Double Higgs Production at Colliders Workshop

Fermilab, September 7th, 2018

Substructure for HH signatures

Marc Antoine Osherson
Rutgers University

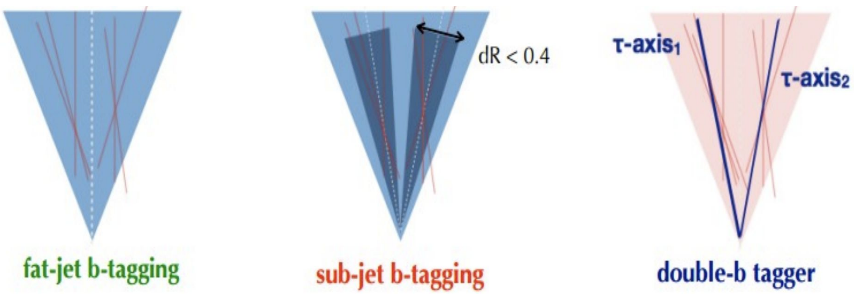
Introduction

- When produced with sufficient momentum, higgs in HH signatures are best reconstructed as single large cone jets: most obvious case H → bb
- Resulting boosted higgses are composite, and can be exploited as such for event categorization, background estimates, etc...
- Properties we can exploit:
  - bb/cc-tagging
  - Jet Mass
  - Two-Pronged Structure
  - Other Aspects?

Introduction

- When produced with sufficient momentum, higgs in HH signatures are best reconstructed as single large cone jets: most obvious case H → bb
- Resulting boosted higgses are composite, and can be exploited as such for event categorization, background estimates, etc...
- Properties we can exploit:
  - **bb/cc-tagging**[1]
  - Jet Mass
  - Two-Pronged Structure
  - Other Aspects?
- Improvements? Future?

Already Covered!

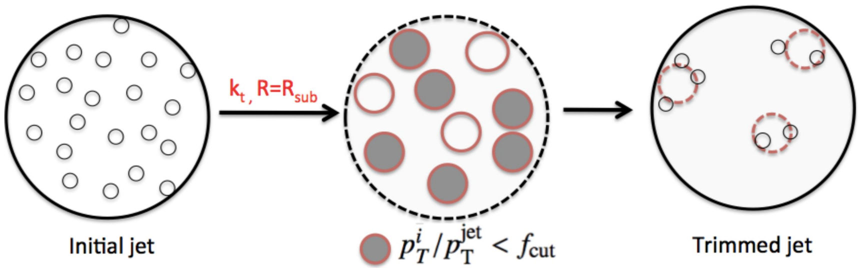


1: https://indico.cern.ch/event/731450/timetable/#17-identifying-boosted-h-bb

Introduction

- When produced with sufficient momentum, higgs in HH signatures are best reconstructed as single large cone jets: most obvious case H → bb
- Resulting boosted higgses are composite, and can be exploited as such for event categorization, background estimates, etc...
- Properties we can exploit:
  - bb/cc-tagging
  - **Jet Mass**[2]
  - Two-Pronged Structure
  - Other Aspects?
- Improvements? Future?

Grooming techniques such as trimming, pruning or soft drop can be used to better reconstruct the higgs mass peak, and suppress multijet backgounds

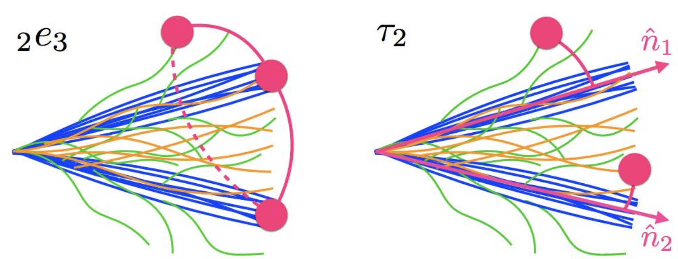


2: image Performance of jet substructure techniques for large-R jets in proton-proton collisions at √s=7 TeV using the ATLAS detector

Introduction

- When produced with sufficient momentum, higgs in HH signatures are best reconstructed as single large cone jets: most obvious case H → bb
- Resulting boosted higgses are composite, and can be exploited as such for event categorization, background estimates, etc...
- Properties we can exploit:
  - bb/cc-tagging
  - Jet Mass
  - **Two-Pronged Structure**[3]
  - Other Aspects?
- Improvements? Future?

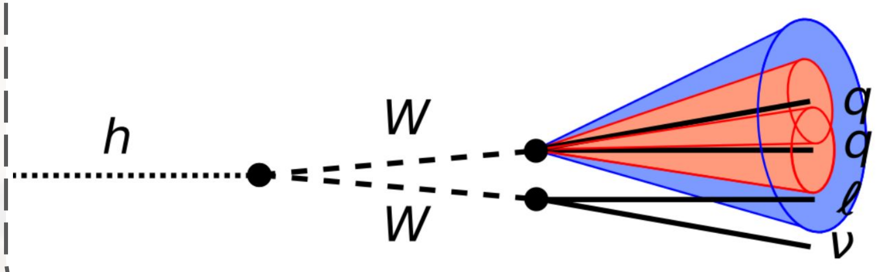
Additional discriminating power by requiring the jet constituents to be clustered in two lobes



3: image New angles on energy correlation functions

- When produced with sufficient momentum, higgs in HH signatures are best reconstructed as single large cone jets: most obvious case H → bb
- Resulting boosted higgses are composite, and can be exploited as such for event categorization, background estimates, etc...
- Properties we can exploit:
  - bb/cc-tagging
  - Jet Mass
  - Two-Pronged Structure
  - **Other Aspects[4]?**
- Improvements? Future?

