QJets

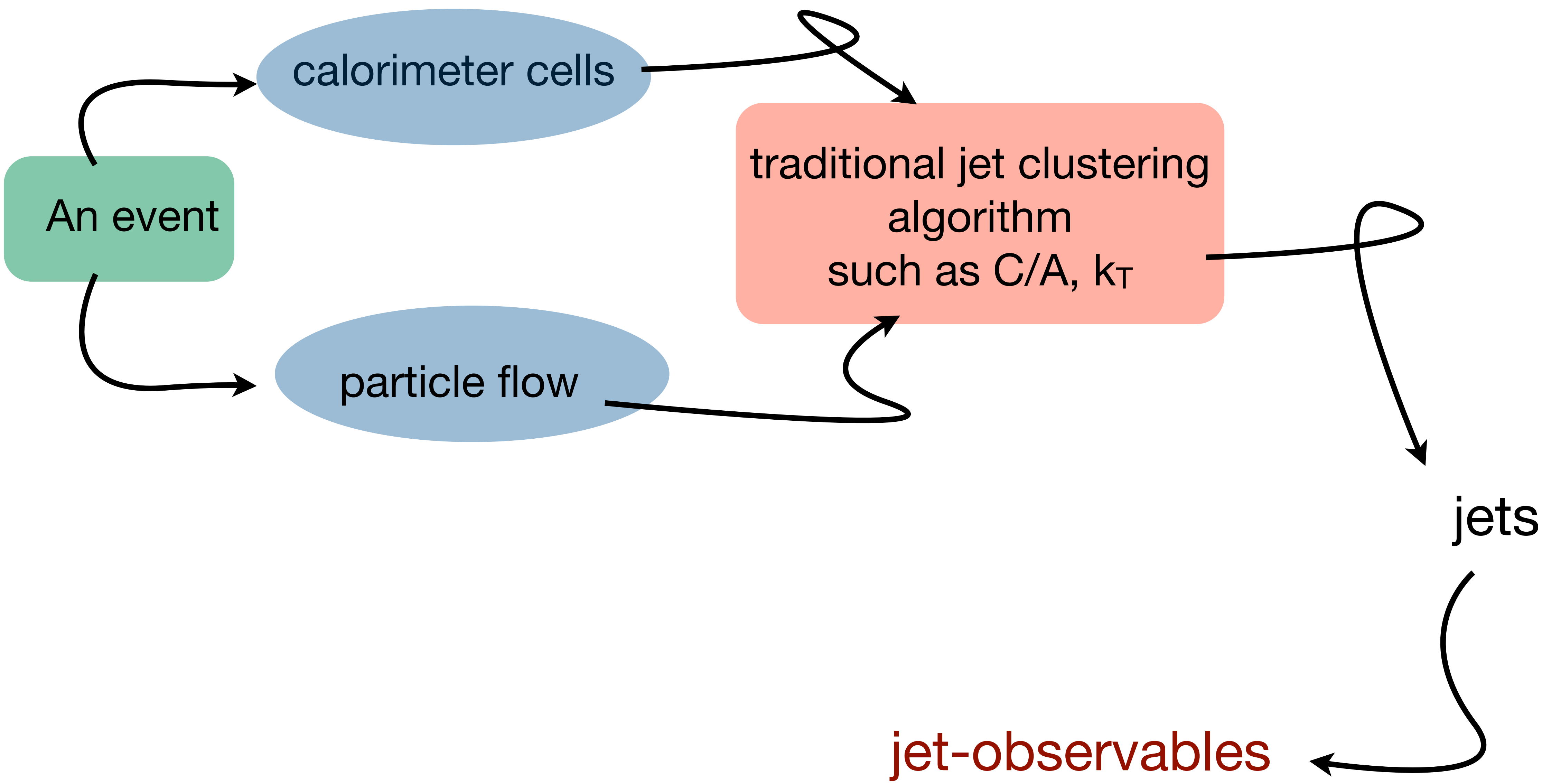
Tuhin S. Roy

Los Alamos National Laboratory

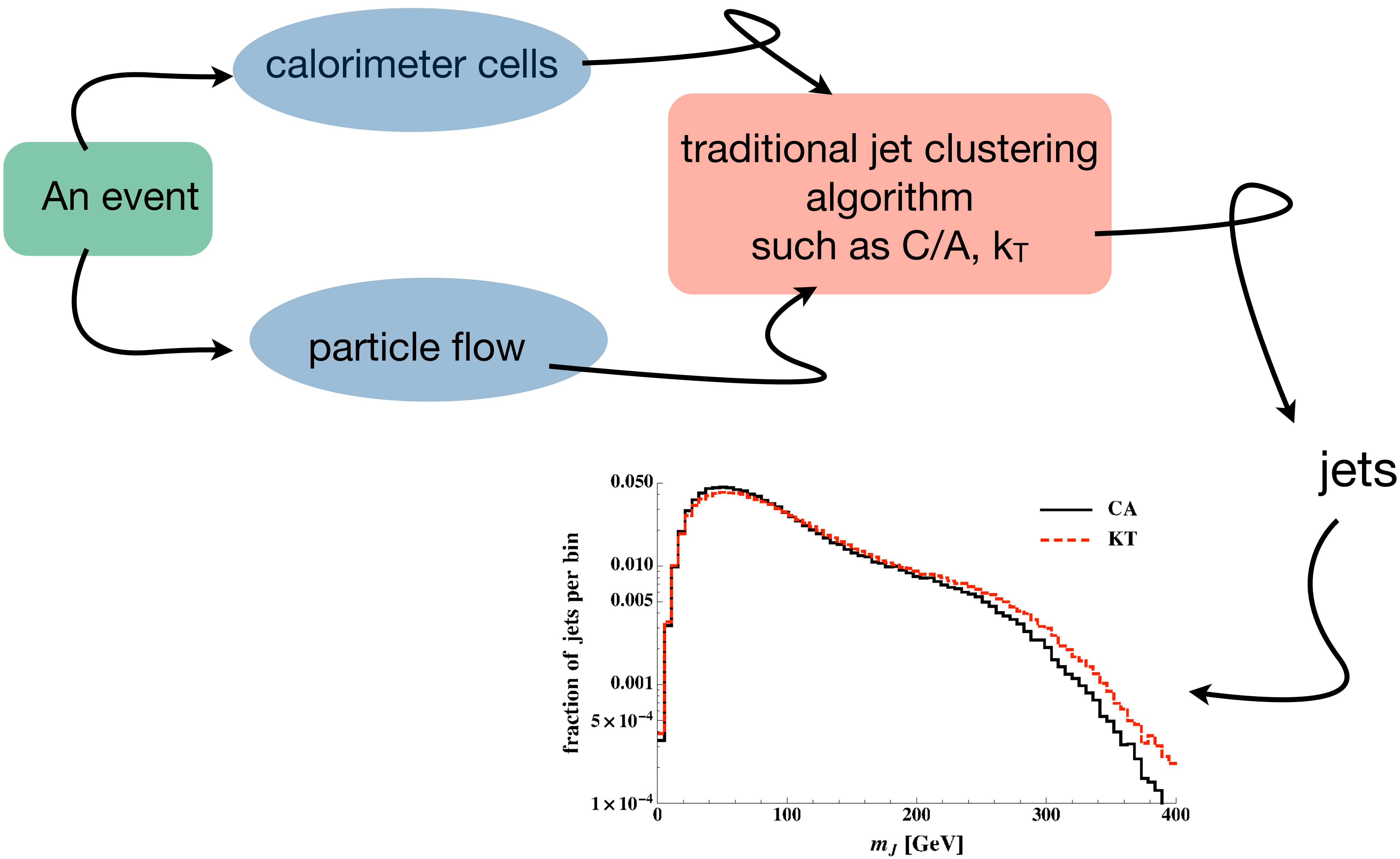
Outline

- **Introduction**
  - The central idea behind Qjets.
  - How Qjets can be used with Pruning
  - The Qjets clustering algorithm
- **QJets + Pruning**
  - Volatility and its uses
    - ➡ results of ATLAS studies.
  - Statistics in the context of Qjets
    - ➡ Applications in stabilizing statistical fluctuations in jet measurements.

