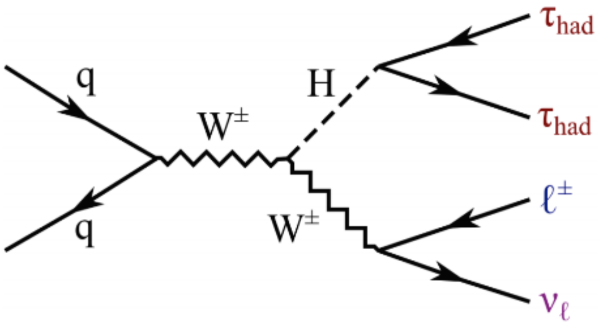
Solutions and Improvement for $VH \longrightarrow \tau\tau$ Run 2 Analysis

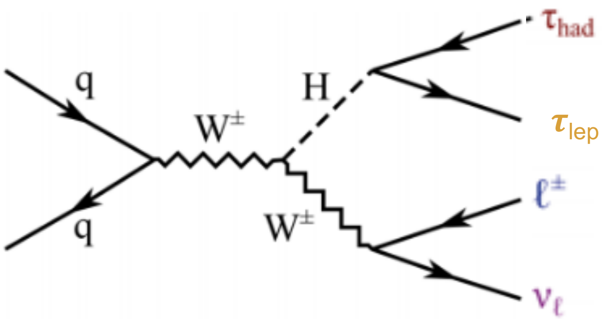
Savannah Thais, Yale University
USLUA Lightning Round, 11/03/2017

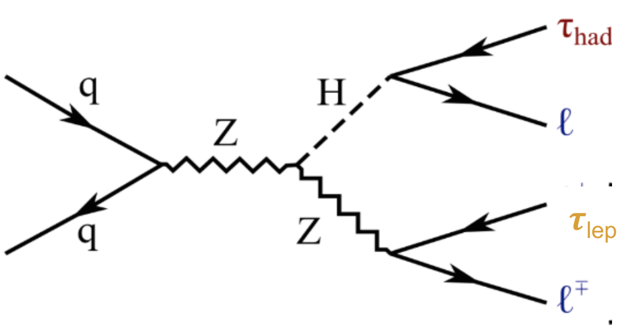
Analysis Introduction

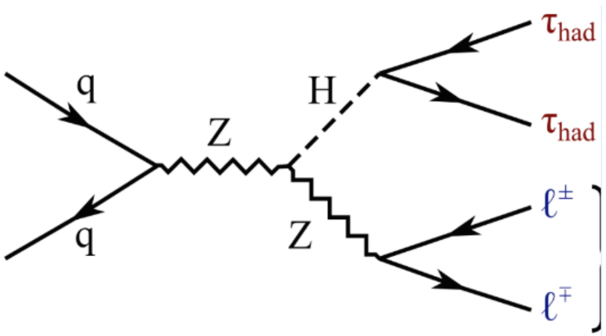
BASICS

- Associated production (VH): W or Z radiates a Higgs
- The Higgs decays to 2 taus
- 4 analysis categories
- All final states have 3-4 objects:
  - 2 taus and 1-2 light leptons
- In ATLAS Sofware:
  - Taus are difficult to trigger on and reconstruct cleanly
  - Light lepton triggers and IDs are highly efficient





