

Northern Illinois University

2022 PhD Progress Review

Elliot Parrish[†] Dhiman Chakraborty[†] Jahred Adelman[†]

†Northern Illinois University, USA

January 20, 2022

Contents



Timeline

Tile Calorimeter

Authorship Task

Data Quality

Maintenance

Test Beams

 $H^\pm o au
u$ Analysis

Theory

Previous Result

Current Analysis

PNN vs BDT

Input Feature Ranking

PNN HPO

Expected Results

Pileup Weights

Continuing Work

Conclusion

Timeline

Timeline to 2022



- Summer 2017: @ NIU
 - Started at NIU
 - Started ATLAS qualification task on the hadronic Tile calorimeter
- **Summer 2018:** @ CERN
 - Qualified as author on ATLAS
 - Start training as Tile Calorimeter
 Data Quality (DQ) Co-Coordinator
 - Tile test beams
 - Machine Learning for High Energy Physics Summer School
 - University of Oxford
 - August 6. 2018 August 12. 2018

- Fall 2018: @ NIU
 - Take over full time as DQ Co-Coordinator
 - Finished collecting Run-2 data
- Summer/Fall 2019: @ CERN
 - Ramped up work on $H^\pm o au
 u$ analysis
 - ATLAS data quality Run-2 paper
 - Reprocessing campaign of Run-2 data starting
 - Moved to CERN long term

Timeline to 2022



- **Spring 2020:** @ CERN
 - Polish National Agency for Academic Exchange: International Scholarship Exchange of PhD Students and Academics
 - February 17, 2020 February 2021. 2020
 - Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland
- **Summer 2020:** @ CERN
 - On detector readout and control electronics maintenance
 - Started $H^{\pm} \rightarrow \tau \nu$ PNN HPO

- Fall 2020: @ CERN
 - Reprocessing monitoring tests
- Spring/Summer 2021: @ CERN
 - Tile ACES (equipment inventory)
 - Trained replacement DQ Co-Coordinator
 - Finalized Run-2 reprocessing preparations
 - APS DPF 2021
- Fall 2021: @ USA
 - Tile test beams
 - Moved from CERN back to US
 - Finished involvement with Tile

Tile Calorimeter

ATLAS Hadronic Tile Calorimeter



- ATLAS is a general purpose particle detector on the Large Hadron Collider at CERN
- The Hadronic Tile calorimeter is the outermost. sub-detector besides the muon system
 - Designed to measure hadron energy by fully absorbing particle showers
 - Highly important in jet reconstruction
 - Scintillating tile with steel absorber
 - Light readout via optical fibers to PMTs



TileCalDiagramATLAS.png TileModuleCrossSection.pn

Tile Calorimeter Calibration Systems



- Cesium
 - Hydraulically moved throughout detector
 - Measures response of scintillator
- Laser
 - Measures response of photomultiplier tubes
 - Can be taken during physics data taking and individual calibration runs
- Charge Injection
 - Four capacitors, two low gain, two high gain
 - Measures response of digitizers and readout electronics
- One centralized database to store all calibration constants

Tile_Calibration_Diagram.png

Centralized ATLAS Conditions Database



- Online and Offline database
- Three main tags within databases
 - UPD1
 - Used to quickly process 10% of events
 - Can only append information
 - UPD4
 - Used after 48 hrs to process full run
 - Can change previous information

- TileCal Robot Authorship Task
 - Web based interface to update conditions database (Tile internal)
 - Previously, had to update all three tags individually
 - Created automated syncing of tags

TileRobotTrippleUpdateScreenShot.png

Tile Data Quality Co-Coordinator



- Took over from Puja Saha at the end of Summer 2018 until Fall 2021
- Workload was dependent on ATLAS data taking schedule
- During data taking
 - 48 hours to investigate issues and take action before full processing of runs
 - Coordinate two shifters (Data Quality Validator and Data Quality Leader)
 - Personally sign off every run and tag with appropriate defects

- During Long Shutdown 2
 - Coordinate two shifters (Data Quality Validator and Data Quality Leader)
 - Fill in when needed
 - Reconstruction codebase was restructured to support multithreading
 - Signed off many tests
 - Finished Summer 2021
 - Reprocessing of Run-2 data
 - Revised Tile channel statuses for 2018, applied changes from previous DQ Co-Coord
 - Finalized Fall 2021

