Towards an understanding of jet substructure

Gavin Salam (CERN)
with Mrinal Dasgupta, Alessandro Fregoso & Simone Marzani

Boost 2013
Flagstaff, Arizona
August 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.

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?

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

This talk → analytical understanding. Why?

Better Insight

Can guide taggers' use in experimental analyses

It may help us design better taggers

This talk → 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

Scope of our study

What kind of events?

To fully understand "Boost" you want to study
all possible signal (W/Z/H/top/…) and QCD jets.

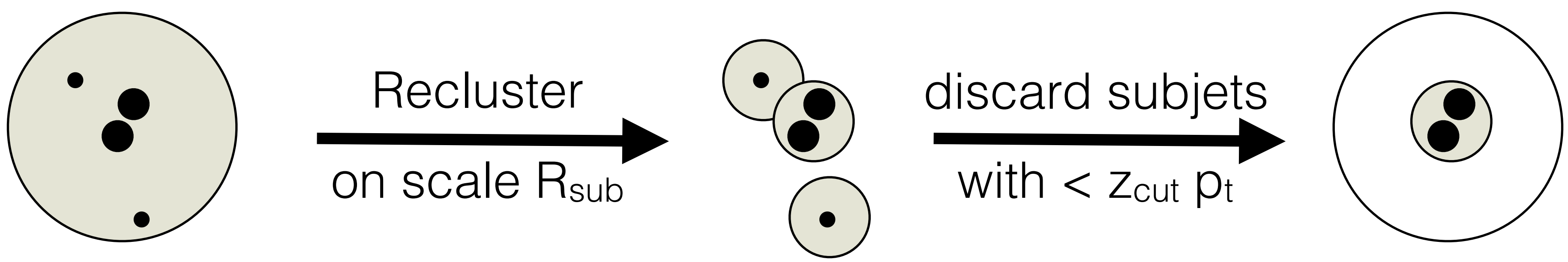
But you need to start somewhere.
We chose the QCD jets because:

(a) they have the richest structure.

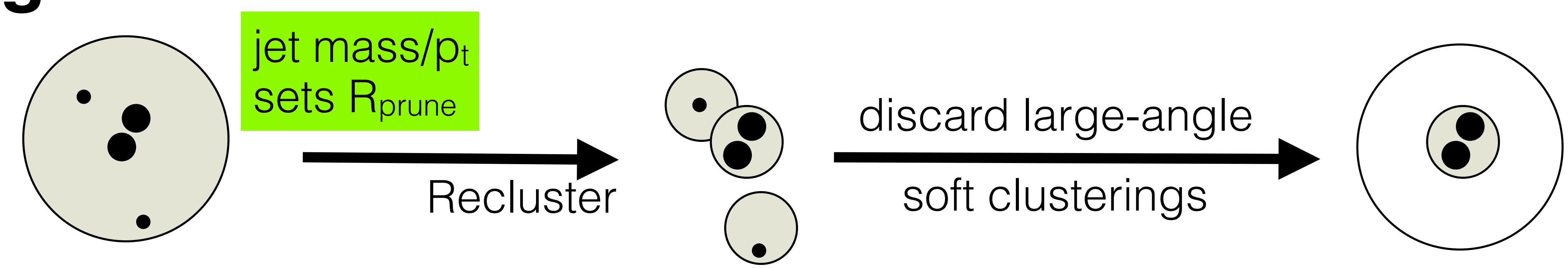
(b) once you know understand the QCD jets,
the route for understanding signal jets becomes clear too.

Cannot possibly study all tools
These 3 are widely used

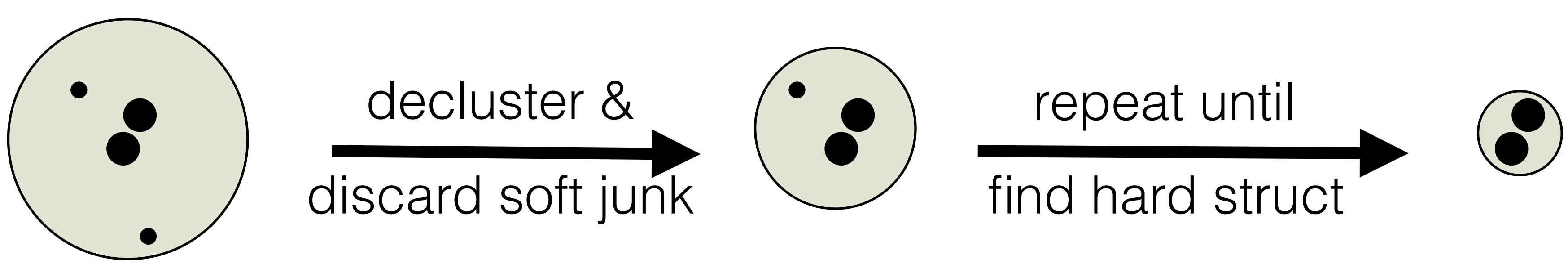
Trimming



Pruning



Mass-drop tagger (MDT, aka BDRS)



First use MC to get a panorama

For phenomenology

Jet mass: *m*

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

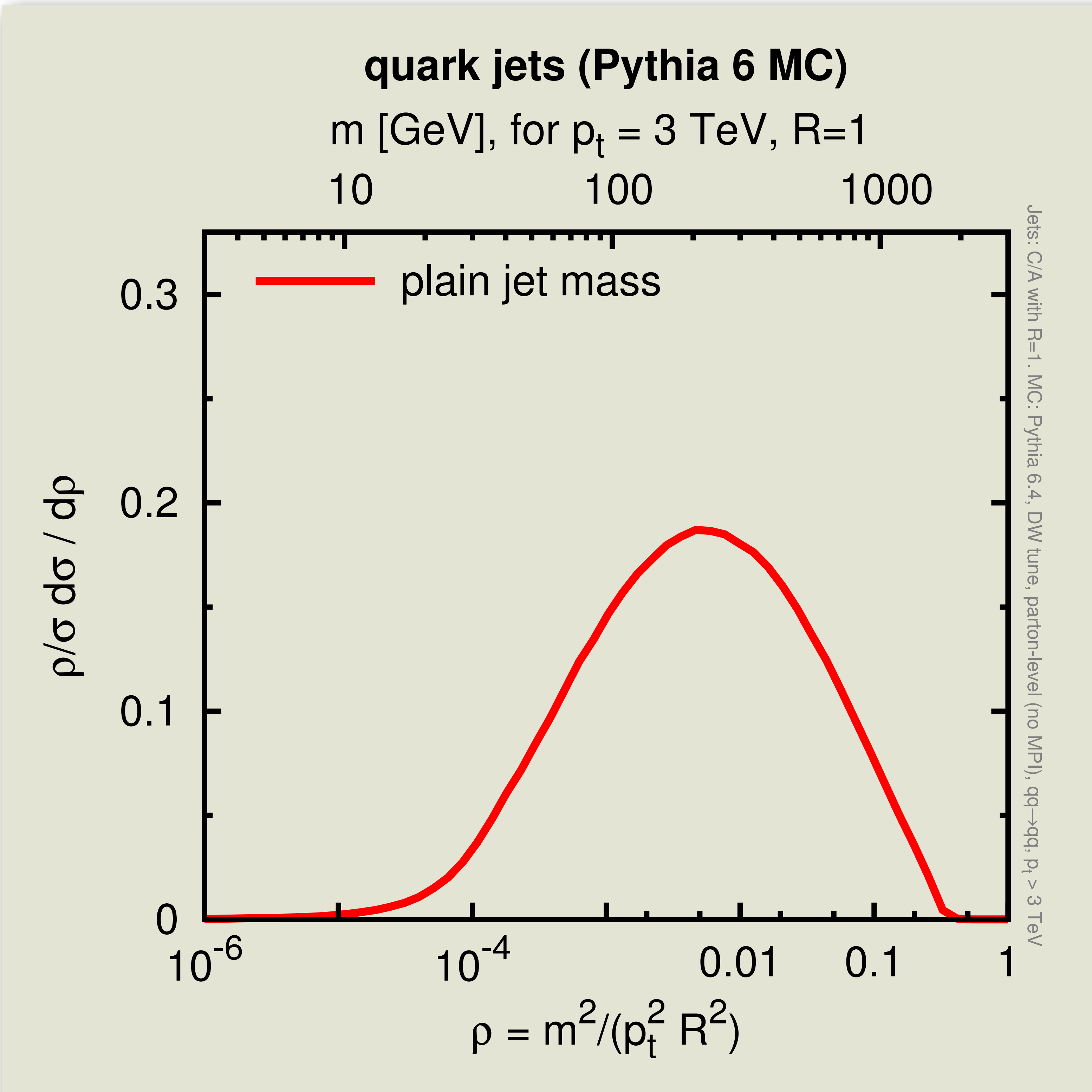
For QCD calculations

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

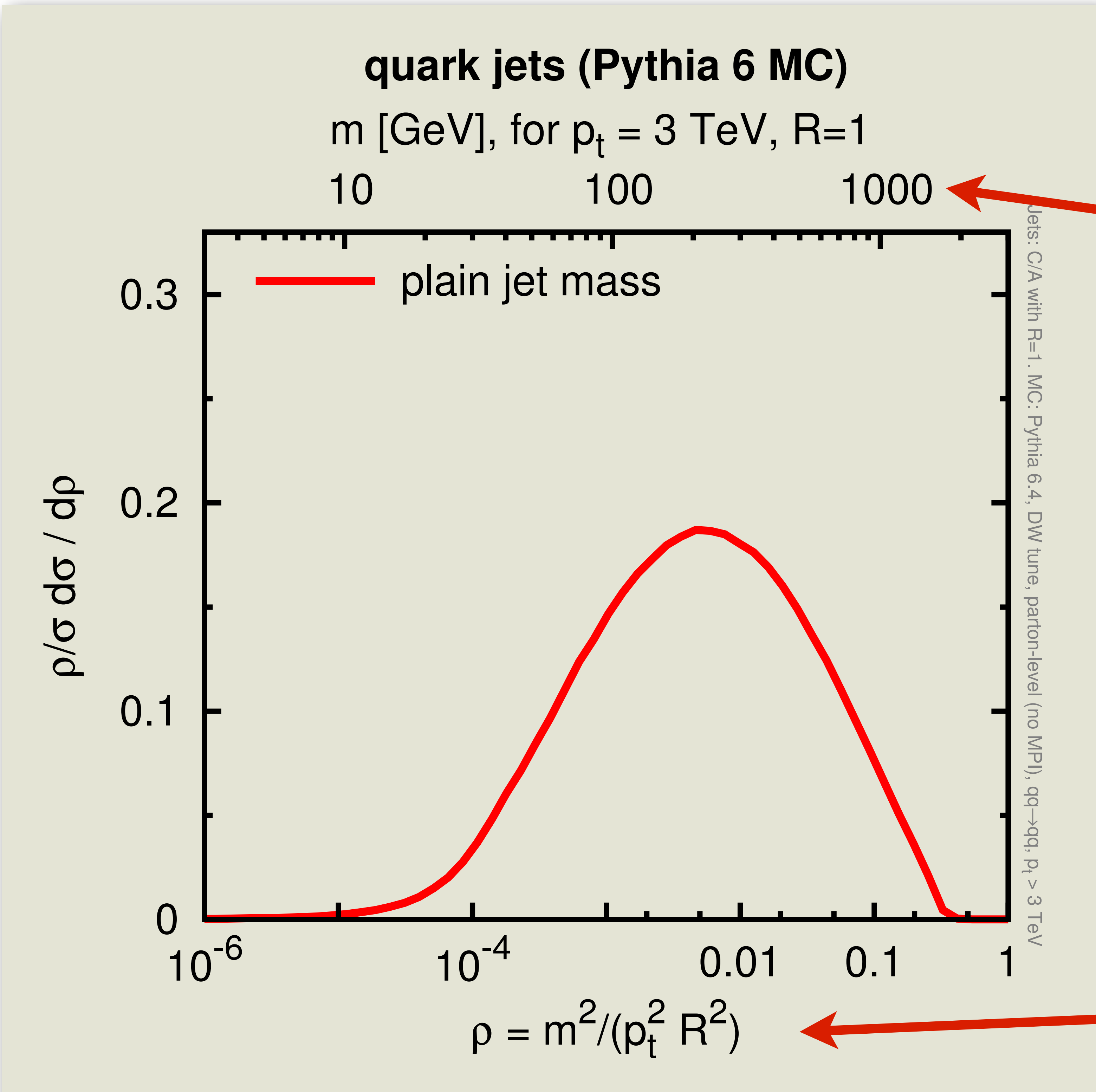
[R is jet opening angle – or radius]

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

Start with "plain" jet mass



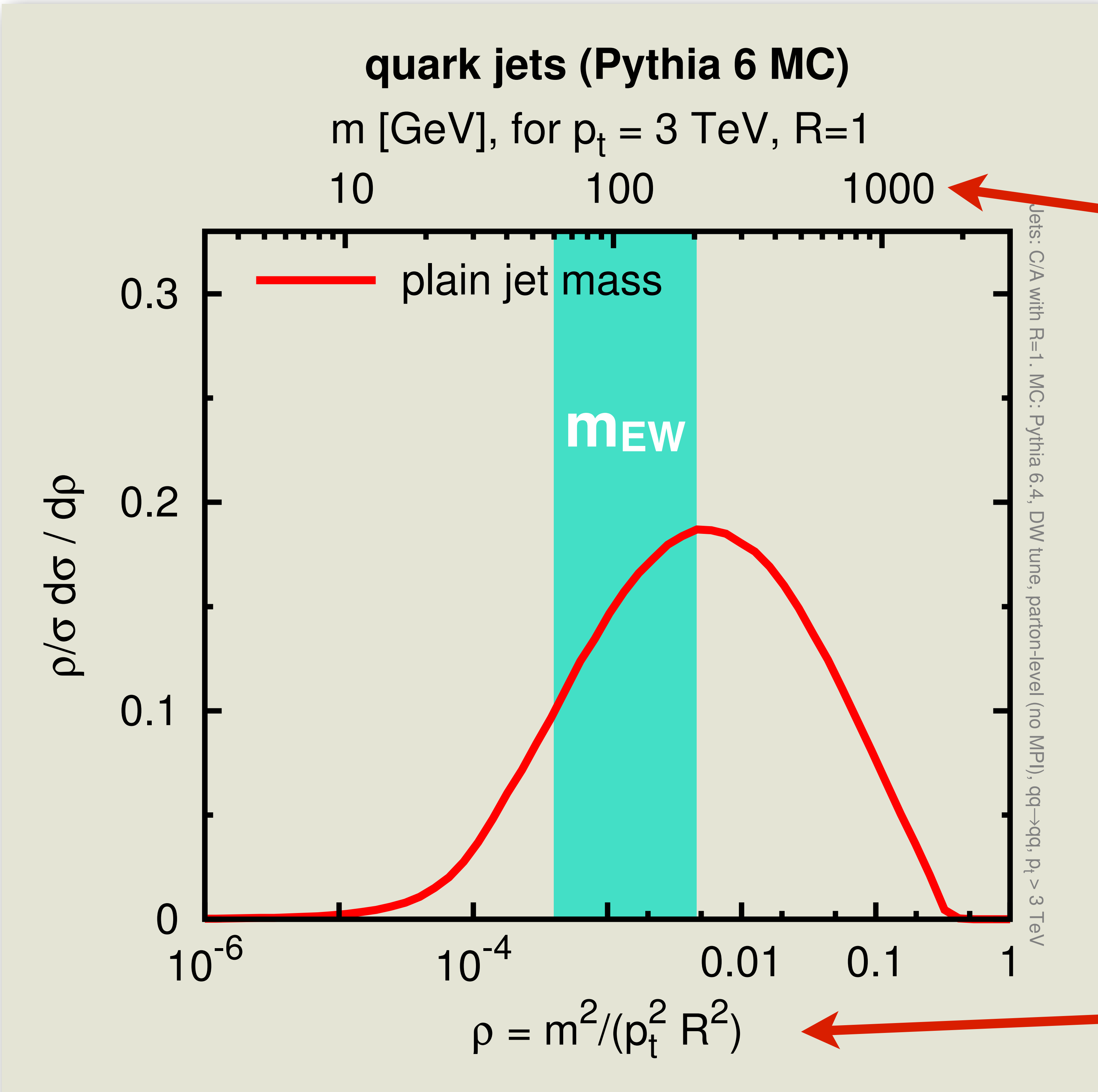
Start with "plain" jet mass



Physical mass for 3 TeV, R=1 jets

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

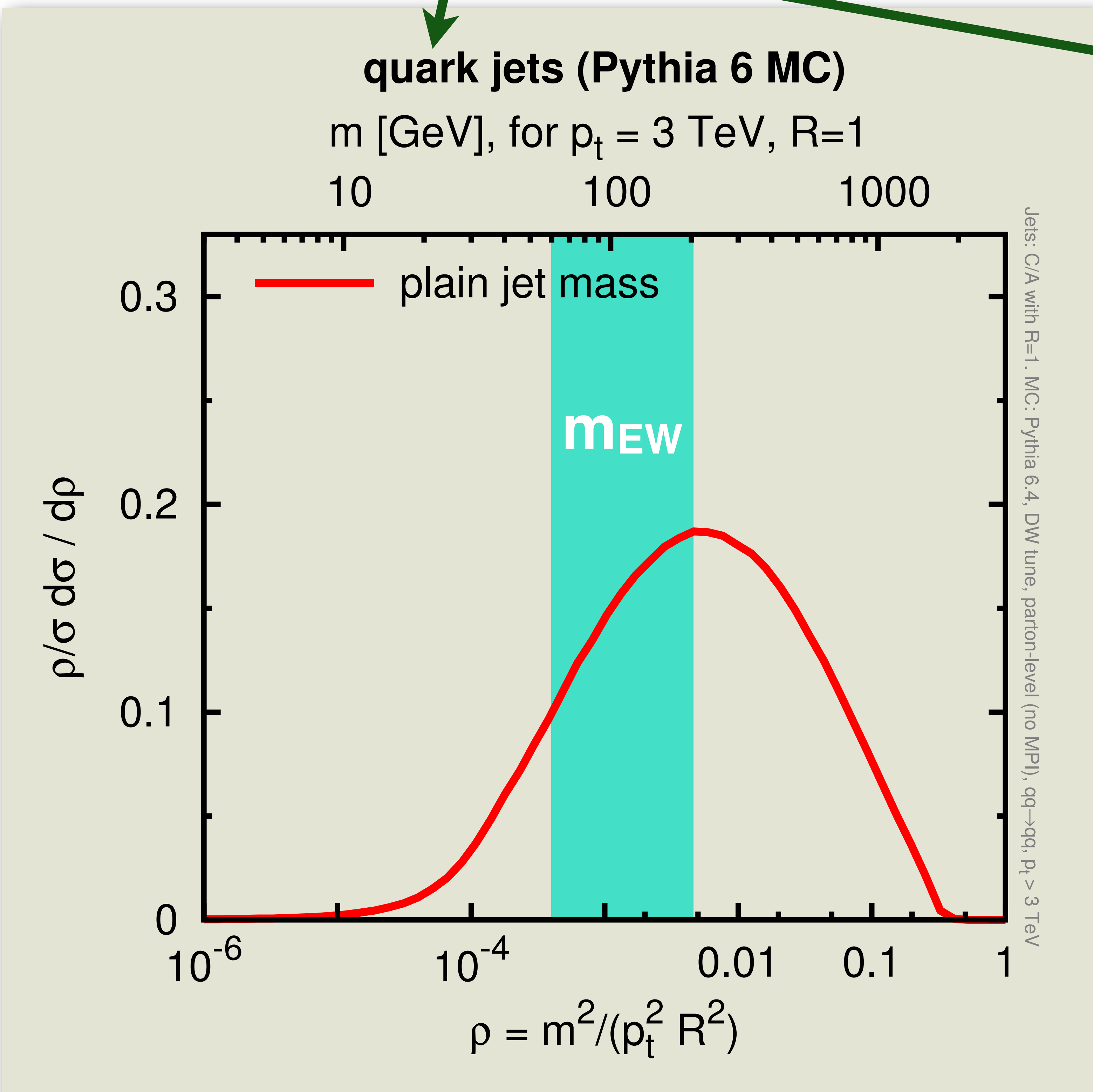
Start with "plain" jet mass



Physical mass for 3 TeV, R=1 jets

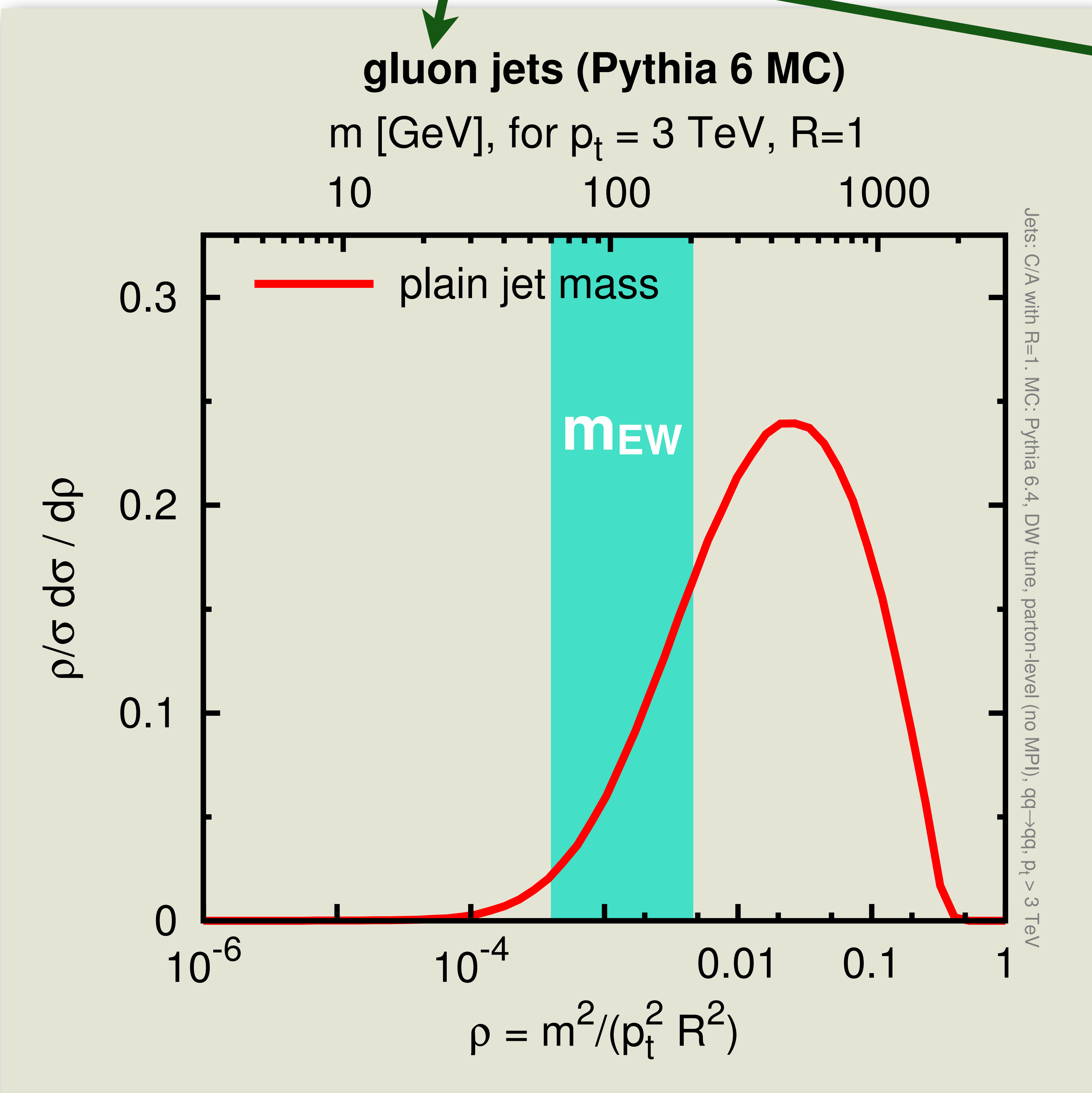
ρ ~ Rescaled mass²
(i.e. the QCD variable)

Start with “plain” jet mass



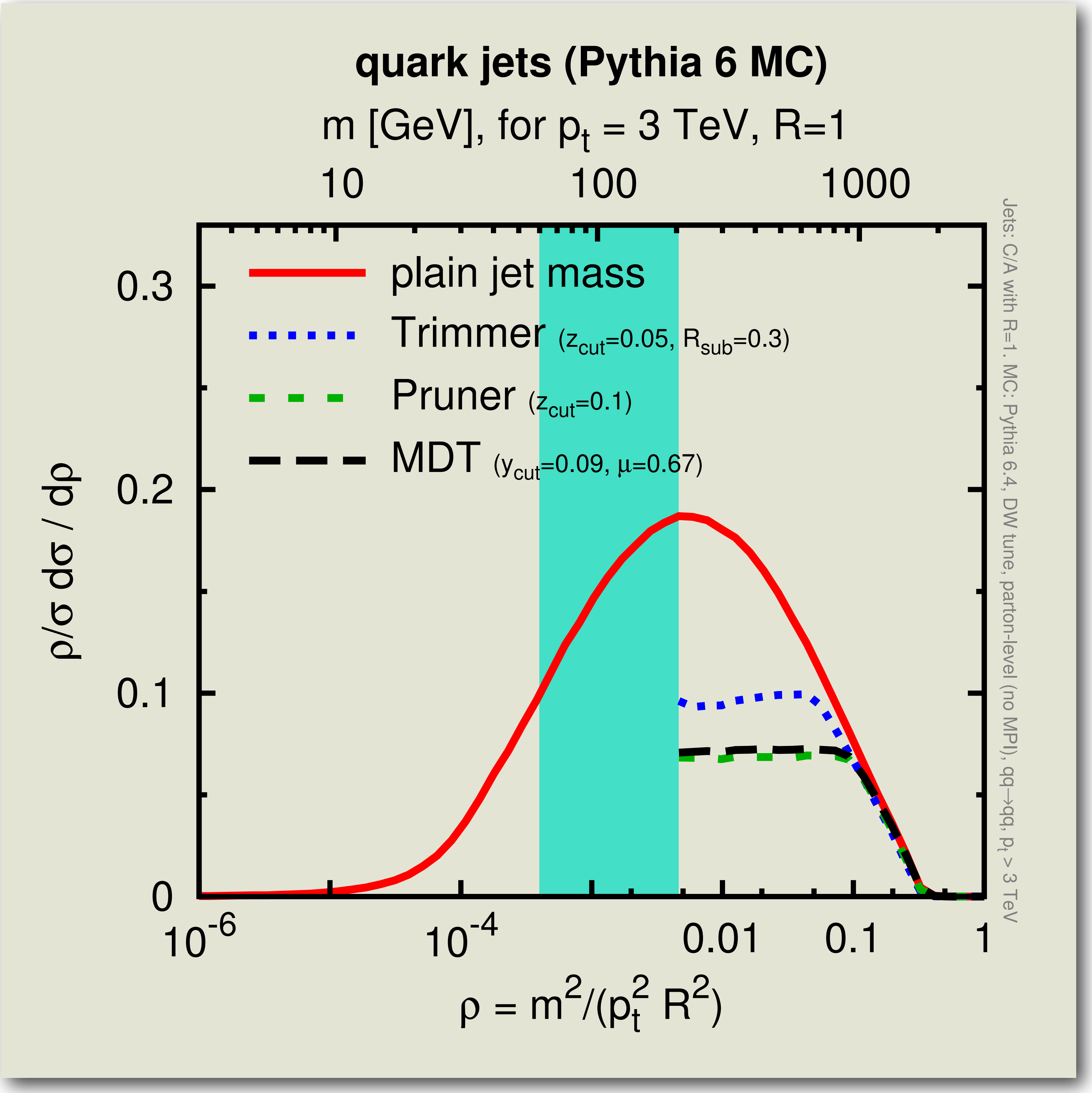
Quark and gluon jets are different, so we treat them separately

Start with "plain" jet mass



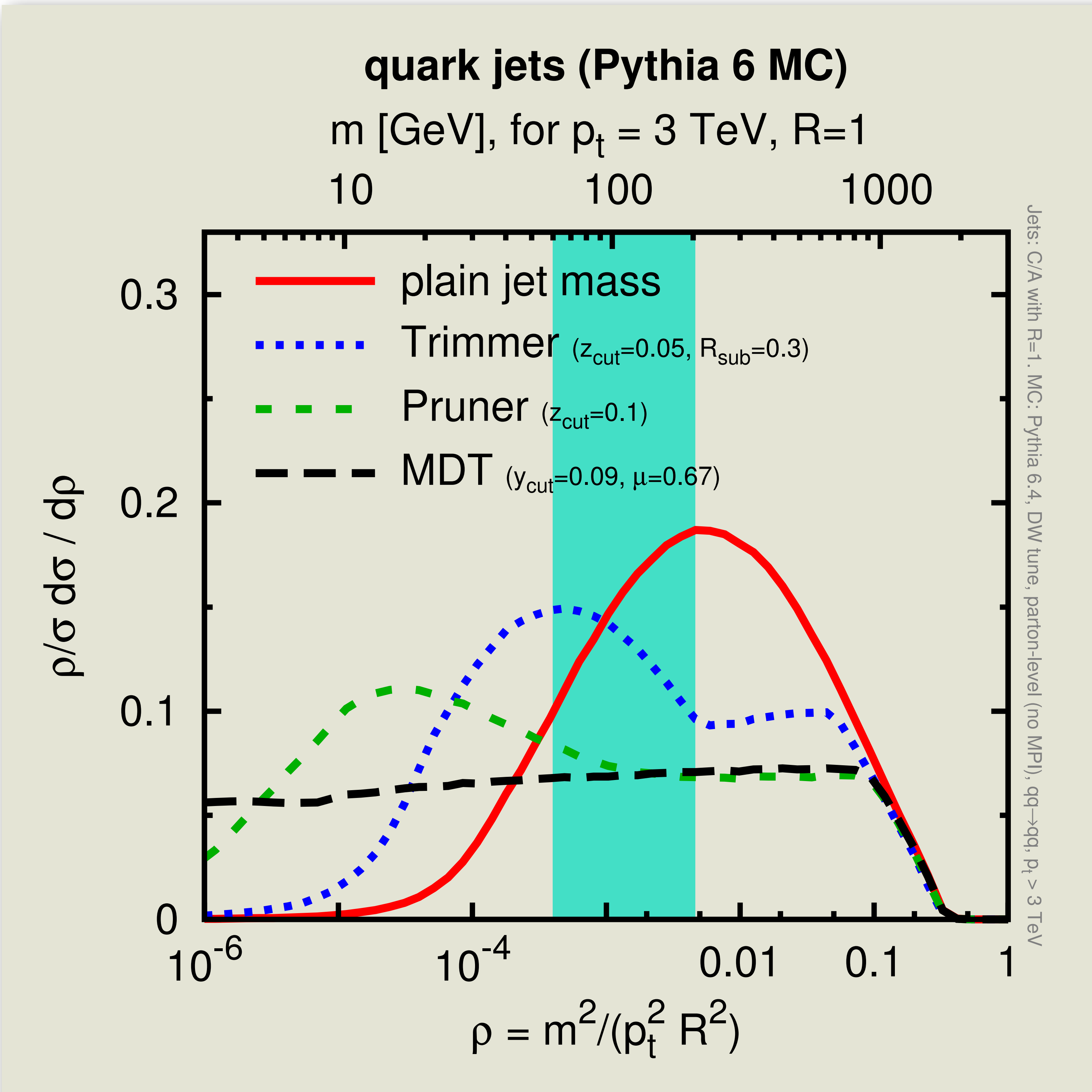
Quark and **gluon** jets are different, so we treat them separately

Now examine "taggers/groomers"



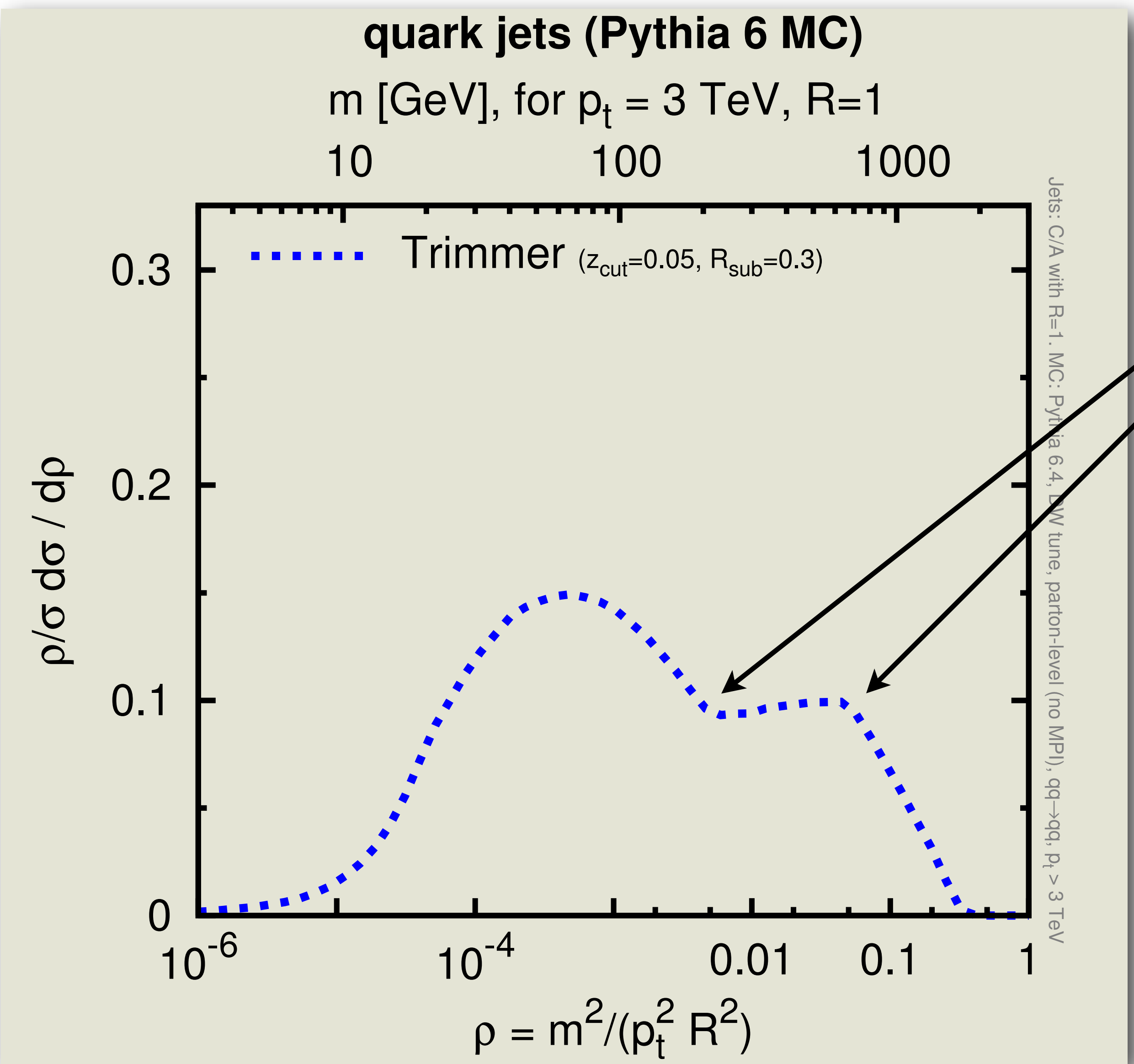
Different taggers can be quite similar

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

The parameters & approximations

Trimming

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

Our approximations

- $\rho \ll 1$
  logs of ρ get resummed
- pretend $R \ll 1$
- $z_{cut} \ll 1$,
  but ($\log z_{cut}$) not large

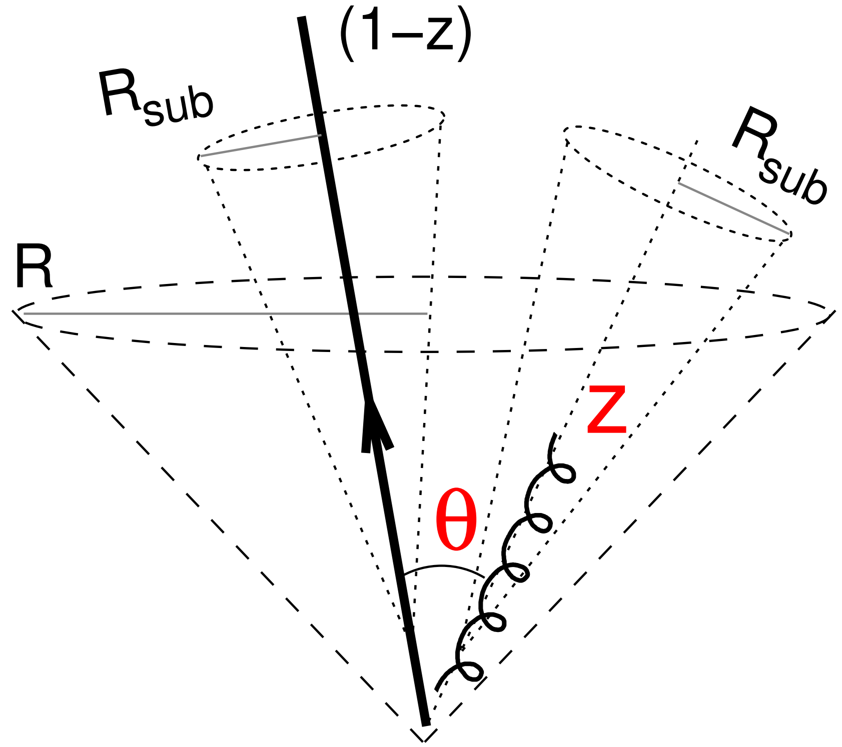
These approximations are not as "wild"as they might sound.

