Subjet structure as a Higgs search tool

Jonathan Butterworth

University College London

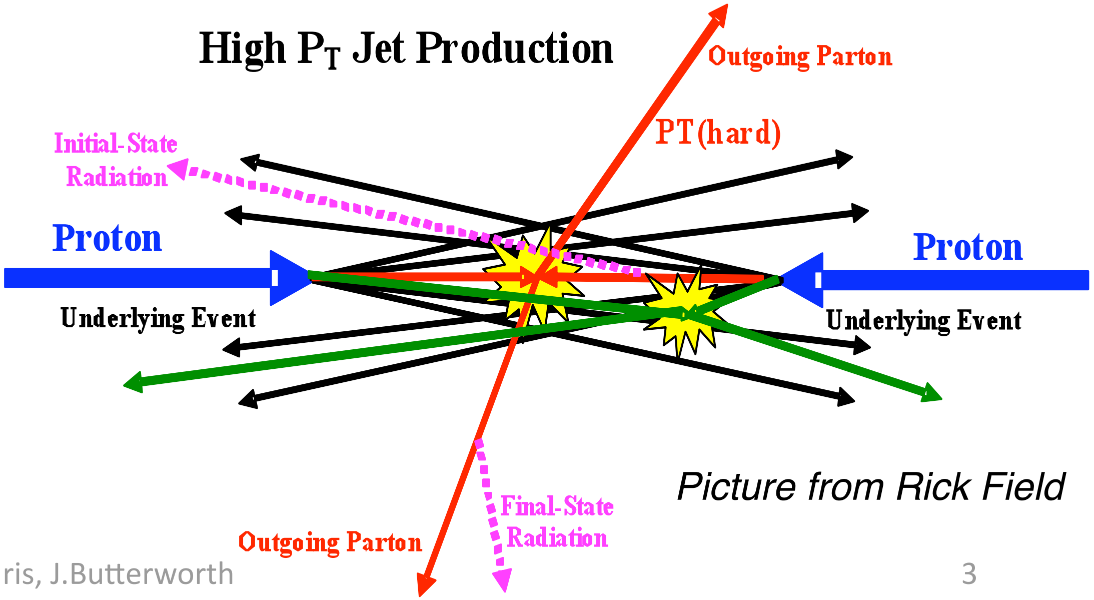
Paris, 30/7/2010

Subjets

- Jets
- Why now?
- Jet substructure and QCD
- Subjets and the Higgs

What is a Jet?

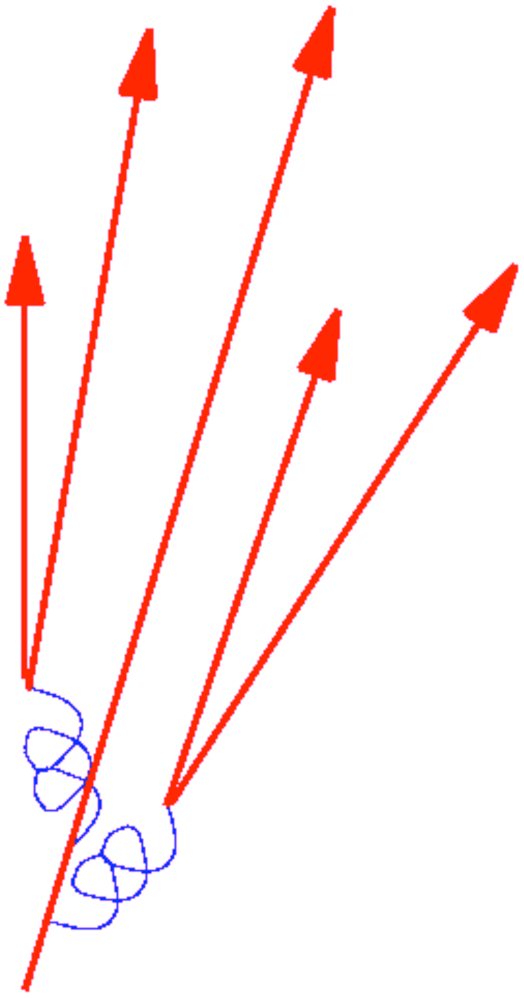
- Protons are made up of quarks and gluons.
- Quarks and gluons are coloured and confined – we only ever see hadrons.
- A jet of hadrons is the signature of a quark or gluon in the final state.
- The gross properties (energy, momentum) reflect the properties of the quark or gluon, and stand out above the rest of the event.
- Jets algorithm.



Picture from Rick Field

What is a Jet?

- Evolution from a hard parton to a jet of partons takes place in a regime where:
  - Energy scale is high enough to use perturbation theory
  - X (momentum fraction of particles) is not very small
  - Collinear logarithms are large
  - Multiplicities can be large
  - **This is largely understood QCD, and can be calculated**
- Hadronisation (non-perturbative) stage has a small effect (sub-GeV level) and is well modelled by tuned Monte Carlo simulation (e.g. Lund string)



Jet Algorithms

- "Cluster" algorithms
  - – Generally start from the smallest objects available, and perform an iterative pair-wise clustering to build larger objects (using either geometric or kinematic properties of the objects)
  - – Sort of inverts the QCD parton shower idea
- Lend themselves most naturally to substructure studies.

Cluster algorithms

- Each has a distance measure, and merges the "closest" objects by this measure until some criteria is reached (could be a specified multiplicity, or "distance")
- Modern ones (k_T, Cambridge, anti-k_T) belong to a general class where the distance parameter is given as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} ,$$

$$d_{iB} = k_{ti}^{2p} ,$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

- p=1 for k_T, 0 for Cam/Aachen, -1 for anti-k_T

k_T algorithm

- *Catani et al Phys Lett B269 (1991); Nucl. Phys. B406 (1993); Ellis and Soper Phys Rev D48 (1993).*
  - p=1

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} ,$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = k_{ti}^{2p} ,$$

- Successively merge objects with low relative k_T
- If the k_T^2 of an object w.r.t the beam is lower than k_T^2 w.r.t anything else in the event divided by R^2, don't merge any more; call it a jet.
- Mimics (inverts) the QCD parton shower.
- Soft stuff merged into the nearest hard stuff.
- Can undo merging. Last merge is the hardest.

Cambridge/Aachen algorithm

- *Dokshitzer, Leder, Morretti, Webber (JHEP 08 (1997) 01; Wobisch and Wengler hep-ph/9907280*
  - p=0

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} ,$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = k_{ti}^{2p} ,$$

- Successively merge objects with low relative Δ.
- Objects with $\Delta^2 > R^2$ not merged
- Can undo merging. Last merge is the furthest away (so is often the softest).

Anti-k_T algorithm

- *Cacciari, Salam, Soyez JHEP 0804:063,2008*
  - p=-1

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2},$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = k_{ti}^{2p},$$

- Successively merge objects with high relative k_T
- d_{ij} is determined soley by the k_T of the harder of i & j, and by Δ. Soft stuff within R^2 of a high k_T object will be merged with it. If two hard jets are close the energy will be shared based on Δ.
- Shape of jet is unaffected by soft radiation.
- Can undo merging but the order is not very meaningful since the hardest object sucks in everything around it regardless of the relative hardness of the splitting.

Why (are subjets suddenly more interesting) now?

Scales in the experiment

1. Proton mass ~ becomes possible to accurately calculate using perturbative QCD

   around 1 - 5 GeV.

2. W, Z mass / electroweak symmetry-breaking scale / Higgs mass if it exists

   around 50-250 GeV

3. Phase space

   A few TeV

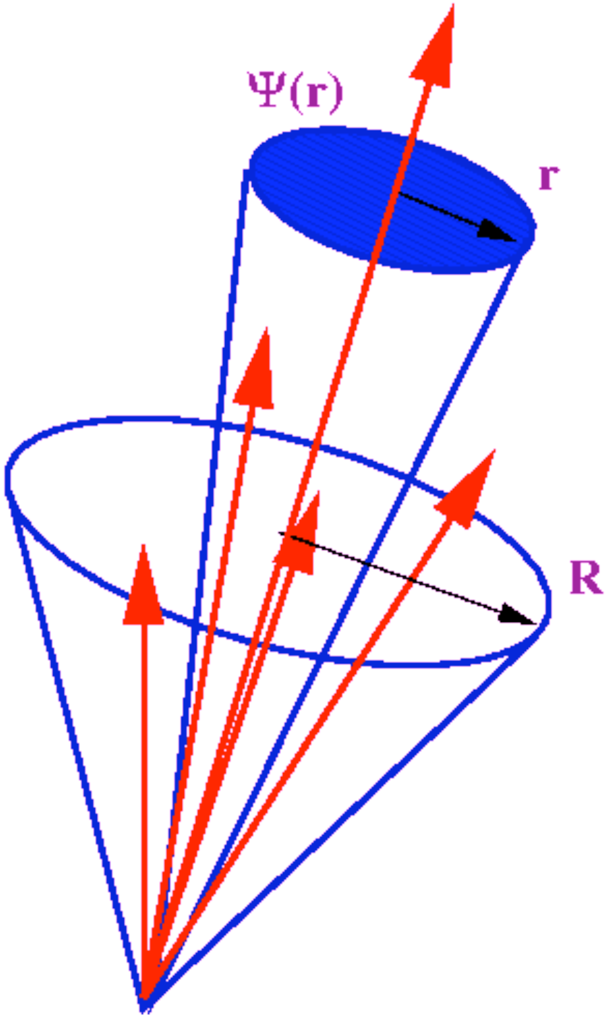
Scales in the experiment

- For the lifetime of the experiment, there will be interesting physics objects around the electroweak scale.
- Production of multiple EW-scale particles (W,Z,H,t…) and jets either directly or in cascade decays
  - Means we need the new calculations, especially Monte Carlos which match many-leg matrix elements to partons showers/resummations.
- Copious production of EW-scale particles well above threshold
  - Means highly collimated decay products, and therefore **interesting sub-jet structure** for hadronic decays.
  - The LHC will be the **first place** we have ever seen this.

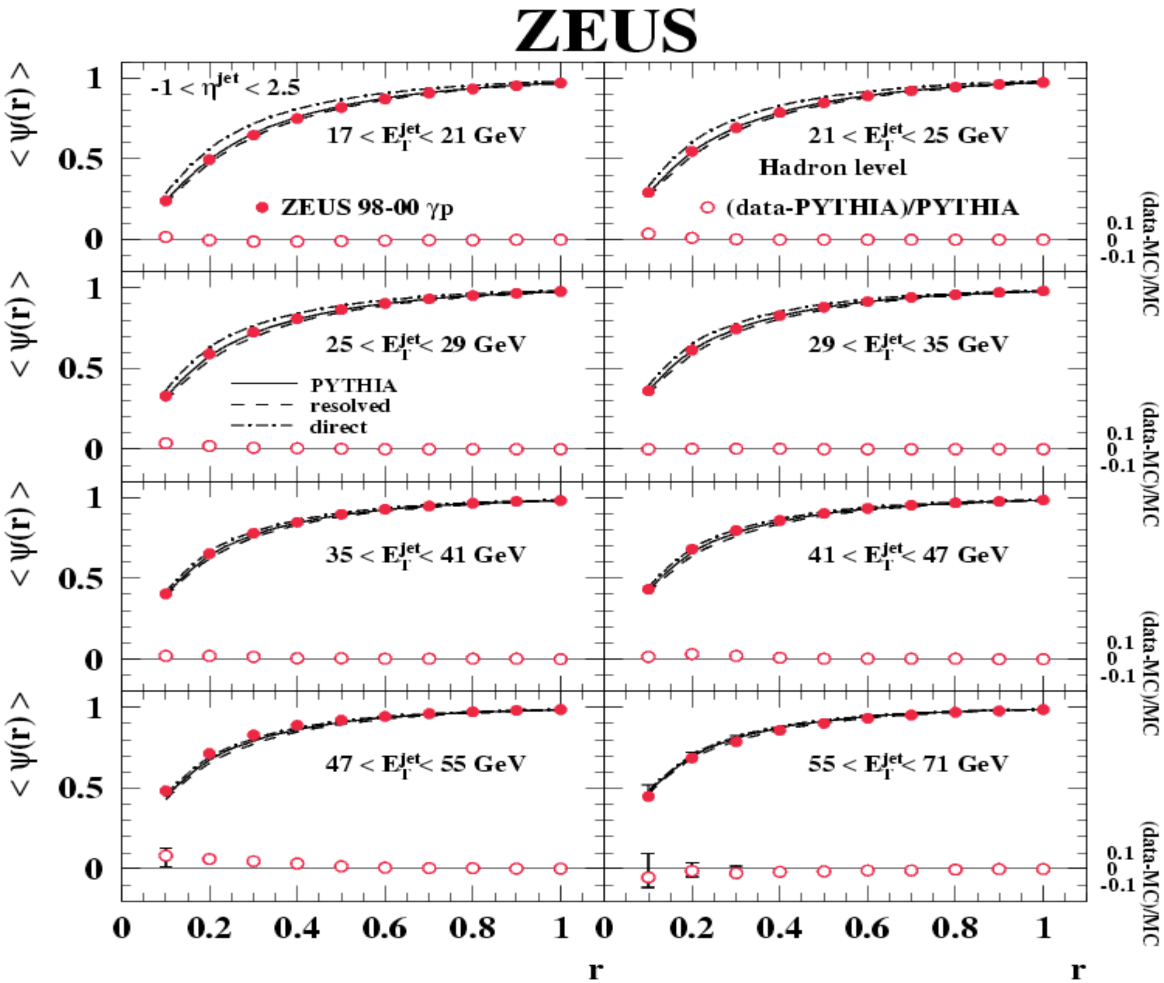
Subjets and QCD

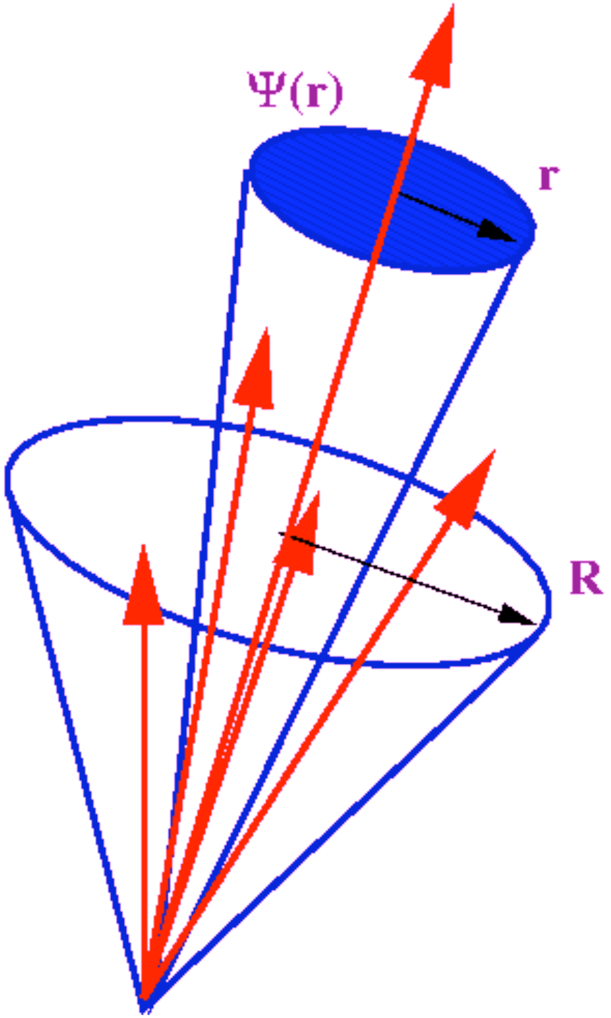
Jet shape

- A way of measuring the energy distribution within a jet
  - Can be defined for any algorithm
  - Generally well modelled by leading-log parton showers
  - Understood well enough to be used to measure α_s