Tile Data Quality Run-2 Performance



- As DQ Co-Coordinator I was responsible for channel and cell status
 - (Un)Masking based on performance
 - Various status flags used in reconstruction

masked_cells_timeline_2018.png

- Tile boasted 100% Data Quality Efficiency for 2018
 - For all of Run-2, we were 99.65% efficient
- Paper detailing ATLAS Run-2 data quality performance published in Journal of Instrumentation
 - DOI: 10.1088/1748-0221/15/04/P04003

DataQualityLosesRun2.png

Tile Maintenance



- At the start of the pandemic, many experts had to leave the CERN area
 - I voluntered to help with transition of new CIS/Maintenance technicians
 - During Long Shutdown 2, Tile performed maintenance
 - Repaired/replaced electronics, PMTs, added cooling isolation valves, etc.
 - I helped install the demonstrator module that is a proof of concept for future Tile upgrades
 - Will remain inside of Tile for Run-3 data collection

OnLBA14.jpg TileElectronicsJalal TileMaintenanceTeamJune2020.png

ACES Database Update



- ATLAS Central Equipment System (ACES)
 - A central database for all racks, cables, detector parts, and general equipment
- Needed to update with latest information since Long Shutdown 2 was coming to an end
 - Myself and Michaela Mlynáriková were tasked with inputting new information
- Finished in Fall 2021

Hard_At_Work.jpg

 USA15 ACES Y.04-16.A1.jp\g04-16.A1_A

Tile Test Beams



- Participated in two test beams for Tile
 - Summer 2018
 - Fall 2021
- In both, collected data for various electronics upgrades
 - Set beam energy and beam type from an offshoot of the Super Proton Synchrotron (SPS)
 - Moved Tile module to collect spatial response
- Helped setup test beam site in 2021
 - Tracing down cables, connected various

TileTestBeamSetup_Michaela_Wili

$H^{\pm} ightarrow au u$ Analysis

Search for $H^{\pm} \to \tau \nu$ using the ATLAS experiment



- Many extensions to the Higgs sector imply the existence of charged scalars (2HDM, NMSMM, Triplet, etc.)
 - $H^{\pm} \rightarrow \tau^{\pm} \nu_{\tau}$ remains significant for high $\tan \beta$
 - $an\!eta$ defined as the ratio of the vacuum expectation values of the two doublets
- At the LHC, theoretical production mode of H^{\pm} is mainly in top-quark decays or in association with a top-quark (t)
 - H^{\pm} production mode is dependent on $m_{H^{\pm}}$
 - Analysis is sensitive for low mass $(m_{H^\pm} < m_t)$, intermediate mass $(m_{H^\pm} \simeq m_t)$, and high mass $(m_{H^\pm} > m_t)$
- ullet Two sub-channels based on the decay mode of associated t

Sensitive at high mass due to higher $W o q\bar{q}$ BR | Sensitive at low mass due to single lepton triggers

t o jets $t o \ell$

HPlus_taunu_tanB.png

Charged_Higgs_BR.png

Analysis Overview



- Search for singly charged H^{\pm} decaying to $\tau_{had} + \nu$ 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 $ au$	Backgrounds w/ fake $ au$
$tar{t}$ estimated with MC	Fake $j ightarrow au$ estimated with data driven fake factor method
V+jets estimated with MC	Fake $\ell ightarrow au$ estimated with MC, validated on $Z ightarrow ee$
VV estimated with MC	

MVA score is used as the final discriminant.

Sub-Channel Sub-Channel										
$\tau + jets$ SR	E _T ^{miss} Trigger		$0 \ \ell \ (e \ or \ \mu)$ $p_T^{\ell} > 20 \ GeV$		≥ 1 b-jets $p_T^{b-jet} > 25$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}} > 150 \; \mathrm{GeV}$	$m_T(au, E_T^{miss}) > 50 \mathrm{GeV}$			
$\tau + \ell \operatorname{SR}$	Single Lepton Trigger		$1 \ell \text{ (e or } \mu\text{)}$ $p_T^{\ell} > 30 \text{ GeV}$		≥ 1 b-jets $p_T^{b-jet} > 25$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}} > 50 \; \mathrm{GeV}$	Opposite sign $ au$ and ℓ			

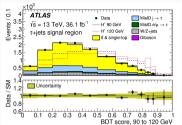


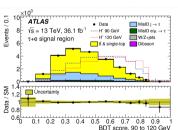


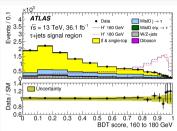


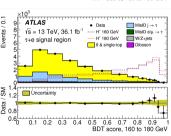
JHEP 09(2018)139 BDT Scores in Signal Regions

