Jet substructure: back to basics

Gavin Salam (CERN)

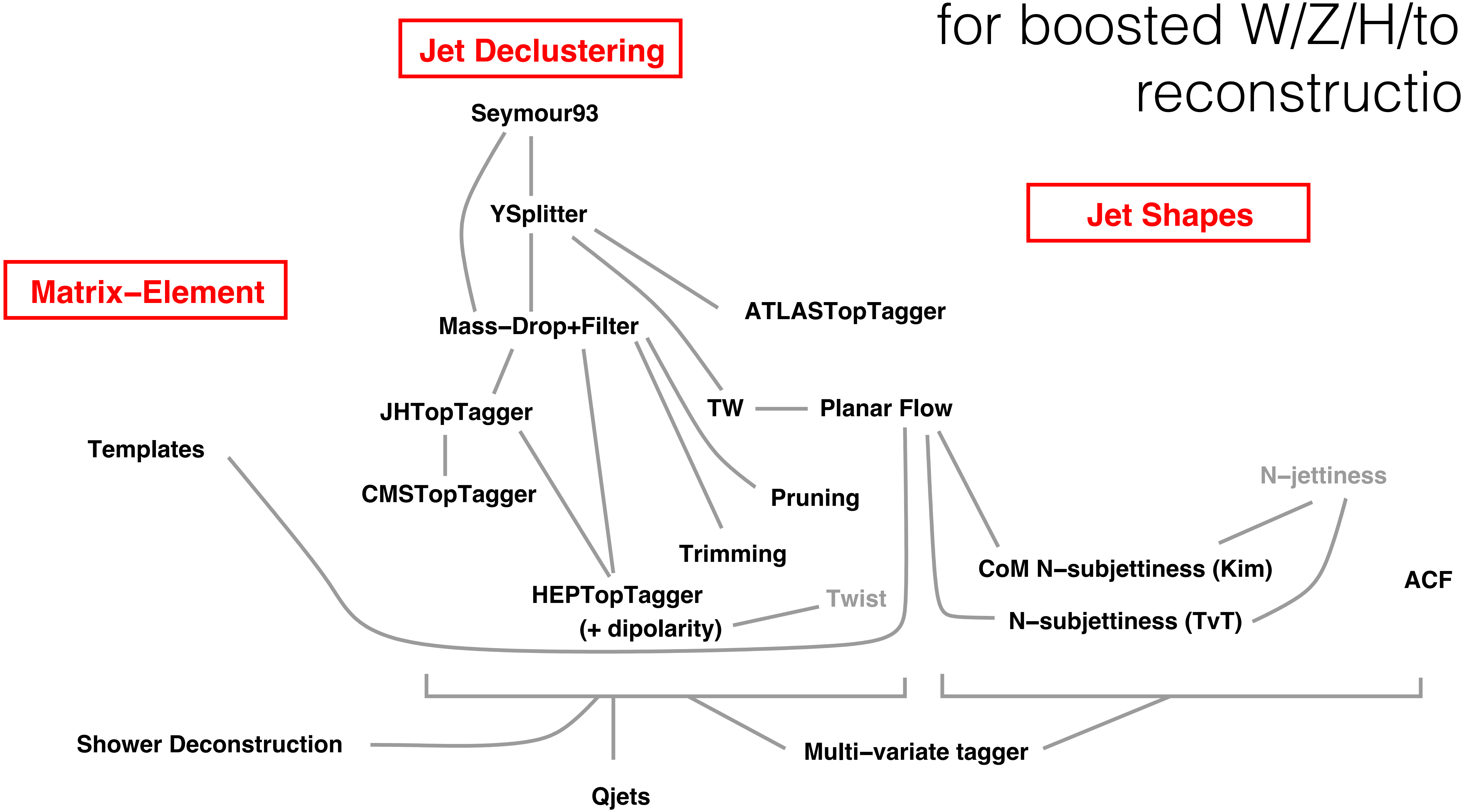
CMS Jet Substructure workshop
CERN, 15 April 2013

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

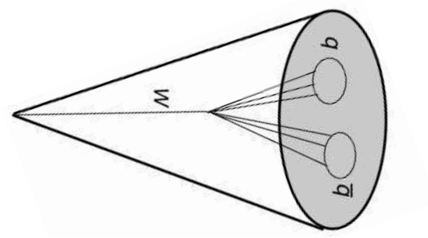
Some of the tools developed for boosted W/Z/H/top reconstruction

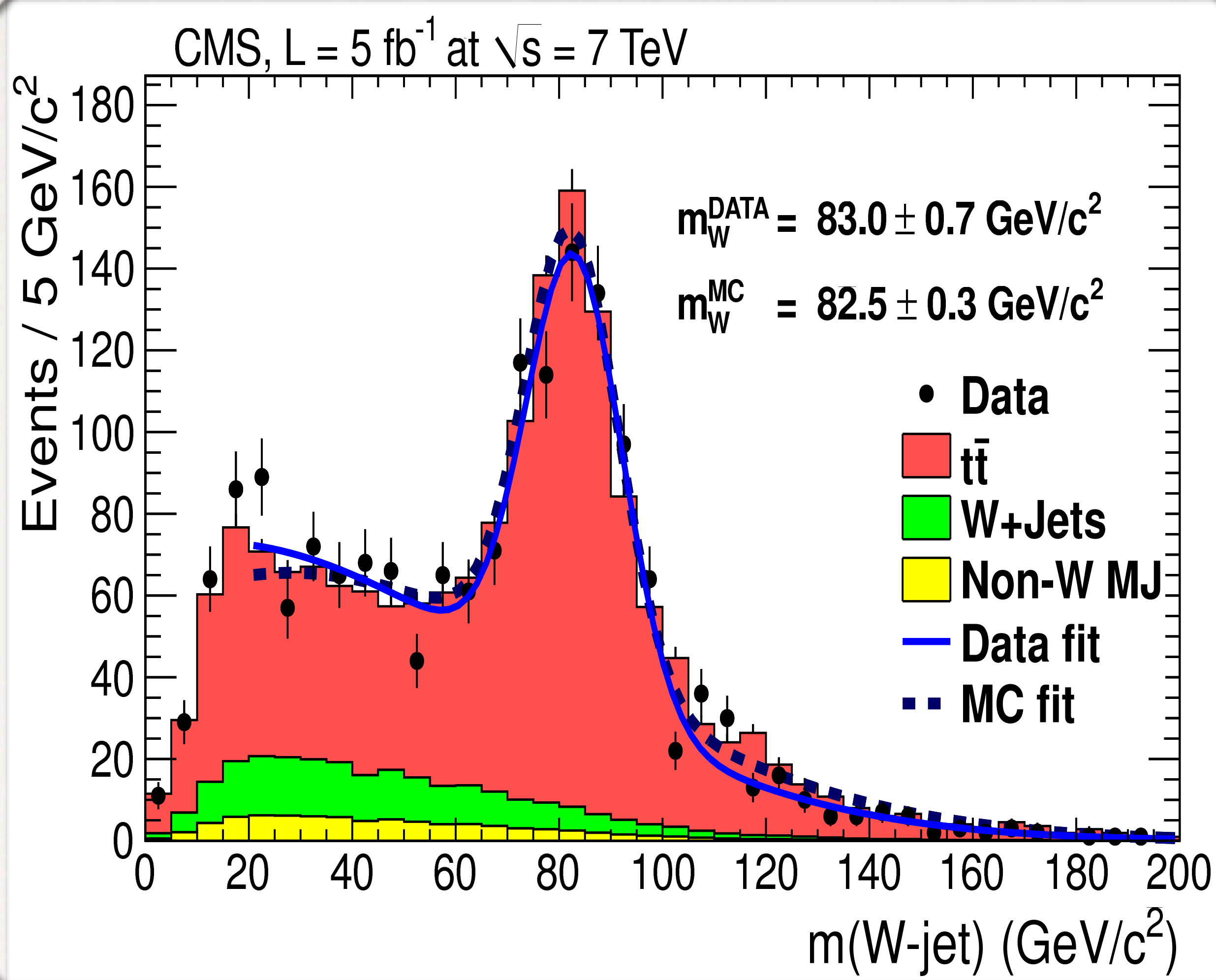


apologies for omitted taggers, arguable links, etc.

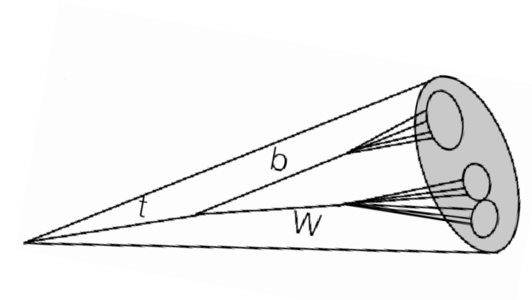
Seeing W's and tops in a single jet

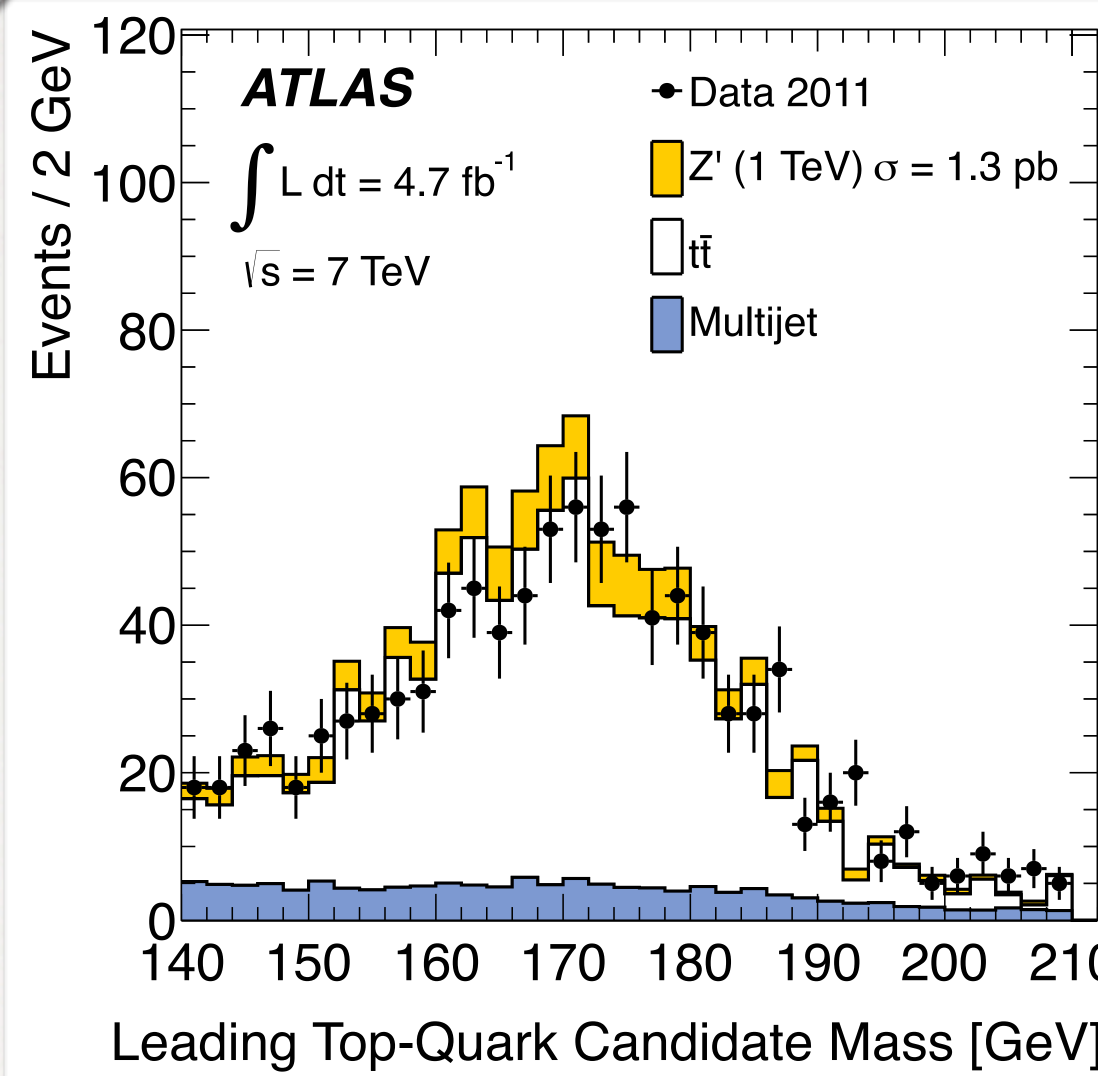
W's in a single jet





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?

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		**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...]

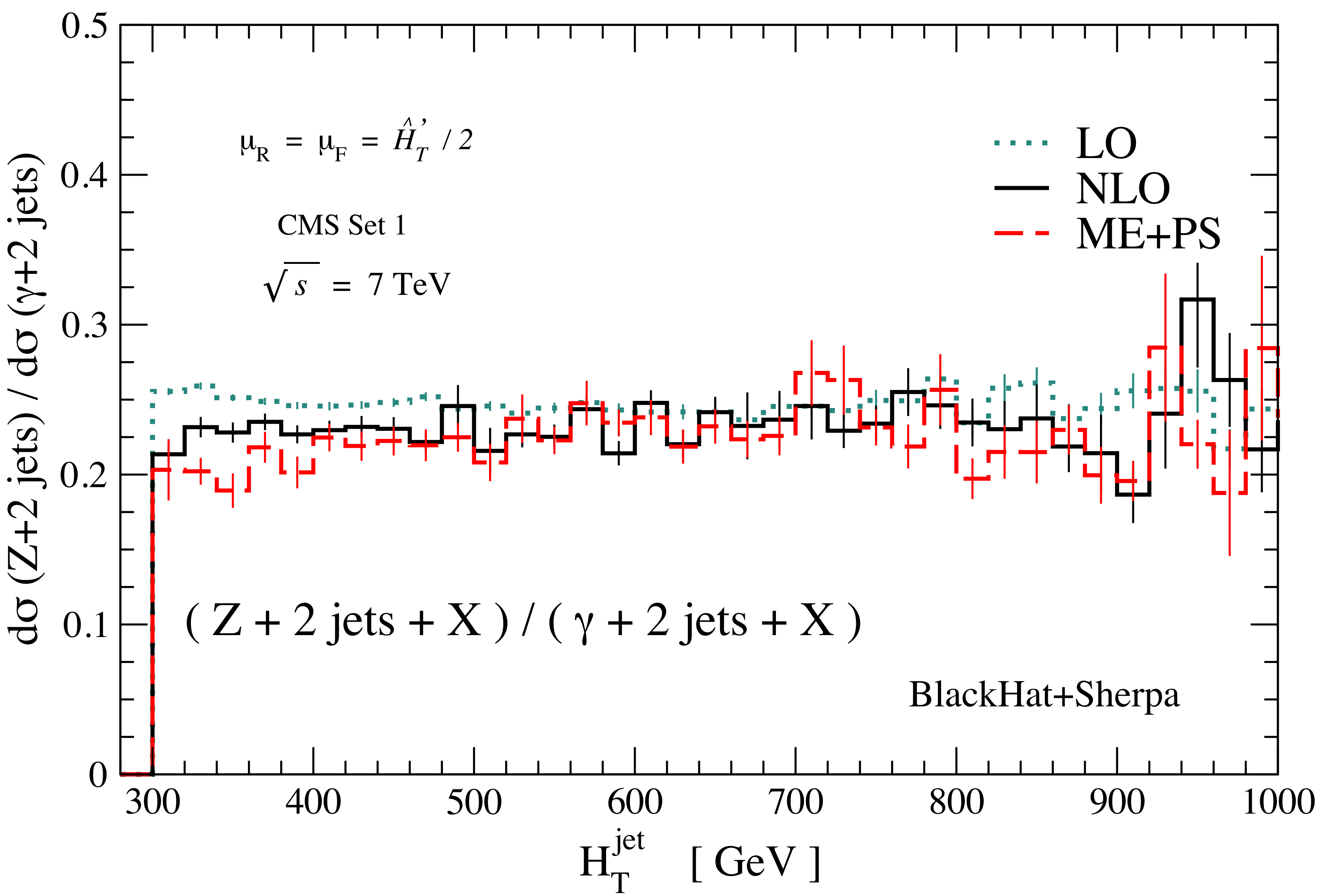
Standard, well understood

Less standard requires case-by-base validation

[Analytic] understanding

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($Z \to \nu\nu$) in multijets is reliably estimated from γ+jets because multiple types of calculations of the ratio agree

What do we know currently?

Boost 2010 proceedings:

The [Monte Carlo] findings discussed above indicate that while [pruning, trimming and filtering] have qualitatively similar effects, there are important differences. For our choice of parameters, pruning acts most aggressively on the signal and background followed by trimming and filtering.

At the time:

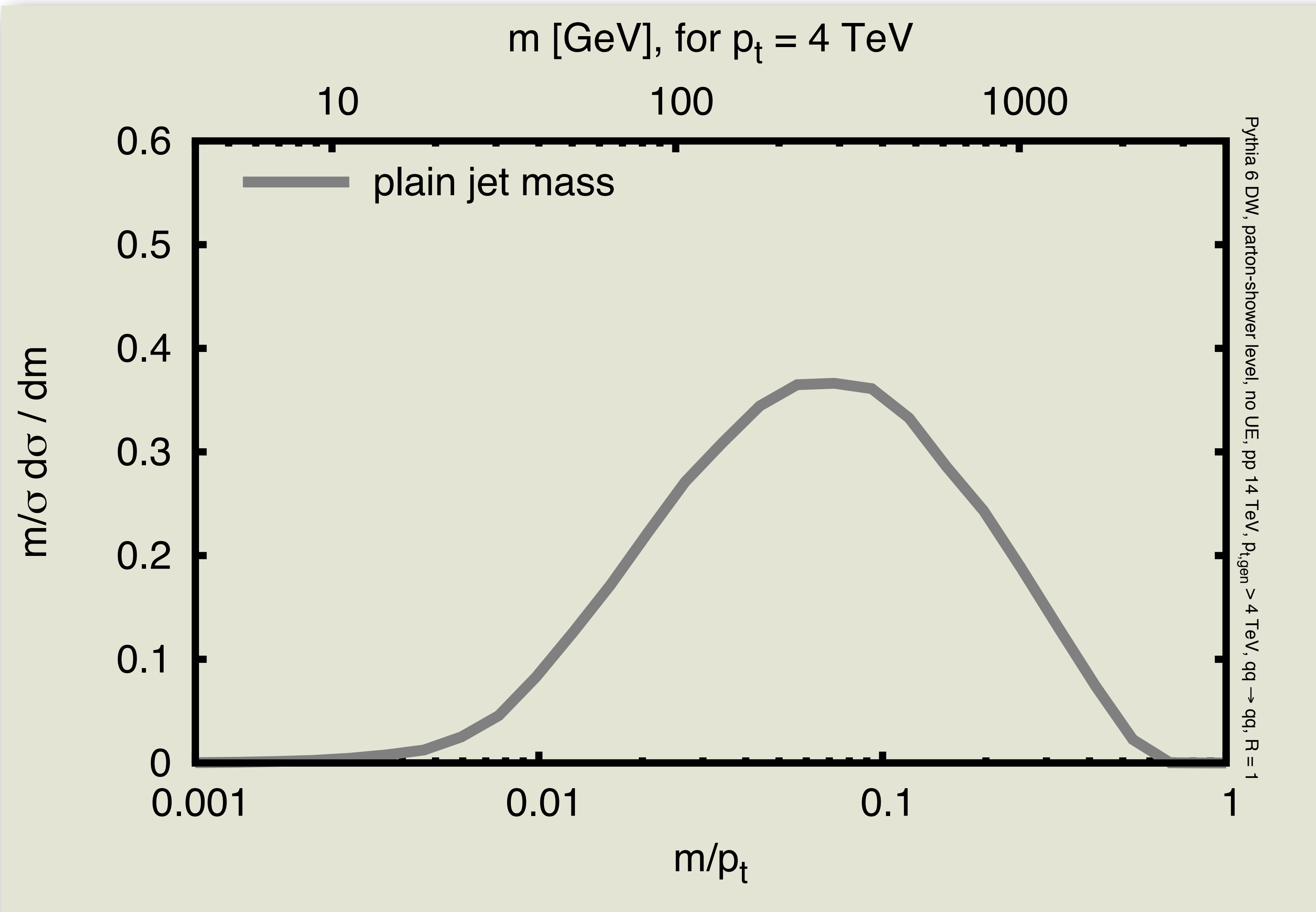
- No clear picture of why the taggers might be similar or different
- No clear picture of how the parameter choices affect the taggers

Today:

- I'll show a selection of **preliminary** lessons from studies for background jets in progress with Dasgupta, Fregoso and Marzani

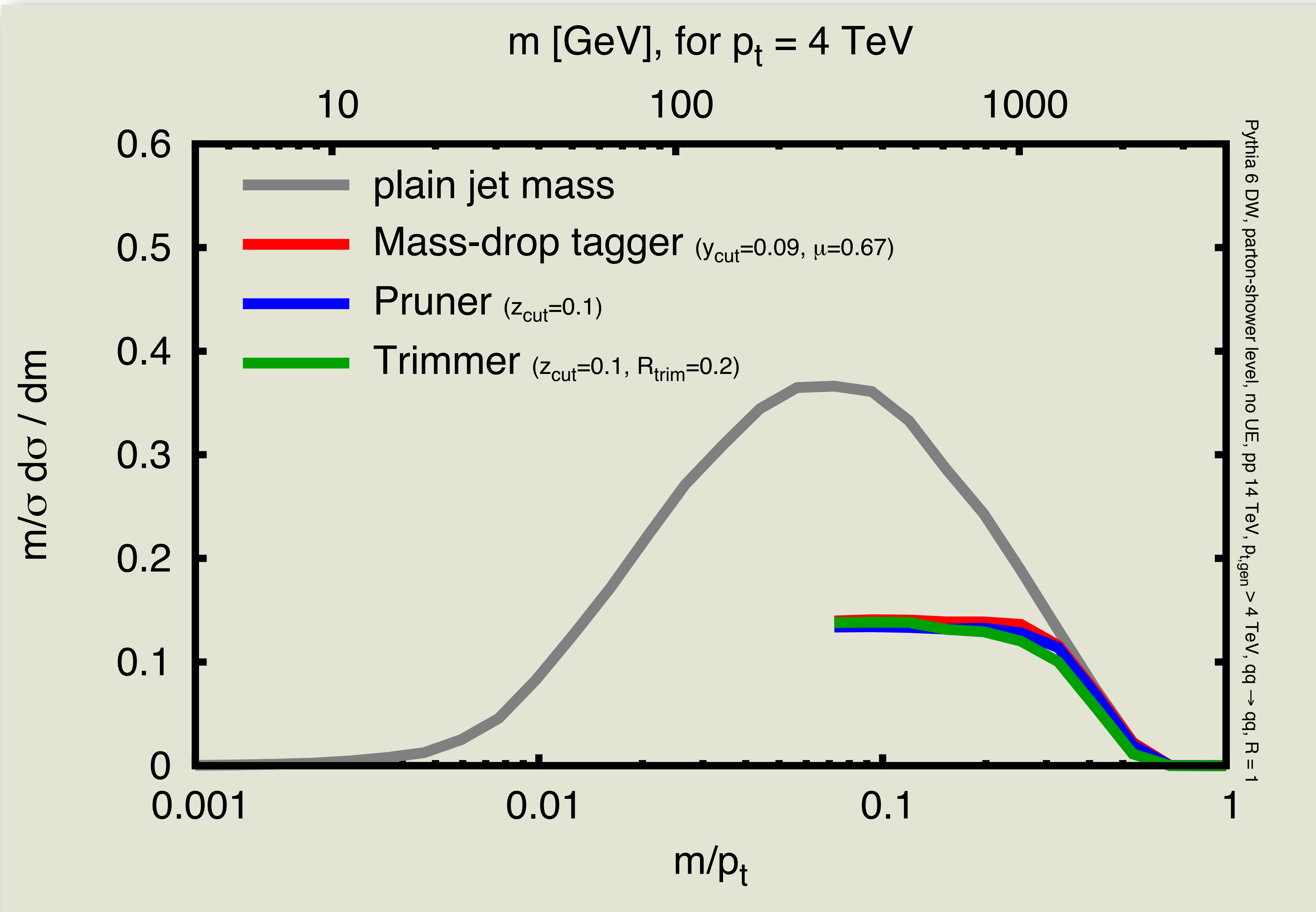
The "right" MC study can already be instructive