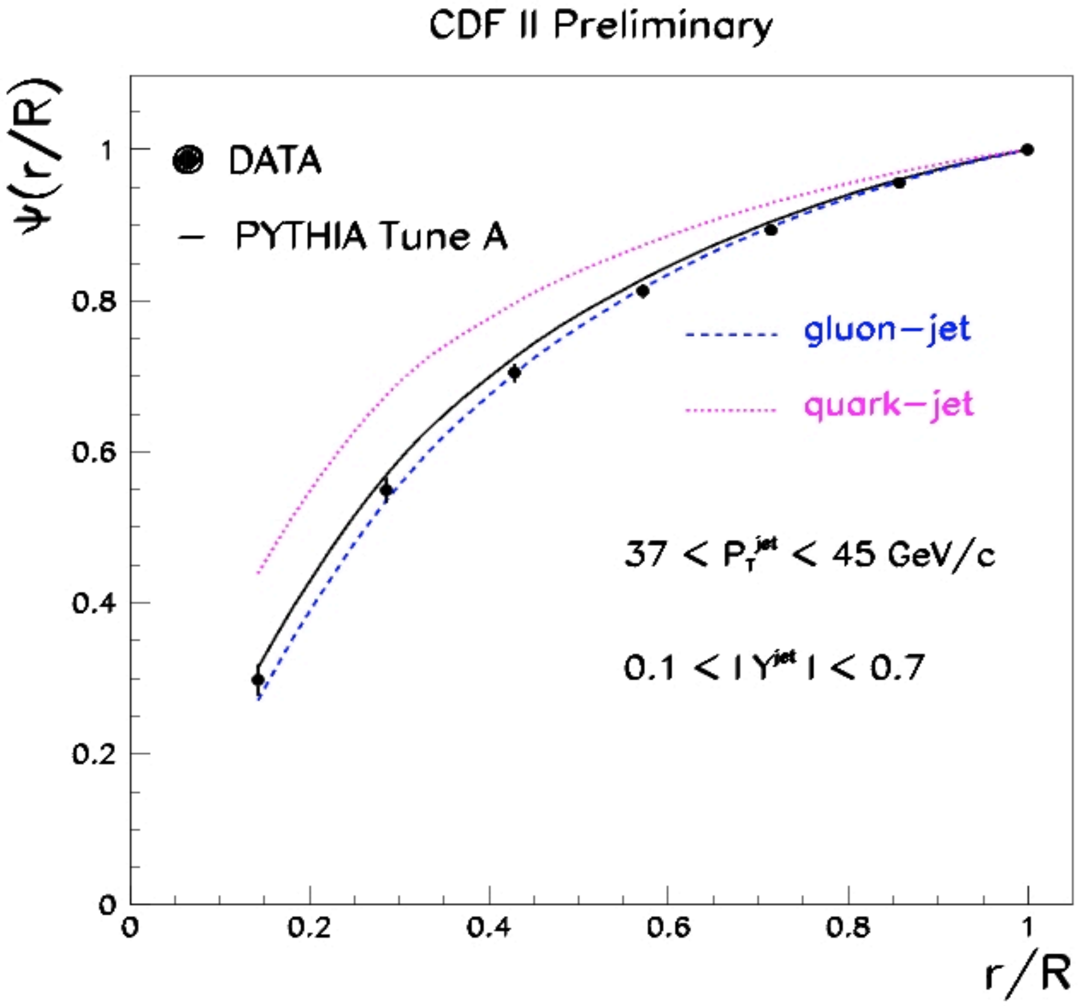
Jet shape

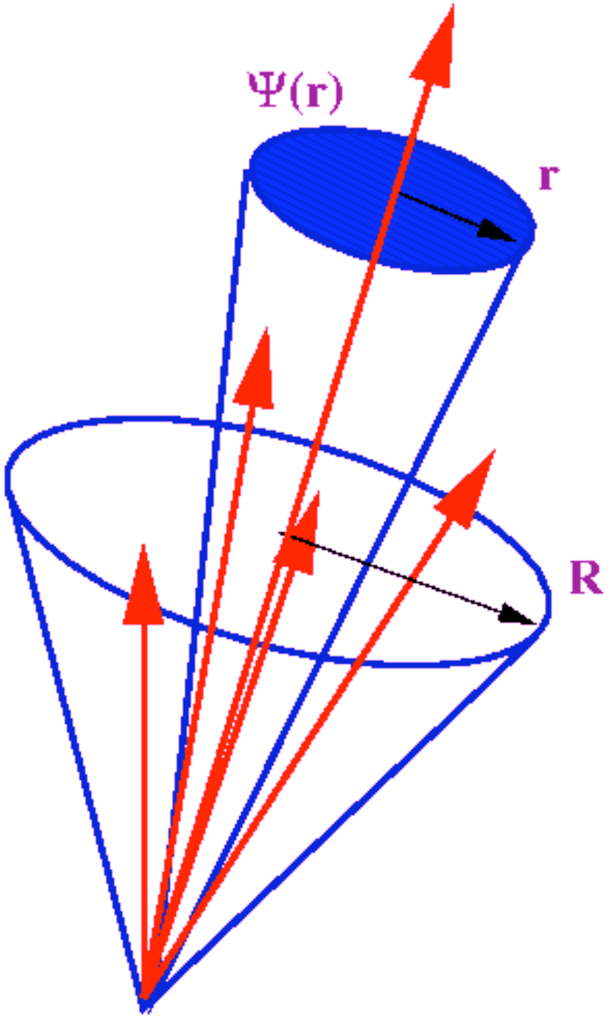




Example: Photoproduction *(ZEUS, Nucl.Phys.B700:3-50,2004.)*

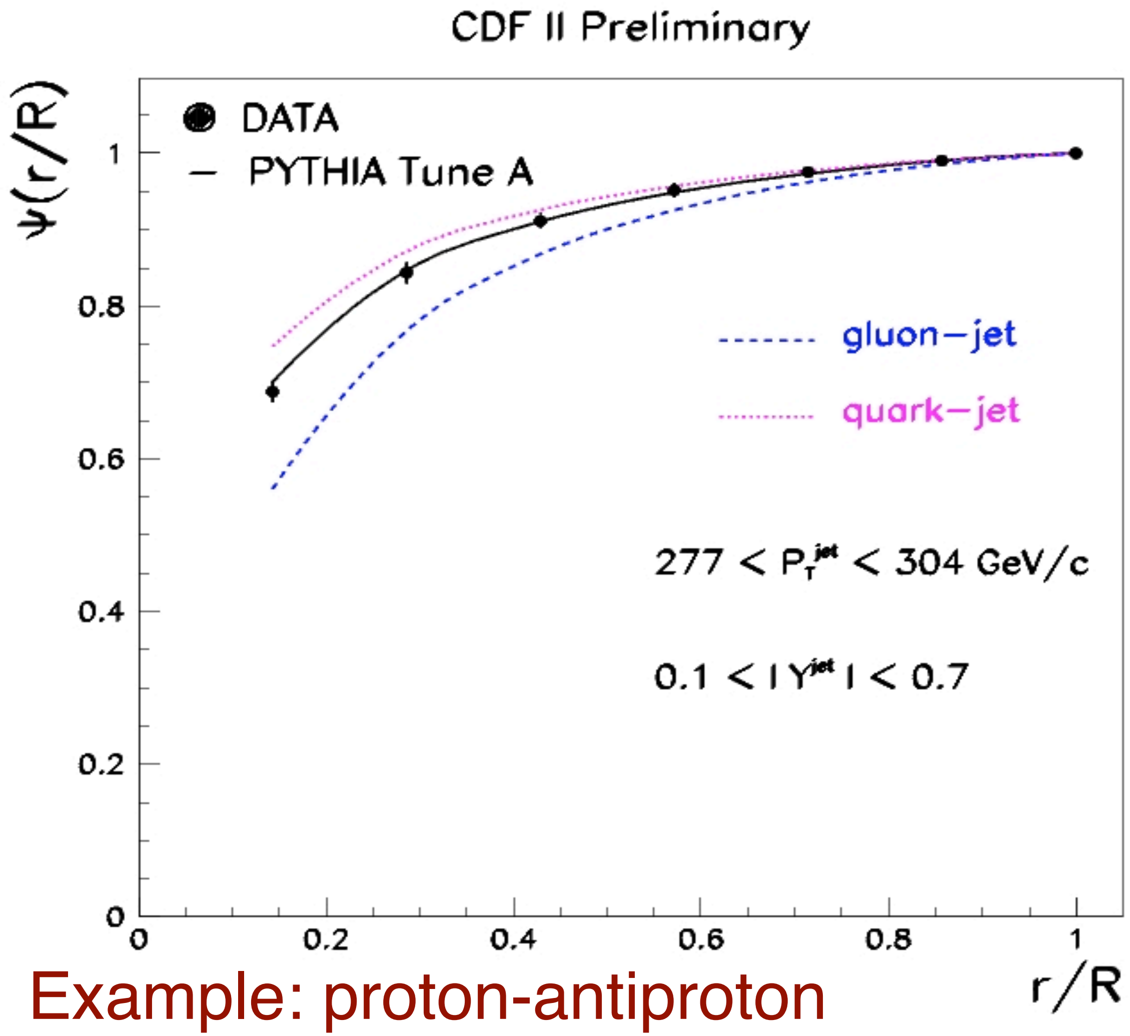
Jet shape

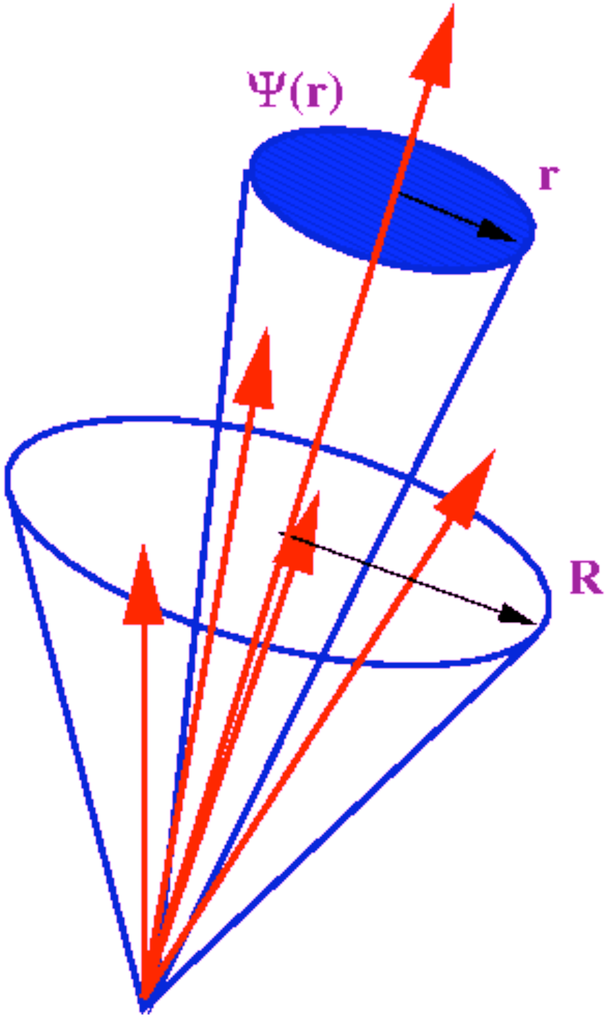




Example: proton-antiproton *(CDF, Phys. Rev. D 71, 112002 (2005) hep-ex/0505013)*

Jet shape

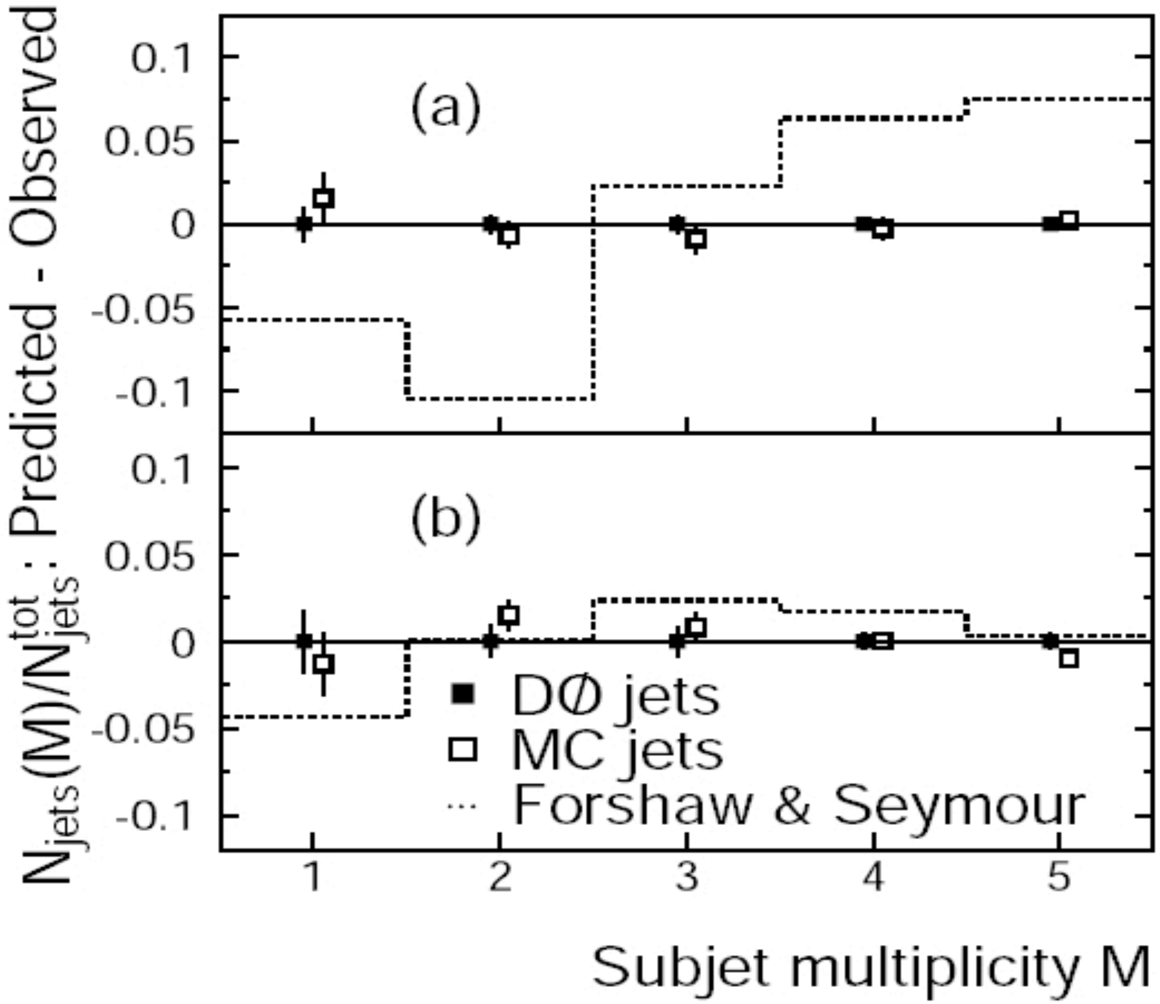




Example: proton-antiproton

(CDF, Phys. Rev. D 71, 112002 (2005) hep-ex/0505013)

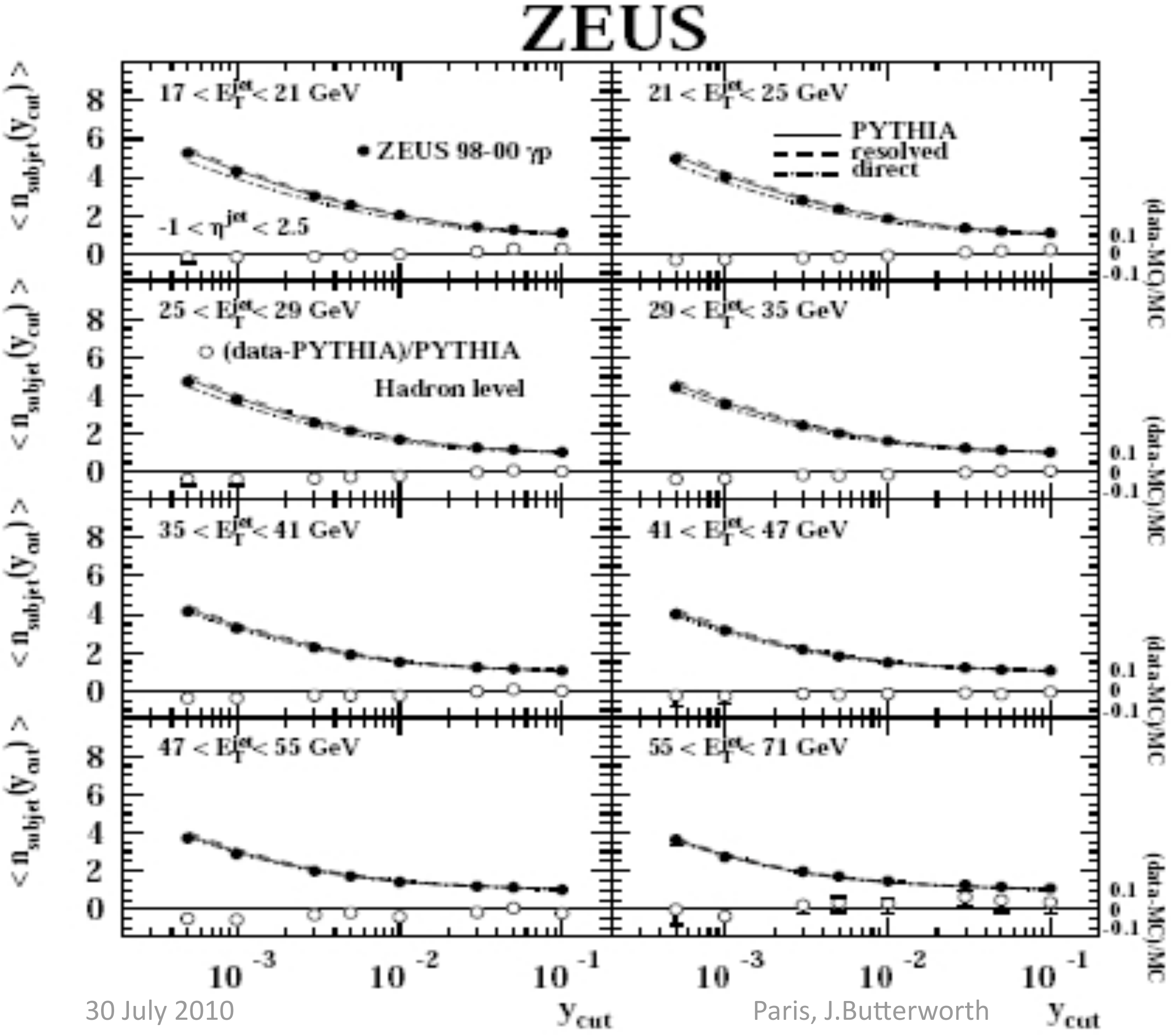
Subjets



- Proton-antiproton Also well described by LL parton shower simulation.