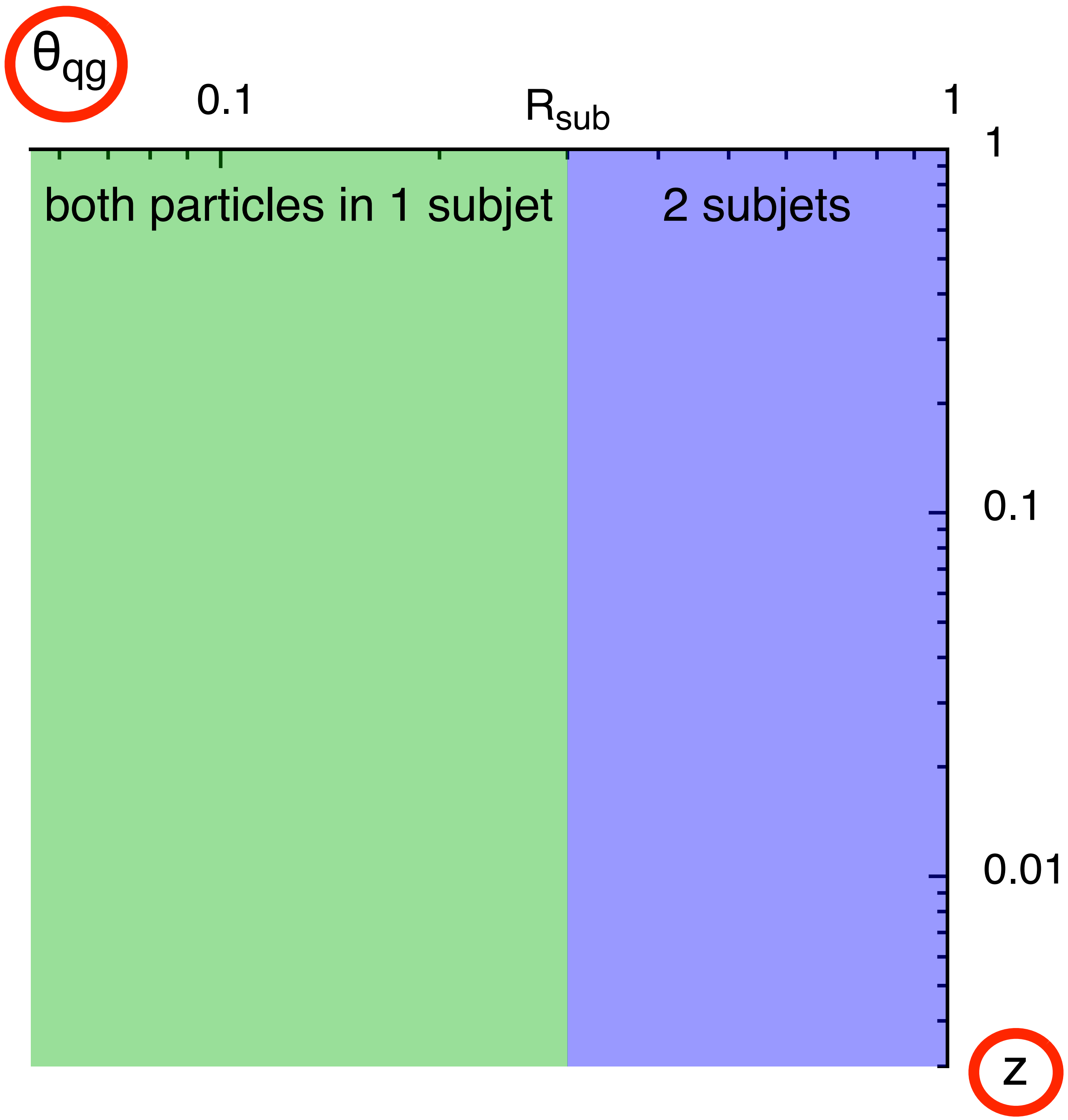
They can also be relaxed.

But our aim for now is to understand the taggers — we leave highest precision calculations till later.

Leading Order — 2-body kinematic plane

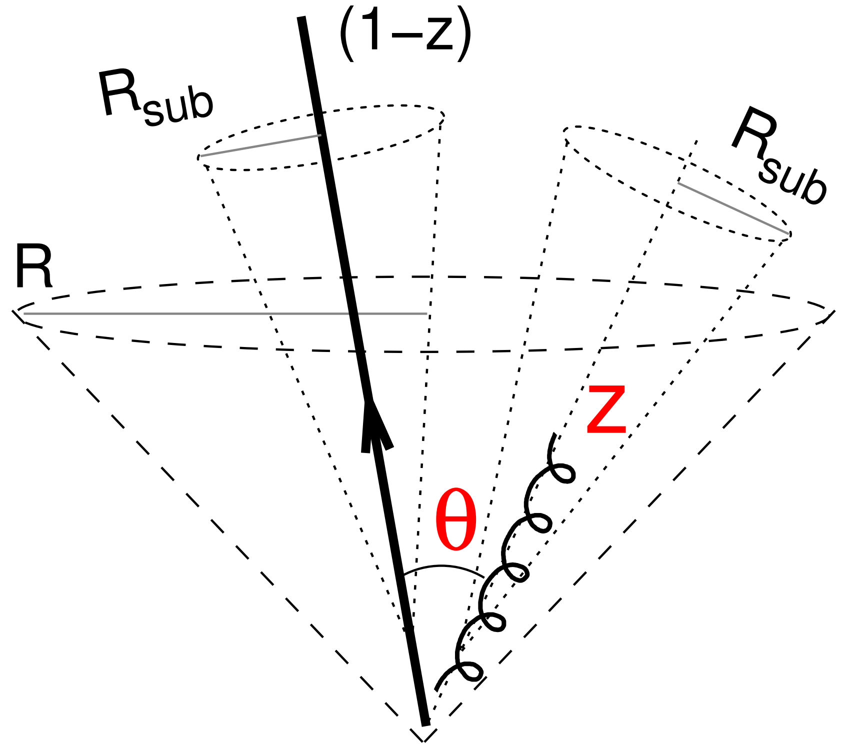
At $O(\alpha_s)$, a quark jet emits a gluon. We study this as a function of the gluon momentum fraction **z** and the quark-gluon opening angle **θ**

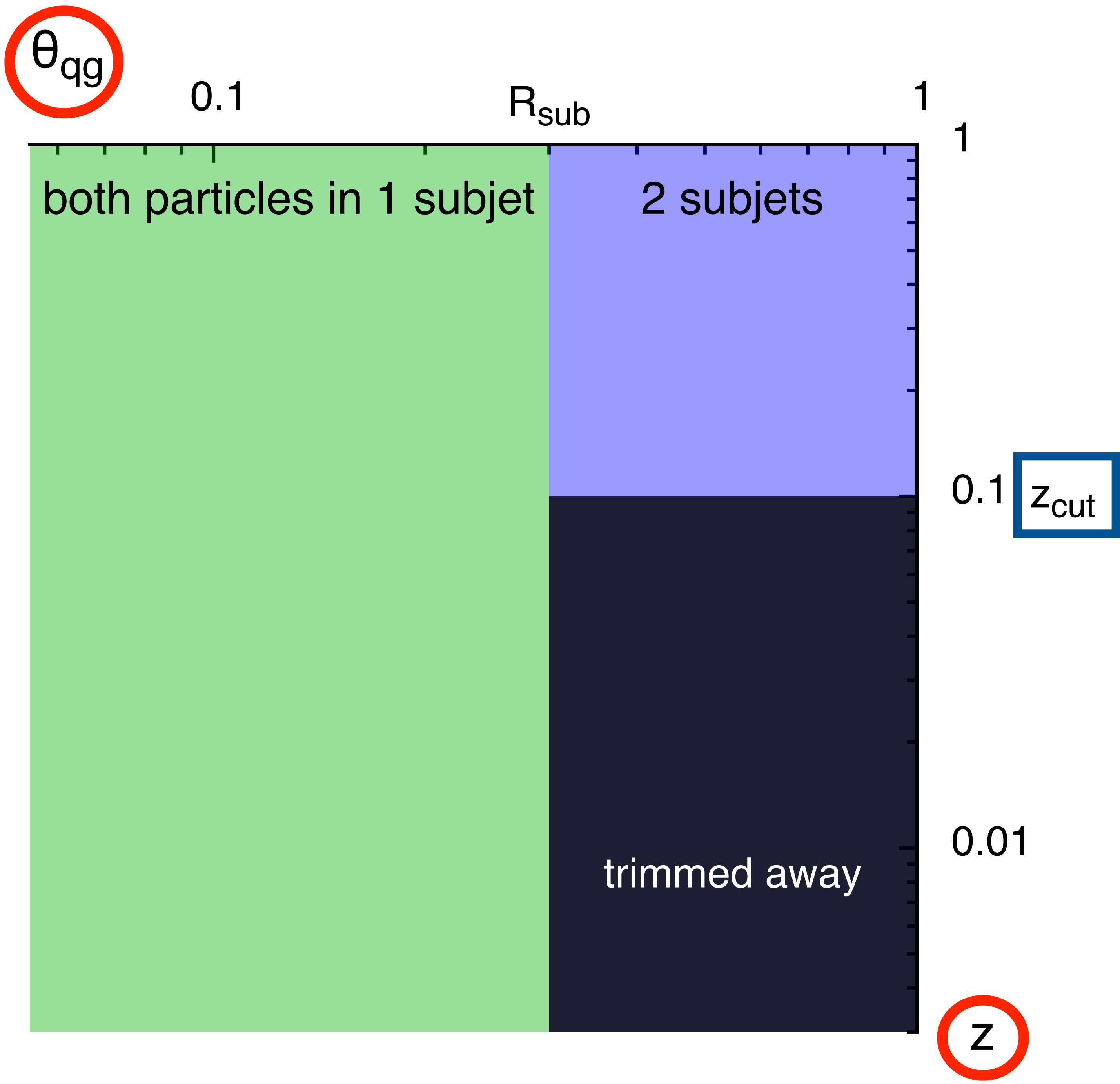




Leading Order — 2-body kinematic plane

At O(α_s), a quark jet emits a gluon. We study this as a function of the gluon momentum fraction **z** and the quark-gluon opening angle **θ**

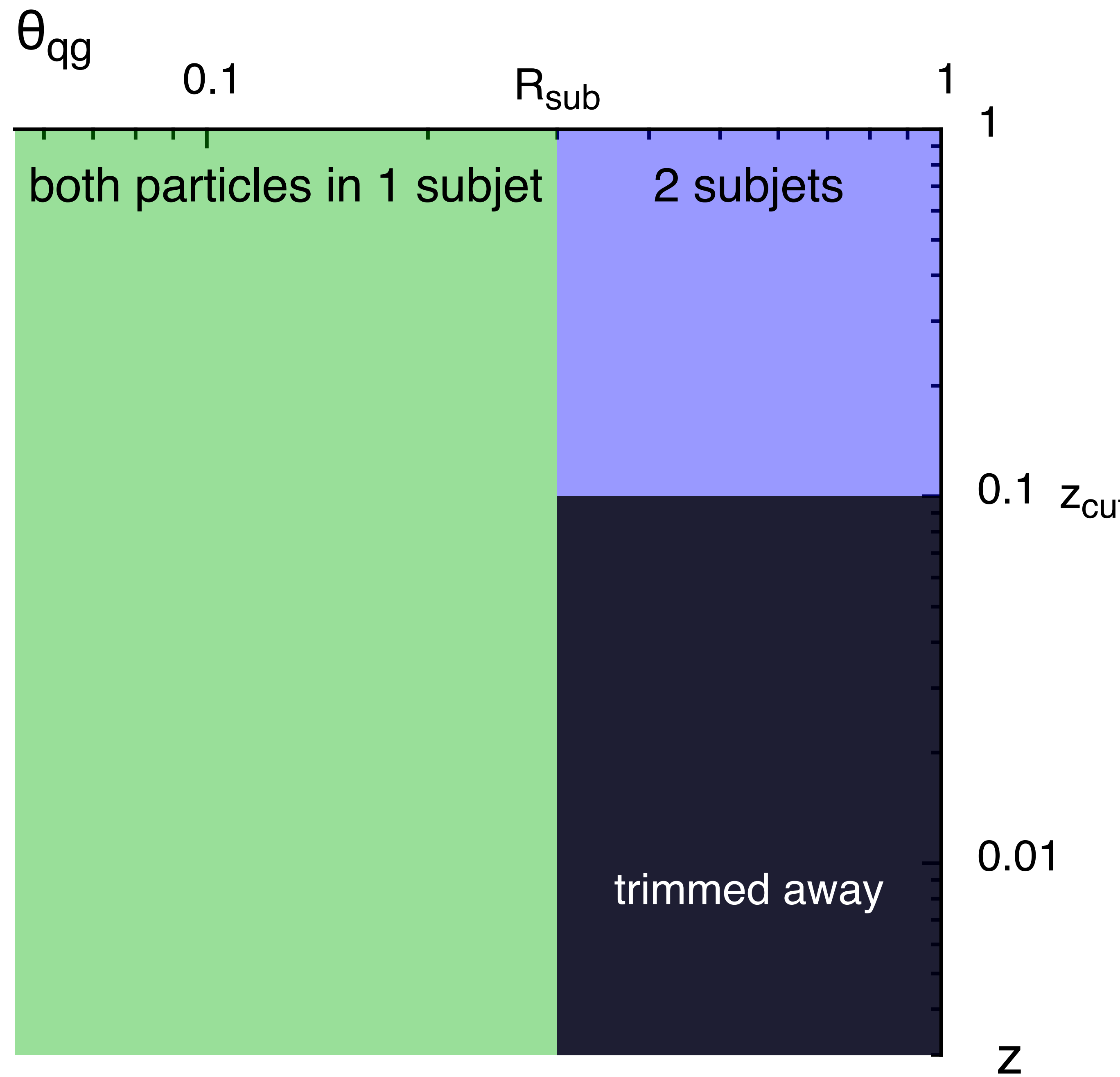




matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane



jet mass

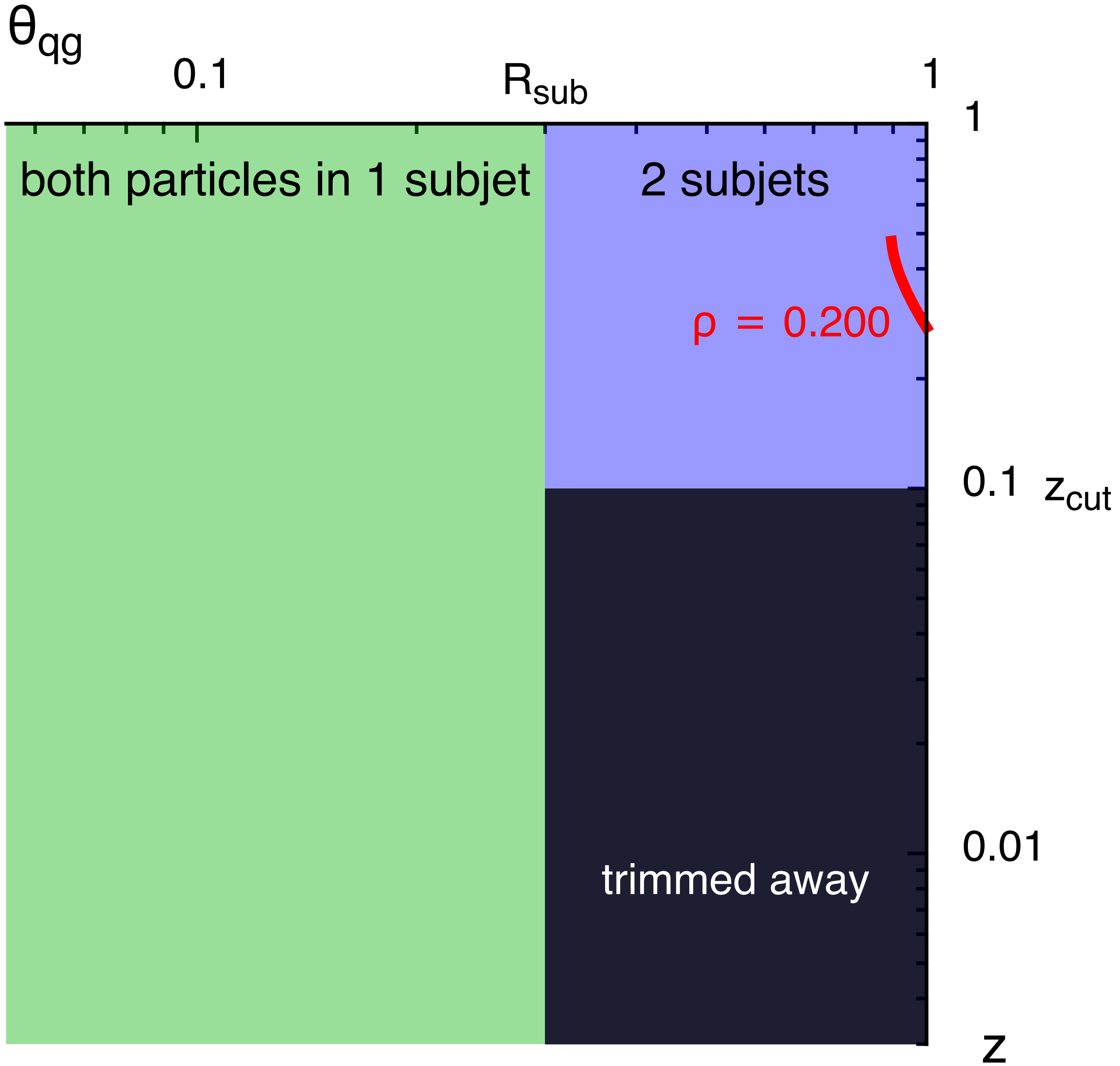
$$\rho = z(1-z)\theta^2$$

length of **fixed-ρ contour** gives LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane



jet mass

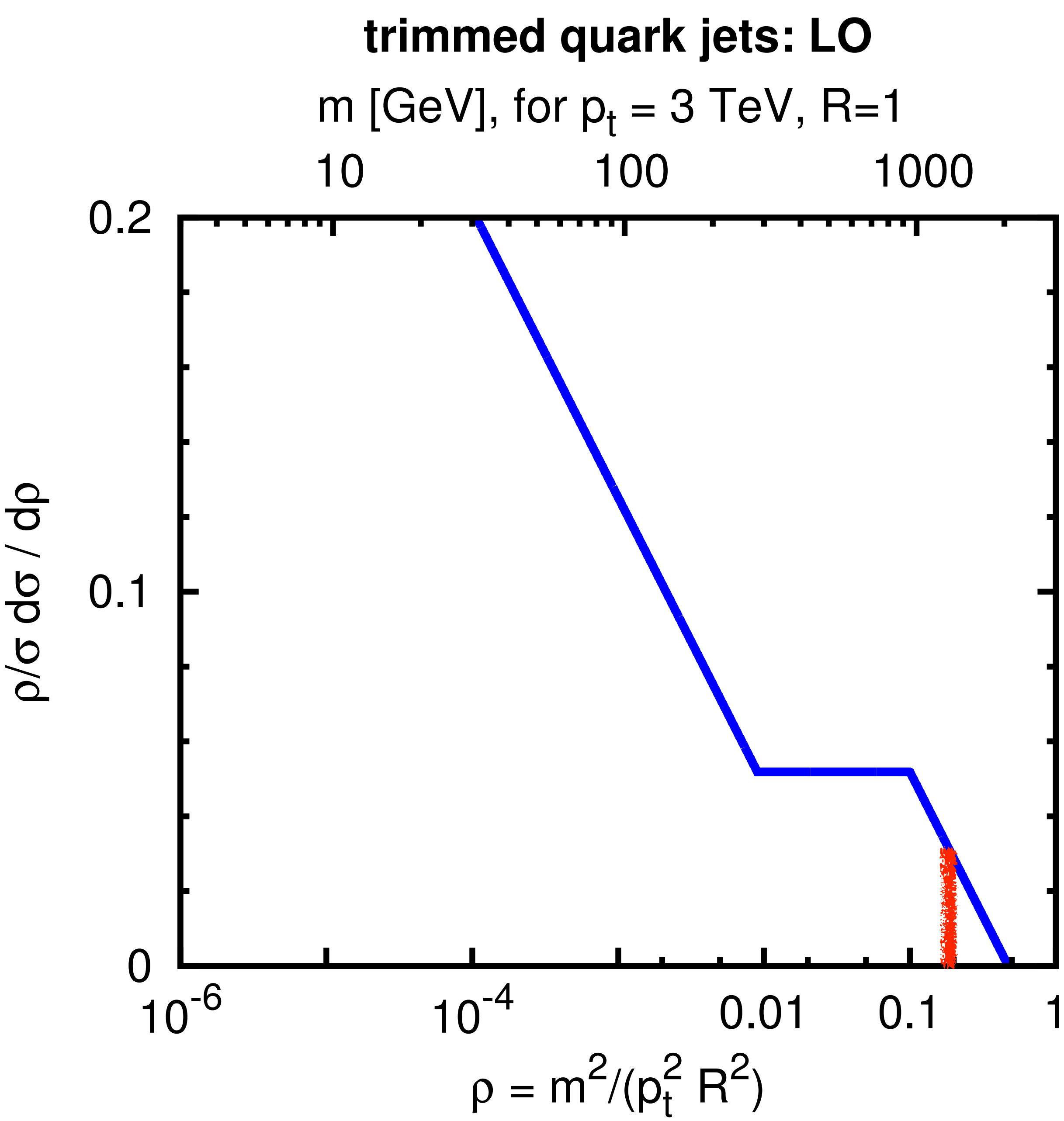
$$\rho = z(1-z)\theta^2$$

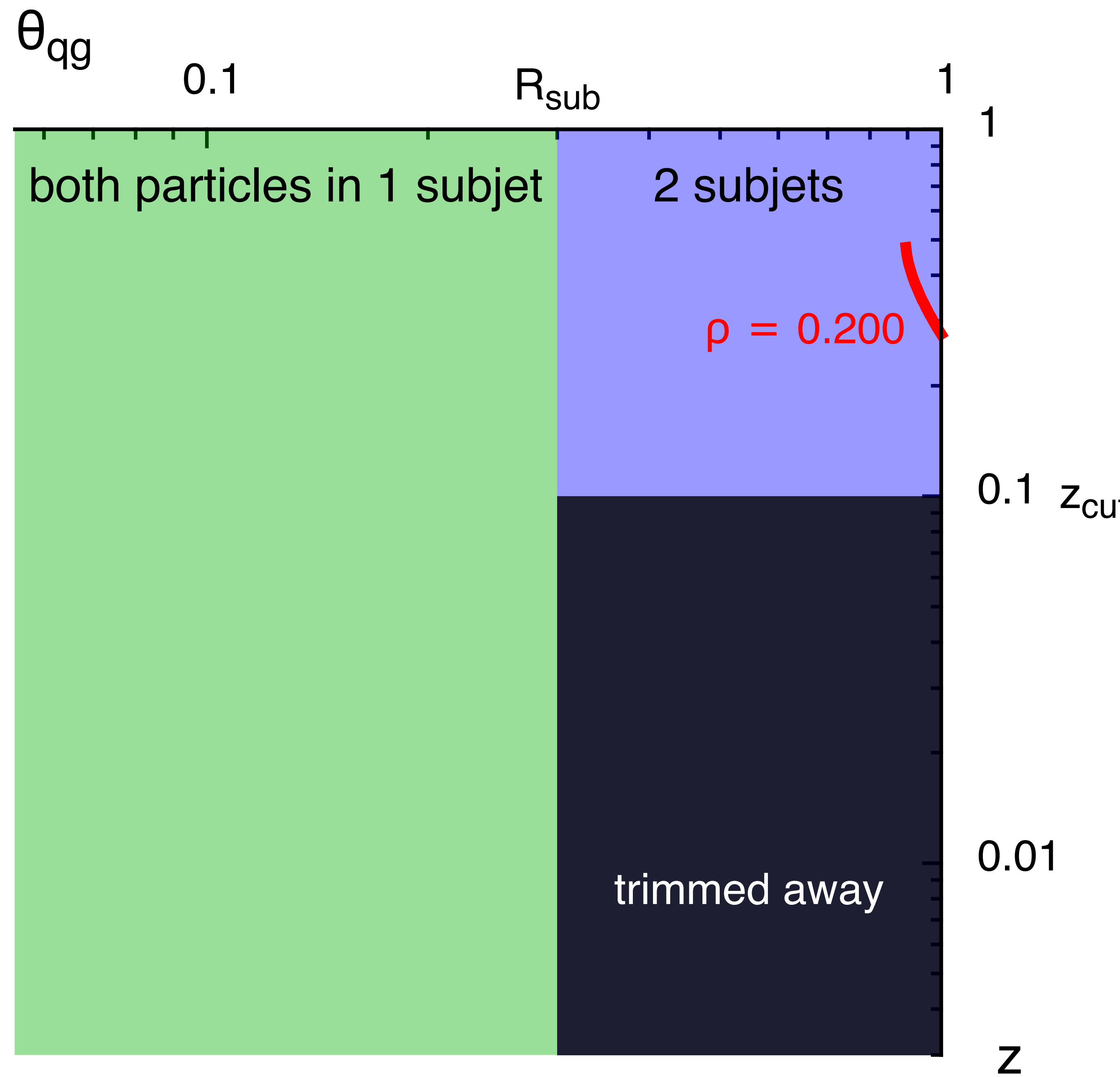
length of **fixed-ρ contour** gives LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

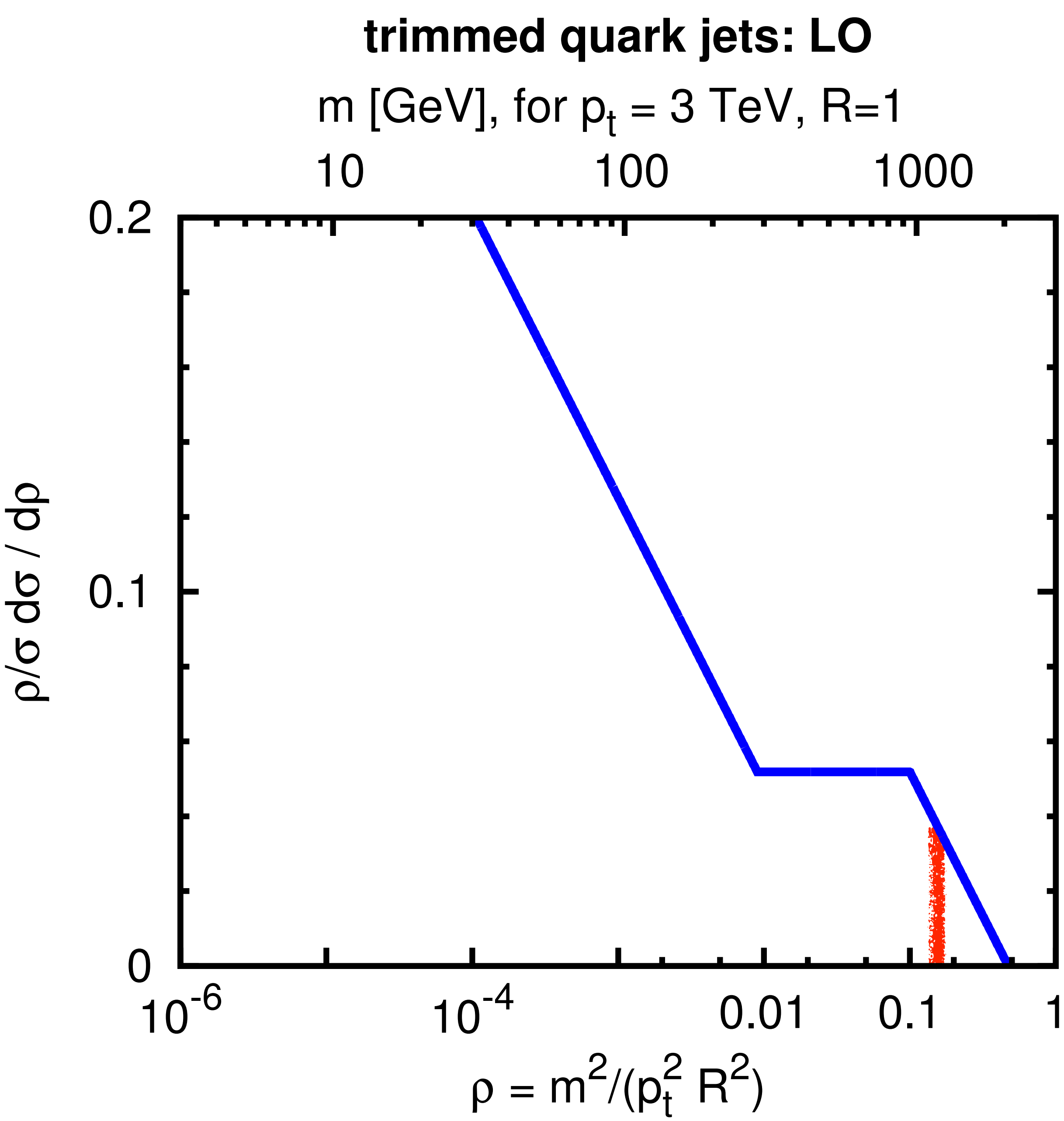
$$\rho = z(1-z)\theta^2$$

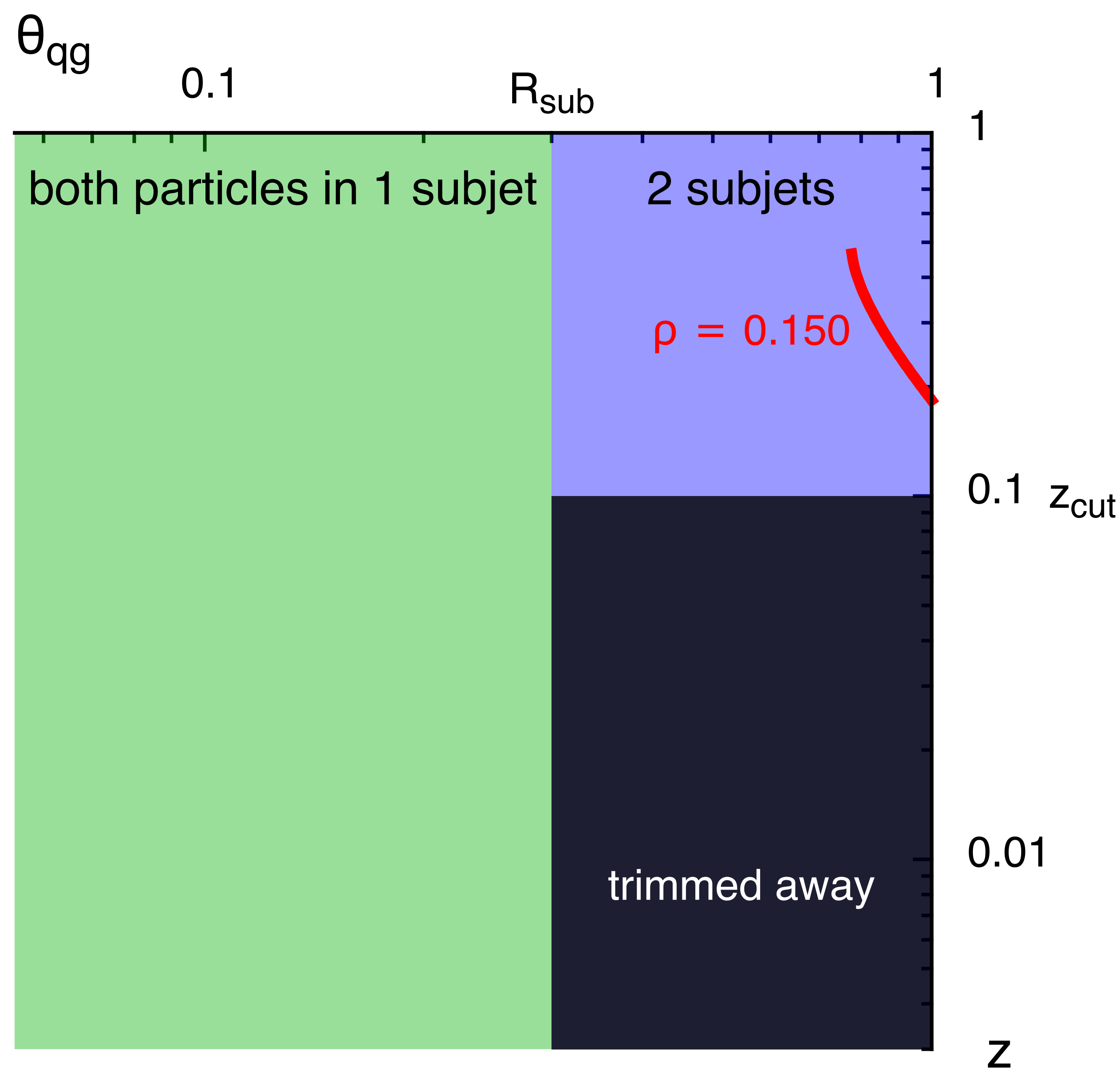
length of **fixed-ρ contour** ~ LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

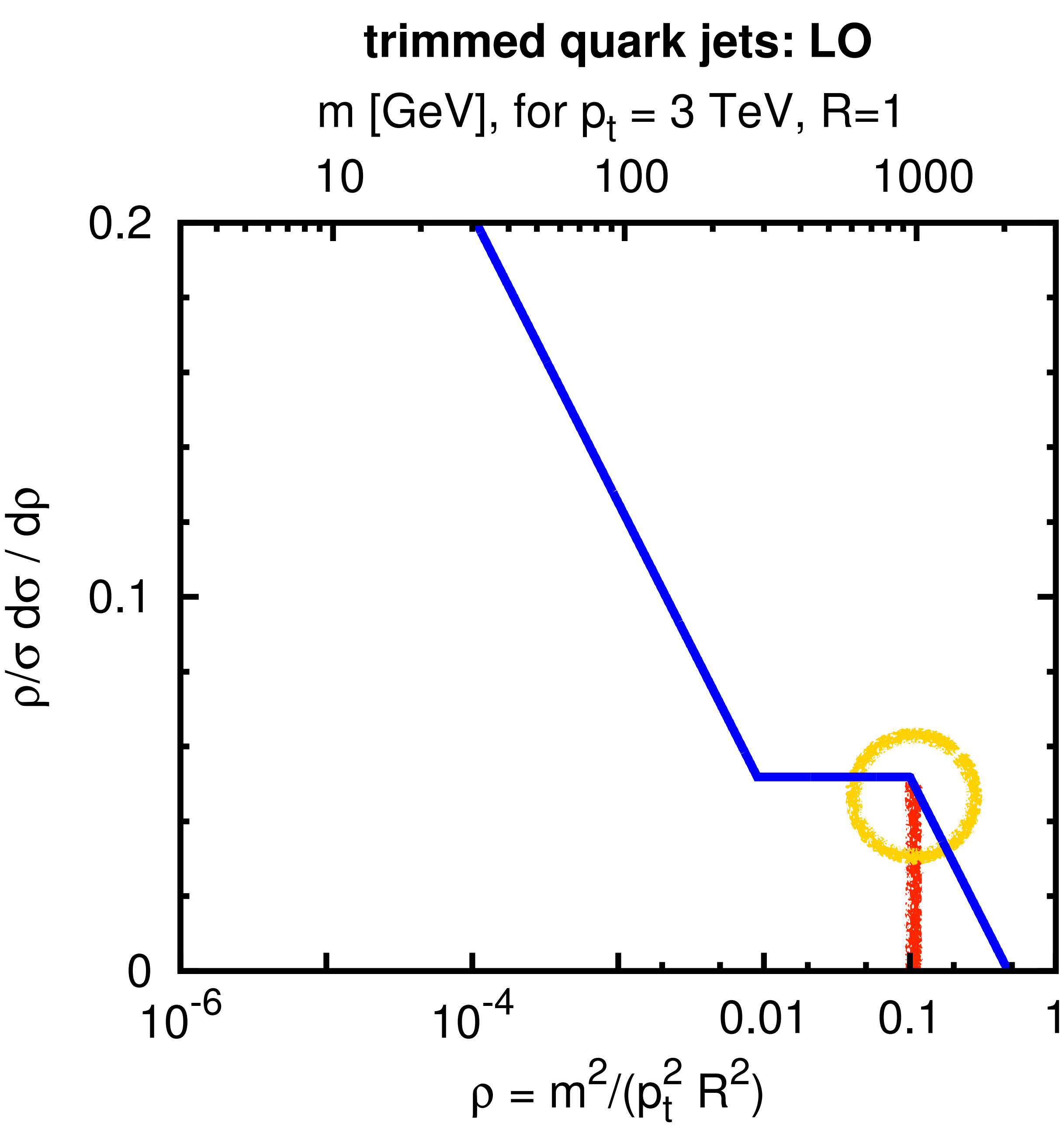
$$\rho = z(1-z)\theta^2$$

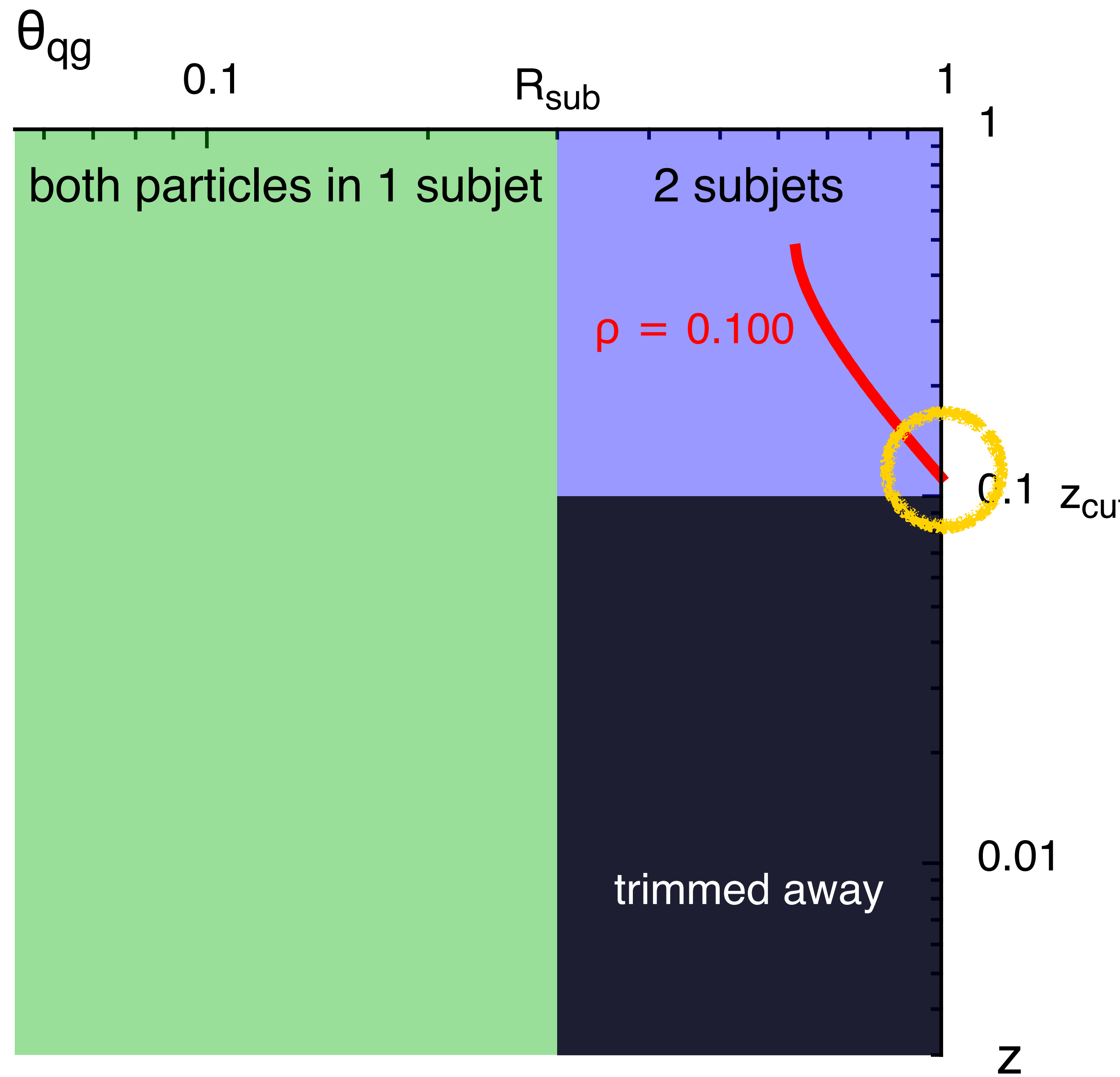
length of **fixed-ρ contour** ~ LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