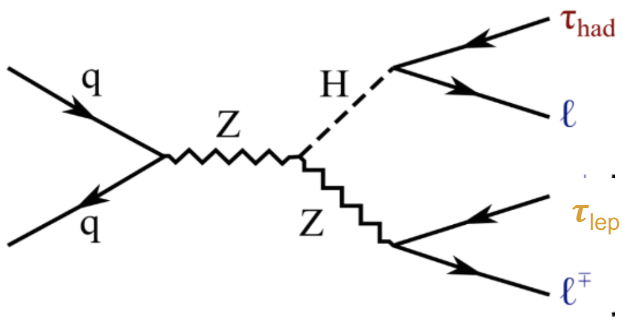




Analysis Introduction

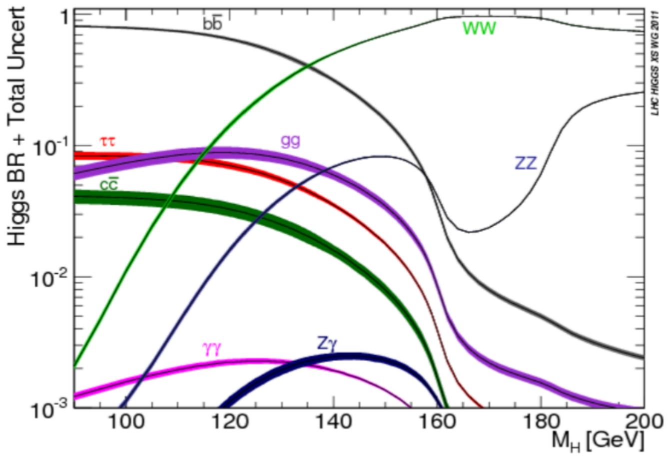
BASICS

- Associated production (VH): W or Z radiates a Higgs
- The Higgs decays to 2 taus
- 4 analysis categories
- All final states have 3-4 objects:
  - 2 taus and 1-2 light leptons
- In ATLAS Sofware:
  - Taus are difficult to trigger on and reconstruct cleanly
  - Light lepton triggers and IDs are highly efficient



MOTIVATION

- $H \longrightarrow \tau\tau$ has a high branching fraction
- Object requirements on 3-4 objects ➜ reduced multi-jet backgrounds
- Increased luminosity ➜ VH is a more viable production mode
- Increased pileup ➜ more difficulty measuring other $H \longrightarrow \tau\tau$ production modes



Very promising channel to observe fermionic Higgs decay!

Analysis Procedure

Select Events	Estimate Backgrounds	Reconstruct Mass Spectrum	Fit Mass Spectrum
Select appropriate events for each of our 4 analysis categories	Calculate background contributions to selected event distributions (major source of uncertainty)	Construct the di-tau mass spectrum accounting for missing mass from neutrinos	Fit the di-tau mass spectrum, compare to SM predictions, obtain final results

Analysis Procedure

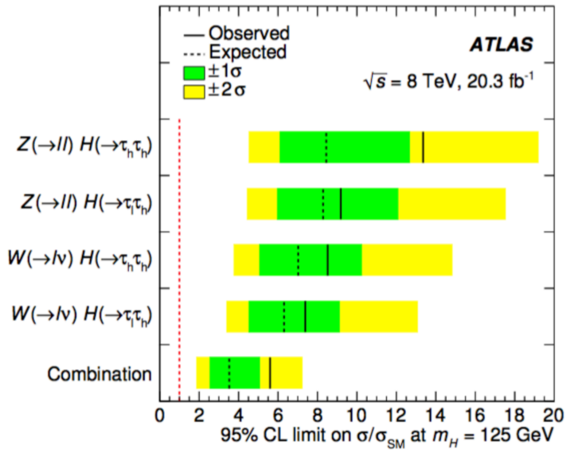
Select Events → **Estimate Backgrounds** → **Reconstruct Mass Spectrum** → **Fit Mass Spectrum**

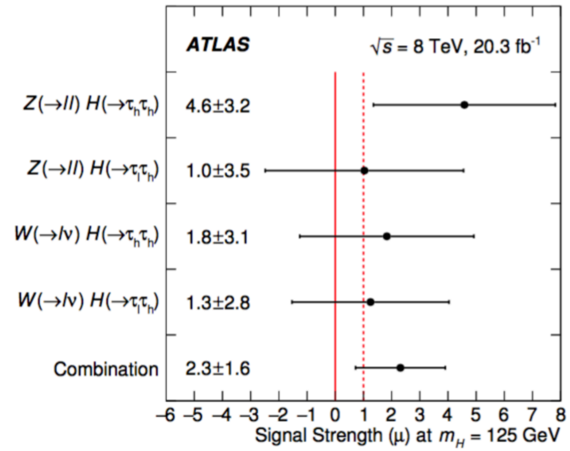
Select appropriate events for each of our 4 analysis categories

Calculate background contributions to selected event distributions (major source of uncertainty)

Construct the di-tau mass spectrum accounting for missing mass from neutrinos

Fit the di-tau mass spectrum, compare to SM predictions, obtain final results





Run 1 combined signal strength: $\mu = 2.3 \pm 1.6$

Event Selection

All categories must have at least 3 identified leptons including at least 1 light lepton