(D0, Phys.Rev.D65:052008,2002.)

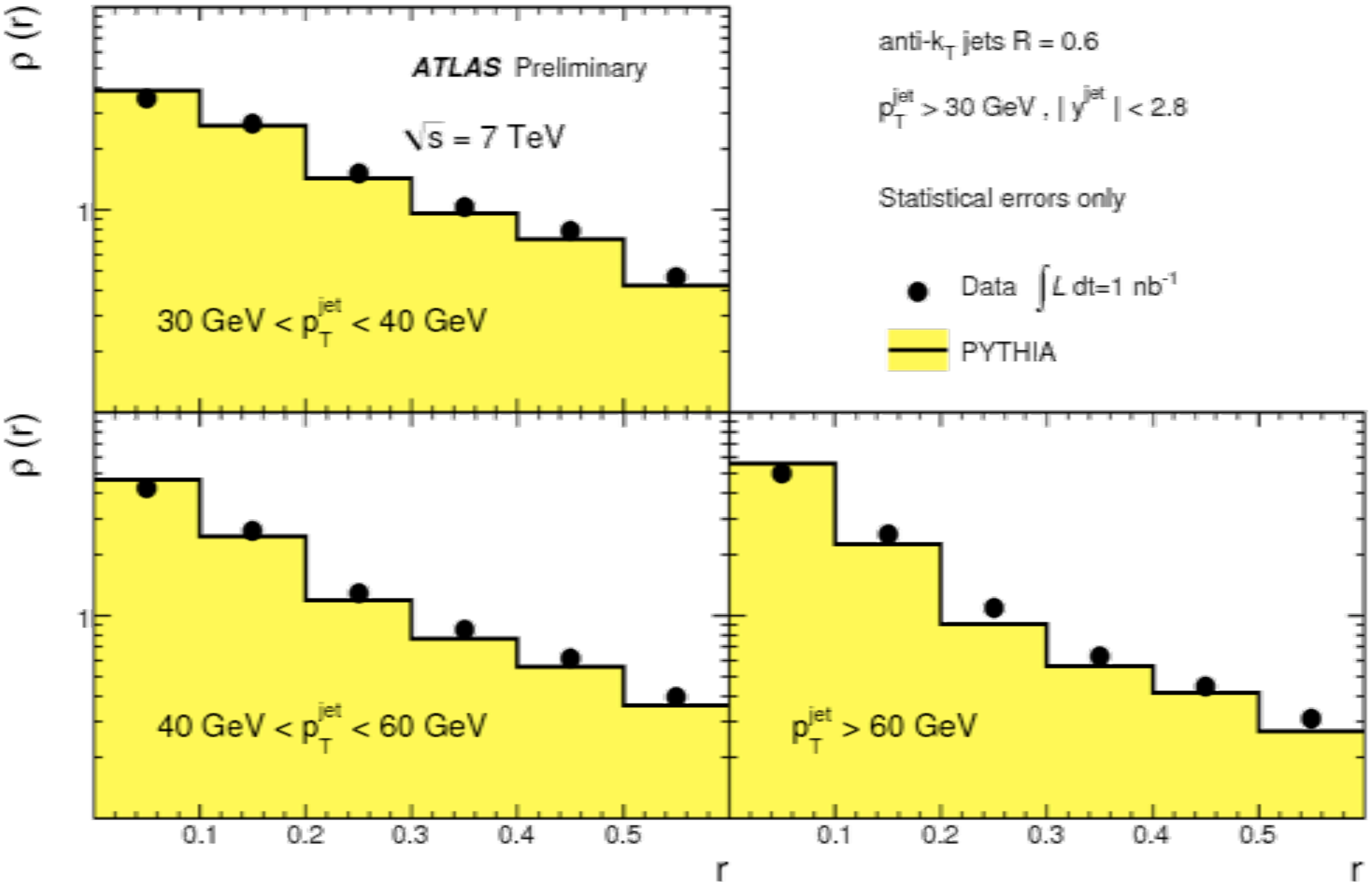
Subjets



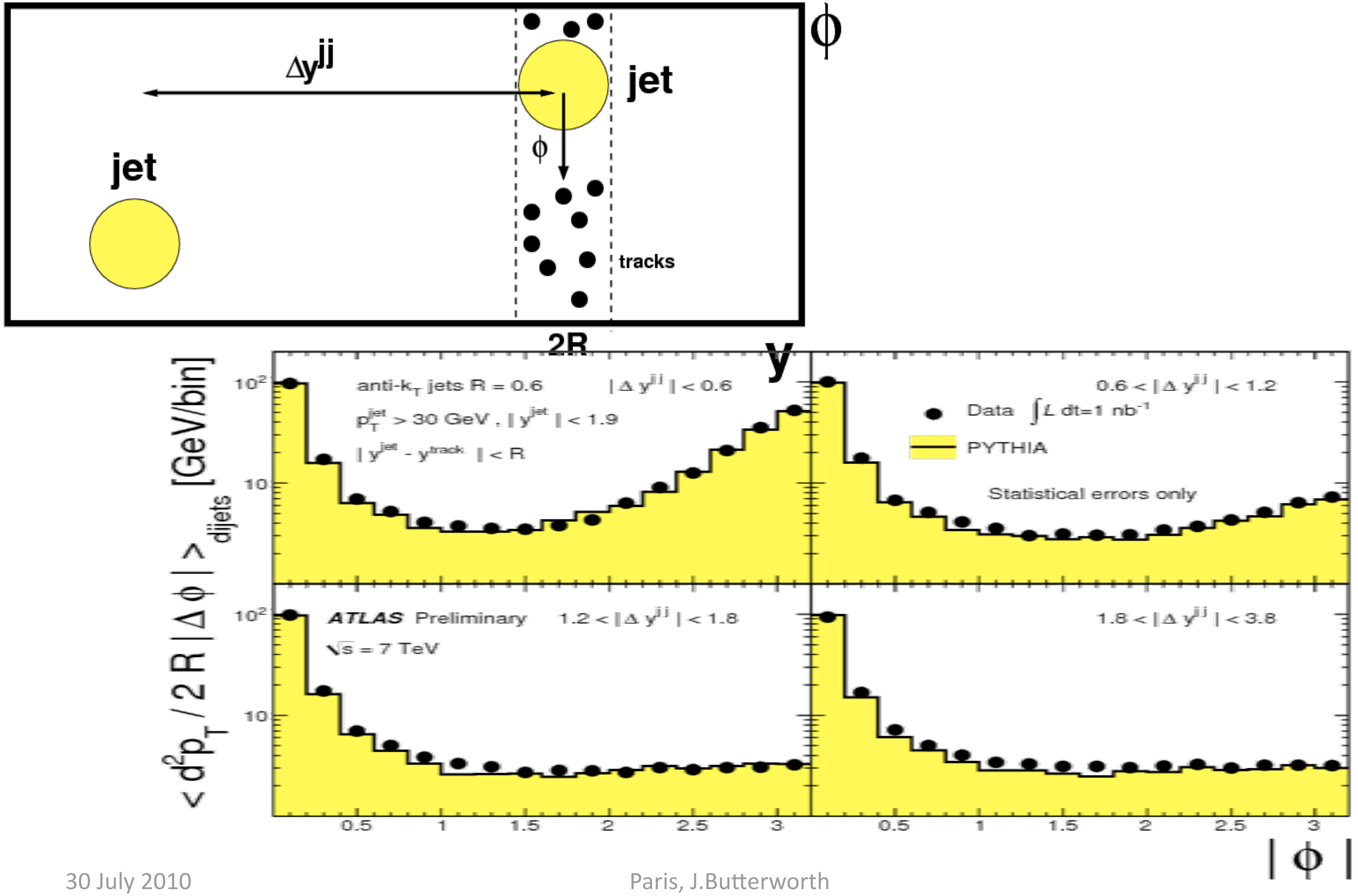
- Photoproduction Also well described by LL parton shower simulation.

(ZEUS, Nucl.Phys.B700:3-50,2004)

Jet substructure in ATLAS data



Jet substructure in ATLAS data



Jet Substructure

- Two goals of the recently developed techniques
  - Improve the single jet mass resolution
  - Background suppression
    - Distinguish between QCD-generated high mass jets and those due to heavy object decays

Improved single jet mass resolution

- First unclustering stages in C/A, throw away softer or more distant partner
  - *JMB, Davison, Rubin, Salam, PRL 100, 242001 (2008).*
  - *Kaplan, Rehermann, Schwartz, Tweedie, PRL 101, 142001 (2008).*
  - *JMB, Ellis, Raklev, Salam, PRL 103, 241803 (2009).*
- "Filtering": Rerun algorithm with tighter distance resolutions
  - *JMB, Davison, Rubin, Salam*
- Variable R parameter
  - *Krohn, Thaler, Wang, JHEP 0906:059,2009.*
- "Pruning" or "Trimming": Remove soft splittings in (re)clustering
  - *S. Ellis, Vermilion, Walsh, PRD 80, 051501 (2009).*
  - *Krohn, Thaler, Wang, arXiv:0912.1342 [hep-ph].*
-

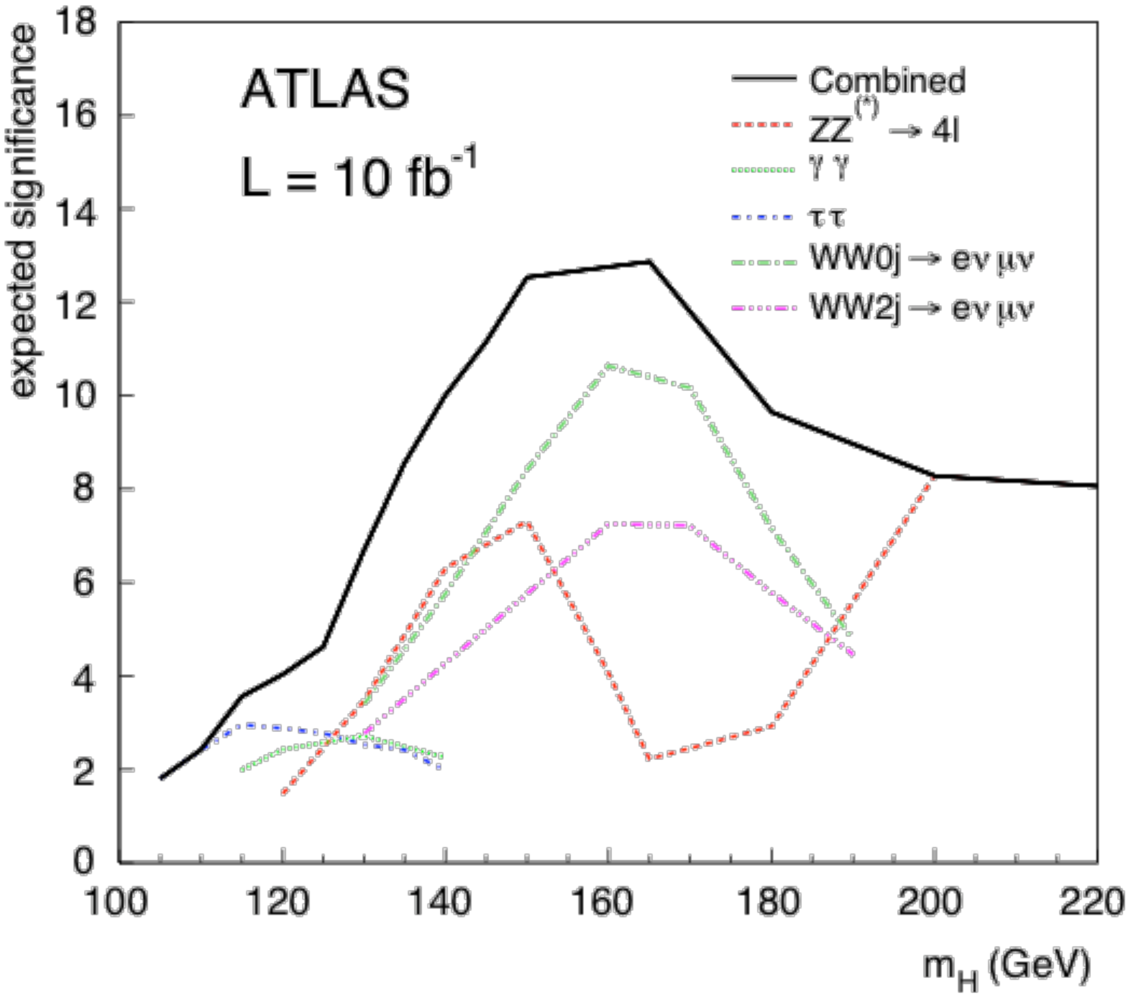
Background Suppression

Distinguish between QCD-generated high mass jets and those due to heavy object decays

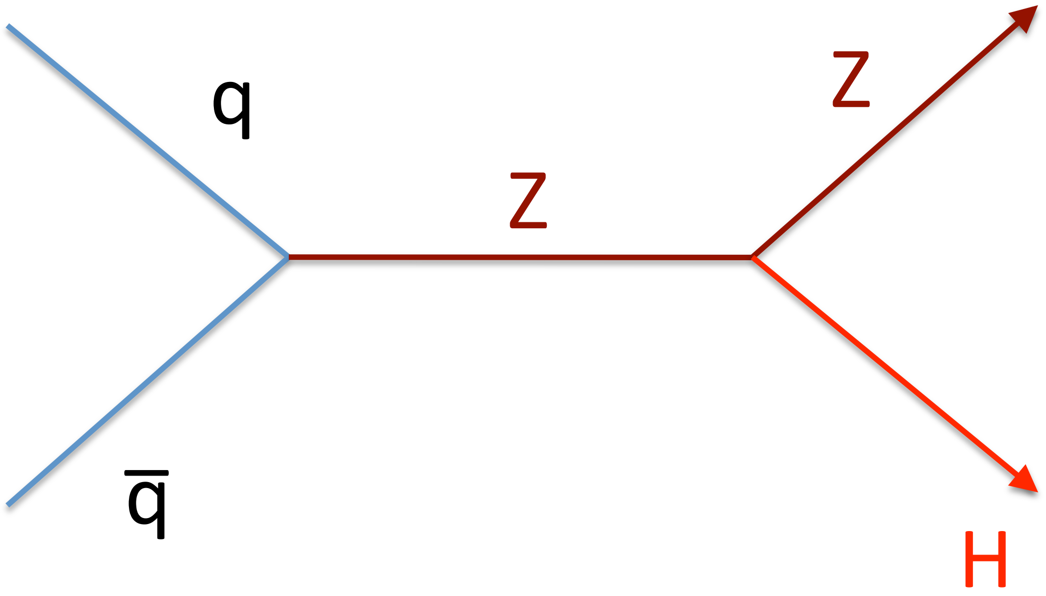
- None-strongly order k_T scale
  - *JMB, Cox, Forshaw, PRD 65; 096014 (2002).*
- Symmetric splitting
  - *Kaplan et al, JMB et al*
- Anomalously large mass drop
  - *JMB et al*
- Analytic jet shapes (planar flow etc)
  - *Almieda et al PRD 79:074017,(2009).*

Low Mass Higgs

- Around 115 GeV no single channel is (was) above 3σ with 10fb^{-1}@14TeV
- Need a combination of channels
- WH, ZH with H ➔bb
  - Principal search channel at Tevatron
  - Not competitive at LHC…

