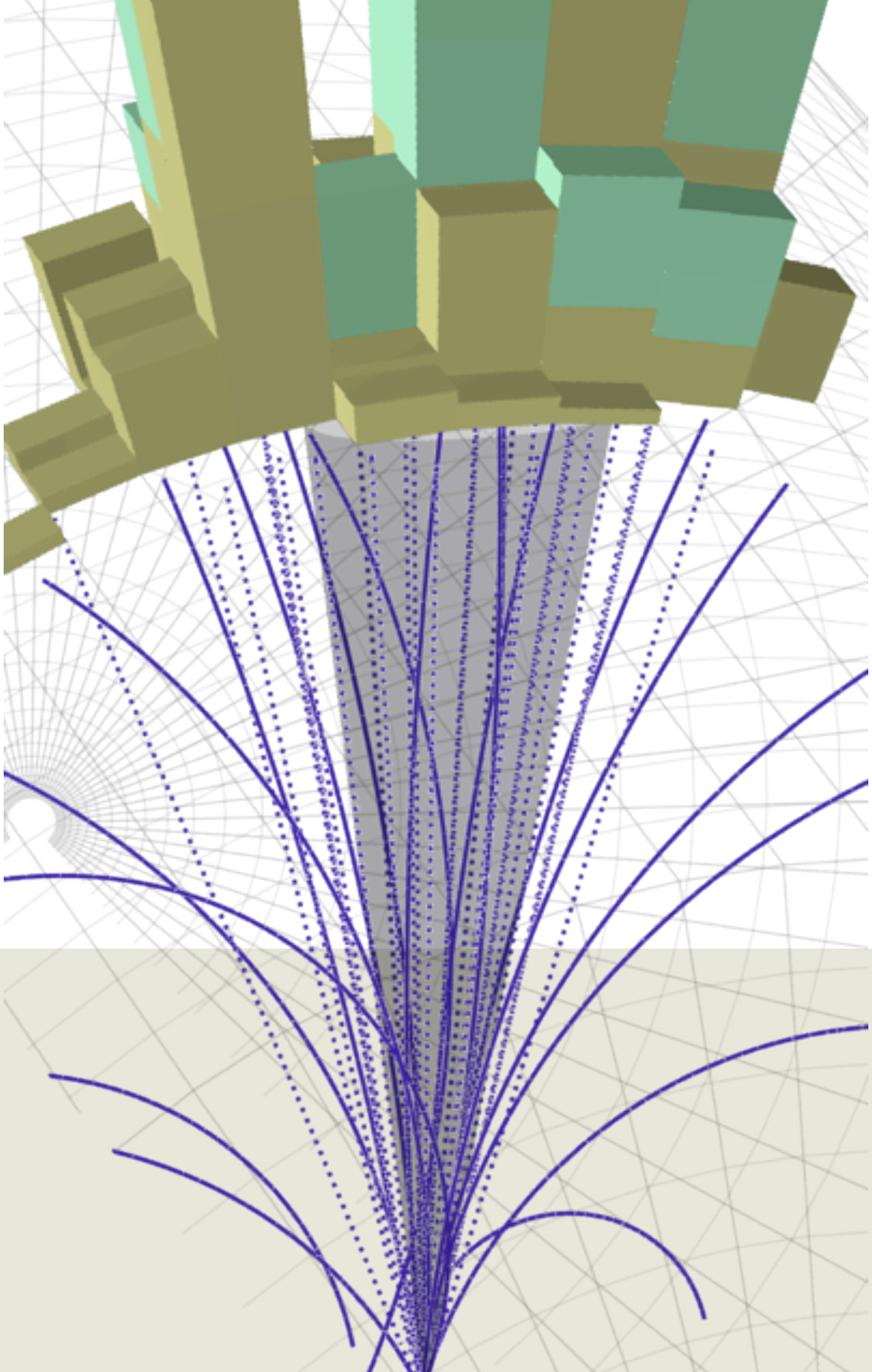
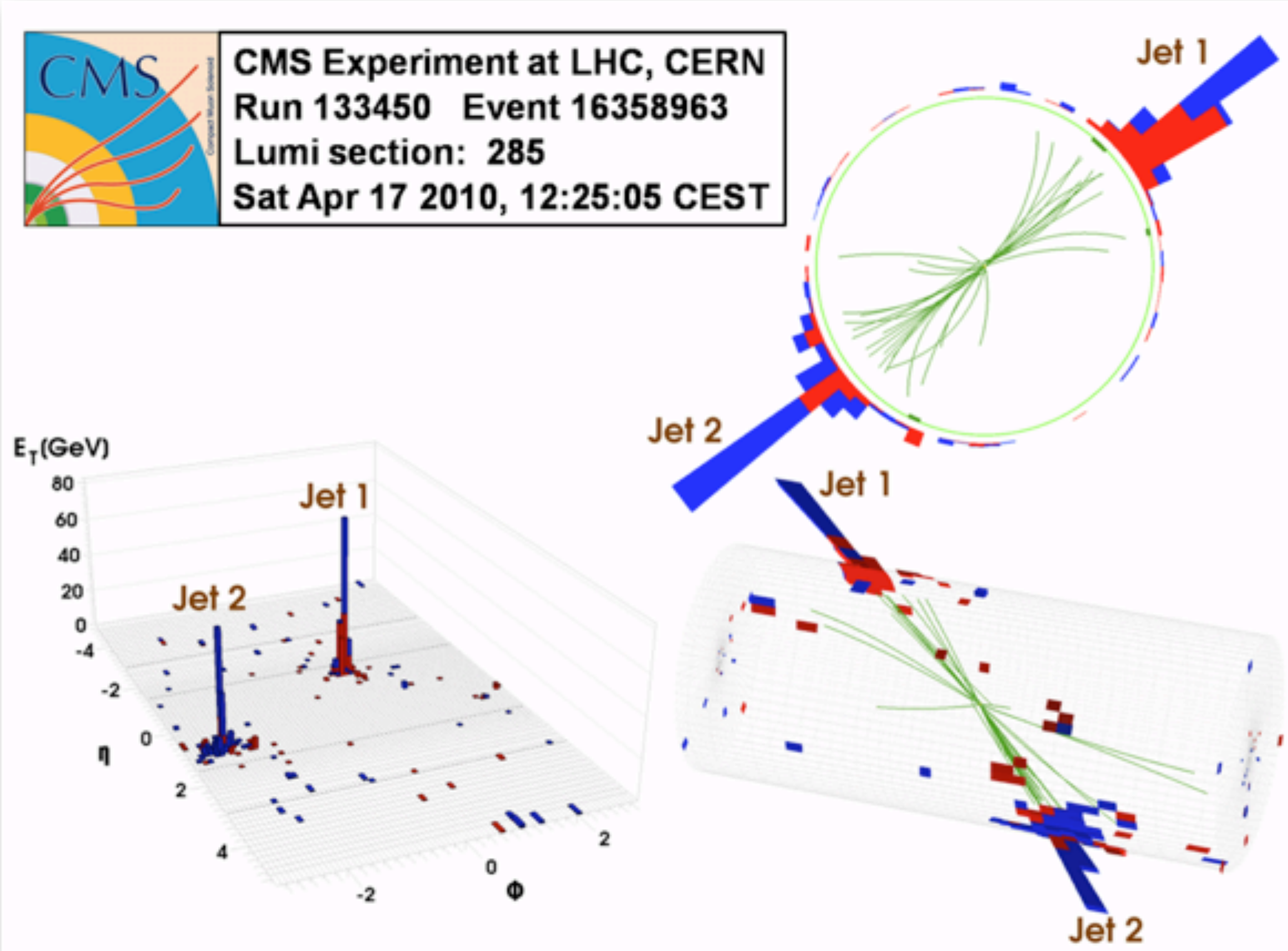
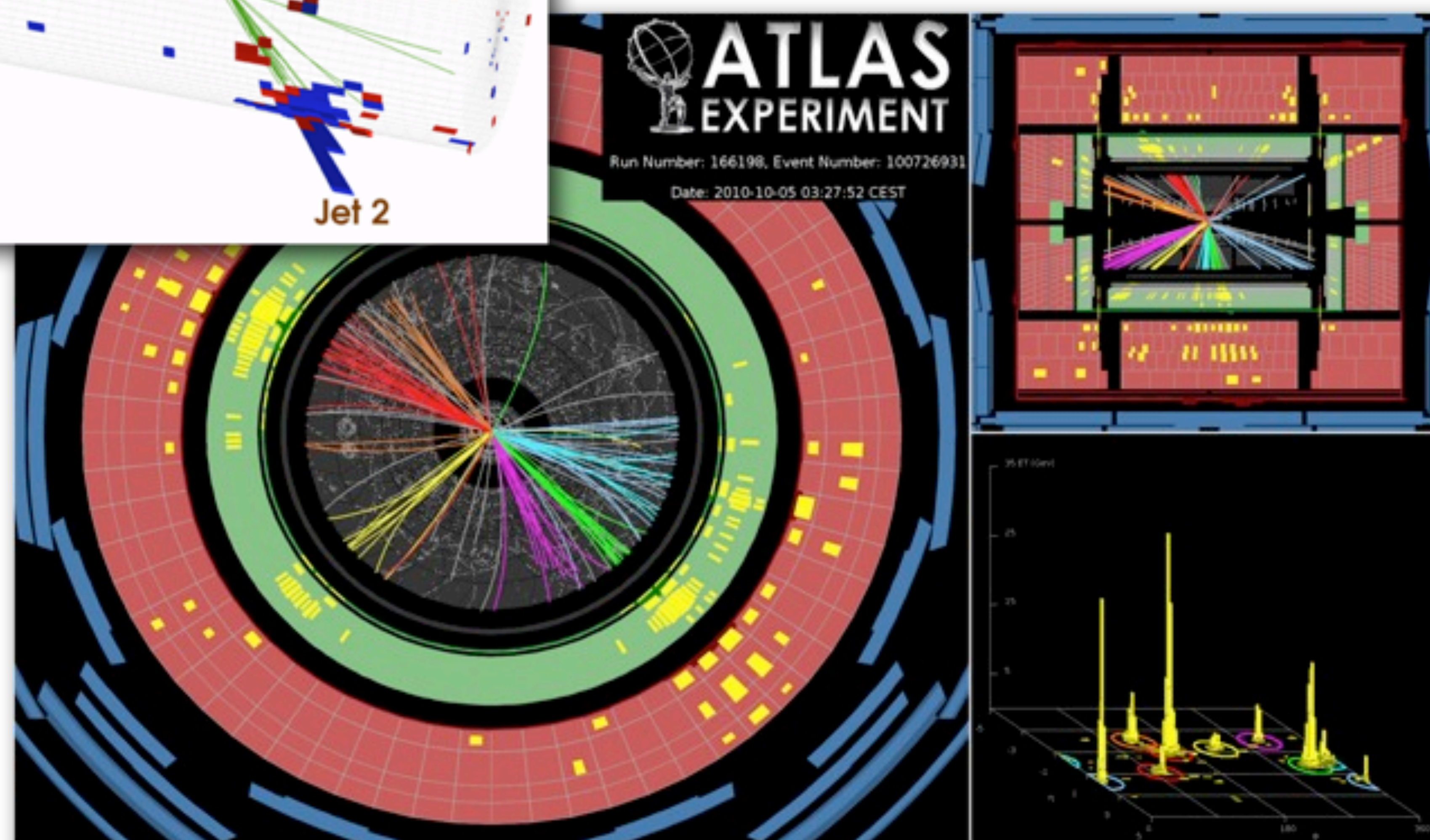


Jet Substructure

Gavin Salam (CERN)

Beyond the LHC
Nordita Workshop
July 2013

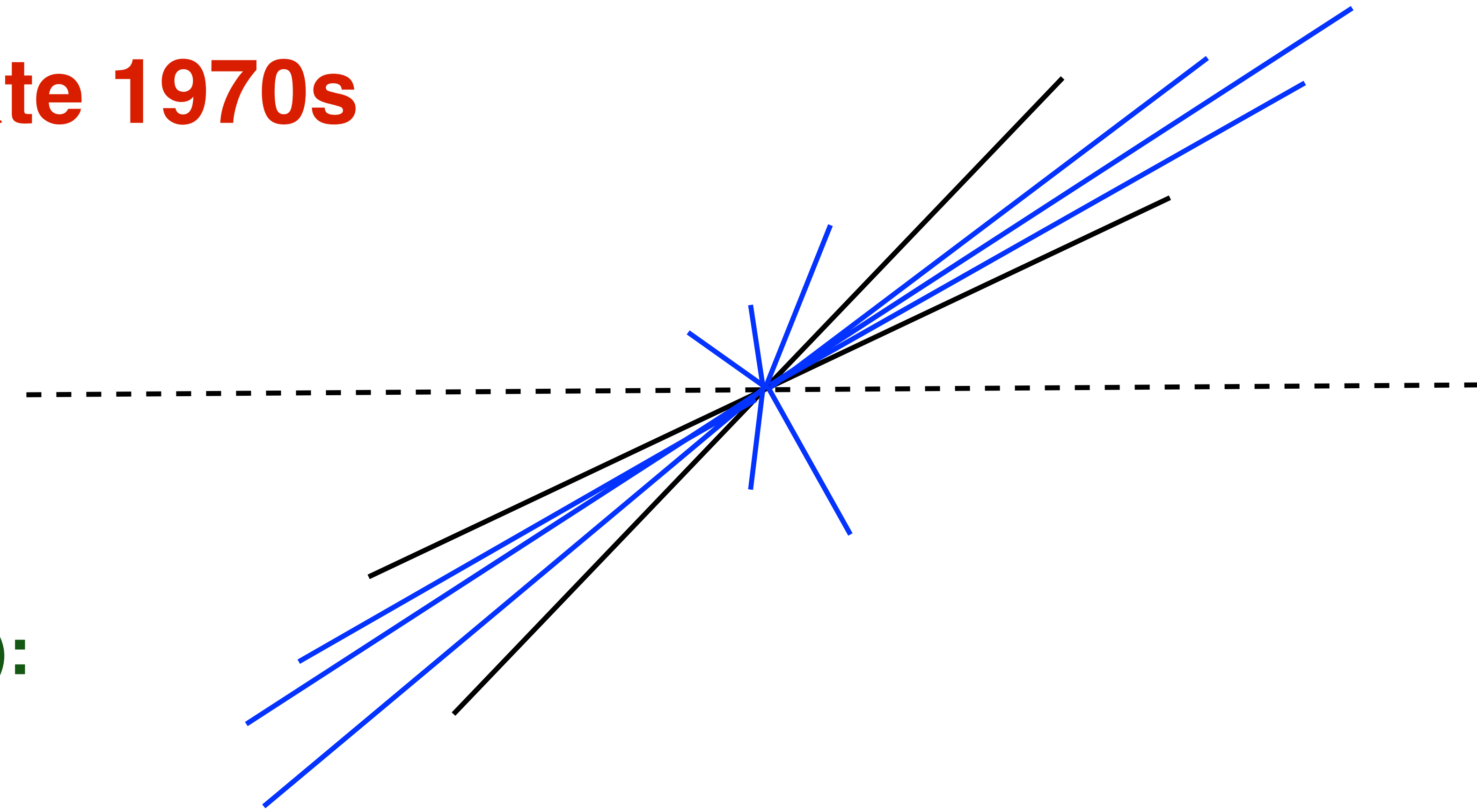




JETS

Collimated, energetic bunches of particles

Jets date back to the late 1970s

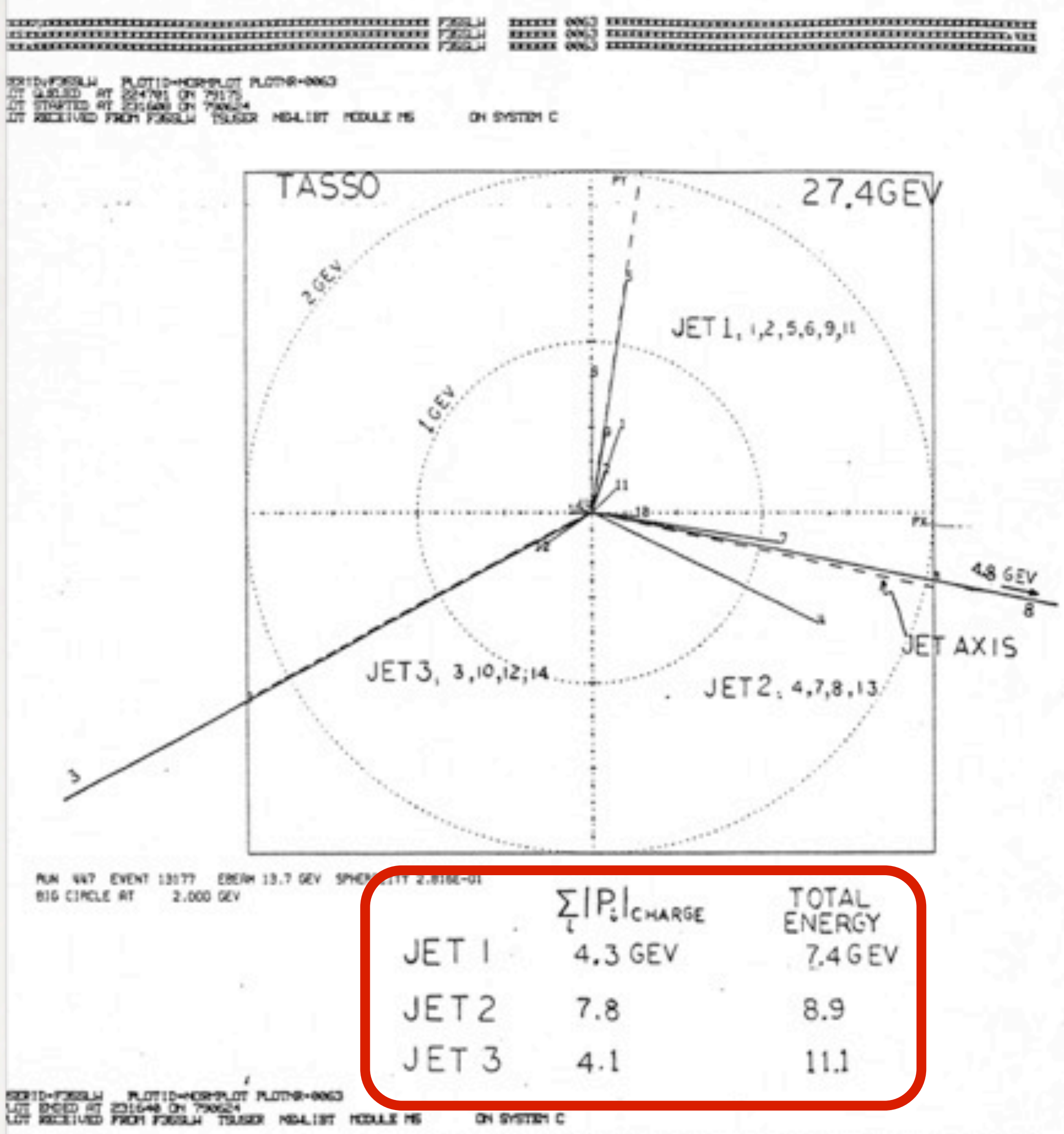


Sterman and Weinberg,
Phys. Rev. Lett. 39, 1436 (1977):

To study jets, we consider the partial cross section $\sigma(E,\theta,\Omega,\varepsilon,\delta)$ for e^+e^- hadron production events, in which all but a fraction $\varepsilon \ll 1$ of the total e^+e^- energy E is emitted within some pair of oppositely directed cones of half-angle $\delta \ll 1$, lying within two fixed cones of solid angle Ω (with $\pi\delta^2 \ll \Omega \ll 1$) at an angle θ to the e^+e^- beam line. We expect this to be measur-

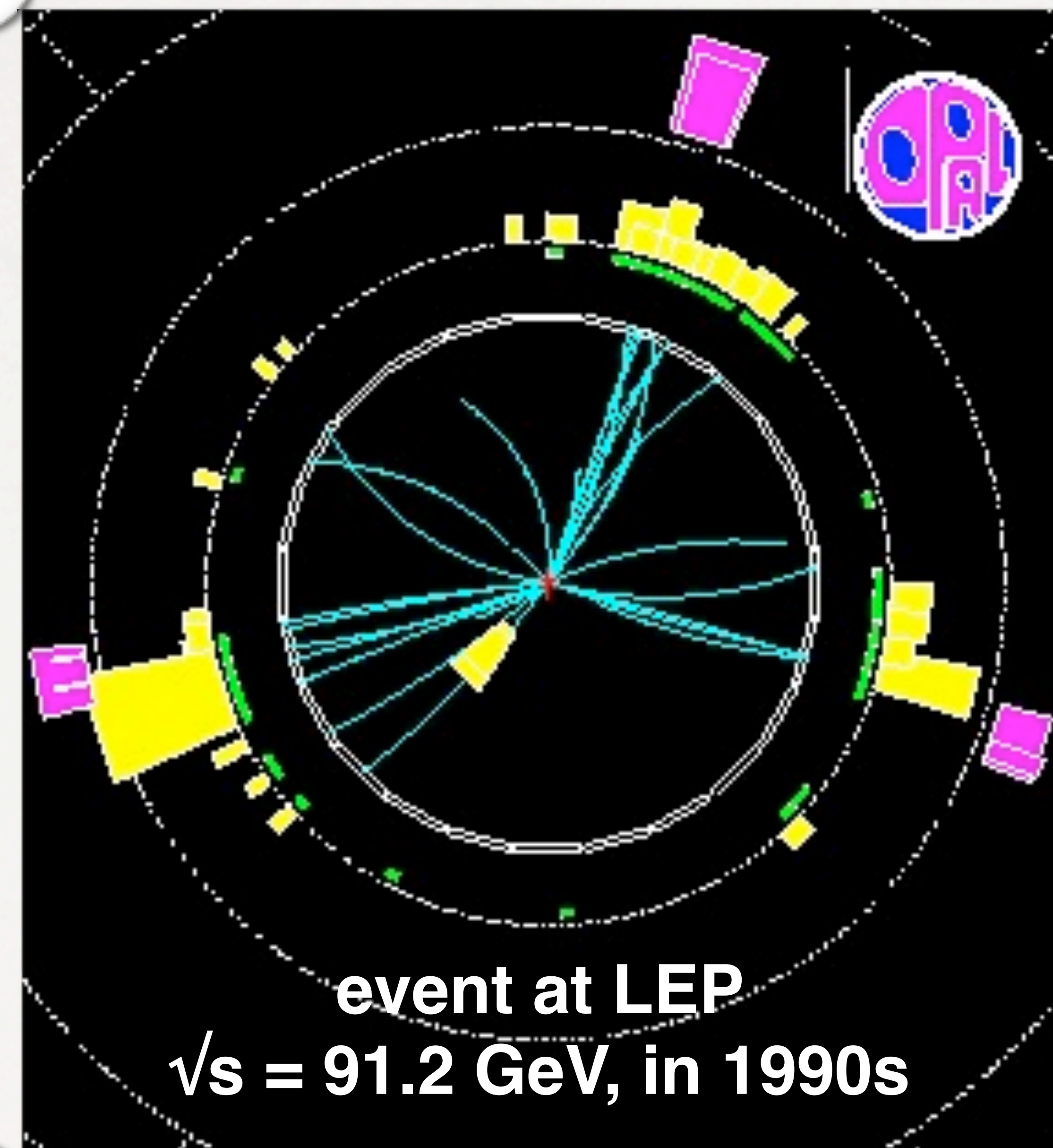
$$\sigma(E,\theta,\Omega,\varepsilon,\delta) = (d\sigma/d\Omega)_0\,\Omega\left[1 - (g_E^2/3\pi^2)\left\{3\ln\delta + 4\ln\delta\,\ln 2\varepsilon + \frac{\pi^3}{3} - \frac{5}{2}\right\}\right]$$

And they've been used and studied at every collider since



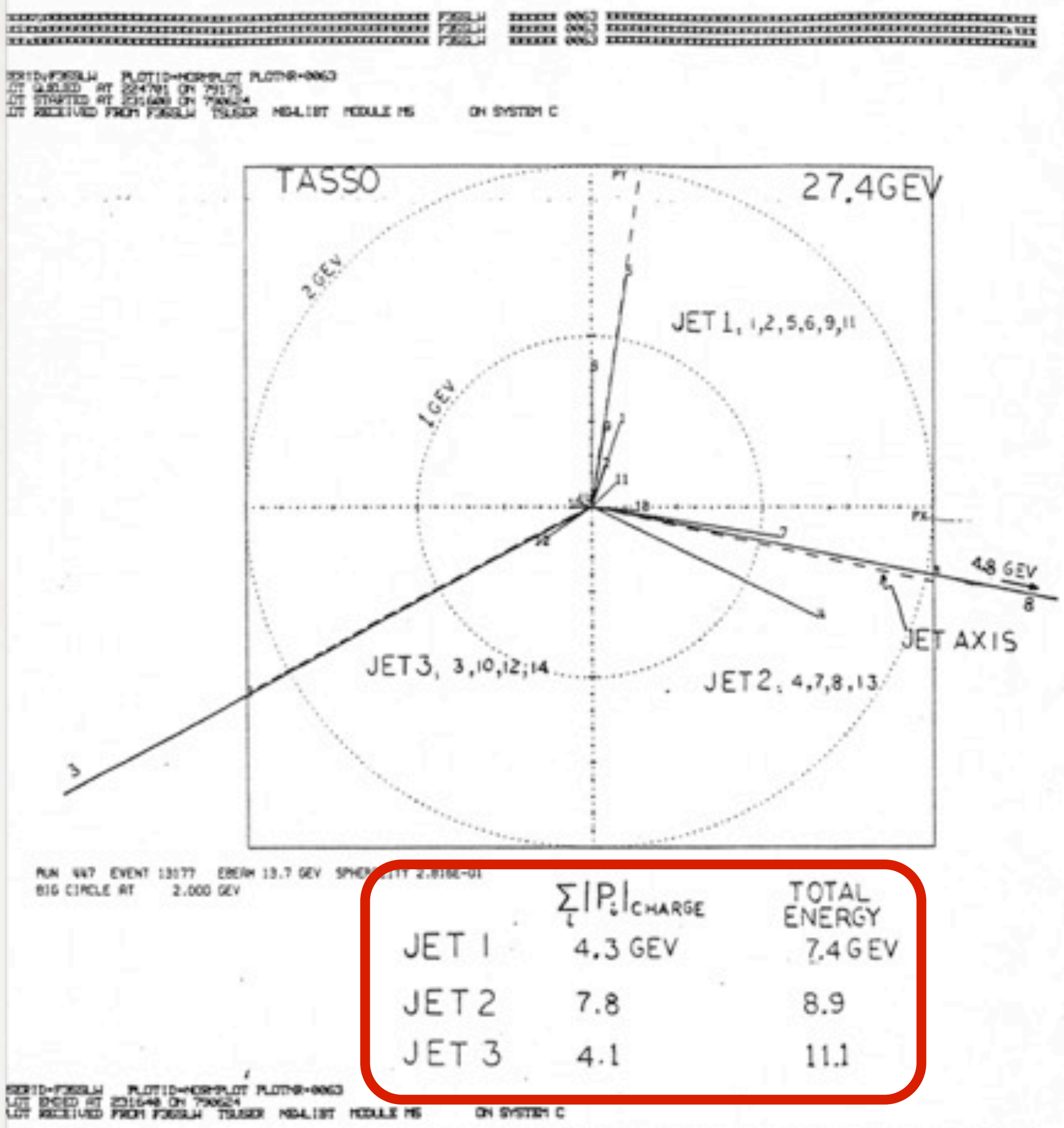
gluon discovery:

event at TASSO
$\sqrt{s}$ = 27.4 GeV, in 1979



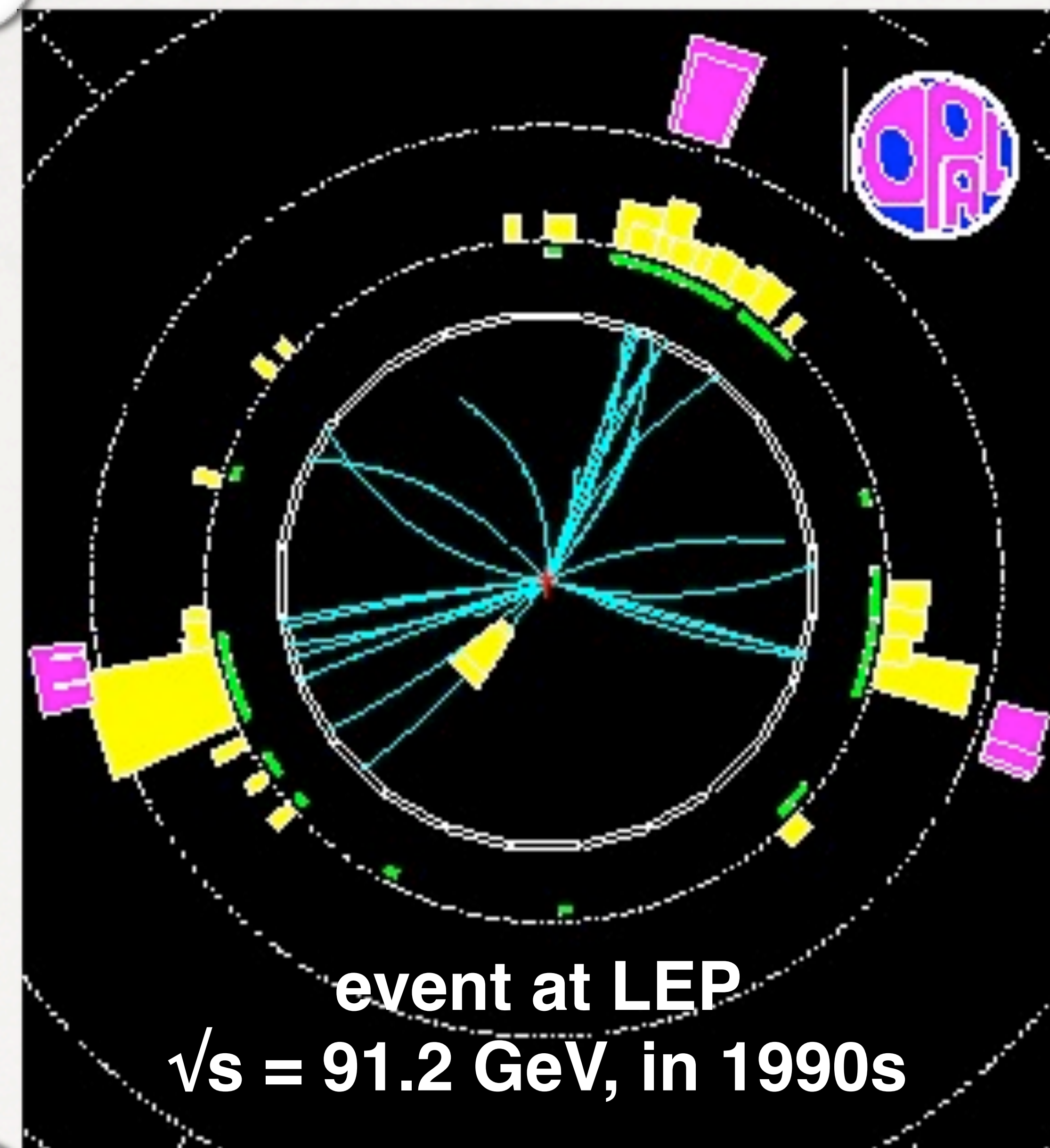
event at LEP
$\sqrt{s}$ = 91.2 GeV, in 1990s

And they've been used and studied at every collider since



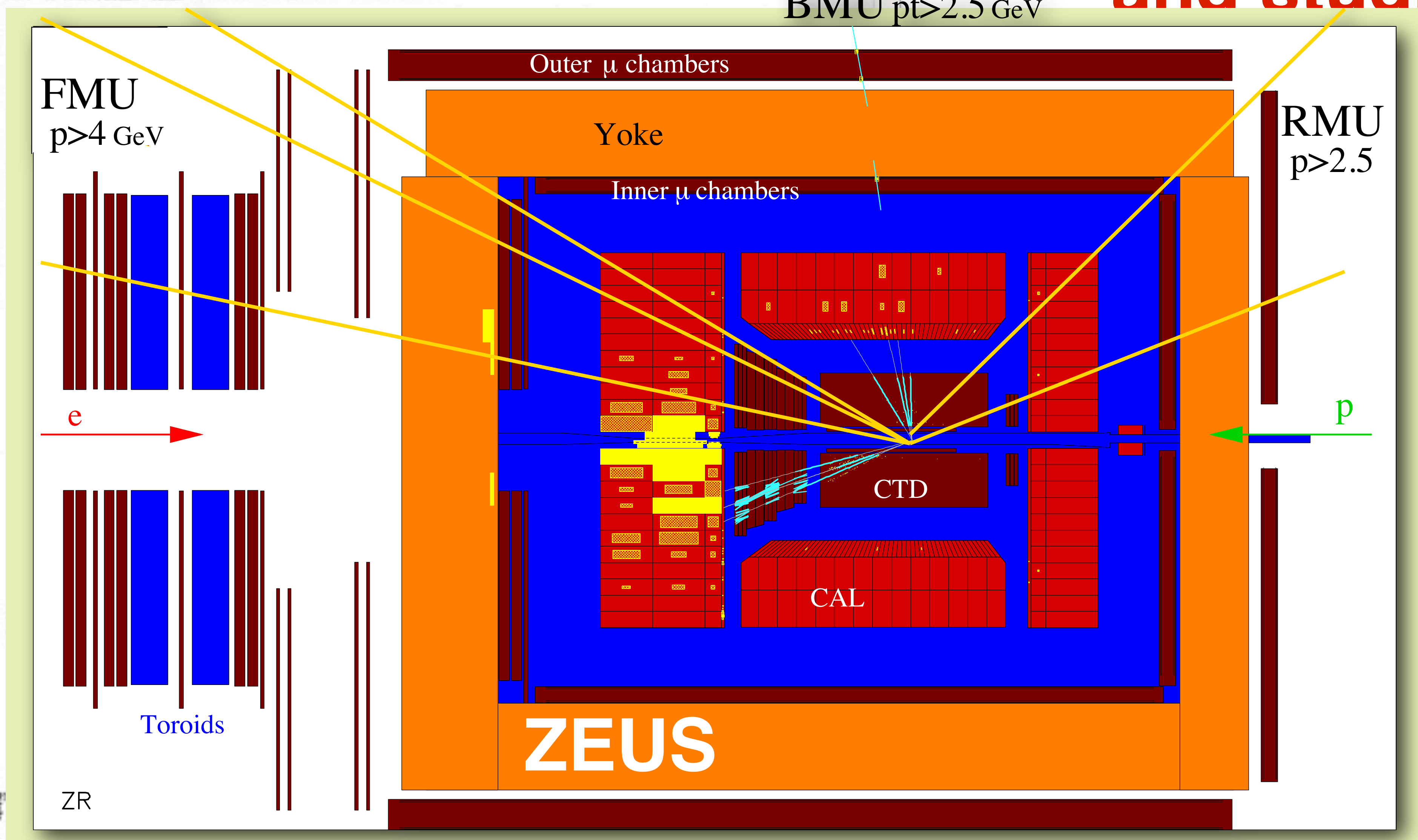
gluon discovery:

event at TASSO
$\sqrt{s}$ = 27.4 GeV, in 1979



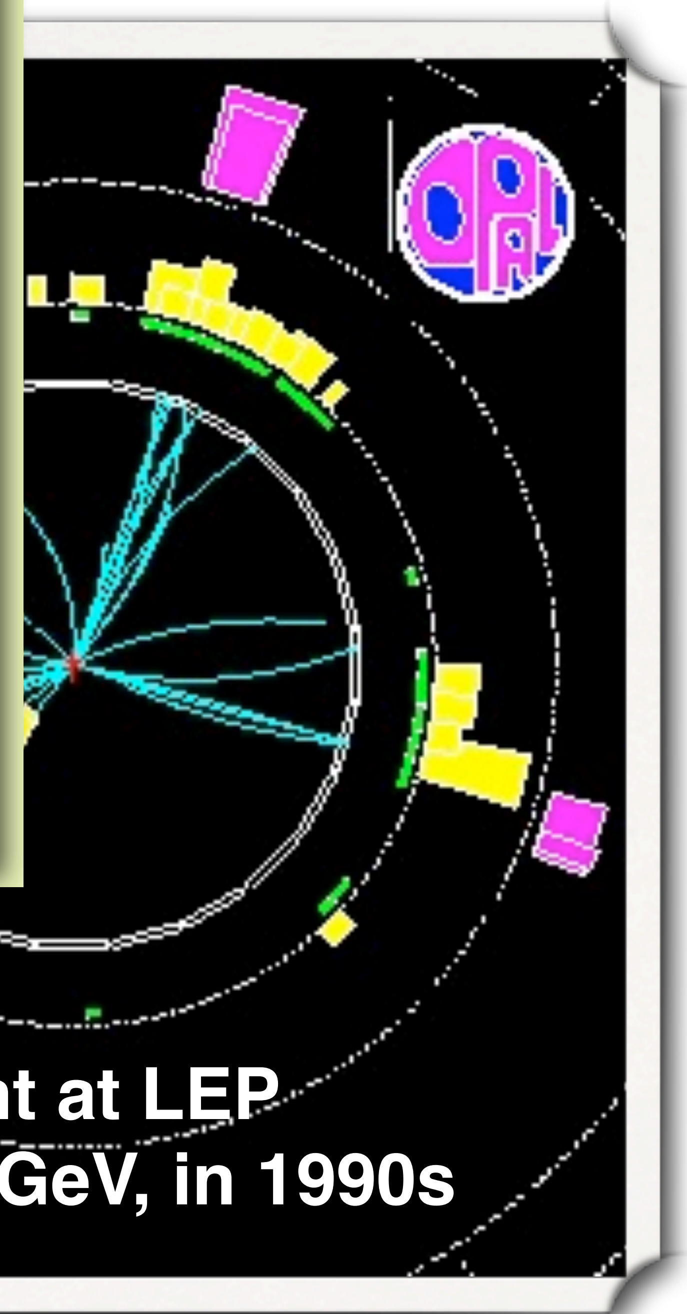
event at LEP
$\sqrt{s}$ = 91.2 GeV, in 1990s

