Jets and their structure

Simone Marzani
Università di Genova &
INFN Sezione di Genova

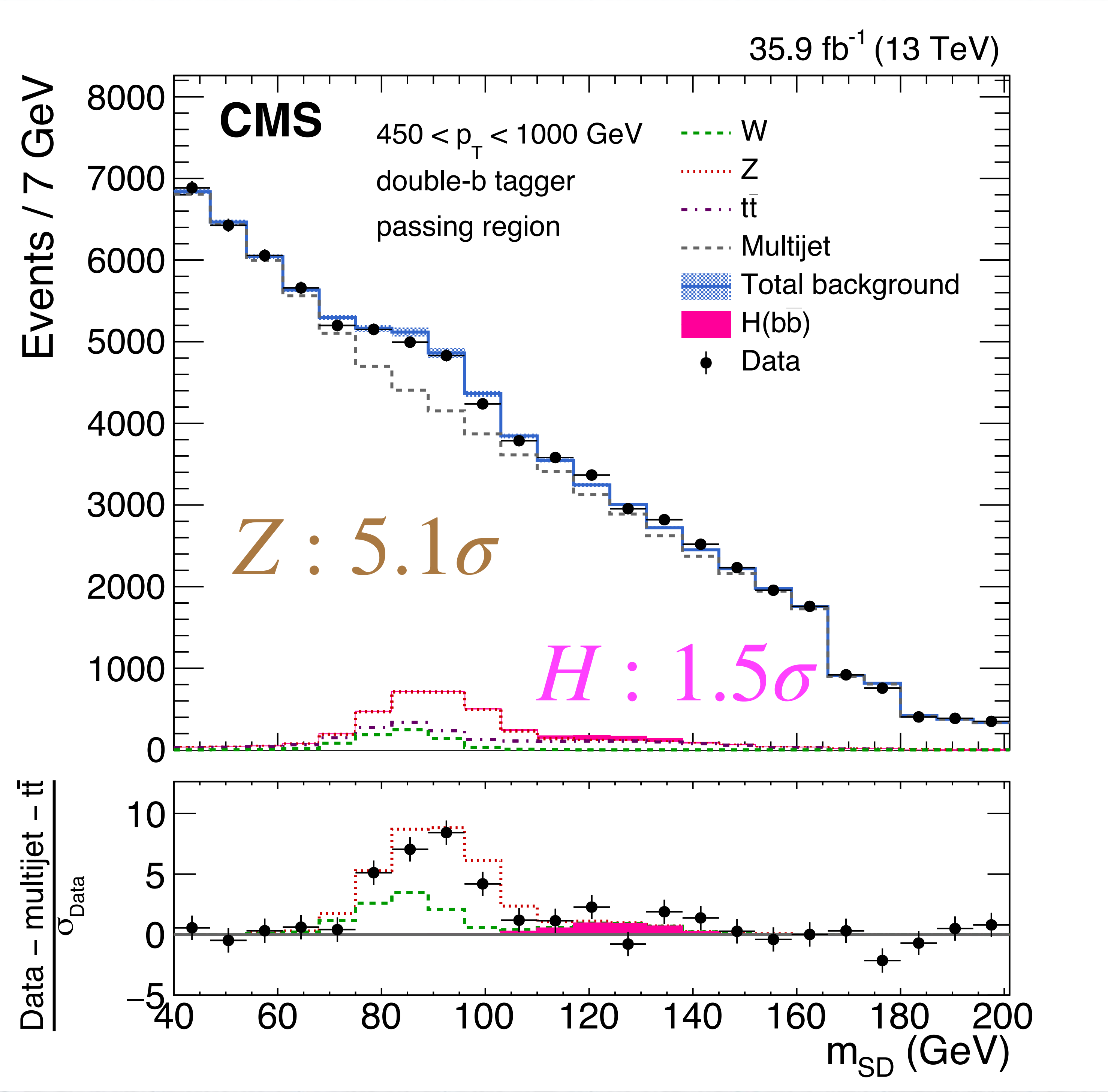
QCD Masterclass
2019

Lecture plan

* lecture 1: jets and jet algorithms
* lecture 2: calculating jet properties
* lecture 3: jet substructure
* lecture 4: more advanced topics & curiosities

Lecture 3: jet substructure

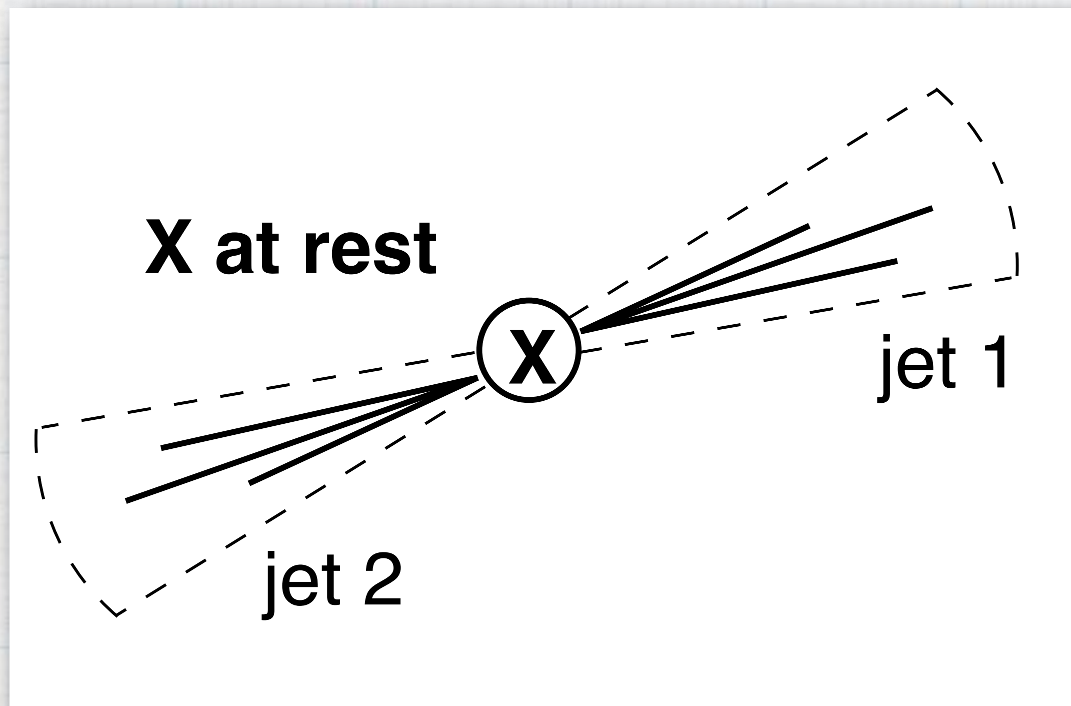
- boosted-objects physics
- grooming and tagging
- calculations for jet substructure

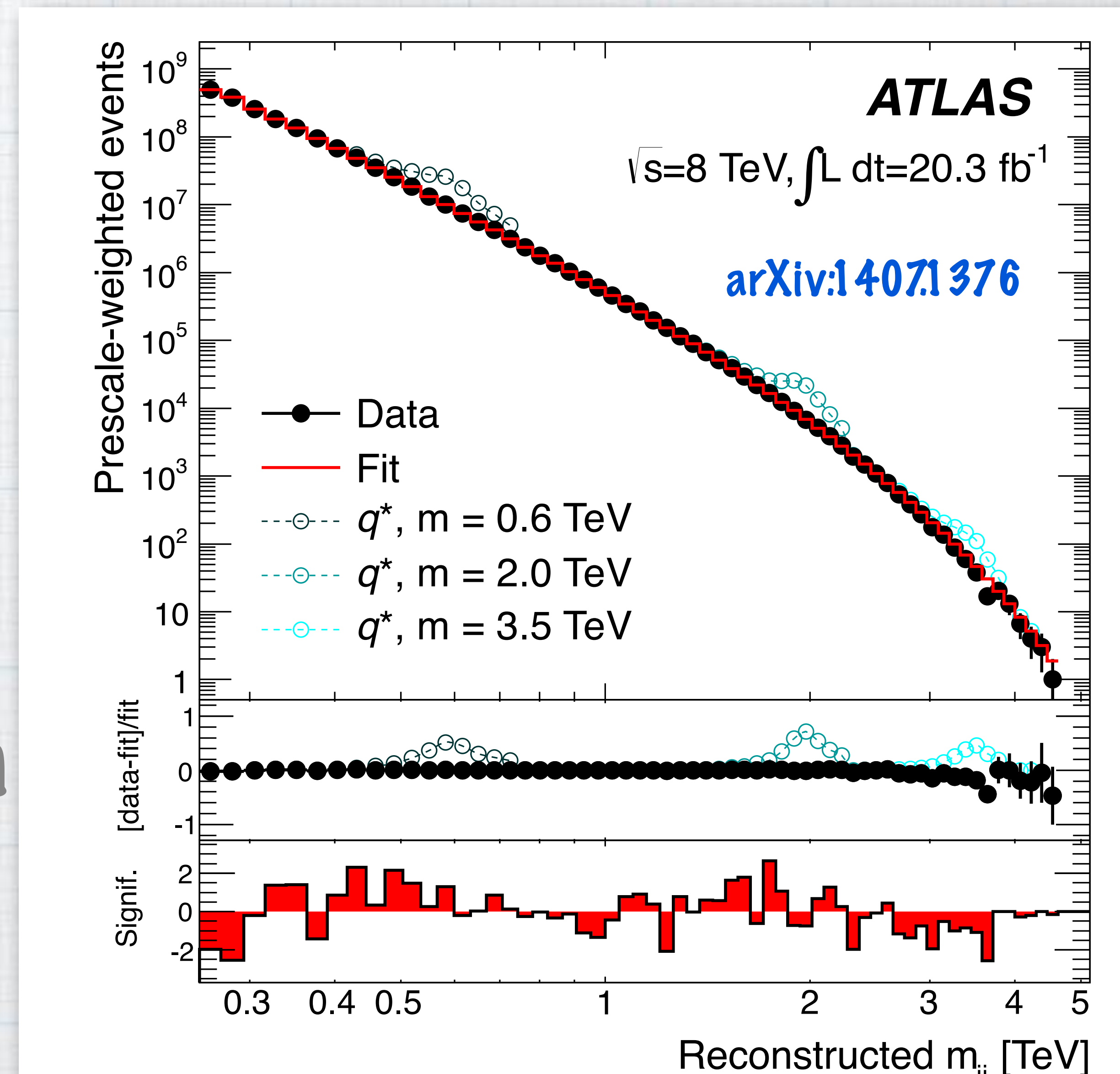


the (ambitious) target of this lecture is to understand this plot

searching for new particles (I)

- Standard analysis: the heavy particle X decays into two partons, reconstructed as two jets



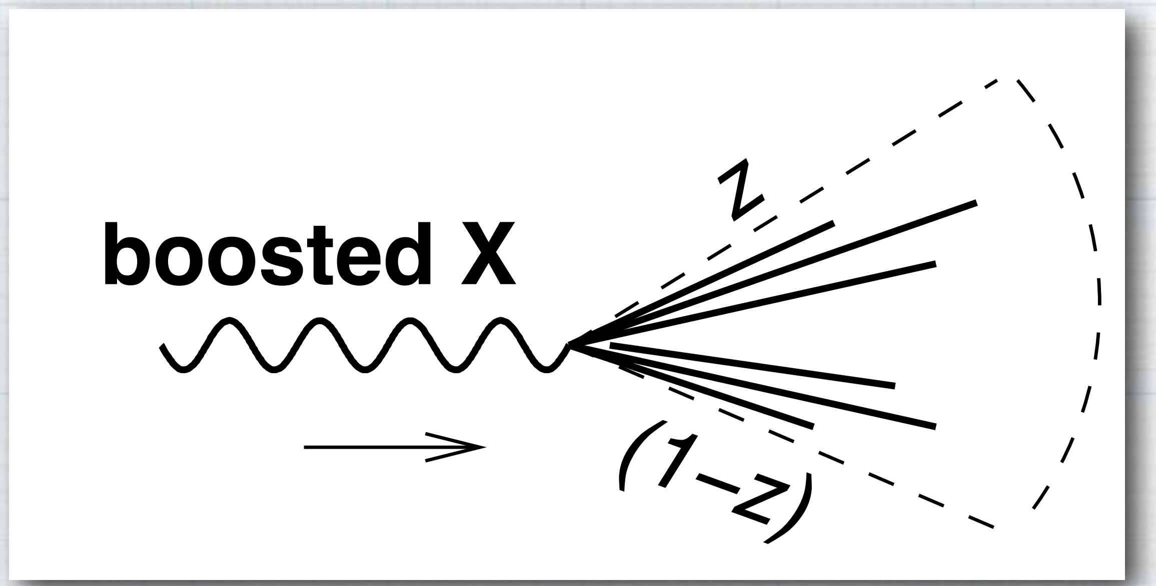


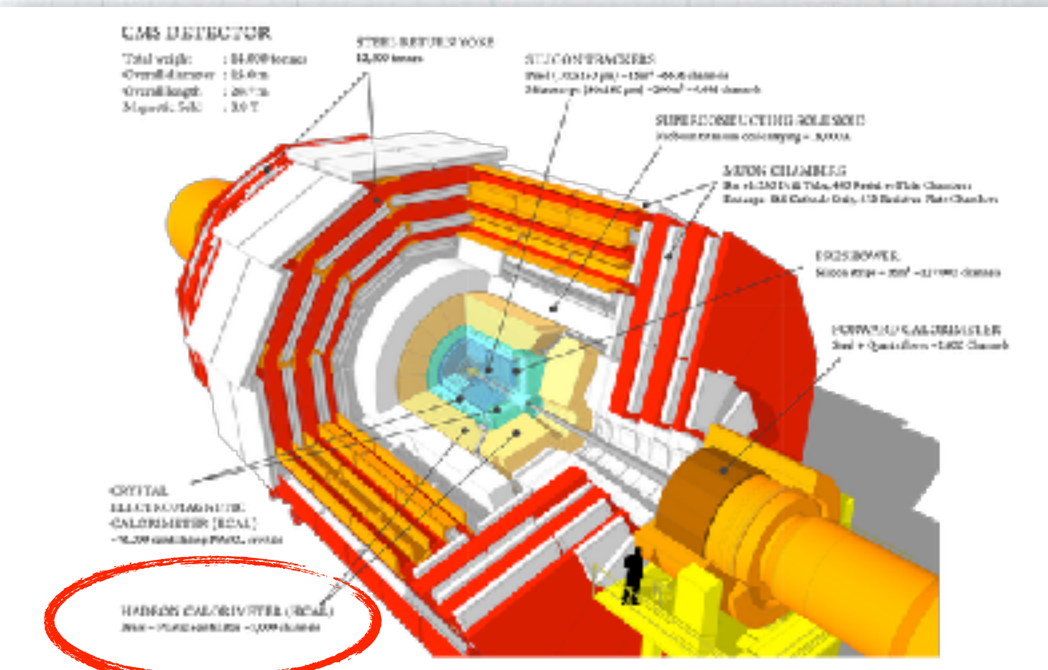
- Look for bumps in the dijet invariant mass distribution

- What about EW-scale particles at the LHC ?

searching for new particles (II)

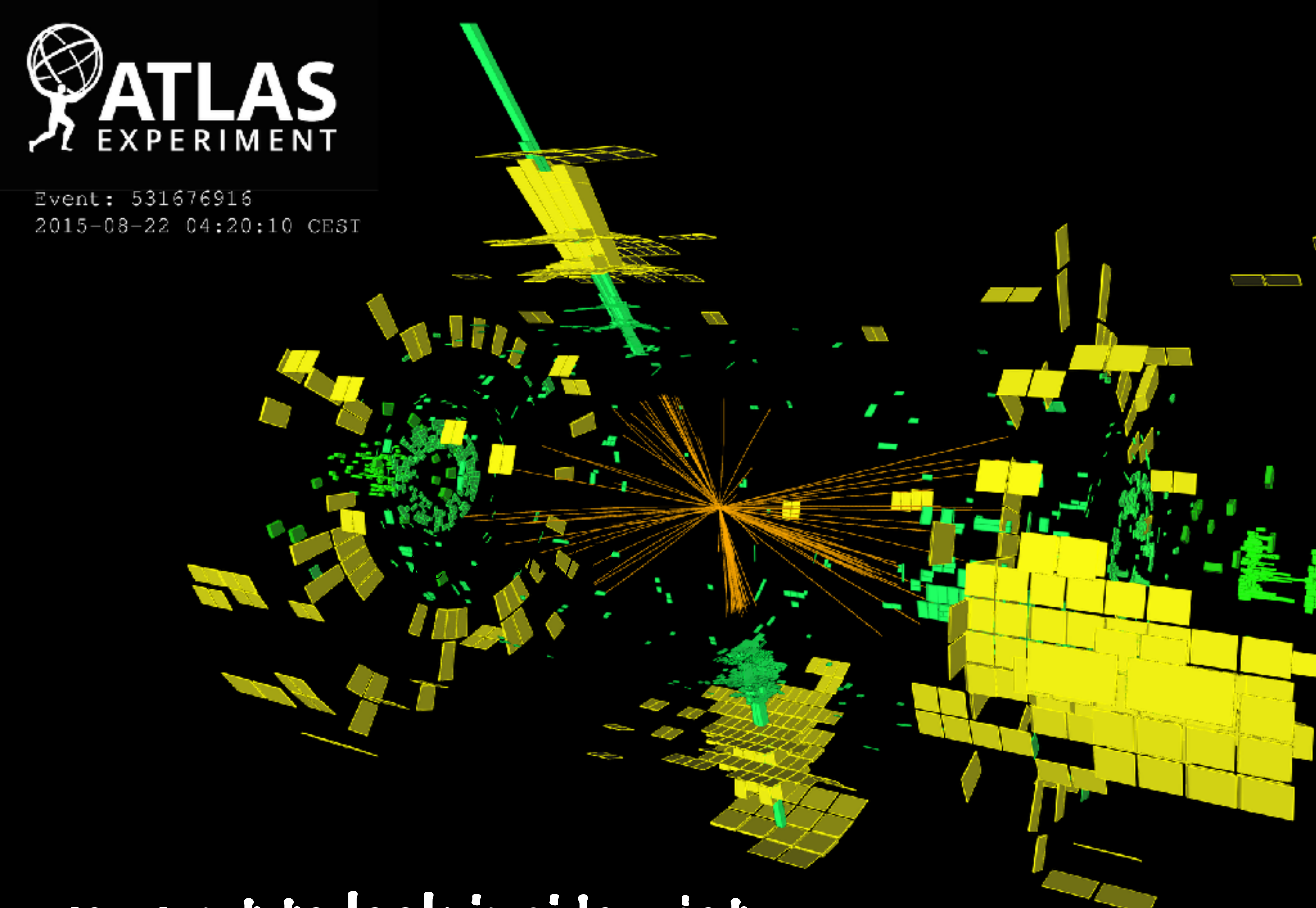
* LHC energy (10^4 GeV) $\gg$ electro-weak scale (10^2 GeV)
* EW-scale particles (new physics, Z/W/H/top) are abundantly produced with a large boost





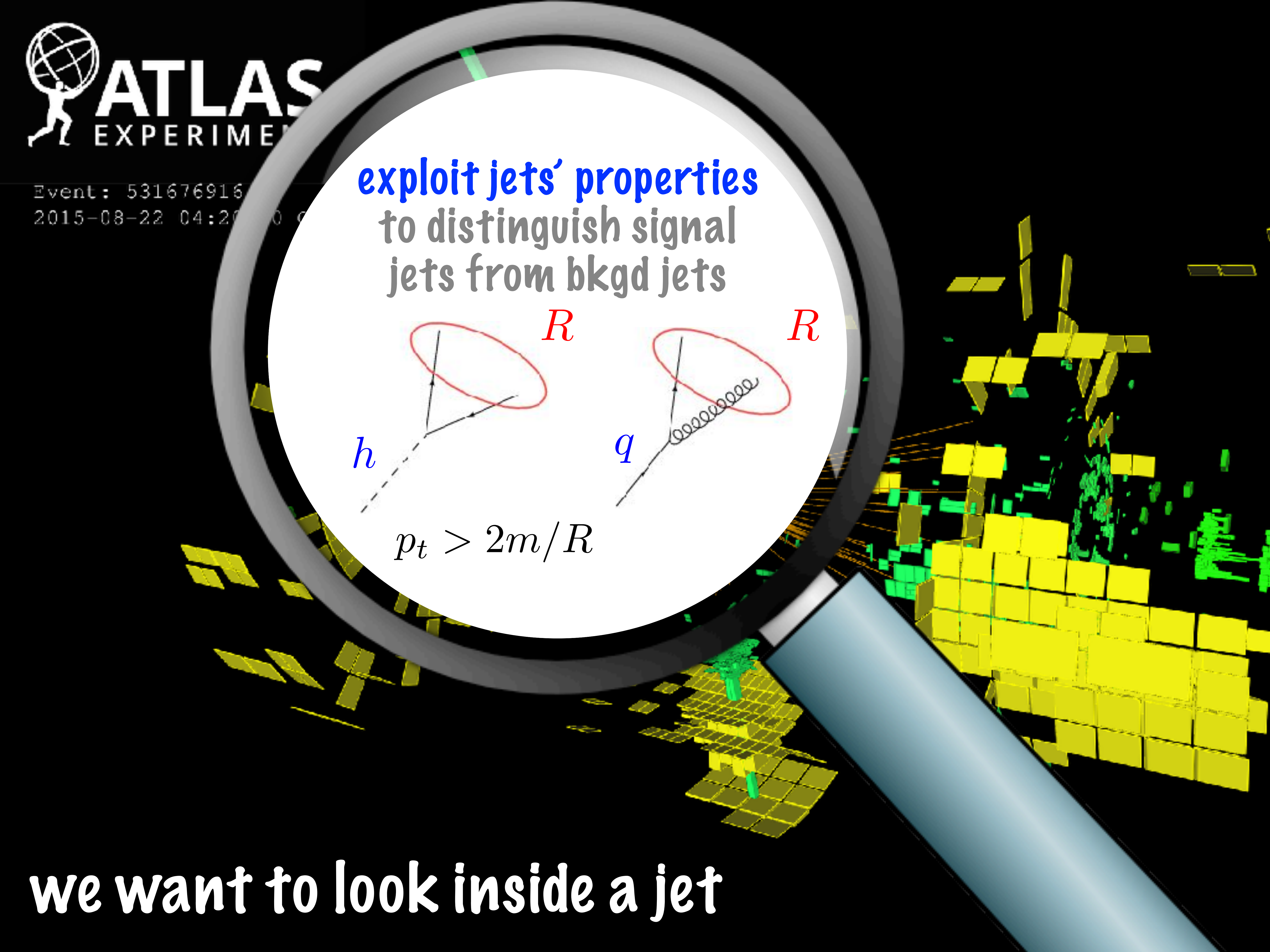
source CMS

* their decay-products are then collimated
* if they decay into hadrons, we end up with localised deposition of energy in the hadronic calorimeter: a jet