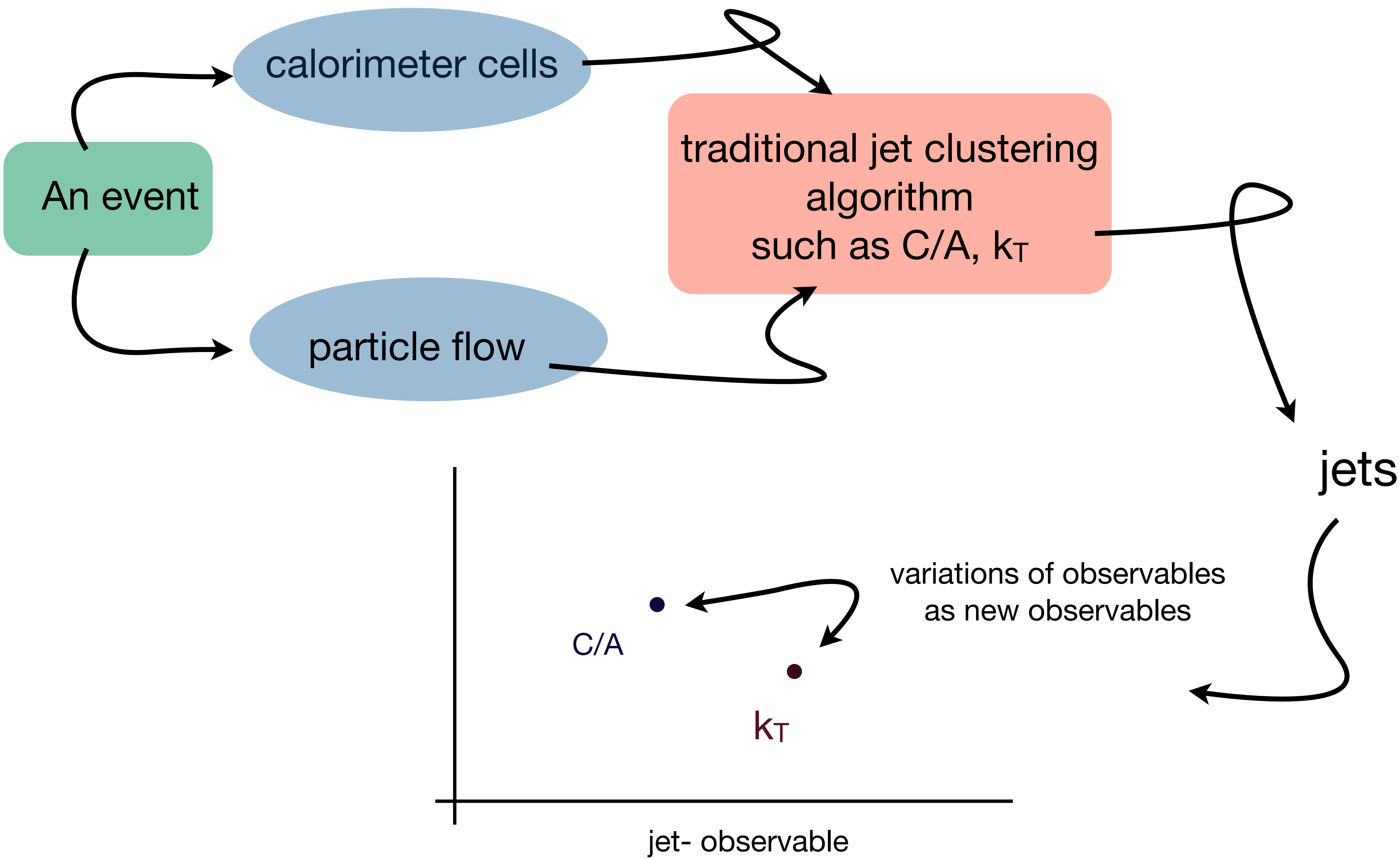
QJets: *The Idea*



QJets: *The Idea*

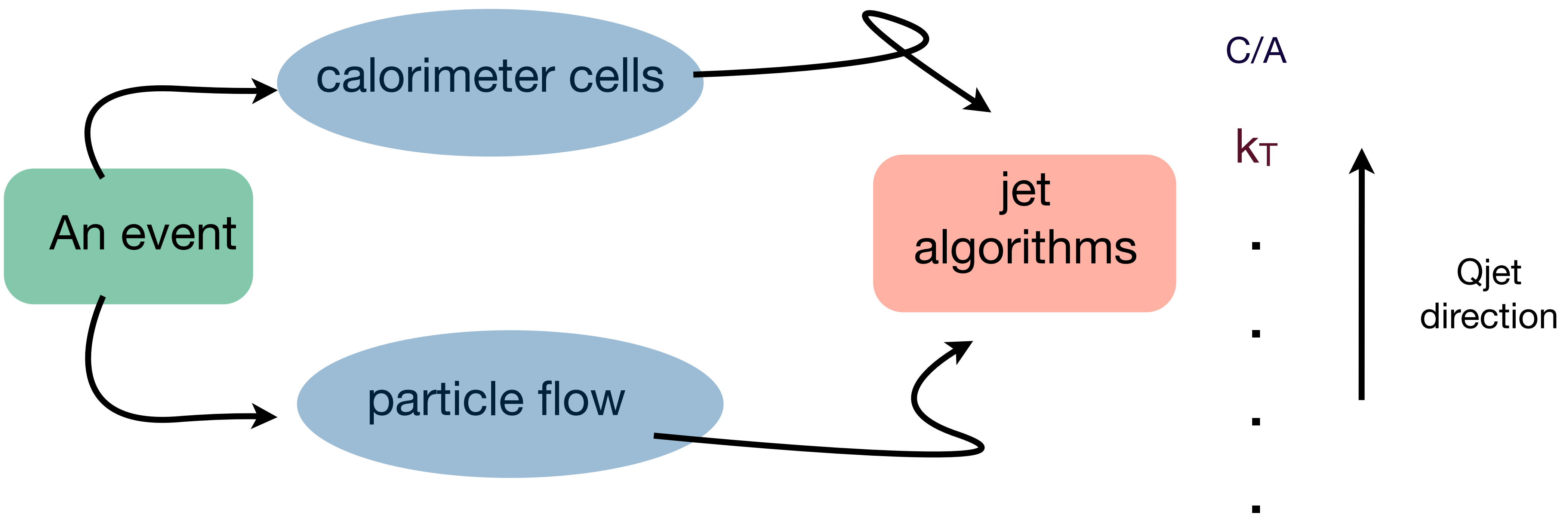


QJets: *The Idea*



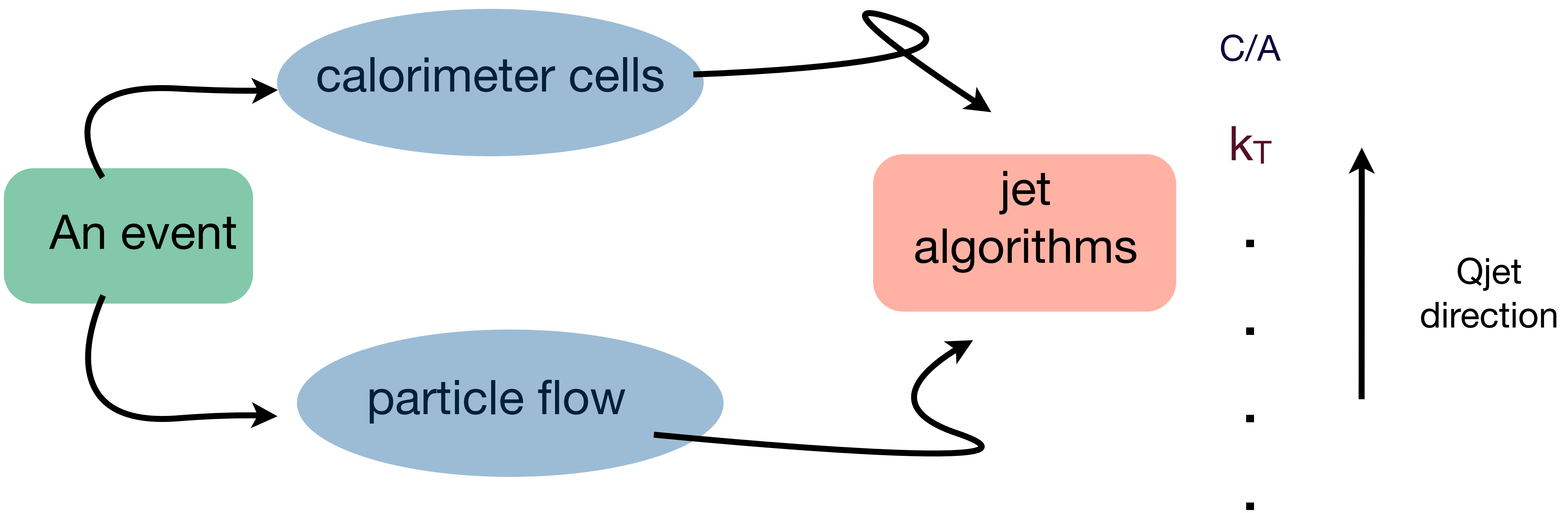
QJets: The Idea

Qjets is an idea that explores the dimension of clustering history



QJets: The Idea

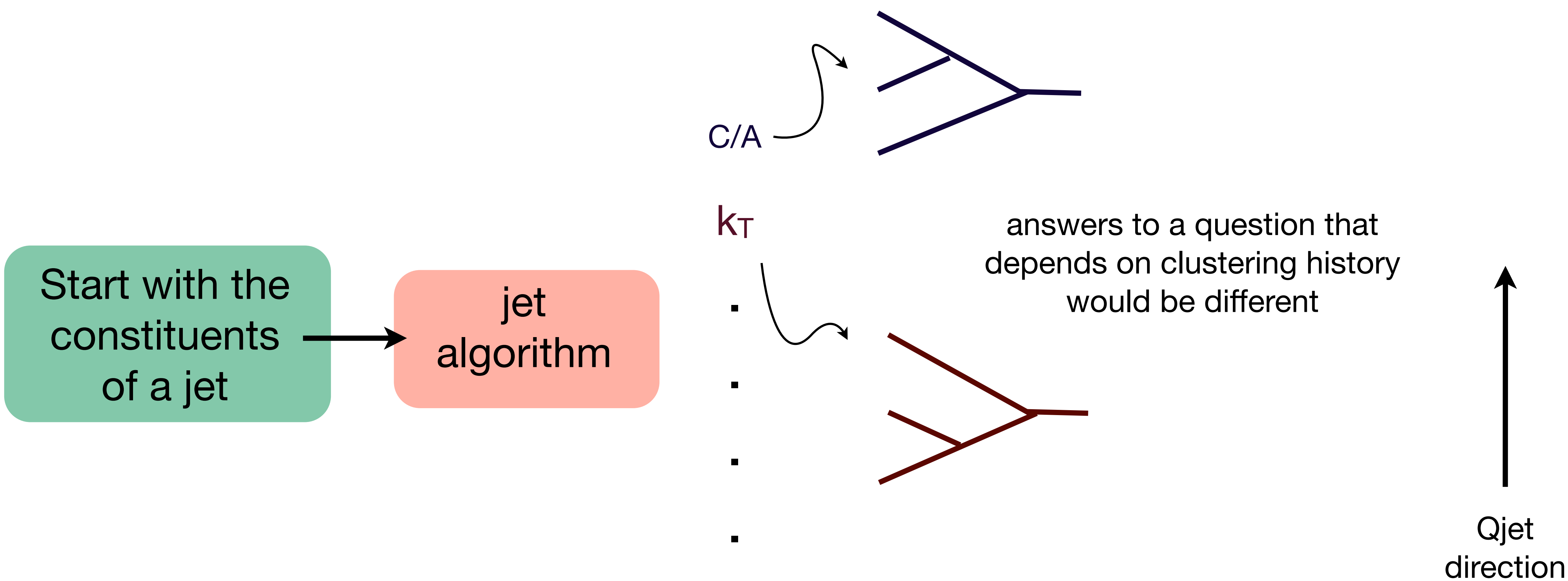
Qjets is an idea that explores the dimension of clustering history