And they've been used and studied at every ... er since



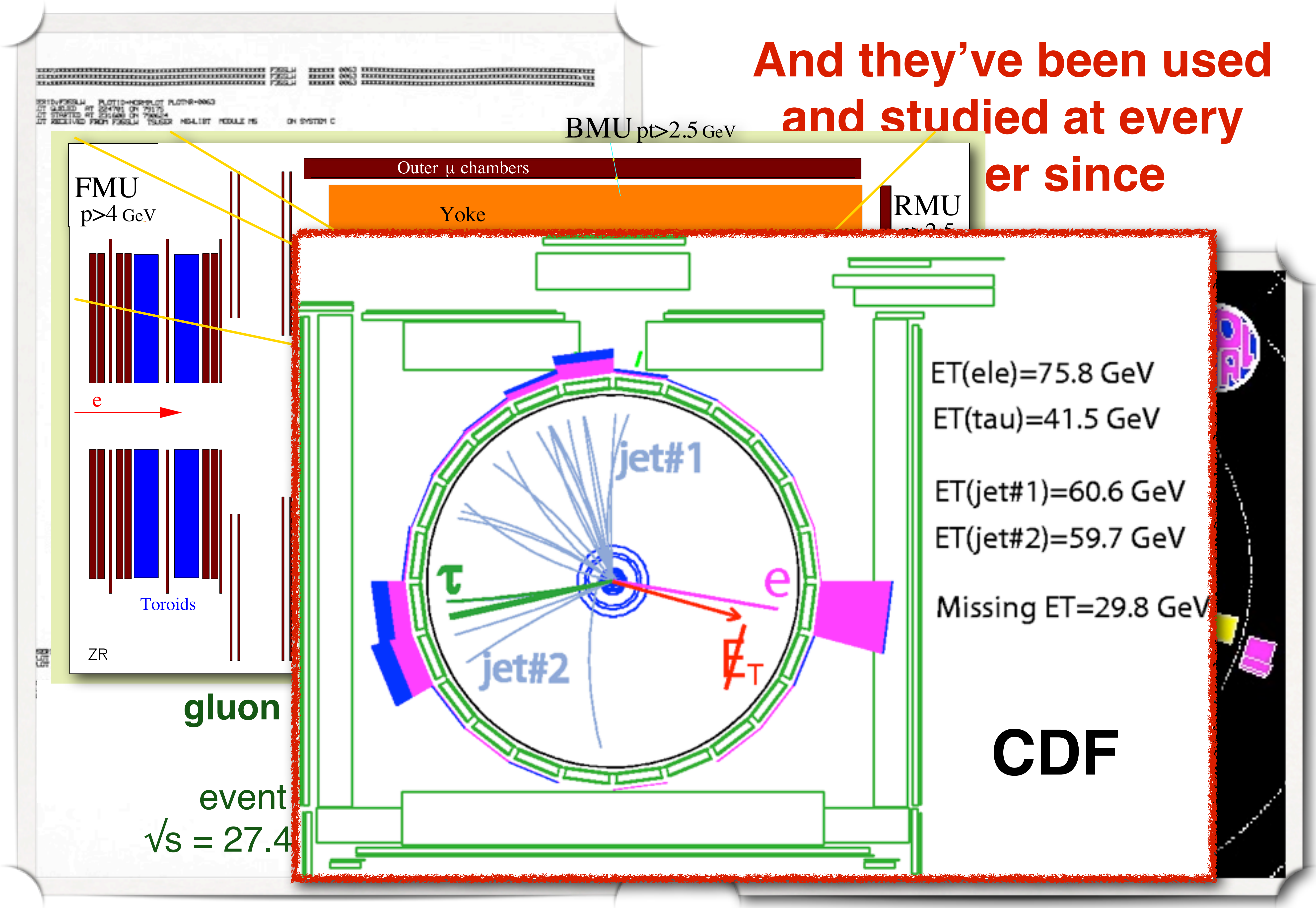
gluon discovery:

event at TASSO
$\sqrt{s}$ = 27.4 GeV, in 1979

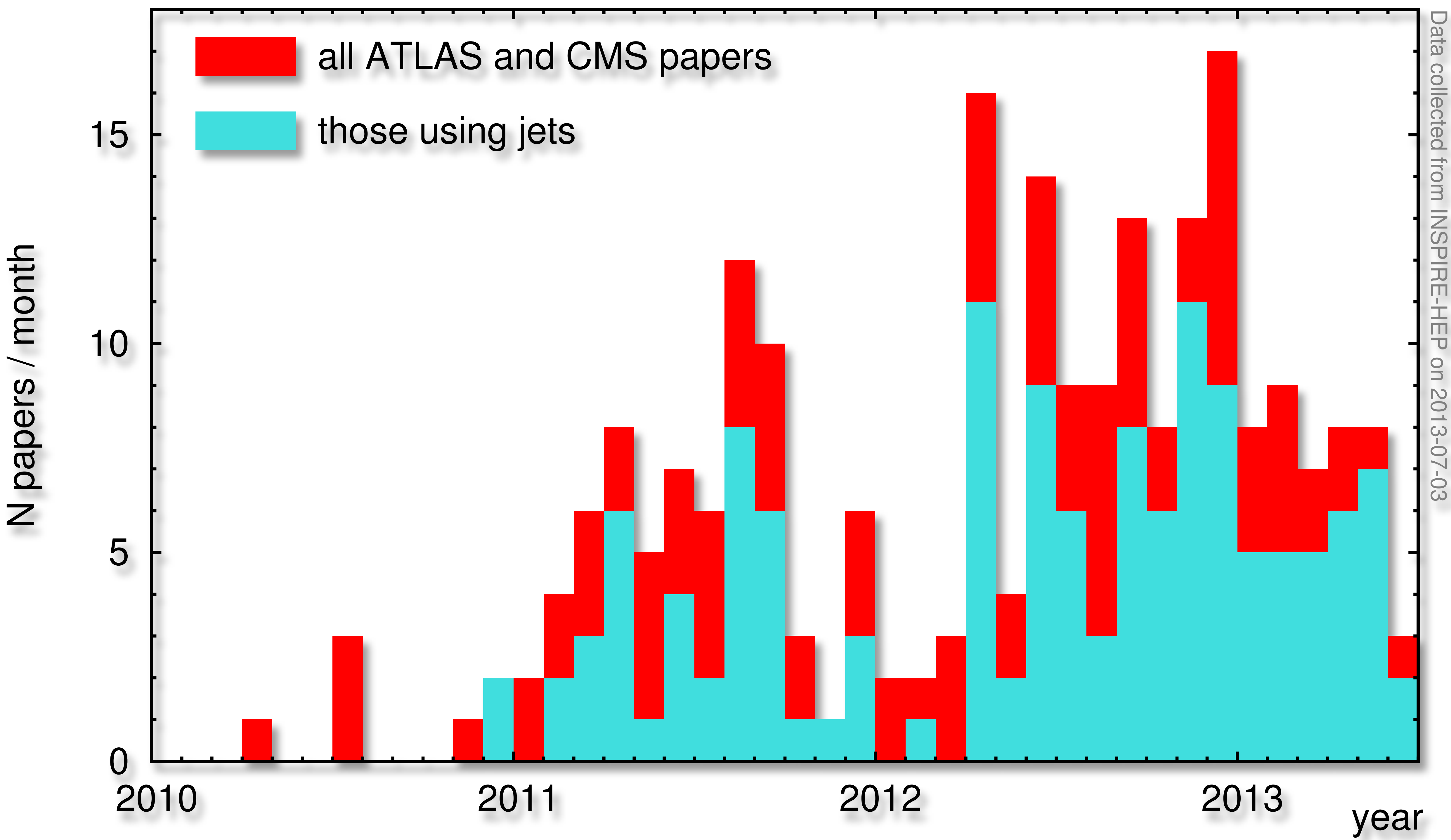


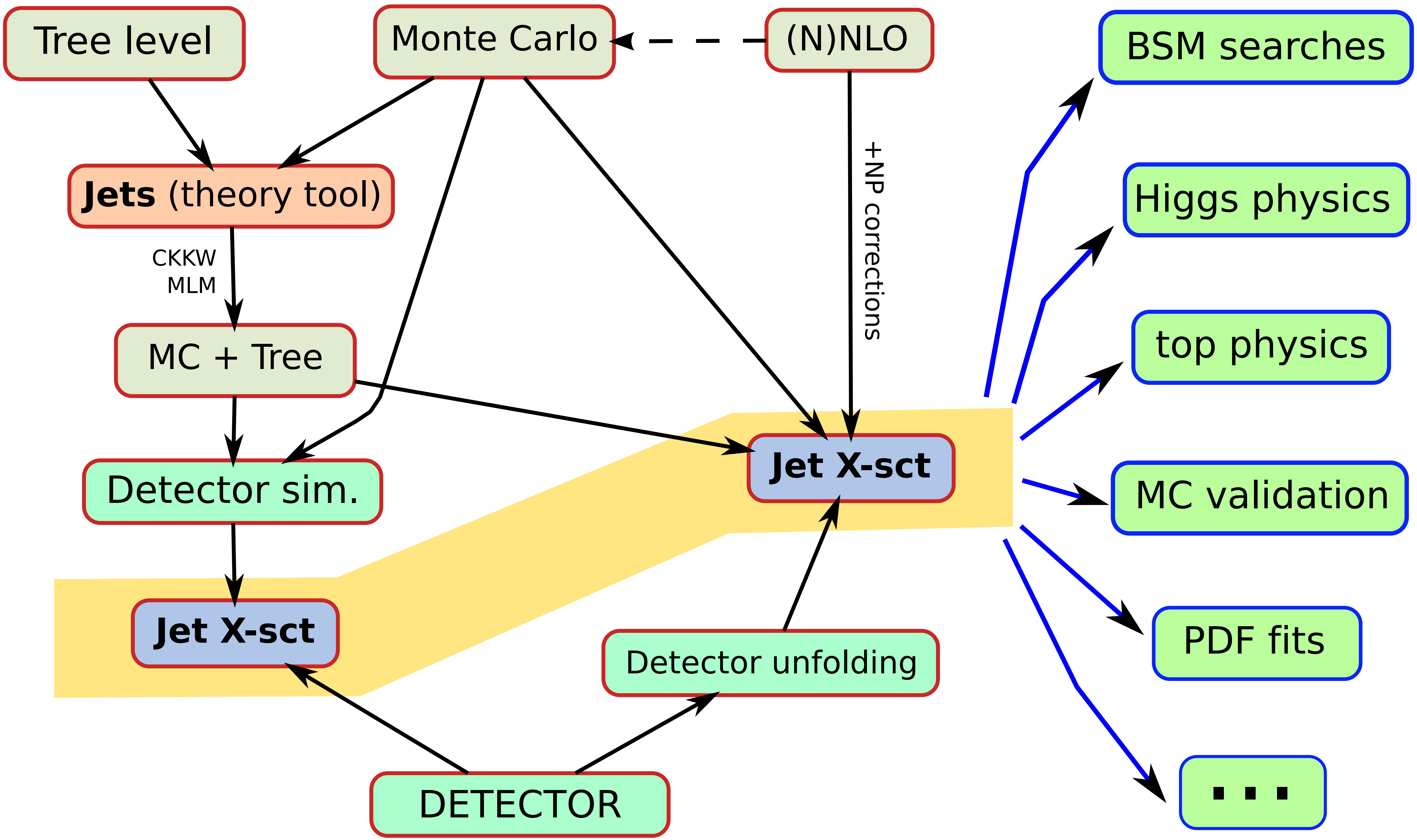
event at LEP
$\sqrt{s}$ = 91.2 GeV, in 1990s

And they’ve been used and studied at every [...]er since



60-70% of recent ATLAS and CMS papers use jets in their analyses




Tree level
Monte Carlo
(N)NLO
BSM searches
Jets (theory tool)
Higgs physics
+NP corrections
CKKW
MLM
MC + Tree
top physics
Jet X-sct
Detector sim.
MC validation
Jet X-sct
PDF fits
Detector unfolding
DETECTOR
...

Most LHC jet uses fall under the (Tevatron-like) category

“a jet is basically a parton”

e.g. from a heavy-object decay, ISR, etc.

If radiation is modelled correctly in the Monte Carlos,
most experimenters don’t even need to think (much) about jets.

Most LHC jet uses fall under the (Tevatron-like) category

"a jet is basically a parton"

e.g. from a heavy-object decay, ISR, etc.

If radiation is modelled correctly in the Monte Carlos,
most experimenters don't even need to think (much) about jets.

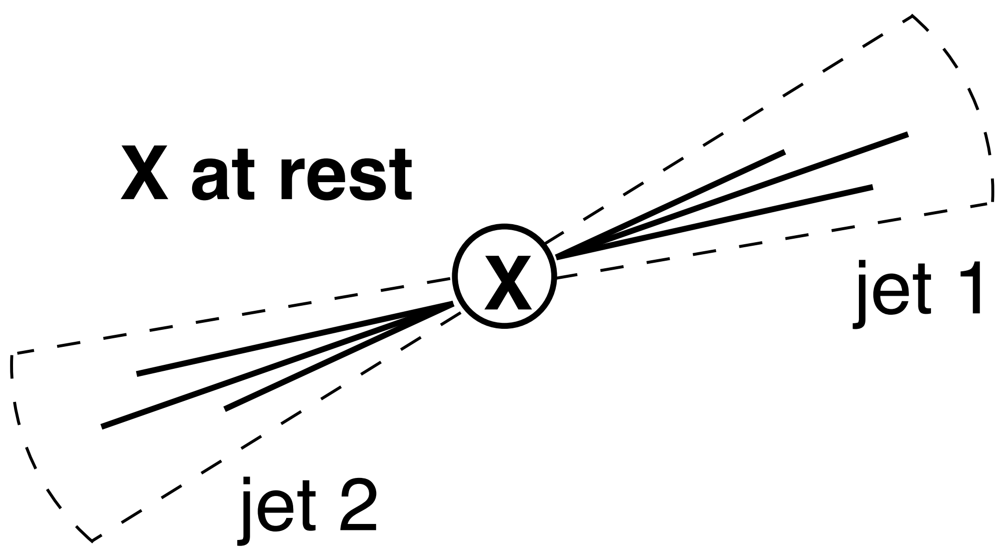
But with new physics proving elusive, we start to need to push analyses to their boundary, e.g.

Enhance sensitivity to small signal/background
Explore very highest p_t's
Learn how to handle complex final states

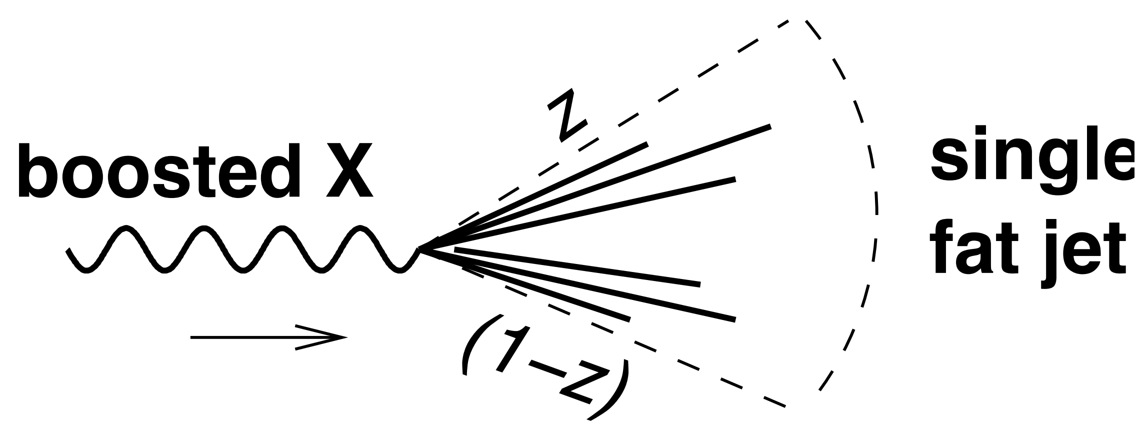
➜ for that, you need advanced jet techniques

Boosted hadronic decays

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

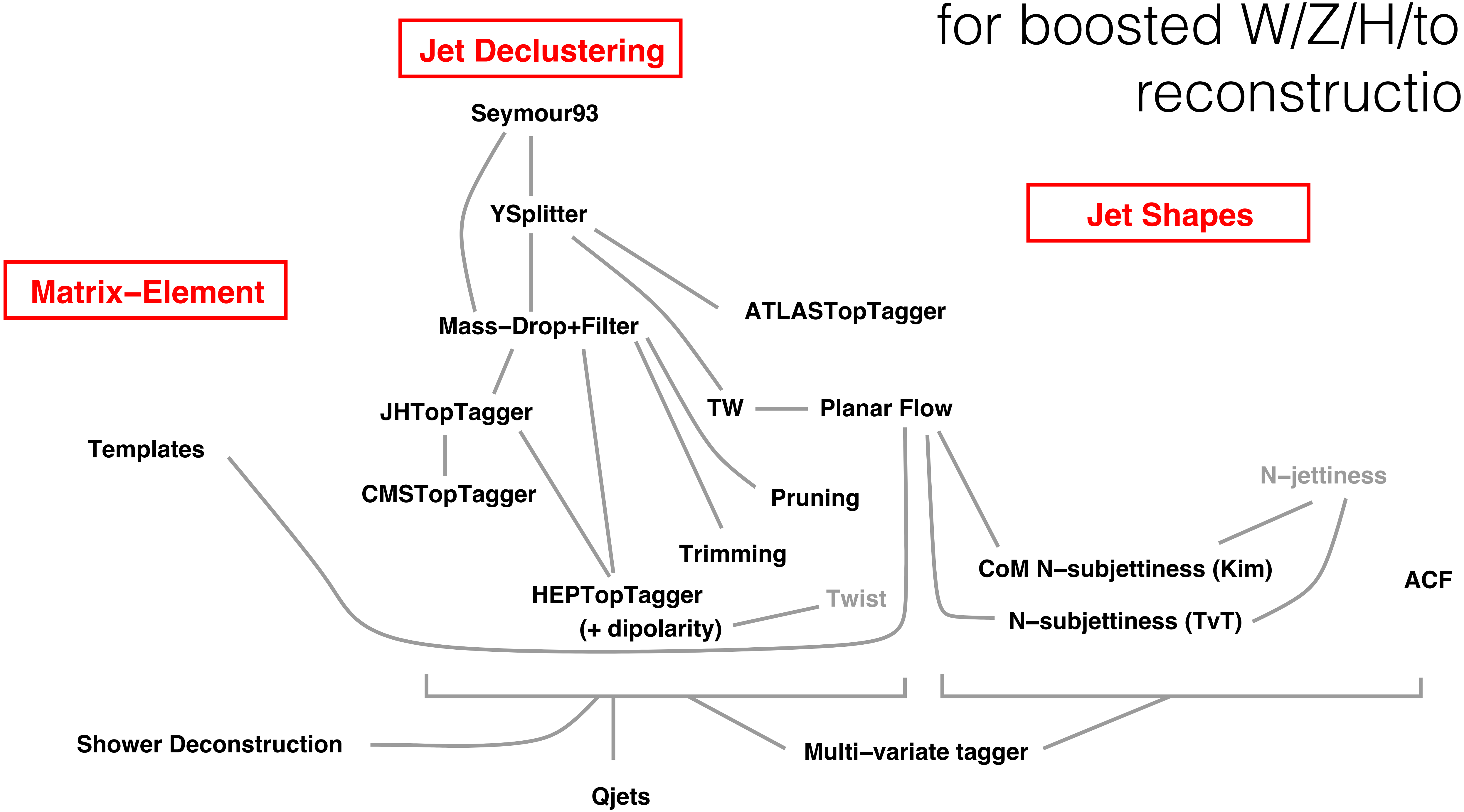


Happens for $p_t \gtrsim 2m/R$

$p_t \gtrsim 320$ GeV for $m = m_W$, $R = 0.5$

Very active research field

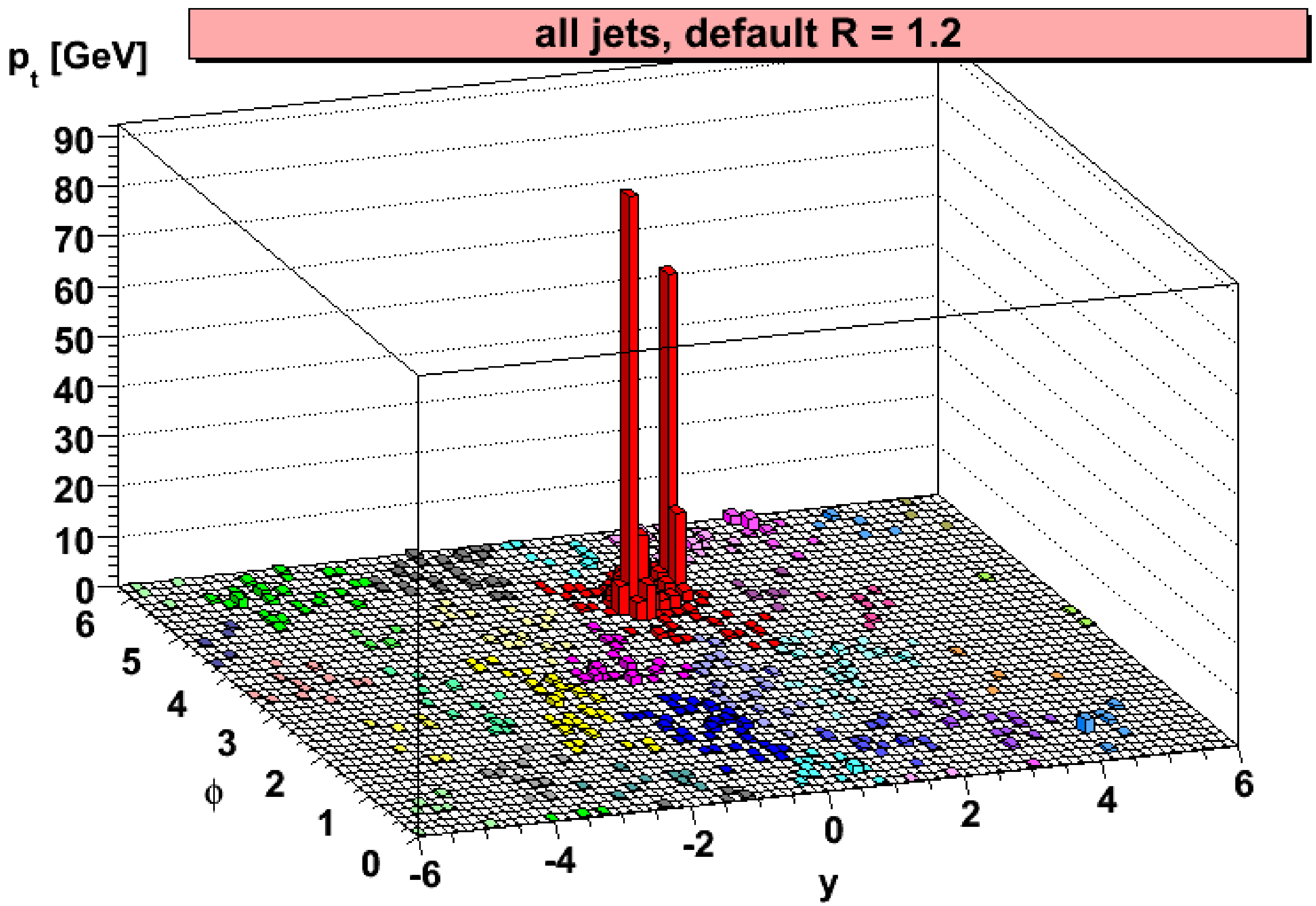
Some of the tools developed for boosted W/Z/H/top reconstruction



apologies for omitted taggers, arguable links, etc.

SIGNAL

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



Zbb BACKGROUND

Cluster event, C/A, R=1.2

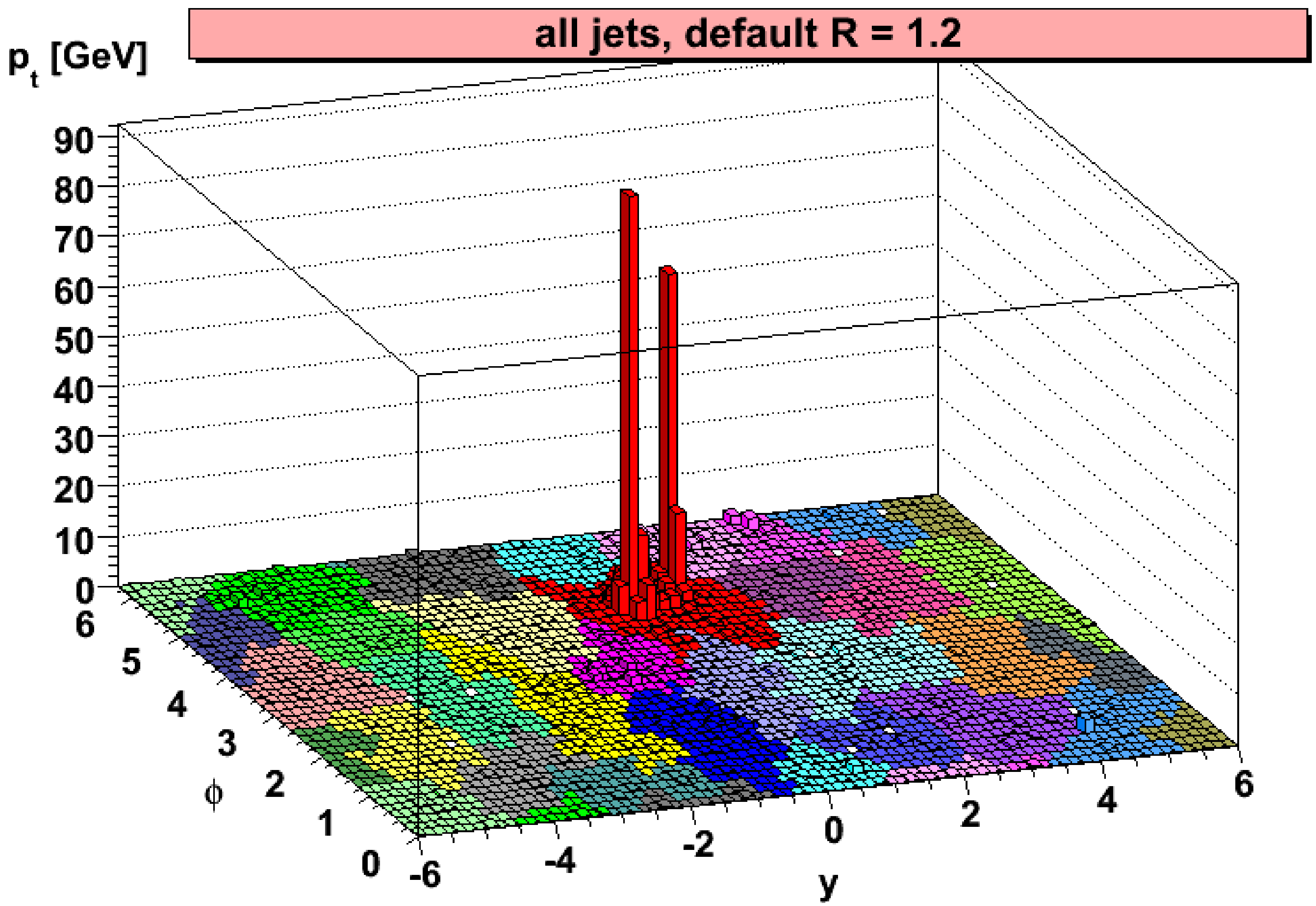
Butterworth, Davison, Rubin & GPS '08

arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

SIGNAL

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



Zbb BACKGROUND

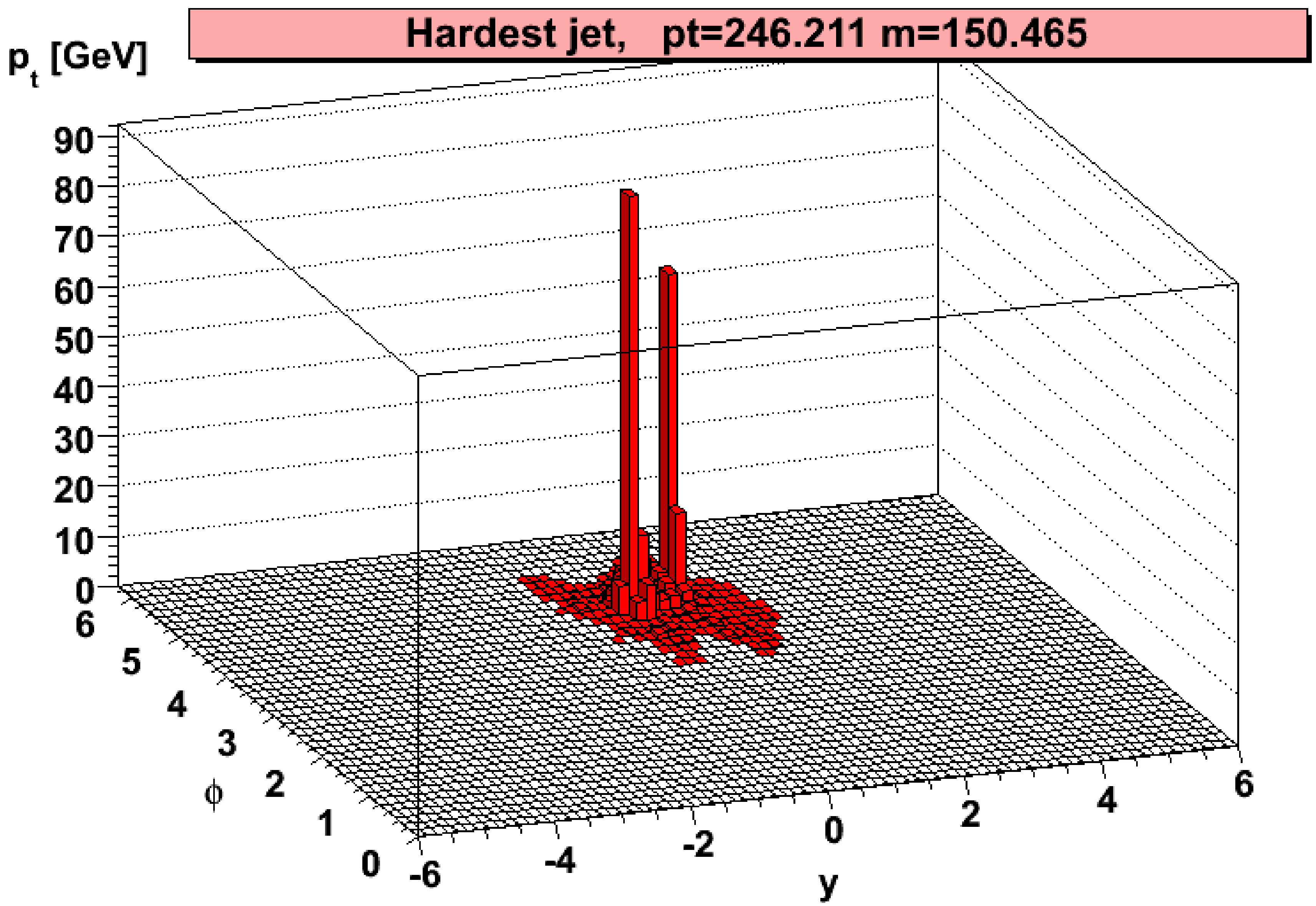
Fill it in, → show jets more clearly

Butterworth, Davison, Rubin & GPS '08

arbitrary norm.

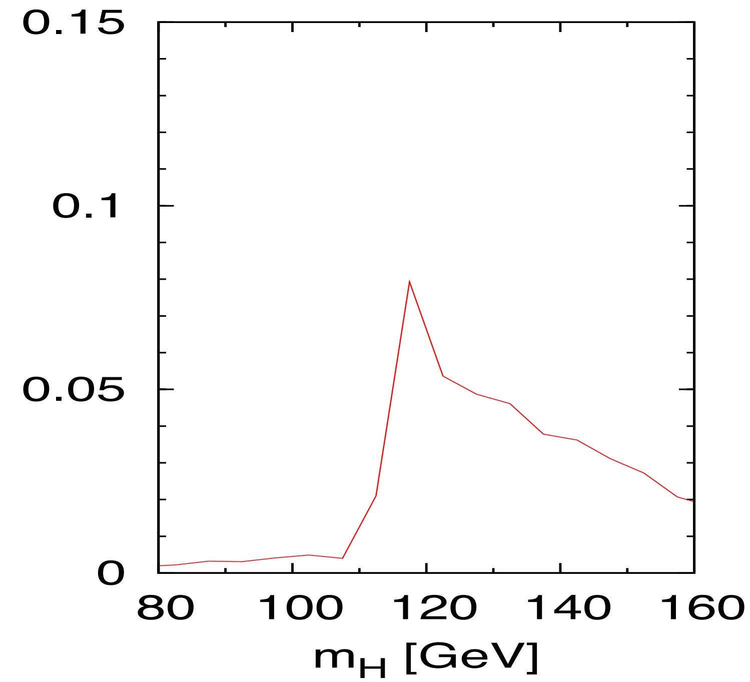
$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



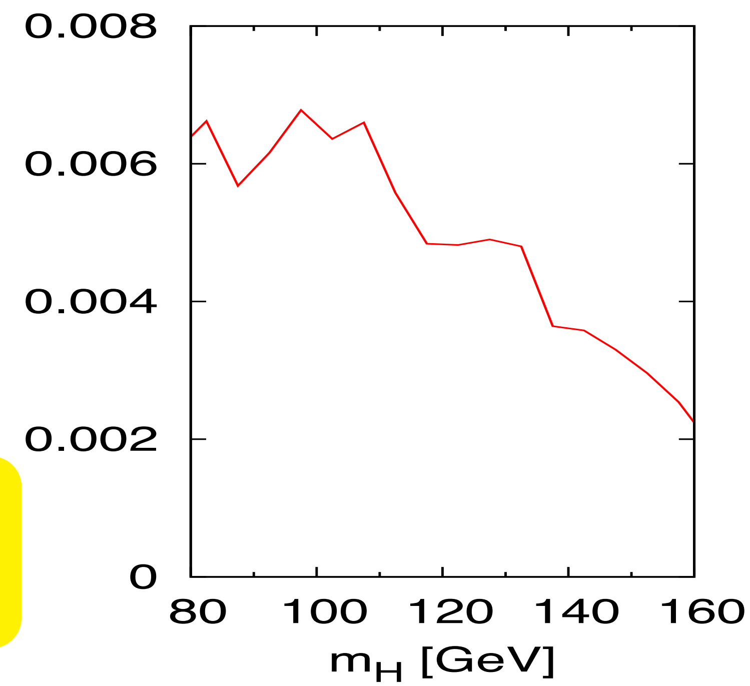
SIGNAL

$200 < p_{tZ} < 250$ GeV



Zbb BACKGROUND

$200 < p_{tZ} < 250$ GeV



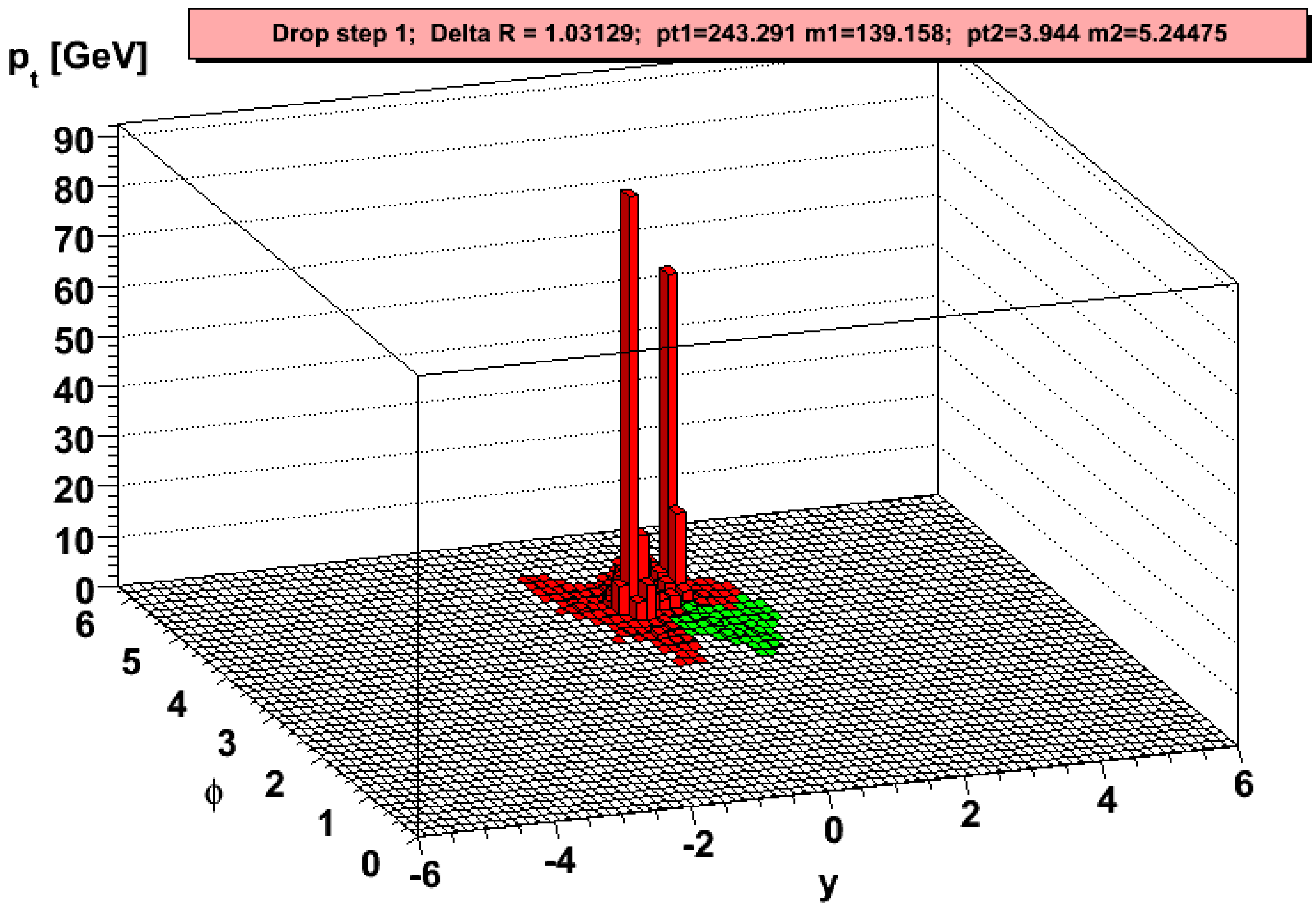
Consider hardest jet, $m = 150$ GeV

Butterworth, Davison, Rubin & GPS '08

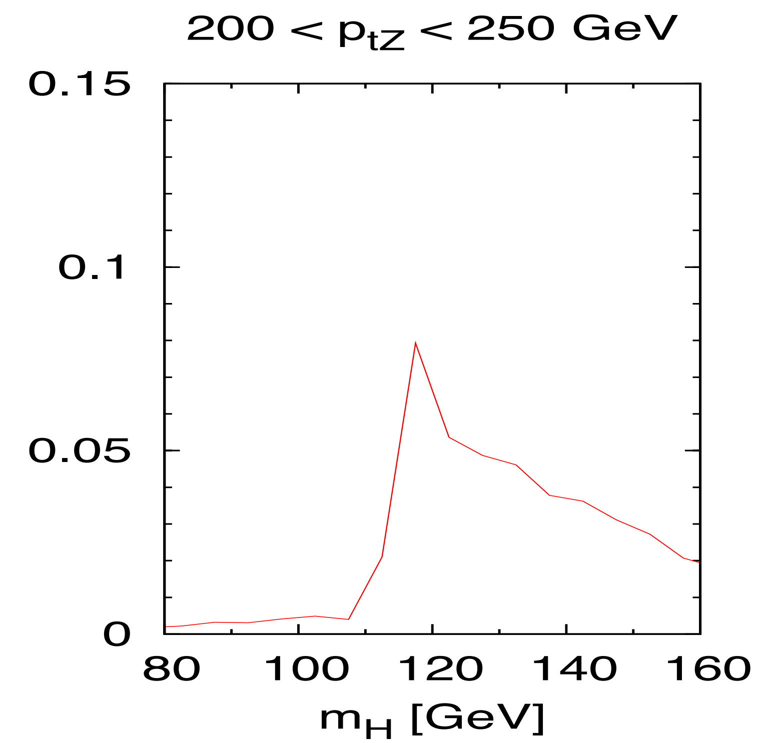
arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

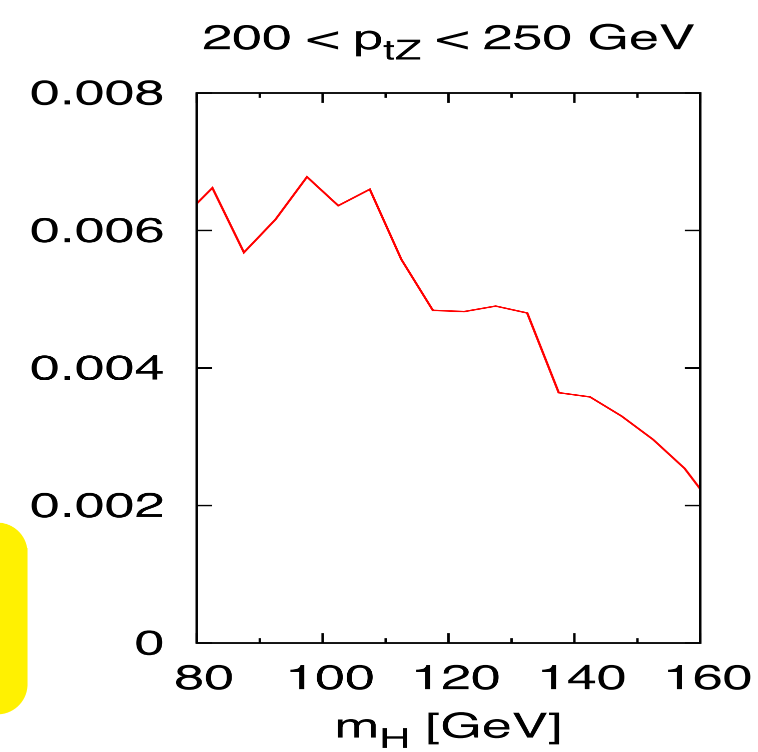
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND



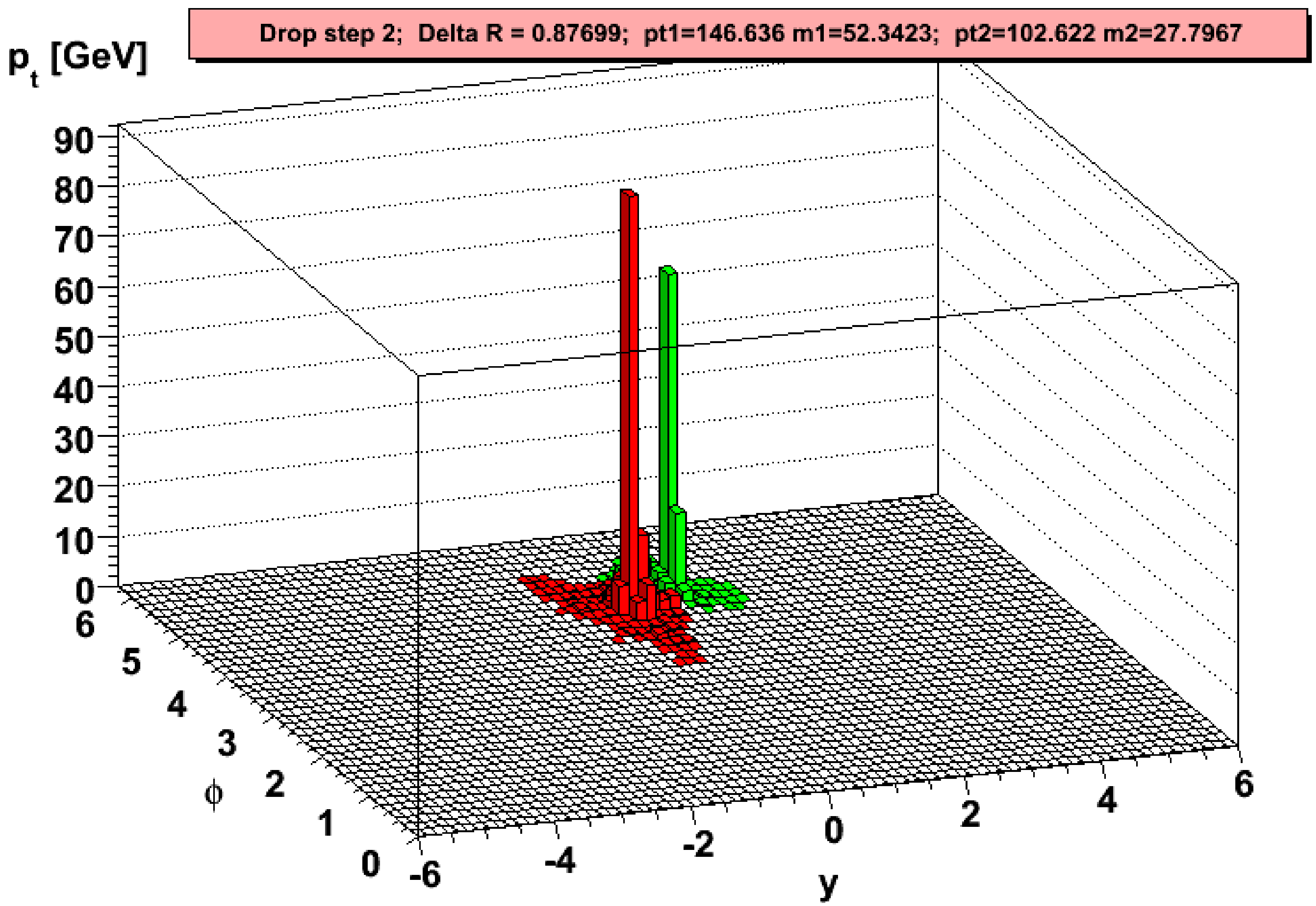
split: $m = 150$ GeV, $\frac{\max(m_1, m2)}{m} = 0.92 \to$ repeat

