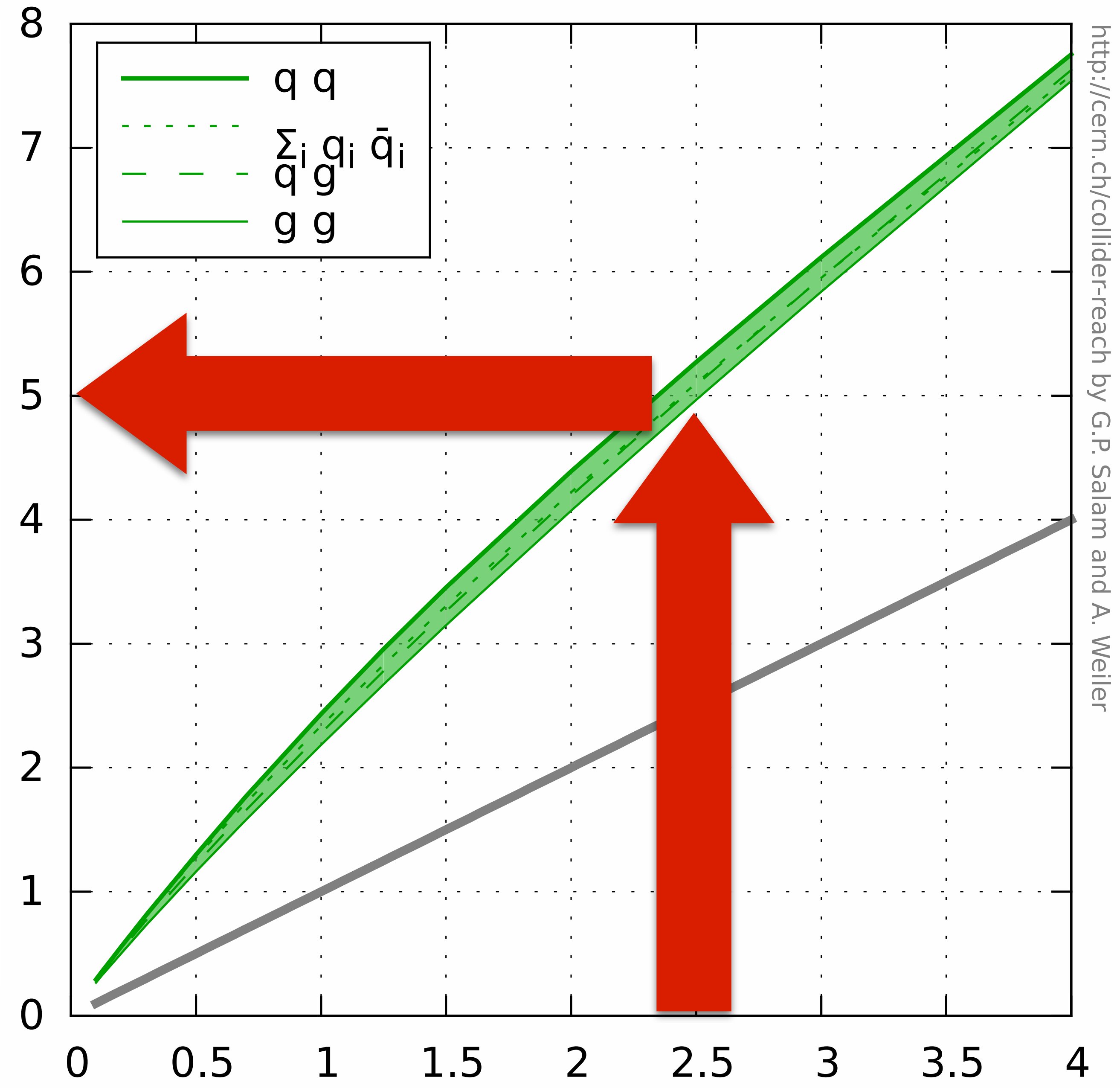
A theorist's perspective on boosted boson tagging

Gavin Salam (CERN-TH)

Boosted Boson Workshop,
CERN, March 25, 2014

14 TeV, after a few years, will roughly double the mass reach of LHC searches



http://cern.ch/collider-reach by G.P. Salam and A. Weiler

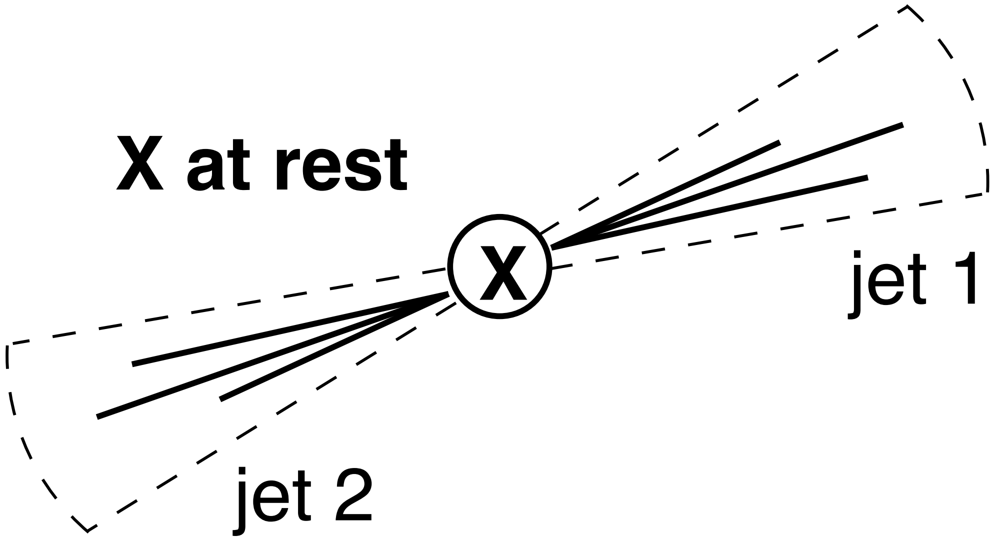
Mass reach [TeV] at 14 TeV, 300 fb^{-1}

Mass reach [TeV] at LHC 8 TeV

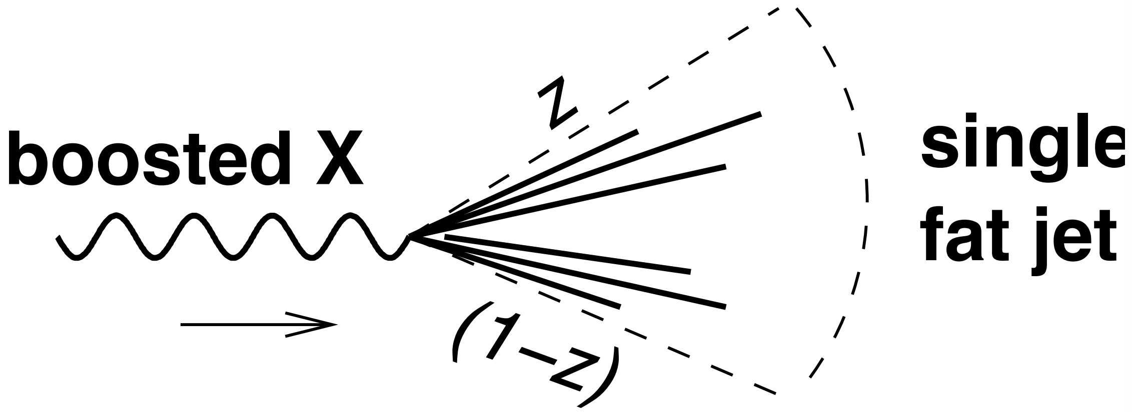
Boosted hadronic decays

(X = W, Z, H, top, new particle)

Normal analyses: two quarks from $X \to q\bar{q}$ reconstructed as two jets



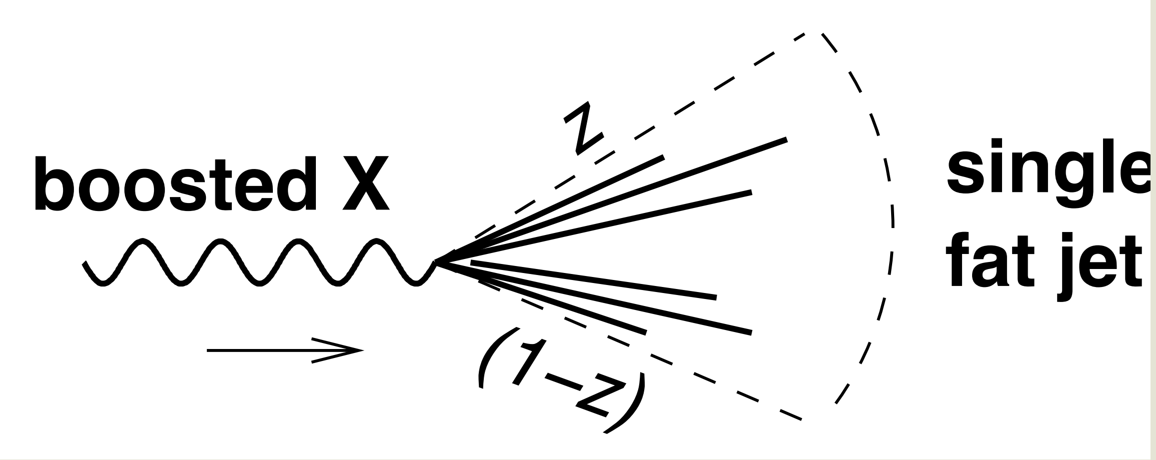
High-p_t regime: EW object X is boosted, decay is collimated, $q\bar{q}$ both in same jet



Happens for $p_t \gtrsim 2m/R$

$p_t > 400$ GeV for $m = m_W$, $R = 0.4$

Most obvious way of detecting a boosted decay is through the mass of the jet

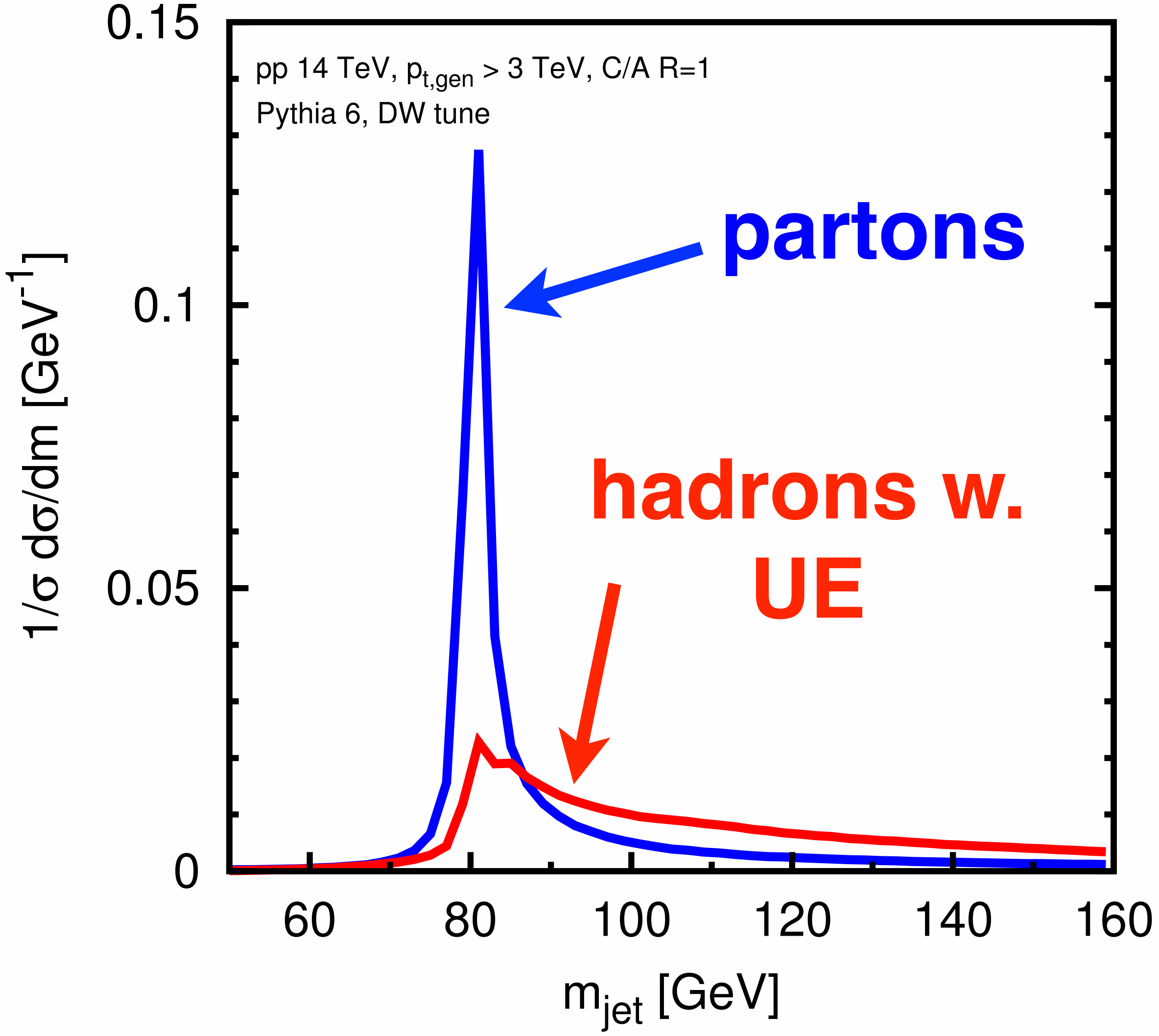


But jet mass is **poor** in practice:

e.g., narrow W resonance highly smeared by QCD radiation

(mainly underlying event/ pileup)

jet mass distribution from W bosons

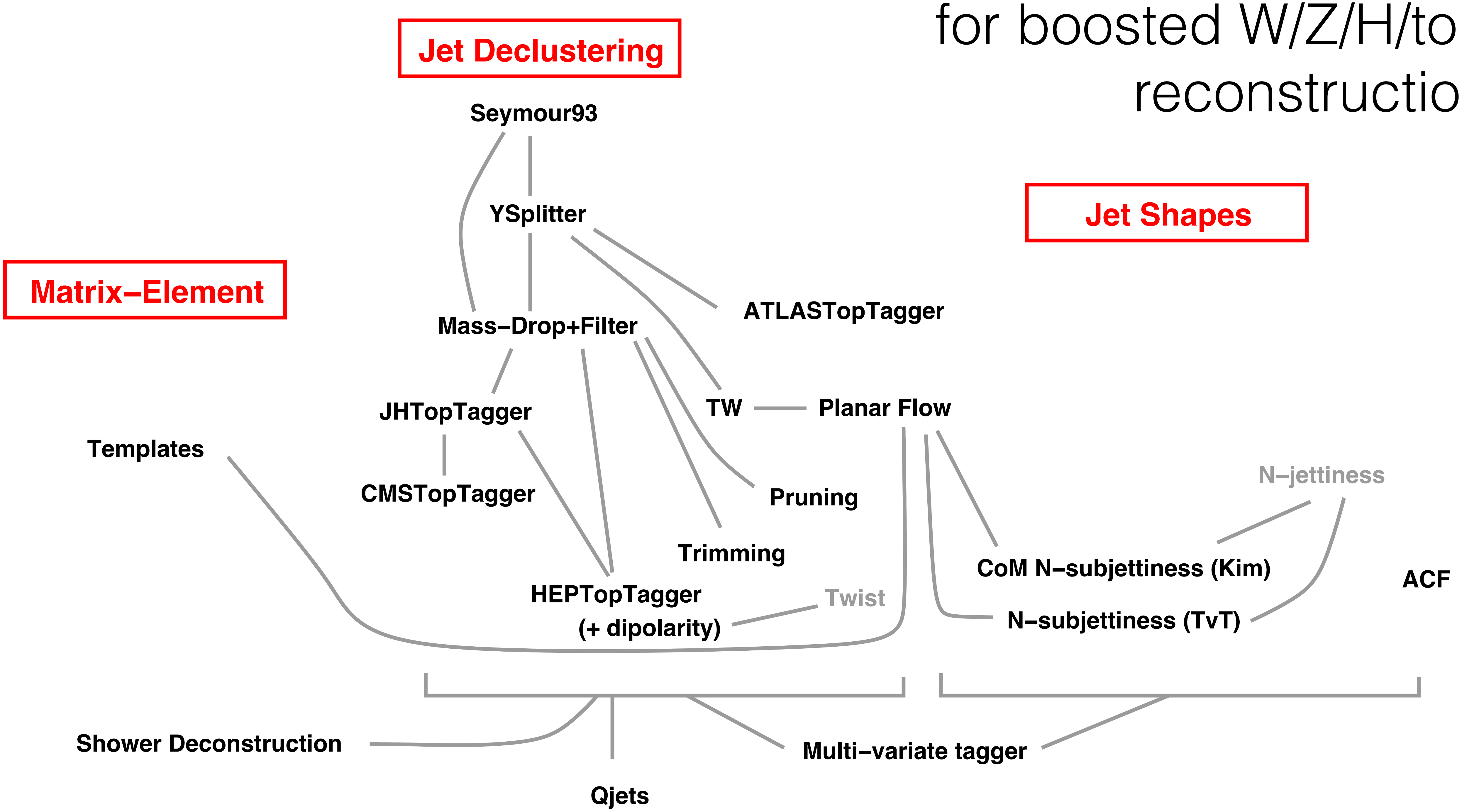


As a field we've devised O(10-20) powerful methods to tag jet substructure.