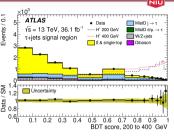


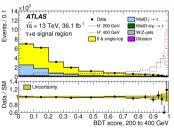






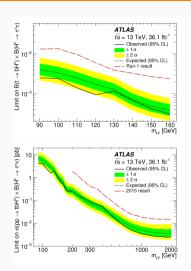


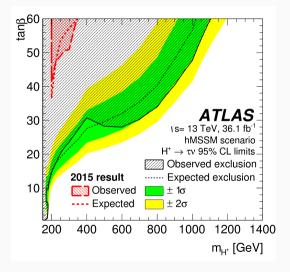




JHEP 09(2018)139 Limits







Updates to analysis since last publication



- Signal mass range extended
 - Previous: $90 \le m_{H^{\pm}} \le 2000 \text{ GeV}$
 - Current: $80 \le m_{H^{\pm}} \le 3000 \text{ GeV}$
- Signal filtering applied in order to effectively increase the statistics in the signal regions
- New analysis framework centered around using modern Machine Learning tools
- Investigating new multivariate analysis techniques
- Updated derivations
 - RNN τ ID recommendations
 - Updated b-tagging recommendations

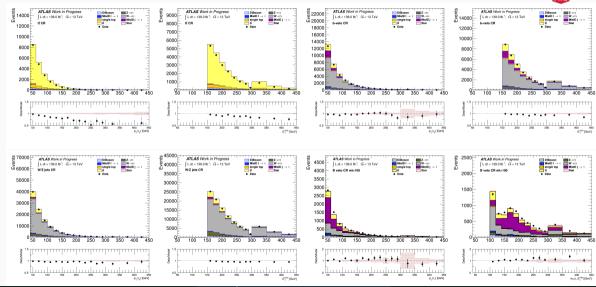
- PFlow jets
- atest Combined Performance recommendations

intlumivstimeRun2.png

Elliot Parrish (NIU)

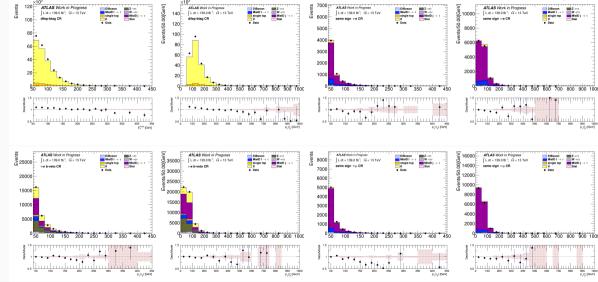
$\tau + jets$ Background Modeling





$\tau + \ell$ Background Modeling





Parameterized Neural Networks



- Parameterized Neural Networks (PNNs) can be trained and evaluated on an entire mass range
 - Detailed information here: arXiv:1601.07913
- Trained using Tensorflow backend for Keras
- ullet m_{H^\pm} is used as an input feature of the model
- Allows one model to be used across entire $m_{H^{\pm}}$ mass range
- ullet Separate models trained on 1 prong and 3 prong au
- $\tau + \ell$ channel is trained inclusive for $\tau + e$ and $\tau + \mu$
- Evaluated on single mass points

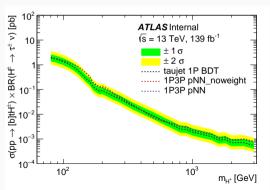
NN_Diagram.png

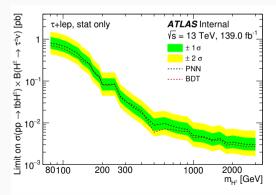
 $H^{\pm} \rightarrow \tau \nu$ Analysis PNN vs BDT

Parameterized Neural Network vs Boosted Decision Tree



- 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





2022 PhD Progress Review

Feature Ranking

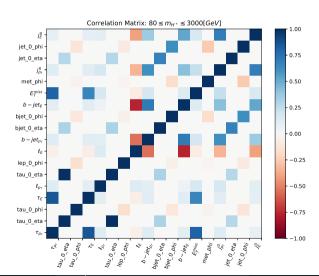


- Linear correlations between input features
 - Neural Networks often do not work very well with linearly correlated variables
- Produced features rankings via Local Interpretable Model-agnostic Explanations (LIME)
 - LIME explains an instance (X) by sampling nearby points with classifier, weighting by distance from X
 - LIME then learns a linear model which explains the decision locally (dashed line)

