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

Dissertation Defense

Elliot Parrish

†Northern Illinois University, USA

October 13, 2022

Contents

Introduction

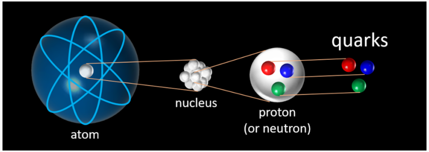
Welcome

- 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!

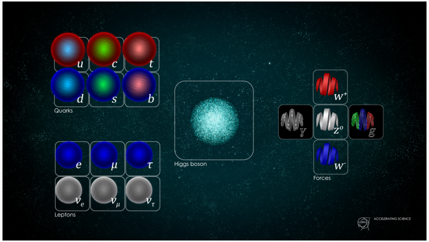
Theory

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
    - Half-integer spin ($s = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, etc.$)
    - Anti-matter is identical to matter except for opposite electromagnetic charge
  - Forces are carried by an exchange of bosons
    - Integer spin ($s = 0, 1, 2, etc.$)
    - Gluon (g) $\rightarrow$ Strong force
    - Photon (γ) $\rightarrow$ Electromagnetism
    - $W^{\pm}, Z^0 \rightarrow$ Weak force

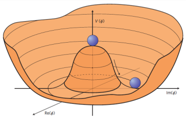


[1]

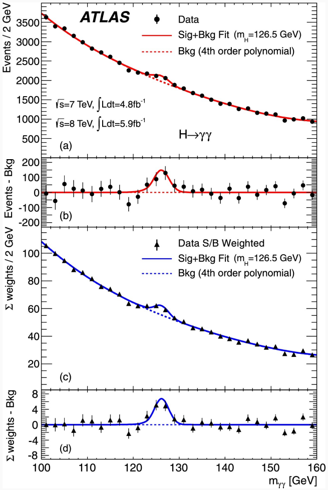


The Higgs Mechanism

- 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



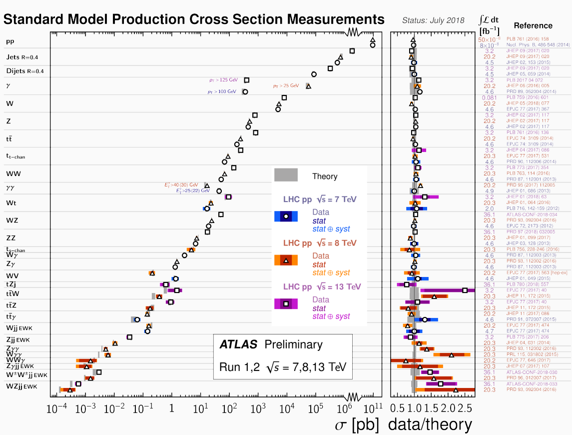
Higgs potential [2]



Higgs discovery [3]

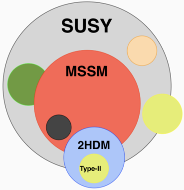
The Standard Model

- 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



Supersymmetry

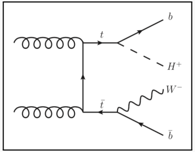
- Hierarchy problem, “unnaturalness”
  - Electroweak scale is ~ 100 GeV
  - Planck scale is $\sim 2.4 \times 10^{18}$ GeV
  - Supersymmetry (SUSY) offers many new particles to occupy the intermediate range
- SUSY proposes a symmetry between fermions and bosons (spin)
  - *Fermion* $\Rightarrow$ *Boson*
  - *Boson* $\Rightarrow$ *Fermion*
- SUSY is a large group of theories
  - Minimal Supersymmetric Standard Model (MSSM) is the smallest SUSY extension to the SM
  - 2 Higgs Doublet Models have two complex doublet scalar fields [4]
    - Two relevant free parameters, $\tan\beta$ and $m_{H^\pm}$
    - $\tan\beta$ is the ratio of the vacuum expectation values of $H^\pm$



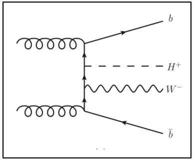
light neutral scalar	h^0
heavy neutral scalar	H^0
neutral pseudoscalar	A^0
two charged scalars	$H^\pm$