Many of the methods have been tried out in searches and work; these kinds of methods will be crucial for searches in the years to come.

Very active research field

Some of the tools developed
for boosted W/Z/H/top
reconstruction



Jet Declustering

Jet Shapes

Matrix–Element

Seymour93

YSplitter

ATLASTopTagger

Mass–Drop+Filter

JHTopTagger

TW

Planar Flow

Templates

CMSTopTagger

Pruning

N–jettiness

Trimming

CoM N–subjettiness (Kim)

ACF

HEPTopTagger
(+ dipolarity)

Twist

N–subjettiness (TvT)

Shower Deconstruction

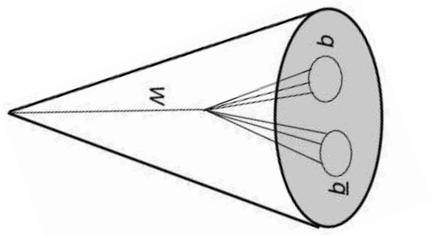
Multi–variate tagger

Qjets

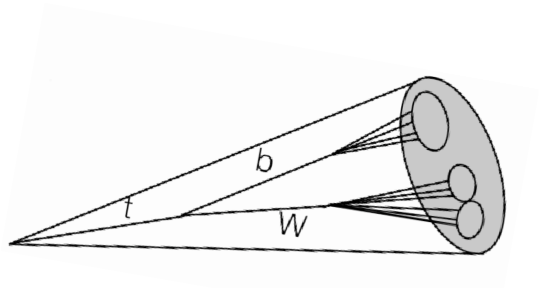
apologies for omitted taggers, arguable links, etc.

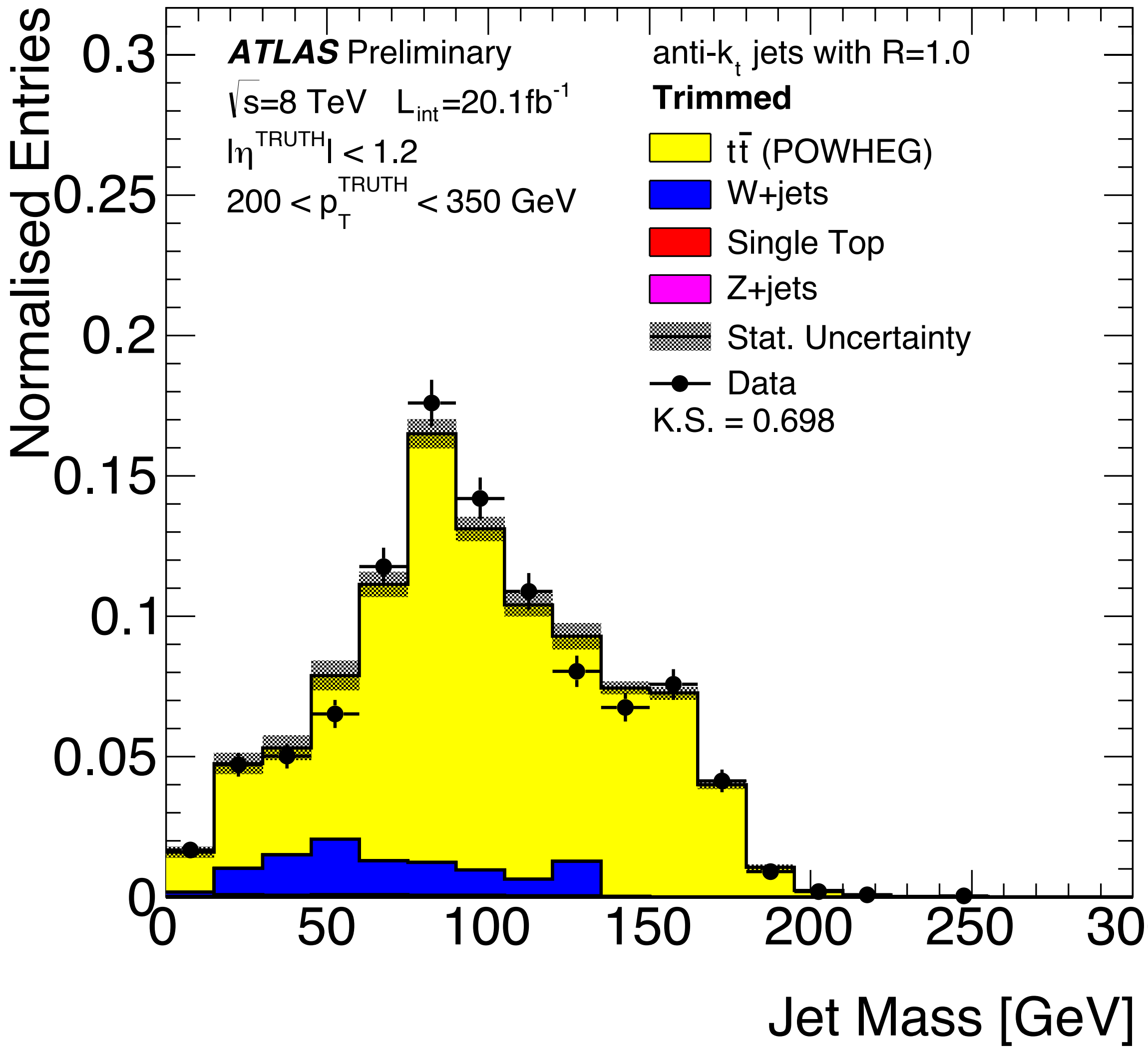
Seeing W’s and tops in a single jet

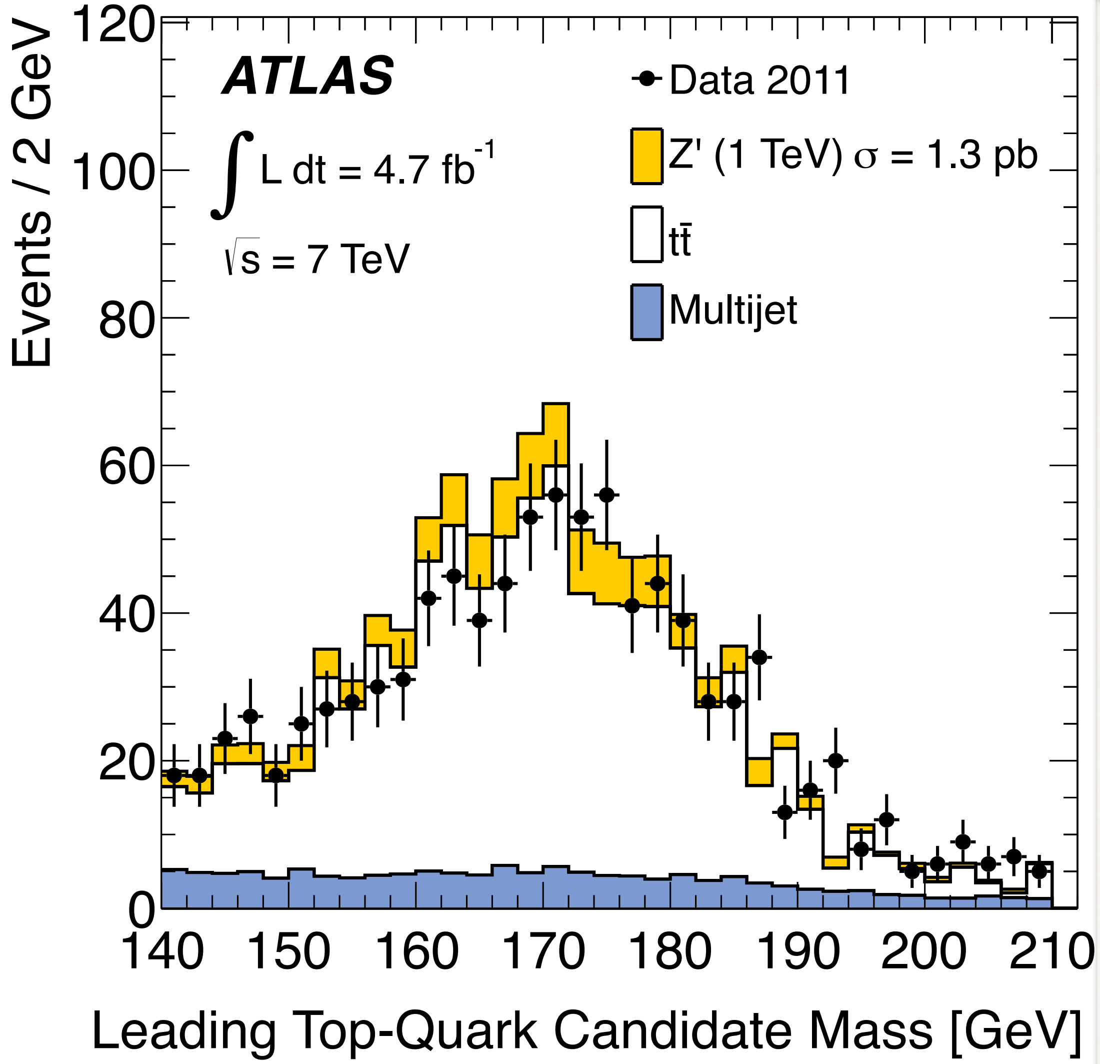
W’s in a single jet



tops in a single jet







As a field we've devised O(10-20) powerful methods to tag jet substructure.

Many of the methods have been tried out in searches and work; these kinds of methods will be crucial for searches in the years to come.

But from outside, the many methods make the field look pretty confusing.

And from inside, I get the impression we don't always know *why or how* the methods work – which is bad if we're looking for robustness.

Is it time to get back to basics?
I.e. to start understanding our tools?

What was the original motivation?

Normal R=0.4/0.5 jet finding fails to find one jet per prong of a boosted [W/Z/H/top/NP] hadronic decay.

We need to make sure that this doesn't prevent us from using EW-scale particles in TeV scale searches.

Question #1:
To what extent are the things we do with "normal" jets (and leptons) mirrored in the things we're doing with "fat" jets?

What have we found out in the meantime?

There's a huge number of things you can do with jet substructure.

Many of the things appear to improve mass resolution, background rejection, etc. [at least in MC simulation]

Question #2:

How should we balance improvements v. "complexity" of method?

What are we comfortable with?

Resolved Analysis

Find one jet/prong

Cut on jet p_t, Δy, ...

What are we comfortable with?

Resolved Analysis

Fat-jet Analysis

Find one jet/prong → Find subjets

Cut on jet p_t, Δy, ... → Cut on subjet z, ΔR, ...

[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]

What are we comfortable with?

Resolved Analysis | **Fat-jet Analysis**

Find one jet/prong → Find subjets

Cut on jet p_t, Δy, ... → Cut on subjet z, ΔR, ...

[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]

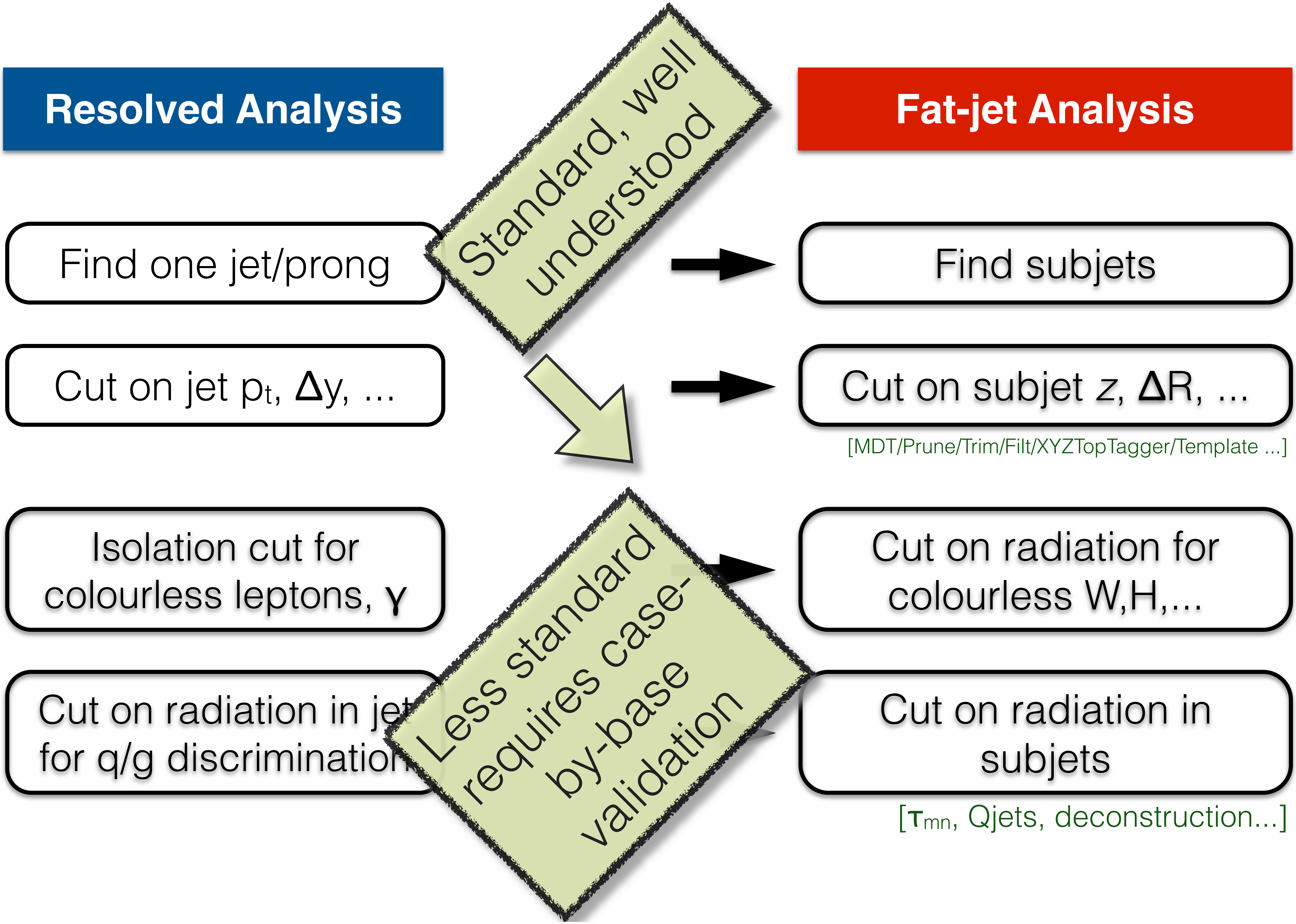
Isolation cut for colourless leptons, γ

Cut on radiation in jet for q/g discrimination

What are we comfortable with?

