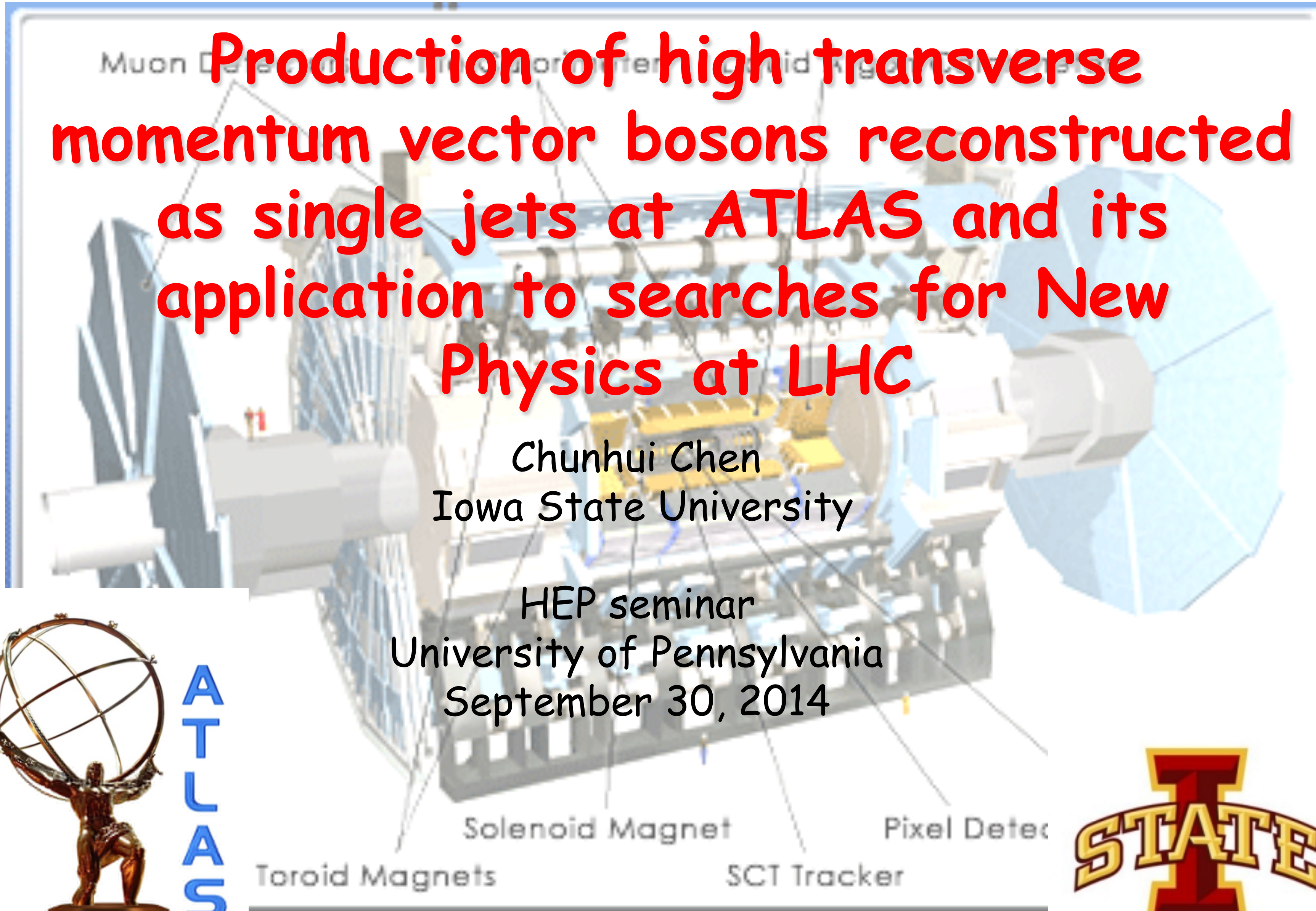
Production of high transverse momentum vector bosons reconstructed as single jets at ATLAS and its application to searches for New Physics at LHC

Chunhui Chen
Iowa State University

HEP seminar
University of Pennsylvania
September 30, 2014

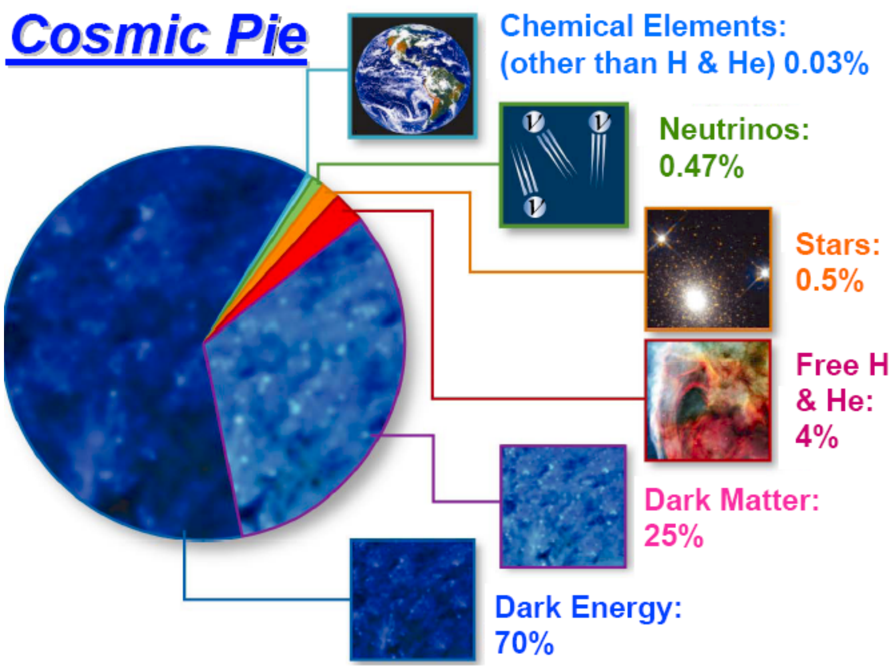


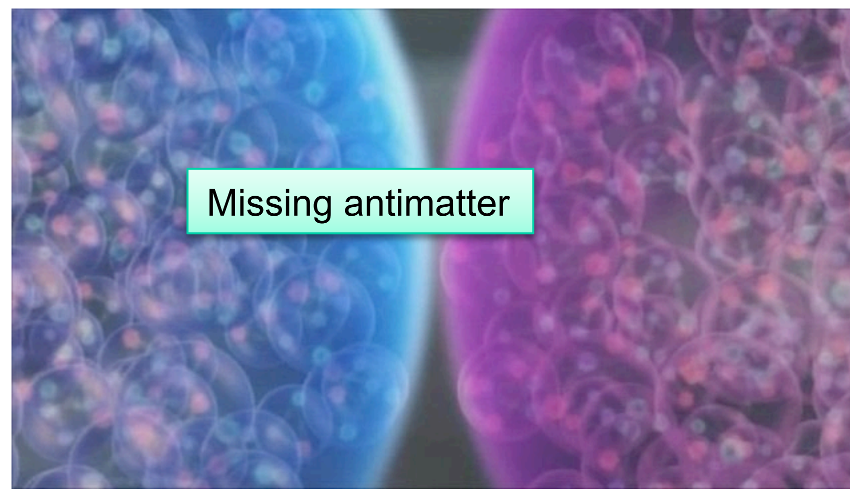
Outline

- Introduction
  - ✓ Boosted hadronically-decaying particle (W, Z and top)
  - ✓ Jet substructure and searches for new physics (NP)
  - ✓ Jet substructure in the center-of-mass frame of jet
- Measurement of boosted hadronic W/Z boson production at ATLAS
- Future application of jet substructure in jet rest frame

The LHC, the experiments and the observation of a Higgs-like boson is a global phenomena

Open questions in particle physics

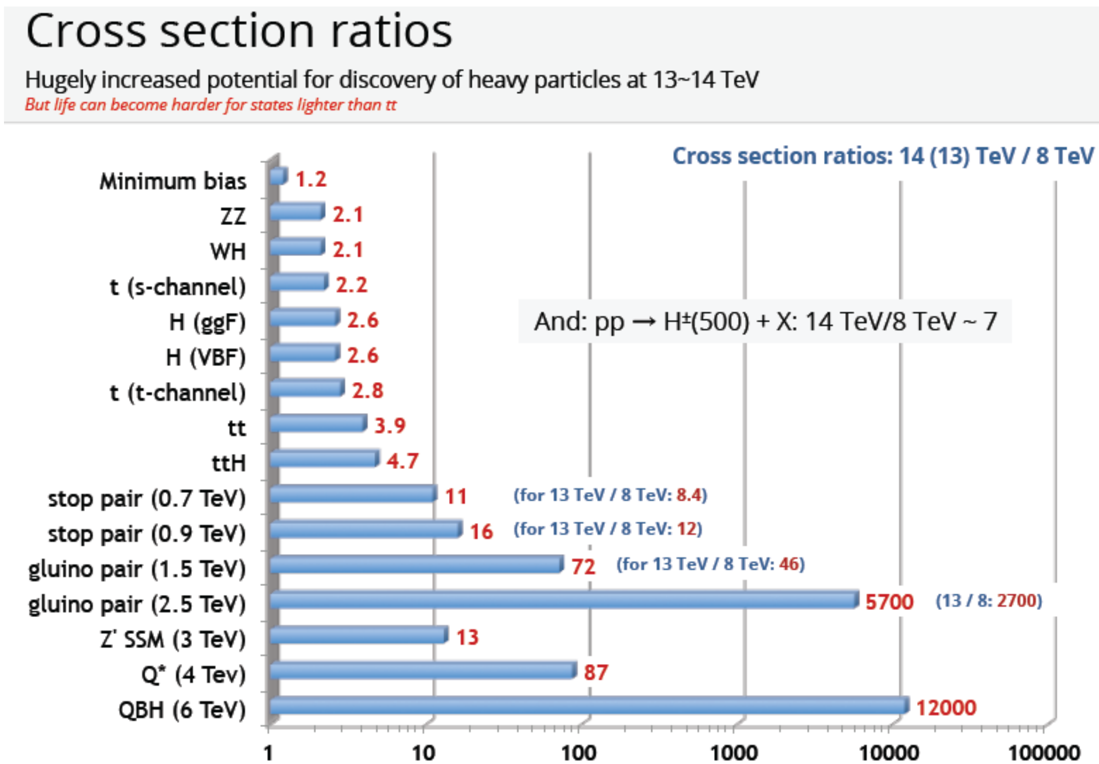




Higgs Discovery at the LHC is just the beginning of an exciting (discovery of new physics) era in high energy physics !

How to look for new physics at LHC

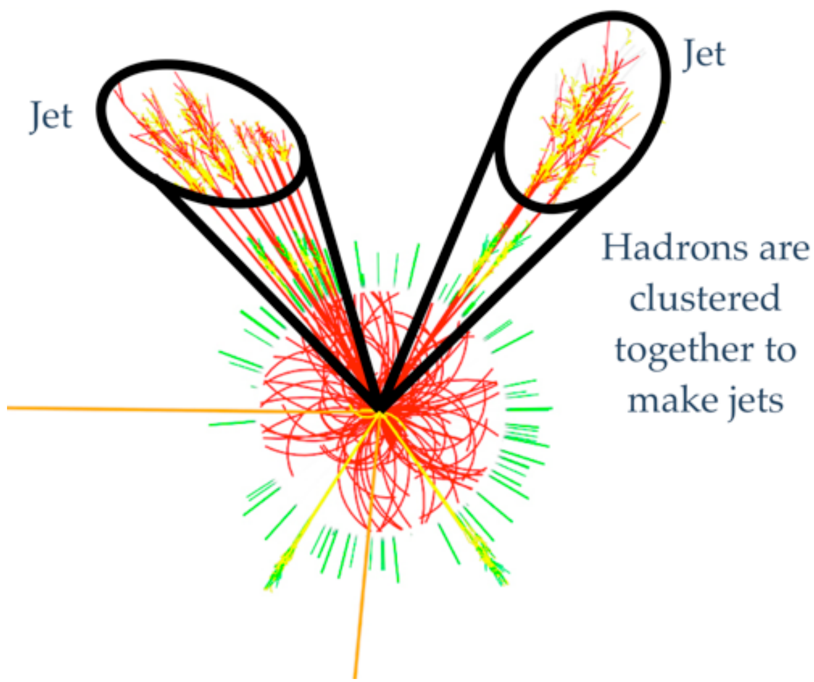
- Repeat searches for NP done in the past (Tavetron, LEP etc)
  - ✓ Well established/sophisticated analysis techniques
  - ✓ Higher production cross section of many NP particles
  - ✓ Higher luminosities (more data)



New experimental analysis techniques, ideas and tools to improve our odds?

Boosted hadronically decaying particle: jet mass & substructure

- A new analysis idea/techniques: boosted object (jet)
  - ✓ Significantly improve sensitivities to search for heavy new particles
    - In the decay final states containing W, Z, Higgs or top quarks
  - ✓ Generate significant theory and experimental interested in LHC
    - Many new theoretical and experimental papers on the subject
    - Annual workshop devoted to boosted object since 2009



Similar to use the charged tracks to identify “stable” particles at lepton & hadron collider

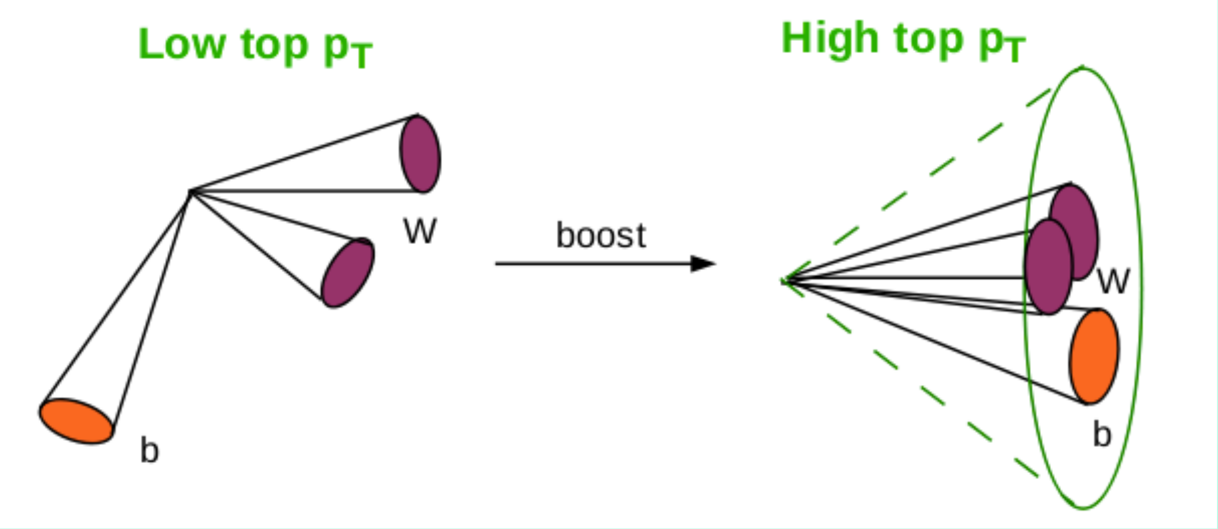
Jet: a collection of energy clusters deposited in calorimeter detector as an experimental signature of the initial parton (quarks & gluons) at the hardon collider

Boosted hadronically decaying particle

- Most NP models predict heavy resonance (~ TeV) decay into W/Z/Top:

$$X \to WW, WZ, t\bar{t}$$

  - ✓ Boosted (high p_T) jets in the final decay states
  - ✓ The hadronic decay product are highly collimated



- Traditional jet reco. relying on one-to-one jet-to-parton assignment not adequate

