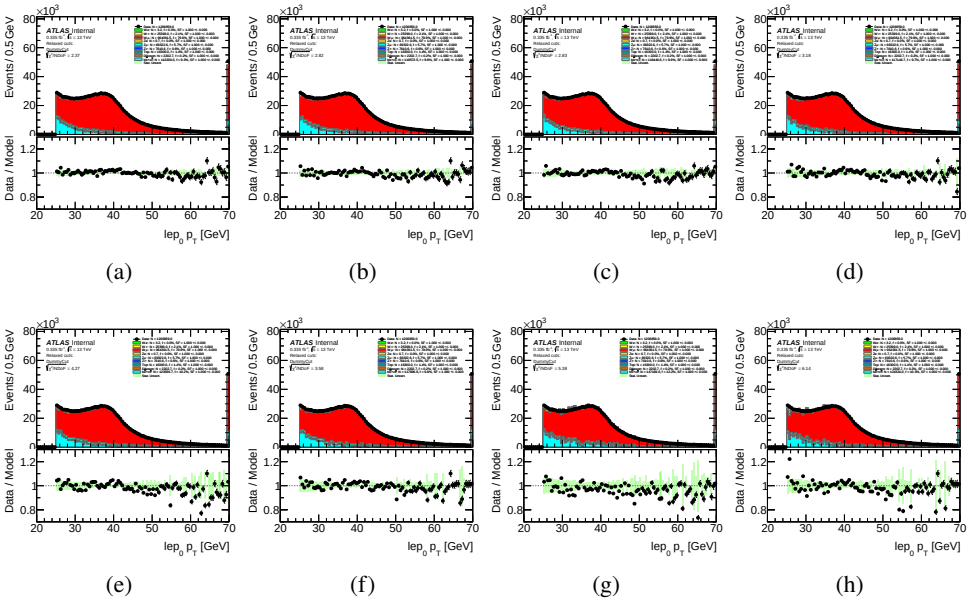
Table 31: Numbers for the muon E_T^{miss} fits on 13 TeV using $b\bar{b} + c\bar{c}$ MC samples. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	29180 ± 592	4.287 ± 0.026	1.000 ± 0.000	719.77	120	2.45	25514 ± 160	1755 ± 14	29180 ± 592	125106 ± 6493
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	23045 ± 524	5.621 ± 0.032	1.000 ± 0.000	1571.79	120	3.62	19741 ± 141	532 ± 7	23045 ± 524	129538 ± 10990
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	17122 ± 451	7.613 ± 0.042	1.000 ± 0.000	2222.31	120	4.30	15432 ± 124	238 ± 5	17122 ± 451	130349 ± 15092
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	13481 ± 400	10.082 ± 0.050	1.000 ± 0.000	3202.30	120	5.17	11810 ± 109	145 ± 5	13481 ± 400	135919 ± 21143
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9429 ± 324	15.306 ± 0.076	1.000 ± 0.000	5017.34	120	6.47	9202 ± 96	89 ± 3	9429 ± 324	144317 ± 32412
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	6993 ± 277	19.906 ± 0.089	1.000 ± 0.000	10730.66	120	9.46	7395 ± 86	64 ± 2	6993 ± 277	139207 ± 52539
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5174 ± 235	28.333 ± 0.123	1.000 ± 0.000	9086.66	120	8.70	5904 ± 77	45 ± 2	5174 ± 235	146599 ± 58169
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4409 ± 212	32.584 ± 0.140	1.000 ± 0.000	12470.62	120	10.19	4854 ± 70	31 ± 2	4409 ± 212	143669 ± 70551

Table 32: Numbers for the muon m_T^W fits on 13 TeV using $b\bar{b} + c\bar{c}$ MC samples. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.

u_T Cut	Slice No.	Iso_{min}	Iso_{max}	N_{FR}^{Data}	N_{FR}^{EW}	N_{CR1}^{Data}	N_{CR1}^{EW}	N_{CR1}^{MJ}	T	α	χ^2	NDoF	$\sqrt{\chi^2/NDoF}$	N_{CR2}^{Data}	N_{CR2}^{EW}	N_{CR2}^{MJ}	N_{SR}^{MJ}
–	0	0.1	0.15	1206859 ± 1099	1085200 ± 343	25514 ± 160	1755 ± 14	29180 ± 592	4.428 ± 0.029	1.000 ± 0.000	446.29	16	5.28	25514 ± 160	1755 ± 14	29180 ± 592	129215 ± 14561
–	1	0.15	0.2	1206859 ± 1099	1085200 ± 343	19741 ± 141	532 ± 7	23045 ± 524	5.920 ± 0.039	1.000 ± 0.000	757.45	16	6.88	19741 ± 141	532 ± 7	23045 ± 524	136427 ± 22225
–	2	0.2	0.25	1206859 ± 1099	1085200 ± 343	15432 ± 124	238 ± 5	17122 ± 451	8.222 ± 0.054	1.000 ± 0.000	1072.60	16	8.19	15432 ± 124	238 ± 5	17122 ± 451	140778 ± 31296
–	3	0.25	0.3	1206859 ± 1099	1085200 ± 343	11810 ± 109	145 ± 5	13481 ± 400	10.778 ± 0.071	1.000 ± 0.000	934.29	16	7.64	11810 ± 109	145 ± 5	13481 ± 400	145295 ± 33774
–	4	0.3	0.35	1206859 ± 1099	1085200 ± 343	9202 ± 96	89 ± 3	9429 ± 324	14.160 ± 0.094	1.000 ± 0.000	805.03	16	7.09	9202 ± 96	89 ± 3	9429 ± 324	133513 ± 33158
–	5	0.35	0.4	1206859 ± 1099	1085200 ± 343	7395 ± 86	64 ± 2	6993 ± 277	21.920 ± 0.146	1.000 ± 0.000	1491.62	16	9.66	7395 ± 86	64 ± 2	6993 ± 277	153289 ± 59519
–	6	0.4	0.45	1206859 ± 1099	1085200 ± 343	5904 ± 77	45 ± 2	5174 ± 235	27.277 ± 0.185	1.000 ± 0.000	1590.97	16	9.97	5904 ± 77	45 ± 2	5174 ± 235	141138 ± 64594
–	7	0.45	0.5	1206859 ± 1099	1085200 ± 343	4854 ± 70	31 ± 2	4409 ± 212	30.474 ± 0.218	1.000 ± 0.000	3818.26	16	15.45	4854 ± 70	31 ± 2	4409 ± 212	134366 ± 100694

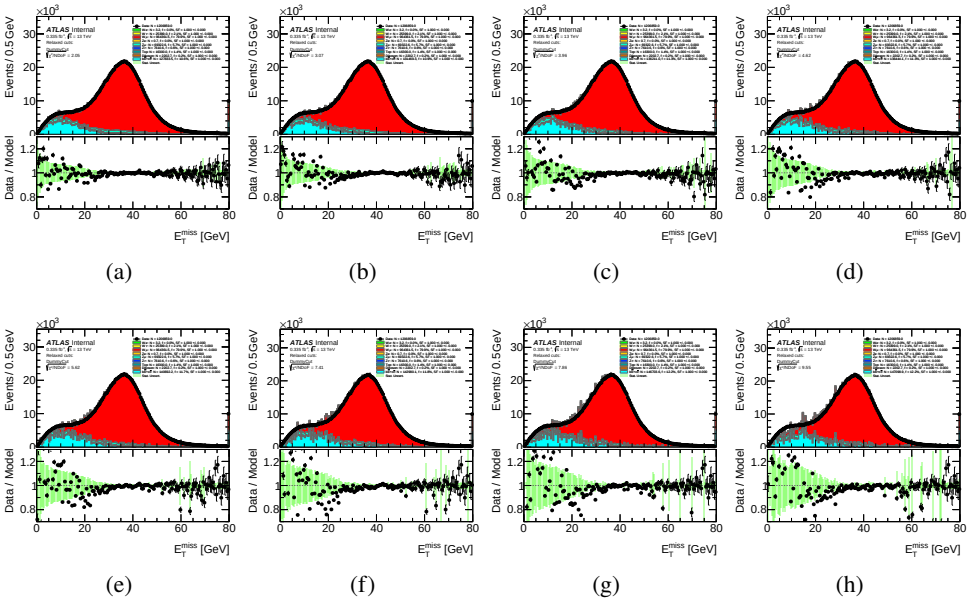
Table 33: Numbers for the muon $|\Delta\phi(lep_0 - MET)|$ fits on 13 TeV using $b\bar{b} + c\bar{c}$ MC samples. Numbers in FR are not affected by Slice No. The uncertainty of the numbers of events are statistical only except N_{SR}^{MJ} which is multiplied by $\sqrt{\chi^2/NDoF}$ and the uncertainty of the parameters are from the statistical uncertainty of the fitting.



(a) (b) (c) (d)

(e) (f) (g) (h)

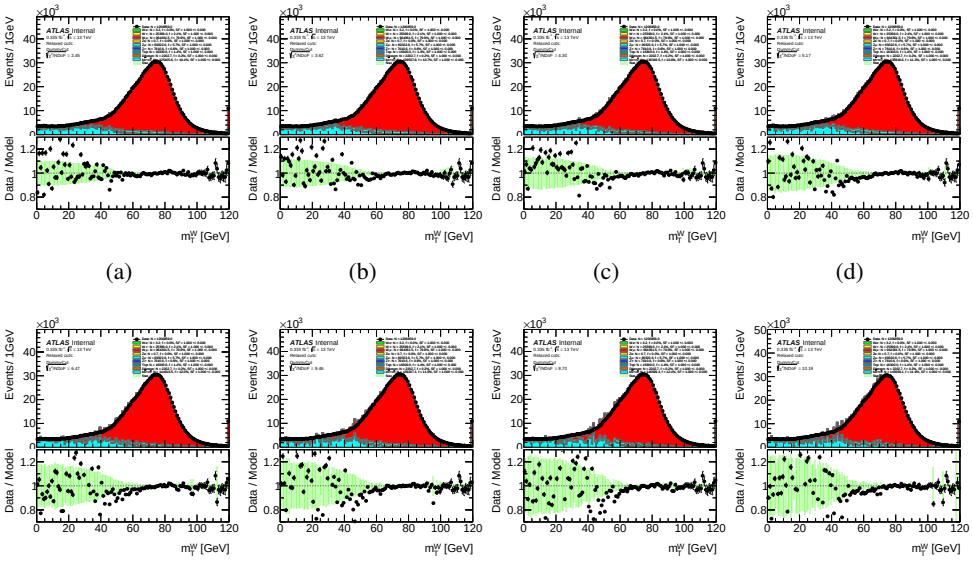
Figure 141: Multi-jet background template fits performed in distributions of p_T^{ℓ} of the muons channel using $b\bar{b} + c\bar{c}$ MC samples, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.15$ to $0.45 < ptvarcone20/p_T < 0.5$.



(a) (b) (c) (d)

(e) (f) (g) (h)

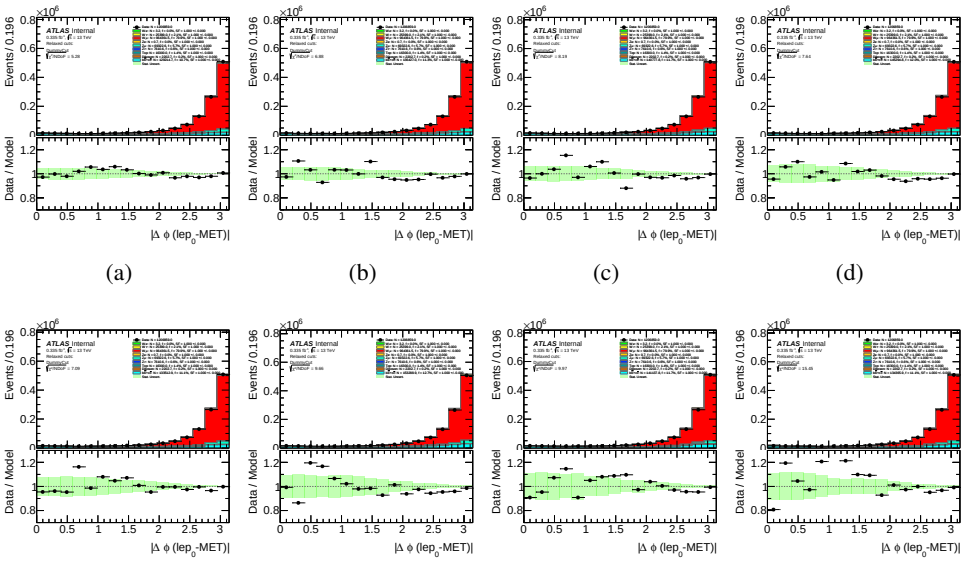
Figure 142: Multi-jet background template fits performed in distributions of E_T^{miss} of the muons channel using $b\bar{b} + c\bar{c}$ MC samples, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.15$ to $0.45 < ptvarcone20/p_T < 0.5$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 143: Multi-jet background template fits performed in distributions of m_T^W of the muons channel using $b\bar{b} + c\bar{c}$ MC samples, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.15$ to $0.45 < ptvarcone20/p_T < 0.5$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 144: Multi-jet background template fits performed in distributions of $|\Delta\phi(lep_0 - MET)|$ of the muons channel using $b\bar{b} + c\bar{c}$ MC samples, columns from (a) to (h) are the fits with the isolation slices from $0.1 < ptvarcone20/p_T < 0.15$ to $0.45 < ptvarcone20/p_T < 0.5$.

E.5 Acceptance correction functions for MJ background

E.5.1 MC sampels

To extract the MJ acceptance correction functions the following MC samples are used.

The properties of the $b\bar{b}$ and $c\bar{c}$ samples at $\sqrt{s}$ = 13 TeV are given in Table 34. The heavy-flavour production cross sections are higher than signal and other backgrounds, but their calculation carries significant uncertainties from the dijet production cross section (calculated at leading order), and the $b/c \rightarrow \mu$ decay fractions and fragmentation functions.

Process	Data set	Generator	σ·BR·ϵ_{filter} [nb]
$b\bar{b} \rightarrow \mu$	361250	Pythia8B+A14+NNPDF23LO	186
$c\bar{c} \rightarrow \mu$	361251	Pythia8B+A14+NNPDF23LO	58

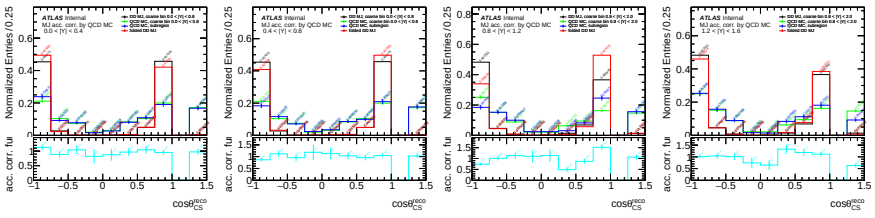
Table 34: MC simulation of $b\bar{b}$ and $c\bar{c}$ decay to muon at $\sqrt{s}$ = 13 TeV.

For the electron channel multijet production was generated using Pythia [8.230] [47] with leading-order matrix elements for dijet production which were matched to the parton shower. Table 35 details the properties of the JF17 samples used.

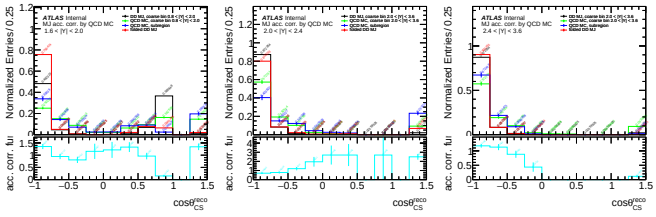
Process	Data set	Generator	σ·BR·ϵ_{filter} [nb]
$JF17$	423300	Pythia8B+A14+NNPDF23LO	180×10^3

Table 35: MC simulation of the MJ JF17 sample $\sqrt{s}$ = 13 TeV.

E.5.2 Electron channel multi-jet background 1D acceptance correction functions

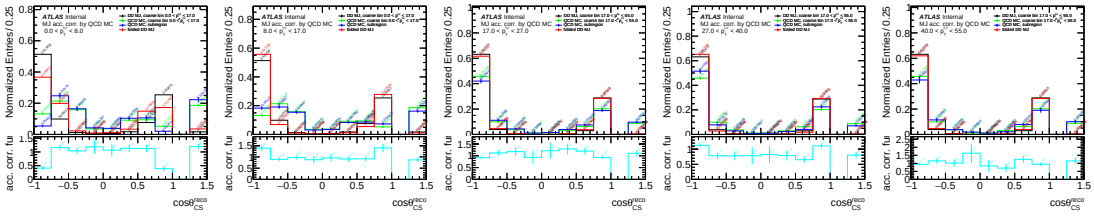


(a) (b) (c) (d)

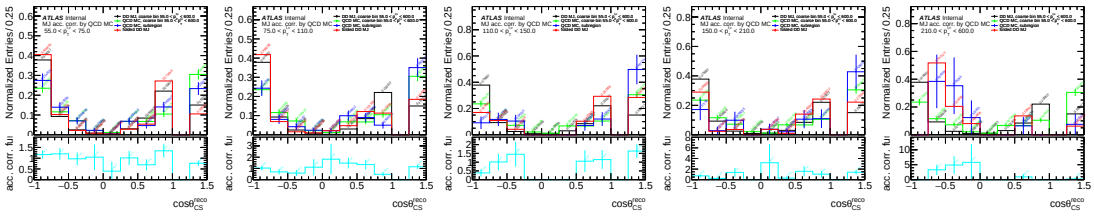


(e) (f) (g)

Figure 145: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $|Y|$ bins for $W^- \to e^-\overline{\nu}$ channel.

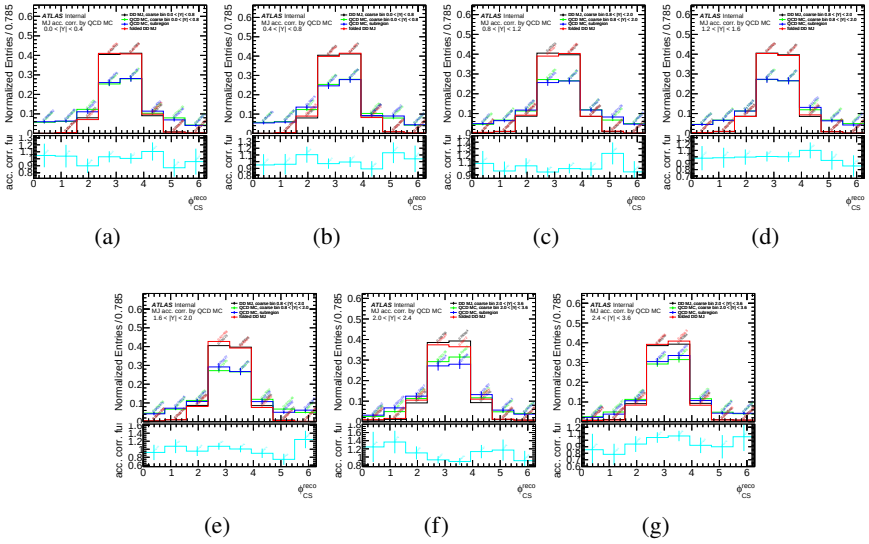


(a) (b) (c) (d) (e)



(f) (g) (h) (i) (j)

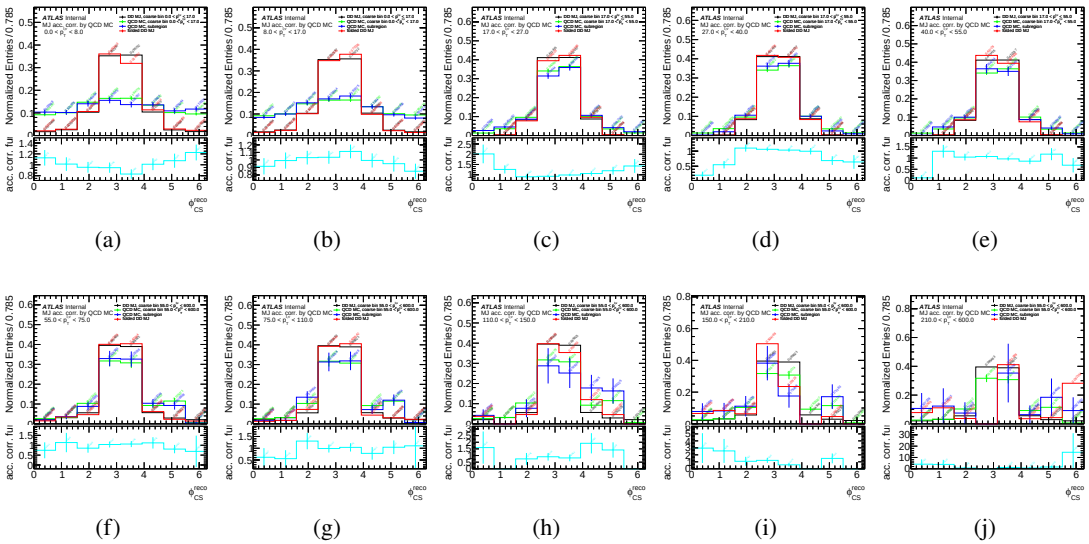
Figure 146: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $p_T^{\ell,\nu}$ bins for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g)

Figure 147: Multi-jet background template acceptance correction functions for ϕ_{CS} for $|Y|$ bins for $W^- \rightarrow e^-\overline{\nu}$ channel.

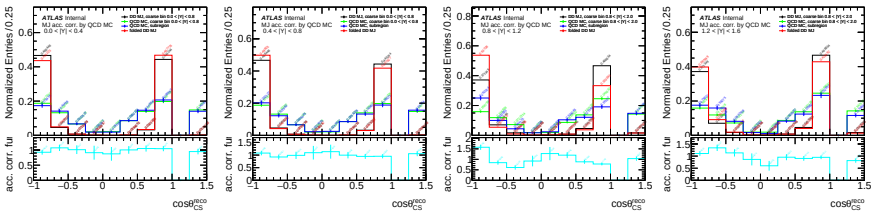


(a) (b) (c) (d) (e)

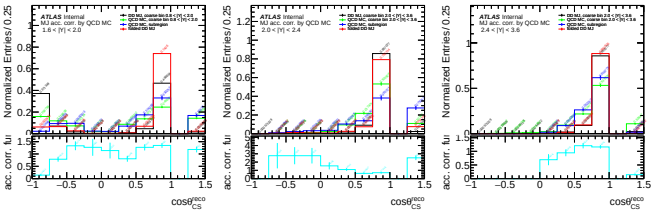
(f) (g) (h) (i) (j)

Figure 148: Multi-jet background template acceptance correction functions for ϕ_{CS} for $p_T^{\ell,\nu}$ bins for $W^- \rightarrow e^-\overline{\nu}$ channel.

E.5.3 Positron channel multi-jet background 1D acceptance correction functions

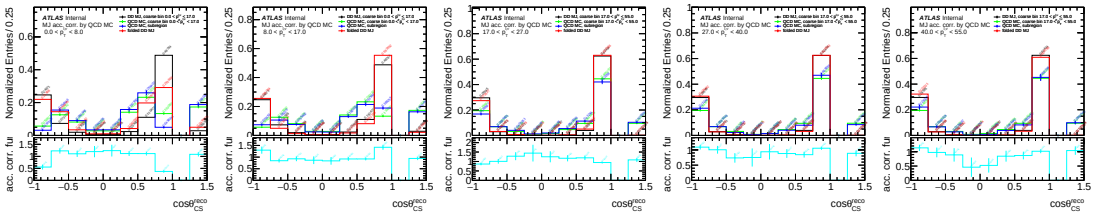


(a) (b) (c) (d)

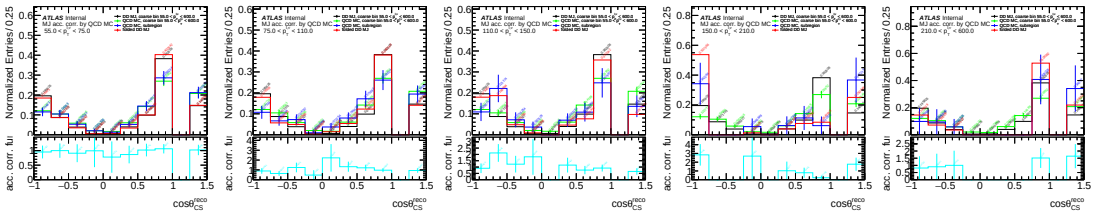


(e) (f) (g)

Figure 149: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $|Y|$ bins for $W^+ \to e^+\nu$ channel.

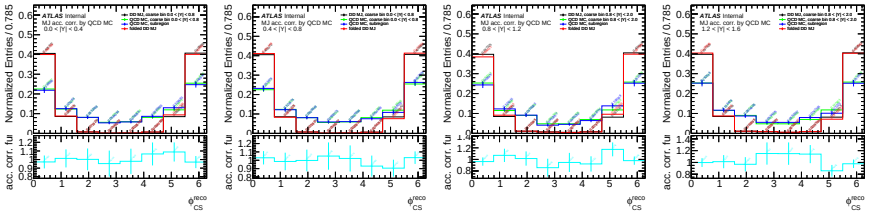


(a) (b) (c) (d) (e)

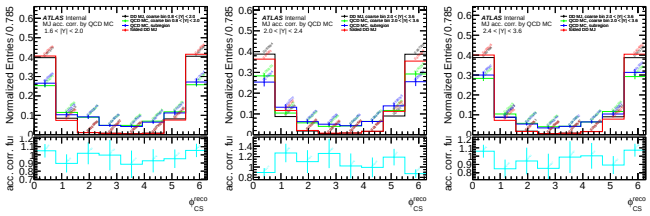


(f) (g) (h) (i) (j)

Figure 150: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $p_T^{\ell,\nu}$ bins for $W^+ \to e^+\nu$ channel.

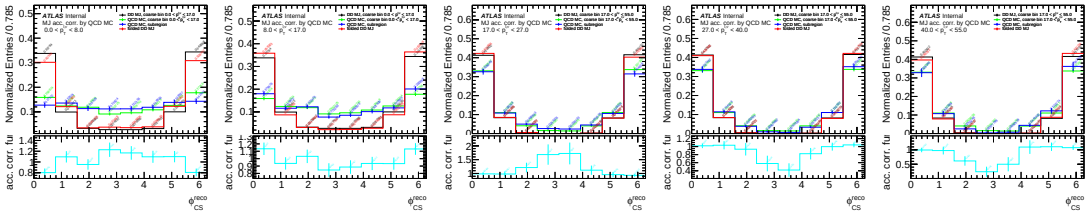


(a) (b) (c) (d)

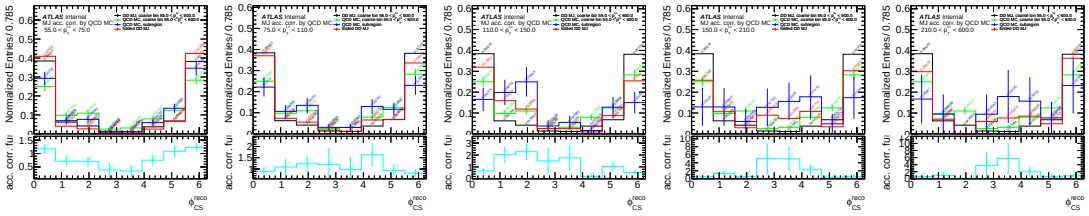


(e) (f) (g)

Figure 151: Multi-jet background template acceptance correction functions for ϕ_{CS} for $|Y|$ bins for $W^+ \rightarrow e^+\nu$ channel.



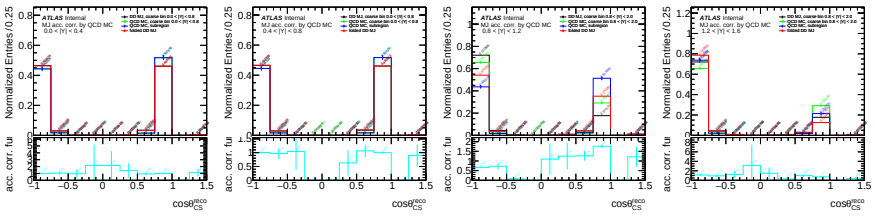
(a) (b) (c) (d) (e)



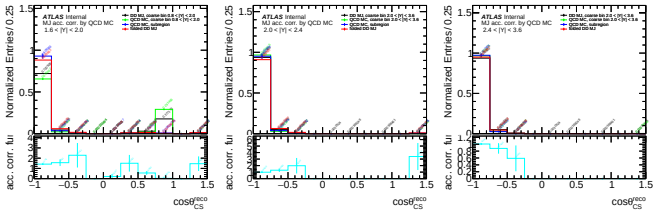
(f) (g) (h) (i) (j)

Figure 152: Multi-jet background template acceptance correction functions for ϕ_{CS} for $p_T^{\ell,\nu}$ bins for $W^+ \rightarrow e^+\nu$ channel.

E.5.4 Muon channel multi-jet background 1D acceptance correction functions

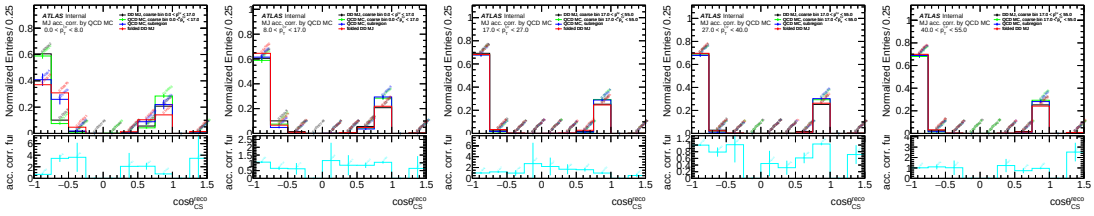


(a) (b) (c) (d)

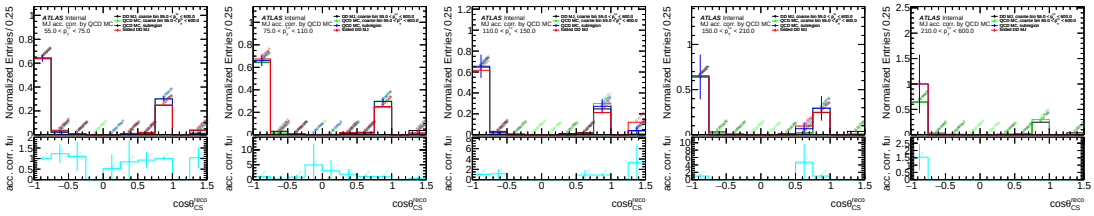


(e) (f) (g)

Figure 153: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $|Y|$ bins for $W^- \to \mu^-\overline{\nu}$ channel.

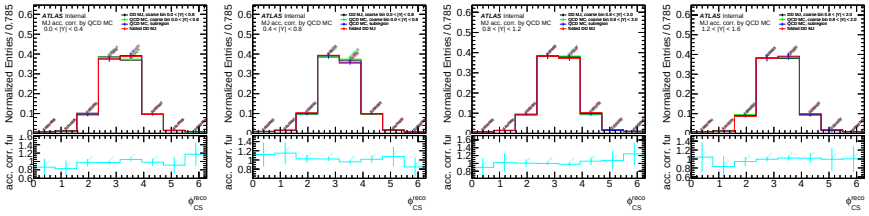


(a) (b) (c) (d) (e)

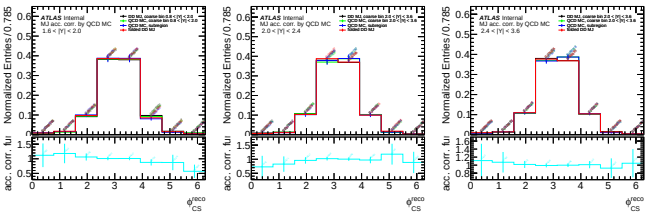


(f) (g) (h) (i) (j)

Figure 154: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $p_T^{\ell,\nu}$ bins for $W^- \to \mu^-\overline{\nu}$ channel.

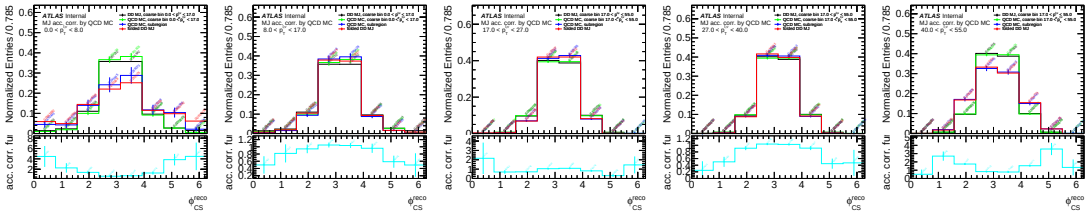


(a) (b) (c) (d)

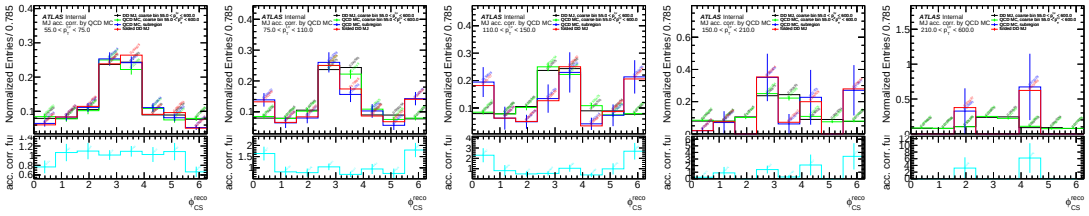


(e) (f) (g)

Figure 155: Multi-jet background template acceptance correction functions for ϕ_{CS} for $|Y|$ bins for $W^- \rightarrow \mu^-\overline{\nu}$ channel.



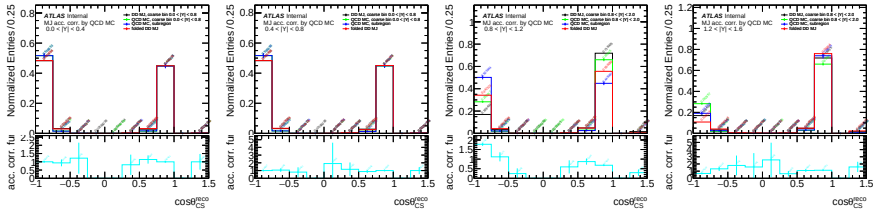
(a) (b) (c) (d) (e)



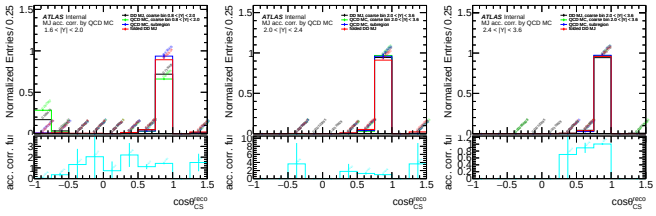
(f) (g) (h) (i) (j)

Figure 156: Multi-jet background template acceptance correction functions for ϕ_{CS} for $p_T^{\ell,\nu}$ bins for $W^- \rightarrow \mu^-\overline{\nu}$ channel.

E.5.5 Anti-muon channel multi-jet background 1D acceptance correction functions

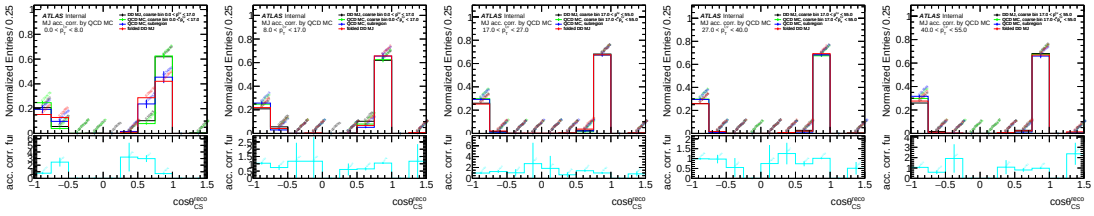


(a) (b) (c) (d)

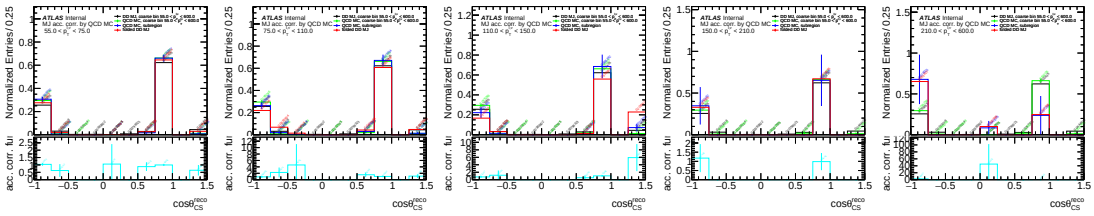


(e) (f) (g)

Figure 157: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $|Y|$ bins for $W^+ \to \mu^+\nu$ channel.

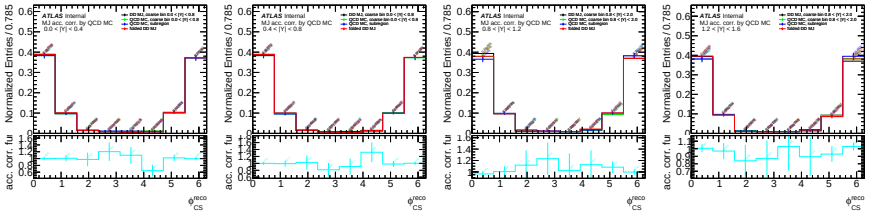


(a) (b) (c) (d) (e)

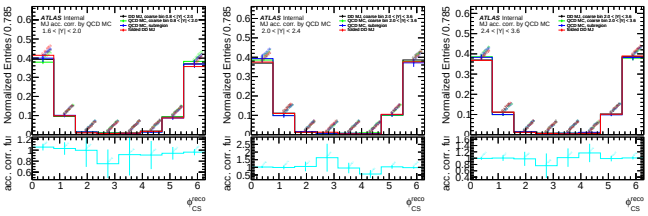


(f) (g) (h) (i) (j)

Figure 158: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ for $p_T^{\ell,\nu}$ bins for $W^+ \to \mu^+\nu$ channel.

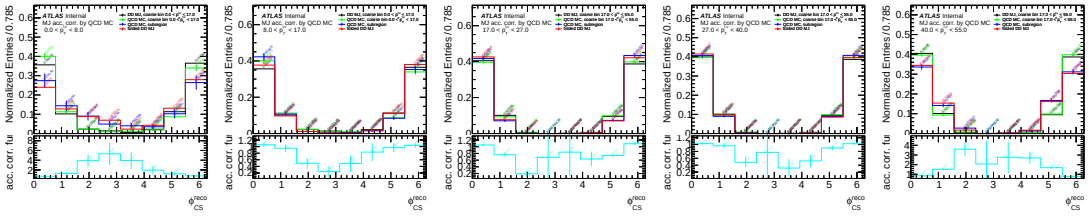


(a) (b) (c) (d)

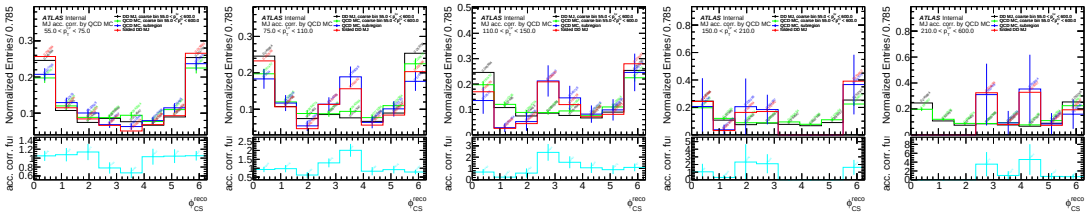


(e) (f) (g)

Figure 159: Multi-jet background template acceptance correction functions for ϕ_{CS} for $|Y|$ bins for $W^+ \rightarrow \mu^+\nu$ channel.



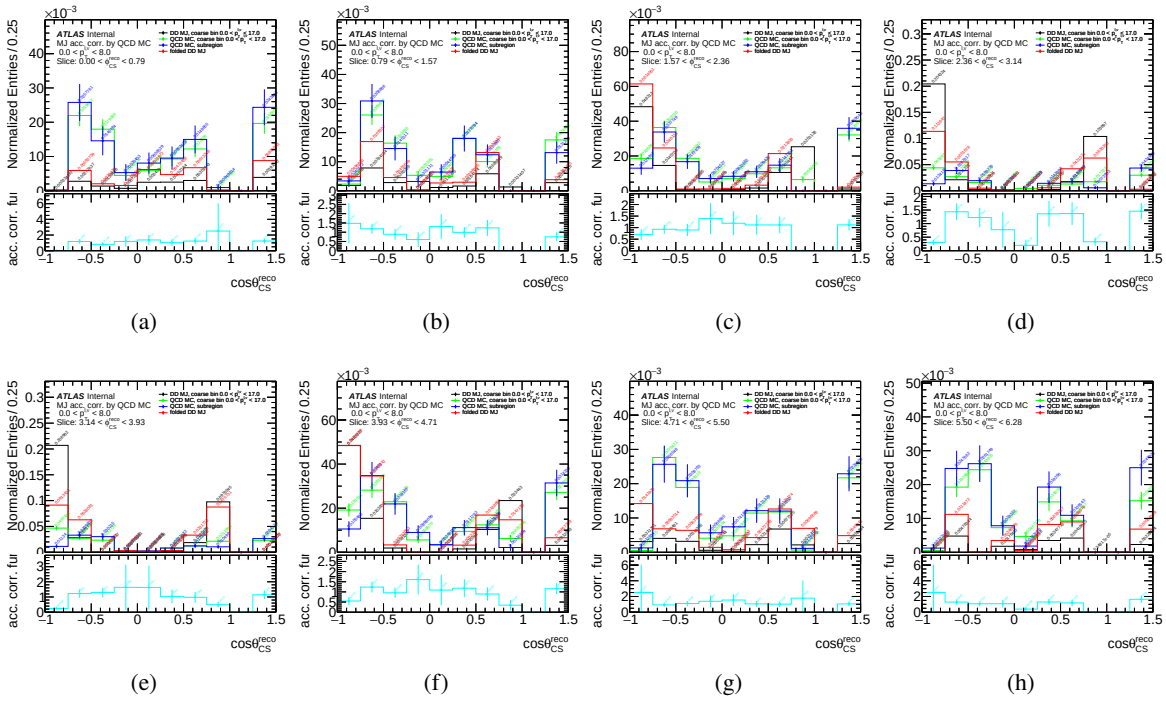
(a) (b) (c) (d) (e)



(f) (g) (h) (i) (j)

Figure 160: Multi-jet background template acceptance correction functions for ϕ_{CS} for $p_T^{\ell,\nu}$ bins for $W^+ \rightarrow \mu^+\nu$ channel.

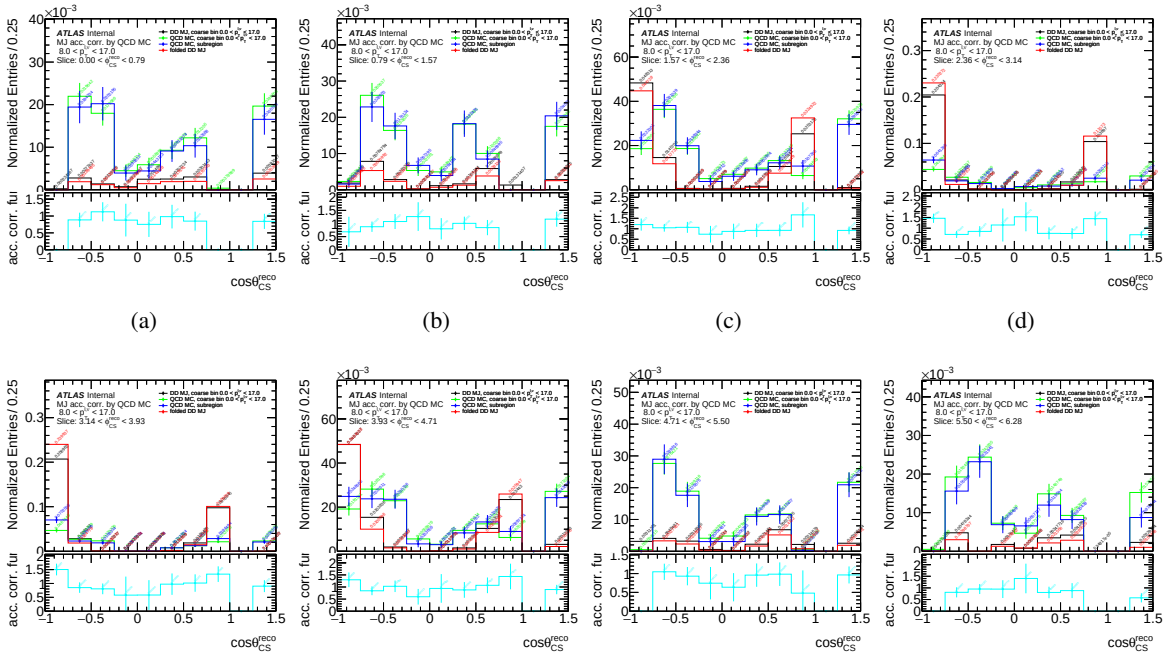
E.5.6 Electron channel multi-jet background 2D acceptance correction functions



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 161: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \rightarrow e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 162: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

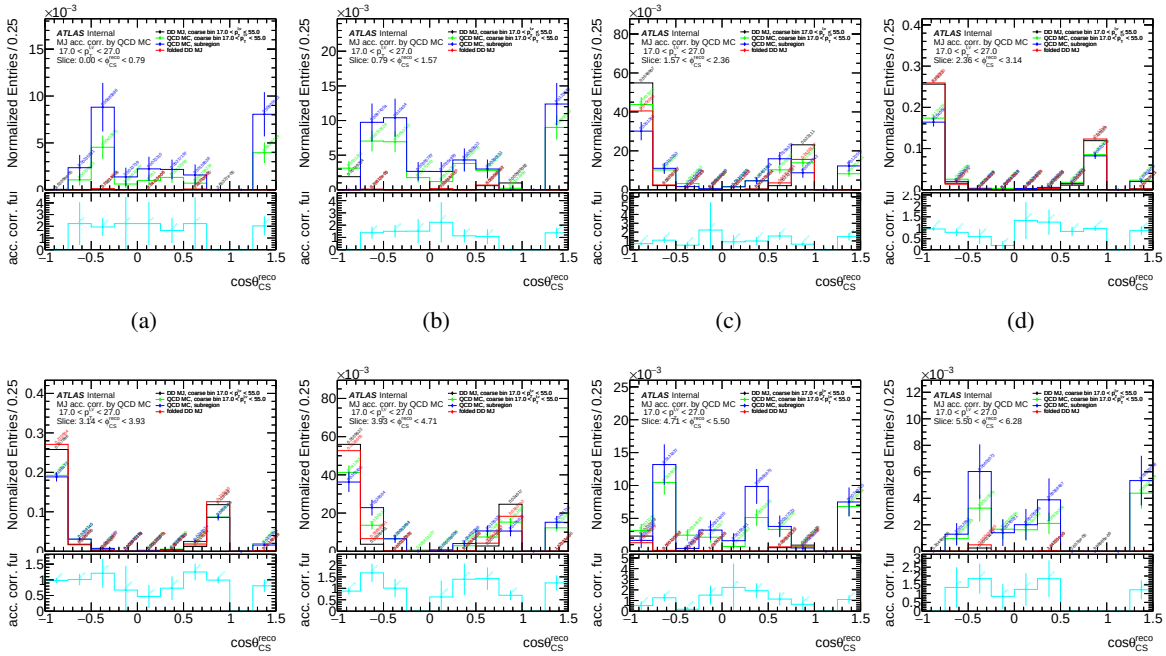


Figure 163: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

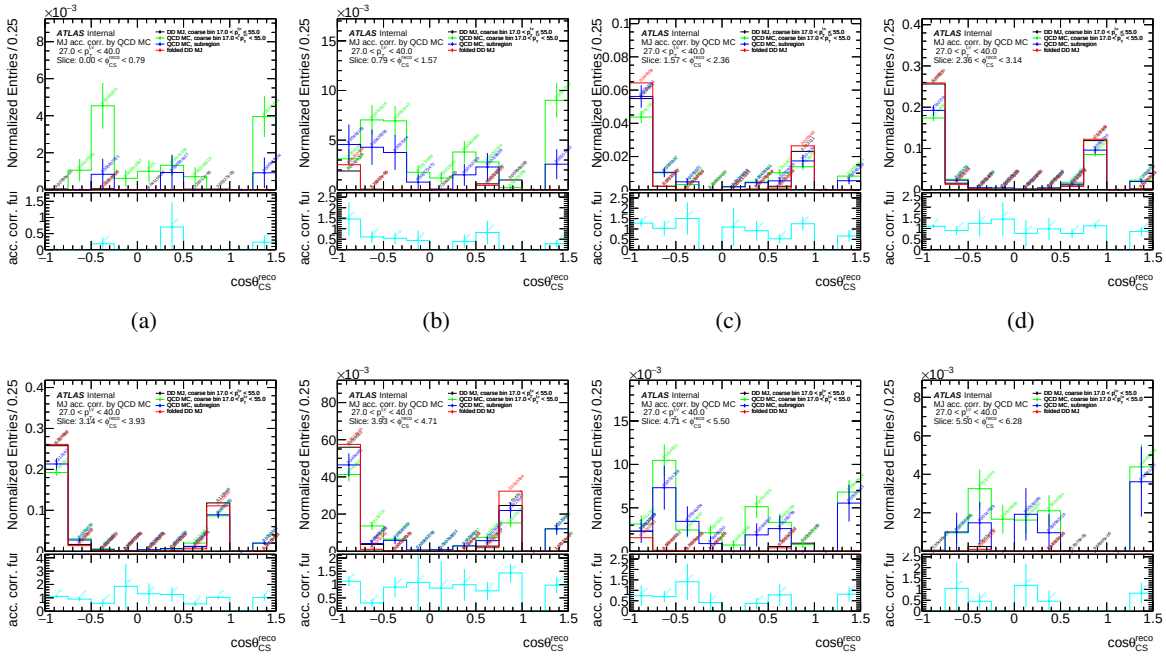


Figure 164: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

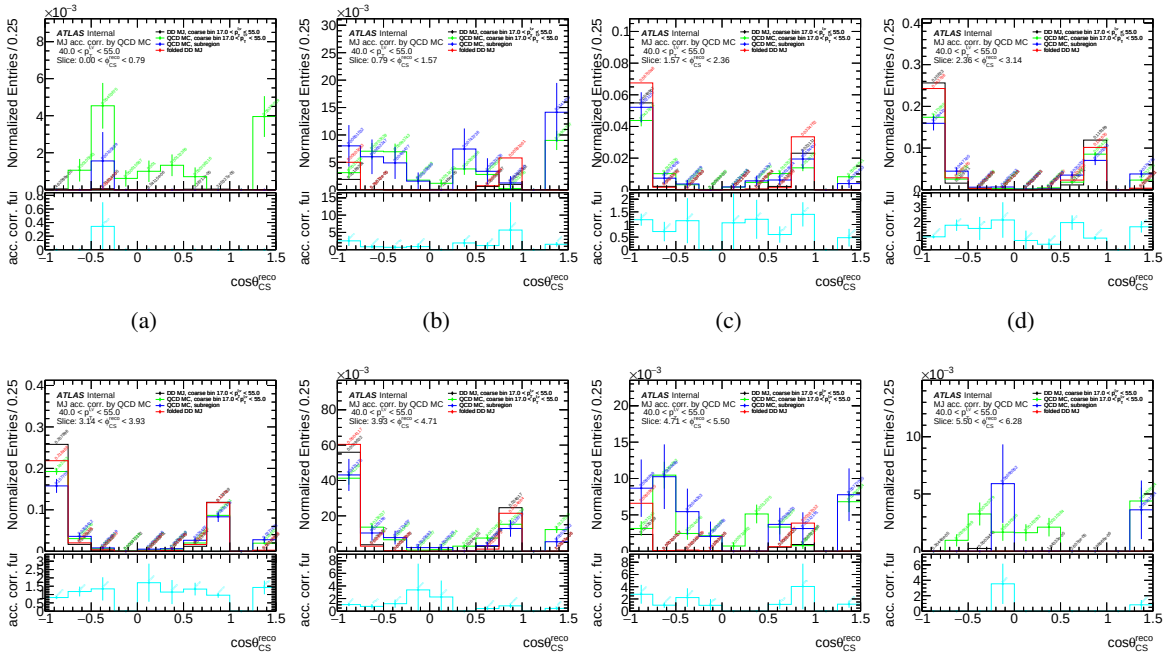


Figure 165: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

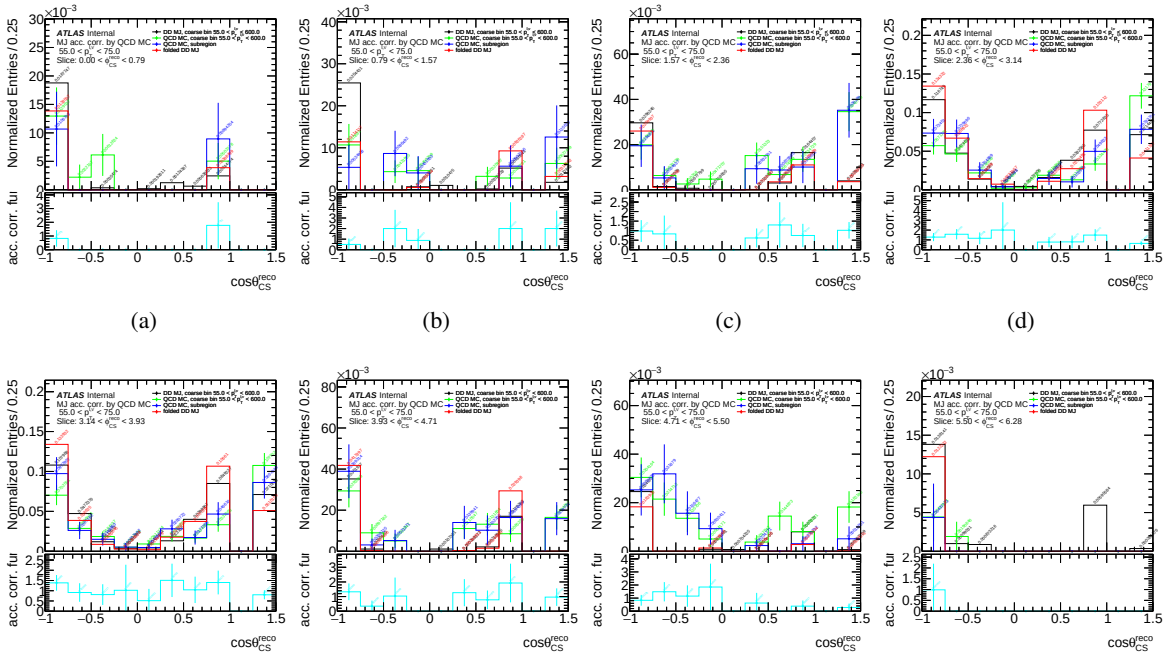


Figure 166: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

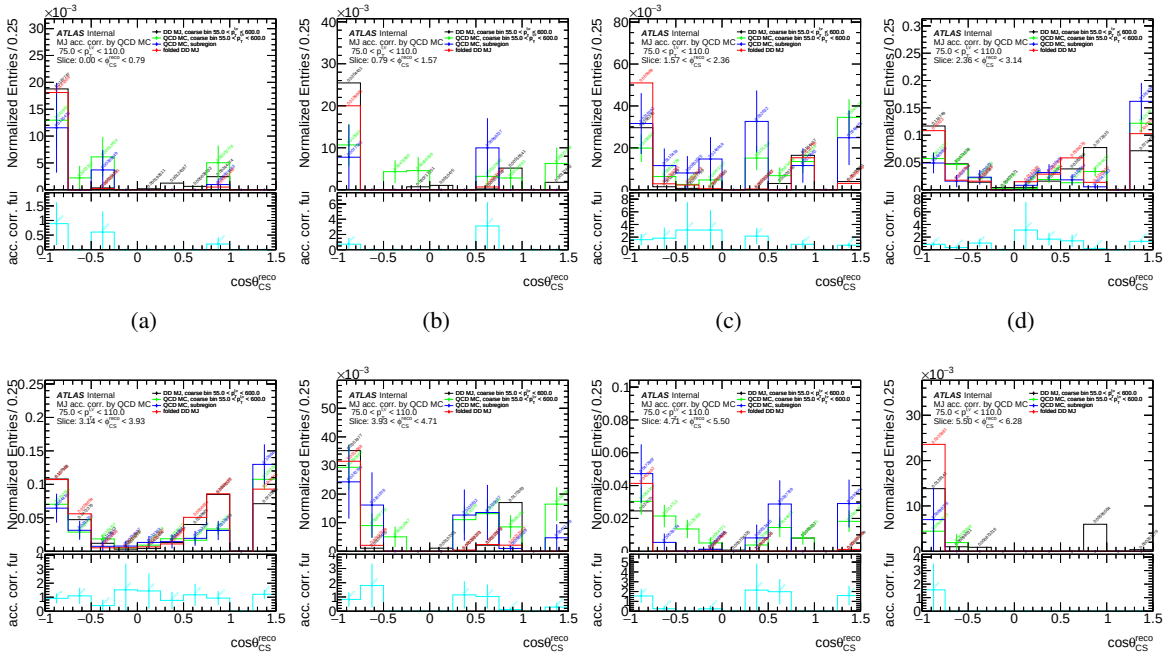


Figure 167: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

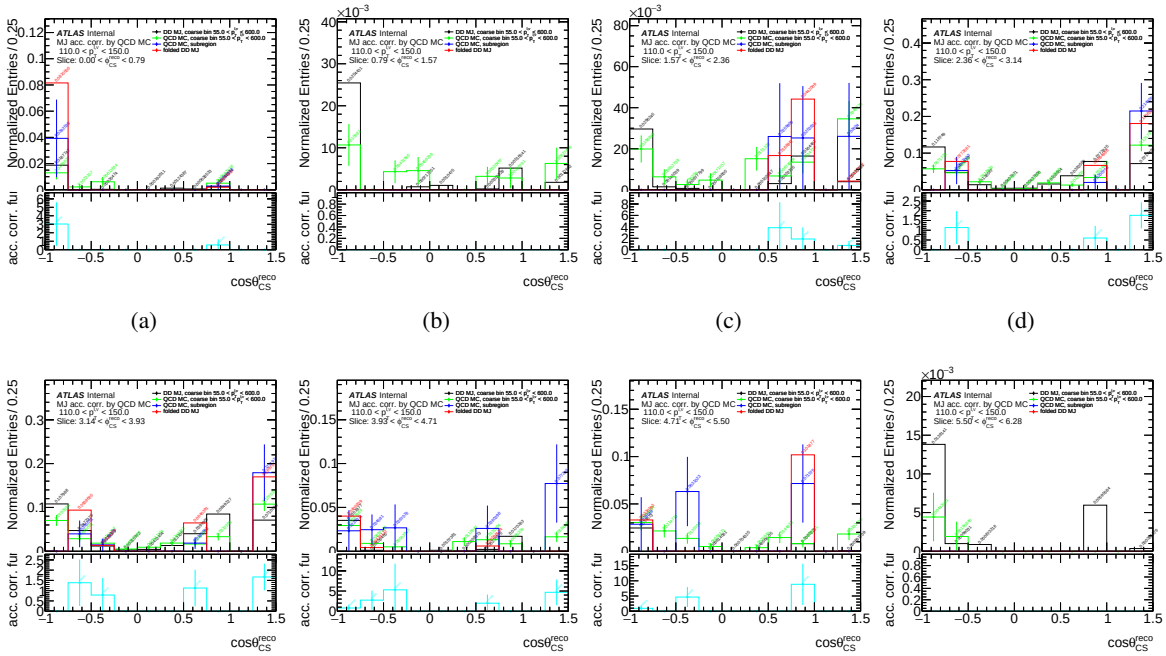


Figure 168: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

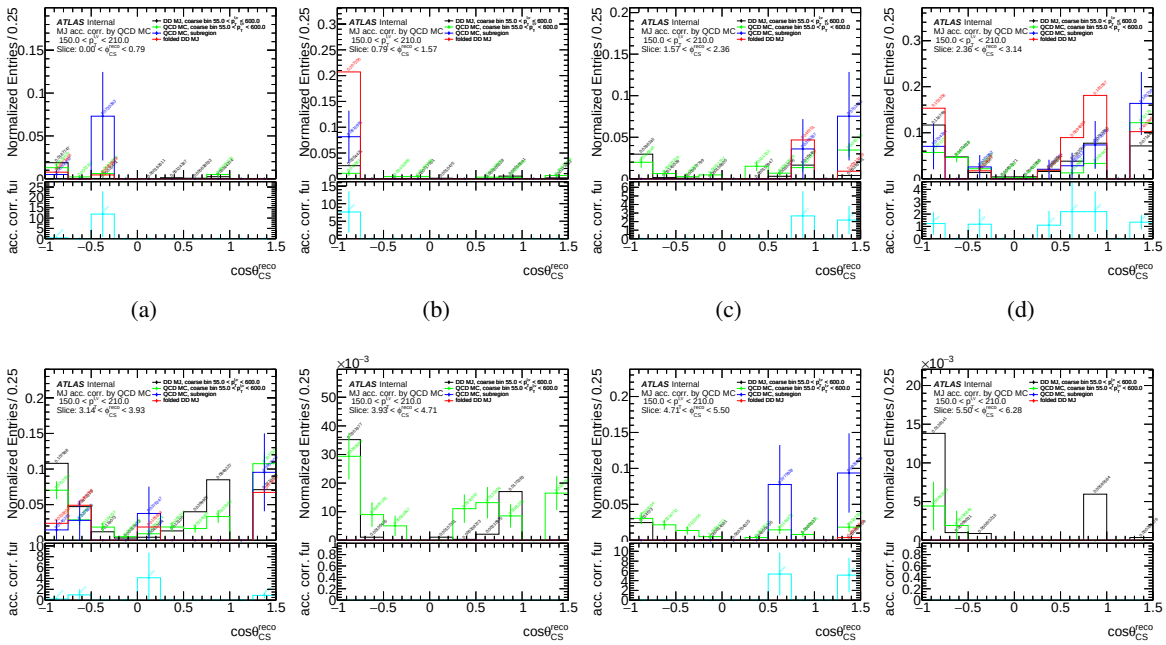


Figure 169: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

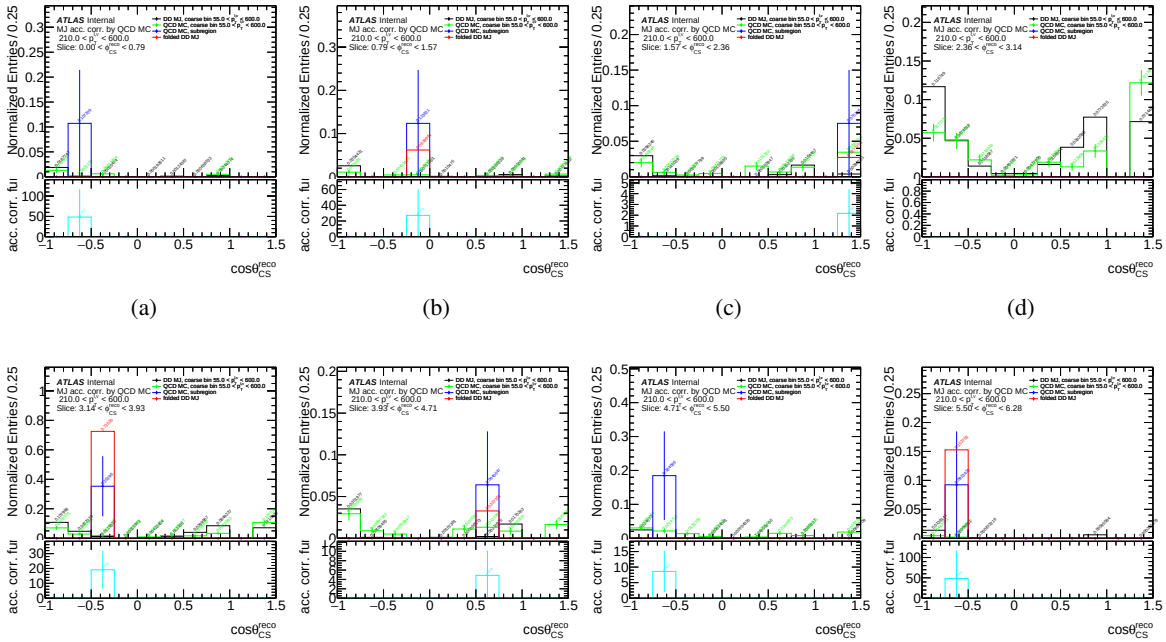


Figure 170: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

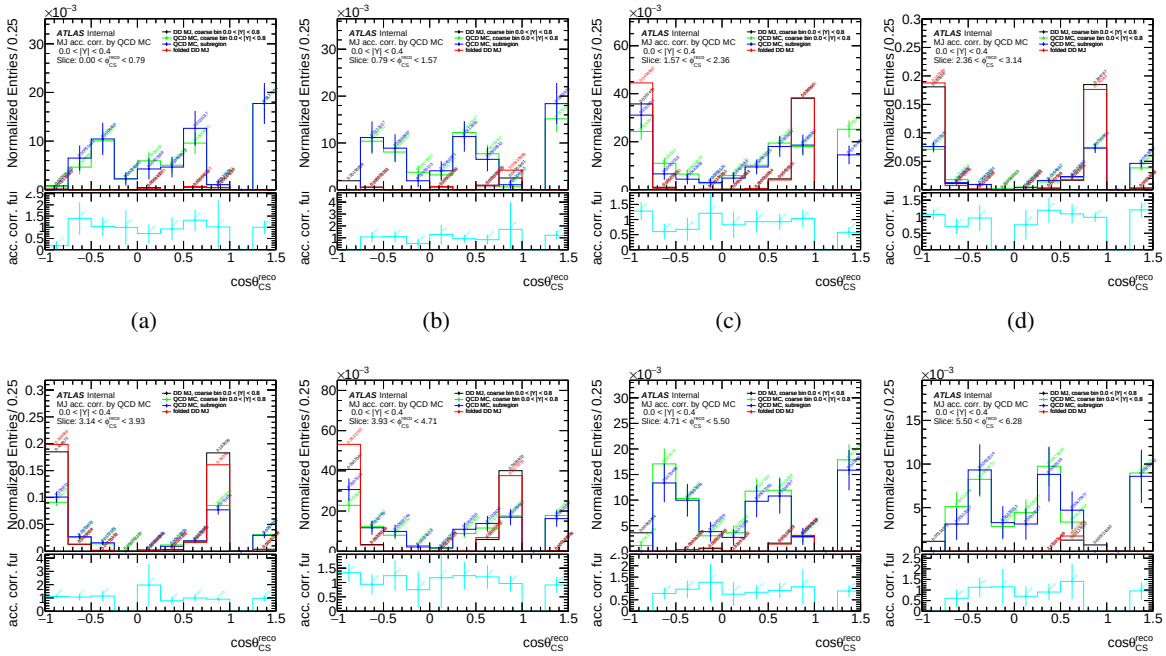


Figure 171: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^- \rightarrow e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

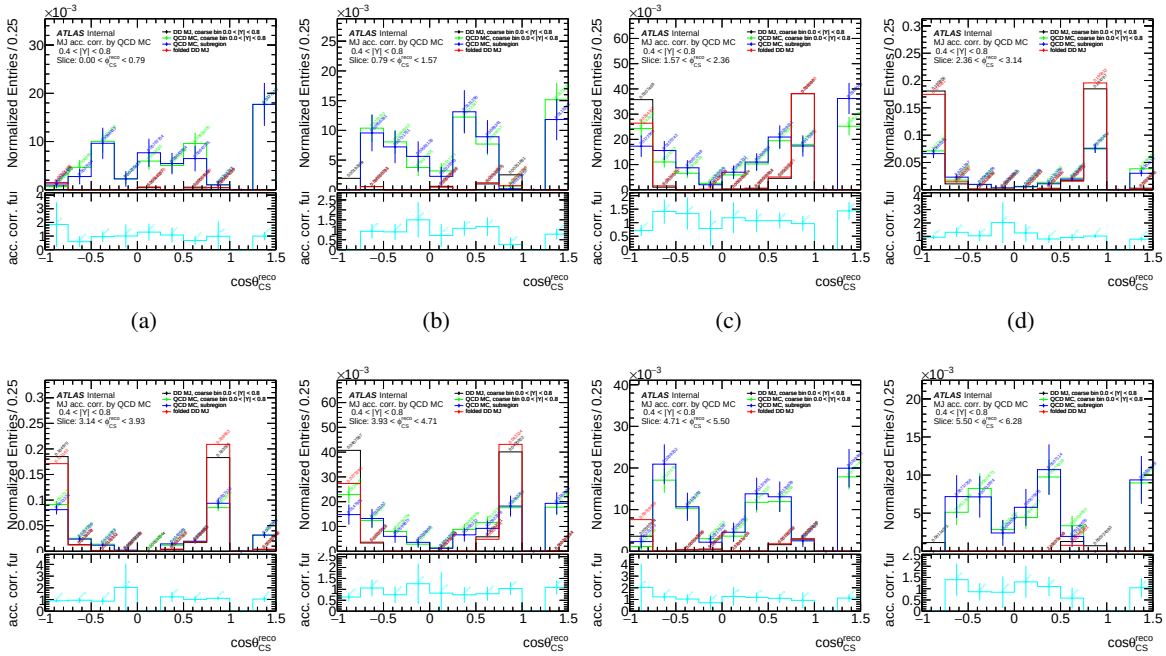
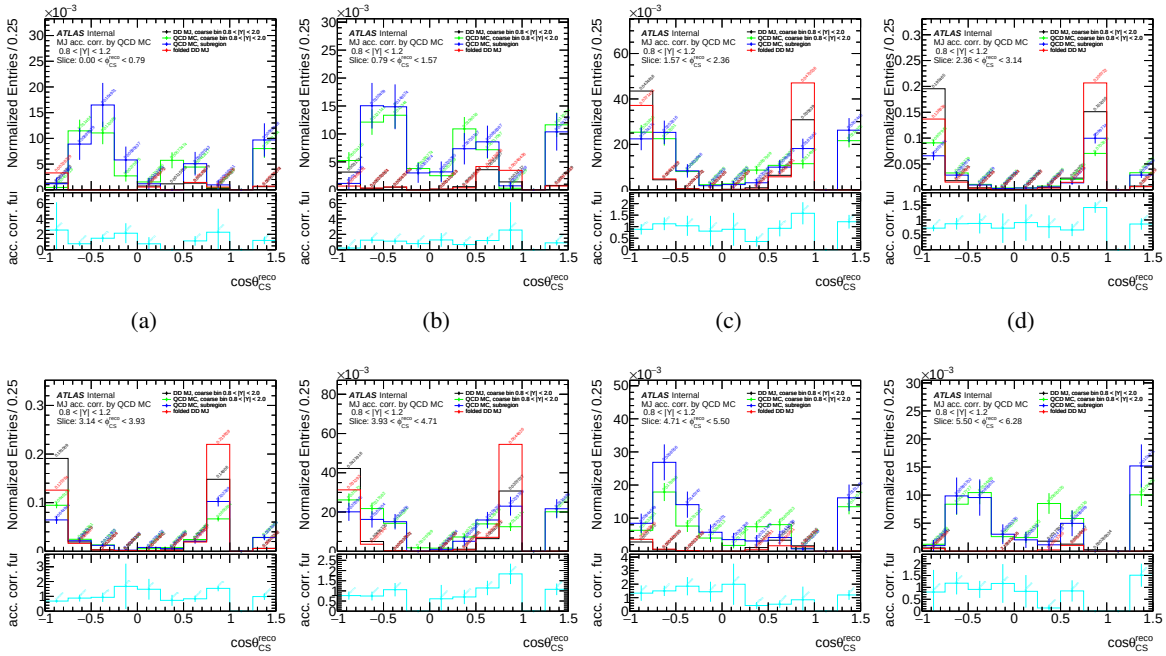


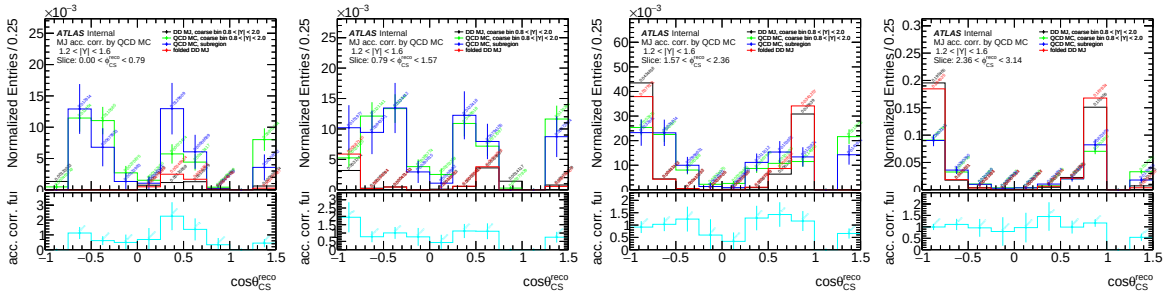
Figure 172: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



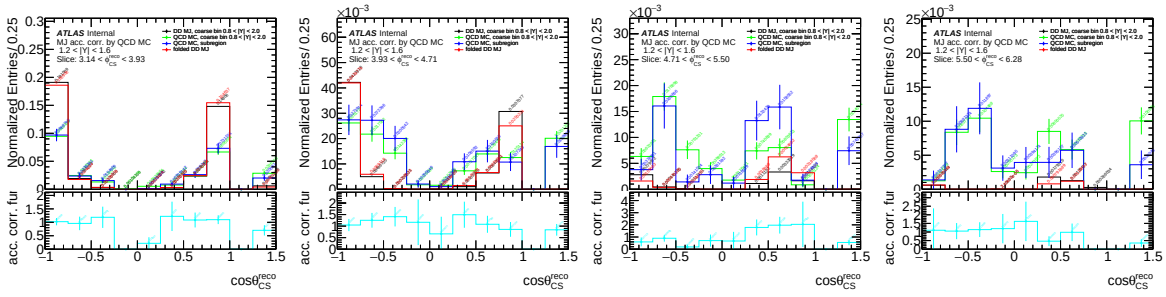
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 173: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 174: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