Resolved Analysis		**Fat-jet Analysis**
Find one jet/prong	→	Find subjets
Cut on jet p_t, Δy, ...	→	Cut on subjet z, ΔR, ... [MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]
Isolation cut for colourless leptons, γ	→	Cut on radiation for colourless W,H,...
Cut on radiation in jet for q/g discrimination	→	Cut on radiation in subjets [τ_{mn}, Qjets, deconstruction...]

What are we comfortable with?



Resolved Analysis

Find one jet/prong

Cut on jet p_t, Δy, ...

Isolation cut for colourless leptons, γ

Cut on radiation in jet for q/g discrimination

Standard, well understood

Less standard requires case-by-base validation

Fat-jet Analysis

Find subjets

Cut on subjet z, ΔR, ...

[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]

Cut on radiation for colourless W,H,...

Cut on radiation in subjets

[τ_{mn}, Qjets, deconstruction...]

analytical understanding: why?

Better Insight

Can guide taggers' use in experimental analyses

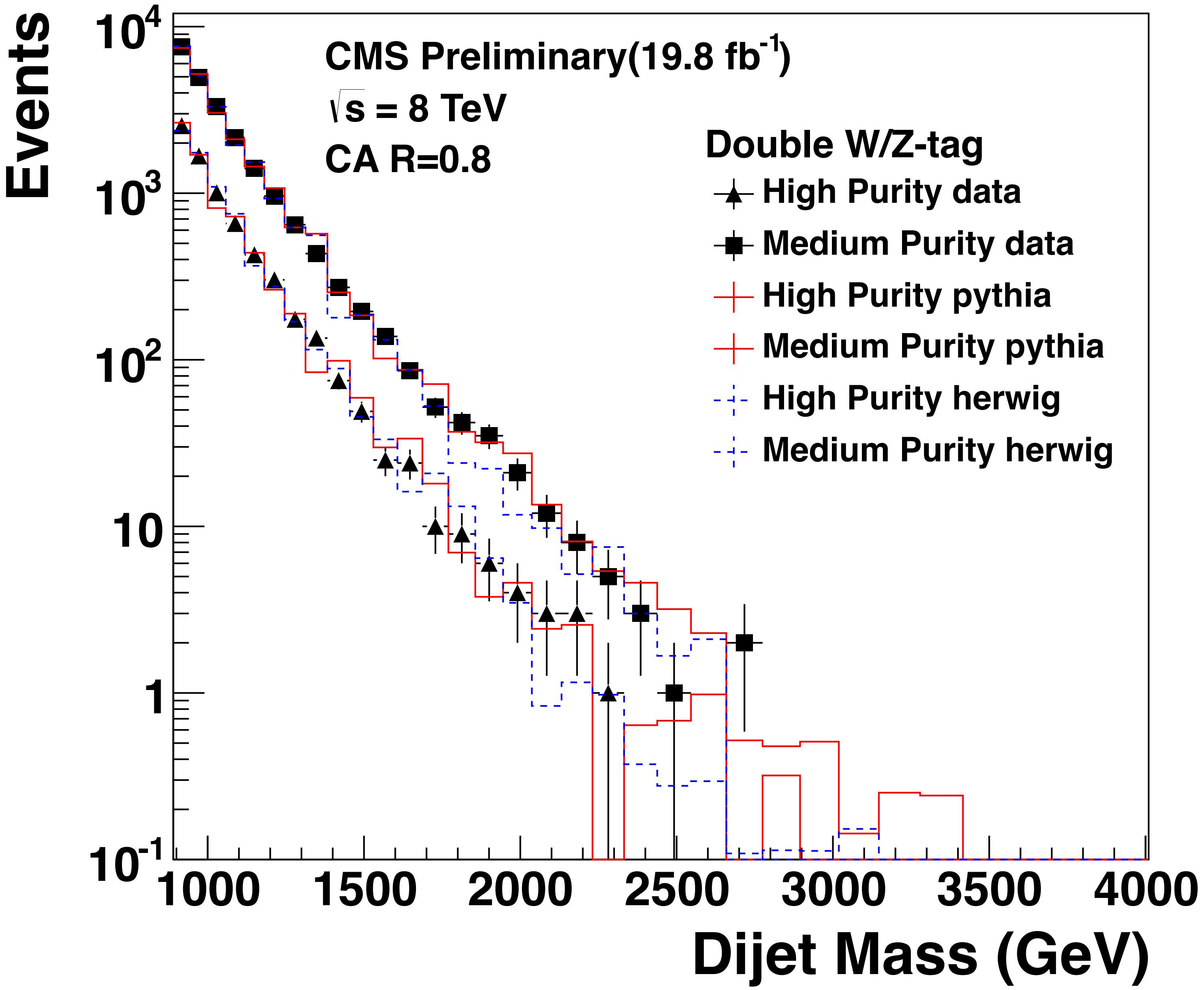
It may help us design better taggers

Robustness

You know what you predict, what you don't

Unlike MC, you have powerful handles for cross-checks & accuracy estimates

There is a "right" answer



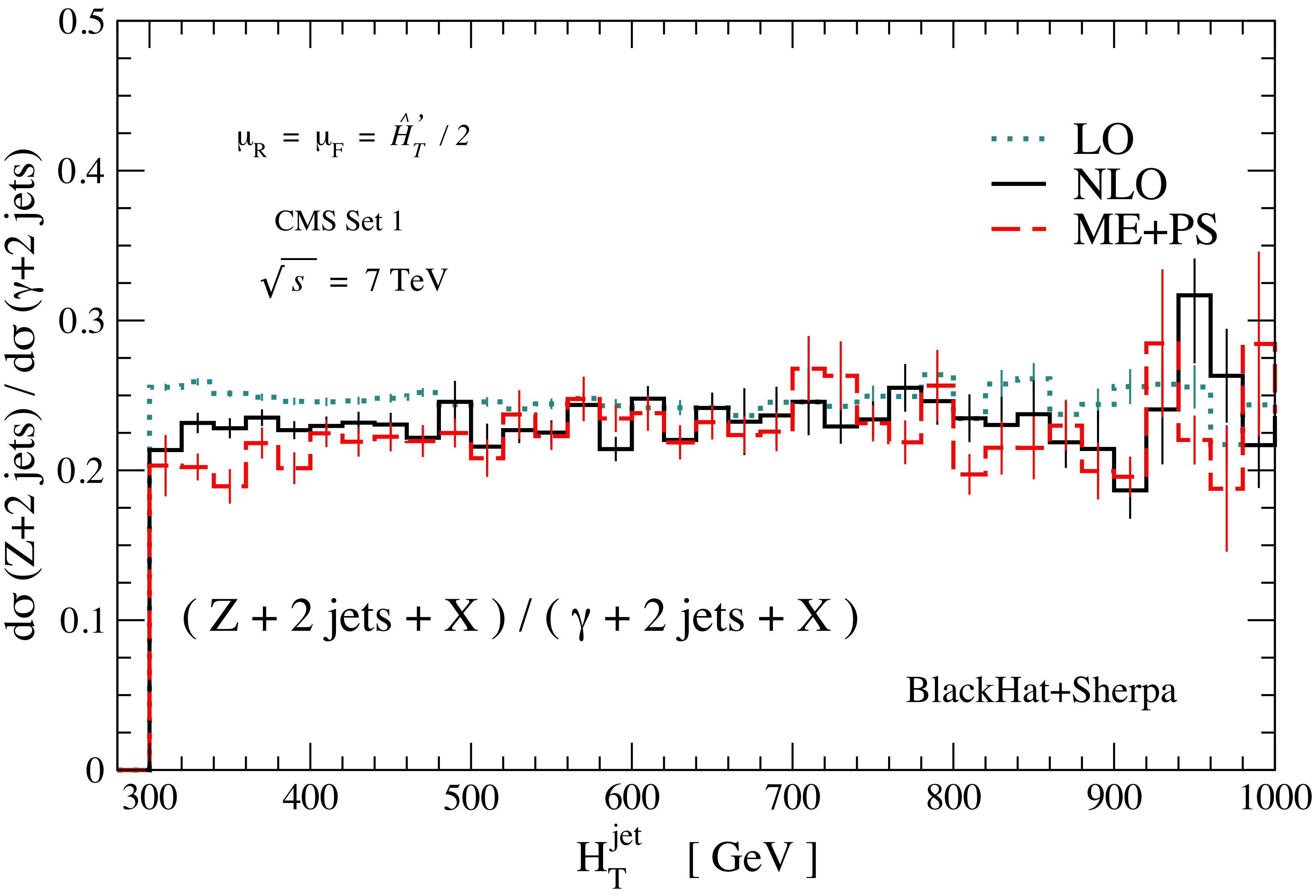
Search for resonances in doubly-tagged dijet events.

Tagging = pruning + tau21 cut

Note different Herwig++ and Pythia6 shapes

Understanding your taggers means you know what tools you can safely use with them

For robustness, you can then choose taggers whose distributions can be predicted in many ways



Just like MET(Z→νν) in multijets is reliably estimated from γ+jets because multiple types of calculations of the ratio agree

The key variables

For phenomenology

Jet mass: *m*

[as compared to W/Z/H or top mass]

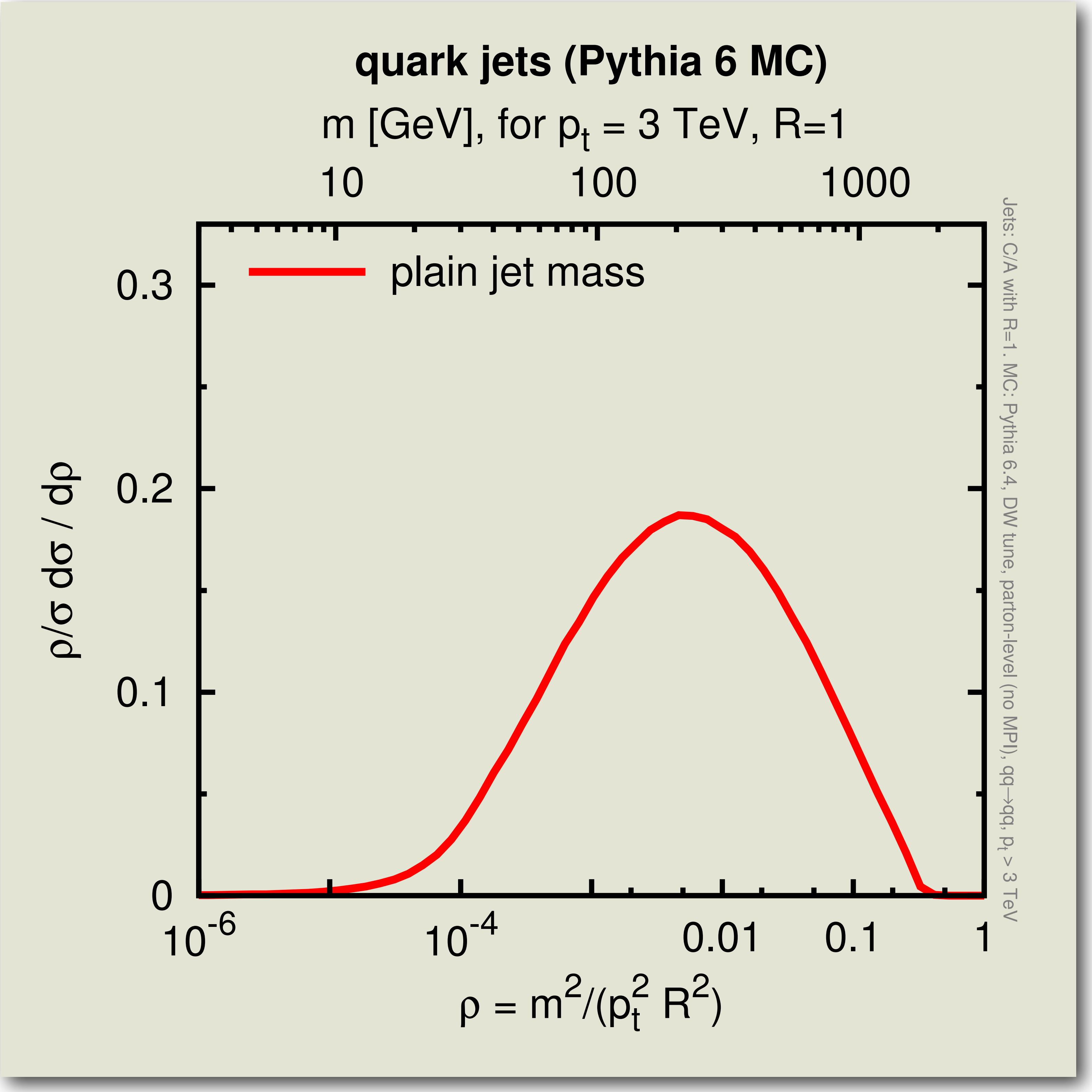
For QCD calculations

$$\rho = \frac{m^2}{p_t^2 R^2}$$

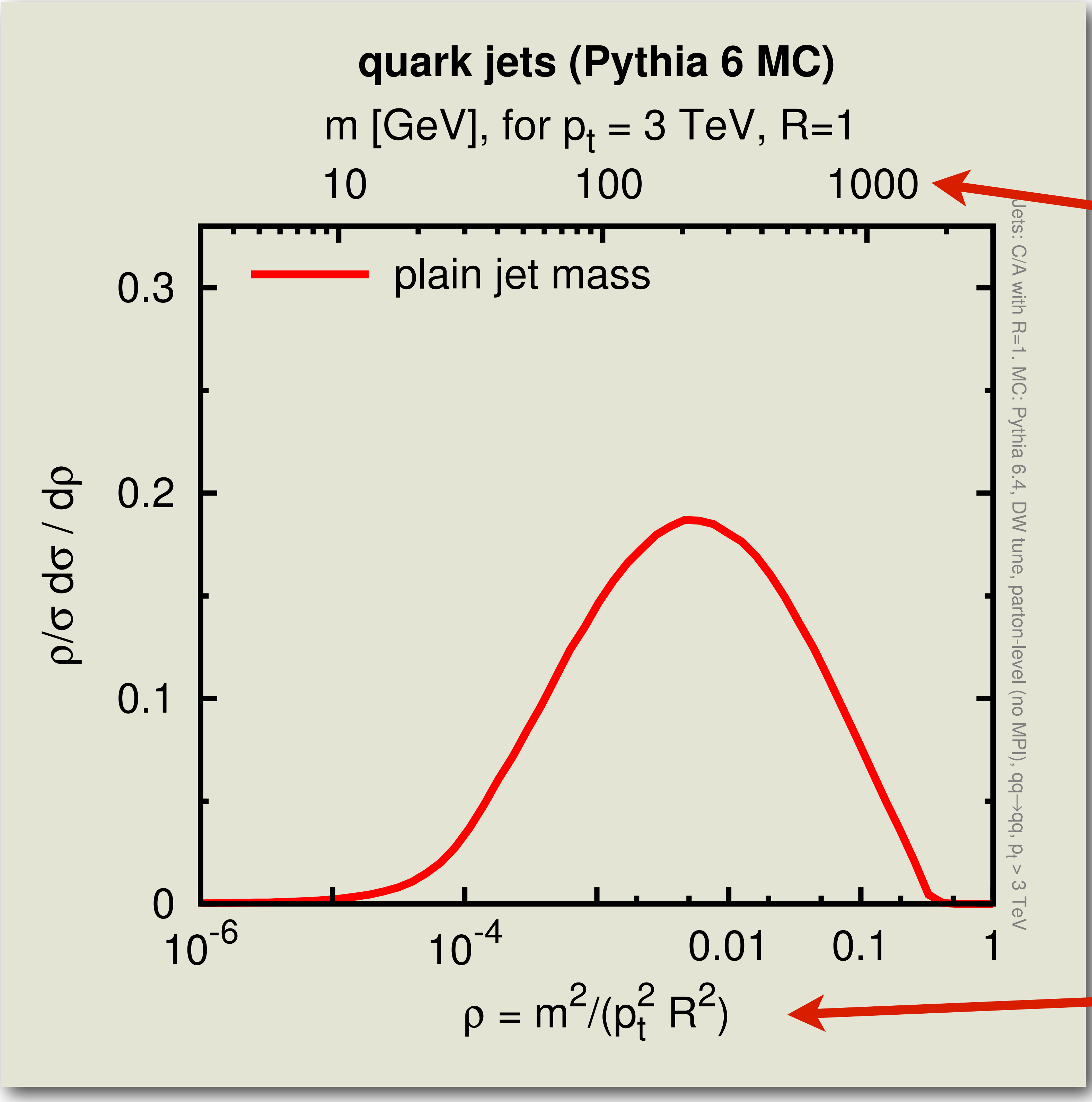
[R is jet opening angle – or radius]