limeToyExplanation.png

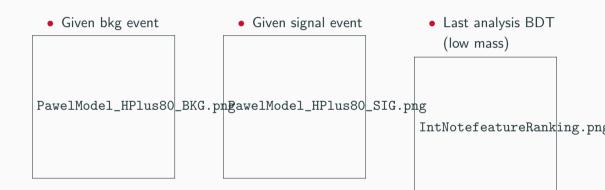
Input Feature Linear Correlations





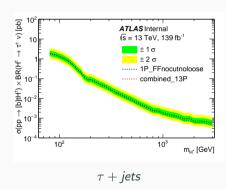
Comparing to Previous Analysis



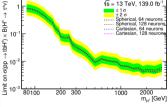


$H^{\pm} \rightarrow \tau \nu$ Expected Limits PNN Studies





→ t[±]v) [pb] τ+lep, stat only ATLAS Internal √s = 13 TeV. 139.0 fb ±1σ ±2σ ... PNN, low level, 3 layers ... PNN, low level, 4 layers ... PNN, low level, 5 layers ... PNN, high level var - †tbH^t) × B(H^t -Limit on o(pp – 80100 200 300 1000 2000 m_H [GeV] ±4) [bb] ATLAS Internal √s = 13 TeV, 139.0 fb ±1σ ±2σ ... Spherical, 64 neurons ... Spherical, 128 neurons ... Cartesian, 64 neurons ... Cartesian, 128 neurone 10



$$\tau + \ell$$

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



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			

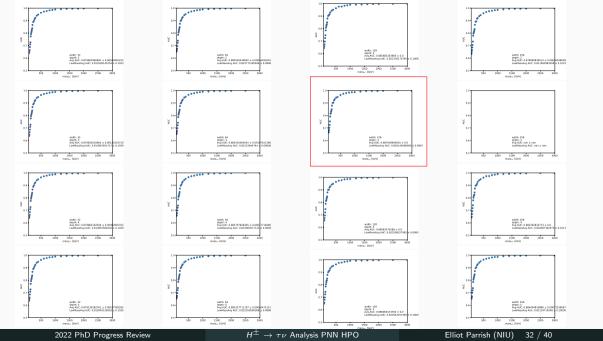
PNN Hyperparameter Optimization Search



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



Elliot Parrish (NIU)



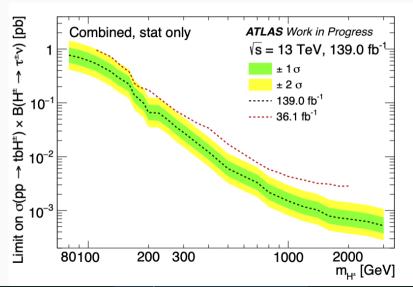
PNN Hyperparameter Optimization Results



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

Preliminary $H^{\pm} \rightarrow \tau \nu$ Expected Limits





Negative Monte Carlo Weights



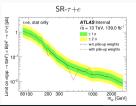
- Limits seem to have spikes and dips at and above the top mass
- Noticed some weird values for our pileup weights
 - NOMINAL_pileup_combined_weight
- Only occurs in signal samples
 - No selections made, pulled directly from signal ntuples

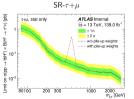
pileup_distributions.png abs_pileup_weights_distribmleupshlengopmgg.png

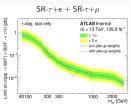
Weight Strategy



- Large absolute values come from very small mc weights
 - $PUW = \frac{X}{mcweight}$
 - Incorrect signal mc weights were applied in high mass region when originally generated
 - Bad weights interact very negatively with pileup reweighting tool
- Pileup Weight (PUW) strategy finalized
 - Applying a cut that encompasses full lower mass PUW distributions
 - $0 \ge PUW \le 2.5$
 - Setting all values outside of threshold to 1







Continuing Work



- MC/Background agreement studies and reweighting are undergoing finalizing
 - ullet Reweighting $tar{t}$ backgrounds due to generator mismodelling of high jet multiplicity
 - Fake factor studies
 - Including systematic uncertainties
- Systematic studies
 - PNN evaluation on systematic files
- Partial unblinding of low PNN score areas
 - Iterating with Editorial Board on strategy and preliminary results
- Finish Internal Note
 - Iterating with Editorial Board
- Make Public Note
- Paper publication
 - Targeted for July, 2022

