

ABSTRACT

SEARCH FOR CHARGED HIGGS BOSONS IN THE $\tau + \ell$ FINAL STATE WITH 36.1 fb^{-1} OF pp COLLISION DATA AT $\sqrt{s} = 13 \text{ TeV}$ WITH THE ATLAS EXPERIMENT

Elliot Wesley Parrish, Ph.D.
Department of Physics
Northern Illinois University, 2022
Dhiman Chakraborty and Jahred Adelman, Director

This dissertation uses 139 fb^{-1} of pp collision data collected at a center of mass energy of $\sqrt{s} = 13 \text{ TeV}$ by the ATLAS detector to search for charged Higgs bosons decaying to a tau lepton and a neutrino ($H^\pm \rightarrow \tau^\pm \nu_\tau$) in association with a leptonically decaying top quark. No significant excess was found, therefore limits are set at the 95% confidence level on the charged Higgs production cross section times the branching fraction into the $\tau^\pm \nu_\tau$ ranging from XX pb to XX fb. These limits are interpreted in the hMSSM benchmark scenario as an exclusion at 95% confidence on $\tan \beta$ as a function of m_{H^\pm} . In this scenario, for $\tan \beta = 60$, the H^\pm mass range up to XXX GeV is excluded, with all values of $\tan \beta$ excluded for $m_{H^\pm} \leq XXX \text{ GeV}$.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

DECEMBER 2022

**SEARCH FOR CHARGED HIGGS BOSONS IN THE $\tau + \ell$ FINAL STATE
WITH 36.1 fb^{-1} OF pp COLLISION DATA AT $\sqrt{s} = 13 \text{ TeV}$ WITH THE
ATLAS EXPERIMENT**

BY

ELLIOT WESLEY PARRISH
© 2022 Elliot Wesley Parrish

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

Dissertation Director:
Dhiman Chakraborty and Jahred Adelman

ACKNOWLEDGEMENTS

DEDICATION

To Dr. Dhiman Chakraborty. Thank you for everything.

TABLE OF CONTENTS

	Page
List of Tables	v
List of Figures.	vii
List of Appendices.	xii
Chapter	
1 Search for Charged Higgs Bosons	1
1.1 Signature and Event Selection	2
1.1.1 Object Definitions.	3
1.1.2 Event Selections	4
1.2 Datasets.	5
1.2.1 Signal Modeling	5
1.3 Background Modeling	7
1.4 Multivariate Analysis Techniques	15
1.4.1 Training.	16
1.4.2 Input Variables Selection.	18
1.4.3 Hyperparameter Optimization	20
1.5 Systematic Uncertainties	25
1.6 Results.	26
Appendices	33
Appendices	34

LIST OF TABLES

Table		Page
1.1	H^\pm production mechanisms based on m_{H^\pm} , dominant H^\pm decay mode, and the main background associated with the diagram.	2
1.2	Definitions of physics objects used in this analysis.	3
1.3	τ +jets and $\tau + \ell$ signal region definitions.	4
1.4	For each H^\pm mass the generator xcross-section ($\sigma \times BR(H^\pm \rightarrow \tau^\pm \nu_\tau)$) is given, as well as the number of generated events for both $\tau + \ell$ and τ +jets subchannels.	6
1.5	Dominant backgrounds from prompt $\tau_{had-vis}$ and fake $\tau_{had-vis}$ candidates. . .	7
1.6	Cross sections for the main SM background samples at $\sqrt{s} = 13$ TeV . Here, ℓ refers to the three lepton families e, μ and τ . All background cross sections are normalized to NNLO predictions, except for diboson events, where the NLO prediction is used. A '*' indicates that the quoted cross section for the sample is neglecting leptonic/hadronic branching ratios.	8
1.7	Control region definitions for the τ +jets subchannel.	9
1.8	Control region definitions for the $\tau + \ell$ subchannel.	10
1.9	List of low level (a) and high level (b) kinematic variables used as input to the PNN in the $\tau + \ell$ subchannel. $\Delta\phi_{X,miss}$ denotes the difference in azimuthal angle between a reconstructed object X ($X = \tau, b-jet, \ell$) and the direction of the missing transverse momentum.	19
1.10	First grid, scanning over activation function and loss function. Binary crossentropy was the chosen loss function, highlighted in red.	21
1.11	Second grid, scanning over width, depth, and dropout value. 0.1 was chosen for the dropout value, highlighted in red.	21
1.12	Third grid, scanning over activation function. LeakyReLU was chosen, highlighted in red.	22

Table	Page
1.13 Fourth grid, scanning over LeakyReLU α value. $\alpha = 0.05$ was chosen, highlighted in red.	22
1.14 Fourth grid, scanning over network width and depth. $width = 128$ and $depth = 3$ was chosen, highlighted in red.	22
1.15 AUCs of final HPO grid	23
1.16 Average AUCs of final HPO grid	23
1.17 Effect on the shape variation and the yields of systematic uncertainties associated with the data-driven fake factor method, used to estimate the $j \rightarrow \tau$ background in the τ +jets and $\tau + \ell$ channel.	25
1.18 Expected event yields for the backgrounds and a hypothetical H^\pm signal after applying all τ +jets selection criteria, and comparison with 139 fb^{-1} of data. All yields are evaluated prior to using the multivariate discriminant and applying the statistical fitting procedure. The values shown for the signal assume a charged Higgs boson mass of 170 GeV and 1000 GeV, with a cross-section times branching fraction $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ corresponding to $\tan\beta = 40$ in the hMSSM benchmark scenario. Statistical uncertainties are quoted. Systematic uncertainties to be added.	26
1.19 Expected event yields for the backgrounds and a hypothetical H^\pm signal after applying all $\tau + \ell$ selection criteria, and comparison with 139 fb^{-1} of data. All yields are evaluated prior to using the multi-variate discriminant and applying the statistical fitting procedure. The values shown for the signal assume a charged Higgs boson mass of 170 GeV and 1000 GeV, with a cross-section times branching fraction $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ corresponding to $\tan\beta = 40$ in the hMSSM benchmark scenario. Statistical uncertainties are quoted. Systematic uncertainties to be added.	27

LIST OF FIGURES

Figure		Page
1.1	Signal acceptance as a function of the charged Higgs boson mass for both the τ +jets (a) and $\tau + \ell$ subchannels (b).....	5
1.2	Mass of $\tau - e$ system in the Zee control region.....	8
1.3	Combined fake factors for the τ +jets b-veto $m_T > 100$ control region, τ +jets signal region, τ +electron(muon) with same-sign control region and the $\tau + \ell$ signal region. Error bars represent systematic uncertainties of the method. .	12
1.4	Comparison between the predicted and the measured E_T^{miss} distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.	13
1.5	Comparison between the predicted and the measured E_T^{miss} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.	14
1.6	Comparison of performance of an optimized BDT and an unopitimized PNN on expected limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ in the τ +jets (a) and $\tau + \ell$ (b) signal regions.	15
1.7	<i>Left</i> , individual networks with input variables (x_1, x_2) , each trained with examples with a single value of some parameter $\theta = \theta_a, \theta_b$. The individual networks are purely functions of the input variables. Performance for intermediate values of θ is not optimal nor does it necessarily vary smoothly between the networks. <i>Right</i> , a single network trained with input variables (x_1, x_2) as well as input parameter θ ; such a network is trained with examples at several values of the parameter θ [3].	17
1.8	The k-fold method for $k = 5$ [7].....	17
1.9	Expected limits comparing a set of high level variables and low level variables with various depths in the PNN architecture. X layers refers to the number of layers in the PNN.	20

Figure	Page
1.10 Final model AUC for each mass point. Individual points correspond to the AUC average over 5 kfold.	24
1.11 LeakyReLU activation function. The associated hyperparameter α is the slope of the negative portion of the function.	25
1.12 PNN score distributions in the signal region of the τ +jets channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.	29
1.13 PNN score distributions in the signal region of the τ +e sub-channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.	30
1.14 PNN score distributions in the signal region of the τ + μ sub-channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.	31
1.15 Expected 95% CL exclusion limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ as a function of the charged Higgs boson mass in 139 fb^{-1} of pp collision data at $\sqrt{s} = 13\text{ TeV}$ in the τ +jets signal region (a), $\tau + \ell$ signal region (b), and the combination of the τ +jets and $\tau + \ell$ signal regions (c). In the case of the expected limits, one- and two-standard-deviation uncertainty bands are also shown. As a comparison, the expected exclusion limits obtained with the dataset collected in 2015 and 2016 [8] are also shown.	32
B.1 Fake factors parameterized as a function of p_T^τ and the number of charged τ decay products (1-prong and 3-prong) obtained in the multi-jet and W+jets CRs. The errors shown represent the statistical uncertainty.	36
B.2 Corrected α_{MJ} values for the $\tau_{\text{had-vis}}$ +jets b-veto $m_T > 100$ control region, $\tau_{\text{had-vis}}$ +jets signal region, $\tau_{\text{had-vis}}$ +electron(muon) with same-sign control region and the $\tau_{\text{had-vis}}$ +lepton signal region. Error bars represent uncertainties due to α_{MJ} fitting using template-fit method.	36

- B.3 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +jets signal region for $p_{\text{T}} \leq 60$ GeV 1-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 37
- B.4 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +jets signal region for $p_{\text{T}} \geq 60$ GeV 1-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 38
- B.5 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +jets signal region for 3-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 39
- B.6 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +lepton signal region for $p_{\text{T}} \leq 50\text{GeV}$ 1-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 40
- B.7 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +lepton signal region for $p_{\text{T}} \geq 50\text{GeV}$ 1-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 41

- B.8 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +lepton signal region for $p_{\text{T}} \leq 60\text{GeV}$ 3-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 42
- B.9 Estimation of α_{MJ} in the $\tau_{\text{had-vis}}$ +lepton signal region for $p_{\text{T}} \geq 60\text{GeV}$ 3-prong $\tau_{\text{had-vis}}$ candidates. Left: templates of discriminating variables for different $\tau_{\text{had-vis}}$ p_{T} and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\text{min}}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$ 43
- C.1 Comparison between the predicted and the measured p_{T}^{τ} distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction. 45
- C.2 Comparison between the predicted and the measured $p_{\text{T}}^{b\text{-jet}}$ distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction. 46
- C.3 Comparison between the predicted and the measured p_{T}^{τ} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction. 48
- C.4 Comparison between the predicted and the measured p_{T}^{ℓ} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction. 49
- C.5 Comparison between the predicted and the measured p_{T}^e distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction. 50

Figure		Page
C.6	Comparison between the predicted and the measured p_T^μ distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.	51
C.7	Comparison between the predicted and the measured p_T^{b-jet} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.	52

LIST OF APPENDICES

Appendix	Page
A TileCal Data Quality	34
B Fake Factors.	35
C Additional Validation Plots	44
C.1 τ +jets Validation Plots.	45
C.2 $\tau + \ell$ Validation Plots	47

CHAPTER 1

SEARCH FOR CHARGED HIGGS BOSONS

This chapter details a search for a charged Higgs boson decaying to a hadronically decaying tau lepton and a neutrino; the phenomenology is discussed in Section ???. This search contains two subchannels, τ +jets and $\tau + \ell$ based on the decay of the associated top quark in the collision event. The τ +jets subchannel ($t \rightarrow Wb, W \rightarrow q\bar{q}$) has a higher branching fraction, leading to higher sensitivity at larger m_{H^\pm} values. The $\tau + \ell$ subchannel ($t \rightarrow Wb, W \rightarrow \ell\nu$) has a much lower branching fraction, but takes advantage of single-lepton triggers which enhance background suppression of QCD jet $\rightarrow \tau$ fakes. This leads to an increased sensitivity at lower m_{H^\pm} values. The extra neutrino in the $\tau + \ell$ decay mode creates extra difficulties in separating signal from background in this subchannel by adding a significant contribution to the E_T^{miss} calculation for the event.

The search described by this dissertation uses a profile likelihood ratio as the test statistic in a simultaneous fit in two control regions and three signal regions. The discriminating variable is chosen to be the output score distribution of a multivariate analysis technique (MVA). In the previous publication described in Section ?? several boosted decision trees (BDT) were used, binned in m_{H^\pm} ; this analysis uses a parameterized neural network (PNN) to classify events as signal-like or background-like.

This chapter discusses in detail the entire analysis, including the signal signatures, event selections, analyzed datasets, modeling of backgrounds, training and evaluation of classifiers, studies of systematic uncertainties, and results.

1.1 Signature and Event Selection

As shown in Figure ??, the production of the H^\pm is dependent on the mass m_{H^\pm} . Table 1.1 shows the production mechanisms for m_{H^\pm} values with respect to the top quark mass m_t as well as the main decay mode (and theoretical constraints), as well as the main source of background. Three mass ranges are defined, low mass $80 \leq m_{H^\pm} \leq 130$ GeV, intermediate mass $140 \leq m_{H^\pm} \leq 190$, and high mass $200 \leq m_{H^\pm} \leq 3000$ GeV. The two subchannels have similar signal signatures with a hard scatter source of E_T^{miss} , one $\tau_{had-vis}$, and at least 1 b -jet from the associated top decay. In the $\tau + \ell$ subchannel there is an extra requirement of a lepton (e or μ). Due to the variable amount of energy available to the final state products based on m_{H^\pm} the event topology changes as a function of m_{H^\pm} . As described in Section ??, classifiers are trained and evaluated in m_{H^\pm} bins to account for the varying event topology.

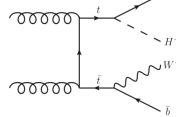
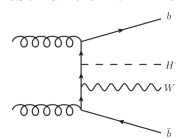
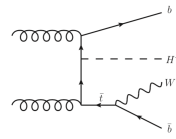
H^\pm Mass	Production Mechanism	Decay	Main Background
$m_{H^\pm} < m_t$	double-resonant $t \rightarrow H^\pm b$ (LO) 	$H^\pm \rightarrow \tau^\pm \nu_\tau$ (low $\tan \beta \Rightarrow H^\pm \rightarrow cs$ or $H^\pm \rightarrow cb$)	$t\bar{t}$, single-top
$m_{H^\pm} \simeq m_t$	non-resonant $t \rightarrow H^\pm b$ (LO) 	$H^\pm \rightarrow \tau \nu$	$t\bar{t}$, single-top
$m_{H^\pm} > m_t$	interferences taken into account single-resonant $gg \rightarrow tbH^\pm$ (NLO) 	$H^\pm \rightarrow tb$ ($\cos(\beta - \alpha) \simeq 0$ and large $\tan(\beta) \Rightarrow H^\pm \rightarrow \tau \nu$ $BR(H^\pm \rightarrow \tau^\pm \nu_\tau) \simeq 10 - 15\%$)	multi-jet

Table 1.1: H^\pm production mechanisms based on m_{H^\pm} , dominant H^\pm decay mode, and the main background associated with the diagram.

1.1.1 Object Definitions

Table 1.2 shows the identification requirements on all objects used in the analysis. In both subchannels $\tau_{had-vis}$ candidates are required to fit the medium working point described in Section ?? that corresponds to a 75% efficiency for 1-prong and 60% efficiency for 3-prong $\tau_{had-vis}$ identification, an $|\eta|$ cut of < 2.3 that also excludes the gap and crack region of the ATLAS calorimeters at $1.37 < |\eta| < 1.52$, an overlap removal with electrons is also performed. For the τ +jets subchannel, the $\tau_{had-vis}$ p_T is required to be greater than 40 GeV and greater than 30 GeV for the $\tau + \ell$ subchannel. Although muons and electrons are not part of the τ +jets signal final state, a loose identification and isolation requirement is used to veto events; while the $\tau + \ell$ subchannel requires there to be either an electron or a muon that passes the tight identification and isolation requirements as well as a p_T above 30 GeV. The jets in candidate events are required to have greater than 25 GeV in p_T and are made with the anti- k_t algorithm with $R=0.4$. Jets tagged as b -jets are done so at a 70% efficient working point using the DL1r tagger described in Section ??.

Object	τ +jets	$\tau + \ell$
$\tau_{had-vis}$	Leading reconstructed τ (regardless of its ID), mediumID*, $p_T > 40$ GeV, $ \eta ^{***} < 2.3$, e OLR	Leading reconstructed τ (regardless of its ID), mediumID*, $p_T > 30$ GeV, $ \eta ^{***} < 2.3$, e OLR
e	LoseLLH, $p_T > 20$ GeV, $ \eta ^{***} < 2.47$, Loose isolation, IP cuts	TightLLH, $p_T > 30$ GeV, $ \eta ^{***} < 2.47$, Tight isolation, IP cuts
μ	LooseID, $p_T > 20$ GeV, $ \eta < 2.5$, Loose isolation, IP cuts	TightID, $p_T > 30$ GeV, $ \eta < 2.5$, Tight isolation, IP cuts
jet	AntiKt4EMPFlow, $p_T > 25$, GeV $ \eta < 2.5$, JVT** > 0.59 , Btag=70%, DL1r	AntiKt4EMPFlow, $p_T > 25$ GeV, $ \eta < 2.5$, JVT** > 0.59 , Btag=70%, DL1r

Table 1.2: Definitions of physics objects used in this analysis.

1.1.2 Event Selections

Each subchannel signal region has stricter requirements than those described in Section 1.2. Table 1.3 details these selections. The channels differ in the triggers used; the τ +jets subchannel relies on E_T^{miss} triggers while the $\tau + \ell$ subchannel relies on single lepton triggers. Due to the difficulty of separating signal from background the τ +jets subchannel has a higher p_T cut on the $\tau_{had-vis}$ of 40 GeV as opposed to the $\tau + \ell$ value of 30 GeV. In addition, a higher value of E_T^{miss} of 150 GeV is required for the τ +jets subchannel. A value of 50 GeV is also required of the transverse mass m_T defined as

$$m_T = \sqrt{2p_T^\tau E_T^{\text{miss}}(1 - \cos\Delta\phi_{\tau, E_T^{\text{miss}}})} \quad (1.1)$$

The $\tau + \ell$ has no such requirement, but does require the $\tau_{had-vis}$ and lepton to have opposite electromagnetic charge. A set of orthogonal control regions are defined for each subchannel to verify proper background modelling and are described in Section 1.3. The acceptance of signal in the signal regions defined in 1.3 is shown in Figure 1.1.

