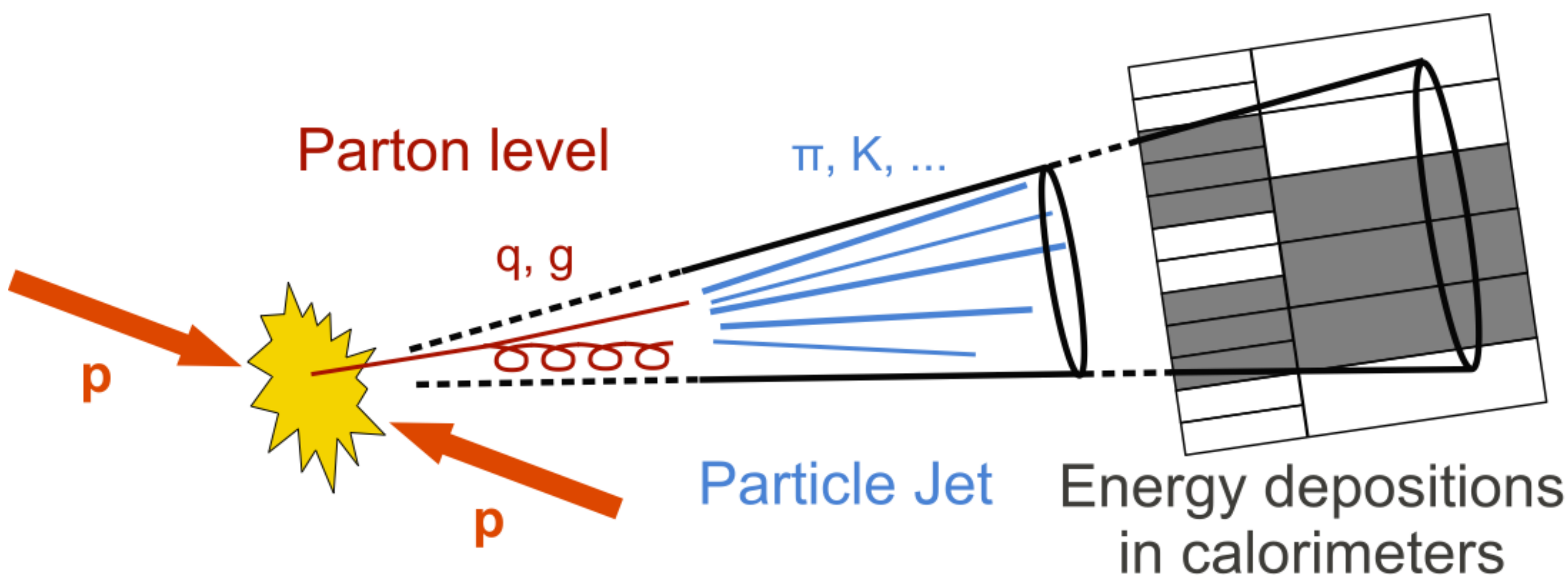
Jet Physics and Machine Learning: Lecture 3 : Jet substructure

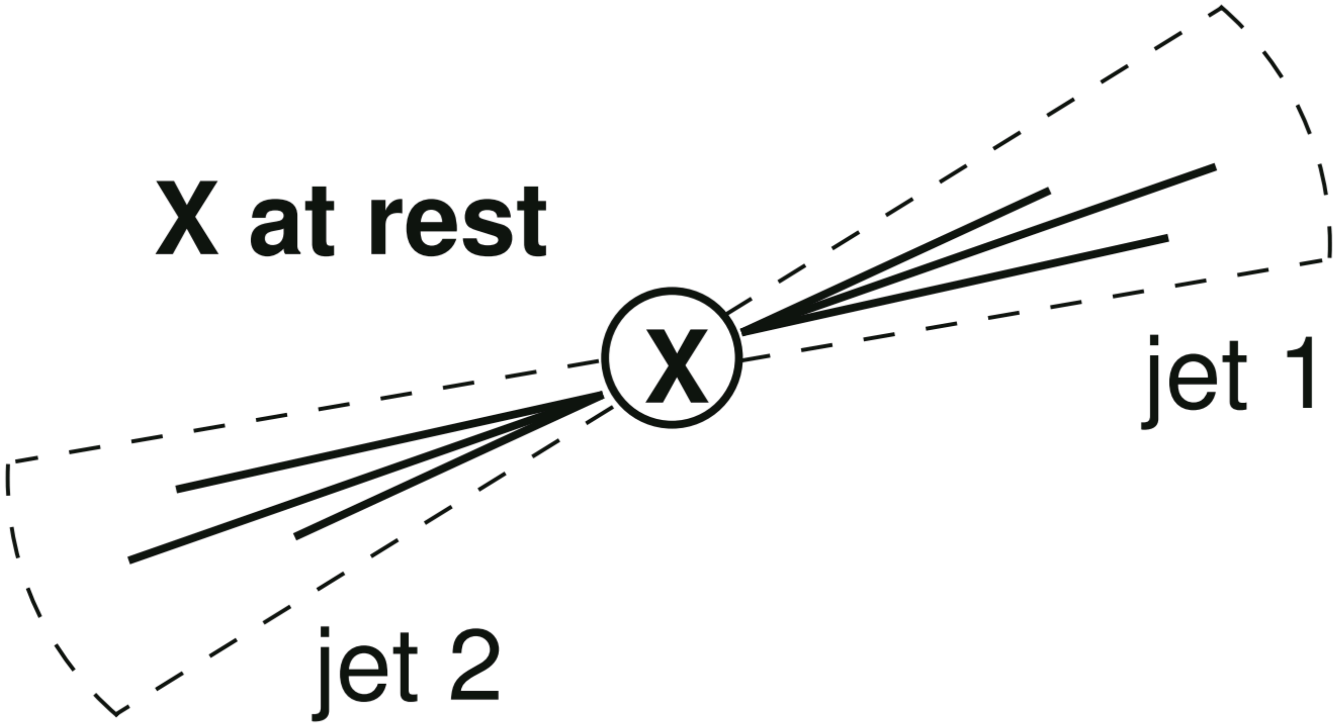
Sanmay Ganguly
IIT-Kanpur
sanmay@iitk.ac.in

27/02/2025

IMSc Spring School on High Energy Physics - 2025

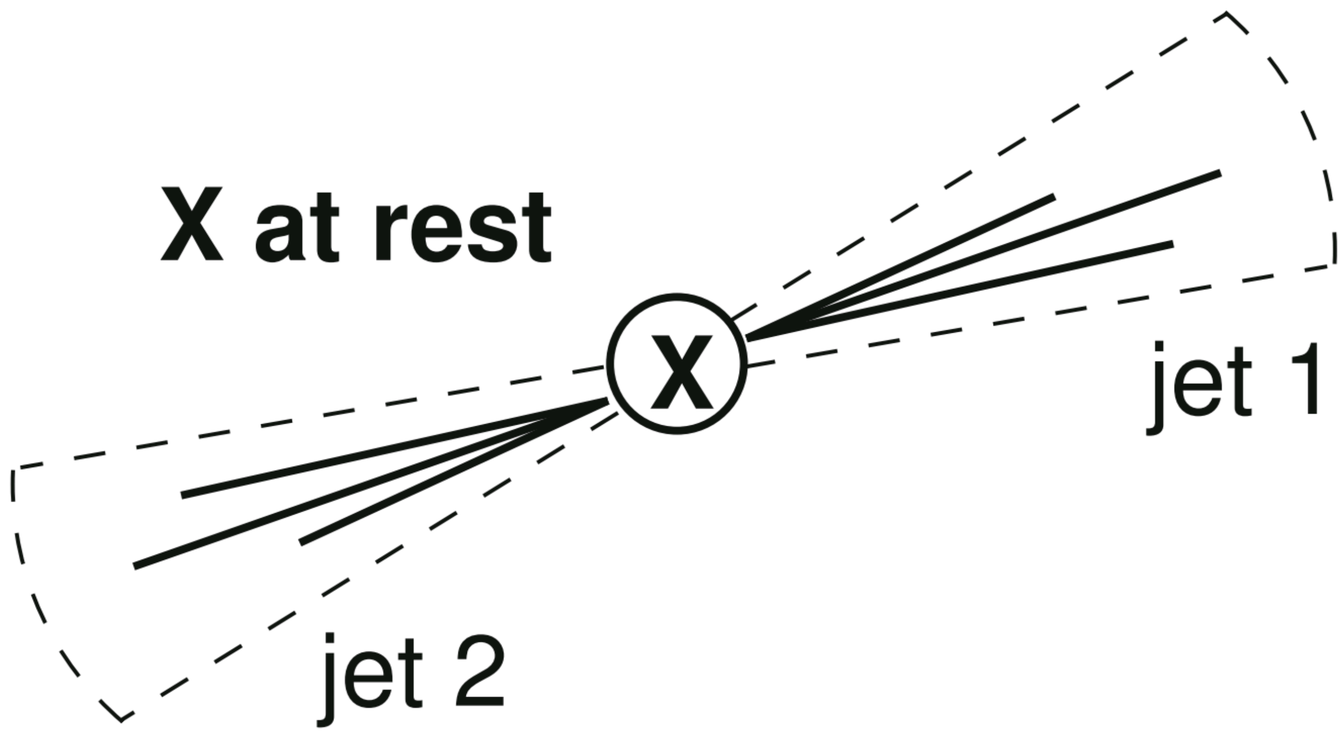


Hadronic decay of a heavy particle

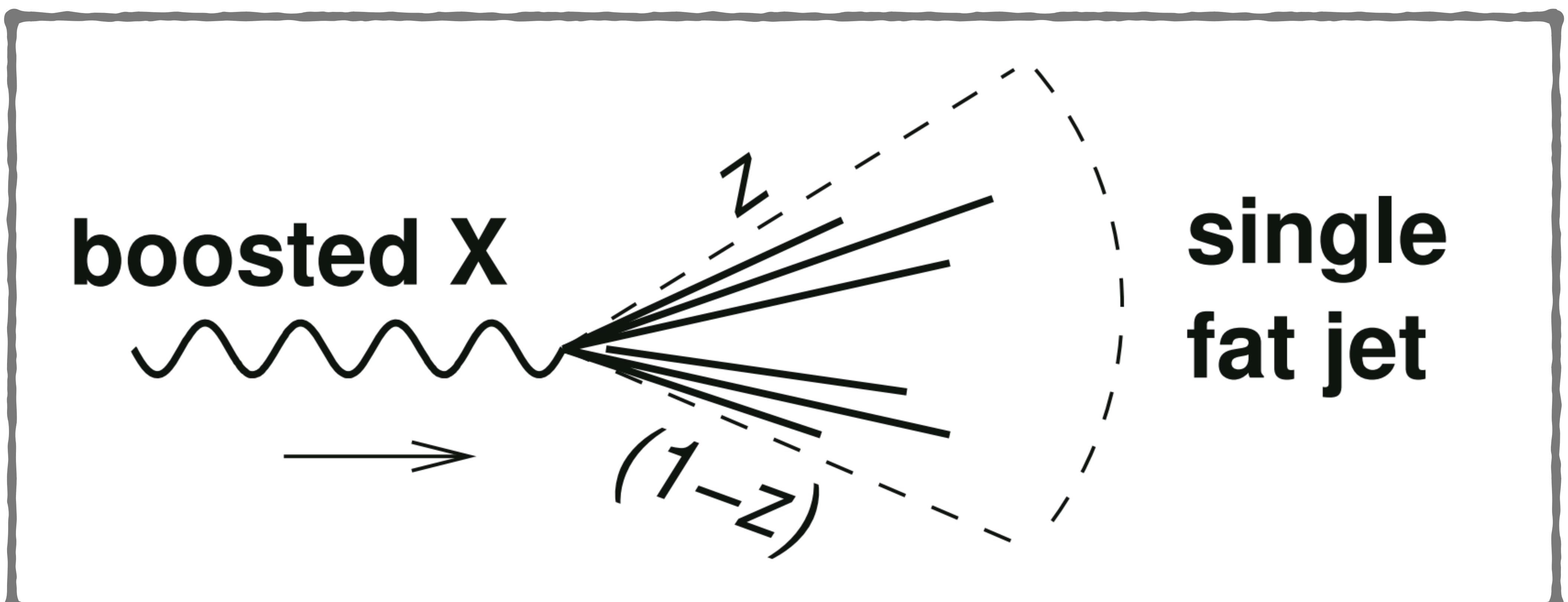


A heavy particle X decaying at rest

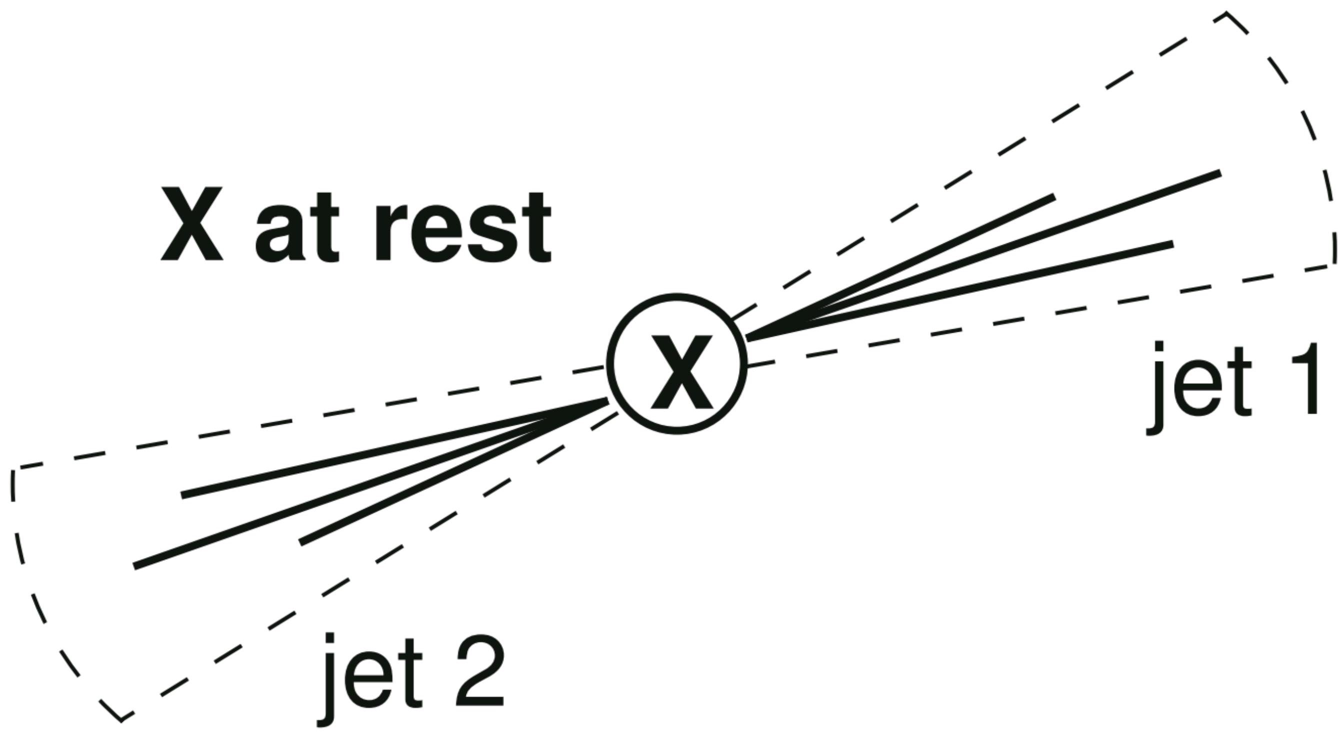
Hadronic decay of a heavy particle



A heavy particle X decaying at rest

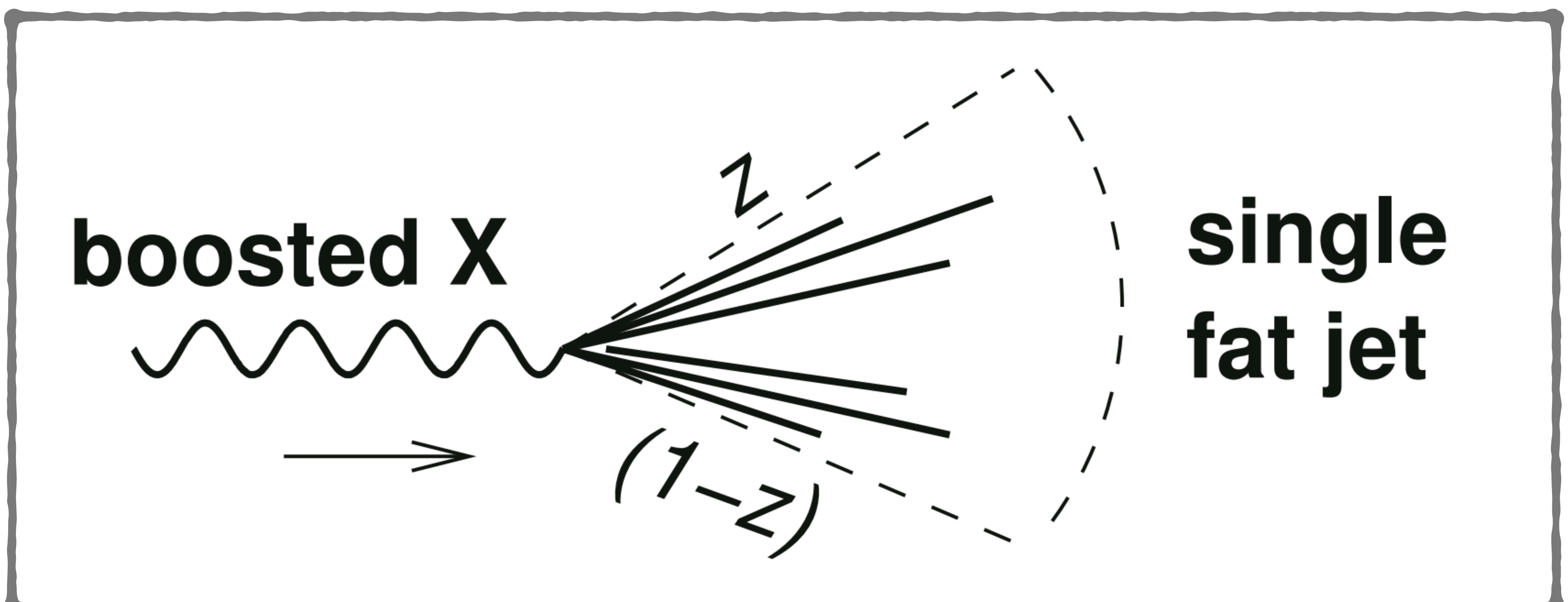


Hadronic decay of a heavy particle

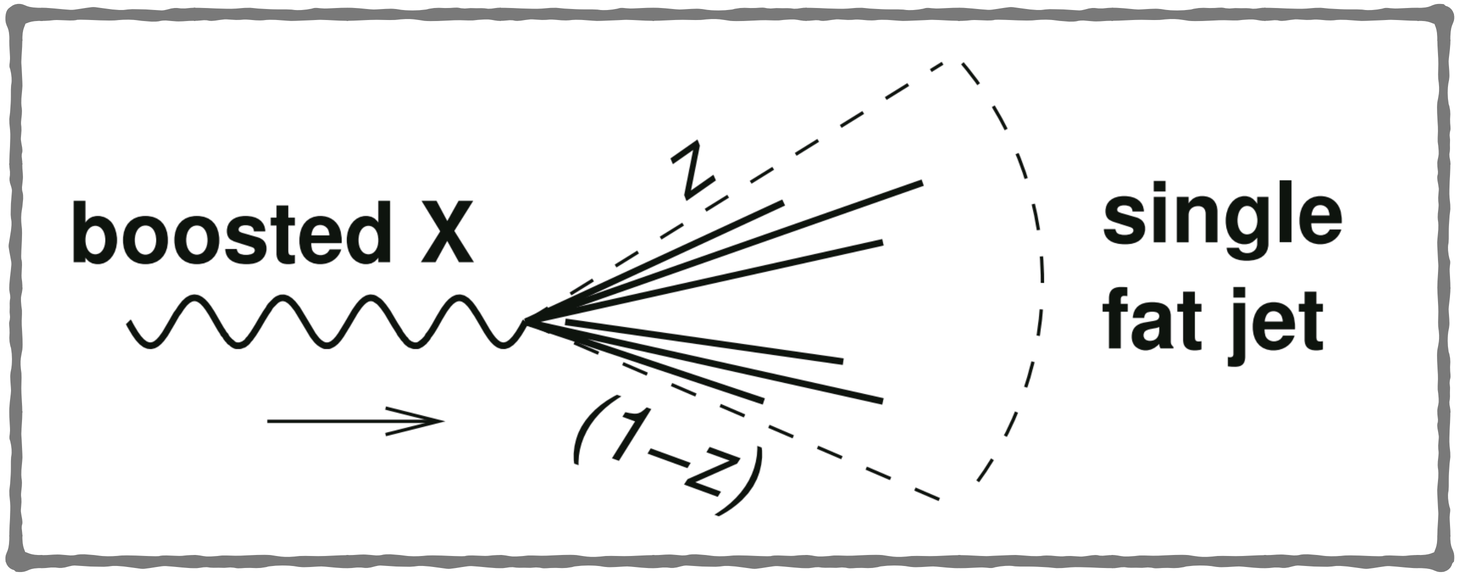


A heavy particle X decaying at rest

$$\Delta R \approx 2m/p_T$$



Hadronic decay of a heavy particle



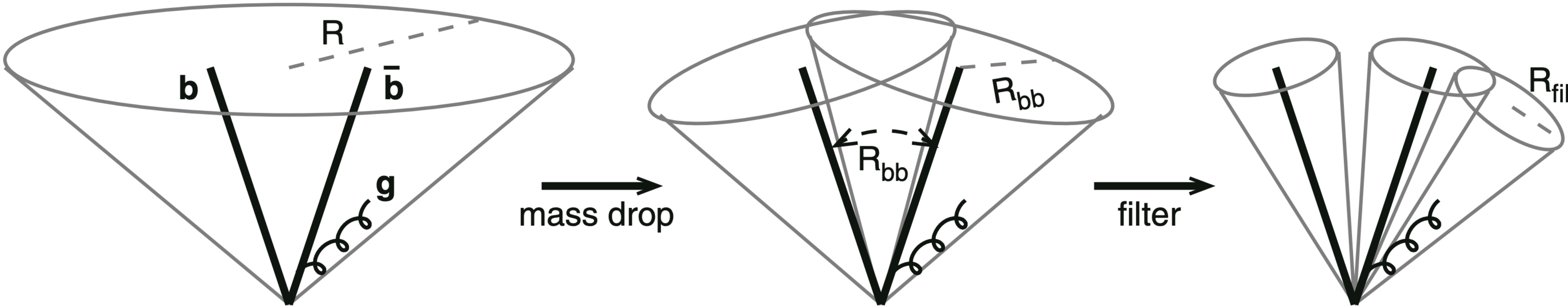
Electroweak particle X decaying hadronically into two jets.