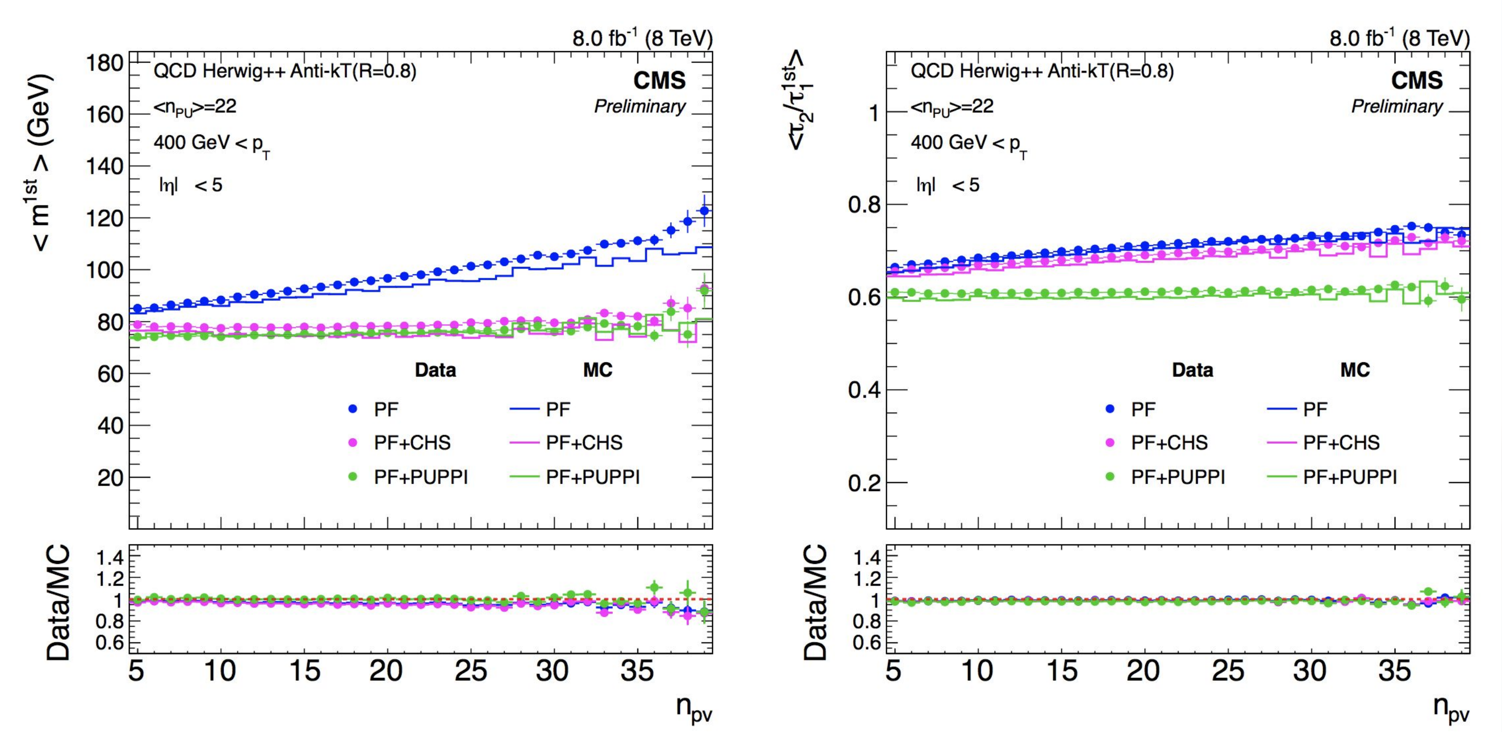
H → bb isn't the whole story! Many higgs branchings to reconstruct:



4: image thanks **Kira Abeling**

Building Large Jets at CMS

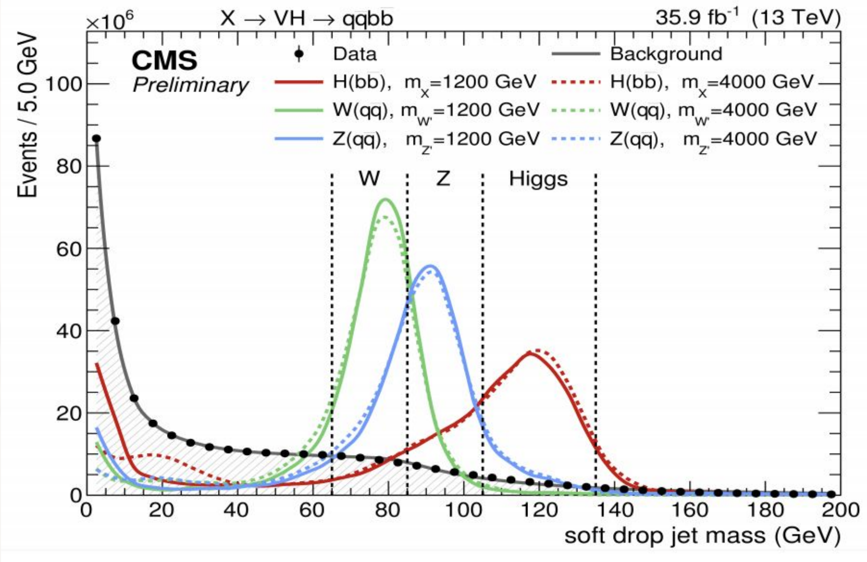
- CMS jets are built from the Particle Flow Algorithm which uses information for every element of the detector.
  - Pileup removed with either CHS or PuPPI
  - PF candidates are natural inputs for substructure measurements



[1]: CMS PAS Study of Pileup Removal Algorithms for Jets

Jet Mass Algorithms at CMS

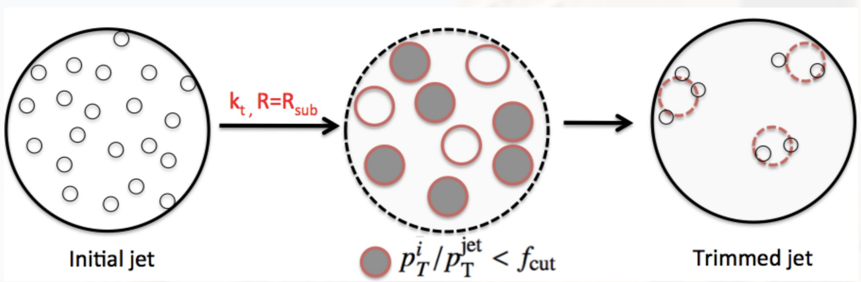
- Two dominant grooming techniques at CMS: Pruning[1] & Soft Drop[2]
- Pruning = Recluster
  - Cluster the jet from its constituents, pausing at each pair of proto-jets to throw out those which fail p_T fraction or ΔR requirements
- Soft Drop = Decluster
  - Break the jet into its last two constituents
  - Discard half if it fails p_T fraction (ΔR requirements)

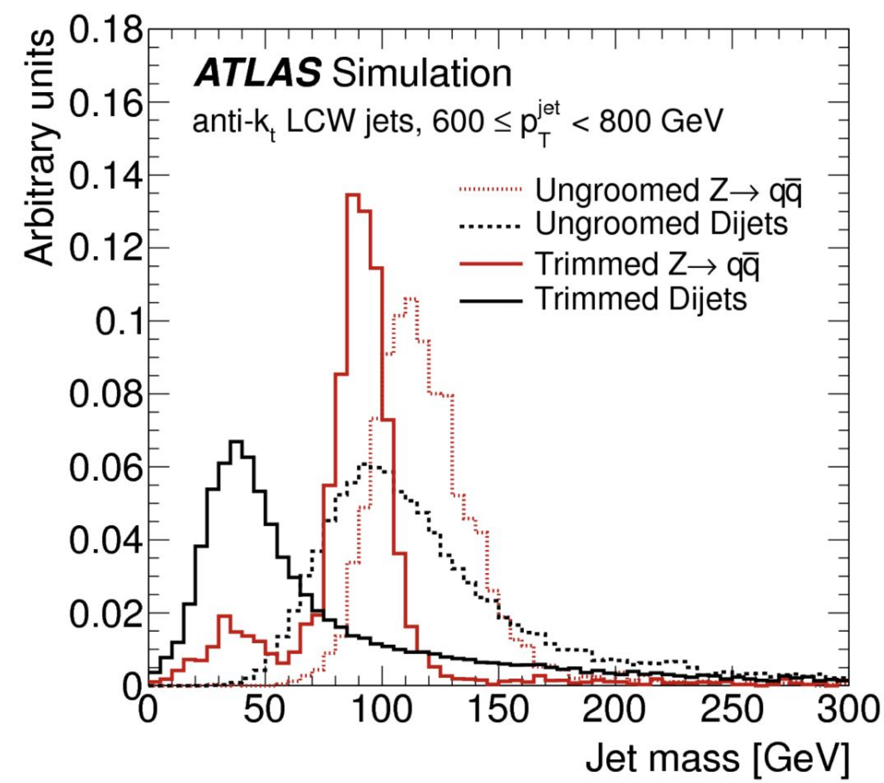


1: pruned mass paper Techniques for improved heavy particle searches with jet substructure
2: soft drop paper Soft Drop 3: image Search for heavy resonances that decay into a vector boson and a Higgs boson