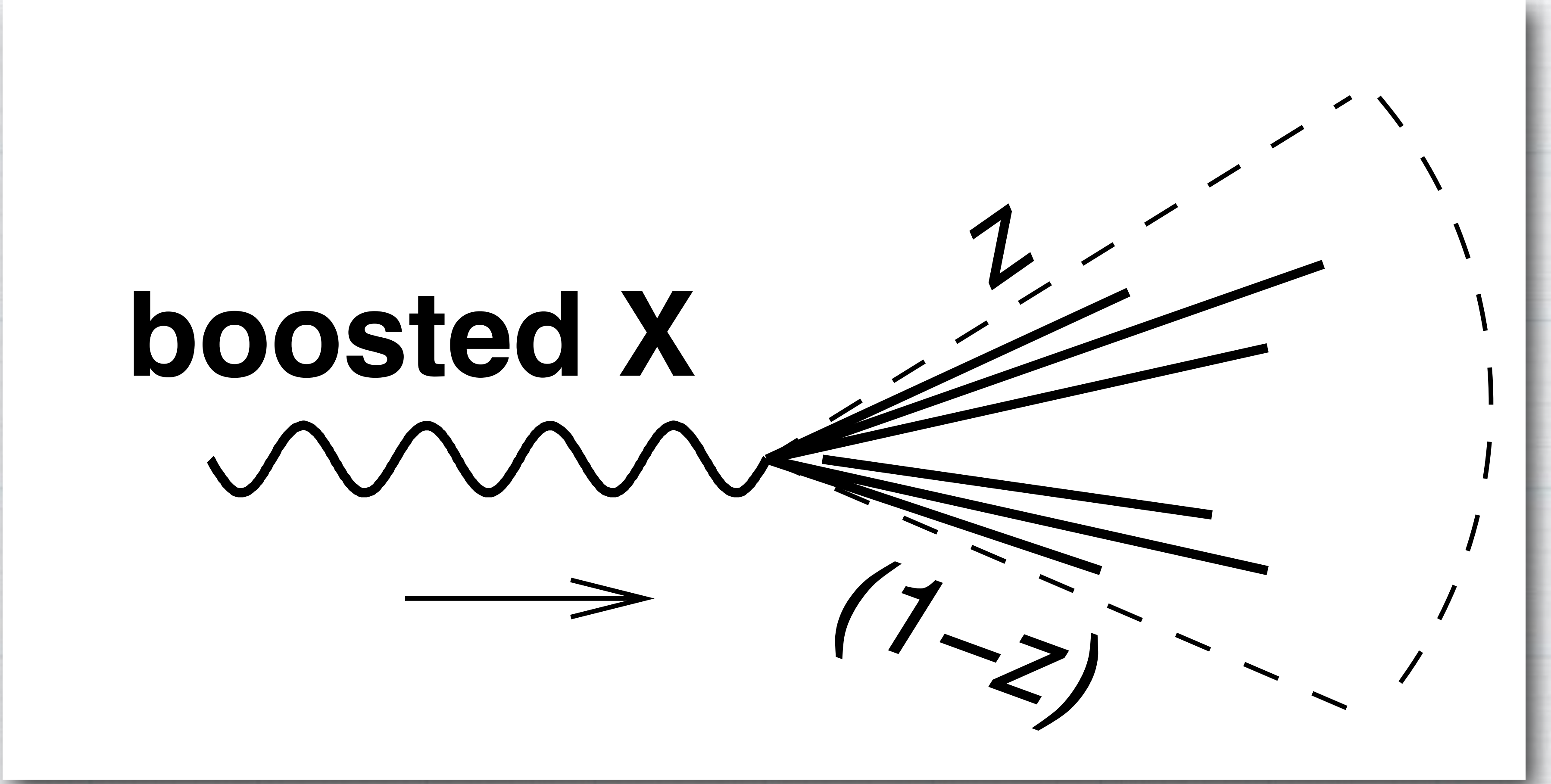
we want to look inside a jet


ATLAS
EXPERIME
Event: 531676916
2015-08-22 04:2
0 CEST
p_t
we want to look inside a jet


exploit jets' properties
to distinguish signal
jets from bkgd jets
R
R
h
q
$p_t > 2m/R$
we want to look inside a jet

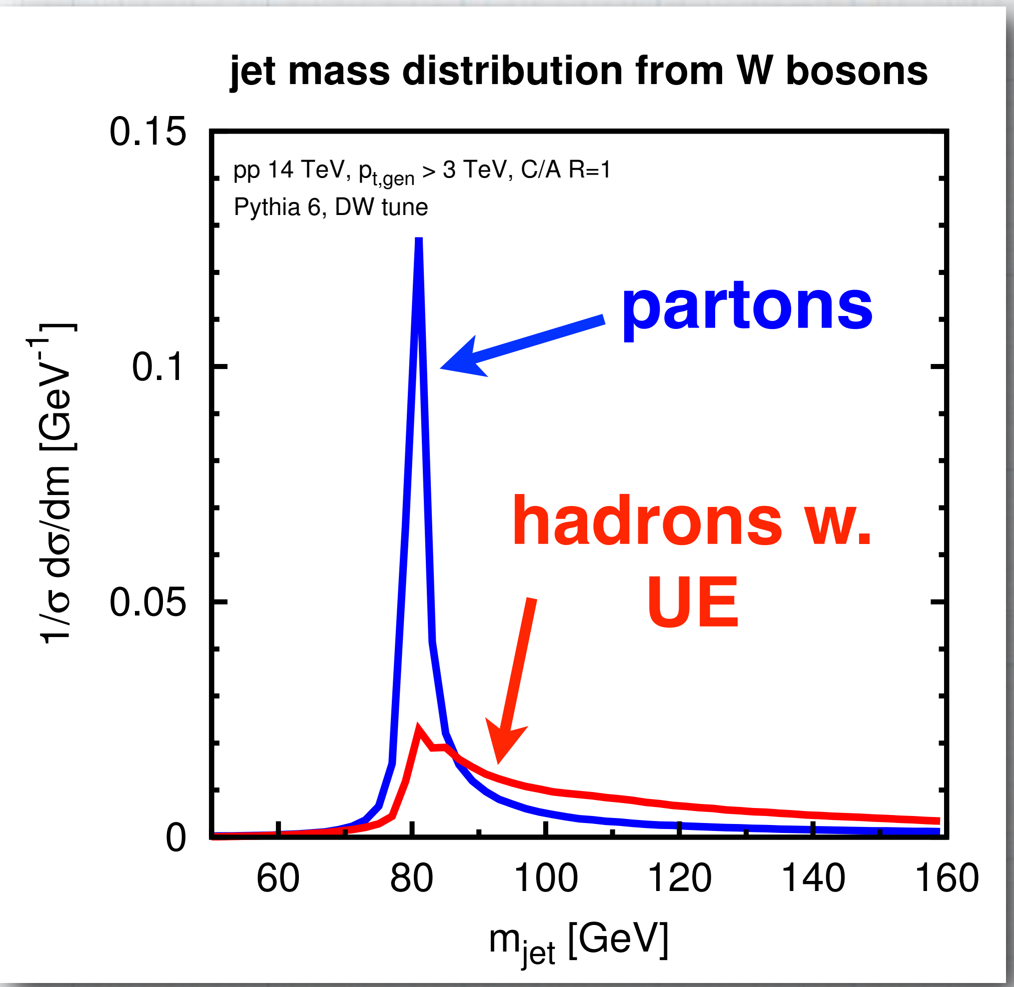
signal-jet mass

* first jet-observable that comes to mind
* signal jets should have a mass distribution peaked near the resonance



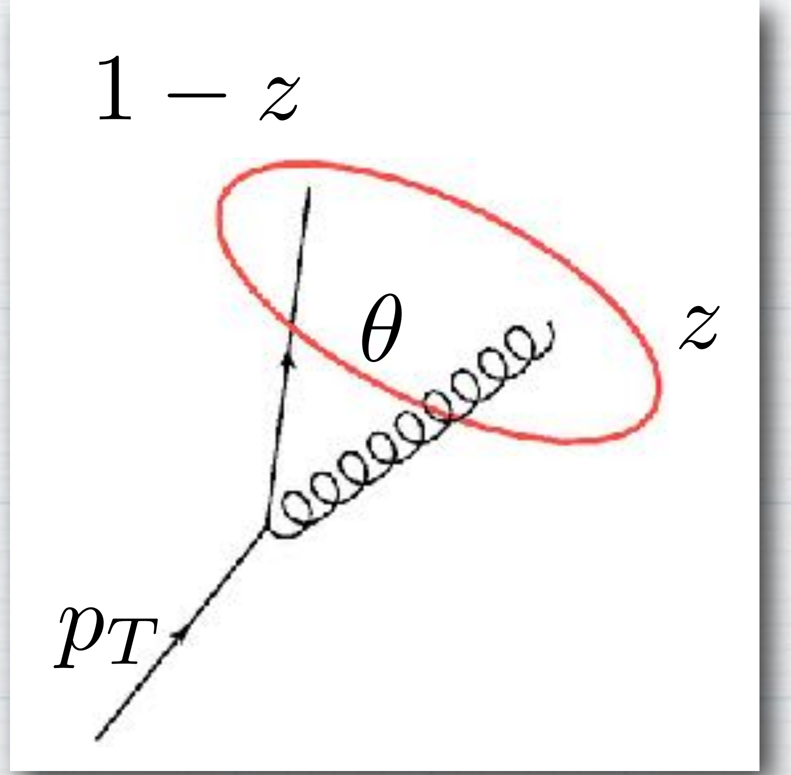
signal-jet mass

* first jet-observable that comes to mind
* signal jets should have a mass distribution peaked near the resonance
* however, that's a simple partonic picture
* perturbative and non-pert. emissions from the qqb pair broadens and shift the peak
* underlying event and pile-up typically enhance the jet mass



QCD-jet mass

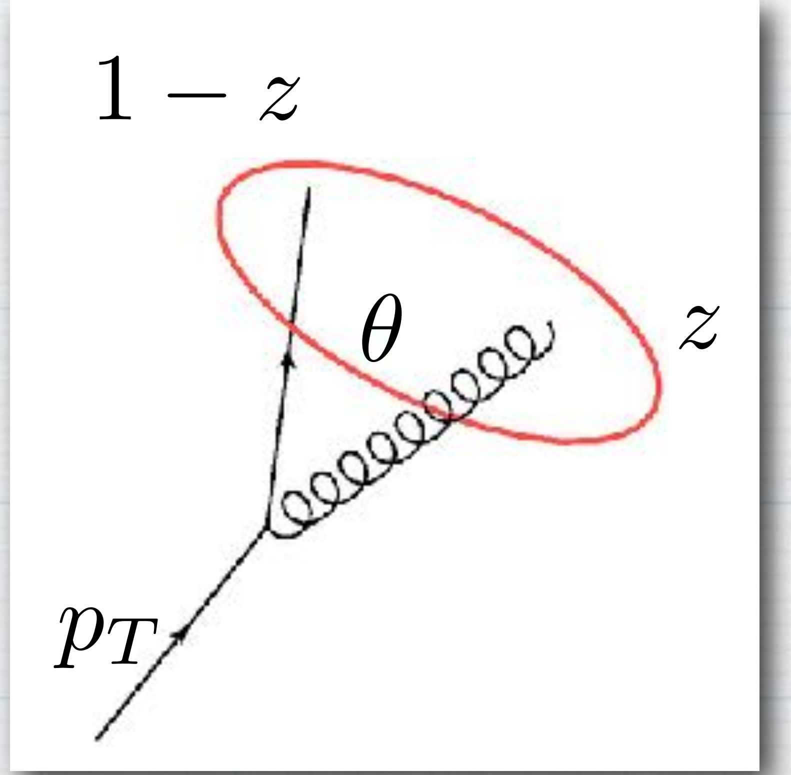
* first jet-observable that comes to mind
* background (QCD) jets acquire mass through showering



$$m^2 = 2p_q \cdot p_g \simeq z(1-z)\theta^2 p_T^2$$

QCD-jet mass

* first jet-observable that comes to mind
* background (QCD) jets acquire mass through showering

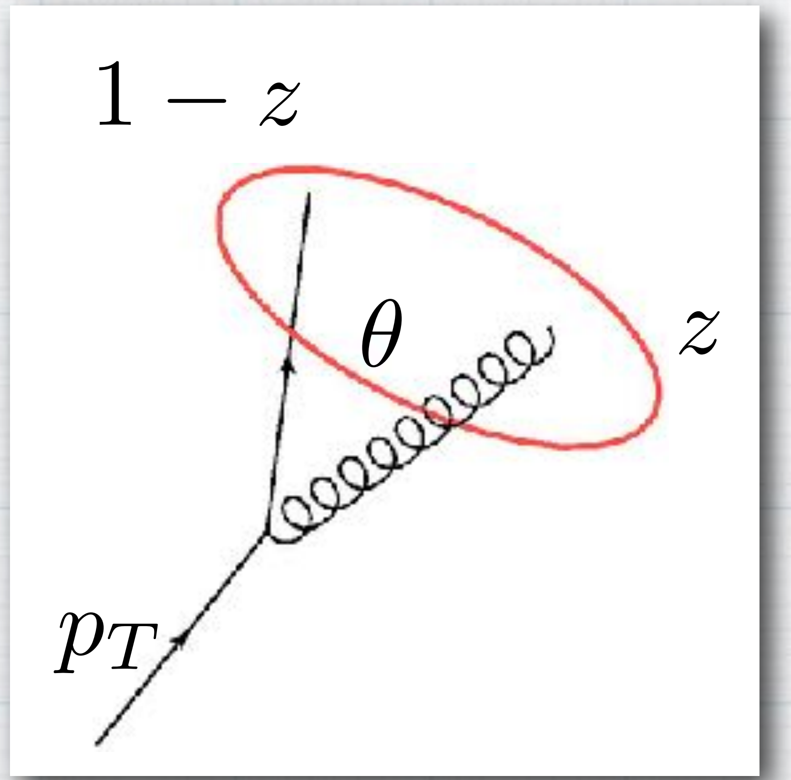


$$m^2 = 2p_q \cdot p_g \simeq z(1-z)\theta^2 p_T^2$$

$$\langle m^2 \rangle \simeq \frac{\alpha_s}{2\pi} p_T^2 \int_0^{R^2} \frac{d\theta^2}{\theta^2} \int_0^1 dz z(1-z)\theta^2 P_{gq}(z)$$

QCD-jet mass

* first jet-observable that comes to mind
* background (QCD) jets acquire mass through showering



$$m^2 = 2p_q \cdot p_g \simeq z(1-z)\theta^2 p_T^2$$

$$\langle m^2 \rangle \simeq \frac{\alpha_s}{2\pi} p_T^2 \int_0^{R^2} \frac{d\theta^2}{\theta^2} \int_0^1 dz z(1-z)\theta^2 P_{gq}(z)$$

$$= \frac{\alpha_s C_F}{\pi} p_T^2 R^2 \int_0^1 dz z(1-z) \frac{2-2z+z^2}{2z}$$

mass grows with p_T

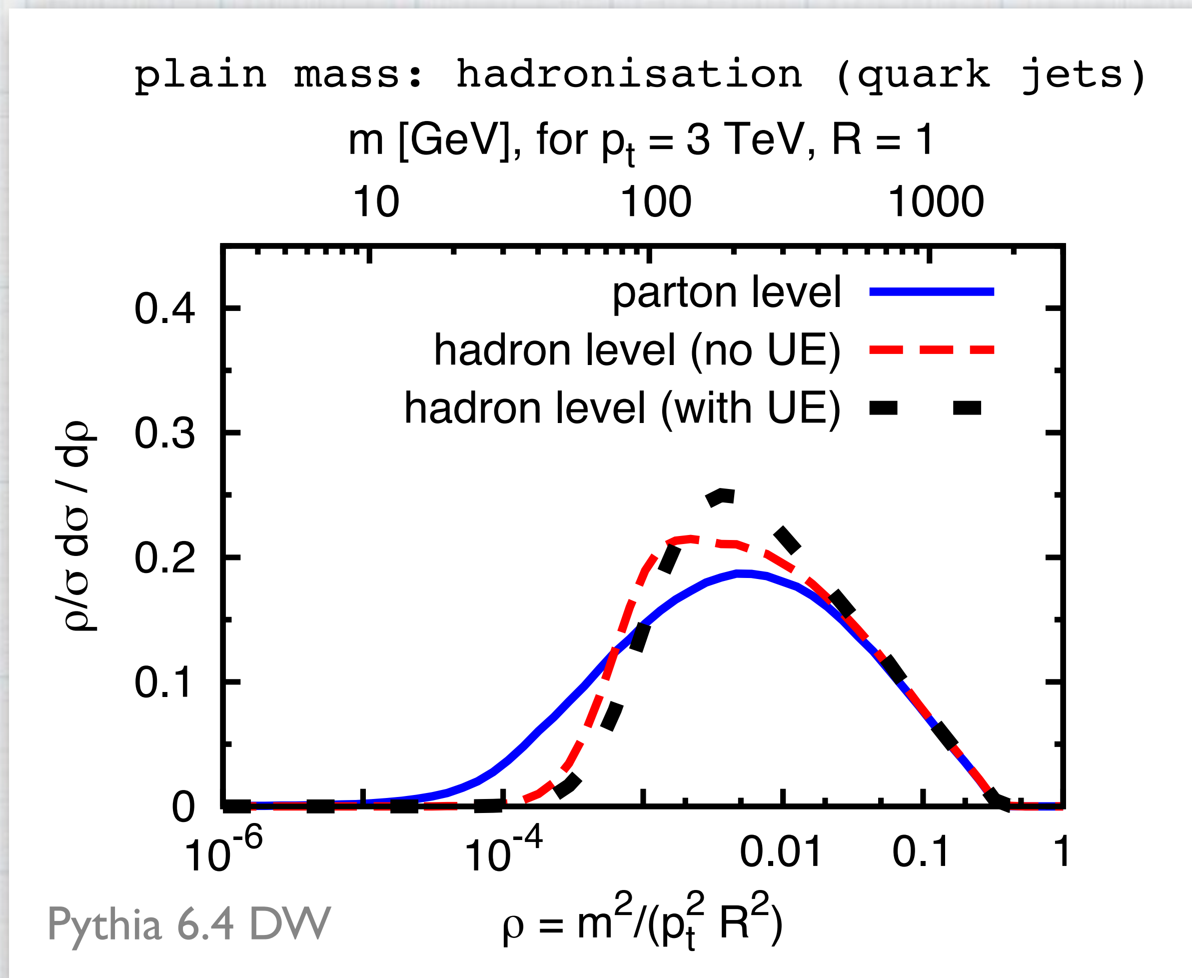
$= 3/8$

homework 5

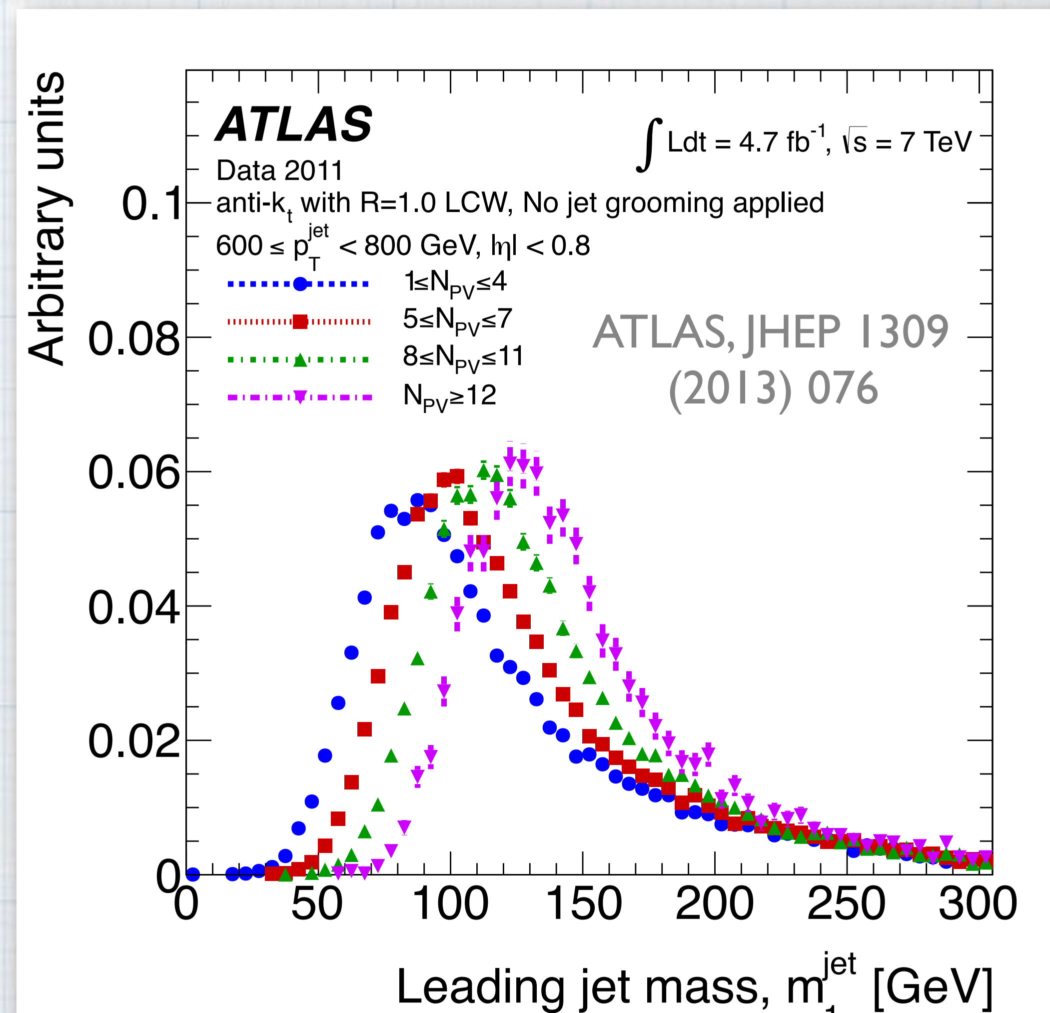
* Gluon splitting into bottom quarks $g \rightarrow bb$ is important for $H \rightarrow bb$ studies. What's its average mass? (take m_b=0)

QCD-jet mass: NP effects

* first jet-observable that comes to mind
* background (QCD) jets receive important non-pert contributions



hadronisation and UE



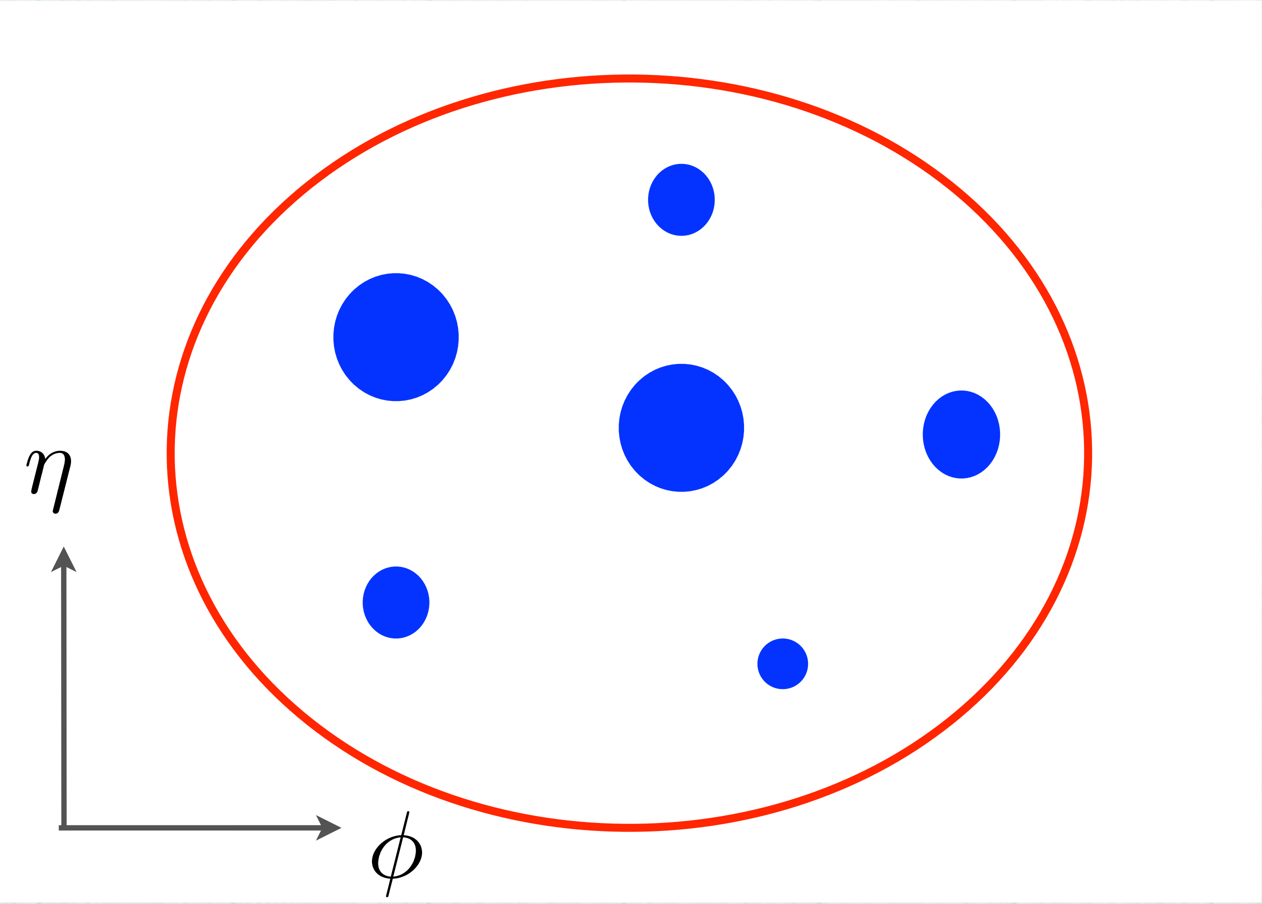
pile-up (data!)