it is a challenging task -- let us start with something simpler

QJets: The Idea

Qjets is an idea that explores the dimension of clustering history



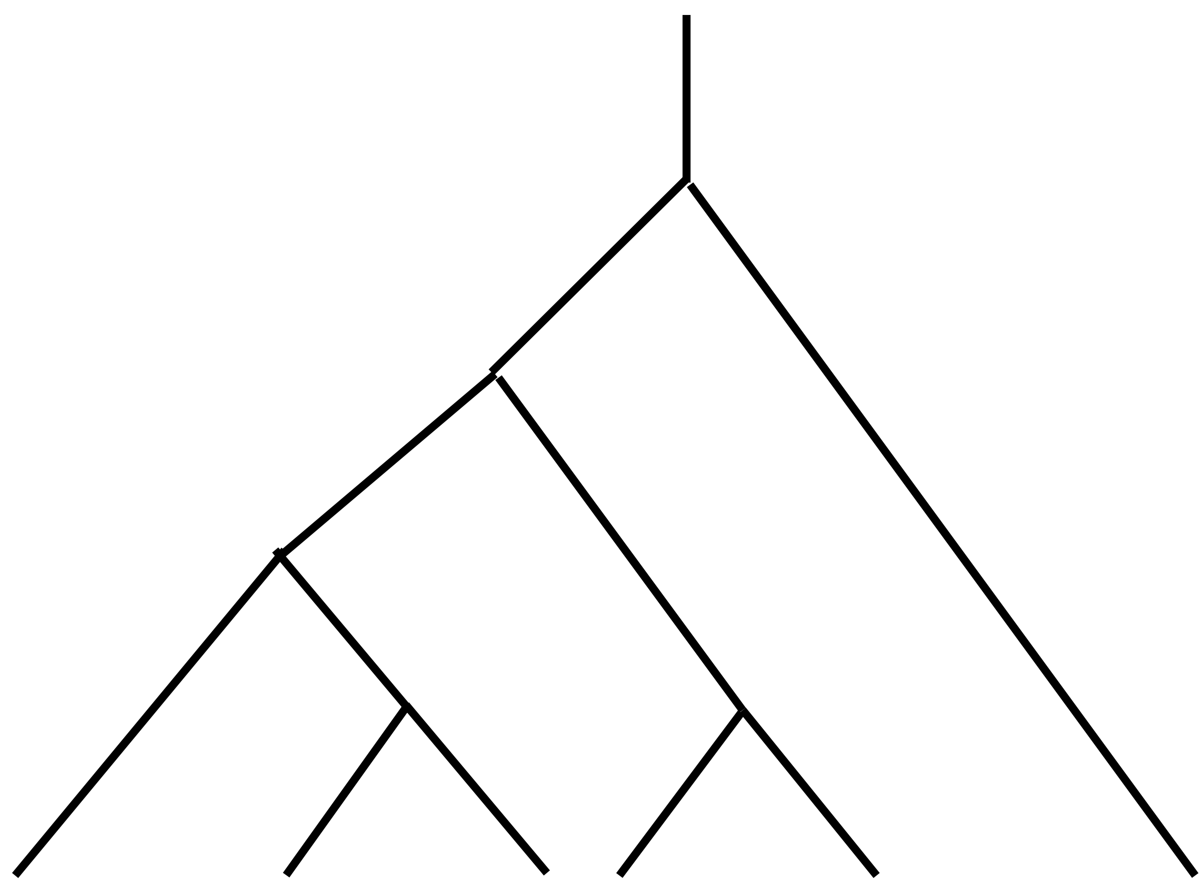
it is a challenging task -- let us start with something simpler

Qjets give different answers to
a question that
depends on clustering history

ex: pruning

Pruning

Start with the constituents of a given jet and rebuild the jet along C/A or k_T



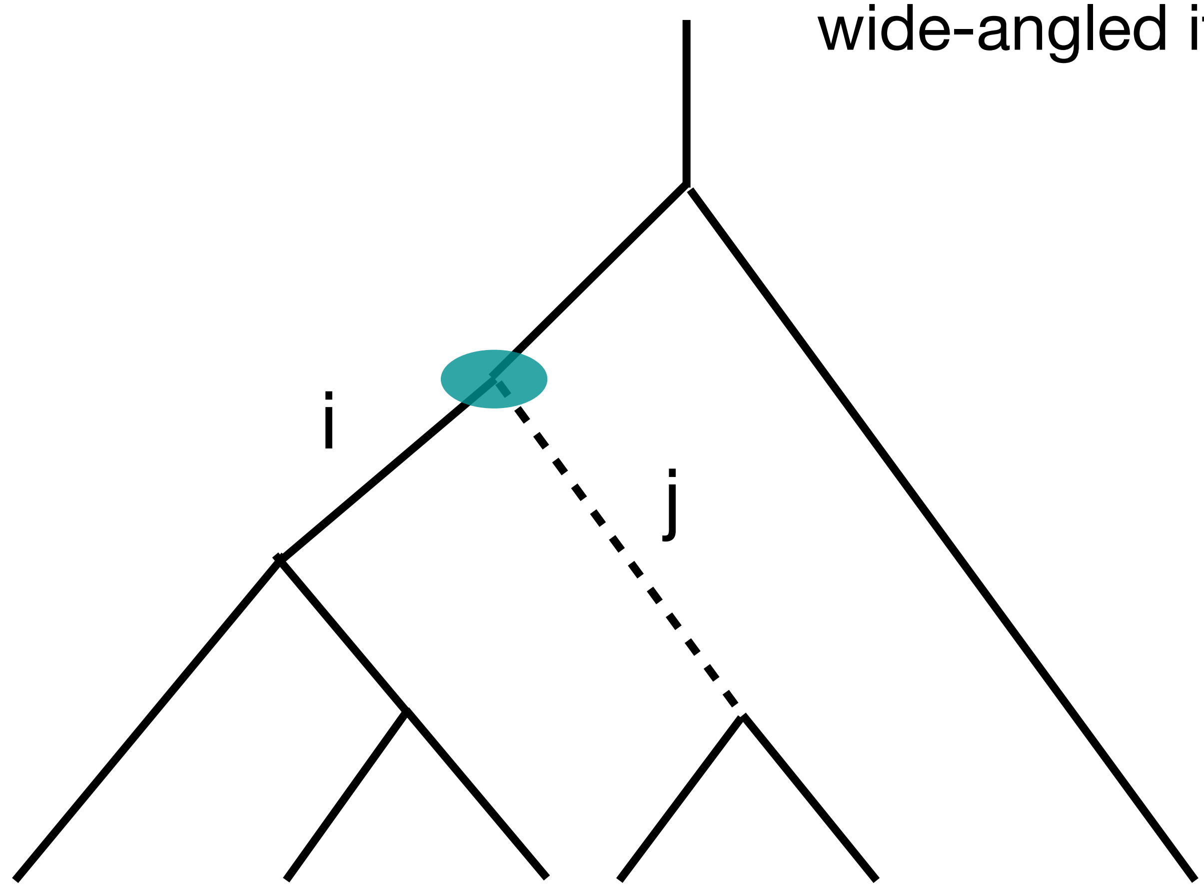
Pruning

At every step of clustering check whether the branch to be added is soft **and** wide angled.

- if yes discard the softer four-vector.

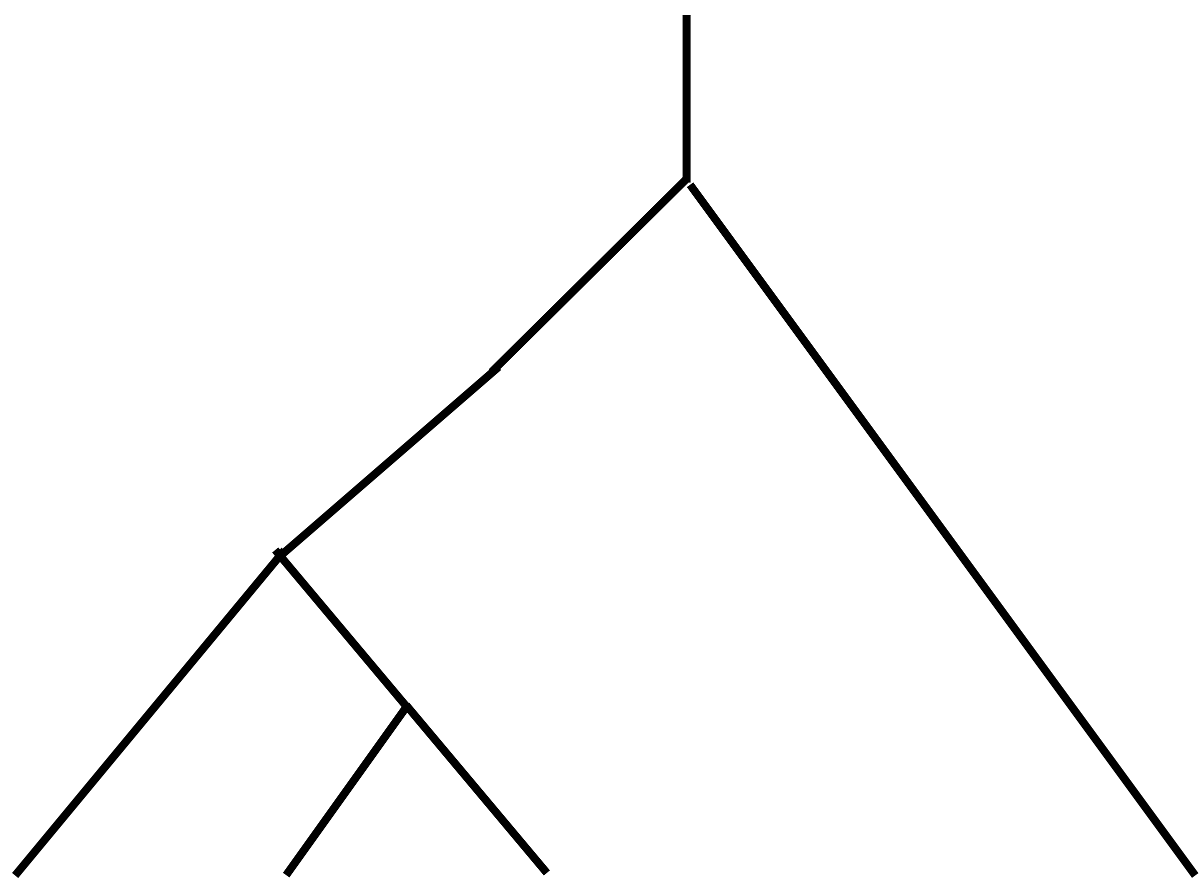
soft if: $\dfrac{\min\left(p_{T_i}, p_{T_j}\right)}{|\vec{p_{T_i}} + \vec{p_{T_j}}|} < z_{\rm cut}$

wide-angled if: $\Delta R_{ij} > D_{\rm cut}$



Pruning

Pruned Jet



Pruning

- Four-vectors that are pruned are actually branches of the tree.