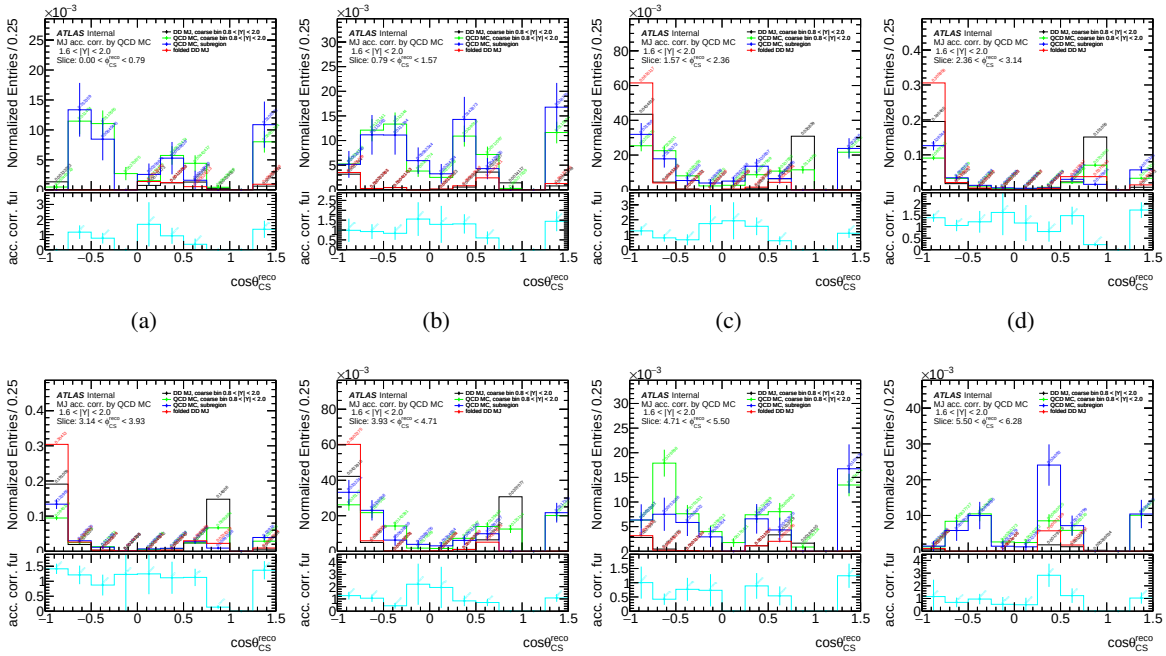


Figure 175: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^- \rightarrow e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

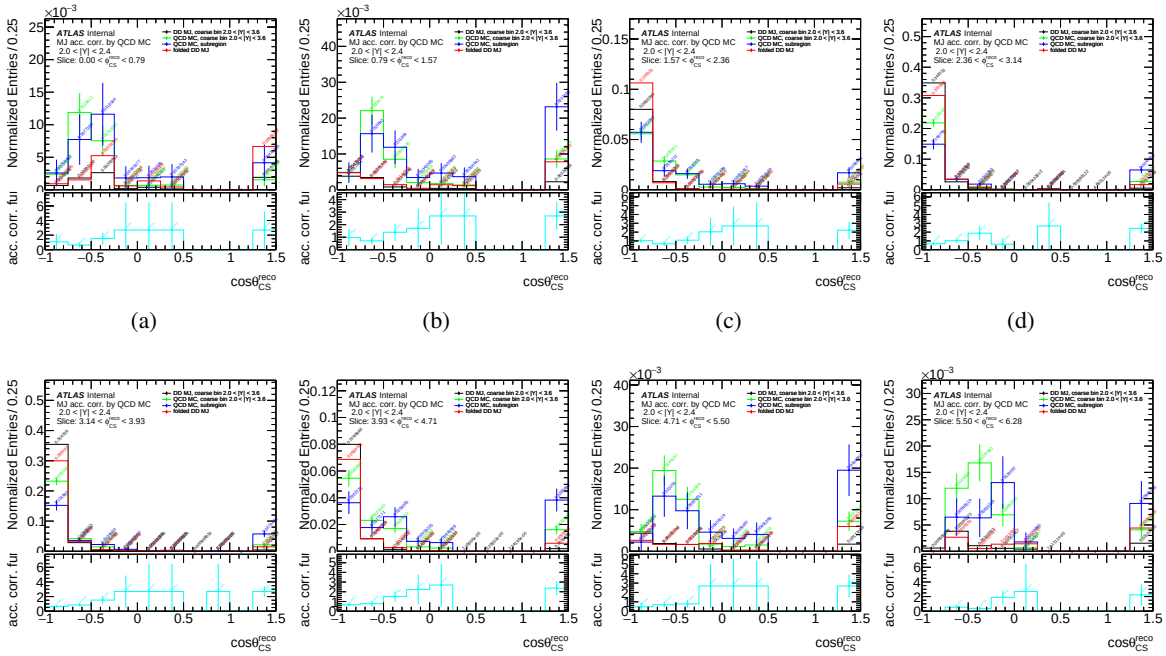
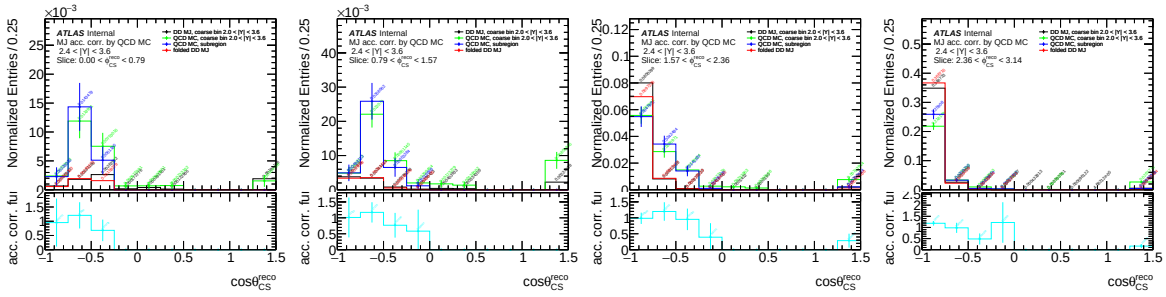


Figure 176: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 177: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^- \to e^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

E.5.7 Positron channel multi-jet background 2D acceptance correction functions

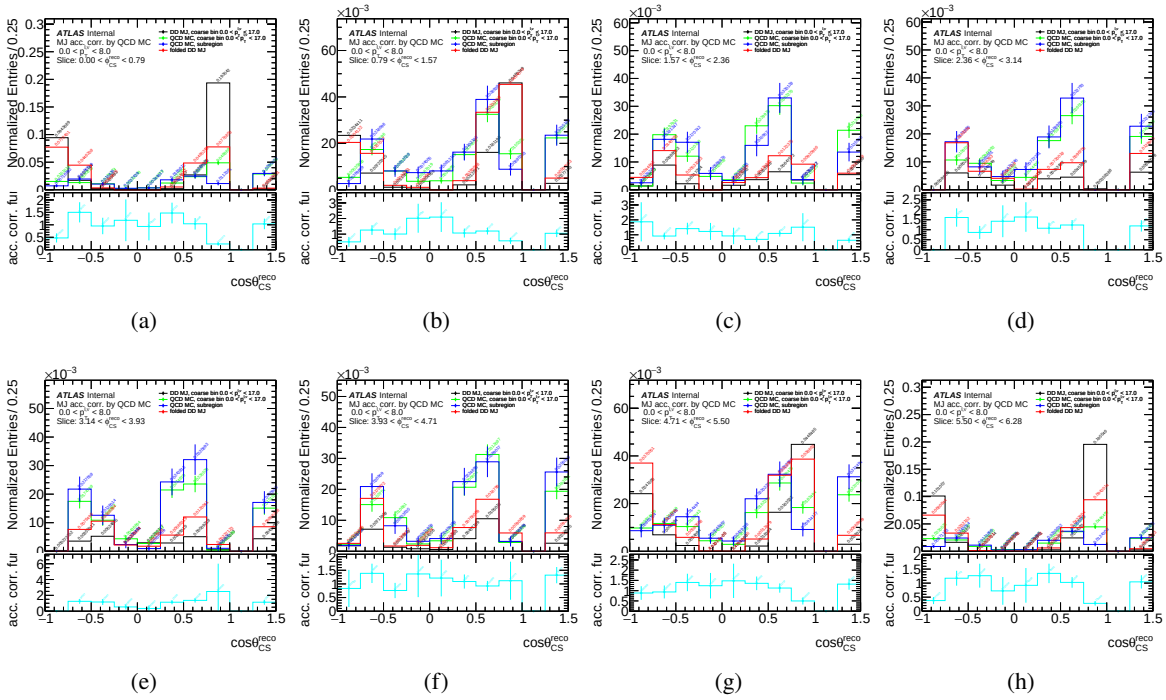


Figure 178: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

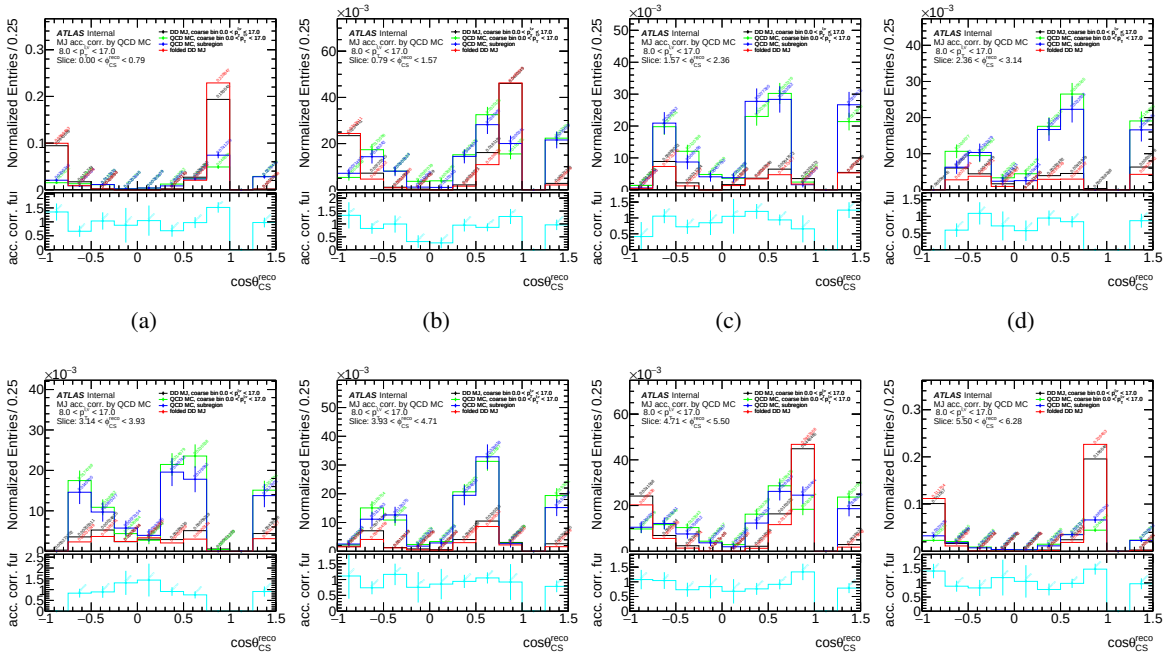


Figure 179: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

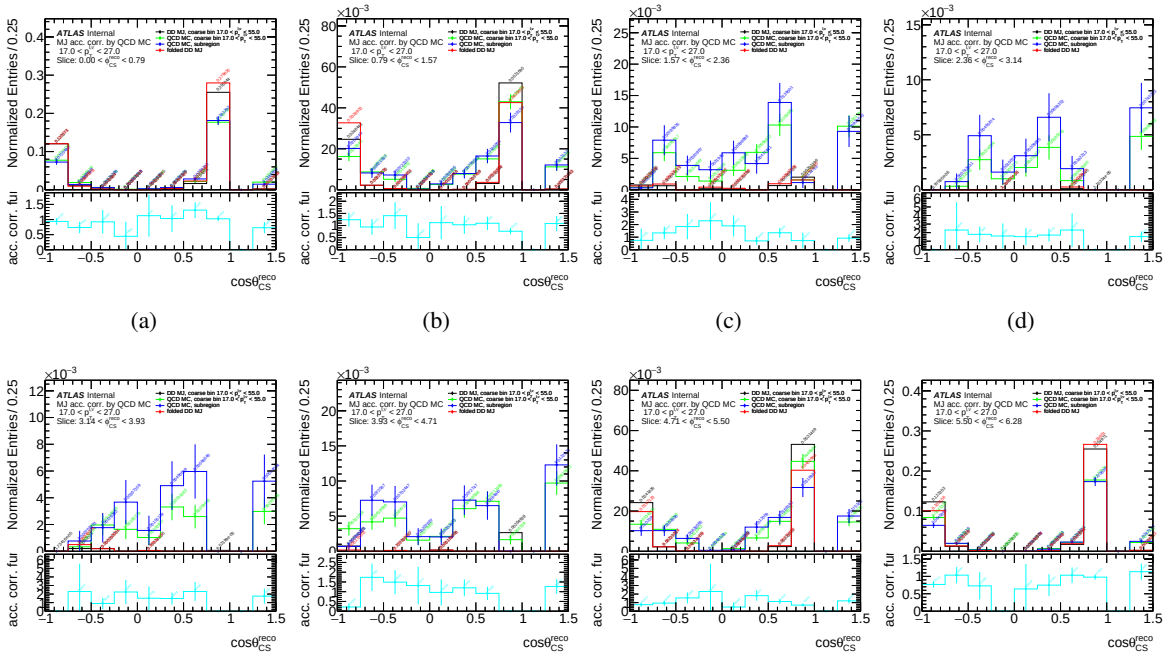


Figure 180: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

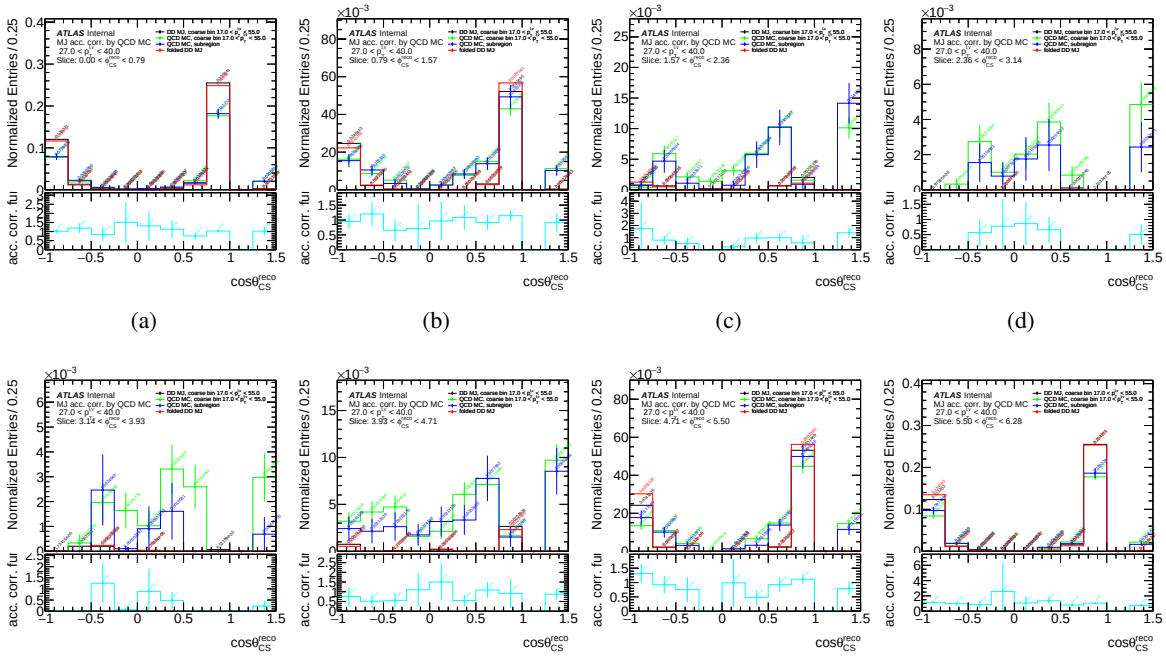
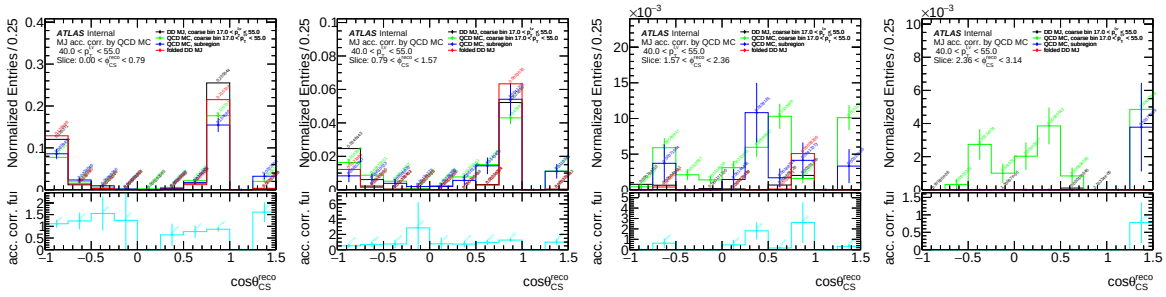
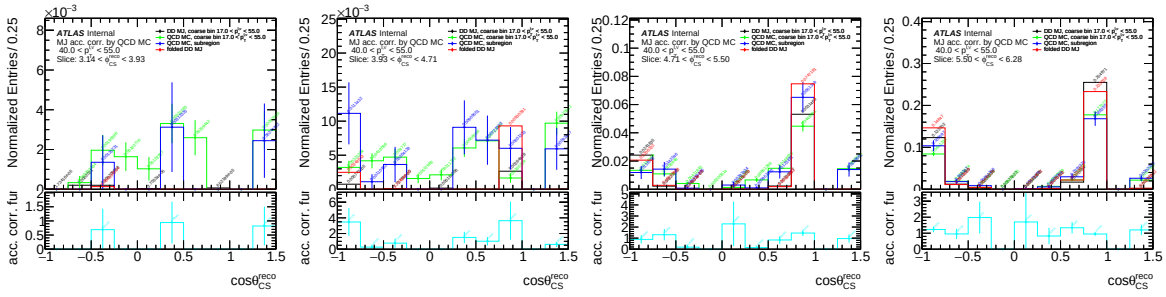


Figure 181: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 182: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

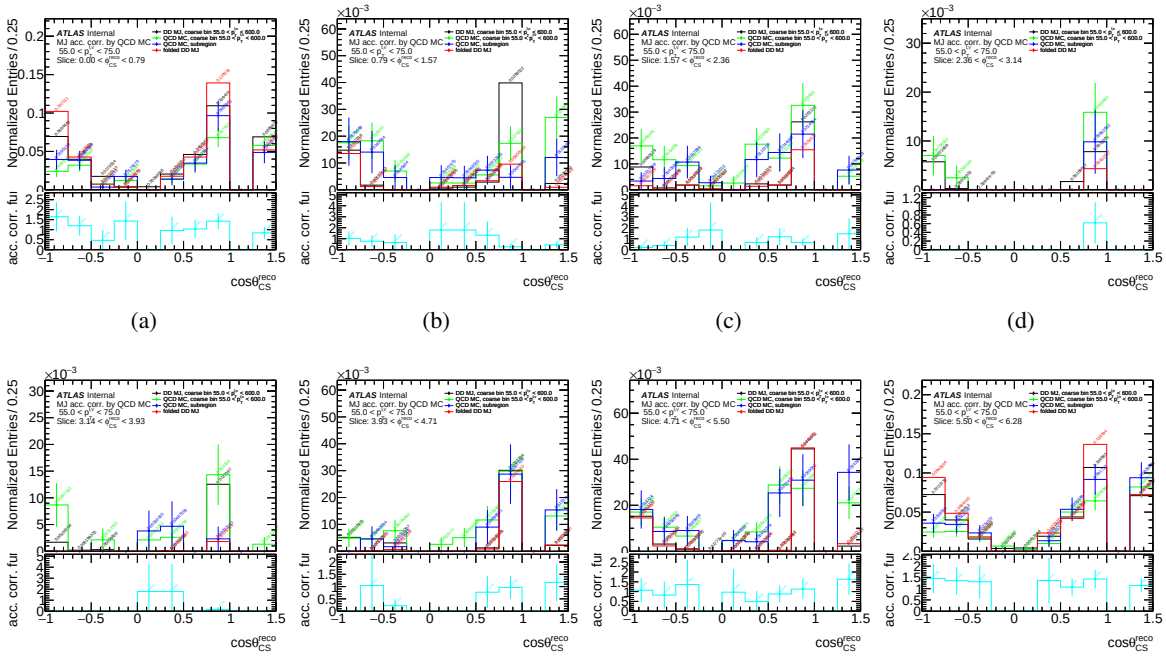


Figure 183: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

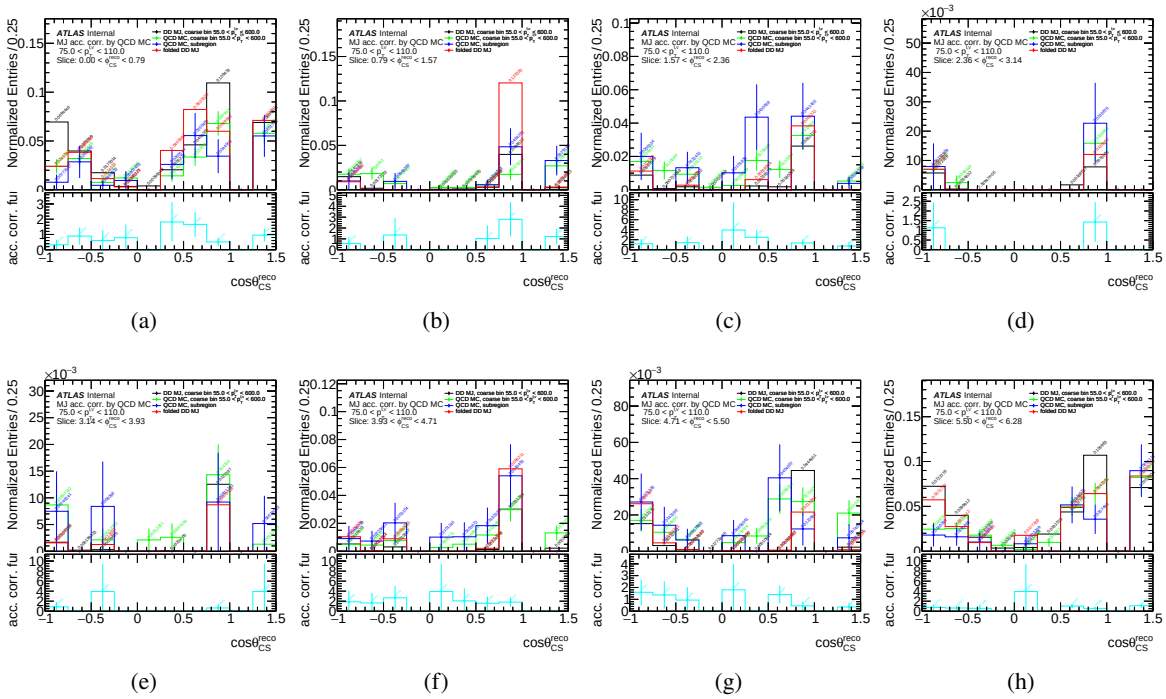
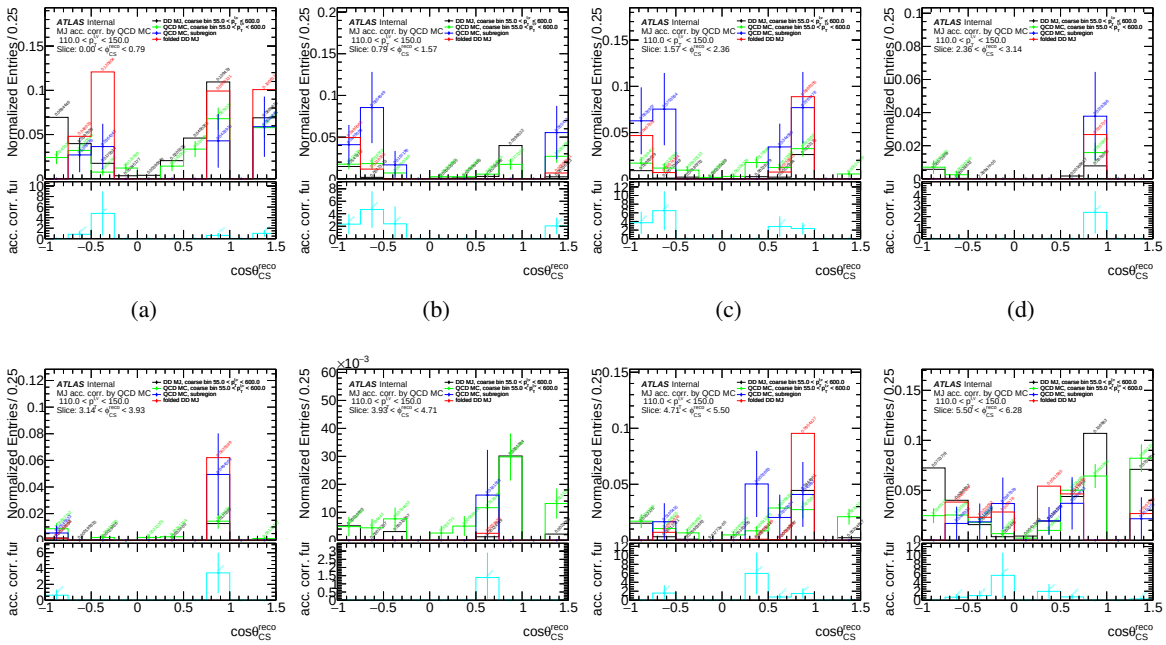


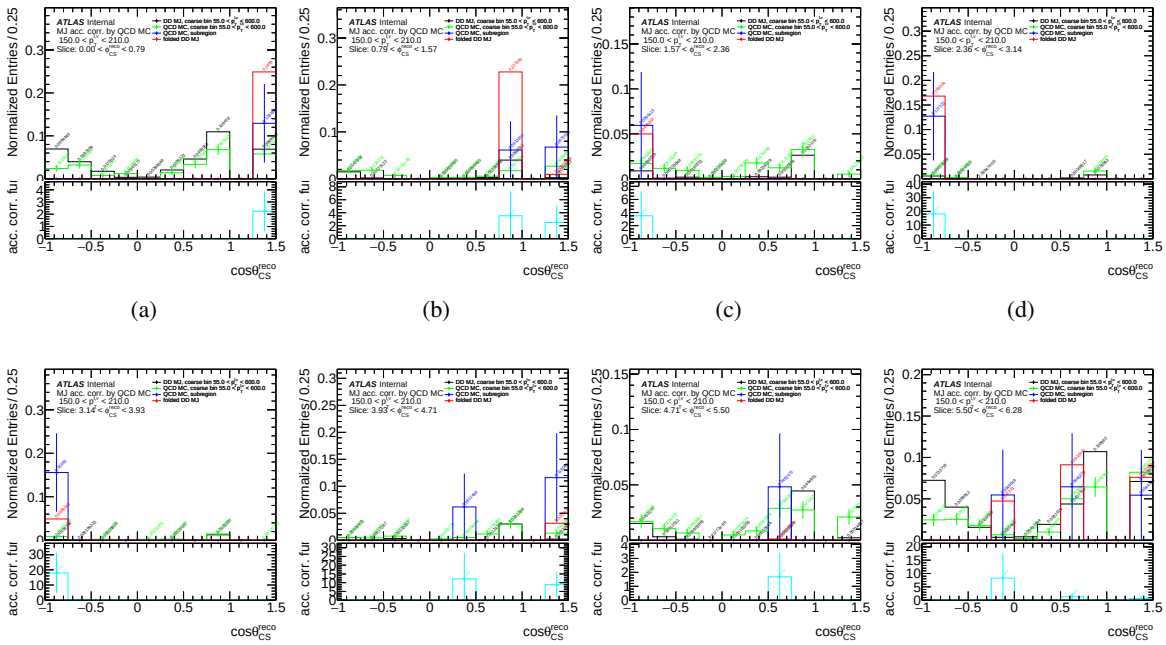
Figure 184: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 185: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 186: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

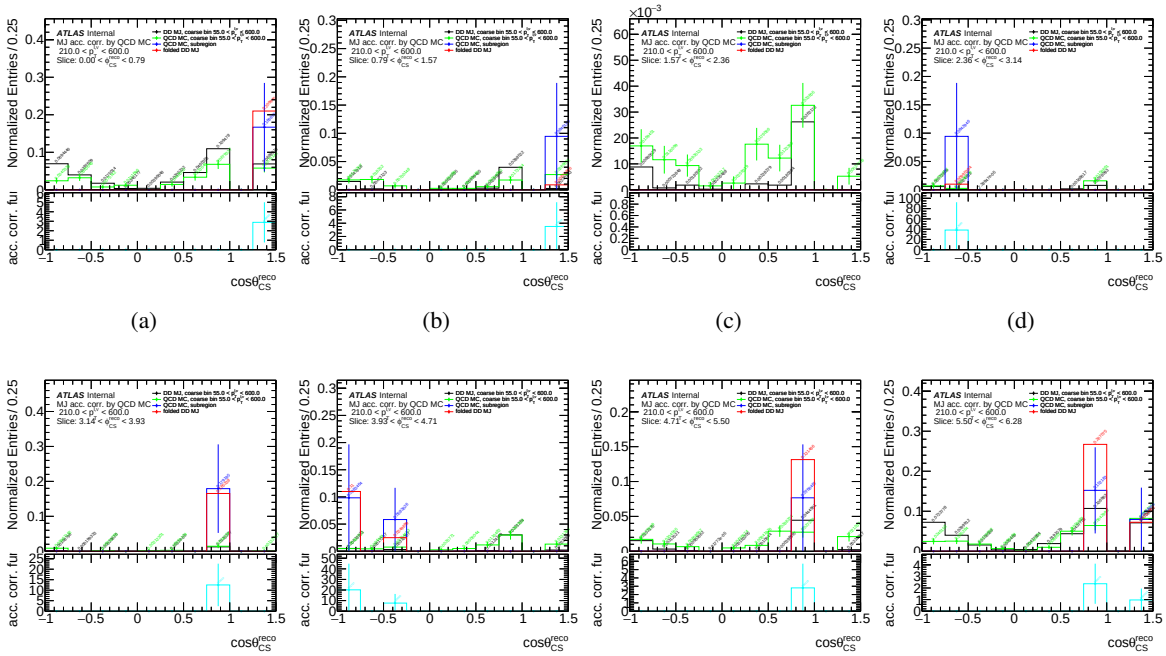


Figure 187: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

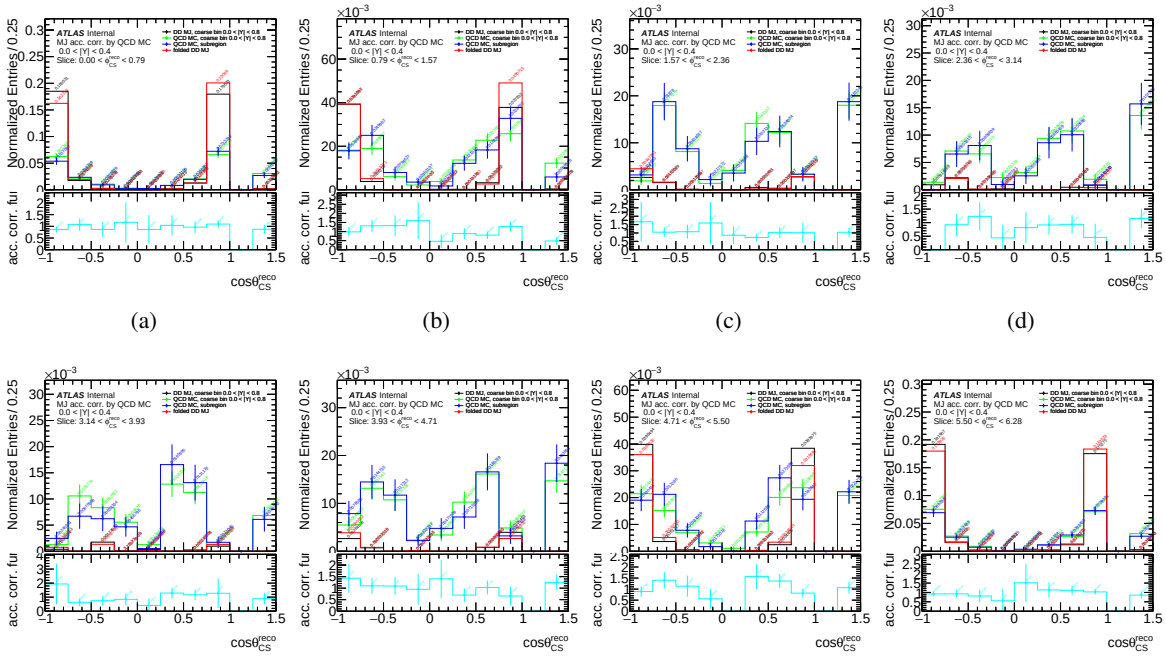
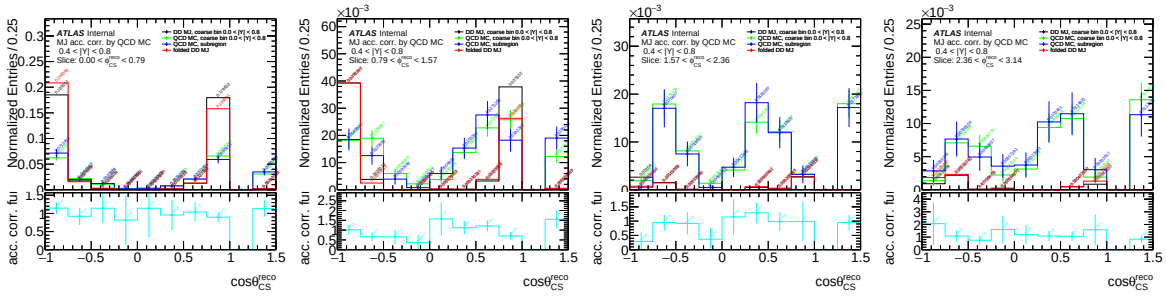
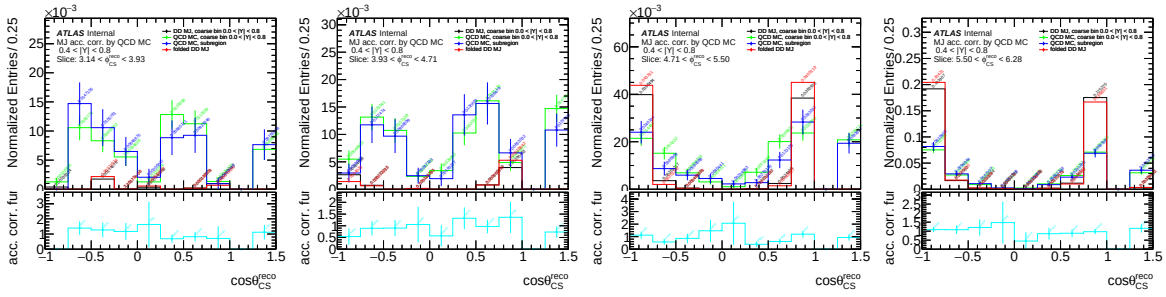


Figure 188: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

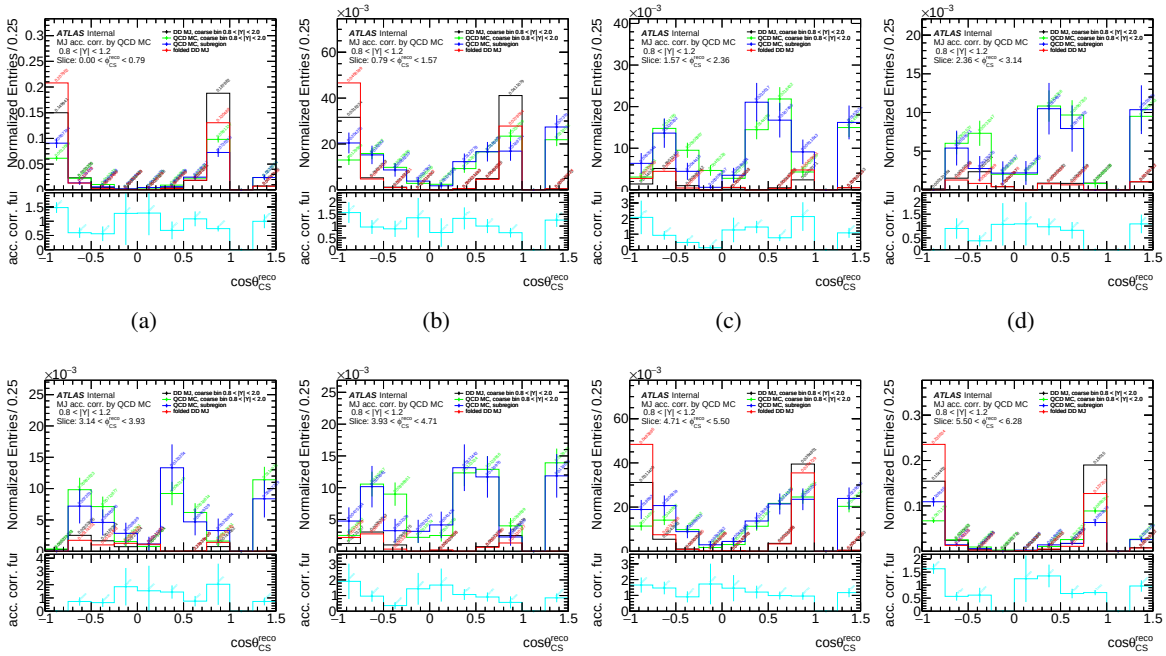


(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 189: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 190: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

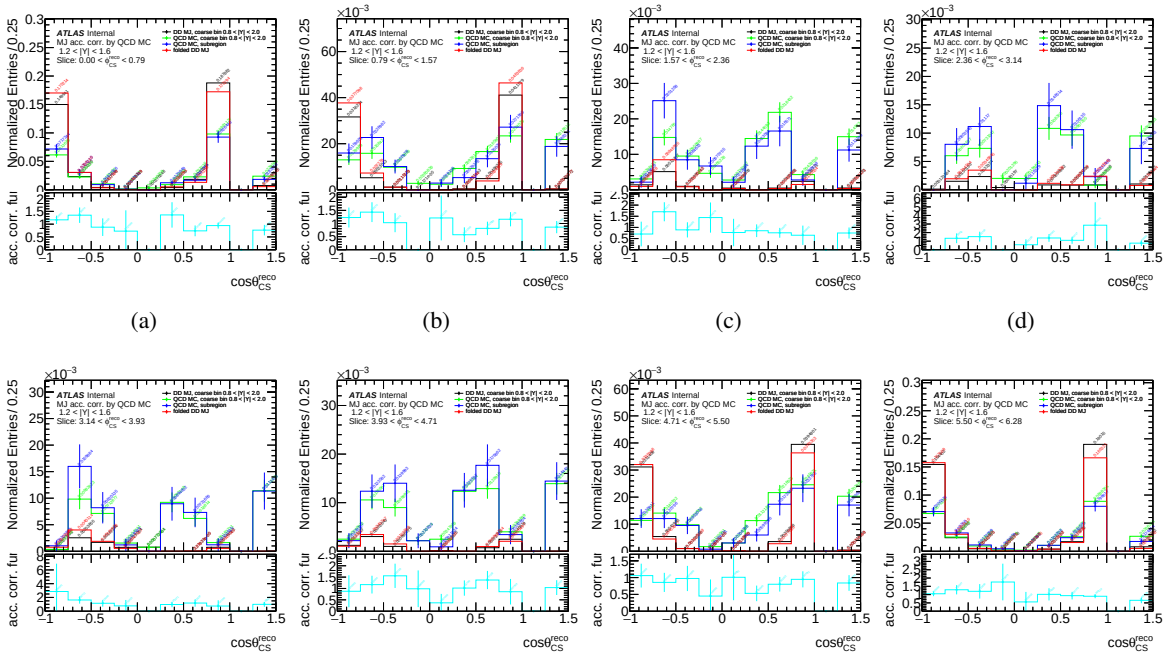


Figure 191: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

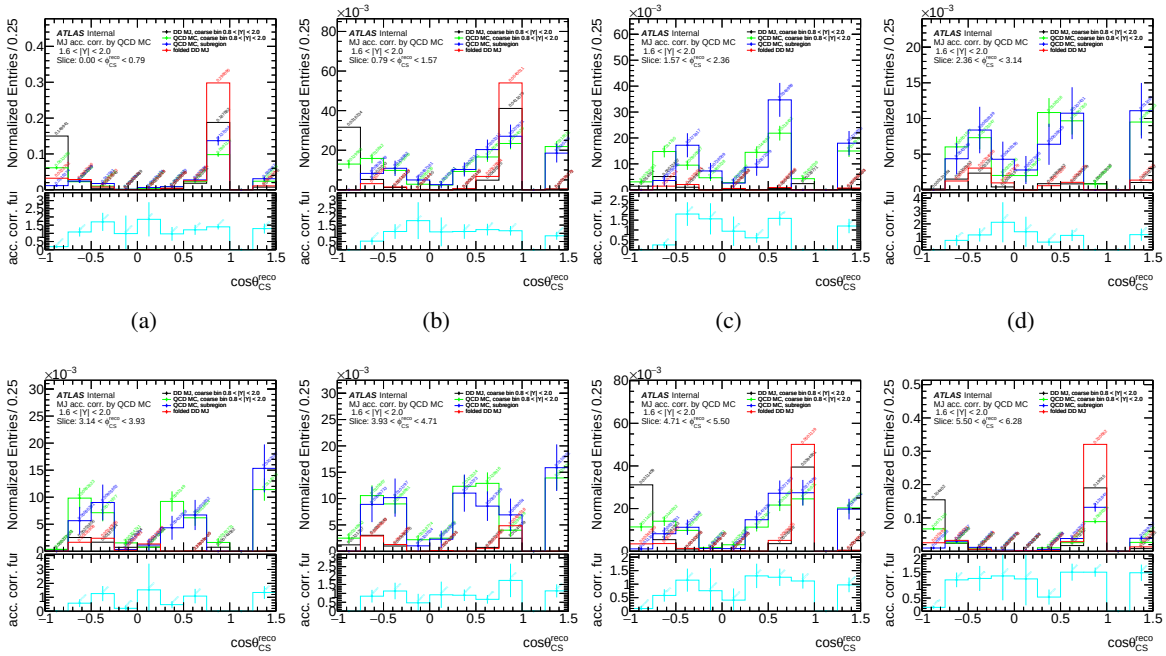
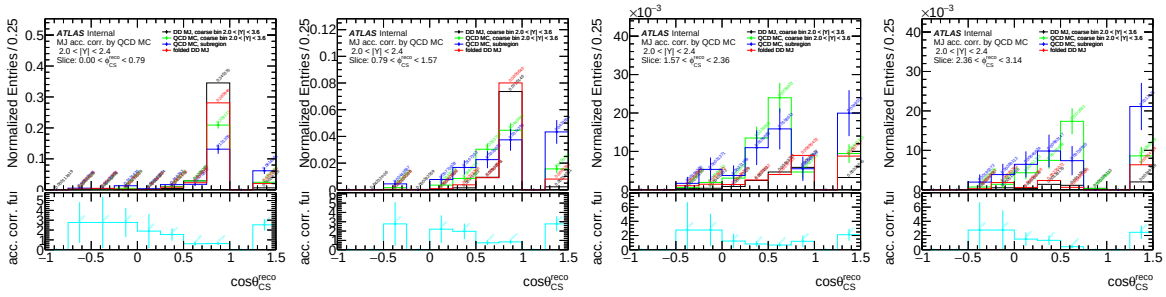


Figure 192: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 193: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

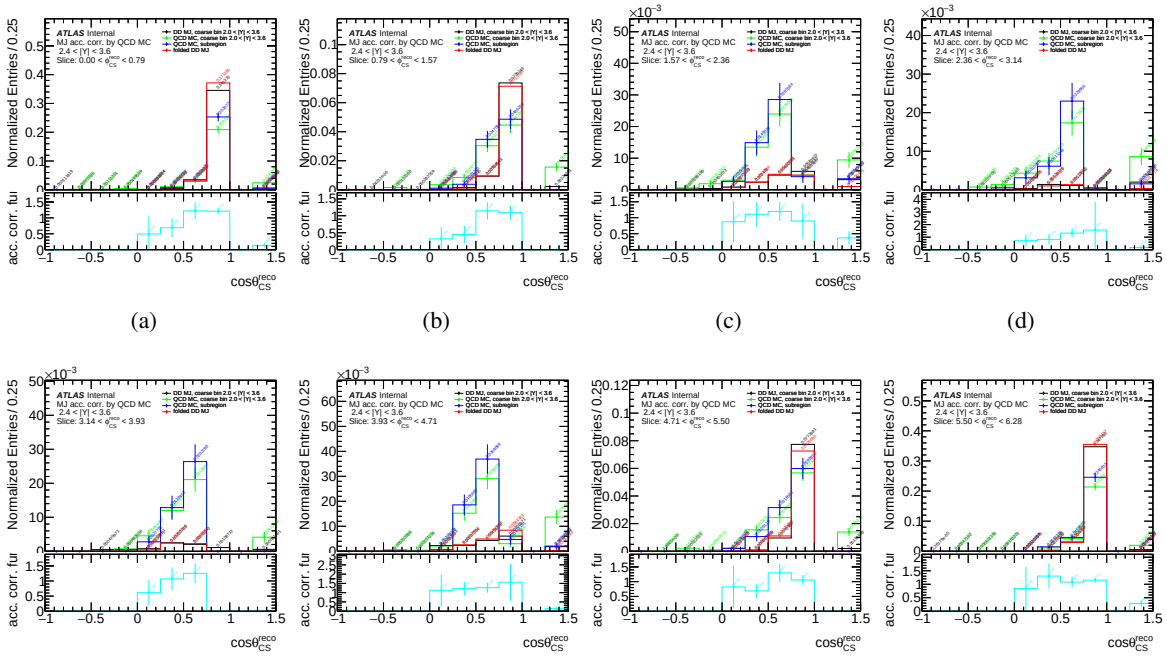


Figure 194: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^+ \to e^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

E.5.8 Muon channel multi-jet background 2D acceptance correction functions

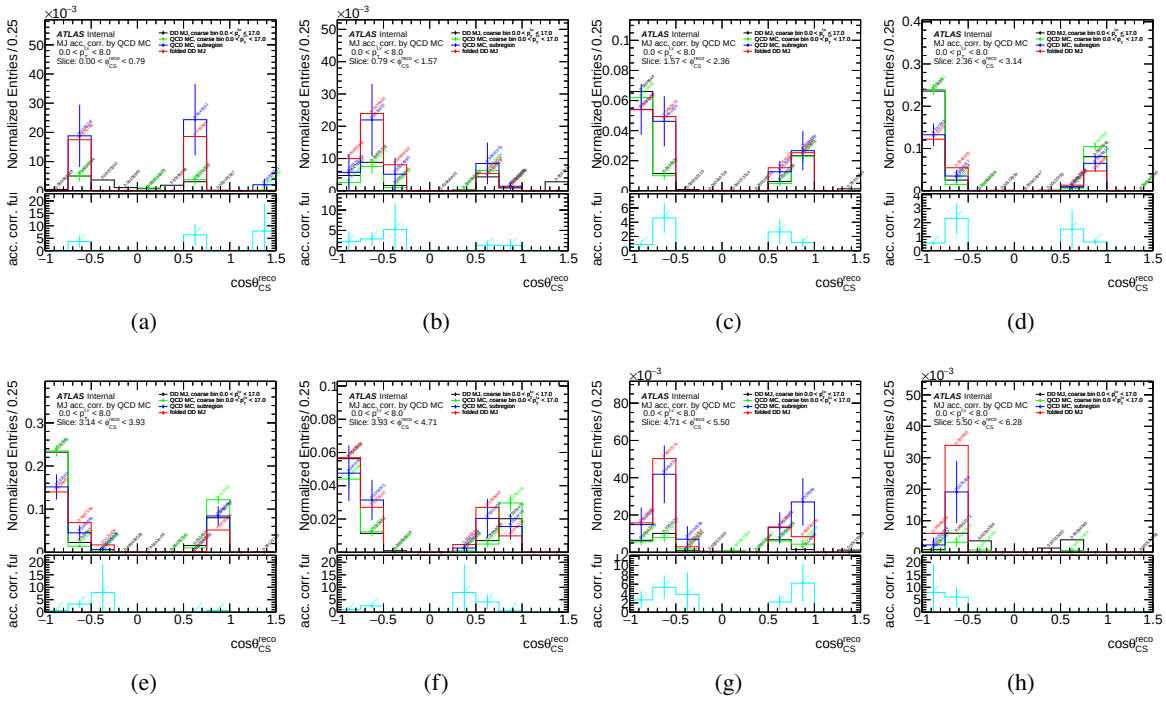
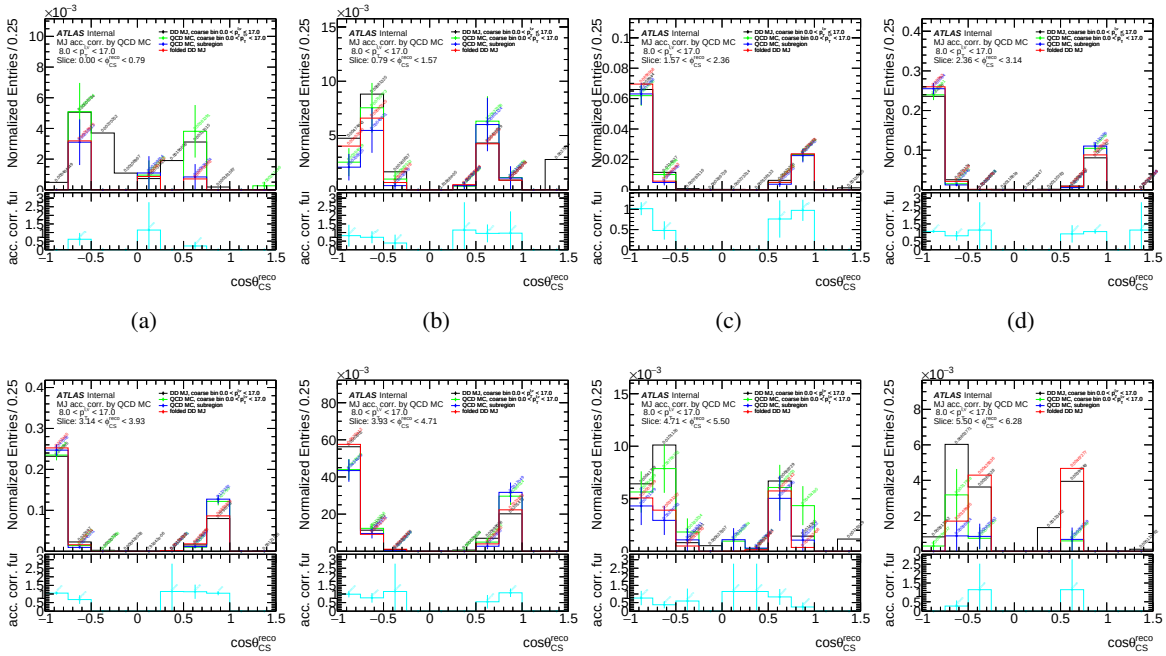


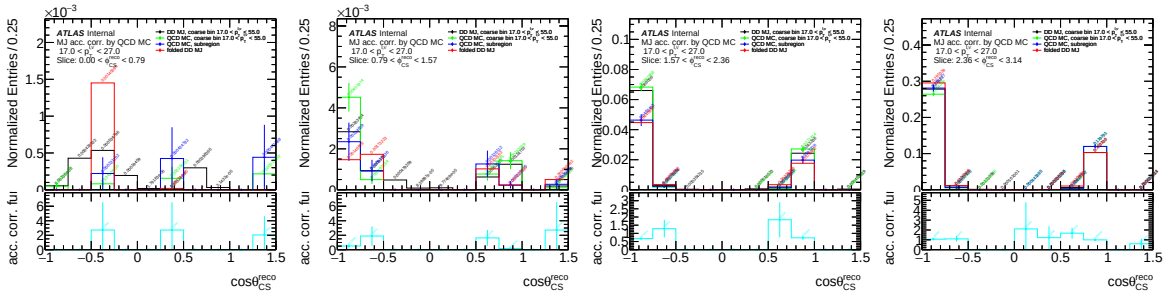
Figure 195: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



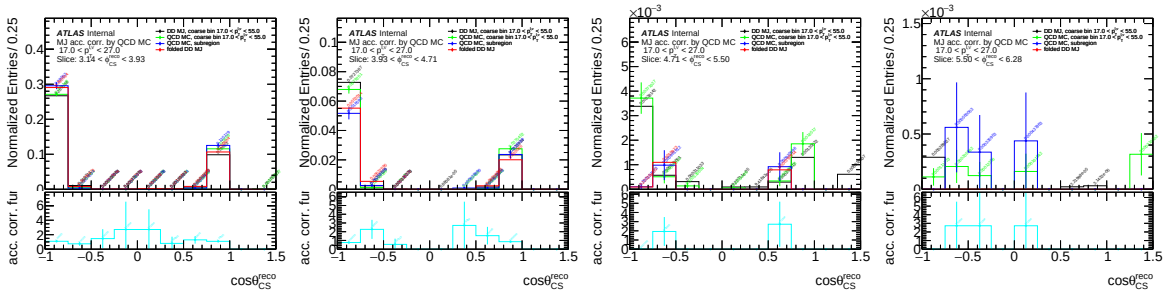
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 196: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 197: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

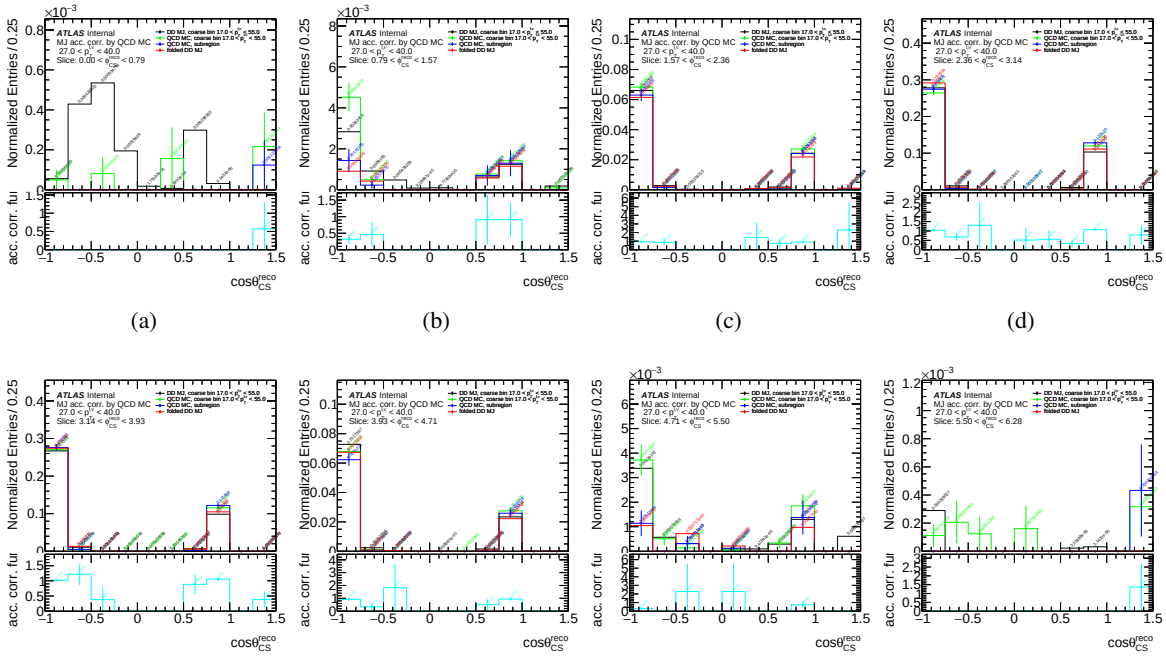


Figure 198: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

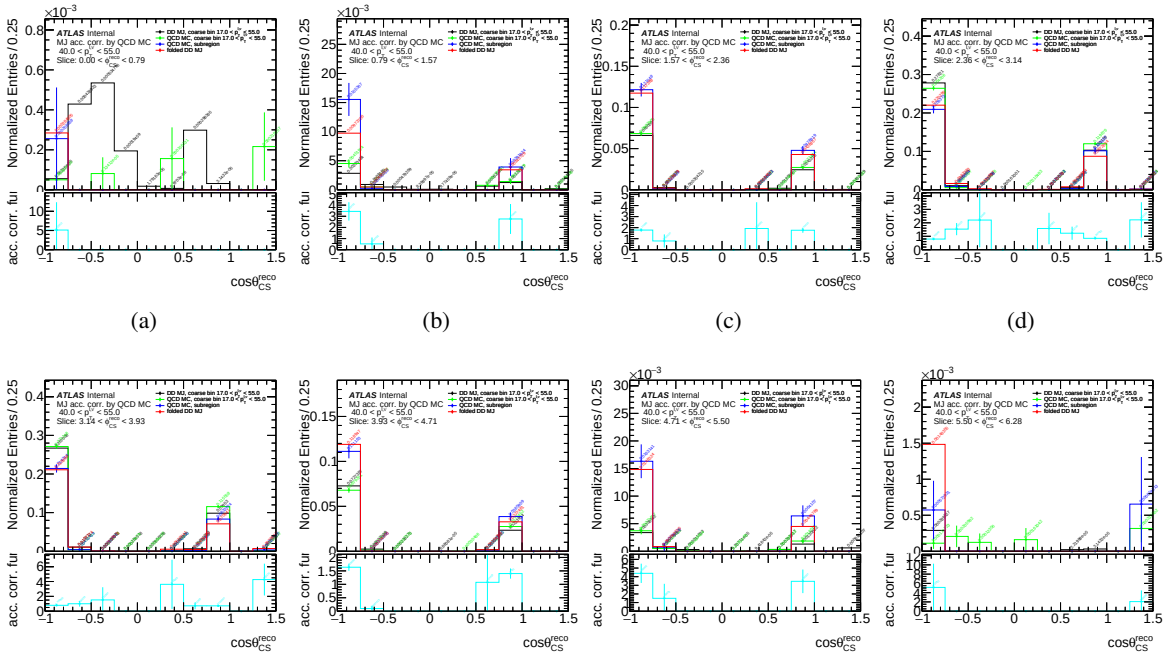
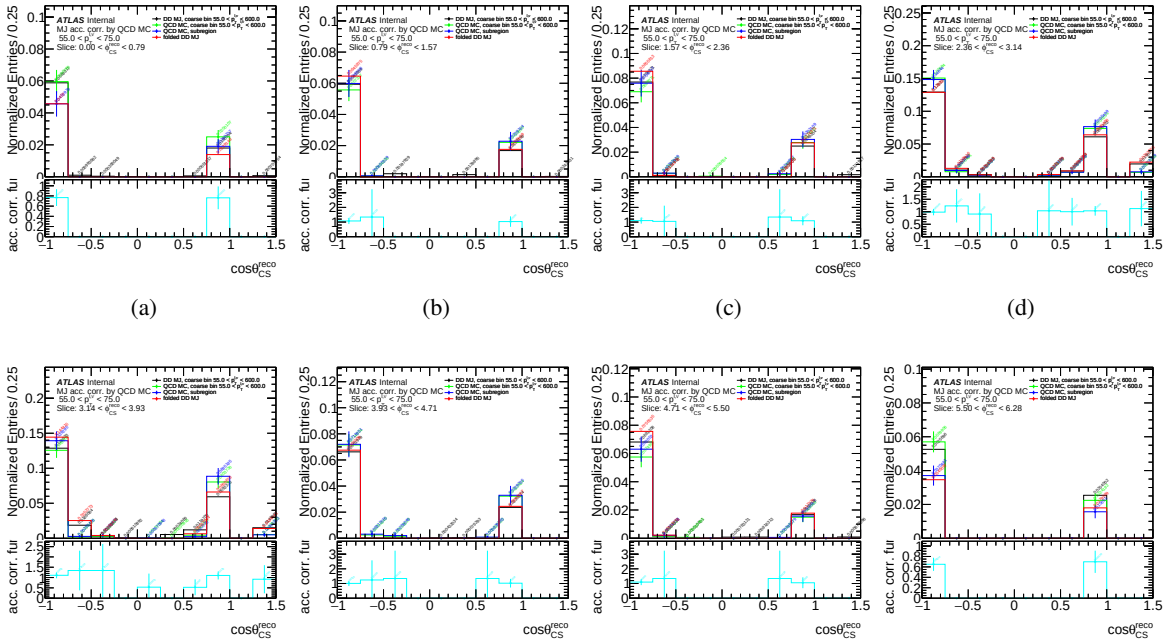


Figure 199: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 200: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

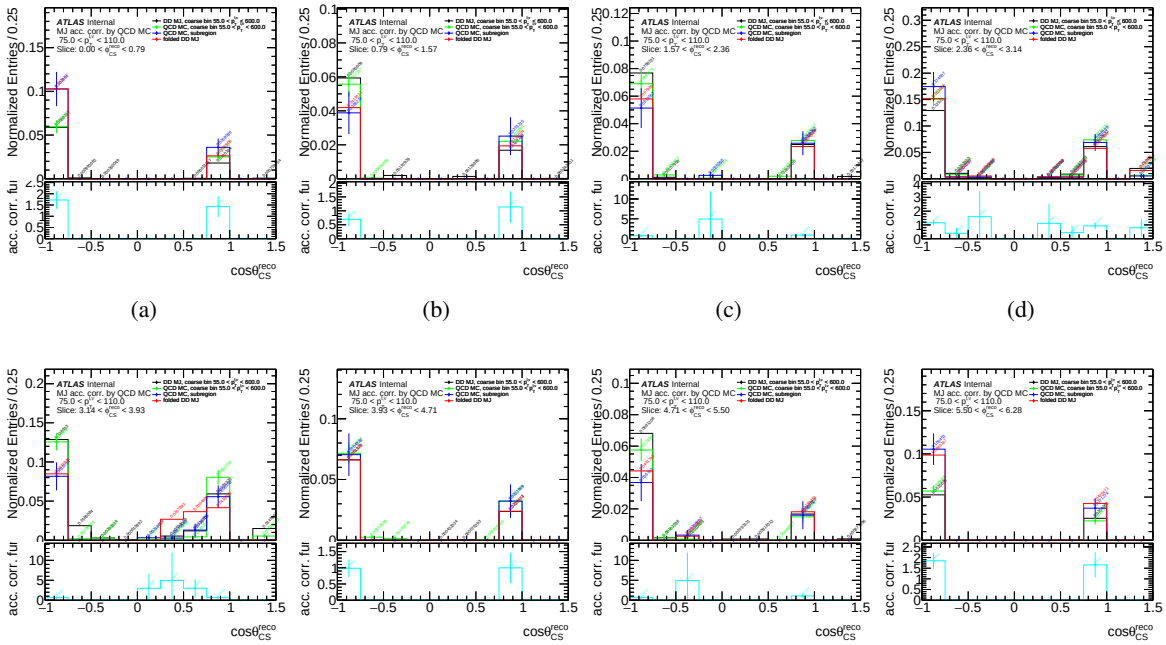


Figure 201: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

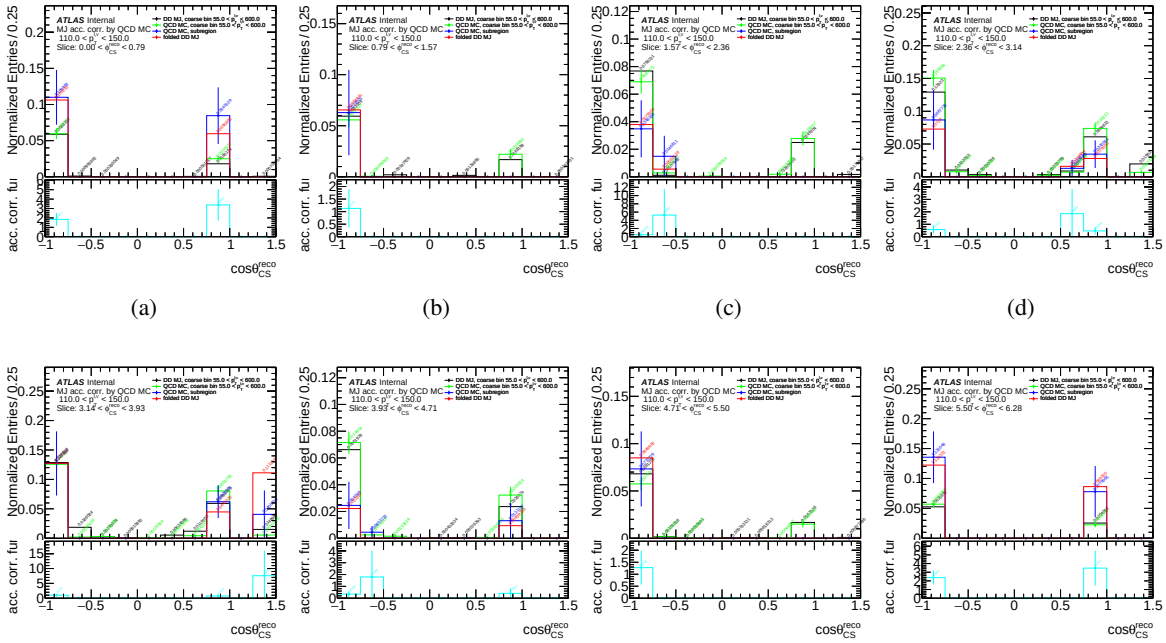
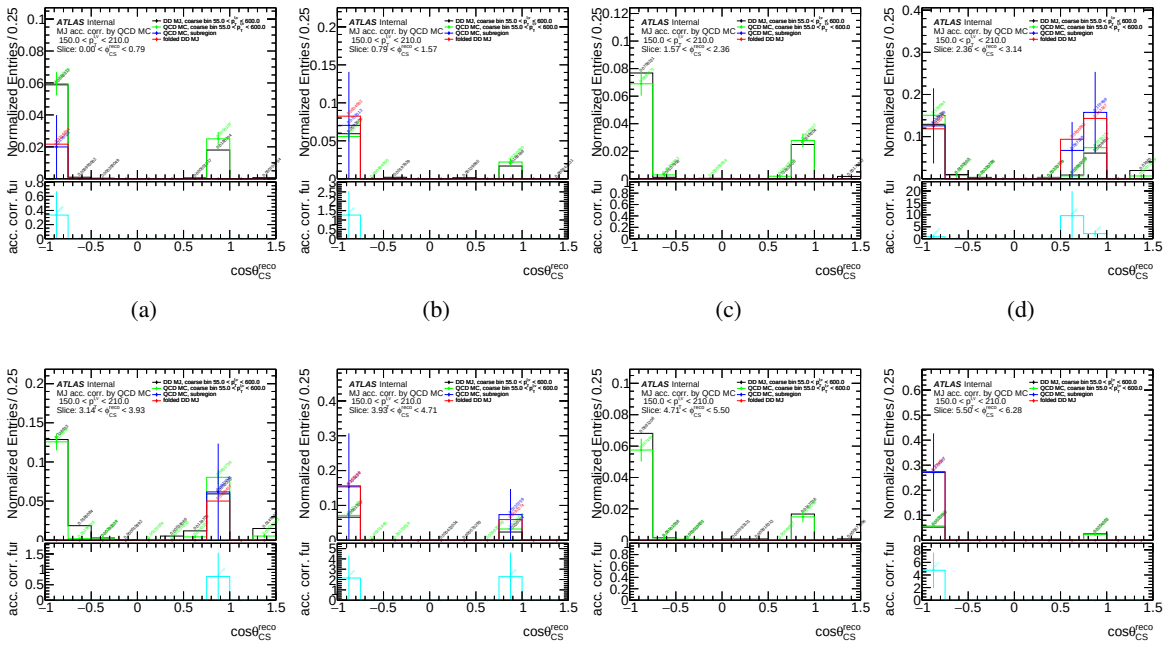


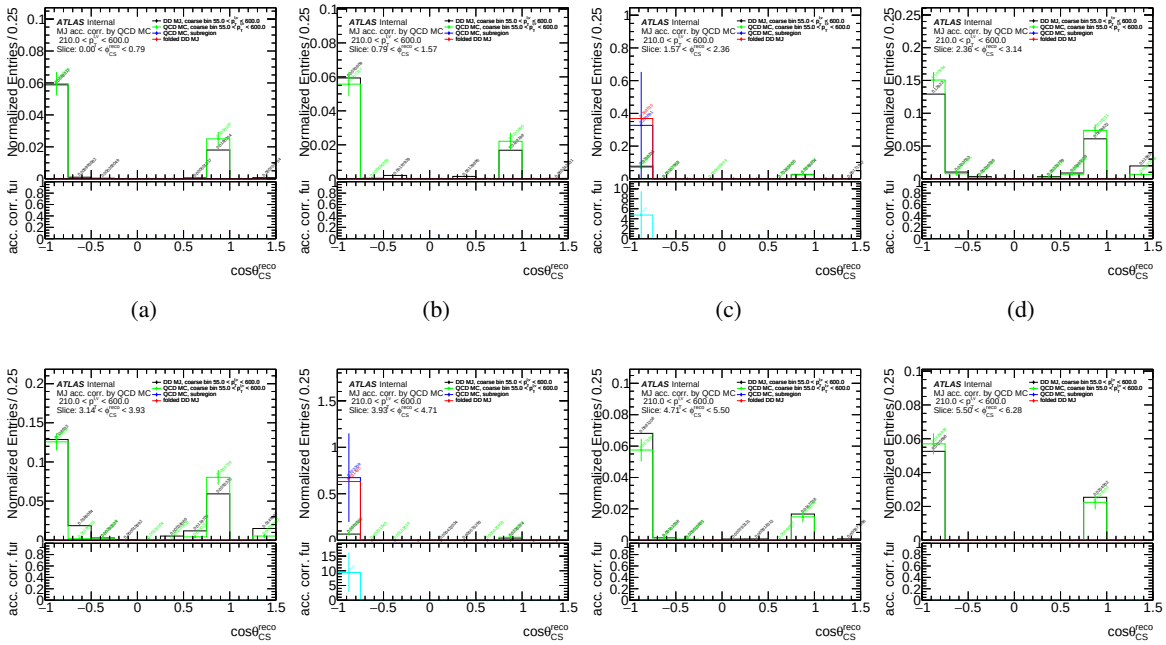
Figure 202: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 203: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 204: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

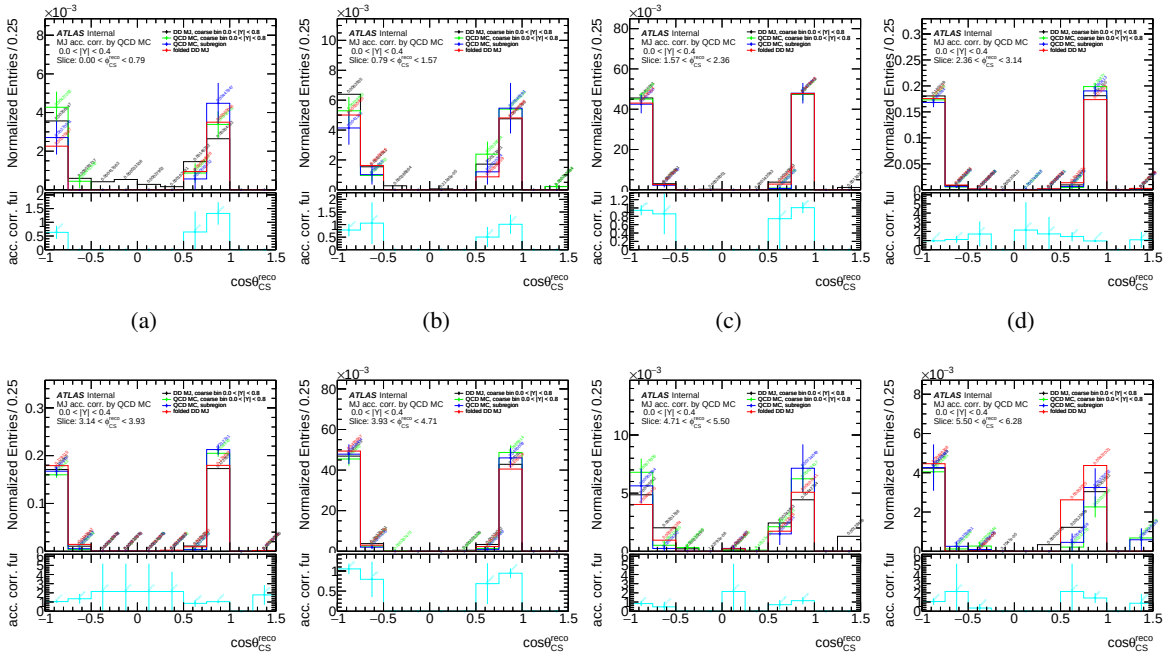


Figure 205: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

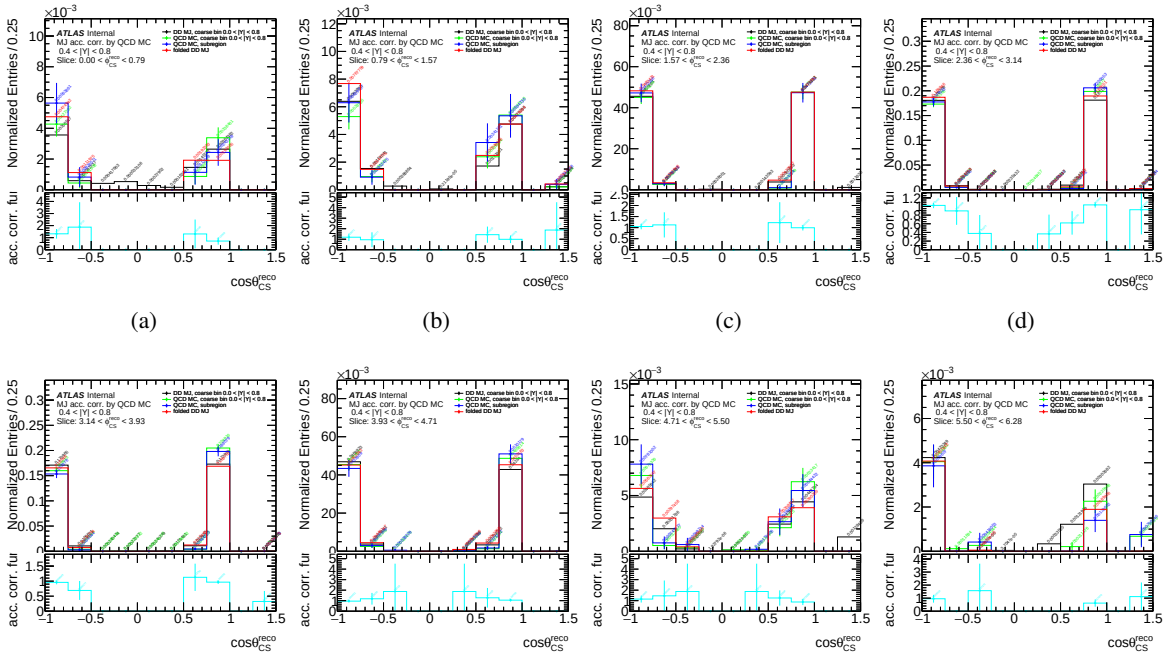
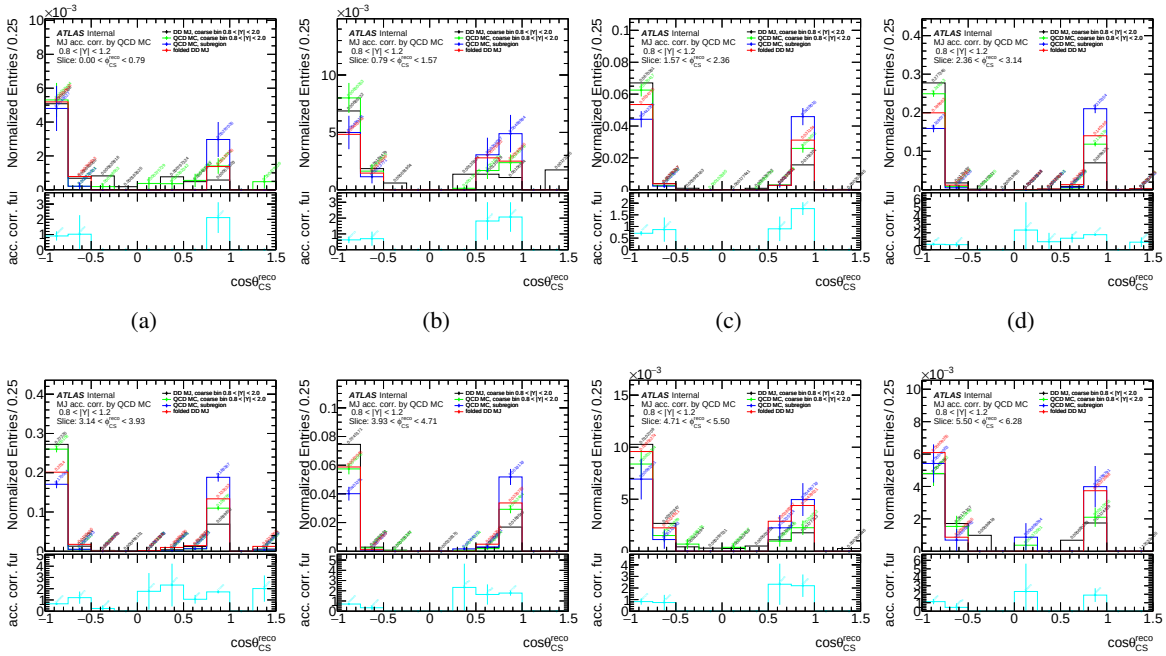