$\tau + jets$ SR	$\tau + \ell$ SR
E_T^{miss} Trigger (mostly HLT_xe110)	Single lepton trigger (e or μ)
1 $\tau_{had-vis}$; $p_T^\tau > 40$ GeV	1 $\tau_{had-vis}$; $p_T^\tau > 30$ GeV
0 ℓ (e or μ) ; $p_T^\ell > 20$ GeV	1 ℓ (e or μ) ; $p_T^\ell > 30$ GeV
≥ 3 jets ; $p_T^j > 25$ GeV	≥ 1 jet ; $p_T^j > 25$ GeV
≥ 1 b -jets; $p_T^{b-jet} > 25$ GeV	≥ 1 b -jets; $p_T^{b-jet} > 25$ GeV
$E_T^{\text{miss}} > 150$ GeV	$E_T^{\text{miss}} > 50$ GeV
$m_T(\tau, E_T^{\text{miss}}) > 50$ GeV	Opposite sign τ and ℓ

Table 1.3: τ +jets and $\tau + \ell$ signal region definitions.

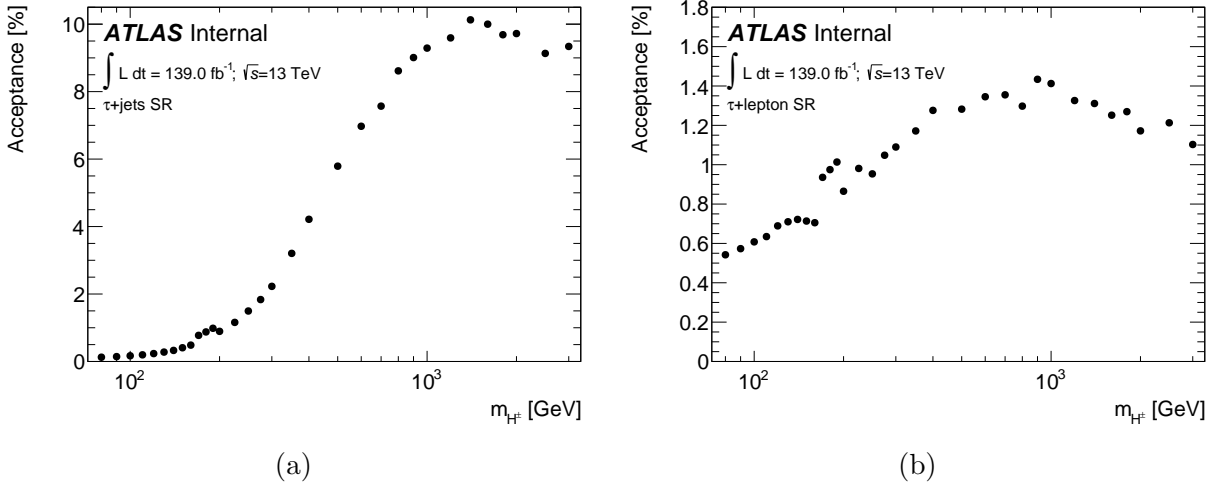


Figure 1.1: Signal acceptance as a function of the charged Higgs boson mass for both the τ +jets (a) and $\tau + \ell$ subchannels (b).

1.2 Datasets

This analysis uses the full Run-2 ATLAS dataset collected between 2015 and 2018 corresponding to $139.0 \pm 2.4 \text{ fb}^{-1}$ [1]. The datasets used are required to be included in the ATLAS “Good Run Lists” (GRLs), meaning they have passed nominal data quality checks with all detector subsystems operating within normal conditions. Further event cleaning is applied that removes events in which a reconstructed jet originated from detector noise or non-collision backgrounds. The collection of data throughout Run-2 can be seen in Figure ??.

1.2.1 Signal Modeling

Monte Carlo (MC) simulations of H^\pm signal events are generated at varying orders dependent on m_{H^\pm} . In all cases, the 2HDM Type II model described in Section ?? is assumed and the generator MadGraph is used. The lower mass range corresponding to $m_{H^\pm} < 140 \text{ GeV}$

where a H^\pm takes the place of a W^\pm in a top decay is generated at LO. The intermediate mass range of $140 \leq m_{H^\pm} < 200$ GeV is generated at LO, taking into account the non-resonant, single-top resonant and double-resonant diagrams and their interferences. In this mass range, the final state contains one H^\pm , one W^\pm , and two b quark. For charged Higgs masses of 200 GeV and above, the H^\pm is produced in association with a top quark and is generated at NLO. In all cases, the parton generator is interfaced with Pythia 8 using the A14 underlying event tuning parameters [2]. Table 1.4 shows the cross section and raw number of events generated for each m_{H^\pm} point for both subchannels.

m_{H^\pm} [GeV]	σ [pb]	$\tau + \ell$ Generated Events	τ +jets Generated Events
80	61.639	220k	110k
90	52.823	220k	110k
100	43.777	220k	110k
110	34.770	220k	110k
120	26.092	220k	110k
130	18.069	220k	110k
140	15.023	220k	220k
150	7.681	220k	220k
160	2.665	220k	220k
170	0.63748	220k	220k
180	0.52979	220k	220k
190	0.47201	220k	220k
200	0.55632	110k	220k
225	0.44081	110k	220k
250	0.3573	110k	220k
275	0.28592	110k	220k
300	0.23373	110k	220k
350	0.15774	110k	220k
400	0.10818	110k	220k
500	0.054139	110k	220k
600	0.02847	110k	220k
700	0.015764	110k	220k
800	0.009067	110k	220k
900	0.005324	110k	220k
1000	0.003271	110k	220k
1200	0.001311	110k	220k
1400	0.000558	110k	220k
1600	0.000252	110k	220k
1800	0.000120	110k	220k
2000	0.0000587	110k	220k
2500	0.0000111	110k	220k
3000	0.00000234	110k	220k

Table 1.4: For each H^\pm mass the generator xcross-section ($\sigma \times BR(H^\pm \rightarrow \tau^\pm \nu_\tau)$) is given, as well as the number of generated events for both $\tau + \ell$ and τ +jets subchannels.

1.3 Background Modeling

The main sources of backgrounds are shown in Table 1.5, separated between backgrounds with a prompt $\tau_{had-vis}$ in the hard scatter process and those that arise from the misidentification of other physics objects as a $\tau_{had-vis}$. The cross section of all simulated background samples and the relevant generators can be seen in Table 1.6. Control regions that are designed to be orthogonal to the signal region are created for both subchannels in order to study the modeling of the backgrounds. These control regions are defined by the cuts in Table 1.7 (τ +jets) and Table 1.8 ($\tau + \ell$). For the $\tau + \ell$ subchannel the Same Sign and b-veto control regions are further split into two control regions, one that requires a μ in the event and another that requires an electron.

Backgrounds w/ prompt $\tau_{had-vis}$	Backgrounds w/ fake τ
$t\bar{t}$ estimated with MC	Fake $j \rightarrow \tau$ estimated with data driven fake factor method
$W(Z) + jets$ estimated with MC	Fake $\ell \rightarrow \tau$ estimated with MC, validated on $Z \rightarrow ee$
Diboson estimated with MC	

Table 1.5: Dominant backgrounds from prompt $\tau_{had-vis}$ and fake $\tau_{had-vis}$ candidates.

As seen in Table 1.5 misidentified objects appearing as $\tau_{had-vis}$ candidates comprise a significant portion of the total background. Fakes arising from $\ell \rightarrow \tau$ misidentification are well modeled in MC simulations and are reweighted with scale factors provided by the ATLAS τ combined performance group. The mass of the $\tau_{had-vis}$ electron system can be seen in Figure 1.2 as verification of fake $\ell \rightarrow \tau$ modeling. However, fakes due to $j \rightarrow \tau$ misidentification are not well modeled in MC simulations due to a poor misunderstanding of systematic uncertainties associated with the fake $\tau_{had-vis}$ object and limited statistics of simulated events. Instead, a data driven method is used to extract a scaling constant referred to as a fake factor.

Background process	Generator & parton shower	Cross section number(s) [pb]
$t\bar{t}$ with at least one lepton ℓ	Powheg & Pythia8	729.77*
Single top-quark t -channel	Powheg & Pythia8	59.17*
Single top-quark s -channel		3.29*
Single top-quark Wt -channel		83.83
$W(\ell\nu) + \text{jets}$	Sherpa 2.2.1	2.0×10^4 2.0×10^4 2.0×10^4
$Z/\gamma^*(\ell\ell, \nu\nu) + \text{jets}$	Sherpa 2.2.1	2.1×10^3 2.1×10^3 2.1×10^3
WW	Powheg & Pythia8	54.81
WZ		16.34
ZZ		8.94

Table 1.6: Cross sections for the main SM background samples at $\sqrt{s} = 13$ TeV . Here, ℓ refers to the three lepton families e , μ and τ . All background cross sections are normalized to NNLO predictions, except for diboson events, where the NLO prediction is used. A ‘*’ indicates that the quoted cross section for the sample is neglecting leptonic/hadronic branching ratios.

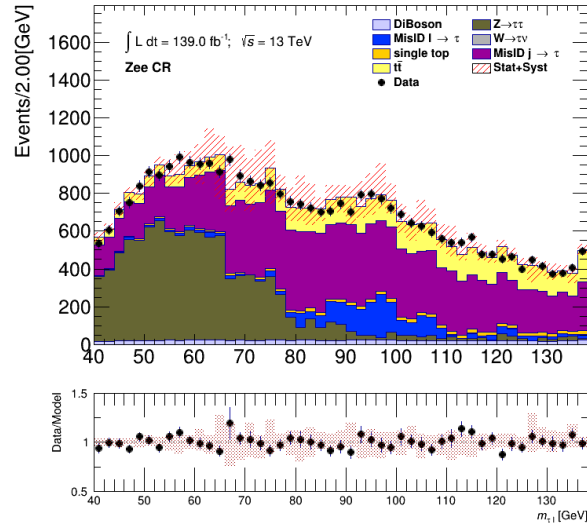


Figure 1.2: Mass of τ - e system in the Zee control region.

<div>$t\bar{t}$ Control Region</div> <div> $1 \tau_{had-vis}$ $p_T^\tau > 40 \text{ GeV}$ $\geq 3 \text{ jets}$ $\geq 2 \text{ } b\text{-jets}$ $E_T^{miss} > 150 \text{ GeV}$ $m_T(\tau, E_T^{miss}) < 100 \text{ GeV}$ </div>	<div>b-veto Control Region</div> <div> $1 \tau_{had-vis}$ $p_T^\tau > 40 \text{ GeV}$ $\geq 3 \text{ jets}$ $p_T^{jet} > 25 \text{ GeV}$ $E_T^{miss} > 150 \text{ GeV}$ $m_T(\tau, E_T^{miss}) > 50 \text{ GeV}$ $b \text{ veto}$ $\ell \text{ veto}$ </div>
(a) $t\bar{t}$ modeling	(b) Close to signal region
<div>W+Jets Control Region</div> <div> $1 \tau_{had-vis}$ $p_T^\tau > 40 \text{ GeV}$ $\geq 3 \text{ jets}$ $p_T^{jet} > 25 \text{ GeV}$ $E_T^{miss} > 150 \text{ GeV}$ $m_T(\tau, E_T^{miss}) > 100 \text{ GeV}$ $b \text{ veto}$ $\ell \text{ veto}$ </div>	<div>b-veto $m_T \geq 100$ Control Region</div> <div> $1 \tau_{had-vis}$ $p_T^\tau > 40 \text{ GeV}$ $\geq 3 \text{ jets}$ $p_T^{jet} > 25 \text{ GeV}$ $E_T^{miss} > 150 \text{ GeV}$ $m_T(\tau, E_T^{miss}) > 100 \text{ GeV}$ $b \text{ veto}$ $\ell \text{ veto}$ </div>
(c) W+Jets modeling	(d) Fake $j \rightarrow \tau$ enriched region

Table 1.7: Control region definitions for the τ +jets subchannel.

<div style="border: 1px solid black; padding: 5px; text-align: center;"> Dilepton-btag CR </div> <div style="border: 1px solid black; padding: 5px; text-align: center;"> τ veto $n \geq 1$ jets $p_T^{jet} > 25 GeV$ ≥ 1 b-jets $E_T^{miss} > 50 GeV$ $1 e$ 1μ </div>	<div style="border: 1px solid black; padding: 5px; text-align: center;"> b-veto CR </div> <div style="border: 1px solid black; padding: 5px; text-align: center;"> $1 \tau_{had-vis}$ $p_T^\tau > 30 GeV$ $1 e(\mu)$ Veto $\mu (e)$ Opposite sign $\tau e (\mu)$ ≥ 1 jets $p_T^{jet} > 25 GeV$ $E_T^{miss} > 50 GeV$ 1 tight $e (\mu)$ </div>
(a) $t\bar{t}$ and single top modeling	(b) Close to signal region
<div style="border: 1px solid black; padding: 5px; text-align: center;"> Zee CR </div> <div style="border: 1px solid black; padding: 5px; text-align: center;"> $1 \tau_{had-vis}$ $p_T^\tau > 30 GeV$ veto μ Opposite sign τe ≥ 1 jets $p_T^{jet} > 25 GeV$ b-jet veto $E_T^{miss} > 50 GeV$ $1 e$ $40 < mass(\tau, e) < 140 GeV$ </div>	<div style="border: 1px solid black; padding: 5px; text-align: center;"> Same Sign CR </div> <div style="border: 1px solid black; padding: 5px; text-align: center;"> $1 \tau_{had-vis}$ $p_T^\tau > 30 GeV$ Same sign $\tau e(\mu)$ Veto $\mu (e)$ $\geq 1 jets$ $p_T^{jet} > 25 GeV$ $E_T^{miss} > 50 GeV$ 1 tight $e (\mu)$ </div>
(c) Fake $\ell \rightarrow \tau$ enriched region	(d) Fake $j \rightarrow \tau$ enriched region

Table 1.8: Control region definitions for the $\tau + \ell$ subchannel.

In the $\tau + \ell$ final state a significant portion of $j \rightarrow \tau$ fakes come from misidentifying $\tau_{had-vis}$ candidates in W+Jets events that contain a true ℓ in the W decay and have a misidentified jet as a $\tau_{had-vis}$. Fakes of this manner also arise from QCD-like multi-jet interactions. The fake factor method used to estimate the amount of expected fake $\tau_{had-vis}$ objects that pass the $\tau_{had-vis}$ identification procedure described in Section ???. This method applies weights, or fake factors, to a subset of "anti- $\tau_{had-vis}$ " objects that have failed the selection and identification criteria in the signal region. A control region is defined to be rich in anti- $\tau_{had-vis}$ objects, where the $\tau_{had-vis}$ candidates fail the loose τ working point but have a τ identification RNN score of greater than 0.01. The fake factor and number of events with misidentified $\tau_{had-vis}$ objects (N_τ) are defined as:

$$FF = \frac{N^{\tau-id}}{N^{anti-\tau-id}} \quad (1.2)$$

$$N_{fakes}^\tau = N_{fakes}^{anti-\tau} \times FF$$

Both of these values are then corrected for $\tau_{had-vis}$ candidates matching a true hadronic τ at generator level:

$$N^{\tau-id} = N^{\tau-id}(Data) - N^{\tau-id}(MC) \quad (1.3)$$

$$N_{fakes}^{anti-\tau} = N_{fakes}^{anti-\tau}(Data) - N_{true}^{anti-\tau}(MC)$$

Two control regions are created, one to capture the multi-jet (MJ) fakes and the other to study the W+jets fakes. The MJ control region uses the τ +jets signal region definition with an additional b-veto and an $E_T^{\text{miss}} < 80$ GeV cut. The W+jets control region uses the $\tau + \ell$ signal region definition with a b-veto, no E_T^{miss} cut, and a cut on the transverse mass of the ℓ - E_T^{miss} system of $60 < m_T(\ell, E_T^{\text{miss}}) < 160$ GeV. The FF in the signal region is defined as

$$FF_{sig} = \alpha_{MJ} \times FF_{MJ} + (1 - \alpha_{MJ}) \times FF_{W+jets} \quad (1.4)$$

where α is taken from a template fit of the τ -ID score distributions of the anti- τ s using template shapes from the anti- τ distributions in the MJ and W+jets control regions. In the signal regions, the number of events containing fake- $\tau_{had-vis}$ candidates is defined as

$$N_{fake-\tau} = FF_{sig} \times N_{anti-\tau} \quad (1.5)$$

Figure 1.3 shows FF plotted in each control region for 1-prong and 3-prong $\tau_{had-vis}$ binned in $p_T\tau$; extracted α values and their fits can be seen in Appendix B.

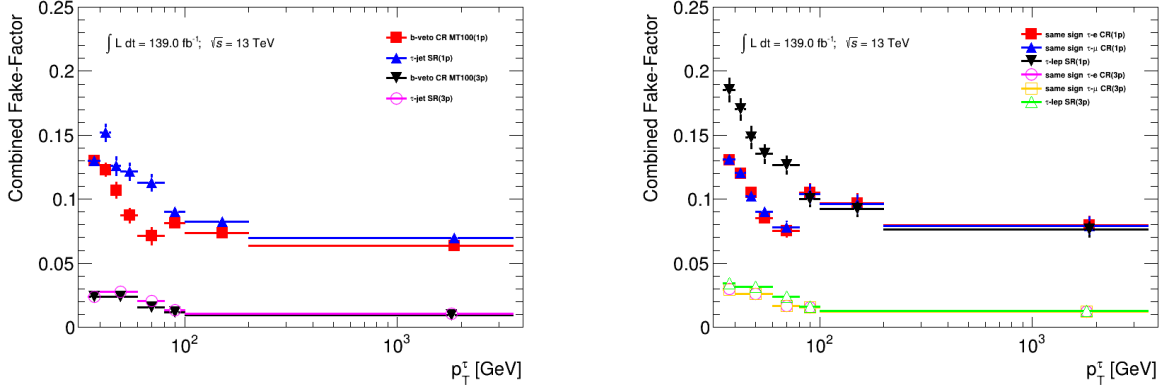


Figure 1.3: Combined fake factors for the τ +jets b-veto $m_T > 100$ control region, τ +jets signal region, τ +electron(muon) with same-sign control region and the $\tau + \ell$ signal region. Error bars represent systematic uncertainties of the method.

To verify background modeling, the E_T^{miss} distributions in each of the control regions are plotted with final scale factors including fake factors in Figure 1.4 (τ +jets) and Figure 1.5 ($\tau + \ell$). These plots include a ratio of reconstructed data events and simulated MC events bin by bin to ensure proper modeling across variable shapes. More background modeling plots can be seen in Appendix C.