$W \longrightarrow \mu\nu/e\nu$, $H \longrightarrow \tau_{lep}\tau_{had}$

- Exactly one isolated muon and electron
- Exactly one τ_{had} passing medium BDT ID
- $p_T(\tau_{had}) > 25$ GeV
- Same-charge e and μ, opposite charge τ_{had}
- Events with b-tagged jets with $p_T > 30$ GeV are vetoed
- $|p_T(\tau_{had})| + |p_T(\mu)| + |p_T(e)| > 80$ GeV
- $\Delta R(\tau_{had}, \tau_{lep}) < 3.2$

$W \longrightarrow \mu\nu/e\nu$, $H \longrightarrow \tau_{had}\tau_{had}$

- Exactly one isolated muon or electron
- Exactly two opposite charge τ_{had} passing medium BDT ID
- $p_T(\tau_{had}) > 20$ GeV
- $|p_T(\tau_{had1})| + |p_T(\tau_{had2})| > 100$ GeV
- $m_T(l, E_T^{miss}) > 20$ GeV
- $0.8 < \Delta R(\tau_{had1}, \tau_{had2}) < 2.8$
- Events with b-tagged jets with $p_T > 30$ GeV are vetoed

$Z \longrightarrow \mu\mu/ee$, $H \longrightarrow \tau_{lep}\tau_{had}$

- Exactly three electrons or muons
- One opposite-charge and same-flavor lepton pair with invariant mass $80 < m_{ll} < 100$ GeV
- Exactly one τ_{had} passing medium BDT ID with opposite charge to the lepton assigned to the Higgs
- $p_T(\tau_{had}) > 20$ GeV
- $|p_T(\tau_{had})| + |p_T(\tau_{lep})| > 60$ GeV

$Z \longrightarrow \mu\mu/ee$, $H \longrightarrow \tau_{had}\tau_{had}$

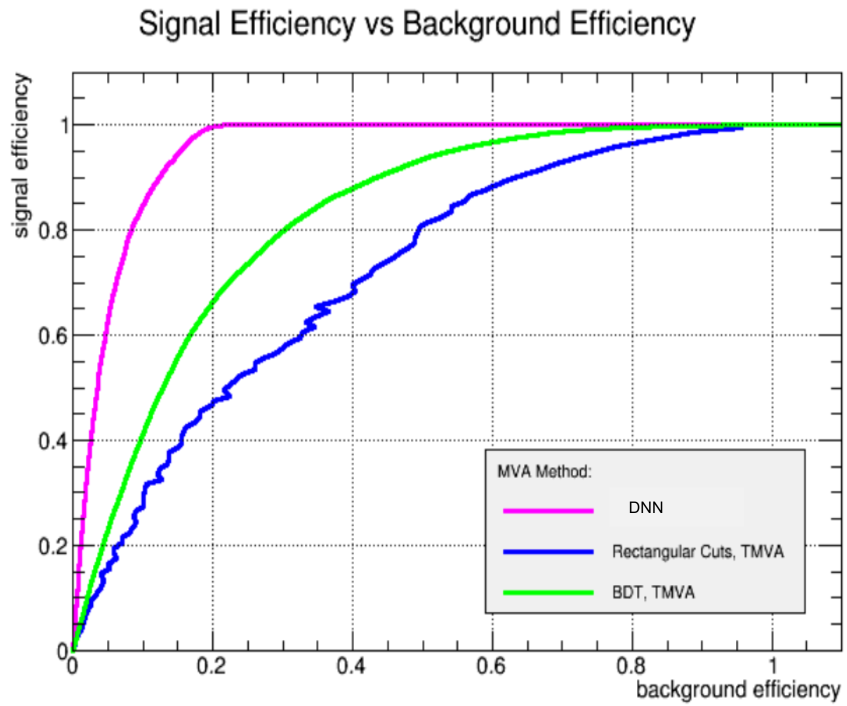
- Exactly two electrons or muons of opposite charge
- Exactly two opposite charge τ_{had} passing medium BDT ID
- $p_T(\tau_{had}) > 20$ GeV
- $60 < m_{ll} < 120$ GeV
- $|p_T(\tau_{had1})| + |p_T(\tau_{had2})| > 88$ GeV

Event Selection Improvements

- Could improve event selection efficiency with Machine Learning
  - Either an individual algorithm for each analysis category
  - Or algorithm with multiple classification outputs
- Can train algorithm using same variables from Run 1 event selection
  - Particle p_T, triggers, number of leptons, lepton charge, lepton quality
  - Or additional variables that were not appropriate for cuts