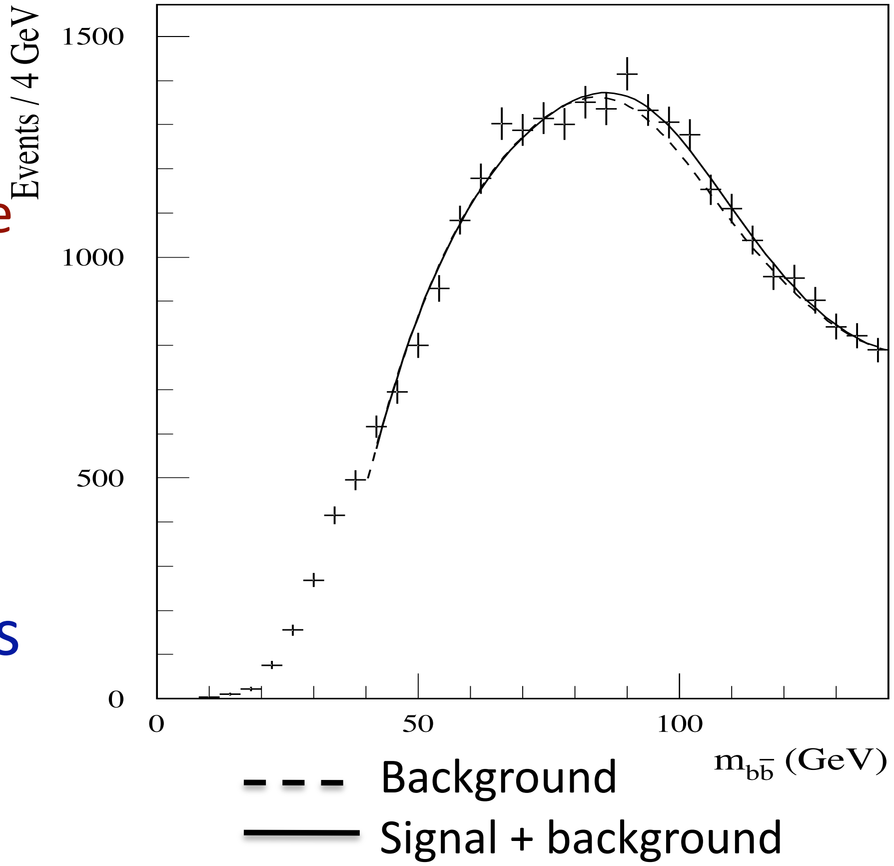


Higgs + (W or Z)



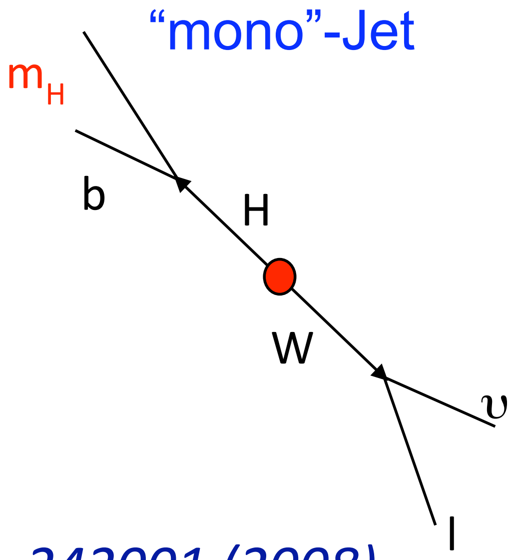
Higgs + (W or Z)

- Example: ATLAS Physics TDR (1999)
  - Poor acceptance
  - Cuts introduce artificial mass scale into the background
  - Top anti-top has a similar mass scale
  - Large combinatorial background
- Signal swamped by backgrounds
  - “very difficult … even under the most optimistic assumptions”



High p_T Higgs and Vector Boson

- By requiring that the Higgs and Vector Boson have a high transverse momentum, we lose a factor of ~20 in cross section
  - However, much of this would have failed other analysis cuts anyway
  - Background cross sections fall by a bigger factor (typically t-channel not s-channel)
- W/Z and H are all central
  - Better b-tagging, better jet resolution
- W/Z and H decay products collimated
  - Simpler topology, fewer combinatorials
  - Difficult for tops to fake this
- Z ➔ neutrinos becomes visible
  - High missing E_T
- *JMB, Davison, Rubin, Salam, Phys. Rev. Lett. 100, 242001 (2008)*

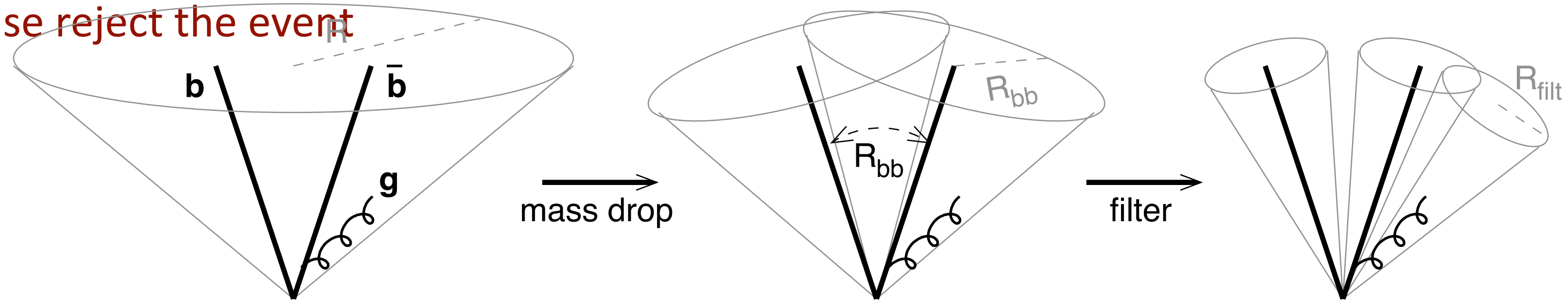


Sub-jet analysis

- Cambridge/Aachen algorithm
  - Dokshitzer et al ‘97, Wengler and Wobisch ‘98
- Like “k_T without the k_T”
  - Work out $\Delta R_{ij} = \sqrt{(\Delta\phi^2 + \Delta y^2)}$ between all pairs of objects
  - Recombine the closest pair
  - Repeat until all objects are separated by $\Delta R_{ij} > R$
- We tried several values for R;
  - Main value chosen: $R = 1.2$
  - best value depends on p_T cut
  - Sensitivity not strongly dependent on the p_T / R combination
- Having clustered an event this way, can then work through backwards to analyse a particular jet.

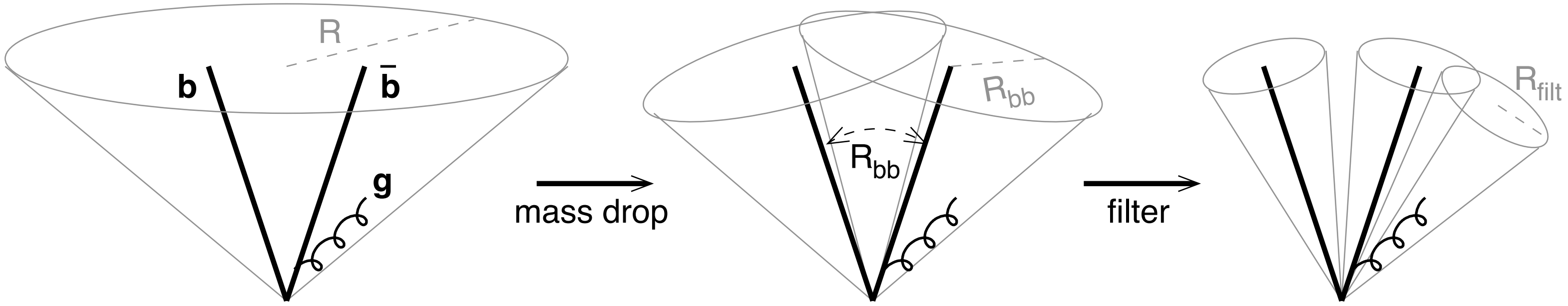
Sub-jet analysis

1. Start with Higgs candidate jet (highest p_T jet in acceptance) with mass m)
2. Undo last stage of clustering (reduce radius to R_{12})
   $J \rightarrow J_1, J_2$
3. If $\max(m_1,m_2) < 2m/3$
   Call this a “mass drop”. This fixes the optimal radius for reconstructing the Higgs decay. Keep the jet J and call it the Higgs candidate.
   Else, go back to 2
4. Require $Y_{12} > 0.09$
   Dimensionless rejection of asymmetric QCD splitting
   Else reject the event
5. Require J_1, J_2 to each contain a b-tag
   Else reject the event

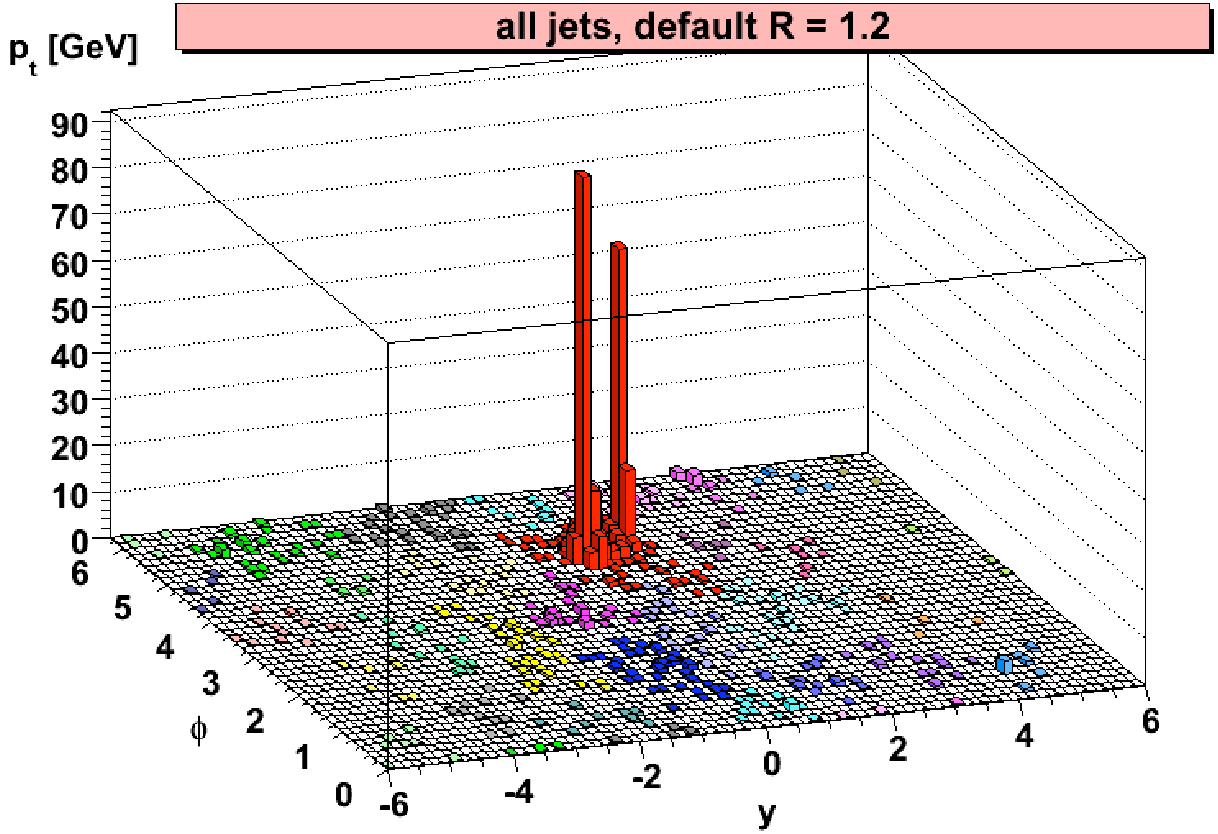


Sub-jet analysis

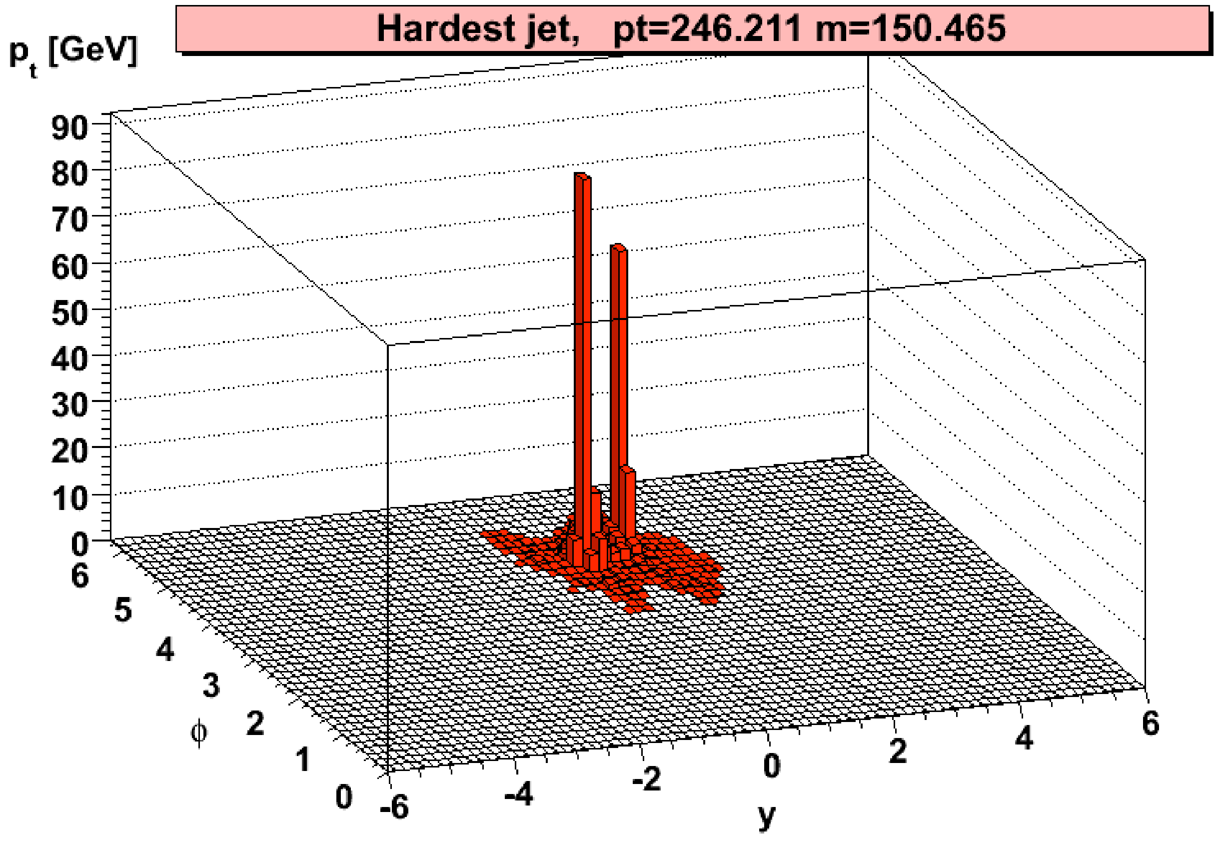
6. Define $R_{filt} = \min(0.3, R_{bb}/2)$
   - Make use event-by-event of the known Higgs decay radius
   - Angular ordering means this is the characteristic radius of QCD radiation from Higgs products
   - Stuff outside of this is likely to be underlying event and pileup.
7. Recluster, with Cambridge/Aachen, $R = R_{filt}$
8. Take the 3 hardest subjets and combine to be the Higgs
   - b, anti-b and leading order final state gluon radiation
9. Plot the mass