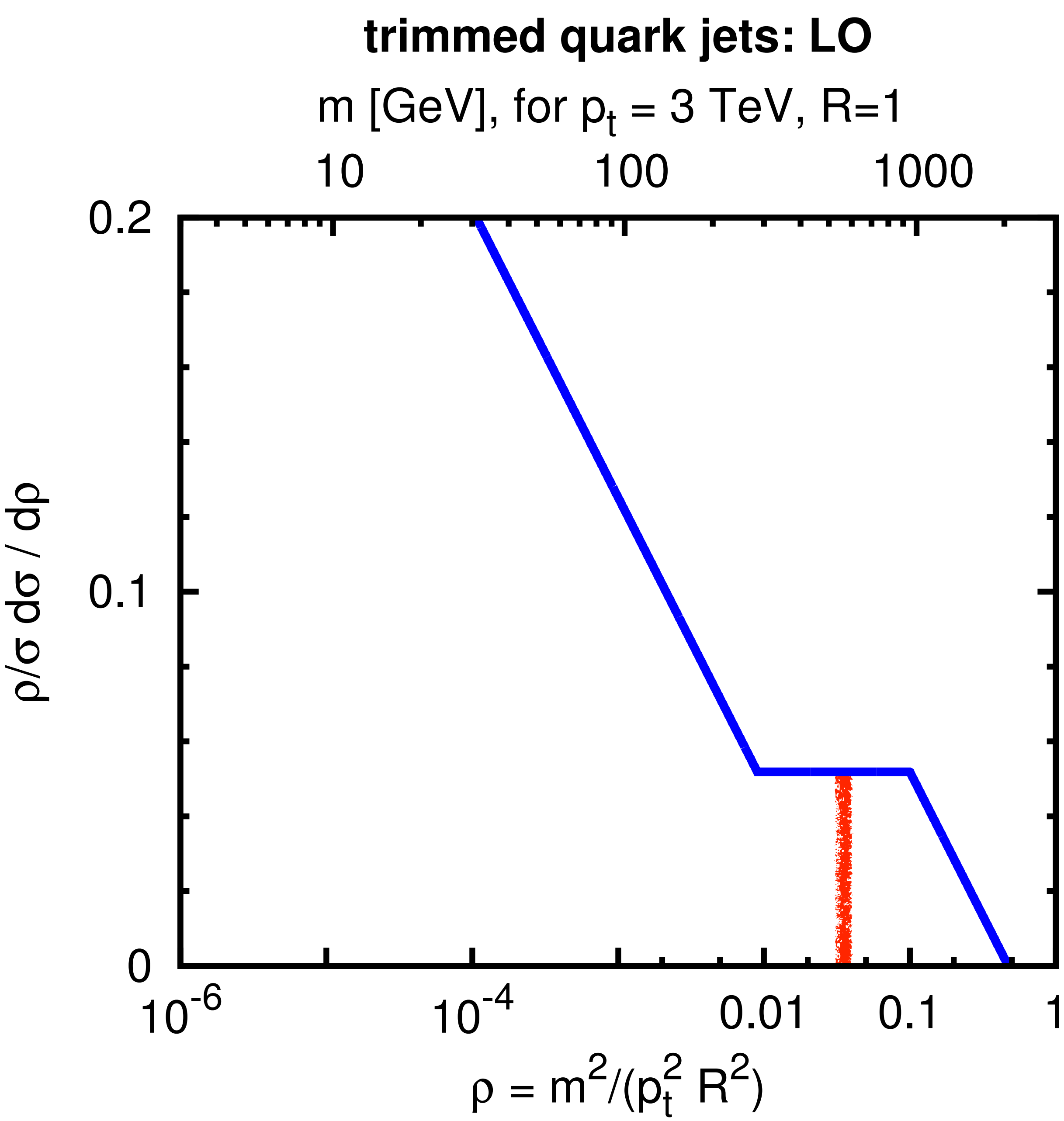
$$\rho = z(1-z)\theta^2$$

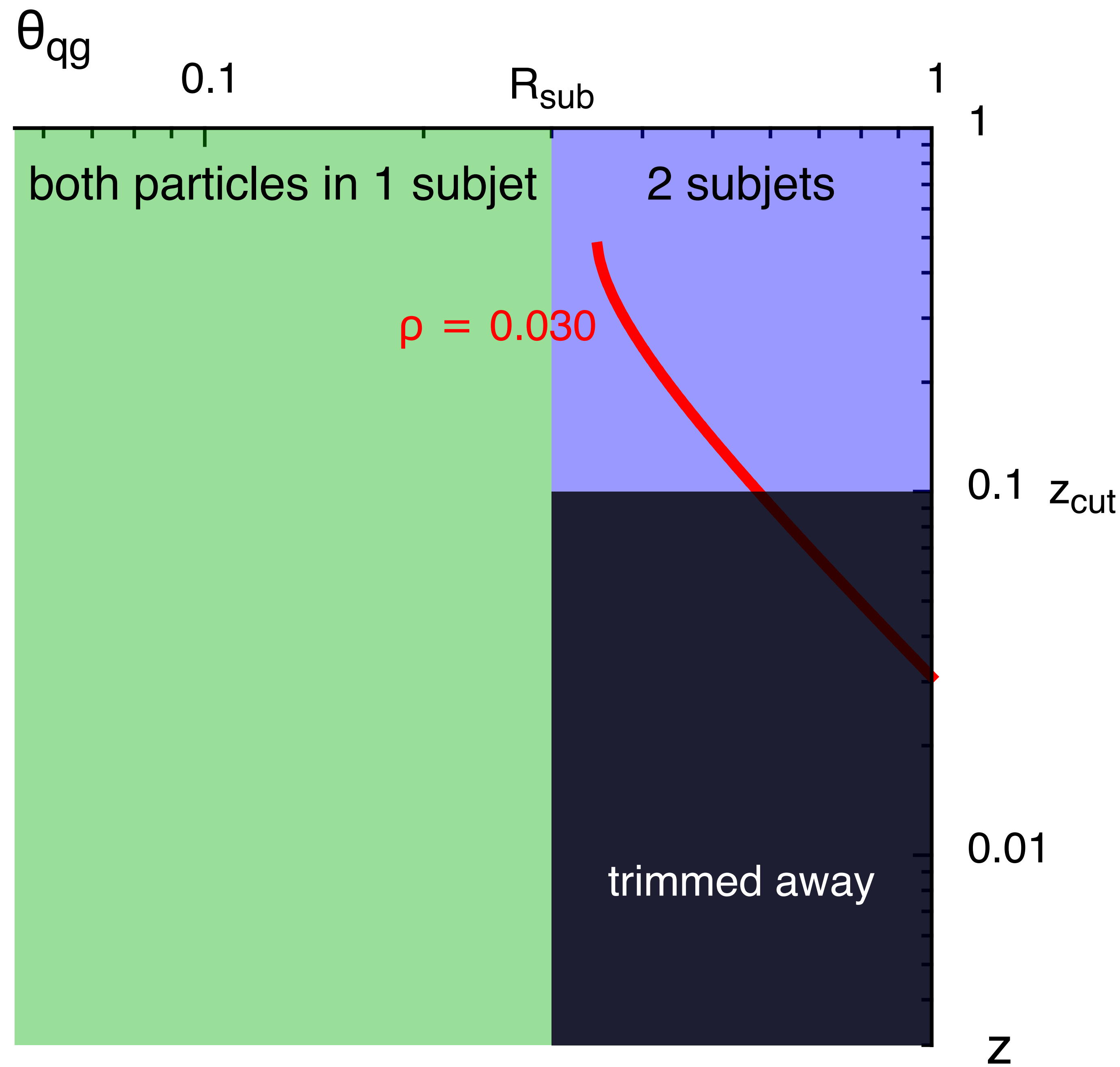
length of **fixed-ρ contour** ~ LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

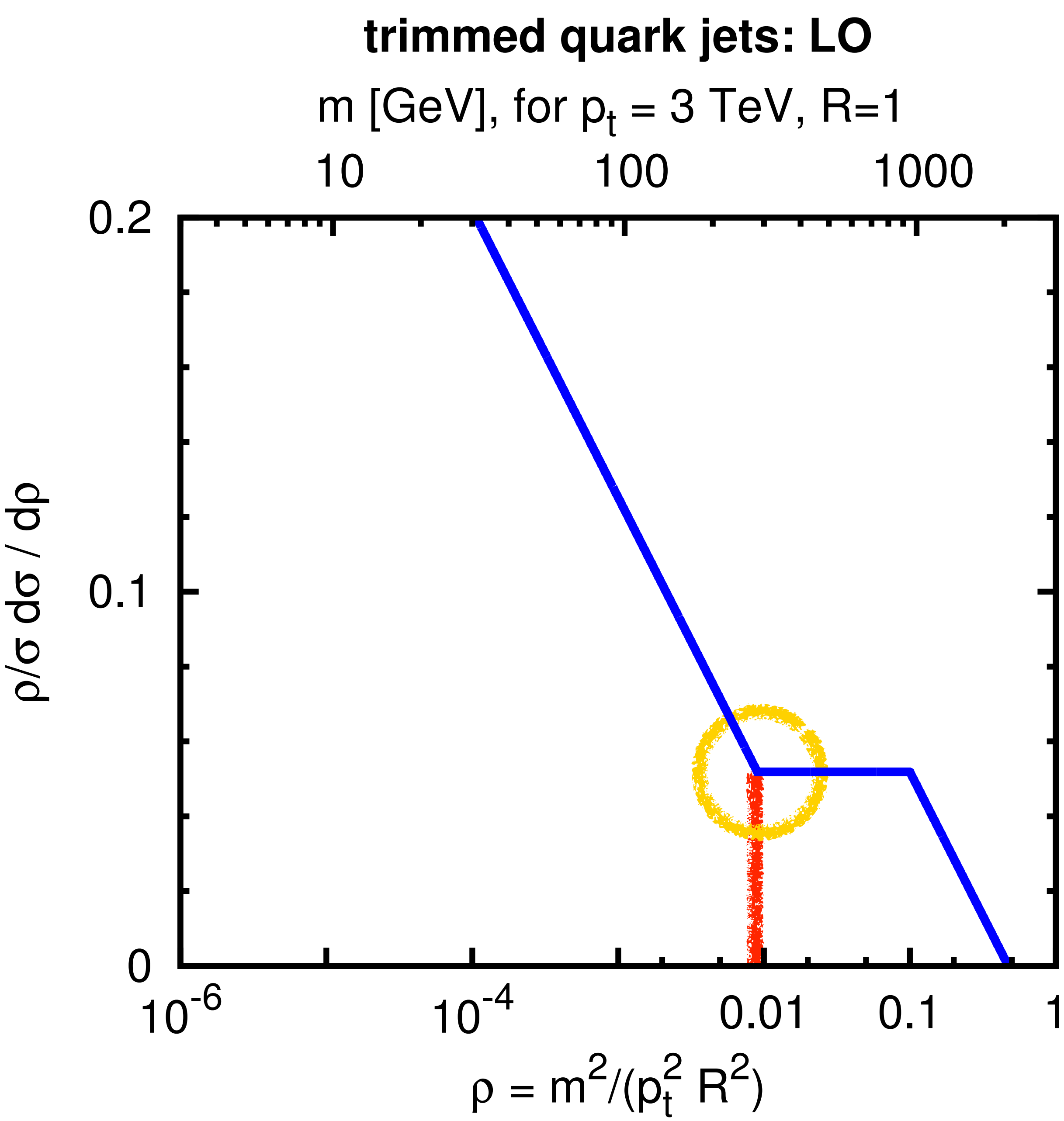
$$\rho = z(1-z)\theta^2$$

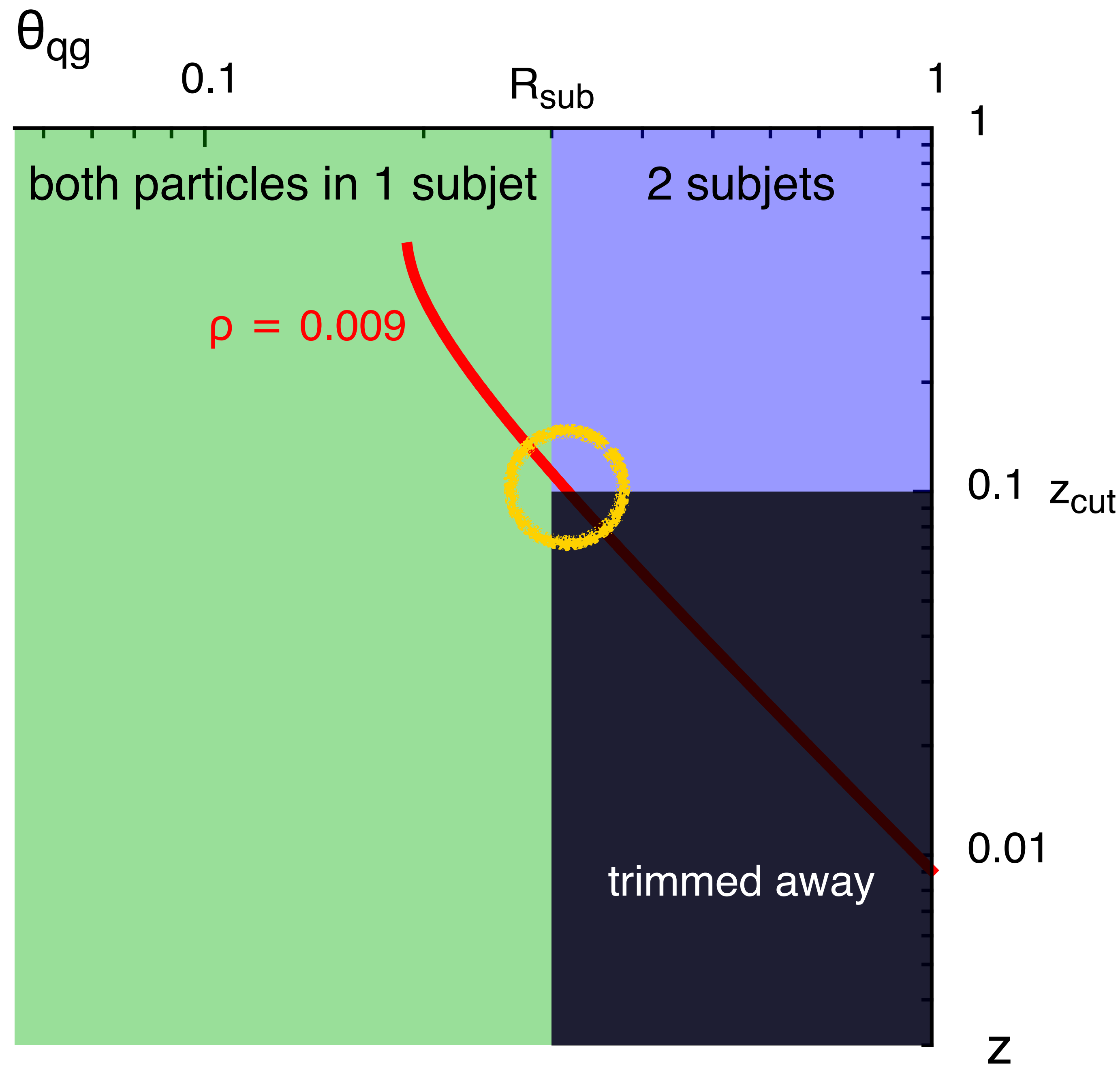
length of **fixed-ρ contour** ~ LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

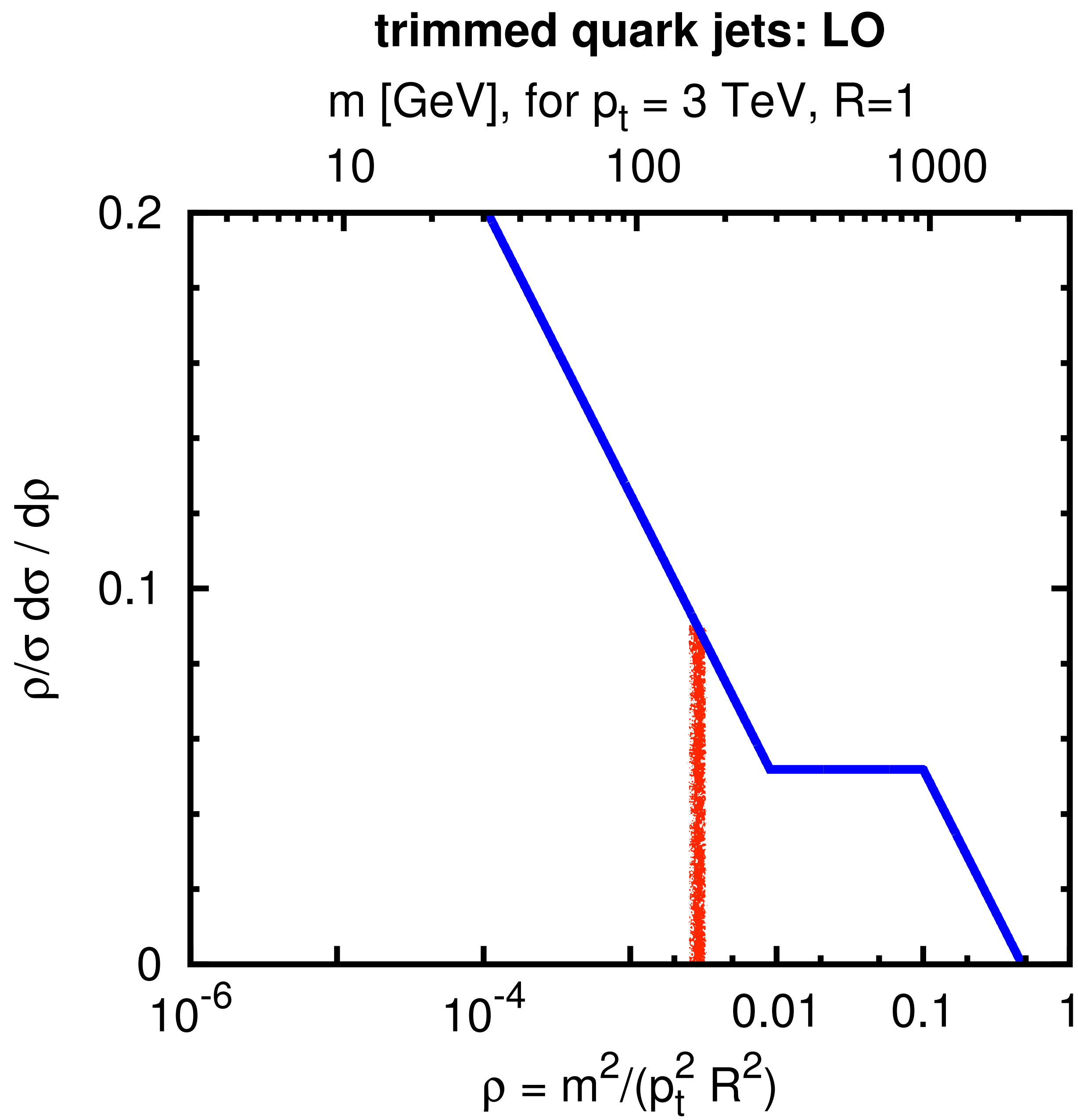
$$\rho = z(1-z)\theta^2$$

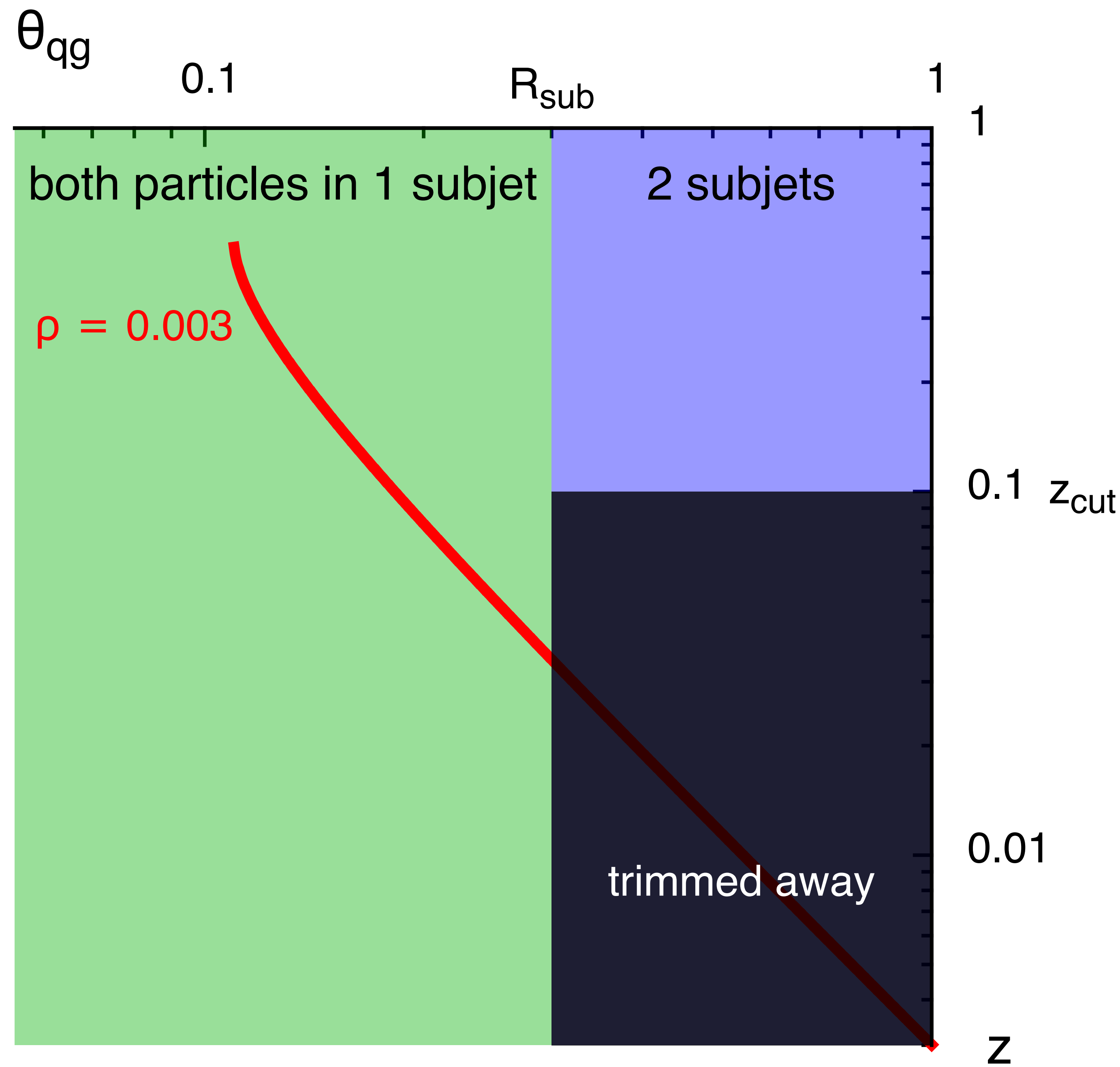
length of **fixed-ρ contour** ~ LO differential cross section

matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane





jet mass

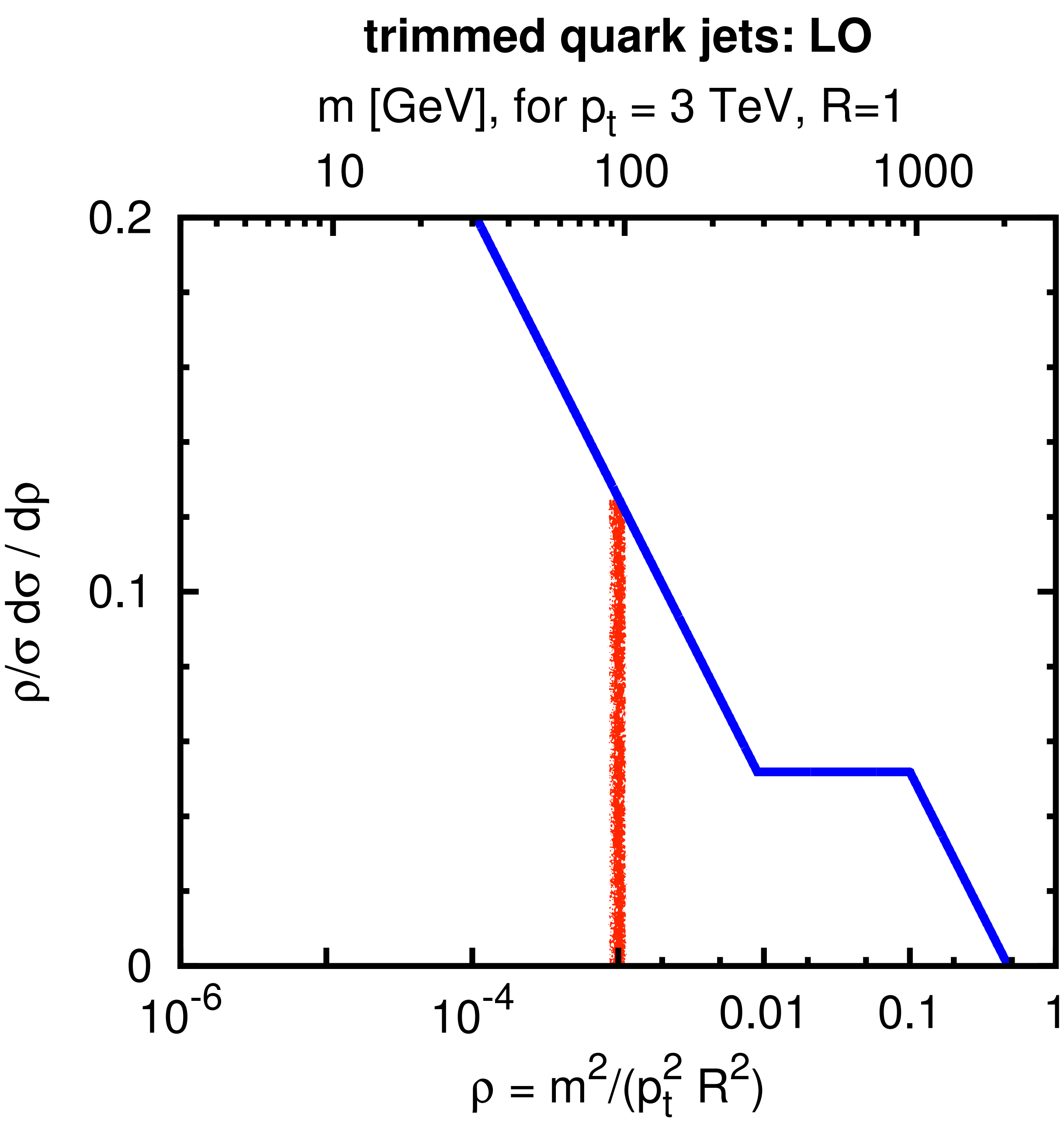
$$\rho = z(1-z)\theta^2$$

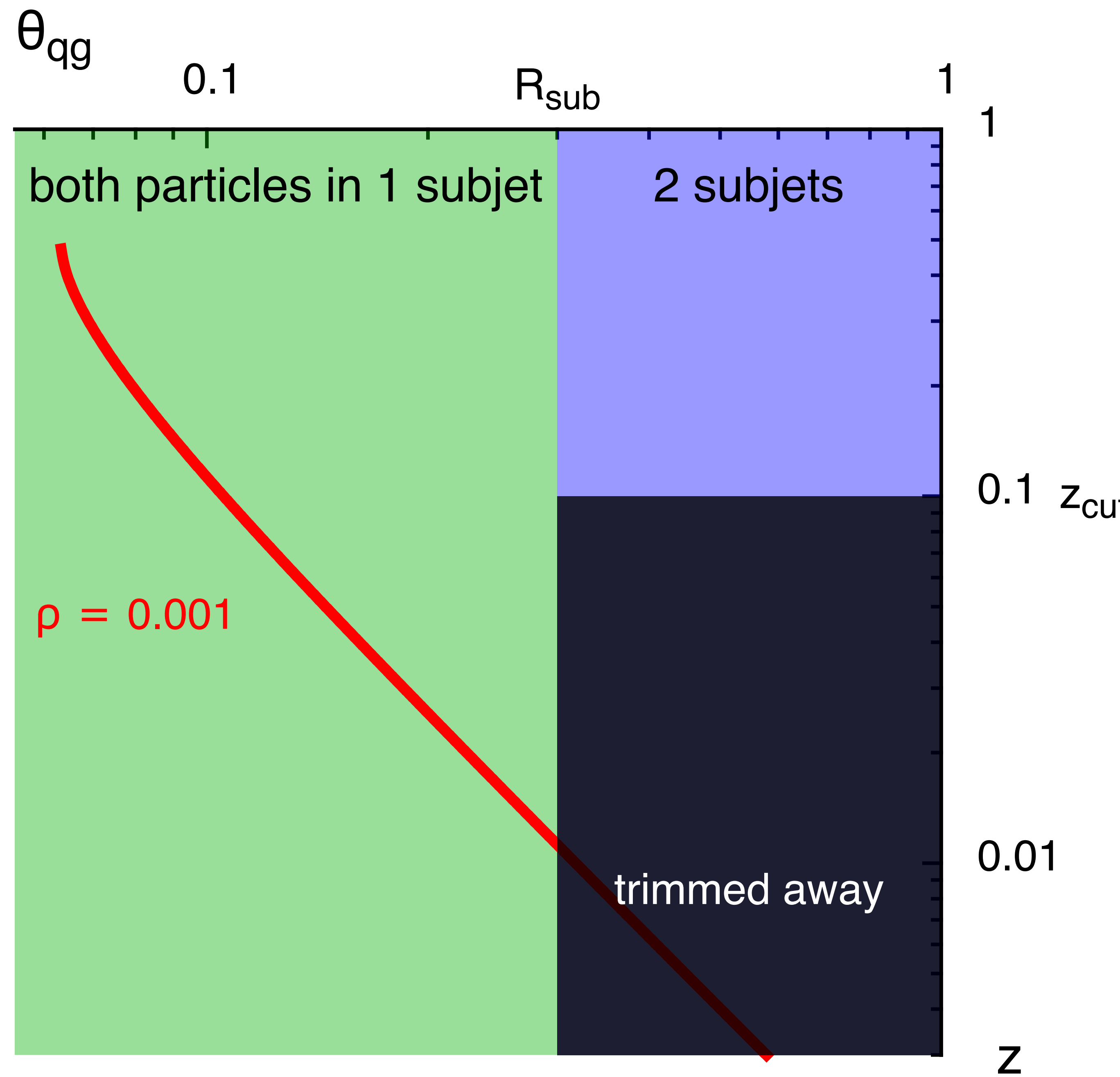
length of **fixed-ρ contour** ~ LO differential cross section