Event Selection Improvements

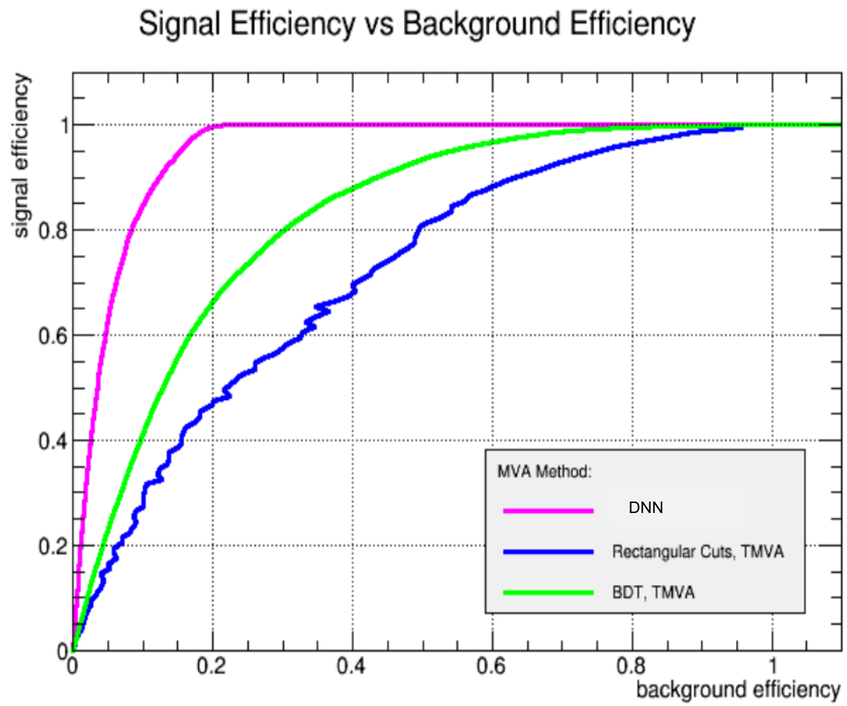
- Could improve event selection efficiency with Machine Learning
  - Either an individual algorithm for each analysis category
  - Or algorithm with multiple classification outputs
- Can train algorithm using same variables from Run 1 event selection
  - Particle p_T, triggers, number of leptons, lepton charge, lepton quality
  - Or additional variables that were not appropriate for cuts



Previous study for $H \rightarrow ZZ_d$ event selection using deep neural network

Event Selection Improvements

- Could improve event selection efficiency with Machine Learning
  - Either an individual algorithm for each analysis category
  - Or algorithm with multiple classification outputs
- Can train algorithm using same variables from Run 1 event selection
  - Particle p_T, triggers, number of leptons, lepton charge, lepton quality
  - Or additional variables that were not appropriate for cuts



Previous study for $H \rightarrow ZZ_d$ event selection using deep neural network

Improving background rejection would reduce dependency on fake rate measurements!

Background Estimation

This analysis has two types of backgrounds:

1. Irreducible backgrounds from diboson, $Z \longrightarrow \tau\tau$, and ttbar (contributions measured in MC)
2. Backgrounds from non-prompt leptons/taus and objects faking leptons/taus (contributions measured in data)