$$m_X^2 = E_X^2 \; z(1-z) \; 2(1 - cos\theta)$$

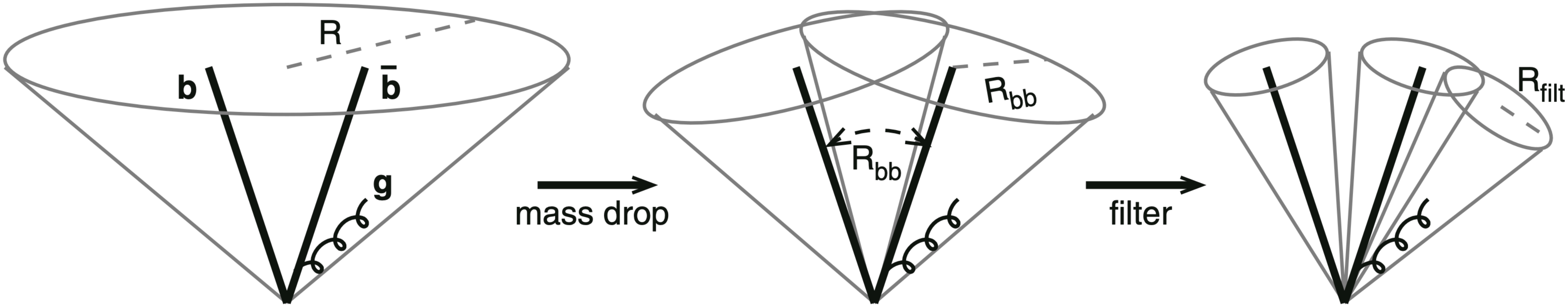
The two prongs end up in a single jet if :

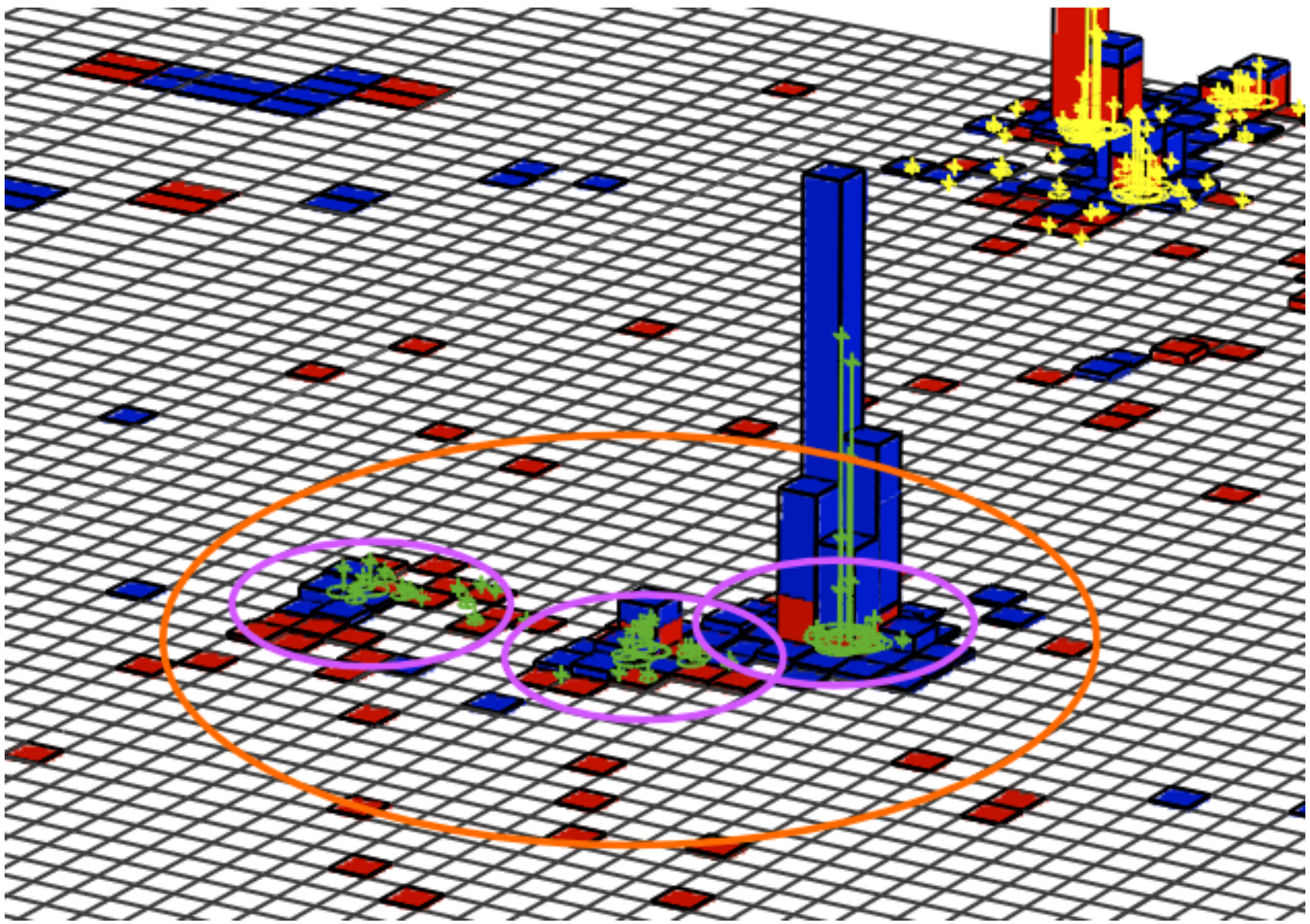
$$\Delta R \sim \frac{m}{p_T} \frac{1}{\sqrt{z(1-z)}} \sim \frac{2m}{p_T}$$

Initially in the context of $H \to b\bar{b}$ / $t \to bW$

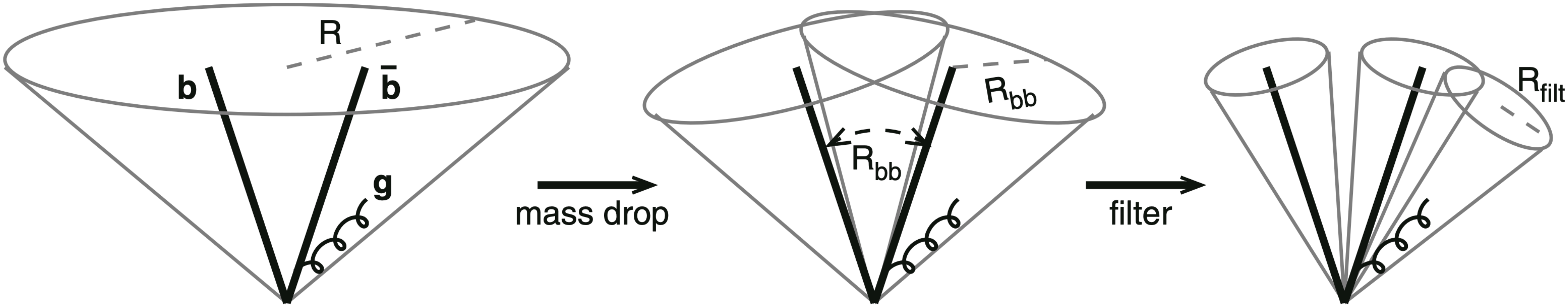


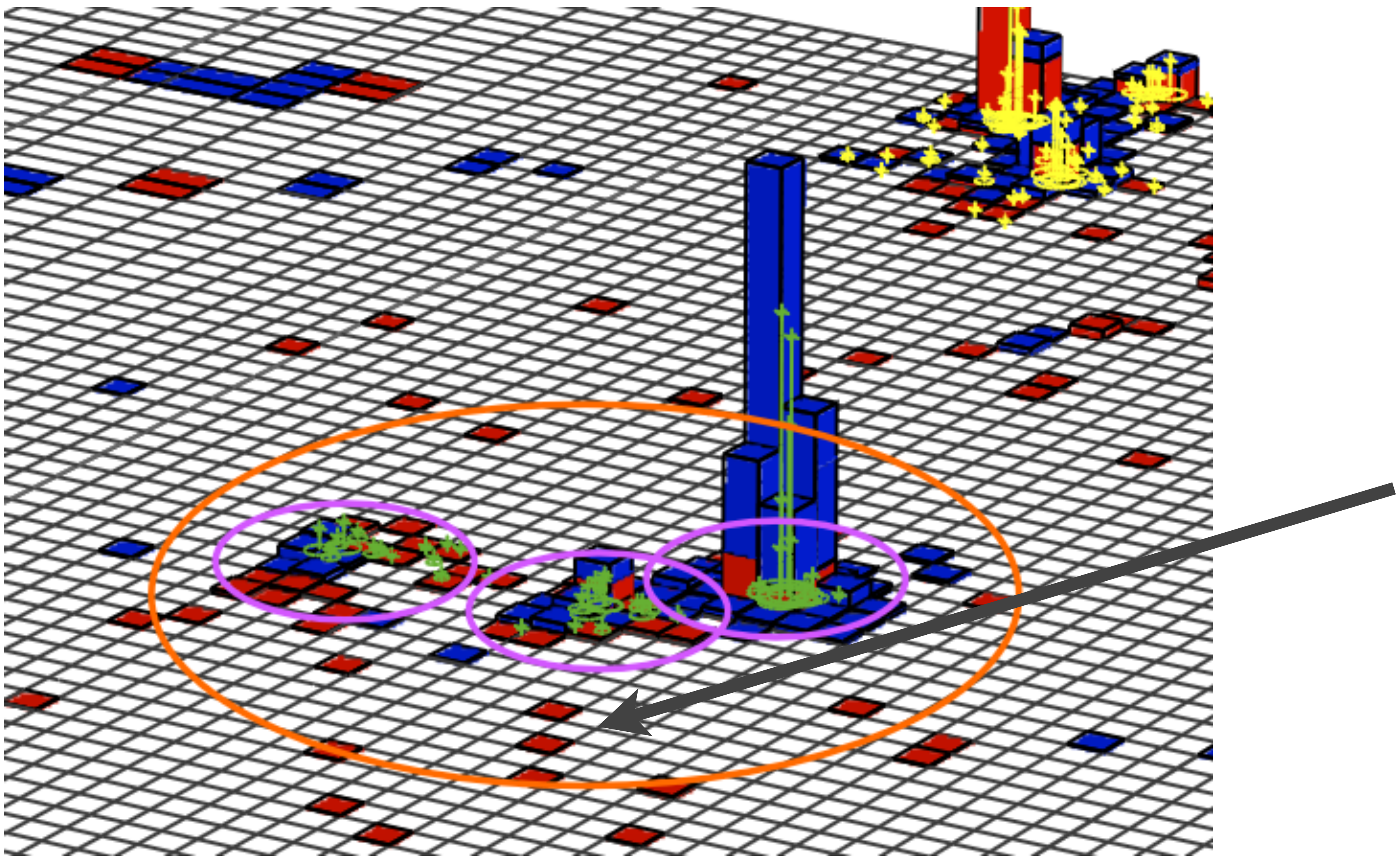
Initially in the context of $H \to b\bar{b}$ / $t \to bW$



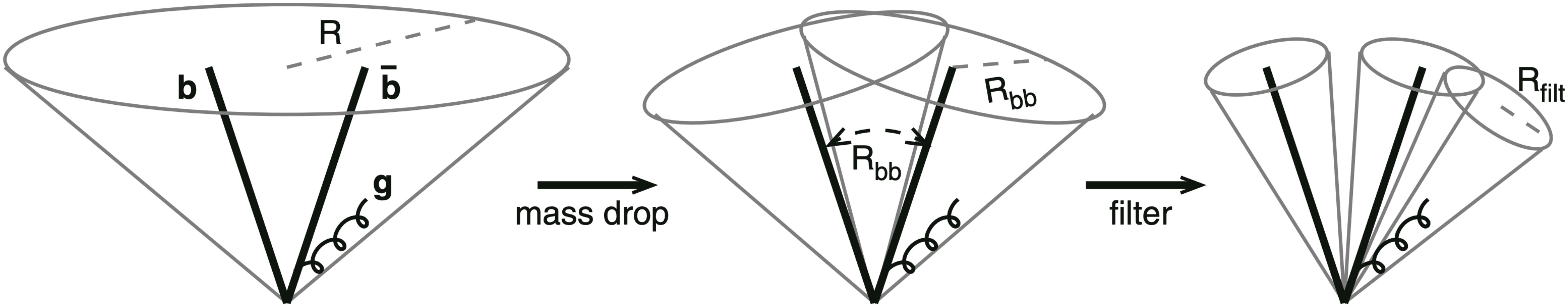


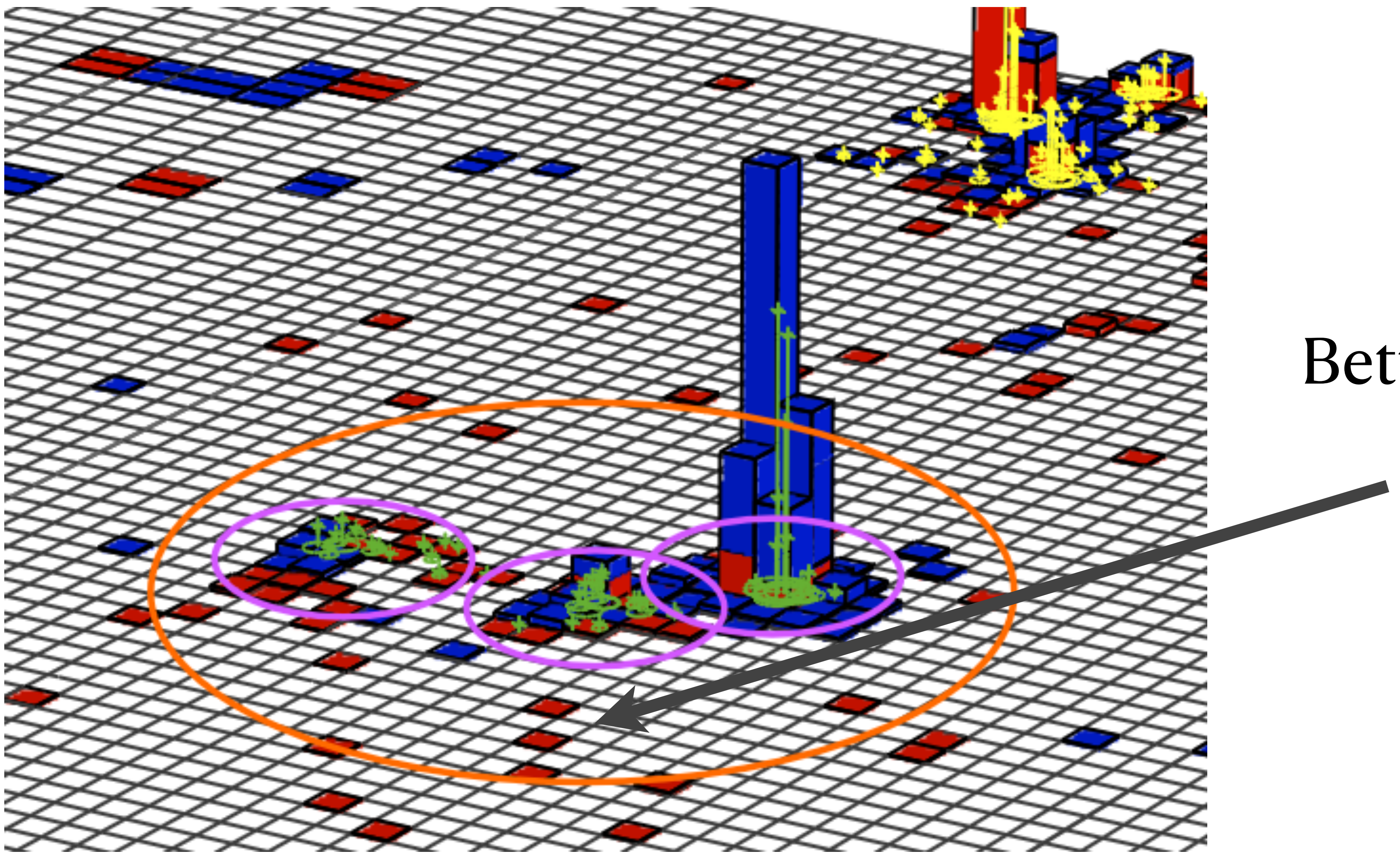
Initially in the context of $H \to b\bar{b}$ / $t \to bW$