Butterworth, Davison, Rubin & GPS '08

arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

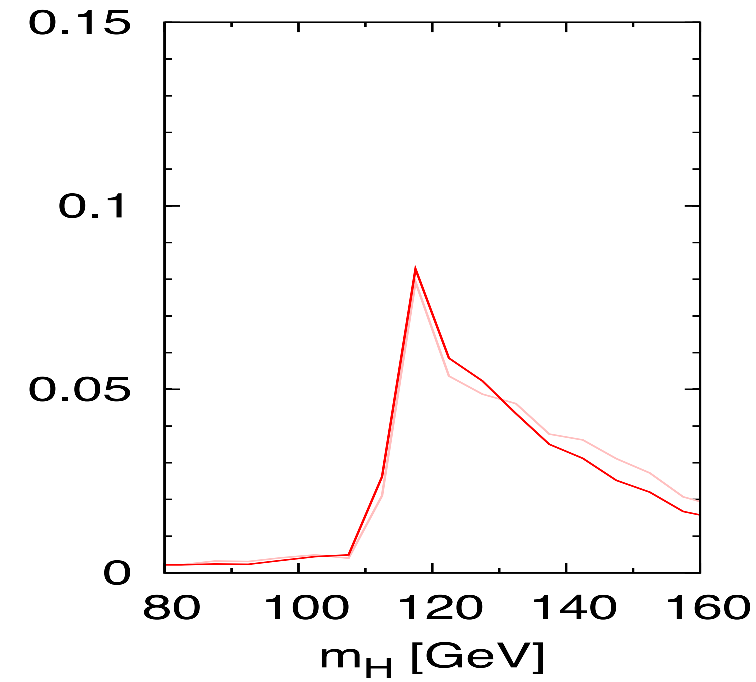


split: $m = 139$ GeV, $\frac{\max(m_1, m2)}{m} = 0.37 \to$ mass drop

Butterworth, Davison, Rubin & GPS '08

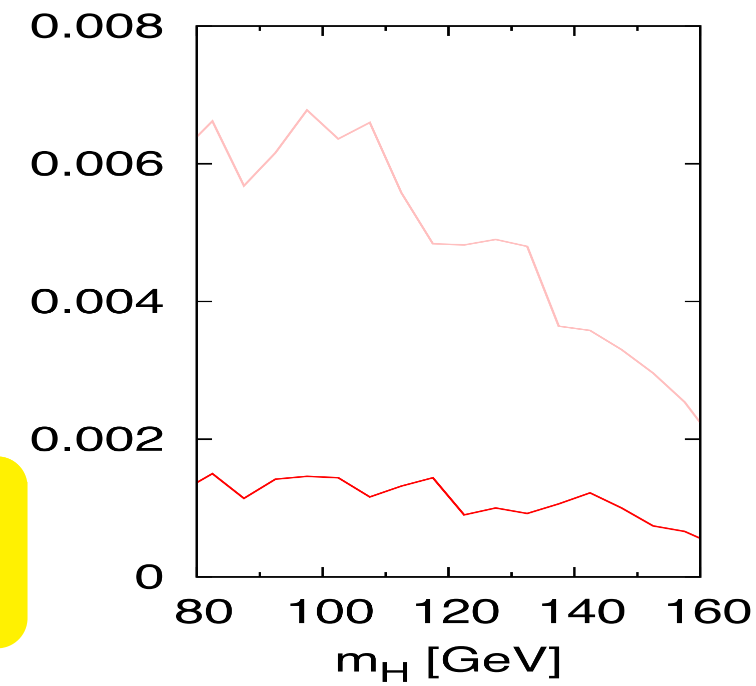
SIGNAL

$200 < p_{tZ} < 250$ GeV



Zbb BACKGROUND

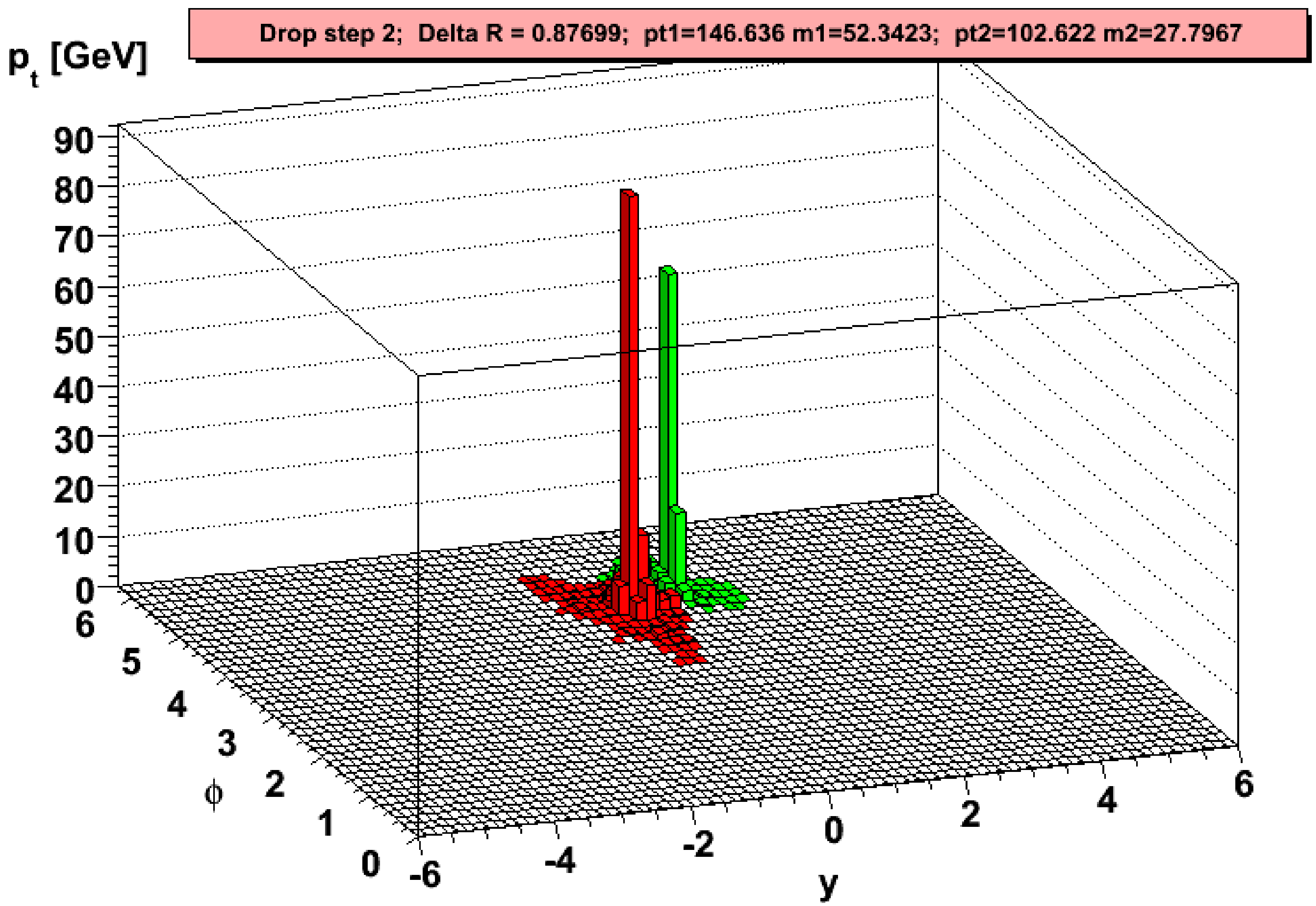
$200 < p_{tZ} < 250$ GeV

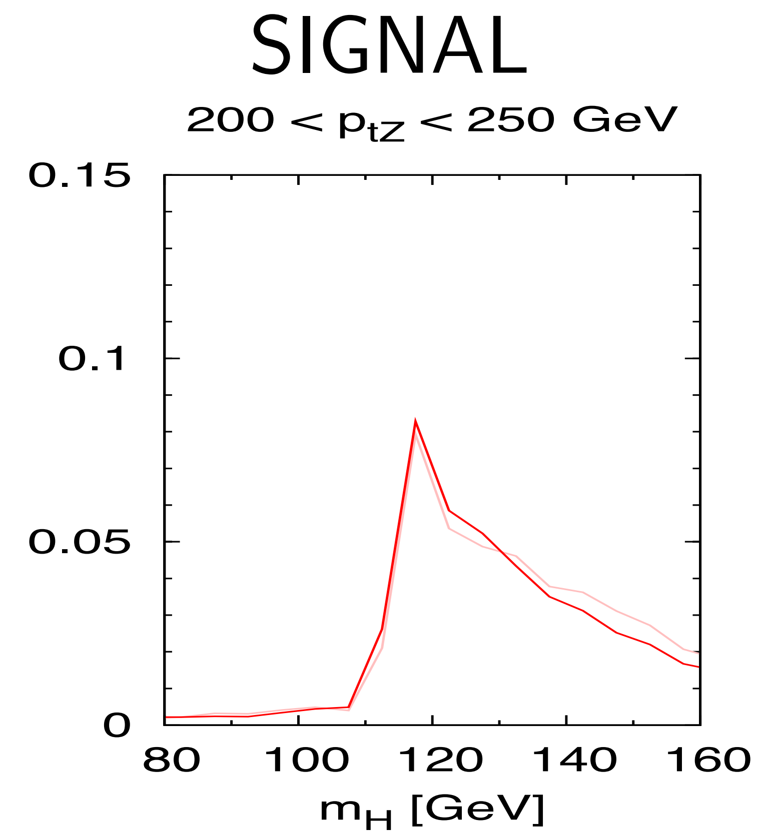


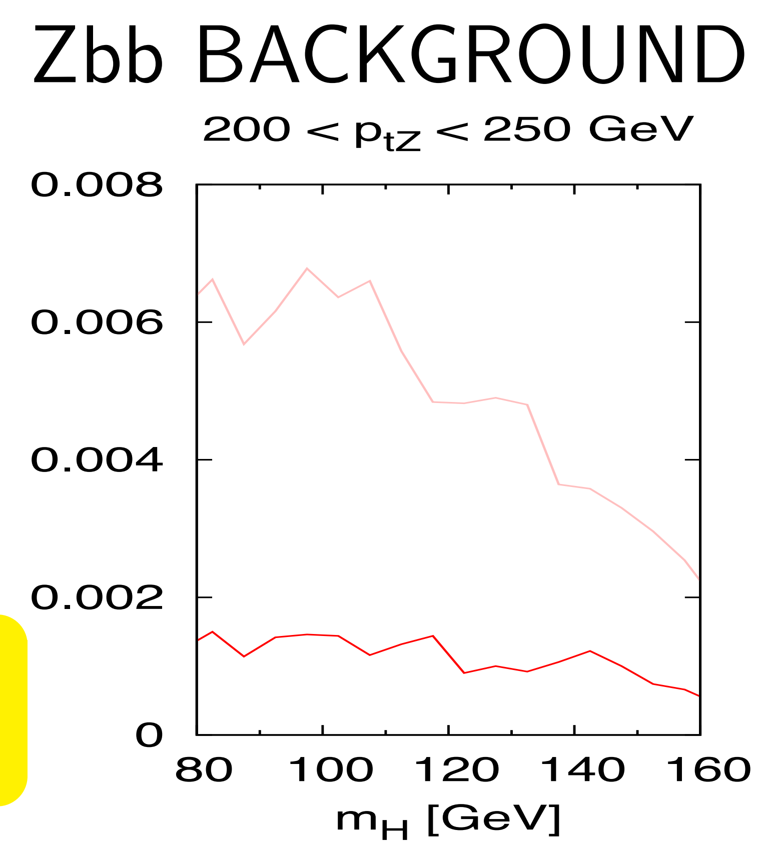
arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3







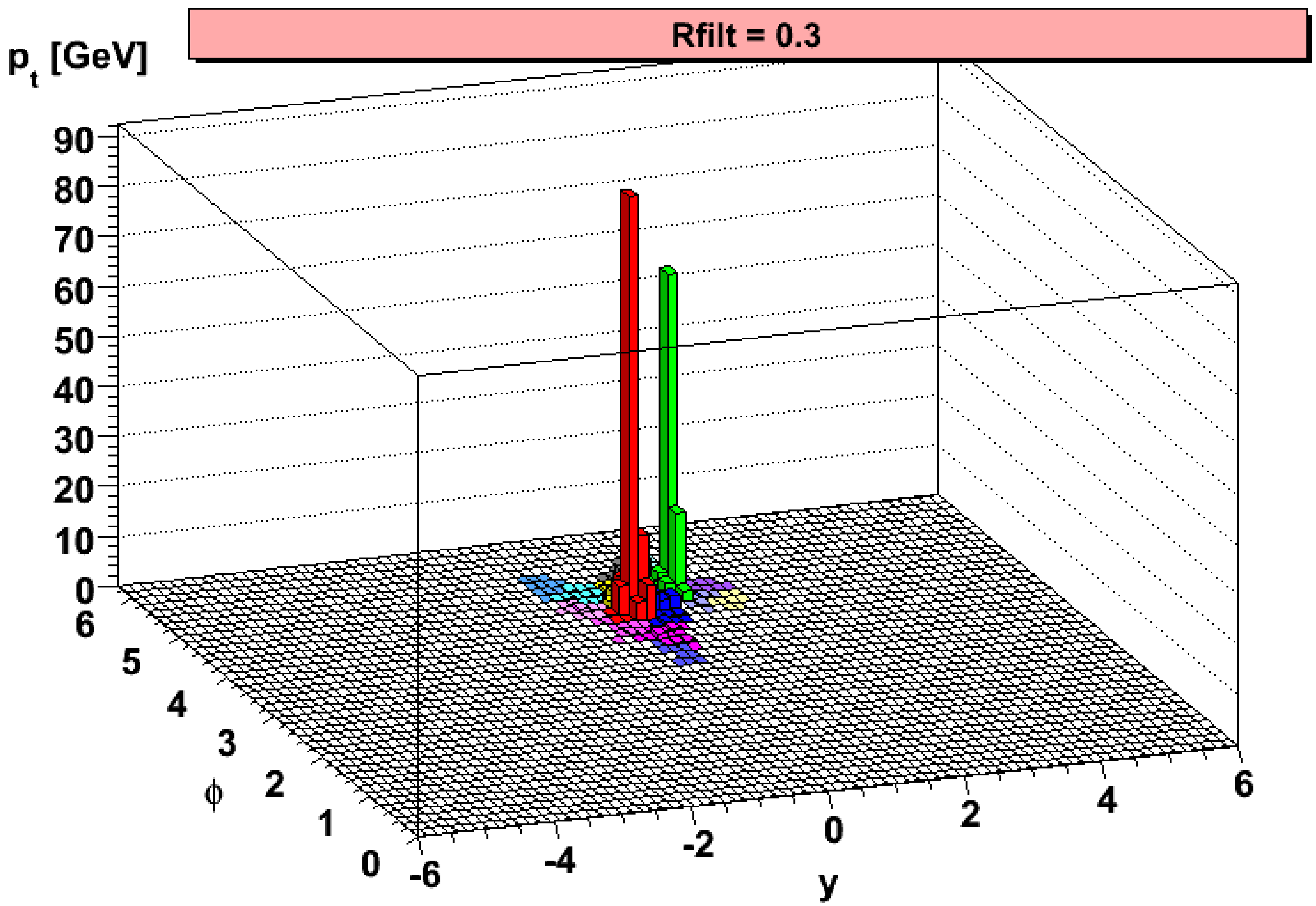
check: $y_{12} \simeq \frac{p_{t2}}{p_{t1}} \simeq 0.7 \to$ OK + 2 b-tags (anti-QCD)

Butterworth, Davison, Rubin & GPS '08

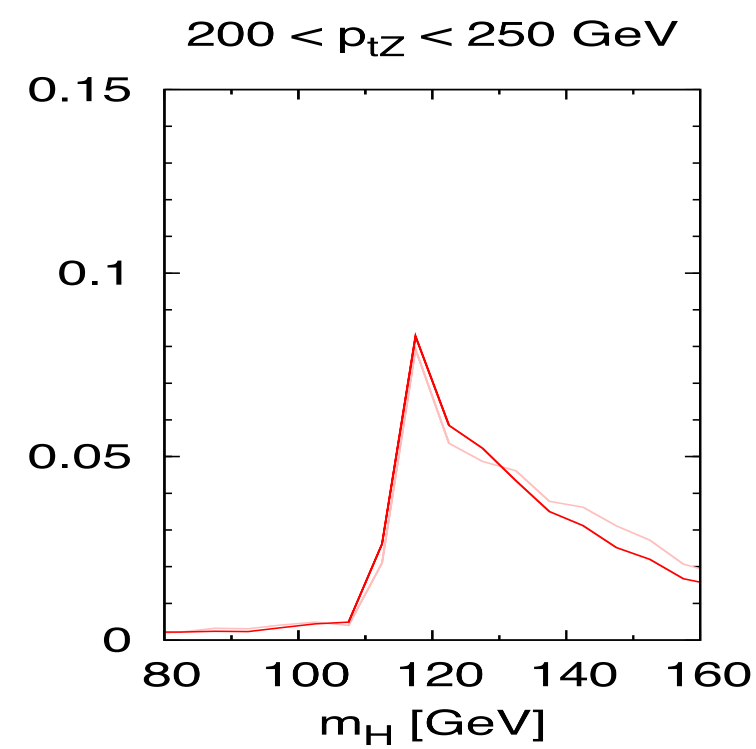
arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

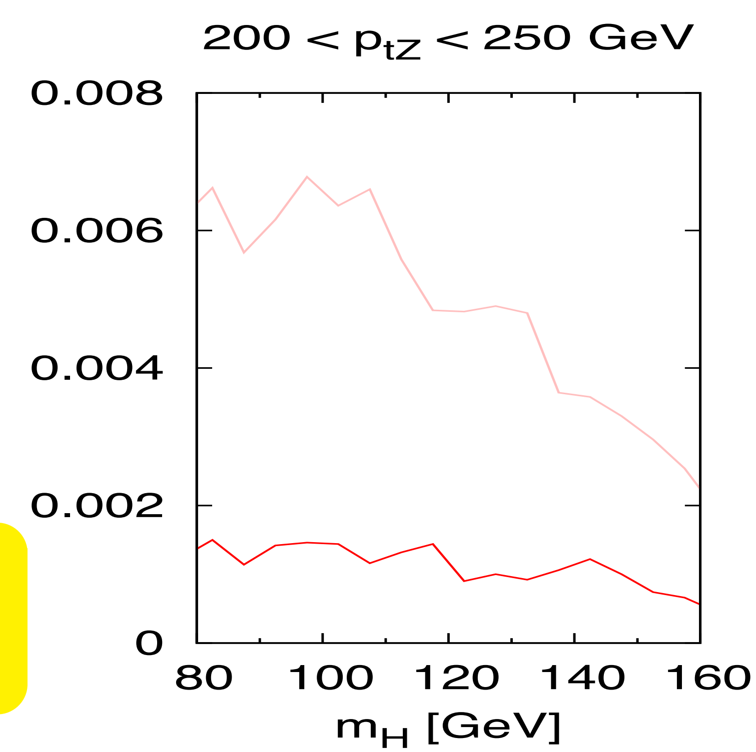
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND



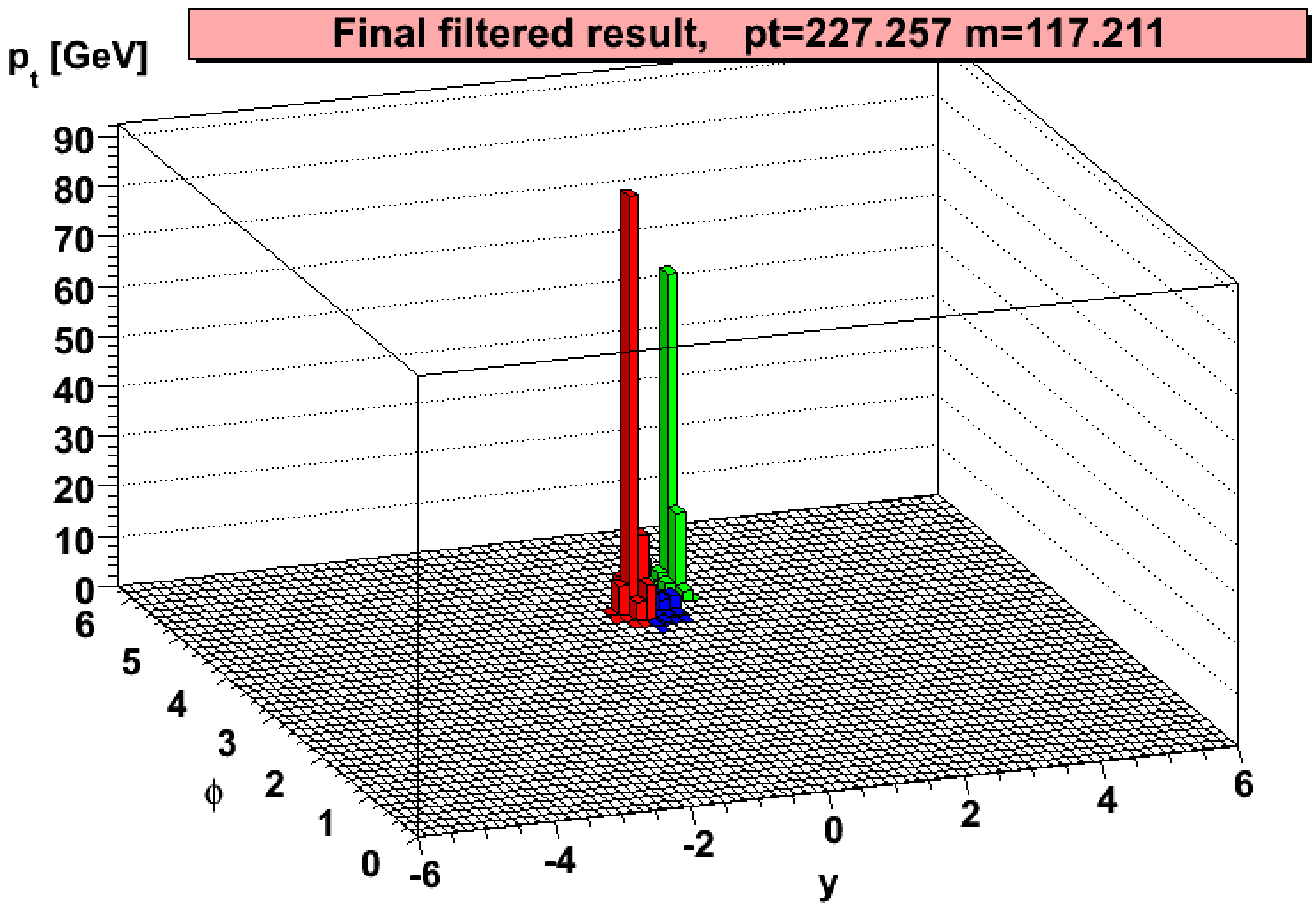
$R_{filt} = 0.3$

Butterworth, Davison, Rubin & GPS '08

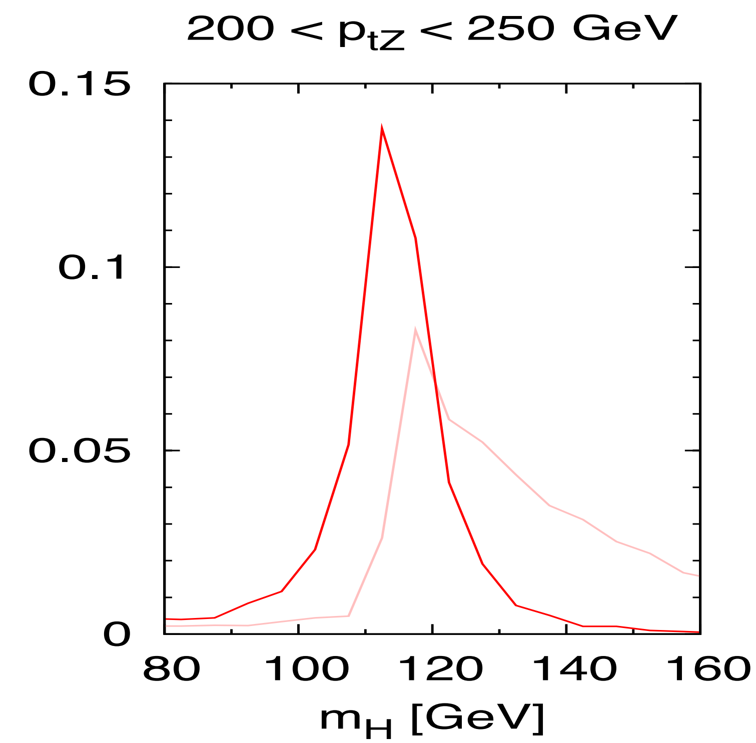
arbitrary norm.

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

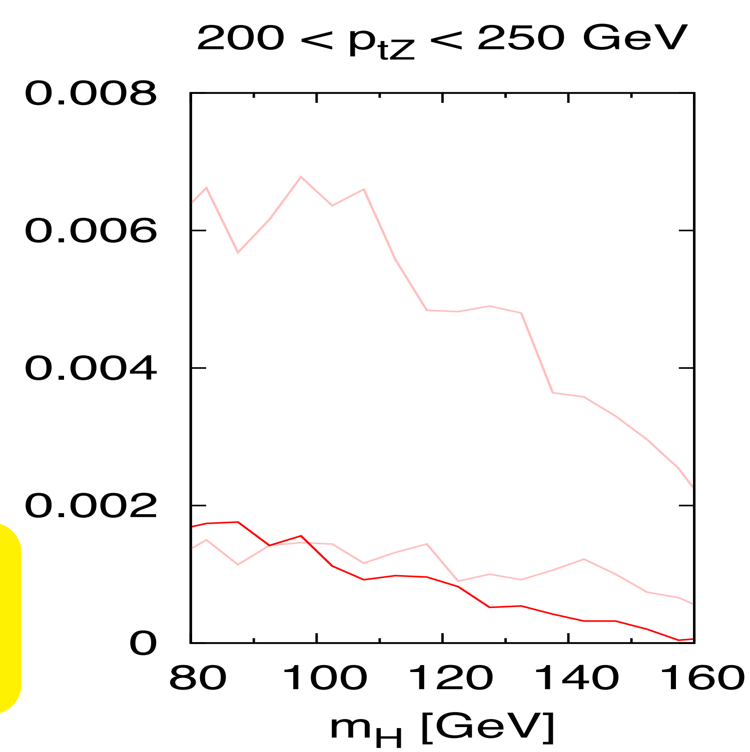
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND

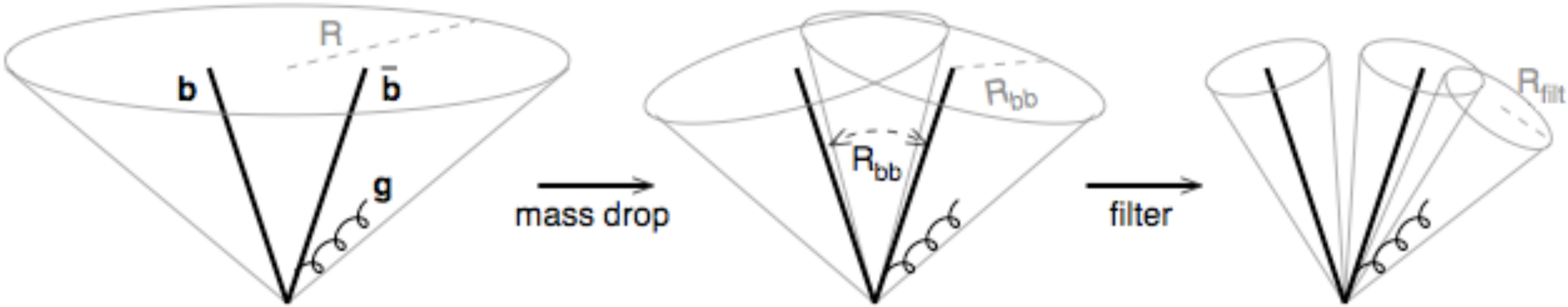


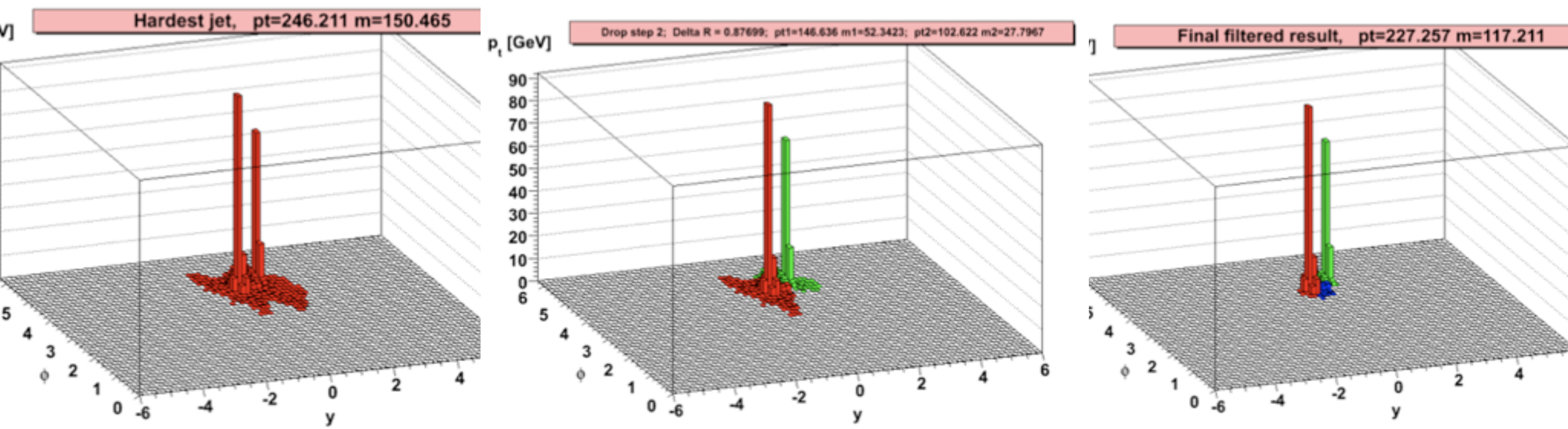
$R_{filt} = 0.3$: take 3 hardest, **m = 117 GeV**

Butterworth, Davison, Rubin & GPS '08

arbitrary norm.

$pp \to ZH \to \nu\nu b\bar{b}$



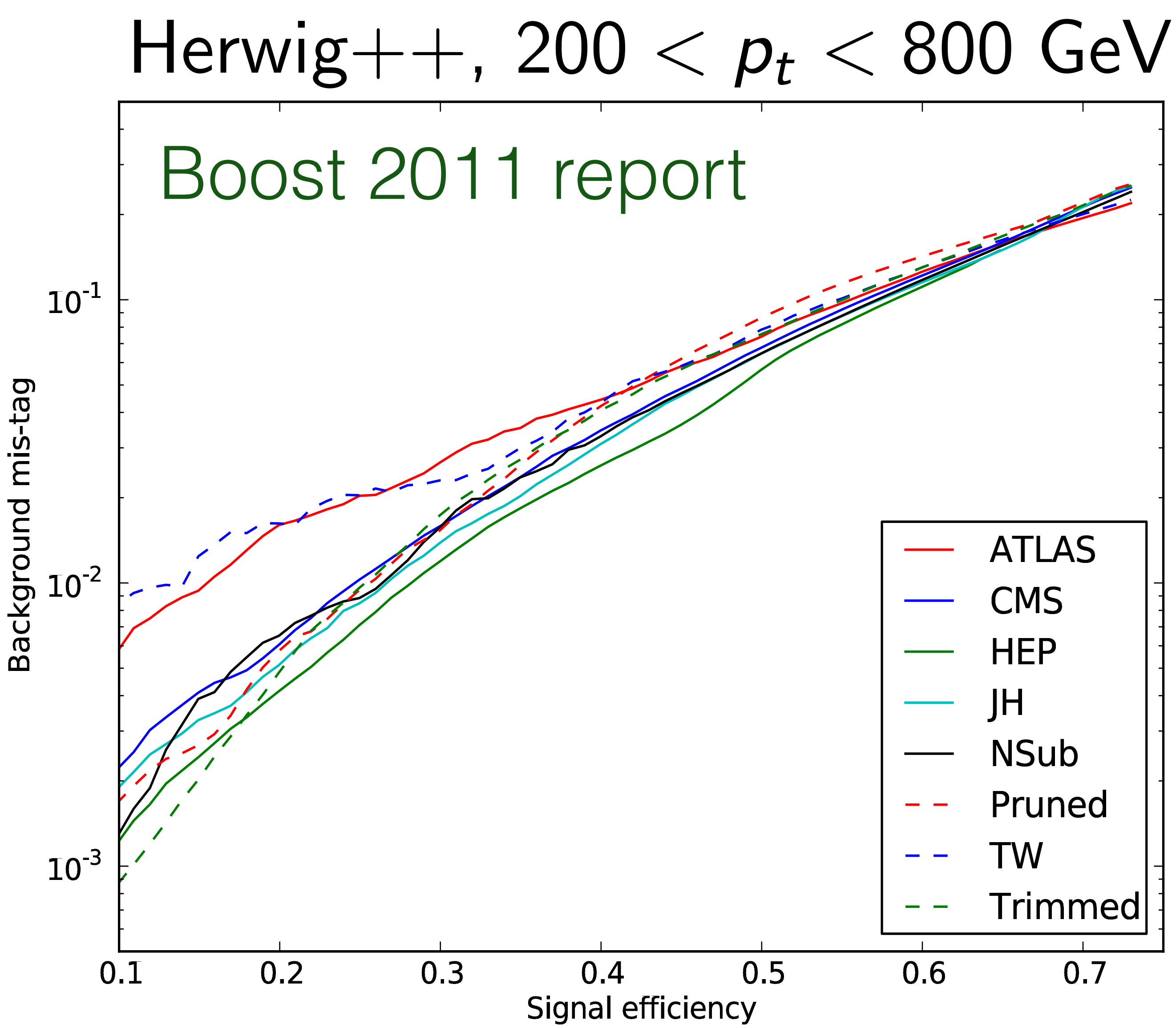


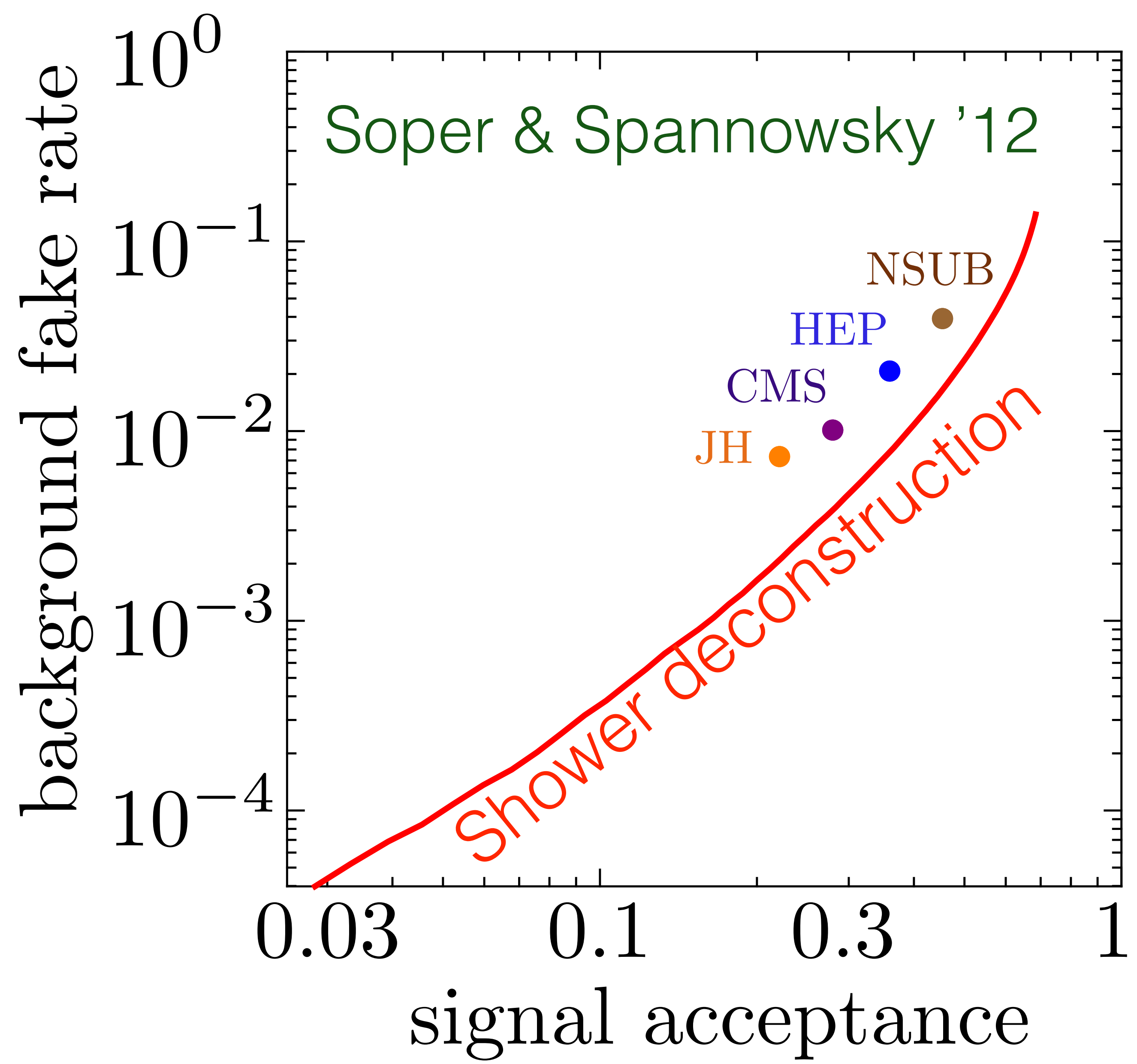
Cluster with a large R

Undo the clustering into subjets, until a large mass drop is observed

Re-cluster with smaller R, and keep only 3 hardest jets

E.g. theorists' comparisons of top taggers





Extensive experimental work

ATLAS Public Results

- Large-R, groomed jets with pile-up
- Large-R jets with substructure
- Quark/gluon jets (see also this link)
- Jet substructure at LHC7
- Jet properties for boosted searches

Resonance searches

- Boosted top (hadronic)
- Boosted top (semileptonic)
- Three-jet resonance (gluino RPV)
- Two-jet resonance (sgluon)

CMS Public Results

- Jet substructure in CMS
- Subjet multiplicity
- Jet mass and grooming

Resonance searches:

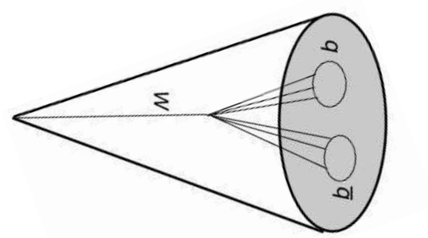
- Boosted top (hadronic)
- Boosted top (semileptonic)
- Boosted W/Z

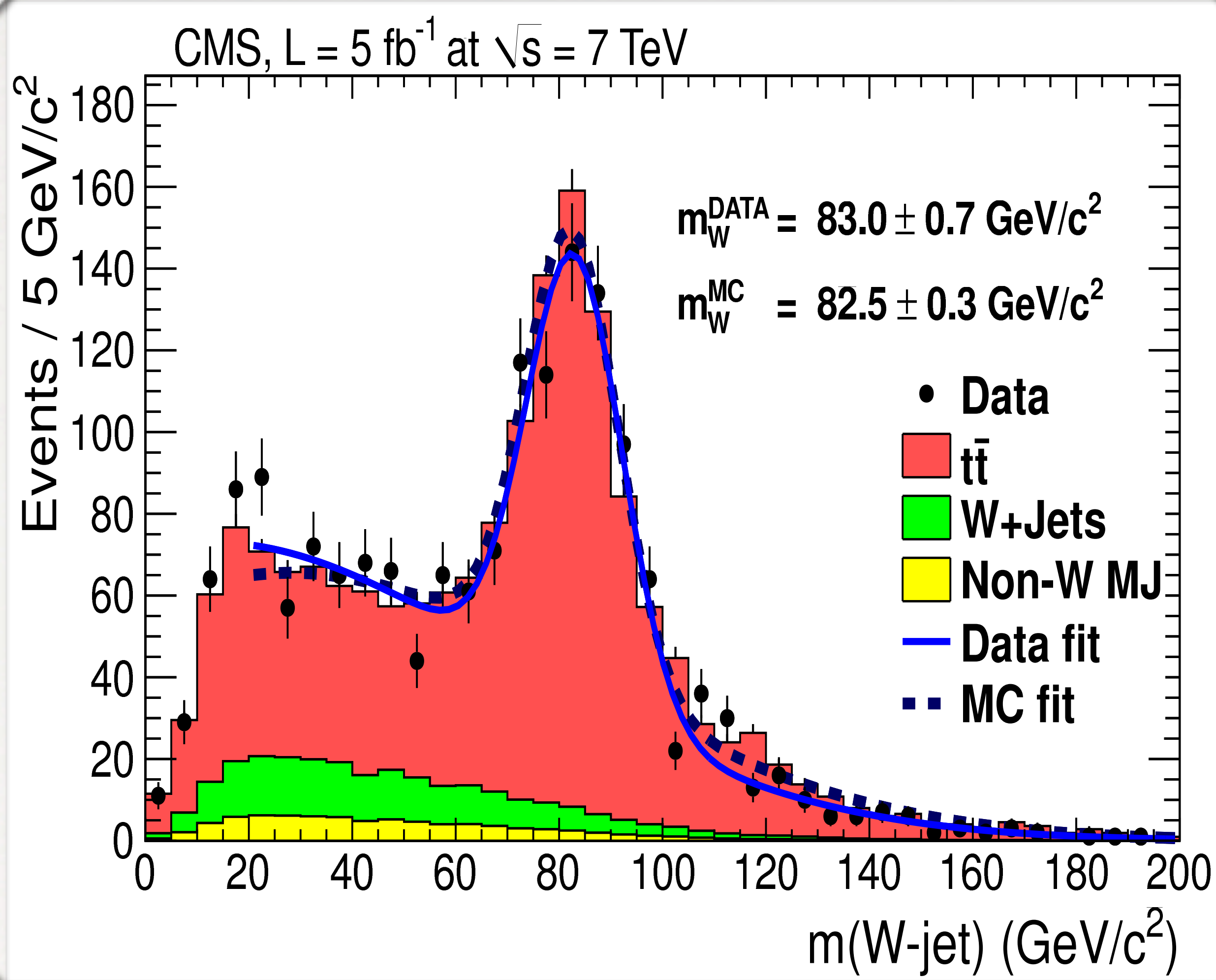
From a list compiled for a recent workshop at Perimeter Institute

Many more analyses in the pipeline

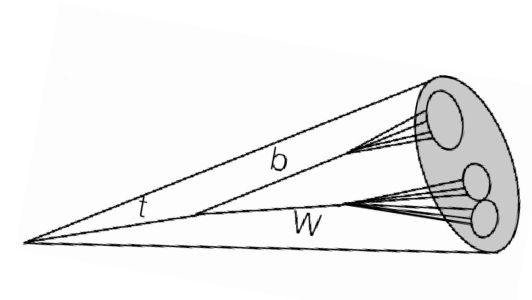
Seeing hadronic W's and tops in a single jet