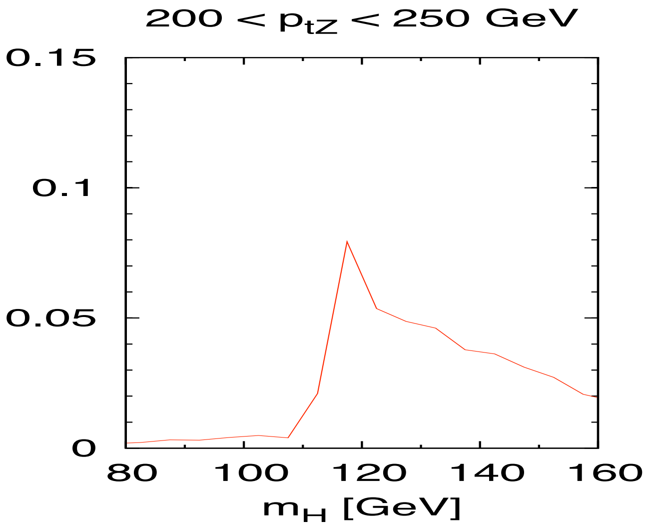


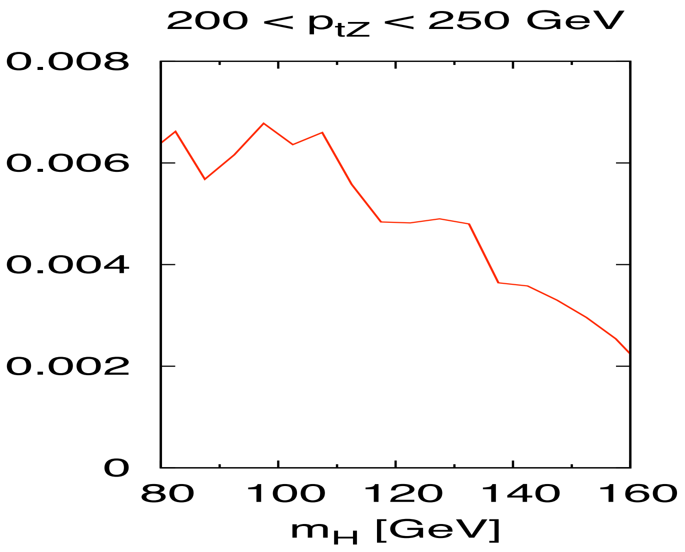
Improved subjet analysis



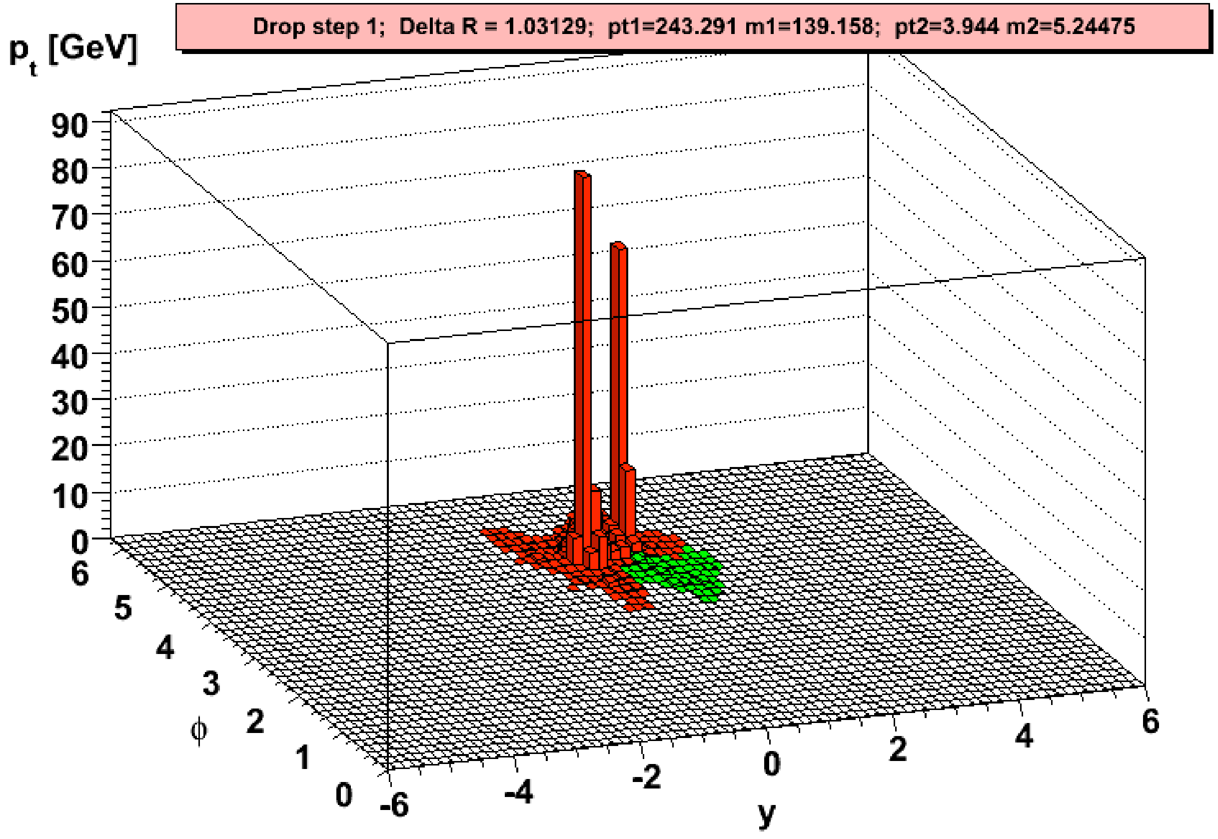
Improved subjet analysis

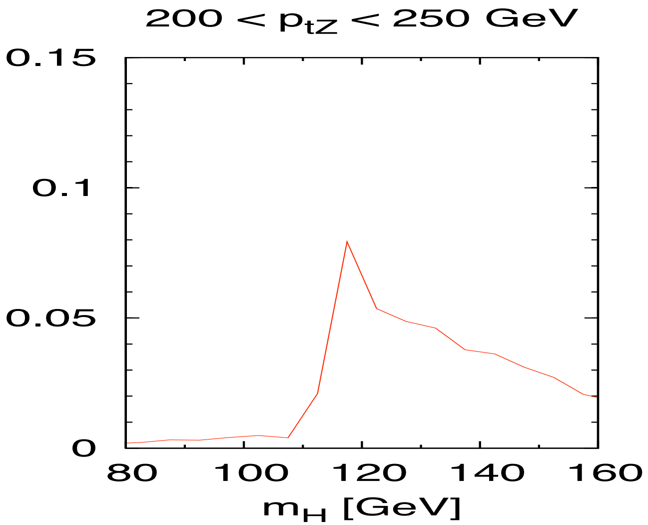


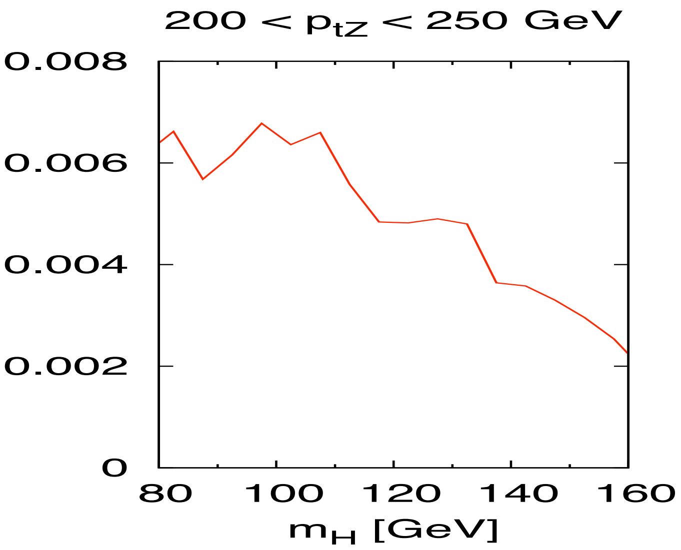




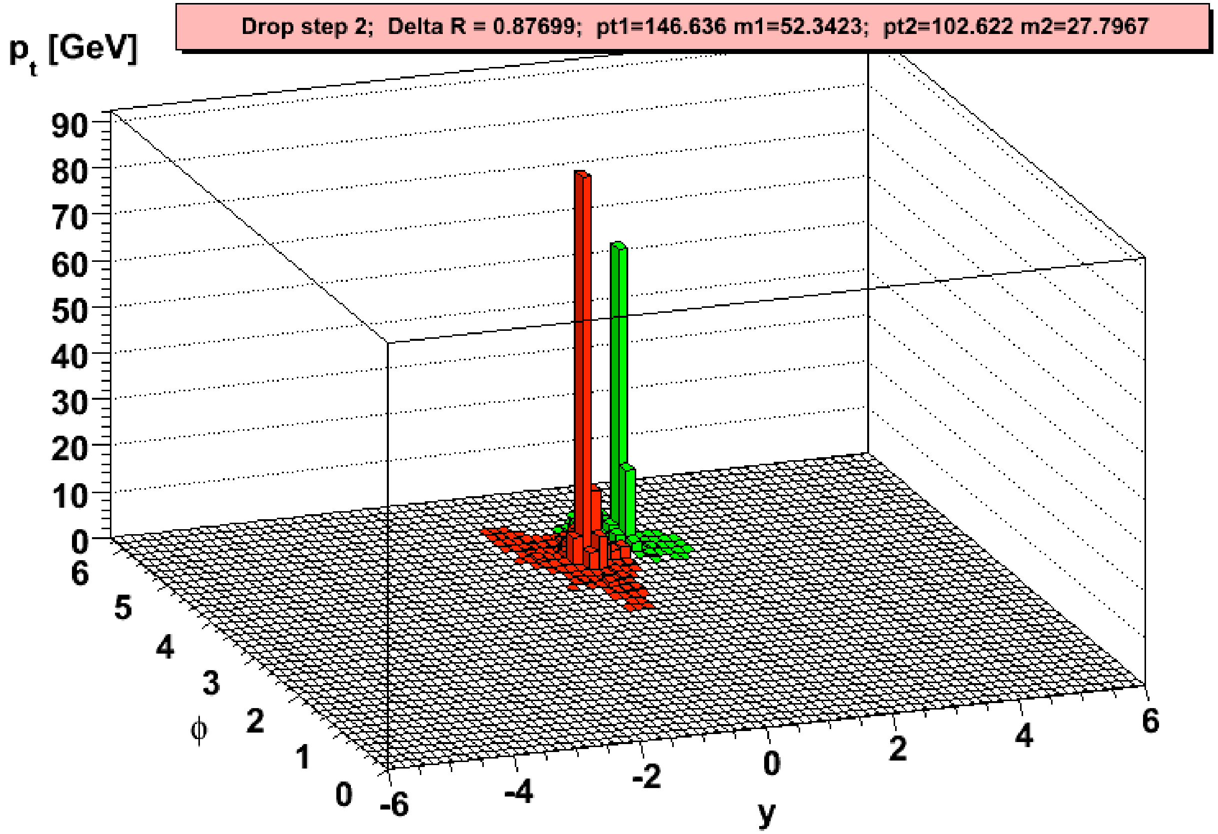
Improved subjet analysis

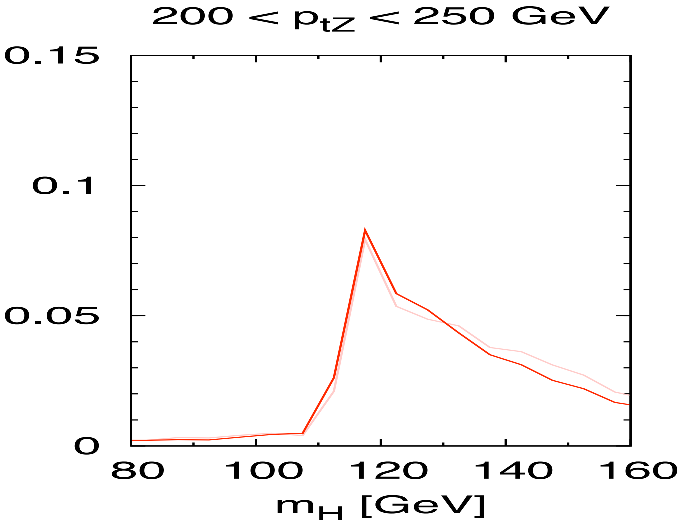


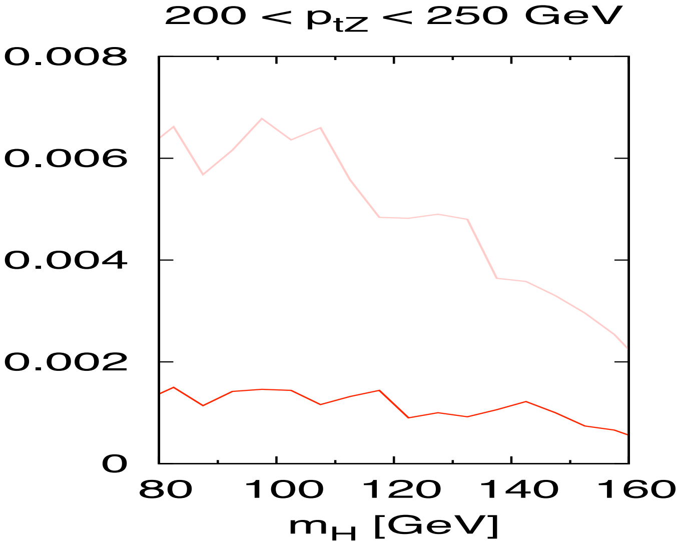




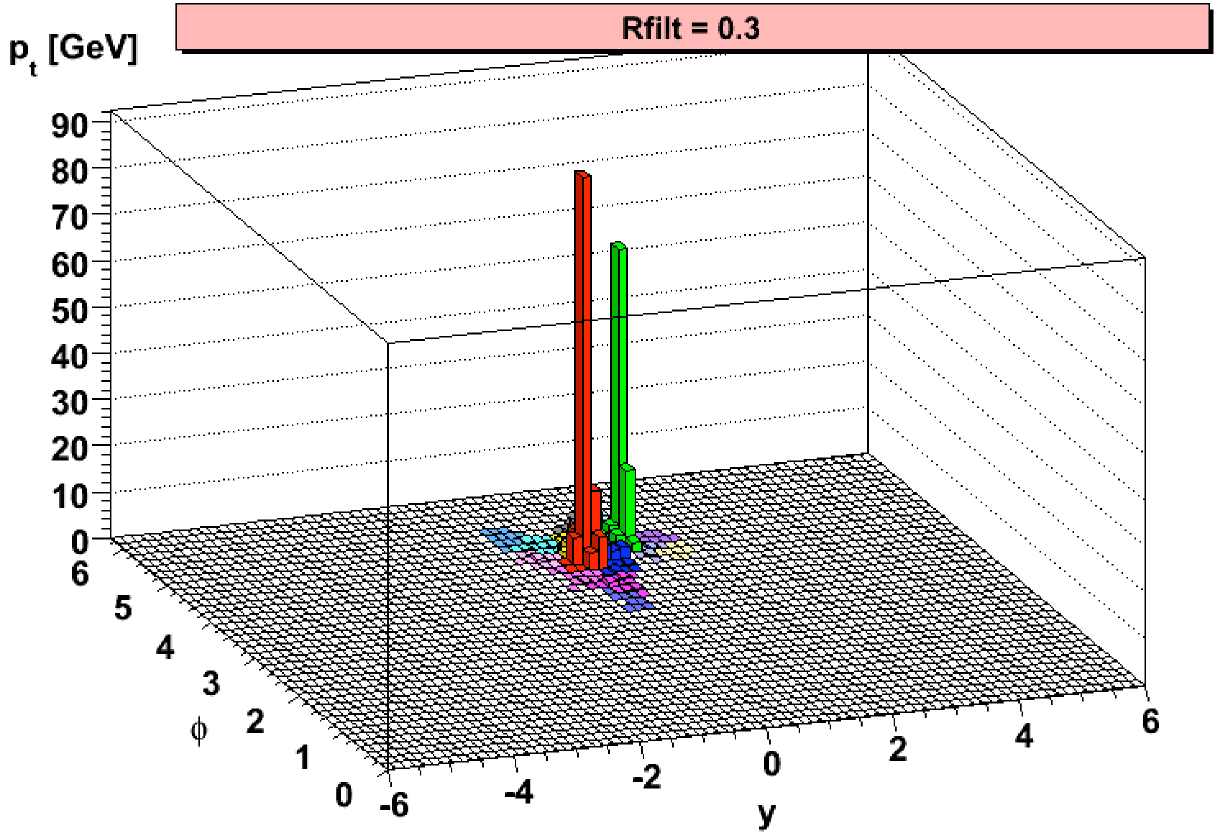
Improved subjet analysis

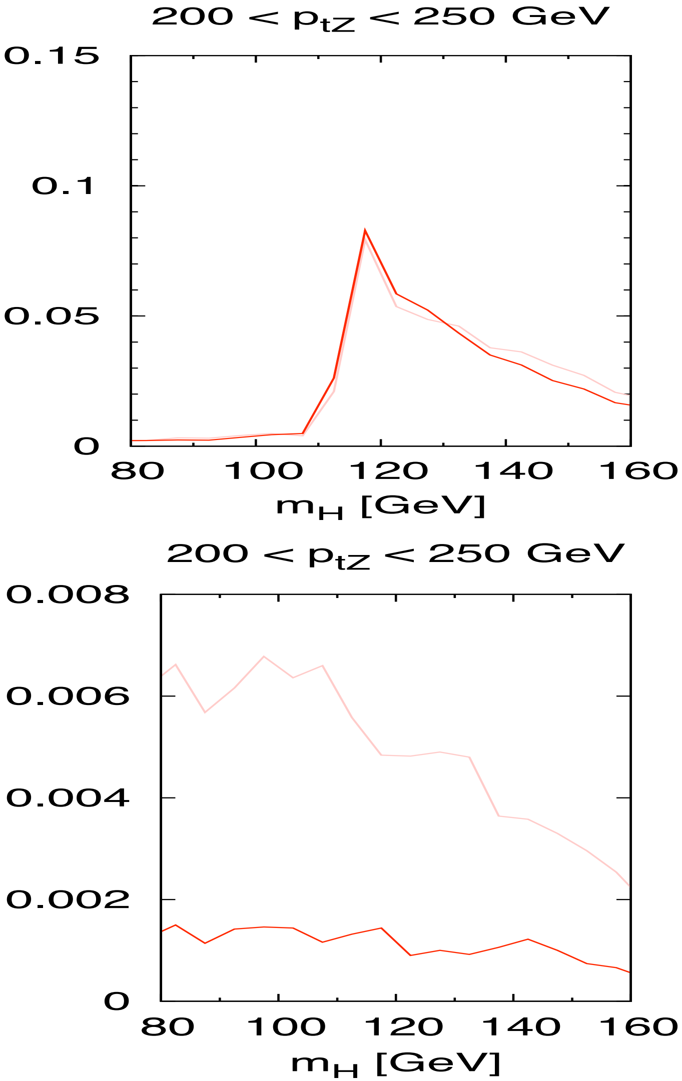




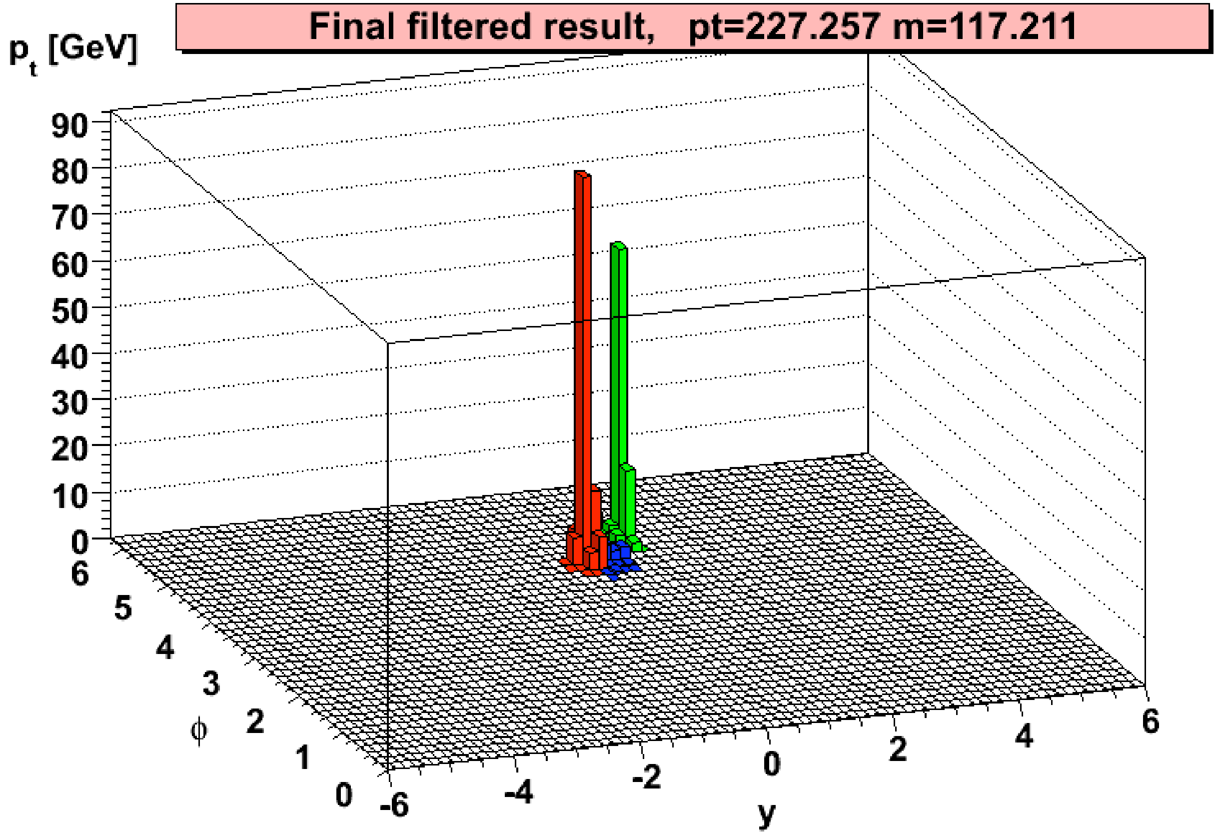


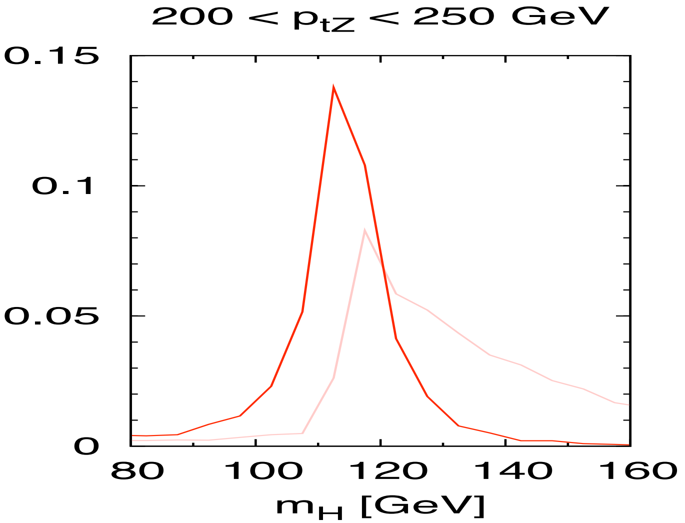
Improved subjet analysis

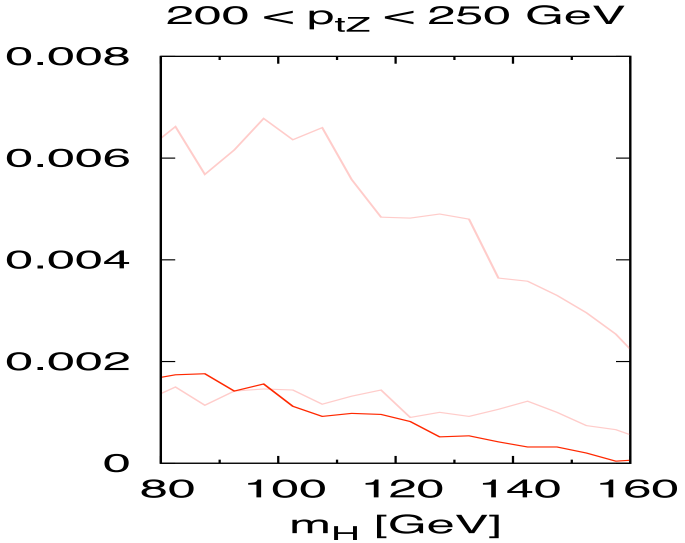




Improved subjet analysis



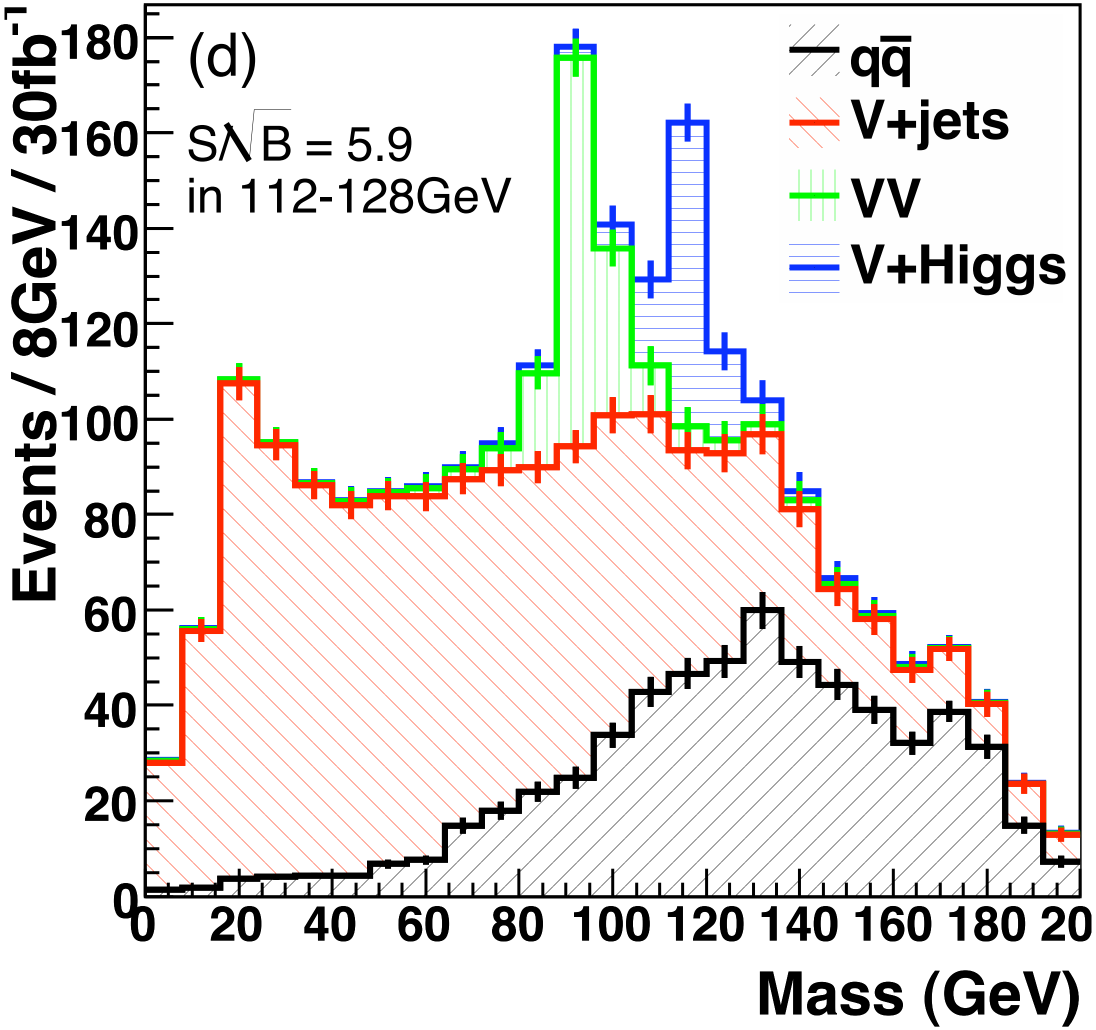




Analysis Overview

- Consider three cases
  - HZ, Z ➔ ee, $\mu\mu$
  - HZ, Z ➔ $\nu\nu$
  - HW, W➔ e/μ + ν
- Three non-overlapping selections
  - l + missing E_T + jet **("Leptonic W case")**
  - $l^+ l^-$ + jet **("Leptonic Z case")**
  - Missing E_T + jet **("Z ➔ neutrinos case")**
- Common cuts
  - p_T Higgs candidate > 200 GeV, p_T VB candidate > 200 GeV
  - $|\eta| < 2.5$ (Higgs candidate and leptons)
