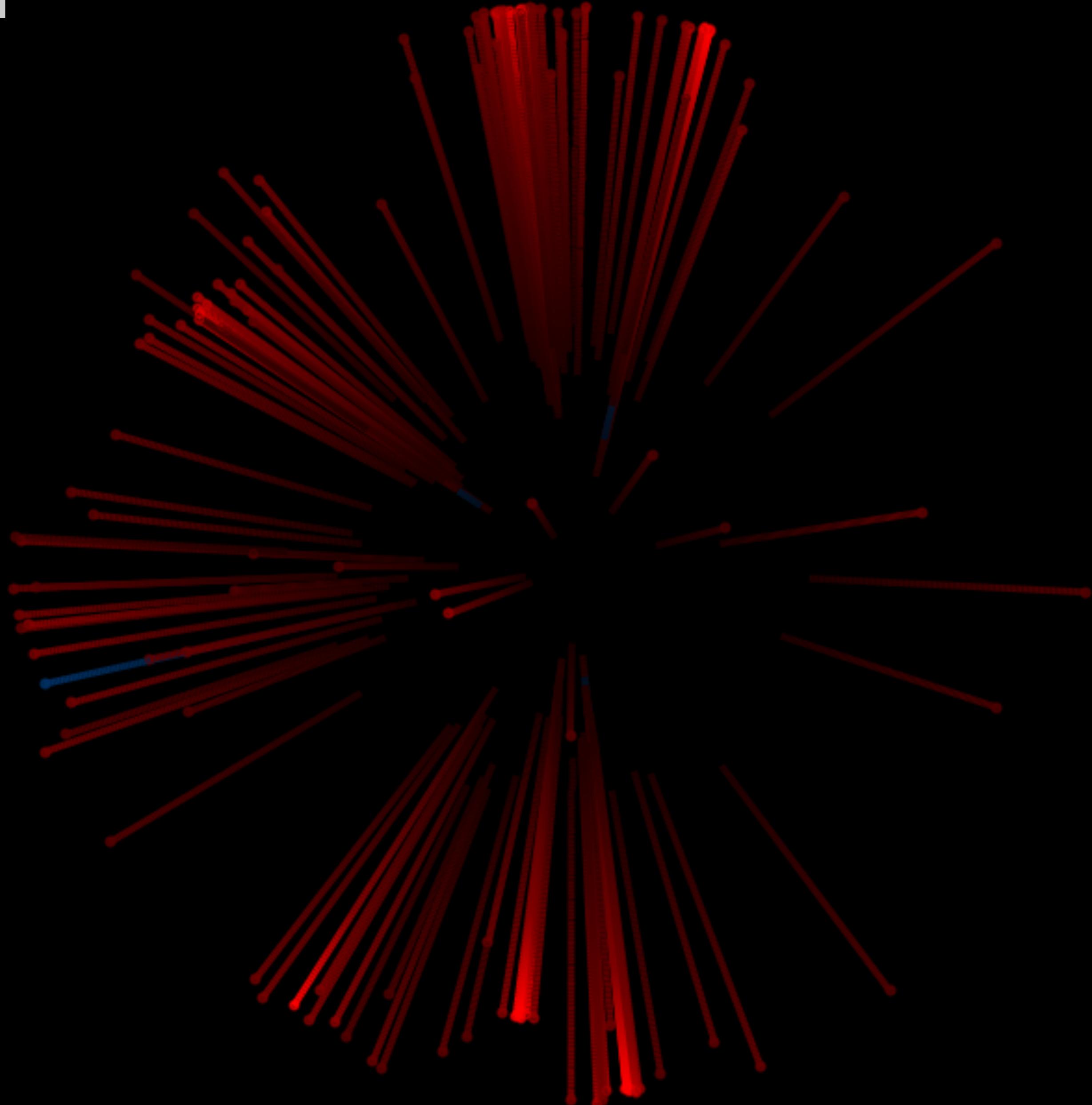


# Jet Reconstruction Theory (2)

PREFIT20: PRrecision Effective  
Field Theory School  
Hamburg, March 2020



**Gavin Salam\***  
Rudolf Peierls Centre for  
Theoretical Physics  
& All Souls College, Oxford

*\* on leave from CERN and CNRS*



- jets come about because of soft-collinear fragmentation & hadronisation
- today's jet reconstruction algorithms are (almost always) sequential recombination algorithms
- when using jets, think about the underlying scattering processes in signal and background (e.g. s- v. t-channel)
- be aware of how jet radius and  $p_t$  cuts may have been optimised for specific processes;
- default choices at ATLAS and CMS  $R=0.4$  and  $p_t > O(20-30)$  GeV; good for many things, but not everything

# Two things that make jets@LHC special

1. The large hierarchy of scales:

$$\sqrt{s} \gg M_{EW}$$

and associated **jet substructure** (this lecture)

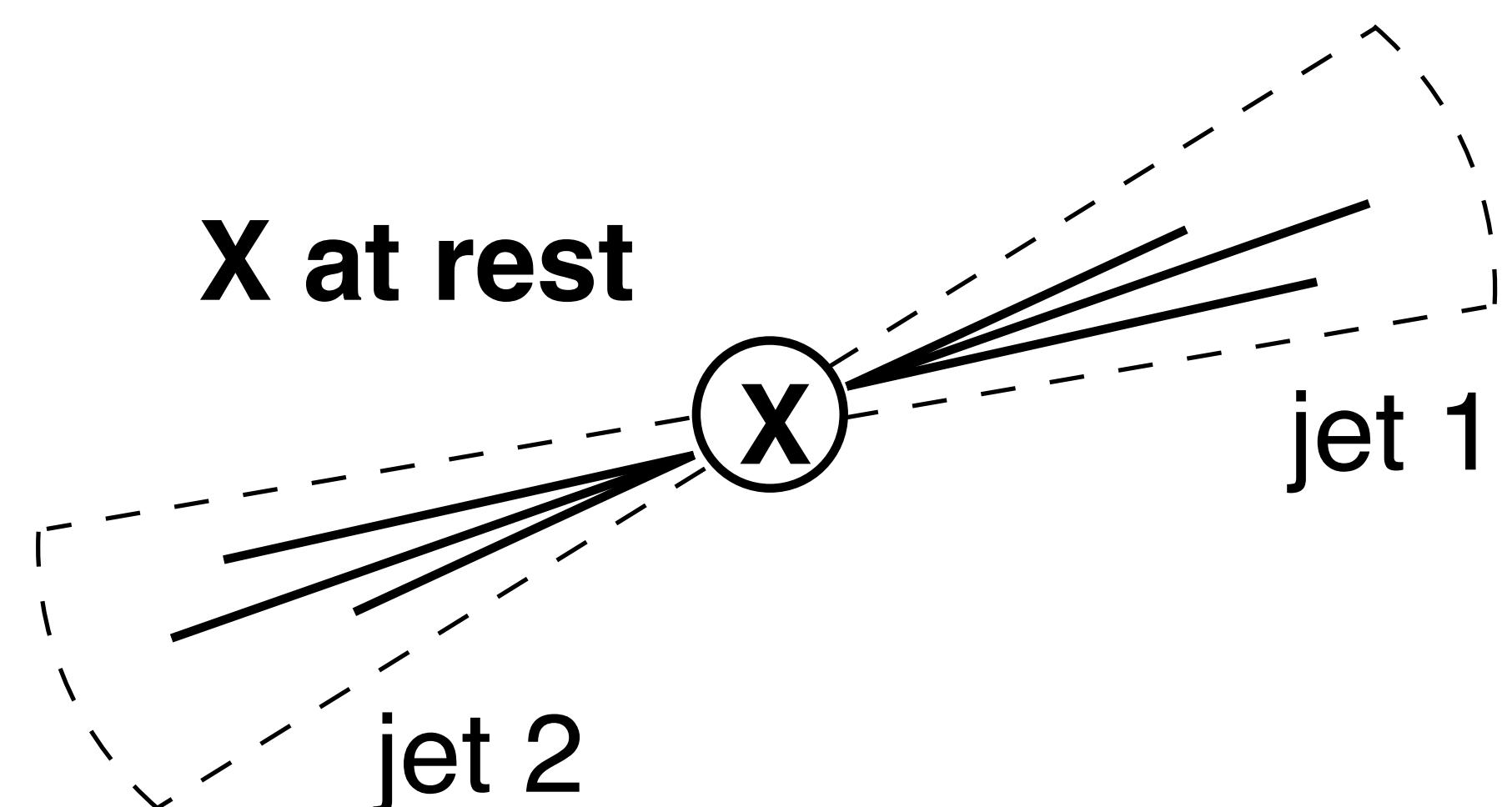
2. The huge pileup (cf. Steven Schramm's lectures):

$$n_{\text{pileup}} \sim 20 - 60 \ (\rightarrow 140 \text{ at HL-LHC})$$

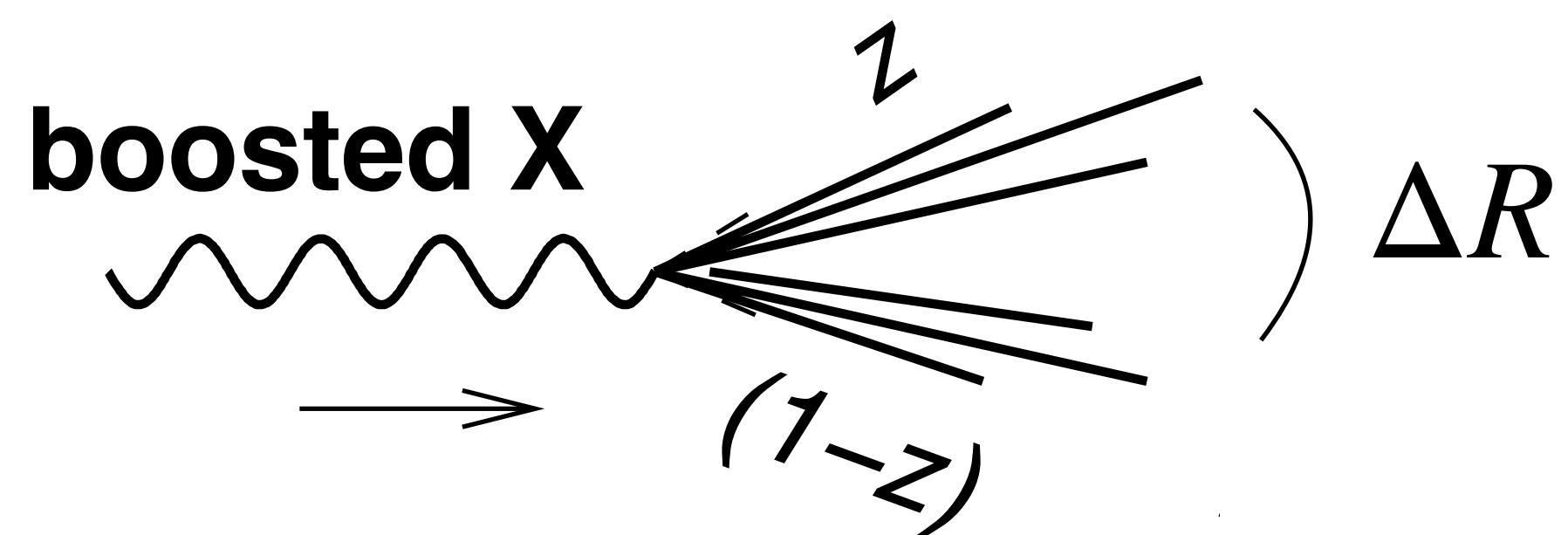
*[These involve two opposite extremes: low  $p_t$  and high  $p_t$ , which nevertheless talk to each other]*

# Boosted hadronic decays of EW-scale objects

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



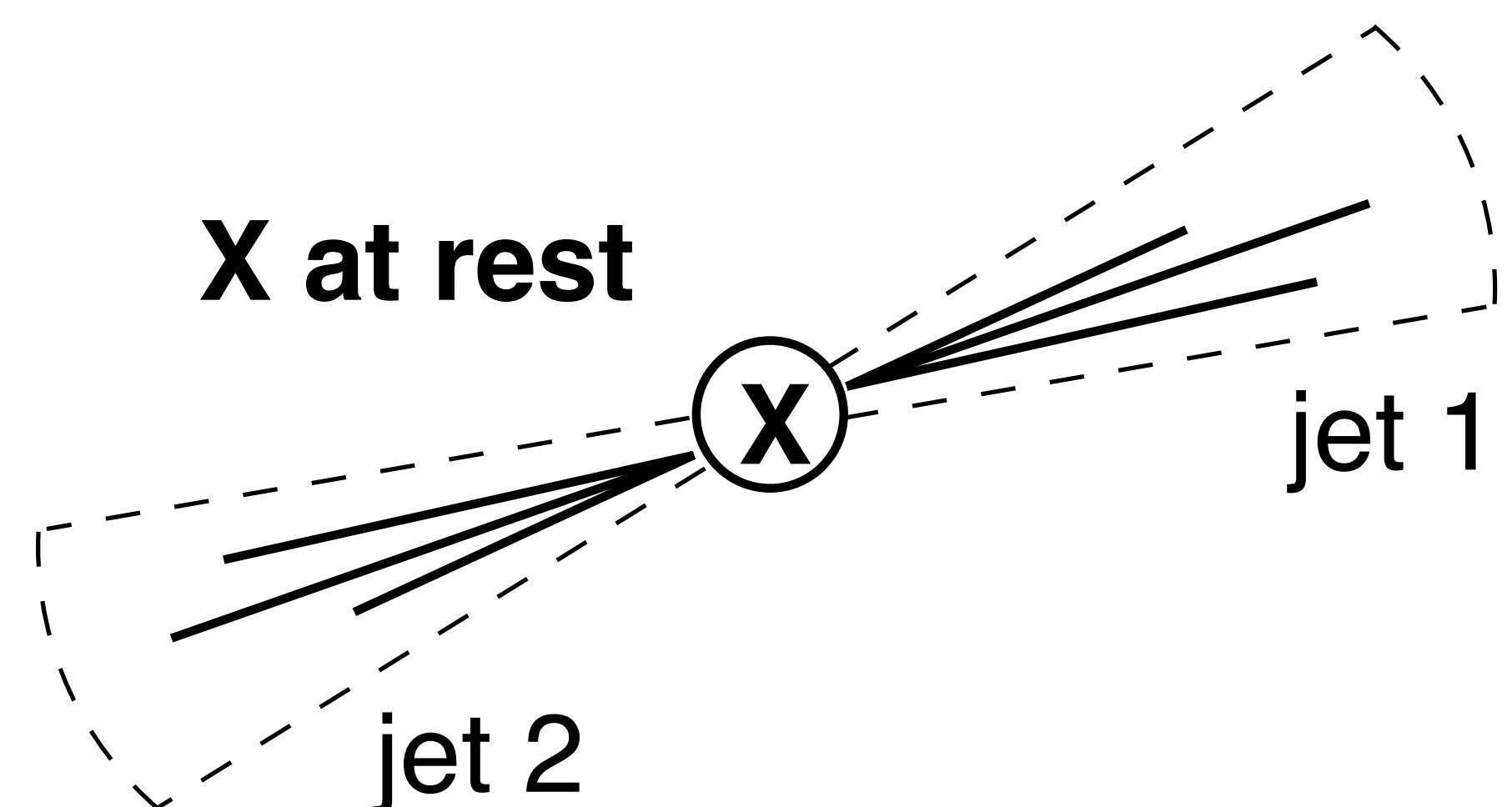
**High- $p_t$  regime: EW object X is boosted, decay is collimated,  $q\bar{q}$  both in same jet**



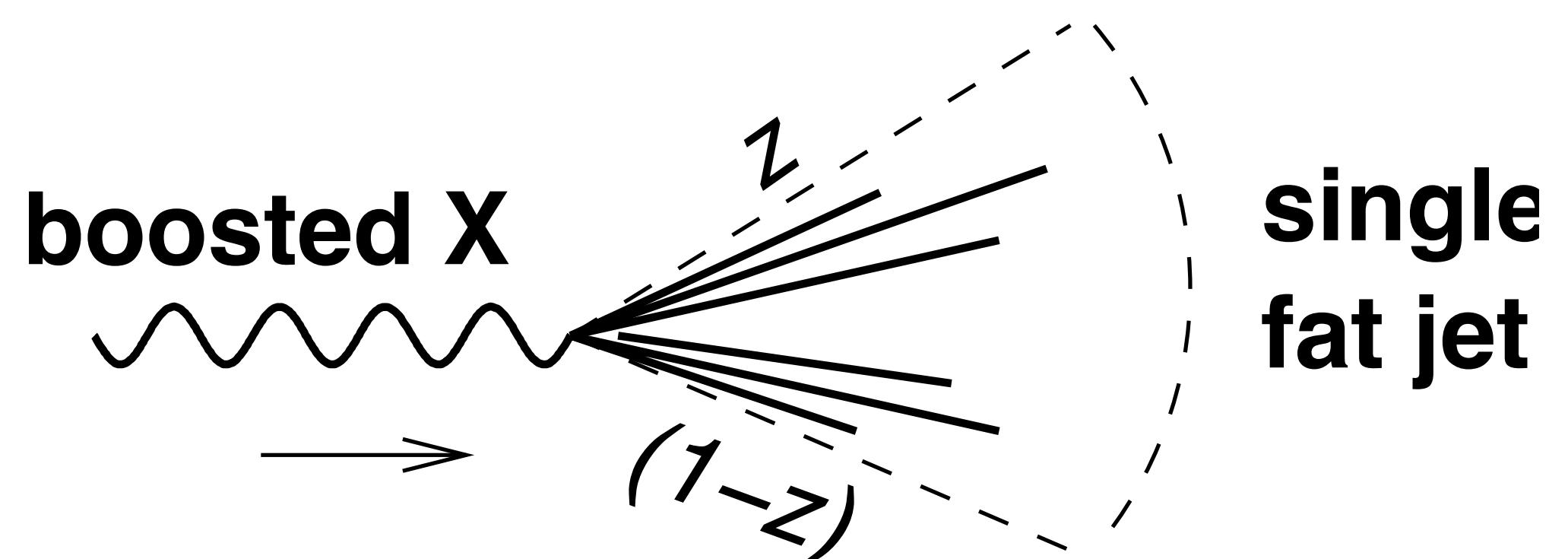
$$\begin{aligned} m_X^2 &= 2 z E_X (1 - z) E_X (1 - \cos \theta) \\ &\simeq z(1 - z) p_{tX}^2 \Delta R^2 \end{aligned}$$

# Boosted hadronic decays of EW-scale objects

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



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

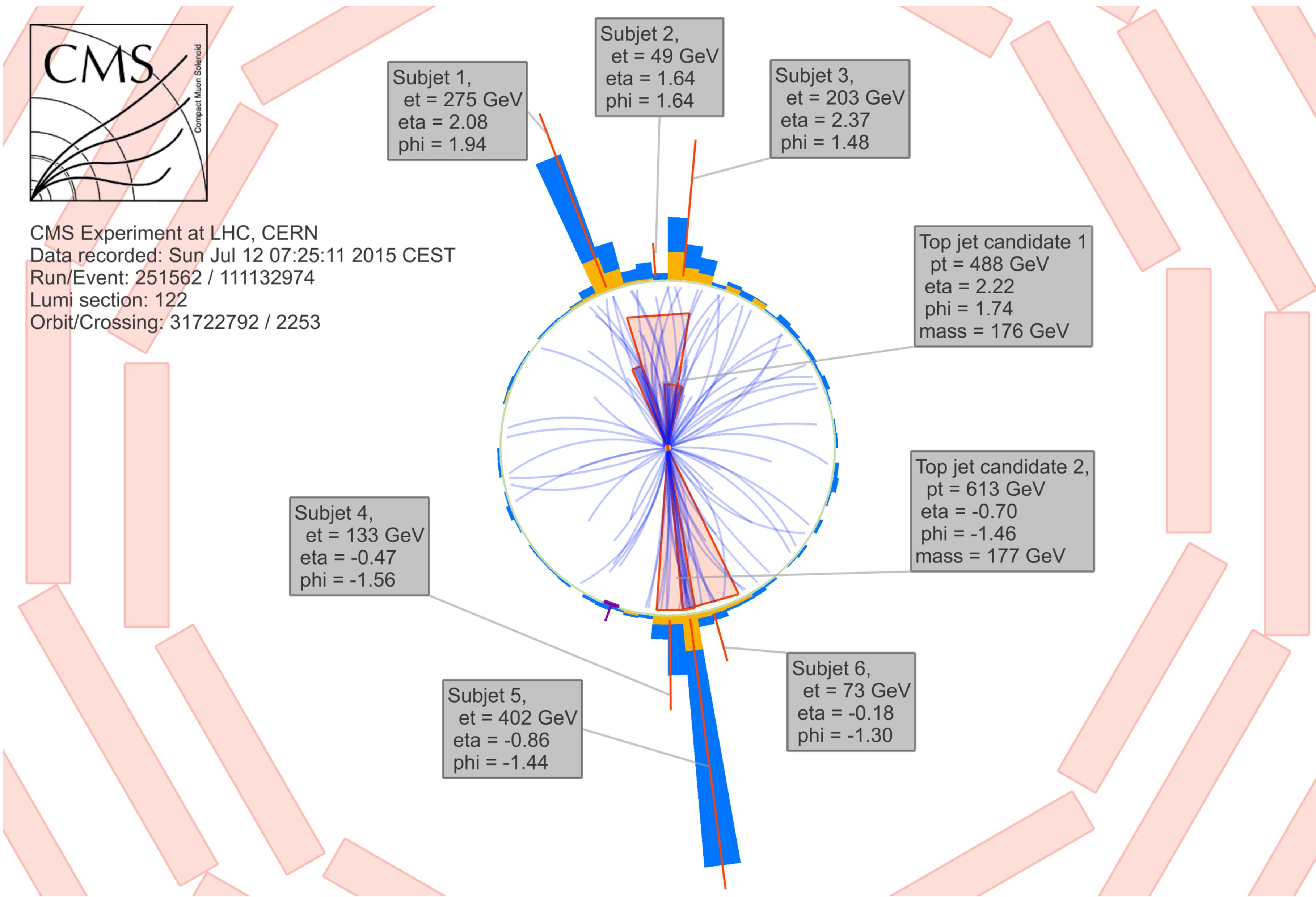


$$\text{Happens for } p_{tX} > \frac{m_X}{R} \frac{1}{\sqrt{z(1-z)}} \gtrsim \frac{2m_X}{R}$$

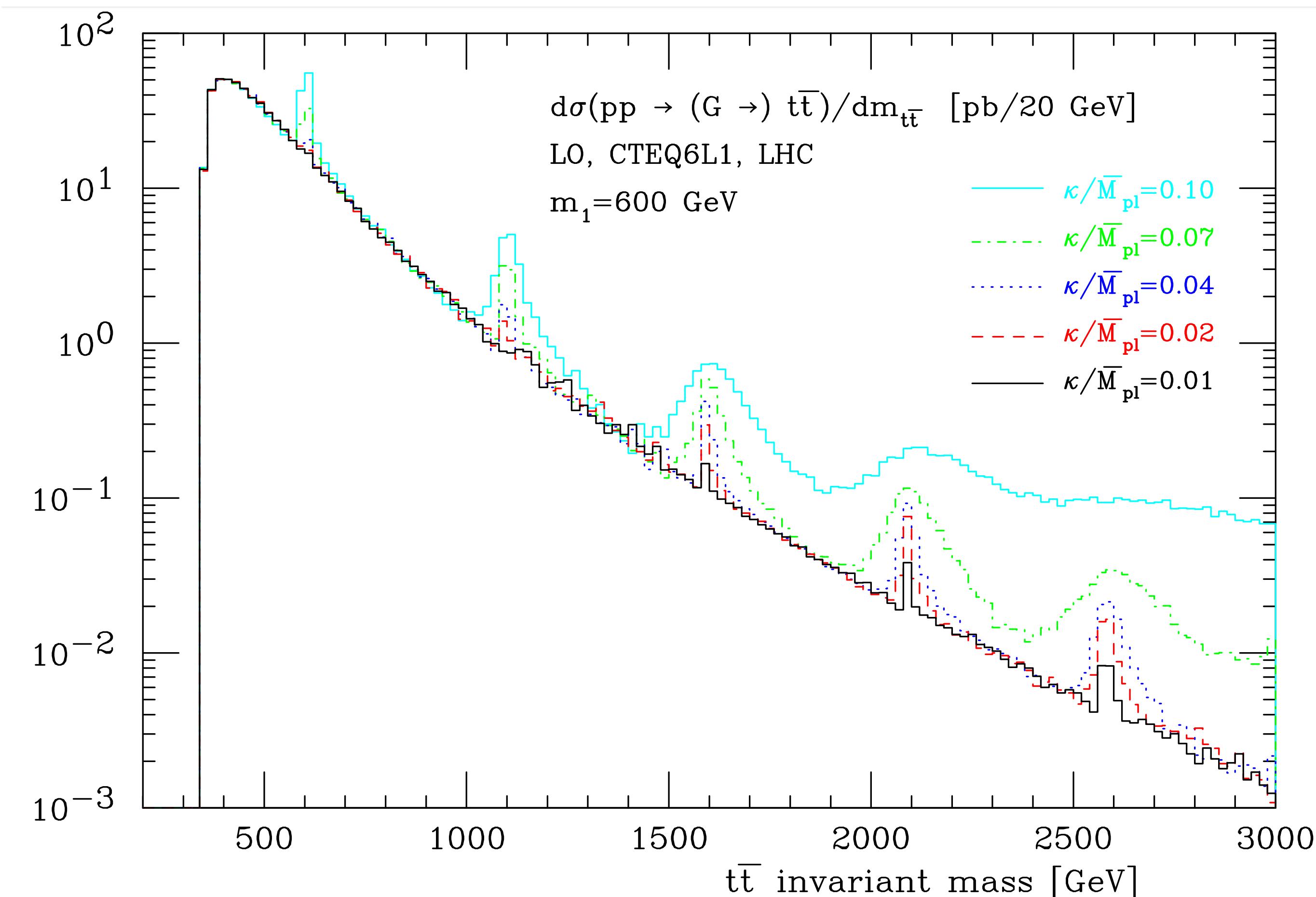
e.g. for a W boson, and R=0.4,  $p_t \gtrsim 400 \text{ GeV}$

*a  $pp$  collision that produces a high  $p_t$  top-antitop pair, resulting in two “top-jets”, each with subjets*

*Such events probe point-like nature of top quarks to TeV scale & allow you to search for new  $t\bar{t}$ bar resonances*



e.g. ttbar resonances



RS KK resonances  $\rightarrow t\bar{t}$ , from Frederix & Maltoni, 0712.2355

NB: QCD dijet spectrum is  $\sim 10^3$  times  $t\bar{t}$

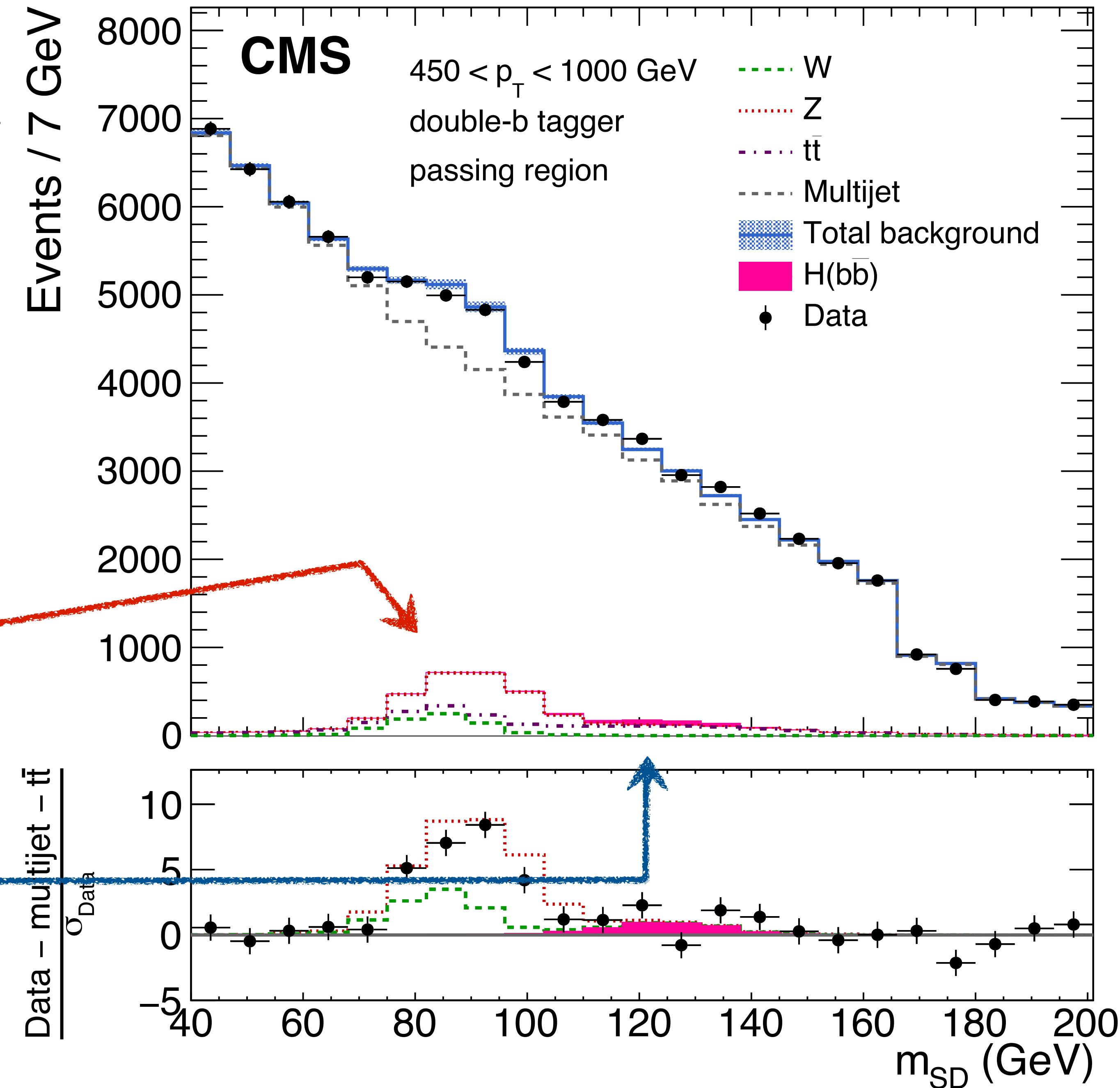
# high $p_T$ Higgs

We wouldn't trust electromagnetism if we'd only tested at one length/momentum scale.

New Higgs interactions need testing at both low and (here) high momenta.

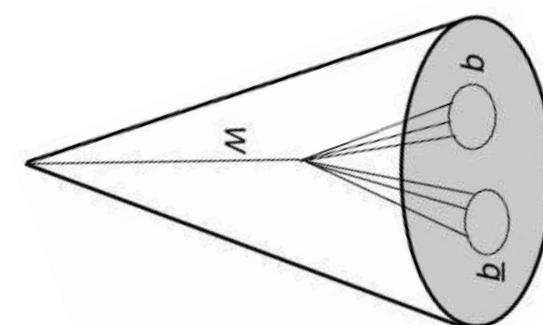
high- $p_T$   $Z \rightarrow bb$  ( $5\sigma$ )

high- $p_T$   $H \rightarrow bb$  ( $\sim 1\sigma$ )

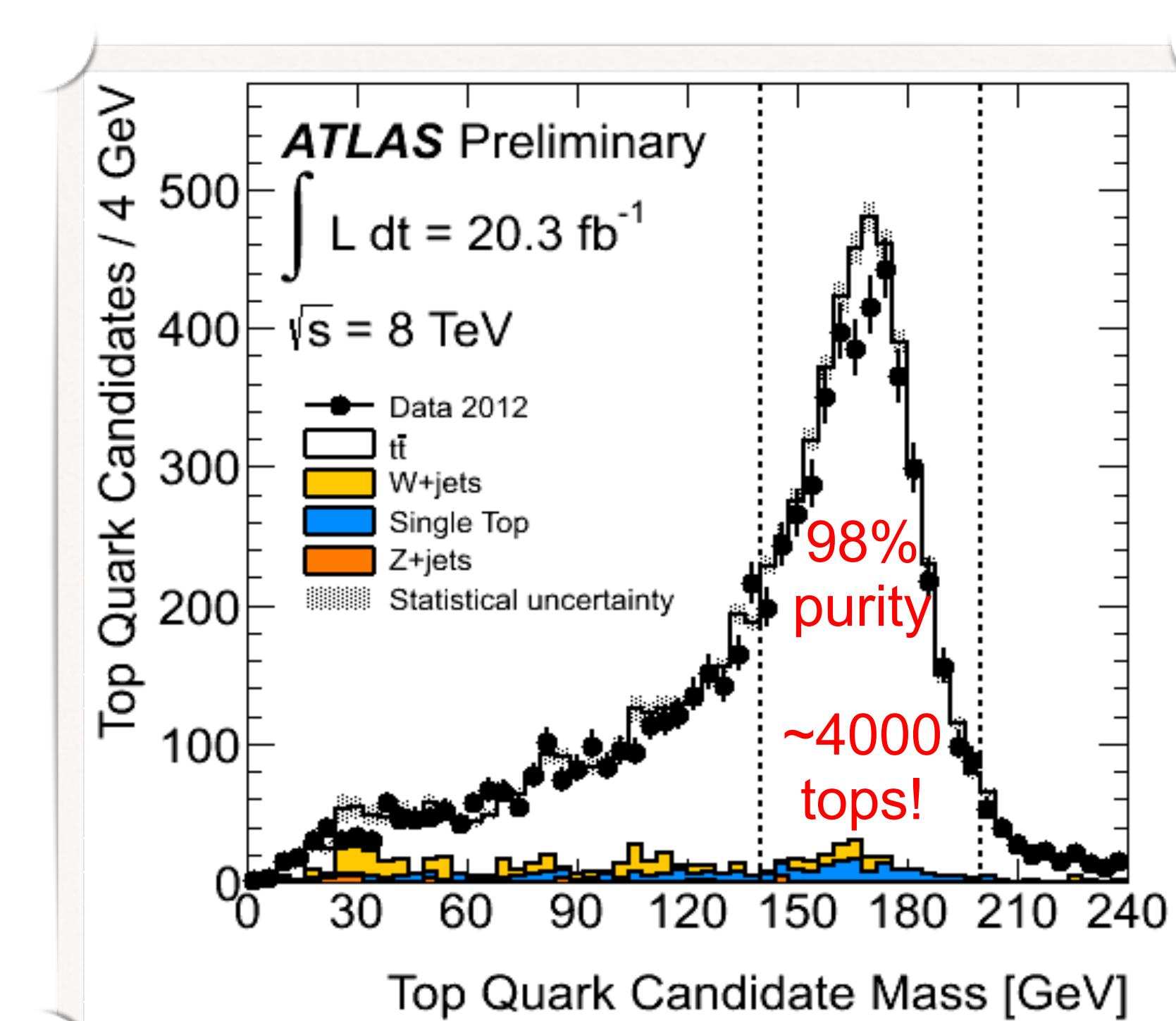
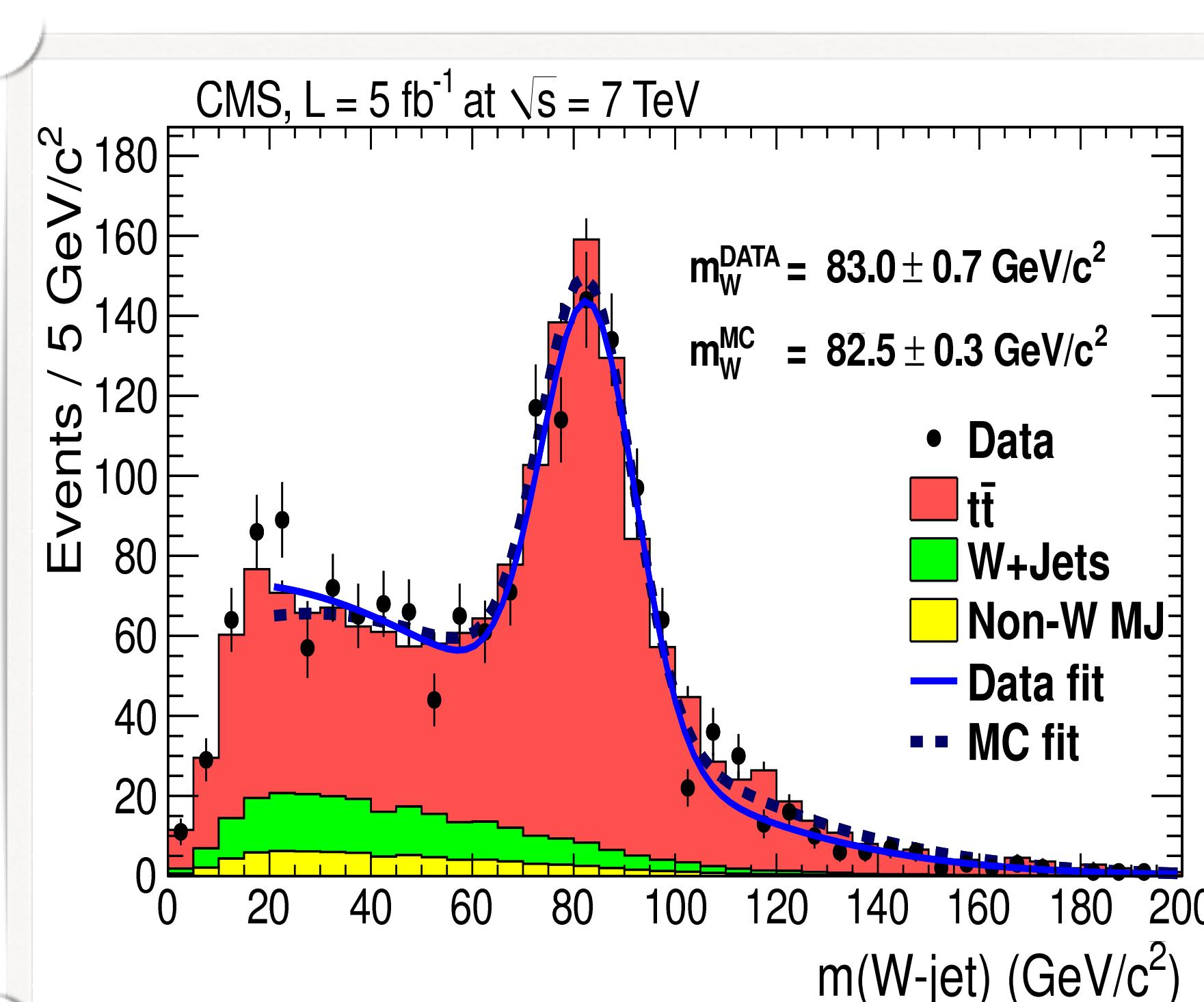
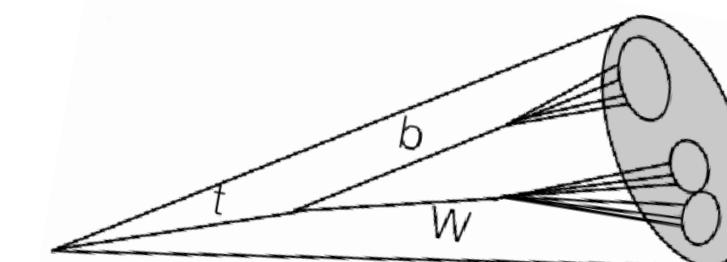


# Seeing W's and tops in a single jet

## W's in a single jet

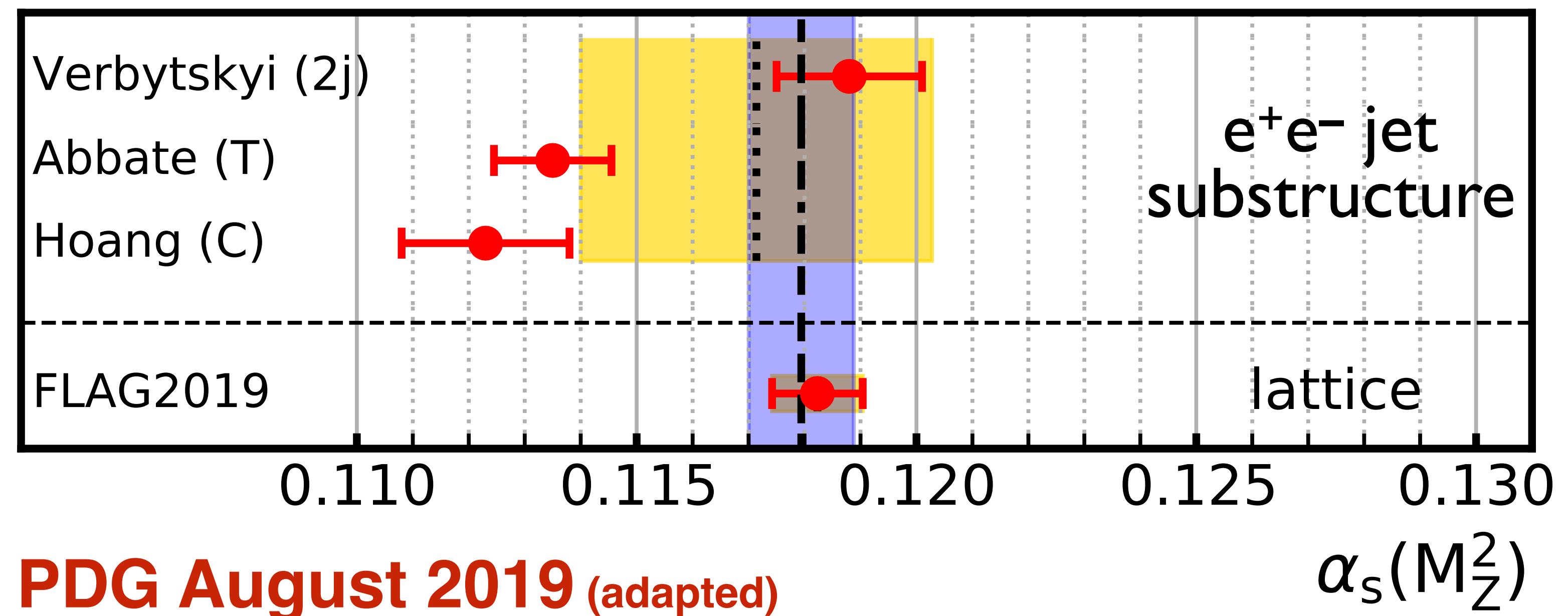


## tops in a single jet



# jet substructure & strong coupling determinations

## selected $\alpha_s$ determinations (most precise + lattice)



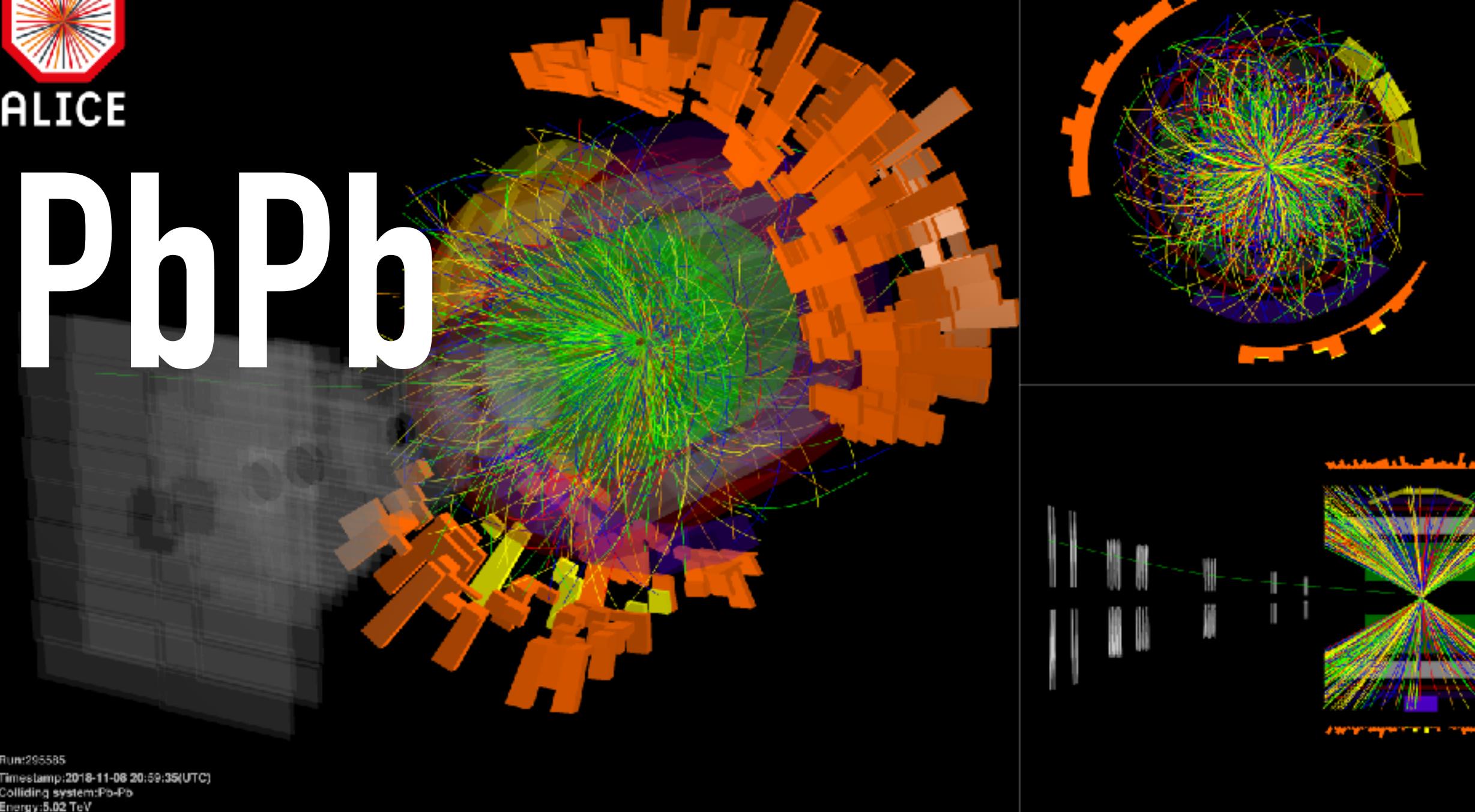
- some of the most smallest quoted errors come from  $e^+e^-$  jet substructure (better known as jet shapes and rates)
- though they don't agree with each other (and C & T don't agree with best lattice results)



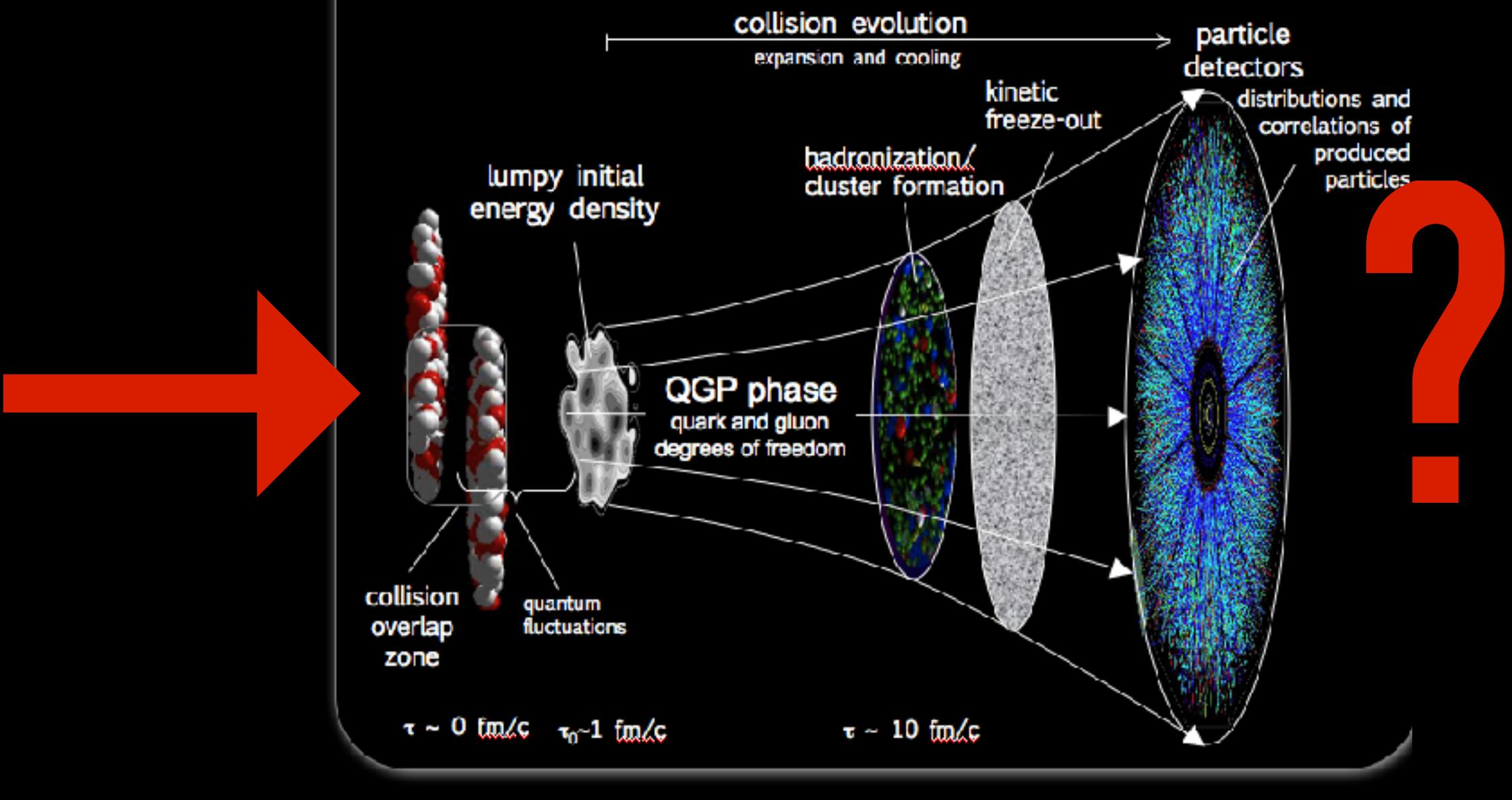
ALICE

# PbPb

Run:295585  
Timestamp:2018-11-08 20:59:35(UTC)  
Colliding system:Pb-Pb  
Energy:5.02 TeV



## Heavy-ion Collisions: Rapid Expansion

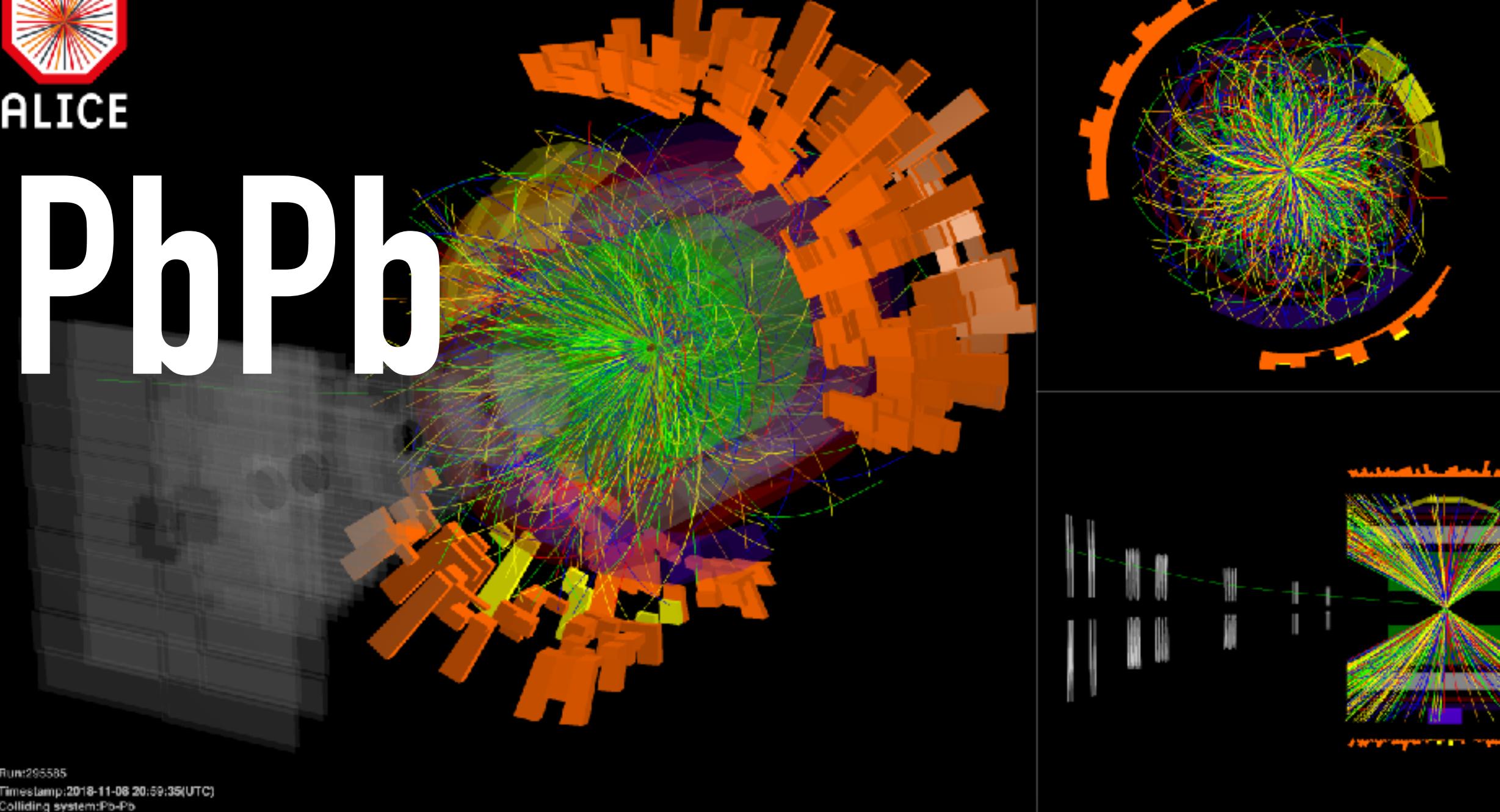


?

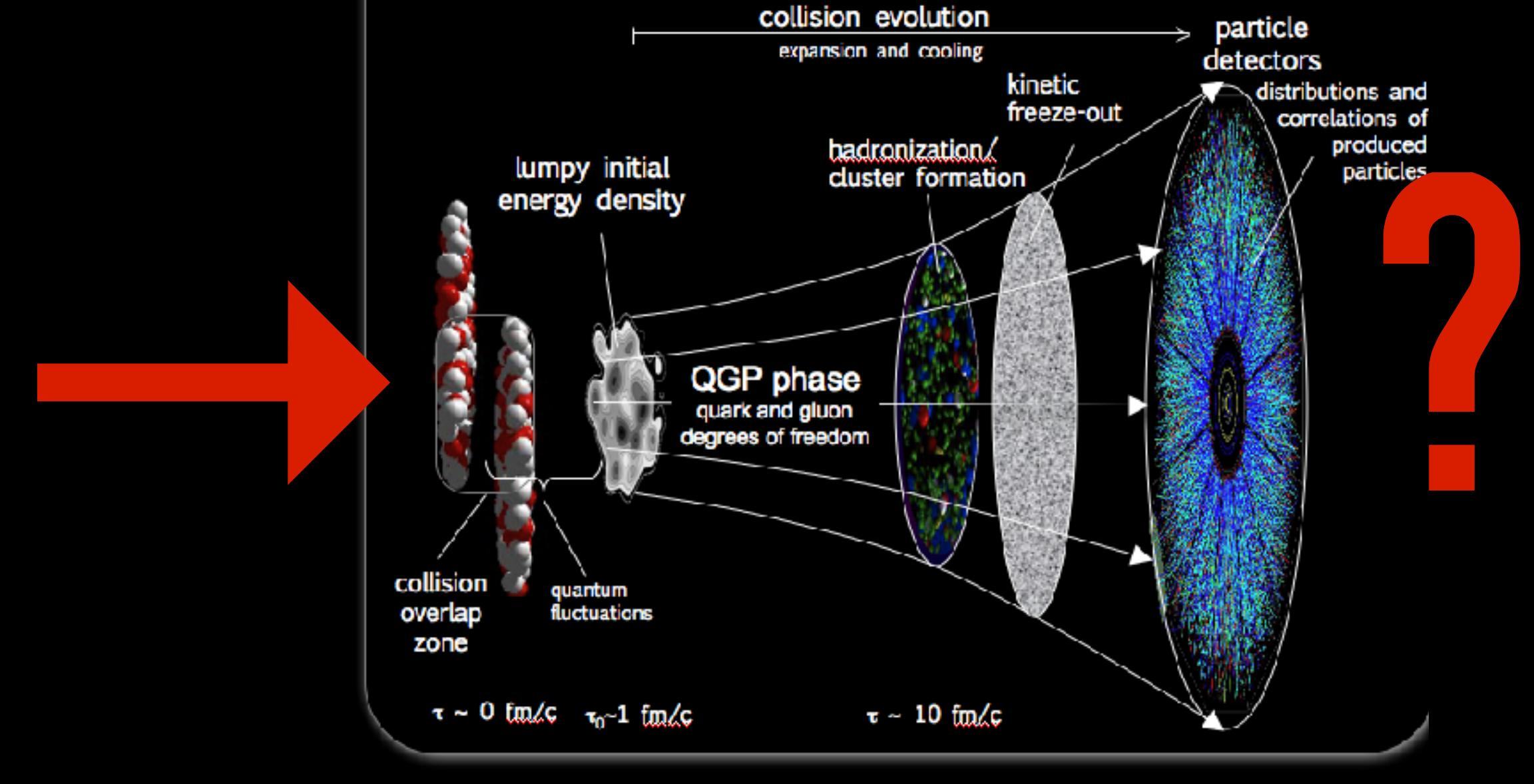


ALICE

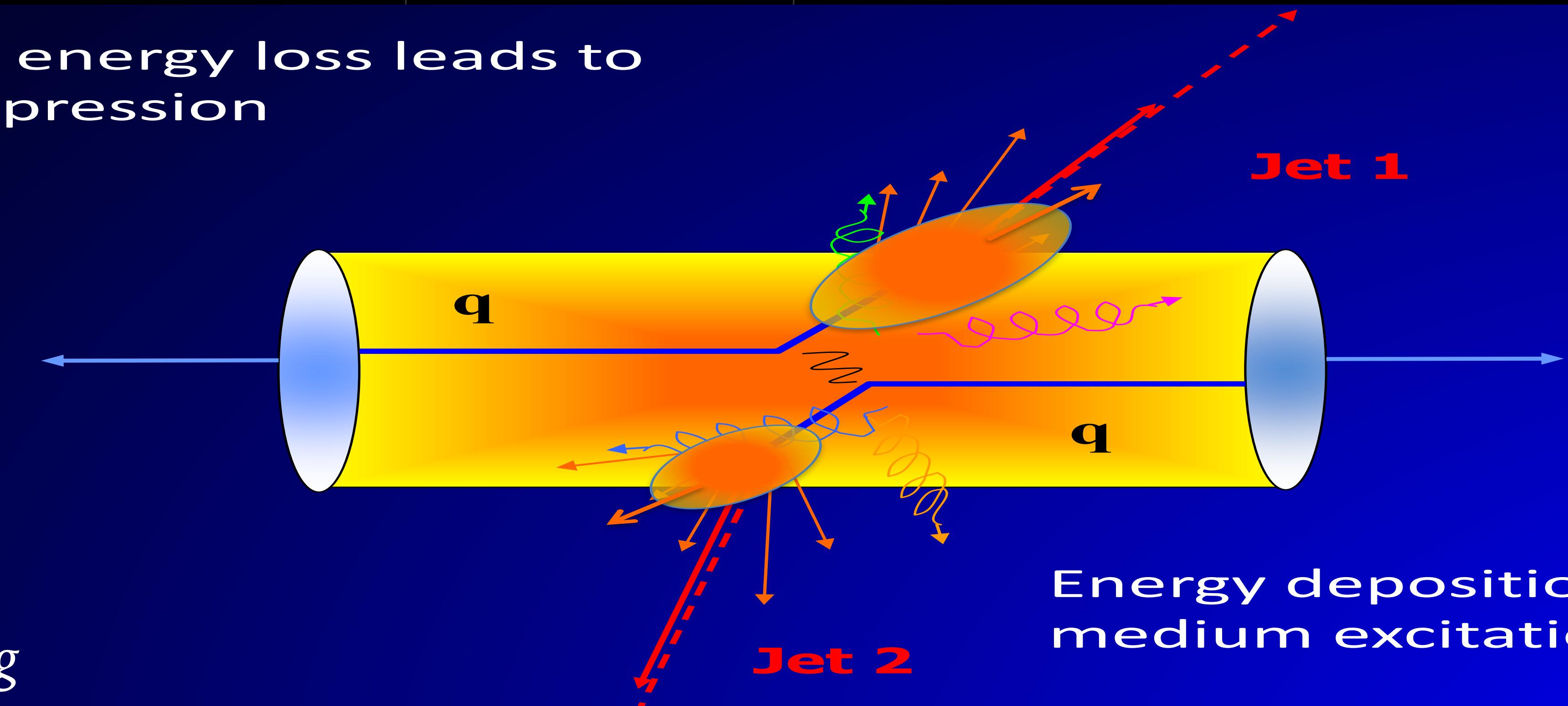
# PbPb



Heavy-ion Collisions: Rapid Expansion



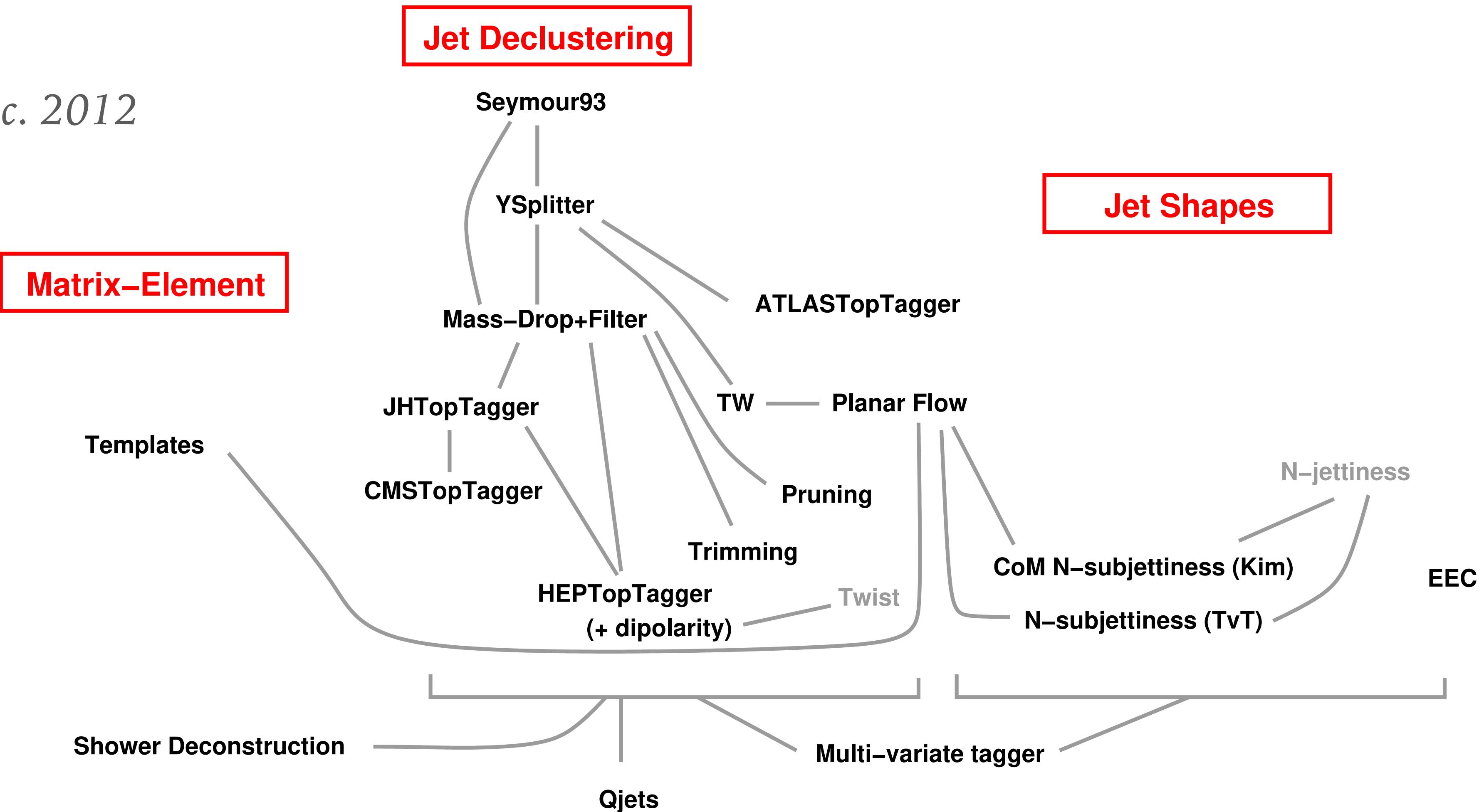
Parton energy loss leads to jet suppression



from X-N. Wang

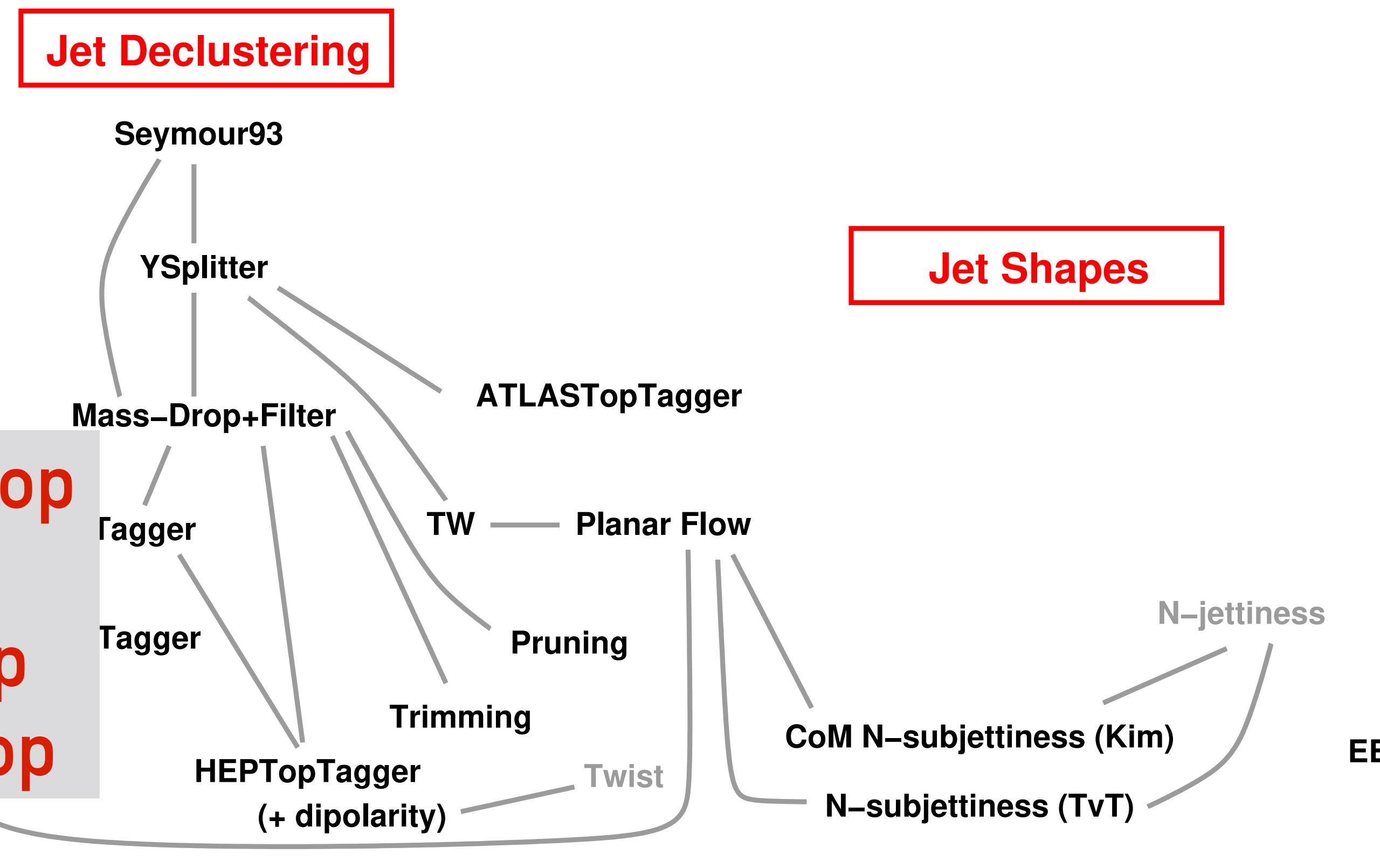
# jet substructure field is full of activity

c. 2012



jet substructure field is full of activity

c. 2018

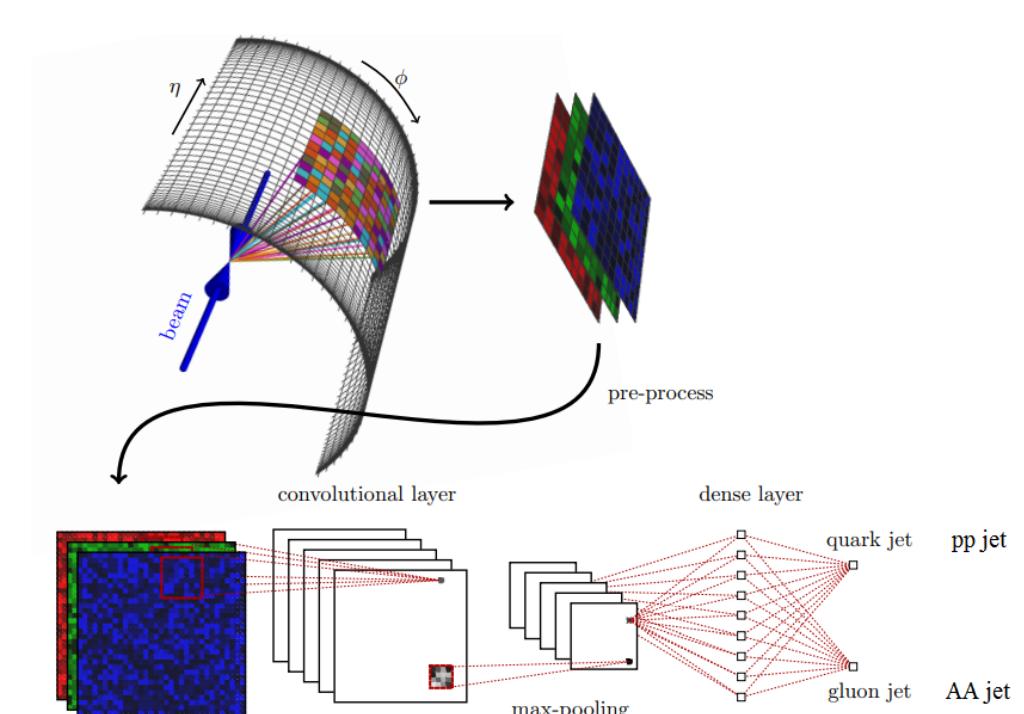


# classification without labels weak supervision

# machine learning

## DNN, CNN,

## RNN, LSTM, etc



$C_n, D_n, v e_n^{(\beta)}, M_n, N_n,$   
 $U_n, EFPs$

etc.

## Tagging on core n-prong structure

similar to asking for 2 jets in  $Z'$  search or 3 jets for a hadronic top

## Tagging on energy flow pattern

not typically done in normal jet usage (closest analogue is jet veto in  $H \rightarrow WW$  analyses)

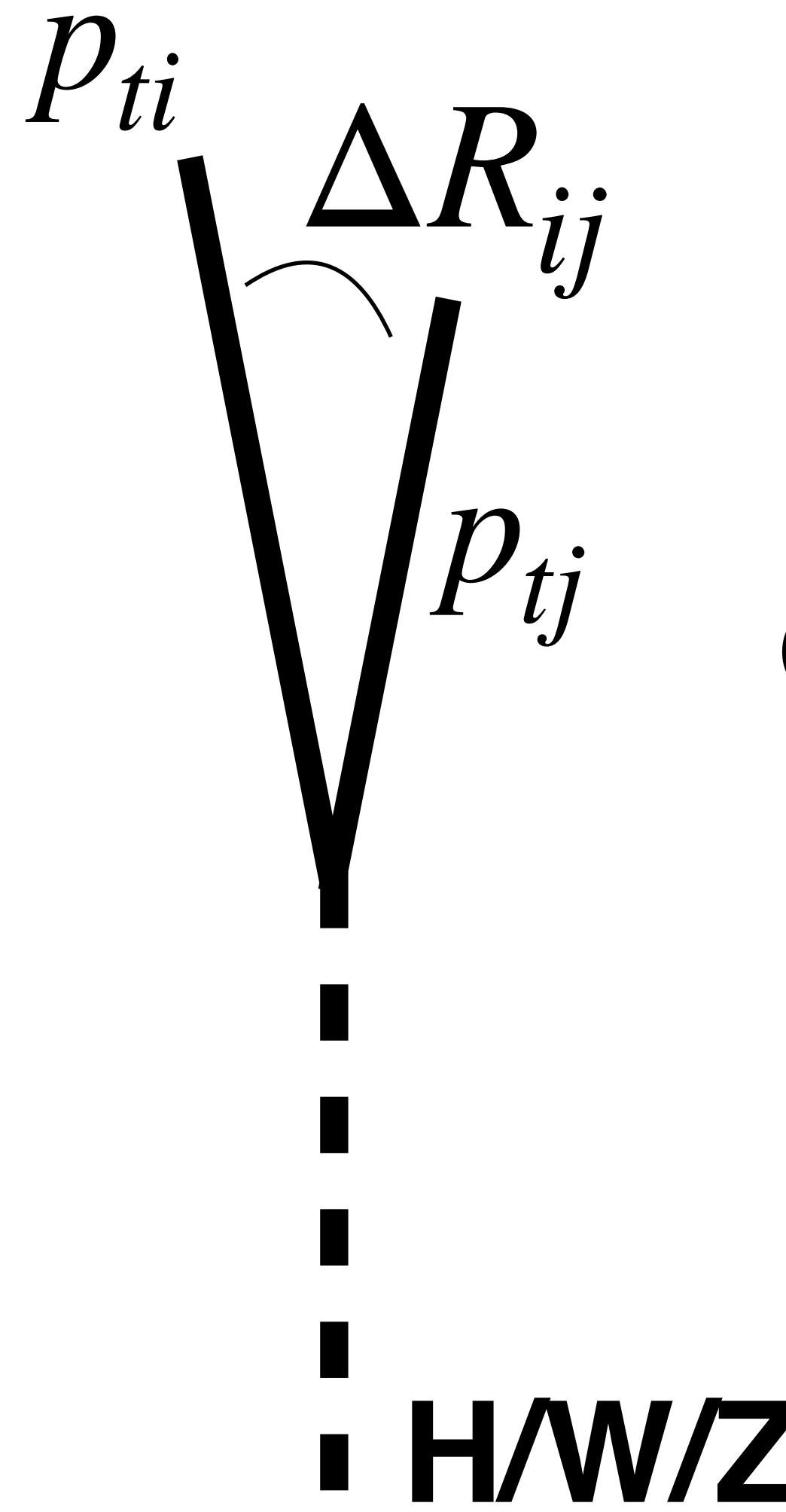
## Grooming

improves signal mass resolution (removing pileup, etc.), without significantly changing background & signal event numbers

# Tagging two-prong decays

## 1. cutting on hard kinematics

# “signals” (e.g. H decay) v. QCD backgrounds



**[signal]**

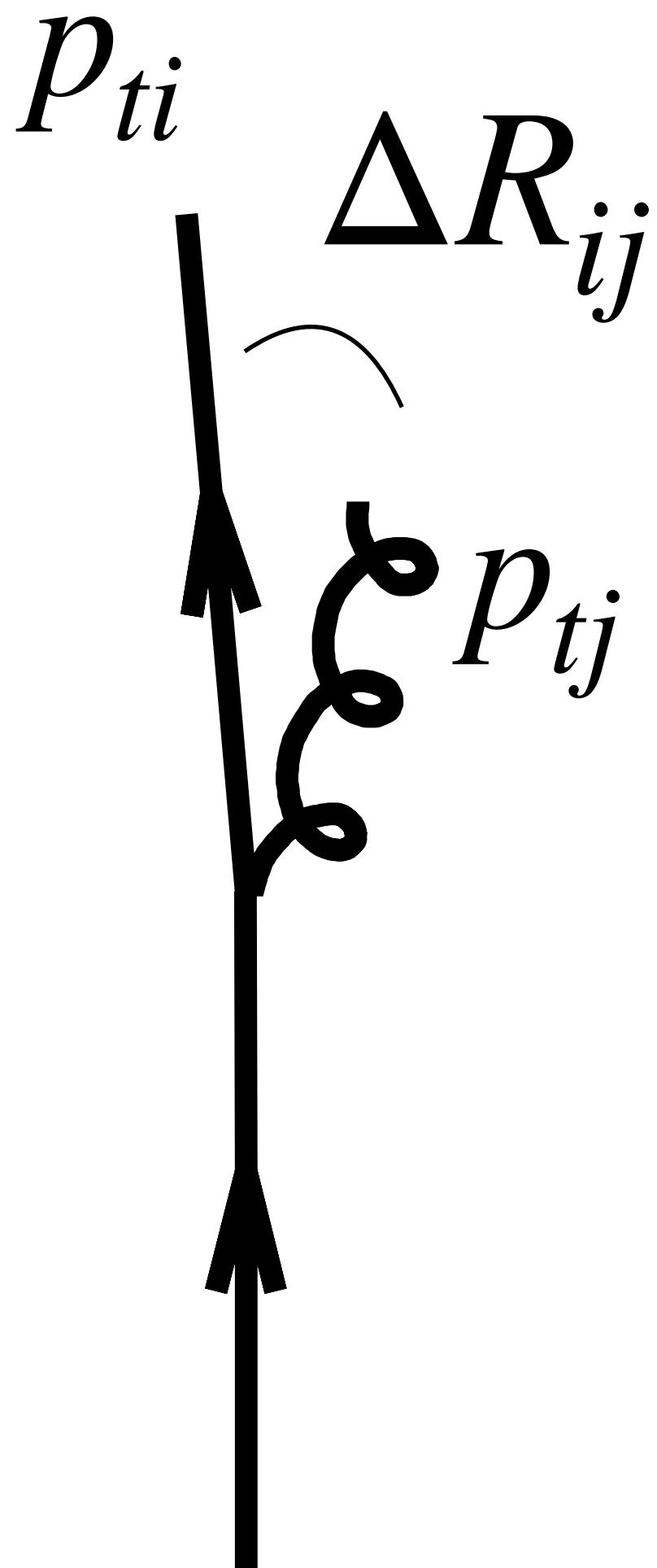
## features in common

Higgs decay dominantly produces 2-prong structure.