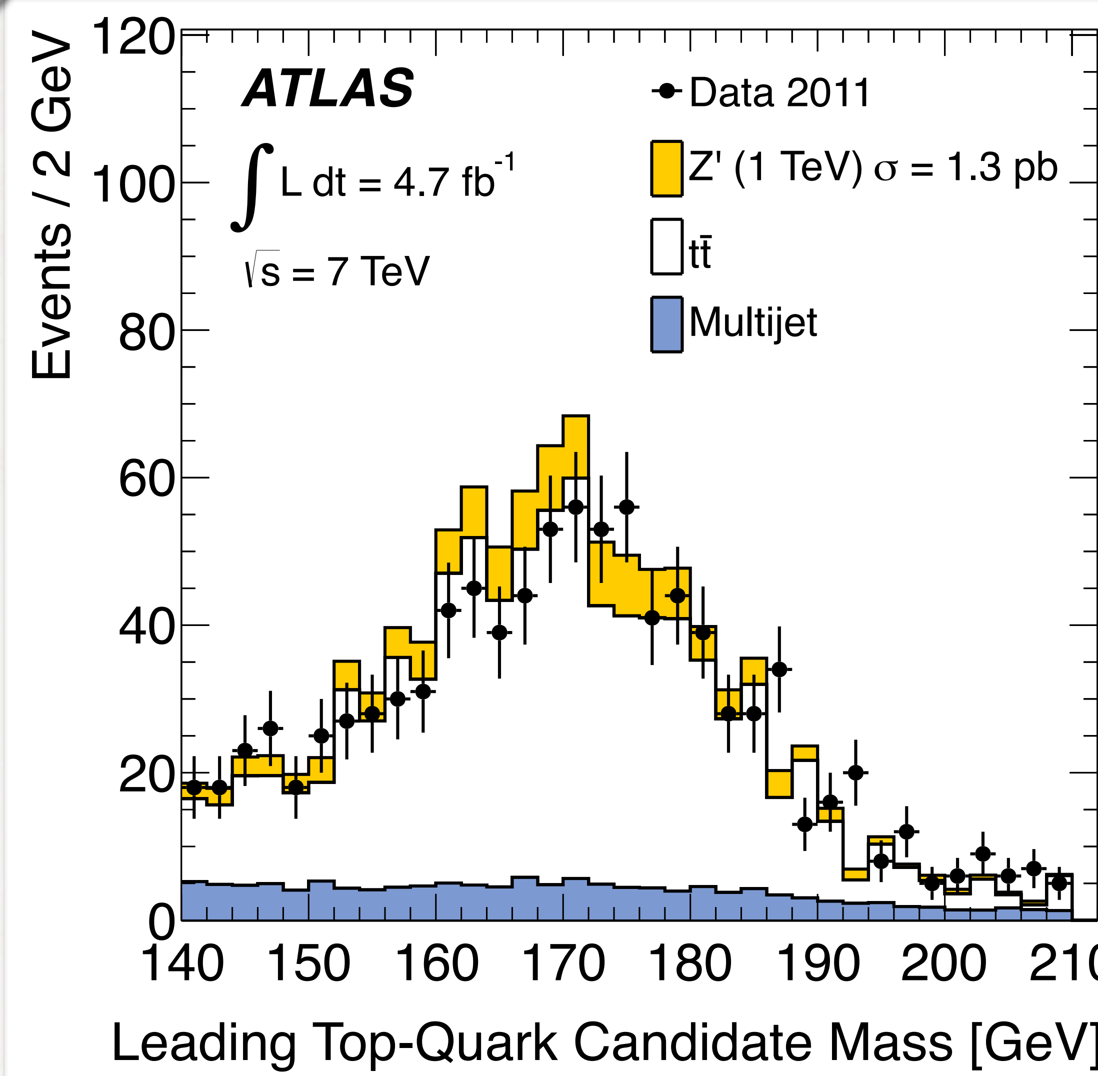
W's in a single jet





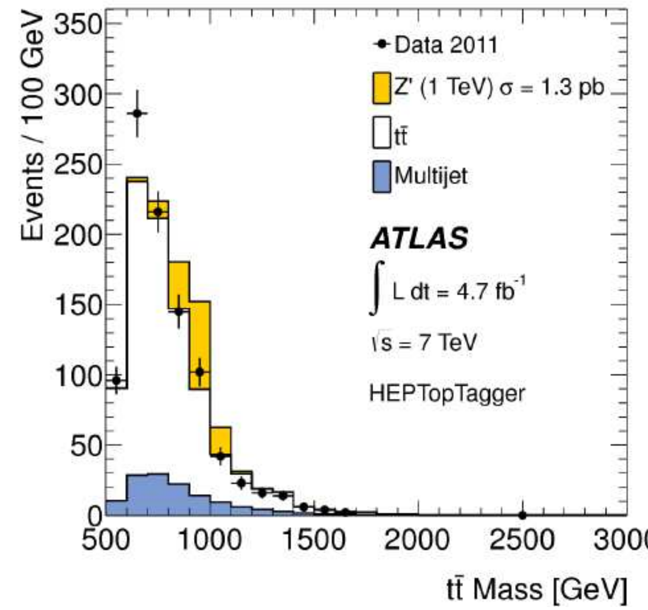
tops in a single jet



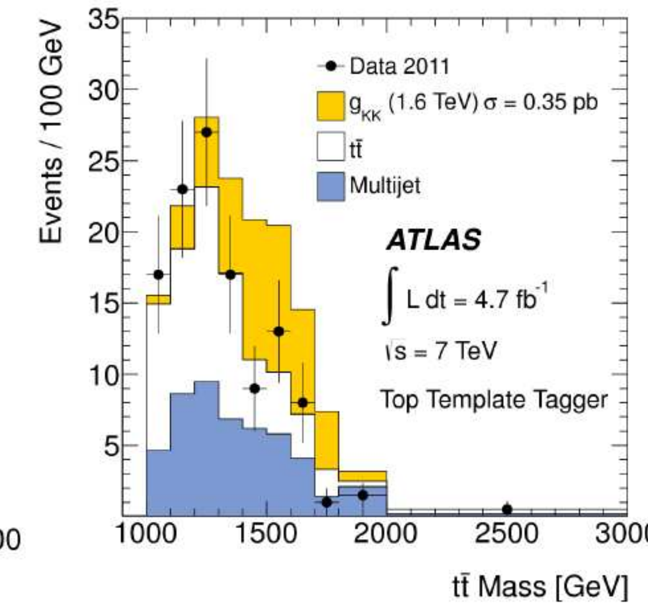


Searches with substructure tools

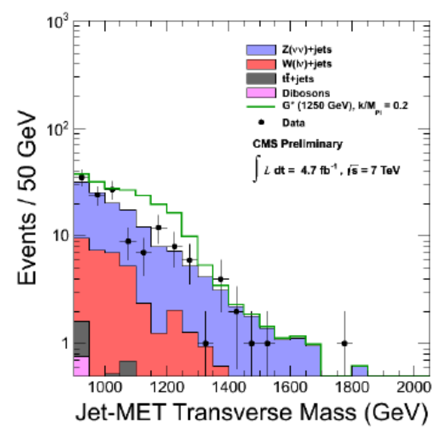
HEPTopTagger



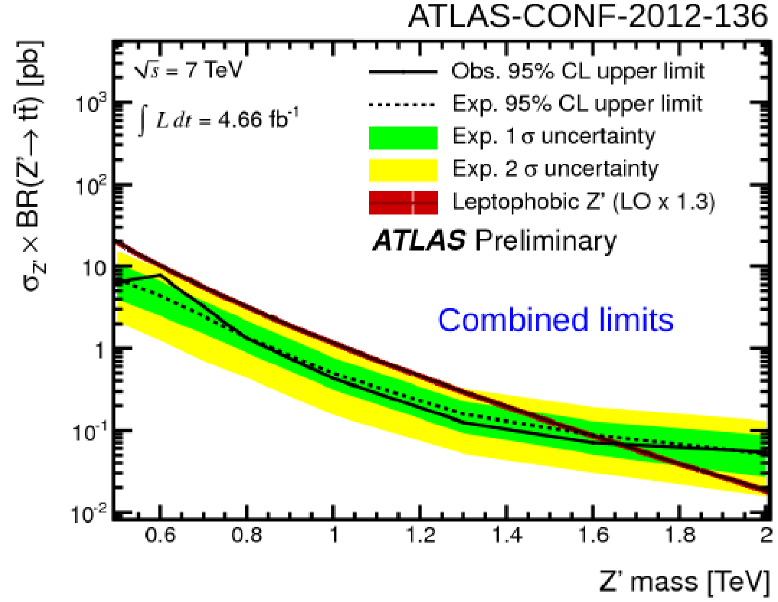
Top Templating

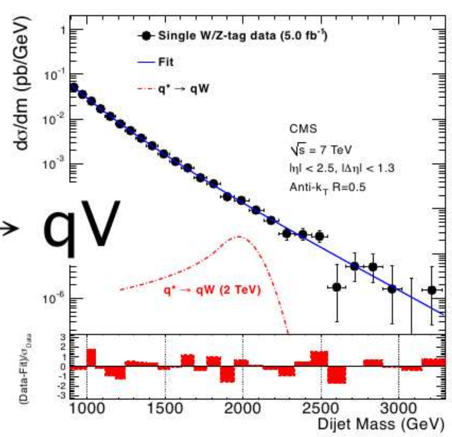


$X \to ZZ \to 2\nu 2q$

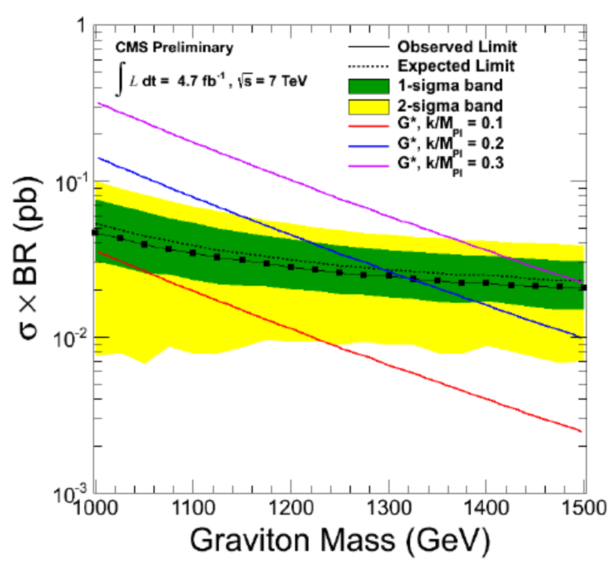


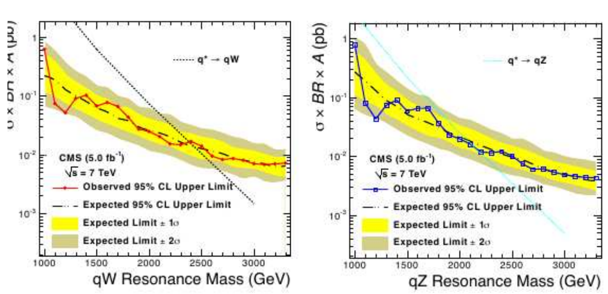
ATLAS-CONF-2012-136





$X \to qV$





A range of techniques being used for varied BSM scenarios

[Analytic] understanding

arXiv:1307.0007

Dasgupta, Fregoso, Marzani & GPS

+Dasgupta, Fregoso, Marzani & Powling, 1307.0013

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

For phenomenology

Jet mass: *m*

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

For QCD calculations

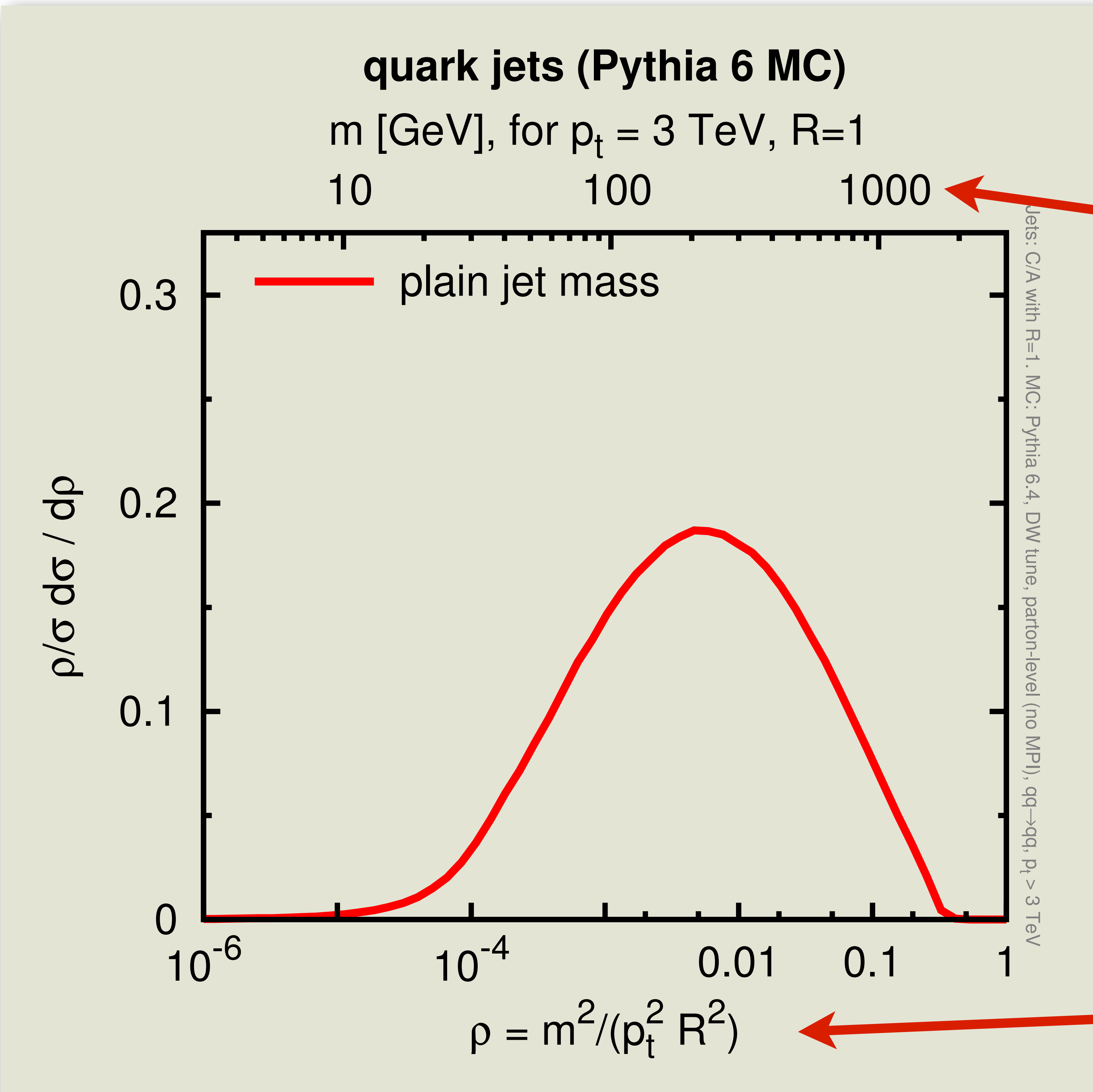
$$\rho = \frac{m^2}{p_t^2 R^2}$$

[R is jet opening angle – or radius]

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

The “right” MC study can already be instructive

(testing on quark [background] jets)

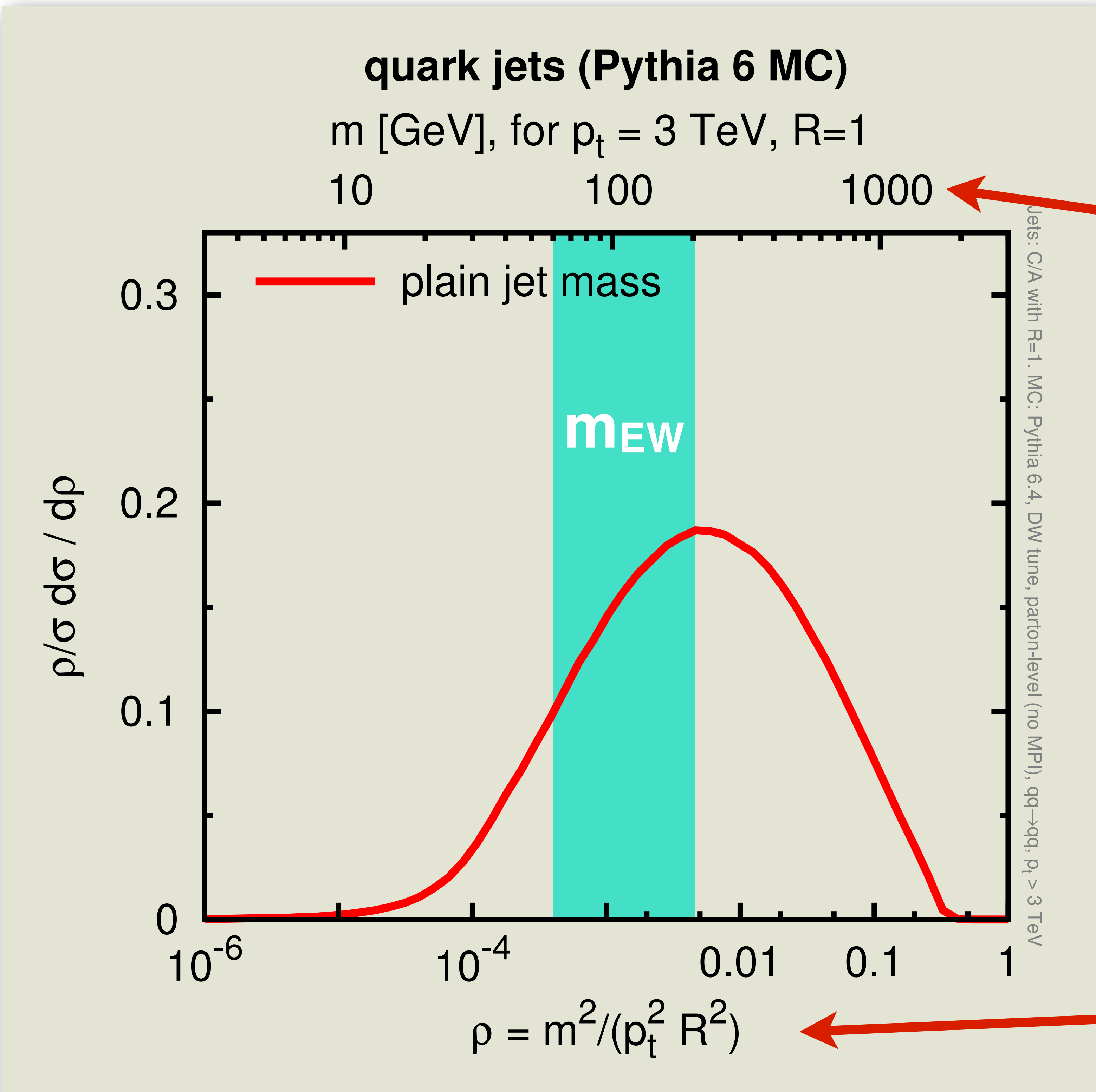


Physical mass for 3 TeV, R=1 jets

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

The “right” MC study can already be instructive

(testing on quark [background] jets)

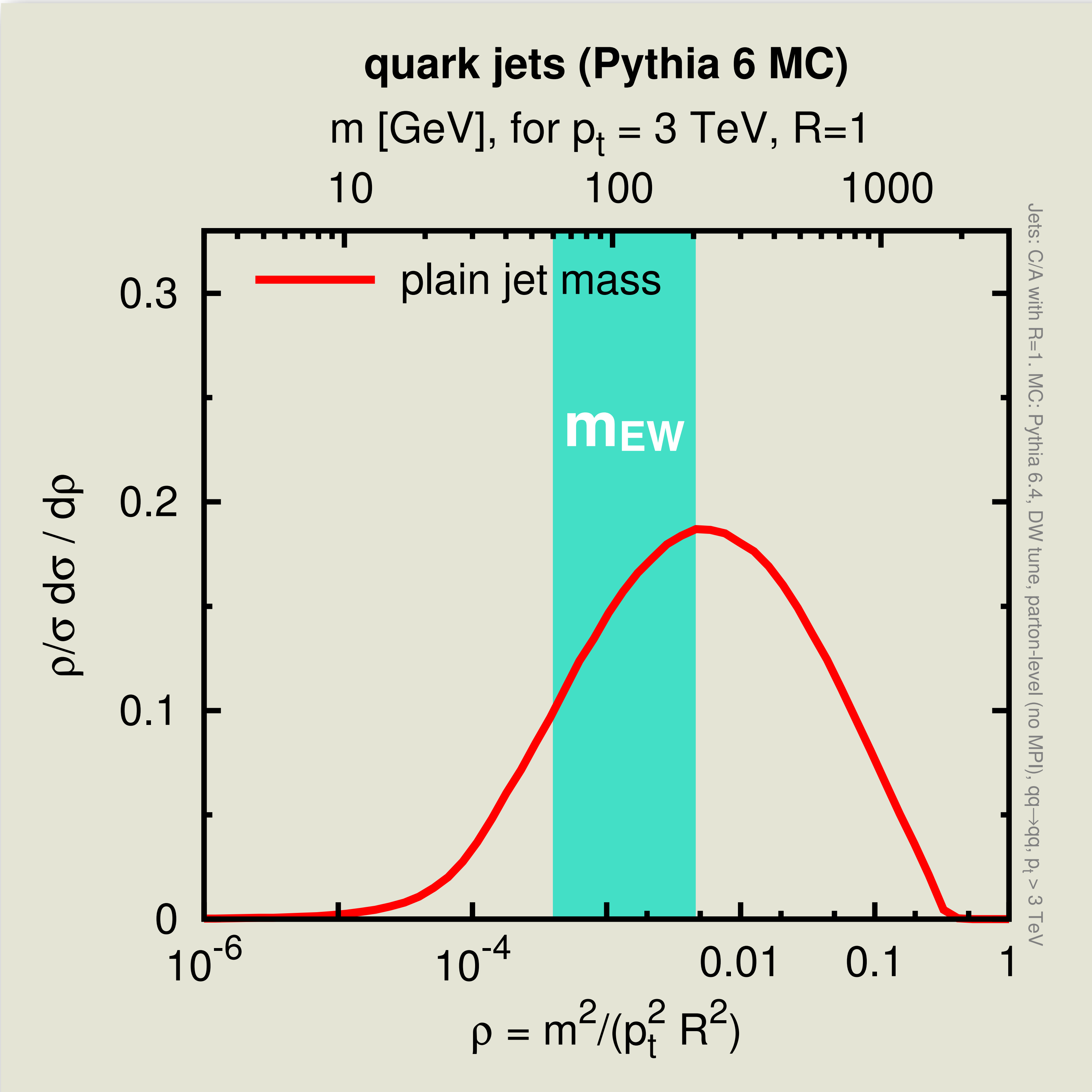


Physical mass for 3 TeV, R=1 jets

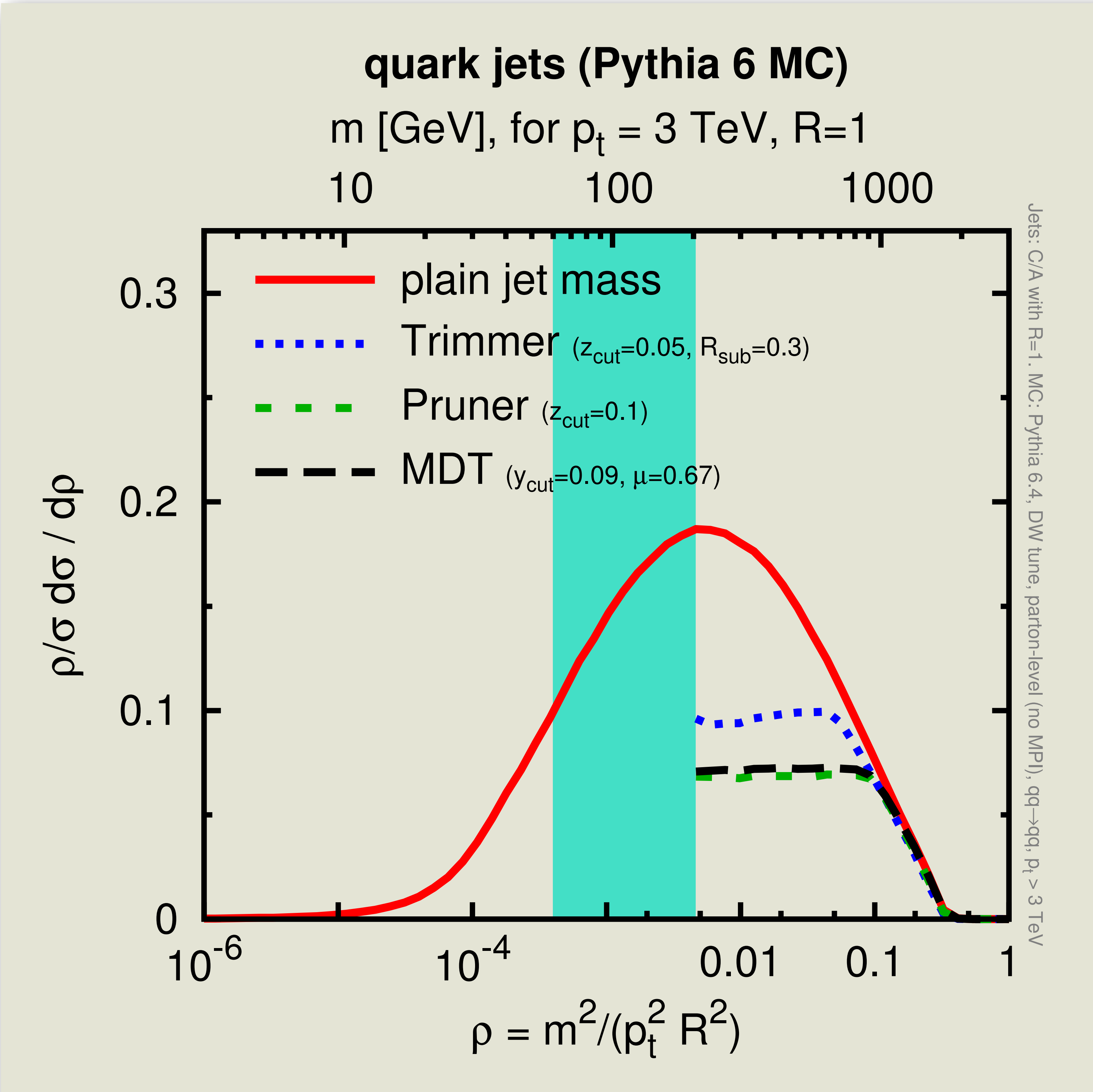
ρ ~ Rescaled mass^2

(i.e. the QCD variable)

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

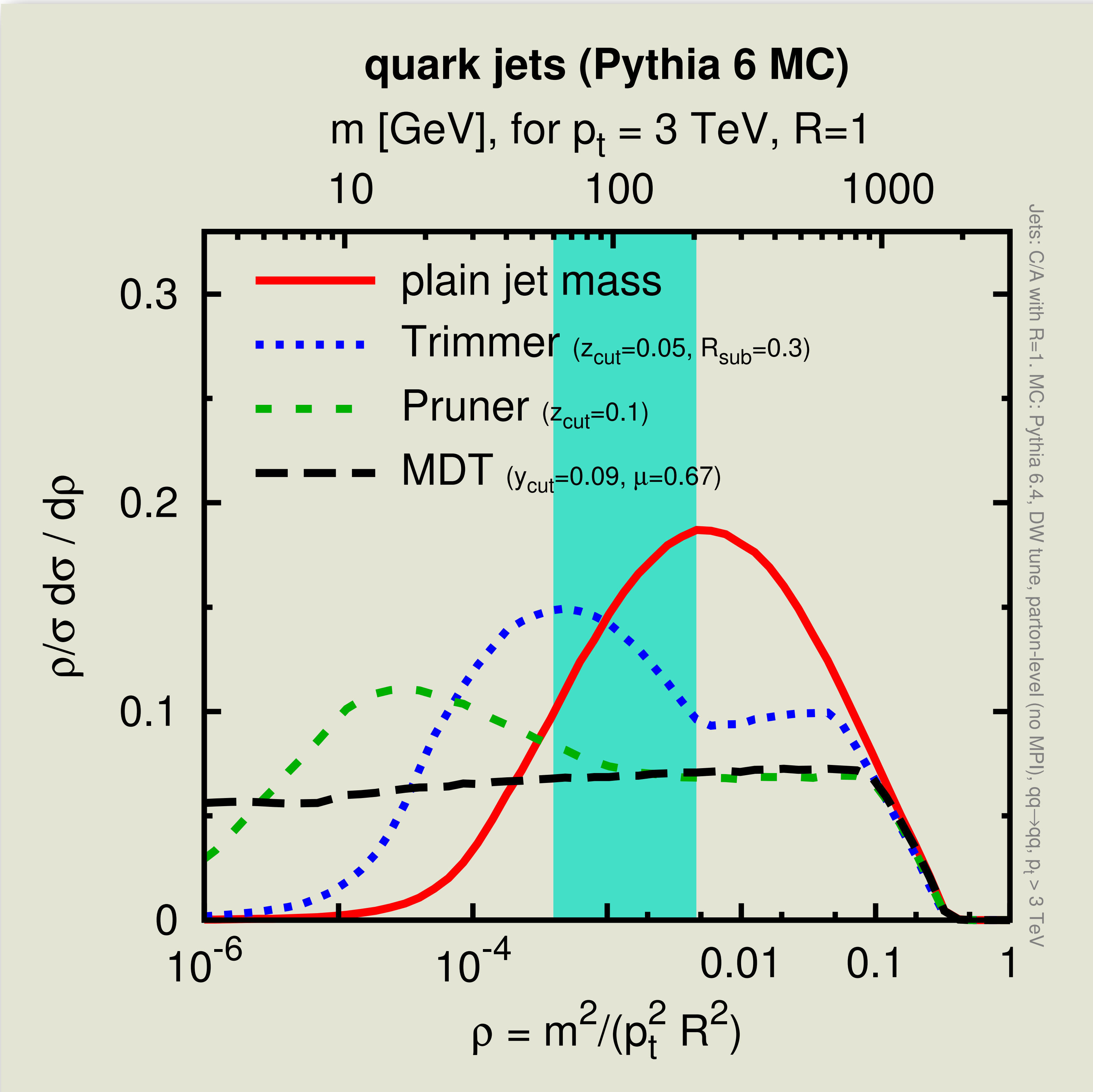


The “right” MC study can already be instructive
(testing on quark [background] jets)



Different taggers can be quite similar

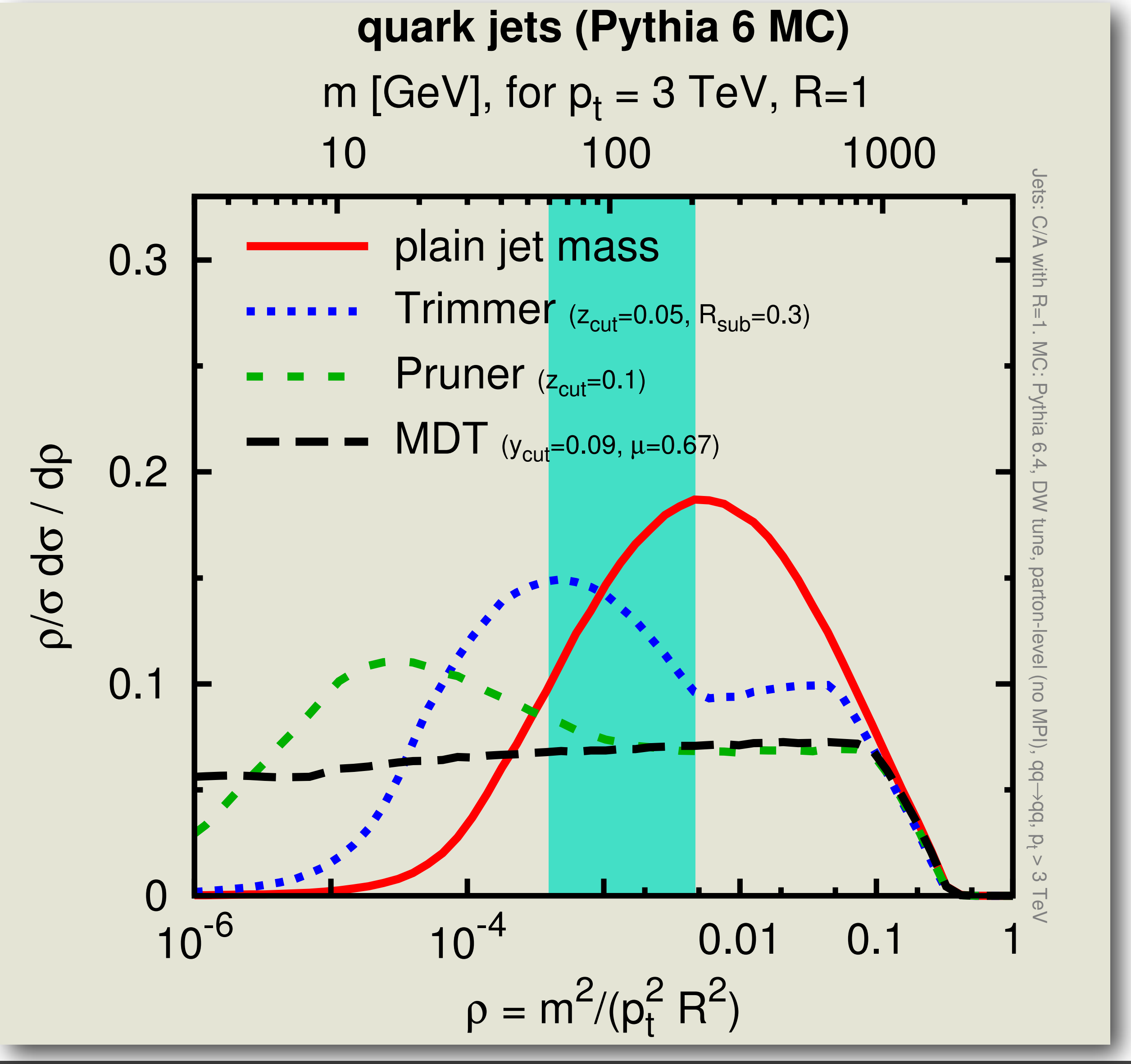
The "right" MC study can already be instructive
(testing on quark [background] jets)



But only for a limited range of masses

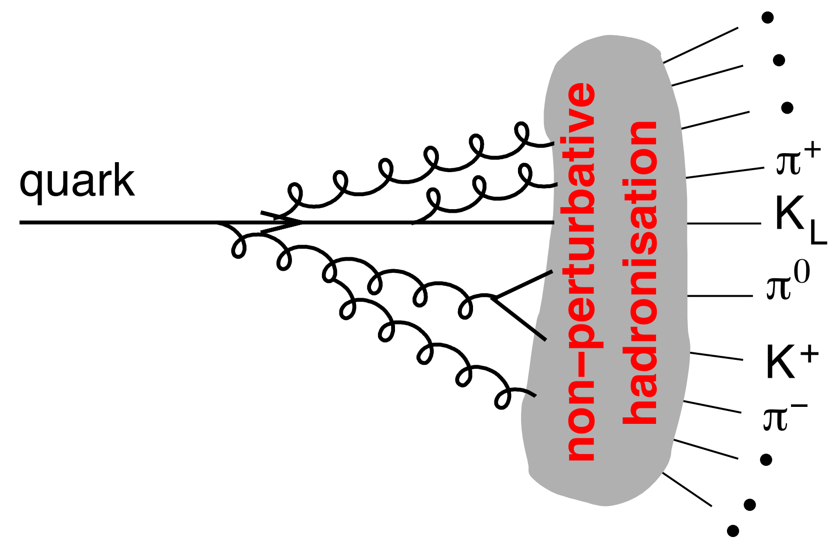
Do we care about such differences?

- 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_{sub}, etc. parameters?
- Are these structures telling us something we might want to know about the taggers? E.g. how to improve them?

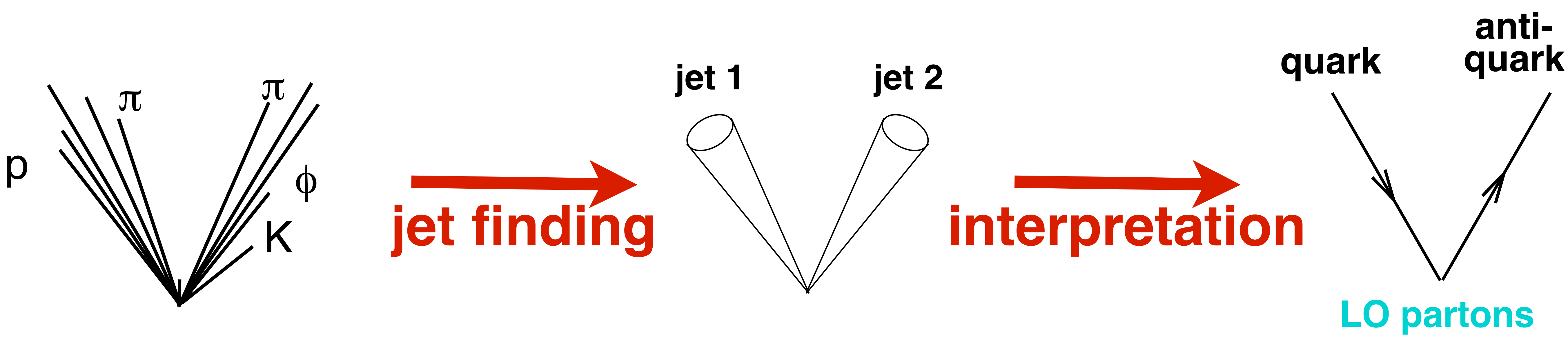


This calls for analysis and calculation

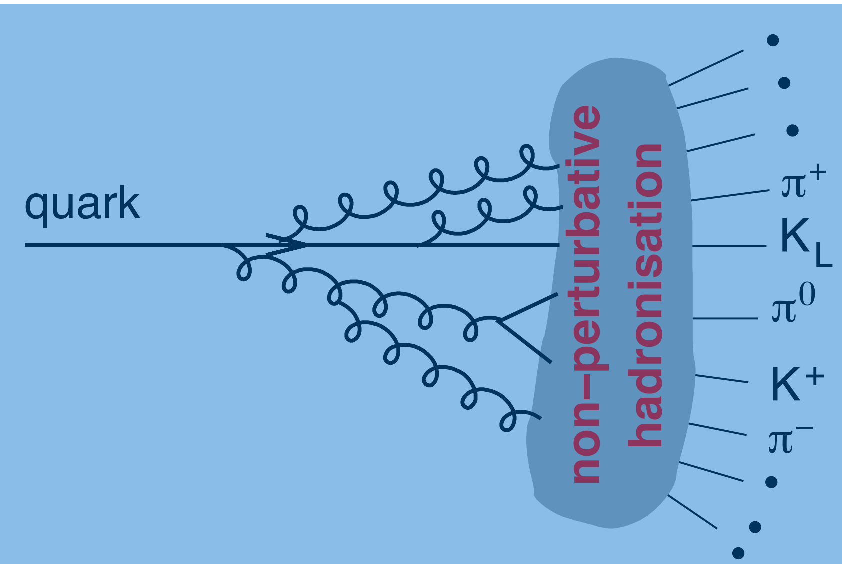
How jets come to have the structure they do



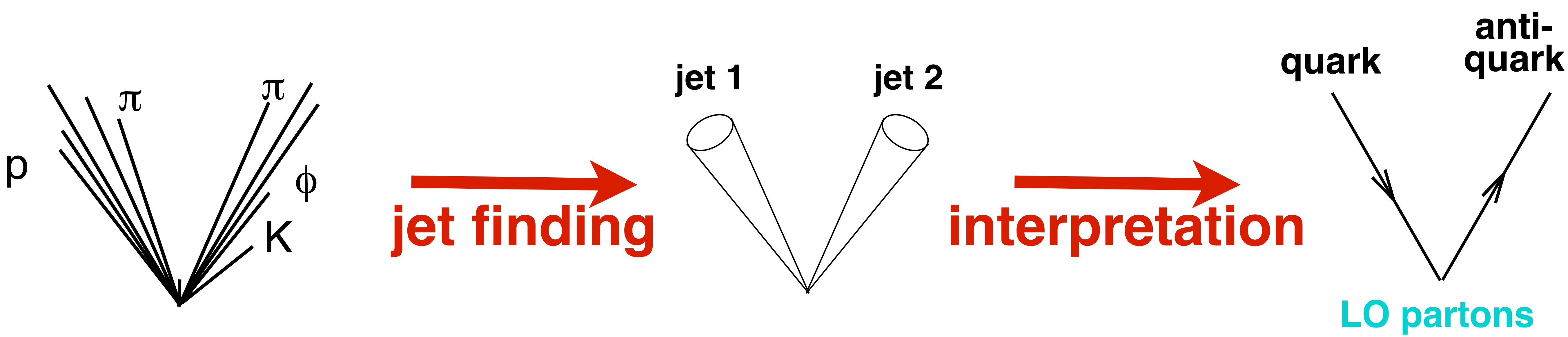
How we "reconstruct" jets



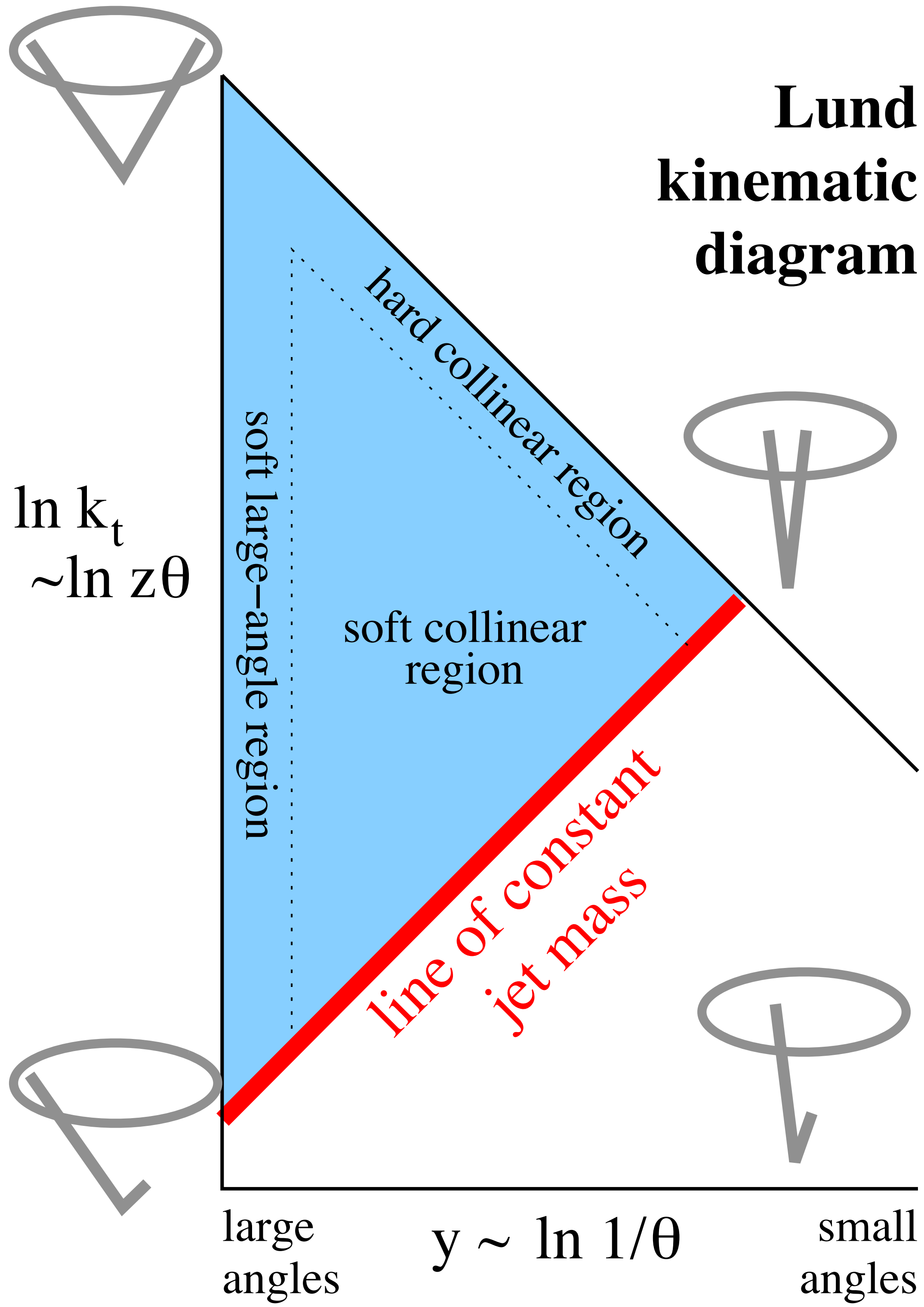
How jets come to have the structure they do