Building Large Jets at ATLAS

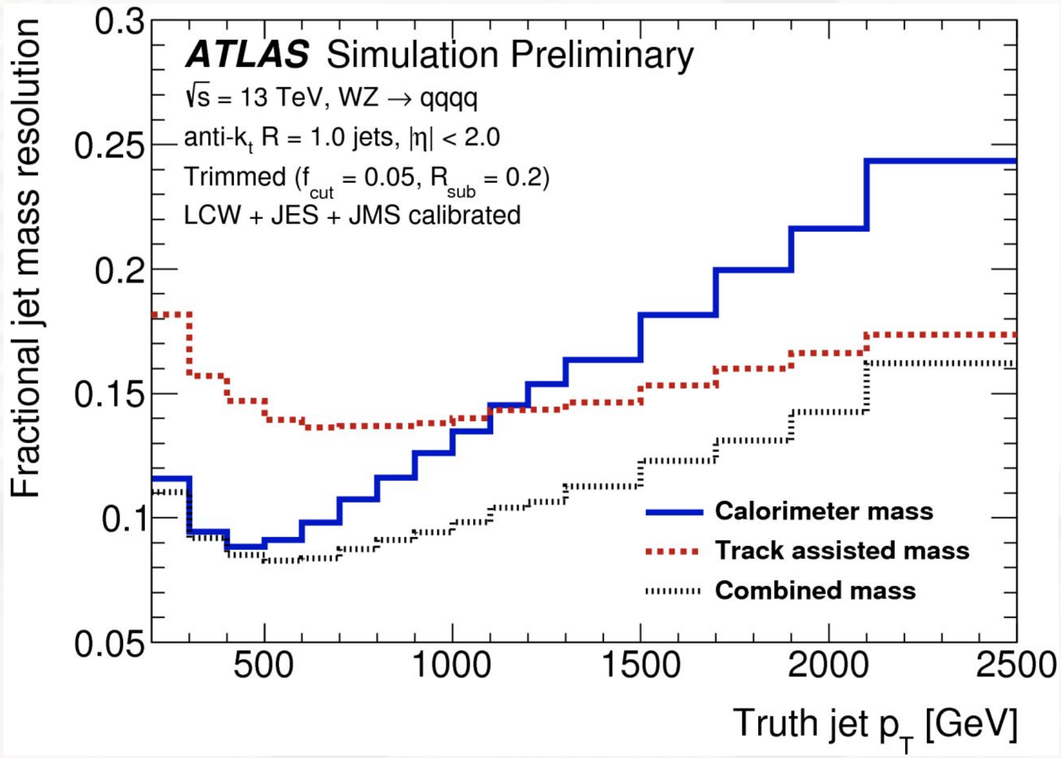
- ATLAS builds similar groomed jets from its calo-clusters.
- Similar idea, different algorithm: Trimming[1]
  - Recluster jet into R =0.2 subjets, discard subjets based on a p_T requirement.





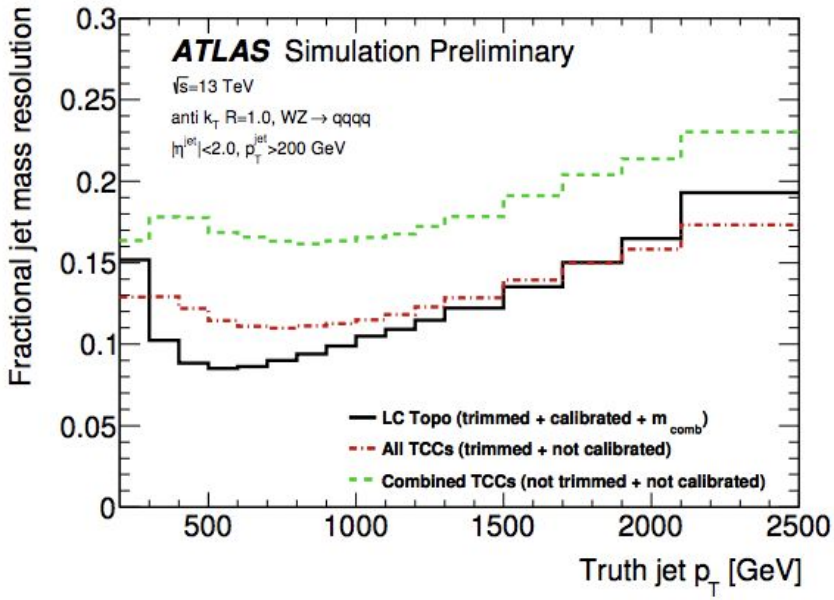
1: paper Jet Trimming
2: image Performance of jet substructure techniques for large-R jets in proton-proton collisions at √s=7 TeV using the ATLAS detector

- Track information is then incorporated in two stages: Track Assisted Mass and the Combined Mass.
  - Track Assisted Mass $m^{TA} \equiv m^{track} \times \frac{p_T^{calo}}{p_T^{track}}$
  - **Combined Mass** is a linear combination of the calo and TA mass:

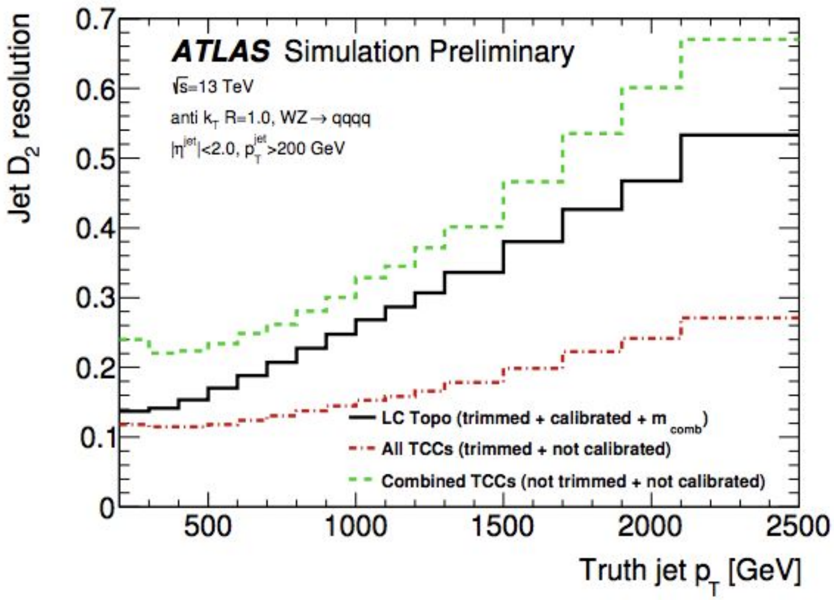


Jet Improvements at ATLAS

- Recent developments at ATLAS promise improvements in the mass reconstructions.
  - Track-Calo Clusters: Combine the excellent energy resolution of the Calorimeter with the angular resolution of the Tracker into PF like 4-vectors:



(a) Jet mass resolution, IQR^r

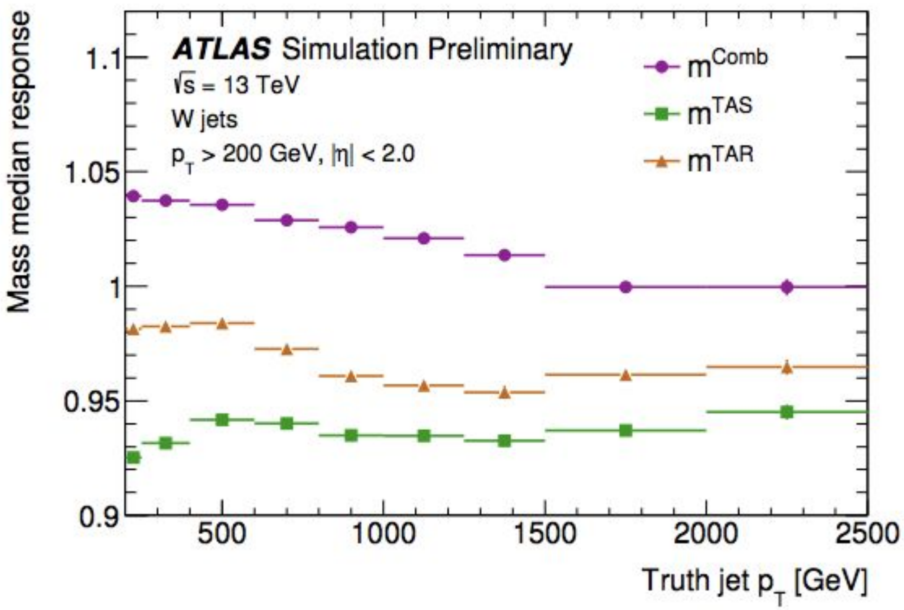


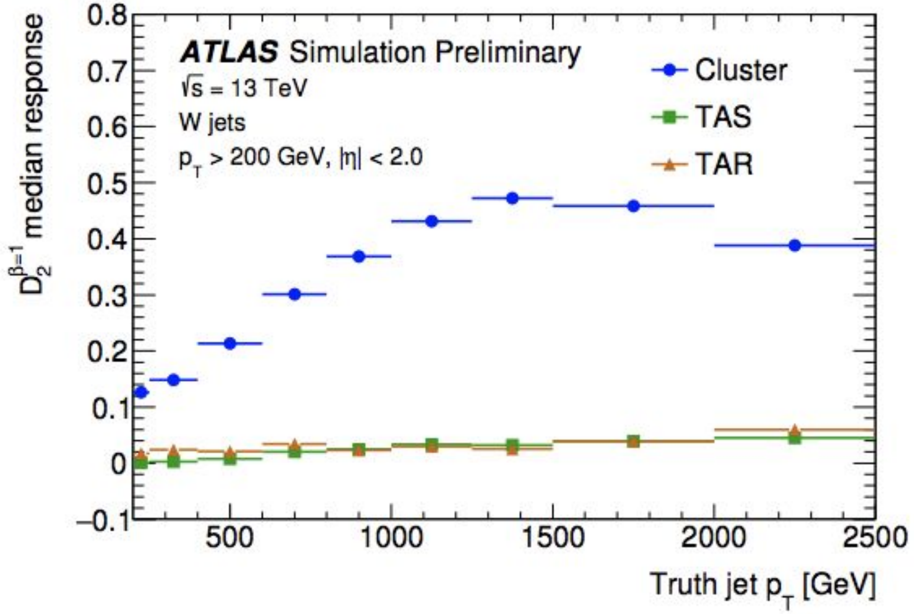
(b) Jet D_2 resolution, IQR^d

1: ATLAS Result [Improving jet substructure performance in ATLAS using Track-CaloClusters]

Jet Improvements at ATLAS

- Recent developments at ATLAS promise improvements in the mass reconstructions.
  - Track Assisted Reclustered Mass: Build jets (R = 0.2) clusters, and calibrate them as with other jets.
  - Recluster these to a large jet, match and rescale tracks!

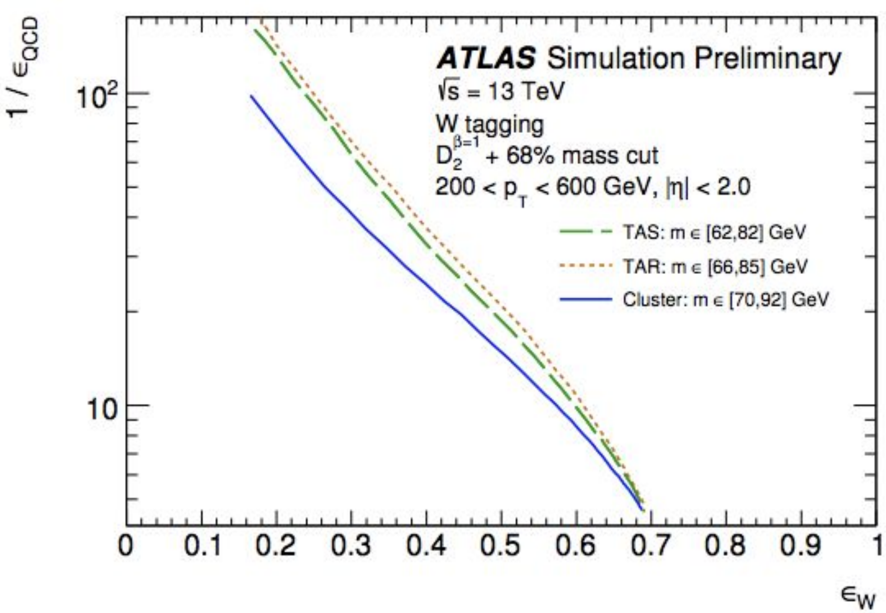


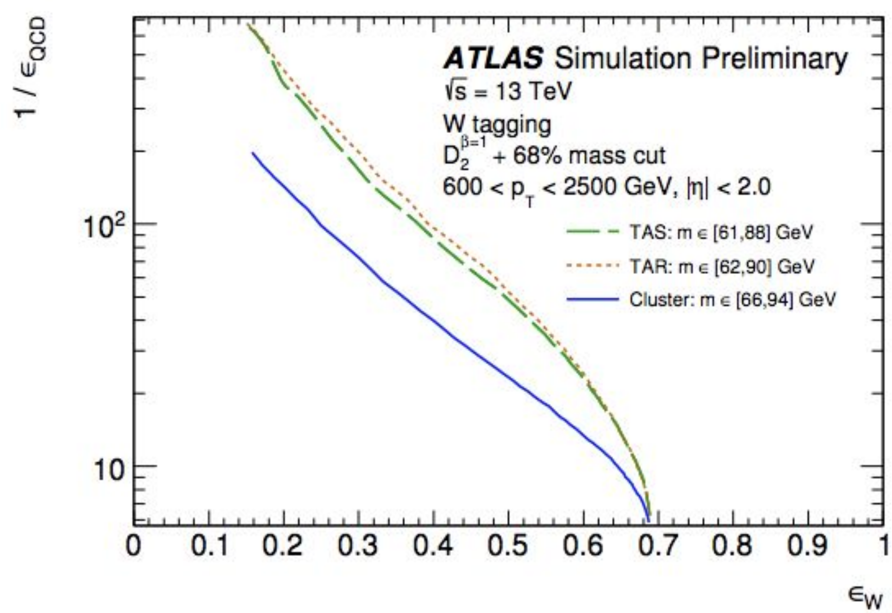


1: ATLAS result Track assisted techniques for jet substructure

Jet Improvements at ATLAS

- Recent developments at ATLAS promise improvements in the mass reconstructions.
  - Track Assisted Reclustered Mass: Build jets (R = 0.2) clusters, and calibrate them as with other jets.
  - Recluster these to a large jet, match and rescale tracks!