Quark/gluon-induced jet can also branch to 2-prong structure

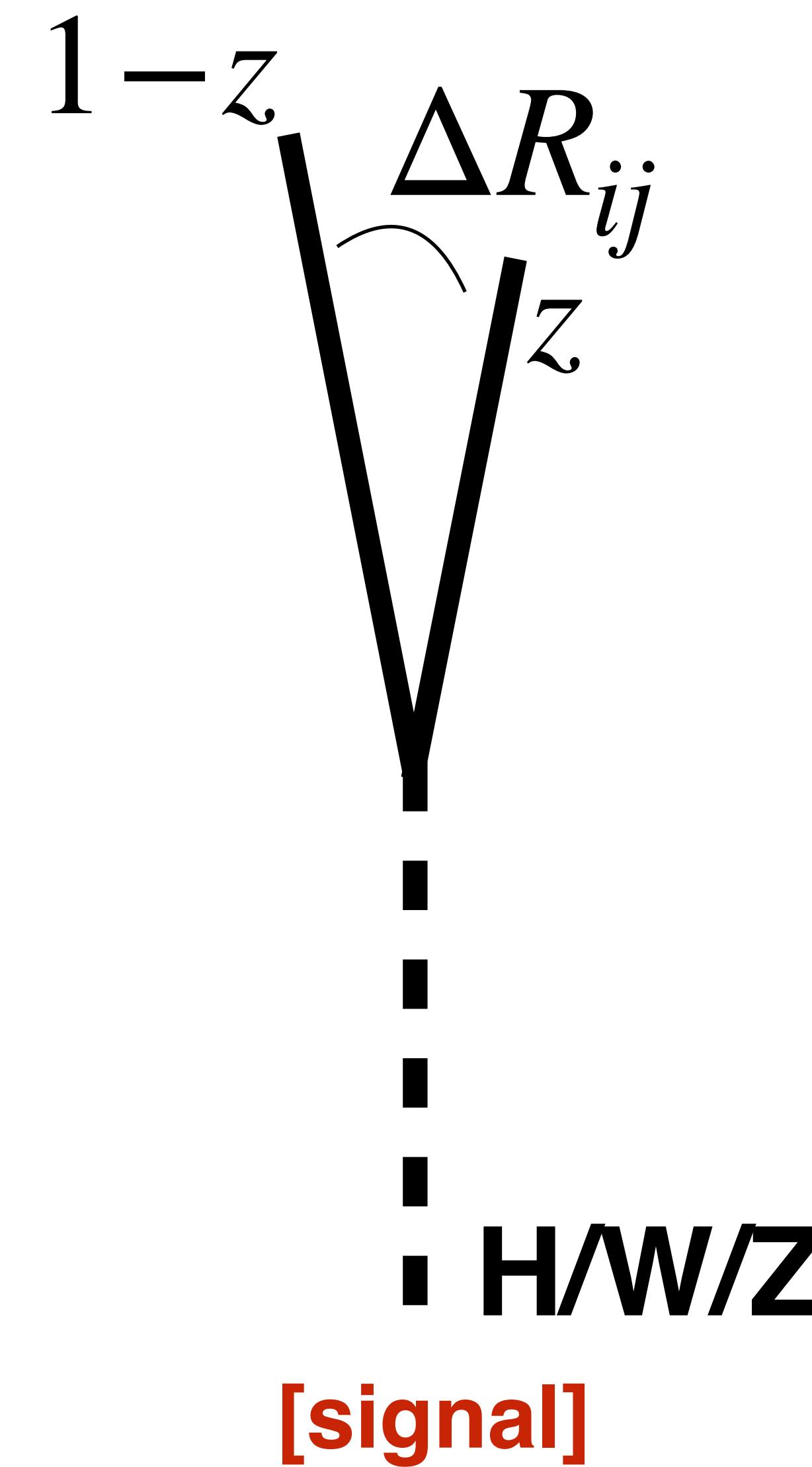
Resulting jet mass is

$$m^2 \simeq p_{ti} p_{tj} \Delta R_{ij}^2$$



**[background]**

# “signals” (e.g. H decay) v. QCD backgrounds



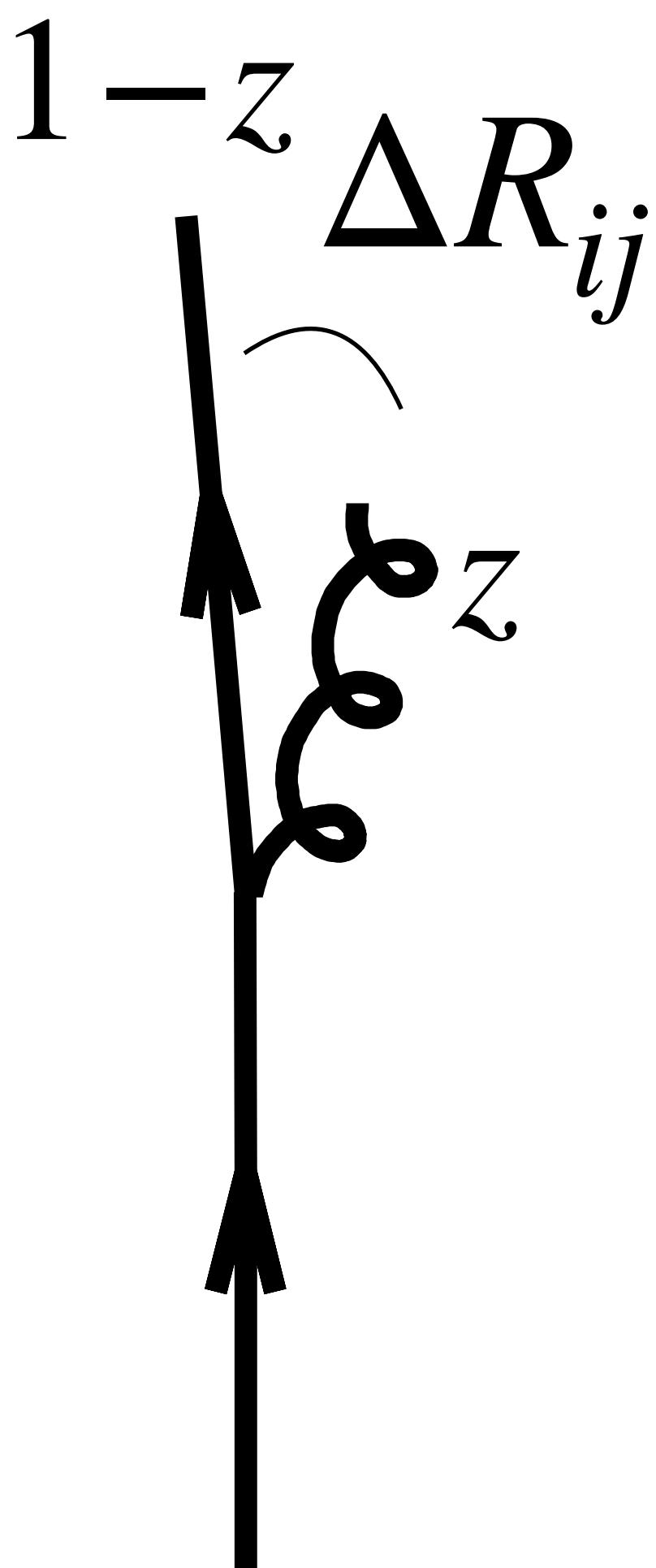
features where they differ

energy-sharing between prongs

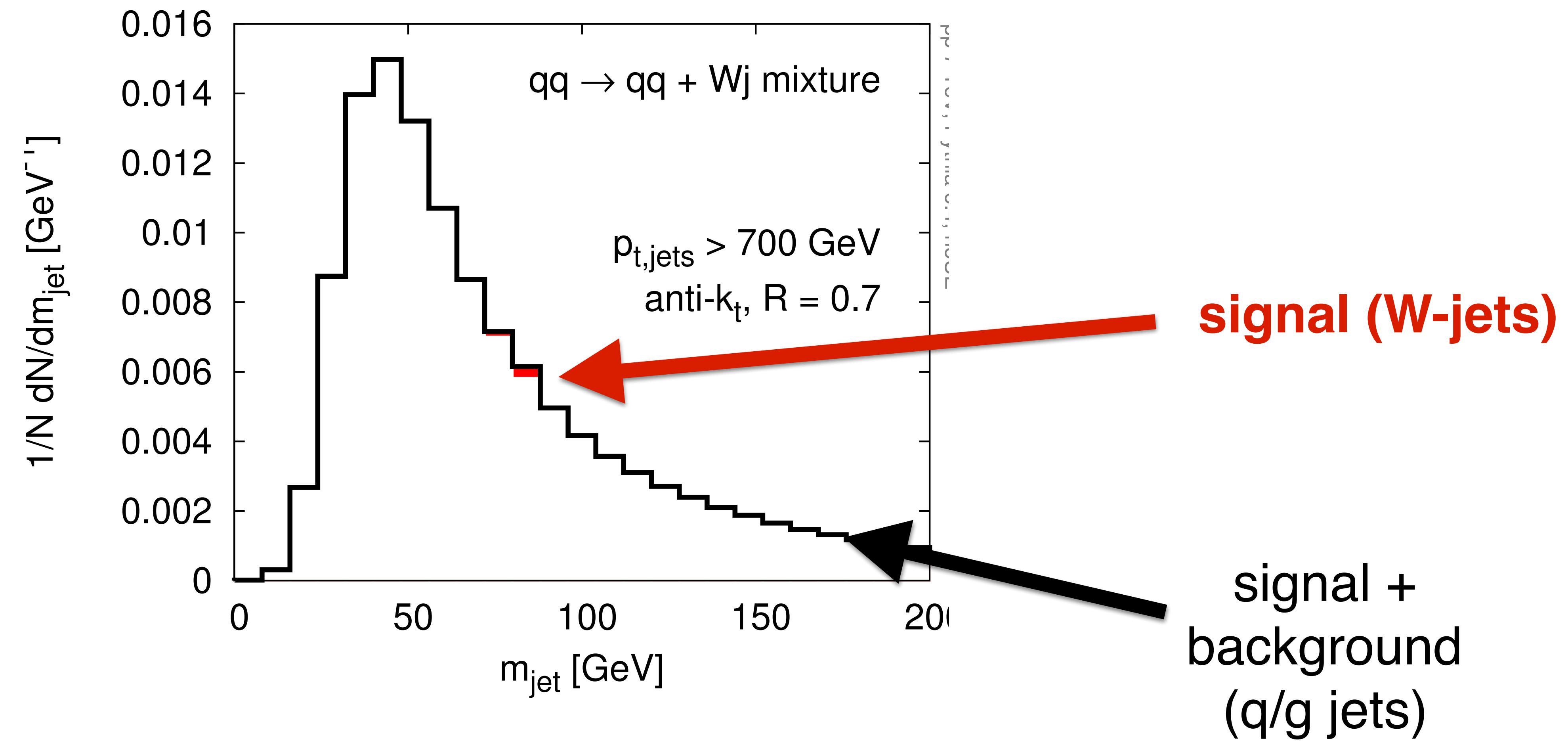
Higgs:

$$\frac{dP}{dz} \propto \text{const}$$

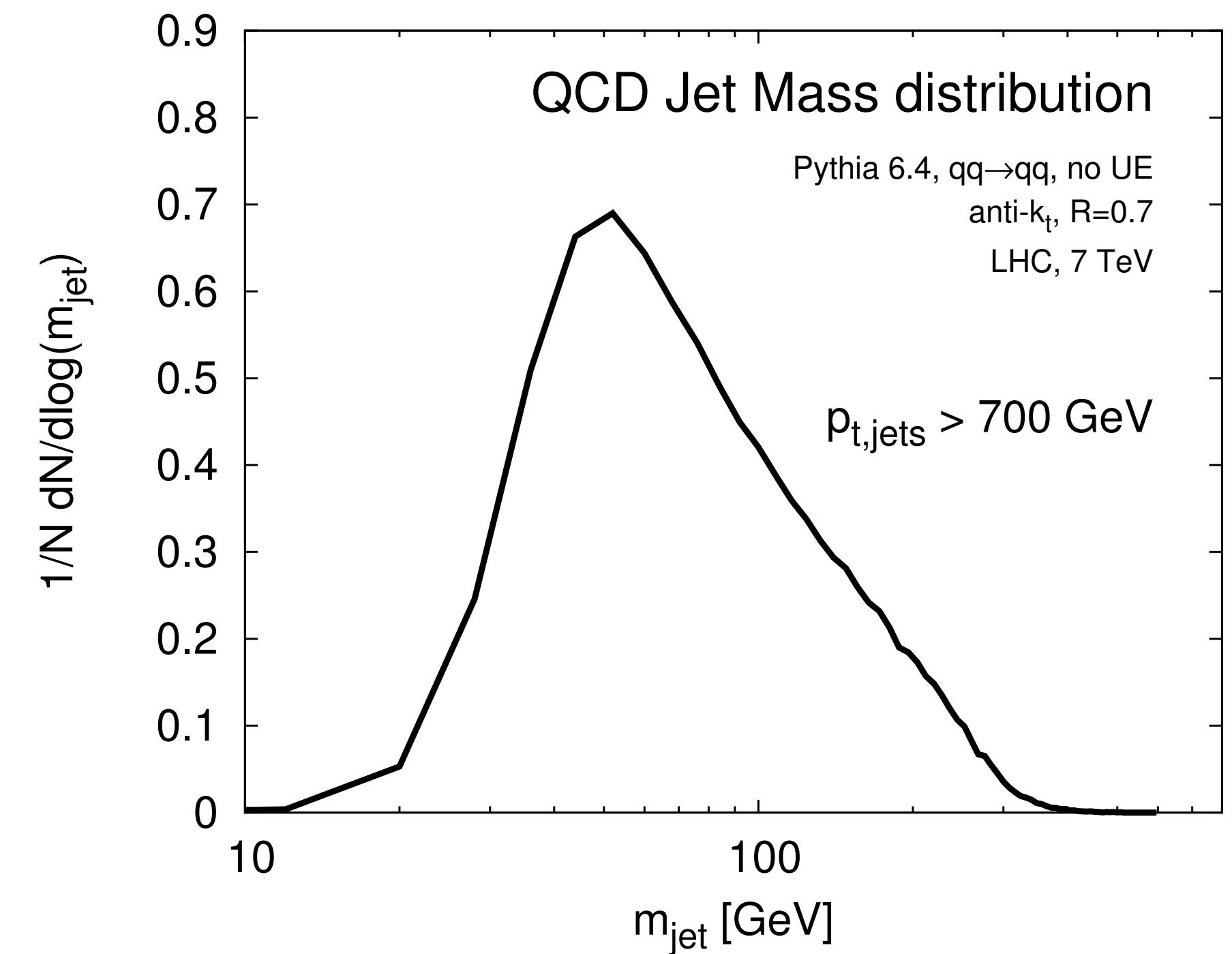
gluon emission:  $\frac{dP}{dz} \propto \frac{1}{z}$



# Look at a jet-mass distribution



# Inside the jet mass

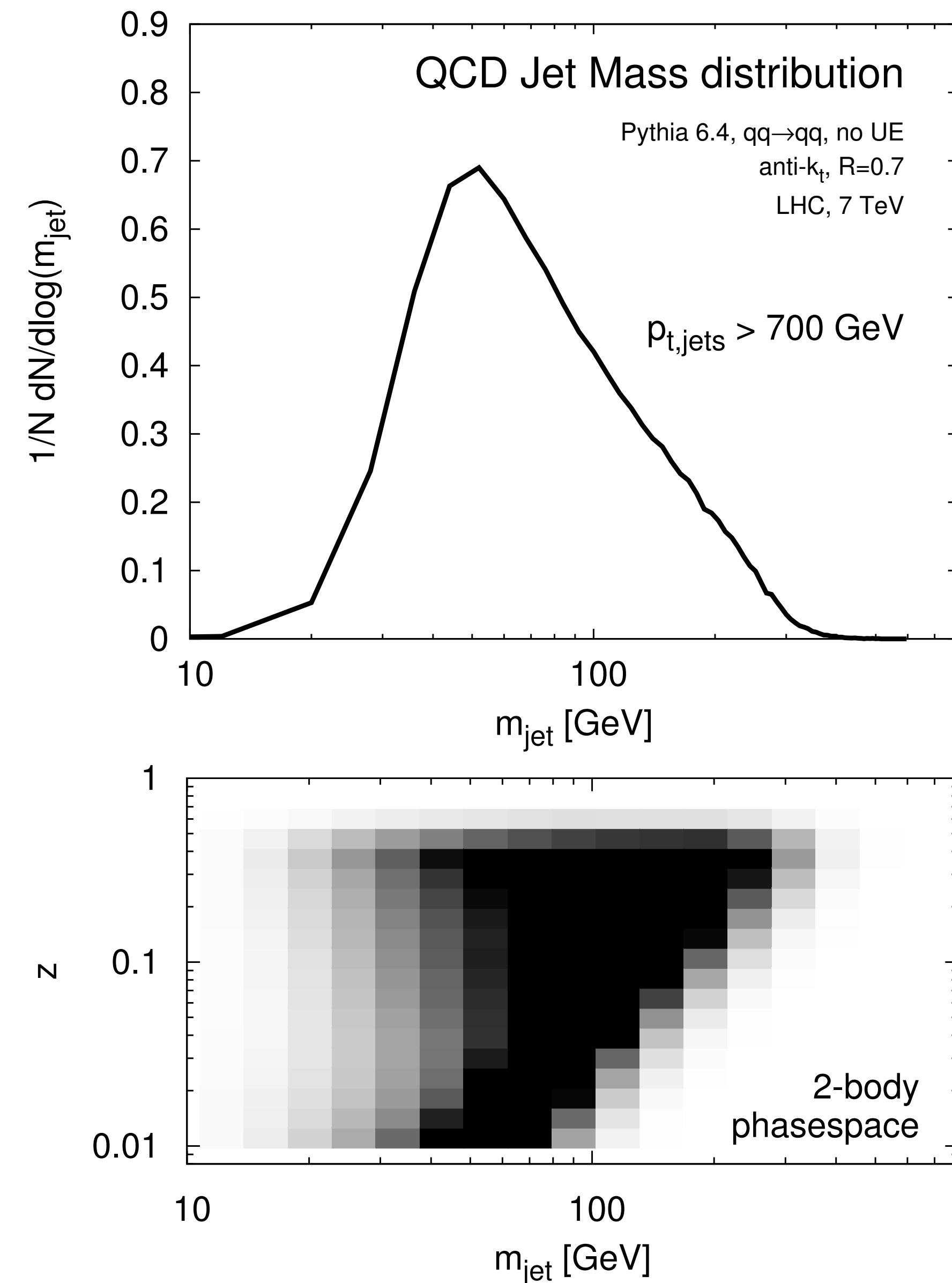


QCD jet mass distribution has the approximate

$$\frac{dN}{d \ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

Work from '80s and '90s

# Inside the jet mass



QCD jet mass distribution has the approximate

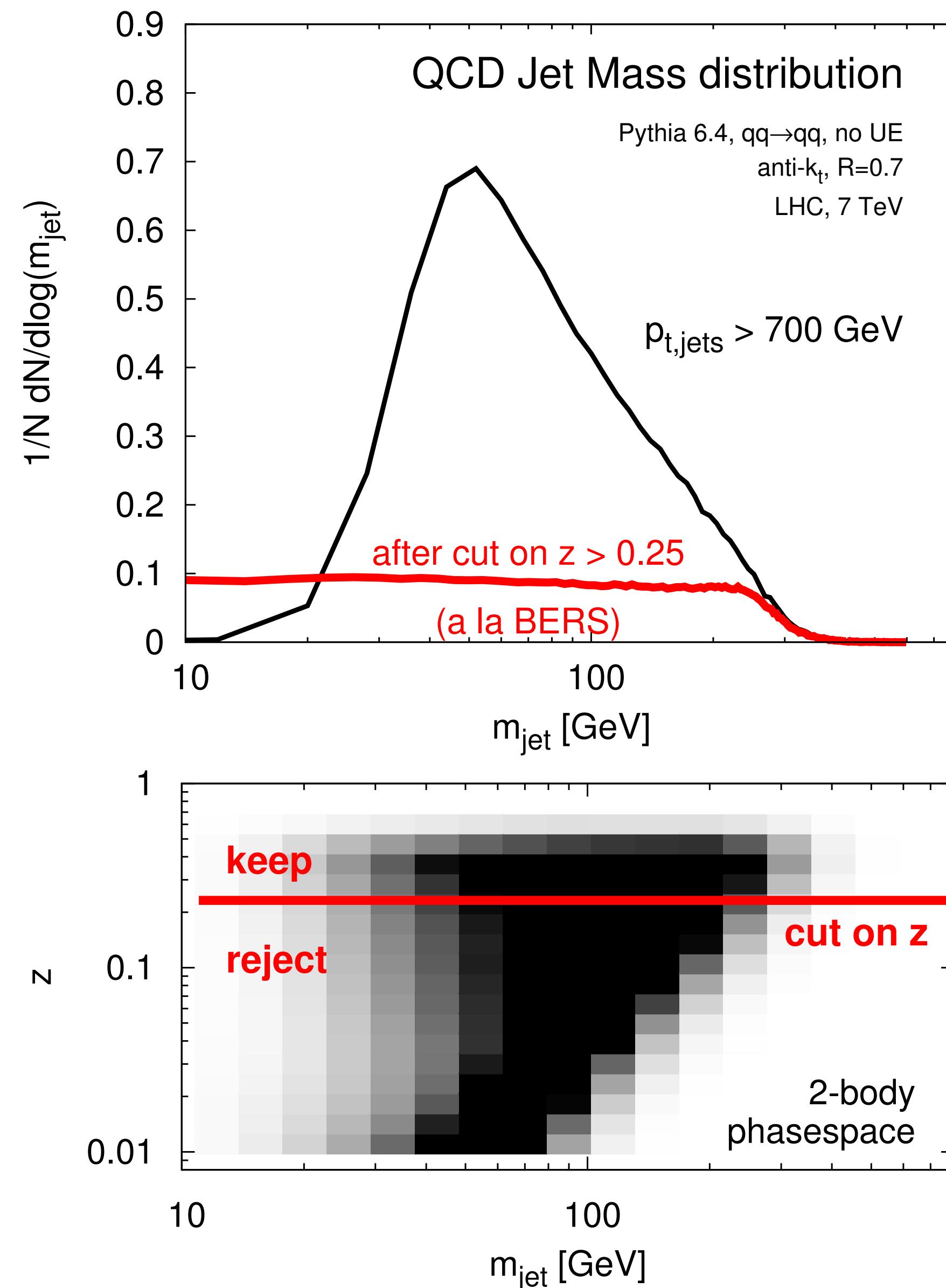
$$\frac{dN}{d \ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

Work from '80s and '90s

The logarithm comes from integral over soft divergence of QCD:

$$\int \frac{1}{\frac{m^2}{p_t^2 R^2}} \frac{dz}{z}$$

# Inside the jet mass



QCD jet mass distribution has the approximate

$$\frac{dN}{d \ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

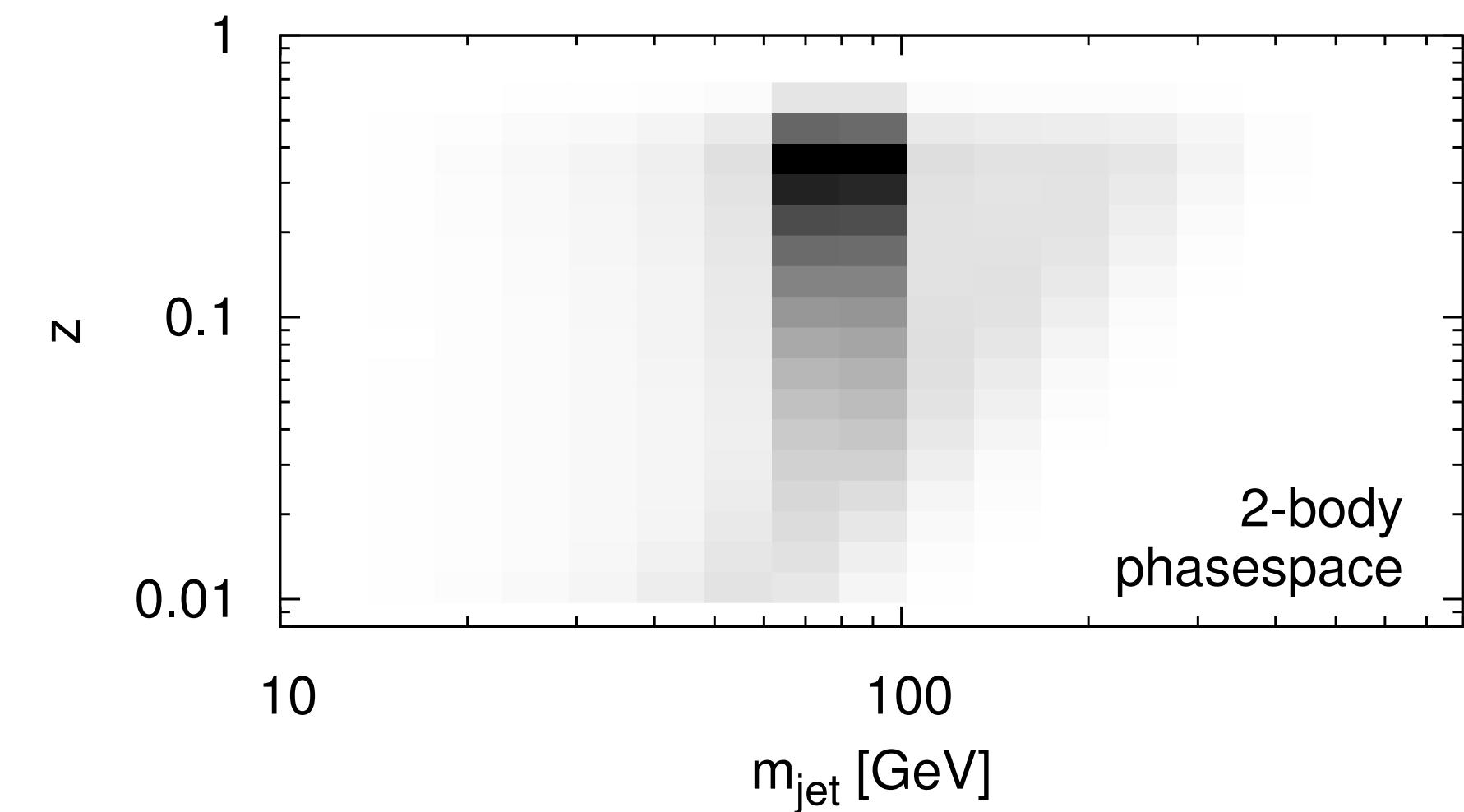
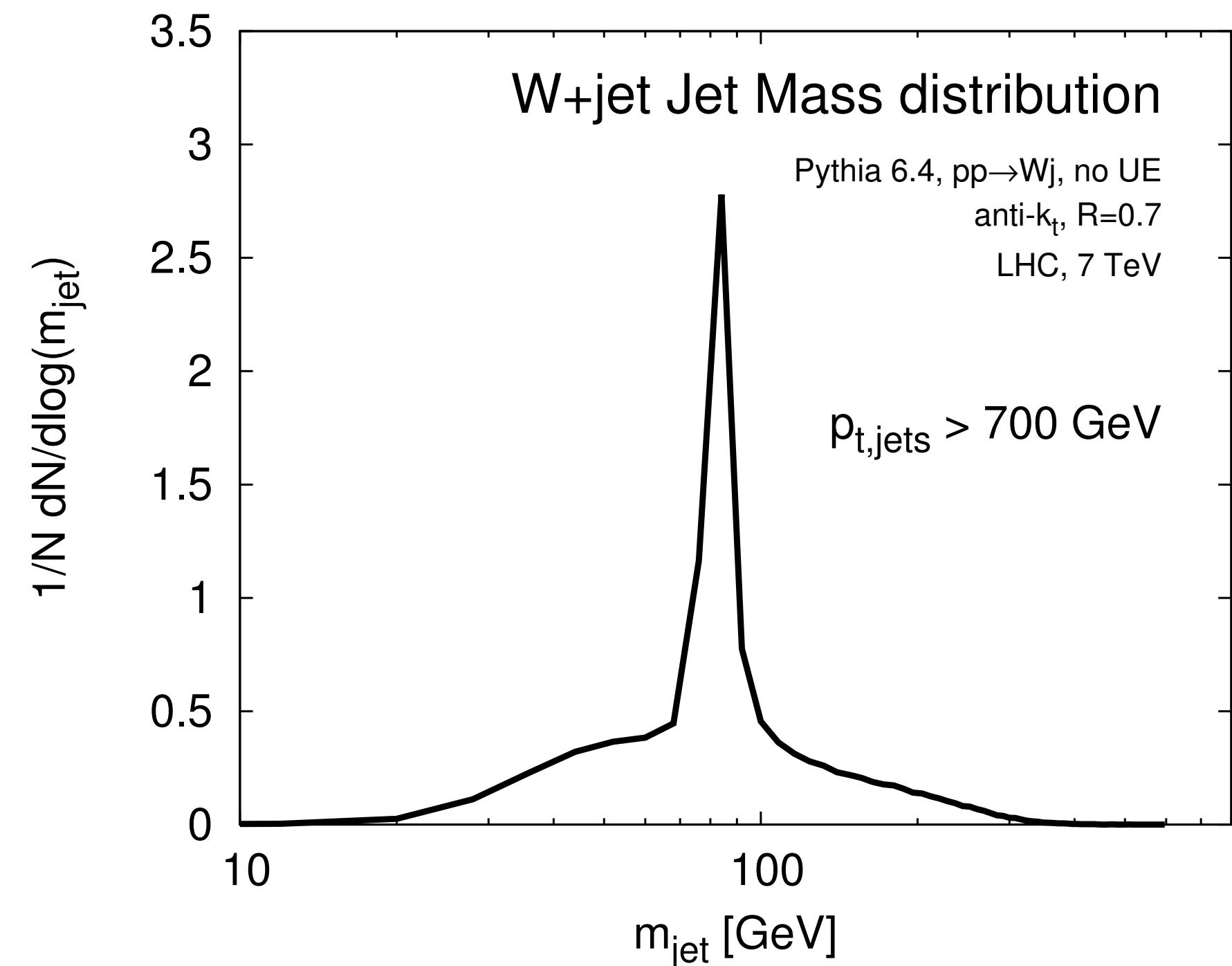
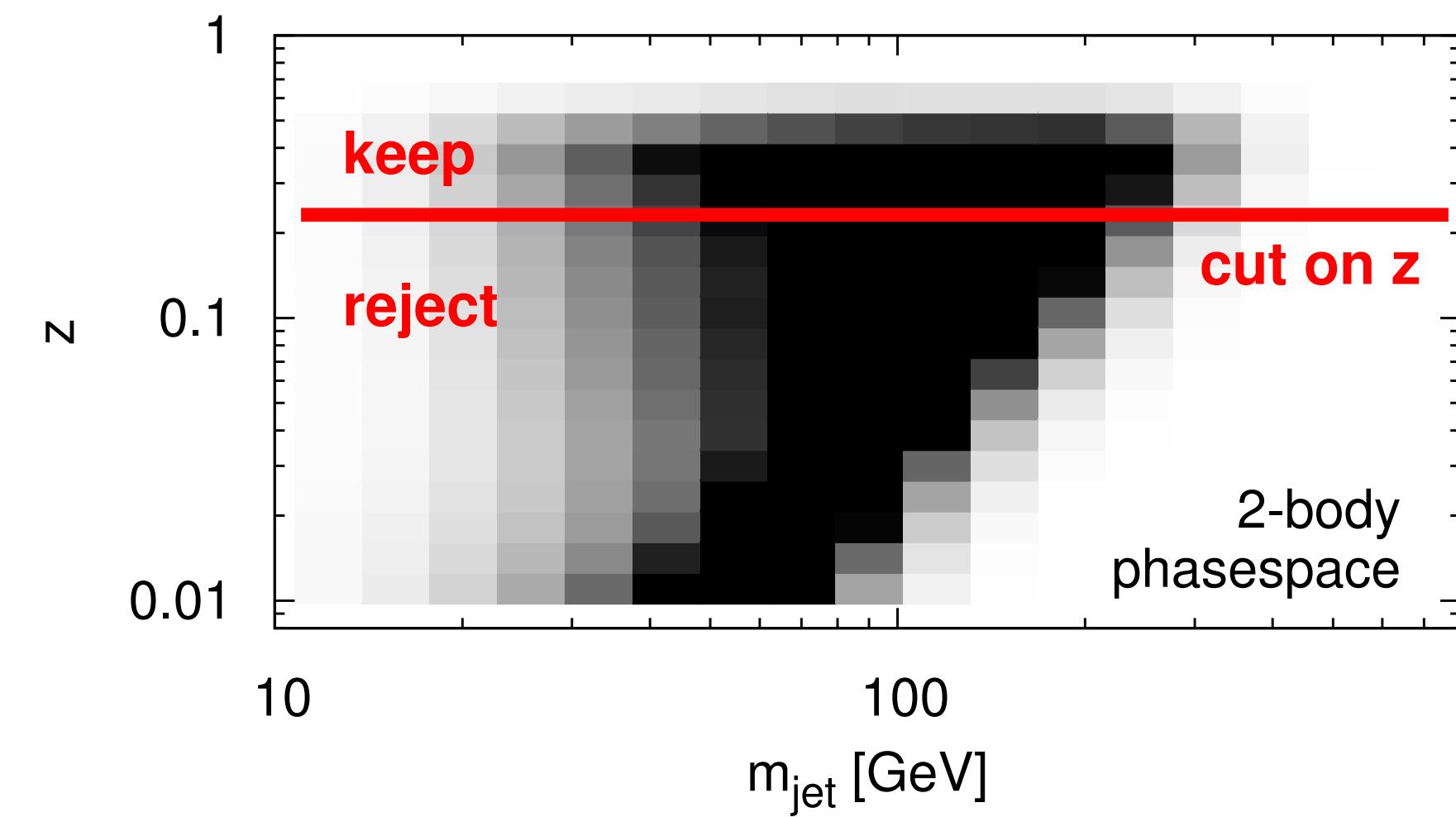
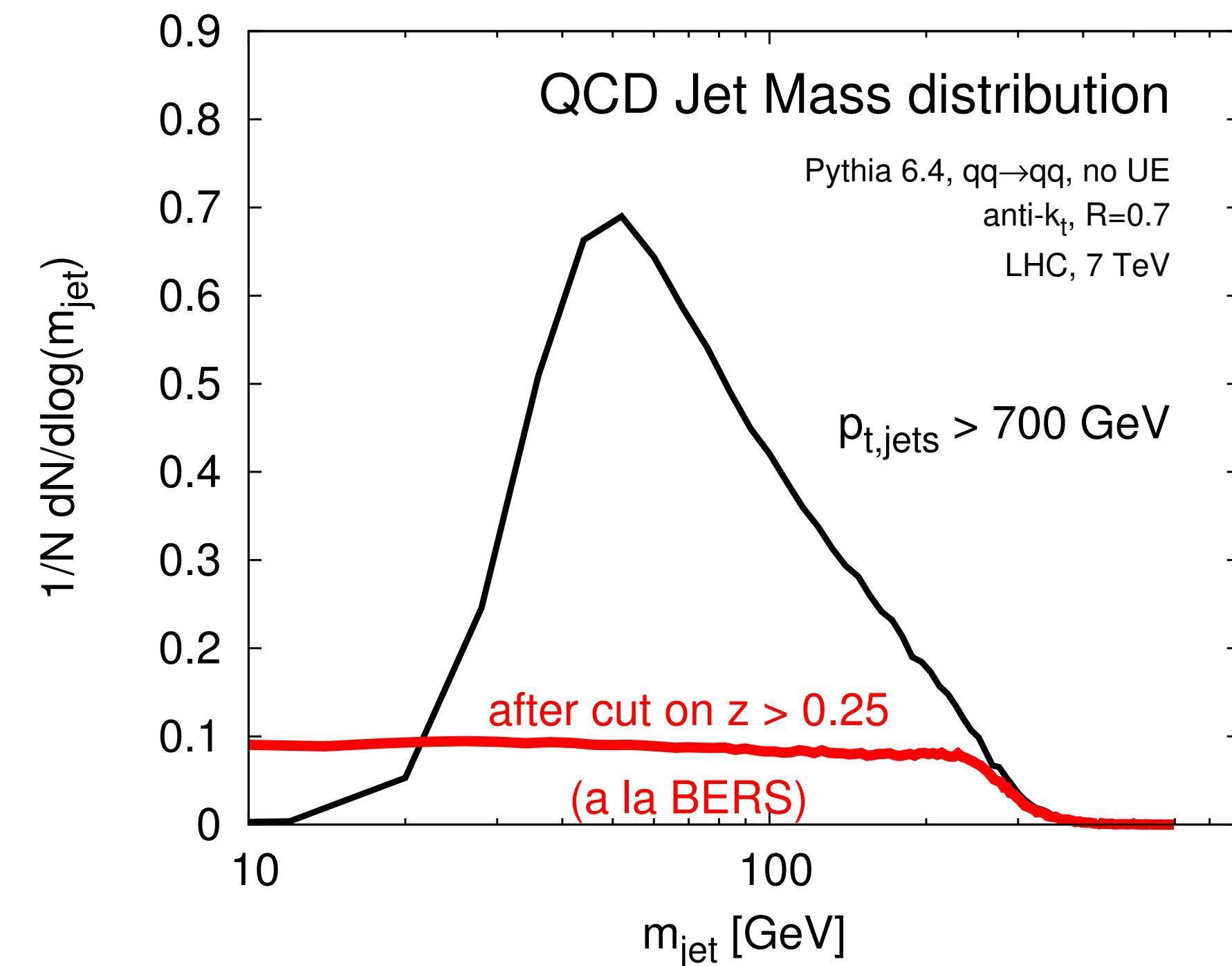
Work from '80s and '90s

The logarithm comes from integral over soft divergence of QCD:

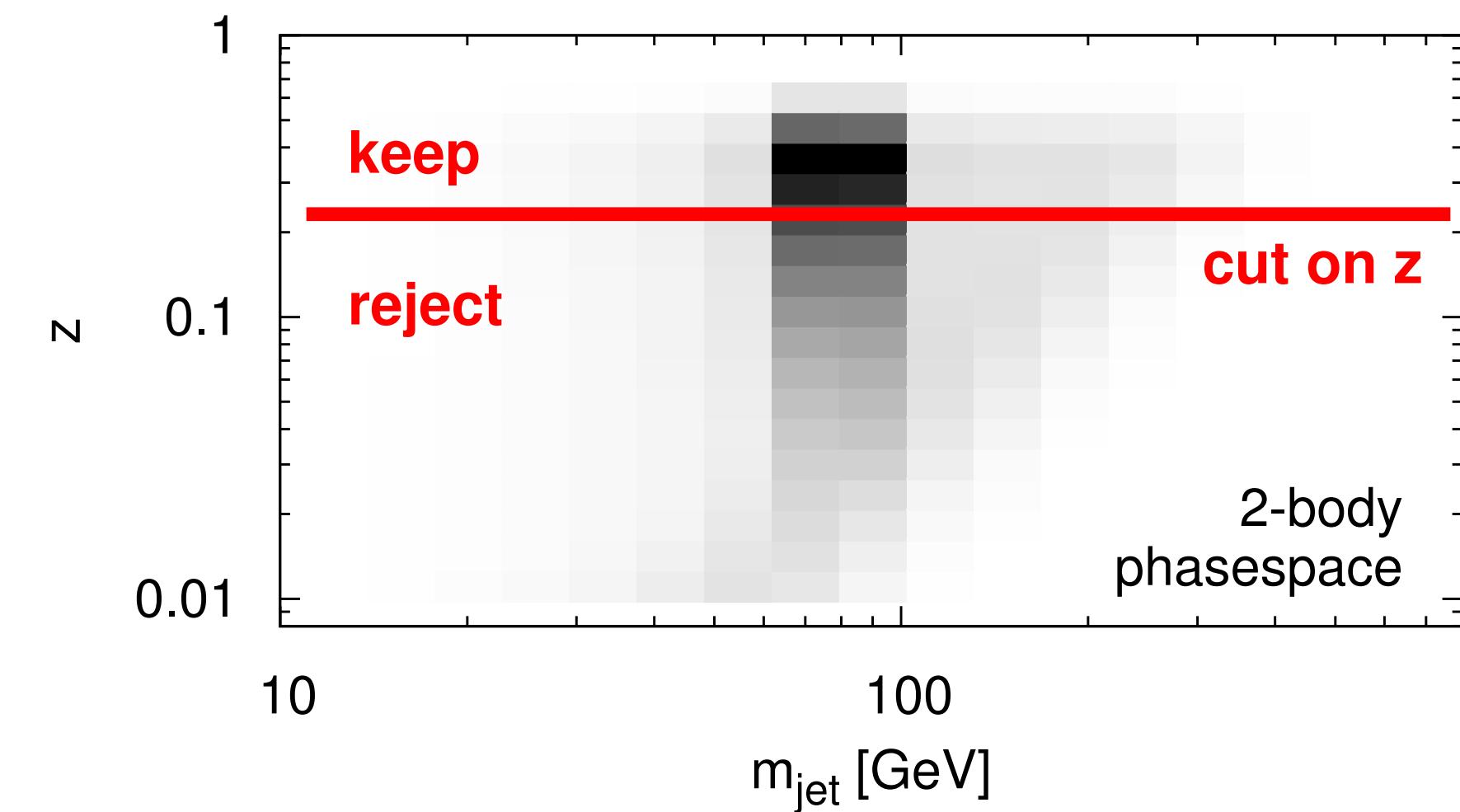
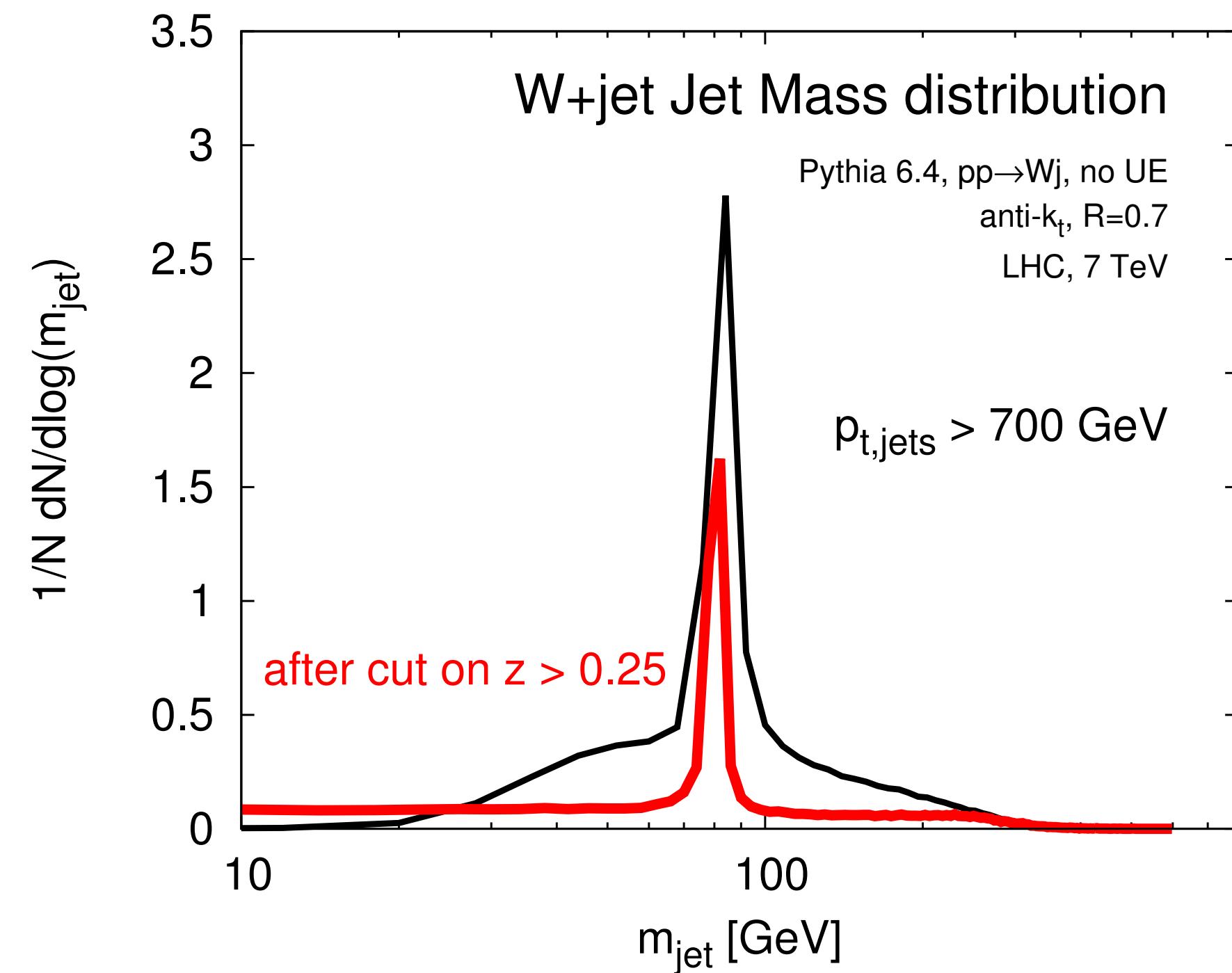
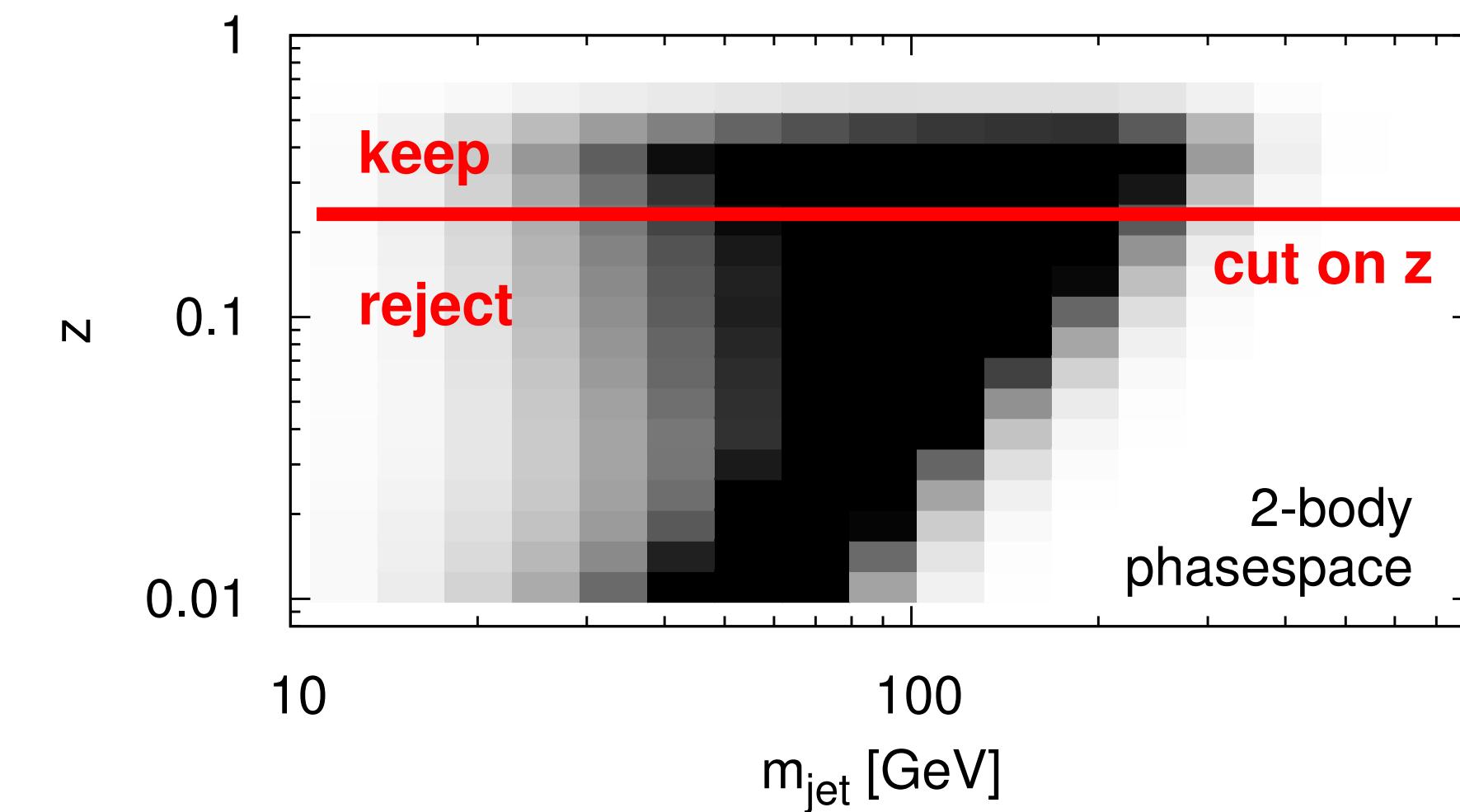
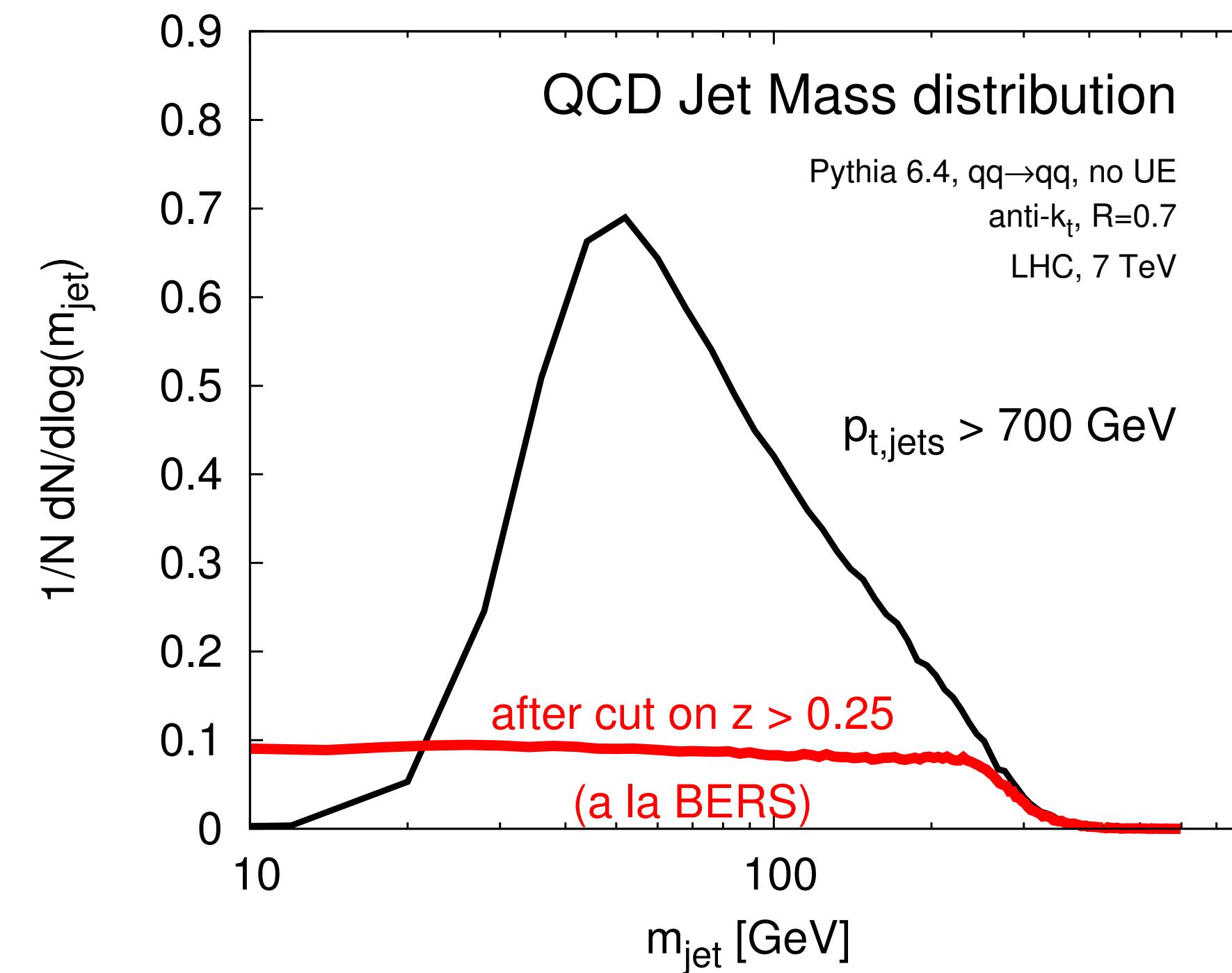
$$\int_{\frac{m^2}{p_t^2 R^2}}^{\frac{1}{2}} \frac{dz}{z}$$

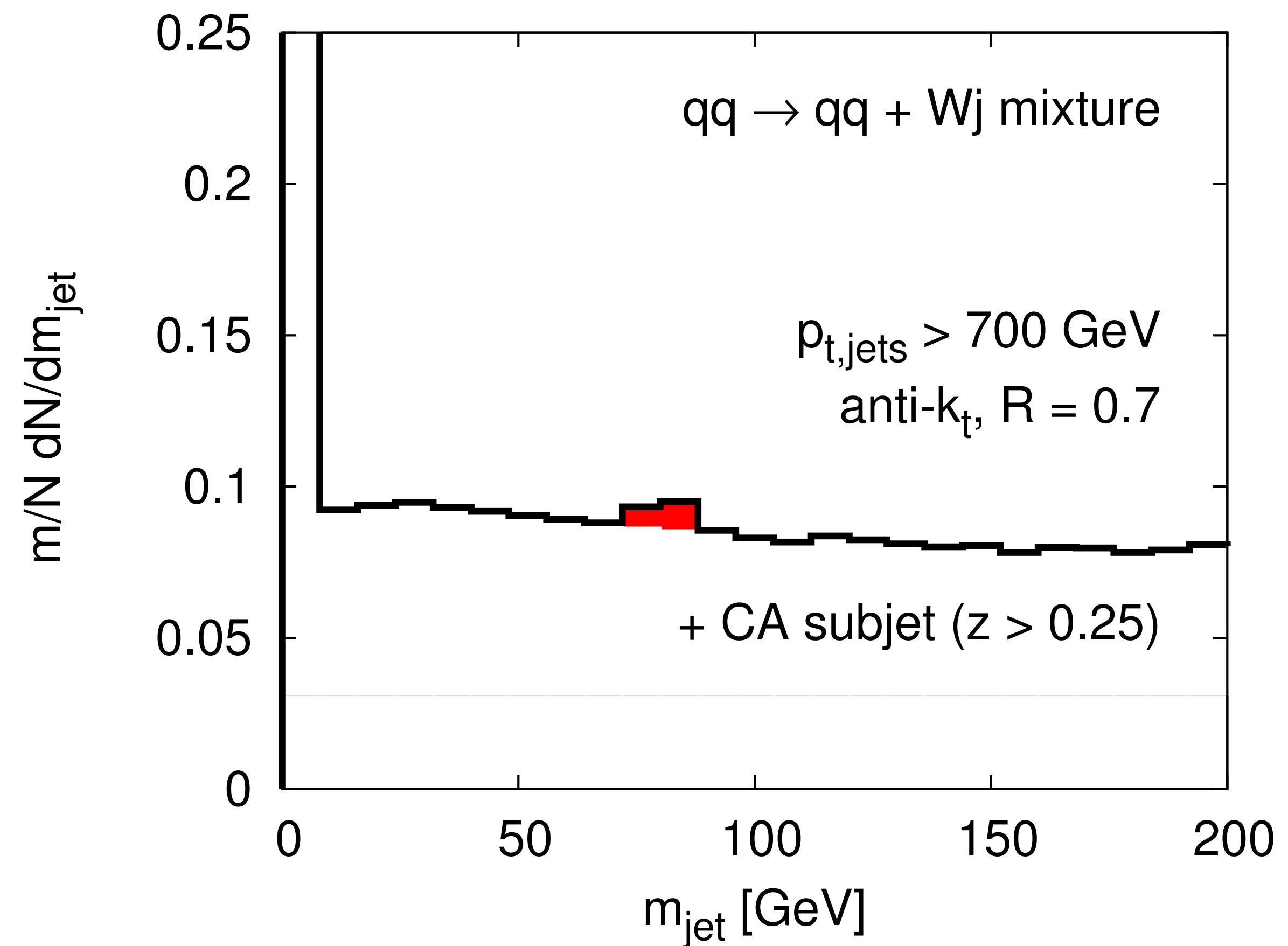
A hard cut on  $z$  reduces QCD background & simplifies its shape

# Inside the jet mass



# Inside the jet mass





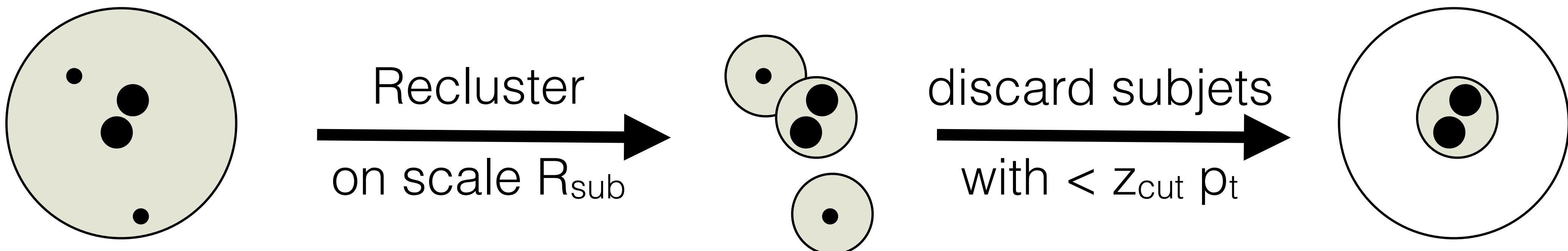
Signal + bkgd  
after cut on  $z$

cut on  $z$  helps “tag” the W-boson signal

# Tagging two-prong decays

## 2. identifying hard kinematics

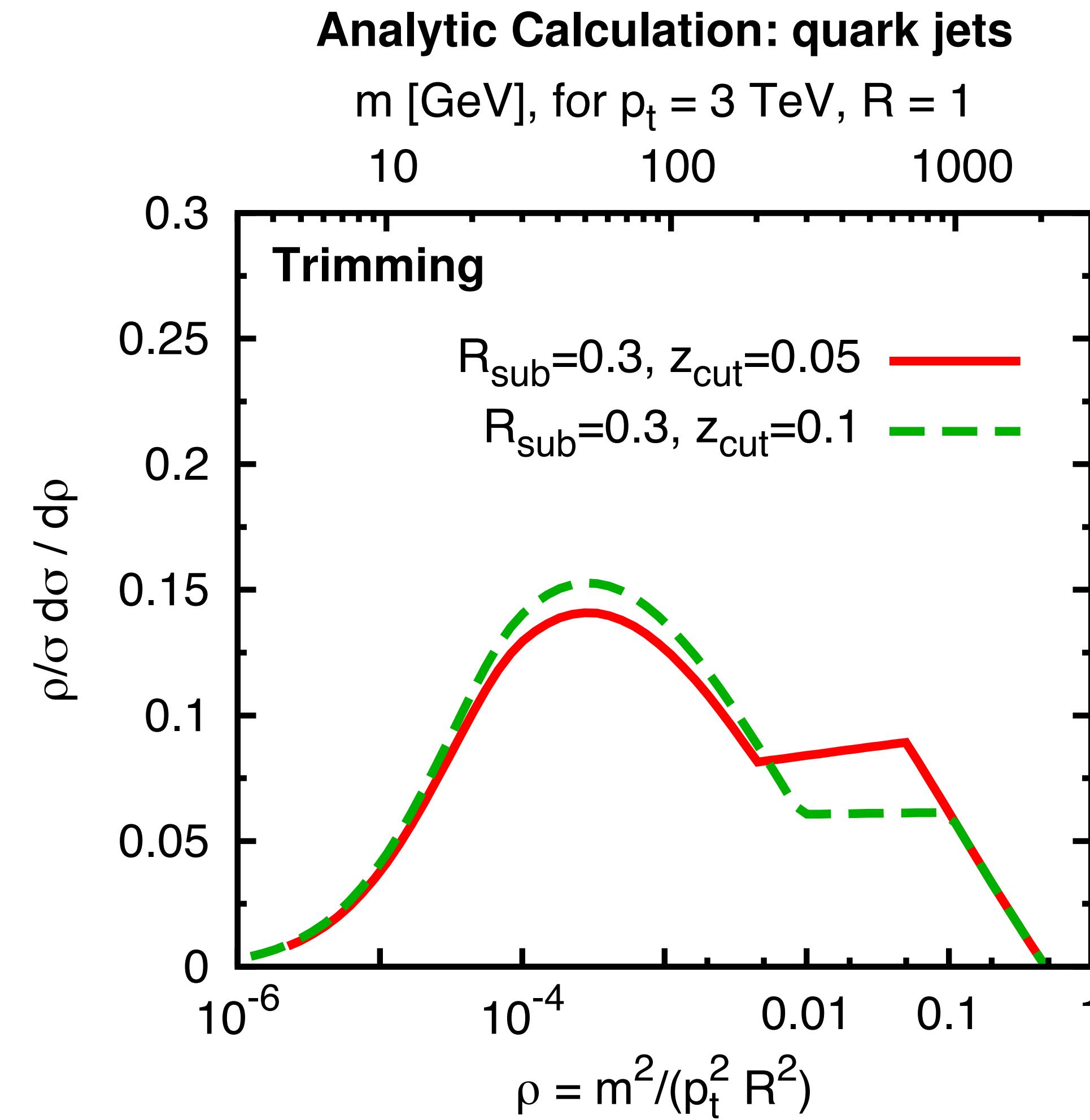
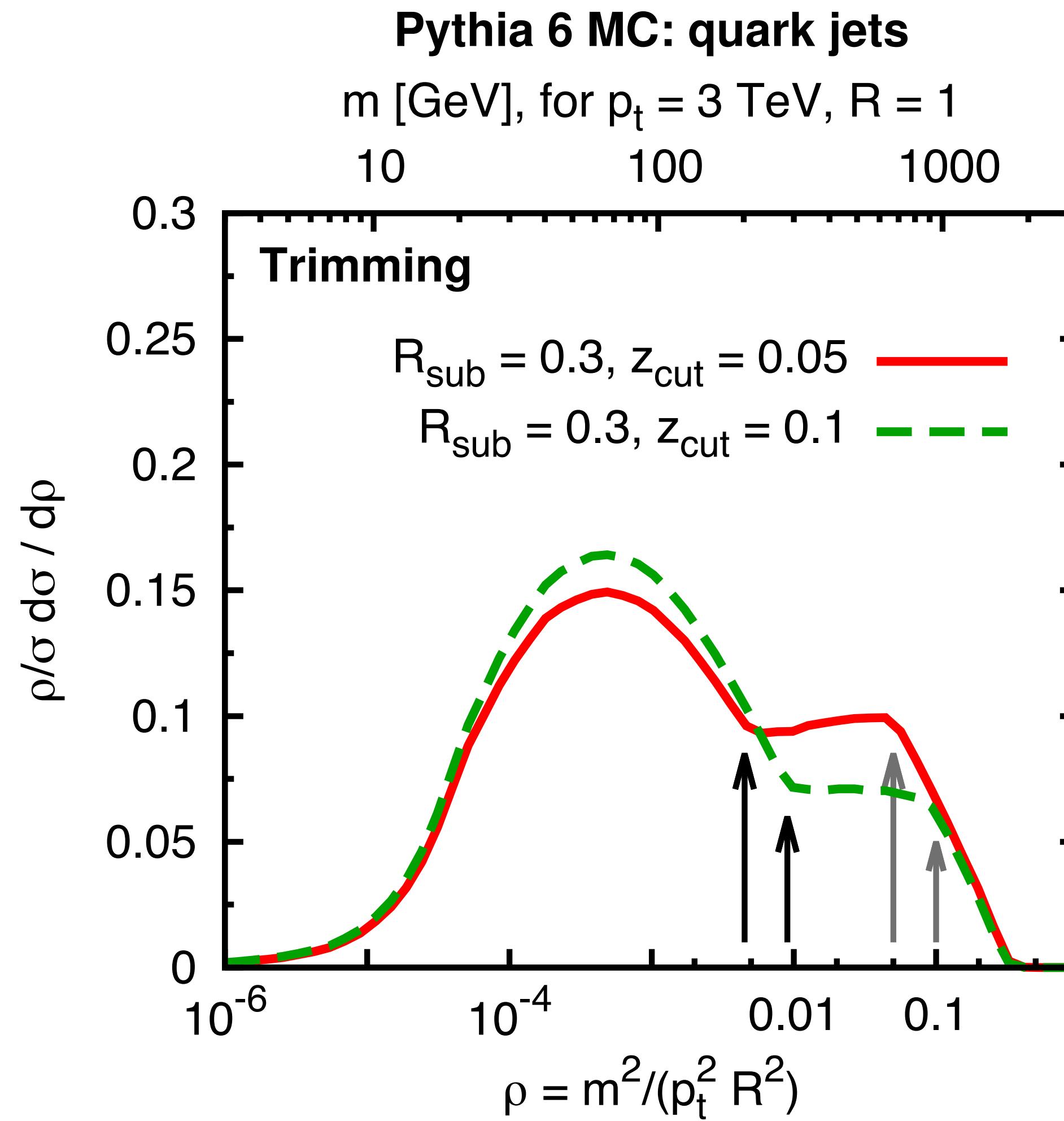
- Take all particles in a jet of radius  $R$
- Recluster them into subjets with a jet definition using  $R_{sub} < R$
- Keep only subjets with  $p_t^{subjet} > z_{cut} p_t^{jet}$
- Recombine them into a single jet



Krohn, Thaler & Wang [arXiv:0912.1342](https://arxiv.org/abs/0912.1342)

widely used in ATLAS

# Trimming: QCD jet mass distribution



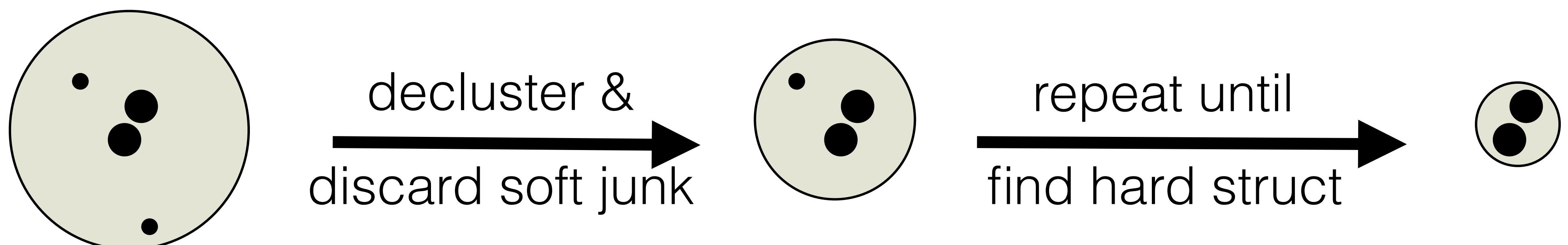
[arXiv:1307.0007](https://arxiv.org/abs/1307.0007)  
Dasgupta,  
Fregoso,  
Marzani & GPS

3 scales: two kinks (from  $z_{\text{cut}}$  &  $R_{\text{sub}}$ ) and one Sudakov peak

# Soft Drop ( $\beta=0$ variant)

1. Recluster jet with Cambridge/Aachen algorithm
2. Undo last step of clustering to give two subjets  $i, j$  (with  $p_{ti} < p_{tj}$ )
3. If  $p_{ti} > z_{cut}(p_{ti} + p_{tj})$  stop
4. otherwise discard  $i$ , go back to step 2 to decluster  $j$

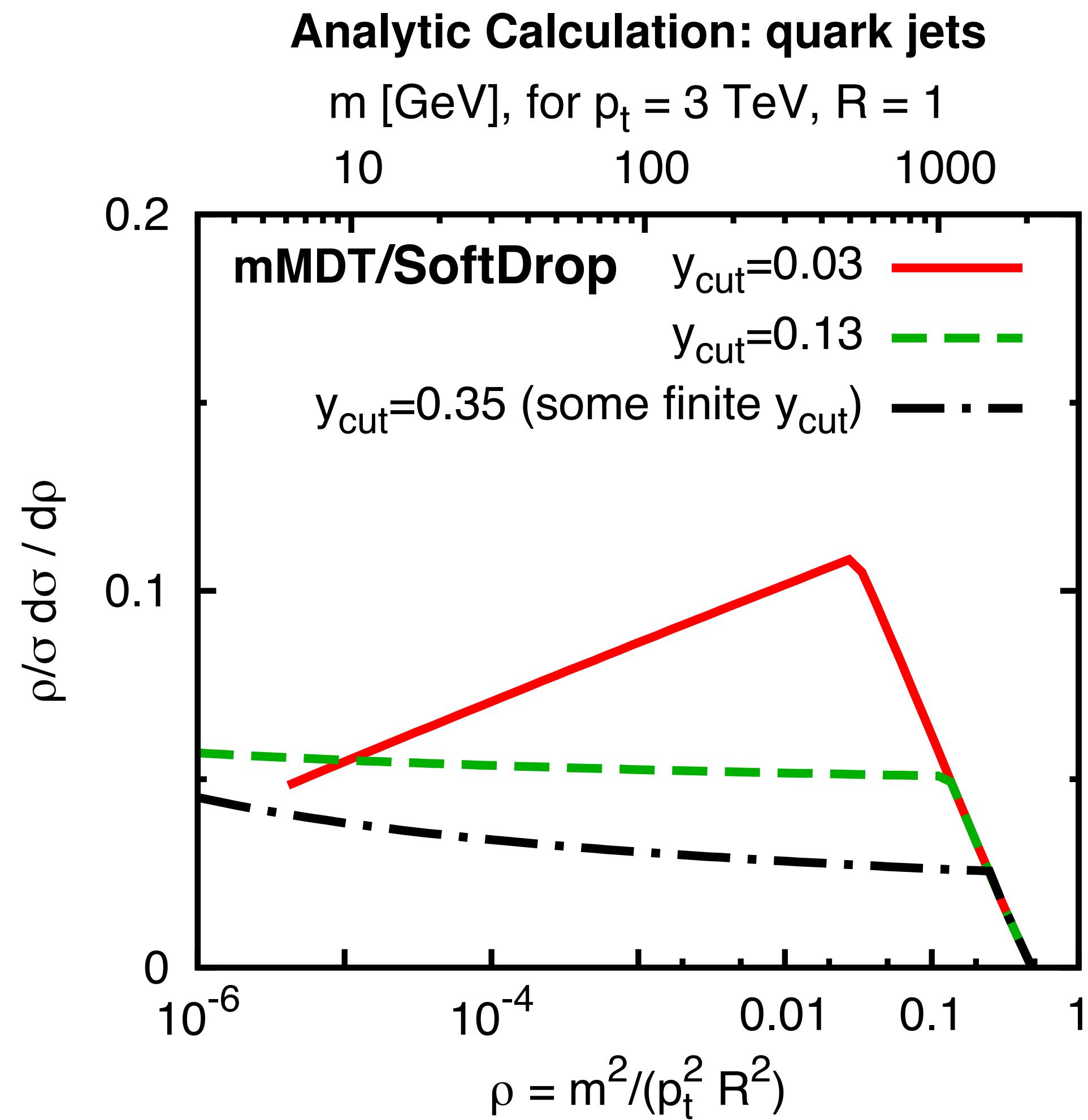
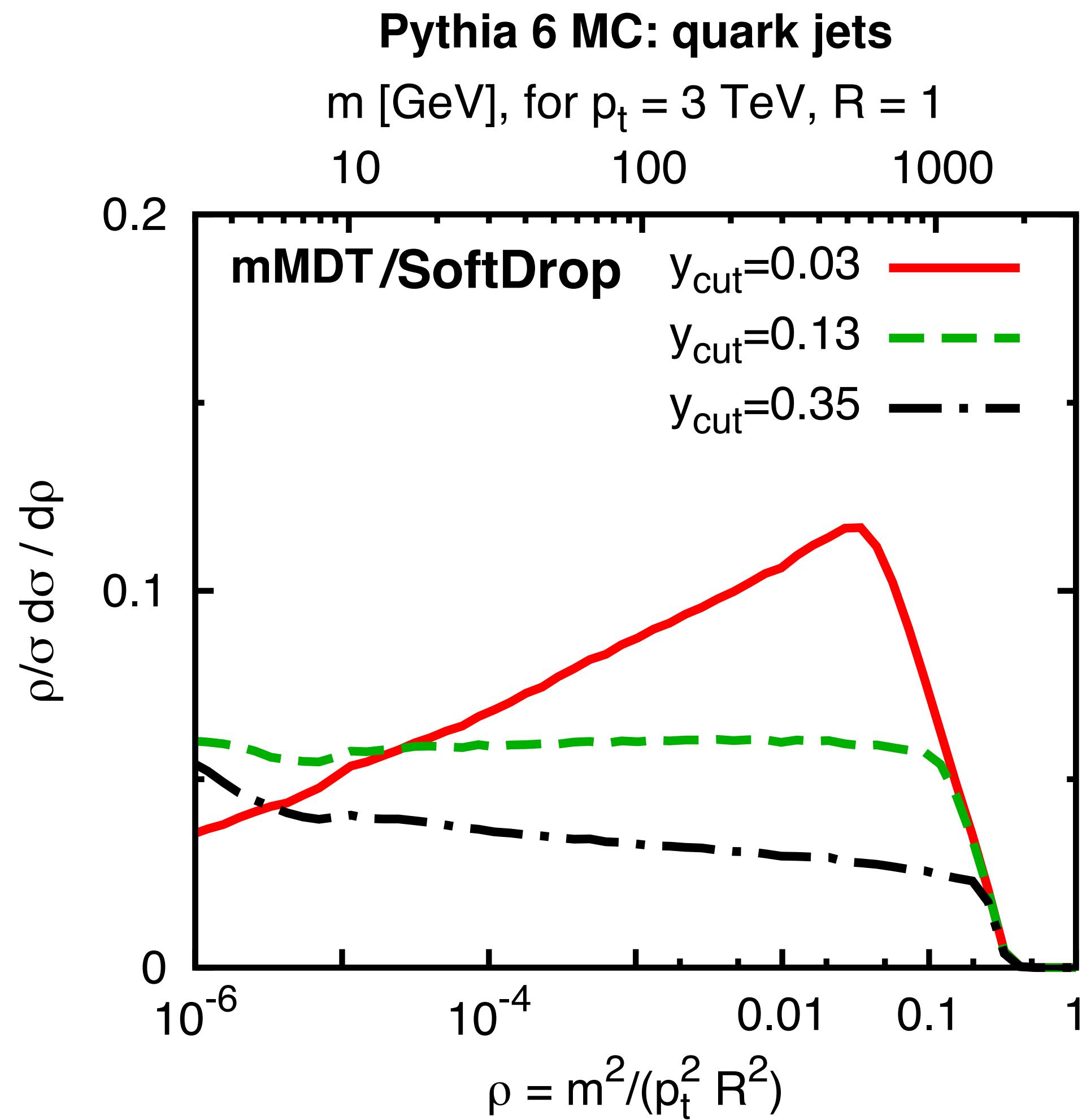
**no manually specified R<sub>cut</sub>**  
**uses internal structure of jet to auto-zoom into right angular scale**



Dasgupta, Fregoso, Marzani & GPS [[arXiv:1307.0007](https://arxiv.org/abs/1307.0007)],  
descended from mass-drop tagger, Butterworth, Davison, Rubin & GPS [[arXiv:0802.2470](https://arxiv.org/abs/0802.2470)]

widely used  
in CMS

# Soft Drop ( $\beta=0$ variant)



smooth distribution, apart from one kink (induced by  $z_{\text{cut}}$ )

## Plain jet mass

$$\Sigma(\rho) = e^{-D(\rho)} \cdot \frac{e^{-\gamma_E D'(\rho)}}{\Gamma(1 + D'(\rho))} \cdot \mathcal{N}(\rho)$$

$$D(\rho) \simeq \frac{\alpha_s C_F}{\pi} \left[ \frac{1}{2} \ln^2 \frac{1}{\rho} - \frac{3}{4} \ln \frac{1}{\rho} + \mathcal{O}(1) \right], \quad (\text{fixed coupling approx.}),$$

**Note  $\ln^2 \rho$**

## SoftDrop(mMDT) jet mass

$$\Sigma^{(\text{mMDT})}(\rho) = \exp \left[ -\frac{\alpha_s C_F}{\pi} \left( \ln \frac{y_{\text{cut}}}{\rho} \ln \frac{1}{y_{\text{cut}}} - \frac{3}{4} \ln \frac{1}{\rho} + \frac{1}{2} \ln^2 \frac{1}{y_{\text{cut}}} \right) \right], \quad (\text{for } \rho < y_{\text{cut}}),$$

**Note  $\ln \rho$**

in fixed-coupling approximation for simplicity

1. Recluster jet with Cambridge/Aachen algorithm
2. Undo last step of clustering to give two subjets  $i, j$   
(with  $p_{ti} < p_{tj}$ )

$$3. \text{ If } p_{ti} > z_{cut} (p_{ti} + p_{tj}) \left( \frac{\Delta R_{ij}}{R_0} \right)^\beta$$

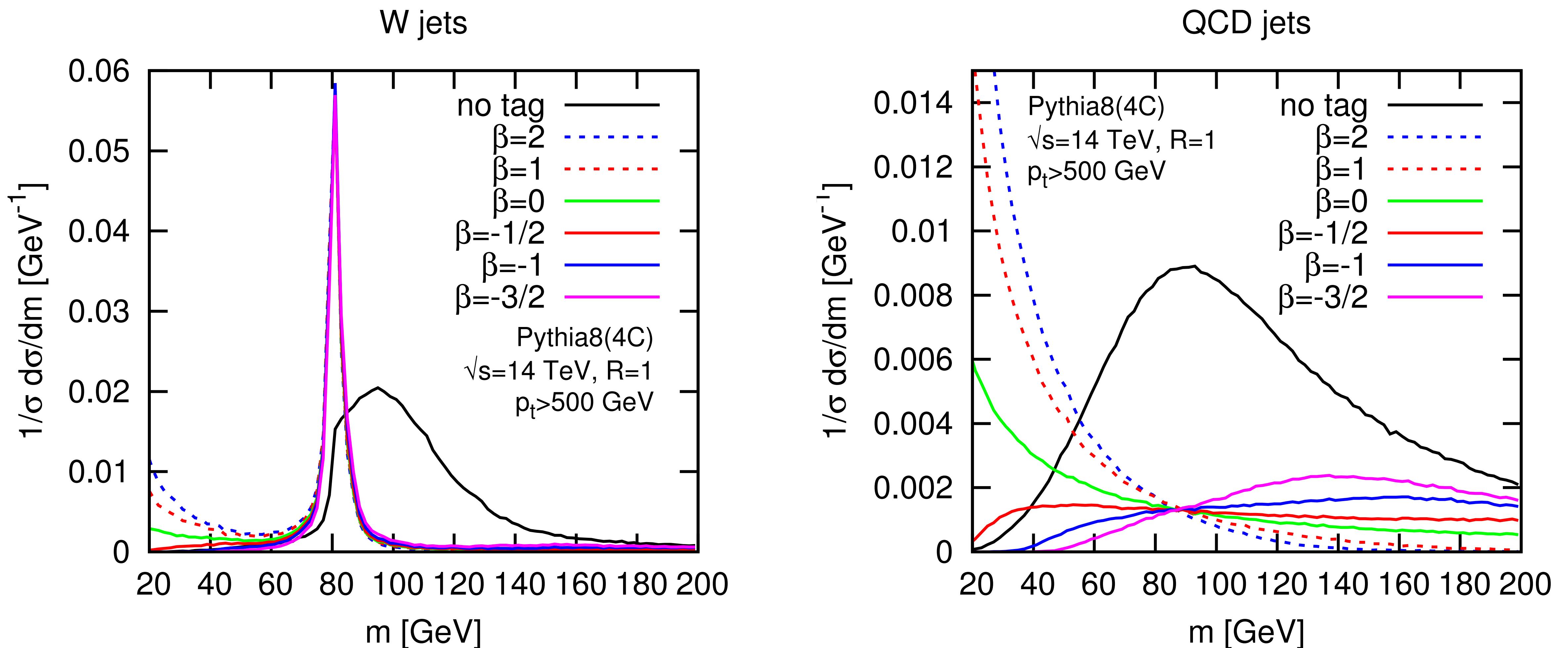
4. otherwise discard  $i$ , and go back to step 2, now with  $j$
5. If  $j$  is a singleton either discard it (tagging mode) or leave  $j$  as the final “groomed” jet

$\beta > 0$  cuts less aggressively as angles decrease

$\beta < 0$  cuts more aggressively as angles decrease

Larkoski, Marzani, Soyez & Thaler [arXiv:1402.2657](https://arxiv.org/abs/1402.2657)

# SoftDrop (with $\beta$ )

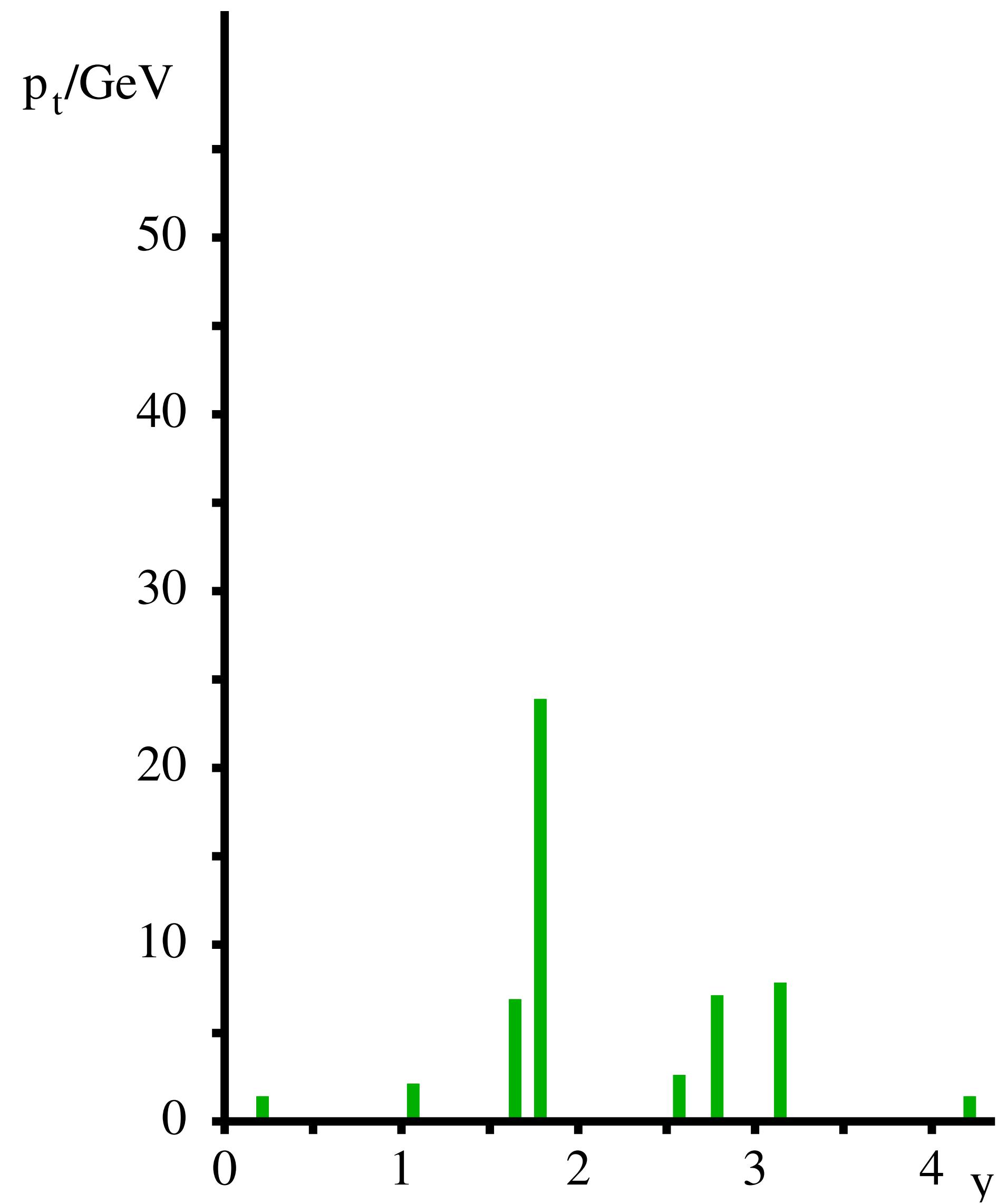


**NB:  $z_{\text{cut}}$  chosen to keep signal efficiency fixed at 35% for all  $\beta$**

# Tagging two-prong decays

3. why the Cambridge/  
Aachen algorithm?