beyond the mass: substructure

* need to go beyond the mass and exploit jet substructure : grooming and tagging:
  * clean the jets up by removing soft radiation
  * identify the features of hard decays and cut on them

beyond the mass: substructure

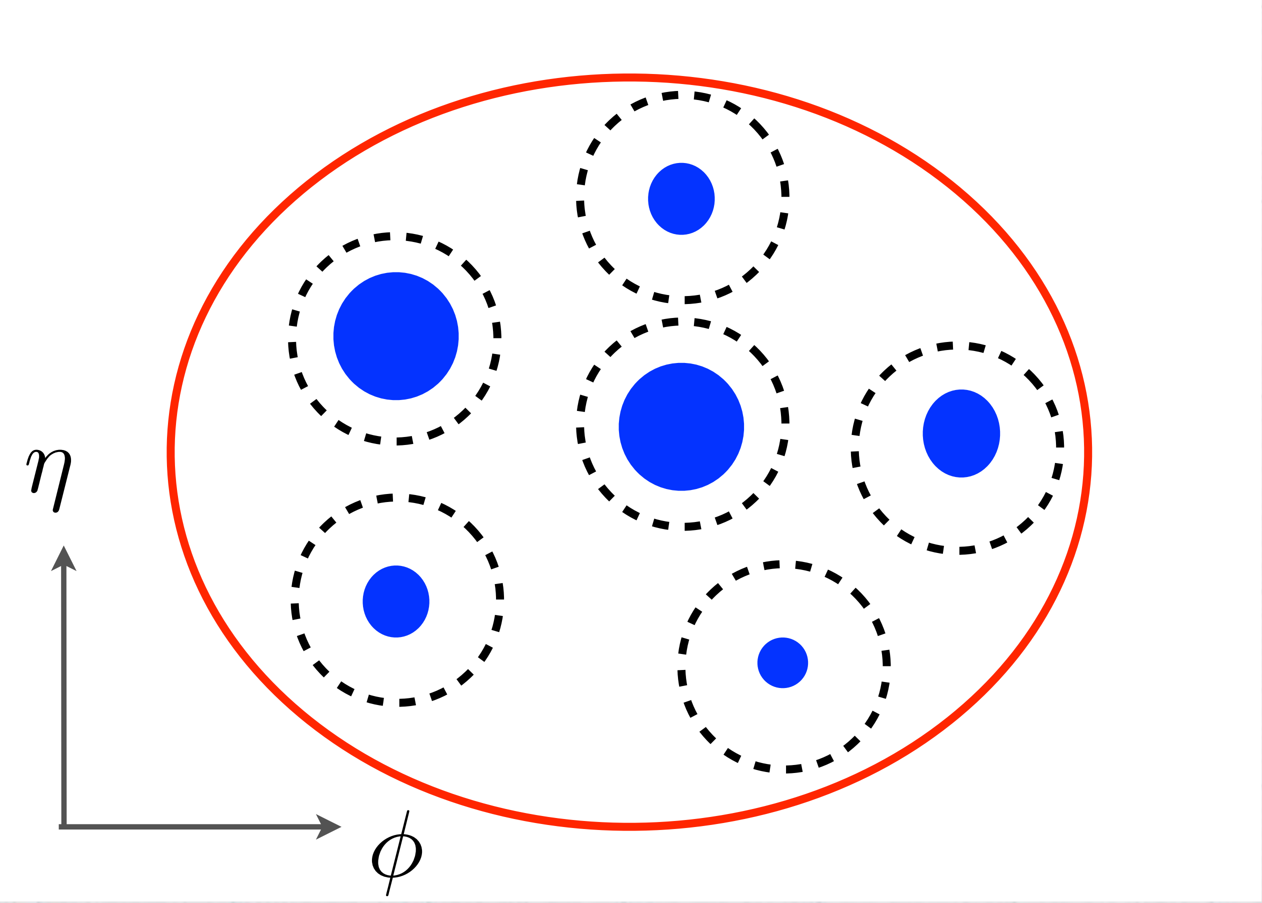
* need to go beyond the mass and exploit jet substructure : grooming and tagging:
  * clean the jets up by removing soft radiation
  * identify the features of hard decays and cut on them



core-idea for grooming:

beyond the mass: substructure

* need to go beyond the mass and exploit jet substructure : **grooming** and **tagging**:
  * clean the jets up by removing soft radiation
  * identify the features of hard decays and cut on them

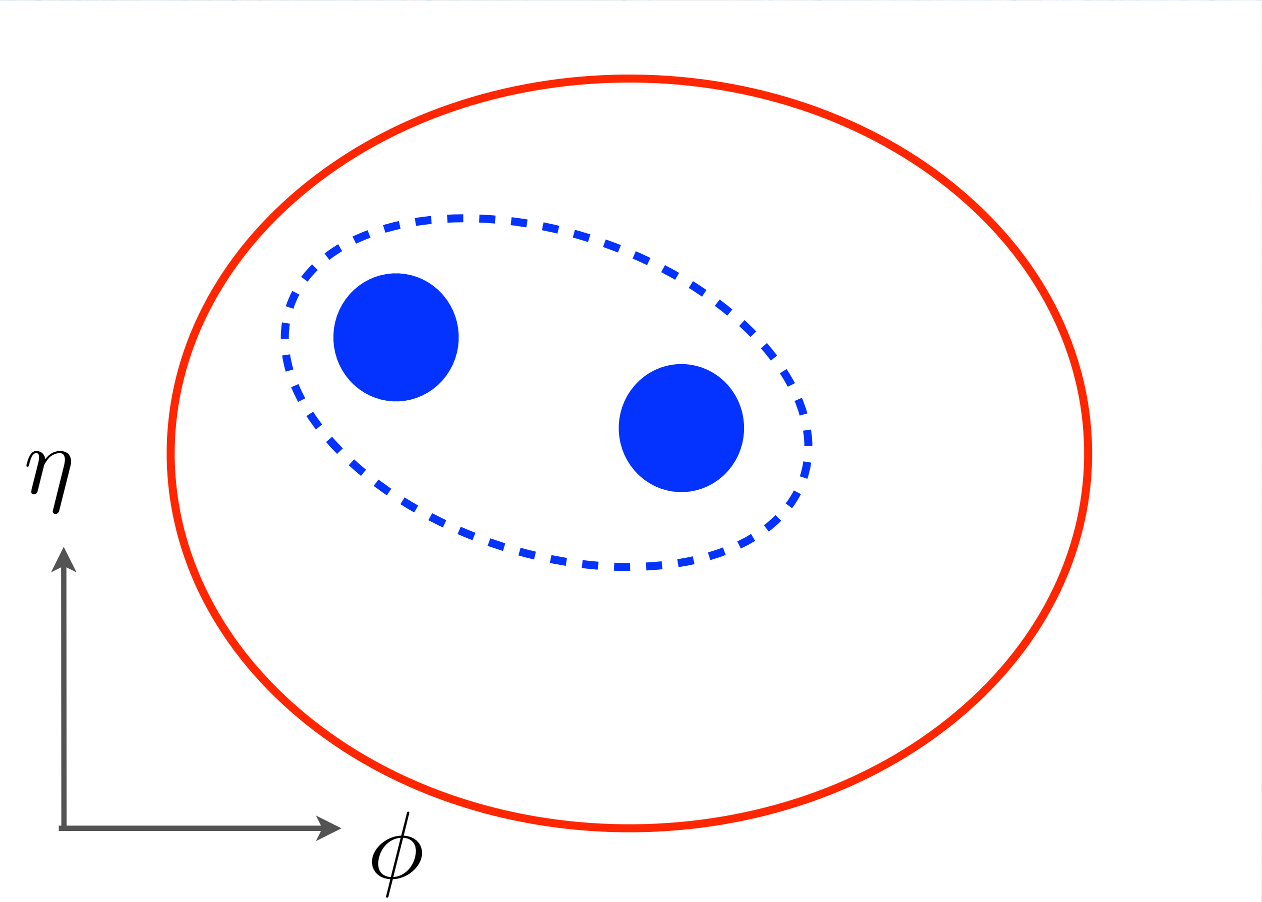


core-idea for grooming:

* identify the "right" angular scale

beyond the mass: substructure

* need to go beyond the mass and exploit jet substructure : grooming and tagging:
    * clean the jets up by removing soft radiation
    * identify the features of hard decays and cut on them



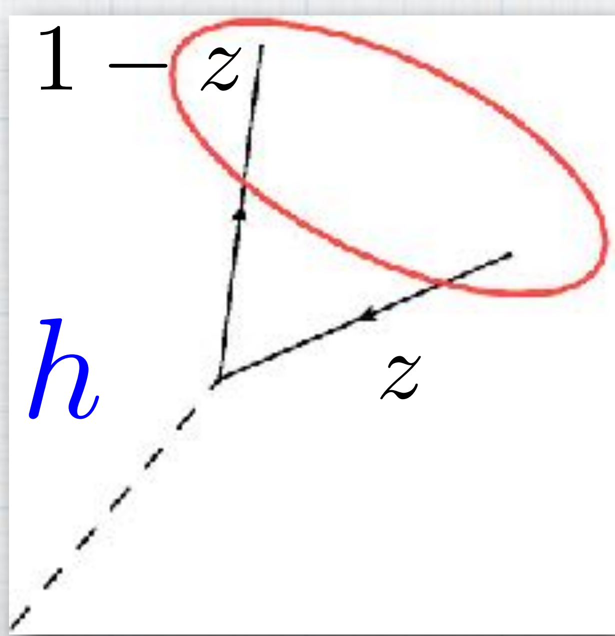
core-idea for grooming:

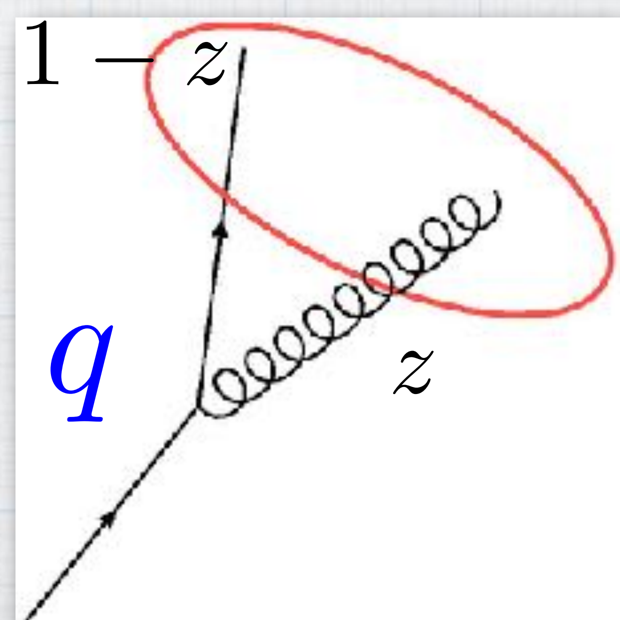
* identify the "right" angular scale
* throw away what is soft & large angle
* left with a groomed jet

beyond the mass: substructure

* need to go beyond the mass and exploit jet substructure : grooming and tagging:
    * clean the jets up by removing soft radiation
    * identify the features of hard decays and cut on them

core-idea for 2-body tagging:





beyond the mass: substructure

* need to go beyond the mass and exploit jet substructure : grooming and tagging:

  * clean the jets up by removing soft radiation

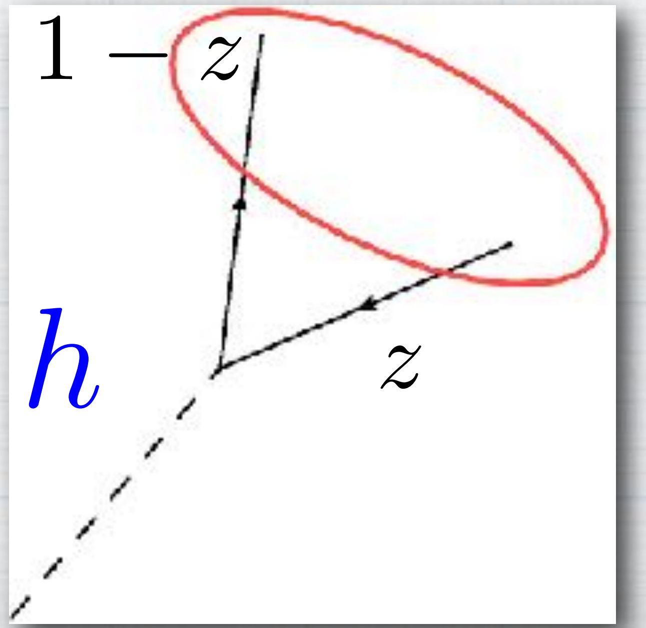
  * identify the features of hard decays and cut on them

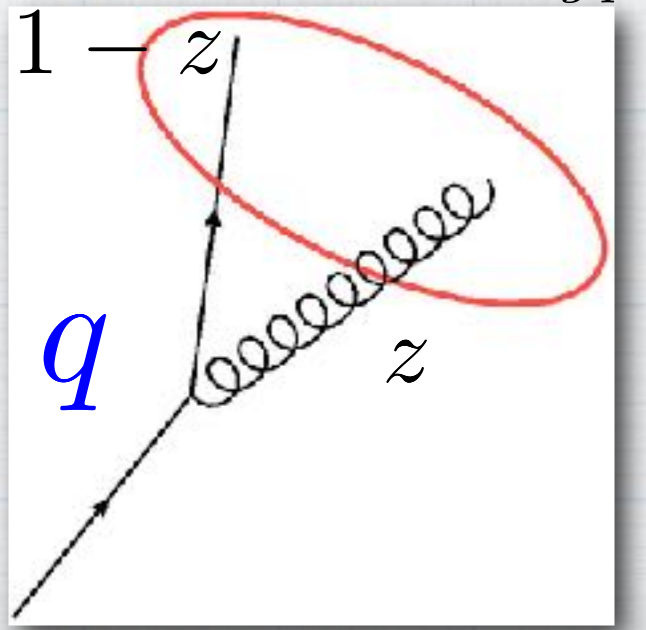
core-idea for 2-body tagging: $\min(z, 1-z) > z_{\text{cut}}$

$$P_{gq} = C_F \frac{1+(1-z)^2}{z}$$

$P_{h \to q\bar{q}} = 1$

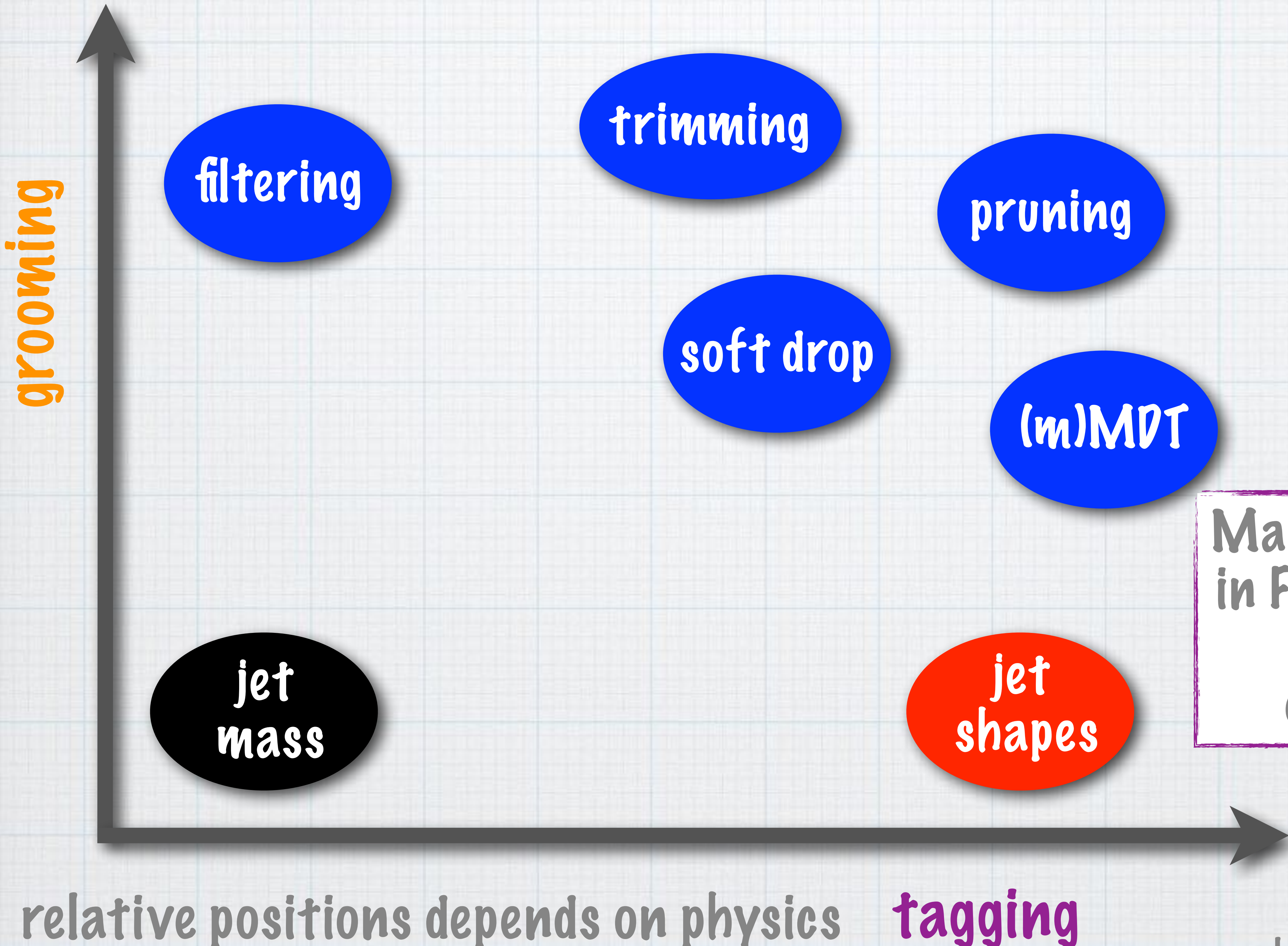
symmetric sharing of the energy





asymmetric sharing of the energy

grooming & tagging landscape

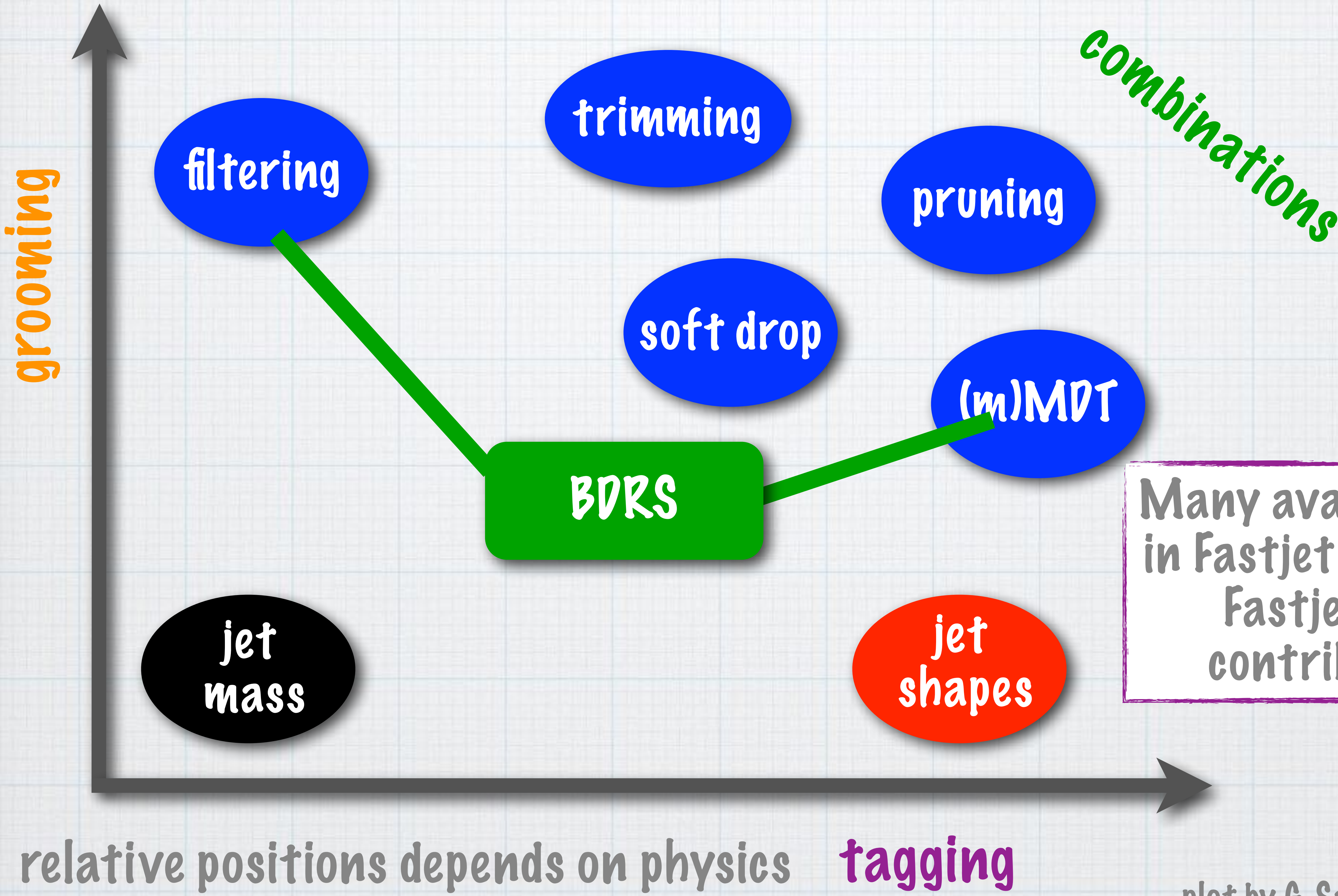


Many available in Fastjet or as Fastjet contribs

relative positions depends on physics context, kinematics, etc.

plot by G. Salam

grooming & tagging landscape

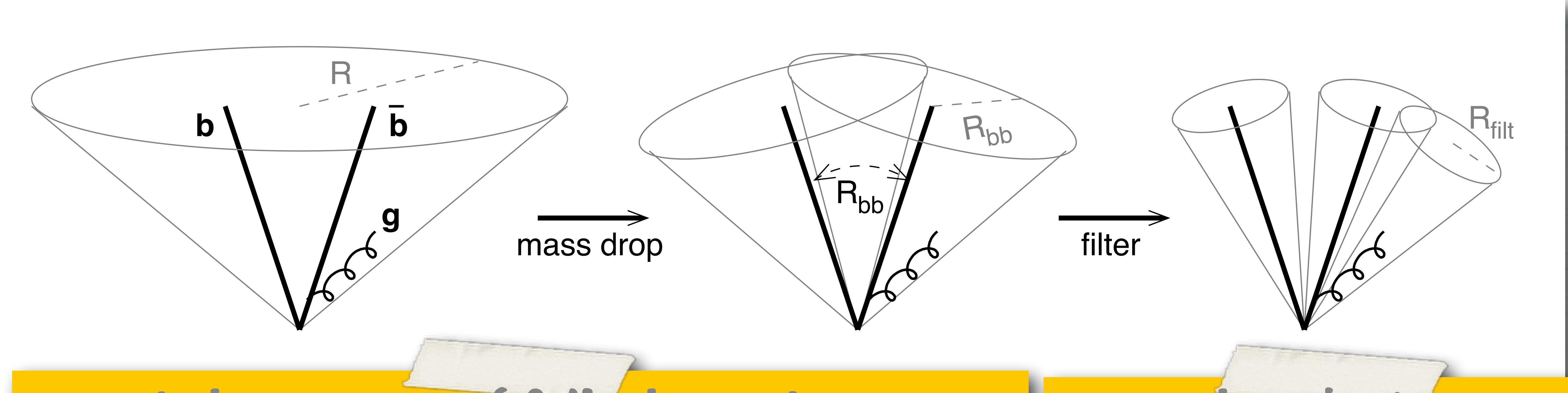


Many available in Fastjet or as Fastjet contribs

relative positions depends on physics context, kinematics, etc.

