



Northern Illinois
University

Search for Charged Higgs Bosons in the $\tau + \ell$ Final State with 139 fb^{-1} of pp Collision Data at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS Experiment

Dissertation Defense

Elliot Parrish[†]

[†]Northern Illinois University, USA

October 13, 2022

Introduction

ATLAS

Theory

Simulation

The Standard Model

$$H^\pm \rightarrow \tau^\pm \nu_\tau$$

Higgs Boson

Background Modeling

New Physics

MVA

Charged Higgs Bosons

PNN HPO

Experimental Apparatus

Results

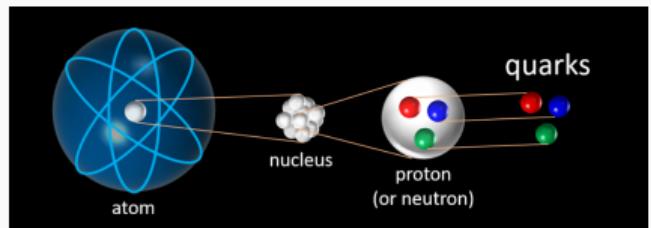
LHC

Conclusion

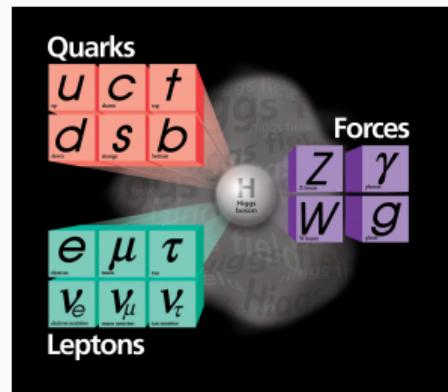
- This defense will take ~ 1 hour
 - I will walk you through the work that is contained in my PhD dissertation
 - After the presentation is complete, the committee and I will address comments privately
 - When we are done, I will return, the committee will discuss among themselves then return
- General Guidelines
 - Please remain muted unless you are speaking
 - There will be time at the end for questions, but feel free to interrupt if there is something urgent
- Thank you for attending!

What are we made of?

- The scientific field of particle physics seeks to explain the building blocks of the universe
 - How many fundamental particles are there?
 - How do they interact with each other?
- The Standard Model of Particle Physics (SM)
 - Matter is comprised of fermions
 - Quarks combine to create hadrons (protons, neutrons, $\pi^{\pm, 0}$, etc)
 - Anti-matter is identical to matter except for opposite electromagnetic charge
 - Forces are carried by an exchange of bosons
 - Gluon (g) → Strong force
 - Photon (γ) → Electromagnetism
 - W^\pm, Z^0 → Weak force

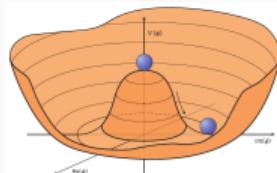


[1] [2]

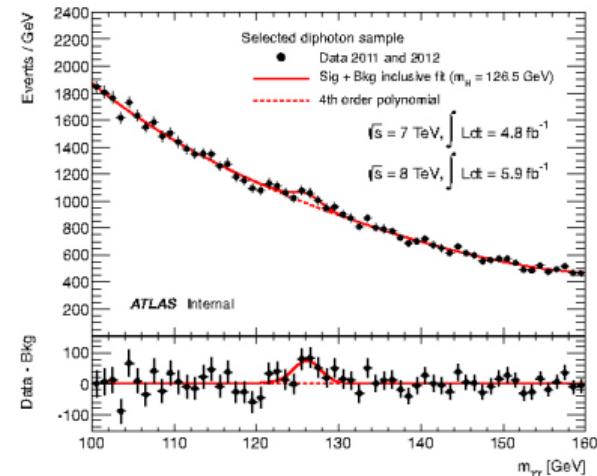


The Higgs Boson

- Theorized by Higgs, Englert, and Brout in 1964
 - Complex scalar doublet ($s = 0$)
 - Non-zero vacuum expectation value
- Interactions with Higgs field give particles mass
- Discovered jointly by the ATLAS and CMS collaborations in 2012



Higgs potential [3]

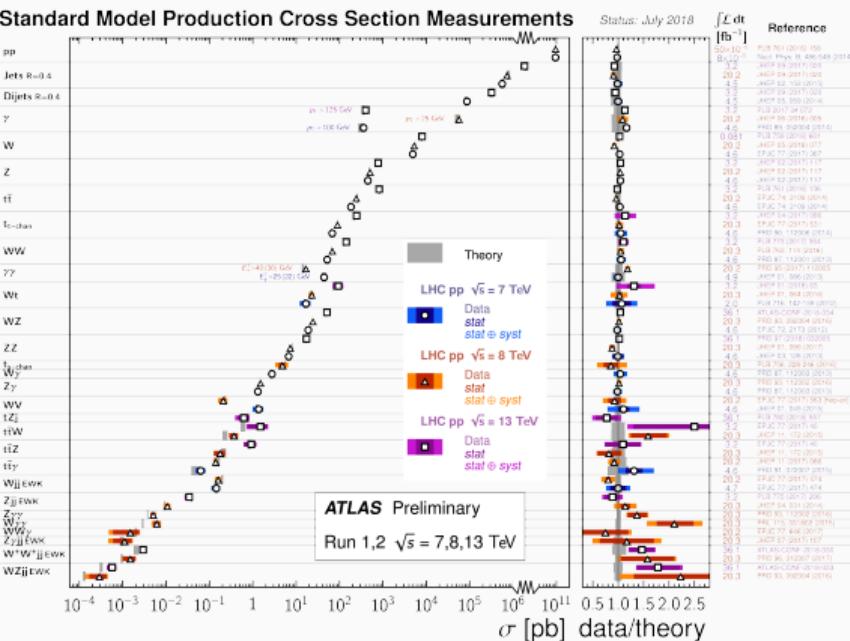


Higgs discovery [4]

The Standard Model

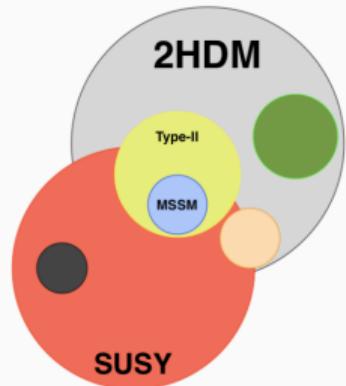


- Predicts the probabilities of particles decaying to others (among many other things)
 - Has been thoroughly tested
 - Measurements agree to a high degree of accuracy
 - Not a complete theory
 - Gravity
 - Matter-antimatter asymmetry in the universe
 - Predicted neutrino masses are 0
 - Observed neutrino mixing says otherwise



Beyond the Standard Model

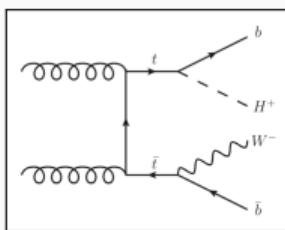
- Hierarchy problem, “unnaturalness”
 - Electroweak scale is ~ 100 GeV
 - Planck scale is $\sim 2.4 \times 10^{18}$ GeV
- 2 Higgs Doublet Models and Supersymmetry (SUSY) are large groups of theories
 - 2HDM have two complex doublet scalar fields [5]
 - Two relevant free parameters, $\tan \beta$ and m_{H^\pm}
 - $\tan \beta$ is the ratio of the vacuum expectation values of H^\pm
 - SUSY proposes a symmetry between fermions and bosons
 - Many new possible particles
 - Minimal Supersymmetric Standard Model (MSSM) is the smallest SUSY extension to the SM



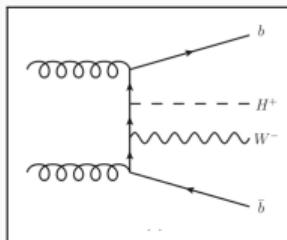
light neutral scalar	h^0
heavy neutral scalar	H^0
neutral pseudoscalar	A^0
two charged scalars	H^\pm

Charged Higgs Bosons

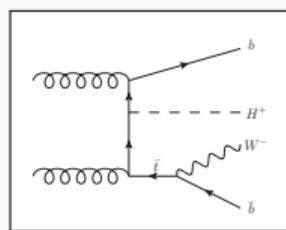
- At the LHC, theoretical production mode of m_{H^\pm} is mainly in top-quark decays or in association with a top-quark (t)
 - H^\pm production mode is dependent on m_{H^\pm}



$$m_{H^\pm} < m_t$$



$$m_{H^\pm} \simeq m_t$$



$$m_{H^\pm} > m_t$$

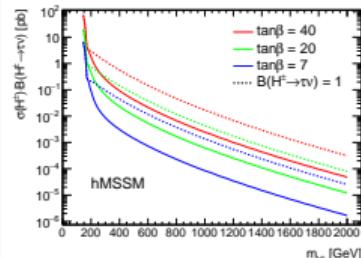
- $H^\pm \rightarrow \tau^\pm \nu_\tau$ decay channel remains significant for high $\tan \beta$
- Two sub-channels based on the decay mode of associated t

$t \rightarrow jets$

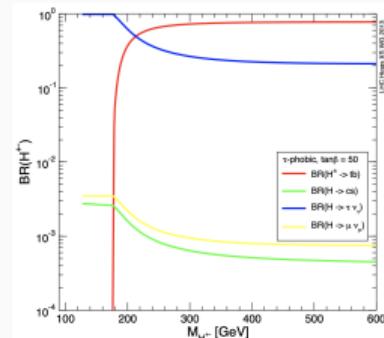
Sensitive at high mass due to higher $W \rightarrow q\bar{q}$ BR

$t \rightarrow \ell$

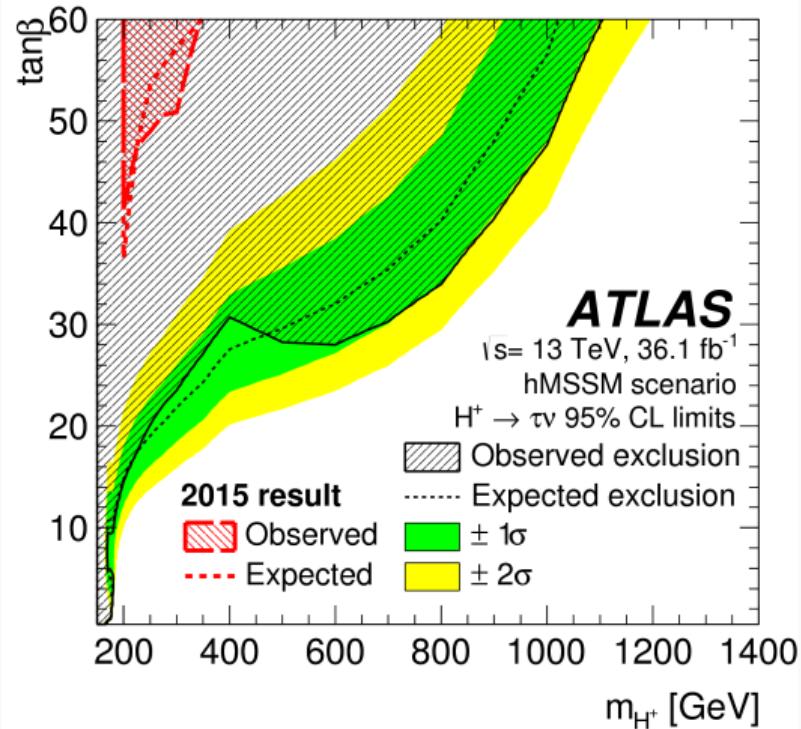
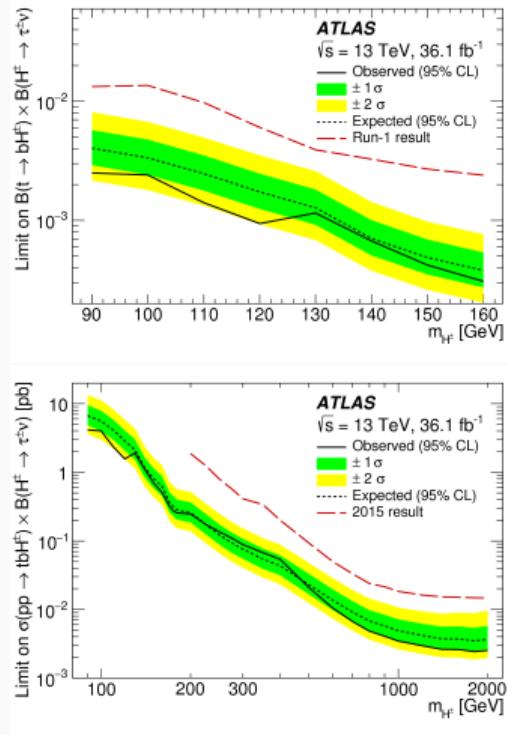
Sensitive at low mass due to easier triggering