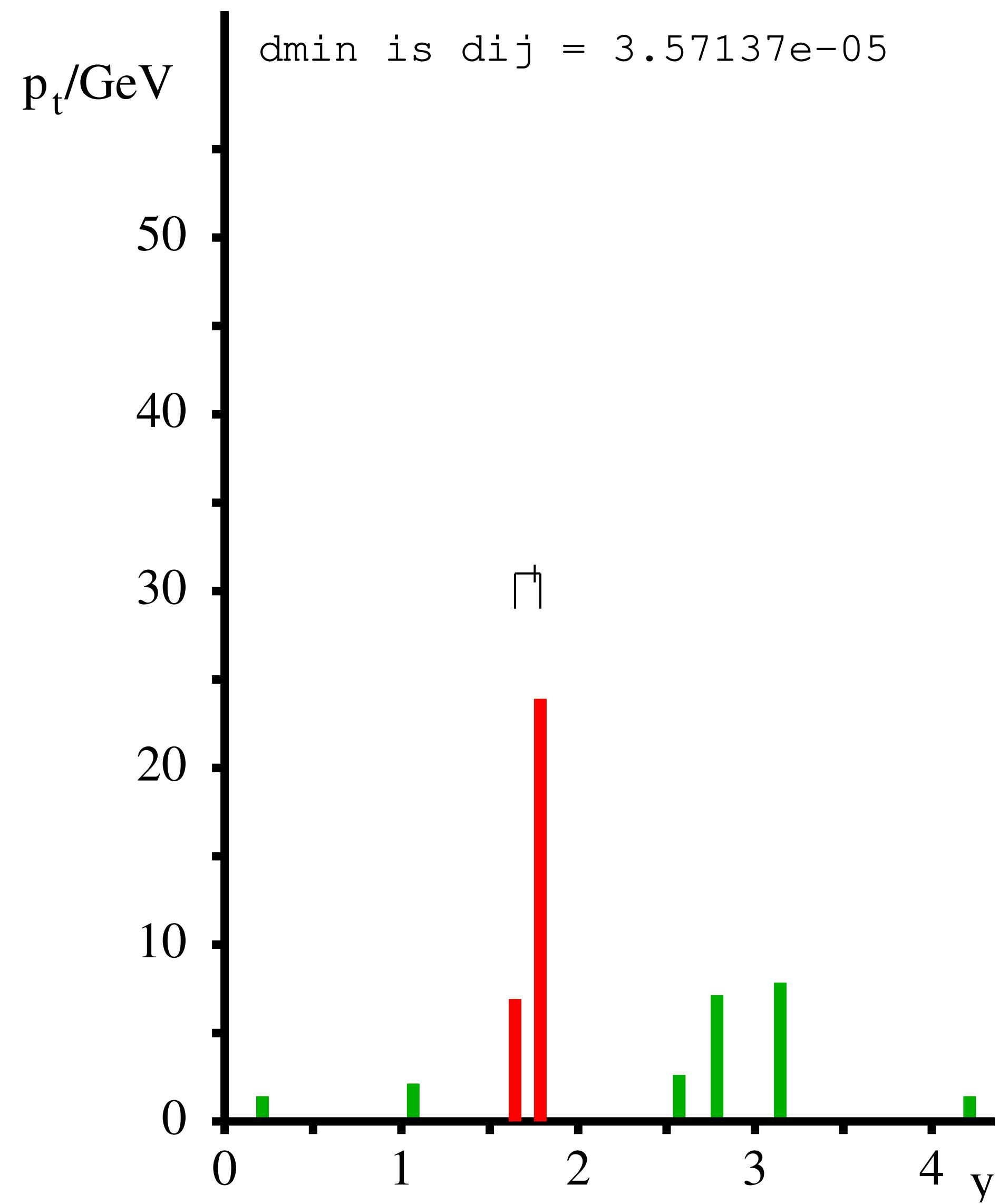
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

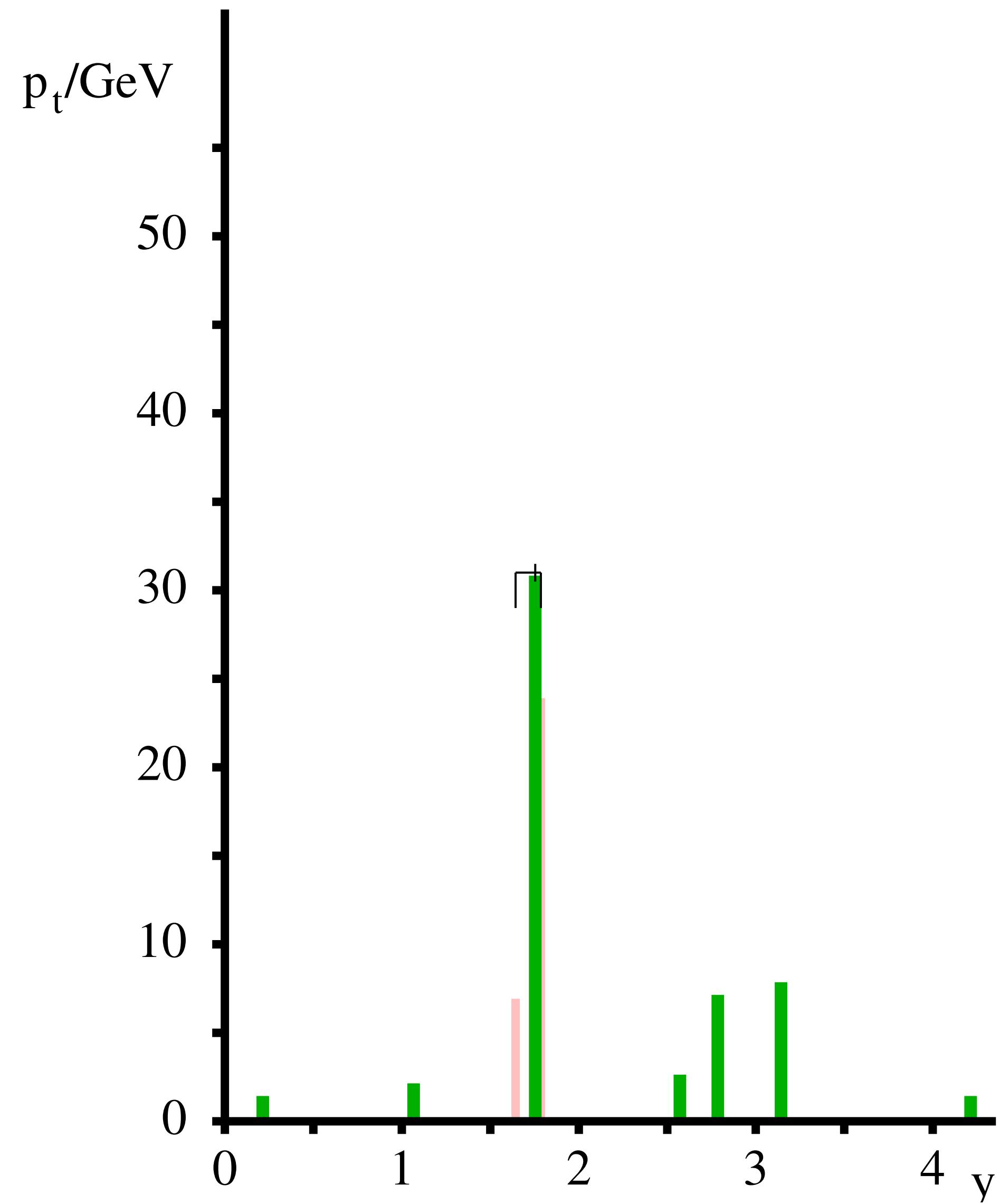
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

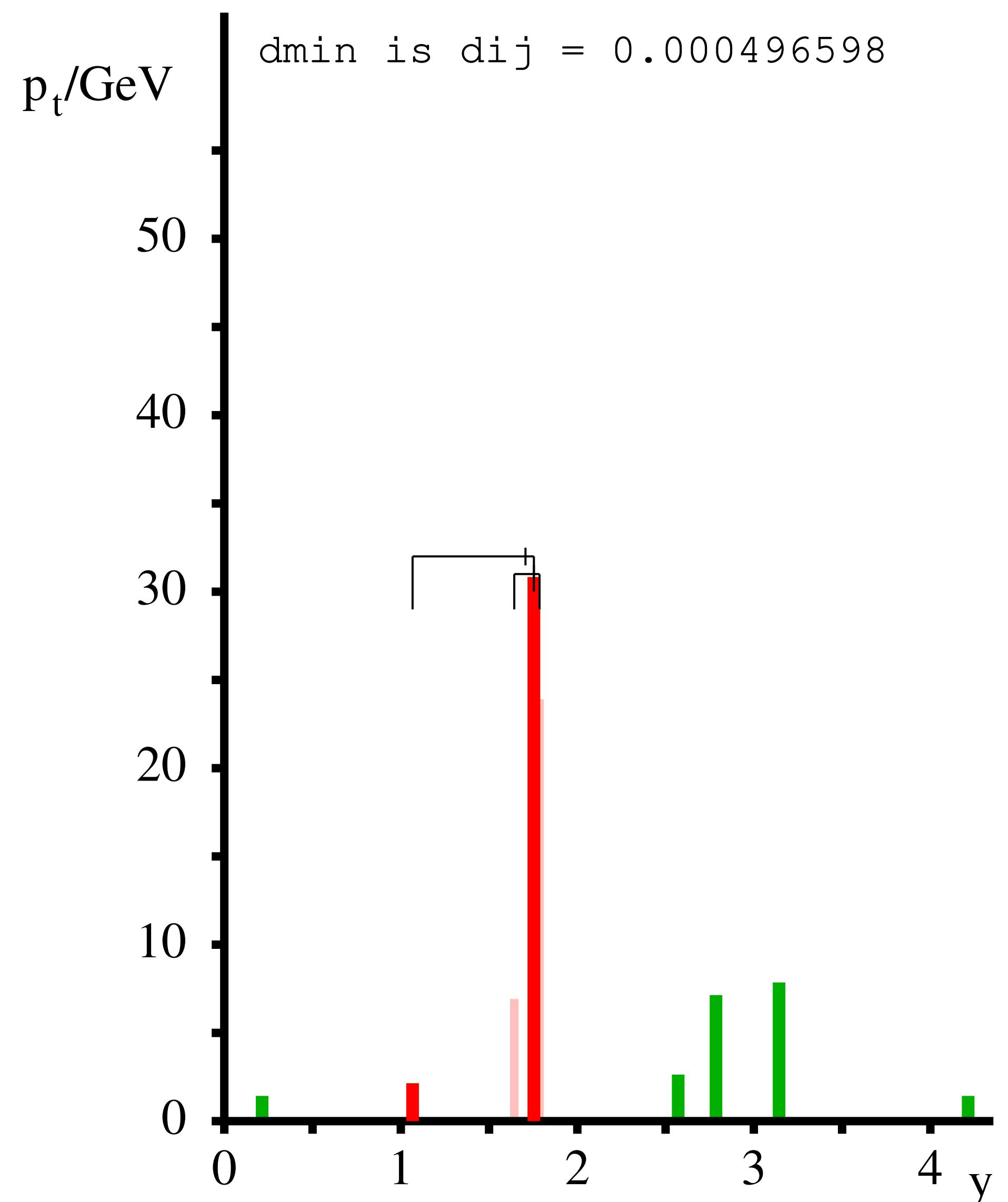
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

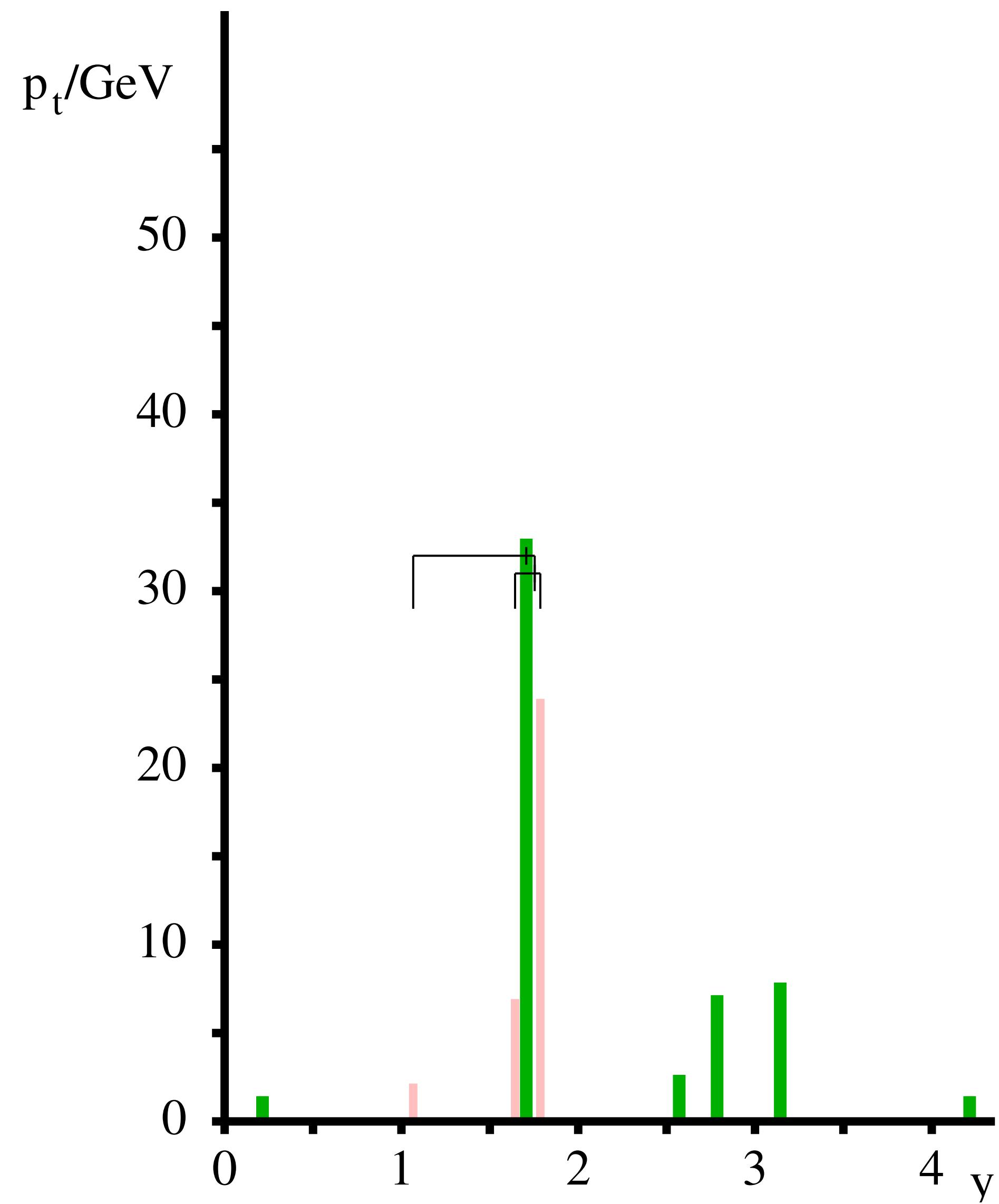
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

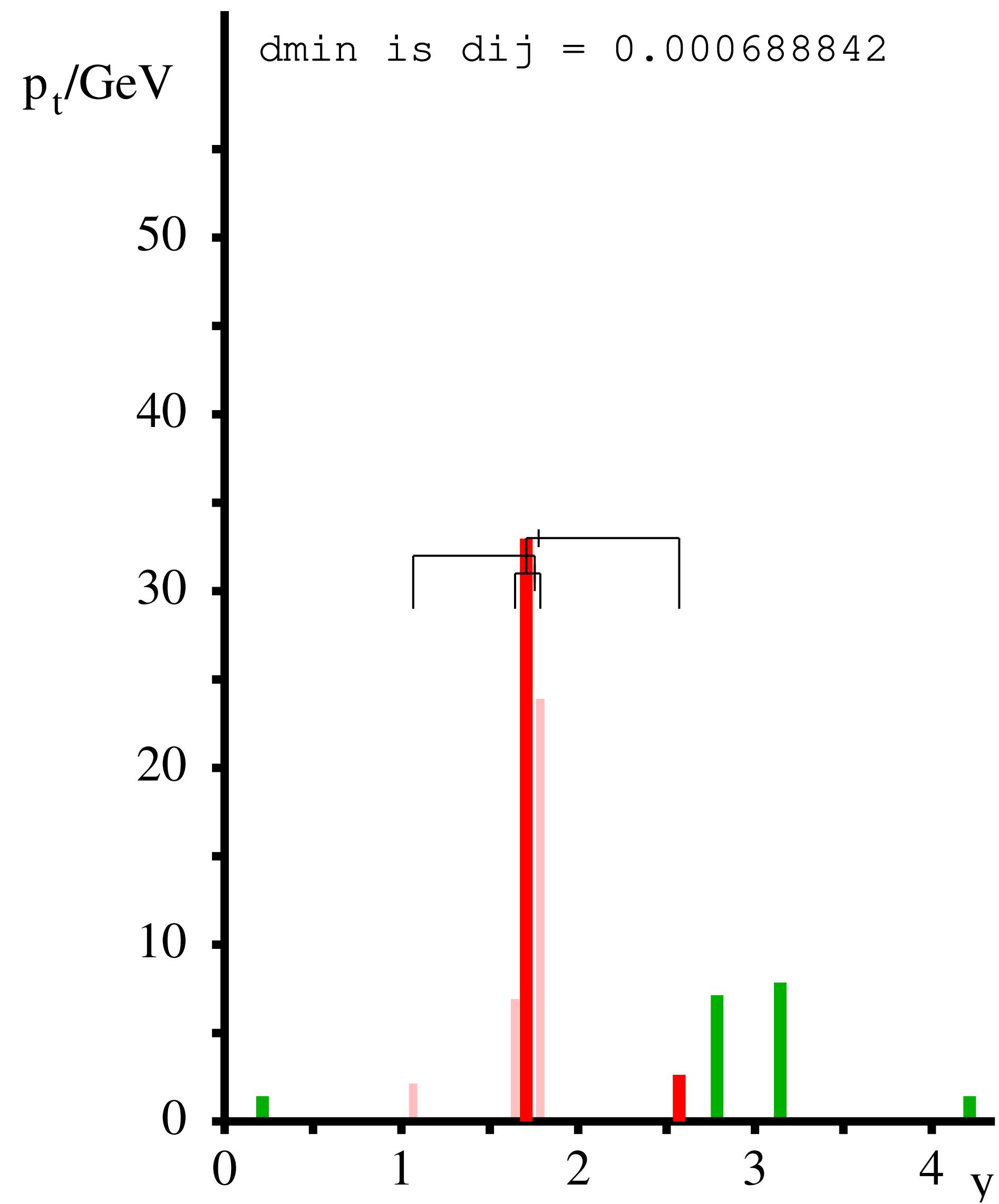
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

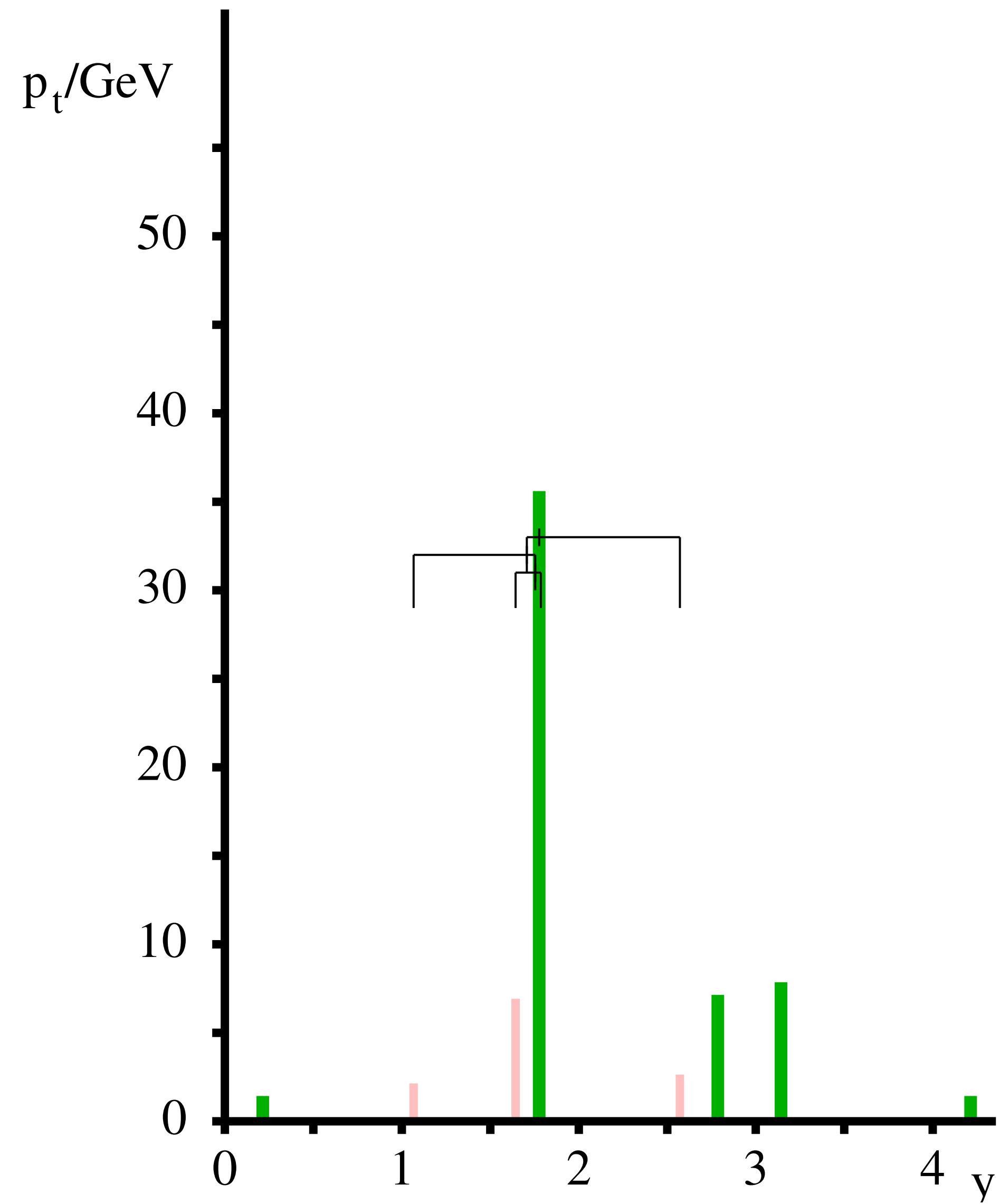
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

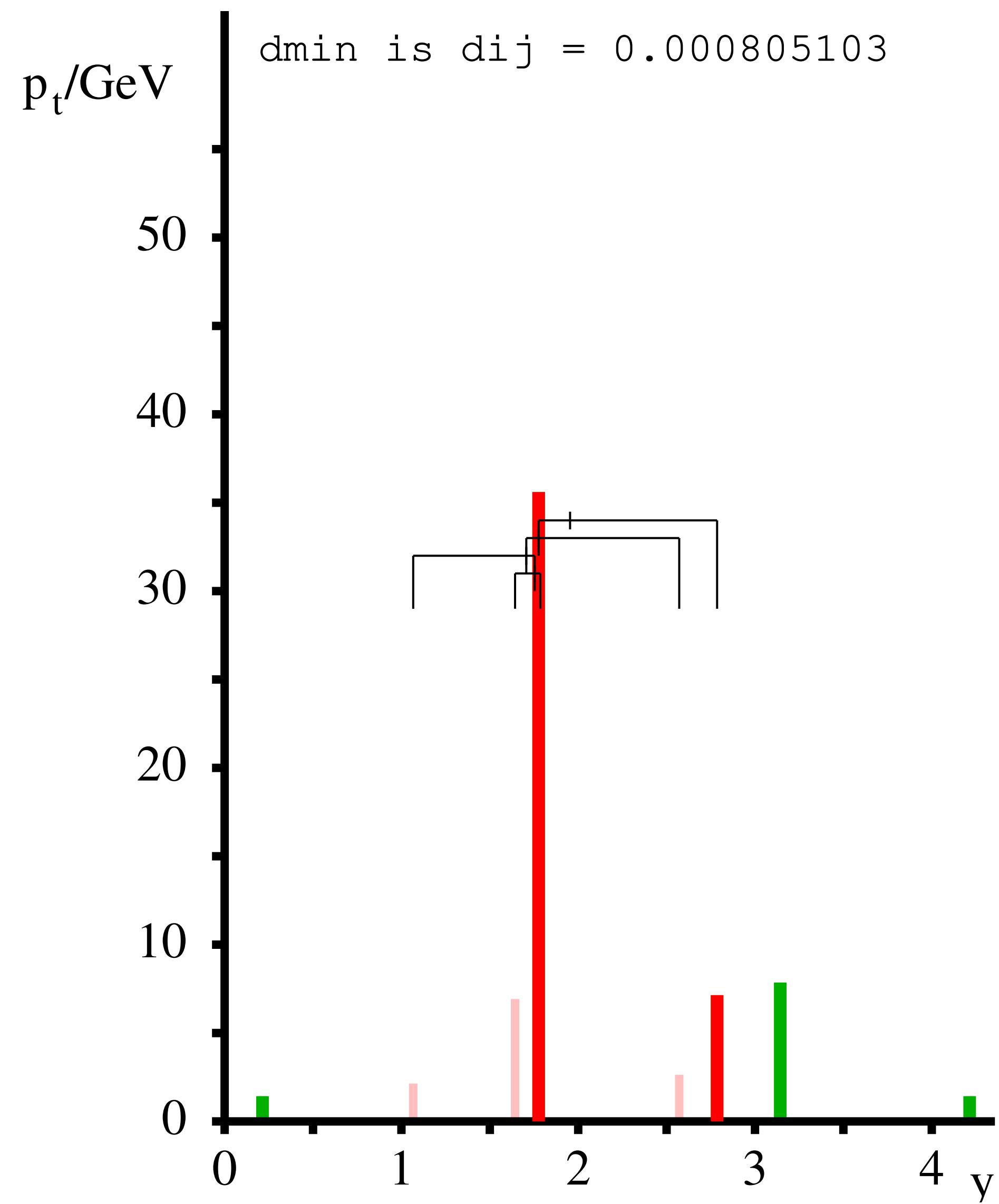
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

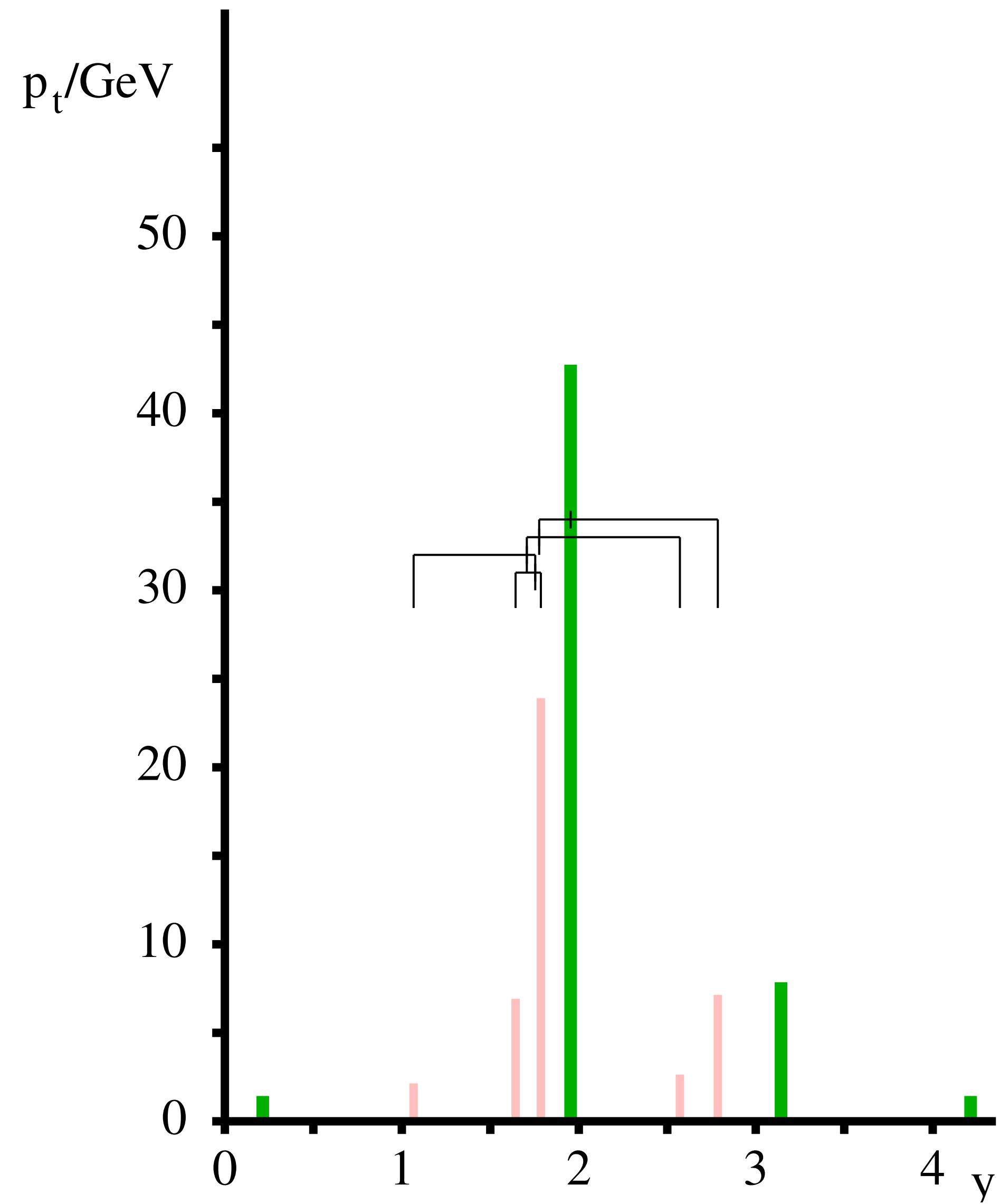
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

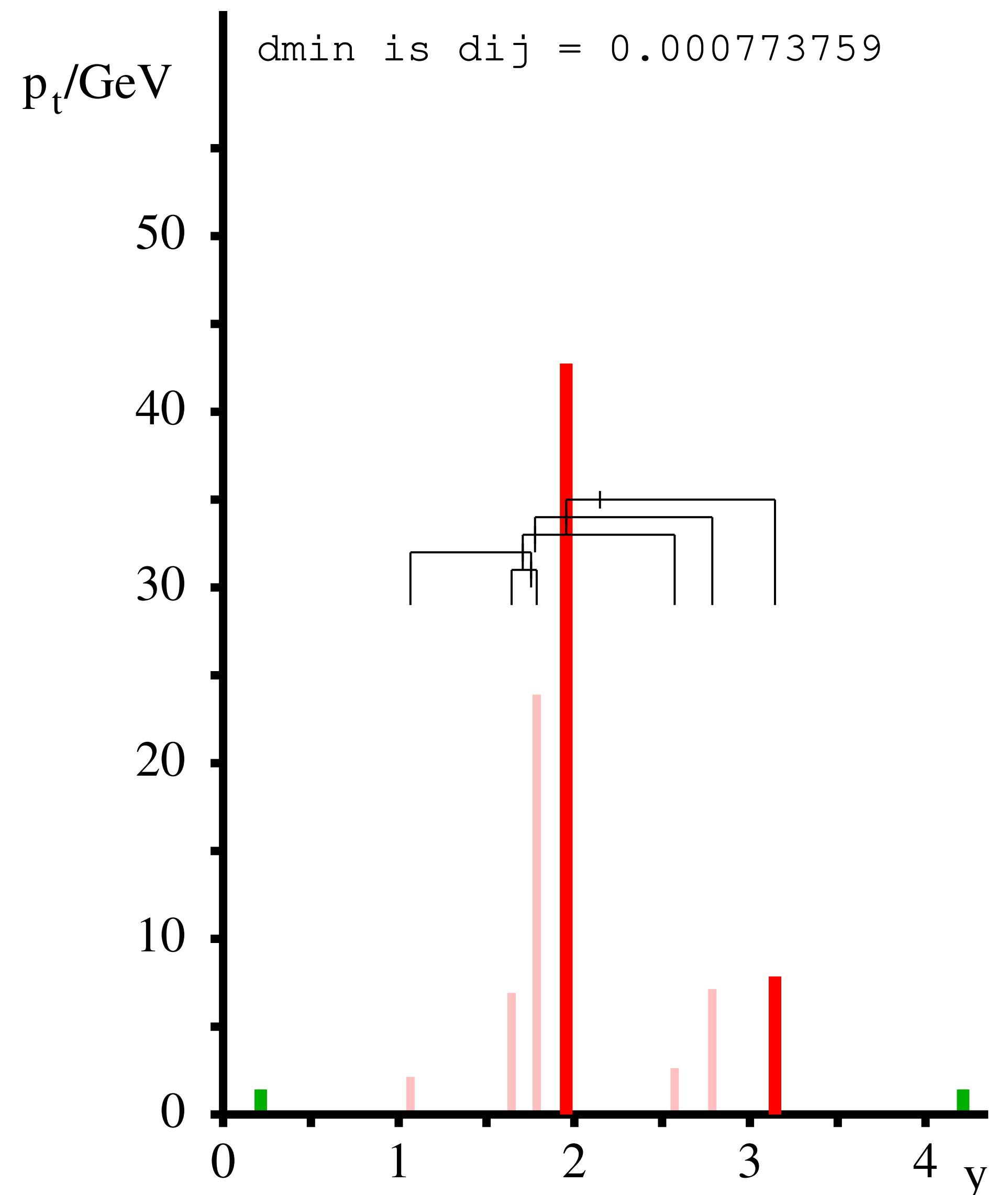
# Identifying jet substructure: try out anti- $k_t$



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

# Identifying jet substructure: try out anti- $k_t$

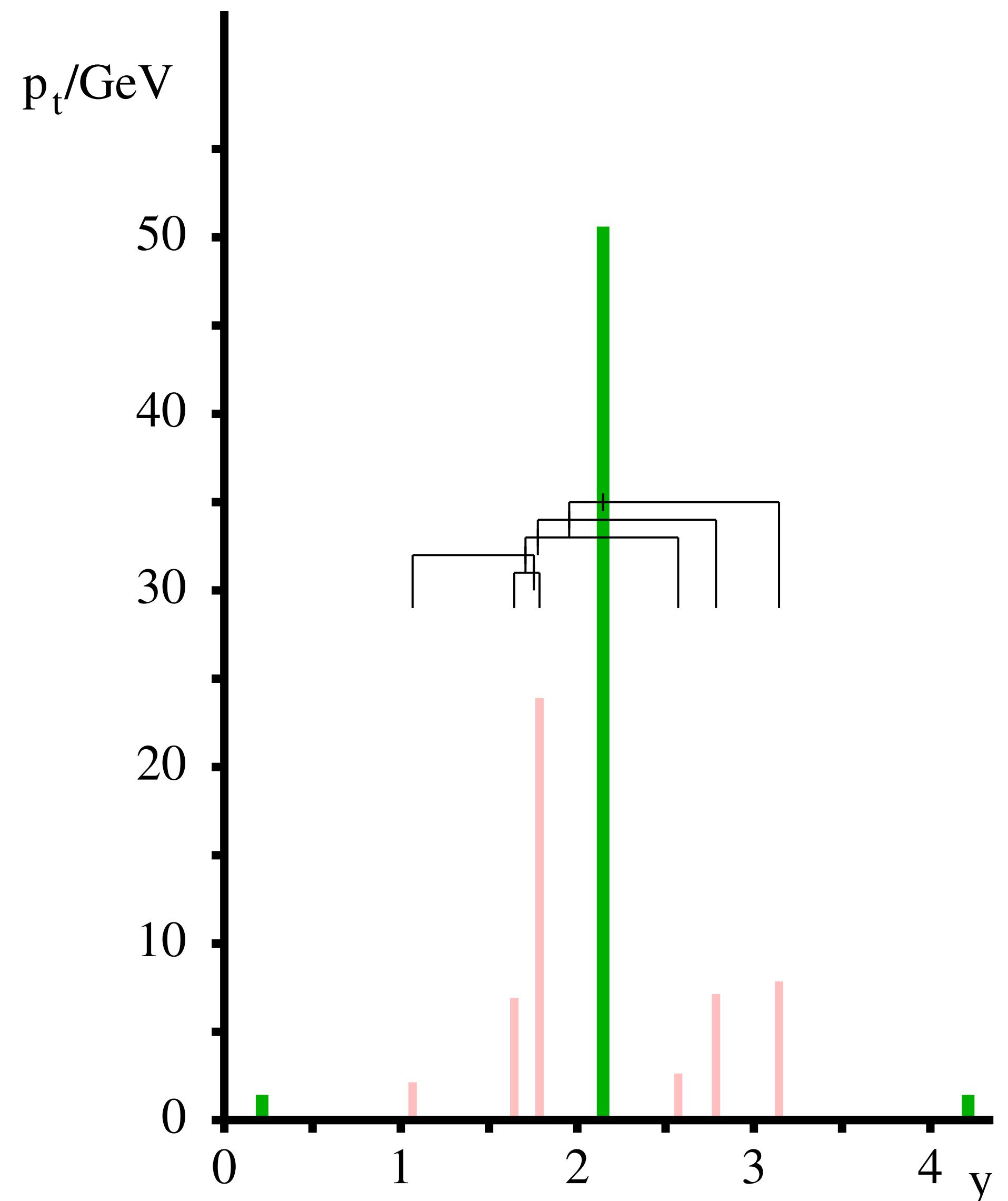


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$

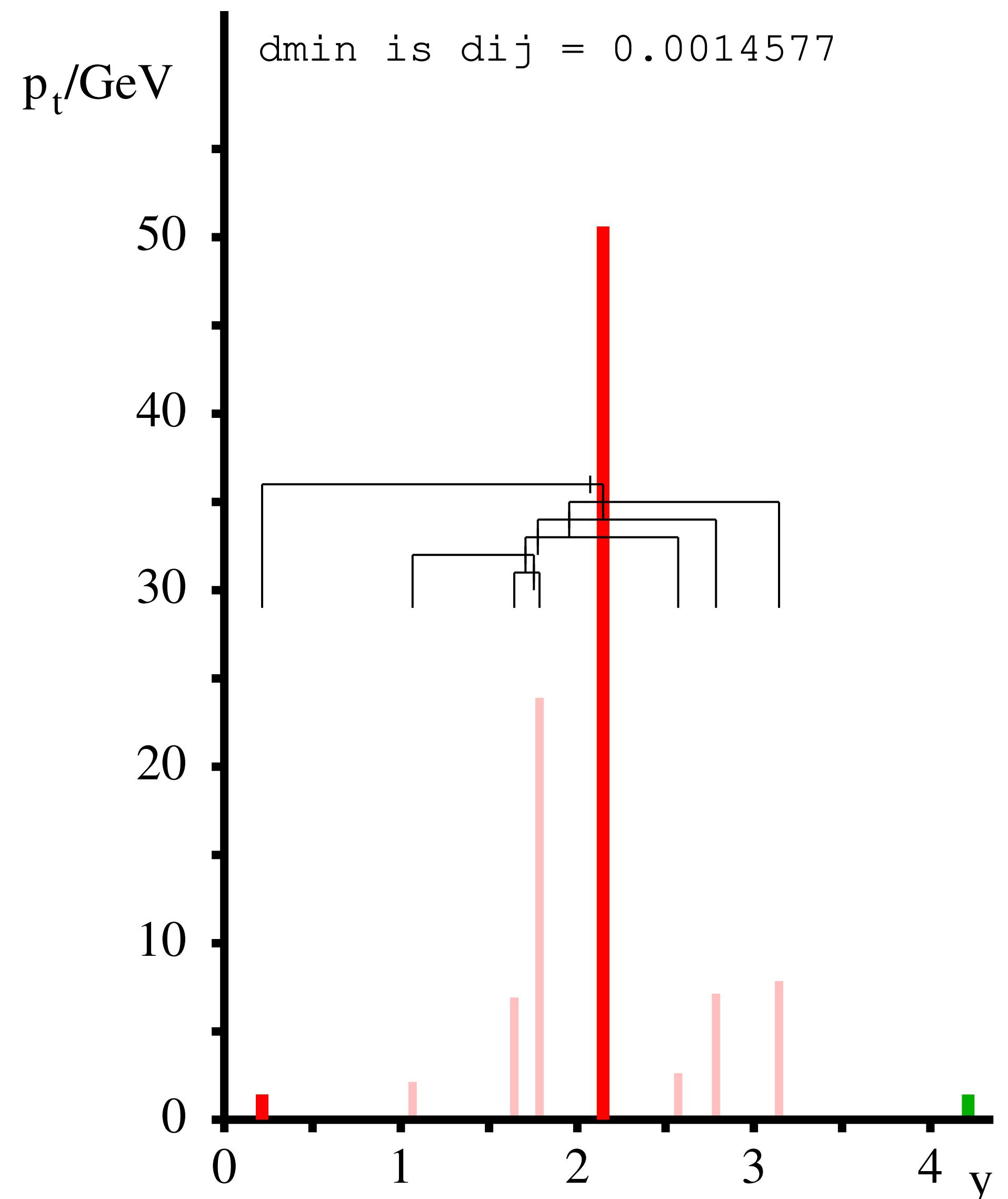


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$

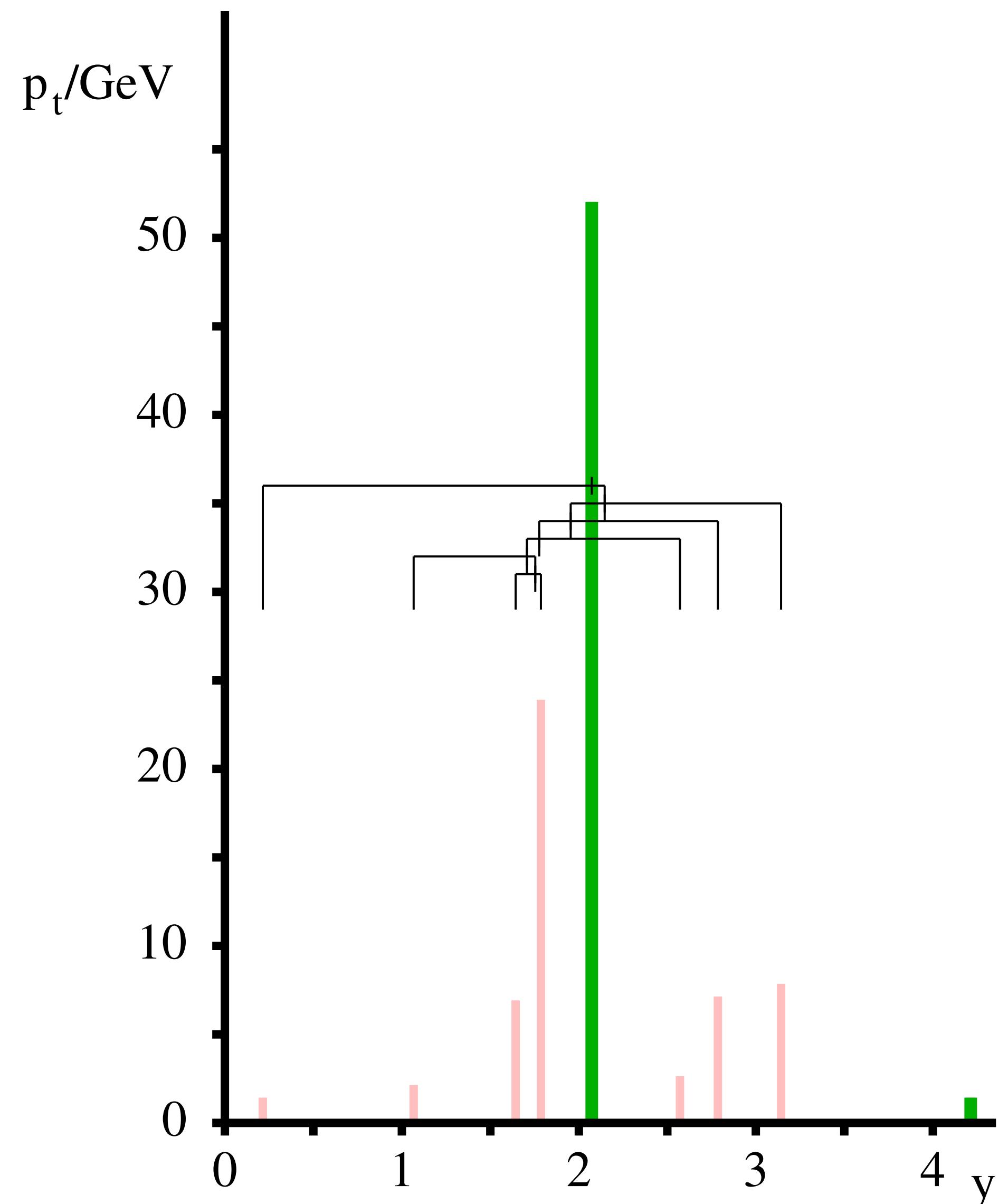


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$

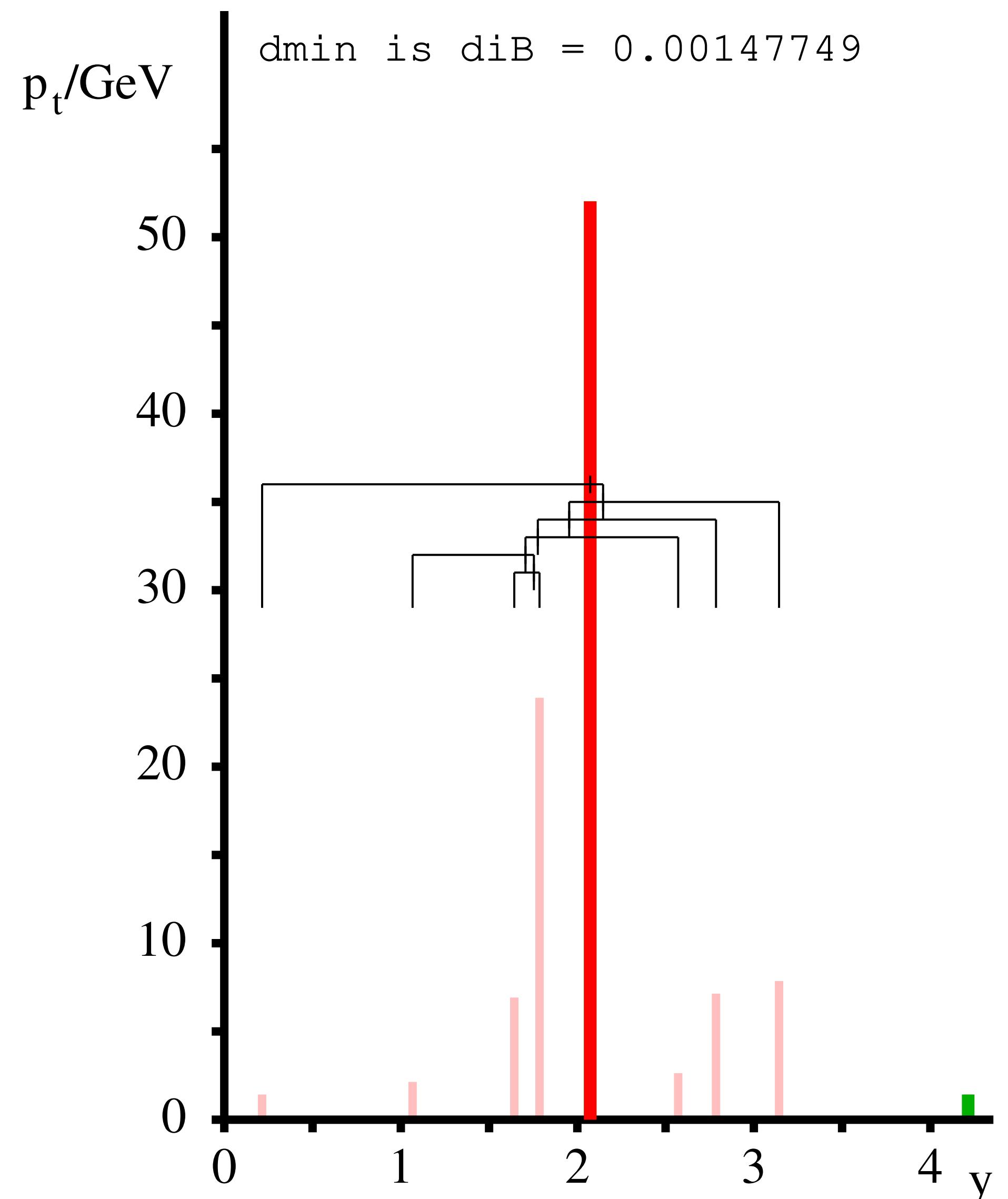


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

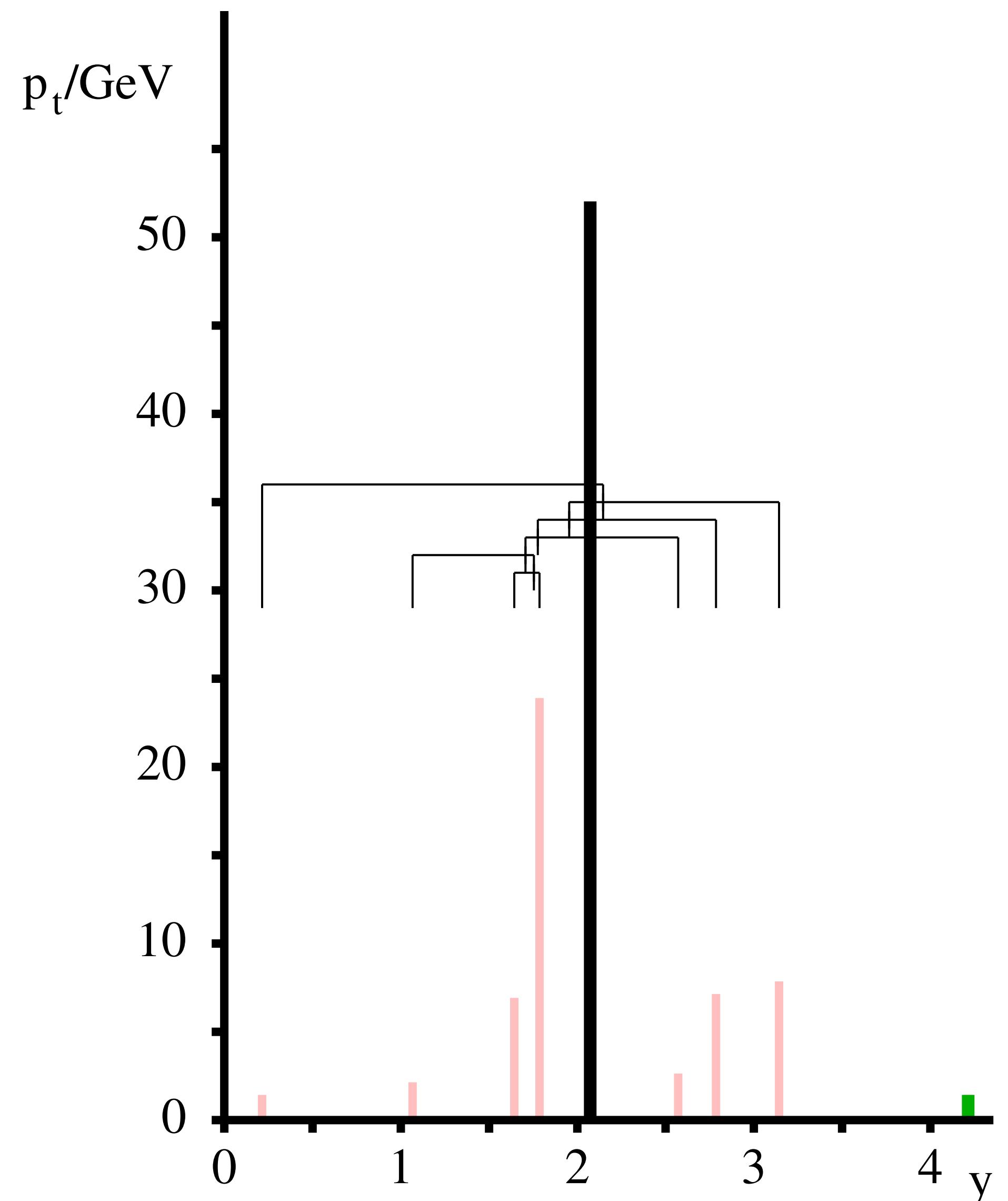
This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$



# Identifying jet substructure: try out anti- $k_t$

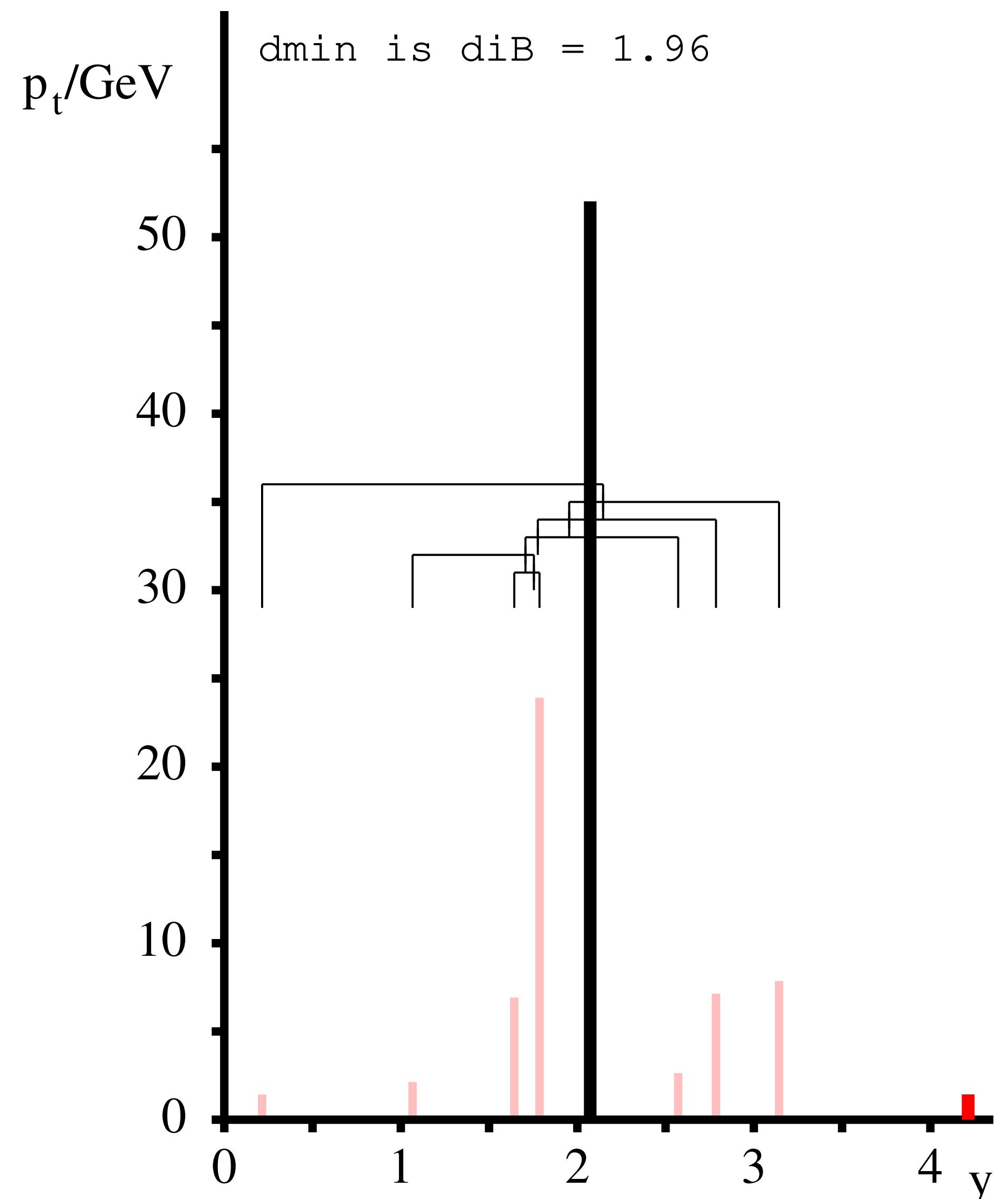


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$

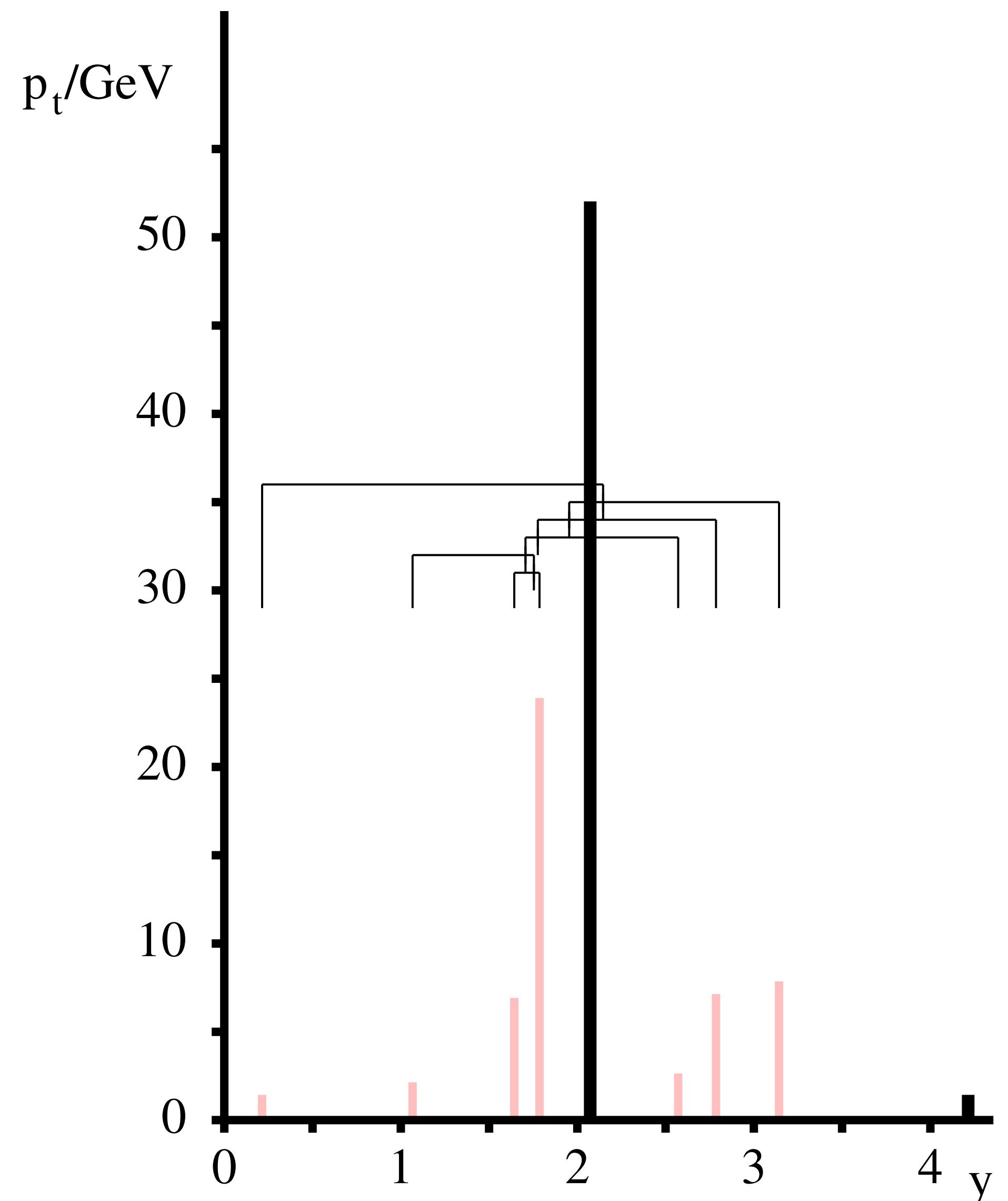


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

# Identifying jet substructure: try out anti- $k_t$

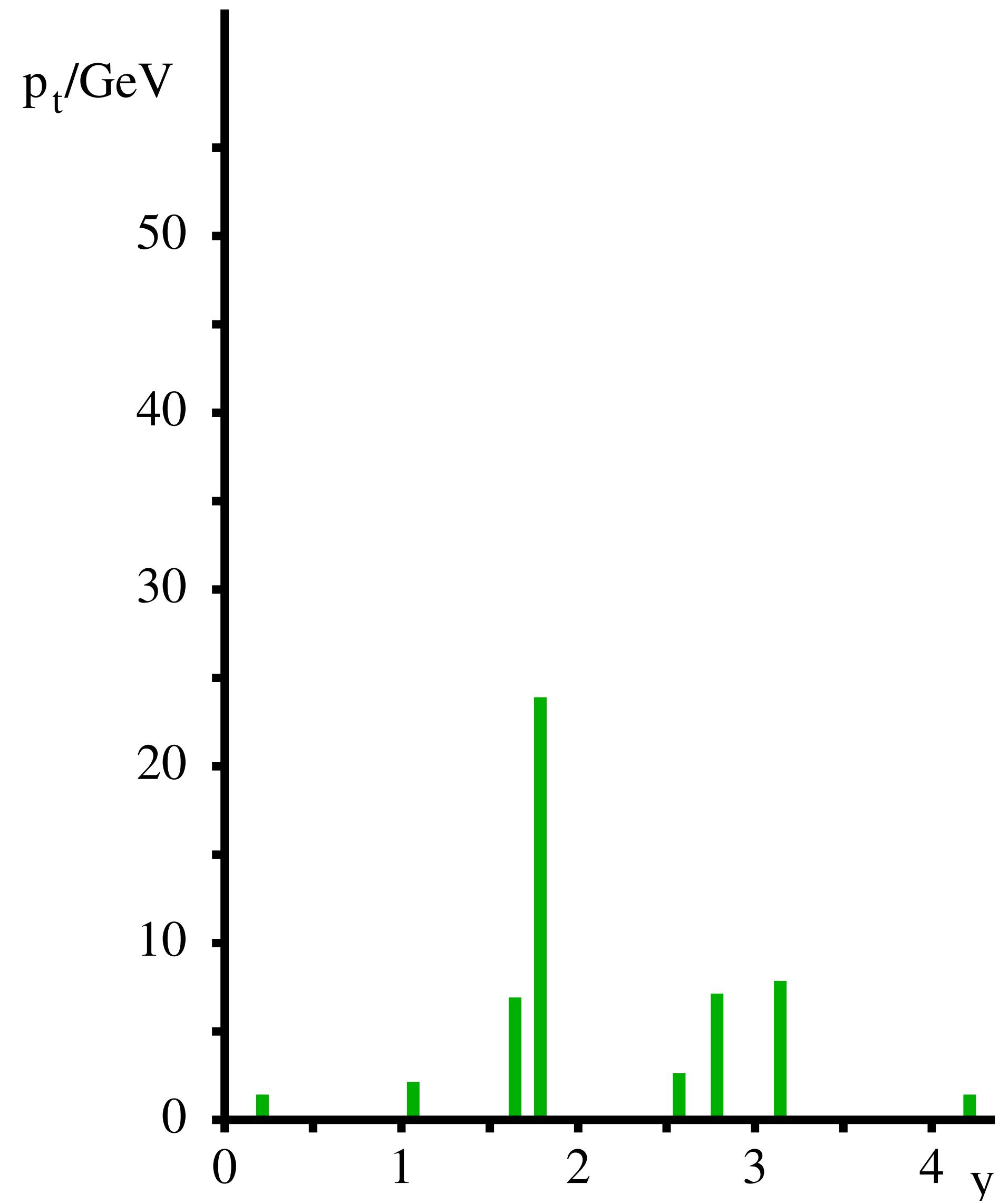


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jets

Anti- $k_t$  gradually makes its way through the secondary blob  $\rightarrow$  no clear ID of substructure associated with 2nd parton

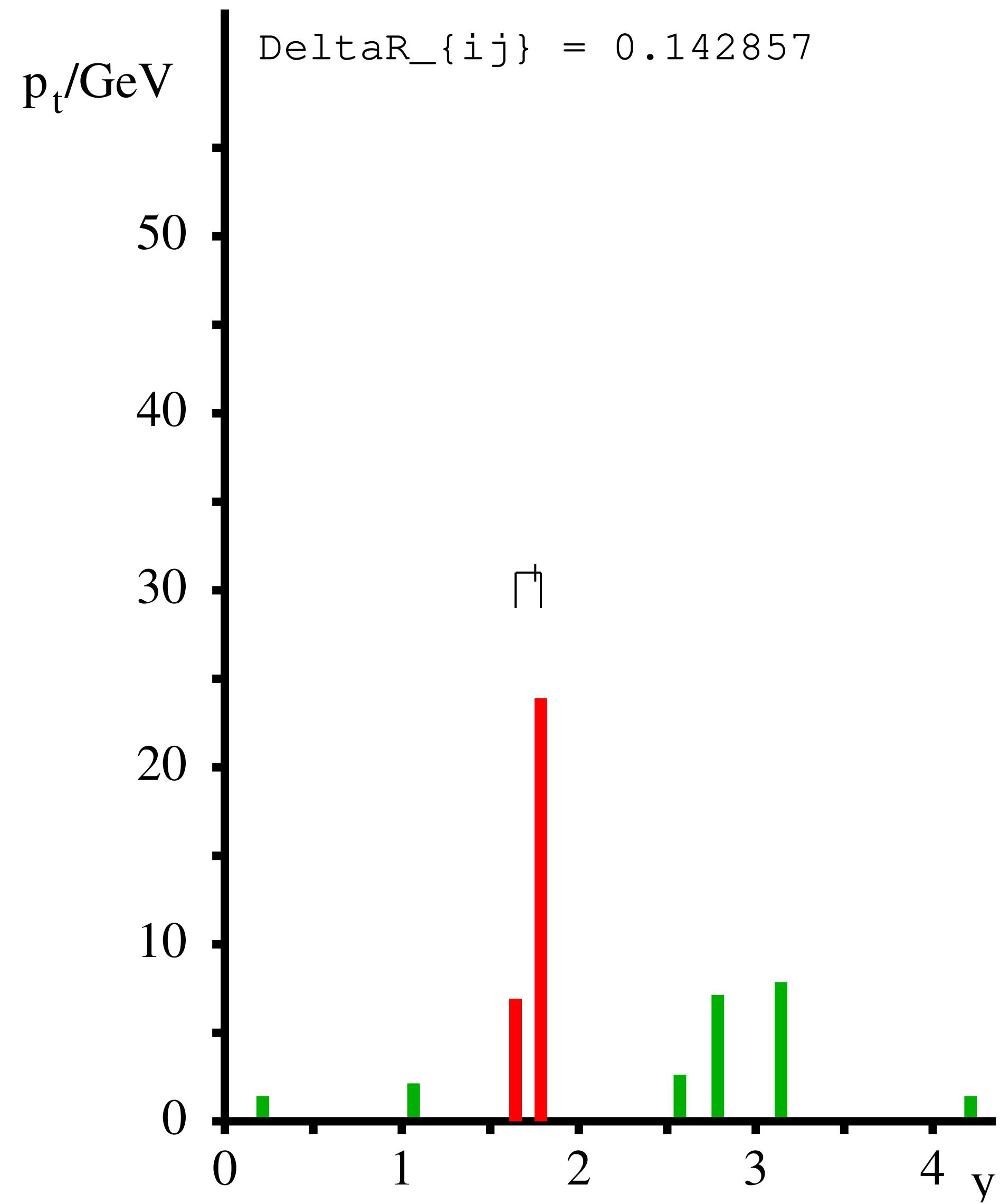
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

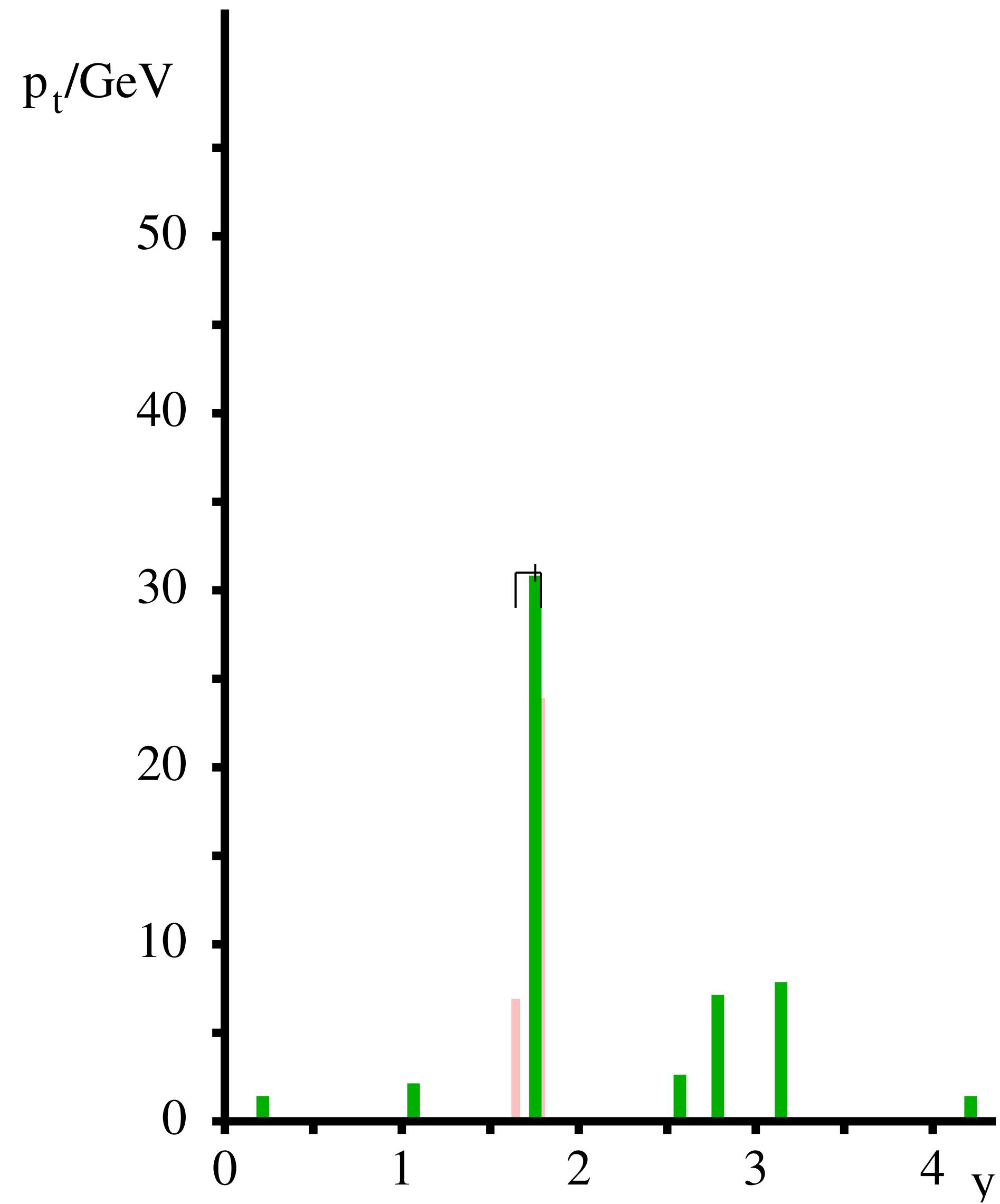
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

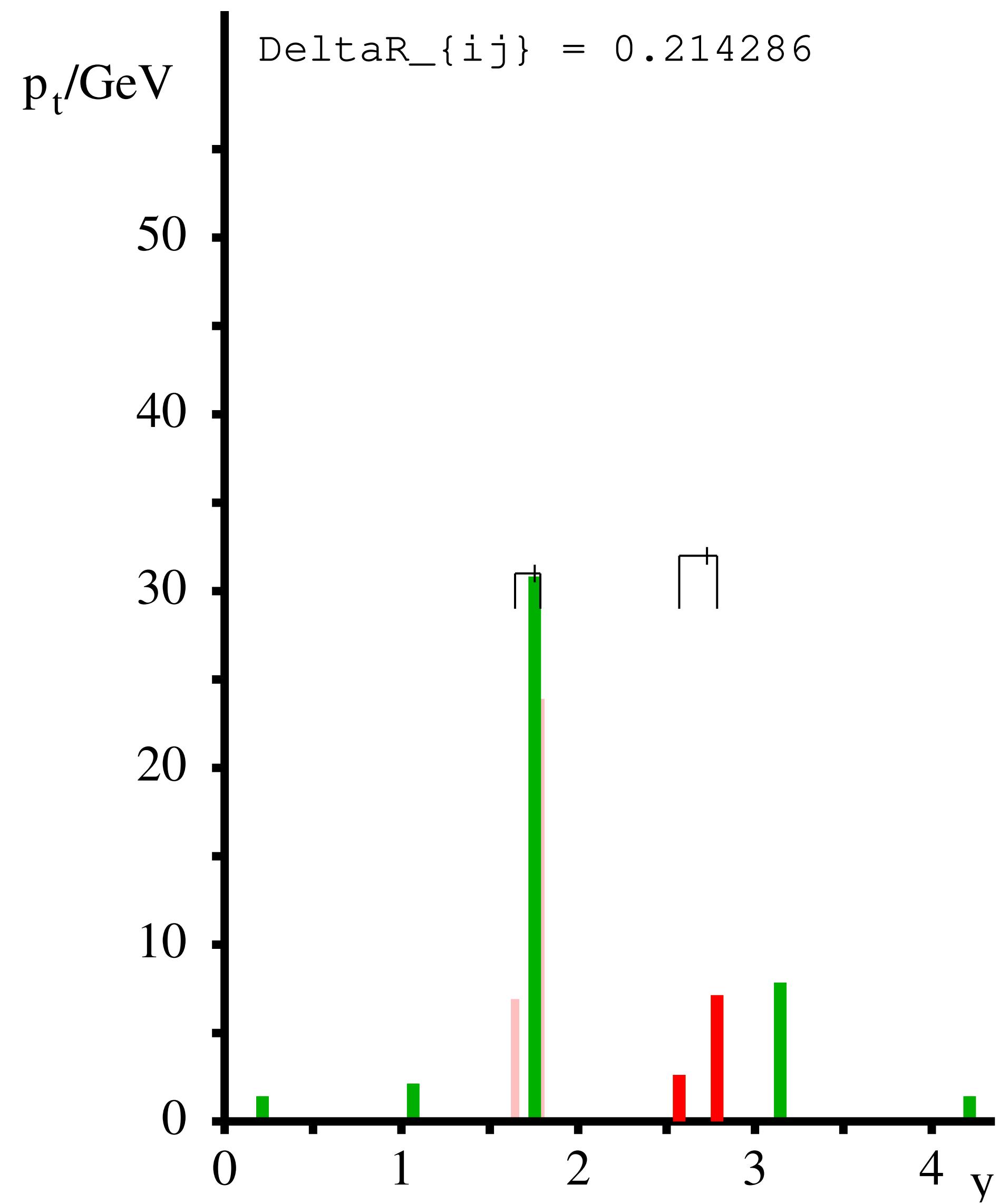
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