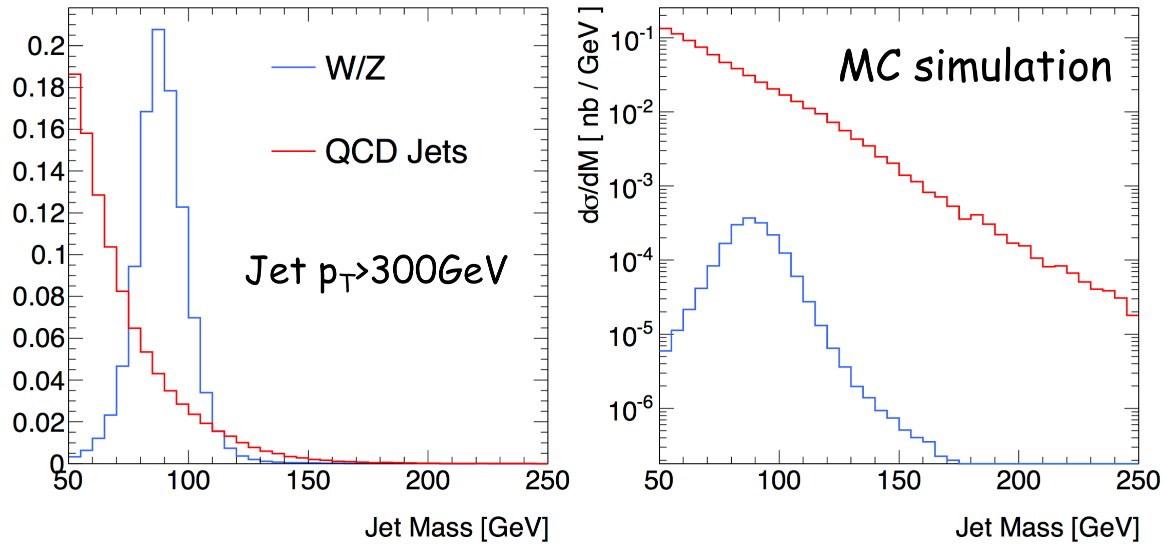
- Solution: reconstruct multi quarks in a single jet (boosted W/Z/Top jet)
  - ✓ 2 quarks for W/Z decay, 3 quarks for top decay
  - ✓ Better (only way) to search for certain NP models

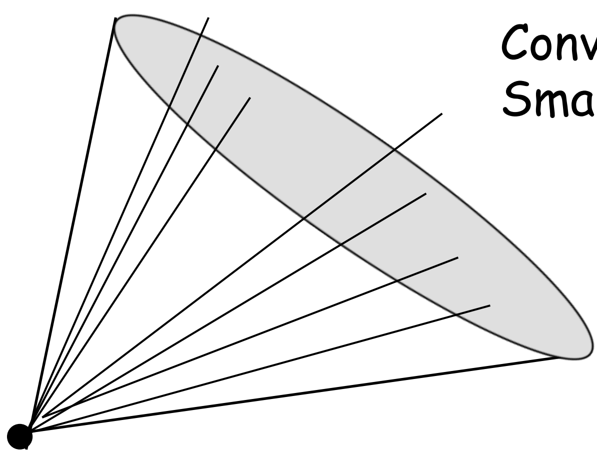
How to identify a Boosted W/Z/t jet

Jet mass: invariant mass of jet

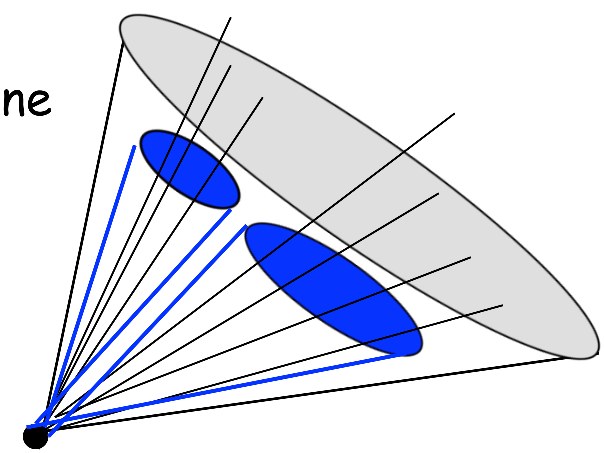
Problem: QCD jet (1 non top quark or gluon) has non-zero mass, its production a few orders higher !

Solution: jet substructure, A active research area in last a few years





Conventional method:
Smaller jet inside big one

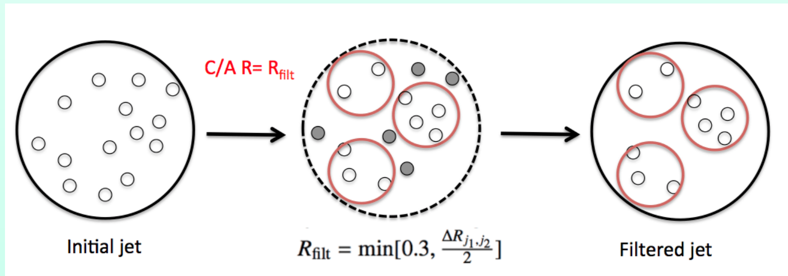


Exam the jet energy cluster information in the lab frame

Popular jet substructure methods

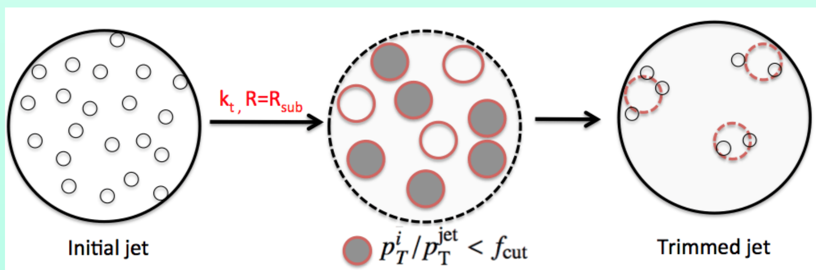
- Jet shape variables: N-subjettiness, momentum balance ……
- Jet grooming: 3 related techniques to reinterpret the jet constituents to improve jet substructure resolution, reduce background and impact of underlying event & pile-up

✓ Filtering:



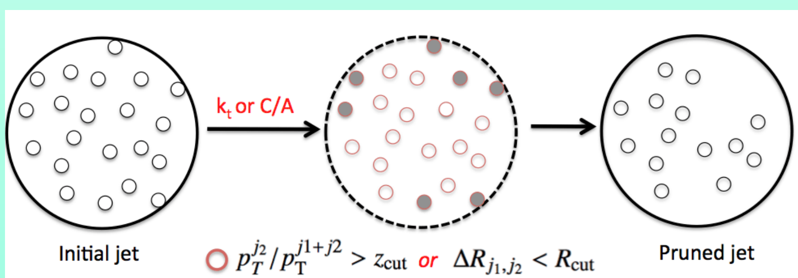
Remove constituents that are outside of subjets

✓ Trimming:



Compares p_T(constituents) with p_T(jet) – removes soft components which are primarily from UI & PU

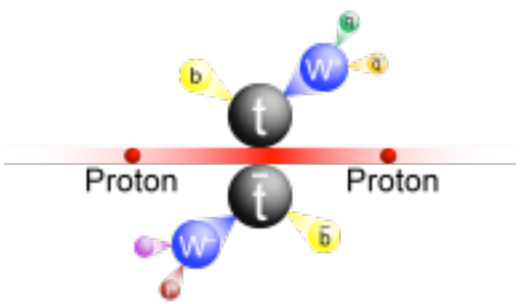
✓ Pruning:

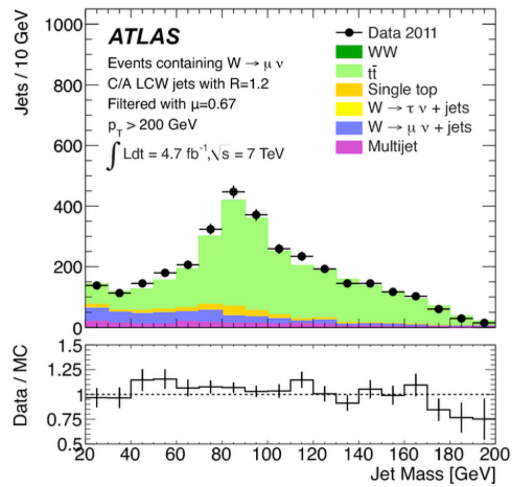


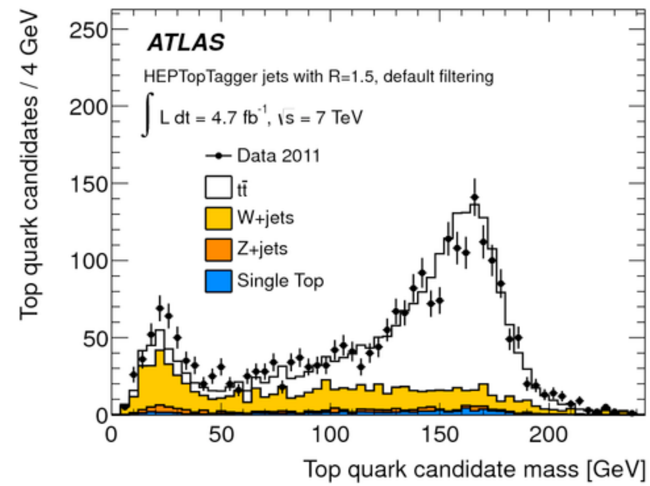
Similar to trimming but occurs during jet reconstruction ⇒ does not require subjet reconstruction

Pedagogical intro: arXiv:1302.0260

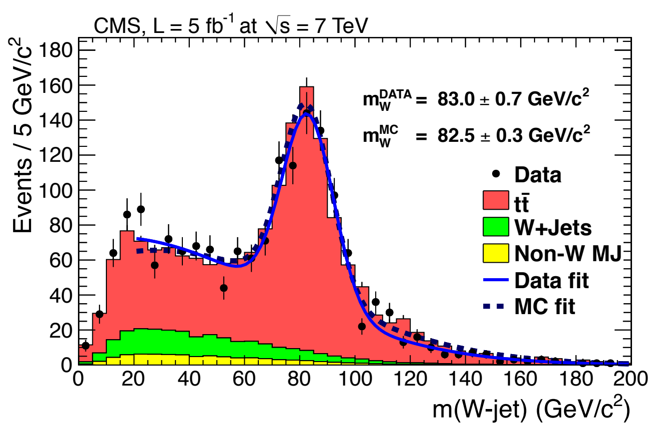
Hadronic W and top signal at LHC

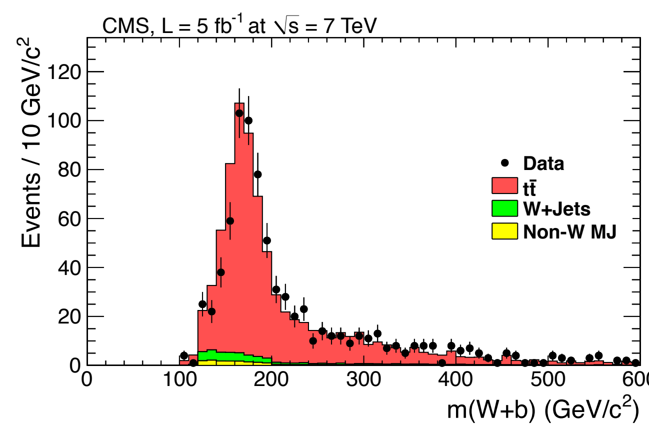






Standard jet size
ATLAS: 0.4(0.6)
CS: 0.5(0.7)



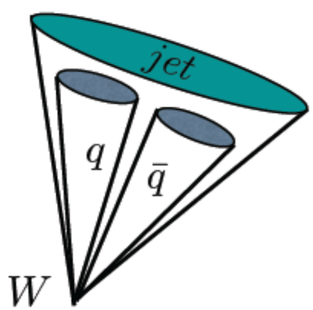


Use jet with large cone size: not exactly a highly boosted single jet
Need lepton and b jet requirement to reduce QCD jet bg from multijet production

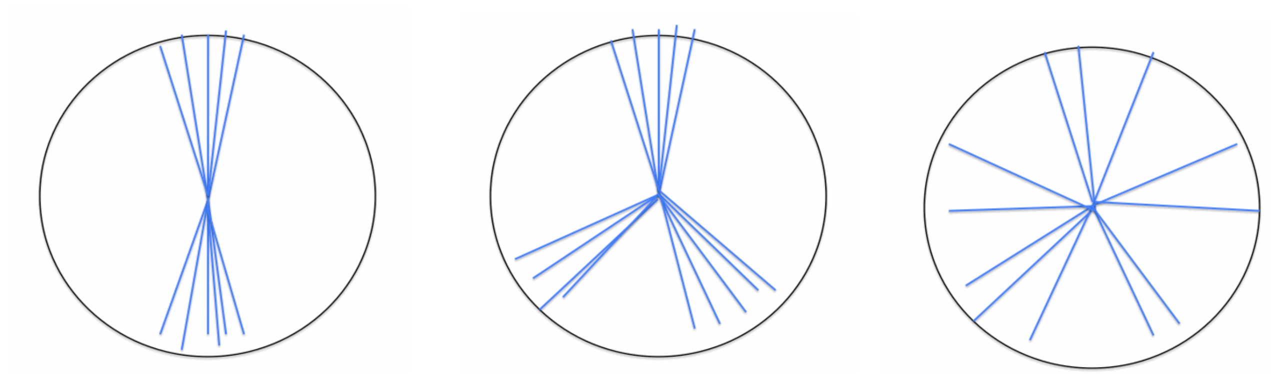
Can we extract boosted hadronic W/Z signal from QCD jet background
using jet substructure only

Jet Substructure in CM Frame of the jet

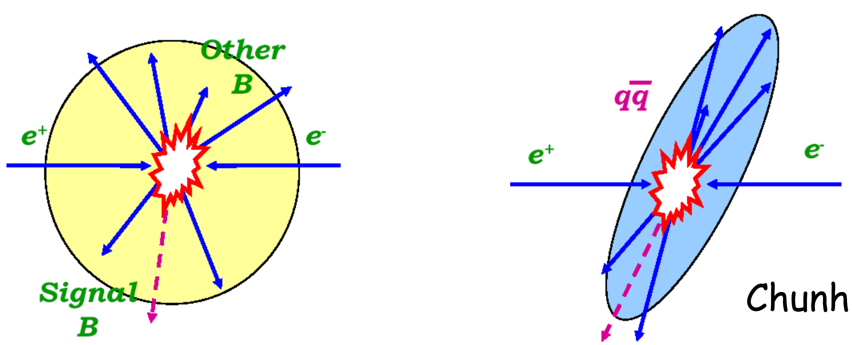
- Existing jet substructure algorithms based on energy clusters in lab frame
- New proposal: study distribution of jet clusters in center-of-mass frame of jet
  - ✓ Jet CM frame: jet 4 momentum = $(0,0,0,m_{jet})$
  - ✓ Nearby clusters in lab frame may not be close in CM frame
  - ✓ Using full momentum information of the energy clusters