- Pruned jets depend crucially on the tree-structure or the clustering algorithm used to construct the jet.

QJets Clustering

Steve Ellis, Andrew Hornig, David Krohn,
Matt Schwartz and TSR
arXiv:1201.1914
Phys.Rev.Lett. 108 (2012)

- As in a sequential recombination algorithm, assign every pair of four-vectors a distance measure d_{ij}.

- However, unlike a normal sequential algorithm (where the pair with the smallest measure is clustered), here a given pair is randomly selected for merging with probability

rigidity parameter

$$\Omega_{ij} = \frac{1}{N} \exp\left(-\alpha \frac{d_{ij}}{d_{\min}}\right)$$

$$d_{\min} = \mathrm{Min}\left(\{d_{ij}\}\right)$$

QJets Clustering

$$\Omega_{ij} = \frac{1}{N} \exp\left(-\alpha \; \frac{d_{ij}}{d_{\min}}\right)$$

d_{ij} : we take C/A or kT measure

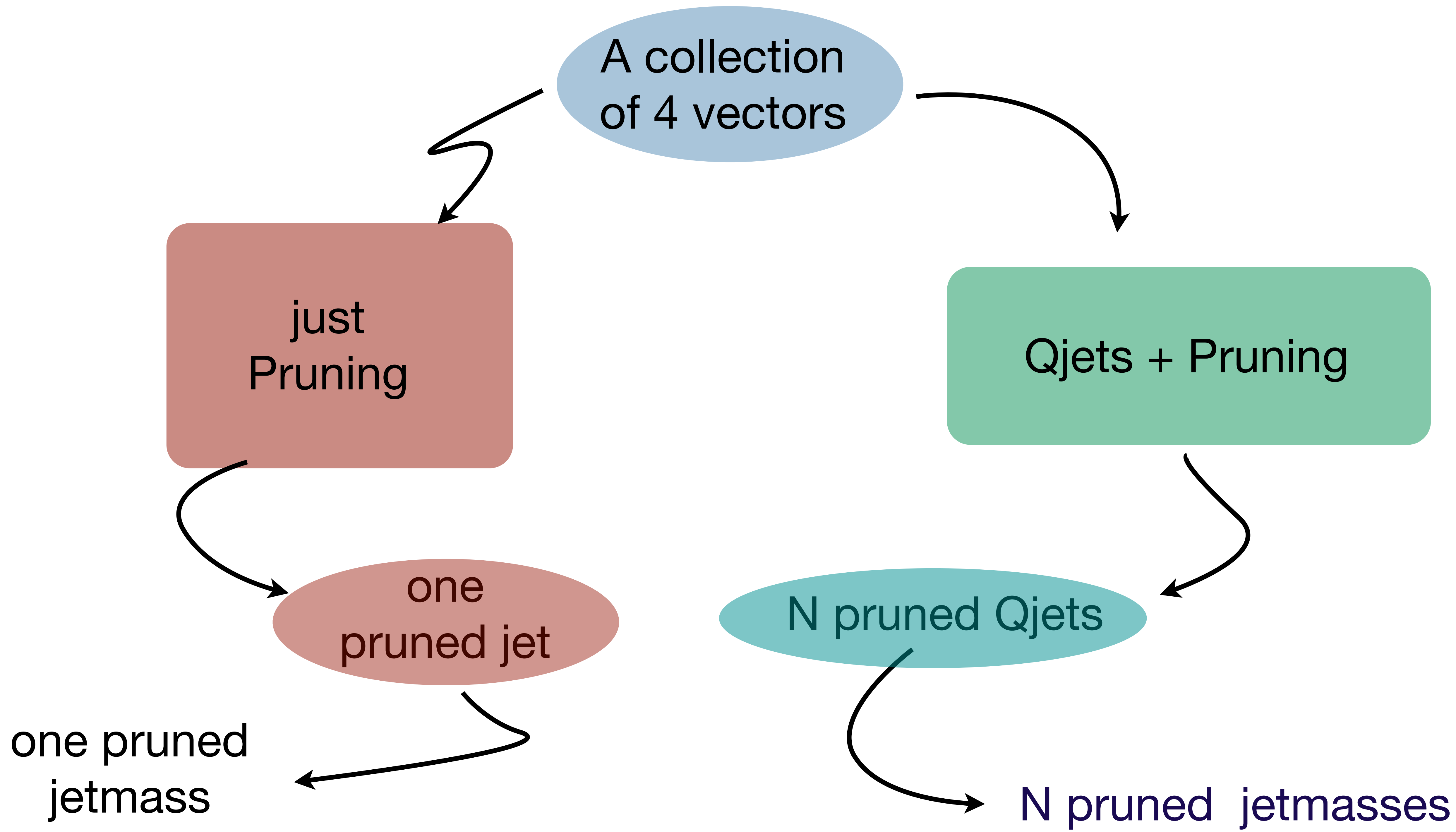
$\alpha \to \infty$ Classical regime: only path corresponding to d_{min} is selected