Figure 206: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 207: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

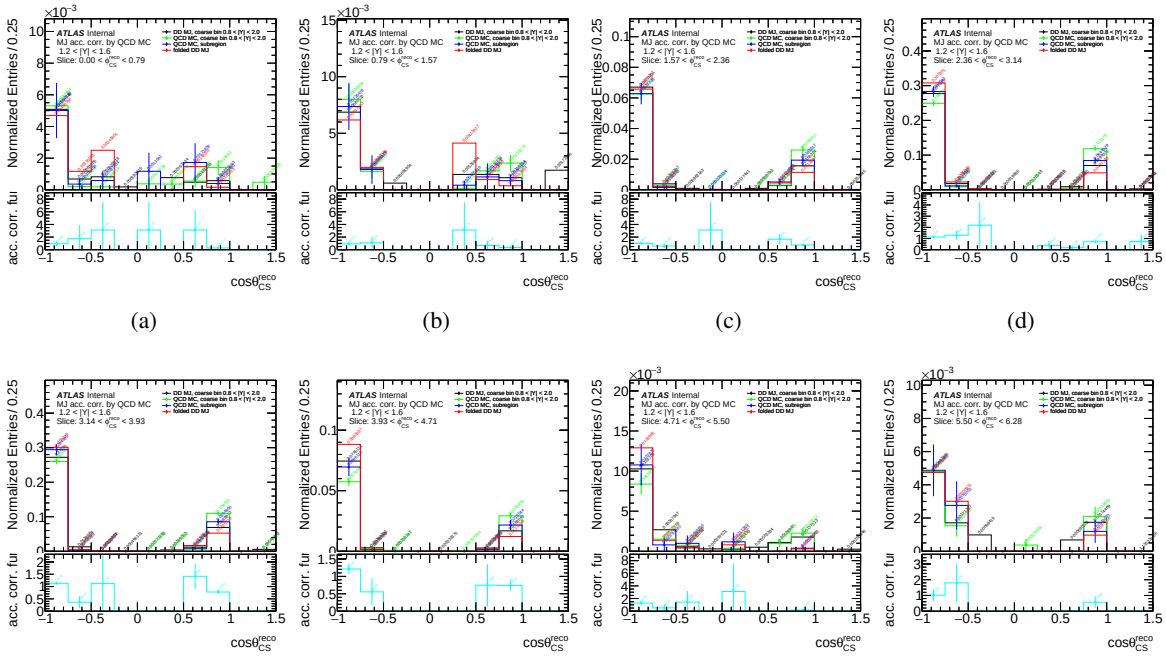
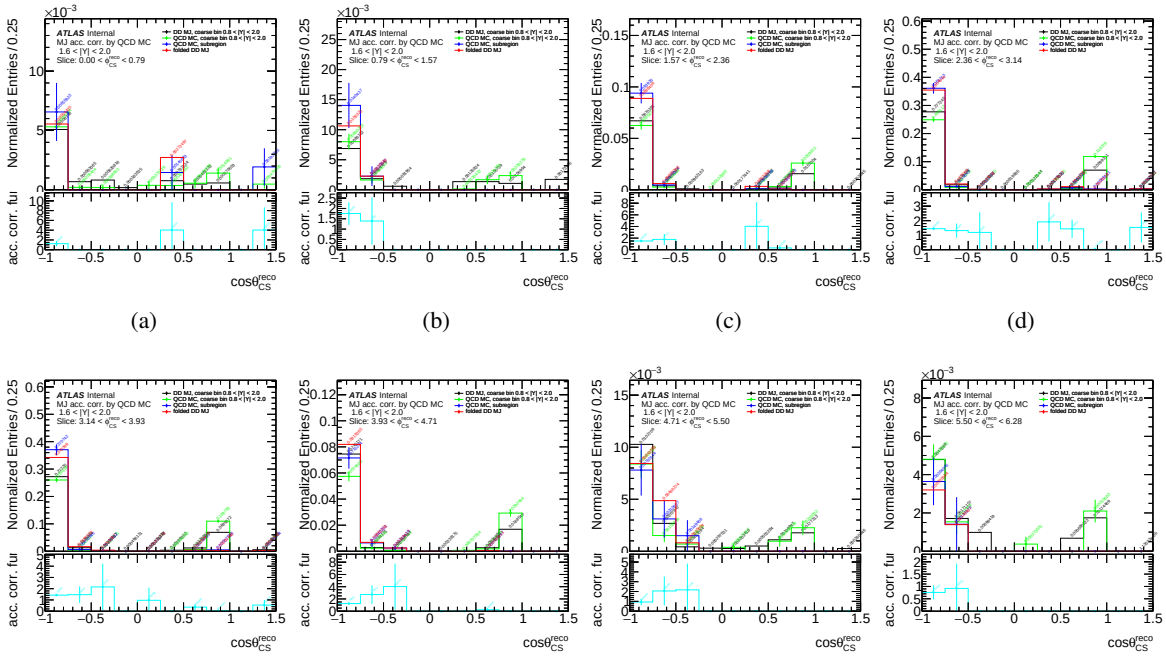


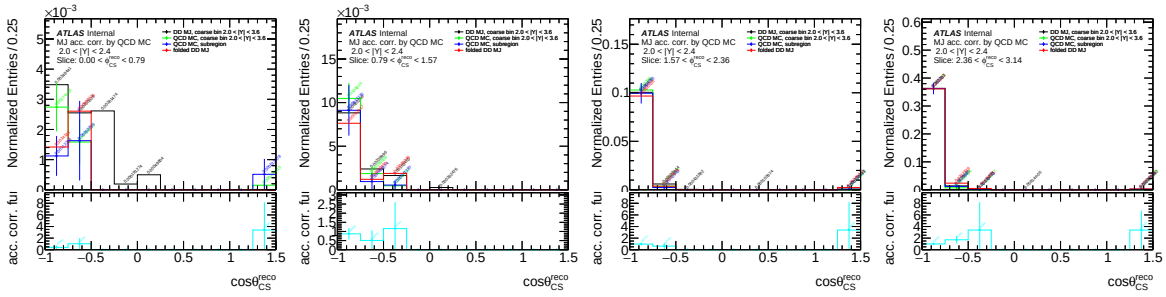
Figure 208: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^- \rightarrow \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



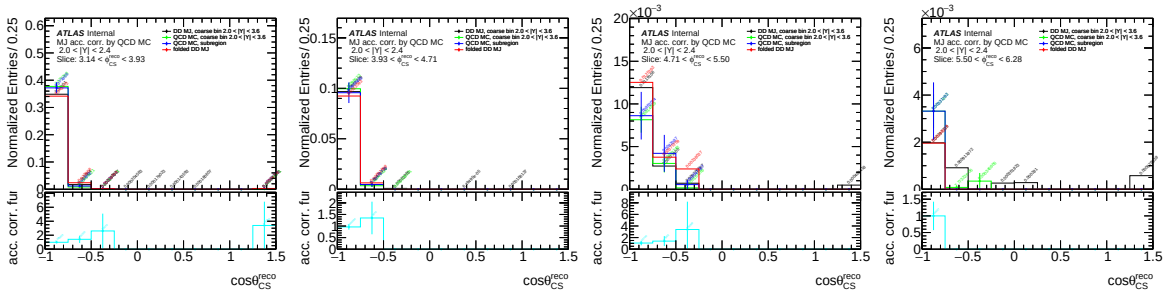
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 209: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 210: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^- \to \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

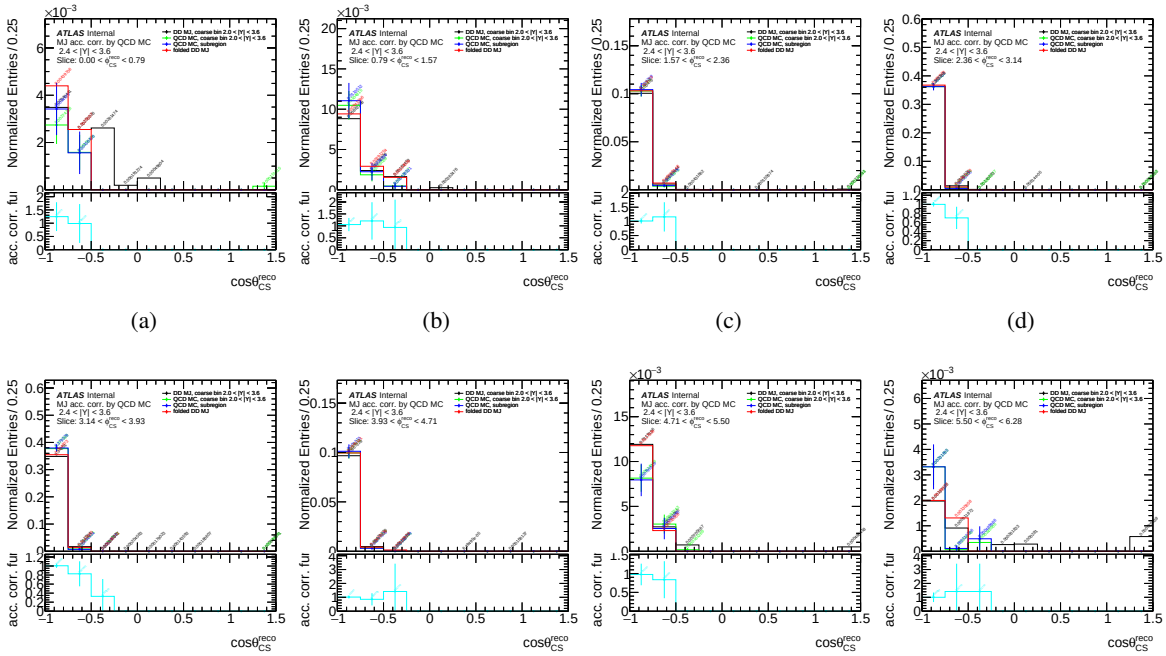


Figure 211: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^- \rightarrow \mu^-\overline{\nu}$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

E.5.9 Anti-muon channel multi-jet background 2D acceptance correction functions

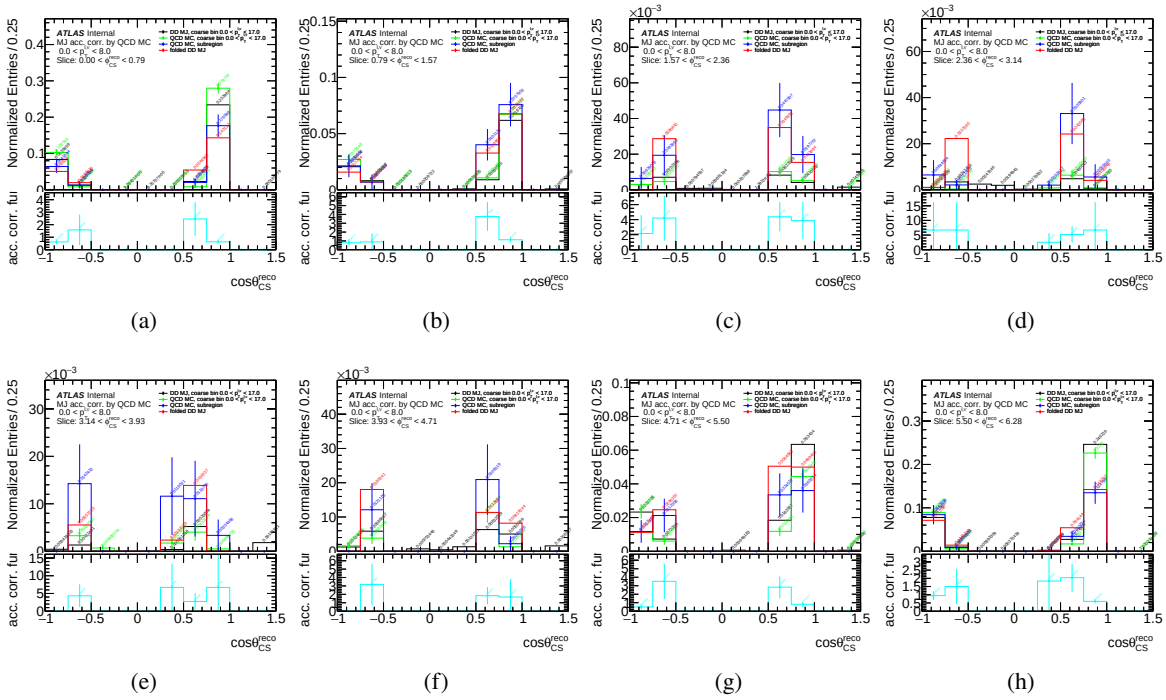


Figure 212: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

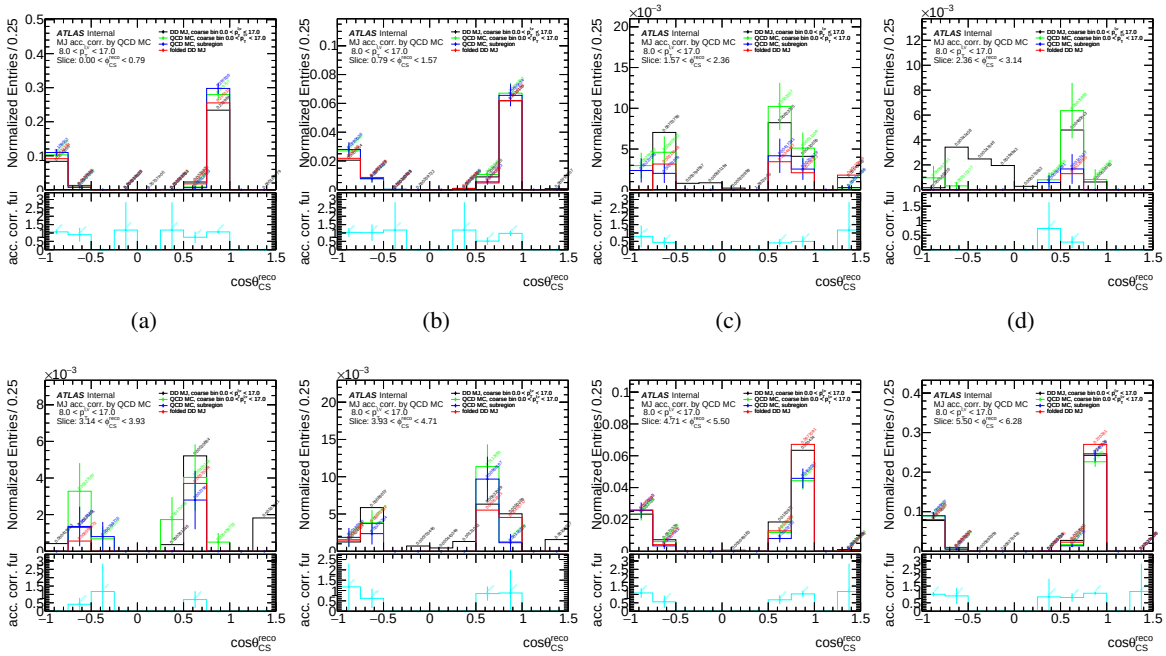
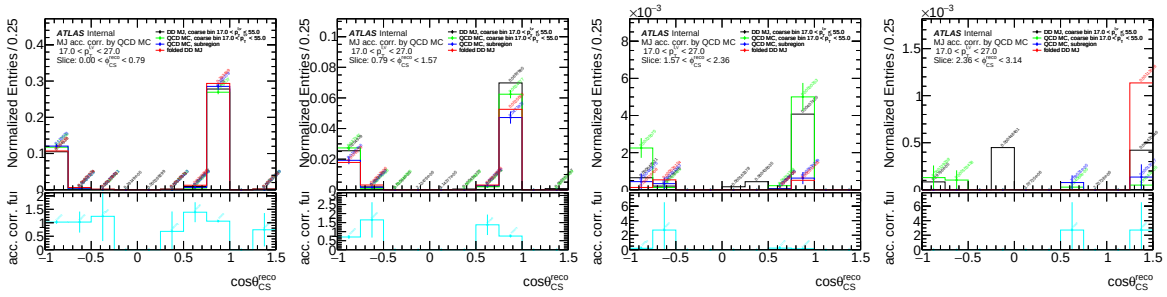
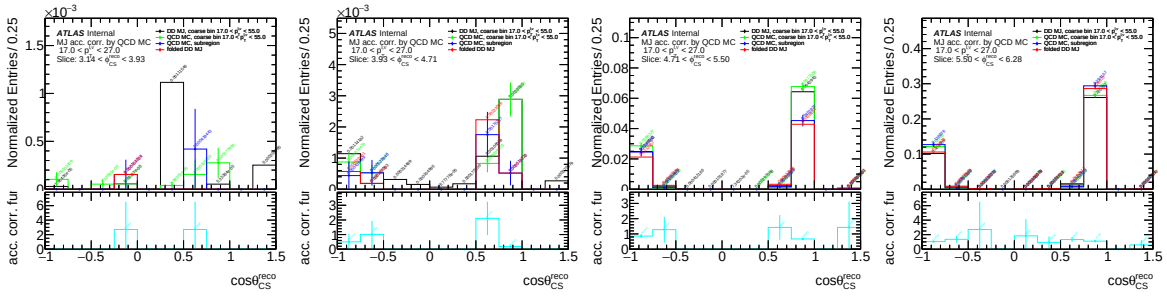


Figure 213: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 214: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

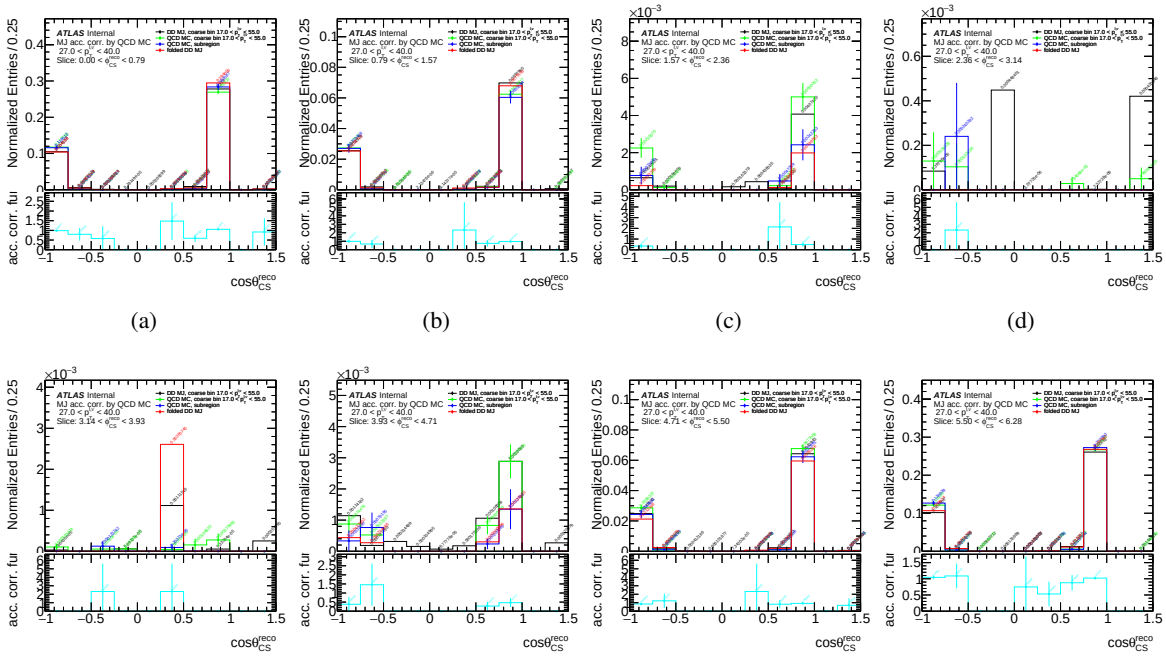


Figure 215: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

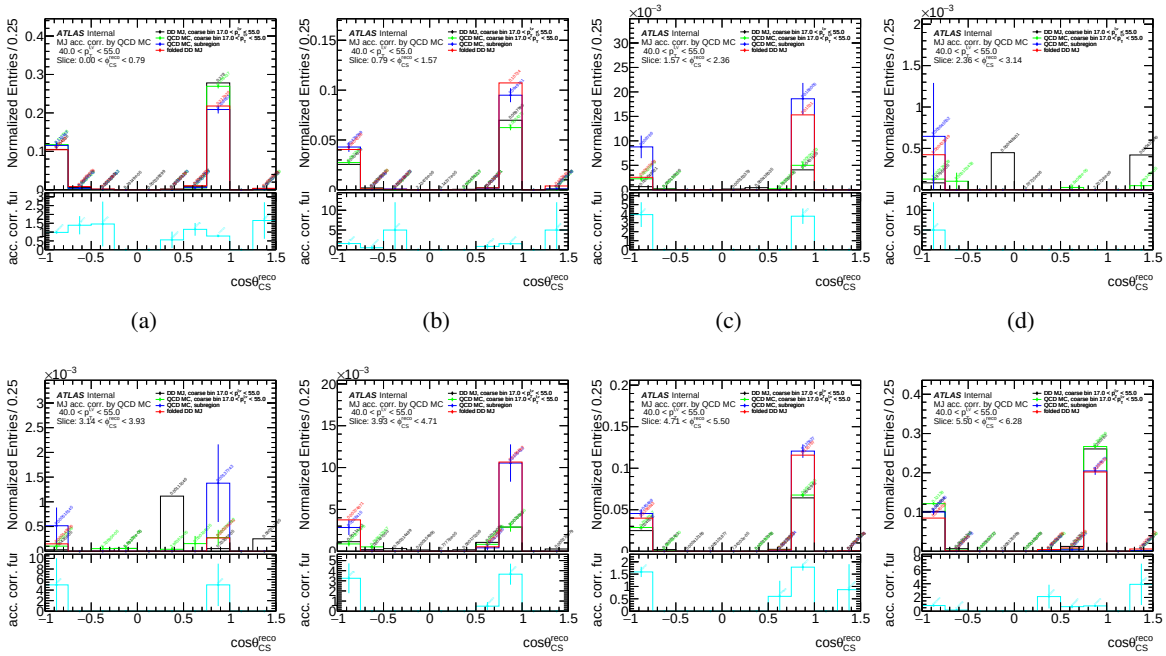
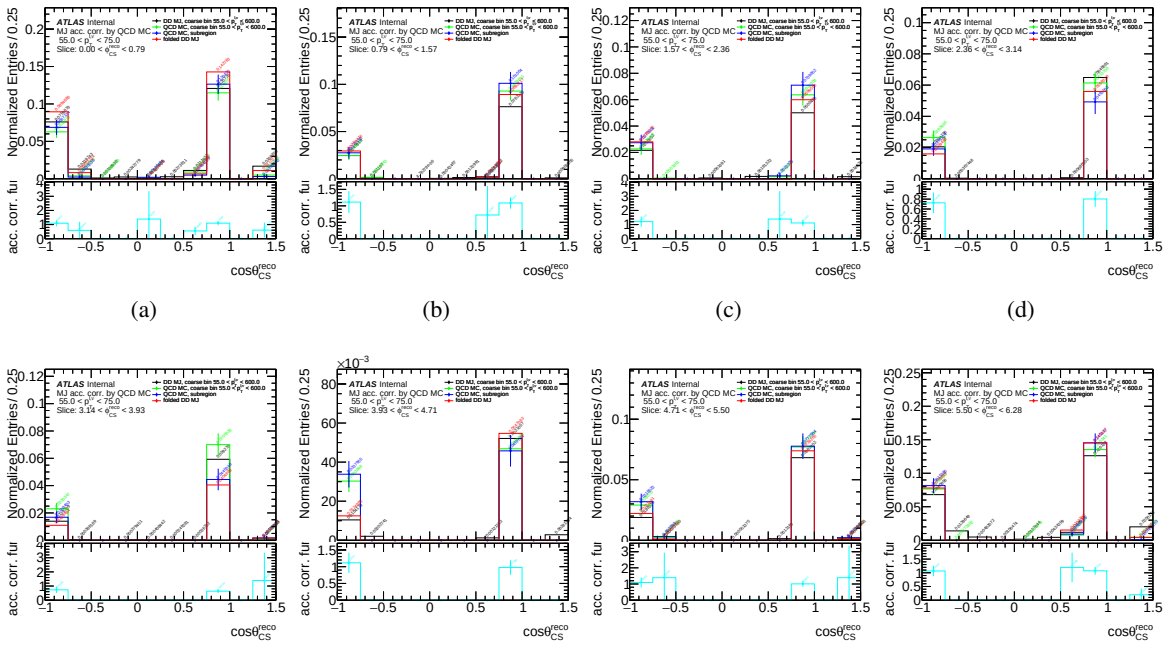


Figure 216: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 217: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

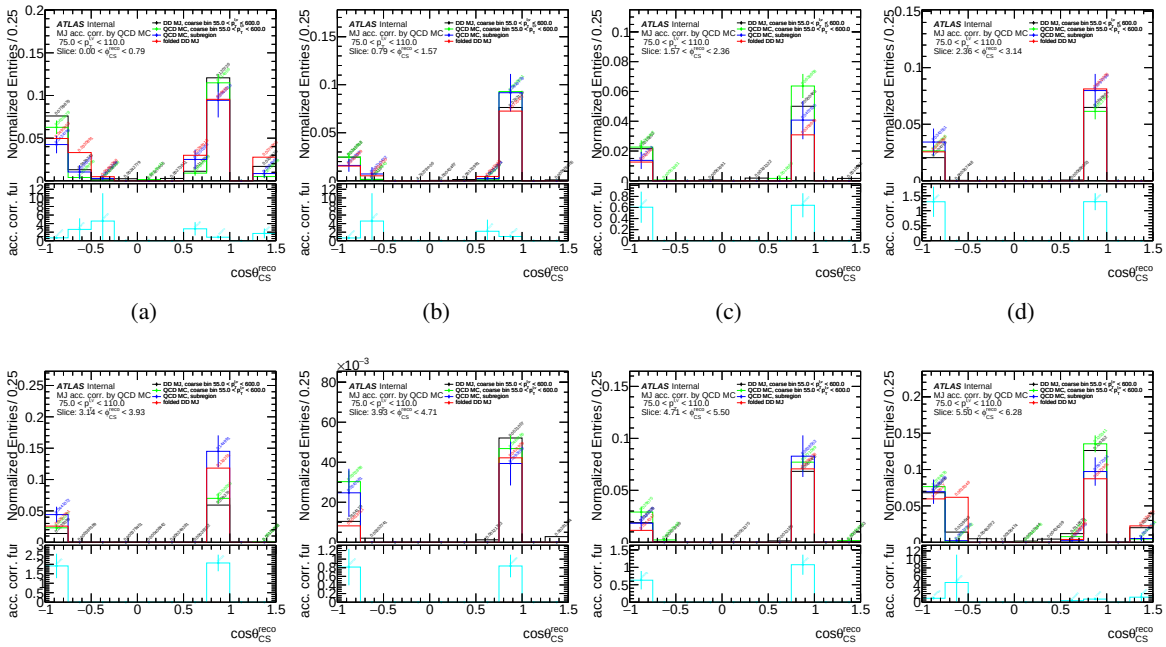


Figure 218: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

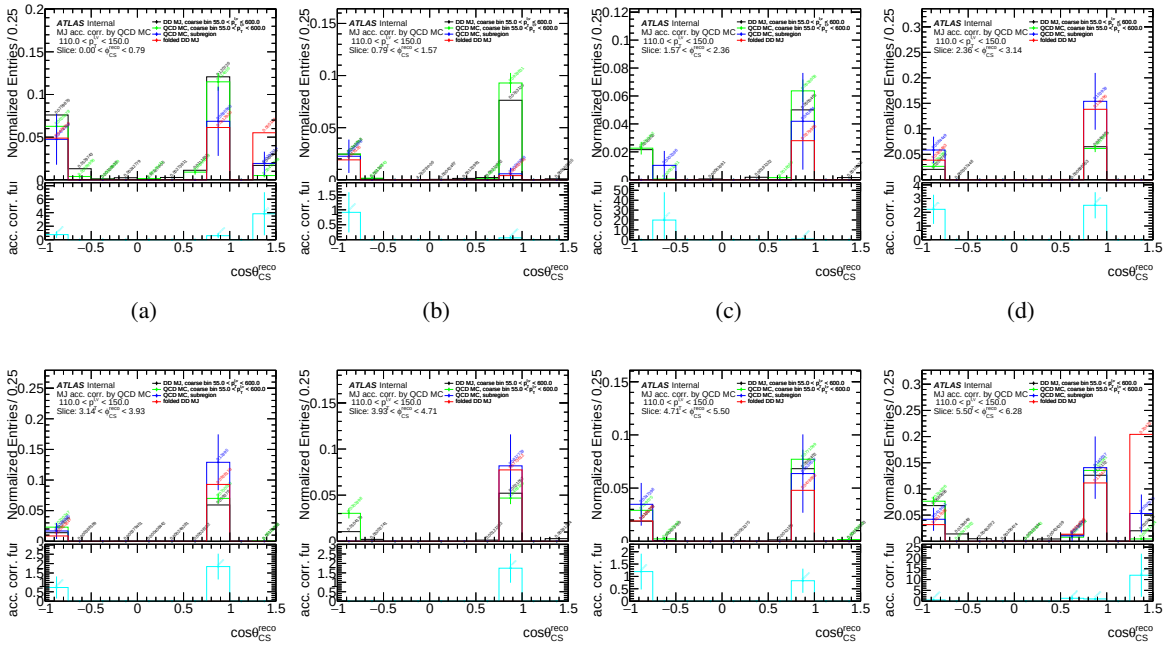


Figure 219: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

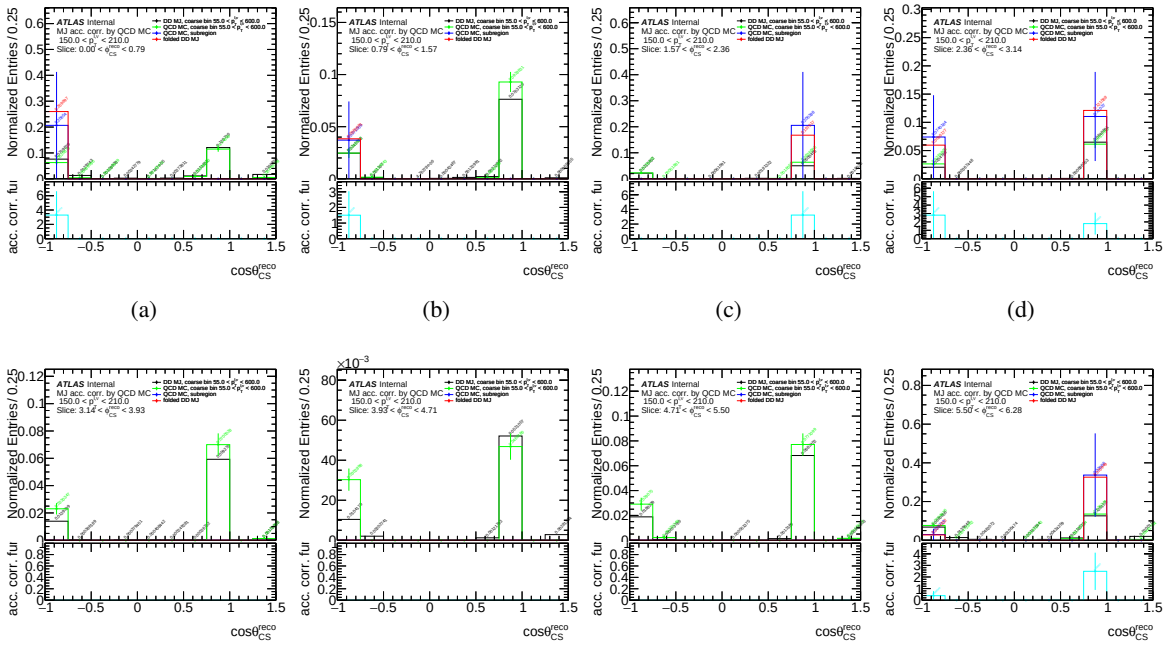


Figure 220: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

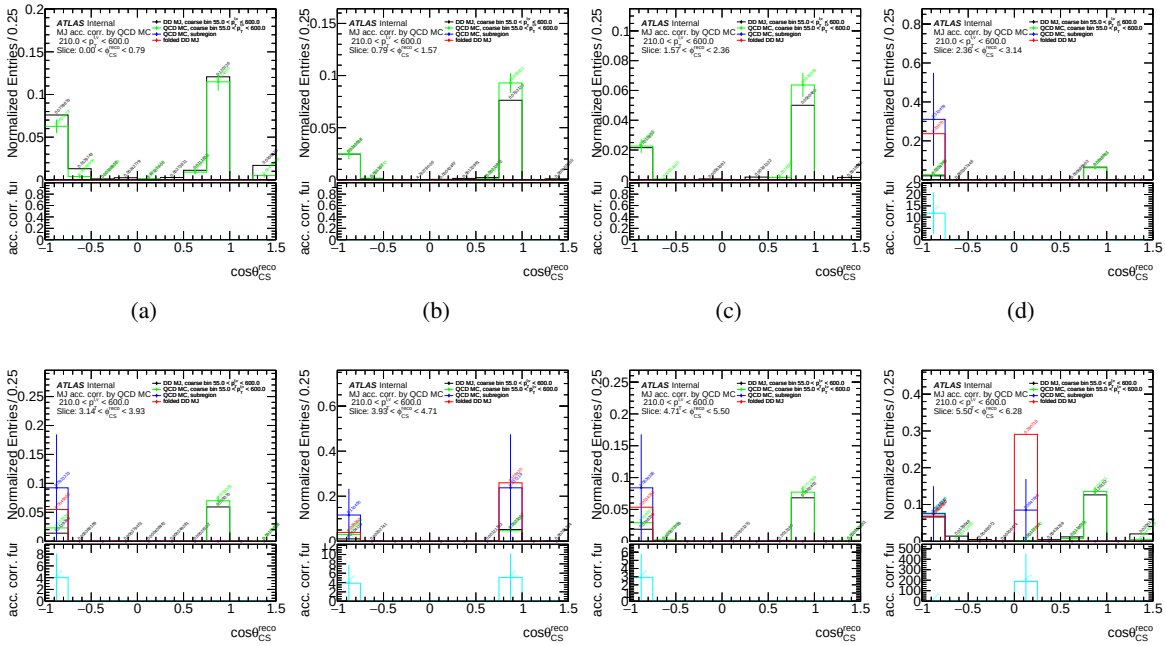
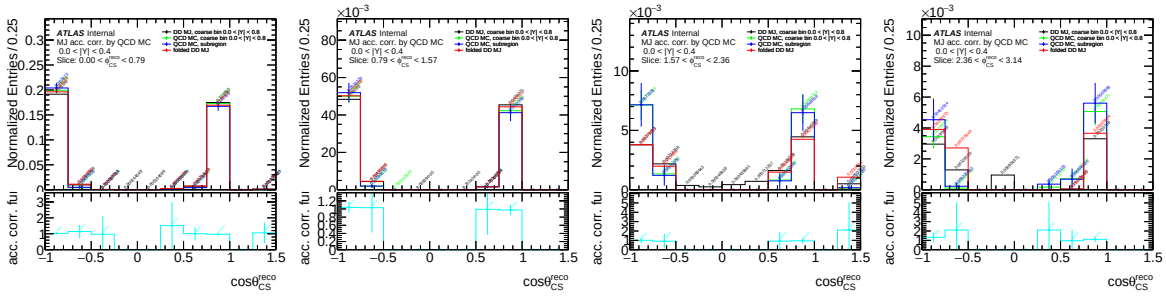
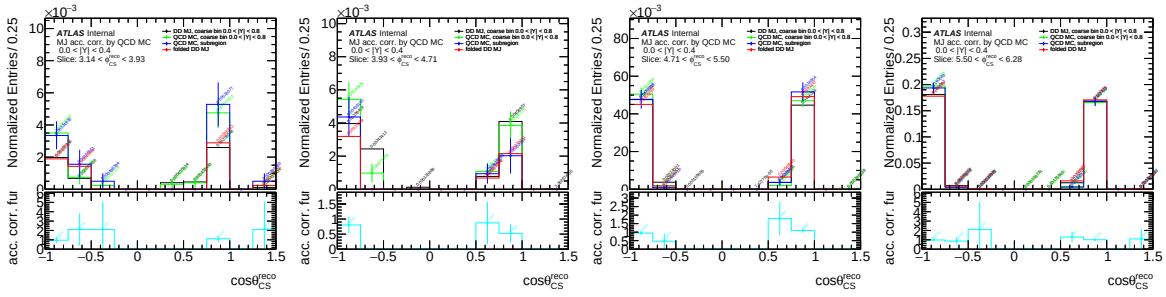


Figure 221: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

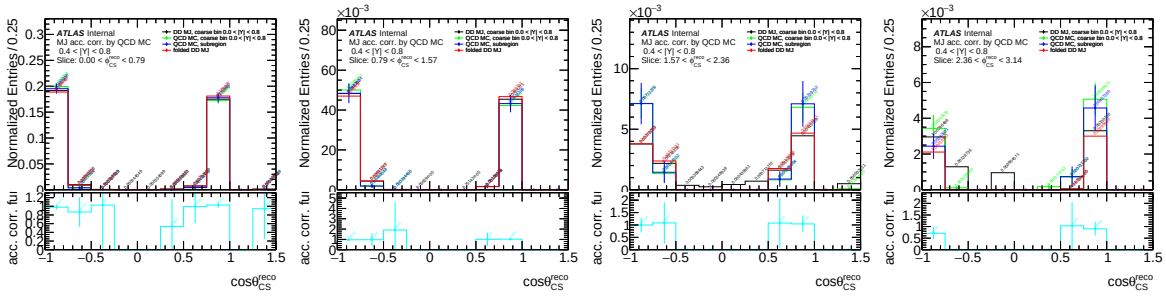


(a) (b) (c) (d)

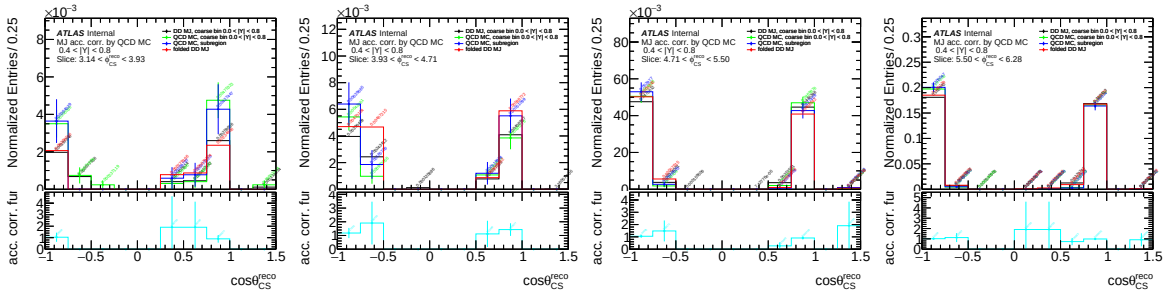


(e) (f) (g) (h)

Figure 222: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

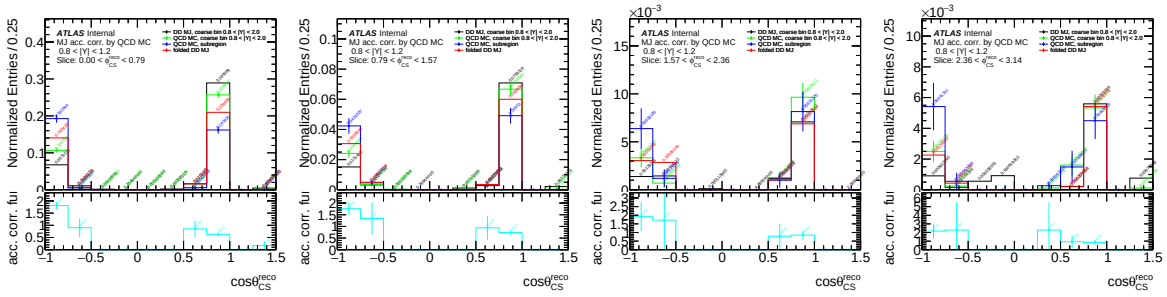


(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 223: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 224: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

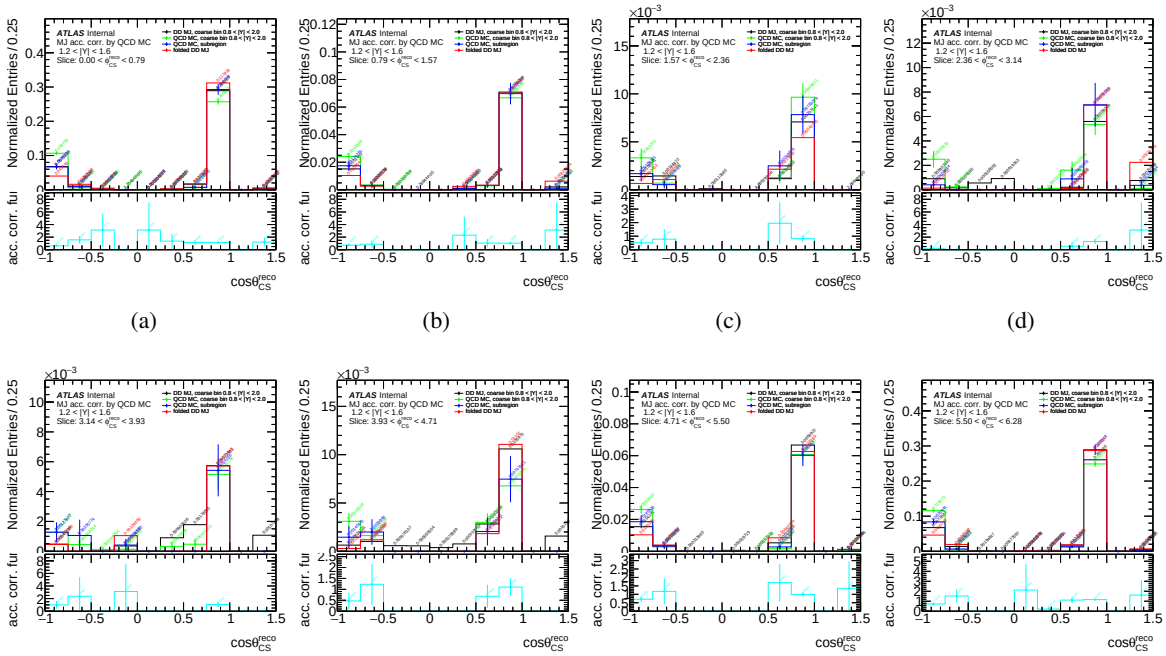


Figure 225: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

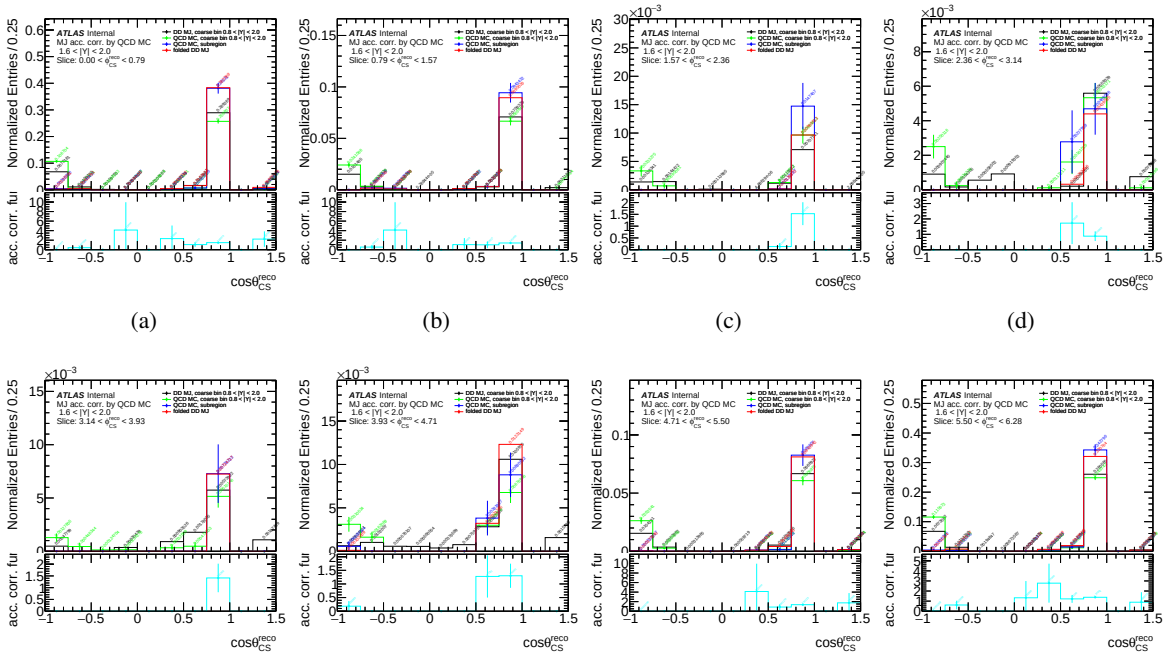
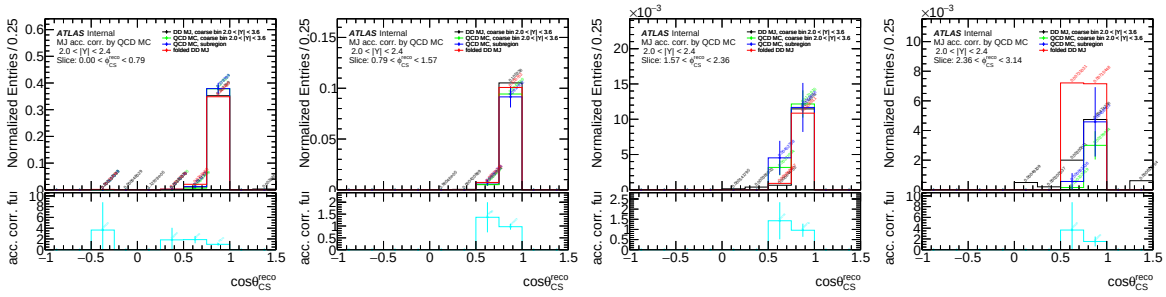
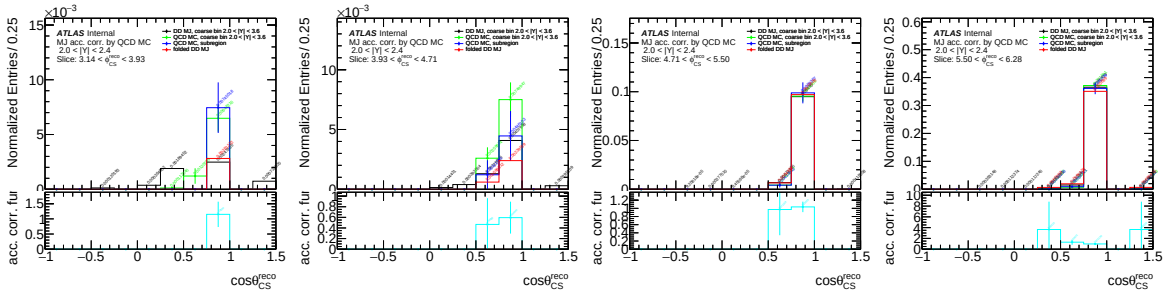


Figure 226: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 227: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

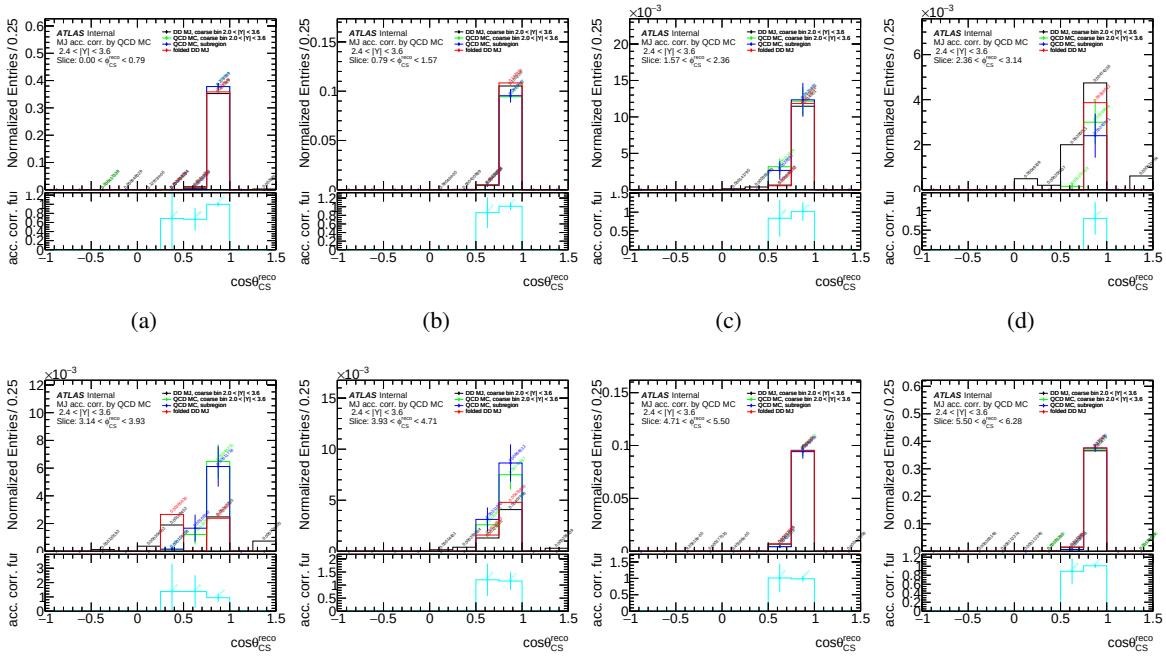


Figure 228: Multi-jet background template acceptance correction functions for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^+ \to \mu^+\nu$ channel. Distributions are normalized over ϕ_{CS} slices. Error bands represents statistical uncertainties.

E.6 MJ background 2D templates

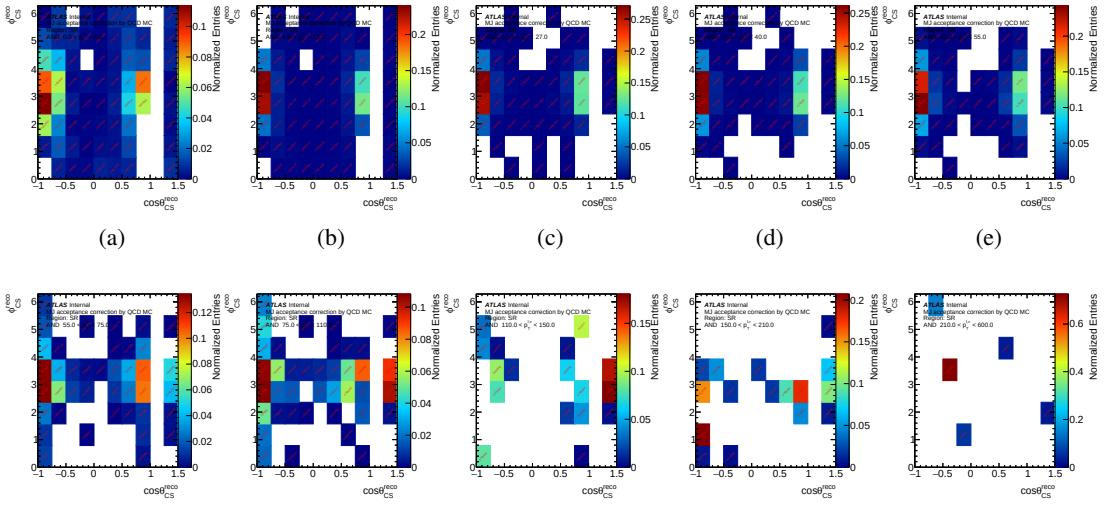


Figure 229: Multi-jet background 2D templates as a function of $p_T^{\ell,\nu}$ bins for $W^- \to e^-\overline{\nu}$ channel.

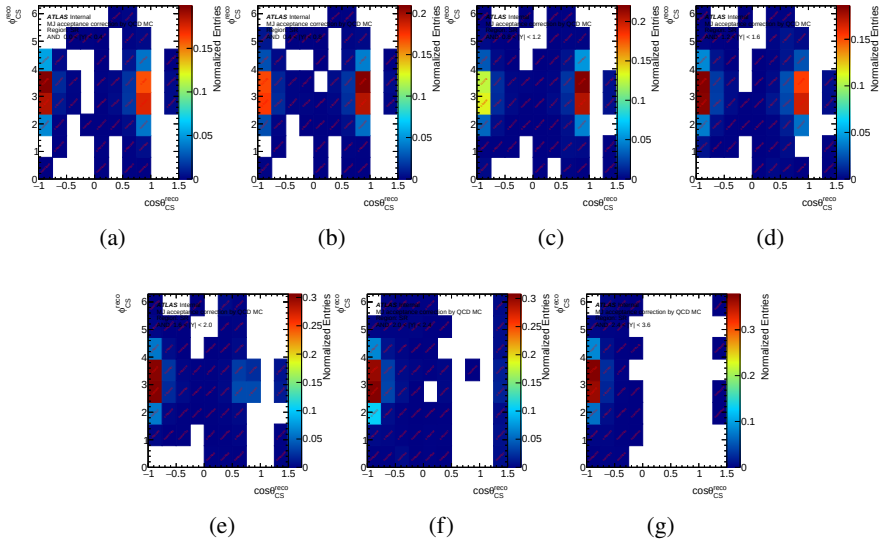
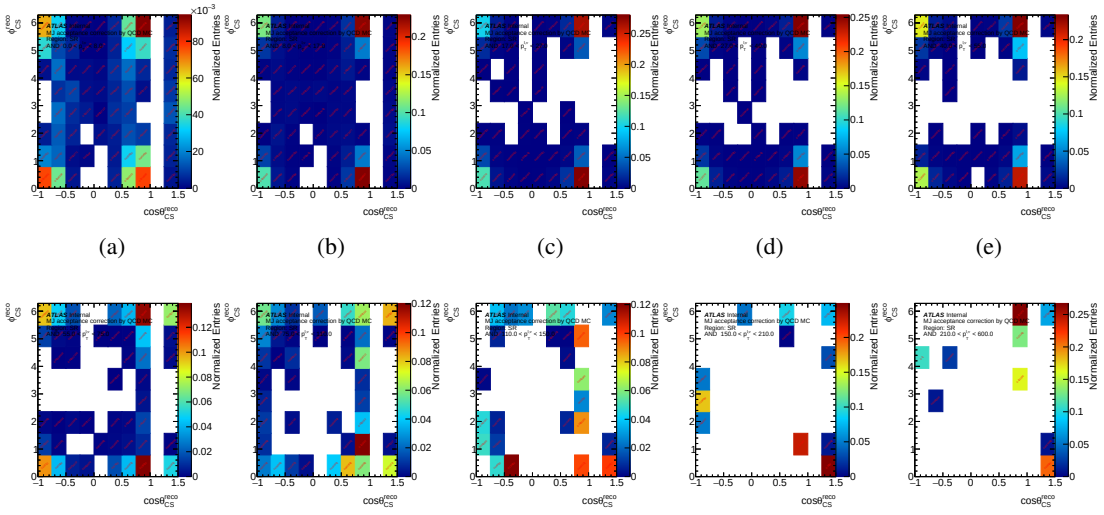


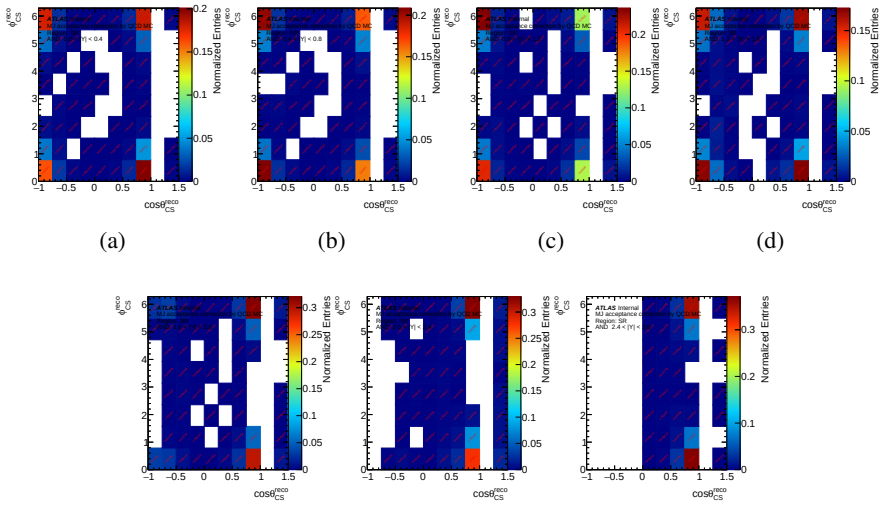
Figure 230: Multi-jet background 2D templates as a function of $|Y|$ bins for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

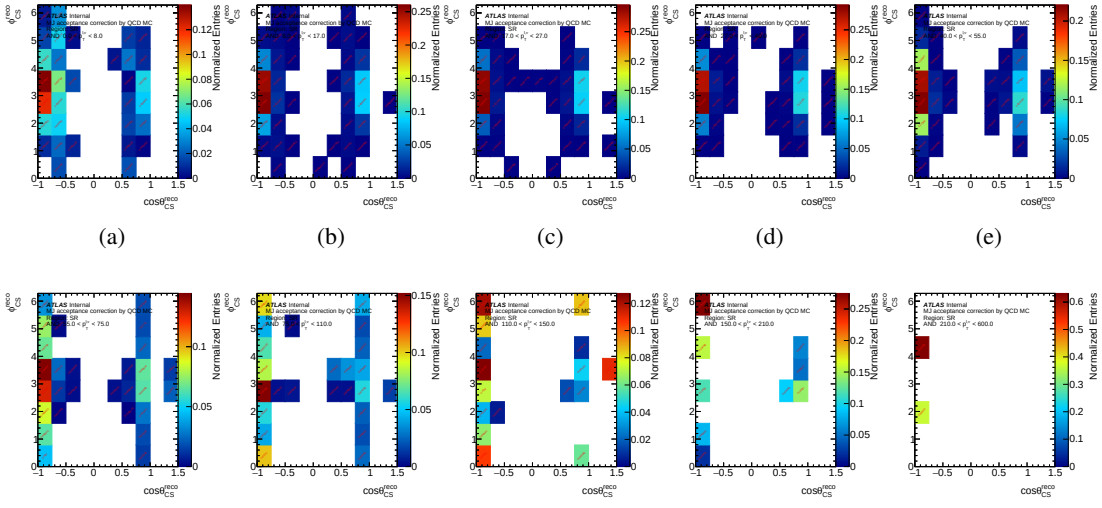
Figure 231: Multi-jet background 2D templates as a function of $p_T^{\ell,\nu}$ bins for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g)

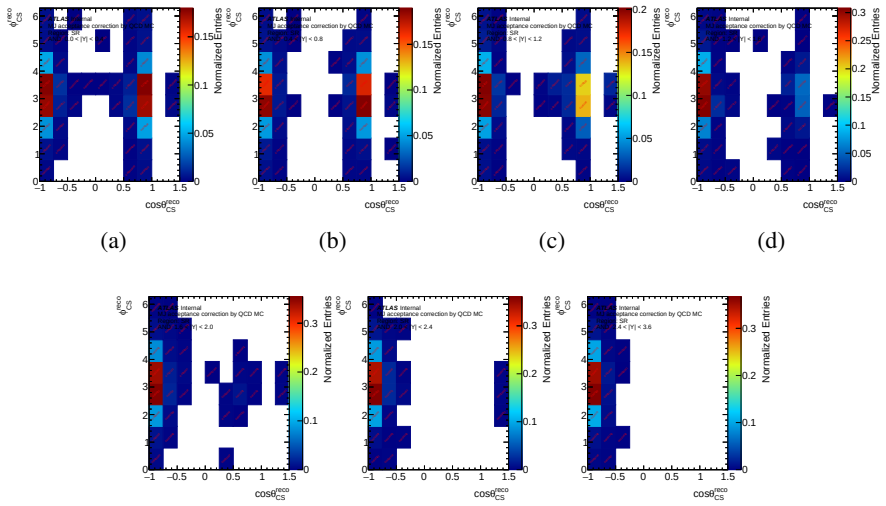
Figure 232: Multi-jet background 2D templates as a function of $|Y|$ bins for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

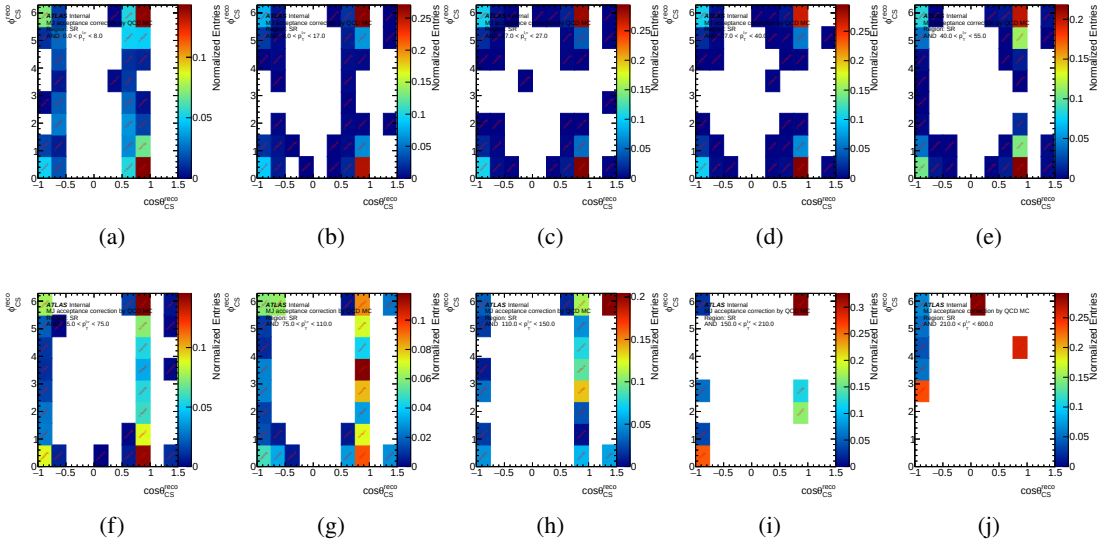
Figure 233: Multi-jet background 2D templates as a function of $p_T^{\ell,\nu}$ bins for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g)

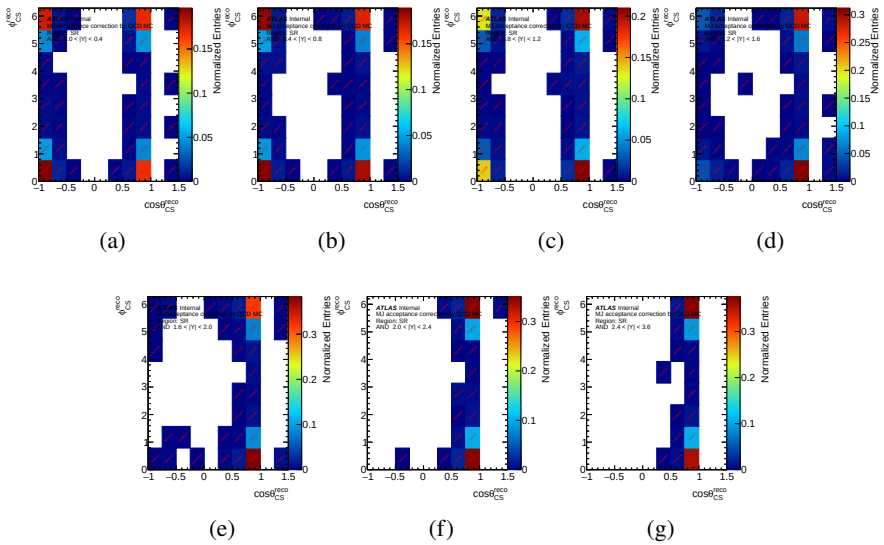
Figure 234: Multi-jet background 2D templates as a function of $|Y|$ bins for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 235: Multi-jet background 2D templates as a function of $p_T^{\ell,\nu}$ bins for $W^+ \to \mu^+\nu$ channel.



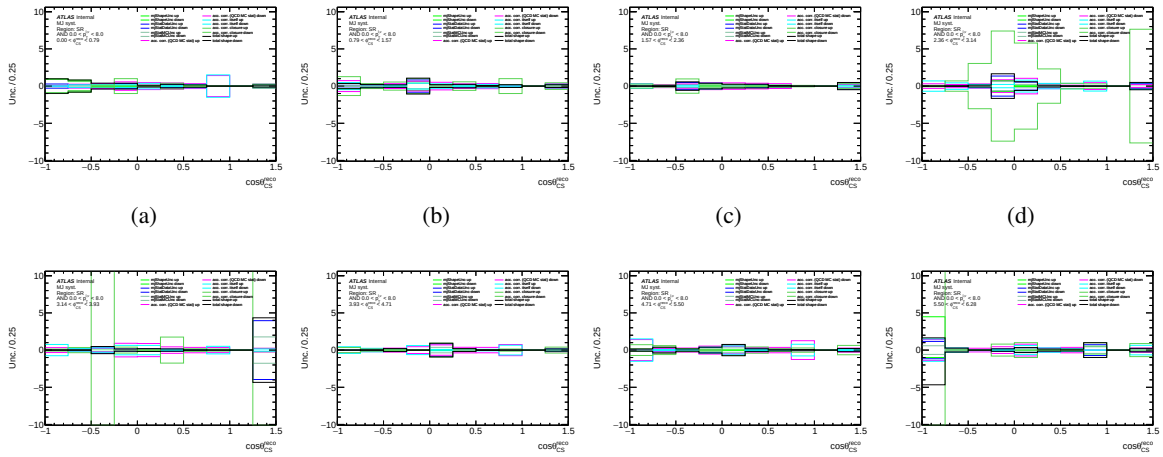
(a) (b) (c) (d)

(e) (f) (g)

Figure 236: Multi-jet background 2D templates as a function of $|Y|$ bins for $W^+ \to \mu^+\nu$ channel.