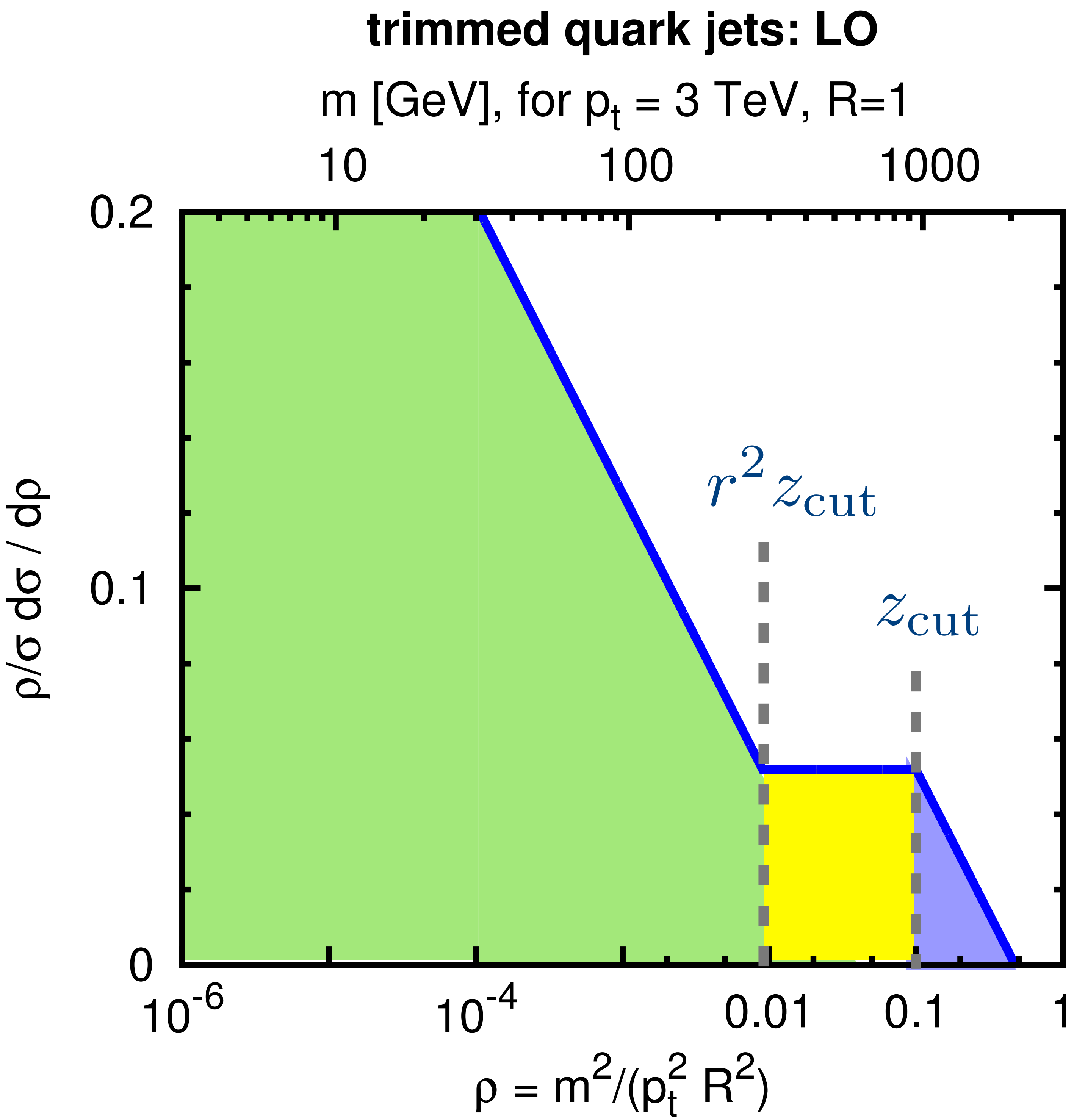
matrix element

$$\frac{\alpha_s C_F}{\pi}\frac{d\theta^2}{\theta^2}\frac{dz}{z}$$

emission probability ~ constant in log θ – log *z* plane







$$\frac{\rho}{\sigma}\frac{d\sigma^{(\mathrm{trim,LO})}}{d\rho} =$$

$$\frac{\alpha_s C_F}{\pi}\left(\ln\frac{r^2}{\rho} - \frac{3}{4}\right)$$

$$\frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{z_{\rm cut}} - \frac{3}{4}\right)$$

$$\frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{\rho} - \frac{3}{4}\right)$$

$$r = \frac{R_{\rm sub}}{R}$$

continue with all-order resummation of terms

$$\alpha_s^n \ln^m \rho$$

Inputs

QCD pattern of multiple soft/collinear emission

Analysis of taggers' behaviour for 1, 2, 3, … n, emissions

→ Establish which simplifying approximations to use for tagger & matrix elements

→ **Output**

approx. formula for tagger's mass distribution for $\rho \ll 1$

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} = \sum_{n=1}^{\infty} c_{nm}\,\alpha_s^n \ln^m \rho$$

keeping only terms with largest powers of ln ρ, e.g. $m = 2n$, $2n$-1

Trimming

$$\rho^{\text{trim}}(k_1, k_2, \dots k_n) \simeq \sum_i^n \rho^{\text{trim}}(k_i)$$

$$\sim \max_i\{\rho^{\text{trim}}(k_i)\}$$

Trimmed jet reduces (~) to sum of trimmed emissions

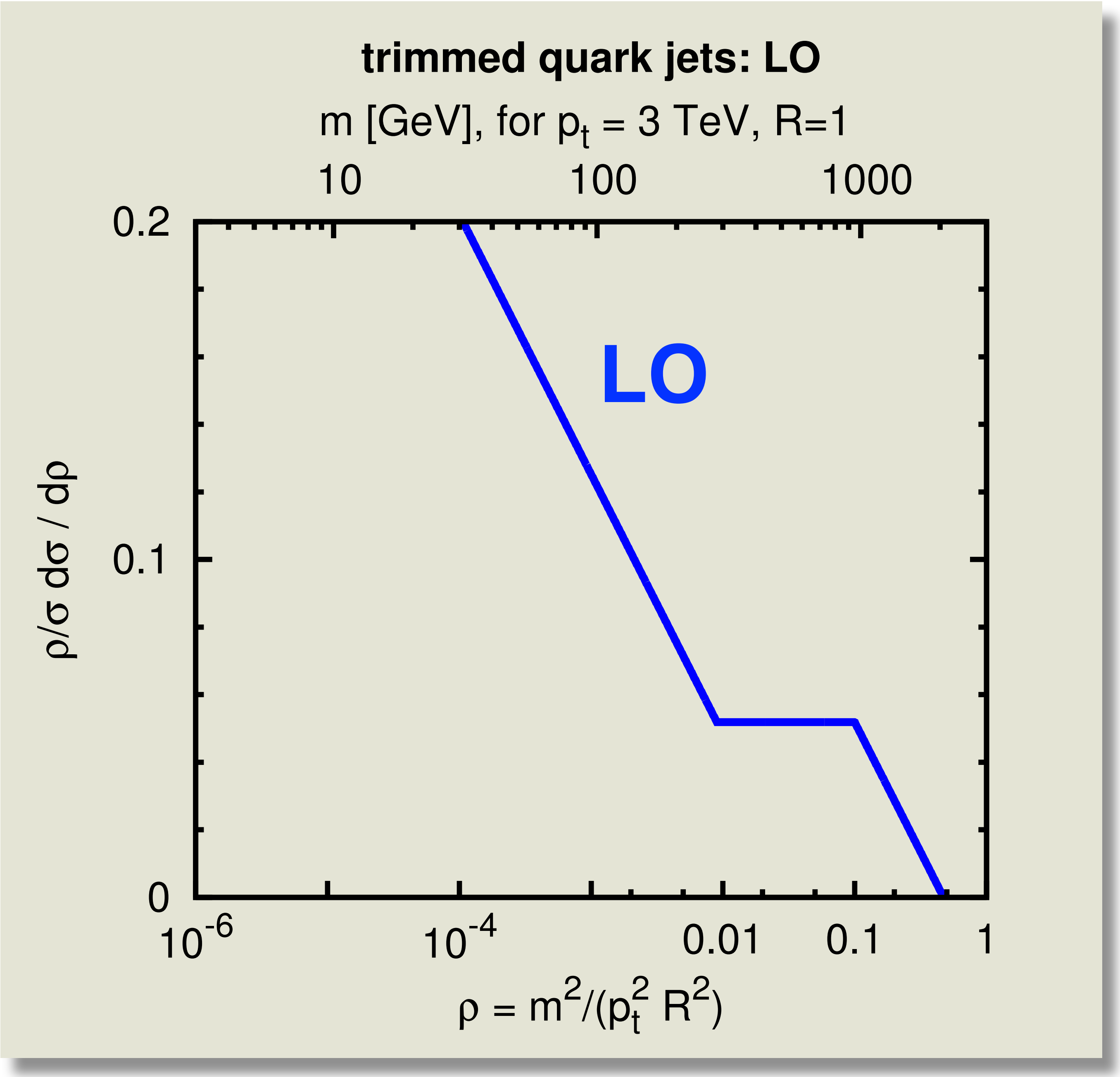
Matrix element

$$\sum_n \frac{1}{n!} \prod_i^n \frac{d\theta_i^2}{\theta_i^2} \frac{dz_i}{z_i} \frac{\alpha_s(\theta_i z_i p_t^{\text{jet}}) C_F}{\pi}$$

can use QED-like independent emissions, as if gluons don't split

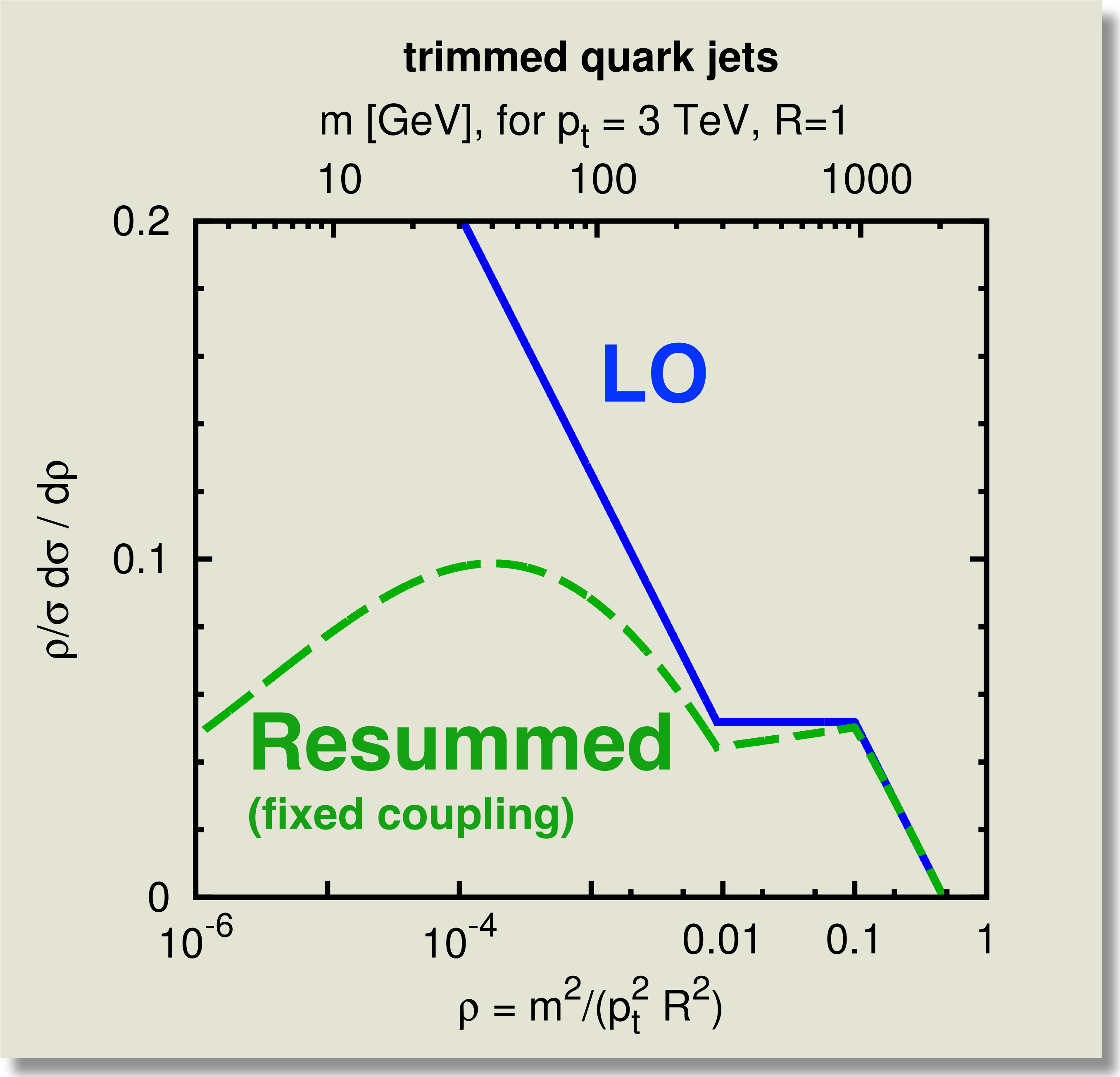
Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho}$$



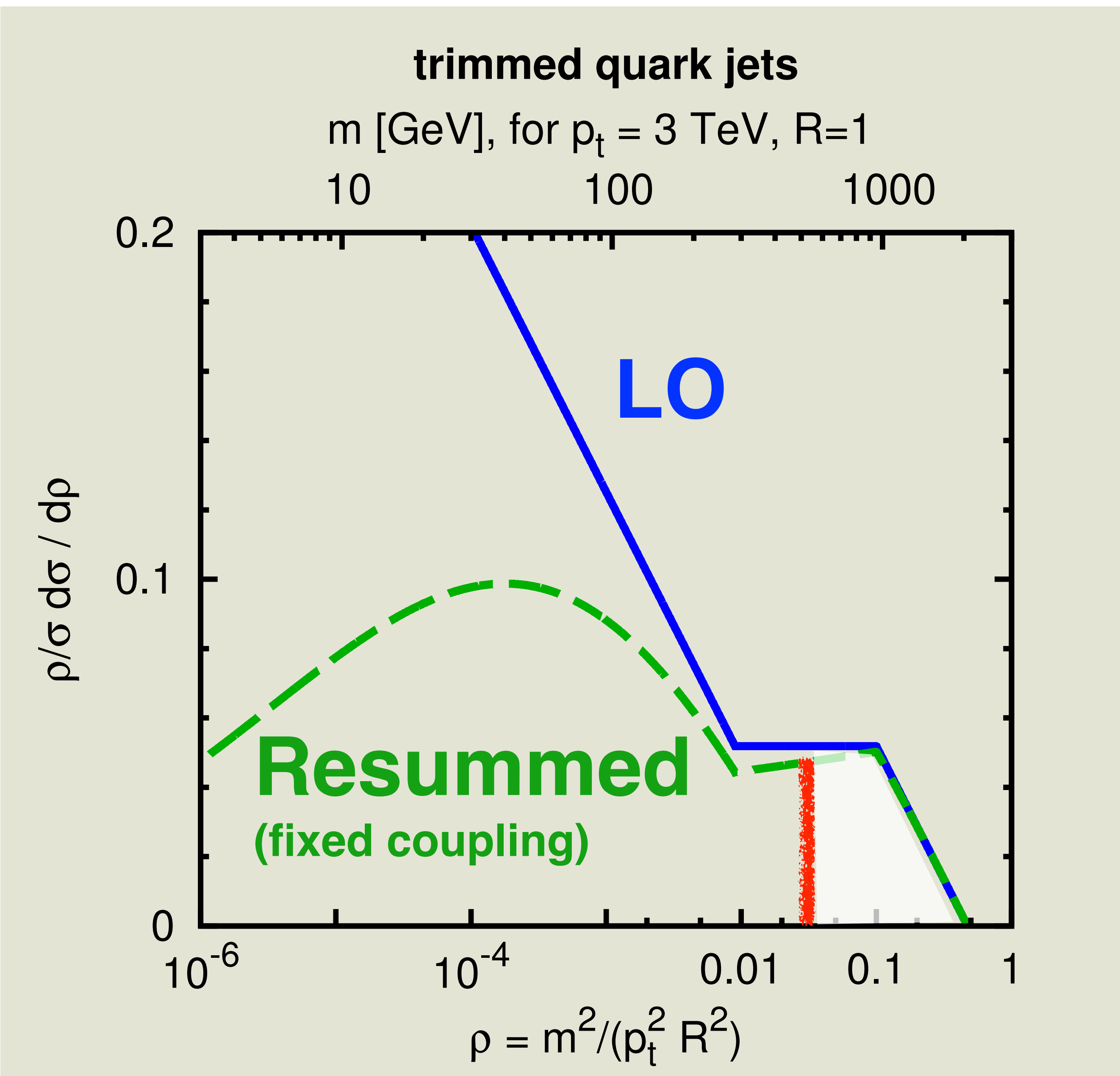
Trimming at all orders

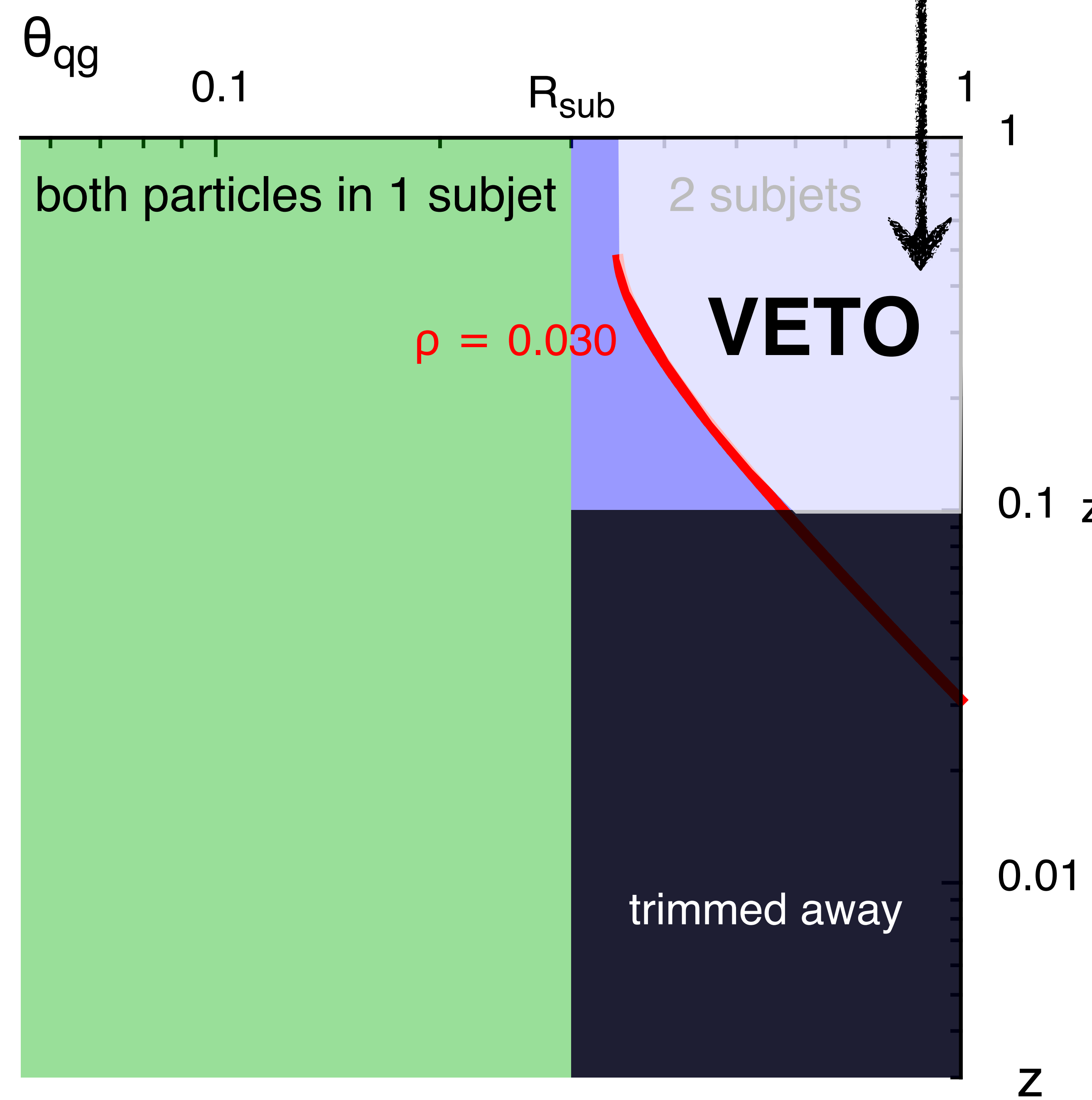
$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma} \frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$



