Jet Radiation Radius and Pileup

arXiv: 1309.xxxx
Zhenyu Han
University of Oregon

8/15/2013, BOOST2013 @ Flagstaff, AZ

Identifying boosted objects in a contaminated environment

- High multiplicity events (e.x., SUSY cascades)
- W’s in top decays (contaminated by b jets)
- Initial state radiation
- Underlying event
- Pileups

Outline

- Review of W jet tagging
- Difficulties with pileup
    - Jet radiation radius
    - Application in W jet tagging
- Outlook and conclusion

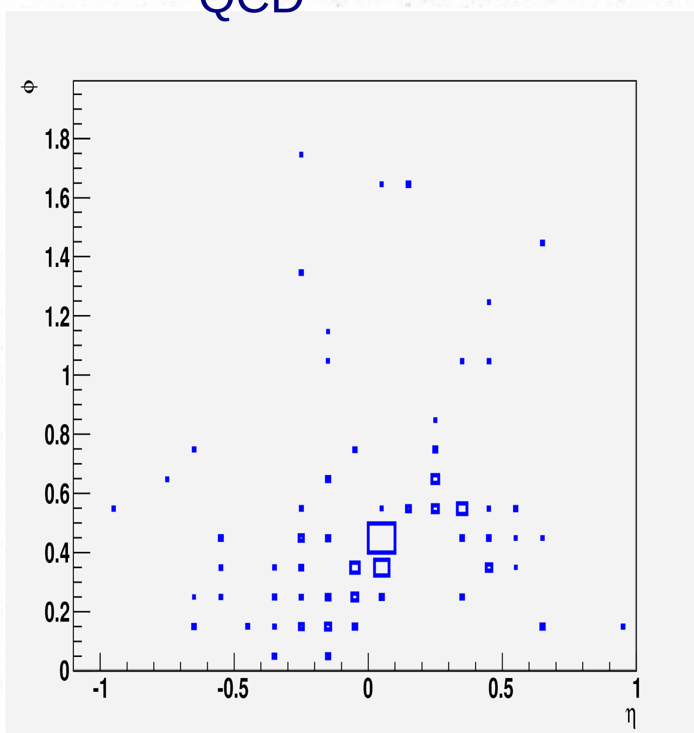
W jet tagging

- Differences between w jets and quark/gluon jets
  - W mass peak vs QCD continuum
  - 2 balanced hard subjets vs hierarchical momenta
  - Color singlet vs triplet/octet: different radiation patterns
    - W: radiation concentrated in a small cone
    - QCD: radiation diffused.

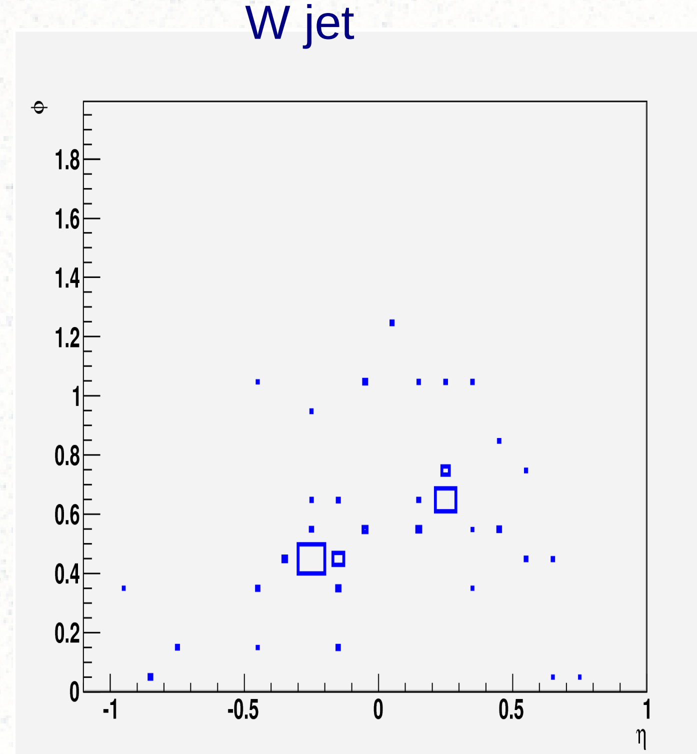
The three differences are (almost) uncorrelated.
Use all of them!

W jets vs QCD jets

QCD



W jet



Group the energy in 0.1x0.1 bins on (eta, phi) plane, jets found using R=1.2, C/A

QCD jet from W+j->lvj, W-jet from WW->lvjj, Madgraph+Pythia 8

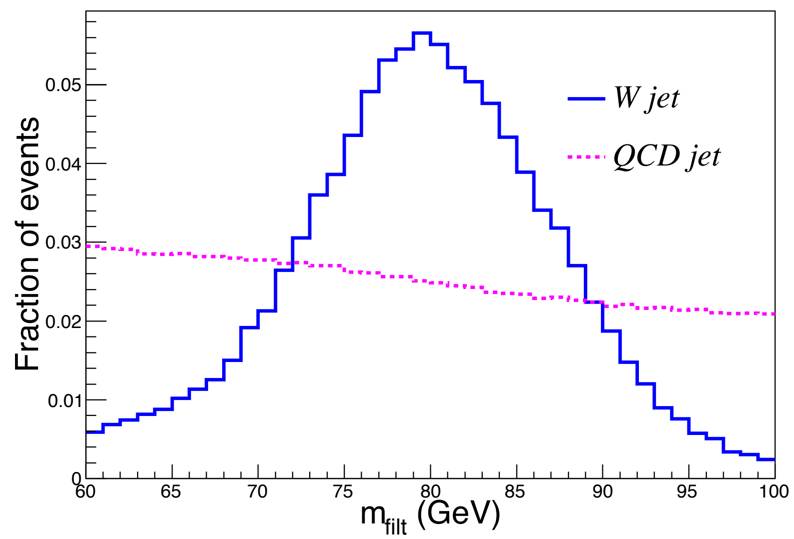
Jet grooming algorithms

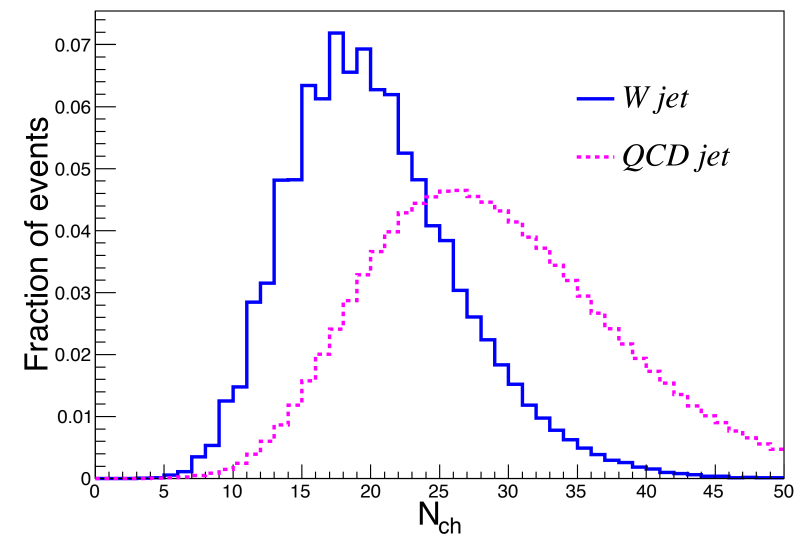
- Eliminate soft radiations, identify the hard subjets
- Mass peak well reconstructed
- Radiation information lost

Jet shape/radiation variables

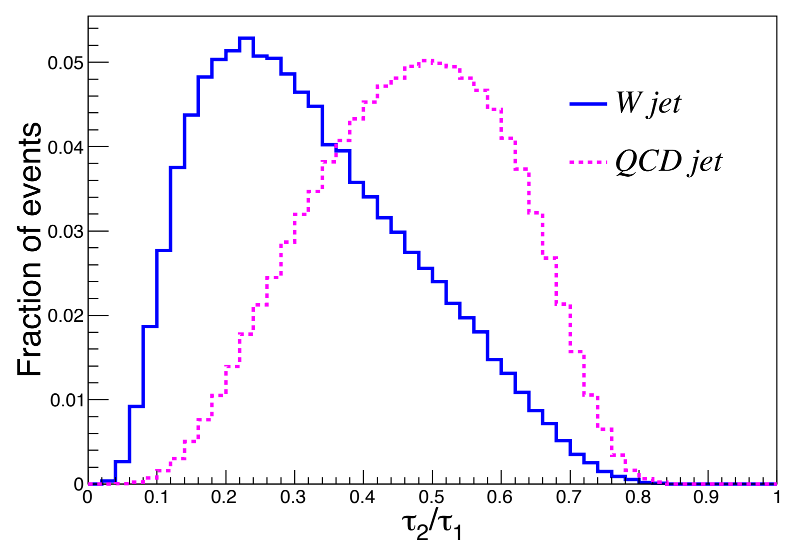
- Variables sensitive to radiation.
  - N-subjettiness
  - Charged particle multiplicity
  - R-cores: measure how mass/pt change according to the jet radius
- Combine with jet grooming methods

Jet shape variables
(without pileups)





Han, 2011

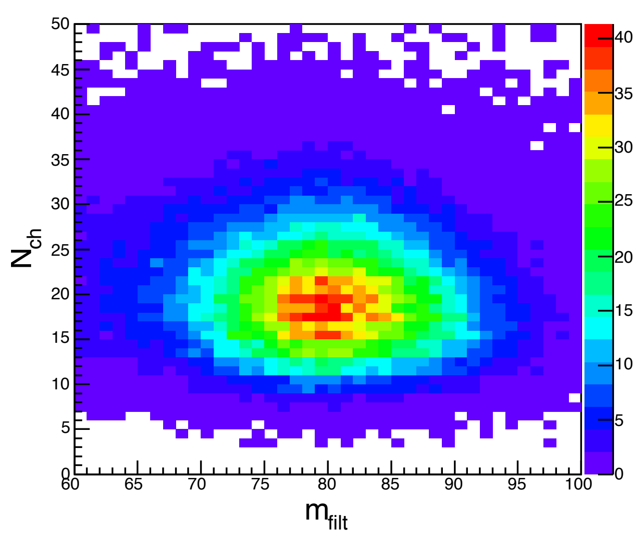


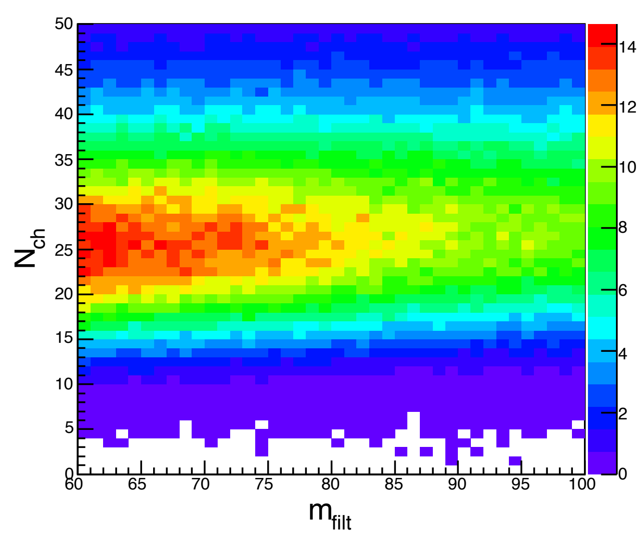
PT=500GeV. After an overall cut on jet mass after filtering/mass drop. Nch and τ_{21} calculated before filtering. mfilt calculated from calorimeter cells (0.1x0.1 binning), τ_{21} from tracks alone.