Charged Higgs Bosons

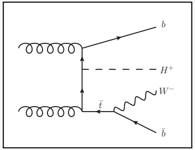
- At the LHC, theoretical production mode of $m_{H^\pm}$ is mainly in top-quark decays or in association with a top-quark (t)
  - $H^\pm$ production mode is dependent on $m_{H^\pm}$



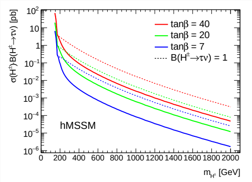
$m_{H^\pm} < m_t$



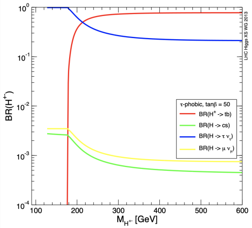
$m_{H^\pm} \simeq m_t$



$m_{H^\pm} > m_t$



[5]

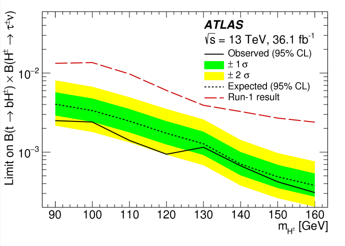


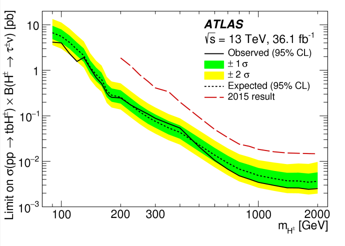
[6]

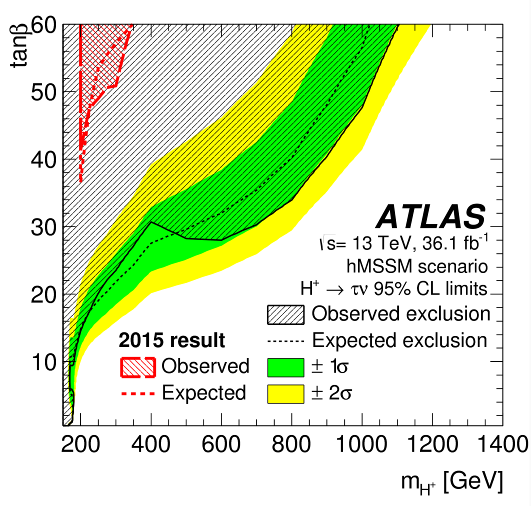
- $H^\pm \to \tau^\pm \nu_\tau$ decay channel remains significant for high $\tan\beta$
- Two sub-channels based on the decay mode of associated t

$t \to jets$	$t \to \ell$
Sensitive at high mass due to higher $W \to q\bar{q}$ BR	Sensitive at low mass due to easier triggering

JHEP 09(2018)139 Limits on

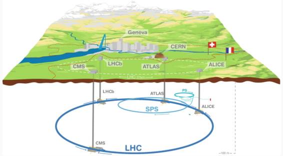




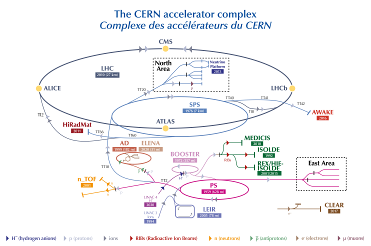


Experimental Apparatus

Large Hadron Collider



- Largest particle collider ever built
- Highest energy particle collider
  - Center of mass energy of 13 TeV
- 27 km circular hadron collider located at CERN outside of Geneva, Switzerland



The ATLAS Detector

-

Inner Detector

-

Calorimeters

-

Tile Calorimeter

-

Muon System

-

Trigger System

-

Simulation

Monte Carlo Simulation

-

Event Reconstruction

Particle Identification

-

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

Analysis Overview

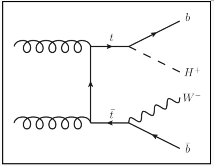
- Search for singly charged $H^{\pm}$ decaying to $\tau^{\pm}\nu_{\tau}$ over a wide mass range
  - Low mass ($m_{H^{\pm}} < m_t$):
  - Intermediate mass* ($m_{H^{\pm}} \simeq m_t$)
  - High mass ($m_{H^{\pm}} > m_t$)
- Dominant backgrounds