$\alpha > 0$ physical regime: physical paths are preferred

$\alpha \to 0$ democratic regime: all paths have same weight

$\alpha < 0$ unphysical regime: physical paths are de-weighted

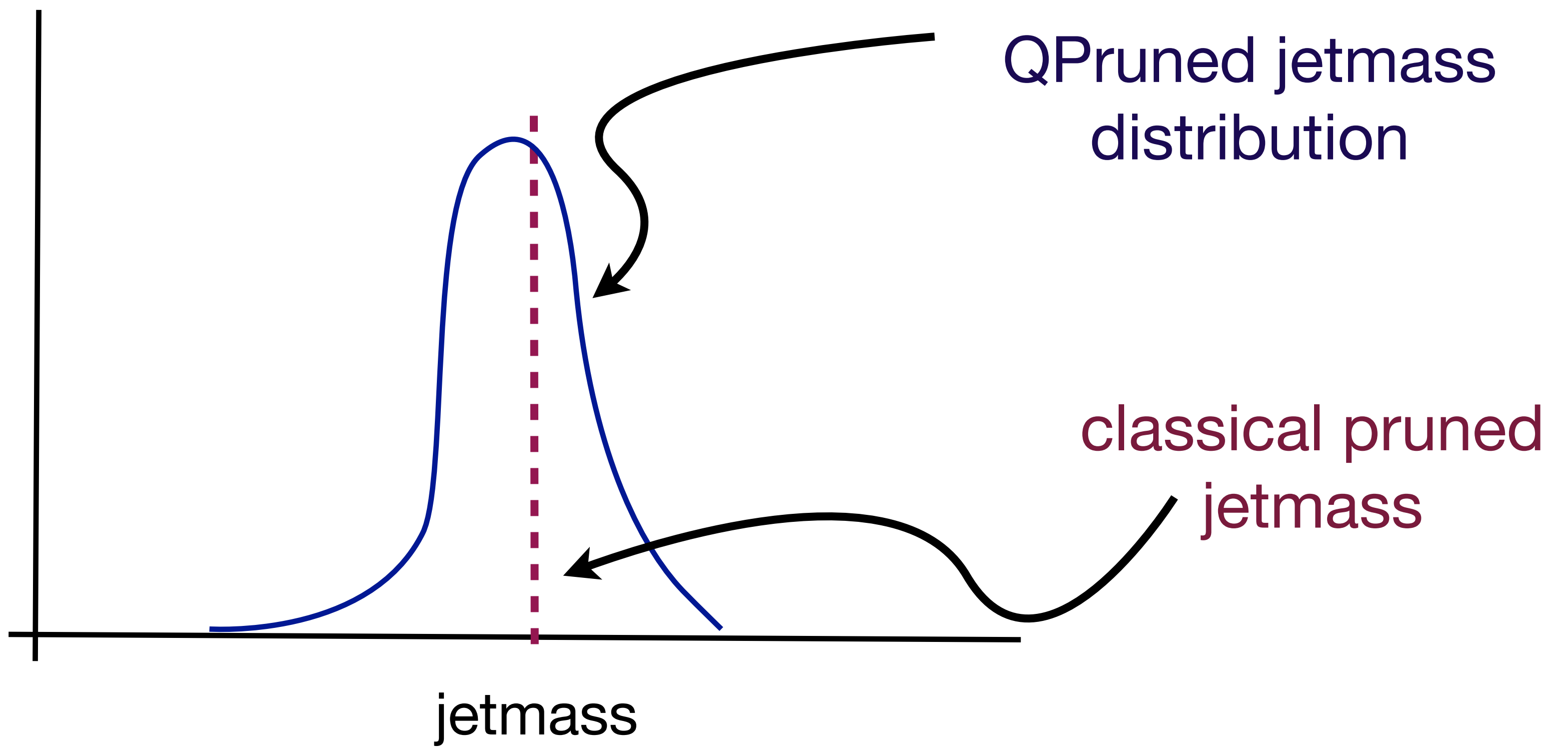
QJets + Pruning



QJets + Pruning

Ex. W decaying to two quarks

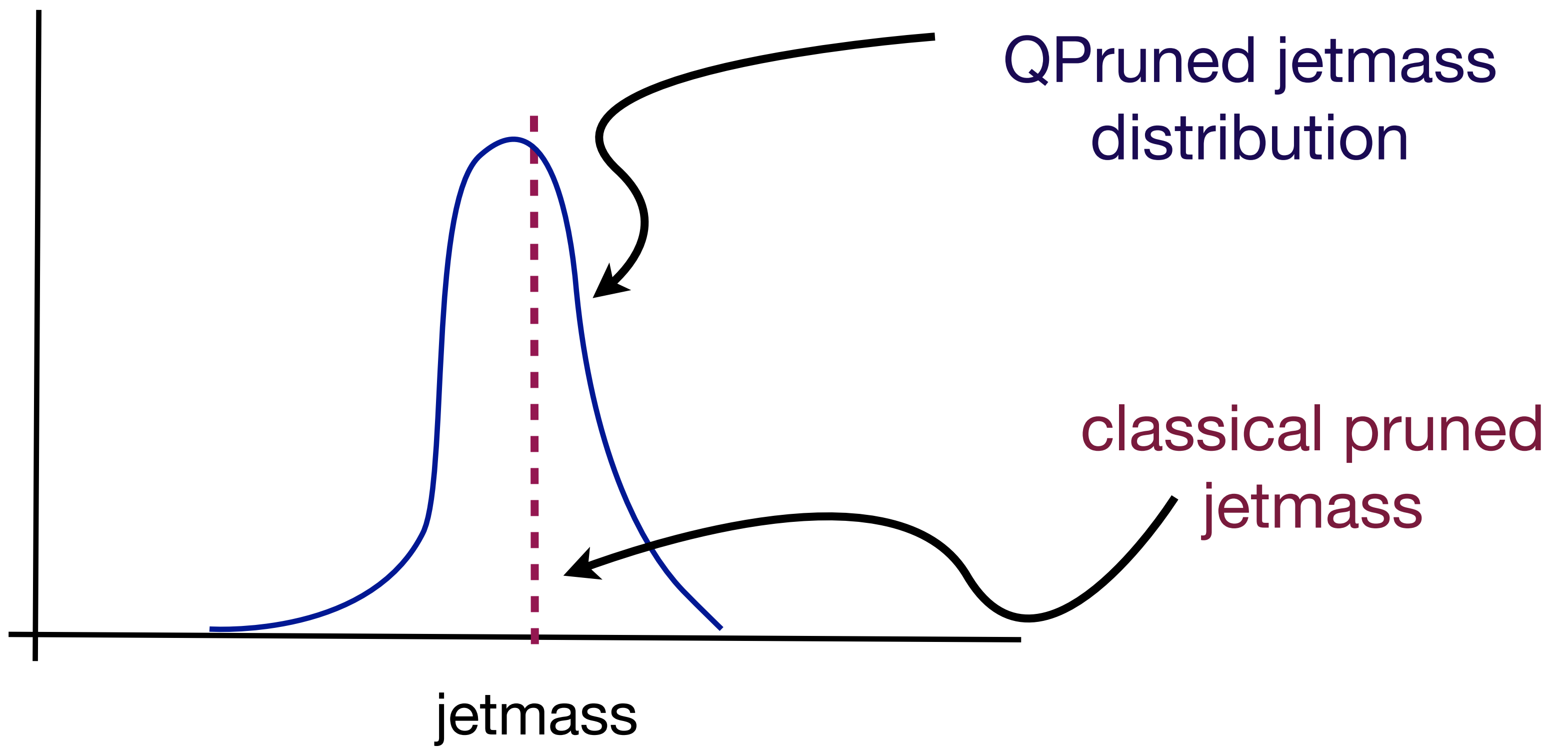
The original jet is made using anti-kT algorithm with R = 0.7 and pT > 500GeV



QJets + Pruning

Ex. W decaying to two quarks

The original jet is made using anti-kT algorithm with R = 0.7 and pT > 500GeV



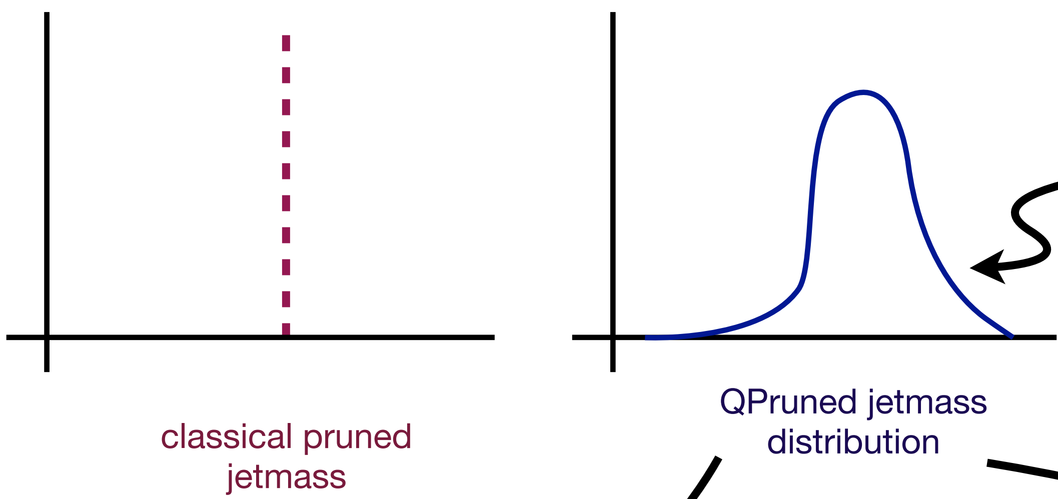
How can this distribution be used?

Outline

- Introduction
  - The central idea behind Qjets.
  - How Qjets can be used with Pruning
  - The Qjets clustering algorithm
- QJets + Pruning
  - Volatility and its uses
    - ➡ results of ATLAS studies.
  - Statistics in the context of Qjets
    - ➡ Applications in stabilizing statistical fluctuations in jet measurements.

QPruning vs. Pruning

Let us take a sample jet



How can this distribution be used?

Simply use the shape of the distribution to discriminate signal from background

Application in signal discovery

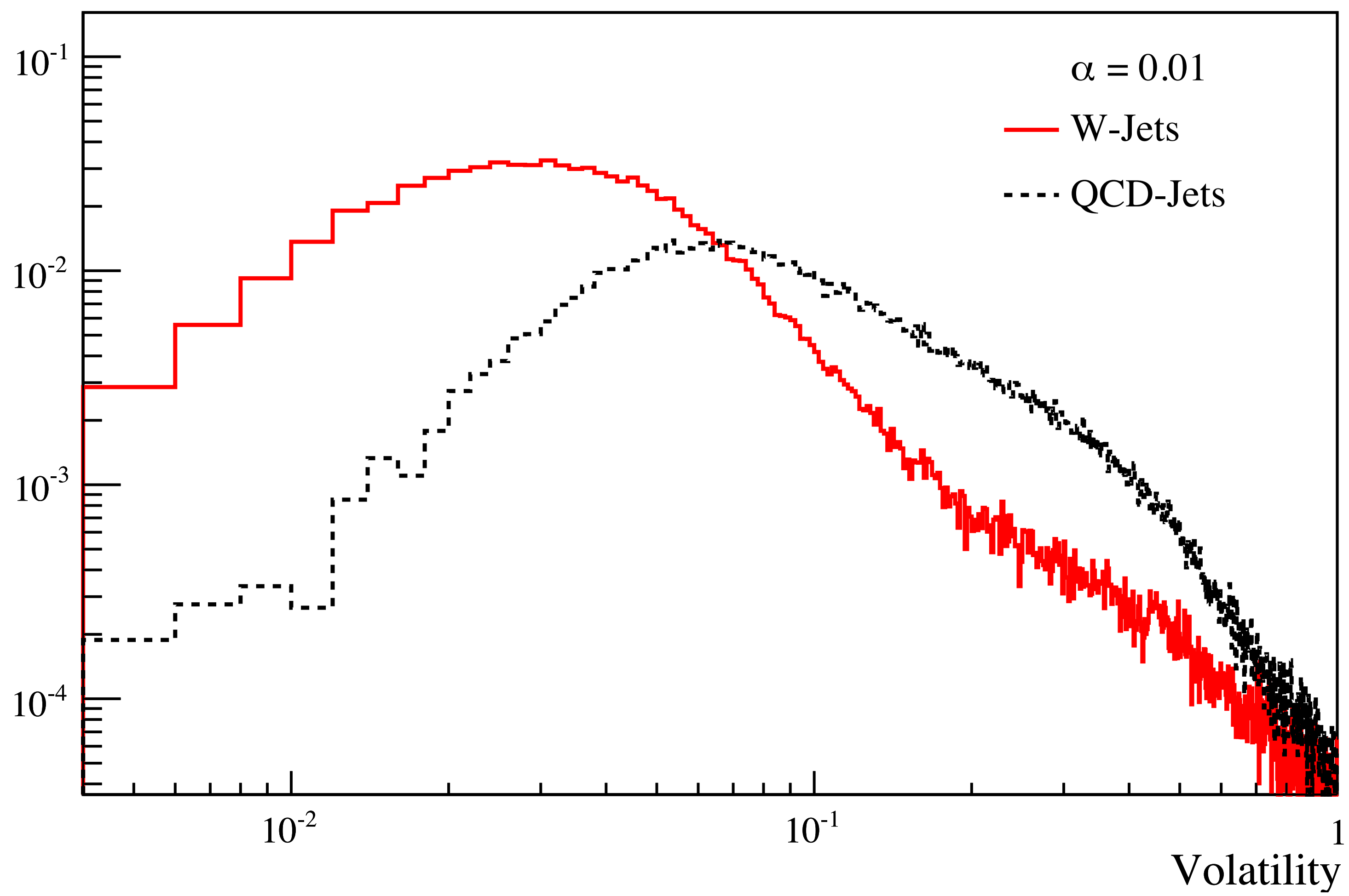
Use the distribution to reduce statistical fluctuations in measurements

Application in determination of cross-section, mass etc.