Initially in the context of $H \to b\bar{b}$ / $t \to bW$

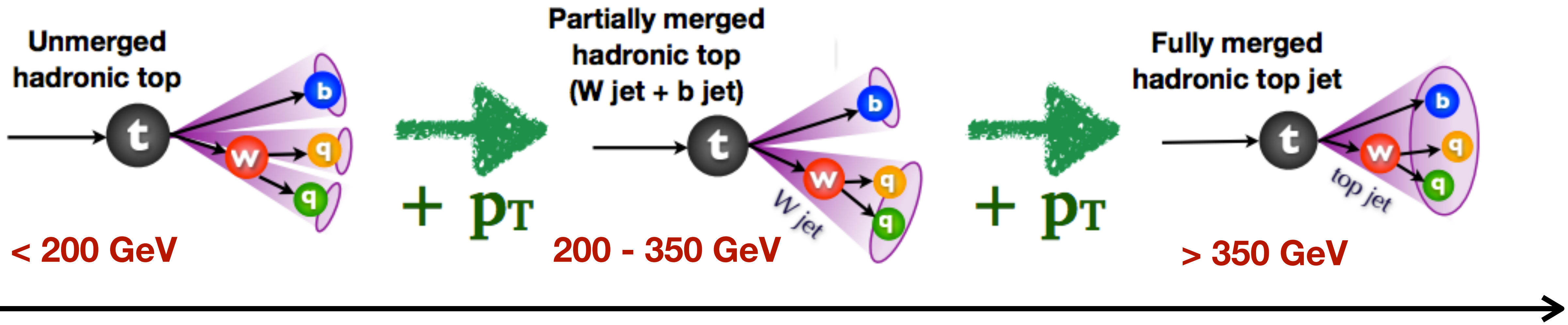


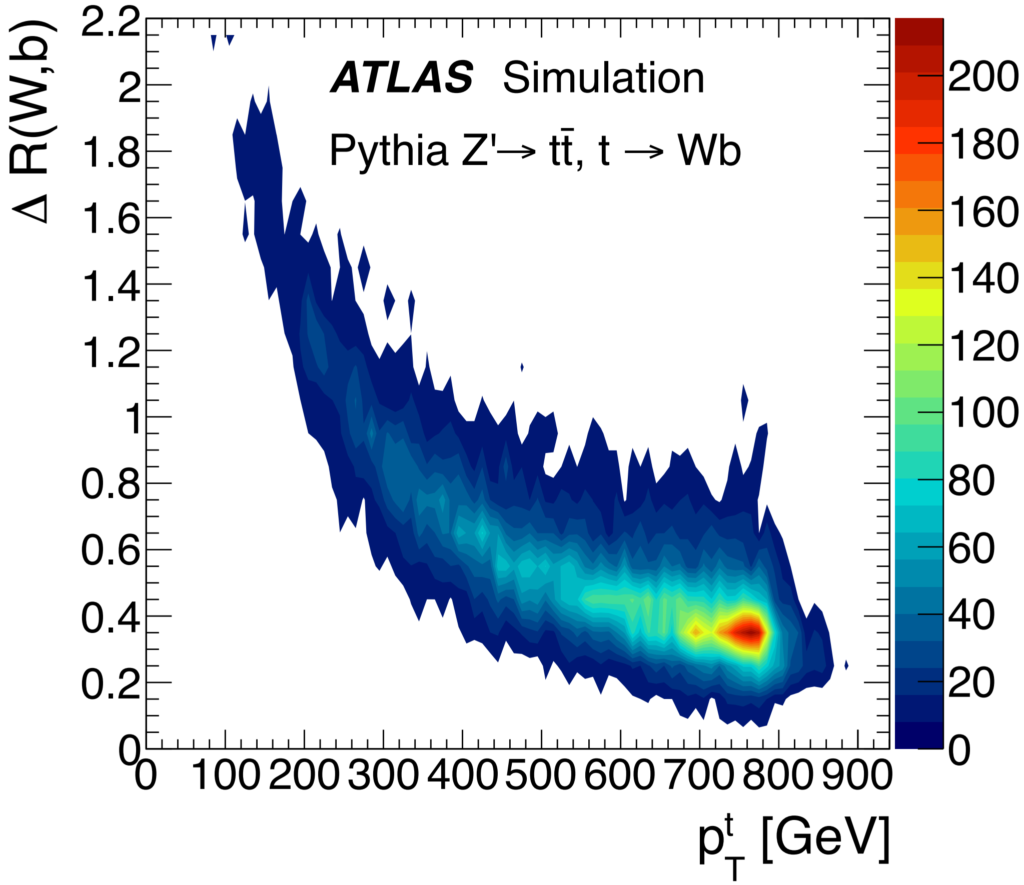


Better we get rid of this

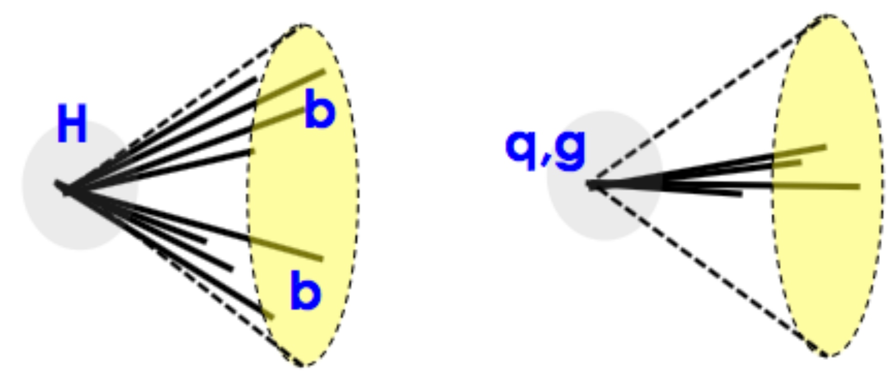
Feature of boosted particle decay

for $\Delta R \sim 1.0$





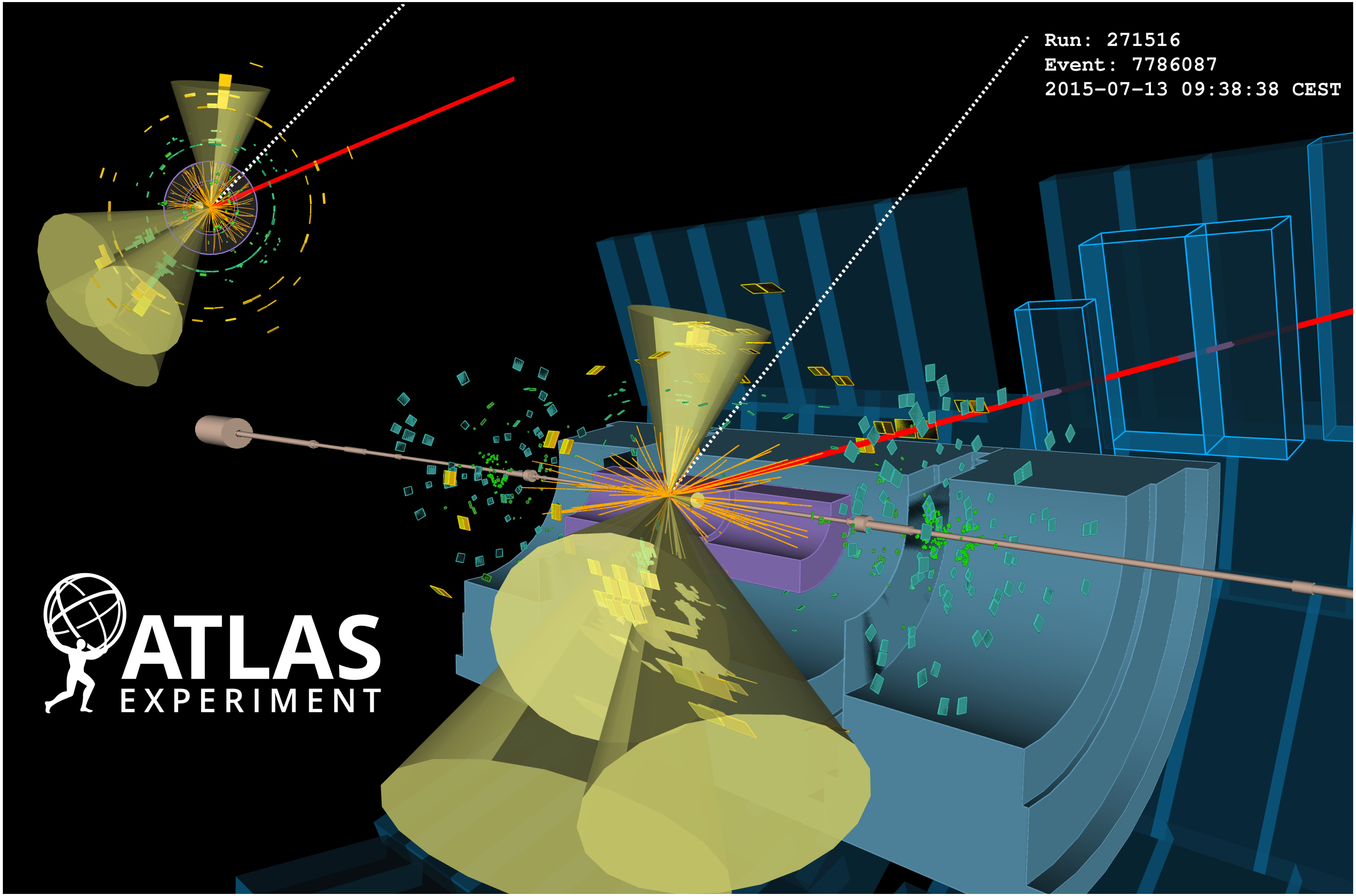
$\Delta R \sim 2m/p_T$



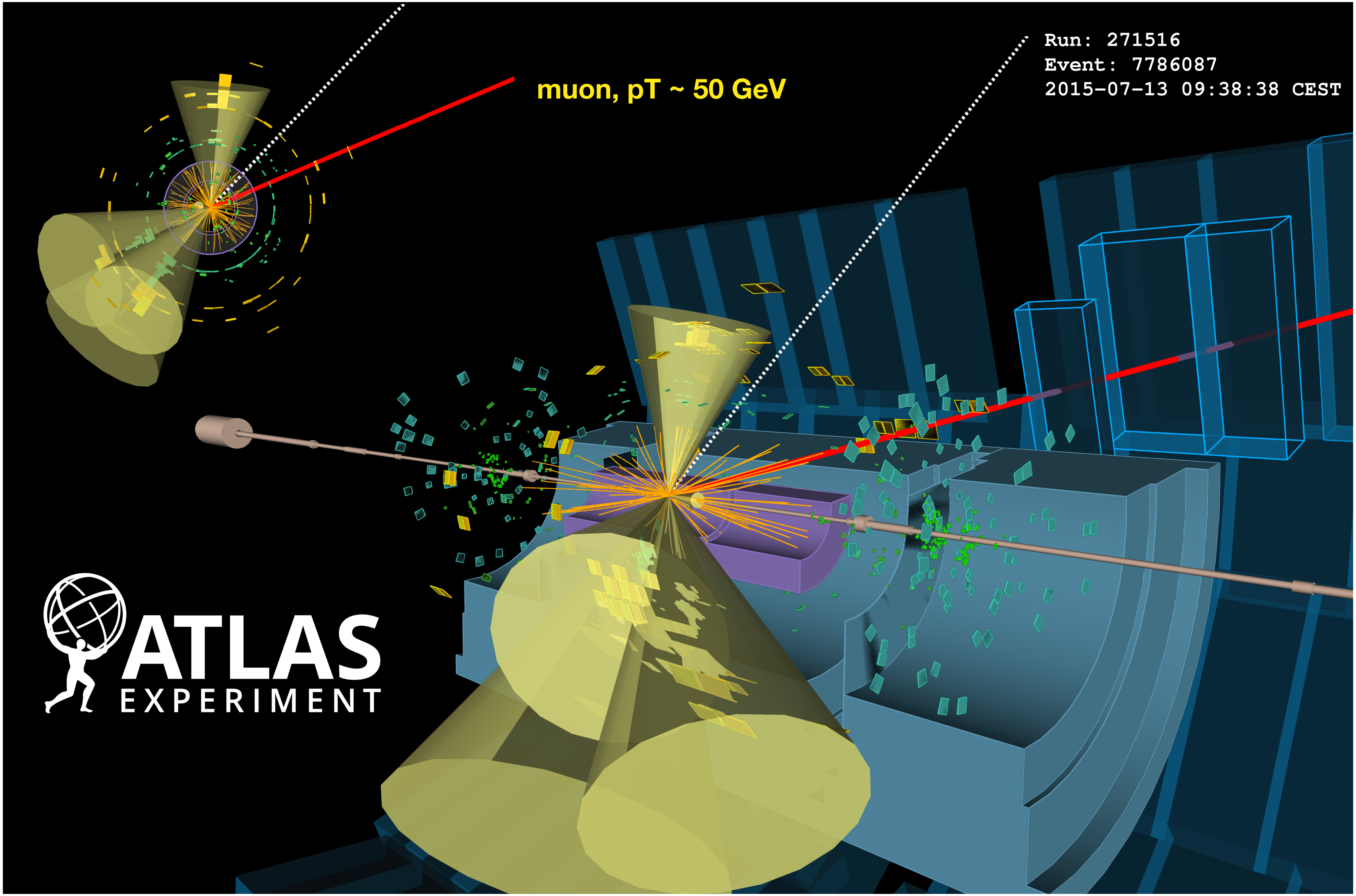
- Large R jets originating from decay of boosted heavy particle have different characteristics compared to fat jets from light flavor quarks or gluons.

- The major challenge with large R jets are heavy contamination due to pileup & large QCD background.