W/Z->jj | t->Wb->jjj | QCD

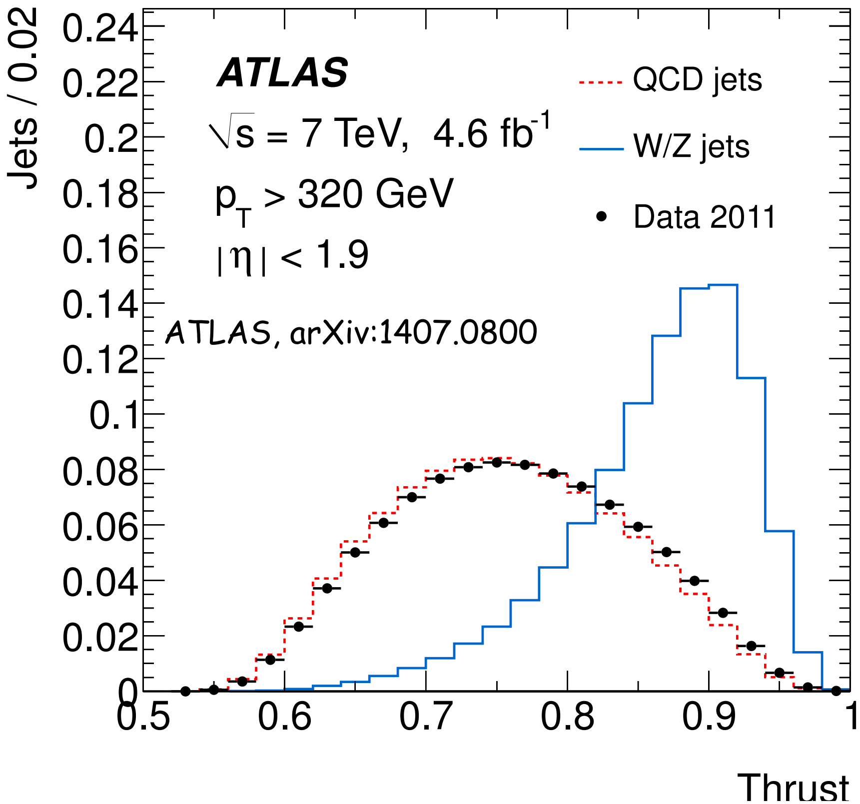


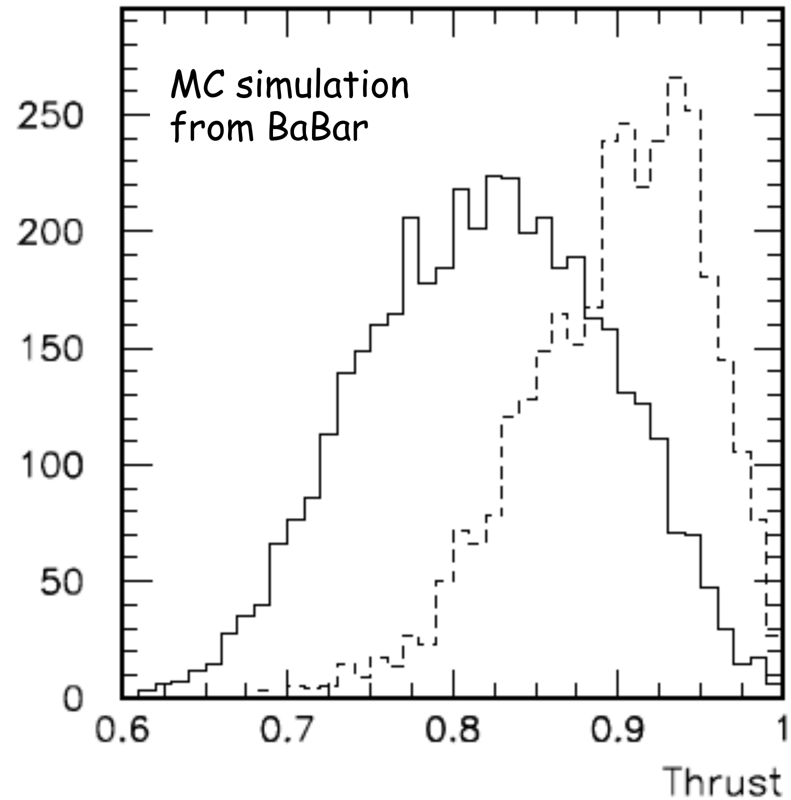
- Lesson learned from e^+e^- collider:



Chunhui Chen, PRD 85,052005 (2012)

Jet shape variables in jet rest farme



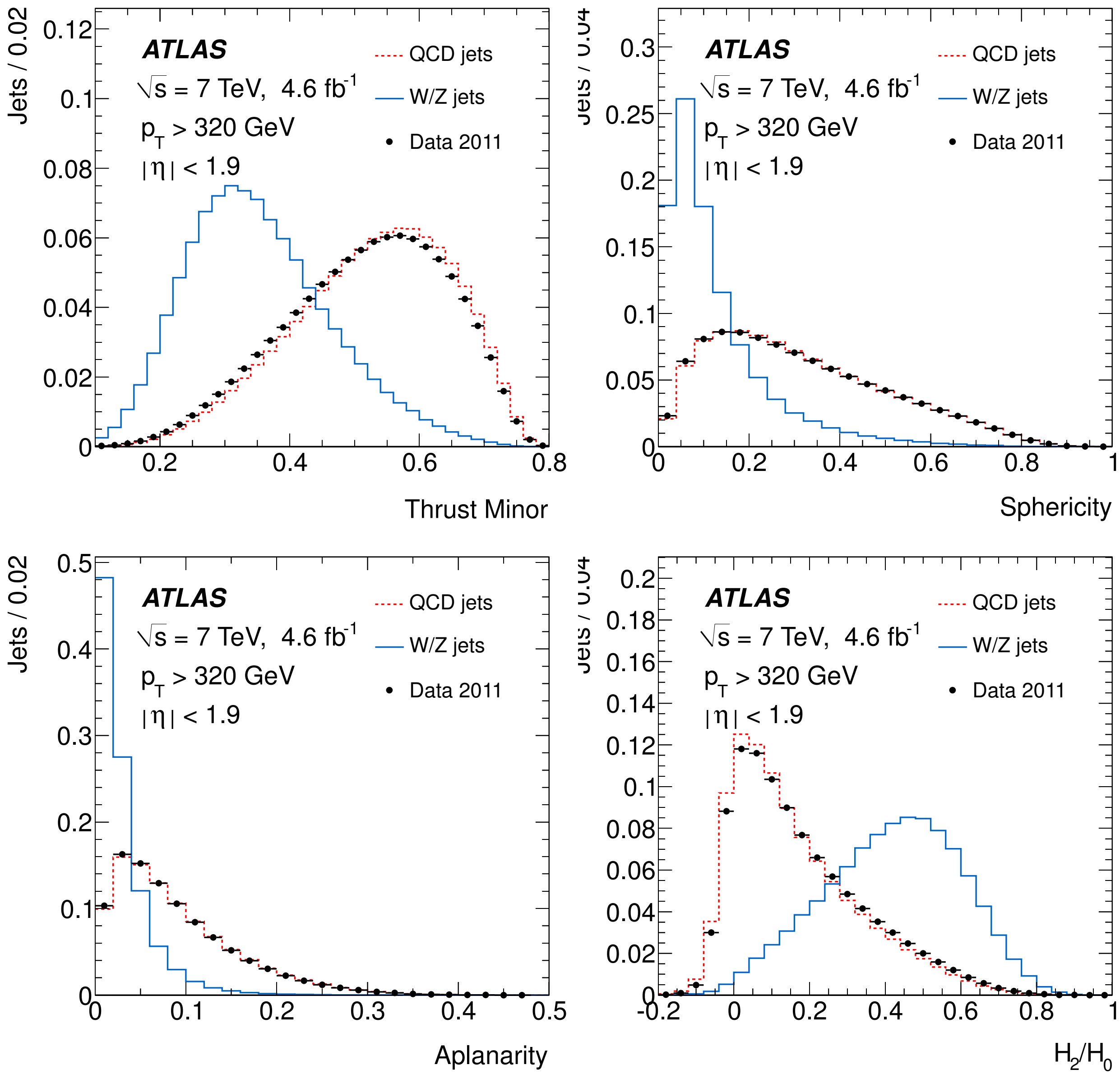


- Thrust: The thrust axis [39, 40] of a jet in its center-of-mass frame, $\hat{T}$, is defined as the direction which maximizes the sum of the longitudinal momenta of the energy clusters. The thrust, T, is related to this direction and is calculated as:

$$T = \frac{\sum_i |\hat{T} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}, \quad (2)$$

where $\vec{p}_i$ is the momentum of each energy cluster in the jet rest frame. The allowed range of T is between 0.5 and 1, where $T = 1$ corresponds to a highly directional distribution of the energy clusters, and $T = 0.5$ corresponds to an isotropic distribution.

Jet shape variables in jet rest farme



Commonly used in e^+e^- collider

Many are originally introduced at hadron colliders

Variables are correlated

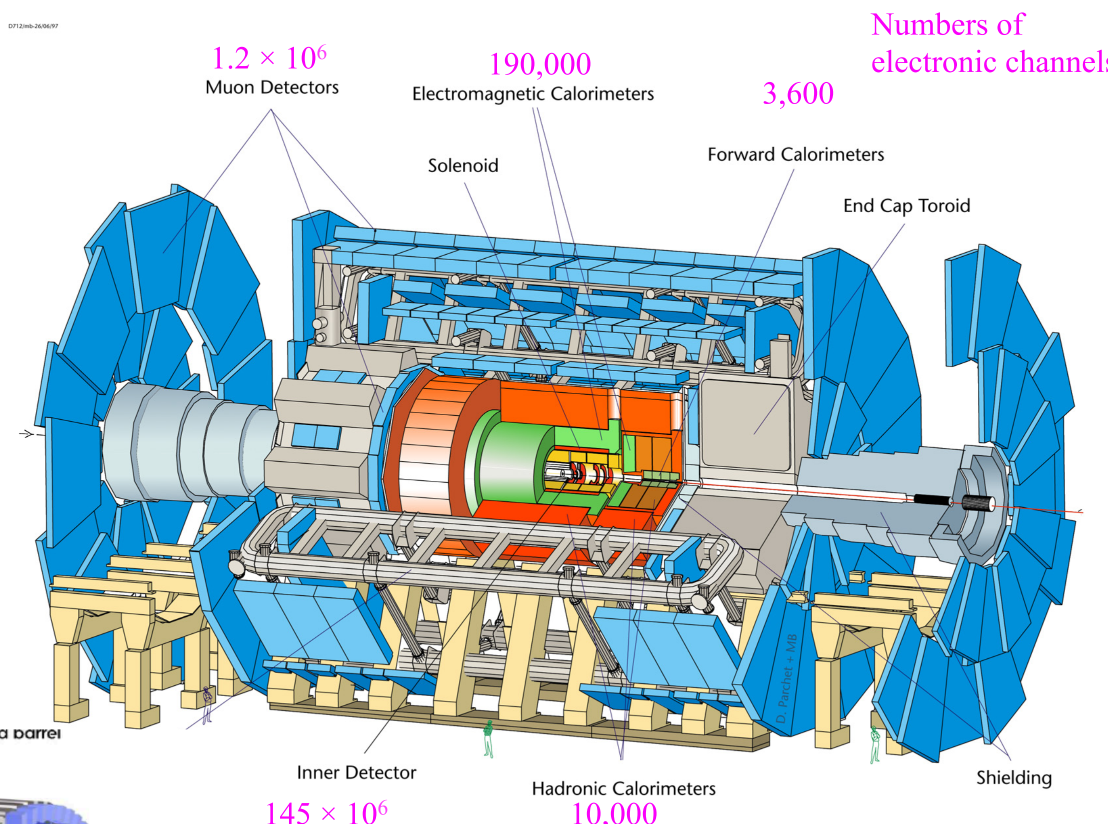
Correlation smaller than jet shape variables in lab frame

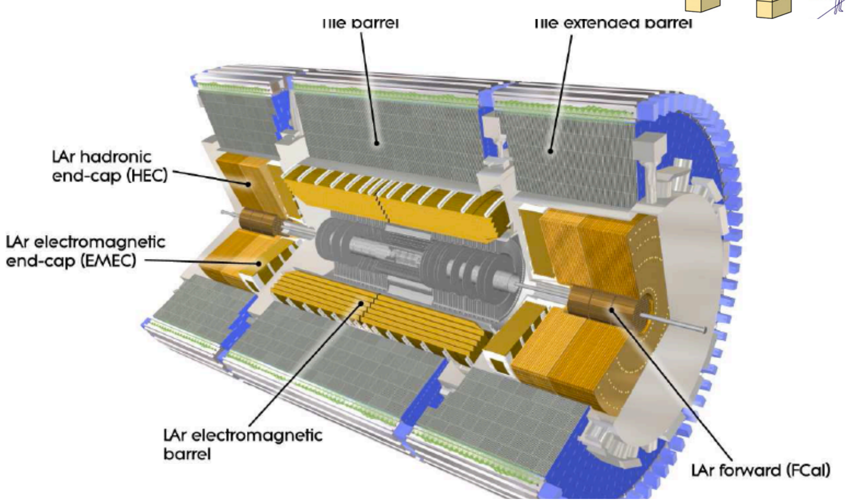
Original paper references can be found at: : C.Chen, PRD 85,052005 (2012)

Measurement of the production cross section of boosted hadronic W/Z reconstructed in single jets at 7TeV using the ATLAS detector

-- Using substructure in jet rest frame

The ATLAS Detector





Only calorimeter detectors used for the analysis: Reconstruction of energy deposition (clusters) deposited by hadrons , calculate jet momentum, jet mass, and jet substructure

Event selection

- Select events that pass the trigger requirements at 7TeV (4.6fb^{-1}):
  - ✓ A jet with p_T>100GeV at level-1
  - ✓ EF level : Scalar sum of jets (p_T>30GeV, |η|<3.2) larger than 350/400 GeV
  - ✓ 100% for offline signal selection

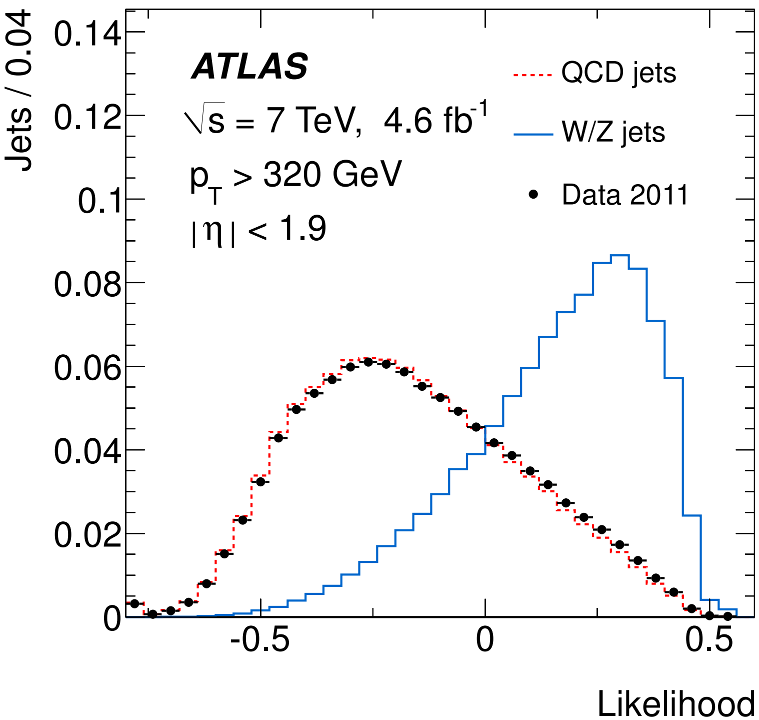
- Select jets as hadronic W/Z candidate (~2.5% multiple candidates/event)
  - ✓ Jet reconstructed with Anti k_T R=0.6
  - ✓ p_T>320GeV, |η|<1.9 and 50<m_{jet}<140GeV
  - ✓ Likelihood cut combining event shape variables in the CM frame to further reduce the QCD background: sphericity, aplanarity and thrust minor
    - Smaller correlation with jet mass
    - Conservative approach, more variables available

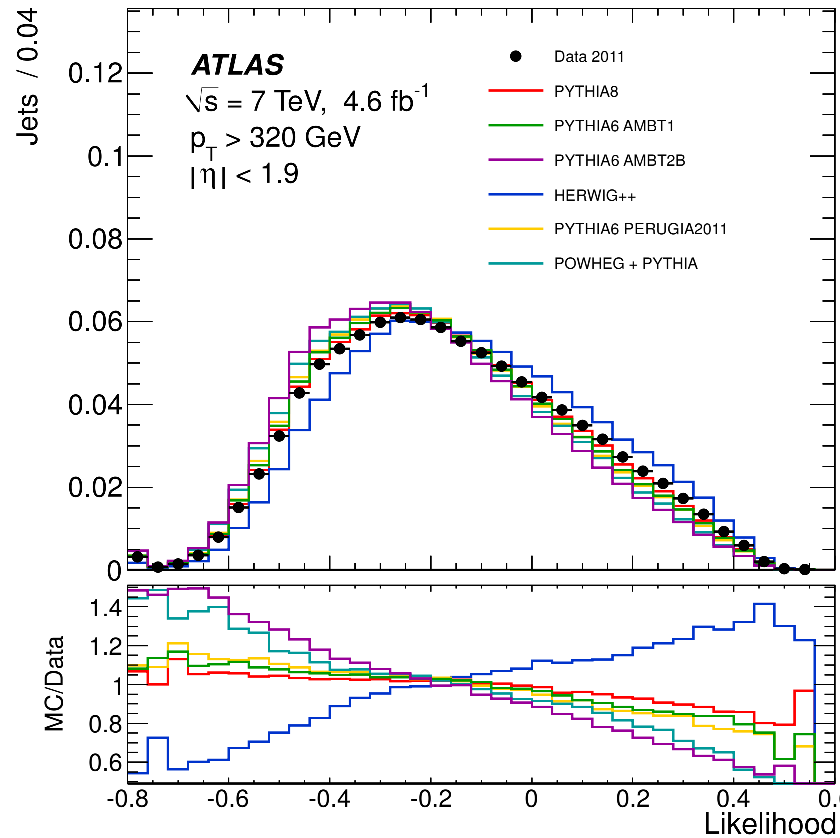
- Data driving analysis with MC as cross check:
  - ✓ Hadronic W/Z signal MC: Herwig, QCD jet bg MC: Pythia8
  - ✓ Many other MC used for cross check and sys error estimate