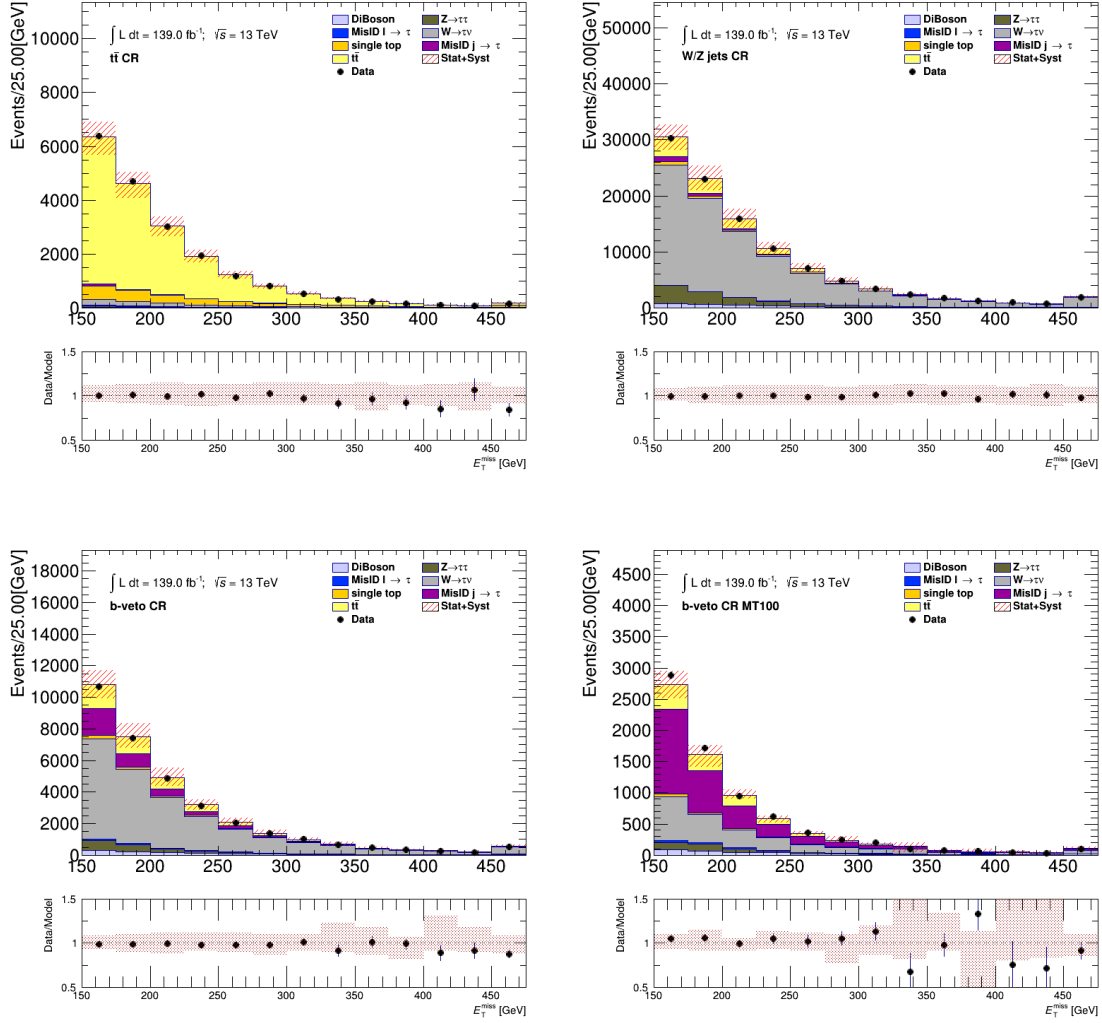


Figure 1.4: Comparison between the predicted and the measured E_T^{miss} distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

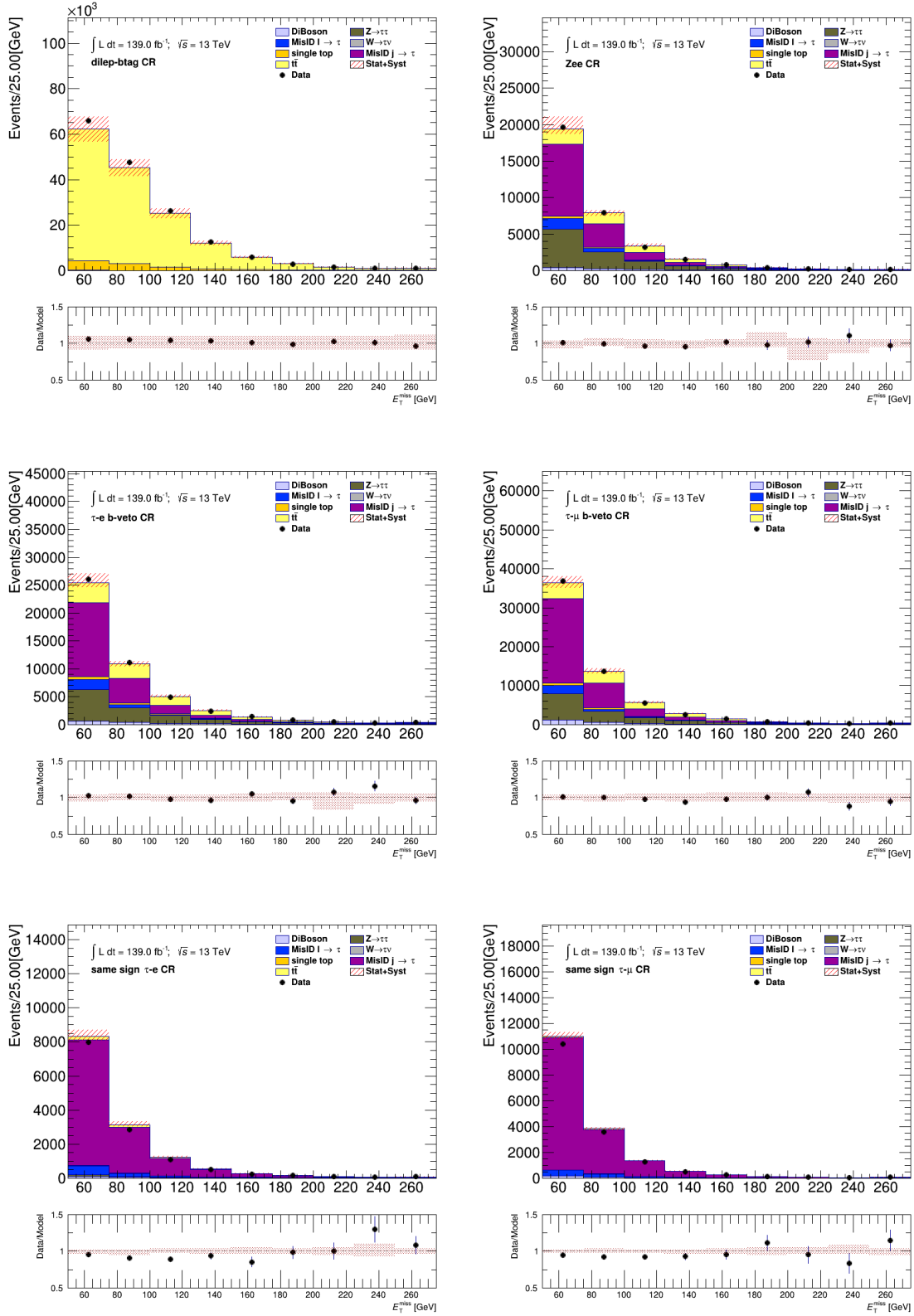


Figure 1.5: Comparison between the predicted and the measured E_T^{miss} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

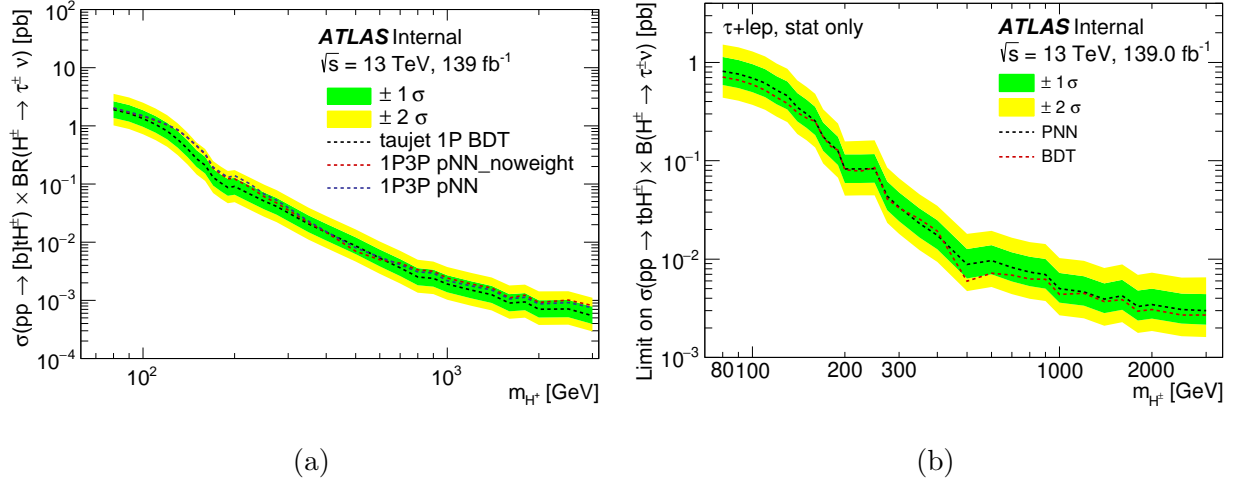


Figure 1.6: Comparison of performance of an optimized BDT and an unoptimized PNN on expected limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ in the τ +jets (a) and $\tau + \ell$ (b) signal regions.

1.4 Multivariate Analysis Techniques

Once variables distributions are properly scaled and data/MC agreement is verified, multivariate analysis techniques are employed to separate signal-like events from background-like events in the signal regions. In the previous publication (described in Section ??), BDTs binned in m_{H^\pm} were used as the classifier, this publication use one PNN for the entire m_{H^\pm} spectrum. BDTs excel at separating linear correlations, whereas neural networks take advantage of nonlinear correlations. In the case of a PNN the parameterized variable, here m_{H^\pm} , is taken as an input to the network in addition with other input variables. PNNs offer the advantage of having one classifier model that can evaluate at any m_{H^\pm} value by learning how the signal event topology changes as m_{H^\pm} varies [3]. For illustrative purposes, expected limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ in both subchannels is shown comparing an optimized BDT and an unoptimized PNN in Figure 1.6. It is seen that the PNN performs similarly to the BDTs used in the previous analysis. A PNN was chosen as the discriminator.

A neural network (NN) is a computing system loosely inspired by the human brain. NNs combine adaptive nonlinear basis functions in an attempt to perform a task, classification in the context of this dissertation. A NN contains layers of nodes connected to each other with an associated weight and threshold. As long as a node has output greater than the given threshold value, data will flow through that node. Otherwise, that node is not activated and data are not sent to the next layer. The NN as a whole relies on a process called training where the node weights are varied, an accuracy is calculated based on a given loss function, the weights are then varied again and the process repeats. This is done until a preferred accuracy is reached; the final node weights are saved and new data can be evaluated. A diagram of a PNN can be seen in Figure 1.7, where the parameterized input is labeled as θ . The learned function of a NN can be written as:

$$y(x) = w_0^2 + \sum_{m=1}^M [w_m^2 \cdot h(w_{0m}^1 + \sum_{k=1}^D w_{km}^1 x_k)] \quad (1.6)$$

where w is the neuron weights, M is the number of basis functions being combined, D is the number of inputs and h is the activation function.

This analysis uses four PNNs, events with 1-prong τ and 3-prong τ are divided into separate datasets within both subchannels.

1.4.1 Training

The training of the PNNs used in this dissertation are done with the Keras [4] library using the TensorFlow [5] library as backend. In order to increase the significance of training statistics and protect from overtraining, the k -fold method is used. Overtraining occurs when a NN has been fine tuned to have a high accuracy with a specific dataset and does not generalize to other datasets. To protect against this, dropout is used [6]. The k -fold method

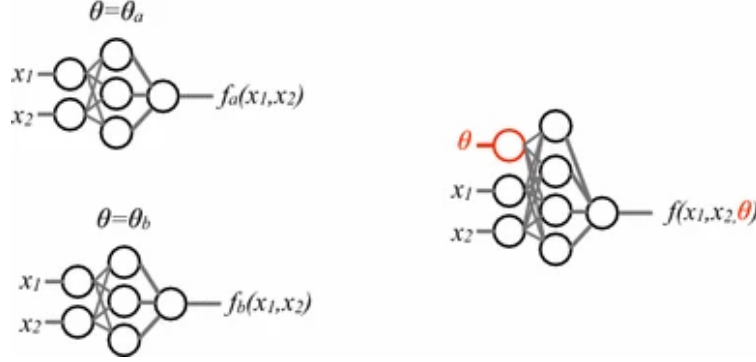


Figure 1.7: *Left*, individual networks with input variables (x_1, x_2) , each trained with examples with a single value of some parameter $\theta = \theta_a, \theta_b$. The individual networks are purely functions of the input variables. Performance for intermediate values of θ is not optimal nor does it necessarily vary smoothly between the networks. *Right*, a single network trained with input variables (x_1, x_2) as well as input parameter θ ; such a network is trained with examples at several values of the parameter θ [3].

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Background
Partition 1	Evaluation	Train	Train	Train	Train	Fold 1
Partition 2	Train	Evaluation	Train	Train	Train	Fold 2
Partition 3	Train	Train	Evaluation	Train	Train	Fold 3
Partition 4	Train	Train	Train	Evaluation	Train	Fold 4
Partition 5	Train	Train	Train	Train	Evaluation	Fold 5

Figure 1.8: The k-fold method for $k = 5$ [7].

divides input training samples into k equally populated subsets. The k -th subset is trained on the other $k - 1$ subsets and evaluated on the k -th subset. Figure 1.8 shows a pictorial representation of the k -fold method. $k = 5$ is chosen in this analysis.

A single PNN training is performed on all m_{H^\pm} values at once, with the m_{H^\pm} value being taken as an input variable. For signal events, the m_{H^\pm} value from the MC generator is given; background events are replicated 32 times (the number of simulated m_{H^\pm} points is 32) and each m_{H^\pm} value is given for each set. To avoid biasing the training due to varying statistics at each m_{H^\pm} value, the background events are weighted by a factor of $w = N_S^i / N_B^i$ where i corresponds to a given m_{H^\pm} value, N_S^i , and N_B^i are the number of signal and background

events respectively. When the PNN is evaluated, the m_{H^\pm} value is chosen and the output is used as the discriminant at that m_{H^\pm} .

1.4.2 Input Variables Selection

The choice of input variables to the PNNs is critical to the performance of the analysis. Several sets of variables were compared using expected limits as the figure of merit. All studies were performed in the $\tau + \ell$ signal region, as this region proves the most difficult challenge to separate signal-like events from background-like events. So called low level variables, consisting of the four vector components of the main physics objects in each event, were compared against high level variables; high level meaning they are calculated quantities from the low level variables. Tables of the low level and high level variables are shown in Table 1.9.

The variable $m_{H^\pm}^{Truth}$ corresponds to the m_{H^\pm} value the training and evaluation is performed at. In both cases, the variable Υ is used. Υ is a measure of the $\tau_{had-vis}$ polarization, computed by taking the asymmetry of energies carried by the charged and neutron pions from the 1-prong τ decay measured in the laboratory frame. Υ is defined as

$$\Upsilon = \frac{E_T^{\pi^\pm} - E_T^{\pi^0}}{E_T^\tau} \approx 2 \frac{p_T^{\tau-track}}{p_T^\tau} - 1 \quad (1.7)$$

where $p_T^{\tau-track}$ is the transverse momentum of the track associated with the 1-prong $\tau_{had-vis}$ candidate. As such, Υ is only defined for 1-prong $\tau_{had-vis}$ candidates. As demonstrated in the previous analysis, Υ provides a large contribution to signal-backgrounds separation at charged Higgs masses below 400 GeV [8].

An estimate of the impact of low level vs high level variables on the expected limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ is shown in 1.9. The set of low level variables was chosen

Low Level Input Variables

p_{T}^{τ}	η^{τ}	ϕ^{τ}	E^{τ}
p_{T}^{ℓ}	η^{ℓ}	ϕ^{ℓ}	E^{ℓ}
p_{T}^{b-jet}	η^{b-jet}	ϕ^{b-jet}	E^{b-jet}
p_{T}^{jet}	η^{jet}	ϕ^{jet}	E^{jet}
$E_{\text{T}}^{\text{miss}}$	$\phi_{\text{T}}^{E^{\text{miss}}}$	$p_{\text{T}}^{j_1}$	Υ
$m_{H^{\pm}}^{\text{Truth}}$			

(a) Low level variables

High Level Input Variables

$E_{\text{T}}^{\text{miss}}$
p_{T}^{τ}
p_{T}^{b-jet}
p_{T}^{ℓ}
$\Delta\phi_{\tau, \text{miss}}$
$\Delta\phi_{b-jet, \text{miss}}$
$\Delta\phi_{\ell, \text{miss}}$
$\Delta R_{\tau, \ell}$
$\Delta R_{b-jet, \ell}$
$\Delta R_{b-jet, \tau}$
$\Delta\phi_{\tau, \text{miss}} / \Delta\phi_{jet, \text{miss}}$
Υ
$m_{H^{\pm}}^{\text{Truth}}$

(b) High level variables

Table 1.9: List of low level (a) and high level (b) kinematic variables used as input to the PNN in the $\tau + \ell$ subchannel. $\Delta\phi_{X, \text{miss}}$ denotes the difference in azimuthal angle between a reconstructed object X ($X = \tau, b-jet, \ell$) and the direction of the missing transverse momentum.