Background Estimation

Backgrounds from non-prompt leptons/taus and objects faking leptons/taus

Calculate this background contribution using fake rate measured in data

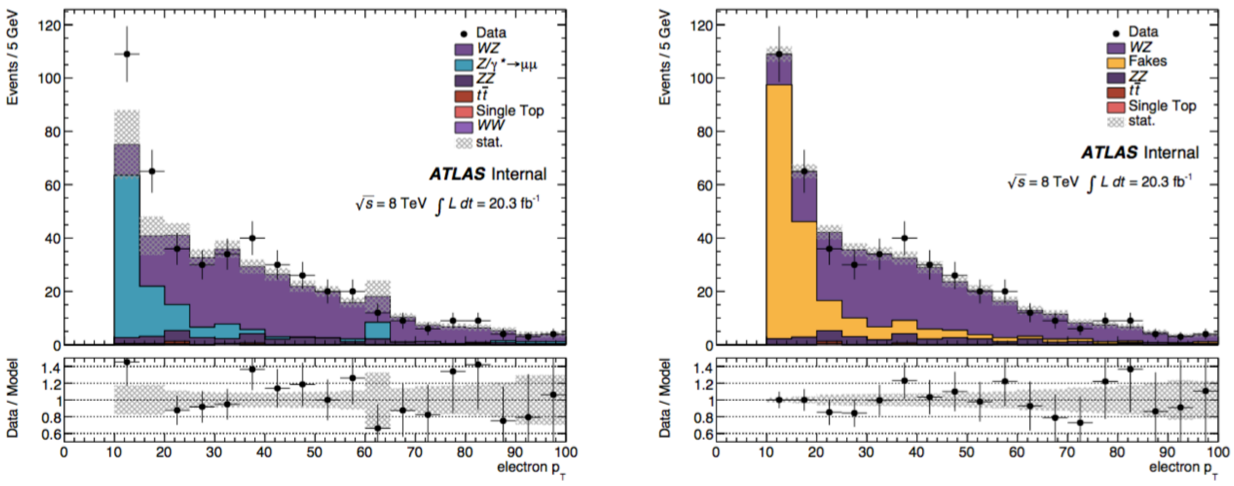
1. Select a region with kinematics similar to Signal Region (SR) but enriched in objects failing ID criteria (anti-selected objects)
2. Measure the frequency at which object candidates in this region pass the object selection criteria (fake rate, r)
3. Use r calculate the scaling factor between events in fake enriched region to events in SR (fake factor, f=r/(1-r))
4. Apply f as an additional bin weighting in SR to account for fake/non-prompt background contribution to final results

Background Estimation

Backgrounds from non-prompt leptons/taus and objects faking leptons/taus

Calculate this background contribution using fake rate measured in data

1. Select a region with kinematics similar to Signal Region (SR) but enriched in objects failing ID criteria (anti-selected objects)
2. Measure the frequency at which object candidates in this region pass the object selection criteria (fake rate, r)
3. Use r calculate the scaling factor between events in fake enriched region to events in SR (fake factor, f=r/(1-r))
4. Apply f as an additional bin weighting in SR to account for fake/non-prompt background contribution to final results



(a) MC estimate (b) FF estimate

Electron fake rate closure test

Background Estimation Improvements

Taus

- Large systematic uncertainty in Run 1
- In part due to unknown quark/gluon initiated fraction
- Can improve by analyzing all tau types together using PanTau

Background Estimation Improvements

Taus

- Large systematic uncertainty in Run 1
- Mainly due to unknown quark/gluon initiated fraction
- Can improve by analyzing all tau types together using PanTau

Electrons

- Only used in W,τ_{lep} τ_{had} channel in Run 1
- ATLAS has new data-driven electron ID which reduces non-prompt selection rate
- Further improved image based ID under construction
- More ID operating points → more options for selection criteria

Background Estimation Improvements

Taus

- Large systematic uncertainty in Run 1
- In part due to unknown quark/gluon initiated fraction
- Can improve by analyzing all tau types together using PanTau

Electrons

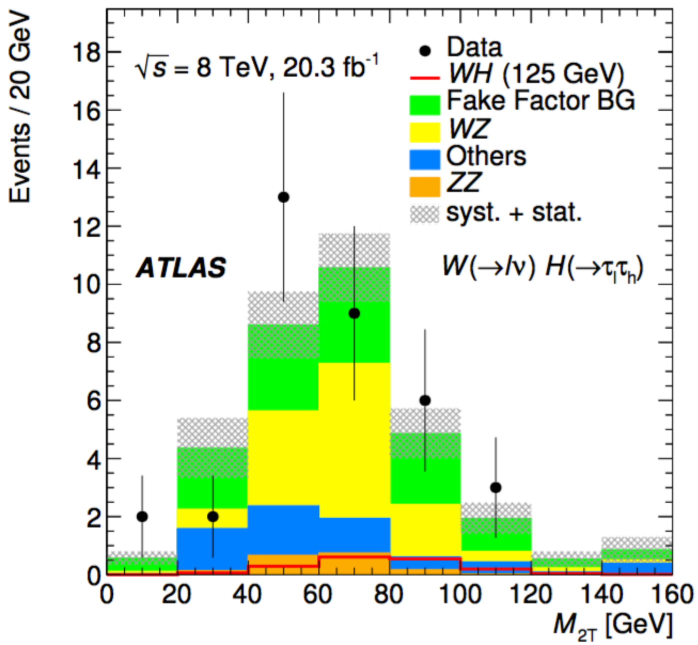
- Only used in $W,\tau_{lep}\tau_{had}$ channel in Run 1
- ATLAS has new data-driven electron ID which reduces non-prompt selection rate
- Further improved image based ID under construction
- More ID operating points → more options for selection criteria