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

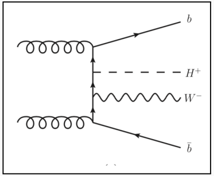
- MVA score is used as the final discriminant

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	$E_T^{miss} > 150$ GeV	$m_T(\tau, E_T^{miss}) > 50$ 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	$E_T^{miss} > 50$ GeV	Opposite sign τ and ℓ

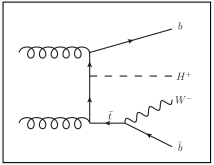
*: First time probed experimentally JHEP 09(2018)139



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



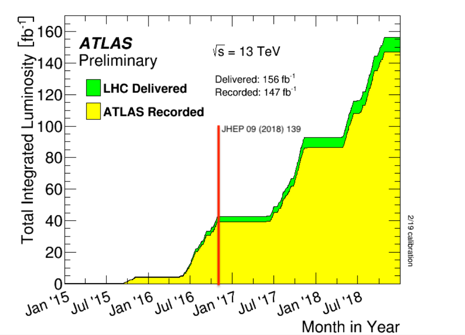
$m_{H^{\pm}} \simeq m_t$



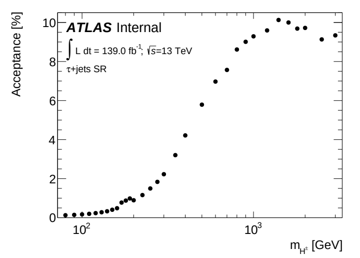
$m_{H^{\pm}} > m_t$

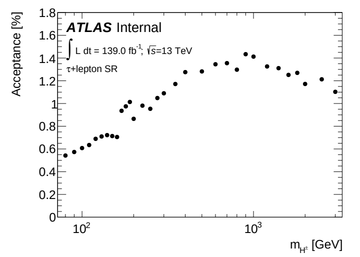
Updates to analysis since last publication

- 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
- Updated derivations
  - RNN τ ID recommendations
  - Updated b-tagging recommendations
  - PFlow jets
  - Latest Combined Performance recommendations



Signal Acceptance





Background Estimation: $j \to \tau$ Fakes

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

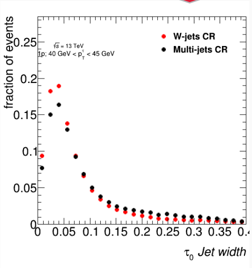
Multi-Jet (gluon enriched)	W+Jets (quark enriched)
E_T^{miss} or Multi-Jet trigger	Single lepton triggers
$\geq 1\tau_{had}$, $p_T^{\tau} > 30$ GeV	$\geq 1\tau_{had}$, $p_T^{\tau} > 30$ GeV
≥ 3 jets	≥ 1 jet
0 b-jets	0 b-jets

- 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

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

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

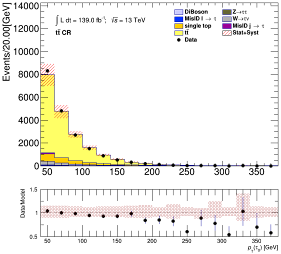
$$N_{fakes}^{\tau_{had-vis}} = \sum_i N^{SR_i}_{\bar{\tau}_{had-vis}} FF_i$$

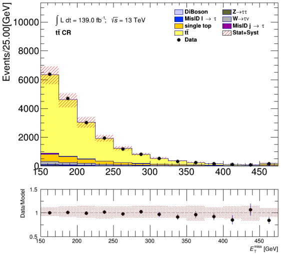


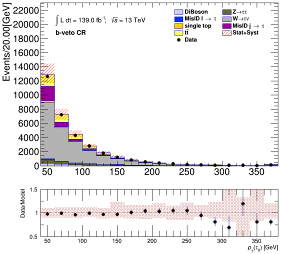
- $\bar{\tau}_0$ jet width used in α fitting of 1-prong and 3-prong $\bar{\tau}$

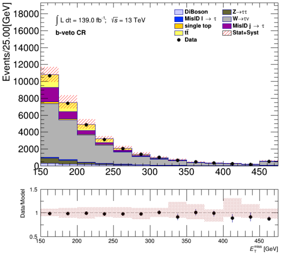
$\bar{\tau}$ ID
RNN Score > 0.01
Not loose