(testing on background [quark] jets)



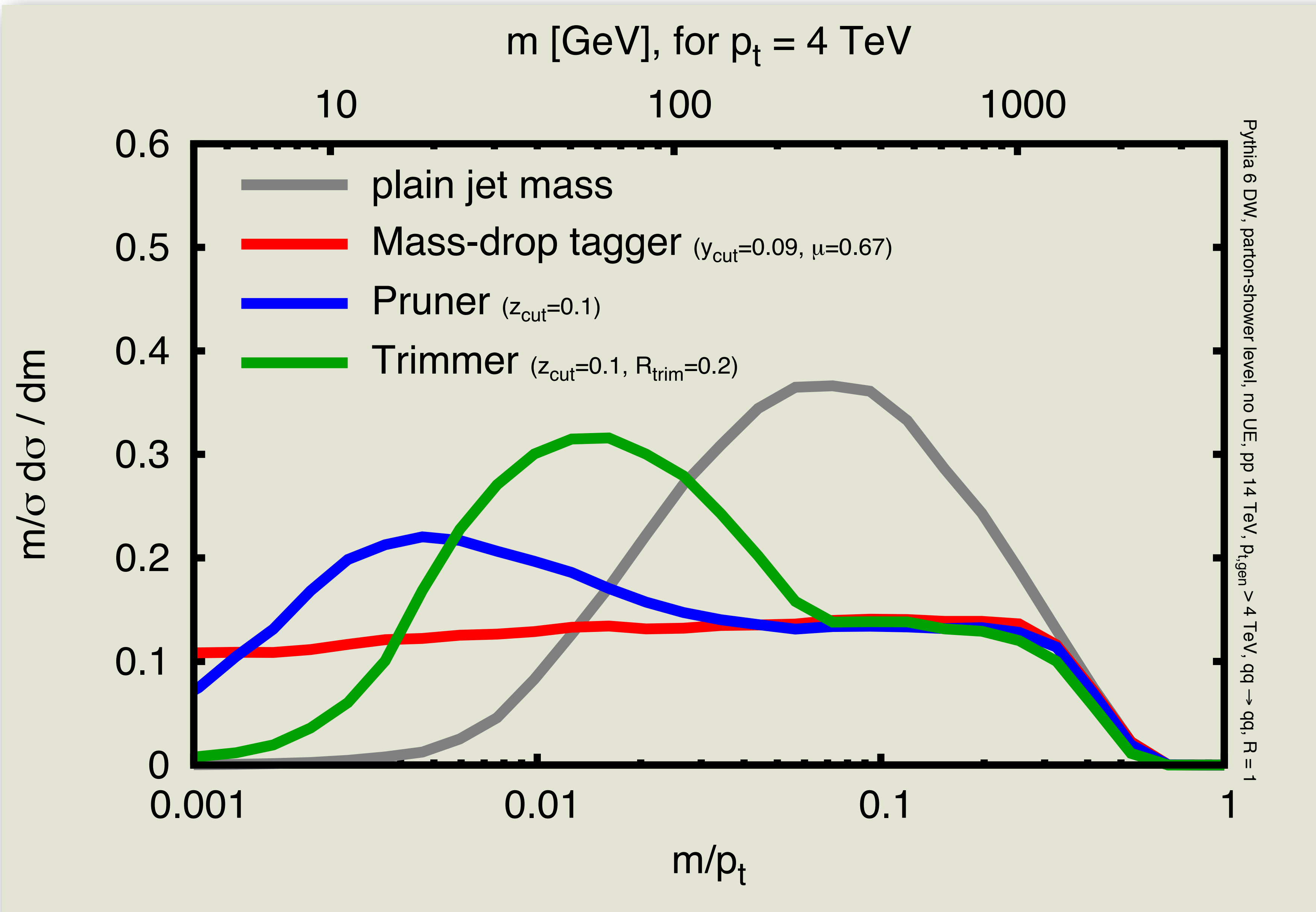
The "right" MC study can already be instructive

(testing on background [quark] jets)



Different taggers are apparently quite similar

The "right" MC study can already be instructive
(testing on background [quark] jets)

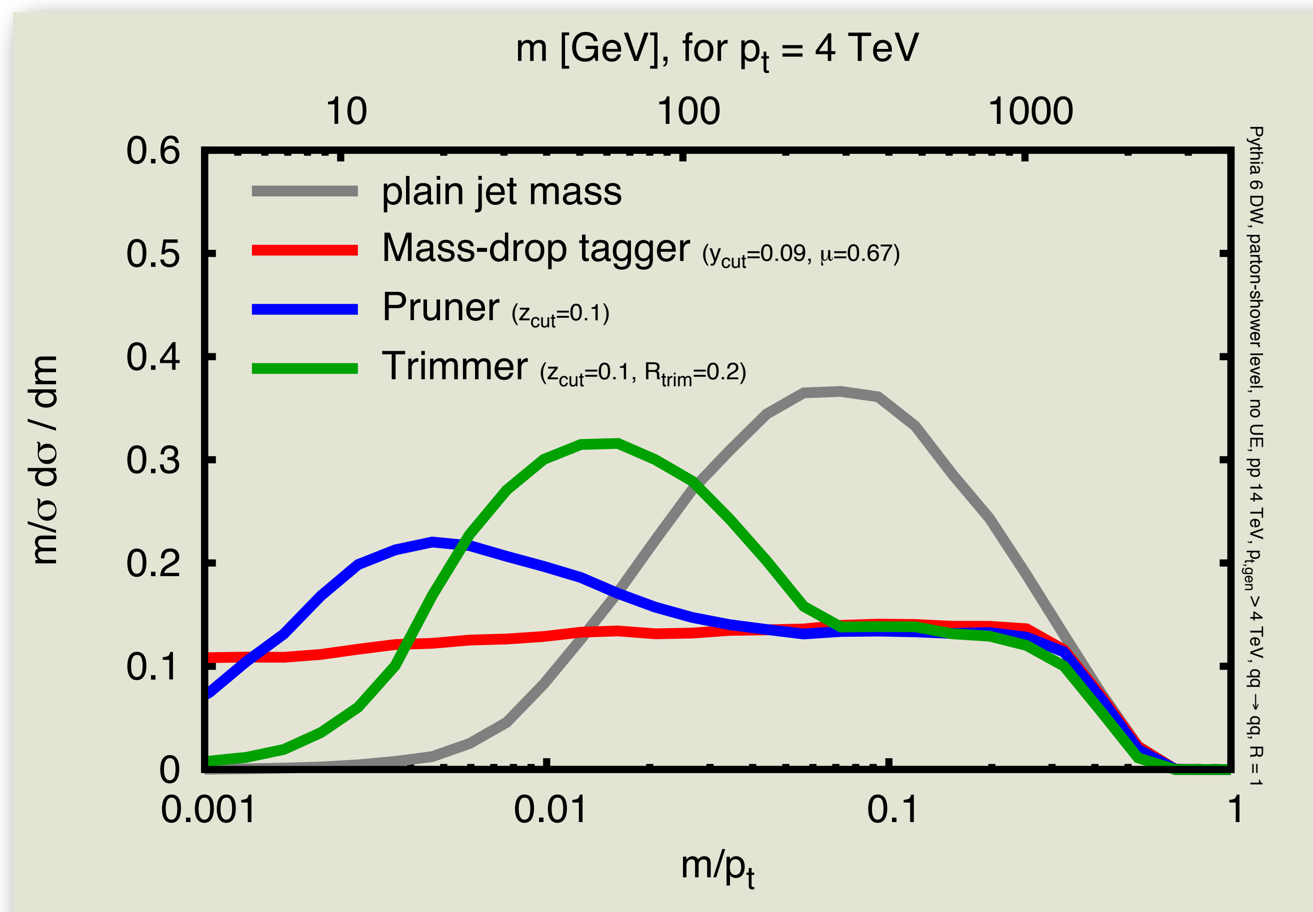


But only for a limited range of masses

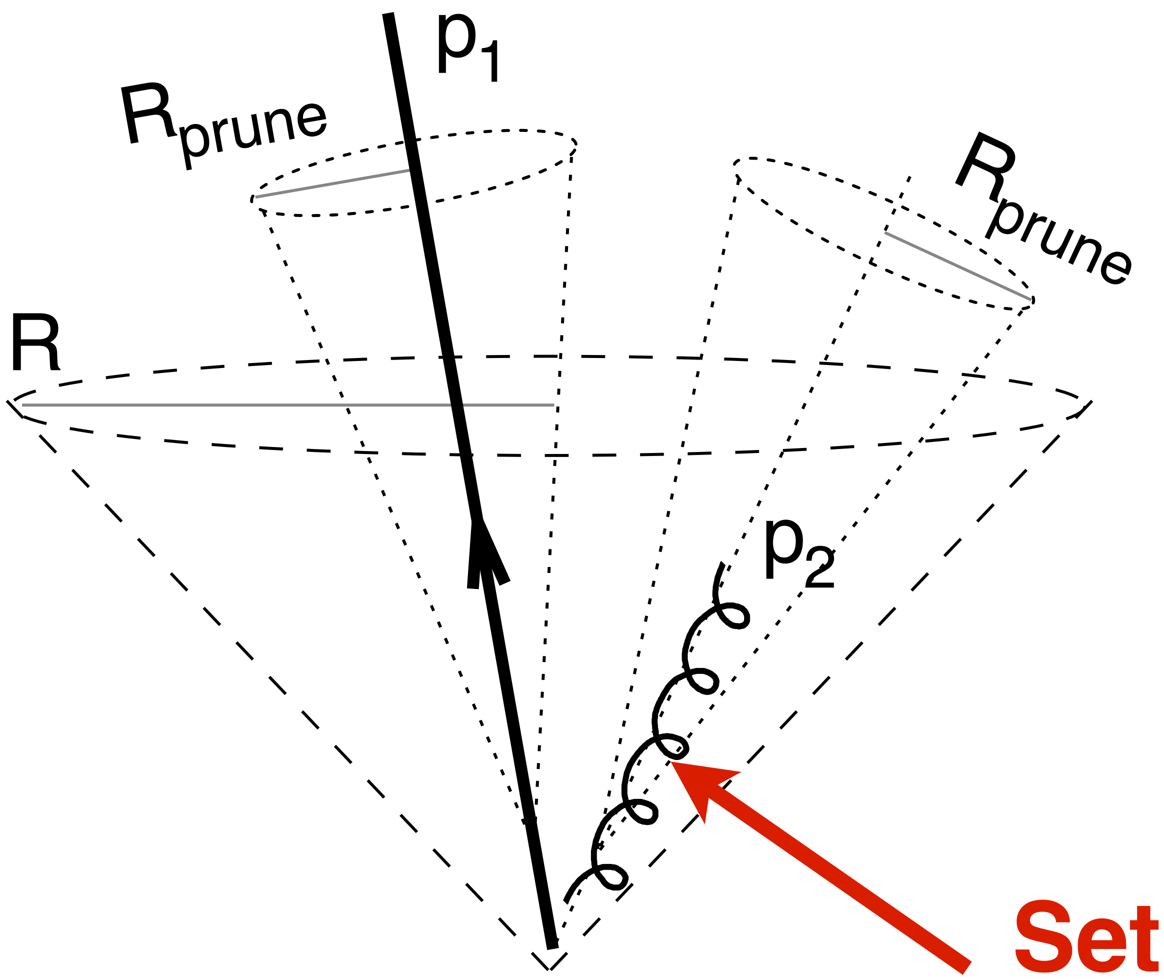
Do we care about such differences?

- Think data-driven backgrounds: kinks can seriously mess with you [especially if you got used to their being absent, e.g. from moderate p_t tests]
- How do these structures depend on the z_{cut}, y_{cut}, R_{trim}, etc. parameters?
- Are these structures telling us something we might want to know about the taggers? E.g. how to improve them?

This calls for analysis and calculation



Pruning [7, 8] takes an initial jet, and from its mass deduces a pruning radius $R_{\text{prune}} = R_{\text{fact}} \cdot \frac{2m}{p_t}$, where R_{fact} is a parameter of the tagger. It then reclusters the jet and for every clustering step, involving objects a and b, it checks whether $\Delta_{ab} > R_{\text{prune}}$ and $\min(p_{ta}, p_{tb}) < z_{\text{cut}} p_{t,(a+b)}$, where z_{cut} is a second parameter of the tagger. If so, then the softer of the a and b is discarded. Otherwise a and b are recombined as usual. Clustering then proceeds with the remaining objects, applying the pruning check at each stage.



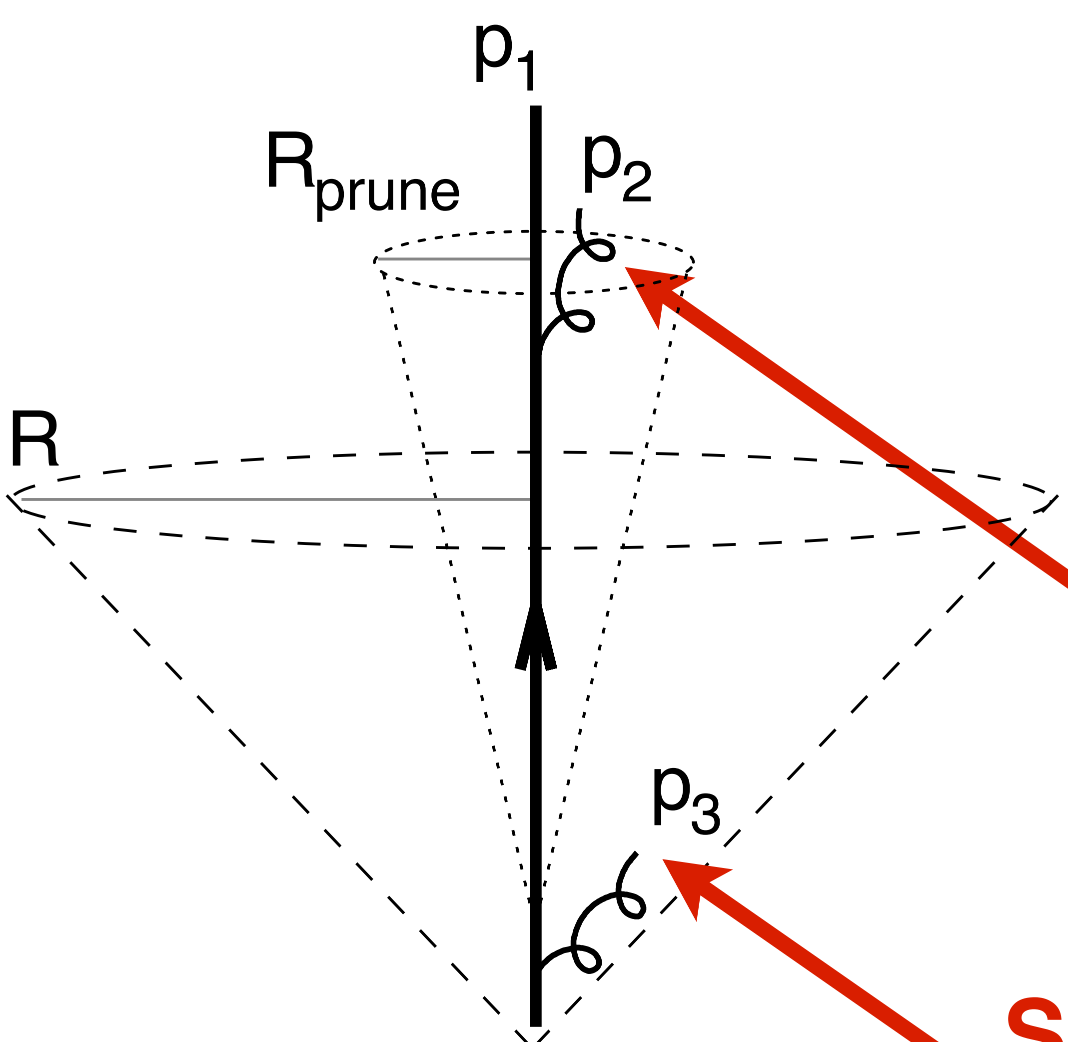
What pruning is meant to do:

Choose an R_{prune} such that different hard prongs (p_1, p_2) end up in different hard subjets.

Discard any softer radiation.

Sets pruning radius, & hard enough to end up as subjet

Pruning [7, 8] takes an initial jet, and from its mass deduces a pruning radius $R_{\text{prune}} = R_{\text{fact}} \cdot \frac{2m}{p_t}$, where R_{fact} is a parameter of the tagger. It then reclusters the jet and for every clustering step, involving objects a and b, it checks whether $\Delta_{ab} > R_{\text{prune}}$ and $\min(p_{ta}, p_{tb}) < z_{\text{cut}} p_{t,(a+b)}$, where z_{cut} is a second parameter of the tagger. If so, then the softer of the a and b is discarded. Otherwise a and b are recombined as usual. Clustering then proceeds with the remaining objects, applying the pruning check at each stage.



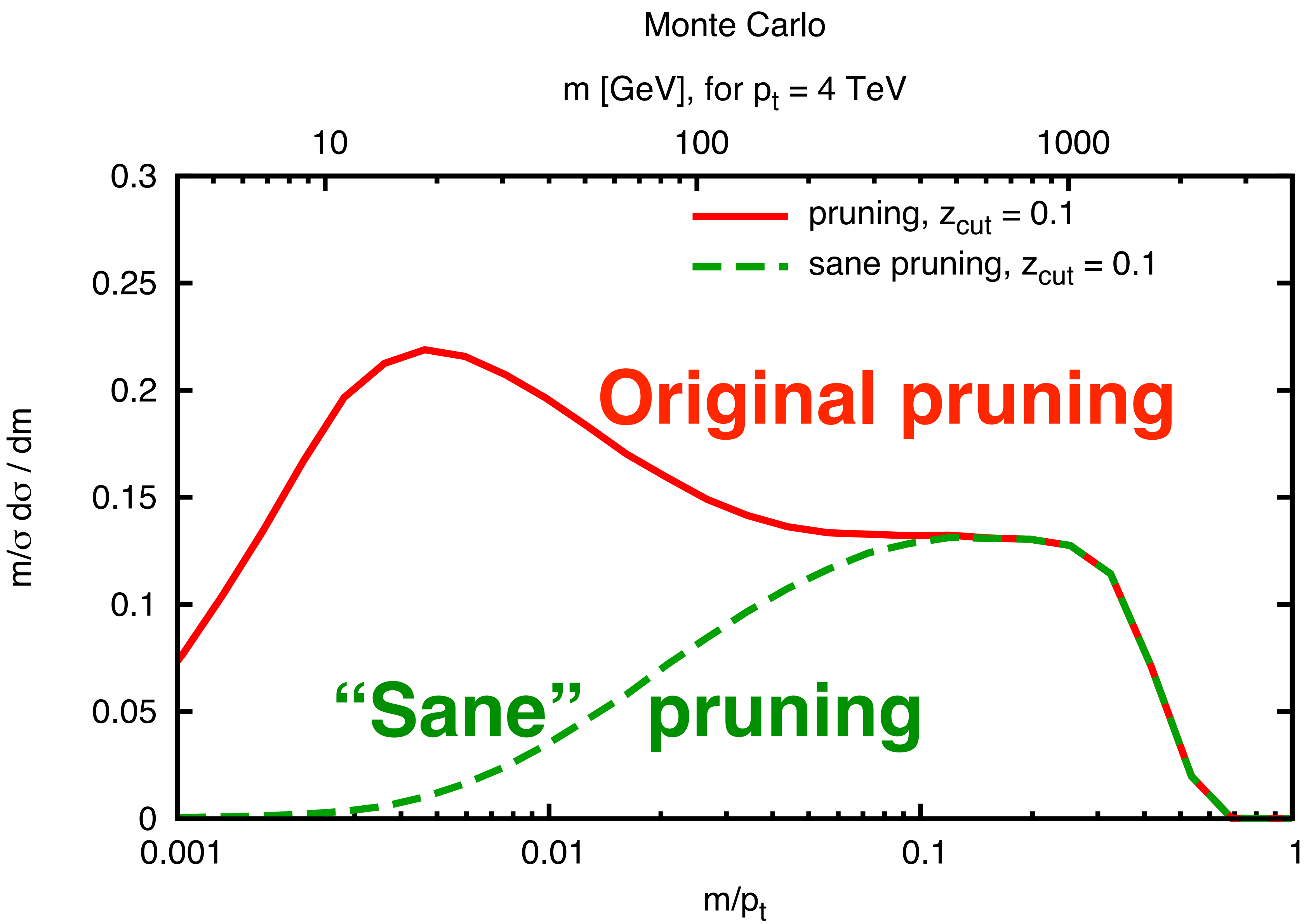
What pruning sometimes does

Chooses R_{prune} based on a soft p_3 (dominates total jet mass), and leads to a single narrow subjet whose mass is also dominated by a soft emission (p_2, within R_{prune} of p_1, so not pruned away).