How we "reconstruct" jets



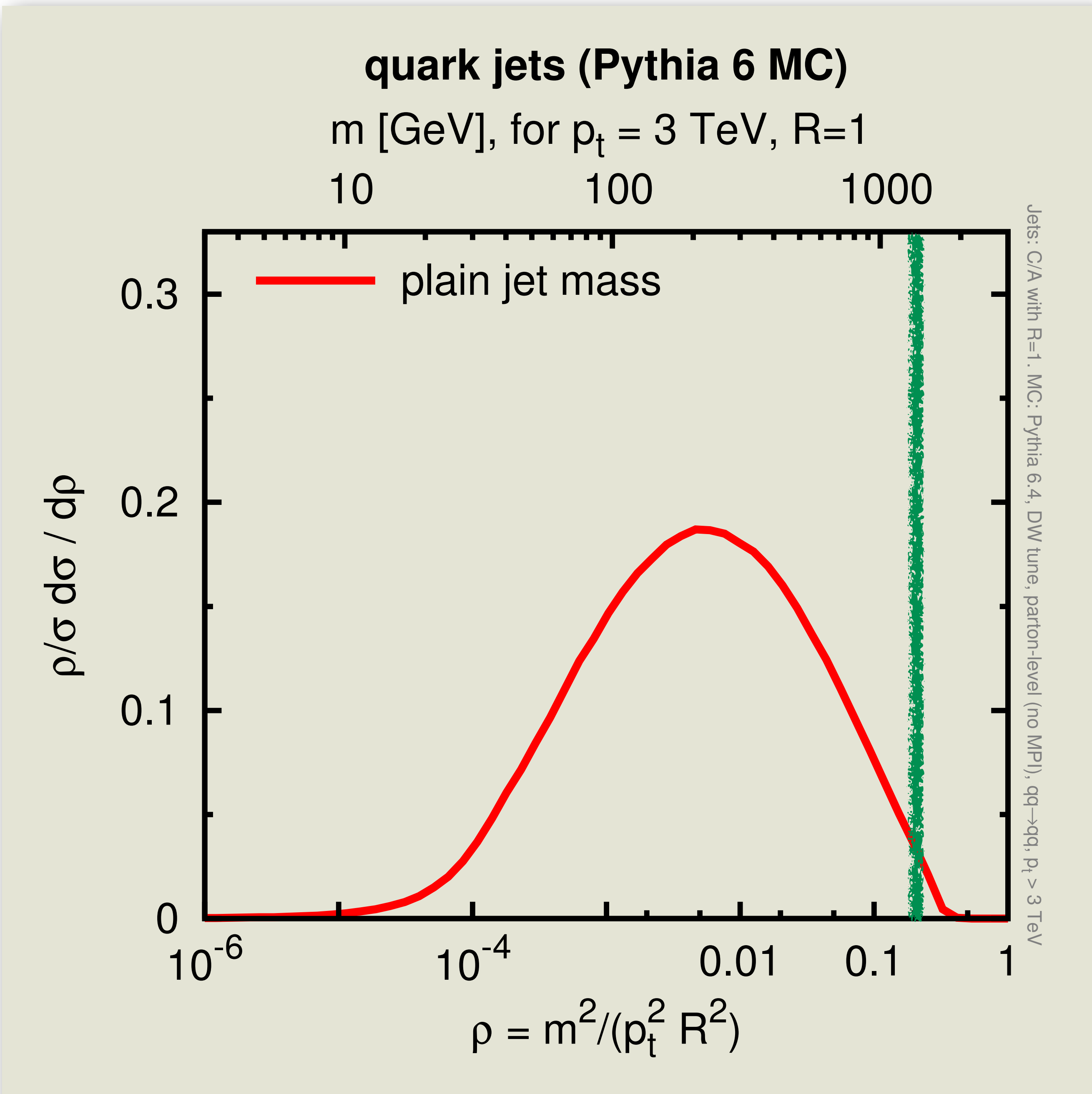
log of transverse momentum of emission wrt jet axis

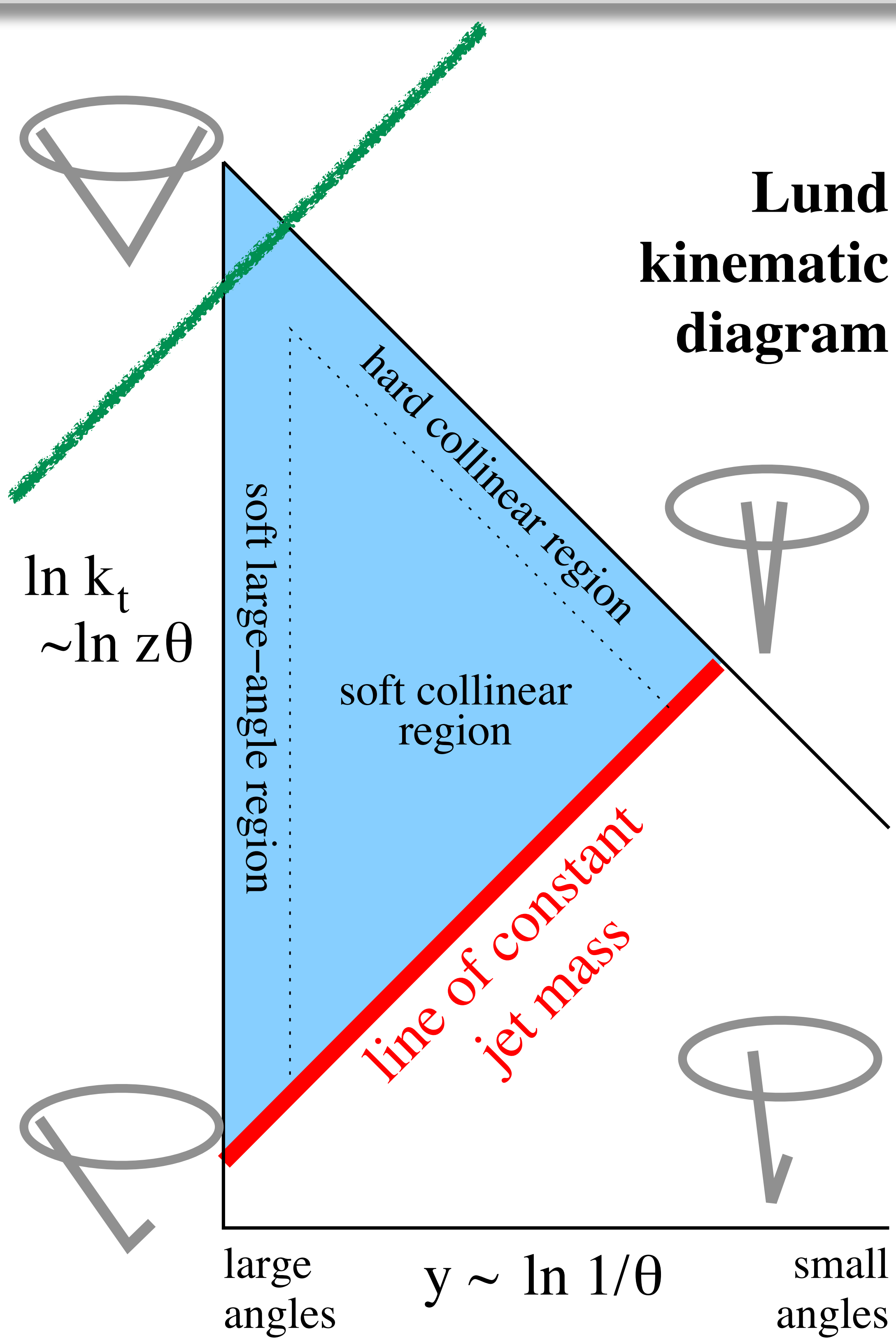


for QCD emissions

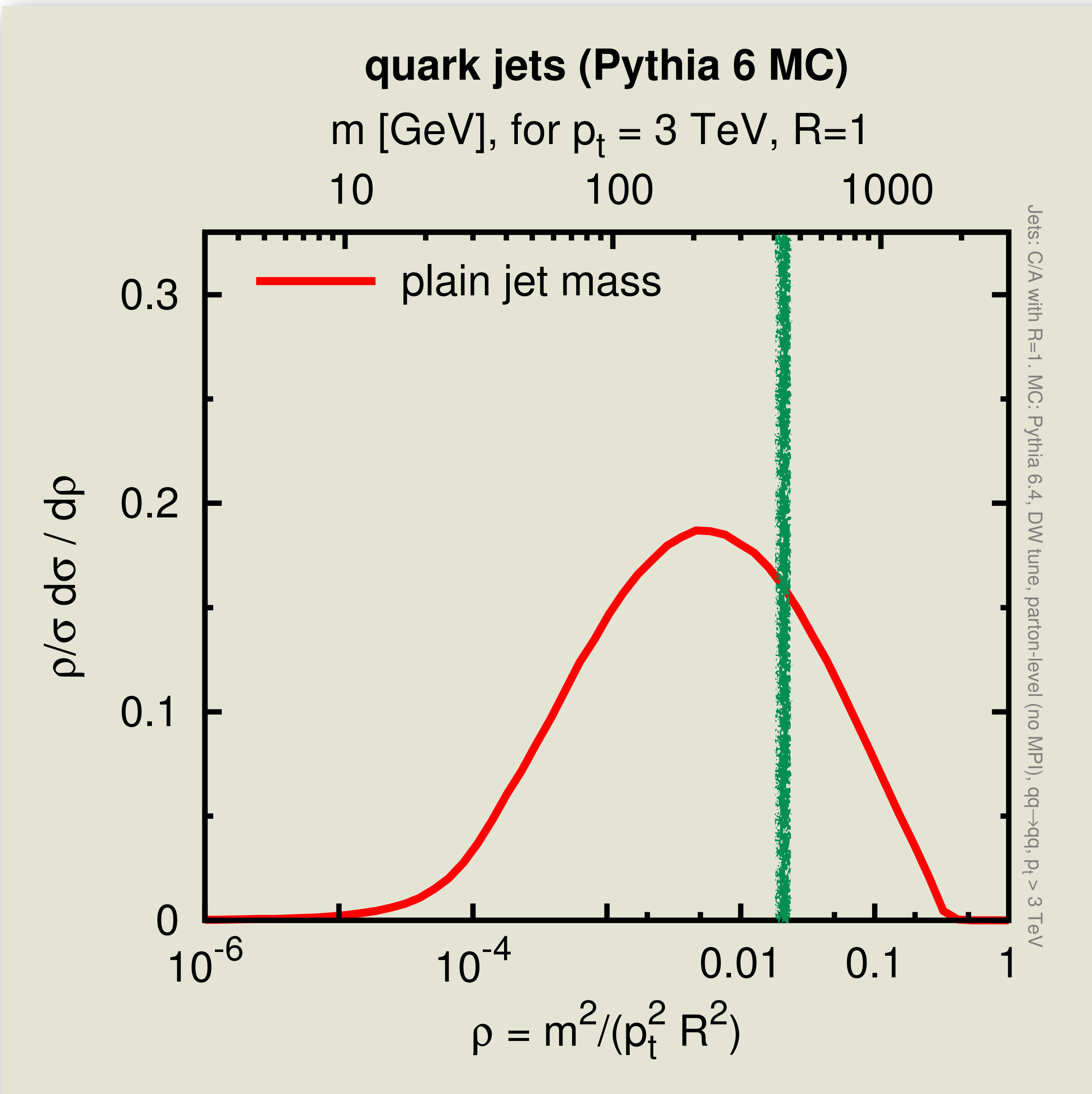
log of 1/(emission angle)

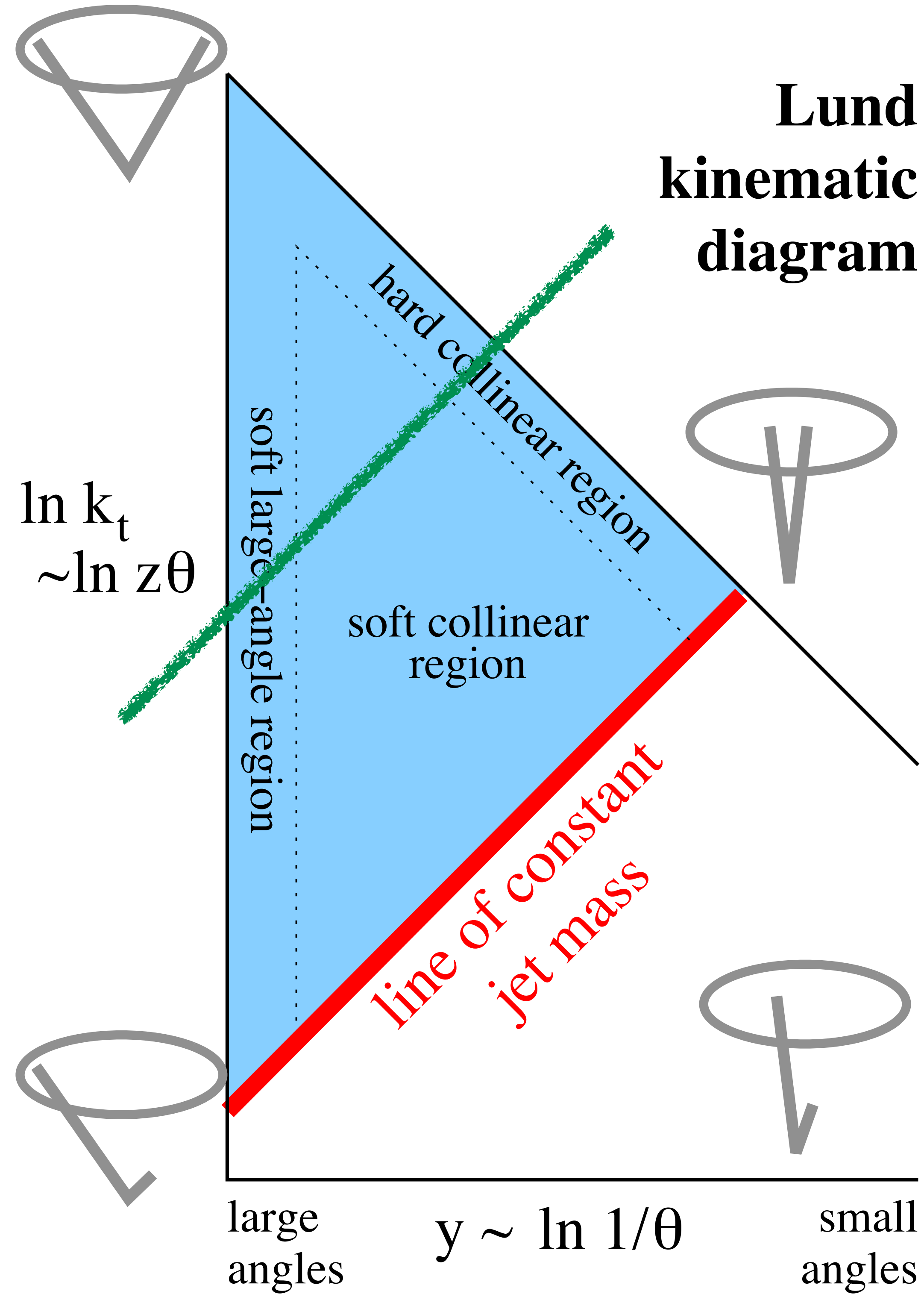
kinematic plane and jet masses





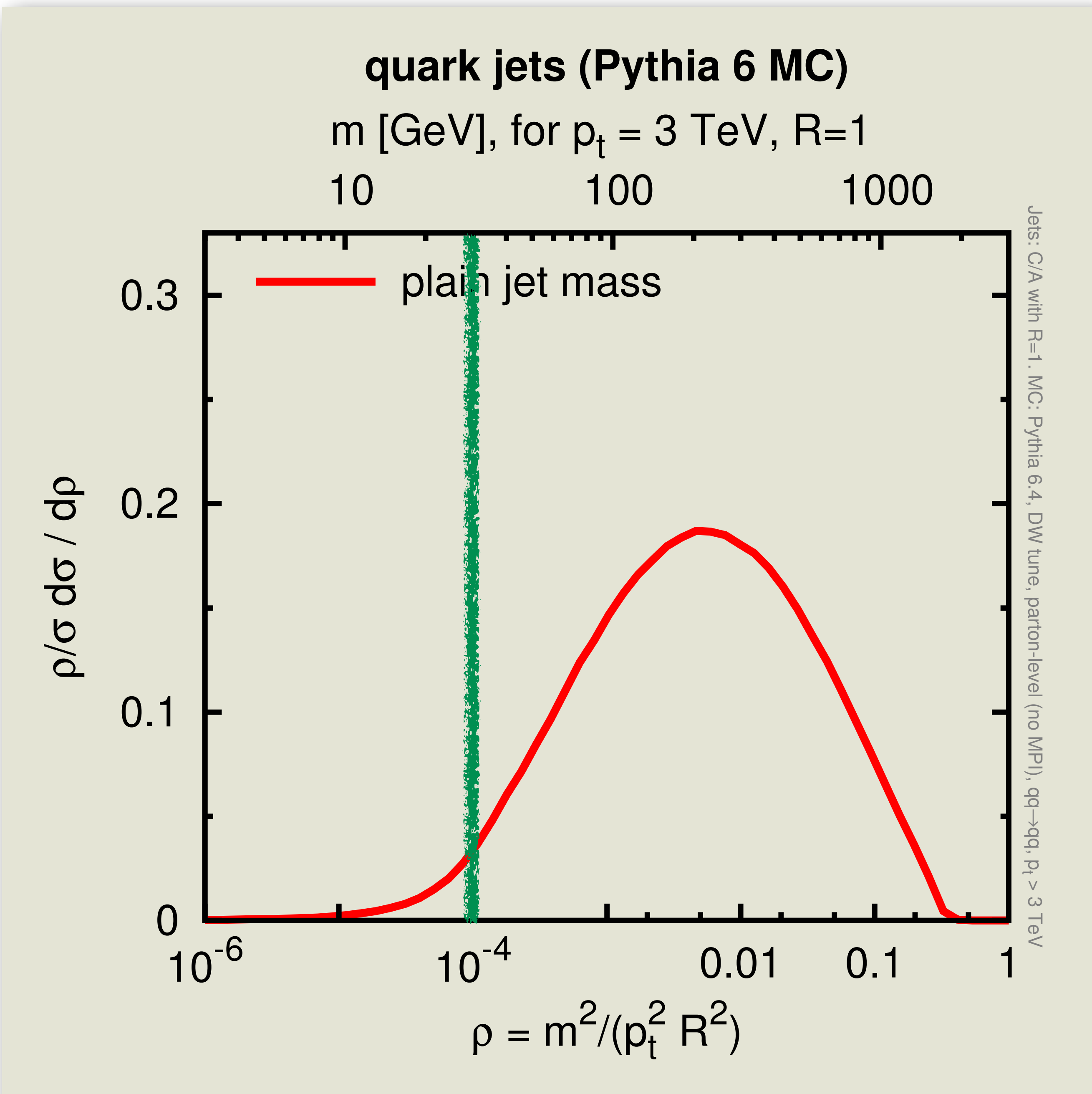
kinematic plane and jet masses

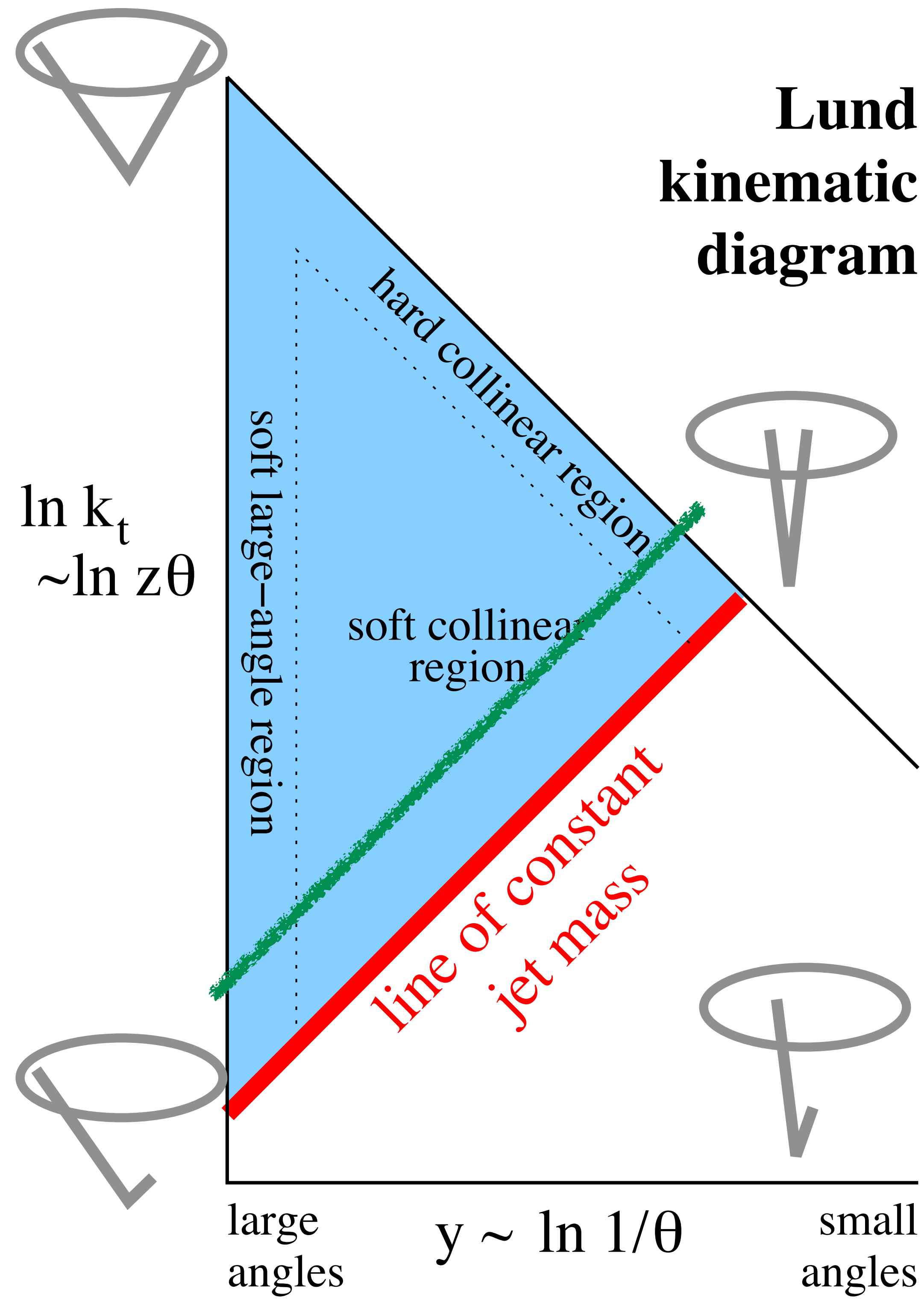




increase in phasespace

kinematic plane and jet masses



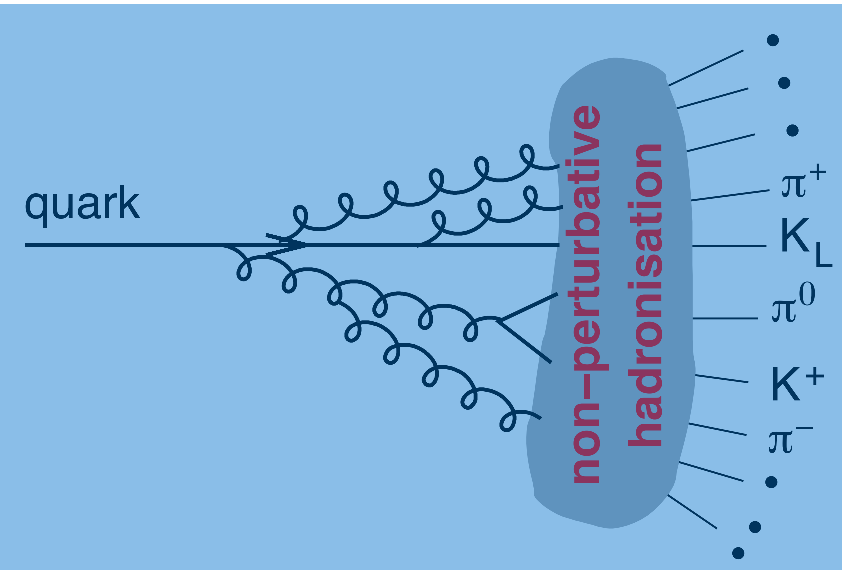


Sudakov suppression ← **increase in phasespace** ←

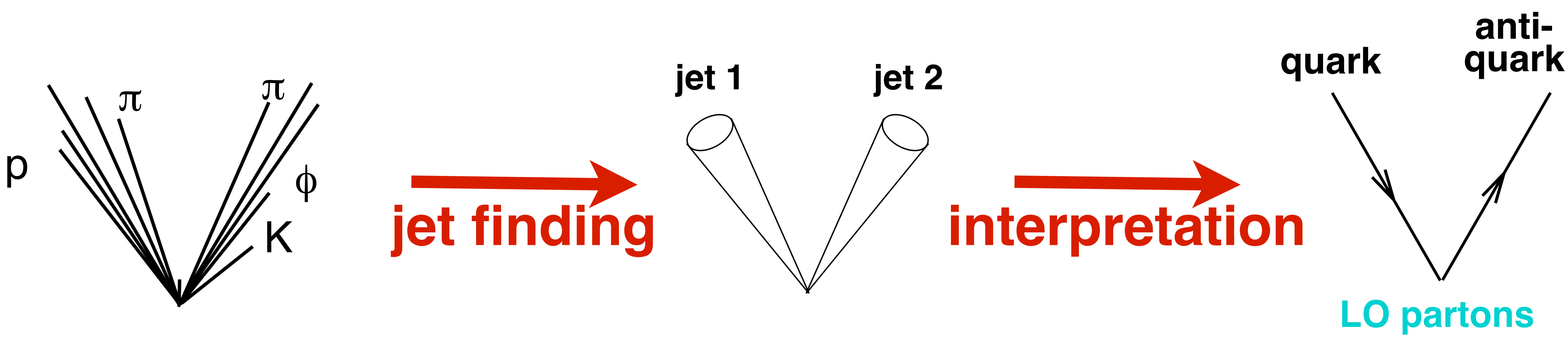
Some key calculations related to jet mass

- Catani, Turnock, Trentadue & Webber, '91: **heavy-jet mass in e^+e^-**
- Dasgupta & GPS, '01: **hemisphere jet mass in e^+e^-** (and DIS)
- Appleby & Seymour, '02
  Delenda, Appleby, Dasgupta & Banfi '06: **impact of jet boundary**
- Gehrmann, Gehrmann de Ridder, Glover '08; Weinzierl '08
  Chien & Schwartz '10: **heavy-jet mass in e^+e^- to higher accuracy**
- Dasgupta, Khelifa-Kerfa, Marzani & Spannowsky '12,
  Chien & Schwartz '12,
  Jouttenus, Stewart, Tackmann, Waalewijn '13:
  jet masses at hadron colliders

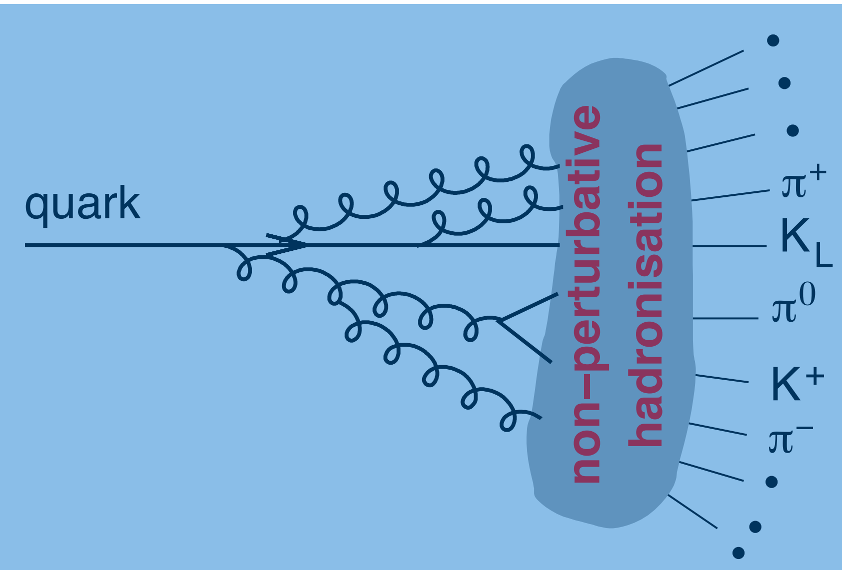
How jets come to have the structure they do



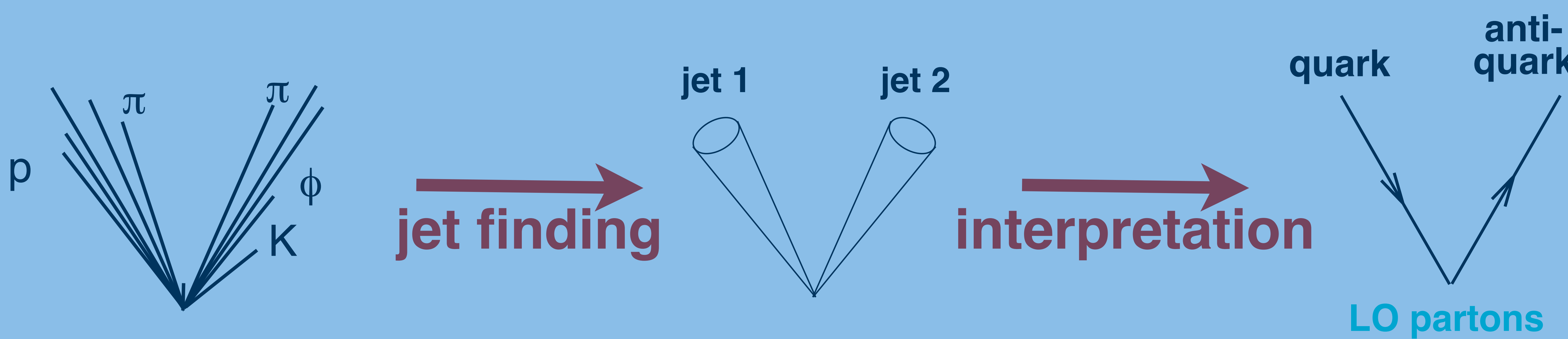
How we “re(de)construct” jets



How jets come to have the structure they do



How we "re(de)construct" jets



Take all particles in a jet of radius R and recluster them into **subjets** with a jet definition with radius

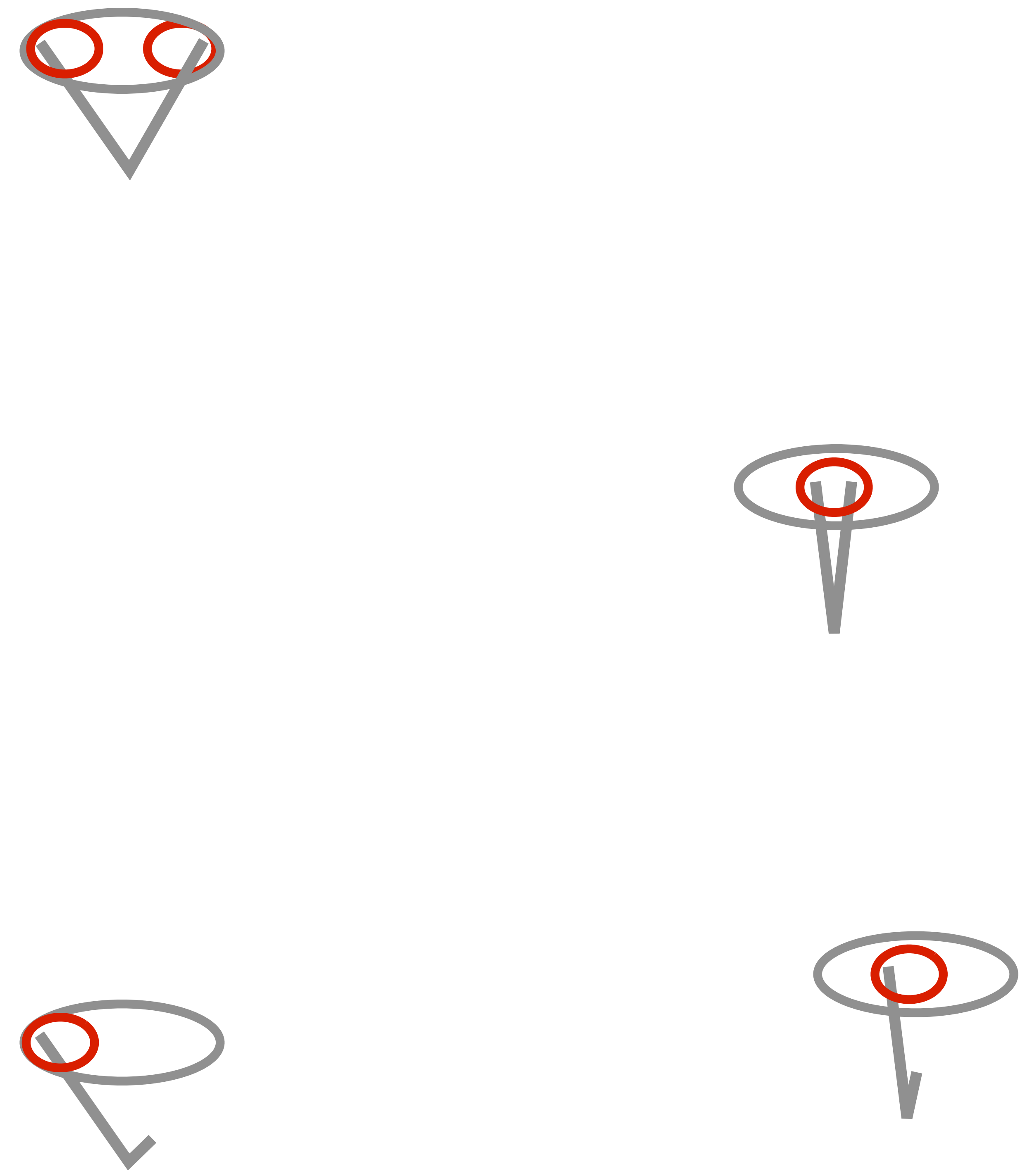
$R_{sub} < R$

The subjets that satisfy the condition

$p_t^{(subjet)} > z_{cut}\, p_t^{(jet)}$

are kept and merged to form the trimmed jet.

Krohn, Thaler & Wang '09



Take all particles in a jet of radius R and recluster them into **subjets** with a jet definition with radius

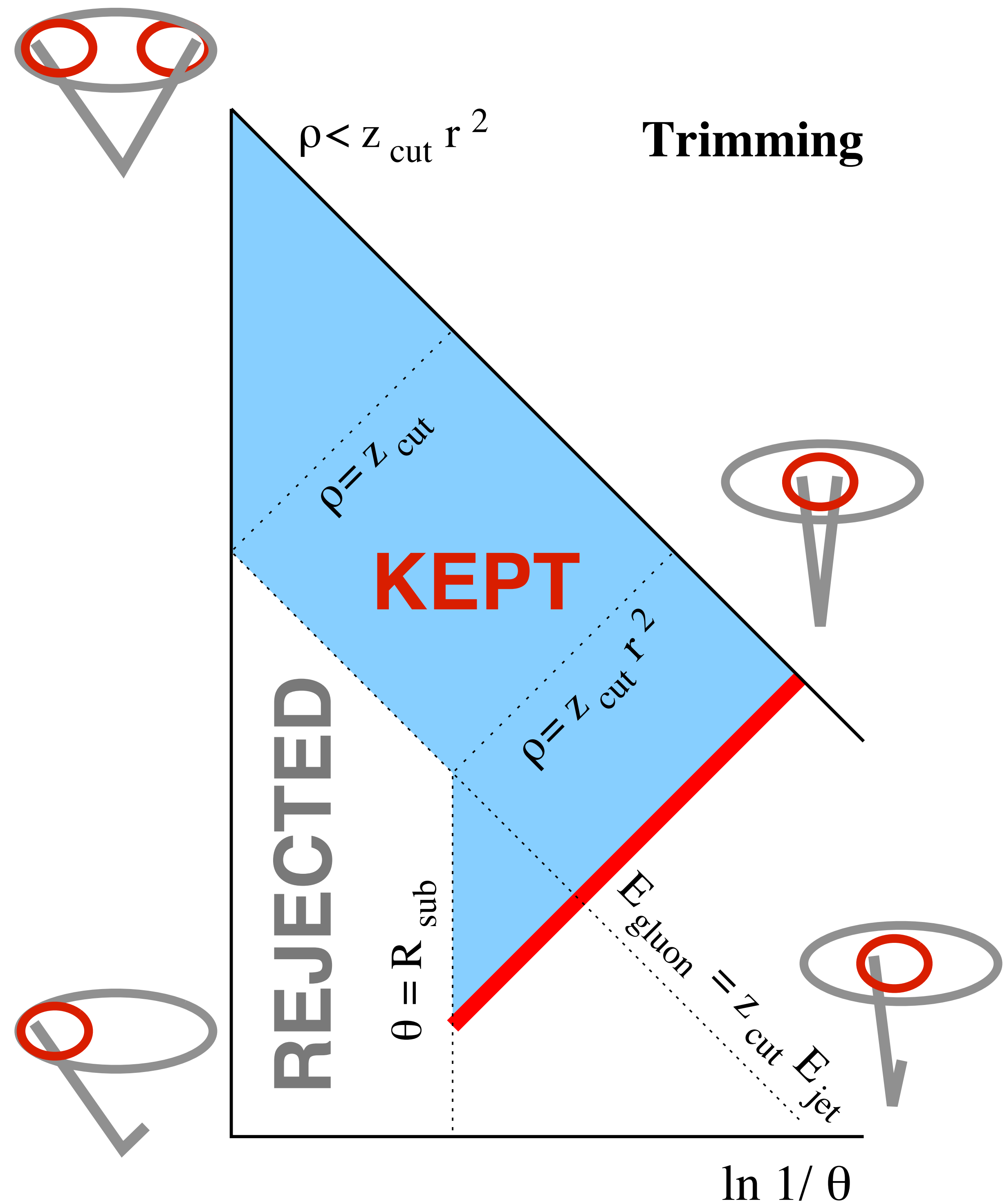
$R_{sub} < R$

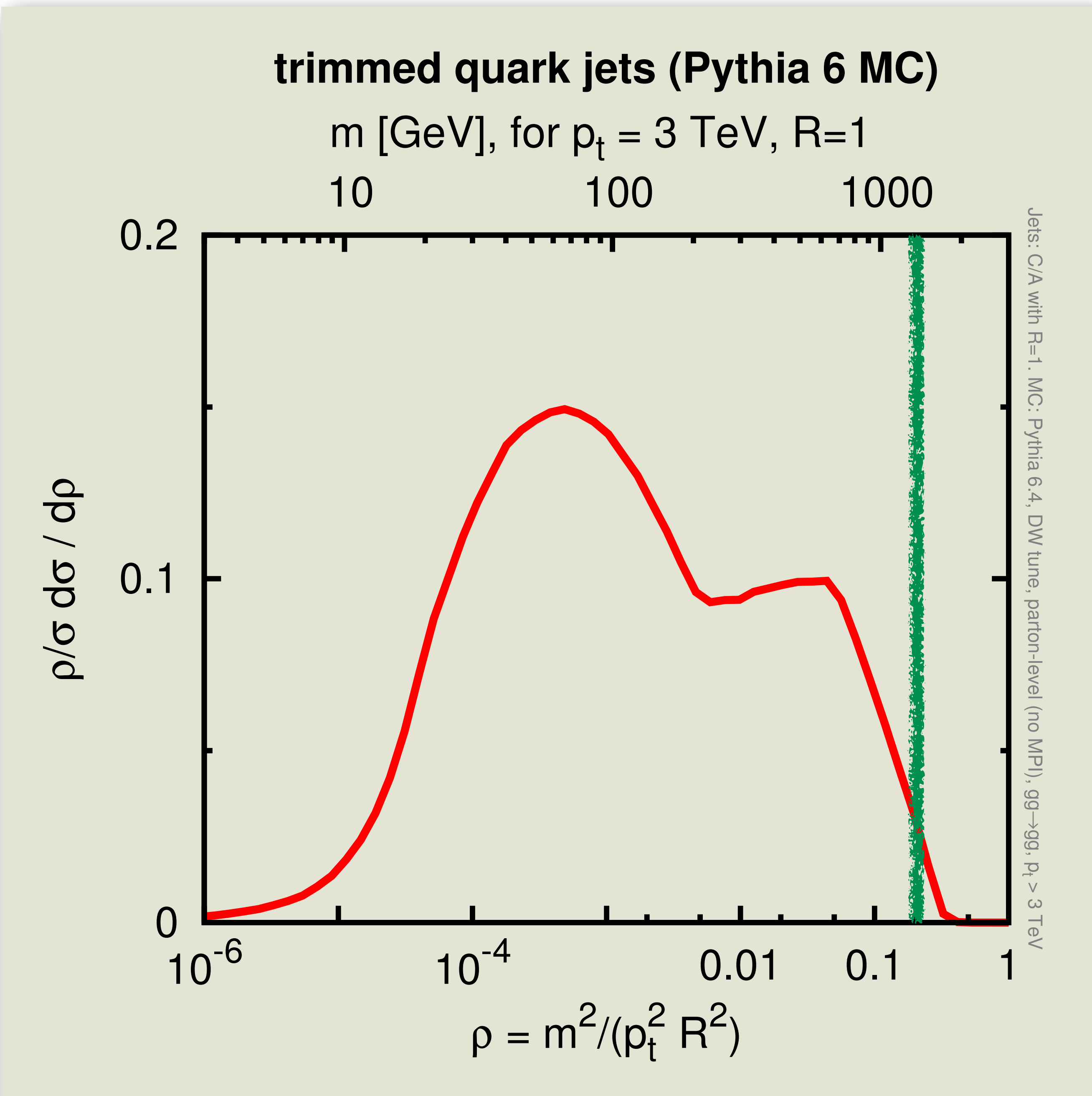
The subjets that satisfy the condition

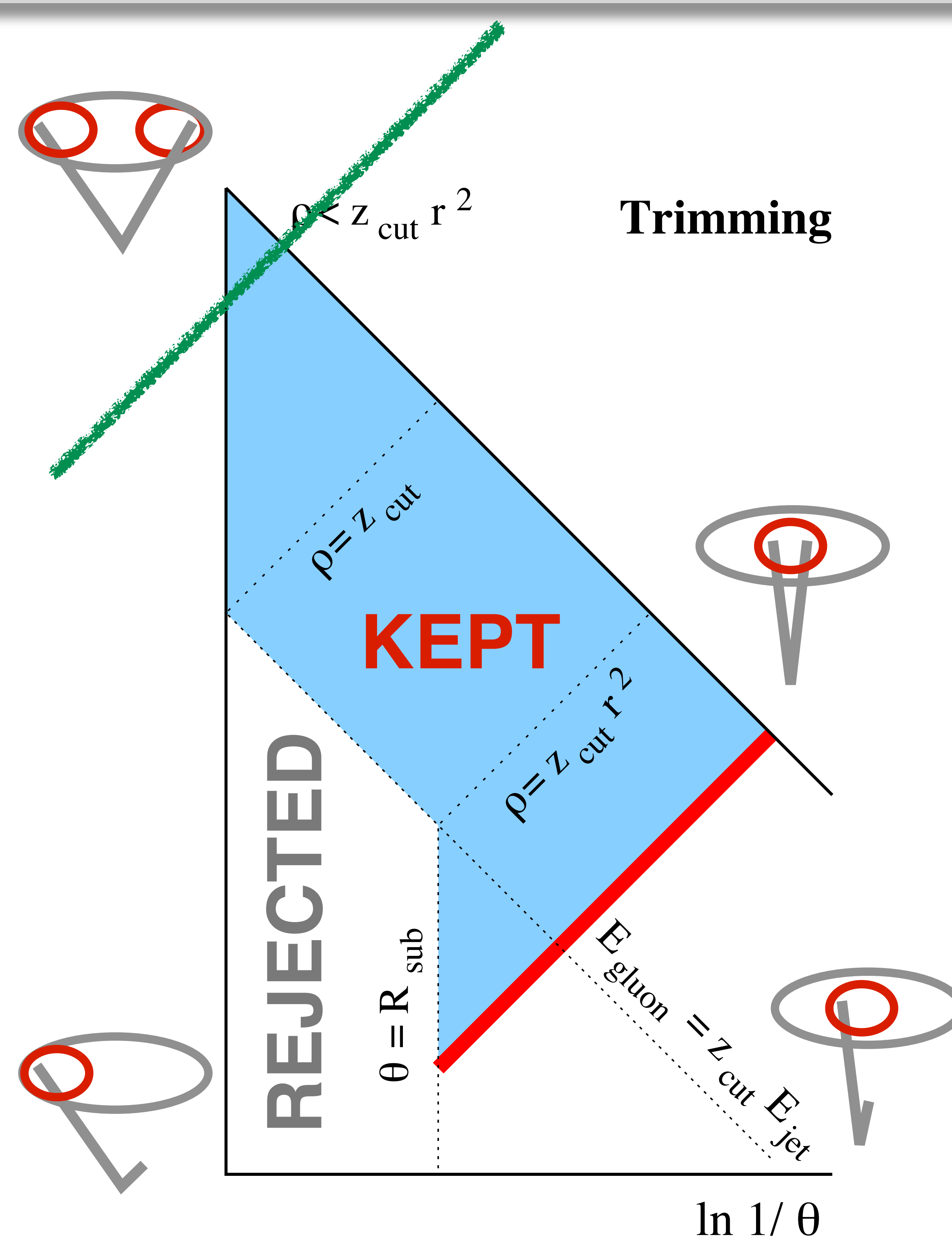
$p_t^{(subjet)} > z_{cut}\, p_t^{(jet)}$