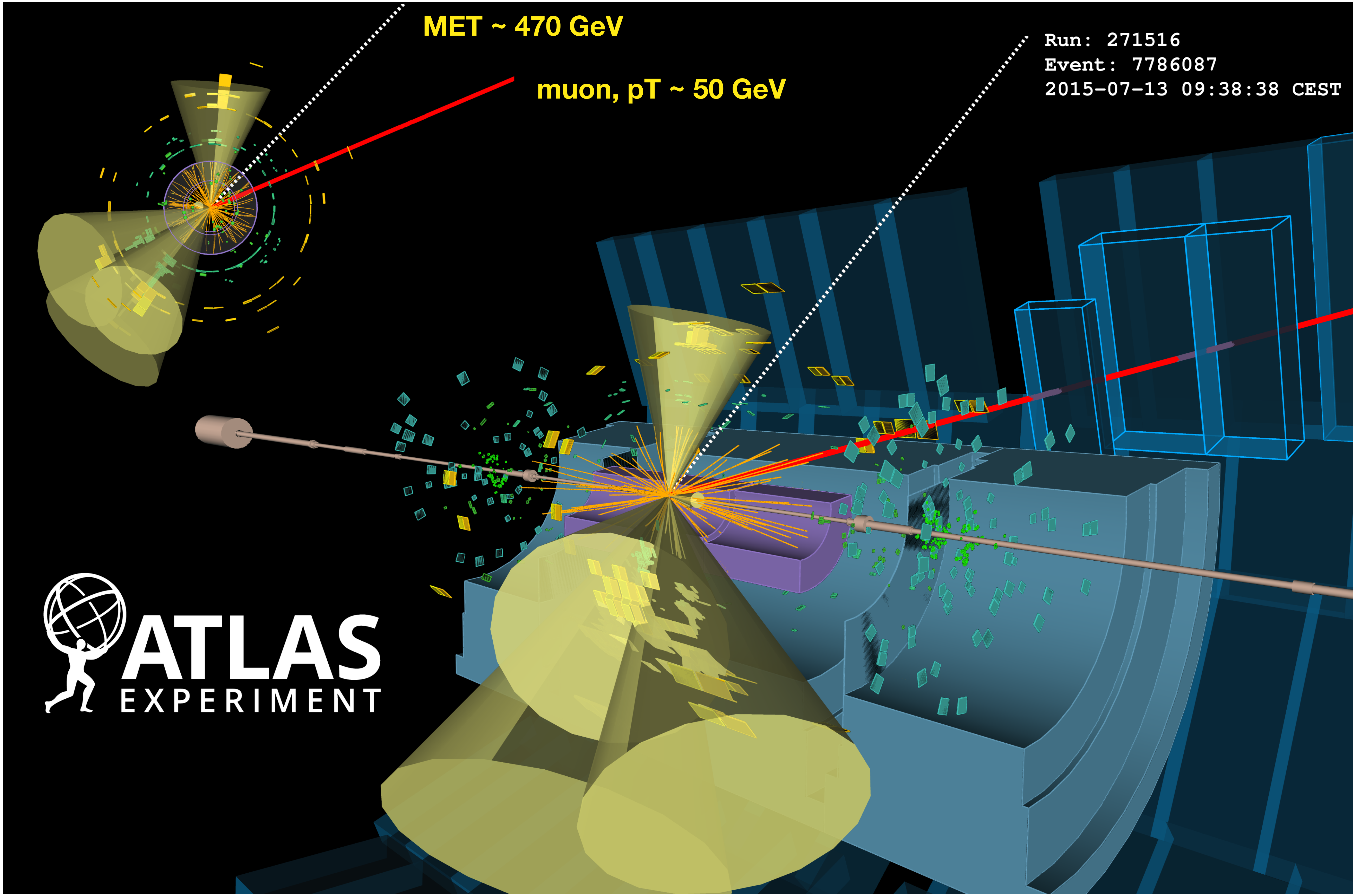
Event with boosted jets



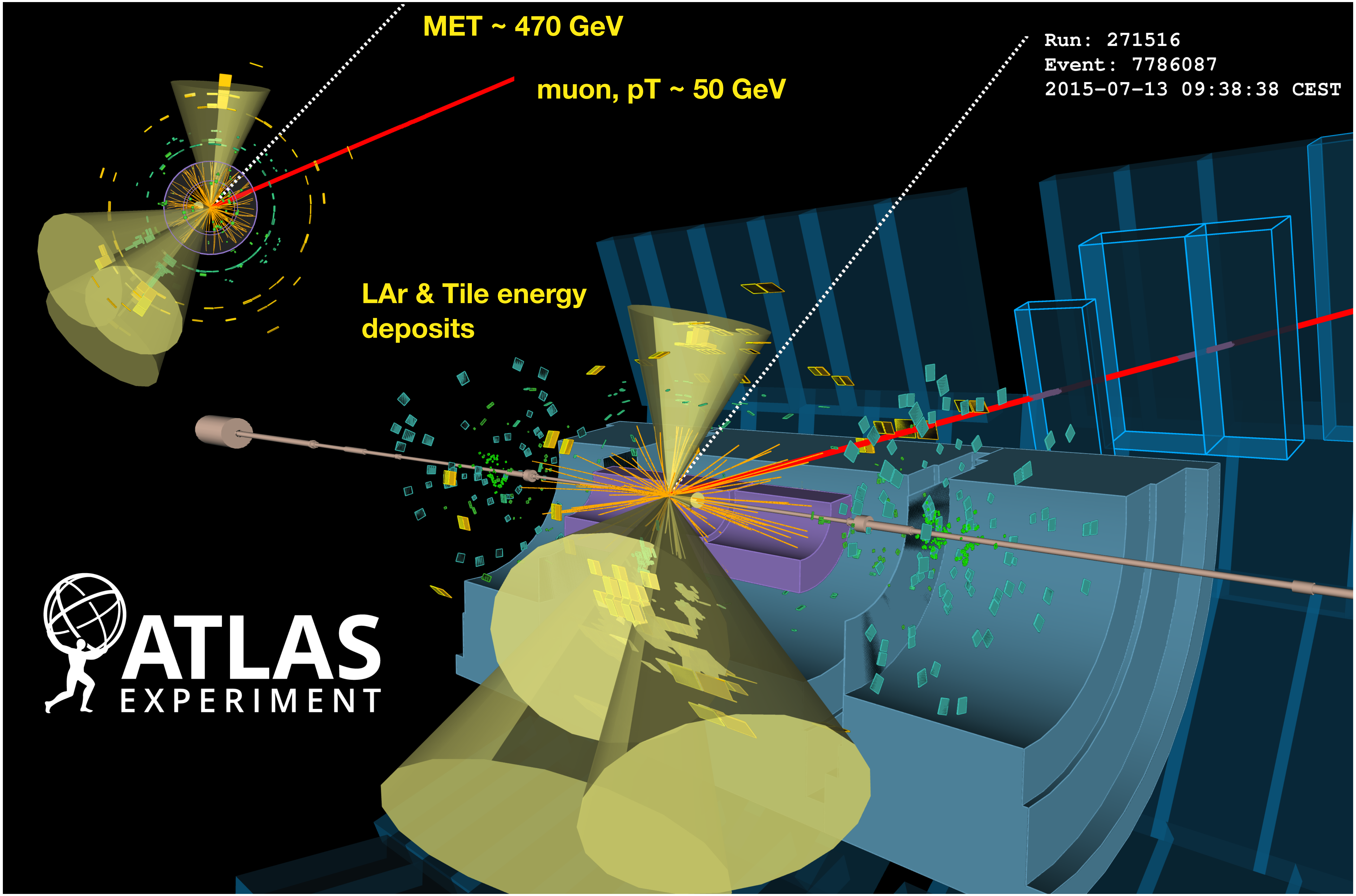
Event with boosted jets



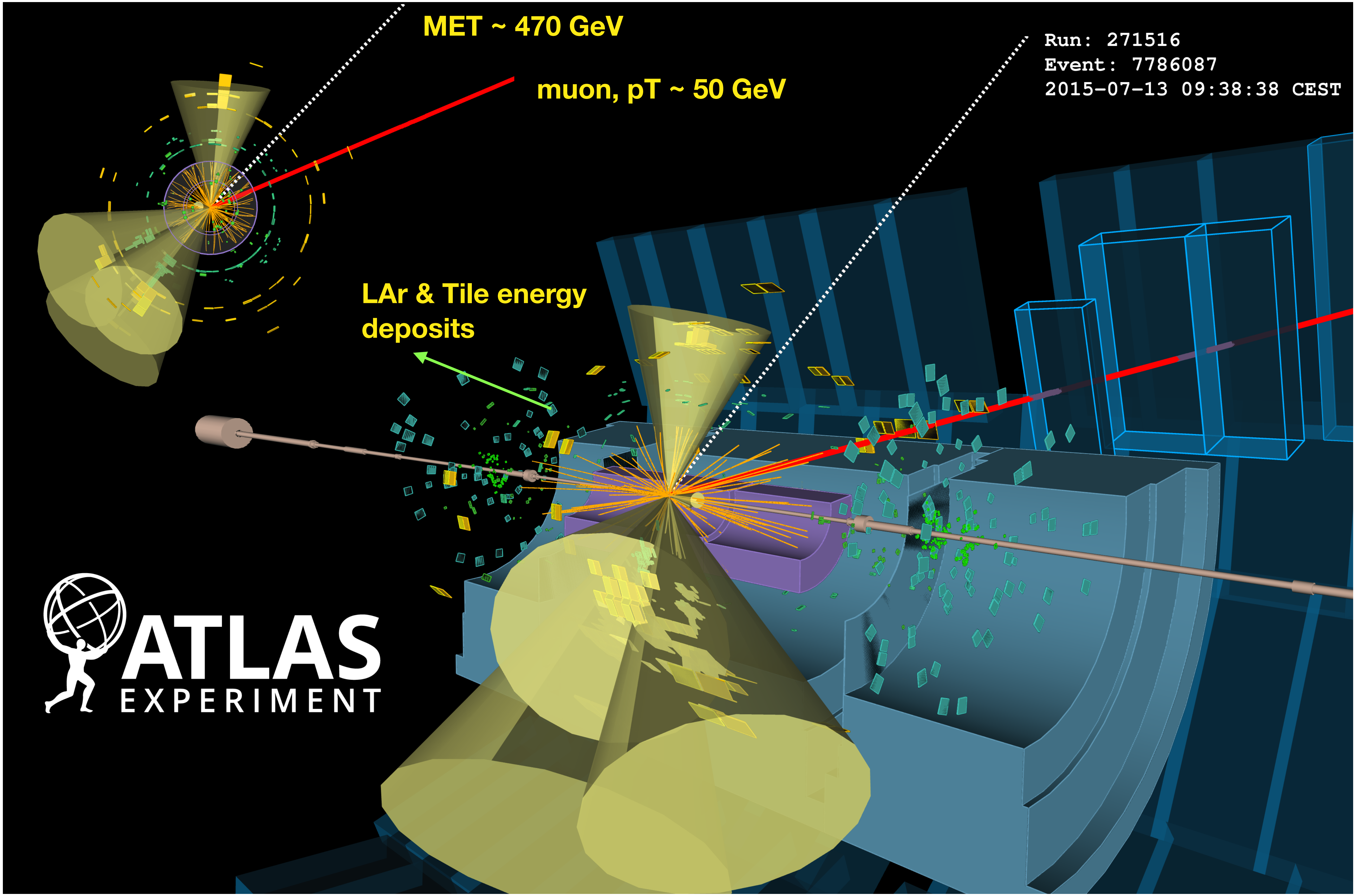
Event with boosted jets



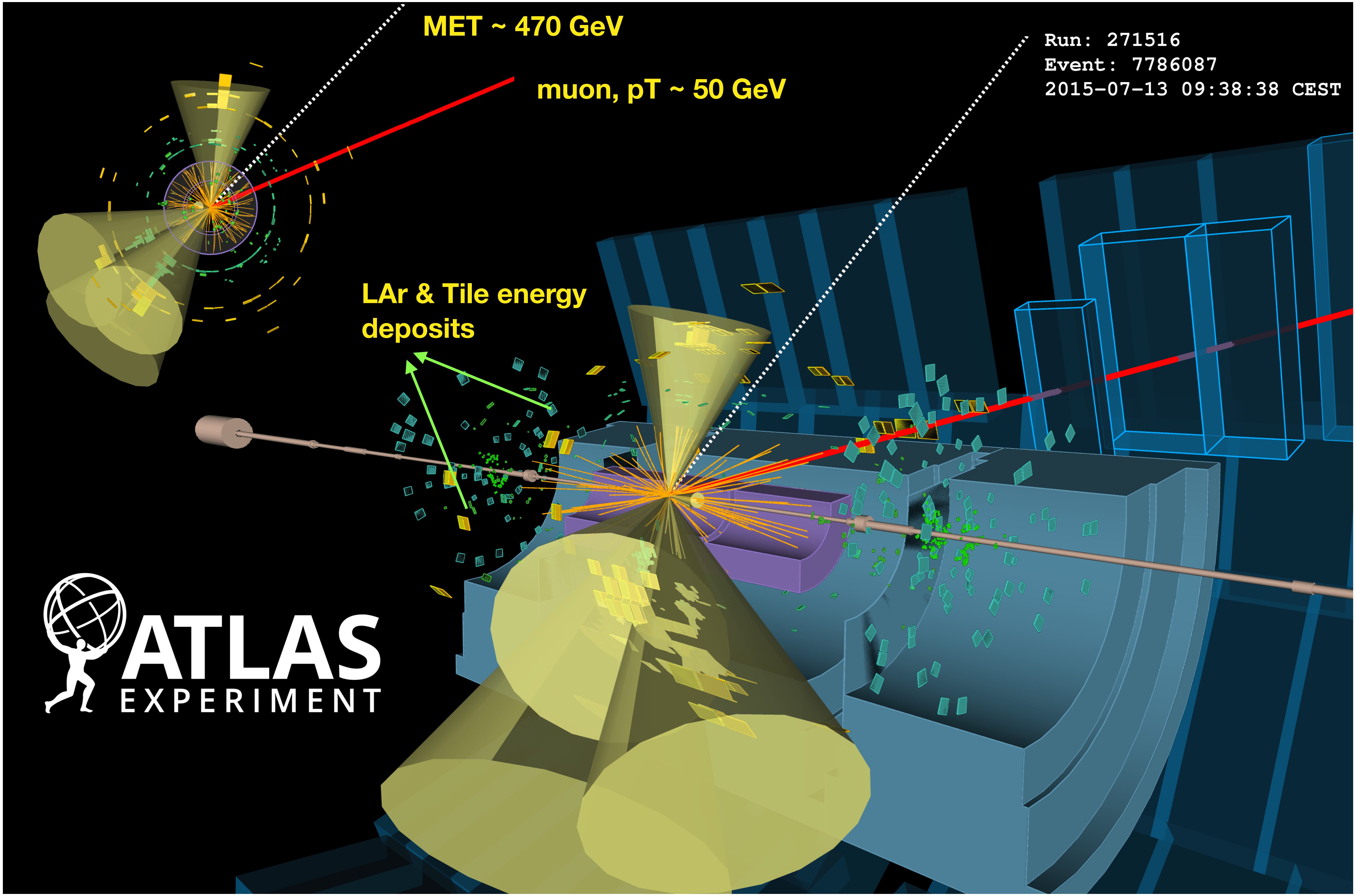
Event with boosted jets



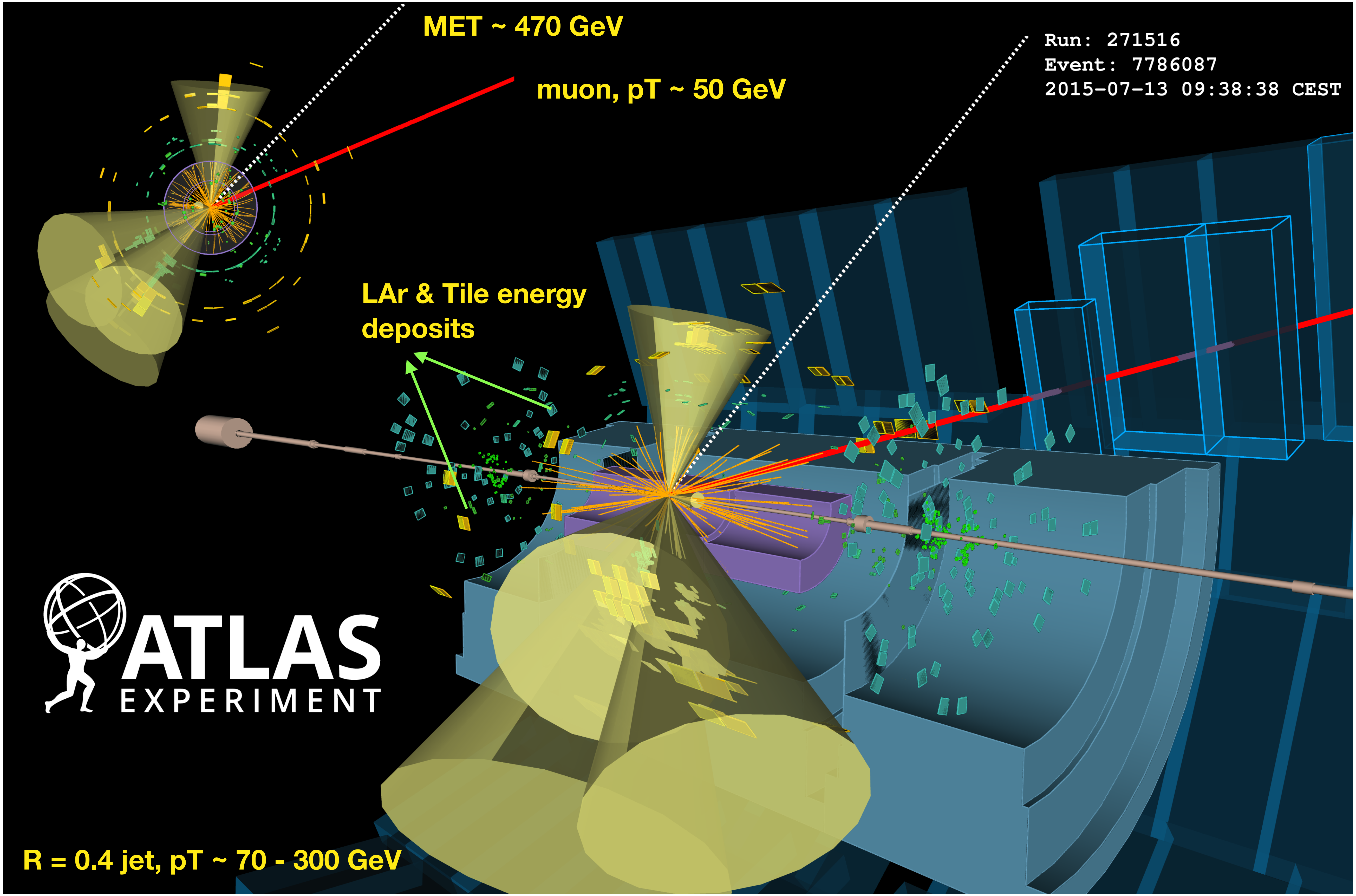
Event with boosted jets



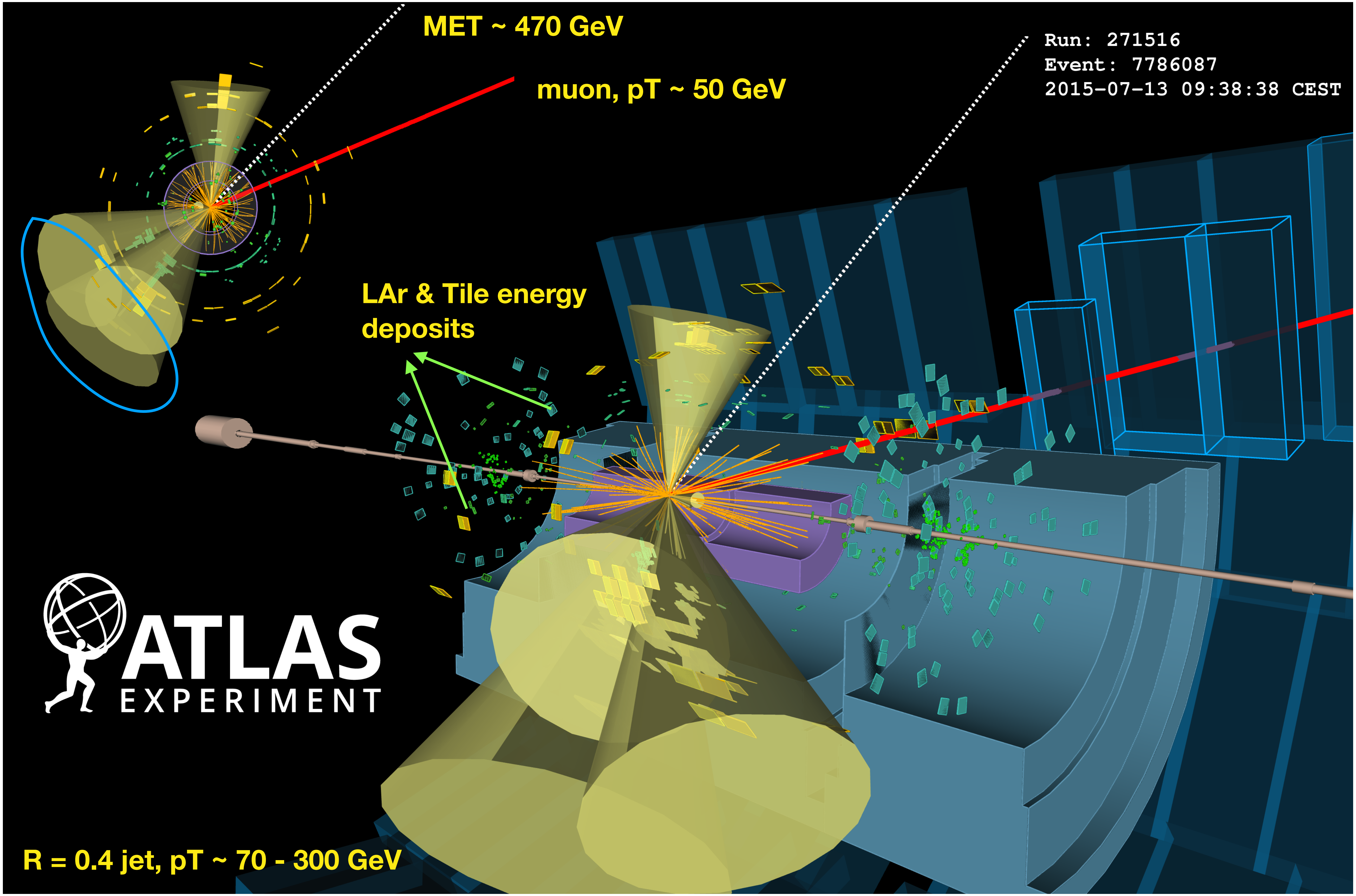
Event with boosted jets



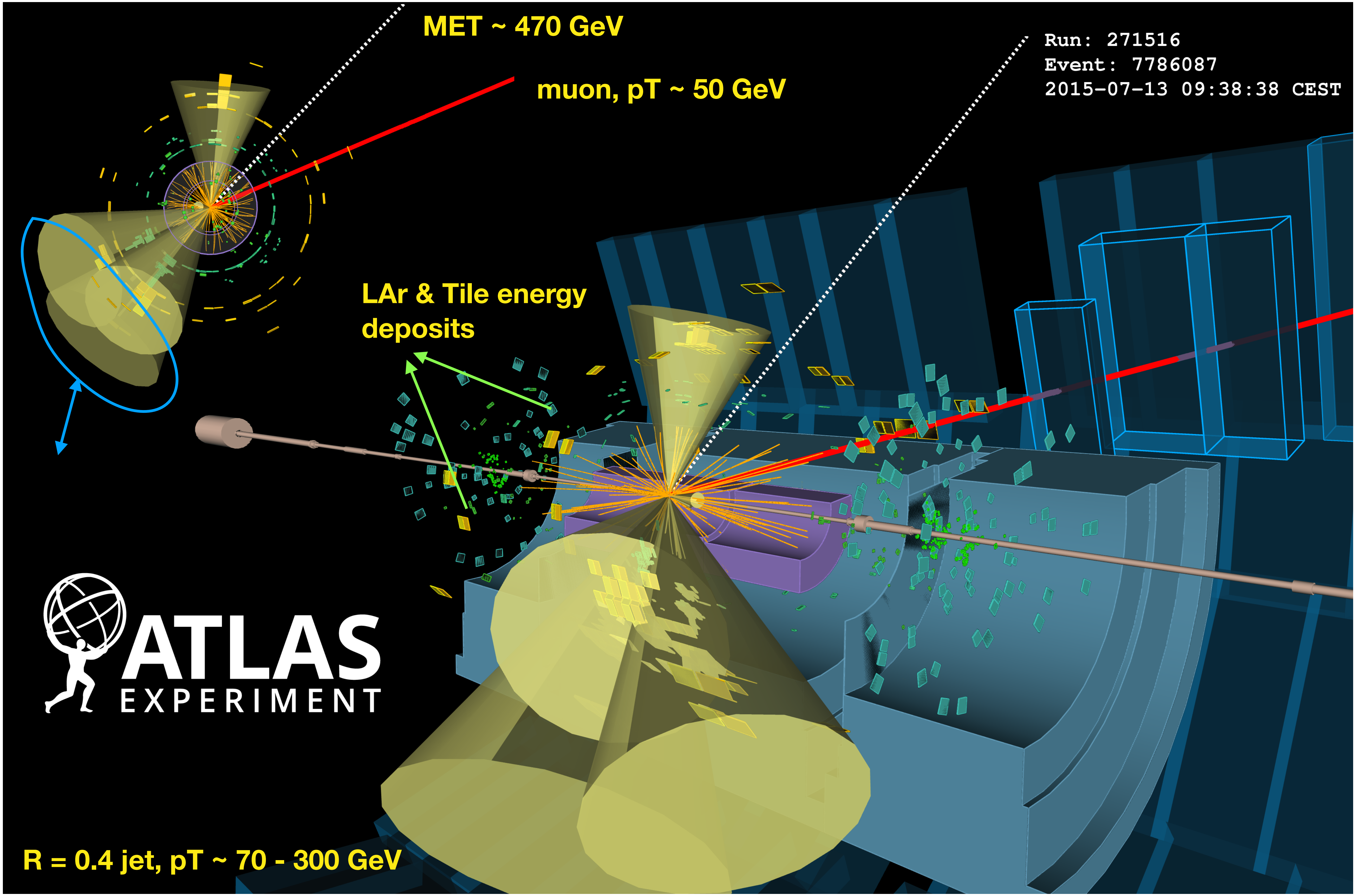
Event with boosted jets



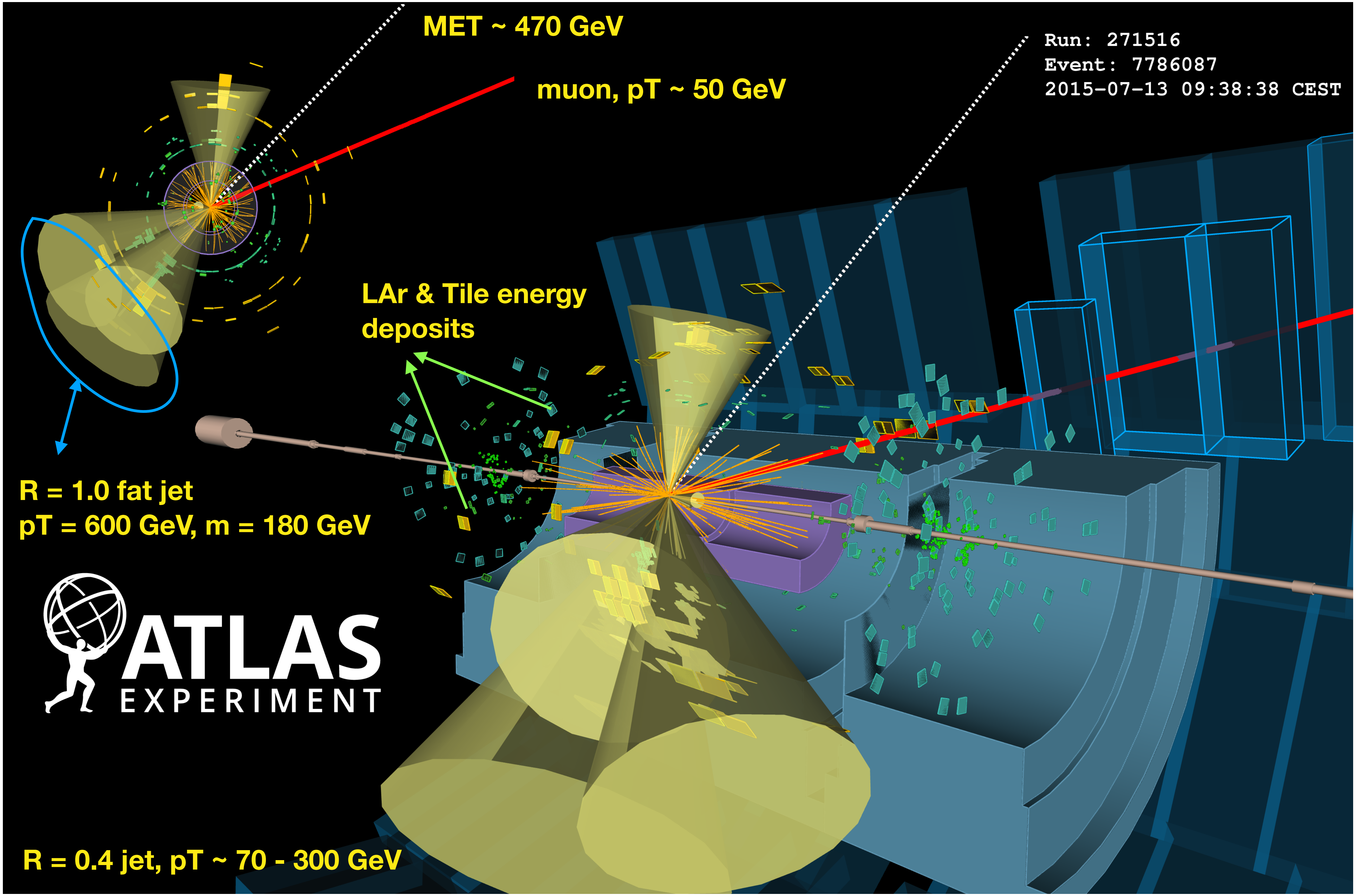
Event with boosted jets



Event with boosted jets



Event with boosted jets



Jet mass : the 0-th jet substructure

The variable of interest is $\frac{1}{\sigma}\frac{d\sigma}{dv}$, where $v = \frac{m_j^2}{E_j^2}$

Jet mass : the 0-th jet substructure

The variable of interest is $\frac{1}{\sigma}\frac{d\sigma}{dv}$, where $v = \frac{m_j^2}{E_j^2}$

$$\frac{v}{\sigma}\frac{d\sigma^{(\mathrm{plain})}}{dv} = \frac{\alpha_s}{\pi}\left(a_{12}L + a_{11} + \mathcal{O}(v)\right) + \left(\frac{\alpha_s}{\pi}\right)^2\left(a_{24}L^3 + a_{23}L^2 + a_{22}L + \mathcal{O}\left(v^0\right)\right) + \mathcal{O}\left(\alpha_s^3\right), \quad L \equiv \ln\frac{R^2}{v}$$

Jet mass : the 0-th jet substructure

The variable of interest is $\frac{1}{\sigma}\frac{d\sigma}{dv}$, where $v = \frac{m_j^2}{E_j^2}$

$$\frac{v}{\sigma}\frac{d\sigma^{(\text{plain})}}{dv} = \frac{\alpha_s}{\pi}\left(a_{12}L + a_{11} + \mathcal{O}(v)\right)$$

In small v limit

$$+ \left(\frac{\alpha_s}{\pi}\right)^2\left(a_{24}L^3 + a_{23}L^2 + a_{22}L + \mathcal{O}\left(v^0\right)\right) + \mathcal{O}\left(\alpha_s^3\right), \quad L \equiv \ln\frac{R^2}{v}$$

Jet mass : the 0-th jet substructure

The variable of interest is $\frac{1}{\sigma}\frac{d\sigma}{dv}$, where $v = \frac{m_j^2}{E_j^2}$

$$\frac{v}{\sigma}\frac{d\sigma^{(\text{plain})}}{dv} = \frac{\alpha_s}{\pi}\left(a_{12}L + a_{11} + \mathcal{O}(v)\right) + \left(\frac{\alpha_s}{\pi}\right)^2\left(a_{24}L^3 + a_{23}L^2 + a_{22}L + \mathcal{O}\left(v^0\right)\right) + \mathcal{O}\left(\alpha_s^3\right), \; L \equiv \ln\frac{R^2}{v}$$

In small v limit

$$\Sigma(v) = \frac{1}{\sigma}\int^{v} dv' \frac{d\sigma}{dv'}.$$

Jet mass : the 0-th jet substructure

The variable of interest is $\frac{1}{\sigma}\frac{d\sigma}{dv}$, where $v = \frac{m_j^2}{E_j^2}$

$$\frac{v}{\sigma}\frac{d\sigma^{(\text{plain})}}{dv} = \frac{\alpha_s}{\pi}\left(a_{12}L + a_{11} + \mathcal{O}(v)\right) + \left(\frac{\alpha_s}{\pi}\right)^2\left(a_{24}L^3 + a_{23}L^2 + a_{22}L + \mathcal{O}\left(v^0\right)\right) + \mathcal{O}\left(\alpha_s^3\right), \quad L \equiv \ln\frac{R^2}{v}$$