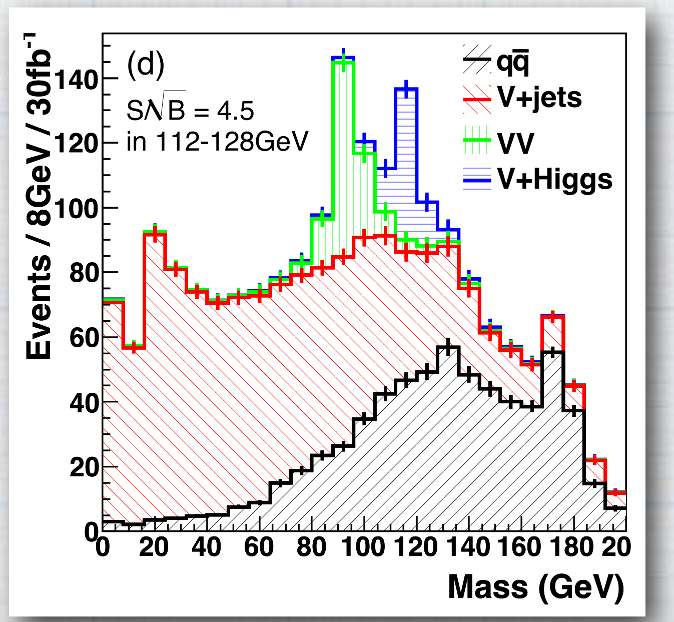
plot by G. Salam

BDRS method for H→bb



- undo last stage of C/A clustering.
- if there's symmetric sharing of energy & significant mass drop, then tag the jet.
- otherwise iterate

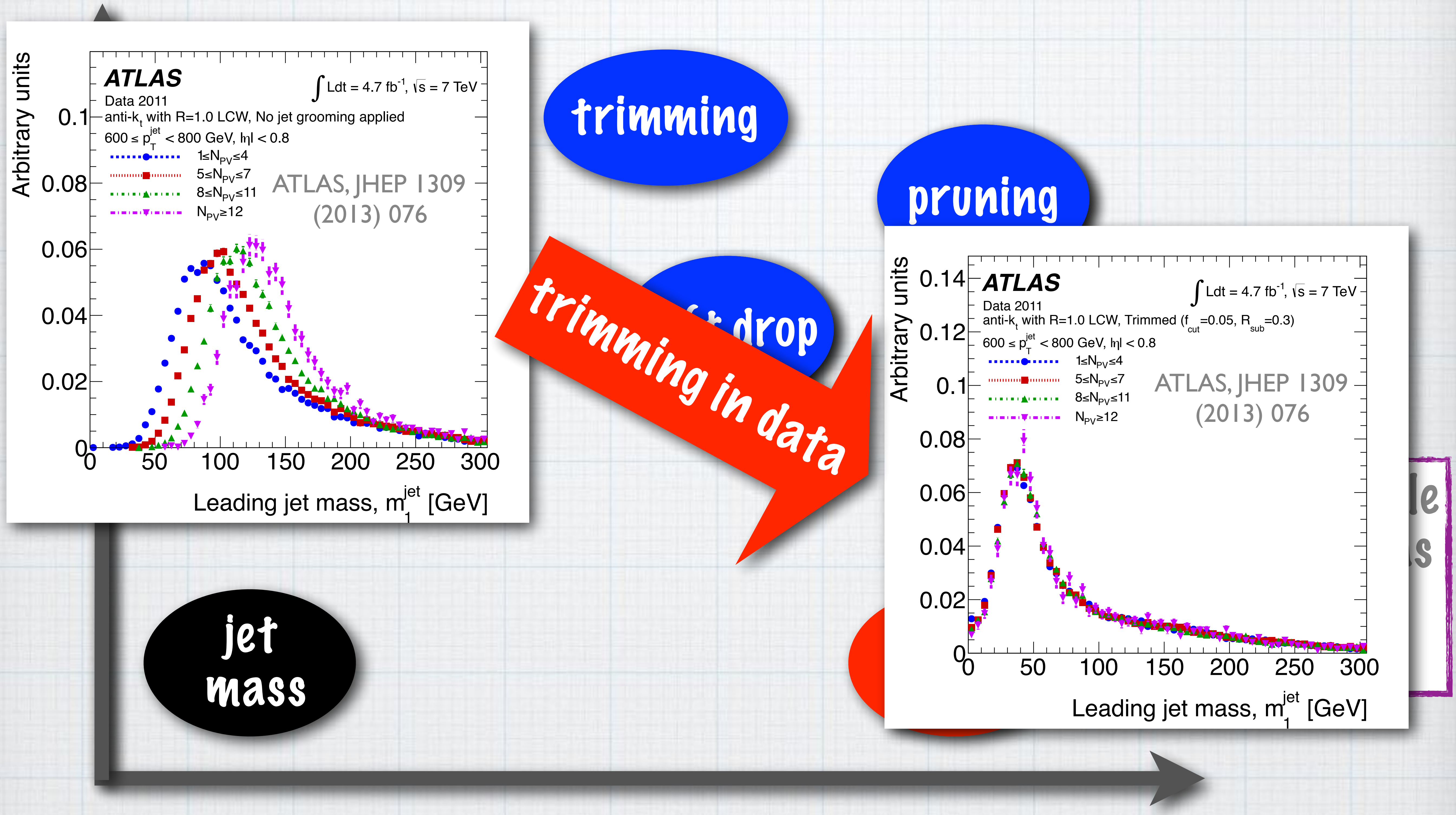
- resolve the jet on a smaller radius
- keep the 3 hardest subjets



* this study resurrected an "impossible" channel
* still very difficult at the LHC !
* it sparked interest in this field !

Butterworth, Davison, Rubin and Salam (2008)

grooming & tagging landscape



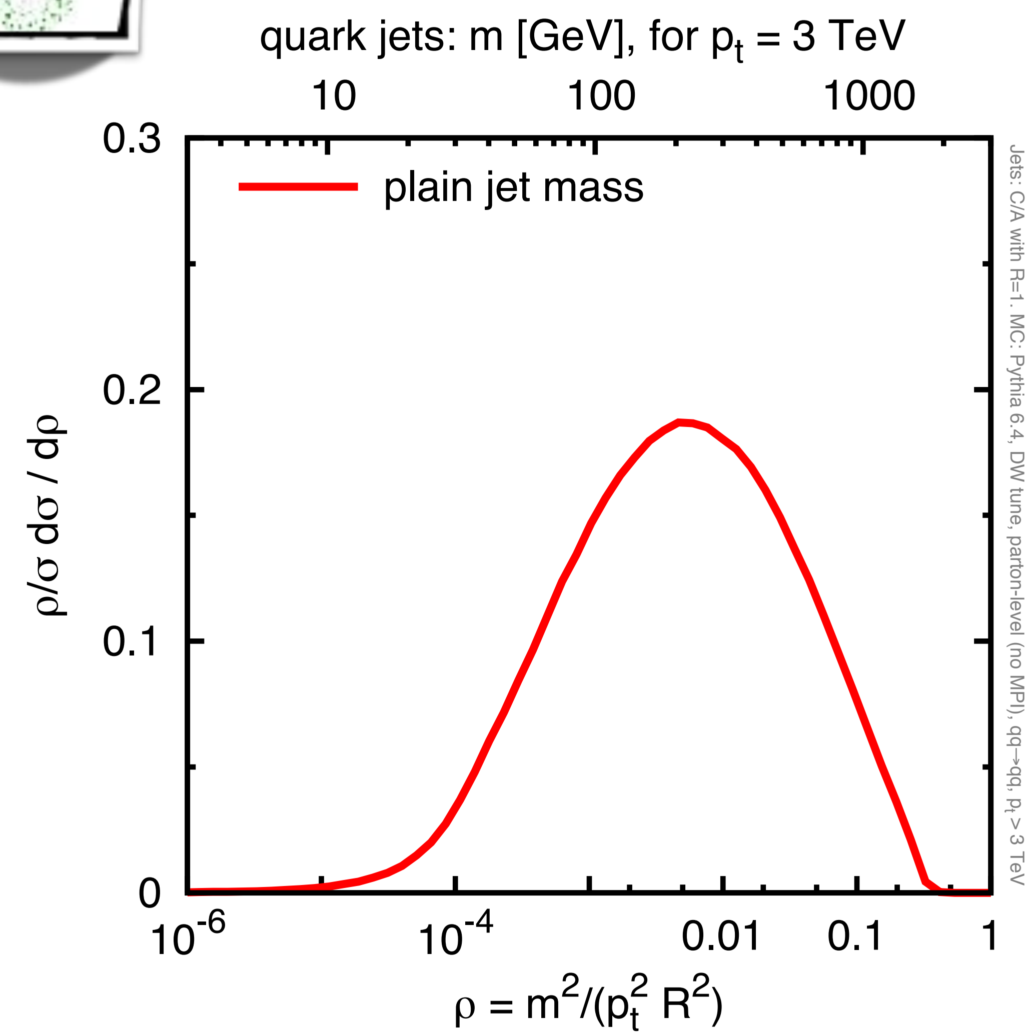
relative positions depends on physics context, kinematics, etc.

tagging

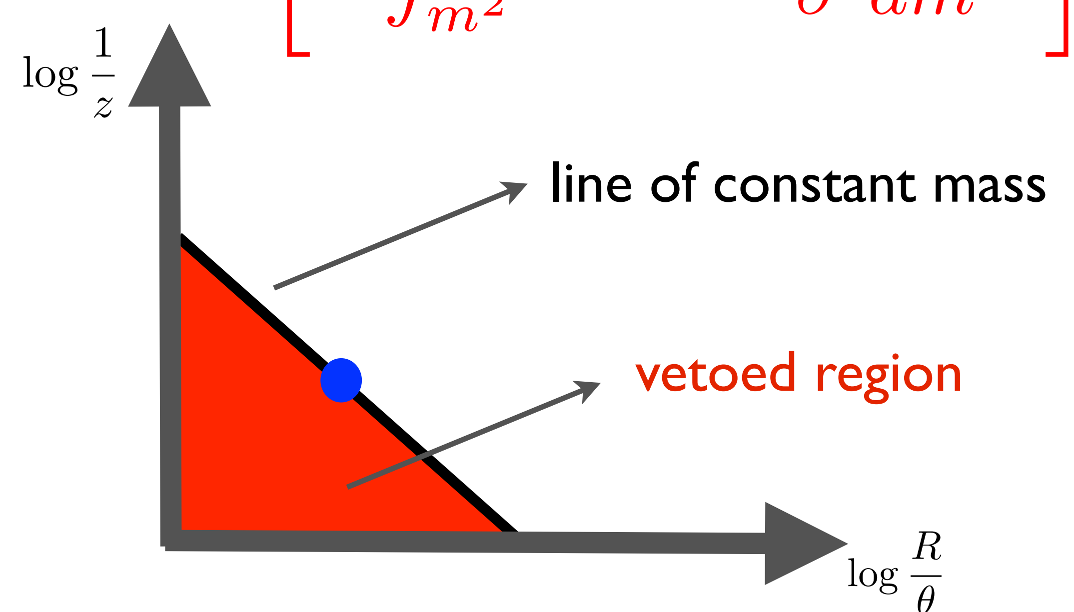
plot by G. Salam

recap: the jet mass

$\sigma_{res} = g_0 \exp[g_1(\alpha_s L)/\alpha_s + g_2(\alpha_s L) + \alpha_s g_3(\alpha_s L) + \dots]$



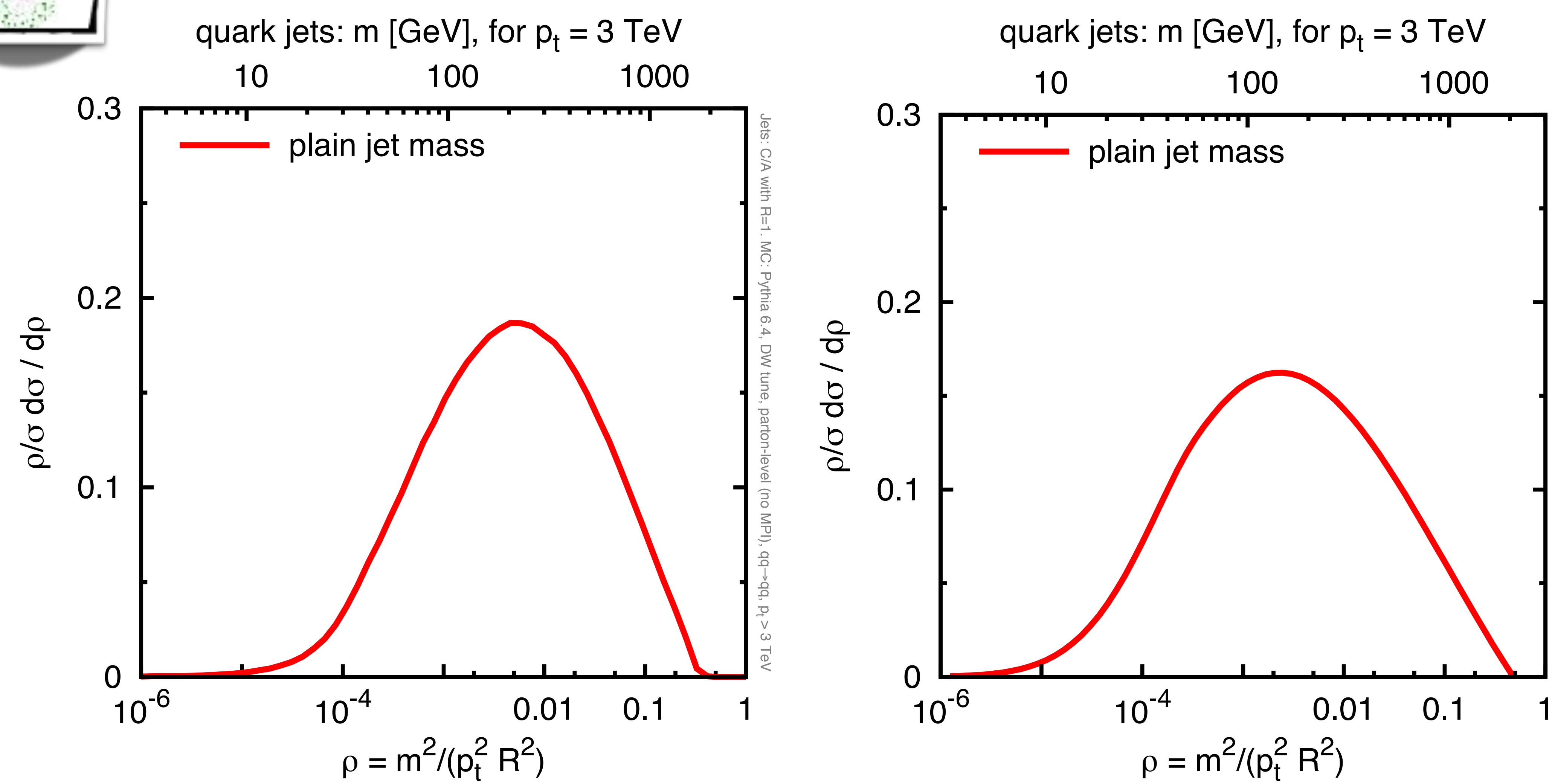
$$\frac{1}{\sigma}\frac{d\sigma^{\rm LL}}{dm^2} = \frac{1}{\sigma}\frac{d\sigma^{\rm LO}}{dm^2} \times \exp\left[-\int_{m^2}^{p_T^2 R^2} dm'^2 \frac{1}{\sigma}\frac{d\sigma^{\rm LO}}{dm'^2}\right]$$



- all-order leading logs: veto emissions which would give too big a mass
- exponential that gives the no-emission probability
- jet mass distributions exhibits double logs

recap: the jet mass

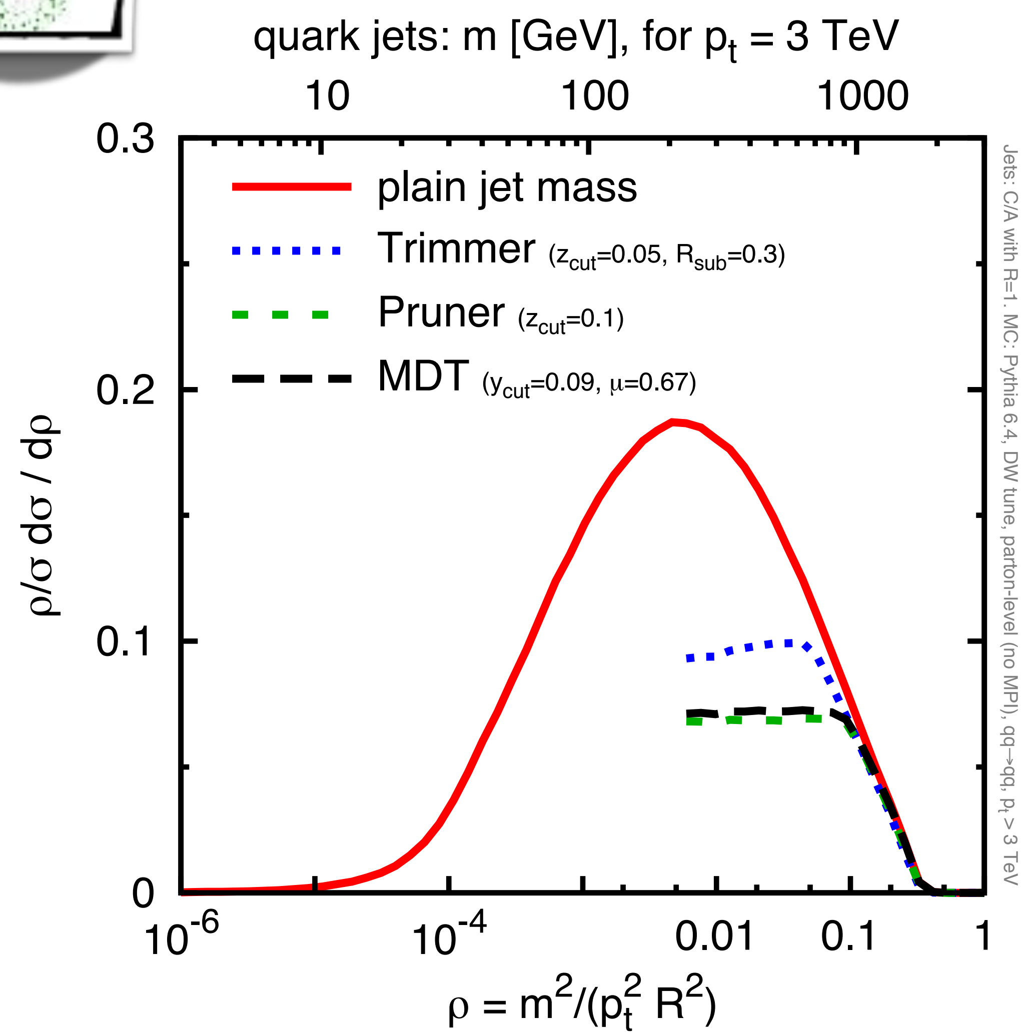
$\sigma_{res} = g_0 \exp[g_1(\alpha_s L)/\alpha_s + g_2(\alpha_s L) + \alpha_s g_3(\alpha_s L) + \dots]$



* all-order leading logs: veto emissions which would give too big a mass
* exponential that gives the no-emission probability
* jet mass distributions exhibits double logs

and now groomed masses

$\sigma_{res} = g_0 \exp[g_1(\alpha_s L)/\alpha_s + g_2(\alpha_s L) + \alpha_s g_3(\alpha_s L) + \dots]$

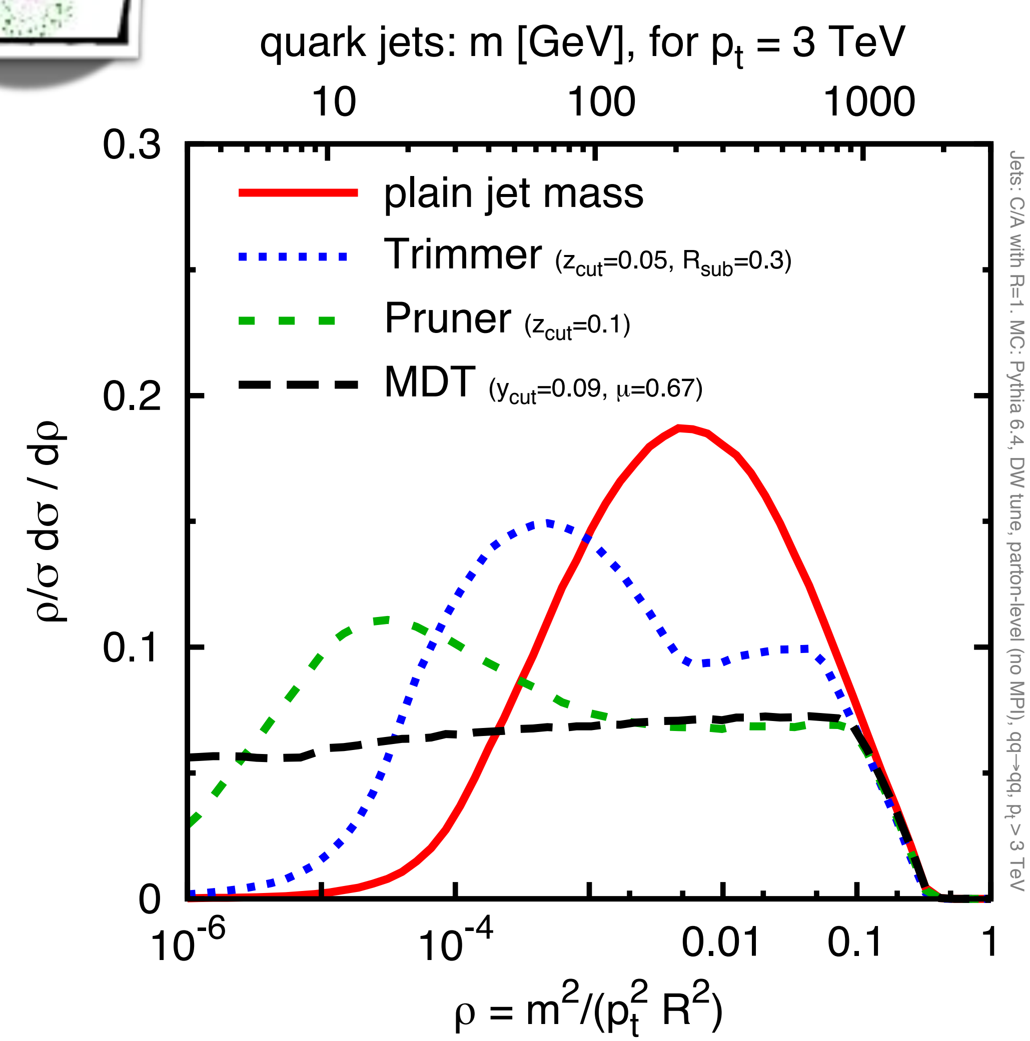


?