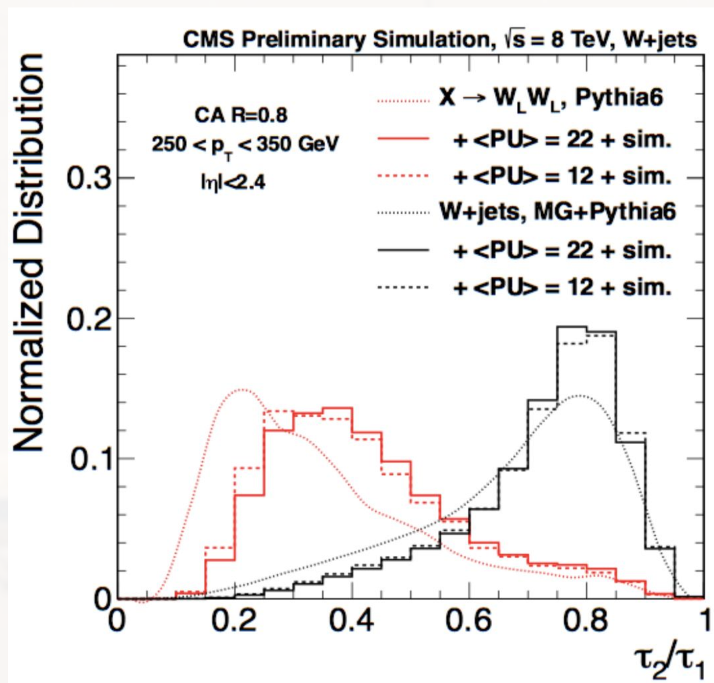


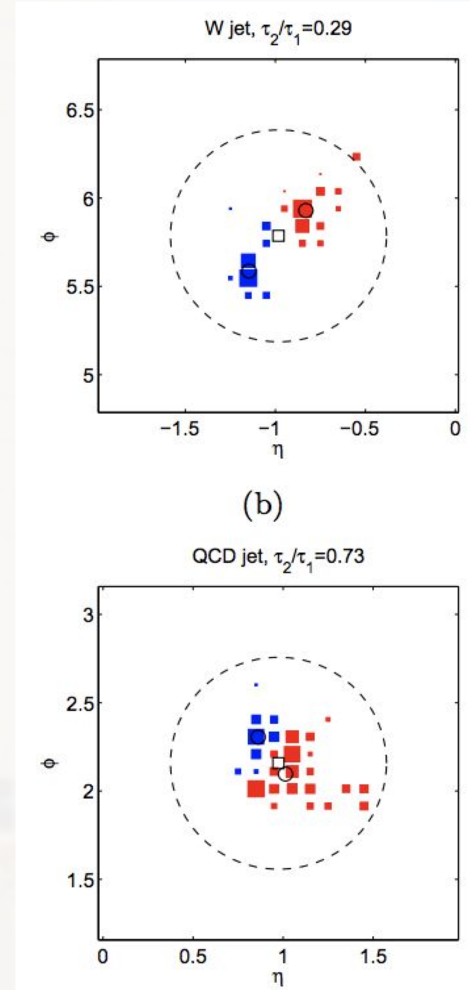
1: ATLAS result Track assisted techniques for jet substructure

- Dominant substructure variable for Higgs searches at CMS is the so-called *N-subjettiness*[1].

$$\tau_N \equiv \frac{1}{d_0}\sum_k p_{T,k} min(\Delta R_{1,k}, \Delta R_{2,k}, \cdots \Delta R_{N,k},)$$

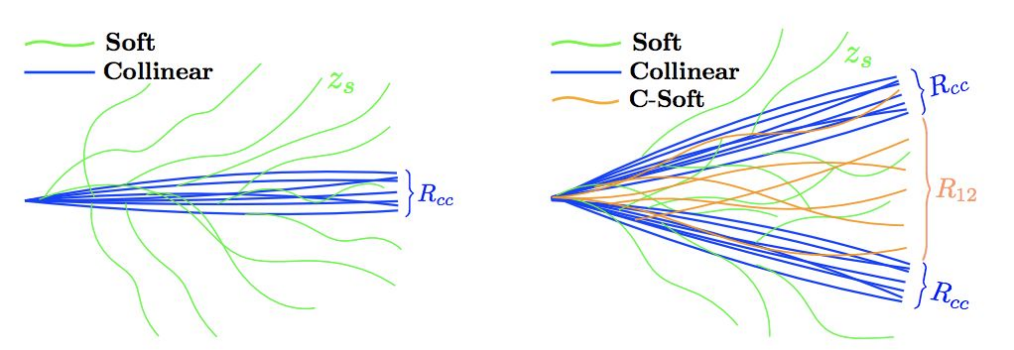
- Ratio of τ_N variables serve as strong discriminants.
- Can be correlated to kinematic properties of the event.

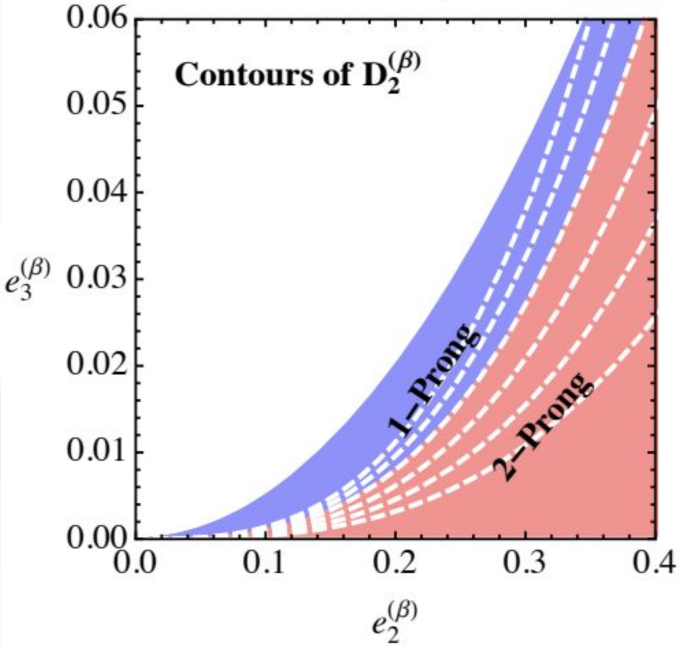




Substructure at ATLAS

- Different approach: Energy Correlation Functions
- Better theoretical motivation
- Similar discrimination





$$e_2^{(\beta)} = \frac{1}{p_{TJ}^2} \sum_{1 \le i < j \le n_J} p_{Ti} p_{Tj} R_{ij}^{\beta} ,$$

$$e_3^{(\beta)} = \frac{1}{p_{TJ}^3} \sum_{1 \le i < j < k \le n_J} p_{Ti} p_{Tj} p_{Tk} R_{ij}^{\beta} R_{ik}^{\beta} R_{jk}^{\beta}$$

$$D_2^{(\beta)} \equiv \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3}$$