Conclusion

Thank You



CERN_globe.jpeg





WillAndMeAtATLAS.jp

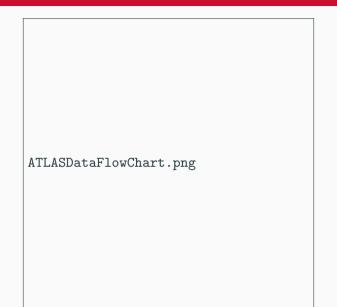
ATLAS_Tour.jpg

2022 PhD Progress Review

Backup

Tile

ATLAS Data Flow



Signals

H[±] Signals

H^{\pm} Mass	Production Mechanism	Decay	Main Background	
$m_{H^\pm} < m_t$	double_resonant $t o H^\pm b$ (LO) double_resonant_production_le	$H^\pm o au u$ (low $tan(eta) \implies H^\pm o cs$ or $H^\pm o cb$)	$tar{t}$, single-top	
$m_{H^\pm} \simeq m_t$	b non-resonant $t o H^\pm b$ (LO) b non_resonant_production_interior interferences taken into account	rm qd iate _t mass.png	$tar{t}$,single-top	
	single-resonant $gg o tbH^\pm$ (NLO)			
$m_{H^\pm} > m_t$	single_resonant_production_l	$H^\pm \to tb$ $(\cos(\beta - \alpha) \simeq 0 \text{ and large } \tan(\beta) \implies H^\pm \to \tau \nu$ $BR(H^\pm \to \tau \nu) \simeq 10 - 15\%$)	multi-jet	

Branching Ratio of H^{\pm}

Charged_Higgs_BR.png

Object Definitions



Object Selection

Object Selection

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

• τ mediumID*

• 1-prong: 75% ID eff

• 3-prong: 60% ID eff

JVT**

 \bullet $p_T < 60 \; {\rm GeV}$

• $|\eta| < 2.4$

|η|***

 $\begin{array}{ll} \bullet & 1.37 < |\mathit{eta}| < 1.52 \\ & \mathsf{excluded} \end{array}$

Region Definitions

Signal Region Definitions

au+jets SR	$ au + \ell$ SR					
E_T^{miss} Trigger (mostly HLT_xe110)	Single lepton trigger (e or μ)					
1 hadronic $ au$	1 hadronic $ au$					
$ ho_T^ au >$ 40 GeV	$p_T^{ au} > 30 \; GeV$					
0 ℓ (e or μ)	$1~\ell$ (e or μ)					
$p_T^\ell > 20 \; GeV$	$p_T^\ell > 30 \; GeV$					
≥ 3 jets	≥ 1 jet					
$p_T^j > 25 \text{ GeV}$	$p_T^j > 25 \text{ GeV}$					
≥ 1 b-jets	≥ 1 b-jets					
$p_T^{b-jet} > 25 \text{ GeV}$	$p_T^{b-jet} > 25 \text{ GeV}$					
$E_{T}^{miss} > 150 \; GeV$	$E_{\rm T}^{\rm miss} > 50~{ m GeV}$					
$m_T(au, E_T^{miss}) > 50 \text{ GeV}$	Opposite sign $ au$ and ℓ					

$\tau + jets$ Region Definitions

 $t\bar{t}$ **CR** ($t\bar{t}$ modeling)

1 hadronic τ

 $p_T^{\tau} > 40 \text{ GeV}$

 \geq 3 jets

> 2 b-jets

 $E_{\rm T}^{\rm miss} > 150 \, GeV$

 $m_T(\tau, E_T^{miss}) < 100 \text{ GeV}$

b-veto CR (Close to SR)

1 hadronic τ

 $p_{T}^{\tau} > 40 \; GeV$

> 3 jets

 $p_{\tau}^{jet} > 25 \; GeV$

 $E_{\rm T}^{\rm miss} > 150 \, GeV$

 $m_T(\tau, E_T^{miss}) > 50 \text{ GeV}$

b veto

 ℓ veto

W+Jets CR (W+Jets modeling)

1 hadronic τ

 $p_T^{\tau} > 40 \; GeV$

≥ 3 jets $p_{\tau}^{jet} > 25 \text{ GeV}$

 $E_{\rm T}^{\rm miss} > 150 \, GeV$

 $m_T(\tau, E_T^{miss}) > 100 \text{ GeV}$ b veto

ℓ veto

b-veto $m_T \ge 100$ **CR** (Fake enriched region)