* different groomers / taggers appear to behave quite similarly

and now groomed masses

$\sigma_{res} = g_0 \exp[g_1(\alpha_s L)/\alpha_s + g_2(\alpha_s L) + \alpha_s g_3(\alpha_s L) + \ldots]$

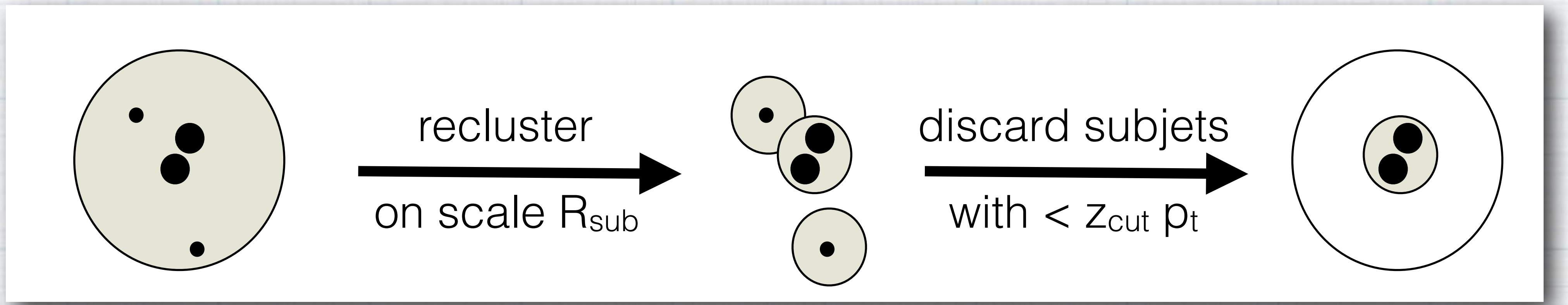


??

- but only for a limited kinematic region!
- complicated algorithm with many parameters
- can we compute groomed mass distributions?

trimming as an example

Krohn, Thaler and Wang (2010)

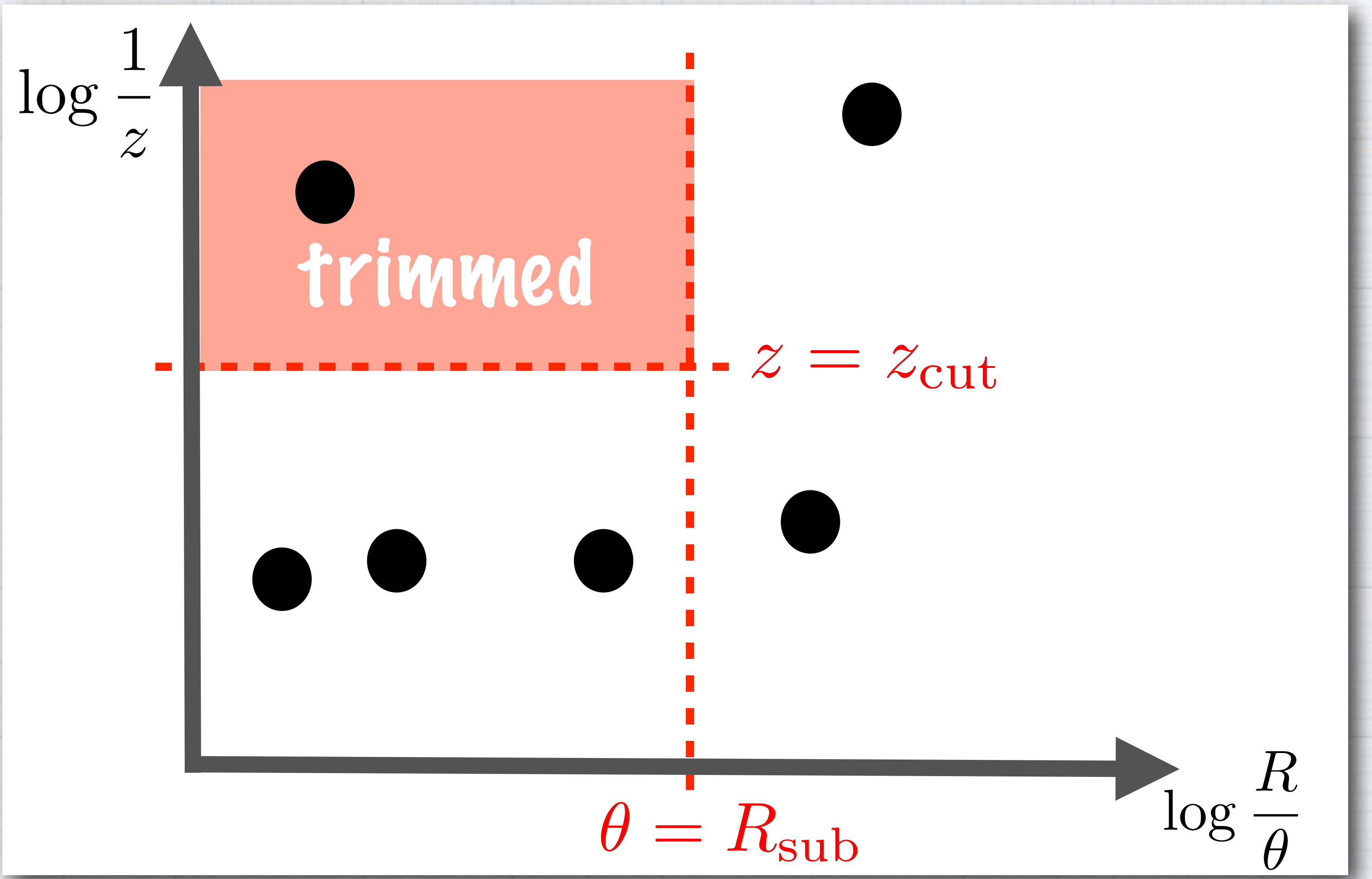


1. take all particles in a jet and re-cluster them with a smaller jet radius $R_{sub} < R$

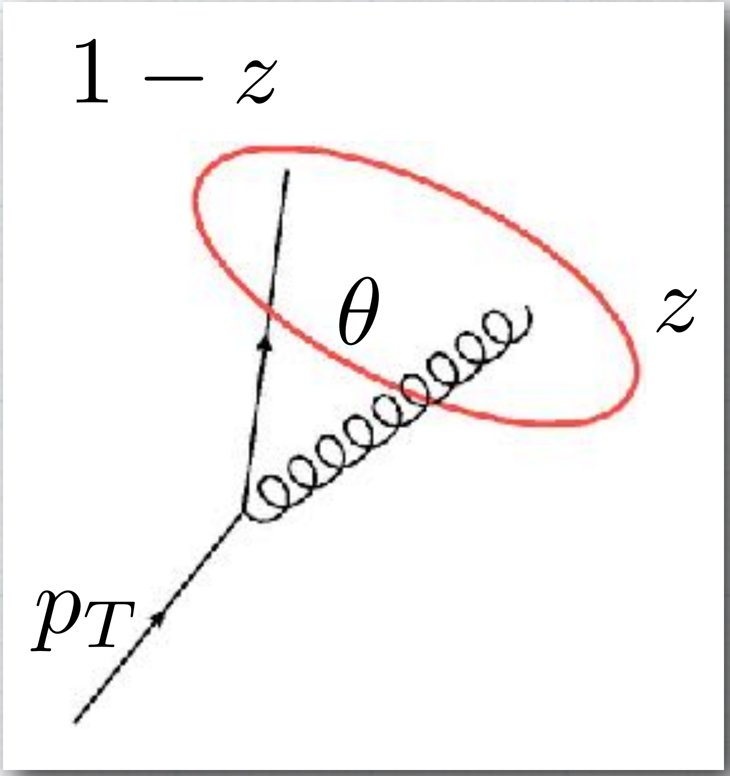
2. keep all subjets for which $p_t^{subjet} > z_{cut}\, p_t$

3. recombine the subjets to form the trimmed jet

trimming

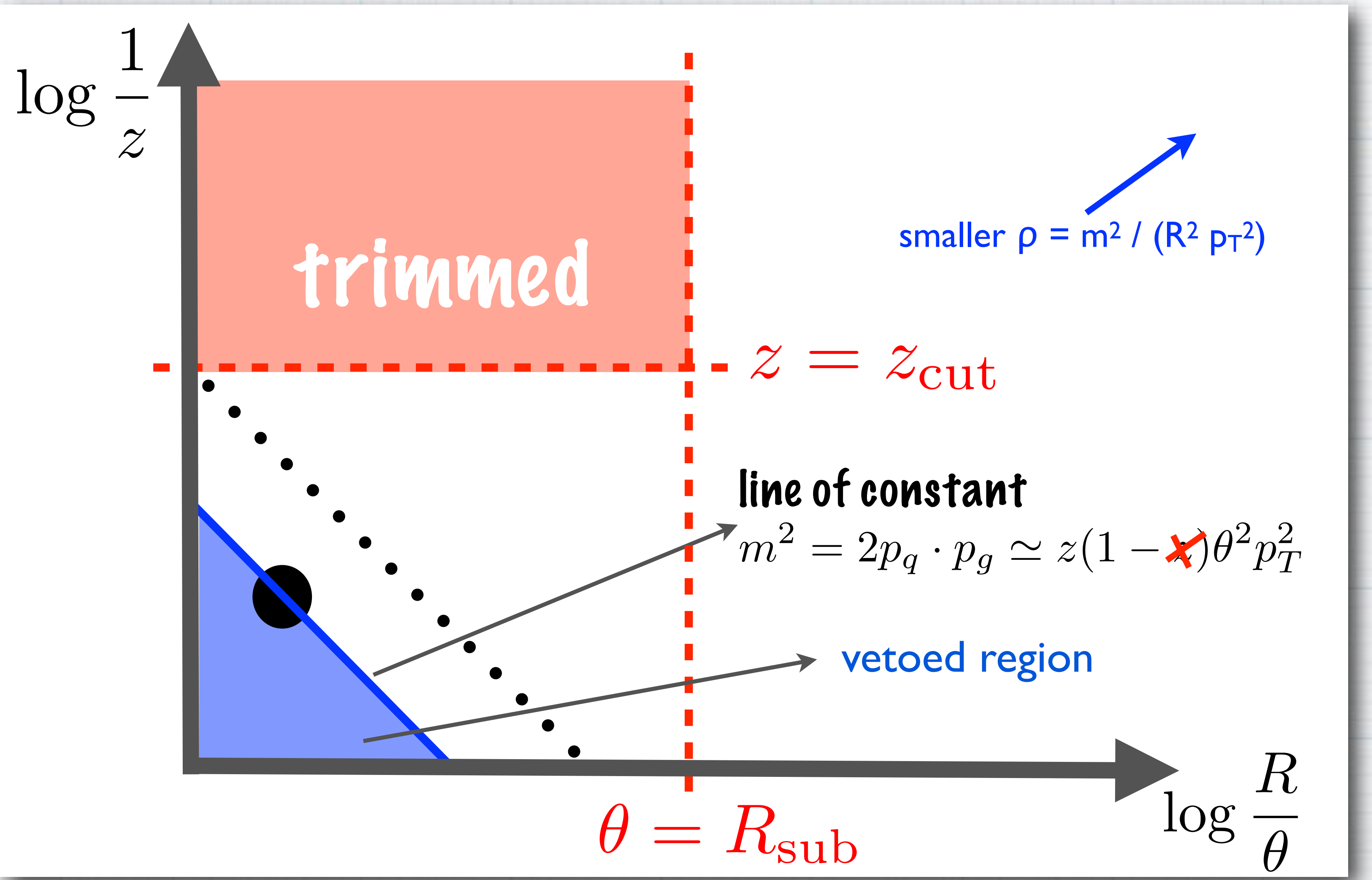


Soft gluons off a hard parton (a quark for definiteness)



* the action of a groomer is to remove some of the allowed phase space (typically soft and soft-collinear)
* what are the consequences for physical observables, e.g. the jet mass ?

trimmed mass at LL

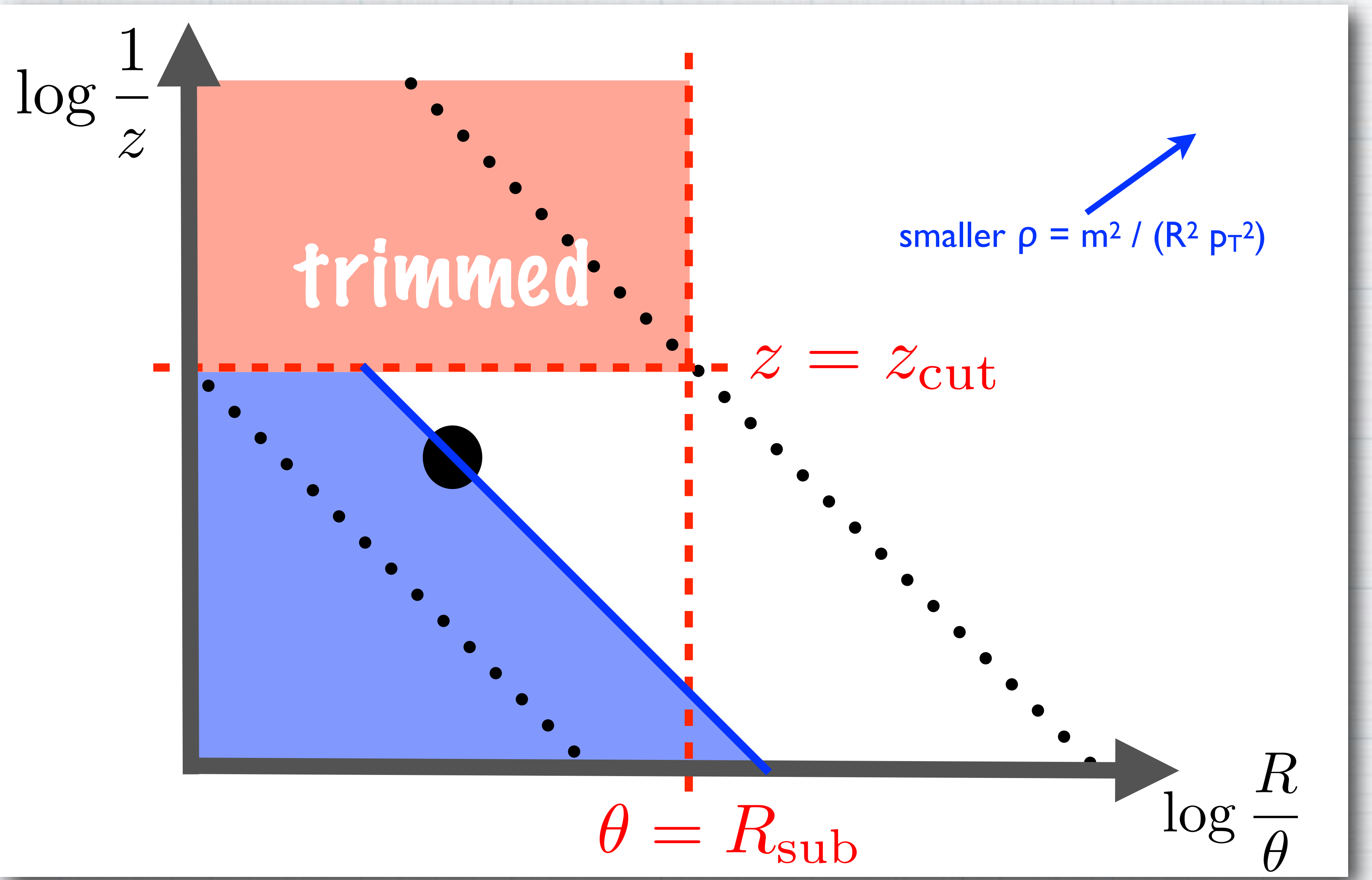


* one emission sets a **mass m**
* veto emissions that would give too big a mass
* **trimming here has no effect**

$$\Sigma(\rho) \equiv \int^{\rho} d\rho' \frac{1}{\sigma} \frac{d\sigma}{d\rho'}$$

$$\Sigma^{(\rm trim)}(\rho) \simeq \exp\left[-\frac{\alpha_s C_F}{2\pi}\left(-\frac{3}{2}\ln\frac{1}{\rho} + \Theta(\rho - z_{\rm cut})\ln^2\frac{1}{\rho} + \right)\right]$$

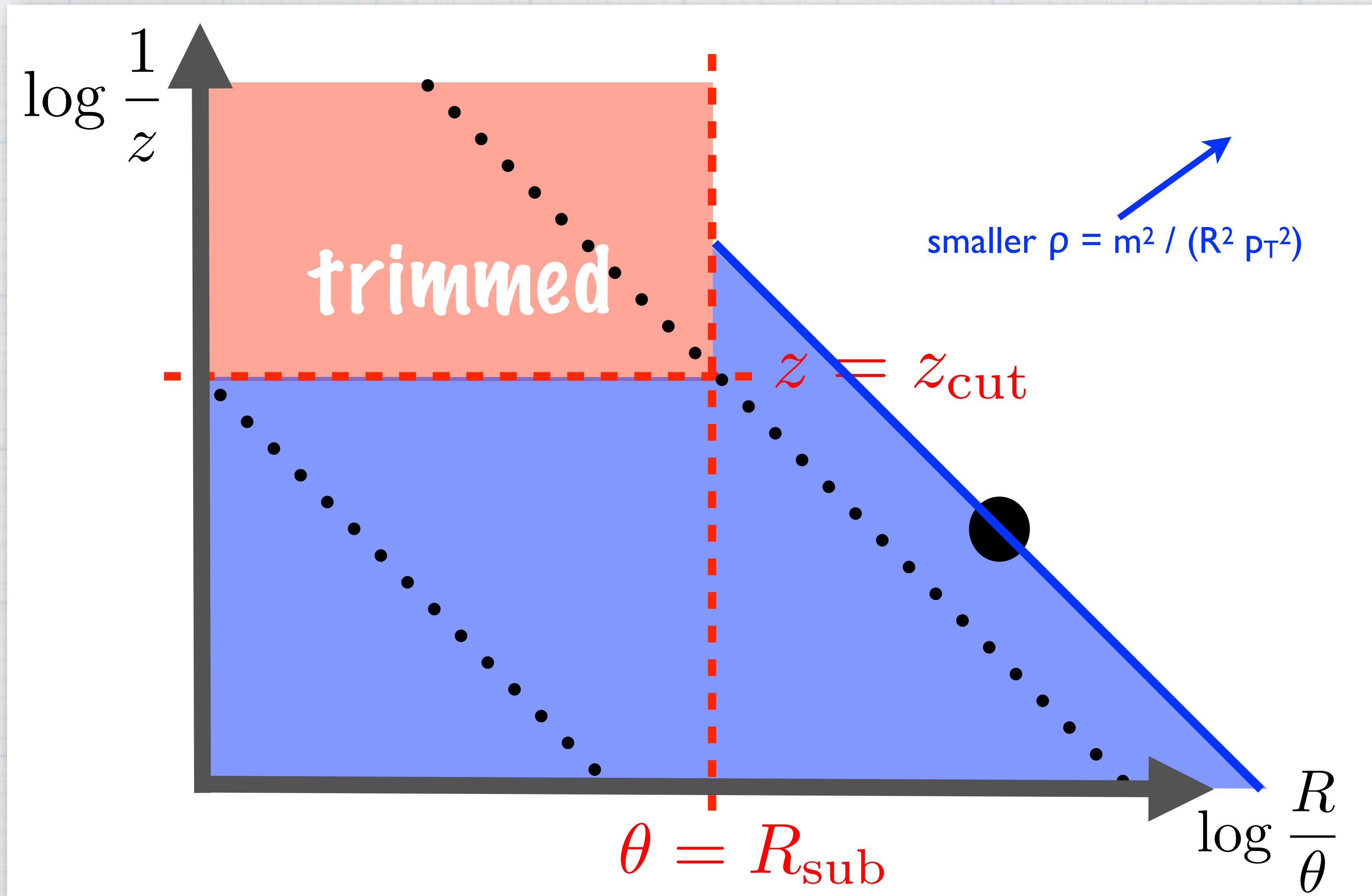
trimmed mass at LL



- first transition at $\rho = z_{cut}$
- soft & soft collinear radiation is trimmed away
- only single logs!