Because **ρ** is invariant under boosts along jet direction

Start with “plain” jet mass



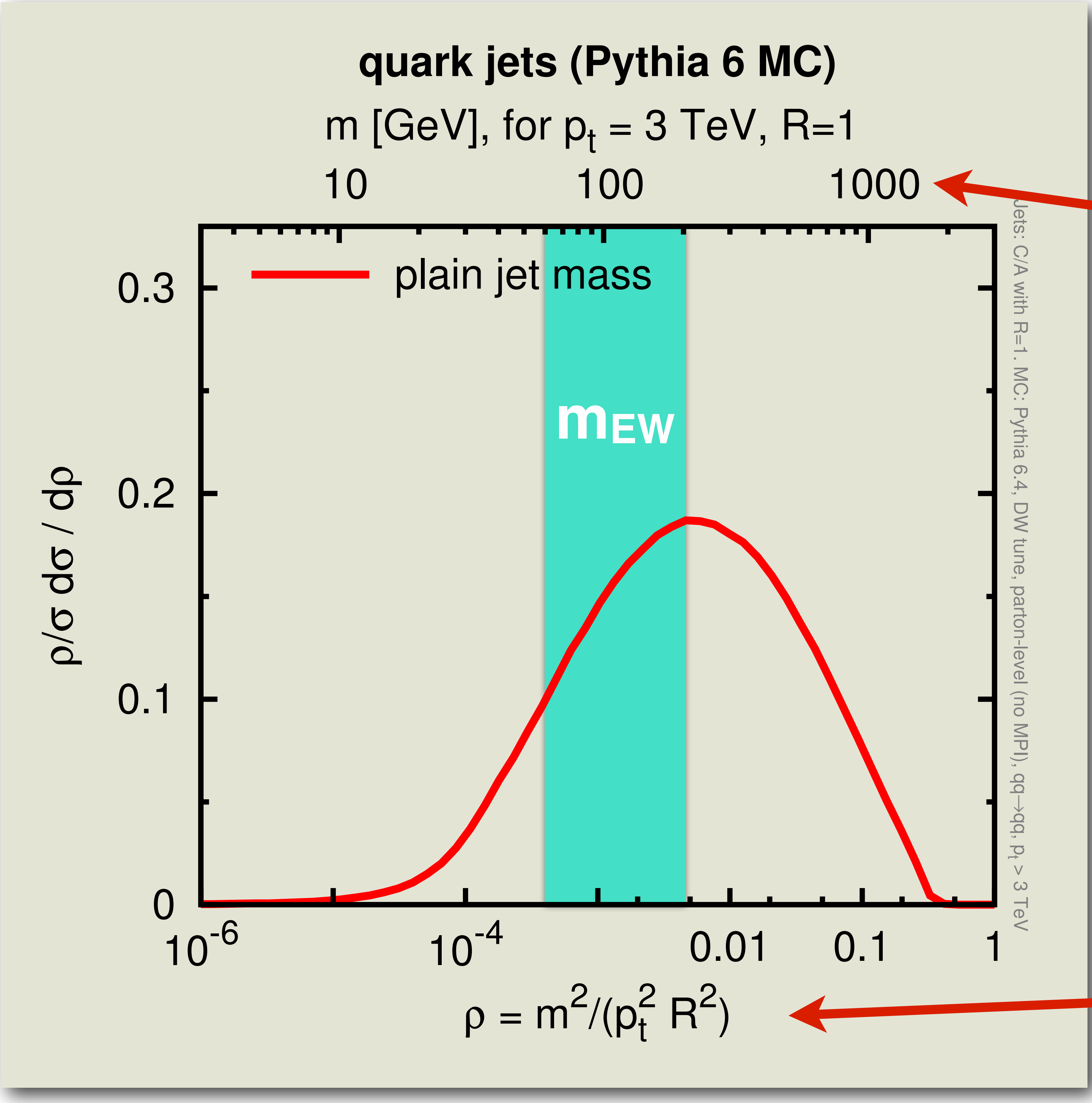
Start with "plain" jet mass



Physical mass for 3 TeV, R=1 jets

ρ ~ Rescaled mass^2
(i.e. the QCD variable)

Start with “plain” jet mass

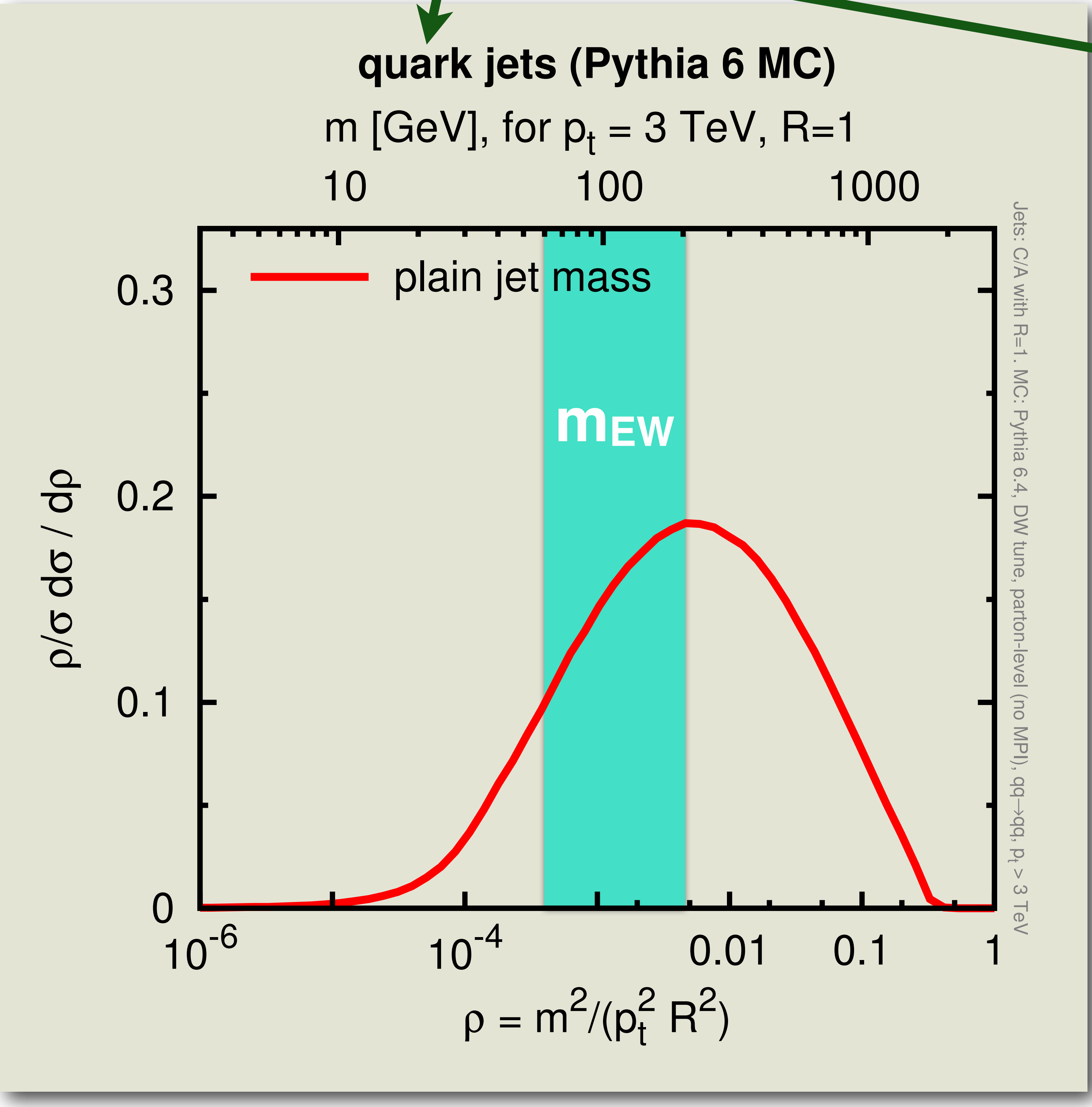


Physical mass for 3 TeV, R=1 jets

ρ ~ Rescaled mass²

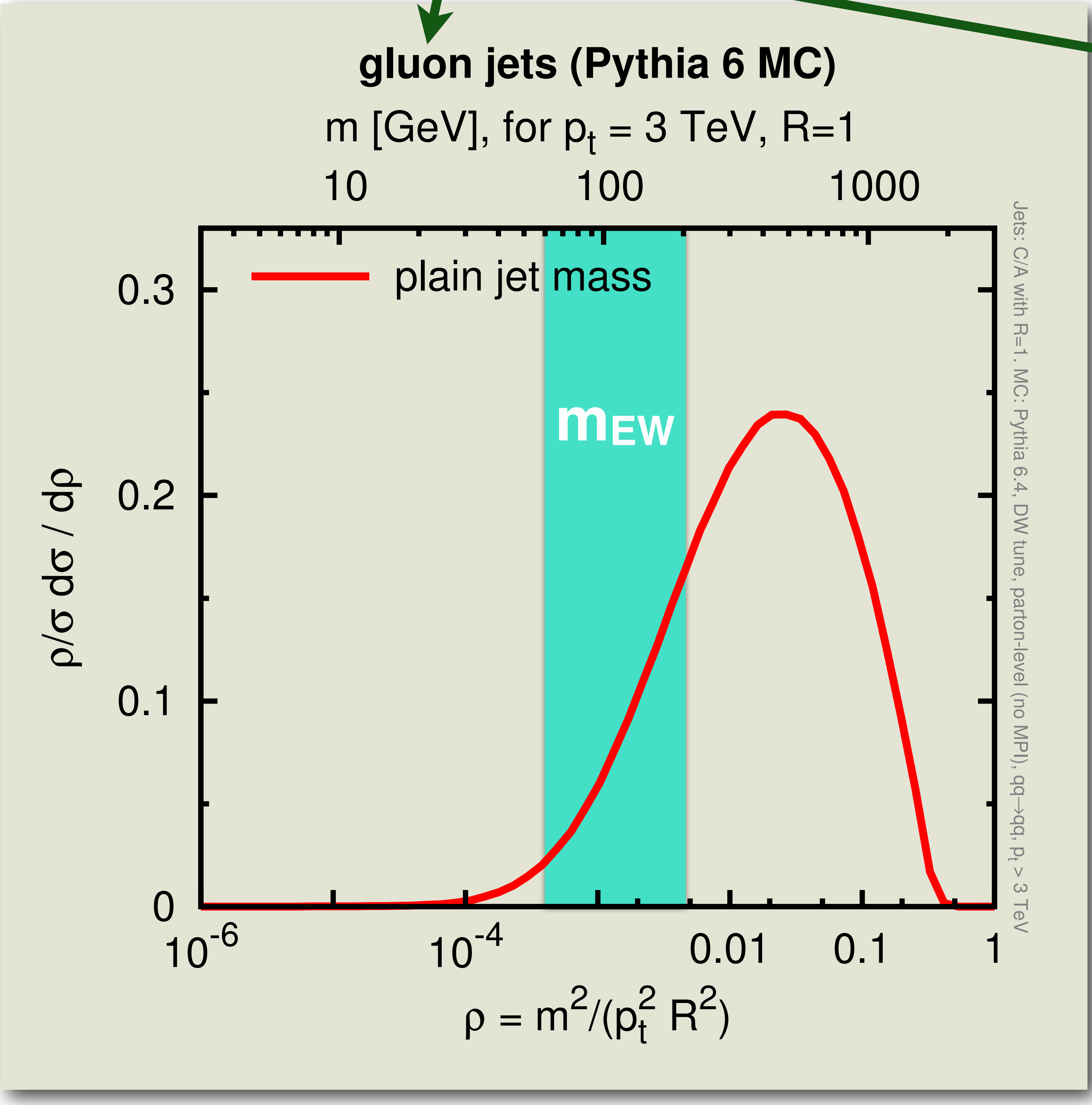
(i.e. the QCD variable)

Start with “plain” jet mass



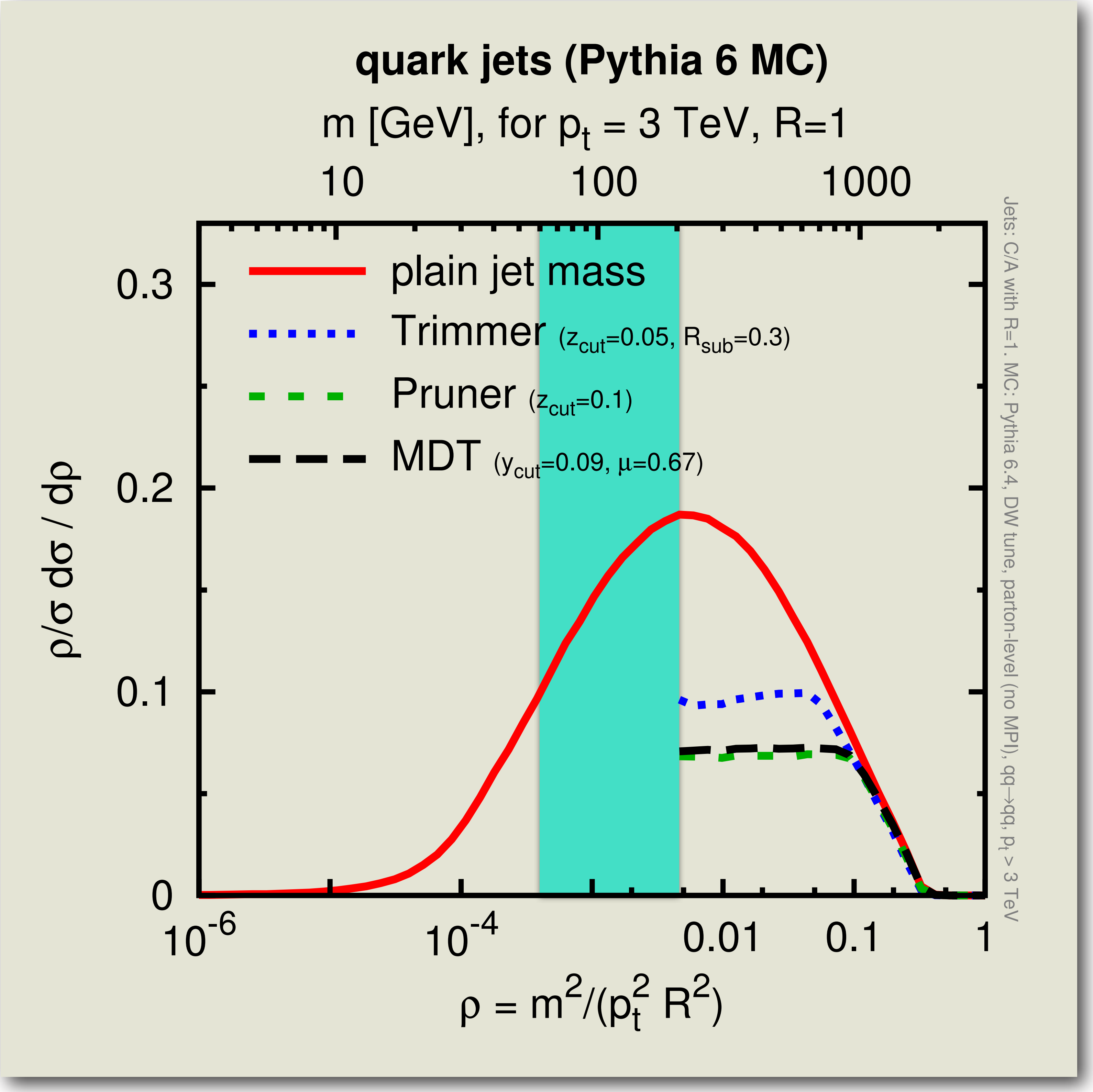
Quark and gluon jets are different, so we treat them separately