1 hadronic au

 $p_{T}^{\tau} > 40 \; GeV$

 \geq 3 jets

 $p_T^{jet} > 25 \text{ GeV}$

 $E_{\rm T}^{\rm miss} > 150 \, GeV$

 $m_T(\tau, E_T^{miss}) > 100 \text{ GeV}$

b veto

 ℓ veto

$\tau + \ell$ Region Definitions

$\textbf{Dilepton-btag} \ \textbf{CR} \ \big(t \bar{t} \ \text{modeling, used in fit} \big)$

 τ veto n > 1 iets

 $p_T^{jet} > 25 GeV$

 ≥ 1 b-jets

 $E_{\rm T}^{\rm miss} > 50 \, GeV$

1 e

 $\perp \mu$

b-veto CR (Close to SR)

1 hadronic au

 $p_T^{ au} >$ 30 GeV

 $1 e(\mu)$

Veto μ (e)

Opposite sign τ $e(\mu)$

 ≥ 1 jets

 $p_T^{jet} > 25 \, GeV$

 $E_{\rm T}^{\rm miss} > 50 \, GeV$

1 tight $e(\mu)$

Zee CR (Fake enriched region)

1 hadronic au

 $p_T^{\tau} > 30 \; GeV$

veto μ

Opposite sign τ e

 ≥ 1 jets

 $p_T^{jet} > 25 \, GeV$

bjet veto

 $E_{\mathrm{T}}^{\mathrm{miss}} > 50\,GeV$

1 e

 $40 < mass(au, e) < 140 \, GeV$

Same Sign CR (Fake enriched region)

1 hadronic τ $p_T^{\tau} > 30 \text{ GeV}$

Same sign $\tau e(\mu)$

Veto μ (e) $n_{-jets} \ge 1$

 $p_T(jet) > 25 GeV$ $E_T^{miss} > 50 GeV$

1 tight $e(\mu)$

PNN



Comparing kFolds (80 GeV, given BKG)



Comparing kFolds (80 GeV, given SIG)

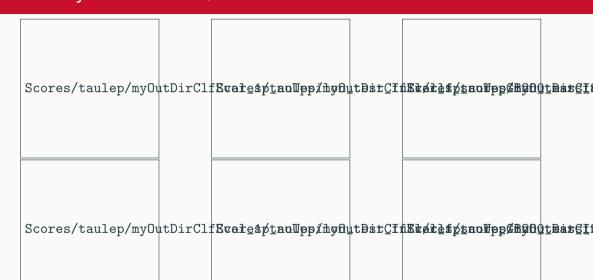


Comparing kFolds (3000 GeV, given SIG)

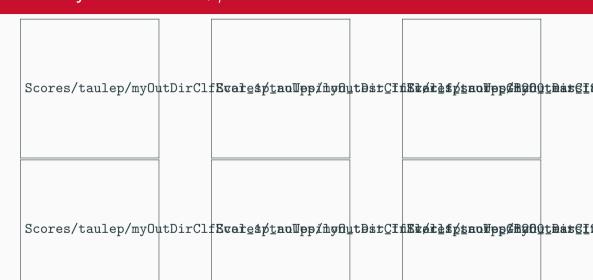


• May have gotten bkg/sig labels backwards :)

Preliminary PNN Results $\tau + e$ SR



Preliminary PNN Results $\tau + \mu$ SR



Pileup Weights

Pileup Weights Distributions

pileup_histo.png	p:

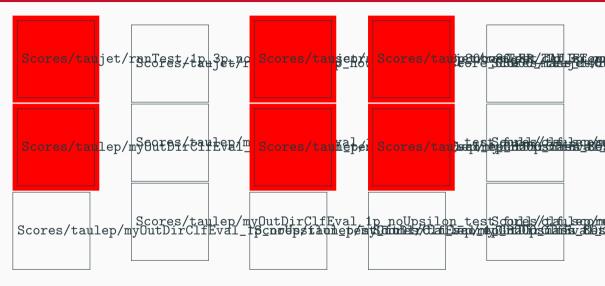
pileup_histo_zoomed.png

Pileup Weights Statistics

	80	90	100	110	120	130	140	150	160	170	180	190	200	225	250	275	300
count	80844.000000	81658.000000	83016.000000	84173.000000	86004.000000	86958.000000	85136.000000	84247.000000	82619.000000	84875.000000	86097.000000	86257.000000	40284.000000	40886.000000	41792.000000	42556.000000	42978.000000
mean	1.001054	1.001780	1.000126	1.001323	1.000811	1.002053	0.999559	0.998034	0.996655	0.999633	0.999488	0.999026	0.508634	1.010265	1.099422	1.014135	1.014974
std	0.253943	0.253156	0.254526	0.253217	0.252664	0.251347	0.255629	0.257582	0.258068	0.254013	0.255667	0.255854	17.490118	0.288834	6.412658	0.339309	0.356573
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	-610.129272	-0.445639	-1.273962	-1.206831	-1.760294
25%	0.920860	0.921549	0.921623	0.921557	0.921378	0.921159	0.921536	0.920747	0.921333	0.923230	0.920085	0.920768	0.900776	0.885019	0.934779	0.922488	0.943855
50%	1.079184	1.081243	1.079326	1.079540	1.080067	1.080879	1.079532	1.079685	1.080971	1.079383	1.080233	1.080472	1.060562	1.064517	1.041212	1.063800	1.046314
75%	1.148908	1.155570	1.150078	1.156189	1.149427	1.150257	1.149785	1.149956	1.150023	1.149662	1.151101	1.150765	1.165013	1.178183	1.187810	1.207834	1.167239
95%	1.205099	1.203470	1.204149	1.204069	1.203694	1.203698	1.205689	1.203701	1.204432	1.204153	1.202932	1.205496	1.315114	1.390322	1.320100	1.326961	1.355505
max	2.409515	2.389924	2.411615	2.404698	2.393531	2.405815	2.386079	2.406105	2.400672	2.398967	2.392730	2.376339	28.892088	3.133942	464.072968	8.240099	17.410671

	350	400	500	600	700	800	900	1000	1200	1400	1600	1800	2000	2500	3000
count	43798.000000	44565.000000	45412.000000	45469.000000	45321.000000	45464.000000	45156.000000	43913.000000	44639.000000	43592.000000	43273.000000	42691.000000	42769.000000	41885.000000	40783.000000
mean	1.006303	1.023722	0.591891	1.020003	1.006465	1.008190	0.926411	1.016459	1.021096	1.018929	1.011275	1.021495	1.006681	1.027024	1.054152
std	0.325241	0.590001	28.165703	0.393464	0.305804	1.520343	4.324449	0.346823	0.410831	0.771618	0.514045	0.328171	0.453343	0.314811	2.333855
min	-7.745956	-0.068810	-1897.117554	-3.842706	-3.495028	-127.155434	-209.416672	-5.508759	-0.557234	-40.070904	-15.846775	-1.963920	-11.616090	-1.816558	-2.588256
25%	0.880324	0.922226	0.900143	0.933184	0.905355	0.907488	0.940907	0.900577	0.924895	0.932238	0.891818	0.927121	0.909429	0.928094	0.920930
50%	1.067698	1.043523	1.057348	1.062986	1.040850	1.022655	1.068709	1.080882	1.039005	1.053244	1.050301	1.053411	1.040414	1.068602	1.037835
75%	1.176714	1.174688	1.188619	1.166802	1.191314	1.184892	1.165839	1.181656	1.163044	1.164395	1.189937	1.157566	1.193018	1.166120	1.161961
95%	1.360068	1.336481	1.326447	1.296390	1.347952	1.450452	1.356199	1.408593	1.391922	1.416487	1.371108	1.371884	1.398844	1.420707	1.432250
max	3.044234	19.576925	1.687775	8.729278	3.958143	9.343400	2.870858	5.665573	17.634136	26.660934	20.727751	5.528868	3.841987	4.248174	175.623535

Preliminary PNN Results with No Pileup Weights



Preliminary PNN Results with No Pileup Weights



PNN Hyperparameter Search

Activation and Loss Functions



LeakyReLu

Hinge_LossFunction.png Hinge Loss Hinge_Squared_lossFunction
Squared Hinge Loss

Binary_Crossentropy_Loss!
Binary Cross

Entropy Loss

PNN Hyperparameter Search Results

$j \rightarrow \tau$ Fakes

Background Estimation: $j \rightarrow \tau$ Fakes

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

Multi-Jet (gluon enriched)	W+Jets (quark enriched)	
E _T ^{miss} or Multi-Jet trigger	Single lepton triggers	
≥ 3 jets	≥ 1 jet	
0 b-jets	0 b-jets	
0 ℓ	$1 \ell, p_T^{\ell} > 30 \text{ GeV}$	
$E_T^{miss} < 80 \text{ GeV}$	1 τ	
$m_T(au, E_T^{miss}) > 50 \text{ GeV}$	$60 < m_T(\ell, E_T^{miss}) < 160 \text{ GeV}$	