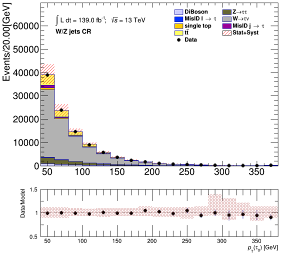
$\tau + jets$ Background Modeling

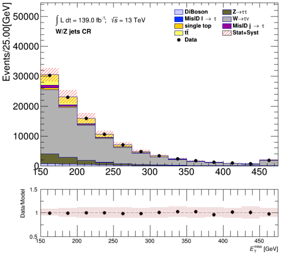


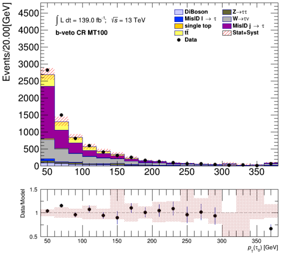


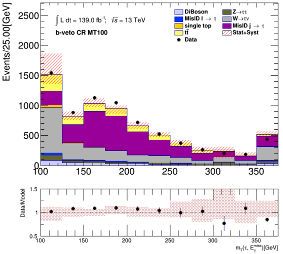




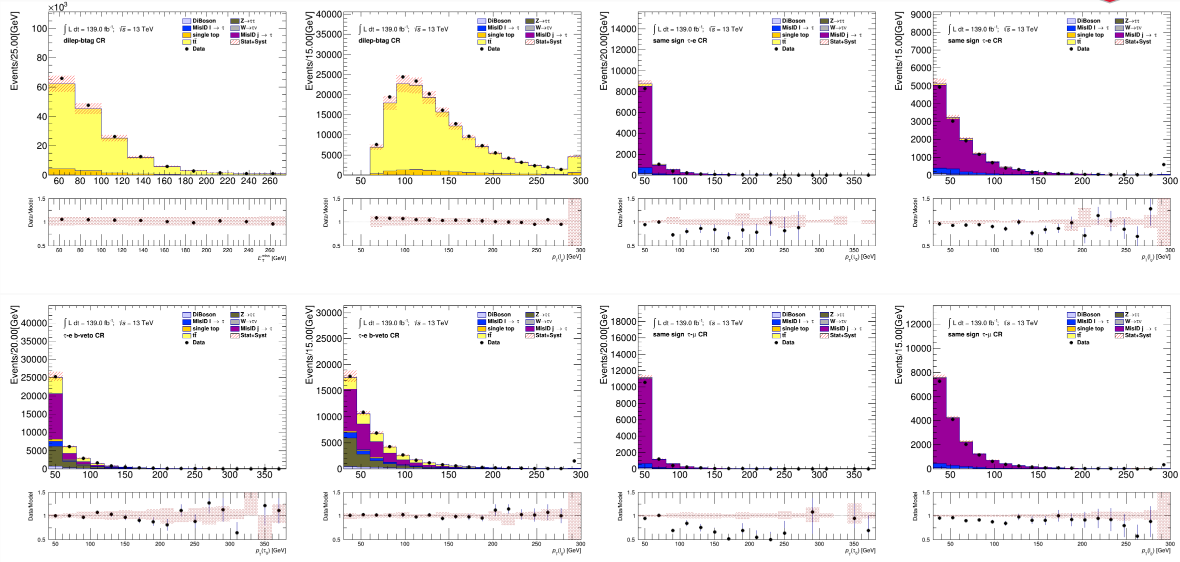






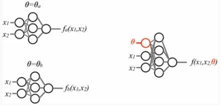


$\tau + \ell$ Background Modeling



Multivariate Analysis Techniques

Parameterized Neural Networks



[8]