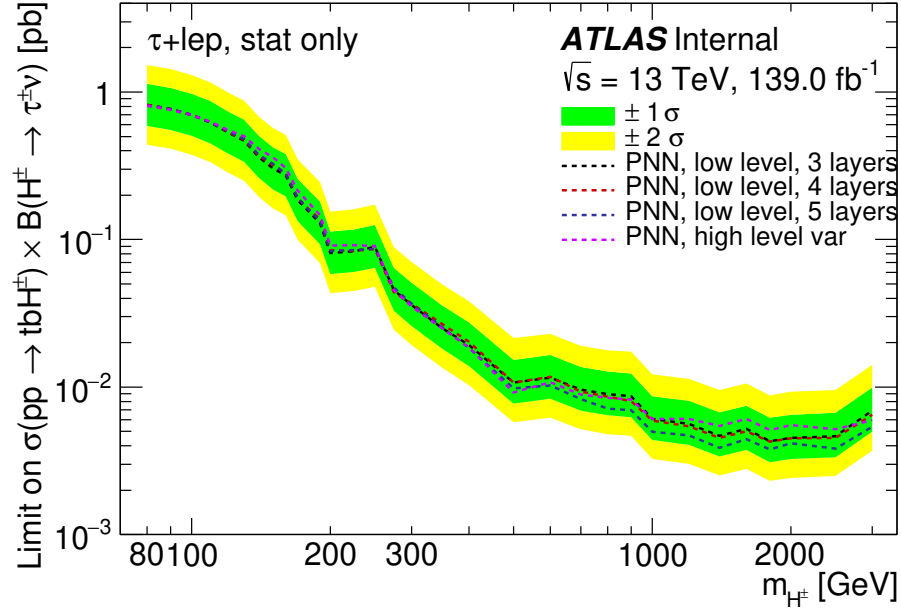


Figure 1.9: Expected limits comparing a set of high level variables and low level variables with various depths in the PNN architecture. X layers refers to the number of layers in the PNN.

as performance was similar in low m_{H^\pm} and greater in high m_{H^\pm} . A optimization of the number of layers in the PNN and several other parameters of the PNN is discussed in detail in Section 1.4.3.

1.4.3 Hyperparameter Optimization

In order to optimize the PNN, a scan of hyperparameters and network architecture was done, referred to as hyperparameter optimization (HPO). A calculated area under the curve, (AUC), was used as the figure of merit. As in the normal training scheme, the k -fold method with $k = 5$ was used to keep background modelling and classifier training statistically independent. To prevent overtraining, the early stopping method was used with

$\Delta_{min} = 0.00001$ and a patience of 10, with the best weights kept to calculate the AUC. To optimize for PNN performance in low H^\pm mass points, a separate average taking into account only H^\pm mass values between 80 and 500 GeV was used as the final figure of merit. In an effort to keep the computational needs low, several small grids of hyperparameters and architecture structures were scanned. Tables 1.10 - 1.14 show the hyperparameter grids that were searched. Here, width refers to the number of neurons per layer and depth is the number of layers. The final hyperparameter from each grid search is highlighted in red. The results of the final grid search can be seen in Tables 1.15 and 1.16; the quoted errors are taken from the different k -folds. The AUC values for each m_{H^\pm} point for the final chosen model are shown in Figure 1.10.

Parameter			
activation function	softsign	relu	LeakyReLU
loss function	binary crossentropy	mean squared error	mean absolute error
width	32		
depth	10		
batch size	1025		

Table 1.10: First grid, scanning over activation function and loss function. Binary crossentropy was the chosen loss function, highlighted in red.

Parameter			
width	8	16	32
depth	3	5	10
dropout	0.1	0.3	
activation function	softsign		
loss function	binary crossentropy		
batch size	1024		

Table 1.11: Second grid, scanning over width, depth, and dropout value. 0.1 was chosen for the dropout value, highlighted in red.

Parameter			
width	32	64	128
depth	2	3	4
activation function	softsign	relu	LeakyReLU
dropout	0.1		
batch size	1024		
loss function	binary crossentropy		

Table 1.12: Third grid, scanning over activation function. LeakyReLU was chosen, highlighted in red.

Parameter				
width	32	64	128	
depth	2	3	4	
α	0.01	0.05	0.001	0.005
batch size	1024			
dropout	0.1			
activation function	LeakyReLU			
loss function	binary crossentropy			

Table 1.13: Fourth grid, scanning over LeakyReLU α value. $\alpha = 0.05$ was chosen, highlighted in red.

Parameter				
width	32	64	128	256
depth	2	3	4	5
batch size	1024			
dropout	0.1			
activation function	LeakyReLU			
batch size	1024			
α	0.05			
loss function	binary crossentropy			

Table 1.14: Fourth grid, scanning over network width and depth. $width = 128$ and $depth = 3$ was chosen, highlighted in red.

width	depth	80	150	250	500	Avg	LowMassAvg
128	3	0.666137 \pm 0.000000	0.814508 \pm 0.000000	0.903123 \pm 0.000000	0.963256 \pm 0.000000	0.887638 \pm 0.000000	0.826145 \pm 0.096754
128	5	0.649154 \pm 0.000000	0.804344 \pm 0.000000	0.907763 \pm 0.000000	0.962846 \pm 0.000000	0.886060 \pm 0.000000	0.823542 \pm 0.100037
128	4	0.659330 \pm 0.000000	0.811707 \pm 0.000000	0.901186 \pm 0.000000	0.963811 \pm 0.000000	0.885833 \pm 0.000000	0.823208 \pm 0.099379
128	2	0.644392 \pm 0.000000	0.807016 \pm 0.000000	0.907517 \pm 0.000000	0.963076 \pm 0.000000	0.885685 \pm 0.000000	0.823139 \pm 0.100649
64	4	0.657593 \pm 0.005023	0.807977 \pm 0.001327	0.905193 \pm 0.004490	0.965553 \pm 0.001622	0.885708 \pm 0.000177	0.823001 \pm 0.099420
64	2	0.652767 \pm 0.006639	0.805184 \pm 0.002345	0.905695 \pm 0.003172	0.965077 \pm 0.000726	0.885537 \pm 0.000443	0.822775 \pm 0.099628
64	5	0.653787 \pm 0.005006	0.804417 \pm 0.001933	0.905833 \pm 0.003671	0.965293 \pm 0.001398	0.885338 \pm 0.000545	0.822360 \pm 0.099660
64	3	0.652007 \pm 0.006721	0.805076 \pm 0.001760	0.904237 \pm 0.004398	0.964922 \pm 0.001898	0.885317 \pm 0.001074	0.822335 \pm 0.099360
256	5	0.653576 \pm 0.000963	0.804396 \pm 0.003342	0.903638 \pm 0.004193	0.964415 \pm 0.002172	0.884405 \pm 0.000175	0.821347 \pm 0.100307
256	4	0.643401 \pm 0.000000	0.801775 \pm 0.000000	0.901747 \pm 0.000000	0.961914 \pm 0.000000	0.882293 \pm 0.000000	0.818097 \pm 0.101322
32	3	0.636902 \pm 0.009356	0.794963 \pm 0.004126	0.897744 \pm 0.003173	0.963498 \pm 0.002178	0.879826 \pm 0.001226	0.813868 \pm 0.103095
32	4	0.638362 \pm 0.003653	0.793516 \pm 0.003269	0.898635 \pm 0.003664	0.963582 \pm 0.001635	0.879864 \pm 0.000928	0.813853 \pm 0.103071
32	2	0.639871 \pm 0.005791	0.792428 \pm 0.002366	0.898305 \pm 0.003283	0.962854 \pm 0.002313	0.879603 \pm 0.000405	0.813528 \pm 0.102342
32	5	0.634979 \pm 0.007666	0.793076 \pm 0.005599	0.898086 \pm 0.002238	0.962539 \pm 0.000514	0.879239 \pm 0.001076	0.812845 \pm 0.103520
256	2	0.632035 \pm 0.004384	0.797129 \pm 0.000014	0.893944 \pm 0.003417	0.958731 \pm 0.001777	0.878091 \pm 0.000165	0.811994 \pm 0.102326

Table 1.15: AUCs of final HPO grid

width	depth	Avg	LowMassAvg
128	3	0.887638 \pm 0.000000	0.826145 \pm 0.096754
128	5	0.886060 \pm 0.000000	0.823542 \pm 0.100037
128	4	0.885833 \pm 0.000000	0.823208 \pm 0.099379
128	2	0.885685 \pm 0.000000	0.823139 \pm 0.100649
64	4	0.885708 \pm 0.000177	0.823001 \pm 0.099420
64	2	0.885537 \pm 0.000443	0.822775 \pm 0.099628
64	5	0.885338 \pm 0.000545	0.822360 \pm 0.099660
64	3	0.885317 \pm 0.001074	0.822335 \pm 0.099360
256	5	0.884405 \pm 0.000175	0.821347 \pm 0.100307
256	4	0.882293 \pm 0.000000	0.818097 \pm 0.101322
32	3	0.879826 \pm 0.001226	0.813868 \pm 0.103095
32	4	0.879864 \pm 0.000928	0.813853 \pm 0.103071
32	2	0.879603 \pm 0.000405	0.813528 \pm 0.102342
32	5	0.879239 \pm 0.001076	0.812845 \pm 0.103520
256	2	0.878091 \pm 0.000165	0.811994 \pm 0.102326

Table 1.16: Average AUCs of final HPO grid

Figure 1.10: Final model AUC for each mass point. Individual points correspond to the AUC average over 5 kfold.

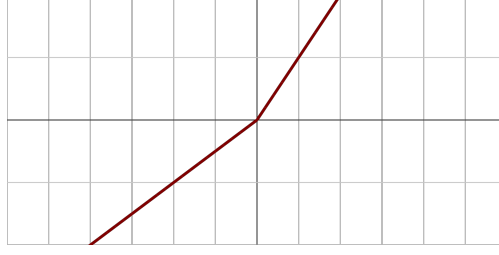


Figure 1.11: LeakyReLU activation function. The associated hyperparameter α is the slope of the negative portion of the function.

1.5 Systematic Uncertainties

Systematic uncertainties have a variety of sources and are discussed in detail here. Detector-related systematic uncertainties from the reconstruction and identification of leptons and $\tau_{had-vis}$ objects, simulation of the electron and muon triggers, reconstruction of E_T^{miss} , and energy/momentum scale and resolution of all physics objects are studied by varying selection cuts by ± 1 standard deviation. The difference in event yields is then taken as a systematic error and summed in quadrature with all other sources of error to give the final quoted errors. Systematic uncertainties arising from the data-driven fake factor method are shown in Table 1.17.

Source of uncertainty	$\tau + \text{jets}$		$\tau + \ell$	
	Effect on yield	Shape	Effect on yield	Shape
Fake factors: statistical uncertainties	3.9%	✓	3.2%	✓
Fake factors: True $\tau_{had-vis}$ in the anti- $\tau_{had-vis}$ CR	+3.4%	✓	+4%	✓
	-3.2%		-4.3%	✓
Fake factors: tau RNN Identification SF	2.7%	✓	2.7%	✓
Fake factors: α_{MJ} uncertainty	3.6%	✓	1.9%	✓
Fake factors: Smirnov transform	0%	✓	0%	✓
Fake factors: heavy flavor jet fraction	5%	✓	5%	✓

Table 1.17: Effect on the shape variation and the yields of systematic uncertainties associated with the data-driven fake factor method, used to estimate the $j \rightarrow \tau$ background in the $\tau + \text{jets}$ and $\tau + \ell$ channel.

Table 1.18: Expected event yields for the backgrounds and a hypothetical H^\pm signal after applying all τ +jets selection criteria, and comparison with 139 fb^{-1} of data. All yields are evaluated prior to using the multivariate discriminant and applying the statistical fitting procedure. The values shown for the signal assume a charged Higgs boson mass of 170 GeV and 1000 GeV, with a cross-section times branching fraction $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ corresponding to $\tan\beta = 40$ in the hMSSM benchmark scenario. Statistical uncertainties are quoted. **Systematic uncertainties to be added.**

Sample	Event yields $\tau_{had-vis} + \text{jets}$			
True τ_{had}				
$t\bar{t}$	18369.33	\pm 48.16	\pm XXX	
Single-top-quark	2276.08	\pm 16.69	\pm XXX	
$W \rightarrow \tau\nu$	1972.76	\pm 23.54	\pm XXX	
$Z \rightarrow \tau\tau$	241.05	\pm 5.47	\pm XXX	
Diboson (WW, WZ, ZZ)	133.30	\pm 4.67	\pm XXX	
Misidentified $e, \mu \rightarrow \tau_{had-vis}$	327.51	\pm 6.82	\pm XXX	
Misidentified jet $\rightarrow \tau_{had-vis}$	2490.58	\pm 17.35	\pm XXX	
All backgrounds	25810.61	\pm 59.60	\pm XXX	
H^\pm (170 GeV), hMSSM $\tan\beta = 40$	4330.22	\pm 36.75	\pm XXX	
H^\pm (1000 GeV), hMSSM $\tan\beta = 40$	31.15	\pm 0.14	\pm XXX	
Data	XXX			

Theoretical uncertainties for signal and $t\bar{t}$ background were considered in the last publication; at the time of writing this dissertation the simulations are being produced and therefore are not included.

Systematic % difference wrt nominal is being finalized. Add in when done.

1.6 Results

The expected event yields for backgrounds and signal¹ are summarized in Table 1.18 (τ +jets) and Table 1.19 ($\tau + \ell$).

The test statistic \tilde{q}_μ [9] is used to test the agreement of the data with the background-only and signal+background hypotheses. The test statistic is based on a profile likelihood

¹At the time of writing, the analysis is still blinded so Data is not included.

Table 1.19: Expected event yields for the backgrounds and a hypothetical H^\pm signal after applying all $\tau + \ell$ selection criteria, and comparison with 139fb^{-1} of data. All yields are evaluated prior to using the multi-variate discriminant and applying the statistical fitting procedure. The values shown for the signal assume a charged Higgs boson mass of 170 GeV and 1000 GeV, with a cross-section times branching fraction $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ corresponding to $\tan\beta = 40$ in the hMSSM benchmark scenario. Statistical uncertainties are quoted. **Systematic uncertainties to be added.**

Sample	Event yields $\tau+e$			Event yields $\tau+\mu$		
True τ_{had}						
$t\bar{t}$	43814.66	± 76.84	$\pm XXX$	44490.69	± 75.96	$\pm XXX$
Single-top-quark	3260.70	± 20.81	$\pm XXX$	3874.57	± 22.06	$\pm XXX$
$W \rightarrow \tau\nu$	2.41	± 0.56	$\pm XXX$	0.07	± 0.12	$\pm XXX$
$Z \rightarrow \tau\tau$	913.61	± 20.42	$\pm XXX$	845.89	± 22.07	$\pm XXX$
Diboson (WW, WZ, ZZ)	73.21	± 1.53	$\pm XXX$	81.32	± 1.53	$\pm XXX$
Misidentified $e, \mu \rightarrow \tau_{\text{had-vis}}$	1096.64	± 24.36	$\pm XXX$	1074.28	± 15.90	$\pm XXX$
Misidentified jet $\rightarrow \tau_{\text{had-vis}}$	8773.81	± 37.64	$\pm XXX$	8558.21	± 37.23	$\pm XXX$
All backgrounds	57935.04	± 93.63	$\pm XXX$	58925.03	± 91.57	$\pm XXX$
H^\pm (170 GeV), hMSSM $\tan\beta = 40$	2543.83	± 28.12	$\pm XXX$	2877.49	± 28.41	$\pm XXX$
H^\pm (1000 GeV), hMSSM $\tan\beta = 40$	2.34	± 0.03	$\pm XXX$	2.46	± 0.03	$\pm XXX$
Data	XXX			XXX		

ratio where the binned likelihood function $\mathcal{L}(\mu, \theta)$ is constructed as the product of Poisson probability terms over all bins and regions. The likelihood ratio is the ratio between the conditional maximum-likelihood estimator of the nuisance parameters, θ , for a given signal hypothesis μ and the unconditional maximum-likelihood estimator for μ and the nuisance parameters. \tilde{q}_μ is defined as:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(0, \hat{\theta}(0))}, & \hat{\mu} < 0 \\ -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \quad (1.8)$$

The fit is performed on the PNN score distributions in the three signal regions, $\tau+\text{jets}$, $\tau+e$, $\tau+\mu$, and the dilepton-btag control region which is enriched in the dominant $t\bar{t}$ background. Pre-fit PNN score distributions are shown in Figures 1.12, 1.13, and 1.14. At the time of writing this dissertation, the analysis is still blinded. Assuming the fit agrees

with the background-only hypothesis expected limits of $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ are calculated. Exclusion limits are set at the 95% confidence level (CL) using the CL_s procedure [10]. The expected exclusion limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau\nu)$ can be seen in Figure 1.15.

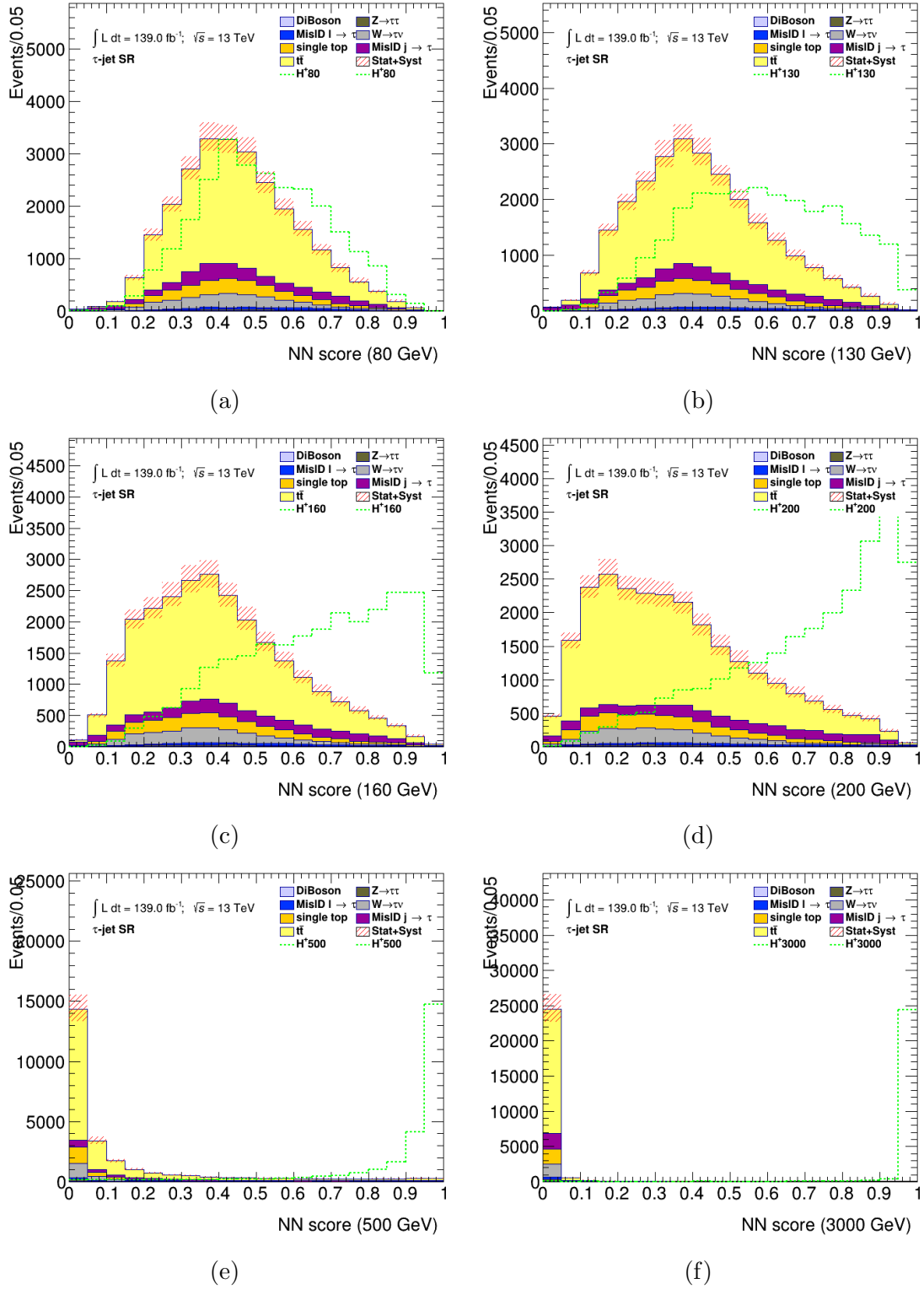


Figure 1.12: PNN score distributions in the signal region of the τ +jets channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.