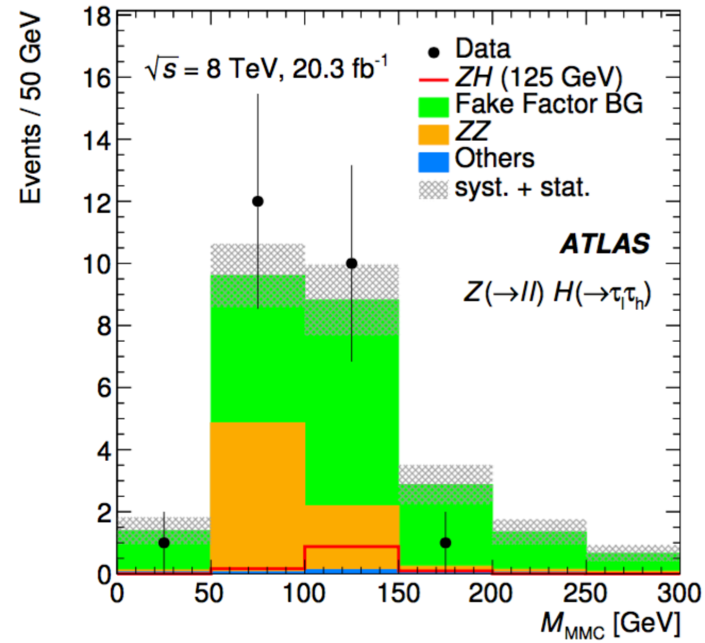
Muons

- Fake rate was unusable in Run 1 due to trigger/ID inconsistencies
- Now using isolation as selection criteria rather than ID

Mass Reconstruction

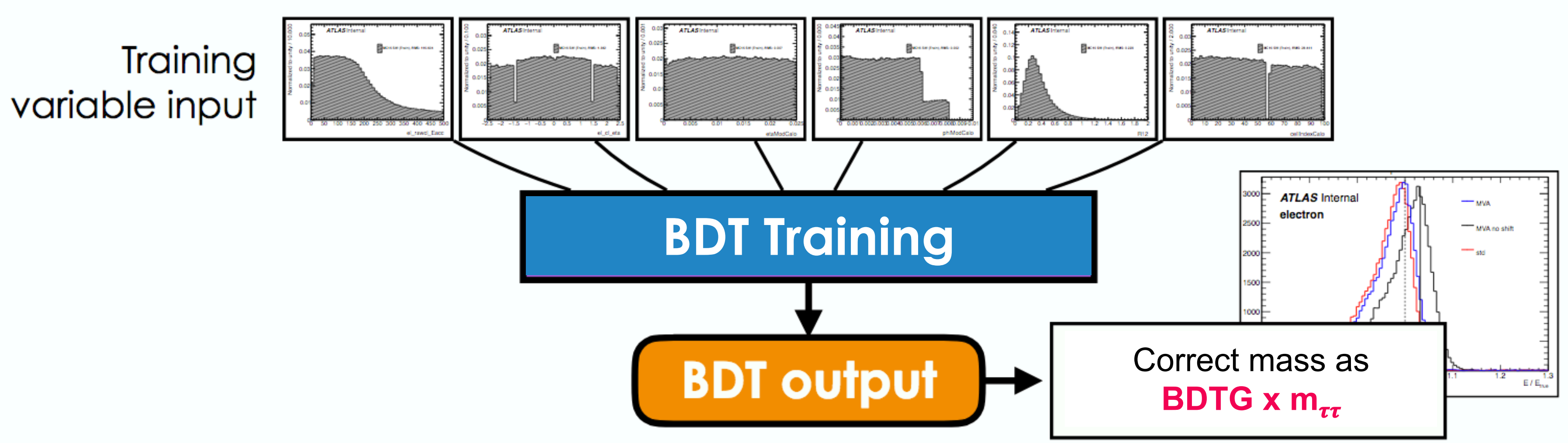
- When constructing final di-tau mass spectrum must account for missing mass from neutrinos
- Done separately for ZH (neutrinos only from tau decays) and WH (additional neutrino from W decay)
  - Technique for WH only gives lower bound on di-tau mass
- This procedure is imprecise → introduces additional uncertainty