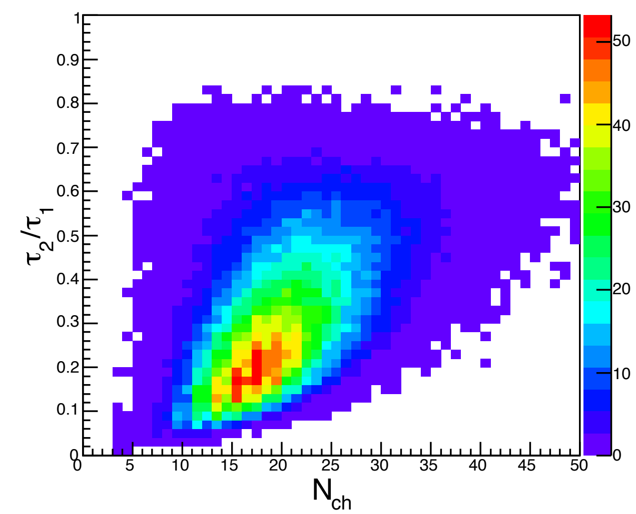
Correlations

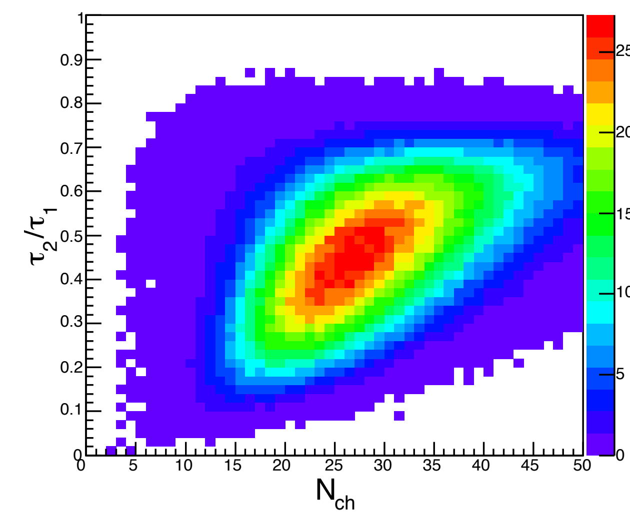
W

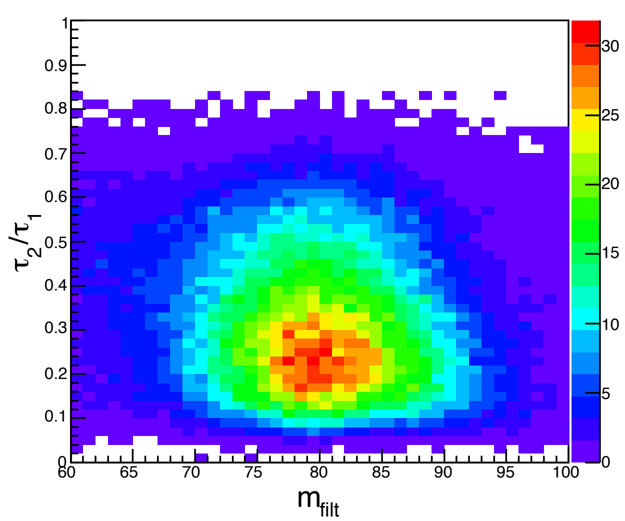
QCD

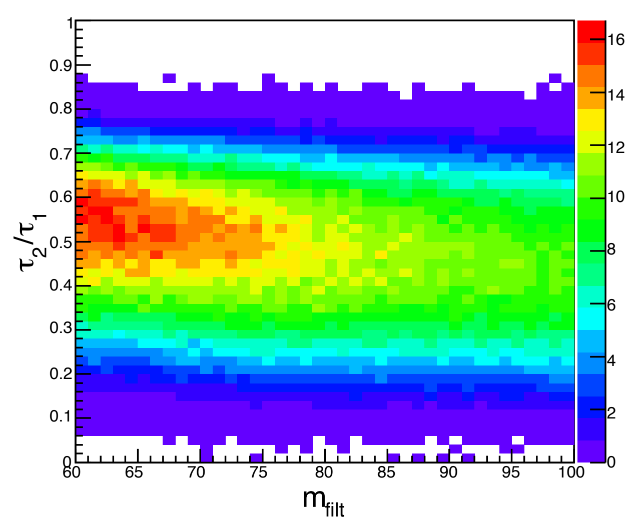












correlation coefficients

W

	$m_{\rm filt}$	$N_{\rm ch}$	τ_2/τ_1
$m_{\rm filt}$	1	-0.08	-0.12
$N_{\rm ch}$	-0.08	1	0.51
τ_2/τ_1	-0.12	0.51	1

QCD

	$m_{\rm filt}$	$N_{\rm ch}$	τ_2/τ_1
$m_{\rm filt}$	1	0.07	-0.14
$N_{\rm ch}$	0.07	1	0.50
τ_2/τ_1	-0.14	0.50	1

A few comments on N-subjettiness

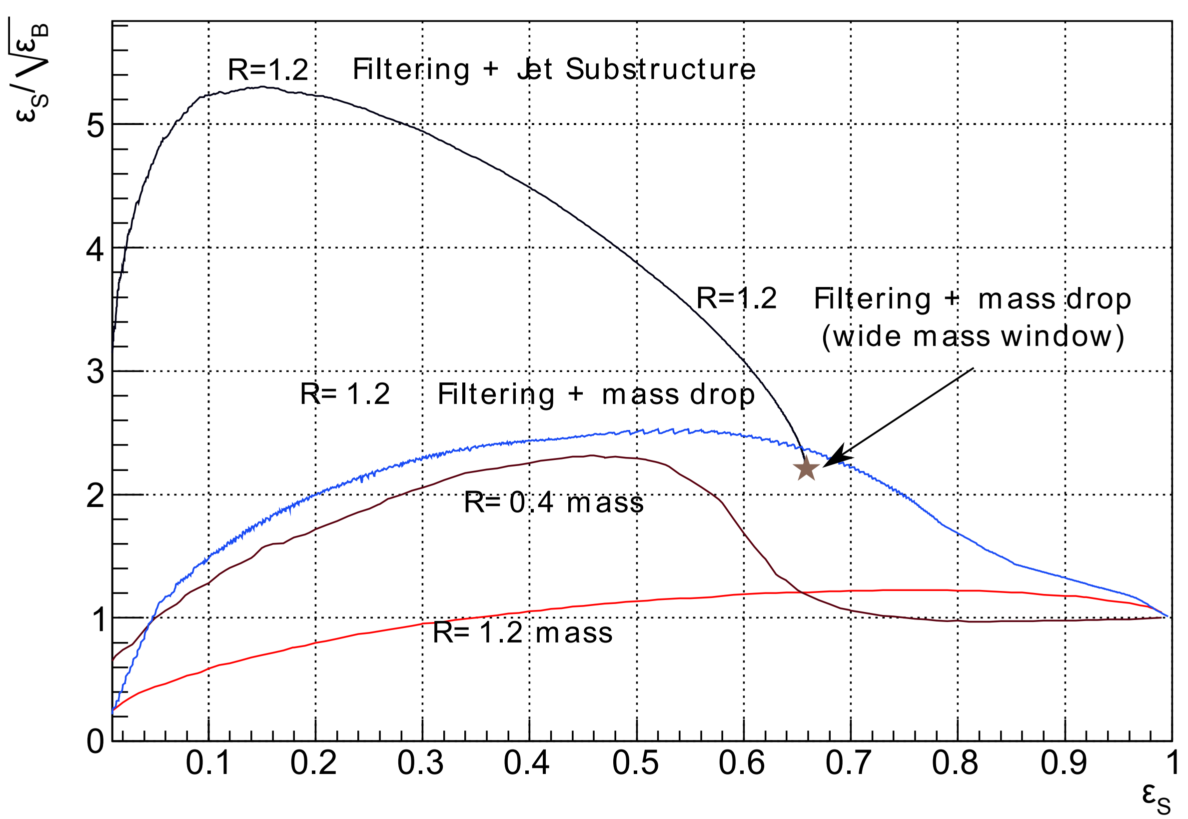
- τ_2/τ_1 (τ_{21}) relevant to W tagging. Two ways to get a large τ_1:
  - 2 (or more) hard subjets (correlated to jet grooming)
  - Significant amount of diffused radiation (small correlation to the hard splitting scale)
    - Similar to (fat) jet mass
- "Best way" to use N-subjettiness:
  - use jet grooming to 'factorize out' the hard splitting,
  - After grooming, τ_{21} becomes a variable measuring the amount of radiation.

Combined performance
(without pileup)

- Combine variables sensitive to the hard splitting (mass after filtering/mass drop) and those sensitive to the color structure.