  - $p_T > 30$ GeV, $|\eta| < 2.5$ (leptons)
  - No extra b jet ($p_T > 30$ GeV, $|\eta| < 2.5$) or lepton passing these cuts.

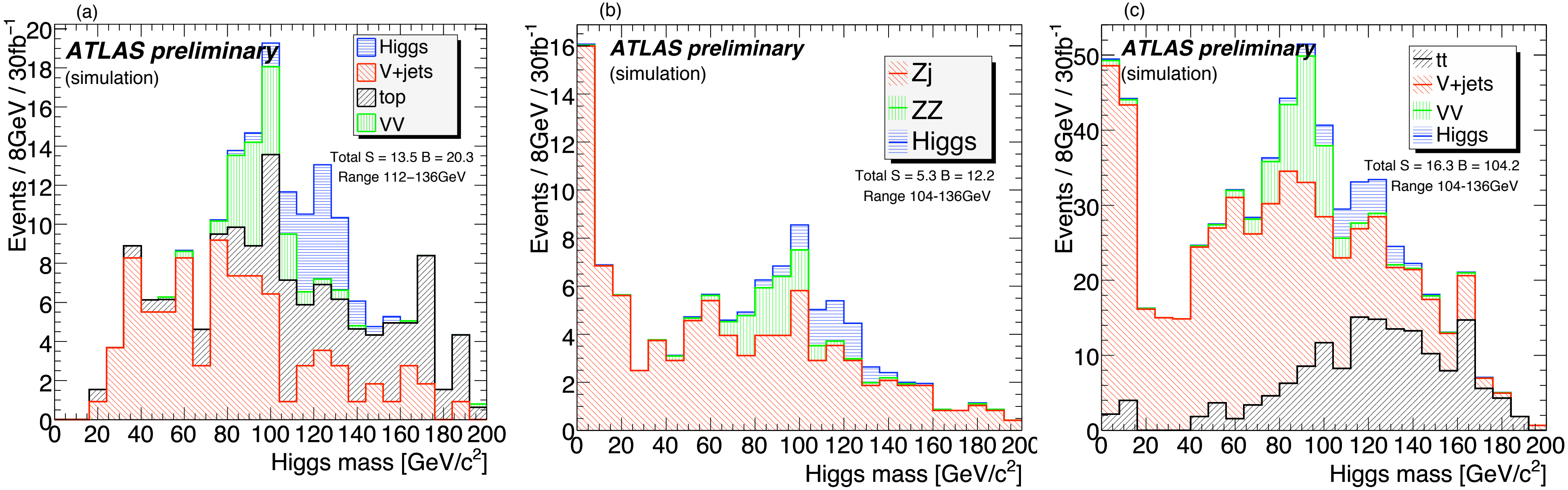
Combined particle-level result



- Note excellent Z peak for calibration
- 5.9 σ; potentially very competitive
- bb branching information critical for extracting Higgs properties
  - *“Measuring the Higgs sector” Lafaye, Plehn, Rauch, D.Zerwas, Duhrssen, arXiv:0904.3866 [hep-ph]*
- Studies within ATLAS are promising and nearly public.

Fully simulated detector

- Included trigger, real ATLAS b-tagging algorithm, detailed tracking & calorimeter
- Also include Wt background omitted from initial study.
- Also included study of Wbb ME vs Wg->Wbb
- Slight degradation w.r.t particle level, but still very promising



High p_T top and tt resonances

- *Kaplan, Rehermann, Schwartz, Tweedie*
- Use C/A technique, optimise b-tagging, use helicity information from top decay

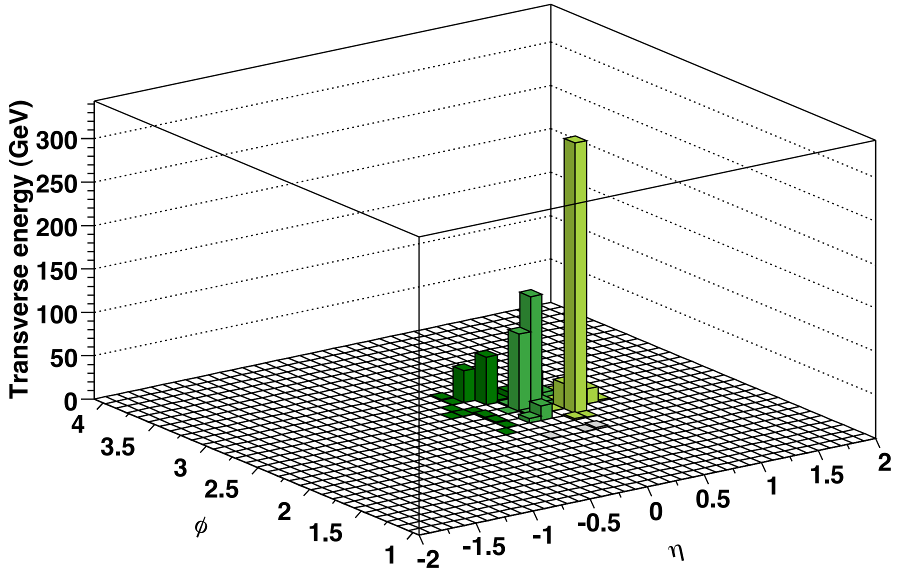
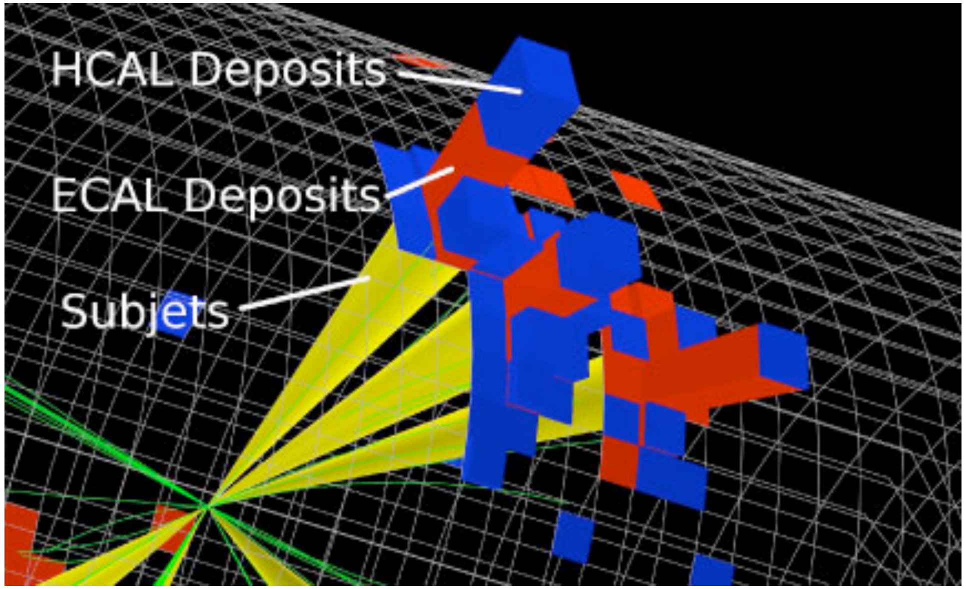
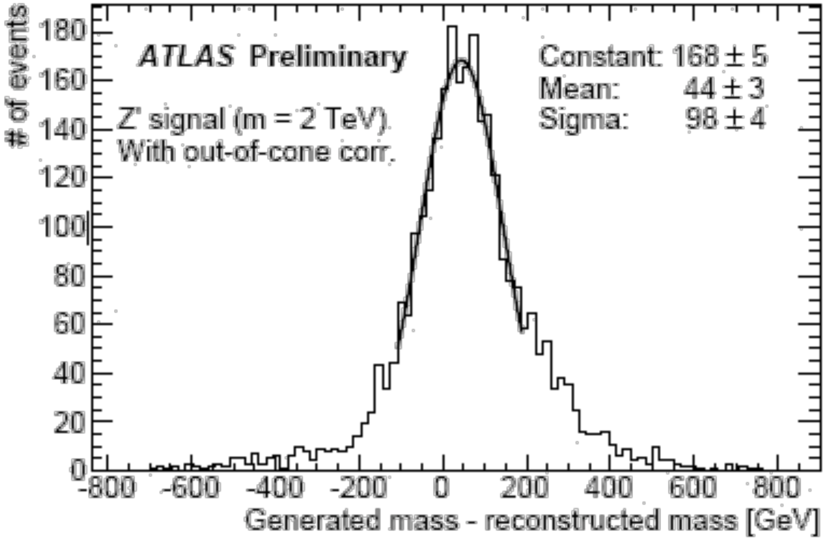


FIG. 1: A typical top jet with a p_T of 800 GeV at the LHC. The three subjets after top-tagging are shaded separately.