- 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

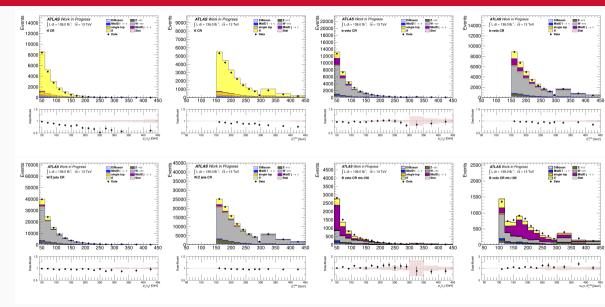
$$FF_i^{comb} = \alpha_i^{MJ} FF_i^{MJ} + [1 - \alpha_i^{MJ}] FF_i^{W+jets}$$

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

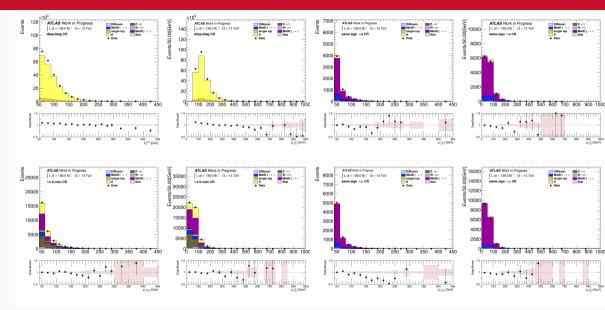
$$N_{fakes}^{ au_{had-vis}} = \sum_{i} N_{ar{ au}_{had-vis}}^{SR_i} FF_i$$

Background Modeling

$\tau + jets$ Background Modeling



$\tau + \ell$ Background Modeling



Preliminary PNN Results $\tau + \mathit{jets}$ SR



Fit

Fit Model

- Workspaces prepared with WSMaker
- Statistical uncertainty only considered for now
- MVA classifier score used as the discriminating variable
 - Fitting the profile likelihood ratio
- Fitted Regions
 - ullet $au+\ell$ Dilepton-btag CR is fit with a single bin (normalization only)

Control Region	Signal Regions	
t ar t	au + jets SR	
	au + e SR	
	$ au + \mu$ SR	

Binning is yet to be decided

Previous Analysis

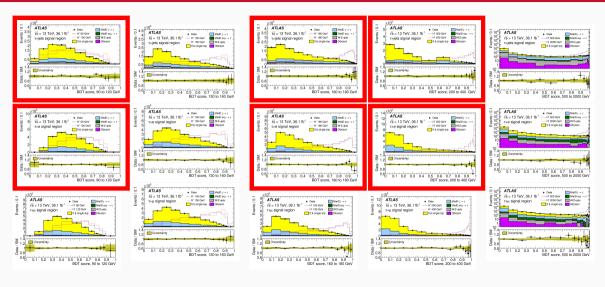
Previous Multivariate Analysis

- Background modeling and classifier training kept statistically independent via the k-fold method
- ullet BDTs (FastBDT via TMVA) binned in m_{H^\pm}
 - 90-120, 130-160, 160-180, 200-400, 500-2000 GeV
 - Trained separately on $\tau + jets$ and $\tau + \ell$
 - Independent trainings for 1 or 3 au tracks
 - $\tau + \ell$ trained inclusively, evaluated separately on $\tau + e$ and $\tau + \mu$
- At low $m_{H^{\pm}}$, τ polarization is important
 - $\bullet \quad \Upsilon = \frac{{E_T^\pi}^\pm {E_T^\pi}^0}{E^T} \approx 2 \frac{{p_T^\tau}^{-track}}{p^T} 1$
 - Only defined for 1 track au candidates
- At high, $m_{H^{\pm}} \Upsilon$ becomes less important
 - $E_{\rm T}^{\rm miss}$, $p_{\rm T}^{\tau}$, and $\Delta\phi_{\tau, {
 m miss}}$ have higher discriminating power



MVA input variable	$\tau + jets$	$\tau + \ell$
Emiss	✓	✓
p_T^{τ}	✓	✓
p_T^{b-jet}	✓	✓
$\Delta \phi(\tau, E_T^{miss})$	✓	✓
$\Delta \phi(b-jet,E_T^{miss})$	✓	✓
$\Delta R(b-jet, \tau)$	✓	
p_T^I		✓
Ad(I Emiss)		/

Previous BDT Scores in Signal Regions



Previous BDT Scores in Signal Regions