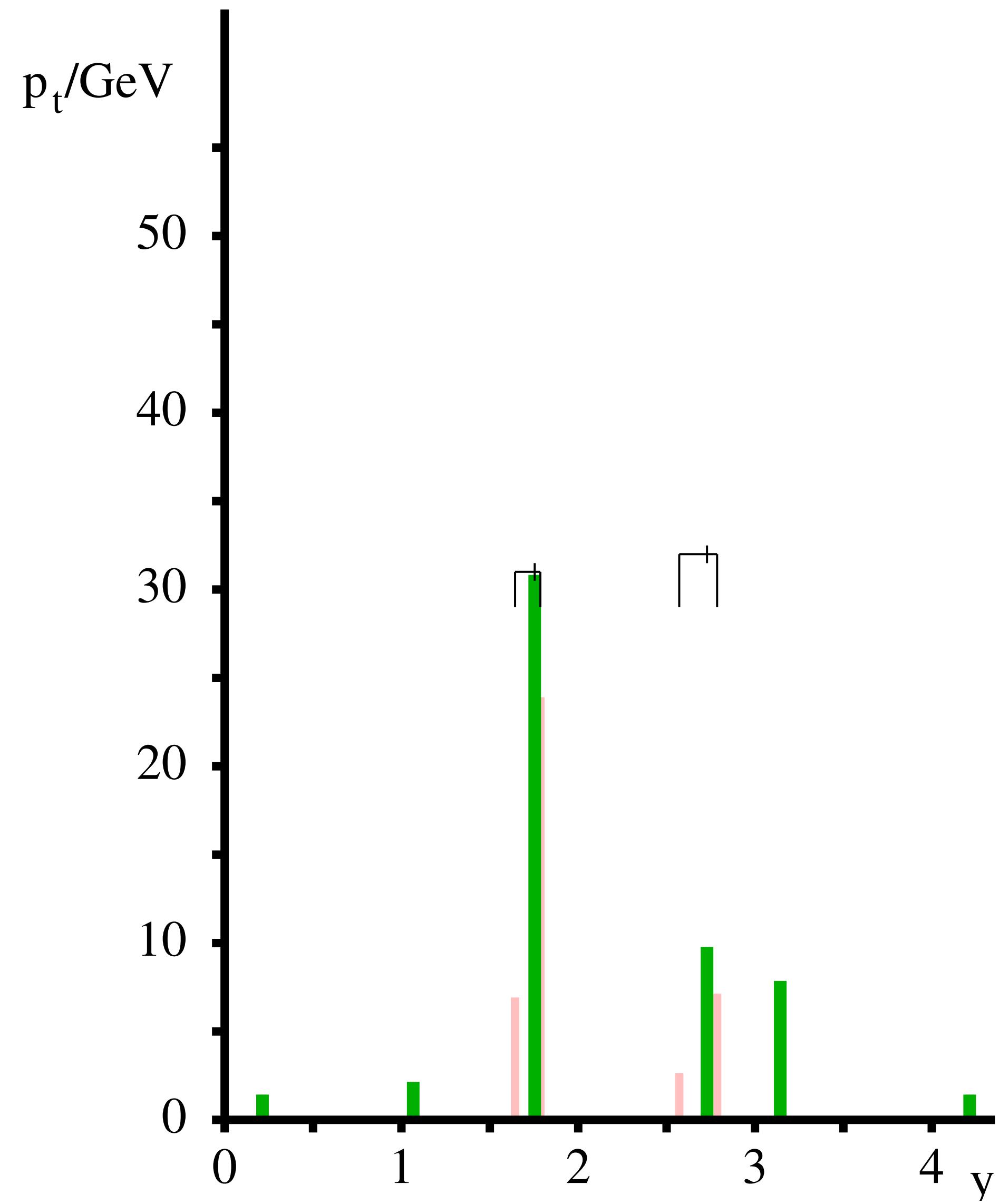
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

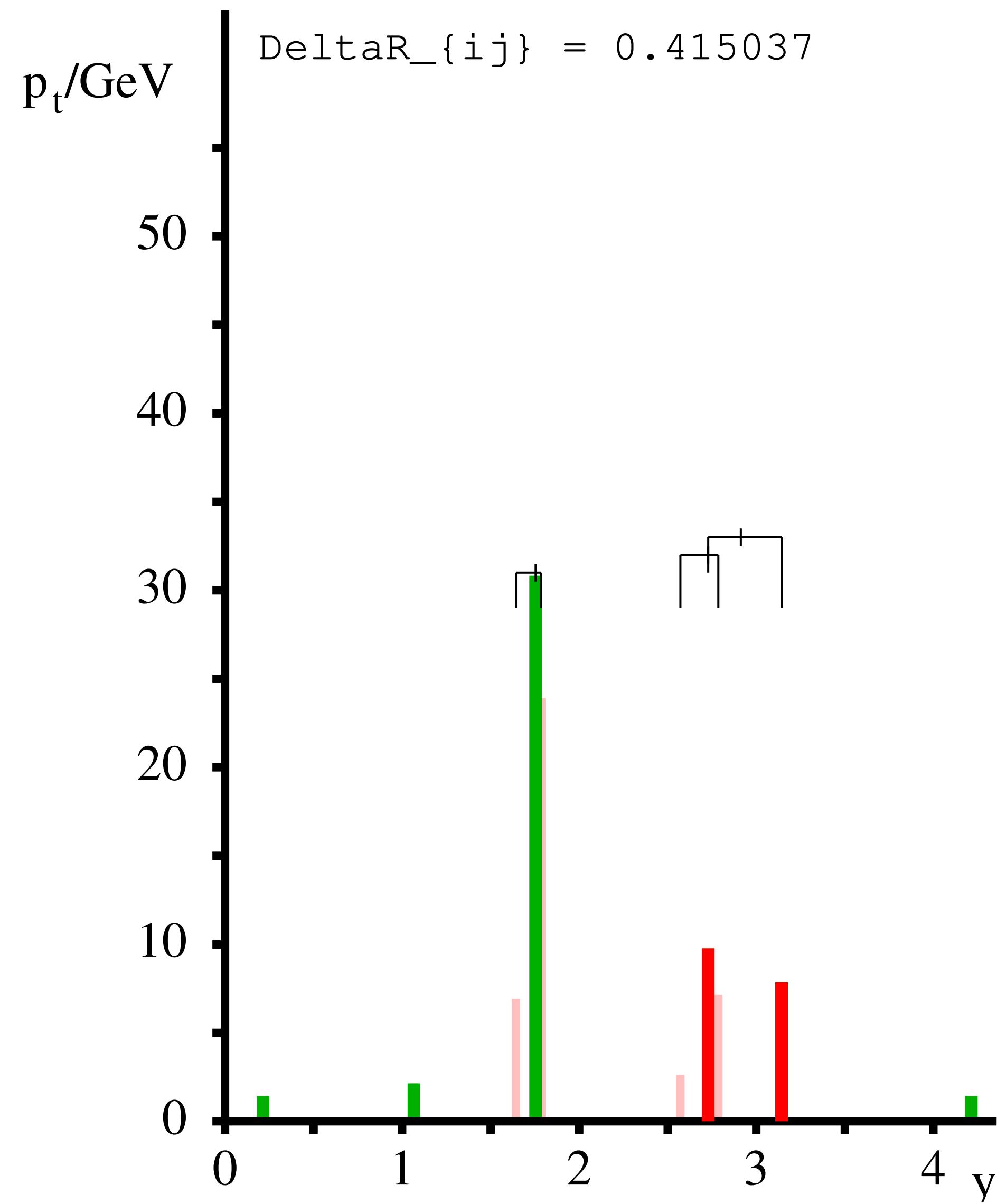
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

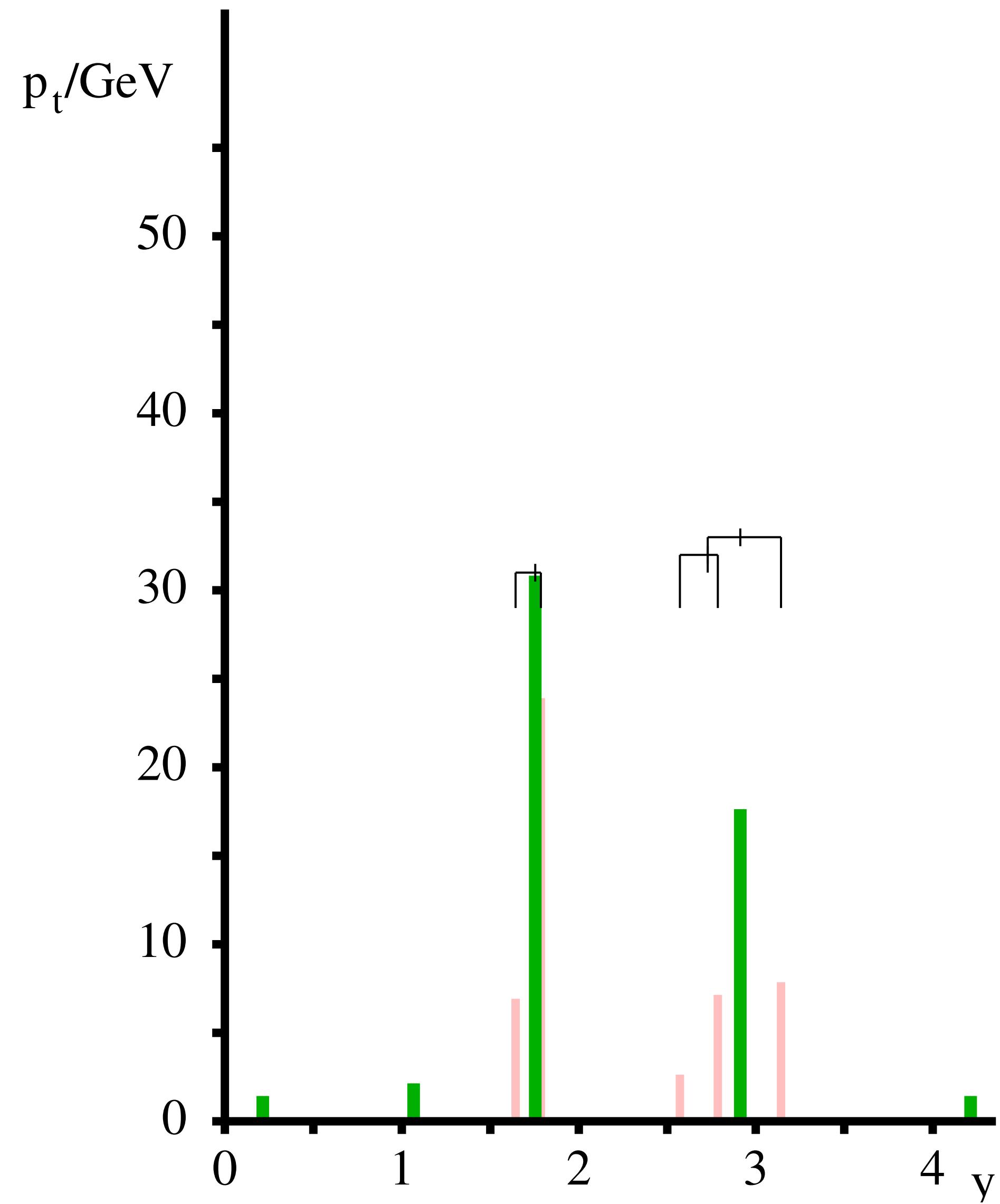
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

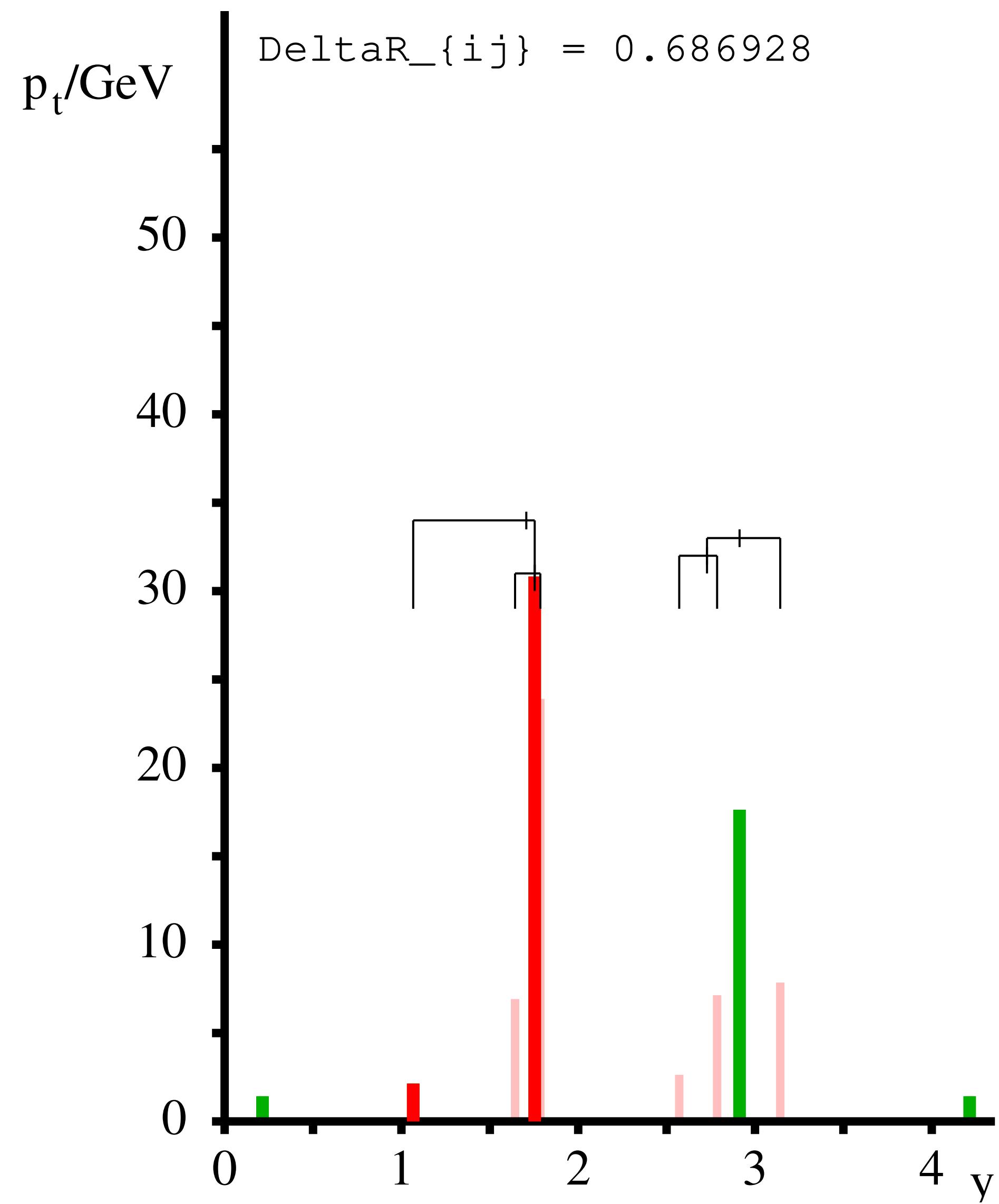
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

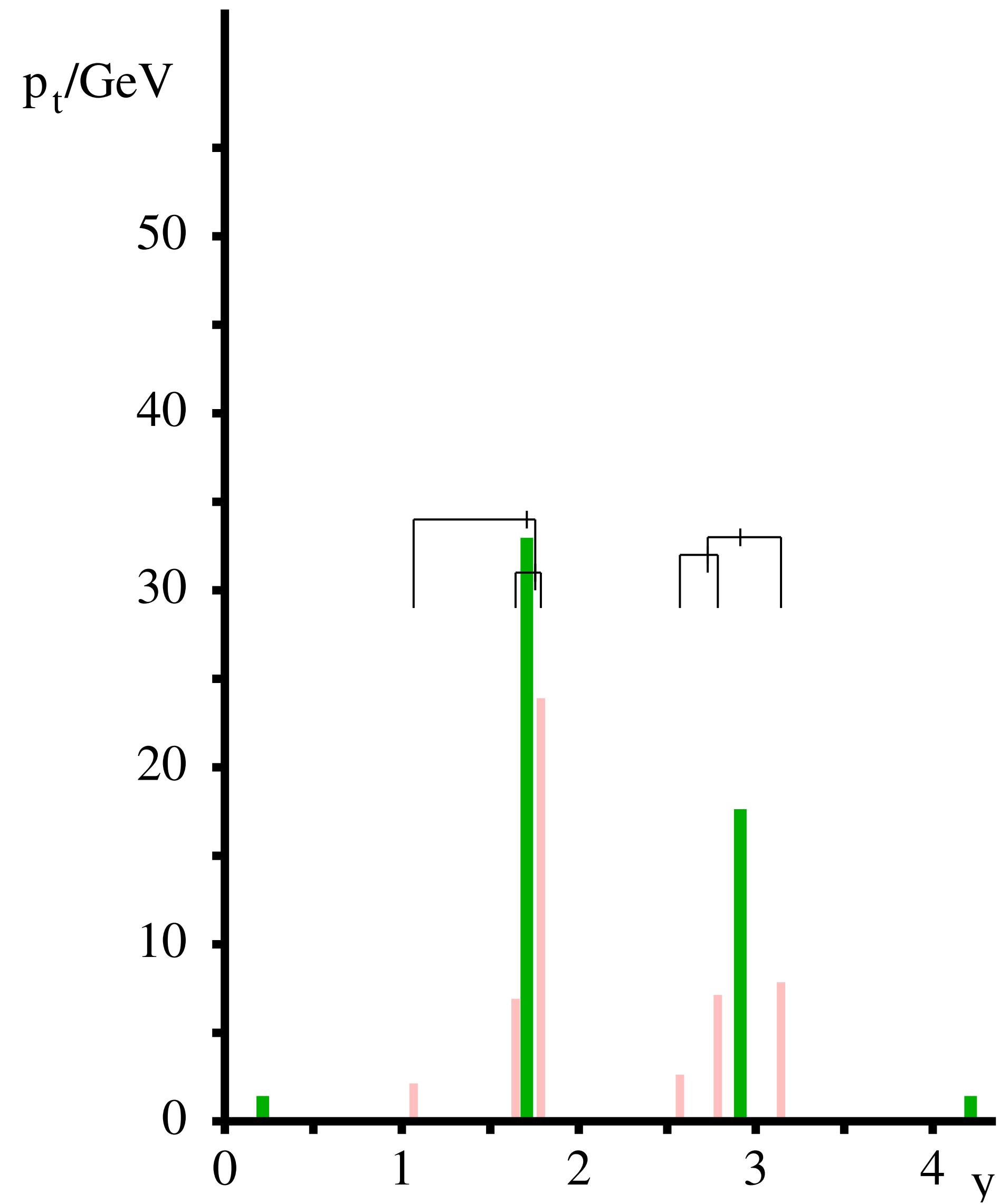
# Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

# Identifying jet substructure: Cam/Aachen

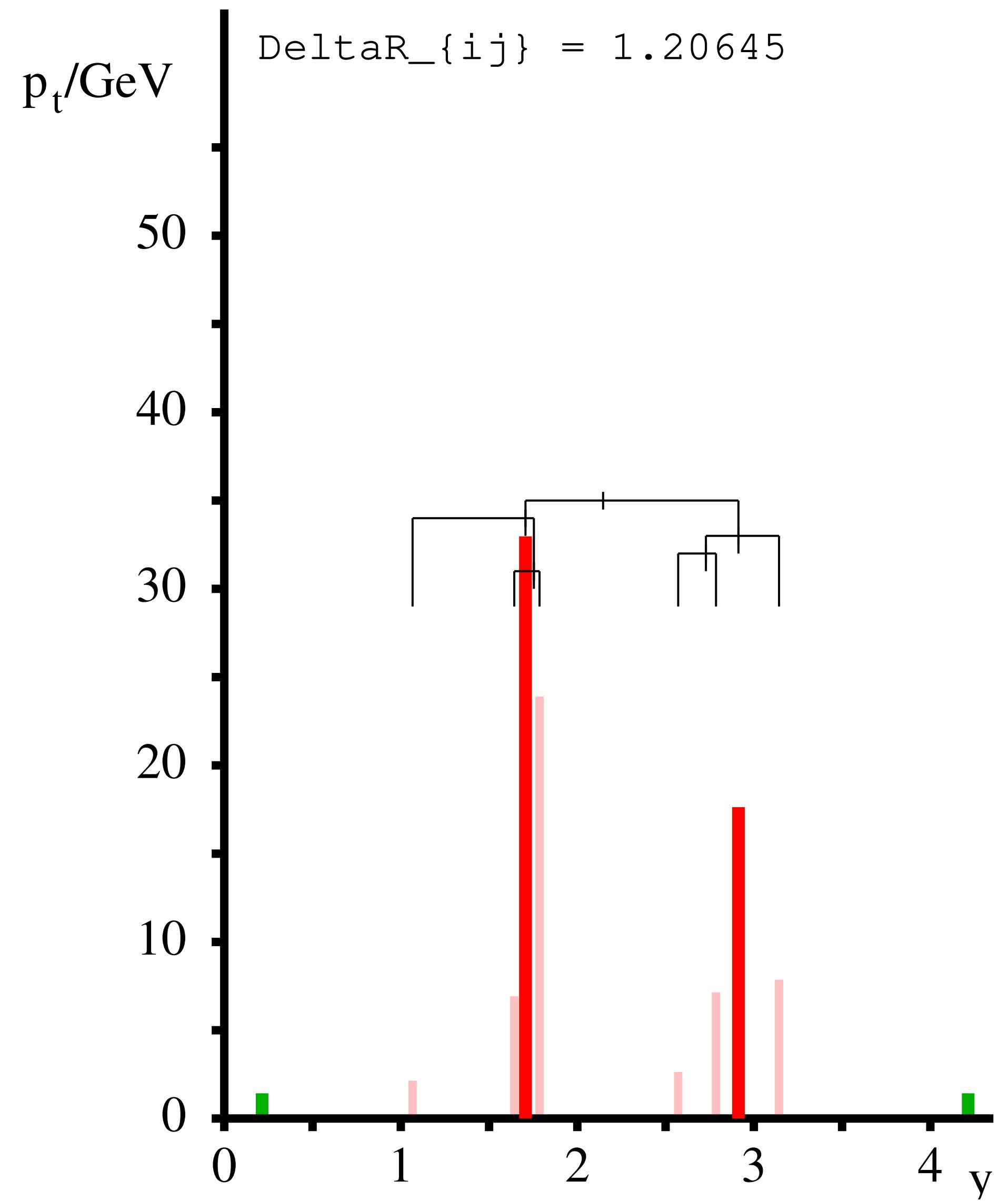


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

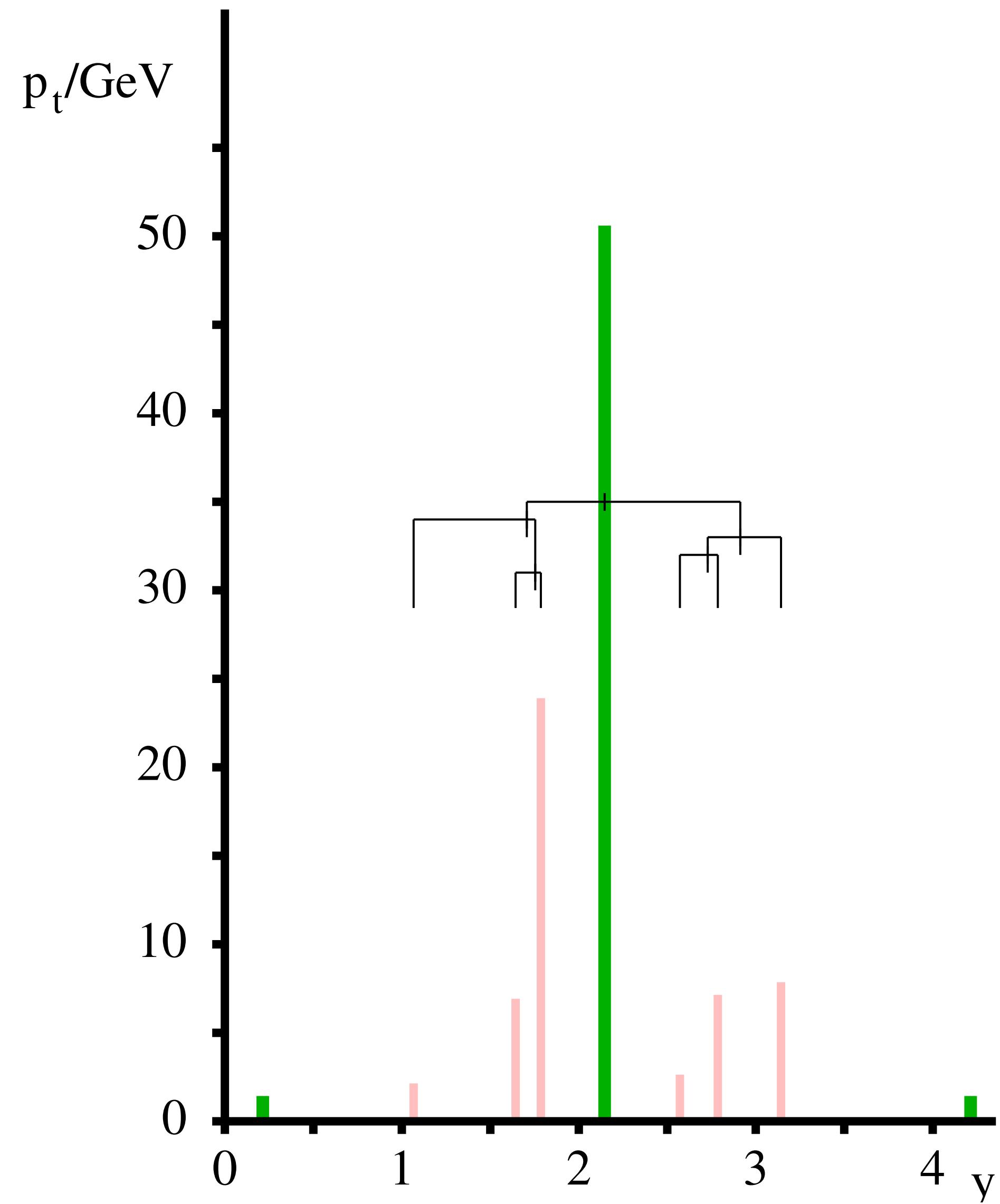


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

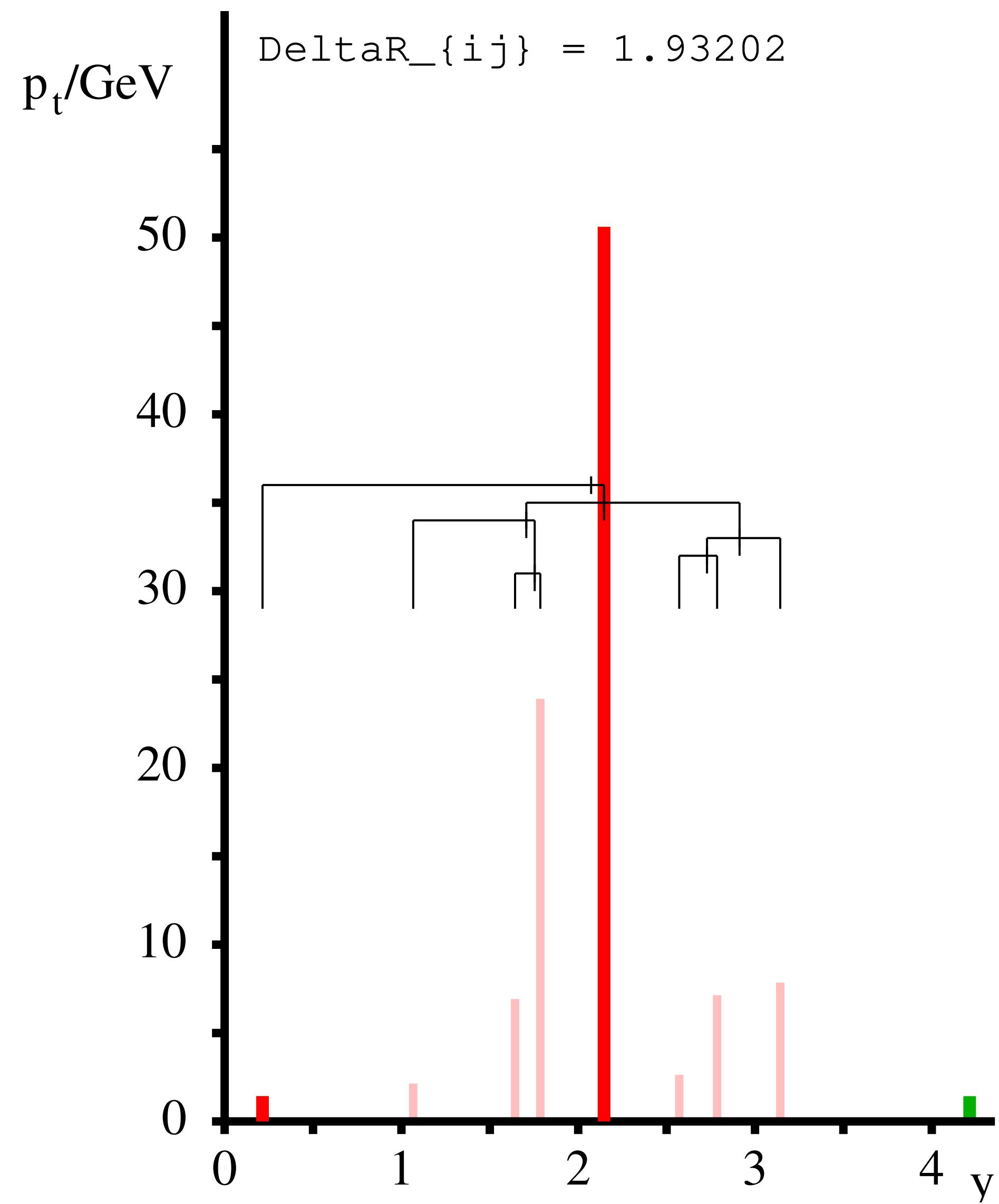


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

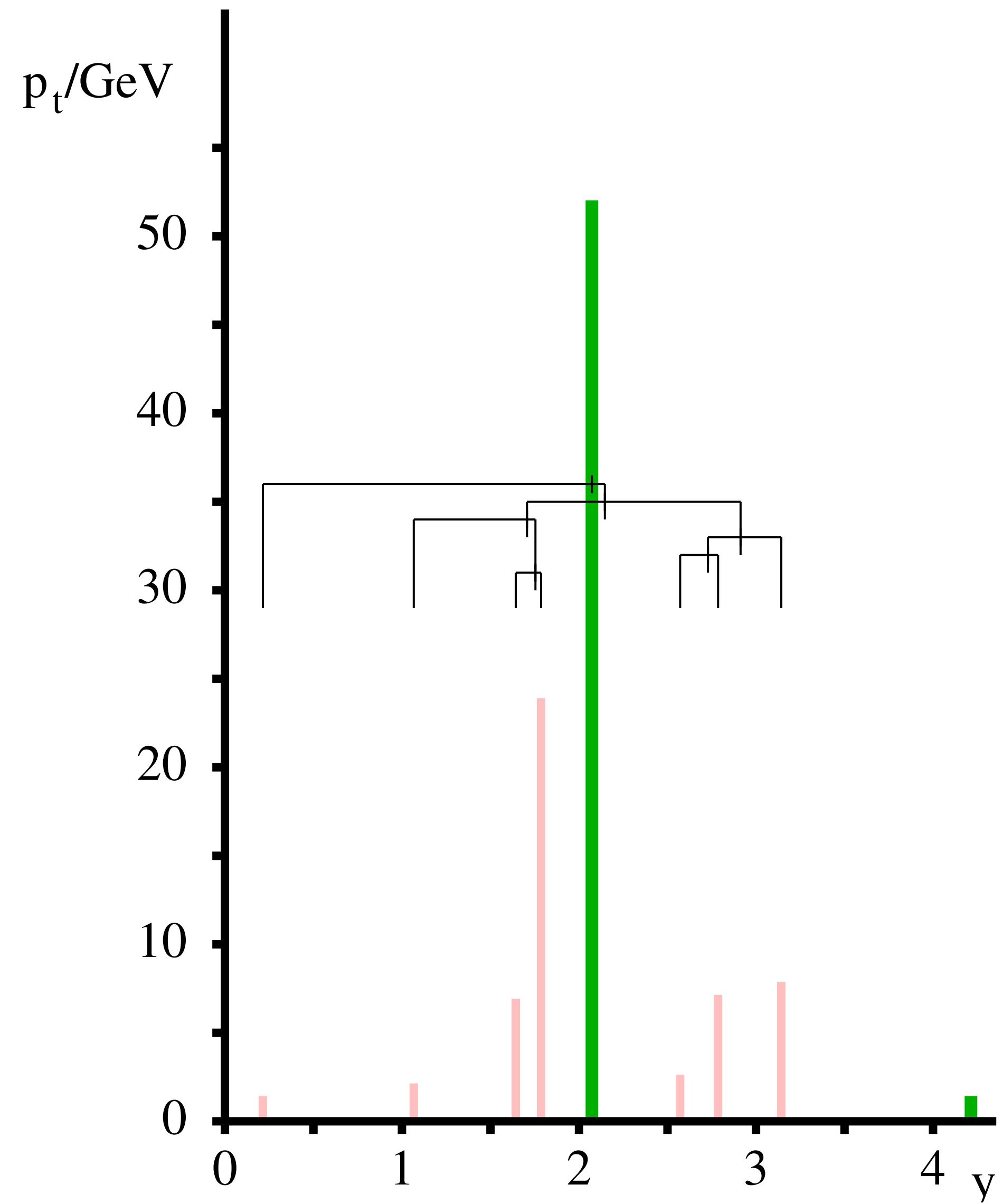


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

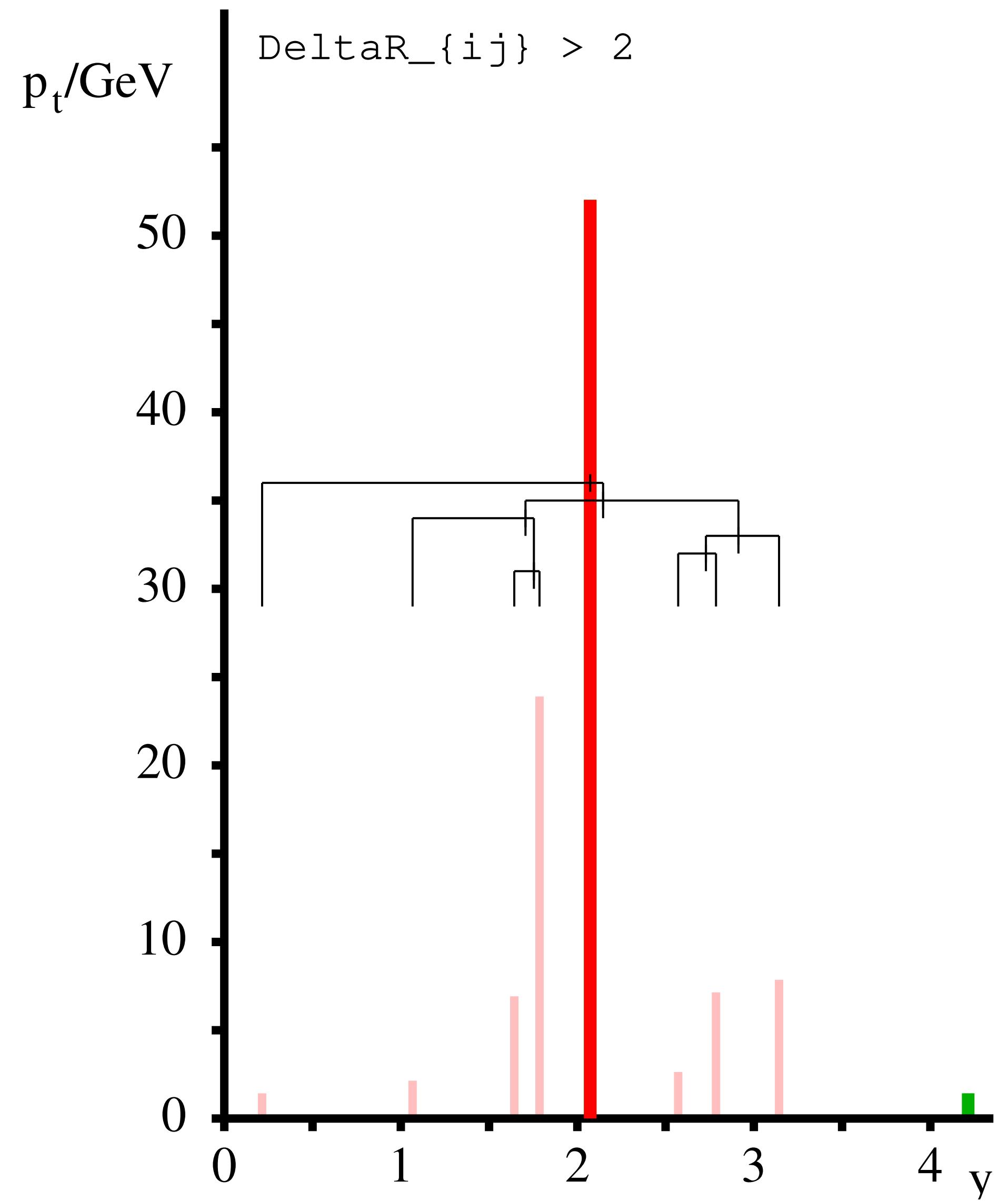


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

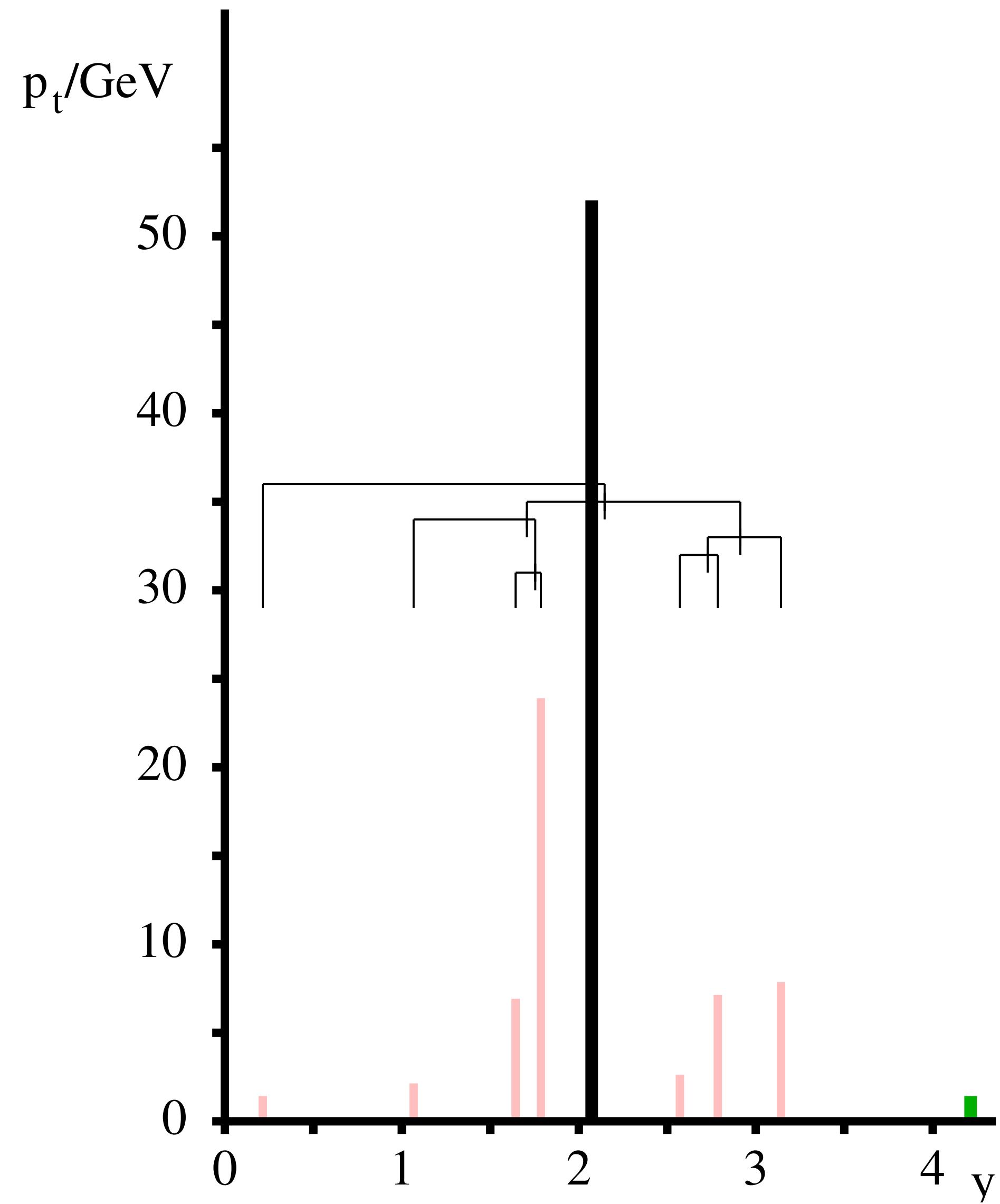


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

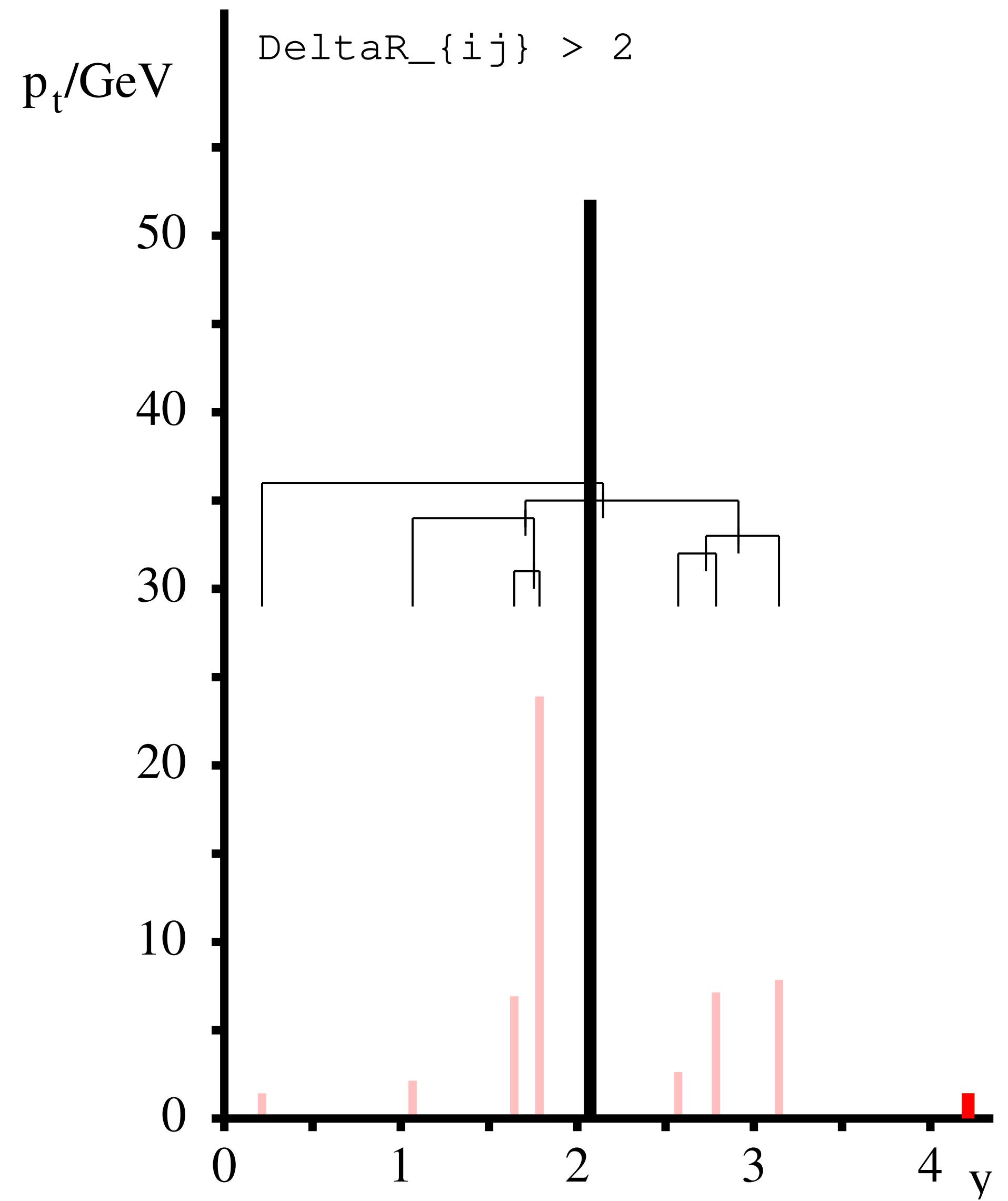


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

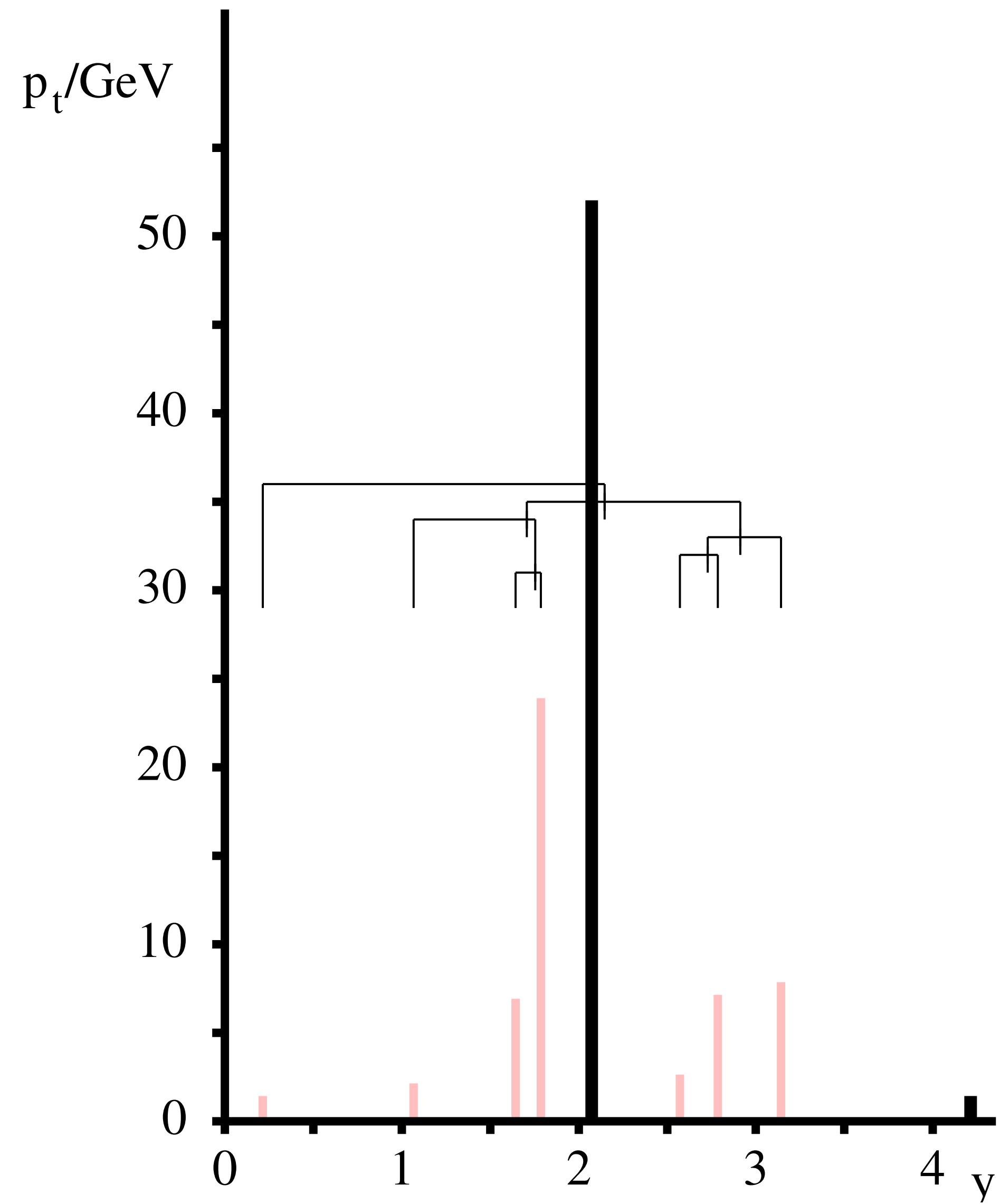


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen

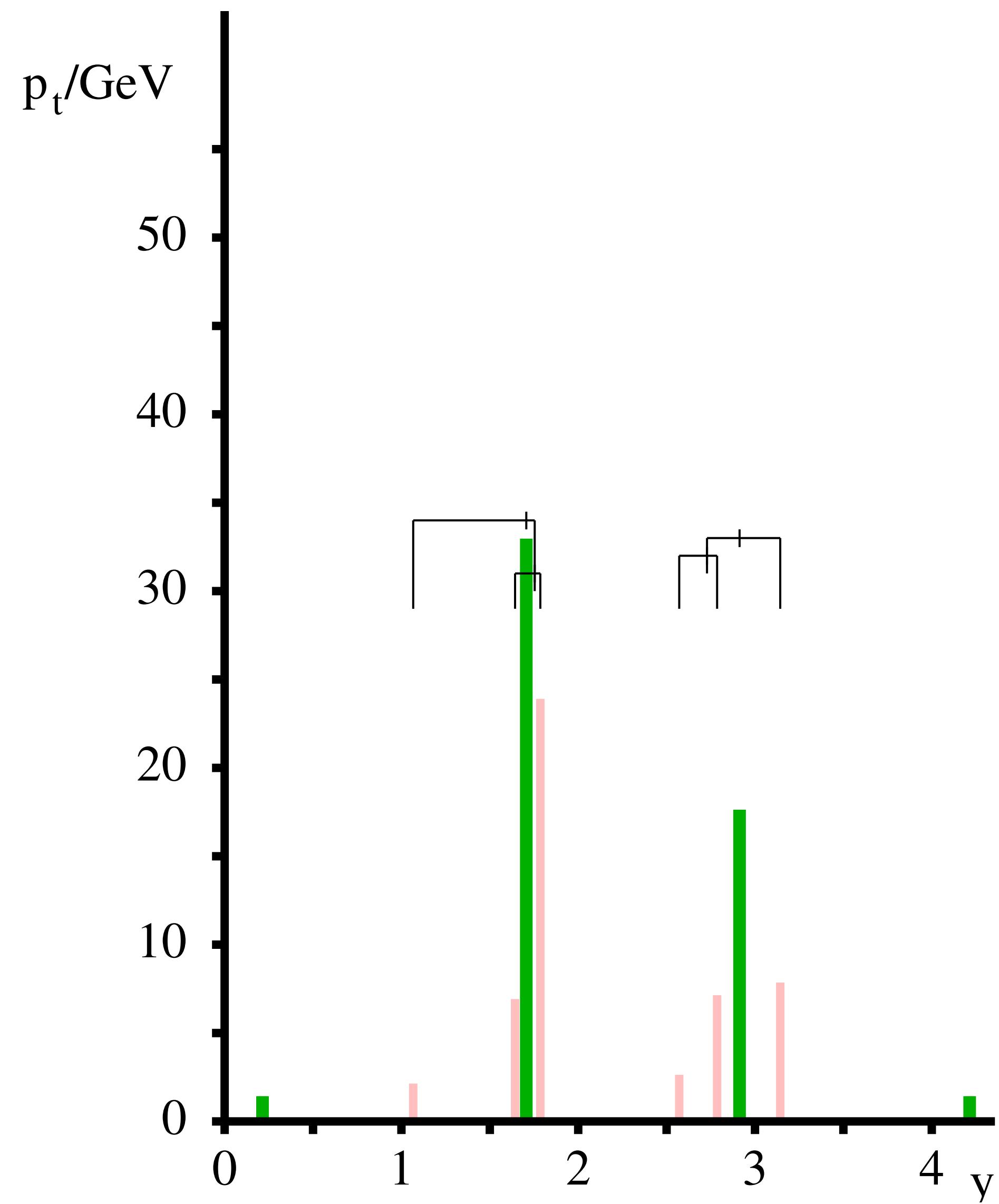


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

# Identifying jet substructure: Cam/Aachen



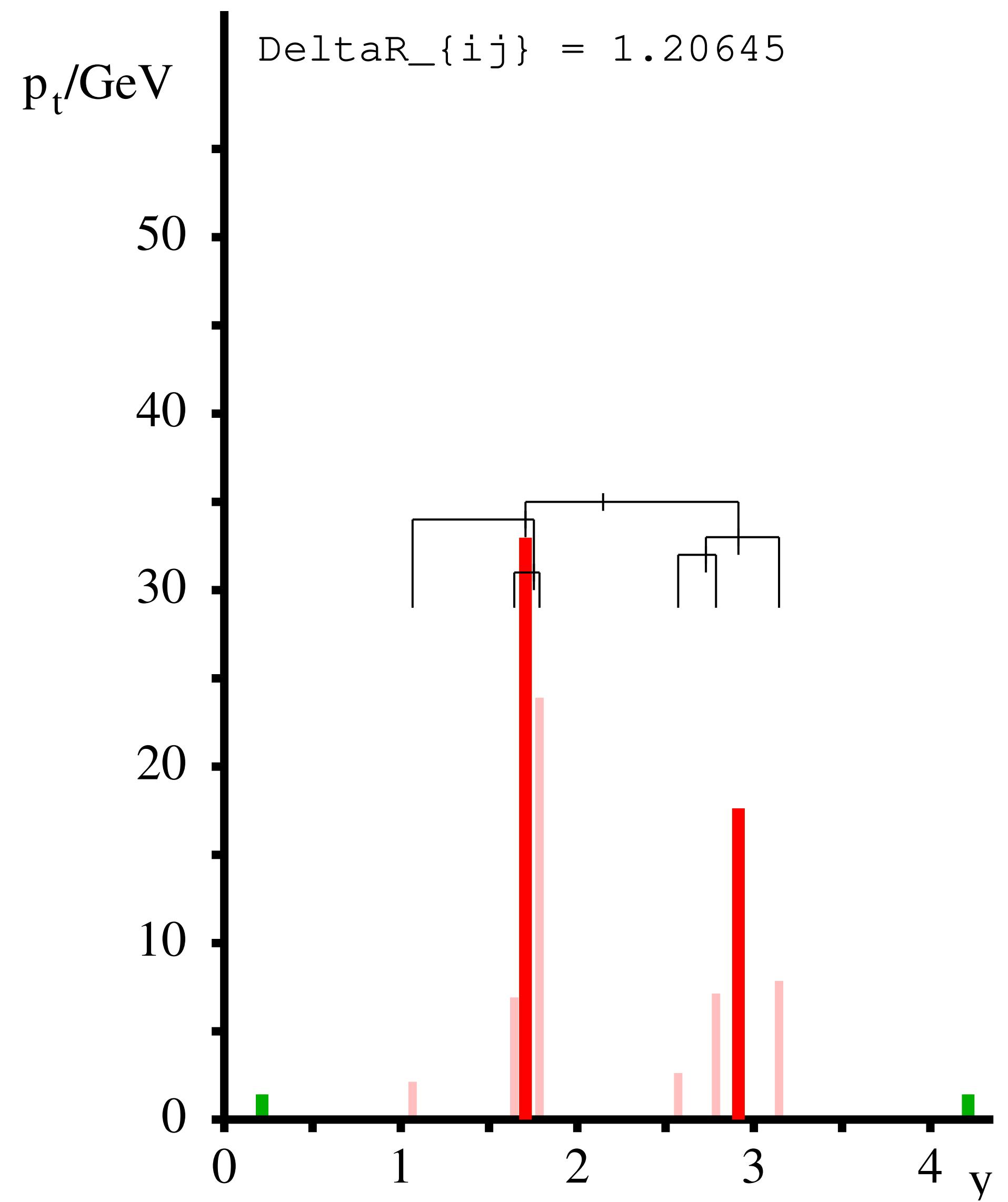
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



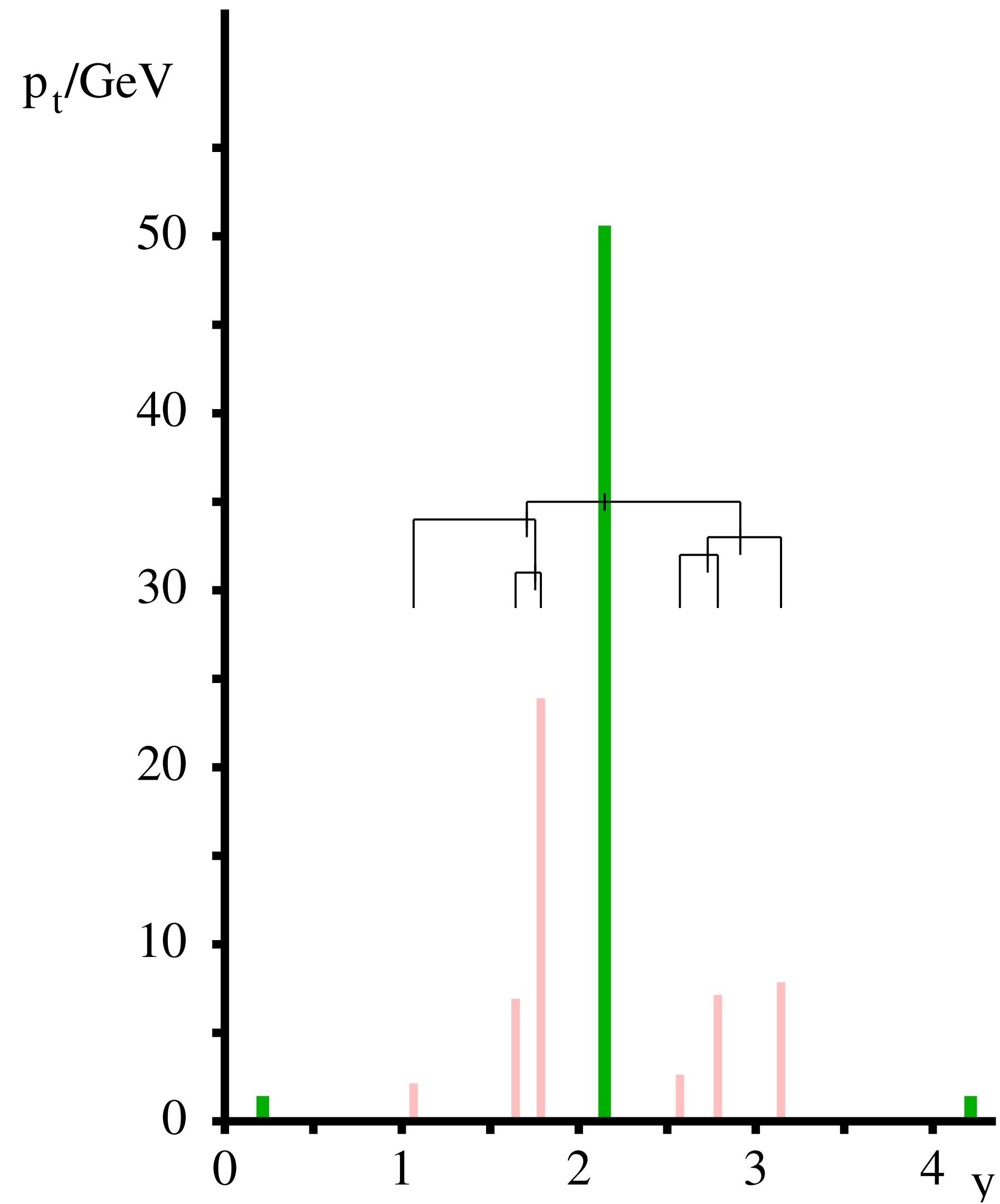
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



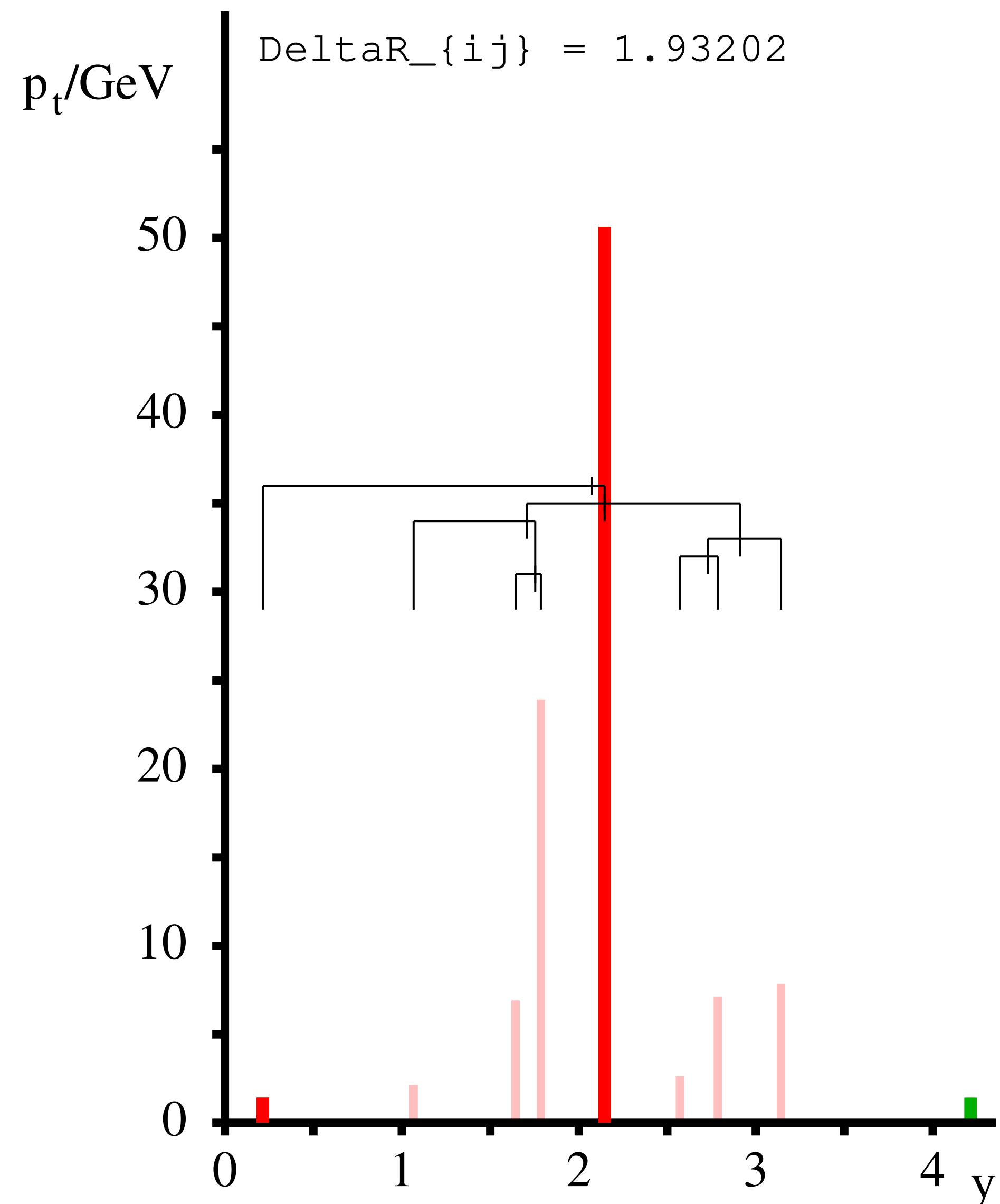
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



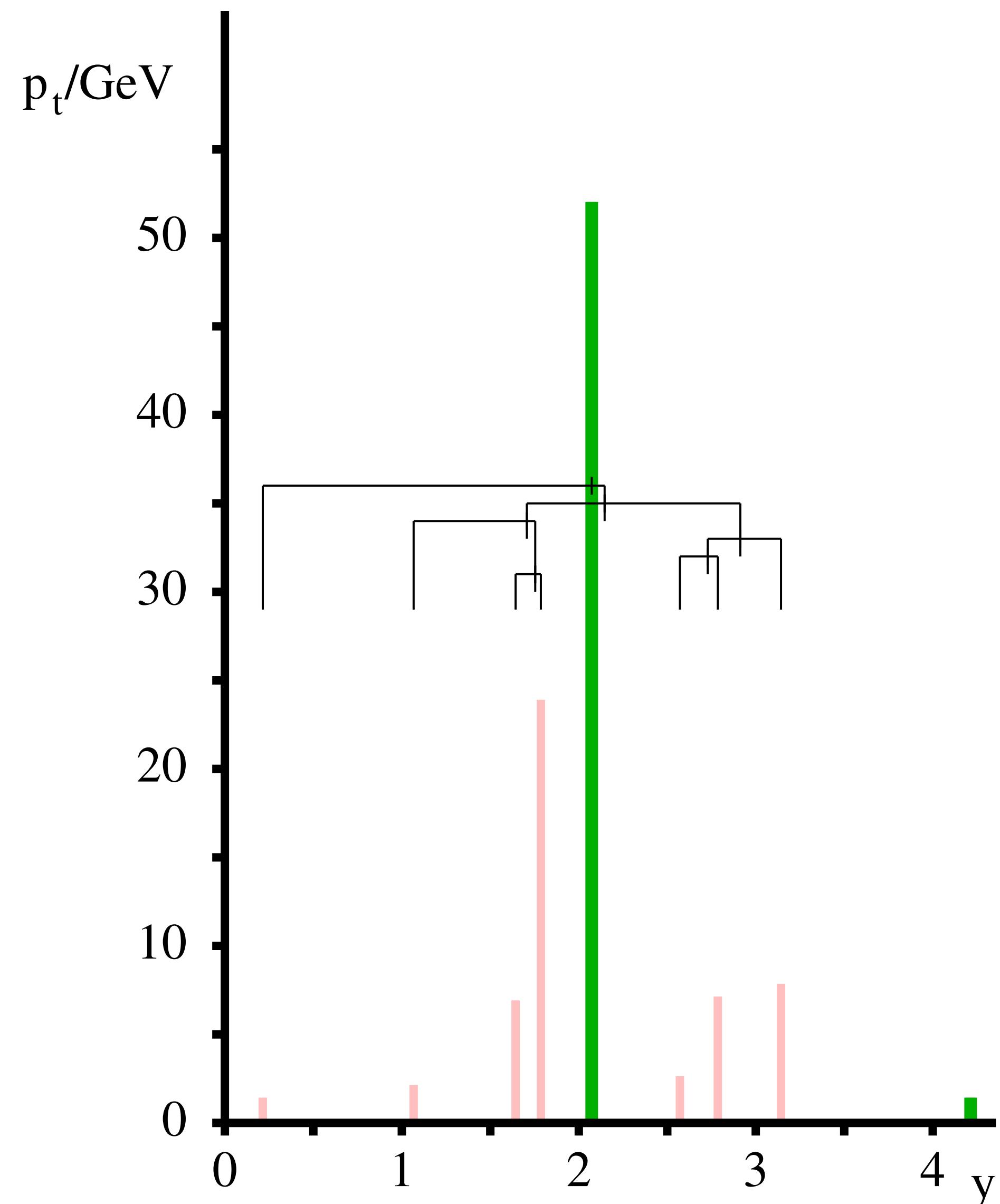
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



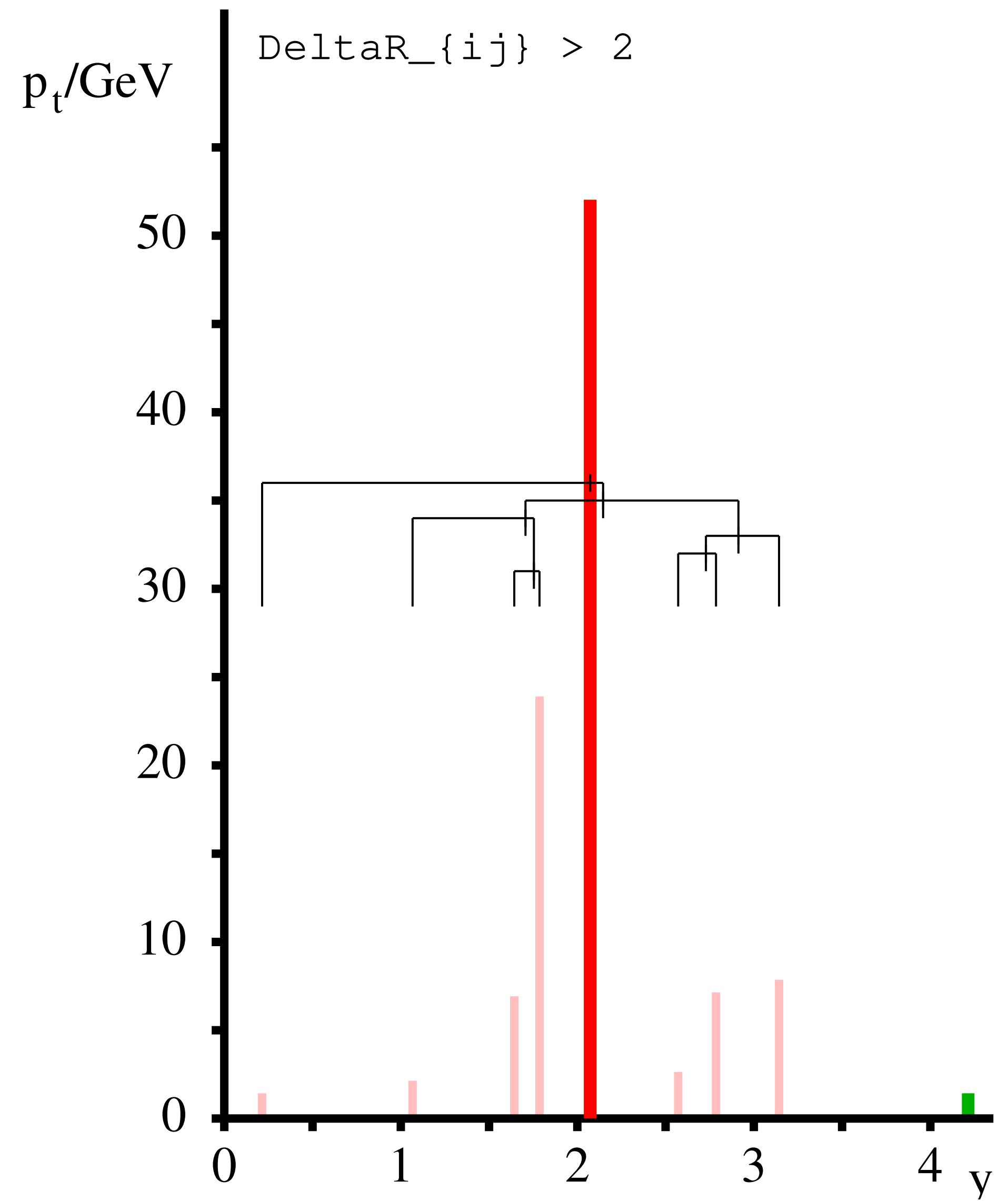
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard "blobs"

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



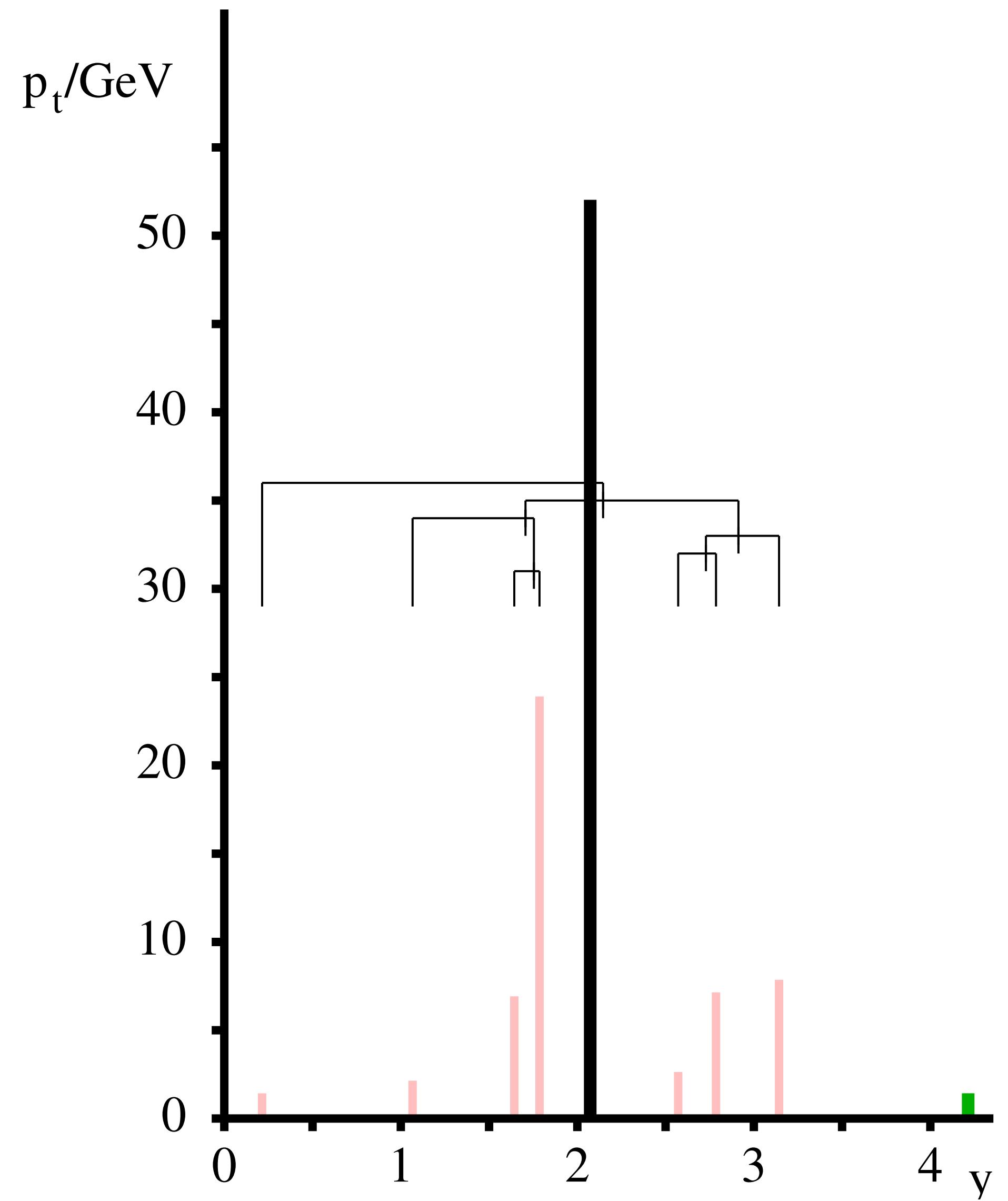
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



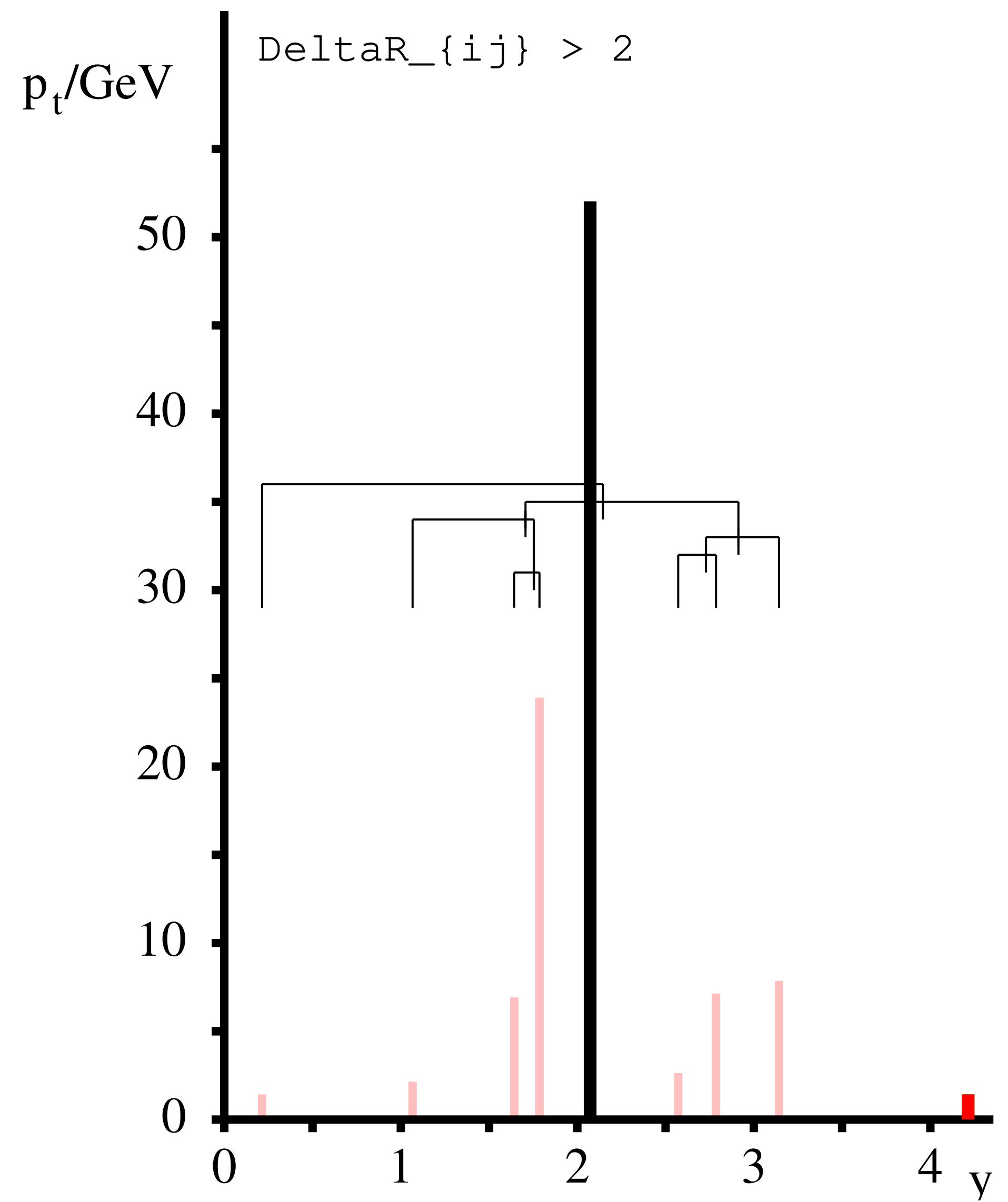
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard "blobs"