Start with “plain” jet mass



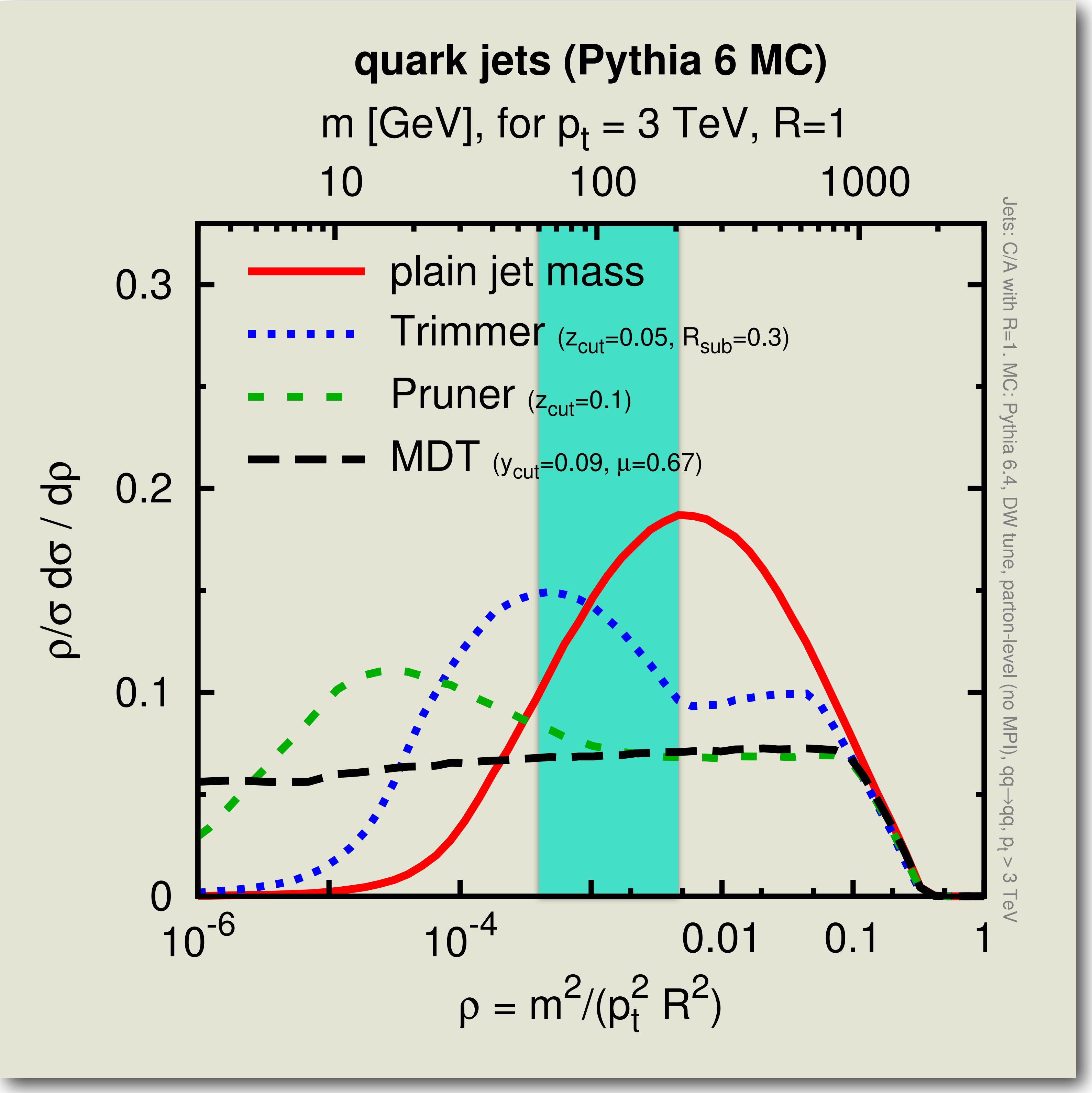
Quark and **gluon** jets are different, so we treat them separately

Now examine “taggers/groomers”



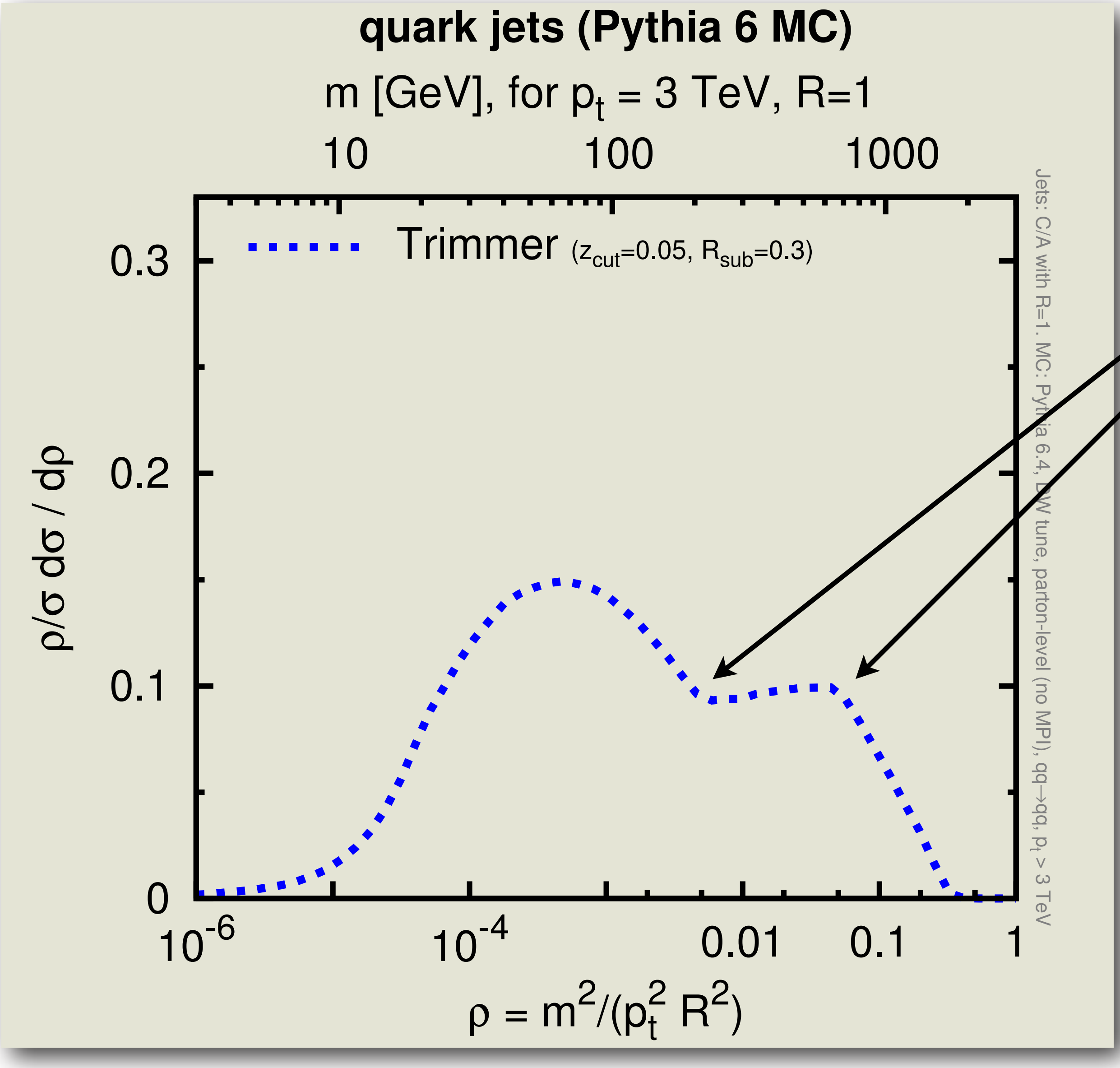
Different taggers
can be
quite similar

Now examine "taggers/groomers"



But only for a limited range of masses

What do we want to find out?



Where exactly are the kinks?
How do their locations depend
on z_{cut}, R_{sub}?

Kinks are especially dangerous for data-driver backgrounds

What physics is relevant in the different regions?

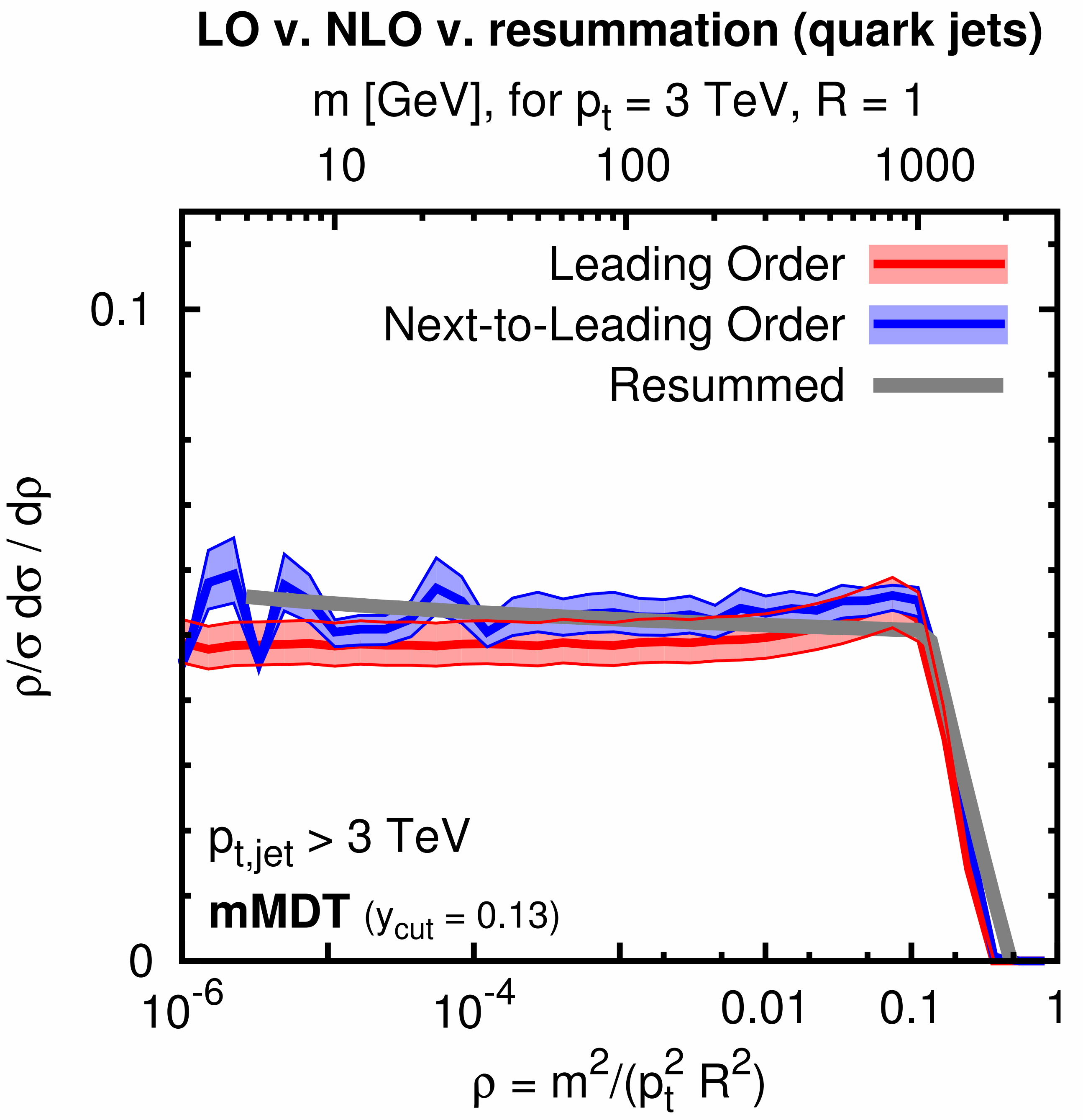
Because then you have an idea of how well you control it

And maybe you can make better taggers

Does tagging/grooming make precise QCD impossible?