- Measurement: fiducial cross section of W and Z production
  - ✓ Extract hadronic W/Z signal by fitting m_{jet} distribution
  - ✓ Can statistically distinguish W & Z (large stat error due to m_{jet} resolution)

$$\sigma_{W+Z} = \sigma_W(p_{\rm T} > 320\,{\rm GeV}, |\eta| < 1.9) \times \mathcal{B}(W \to q\bar{q}') + \sigma_Z(p_{\rm T} > 320\,{\rm GeV}, |\eta| < 1.9) \times \mathcal{B}(Z \to q\bar{q})$$

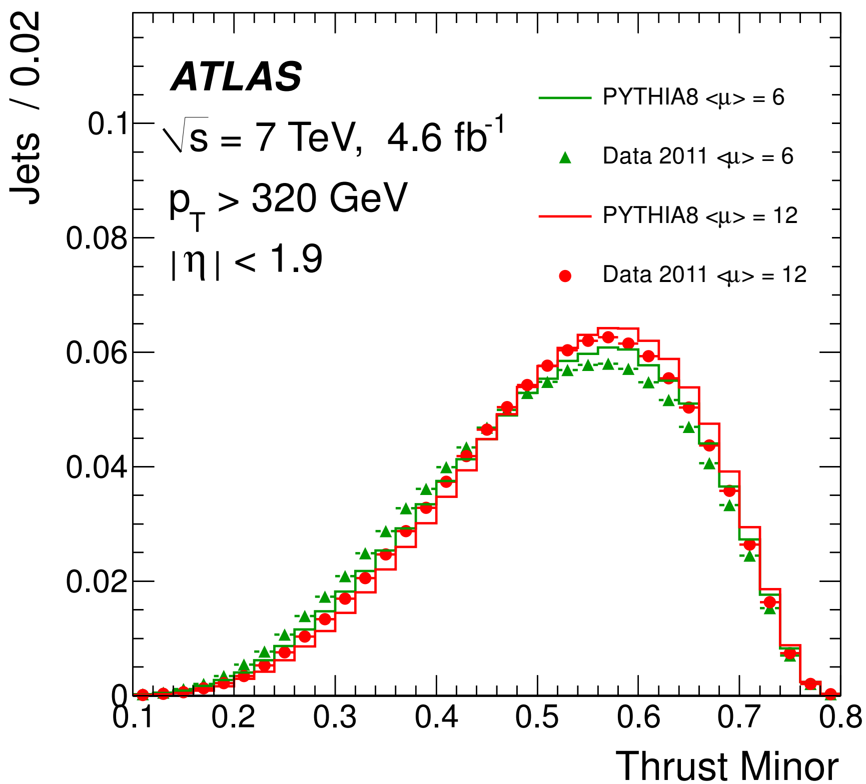
Likelihood variable selection

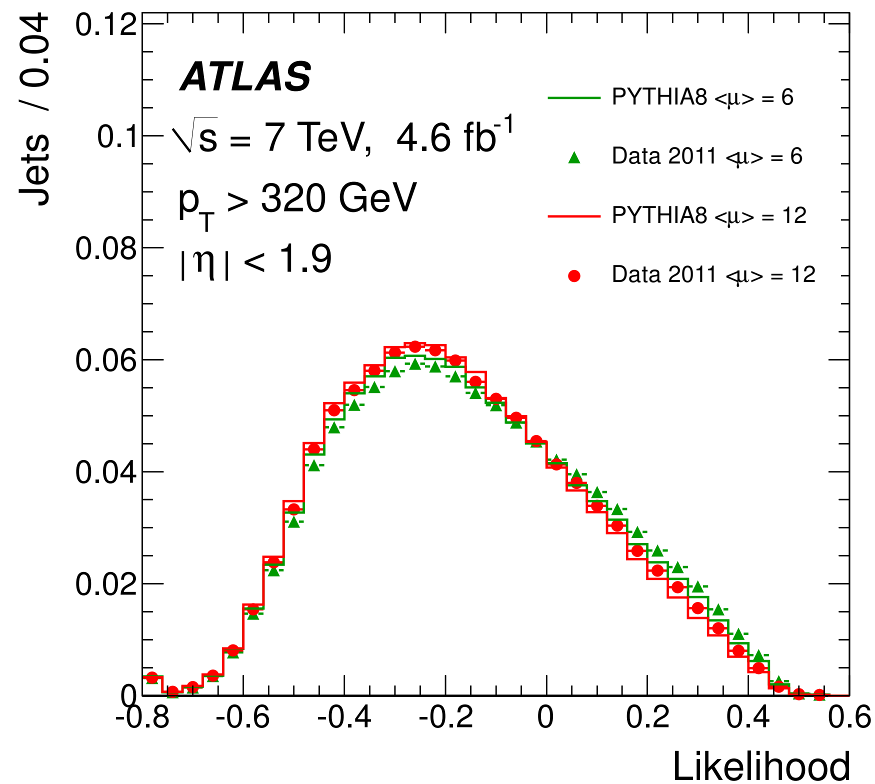




- Dominated by QCD jets after initial selection
  - ✓ Using LH variable to reject QCD background
  - ✓ Optimize to maximize S/sqrt(S+B): ~55% signal eff & reject ~90% bg
- Distribution well modeled by default MC simulation
  - ✓ Different MC simulation to estimate sys uncertainty of the analysis

Jet substructure with different pileup





- Some dependence of pileup conditions
- Well modeled by the MC simulation

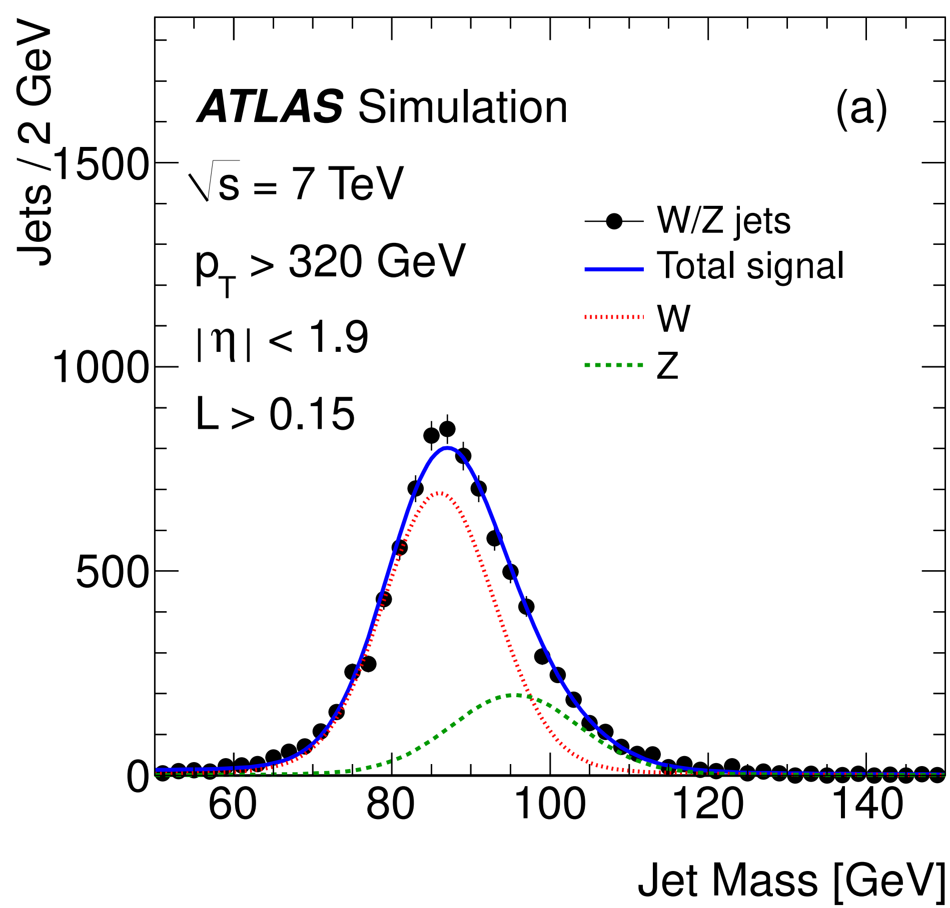
Signal PDF

- Extract hadronic W/Z signal by fitting m_{jet} distribution
- Probability density function of hadronic W/Z signals
  - ✓ Two Breit-Wigner functions convoluted by Gaussian functions

$$S_W(m_{\rm jet}) = F_{\rm BW}(m_{\rm jet} : m, \Gamma_W) \otimes G(m_{\rm jet} : m, \sigma_W),$$
$$S_Z(m_{\rm jet}) = F_{\rm BW}(m_{\rm jet} : m, \Gamma_Z) \otimes G(m_{\rm jet} : m, \sigma_Z),$$
$$F_{\rm BW}(m : \bar{m}, \Gamma) = \frac{1}{2\pi} \frac{\Gamma}{(m_{\rm jet} - \bar{m})^2 + \Gamma^2/4},$$
$$G(x : \bar{x}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right].$$
$$\bar{m}_W = m_W + m_W^{\rm offset},$$
$$\bar{m}_Z = m_W + \Delta m_{WZ} + m_Z^{\rm offset},$$

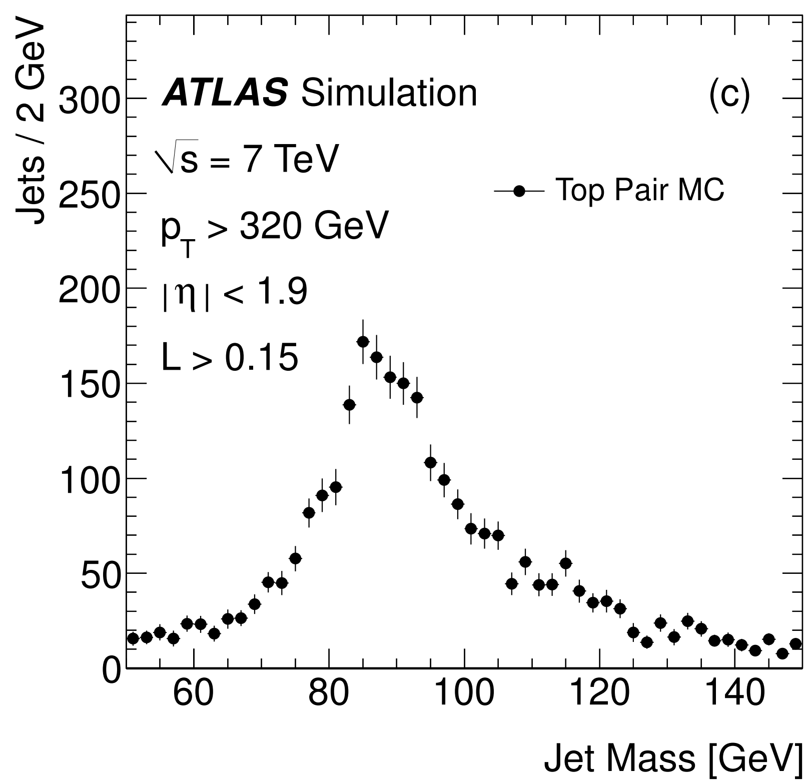
All parameters are fixed to MC predicted values:

1. Means of BW functions (W/Z mass)
2. Width of BW function (W/Z width) fixed to PDG
3. W and Z resolutions (Gauss width)
4. Relative fraction of W/Z signal yield (MCFM calculation)



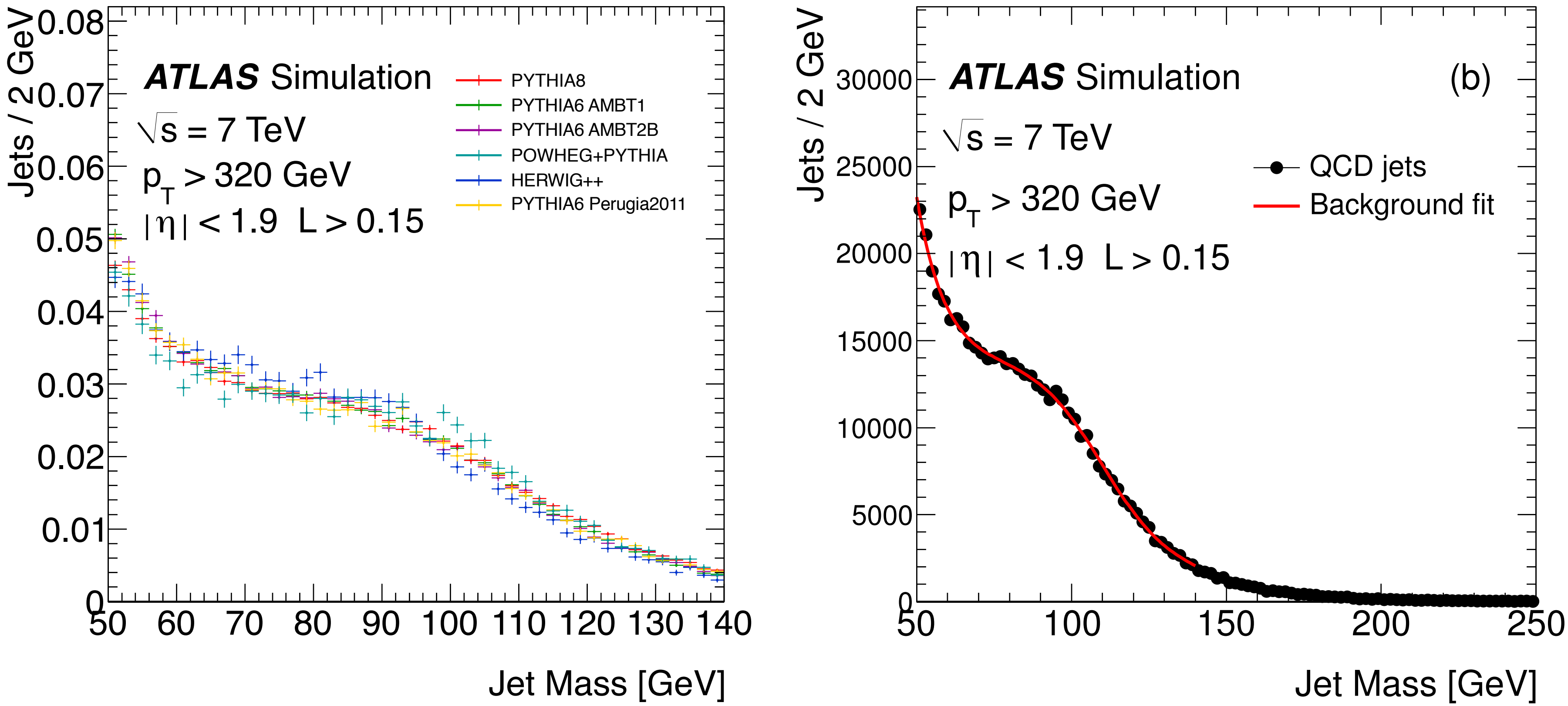
Peaking background PDF

- SM diboson production: WW,WZ,Wγ and Zγ
  - ✓ Identical PDF as signal, very small contribution,
  - ✓ Not explicitly included in fit, deduct contribution from fitted signal yields
  - ✓ Sys error: theoretical prediction of cross section
- SM single top production:
  - ✓ Dominated by Wt production, very small contribution,
  - ✓ Not explicitly included in fit,
  - ✓ Deduct contribution from fitted yields
  - ✓ Sys error: theory cross section
- SM top pair production:
  - ✓ Model using 1-D histogram from MC
  - ✓ Broader distribution to signal PDF
    - Nearby b jet
  - ✓ Small contribution to signal
  - ✓ Sys uncertainties:
    - Theory cross section
    - Different MC simulation



Combinatorial (QCD jet) background PDF

- The dominated background
  - ✓ Not well predicted by MC simulation



- Can be described by the same analytical function (different parameter values)
  - ✓ Different MC simulation, different selection (LH, p_T, jet cone size)
  - ✓ More than 100 different variations
  - ✓ Verified with the control sample from the data (see later slides)