$$\Sigma^{(\rm trim)}(\rho) \simeq \exp\left[-\frac{\alpha_s C_F}{2\pi}\left(-\frac{3}{2}\ln\frac{1}{\rho} + \Theta(\rho - z_{\rm cut})\ln^2\frac{1}{\rho} + \right.\right.$$
$$\left.\left. +\Theta(z_{\rm cut}-\rho)\left(\ln^2\frac{1}{z_{\rm cut}} + 2\ln\frac{z_{\rm cut}}{\rho}\ln\frac{1}{z_{\rm cut}}\right)\right)\right]$$

trimmed mass at LL

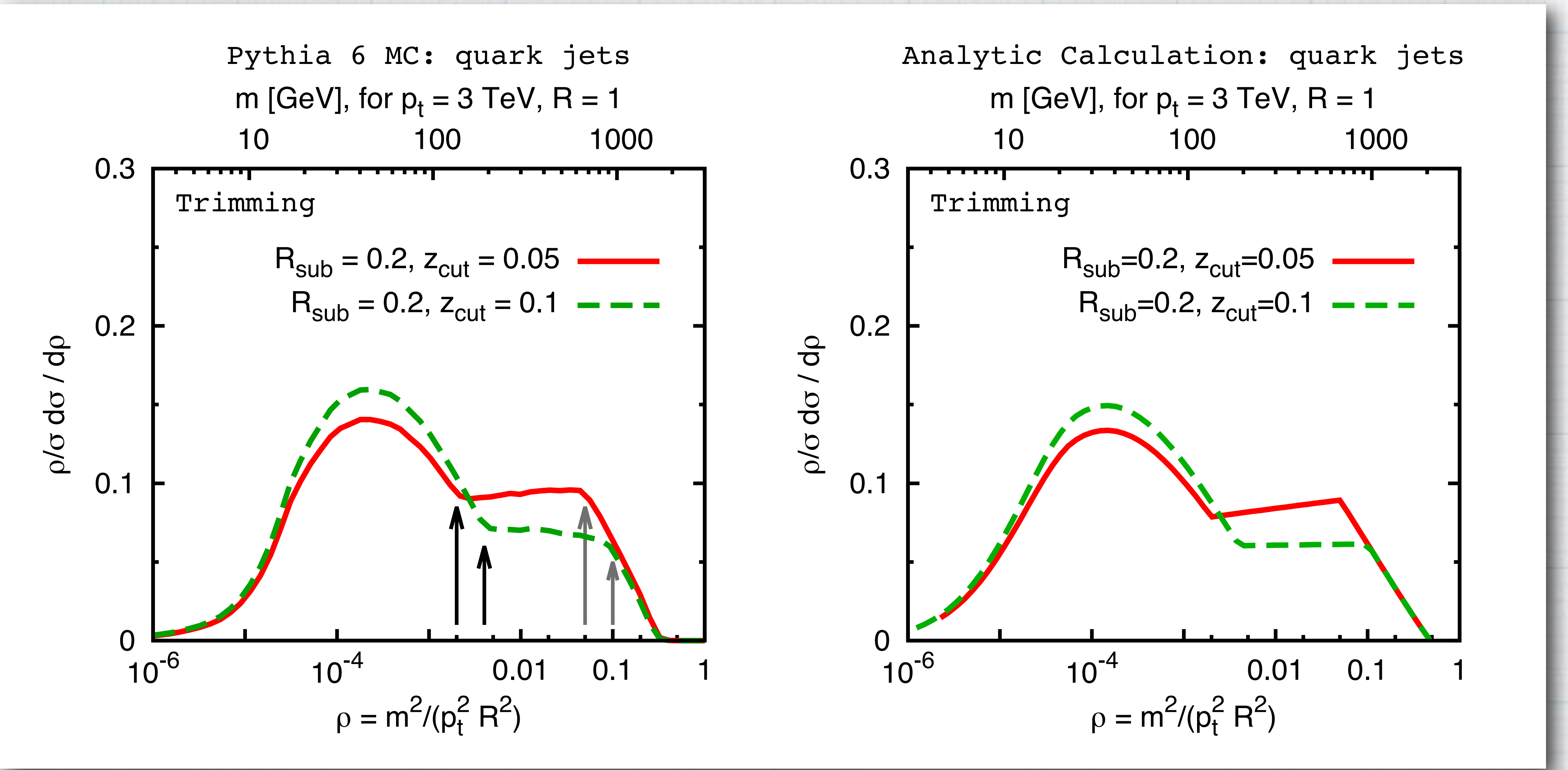


* second transition at $\rho = R_{sub}^2/R^2\, z_{cut}$
* soft & soft-collinear radiation below angular resolution isn't trimmed away
* **back to double logs** (same as plain mass)

$$\Sigma^{(\rm trim)}(\rho) \simeq \exp\left[-\frac{\alpha_s C_F}{2\pi}\left(-\frac{3}{2}\ln\frac{1}{\rho} + \Theta(\rho - z_{\rm cut})\ln^2\frac{1}{\rho} + \right.\right.$$
$$\left.\left. +\Theta(z_{\rm cut}-\rho)\left(\ln^2\frac{1}{z_{\rm cut}} + 2\ln\frac{z_{\rm cut}}{\rho}\ln\frac{1}{z_{\rm cut}}\right) + \Theta(z_{\rm cut} r^2 - \rho)\ln^2\frac{z_{\rm cut} r^2}{\rho}\right)\right]$$

trimmed mass: MC vs analytics

Modified LL (MLL): LL + hard collinear + running coupling

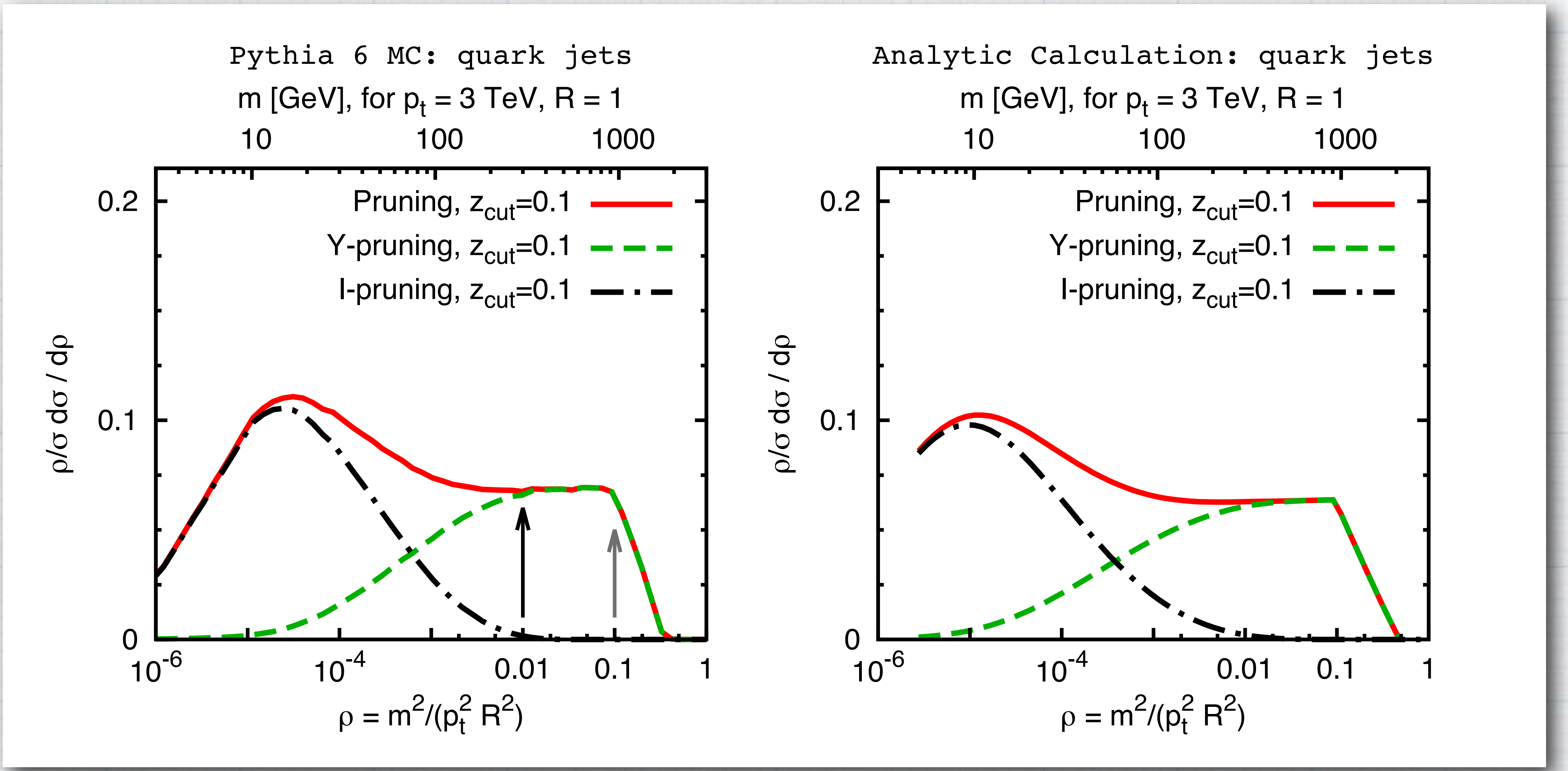


* trimming is active (and aggressive) for $z_{cut} < \rho < R_{sub}^2/R^2\, z_{cut}$
* not active below because of fixed R_{sub}

pruned mass: MC vs analytics

Ellis, Vermilion, Walsh (2010)

Modified LL (MLL): LL + hard collinear + running coupling

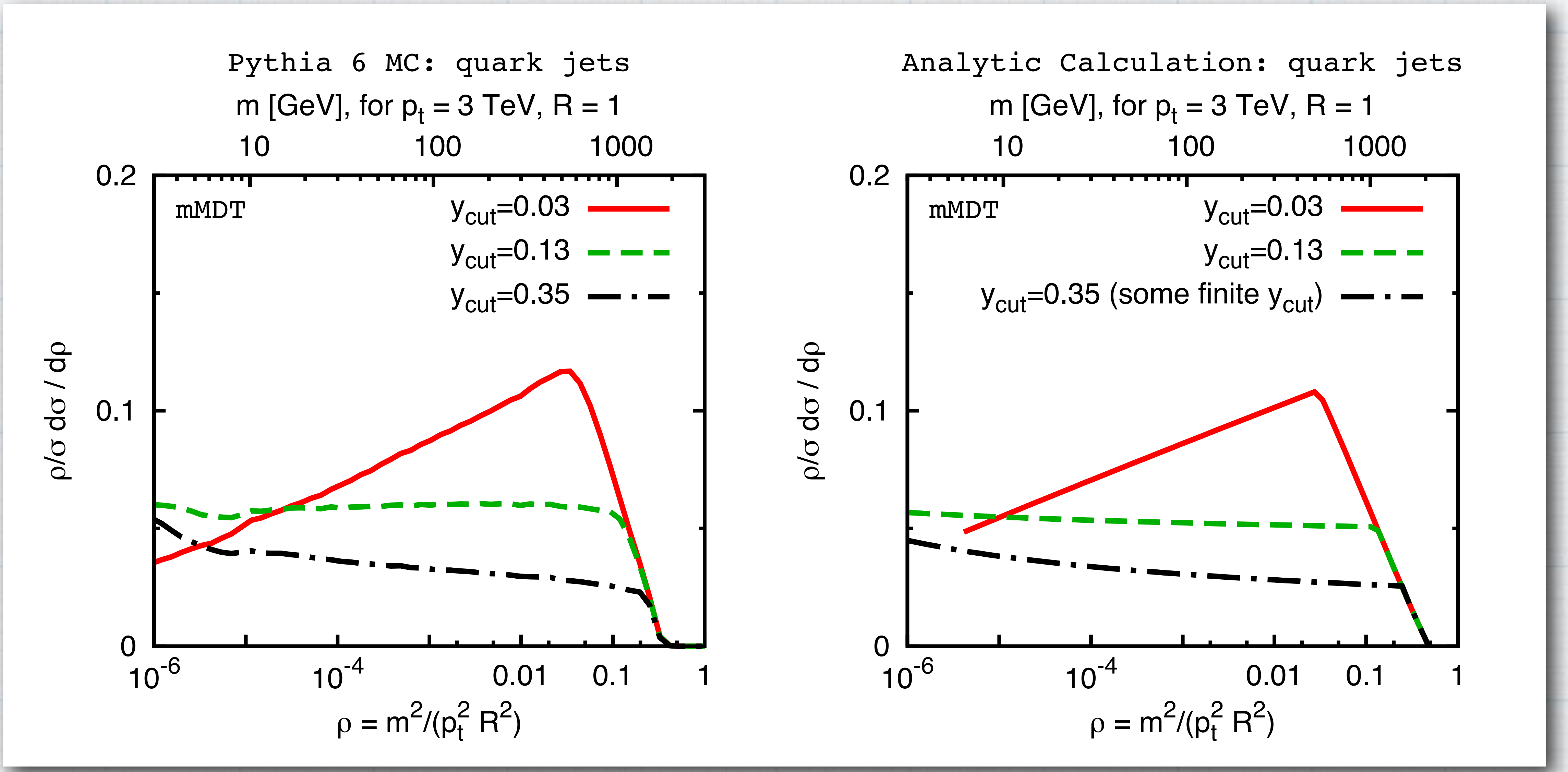


* more complex structure, no simple exponentiation
* single logs for $z_{cut}^2 < \rho < z_{cut}$

mMDT mass: MC vs analytics

Dasgupta, Fregoso, SM, Salam (2013)

Modified LL (MLL): LL + hard collinear + running coupling



* MDT has only single logs at LO
* modified MDT maintains this feature to all orders

homework 6

Show that the leading-order mass distributions for **MDT** and **pruning** are single-logarithmic. (This doesn't hold at higher orders!). Use the definition below

1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2 ($m_1 > m_2$)
2. If $m_1 < \mu m$ (mass drop) and the splitting was not too asymmetric, ie

$$\frac{\min(p_{ti}, p_{tj})}{p_{ti} + p_{tj}} > z_{\text{cut}}$$

tag the jet.
3. Otherwise redefine $j = j_1$ and iterate.

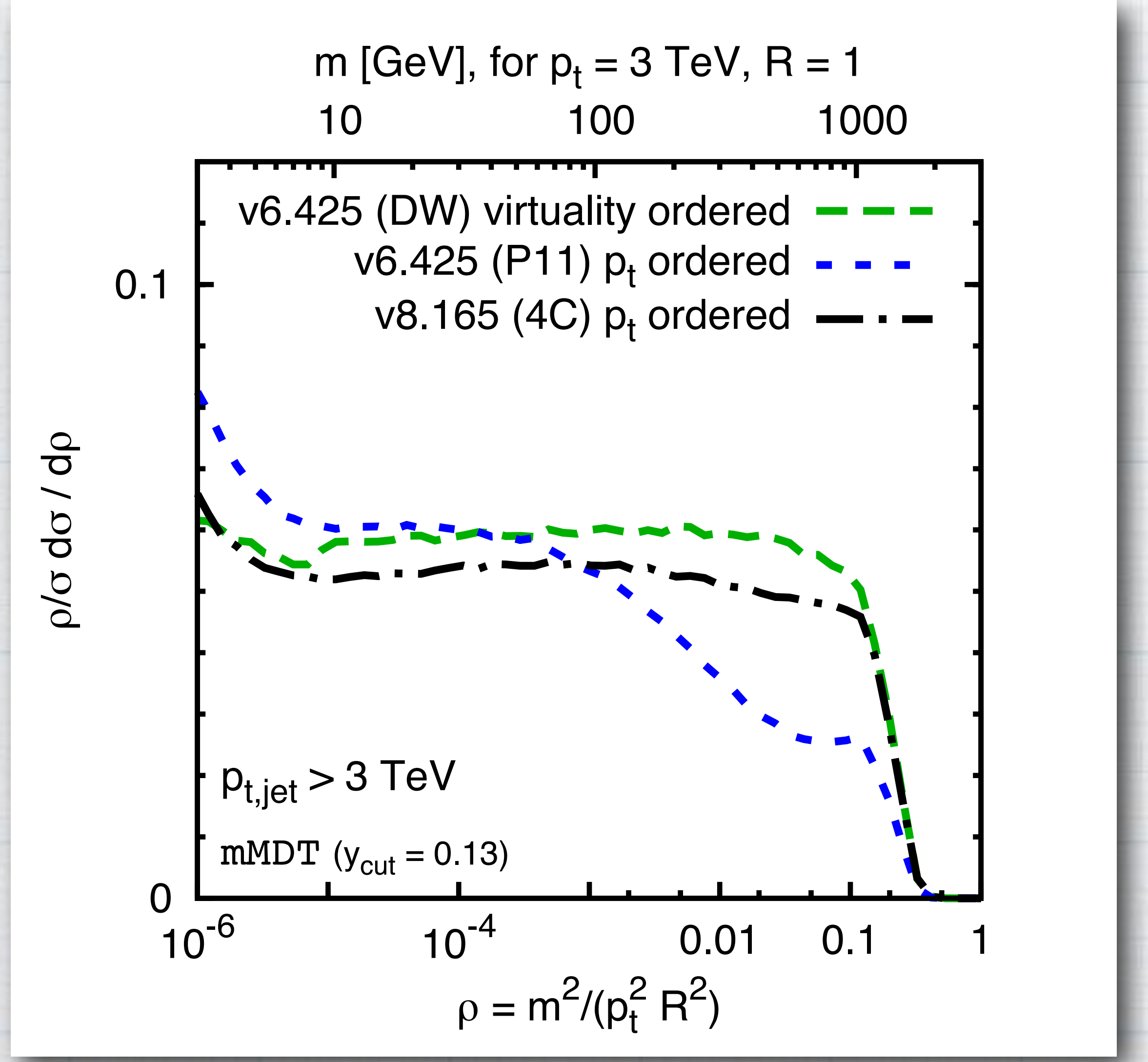
1. From an initial jet with mass m define the pruning radius $R_{prune} = m / p_t$
2. Re-cluster the jet, vetoing recombination for which: $d_{ij} > R_{prune}$ and

$$\frac{\min(p_{ti}, p_{tj})}{p_{ti} + p_{tj}} > z_{\text{cut}}$$

[Hint] Consider as in the lecture the emission of a collinear gluon off a quark. Take the small-z_{cut} limit to simplify your expressions.

analytics to check MCs

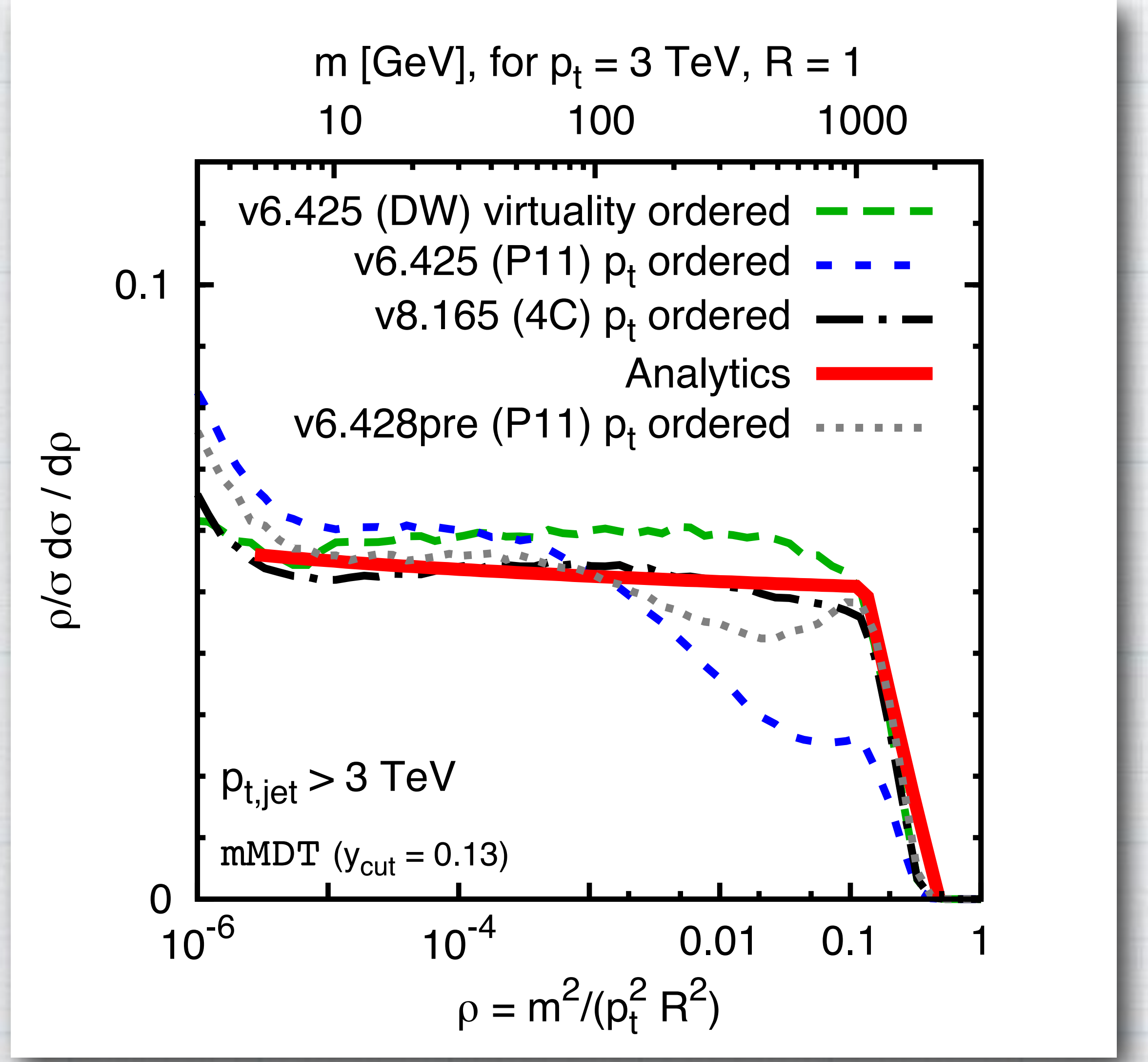
* so far we have always compared to a single MC simulation
* how solid are MC descriptions ?



* take the spread as the uncertainty ?
* but we can also add the analytic calculation

analytics to check MCs

* so far we have always compared to a single MC simulation
* how solid are MC descriptions ?



* take the spread as the uncertainty ?
* but we can also add the analytic calculation

* problem in the shower: fixed by the Authors in the 6.428pre version

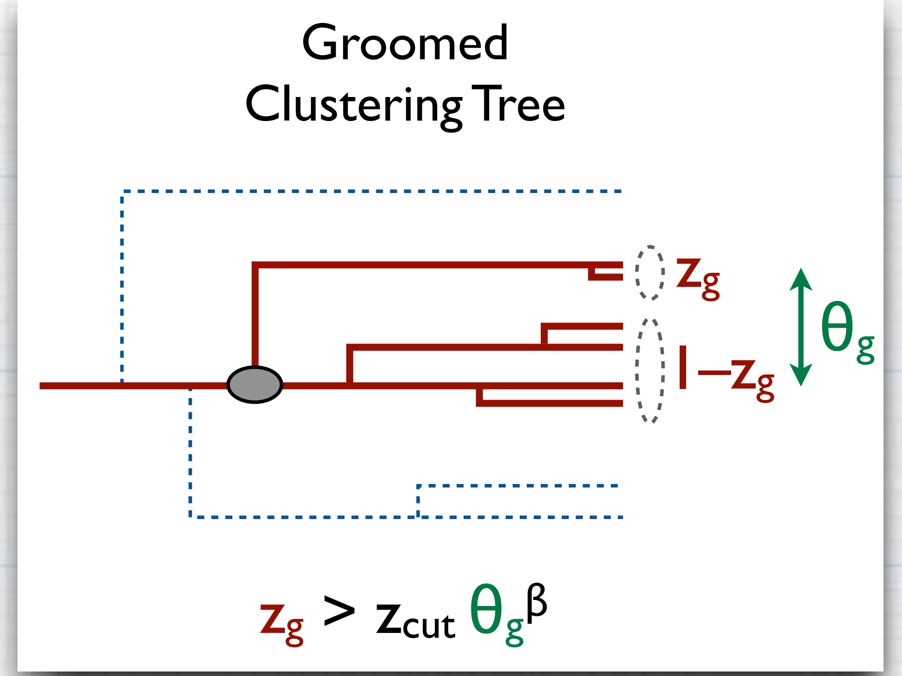
analytic understanding at work: soft drop

Larkoski, SM, Soyez and Thaler (2014)

1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2.

2. If $$\frac{\min(p_{T1}, p_{T2})}{p_{T1}+p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$$

   then deem j to be the soft-drop jet.