Matt Schwartz @ Boost 2012

mMDT resummation v. fixed order

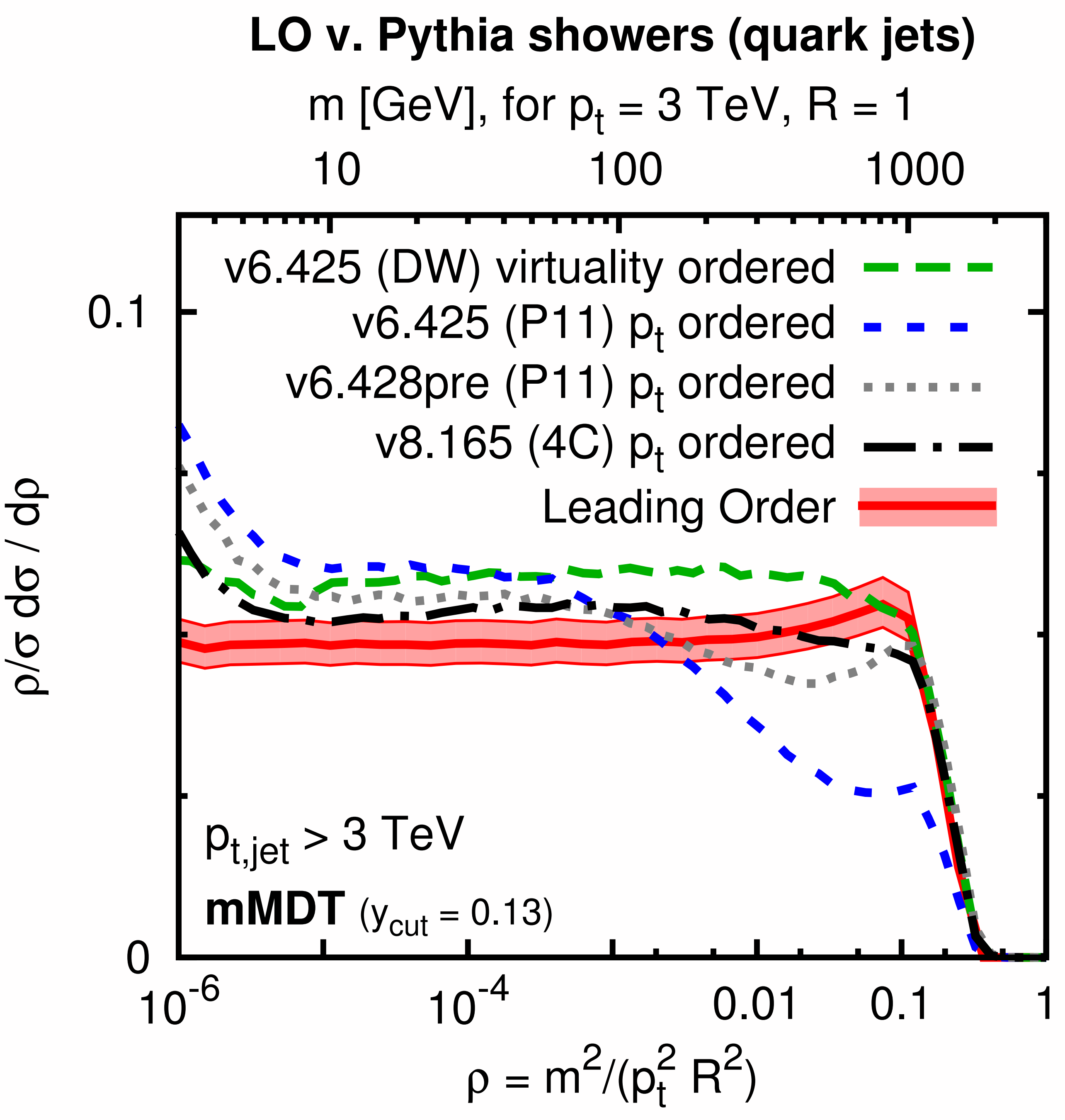


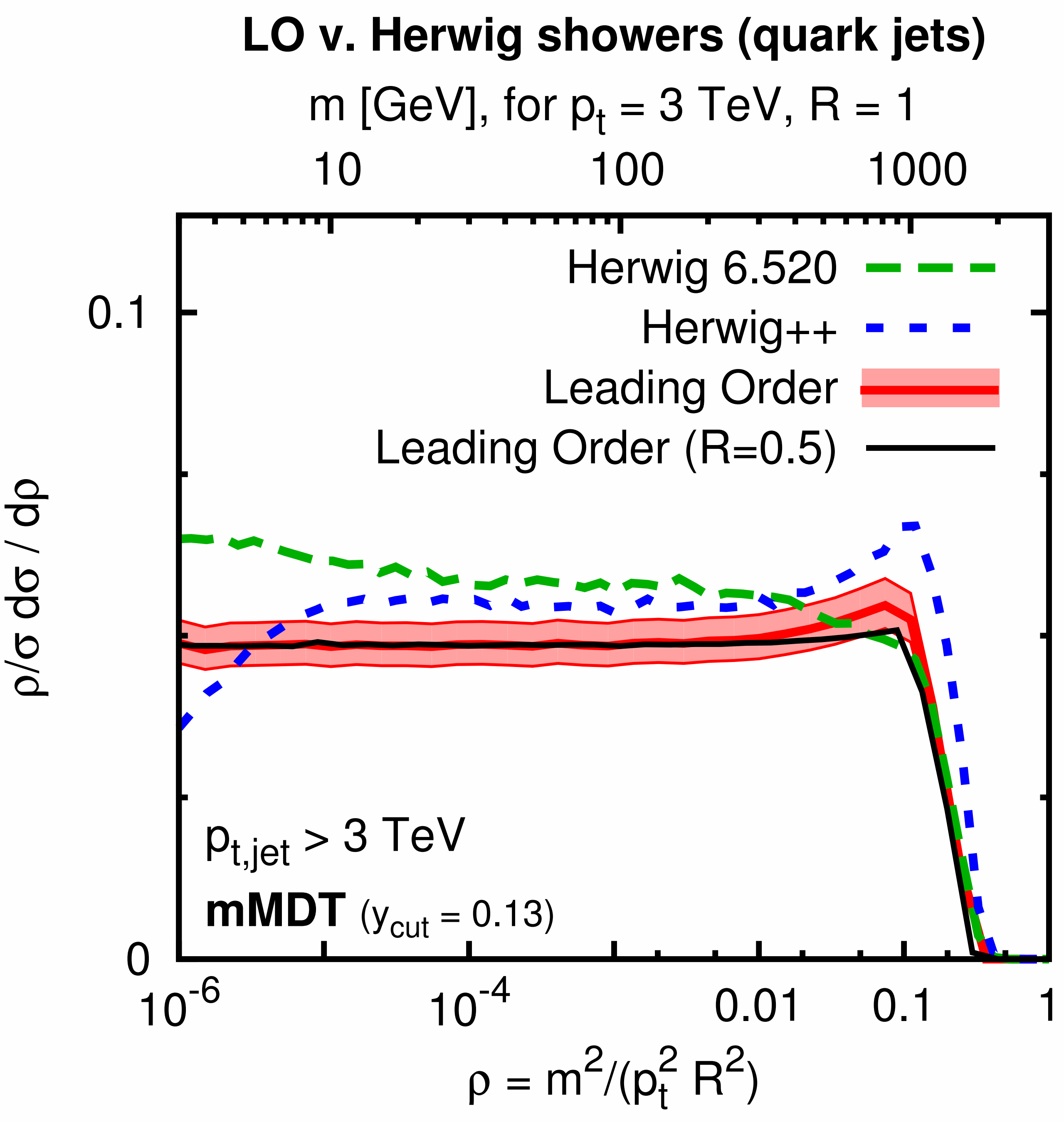
NLO from NLOJet++

Because we only have single logs, fixed-order is valid over a broader than usual range of scales

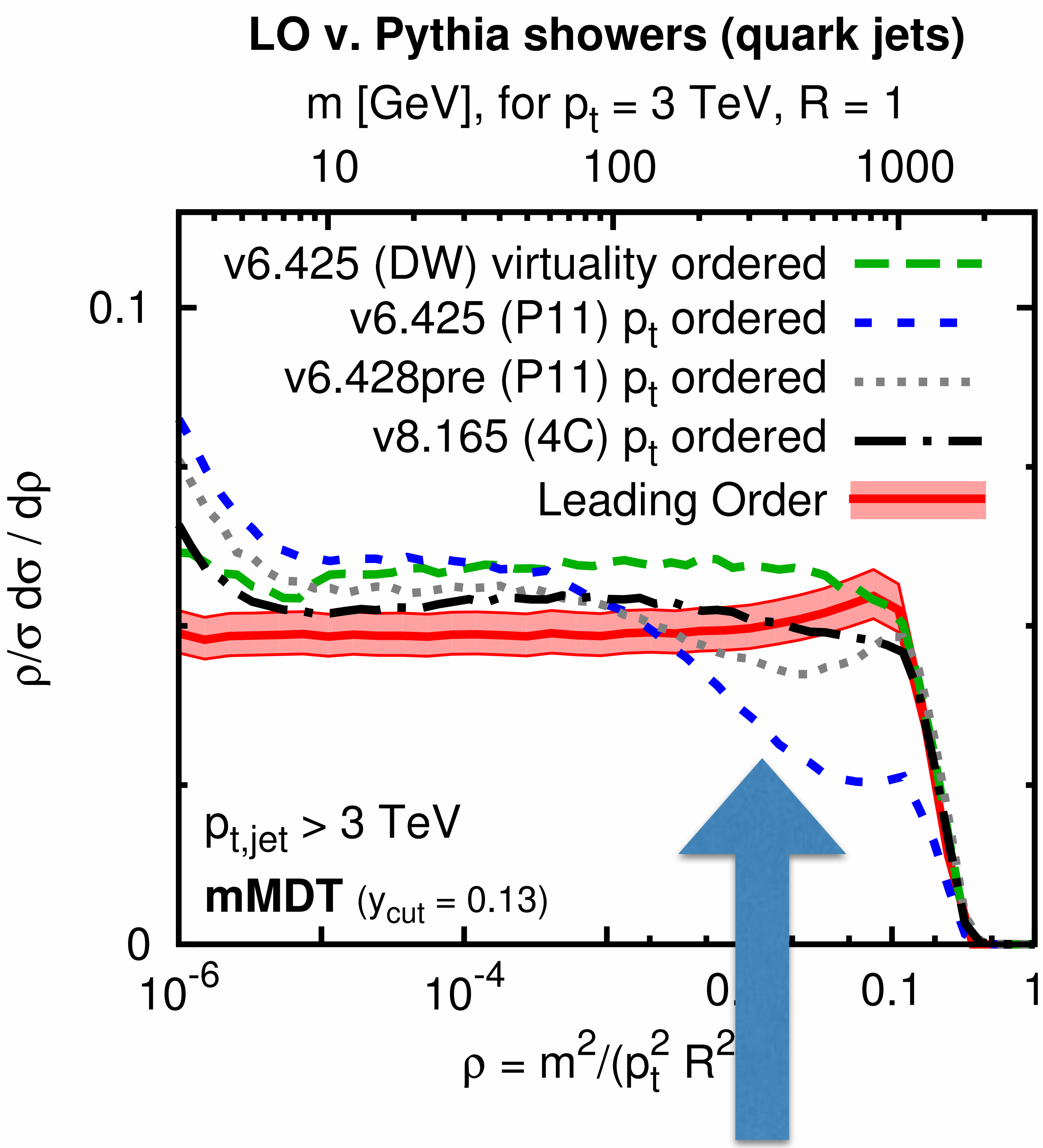
(helped by fortuitous cancellation between running coupling and single-log Sudakov)

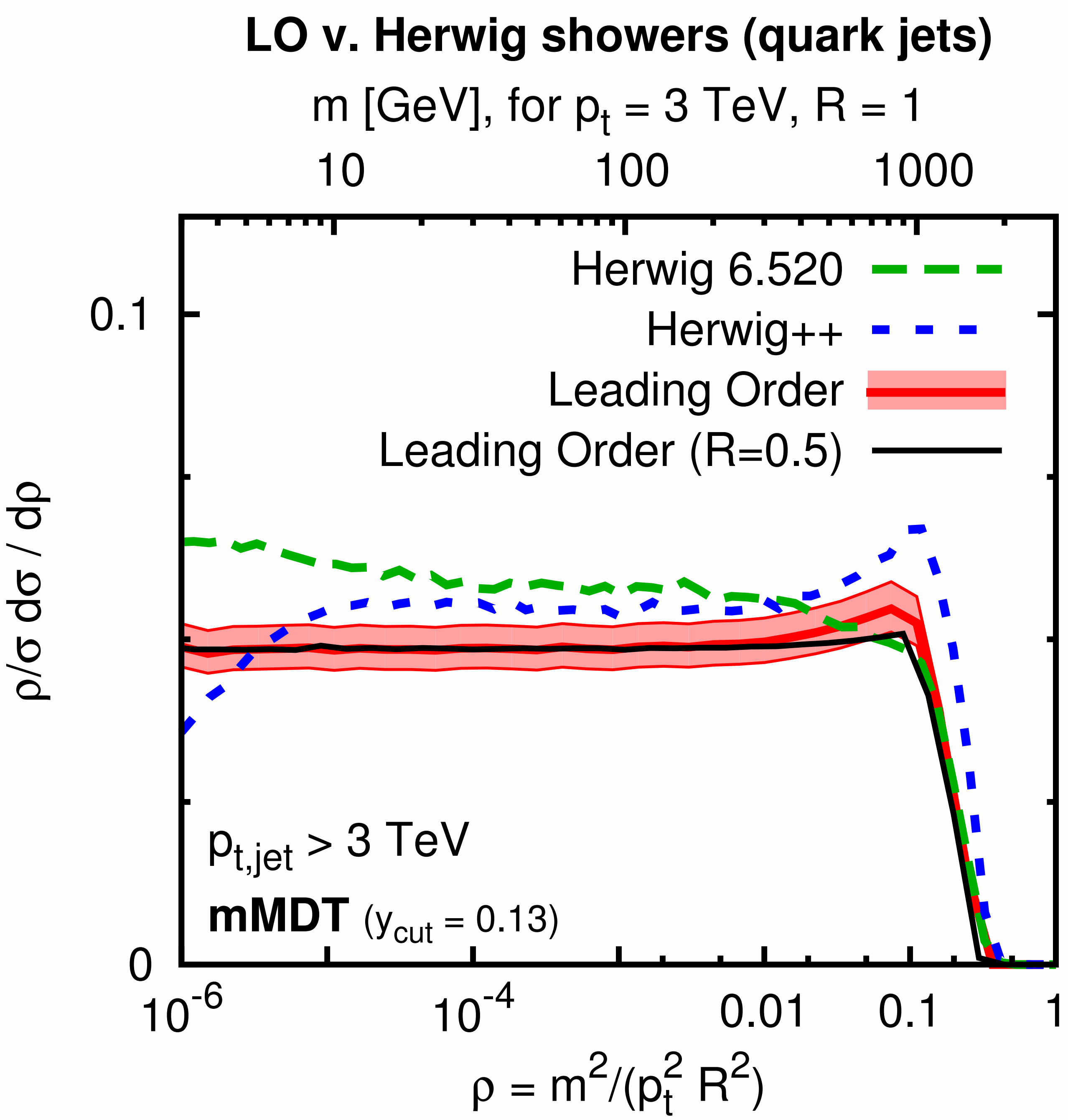
mMDT: comparing many showers


LO v. Pythia showers (quark jets)


LO v. Herwig showers (quark jets)