[6]

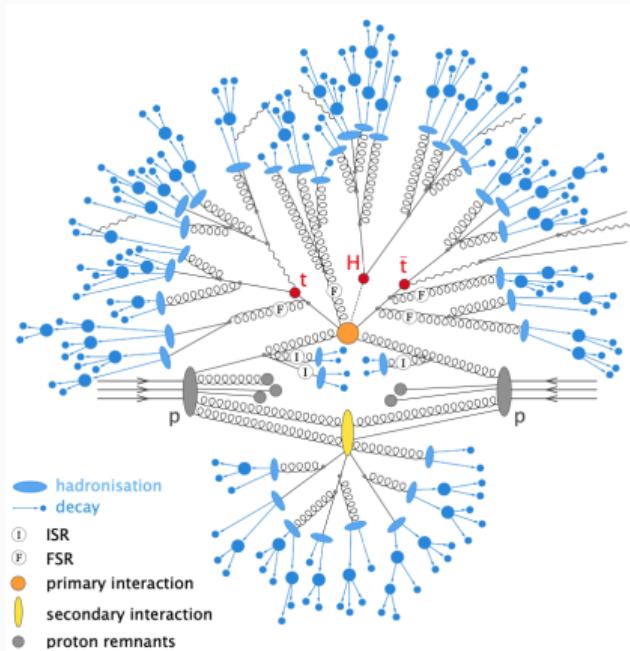


[7]



Particle Collisions

- Most particles we are interested in are not stable and do not exist in our environment
- Collide particles to study them and potentially new particles
 - $E^2 = m^2 c^4 + p^2 c^2$
- Hadronization
 - Quarks cannot exist on their own
 - Must combine to form bound states
 - Combine shower into a jet



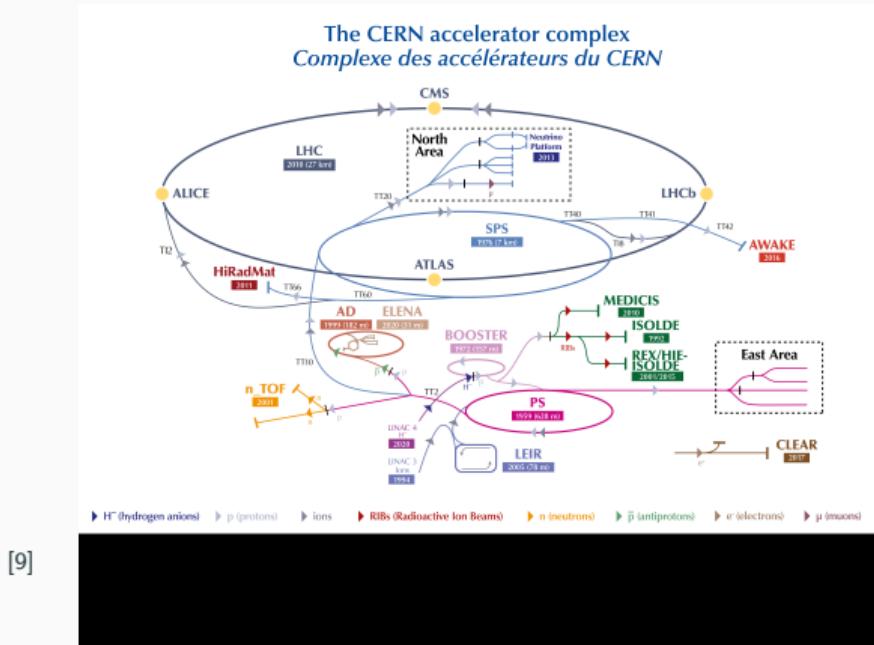
[8]

Large Hadron Collider



- Largest particle collider ever built
- Highest energy particle collider
- Located at CERN outside of Geneva, Switzerland
- Four main collision points
 - **ATLAS**, ALICE, CMS, LHCb

Circumference	26,659 m
Magnet operating temperature	1.9 K
Beam energy	6.5 TeV (13 CM TeV)
Protons per bunch	1.2×10^{11}
Bunches per beam	2808
Speed of bunches	> 1,000,000,000 km/hr ($\simeq 99.999\% c$)



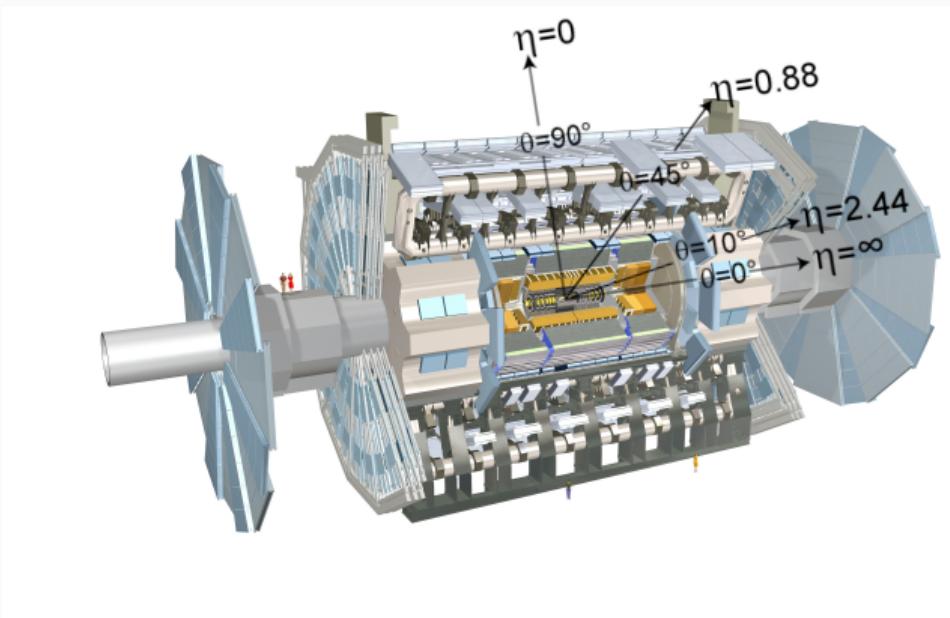
[9]

[10]

The ATLAS Detector



- General purpose particle detector
 - Magnet System
 - Tracker
 - Calorimeters
 - Muon Spectrometer
- Coordinate system origin at interaction point
 - r is radial distance
 - $\eta \equiv -\ln(\tan(\frac{\theta}{2}))$
 - ϕ is azimuthal angle
- Complex machine with
 $\sim 3,000$ scientific authors

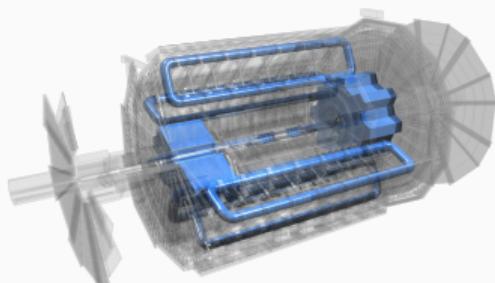


[11] [12]

Magnet System and Inner Detector

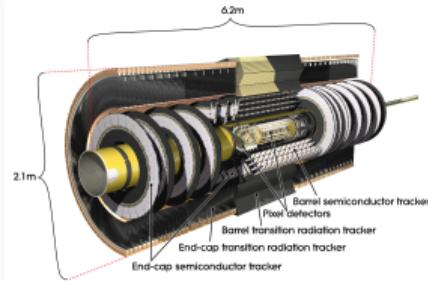
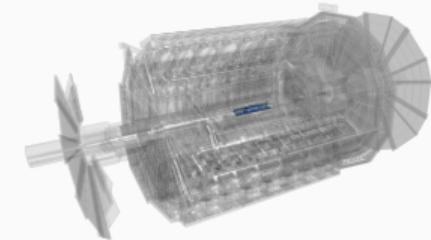
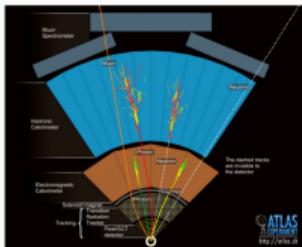


- Central solenoid
 - 2 T magnetic field
 - Bends charged particles in transverse plane
- Toroid system
 - 3.9 T magnetic field
 - Bends charged particles (μ) along beam axis



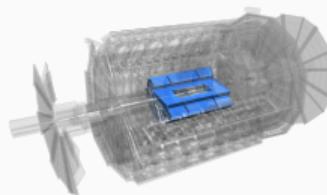
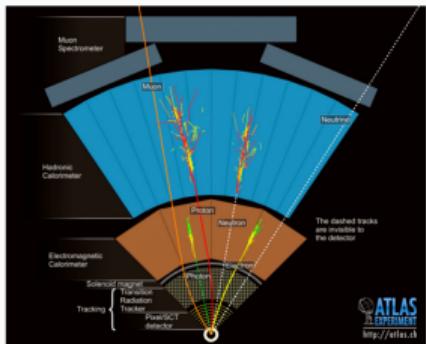
[11]

- Tracks trajectories of charged particles
 - Used to measure momentum in the transverse plane (p_T)

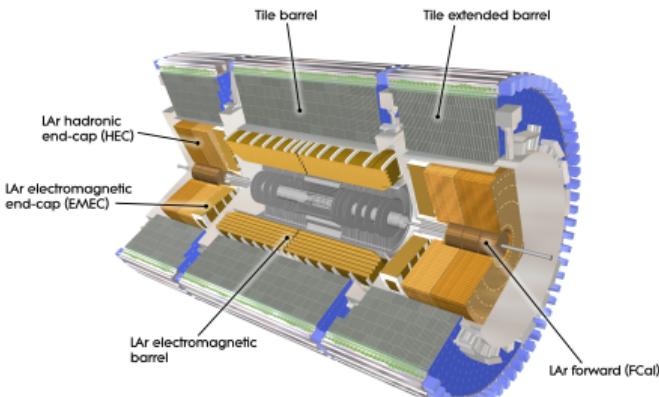


Calorimeters

- Measures energy of charged particles
 - Fully absorb particles
 - Liquid argon electromagnetic calorimeter
 - Scintillating tile hadronic calorimeter



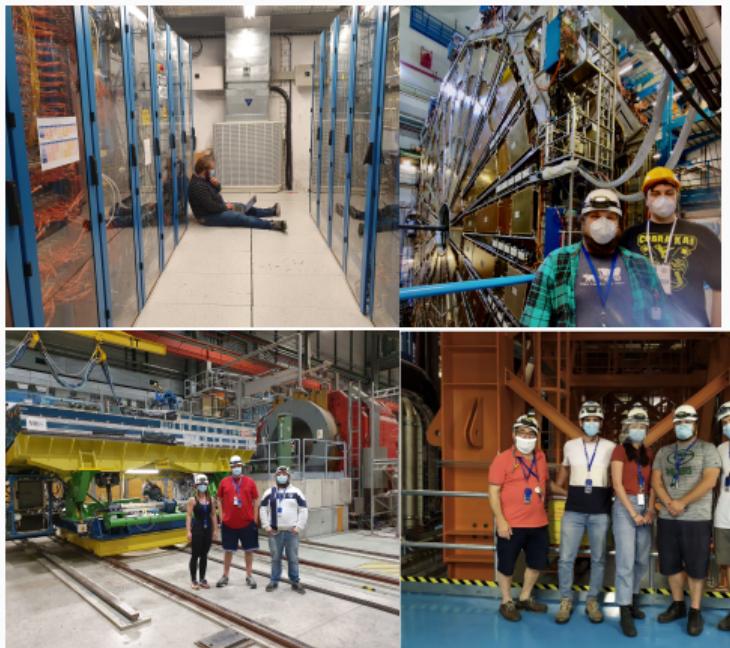
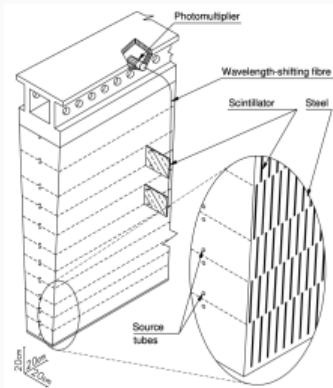
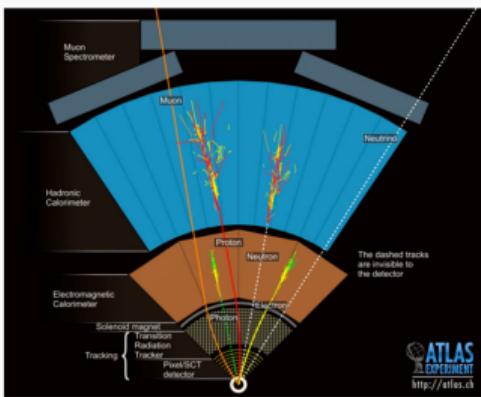
[11]