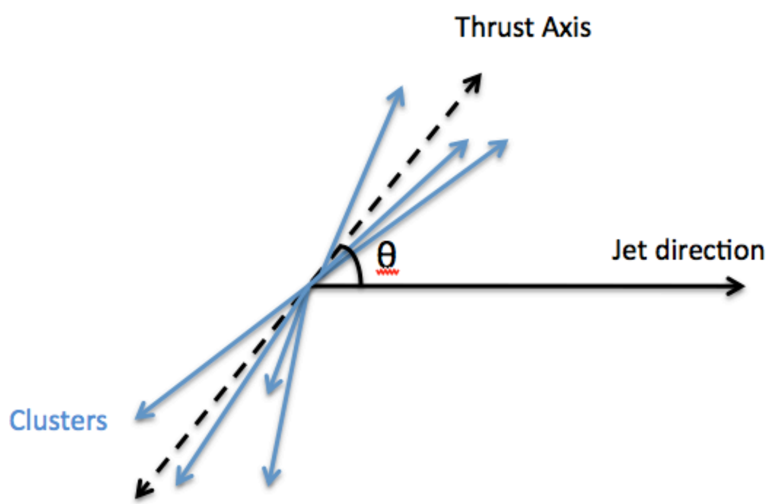
Combinatorial (QCD jet) background PDF

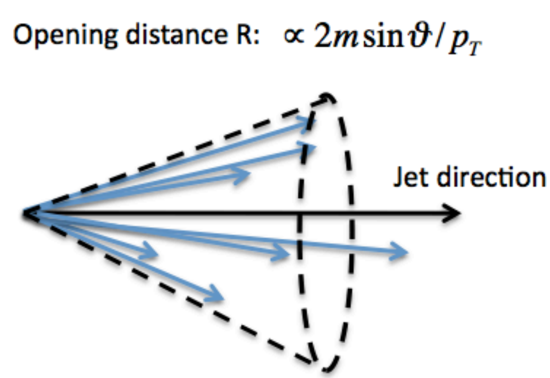
- Analytical PDF: all parameters are free in the fit:
  - ✓ 2 exponential functions + 1 sigmoid function

$$S_{\mathrm{QCD}}(m_{\mathrm{jet}}) = f_{\mathrm{E}} \cdot \mathrm{E}(\mathrm{m_{jet}} : \mathrm{m_0}, \sigma_{\mathrm{m}}) + \mathrm{f_1} \cdot \mathrm{C_1} \exp(\mathrm{a_1} \cdot \mathrm{m_{jet}}) + (1 - \mathrm{f_E} - \mathrm{f_1}) \cdot \mathrm{C_2} \exp(\mathrm{a_2} \cdot \mathrm{m_{jet}}),$$

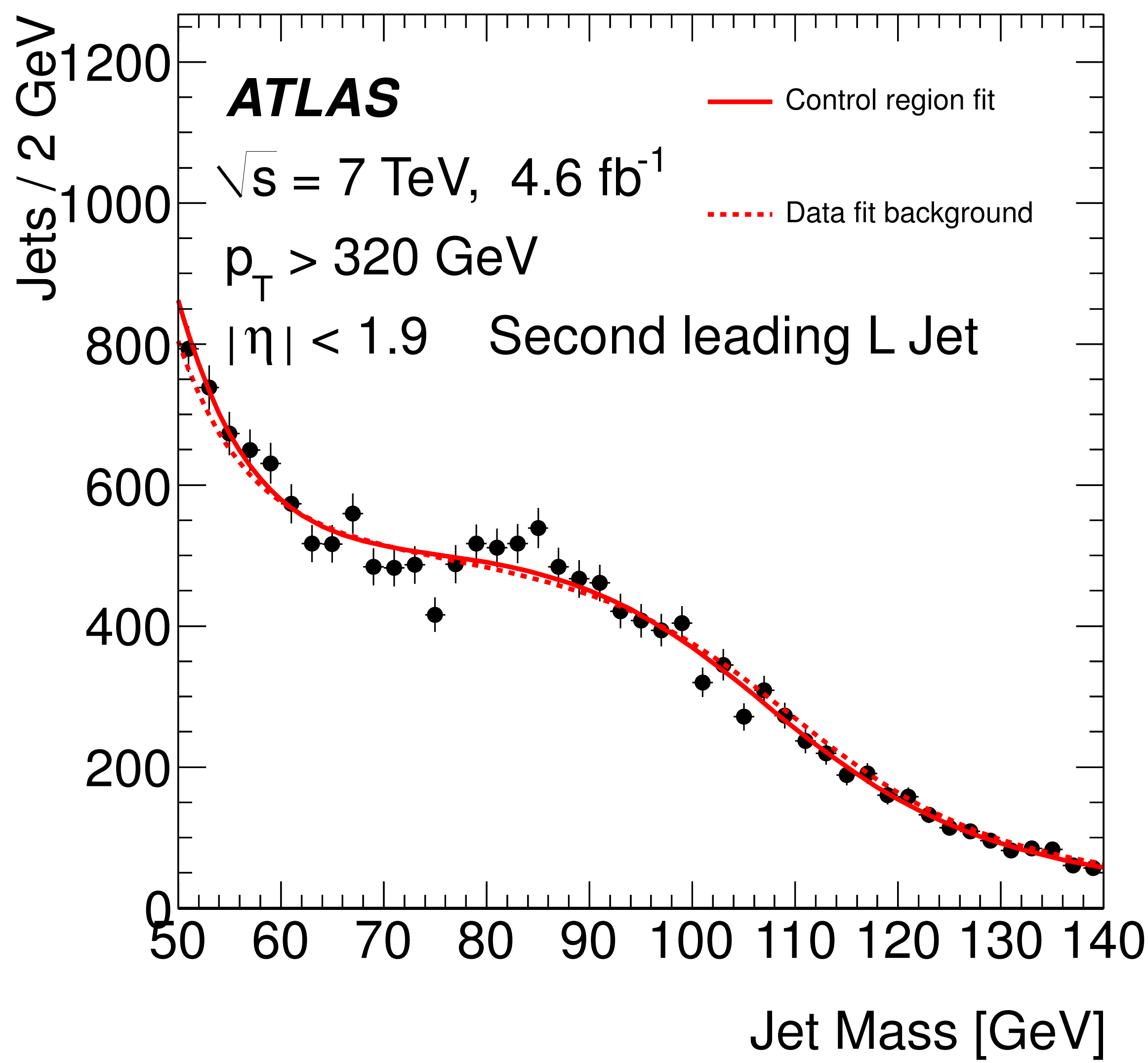
$$E(\bar{m}) = \bar{m}/\sqrt{1+\bar{m}^2} \quad \text{and} \quad \bar{m} = (m_{\mathrm{jet}} - m_0)/\sigma_m$$

- The sigmoid function (shoulder structure) is caused by kinematic effects
  - ✓ Variation with respect to jet p_T and jet cone size well produced in MC
    - Large p_T, higher threshold of the shoulder
    - Large jet cone size, higher threshold of the shoulder

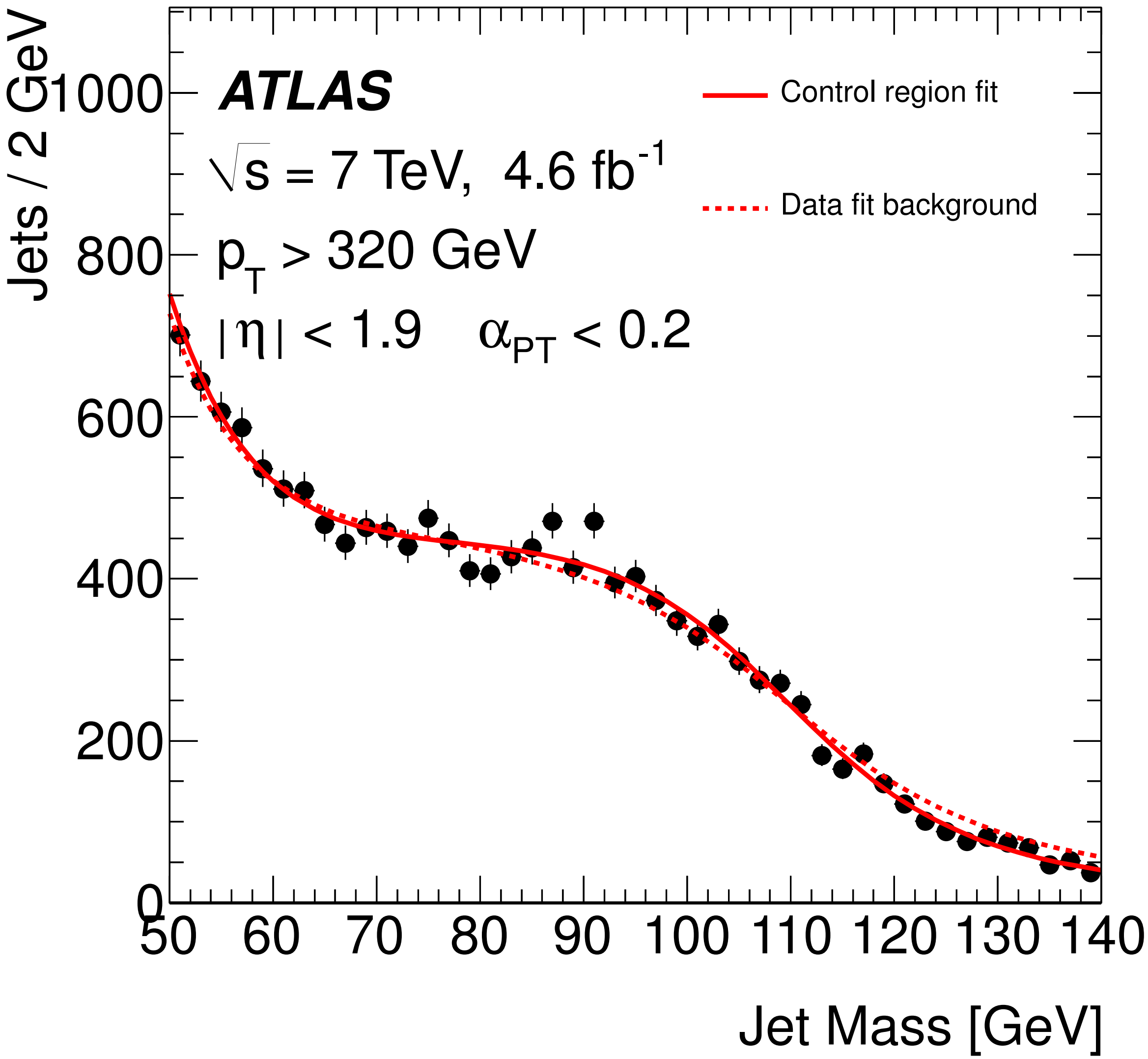




Test background PDF using control data



Test background PDF using control data



$$\alpha = p_T^{bal}/M_{dijet}$$

p^{bal}_T is the transverse momentum of the best balancing jet

M_{dijet} is the invariant mass of the balancing jet and candidate jet

Weak correlation with m_{jet} and substructure

Less than 1% signal contamination

Maximum likelihood fit

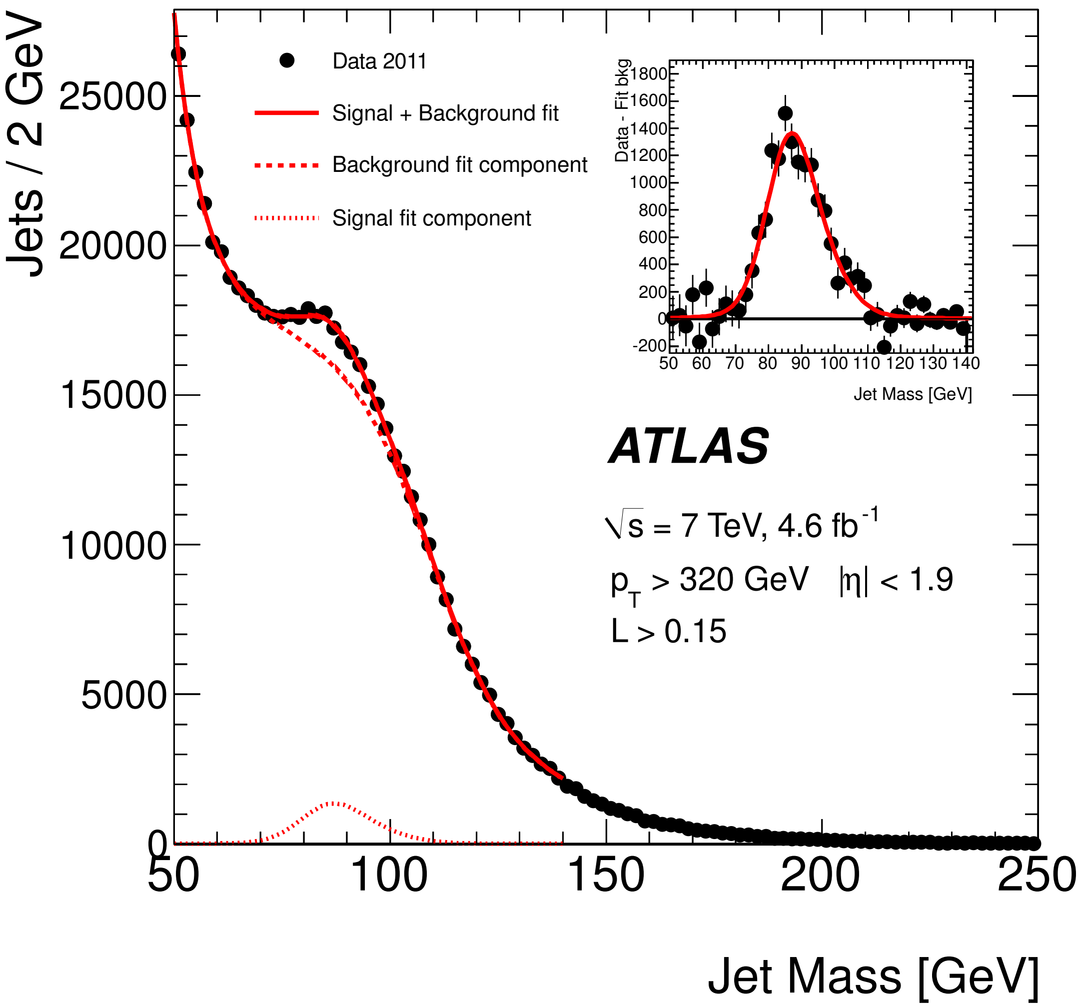
- Binned maximum likelihood fit:
  - ✓ Fit hadronic W/Z signal yield and background parameters from data

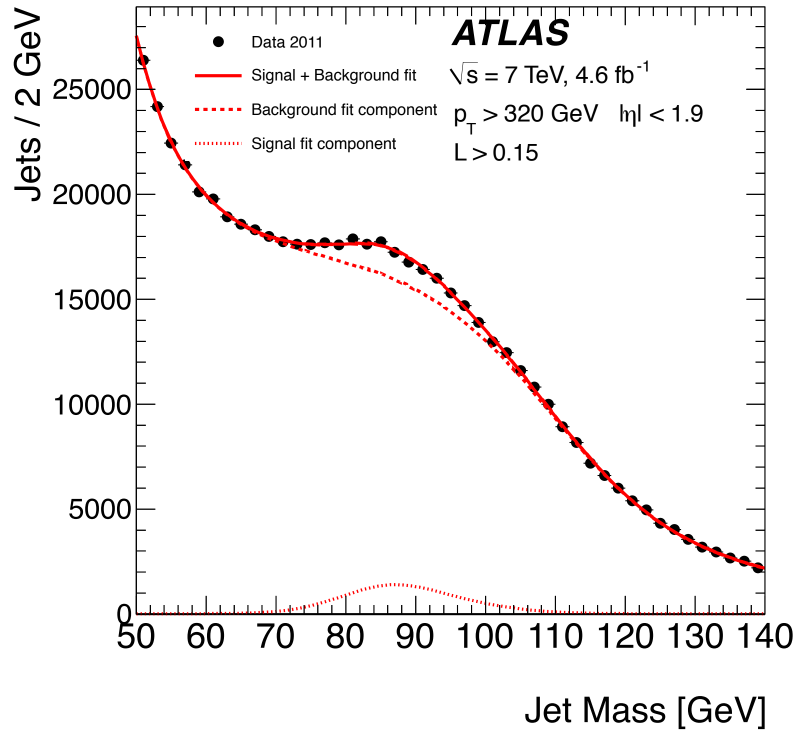
$$\mathcal{L} \equiv \prod_{i=1..n} \left\{ f_{\rm sig} \times [f_W \cdot S_W + (1 - f_W) \cdot S_Z] + f_{t\bar{t}} \times S_{t\bar{t}} + (1 - f_{\rm sig} - f_{t\bar{t}}) \times S_{\rm QCD} \right\}_i$$

name	Description	Comments
	Signal	
$f_{\rm sig}$	Combined signal fraction	Free parameter
f_W	Relative fraction of W-jets of signal yield	Fixed to MC prediction
m_W	W boson pole mass	Fixed to PDG value
Γ_W	Intrinsic width W boson	Fixed to PDG value
Γ_Z	Intrinsic width Z boson	Fixed to PDG value
σ_W	Detector resolution of reconstructed W mass	Fixed to MC prediction
σ_Z	Detector resolution of reconstructed Z mass	Fixed to MC prediction
	QCD	
$f_{\rm E}$	Fraction of the sigmoid component in QCD PDF	Free parameter
f_1	Fraction of the first exponential component in QCD PDF	Free parameter
m_0	Inflection point of the Sigmoid function in QCD PDF	Free parameter
σ_m	Curvature at inflection point of the Sigmoid function in QCD DPF	Free parameter
a_1	Slope of the first exponential component in QCD PDF	Free parameter
a_2	Slope of the second exponential component in QCD PDF	Free parameter
	Other background	
$f_{t\bar{t}}$	Fraction of $t\bar{t}$ background	Fixed to MC prediction

Table 4: List of parameters used in the default fit.