Sets pruning radius, but gets pruned away

A simple fix: "sane" pruning

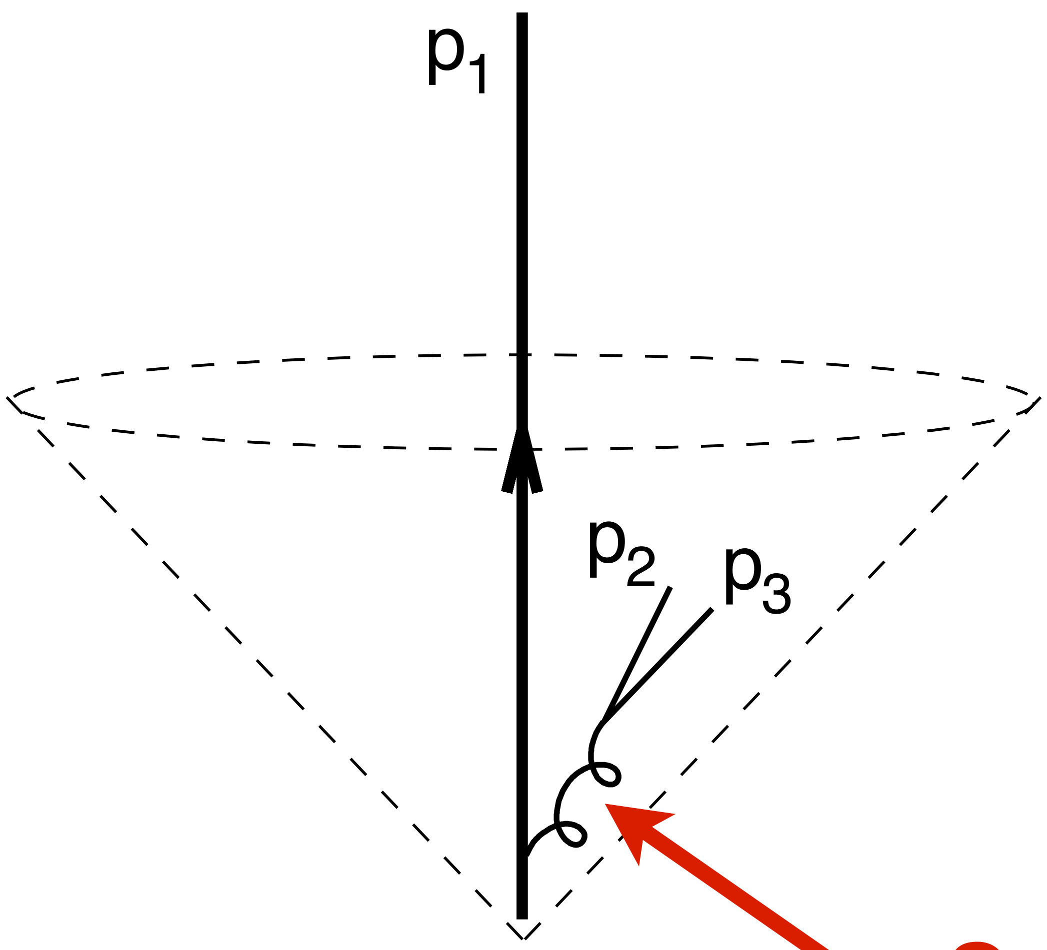
Require at least one successful merging with $\Delta R > R_{prune}$ and $z > z_{cut}$



"sane" pruning is effectively placing an isolation cut on radiation around the tagged (colour-neutral) object

Mass drop tagger

1. Break the jet j into two subjets by undoing its last stage of clustering. Label the two subjets j_1, j_2 such that $m_{j_1} > m_{j_2}$.
2. If there was a significant mass drop, $m_{j_1} < \mu m_j$, and the splitting is not too asymmetric, $y = \min(p_{tj_1}^2, p_{tj_2}^2)\Delta R_{j_1 j_2}^2 / m_j^2 > y_{\rm cut}$, then deem j to be the tagged jet.
3. Otherwise redefine j to be equal to j_1 and go back to step 1 (unless j consists of just a single particle, in which case the original jet is deemed untagged).



What MDT does wrong:

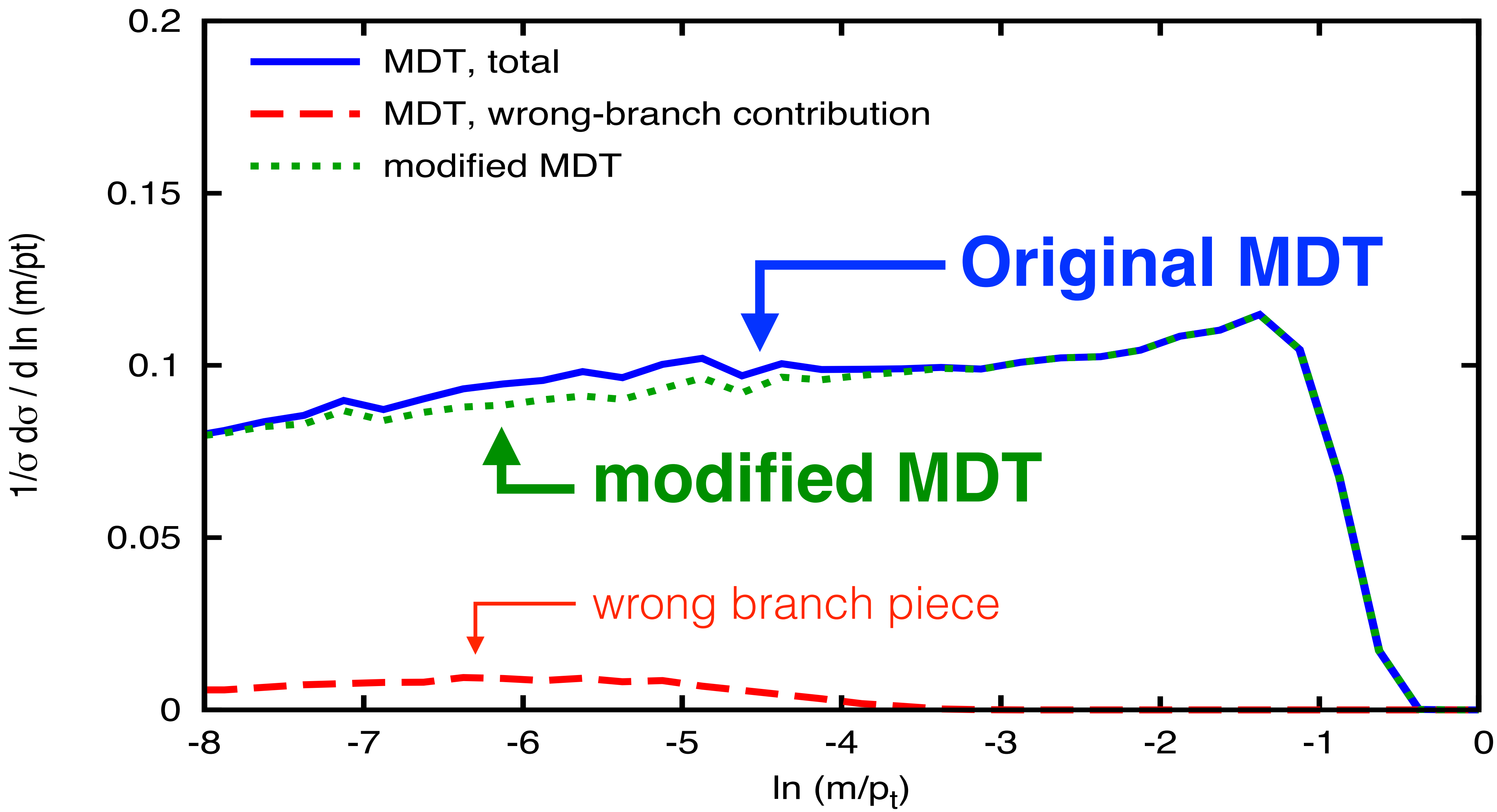
Can follow a soft branch ($p_2+p_3 < y_{cut}\, p_{jet}$) with "accidental" small mass, when the "right" answer was that the (massless) hard branch had no substructure

Subjet is soft, but has more substructure than hard subjet

A simple fix for "modified" Mass Drop Tagger:

When recursing, follow branch with larger ($m^2+p_t^2$)

(rather than the one with larger m)



What about analytic calculations of the taggers?

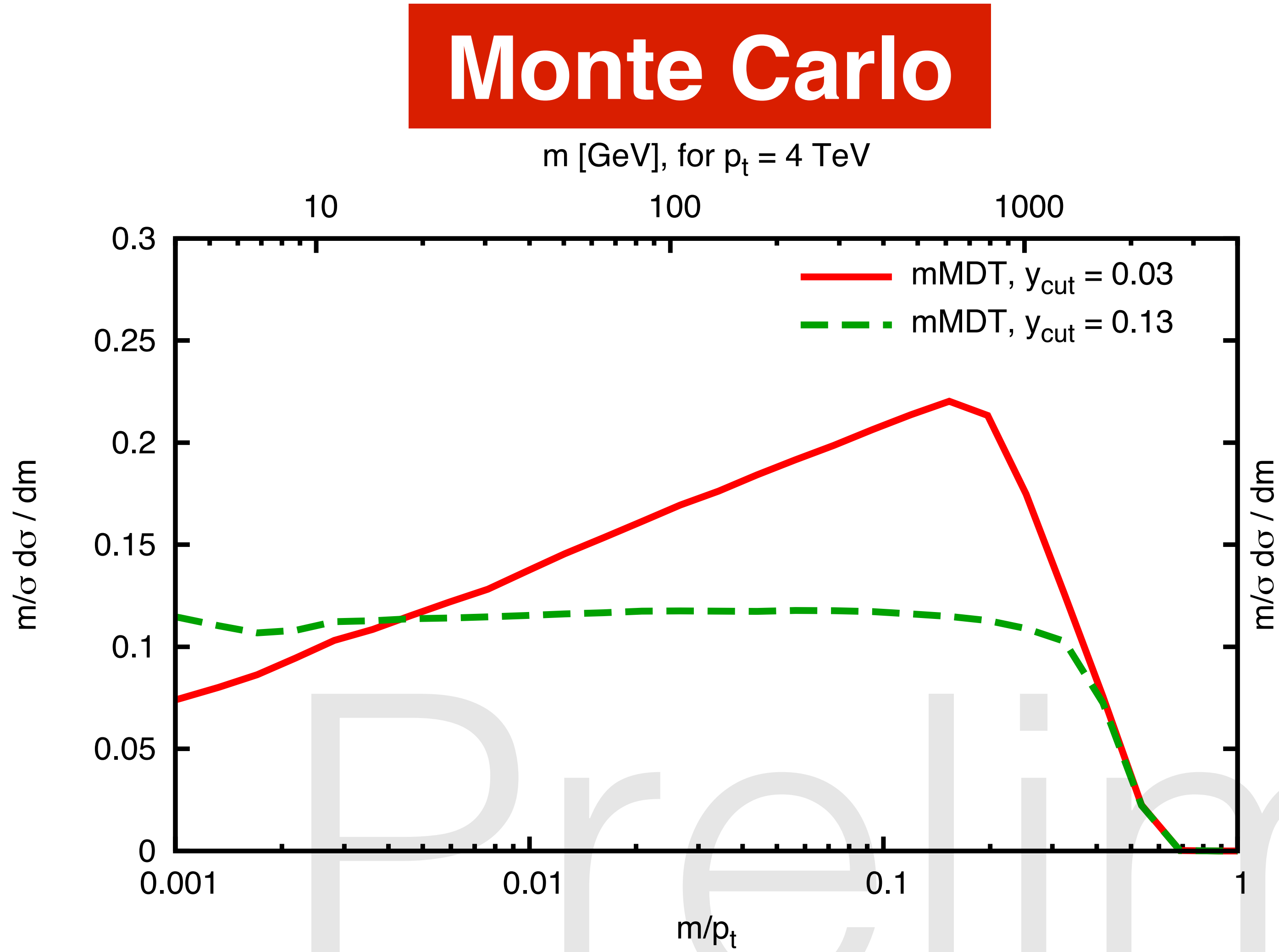
Simpler → More complex

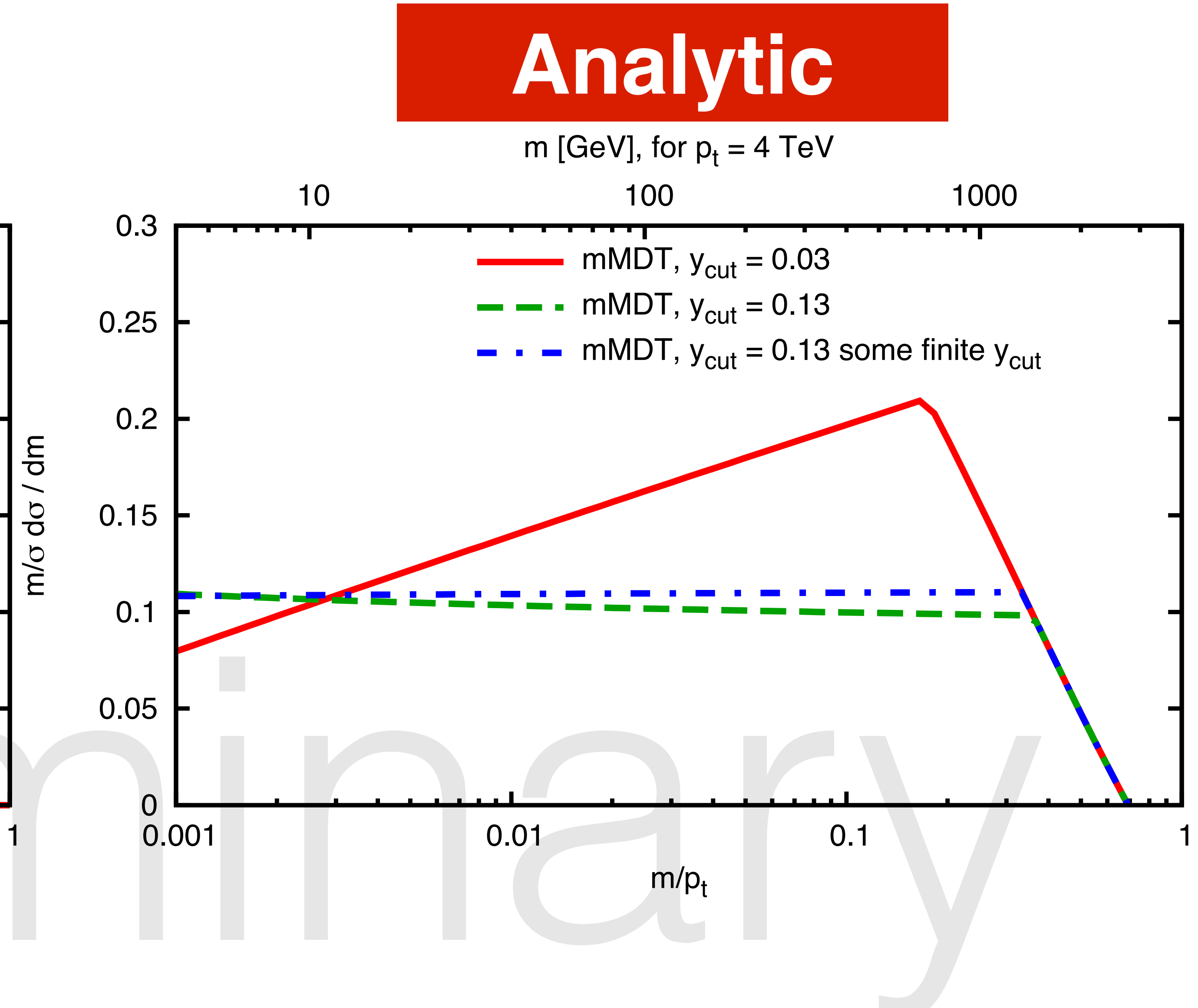
Dasgupta et al (full NLL, including non-global part)
Chien et al (partial NNLL)
Jouttenus et al (partial NNLL / exclusive jet events)

mMDT > plain mass ~ trimming > sane pruning > full pruning » ~~MDT~~

Dasgupta, Fregoso, Marzani & GPS, forthcoming:
LL in all cases, plus some subleading logs
[NB: LL doesn't mean the same thing in all cases!)

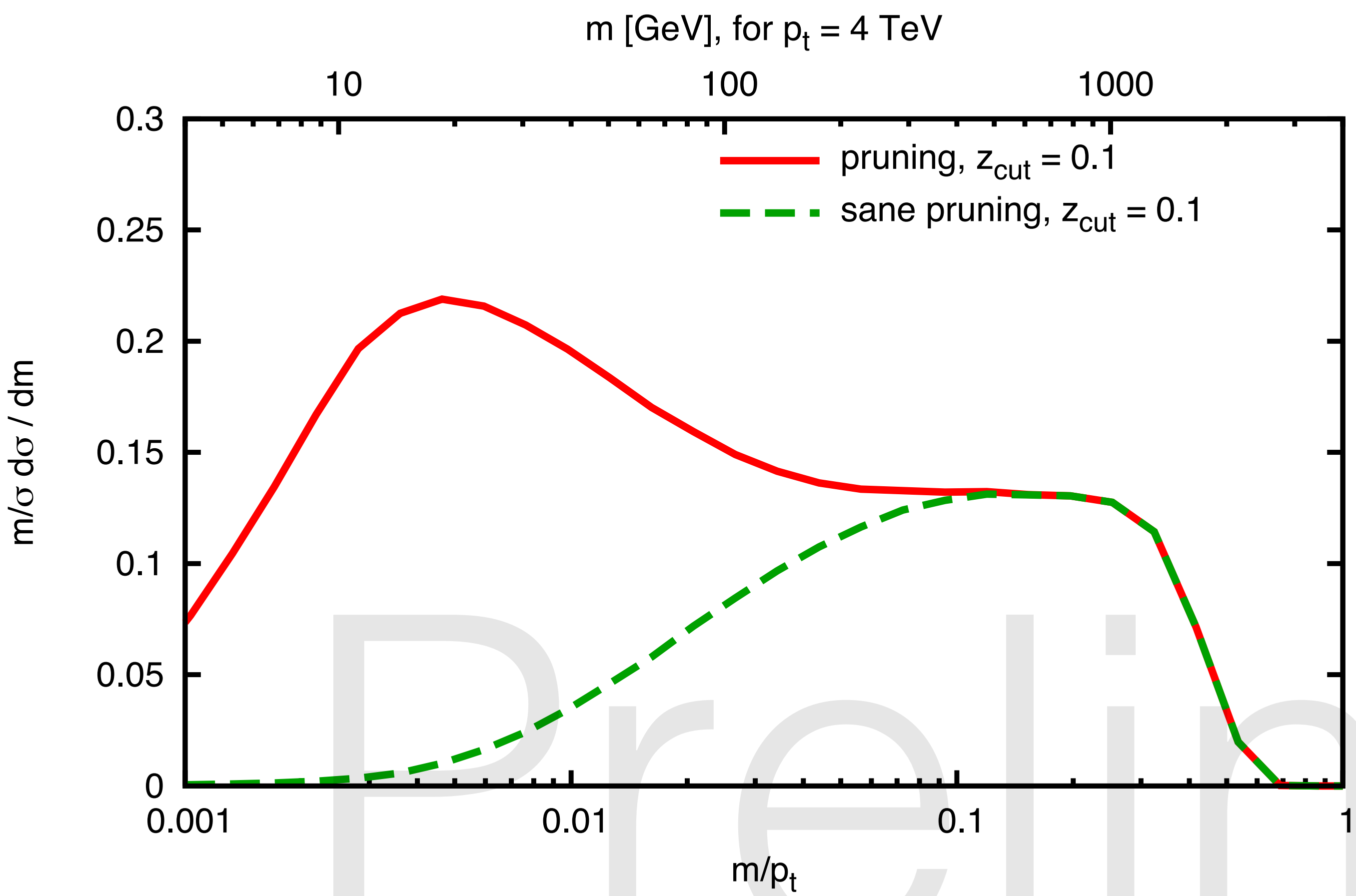
Modified Mass Drop Tagger



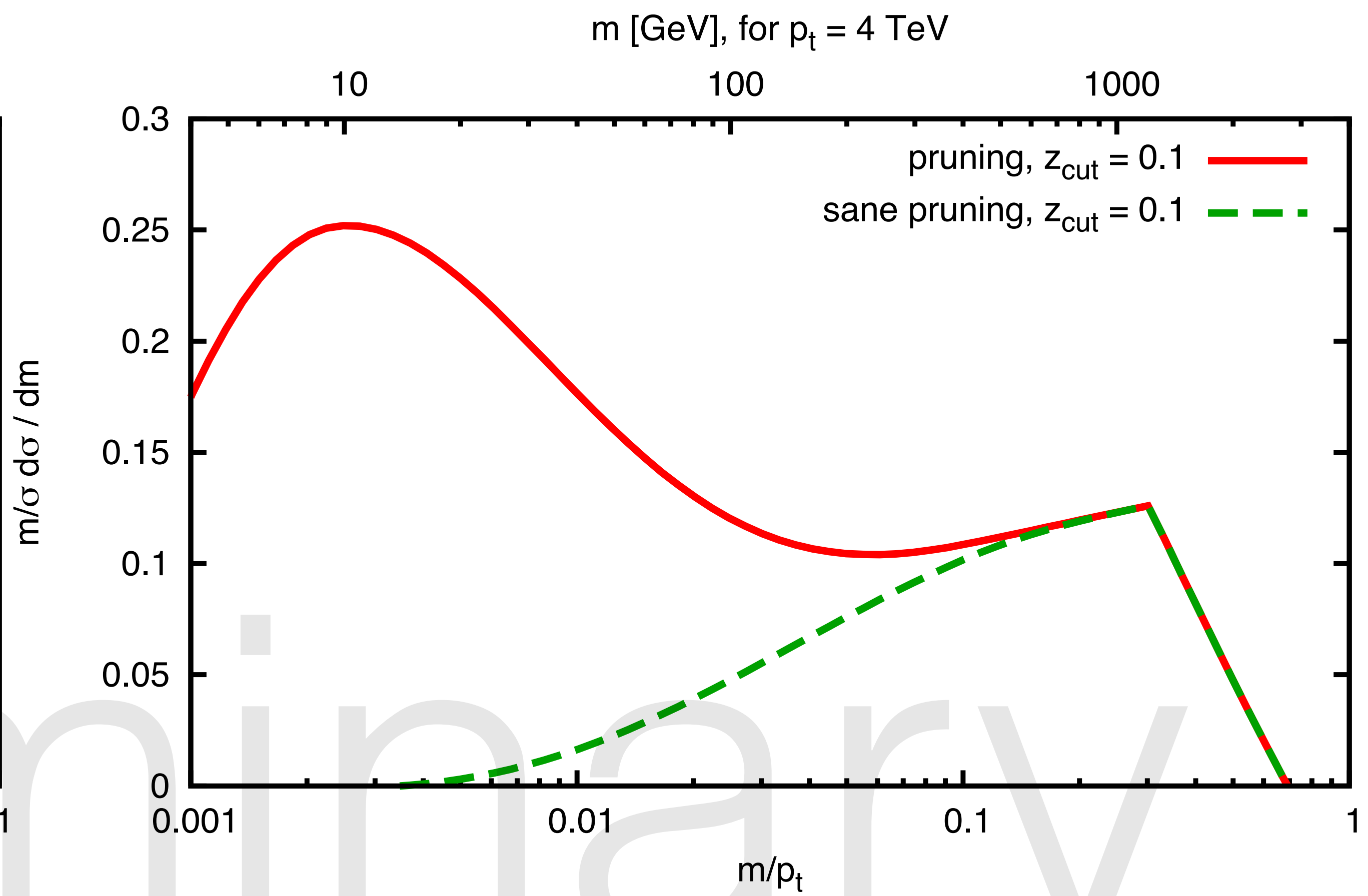


[mMDT is closest we have to a scale-invariant tagger, though exact behaviour depends on q/g fractions]