3. Otherwise redefine j to be the harder subjet and iterate.

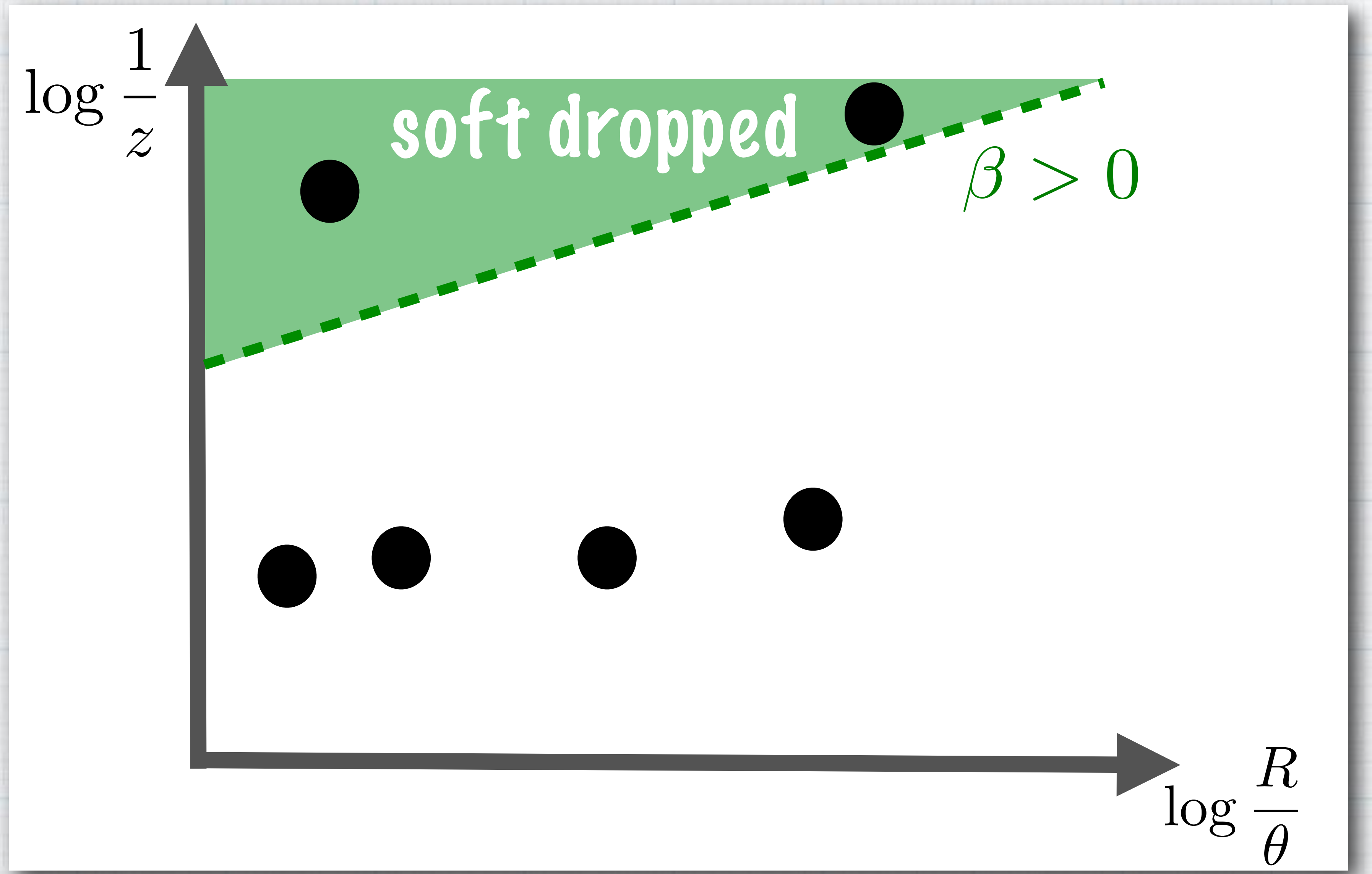
1-prong jets can be either kept (grooming mode) or discarded (tagging mode)

* generalisation of the (modified) Mass Drop procedure
* no mass drop condition (not so important)
* mMDT recovered for β=0
* some inspiration from semi-classical jets

Butterworth, Davison, Rubin and Salam (2008)
Dasgupta, Fregoso, SM and Salam (2013)

Tseng and Evans (2013)

soft drop as a groomer

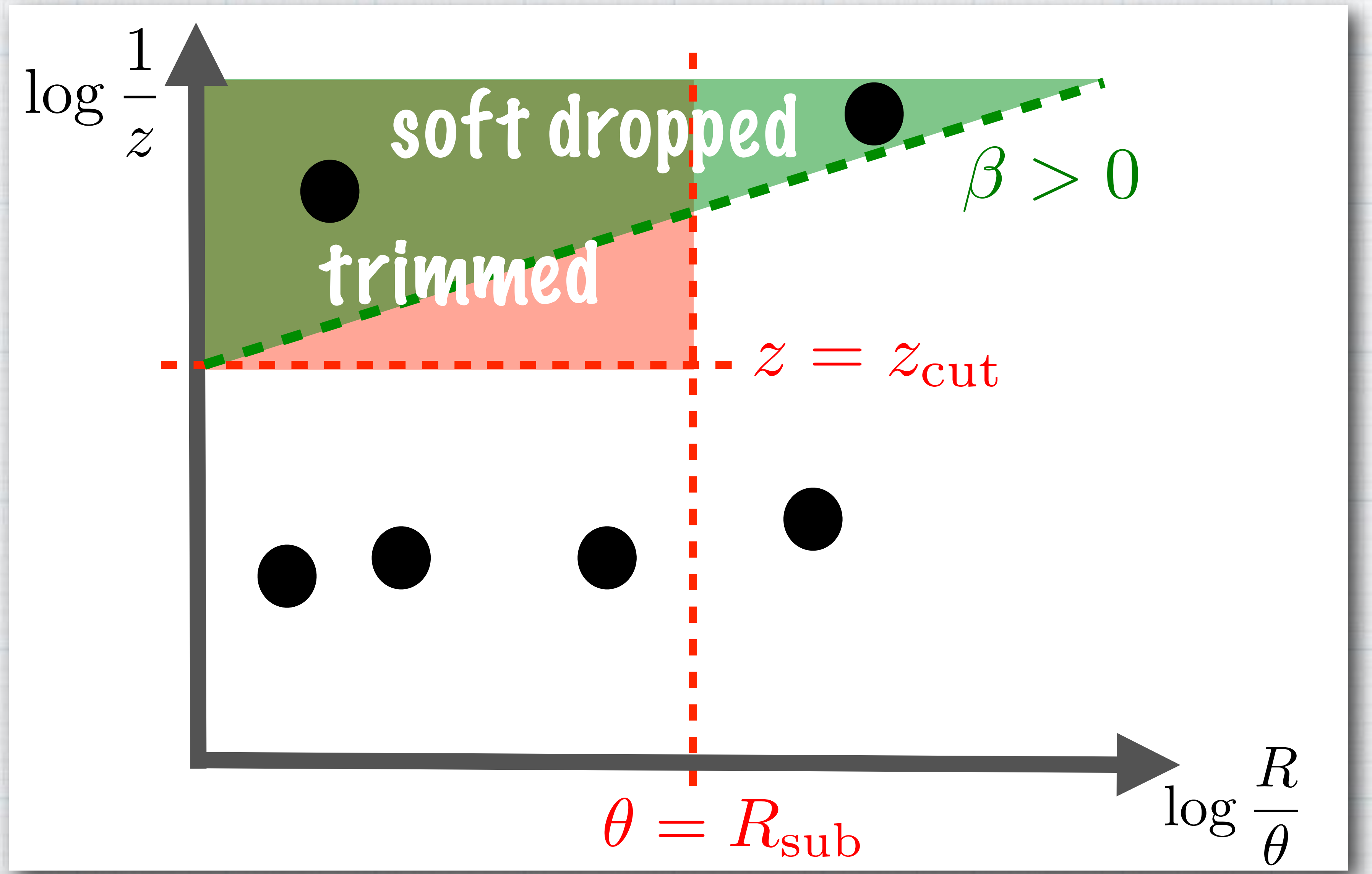


* useful to consider the soft-gluon phase space
* soft-drop condition becomes

$$z > z_{\mathrm{cut}} \left(\frac{\theta}{R} \right)^{\beta}$$

* soft drop always removes soft radiation entirely (hence the name)
* for β>0 soft-collinear is partially removed

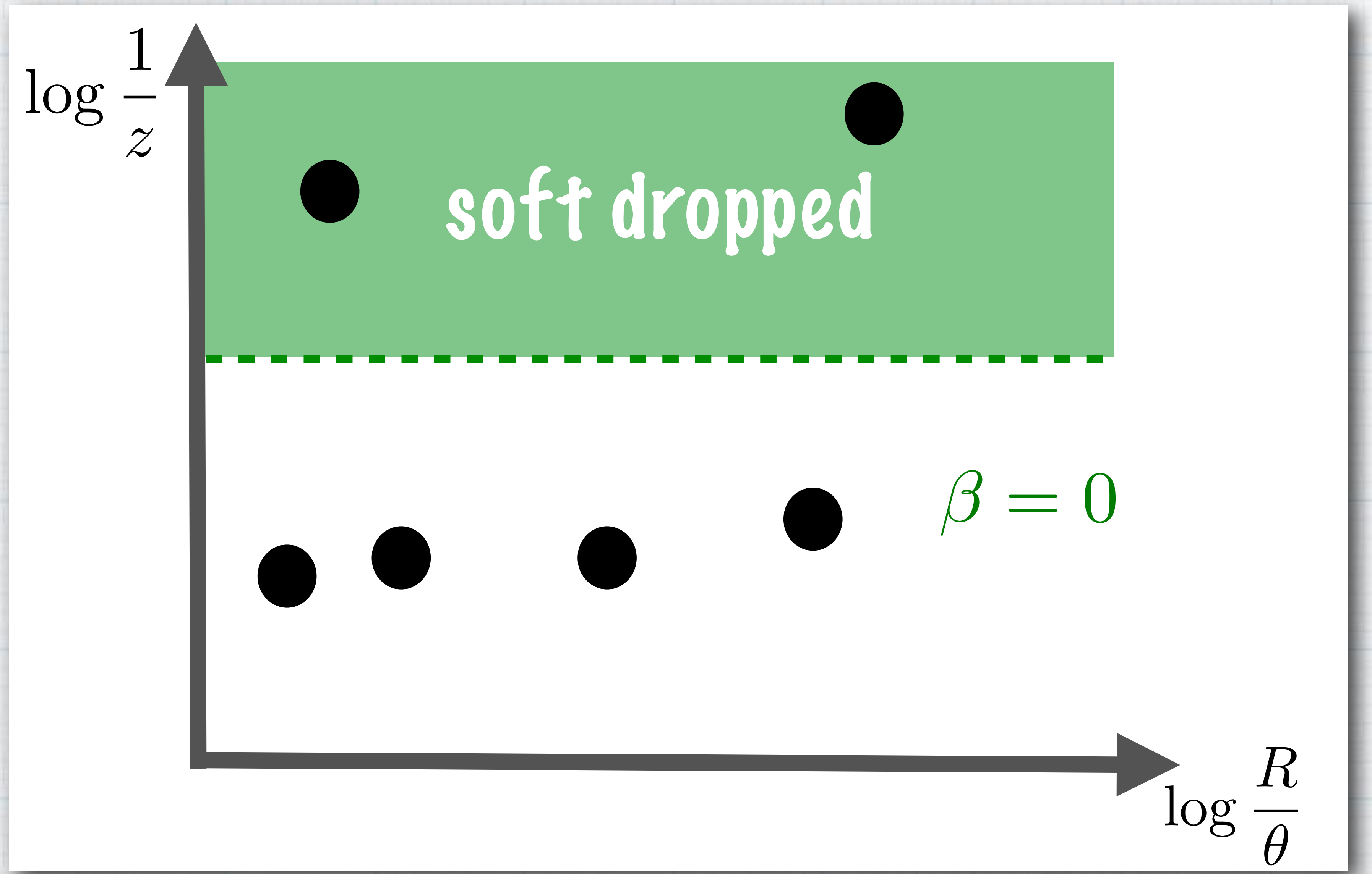
soft drop vs trimming



- trimming had an abrupt change of behaviour due to fixed R_{sub}
- in soft-drop angular resolution controlled by the exponent β
- phase-space appears smoother

Soft drop in grooming mode (β>0) works as a dynamical trimmer

soft drop and mMDT

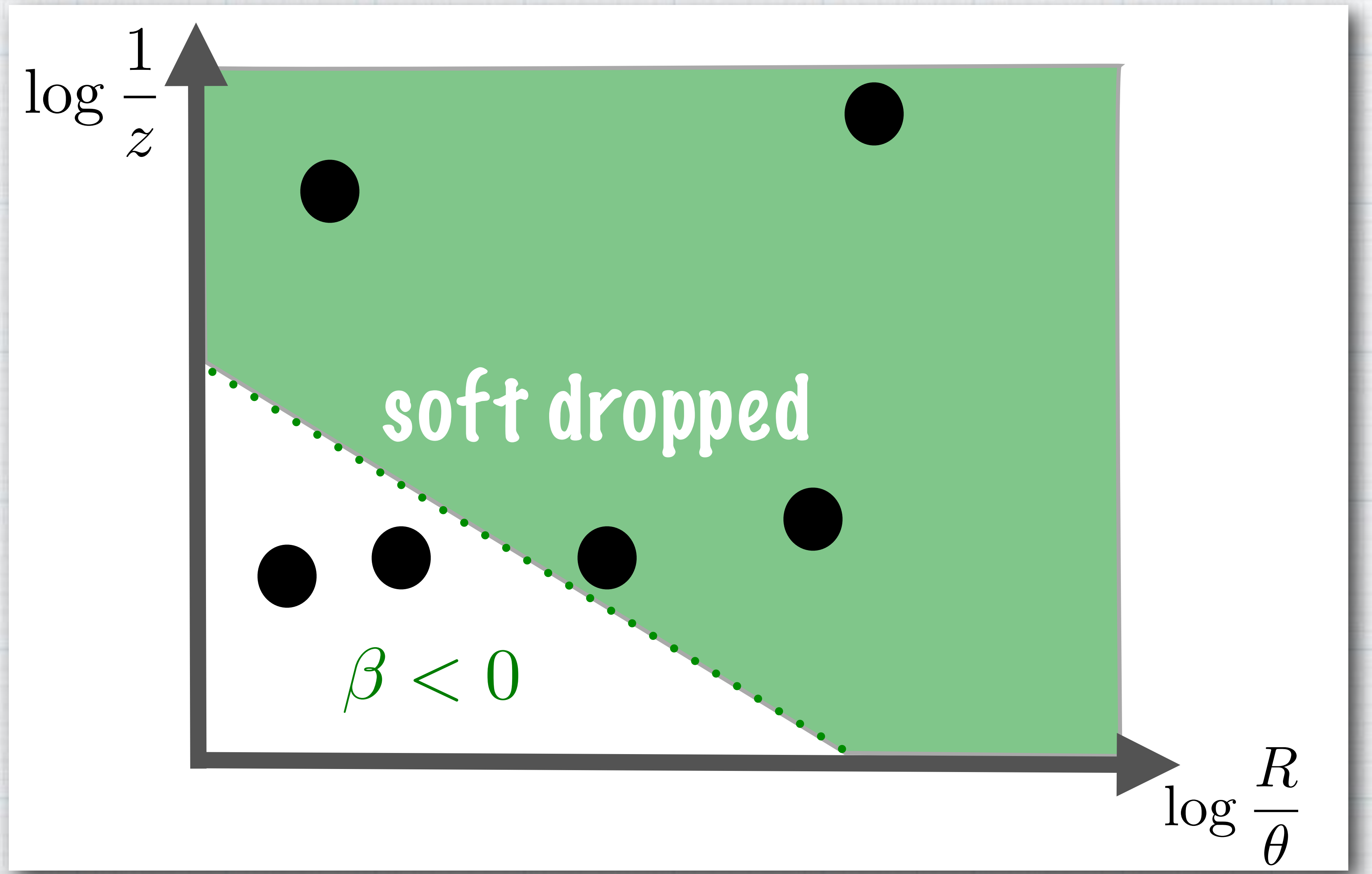


* soft-drop condition becomes

$$z > z_{\text{cut}} \left(\frac{\theta}{R}\right)^{\beta}$$

* for β=0 soft-collinear is also entirely removed (mMDT limit)

soft drop as a tagger

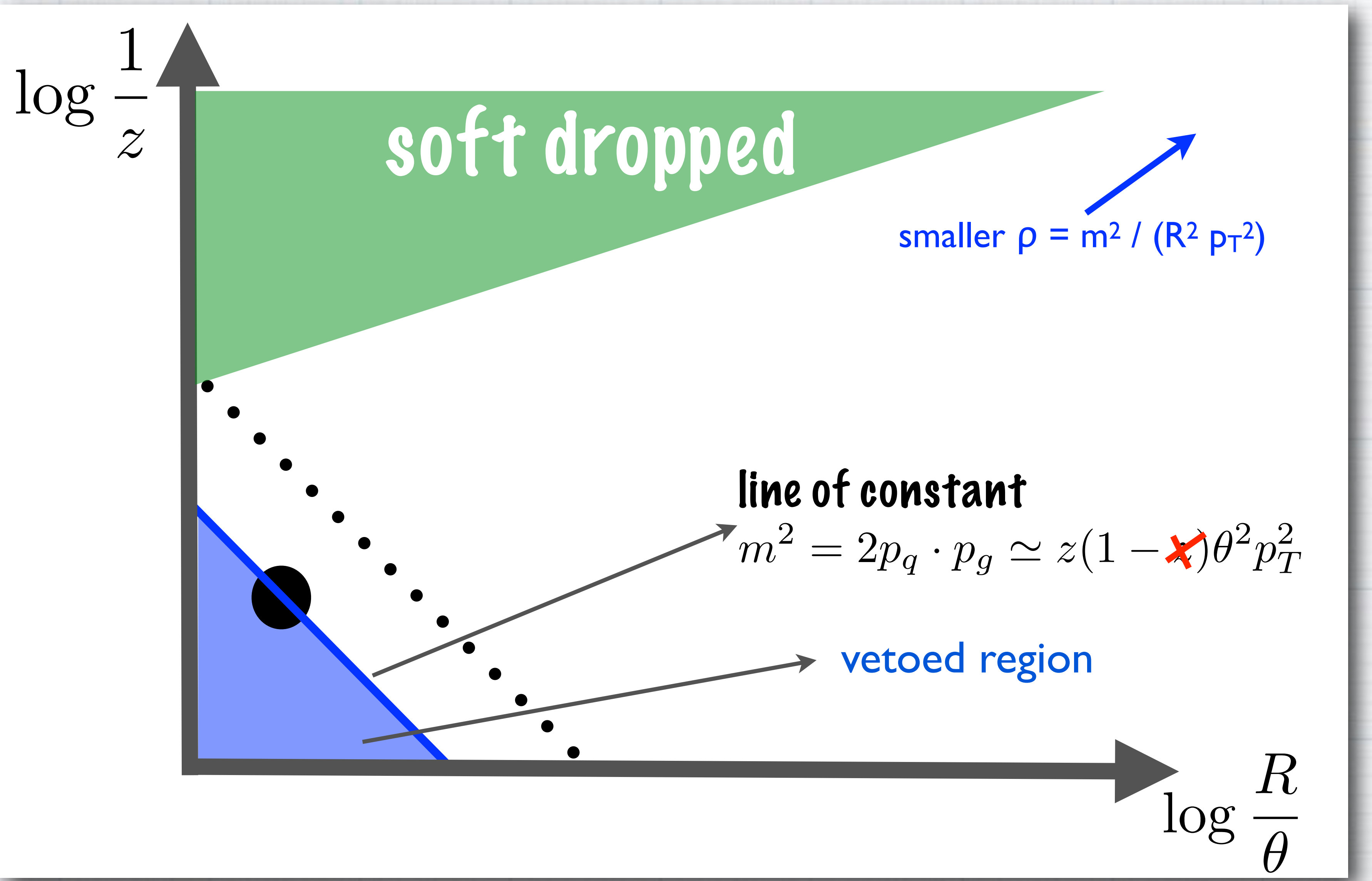


* soft-drop condition becomes

$$z > z_{\text{cut}} \left(\frac{\theta}{R} \right)^{\beta}$$

* for β<0 some hard-collinear is also partially removed

soft-drop mass at LL

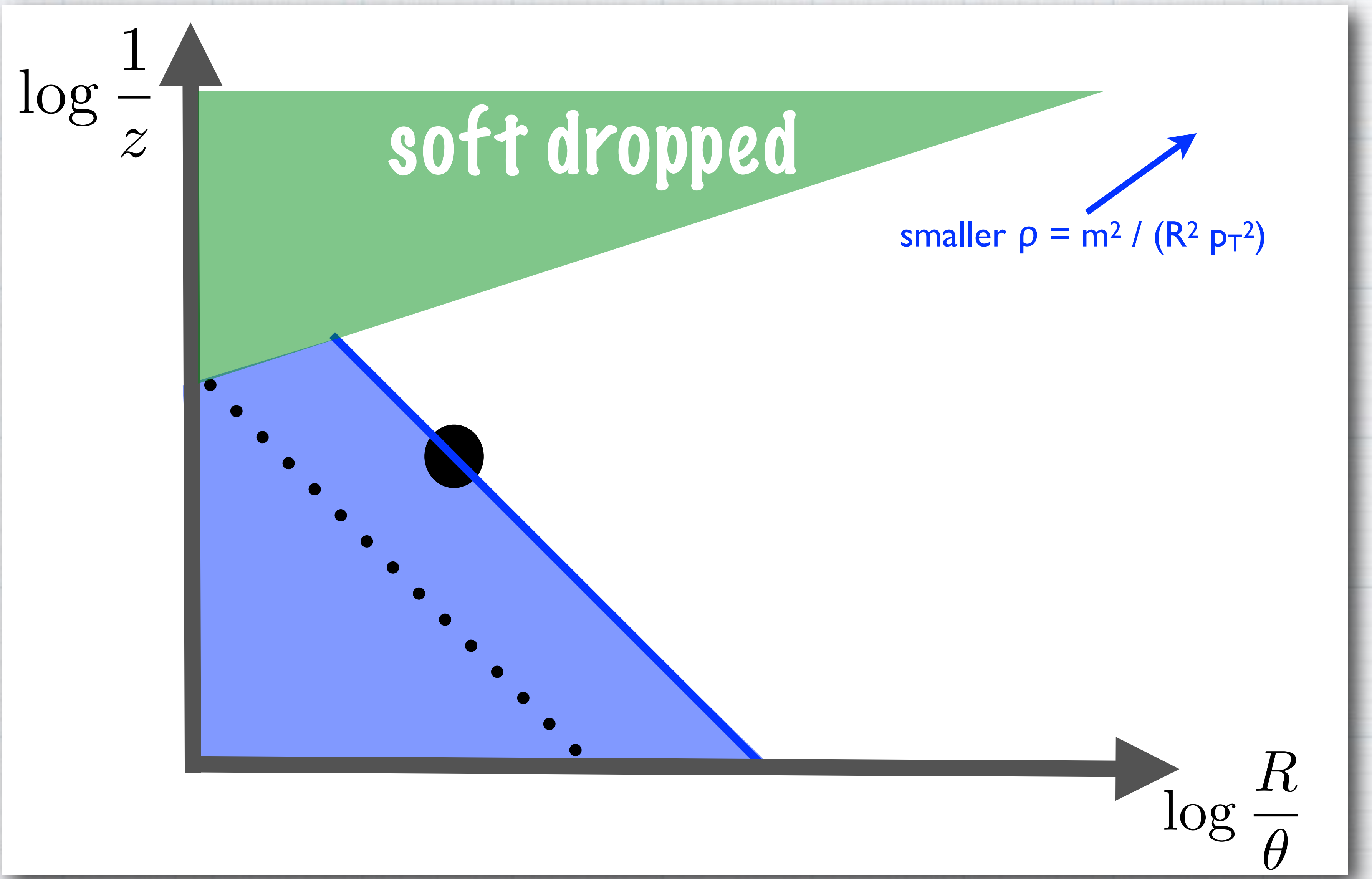


- one emission sets a **mass m**
- veto emissions that would give too big a mass
- **soft drop here has no effect**