High p_T top and tt resonances

- Ysplitter technique used by ATLAS (ATLA-PHYS-PUB-2009-081)
- Improved C/A technique used by CMS ----------->
- Both feasible, and some kind of subjet analysis is required to obtain best sensitivity.





Higgs and top together

- Combine techniques for top and Higgs tagging to improve sensitivity to Higgs in ttH channel
  - *Plehn, Salam, Spannowsky arXiv:0910.5472 [hep-ph]*
- Also use Higgs techniques in new physics events
  - *Kribs, Martin, Roy, Spannowsky arXiv: 0912.4731 [hep-ph]*

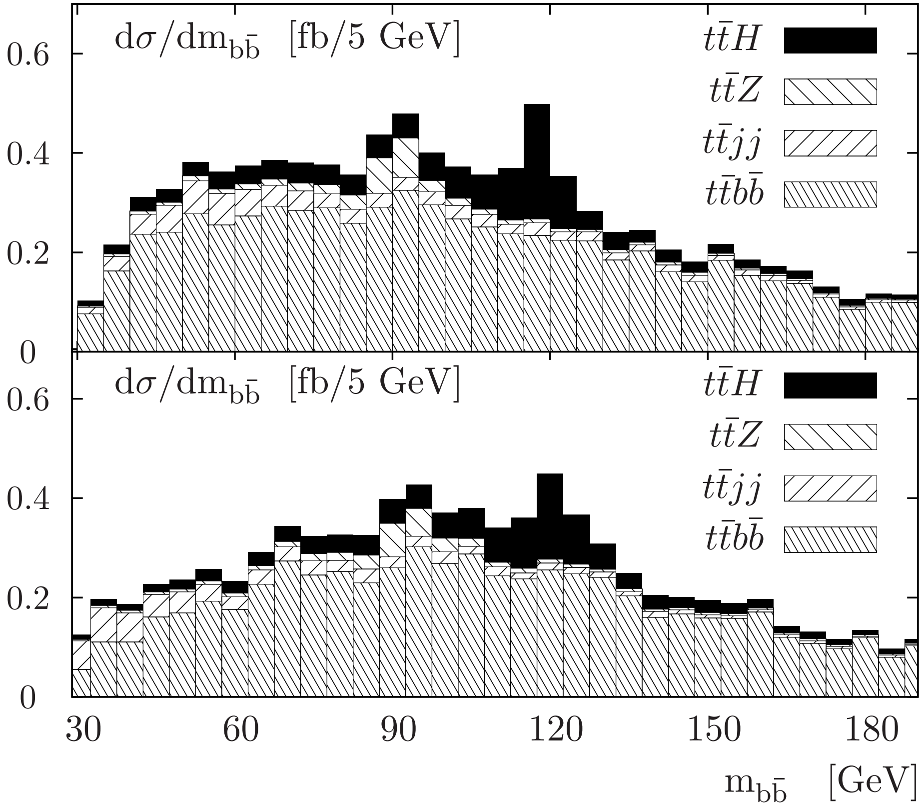


FIG. 3: Reconstructed bottom-pair mass $m_{bb}^{\rm rec}$ for signal ($m_H = 120$ GeV) and backgrounds without (upper) and including (lower) underlying event. The distributions shown include three b tags.

Summary

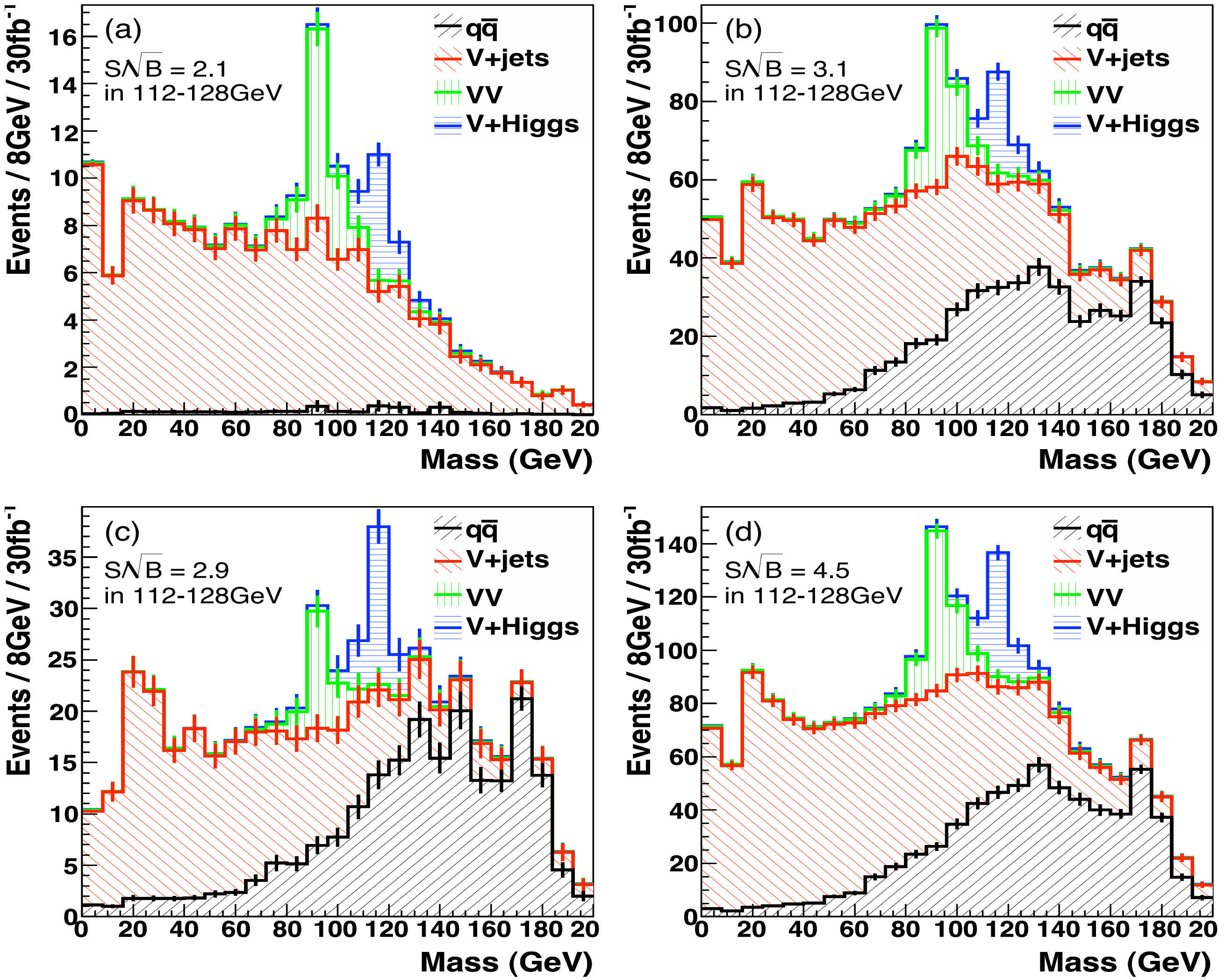
- Jet finding, jet mass, sub-jet technology, and the associated understanding of QCD, have come a long way since (and because of) the previous round of colliders.

- Subjet analysis has dramatic benefits in H->bb search channels; looks very promising and practical, even after full detector simulation. (7 TeV needs more investigation)

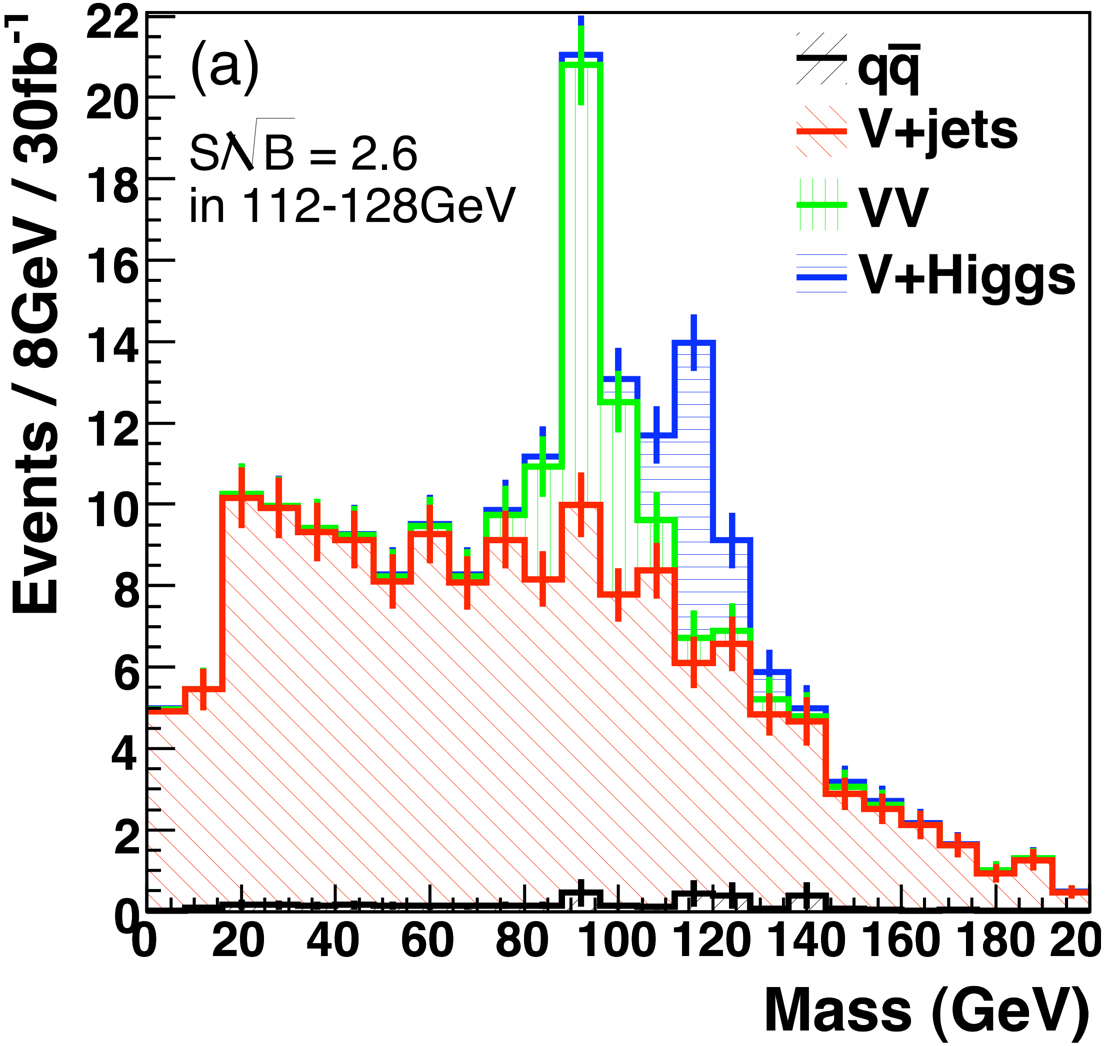
- At the LHC we have interesting physics at $\mathcal{O}$(100 GeV), and phase space open at $\mathcal{O}$(1 TeV). This means that a single jet often contains interesting physics. We probably haven't yet appreciated all the consequences of this qualitatively new feature of physics beyond the EWSB scale

EXTRAS

Extra

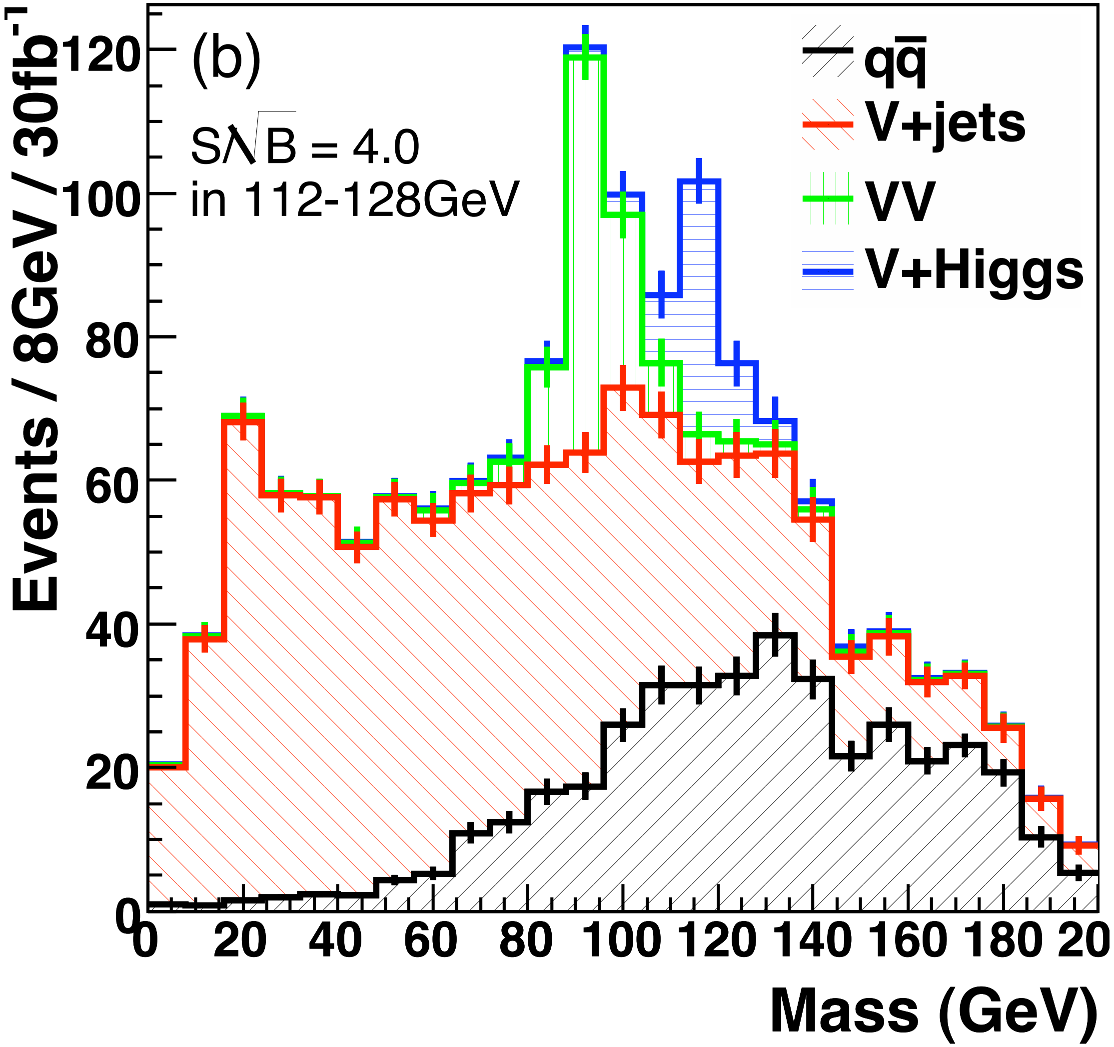

(a)
S/√B = 2.1
in 112-128GeV
(b)
S/√B = 3.1
in 112-128GeV
(c)
S/√B = 2.9
in 112-128GeV
(d)
S/√B = 4.5
in 112-128GeV
qq̄
V+jets
VV
V+Higgs
Events / 8GeV / 30fb⁻¹
Mass (GeV)
0
20
40
60
80
100
120
140
160
180
20