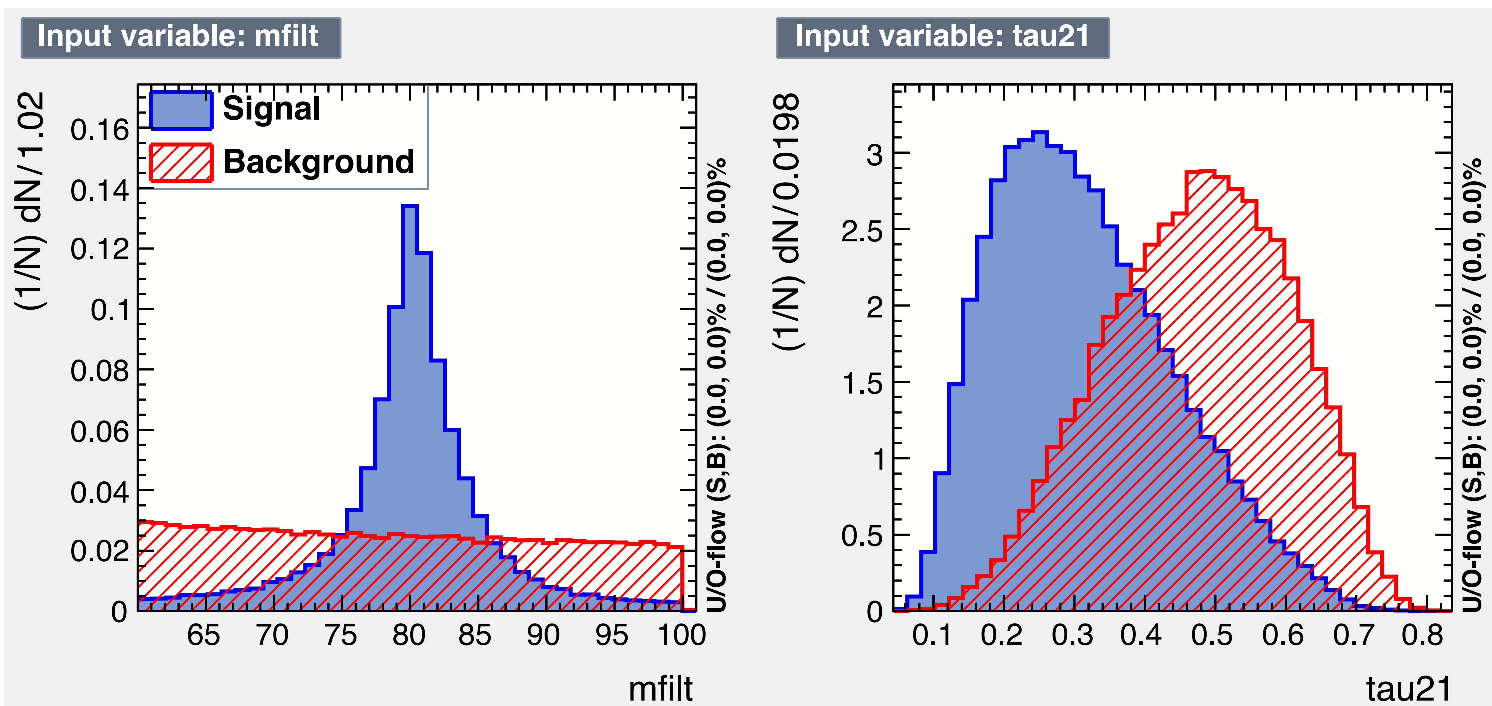
- Combining filtering and τ_2/τ_1(Nch) improves the significance (S/sqrt(B)) by a factor of 1.5(1.4), over filtering alone; Combining all three gives a factor of 1.6 (*Han, 2011*).

- 'Ultimate' performance with MVA:

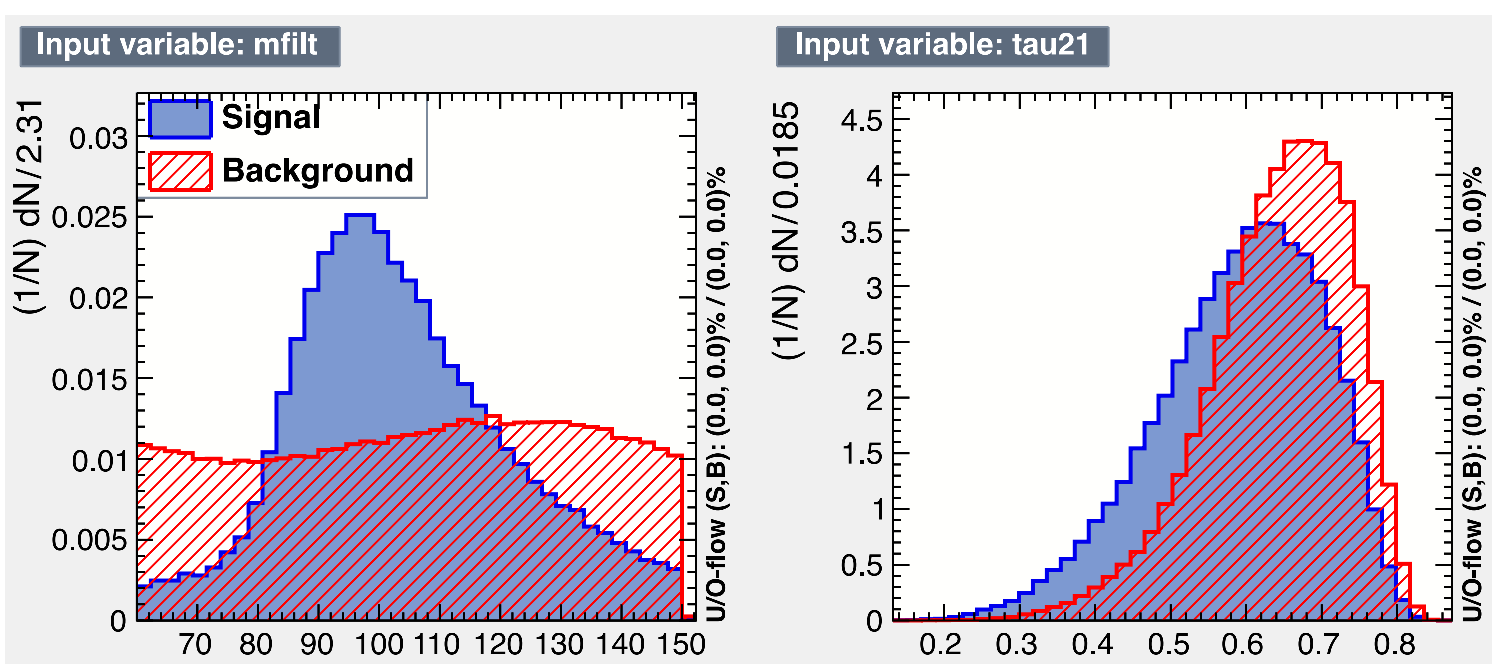


Cui, Han and Schwartz, 2010

Difficulties from pileup

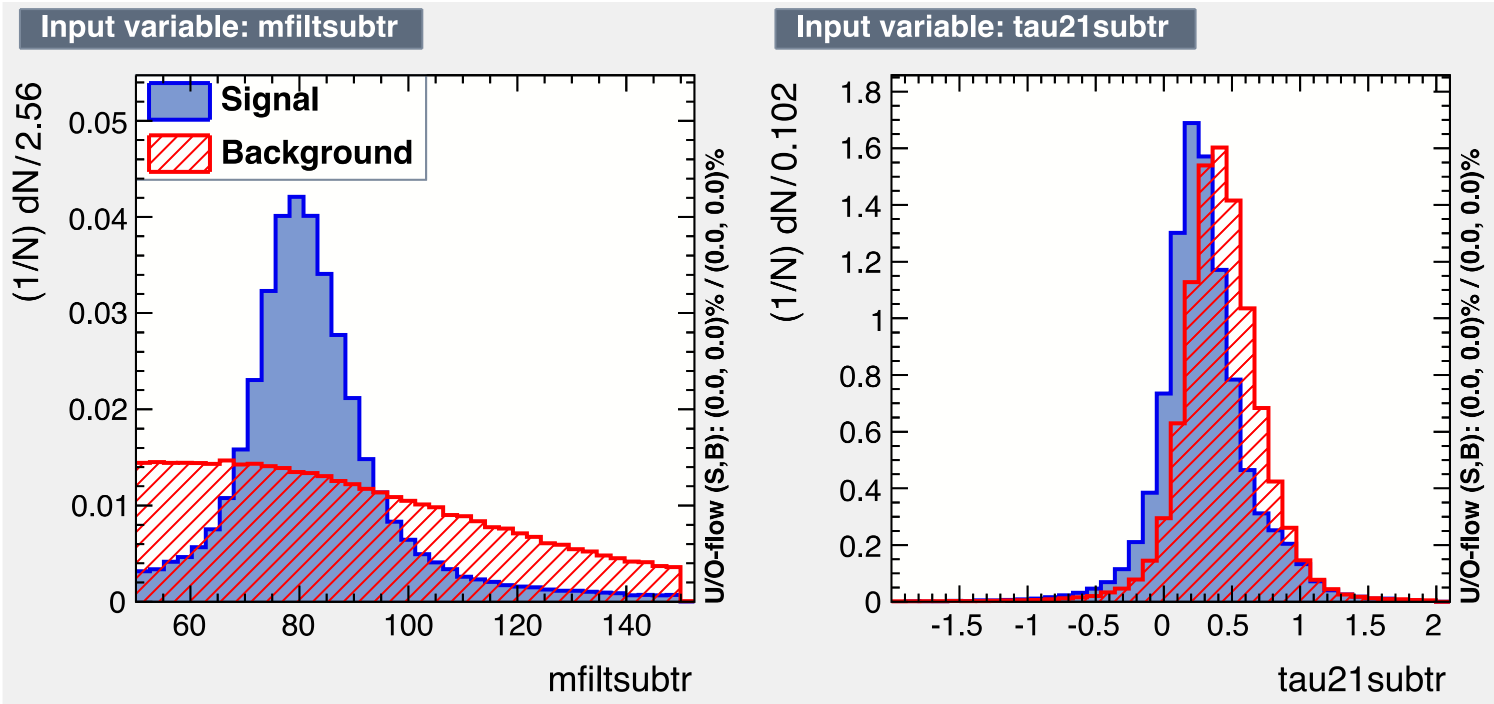


No Pileup, R=1.2, >300GeV jets
Pythia 8: WW, QCD dijet
particle level
Best improvement in S/sqrt(B)
filtering: 2.7
filtering+τ_{21}: 4.8



60 Pileup events (Pythia 8 SoftQCD),
R=1.2 >300GeV jets
Best improvement in S/sqrt(B)
filtering: 1.47
filtering+τ_{21}: 1.54

Pileup Subtraction (Soyes, Salam, Kim, Dutta, Cacciari, 2012)



ρ & ρ_m subtracted

Best improvement in S/sqrt(B):