1: paper Power counting to better jet observables

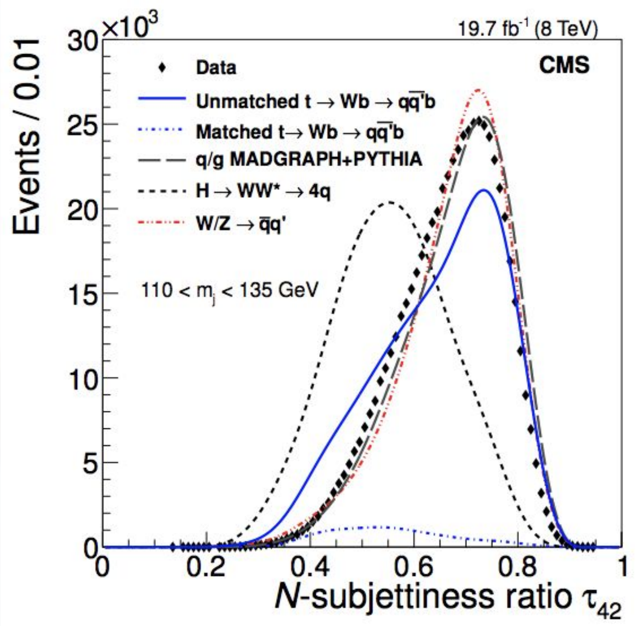
Exotic Jet Topologies

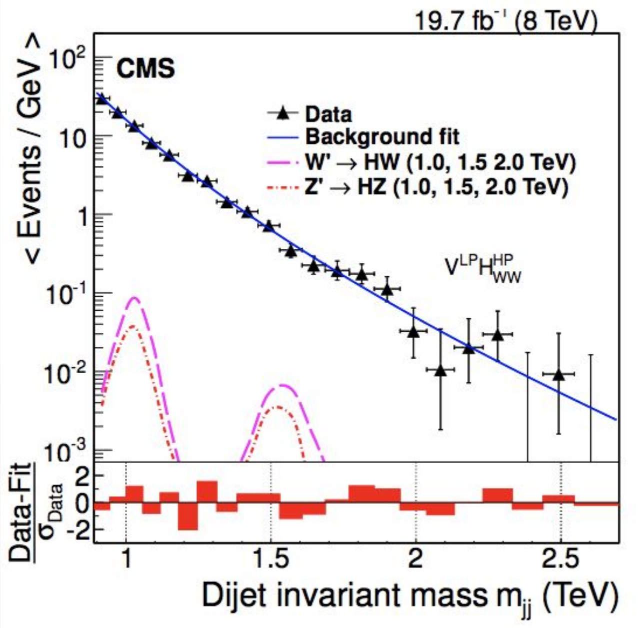
- $H \rightarrow bb$ isn't the whole story! Full exploration of HH signatures should consider other "exotic" decays of the SM higgs.
  - Already saw $WW^* \rightarrow qql\nu$[1]
  - ZZ^* and WW^* can go to 4q
    - Do we have the technology to reconstruct this: **yes**[2] (and we have had for some time)
  - τ_4 used in recent CMS result (SUSY signature)[3]

Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at $\sqrt{s} = 8$ TeV

1: https://indico.cern.ch/event/731450/timetable/#41-dedicated-object-reconstruc
2: Link to JHEP paper 3: CMS Result (submitted) https://arxiv.org/pdf/1806.01058.pdf

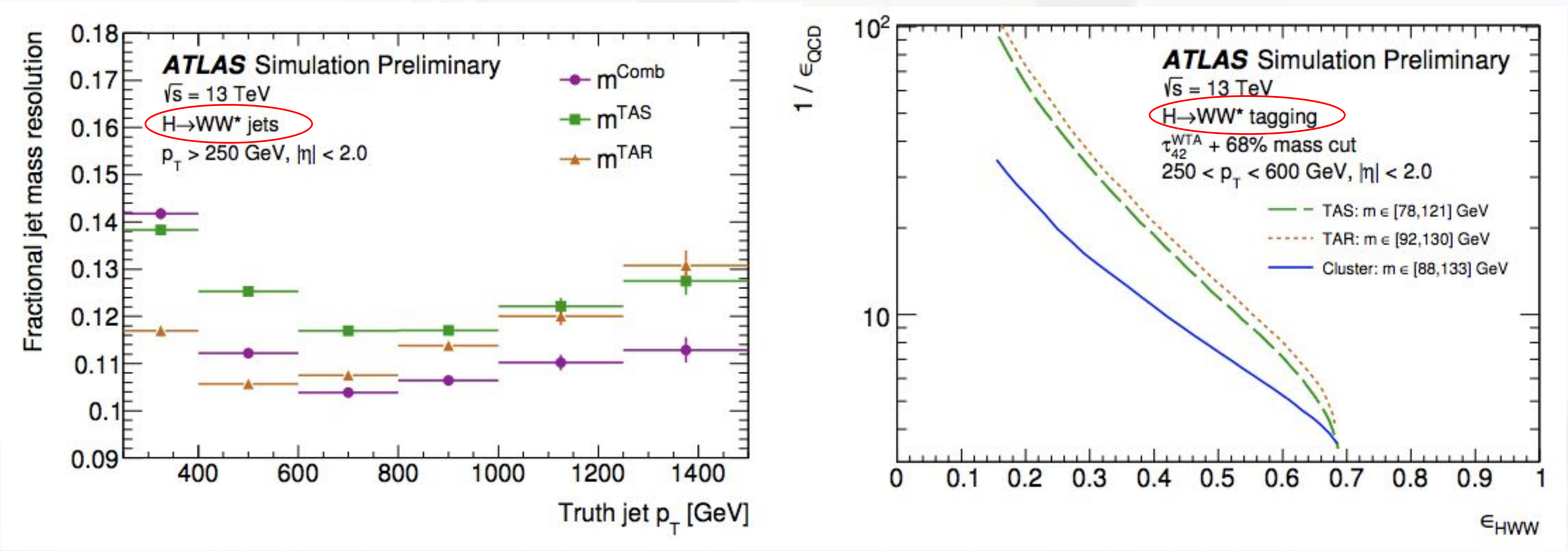
- Search for X → VH, considers all hadronic H → bb/4q
  - H → bb dominates significance at low masses, but WW* category contributes at high masses where background are naturally lower.
  - Uncertainties for four-pronged variables were large.
- Equivalent HH signatures not public, but in the works.





Exotic Jet Topologies

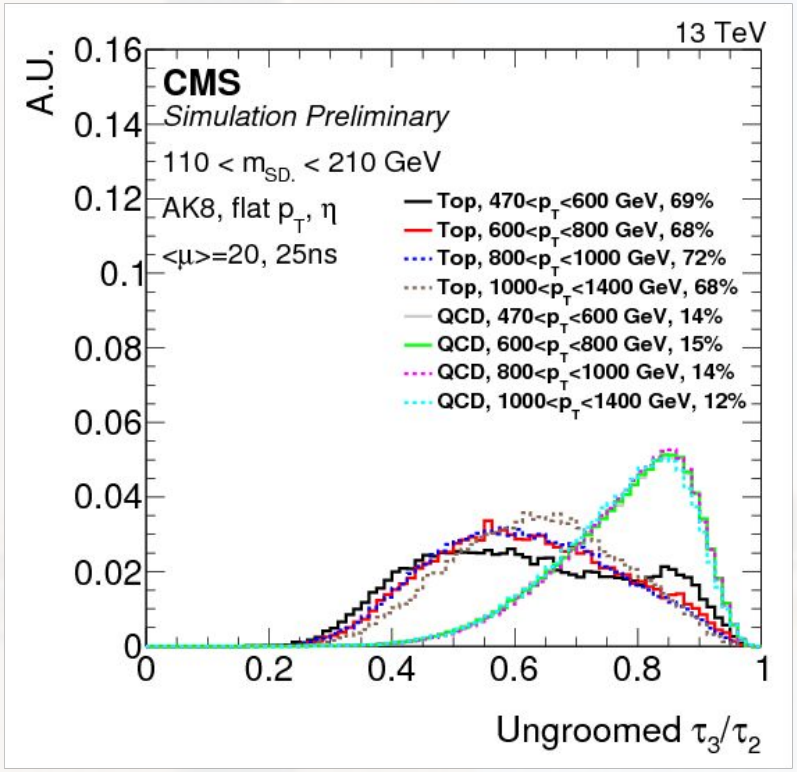
- While no recent public results, such decays are on the experimental radar! ATLAS's Track Assisted Reclustering evaluated the effect on H→ WW* jets.