E.7 Multi Jet shape systematics breakdown

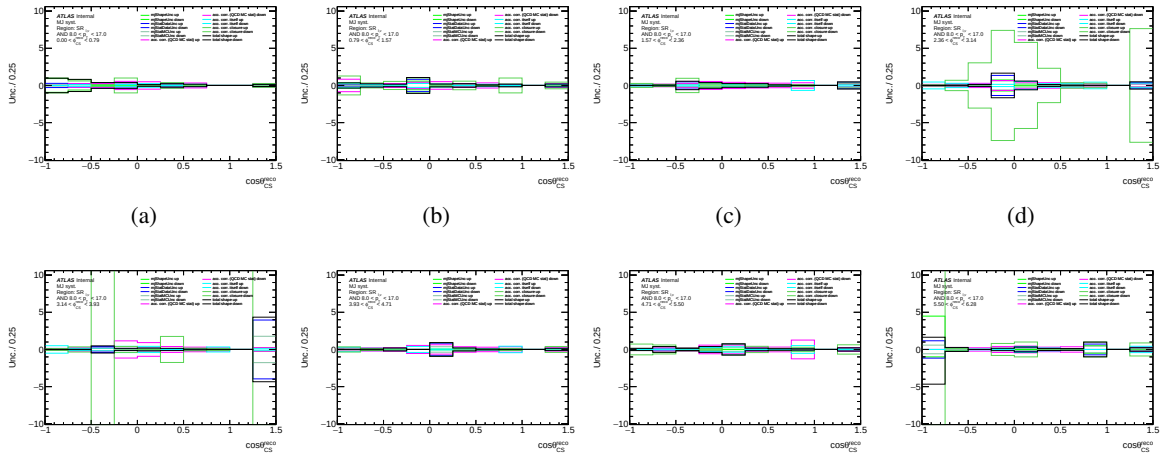
E.7.1 Electron channel



(a) (b) (c) (d)

(e) (f) (g) (h)

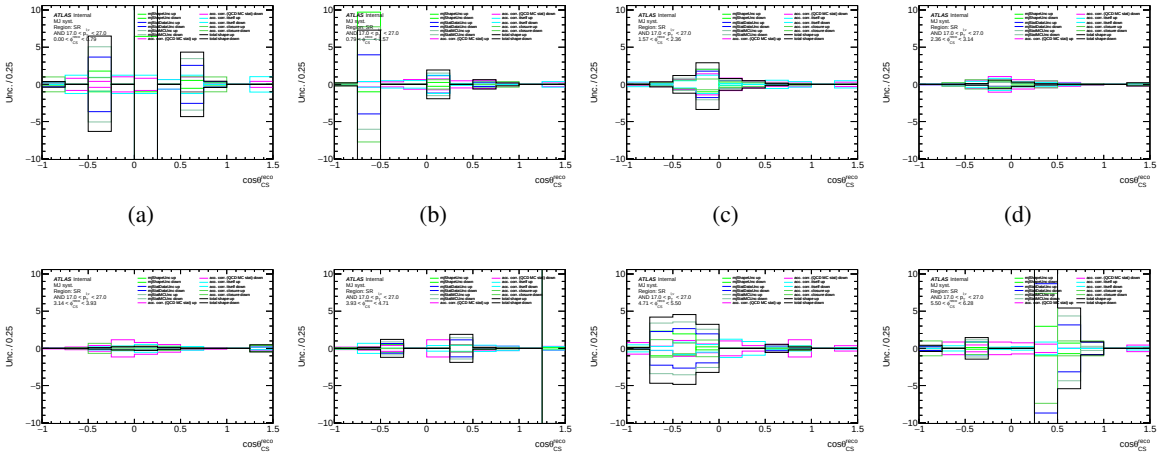
Figure 237: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \rightarrow e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

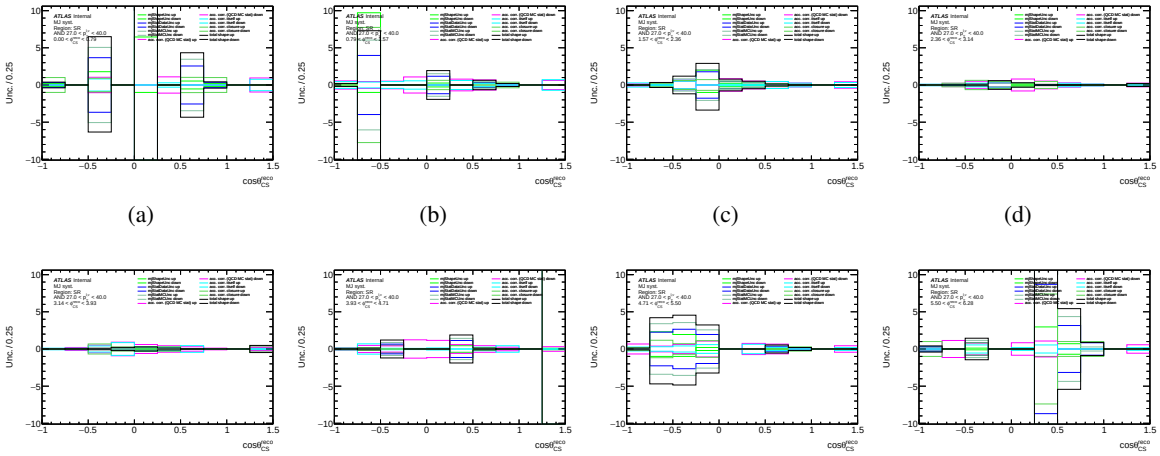
Figure 238: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \rightarrow e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

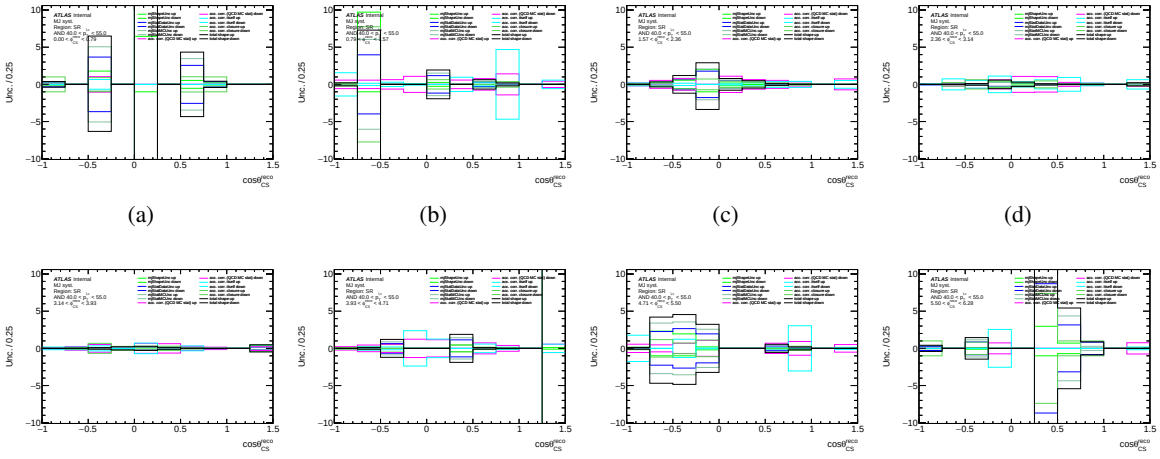
Figure 239: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

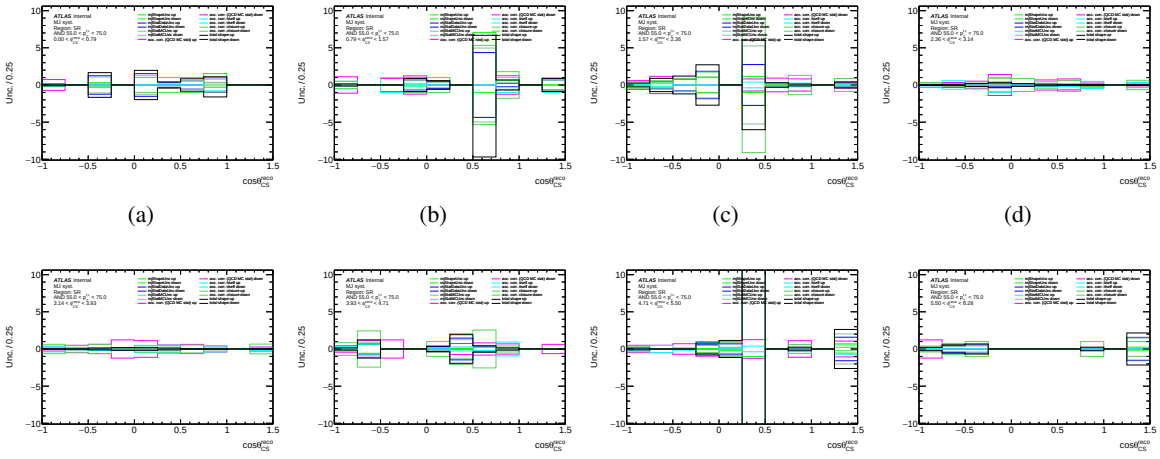
Figure 240: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

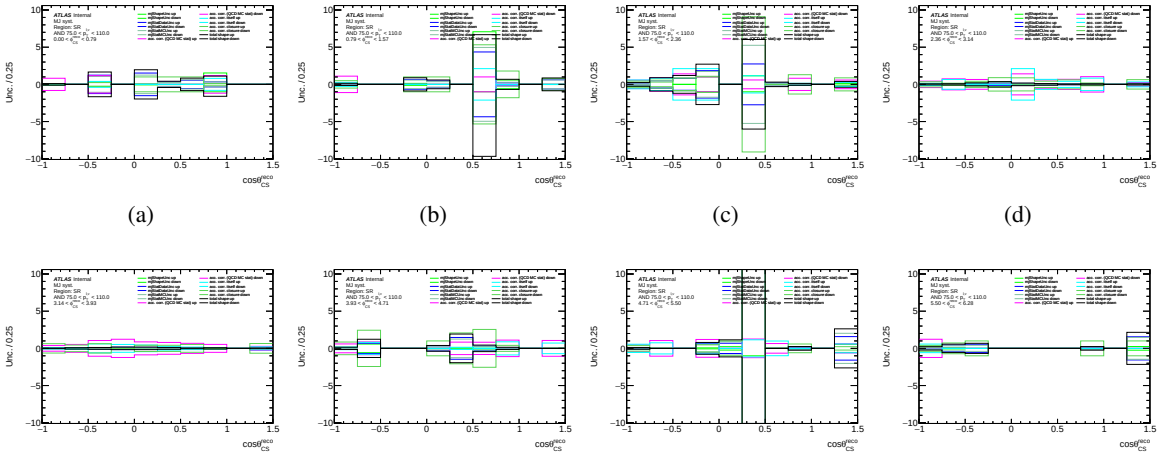
Figure 241: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

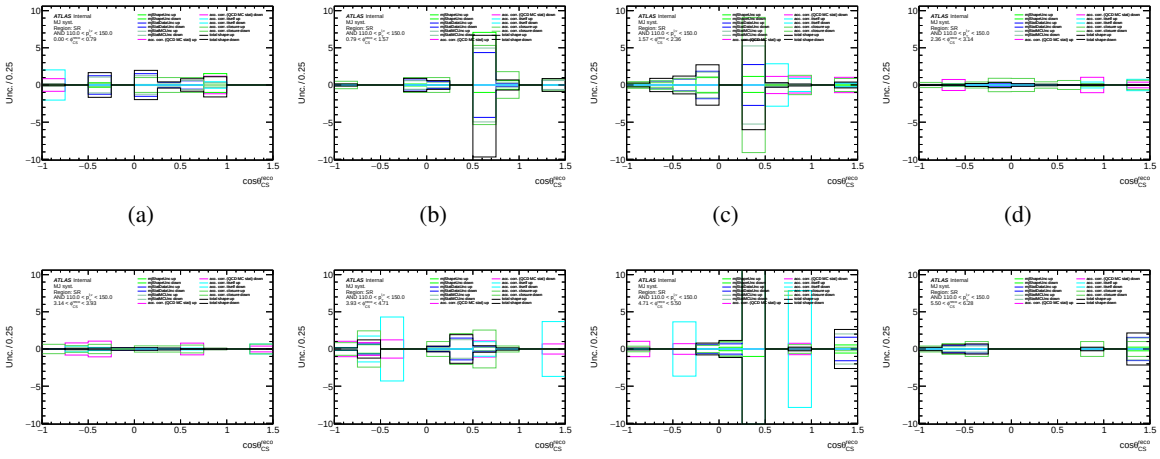
Figure 242: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

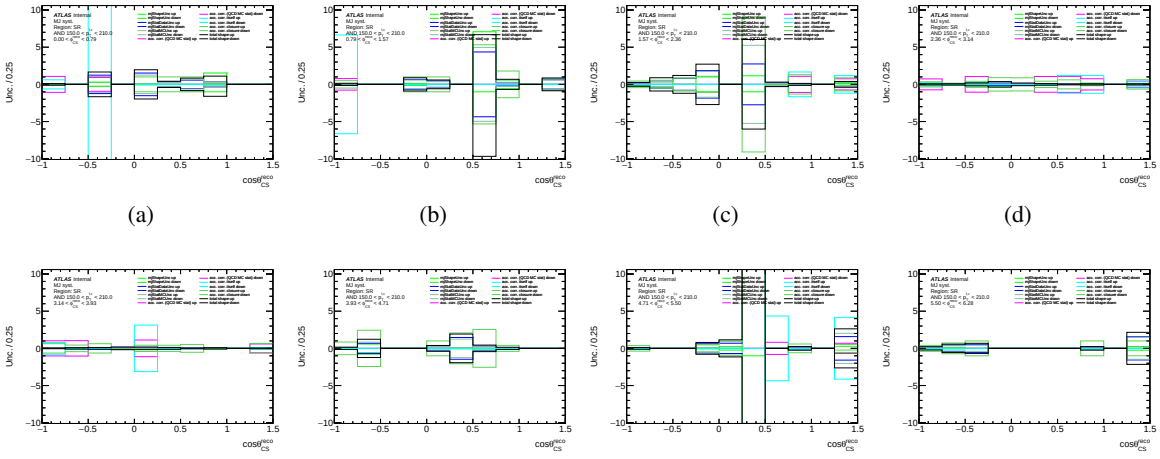
Figure 243: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

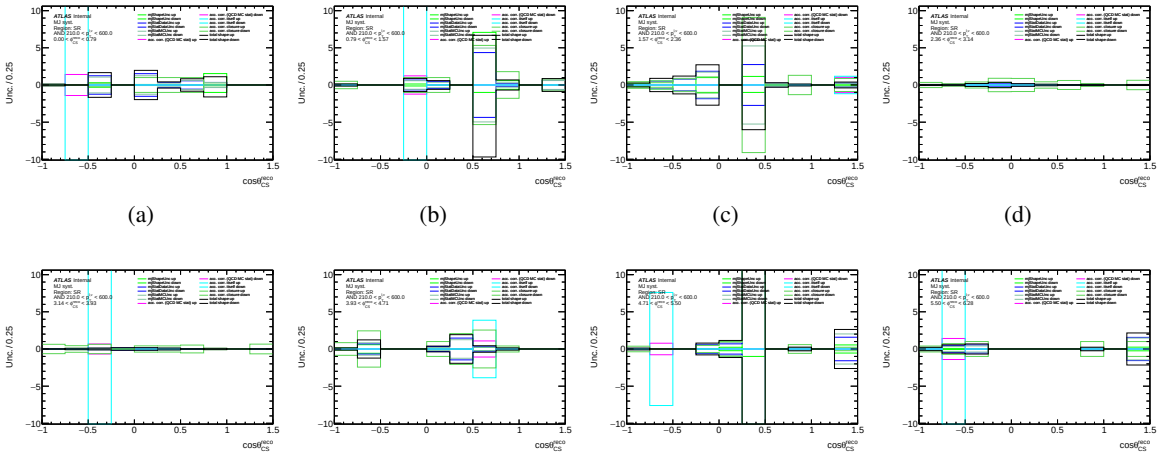
Figure 244: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

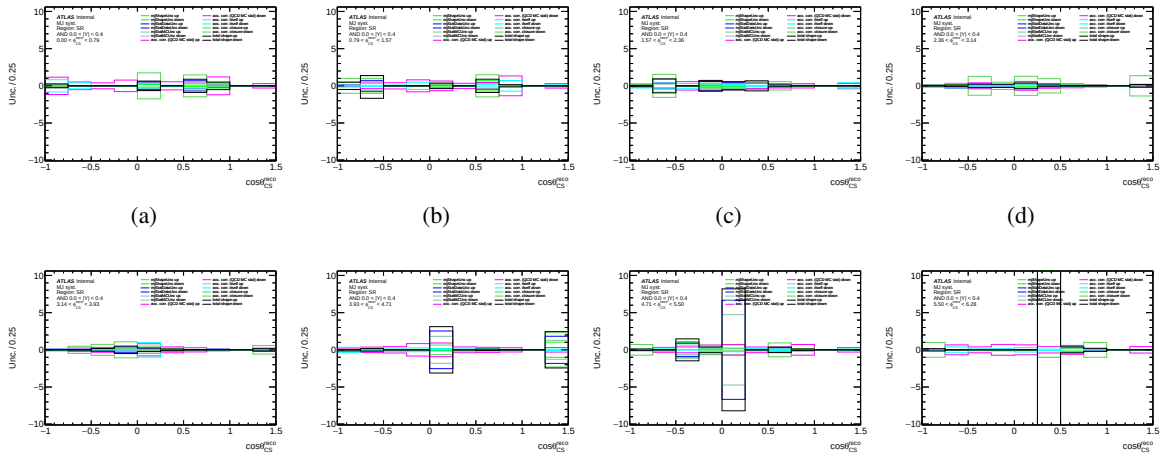
Figure 245: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

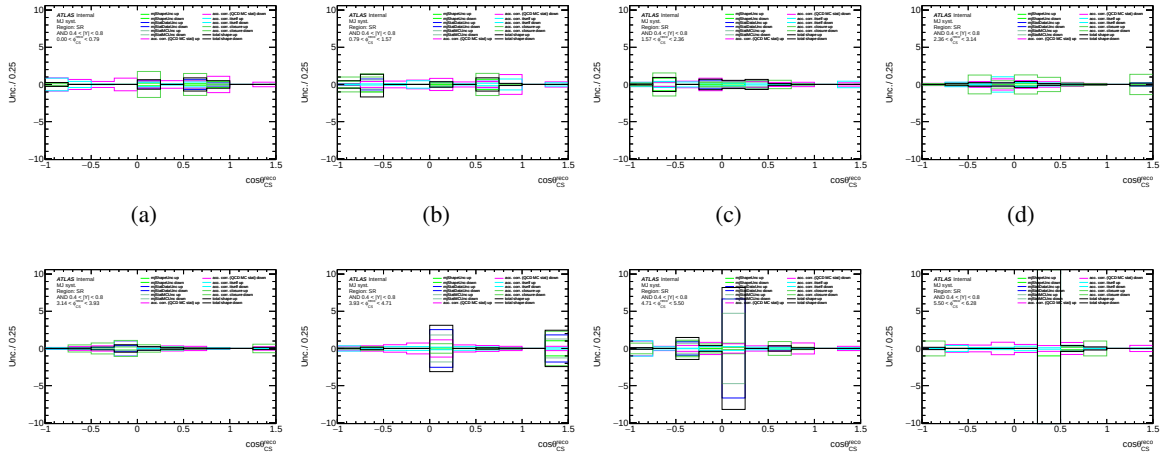
Figure 246: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

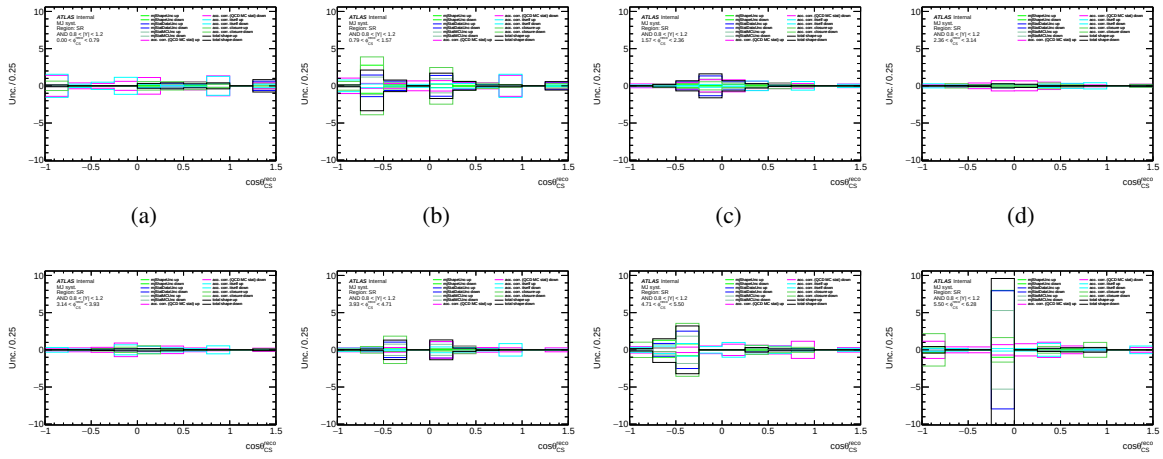
Figure 247: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

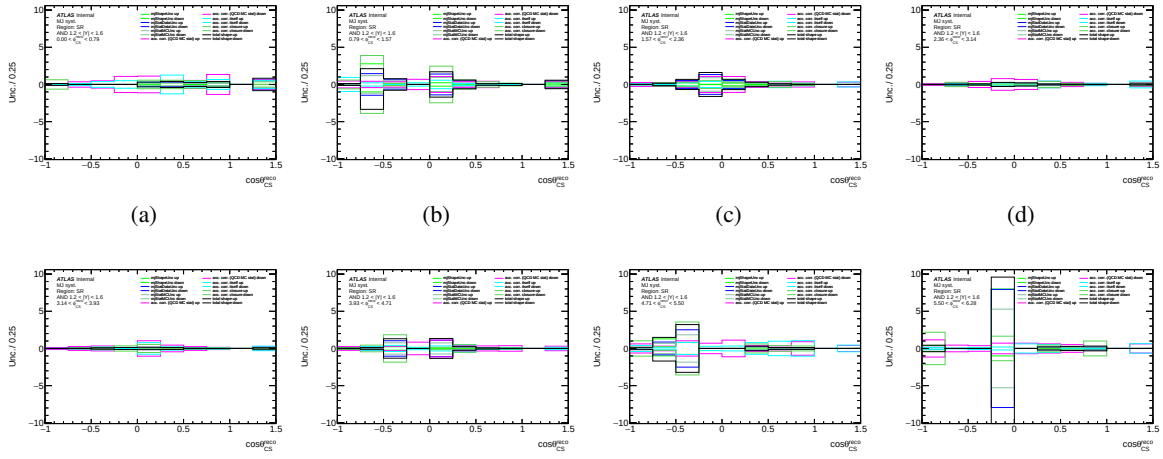
Figure 248: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

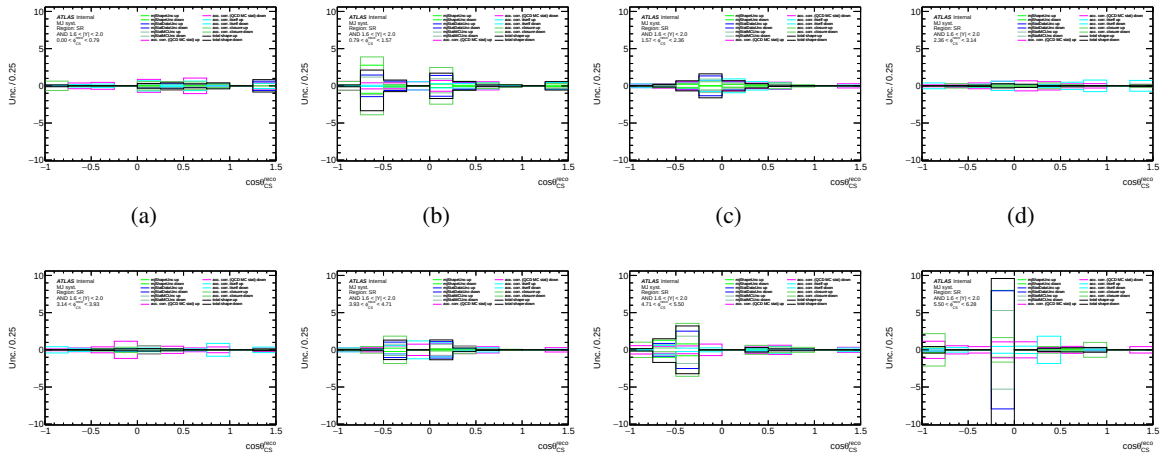
Figure 249: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

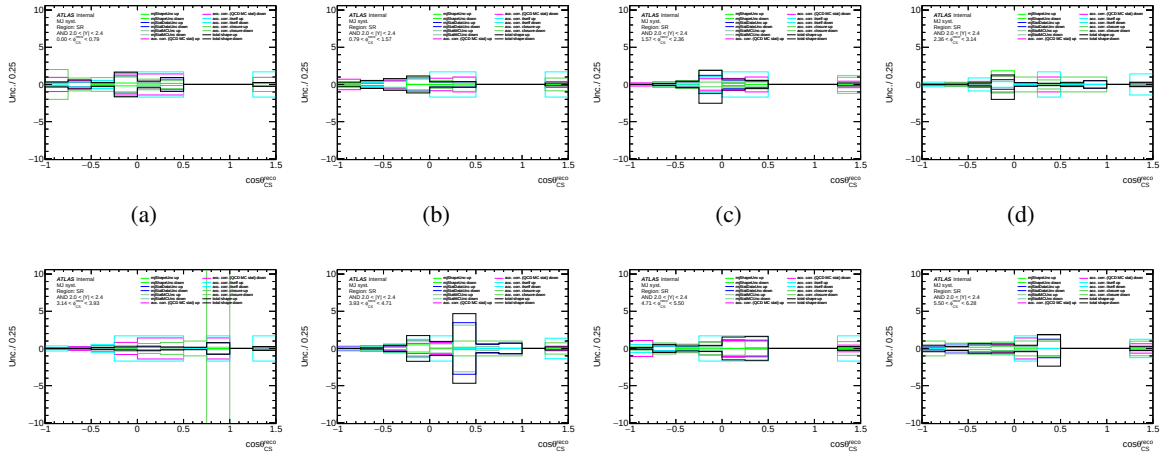
Figure 250: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

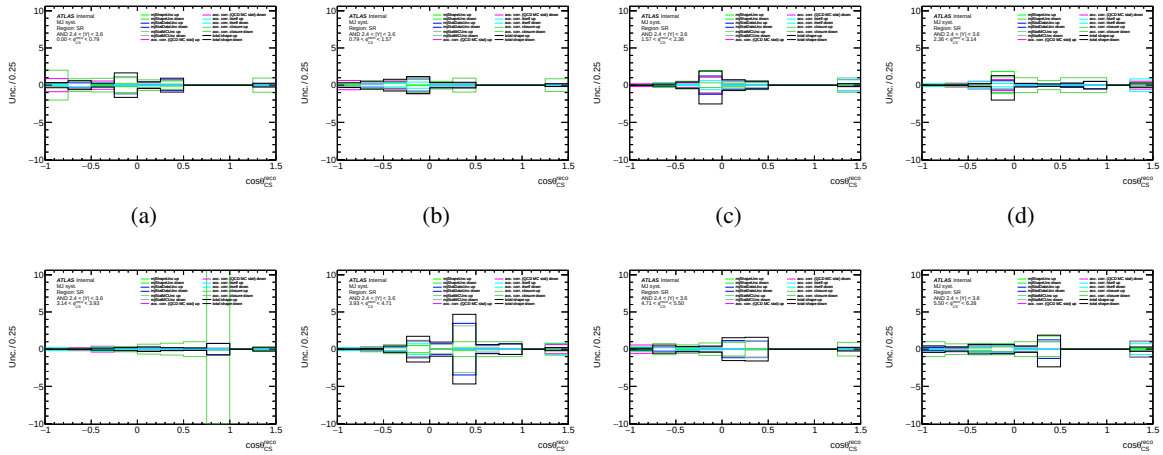
Figure 251: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 252: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^- \to e^-\overline{\nu}$ channel.

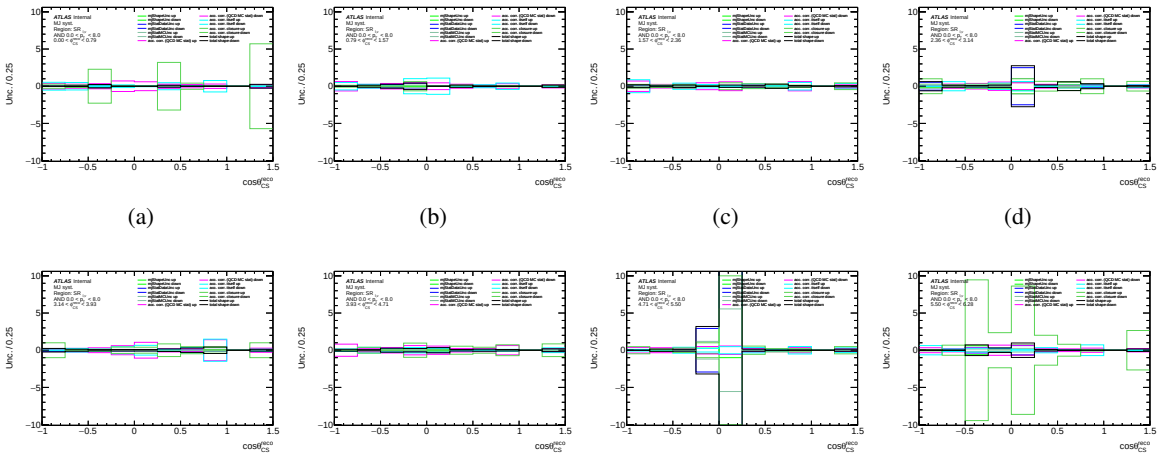


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 253: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^- \to e^-\overline{\nu}$ channel.

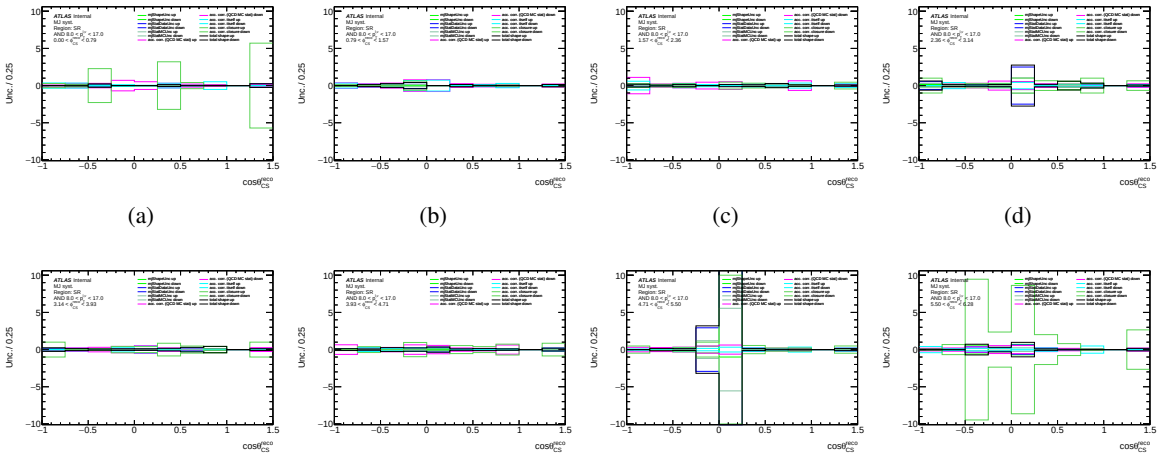
E.7.2 Positron channel



(a) (b) (c) (d)

(e) (f) (g) (h)

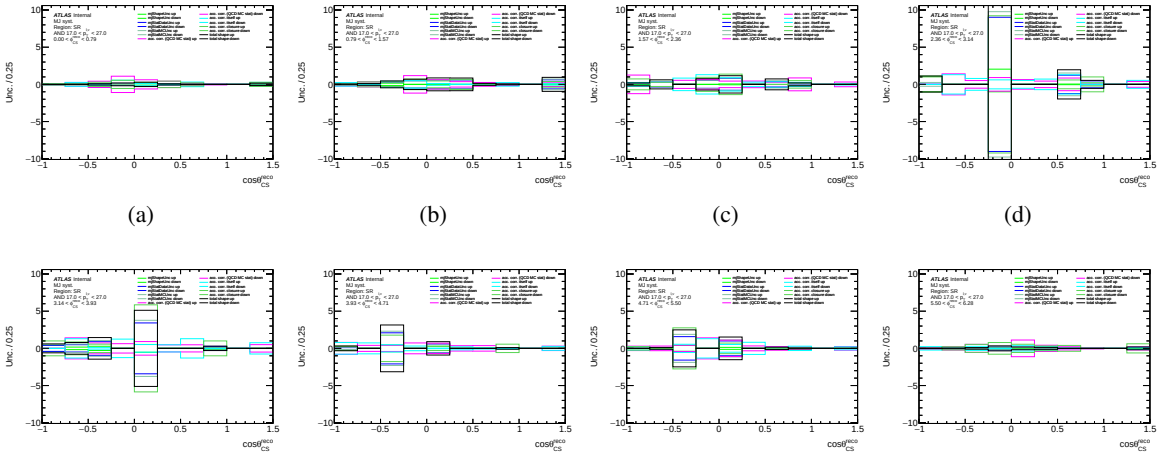
Figure 254: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

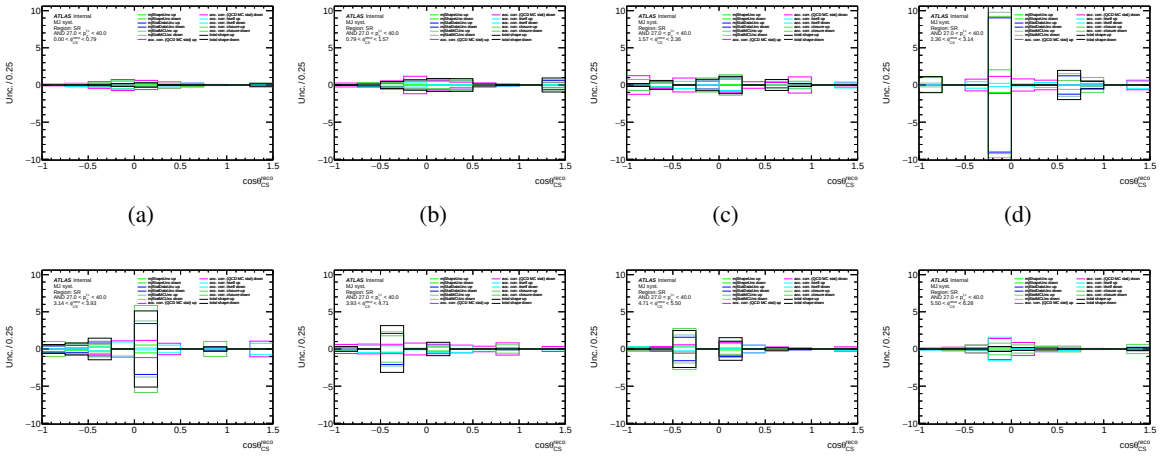
Figure 255: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

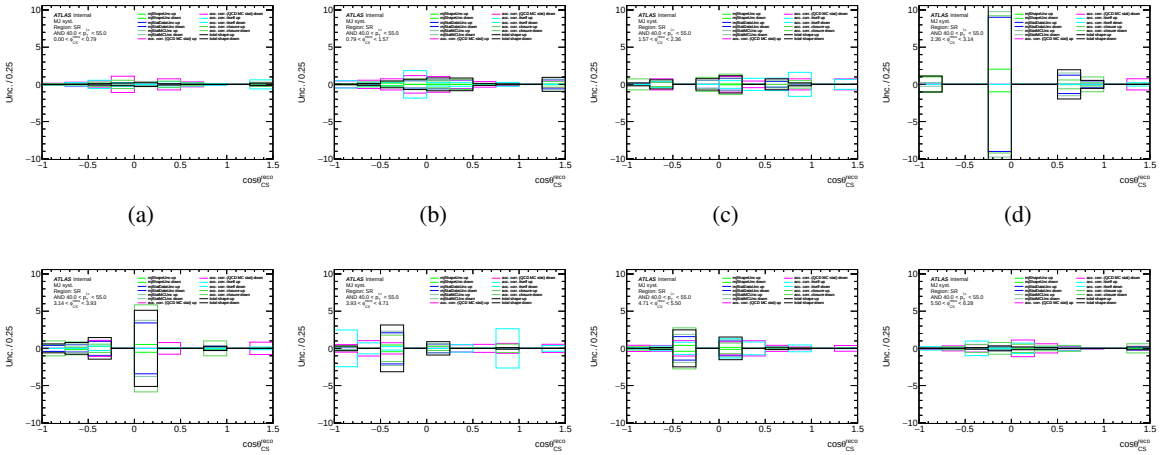
Figure 256: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

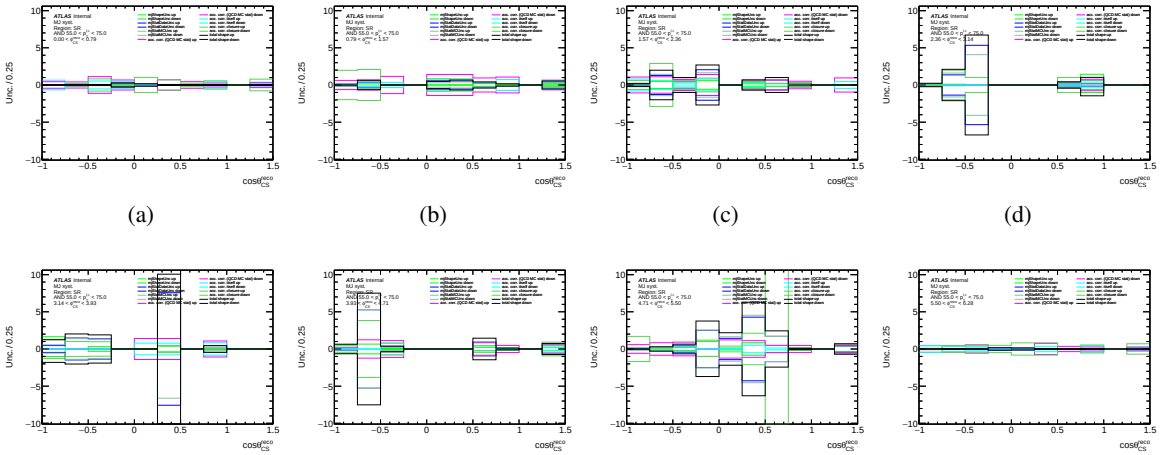
Figure 257: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

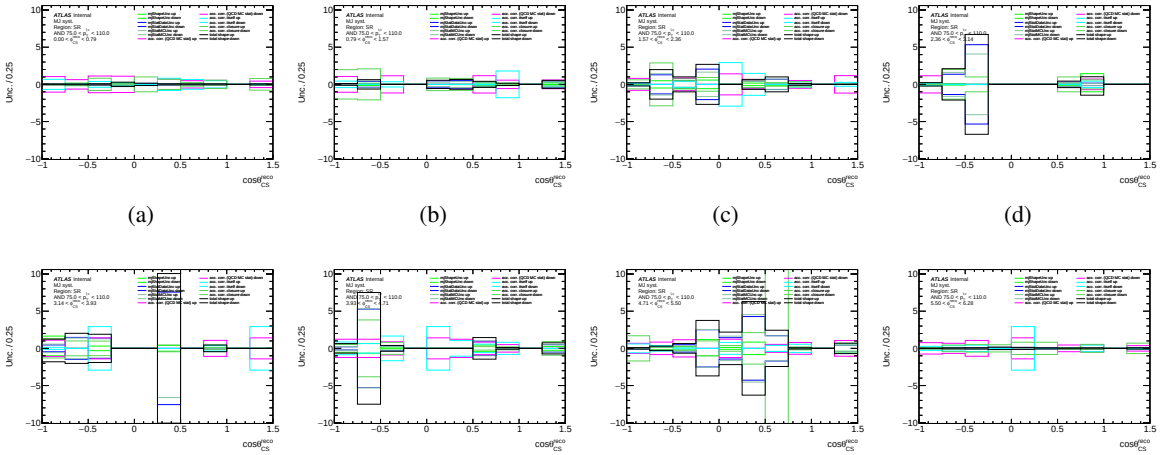
Figure 258: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

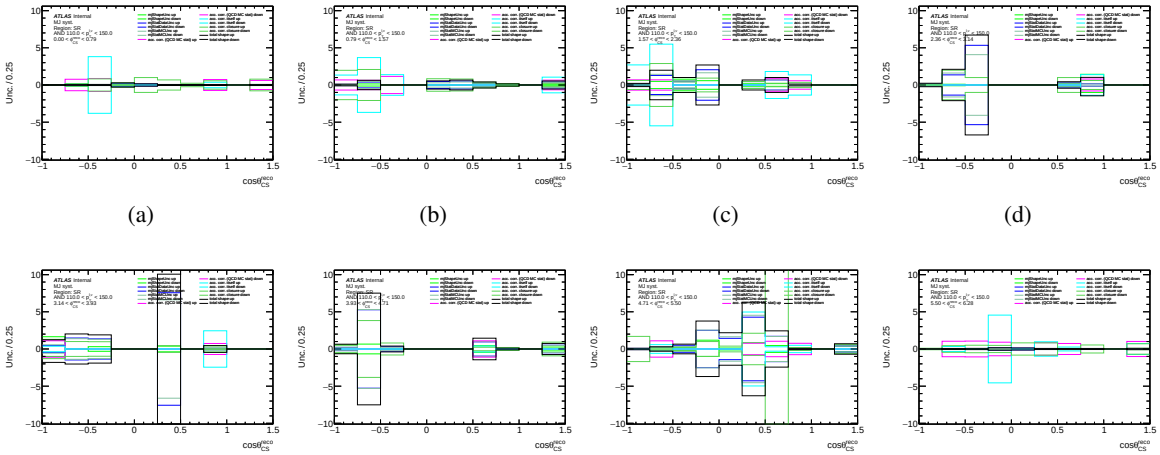
Figure 259: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

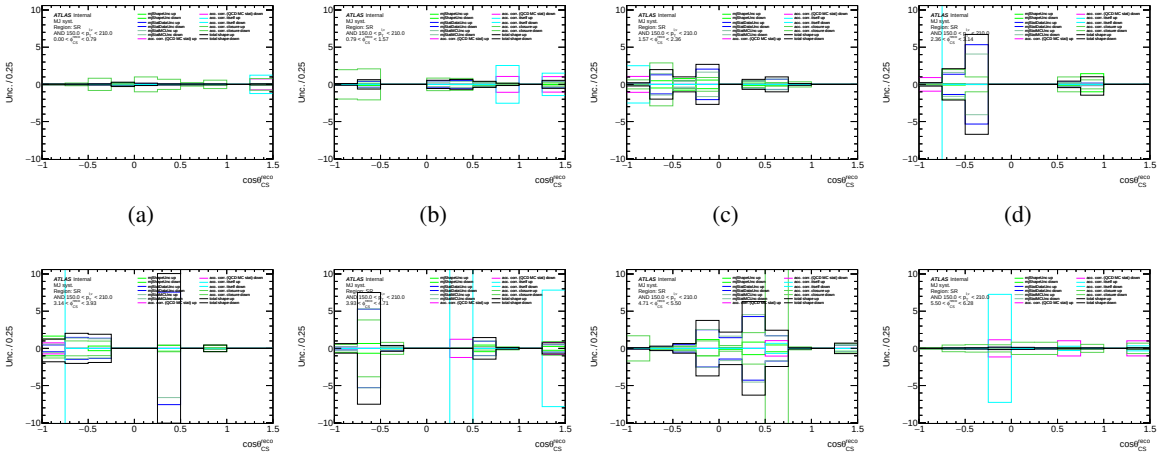
Figure 260: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

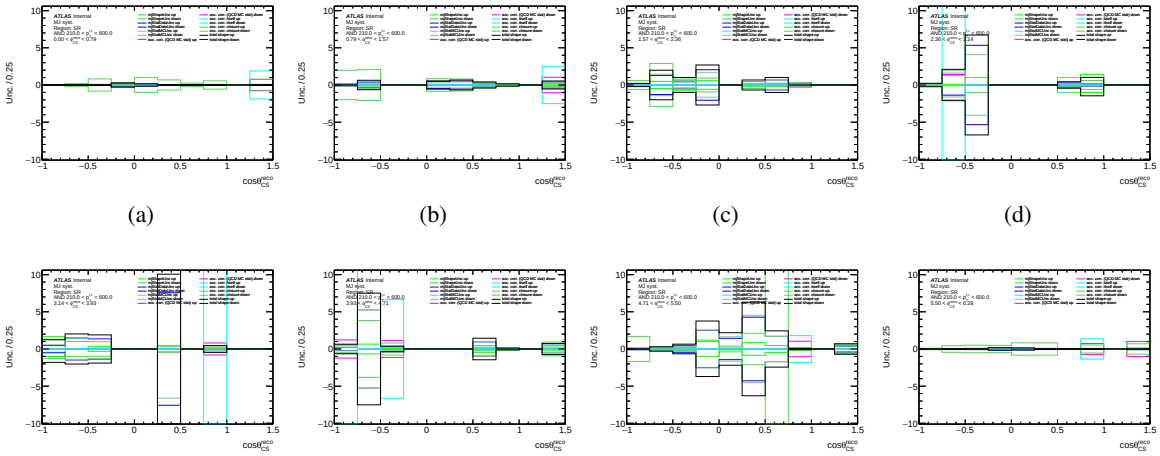
Figure 261: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \rightarrow e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

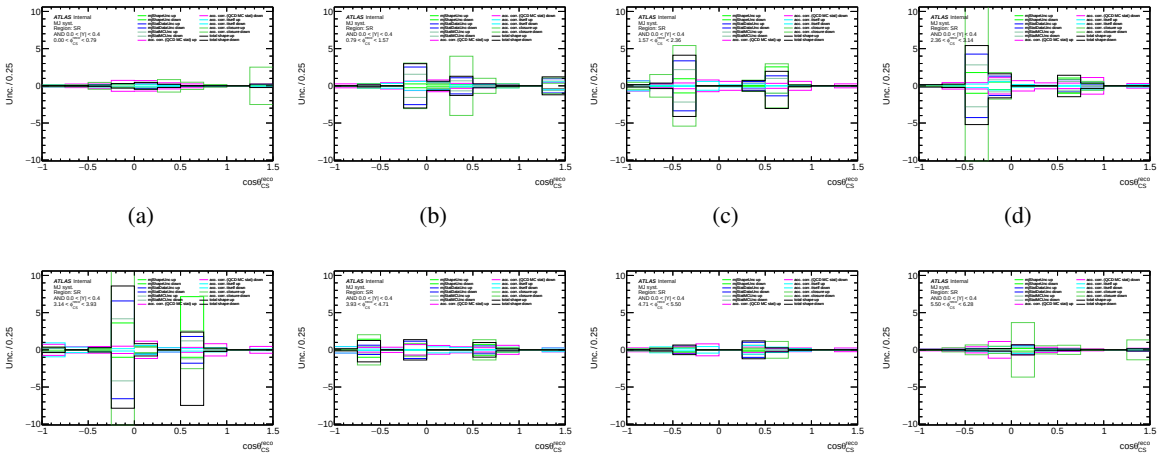
Figure 262: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

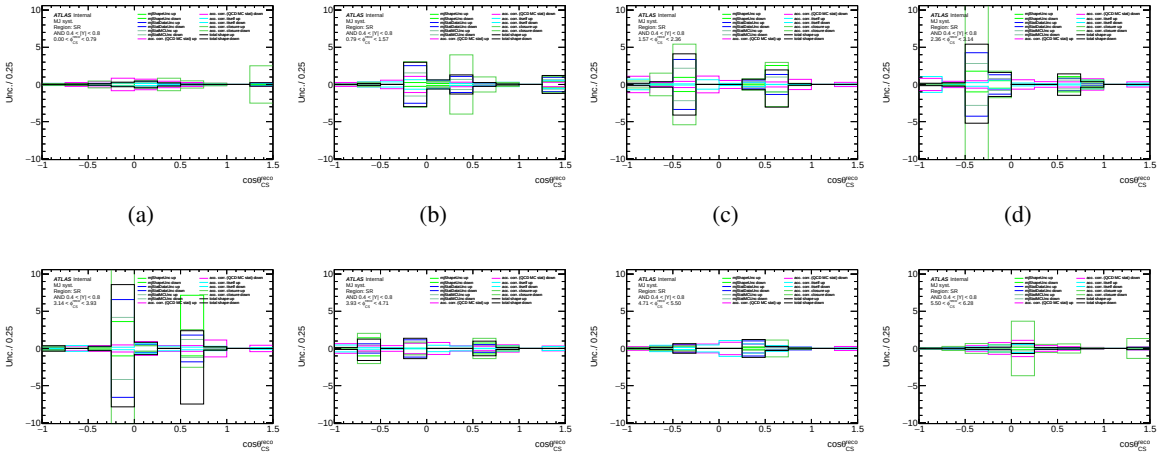
Figure 263: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

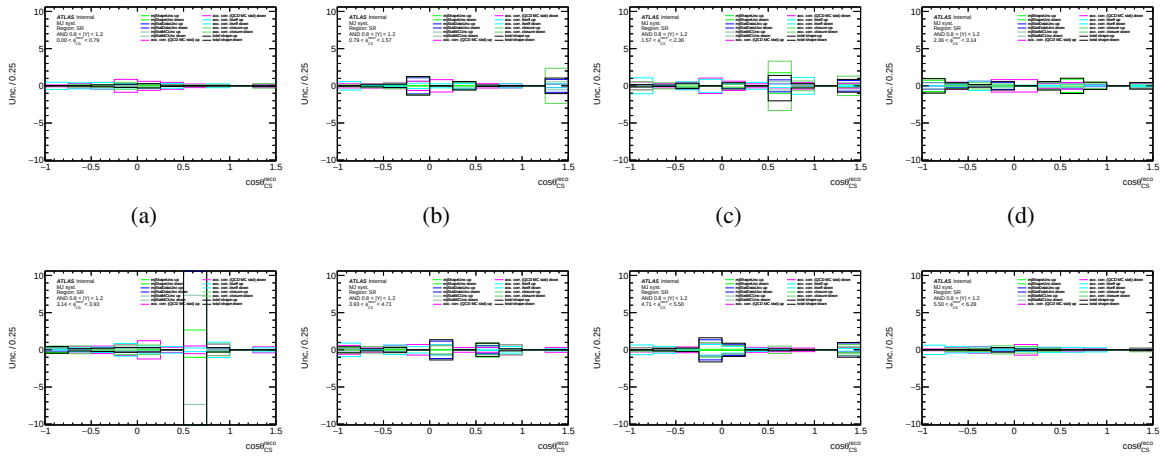
Figure 264: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

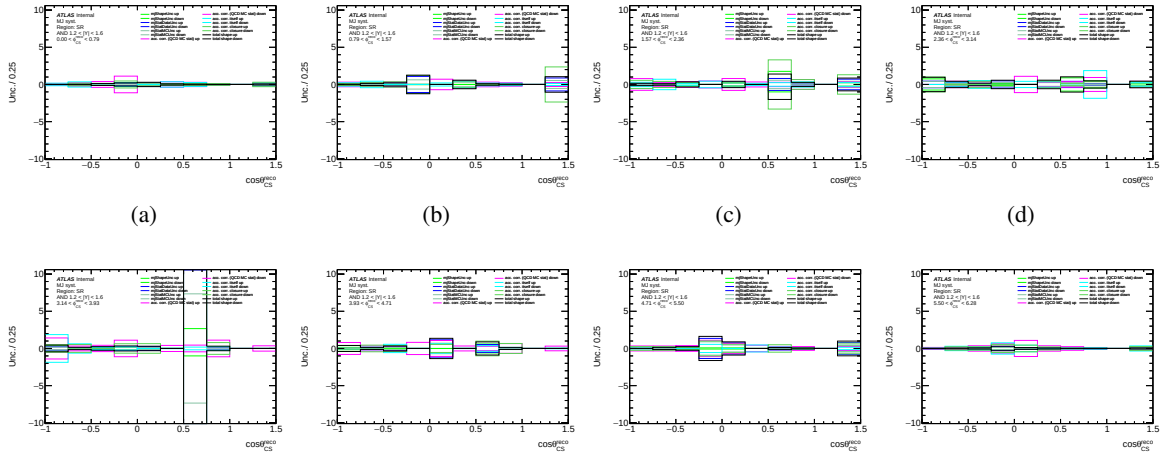
Figure 265: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

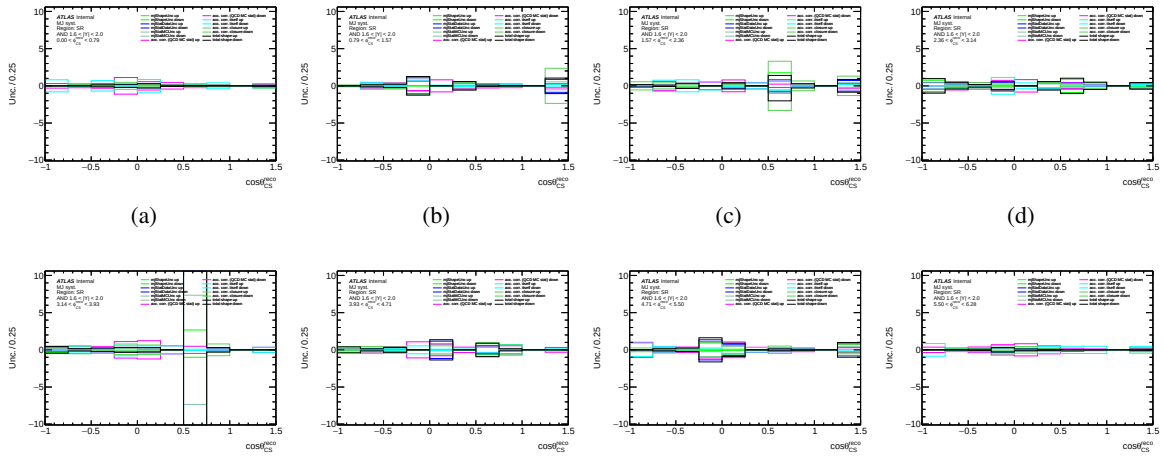
Figure 266: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

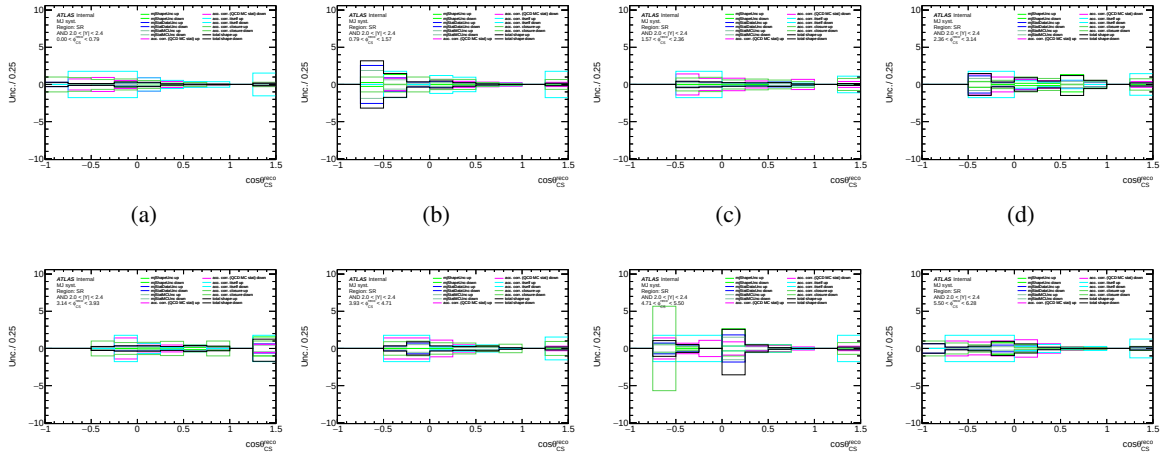
Figure 267: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

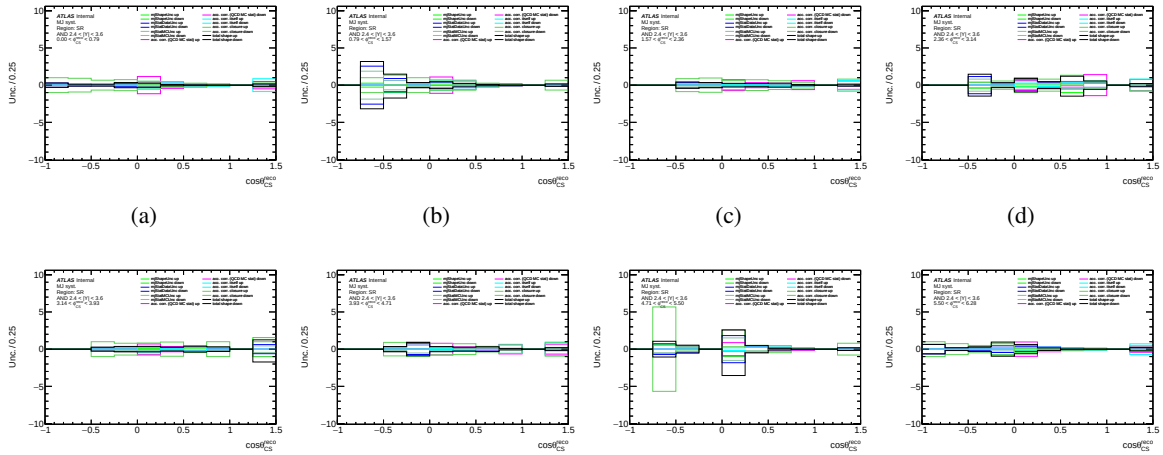
Figure 268: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 269: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^+ \to e^+\nu$ channel.

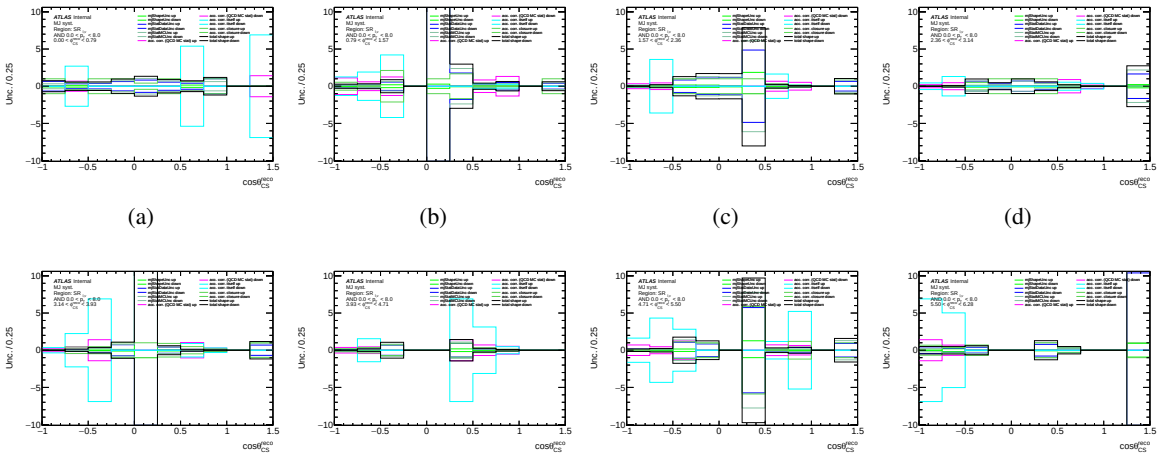


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 270: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^+ \to e^+\nu$ channel.

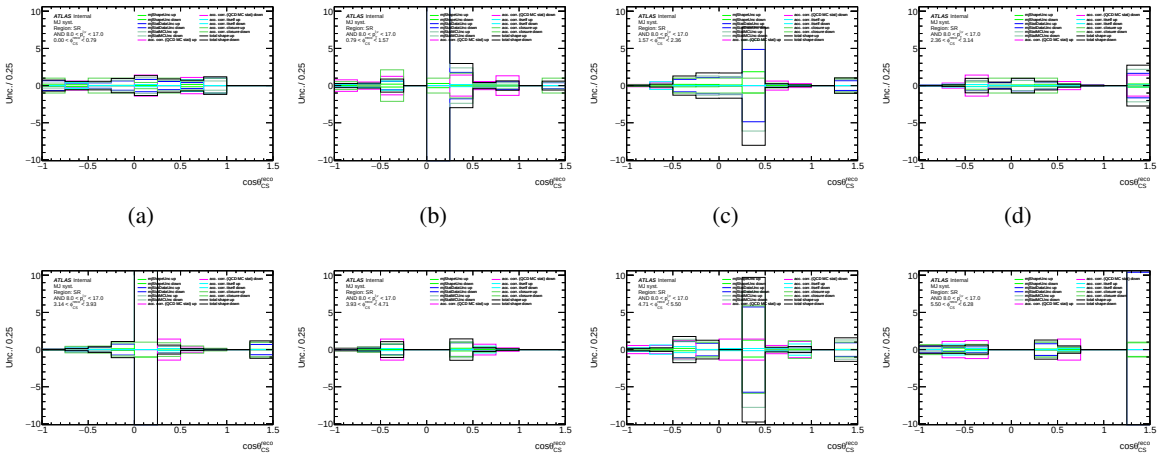
E.7.3 Muon channel



(a) (b) (c) (d)

(e) (f) (g) (h)

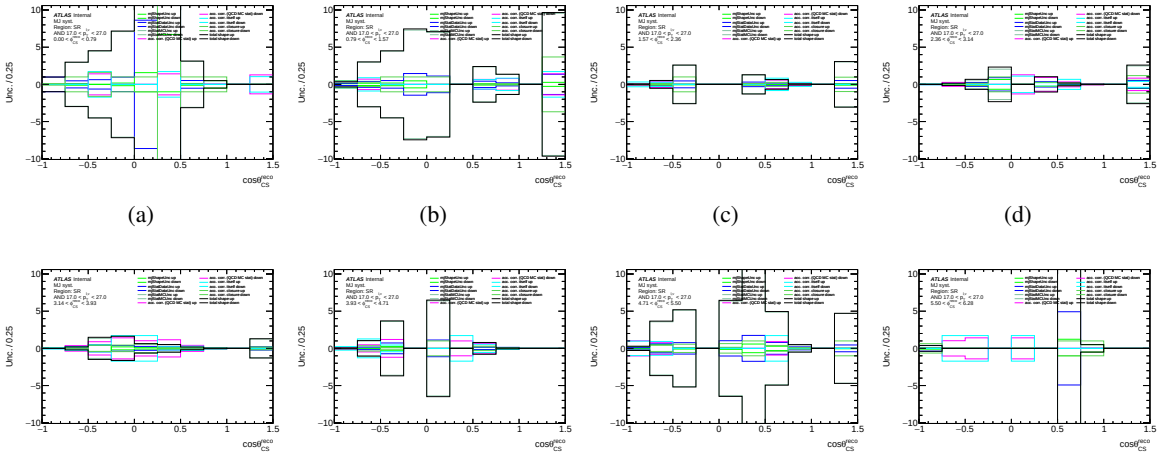
Figure 271: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

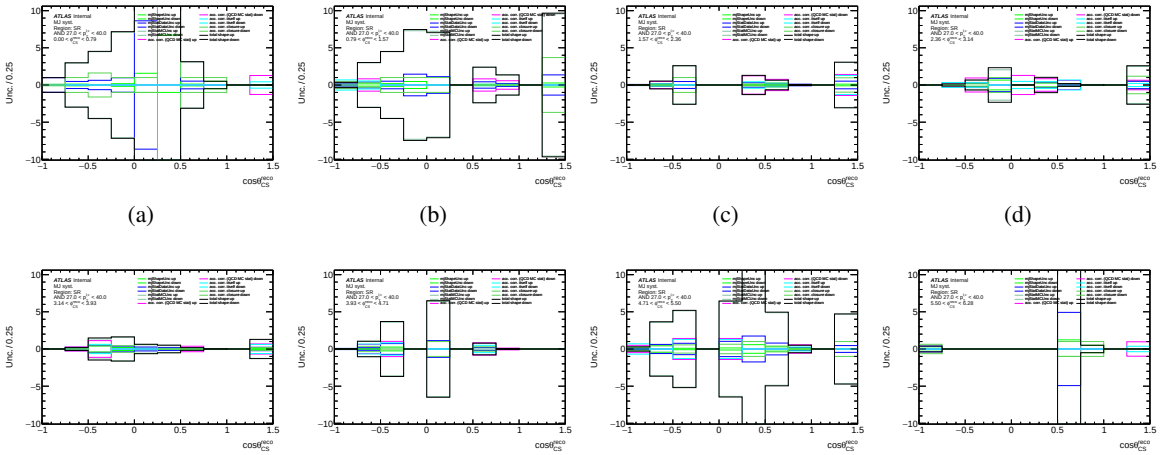
Figure 272: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

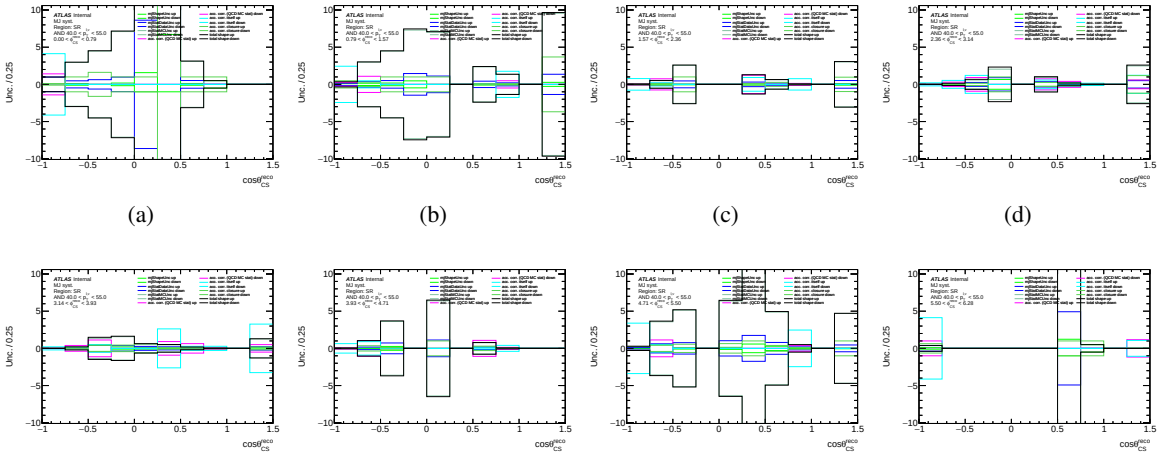
Figure 273: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

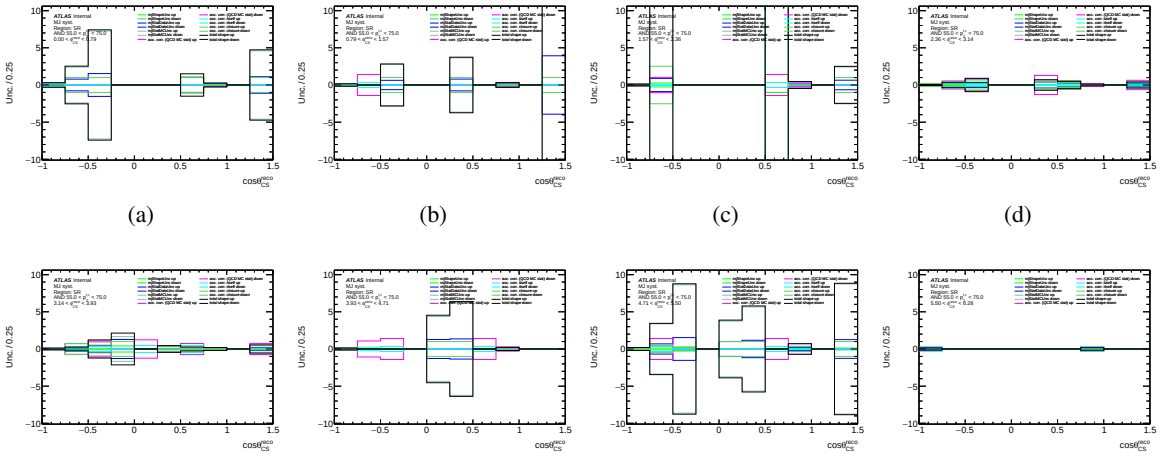
Figure 274: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

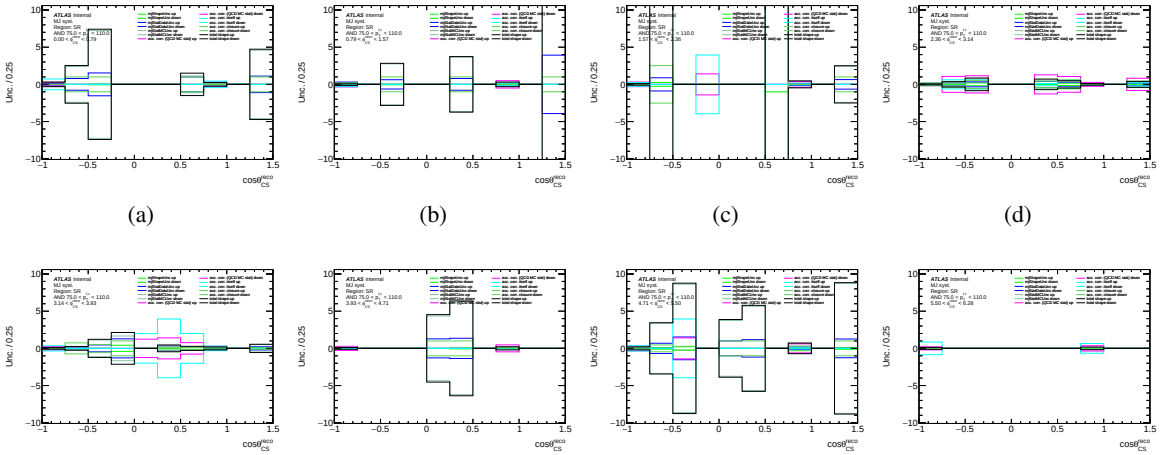
Figure 275: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

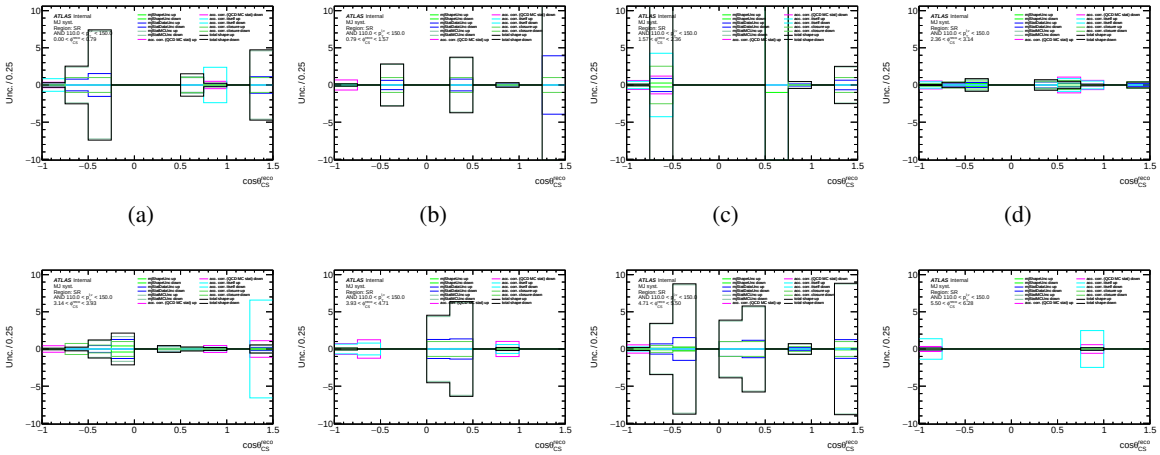
Figure 276: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

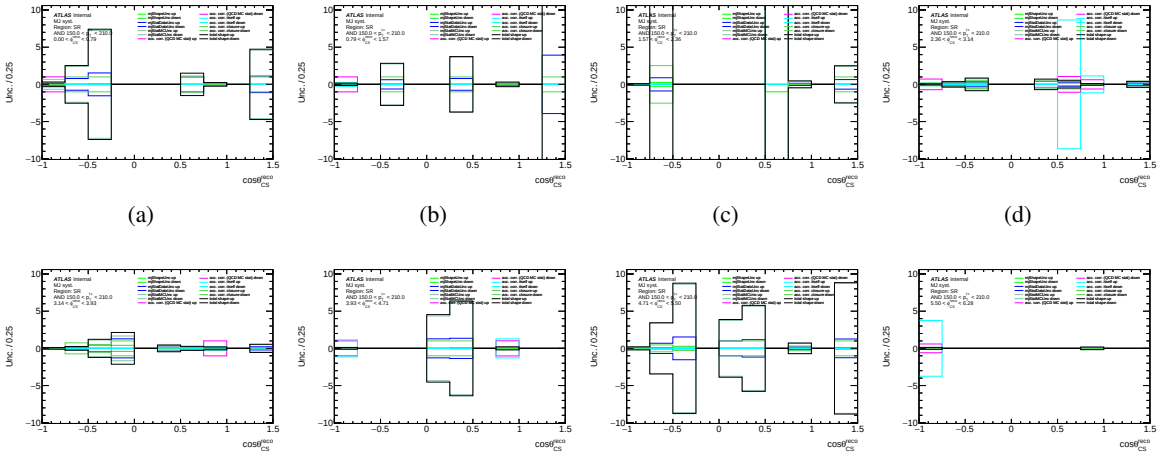
Figure 277: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

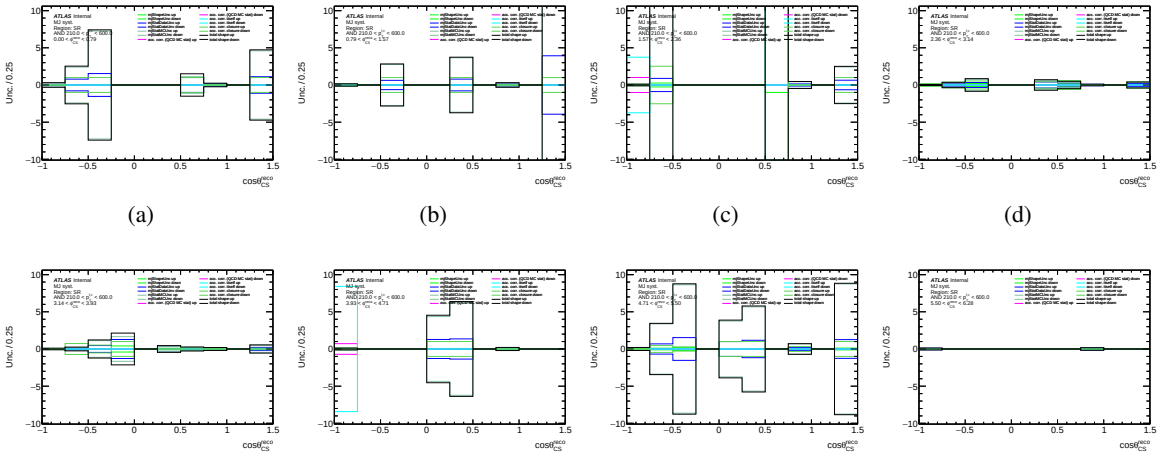
Figure 278: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

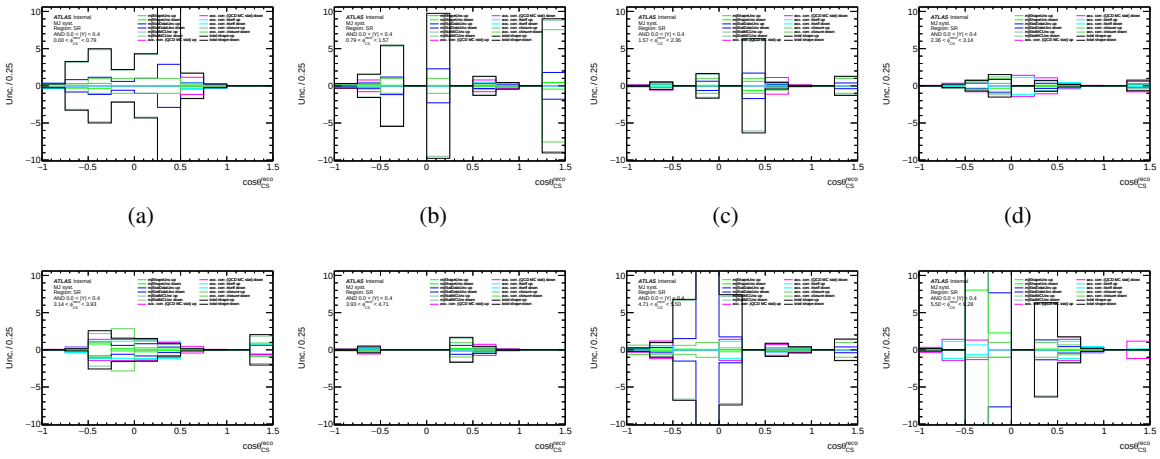
Figure 279: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

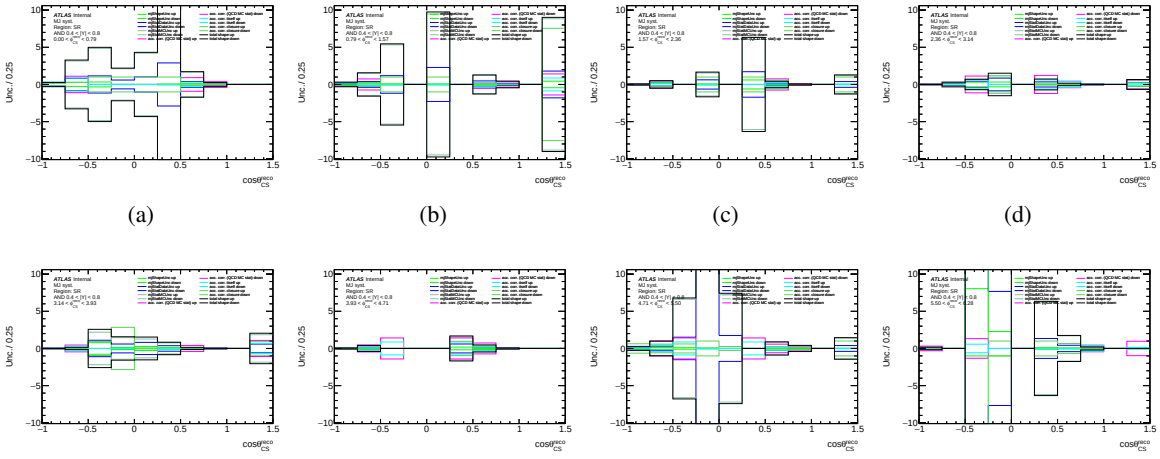
Figure 280: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

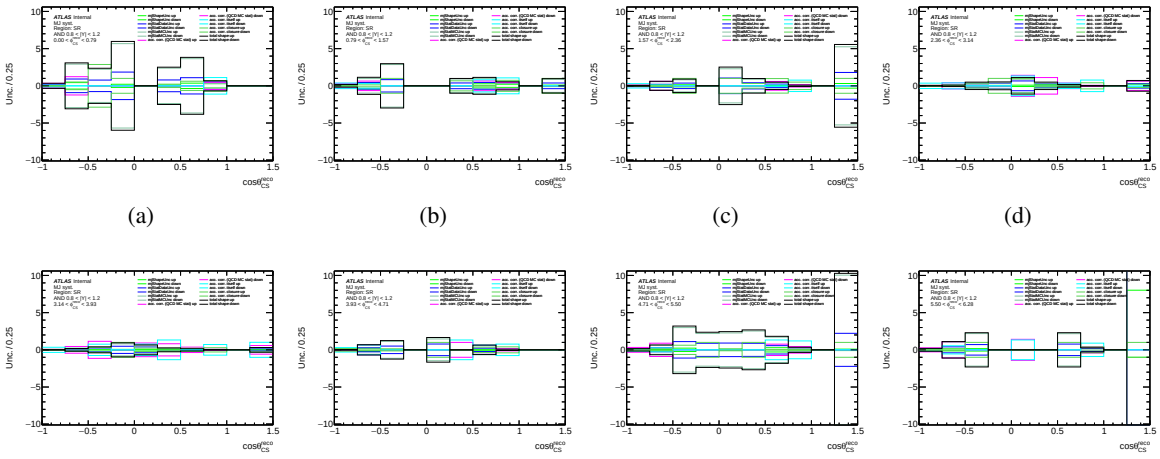
Figure 281: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

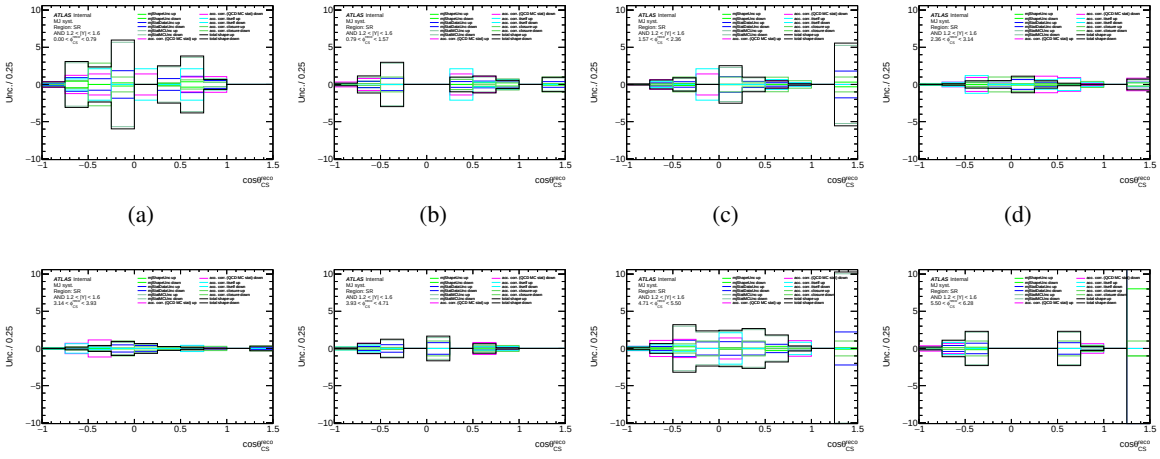
Figure 282: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^- \to \mu^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