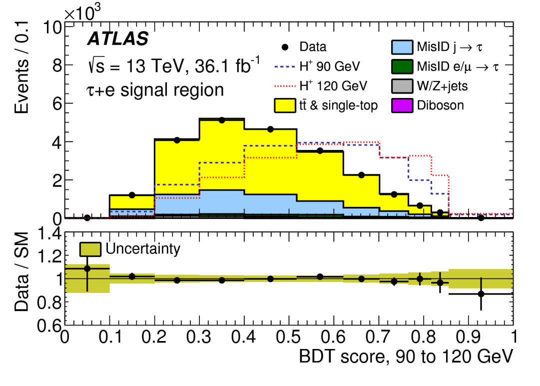
- Background modeling and classifier training kept statistically independent via the k-fold method
  - For this analysis $k = 5$
- Trained using MC for backgrounds with true τ
- Data-driven $j \to \tau$ fakes included for $\geq$ 500 GeV
- Parameterized Neural Networks (PNNs) can be trained and evaluated on an entire mass range
  - Detailed information here: arXiv:1601.07913
- Separate models trained on 1 prong and 3 prong τ
- $\tau + \ell$ channel is trained inclusive for τ+e and $\tau+\mu$

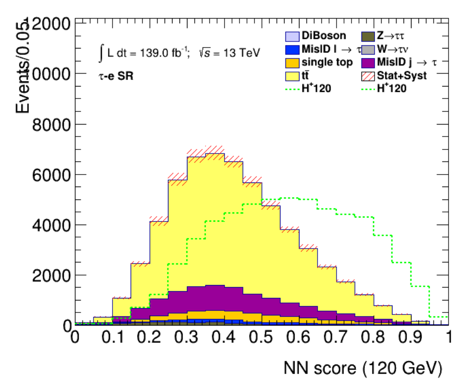
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Background
Partition 1	Evaluation	Train	Train	Train	Train	Fold 1
Partition 2	Train	Evaluation	Train	Train	Train	Fold 2
Partition 3	Train	Train	Evaluation	Train	Train	Fold 3
Partition 4	Train	Train	Train	Evaluation	Train	Fold 4
Partition 5	Train	Train	Train	Train	Evaluation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

$\tau + \ell$ **Input Variables**		
p_T^τ	η^τ	ϕ^τ
p_T^ℓ	η^ℓ	ϕ^ℓ
p_T^{j0}	η^{j0}	ϕ^{j0}
E_T^{miss}	$\phi^{E_T^{miss}}$	p_T^{j1}
$m_{Truth}^{H^\pm}$	Υ^τ	

$\tau +$ *jets* **Input Variables**		
p_T^τ	η^τ	ϕ^τ
p_T^{j0}	η^{j0}	ϕ^{j0}
p_T^{j1}	η^{j1}	ϕ^{j1}
p_T^{j2}	η^{j2}	ϕ^{j2}
E_T^{miss}	$\phi^{E_T^{miss}}$	$m_{Truth}^{H^\pm}$

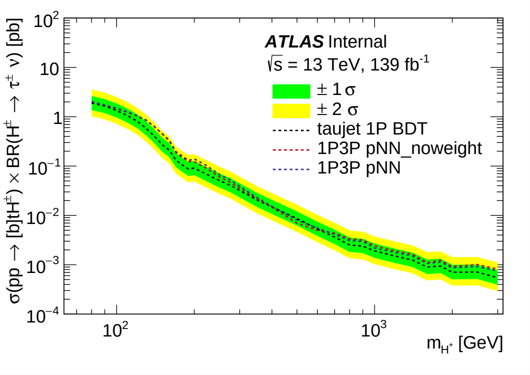
Boosted Decision Tree vs Parameterized Neural Network

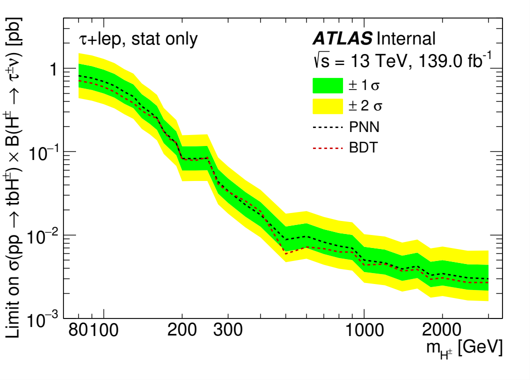




Boosted Decision Tree vs Parameterized Neural Network

- One model for entire mass range makes analysis less computationally expensive
- PNN can be evaluated on mass points that are not simulated
- Limits are preliminary, many fixes and changes have been implemented since