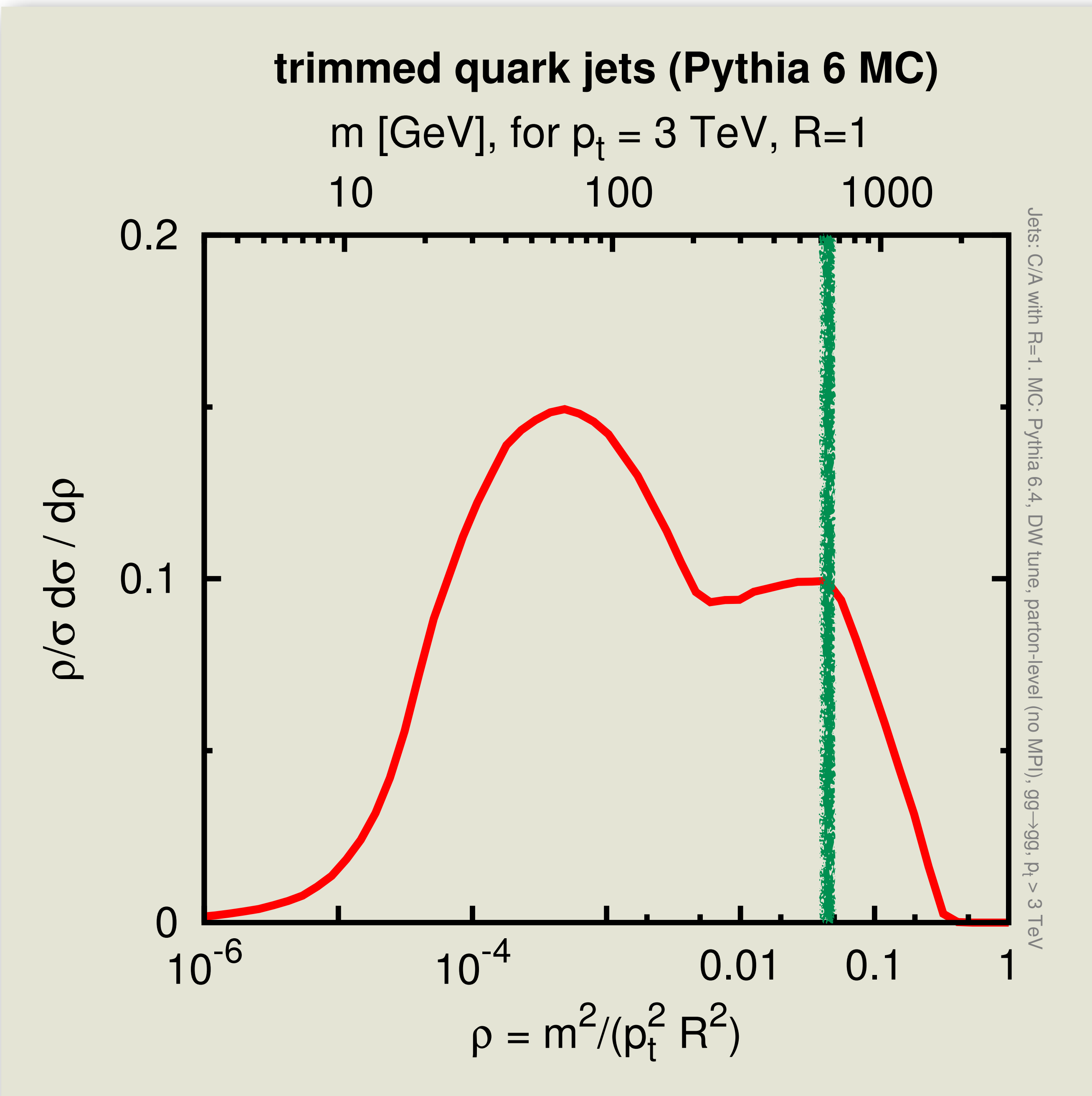
are kept and merged to form the trimmed jet.

Krohn, Thaler & Wang '09

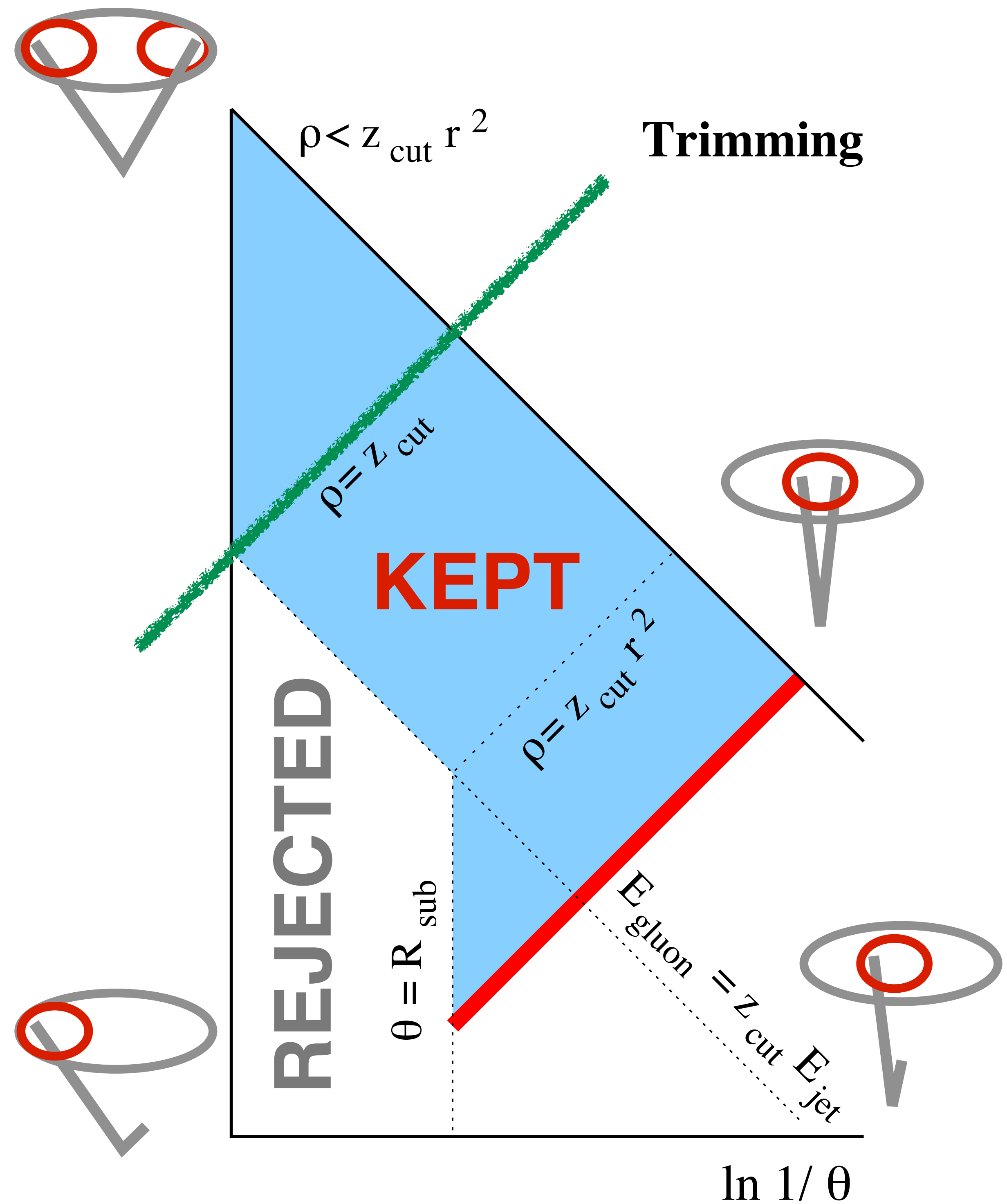



trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\, d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV

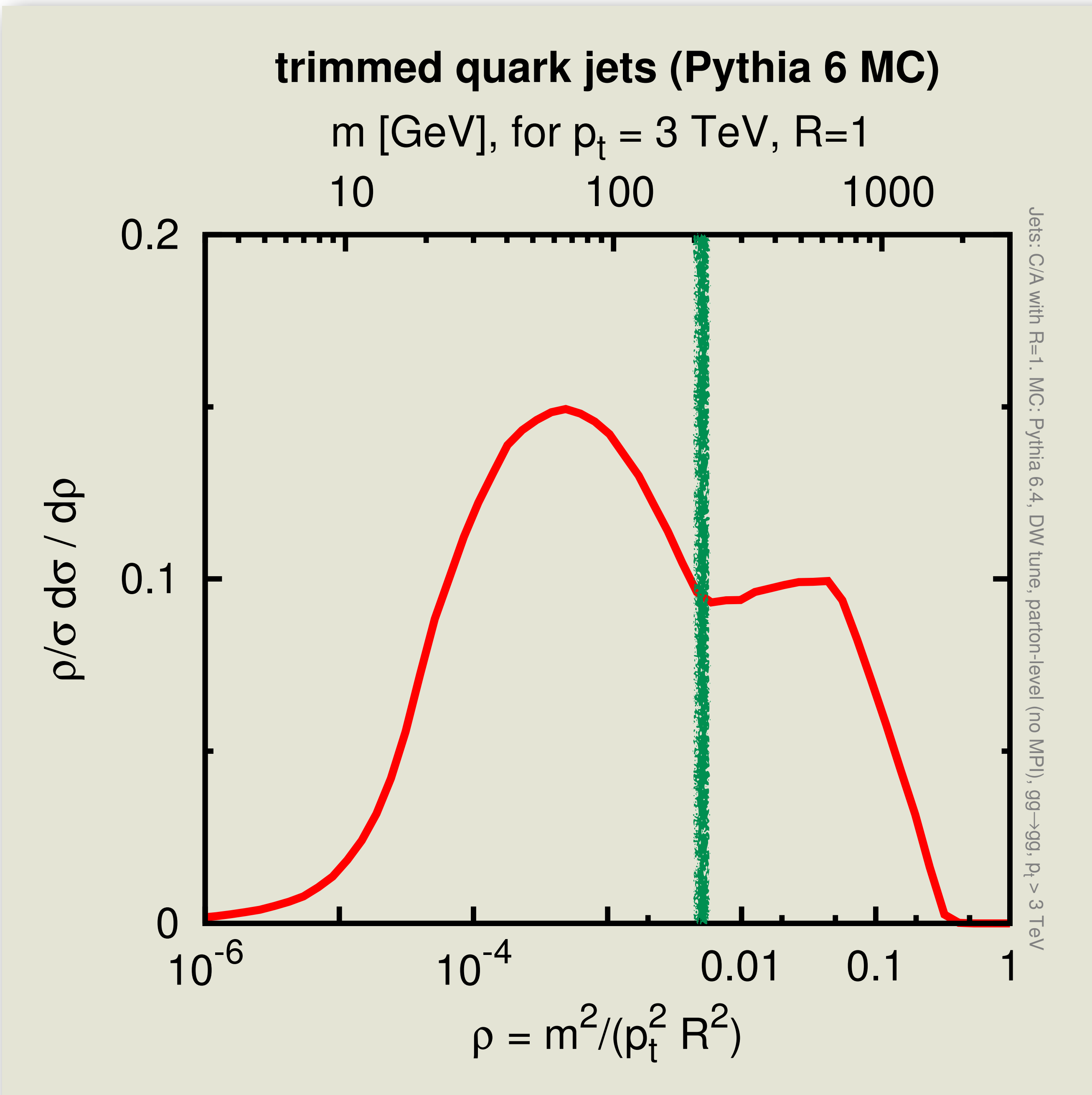

$\rho < z_{cut}\, r^2$
Trimming
$\rho = z_{cut}$
KEPT
$\rho = z_{cut}\, r^2$
REJECTED
$\theta = R_{sub}$
$E_{gluon} = z_{cut}\, E_{jet}$
ln 1/ θ



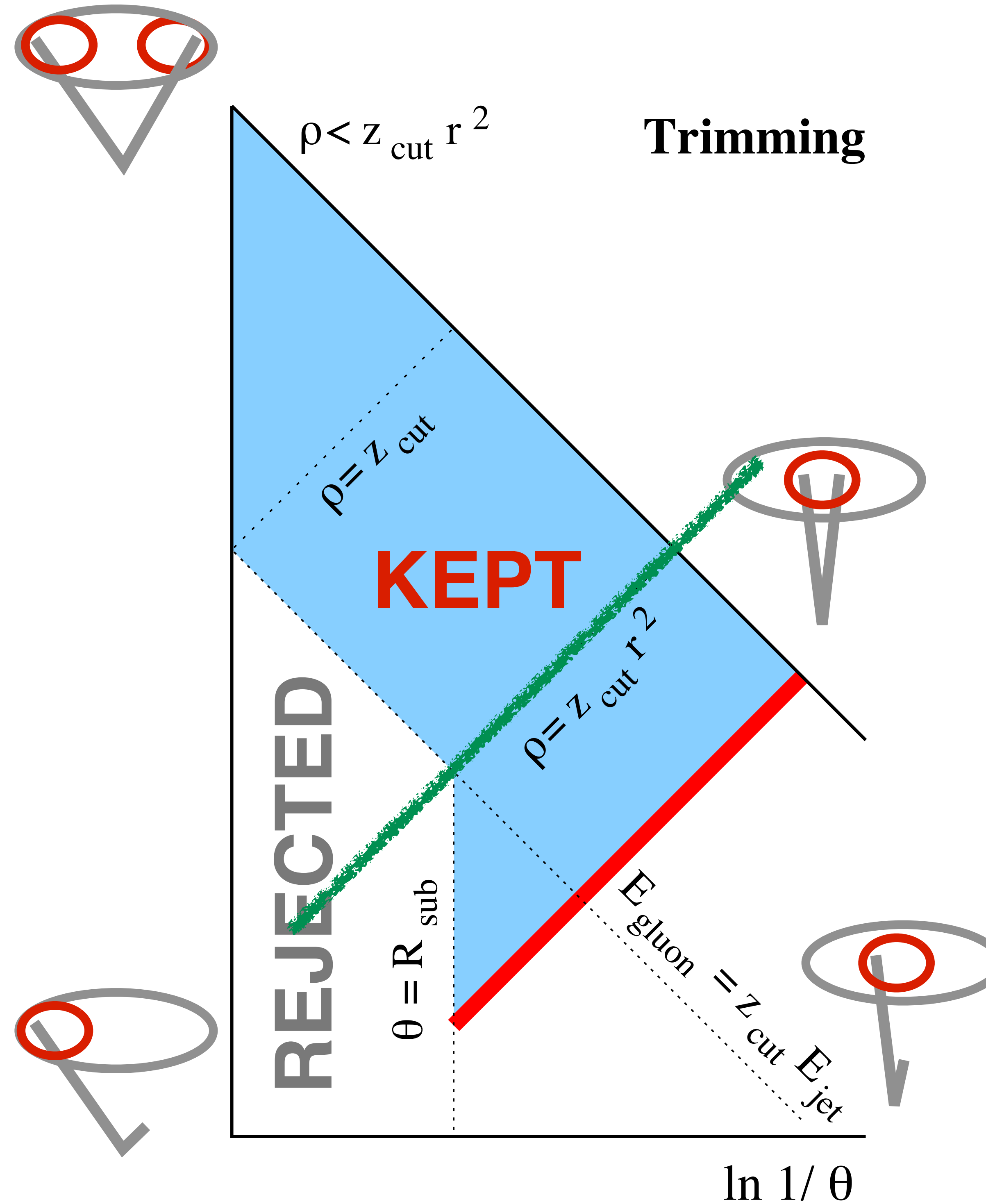
trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma \, d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV



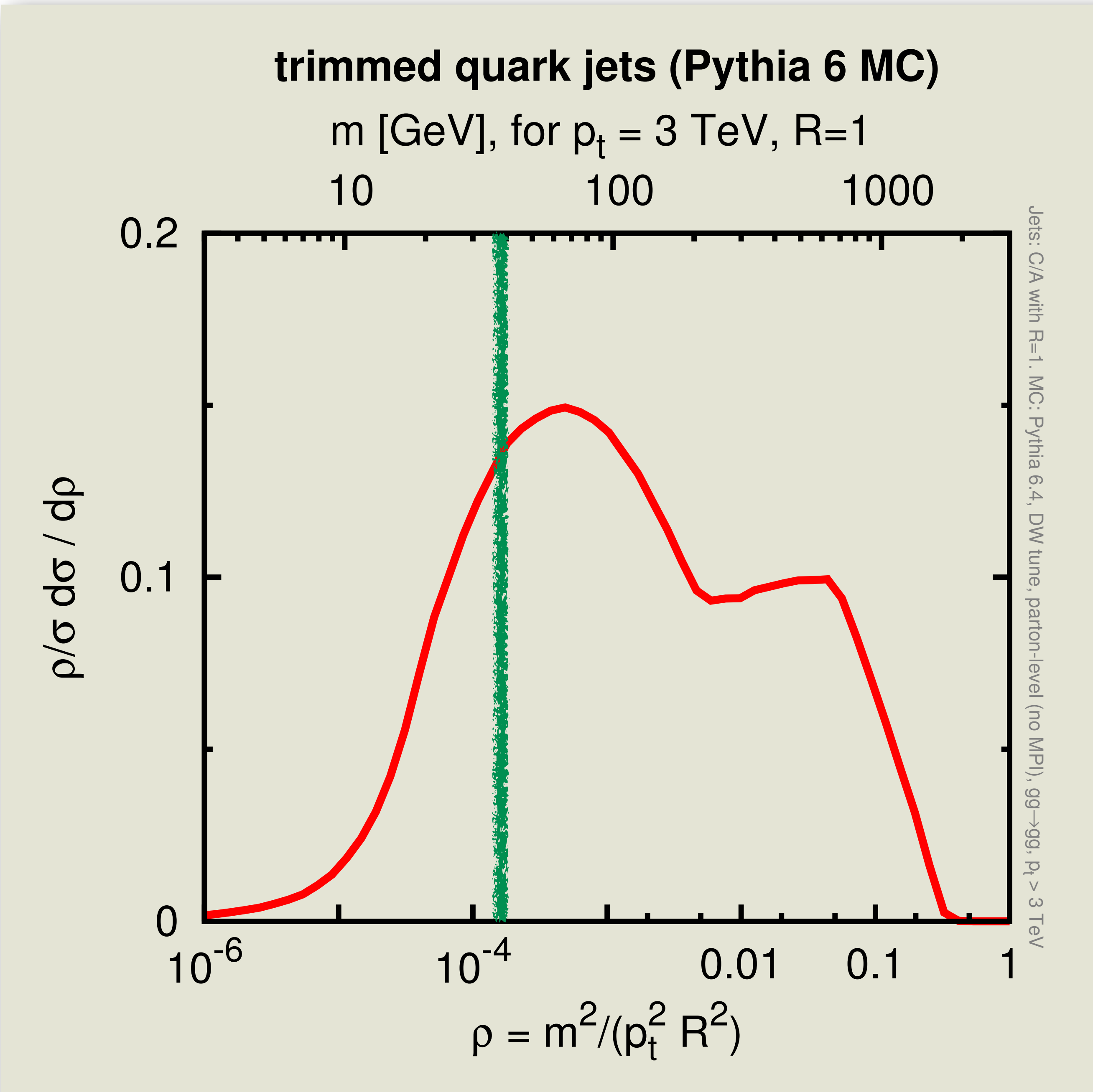
$\rho < z_{cut} r^2$
Trimming
$\rho = z_{cut}$
KEPT
$\rho = z_{cut} r^2$
REJECTED
$\theta = R_{sub}$
$E_{gluon} = z_{cut} E_{jet}$
ln 1/ θ

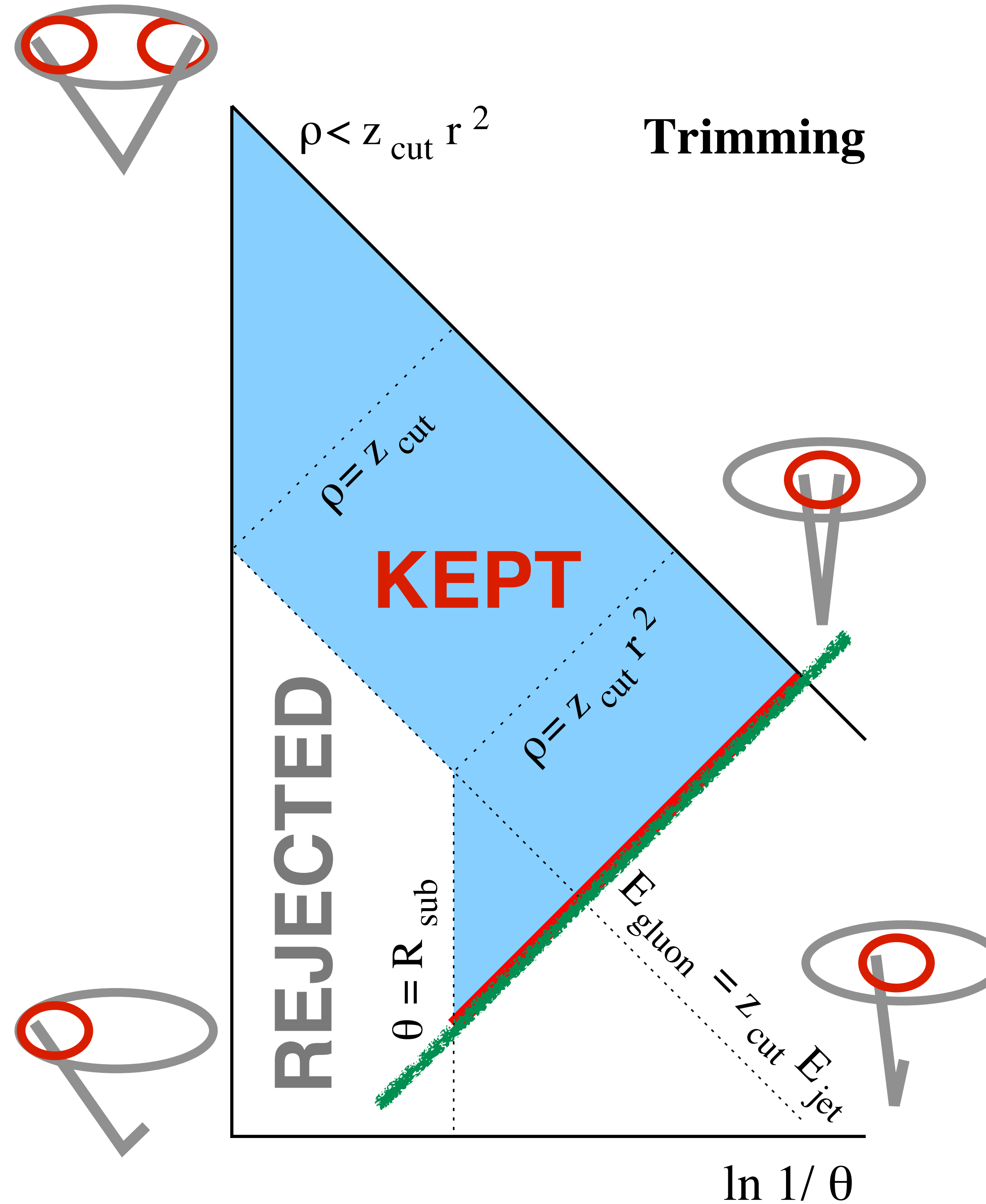


trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\ d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV

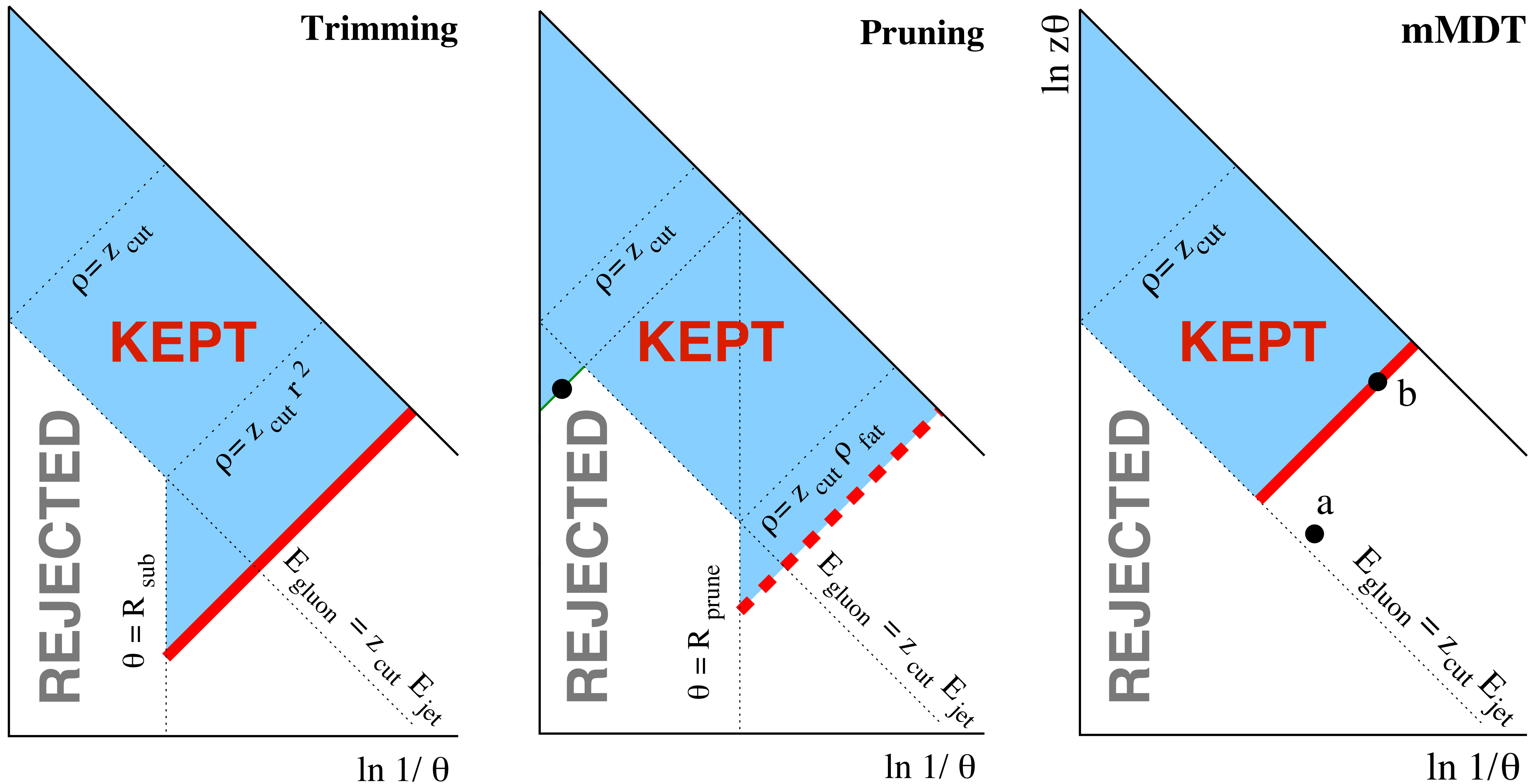


$\rho < z_{cut}\, r^2$
Trimming
$\rho = z_{cut}$
KEPT
$\rho = z_{cut}\, r^2$
REJECTED
$\theta = R_{sub}$
$E_{gluon} = z_{cut}\, E_{jet}$
ln 1/ θ

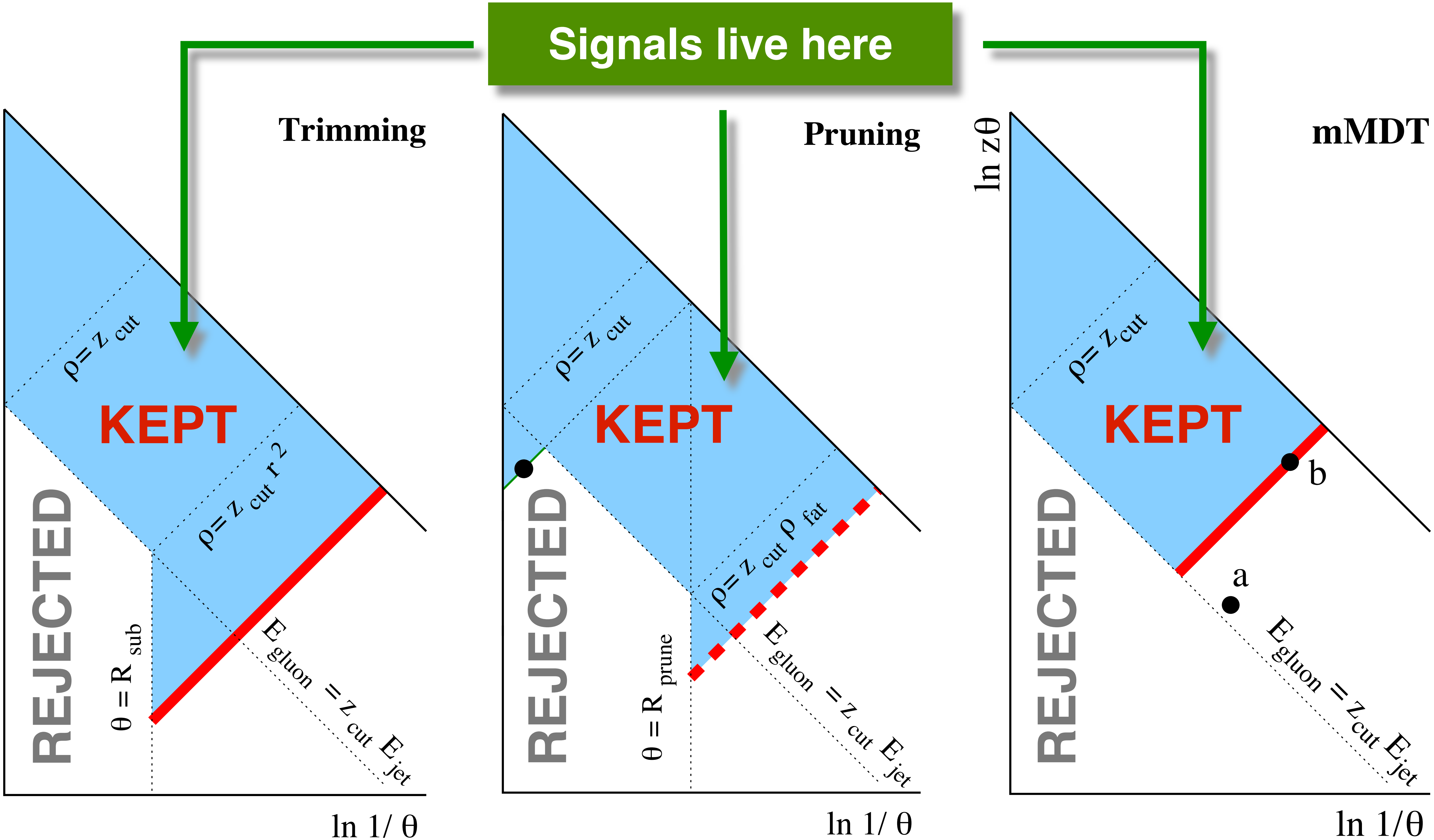

trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\, d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV


$\rho < z_{cut}\, r^2$
Trimming
$\rho = z_{cut}$
KEPT
$\rho = z_{cut}\, r^2$
REJECTED
$\theta = R_{sub}$
$E_{gluon} = z_{cut}\, E_{jet}$
ln 1/ θ

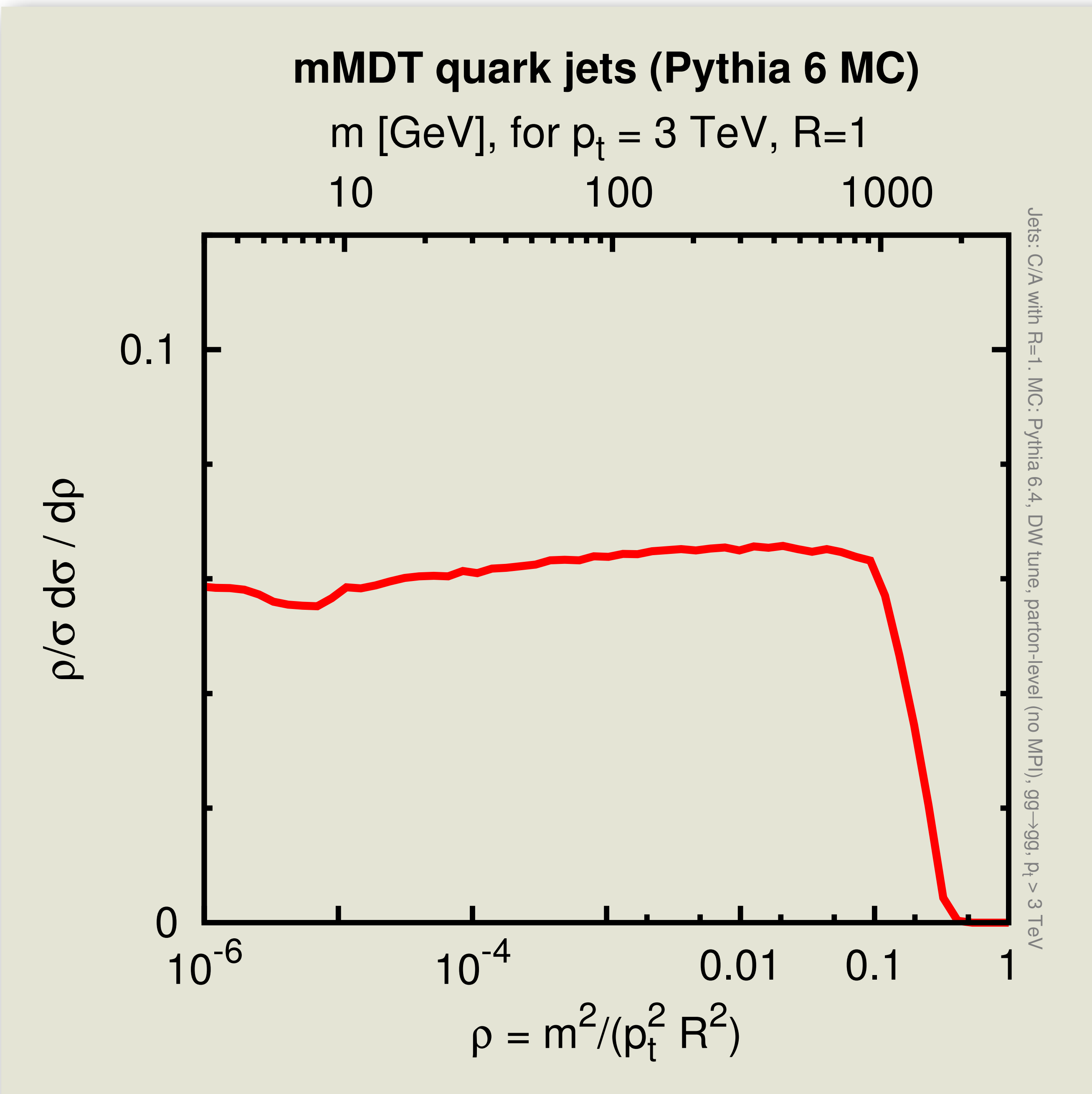
Kinematic regions for different taggers

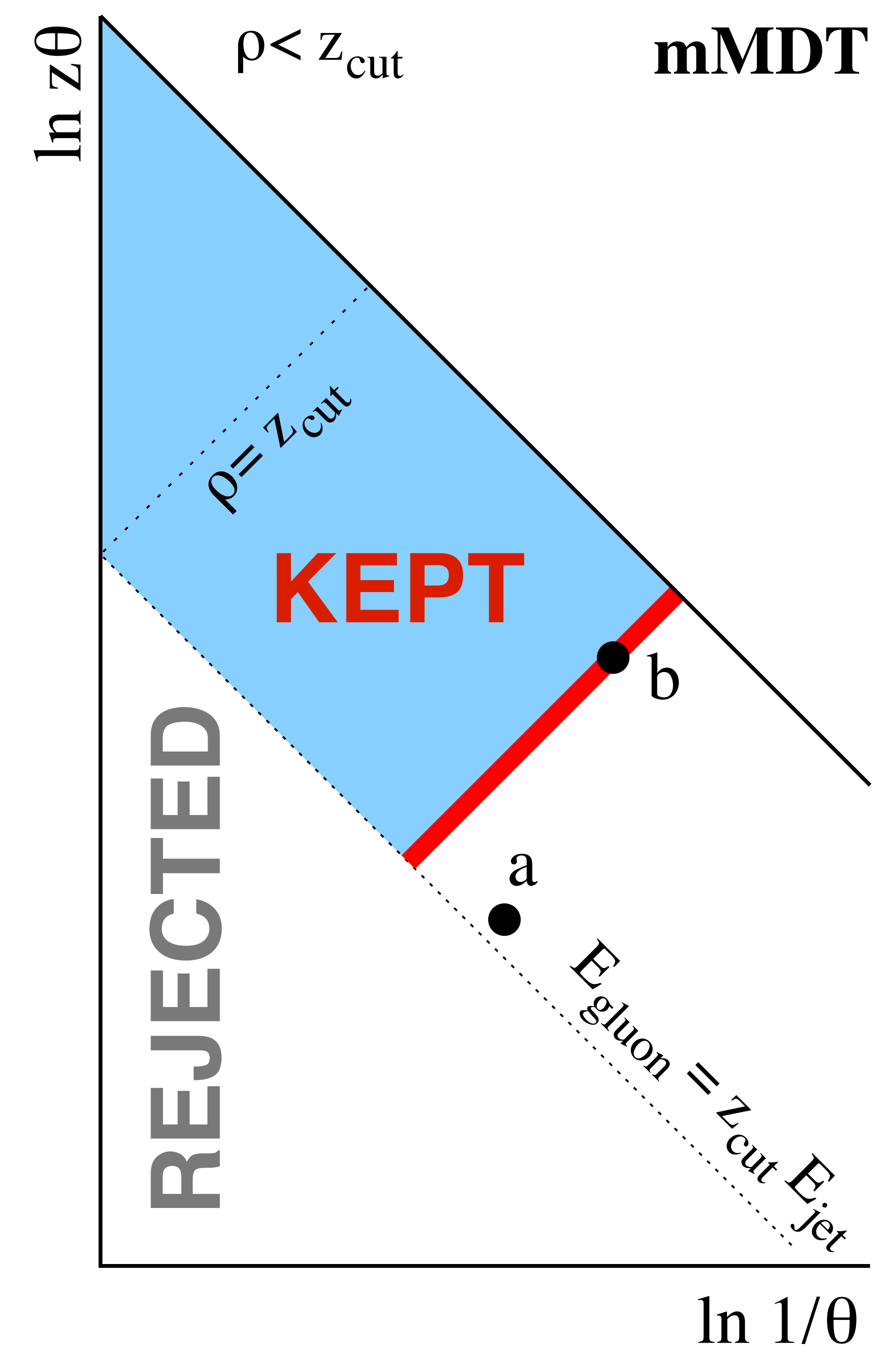


Kinematic regions for different taggers



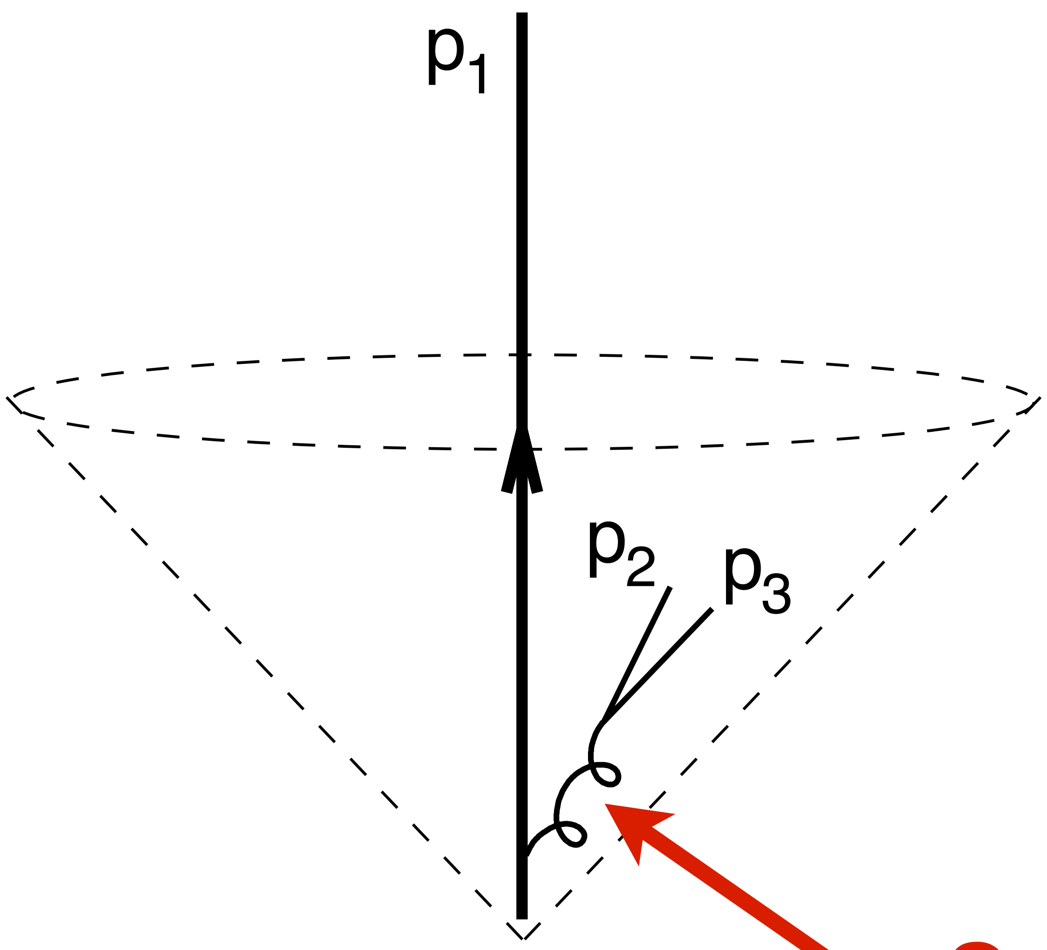
simple kinematic region → simple mass dist^{n}