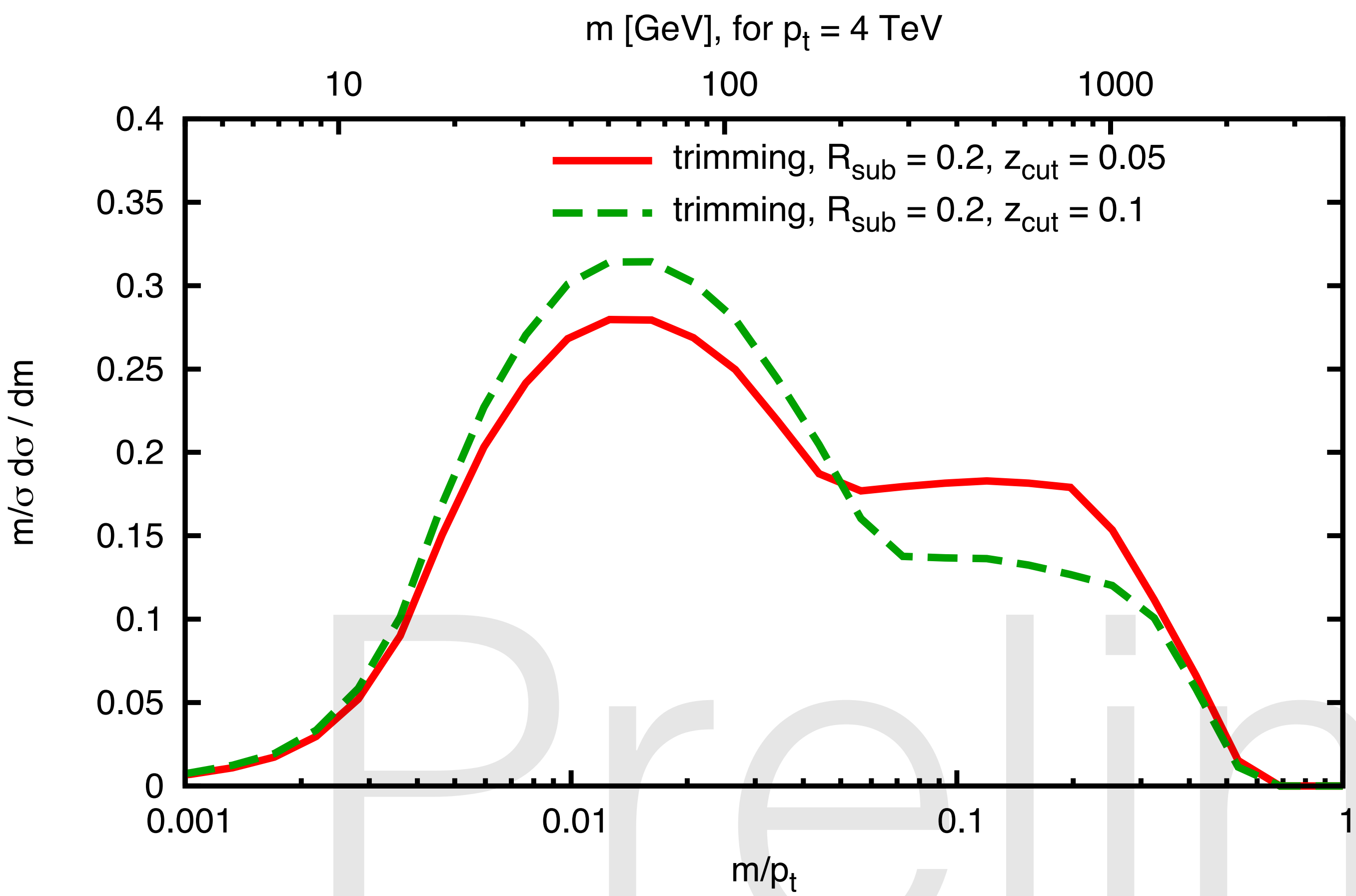
Monte Carlo



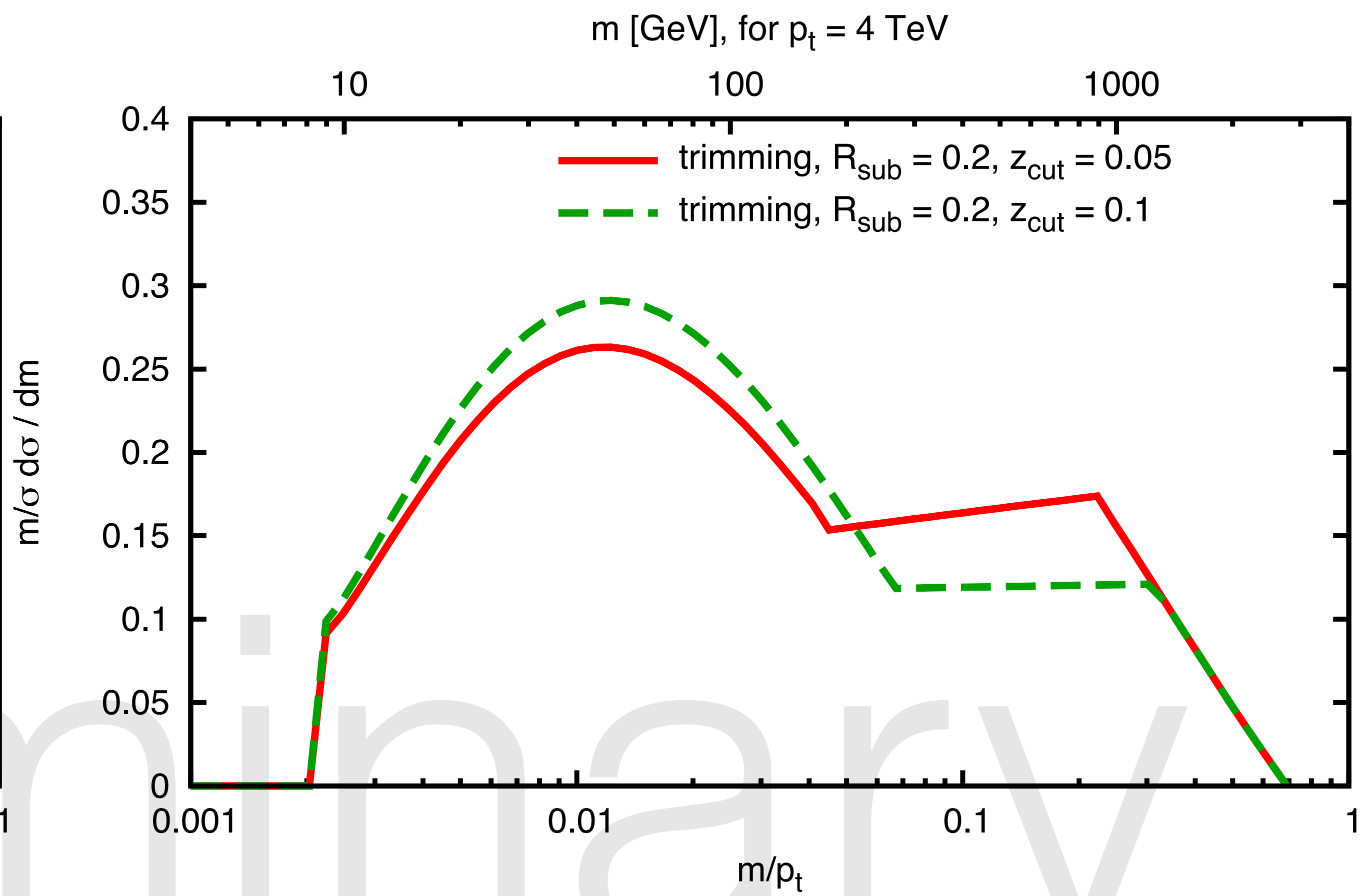
Analytic



Monte Carlo

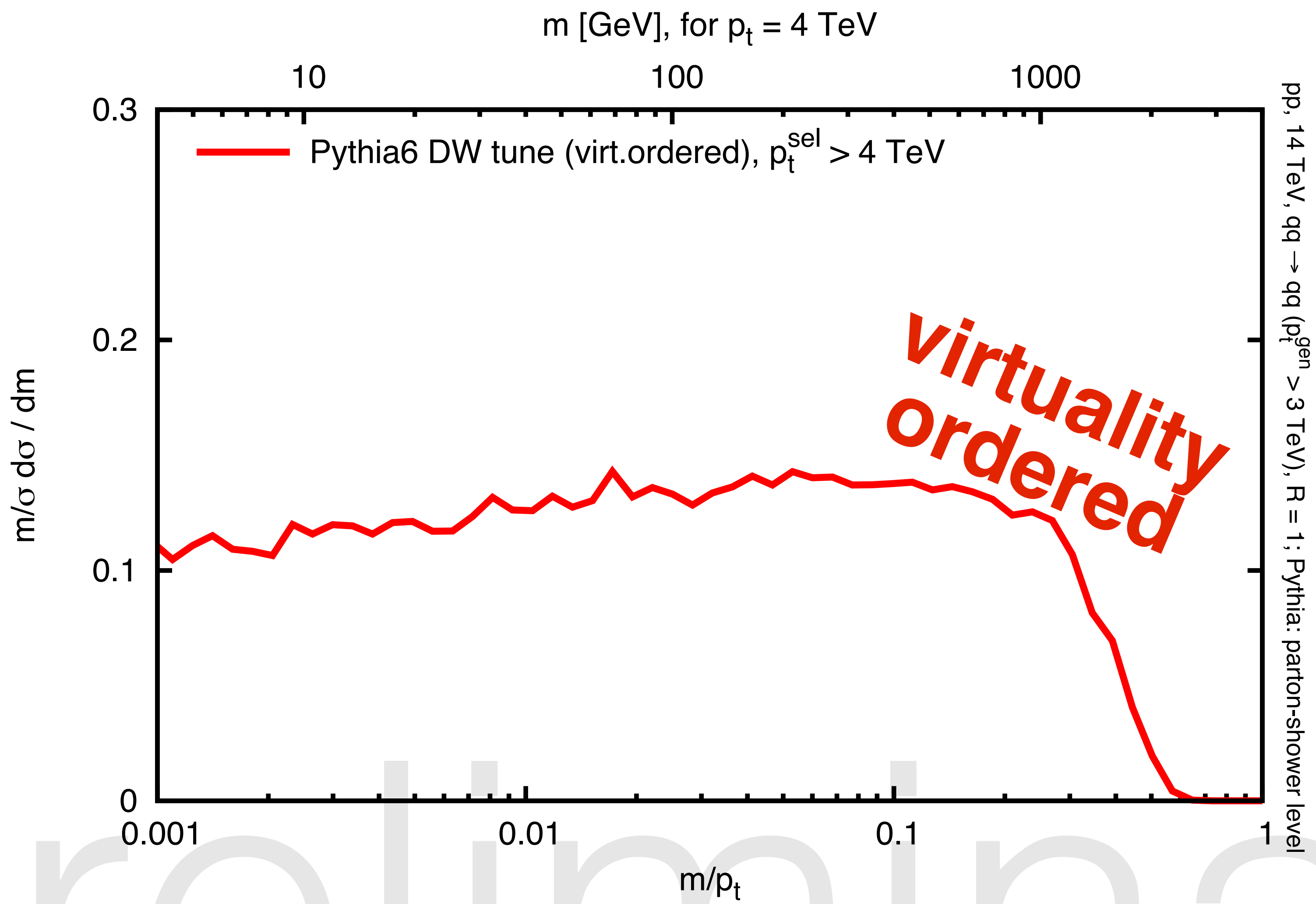


Analytic



Why is this useful?

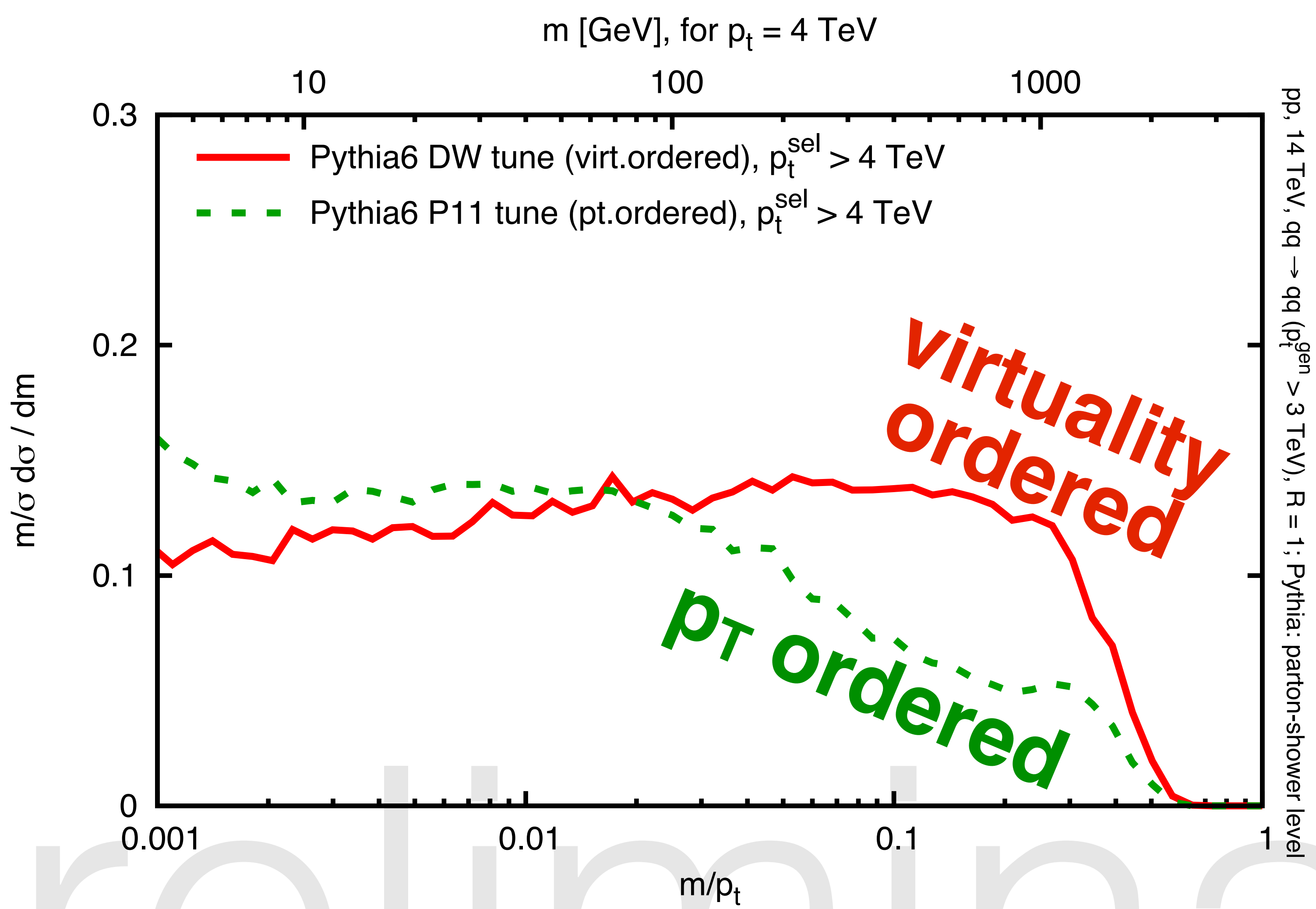
If MC & analytics always agree, why bother with calculations?



Preliminary

Why is this useful?

Different MC showers don't always agree.
That turns into a systematic uncertainty.

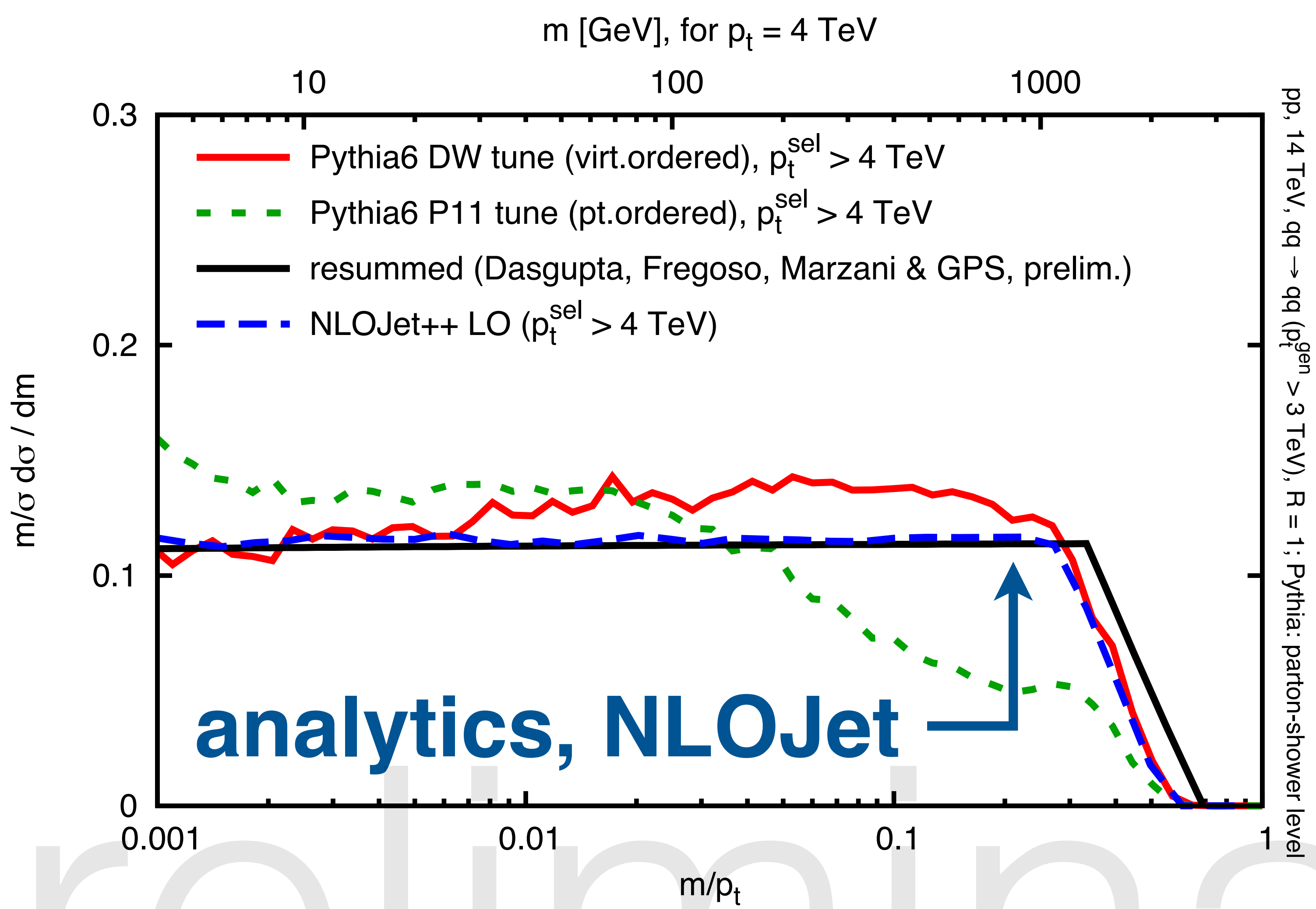


Preliminary

Why is this useful?

Different MC showers don't always agree.
That turns into a systematic uncertainty.

Analytics can tell you which is "right".



Bottom line on "understanding"

- Taggers may be quite simple to write, but potentially involved to understand – tiny details can lead even courageous theorists to tears.
- Contrast this with p_t cuts for standard jet analyses – (mostly) simple
- Still, many taggers/groomers are within calculational reach.
- New "modified" Mass Drop Tagger is especially simple; New "sane" pruning is also interesting
- The better you understand a tagger, the better you can detect signals

Infrared safety

Infrared safety:

When the addition of one soft particle with momentum **ε** changes the outcome of tagging by an amount O(1).

It means that perturbative calculations give ∞

It means that the physics of hard objects may be irremediably contaminated by non-perturbative physics – not good for robustness or calculability!

Was long an issue in hadron-collider jet-finding.
Let's make sure it doesn't come back to haunt us!