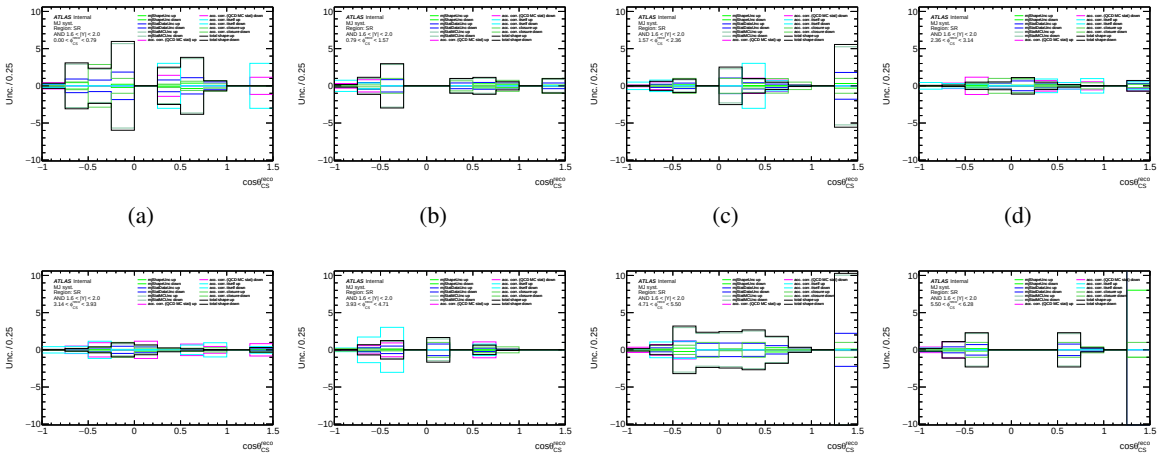
Figure 283: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

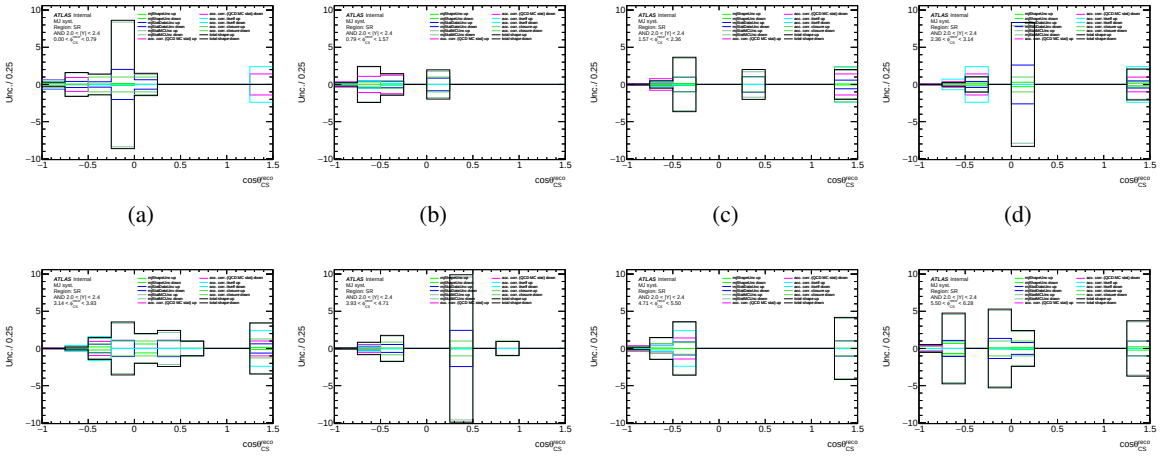
Figure 284: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

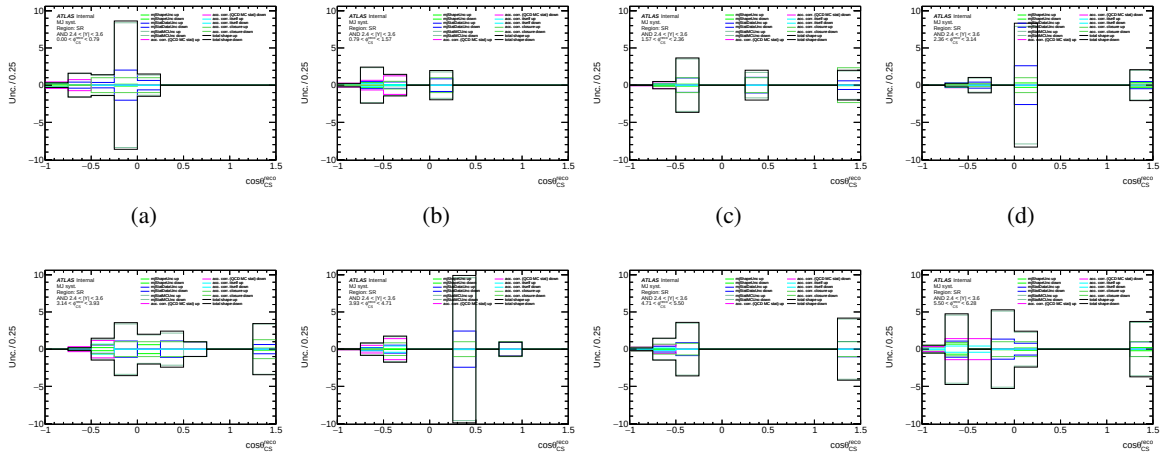
Figure 285: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 286: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

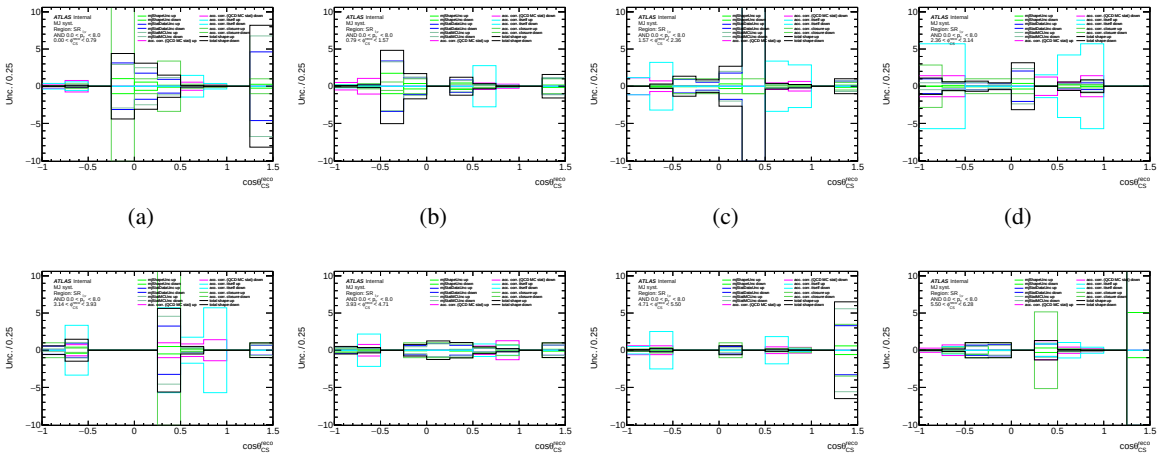


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 287: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

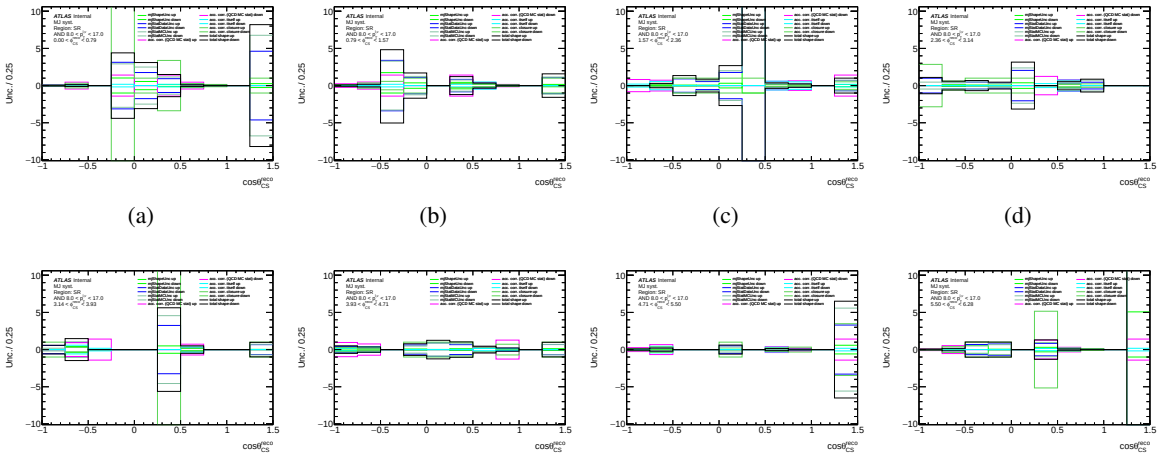
E.7.4 Anti-muon channel



(a) (b) (c) (d)

(e) (f) (g) (h)

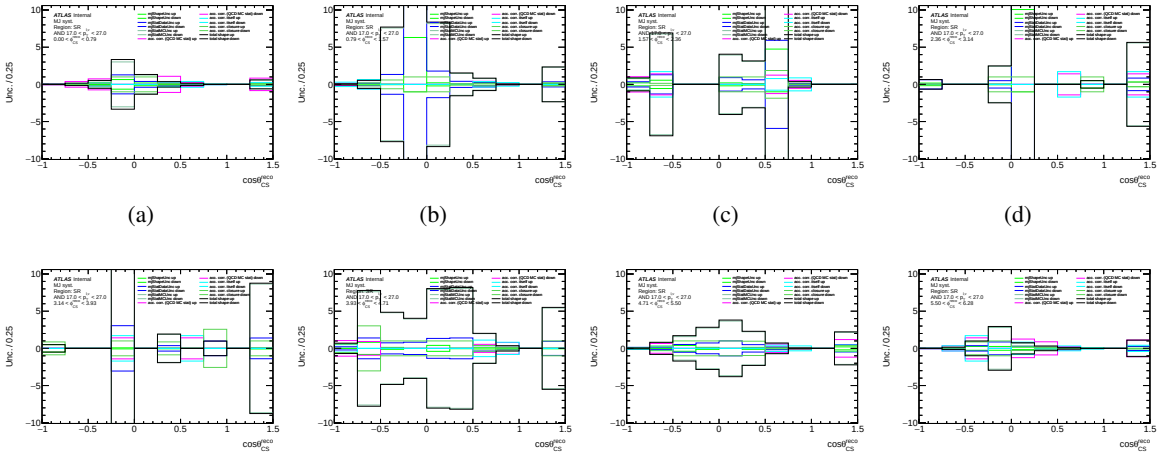
Figure 288: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

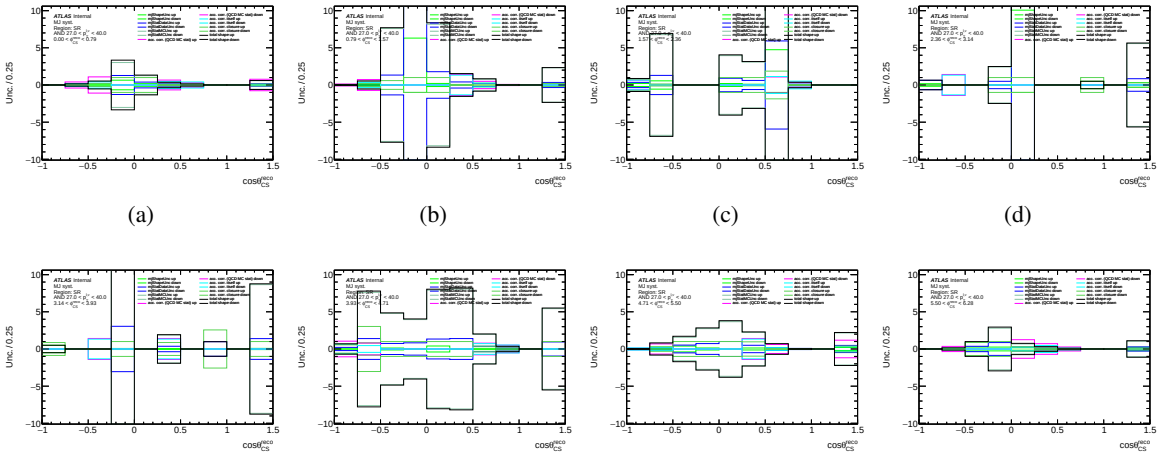
Figure 289: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

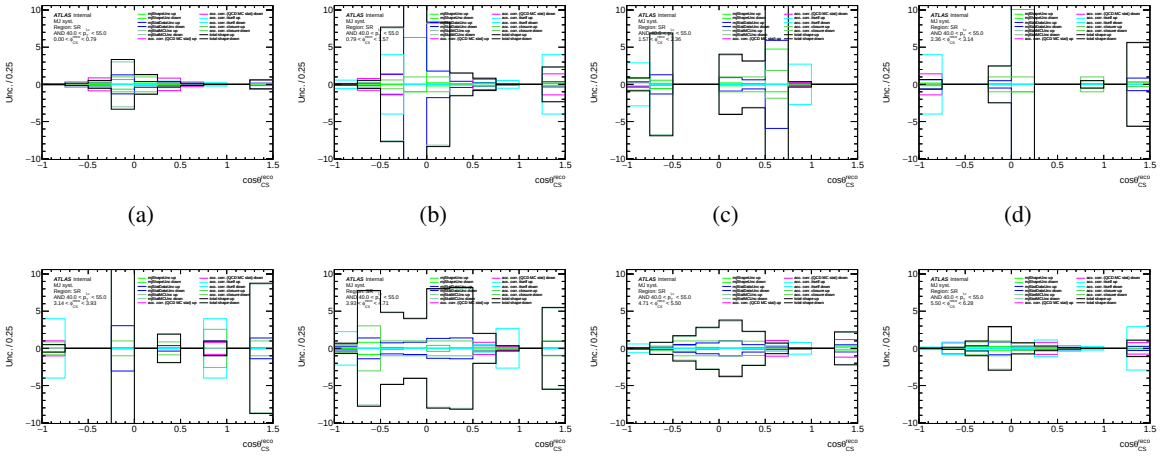
Figure 290: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

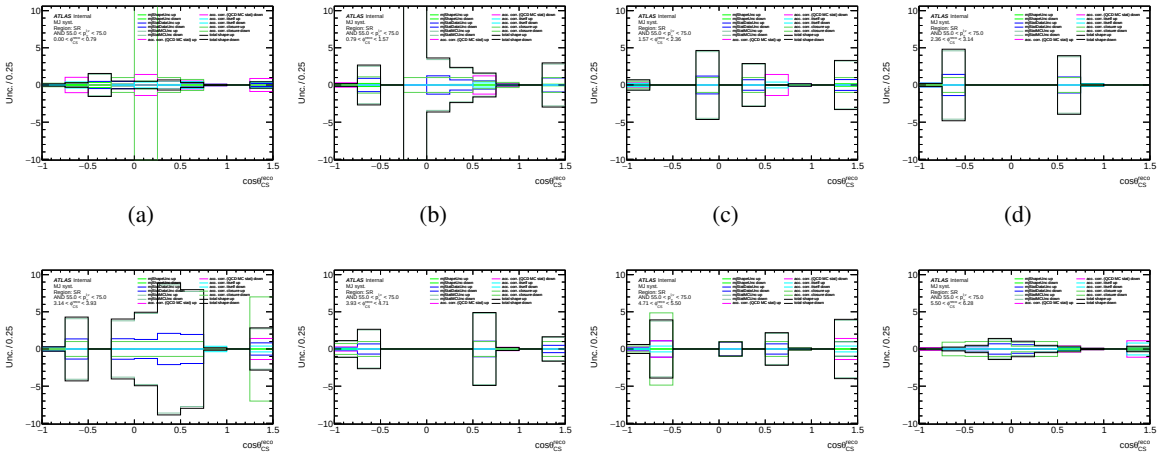
Figure 291: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

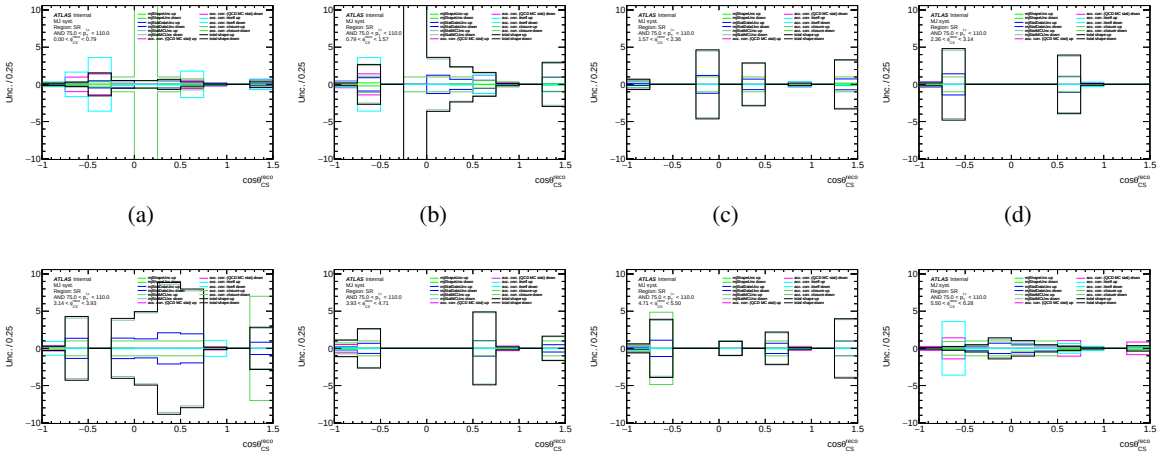
Figure 292: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \rightarrow \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

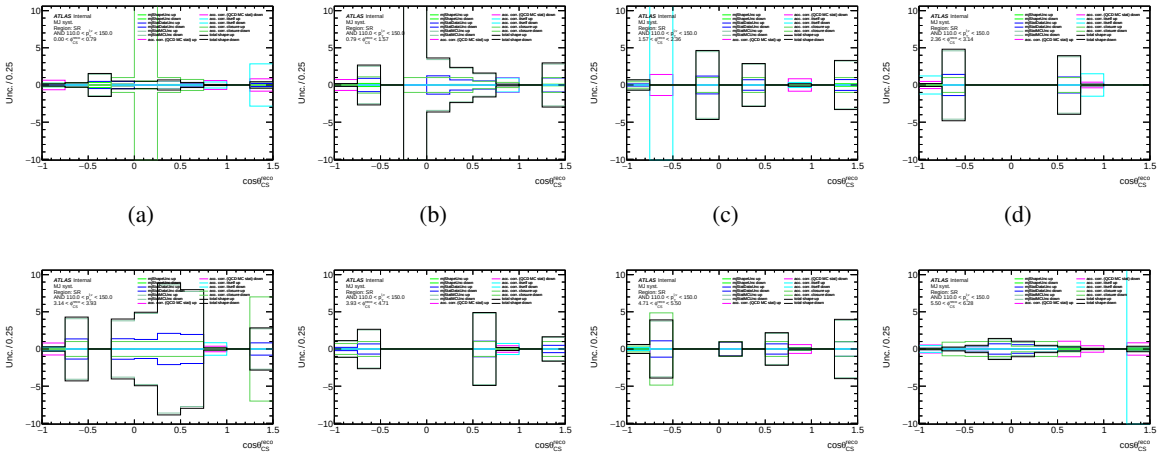
Figure 293: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \rightarrow \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

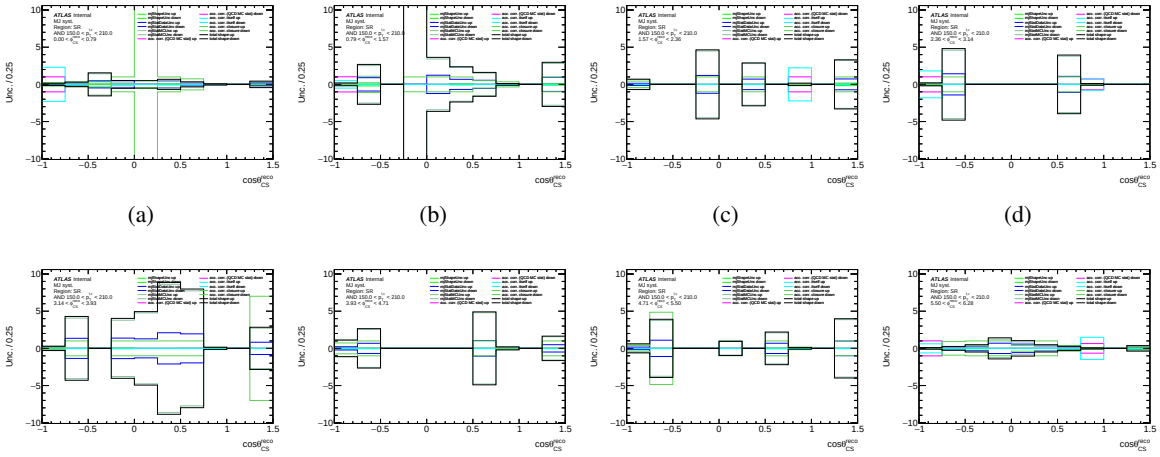
Figure 294: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

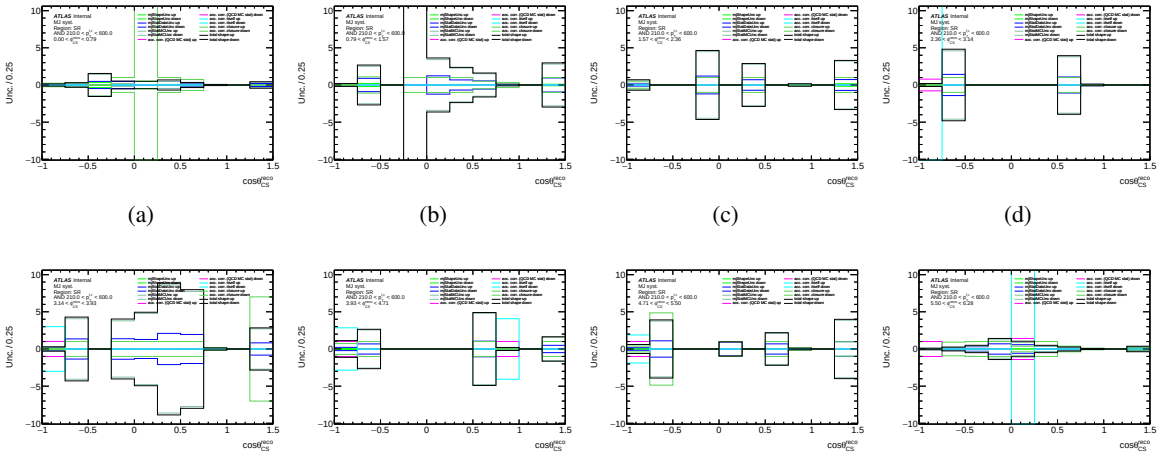
Figure 295: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

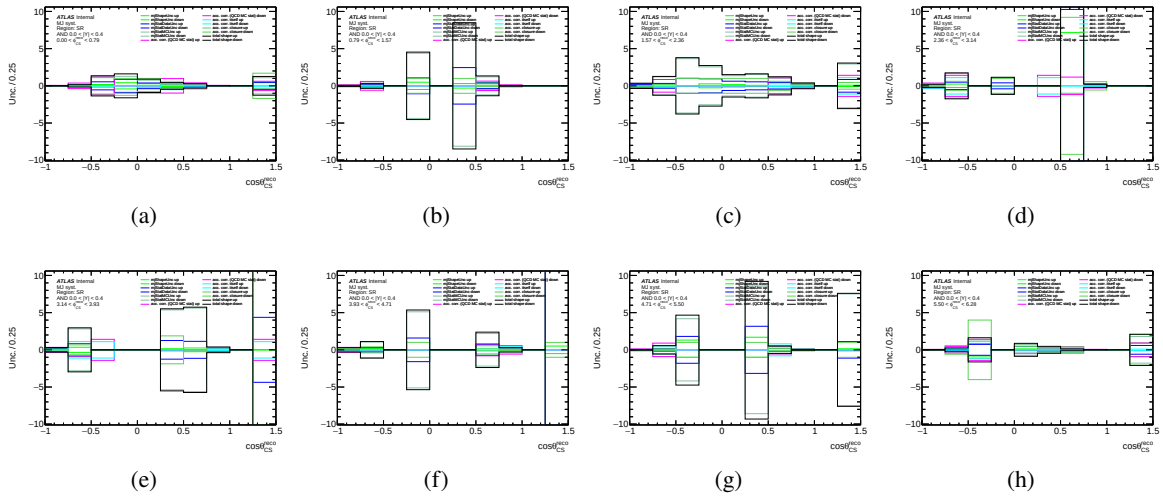
Figure 296: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

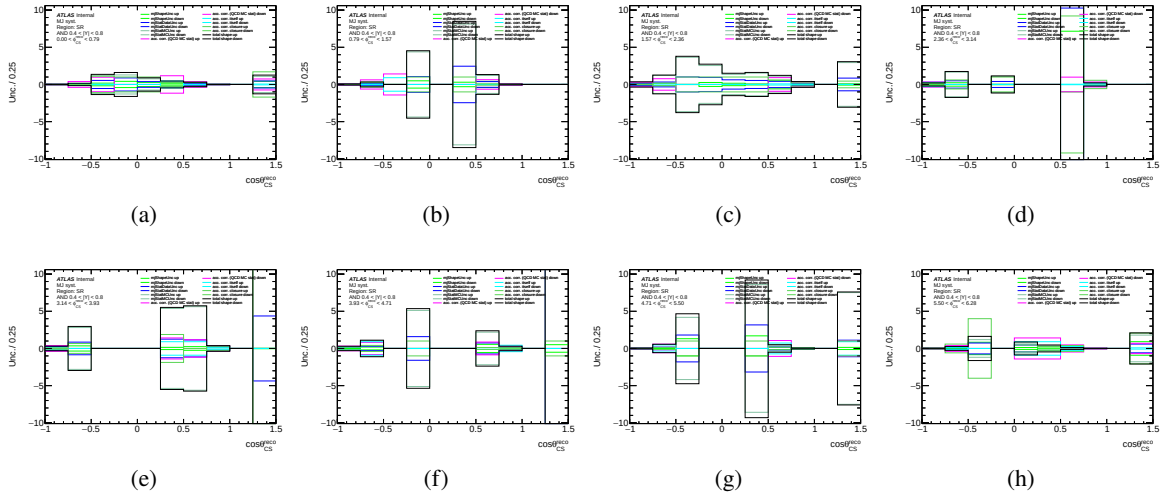
Figure 297: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

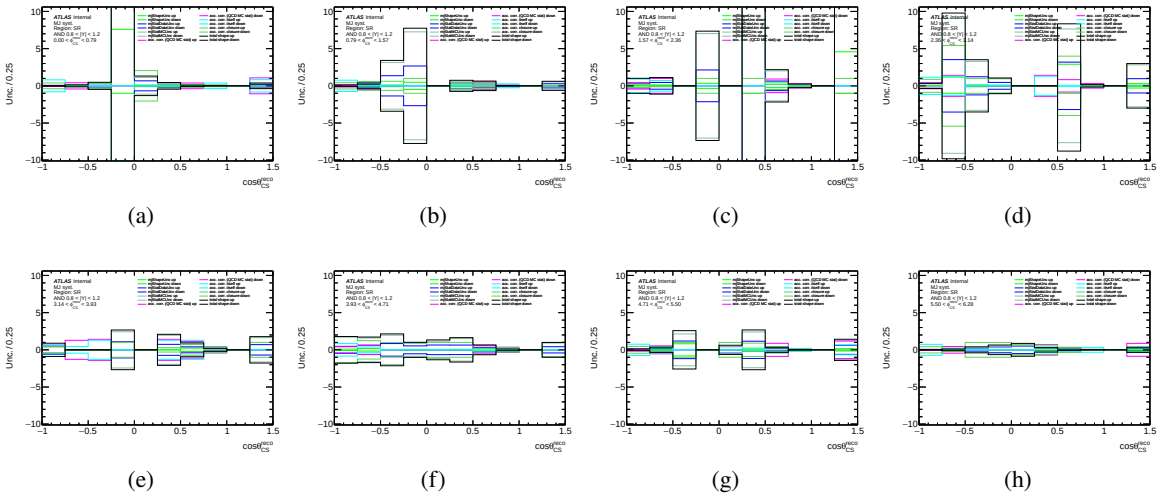
Figure 298: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0 < |Y| < 0.4$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

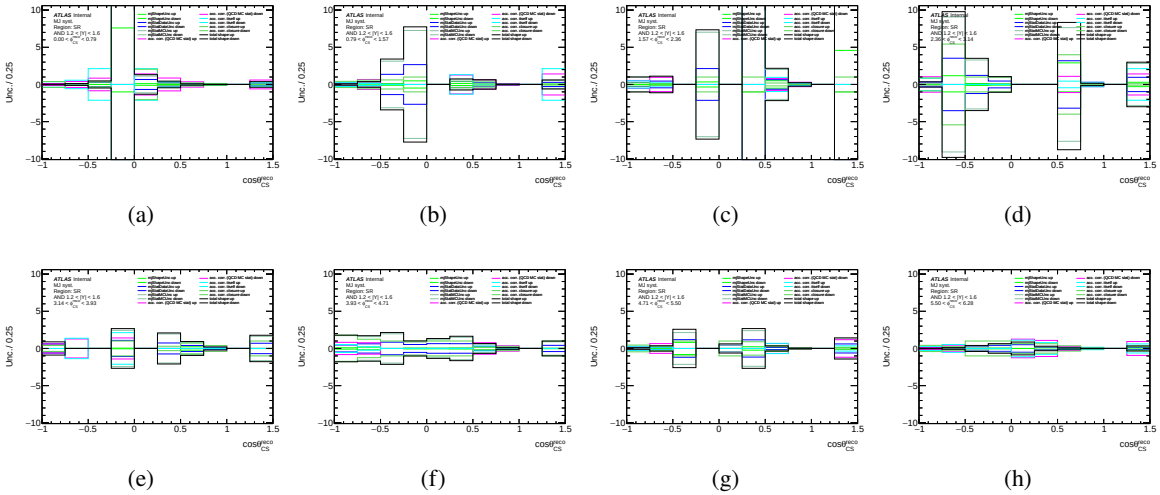
Figure 299: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.4 < |Y| < 0.8$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

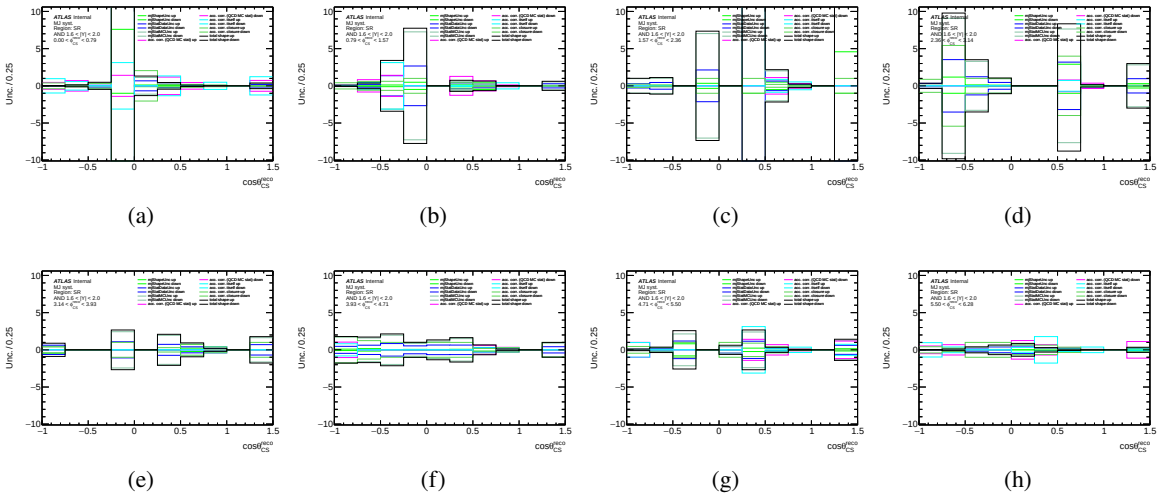
Figure 300: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $0.8 < |Y| < 1.2$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

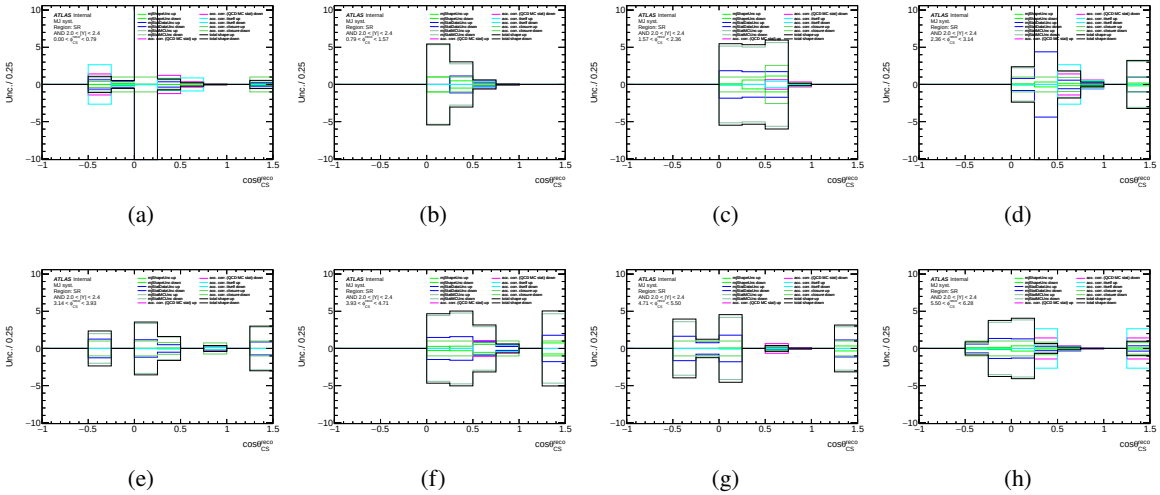
Figure 301: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.2 < |Y| < 1.6$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

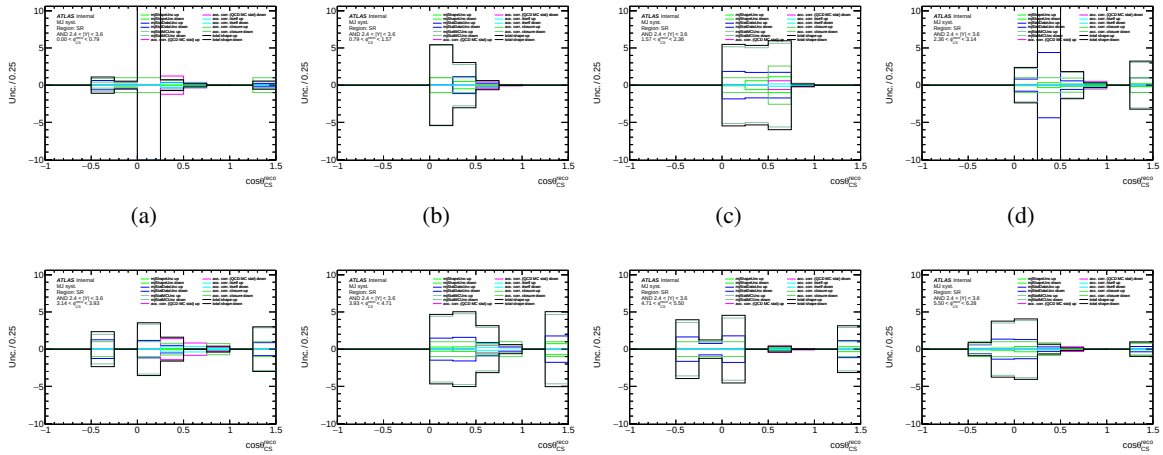
Figure 302: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $1.6 < |Y| < 2.0$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 303: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.0 < |Y| < 2.4$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 304: Multi-jet background shape systematics breakdown for $\cos\theta_{CS}$ as slices of ϕ_{CS} for $2.4 < |Y| < 3.6$ bin for $W^+ \to \mu^+\nu$ channel.

F Control Plots

Since analysis is based on 2D distibutions $\cos\theta_{CS}$ vs. ϕ_{CS}, to show Data/Model agreement we present control plots as slices of ϕ_{CS}. For the $W^- \to e^-\overline{\nu}$ channel Figures 305-314 shows results for $p_T^{\ell,\nu}$ bins and Figures 315-321 shows control plots for $|Y|$ bins. For the $W^- \to \mu^-\overline{\nu}$ channel Figures 322-331 shows results for $p_T^{\ell,\nu}$ bins and Figures 332-338 shows control plots for $|Y|$ bins. Double ratio e/μ plots are shown at Figures 390-399 for $p_T^{\ell,\nu}$ bins and on Figures 400-406 for $|Y|$ bins.

Only the statistical and MJ systematic uncertanties are shown.

F.1 Electron channel

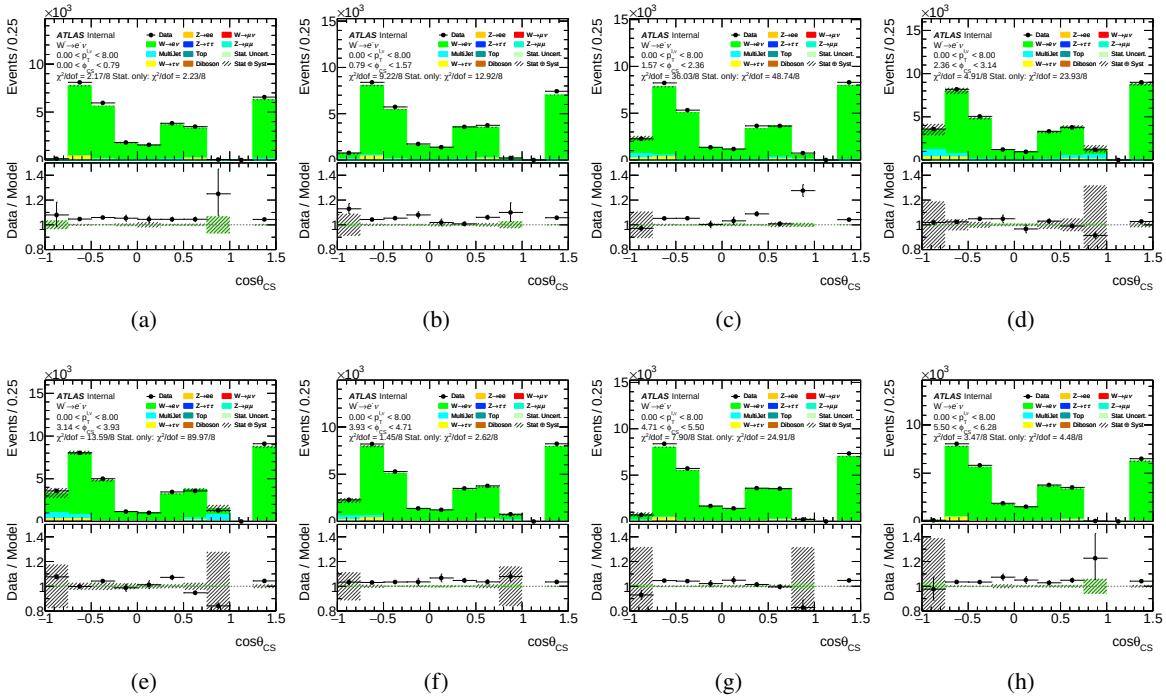
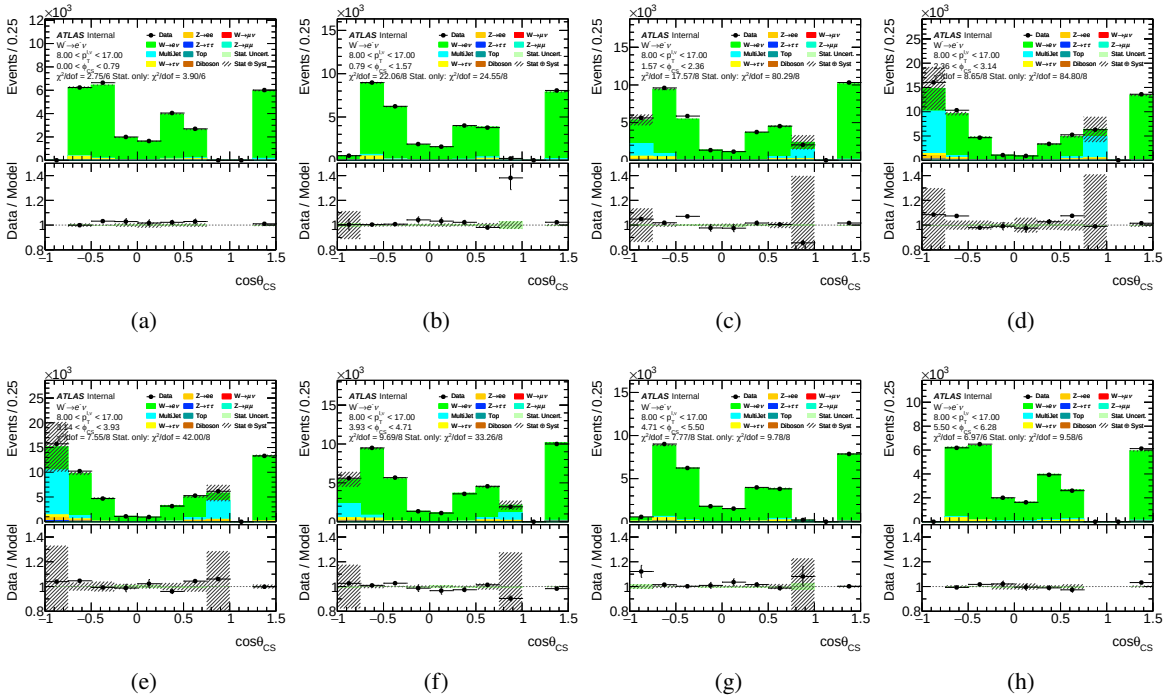


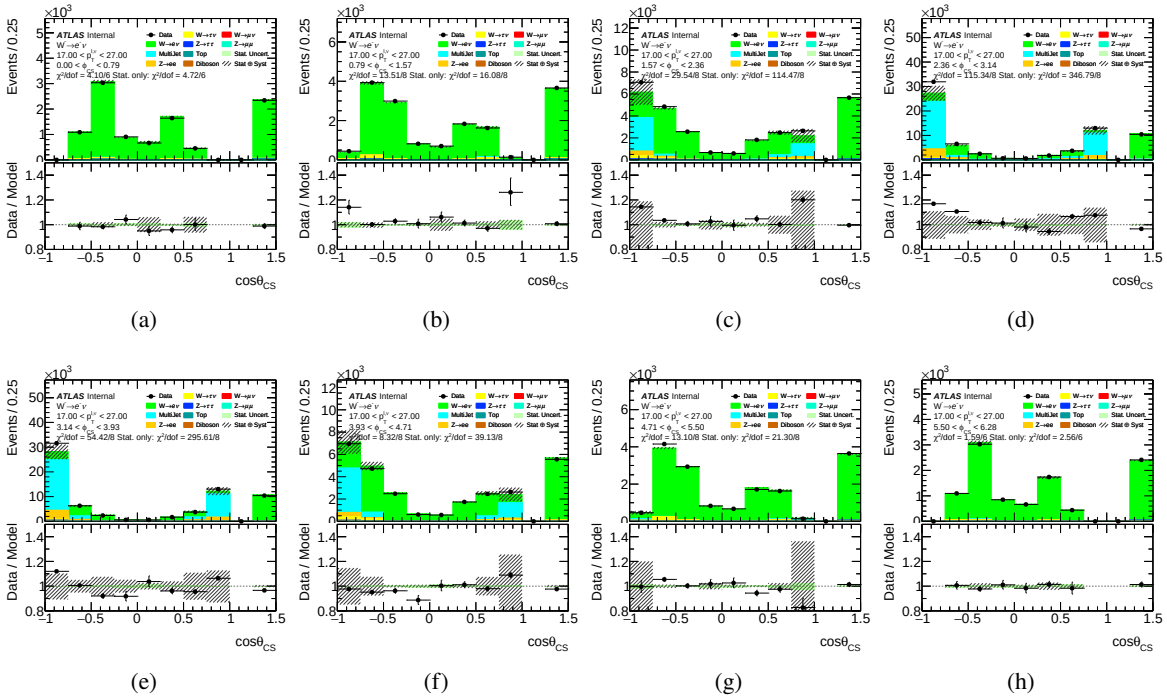
Figure 305: Control plots for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 306: Control plots for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 307: Control plots for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.

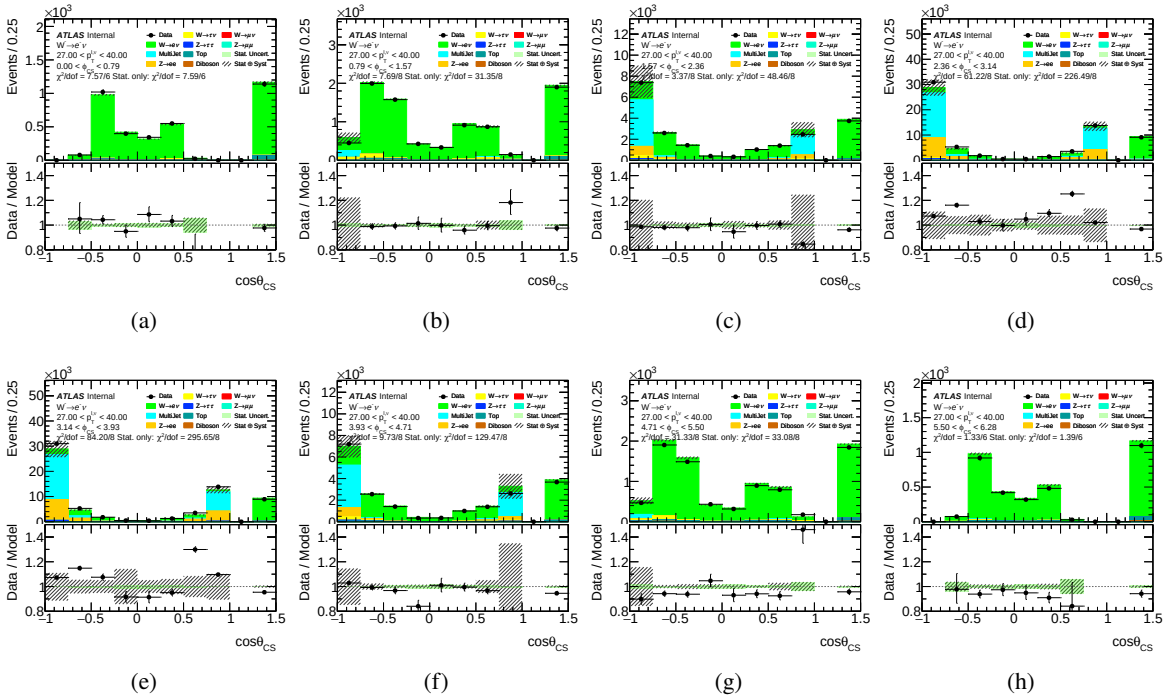


Figure 308: Control plots for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.

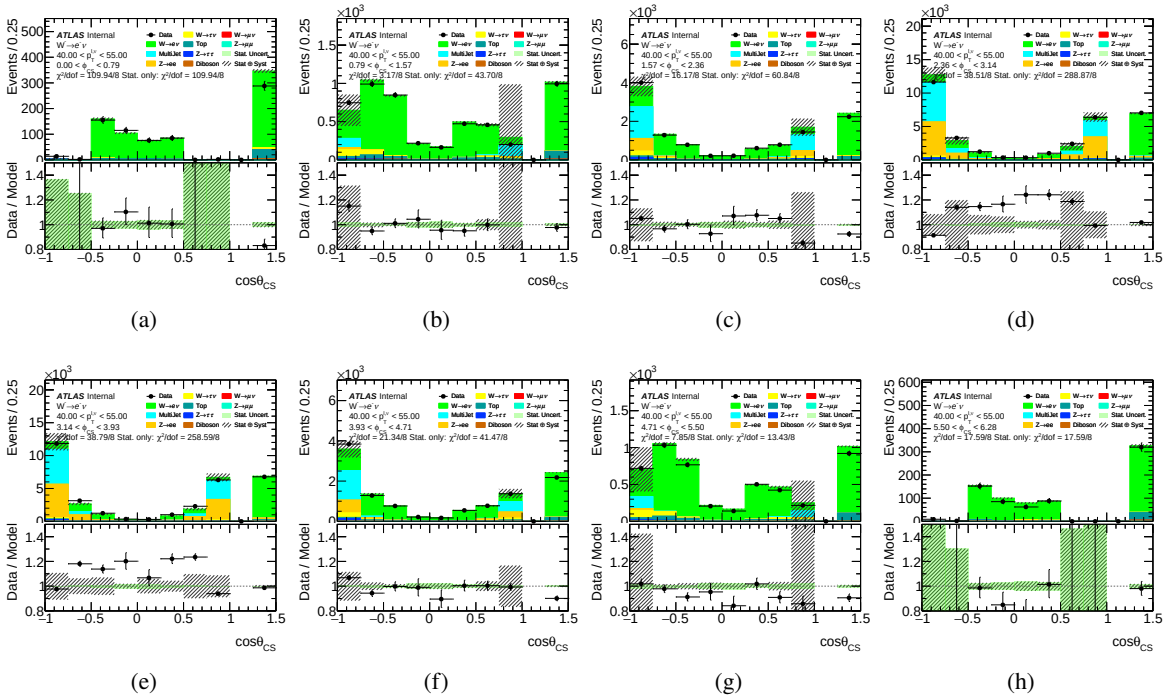


Figure 309: Control plots for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.

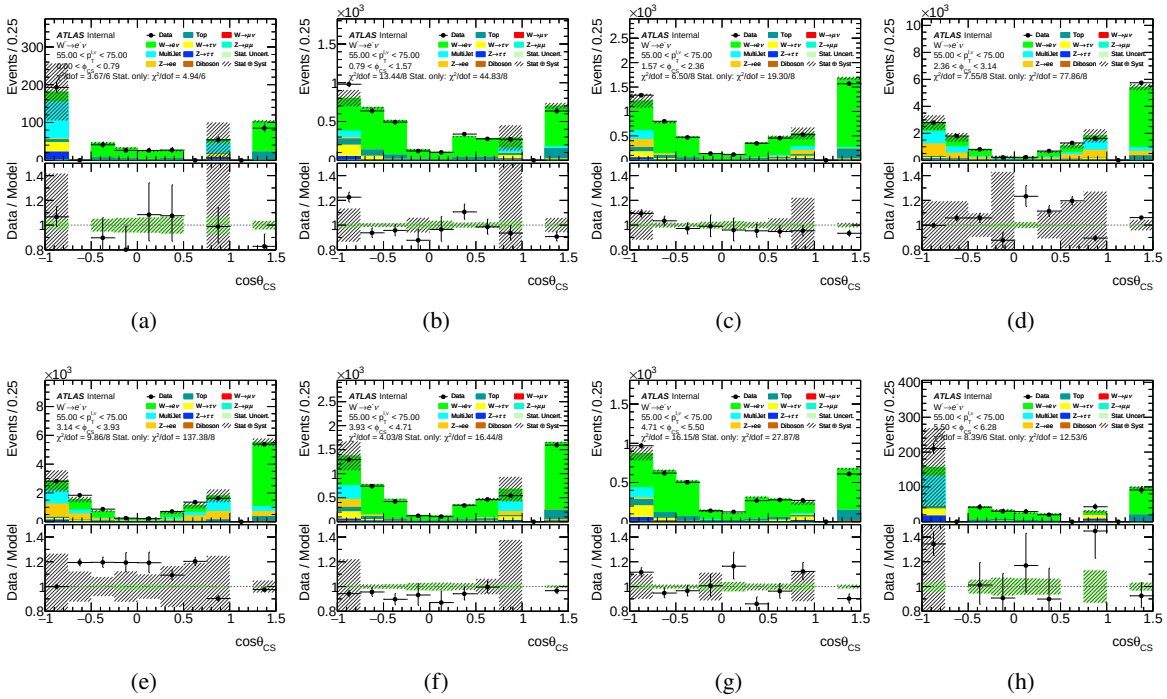


Figure 310: Control plots for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.

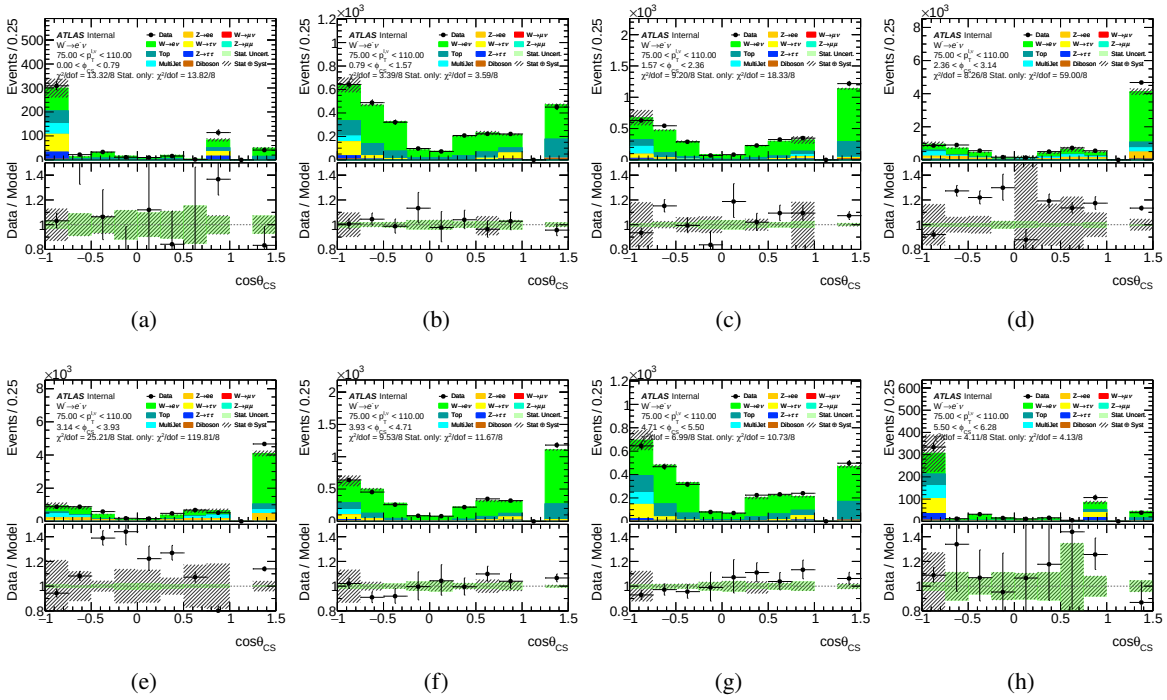
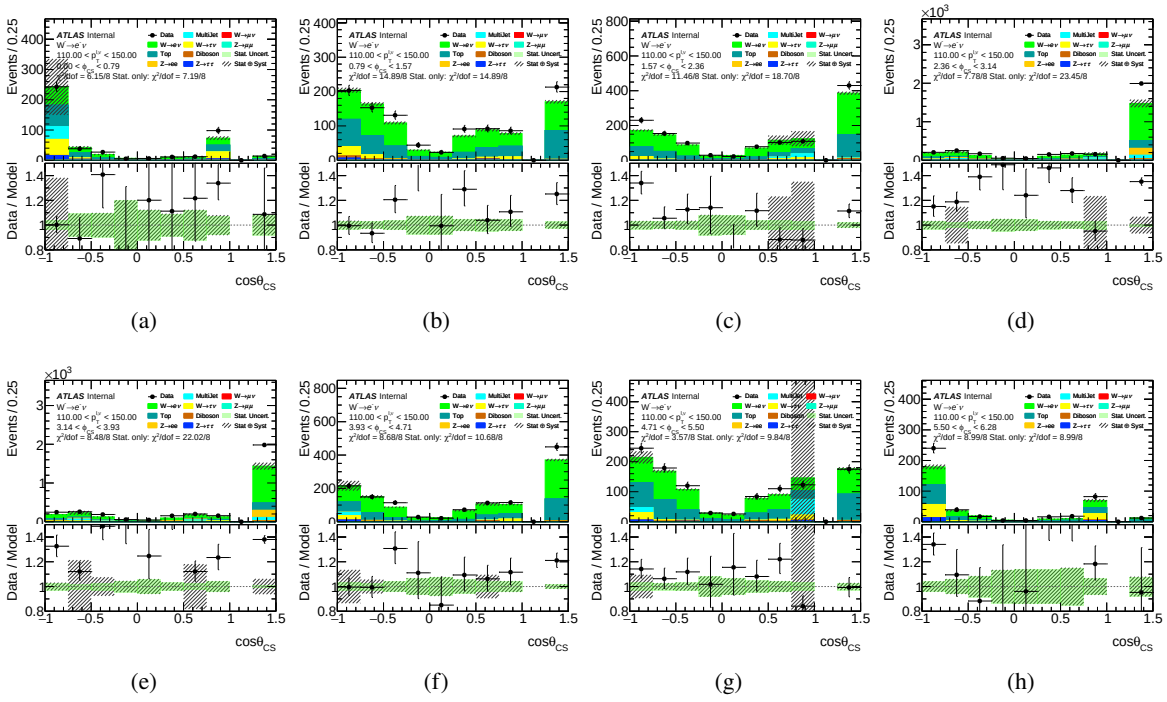


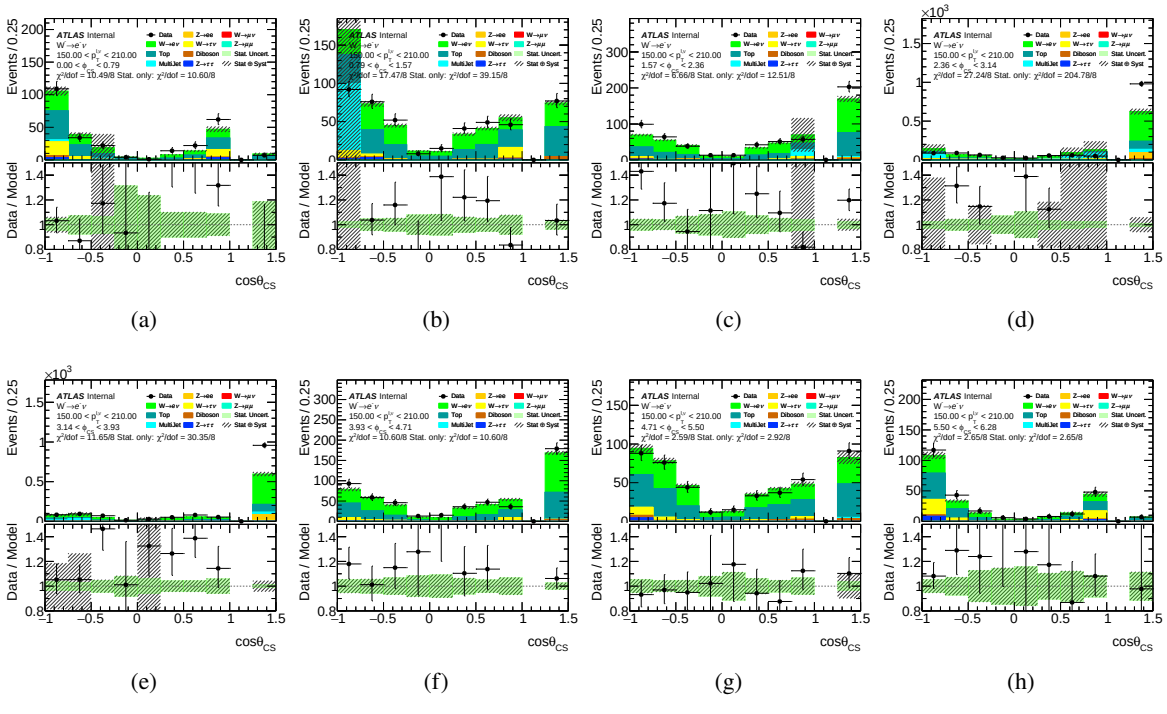
Figure 311: Control plots for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

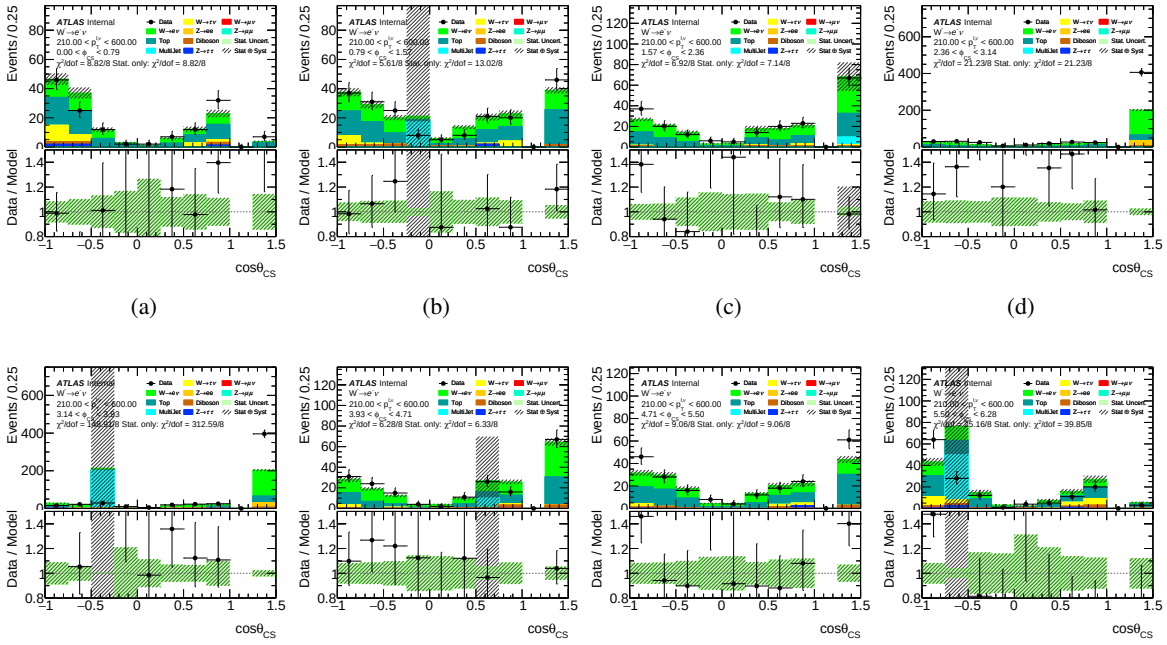
Figure 312: Control plots for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

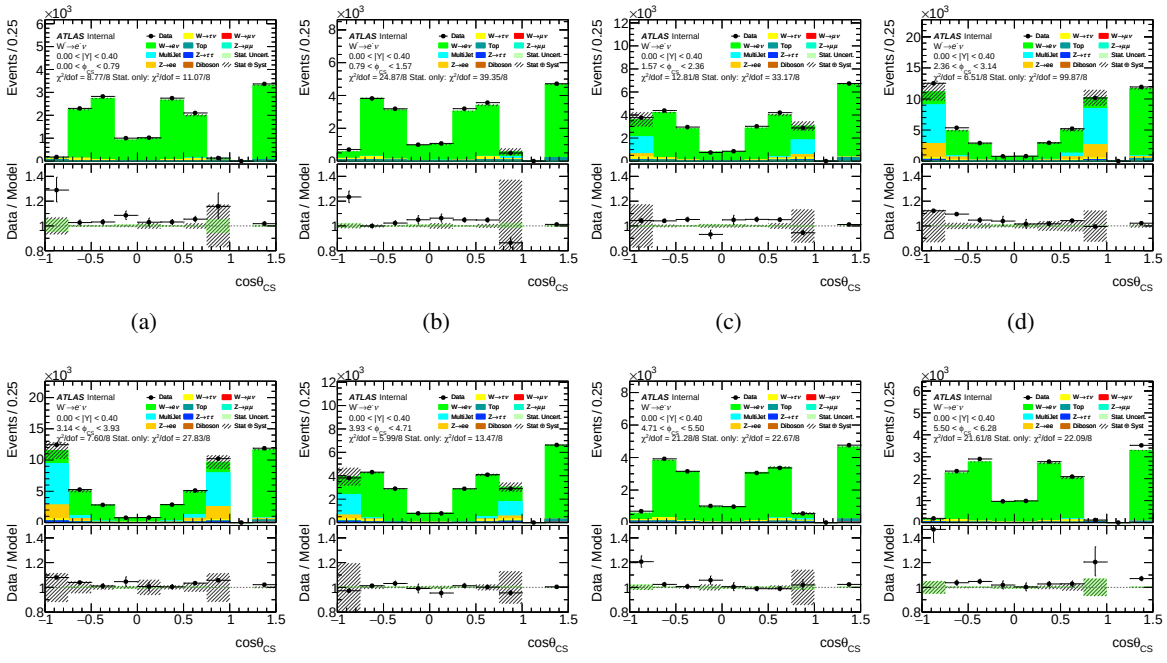
Figure 313: Control plots for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to e^-\overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

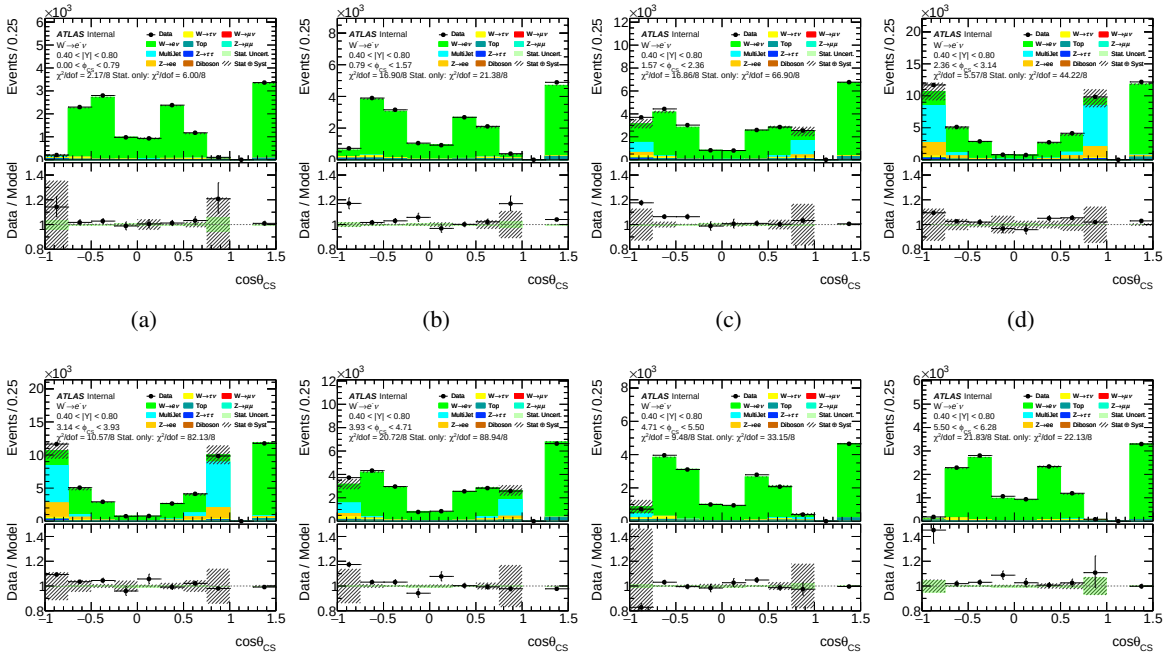
Figure 314: Control plots for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

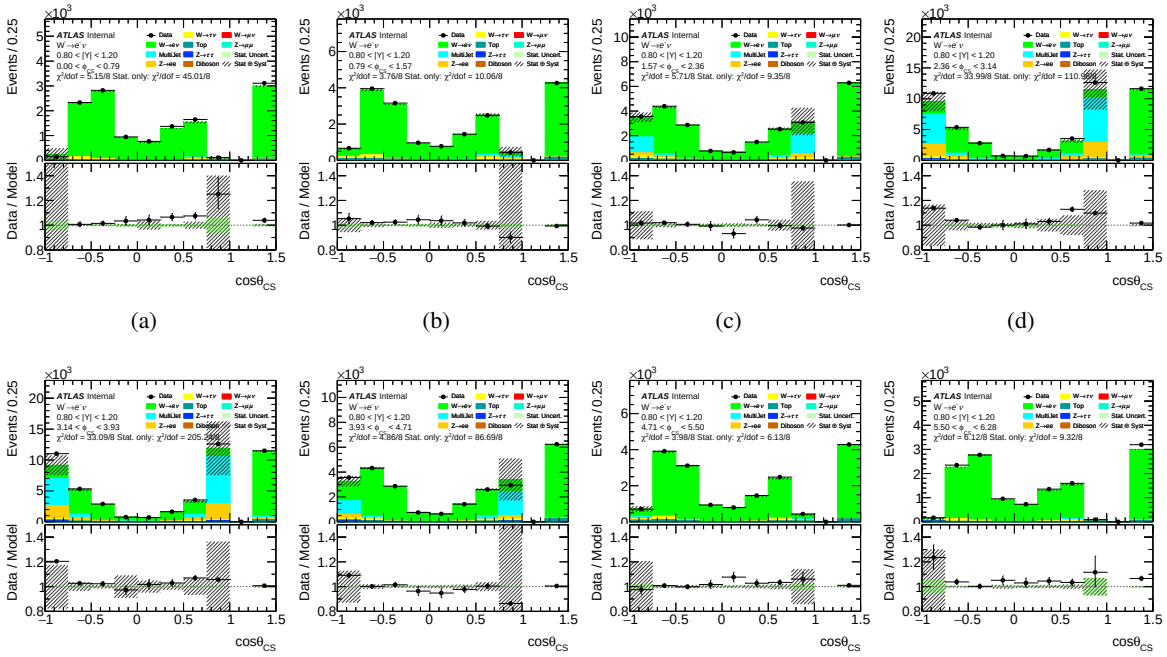
Figure 315: Control plots for $0 < |Y| < 0.4$ bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 316: Control plots for $0.4 < |Y| < 0.8$ bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 317: Control plots for $0.8 < |Y| < 1.2$ bin for $W^- \to e^- \overline{\nu}$ channel.

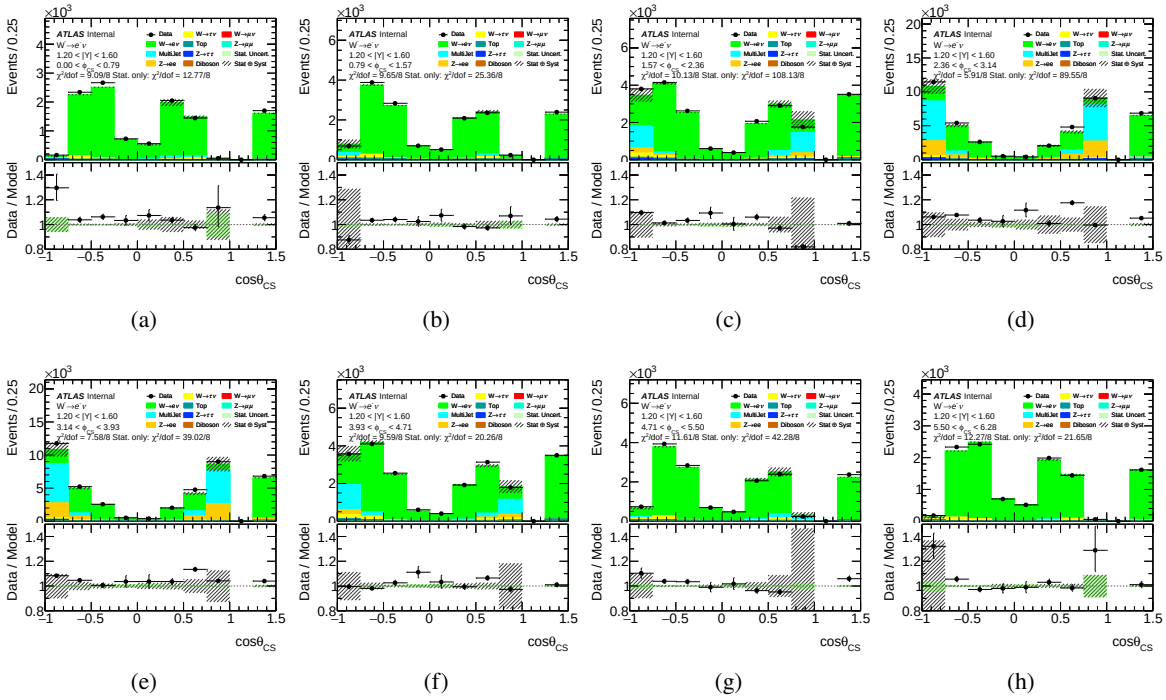


Figure 318: Control plots for $1.2 < |Y| < 1.6$ bin for $W^- \to e^- \overline{\nu}$ channel.