$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

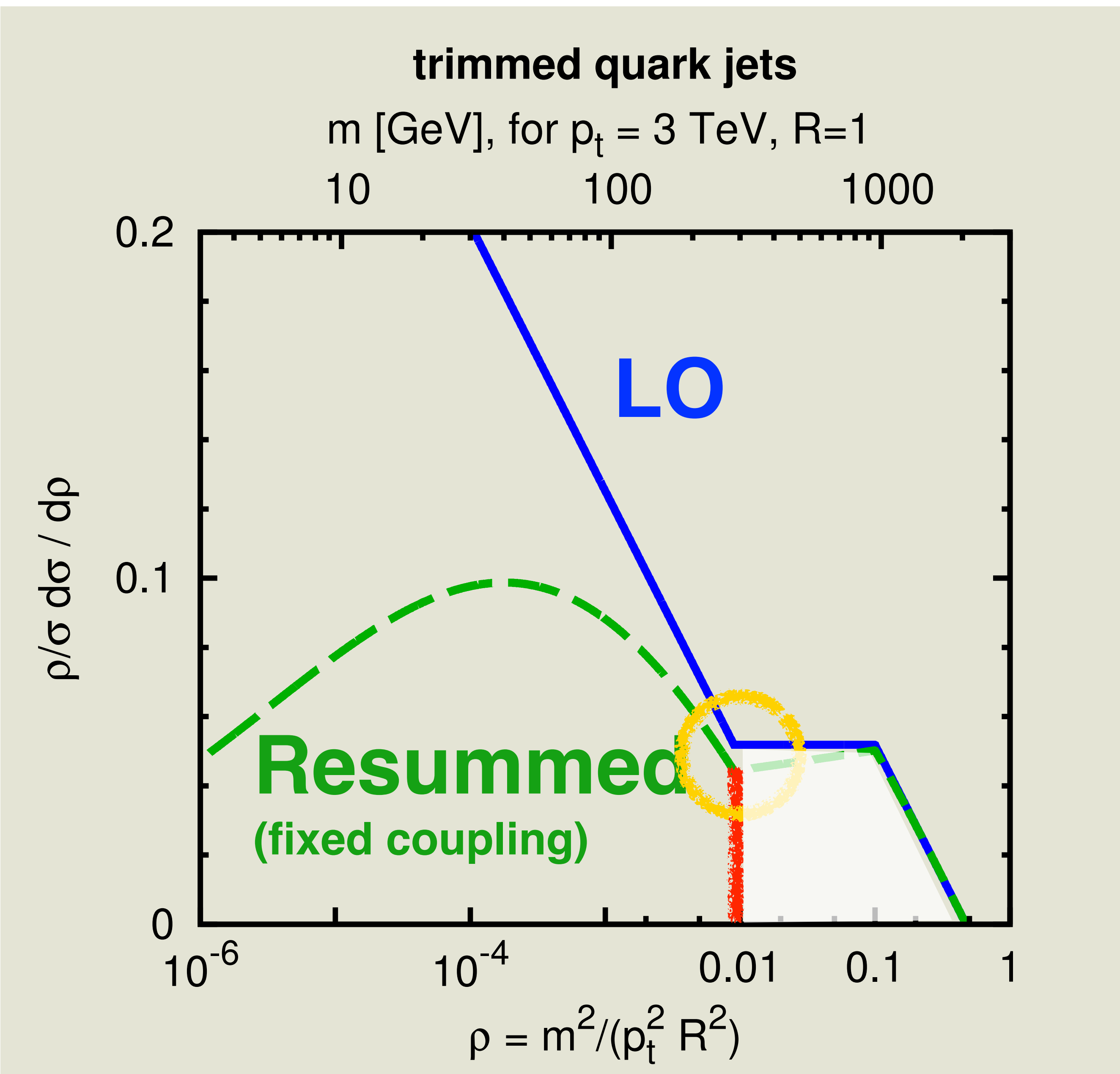


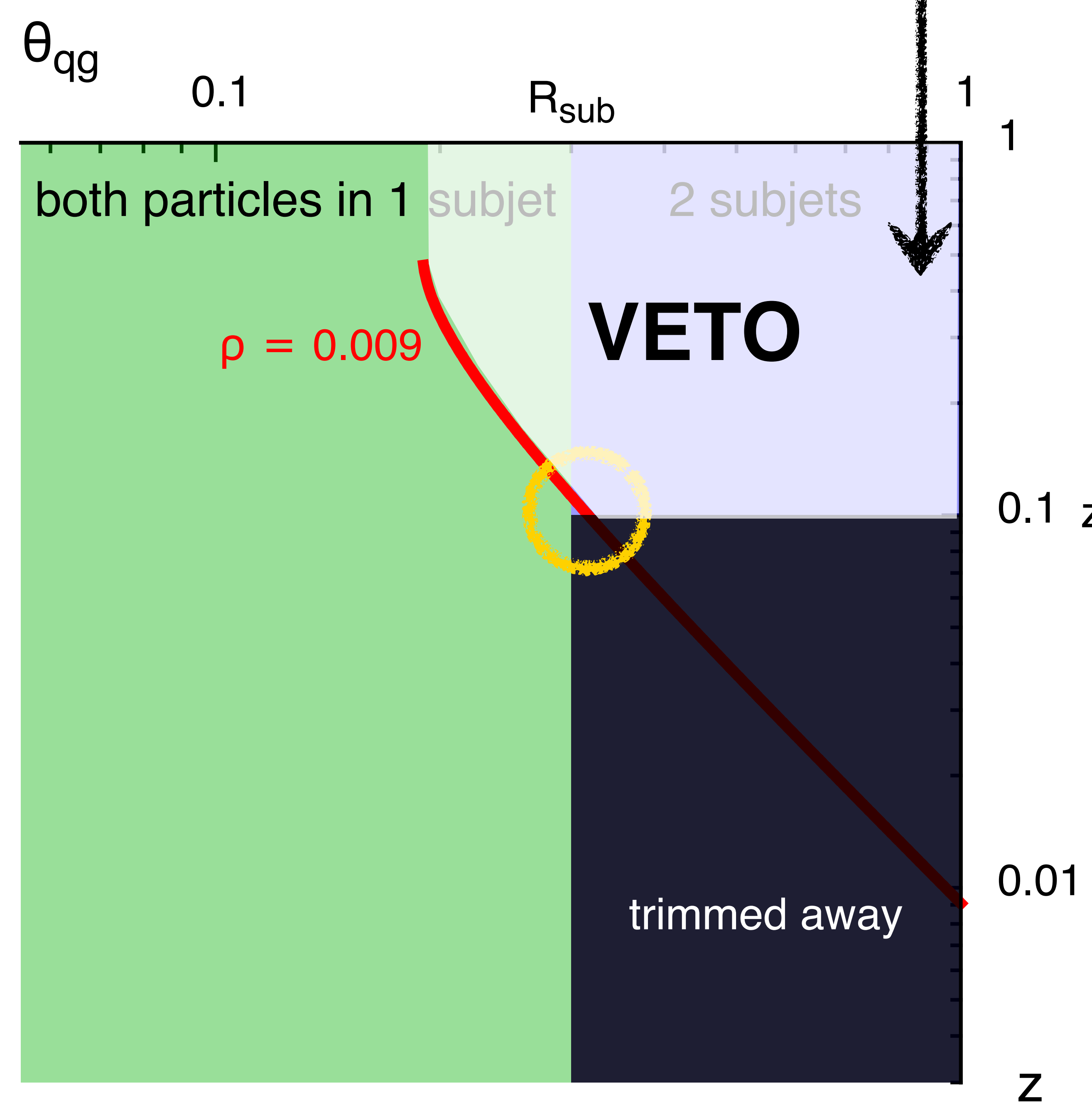


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

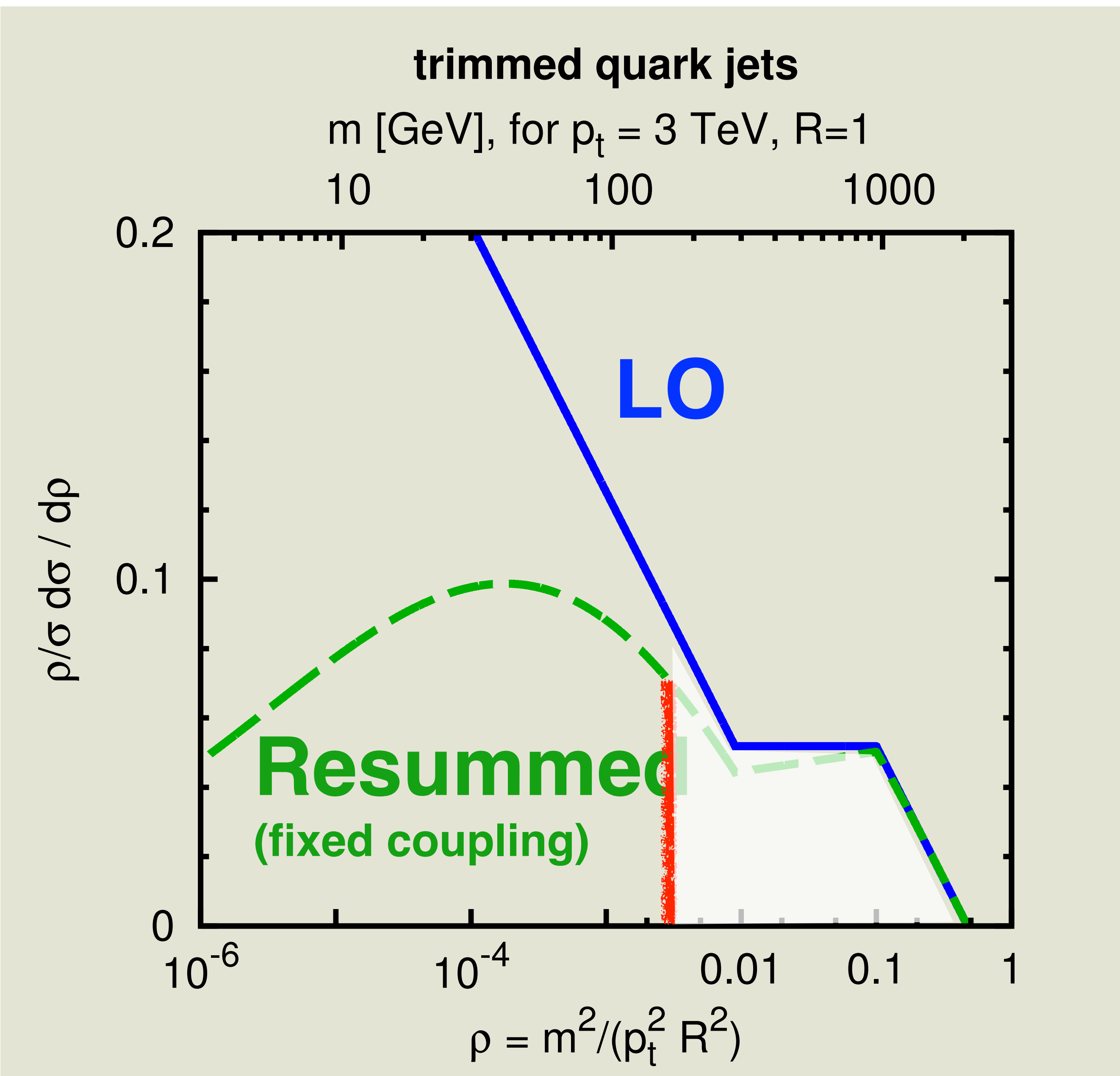


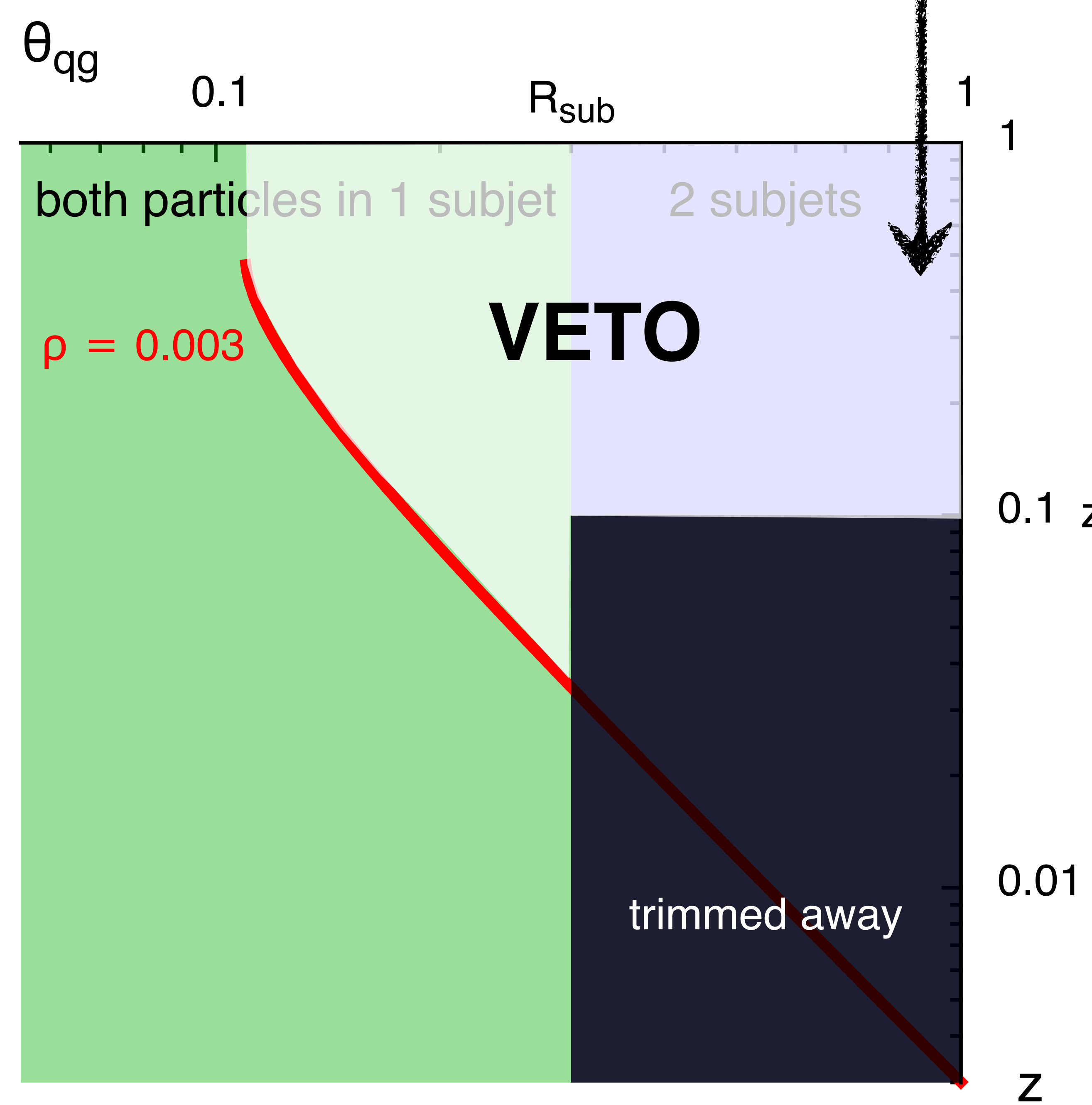


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

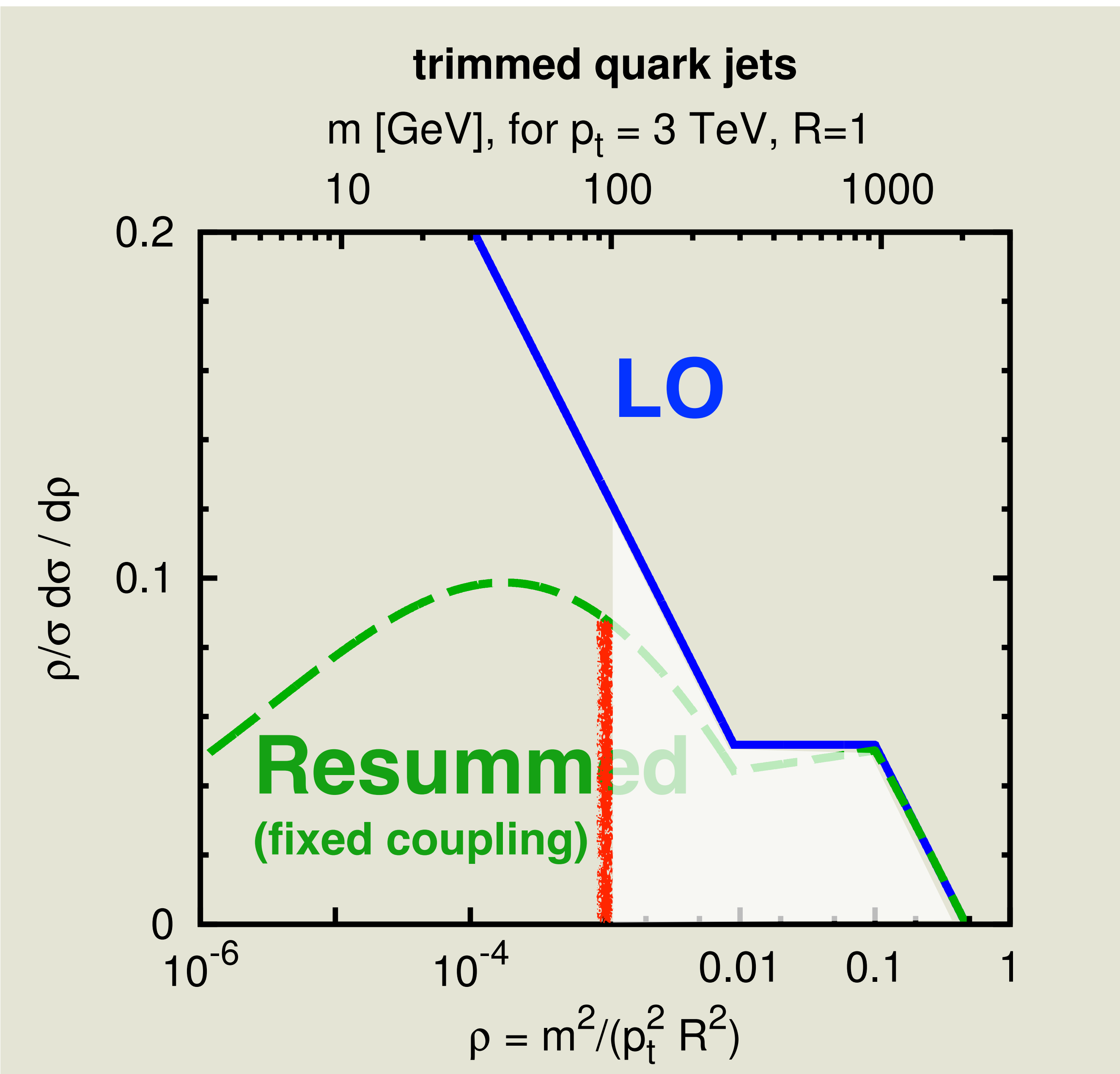


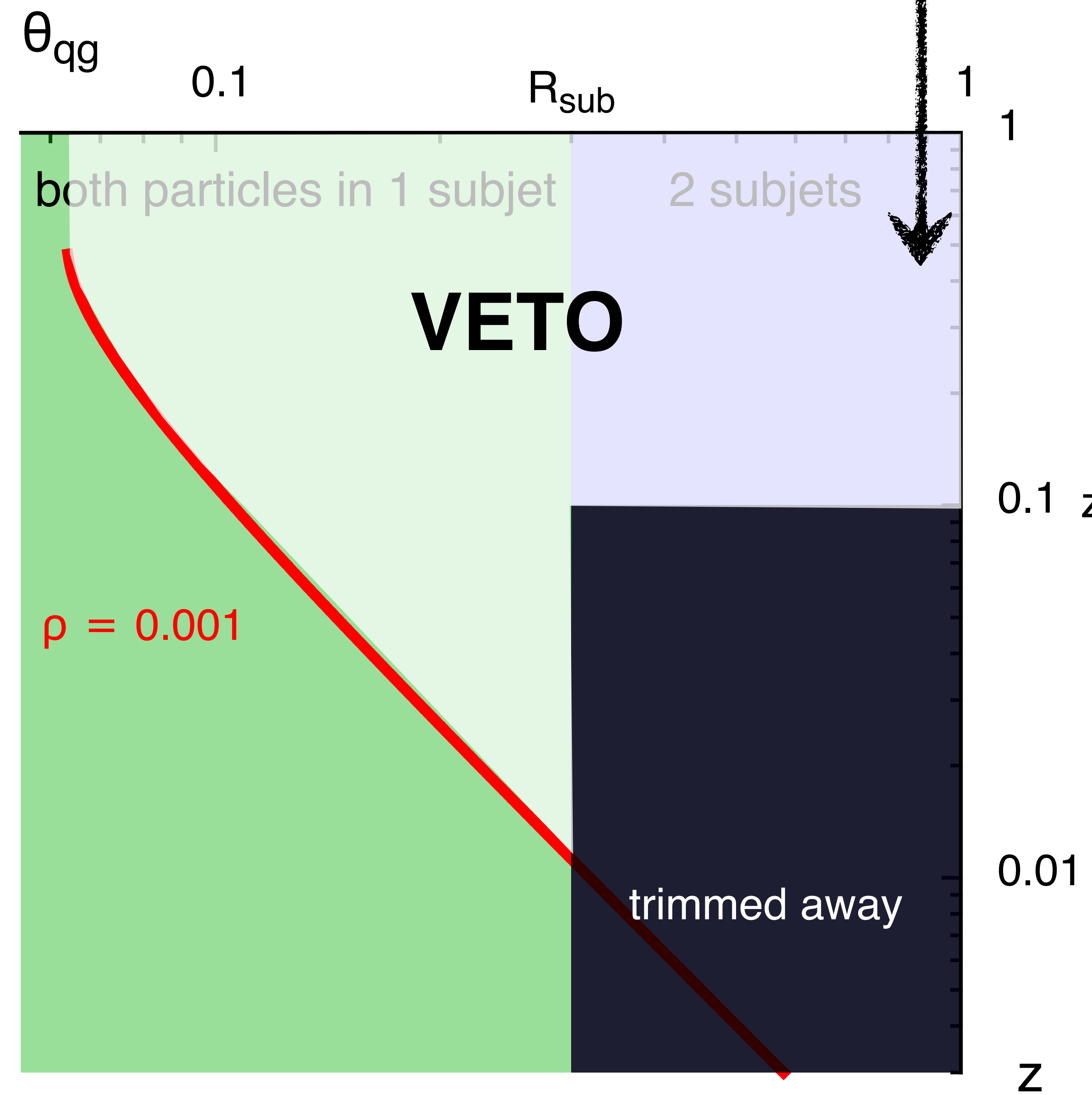


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

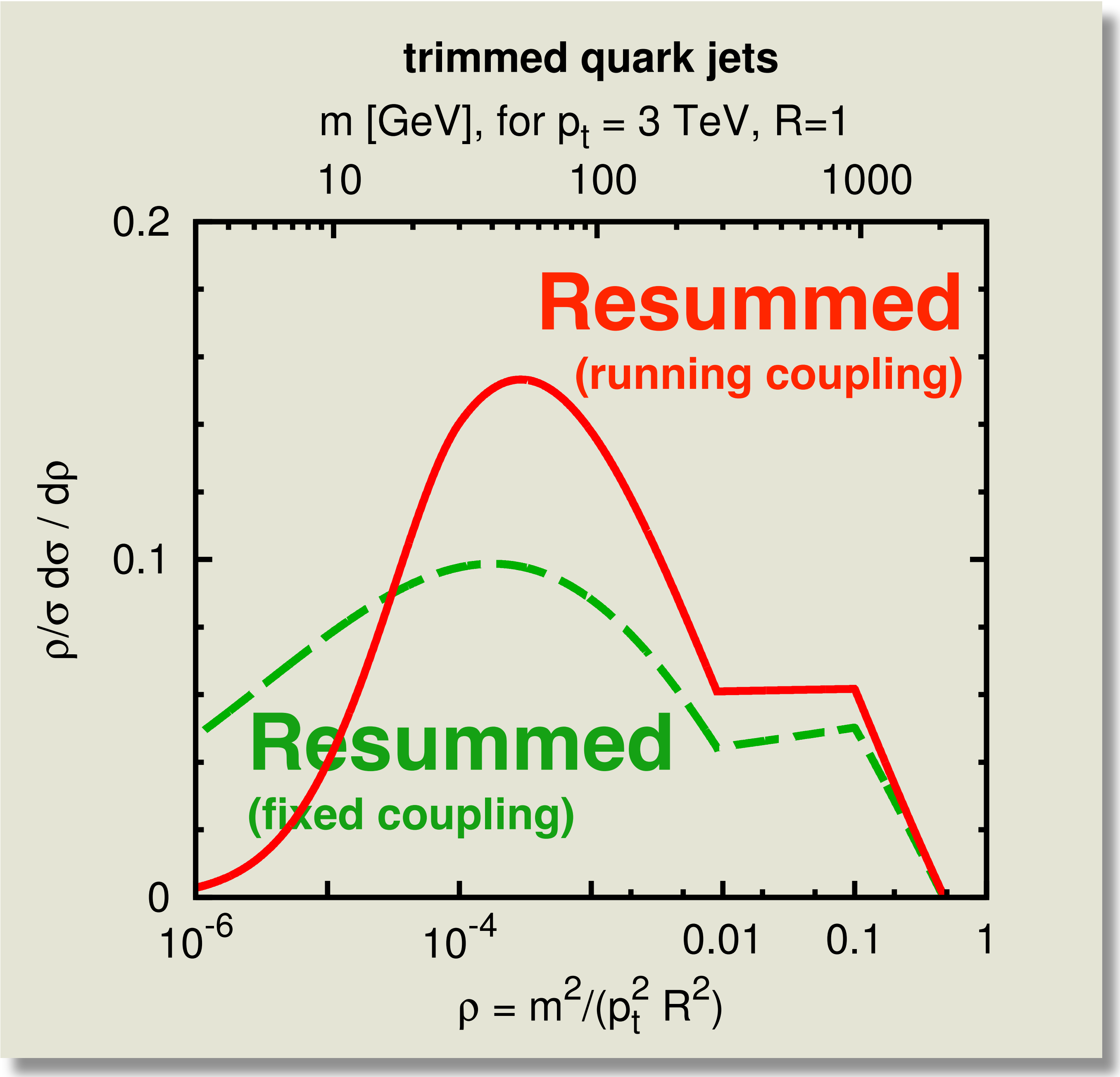
Sudakov





Trimming at all orders

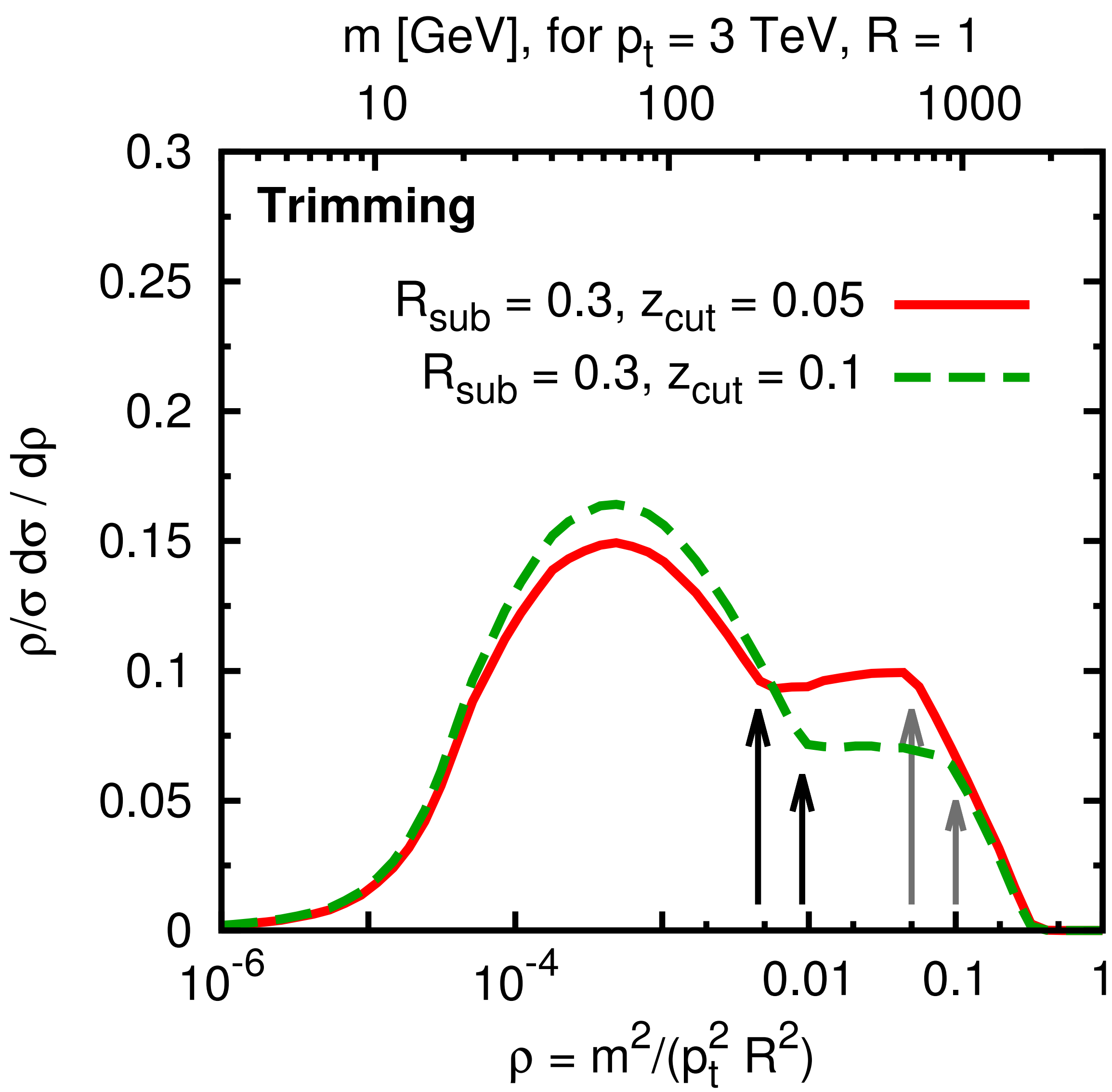
$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$



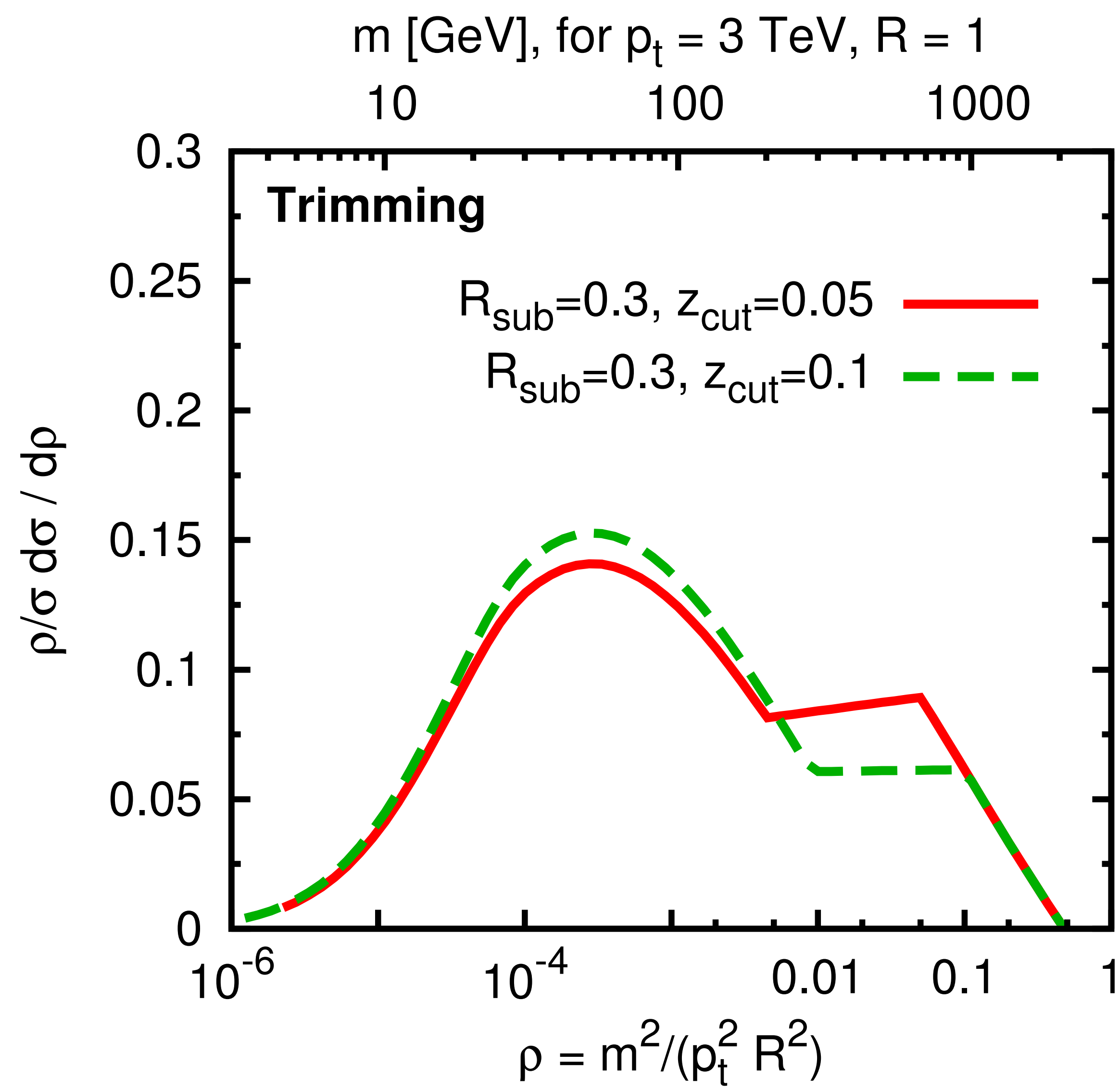
Full resummation also needs treatment of running coupling

Trimming: MC v. analytics

Monte Carlo



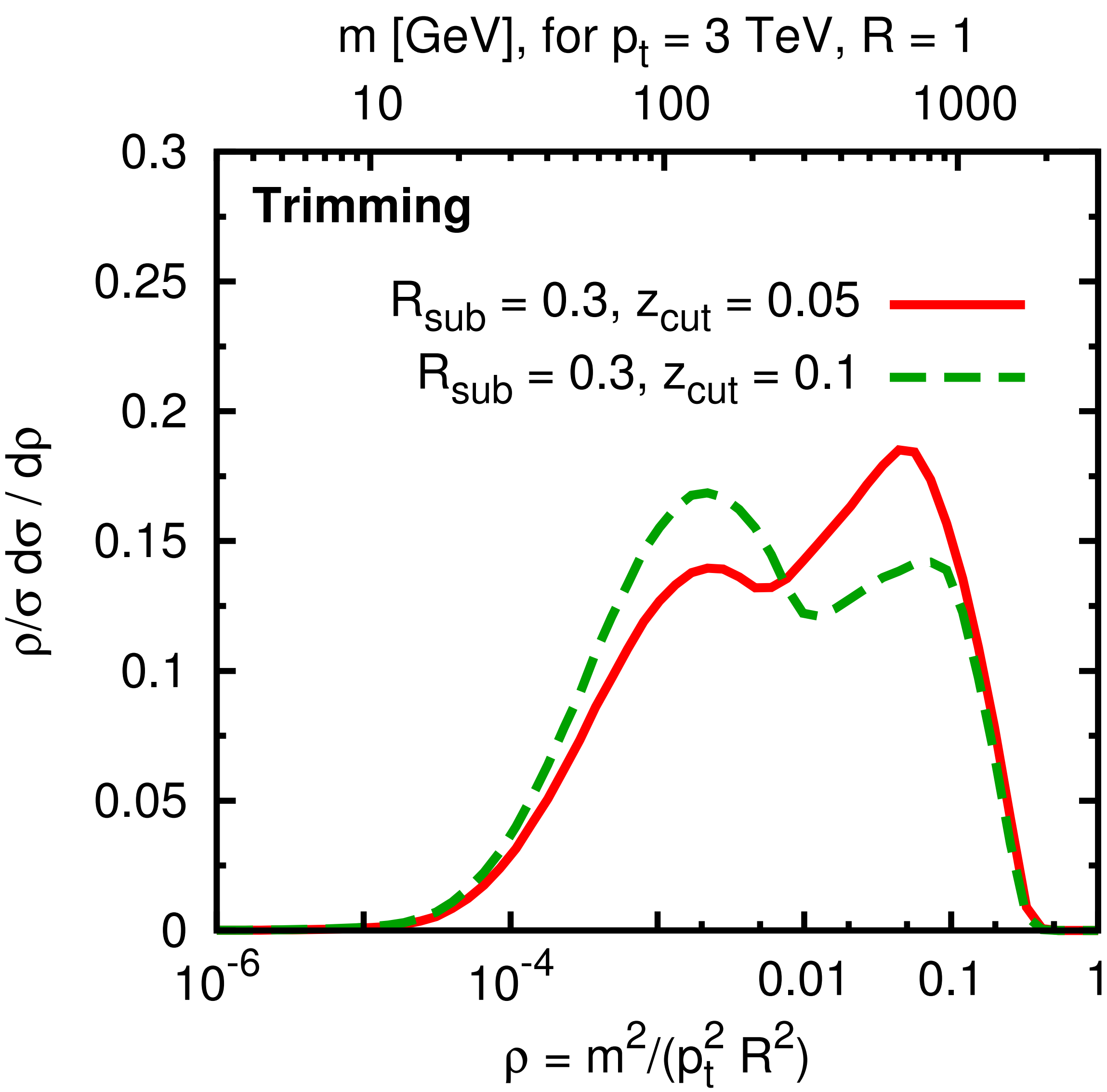
Analytic



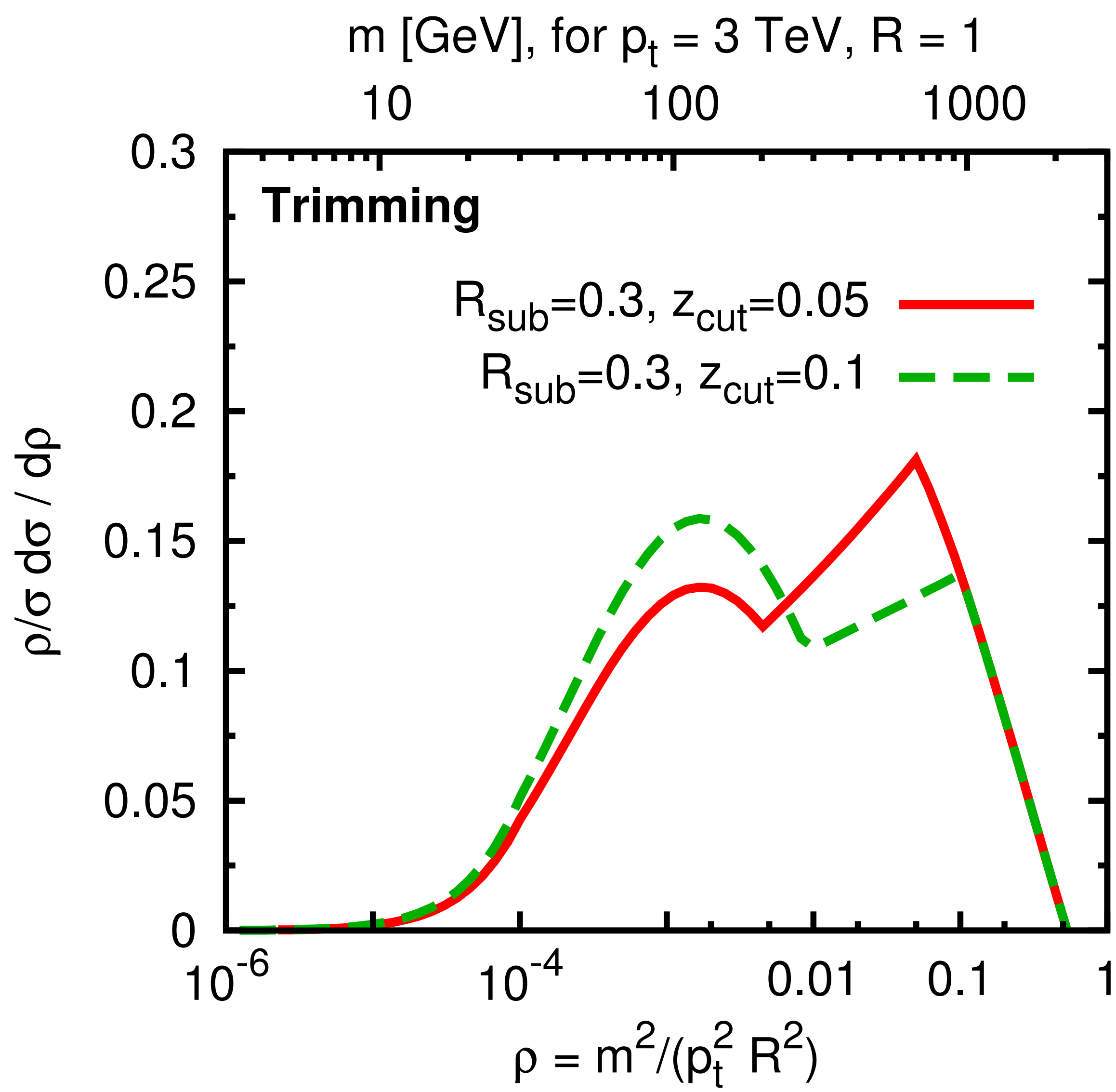
Non-trivial agreement!
(also for dependence on parameters)

Trimming: MC v. analytics

Monte Carlo



Analytic



Non-trivial agreement!
(also for dependence on parameters)

What logs, what accuracy?

Theorists state accuracy for the "cumulant" **Σ(ρ)**:

$$\Sigma(\rho) = \int_0^\rho d\rho' \frac{1}{\sigma}\frac{d\sigma}{d\rho'}$$

Use shorthand L = log 1/**ρ**

Trimming's **leading logs** (LL, in **Σ**) are:

$$\alpha_s L^2,\ \alpha_s^2 L^4,\ \ldots.\ \text{I.e. } \boldsymbol{\alpha_s^n L^{2n}}$$

Just like the jet mass

We also have **next-to-leading logs** (NLL): $\boldsymbol{\alpha_s^n L^{2n-1}}$

What logs, what accuracy?

Theorists state accuracy for the "cumulant" **Σ(ρ)**:

$$\Sigma(\rho) = \int_0^\rho d\rho' \frac{1}{\sigma}\frac{d\sigma}{d\rho'}$$

Use shorthand L = log 1/ρ

Trimming's **leading logs** (LL, in **Σ**) are:

$$\alpha_s L^2,\ \alpha_s^2 L^4,\ \ldots.\ \text{I.e.}\ \boldsymbol{\alpha_s^n L^{2n}}$$

Just like the jet mass

We also have **next-to-leading logs** (NLL): $\boldsymbol{\alpha_s^n L^{2n-1}}$

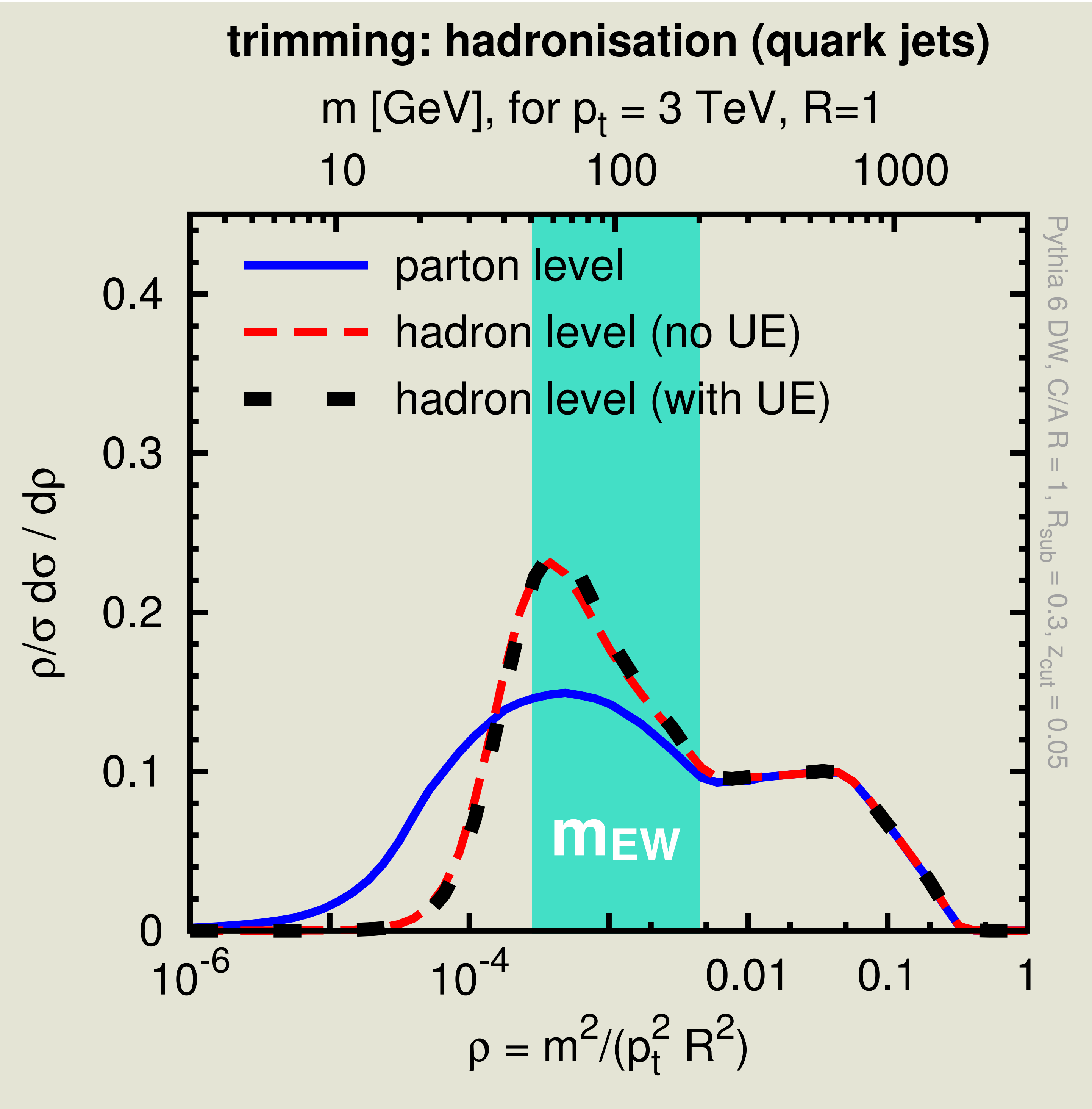
Could we do better? Yes: NLL in ln **Σ**:

$$\ln\Sigma:\ \alpha_s^n L^{n+1}\ \text{and}\ \alpha_s^n L^n$$

Trimmed mass is like plain jet mass (with R → R_{sub}), and this accuracy involves **non-global logs, clustering logs**

Everything so far was at parton level

Are partons sufficient?



hadronisation adds roughly

$$\Delta m^2 \sim \mu_{\text{NP}}\, p_t^{\text{jet}}\, R_{\text{sub}}$$
$$\sim (30\,\text{GeV})^2$$

to the trimmed-jet squared mass

non-perturbative QCD important for EW-scale phenomenology...

It's even worse for plain jet mass but better for other taggers...

Comments

Past years → a **vast trove of ideas** for jet substructure tagging.

But maybe it's **time to** try to **go back to "basics"**
→ detailed understanding about how our methods work.

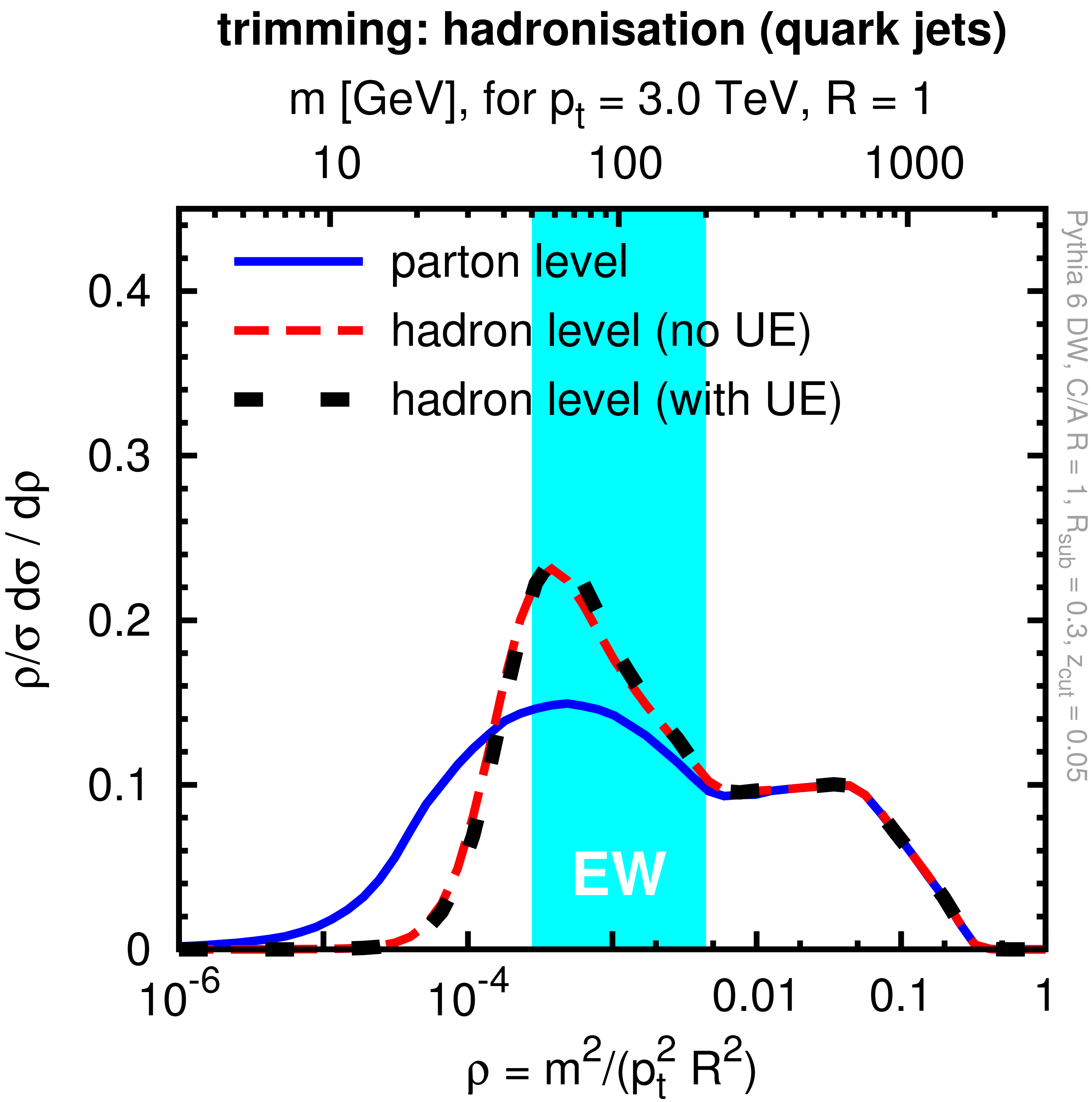
Trimming was a particularly illustrative case:

- has non-trivial structure, relevant for phenomenology
- can mostly be understood from LO calculation & standard resummation techniques — quite similar to jet mass
- non-perturbative effects are relevant

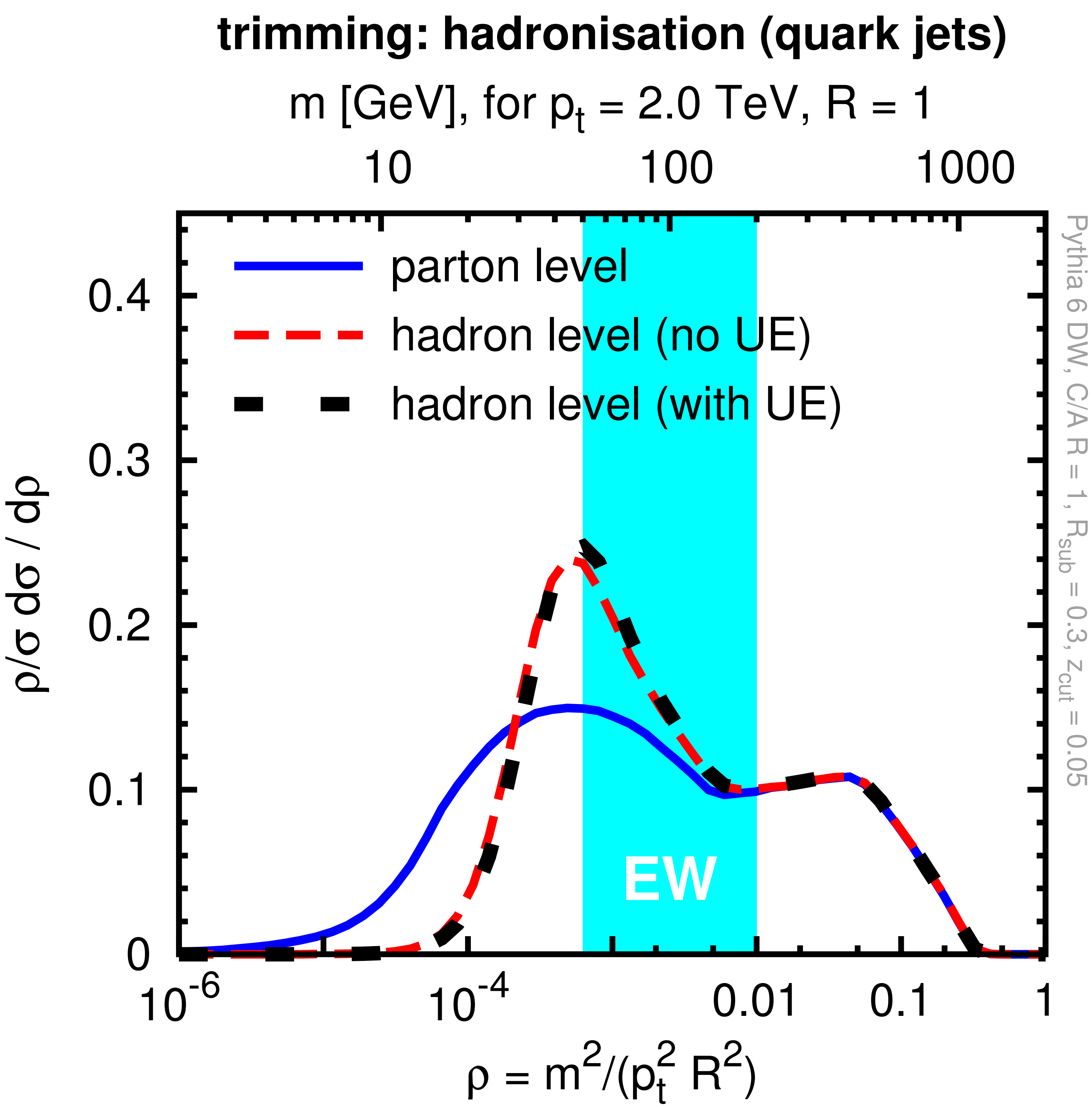
Now over to Simone, who will discuss **pruning** and **MDT**

EXTRAS

Trimming v. jet p_t



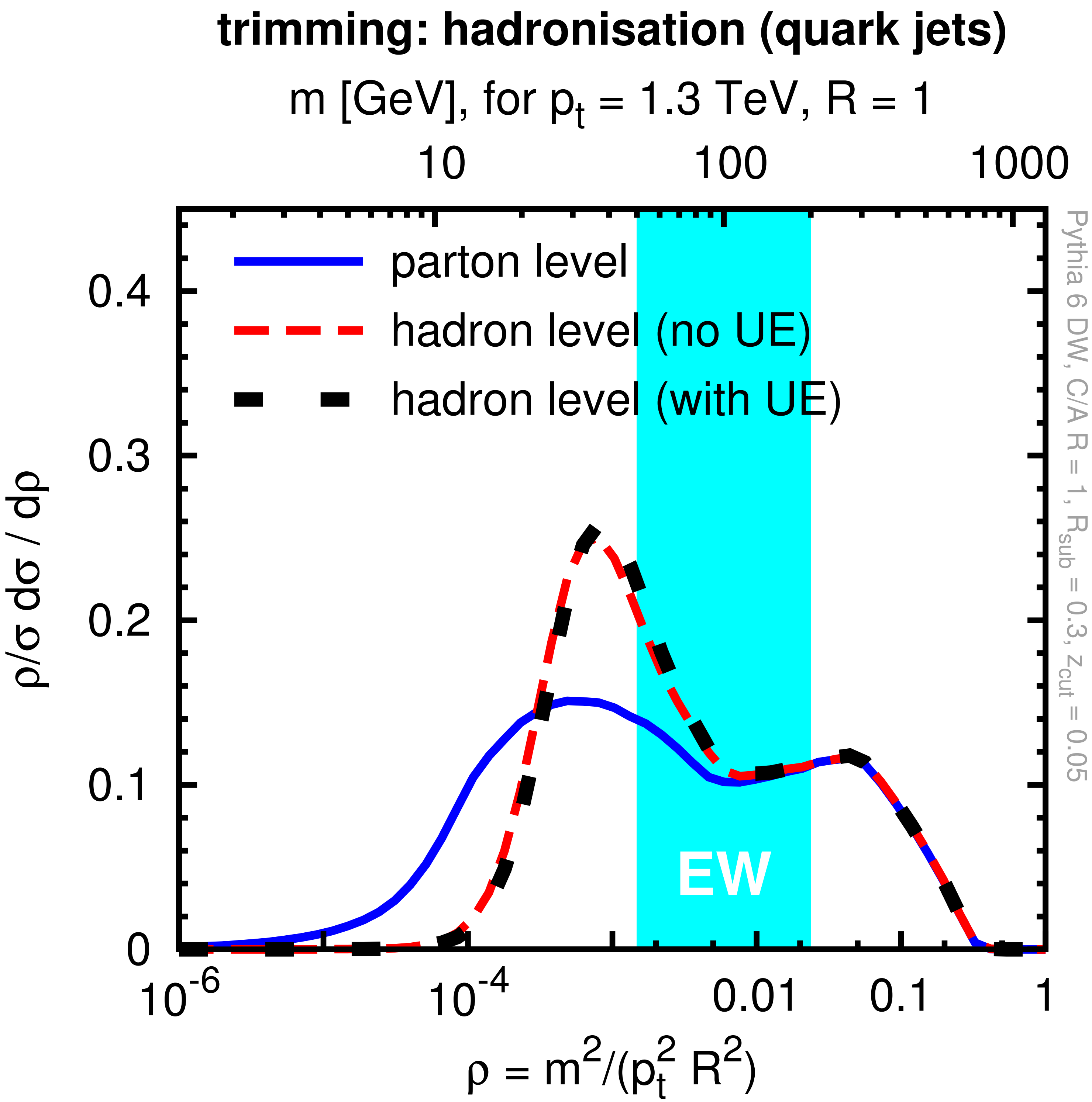
3000 GeV



trimming: hadronisation (quark jets)
m [GeV], for p_t = 2.0 TeV, R = 1
10
100
1000
parton level
hadron level (no UE)
hadron level (with UE)
EW
0.4
0.3
0.2
0.1
0
$\rho/\sigma\ d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Pythia 6 DW, C/A R = 1, R_{sub} = 0.3, z_{cut} = 0.05