In small v limit

$$\Sigma(v) = \frac{1}{\sigma}\int^v dv' \frac{d\sigma}{dv'}.$$

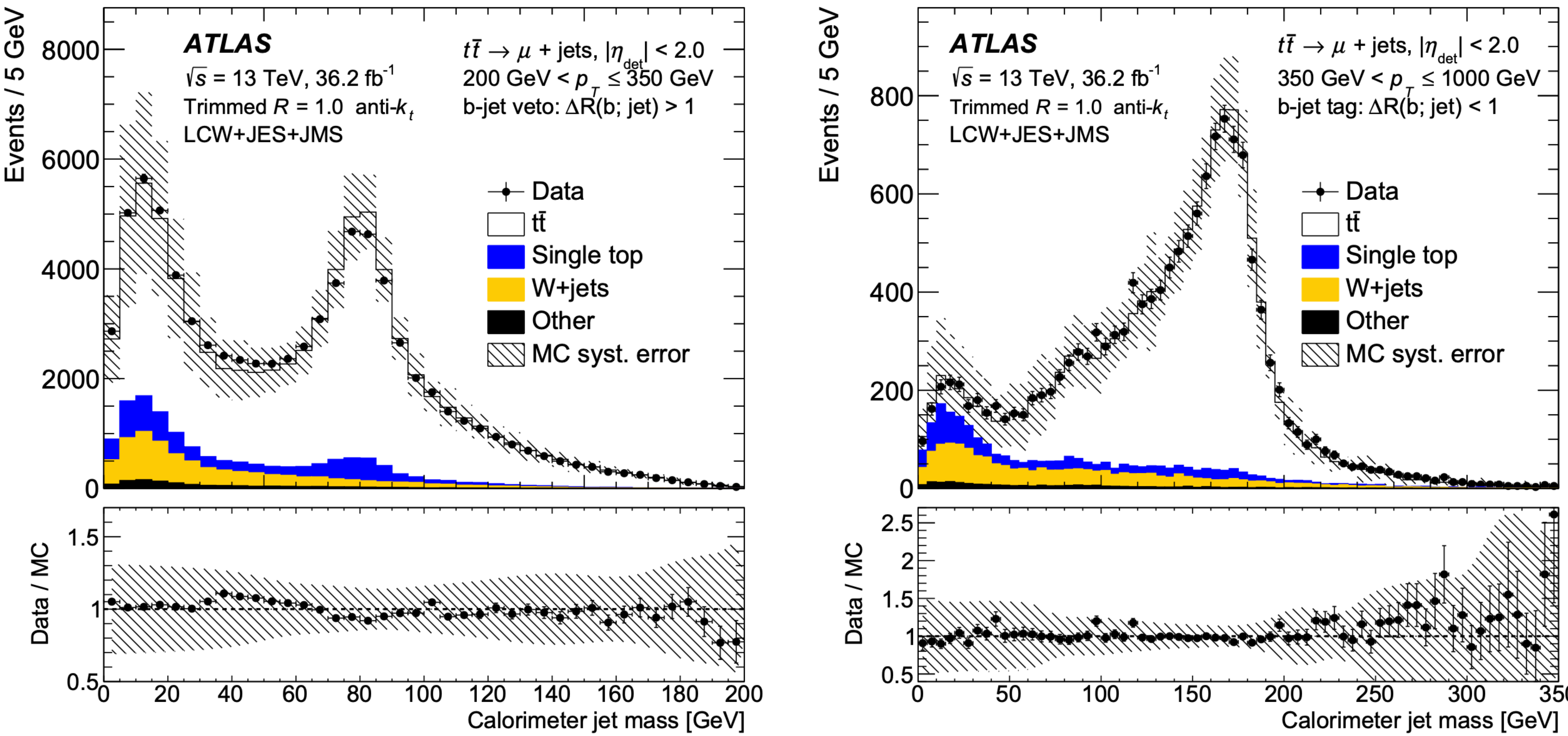
$$a_{12} = C_F,$$

$$a_{11} = -\frac{3C_F}{4},$$

$$a_{24} = -\frac{C_F^2}{2},$$

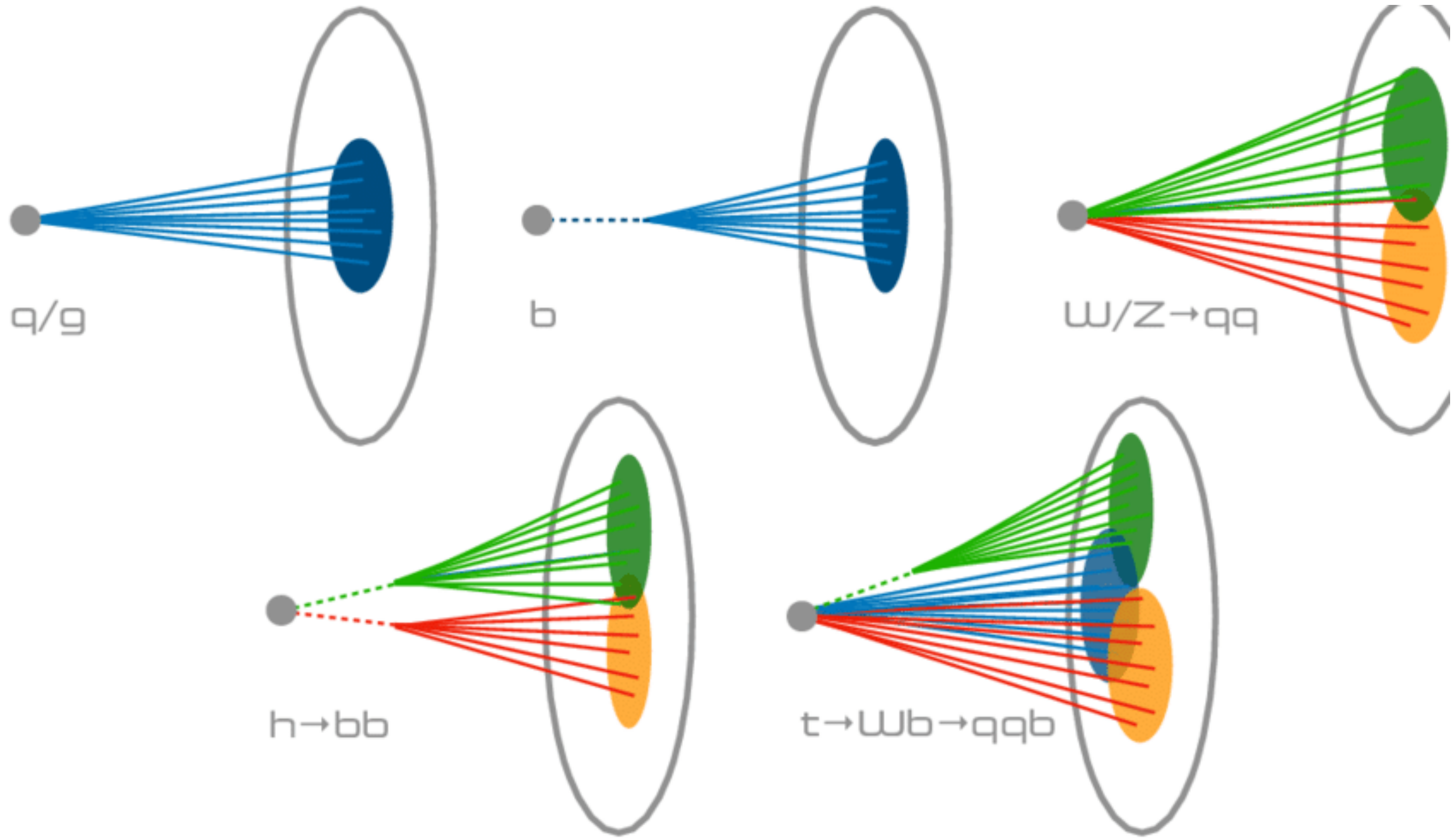
$$a_{23} = \frac{3}{8}C_F\left(3C_F + 4\beta_0\right)$$

Jet mass : the 0-th jet substructure

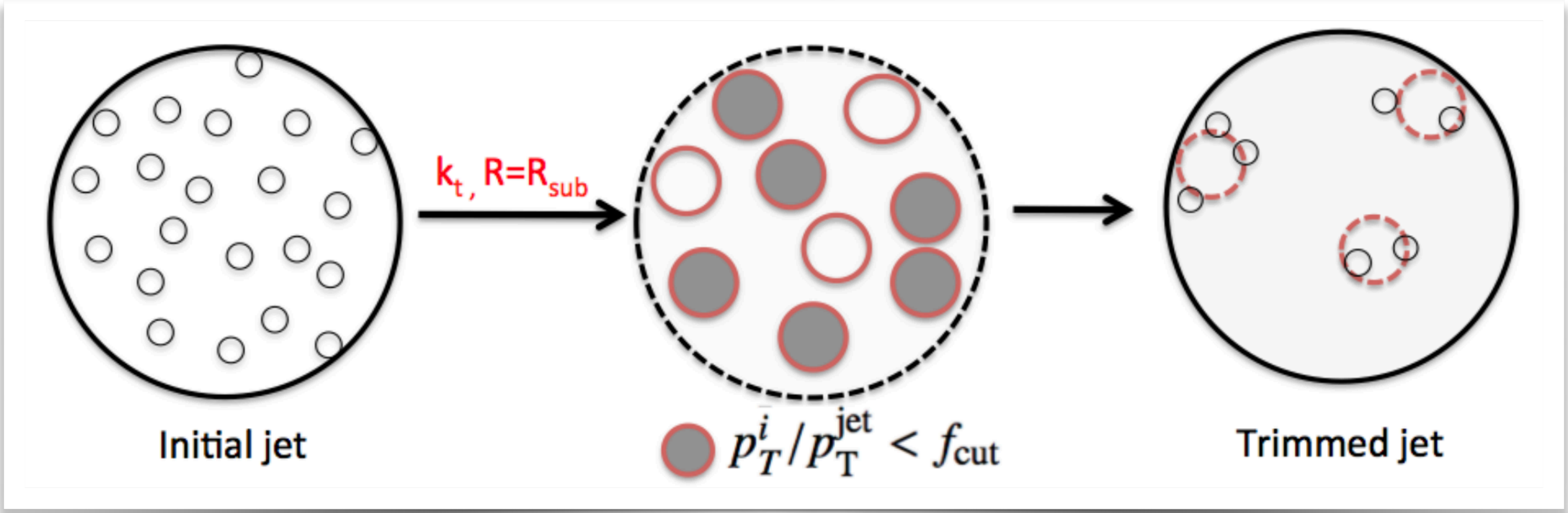


The generic class of substructures

Category I: prong finders. Tools in this category exploit the fact that when a boosted massive object decays into partons, all the partons typically carry a sizeable fraction of the initial jet transverse momentum, resulting in multiple hard cores in the jet. Conversely, quark and gluon jets are dominated by the radiation of soft gluons, and are therefore mainly single-core jets. Prong finders therefore look for multiple hard cores in a jet, hence reducing the contamination from "standard" QCD jets. This is often used to characterise the boosted jets in terms of their "pronginess", i.e. to their expected number of hard cores: QCD jets would be 1-prong objects, W/Z/H jets would be two-pronged, boosted top



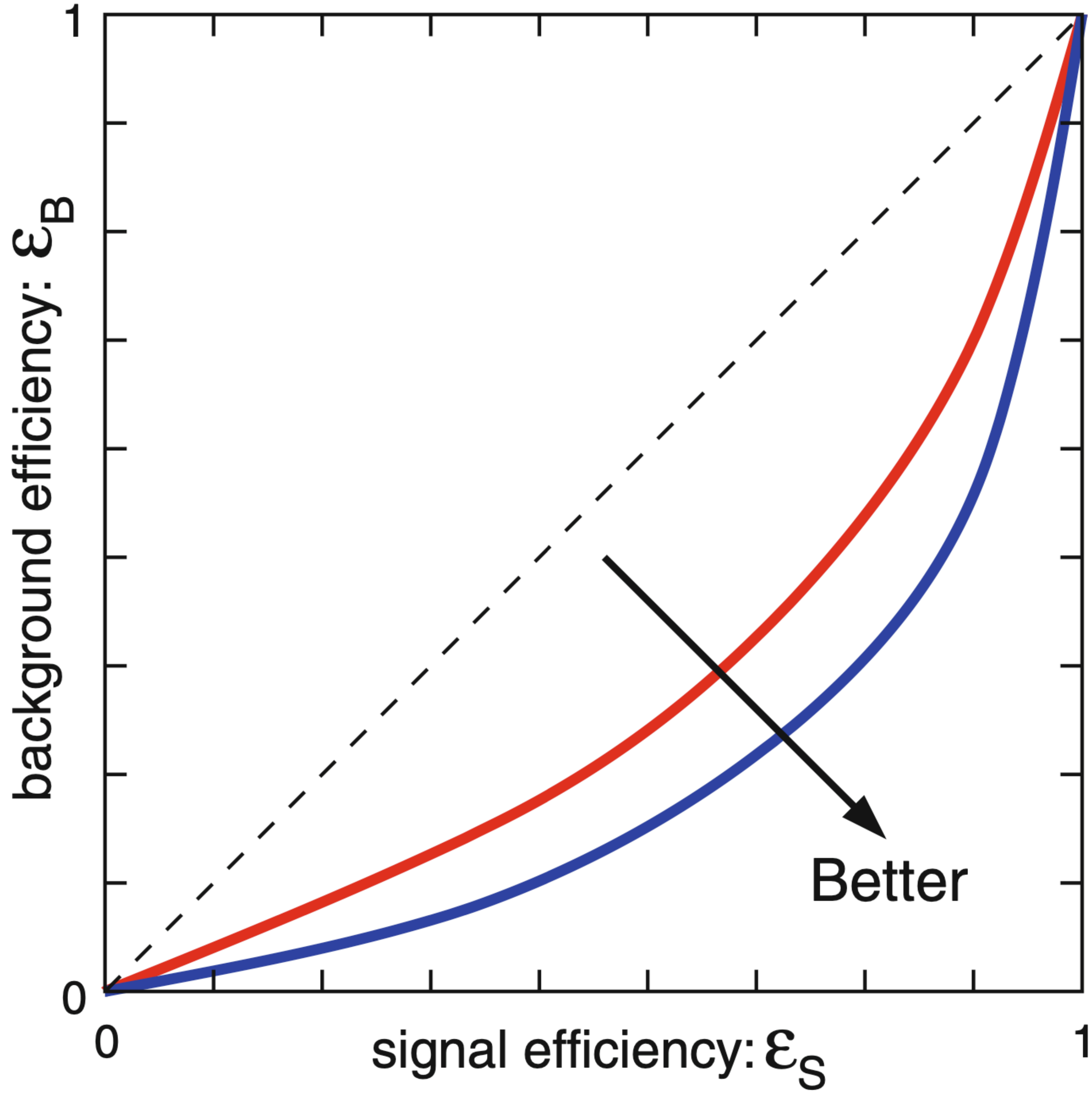
The generic class of substructures



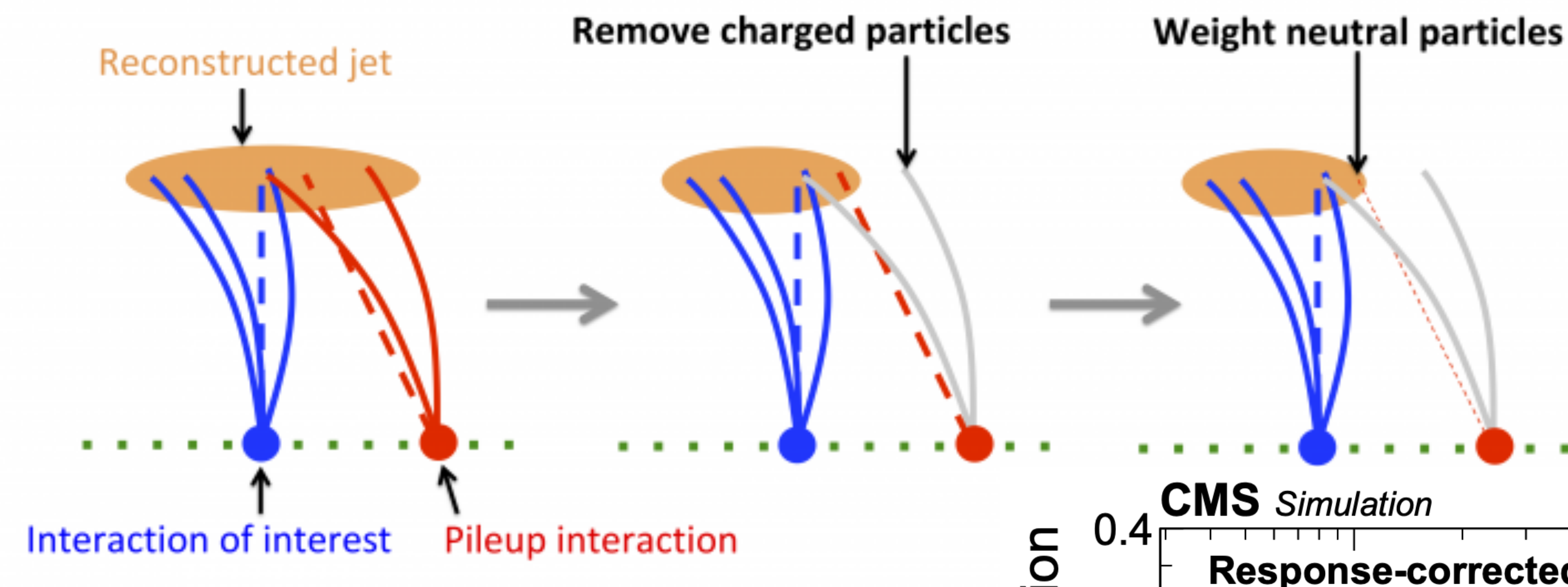
Category III: groomers. There is a third category of widely-used tools related to the fact that one often use large-radius jets for substructure studies. As we have already discussed, because of their large area, these jets are particularly sensitive to soft backgrounds, such as the UE and pileup. "Grooming" tools have therefore been introduced to mitigate the impact of these soft backgrounds on the fat jets. These tools usually work by removing the soft radiation far from the jet axis, where it is the most likely to come from a soft contamination rather than from QCD radiation inside the jet. In many respects, groomers share similarities with prong finders, essentially due to the fact that removing soft contamination and keeping the hard prongs are closely related.

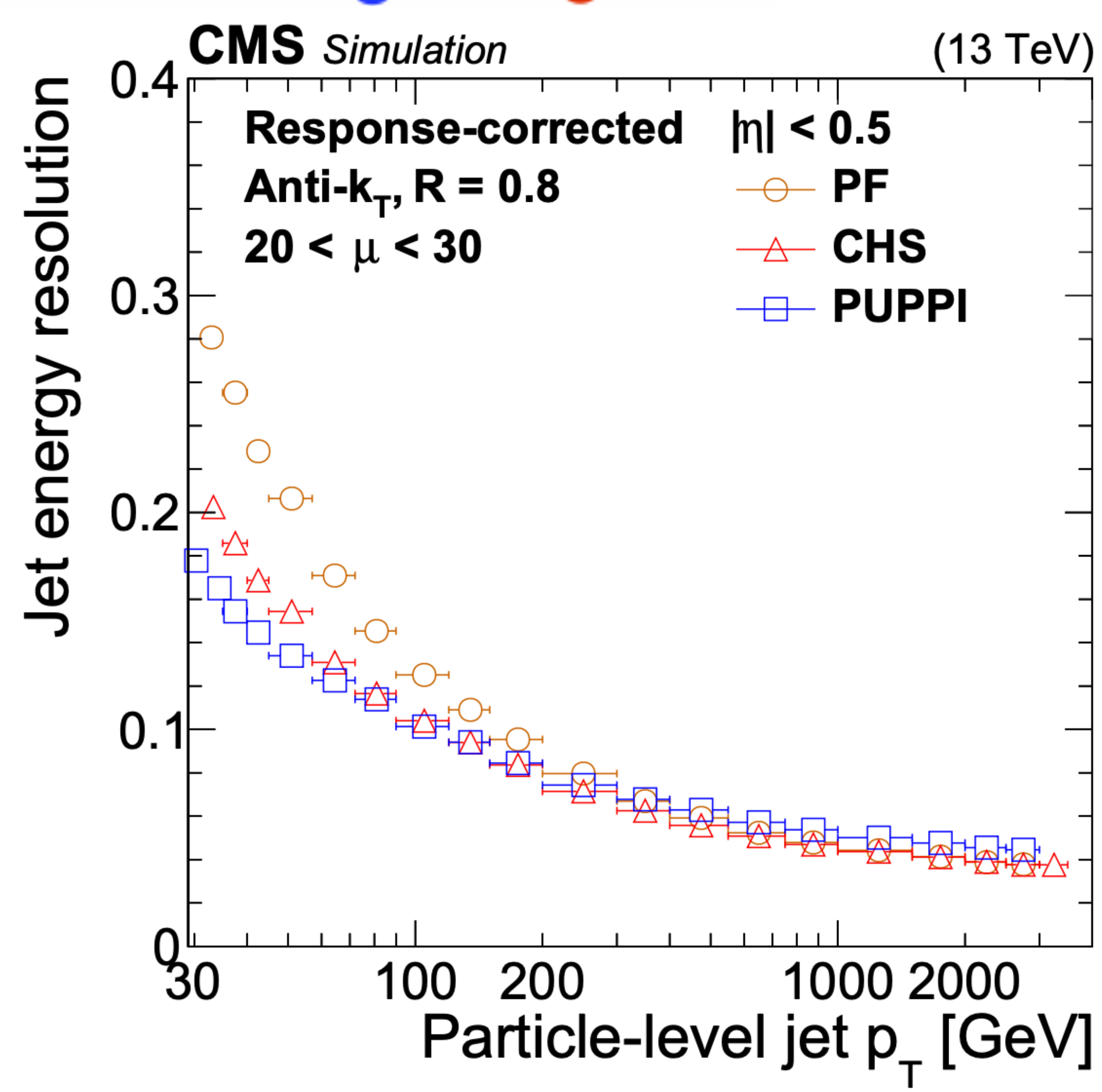
How do we quantify a tagger performance?