PNN Input Variable Selection

Set A 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}
$E_{\mathrm{T}}^{\mathrm{miss}}$	$\phi^{E_{\mathrm{T}}^{\mathrm{miss}}}$	$p_{\mathrm{T}}^{j_1}$
Υ	$m_{H^{\pm}}^{Truth}$	

(a) Set A of input variables

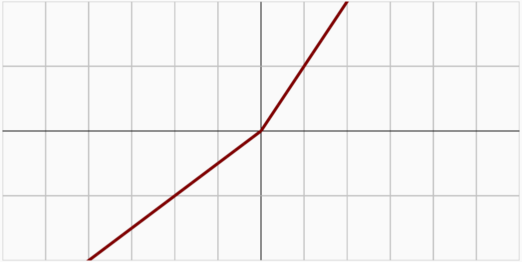
Set B Input Variables
$E_{\mathrm{T}}^{\mathrm{miss}}$
p_{T}^{τ}

PNN Hyperparameter Optimization

- 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 make hyperparameter optimization (HPO) go quicker, ran multiple small grids of hparams
  - 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

PNN Hyperparameter Optimization Search

- 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	$\boxed{\text{binary crossentropy}}$	mean squared error	mean absolute error
width	32		
depth	10		

Parameter			
width	8	16	32
depth	3	5	10
dropout	$\boxed{0.1}$	0.3	
activation function	softsign		
loss function	binary crossentropy		

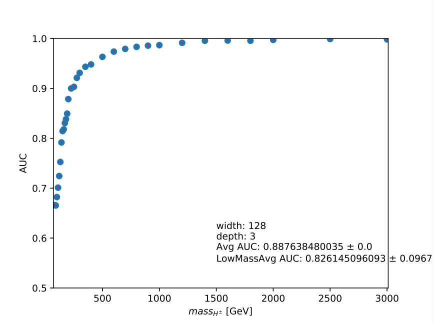
Parameter			
width	32	64	128
depth	2	3	4
dropout	0.1		
activation function	softsign	relu	$\boxed{\text{LeakyReLU}}$
batch size	1025		
loss function	binary crossentropy		

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

Parameter				
width	32	64	$\boxed{128}$	256
depth	2	$\boxed{3}$	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



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
- Smirnov transformation
- Tau RNN Identification SF variation
- Heavy flavor jet fraction

	τ+jets		$\tau+\ell$	
Source of uncertainty	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% -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: Smirnov transform	0%	✓	0%	✓
Fake factors: heavy flavor jet fraction	6%	✓	5.53%	✓