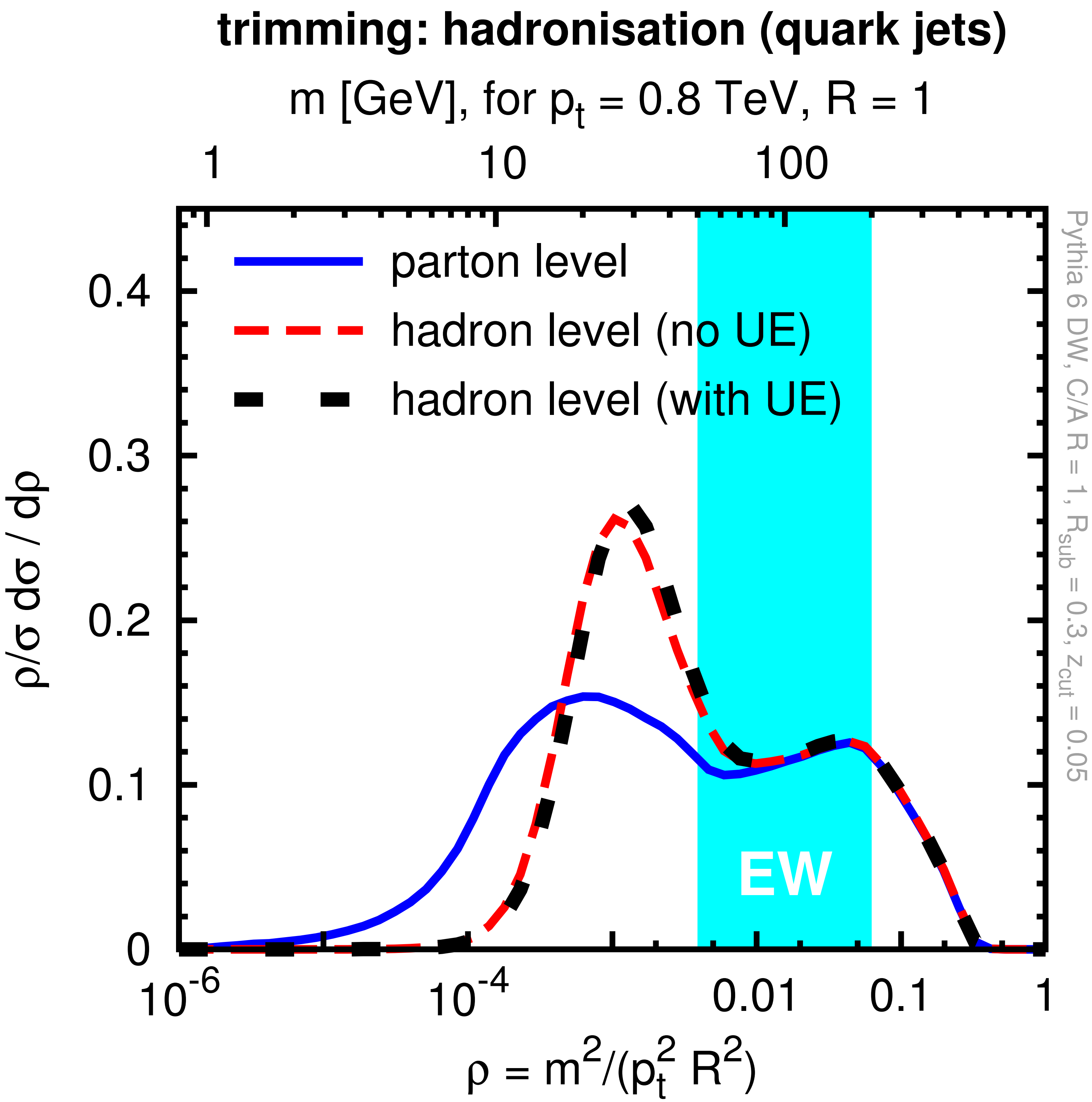
2000 GeV

Trimming v. jet p_t



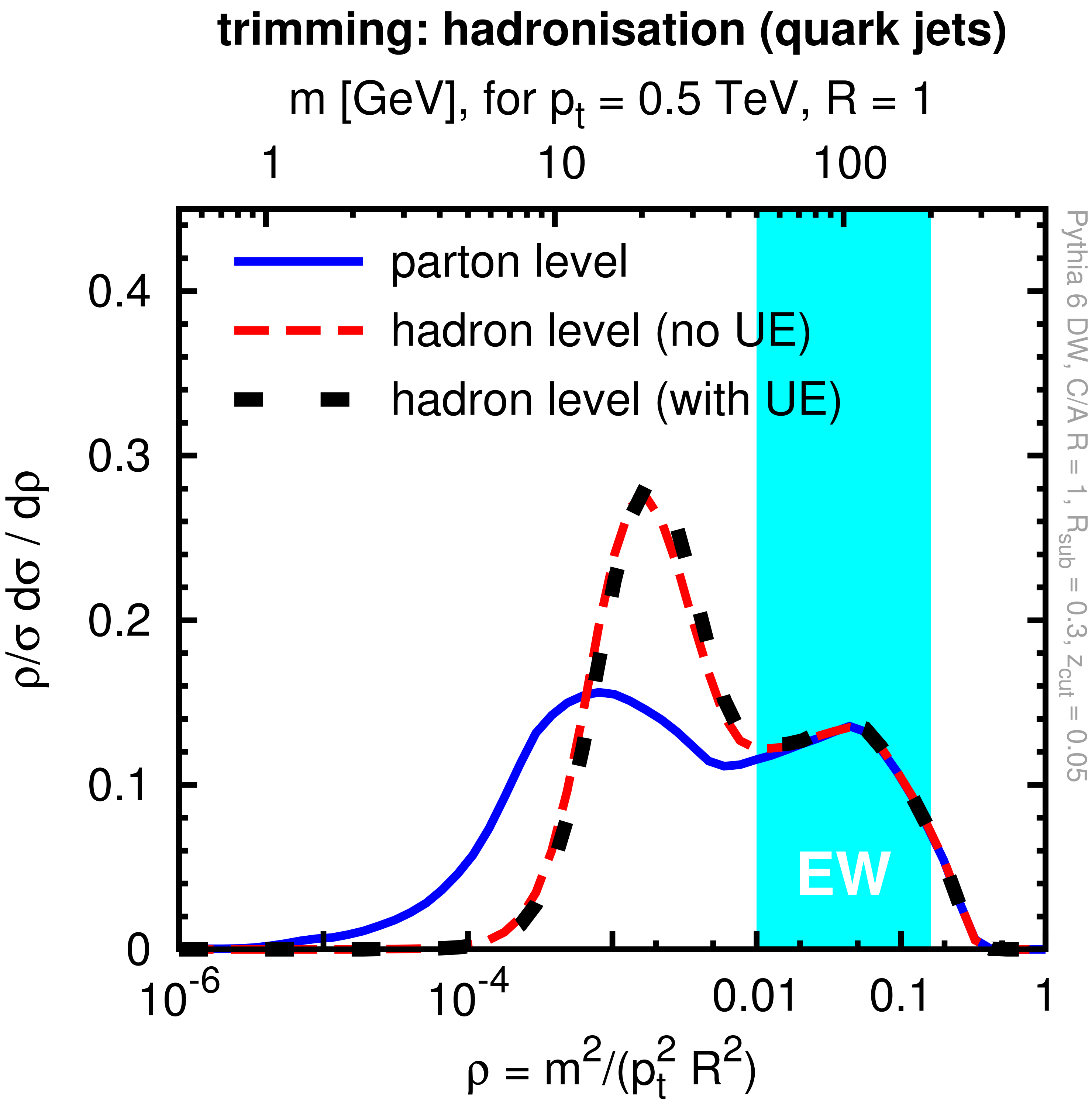
1300 GeV

Trimming v. jet p_t



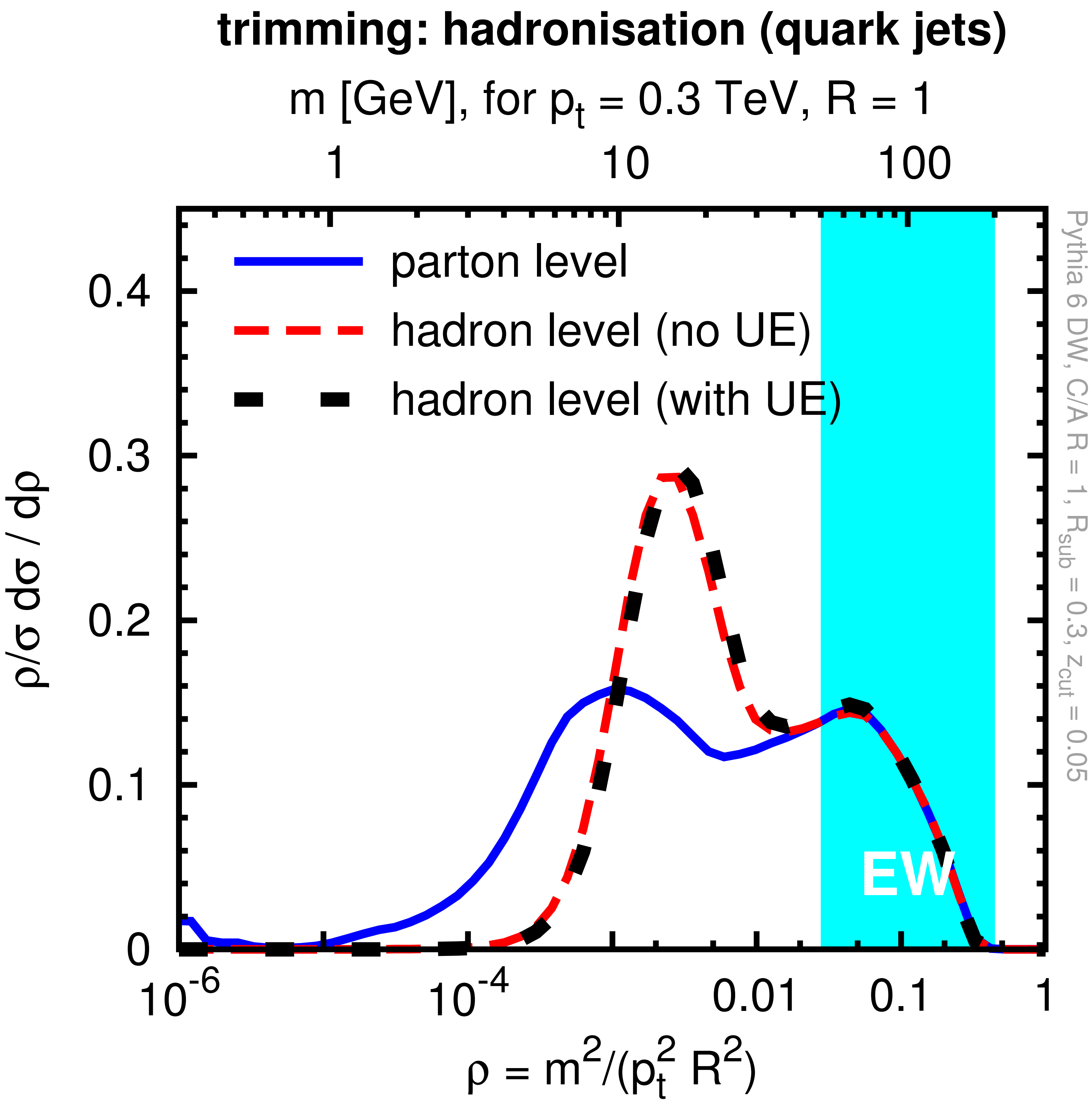
800 GeV

Trimming v. jet p_t



500 GeV

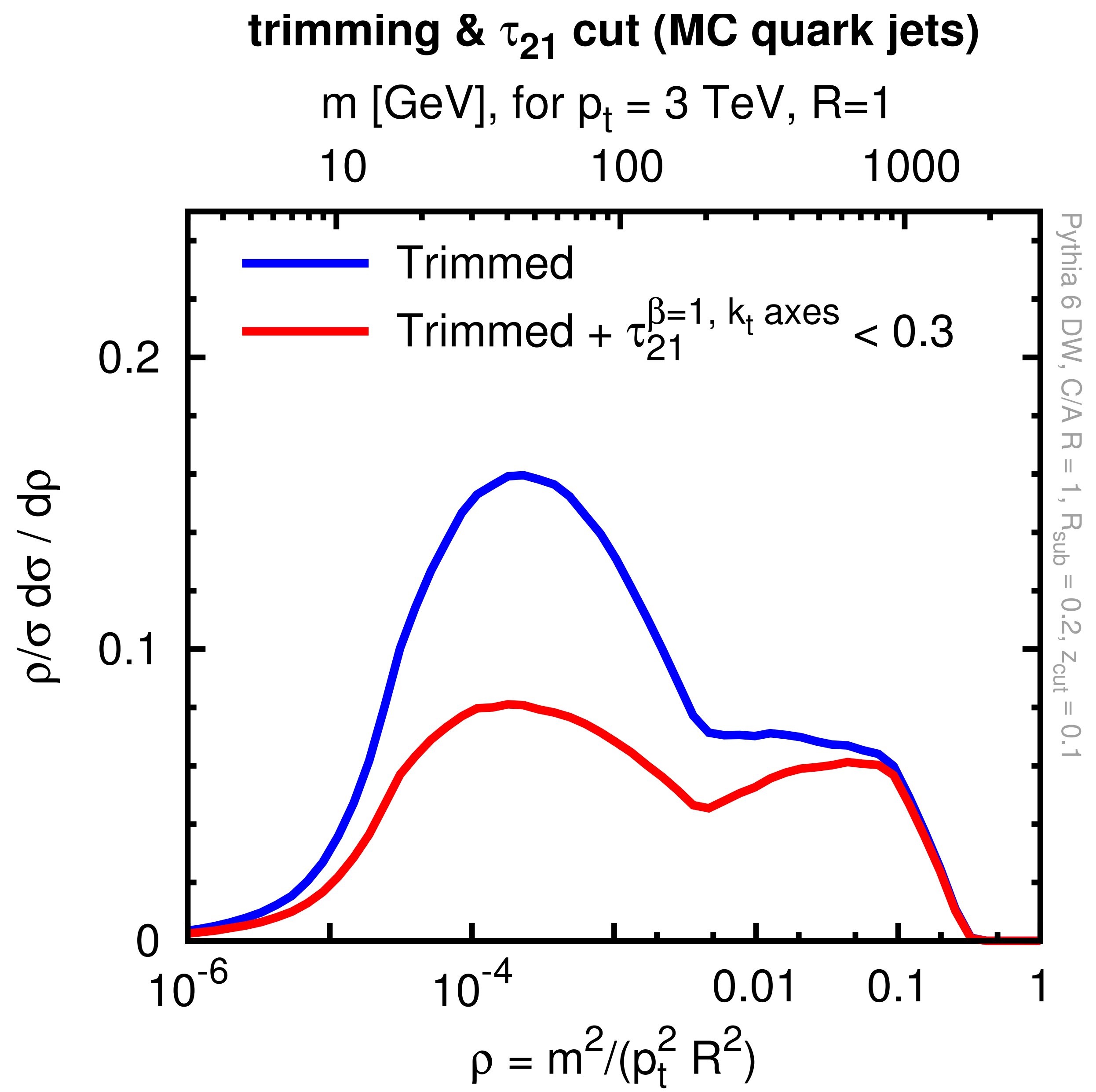
Trimming v. jet p_t



300 GeV

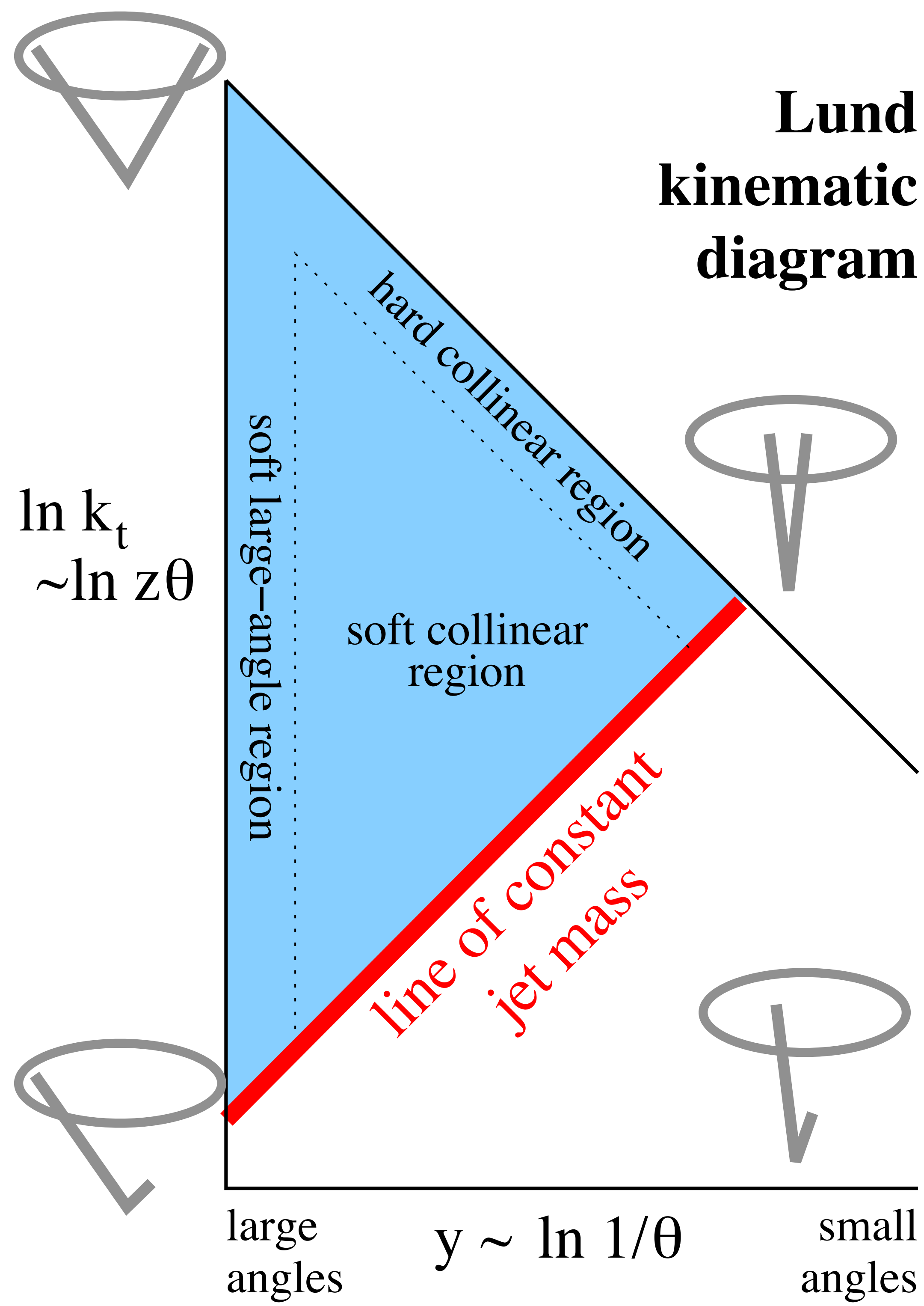
cuts on N-subjettiness, etc.?

- 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



kinematic plane and jet masses

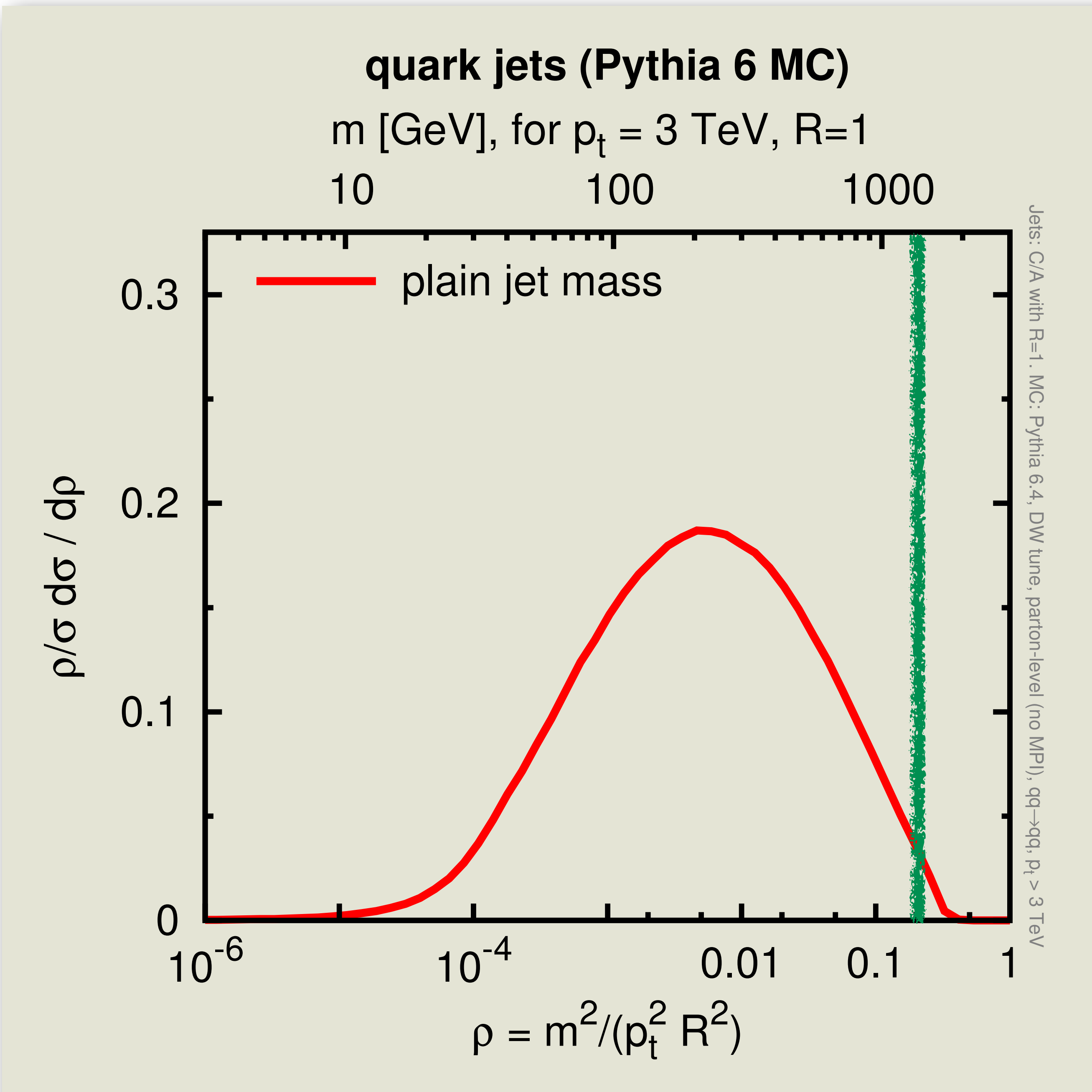
log of transverse momentum of emission wrt jet axis

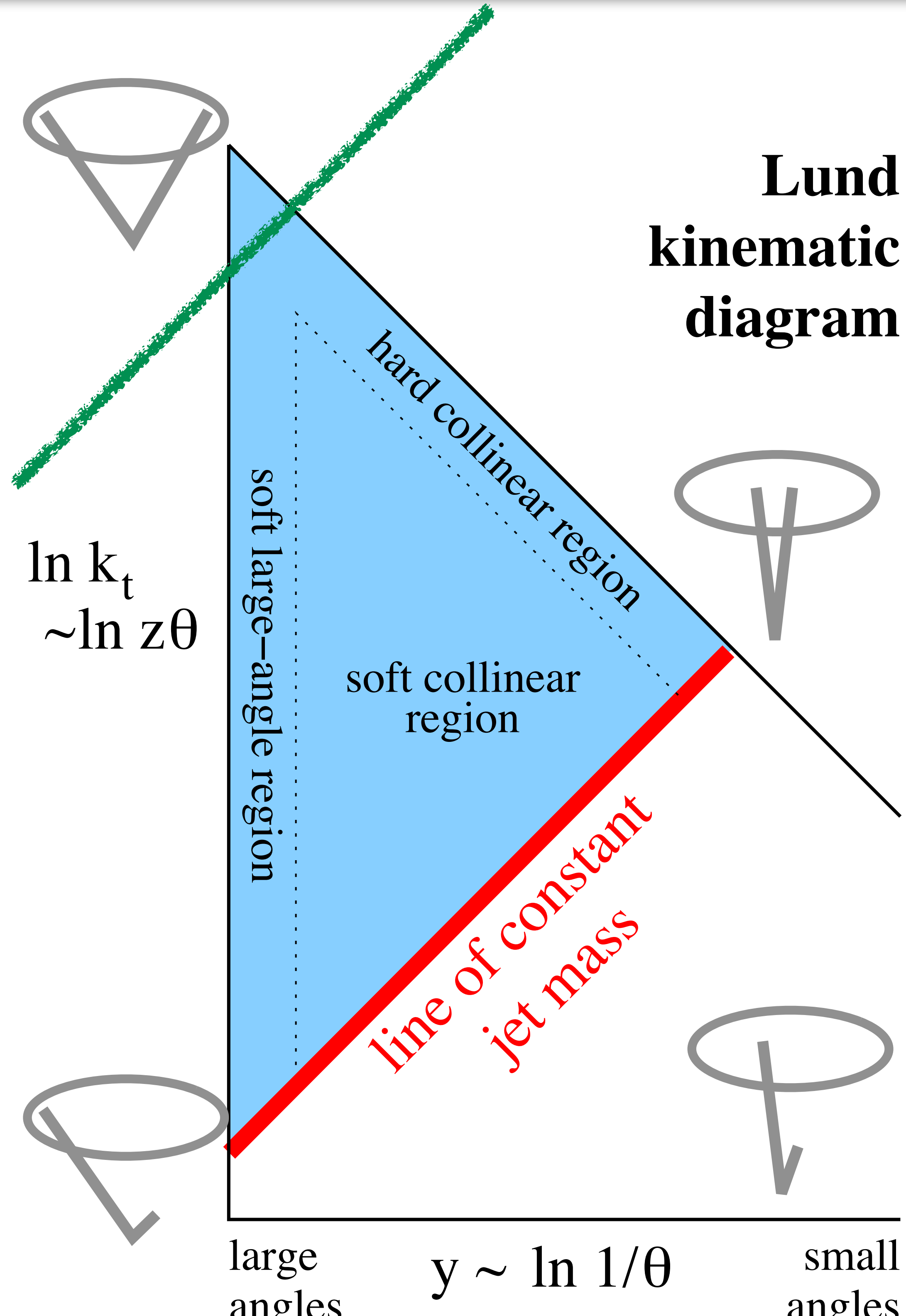


for QCD emissions

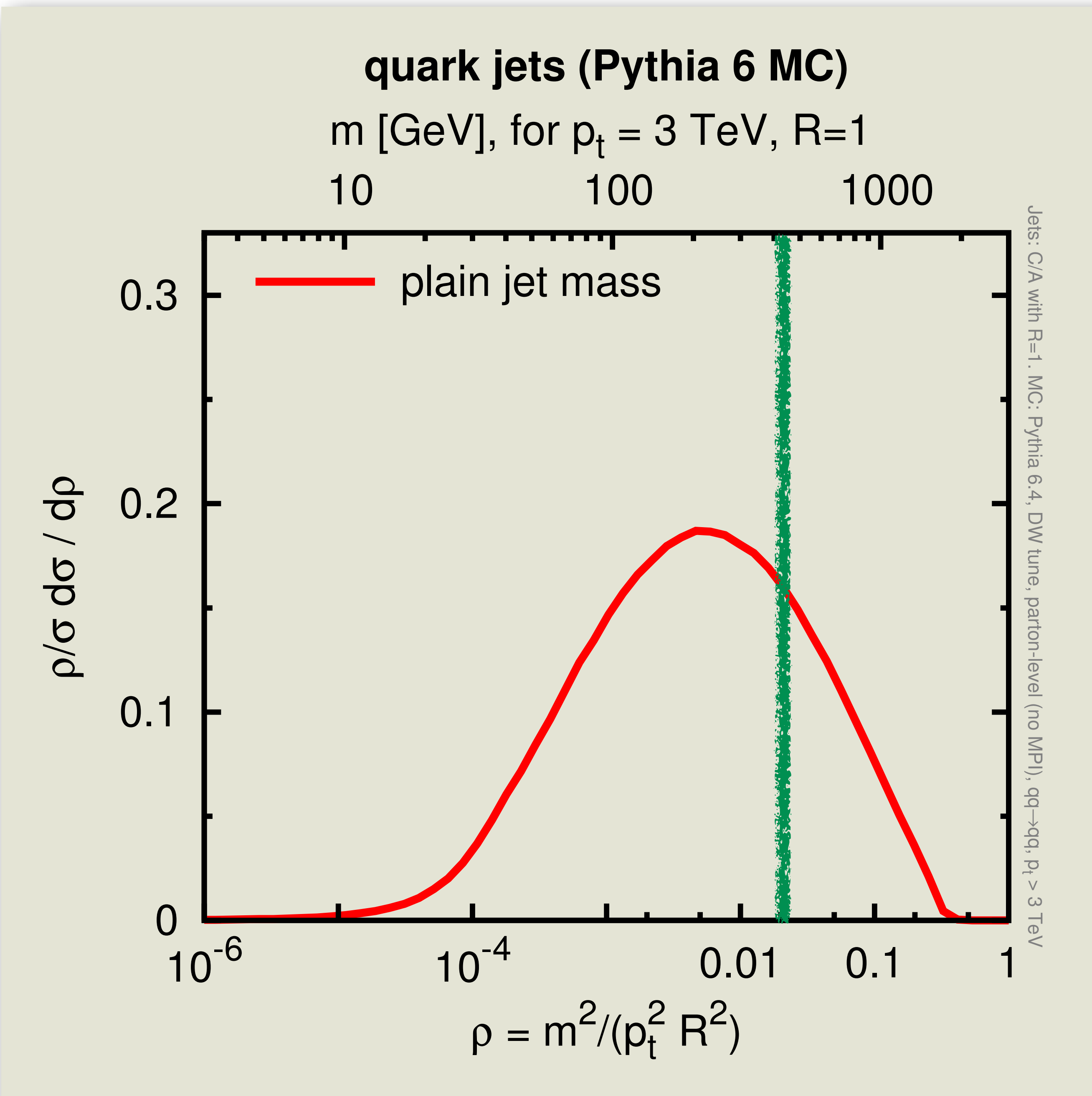
log of 1/(emission angle)