Tile Calorimeter



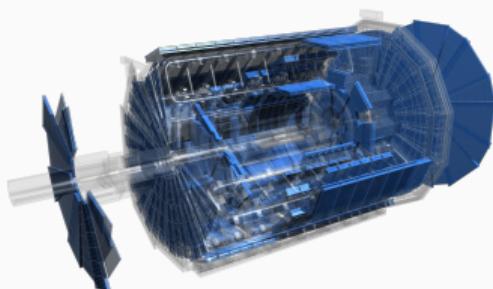
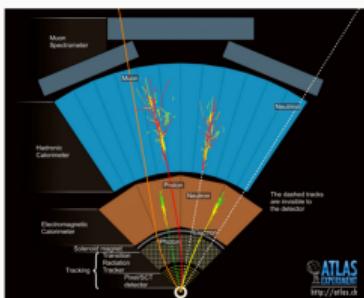
- Scintillating tile with steel absorber
- I served as Data Quality Co-Coordinator for several years
 - In Run-2 99.65% DQ efficiency [13]



Muon System and Trigger



- Muon Spectrometers track trajectory of μ
 - Muons are minimally ionizing
 - μ reach the outermost region of the detector
- Information combined with inner detector to reconstruct μ

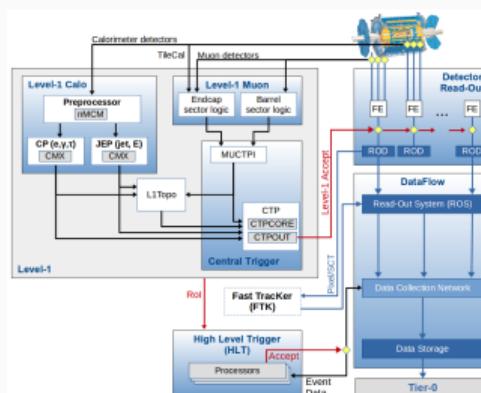


Search for $H^\pm \rightarrow \tau^\pm \nu_\tau$ with ATLAS

Experimental Apparatus ATLAS

Trigger System

- Need to quickly sort through data and decide if a collision is interesting or not
- Mix of hardware and software



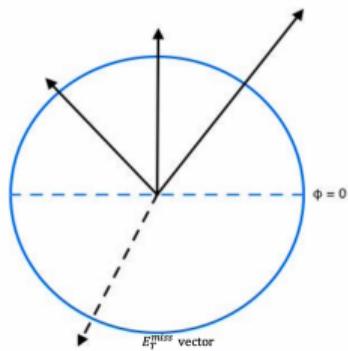
[15]

Elliot Parrish (NIU)

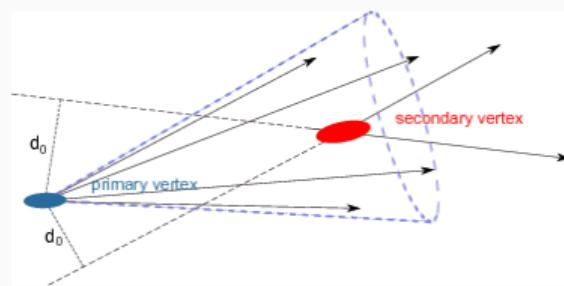
16 / 35

E_T^{miss} and b -jets

- Particles like ν do not leave a signature within ATLAS
- Instead, infer their presence through energy conservation
 - Negative vector sum of all objects in a collision event



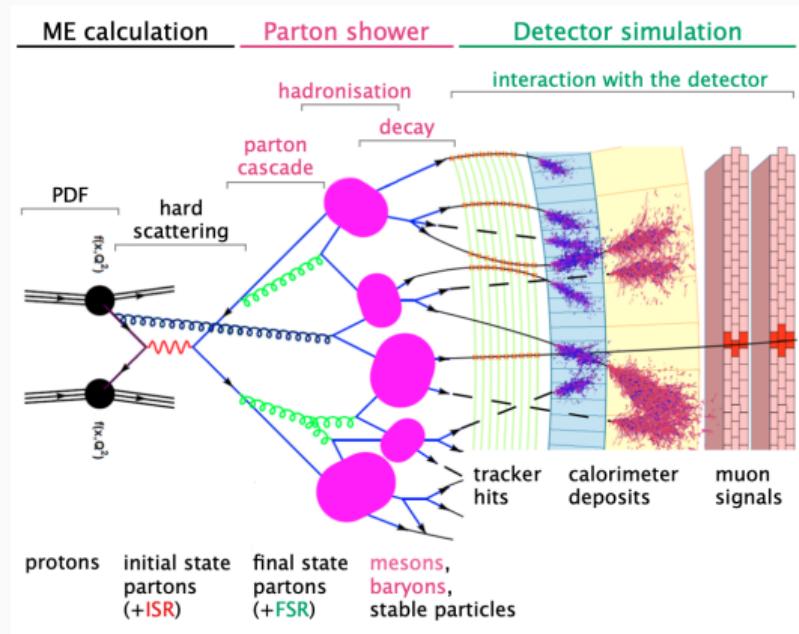
- b quark initiated hadronic showers are identified by a displaced vertex
- Reconstructed as a jet, tagged as a b -jet
 - Algorithms designed to identify b -jets called taggers [16]



[17]

Simulation

- Incredibly detailed simulations are used to create analysis strategy
 - Entire data flow from collision to detector readout is simulated
- Control regions are used to verify simulation agreement with data



[8]

Analysis Overview

- Search for singly charged H^\pm decaying to $\tau^\pm \nu_\tau$ over a wide mass range
 - Low mass ($m_{H^\pm} < m_t$):
 - Intermediate mass* ($m_{H^\pm} \simeq m_t$)
 - High mass ($m_{H^\pm} > m_t$)

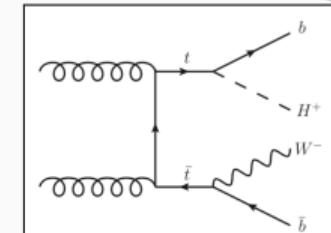
- Dominant backgrounds

Backgrounds w/ prompt hadronic τ	Backgrounds w/ fake τ
$t\bar{t}$ estimated with simulation	Fake $j \rightarrow \tau$ estimated with data driven fake factor method
$V + jets$ estimated with simulation	Fake $\ell \rightarrow \tau$ estimated with simulation, validated on $Z \rightarrow ee$
VV estimated with simulation	

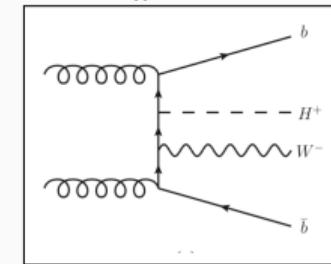
- MVA score is used as the final discriminant

Sub-Channel							
$\tau + jets$ SR	E_T^{miss} Trigger	1 hadronic τ $p_T^\tau > 40$ GeV	0 ℓ (e or μ) $p_T^\ell > 20$ GeV	≥ 3 jets $p_T^j > 25$ GeV	≥ 1 b-jets $p_T^{b-jet} > 25$ GeV	$E_T^{\text{miss}} > 150$ GeV	$m_T(\tau, E_T^{\text{miss}}) > 50$ GeV
$\tau + \ell$ SR	Single Lepton Trigger	1 hadronic τ $p_T^\tau > 30$ GeV	1 ℓ (e or μ) $p_T^\ell > 30$ GeV	≥ 1 jet $p_T^j > 25$ GeV	≥ 1 b-jets $p_T^{b-jet} > 25$ GeV	$E_T^{\text{miss}} > 50$ GeV	Opposite sign τ and ℓ

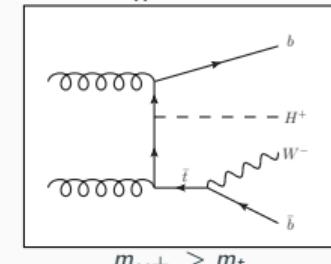
*: First time probed experimentally JHEP 09(2018)139



$$m_{H^\pm} < m_t$$



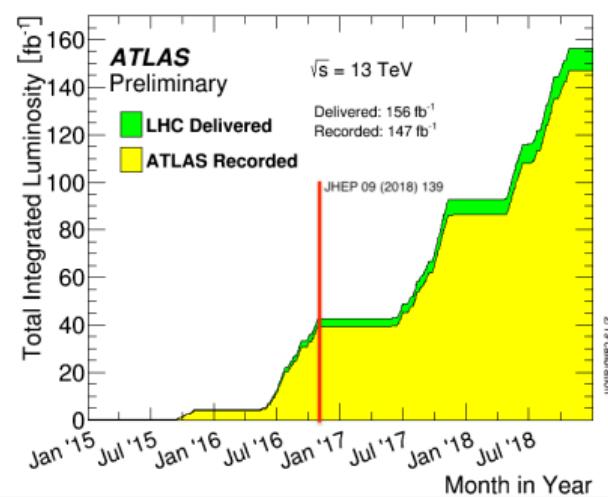
$$m_{H^\pm} \simeq m_t$$



$$m_{H^\pm} > m_t$$

Updates to analysis since last publication

- Signal mass range extended
 - Previous: $90 \leq m_{H^\pm} \leq 2000$ GeV
 - Current: $80 \leq m_{H^\pm} \leq 3000$ GeV
- Signal filtering applied in order to effectively increase the statistics in the signal regions
- New analysis framework centered around using modern Machine Learning tools
- Investigated new multivariate analysis techniques
- Improved particle identification algorithms
- Almost 4× more data

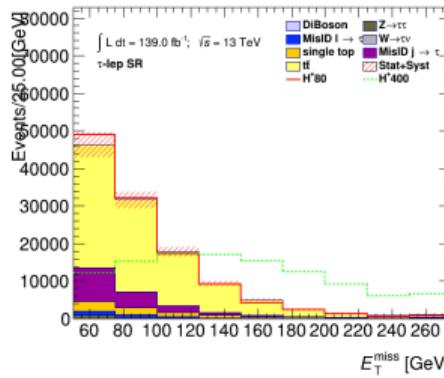
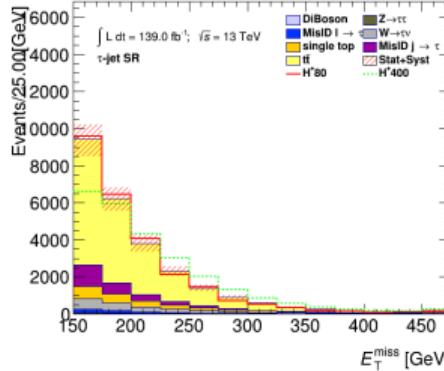


Background Estimation

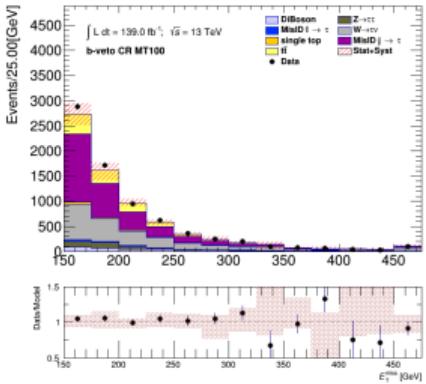
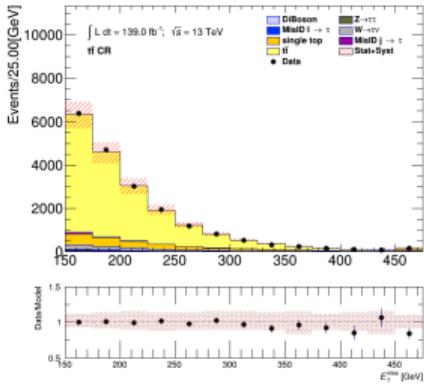
- Control regions defined to verify main sources of background

	$t\bar{t}$ CR	W+Jets CR	b-veto CR	b-veto $m_T > 100$ CR
Number of $\tau_{had-vis}$	1	1	0	0
p_T^τ	> 40 GeV	> 40 GeV	> 40 GeV	> 40 GeV
Number of jets	≥ 3	≥ 3	≥ 3	≥ 3
p_T^{jet}	≥ 25 GeV	≥ 25 GeV	≥ 25 GeV	≥ 25 GeV
Number of b-jets	≥ 2	0	0	0
Number of ℓ	0	0	0	0
E_T^{miss}	> 150 GeV	> 150 GeV	> 150 GeV	> 150 GeV
$m_T(\tau, E_T^{\text{miss}})$	< 100 GeV	< 100 GeV	> 50 GeV	> 100 GeV
Type of modeling	$t\bar{t}$	W+Jets	Close to SR	Fake $j \rightarrow \tau$ enriched