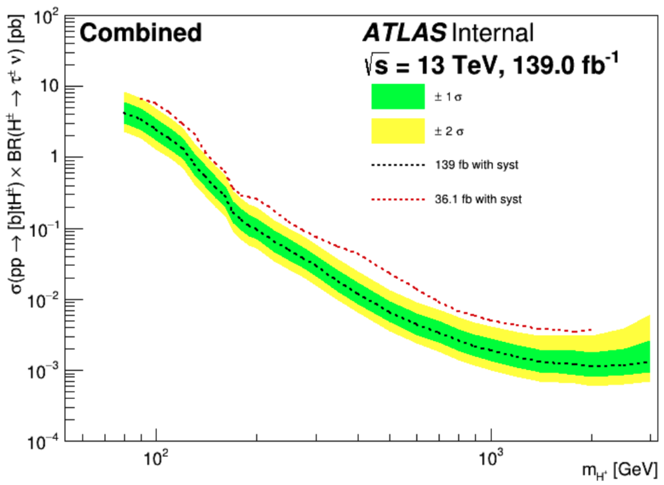
Sources of Systematic Uncertainty

Table with systs

PNN Results

Expected Yields

$H^{\pm} \to \tau\nu$ Expected Limits



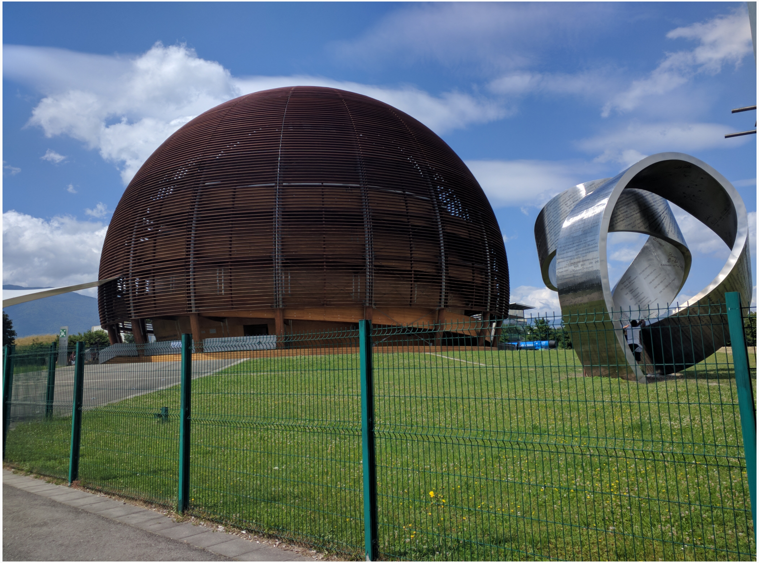
Combined

Conclusion

Conclusions

- New analysis strategy outperforms previous analysis by a factor of XX
- Analysis is still blinded
- Paper will be published soon

Thank You



Bibliography

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] 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.

[3] 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.

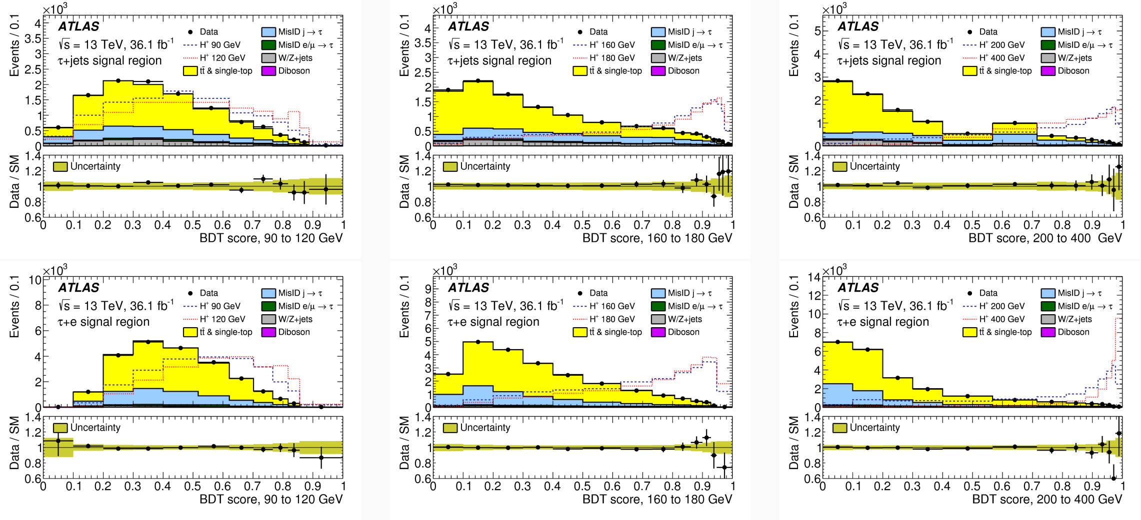
[4] G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, and J. P. Silva, "Theory and phenomenology of two-higgs-doublet models," *Physics Reports*, vol. 516, no. 1–2, pp. 1–102, Jul. 2012, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.02.002. [Online]. Available: http://dx.doi.org/10.1016/j.physrep.2012.02.002.

[5] The ATLAS Collaboration, "Search for charged Higgs bosons decaying via $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ 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*, vol. 09, p. 139, 2018. DOI: 10.1007/JHEP09(2018)139. arXiv: 1807.07915 [hep-ex].