Volatility of a jet

volatility of a jet $$\mathcal{V} = \frac{\omega_p}{m_p}$$

ω_p = width of jetmass distribution
m_p = averaged pruned jetmass



Volatility of a jet: *experimental results*

ATLAS NOTE

ATLAS-CONF-2013-087

August 11, 2013

Performance and Validation of Q-jets at the ATLAS Detector in pp Collisions at $\sqrt{s}$ = 8 TeV in 2012

The ATLAS Collaboration

Available on the CERN CDS information server **CMS PAS JME-13-006**

CMS Physics Analysis Summary

Contact: cms-pog-conveners-jetmet@cern.ch 2013/08/15

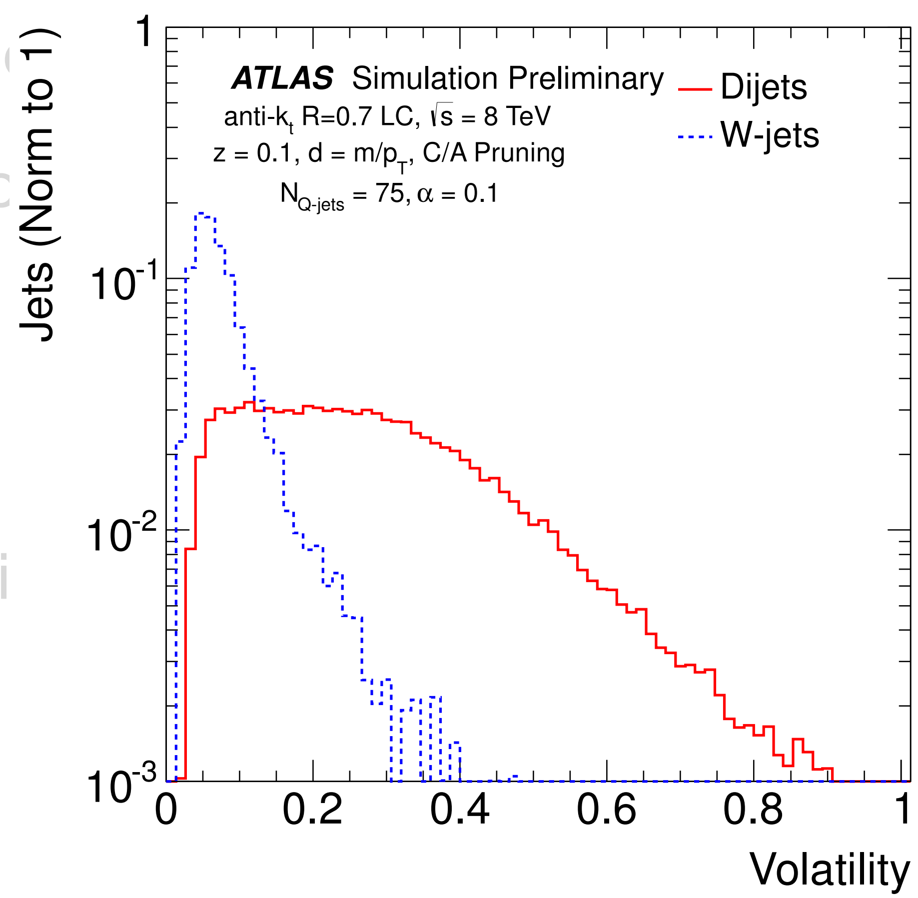
Identifying Hadronically Decaying W Bosons Merged into a Single Jet

The CMS Collaboration

Volatility of a jet: *experimental results*

Summary:

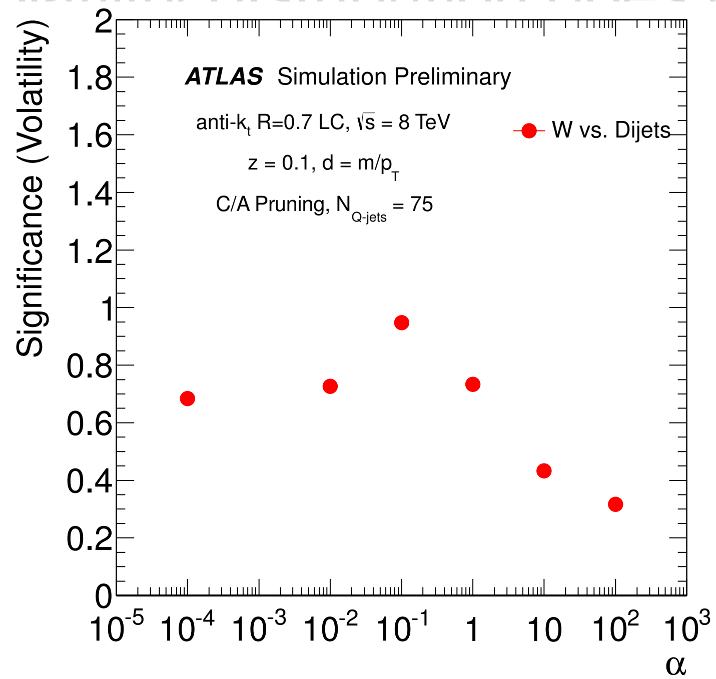
- The volatility distribution for QCD jets and W-jets have been reproduced.
- The optimized separation is
- Volatility distribution does n -up
- good data/MC agreement
- 15 QCD jet rejection at 50%
- Comparable performance wi
- Volatility and N-subjettiness)
  correlated !

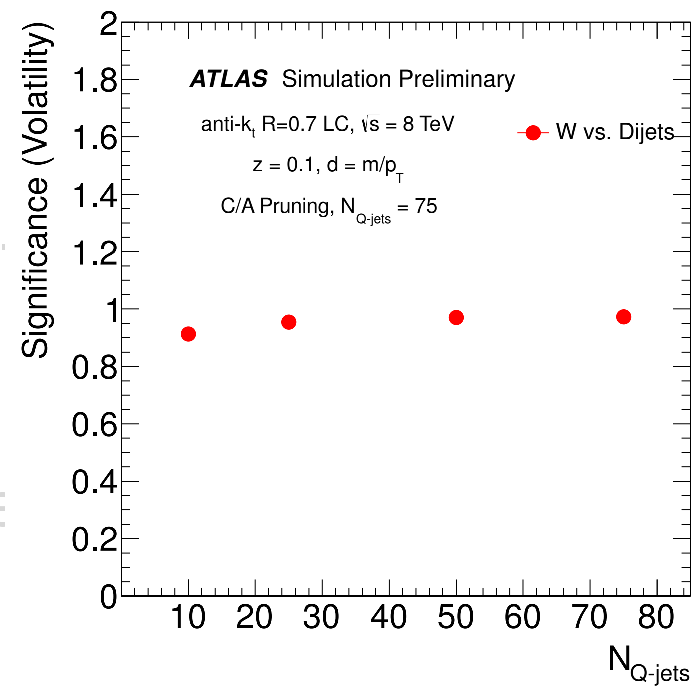


Volatility of a jet: experimental results

Summary:

- The volatility distribution for QCD jets and W-jets have been reproduced.
- The optimized separation is obtained for α=0.1 and N_{Qjets} > 25
- Volatility distribution does not have strong dependence on Pile-up
- g
- 1 … W
- C … with
- V … are … (CMS)
  c

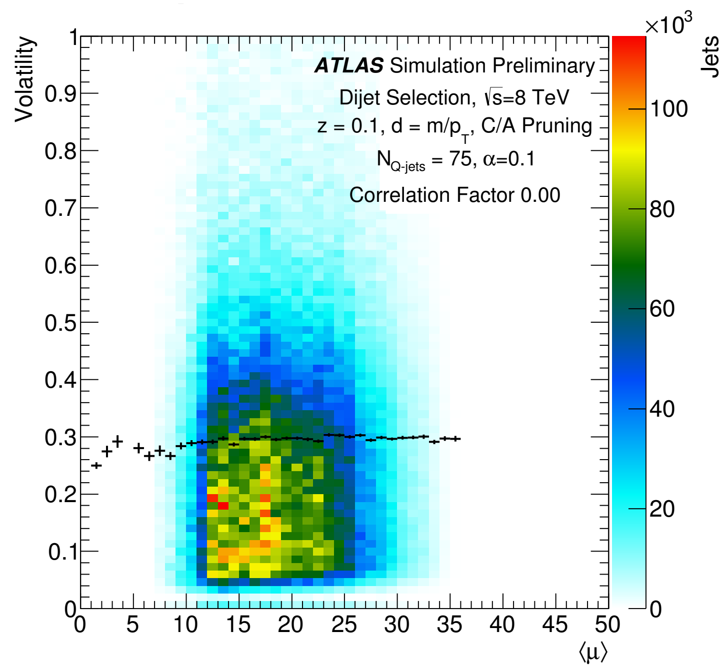




Volatility of a jet: *experimental results*

Summary:

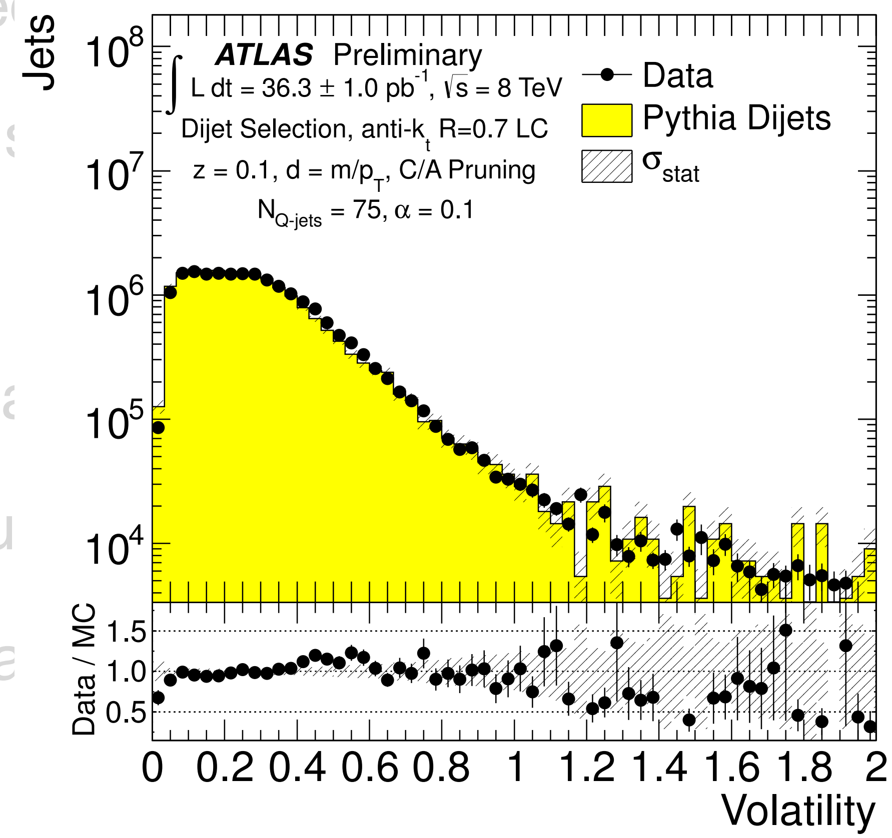
- The volatility distribution for QCD jets and W-jets have been reproduced.
- The optimized separation is obtained for α=0.1 and $N_{Qjets} > 25$
- **Volatility distribution does not have strong dependence on Pile-up**
- good data/MC agre
- 15 QCD jet rejectio ncy
- Comparable perfor s
- Volatility and N-sub S)/ strongly (CMS) correlated !