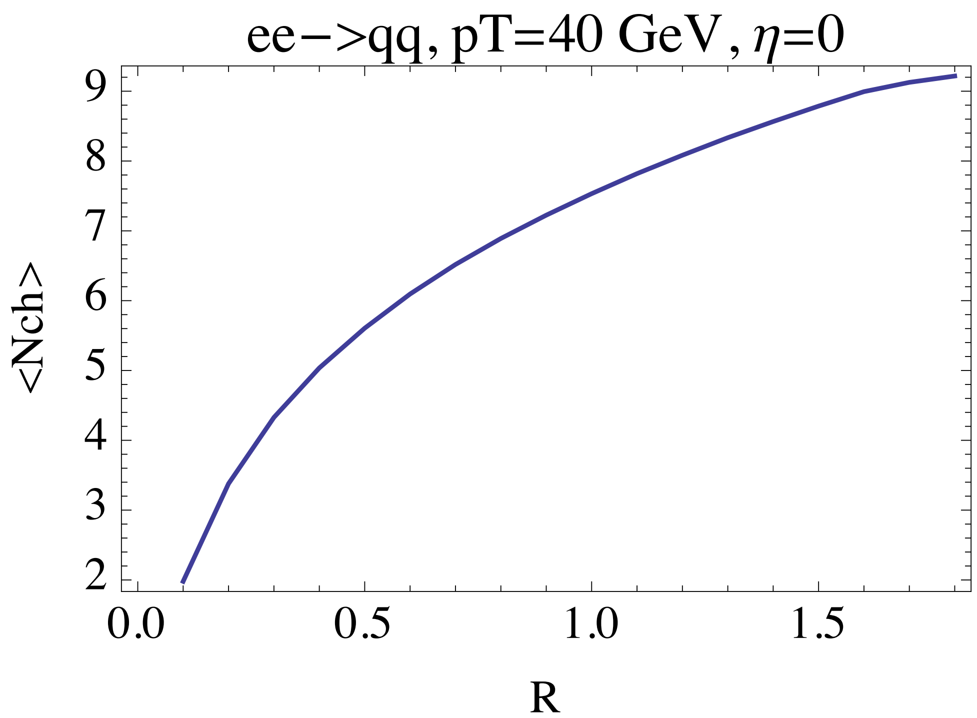
filtering(subtr):1.61

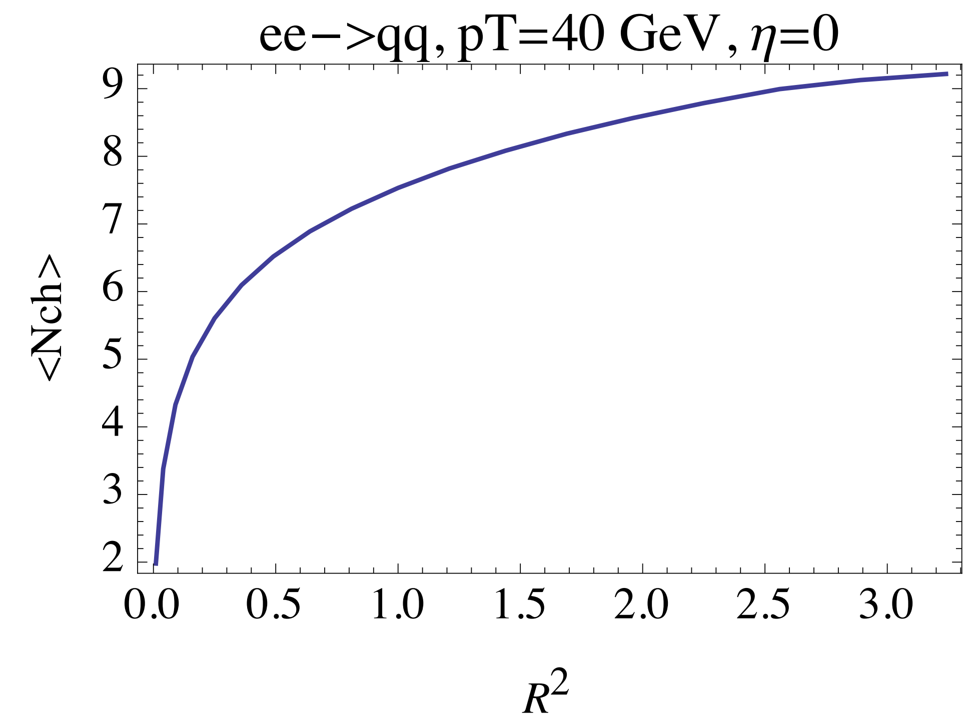
filtering(subtr)+τ_{21}(subtr):1.8

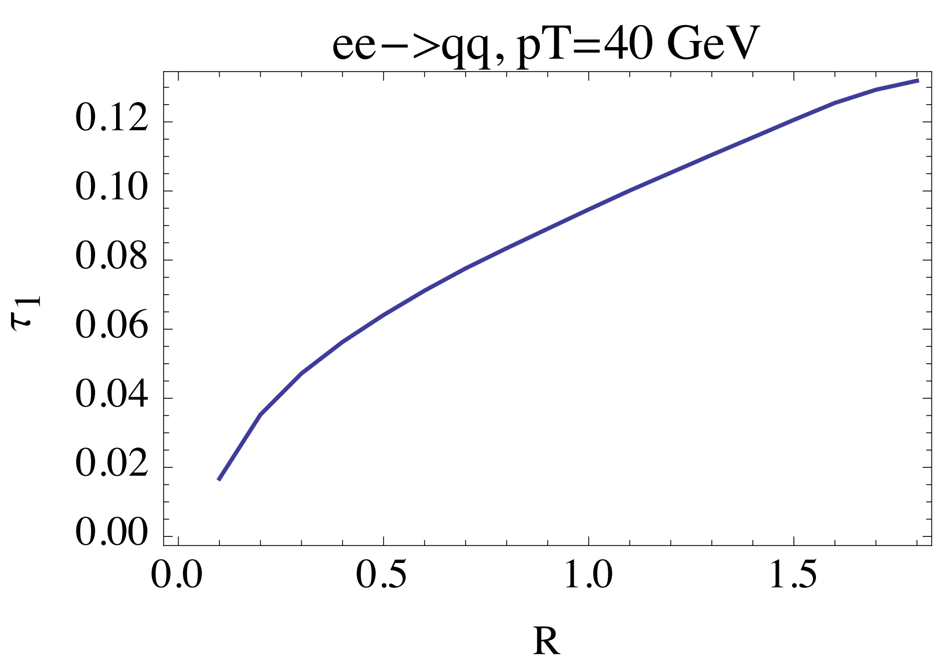
Do we need R=1.2 to calculate τ_{21}?

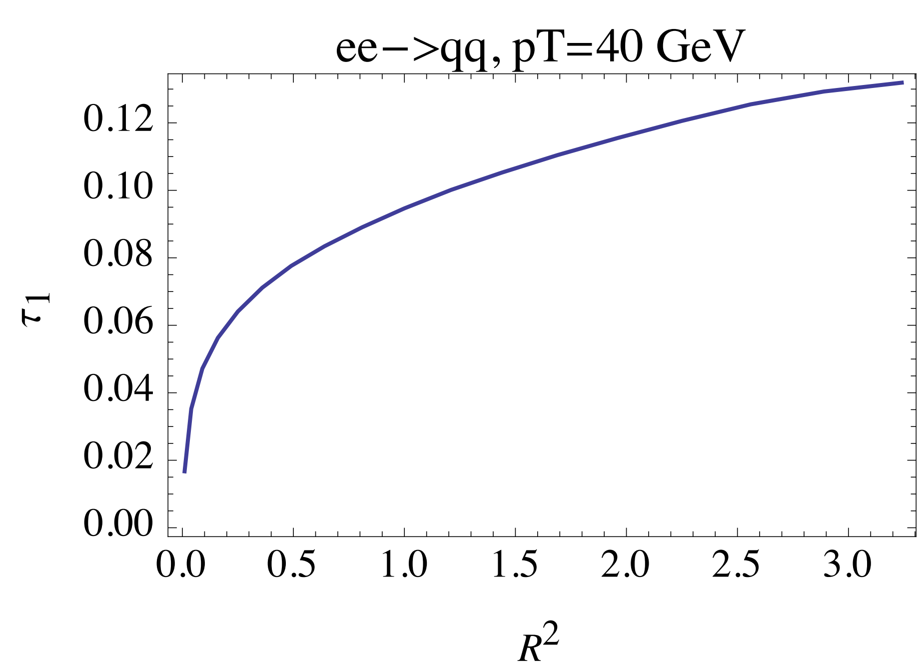
Dijets, PT=40GeV (W)