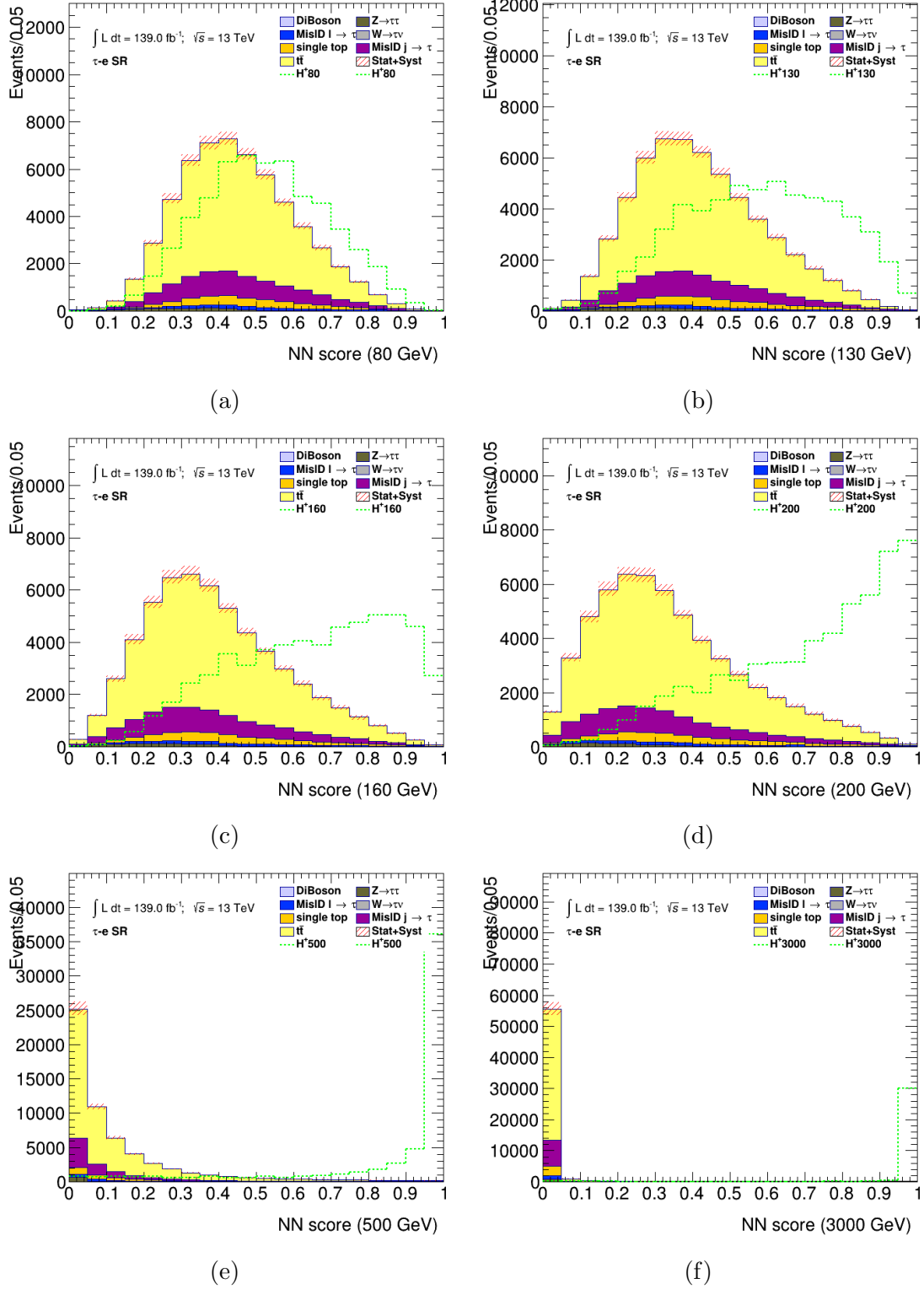


Figure 1.13: PNN score distributions in the signal region of the $\tau+e$ sub-channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.

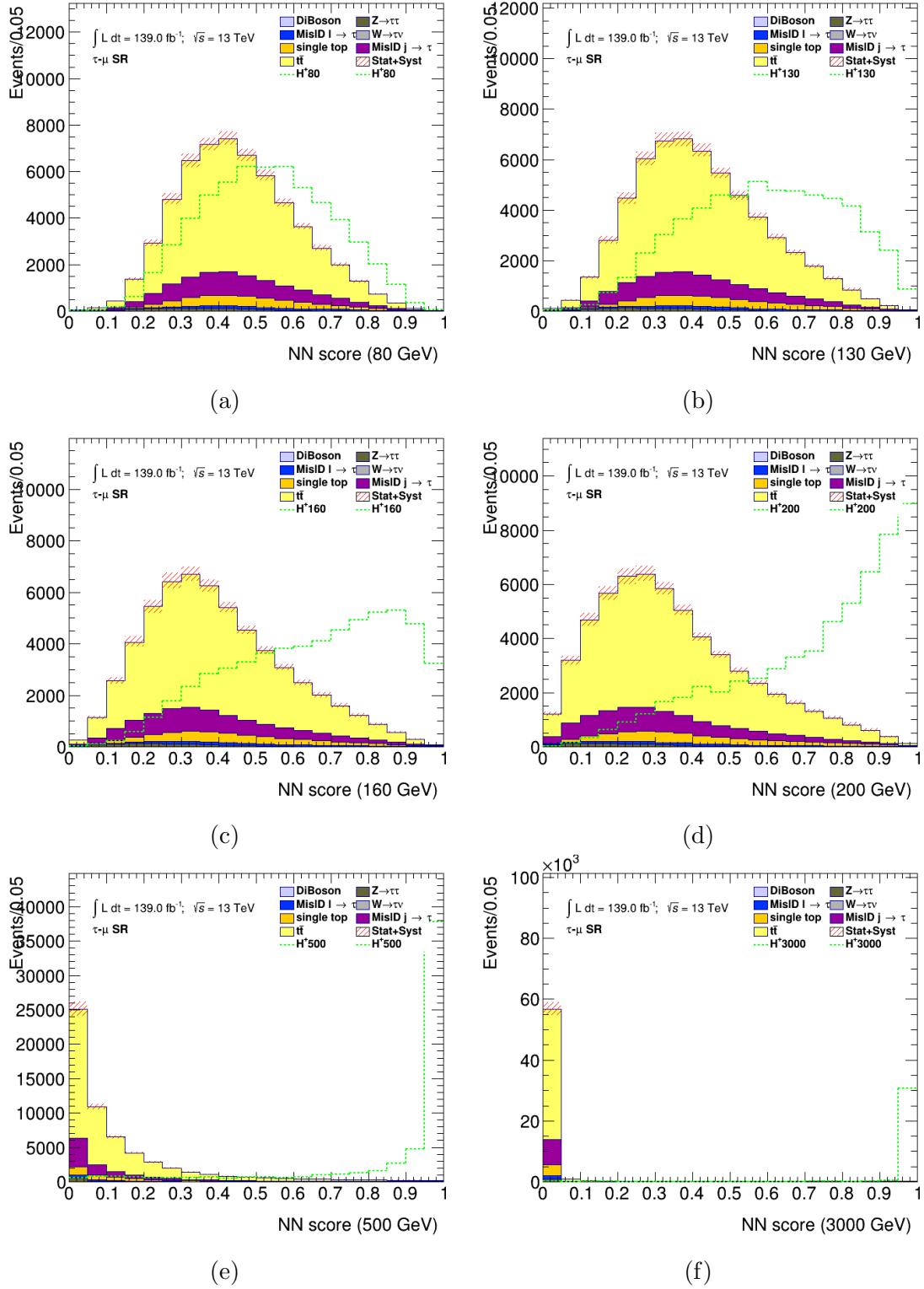
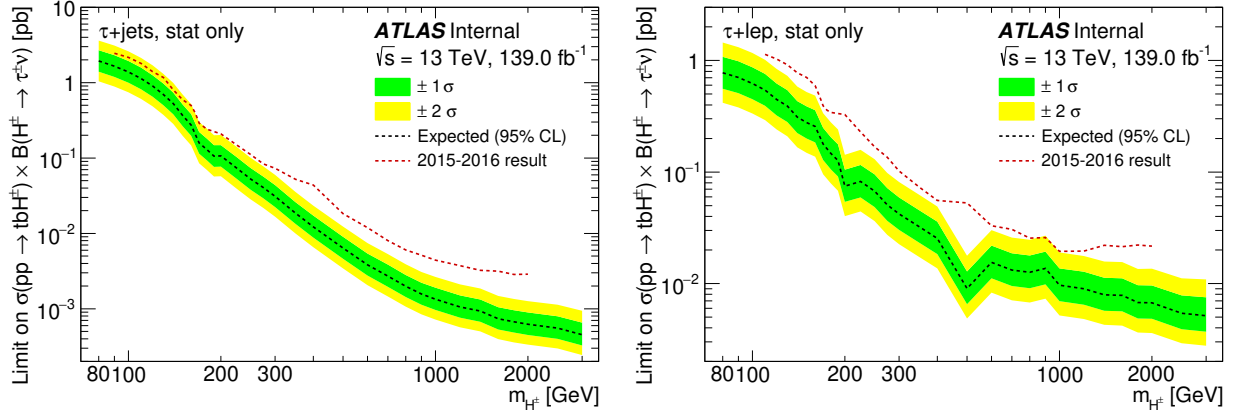
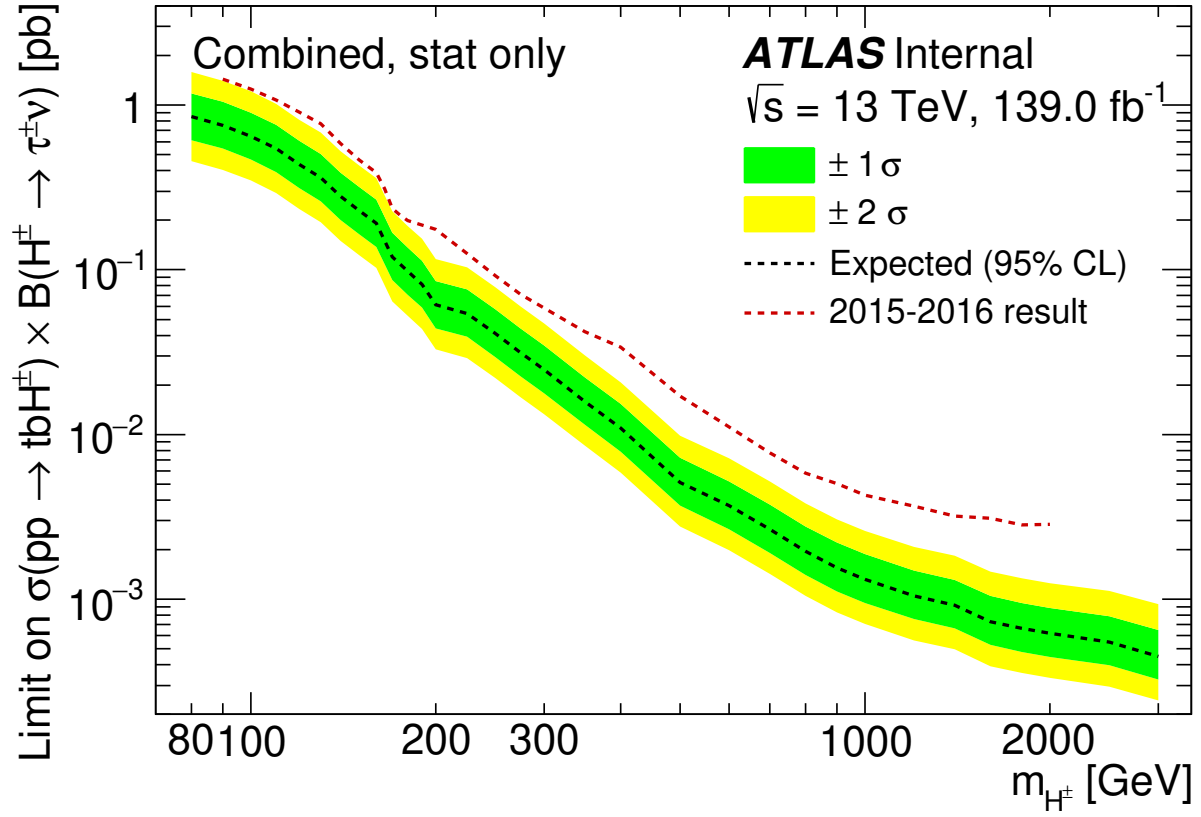


Figure 1.14: PNN score distributions in the signal region of the $\tau+\mu$ sub-channel, for the six charged Higgs boson mass parameters. The lower panel of each plot shows the ratio of data to the SM background prediction. The uncertainty bands include all statistical and systematic uncertainties. The normalization of the signal (shown for illustration) corresponds to the integral of the background.



(a)

(b)



(c)

Figure 1.15: Expected 95% CL exclusion limits on $\sigma(pp \rightarrow tbH^\pm) \times \mathcal{B}(H^\pm \rightarrow \tau^\pm \nu)$ as a function of the charged Higgs boson mass in 139 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV in the τ +jets signal region (a), $\tau + \ell$ signal region (b), and the combination of the τ +jets and $\tau + \ell$ signal regions (c). In the case of the expected limits, one- and two-standard-deviation uncertainty bands are also shown. As a comparison, the expected exclusion limits obtained with the dataset collected in 2015 and 2016 [8] are also shown.

Appendices

APPENDIX A
TILECAL DATA QUALITY

APPENDIX B

FAKE FACTORS

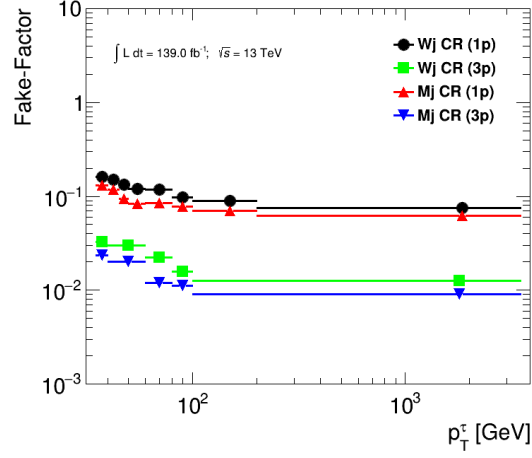


Figure B.1: Fake factors parameterized as a function of p_T^τ and the number of charged τ decay products (1-prong and 3-prong) obtained in the multi-jet and W+jets CRs. The errors shown represent the statistical uncertainty.

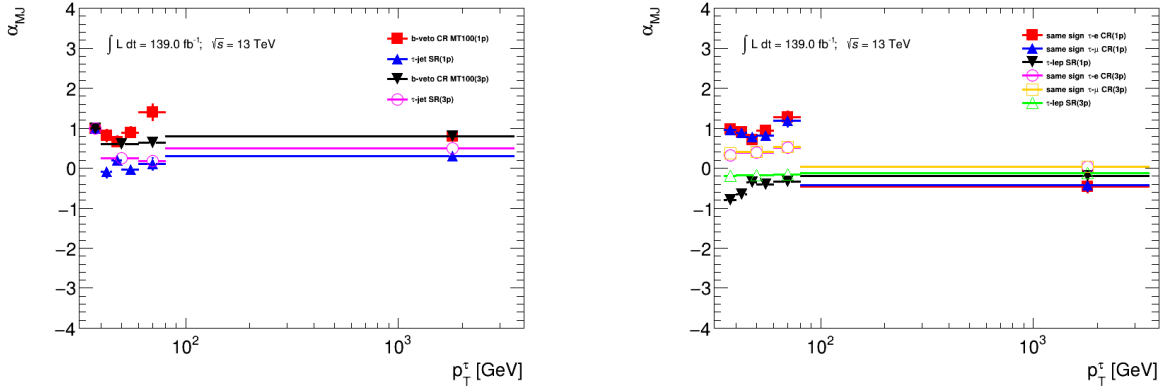


Figure B.2: Corrected α_{MJ} values for the $\tau_{had-vis}$ +jets b-veto $m_T > 100$ control region, $\tau_{had-vis}$ +jets signal region, $\tau_{had-vis}$ +electron(muon) with same-sign control region and the $\tau_{had-vis}$ +lepton signal region. Error bars represent uncertainties due to α_{MJ} fitting using template-fit method.