Volatility of a jet: *experimental results*

Summary:

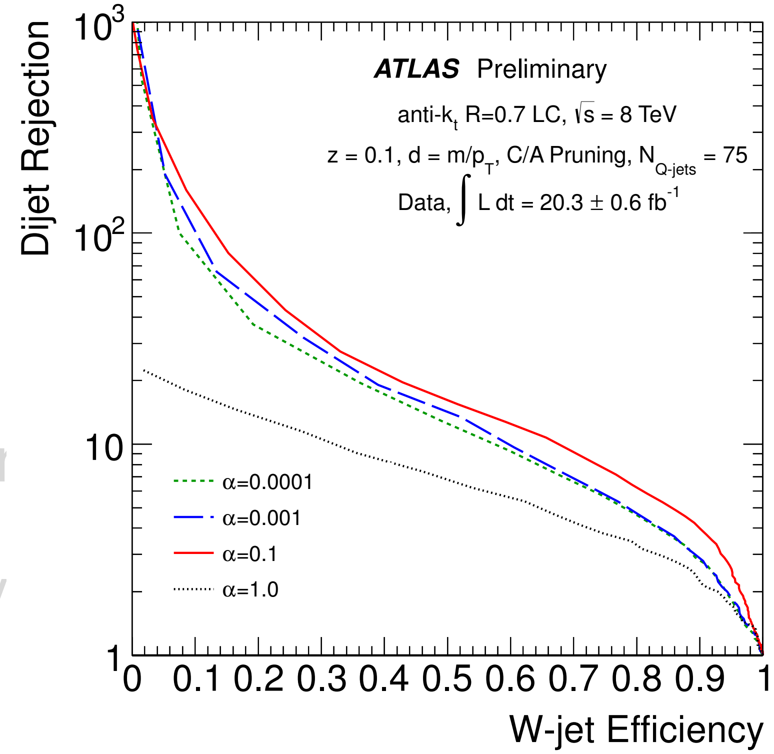
- The volatility distribution for QCD jets and W-jets have been reproduced.
- The optimized separation is obtaine
- Volatility distribution does not have s
- good data/MC agreement
- 15 QCD jet rejection at 50% W-jet ta
- Comparable performance with N-su
- Volatility and N-subjettiness are wea correlated !



Volatility of a jet: *experimental results*

Summary:

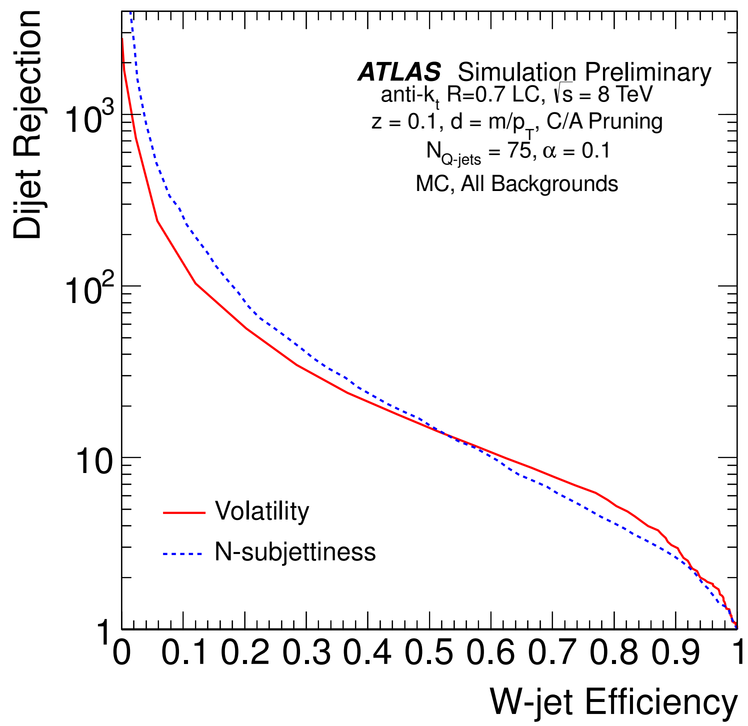
- The volatility distribution for QCD
  reproduced.
- The optimized separation is obtai 5
- Volatility distribution does not hav -up
- good data/MC agreement
- **15 QCD jet rejection at 50% W-jet tag efficiency**
- Comparable performance with N-subjettiness
- Volatility and N-subjettiness are weakly correlated -- suggests useful potential combination

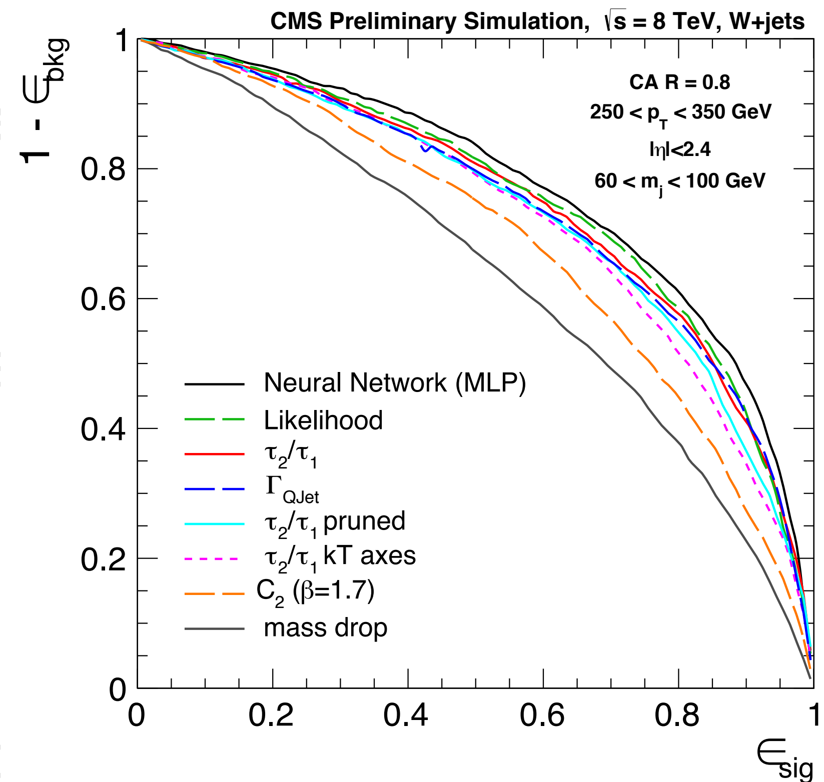


Volatility of a jet: *experimental results*

Summary:

- The v… je…
  reproc…
- The o… ne…
- Volatil… e… o
- good
- 15 QCD jet rejection at 50% W-jet t…

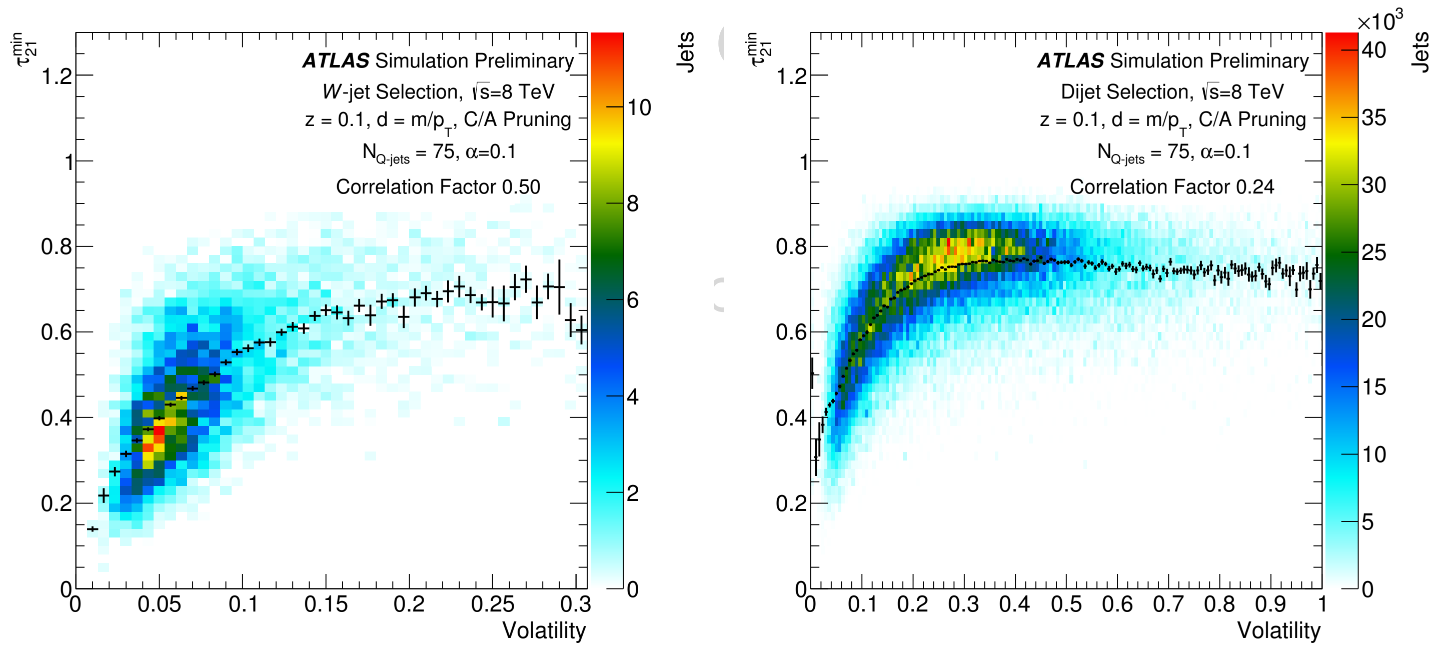




- Comparable performance with N-subjettiness
- Volatility and N-subjettiness are weakly (ATLAS)/ strongly (CMS) correlated !