1: ATLAS result Track assisted techniques for jet substructure

Lessons from Top Tagging

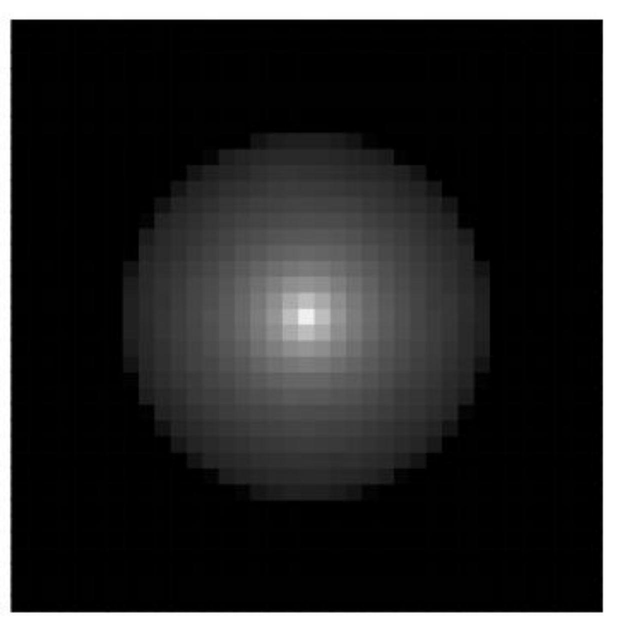
- Considerable improvements in substructure tagging are possible with new machine learning techniques.
- Take the example of top tagging[1]:
- CMS uses N-subjetiness variables coupled with soft-drop mass to identify boosted hadronic Tops -- Very similar to H-tag
- Further improvements possible with dedicated strategies, e.g. HEP Top Tagger (HTT)



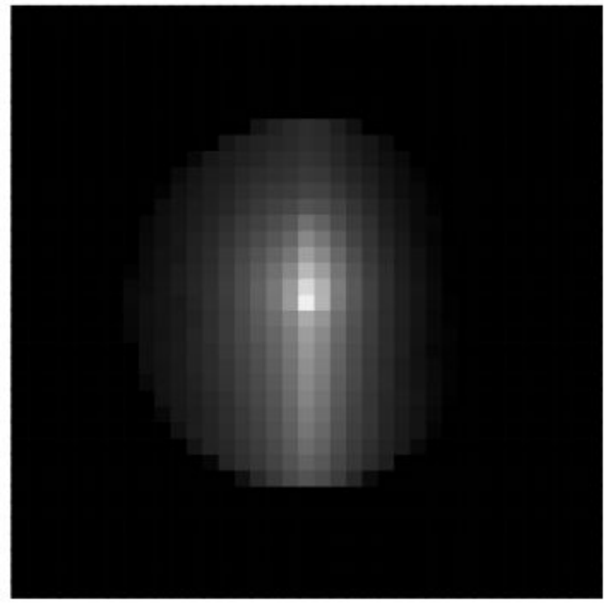
1: paper Pulling Out All the Tops with Computer Vision and Deep Learning
2: figure Top Tagging with New Approaches

Lessons from Top Tagging

- Considerable improvements in substructure tagging are possible with new machine learning techniques.
- Take the example of top tagging[1]:
  - "Jet Image" based CNN to separate QCD from top decays



100,000 QCD Jet Images

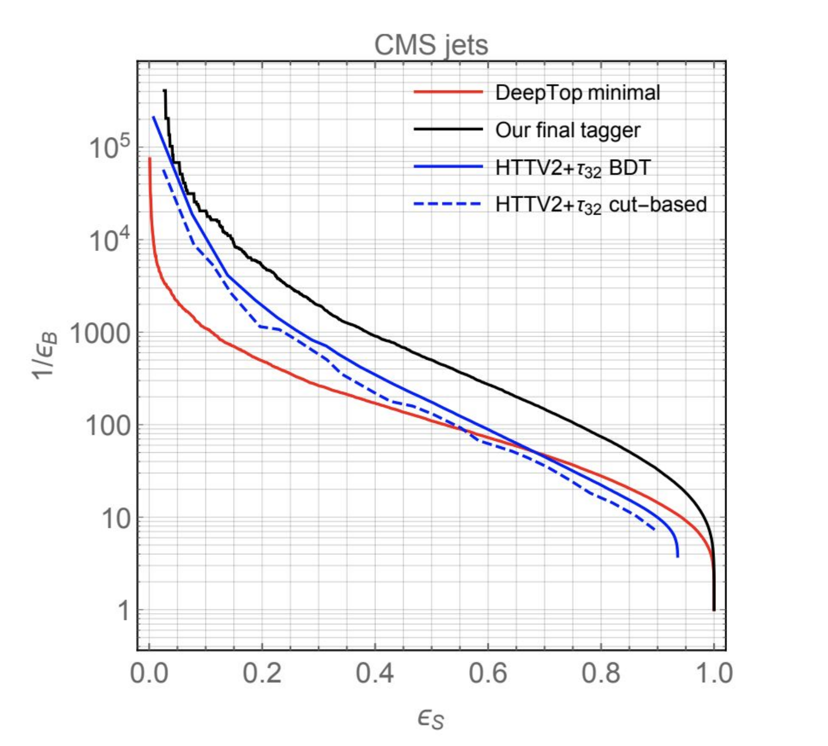


100,000 Top Jet Images

1: paper Pulling Out All the Tops with Computer Vision and Deep Learning

Lessons from Top Tagging

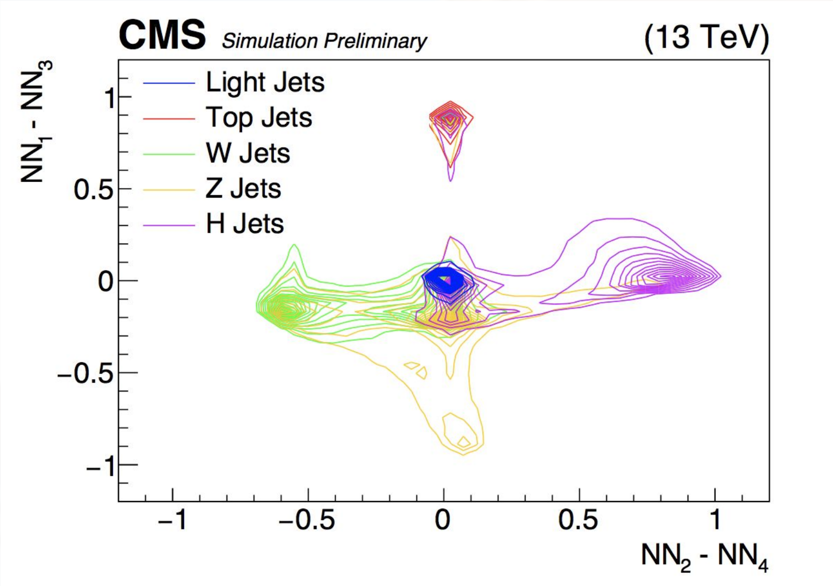
- Considerable improvements in substructure tagging are possible with new machine learning techniques.
- Take the example of top tagging[1]:
  - Large improvement, even without b-tagging information!