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, $y = \min(p_{tj_1}^2, p_{tj_2}^2)\Delta R_{j_1 j_2}^2 / m_j^2 > y_{\text{cut}}$, 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:

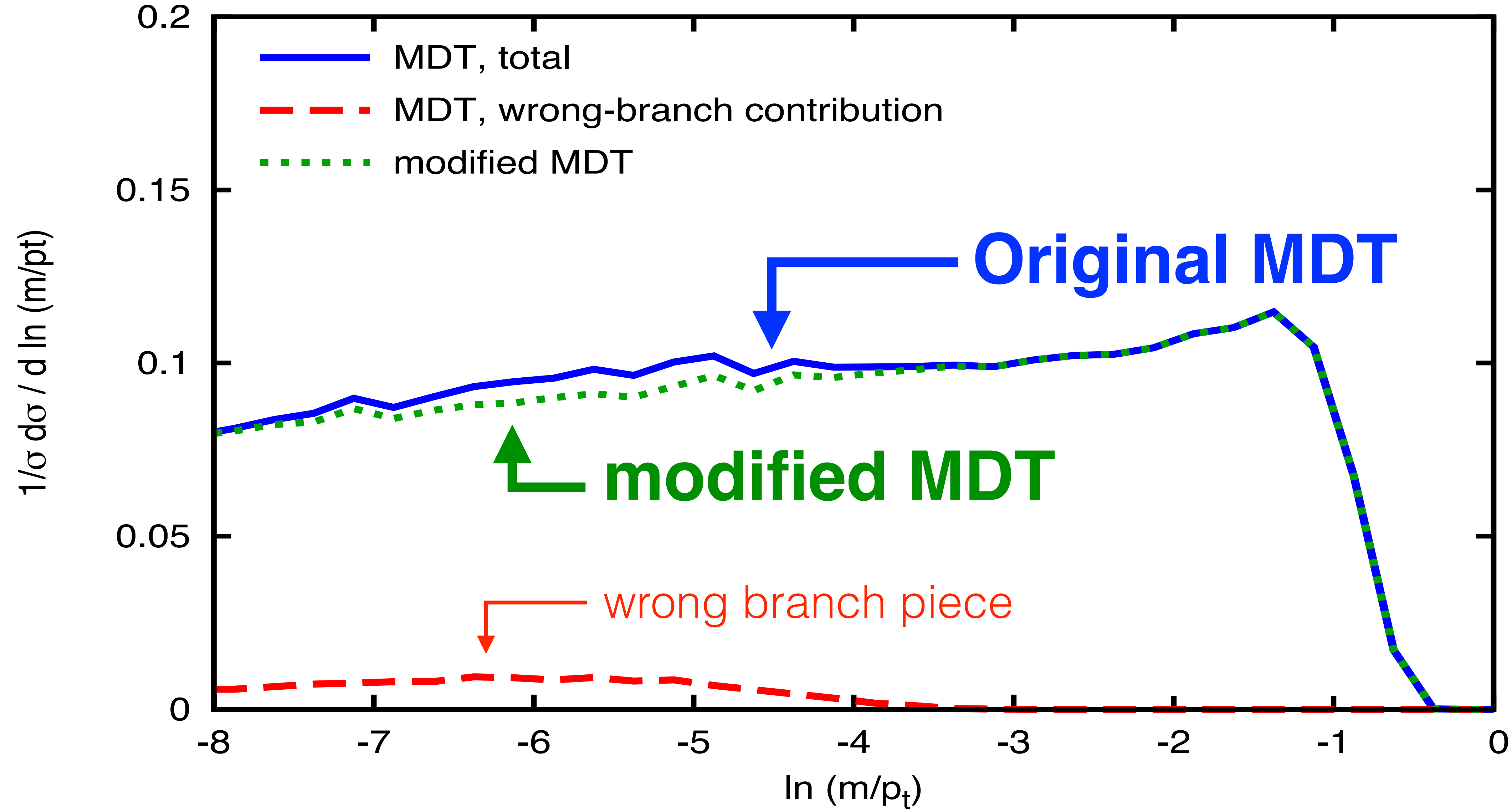
Follows a soft branch ($p_2+p_3 < y_{cut}\ p_{jet}$) with “accidental” small mass, when the “right” answer was that the (massless) hard branch had no substructure

Subjet is soft, but has more substructure than hard subjet

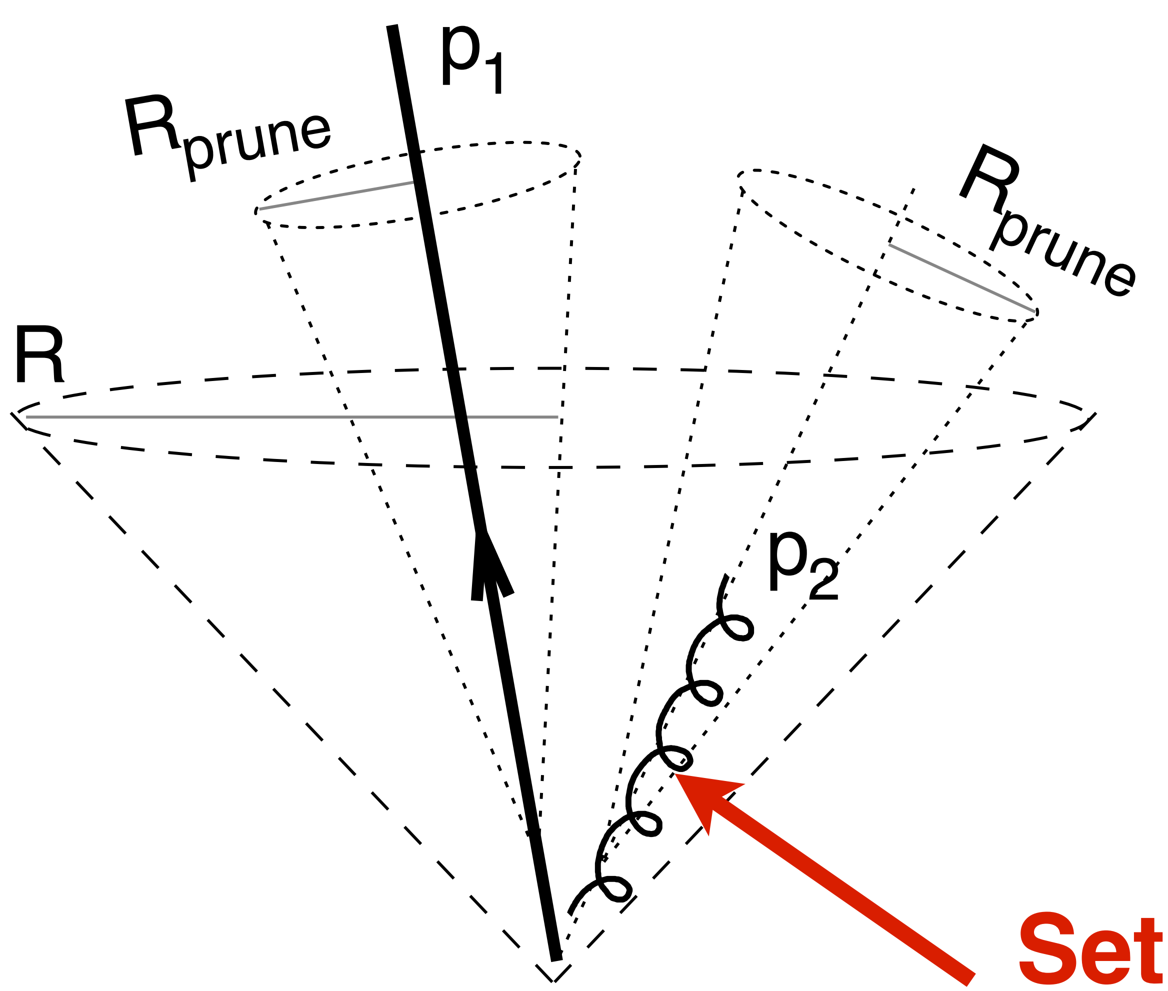
A simple fix: “modified” Mass Drop Tagger:

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

(rather than the one with larger m)



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.



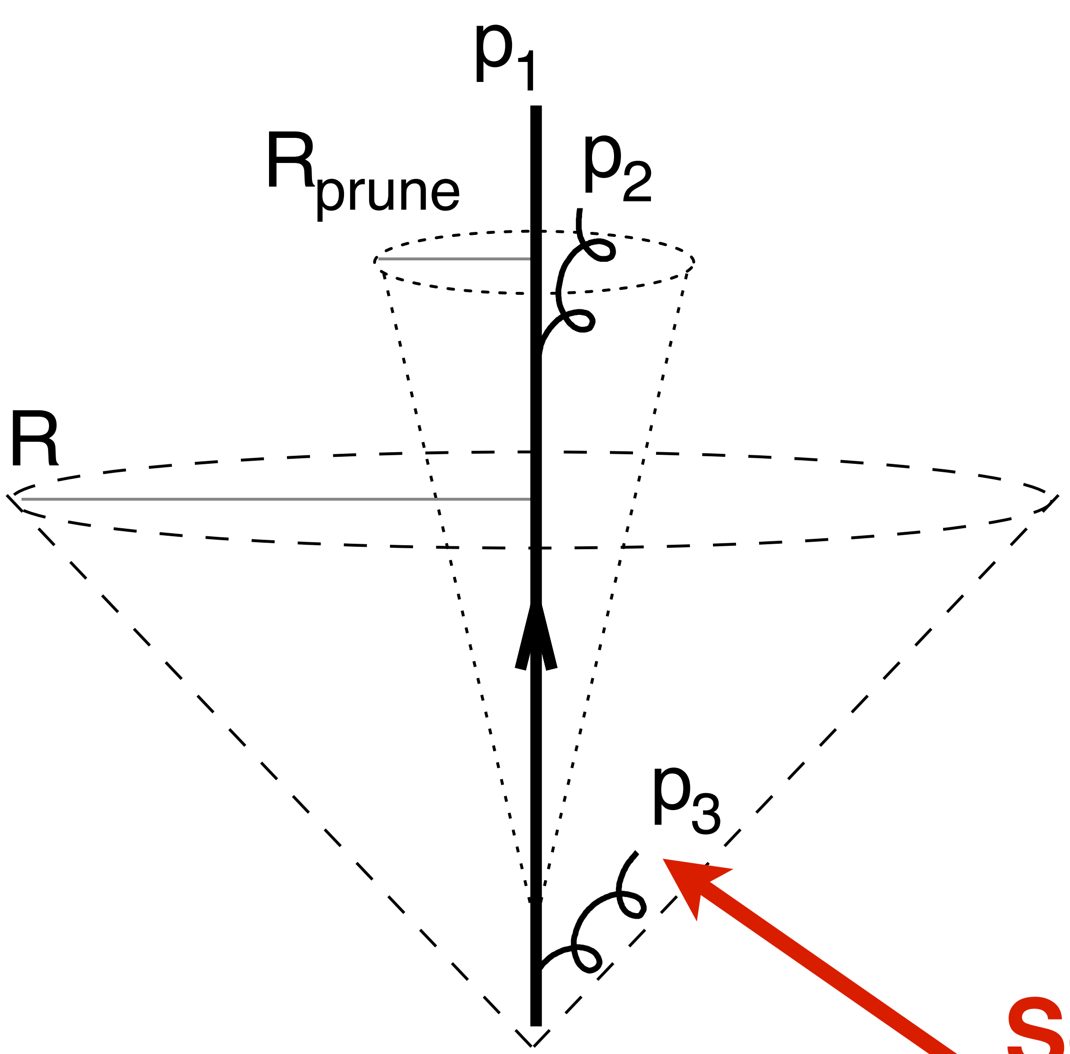
What pruning is meant to do:

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.



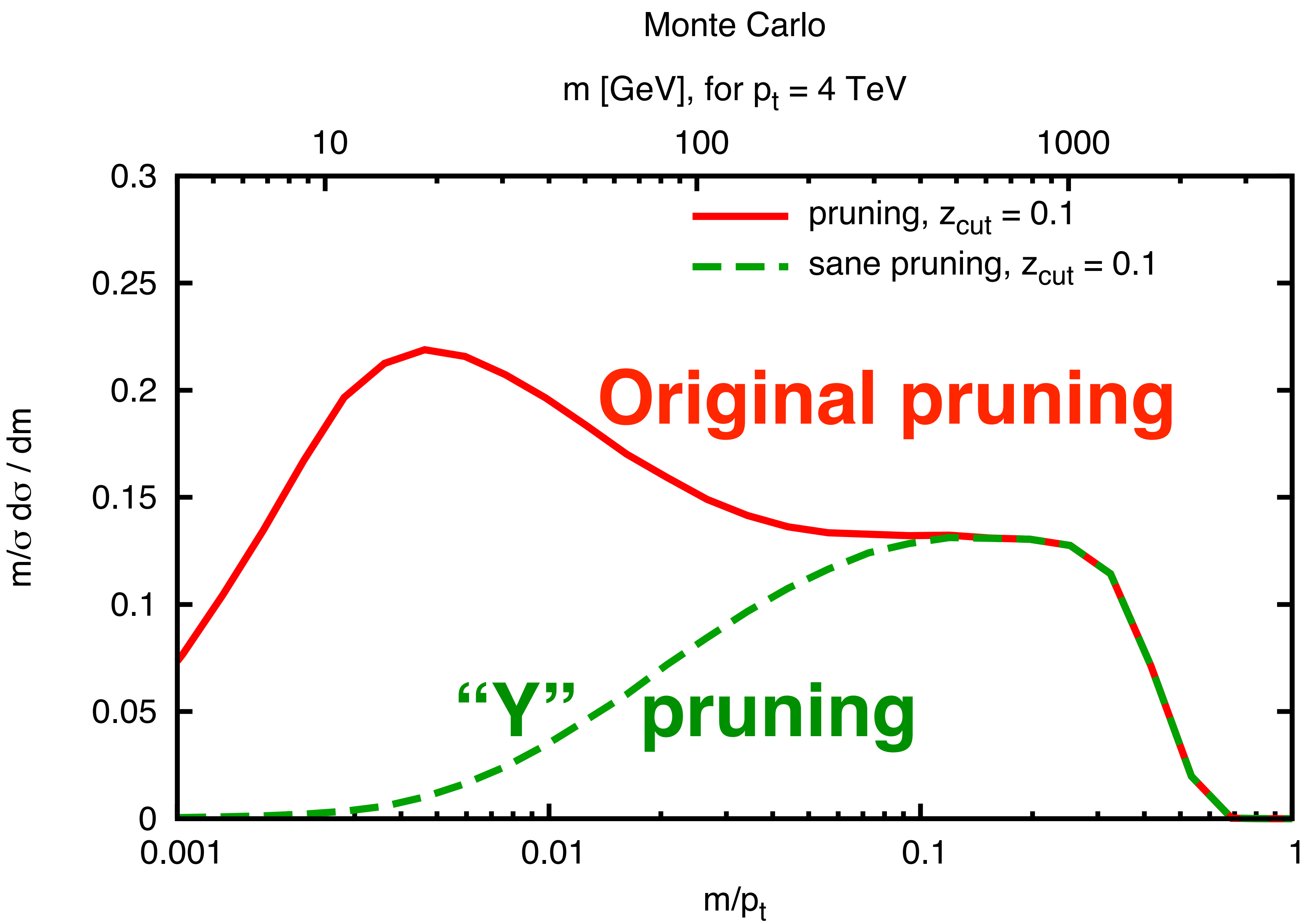
Sets pruning radius, but gets pruned away

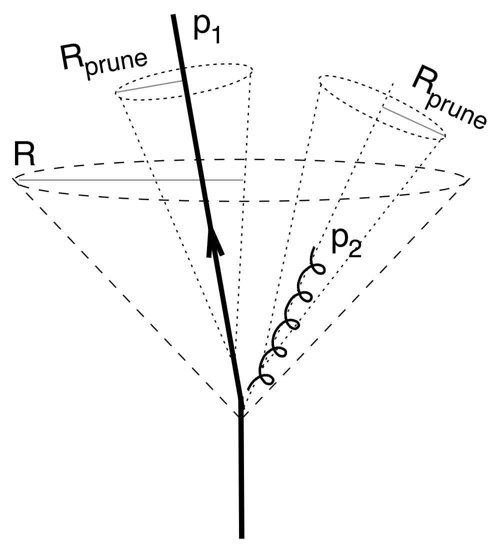
<u>What pruning sometimes does</u>

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).

A simple fix: "Y" pruning

Require at least one successful merging with $\Delta R > R_{prune}$ and $z > z_{cut}$ – forces 2-pronged ("Y") configurations





"Y" pruning is effectively placing an isolation cut on radiation around the tagged object

What about actual calculations of the taggers?

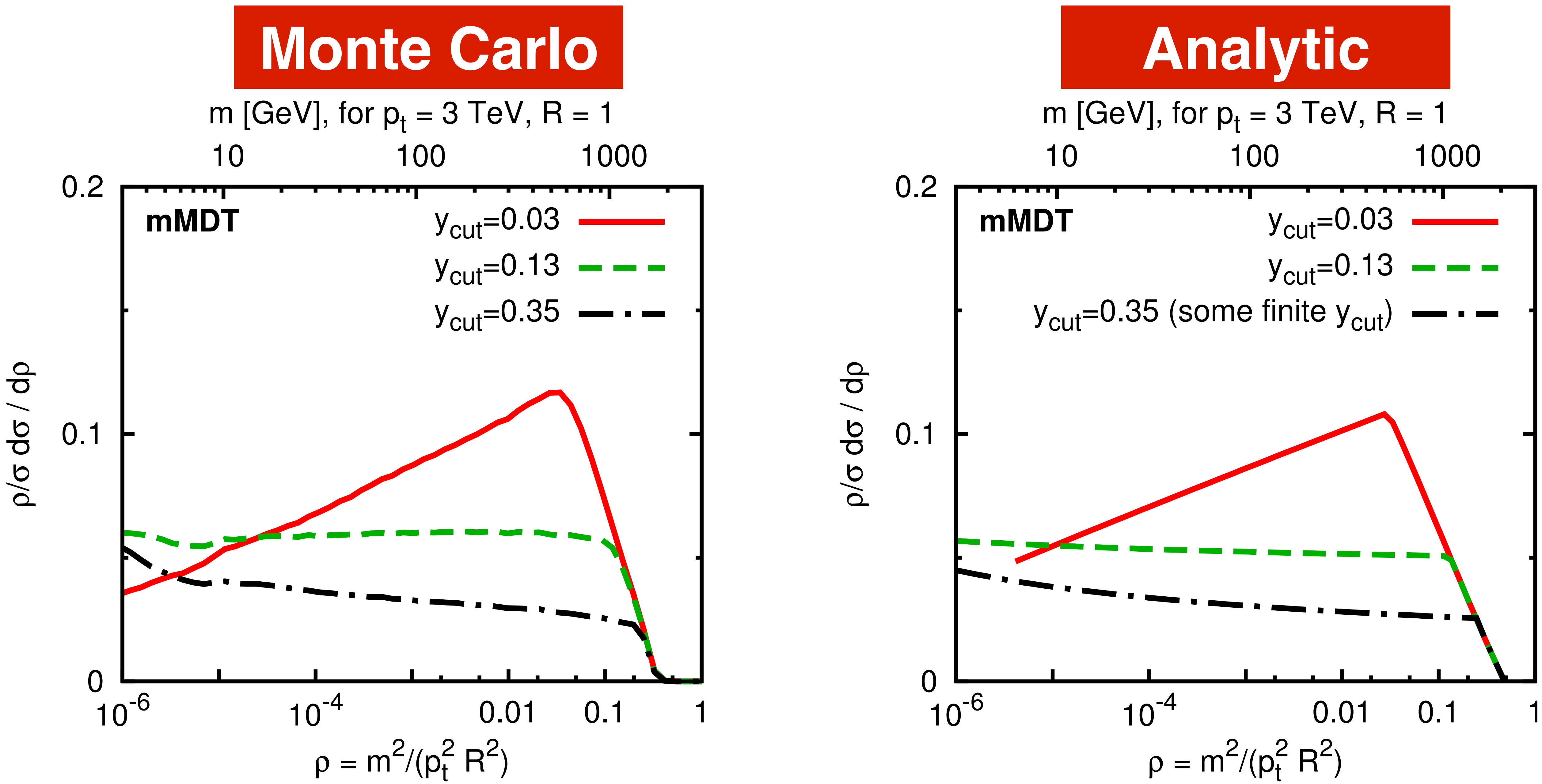
Simpler → More complex

Dasgupta et al (full NLL, including non-global part)
Chien et al (partial NNLL)
Jouttenus et al (they say full NNLL; I remain to be convinced!)

mMDT > plain mass ~ trimming > Y-pruning > full pruning ≫ ~~MDT~~

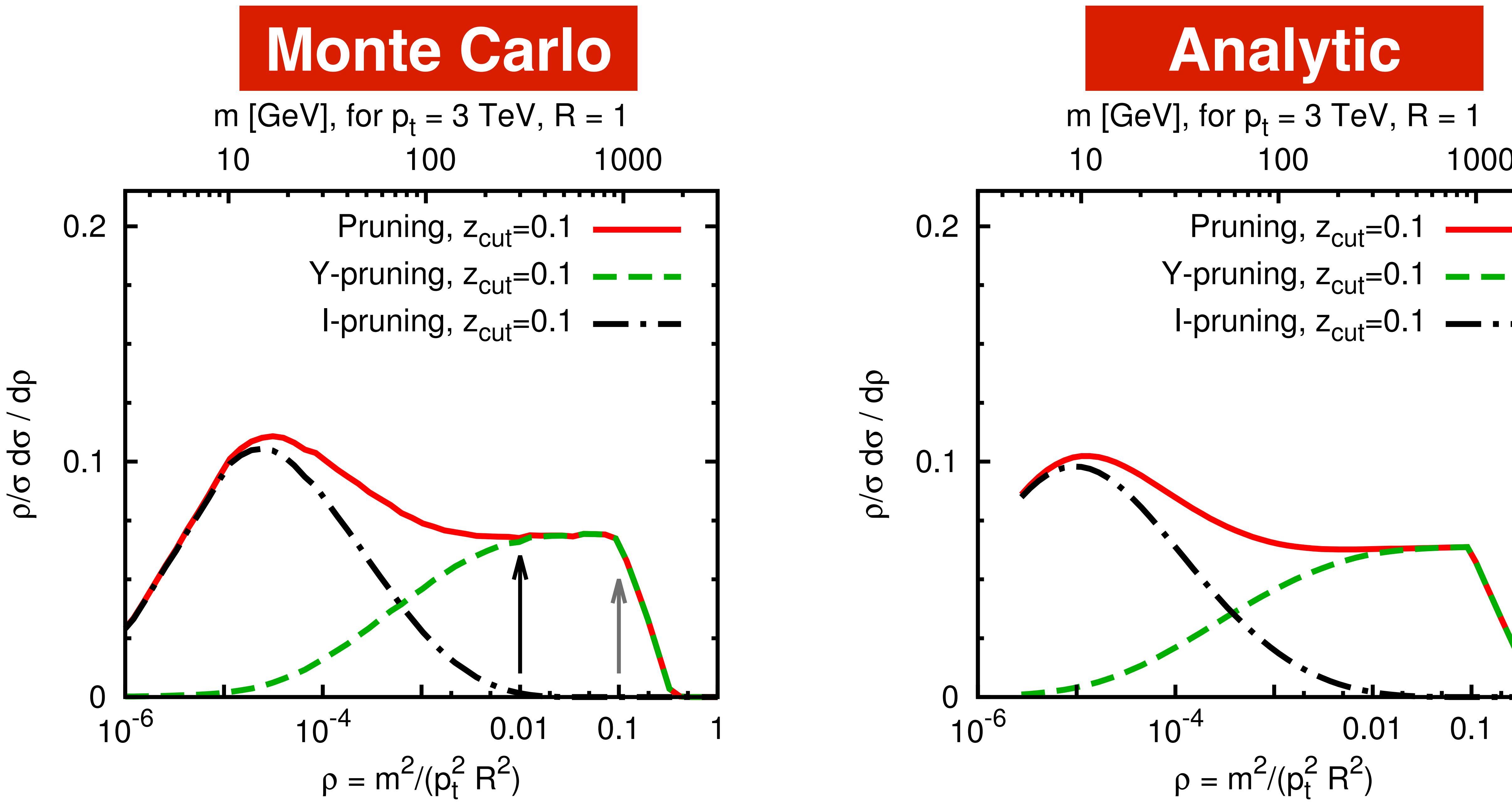
LL in all cases, plus some subleading logs
[NB: LL doesn't mean the same thing in all cases!)

Modified Mass Drop Tagger

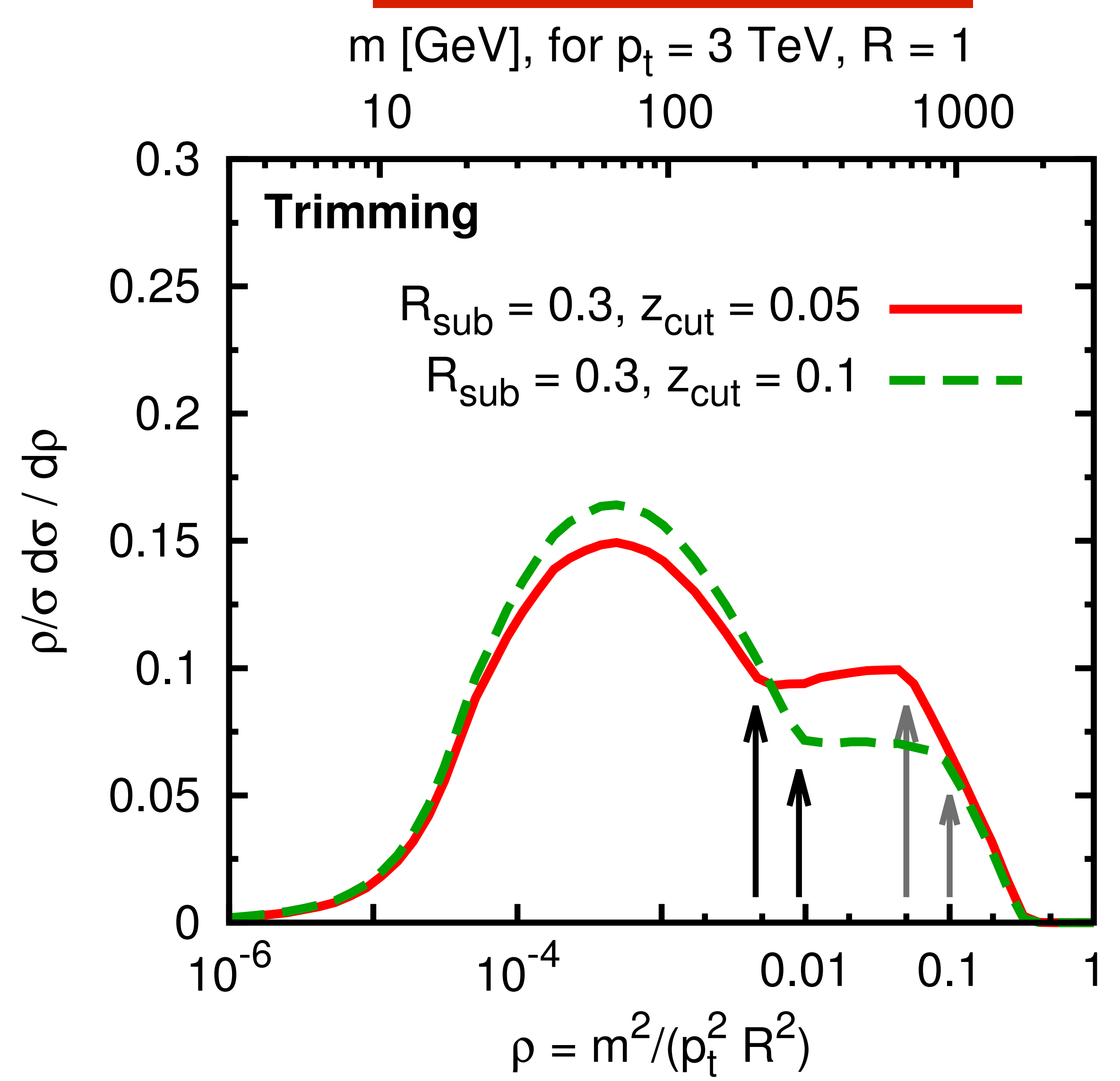


[mMDT is closest we have to a scale-invariant tagger, though exact behaviour depends on q/g fractions]

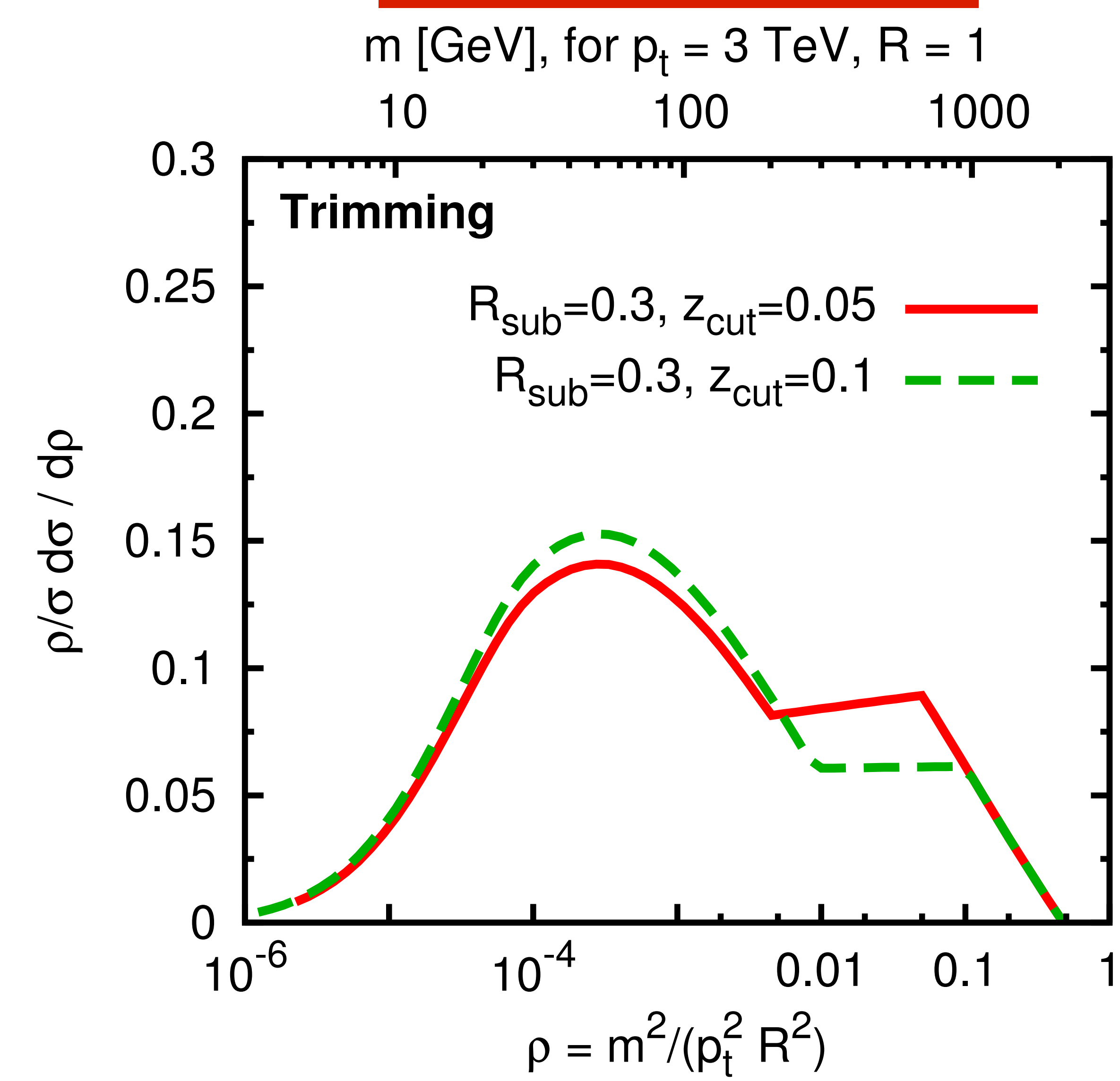
Pruning


Monte Carlo
m [GeV], for p_t = 3 TeV, R = 1
10
100
1000
0.2
0.1
0
Pruning, z_{cut}=0.1
Y-pruning, z_{cut}=0.1
I-pruning, z_{cut}=0.1
ρ/σ dσ / dρ
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Analytic
m [GeV], for p_t = 3 TeV, R = 1
10
100
1000
0.2
0.1
0
Pruning, z_{cut}=0.1
Y-pruning, z_{cut}=0.1
I-pruning, z_{cut}=0.1
ρ/σ dσ / dρ
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$

Monte Carlo



Analytic



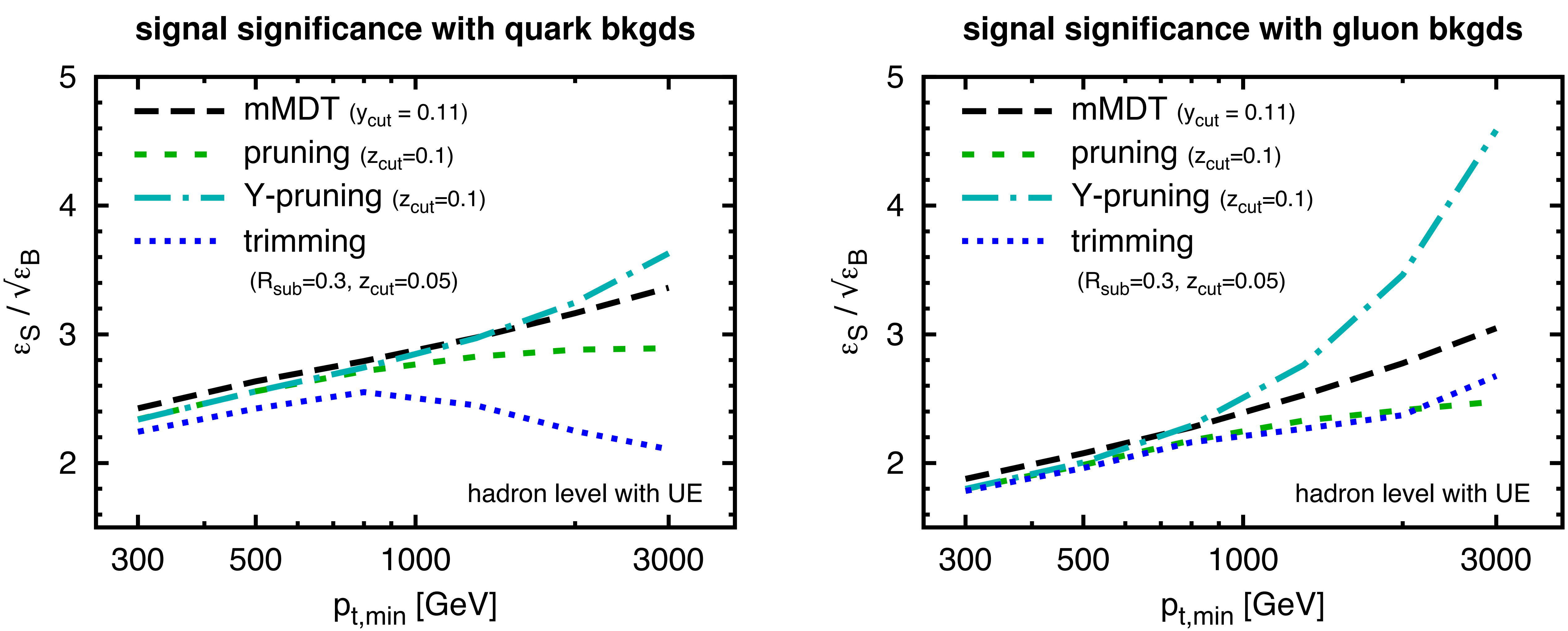
	highest logs	transition(s)	Sudakov peak	NGLs
plain mass	$\alpha_s^n L^{2n}$	—	$L \simeq 1/\sqrt{\bar{\alpha}_s}$	yes
trimming	$\alpha_s^n L^{2n}$	$z_{\mathrm{cut}},\ r^2 z_{\mathrm{cut}}$	$L \simeq 1/\sqrt{\bar{\alpha}_s} - 2\ln r$	yes
pruning	$\alpha_s^n L^{2n}$	$z_{\mathrm{cut}},\ z_{\mathrm{cut}}^2$	$L \simeq 2.3/\sqrt{\bar{\alpha}_s}$	yes
MDT	$\alpha_s^n L^{2n-1}$	$y_{\mathrm{cut}},\ \frac{1}{4}y_{\mathrm{cut}}^2,\ y_{\mathrm{cut}}^3$	—	yes
Y-pruning	$\alpha_s^n L^{2n-1}$	z_{cut}	(Sudakov tail)	yes
mMDT	$\alpha_s^n L^n$	y_{cut}	—	no

NEW

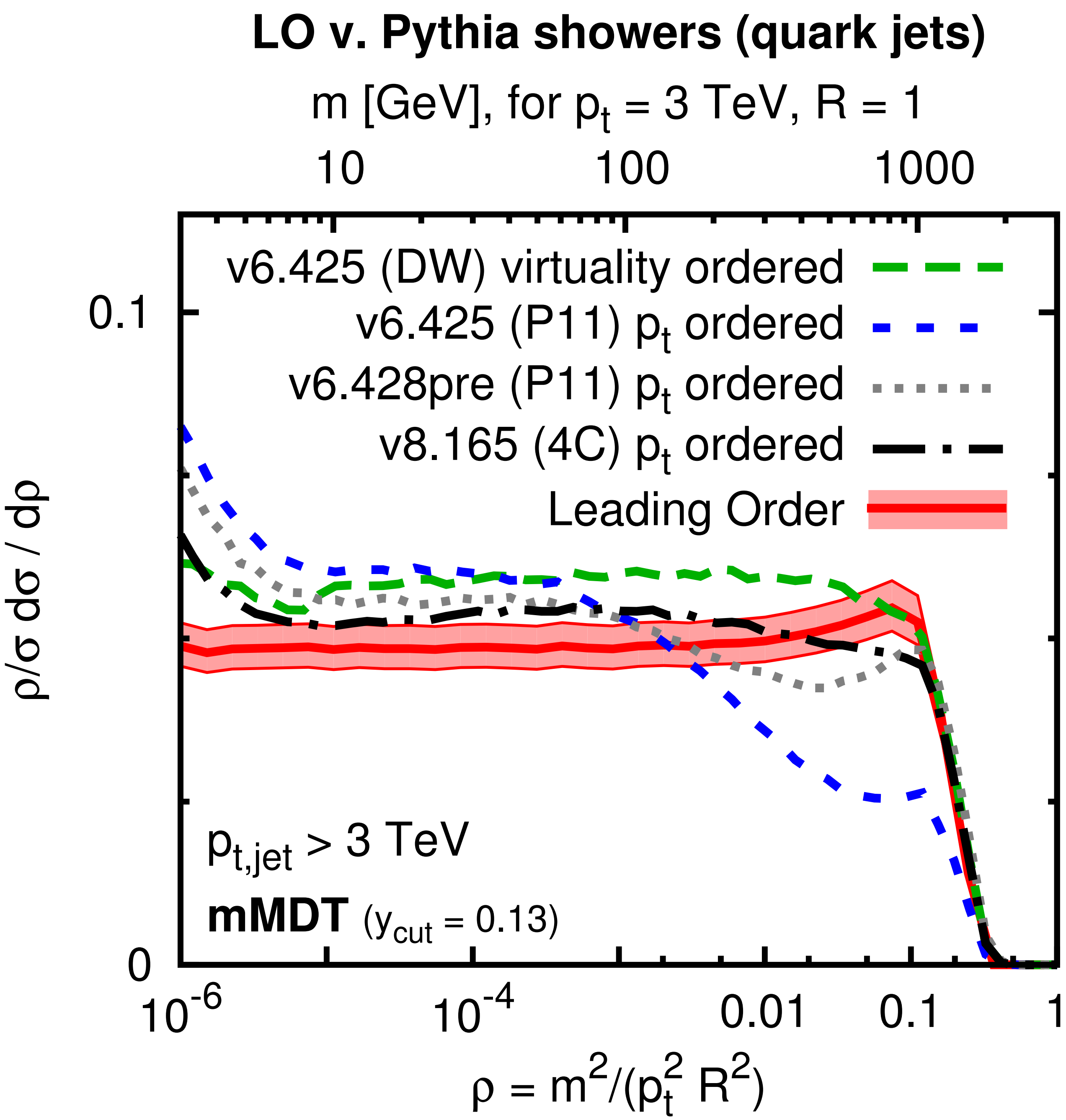
Special: only single logarithms (L = ln ρ)
→ more accurately calculable