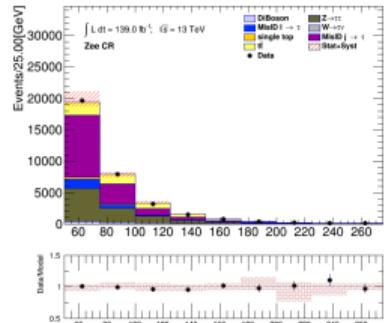
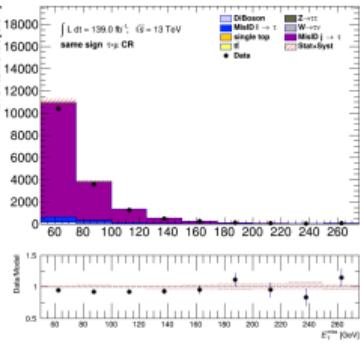
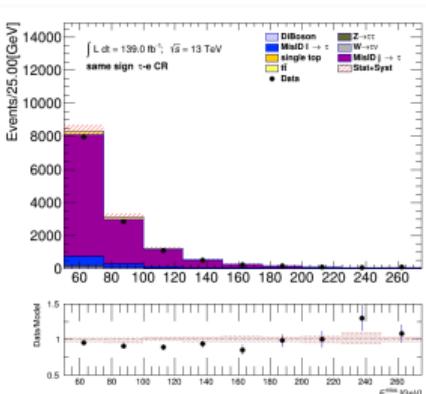
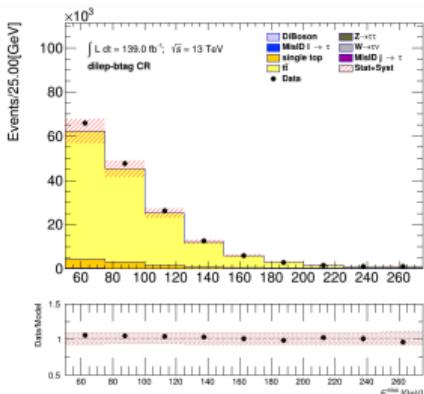
	Dilepton-btag CR	Zee CR	b-veto CR	Same Sign CR
Number of $\tau_{had-vis}$	0	1	0	0
p_T^τ	> 30 GeV	> 30 GeV	> 30 GeV	> 30 GeV
Number of jets	≥ 1	≥ 1	≥ 1	≥ 1
p_T^{jet}	≥ 25 GeV	≥ 25 GeV	≥ 25 GeV	≥ 25 GeV
Number of b-jets	≥ 1	0	0	≥ 1
Number of ℓ	2 (1 e, 1 μ)	1 e	1 tight e (μ)	1 tight e (μ)
E_T^{miss}	> 50 GeV	> 50 GeV	> 50 GeV	> 50 GeV
mass(τ, e)	N/A	$> 40; < 140$ GeV	N/A	N/A
Type of modeling	$t\bar{t}$ and single-top	Fake $\ell \rightarrow \tau$ enriched	Close to SR	Fake $j \rightarrow \tau$ enriched



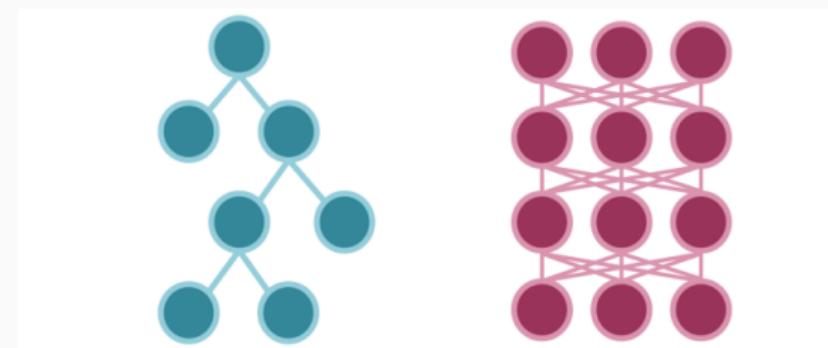
Background Modeling



- E_{τ}^{miss} distributions in control regions
- Simulation agrees with data within errors



- Boosted Decision Tree (BDT)
 - Cascading decisions bin parameter space to optimize accuracy
 - Deterministic approach
 - Excel at exploiting linear correlations
 - Often combined into a random forest



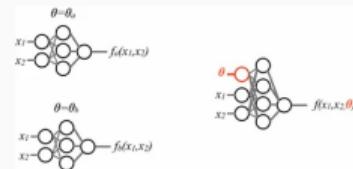
- Neural Networks
 - Network of connected nodes of activation functions connected by weights
 - Probabilistic approach
 - Excel at exploiting non-linear correlations

[18]

Parameterized Neural Networks



- Parameterized Neural Networks (PNNs) can be trained and evaluated on entire m_{H^\pm} range
 - Detailed information here: [arXiv:1601.07913](https://arxiv.org/abs/1601.07913)
- Background modeling and classifier training kept statistically independent via the k-fold method ($k = 5$)
 - Trained using simulation for backgrounds with true τ
 - Data-driven $j \rightarrow \tau$ fakes included for ≥ 500 GeV
 - Separate models trained on 1 prong and 3 prong τ
 - τ polarization used to enhance low m_{H^\pm} performance
 - $\Upsilon = \frac{E_T^{\pi^\pm} - E_T^{\pi^0}}{E_T^\tau} \approx 2 \frac{p_T^{\tau-\text{track}}}{p_T^\tau} - 1$
 - $\tau + \ell$ channel is trained inclusive for $\tau + e$ and $\tau + \mu$



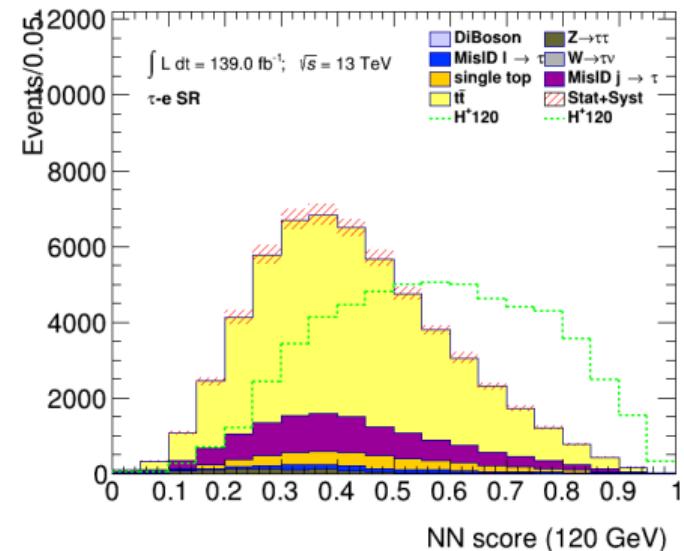
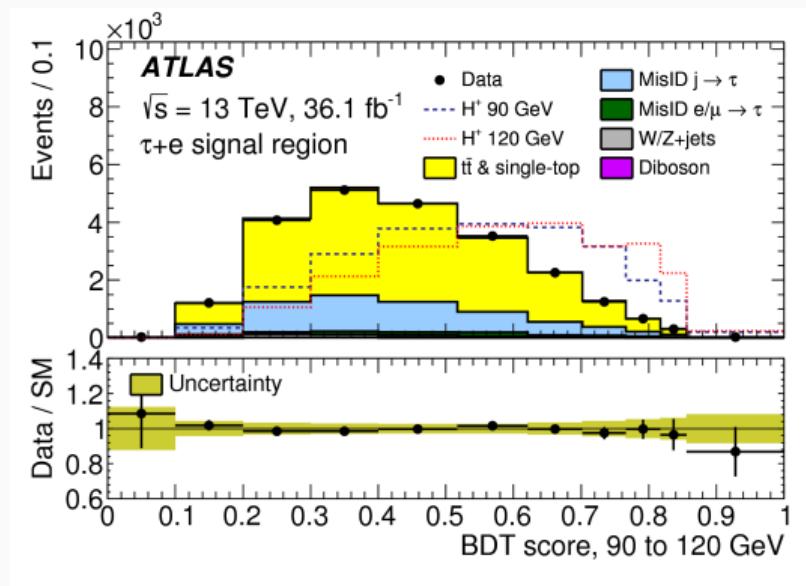
[19]

	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

$\tau + \ell$ Input Variables		
p_T^τ	η^τ	ϕ^τ
p_T^ℓ	η^ℓ	ϕ^ℓ
$p_T^{j_0}$	η^{j_0}	ϕ^{j_0}
E_T^{miss}	$\phi^{E_T^{\text{miss}}}$	p_T^h
$m_{\text{Truth}}^{H^\pm}$	Υ^τ	

$\tau + \text{jets}$ Input Variables		
p_T^τ	η^τ	ϕ^τ
$p_T^{j_0}$	η^{j_0}	ϕ^{j_0}
$p_T^{j_1}$	η^{j_1}	ϕ^{j_1}
$p_T^{j_2}$	η^{j_2}	ϕ^{j_2}
E_T^{miss}	$\phi^{E_T^{\text{miss}}}$	$m_{\text{Truth}}^{H^\pm}$
Υ^τ		

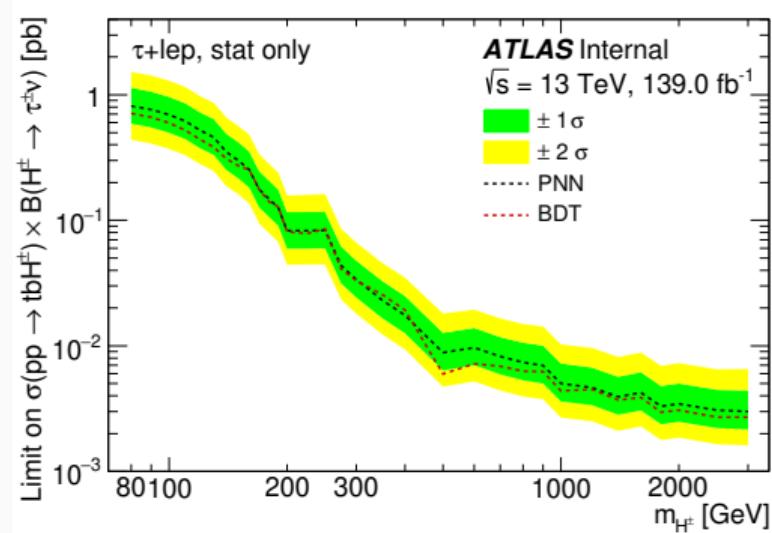
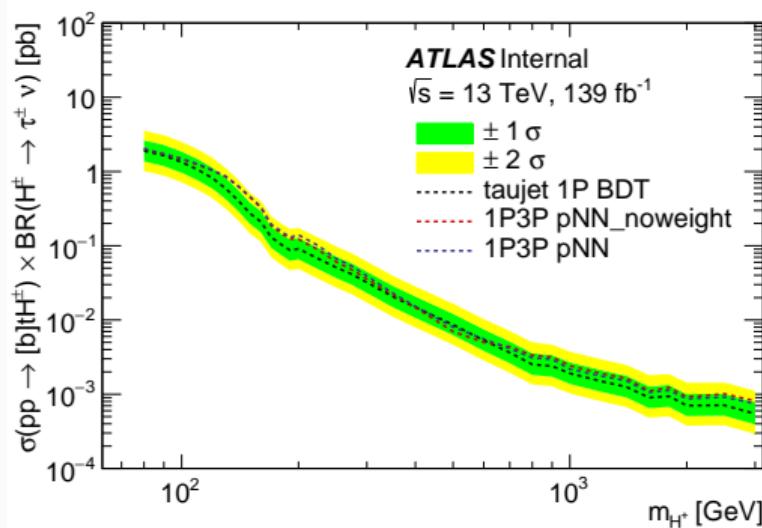
Boosted Decision Tree vs Parameterized Neural Network



Boosted Decision Tree vs Parameterized Neural Network



- One model for entire mass range makes analysis less computationally expensive
- PNN can be evaluated on mass points that are not simulated
- These comparisons were done with an optimized BDT and an unoptimized PNN
- PNN chosen as classifier