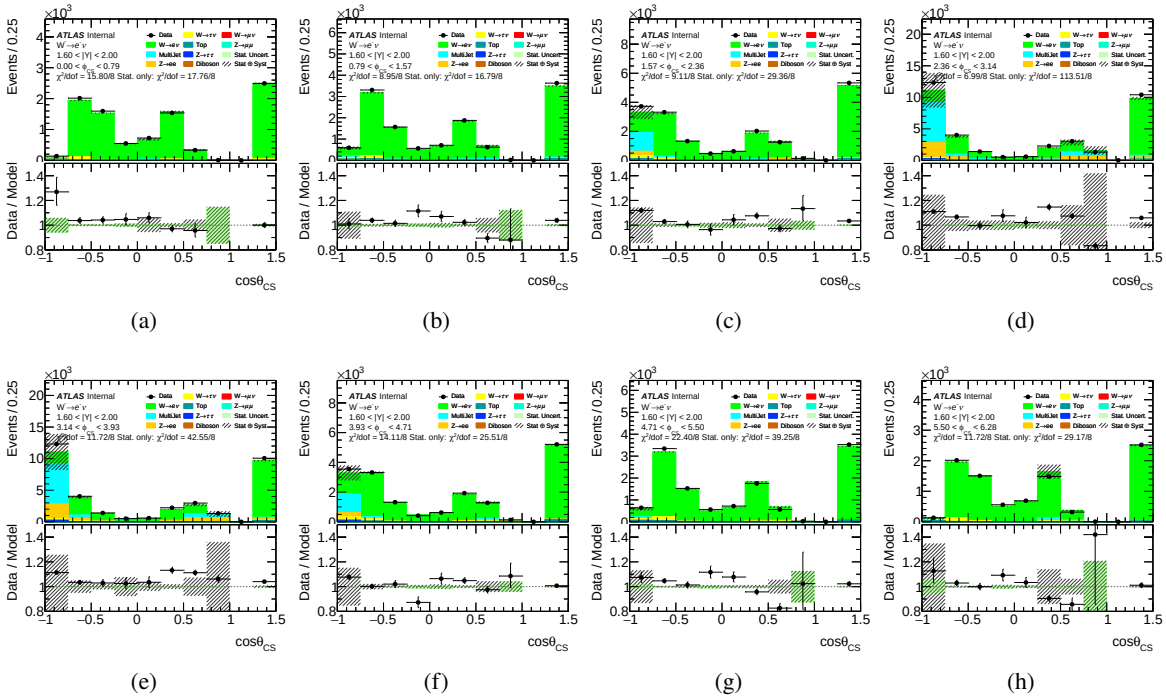
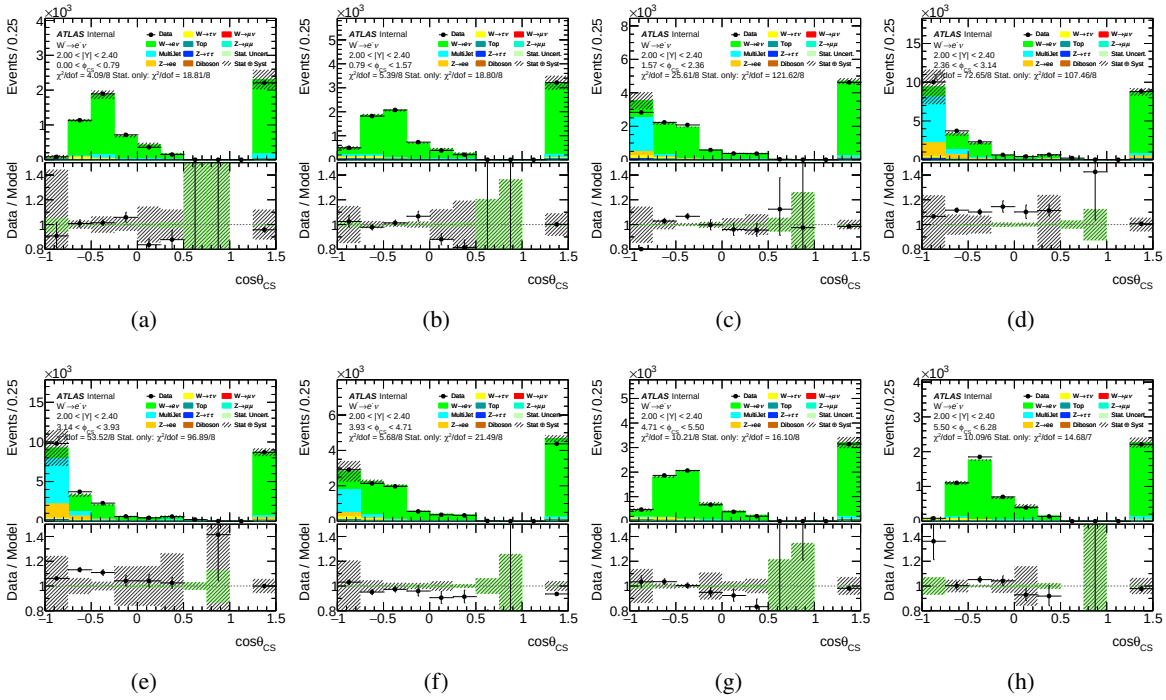


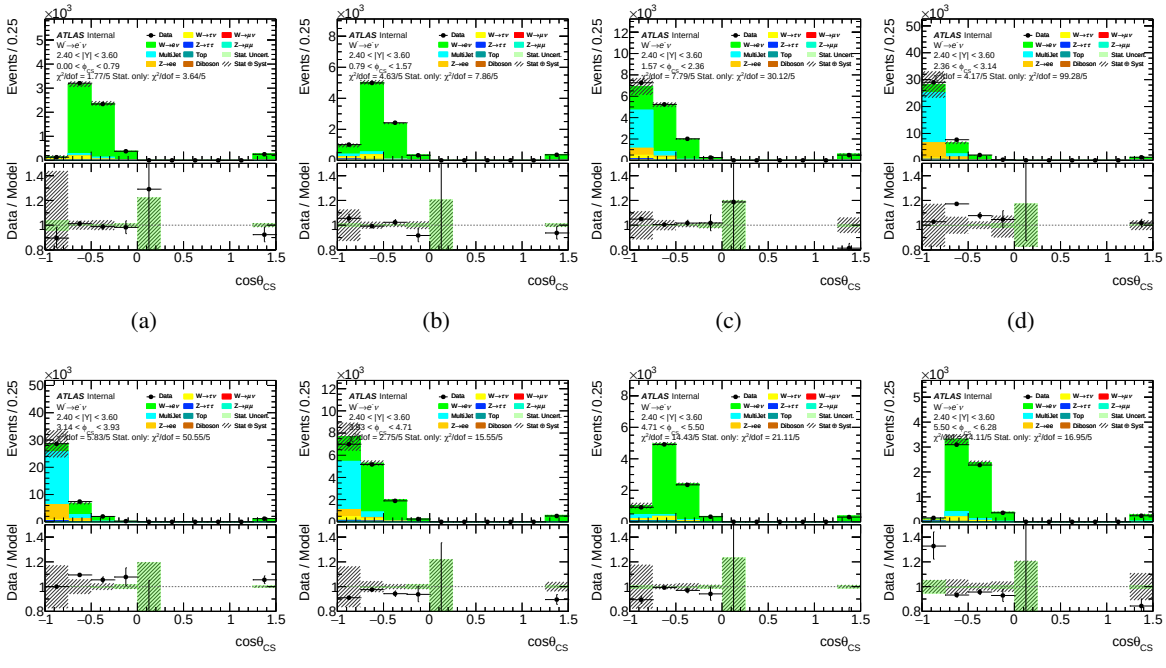
Figure 319: Control plots for $1.6 < |Y| < 2.0$ bin for $W^- \to e^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 320: Control plots for $2.0 < |Y| < 2.4$ bin for $W^- \to e^- \overline{\nu}$ channel.

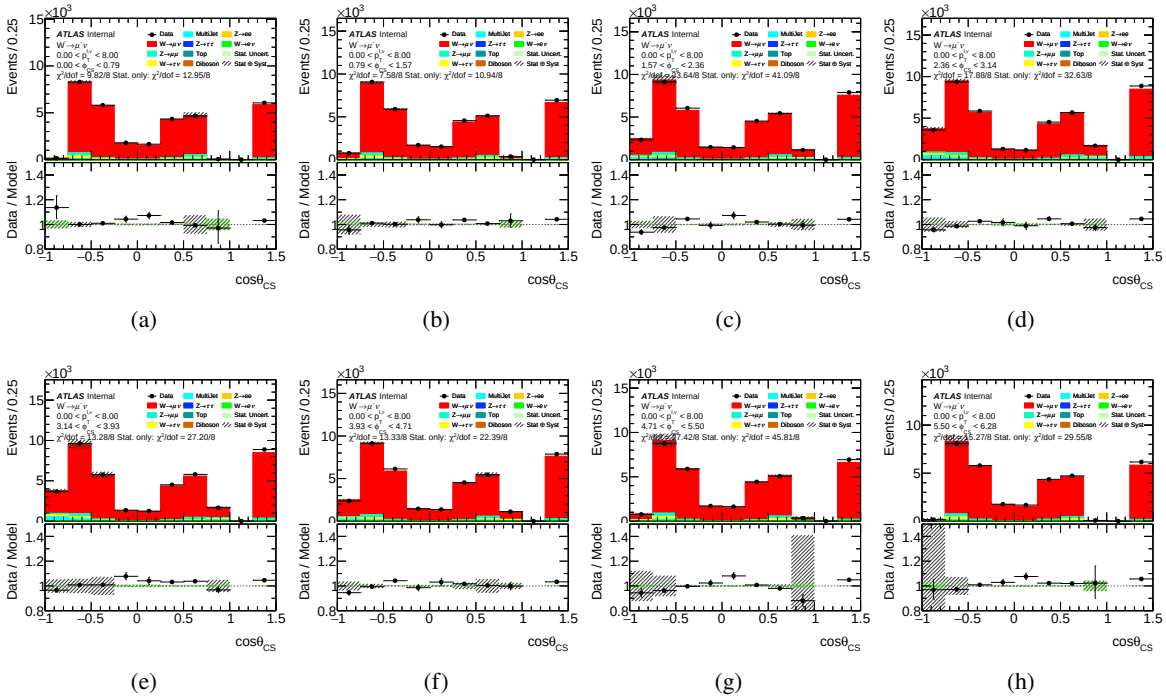


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 321: Control plots for $2.4 < |Y| < 3.6$ bin for $W^- \to e^- \overline{\nu}$ channel.

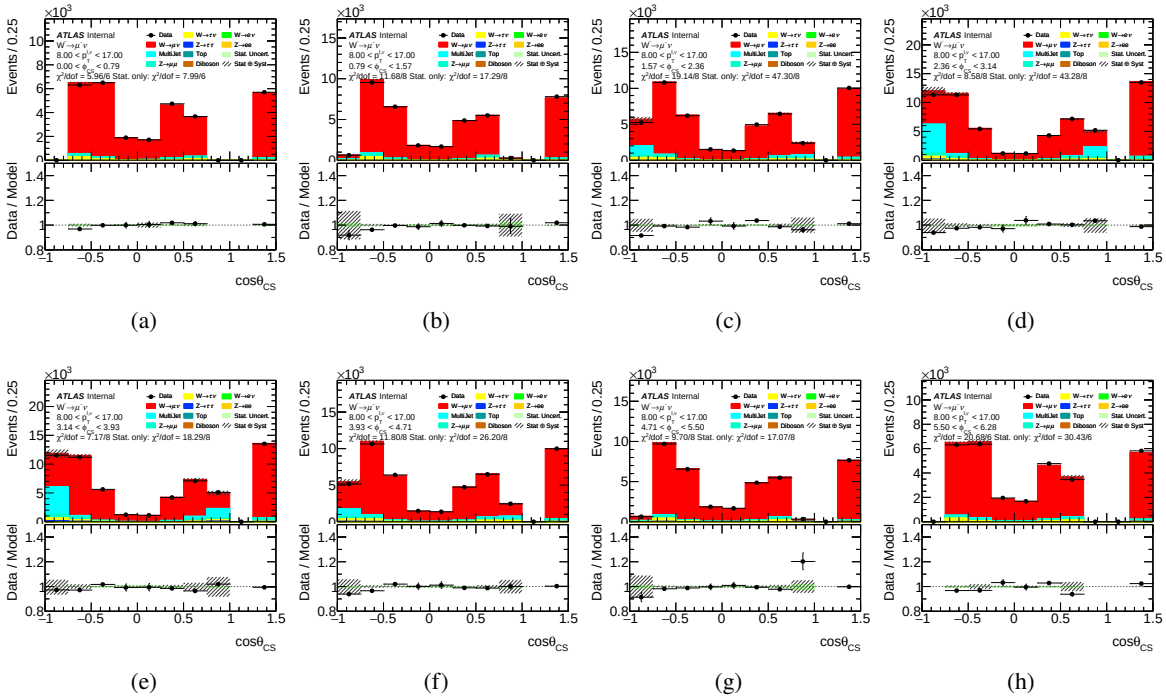
F.2 Muon channel



(a) (b) (c) (d)

(e) (f) (g) (h)

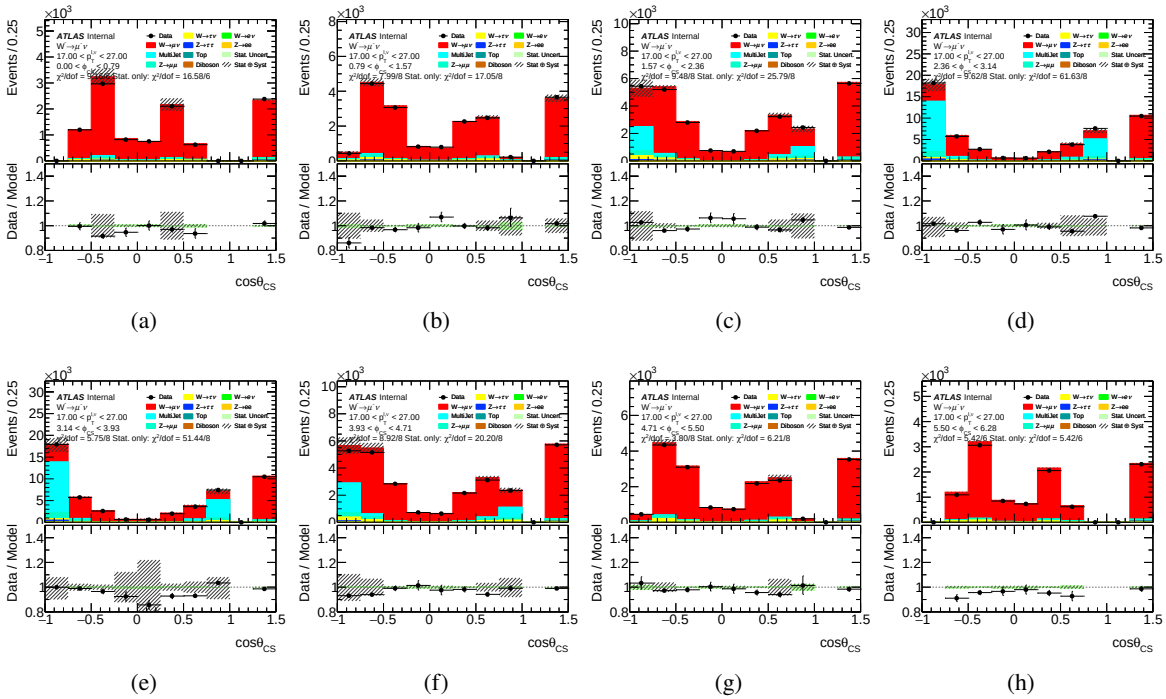
Figure 322: Control plots for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 323: Control plots for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 324: Control plots for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.

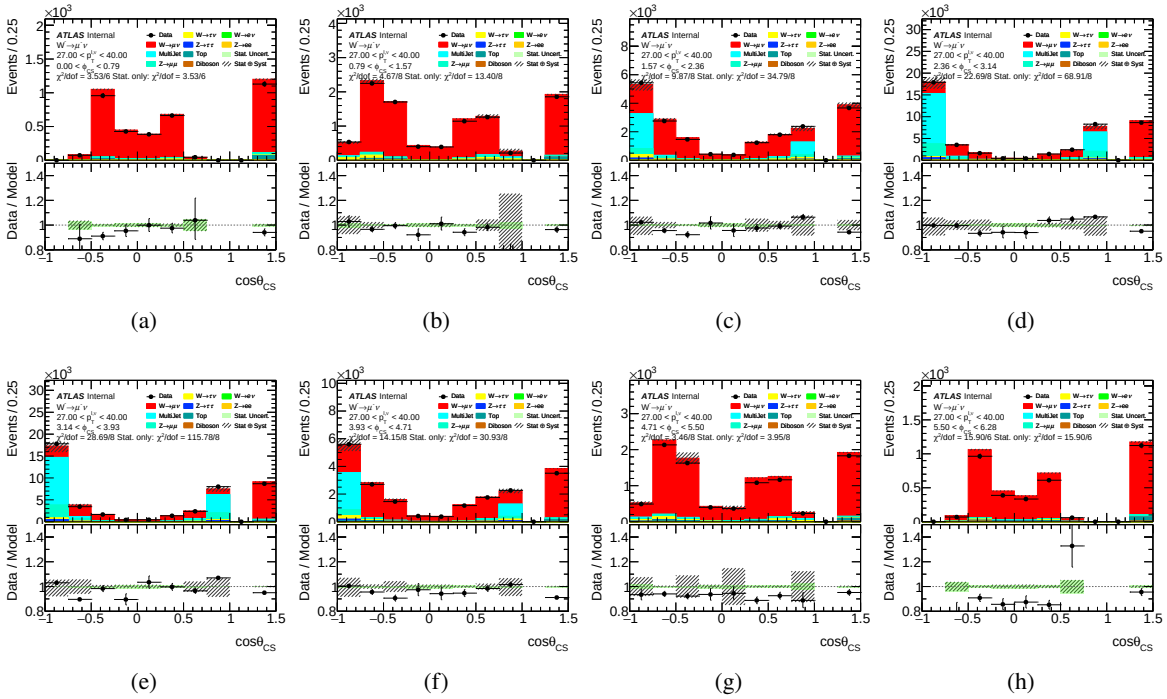


Figure 325: Control plots for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.

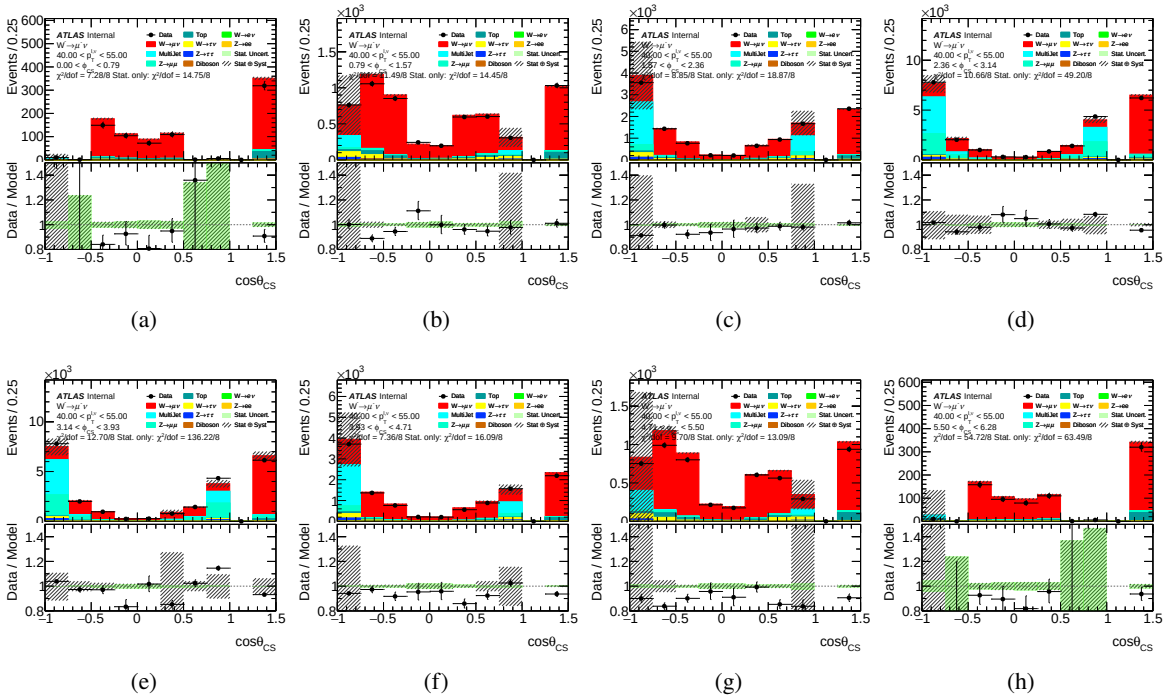
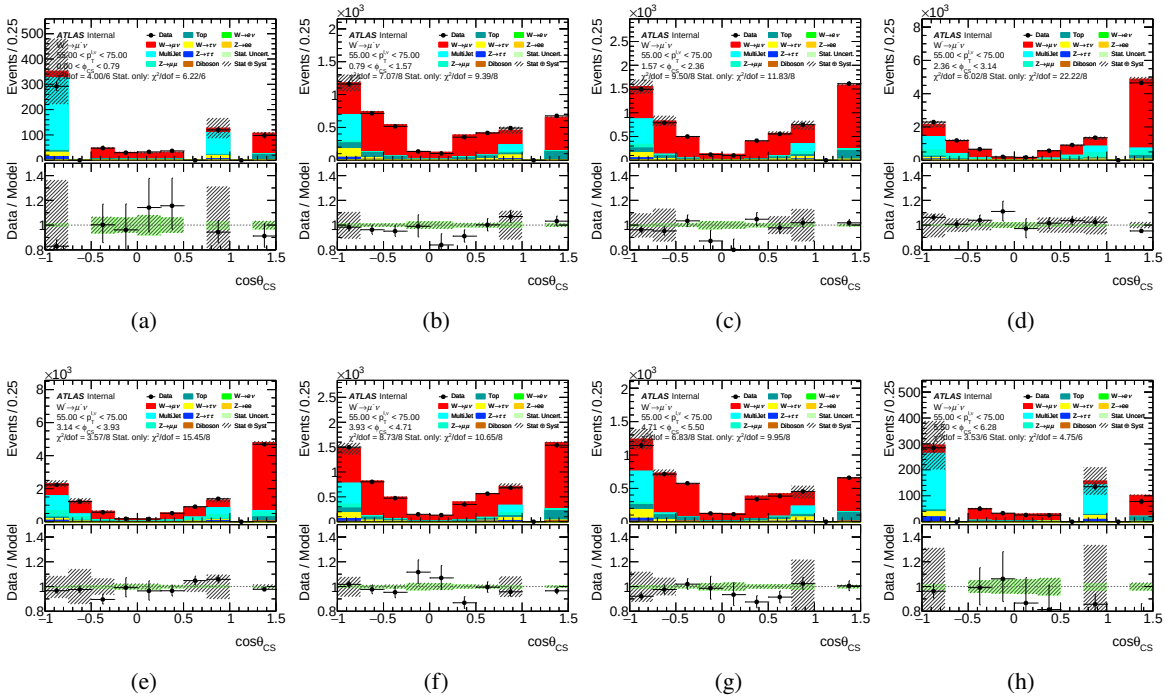


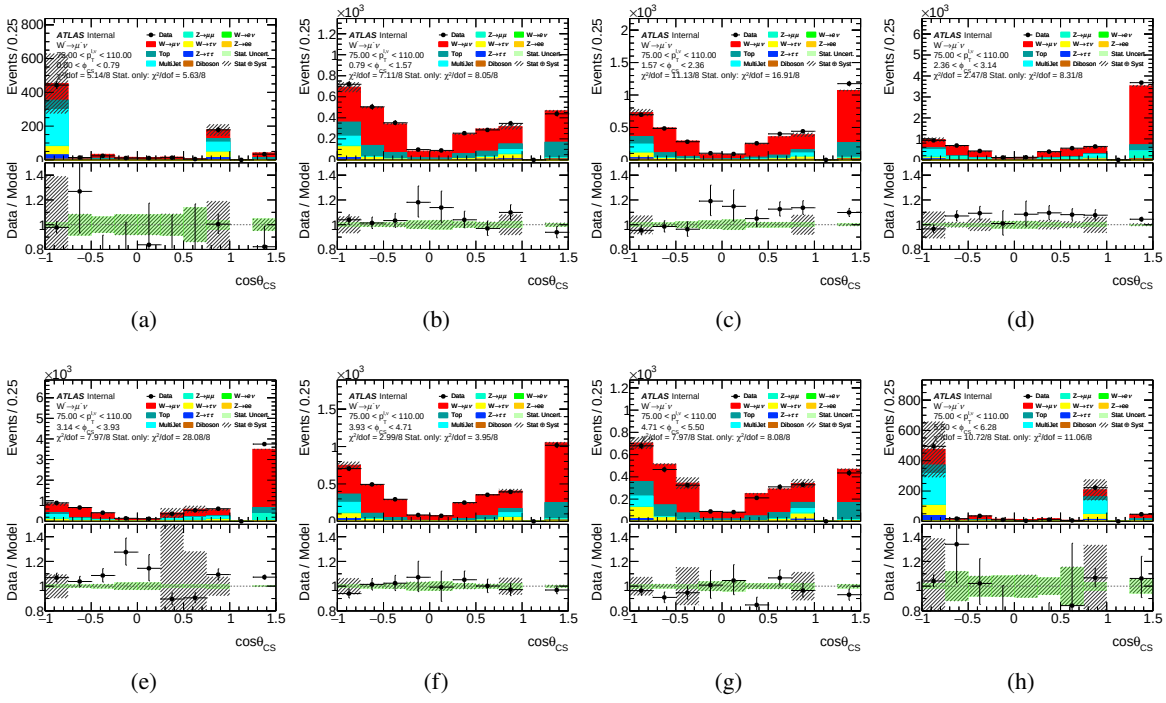
Figure 326: Control plots for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

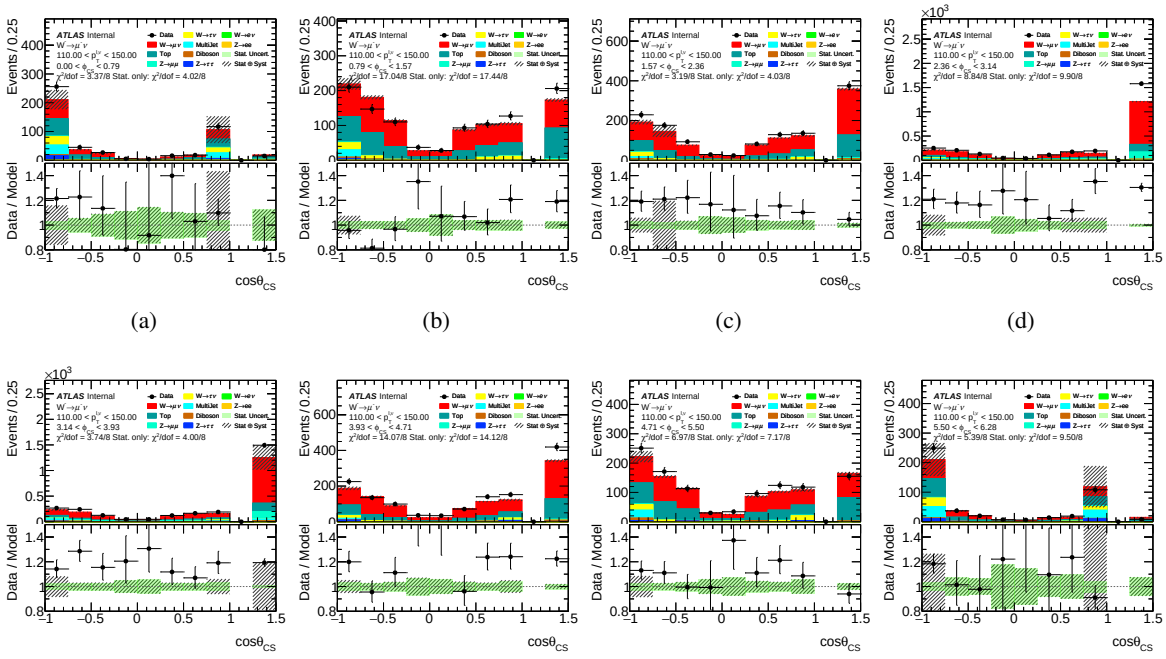
Figure 327: Control plots for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

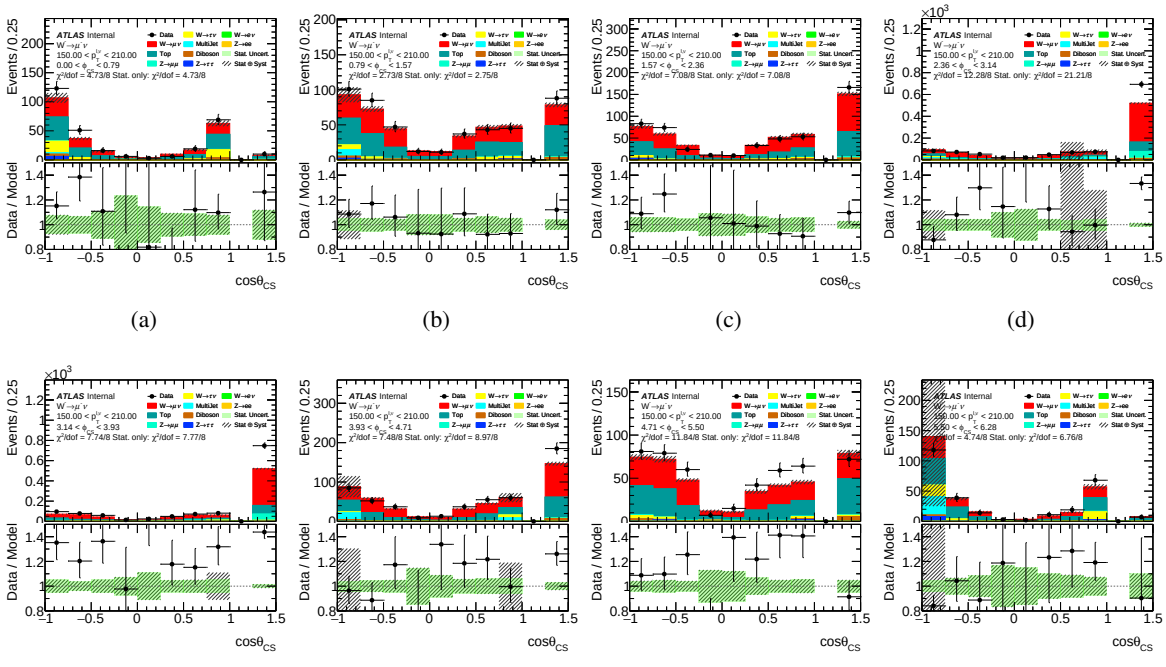
Figure 328: Control plots for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

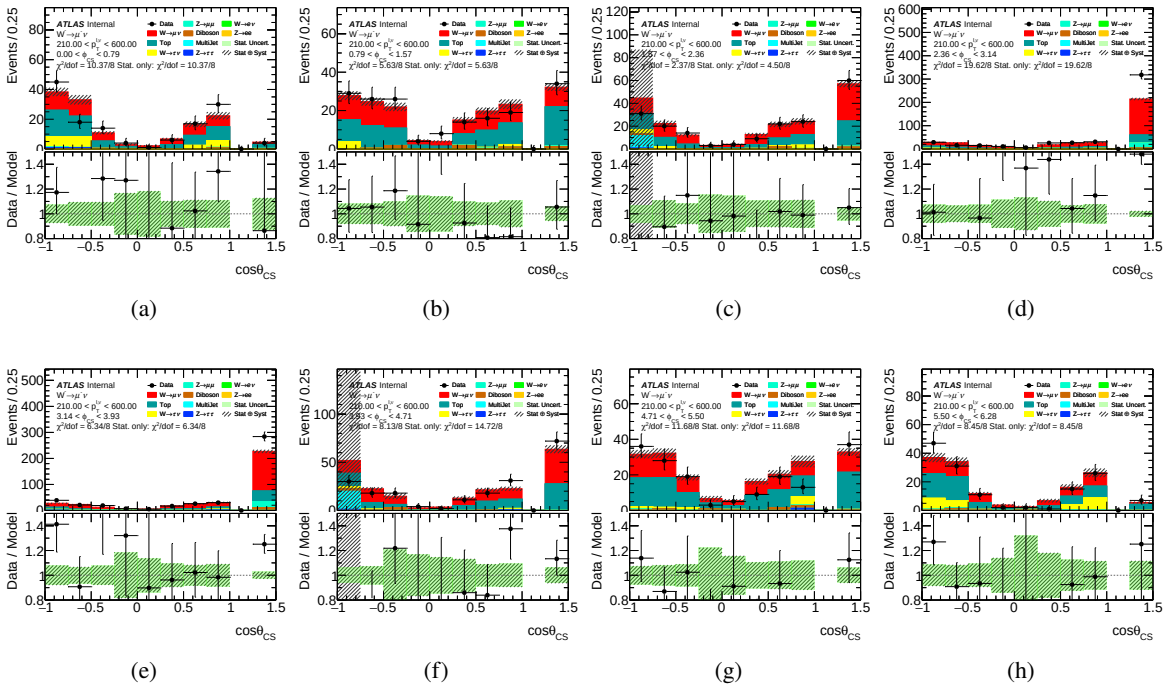
Figure 329: Control plots for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

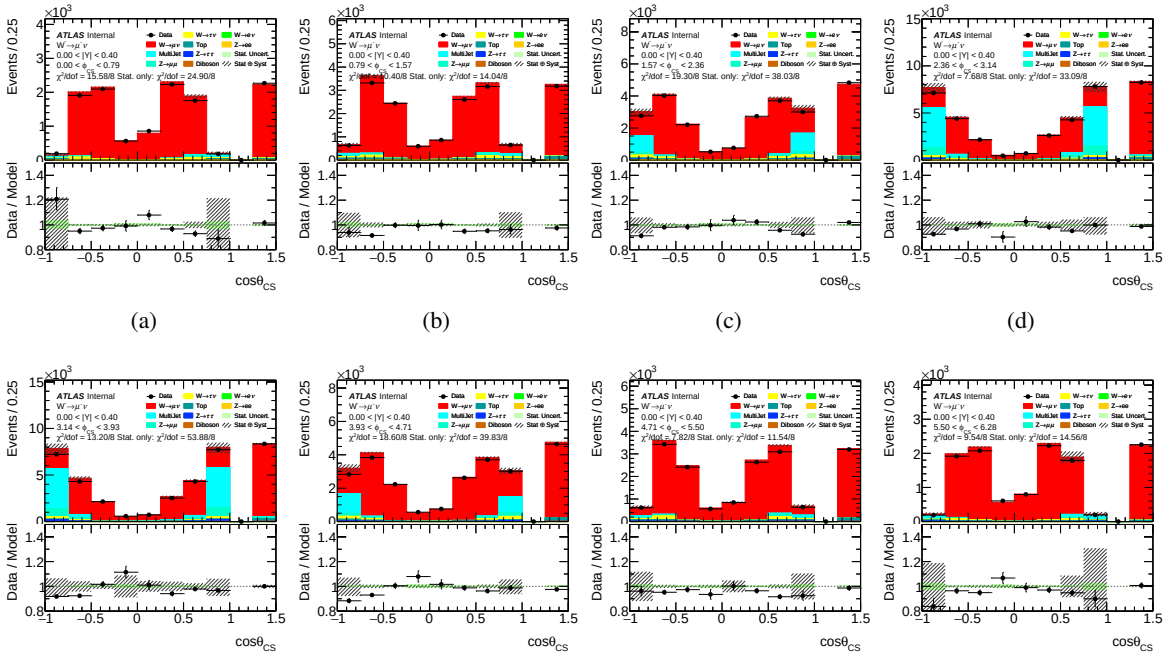
Figure 330: Control plots for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

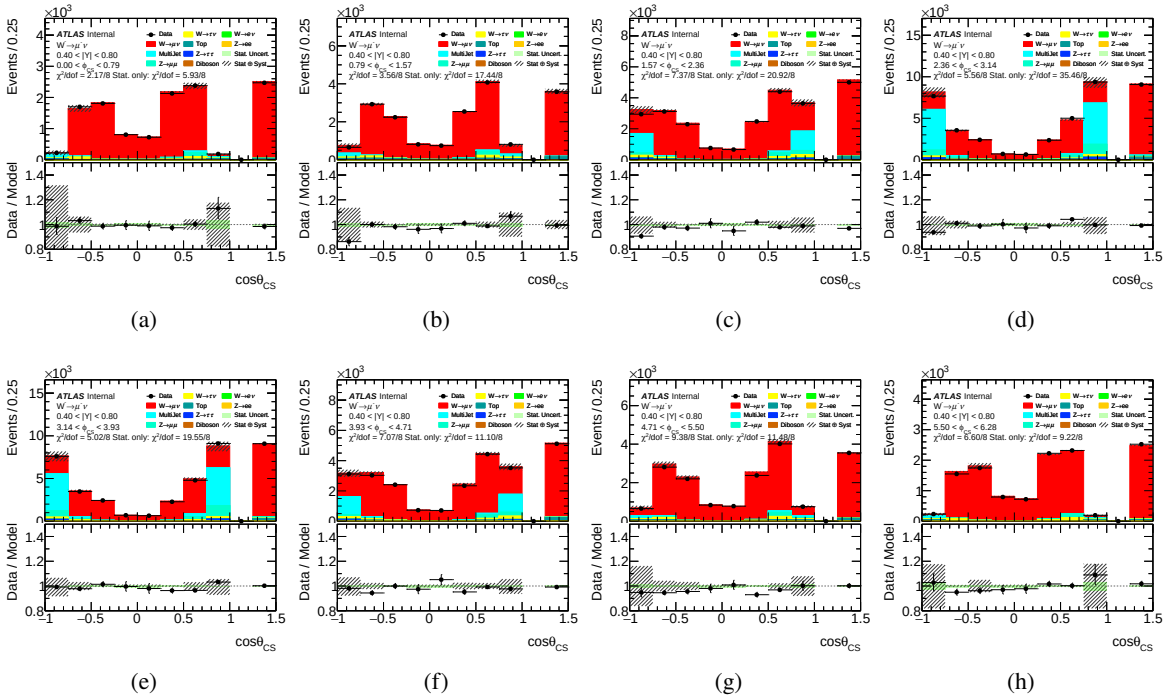
Figure 331: Control plots for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

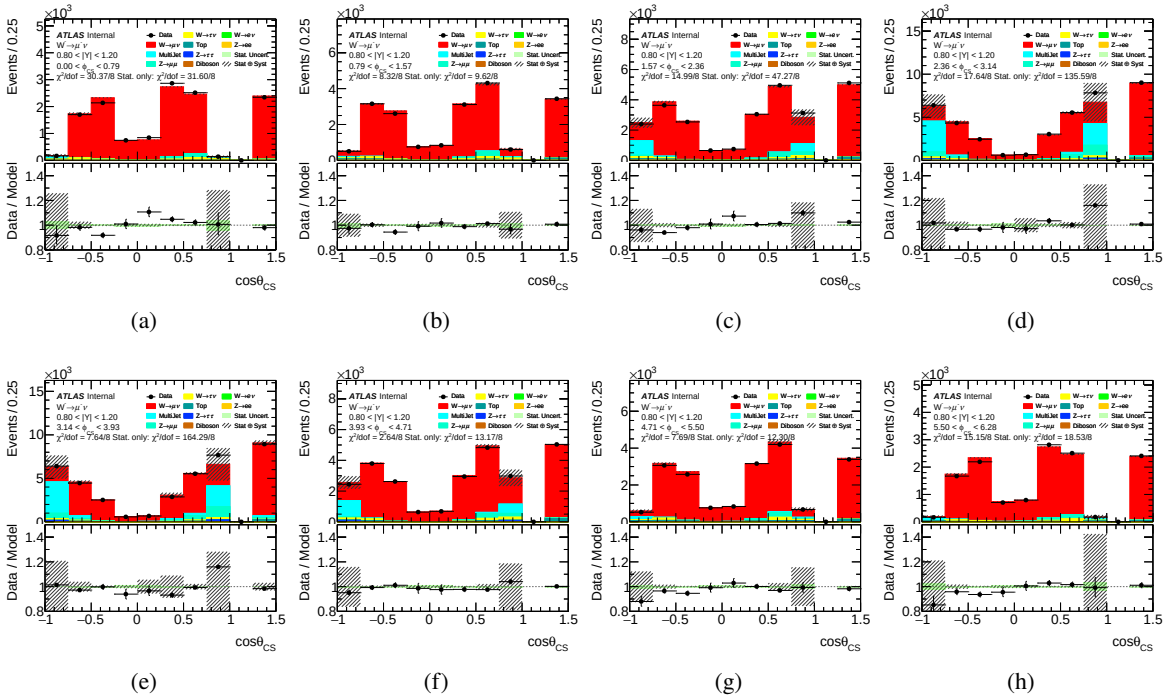
Figure 332: Control plots for $0 < |Y| < 0.4$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 333: Control plots for $0.4 < |Y| < 0.8$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 334: Control plots for $0.8 < |Y| < 1.2$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

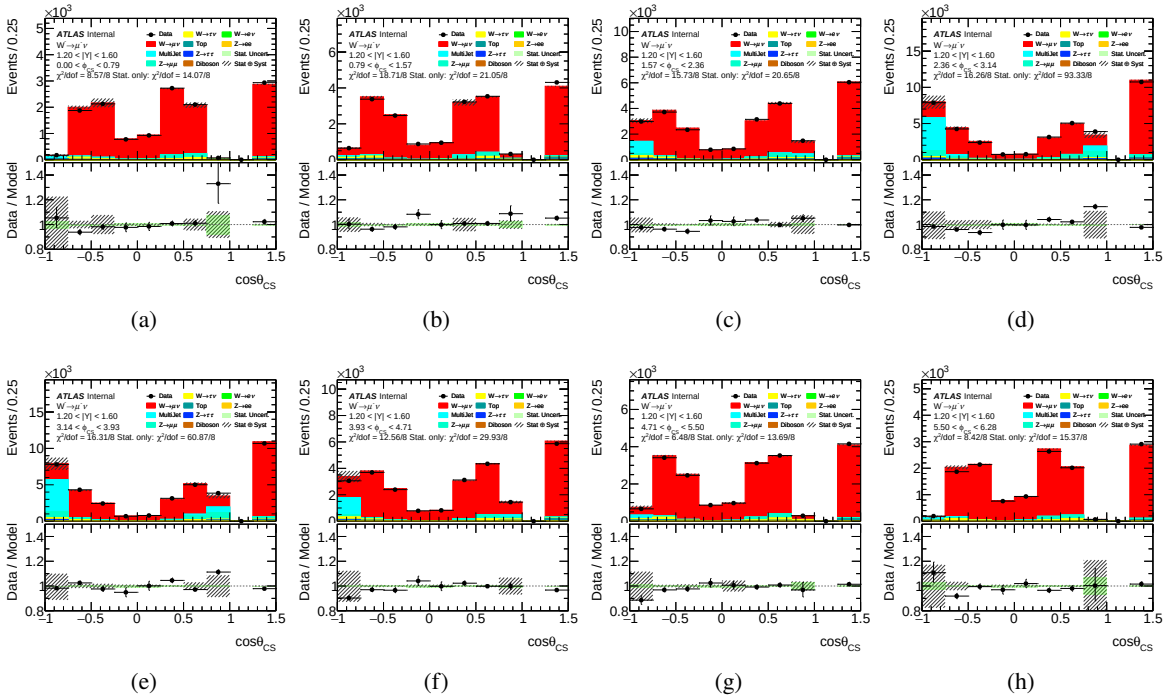


Figure 335: Control plots for $1.2 < |Y| < 1.6$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

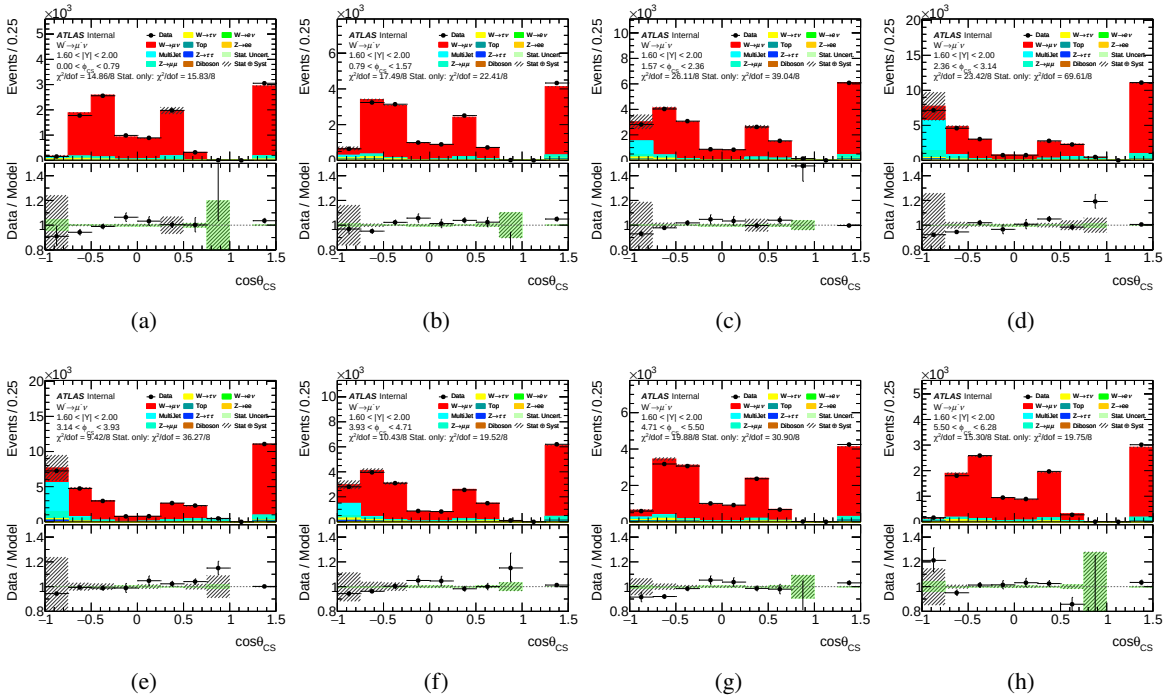
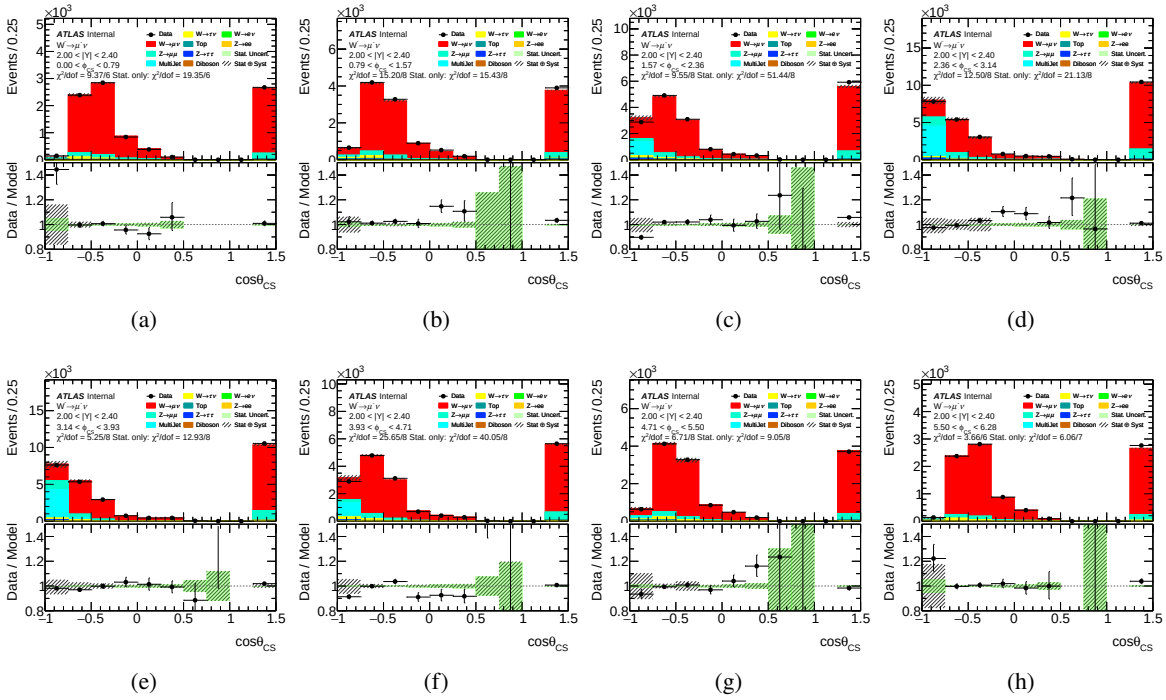


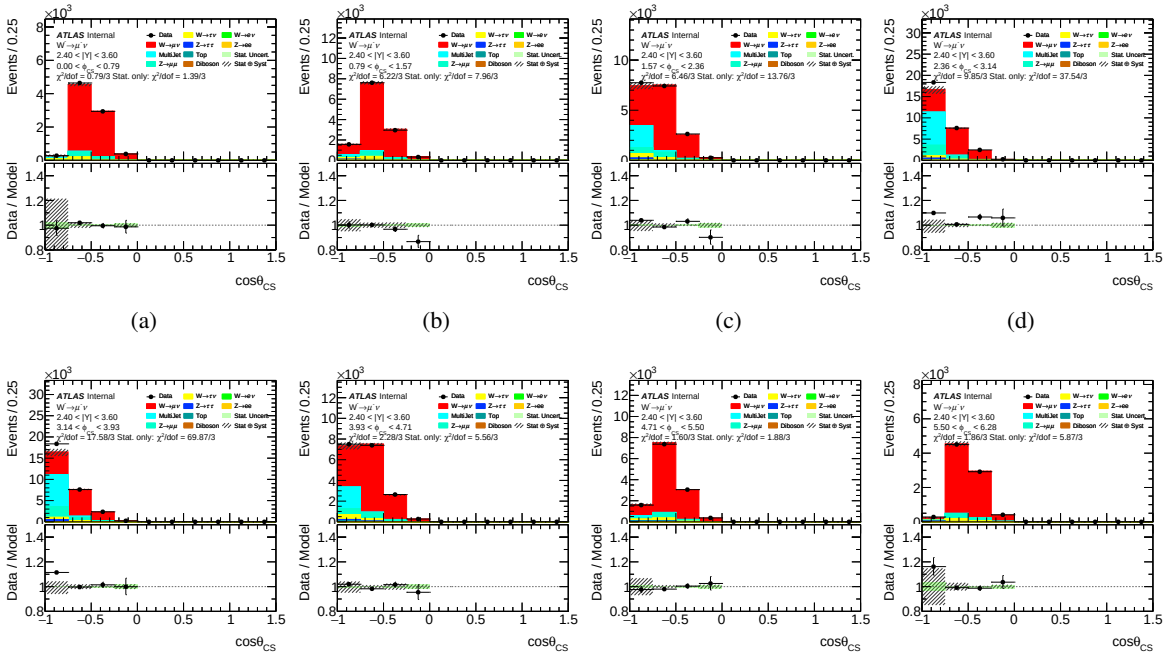
Figure 336: Control plots for $1.6 < |Y| < 2.0$ bin for $W^- \to \mu^- \overline{\nu}$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 337: Control plots for $2.0 < |Y| < 2.4$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

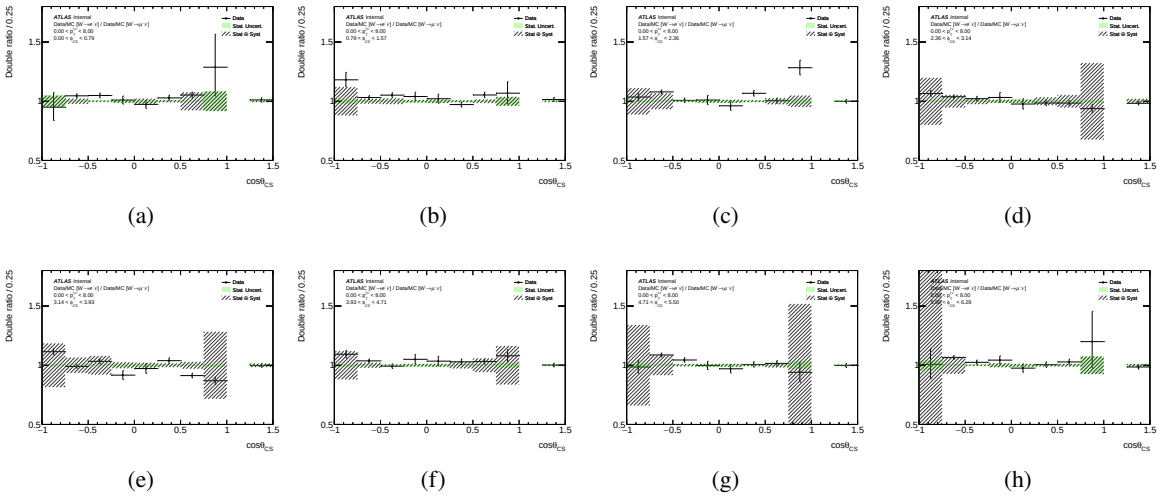


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 338: Control plots for $2.4 < |Y| < 3.6$ bin for $W^- \to \mu^- \overline{\nu}$ channel.

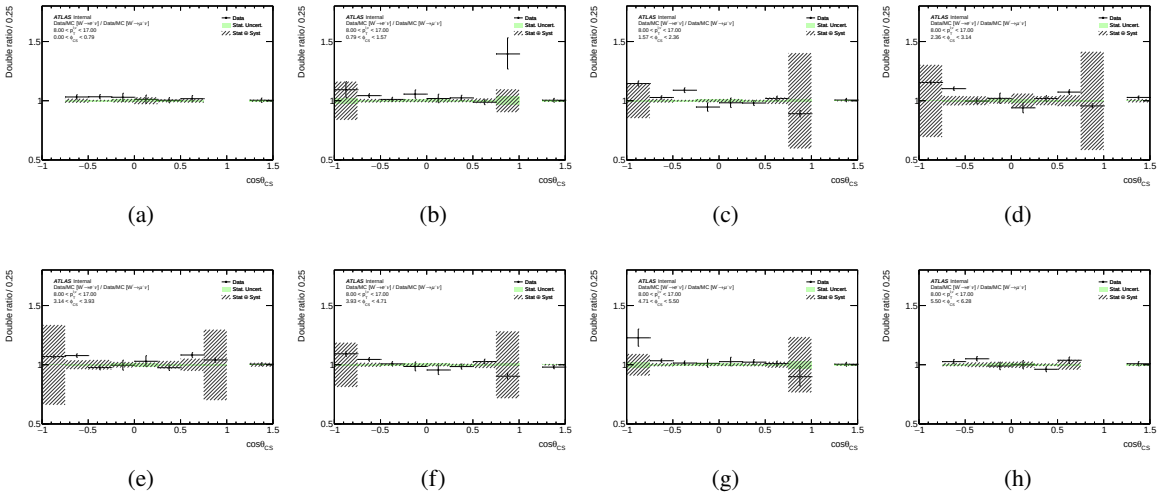
F.3 Double ratio $W^- \to e^-\overline{\nu}$/ $W^- \to \mu^-\overline{\nu}$



(a) (b) (c) (d)

(e) (f) (g) (h)

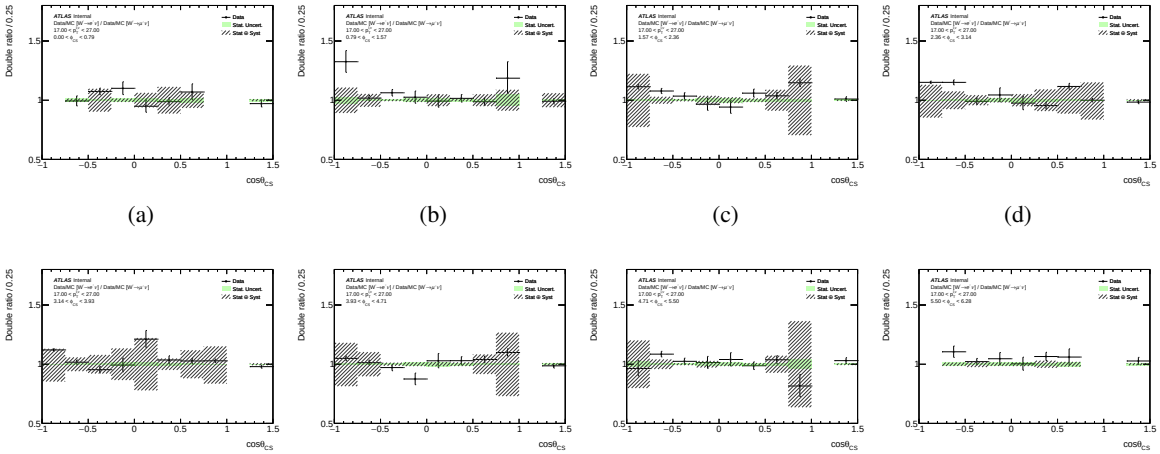
Figure 339: Double ratio plots for e/μ for $0 < p_T^{\ell,\nu} < 8$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

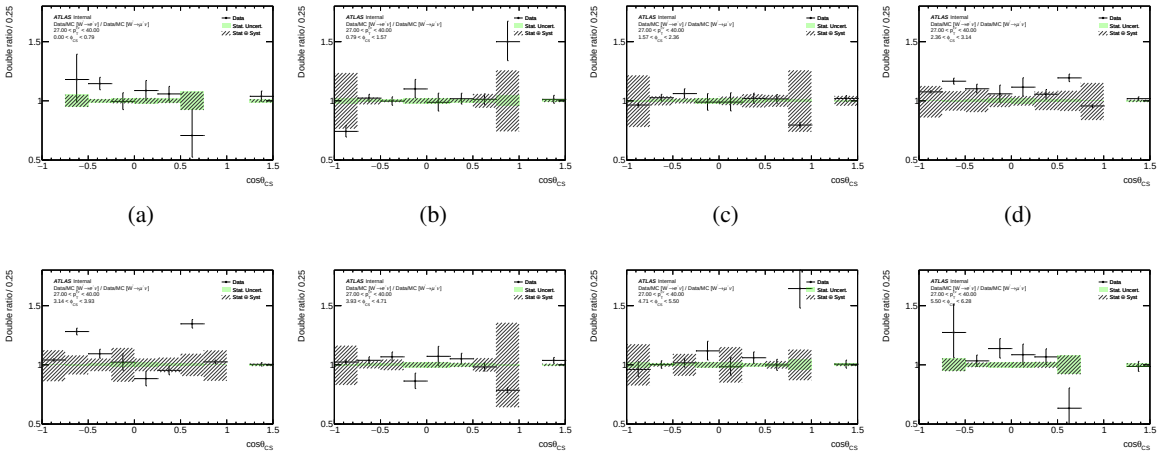
Figure 340: Double ratio plots for e/μ for $8 < p_T^{\ell,\nu} < 17$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

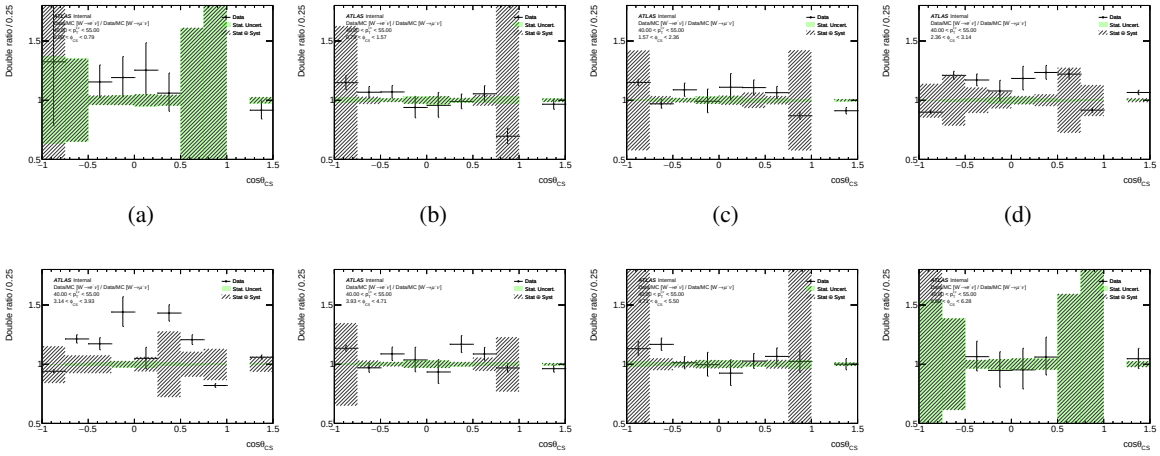
Figure 341: Double ratio plots for e/μ for $17 < p_T^{\ell,\nu} < 27$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

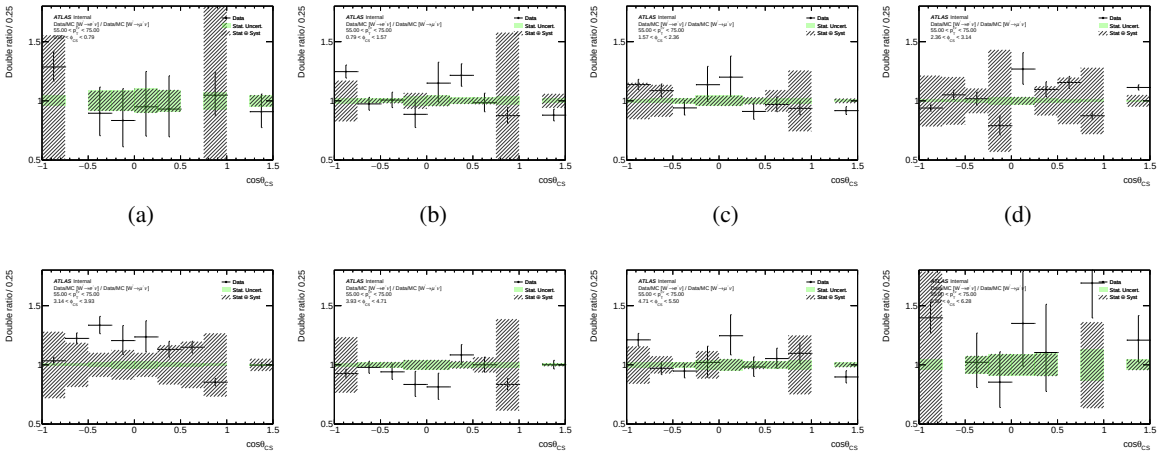
Figure 342: Double ratio plots for e/μ for $27 < p_T^{\ell,\nu} < 40$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

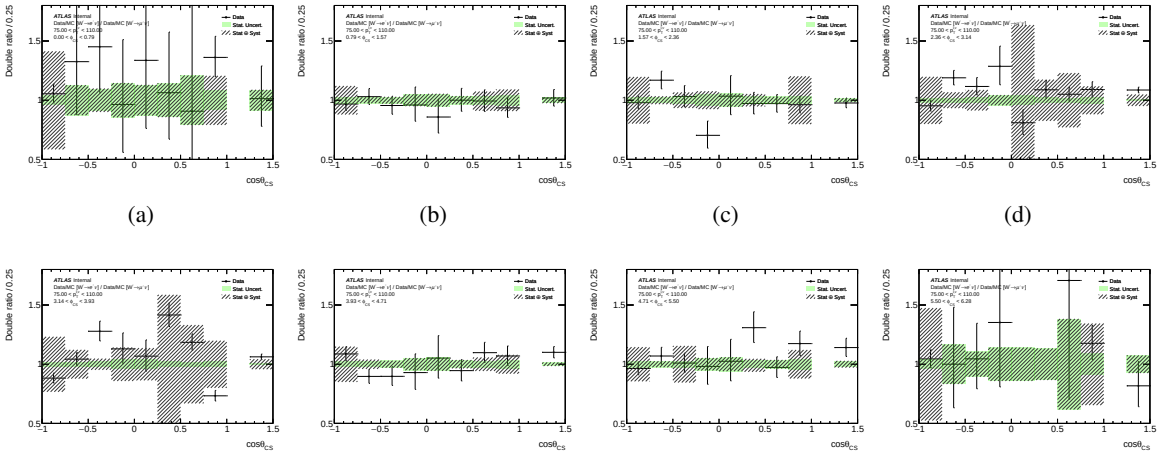
Figure 343: Double ratio plots for e/μ for $40 < p_T^{\ell,\nu} < 55$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

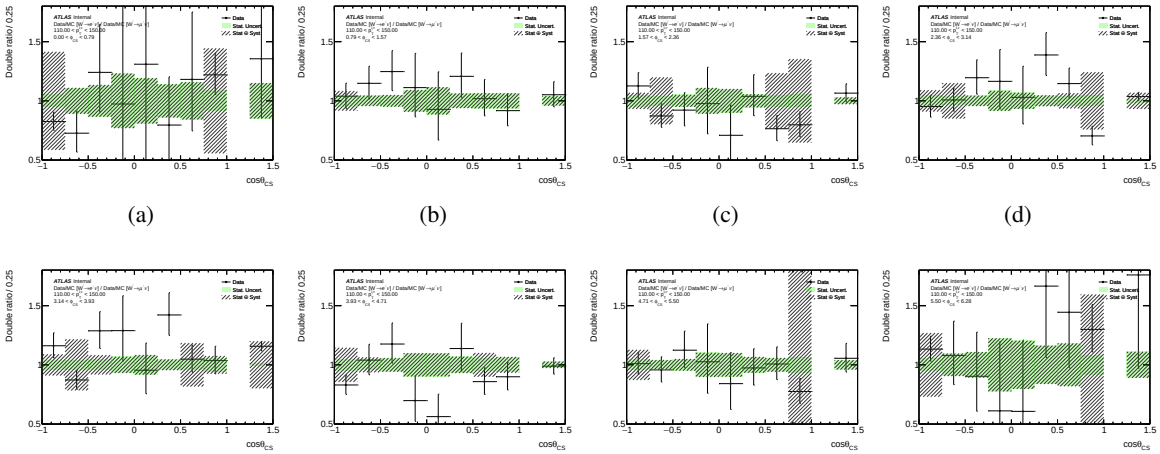
Figure 344: Double ratio plots for e/μ for $55 < p_T^{\ell,\nu} < 75$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

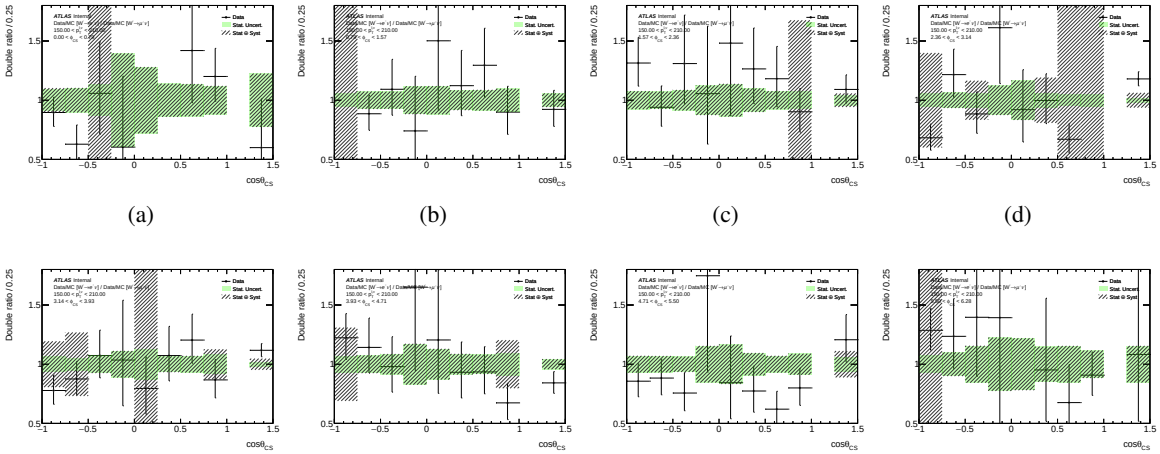
Figure 345: Double ratio plots for e/μ for $75 < p_T^{\ell,\nu} < 110$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

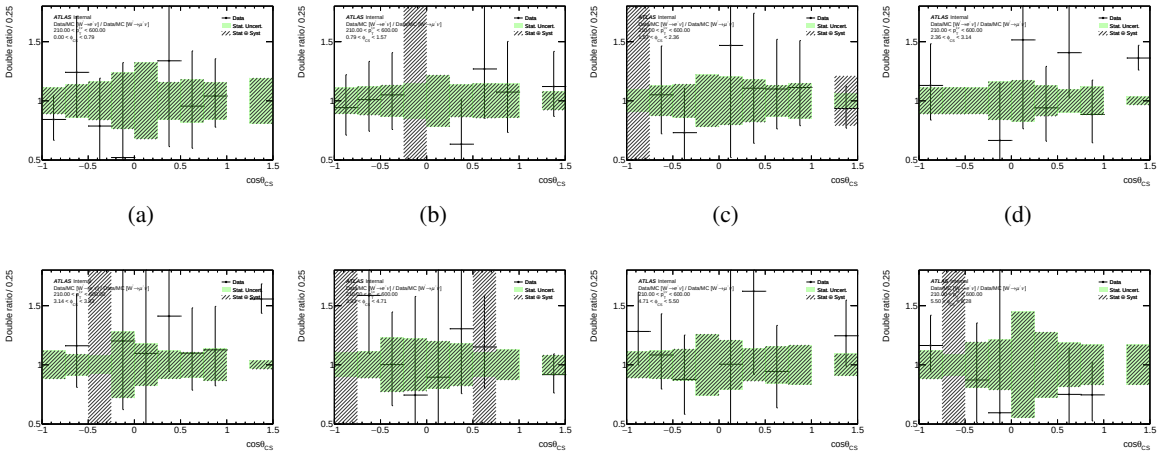
Figure 346: Double ratio plots for e/μ for $110 < p_T^{\ell,\nu} < 150$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

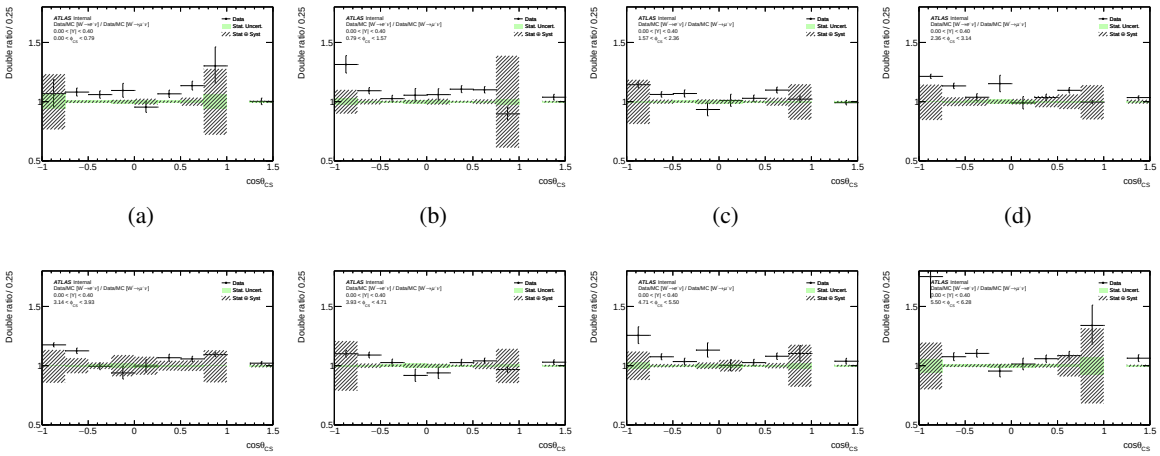
Figure 347: Double ratio plots for e/μ for $150 < p_T^{\ell,\nu} < 210$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

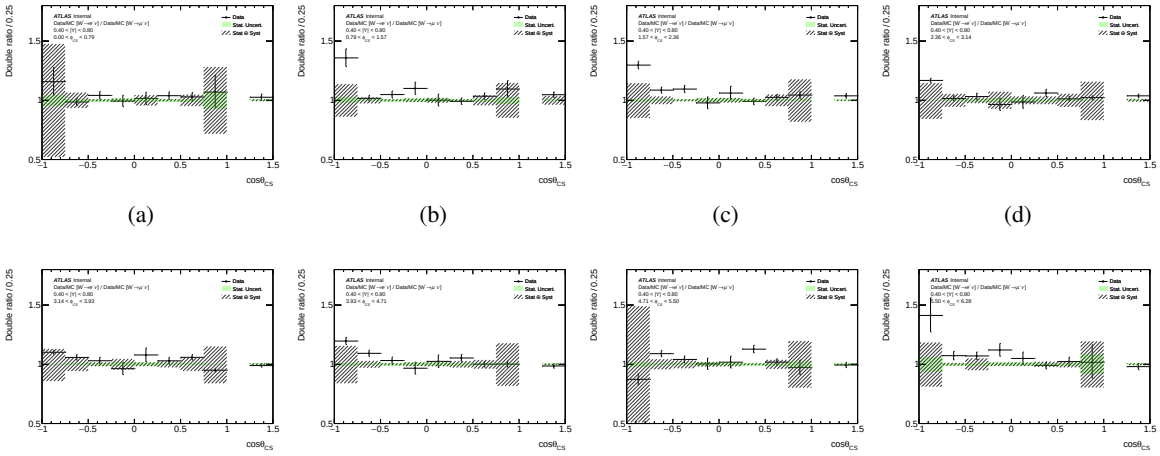
Figure 348: Double ratio plots for e/μ for $210 < p_T^{\ell,\nu} < 600$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

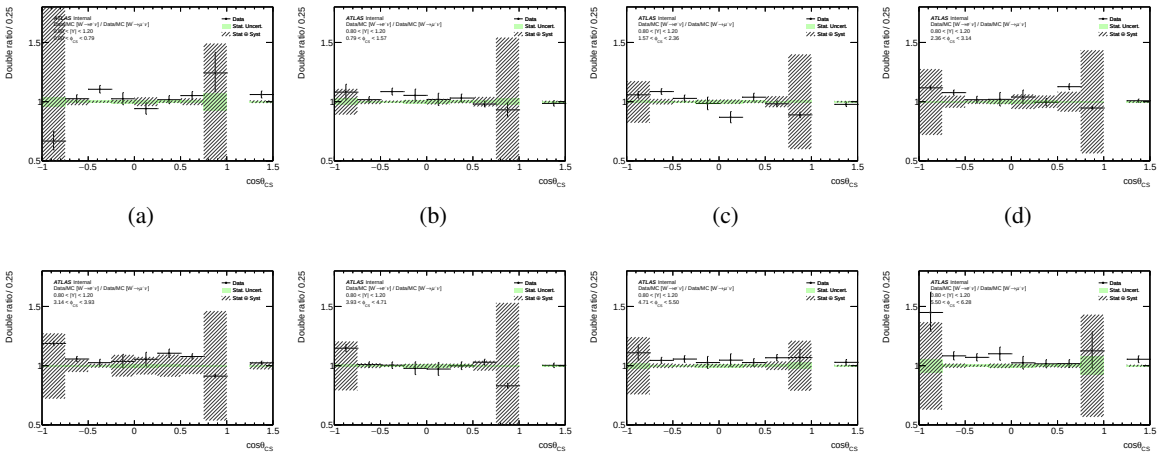
Figure 349: Double ratio plots for e/μ for $0 < |Y| < 0.4$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

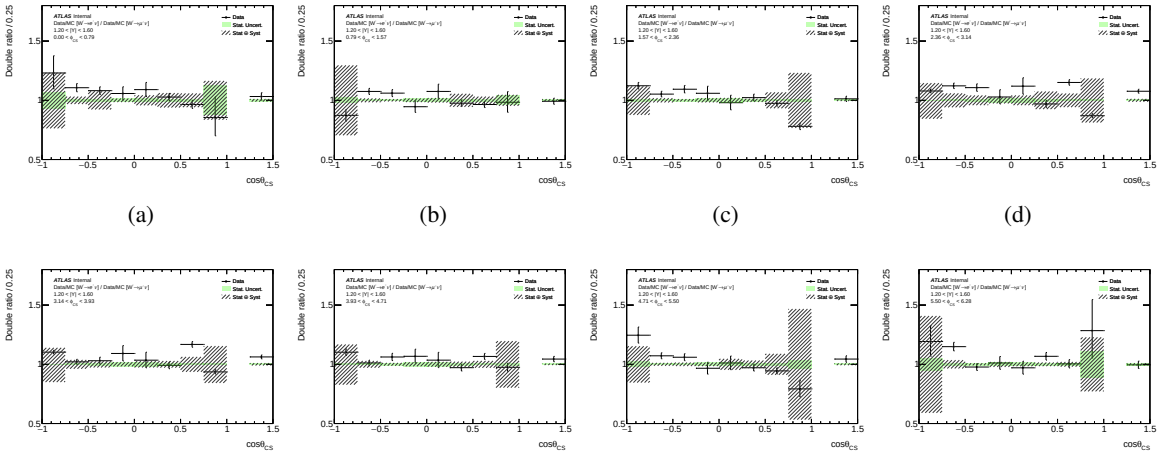
Figure 350: Double ratio plots for e/μ for $0.4 < |Y| < 0.8$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

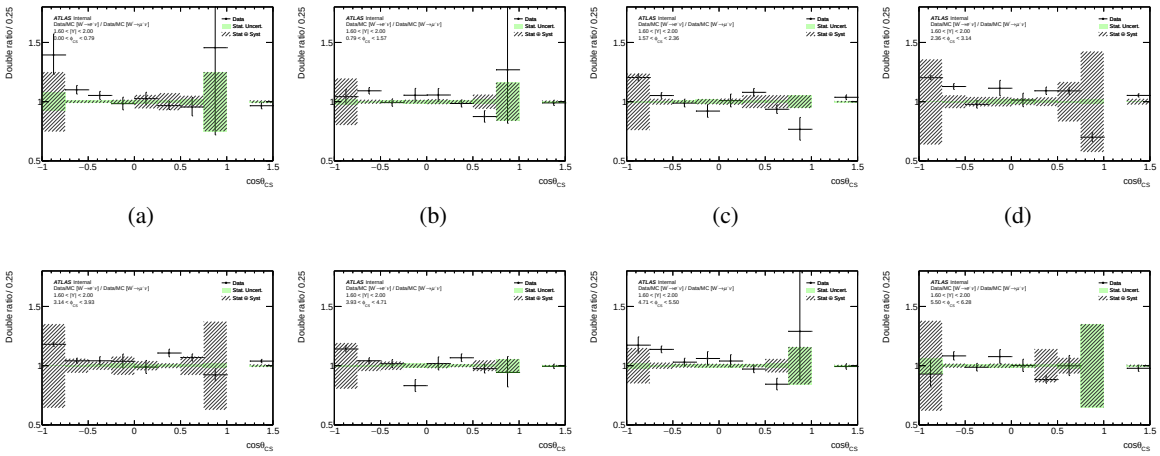
Figure 351: Double ratio plots for e/μ for $0.8 < |Y| < 1.2$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

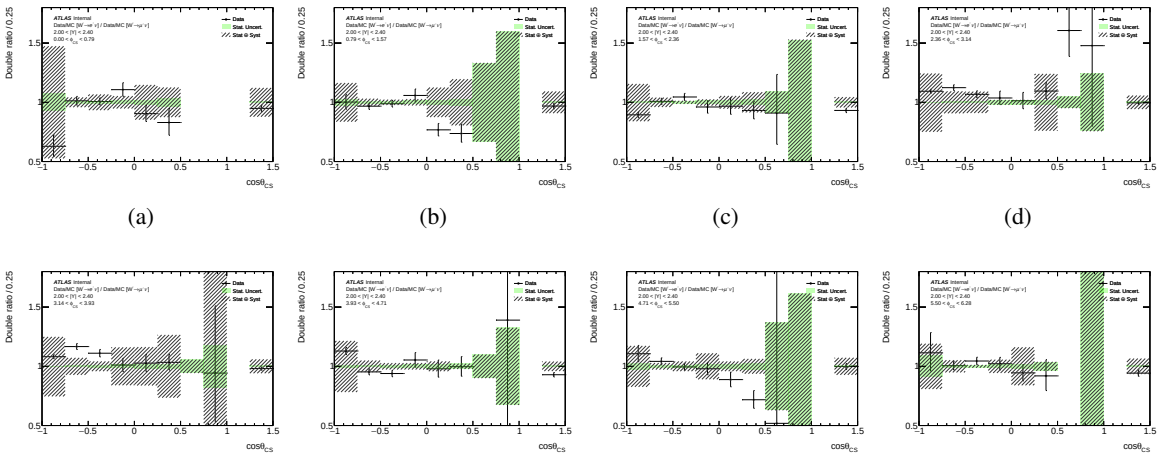
Figure 352: Double ratio plots for e/μ for $1.2 < |Y| < 1.6$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

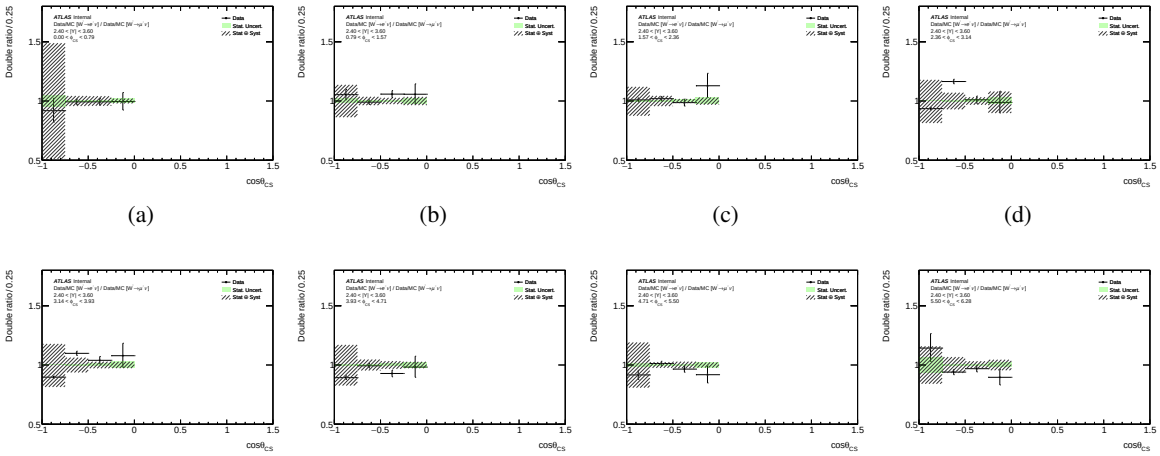
Figure 353: Double ratio plots for e/μ for $1.6 < |Y| < 2.0$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 354: Double ratio plots for e/μ for $2.0 < |Y| < 2.4$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 355: Double ratio plots for e/μ for $2.4 < |Y| < 3.6$ bin.