[6] C. T. Potter, H. Mantler, S. Lehti, I. Puljak, H. Kim, R. Klees, D. De Florian, R. K. Ellis, S. Frixione, J. Thompson, J. Huston, S. Y. Choi, D. Rebuzzi, M. Schulze, J. Yu, R. Rietkerk, R. Di Nardo, M. Zaro, R. Harlander, S. Dittmaier, C. Oleari, A. David, F. Maltoni, L. Kashif, B. Di Micco, M. V. Acosta, P. T. E. Taylor, S. O. Moch, J. Olsen, S. Bolognesi, G. Weiglein, M. Ubiali, M. Rauch, F. J. Tackmann, D. Wackeroth, N. V. Tran, Y. Gao, M. Muhlleitner, P. Vanlaer, X. Liu, A. Whitbeck, M. Zanetti, T. Peraro, D. Del Re, G. Degrassi, J. M. Smillie, P. Monni, E. Mirabella, F. Schissler, K. Mawatari, D. Tommasini, P. Torrielli, S. Heinemeyer, P. Govoni, P. Slavich, M. Kramer, G. Sborlini, R. Tanaka, R. Hernandez-Pinto, J. R. Walsh, G. Passarino, C. Zhang, C. Wagner, G. Ferrera, Z. Trocsanyi, C. Mariotti, G. P. Salam, B. Mellado, M. Kadastik, D. Gillberg, T. J. E. Zirke, F. Krauss, A. Rizzi, P. Bolzoni, Q. Li, C. Neu, M. Schonherr, T. Becher, M. Grazzini, P. Musella, M. Carena, M. V. Garzelli, E. A. Bagnaschi, A. Kardos, W. J. Waalewijn, O. Stal, J. Mazzitelli, D. C. Hall,

Backup

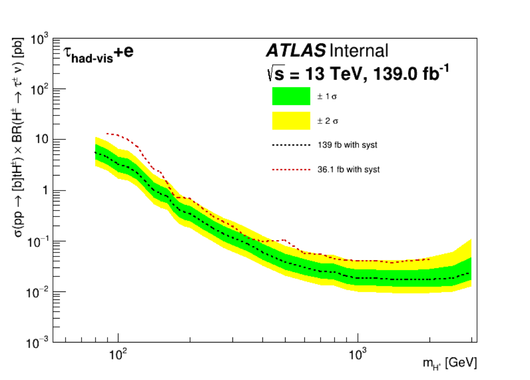
JHEP 09(2018)139 BDT Scores in Signal Regions



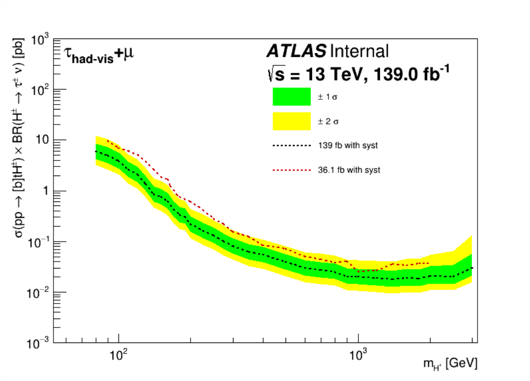
PNN Hyperparameter Optimization Results

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

$H^{\pm} \to \tau\nu$ Expected Limits

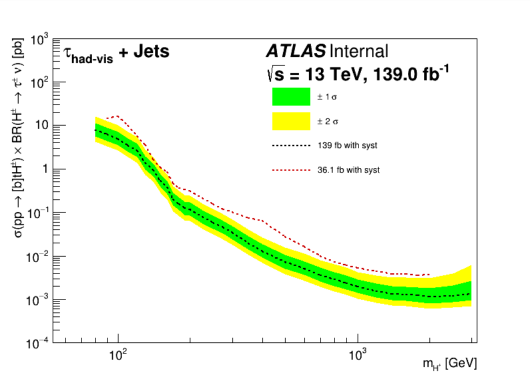


$\tau + e$

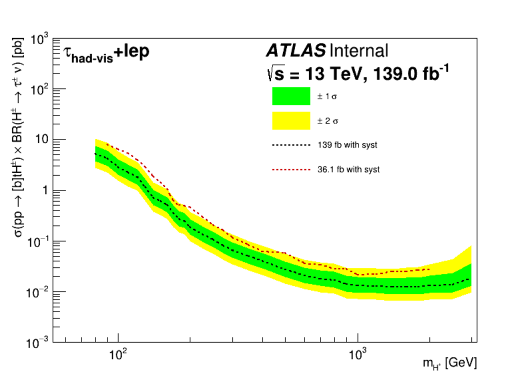


$\tau + \mu$

$H^{\pm} \to \tau\nu$ Expected Limits



$\tau + jets$



$\tau + \ell$