PNN Input Variable Selection



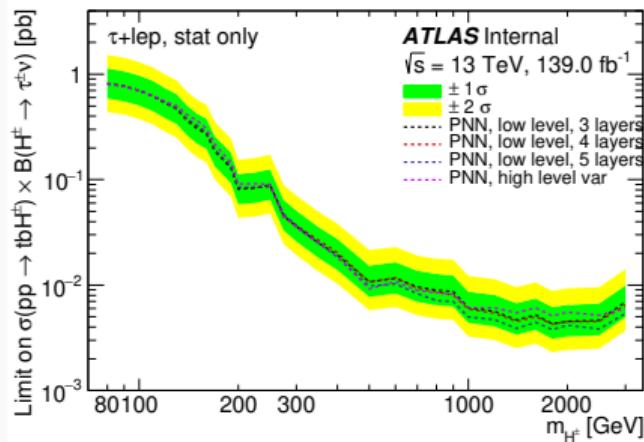
Low Level Input Variables

p_T^τ	η^τ	ϕ^τ
p_T^ℓ	η^ℓ	ϕ^ℓ
p_T^{b-jet}	η^{b-jet}	ϕ^{b-jet}
p_T^{jet}	η^{jet}	ϕ^{jet}
E_T^{miss}	ϕ_E^{miss}	p_T^h
Υ	$m_{H^\pm}^{\text{Truth}}$	

High Level Input Variables

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

- Comparison between raw input variables and engineered variables
- Expected limits in the $\tau + \ell$ subchannel was used as figure of merit
- Low level variables best at high m_{H^\pm}



- Performed in the $\tau + \ell$ sub-channel
- Used area under curve (AUC) of scores as figure of merit
 - Averaged over 5 kfolds, standard deviation is taken from kfolds
- Used early stopping for training
 - $\Delta_{min} = 0.00001$ and a patience of 10
 - Best weights were kept
- To speed up hyperparameter optimization (HPO), ran multiple small grids of hyperparameters
 - Scan over activation functions and loss functions
 - Scan over dropout value
 - Scan over activation function
 - Scan over LeakyReLU α
 - Fixed alpha over more widths and depths
 - AUC from 80 GeV to 500 GeV to optimize for low mass

- LeakyReLU activation function has an α parameter
- Slope of negative portion
 - Prevents neurons from “dying” by allowing negative weight values
- Standard relu is where $\alpha = 0$



PNN Hyperparameter Optimization



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

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

Parameter			
width	32	64	128
depth	2	3	4
dropout	0.1		
activation function	softsign	relu	LeakyReLU
batch size	1025		
loss function	binary crossentropy		

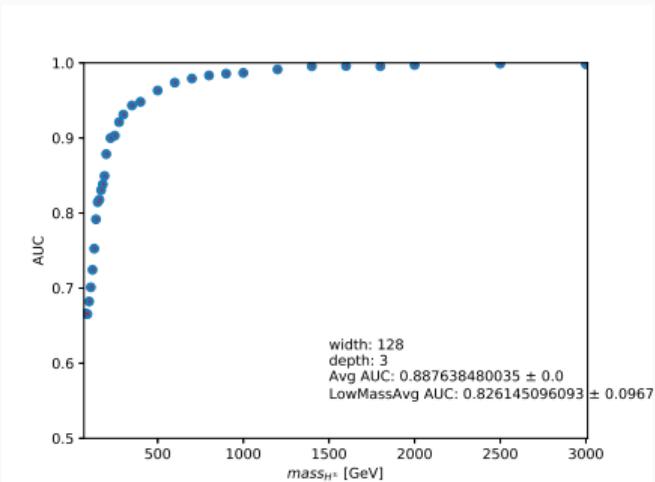
Parameter				
width	32	64	128	
depth	2	3	4	
α	0.01	0.05	0.001	0.005
batch size	1024			
dropout	0.1			
activation function	LeakyReLU			
loss function	binary crossentropy			

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

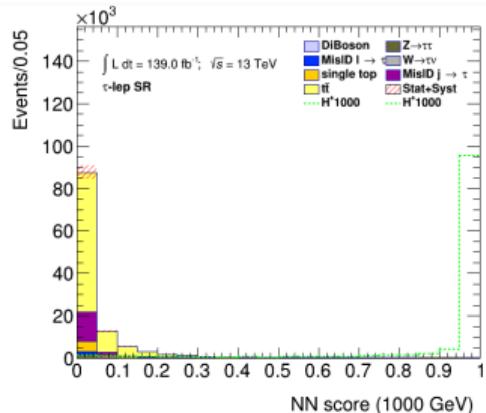
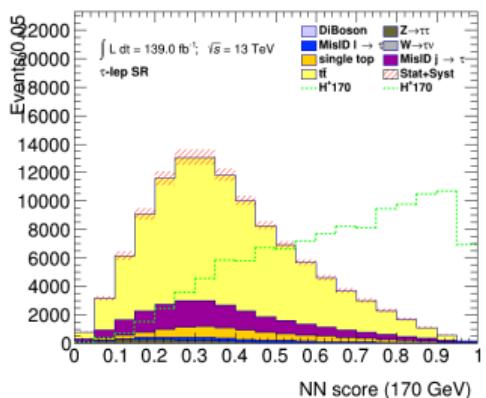
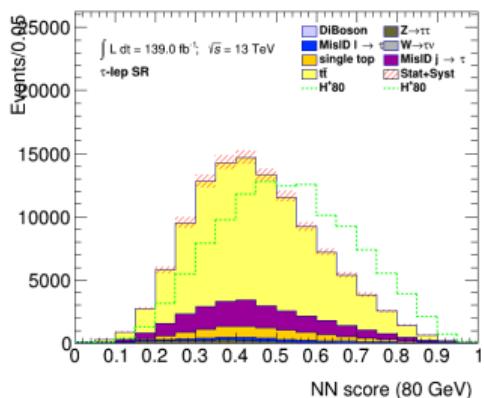
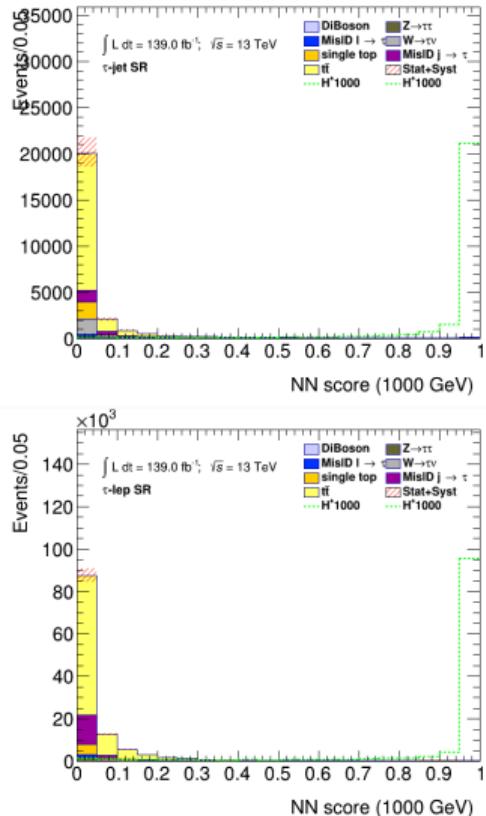
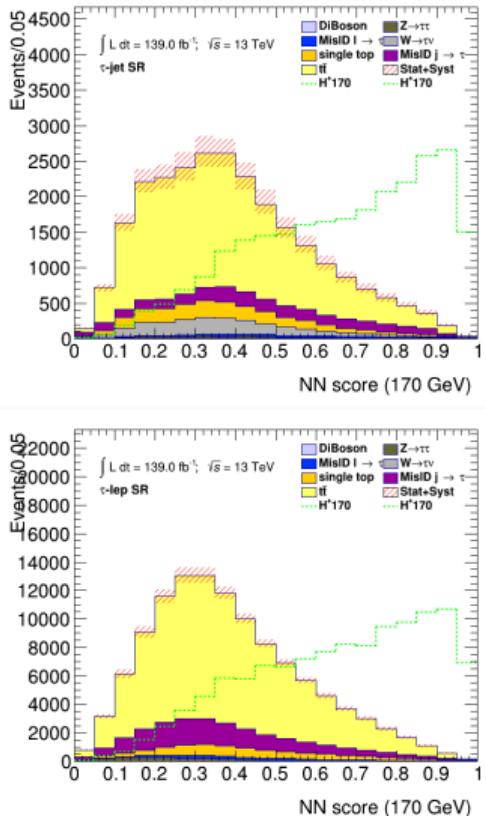
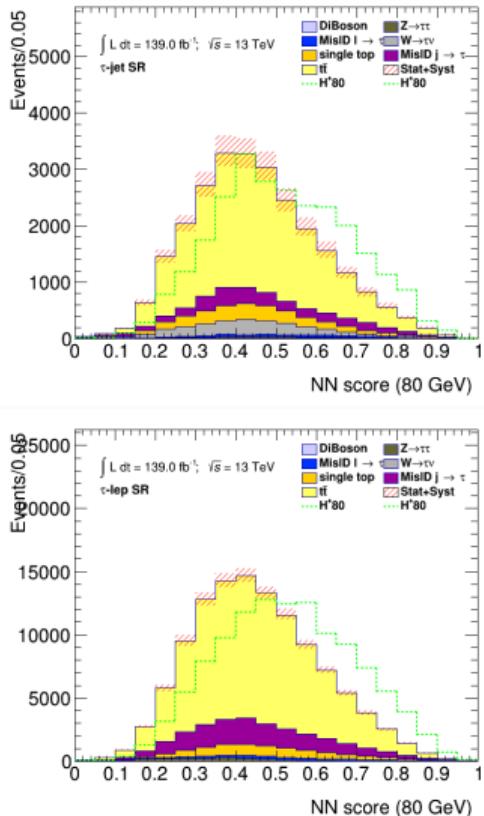
Final PNN Model Performance



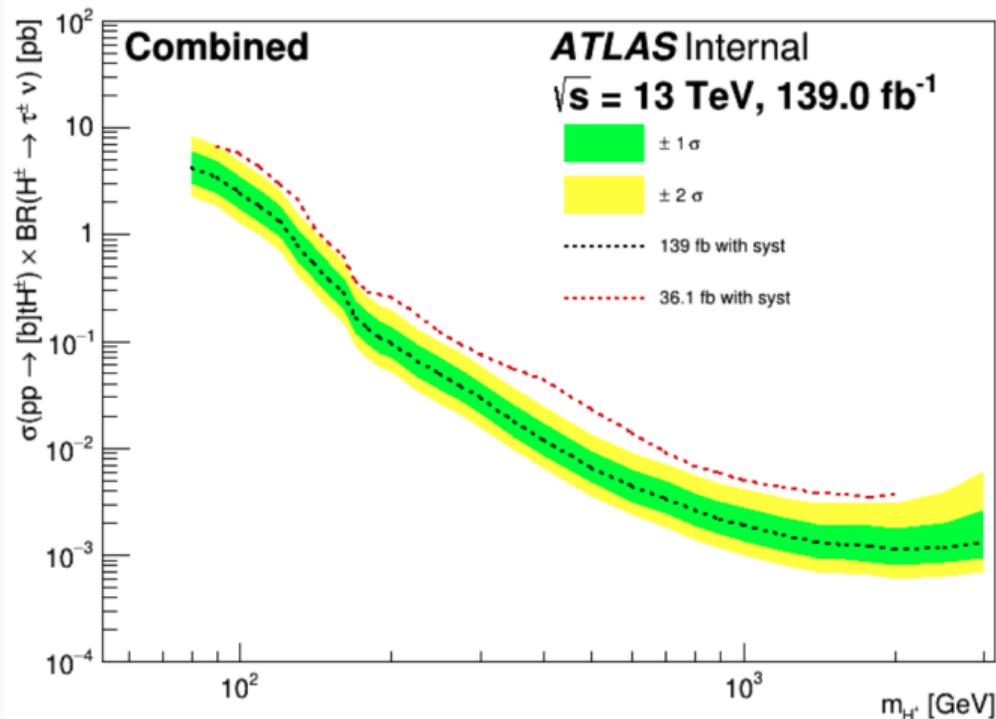
width	depth	Avg	LowMassAvg
128	3	0.8876 ± 0.0000	0.8261 ± 0.0968
128	5	0.8861 ± 0.0000	0.8235 ± 0.1000
128	4	0.8858 ± 0.0000	0.8232 ± 0.0994
128	2	0.8857 ± 0.0000	0.8231 ± 0.1006
64	4	0.8857 ± 0.0002	0.8230 ± 0.0994
64	2	0.8855 ± 0.0004	0.8228 ± 0.0996
64	5	0.8853 ± 0.0005	0.8224 ± 0.0997
64	3	0.8853 ± 0.0011	0.8223 ± 0.0994
256	5	0.8844 ± 0.0002	0.8213 ± 0.1003
256	4	0.8823 ± 0.0000	0.8181 ± 0.1013
32	3	0.8798 ± 0.0012	0.8139 ± 0.1031
32	4	0.8799 ± 0.0009	0.8139 ± 0.1031
32	2	0.8796 ± 0.0004	0.8135 ± 0.1023
32	5	0.8792 ± 0.0011	0.8128 ± 0.1035
256	2	0.8781 ± 0.0002	0.8120 ± 0.1023