Pythia 8 showering and hadronization,
clustered with anti-kt

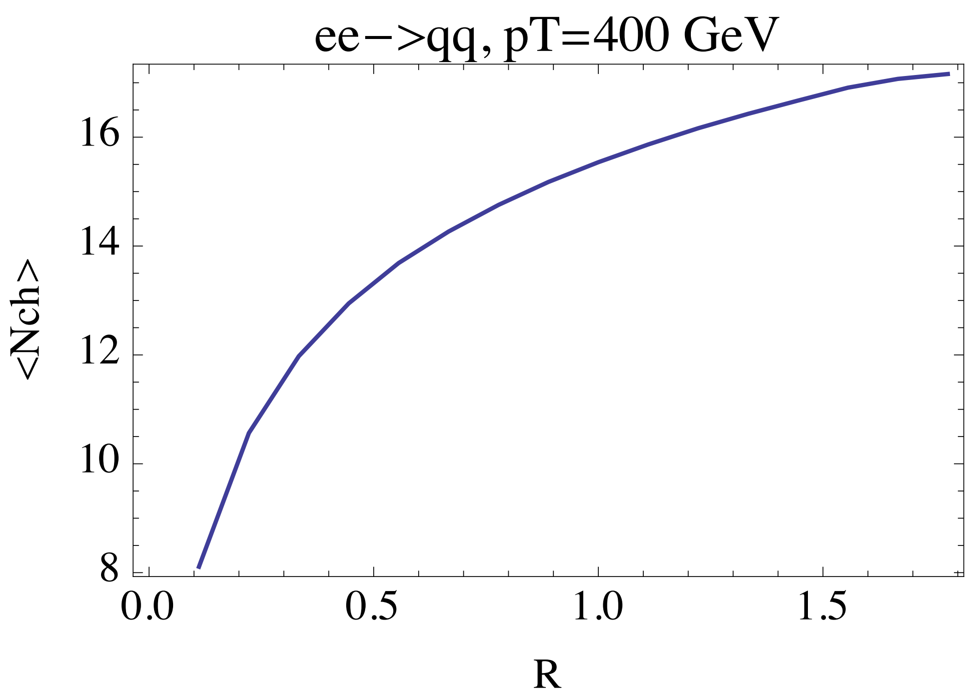


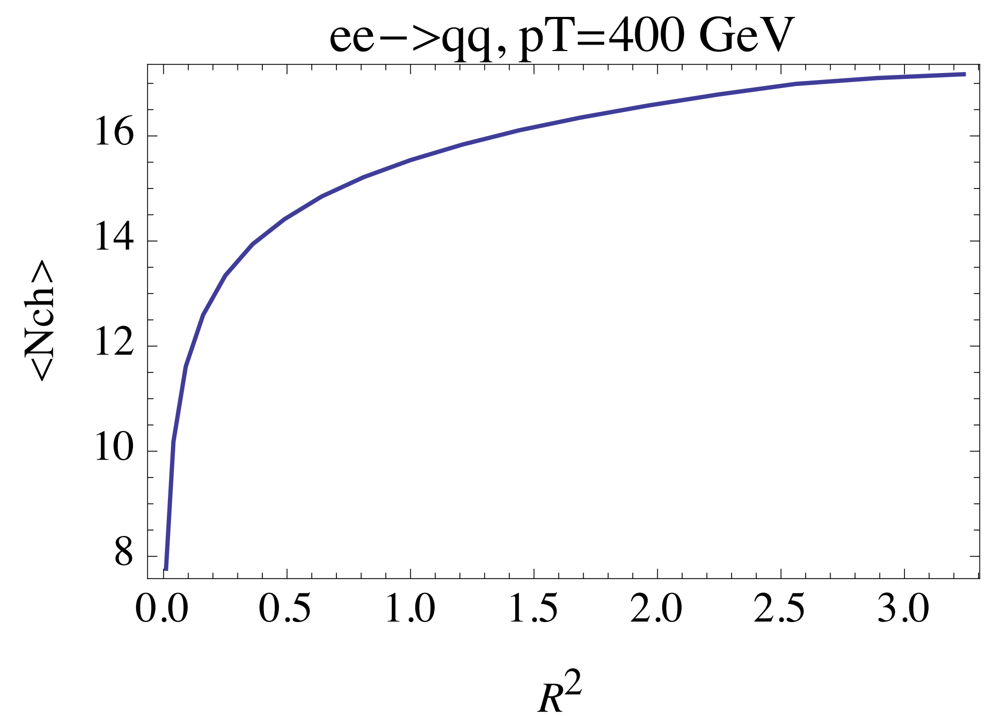


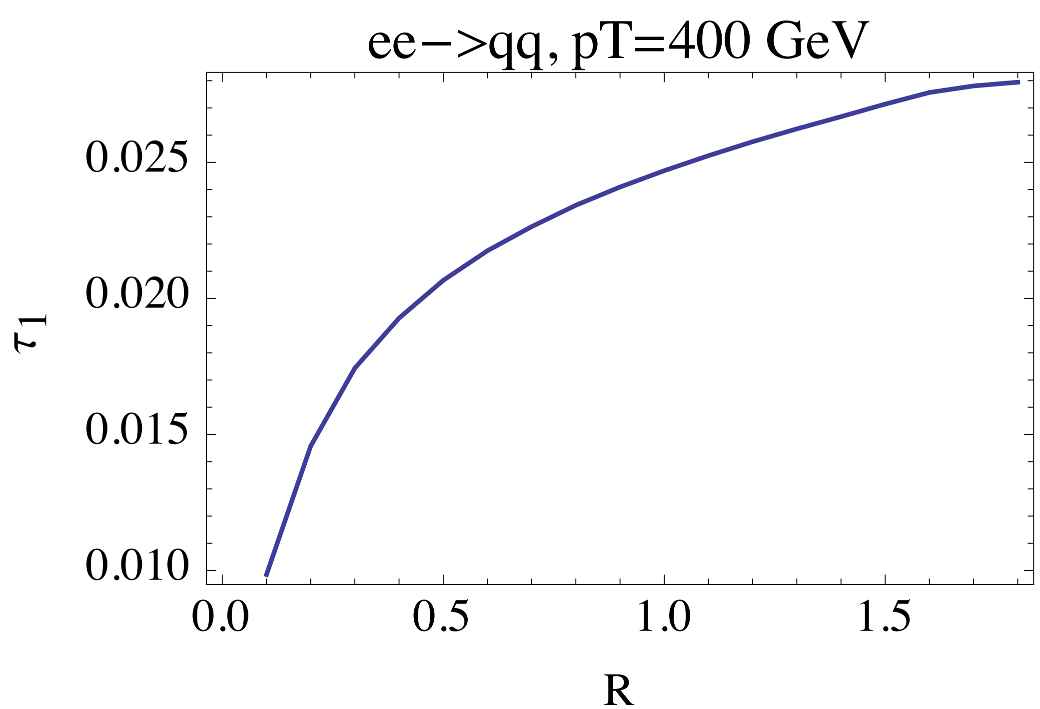


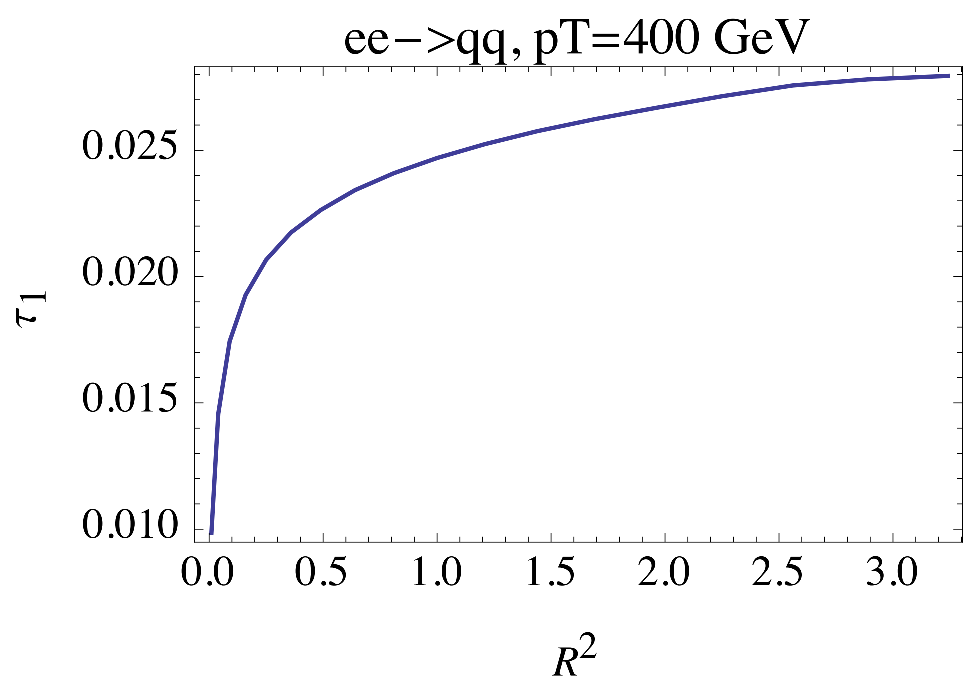


Dijets, PT=400GeV

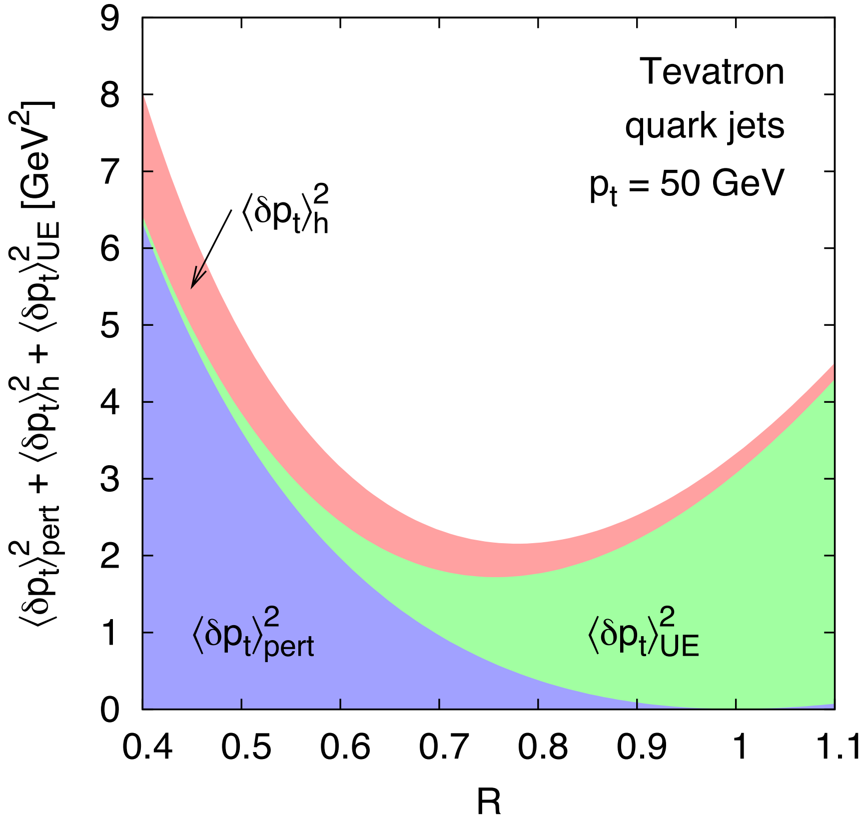




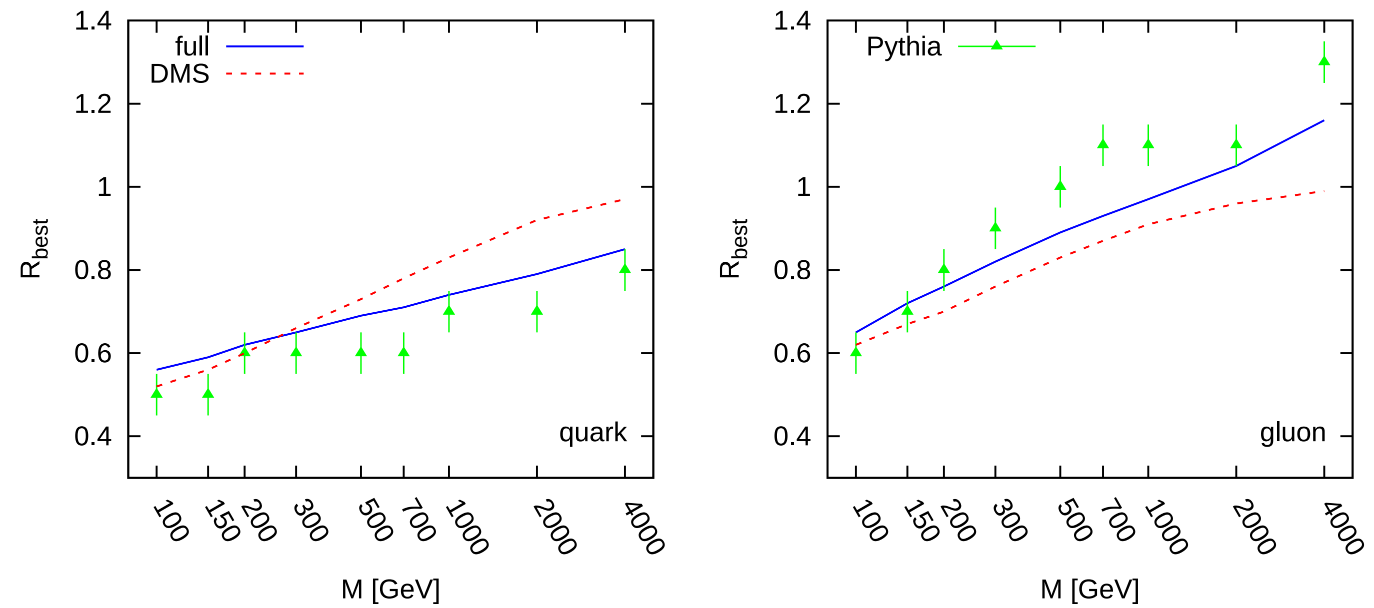




Related works



Dasgupta, Magnea, Salam, 2007



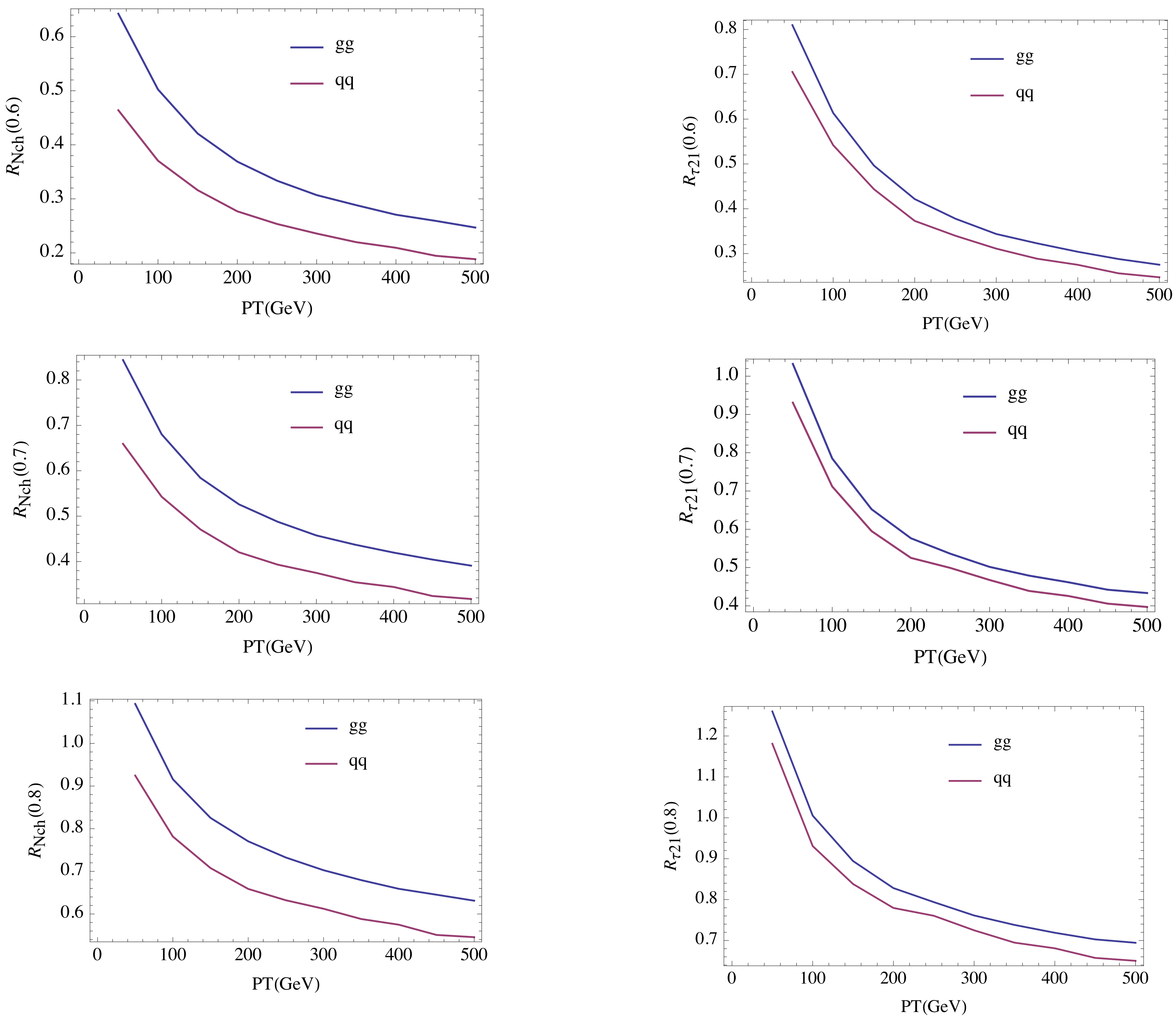
Soyez, 2010

- Best radius for discrimination?
- Radiation variables: τ_{21}, Nch...?
- Effect from pileup?
- Boosted case?

Jet radiation radius

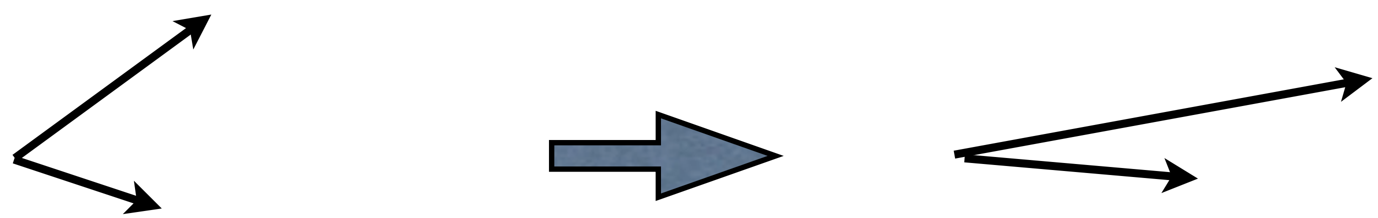
- What's the 'intrinsic size' of a jet?