Pruning followed by a mass-drop cut:

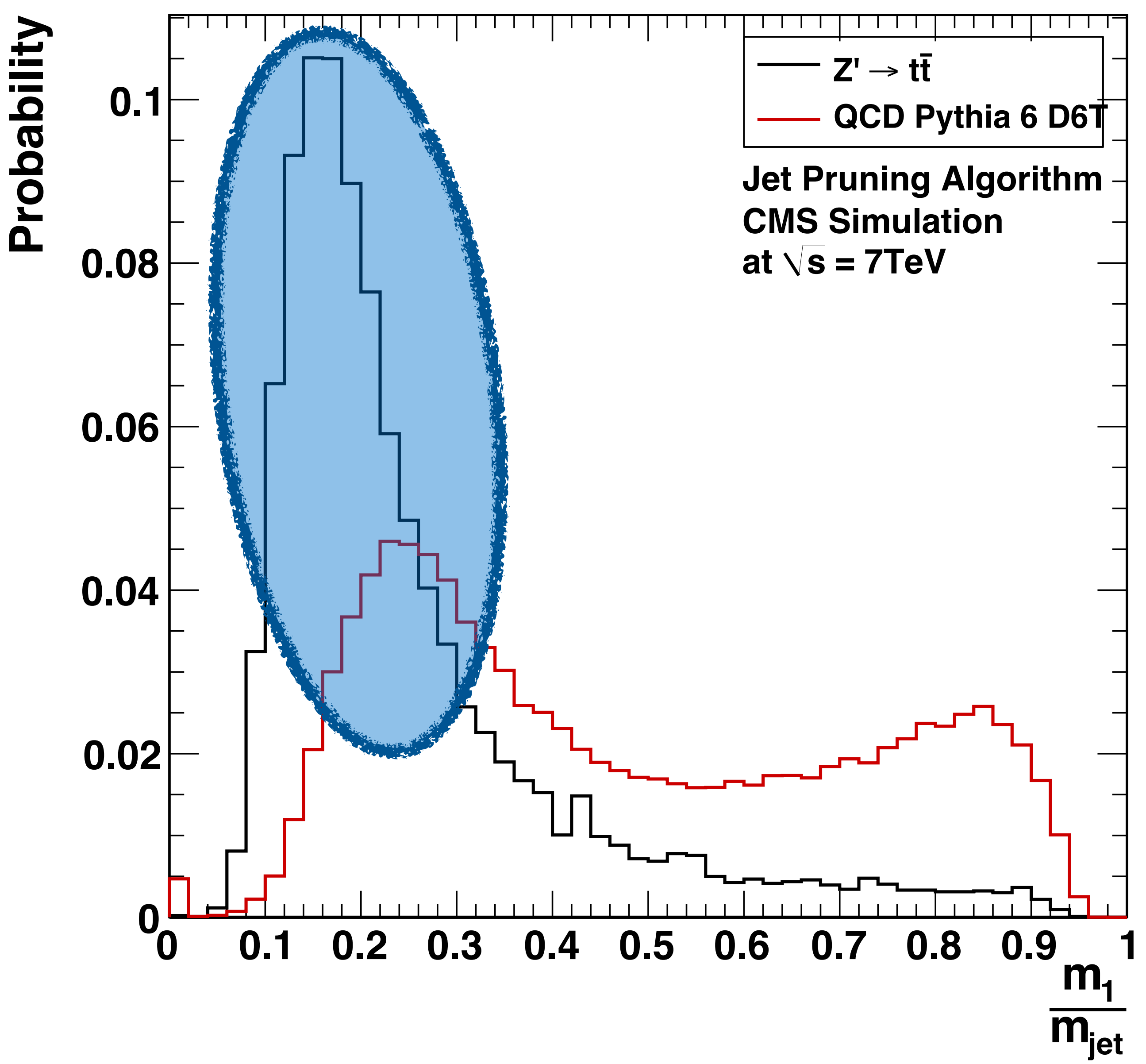
- Cut on the mass drop, $\mu = m_1/m$
  - m_1 is mass of highest pT subjet
  - subjets defined by un-clustering last step

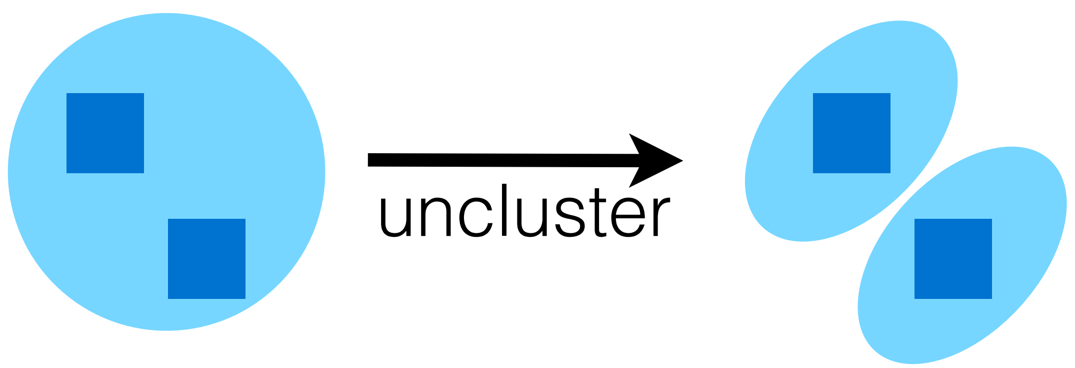
This is infrared unsafe

Pruning followed by a mass-drop cut:

- Cut on the mass drop, $\mu = m_1/m$
  - m_1 is mass of highest pT subjet
  - subjets defined by un-clustering last step

This is infrared unsafe

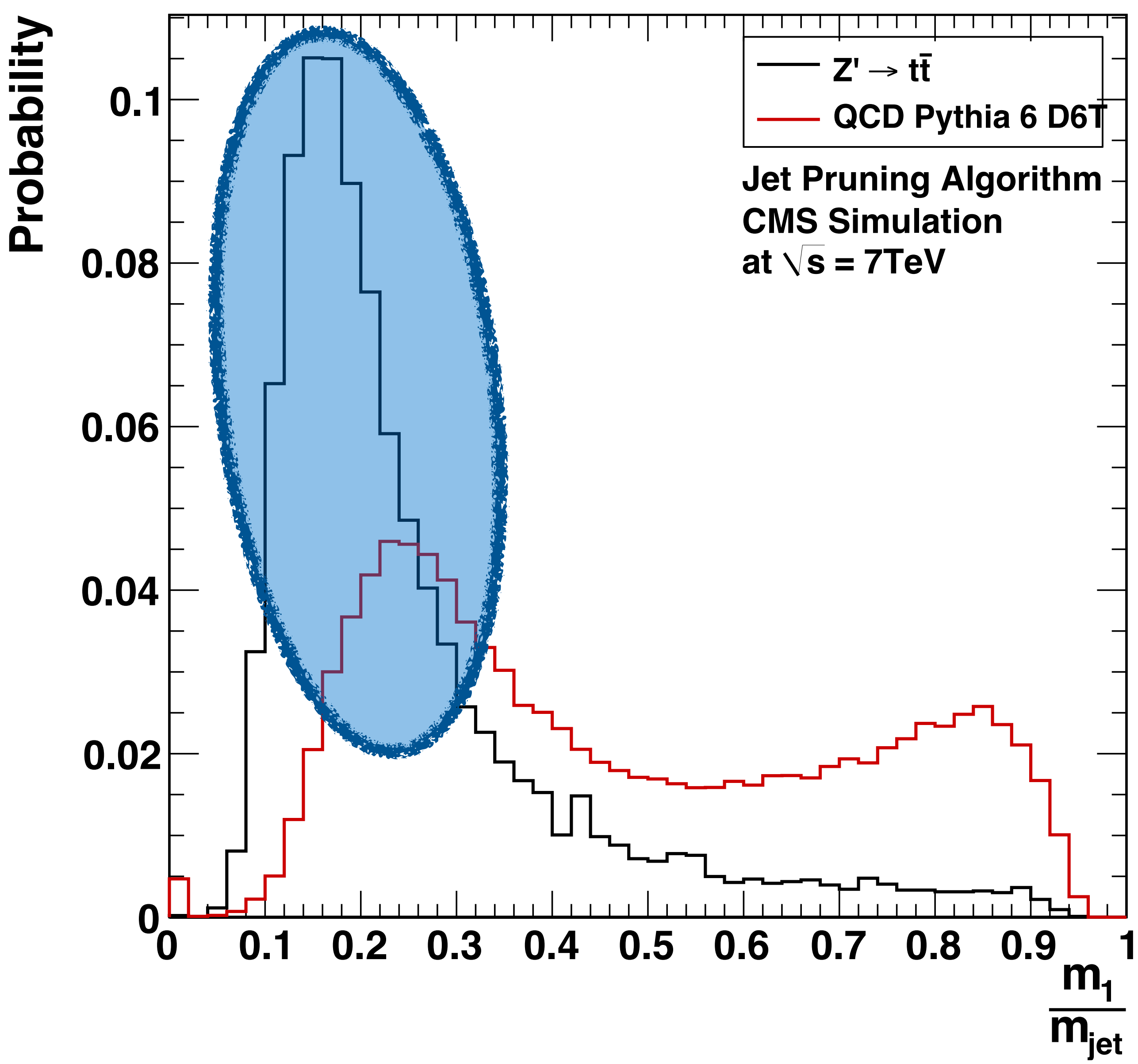


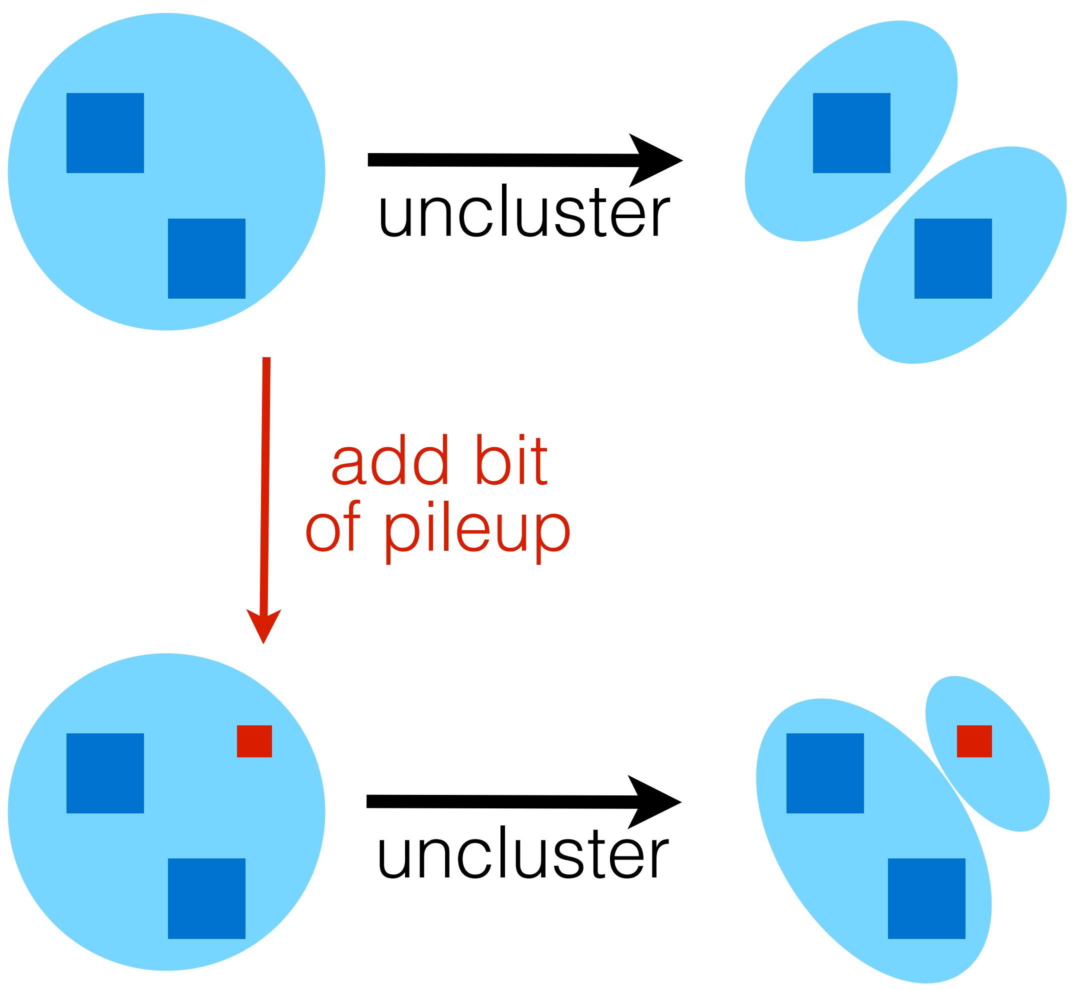


Pruning followed by a mass-drop cut:

- Cut on the mass drop, $\mu = m_1/m$
  - m_1 is mass of highest pT subjet
  - subjets defined by un-clustering last step

This is infrared unsafe

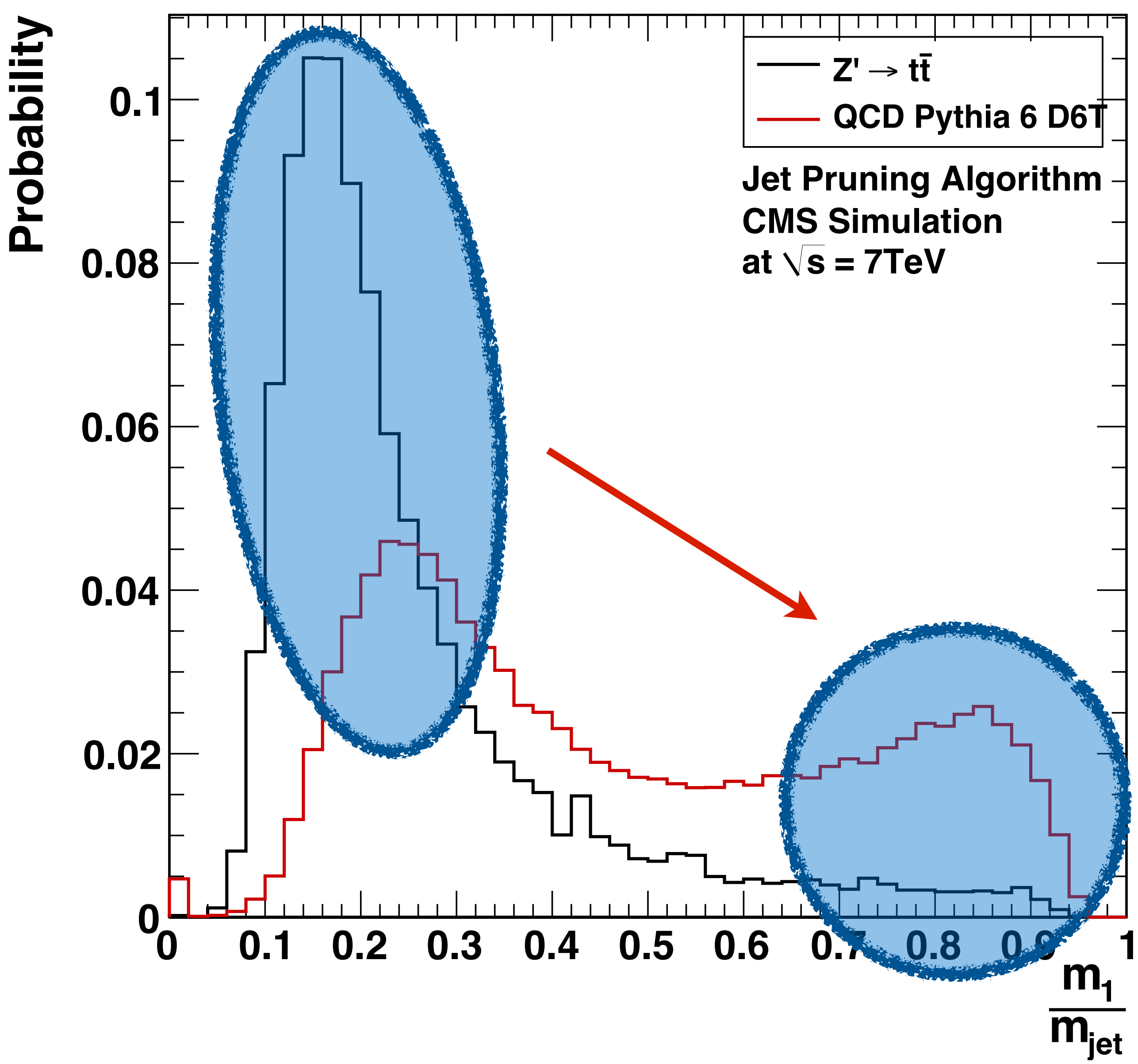


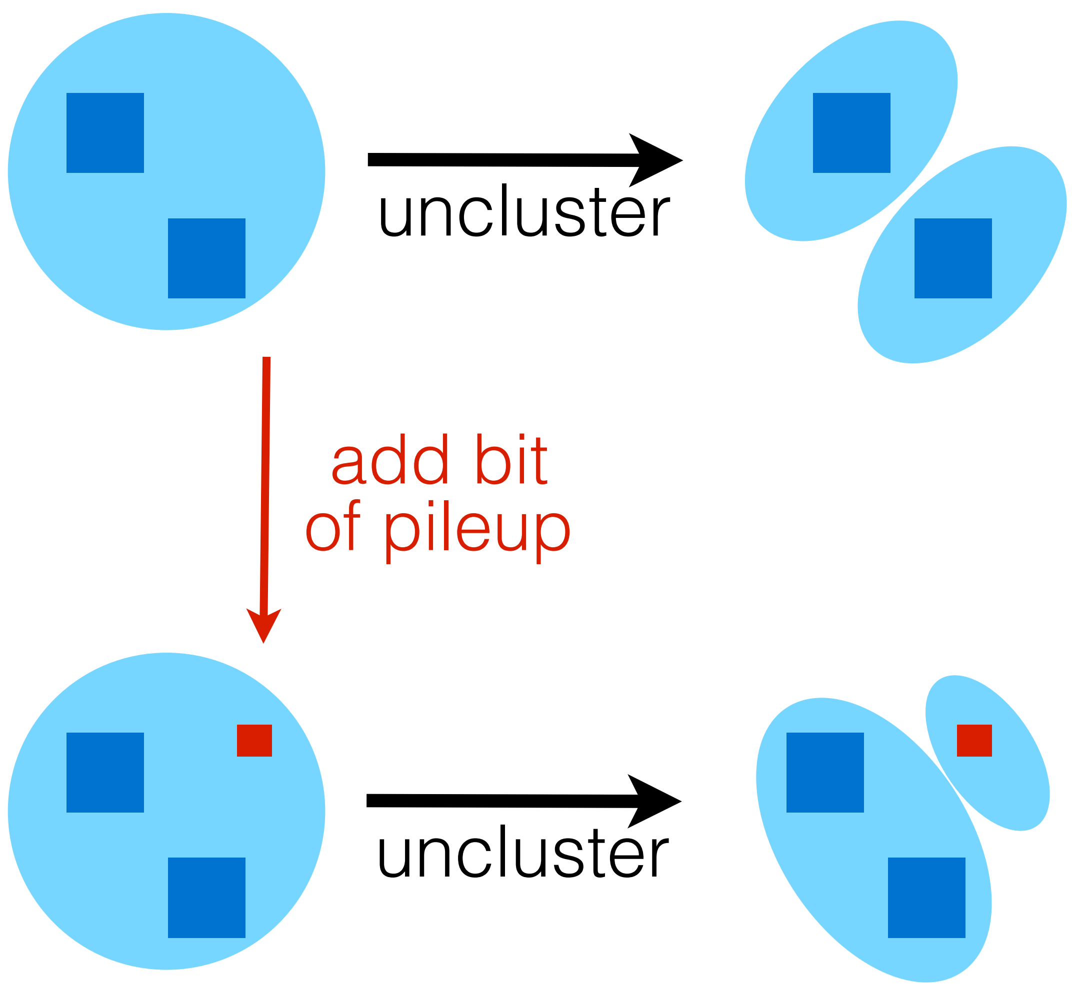


Pruning followed by a mass-drop cut:

- Cut on the mass drop, $\mu = m_1/m$
  - m_1 is mass of highest pT subjet
  - subjets defined by un-clustering last step

This is infrared unsafe

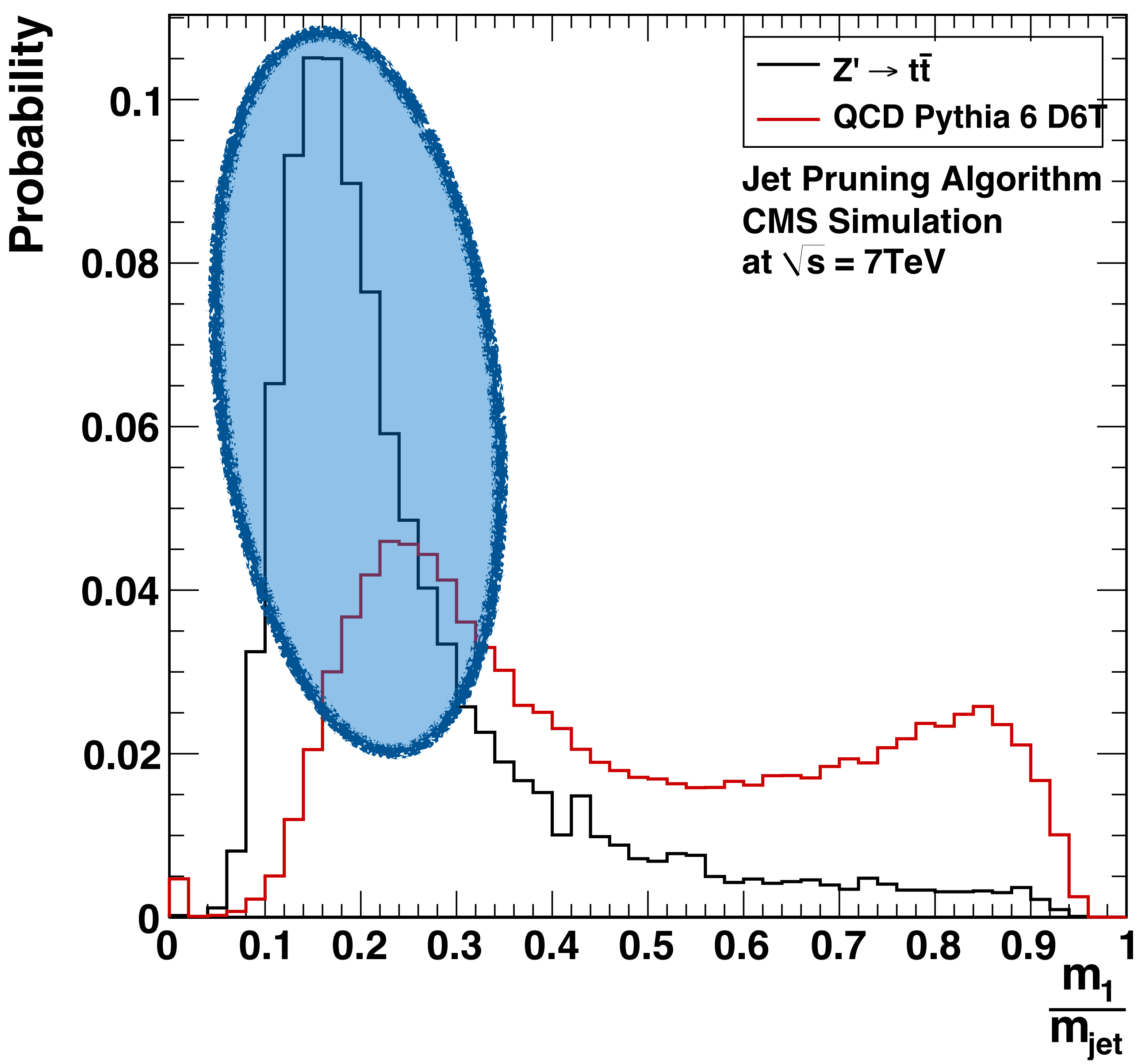


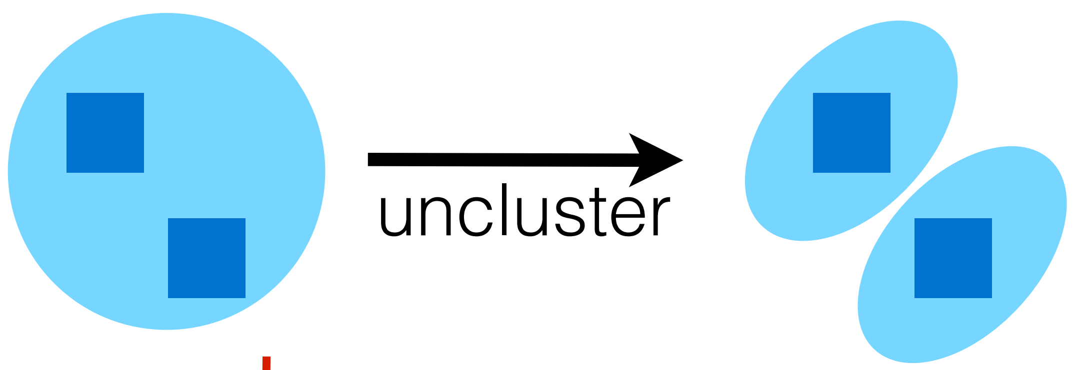


Pruning followed by a mass-drop cut:

- Cut on the mass drop, $\mu = m_1/m$
  - m_1 is mass of highest pT subjet
  - subjets defined by un-clustering last step