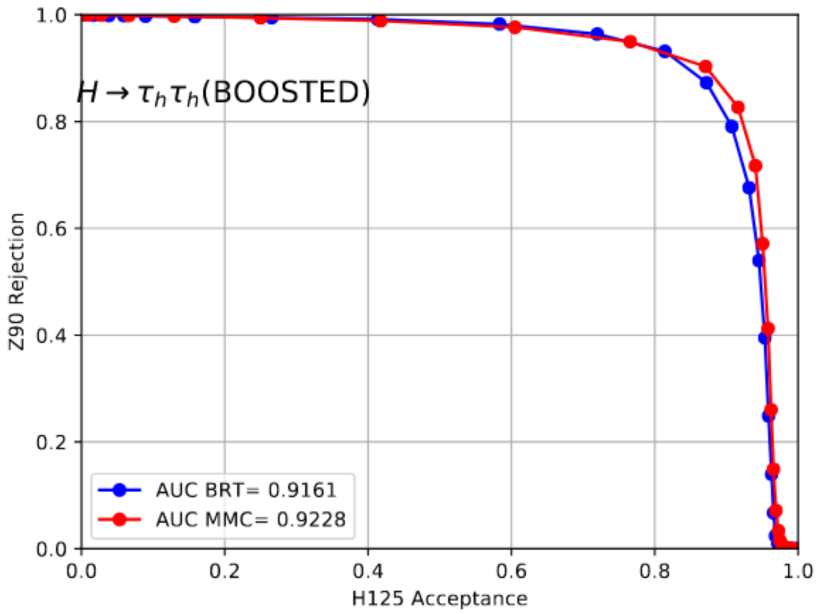
Mass Reconstruction Improvements

- Can use Boosted Decision Trees (BDTs) to improve the missing mass calculation:
  - Train BDT to produce mass correction scale factor using similar input variables as current techniques (p_T, visible mass, angular separation)
  - Can be binned based on event kinematics → more precise



Mass Reconstruction Improvements

- Can use Boosted Decision Trees (BDTs) to improve the missing mass calculation:
  - Train BDT to produce mass correction scale factor using similar input variables as current techniques (p_T, visible mass, angular separation)
  - Can be binned based on event kinematics → more precise



Same procedure has been successful in other applications:

- Electron energy re-calibration in E/Gamma
- $H \longrightarrow \tau\tau$ analysis (BRTs)

Conclusions

- VH $\rightarrow \tau\tau$ is a promising channel for measuring fermionic Higgs decay in ATLAS in Run 2

Conclusions

- VH → $\tau\tau$ is a promising channel for measuring fermionic Higgs decay in ATLAS in Run 2
- Improvements planned for main analysis steps:
  - Event selection: use ML classifier
  - Background estimation: improved particle ID and fake object criteria
  - Mass reconstruction: Use BDTs to calculate mass correction factor

Conclusions

- VH → $\tau\tau$ is a promising channel for measuring fermionic Higgs decay in ATLAS in Run 2
- Improvements planned for main analysis steps:
  - Event selection: use ML classifier
  - Background estimation: improved particle ID and fake object criteria
  - Mass reconstruction: Use BDTs to calculate mass correction factor
- Even more improvements are possible: improving uncertainty calculations, mass spectrum fitting, inclusion of additional analysis categories, and more….

Conclusions

- VH → $\tau\tau$ is a promising channel for measuring fermionic Higgs decay in ATLAS in Run 2
- Improvements planned for main analysis steps:
  - Event selection: use ML classifier
  - Background estimation: improved particle ID and fake object criteria
  - Mass reconstruction: Use BDTs to calculate mass correction factor
- Even more improvements are possible: improving uncertainty calculations, mass spectrum fitting, inclusion of additional analysis categories, and more….
- **Exciting results to come!**

Thanks!
Any questions?

Savannah Thais
savannah.thais@yale.edu