- In a color singlet system in its center of mass frame, R(*x*) is defined as the minimum jet radius that averagely *x* percent of the total 'amount of radiation' is contained in the leading 2 (n) jets.
    - The definition of 'amount of radiation' depends on the variable used.
    - Alternative definition: the minimum jet radius that x percent of the events have x percent of the radiation contained in the leading jets (similar to the h-index definition).

PT dependence
ee->qq, gg



Boosted W

- Shrinking cone size R ~ 1/PT :



$$p_1 \cdot p_2 = p_{10} \cdot p_{20}$$

$$E_1 E_2 - E_1 E_2 \cos\theta = E_{10} E_{20} - E_{10} E_{20} \cos\theta_0$$

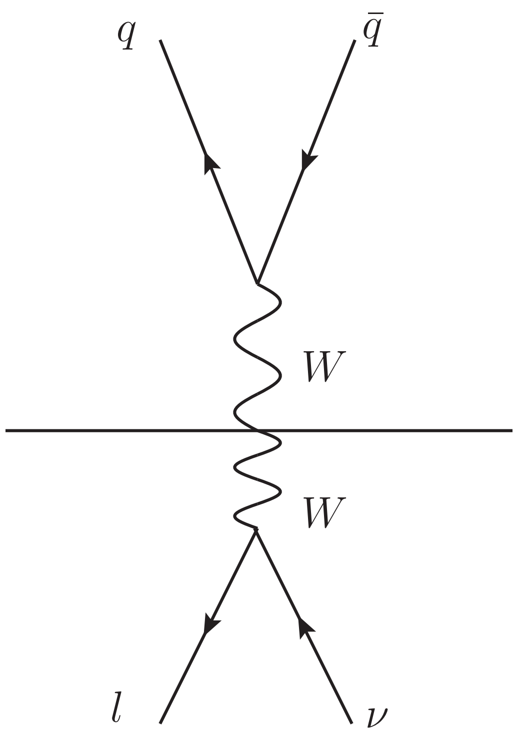
$$\text{boost:} \quad E_1/E_{10} \approx E_2/E_{20}$$

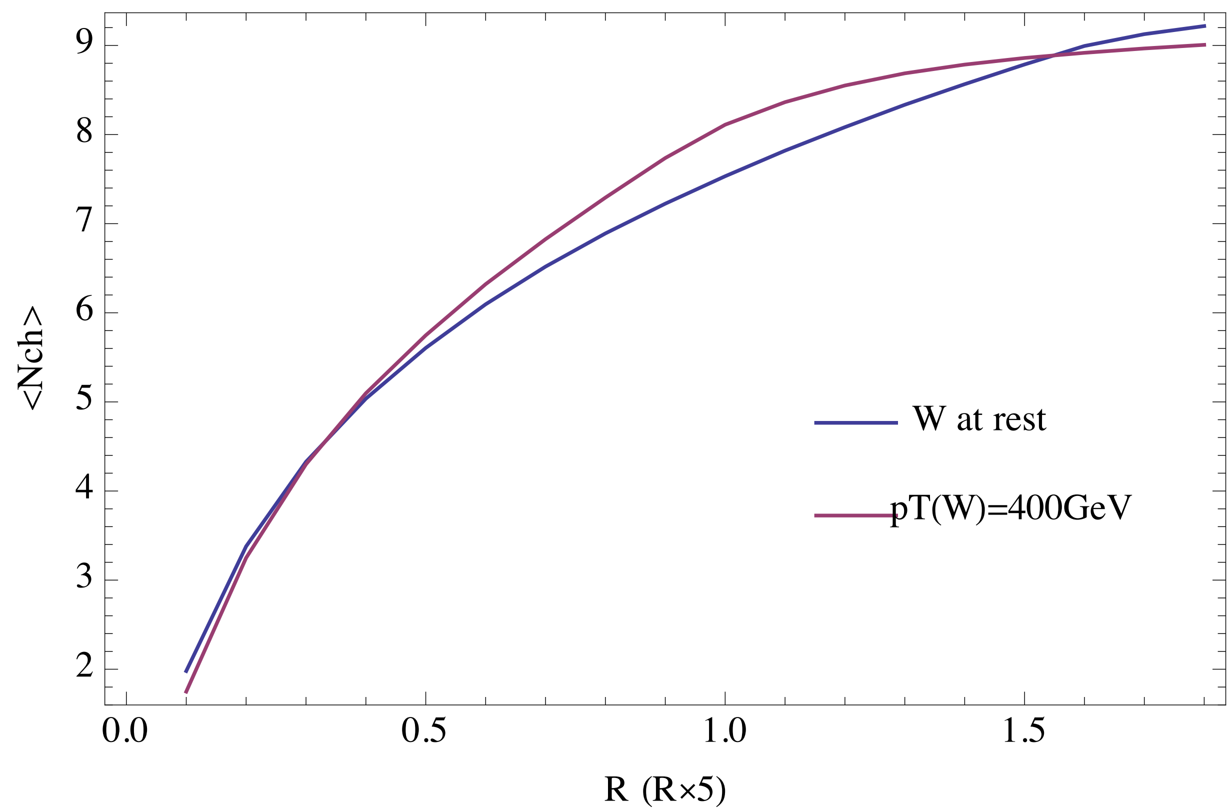
$$\Rightarrow \theta/\theta_0 \approx E_{10}/E_1$$