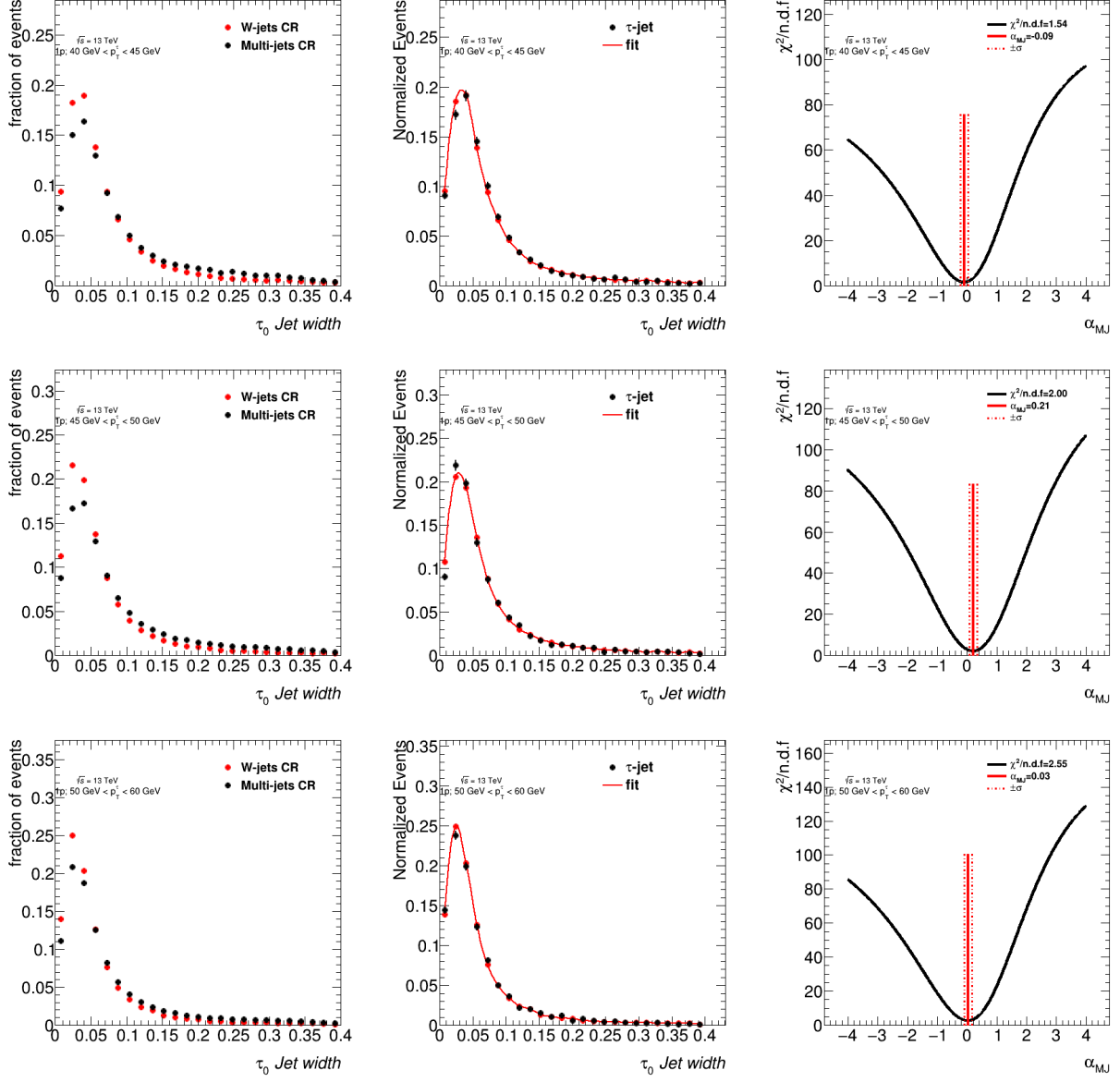


Figure B.3: Estimation of α_{MJ} in the $\tau_{had-vis} + jets$ signal region for $p_T \leq 60$ GeV 1-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{min}/ndf + \sqrt{\frac{2}{ndf}}$.

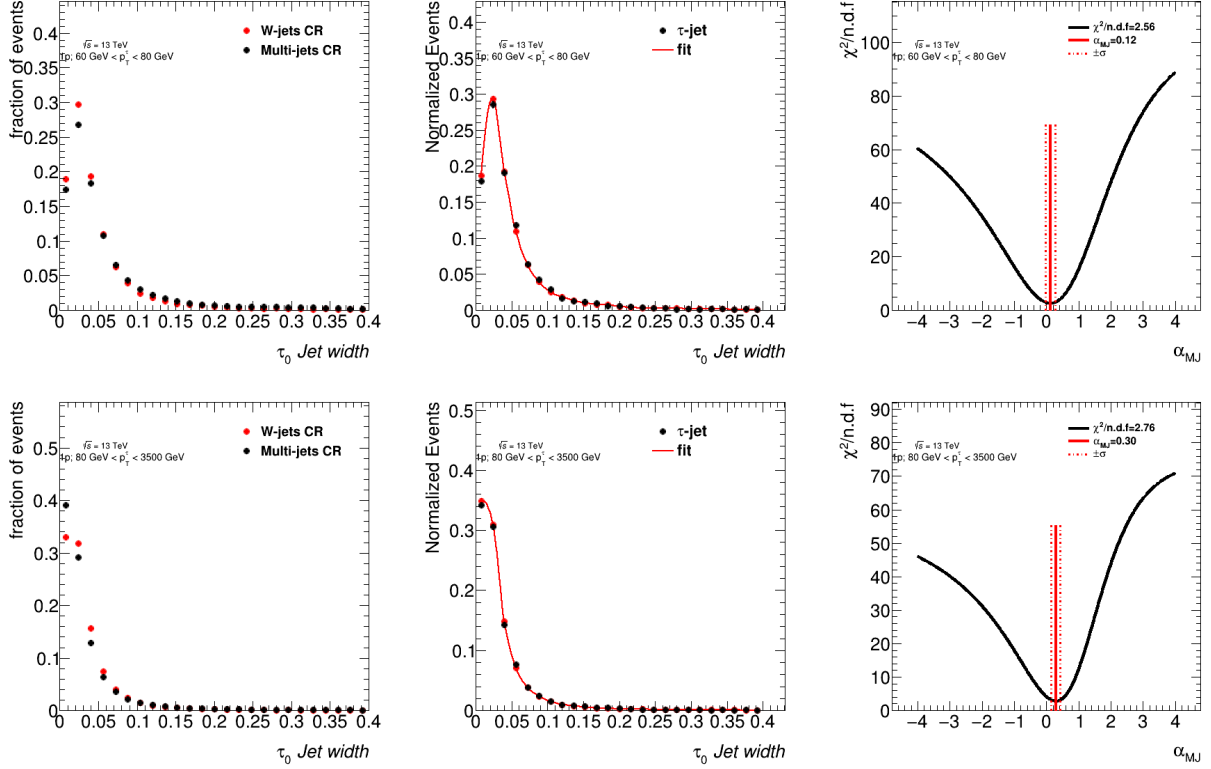


Figure B.4: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{jets}$ signal region for $p_T \geq 60$ GeV 1-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\min}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$.

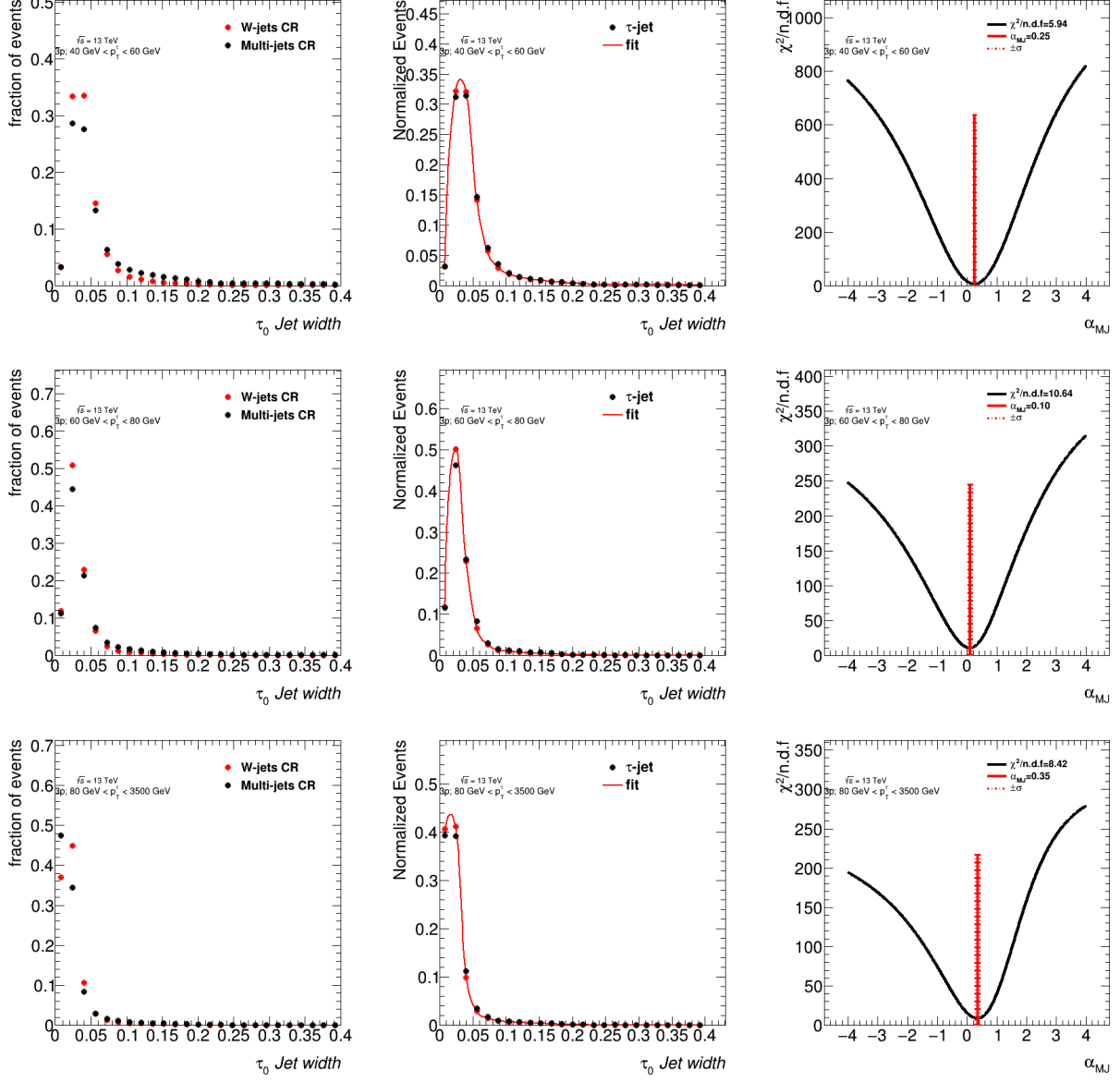


Figure B.5: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{jets}$ signal region for 3-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\min}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$.

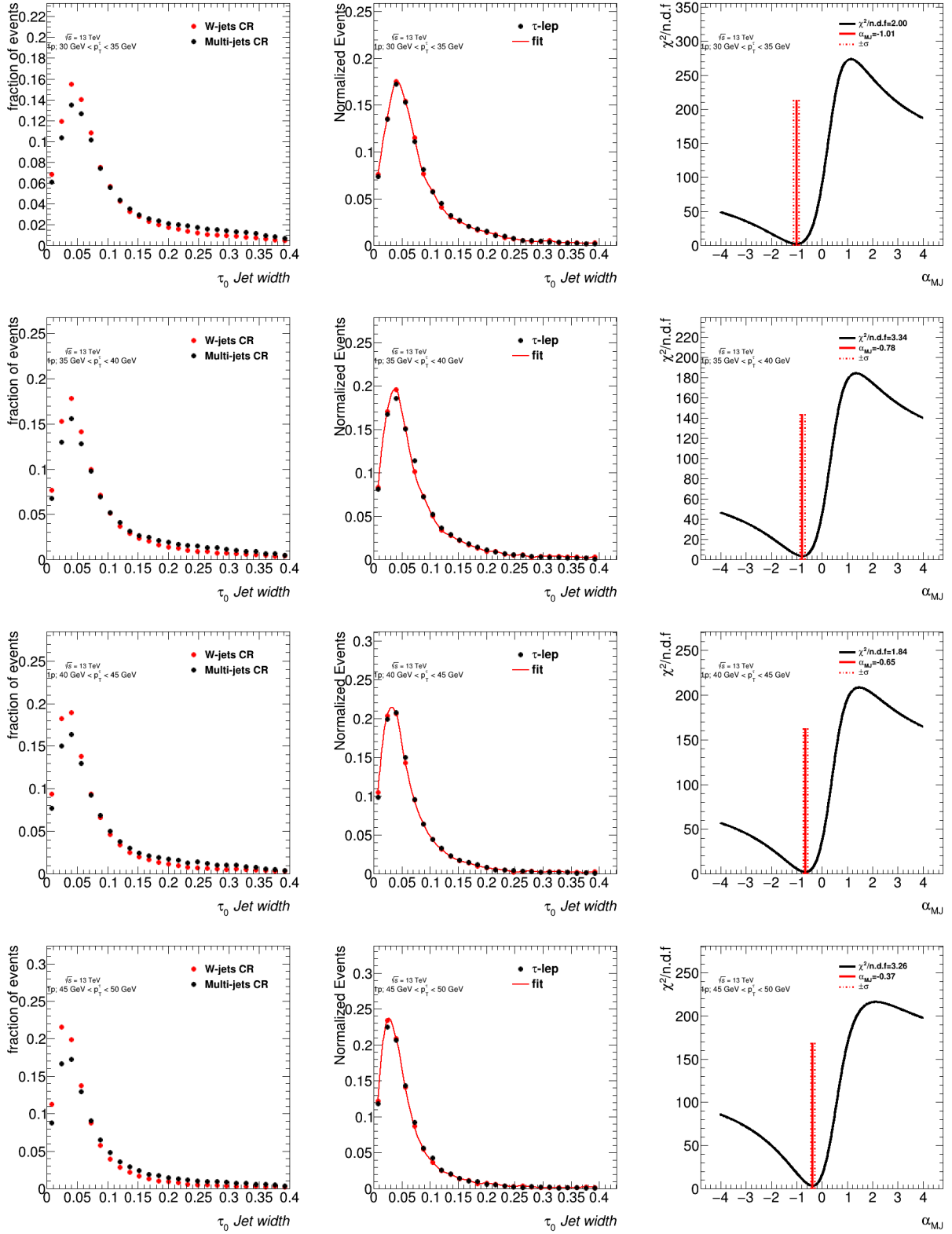


Figure B.6: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{lepton}$ signal region for $p_T \leq 50 \text{ GeV}$ 1-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: $\chi^2/\text{n.d.f.}$ of the fit as

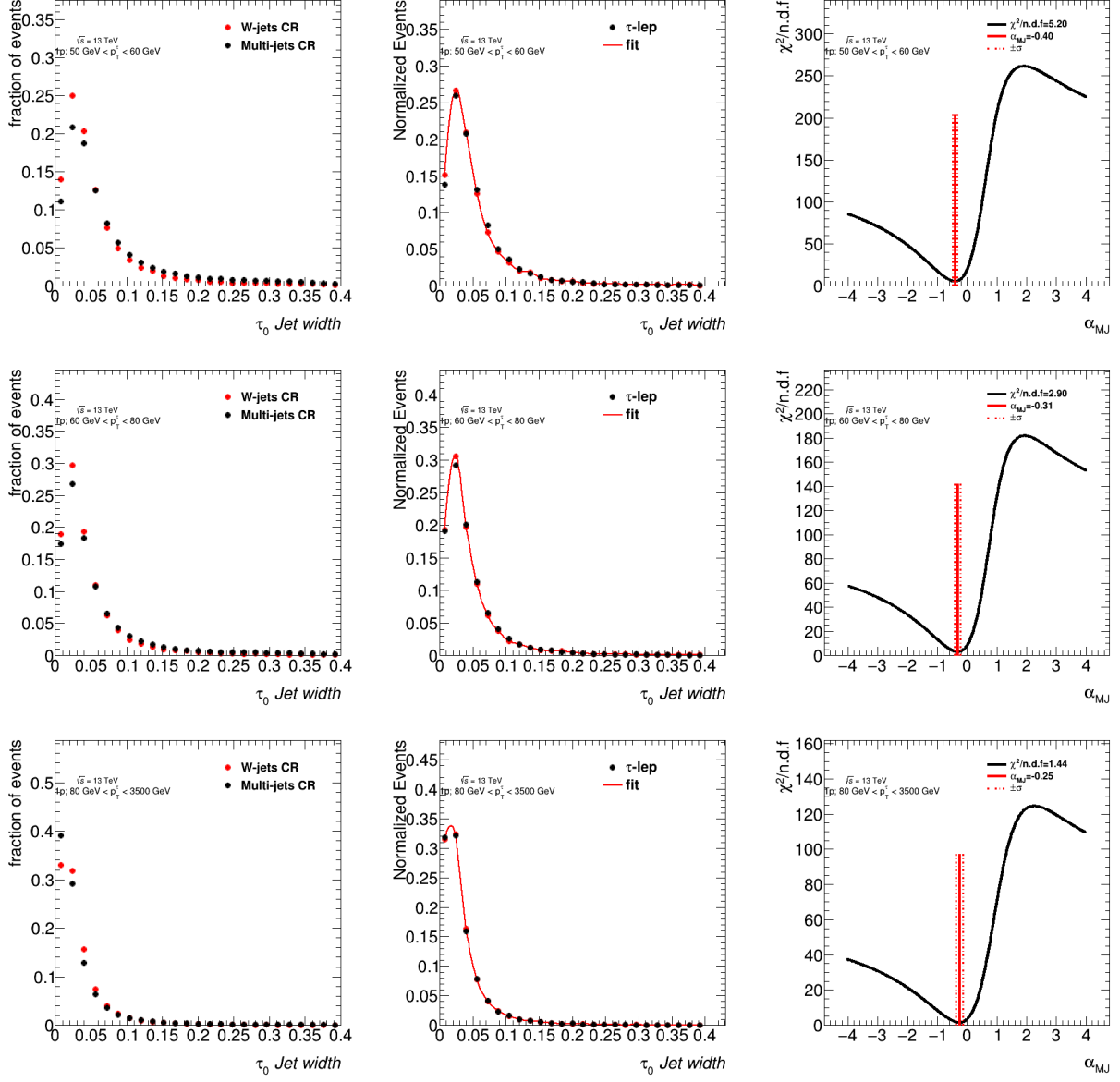


Figure B.7: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{lepton}$ signal region for $p_T \geq 50 \text{ GeV}$ 1-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\min}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$.