$$\Sigma(\rho) \equiv \int^{\rho} d\rho' \frac{1}{\sigma}\frac{d\sigma}{d\rho'}$$

$$\Sigma^{(\text{s.d.})} = \exp\left[-\frac{\alpha_s C_F}{2\pi}\left[-\frac{3}{2}\ln\frac{1}{\rho} + \Theta(\rho > z_{\text{cut}})\ln^2\frac{1}{\rho} + \Theta(z_{\text{cut}} > \rho)\left(\frac{\beta}{2+\beta}\ln^2\frac{1}{\rho} + \frac{2}{2+\beta}\left(\ln^2\frac{1}{z_{\text{cut}}} + 2\ln\frac{z_{\text{cut}}}{\rho}\ln\frac{1}{z_{\text{cut}}}\right)\right)\right]\right]$$

soft-drop mass at LL

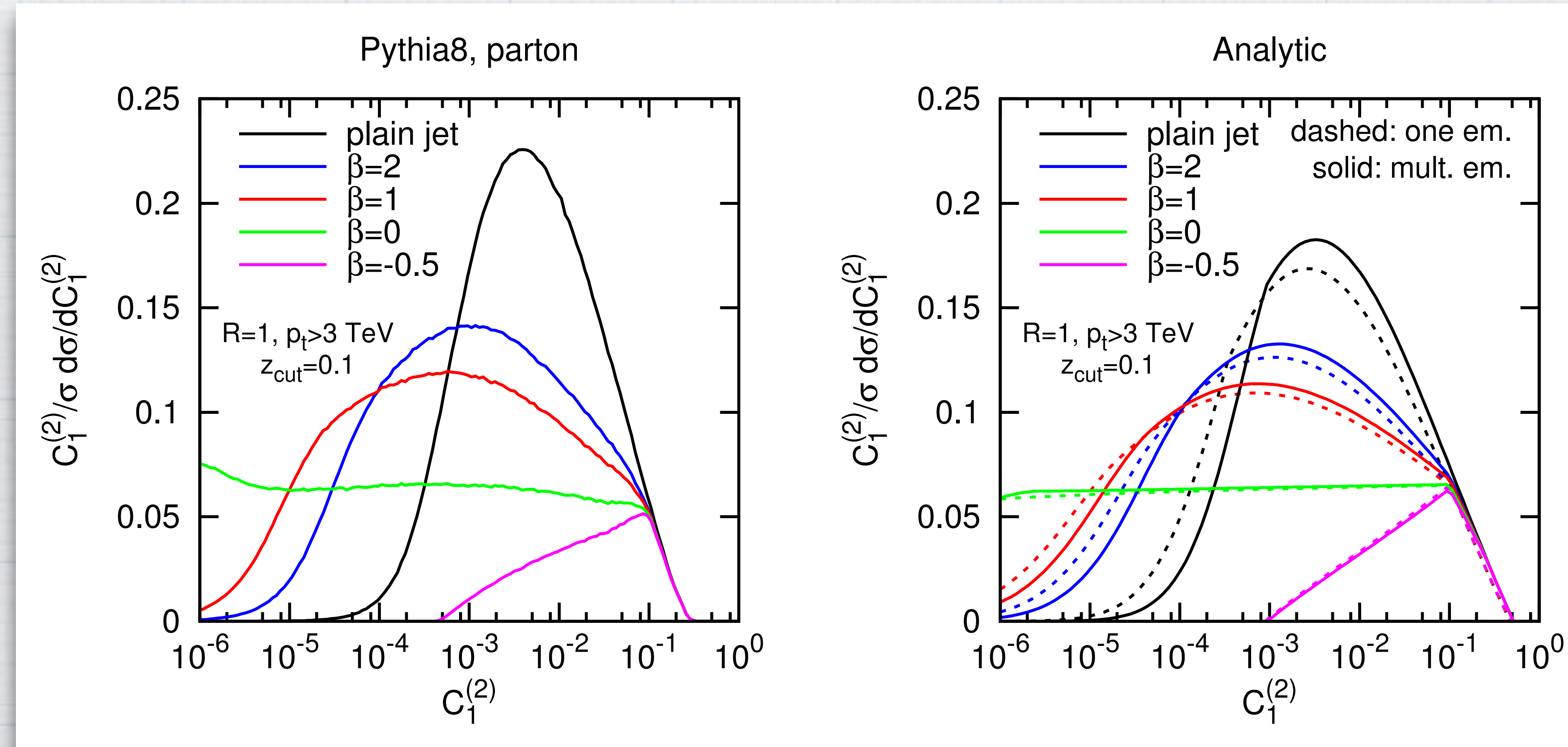


* only one transition point at $\rho = z_{cut}$
* soft & soft collinear radiation is partially removed
* only single logs for β=0!

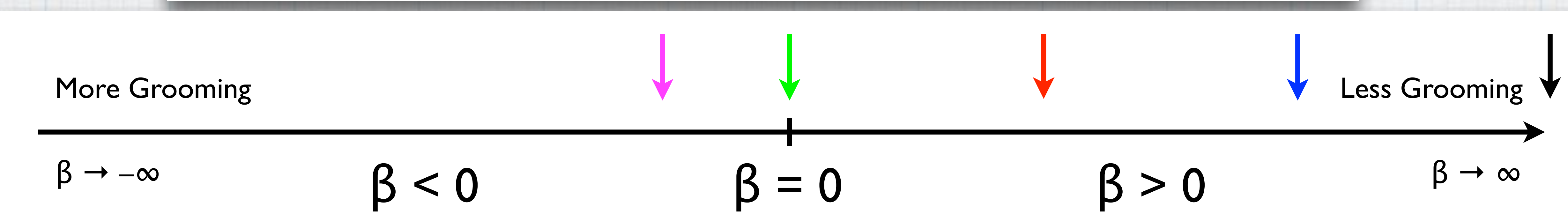
$$\Sigma^{(\text{s.d.})} = \exp\left[-\frac{\alpha_s C_F}{2\pi} \left[-\frac{3}{2} \ln\frac{1}{\rho} + \Theta(\rho > z_{\text{cut}}) \ln^2\frac{1}{\rho} \right.\right.$$
$$\left.\left. + \Theta(z_{\text{cut}} > \rho) \left(\frac{\beta}{2+\beta} \ln^2\frac{1}{\rho} + \frac{2}{2+\beta} \left(\ln^2\frac{1}{z_{\text{cut}}} + 2\ln\frac{z_{\text{cut}}}{\rho} \ln\frac{1}{z_{\text{cut}}} \right) \right) \right] \right]$$

soft-drop mass: MC vs analytics

Modified LL (MLL): LL + hard collinear + running coupling

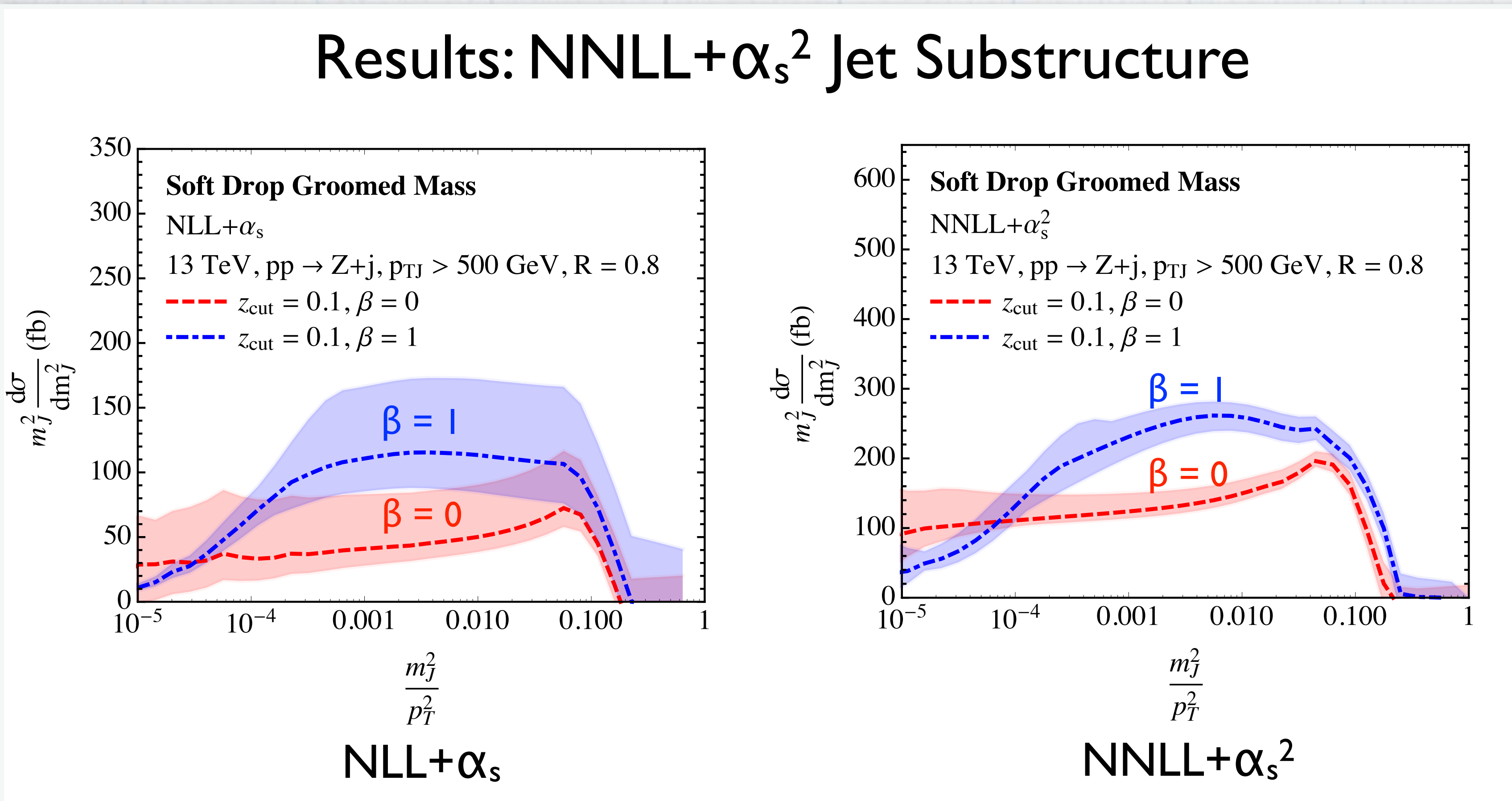


$C_1^{(2)} \simeq m^2/p_T^2$



* smooth distributions
* flatness in bkg can be achieved for β=0
* now standard choice for CMS

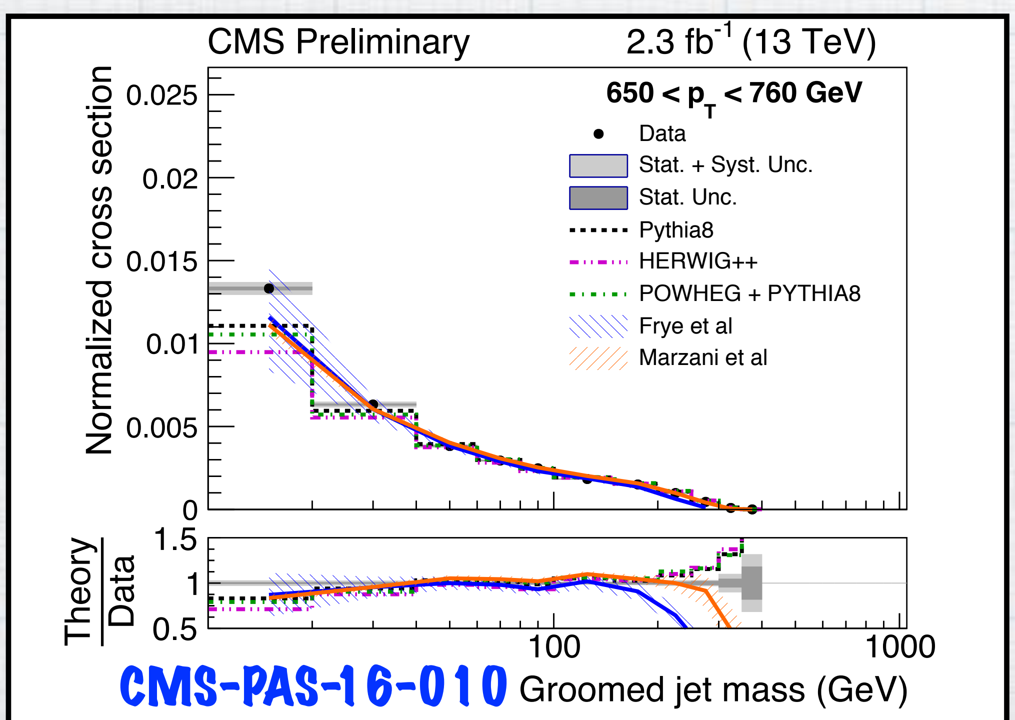
precision jet substructure

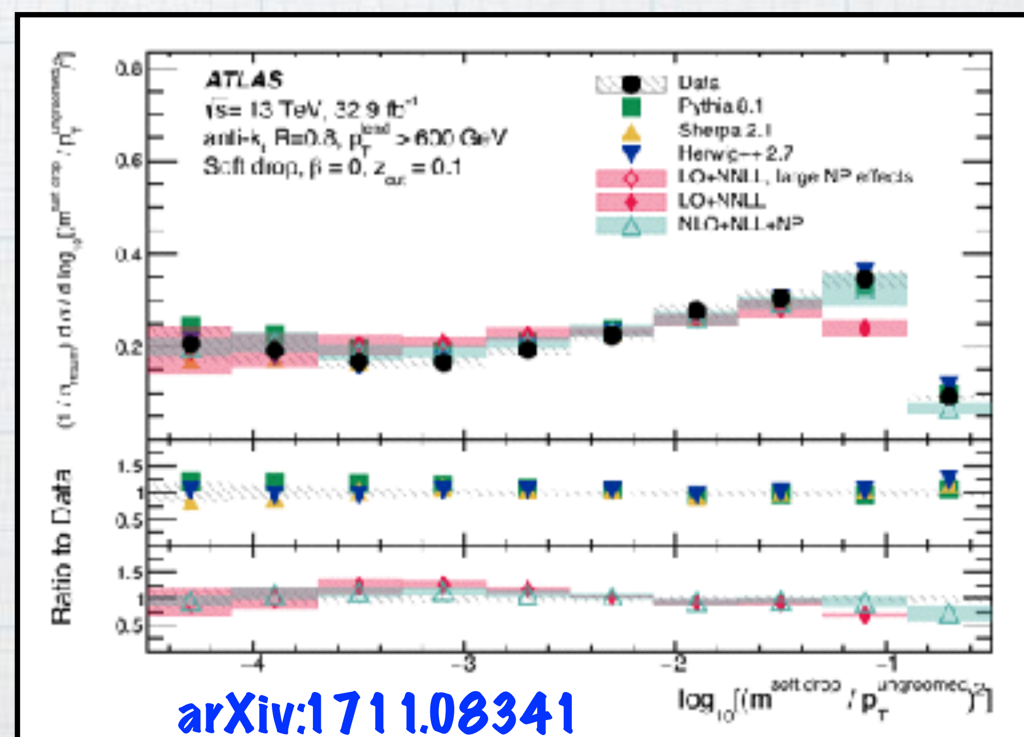


- using SCET, precision pushed to NNLL
- no non-global logs
- no colour correlations

Frye, Larkoski, Schwartz, Yan (2016)

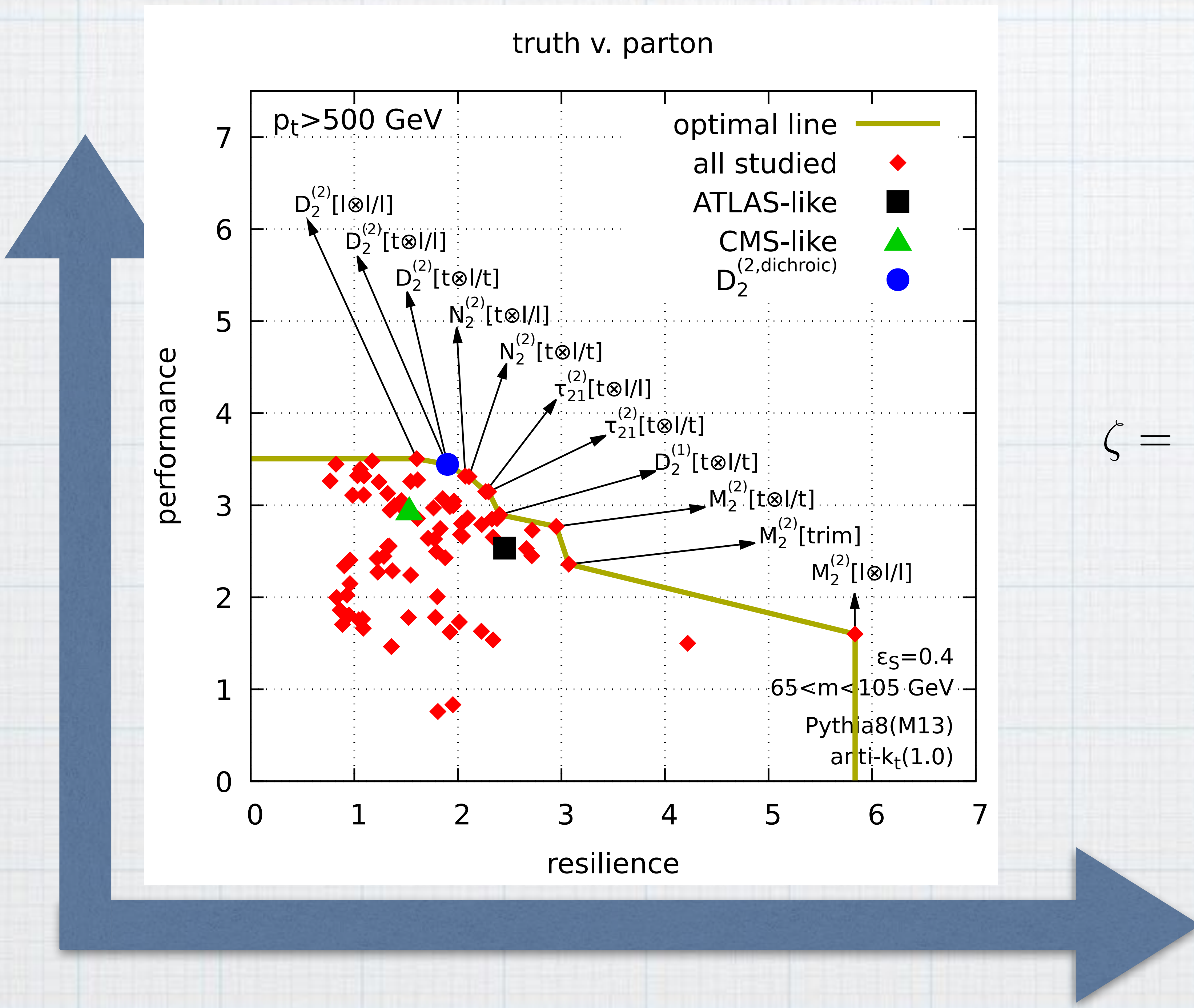
and data!





performance & resilience

more efficient



more robust

$$\zeta = \left(\frac{\Delta\epsilon_S^2}{\langle\epsilon\rangle_S^2} + \frac{\Delta\epsilon_B^2}{\langle\epsilon\rangle_B^2} \right)^{-1/2}$$

$$\Delta\epsilon_{S,B} = \epsilon_{S,B} - \epsilon'_{S,B},$$

$$\langle\epsilon\rangle_{S,B} = \frac{1}{2}\left(\epsilon_{S,B} + \epsilon'_{S,B}\right)$$

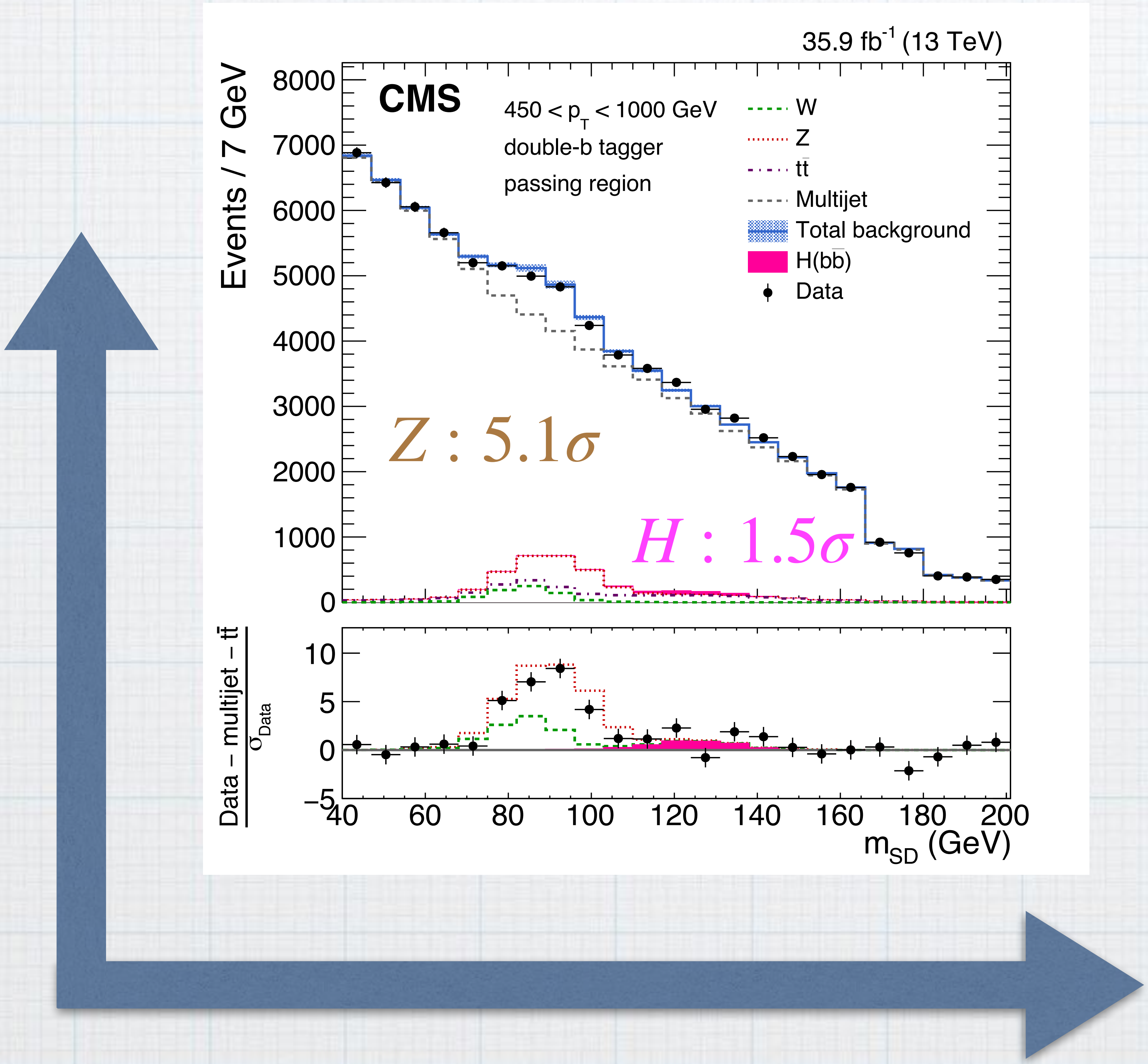
summary of lecture 3

more efficient

* boosted-object physics
* grooming and tagging
* precision substructure physics with soft-drop

more robust

summary of lecture 3



more efficient

more robust