Fit results





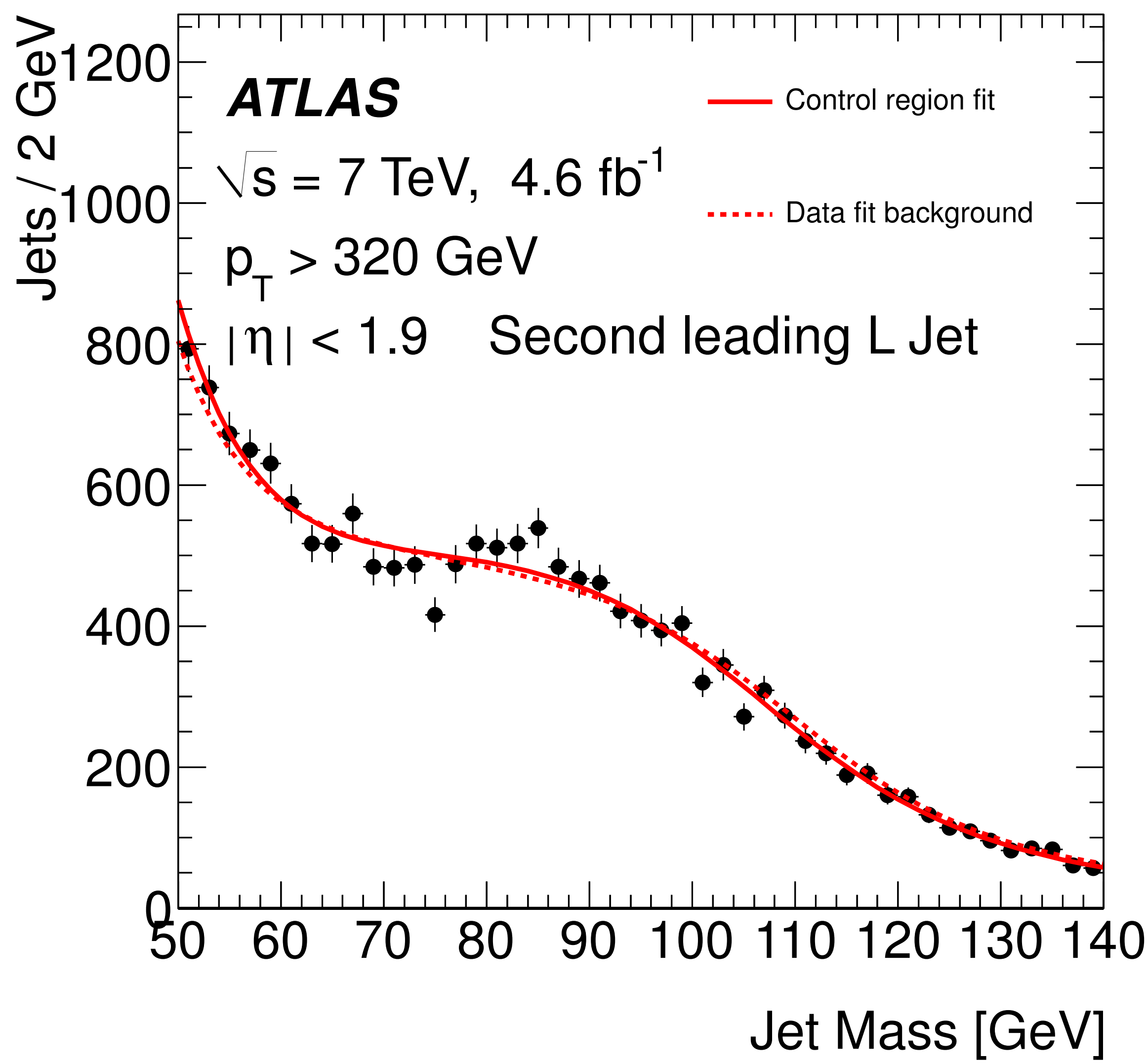
Fit χ^2/ndf=41.4/38
χ^2 probability: 32%

χ^2 probability < 10^{-7} if assuming no signal peak

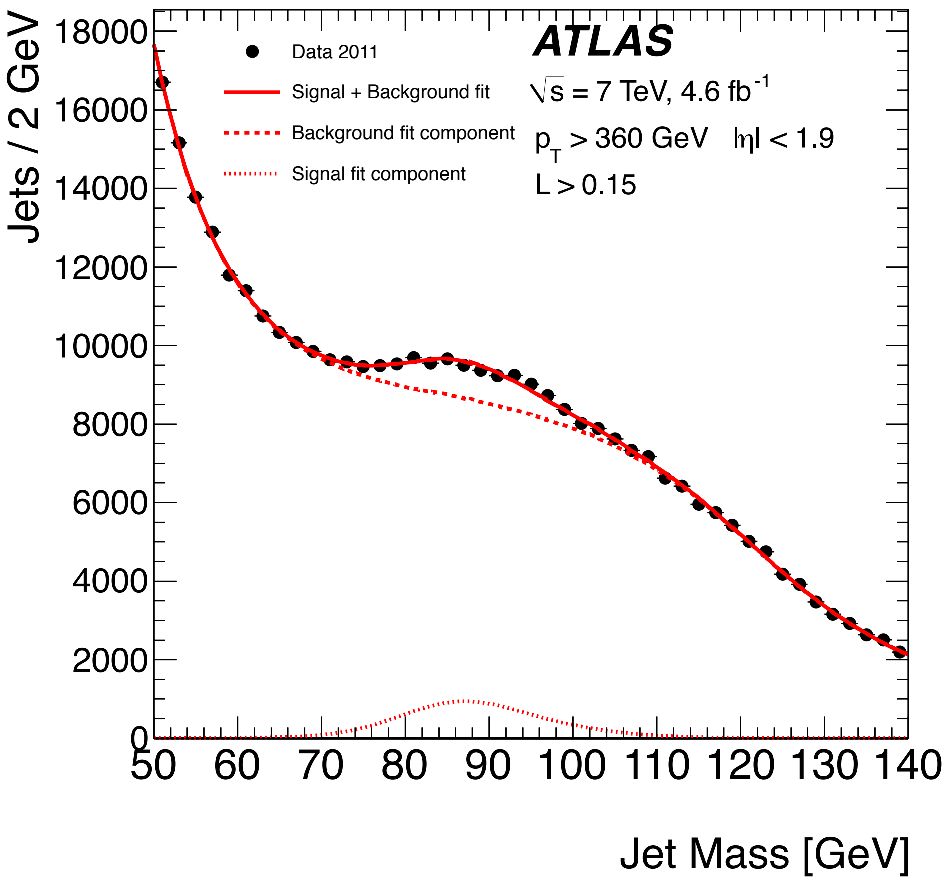
After deduct peaking bg contribution:

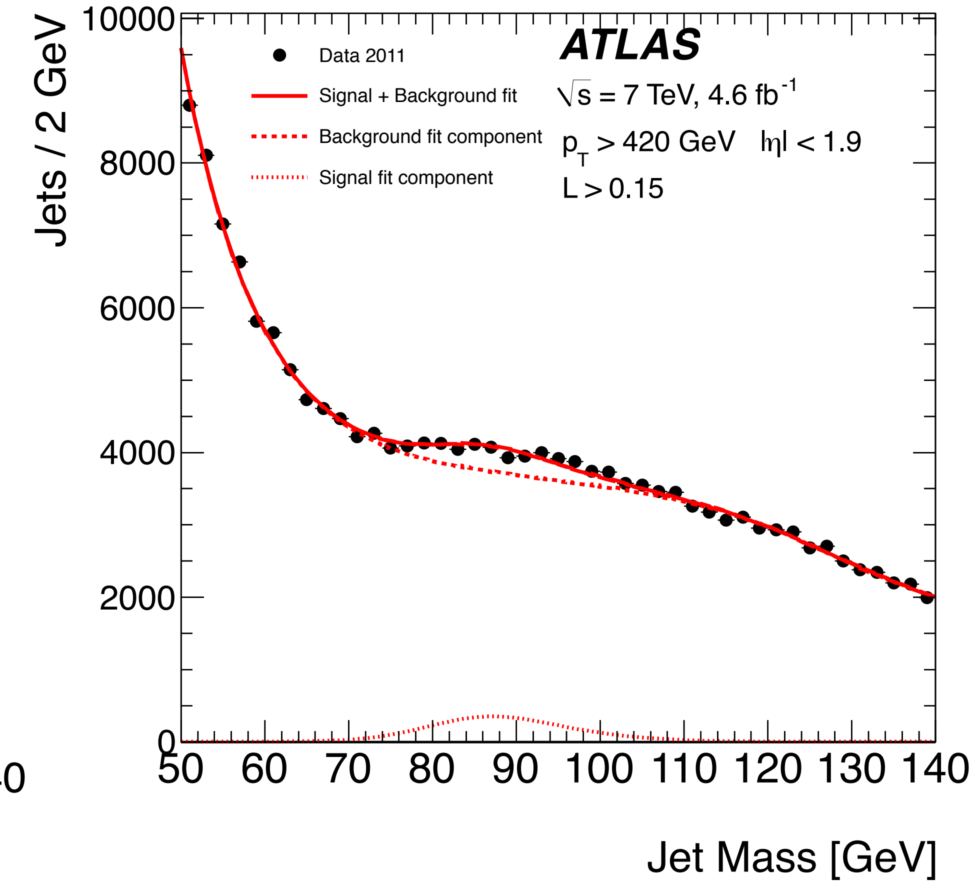
$$N^{W+Z} = 14200 \pm 1300(\mathrm{stat})$$

Test background PDF using control data



Cross check





- Relative signal yield with different p_T cut consistent with MC efficiency calculation
- Change of threshold position of the background consistent with MC expectation

Cross check

- Signal yields of different LH cuts consistent with MC expectation

- Repeat fit with m_W allowed to be float
  - ✓ The difference between m_W and m_Z is fixed to MC

$$\Delta m = m_W^{\rm fit} - m_W^{\rm MC} = -0.45 \pm 0.86 \ {\rm GeV}$$

  - ✓ Similar results with different p_T and LH cuts

- Repeat fit with m_{jet} resolution to be free parameter
  - ✓ A common scale factor of the resolution of $m_{jet}(W)$ and $m_{jet}(Z)$
  - ✓ Fitted scale factor consistent with 1 within statistic uncertainty

- Repeat fit with relative yield of W and Z signal to be floated
  - ✓ Fitted results consistent with MC expectation with large stat error (~10%)
  - ✓ Small impact of total signal yields
  - ✓ Small impact on total signal yield with simultaneous free m_W parameter

- Toy MC studies to verify that no bias in the fit procedure
  - ✓ Toy MC with analytical background PDF
  - ✓ Toy MC with background using control data

Cross section measurement

- Measurement of the cross section

$$\sigma_{W+Z} = N^{W+Z}/(\mathcal{L}\cdot\epsilon) \quad \text{and} \quad \epsilon = N^{W+Z}_{\text{reco}}/N^{W+Z}_{\text{gen}}$$

- Efficiency estimated to be 0.36±0.02 (stat) using MC

- Results:

$$\sigma_{W+Z} = 8.5 \pm 0.8\,\text{pb}$$

- Systematic uncertainty:
  - ✓ Efficiency calculation:
    - Evaluated using different MC
    - ~ 4.4% relative uncertainty for the cross section measurement
  - ✓ Signal yield (see later slides)
    - Dominant systematic uncertainty
    - ~ 18% relative uncertainty of the cross section measurement

Systematic uncertainties

Sources	σ_{W+Z}
MC modelling	4.4 %
Background pdf	8.8 %
Signal pdf	5 %
Jet energy scale	3.7 %
Jet energy resolution	< 1 %
Jet mass scale	2.2 %
Jet mass resolution	12.6 %
$t\bar{t}$ contribution	2.8 %
Single-top and diboson contribution	< 1 %
W and Z relative yield	2.9 %
Luminosity	1.8 %
Total	18 %

Systematic uncertainties

- Evaluation of sys error due to the background PDF:
  - ✓ Different choices of analytical functions of sigmoid function
  - ✓ Additional of 2^{nd} order polynomial functions
  - ✓ Add or remove the exponential functions
  - ✓ Maximum deviation of the signal yield as sys error
- Systematic uncertainty due to signal PDF: different MC generator
- Three independent evaluations of the sys uncertainty due to jet mass resolution
  - ✓ Using MC simulation to estimate sys effects (default)
    - Different parton shower and hadronization model
    - Different materials and geometry in detector GEANT model
    - Different model of interactions of high energy hadron with martials
  - ✓ Let mass resolution scale to float in fit (data driving cross check)
  - ✓ Study the uncertainties of the energy cluster measurements
    - Different passive materials in detector model
    - Measurement uncertainties of cluster energy
    - Measurement uncertainties of cluster positions
    - Propagate it to the mass measurements
- The systematic uncertainties can be reduced with more data

Final result

$$\sigma_{W+Z} = 8.5 \pm 0.8\,(\mathrm{stat.}) \pm 1.5\,(\mathrm{syst.})\,\mathrm{pb}$$

- NLO QCD calculation: MCFM
  - ✓ W/Z+jets calculation
  - ✓ CT10 parton distribution function
  - ✓ Dynamic scale factor $H_T/2$: H_T is scalar sum of particle p_T in the final state

$$\sigma_{W+Z} = 5.1 \pm 0.5\,\mathrm{pb}$$

- Systematic uncertainty of NLO QCD calculation:
  - ✓ Varying factorization and normalization scale from 0.5 to 1
  - ✓ Parton distribution function (small)
  - ✓ Strong coupling constant (small)
- Measurement consistent with NLO QCD calculation within 2 sigma level

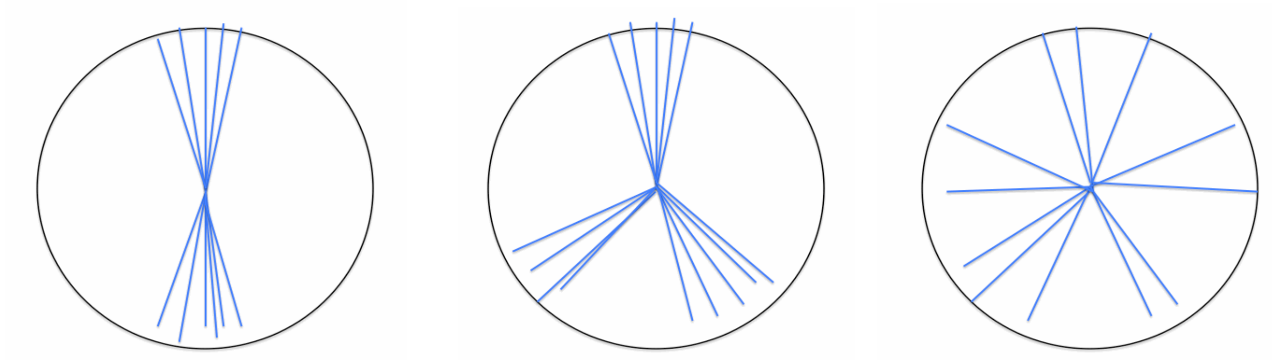
ATLAS, arXiv:1407.0800 (accepted by NJP)