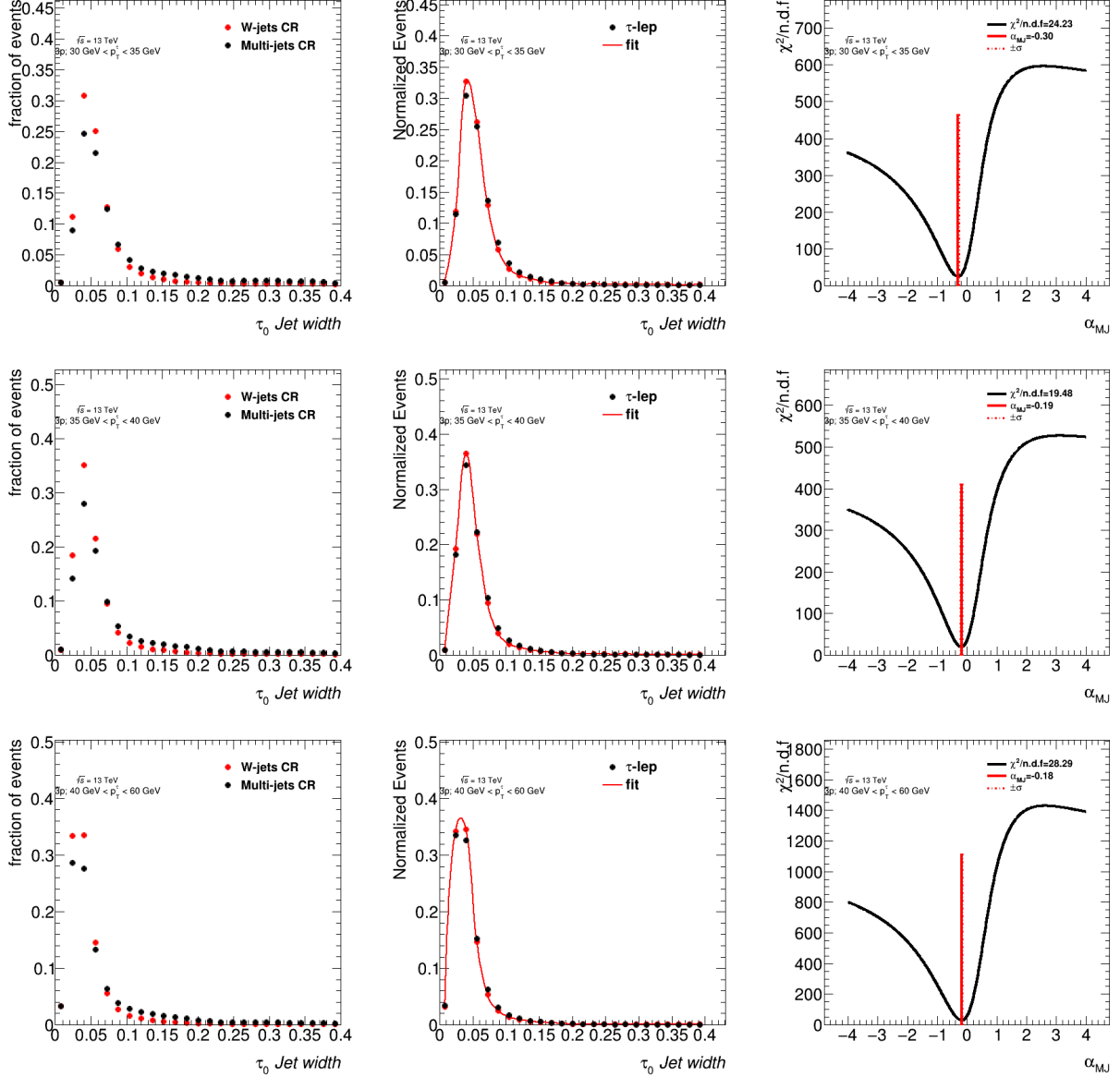


Figure B.8: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{lepton}$ signal region for $p_T \leq 60 \text{ GeV}$ 3-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\min}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$.

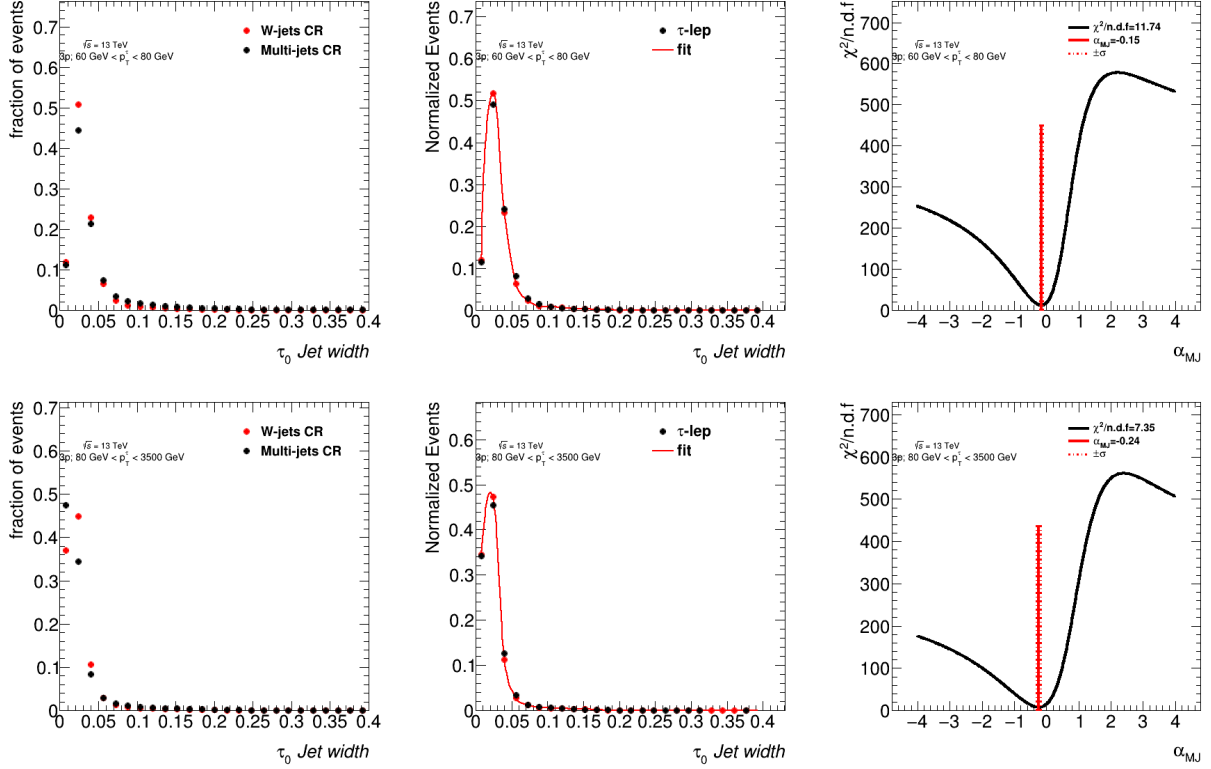


Figure B.9: Estimation of α_{MJ} in the $\tau_{had-vis} + \text{lepton}$ signal region for $p_T \geq 60 \text{ GeV}$ 3-prong $\tau_{had-vis}$ candidates. Left: templates of discriminating variables for different $\tau_{had-vis}$ p_T and n-prong slices. Middle: shape of the discriminating variable obtained in the signal region and fitted shape using the templates measured in the control regions. Right: χ^2/ndf of the fit as a function of α_{MJ} , the error on α_{MJ} is defined by the band at $\chi^2_{\min}/\text{ndf} + \sqrt{\frac{2}{\text{ndf}}}$.

APPENDIX C

ADDITIONAL VALIDATION PLOTS

C.1 τ +jets Validation Plots

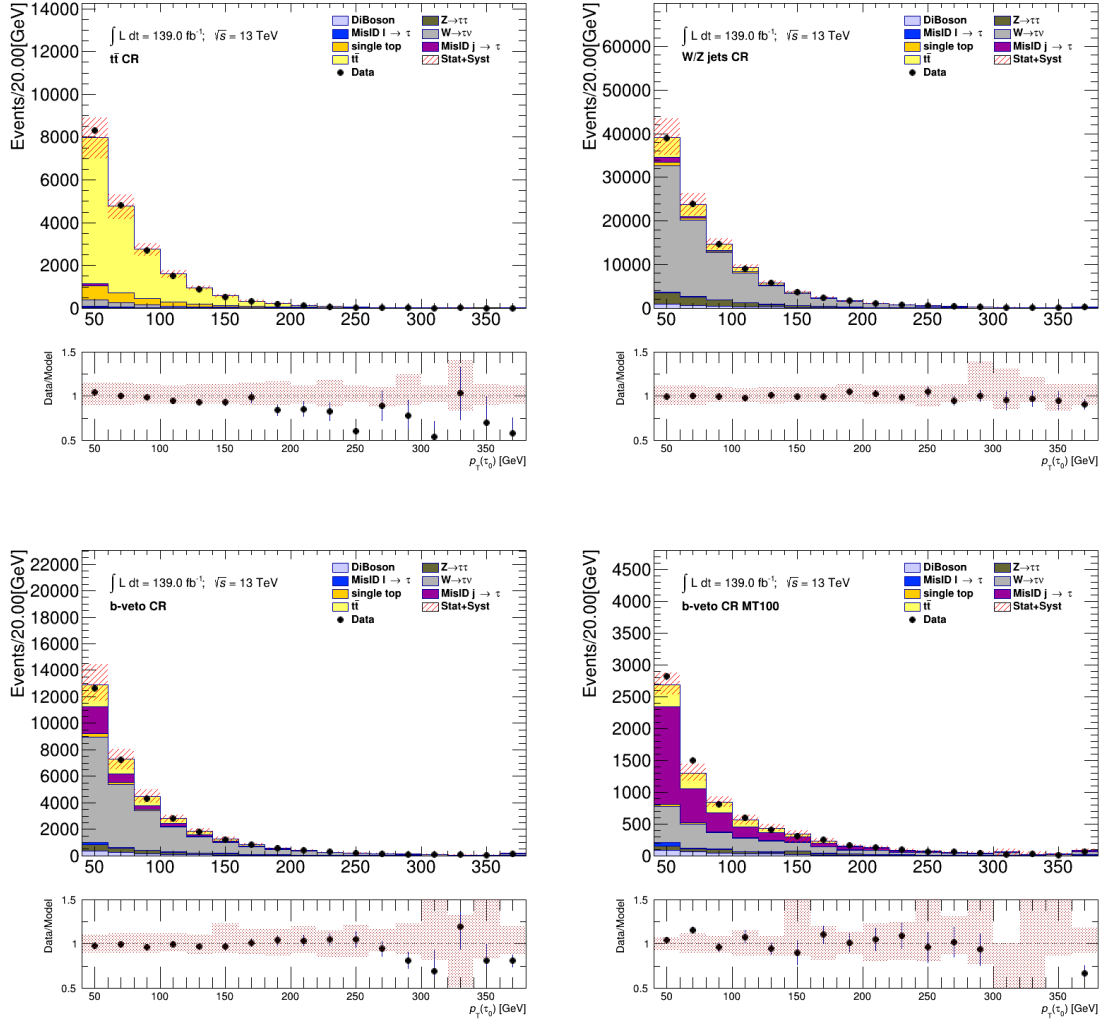


Figure C.1: Comparison between the predicted and the measured p_T^τ distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

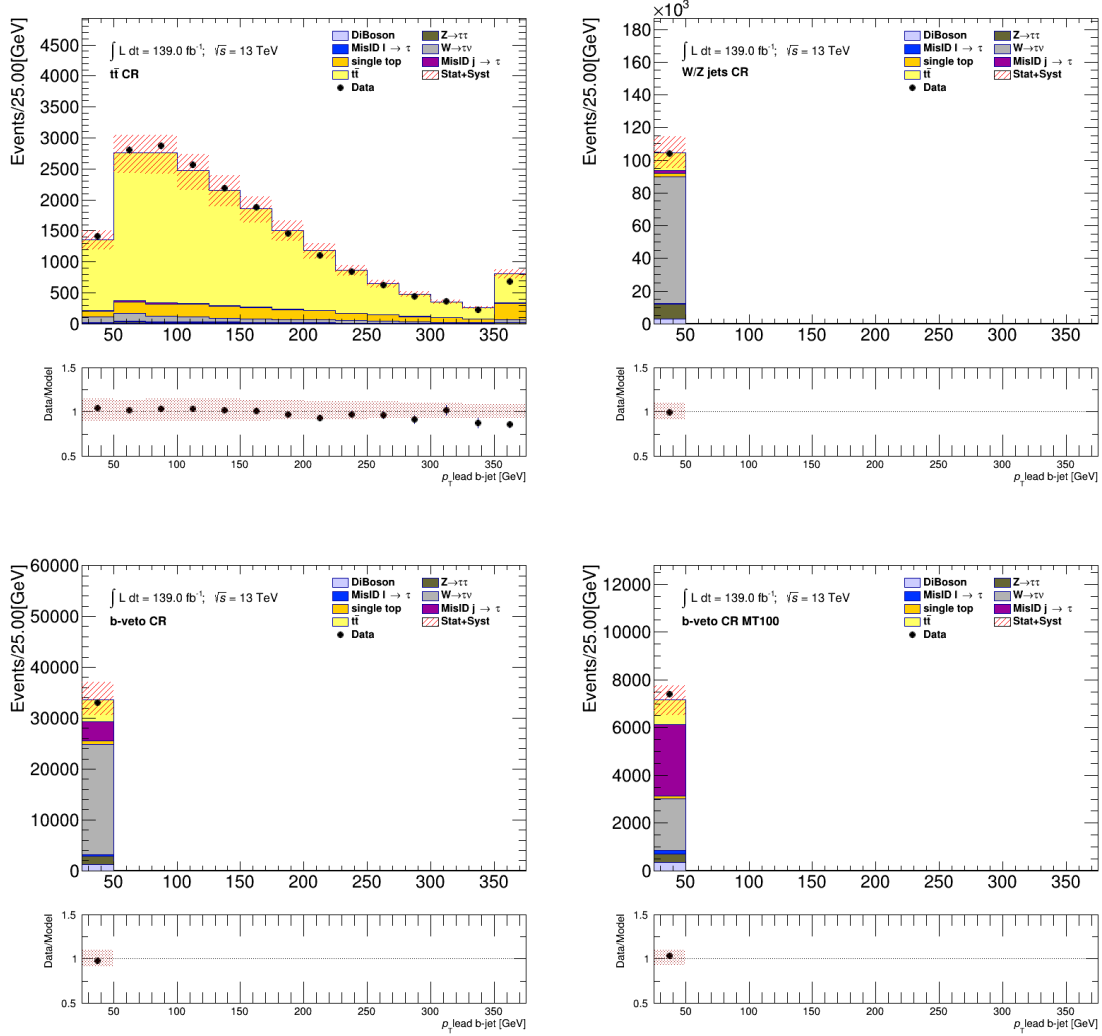


Figure C.2: Comparison between the predicted and the measured p_T^{b-jet} distributions in various control regions defined for the τ +jets channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

C.2 $\tau + \ell$ Validation Plots

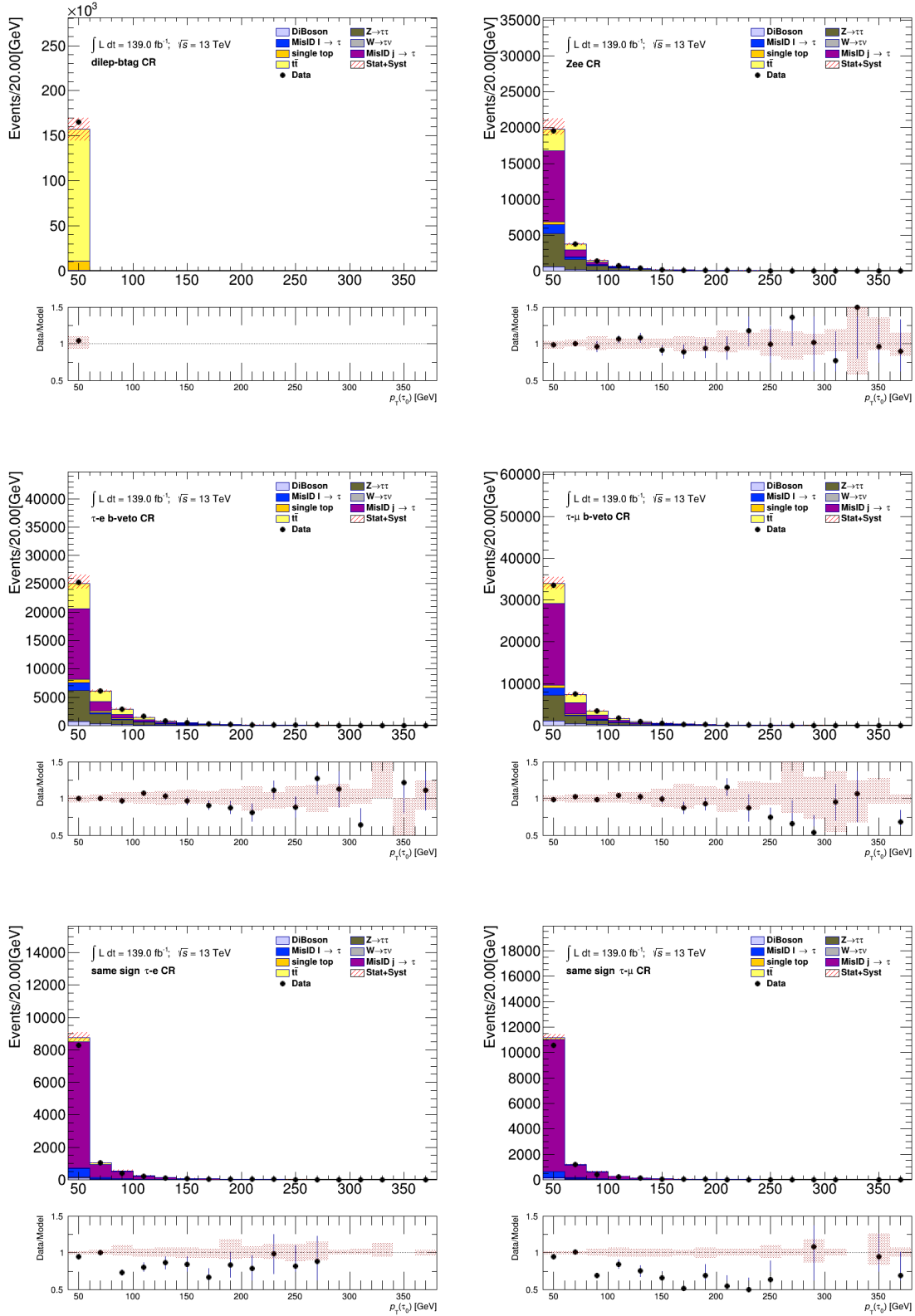


Figure C.3: Comparison between the predicted and the measured p_T^τ distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