When jets merged, jet radius becomes subjet radius

Boosted W

ee->WW

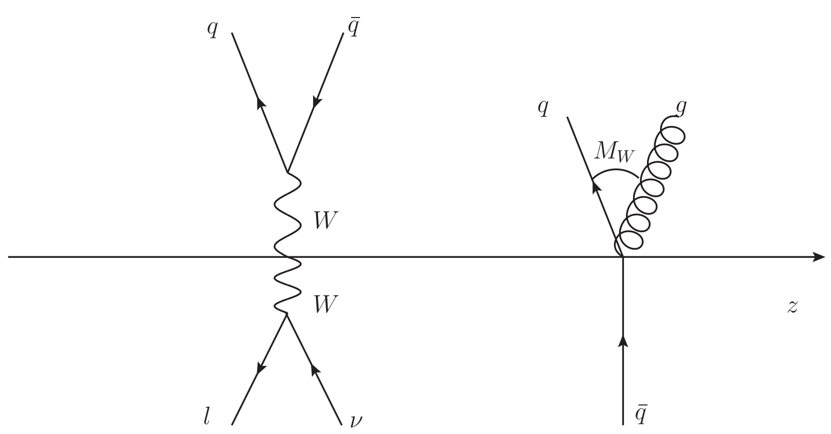


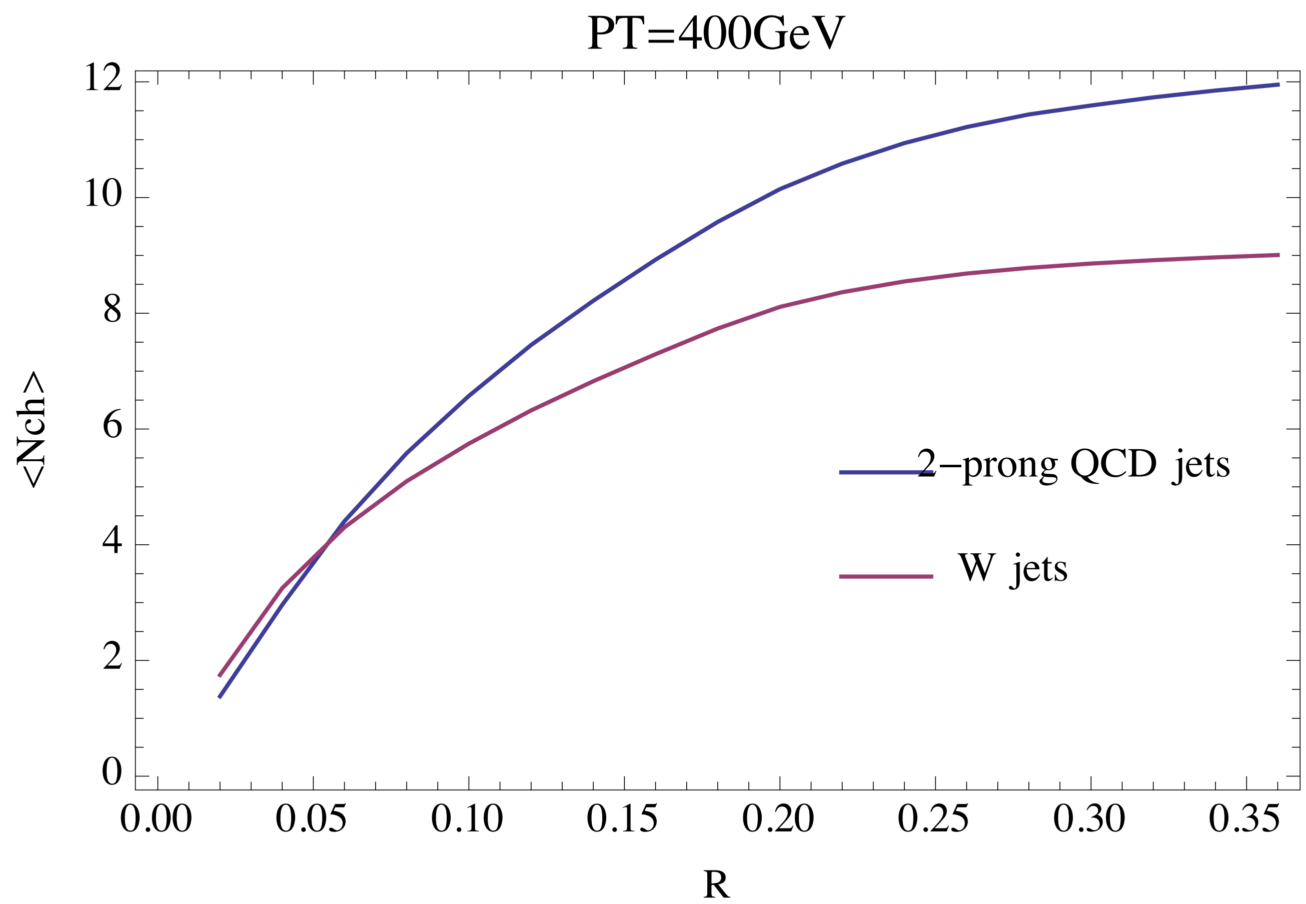


40 GeV *vs* 200 GeV quarks from W decay

Boosted W vs QCD jet

ee->ww, qqg





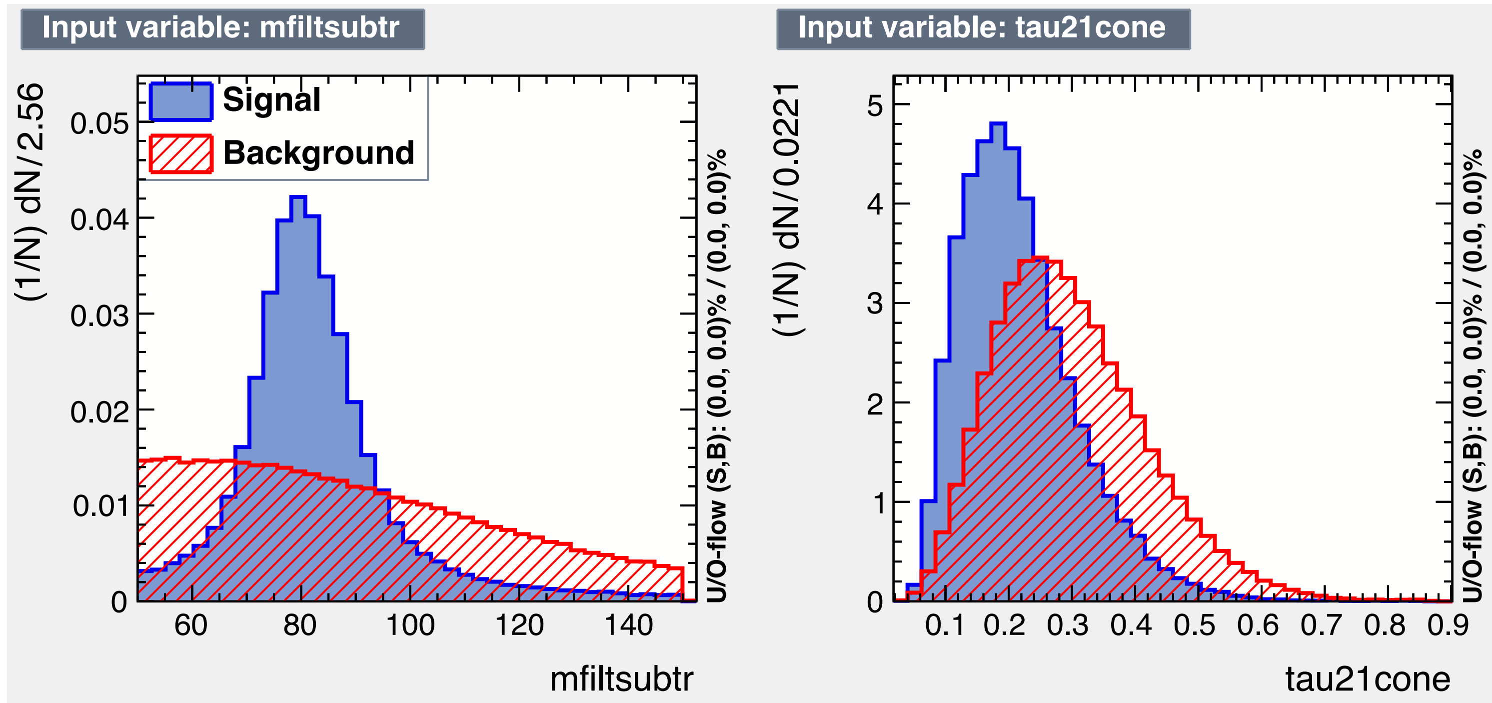
Discrimination procedure

- Cluster to fat, high pt jets
- Calculate subtracted groomed jet mass
- Apply a precut on the groomed jet mass - optional
- Find the axes of the two leading subjets, calculate τ_{21} for jet constituents with a cone around the two axes, (shrinking) cone size determined by