This is infrared unsafe



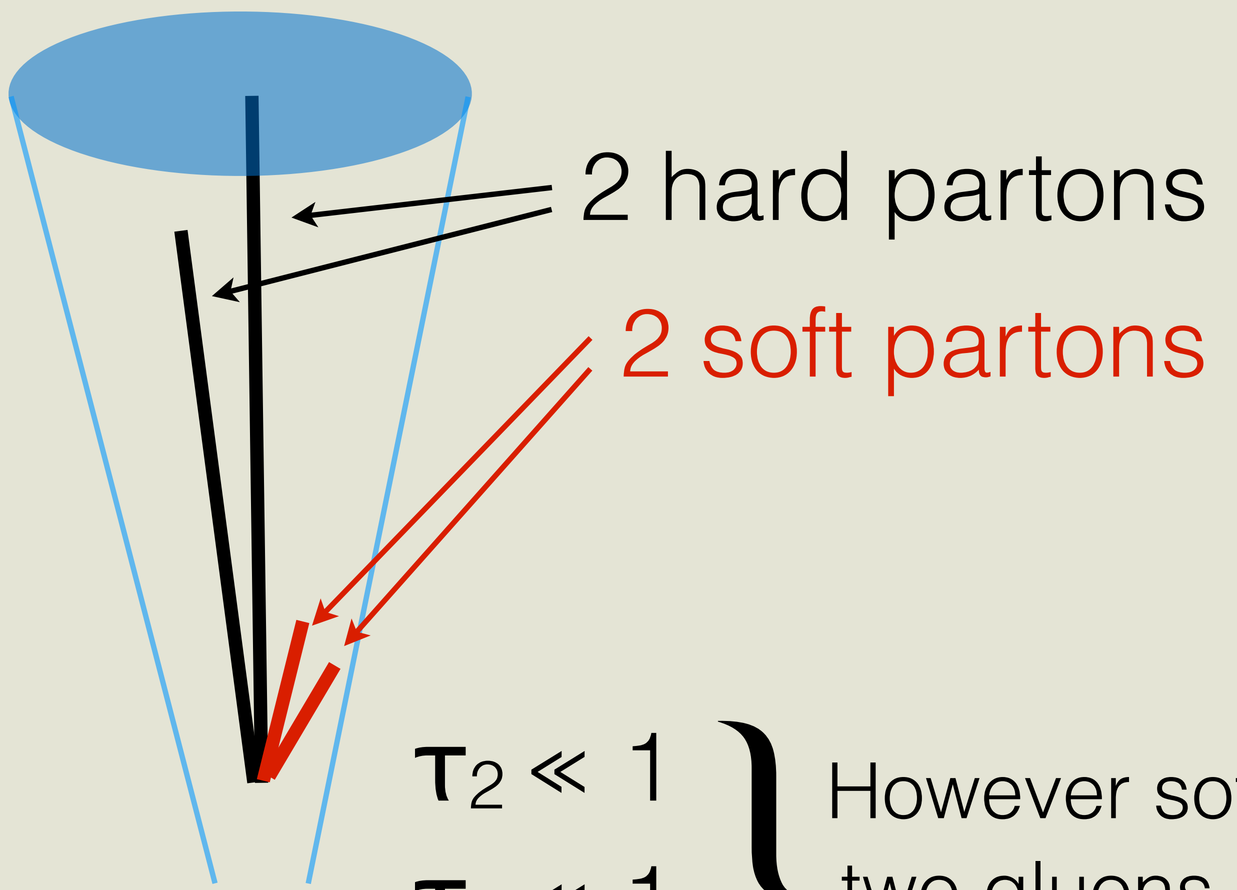


If you use "sane" pruning, the IR safety problem gets cured

IR issues in τ_{23}

N-subjettiness τ_3 / τ_2:

τ_2 measures departure from 2-parton energy flow
τ_3 measures departure from 3-parton energy flow



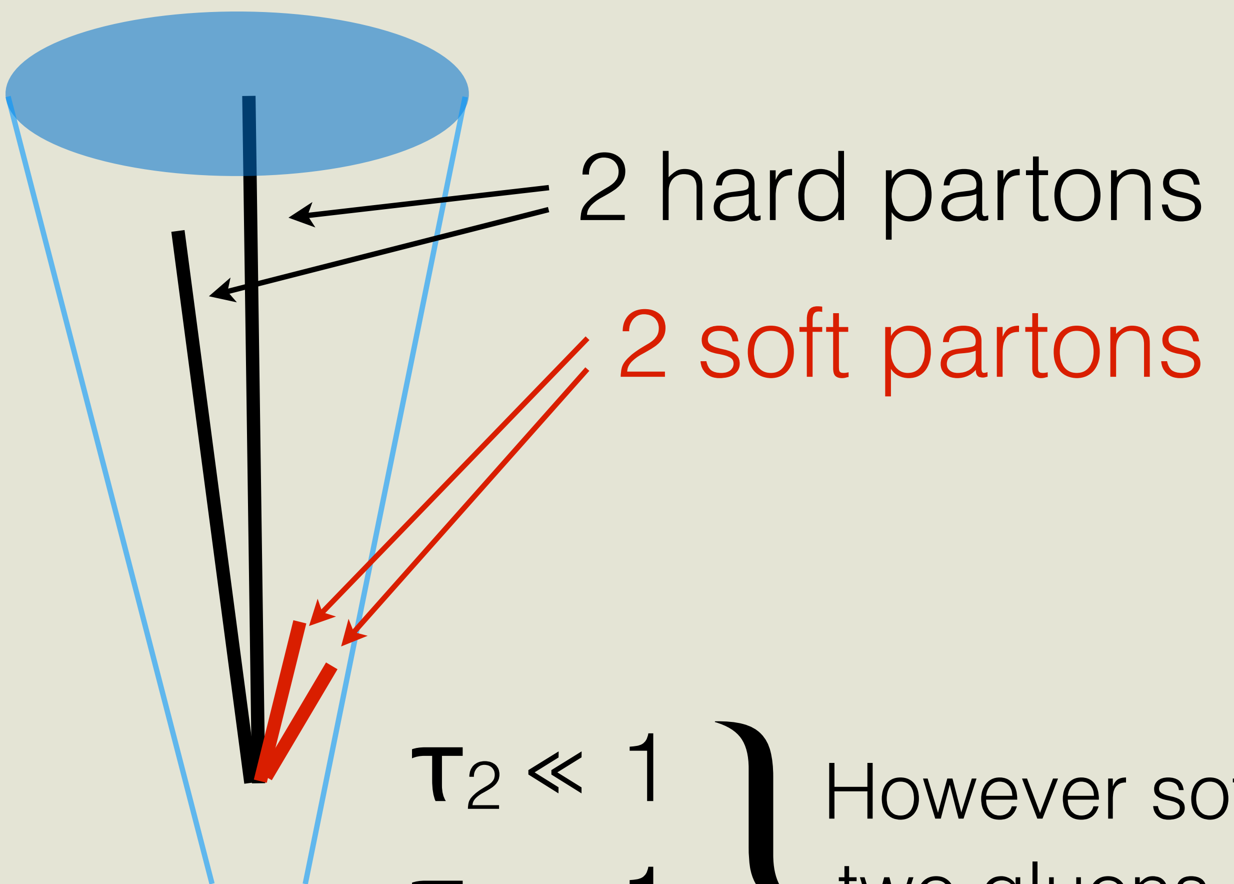
$\tau_2 \ll 1$
$\tau_3 \ll 1$
$\tau_3/\tau_2 \sim 1$

However soft the two gluons, you get finite τ_3/τ_2

That's IR unsafe

N-subjettiness τ_3 / τ_2:

τ_2 measures departure from 2-parton energy flow
τ_3 measures departure from 3-parton energy flow



$\tau_2 \ll 1$
$\tau_3 \ll 1$
$\tau_3/\tau_2 \sim 1$

However soft the two gluons, you get finite τ_3/τ_2

That's IR unsafe

Easily cured with a cut on τ_2 / τ_1, which forces 3rd prong not to be soft.

Extra cut has almost no impact on performance

Cacciari et al '12

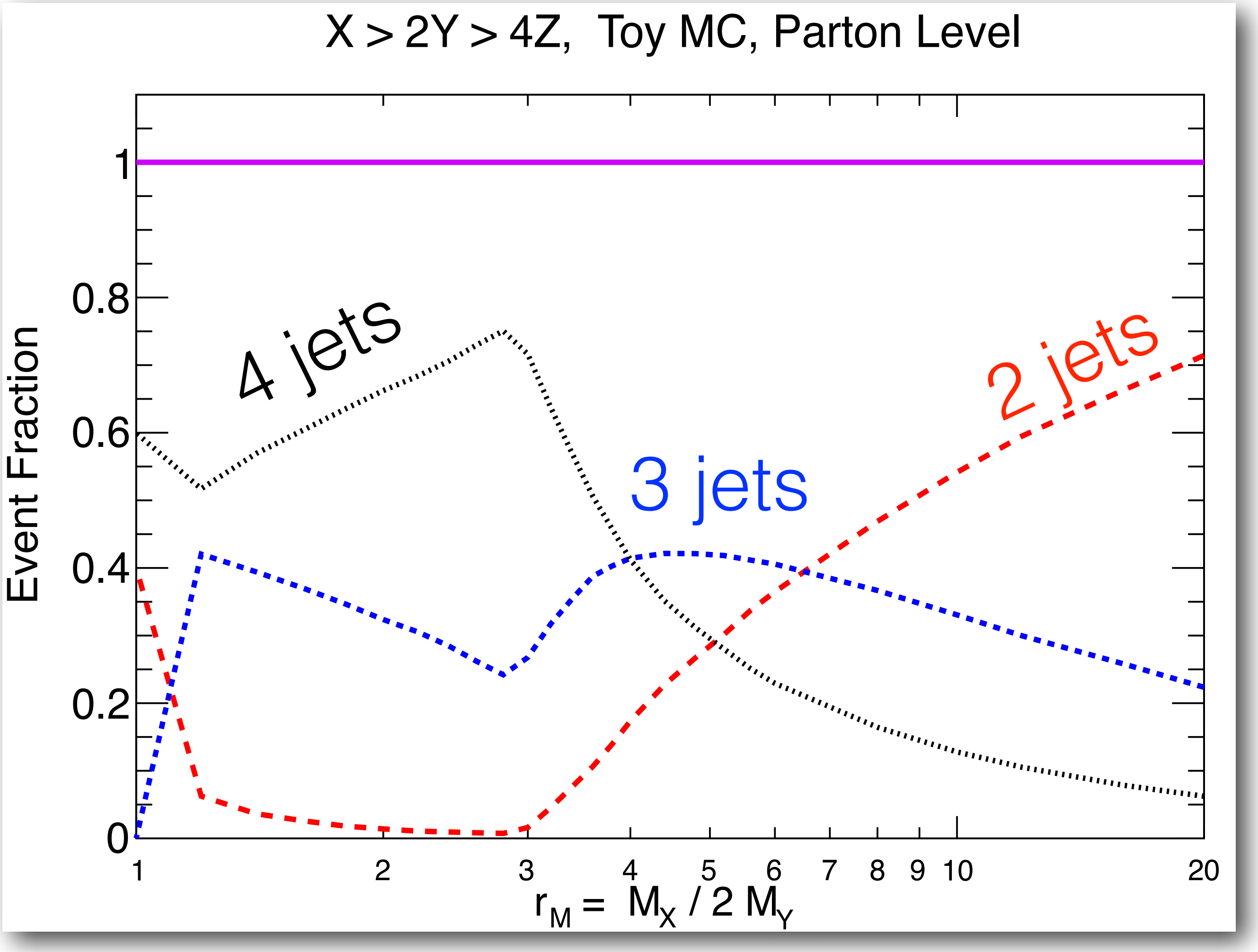
Scale invariant searches

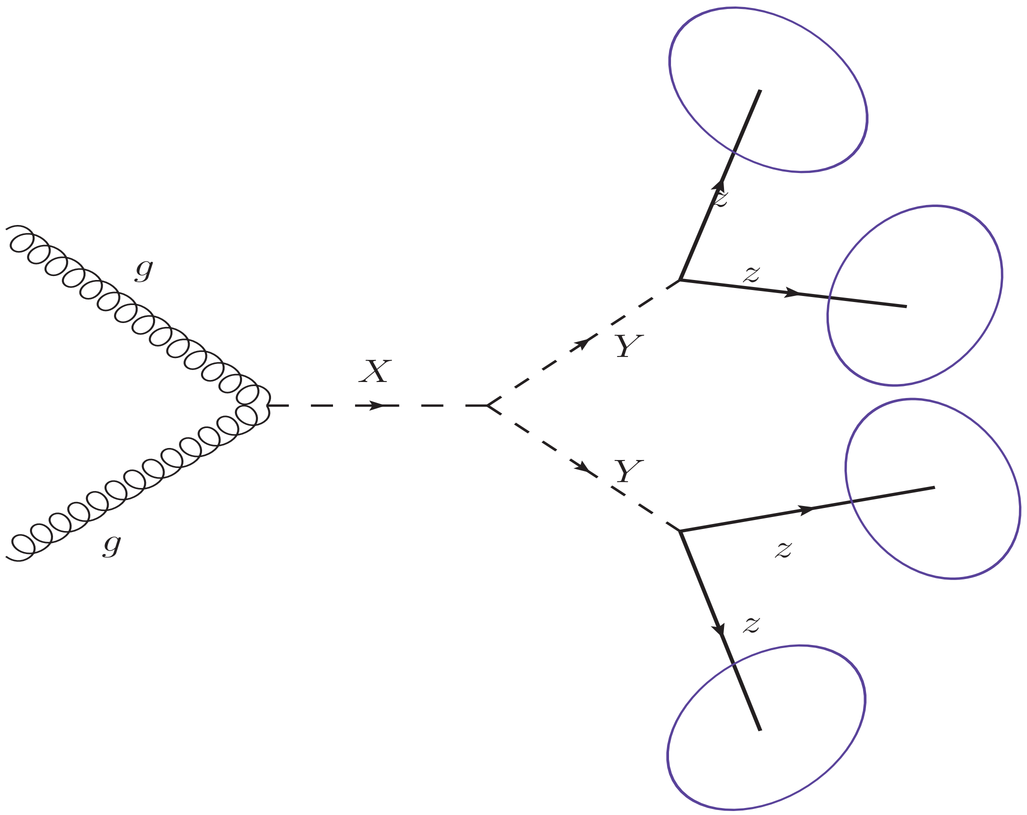
arXiv:1303.6636 with Gouzevitch, Oliveira, Rojo, Rosenfeld & Sanz

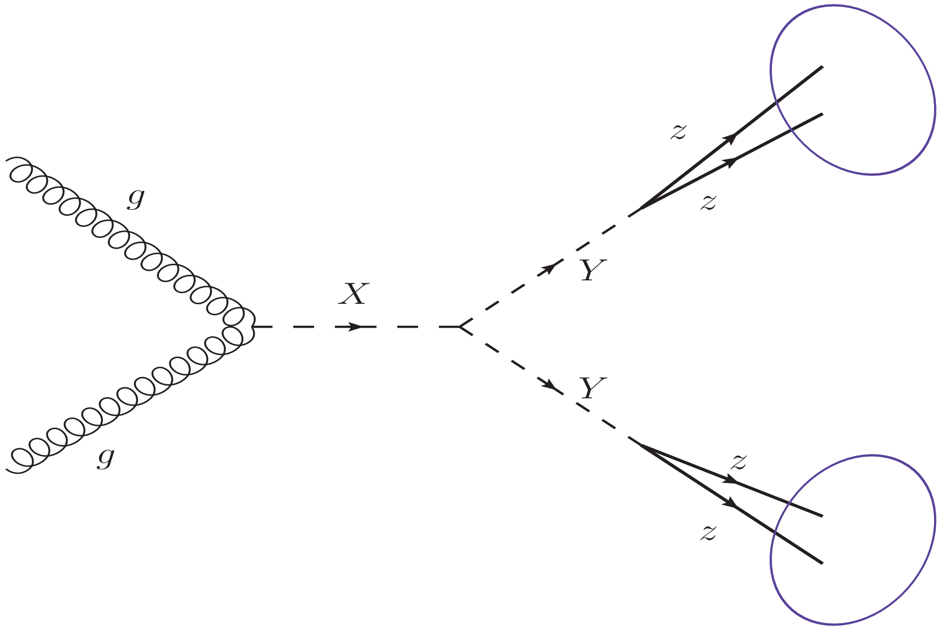
Experiments often have two distinct searches:

Resolved (small-R multi-jet)
Boosted (large-R fat-jet)

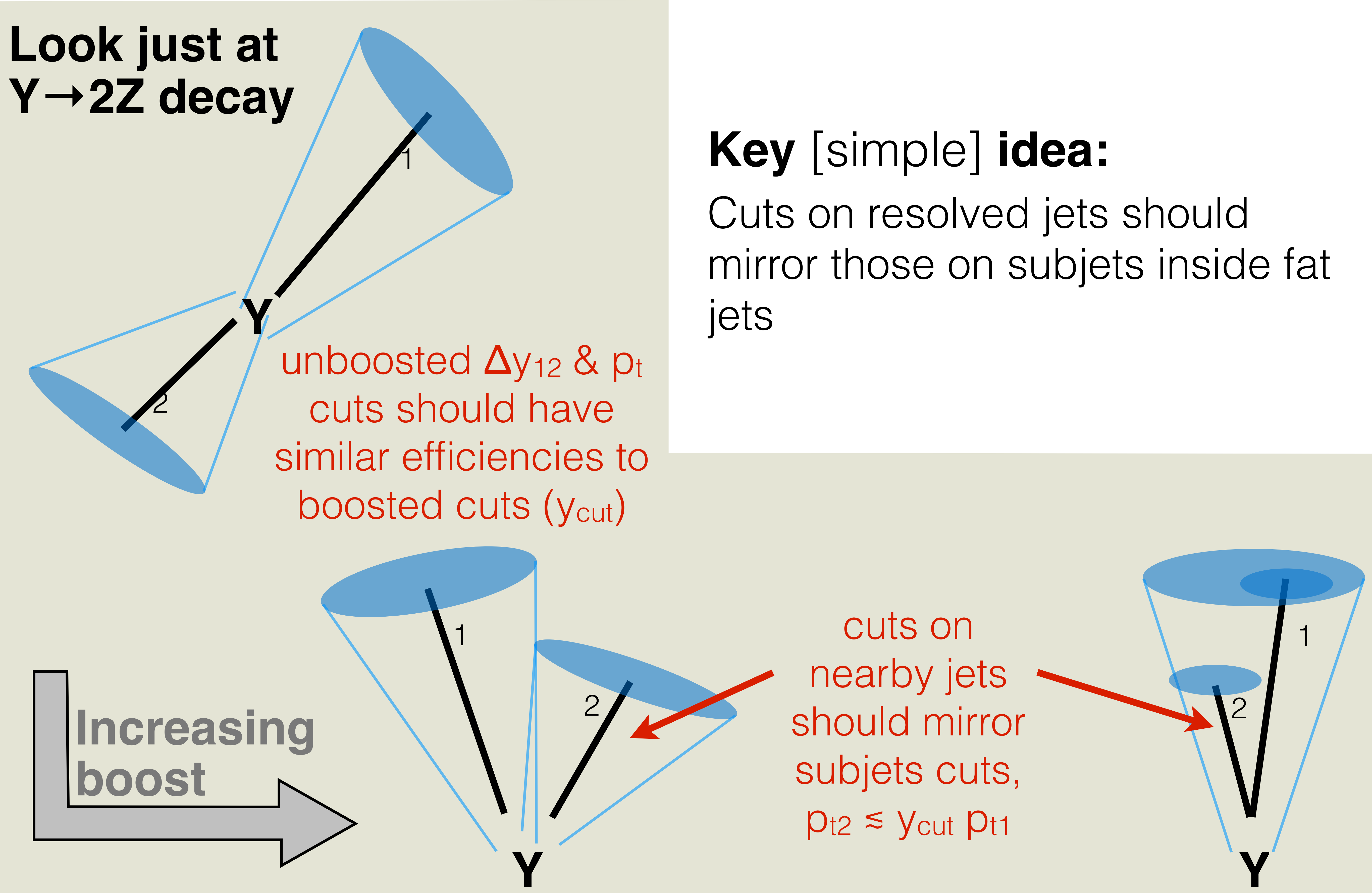
Can resolved and boosted analyses be consistently performed together?