Receiver Operator Characteristics

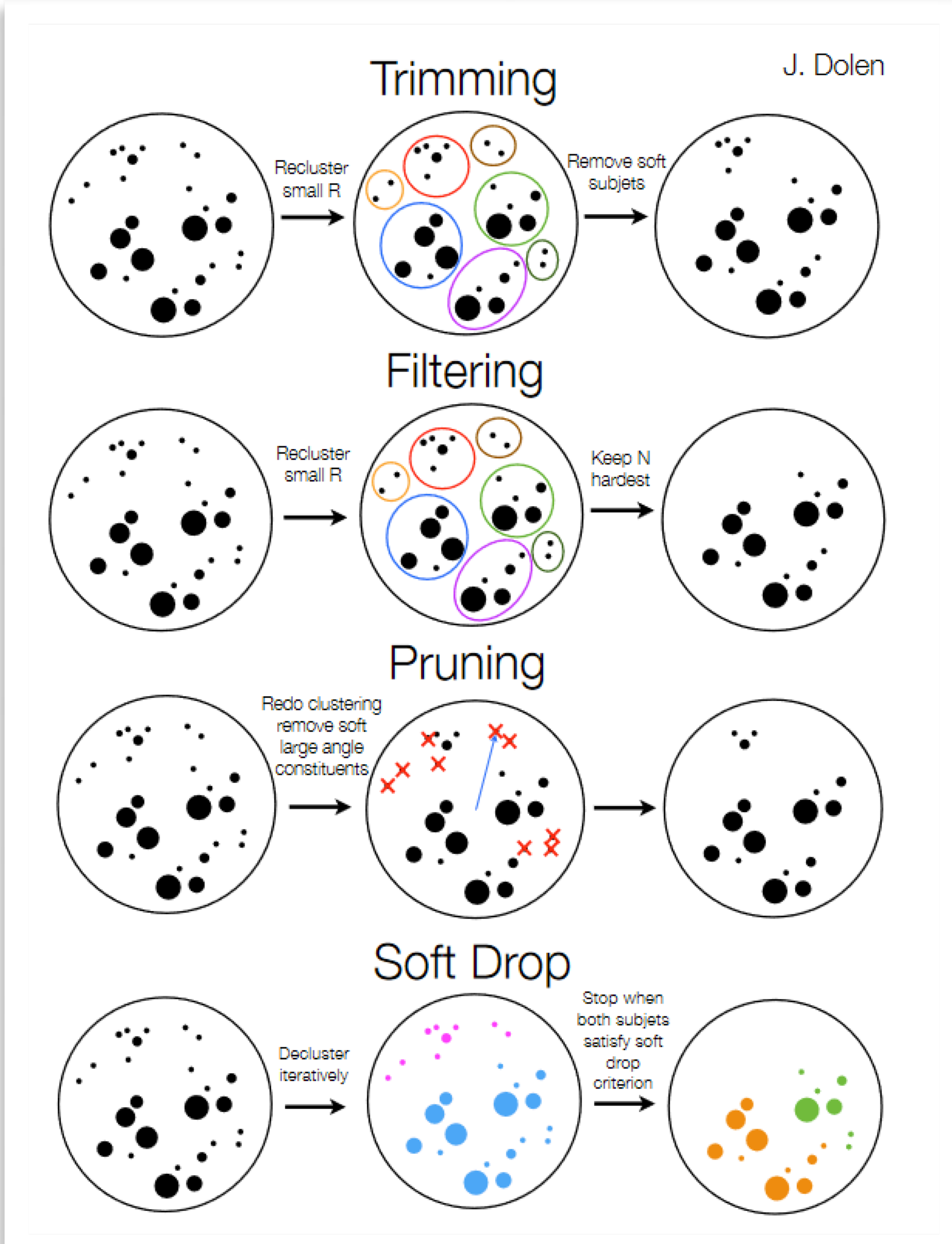


Pileup mitigation methods





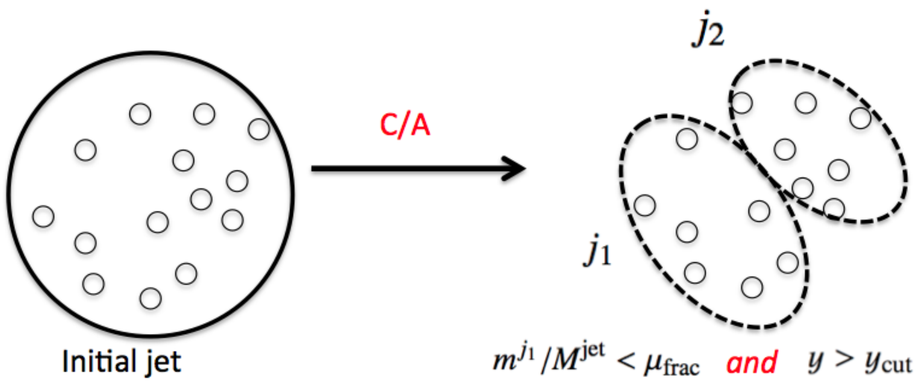
Different methods that evolved over edges

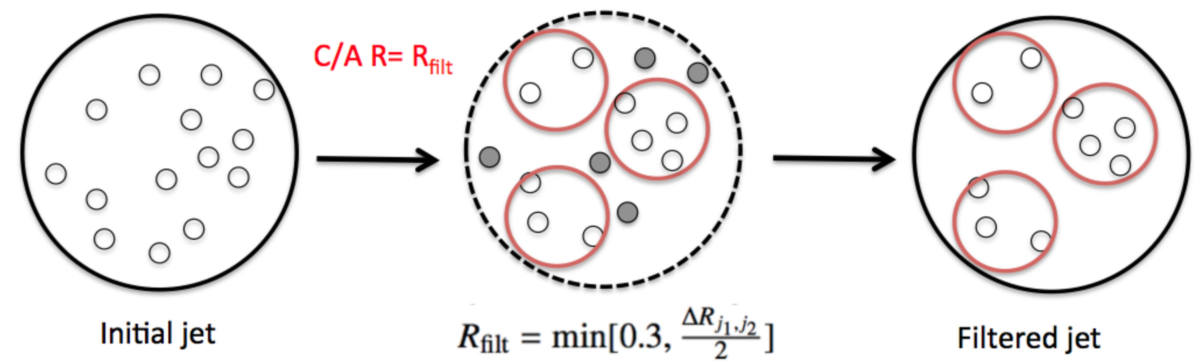


Cleaning large-R jets 1

The first step is to clean the hadronic jets via grooming : mass drop, trimming & pruning

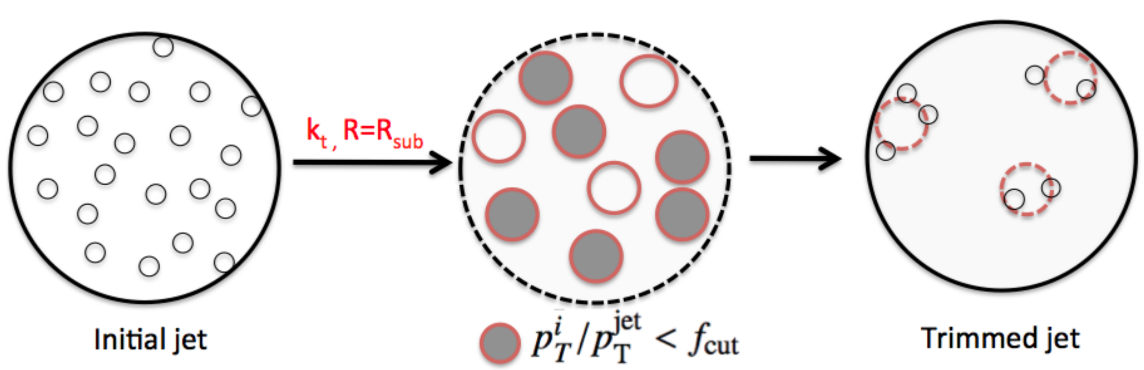
Mass Drop : It seeks to isolate concentrations of energy within a jet by identifying relatively symmetric sub-jets.





$$y = \frac{\min[(p_T^{j1})^2, (p_T^{j2})^2]}{(m_{jet})^2} \times \Delta R^2_{j1,j2}$$

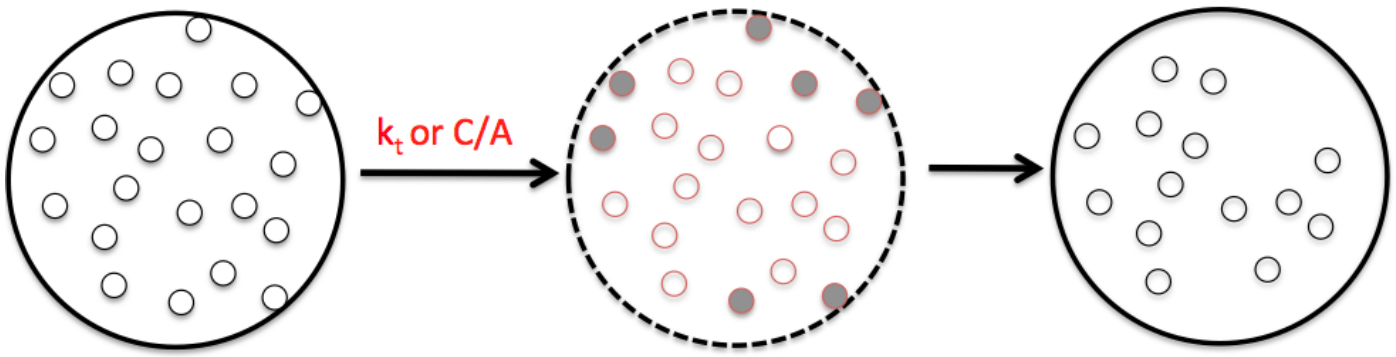
Trimming : Based on the idea that contamination from pileup, ISR, MPI in a jet should be much softer than the hard scattered constituents.



Cleaning large-R jets 2

The popular grooming techniques studied are : mass drop, trimming & pruning

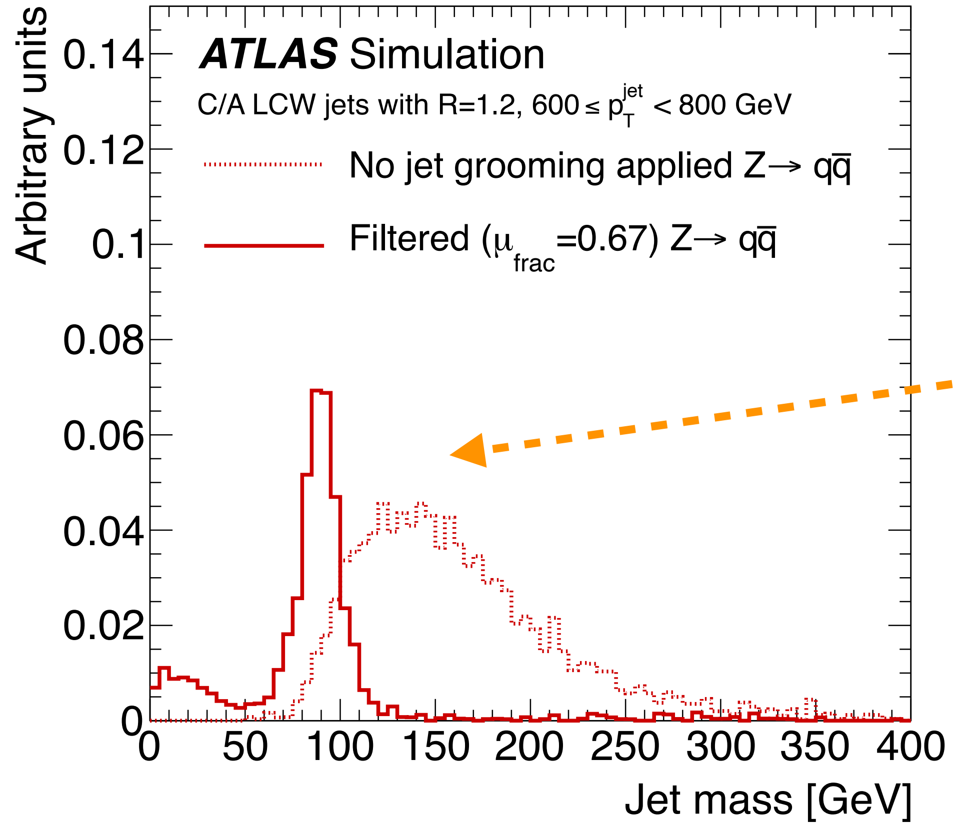
Pruning: Pruning is similar to trimming in terms of rejecting soft radiations. On top of that additional veto is applied on wide angle radiation.



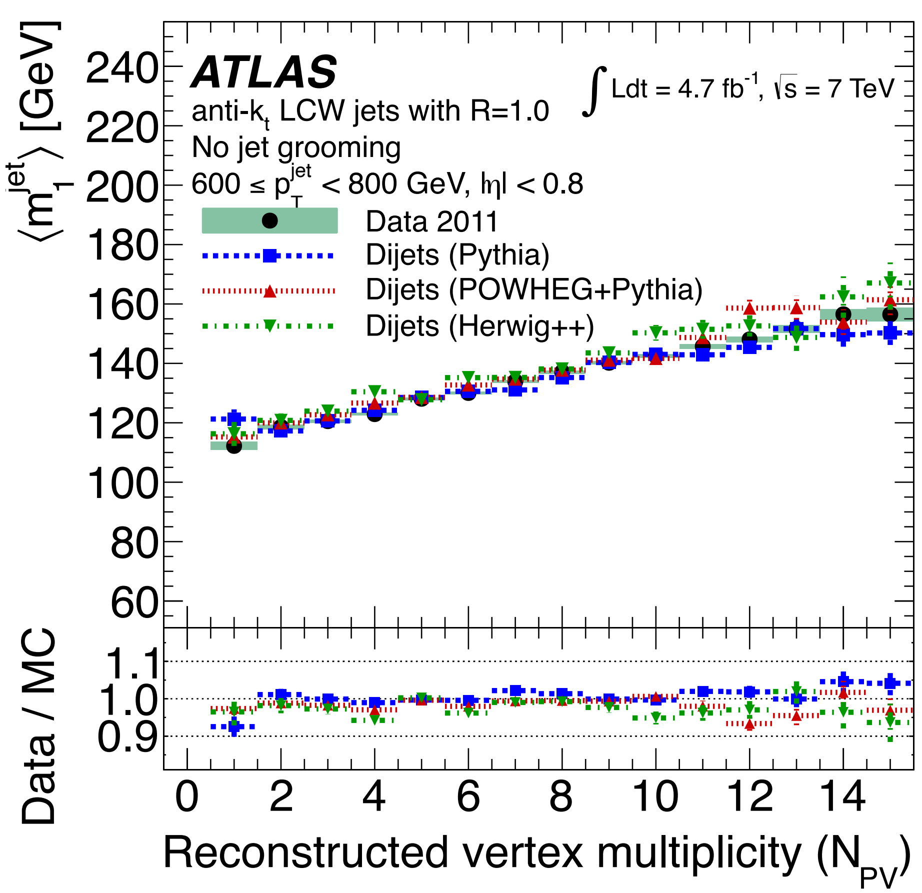
Jet finding algorithms used	Grooming algorithm	Configurations considered
C/A	Mass-Drop Filtering	μ_{frac} = 0.20, 0.33, **0.67**
Anti-k_t and C/A	Trimming	f_{cut} = 0.01, **0.03**, 0.05 R_{sub} = **0.2**, 0.3
Anti-k_t and C/A	Pruning	R_{cut} = 0.1, 0.2, 0.3 z_{cut} = 0.05, 0.1
C/A	HEPTopTagger	(see table 2)

Parameters used in ATLAS analysis for different groomers

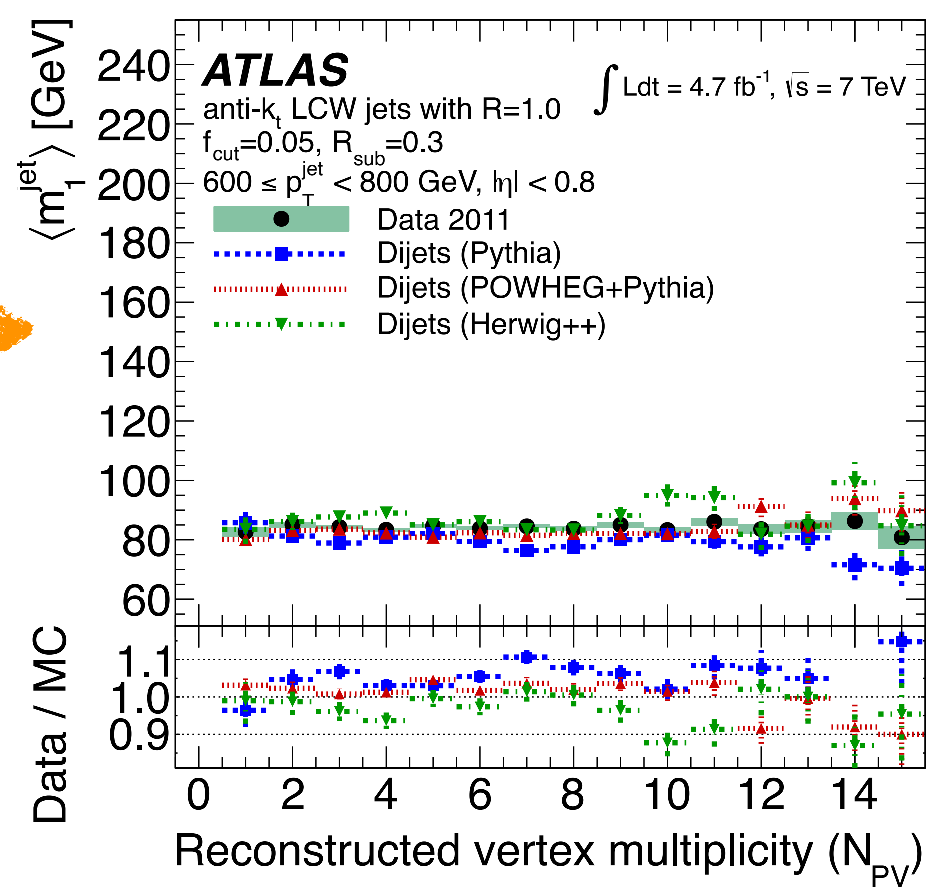
Impact of jet cleaning



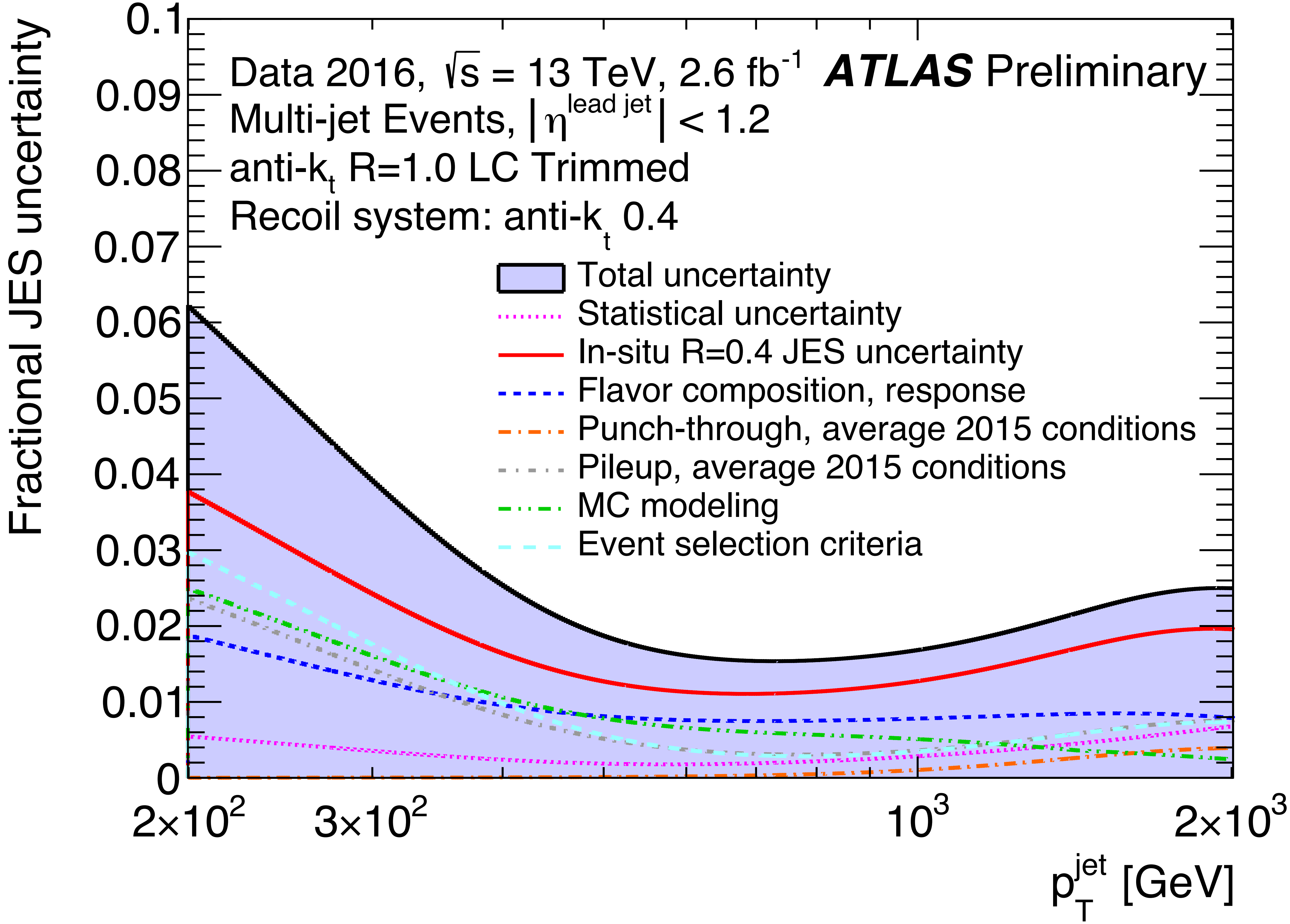
Jet mass peaks are prominent after grooming applied.



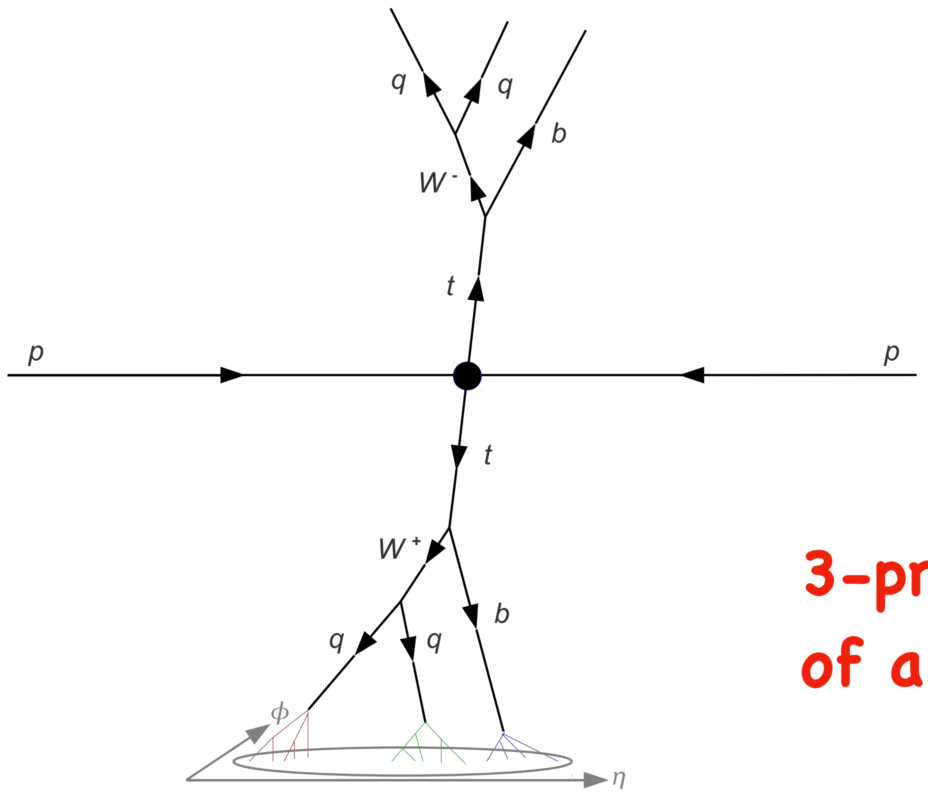
Pileup effects removed by jet grooming

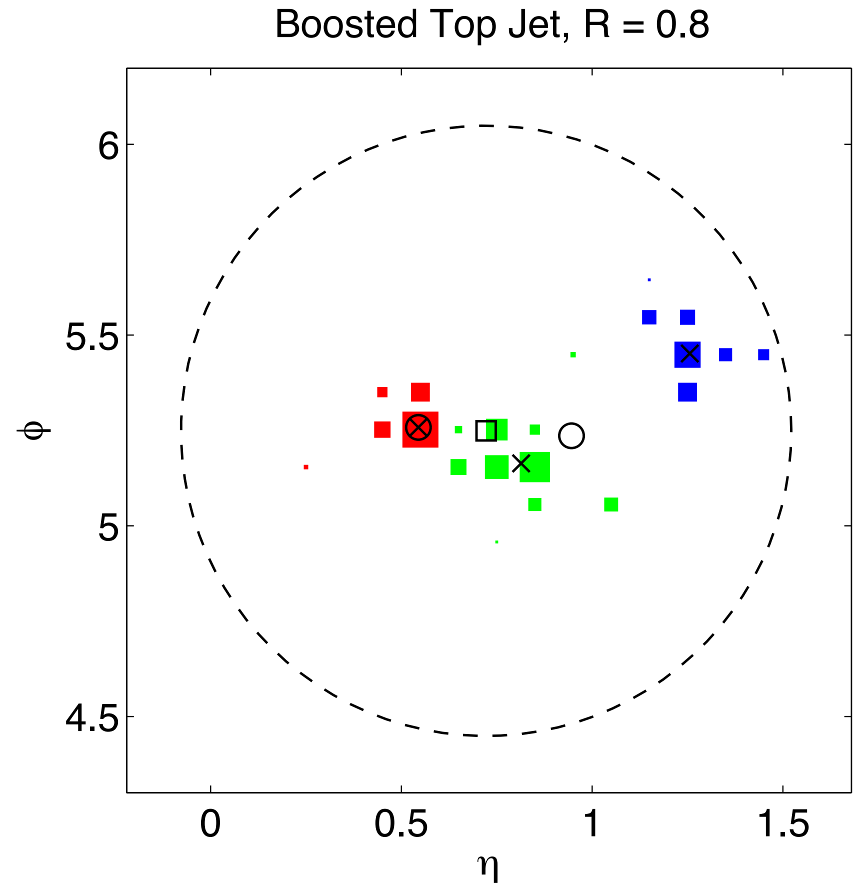


How well we can reconstruct “fat”-jets?