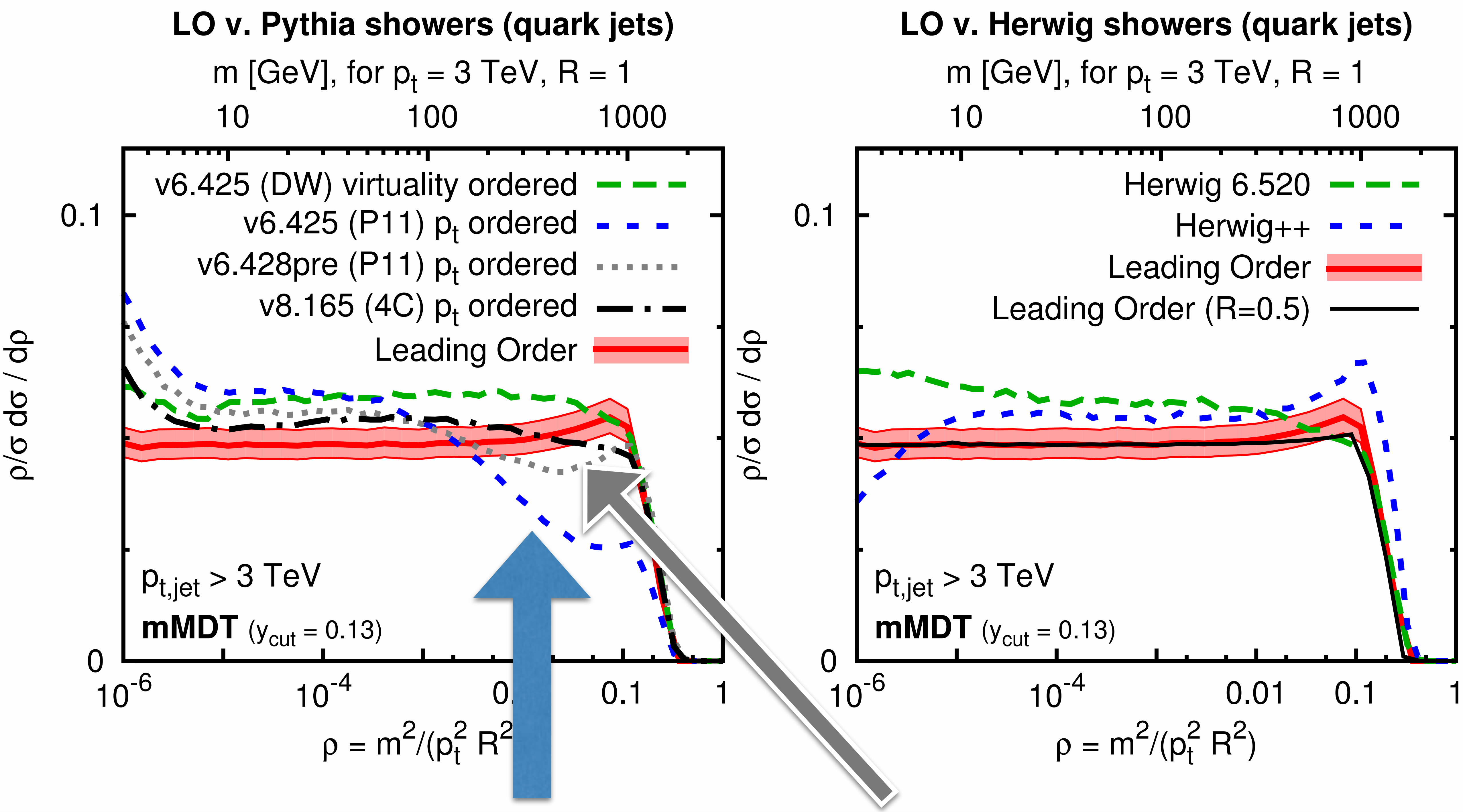
mMDT: comparing many showers





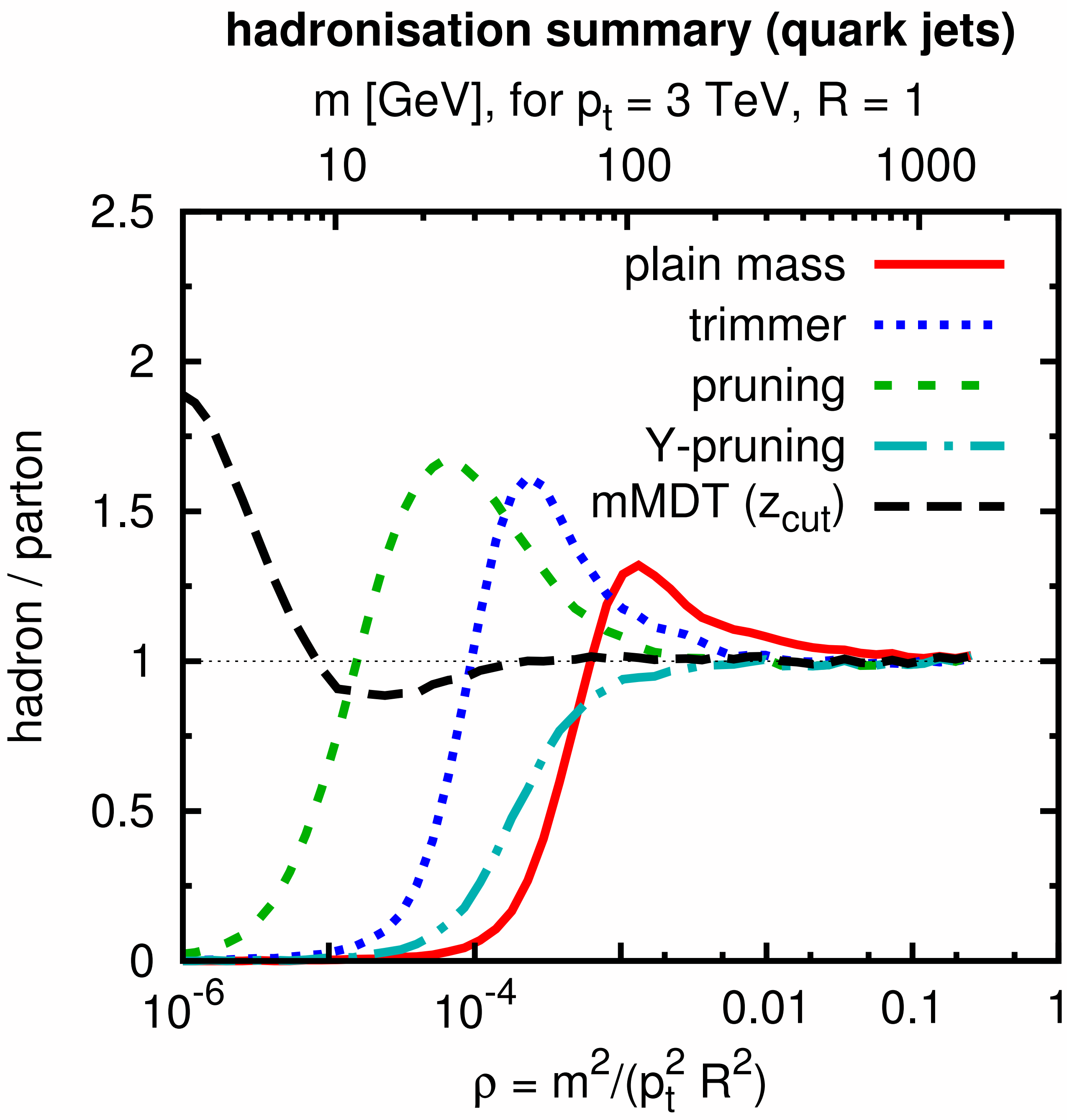
Issue found in Pythia 6 p_t-ordered shower → promptly identified and fixed by Pythia authors!

mMDT: comparing many showers



Issue found in Pythia 6 p_t-ordered shower → promptly identified and fixed by Pythia authors!

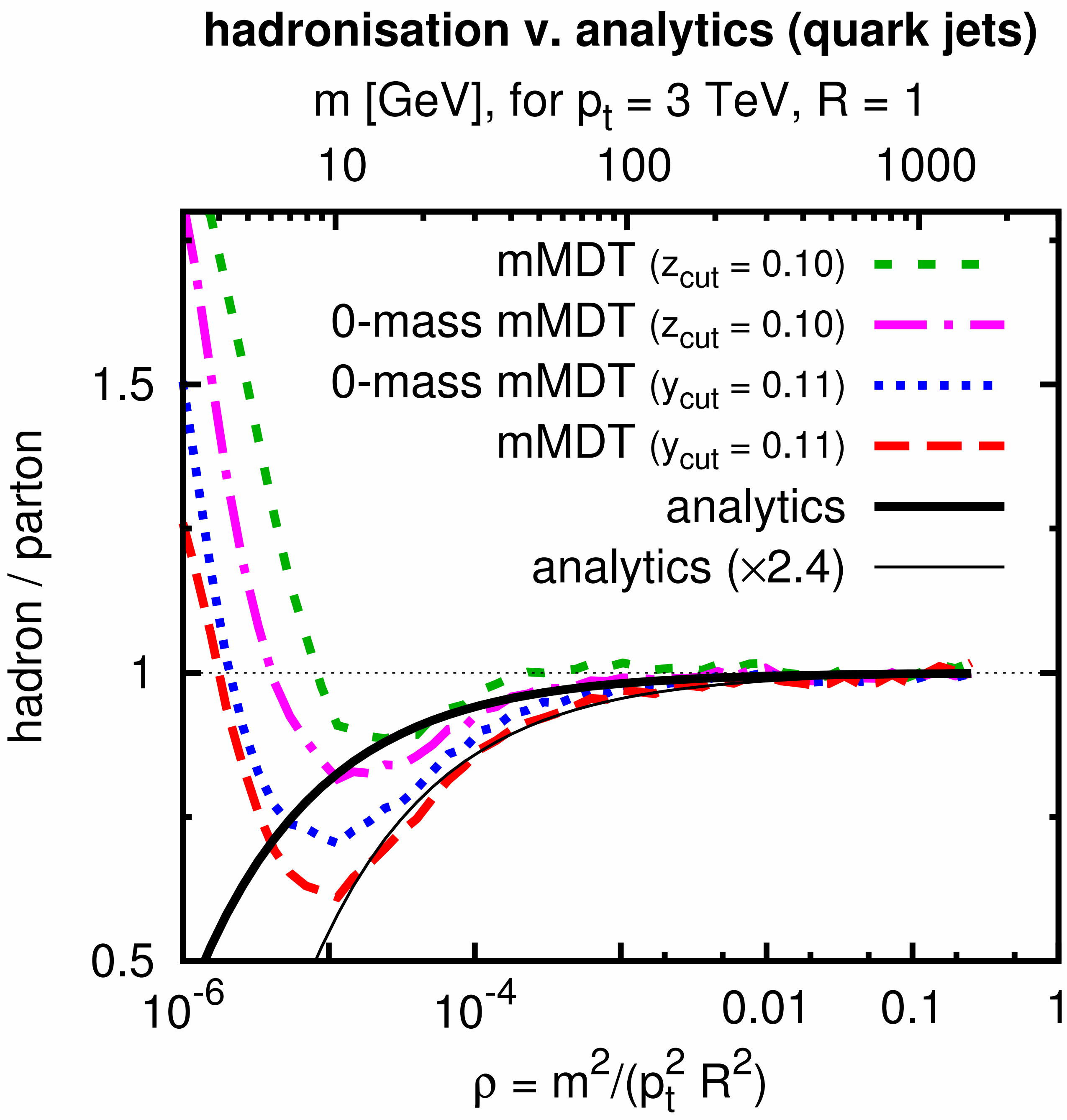
Hadronisation effects



Nearly all taggers have large hadronisation effects:

15 – 60%
for m = 30 – 100 GeV

Hadronisation effects



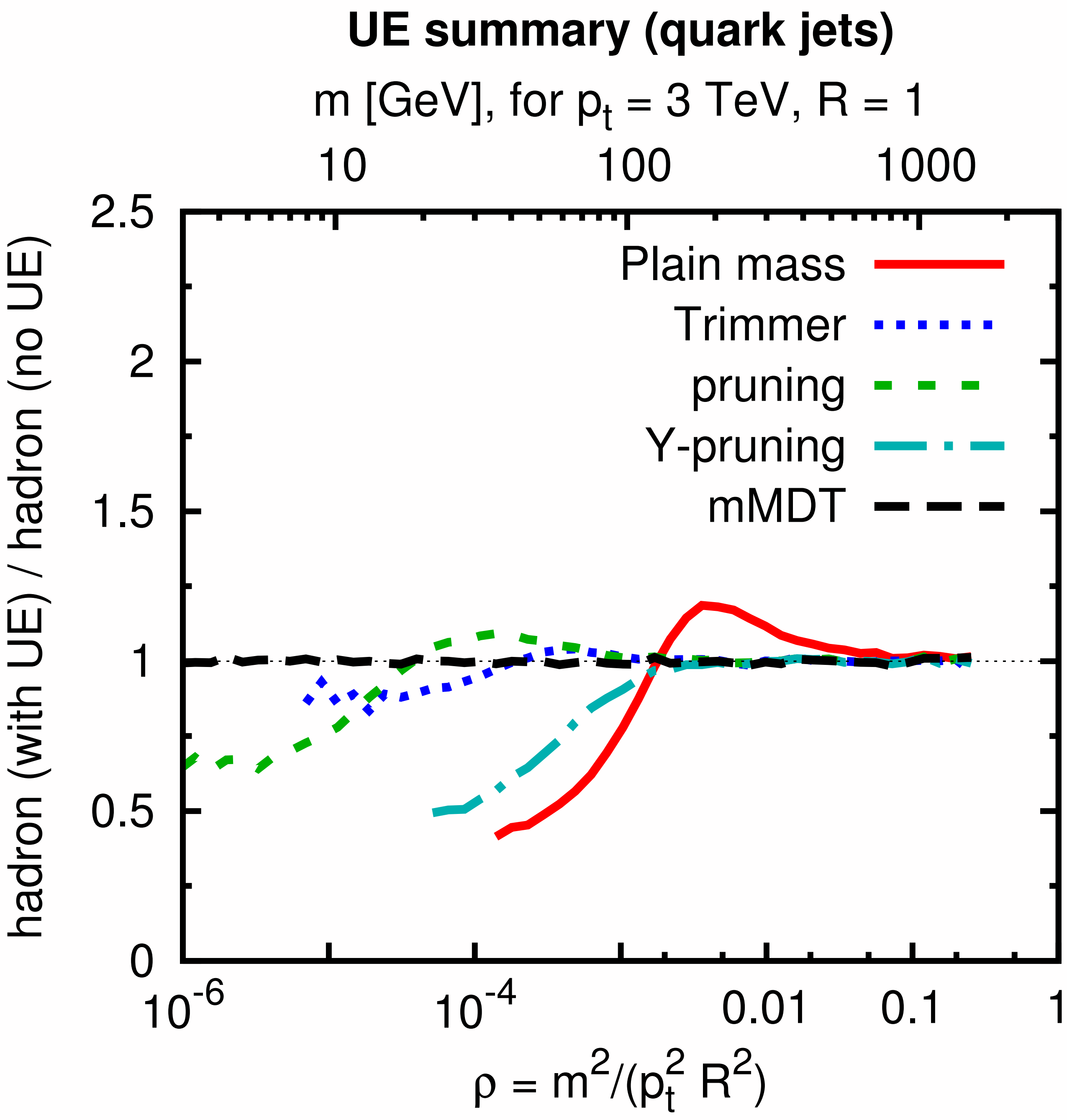
Exception is (m)MDT.

In some cases
just few % effect.

m-dependence of
hadronisation even
understood analytically!

$$\frac{d\sigma}{dm}^{\text{hadron}} \simeq \frac{d\sigma}{dm}^{\text{parton}} \left(1 - c\frac{\Lambda}{m}\right)$$

Underlying Event (UE)