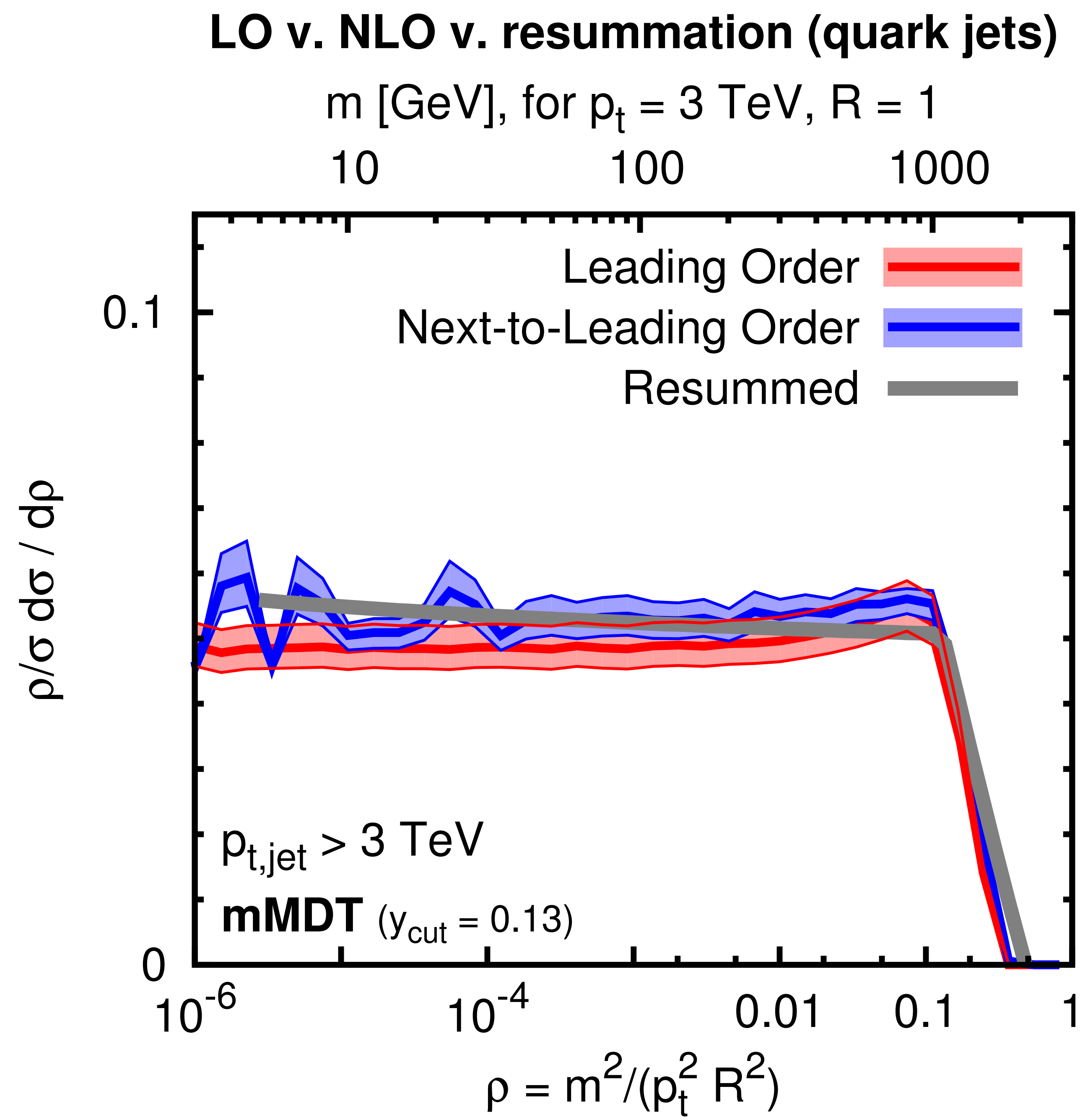
Special: better exploits signal/bkgd differences

Performance for finding signals



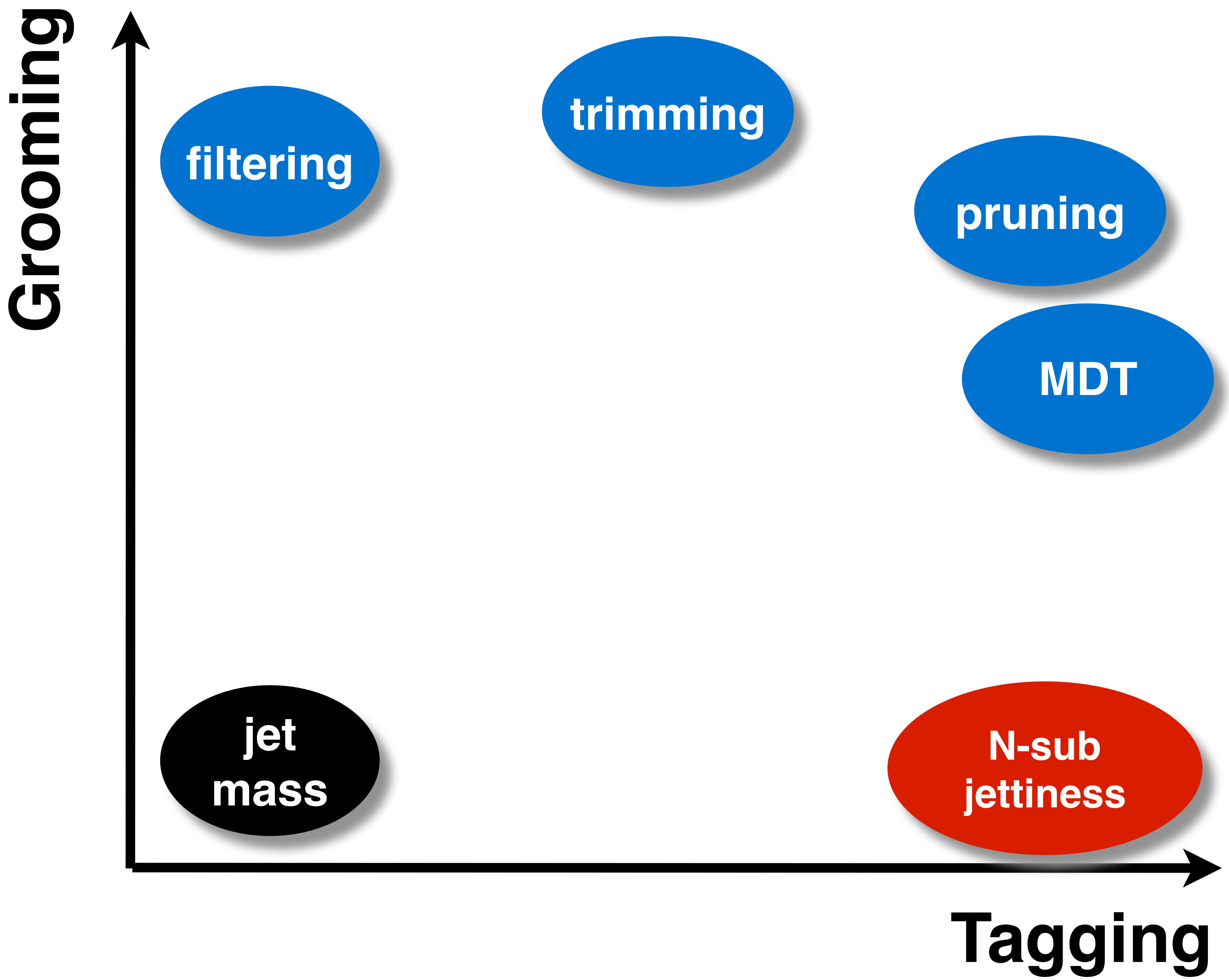
Comparing MC & other tools



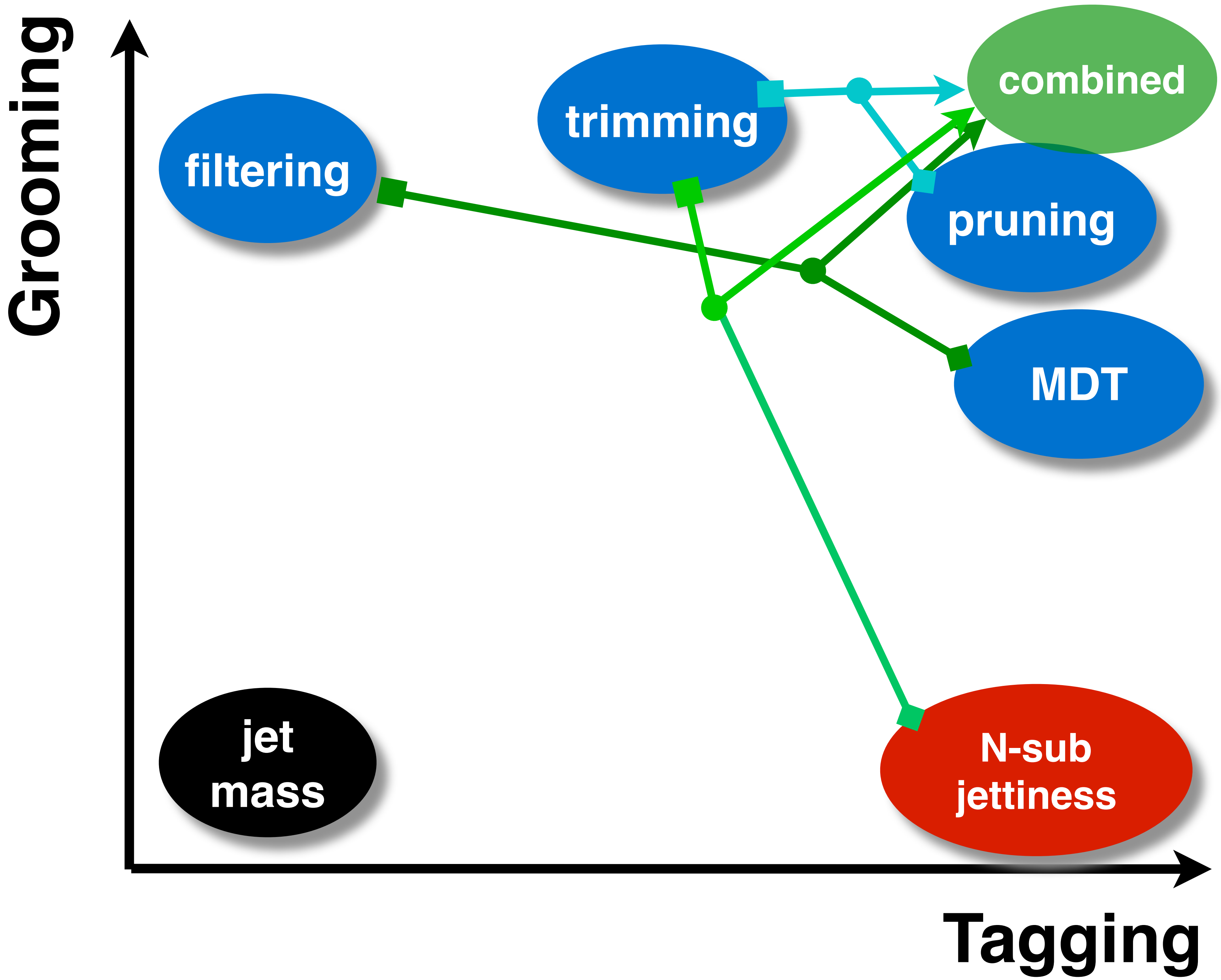


different (2-body) substructure tools

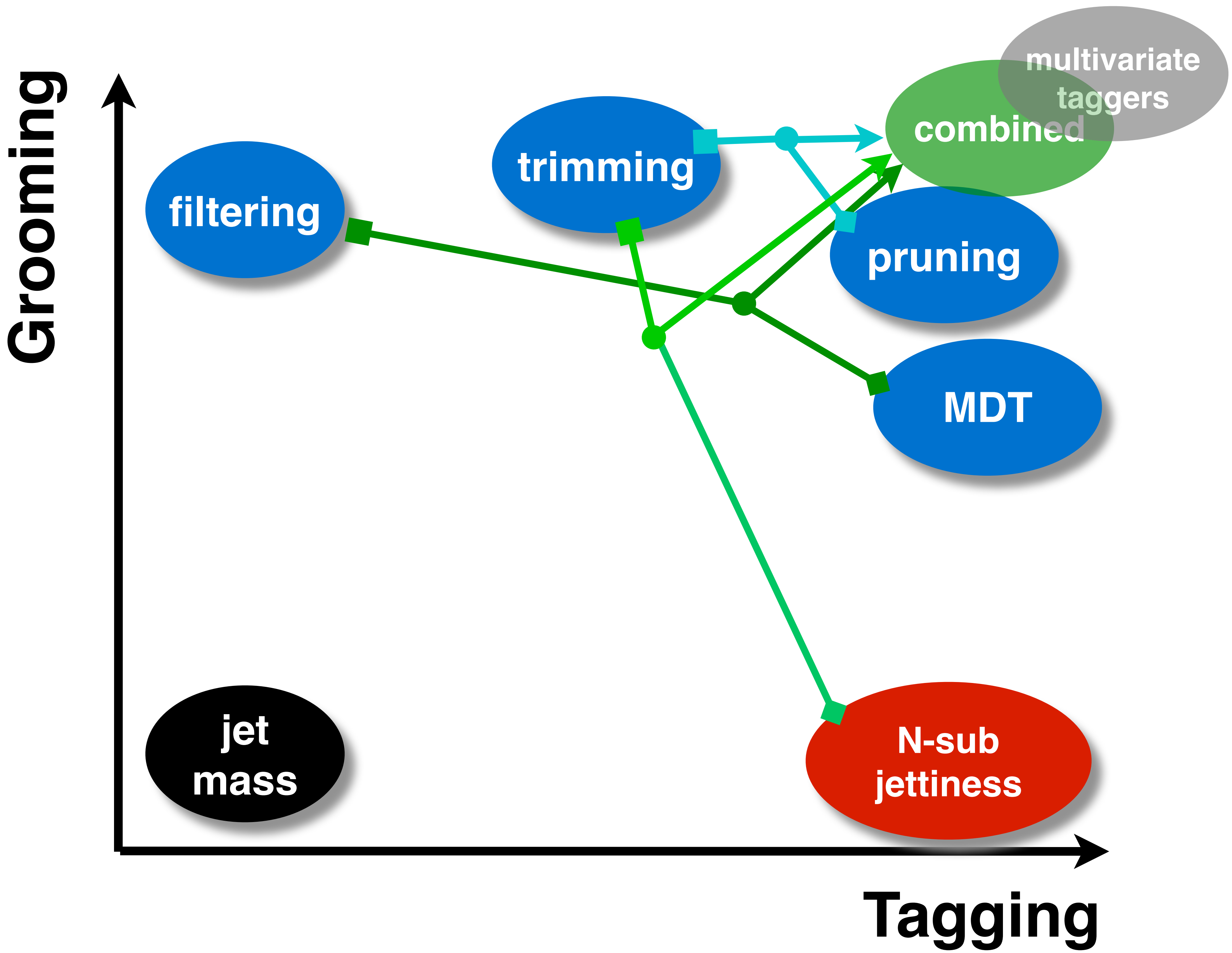
Detailed relative positions depend on physics context
(and are possibly contentious!)



Detailed relative positions depend on physics context
(and are possibly contentious!)

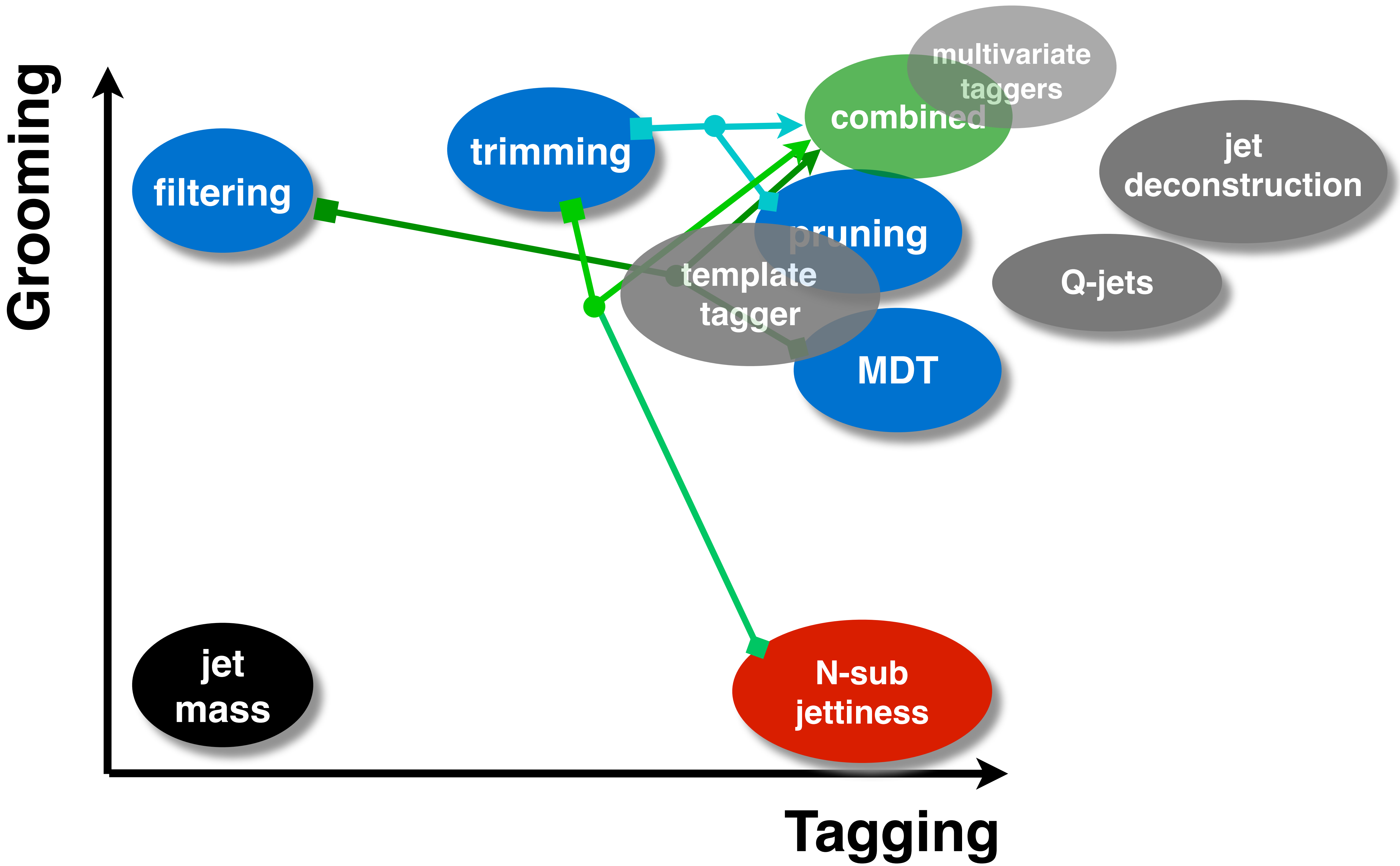


Detailed relative positions depend on physics context
(and are possibly contentious!)



different (2-body) substructure tools

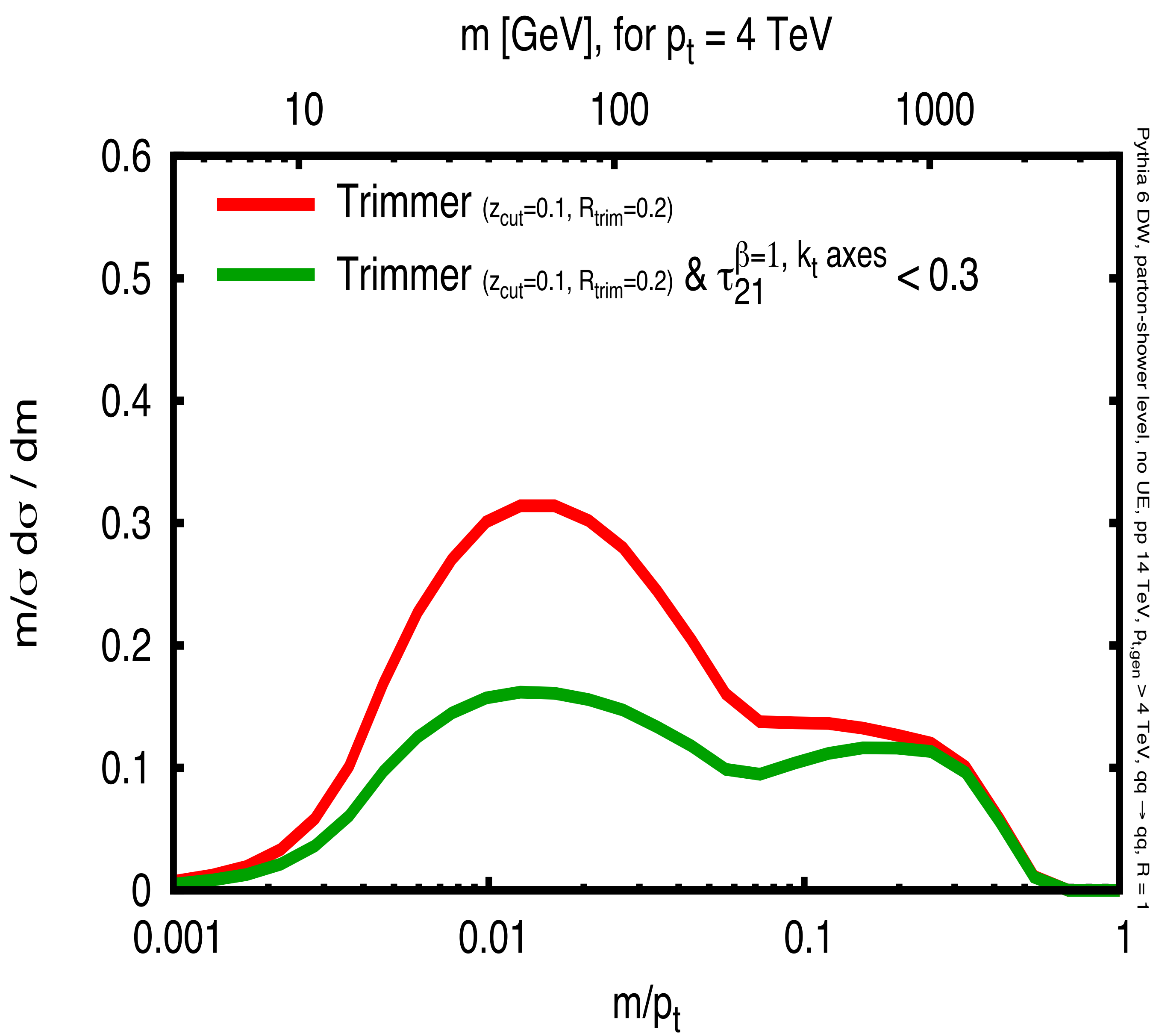
Detailed relative positions depend on physics context
(and are possibly contentious!)



What about cuts on shapes/radiation

E.g. cuts on N-subjettiness, tight mass drop, etc.?

- These cuts are nearly always for a jet whose mass is somehow groomed. All the structure from the grooming persists.
- So tagging & shape must probably be calculated together



Summary

Use of jets beyond the *"jet=parton"* idea is with us today.

That puts a responsibility on theorists to start understanding jet substructure beyond simply running Monte Carlos.

It seems that's feasible, with the potential also to guide development of more powerful and more robust jet tools.

Hopefully, this will help reliably stretch the boundaries of what LHC can do in its searches and measurements!

EXTRAS

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-1.001 (released 23 February 2013), which contains these contributions. Changes relative to earlier versions are briefly described in the NEWS file.

After downloading and unpacking, enter the fjcontrib-1.001/ directory and then run

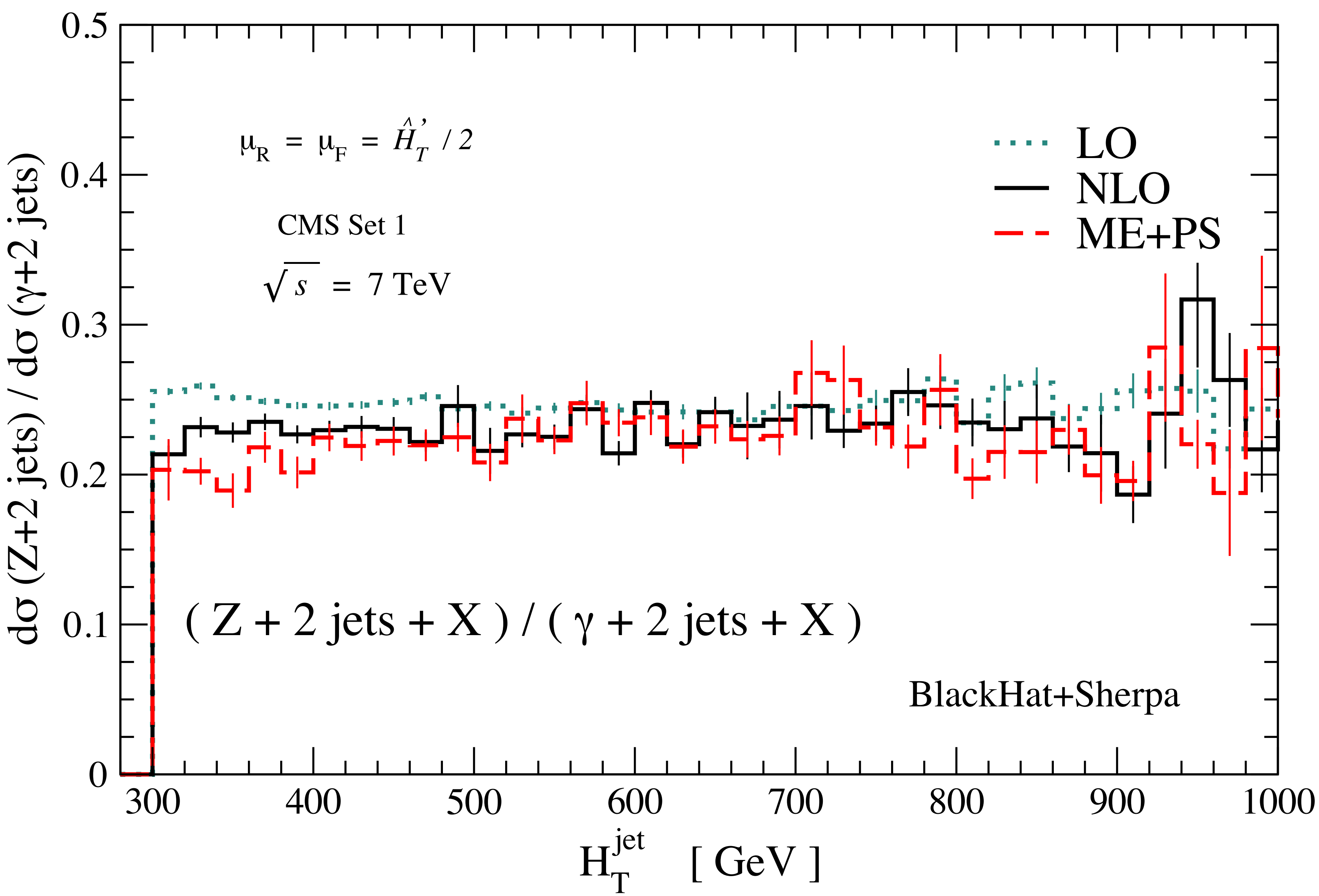
```
./configure [--fastjet-config=FILE] [--prefix=...] [...]
make
make check          # optional
make install
```

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.

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 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

Bottom line on “understanding”

- Taggers may be quite simple to write, but potentially involved to understand.
- 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 – further investigation warranted...

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)

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!

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

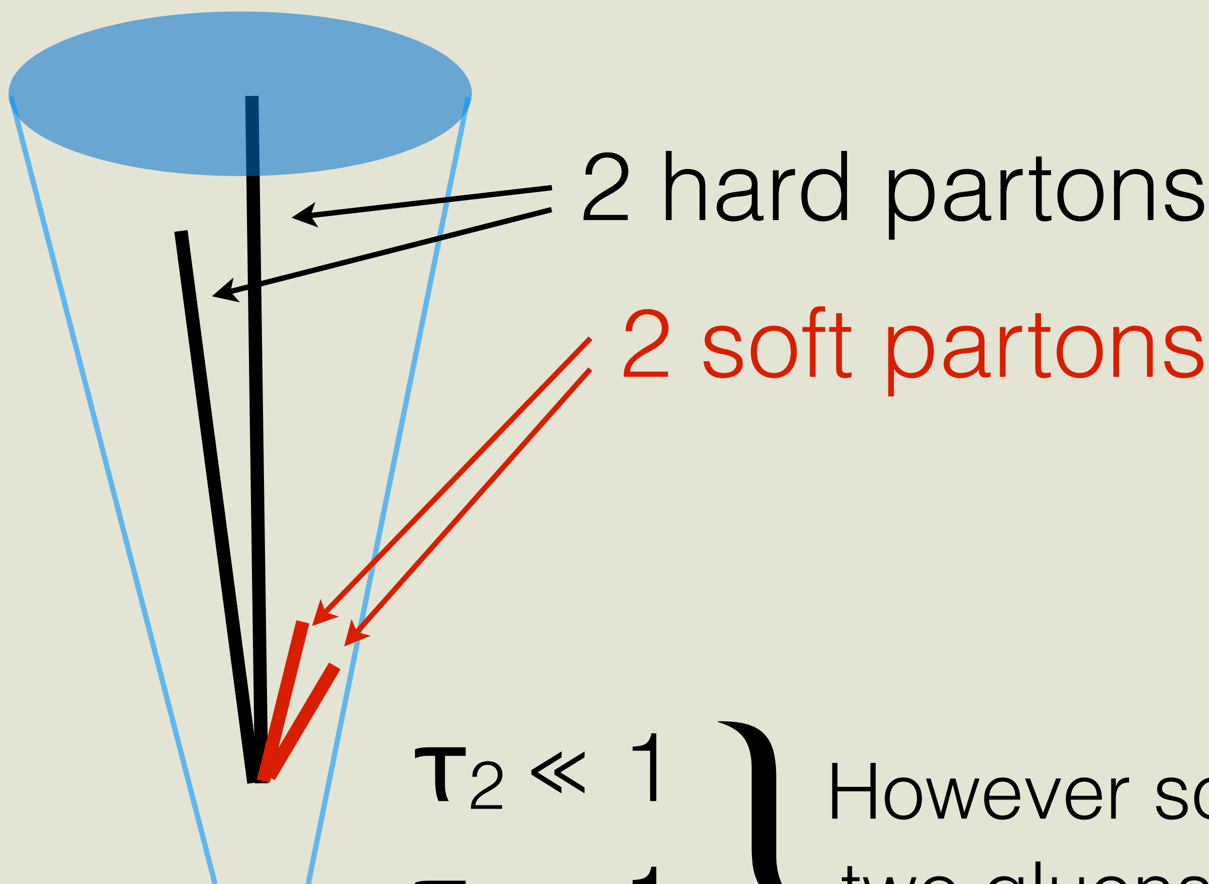
CMS's pruning followed by a mass-drop cut:

see blackboard!

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

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

Pileup in the boosted regime

Pronged top taggers

Some have pileup-***reduction*** built in (HEPTopTagger, Template), essentially by using small (R~0.2–0.3) sub-cones, sometimes dynamically adjusted to the top 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.

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

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

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

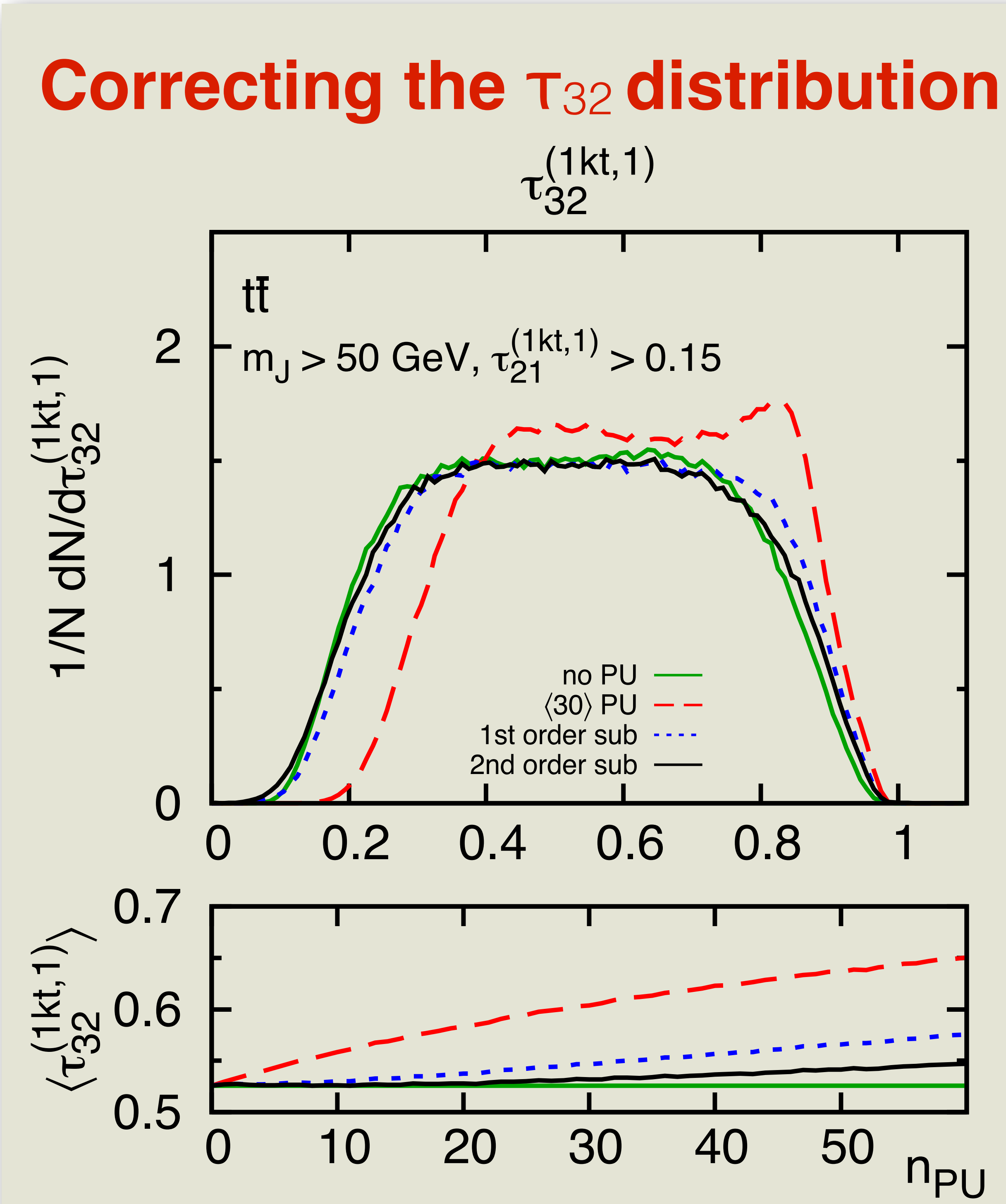
Ghost area

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

Subtracted shape

pileup density

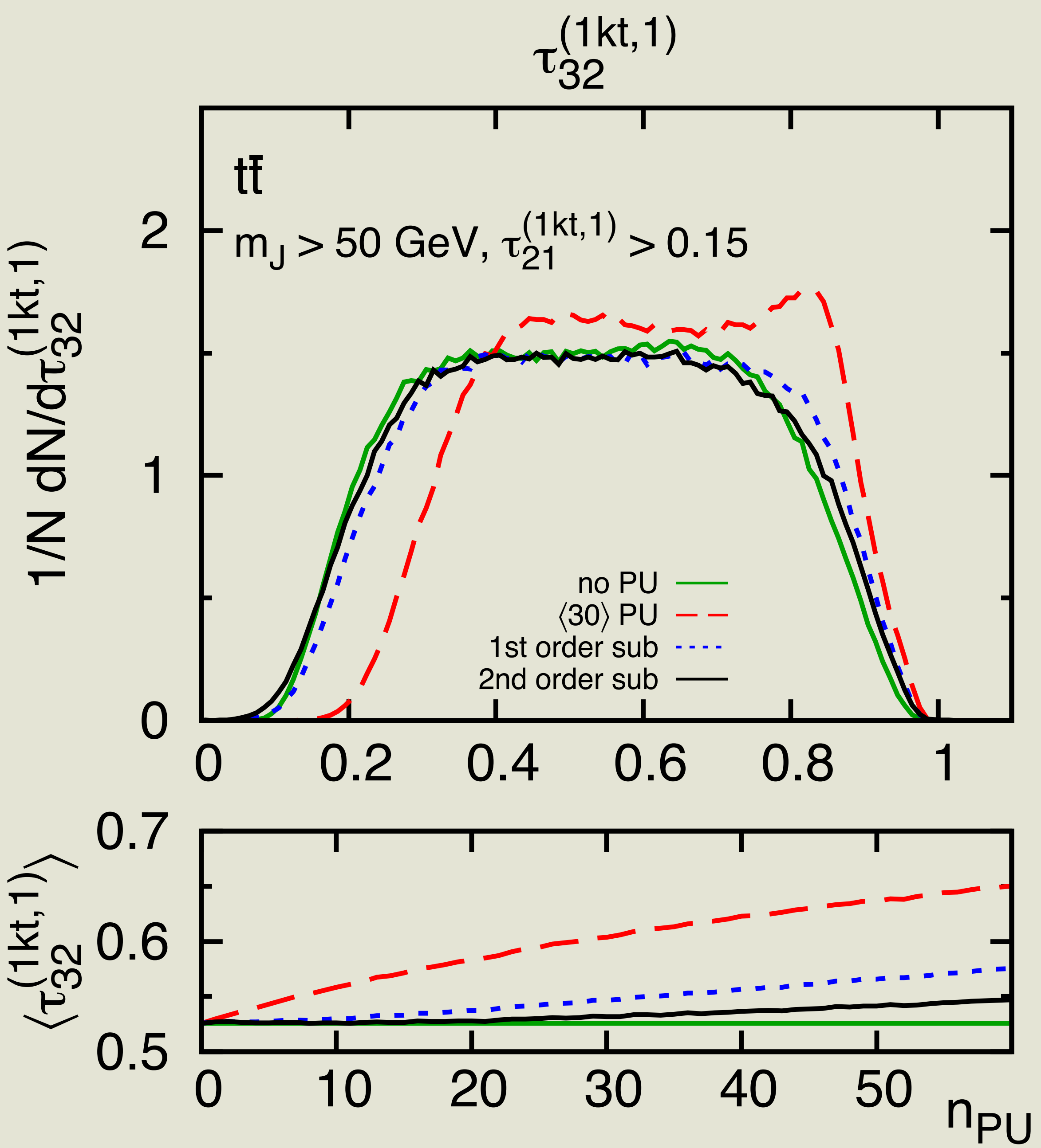
Practical test: τ_{32} and top tagging



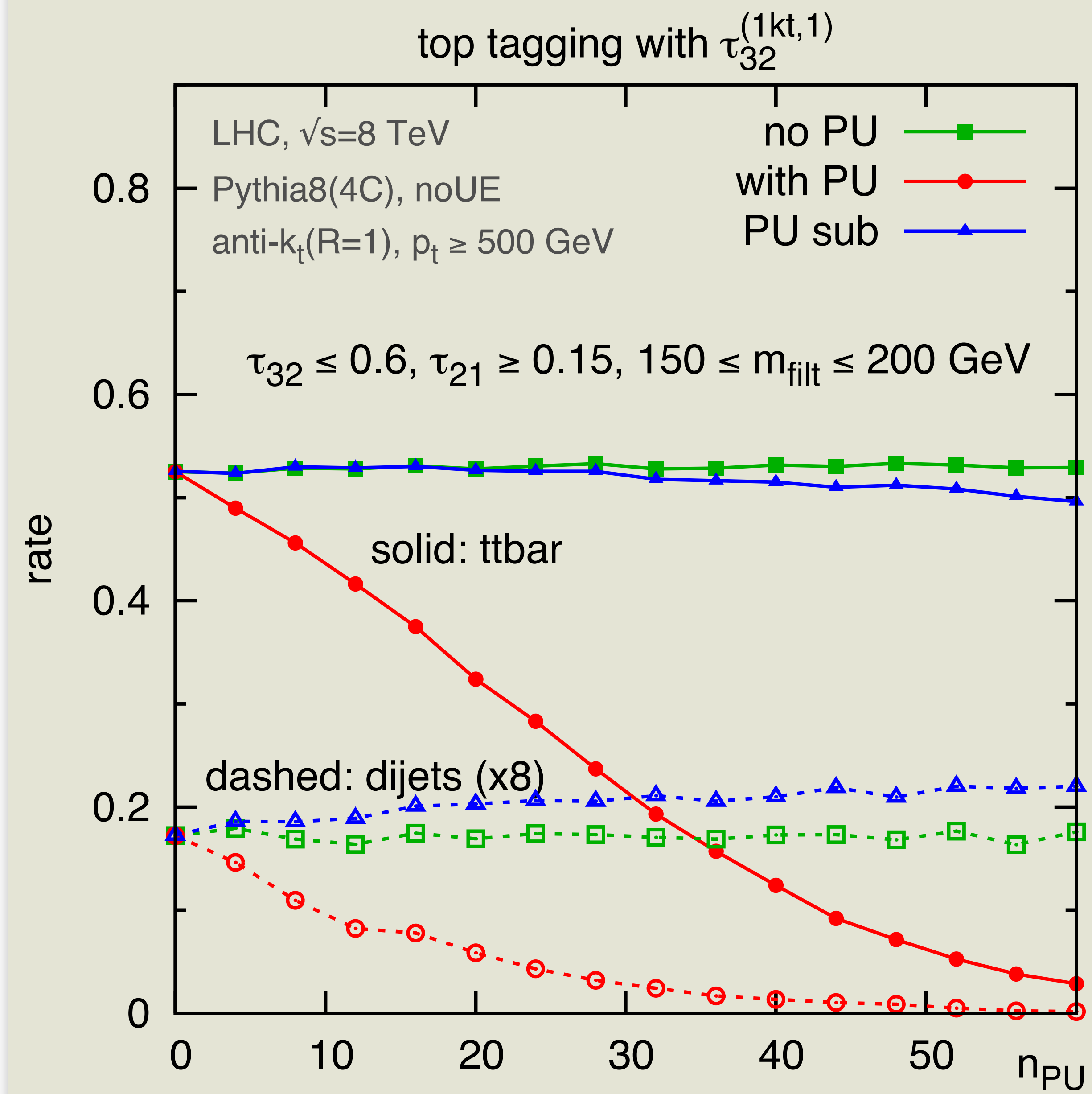
Green: no PU **Red: with PU** **Blue/Black: subtracted**

Practical test: τ_{32} and top tagging

Correcting the τ_{32} distribution



Tagging efficiency



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