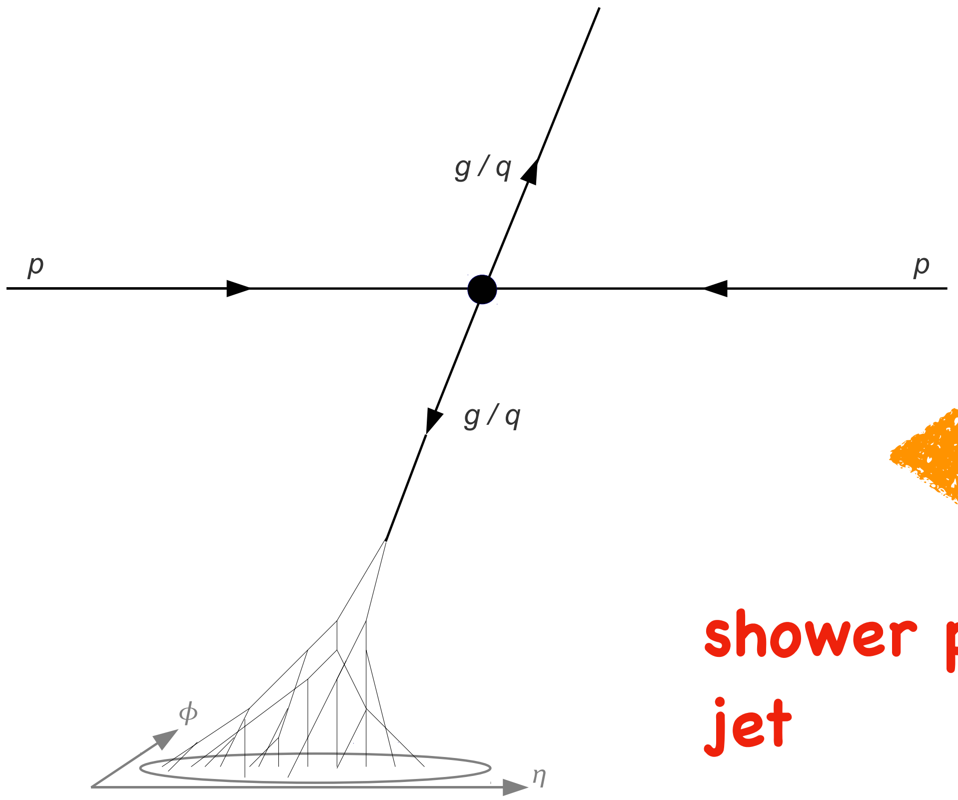


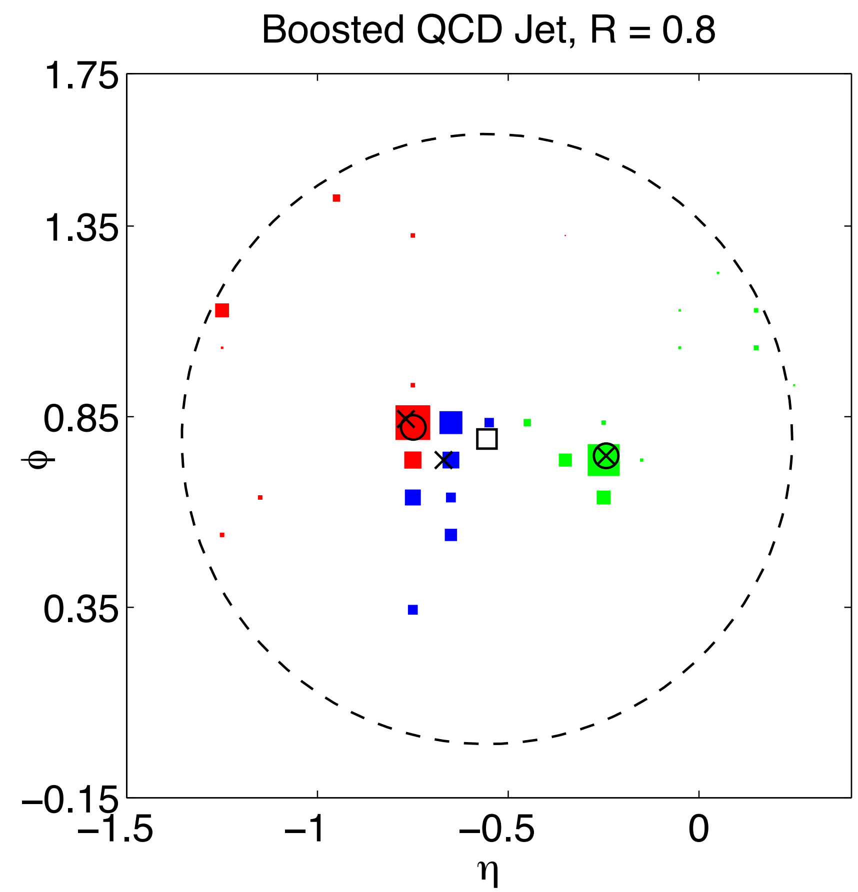
Is grooming sufficient?





3-prong shower pattern of a top like jet





shower pattern of a QCD jet

Variables for identifying boosted tops

Jet Mass: The invariant mass of the jet built from constituents : $\mathbf{m^2 = \left(\sum_{i\in jet} p_i\right)^2}$

K_t Splitting Scale: scale of the last recombination steps in the K_t algorithm.
Large R jet constituents are reclustered using K_t algorithm.

JHEP 1606 (2016) 093

$$\sqrt{d_{ij}} = \min(p_{T,i}, p_{T,j}) \times \Delta R_{ij}$$

- ☑ Sensitive to whether the last recombination steps correspond to the merging of the decay products of massive particles.
- ☑ For hadronic top case : $\sqrt{d_{12}} = m_t/2$ and $\sqrt{d_{23}} = m_W/2$

N-subjettiness: τ_N quantifies how well jets can be described as containing N or fewer subjets.

$$\tau_N = \frac{1}{d_0}\sum_k p_{Tk} \times \Delta R_k^{min}$$

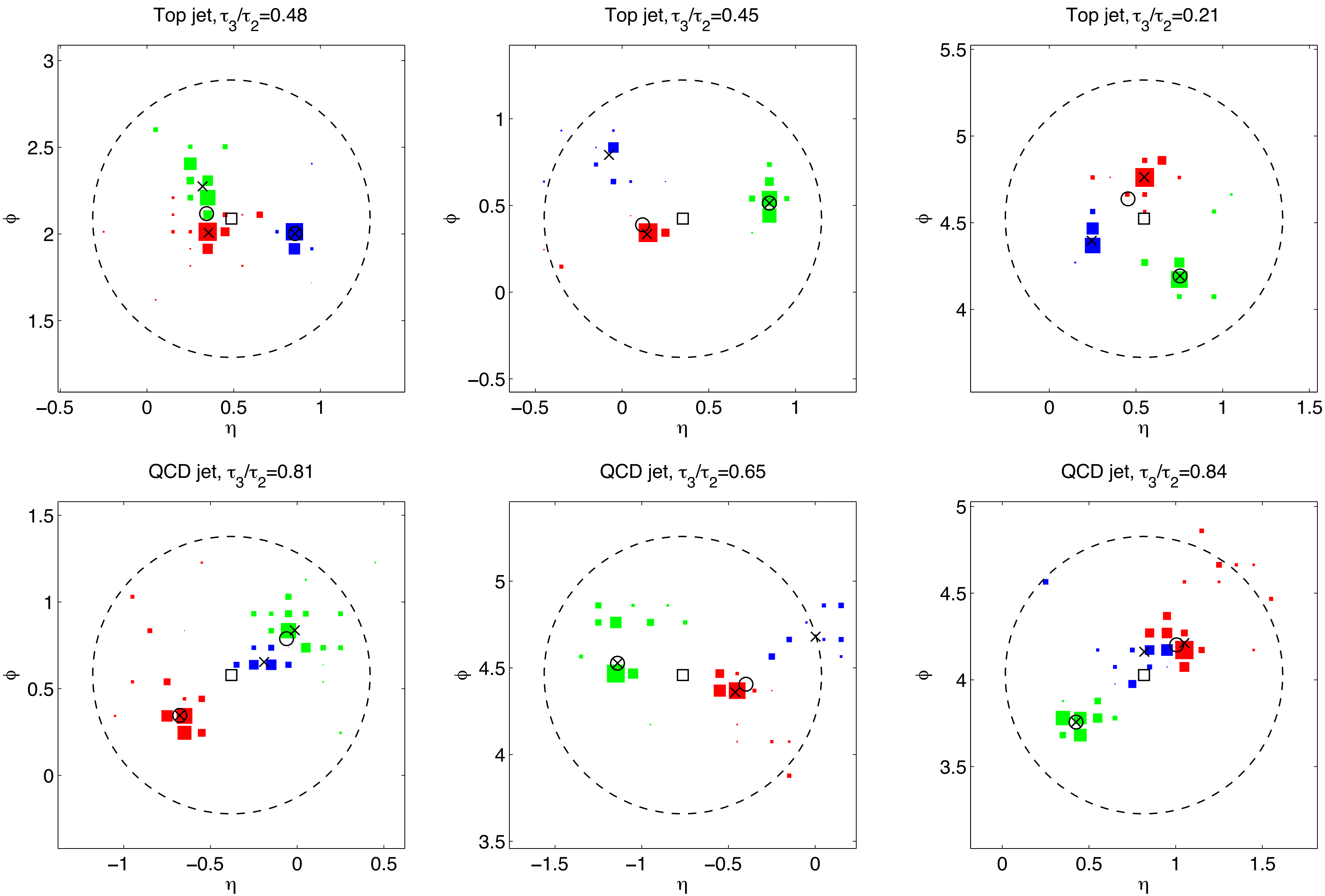
with $d_0 \equiv \sum_k p_{Tk} \times R$. $\tau_{ij} = \tau_i/\tau_j$ with τ_N/τ_{N-1} characterising a N prong fat jet.

Shower Deconstruction: ratio of likelihoods that measure the compatibility of the shower with a top quark decay over the same for a background process.

$$\chi(\{p_i^k\}) = \frac{\sum_{perm} P(\{p_i^k\}|signal)}{\sum_{perm} P(\{p_i^k\}|background)}$$

- ☑ Computation of SD is CPU intensive.
- ☑ Tagging is done by putting a threshold value on SD variable.

Discriminating power of jet shapes



Substructure based taggers

JHEP 1606 (2016) 093

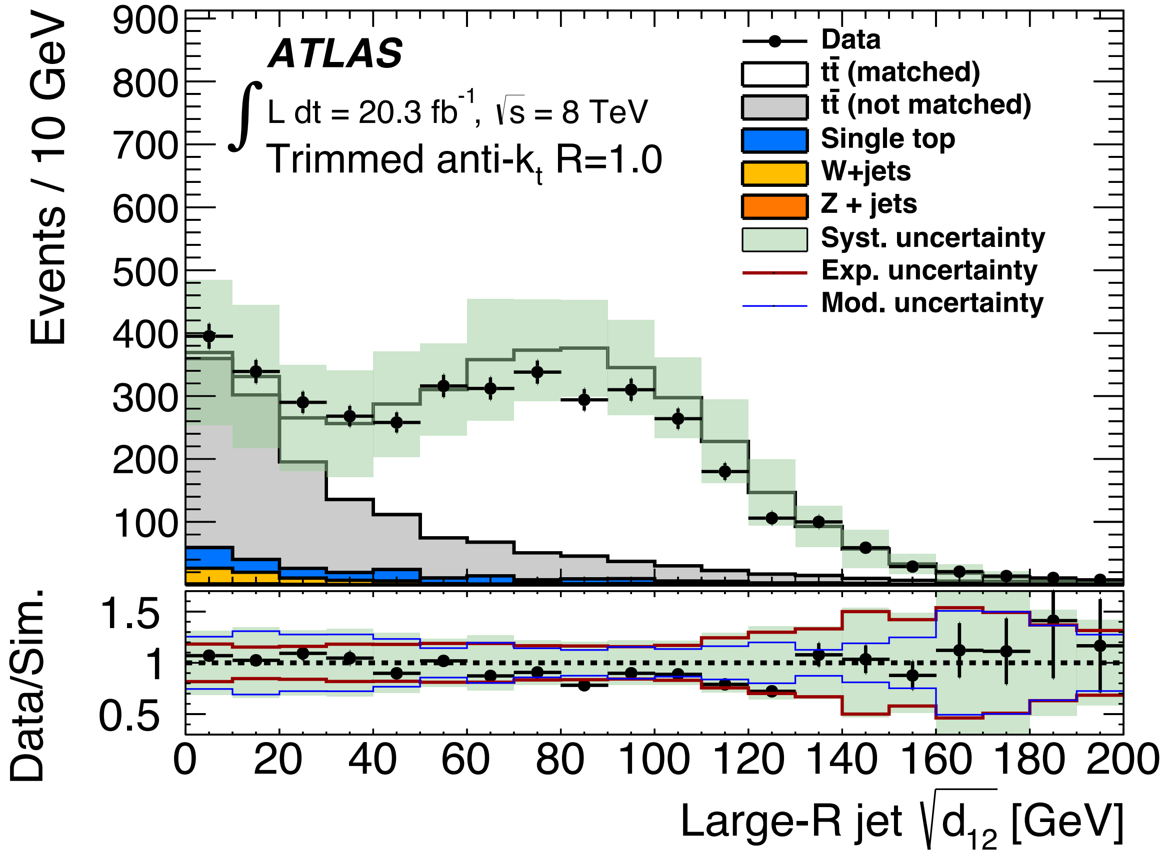
Tagger	Top-tagging criterion
Substructure tagger I	$\sqrt{d_{12}} > 40$ GeV
Substructure tagger II	$m > 100$ GeV
Substructure tagger III	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV
Substructure tagger IV	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 10$ GeV
Substructure tagger V	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 20$ GeV
W' top tagger	$\sqrt{d_{12}} > 40$ GeV and $0.4 < \tau_{21} < 0.9$ and $\tau_{32} < 0.65$

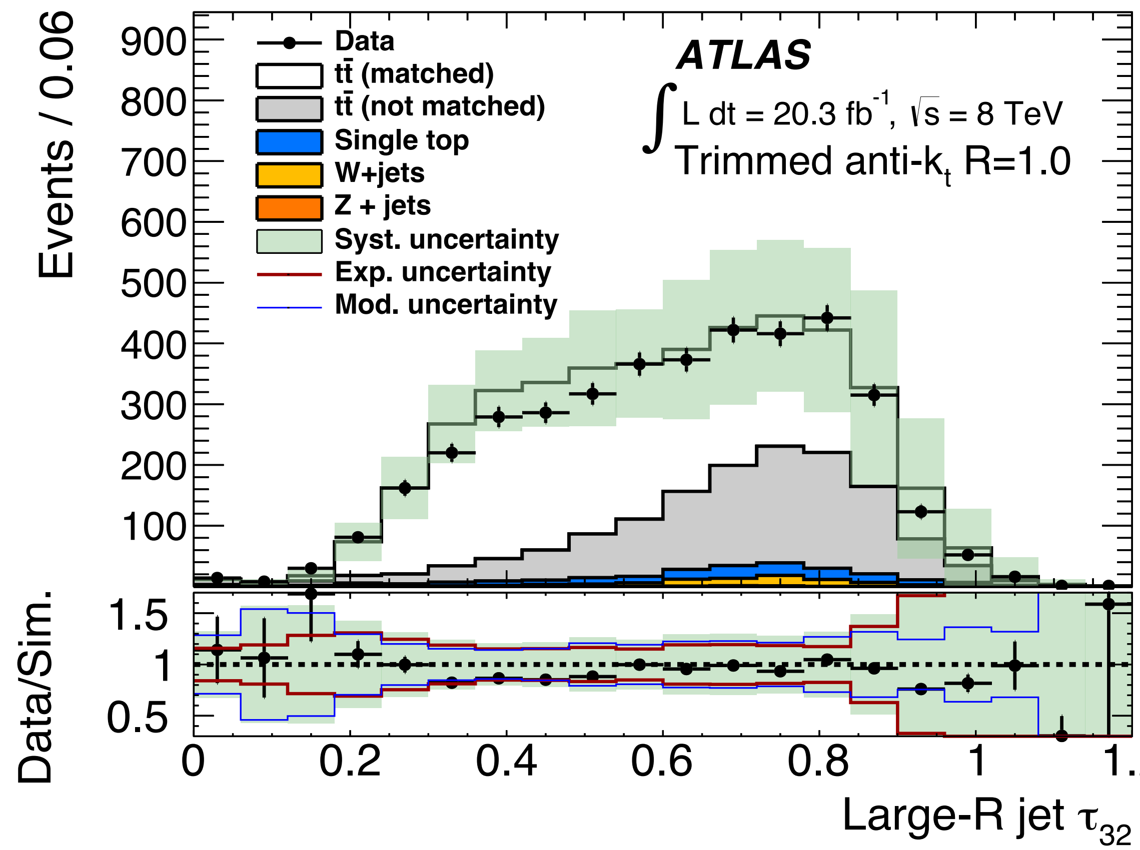
Tagger	Jet algorithm	Grooming	Radius parameter	p_T range	$\lvert\eta\rvert$ range
Tagger I–V W' top tagger Shower Deconstruction	anti-k_t	trimming ($R_{sub} = 0.3$, $f_{cut} = 0.05$)	$R = 1.0$	> 350 GeV	< 2
HEPTopTagger	C/A	none	$R = 1.5$	> 200 GeV	< 2

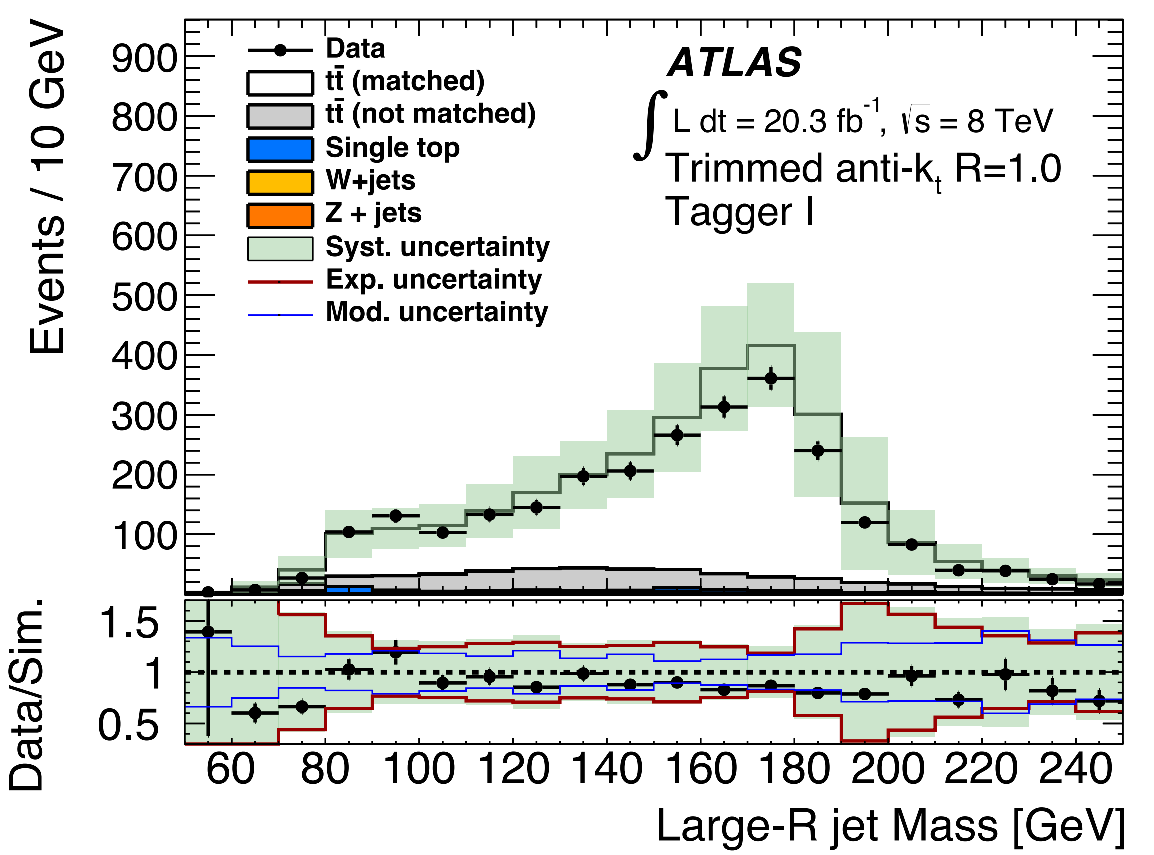
The jet grooming & parameters used by different taggers.