Leptonic Z case



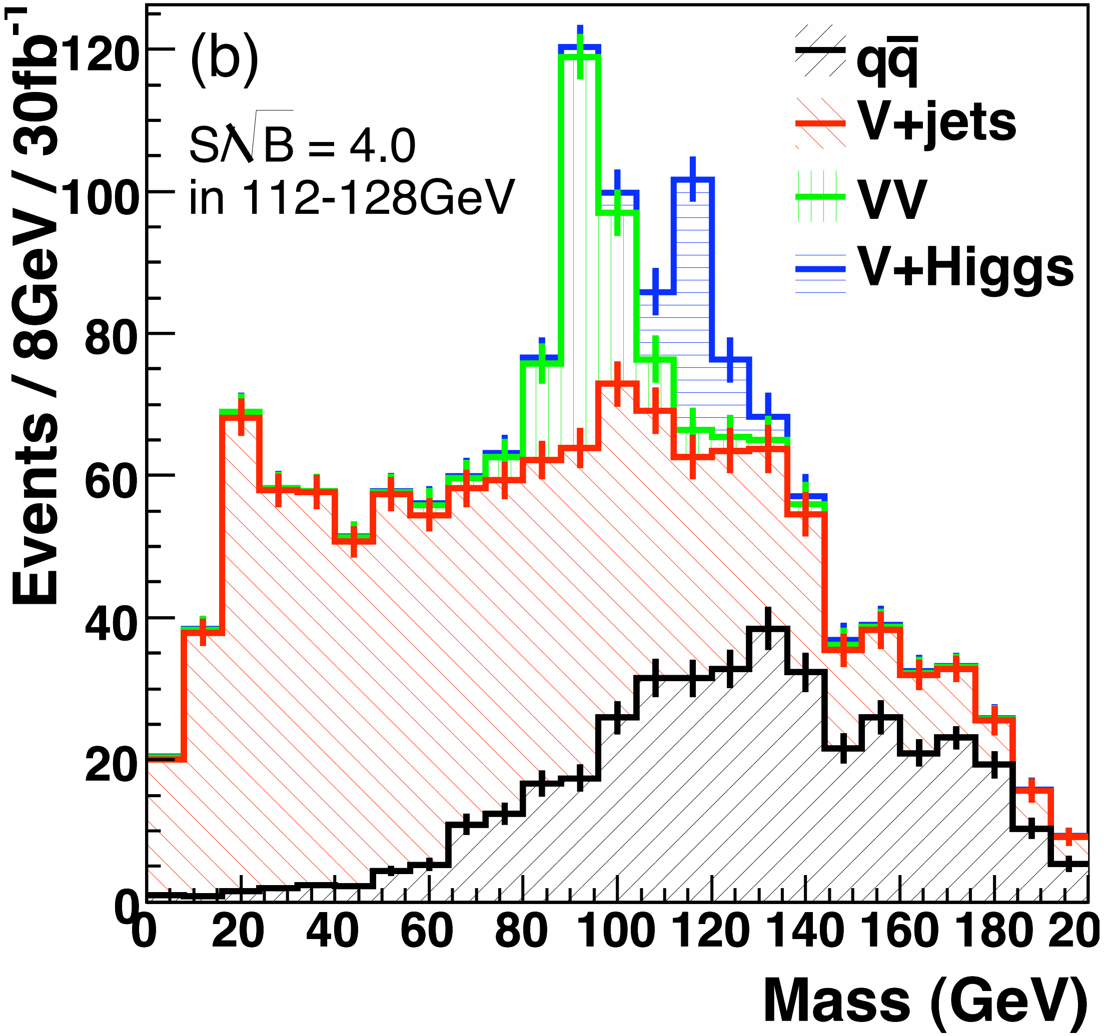
- Common cuts, plus require
    - dilepton mass between 80 and 100 GeV

Z ➔ neutrinos case



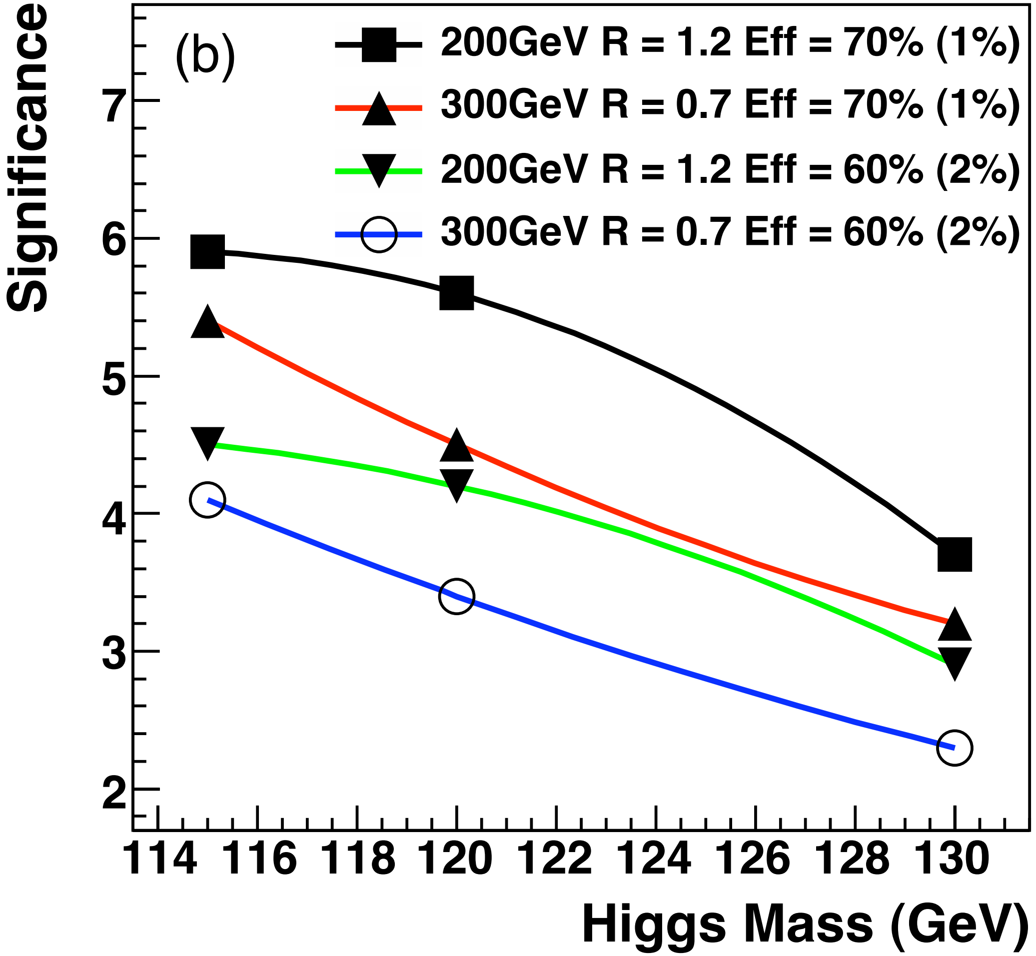
- Common cuts, plus require
    - Missing E_T greater than 200 GeV

Leptonic W case



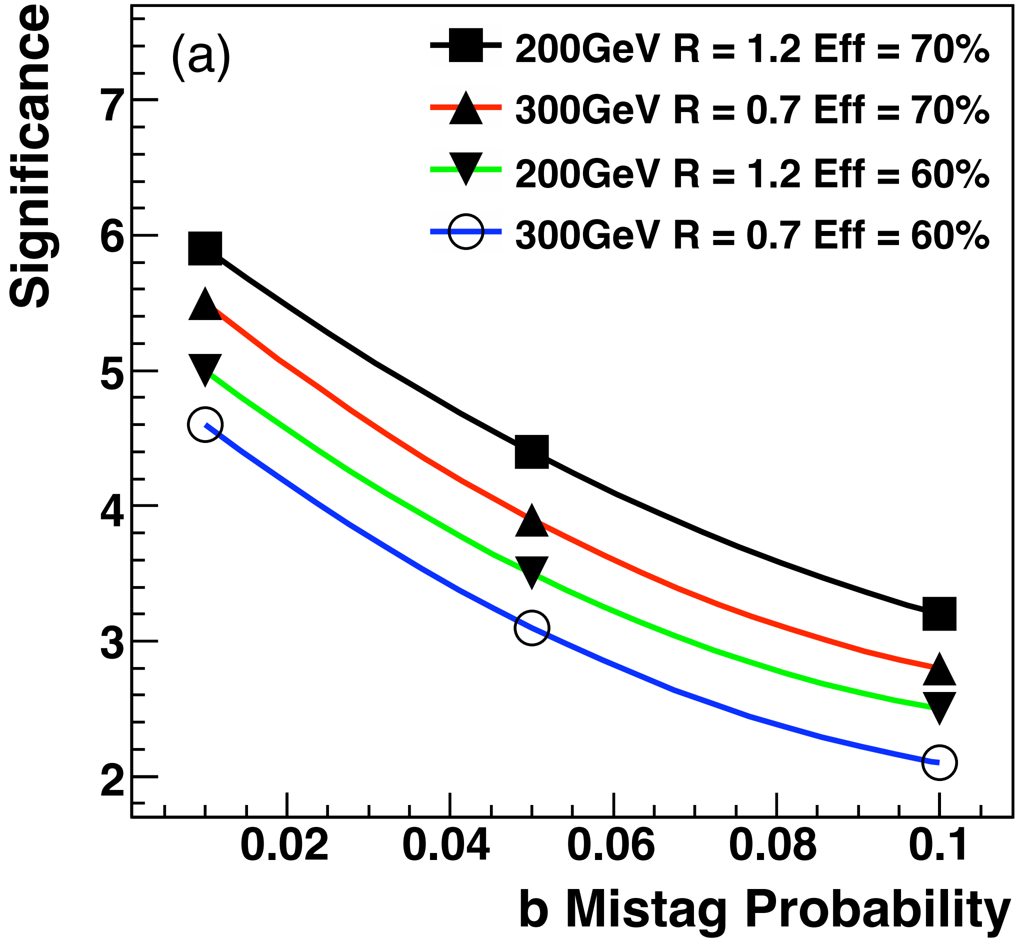
- Common cuts, plus require
  - missing $E_T > 30$ GeV
  - Lepton and missing E_T consistent with a W
  - No extra jets $|\eta| < 3$, $p_T > 30$ GeV

Combined result



- Note excellent Z peak for calibration
- 5.9 σ; potentially very competitive
- Also, unique information on relative coupling of H to Z and W.
- Reasonably good up to about 130 GeV.

Combined result



- Strong dependence of b-tag performance
- Significance about 4.5σ for 60% efficiency/factor 50 rejection
- Also strong dependence on jet mass resolution (not shown)

EXTRAS

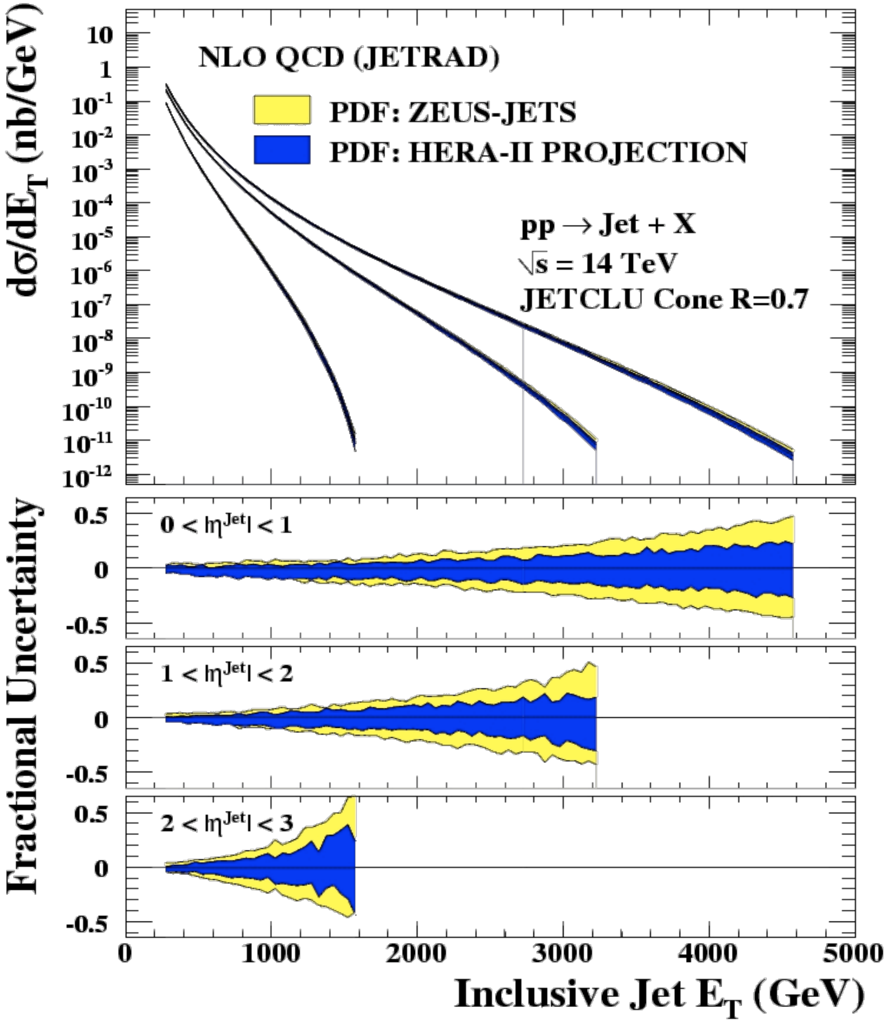
- *Todo:)*
  - *Reduce LSP*
  - *Include ttH*

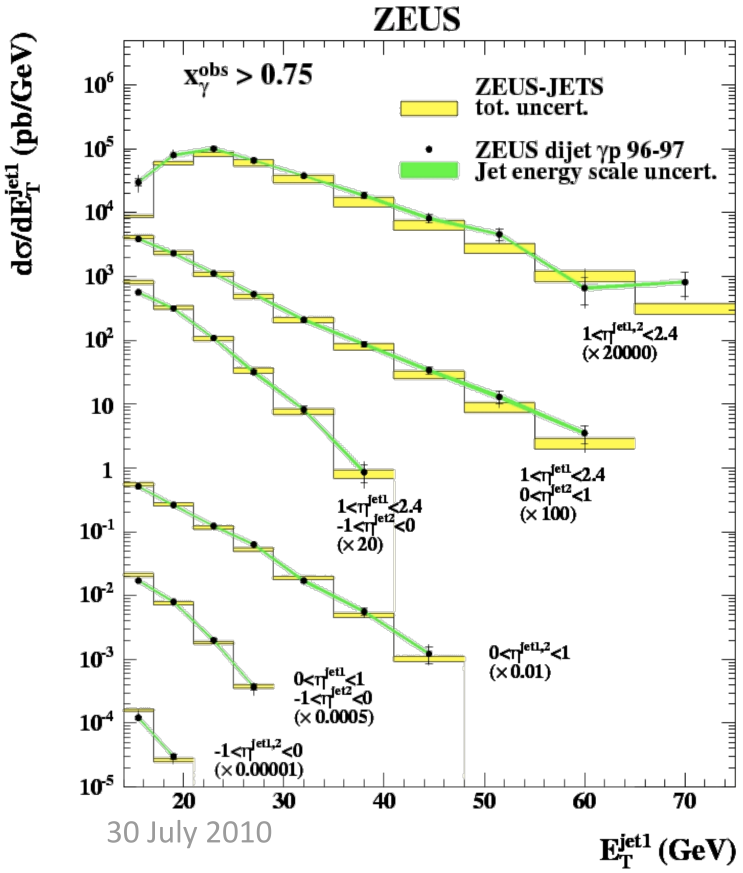
Infrared Safety?

- Adding an arbitrarily soft gluon to the event should not change the jets.
- Non-infrared-safe jet algorithms cannot be used in higher order calculations.
    - so if we use them in the experiment we cannot compare to the best theory (at least without making some intermediate model-dependent correction).
- Actually, infrared instabilities undermine the claim of a jet algorithm to be telling us about the short distance physics.
    - We should also worry about sensitivity to arbitrarily low noise, or arbitrarily soft pions, for example.
- An example of an infrared instability is the “seed” in old cone algorithm. But there are others (e.g. in the splitting/merging)

Precision Application

- ZEUS Jet measurements
- 1% energy scale, k_T algorithm
- Compared to NLO QCD, used in NLO PDF fits





Extrapolation: A. Cooper-Sarkar, C.Gwenlan, C.Targett-Adams, HERA-LHC Workshop, hep-ph/0509220

What is a Jet?

- Jets are not just less-well-measured leptons or “smeared” partons.
  - Hard radiation interference at amplitude level
  - Matching at high scales with Matrix element
  - Matching at low scales with parton densities and hadronisation model
  - potentially useful information in the internal jet structure, and in particle/energy flow between the jets
- Jets have **no existence independent of the algorithm**
  - even if the “algorithm” = event display + physicist

What is a Jet?

- So jet algorithms don't so much **find** a pre-existing jet as **define** one.
- A "jet" (or a pattern of jets) is a complex QCD event shape, designed to reflect as closely as possible the short distance degrees of freedom (quarks, gluons, H, Z, W…)
  - The degrees of freedom themselves are generally not physical observables, but can only be extracted within some theory or model
  - The cross section for quark production in the final state at LHC is *zero* (unless we find something very exciting…)