Application of jet substructure in jet rest frame

Reclustering in the Rest Frame

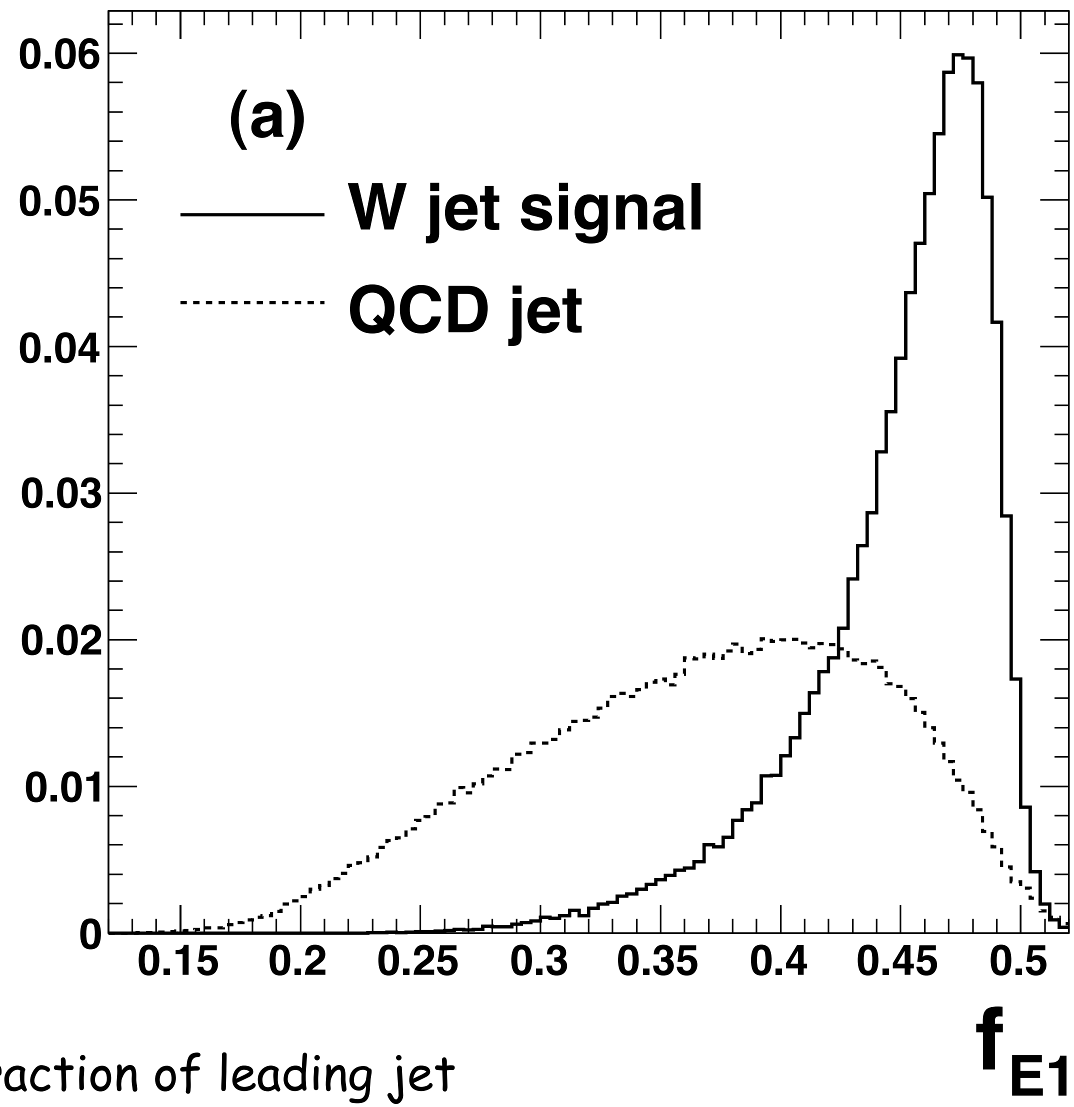
- More improvements/applications based on jet CM substructure
  - ✓ Reclustering (filtering): reconstruct subjets in jet rest frame
  - ✓ Combination of pruning and trimming, tracks information …….

W/Z->jj t->Wb->jjj QCD



- Rerun the jet finding algorithm on the clusters in the CM frame
  - ✓ Fastjet
  - ✓ Jet algorithm similar (not identical) to e^+e^- experiments
    - Tradition jet algorithm based on η and θ not appropriate
    - Combine 2 clusters in $\Delta\Theta<0.6$
    - Angle Θ: angle between 2 clusters

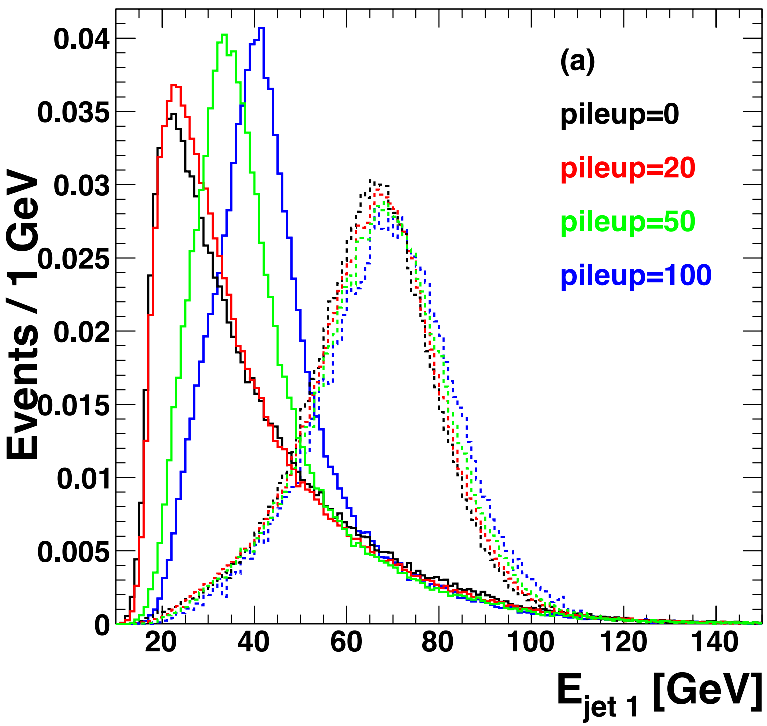
Reclustering in the Rest Frame

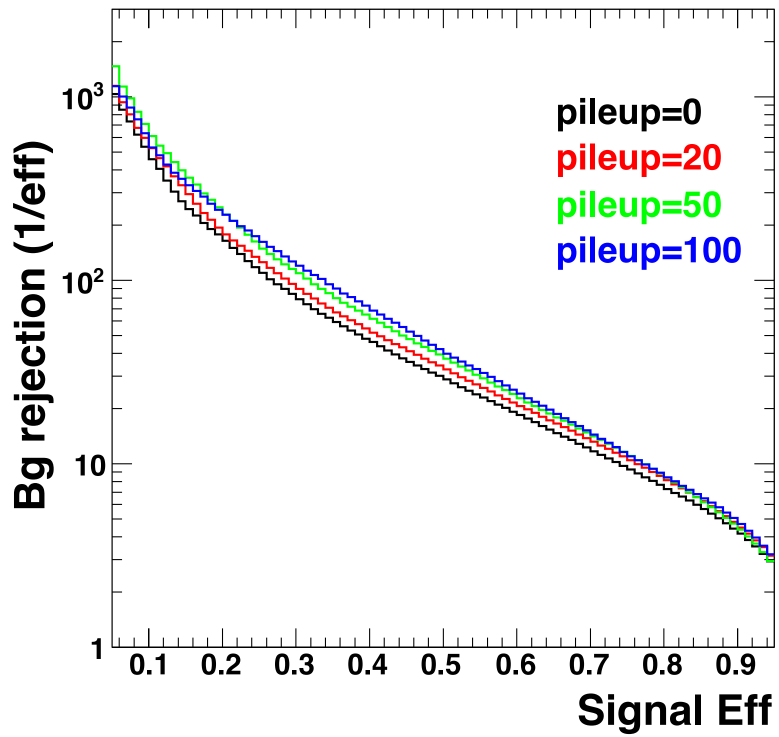


f_{E1}: energy fraction of leading jet

Subjets of boosted top in its CM frame

- SM top pair MC samples: $p_T(top)$>600GeV, W decay hadronically
- At least 3 subjets with E>10GeV



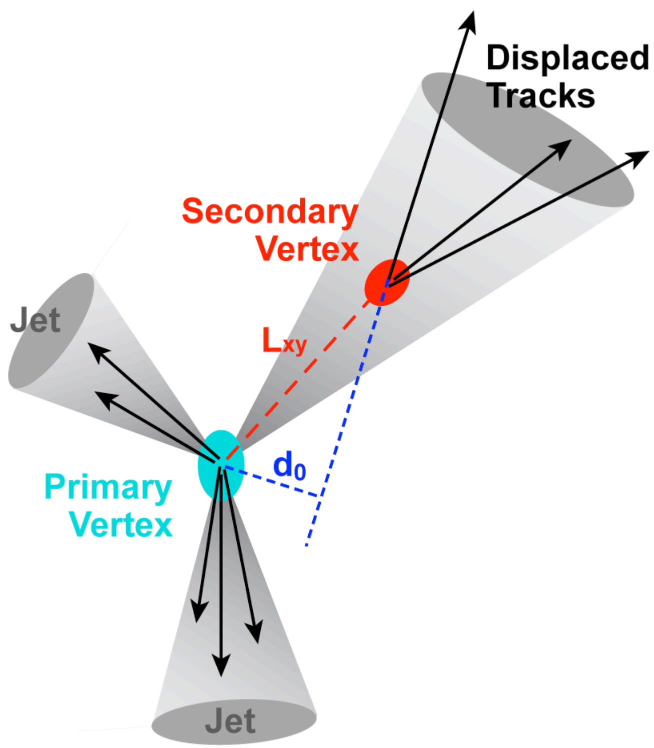


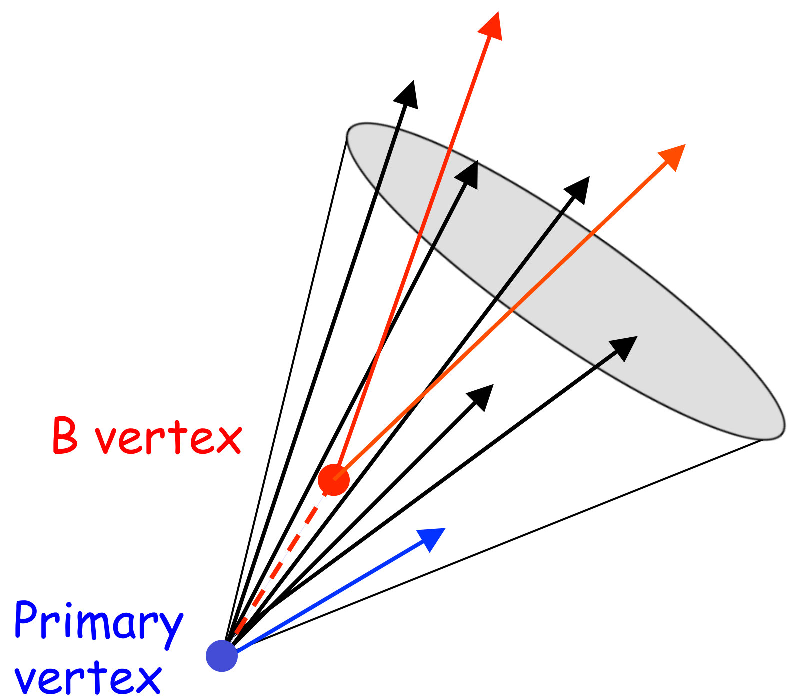
- Many jet substructure variables are correlated
- Multi variable approach to combine different variables
  - ✓ Energies of 3 leading jets, mass combinations ……

Chunhui Chen, PRD 87,074007 (2013)

Identify b quark inside boosted top

- Top quark decays to Wb almost 100%
- Identify b quark (b-tagging) based on its long lifetime



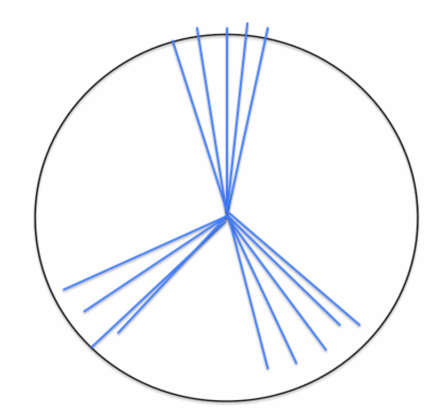


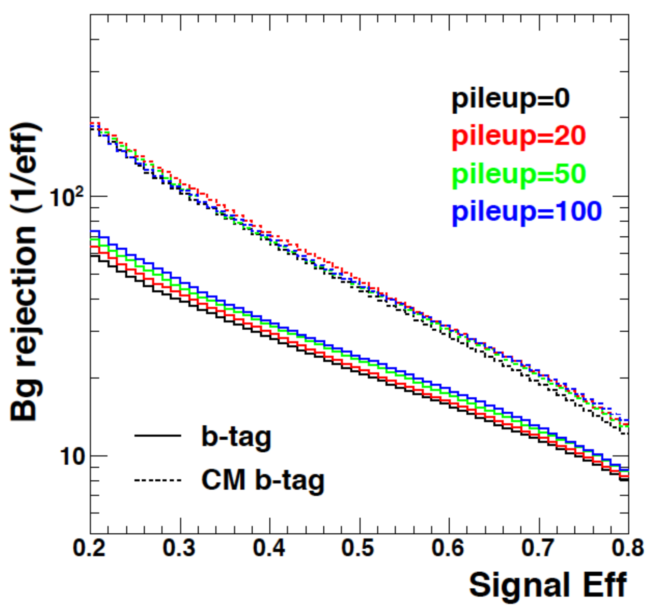
Problem of direct application of b-tagging for boosted top jet: Difficult to disentangle tracks originated by b decays from tracks originated from W decay