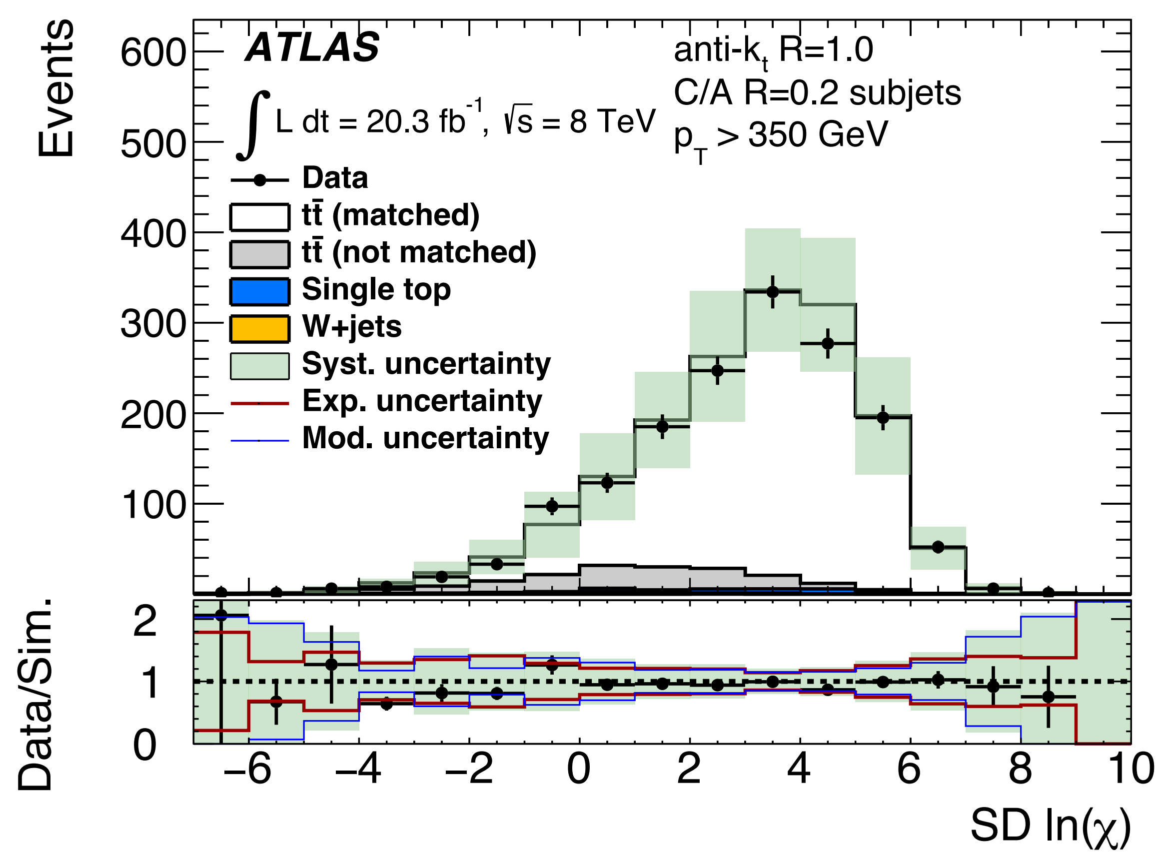
Distribution of tagger variables

JHEP 1606 (2016) 093



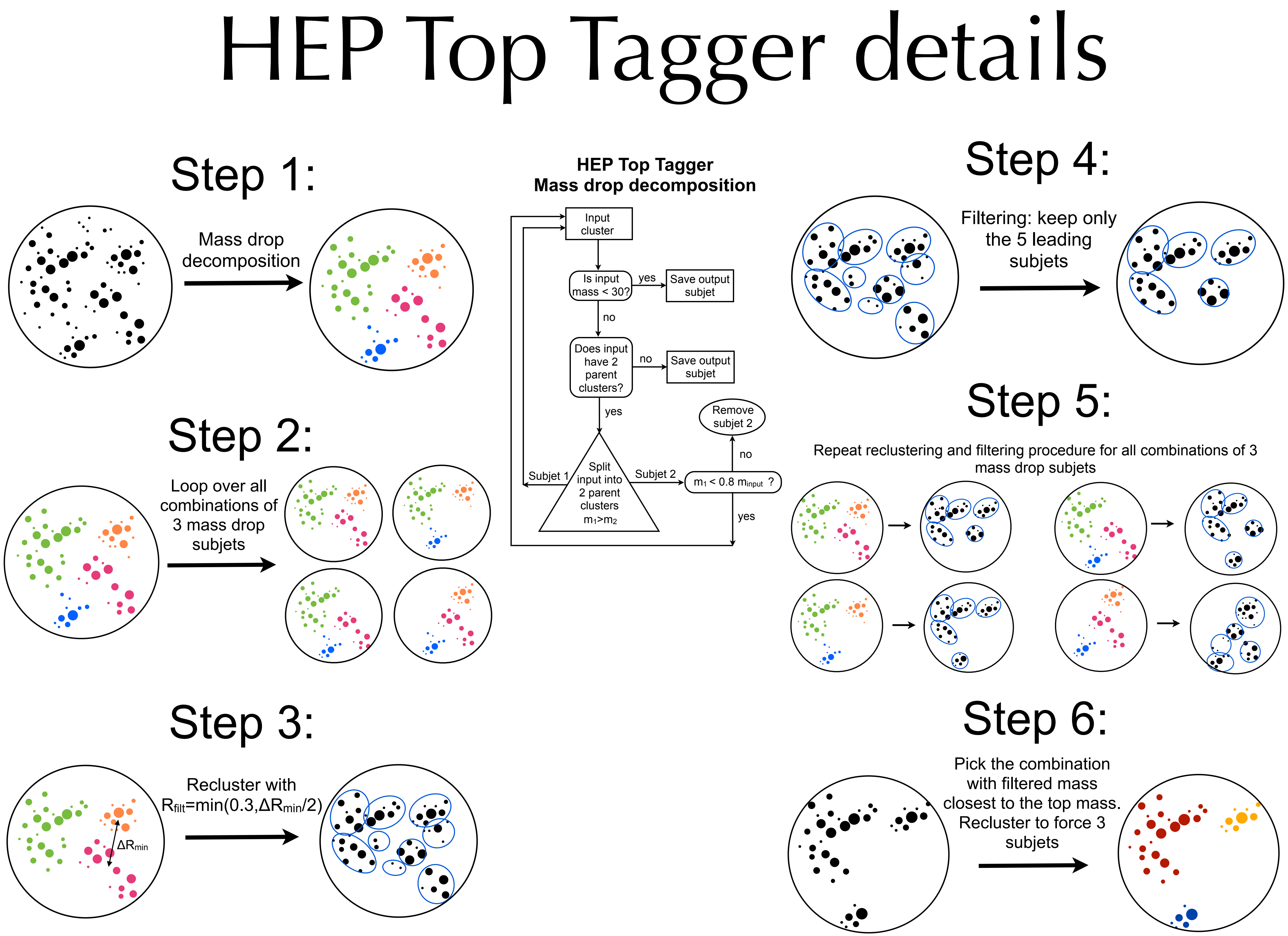




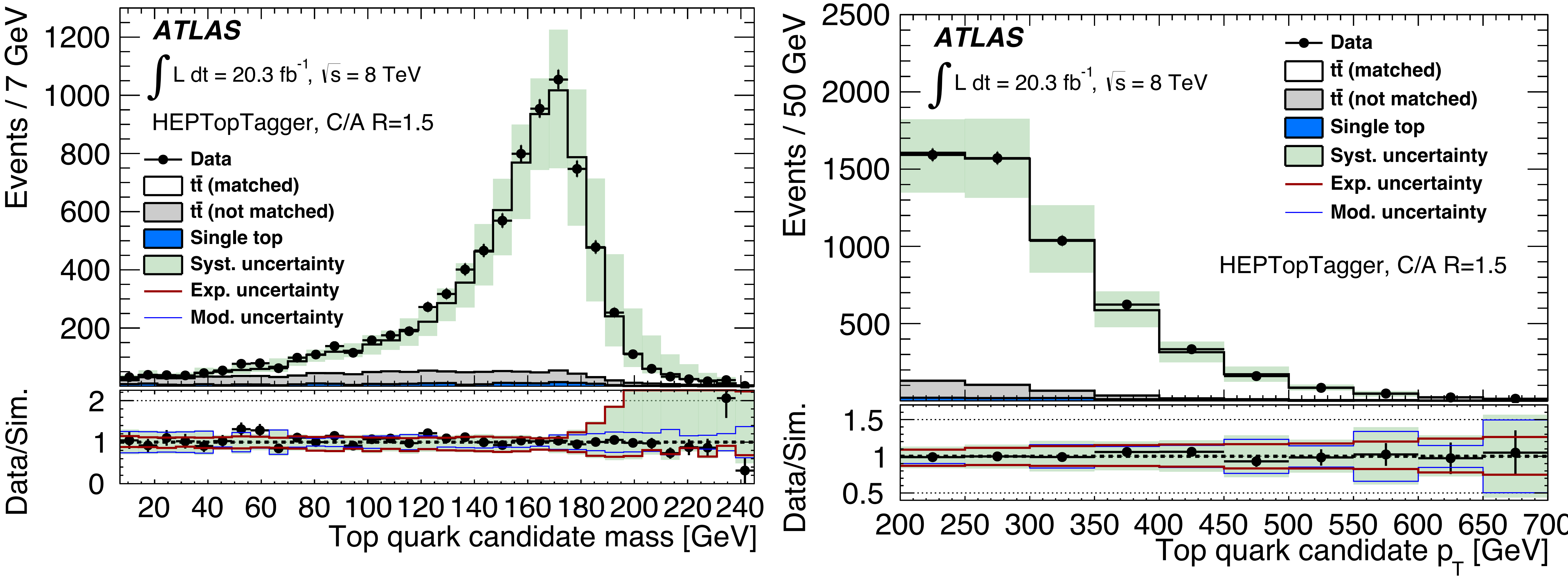


HEP Top-tagger

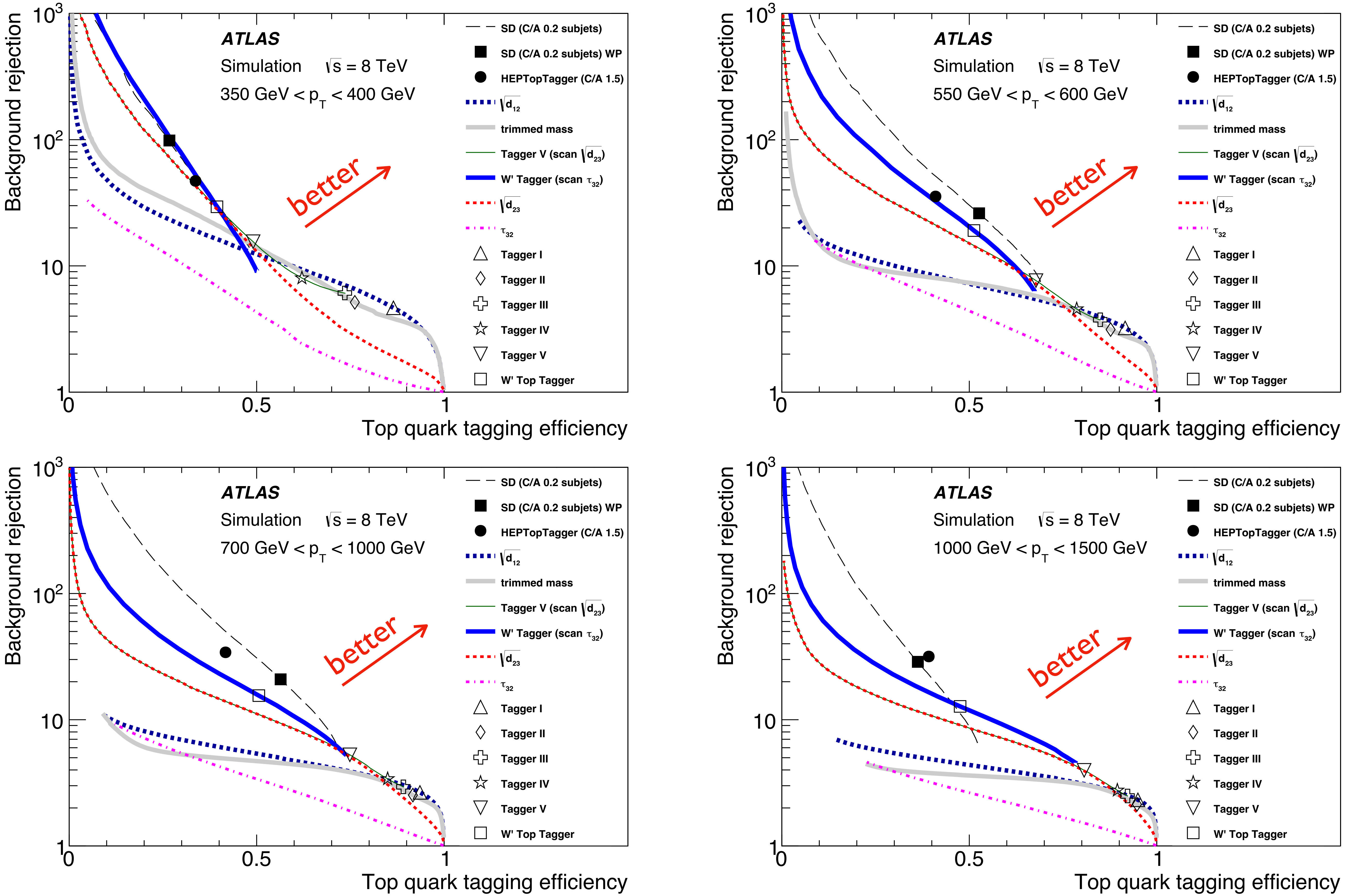
The HEPTopTagger algorithm is designed to identify a top quark with a hadronically decaying W boson daughter over a large multi-jet background



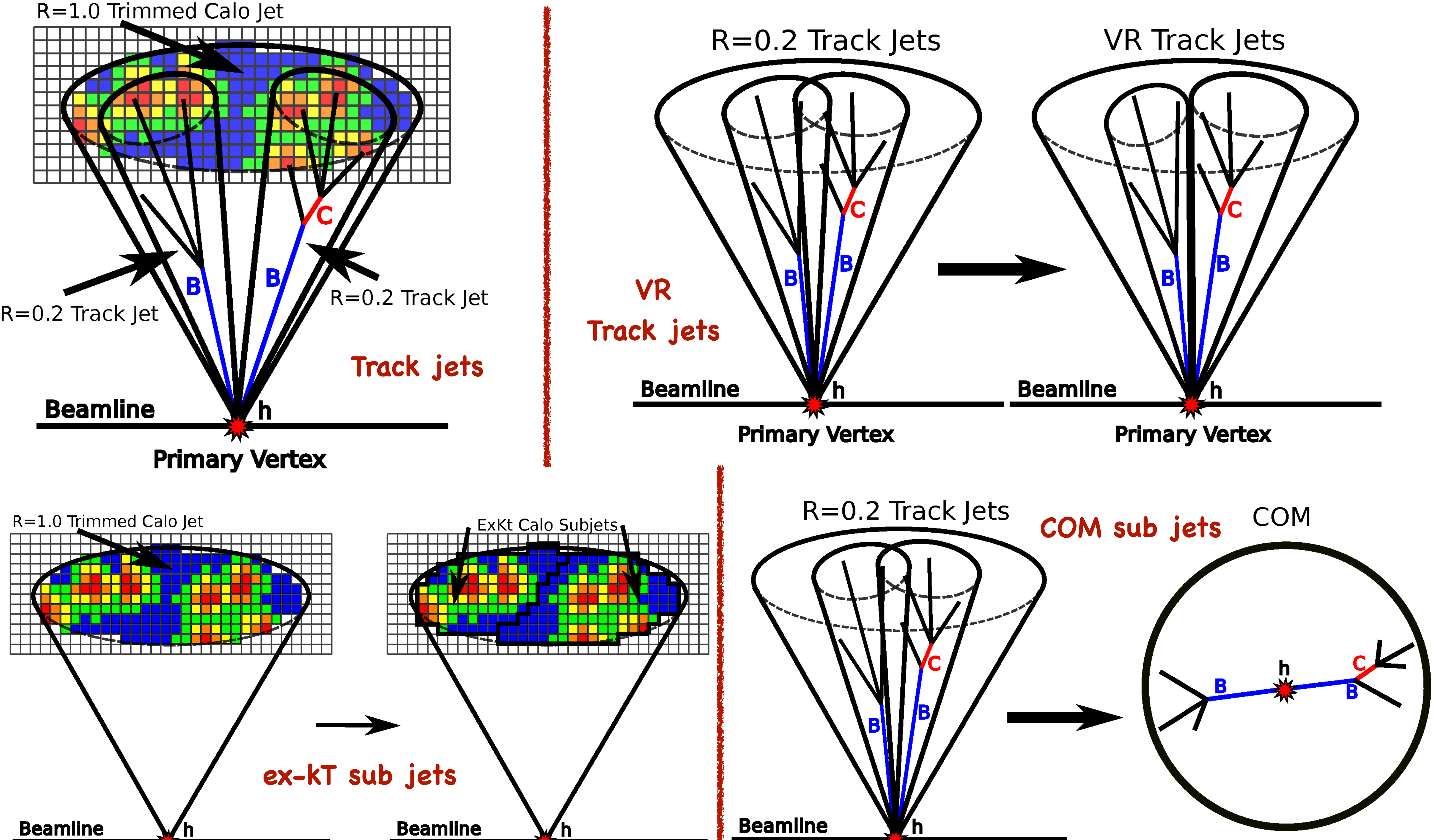
HEP Top-tagger performance



Performance of the tagger

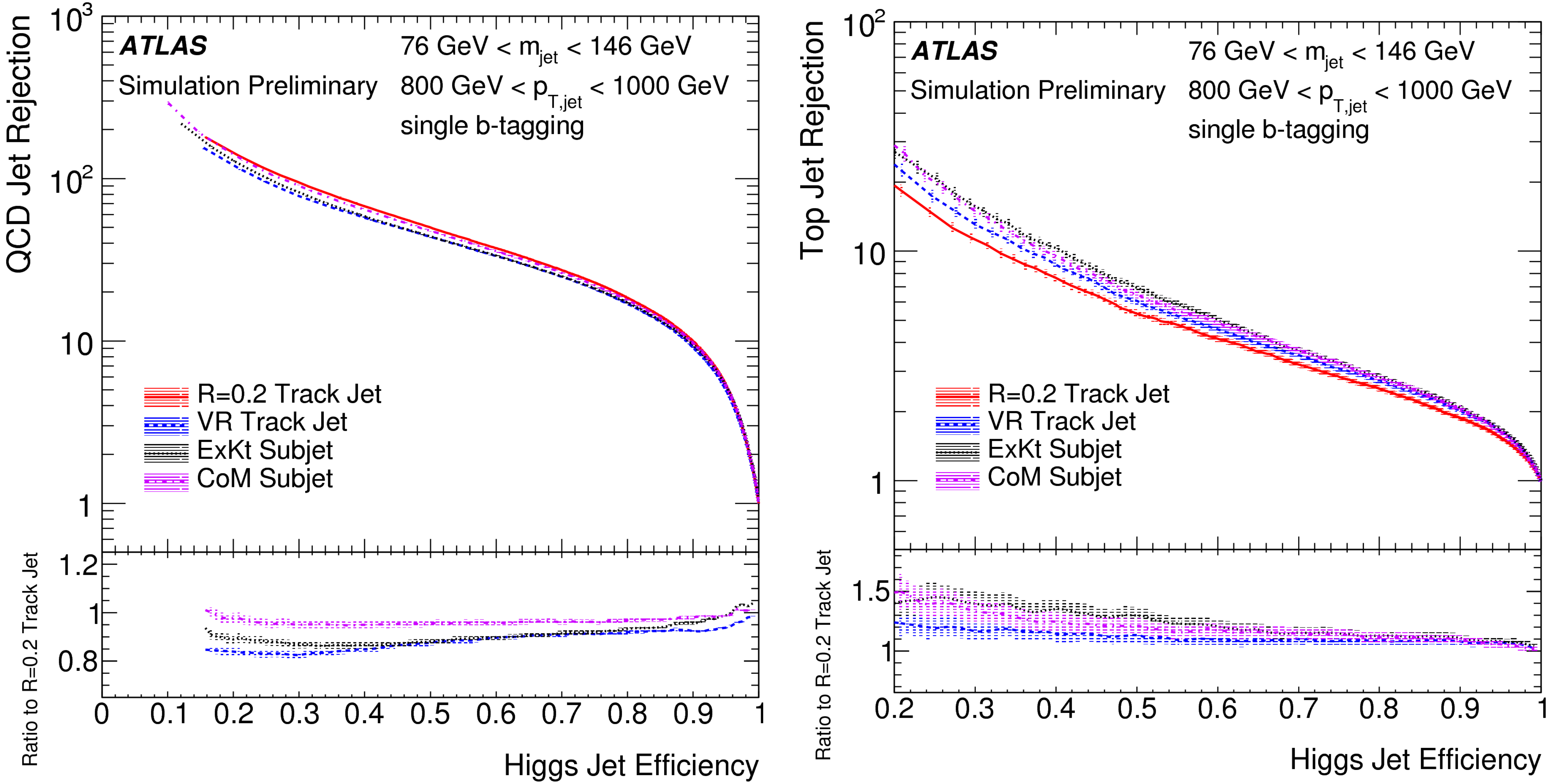


Higgs tagging algorithm



$$y_{ij} = 2 \times \left(1 - cos\theta_{ij}\right)$$

Higgs Tagging performance



Performance of the Higgs taggers