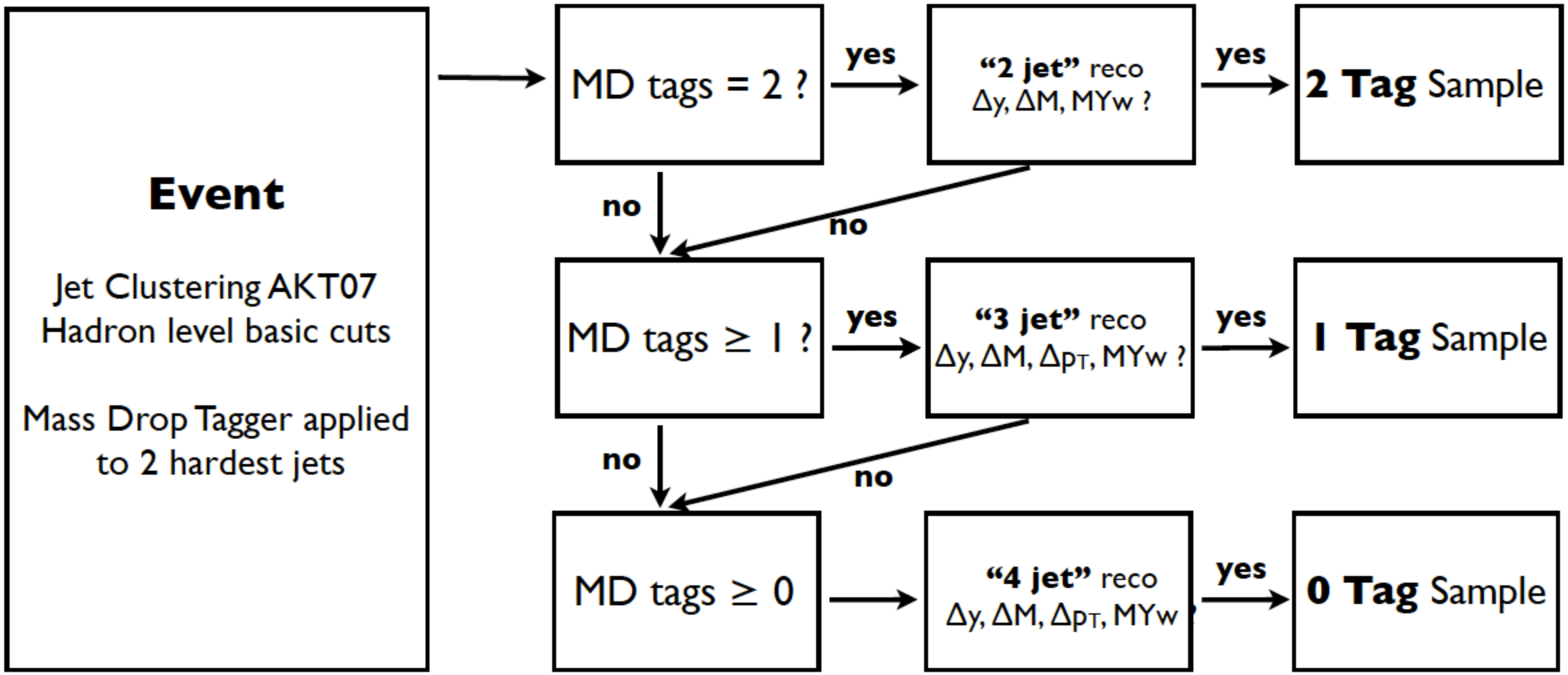


Look just at $Y \to 2Z$ decay



Key [simple] **idea:**

Cuts on resolved jets should mirror those on subjets inside fat jets



$$\left|\frac{2\,(m_{Y1}-m_{Y2})}{m_{Y1}+m_{Y2}}\right| \le f_m \qquad M_Y\,(1-f_m) \le m_{Y1}, m_{Y2} \le M_Y\,(1+f_m)$$

Traditional cuts on resolved jets

$$\Delta y \equiv |y_{Y1} - y_{Y2}| \le \Delta y_{\max} \qquad \Delta y \equiv |y_{Yi,1} - y_{Yi,2}| \le \Delta y^{\rm res}_{\max}$$

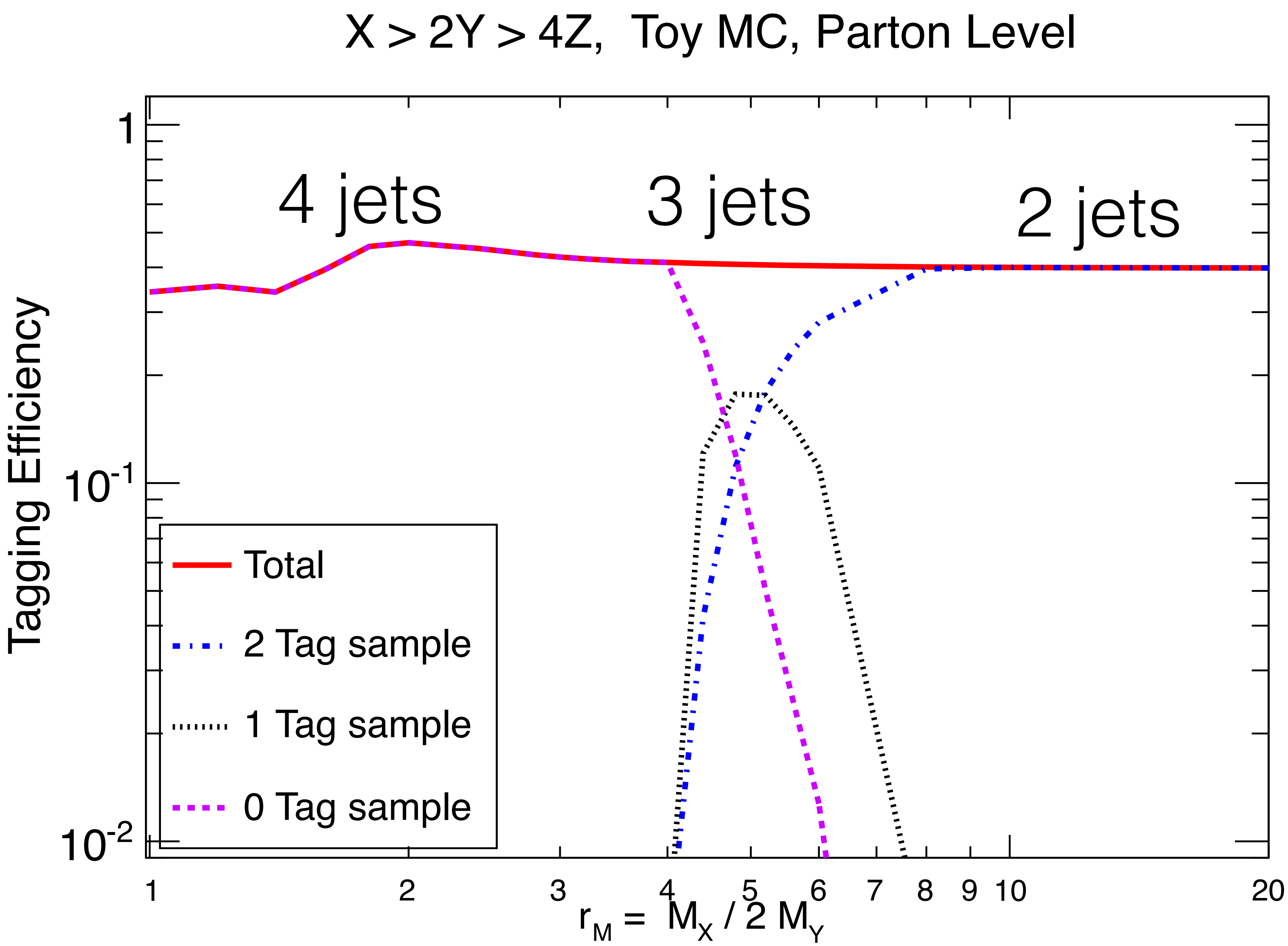
$$\Delta p_T \equiv p_T^{(1)} - p_T^{(2)} \ge (1 - y_{\rm cut})\, p_T^{(1)} \qquad \max\left(m_{Y,1}, m_{Y,2}\right) \le \mu \cdot m_Y$$

Boosted cuts on resolved jets

Jet Reconstruction					
R	$R_{\rm sj}$	$R_{\rm f}$	$n_{\rm filt}$	μ	$y_{\rm cut}$
0.5	1.3	0.3	3	0.67	0.09

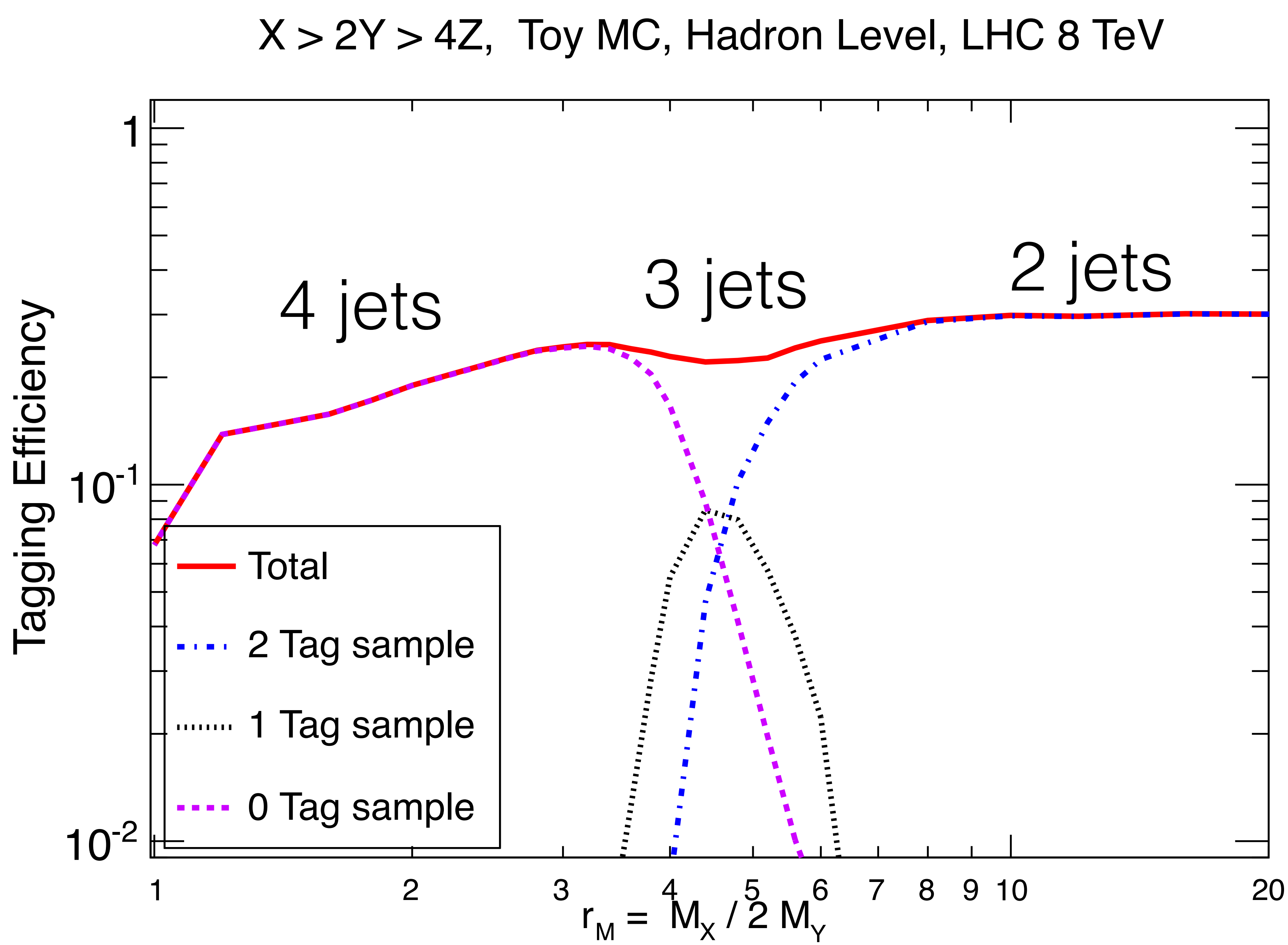
Basic cuts		
$p_T^{\rm min}$	$\lvert y_{\rm max}\rvert$	$H_T^{\rm min}$
25 GeV	5.0	100 GeV

Quality requirements			
M_Y	$\Delta y_{\rm max}$	$\Delta y^{\rm res}_{\rm max}$	f_m
125 GeV	1.3	1.5	0.15



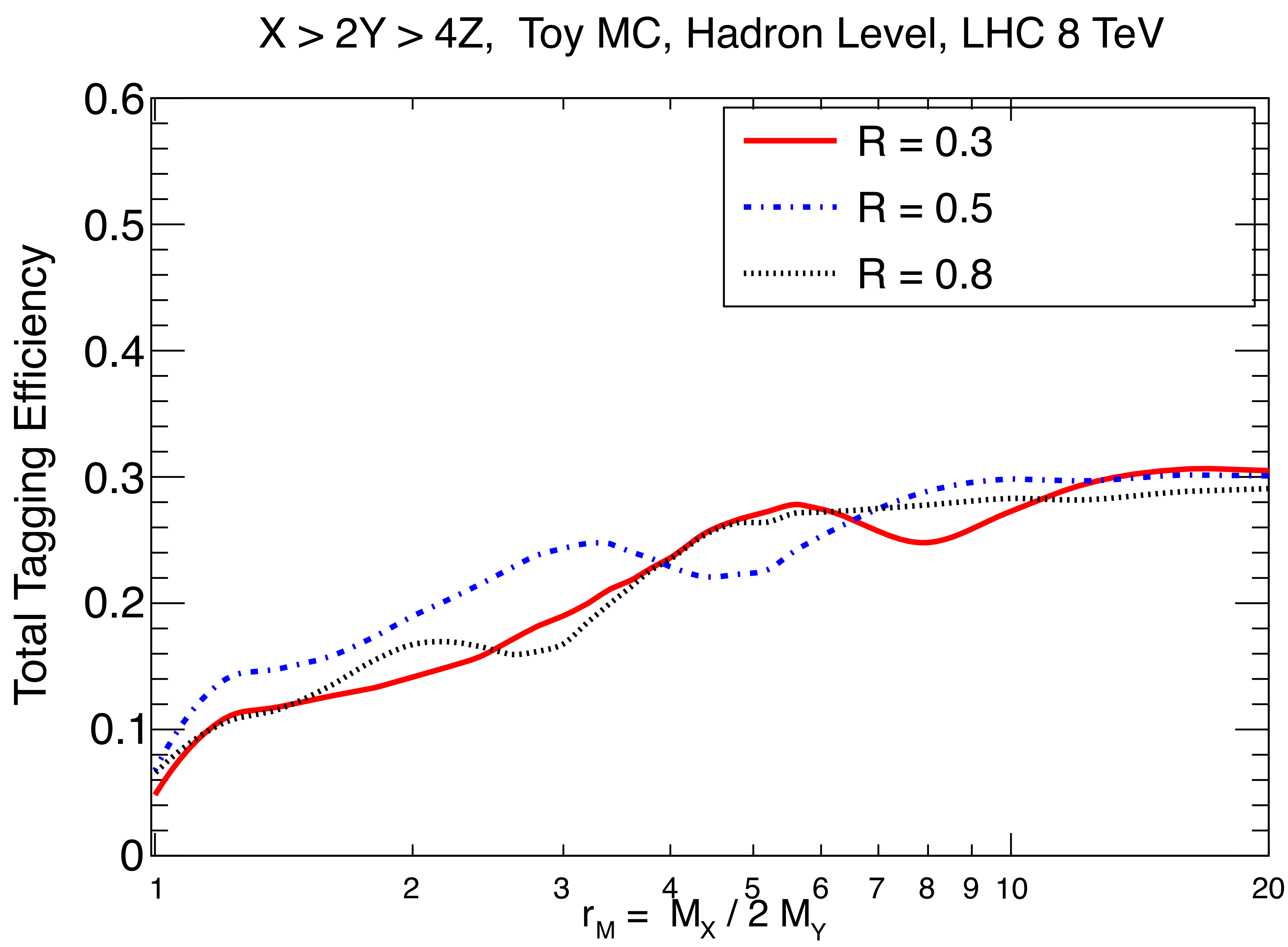
Efficiency roughly stable as X→2Y→4Z goes from 4→3→2 jets

Cuts are close to those that optimise S/√B in all regions



Efficiency roughly stable as X→2Y→4Z goes from 4→3→2 jets

Cuts are close to those that optimise S/√B in all regions



Efficiency roughly independent of R used in clustering

Bottom line:

traditional and substructure techniques can be used together

[an analogous method still needs to be worked out for top]

Pileup in the boosted regime

Pronged taggers

Some have pileup-***reduction*** built in (MassDrop+Filtering, Pruning, HEPTopTagger, Template), essentially by using small (R~0.2–0.3) sub-cones, sometimes dynamically adjusted to the jet p_t

For heavy pileup you will need to supplement them with full pileup ***subtraction*** (e.g. area-based).

[Technically trivial, but so far studied only for filtering & trimming]

Shape-based taggers

Until recently, no clear way of subtracting pileup.

[Grooming reduces PU, but also discards info]

Pileup subtraction for shapes

Cacciari, Dutta, JH Kim, GPS & Soyez '12

n[th] **derivative of shape** wrt ghost momenta

Shape as a function of particle momenta in jet

$$V_{\mathrm{jet}}^{[n]} \equiv A_g^n \frac{d^n}{dr_{t,g}^n} V(\{p_i\}_{\mathrm{jet}})$$

Ghost area

Pileup subtraction for shapes

Cacciari, Dutta, JH Kim, GPS & Soyez '12

n[th] **derivative of shape** wrt ghost momenta

Shape as a function of particle momenta in jet

$$V_{\text{jet}}^{[n]} \equiv A_g^n \frac{d^n}{dr_{t,g}^n} V(\{p_i\}_{\text{jet}})$$

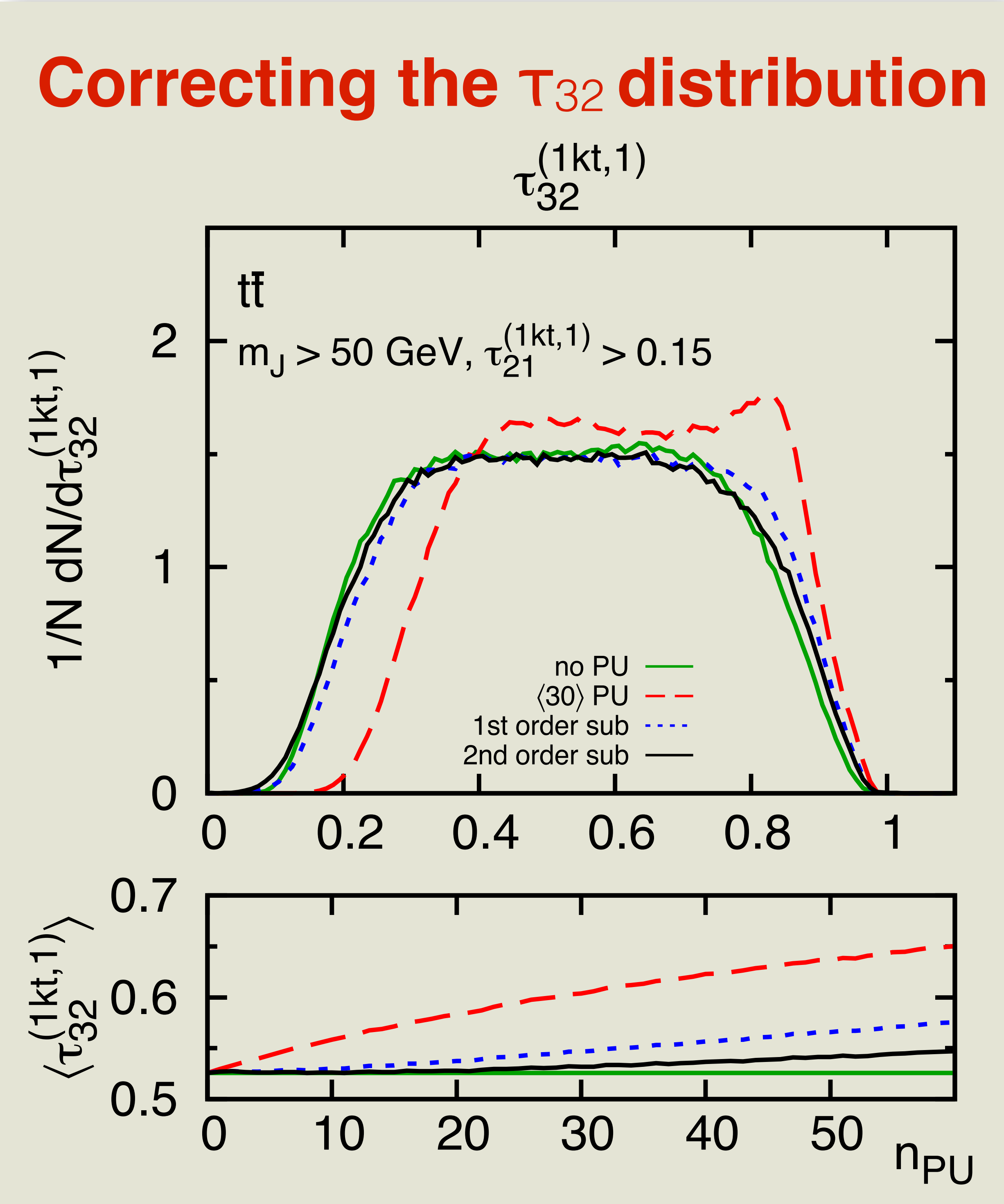
Ghost area

$$V_{\text{jet,sub}} = V_{\text{jet}} - \rho V_{\text{jet}}^{[1]} + \frac{1}{2}\rho^2 V_{\text{jet}}^{[2]} + \cdots$$