PNN Results



$H^\pm \rightarrow \tau\nu$ Expected Limits

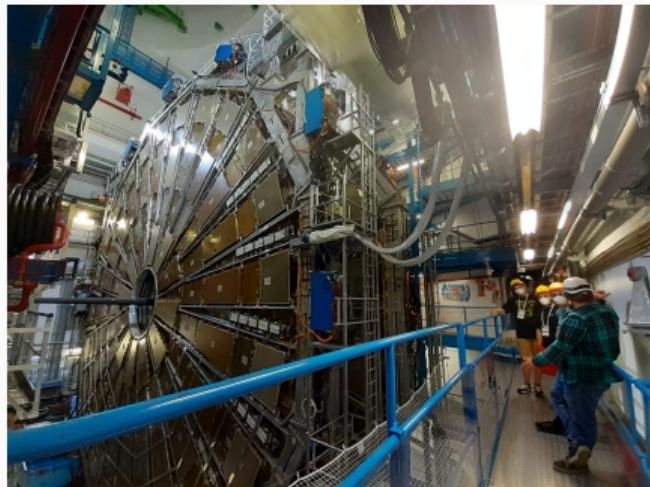


Combined $\tau + \text{jets}$ and $\tau + \ell$ signal regions

Conclusions



- A search for $H^\pm \rightarrow \tau^\pm \nu_\tau$ was improved upon
- Investigated, optimized, and implemented modern machine learning techniques
- New analysis strategy outperforms previous analysis by $> 3\times$
- Analysis is still blinded and work is ongoing towards a publication



Thank You



References i

- [1] D. Hemphill, *The behavior of the primordial universe*, Apr. 2020. [Online]. Available:
https://www.physics.purdue.edu/about/prizes_awards/charlotte_ida_litman_tubis_award/2017_behavior_primordial_universe.html.
- [2] D. Leah, *The standard model: The most successful theory ever*, [Online]. Available: <https://news.fnal.gov/2011/11/the-standard-model-the-most-successful-theory-ever/>.
- [3] J. Ellis, “Higgs Physics,” 117–168. 52 p, Dec. 2013, 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report. DOI: 10.5170/CERN-2015-004.117. arXiv: 1312.5672. [Online]. Available: <https://cds.cern.ch/record/1638469>.

References ii

- [4] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc," *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012.
- [5] G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, and J. P. Silva, "Theory and phenomenology of two-higgs-doublet models," *Physics Reports*, vol. 516, no. 1–2, pp. 1–102, Jul. 2012, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.02.002. [Online]. Available: <http://dx.doi.org/10.1016/j.physrep.2012.02.002>.
- [6] The ATLAS Collaboration, "Search for charged Higgs bosons decaying via $H^\pm \rightarrow \tau^\pm \nu_\tau$ in the $\tau + \text{jets}$ and $\tau + \text{lepton}$ final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS experiment," *JHEP*, vol. 09, p. 139, 2018. DOI: 10.1007/JHEP09(2018)139. arXiv: 1807.07915 [hep-ex].

References iii

- [7] C. T. e. a. Potter, *Handbook of lhc higgs cross sections: 3. higgs properties: Report of the lhc higgs cross section working group*, en, 2013. DOI: 10.5170/CERN-2013-004. [Online]. Available: <http://cds.cern.ch/record/1559921>.
- [8] C. Wanotayaroj, "Search for a Scalar Partner of the Top Quark in the Jets+MET Final State with the ATLAS detector," Presented 25 Oct 2016, Nov. 2016. [Online]. Available: <http://cds.cern.ch/record/2242196>.
- [9] CERN, *Facts and figures about the lhc*, [Online]. Available: <https://home.cern/resources/faqs/facts-and-figures-about-lhc>.
- [10] E. Lopienska, "The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022," , Feb. 2022, General Photo. [Online]. Available: <https://cds.cern.ch/record/2800984>.

References iv

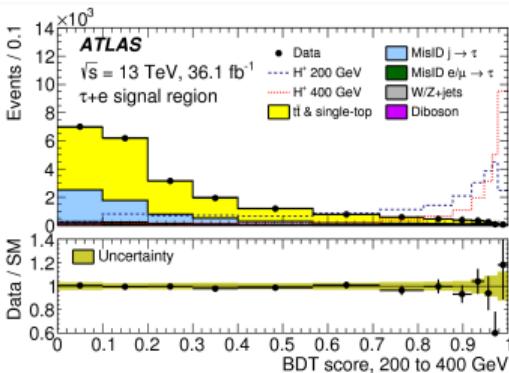
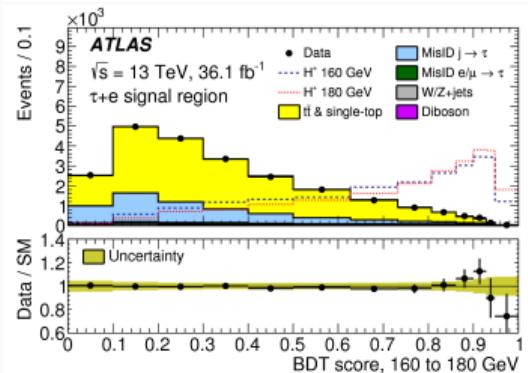
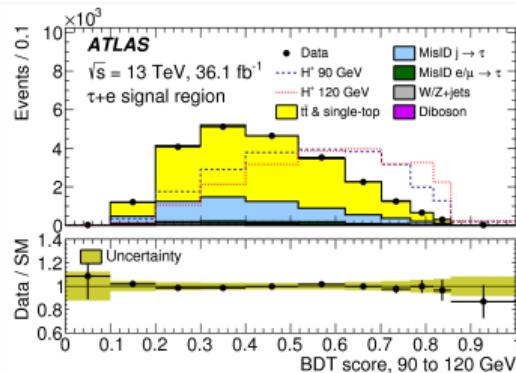
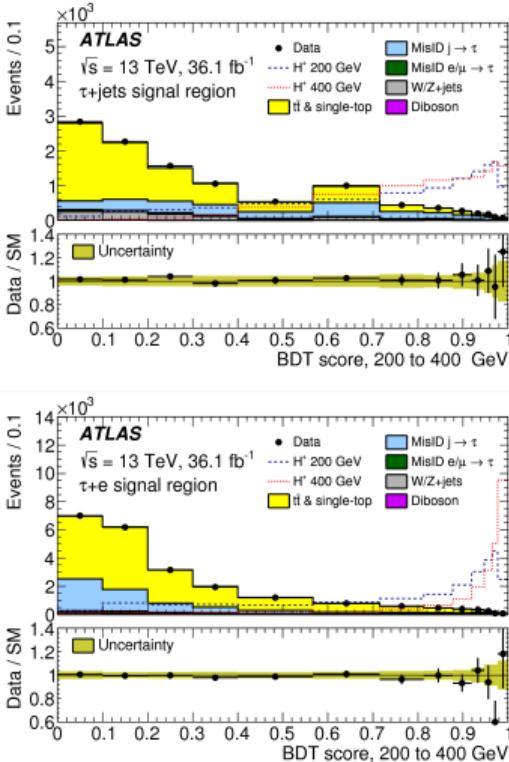
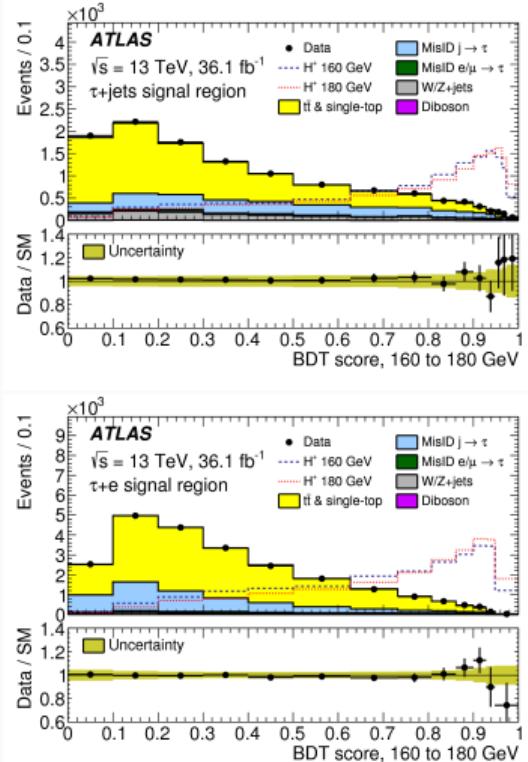
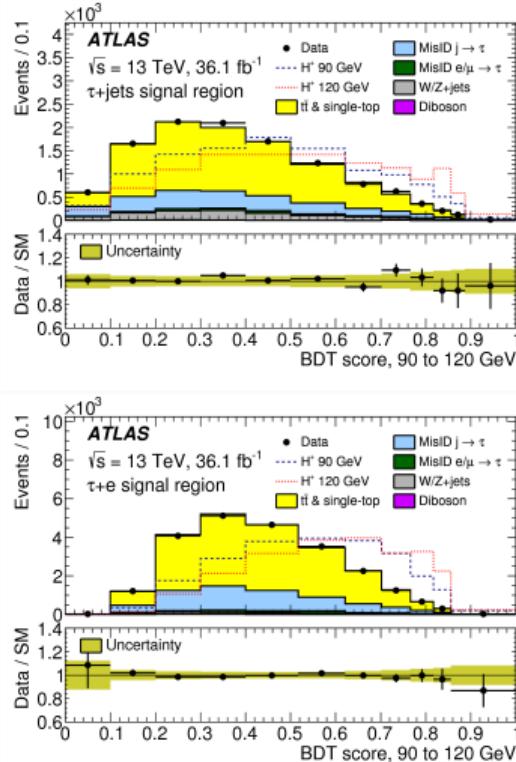
- [11] ATLAS Collaboration, *Atlas schematics*, [Online]. Available: <https://atlas.cern/Resources/Schematics>.
- [12] *Pseudorapidity*. Wikipedia, Feb. 2009. [Online]. Available: <https://en.wikipedia.org/wiki/File:Pseudorapidity2.png>.
- [13] ——, “ATLAS data quality operations and performance for 2015–2018 data-taking,” *Journal of Instrumentation*, vol. 15, no. 04, P04003–P04003, Apr. 2020. DOI: 10.1088/1748-0221/15/04/p04003. [Online]. Available: <https://doi.org/10.1088/1748-0221/15/04/p04003>.
- [14] ——, *ATLAS tile calorimeter: Technical Design Report*, ser. Technical design report. ATLAS. Geneva: CERN, 1996. DOI: 10.17181/CERN.JRBJ.7028. [Online]. Available: <https://cds.cern.ch/record/331062>.

References v

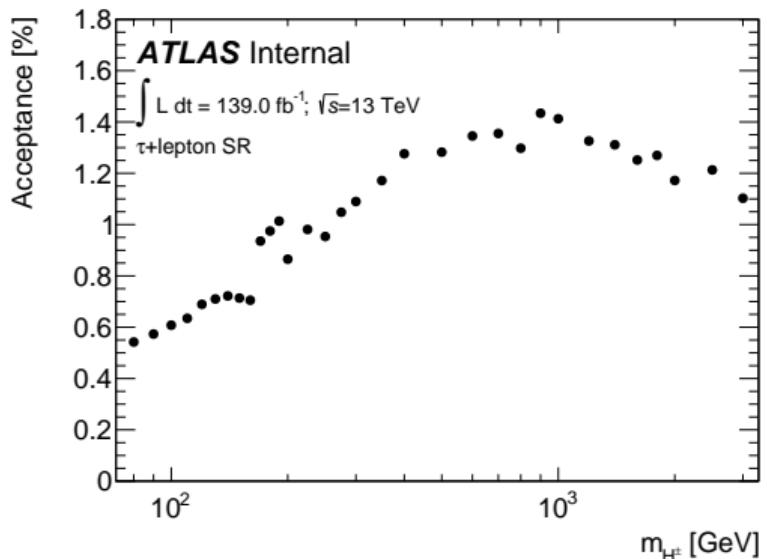
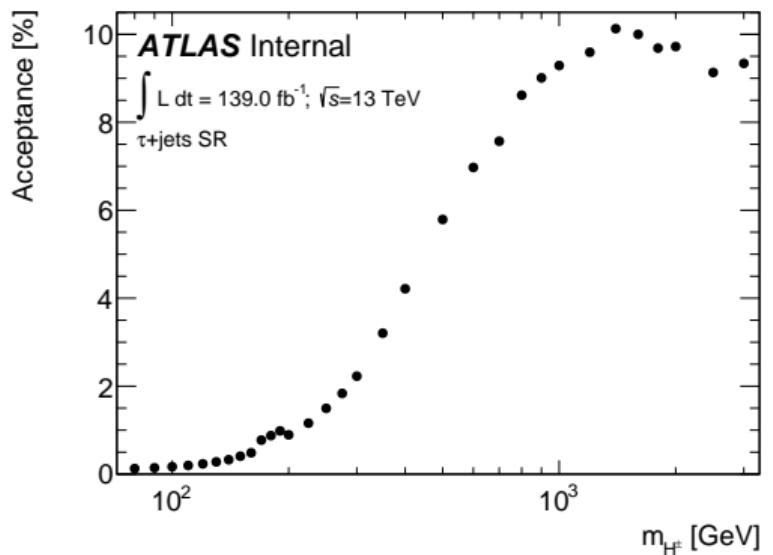
- [15] ——, *Atlas daq approved plots*, [Online]. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ>.
- [16] ——, “Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run,”, Jul. 2017, All figures including auxiliary figures are available at
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013>. [Online]. Available: <http://cds.cern.ch/record/2273281>.
- [17] P. O. Hansson Adrian, “The ATLAS b -jet Trigger,”, Nov. 2011, Comments: 4 pages, 6 figures, conference proceedings for PIC2011. arXiv: 1111.4190. [Online]. Available: <https://cds.cern.ch/record/1397942>.