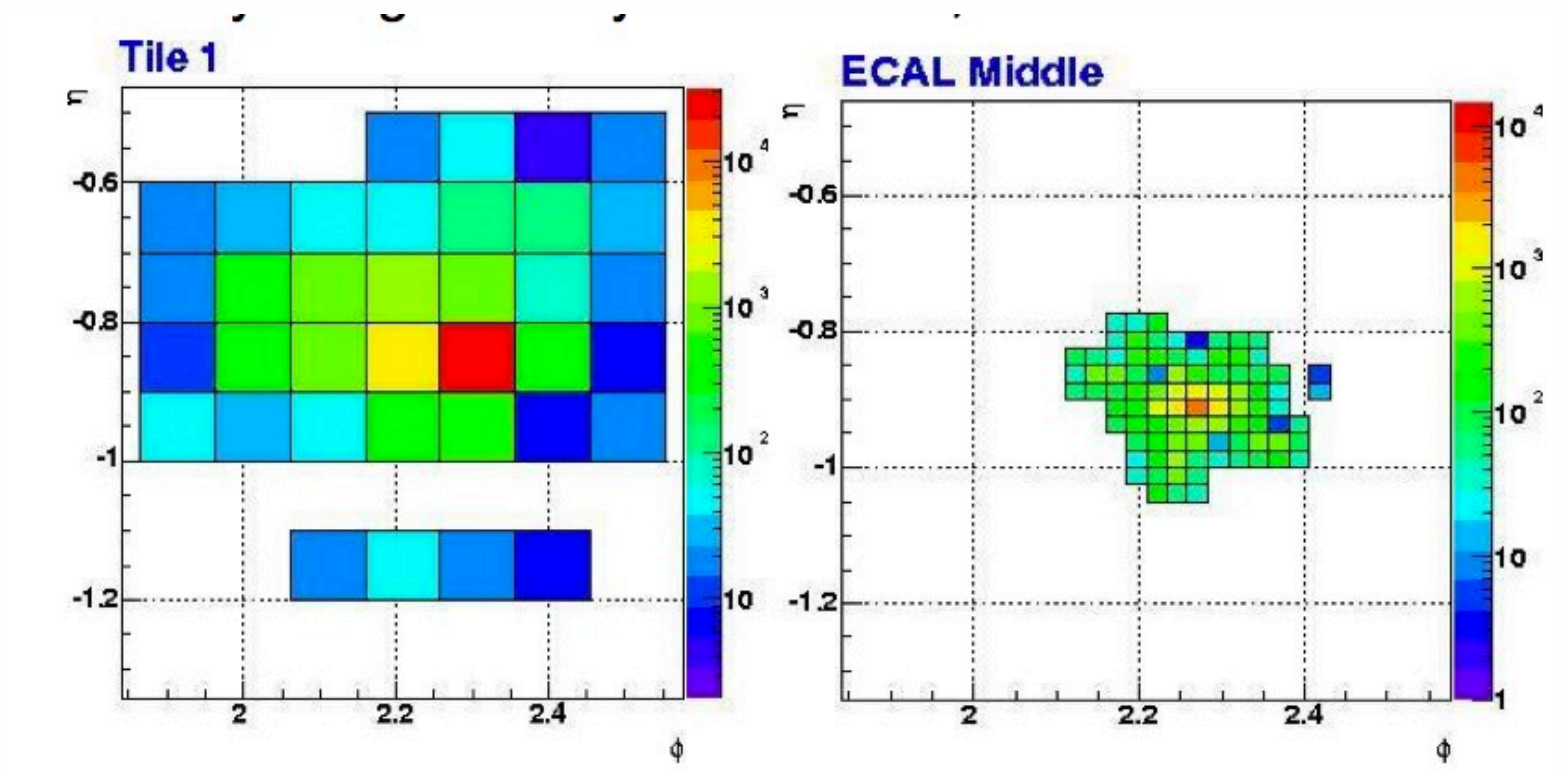


Underlying event impact much reduced relative to jet mass

Almost zero for mMDT

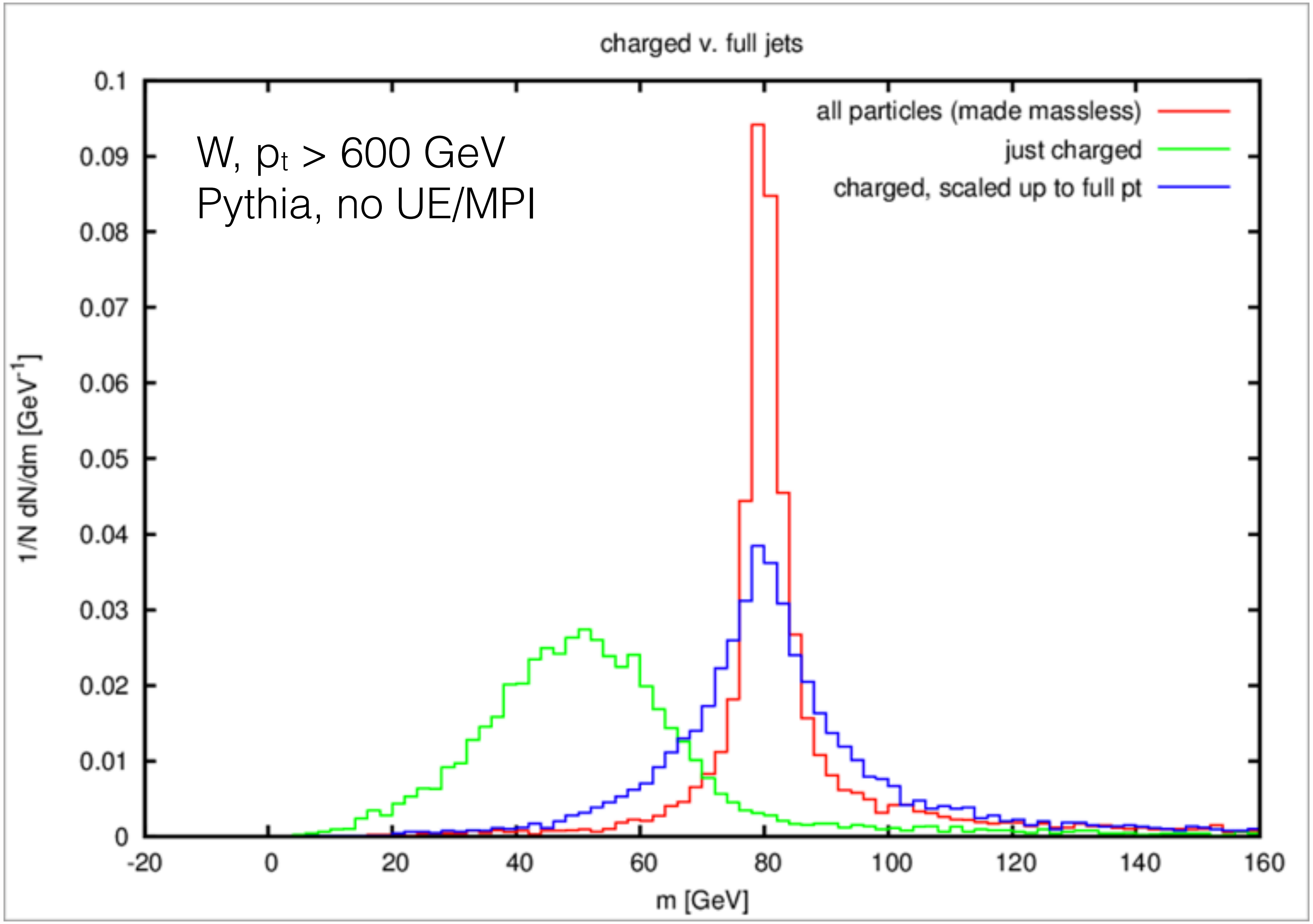
(this depends on jet p_t)

A question from theorists to experimenters: How well can you work around detector granularity at high p_t?



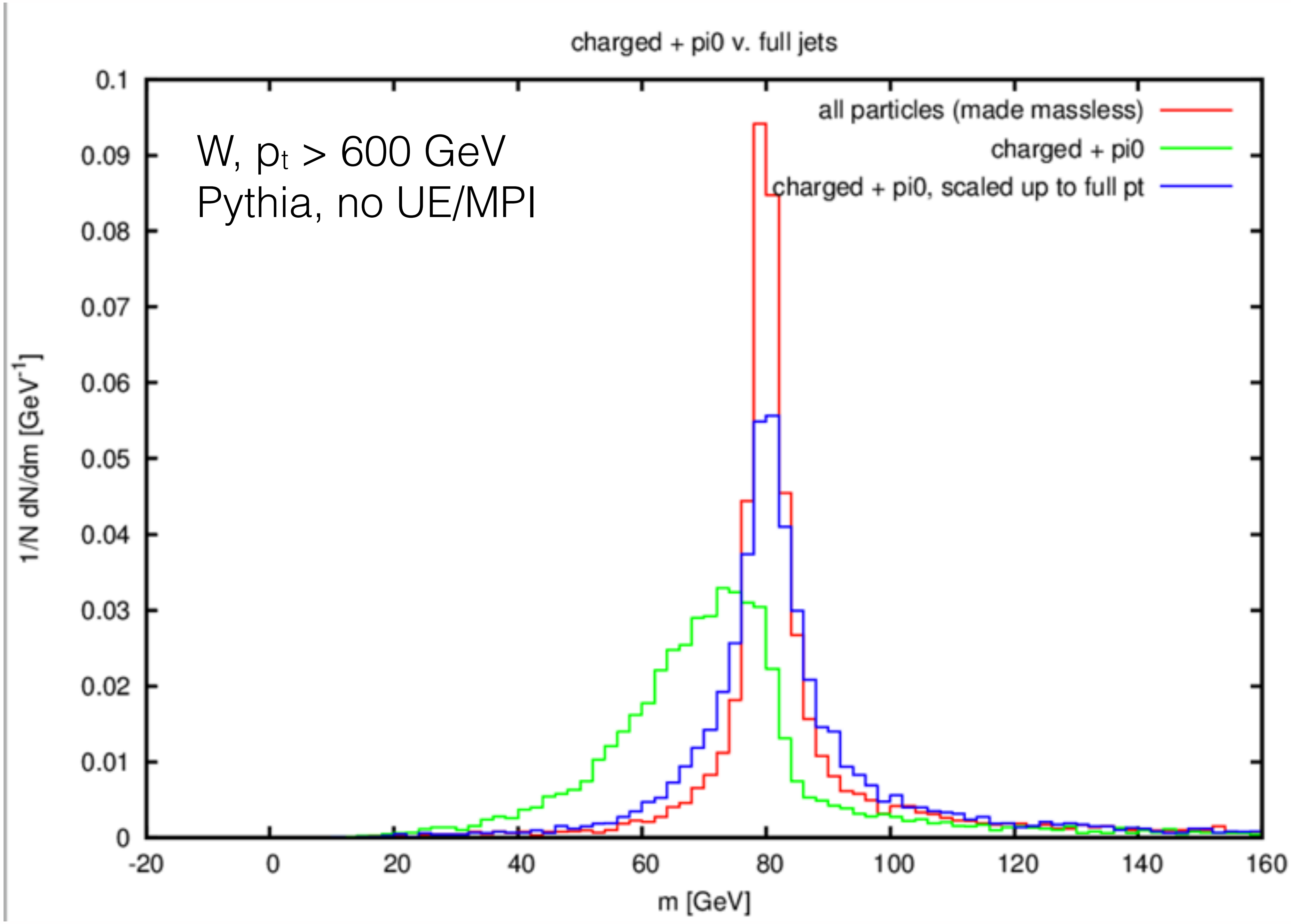
Detector resolution? Use Topocluster/PFlow directly. Or charged tracks scaled to total jet energy?

cf. work by Tweedie et al



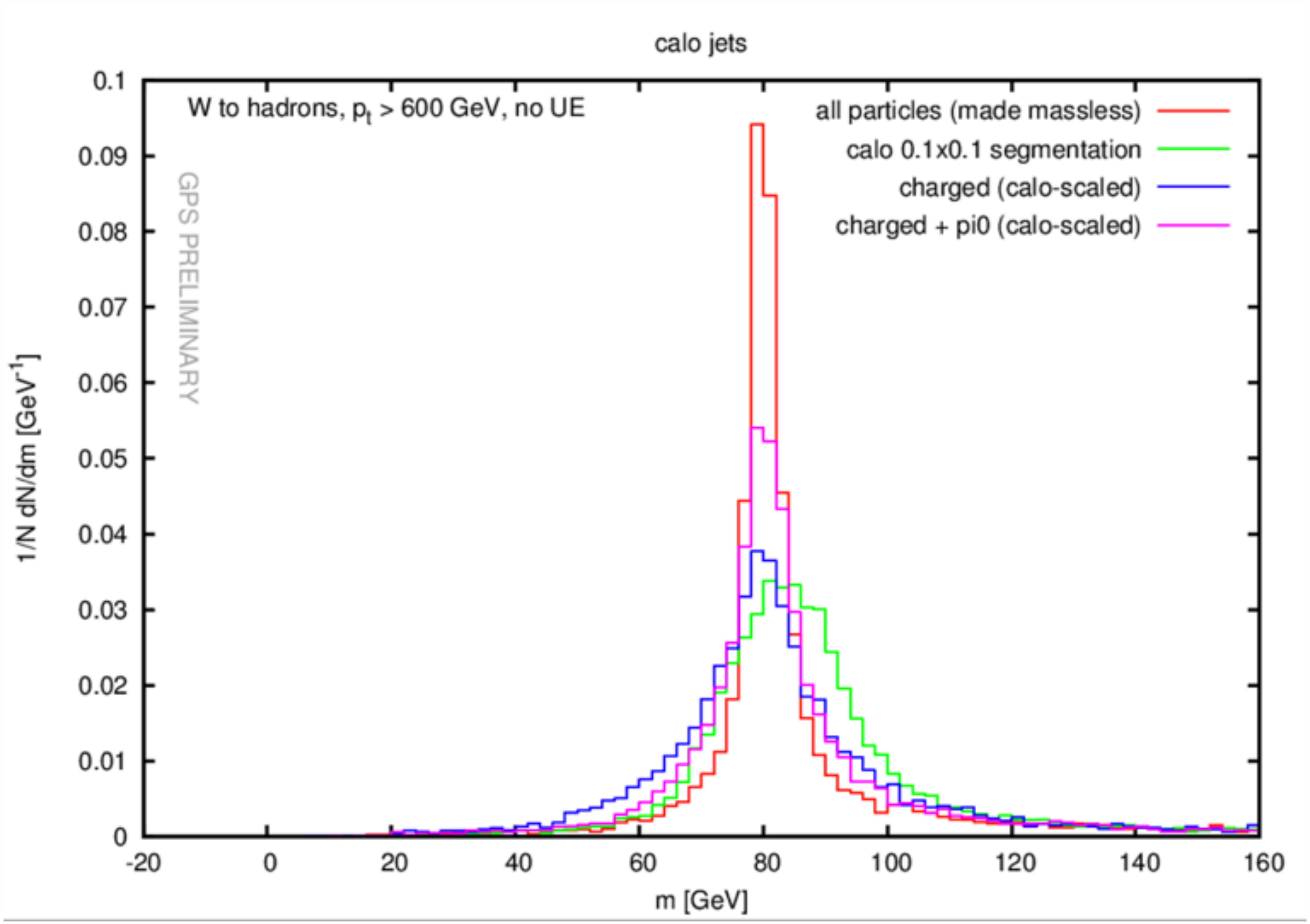
Detector resolution? Use Topocluster/PFlow directly. Or charged tracks + π^0 scaled to total jet energy?

cf. work by Tweedie et al



Detector resolution? Use Topocluster/PFlow directly. Or charged tracks + π^0 scaled to total jet energy?

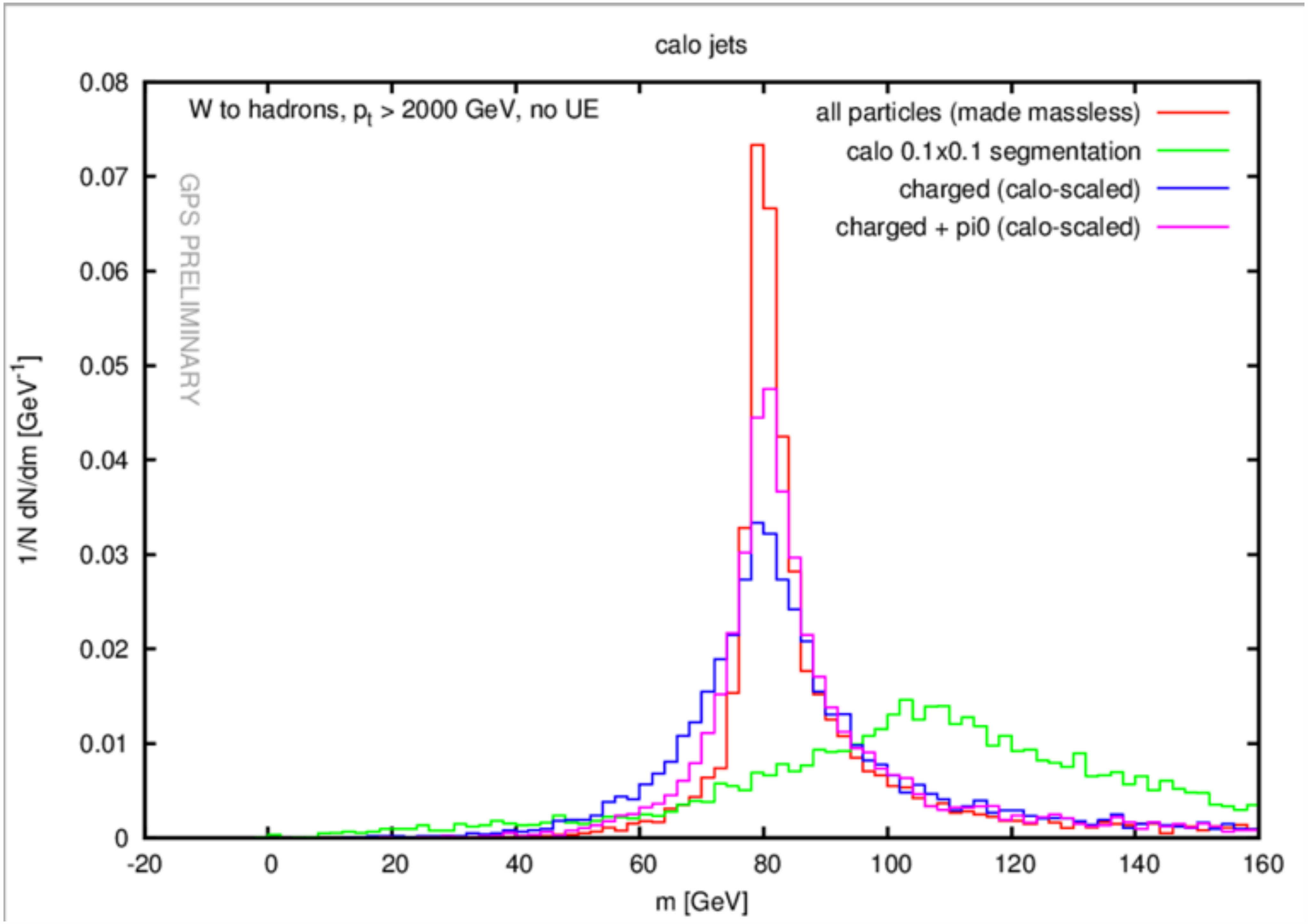
cf. work by Tweedie et al



illustrative "Calorimeter" only has granularity, no energy fluctuations

Detector resolution? Use Topocluster/PFlow directly. Or charged tracks + π^0 scaled to total jet energy?

cf. work by Tweedie et al



illustrative "Calorimeter" only has granularity, no energy fluctuations