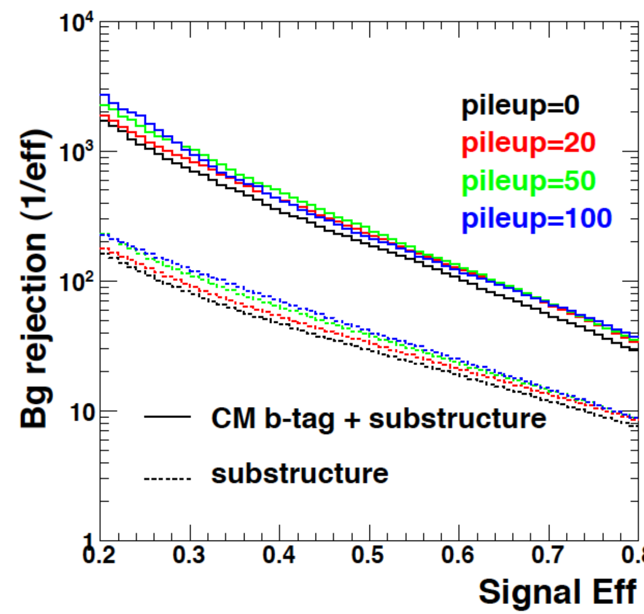
Identify b quark inside boosted top

- Boost charged tracks back into jet rest frame
- Associate tracks with subjets
- Separate tracks originated from different partons: b or W->qq'
- Comparing to direct application of b-tagging
  - ✓ Studies done using impact parameter algorithm b-tagging
  - ✓ Better performance using CM b-tagging
- Combine b-tagging with jet substructure

t->Wb->jjj



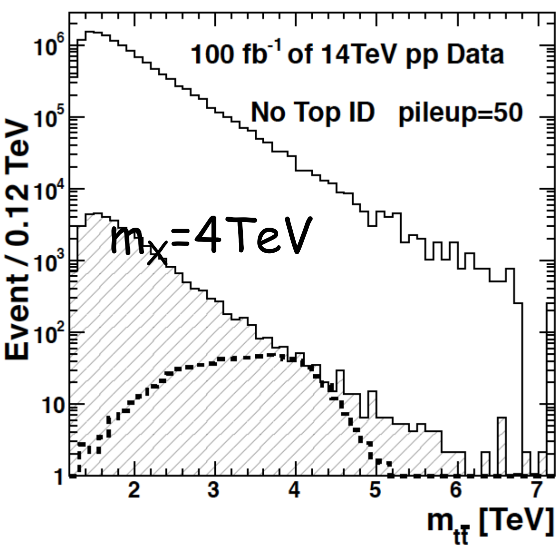


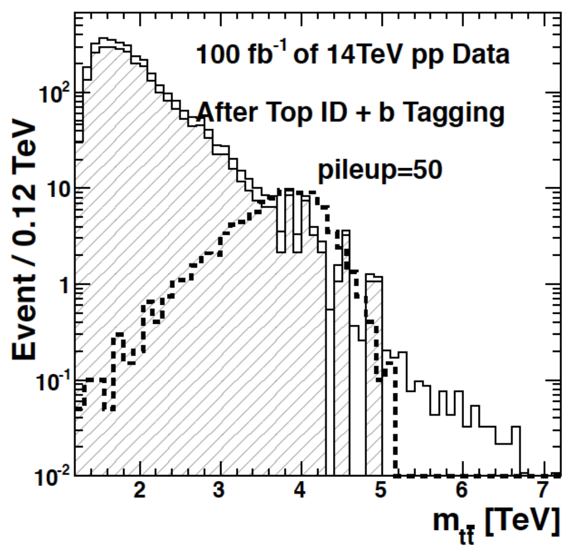


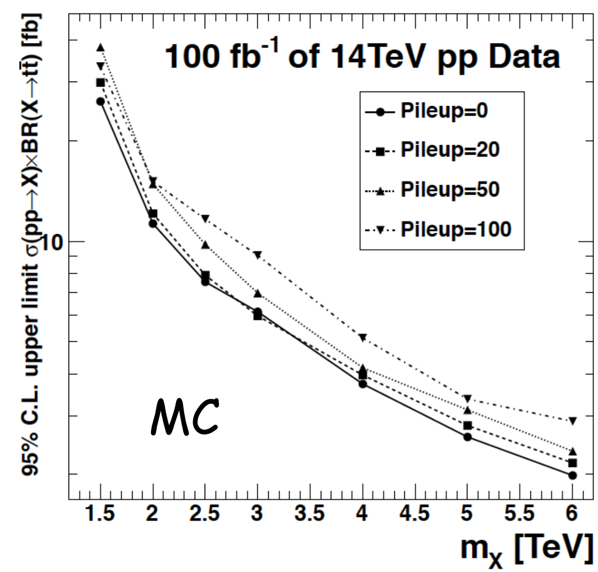
Chunhui Chen, PRD 88,074009 (2013)

Using Top ID to search for X->ttbar

- Heavy resonance decaying into a top pair
- Both top decay hadronically: hadronic W
  - ✓ Dominant bg: SM mutijet production, top pair production
- Choose 2 leading jets as top candidate to form a X candidate
- Assuming effective production cross section of X: 10fb

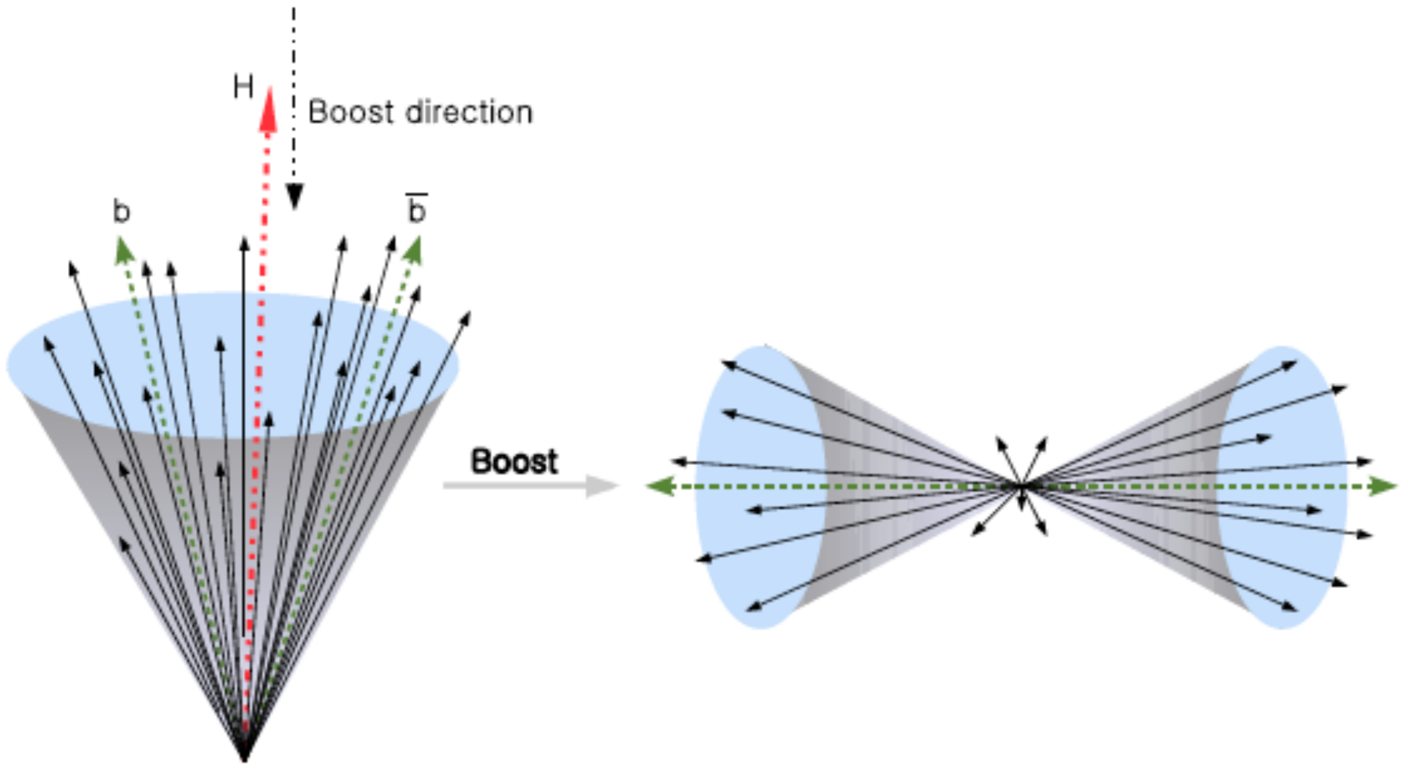






Chunhui Chen, PRD 88,074009 (2013)

Double b tagging in boosted Higgs jet



Summary and Conclusion

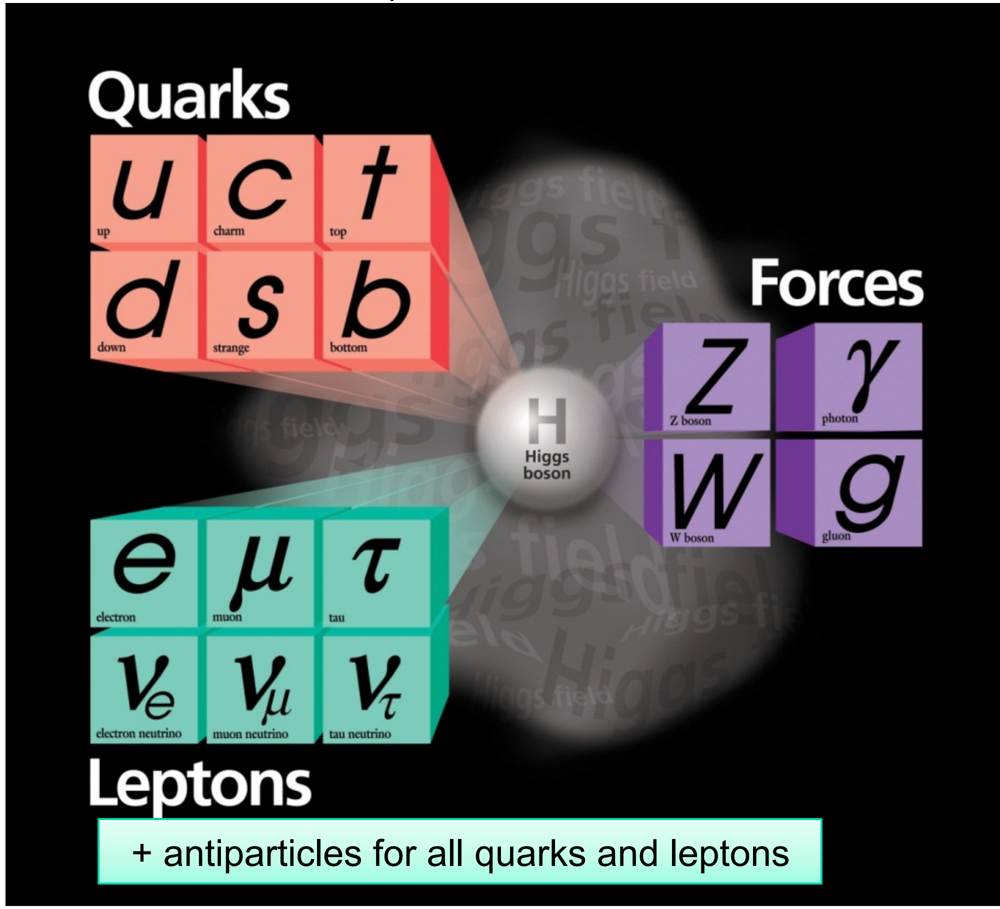
- Boosted hadronic decaying particles a powerful tool to search for NP
  - ✓ Jet mass and jet substructure
- Describe a new approach to identify boosted particle
  - ✓ Based on shape variables/reclustering in jet CM frame
- **First measurement of boosted hadronic W/Z production using single jets**
  - ✓ **Demonstrate power of jet substructure algorithm in jet rest frame**
  - ✓ **The new method complementary to existing jet substructure algorithms**
  - ✓ ATLAS, arXiv:1407.0800 (accepted by NJP)
- Additional improvement/application of jet substructure in jet rest frame
  - ✓ C. Chen: PRD 85,034007(2012); PRD87,074007(2013) and PRD 88,074009 (2013)

Application of the jet substructure in jet rest frame to search for NP using coming 14TeV data at the ATLAS, stay tuned!

Backup

Standard Model of Particle Physics

- Particle Physics (High Energy Physics):
  - ✓ Study fundamental particles and how they interact
- Matter is made of fermions
  - ✓ quarks and leptons
  - ✓ 3 generations
- Forces carried by bosons:
  - ✓ Electromagnetic: γ
  - ✓ Weak: W and Z
  - ✓ Strong: gluons
- Higgs boson:
  - ✓ Give mass to particles



Standard Model + Gravity = Basic Building blocks of our Knowledge

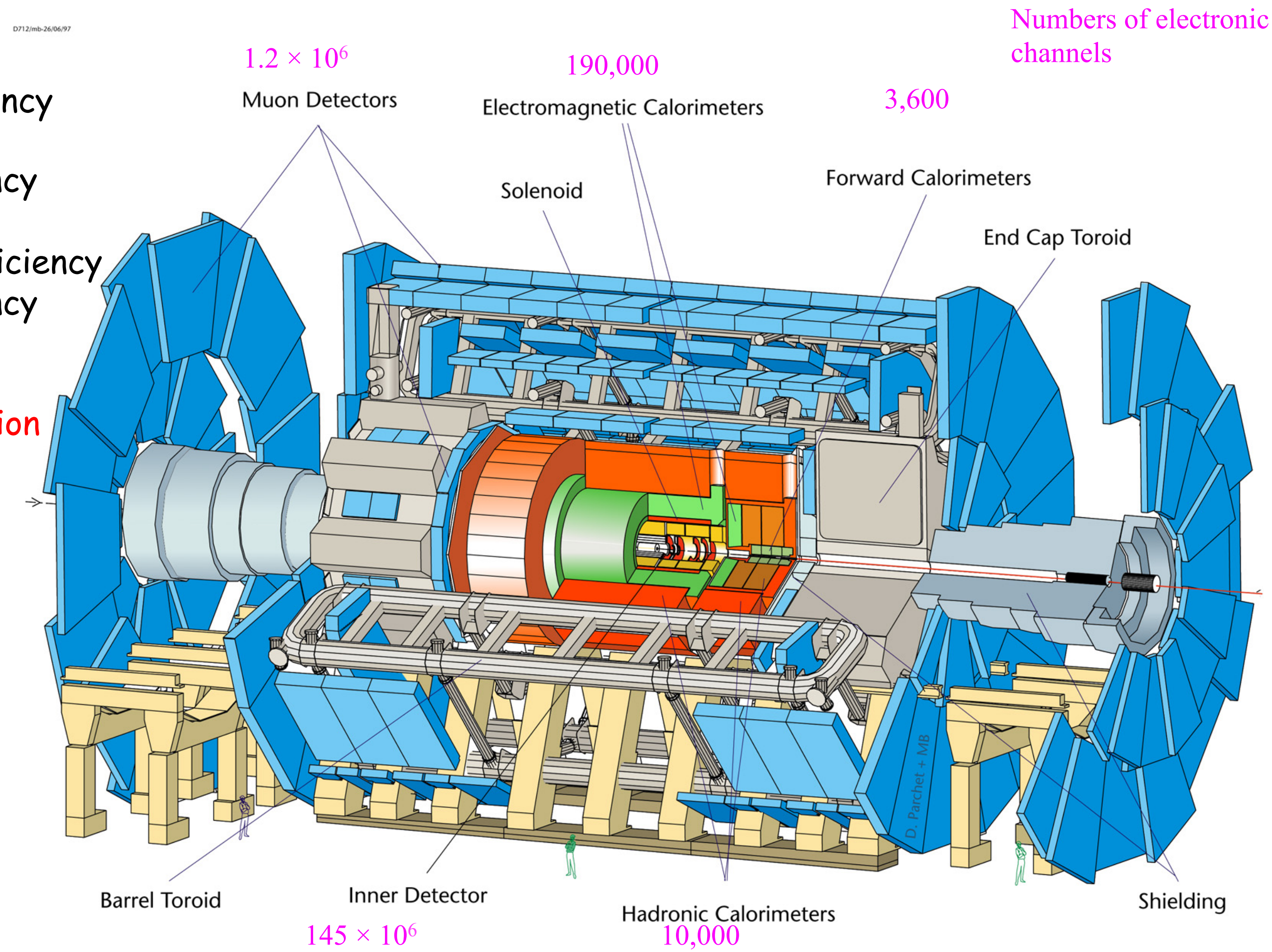
The ATLAS Detector

e: ~75 – 90% efficiency

muon: ~90% efficiency

b tagging: ~57% efficiency
~ 0.2% fake efficiency
from light jets

Efficiency & resolution
dependents on the
selection criteria



46 m long, Overall weight: 7000 Tons

Excellent reconstruction efficiency and resolution:
Electron, muon, track, jets, b-tagging & missing transverse energy

The jet observables

- Single jet mass $m_{\text{jet}} = \sqrt{E_{\text{jet}}^2 - p_{\text{jet}}^2}$
  - Deduced from four-momentum sum of all jet constituents
    - Before and after any grooming
  - Can be reconstructed for any meaningful jet algorithm

- momentum balance $\sqrt{y_f} = \min(p_T^{j1}, p_T^{j2})\Delta R_{12} / m_{12}$
  - Where p_T^{j1} and p_T^{j2} are the transverse momenta of the two leading subjets, ΔR_{12} is their separation and m_{12} is their mass
  - To suppress jets from gluon radiation and splitting, $\sqrt{y_f} > 0.45$

- *N*-subjettiness
  - Measures how well jets can be described assuming *N* sub-jets
    - Degree of alignment of jet constituents with *N* sub-jet axes
  - Sensitive to two- or three-prong decay versus gluon or quark jet
    - Highest signal efficiencies from N-subjettiness ratios τ_{N+1}/τ_N ($\tau_{N+1/N}$ or $\tau_{N+1,N}$)
    - For most analyses in this talk ($W/Z \rightarrow qq$) will use $\tau_{2/1}$