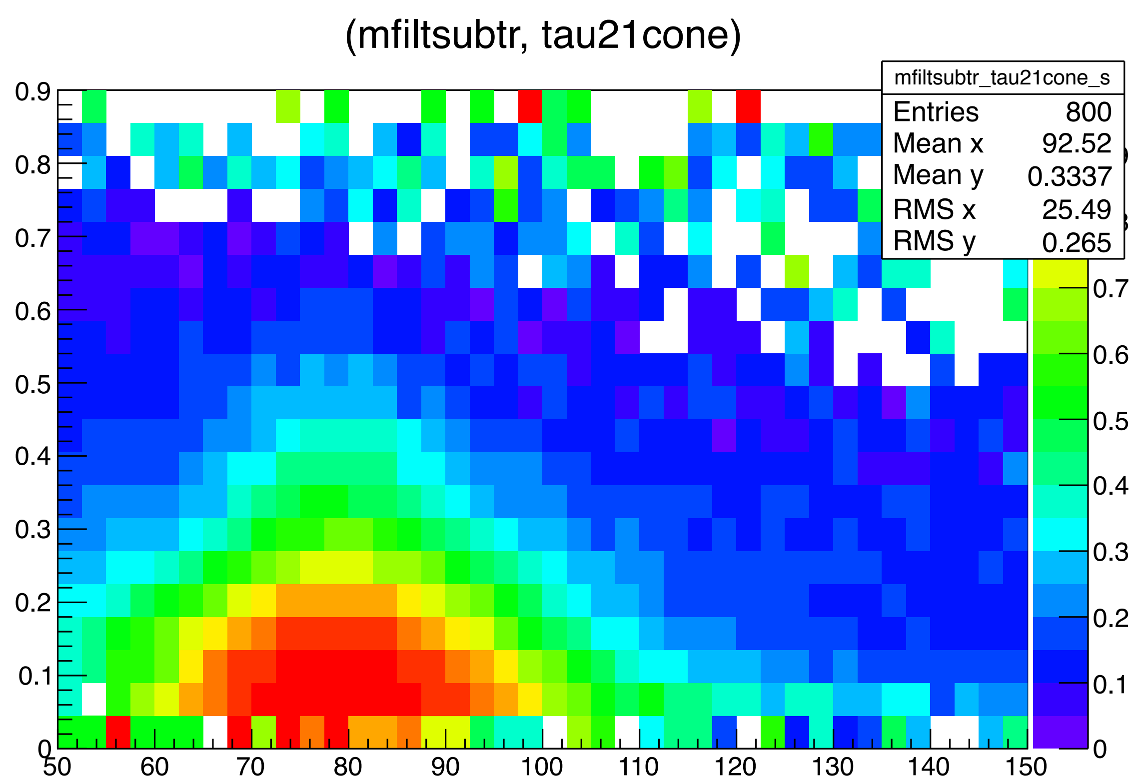
$$R_{\mathrm{sub}} = R_{\mathrm{ref}} \frac{100\mathrm{GeV}}{p_{T,\mathrm{sub}}}$$

- 2D analysis using ($m_{\mathrm{filt,subtr}}$, $\tau_{21,\mathrm{sc}}$).

Performance at the LHC

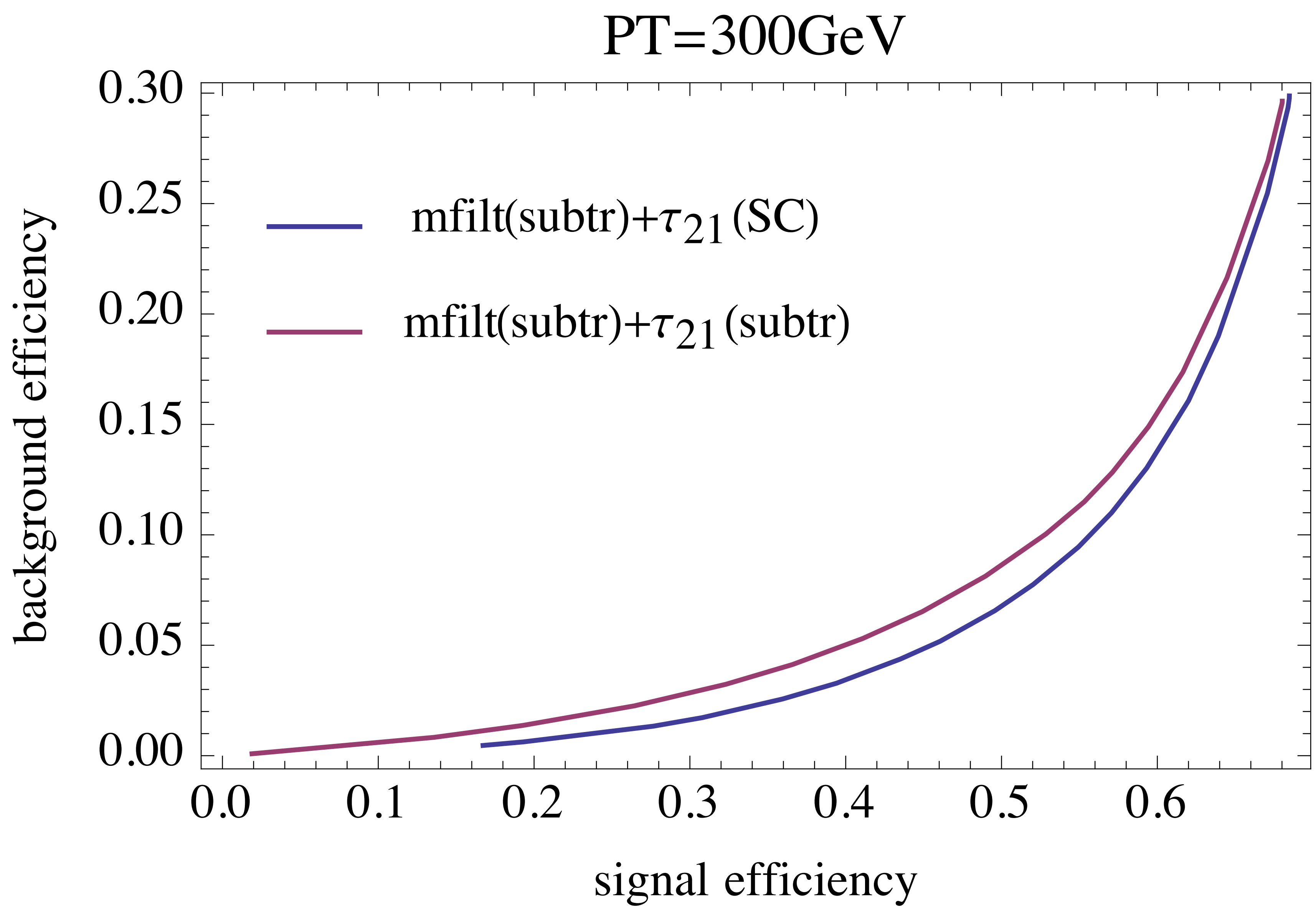


likelihood = S/(S+B)



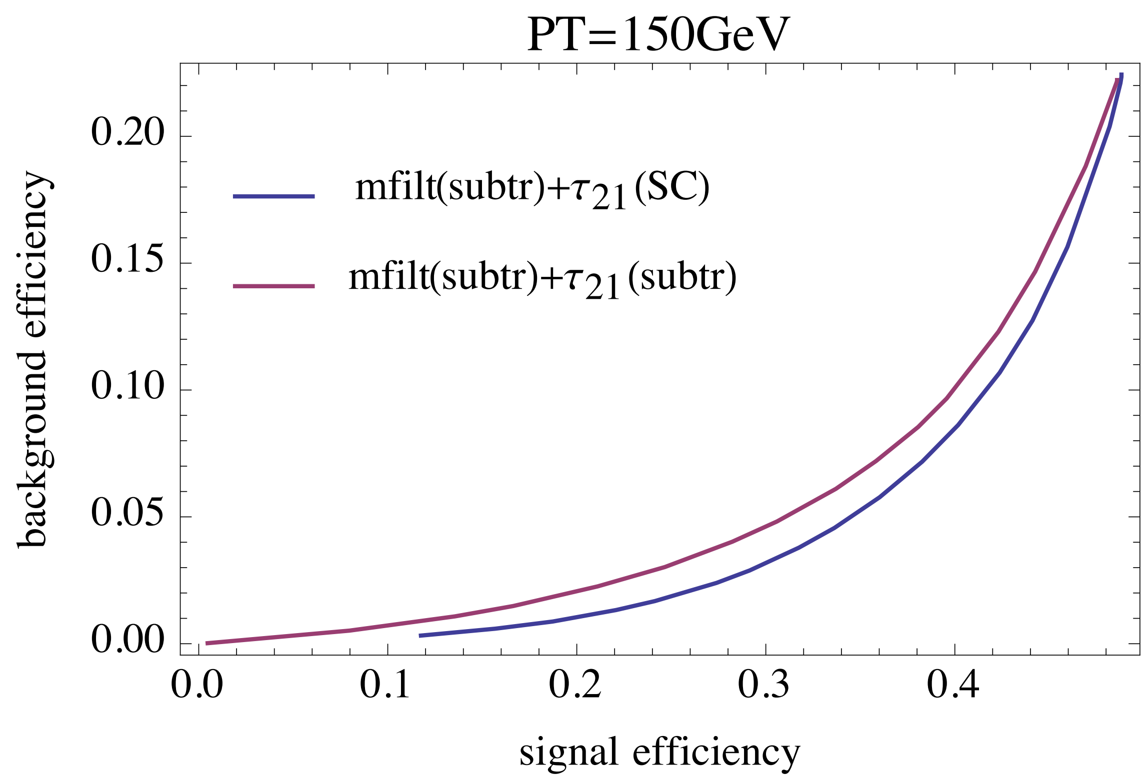
Pythia 8, 60 pileups
WW & dijet
Rfat=1.2, PT>300GeV
Rref=0.3 (at 100GeV)

Performance



Best improvement in S/sqrt(B): 2.39
at eff(S)=0.28, eff(B)=0.013
compare: filtering(subtr) alone: 1.61
filtering(subtr) +τ_{21}(subtr): 1.80

Performance



Best significance 1.77,
filtering(subtr) alone: 1.30
filtering(subtr)+τ21(subtr):1.40

- Suggestion to experimentalists: increase jet radius for the fat jet (to cluster more W's), but reduce the radius for calculating radiation variables.

Future directions

- More sophisticated R_{sub} choices? Other variables? Combine with charged hadron subtraction?
- Other applications
  - Higgs search
  - Top tagging (even useful when using tracks alone)
  - Quark-gluon discrimination
  - Processes with many final state partons
- Theoretical calculations, Monte Carlo validations

Conclusion

- Radiation of a hard parton is concentrated, which can be quantified by jet radiation radius.
- The jet radiation radius is smaller for larger boost
- By selecting a small cone size dependent on the (sub)jet momentum to calculate radiation variables, we can reduce the impact from pileups and improve the tagging efficiencies for boosted objects.