- [18] A. Ye, *When and why tree-based models (often) outperform neural networks*, Sep. 2020. [Online]. Available:
<https://towardsdatascience.com/when-and-why-tree-based-models-often-outperform-neural-networks-ceba9ecd0fd8>.
- [19] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, “Parameterized neural networks for high-energy physics,” *The European Physical Journal C*, vol. 76, no. 5, p. 235, 2016. DOI: 10.1140/epjc/s10052-016-4099-4.
[Online]. Available: <https://doi.org/10.1140/epjc/s10052-016-4099-4>.

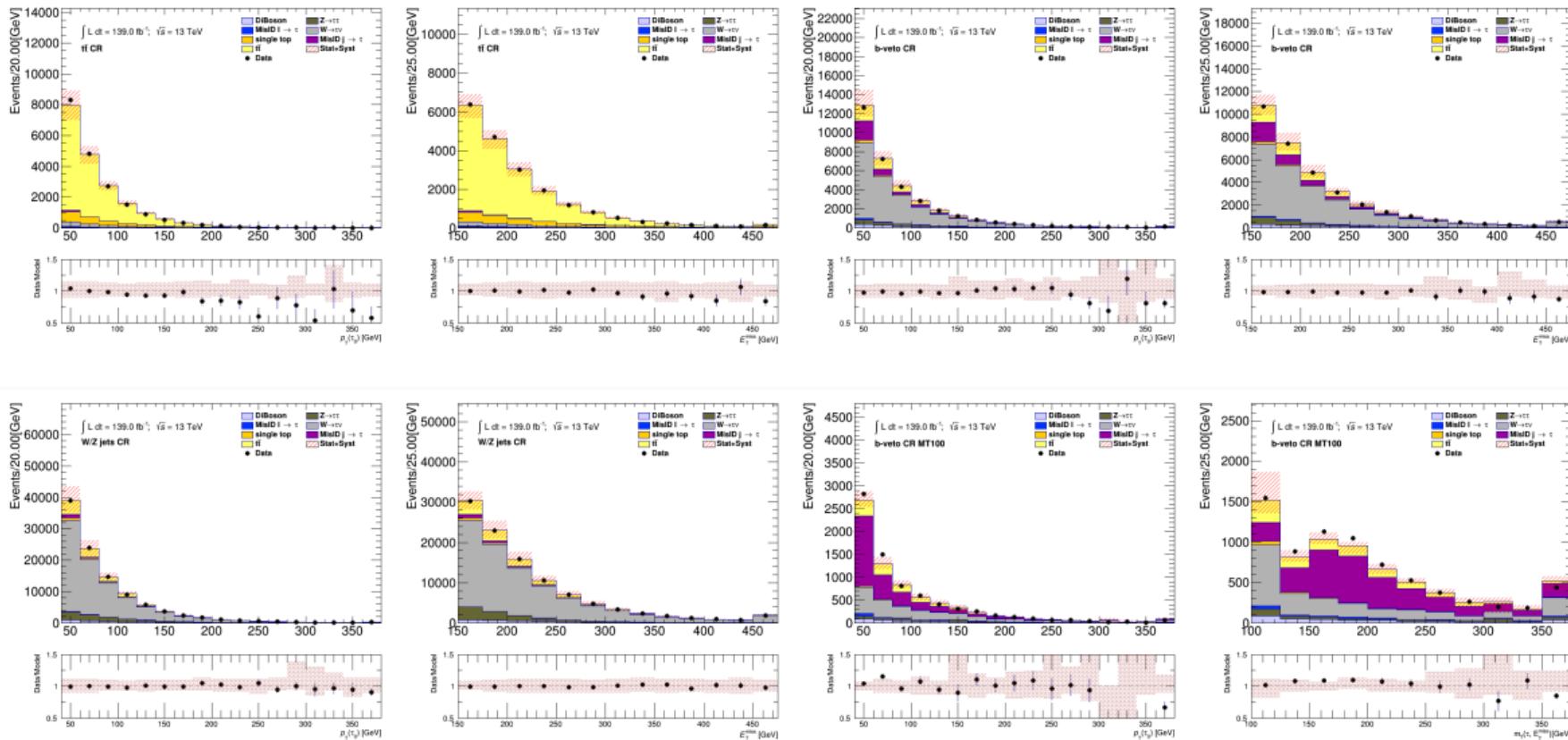
JHEP 09(2018)139 BDT Scores in Signal Regions



Signal Acceptance



$\tau + \text{jets}$ Background Modeling



Background Estimation: $j \rightarrow \tau$ Fakes

- Extract Fake-Factors $FF = \frac{N_{\tau_{had-vis}}^{CR}}{N_{\bar{\tau}_{had-vis}}^{CR}}$ from two orthogonal control regions:

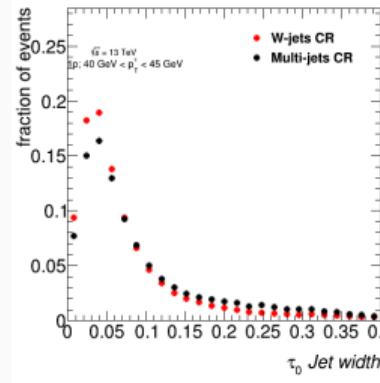
Multi-Jet (gluon enriched)	W+Jets (quark enriched)
E_T^{miss} or Multi-Jet trigger	Single lepton triggers
$\geq 1\tau_{had}, p_T^{\tau} > 30 \text{ GeV}$	$\geq 1\tau_{had}, p_T^{\tau} > 30 \text{ GeV}$
≥ 3 jets	≥ 1 jet
0 b-jets	0 b-jets

- Combine the two Fake-Factors via the template fit method:
 - Find two separate templates for anti- τ in each CR
 - Fit both templates to the shape of the anti- τ in the SR
 - Lowest χ^2 of the fit defines α_{MJ} value and the corresponding error

$$FF_i^{\text{comb}} = \alpha_i^{MJ} FF_i^{MJ} + [1 - \alpha_i^{MJ}] FF_i^{W+jets}$$

- Number of events with fake τ in the signal region is given by:

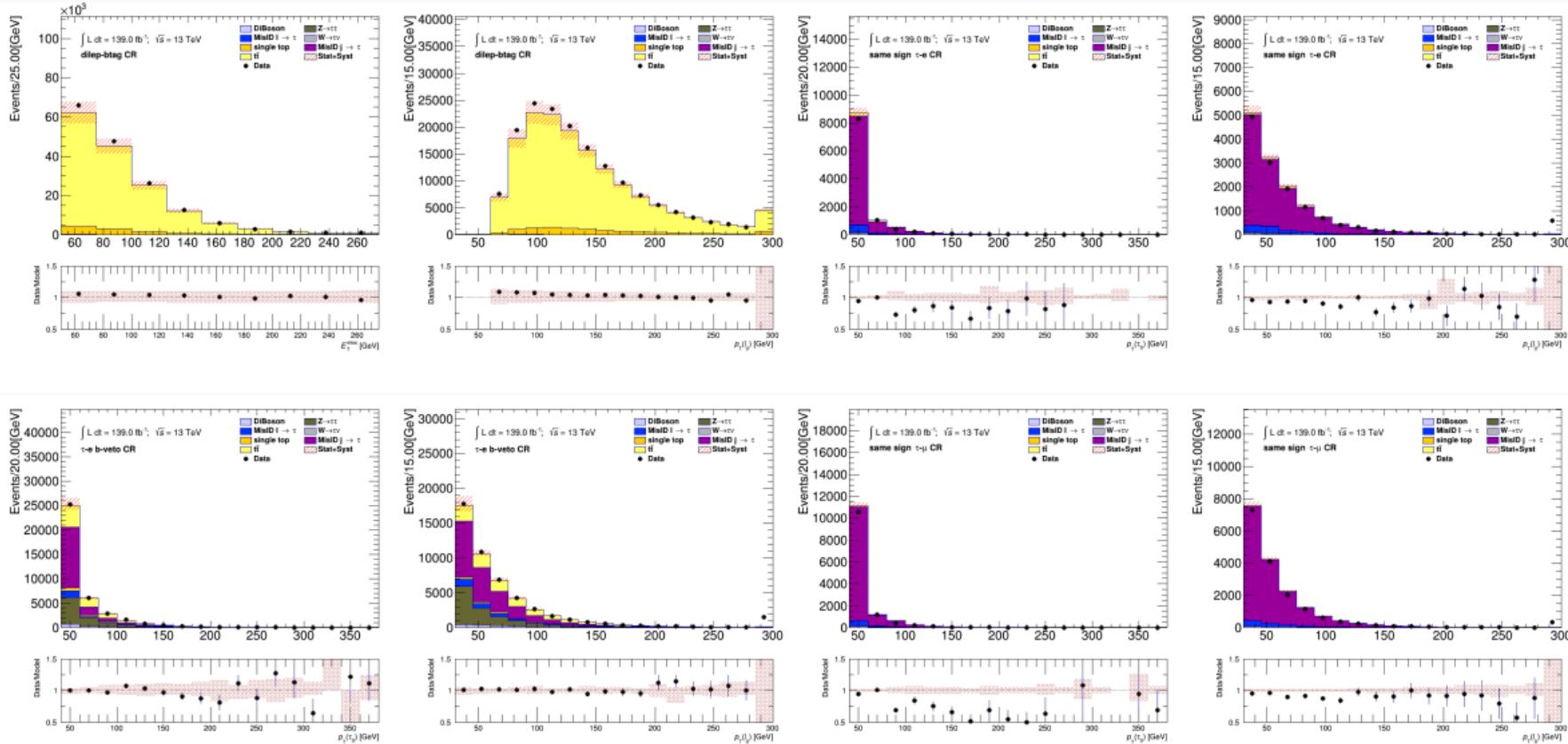
$$N_{\text{fakes}}^{\tau_{had-vis}} = \sum_i N_{\bar{\tau}_{had-vis}}^{SR_i} FF_i$$