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



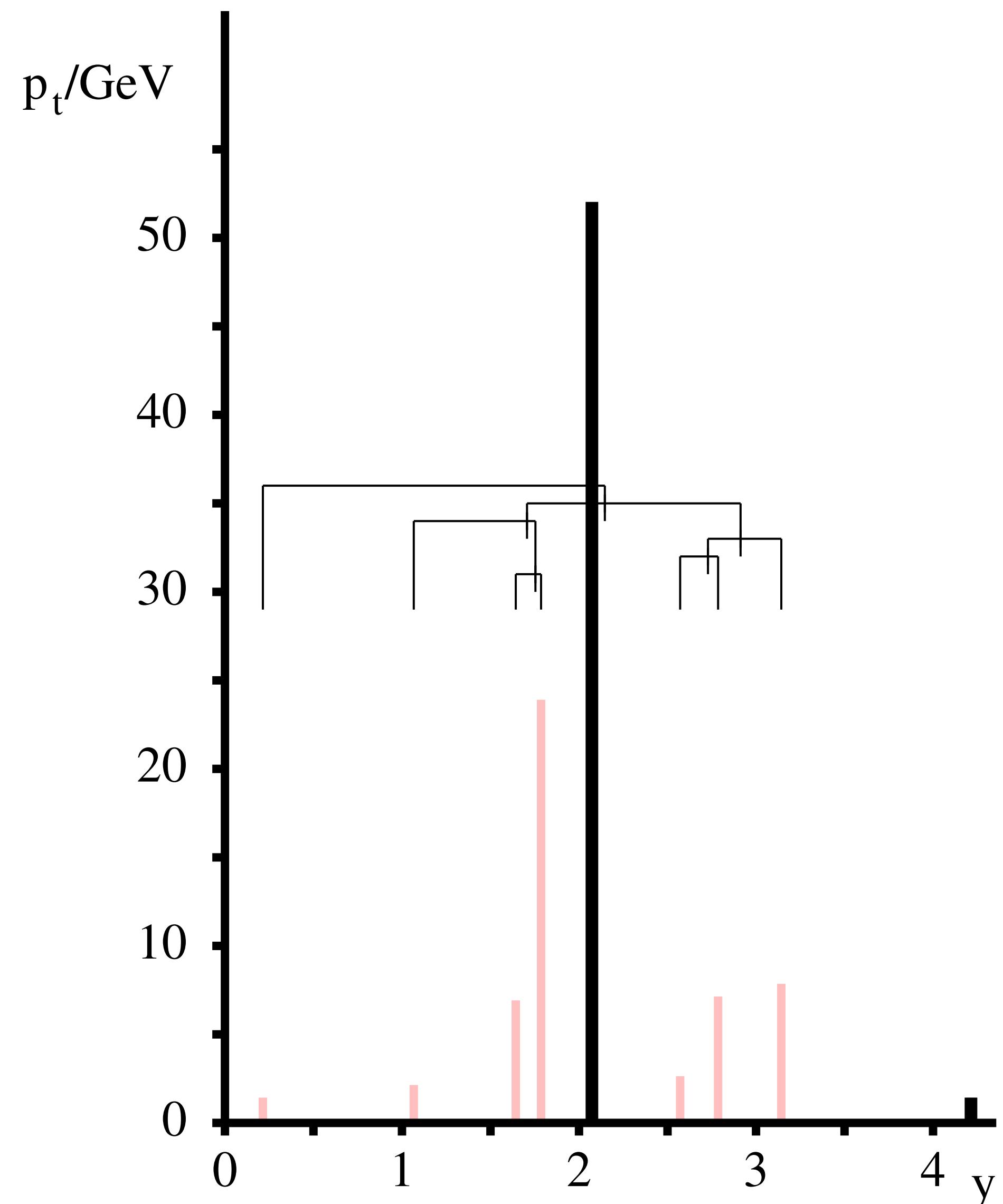
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



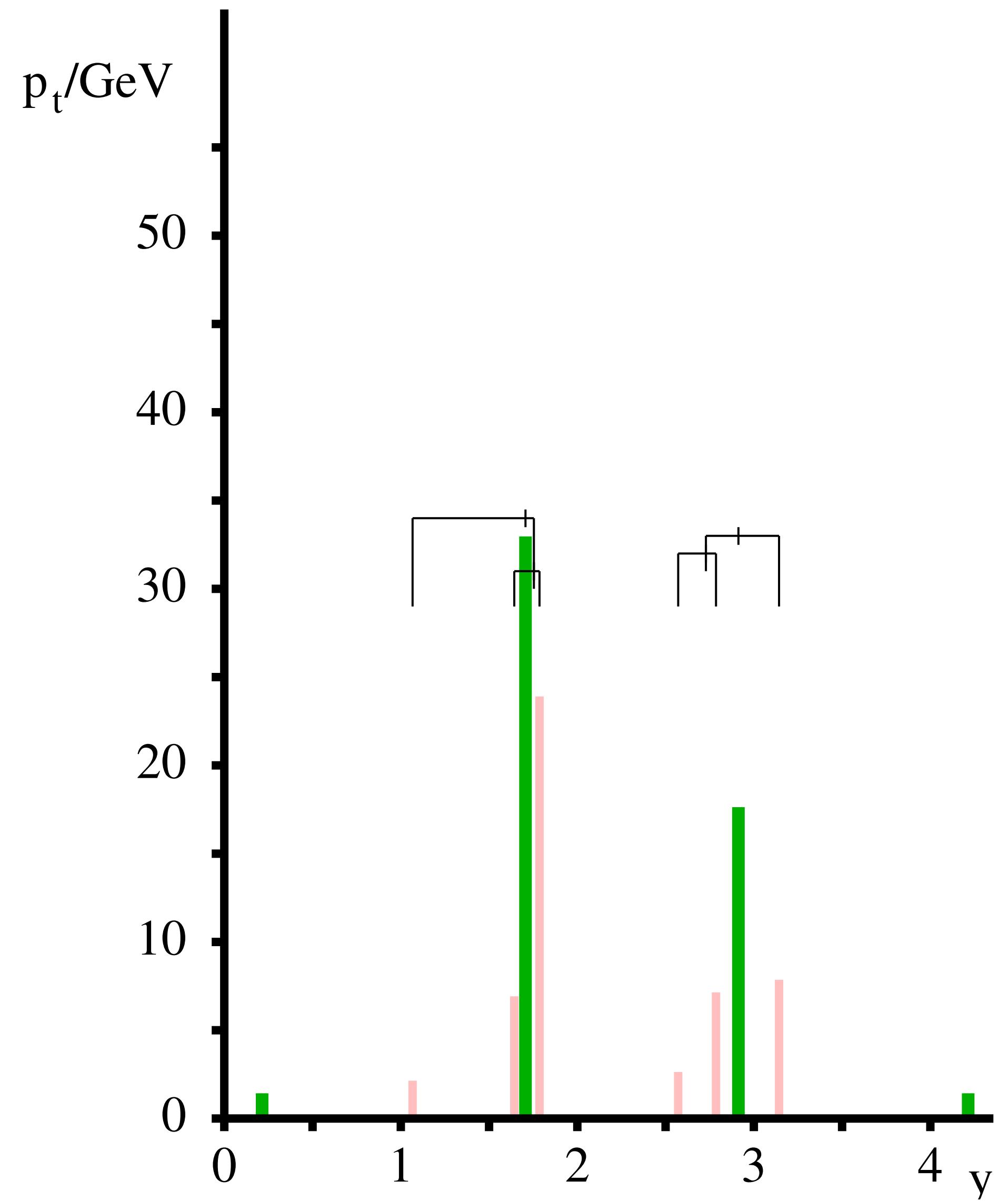
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end

# Identifying jet substructure: Cam/Aachen



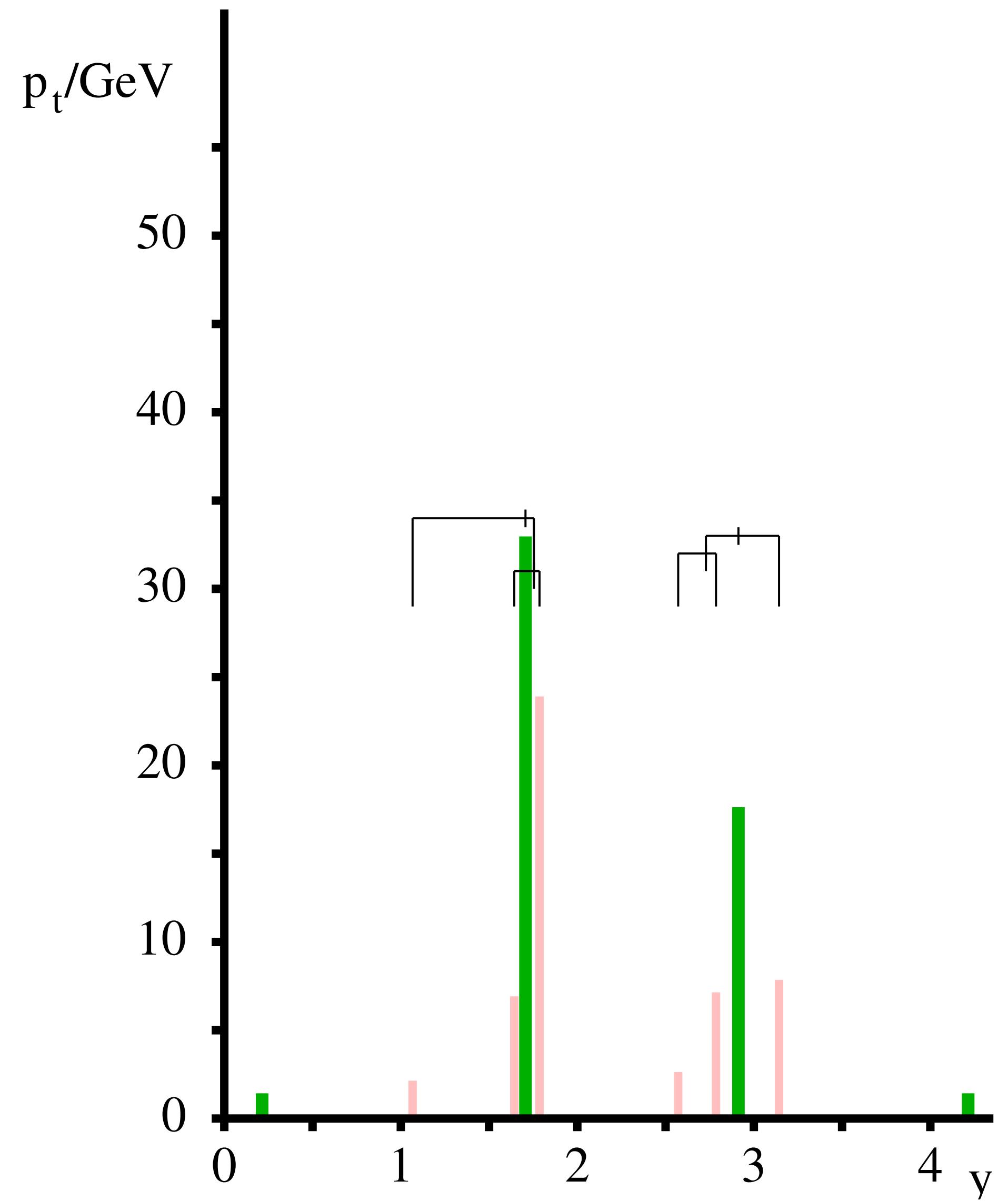
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end **(and is what SoftDrop removes)**

# Identifying jet substructure: Cam/Aachen



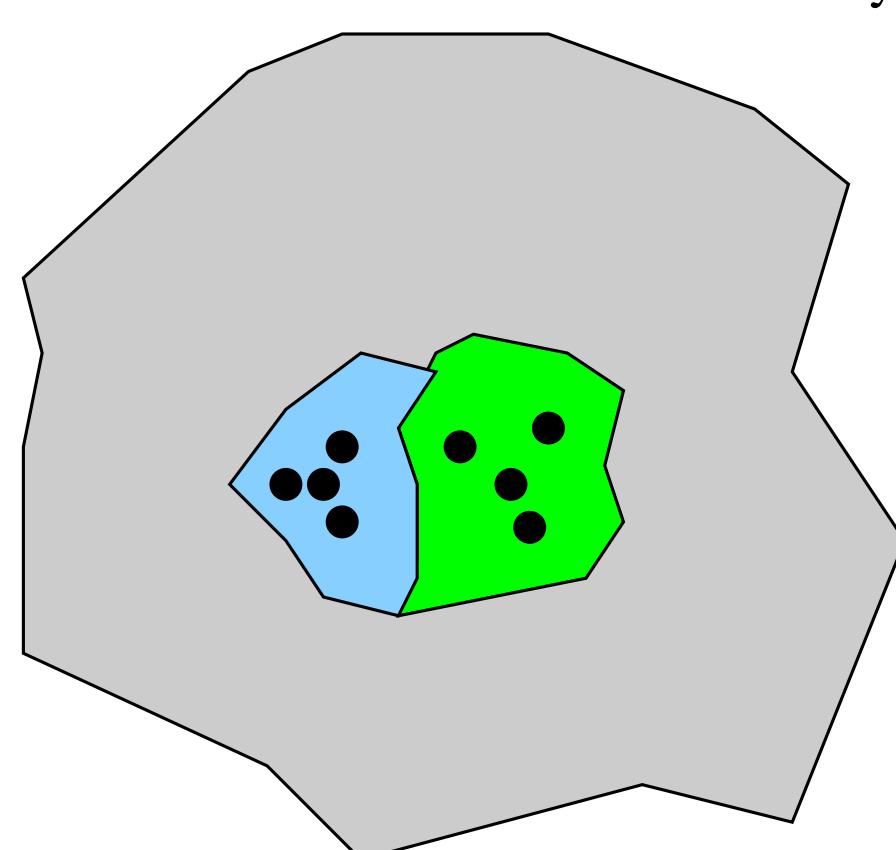
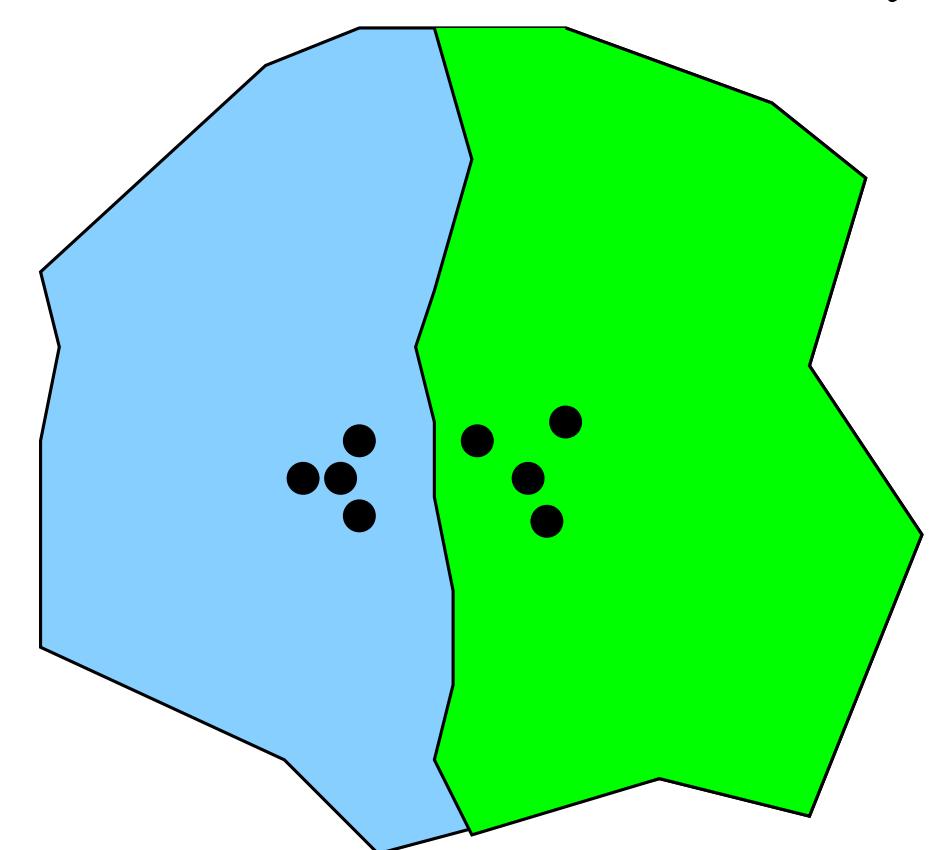
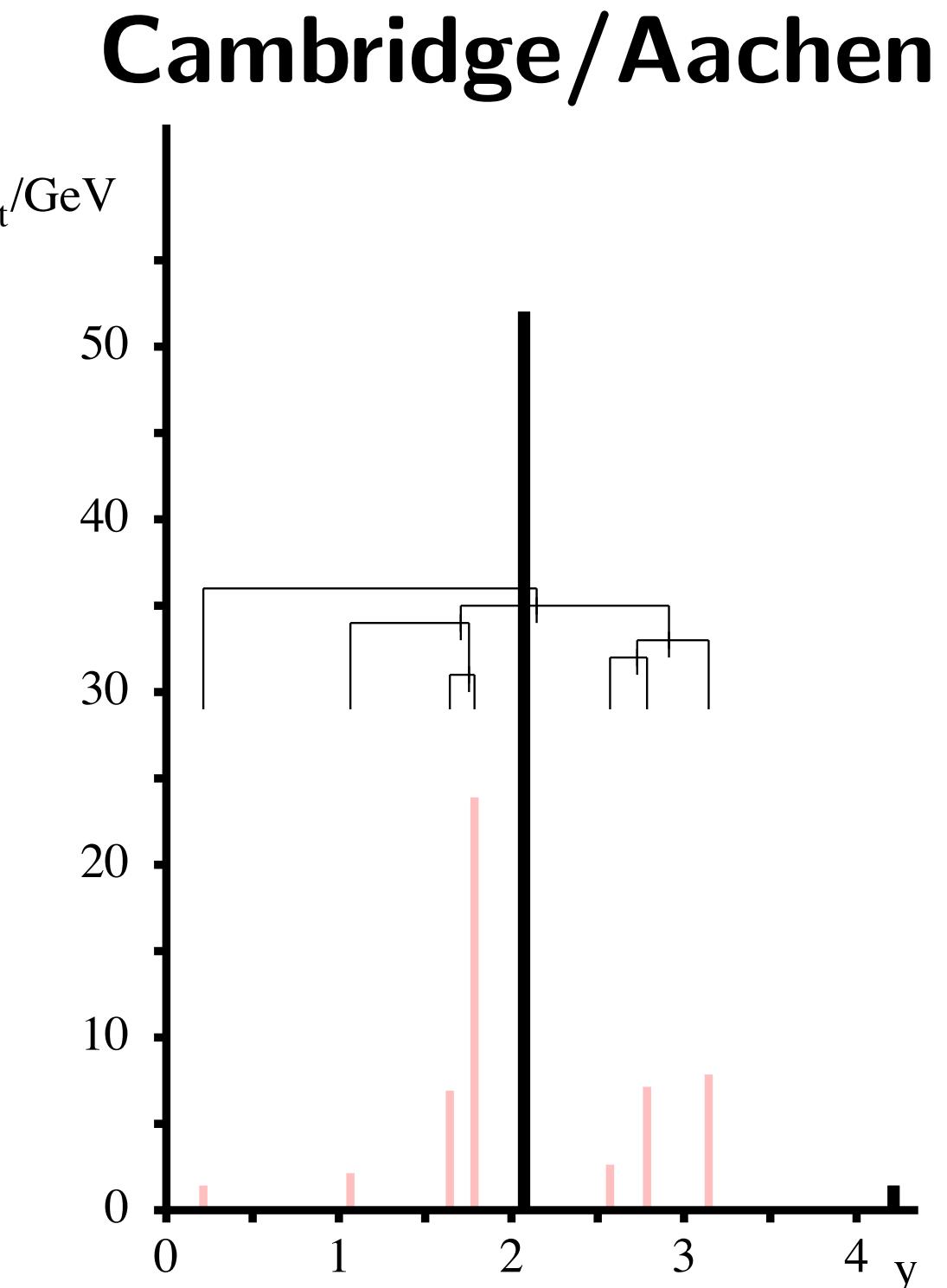
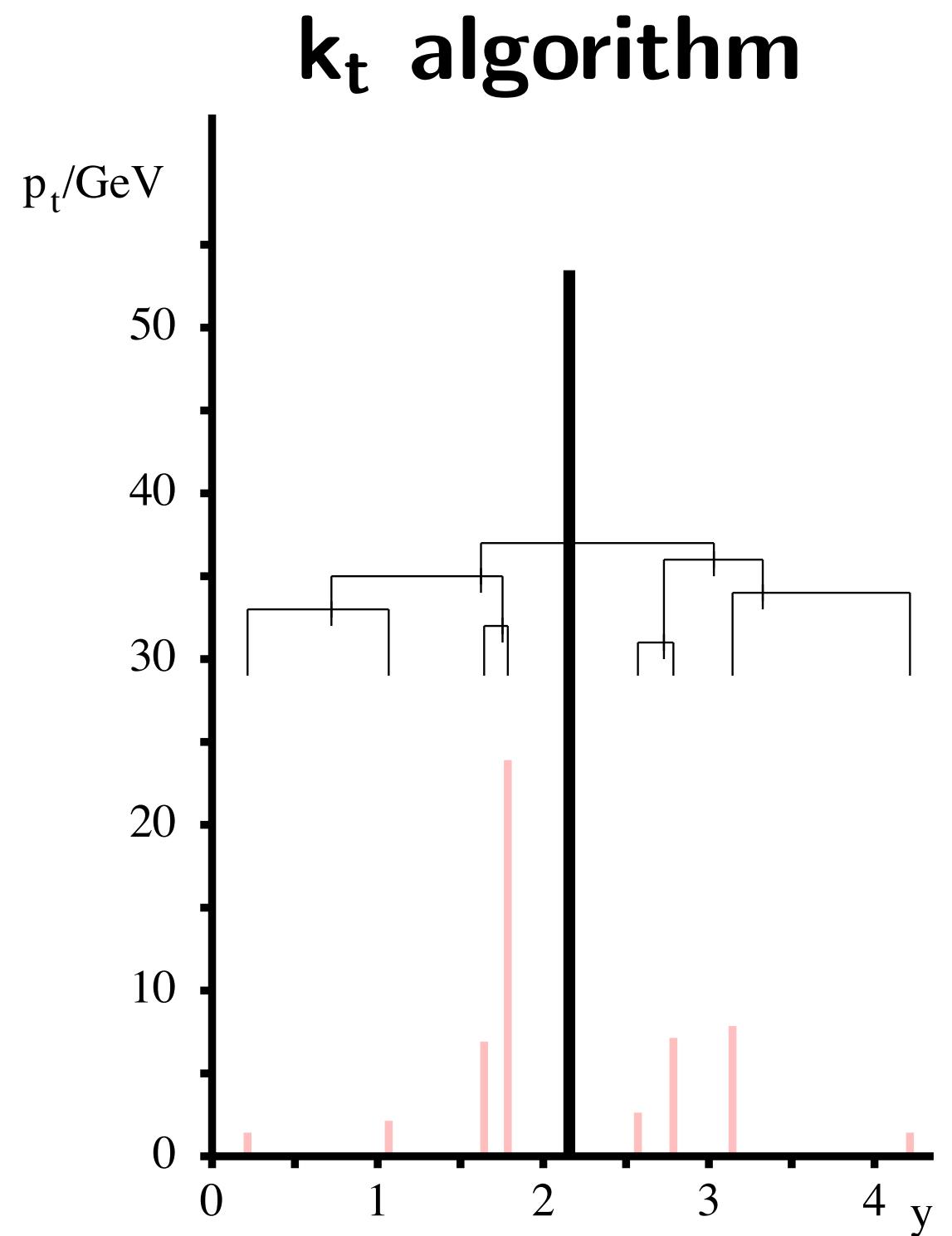
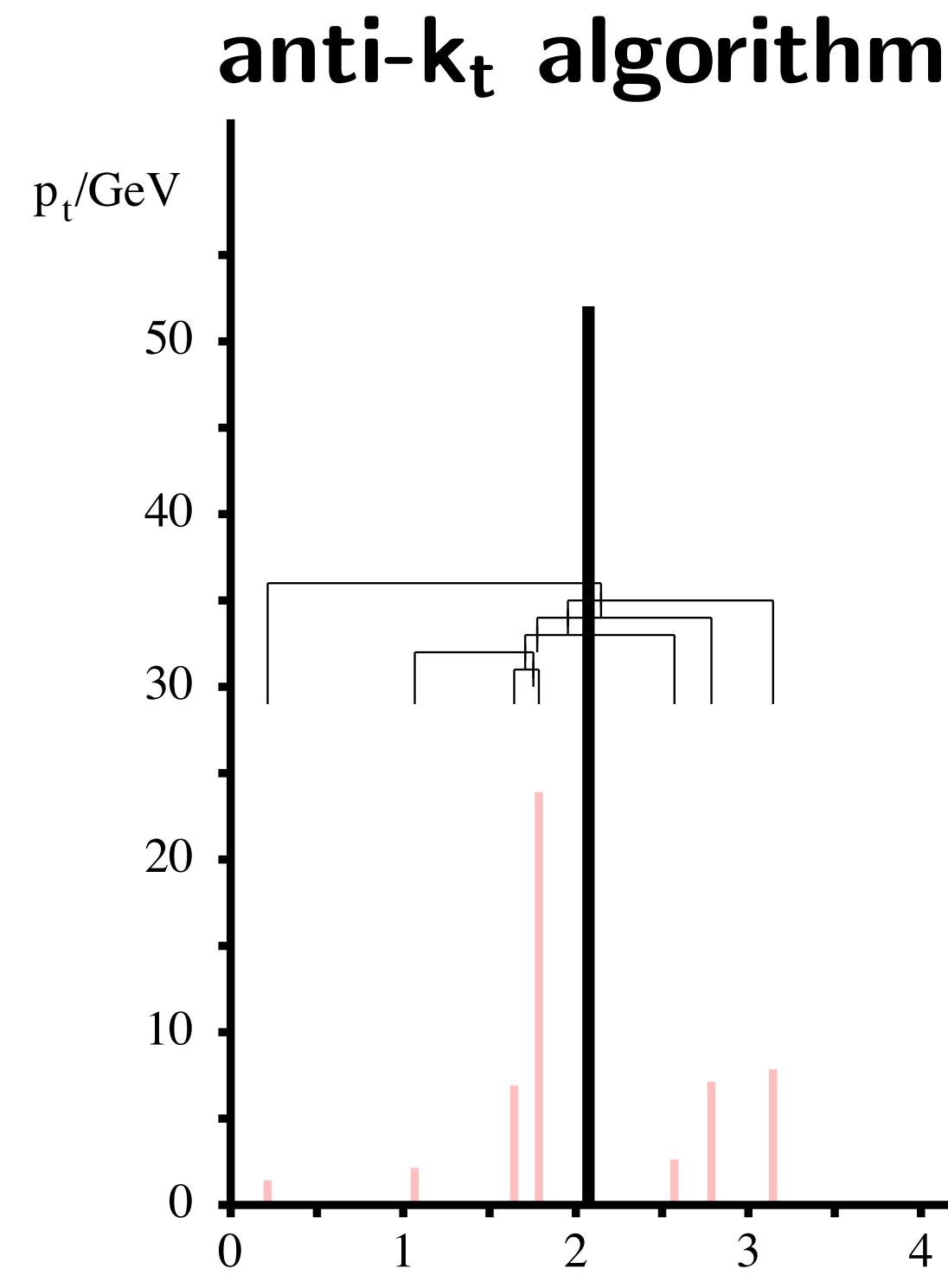
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A algorithm works hierarchically in angle

Clear ID of two separate hard “blobs”

Soft junk gets added in at the end **(and is what SoftDrop removes)**

# Comparing clustering trees



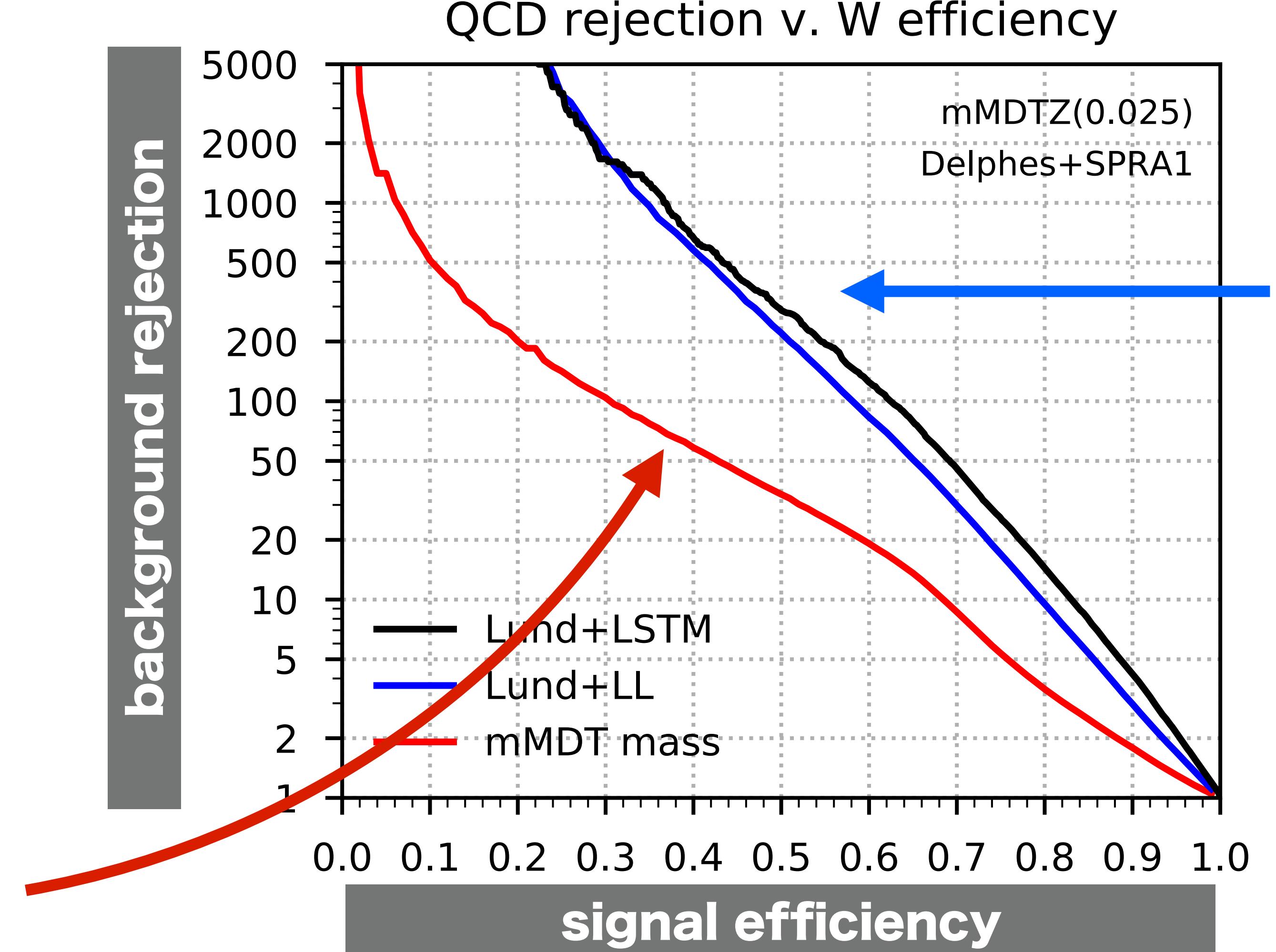
# Tagging two-prong decays

## 4. Looking beyond hard substructure

# using full jet information for W boson tagging

Figure from  
Frédéric Dreyer

*QCD rejection with  
just jet mass  
(SD/mMDT)  
i.e. 2008 tools &  
their descendants*



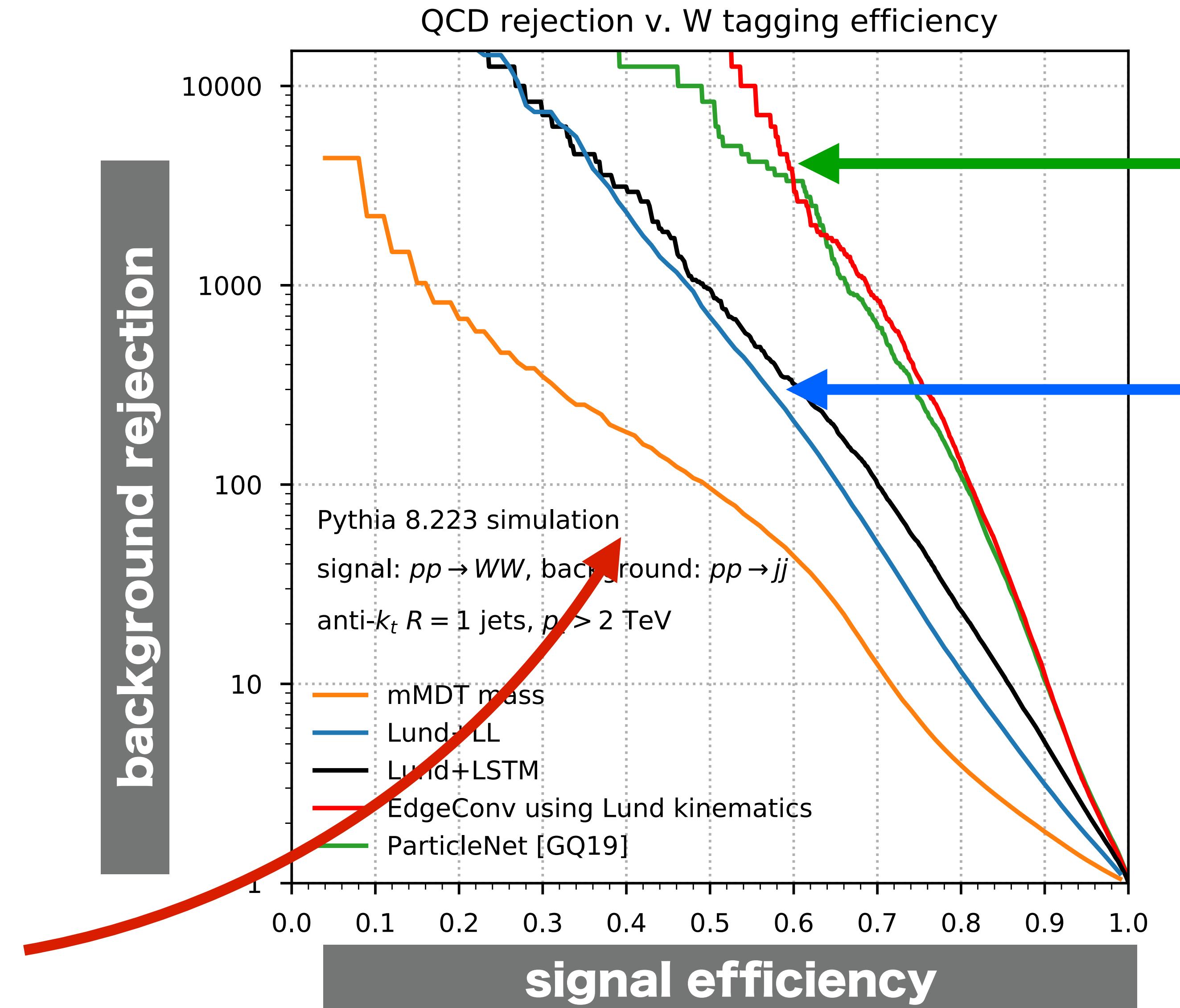
*QCD rejection with use  
of full jet  
substructure  
(2018 ML tools)  
5–10x better*

*First started to be exploited  
by Thaler & Van Tilburg with  
“N-subjettiness” (2010/11)*

# using full jet information for W boson tagging

Figure from  
Frédéric Dreyer

*QCD rejection with  
just jet mass  
(SD/mMDT)  
i.e. 2008 tools &  
their descendants*

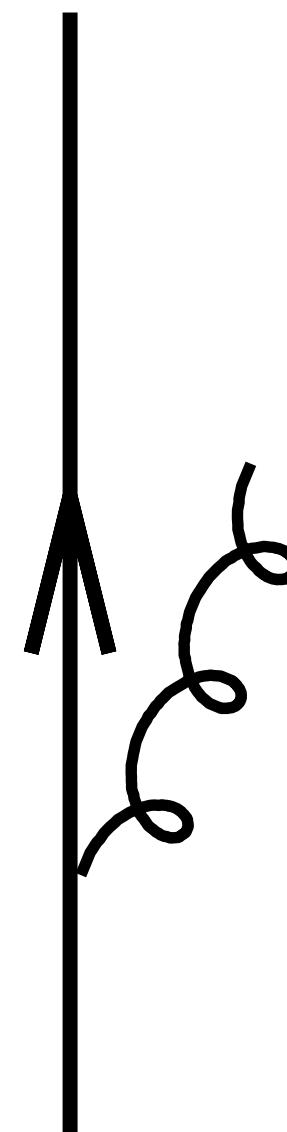


*2019/20 ML tools  
another x10*

*QCD rejection with use  
of full jet  
substructure  
(2018 ML tools)  
5–10x better*

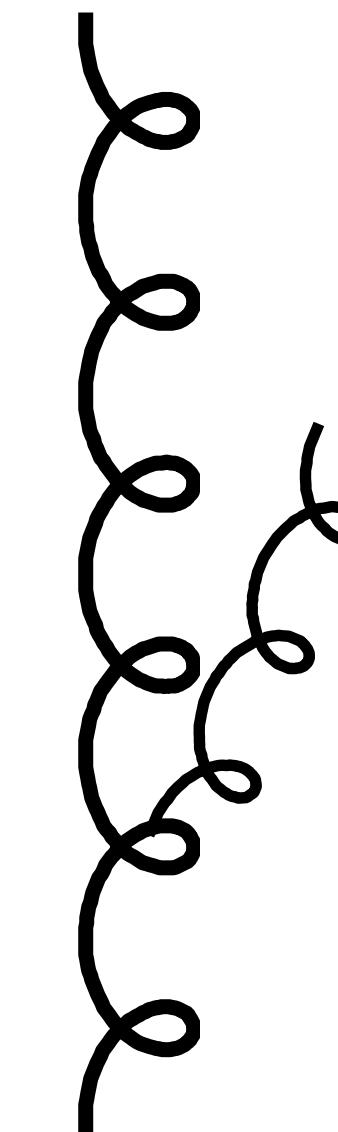
*First started to be exploited  
by Thaler & Van Tilburg with  
“N-subjettiness” (2010/11)*

# Intensity of gluon radiation from a single parton



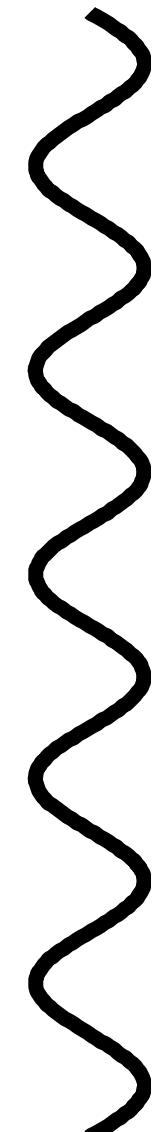
$$\frac{C_F \alpha_s}{\pi} \frac{dE}{E} \frac{d\theta^2}{\theta^2}$$

$$C_F = \frac{4}{3}$$



$$\frac{C_A \alpha_s}{\pi} \frac{dE}{E} \frac{d\theta^2}{\theta^2}$$

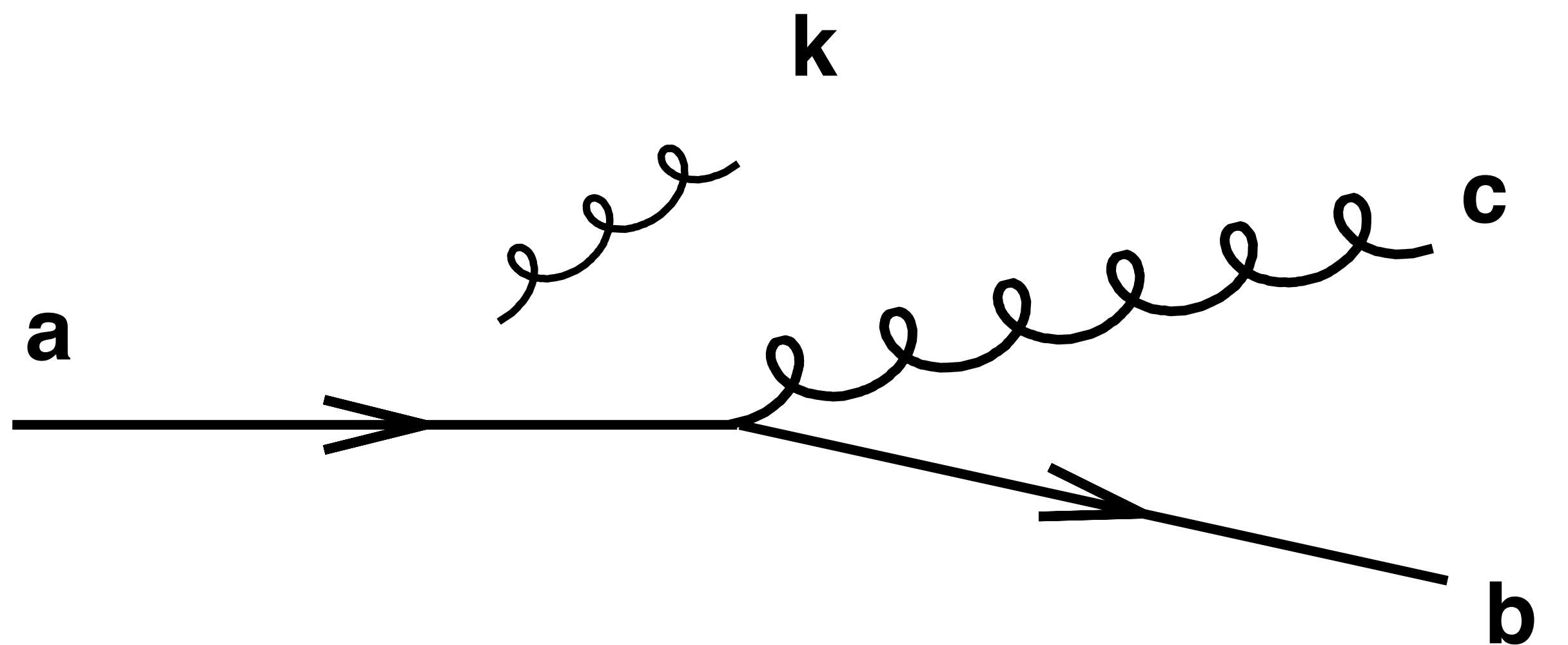
$$C_A = 3$$



$$0$$

Z/W/H don't  
radiate gluons

# radiation from a multi-parton system



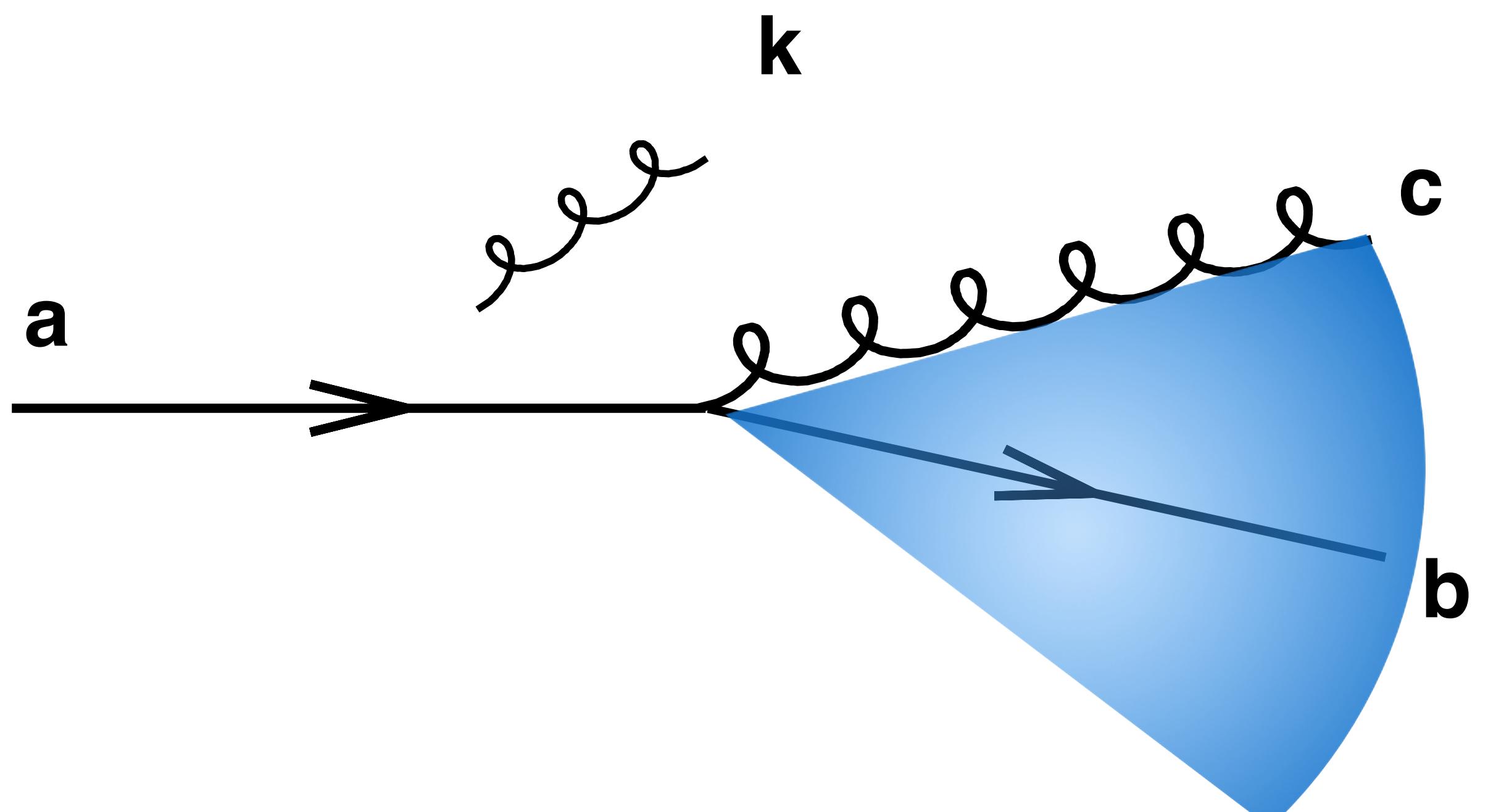
$$C_F \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

$$C_A \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# radiation from a multi-parton system



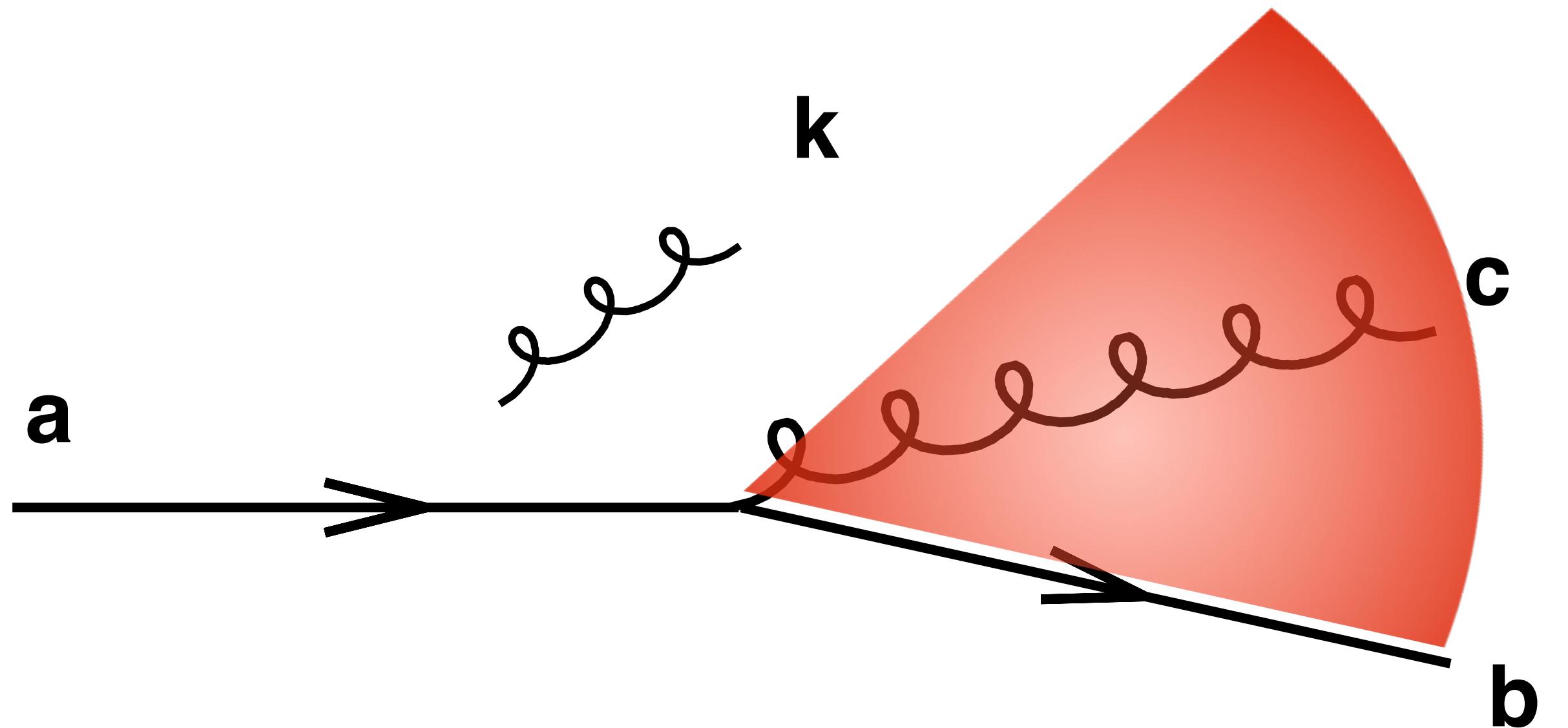
$$C_F \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

$$C_A \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# radiation from a multi-parton system



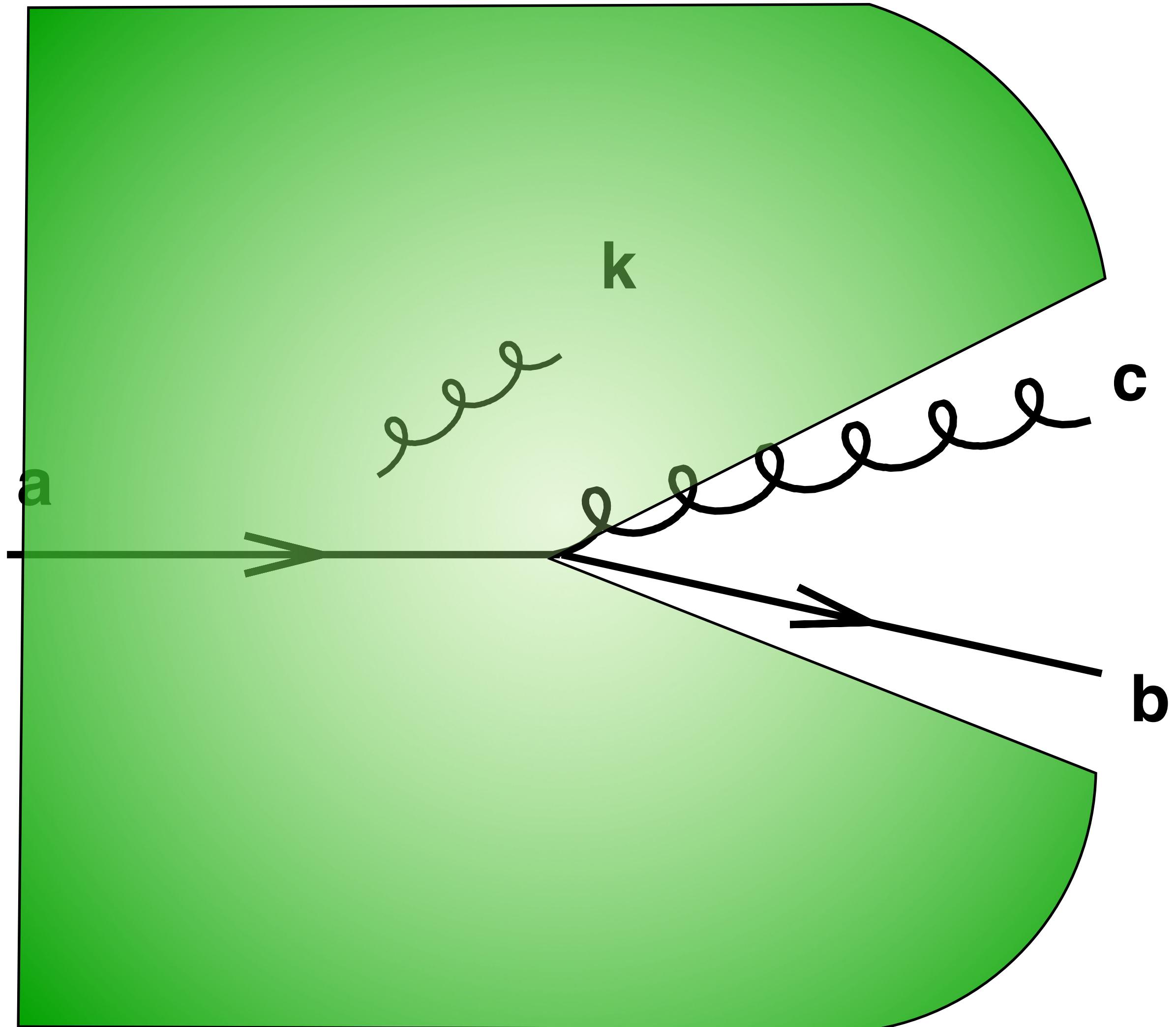
$$C_F \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

$$C_A \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# radiation from a multi-parton system



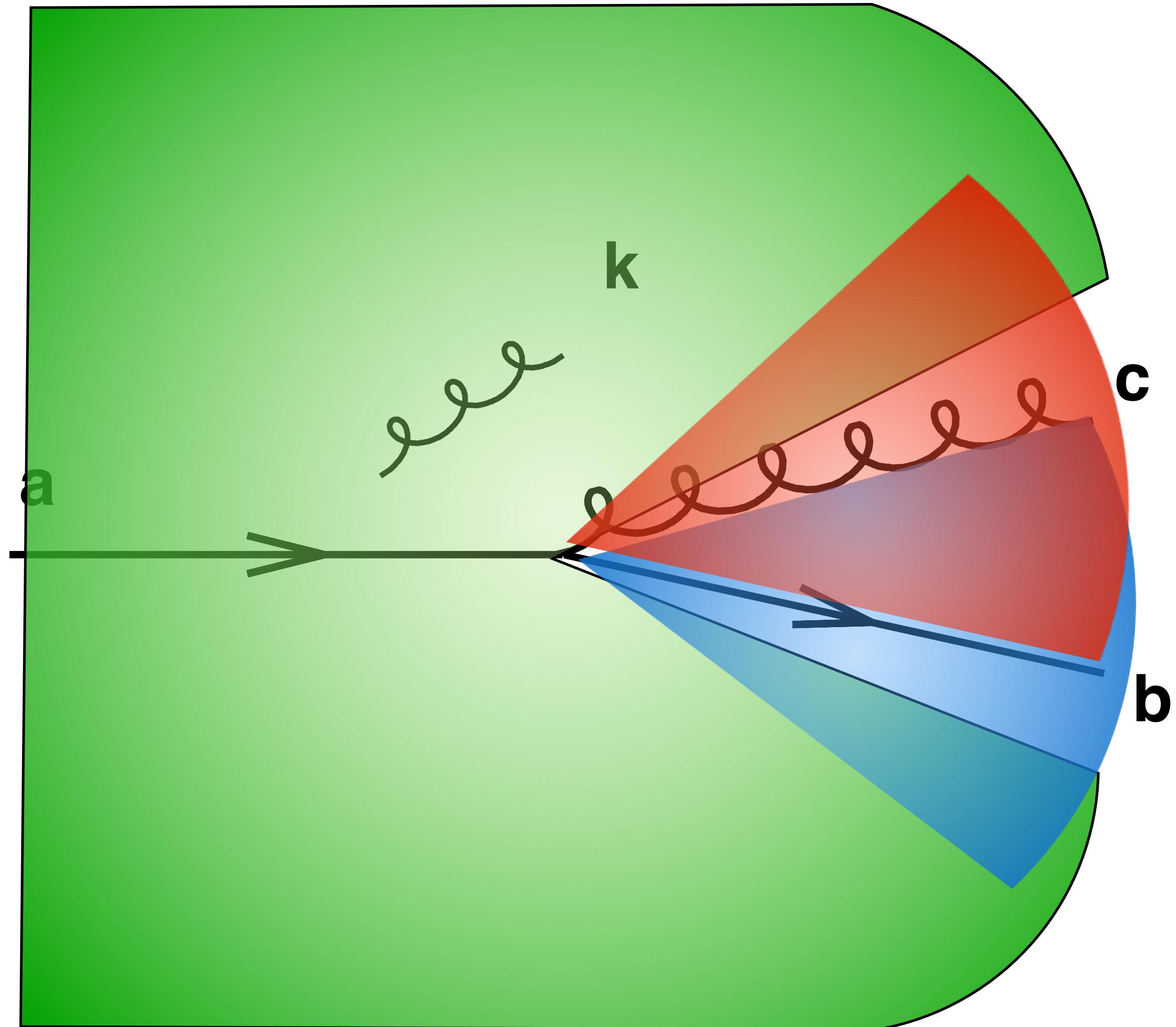
$$C_F \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

$$C_A \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# radiation from a multi-parton system



$$C_F \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

$$C_A \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

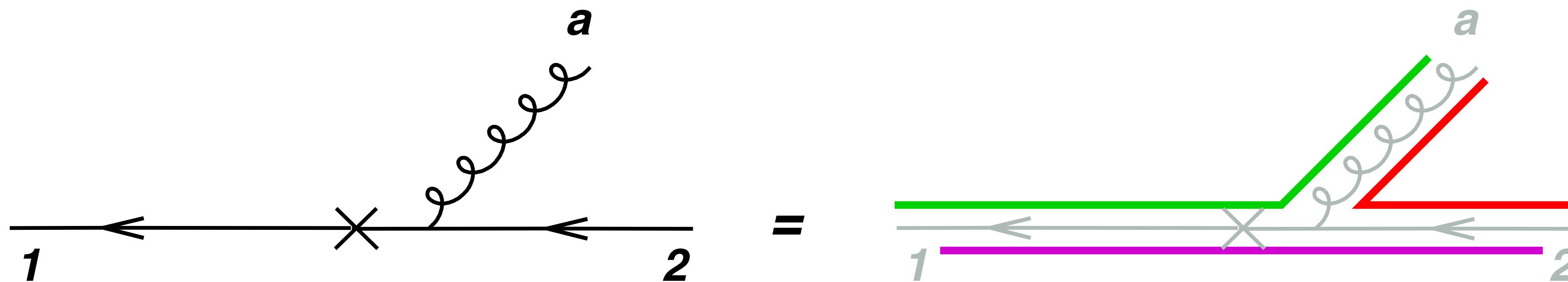
angular ordering approximation

# expression without angular ordering approximation

Write as  $|M_{q\bar{q}a}^2| \times (\text{emission of } b)$

$$|M_{q\bar{q}ab}|^2 = |M_{q\bar{q}a}|^2 2g_s^2 \left( -\frac{1}{2N_c} \frac{(12)}{(1b)(b2)} + \frac{N_c}{2} \frac{(a1)}{(ab)(b1)} + \frac{N_c}{2} \frac{(a2)}{(ab)(b2)} \right)$$

Note structure as *incoherent sum over dipoles*



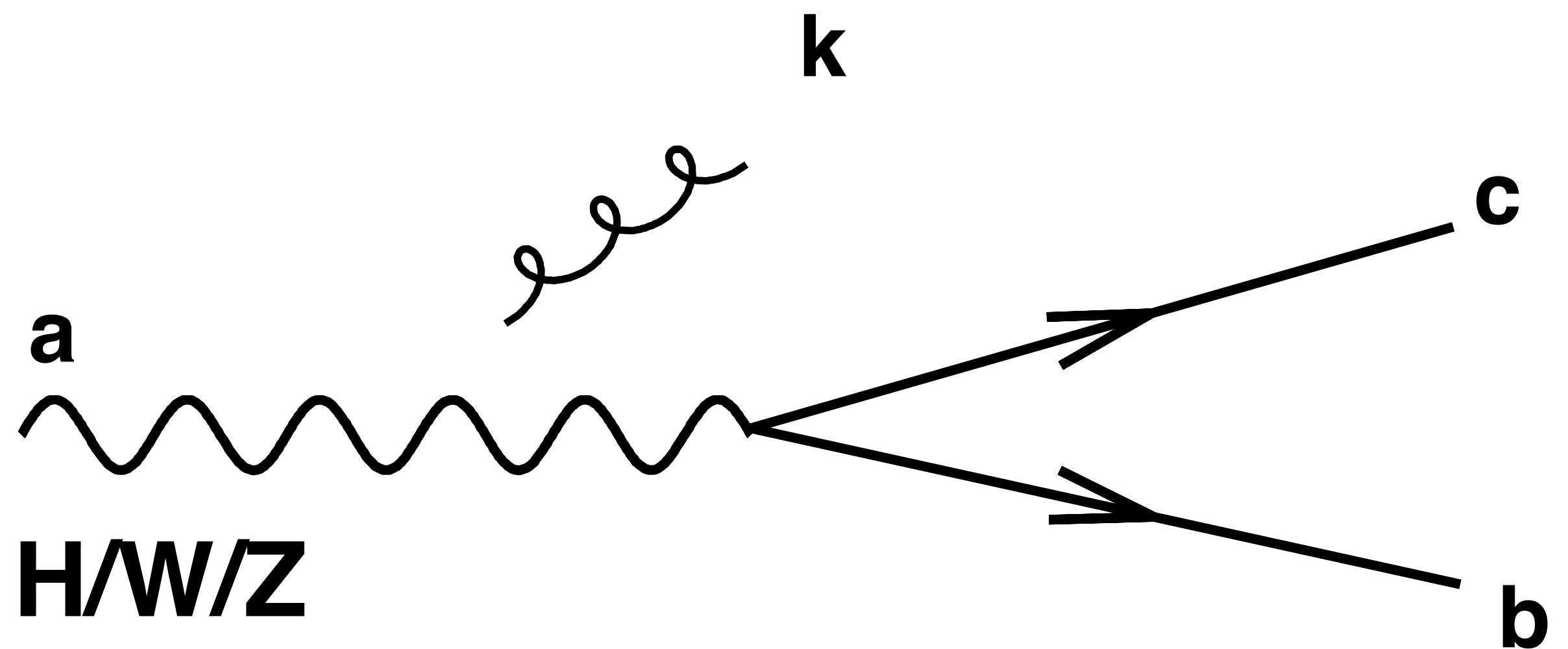
Consistency: off each quark we have

$$-\frac{1}{2N_c} + \frac{N_c}{2} = -\frac{1}{2N_c} + \frac{N_c}{2} = C_F$$

For gluon

$$\frac{N_c}{2} + \frac{N_c}{2} = C_A$$

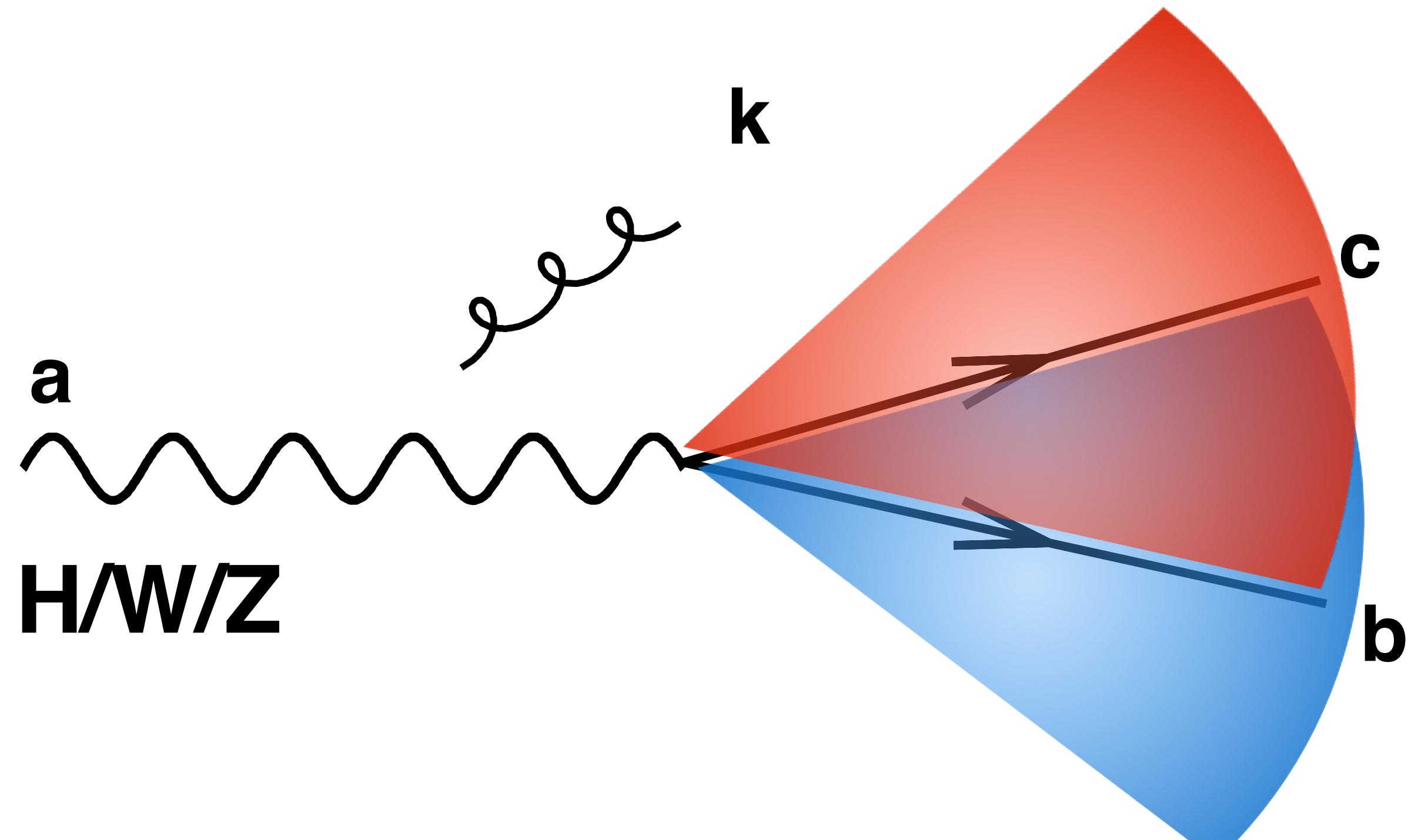
# radiation from a **colour-singlet** multi-parton system



$$0 \times \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$
$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$
$$C_F \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# radiation from a **colour-singlet** multi-parton system



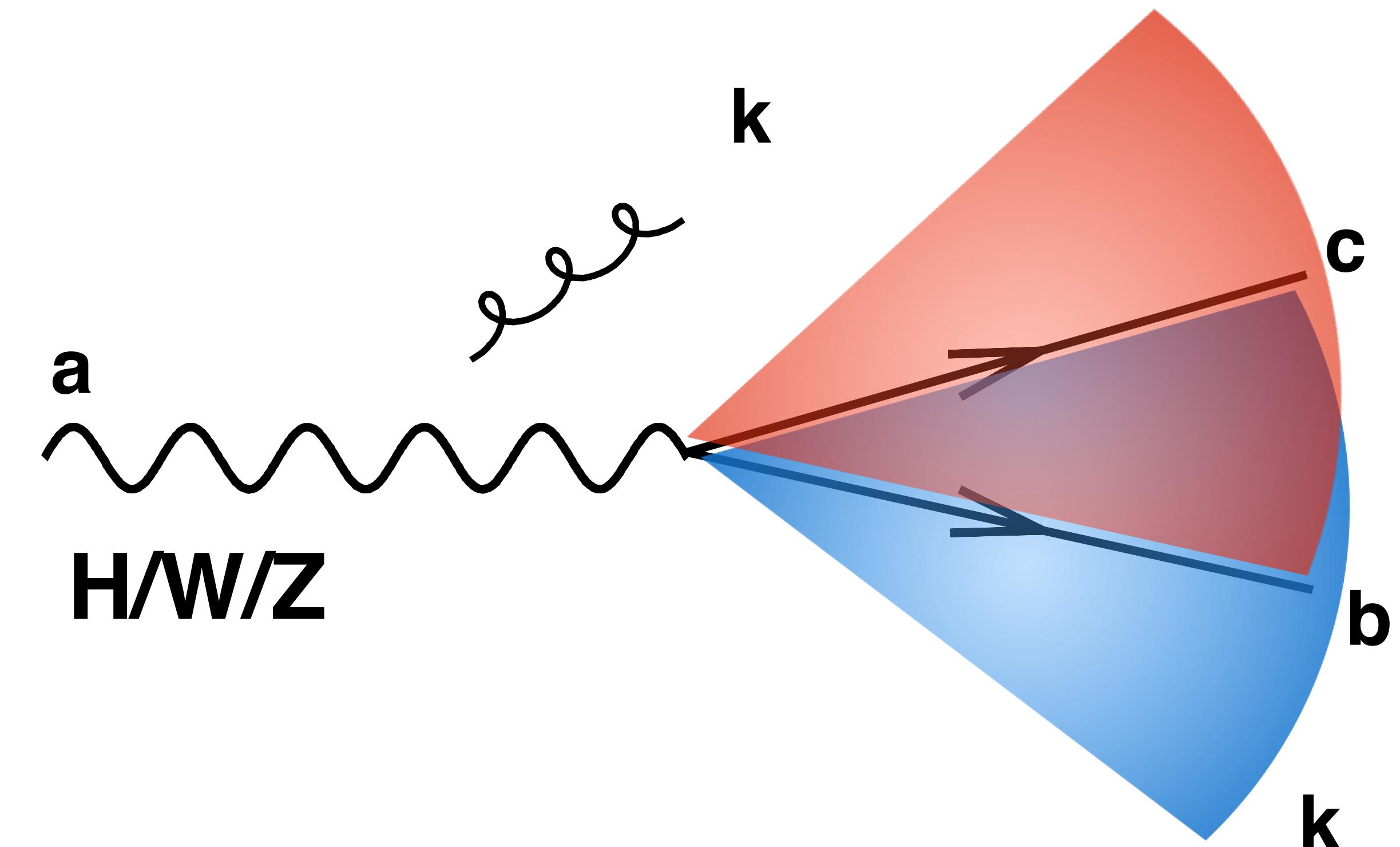
$$0 \times \frac{d\theta_{ak}^2}{\theta_{ak}^2} \Theta(\theta_{ak} \gtrsim \theta_{bc})$$

$$C_F \frac{d\theta_{bk}^2}{\theta_{bk}^2} \Theta(\theta_{bk} \lesssim \theta_{bc})$$

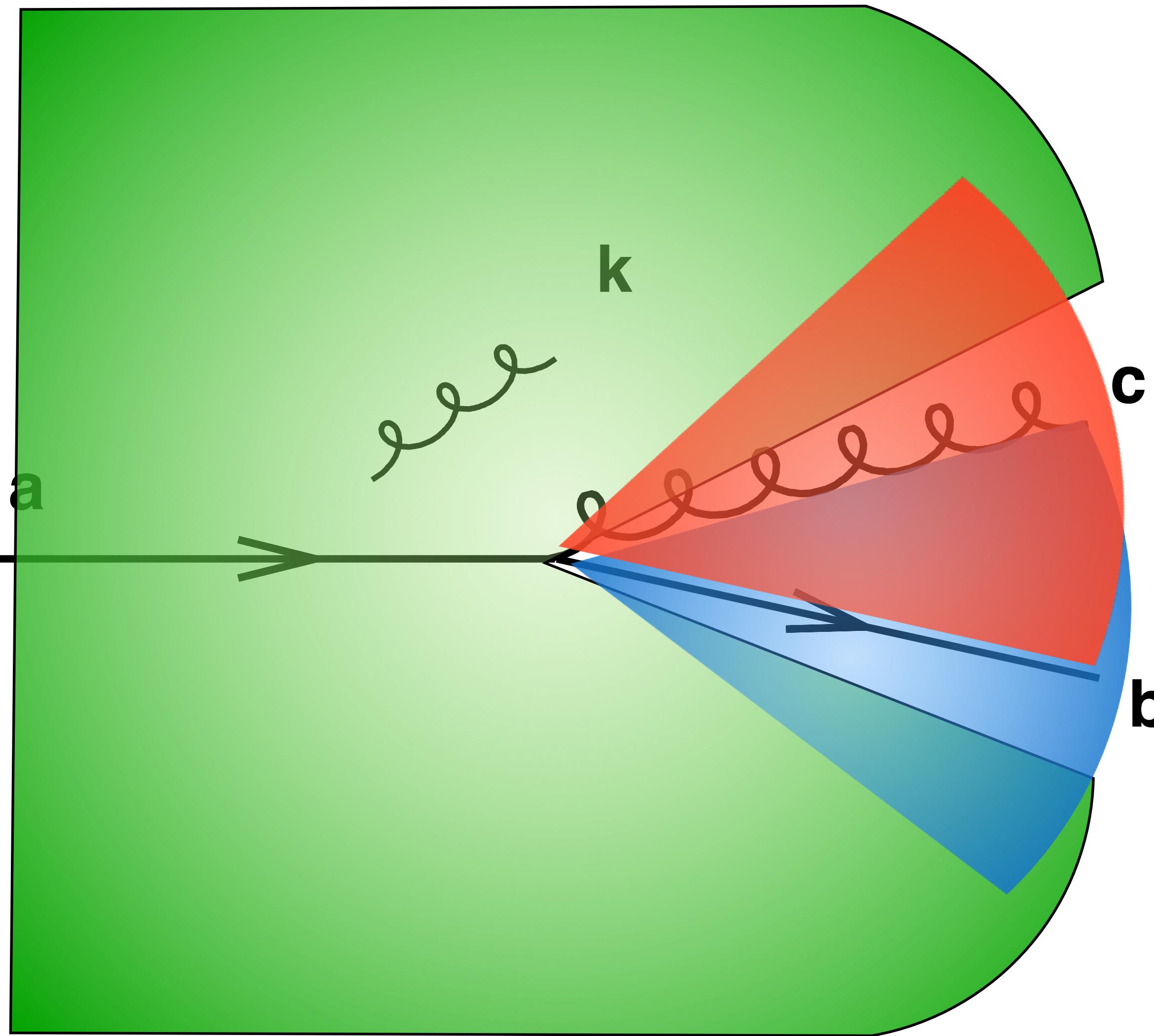
$$C_F \frac{d\theta_{ck}^2}{\theta_{ck}^2} \Theta(\theta_{ck} \lesssim \theta_{bc})$$

angular ordering approximation

# signal vs. background radiation patterns



**signal ( $H/W/Z \rightarrow q\bar{q}$ )**



**background ( $q \rightarrow qg$ )**

observables sensitive to  
radiation pattern

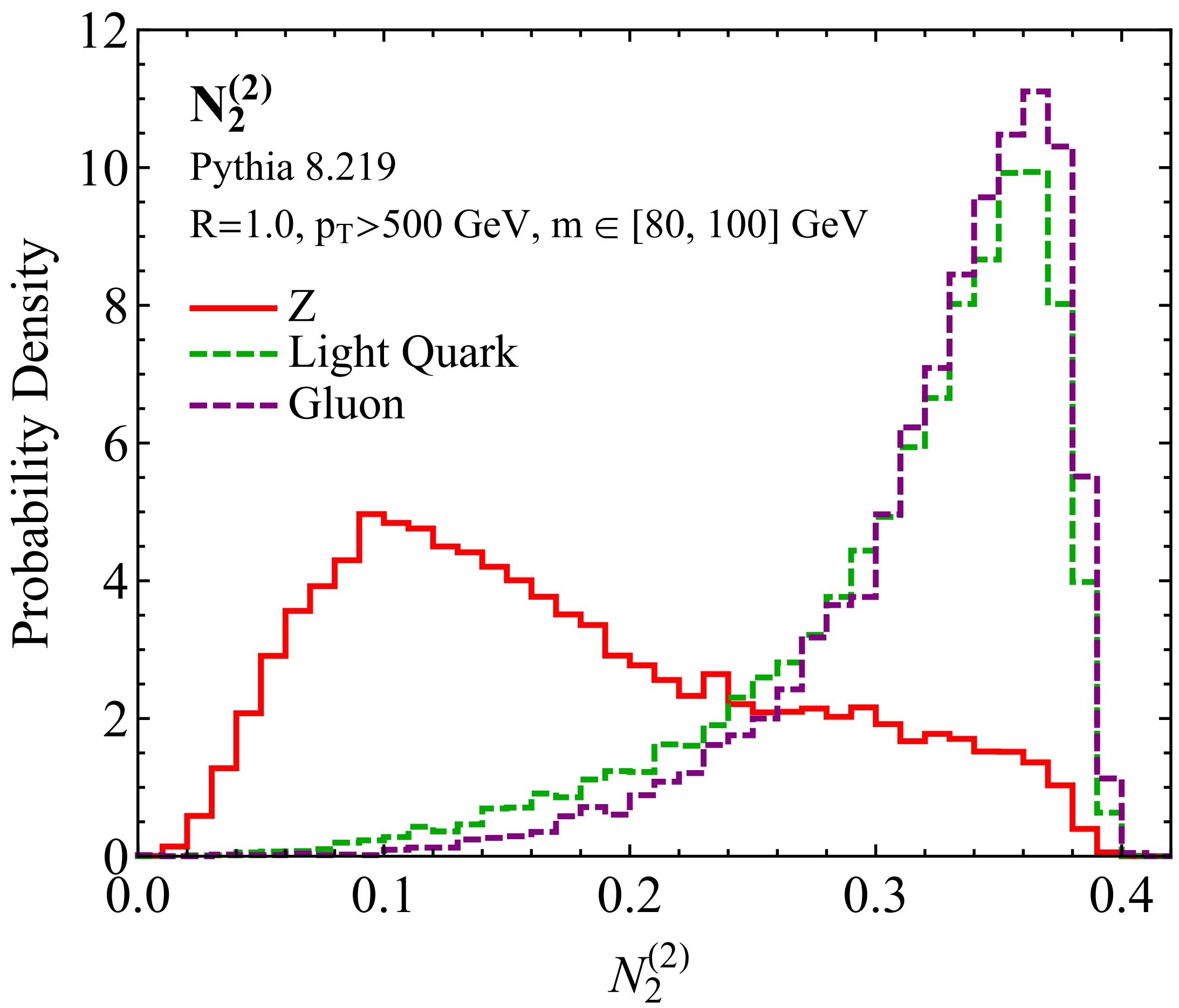
# example of energy flow correlations

$$2e_3^{(\beta)} = \sum_{1 \leq i < j < k \leq n_J} z_i z_j z_k \min \left\{ \theta_{ij}^\beta \theta_{ik}^\beta, \theta_{ij}^\beta \theta_{jk}^\beta, \theta_{ik}^\beta \theta_{jk}^\beta \right\}$$

$${}_1e_2^{(\beta)} \equiv e_2^{(\beta)} = \sum_{1 \leq i < j \leq n_J} z_i z_j \theta_{ij}^\beta,$$

$$N_2 = \frac{2e_3^{(\beta)}}{({}_1e_2^{(\beta)})^2},$$

Moult, Necib & Thaler  
[arXiv:1609.07483](https://arxiv.org/abs/1609.07483)