F.4 Positron channel

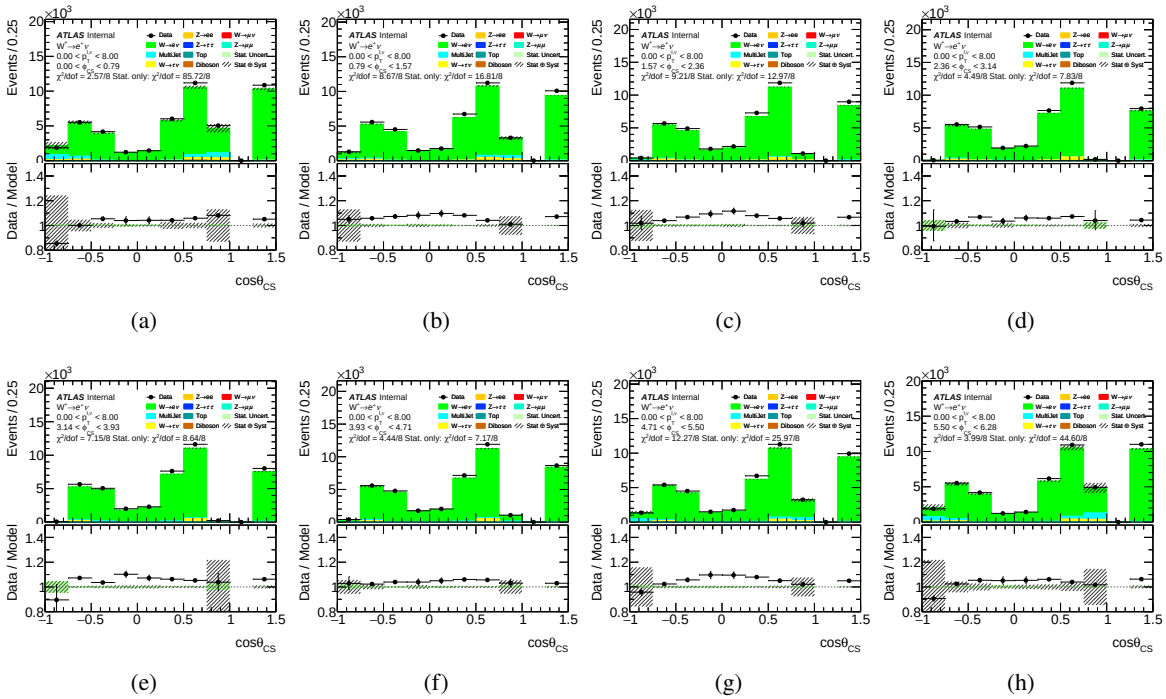


Figure 356: Control plots for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to e^+\nu$ channel.

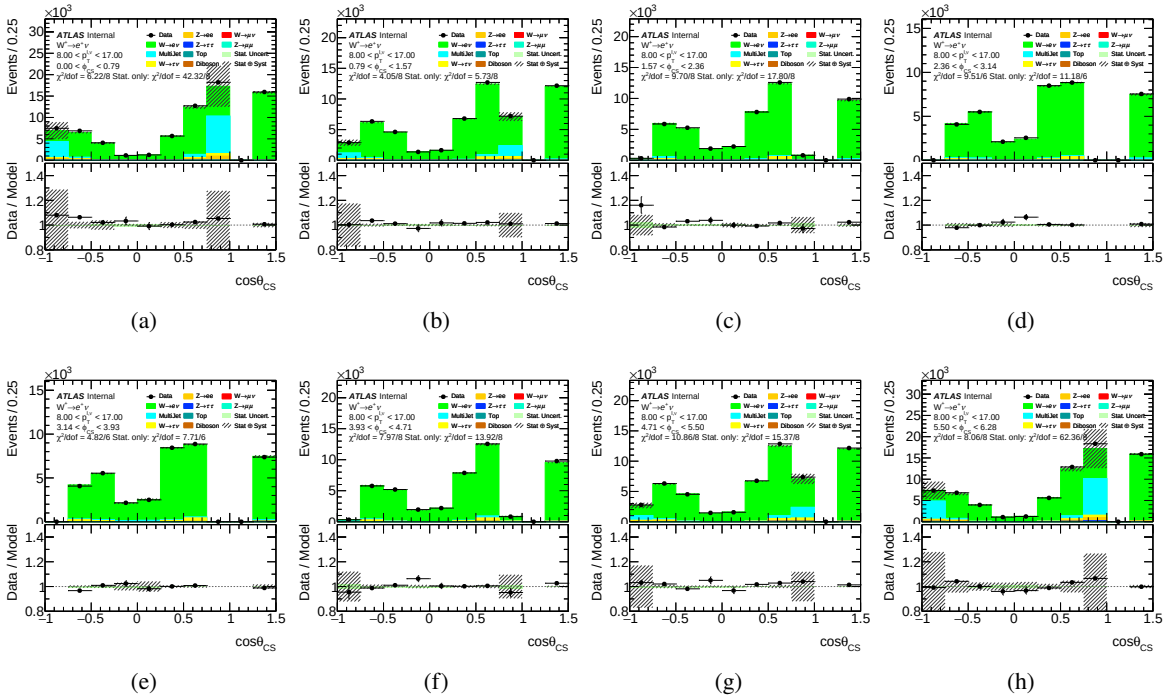


Figure 357: Control plots for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to e^+\nu$ channel.

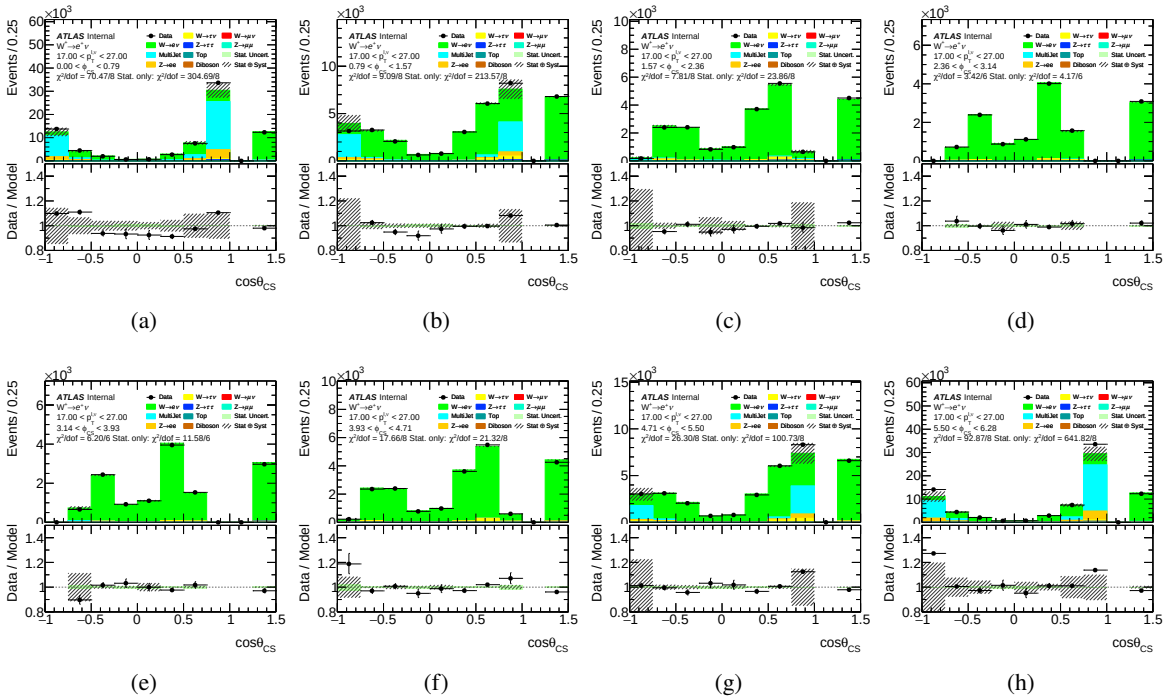
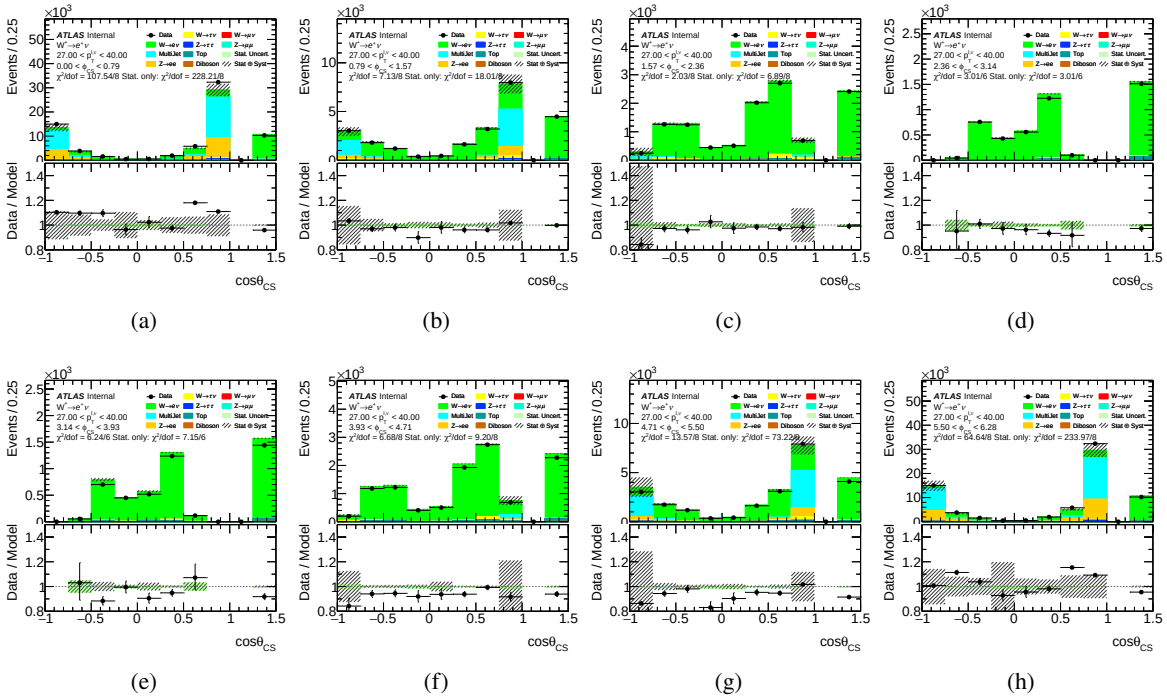


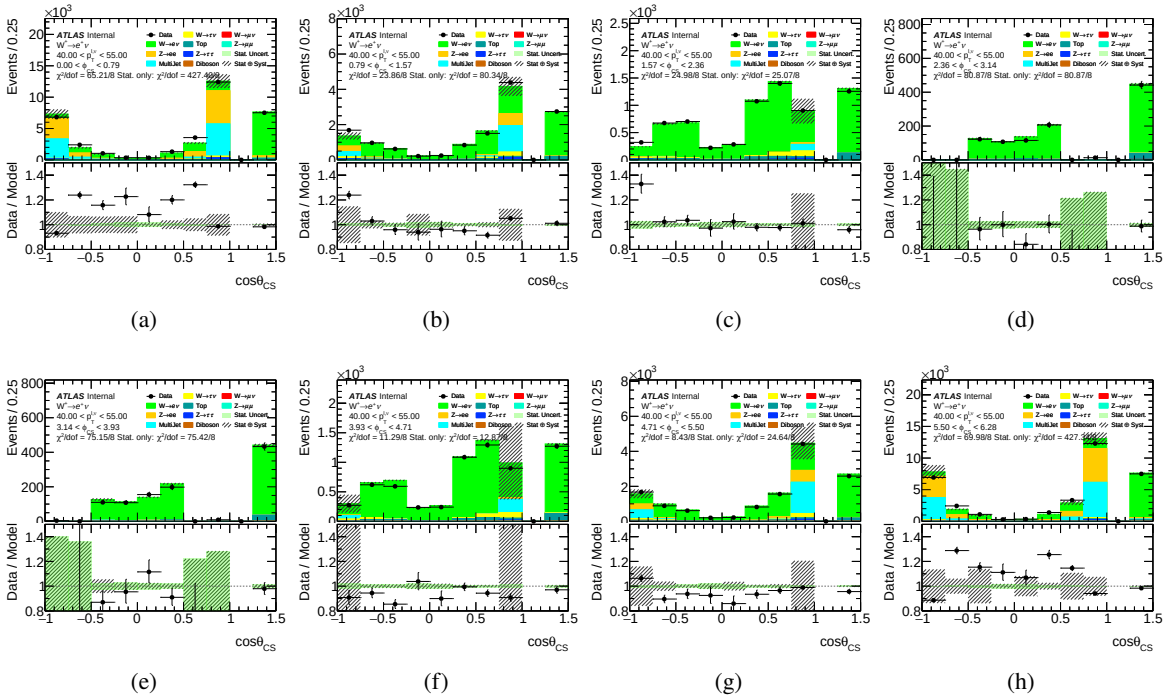
Figure 358: Control plots for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

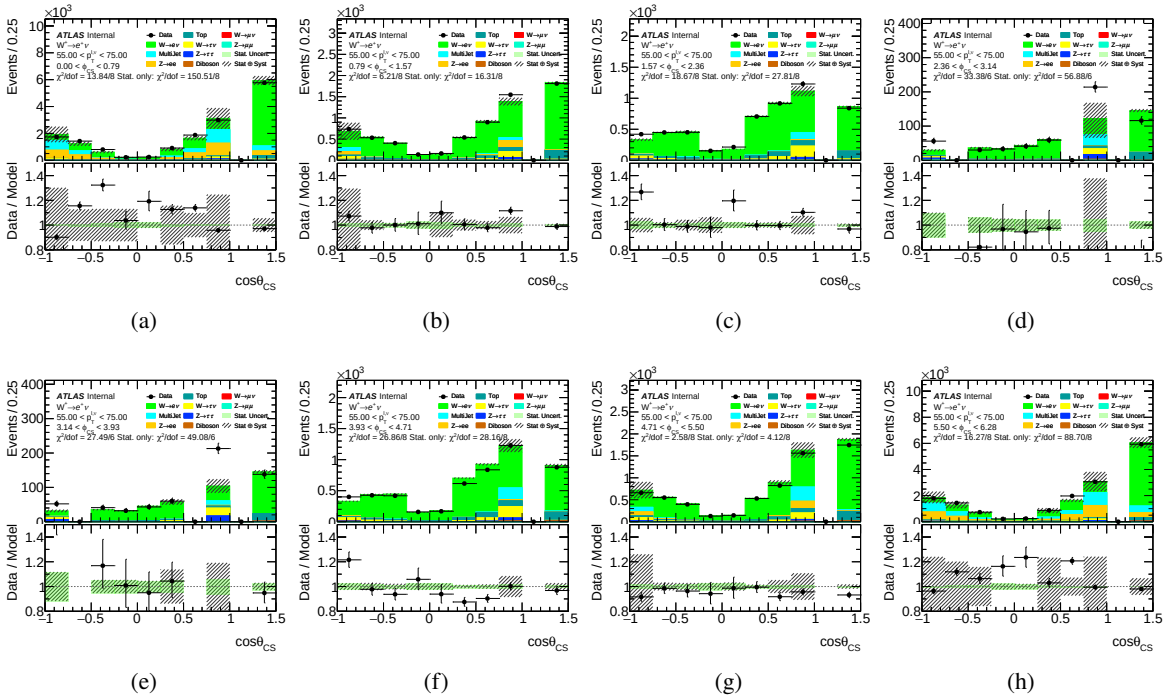
Figure 359: Control plots for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

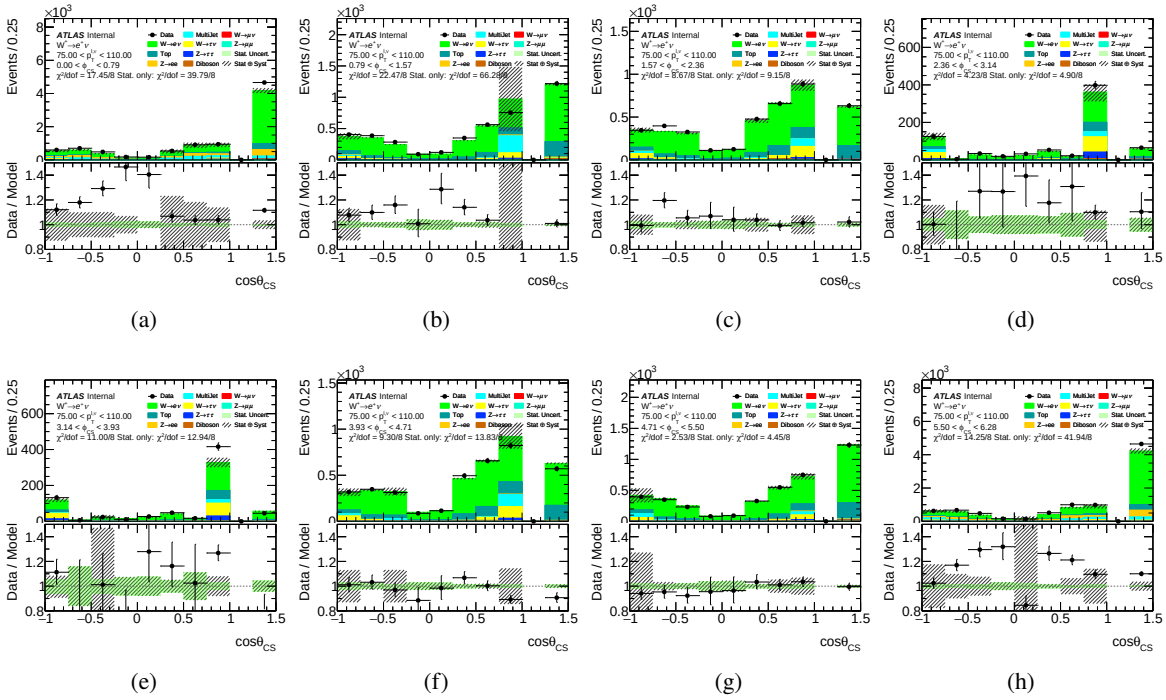
Figure 360: Control plots for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 361: Control plots for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 362: Control plots for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \to e^+\nu$ channel.

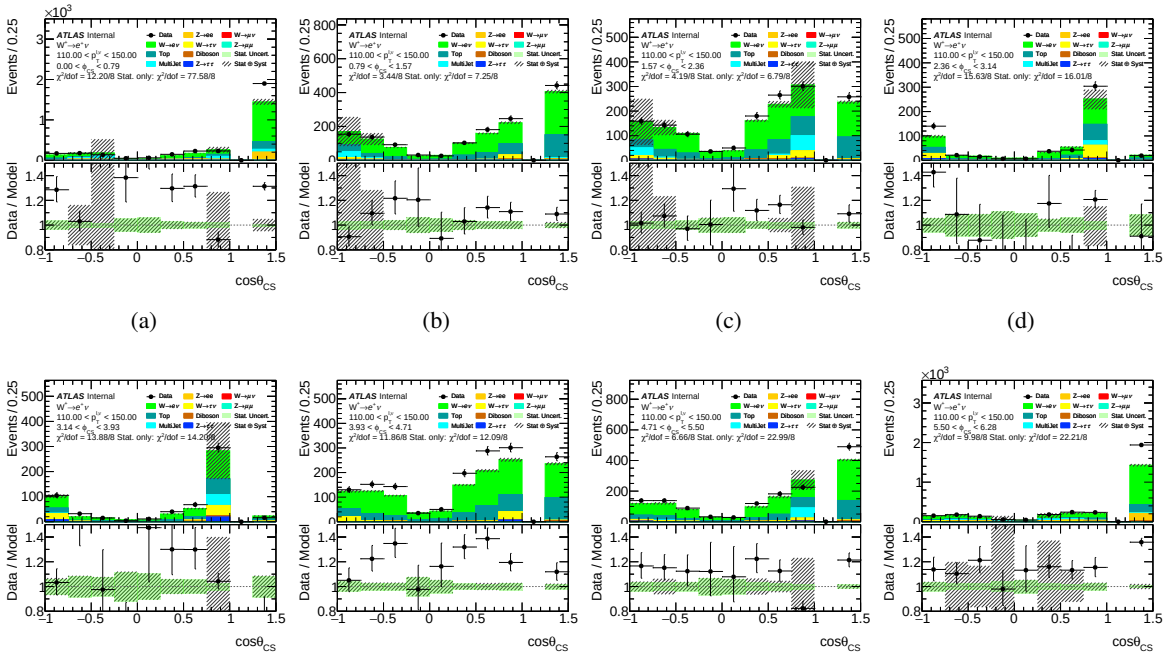


Figure 363: Control plots for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \to e^+\nu$ channel.

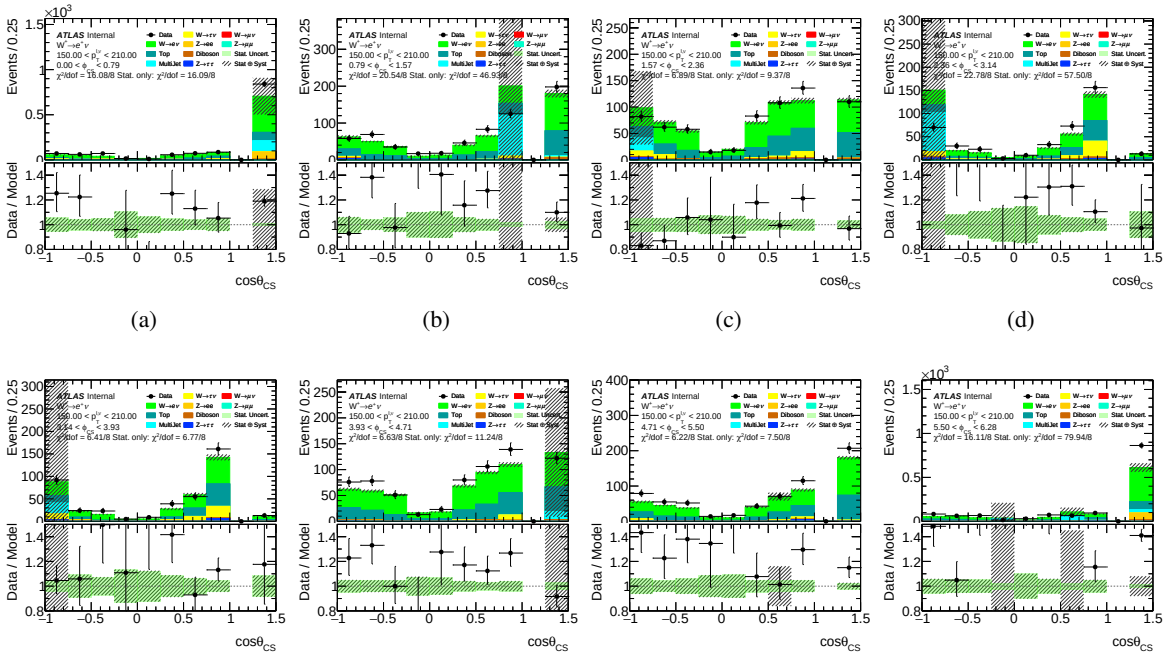
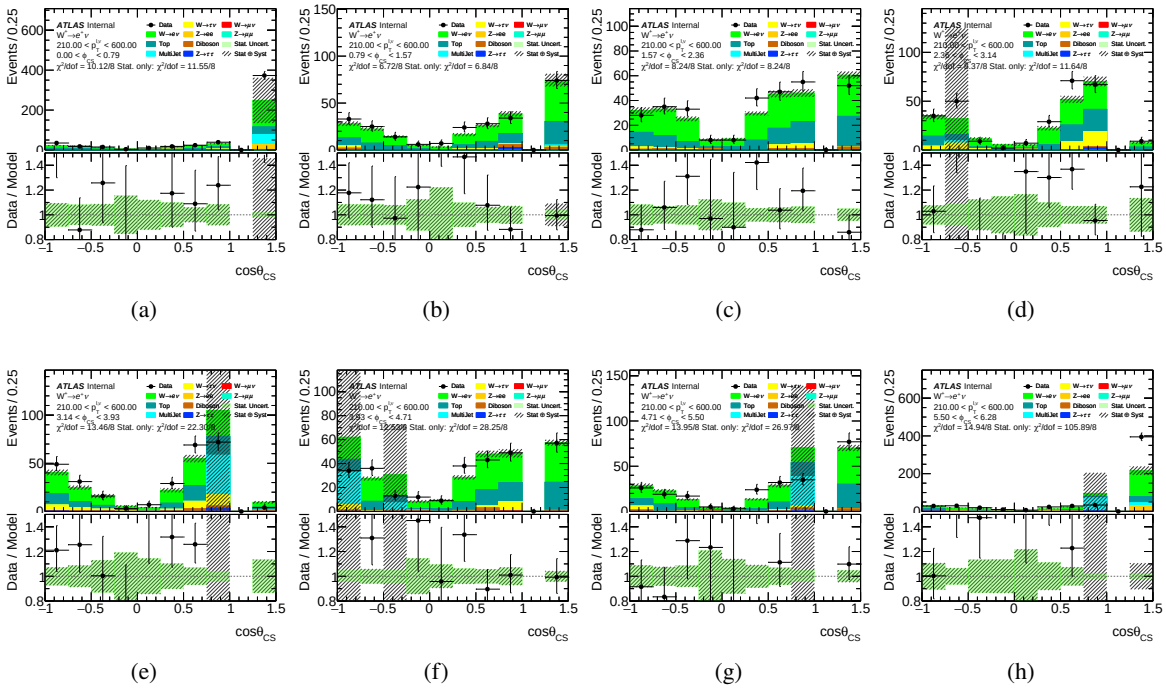


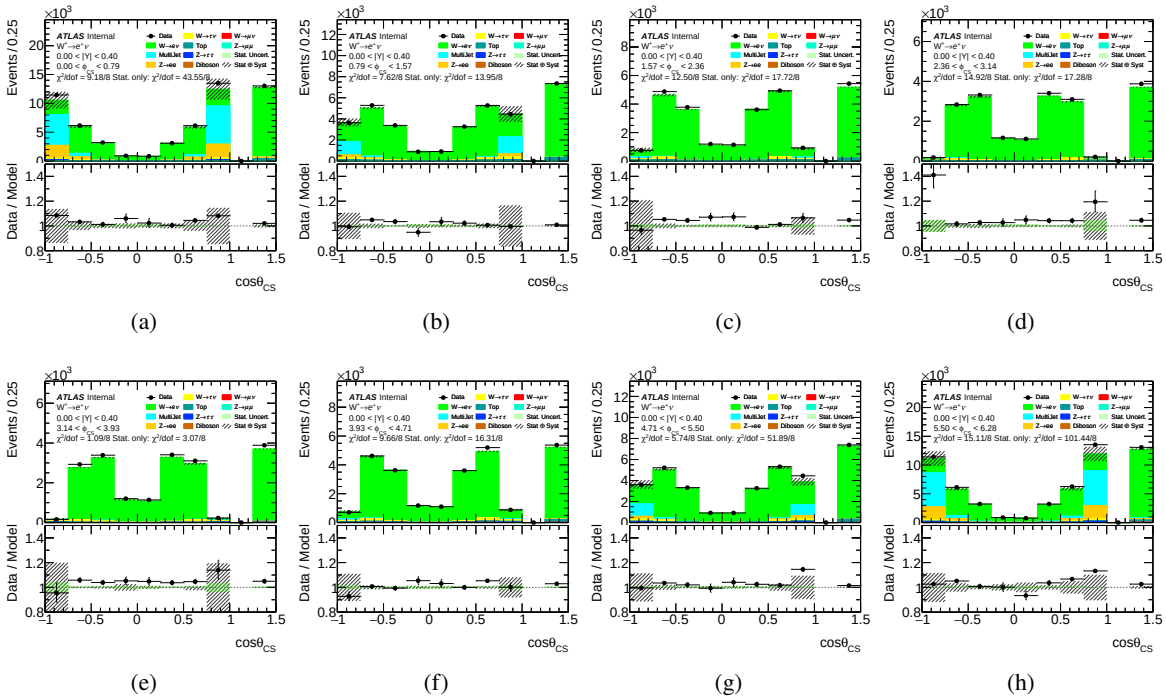
Figure 364: Control plots for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 365: Control plots for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to e^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 366: Control plots for $0 < |Y| < 0.4$ bin for $W^+ \to e^+\nu$ channel.

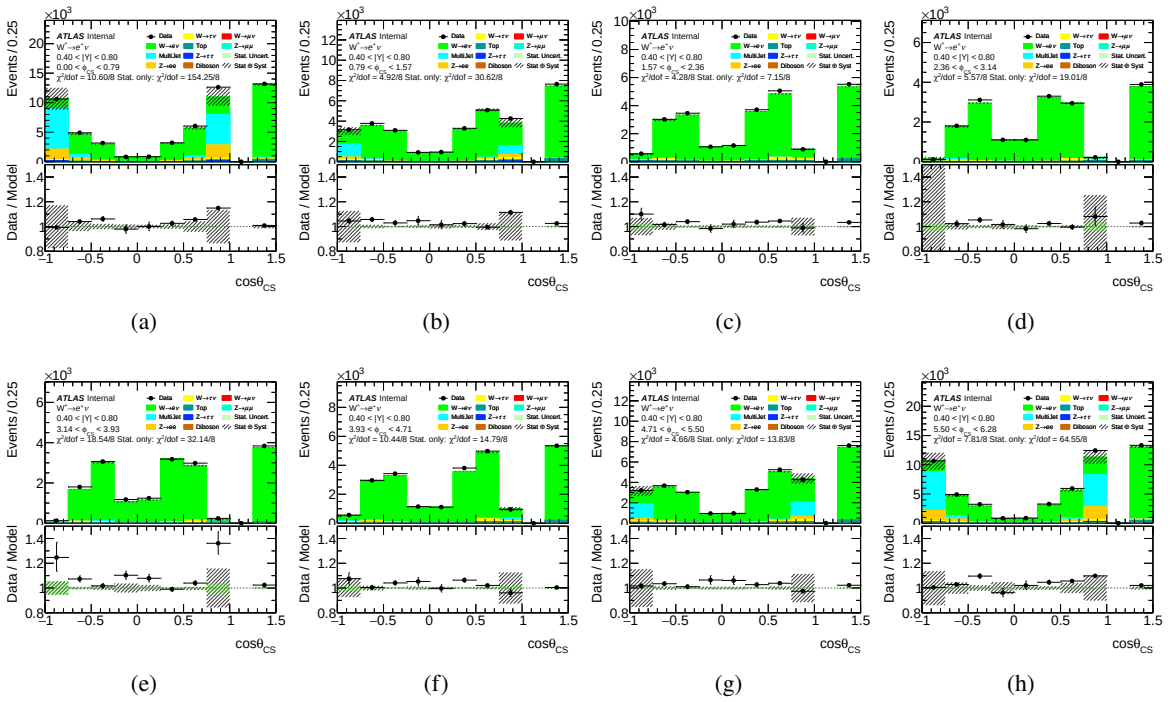


Figure 367: Control plots for $0.4 < |Y| < 0.8$ bin for $W^+ \to e^+\nu$ channel.

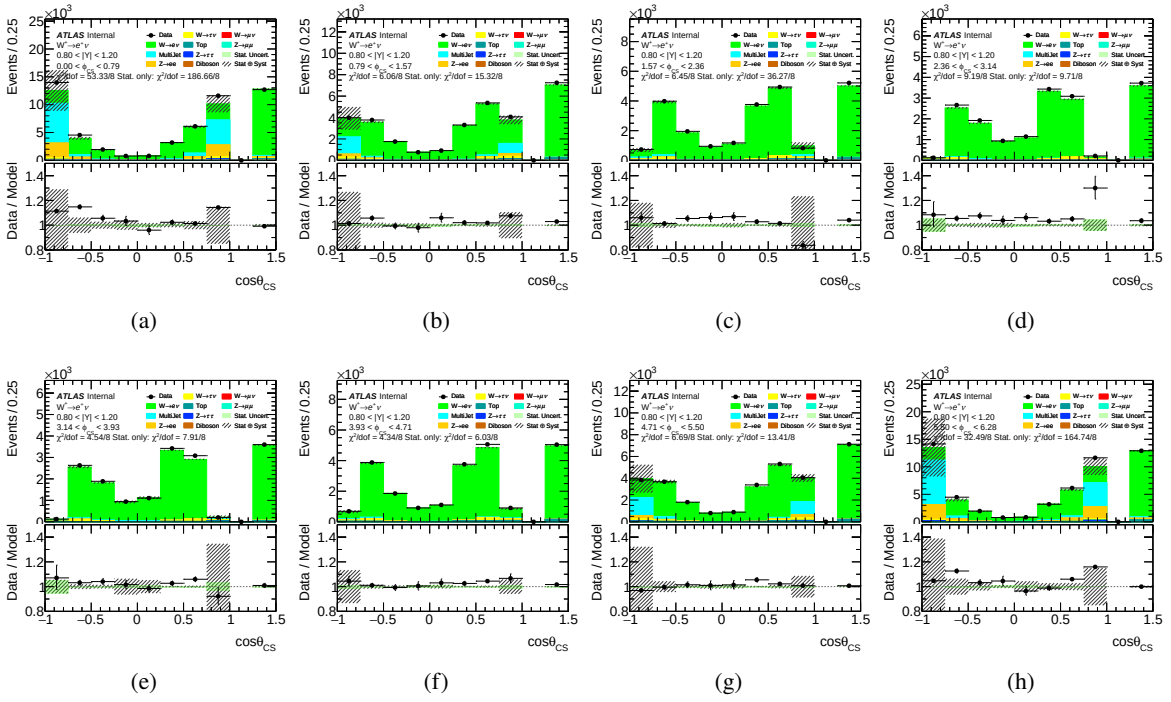


Figure 368: Control plots for $0.8 < |Y| < 1.2$ bin for $W^+ \to e^+\nu$ channel.

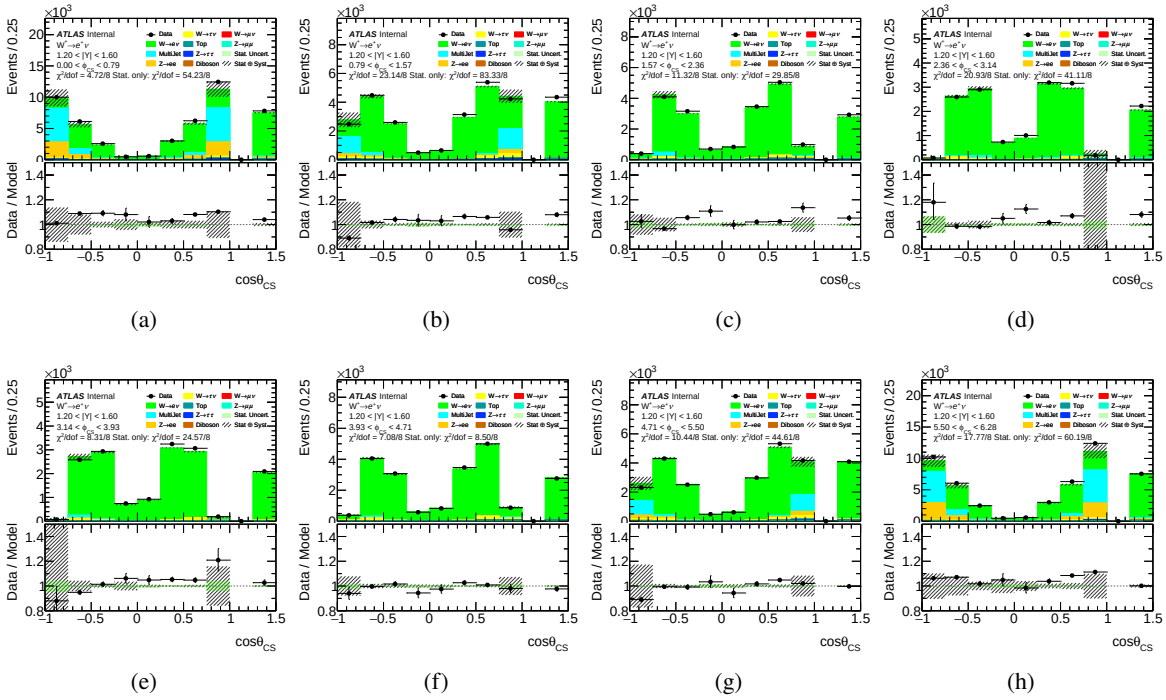


Figure 369: Control plots for $1.2 < |Y| < 1.6$ bin for $W^+ \to e^+\nu$ channel.

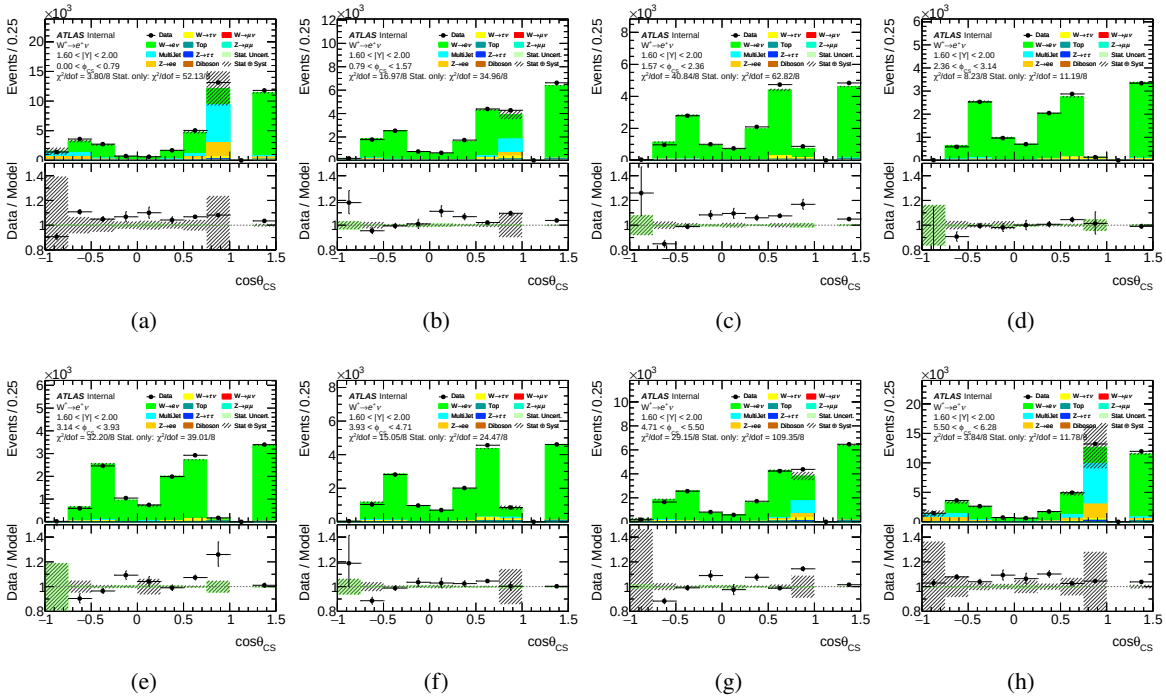


Figure 370: Control plots for $1.6 < |Y| < 2.0$ bin for $W^+ \to e^+\nu$ channel.

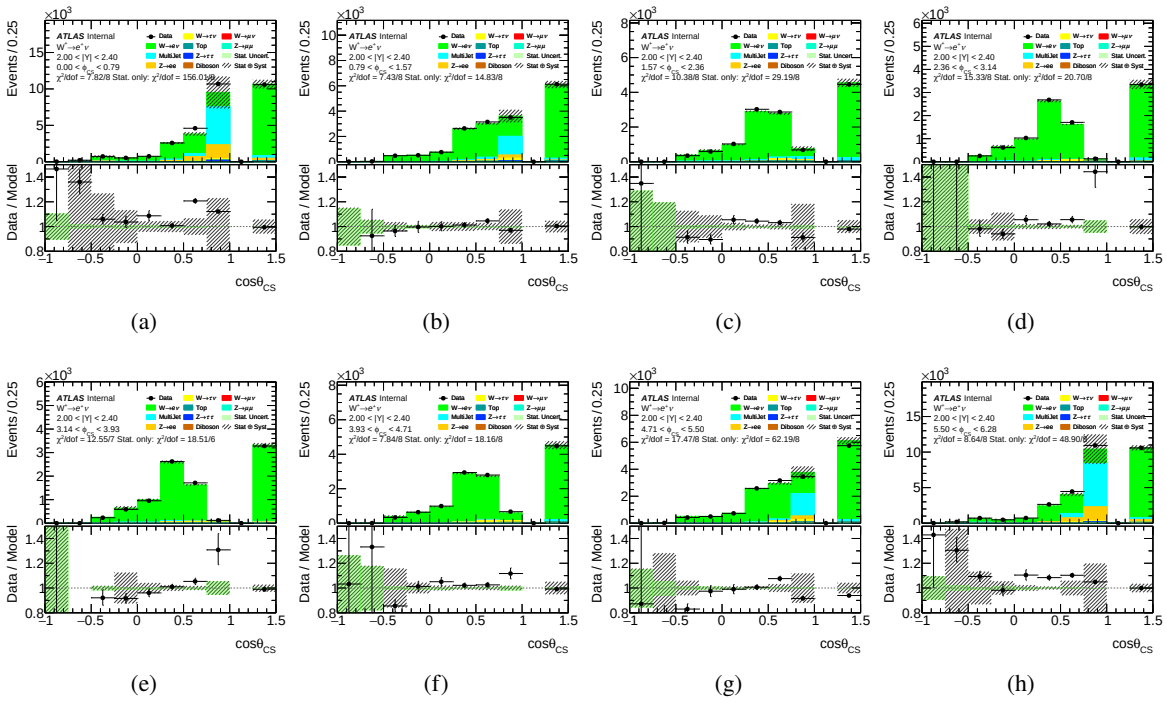


Figure 371: Control plots for $2.0 < |Y| < 2.4$ bin for $W^+ \to e^+\nu$ channel.

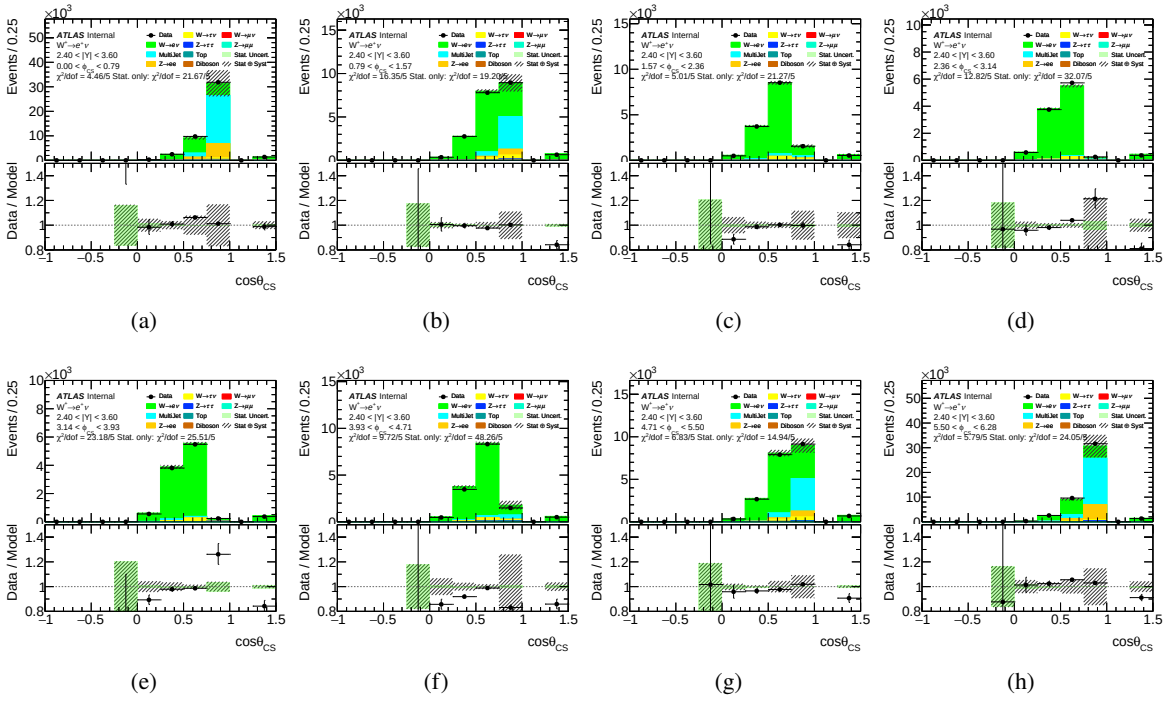
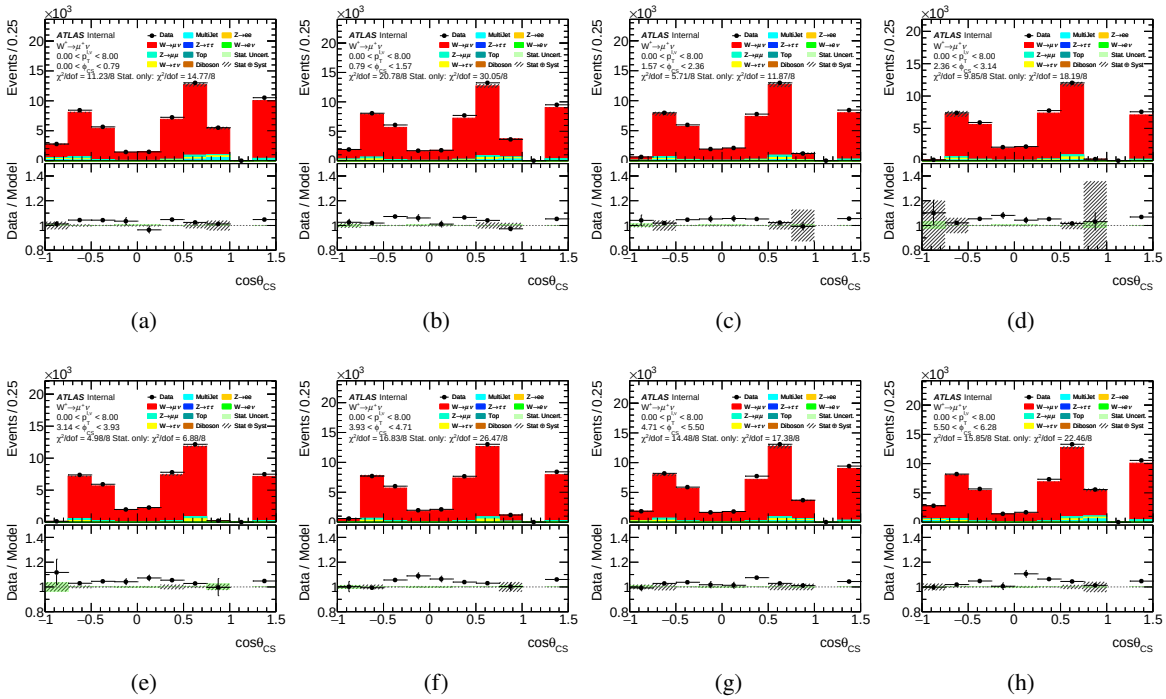


Figure 372: Control plots for $2.4 < |Y| < 3.6$ bin for $W^+ \to e^+\nu$ channel.

F.5 Anti-muon channel



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 373: Control plots for $0 < p_T^{\ell,\nu} < 8$ GeV bin for $W^+ \to \mu^+\nu$ channel.

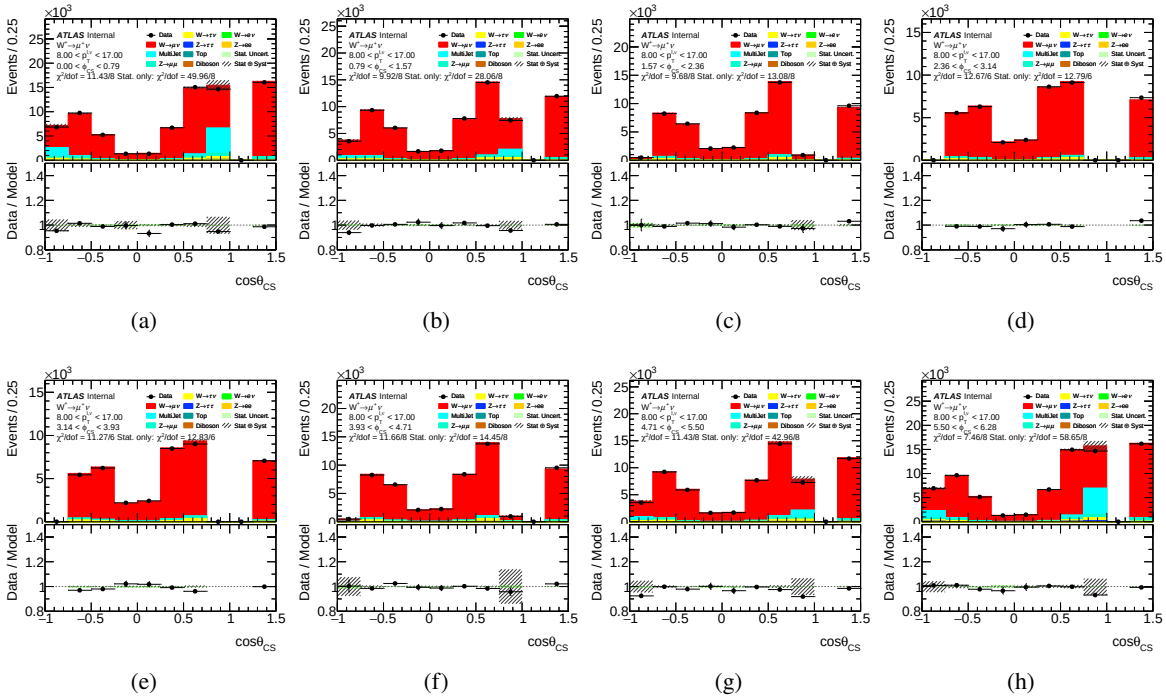


Figure 374: Control plots for $8 < p_T^{\ell,\nu} < 17$ GeV bin for $W^+ \to \mu^+\nu$ channel.

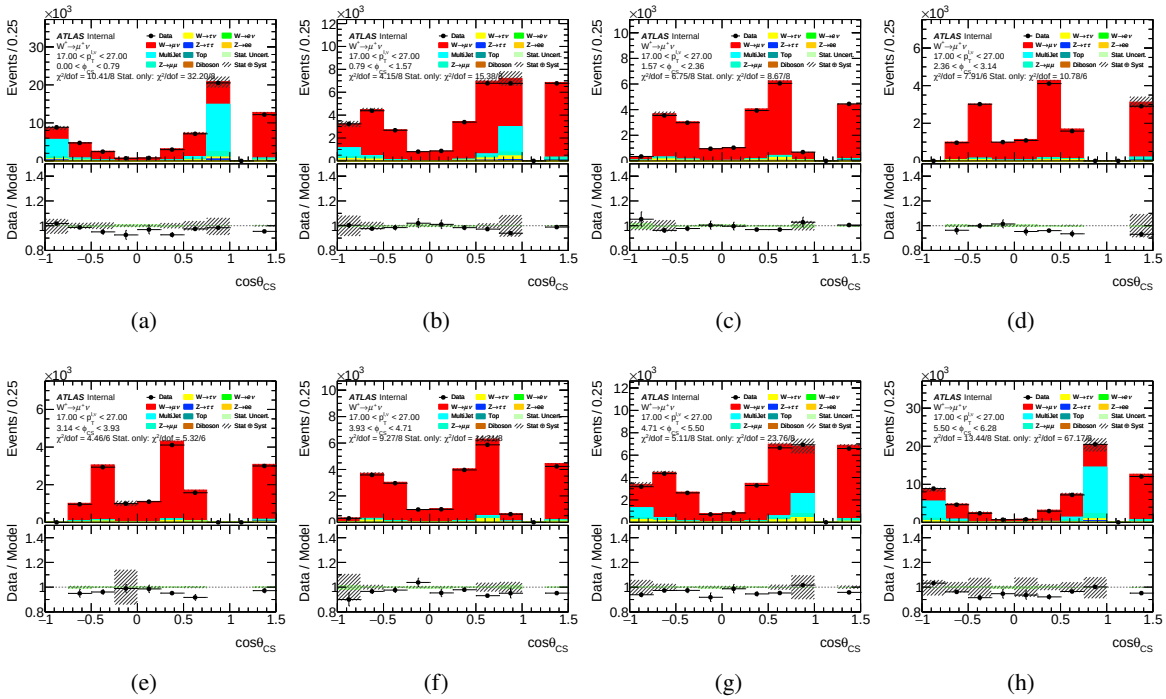
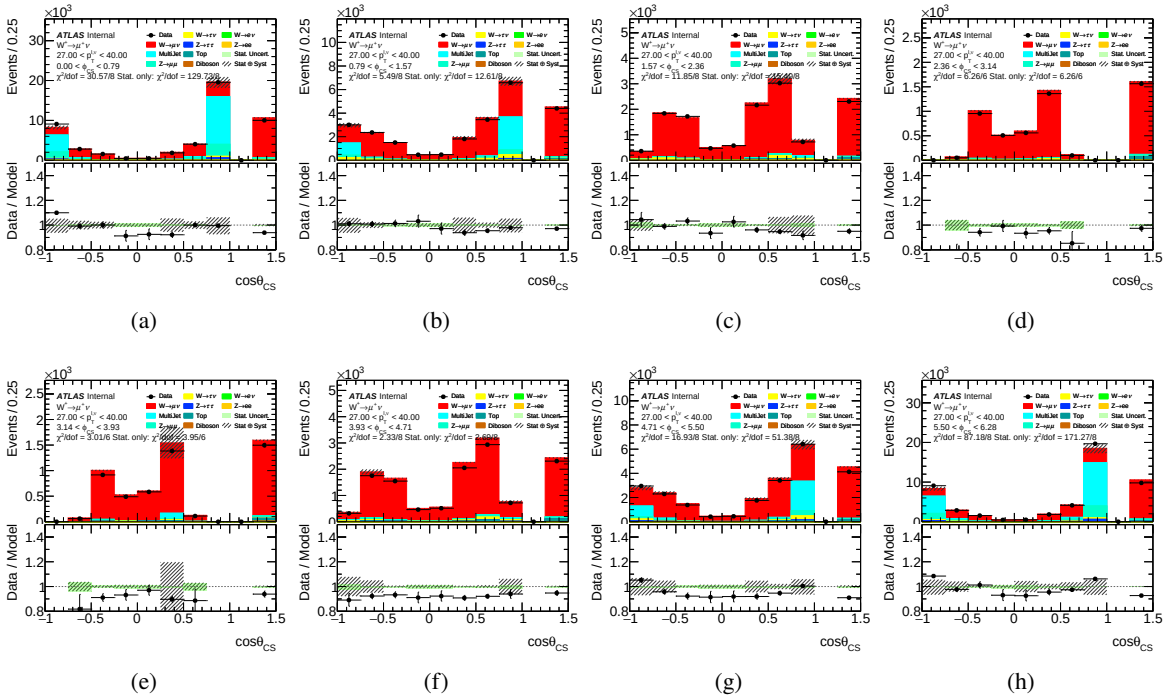


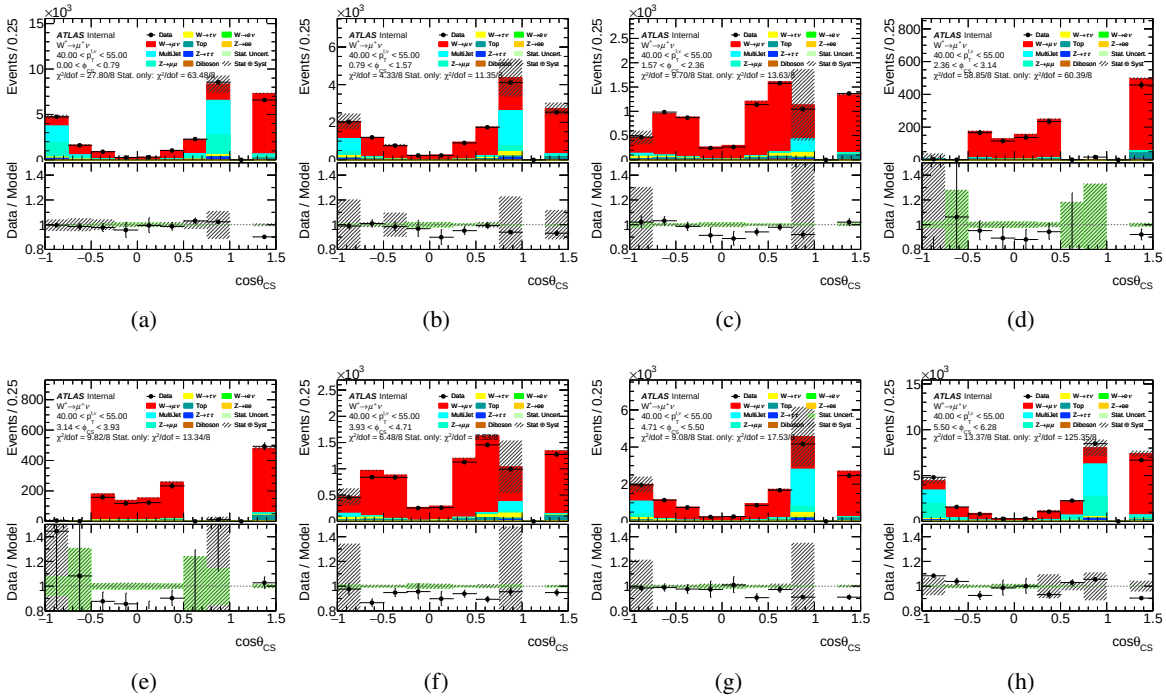
Figure 375: Control plots for $17 < p_T^{\ell,\nu} < 27$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 376: Control plots for $27 < p_T^{\ell,\nu} < 40$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 377: Control plots for $40 < p_T^{\ell,\nu} < 55$ GeV bin for $W^+ \to \mu^+\nu$ channel.

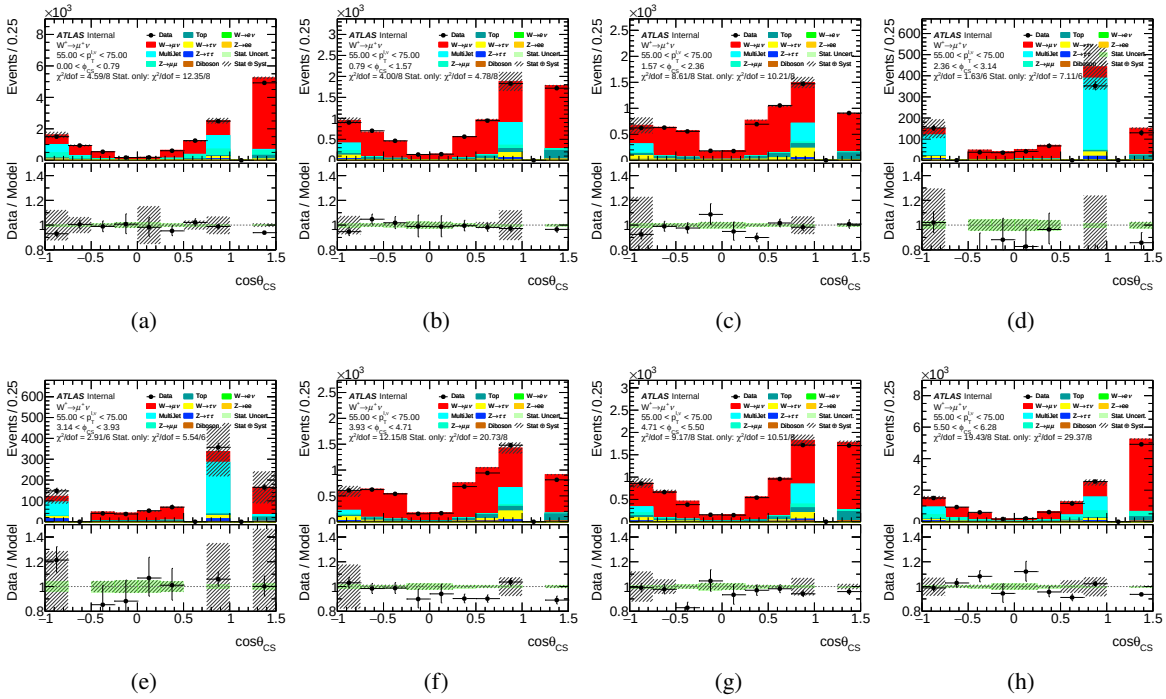


Figure 378: Control plots for $55 < p_T^{\ell,\nu} < 75$ GeV bin for $W^+ \to \mu^+\nu$ channel.

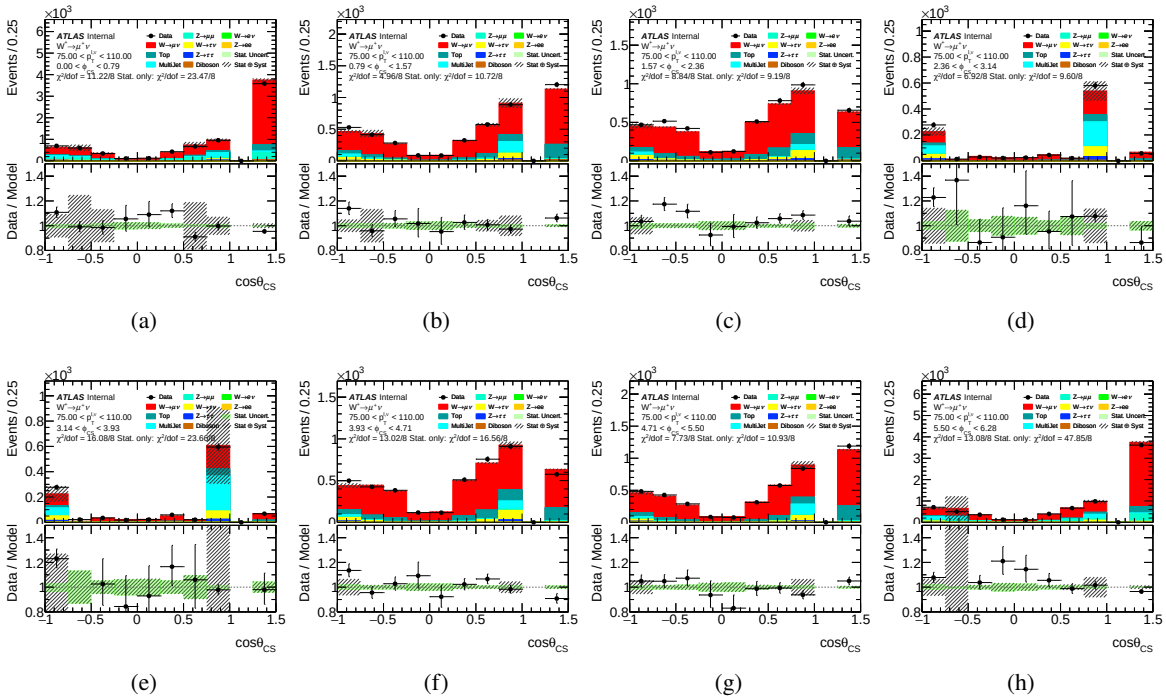
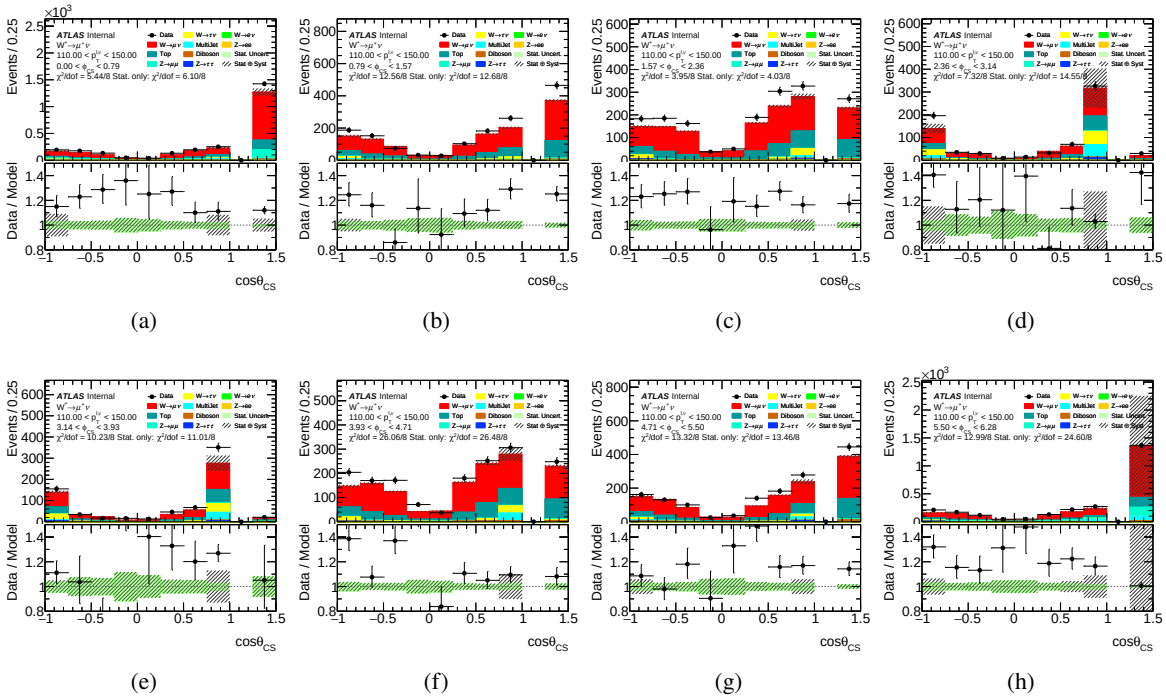


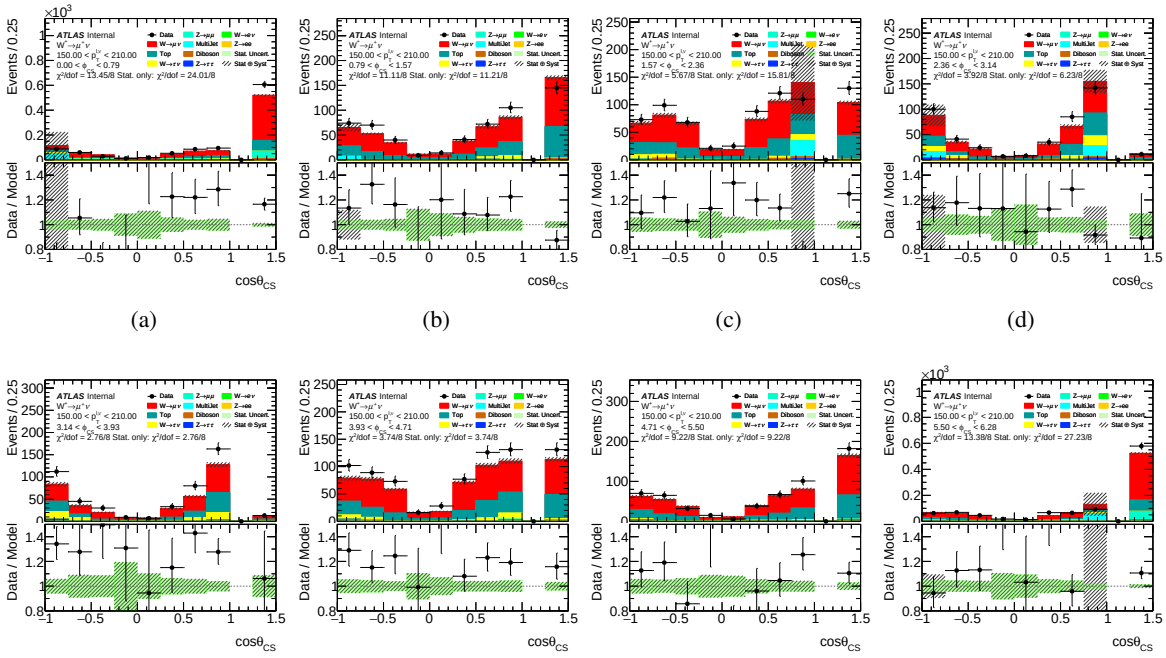
Figure 379: Control plots for $75 < p_T^{\ell,\nu} < 110$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

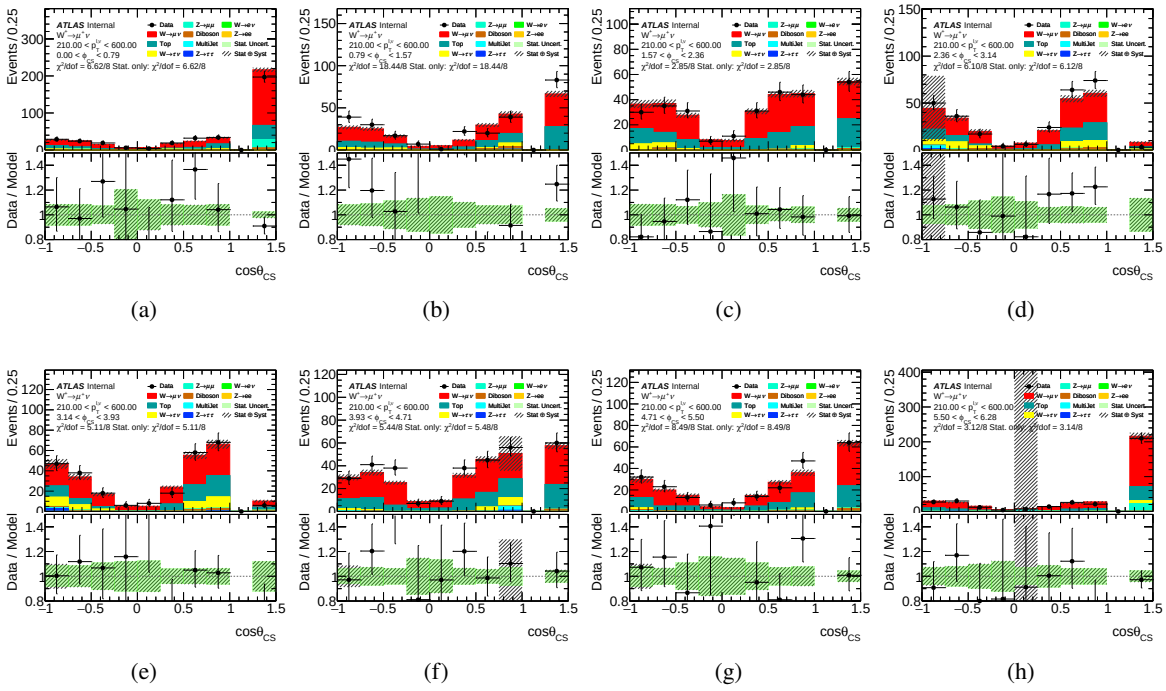
Figure 380: Control plots for $110 < p_T^{\ell,\nu} < 150$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

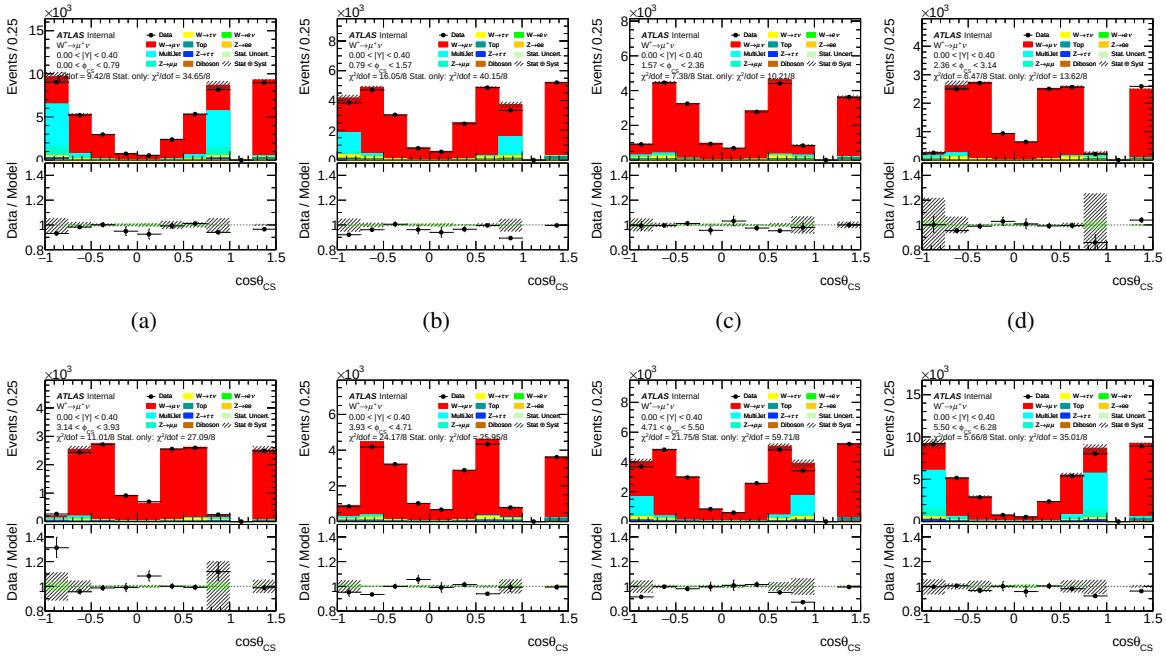
Figure 381: Control plots for $150 < p_T^{\ell,\nu} < 210$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 382: Control plots for $210 < p_T^{\ell,\nu} < 600$ GeV bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 383: Control plots for $0 < |Y| < 0.4$ bin for $W^+ \to \mu^+\nu$ channel.