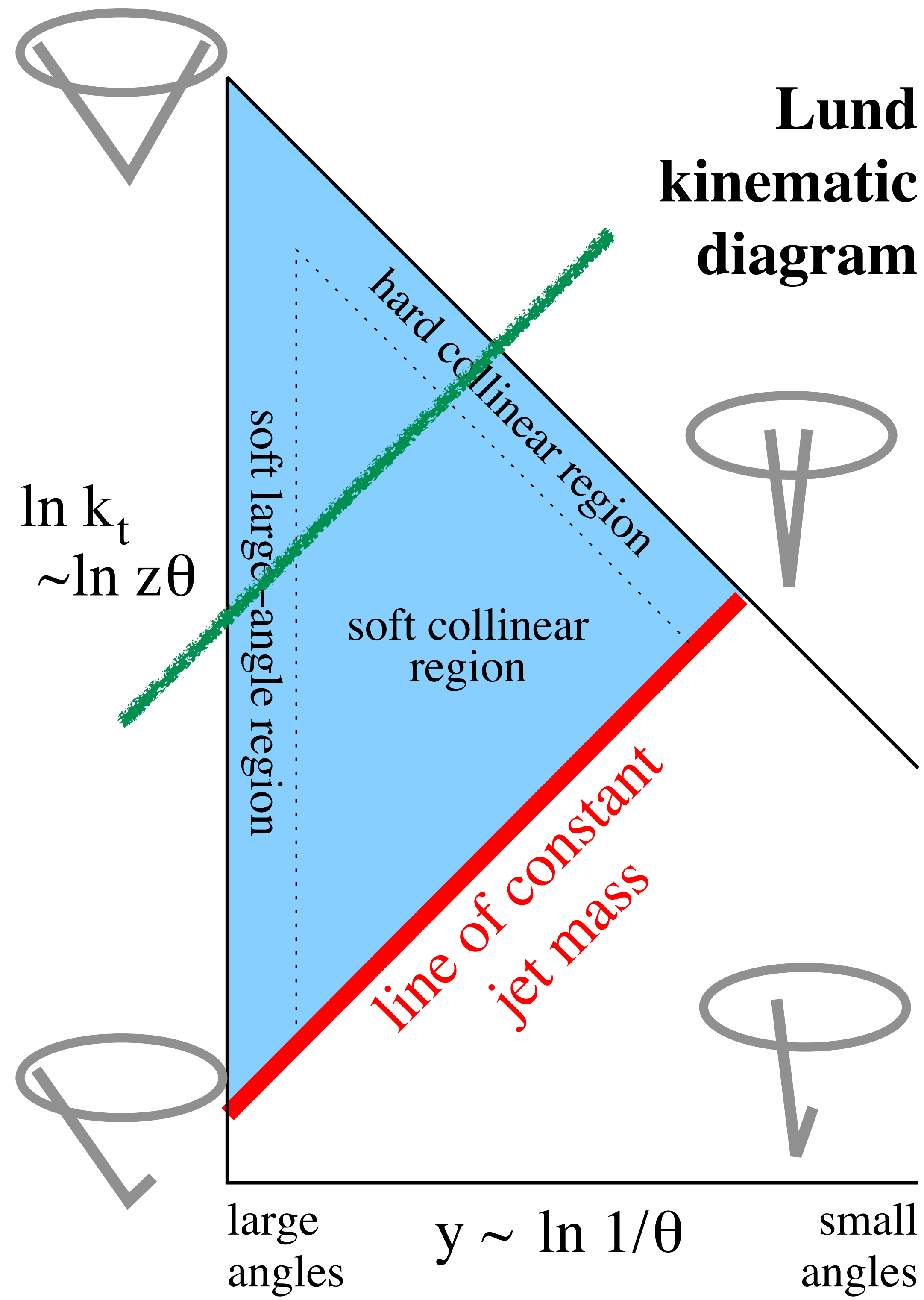
kinematic plane and jet masses





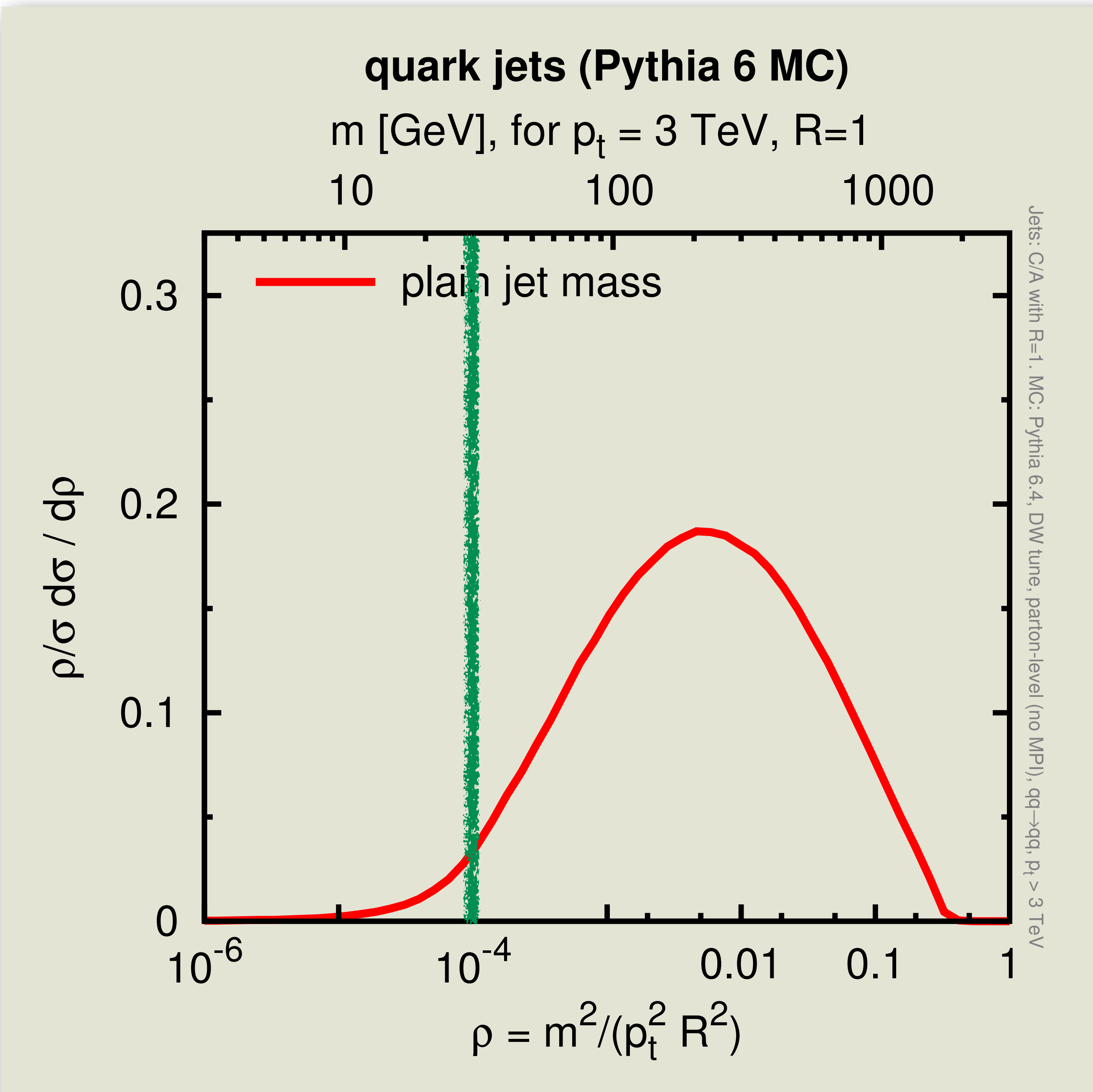
kinematic plane and jet masses

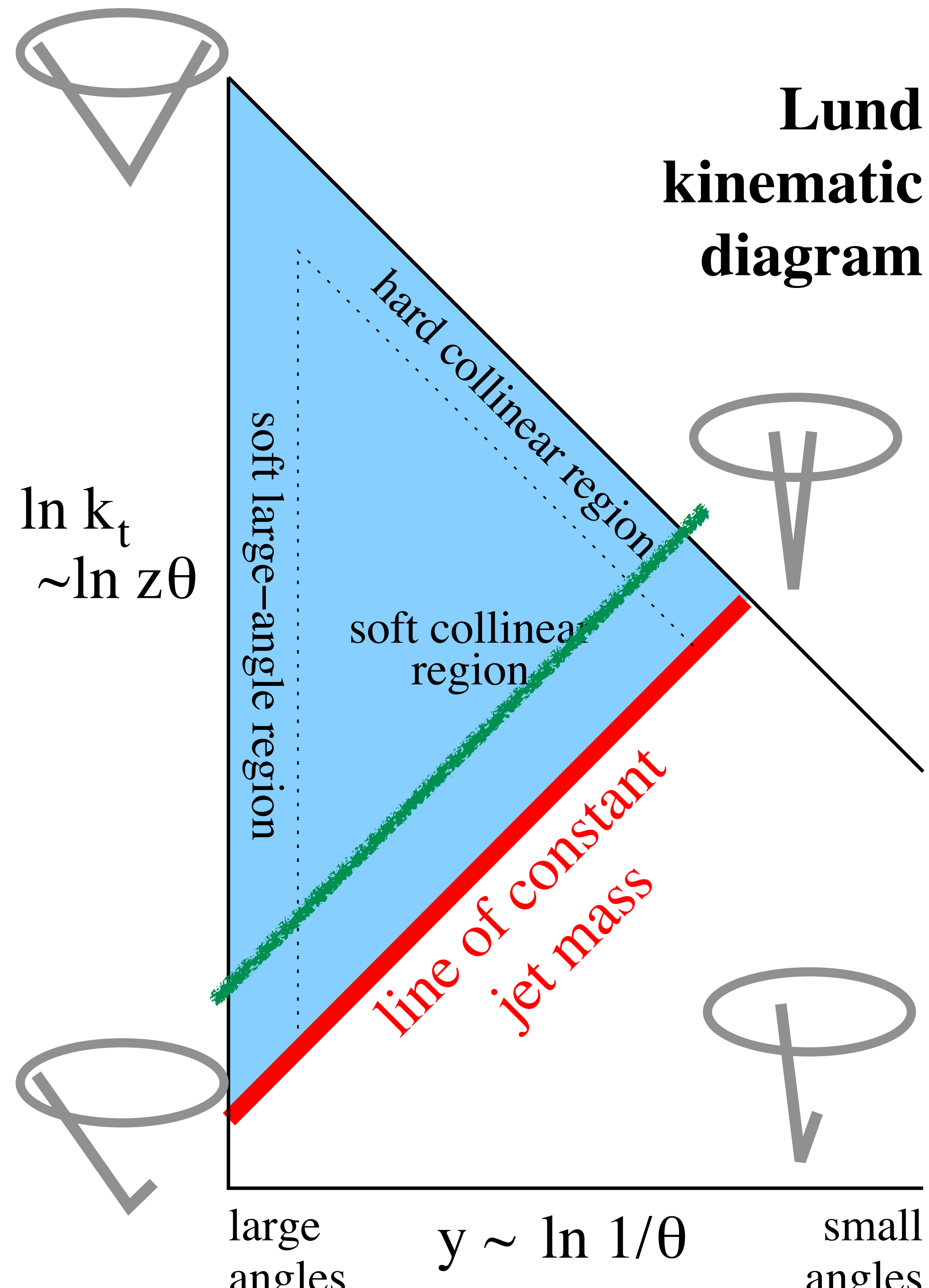




increase in phasespace

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

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

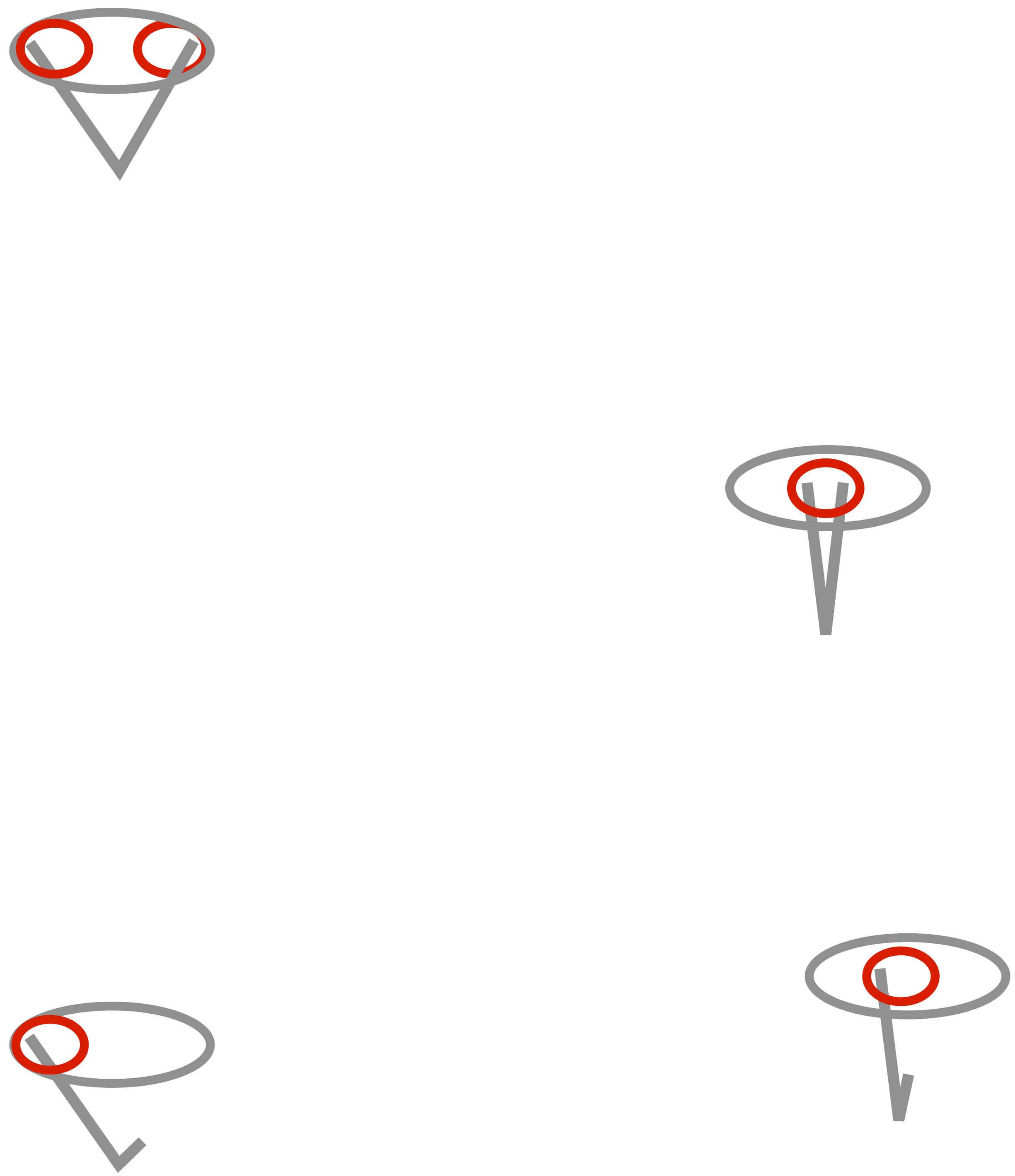
$R_{sub} < R$

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

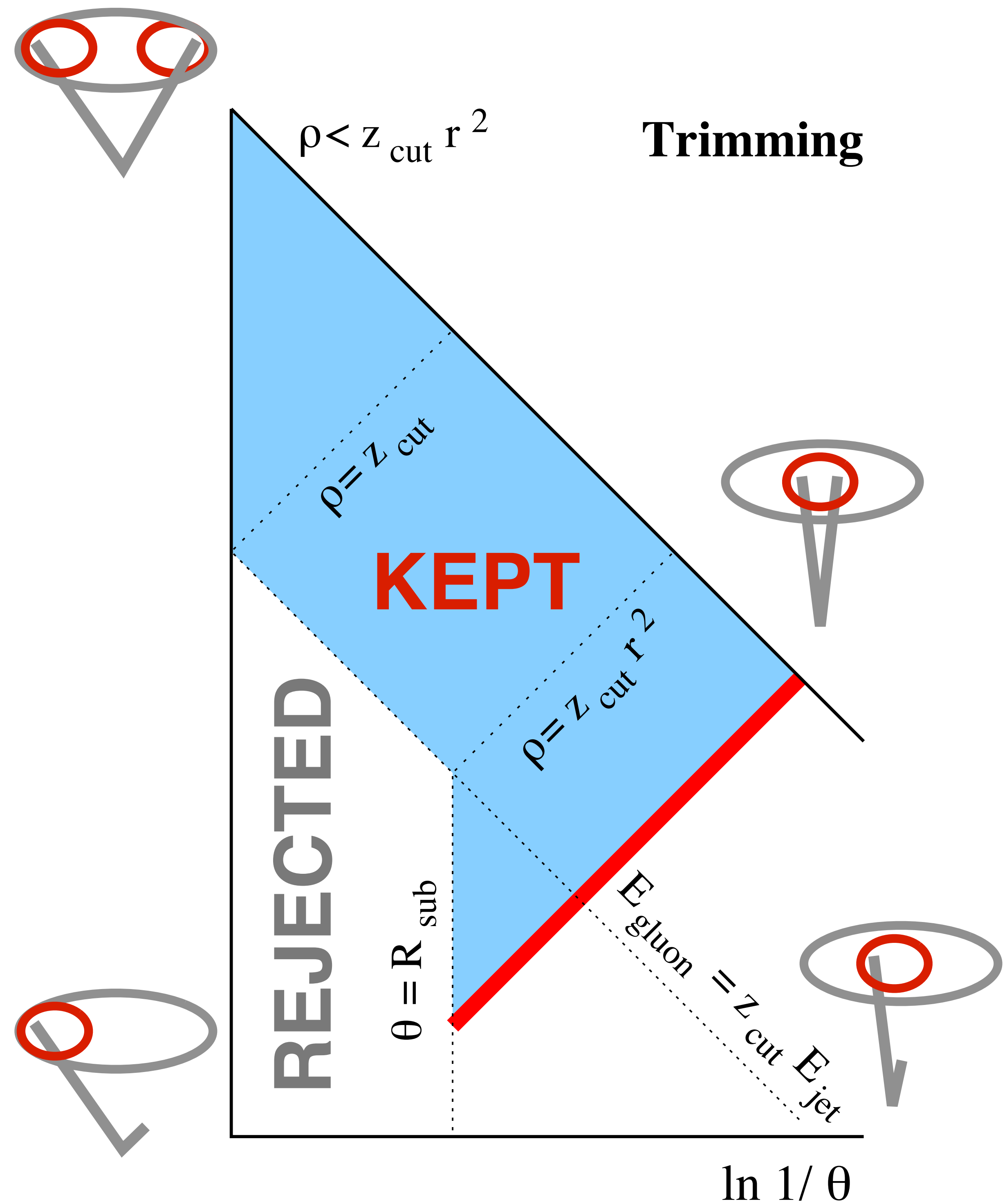
$R_{sub} < R$

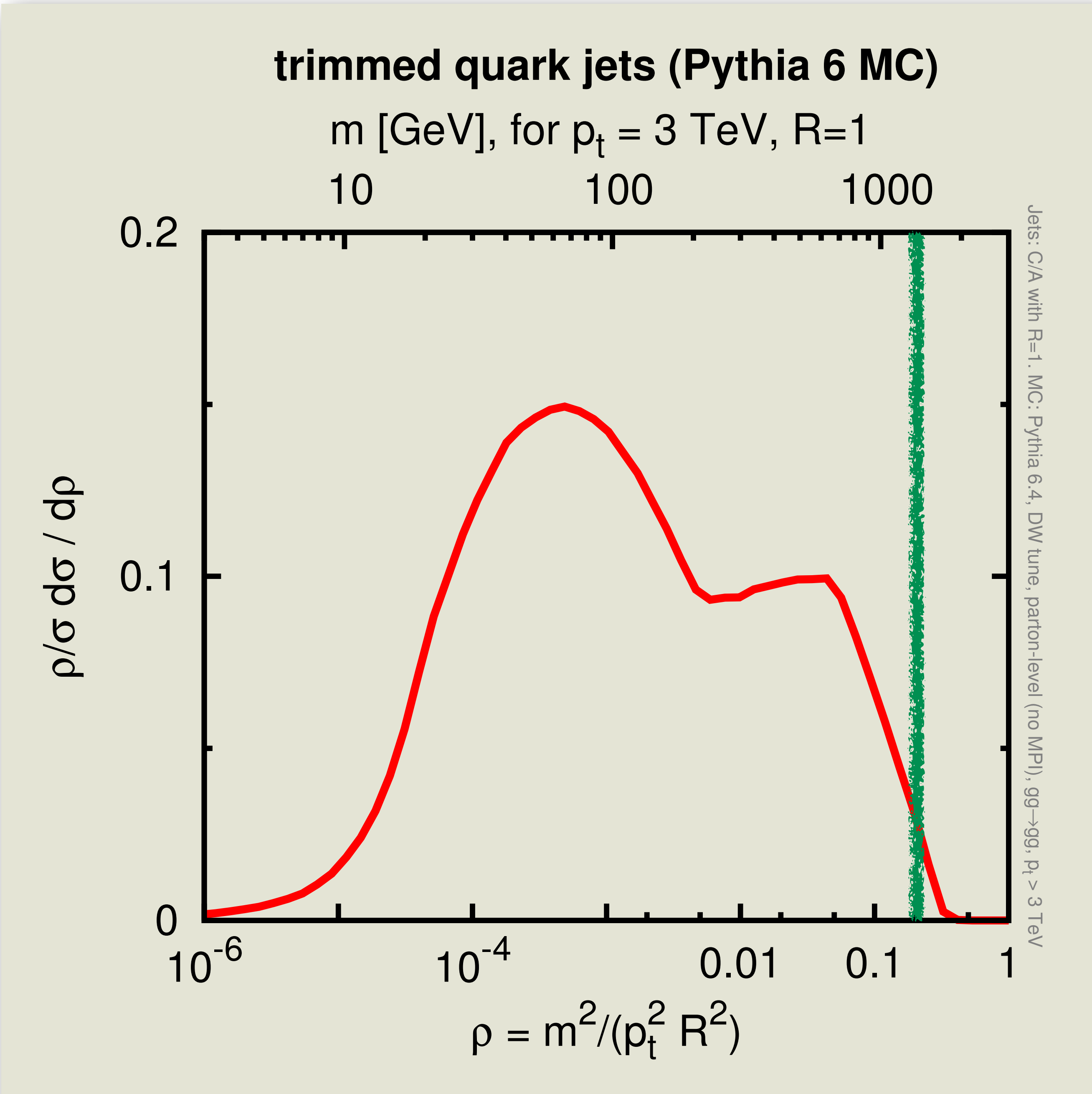
The subjets that satisfy the condition

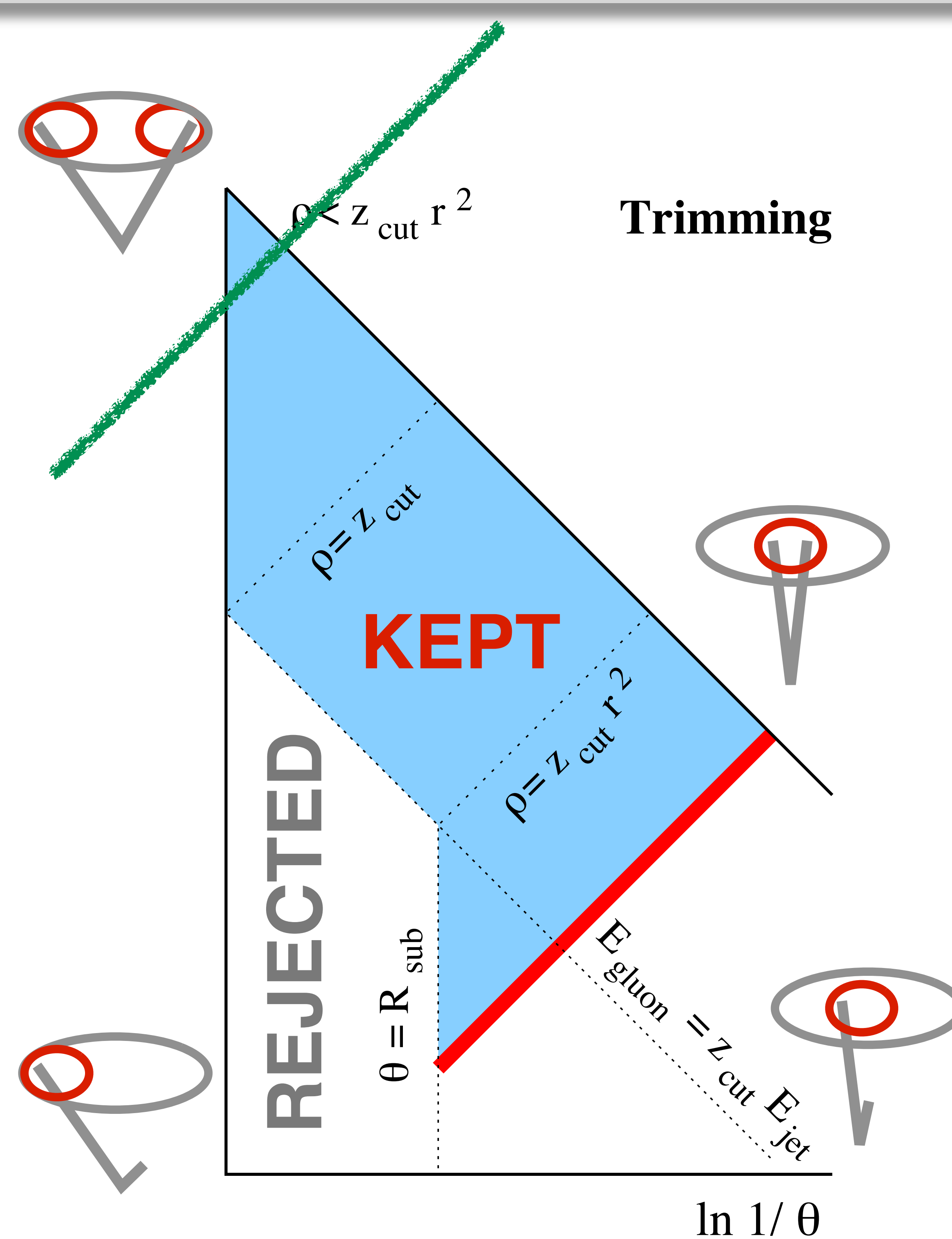
$p_t^{(subjet)} > z_{cut}\, p_t^{(jet)}$

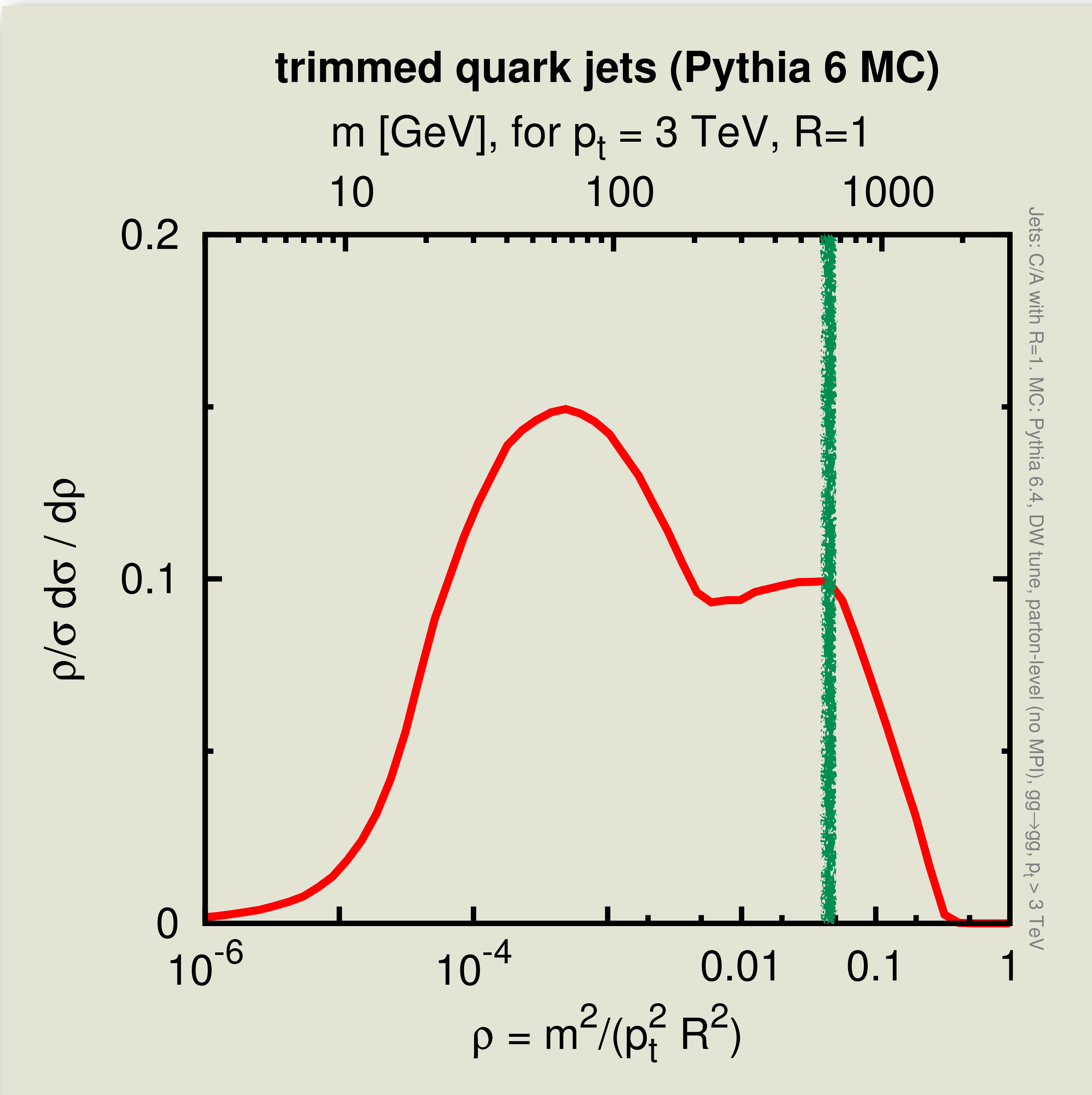
are kept and merged to form the trimmed jet.

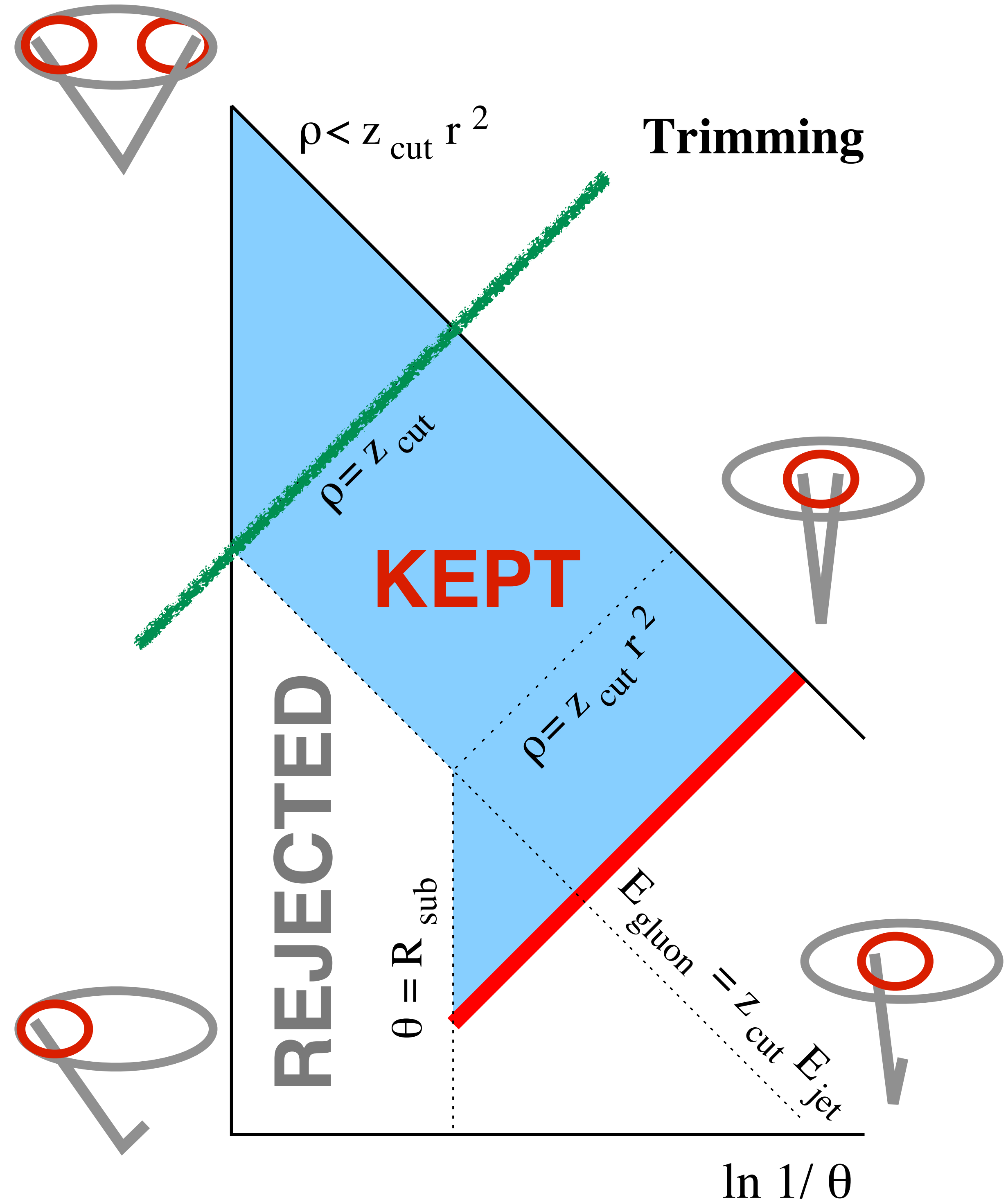
Krohn, Thaler & Wang '09

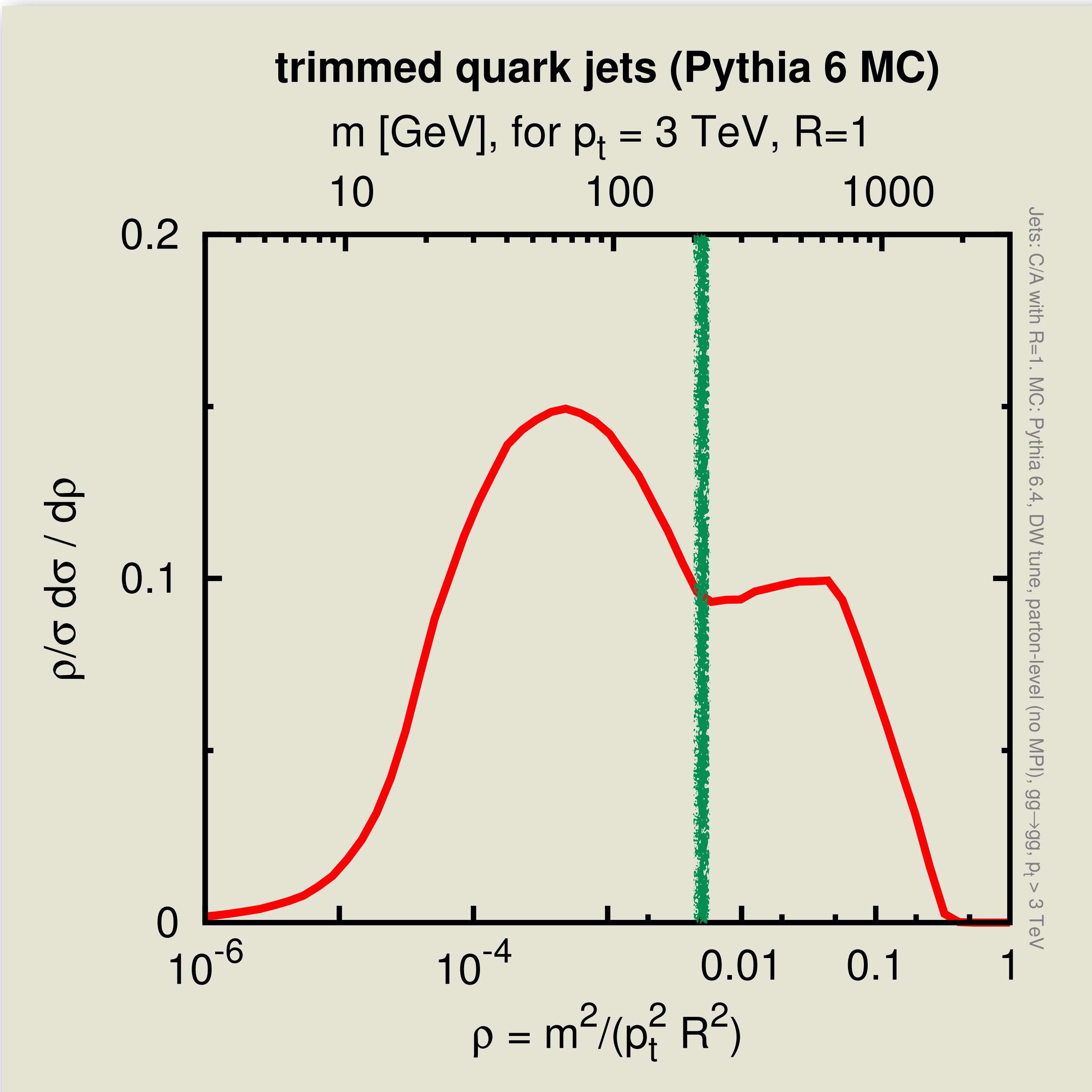


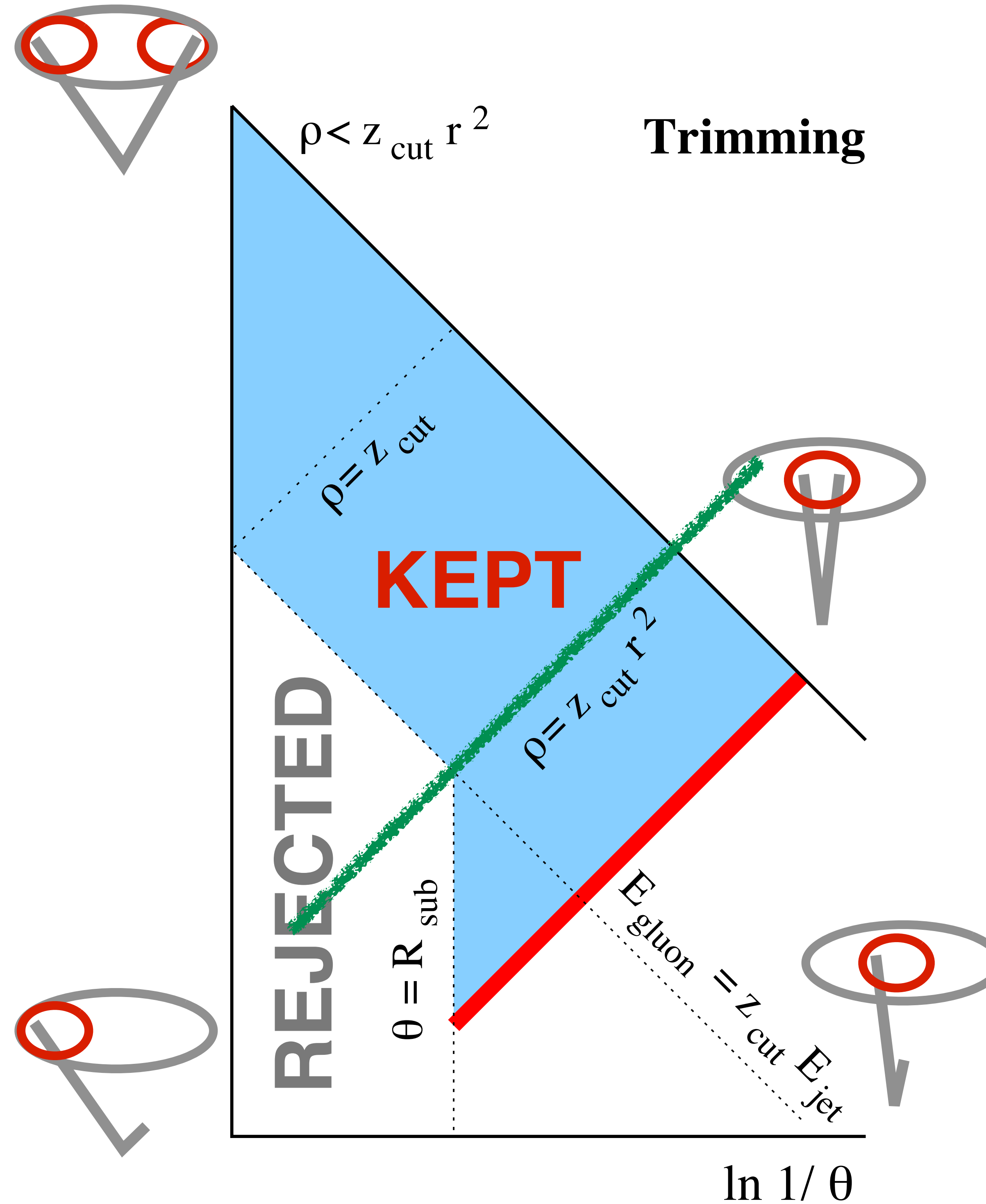

trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\ d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV


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

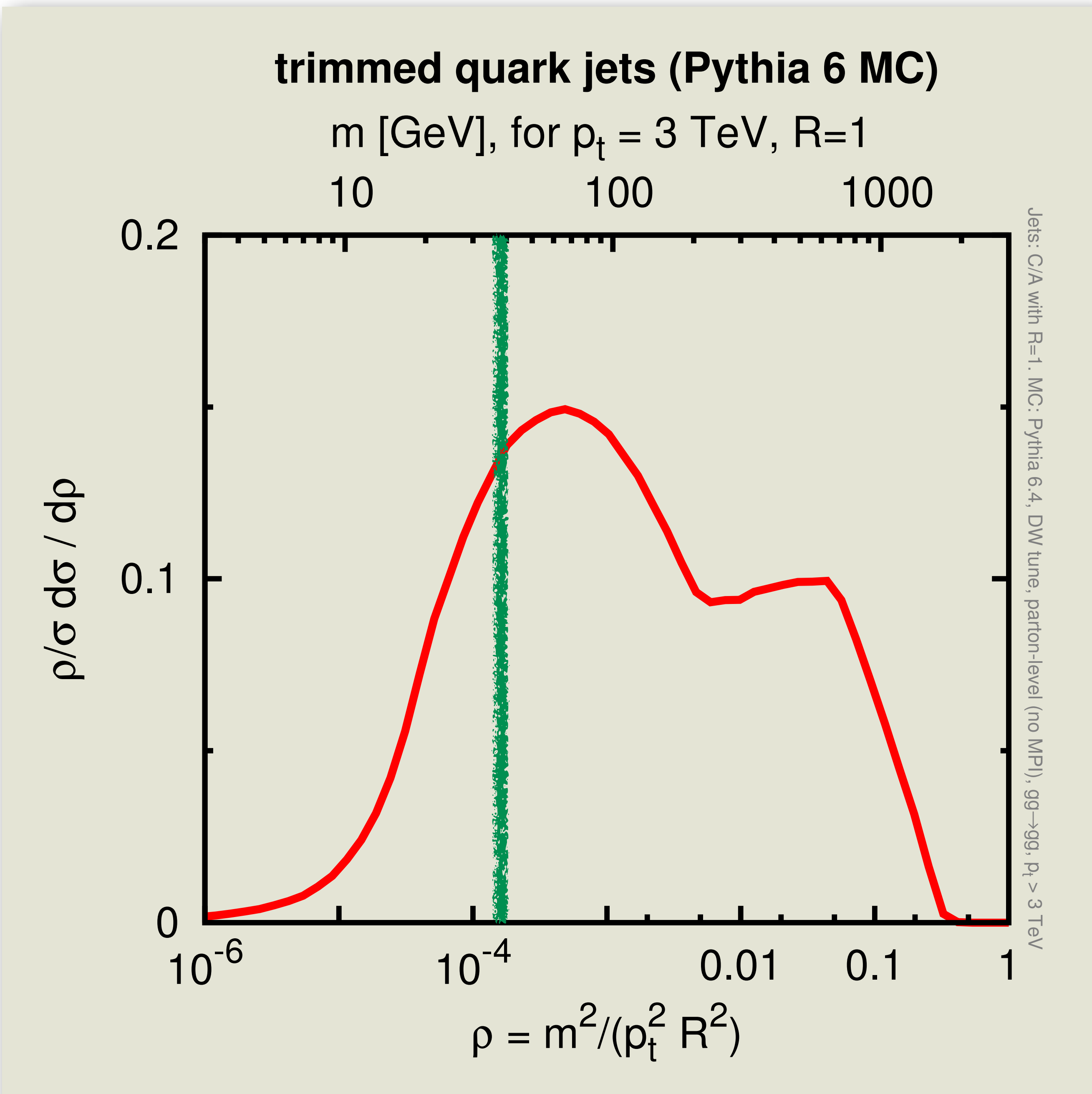

trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\ d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV


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

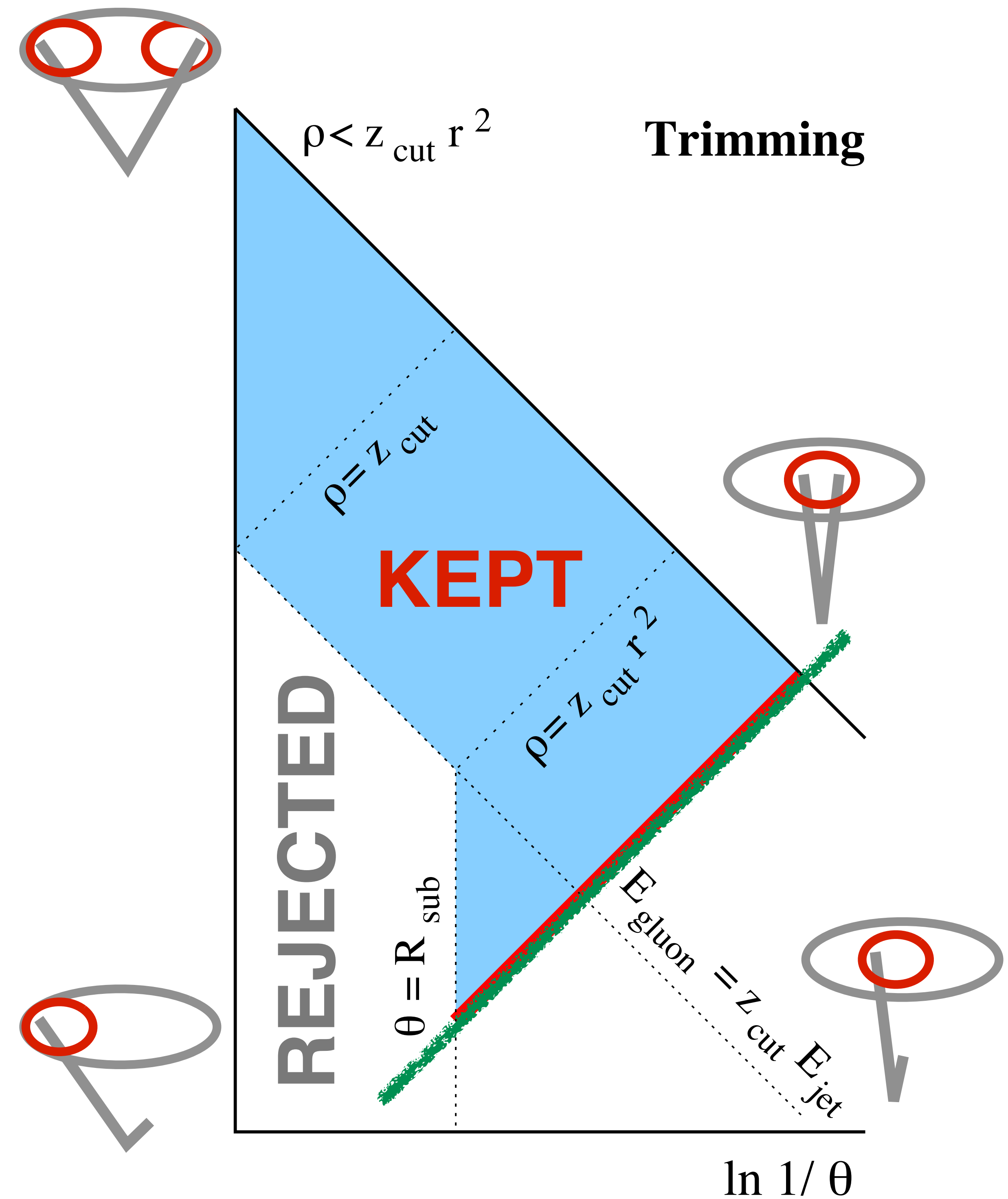

trimmed quark jets (Pythia 6 MC)
m [GeV], for p_t = 3 TeV, R=1
10
100
1000
0.2
0.1
0
$\rho/\sigma\ d\sigma / d\rho$
10^{-6}
10^{-4}
0.01
0.1
1
$\rho = m^2/(p_t^2 R^2)$
Jets: C/A with R=1. MC: Pythia 6.4, DW tune, parton-level (no MPI), gg→gg, p_t > 3 TeV


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

Trimming



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

What about actual calculations of the taggers?

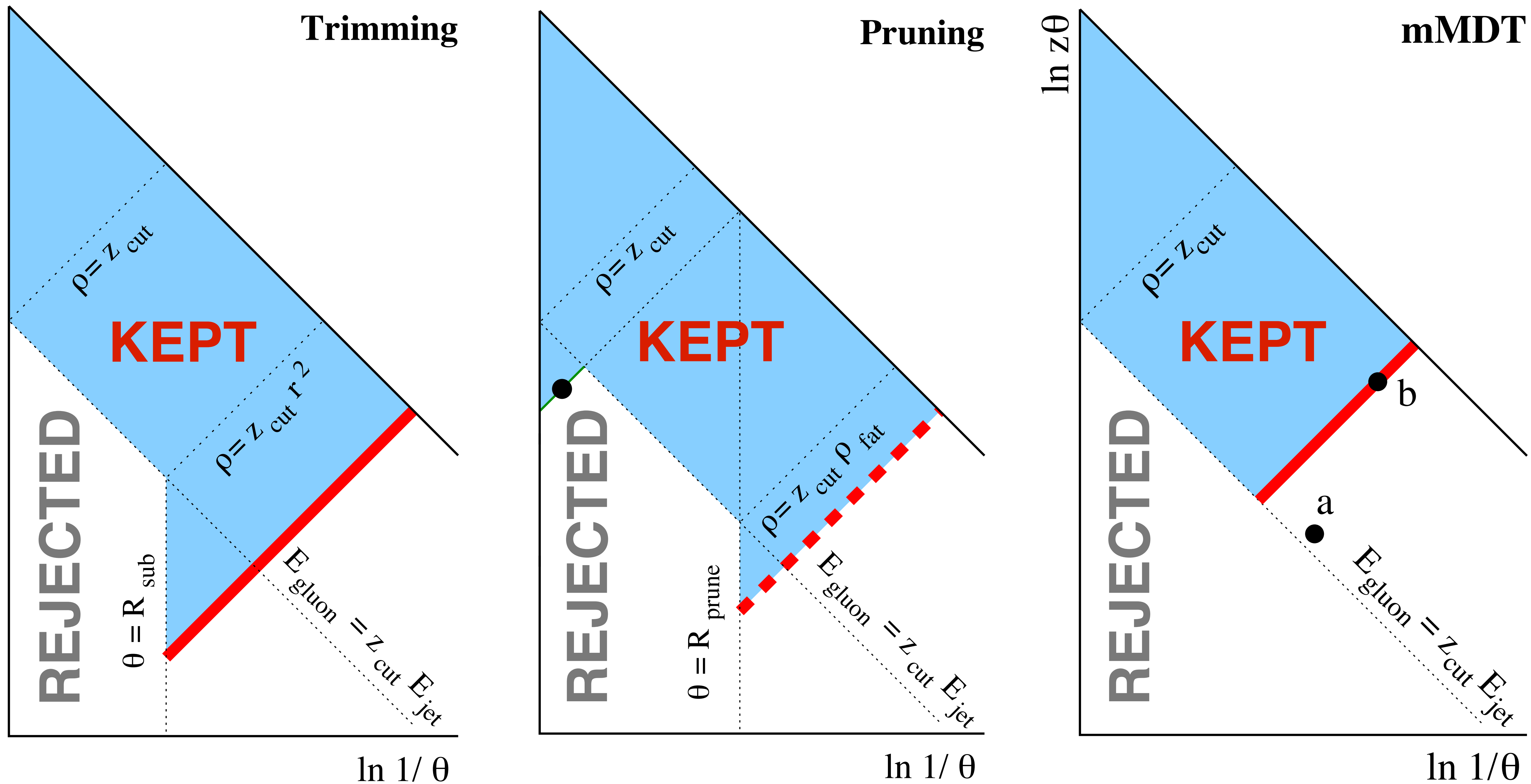
Simpler   More complex

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!)

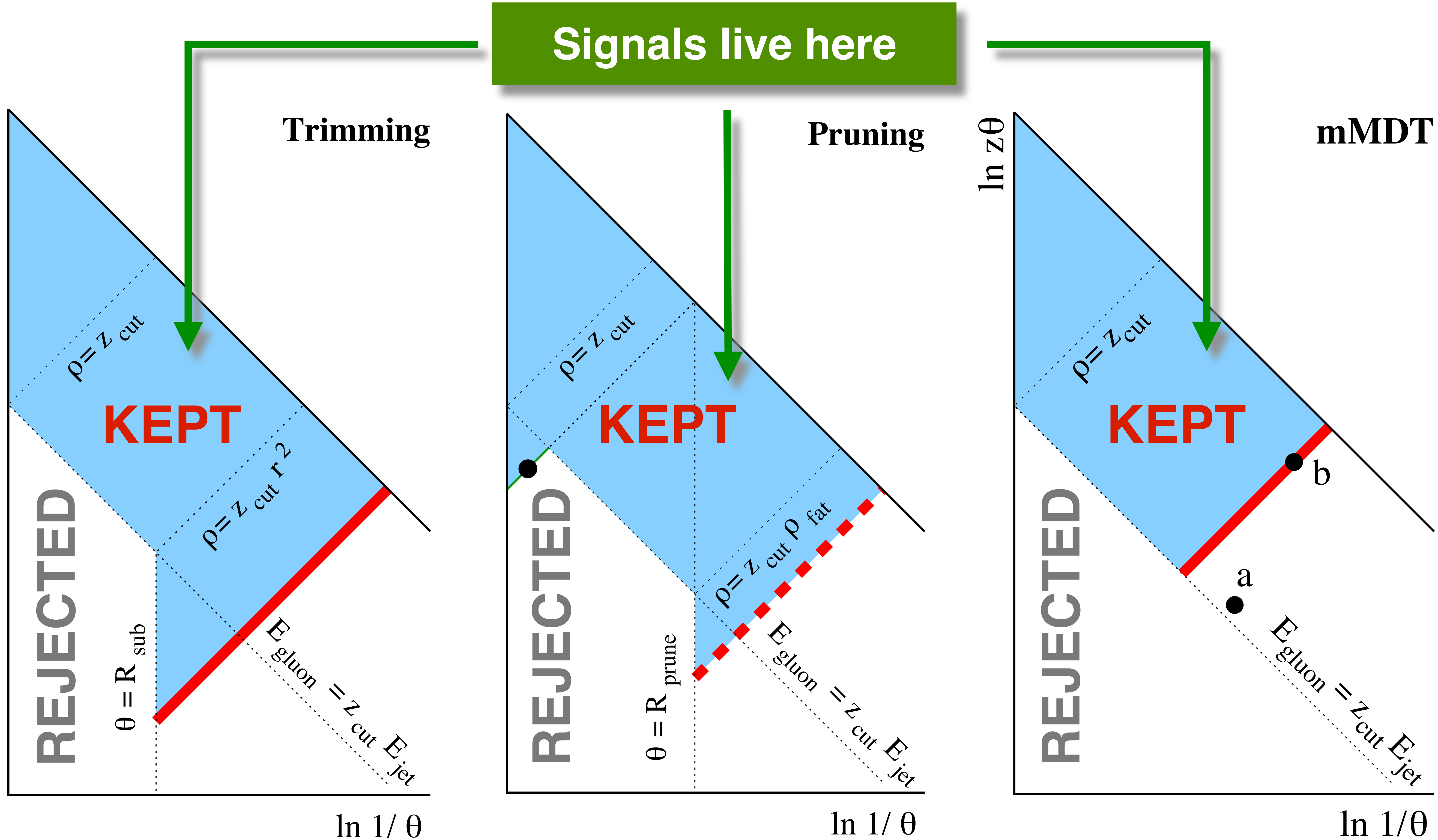
mMDT > plain mass ~ trimming > Y-pruning > full pruning ≫ ~~MDT~~

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

Kinematic regions for different taggers



Kinematic regions for different taggers



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

NEW

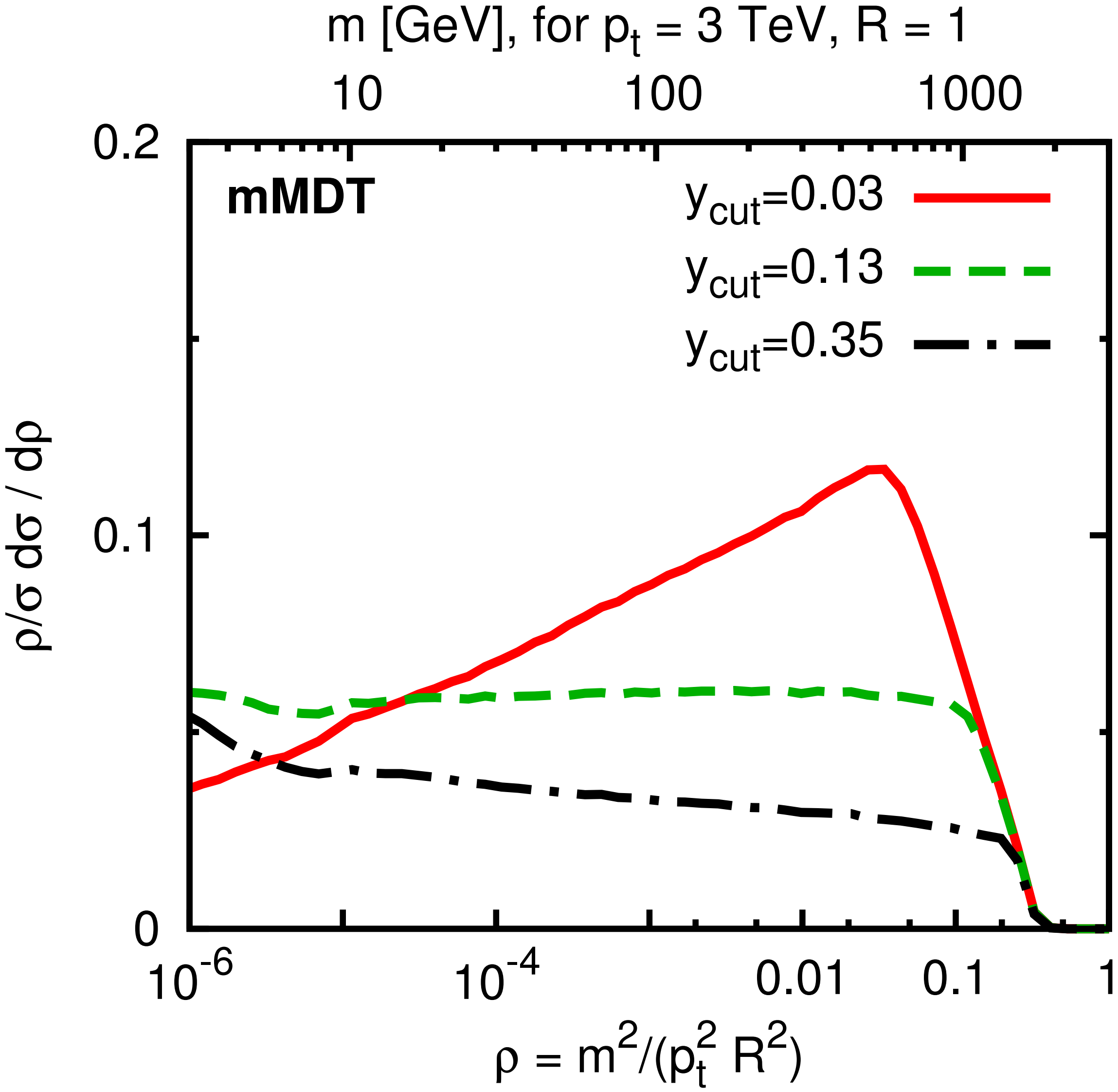
Special: only single logarithms (L = ln ρ)

→ more accurately calculable

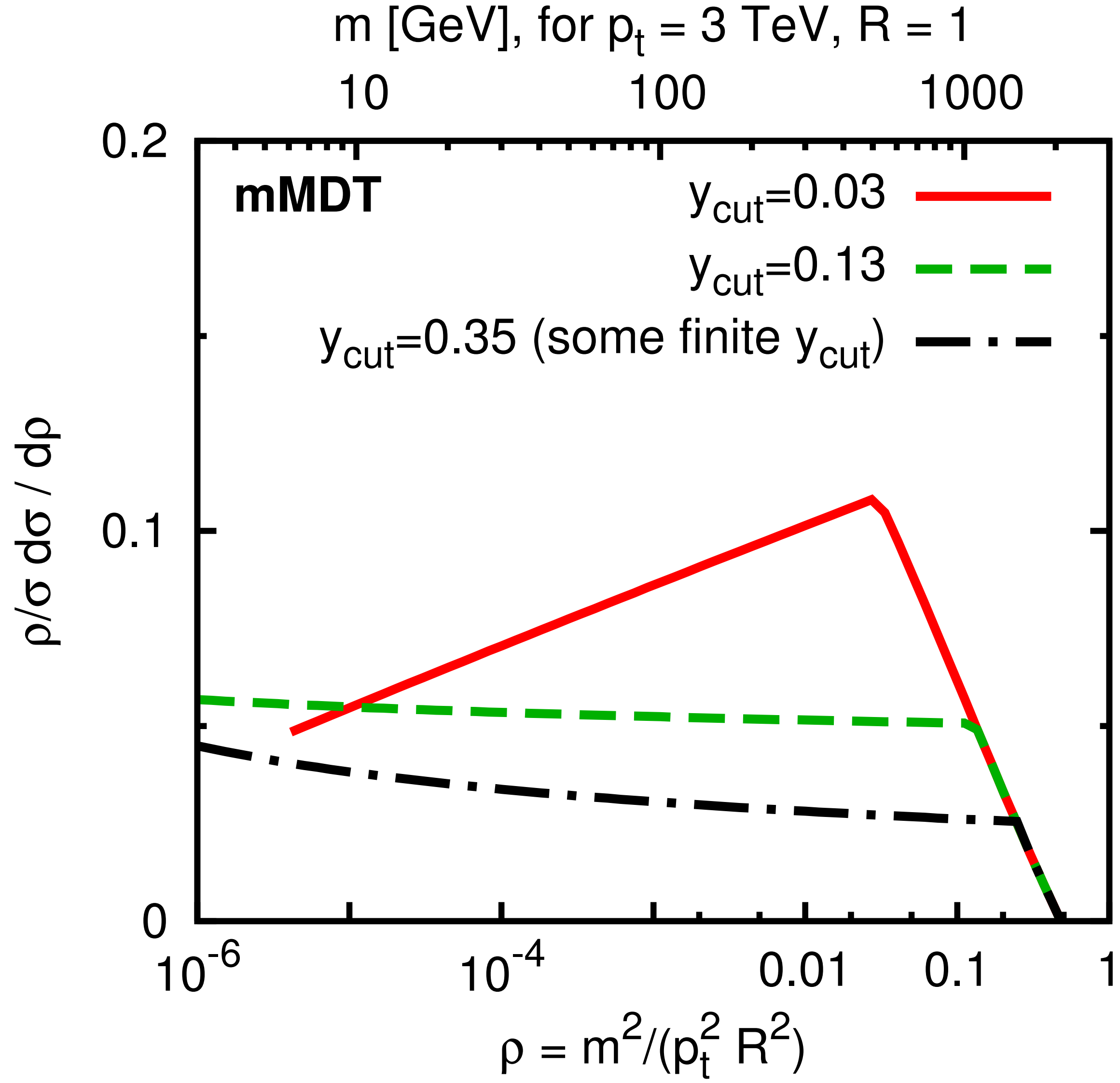
Special: better exploits signal/bkgd differences

Modified Mass Drop Tagger

Monte Carlo



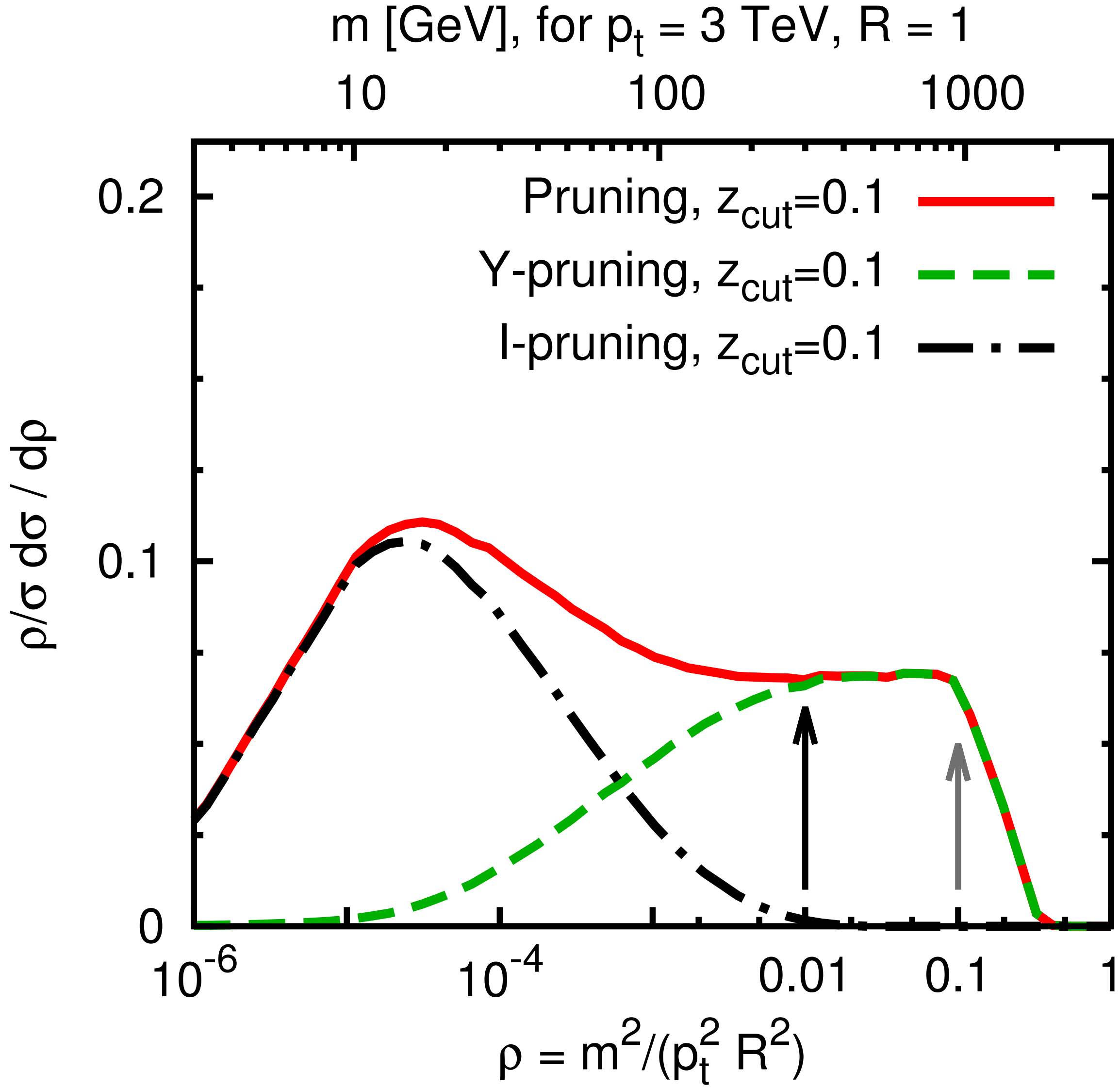
Analytic



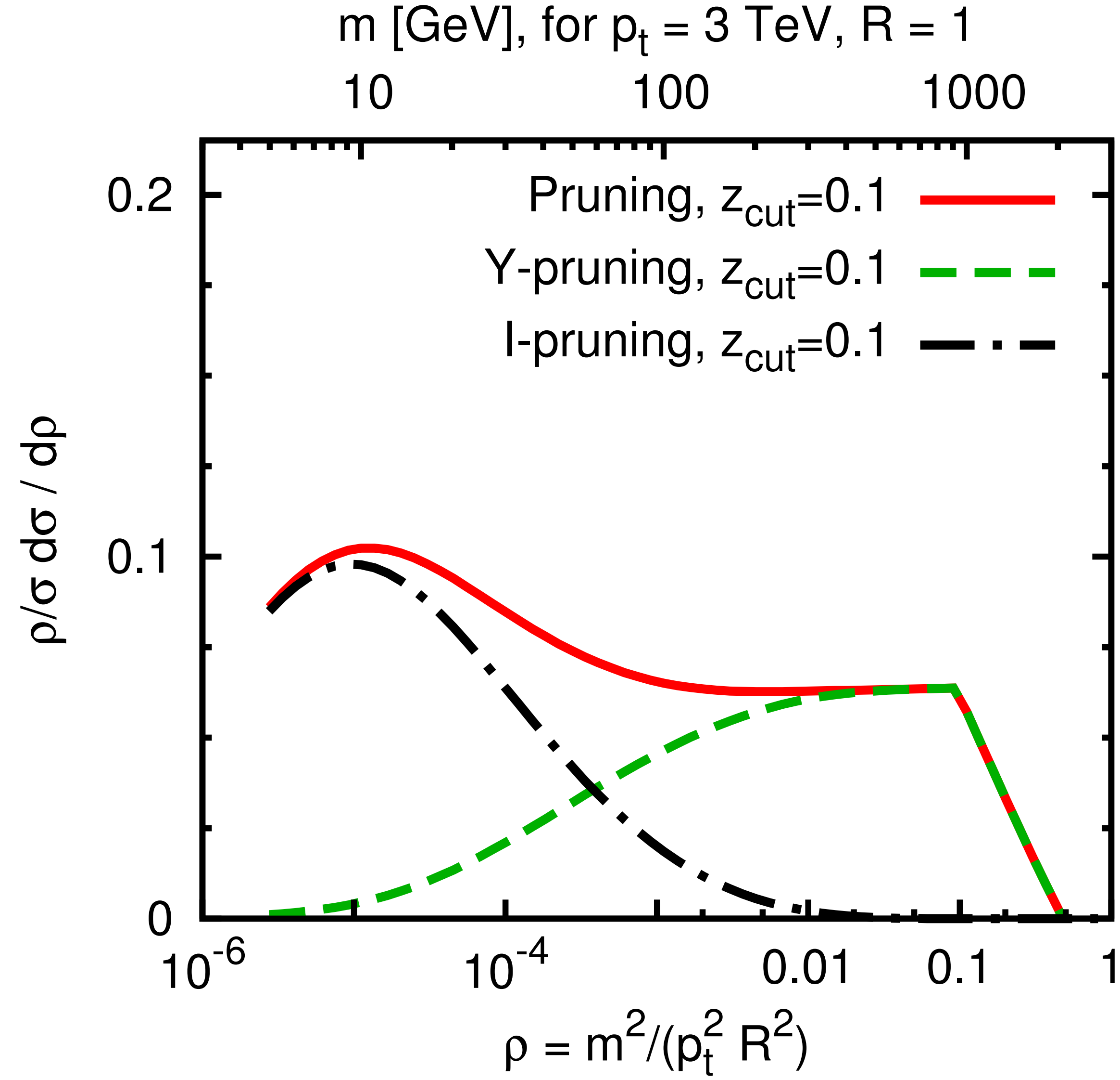
[mMDT is closest we have to a scale-invariant tagger, though exact behaviour depends on q/g fractions]

Pruning

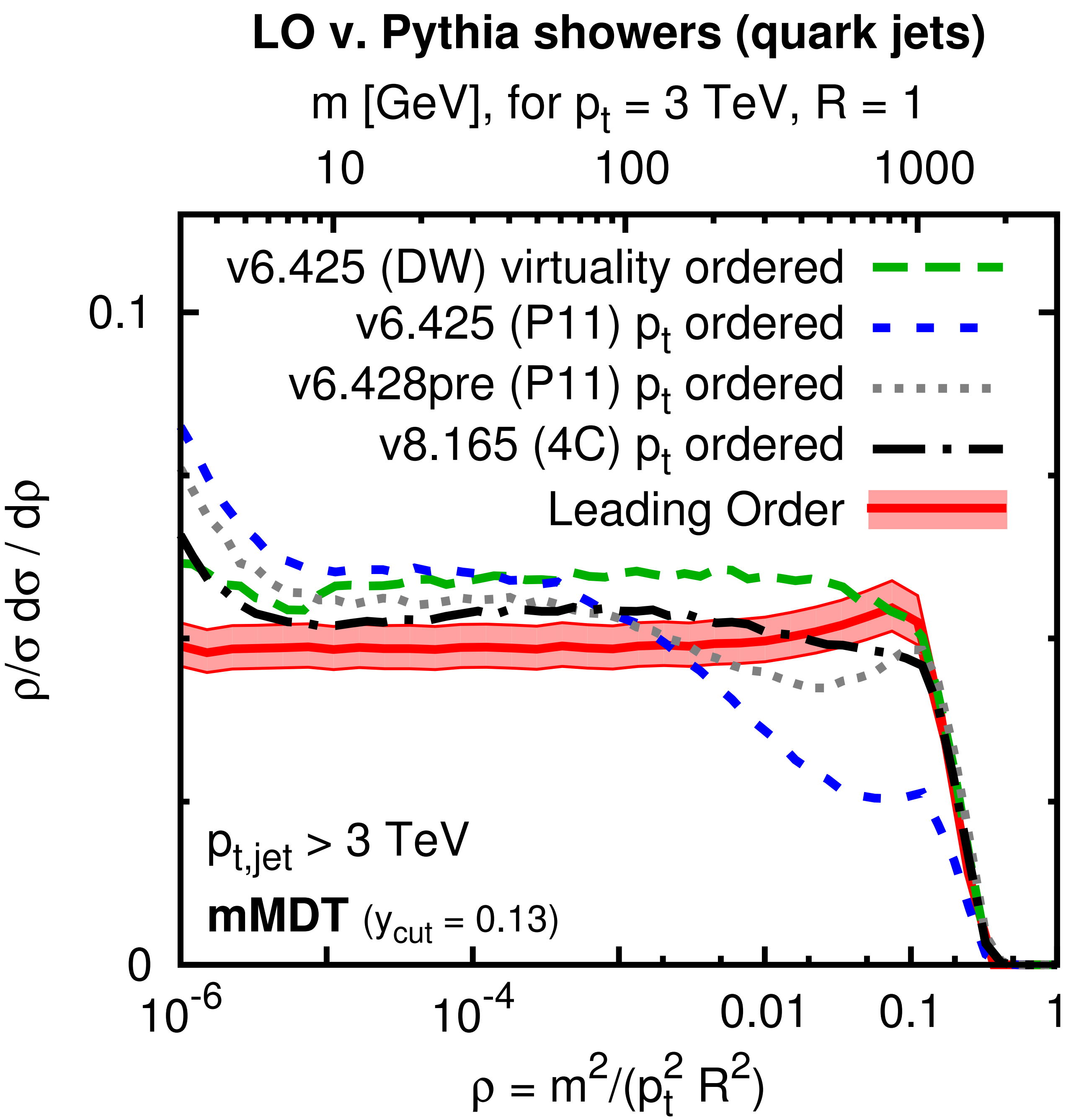
Monte Carlo

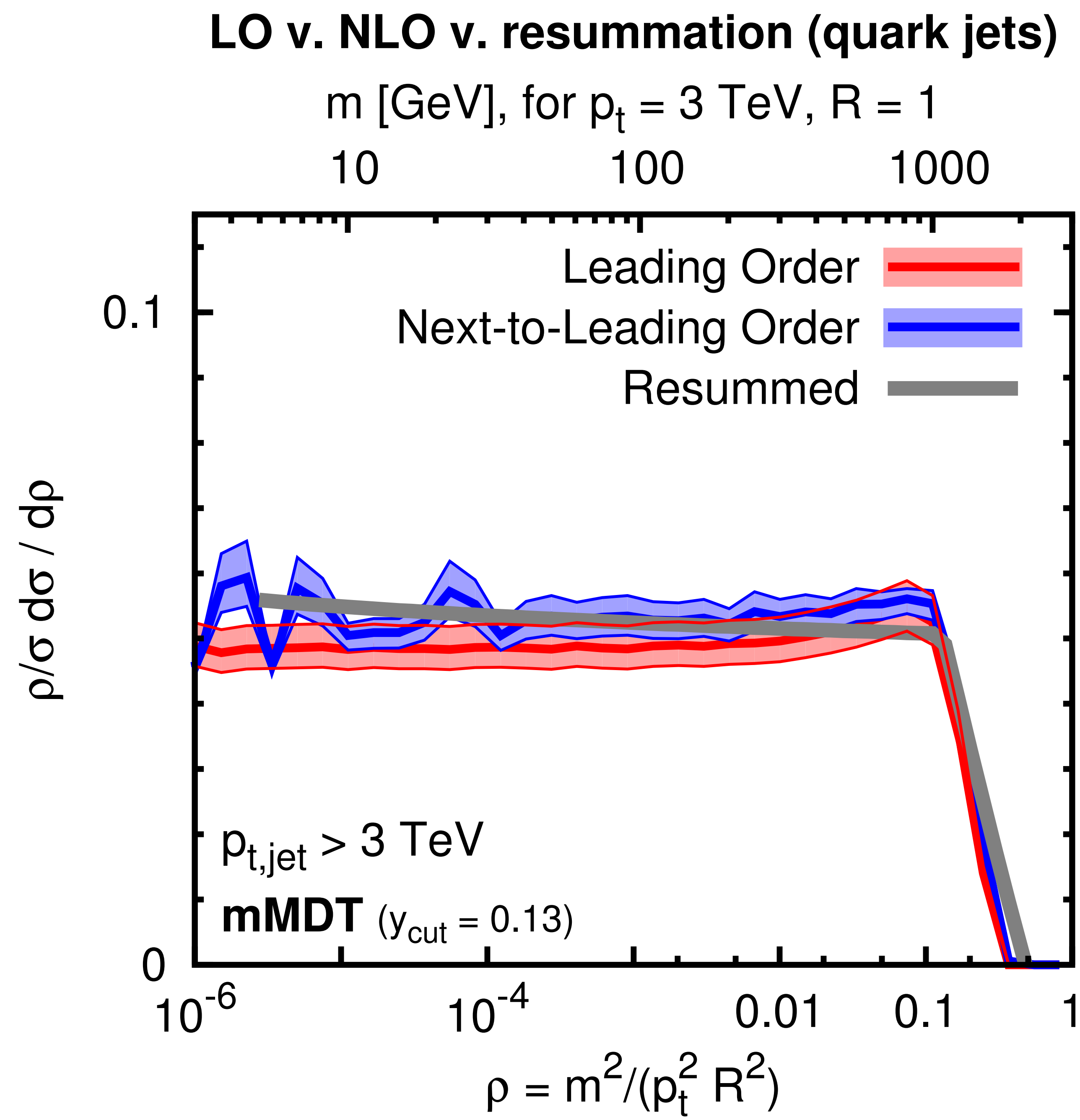


Analytic



Comparing MC & other tools





Performance for finding signals