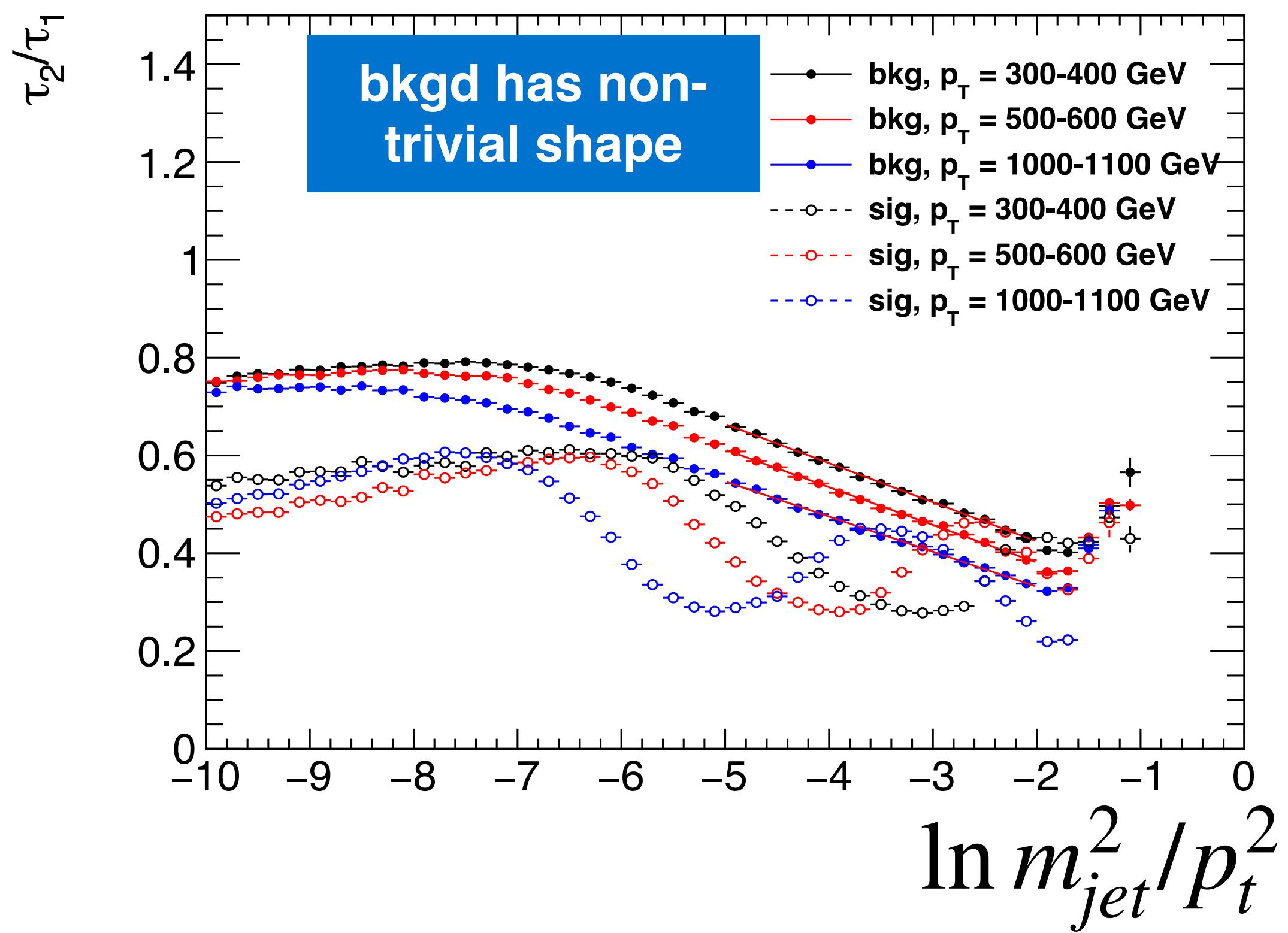


# Decorrelate radiation taggers from jet mass

Original radiation tagger

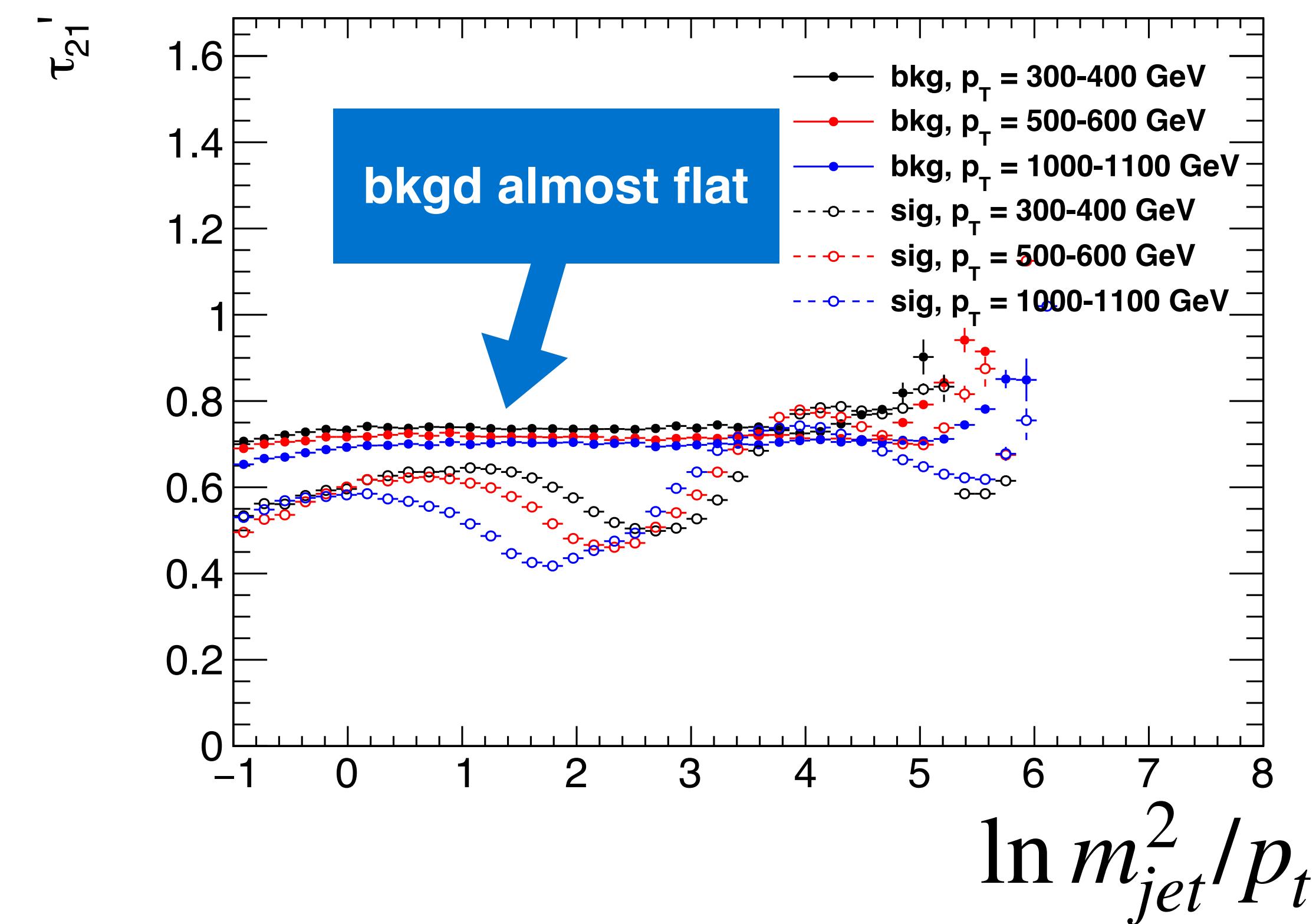
$$\tau_{21} \equiv \tau_2/\tau_1$$

(plot shows avg v. jet mass)



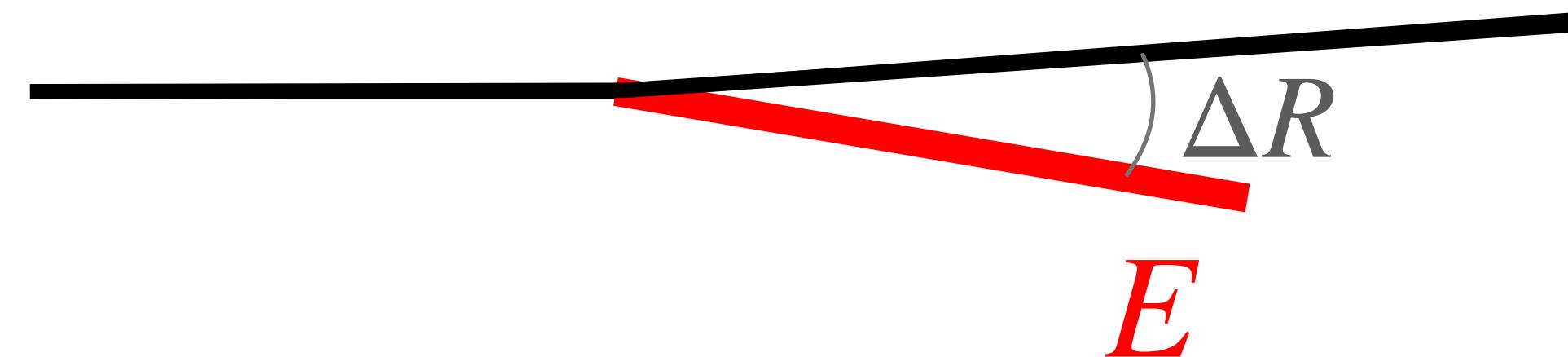
Decorrelated radiation tagger

$$\tau'_{21} \equiv \tau_2/\tau_1 - \text{const} \times \ln m_{jet}^2/p_t$$



Dolen, Harris, Marzani, Rappoccio, Tran, [arXiv:1603.00027](https://arxiv.org/abs/1603.00027)

# Phase space: two key variables (+ azimuth)

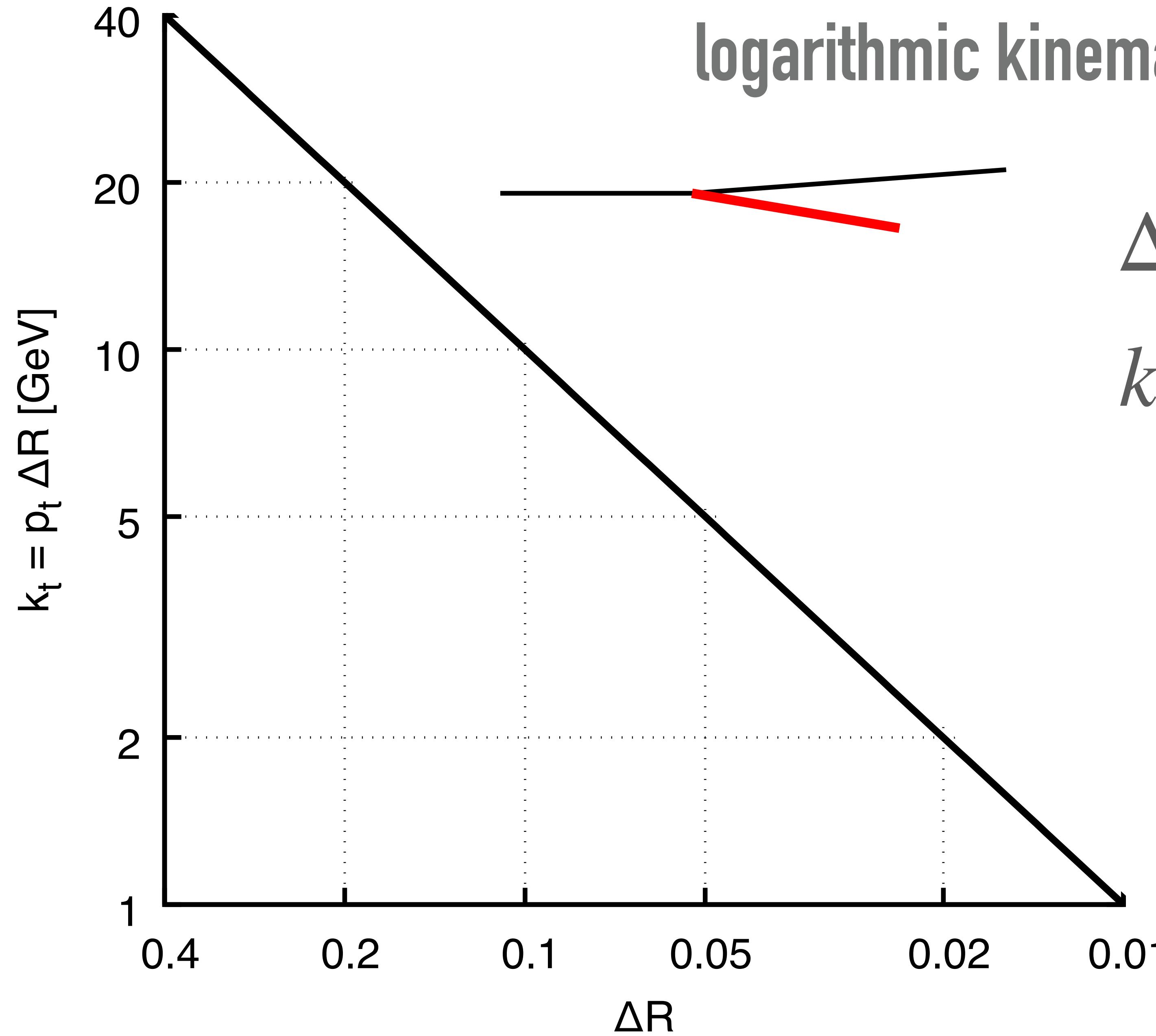


$$\Delta R \text{ (or } \eta = -\ln \tan \frac{\Delta R}{2} \text{)}$$

*$\eta$  is (pseudo)rapidity within the jet*

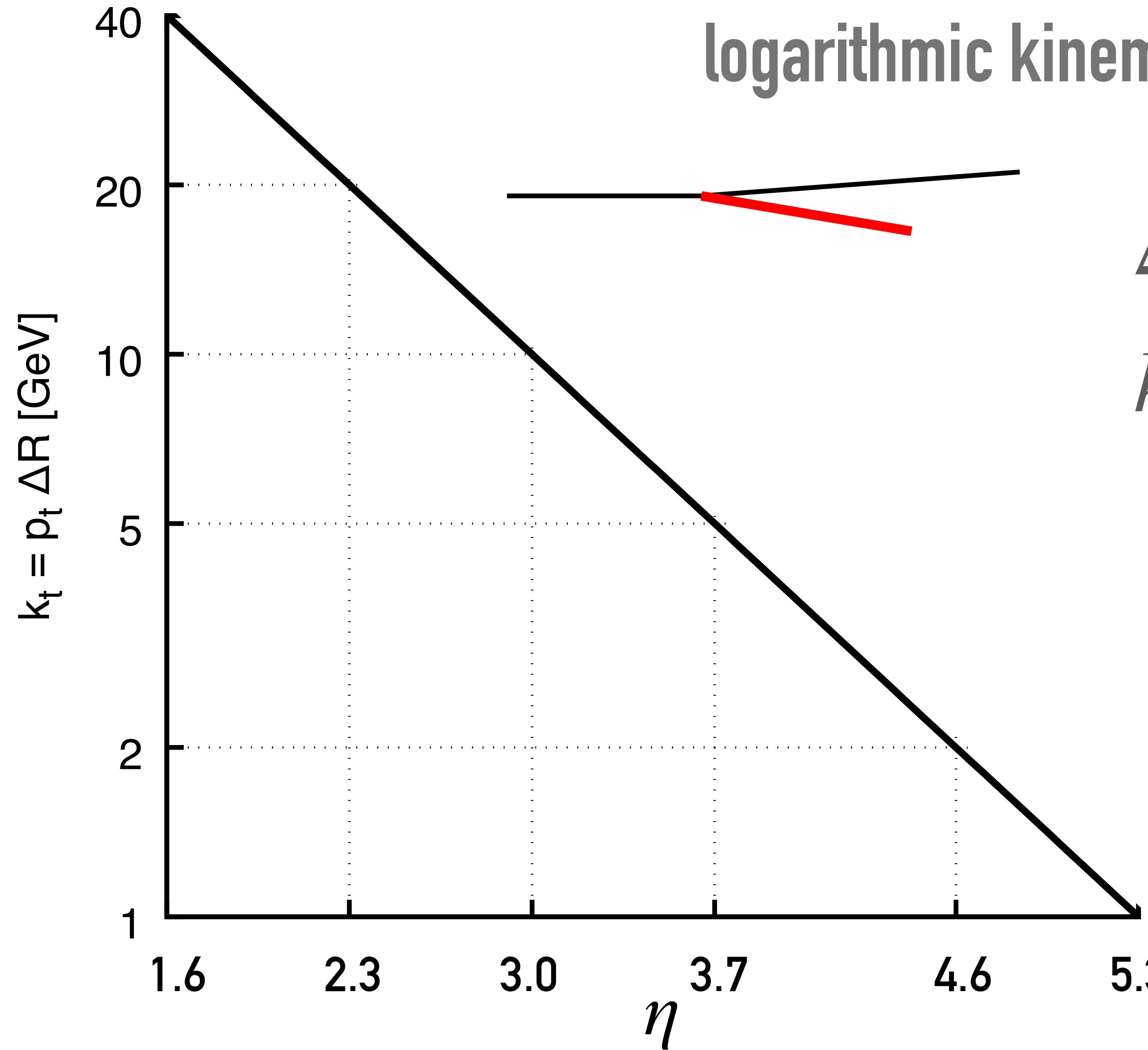
$$k_t = p_t \Delta R$$

*$p_t$  (or  $p_\perp$ ) is transverse momentum within the jet*



Introduced for understanding Parton Shower Monte Carlos by  
B. Andersson, G. Gustafson, L. Lonnblad and Pettersson 1989

The Lund Plane

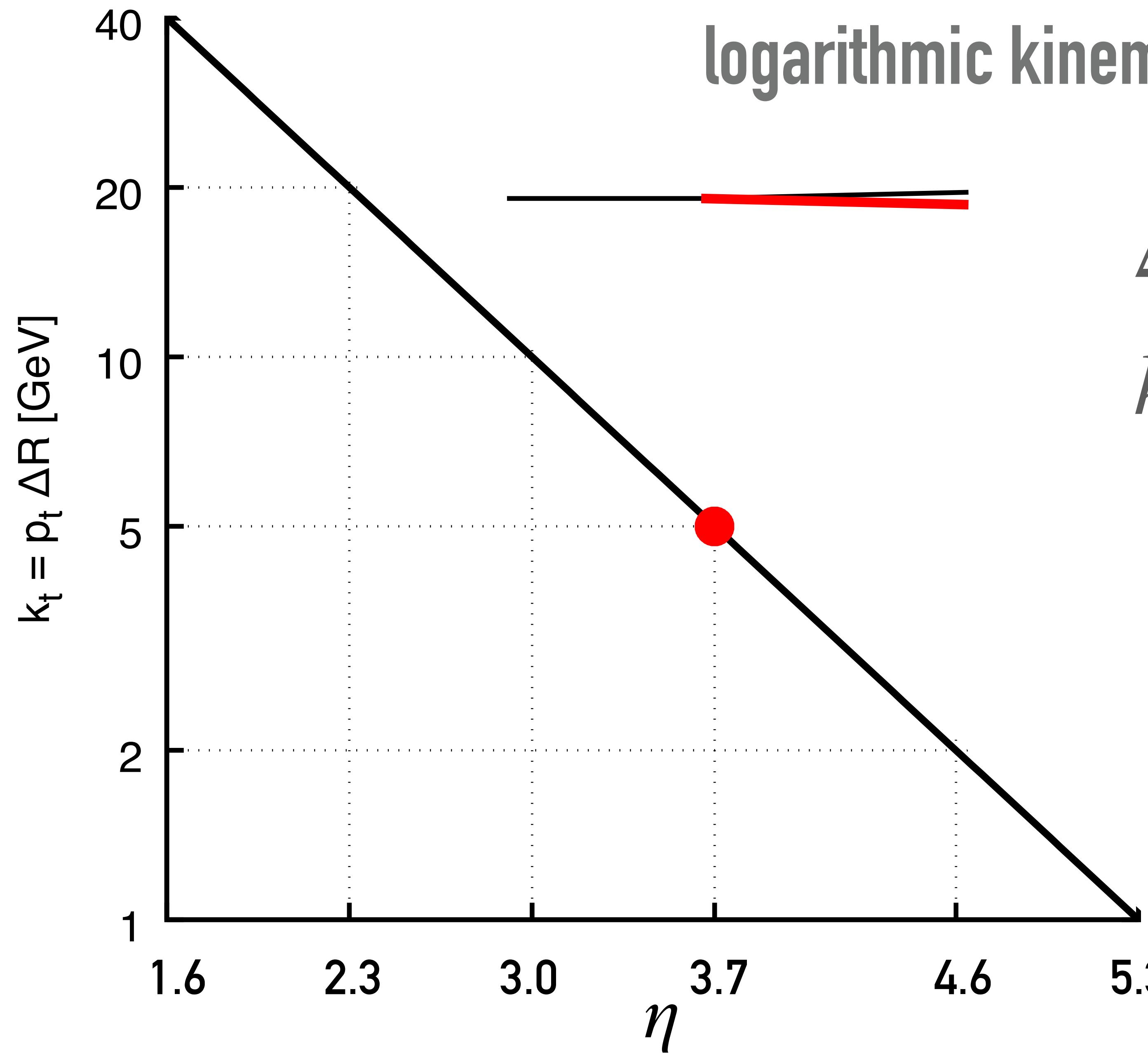


$\Delta R$  (or  $\eta = -\ln \tan \frac{\Delta R}{2}$ )

$k_t = p_t \Delta R$

Introduced for understanding Parton Shower Monte Carlos by  
B. Andersson, G. Gustafson, L. Lonnblad and Pettersson 1989

The Lund Plane

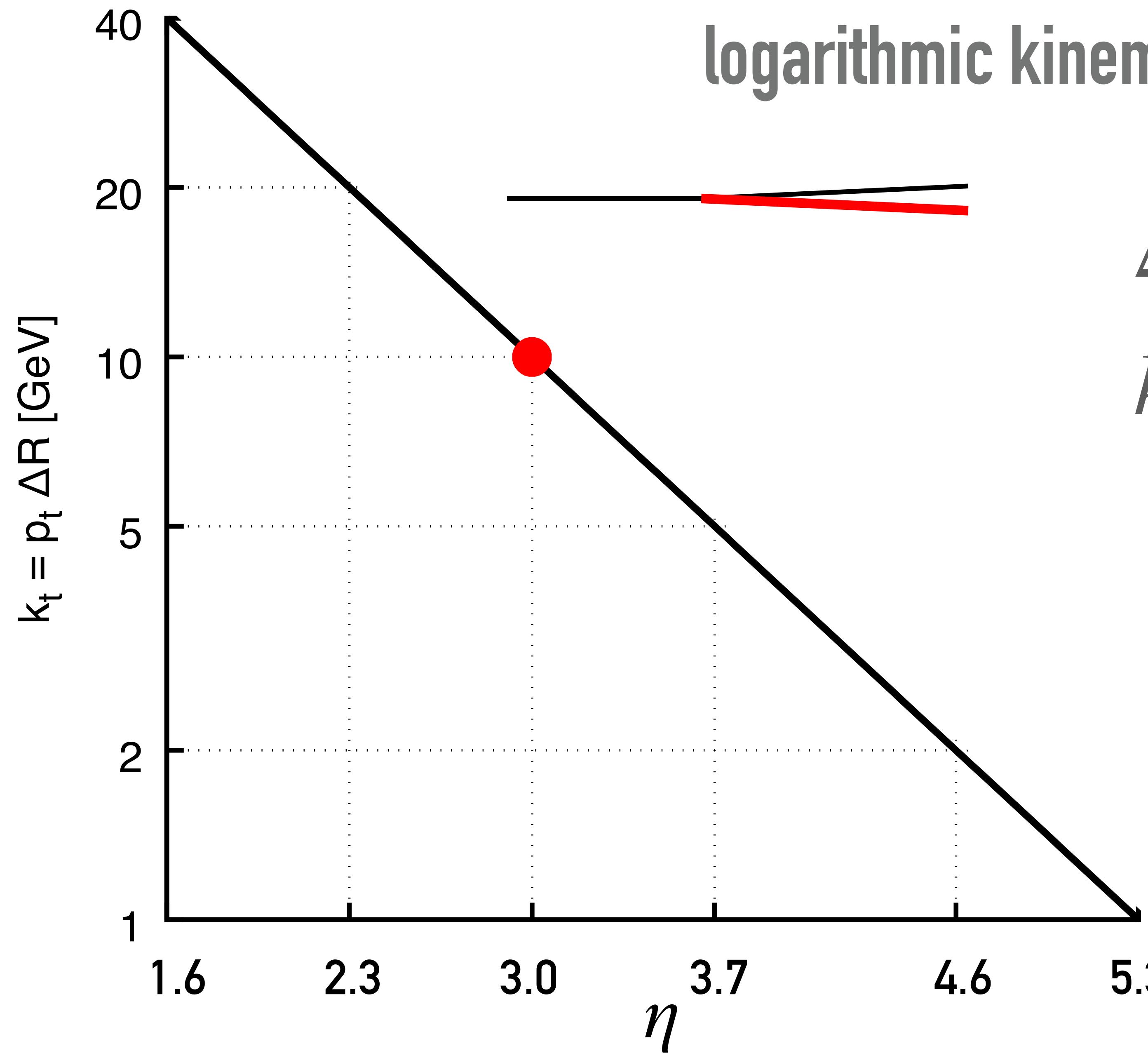


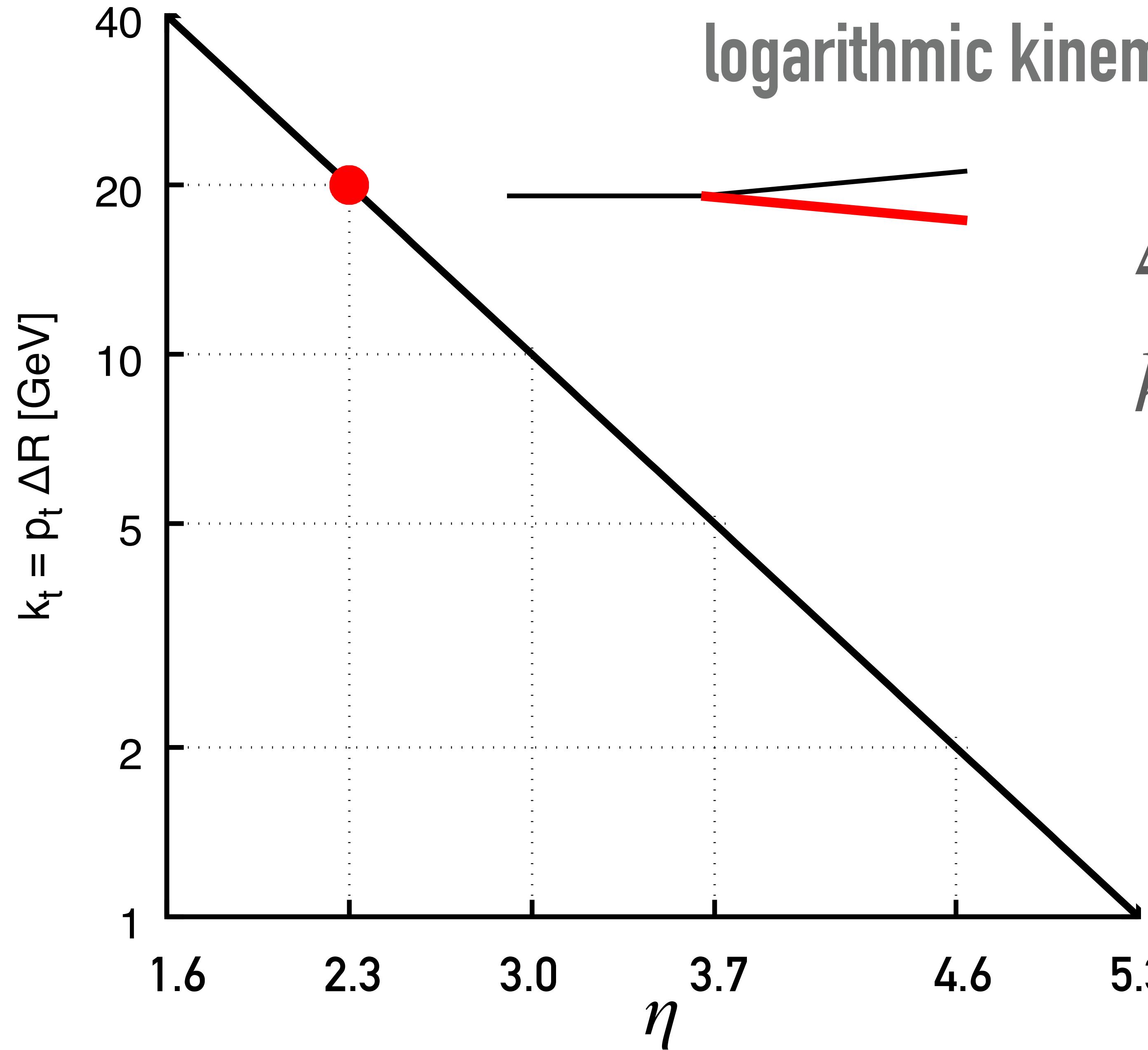
logarithmic kinematic plane whose two variables are

$$\Delta R \text{ (or } \eta = -\ln \tan \frac{\Delta R}{2})$$
$$k_t = p_t \Delta R$$

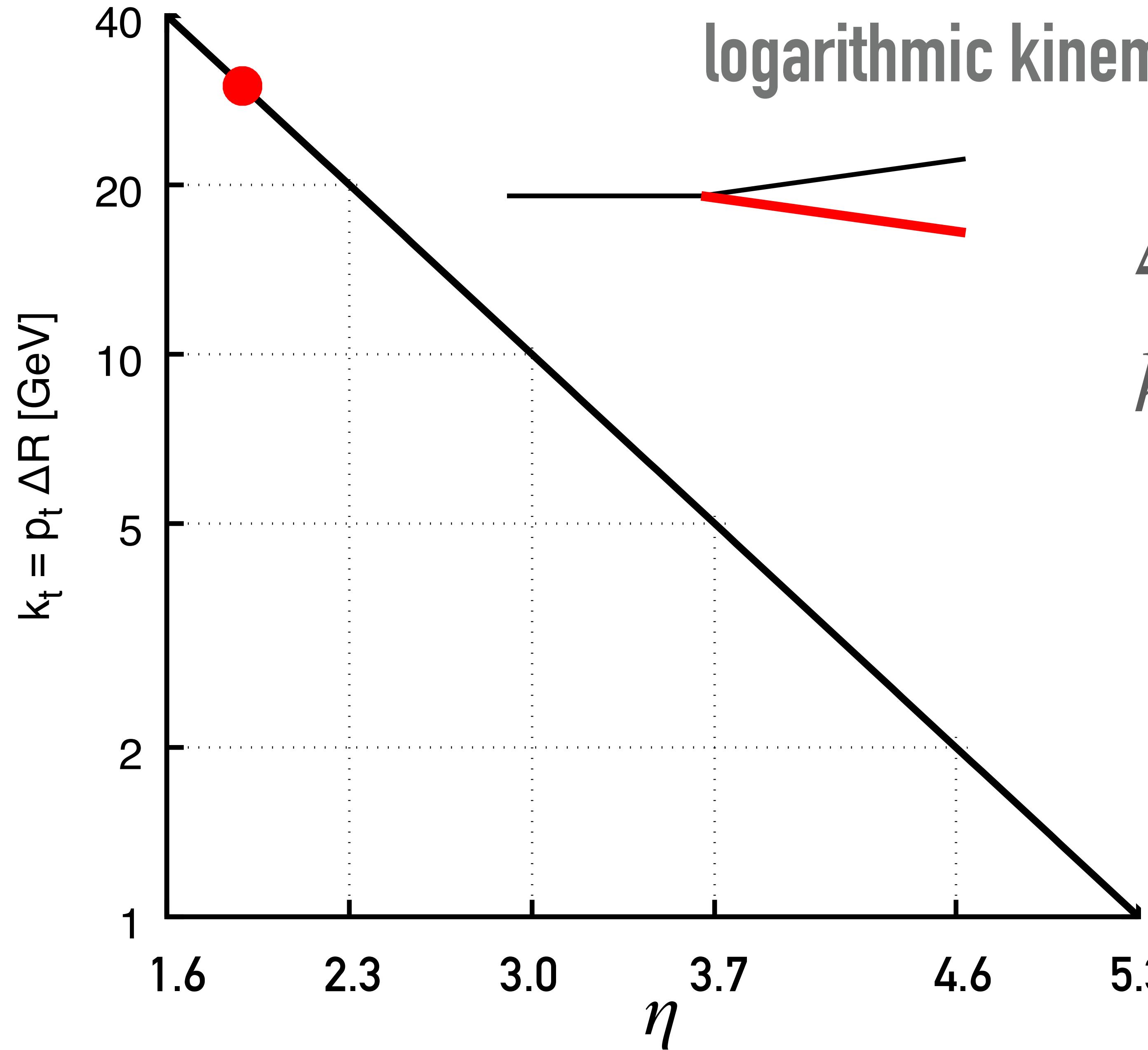
Introduced for understanding Parton Shower Monte Carlos by  
B. Andersson, G. Gustafson, L. Lonnblad and Pettersson 1989

The Lund Plane

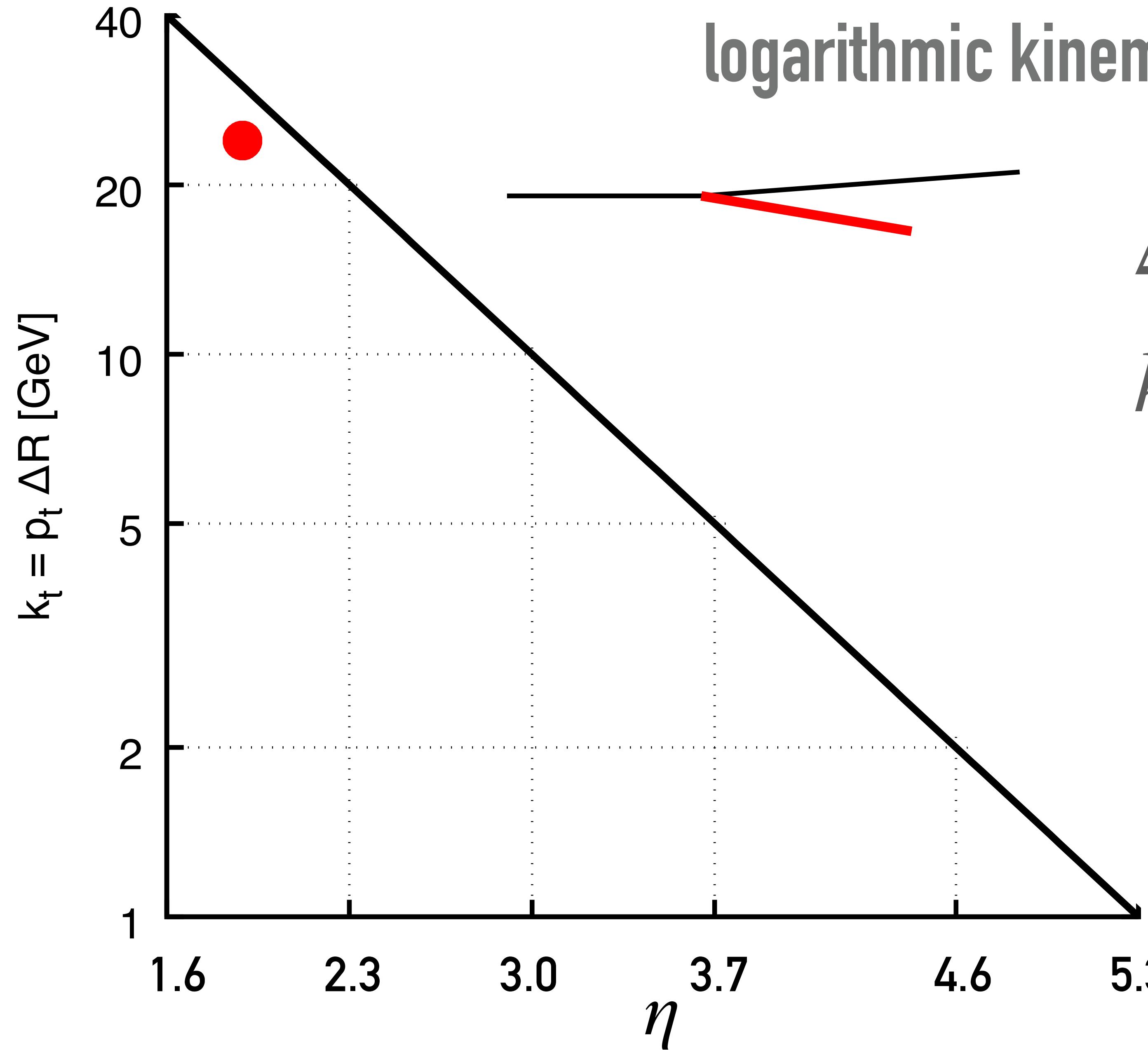




The Lund Plane



The Lund Plane

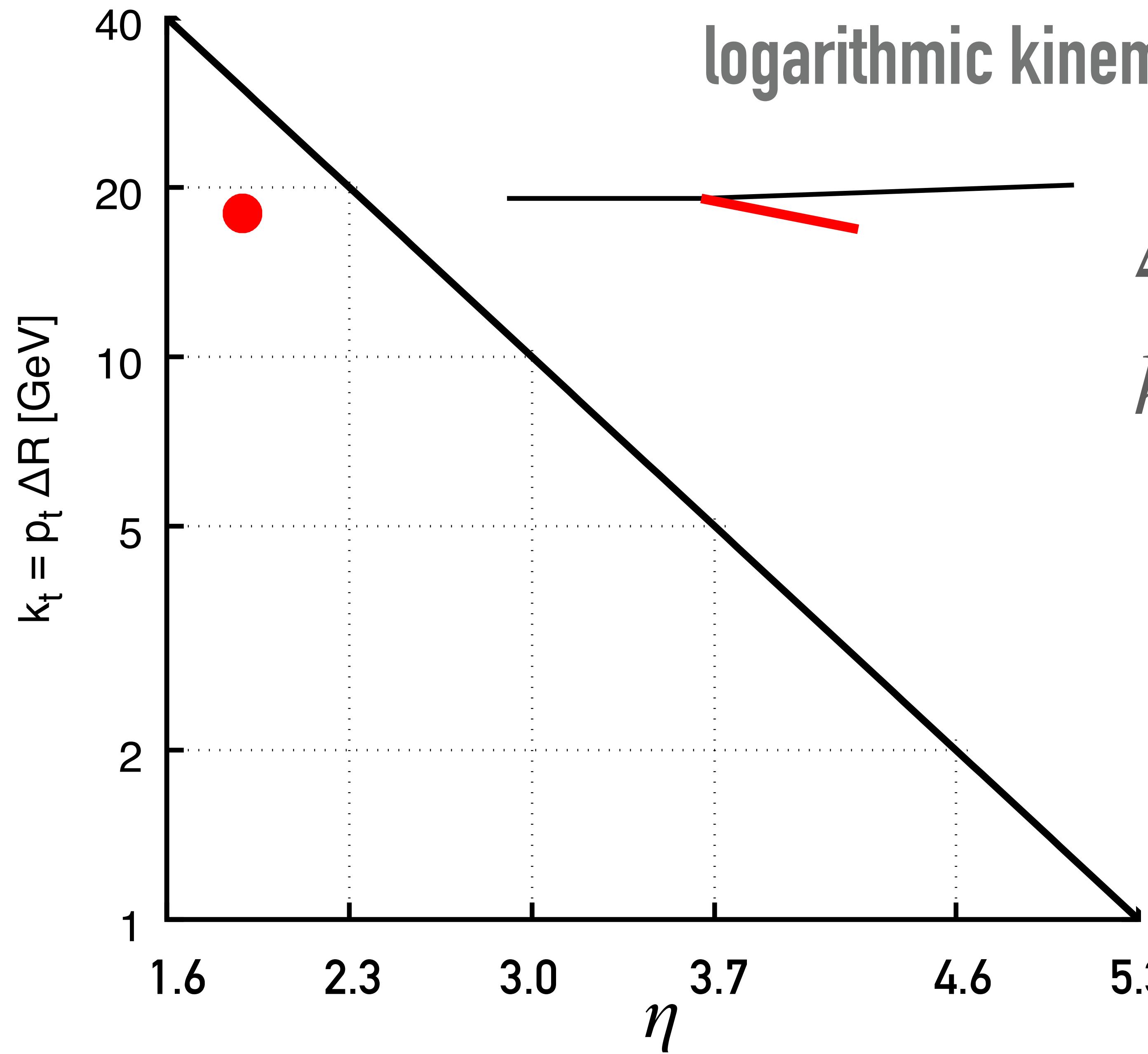


logarithmic kinematic plane whose two variables are

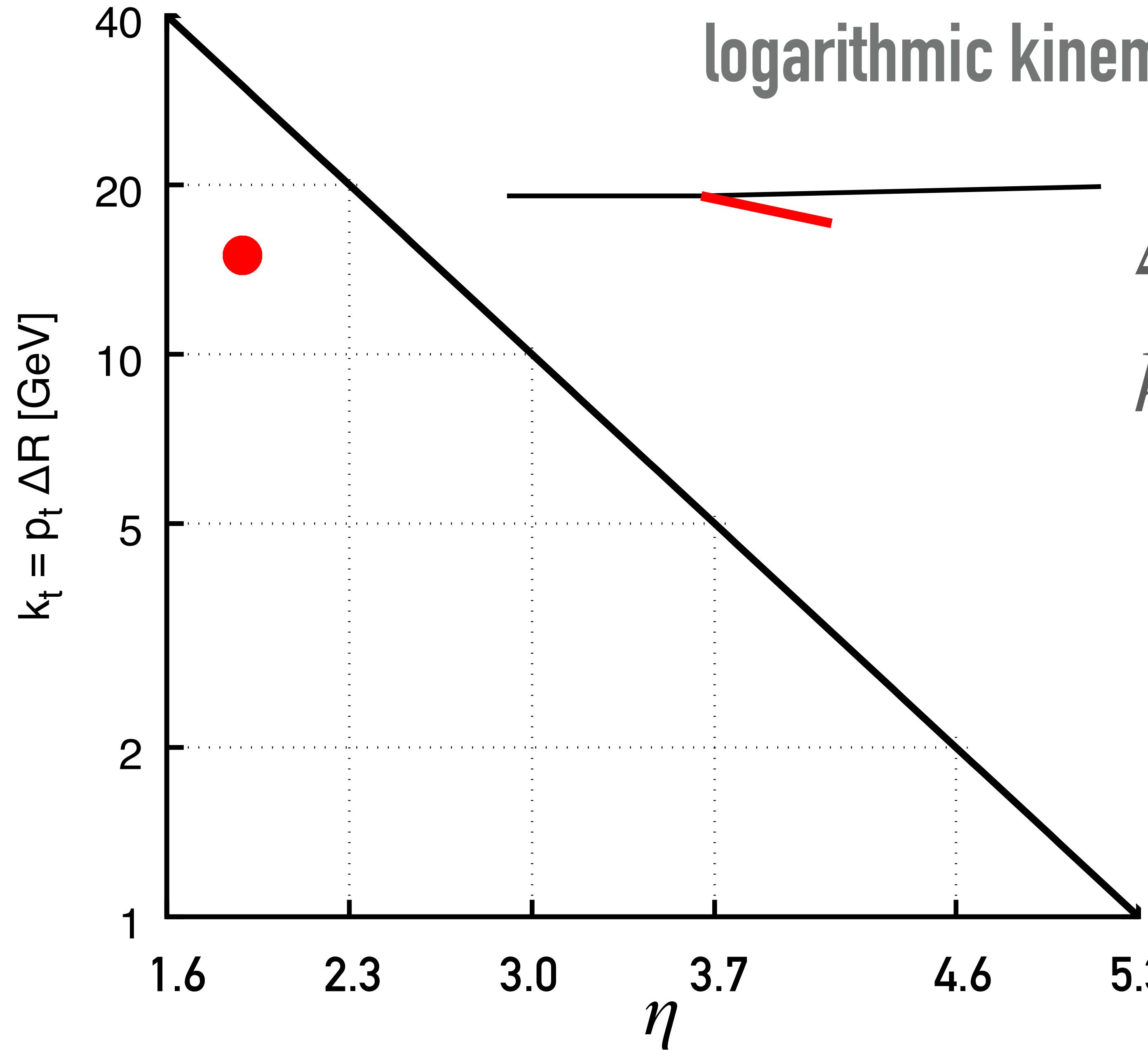
$$\Delta R \text{ (or } \eta = -\ln \tan \frac{\Delta R}{2})$$
$$k_t = p_t \Delta R$$

Introduced for understanding Parton Shower Monte Carlos by  
B. Andersson, G. Gustafson, L. Lonnblad and Pettersson 1989

The Lund Plane

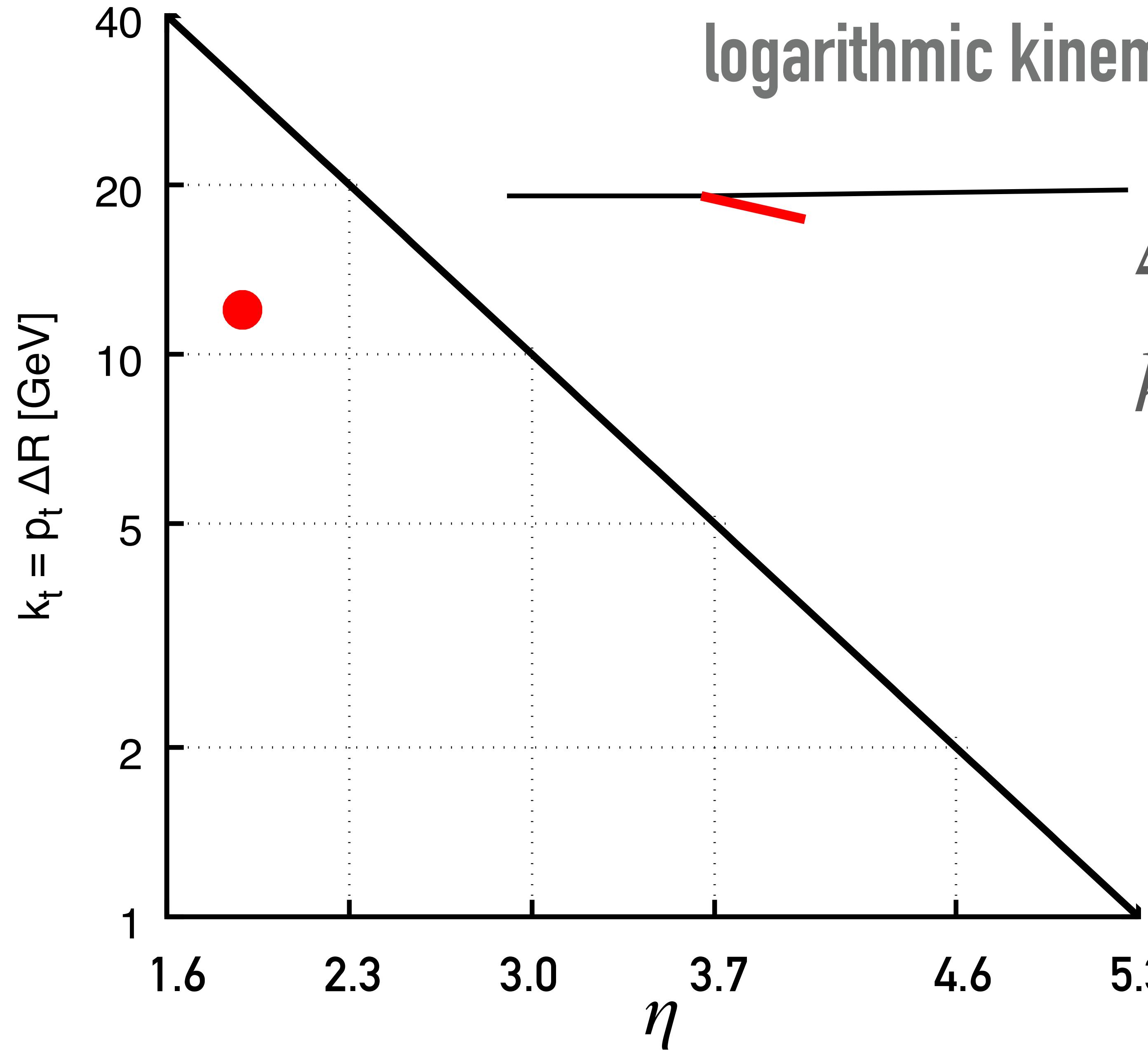


$$\Delta R \text{ (or } \eta = -\ln \tan \frac{\Delta R}{2})$$

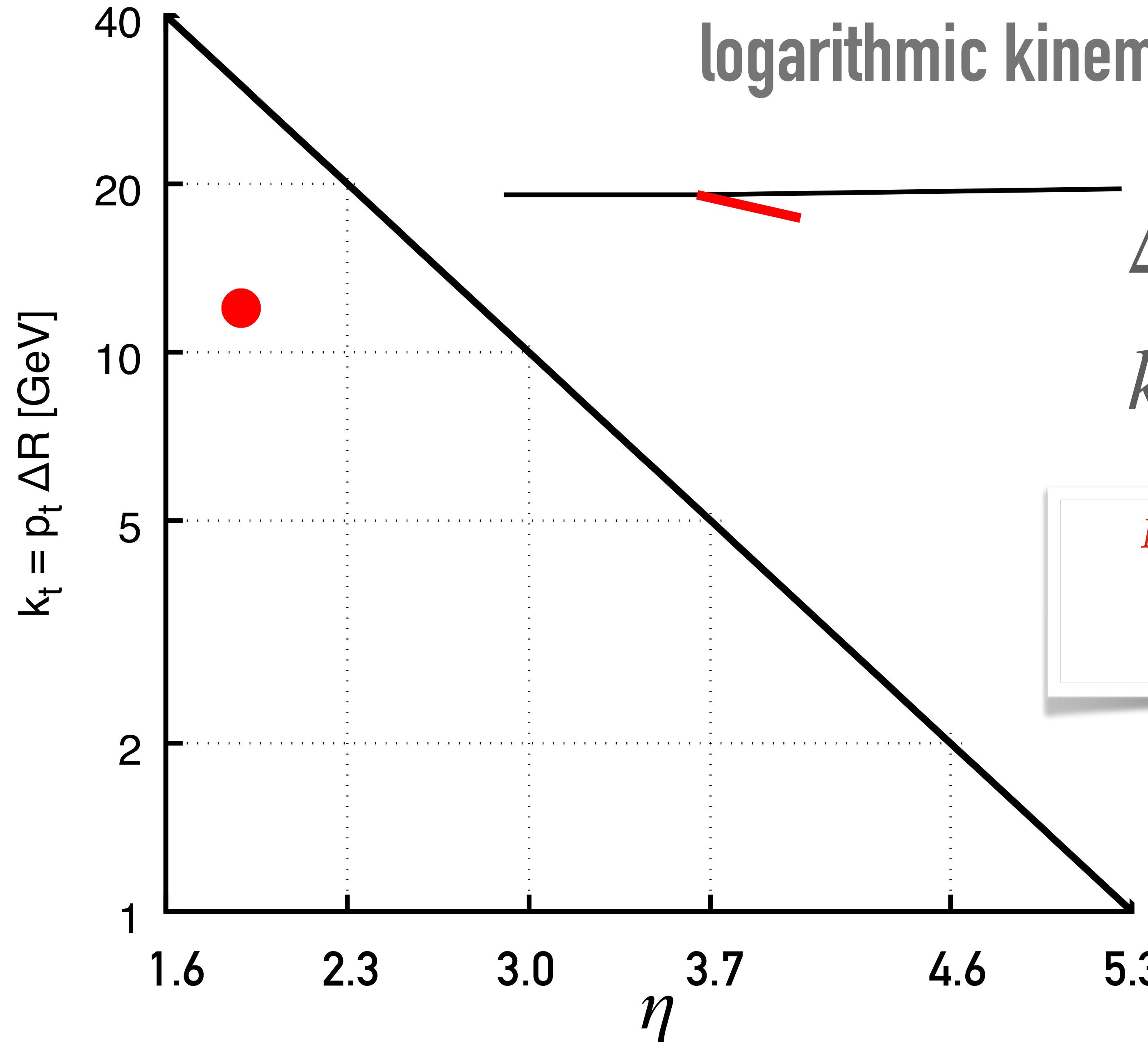


$\Delta R$  (or  $\eta = -\ln \tan \frac{\Delta R}{2}$ )

$k_t = p_t \Delta R$



The Lund Plane



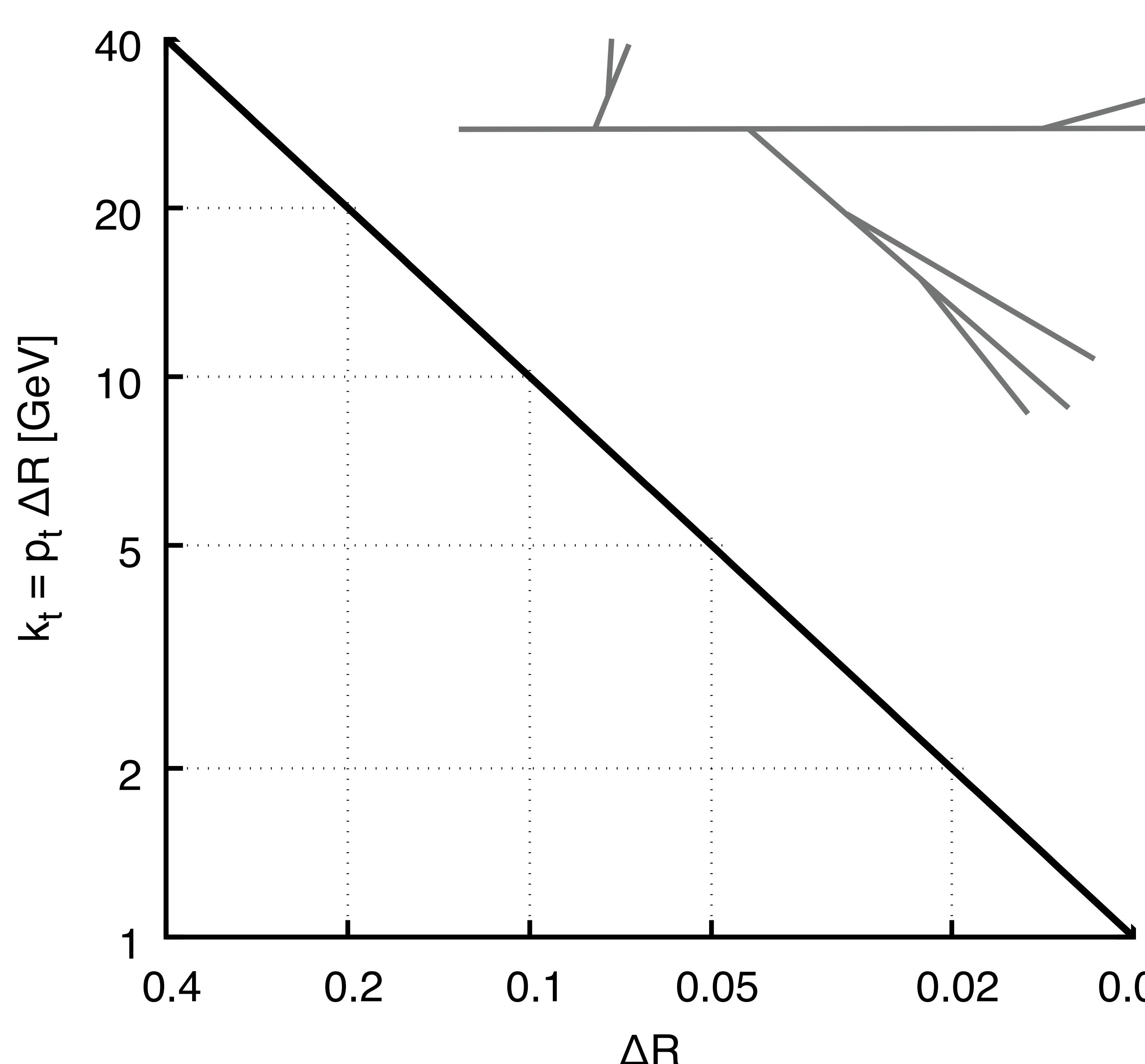
logarithmic kinematic plane whose two variables are

$$\Delta R \text{ (or } \eta = -\ln \tan \frac{\Delta R}{2})$$
$$k_t = p_t \Delta R$$

NB: Lund plane can be constructed event-by-event  
using Cambridge/Aachen jet clustering sequence,  
cf. Dreyer, GPS & Soyez '18

Introduced for understanding Parton Shower Monte Carlos by  
B. Andersson, G. Gustafson, L. Lonnblad and Pettersson 1989

The Lund Plane

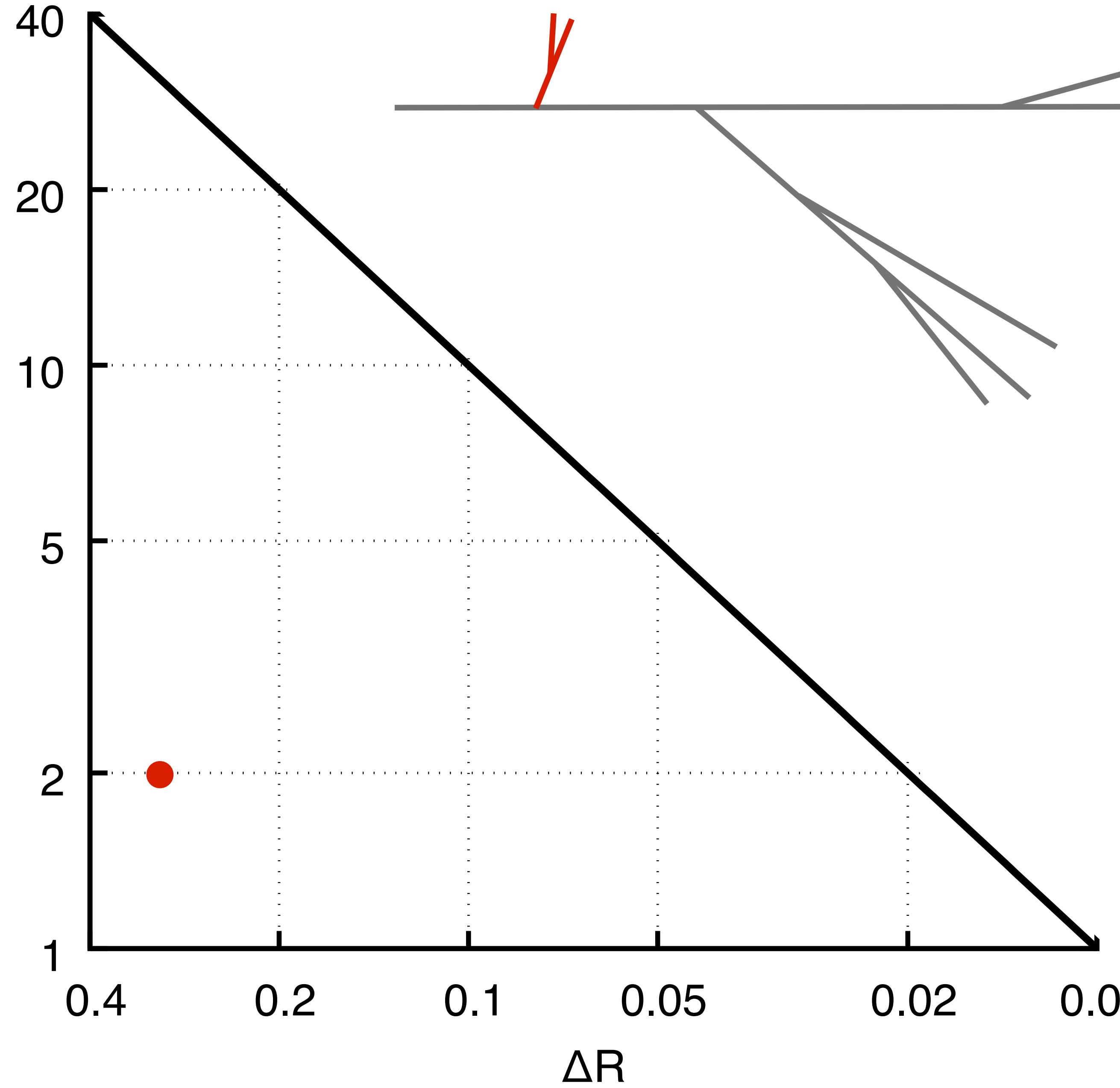


decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane  
repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane

$k_t = p_t \Delta R$  [GeV]



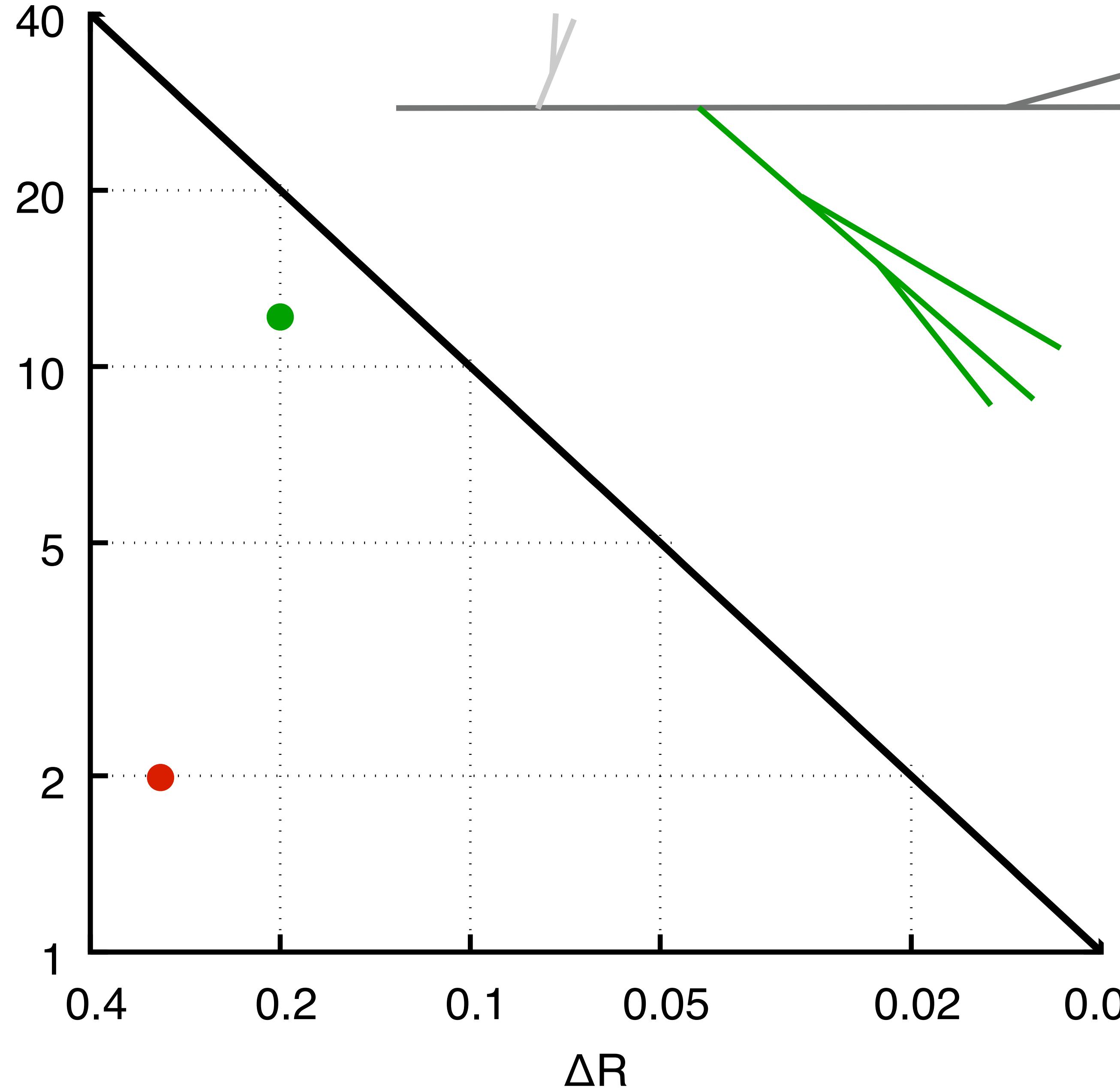
decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane

$k_t = p_t \Delta R$  [GeV]



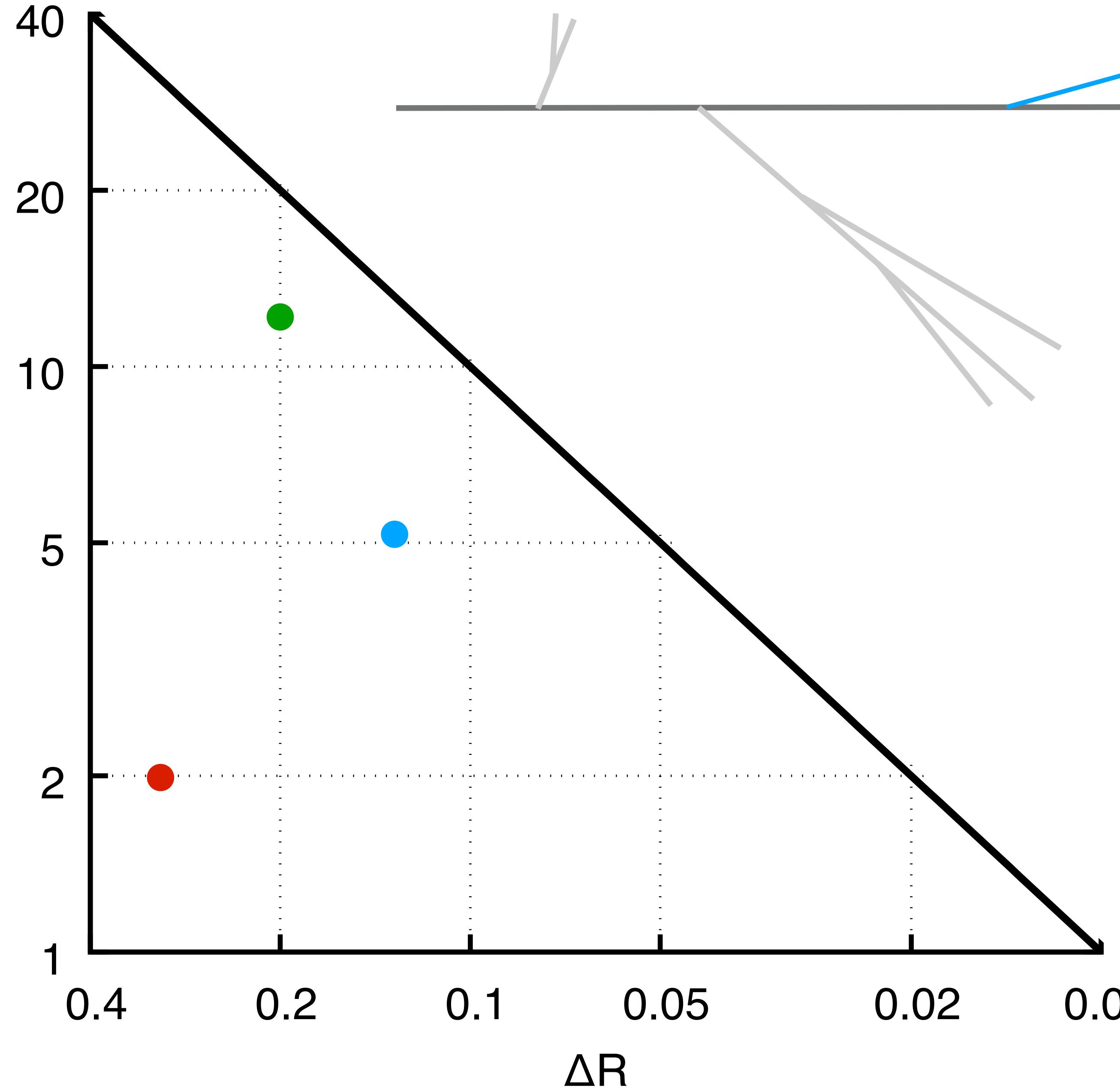
constructing the Lund plane

decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

$k_t = p_t \Delta R$  [GeV]

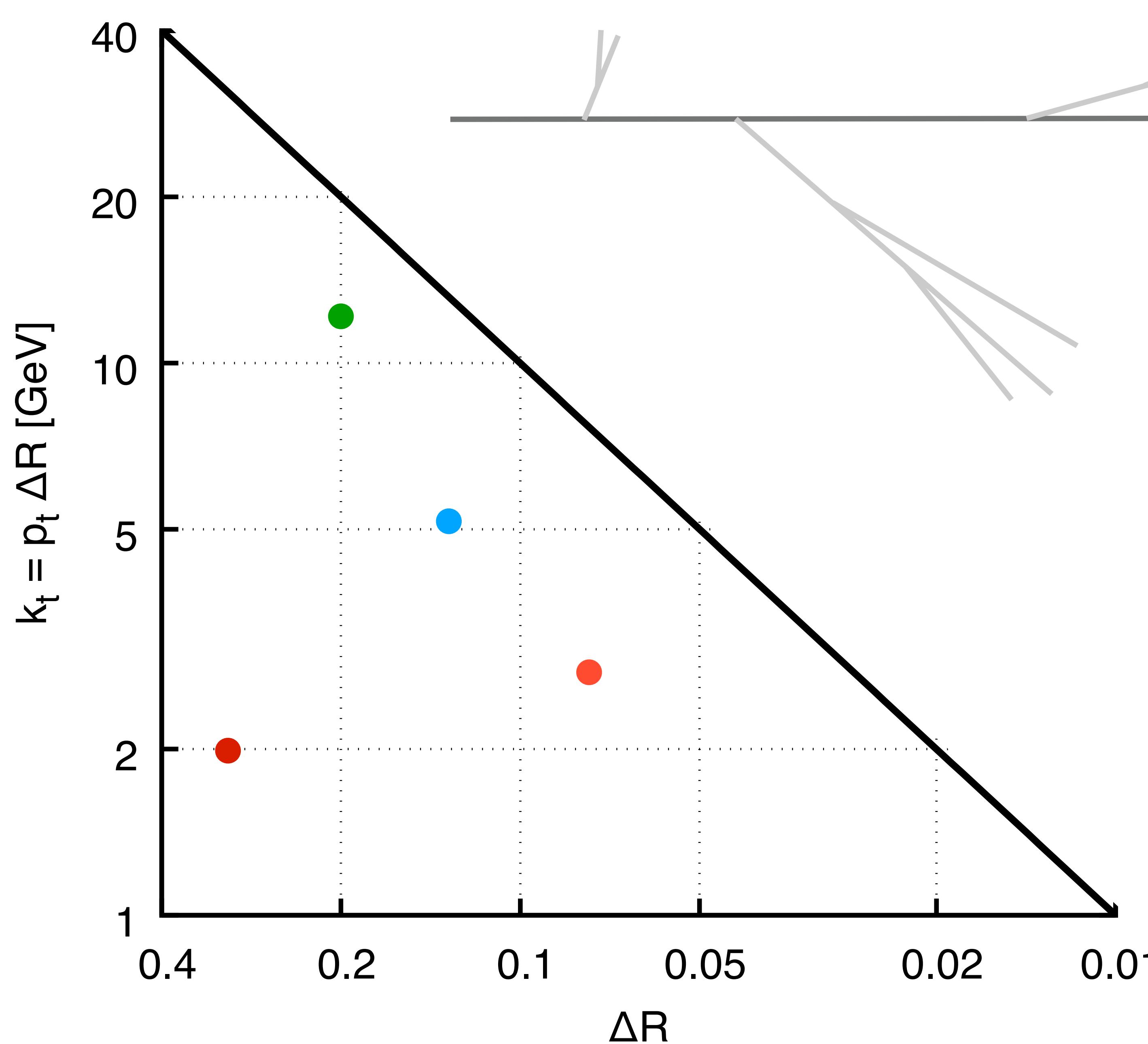


decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane



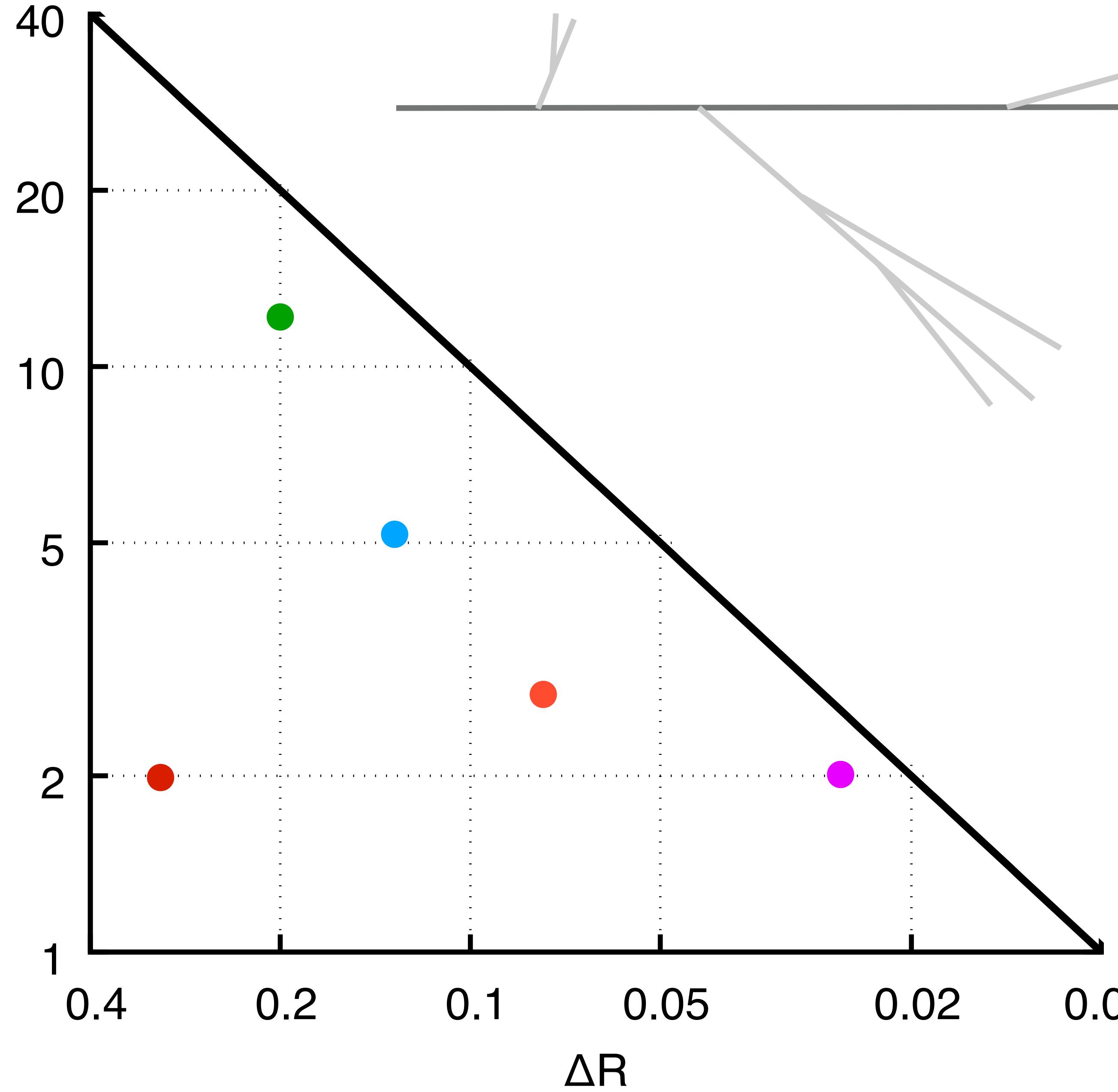
decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane

$k_t = p_t \Delta R$  [GeV]



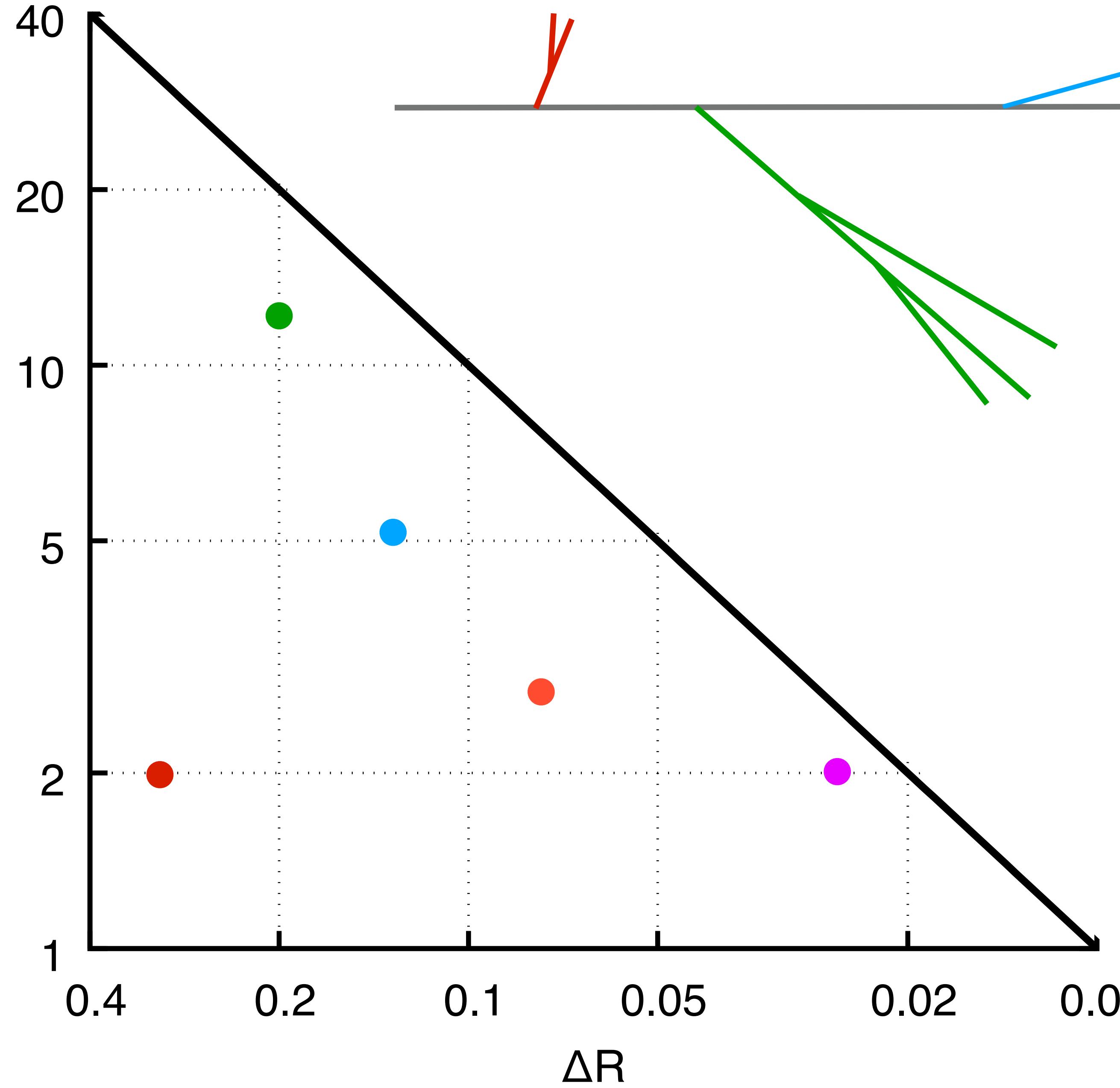
decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane

$k_t = p_t \Delta R$  [GeV]