Subtracted shape

pileup density

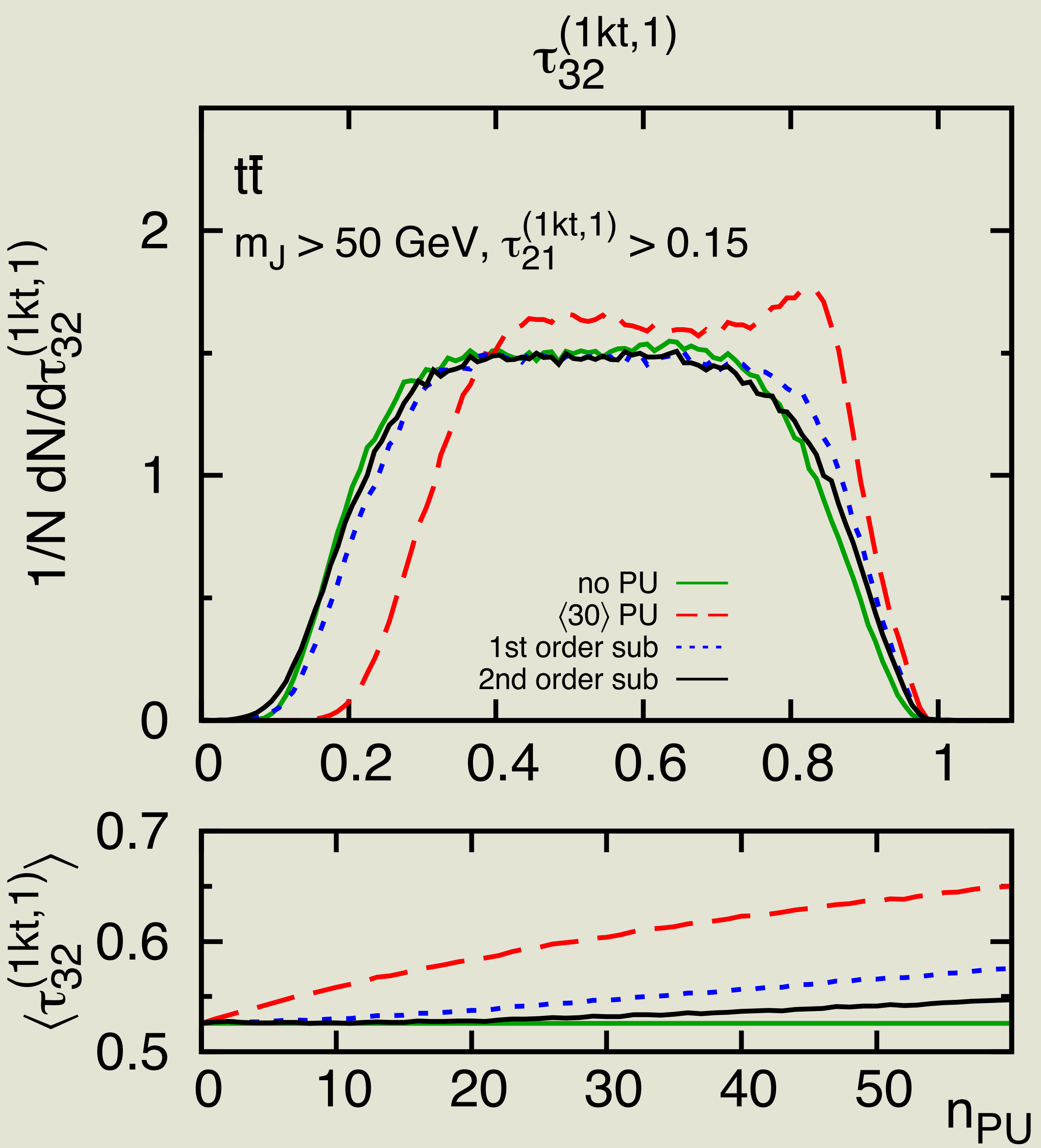
Practical test: τ_{32} and top tagging



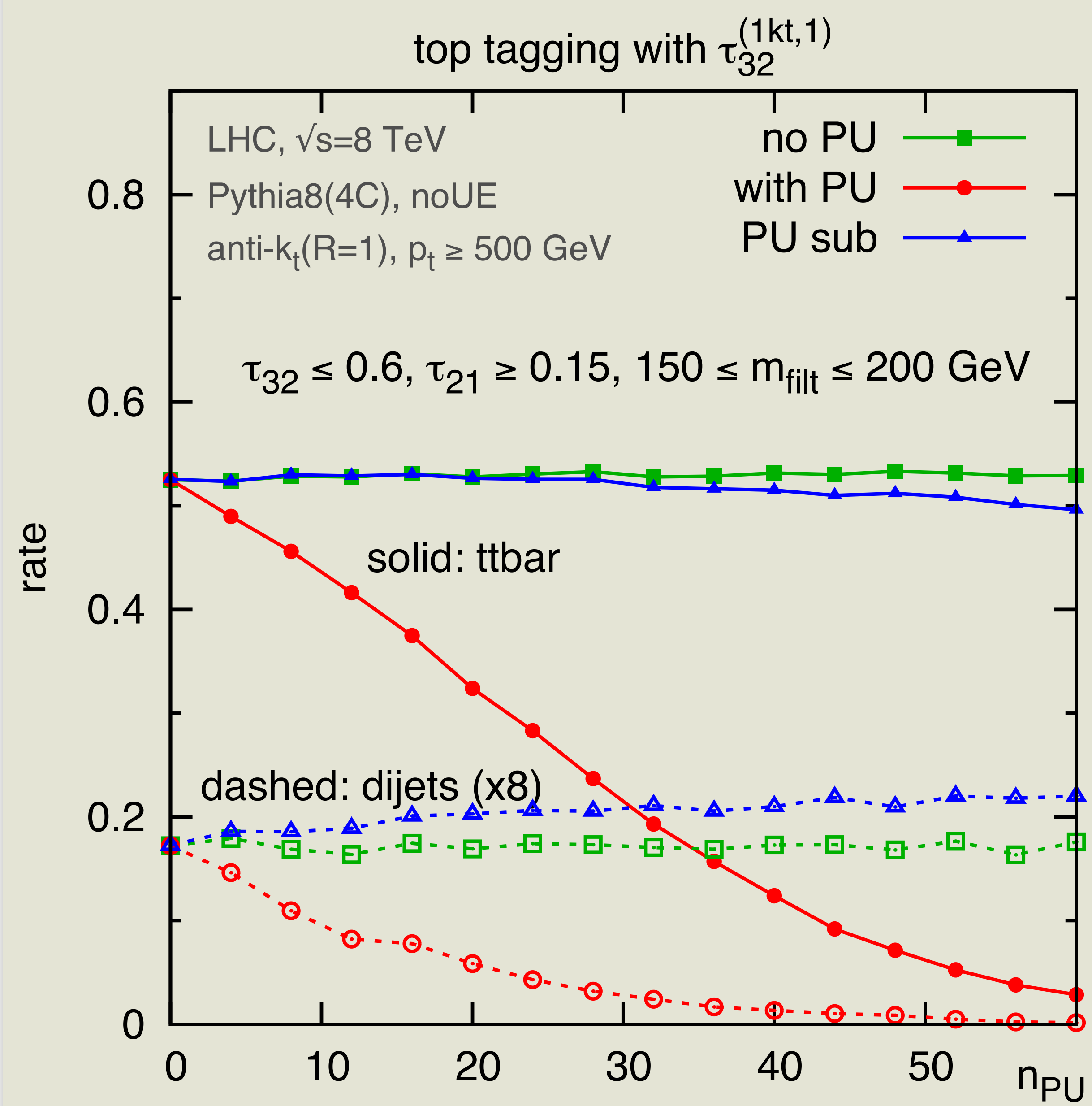
Green: no PU **Red: with PU** **Blue/Black: subtracted**

Practical test: τ32 and top tagging

Correcting the τ32 distribution



Tagging efficiency



Green: no PU **Red: with PU** **Blue/Black: subtracted**

Closing

Other things I would have liked to talk about

Jet deconstruction

Matrix-element method extended to all orders.

Soper & Spannowsky, arXiv:1102.3480, arXiv:1211.3140

Q-jets

Does clustering have to give a unique answer? What if you probe multiple possible clustering histories?

Ellis et al. arXiv:1201.1914, arXiv:1304.2394

Jet substructure by accident

Rather than looking for 16 jets (e.g. in BSM → 4 low-p_t tops), look for O(4) fat jets, each with substructure. May be easier to reliably predict backgrounds

Cohen, Izaguirre, Lisanti & Lou arXiv:1212.1456

Quark–Gluon discrimination

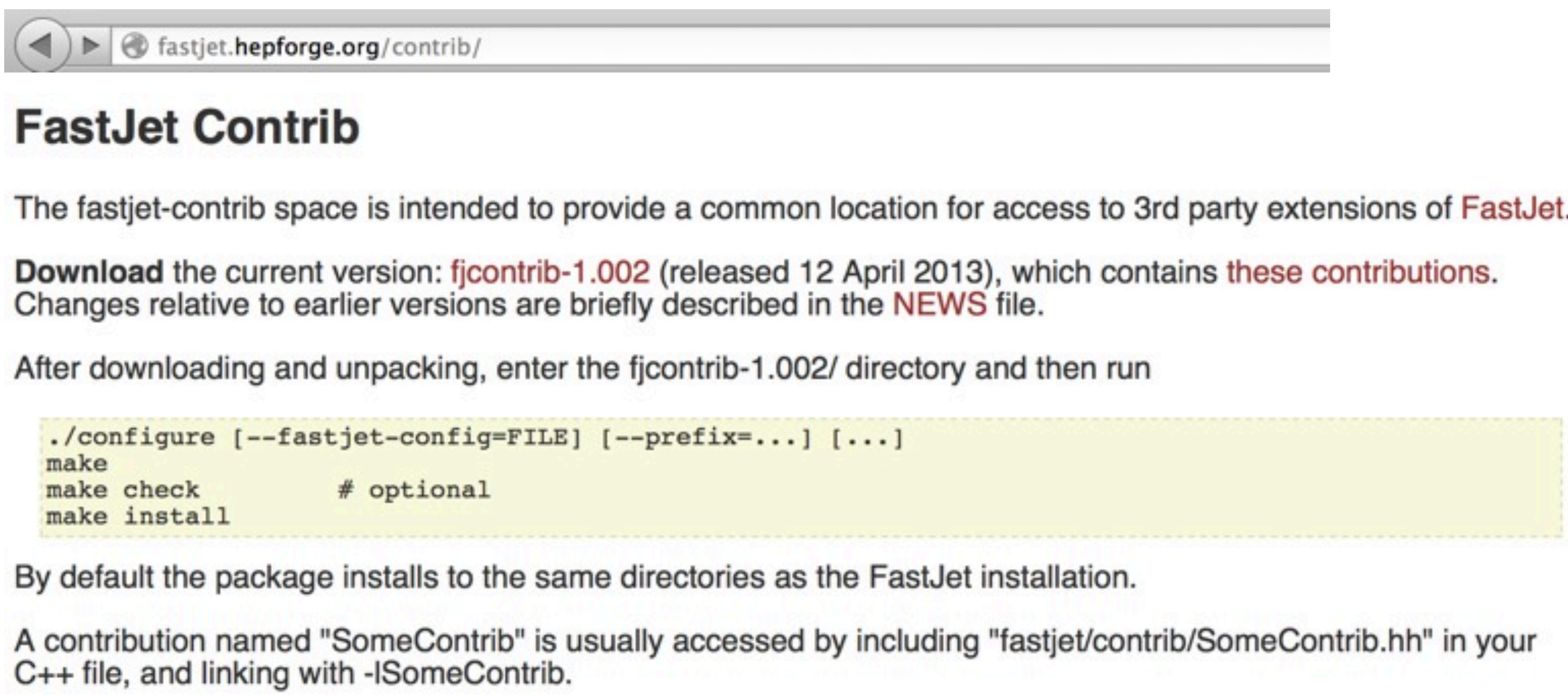
What information are we exploiting? Can we exploit it better?

Gallicchio & Schwartz '12
Larkoski, GPS & Thaler, in preparation

Other things I would have liked to talk about

FastJet Contrib

A space for people to contribute their own jet-tool libraries, to provide users with uniform, regularly updated and reliable access to a broad range of jet tools.


fastjet.hepforge.org/contrib/

FastJet Contrib

The fastjet-contrib space is intended to provide a common location for access to 3rd party extensions of FastJet.

Download the current version: fjcontrib-1.002 (released 12 April 2013), which contains these contributions.
Changes relative to earlier versions are briefly described in the NEWS file.

After downloading and unpacking, enter the fjcontrib-1.002/ directory and then run

```
./configure [--fastjet-config=FILE] [--prefix=...] [...]
make
make check        # optional
make install
```

By default the package installs to the same directories as the FastJet installation.

A contribution named "SomeContrib" is usually accessed by including "fastjet/contrib/SomeContrib.hh" in your C++ file, and linking with -lSomeContrib.

Summary [and points for discussion]

It's time to make the transition to a deep understanding of our tools, the only way of guaranteeing robustness

- Analytical control of "pronged" taggers now seems to be possible [though still early days]
- Taggers can have surprises in store for us – especially when we explore full LHC14 p_t range.
  [They can also be "fixed up", e.g. sane pruning, modified MDT]
- When do we want want to use "radiation-based" taggers?
  [Should resolved analyses always exploit q/g discrimination?]
- Do we need/want continuous resolved–fat-jet analyses?
- Pileup: it's time to start dealing with it systematically in our taggers [beyond just grooming, even as part of grooming]

EXTRAS

Different fat-jet tagger types

Prong based

(e.g. HEPTopTagger, Template Tagger)

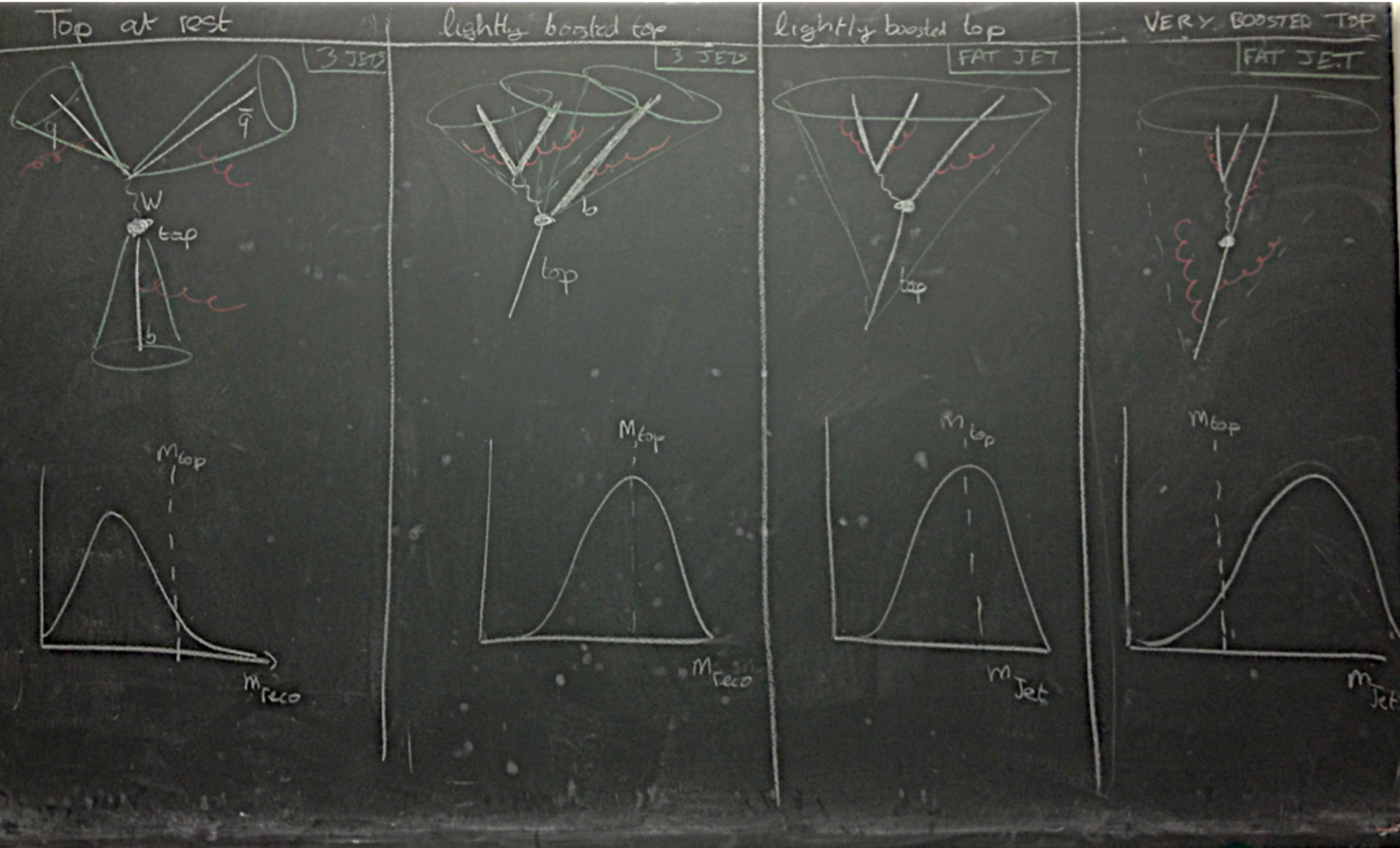
- Identifies prongs
- Requires prongs be consistent with kinematics of $t \to Wb \to$ 3 quarks

Radiation based

(e.g. N-subjettiness = τ_3/τ_2 + mass cut)

- Requires top-mass consistency (maybe with some grooming)
- Exploits weaker radiation from top (3 quarks) than background (1q+2g or 3g)

Top quarks, Tops from quark-jets &Top fat-jets
[and their radiation pattern]

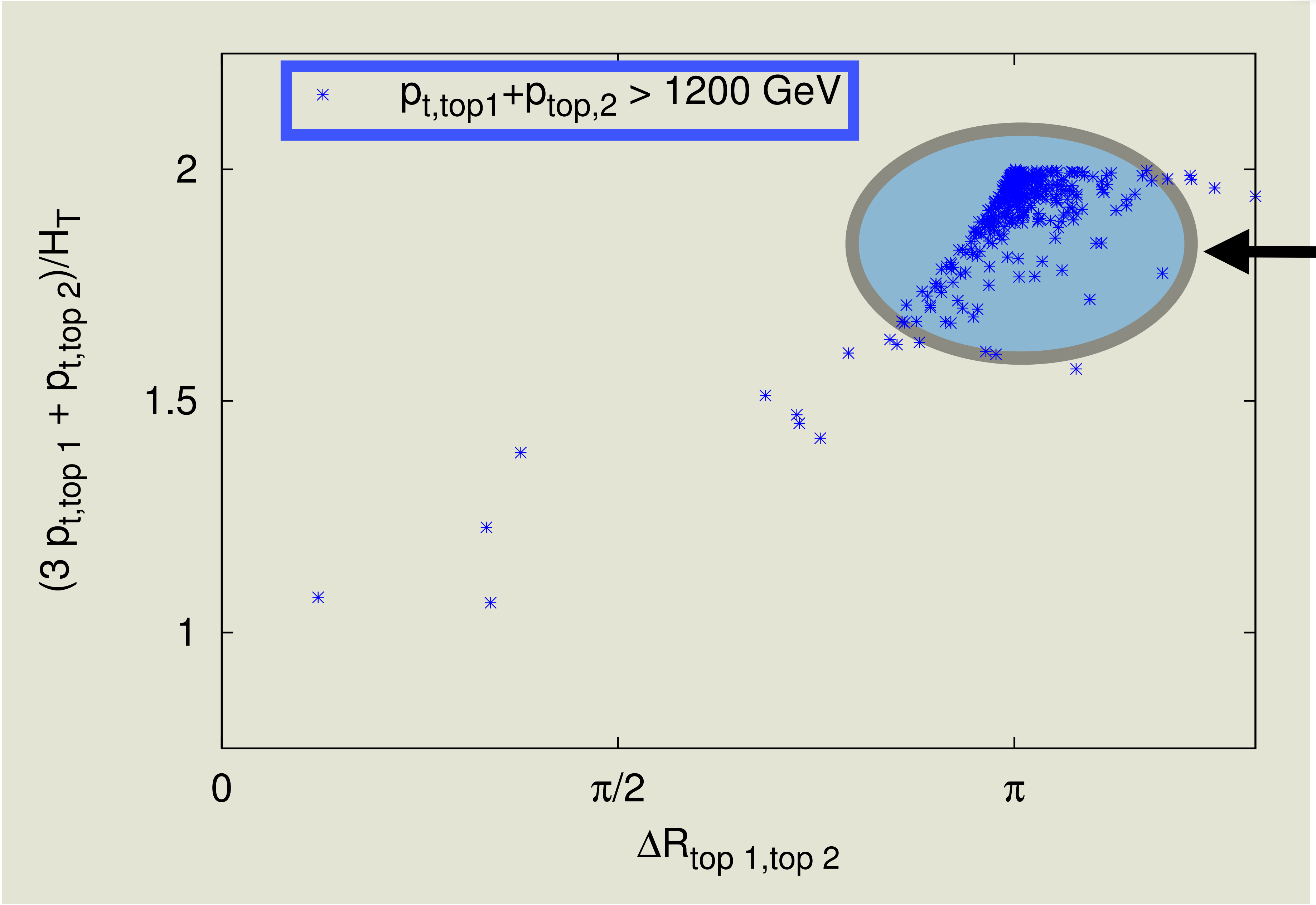


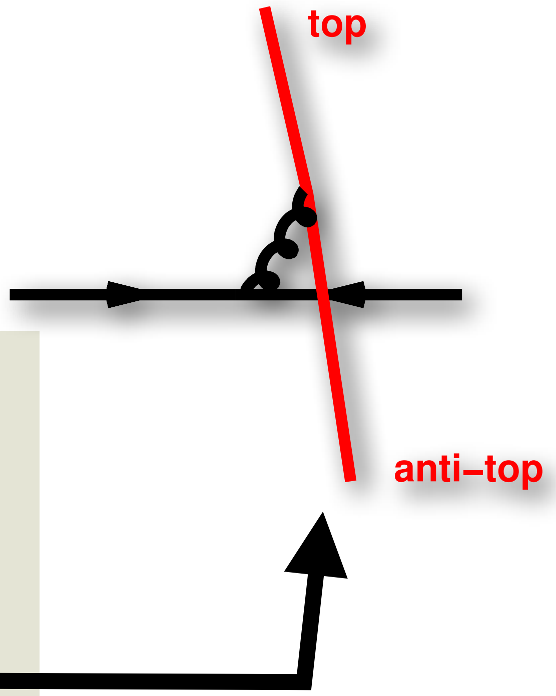
Are top pairs in high-p_t events always back-to-back?

A reminder that top-quarks at LHC are almost “light”

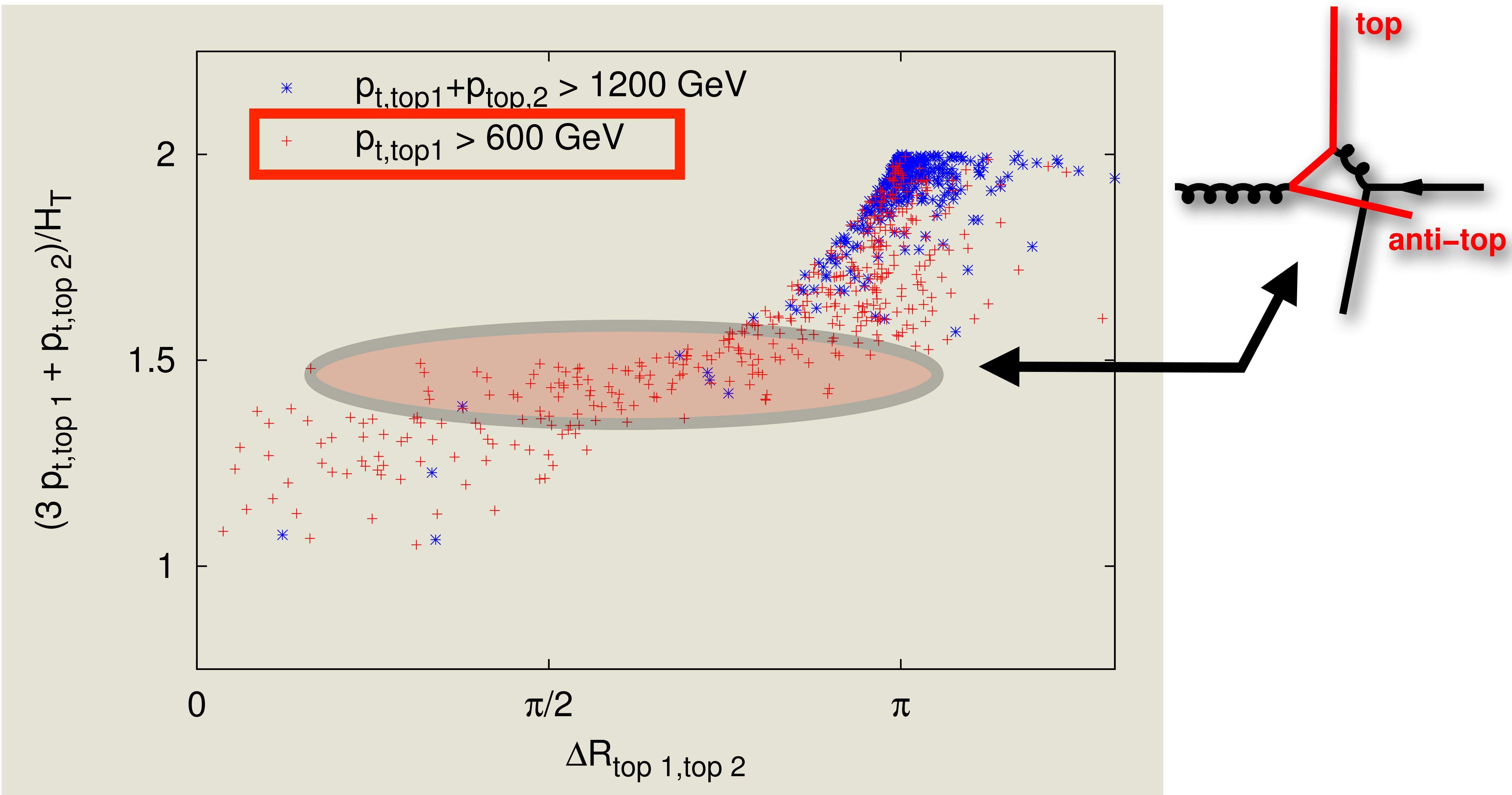
An 8 TeV study with POWHEG, top-pair production, no decay and no parton showering (to keep things simple)

Flavour Creation





Flavour Excitation – tops inside your PDFs



Gluon Splitting