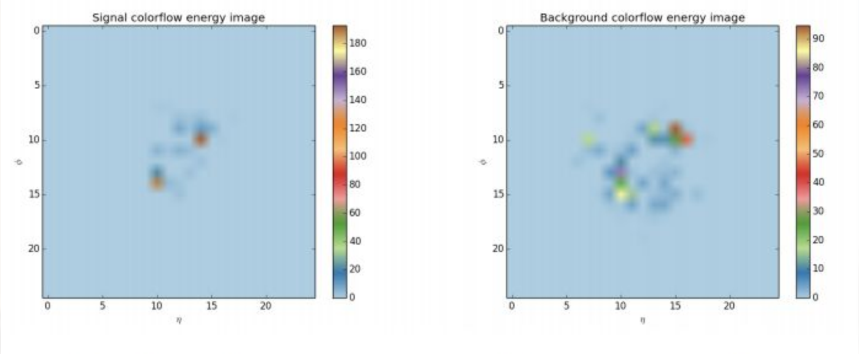


1: paper Pulling Out All the Tops with Computer Vision and Deep Learning

ML for Higgs Tagging

- There are already a number of improved Higgs results, though none currently implemented in public CMS/ATLAS results:
  - Multi-taggers[1]
  - CNN approaches similar to the top taggers[2]
  - New? Seems like we are only scratching the surface!





(a) Signal image (b) Background image

1: CMS DP Note New Developments for Jet Substructure Reconstruction in CMS
2: Report Identifying the Higgs Boson with Convolutional Neural Networks

Challenges in H-tagging

- We do not have a sample of boosted Higgses in data to calibrate our techniques and taggers. Have to extrapolate:
  - g → bb and V decays currently used.
  - Some talk of t → bW+FSR for a 4-pronged tagger.
- Result are large uncertainties on the signals.
- ML could provide some solutions.

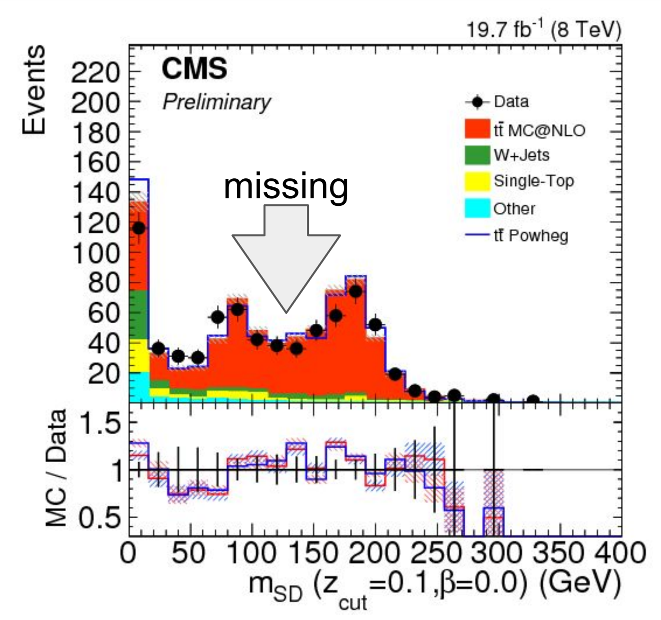


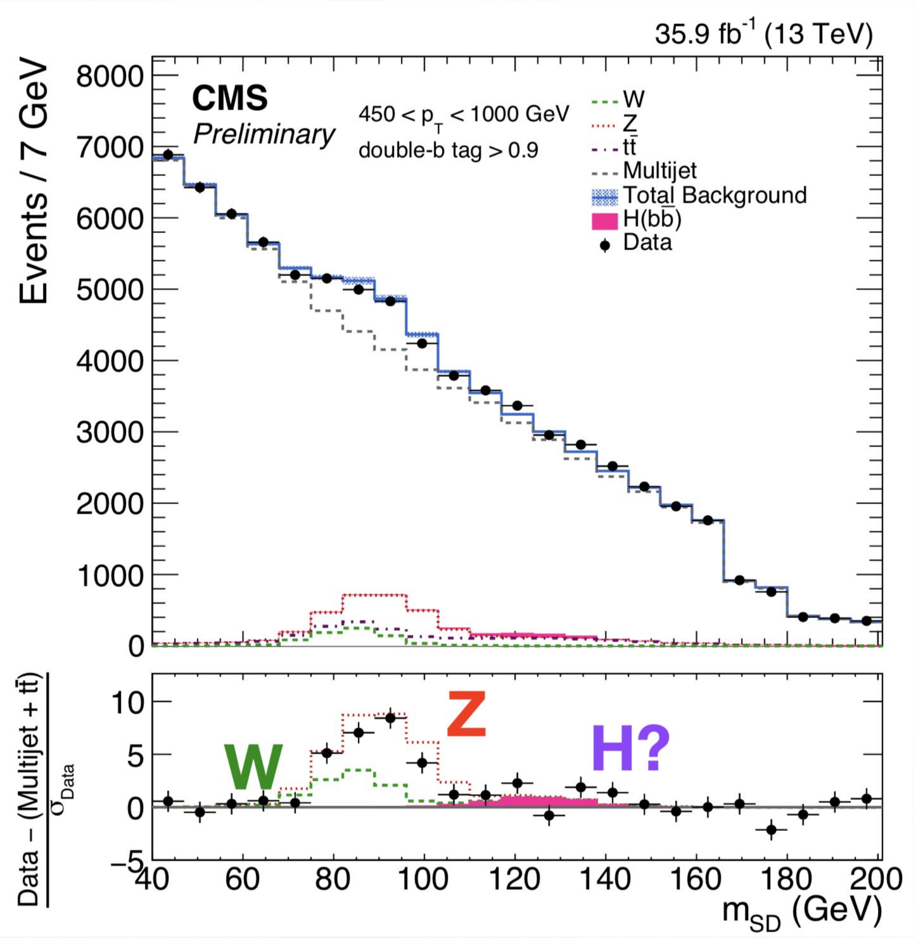
Table 2: Summary of systematic uncertainties in the signal and background yields.

Source	Uncertainty (%)
Signal yield	
Trigger efficiency	1–15
H jet energy scale and resolution	1
H jet mass scale and resolution	2
H jet τ_{21} selection	+30 / −26
H-tagging correction factor	7–20
Double-b tagger discriminator	2–5
Pileup modelling	2
PDF and scales	0.1-2
Luminosity	2.5
Background yield	
$R_{p/f}$ fit	2.6 (LL category) 6.8 (TT category)

1: figure Top Tagging with New Approaches
2: table Search for a massive resonance decaying to a pair of Higgs bosons in the four b quark final state in pp collisions at 13 TeV

Conclusion

- It's a great time to use substructure for HH searches!
  - Current tools are mature, tested, working.
  - New tools are being developed. Lots of room for improvement. Many avenues to explore.
  - Many channels never fully explored.



Thank you