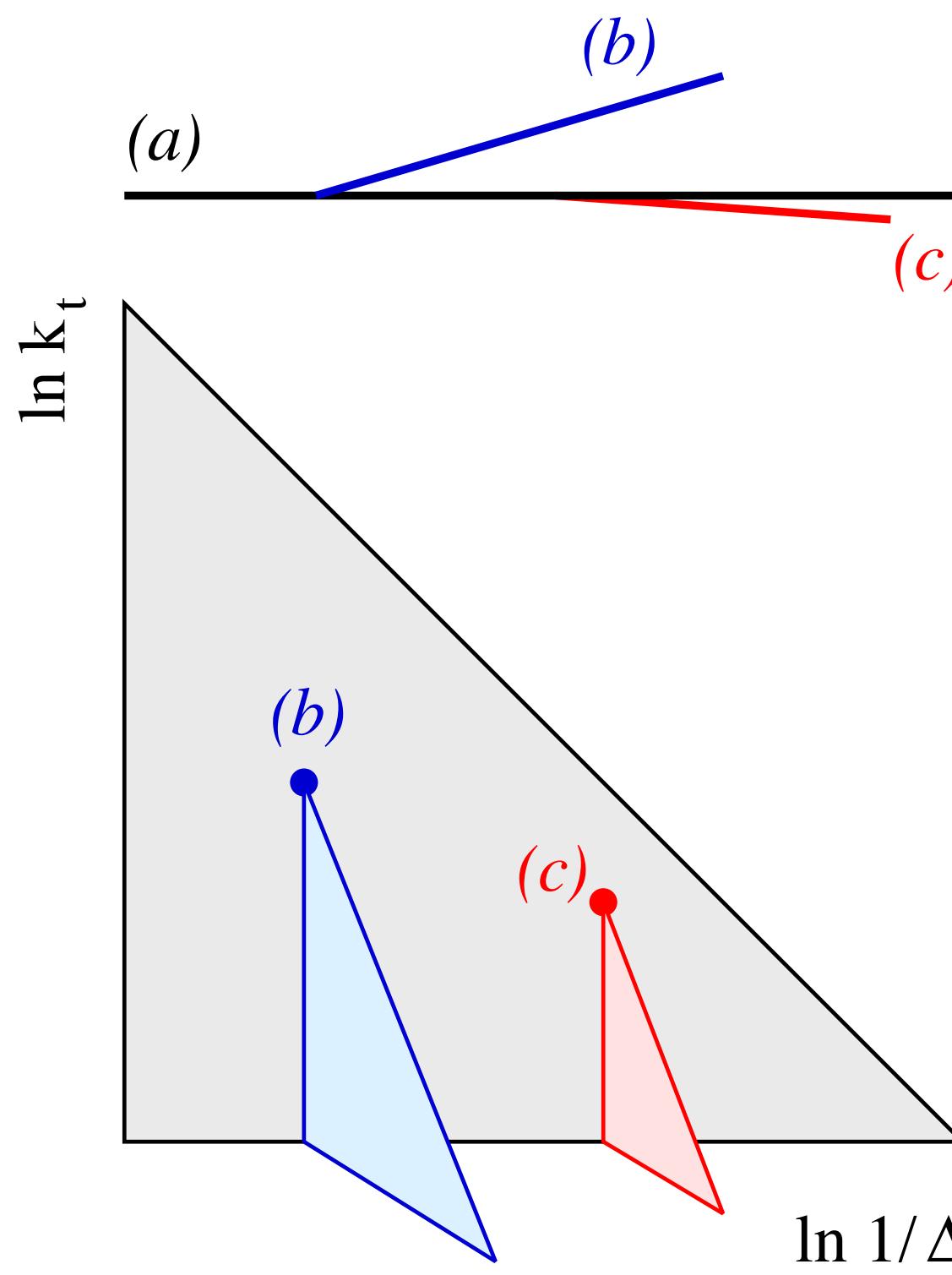
decluster a C/A jet:  
at each step record  $\Delta R, k_t$   
as a point in the Lund plane

repeatedly follow harder branch

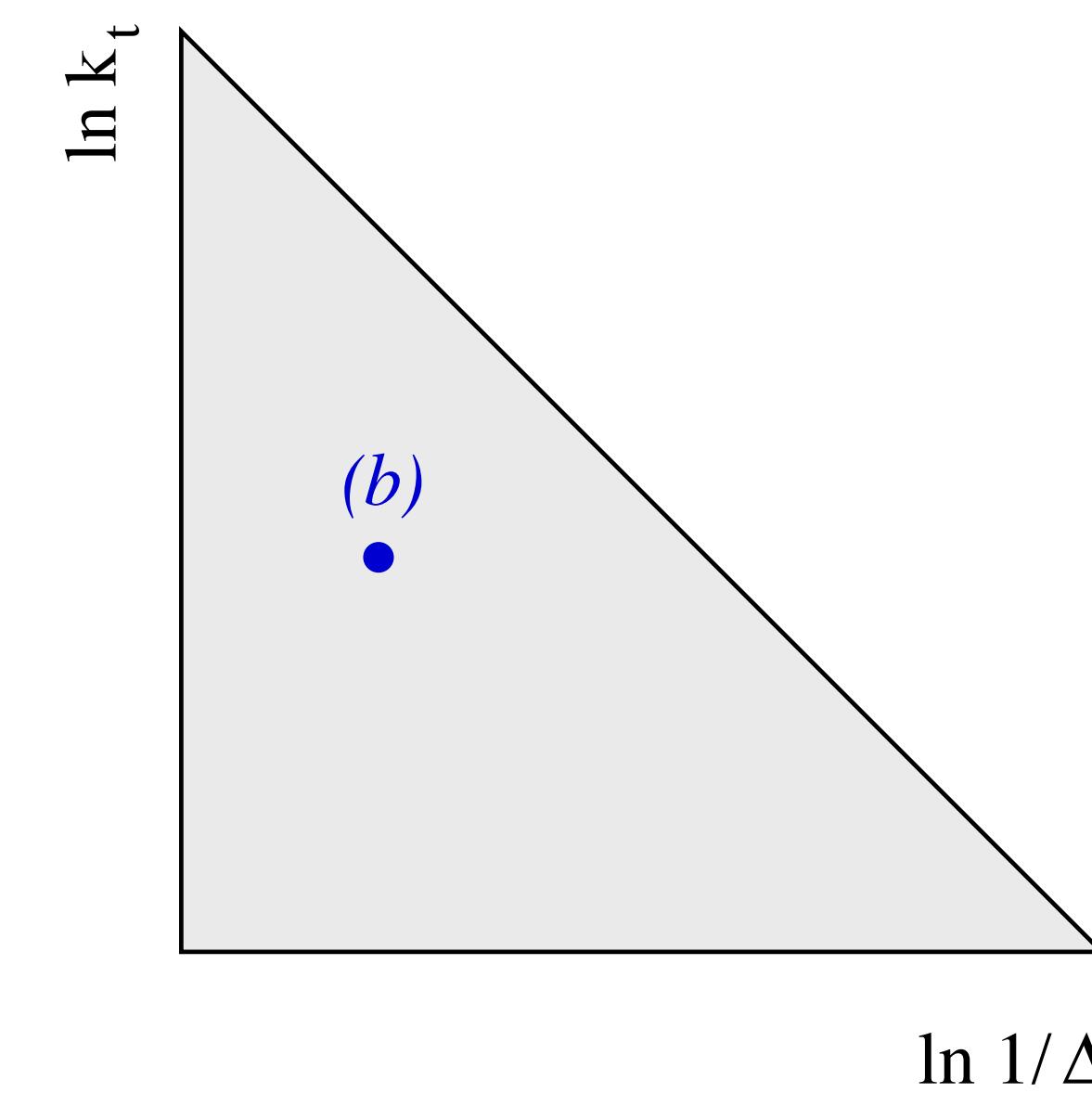
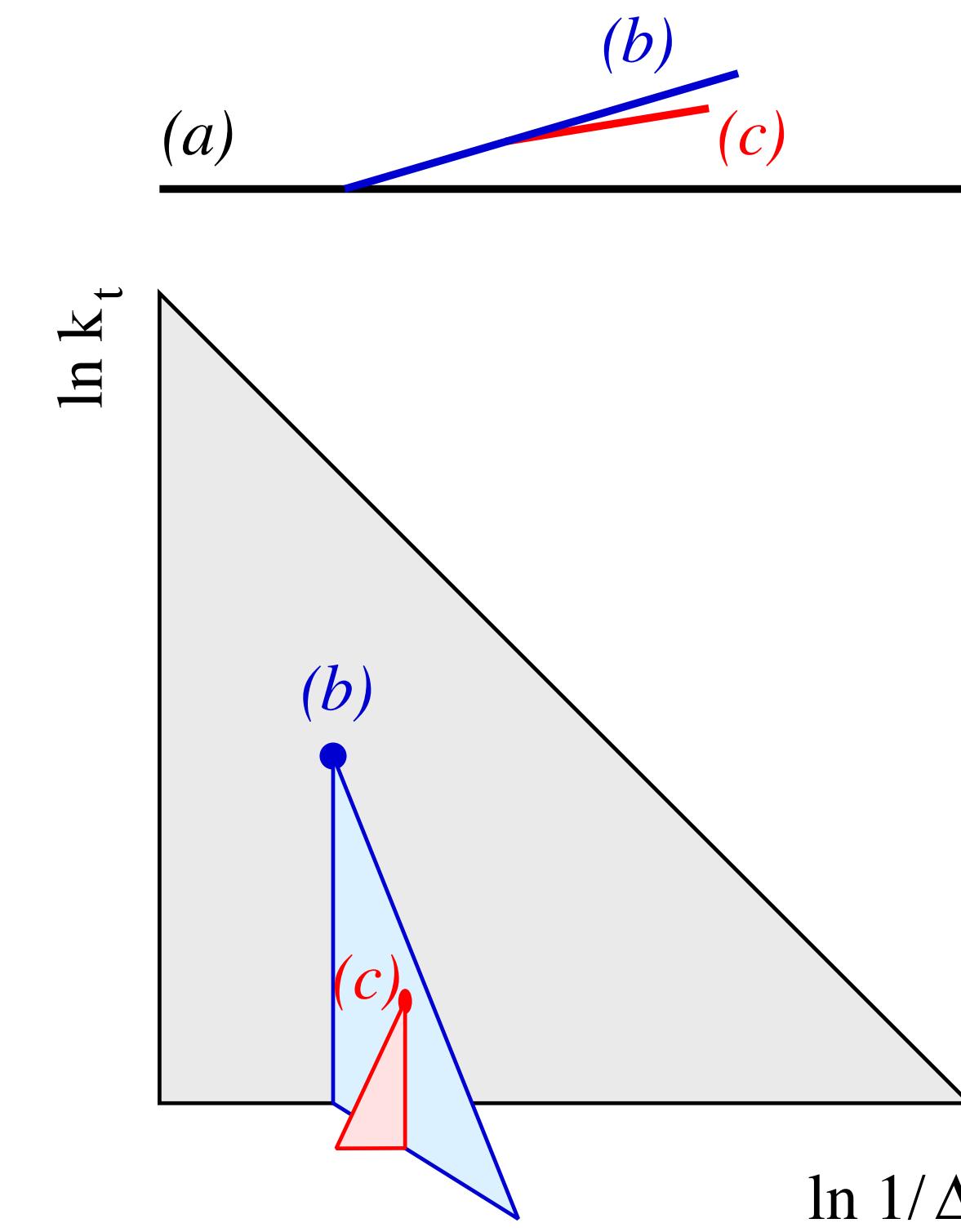
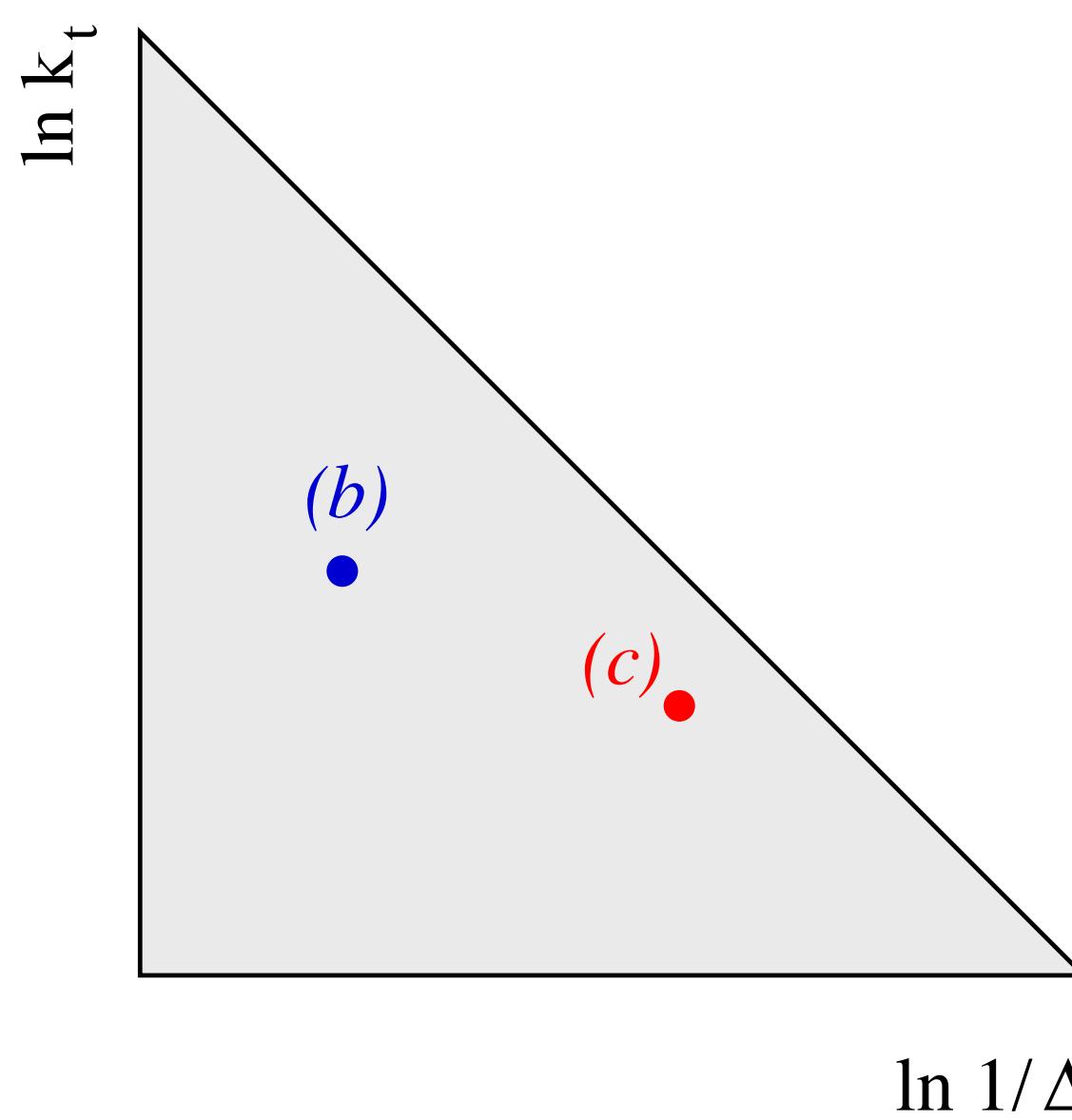
5th heavy-ion workshop @ CERN, [1808.03689](#)  
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)

constructing the Lund plane

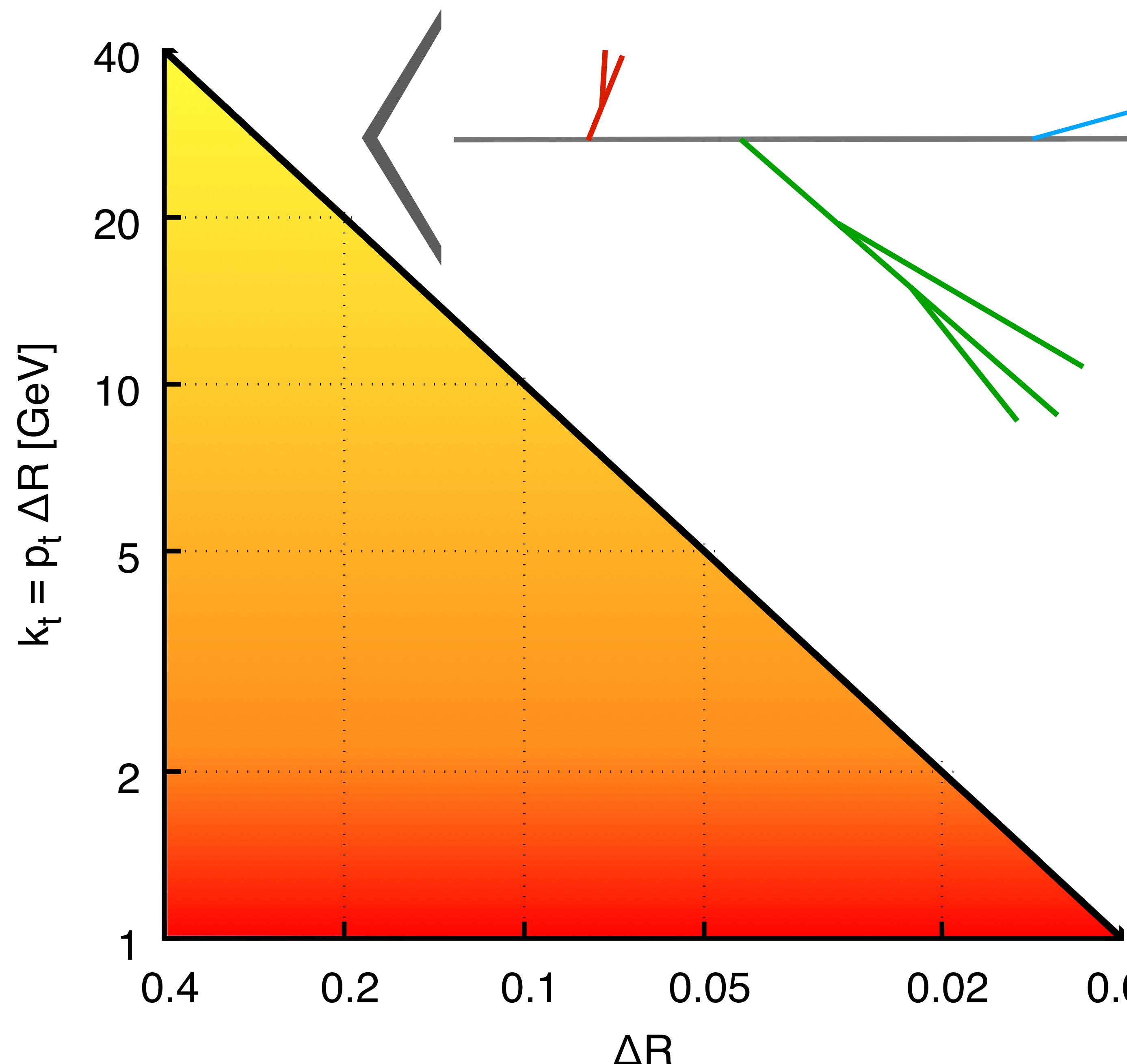
LUND DIAGRAM



PRIMARY LUND PLANE



jet with  $R = 0.4$ ,  $p_t = 200 \text{ GeV}$

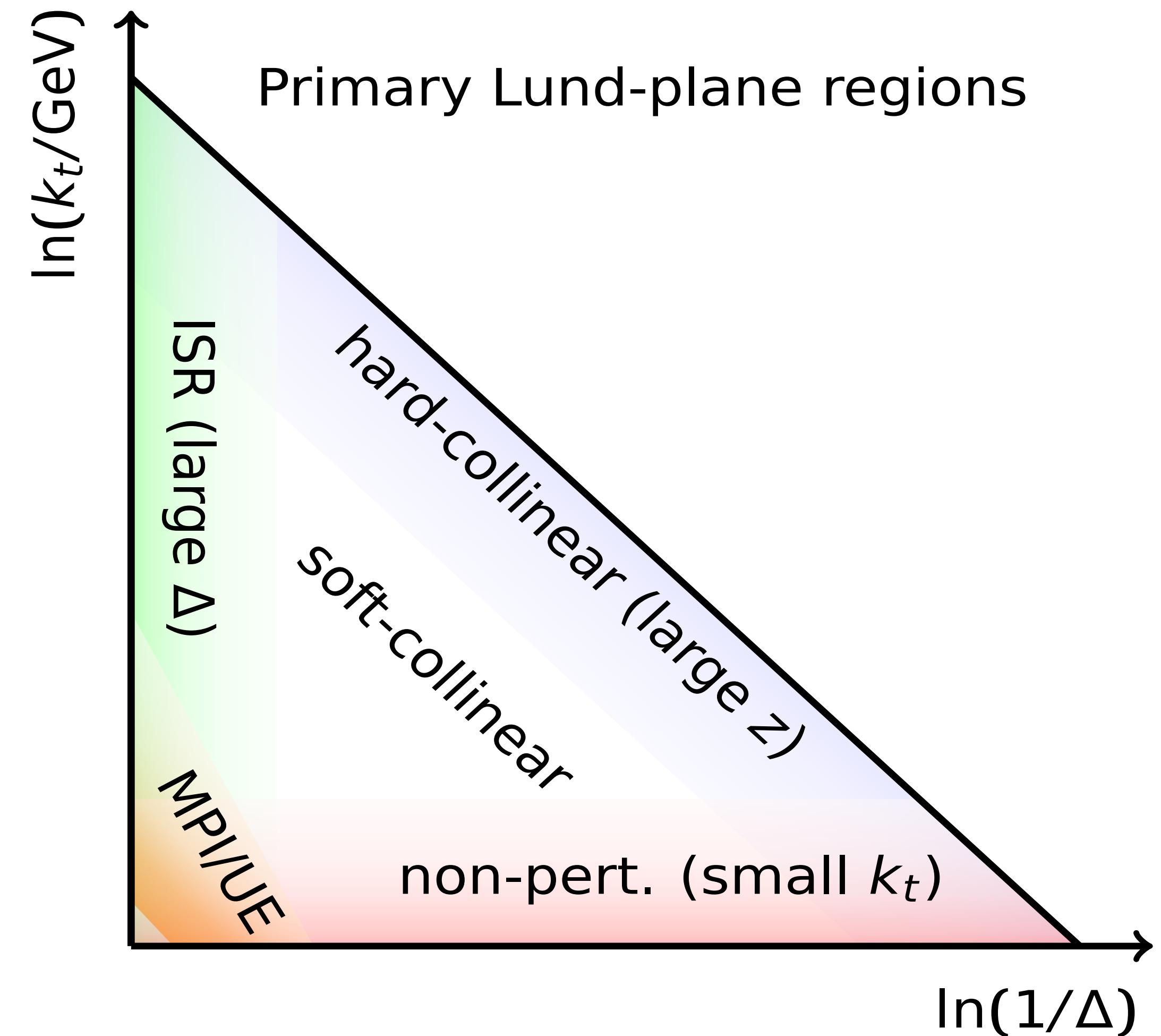
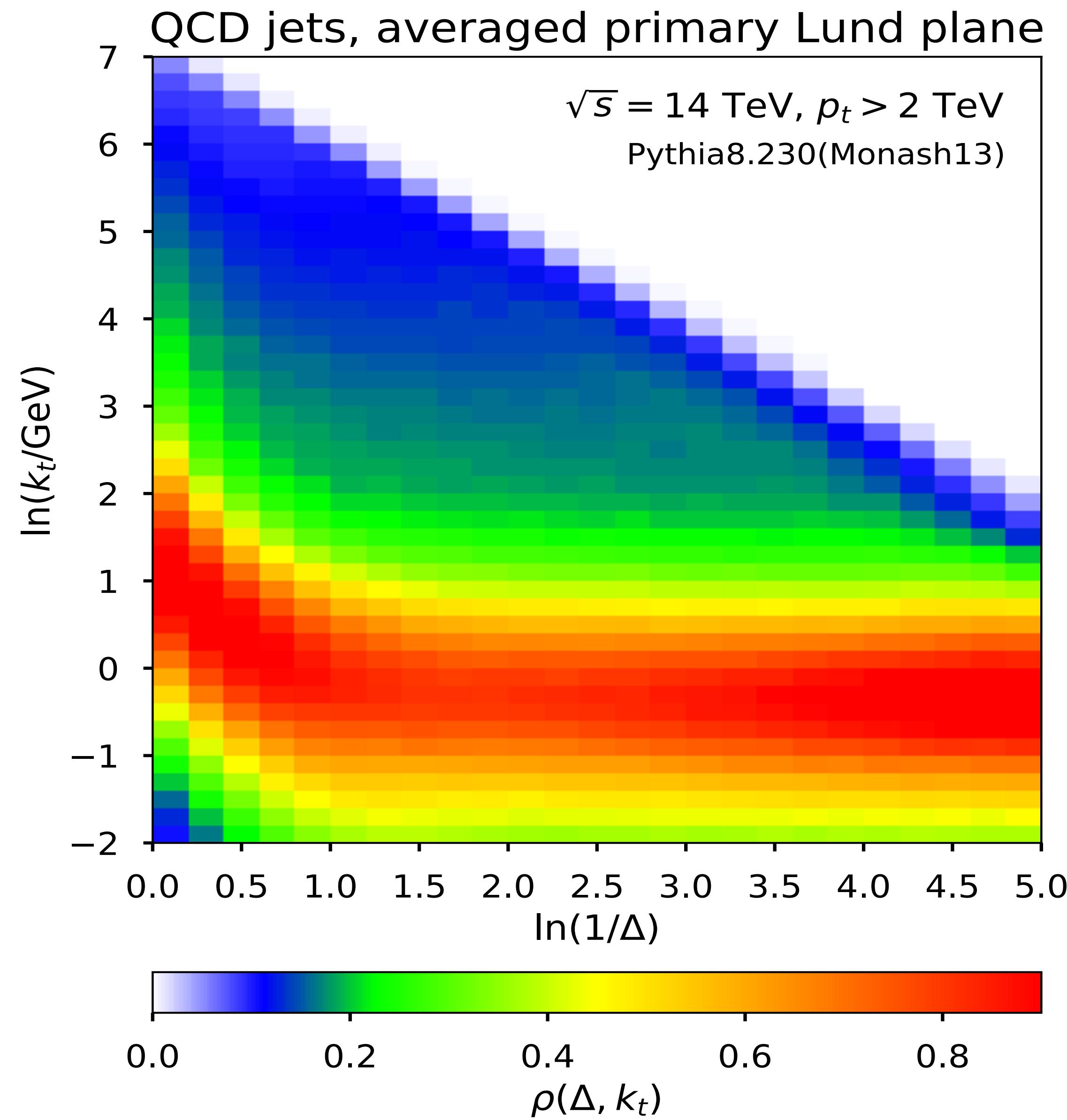


average over many jets:  
**Lund plane density**

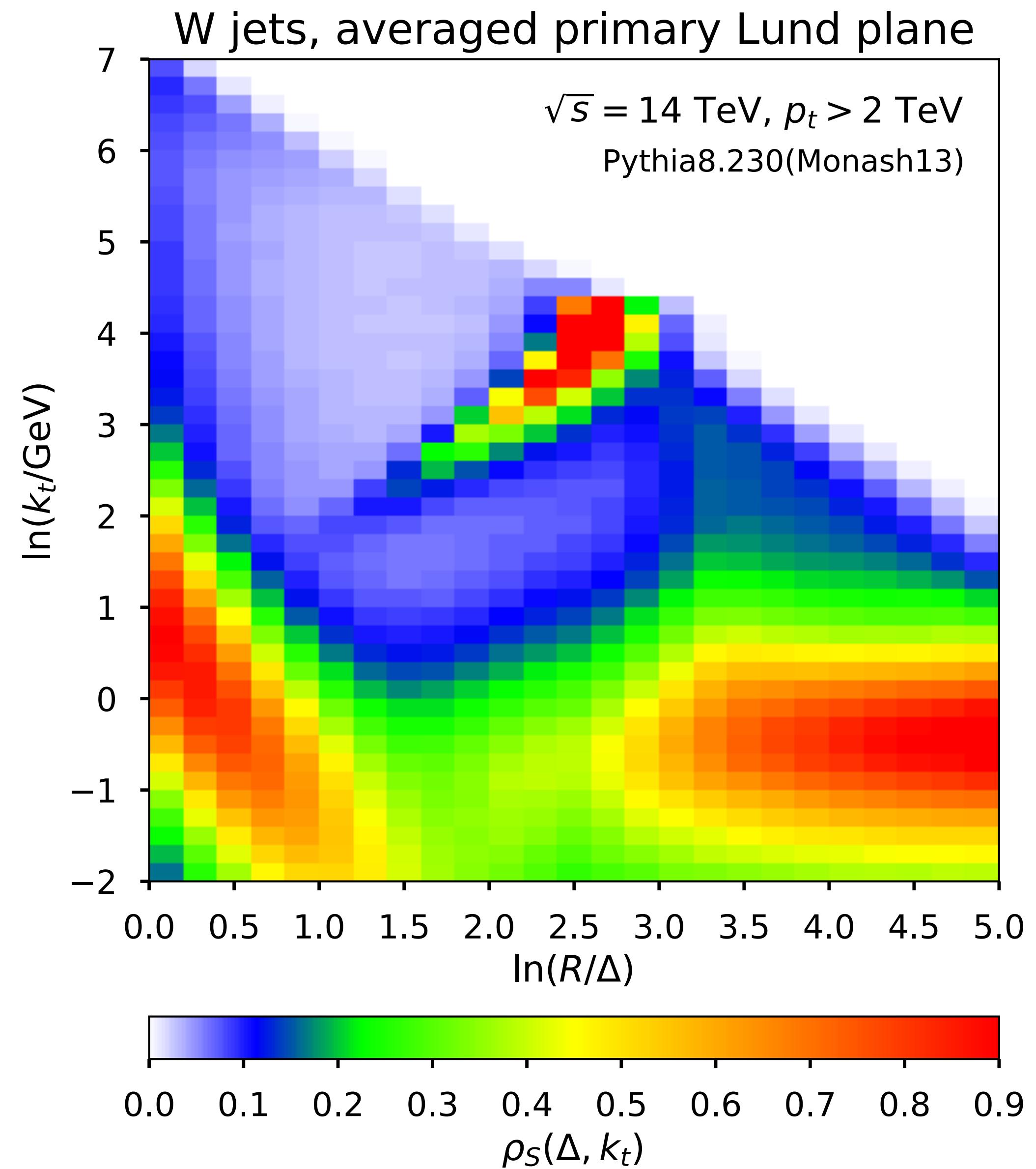
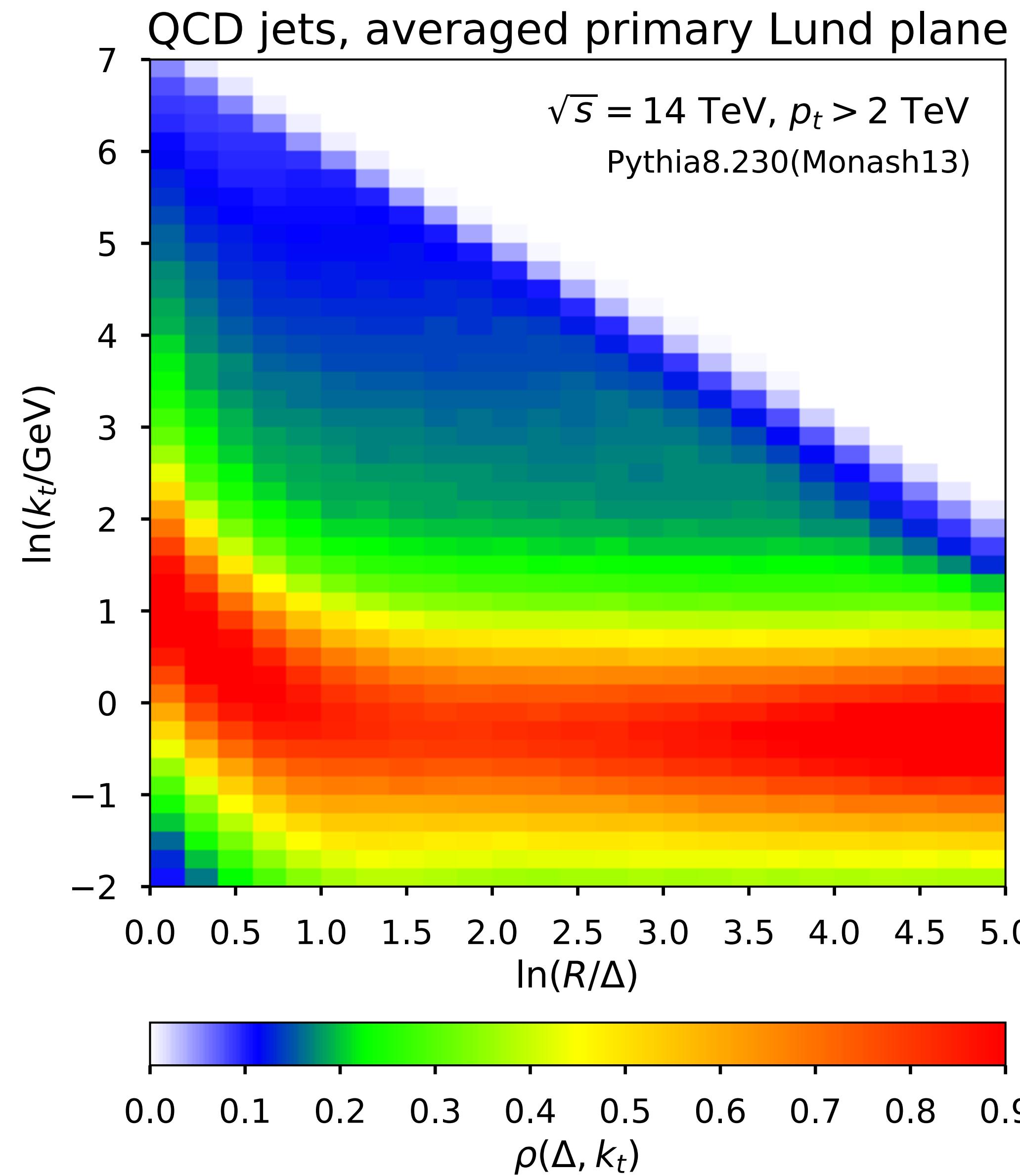
Dreyer, Soyez & GPS, [1807.04758](#) (for pp applications)  
& 5th heavy-ion workshop @ CERN, [1808.03689](#)

constructing the Lund plane

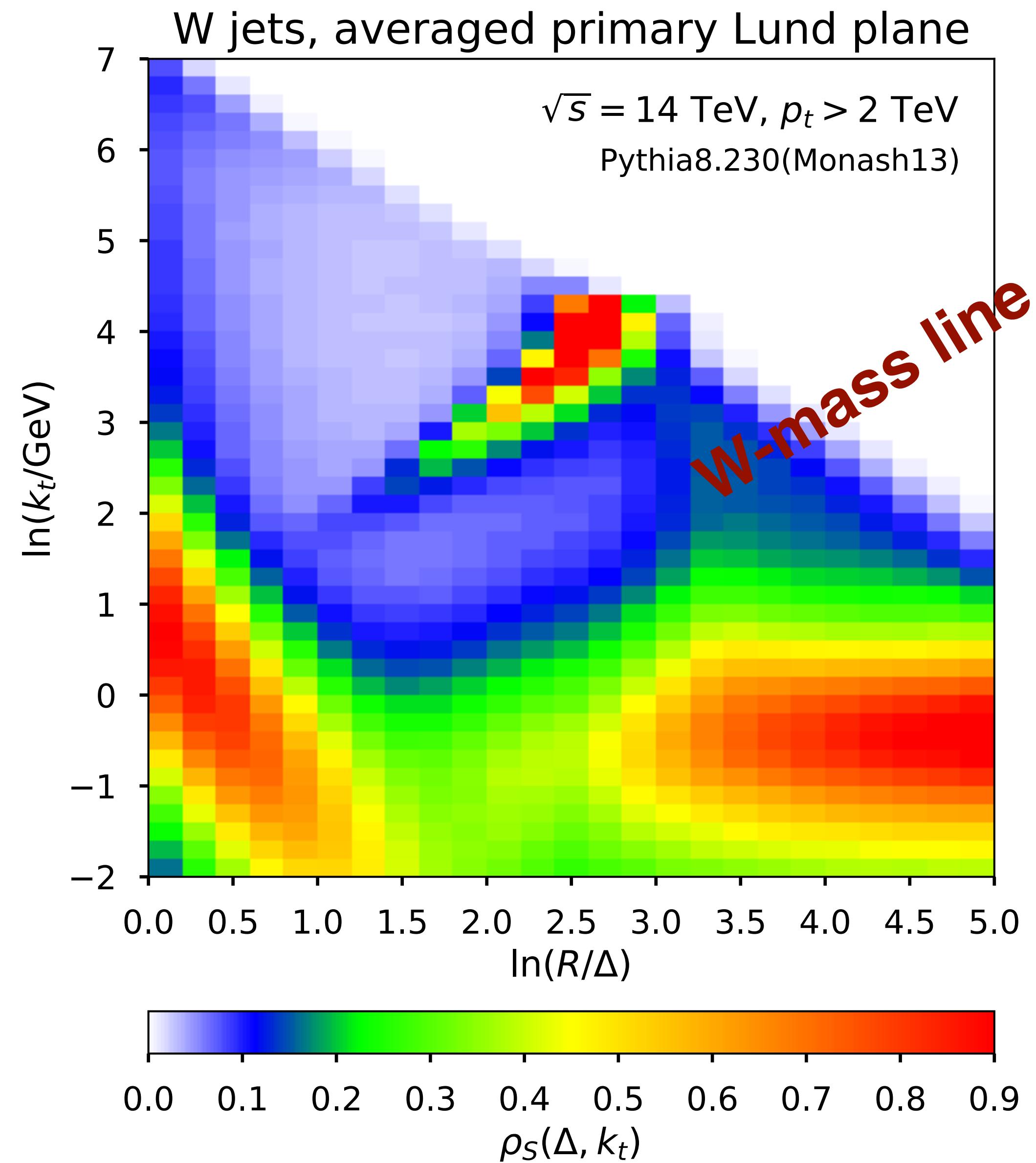
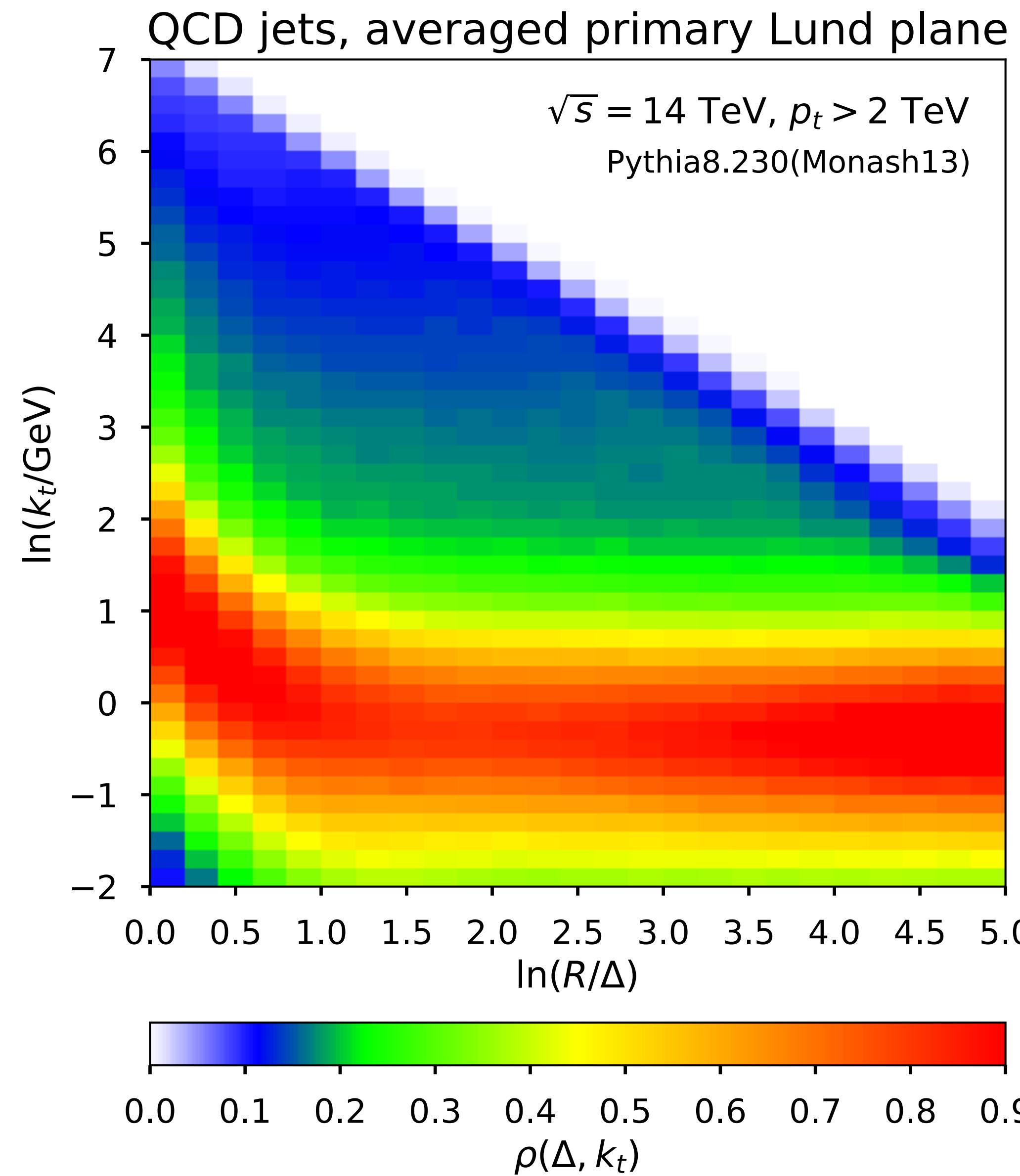
# Regions of the Lund plane



# intrajet energy flow for QCD jets & W jets

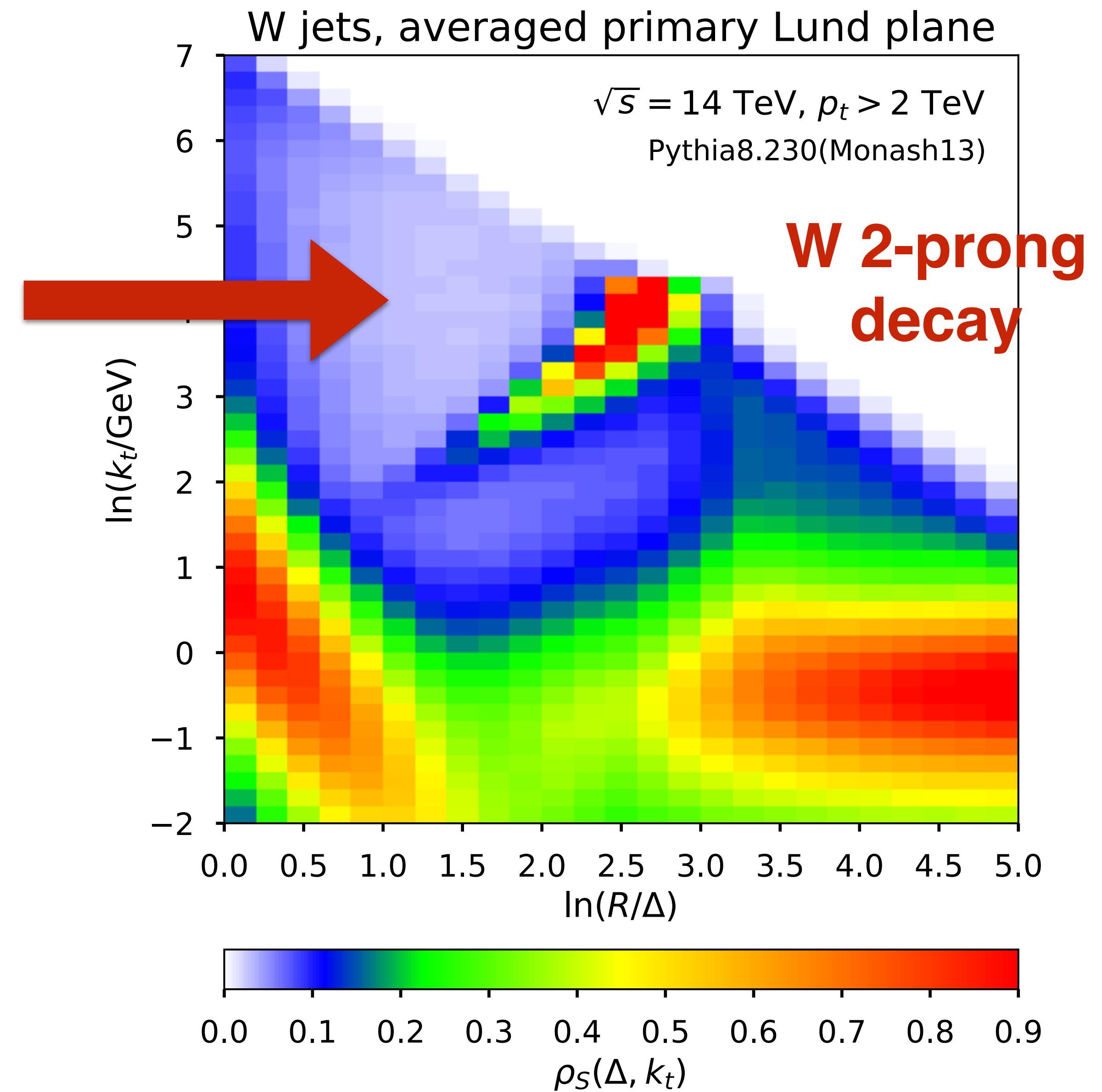


# intrajet energy flow for QCD jets & W jets

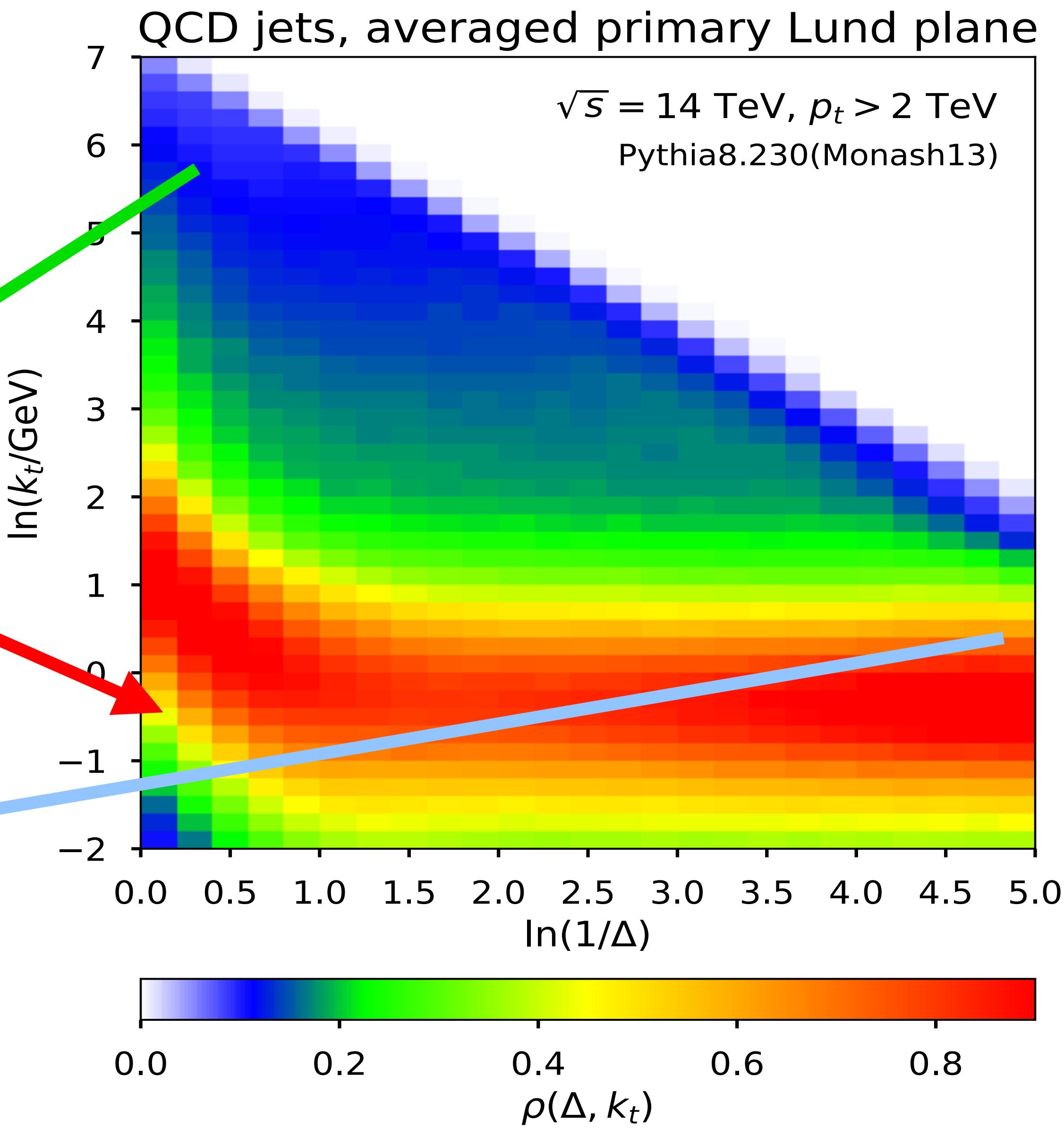
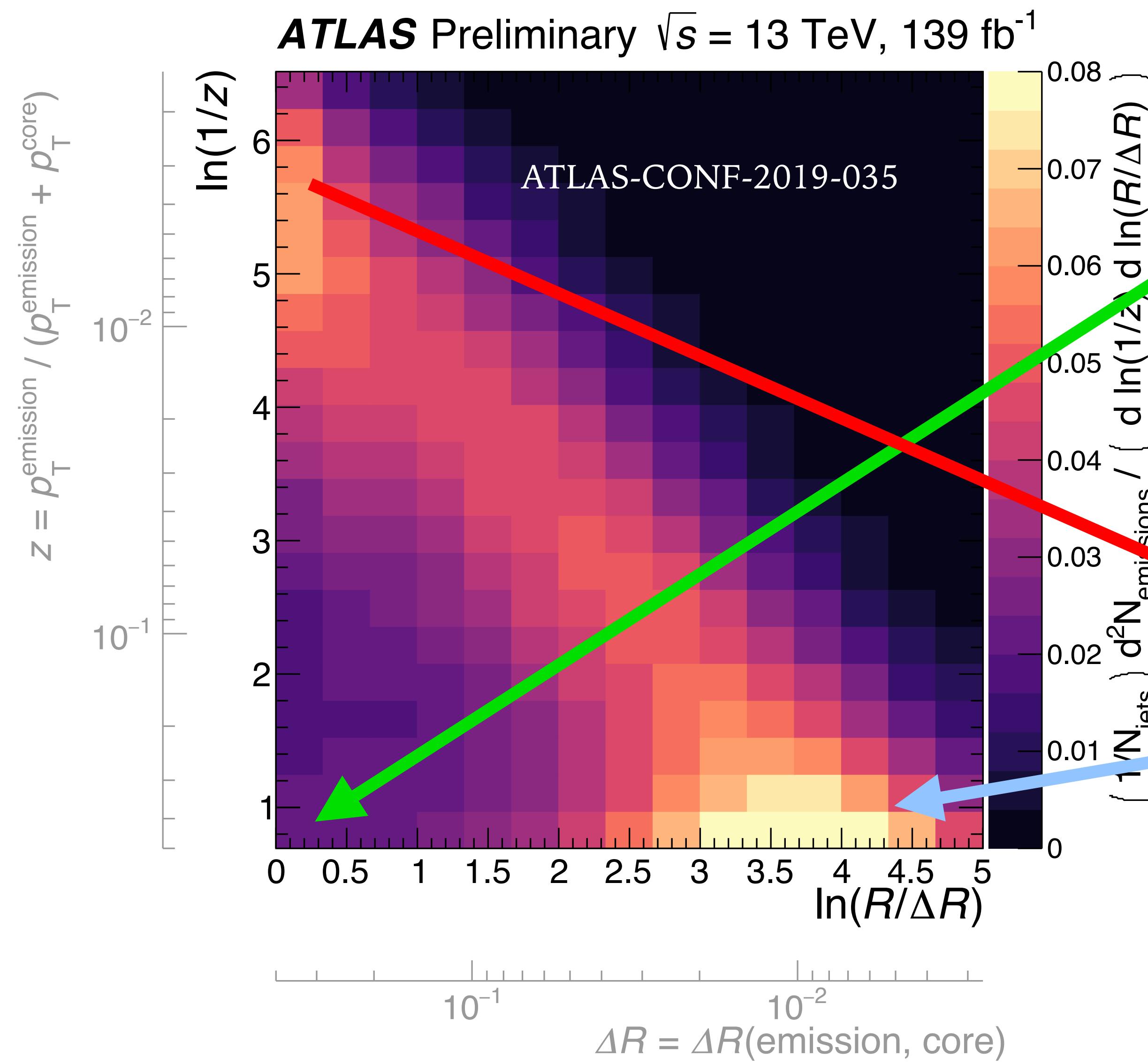


# intrajet energy flow for QCD jets & W jets

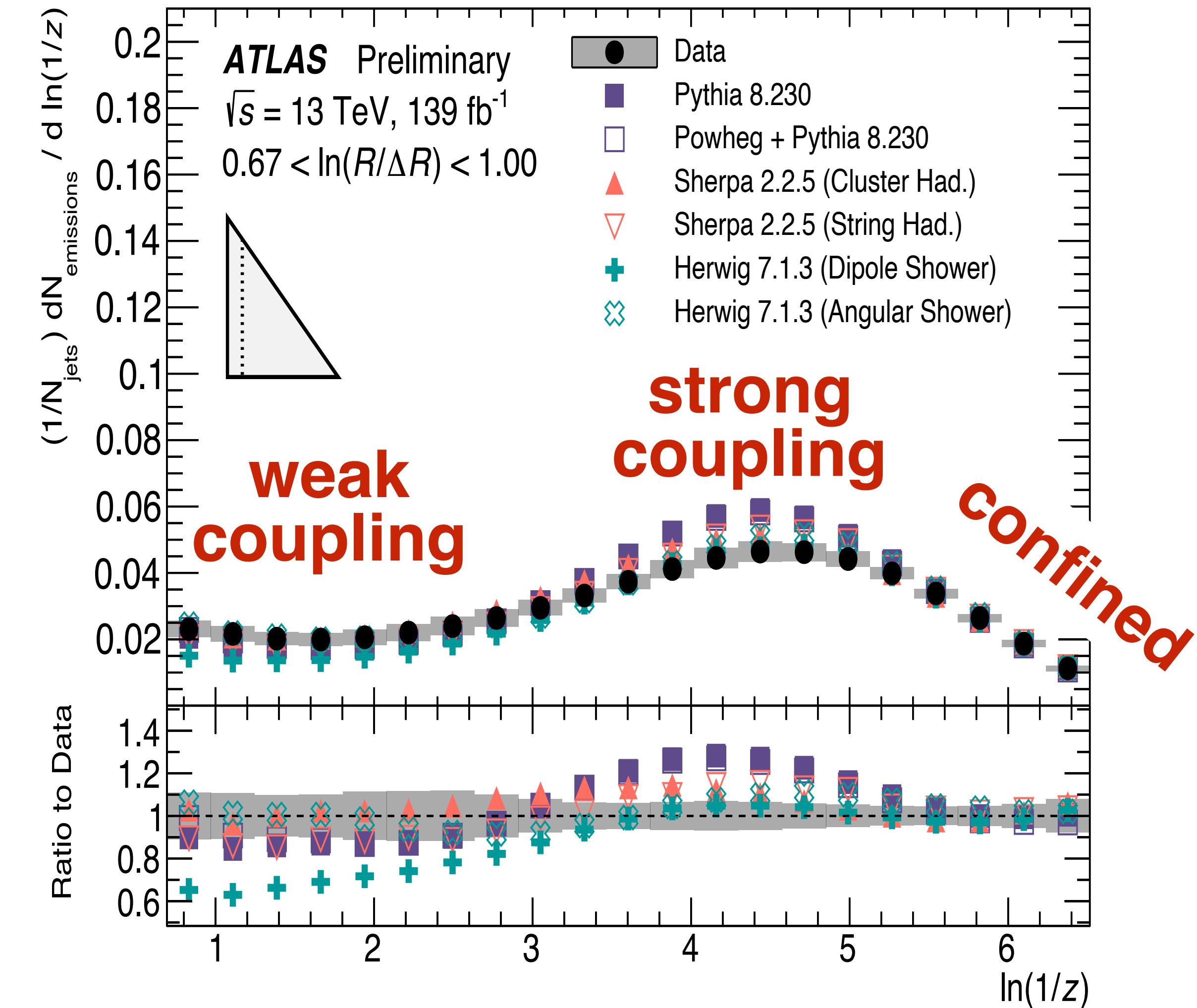
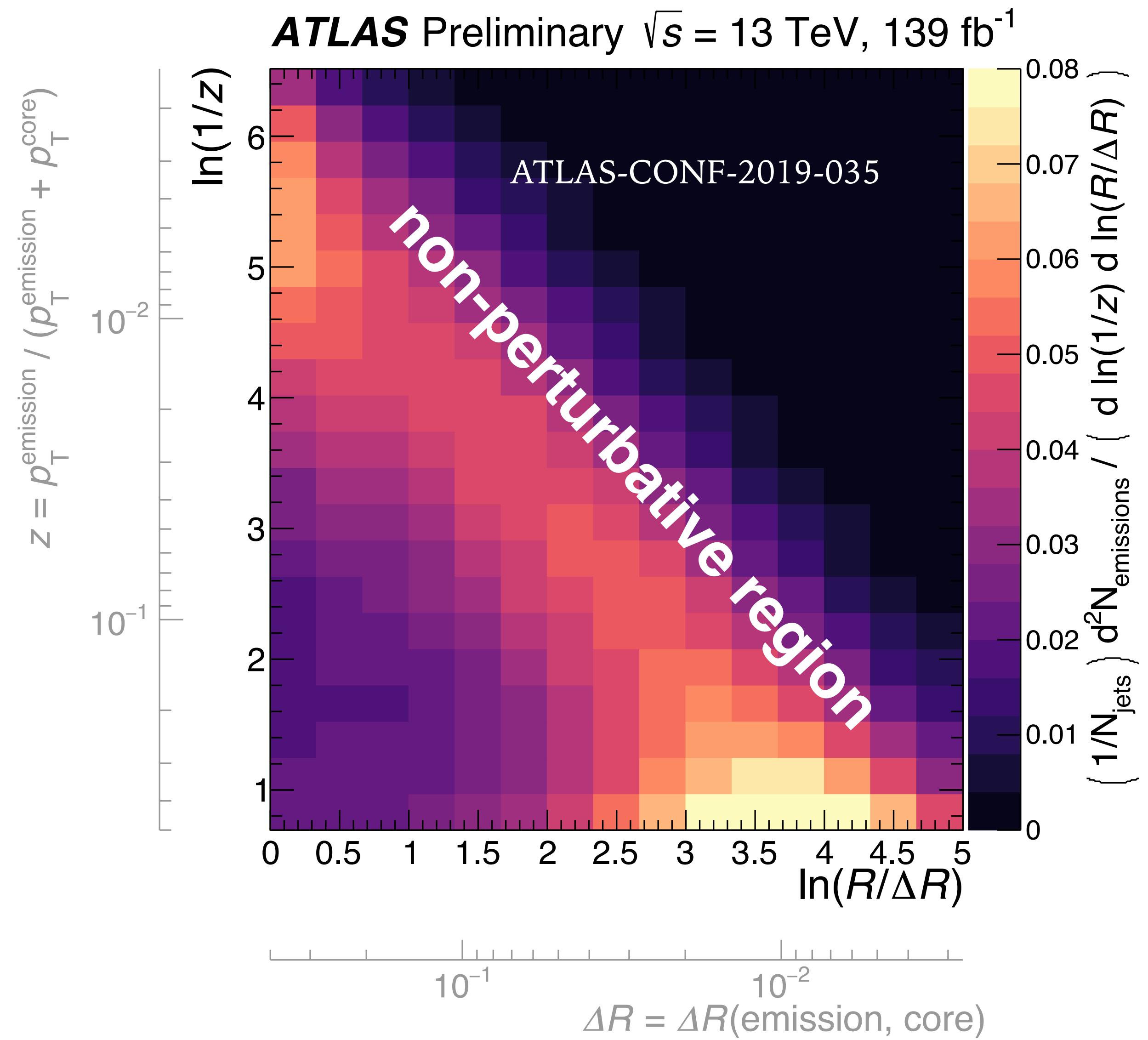
reduced  
emission  
(W colour  
singlet)



# Lund plane measurement

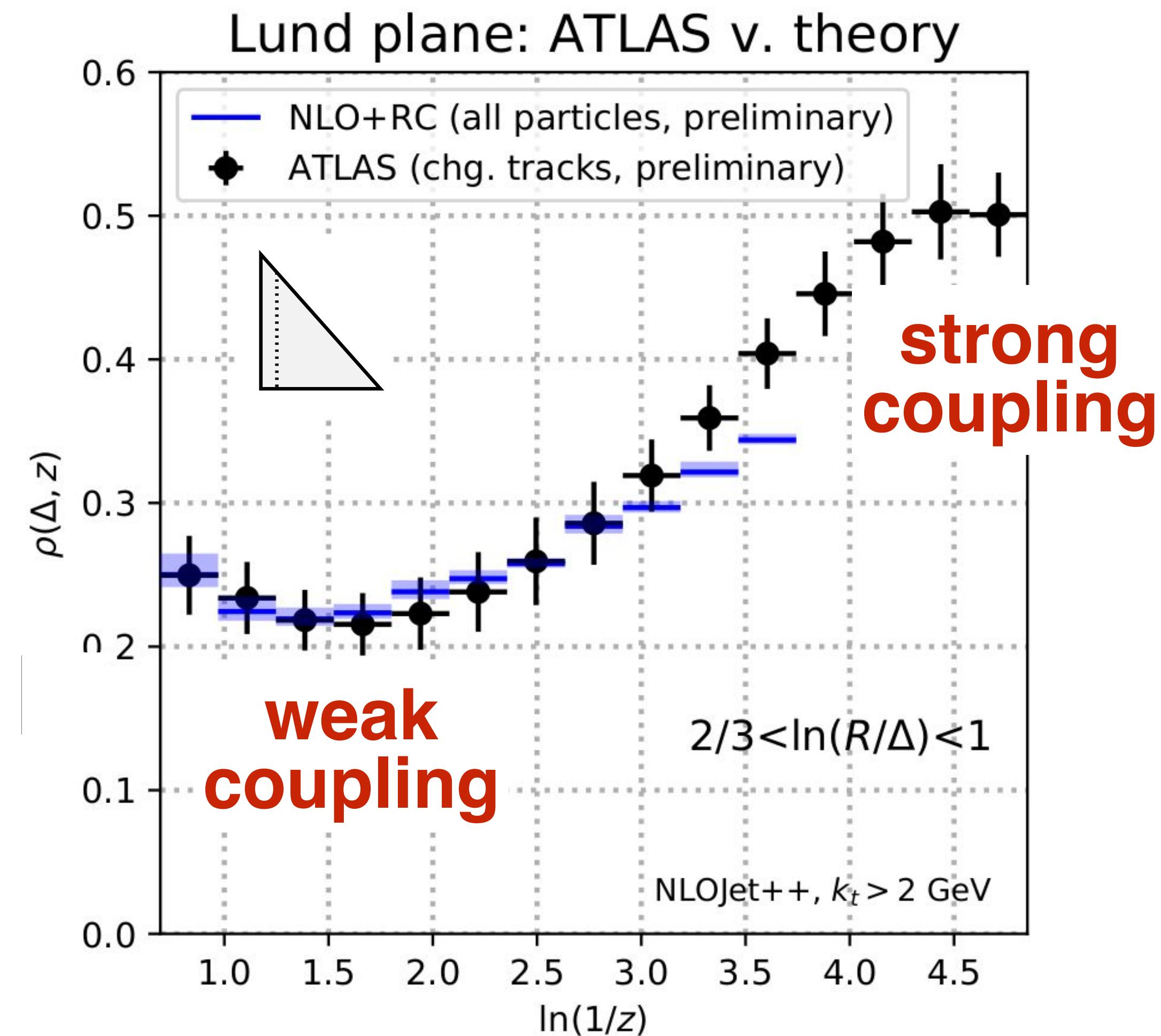
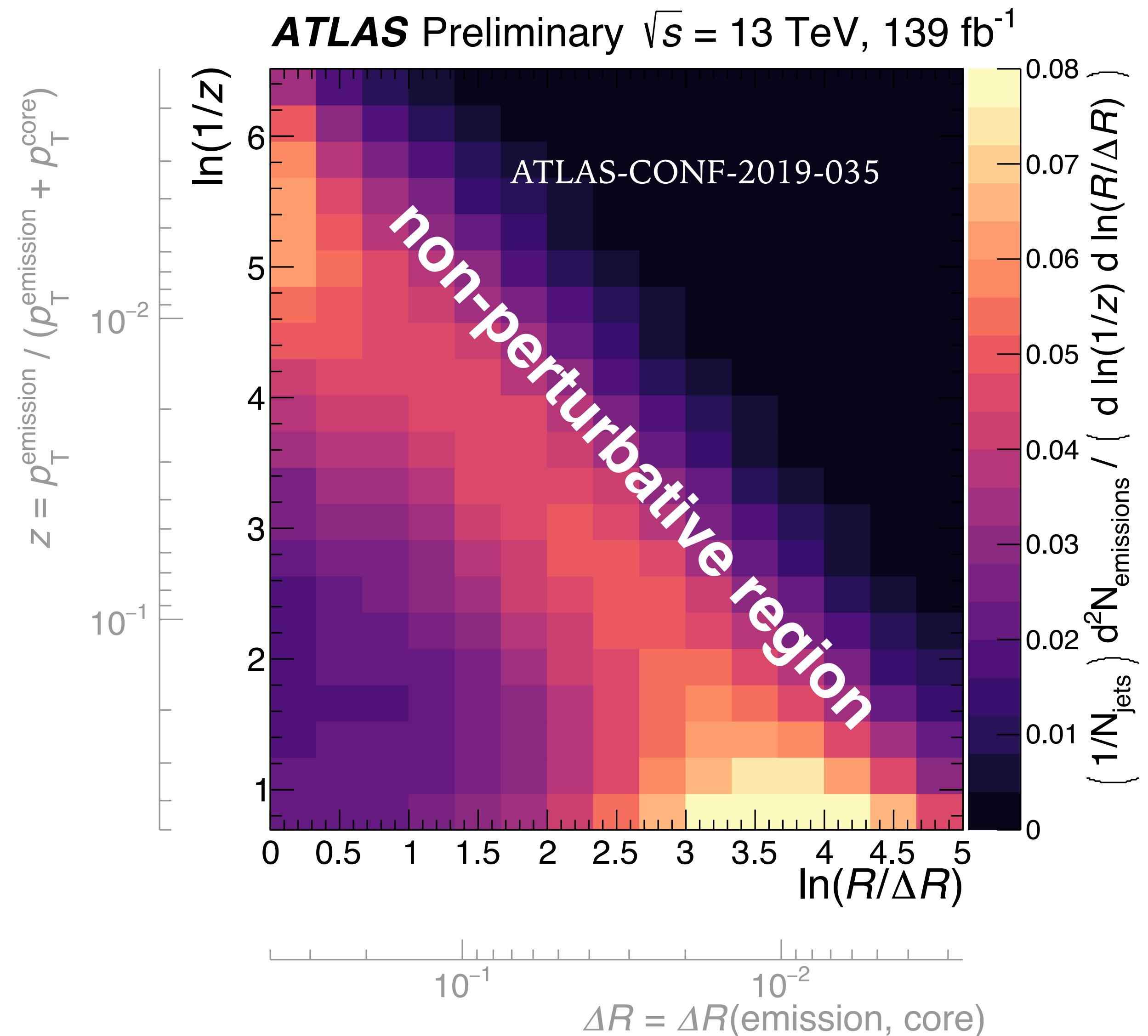


# Lund plane measurement



MC deviates up to 40% wrt data

# Lund plane measurement

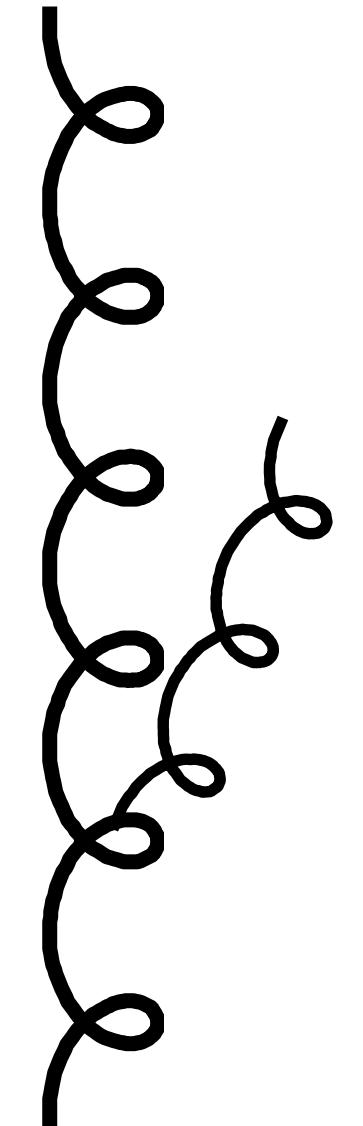
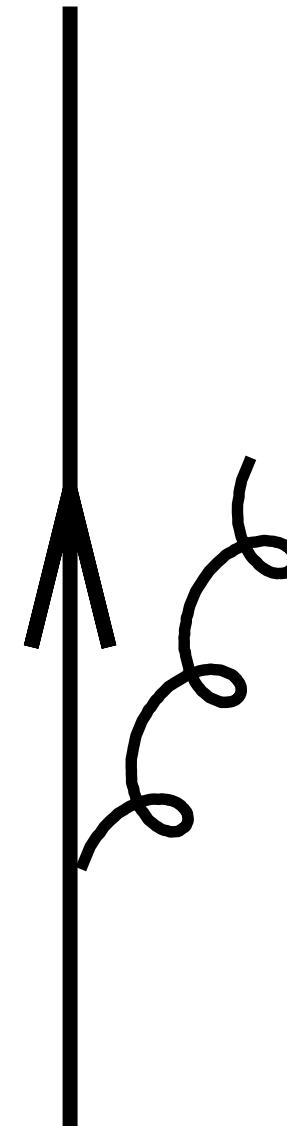


*Dreyer & Soyez prelim @Boost 2019*

# quark/gluon tagging

relies entirely on radiation pattern

# Intensity of gluon radiation from a quark or gluon



$$\frac{C_F \alpha_s}{\pi} \frac{dE}{E} \frac{d\theta^2}{\theta^2}$$

$$C_F = \frac{4}{3}$$

$$\frac{C_A \alpha_s}{\pi} \frac{dE}{E} \frac{d\theta^2}{\theta^2}$$

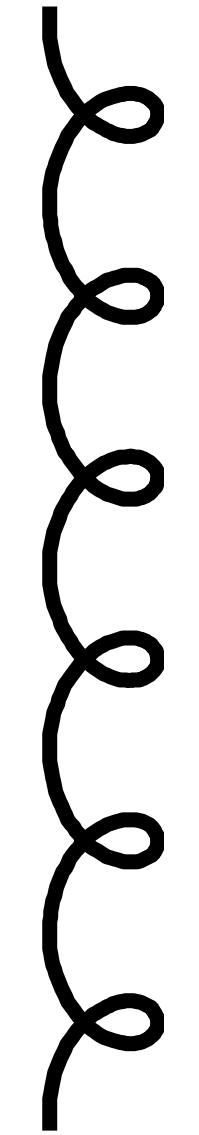
$$C_A = 3$$

Only observable difference between quark and gluon-induced jets is in the amount of radiation that they emit

# quark v. gluon-induced jet is ambiguous

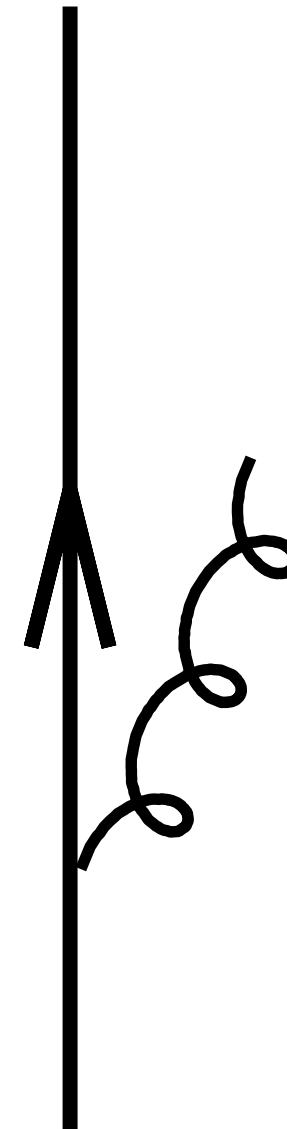


quark-induced  
jet

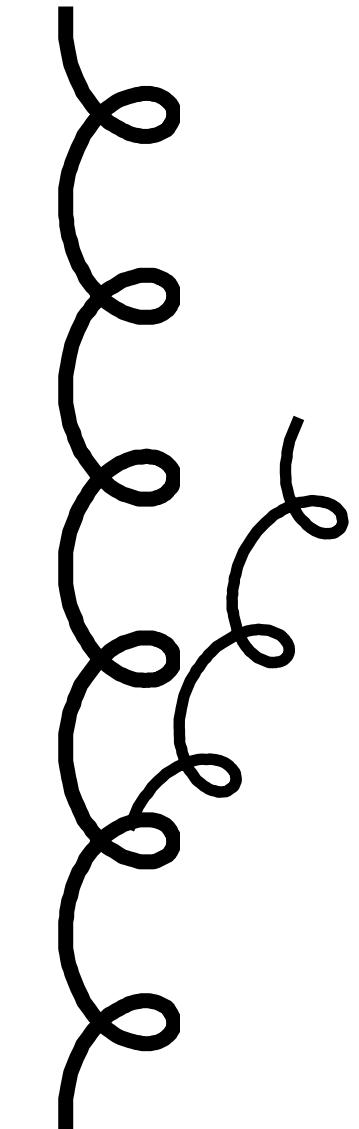


gluon-induced  
jet

# quark v. gluon-induced jet is ambiguous

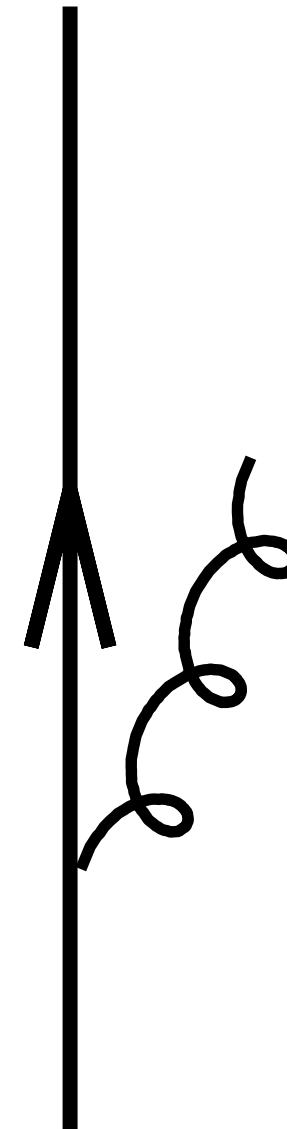


quark-induced  
jet

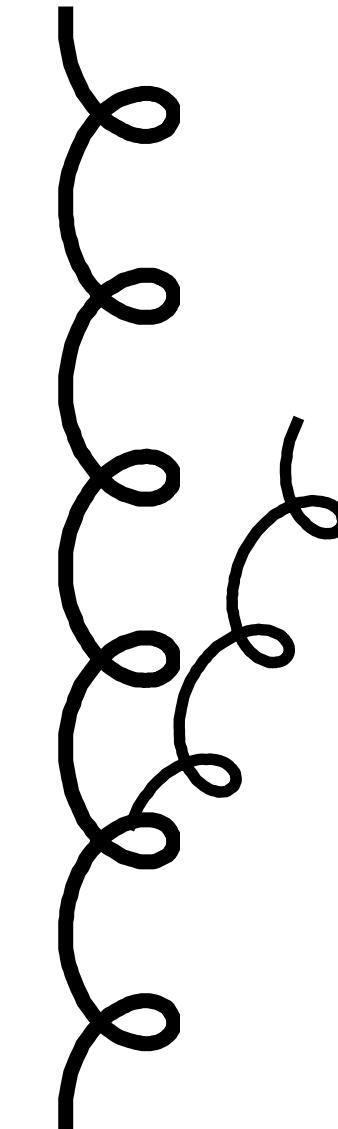


gluon-induced  
jet

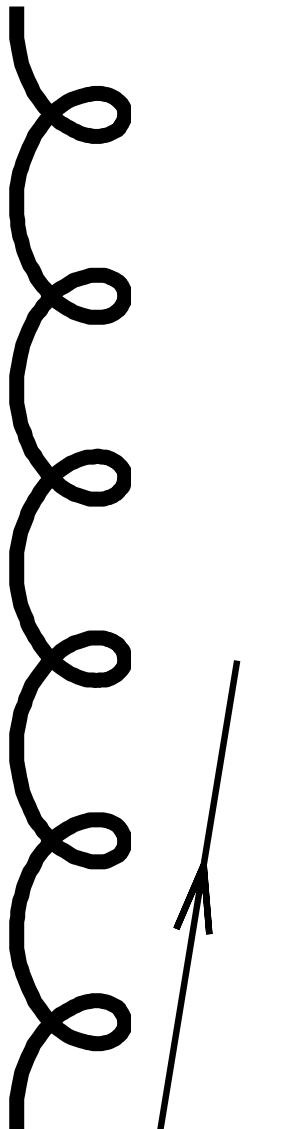
# quark v. gluon-induced jet is ambiguous



quark-induced  
jet



gluon-induced  
jet



quark-induced jet that  
radiated hard gluon?  
or gluon jet with a random  
quark that flew in?  
occurs for fraction  $\alpha_s$  of jets

# Pythia dijet and Z+jet sub-process fractions

dijet sample  $p_t > 250$  GeV

I	I	I	I	I	I
I	Subprocess	I	Number of points	I	Sigma
I	I	I	I	I	I
I	I	I	I	I	I
I	N:o Type	I	Generated	Tried	I
I	I	I	I	I	I
I	0 All included subprocesses	I	1000	6578	I 2.256D-05
I	11 f + f' -> f + f' (QCD)	I	139	828	I 3.000D-06
I	12 f + fbar -> f' + fbar'	I	6	9	I 6.865D-08
I	13 f + fbar -> g + g	I	1	4	I 5.411D-08
I	28 f + g -> f + g	I	553	4209	I 1.193D-05
I	53 g + g -> f + fbar	I	12	47	I 2.856D-07
I	68 g + g -> g + g	I	289	1481	I 7.228D-06
I	I	I	I	I	I