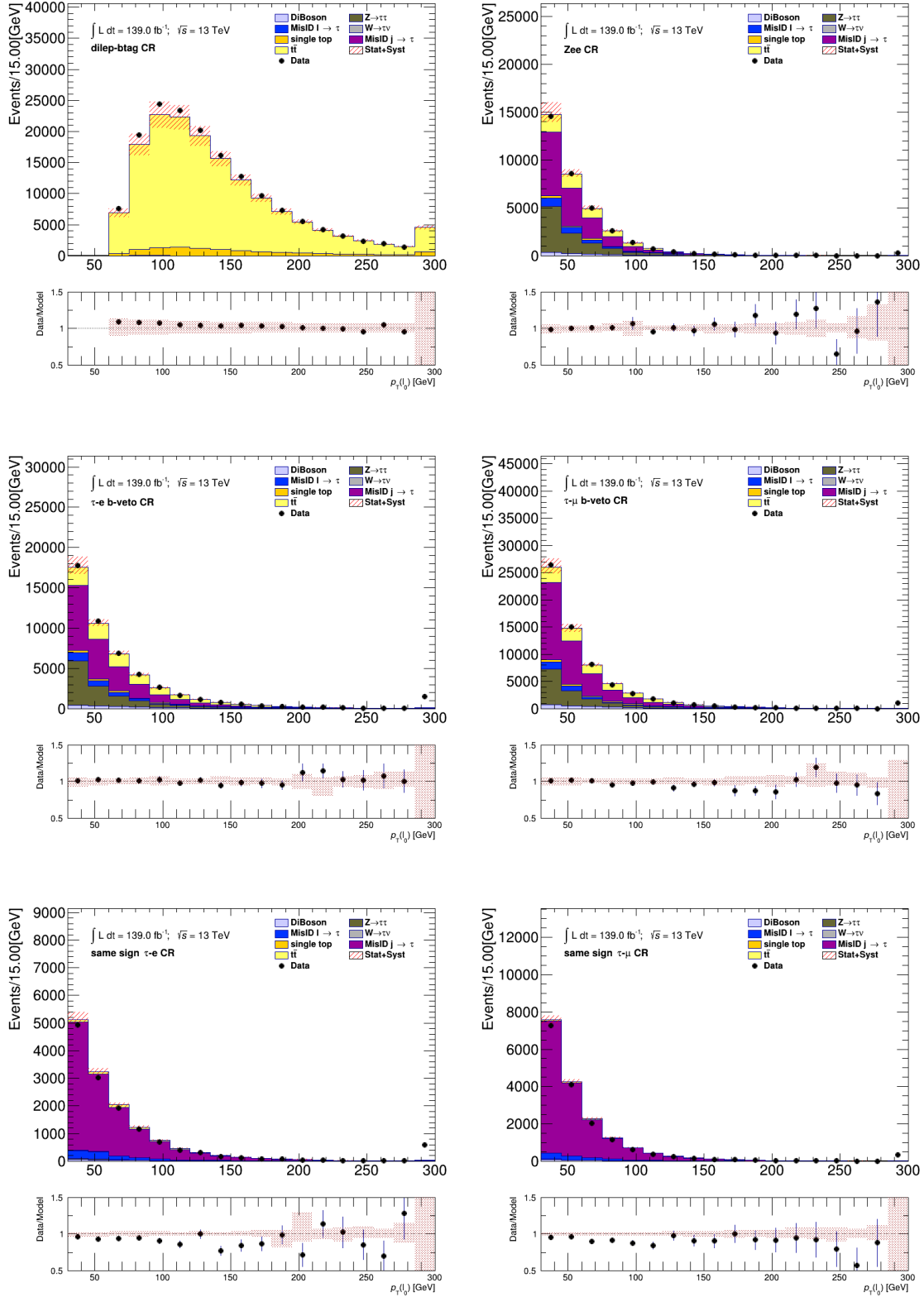


Figure C.4: Comparison between the predicted and the measured p_T^ℓ distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

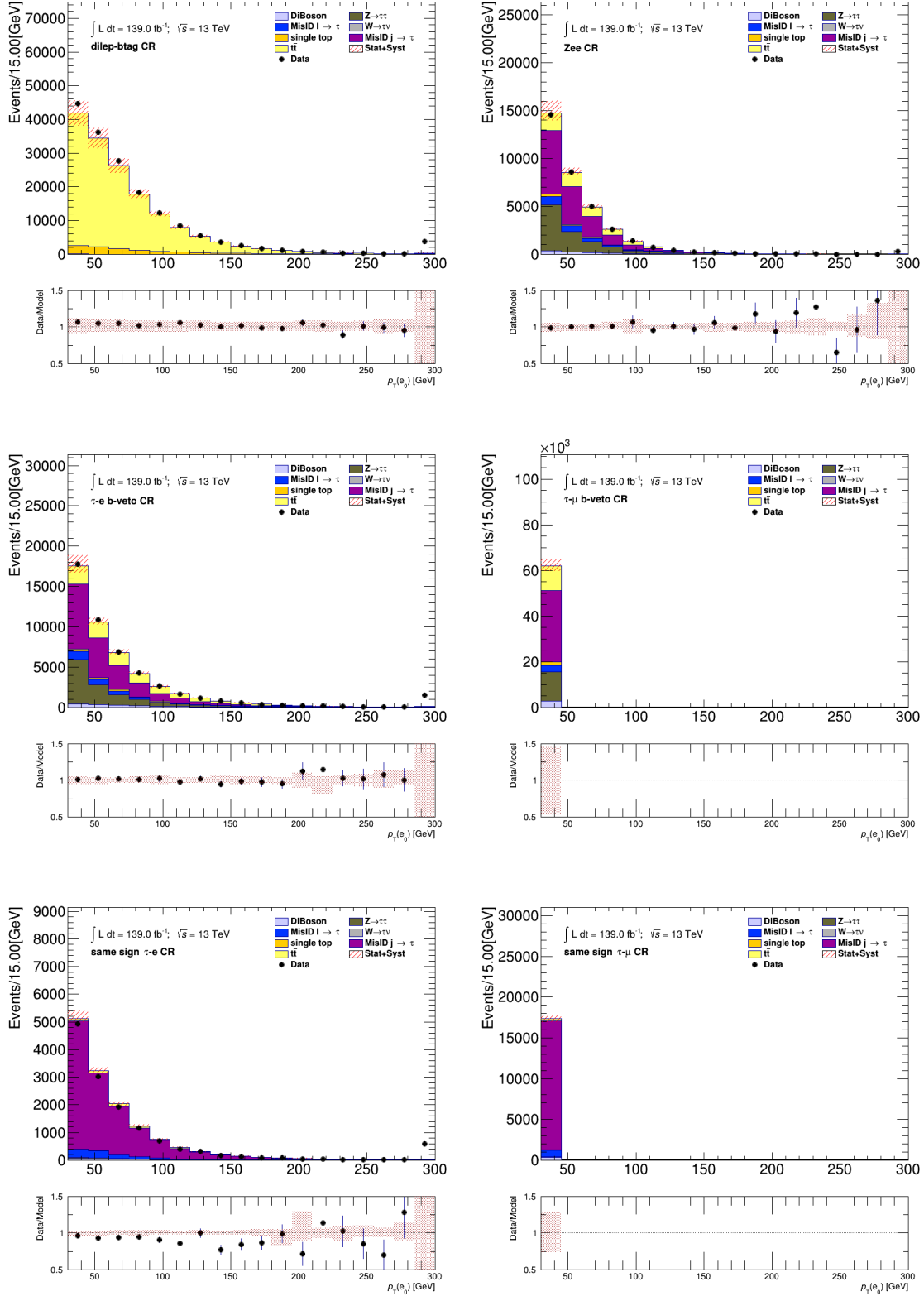


Figure C.5: Comparison between the predicted and the measured p_T^e distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

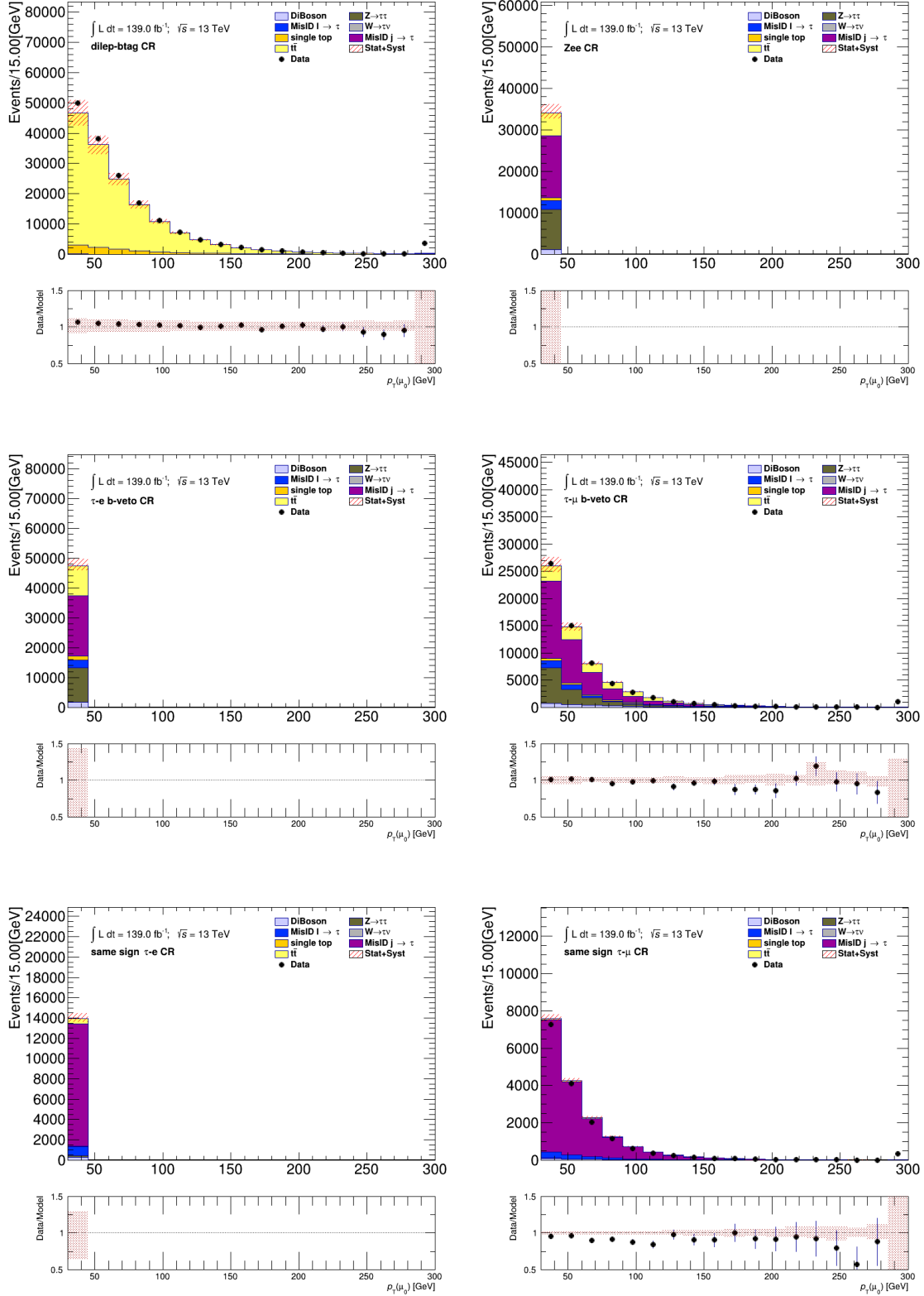


Figure C.6: Comparison between the predicted and the measured p_T^μ distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

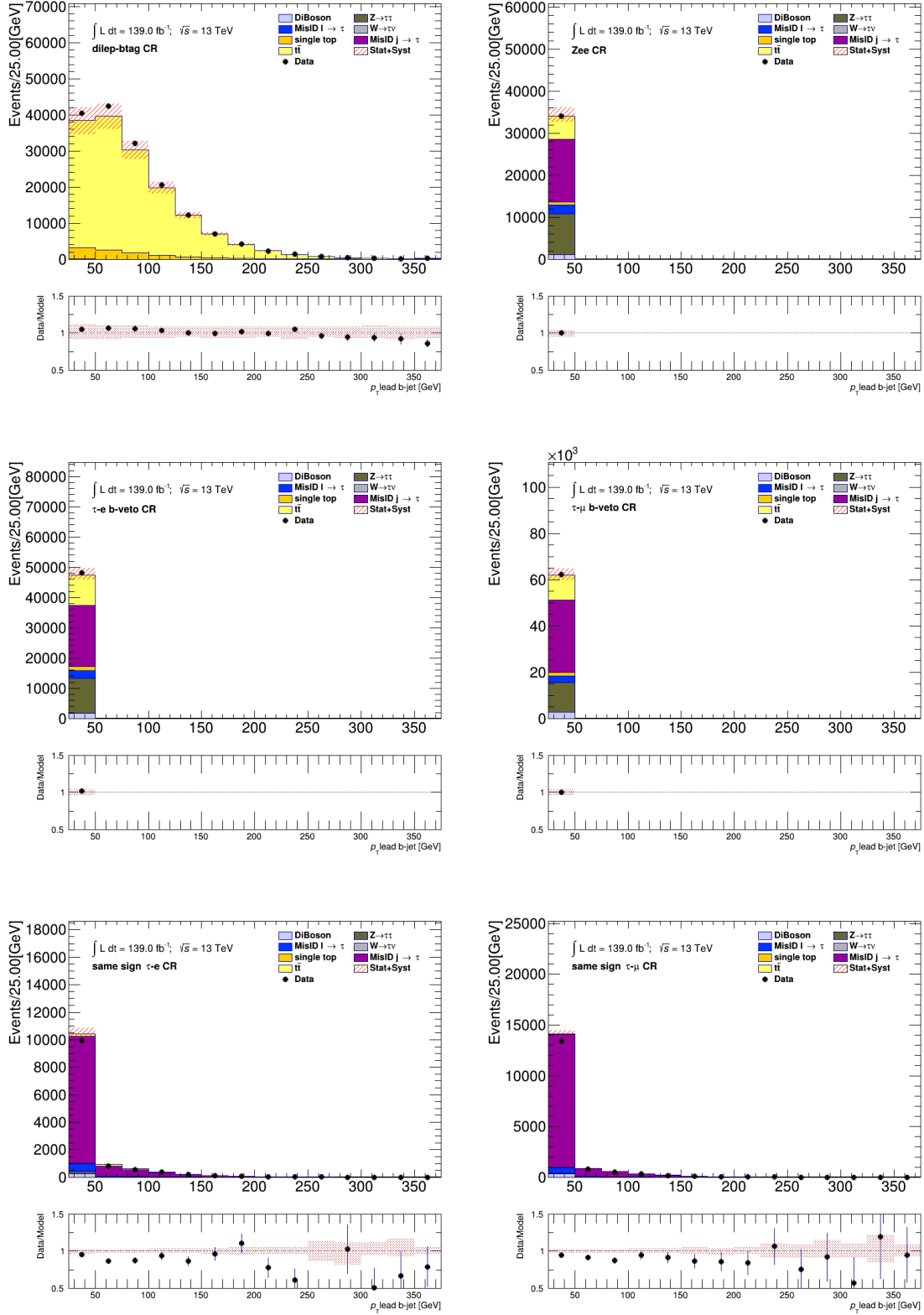


Figure C.7: Comparison between the predicted and the measured p_T^{b-jet} distributions in various control regions defined for the $\tau + \ell$ channel. The uncertainty band includes both statistical and systematic uncertainties on the background prediction.

BIBLIOGRAPHY

- [1] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021>. Geneva: CERN, June 2019.
URL: <http://cds.cern.ch/record/2677054>.
- [2] *A study of optimal parameter setting for MADGRAPH5_AMC@NLO + PYTHIA 8 matched setup*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-048>. Geneva: CERN, Nov. 2015.
URL: <https://cds.cern.ch/record/2103221>.
- [3] P. Baldi et al. “Parameterized neural networks for high-energy physics”. *The European Physical Journal C* 76.5 (2016), p. 235.
DOI: 10.1140/epjc/s10052-016-4099-4.
URL: <https://doi.org/10.1140/epjc/s10052-016-4099-4>.
- [4] URL: <https://keras.io/>.
- [5] Martín Abadi et al.
TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.
Software available from [tensorflow.org](https://www.tensorflow.org/). 2015. URL: <https://www.tensorflow.org/>.
- [6] N. Srivastava et al.
“Dropout: A Simple Way to Prevent Neural Networks from Overfitting”.
Journal of Machine Learning Research 15.56 (2014), pp. 1929–1958.
URL: <http://jmlr.org/papers/v15/srivastava14a.html>.
- [7] B. Burghgrave. “Search for charged Higgs Bosons in the $\tau +$ lepton final state with 36.1 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment”. PhD thesis. Northern Illinois U., 2018.

- [8] T. A. Collaboration. “Search for charged Higgs bosons decaying via $H^\pm \rightarrow \tau^\pm \nu_\tau$ in the τ +jets and τ +lepton final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment”. *JHEP* 09 (2018), p. 139.
DOI: 10.1007/JHEP09(2018)139. arXiv: 1807.07915 [hep-ex].
- [9] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”.
The European Physical Journal C 71.2 (Feb. 2011).
DOI: 10.1140/epjc/s10052-011-1554-0.
URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0>.
- [10] A. L. Read.
“Presentation of search results: Presentation of search results: the CL_s technique”.
Journal of Physics G: Nuclear and Particle Physics 28.10 (Sept. 2002), pp. 2693–2704.
DOI: 10.1088/0954-3899/28/10/313.
URL: <https://doi.org/10.1088/0954-3899/28/10/313>.