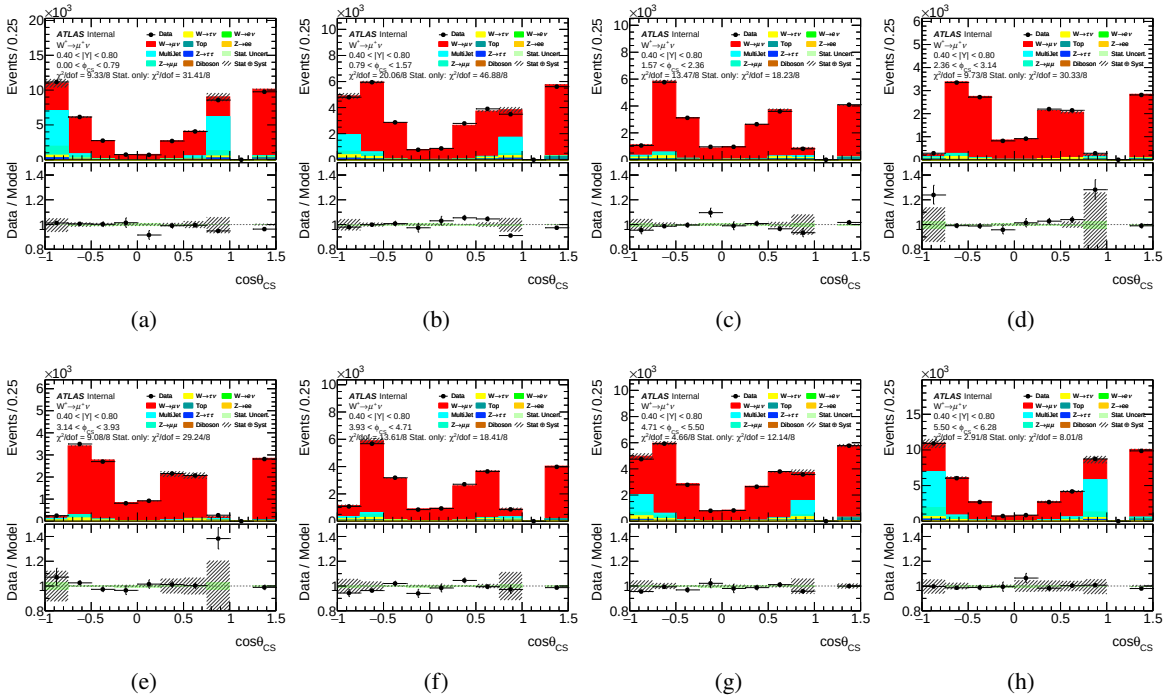


Figure 384: Control plots for $0.4 < |Y| < 0.8$ bin for $W^+ \to \mu^+\nu$ channel.

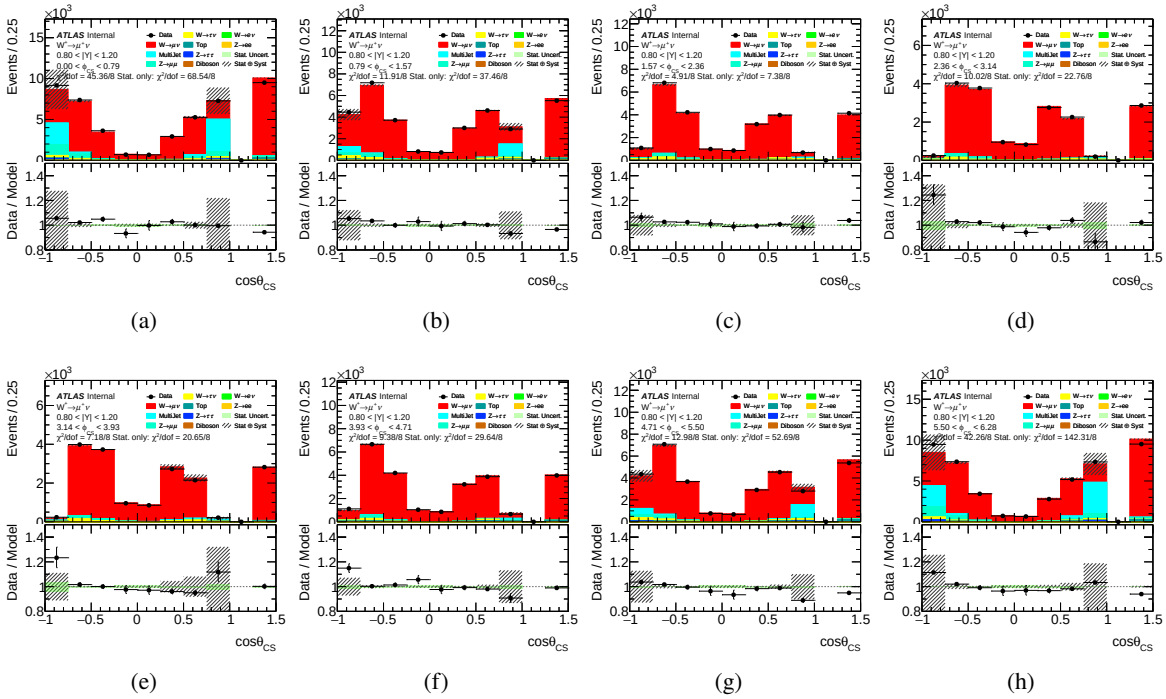
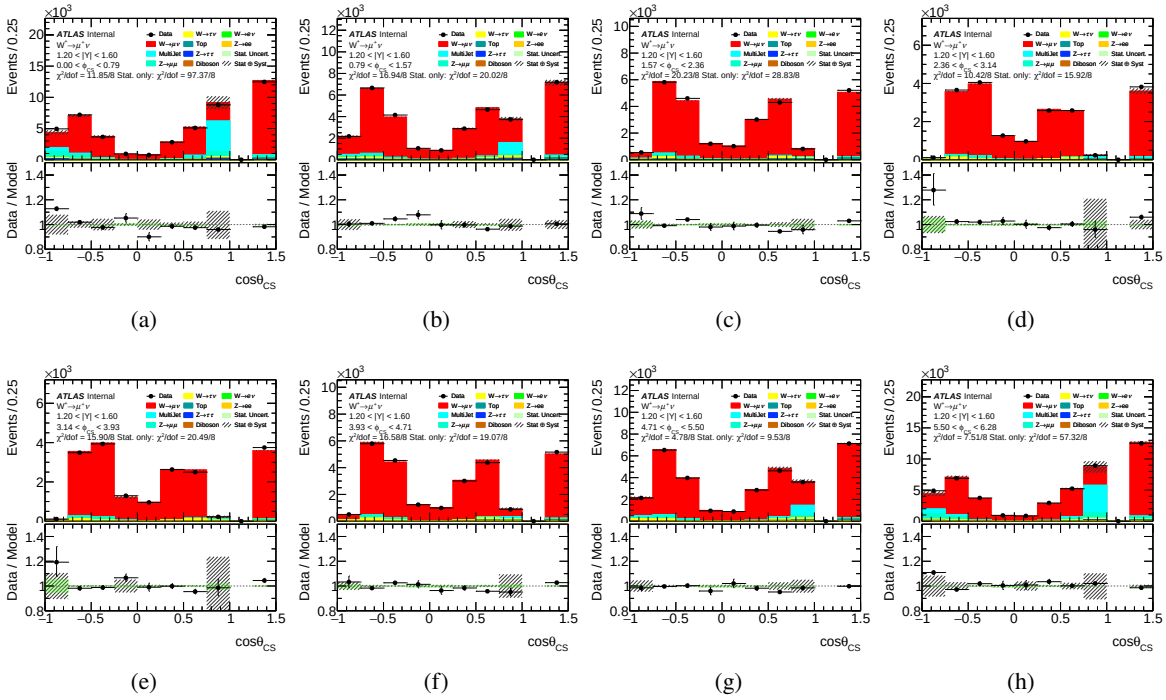


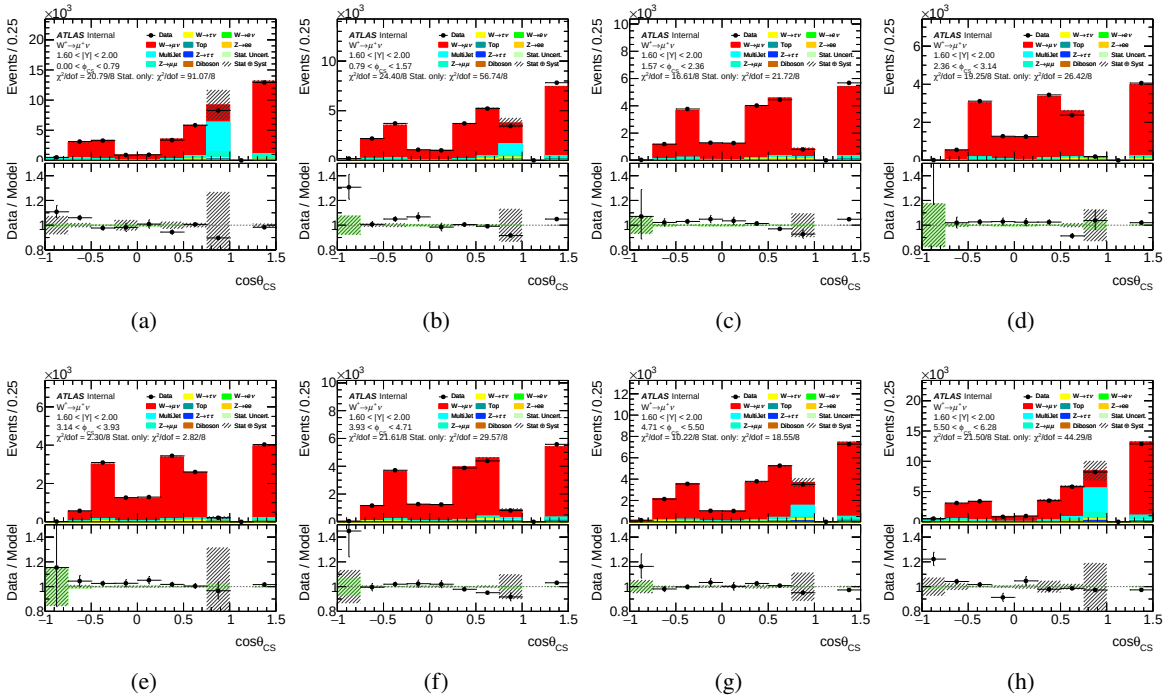
Figure 385: Control plots for $0.8 < |Y| < 1.2$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

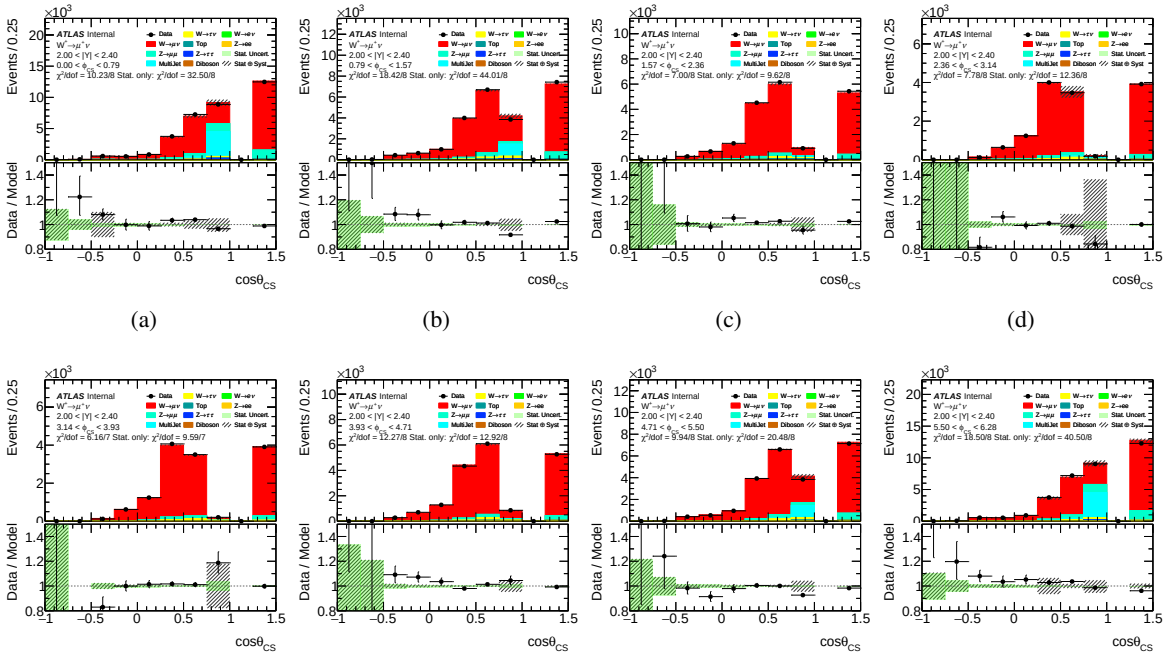
Figure 386: Control plots for $1.2 < |Y| < 1.6$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

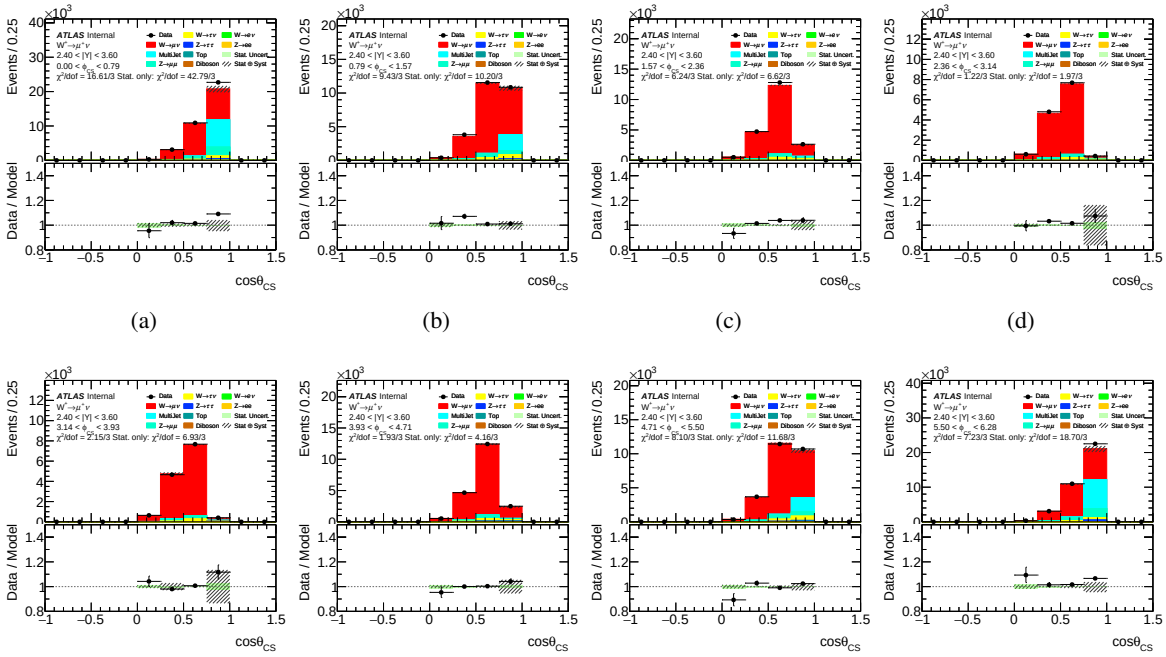
Figure 387: Control plots for $1.6 < |Y| < 2.0$ bin for $W^+ \to \mu^+\nu$ channel.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 388: Control plots for $2.0 < |Y| < 2.4$ bin for $W^+ \to \mu^+\nu$ channel.

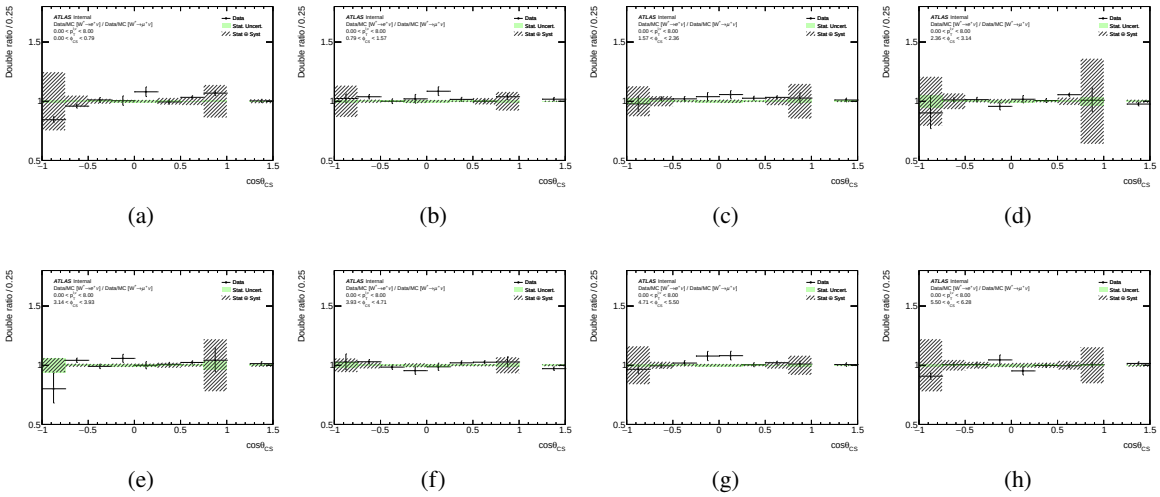


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 389: Control plots for $2.4 < |Y| < 3.6$ bin for $W^+ \to \mu^+\nu$ channel.

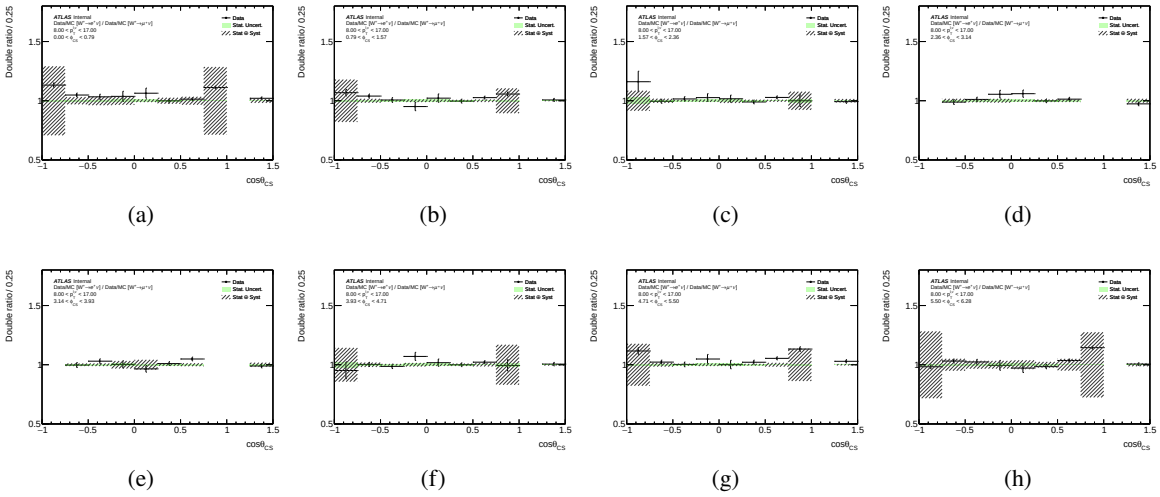
F.6 Double ratio $W^+ \to e^+\nu / W^+ \to \mu^+\nu$



(a) (b) (c) (d)

(e) (f) (g) (h)

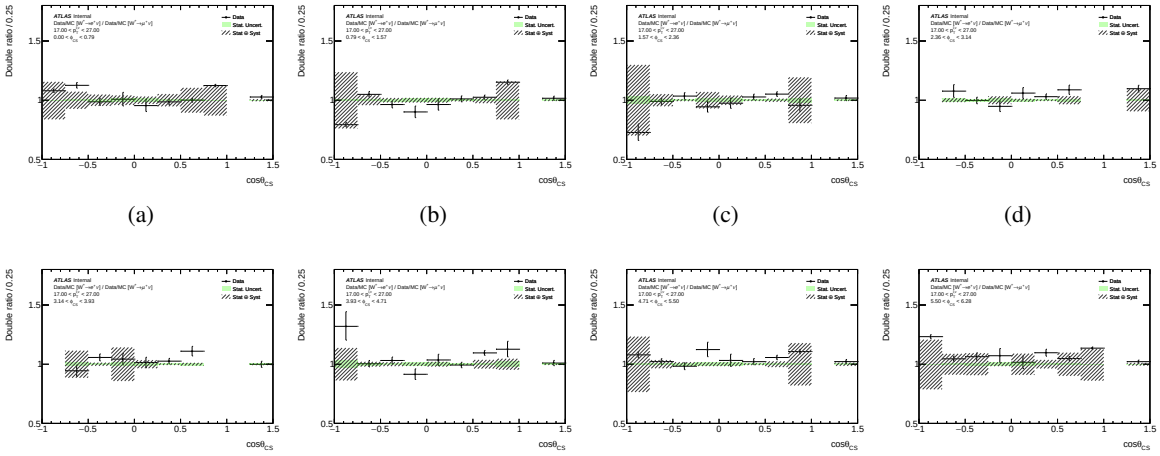
Figure 390: Double ratio plots for e/μ for $0 < p_T^{\ell,\nu} < 8$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

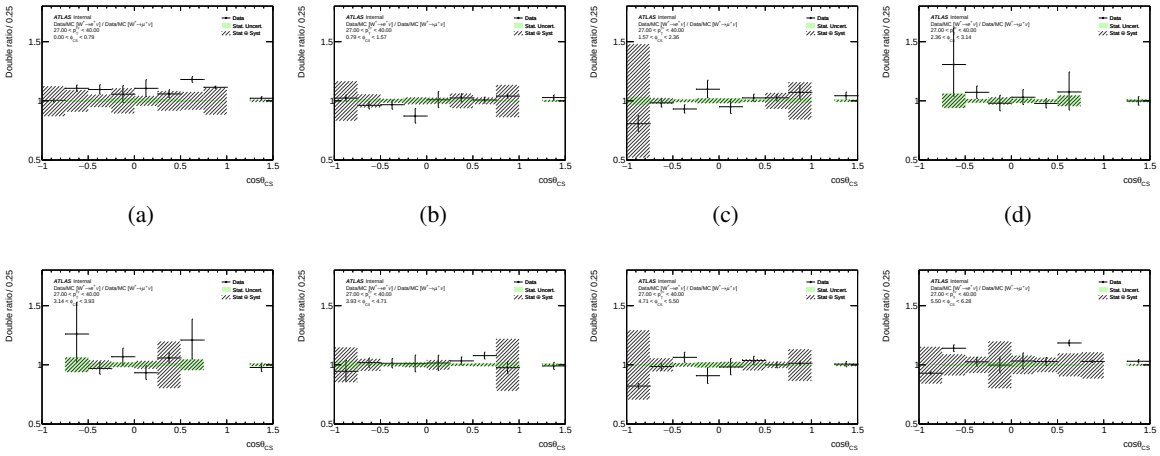
Figure 391: Double ratio plots for e/μ for $8 < p_T^{\ell,\nu} < 17$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

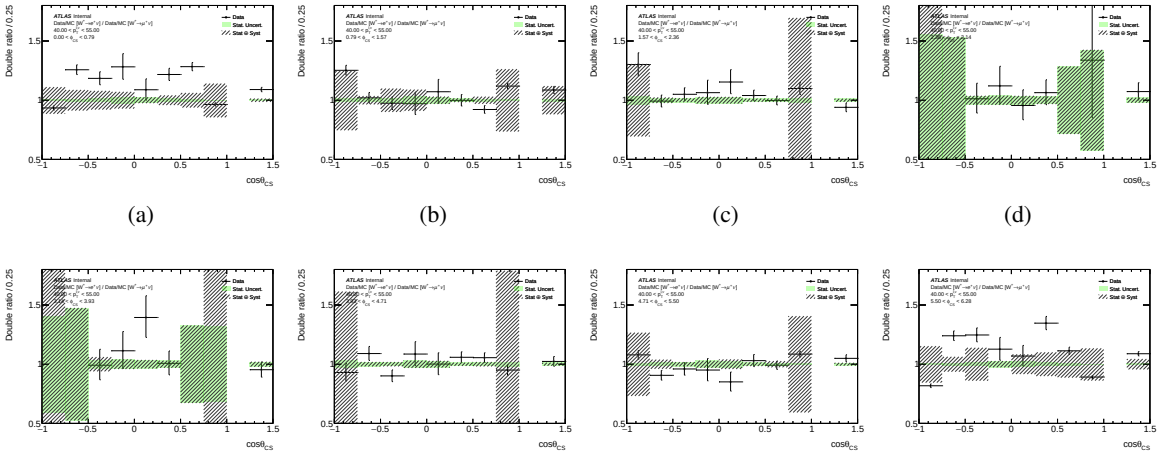
Figure 392: Double ratio plots for e/μ for $17 < p_T^{\ell,\nu} < 27$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

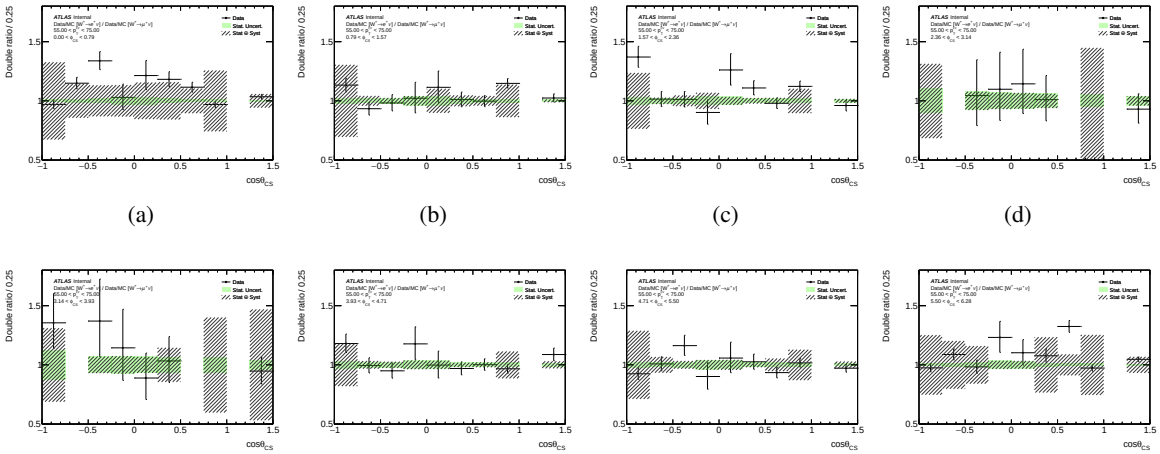
Figure 393: Double ratio plots for e/μ for $27 < p_T^{\ell,\nu} < 40$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

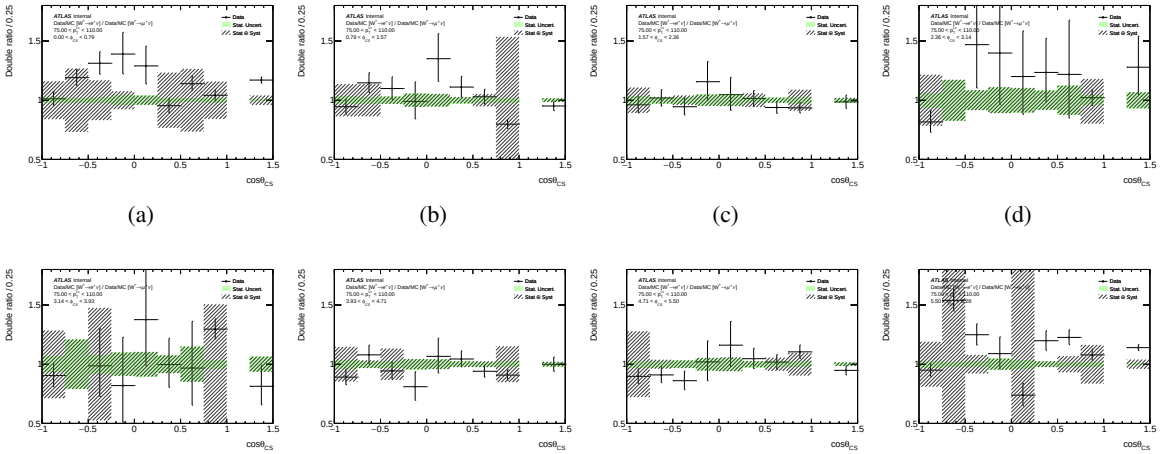
Figure 394: Double ratio plots for e/μ for $40 < p_T^{\ell,\nu} < 55$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

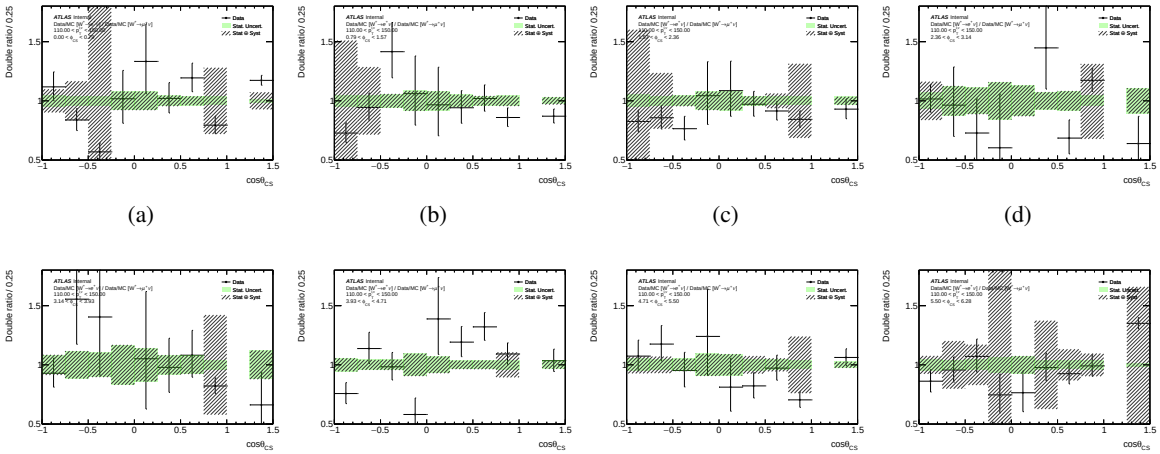
Figure 395: Double ratio plots for e/μ for $55 < p_T^{\ell,\nu} < 75$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

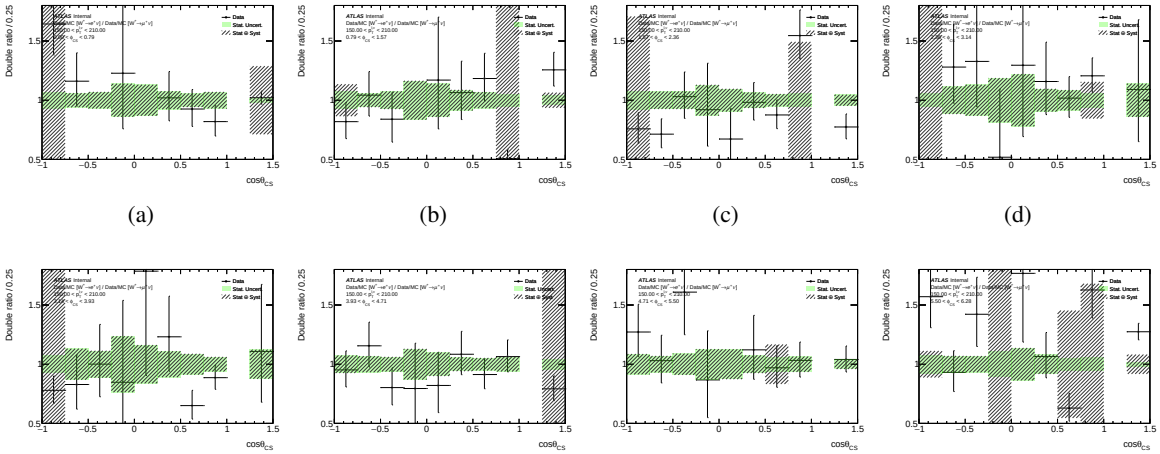
Figure 396: Double ratio plots for e/μ for $75 < p_T^{\ell,\nu} < 110$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

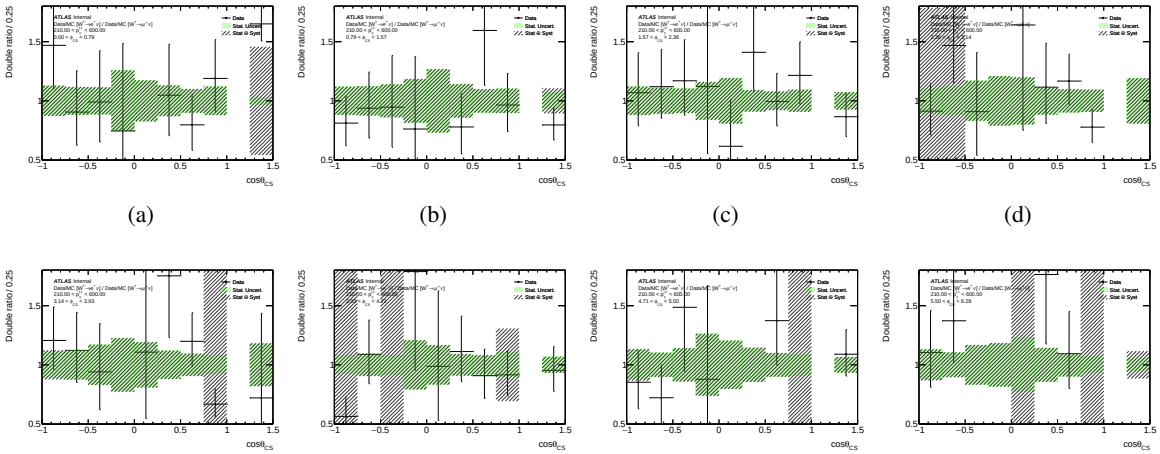
Figure 397: Double ratio plots for e/μ for $110 < p_T^{\ell,\nu} < 150$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

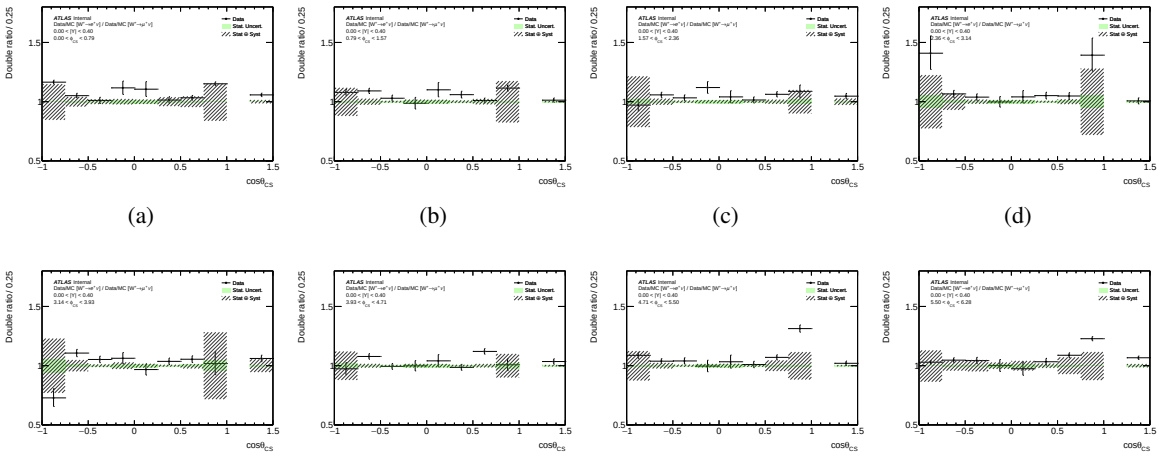
Figure 398: Double ratio plots for e/μ for $150 < p_T^{\ell,\nu} < 210$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

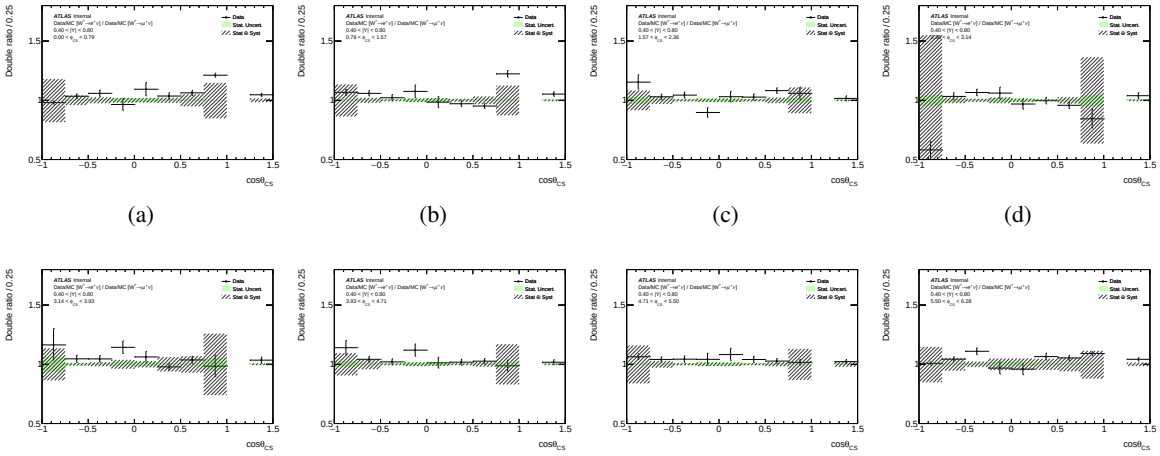
Figure 399: Double ratio plots for e/μ for $210 < p_T^{\ell,\nu} < 600$ GeV bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

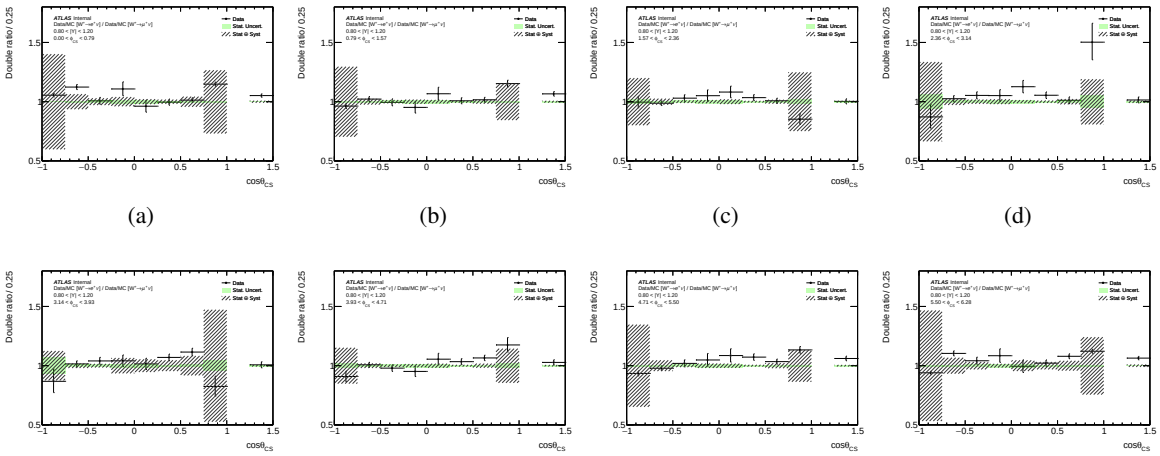
Figure 400: Double ratio plots for e/μ for $0 < |Y| < 0.4$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

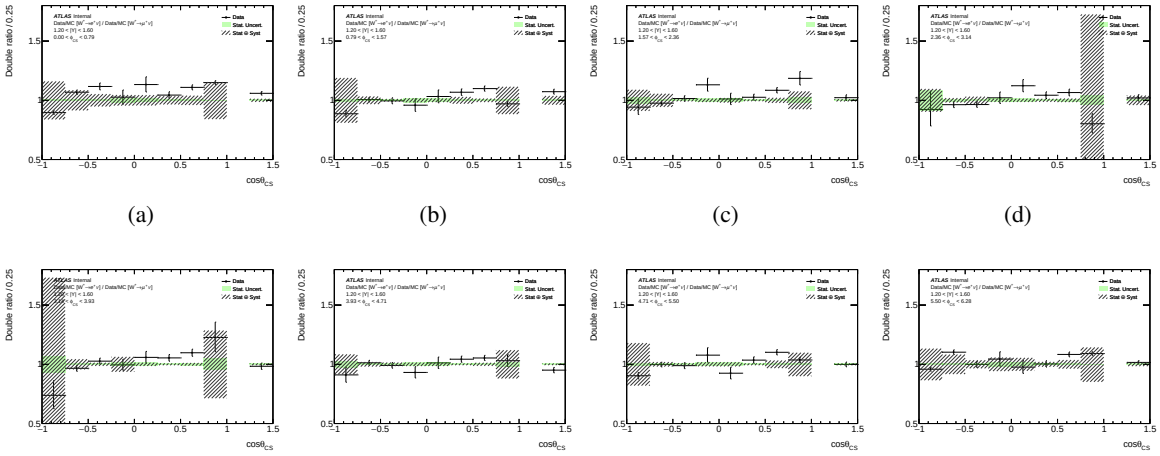
Figure 401: Double ratio plots for e/μ for $0.4 < |Y| < 0.8$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

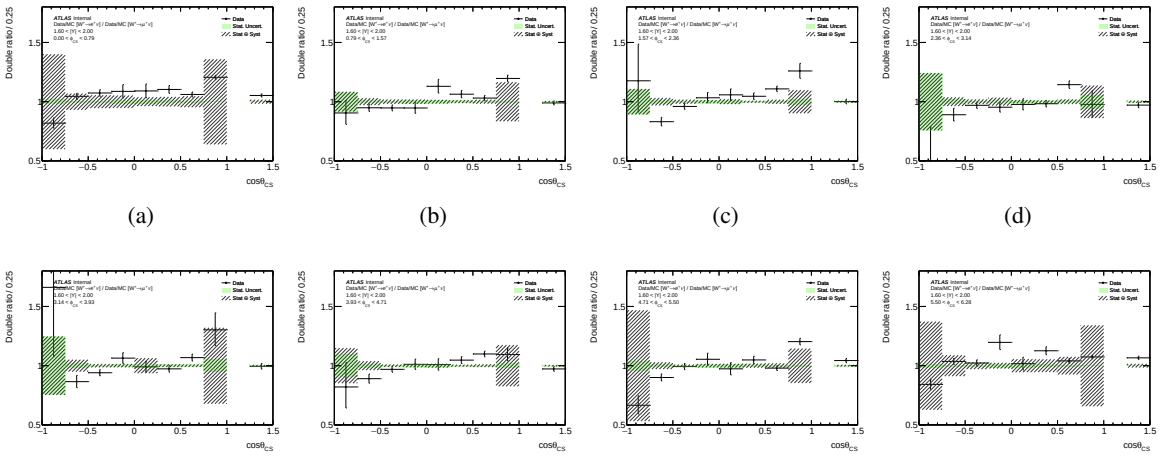
Figure 402: Double ratio plots for e/μ for $0.8 < |Y| < 1.2$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

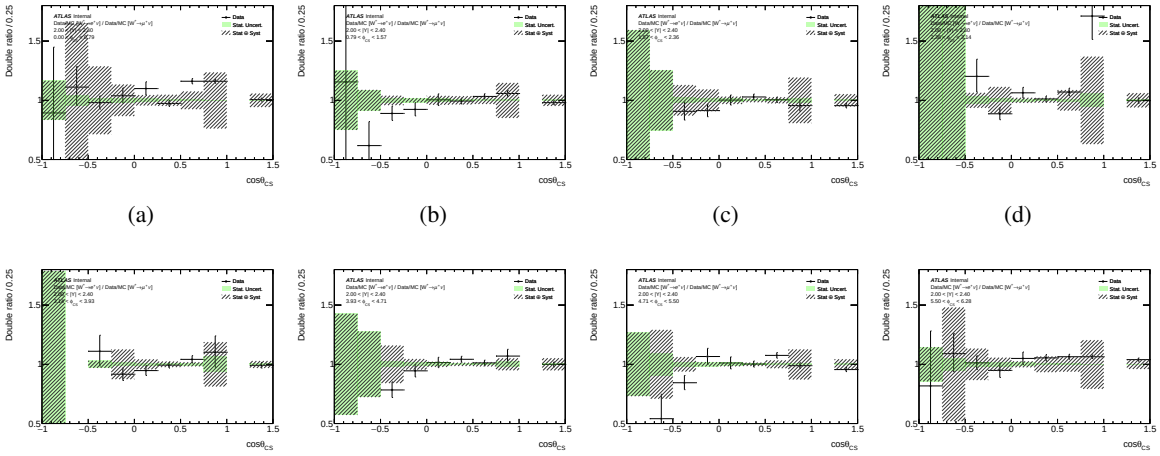
Figure 403: Double ratio plots for e/μ for $1.2 < |Y| < 1.6$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

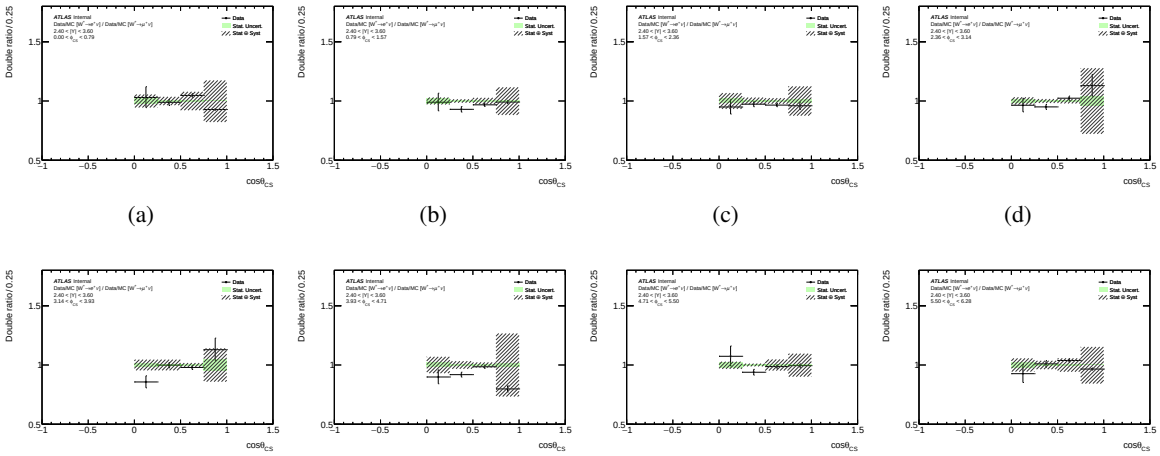
Figure 404: Double ratio plots for e/μ for $1.6 < |Y| < 2.0$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 405: Double ratio plots for e/μ for $2.0 < |Y| < 2.4$ bin.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 406: Double ratio plots for e/μ for $2.4 < |Y| < 3.6$ bin.

G FTL

G.1 Newton's Method

Based on Newton's method which is an iterative root-finding method based on Taylor expansion of a function. In 1D:

$$y \approx f(x_n) + \delta f(x_n)/\delta x \times (x_{n+1} - x_n) \approx 0 \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{[\delta f(x_n)/\delta x]}. \tag{63}$$

To minimize a function $g(x)$, in the case of this analysis $-\ln \mathcal{L}$, we set $f(x) = \delta g(x)/\delta x$ to find $\delta g(x_n)/\delta x = 0$ so x_{n+1} becomes

$$x_{n+1} = x_n - \frac{\delta g(x_n)/\delta x}{\left[\delta^2 g(x_n)/\delta^2 x\right]}. \tag{64}$$

Extrapolating this to N dimensions

$$\begin{aligned} & x \to \vec{x} \\ & \delta g(x_n)/\delta x \to \nabla g(\vec{x}) \\ & 1/\left[\delta^2 g(x_n)/\delta^2 x\right] \to \mathbf{H}^{-1} = \left[\nabla^2 g(\vec{x})\right]^{-1} \\ & \Rightarrow \vec{x}_{n+1} = \vec{x}_n - \mathbf{H}^{-1} \nabla g(\vec{x}) \end{aligned} \tag{65}$$

Can introduce μ parameter as $\vec{x}_{n+1} = \vec{x}_n - \mu \mathbf{H}^{-1} \nabla g(\vec{x})$ and perform line search to find minimum $-\ln \mathcal{L}$ for each iteration n. The bulk of the computation time is spent on computing $\mathbf{H}$ and solving the system of equations for $\mathbf{H}\vec{x}_{\text{corr}} = \nabla g(\vec{x})$.

G.2 Uncertainty Decomposition

The total uncertainty on a given parameter is the square root of the diagonal of the covariance matrix $\mathbf{H}^{-1}$. For a group of parameters 'm' we want to find the component uncertainty of we need to block decompose covariance matrix M:

$$M = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \tag{66}$$

If M is of size $n \times n$ then A is $(n-m) \times (n-m)$, C is $m \times m$, and B is $(n-m) \times m$. The Schur complement of M, $M/C = A - BC^{-1}B^\top$, is the conditional covariance matrix of M given C, which is the covariance matrix with parameters in C fixed to constant. The diagonals of $BC^{-1}B^\top$ gives us the square of the component uncertainty for the parameters A due to the parameters in C. This process is then repeated for each group of parameters.

A multivariate Gaussian likelihood can be built directly from the Hessian $\mathbf{H}$, parameters $\vec{x}$, and their best fit values $\vec{\mu}$ as $-\ln \mathcal{L}(\vec{x}) = 0.5(\vec{x} - \vec{\mu})\,\mathbf{H}\,(\vec{x} - \vec{\mu})^\top$ which could be used to find the uncertainty decomposition for alternative parameterizations.

H ID and Isolation Scale Factors

Relaxing the kinematic selection requirements to not include one for $E_{\mathrm{T}}^{\mathrm{miss}}$ or m_T was chosen to allow for more signal at high p_{T}^{W} but allowed for more MJ background. In order to suppress the MJ an additional TightLLH_d0z0 ID selection was added to electron channels and a calorimeter isolation requirement of topoetcone20$/p_{\mathrm{T}}^{\ell} < 0.05$ for all channels. These selection requirements need associated scale factors (SFs) to account for the discrepancy between MC and data efficiencies so this appendix covers the process of deriving the SFs.

To produce these SFs events are pre-selected to have leptons that pass all of the other selection requirements used in the analysis so that these SFs are applied in addition the SFs used already. During the selection process the events were weighted with the SFs for existing selection requirements like ID, isolation, and calibration. From there the tag and probe method was used with events broken up into two categories,

- Leading lepton - tag — subleading lepton - probe
- Subleading lepton - tag — leading lepton - probe

These were then combined into one sample before calculating the efficiency, $\epsilon = N_{\mathrm{probes}}^{\mathrm{passed}}/N_{\mathrm{probes}}$ for data and MC using Bayesian statistics. The SF is then calculated by taking the ratio of the data efficiency over the MC efficiency. This was done as a function of p_{T}^{ℓ} and η^{ℓ} as seen in Figures 407, 408, and 409 for electron TightLLH_d0z0 ID, electron calorimeter isolation, and muon isolation respectively.

To estimate a systematic uncertainty instead of combining the previously stated tag and probe categories they were processed separately to produce SFs denoted by the red and blue points in Figures 410, 411, and 412. With the separate SFs a weighted average is taken between them denoted by the magenta points in the figures. The systematic uncertainty is then the difference between the SF from combining the samples and the weighted average between them.

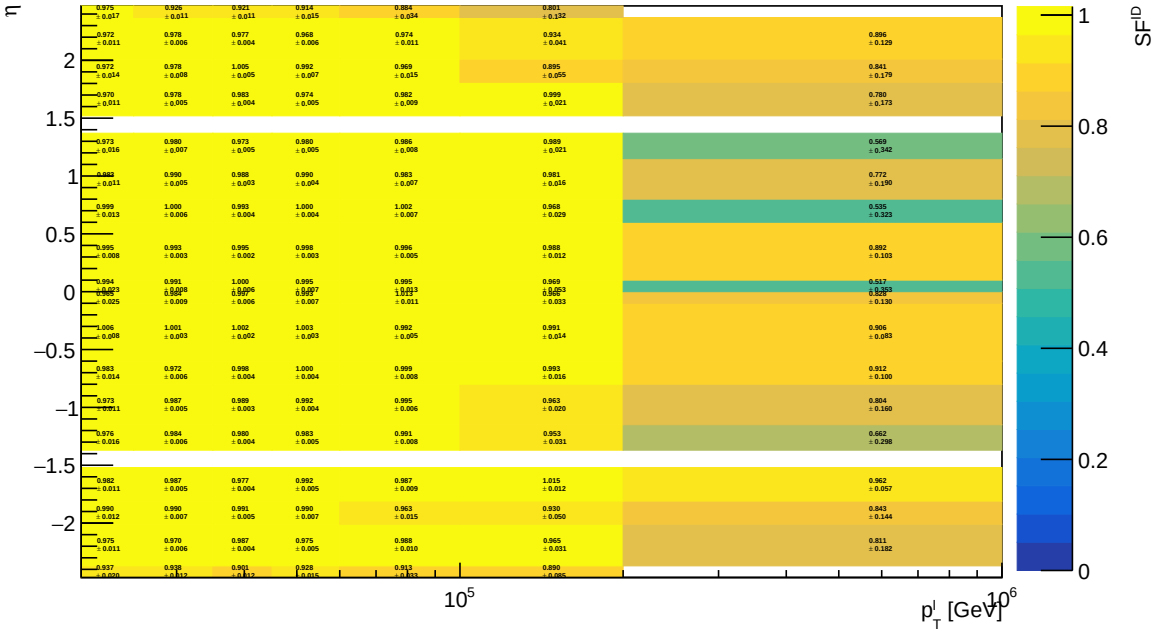


Figure 407: Scale factors derived for TightLLH_d0z0 ID selection requirements for electrons as a function of p_T^e and η^e.

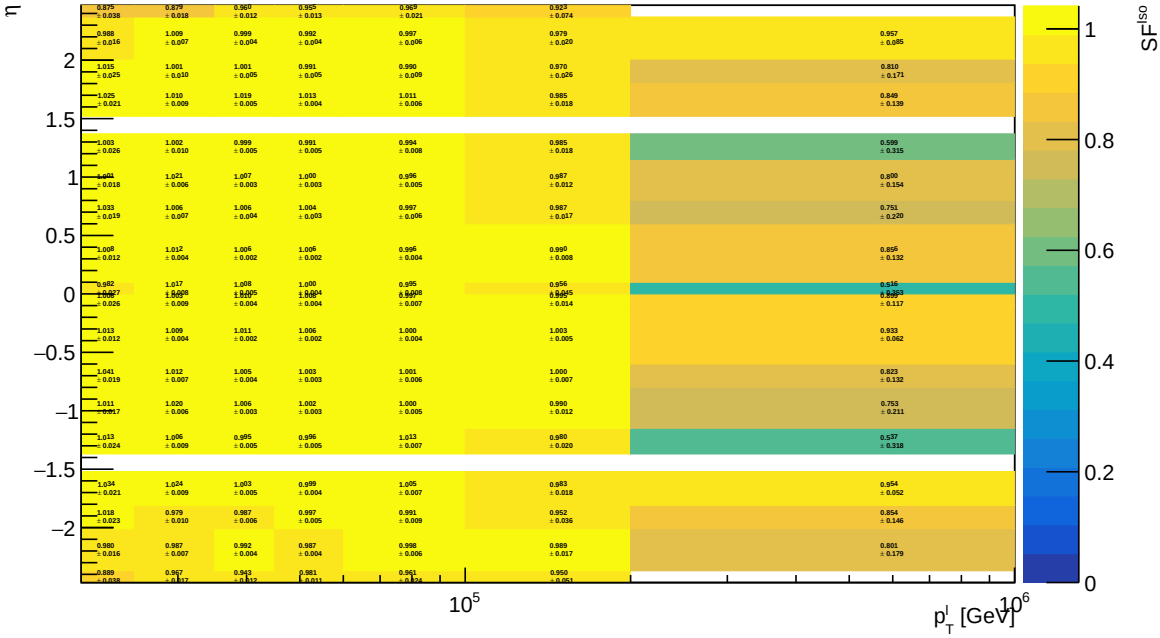


Figure 408: Scale factors derived for topoetcone20/p_T^e < 0.05 isolation selection requirements for electrons as a function of p_T^e and η^e.

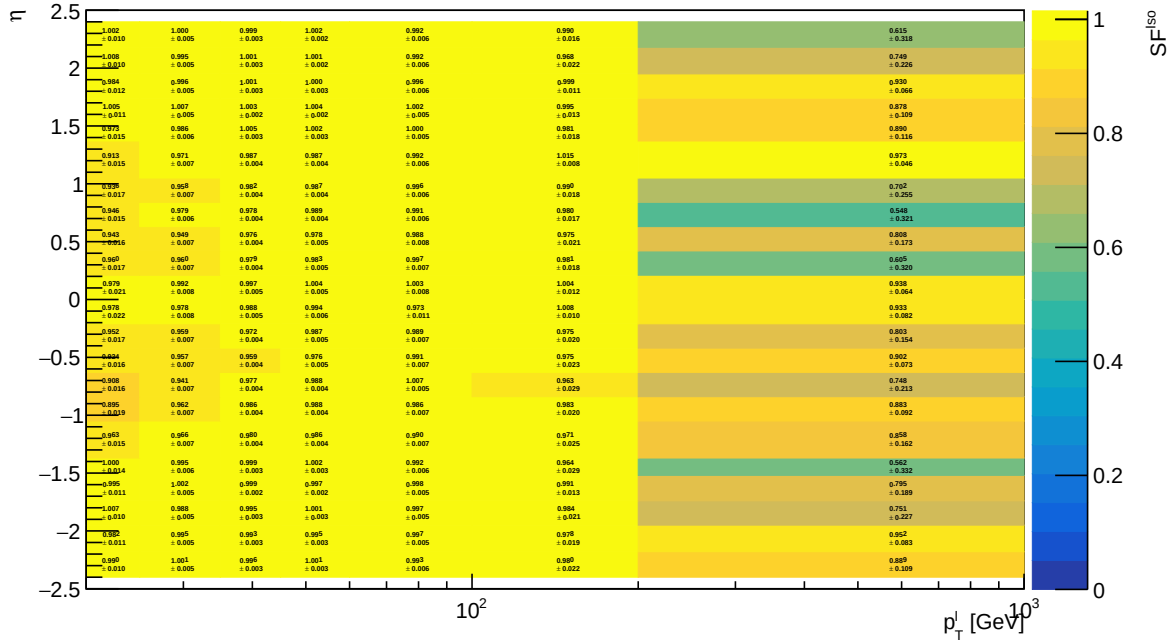
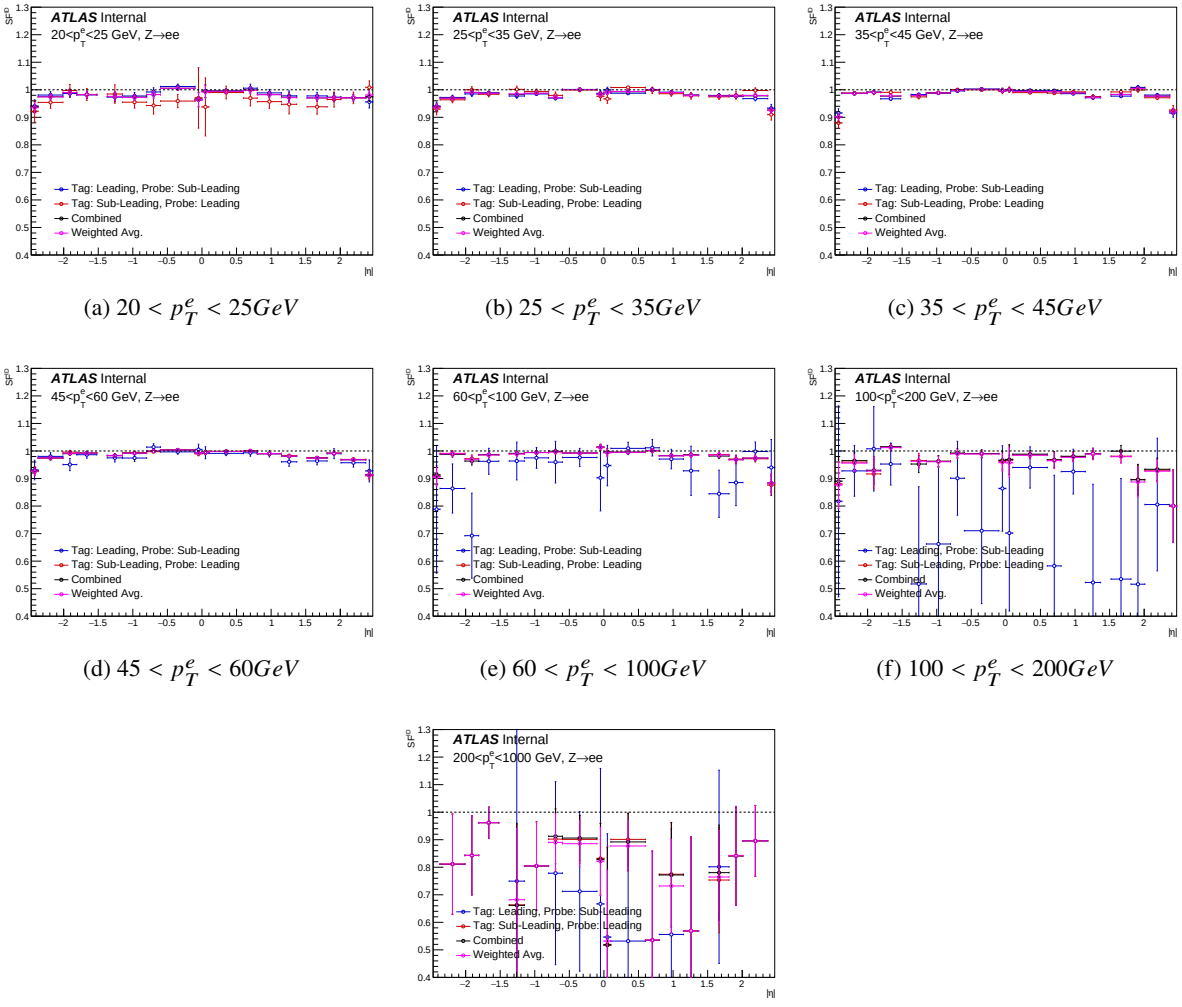


Figure 409: Scale factors derived for topoetcone20/p_T^μ < 0.05 isolation selection requirements for electrons as a function of p_T^μ and η^μ.

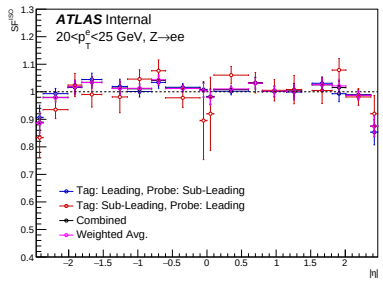


(a) $20 < p_T^e < 25GeV$ (b) $25 < p_T^e < 35GeV$ (c) $35 < p_T^e < 45GeV$

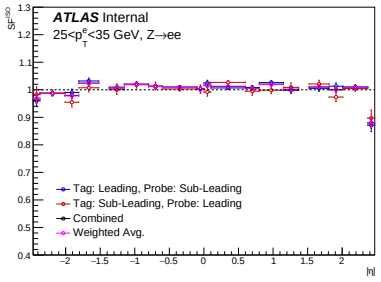
(d) $45 < p_T^e < 60GeV$ (e) $60 < p_T^e < 100GeV$ (f) $100 < p_T^e < 200GeV$

(g) $200 < p_T^e < 1000GeV$

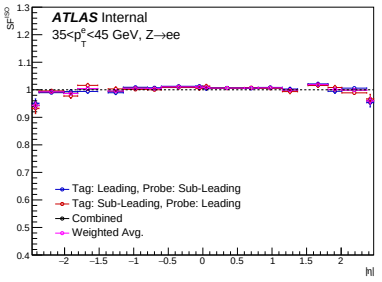
Figure 410: Scale factors derived for TightLLH_d0z0 ID selection requirements for electrons as a function for individual p_T^e slice. The final SFs used are denoted by combined in black and are compared to deriving using each category individually in red and blue and the weighted average of these categories in magenta.



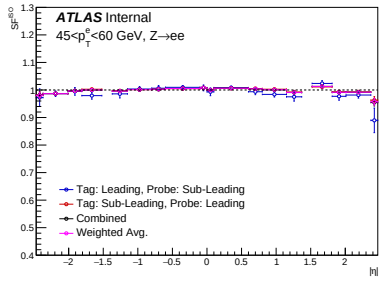
(a) $20 < p_T^e < 25GeV$



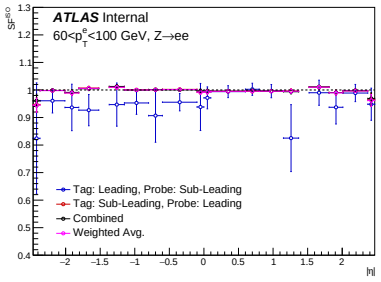
(b) $25 < p_T^e < 35GeV$



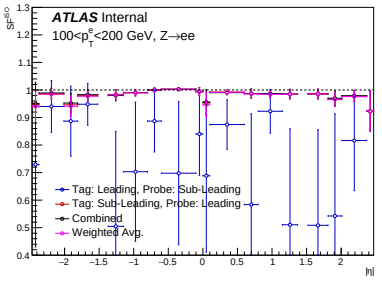
(c) $35 < p_T^e < 45GeV$



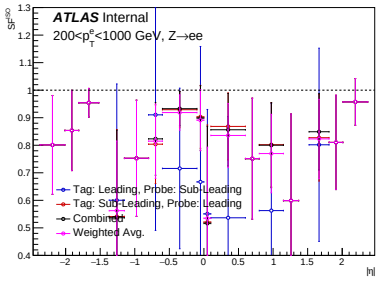
(d) $45 < p_T^e < 60GeV$



(e) $60 < p_T^e < 100GeV$

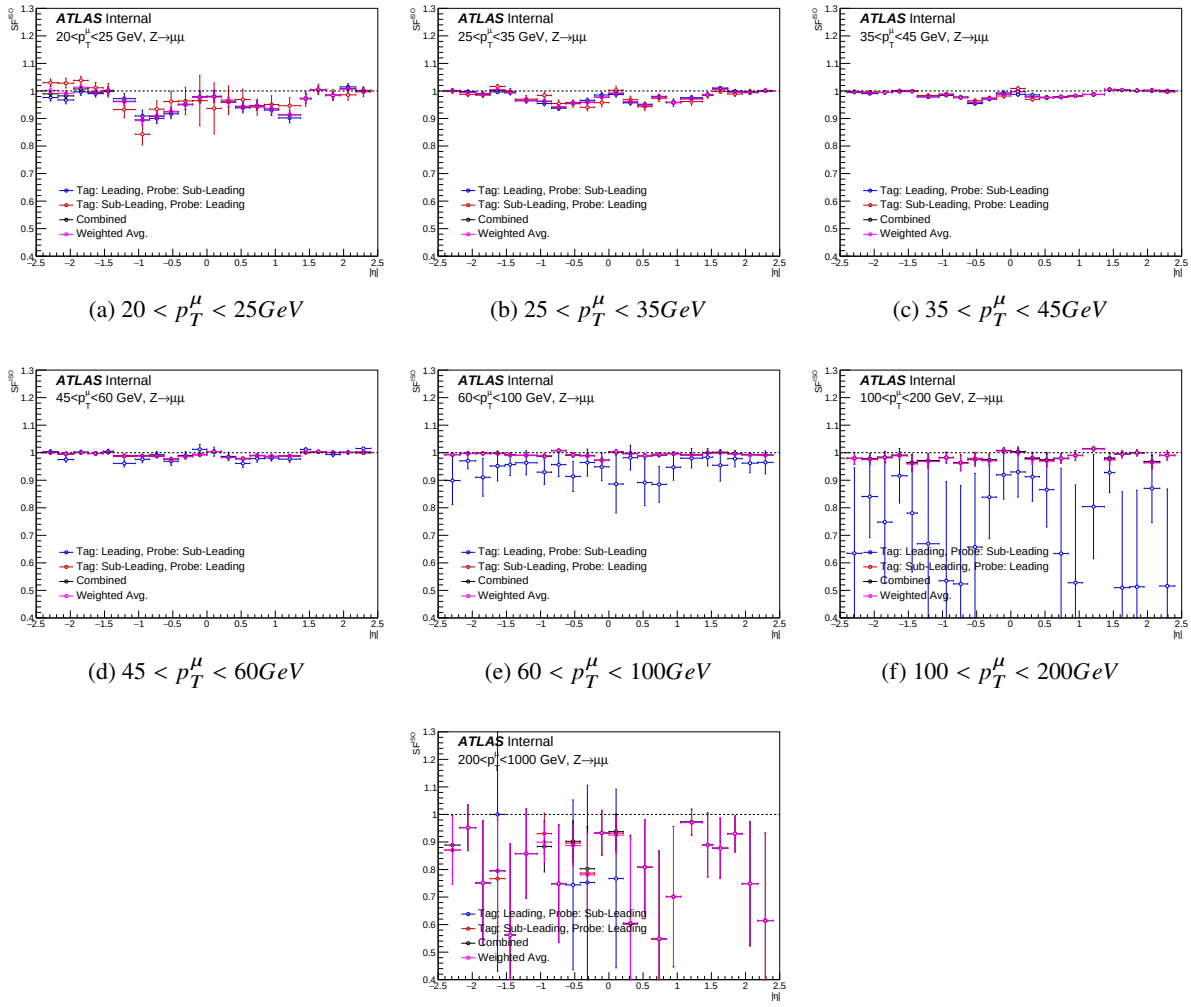


(f) $100 < p_T^e < 200GeV$



(g) $200 < p_T^e < 1000GeV$

Figure 411: Scale factors derived for topoetcone20/p_T^e < 0.05 isolation selection requirements for electrons as a function for individual p_T^e slice. The final SFs used are denoted by combined in black and are compared to deriving using each category individually in red and blue and the weighted average of these categories in magenta.



(a) $20 < p_T^\mu < 25GeV$ (b) $25 < p_T^\mu < 35GeV$ (c) $35 < p_T^\mu < 45GeV$

(d) $45 < p_T^\mu < 60GeV$ (e) $60 < p_T^\mu < 100GeV$ (f) $100 < p_T^\mu < 200GeV$

(g) $200 < p_T^\mu < 1000GeV$

Figure 412: Scale factors derived for topoetcone20/p_T^μ < 0.05 isolation selection requirements for electrons as a function for individual p_T^μ slice. The final SFs used are denoted by combined in black and are compared to deriving using each category individually in red and blue and the weighted average of these categories in magenta.