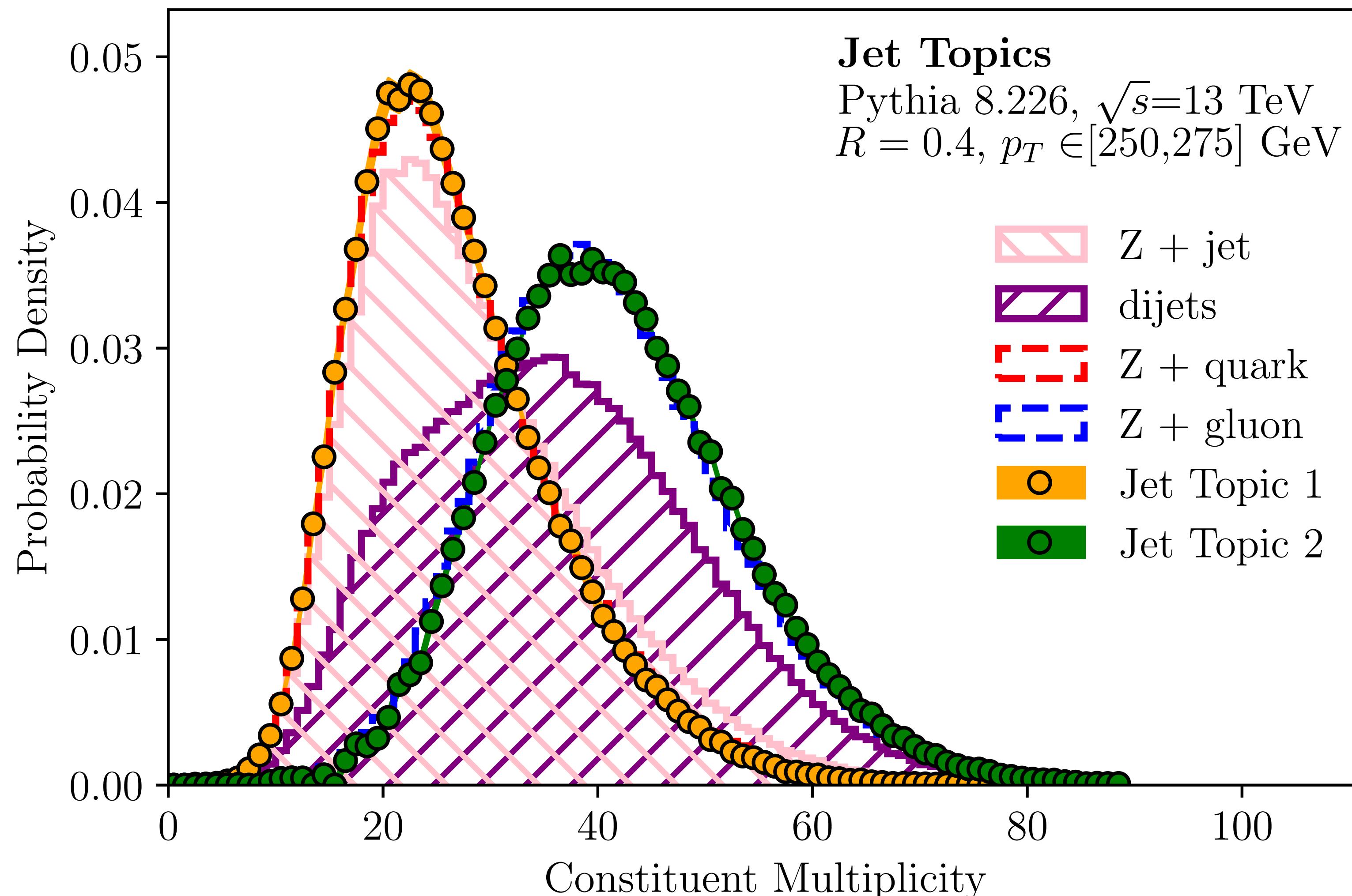
57% gluon jets

Z+jet sample  $p_t > 250$  GeV

I	I	I	I	I	I
I	Subprocess	I	Number of points	I	Sigma
I	I	I	I	I	I
I	I	I	I	I	I
I	N:o Type	I	Generated	Tried	I
I	I	I	I	I	I
I	0 All included subprocesses	I	1000	4111	I 2.589D-08
I	15 f + fbar -> g + Z0	I	151	546	I 3.971D-09
I	30 f + g -> f + Z0	I	849	3565	I 2.192D-08
I	I	I	I	I	I

85% quark jets

# jet “topics”, mutual irreducibility

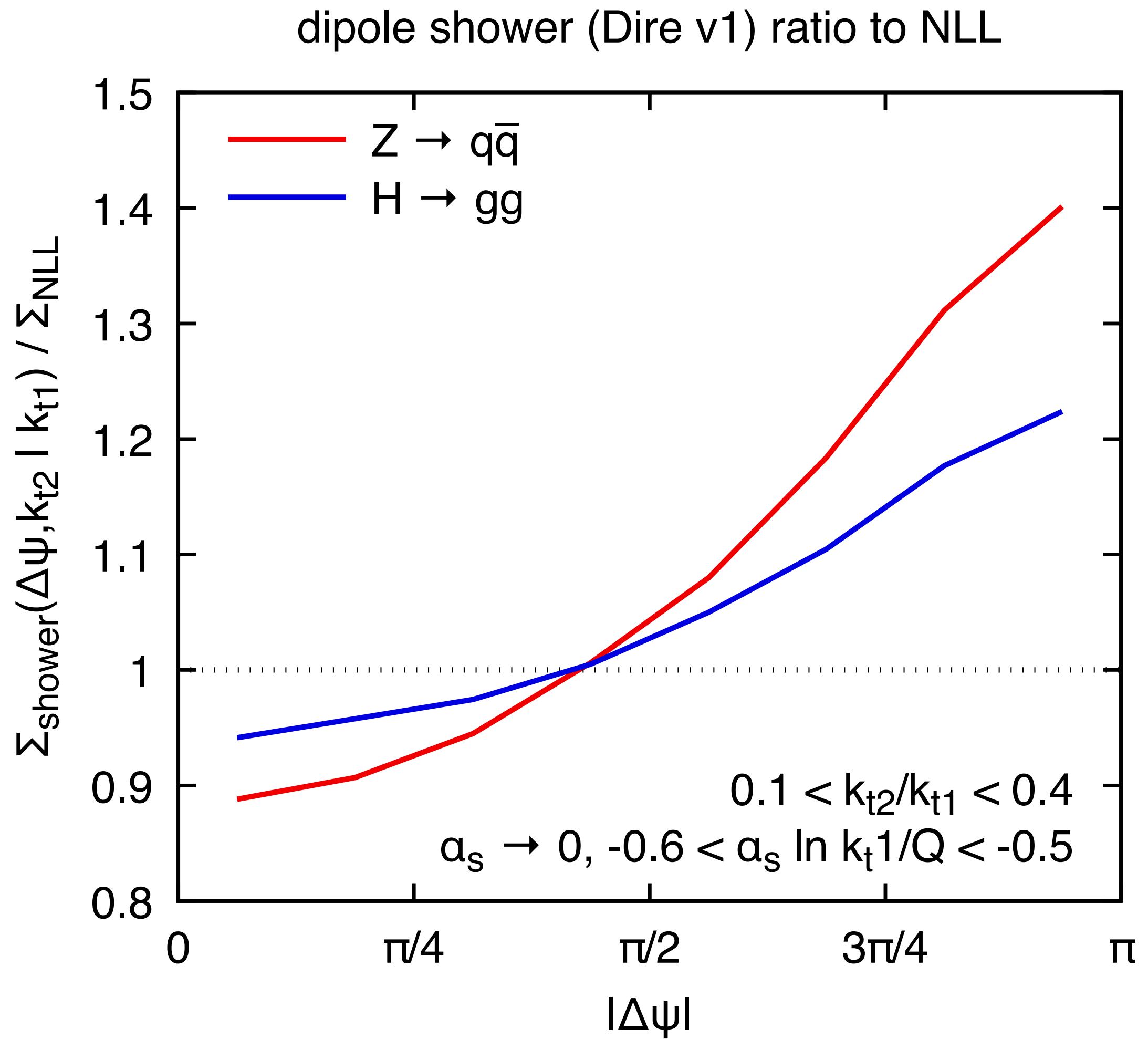


particle multiplicity  $N$   
distribution has region at  
low  $N$  (high  $N$ ) populated  
only by quarks (gluons)  
[mutual irreducibility]

That information, together  
with two samples with  
different (but unknown) q/g  
fractions, is enough to  
operationally define q/g  
fractions in each sample

Metodiev & Thaler, [arXiv:1802.00008](https://arxiv.org/abs/1802.00008) ,  
Komiske, Metodiev & Thaler, [arXiv:1809.01140](https://arxiv.org/abs/1809.01140)

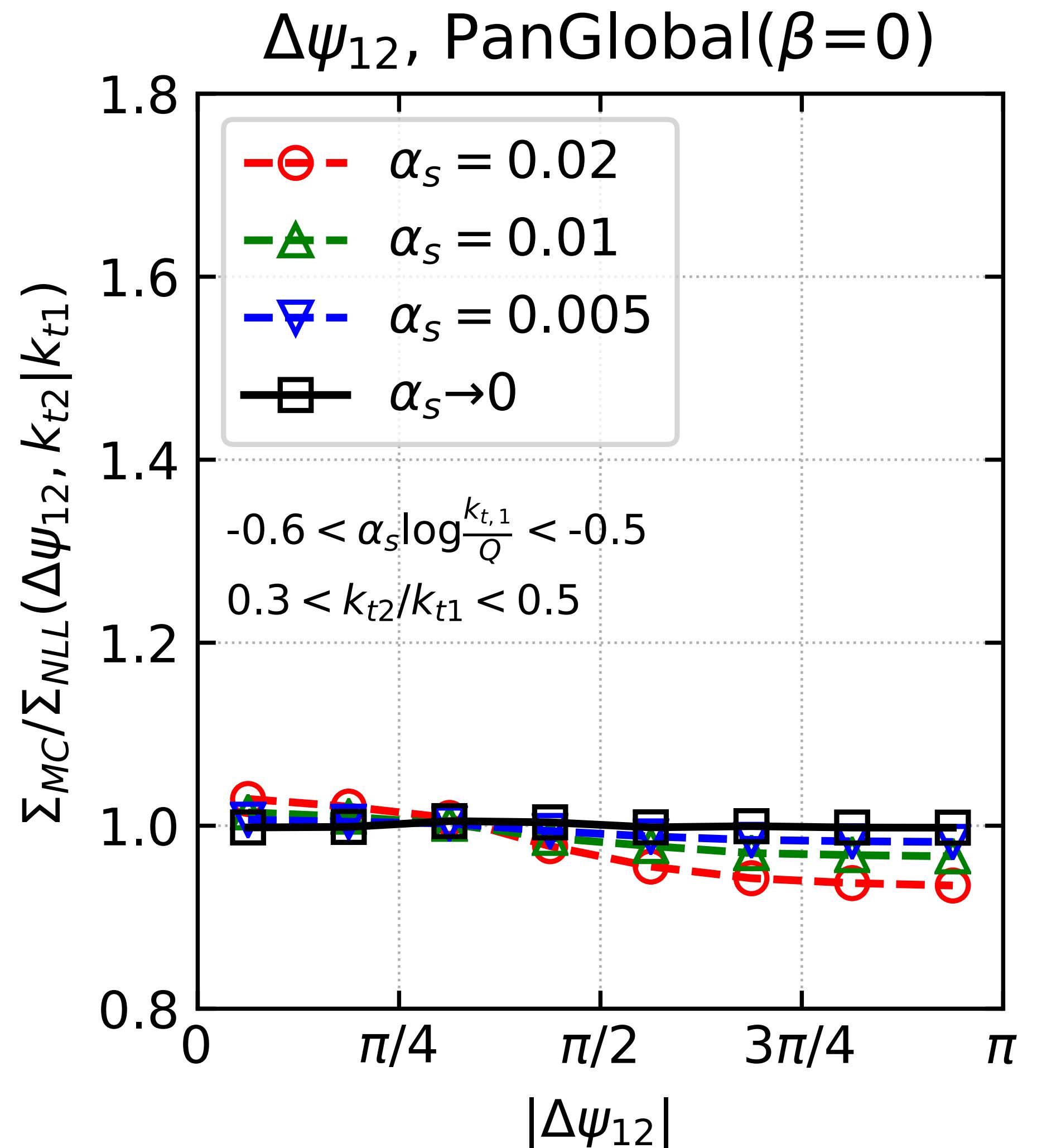
# beware, parton showers are not perfect



common dipole showers display quark/gluon differences (e.g. in some azimuthal observables) that shouldn't be there

especially relevant for machine learning

# beware, parton showers are not perfect



common dipole showers display quark/gluon differences (e.g. in some azimuthal observables) that shouldn't be there

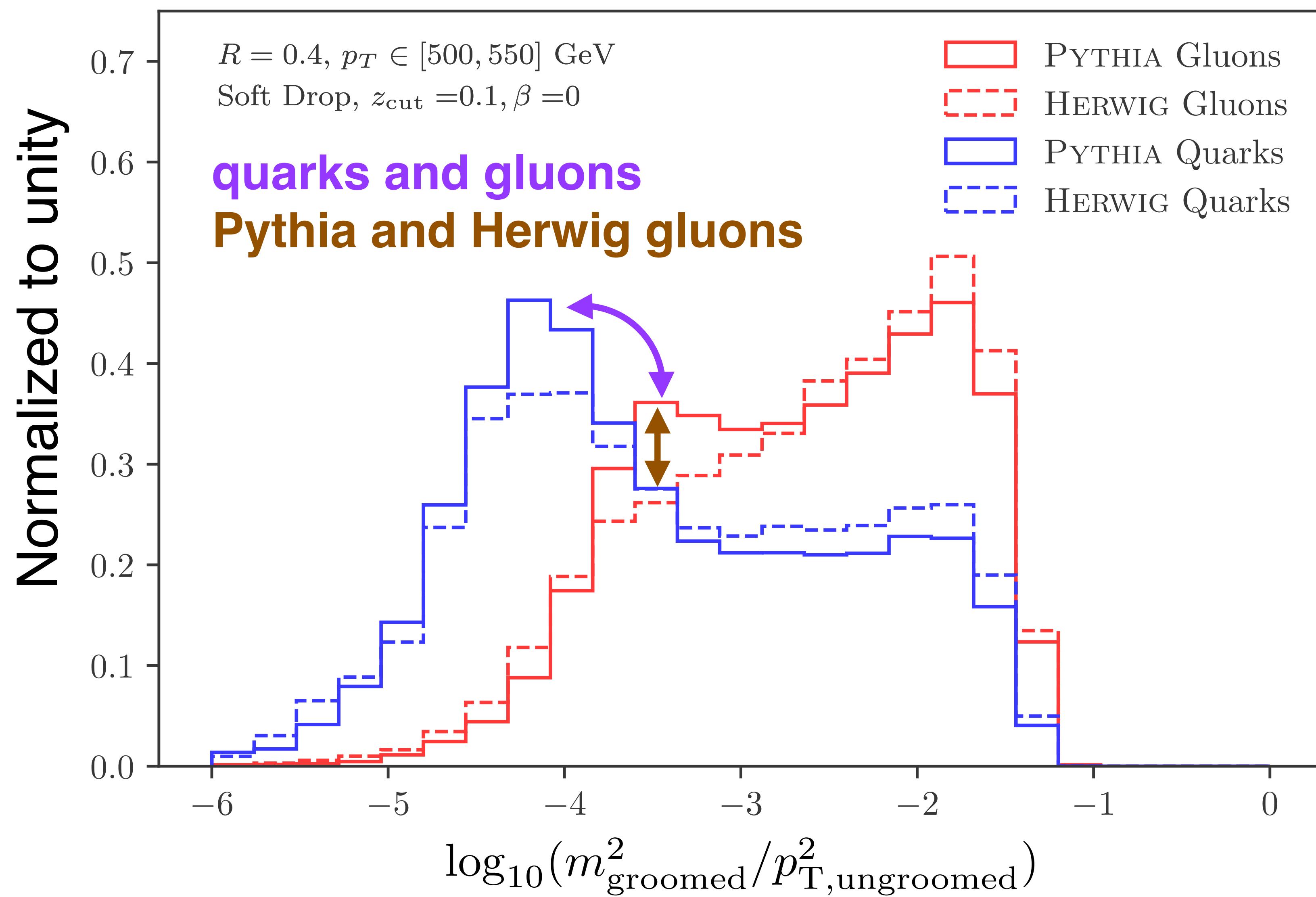
time has come to develop new generation of parton showers with quantifiable accuracy

**especially relevant for machine learning**

- **SoftDrop**: for extracting underlying hard jet substructure (e.g. 2-prong structure). Uses key idea of C/A declustering, gives distributions with well-understood characteristics (e.g. only collinear single logarithms in resummation)
- **Subjettiness / energy-energy-correlations / energy-flow polynomials / Lund Plane structure**: all try to measure the additional radiation flow around the core n-prong structure of a jet (e.g. 2-prong for Higgs decay)
- **Top performance ultimately comes from machine learning**, but we should still aim to understand the information we are using, whether data or MC driven, and questions such as assigning meaning to the categories we use (e.g. quark v gluon-induced jets)

- **SoftDrop**: uses key idea of C/A declustering, gives distributions with well-understood characteristics (e.g. only collinear single logarithms in resummation)
- **Subjettiness / energy-energy-correlations / energy-flow polynomials / Lund Plane structure**: all try to measure the energy flow around the core n-prong structure of a jet (e.g. 2-prong for Higgs decay)
- **Machine learning**: jet substructure is one of the most dynamic playgrounds for ML, with large gains to be had in pulling out all info from jets

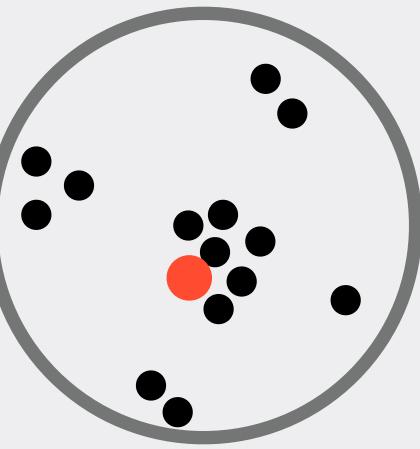
# EXTRAS



- comment on differences
- but at some level they are most prominent at non-perturbative level (which is around -4.3)

*fragmentation  
function*

$$D(z) = \left\langle \sum_{i \in \text{jet}} \delta(z - p_{ti}/p_{t,\text{jet}}) \right\rangle_{\text{jets}}$$

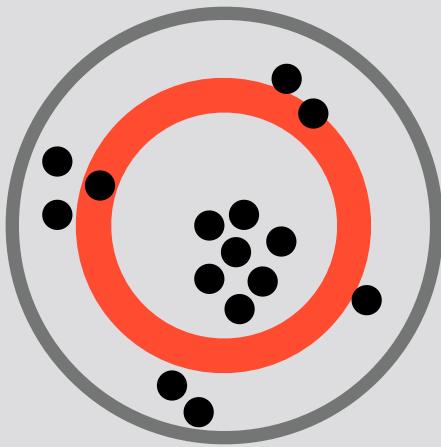
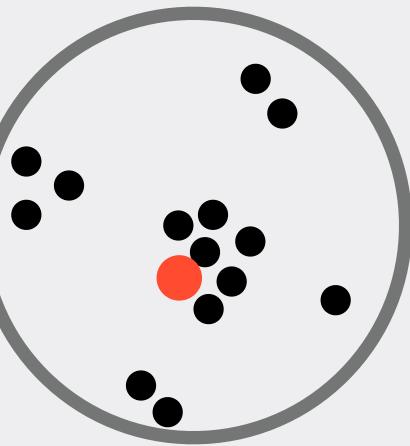


*fragmentation  
function*

*differential  
jet shape*

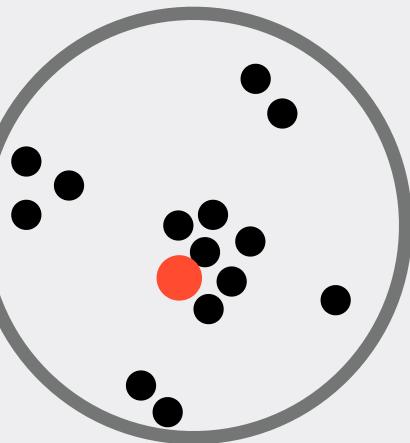
$$D(z) = \left\langle \sum_{i \in \text{jet}} \delta(z - p_{ti}/p_{t,\text{jet}}) \right\rangle_{\text{jets}}$$

$$\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r + \delta r]}} p_{\perp}^{(k)},$$



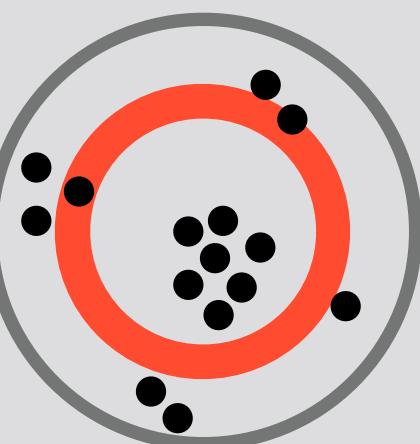
*fragmentation  
function*

$$D(z) = \left\langle \sum_{i \in \text{jet}} \delta(z - p_{ti}/p_{t,\text{jet}}) \right\rangle_{\text{jets}}$$



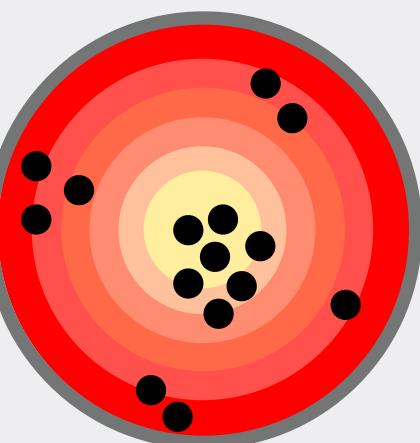
*differential  
jet shape*

$$\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r + \delta r]}} p_{\perp}^{(k)},$$



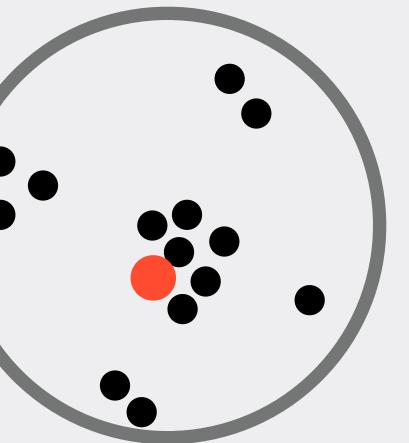
*width=broadening*

$$g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta R_{kJ},$$



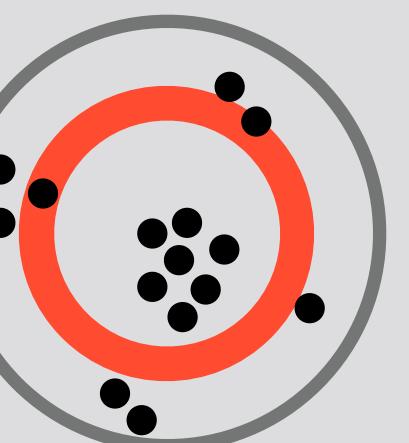
*fragmentation  
function*

$$D(z) = \left\langle \sum_{i \in \text{jet}} \delta(z - p_{ti}/p_{t,\text{jet}}) \right\rangle_{\text{jets}}$$



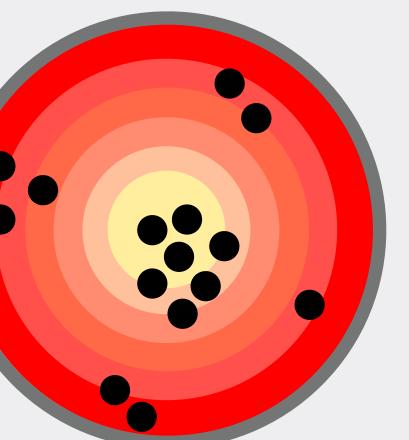
*differential  
jet shape*

$$\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r + \delta r]}} p_{\perp}^{(k)},$$



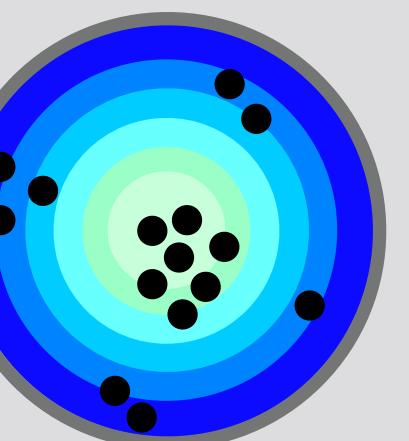
*width=broadening*

$$g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta R_{kJ},$$



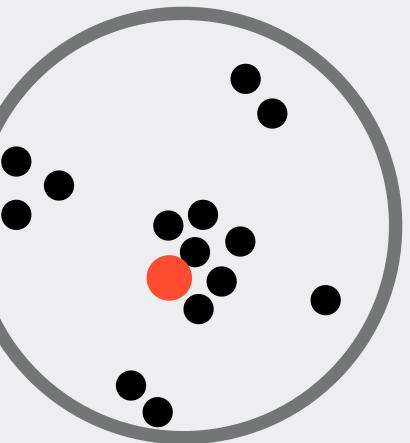
*jet mass, groomed  
& ungroomed*

$$m^2 = \left( \sum_{i \in (\text{sub})\text{jet}} p_i^{\mu} \right)^2$$



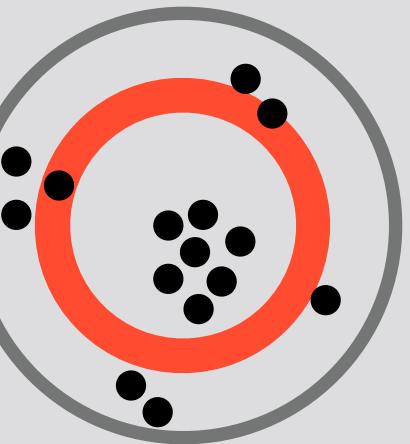
*fragmentation function*

$$D(z) = \left\langle \sum_{i \in \text{jet}} \delta(z - p_{ti}/p_{t,\text{jet}}) \right\rangle_{\text{jets}}$$



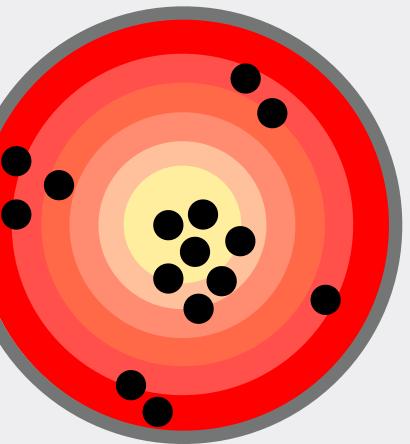
*differential jet shape*

$$\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r + \delta r]}} p_{\perp}^{(k)},$$



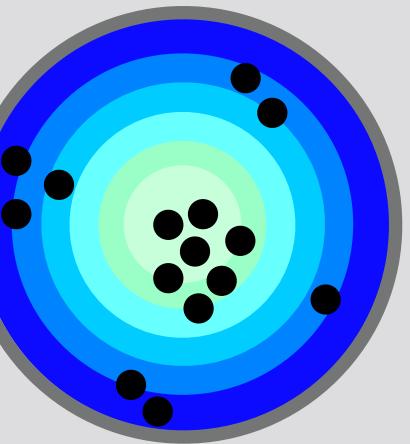
*width=broadening*

$$g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta R_{kJ},$$



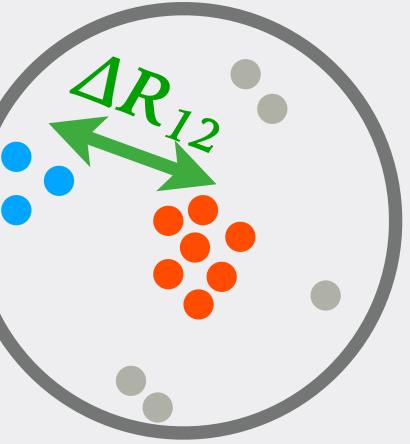
*jet mass, groomed & ungroomed*

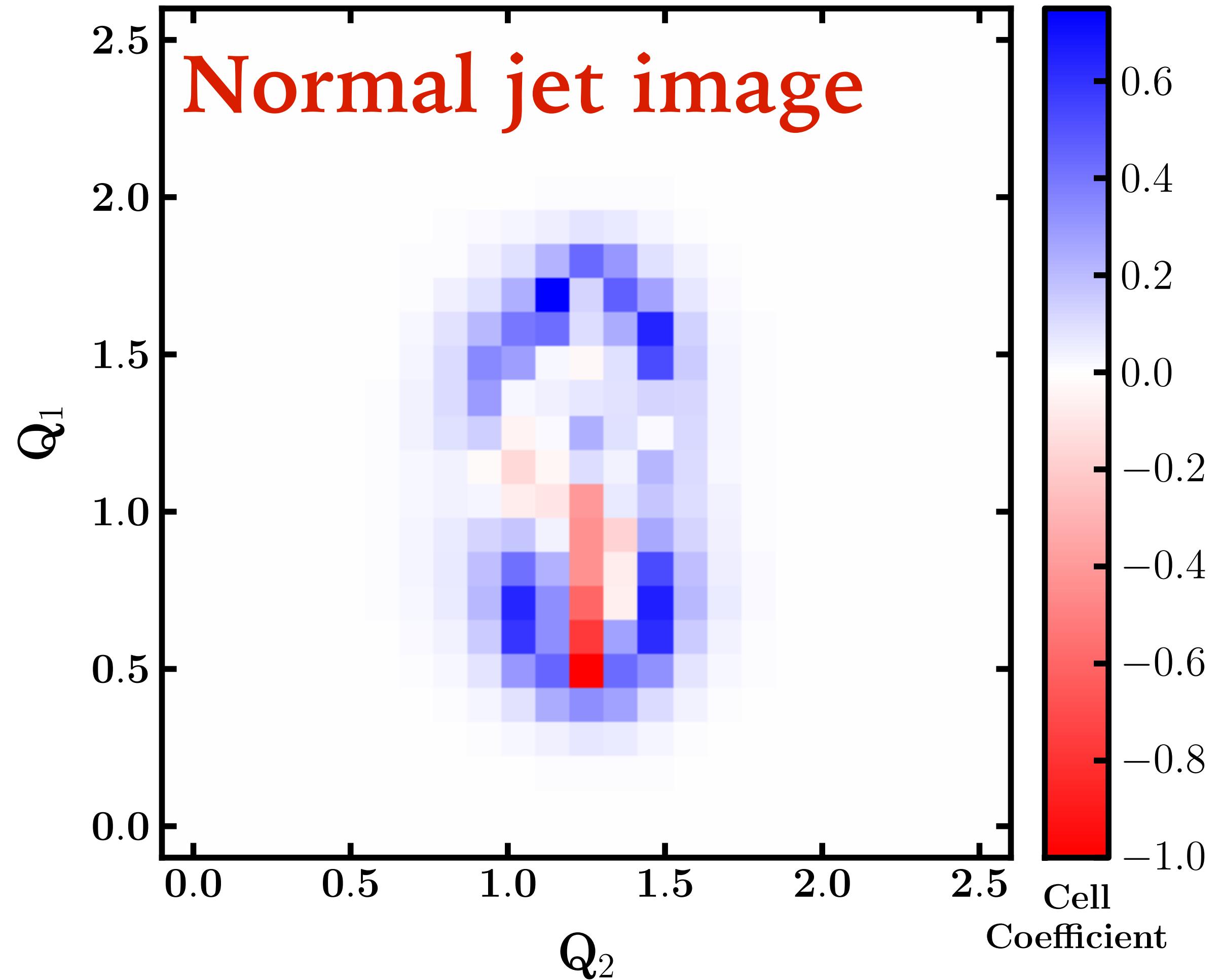
$$m^2 = \left( \sum_{i \in (\text{sub})\text{jet}} p_i^{\mu} \right)^2$$



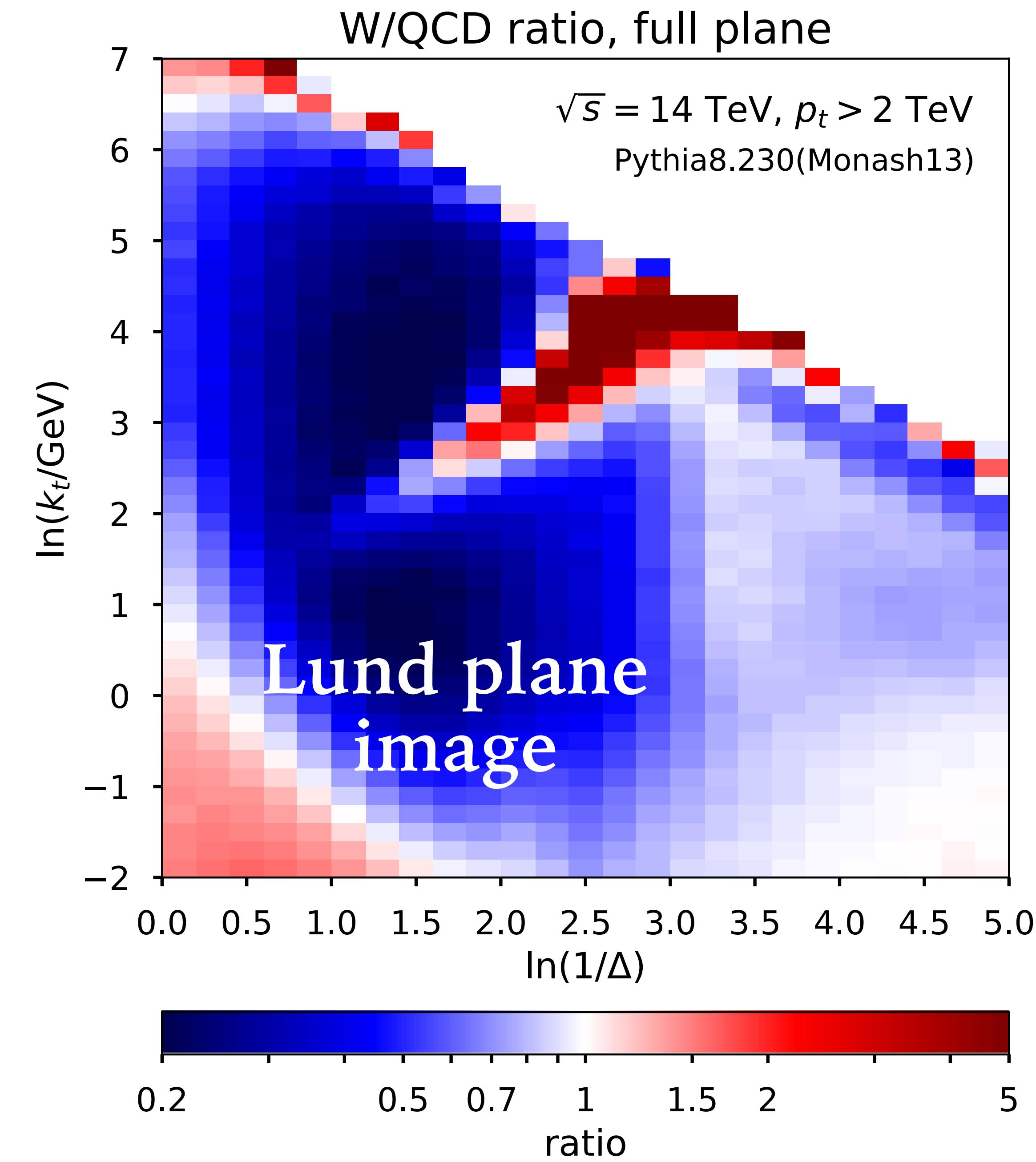
$z_g, \Delta R_{12}$

$$z_g = \frac{\min(p_{\perp,1}, p_{\perp,2})}{p_{\perp,1} + p_{\perp,2}} > z_{\text{cut}} \left( \frac{\Delta R_{1,2}}{R_J} \right)^{\beta}$$



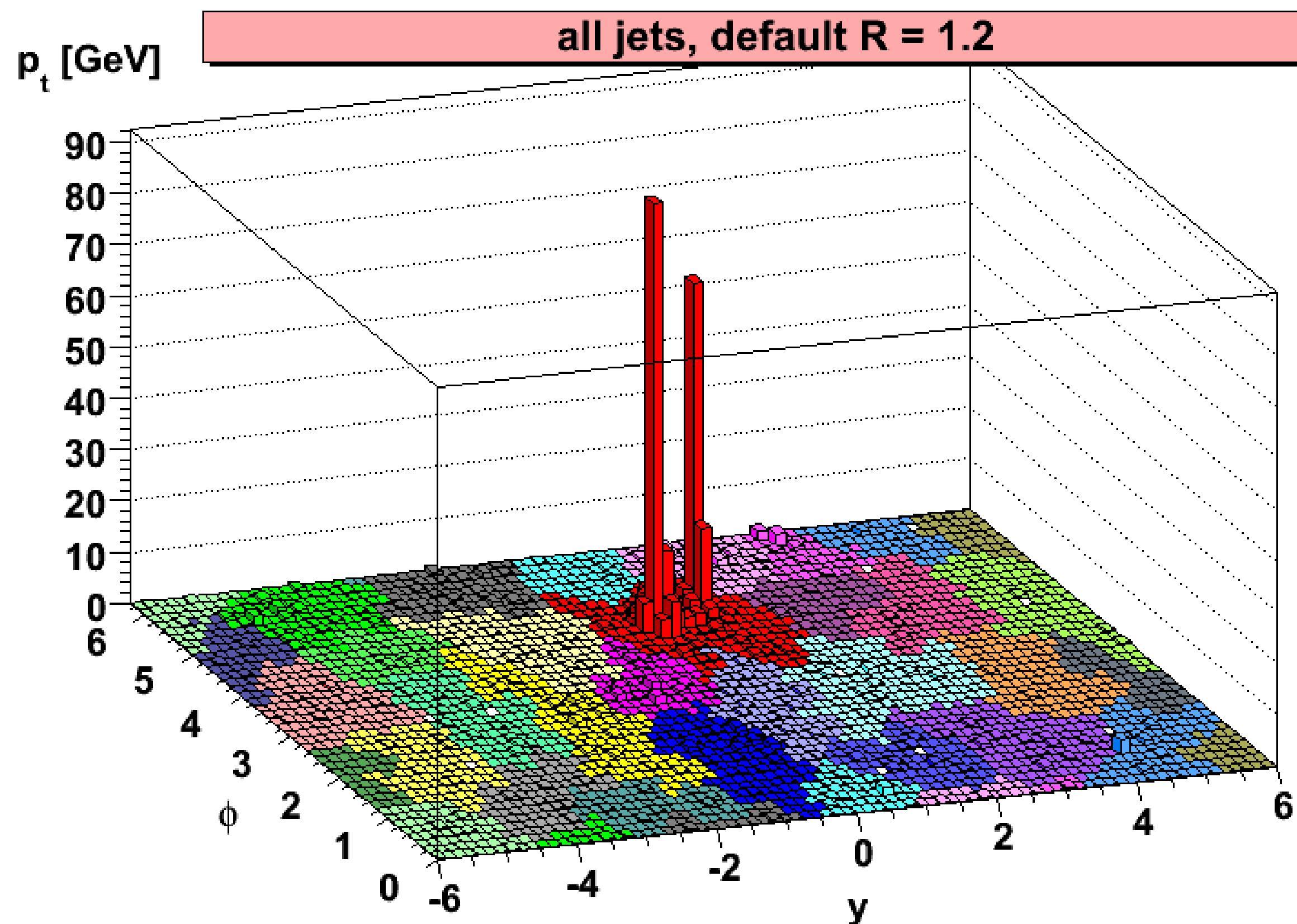


*Cogan, Kagan, Strauss, Schwartzman, 1407.5674*



Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

SIGNAL



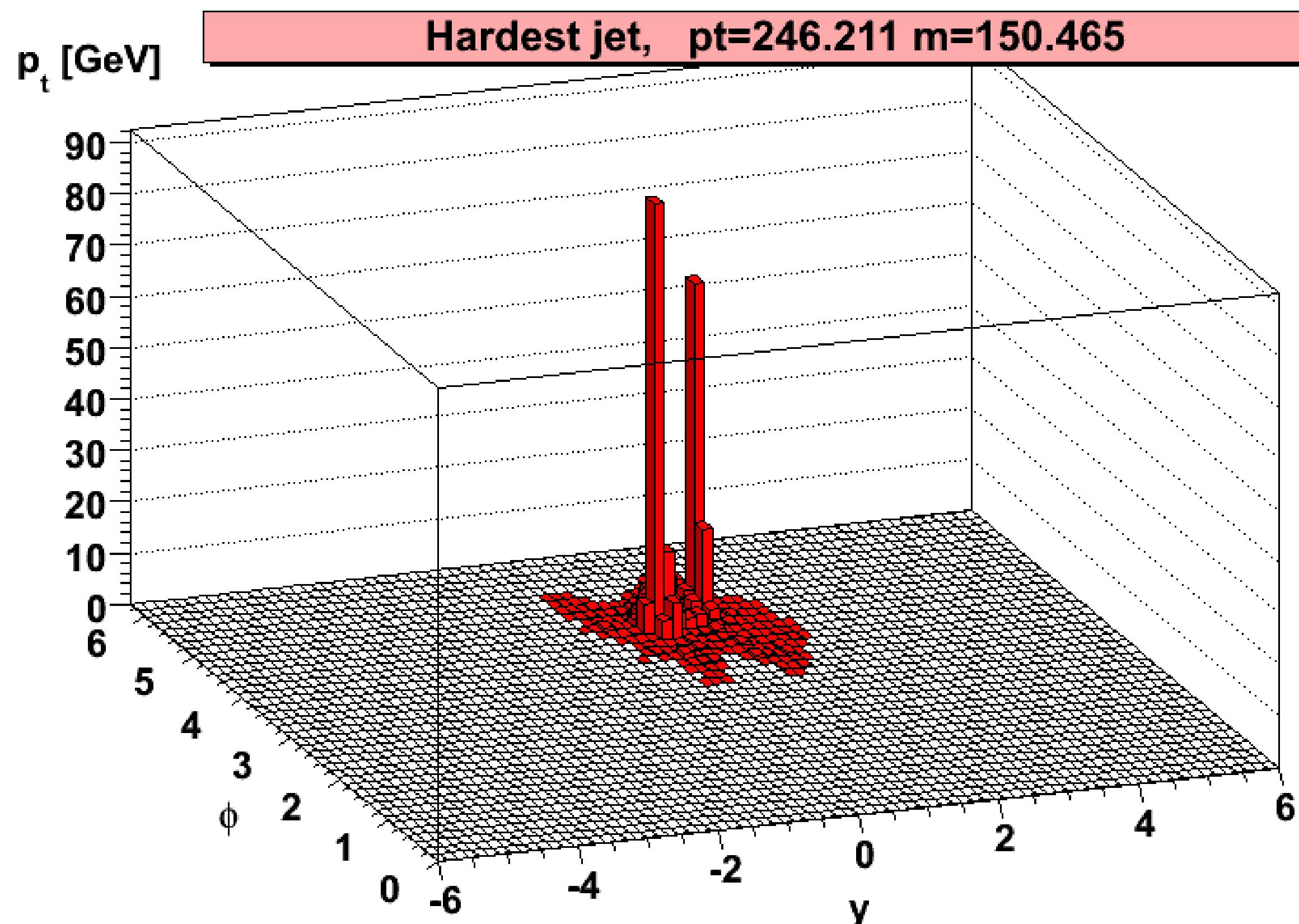
Zbb BACKGROUND

Fill it in,  $\rightarrow$  show jets more clearly

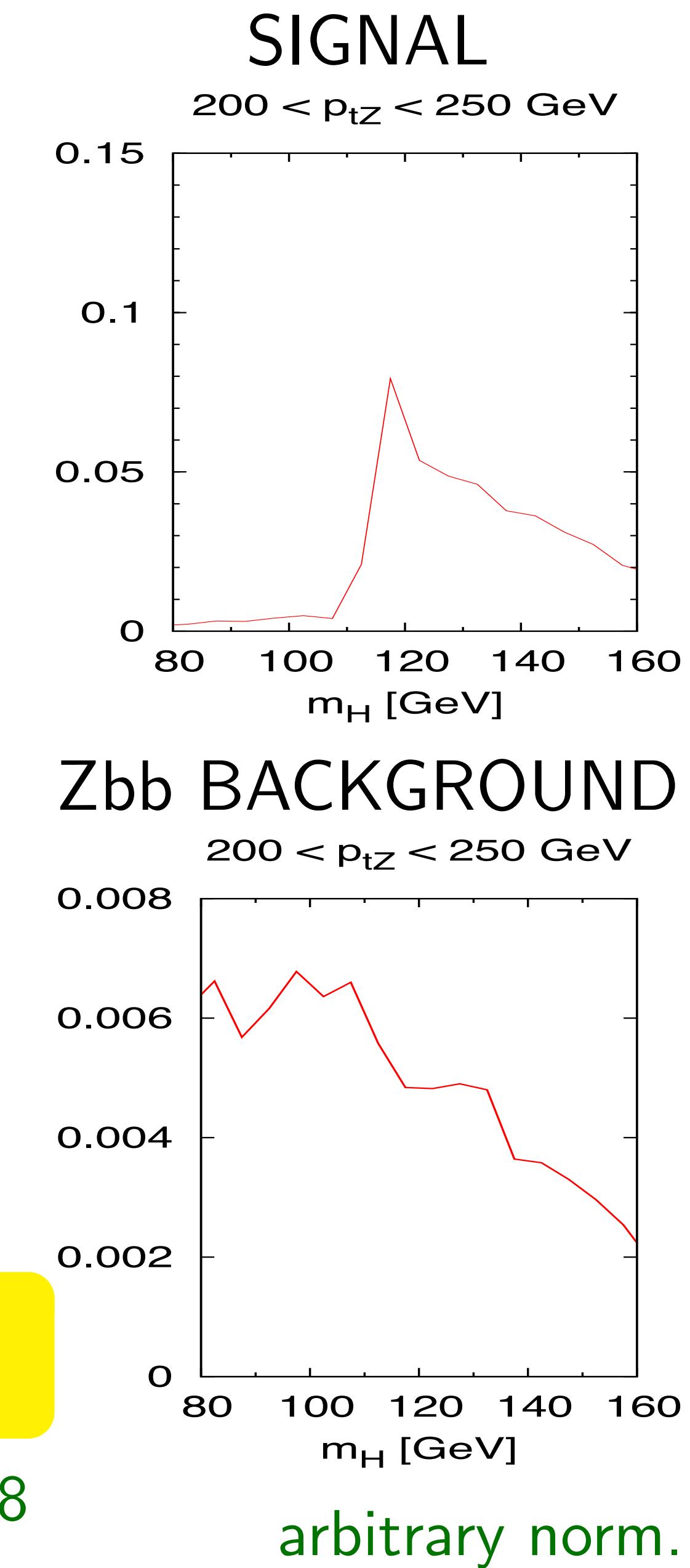
Butterworth, Davison, Rubin & GPS '08

arbitrary norm.

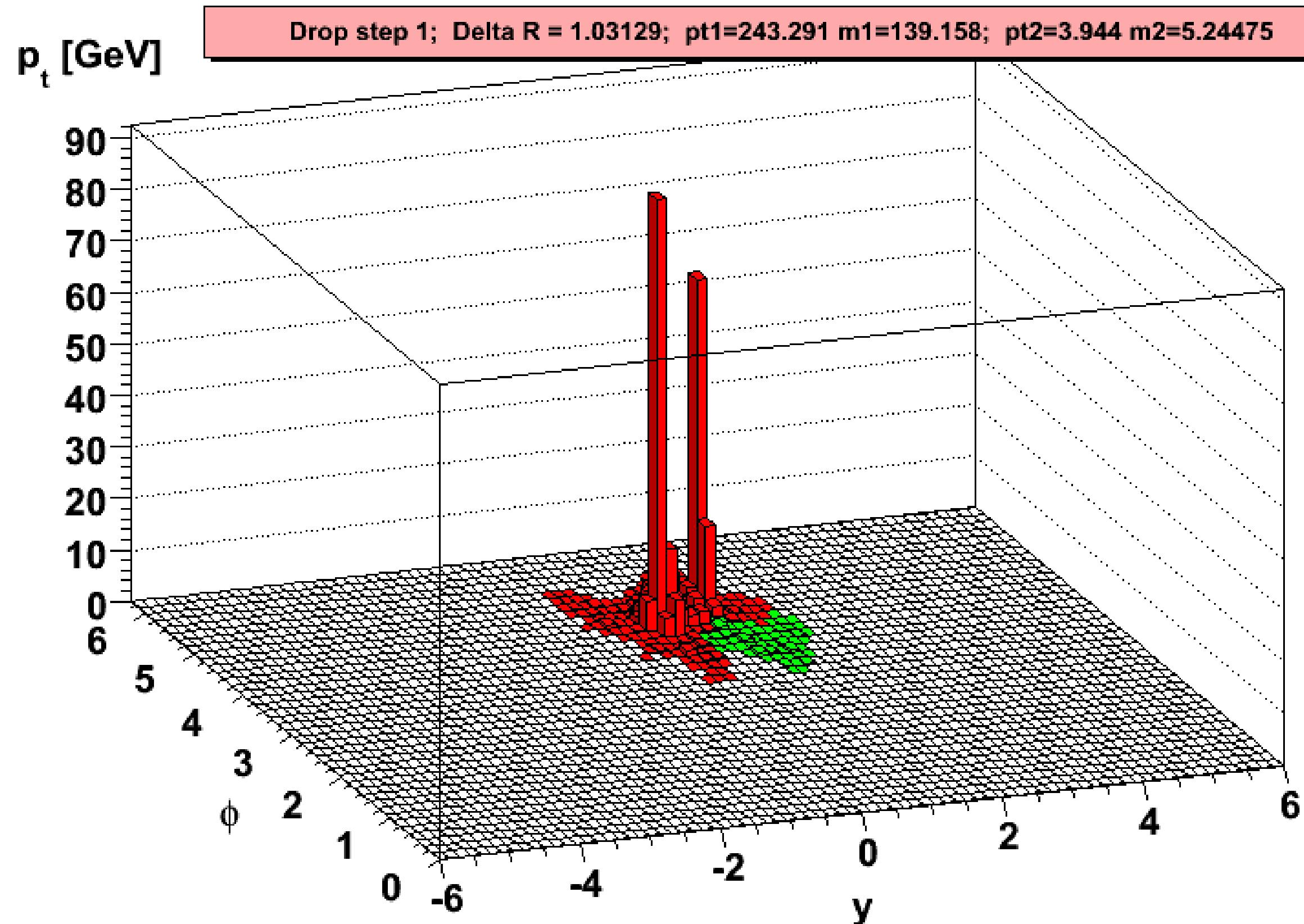
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



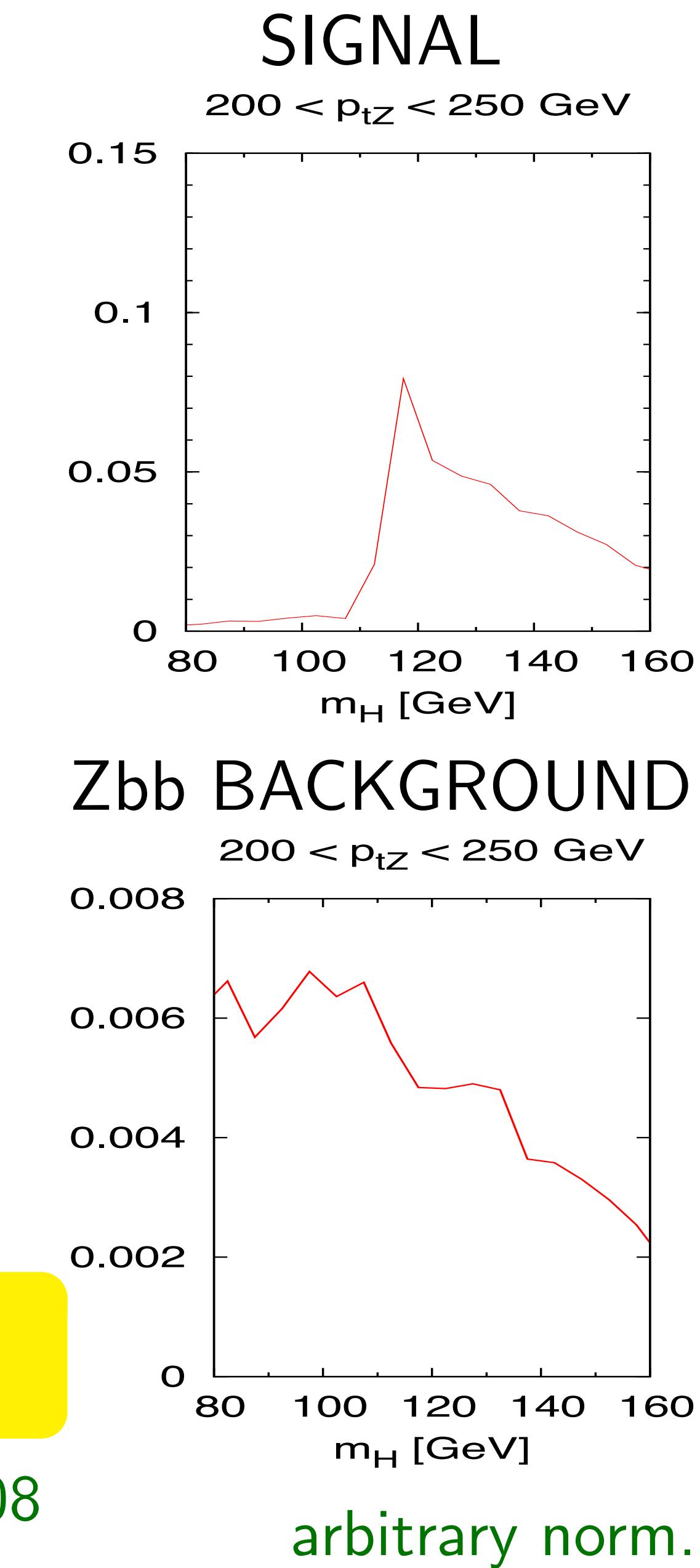
Consider hardest jet,  $m = 150$  GeV



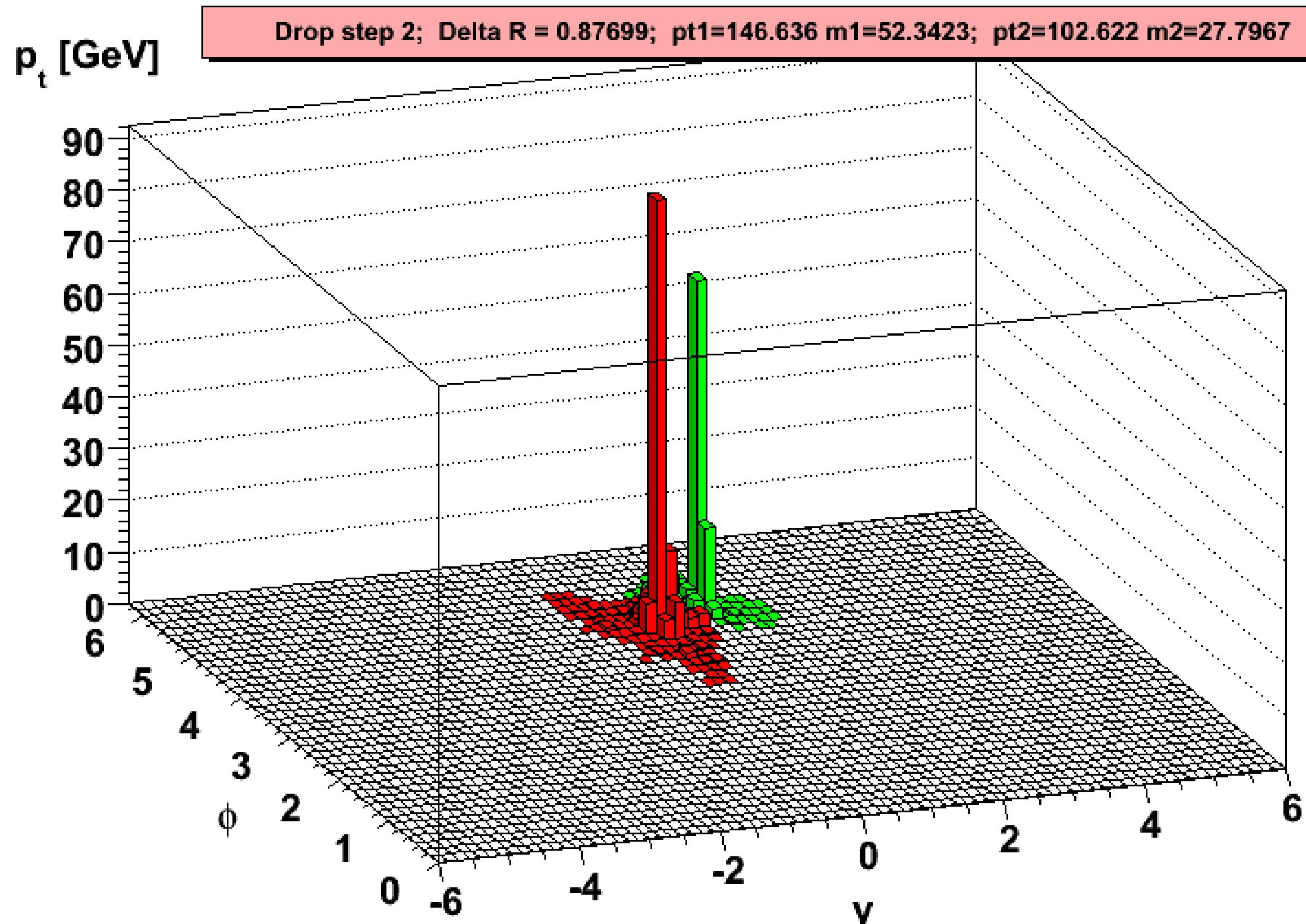
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



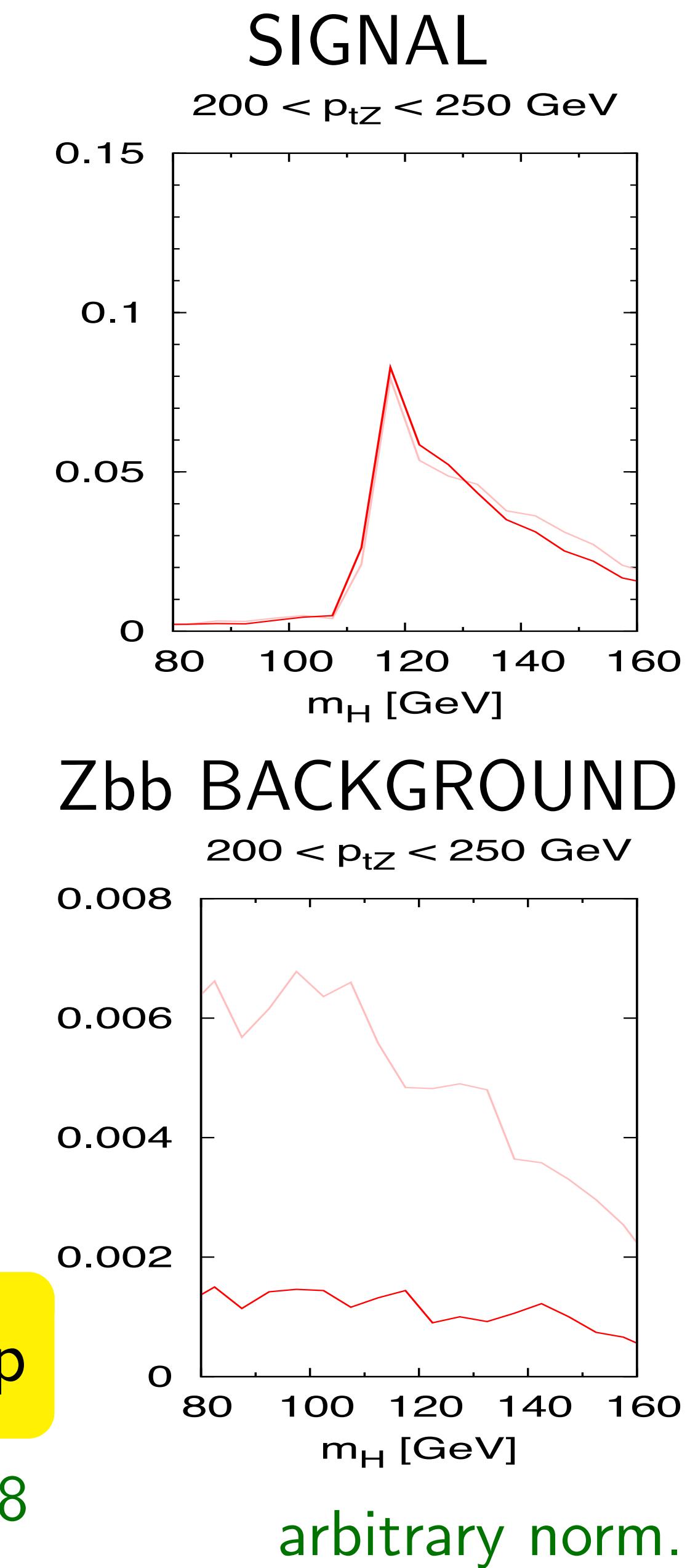
split:  $m = 150$  GeV,  $\frac{\max(m_1, m_2)}{m} = 0.92 \rightarrow$  repeat



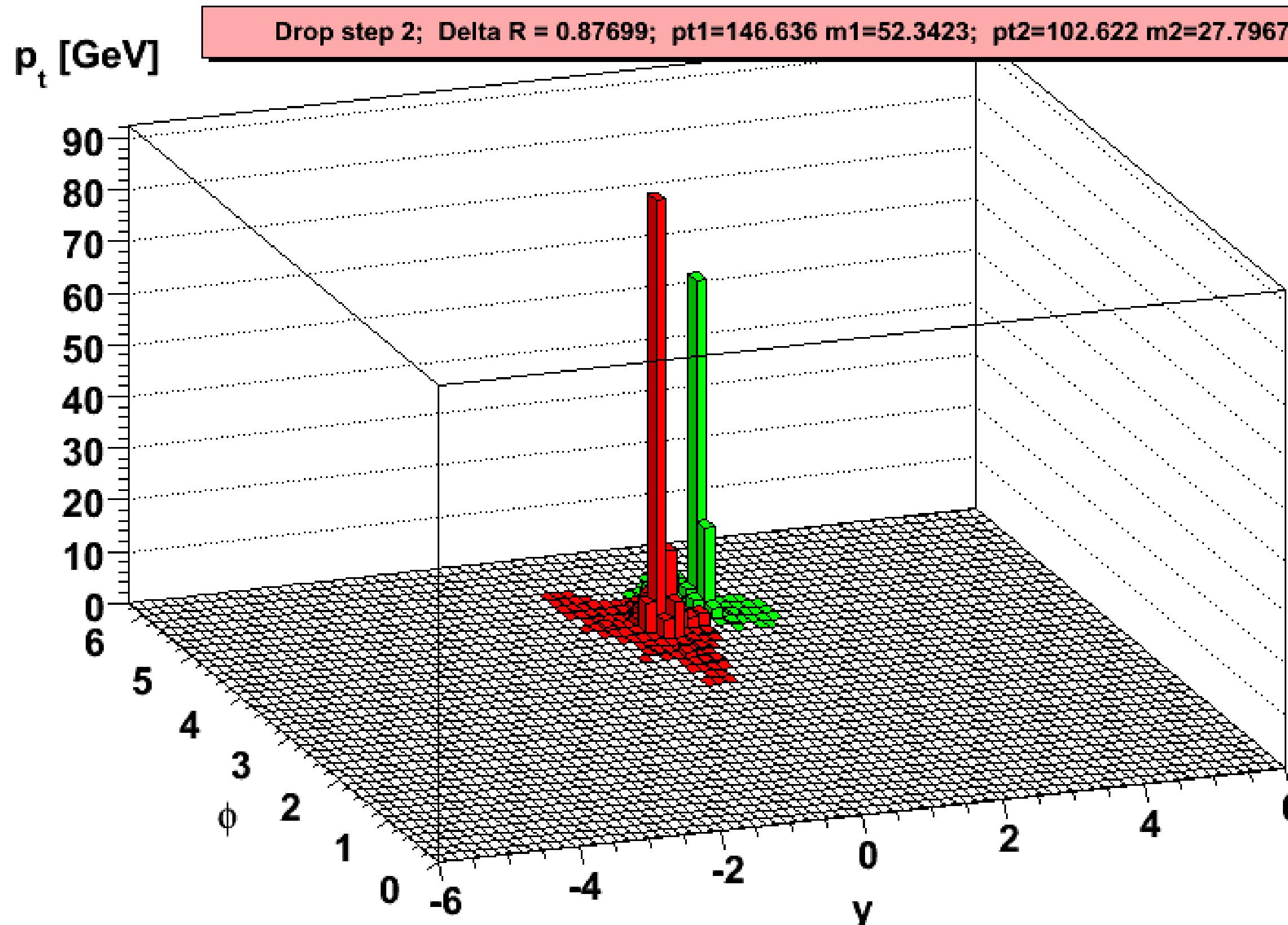
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



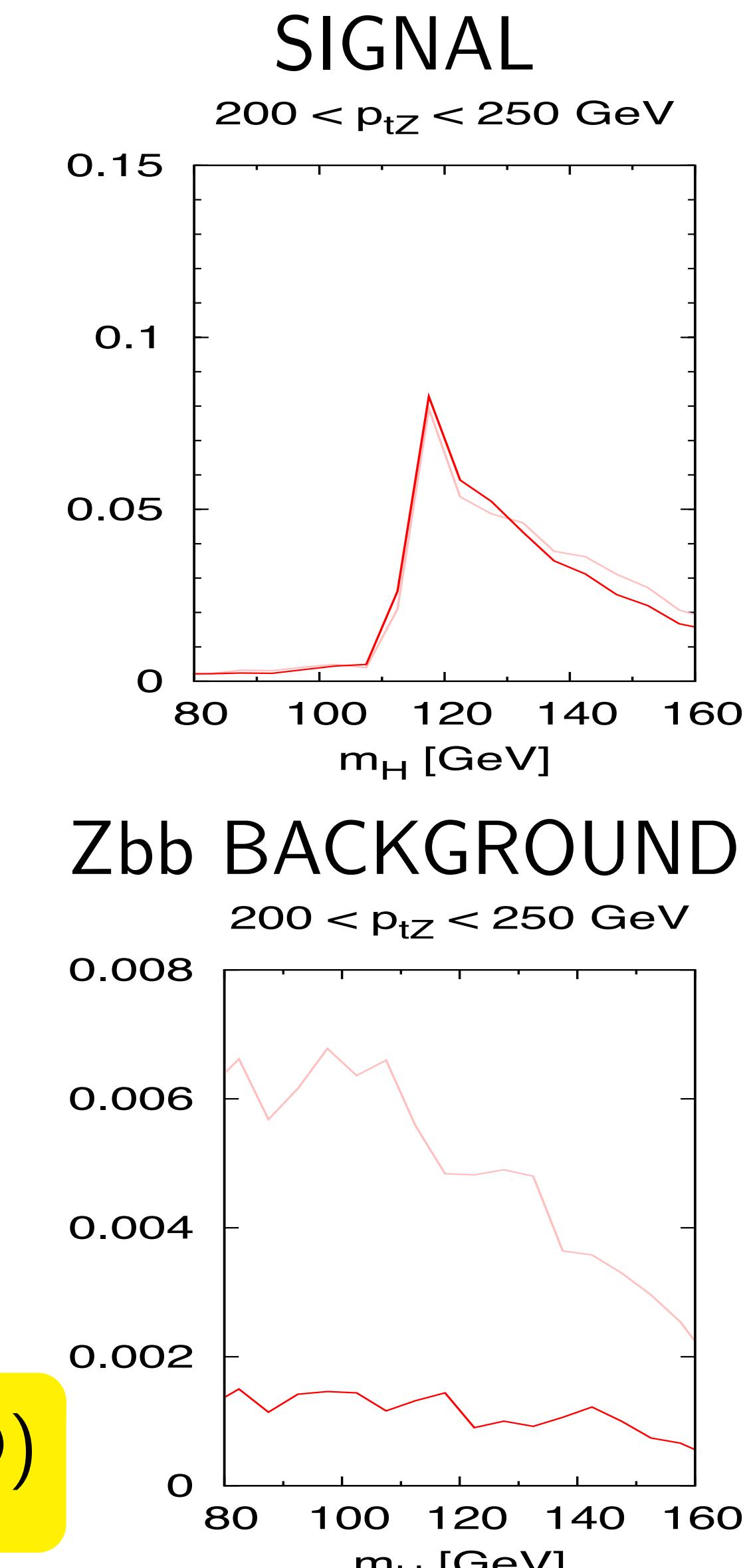
split:  $m = 139$  GeV,  $\frac{\max(m_1, m_2)}{m} = 0.37 \rightarrow$  mass drop



Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

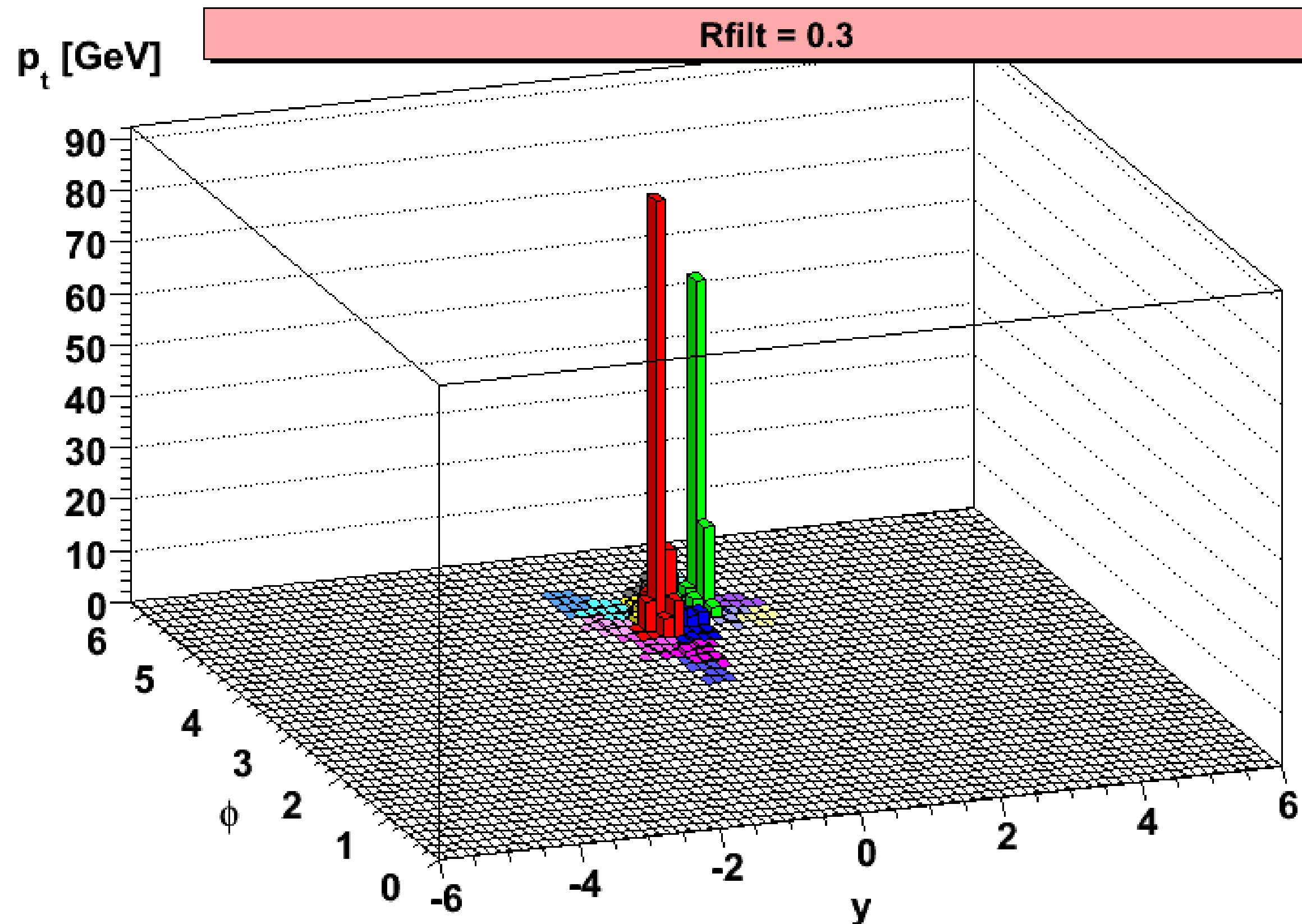


check:  $y_{12} \simeq \frac{p_{t2}}{p_{t1}} \simeq 0.7 \rightarrow$  OK + 2  $b$ -tags (anti-QCD)

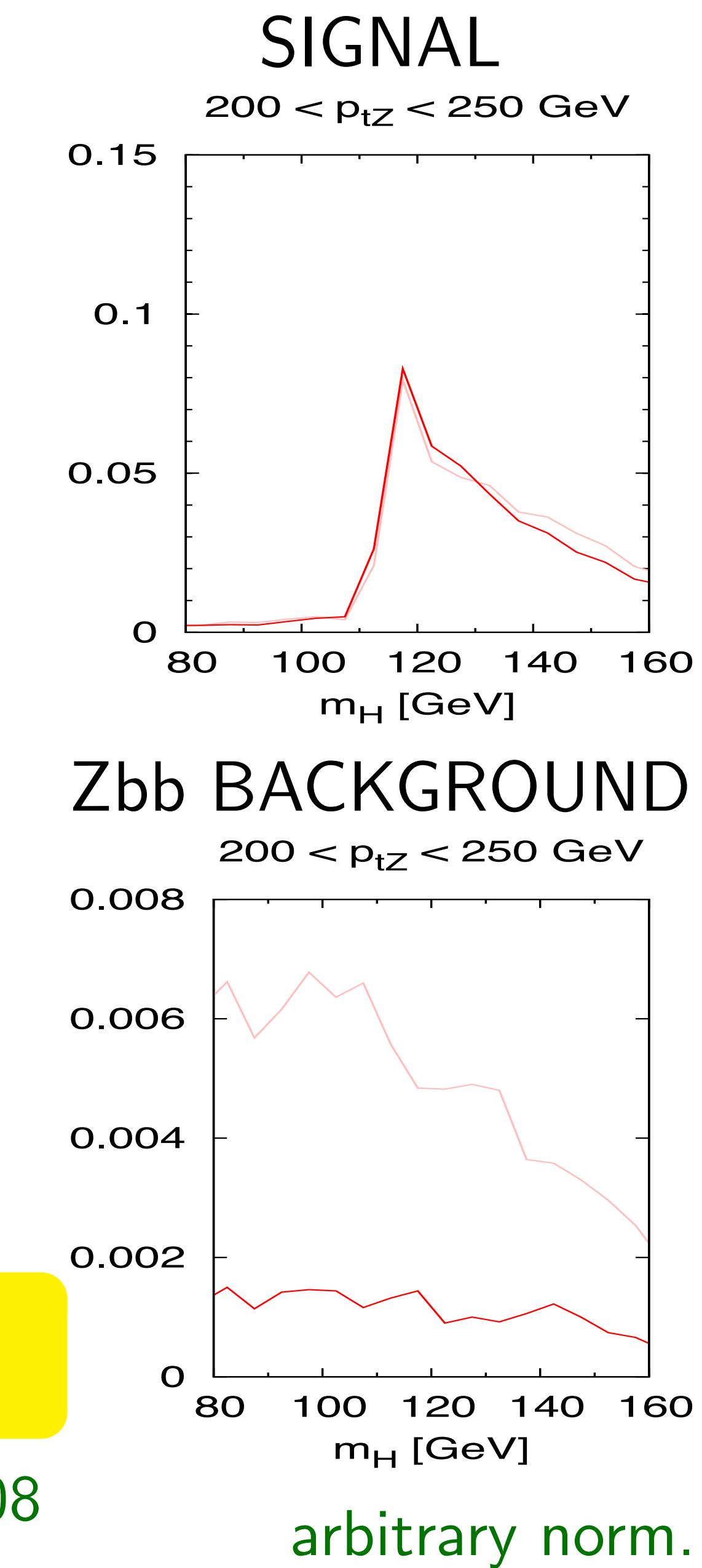


arbitrary norm.

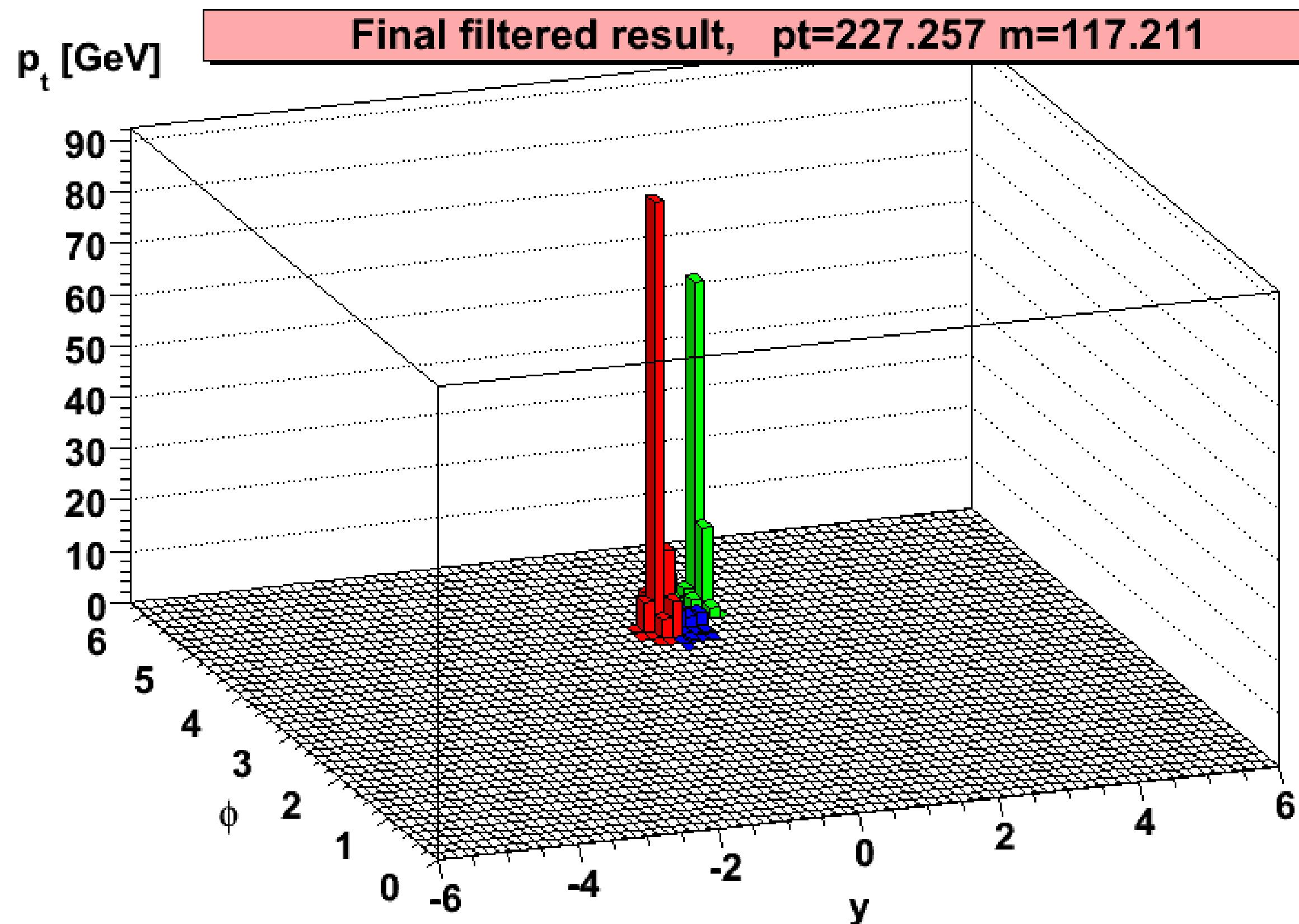
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



$R_{filt} = 0.3$  [ $= \min(\Delta R/2, 0.3)$ ]



Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



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