- $\bar{\tau}_0$ jet width used in α fitting of 1-prong and 3-prong $\bar{\tau}$

$\bar{\tau} ID$
RNN Score > 0.01
Not loose

$\tau + \ell$ Background Modeling



PNN Hyperparameter Optimization Results

width	depth	80	150	250	500	Avg	LowMassAvg
128	3	0.6661 ± 0.0000	0.8145 ± 0.0000	0.9031 ± 0.0000	0.9633 ± 0.0000	0.8876 ± 0.0000	0.8261 ± 0.0968
128	5	0.6492 ± 0.0000	0.8043 ± 0.0000	0.9078 ± 0.0000	0.9628 ± 0.0000	0.8861 ± 0.0000	0.8235 ± 0.1000
128	4	0.6593 ± 0.0000	0.8117 ± 0.0000	0.9012 ± 0.0000	0.9638 ± 0.0000	0.8858 ± 0.0000	0.8232 ± 0.0994
128	2	0.6444 ± 0.0000	0.8070 ± 0.0000	0.9075 ± 0.0000	0.9631 ± 0.0000	0.8857 ± 0.0000	0.8231 ± 0.1006
64	4	0.6576 ± 0.0050	0.8080 ± 0.0013	0.9052 ± 0.0045	0.9656 ± 0.0016	0.8857 ± 0.0002	0.8230 ± 0.0994
64	2	0.6528 ± 0.0066	0.8052 ± 0.0023	0.9057 ± 0.0032	0.9651 ± 0.0007	0.8855 ± 0.0004	0.8228 ± 0.0996
64	5	0.6538 ± 0.0050	0.8044 ± 0.0019	0.9058 ± 0.0037	0.9653 ± 0.0014	0.8853 ± 0.0005	0.8224 ± 0.0997
64	3	0.6520 ± 0.0067	0.8051 ± 0.0018	0.9042 ± 0.0044	0.9649 ± 0.0019	0.8853 ± 0.0011	0.8223 ± 0.0994
256	5	0.6536 ± 0.0010	0.8044 ± 0.0033	0.9036 ± 0.0042	0.9644 ± 0.0022	0.8844 ± 0.0002	0.8213 ± 0.1003
256	4	0.6434 ± 0.0000	0.8018 ± 0.0000	0.9017 ± 0.0000	0.9619 ± 0.0000	0.8823 ± 0.0000	0.8181 ± 0.1013
32	3	0.6369 ± 0.0094	0.7950 ± 0.0041	0.8977 ± 0.0032	0.9635 ± 0.0022	0.8798 ± 0.0012	0.8139 ± 0.1031
32	4	0.6384 ± 0.0037	0.7935 ± 0.0033	0.8986 ± 0.0037	0.9636 ± 0.0016	0.8799 ± 0.0009	0.8139 ± 0.1031
32	2	0.6399 ± 0.0058	0.7924 ± 0.0024	0.8983 ± 0.0033	0.9629 ± 0.0023	0.8796 ± 0.0004	0.8135 ± 0.1023
32	5	0.6350 ± 0.0077	0.7931 ± 0.0056	0.8981 ± 0.0022	0.9625 ± 0.0005	0.8792 ± 0.0011	0.8128 ± 0.1035
256	2	0.6320 ± 0.0044	0.7971 ± 0.0000	0.8939 ± 0.0034	0.9587 ± 0.0018	0.8781 ± 0.0002	0.8120 ± 0.1023

Fake-factor method uncertainties

Sources of systematic uncertainties associated with the FF method:

- Statistical uncertainties
- True τ contamination in the anti- τ CR
- α_{MJ} fitting procedure uncertainty
- Tau RNN Identification SF variation
- Heavy flavor jet fraction

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 $\tau_{\text{had-vis}}$ in the anti- $\tau_{\text{had-vis}}$ CR	+3.4% -3.2%	✗	+4% -4.3%	✗
Fake factors: tau RNN Identification SF	2.7%	✓	2.7%	✓
Fake factors: α_{MJ} uncertainty	3.6%	✗	1.9%	✗
Fake factors: heavy flavor jet fraction	6%	✓	5.53%	✓

Sources of Systematic Uncertainty

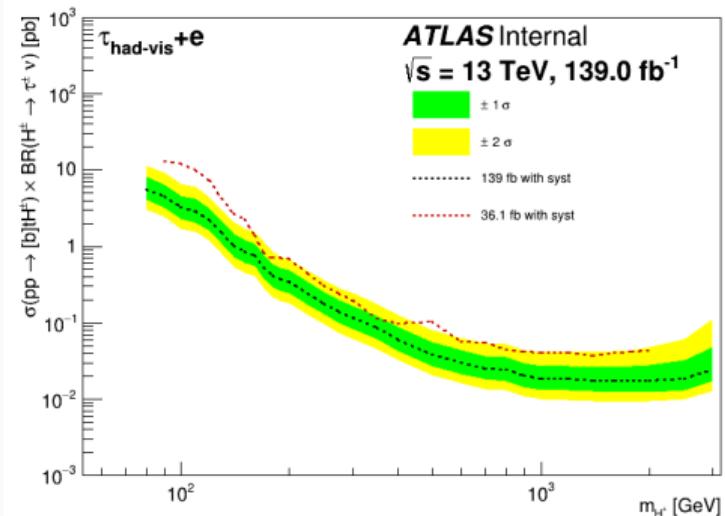
Source	Impact on the expected event yield (%)					
	τ -jets		$\tau+e$		$\tau+\mu$	
	$t\bar{t}$	$H^\pm 200 \text{ GeV}$	$t\bar{t}$	$H^\pm 200 \text{ GeV}$	$t\bar{t}$	$H^\pm 200 \text{ GeV}$
$T_{had-vis}$ reconstruction efficiency	± 1.24	± 1.22	± 1.23	$+1.22$ -1.23	± 1.23	± 1.22
$T_{had-vis}$ -id	± 1.79	± 0.52	± 1.40	± 0.50	± 1.40	± 0.48
$T_{had-vis}$ energy scale	$+2.53$ -2.80	$+2.00$ -1.66	$+1.60$ -1.44	$+1.28$ -1.66	$+1.53$ -1.39	$+1.72$ -1.46
$T_{had-vis}$ energy scale (detector)	$+1.96$ -1.55	$+1.64$ -1.49	$+0.23$ -0.21	$+1.15$ -1.08	$+0.16$ -0.55	$+0.49$ -1.5
$T_{had-vis}$ energy scale (in-situ)	$+144$ -1.43	$+0.22$ -0.74	$+1.17$ -1.20	$+0.74$ -0.63	$+1.14$ -1.15	$+0.54$ -0.37
$T_{had-vis}$ energy scale (model)	$+0.56$ -0.61	-0.06 -0.21	$+0.23$ -0.18	$+1.15$ -1.08	$+0.16$ -0.55	$+0.49$ -1.50
$T_{had-vis}$ energy scale (physics list)	$+1.27$ -1.26	-0.72 -0.65	$+0.74$ -0.65	$+0.67$ -0.25	$+0.72$ -0.63	$+0.83$ -0.60
jet uncertainties	$+7.38$ -8.39	$+6.51$ -9.06	$+3.41$ -3.31	$+4.49$ -2.78	$+3.18$ -3.24	$+3.67$ -2.96
E_T^{miss} soft term scale/resolution	$+1.31$ -1.12	$+1.15$ -1.49	$+0.29$ -0.24	$+0.88$ -0.34	$+0.30$ -0.23	$+0.21$ -0.11
trigger	$+1.23$ -1.61	± 0.03	0	0	$+0.55$ -0.56	± 56
e-id	0	0	± 0.71	± 0.73	0	0
μ -id/reconstruction/isolation	0	0	0 -0.01	0 -0.11	$+0.97$ -1.40	$+1.00$ -2.94
μ MS	0	0	0	0	$+0.09$ -0.12	$+0.40$ -0.34

Expected Yields

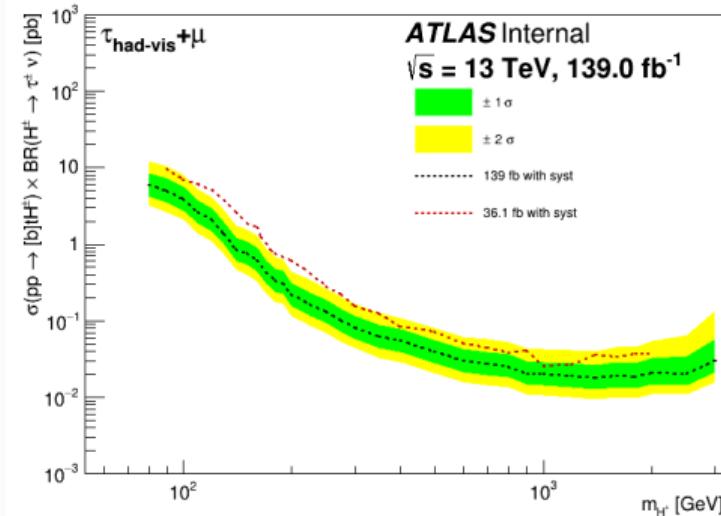
- Analysis is still blinded
 - Ongoing work with the internal ATLAS review process
- Expected number of events from each background source and two arbitrary m_{H^\pm} points
- PNN is used to separate signal from background

Sample	Event yields $+j$ -jets		Event yields $\tau + e$		Event yields $\tau + \mu$				
$t\bar{t}$	18443.27	\pm 48.35	+1545.67 -1697.11	43813.50	\pm 76.85	+1749.82 -1833.87	44486.48	\pm 75.33	+1811.78 -1907.08
Single-top-quark	2284.34	\pm 17.39	+184.69 -207.49	3260.52	\pm 20.81	+124.27 -134.66	3873.35	\pm 22.06	+158.03 -165.92
$W \rightarrow \tau\nu$	1979.17	\pm 23.63	+179.85 -229.80	2.41	\pm 0.56	+0.22 -2.15	0.07	\pm 0.12	+0.08 -0.16
$Z \rightarrow \tau\tau$	242.12	\pm 5.50	+24.27 -32.88	913.55	\pm 20.42	+64.56 -149.42	845.89	\pm 22.07	+88.71 -111.03
Diboson (WW, WZ, ZZ)	133.76	\pm 4.69	+9.47 -12.61	72.64	\pm 1.52	+5.25 -3.91	80.81	\pm 1.53	+5.40 -6.45
Misidentified $e, \mu \rightarrow \tau_{had-vis}$	328.89	\pm 6.85	+25.60 -34.58	1083.97	\pm 24.33	+41.65 -73.42	1060.30	\pm 15.84	+43.44 -70.69
Misidentified jet $\rightarrow \tau_{had-vis}$	2506.28	\pm 17.39	+130.53 -133.40	8662.43	\pm 37.49	+450.65 -470.53	8426.64	\pm 37.12	+440.04 -459.88
All backgrounds	25917.83	\pm 59.82	+1572.87 -1730.97	57809.03	\pm 93.57	+1812.82 -1846.46	58773.63	\pm 90.99	+1873.75 -1970.06
H^\pm (170 GeV), hMSSM $\tan\beta = 40$	1075.81	\pm 9.12	+82.89 -79.19	598.17	\pm 6.60	+20.84 -22.52	702.21	\pm 6.93	+22.39 -16.12
H^\pm (1000 GeV), hMSSM $\tan\beta = 40$	12910.36	\pm 59.30	+784.57 -720.17	938.90	\pm 13.25	+48.99 -37.72	1024.06	\pm 13.21	+48.42 -57.01

$H^\pm \rightarrow \tau\nu$ Expected Limits

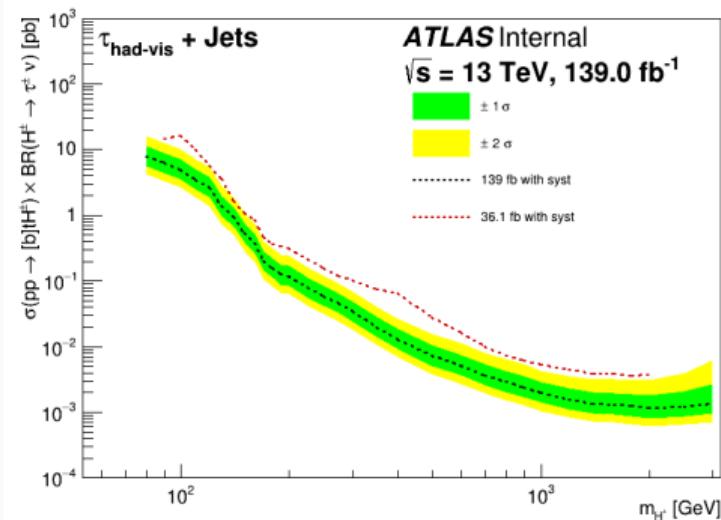


$\tau + e$

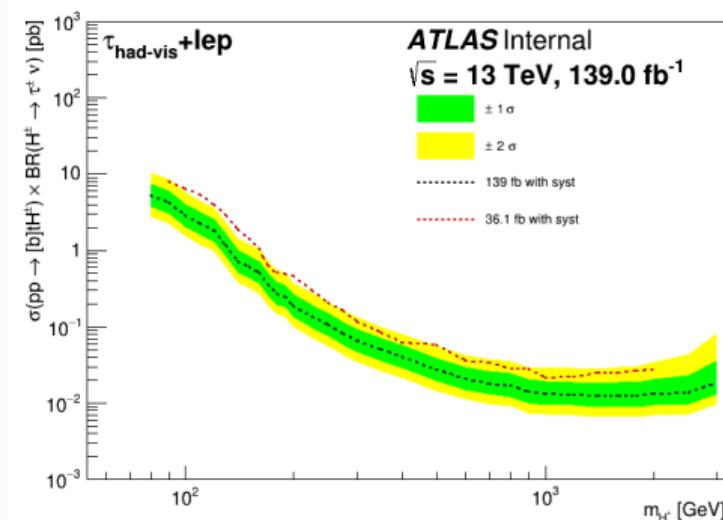


$\tau + \mu$

$H^\pm \rightarrow \tau\nu$ Expected Limits



$\tau + jets$



$\tau + \ell$