Volatility of a jet: experimental results

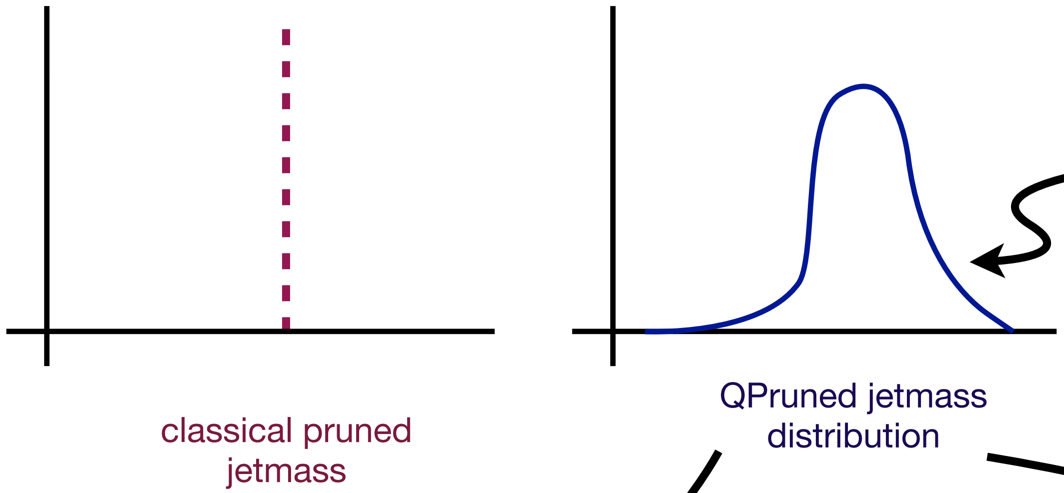
Summary:



- 15 QCD jet rejection at 50% W-jet tag efficiency
- Comparable performance with N-subjettiness
- Volatility and N-subjettiness are weakly (ATLAS)/ strongly (CMS) correlated !

QPruning vs. Pruning

Let us take a sample jet



How can this distribution be used?

Simply use the shape of
the distribution
to discriminate signal from
background

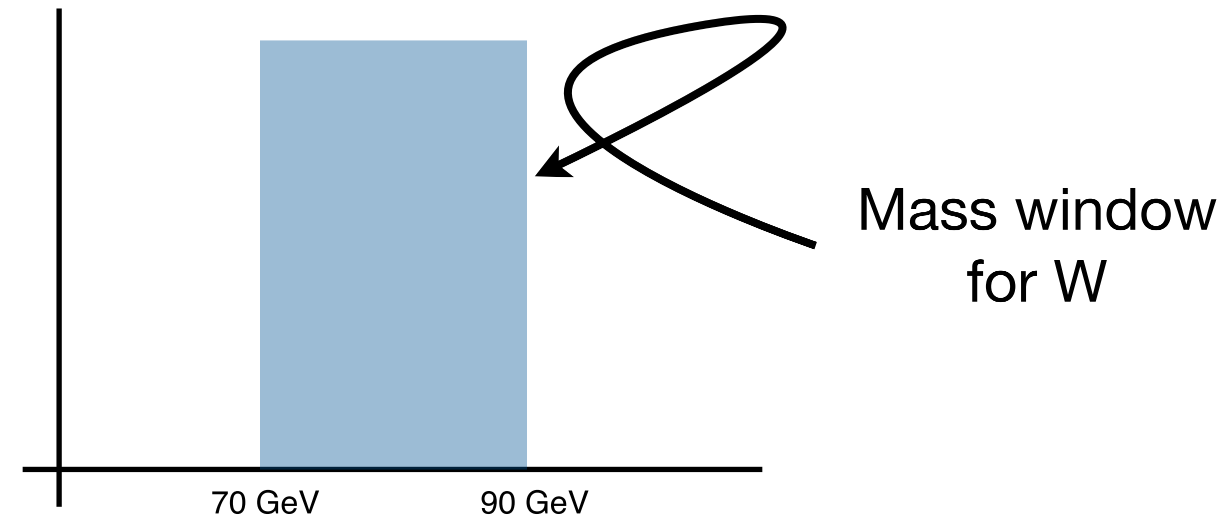
Application in signal discovery

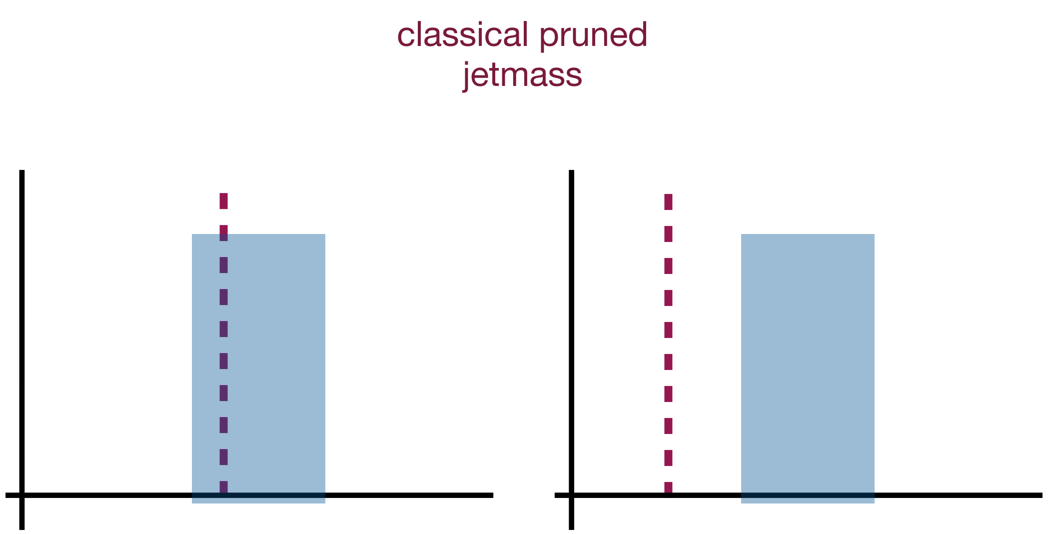
Use the distribution
to reduce statistical
fluctuations in
measurements

Application in determination of
cross-section, mass etc.

QPruning vs. Pruning

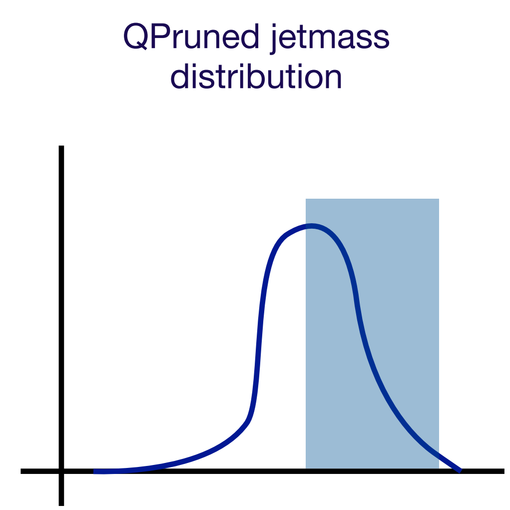
Consider candidates for a W jet





pruned mass is
either in or out of the bin

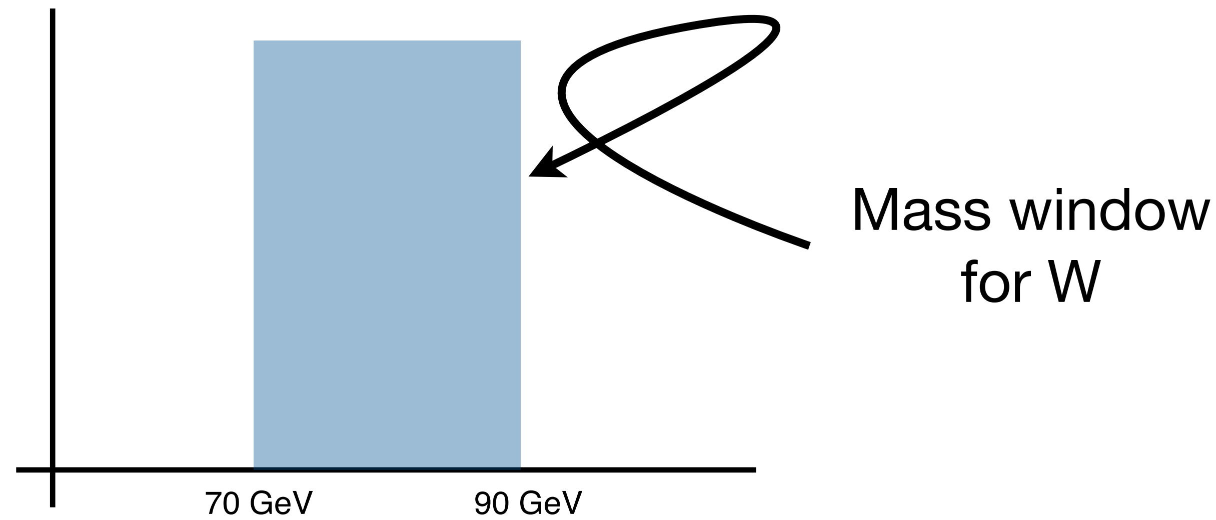
tagging efficiency is either 0 or 1

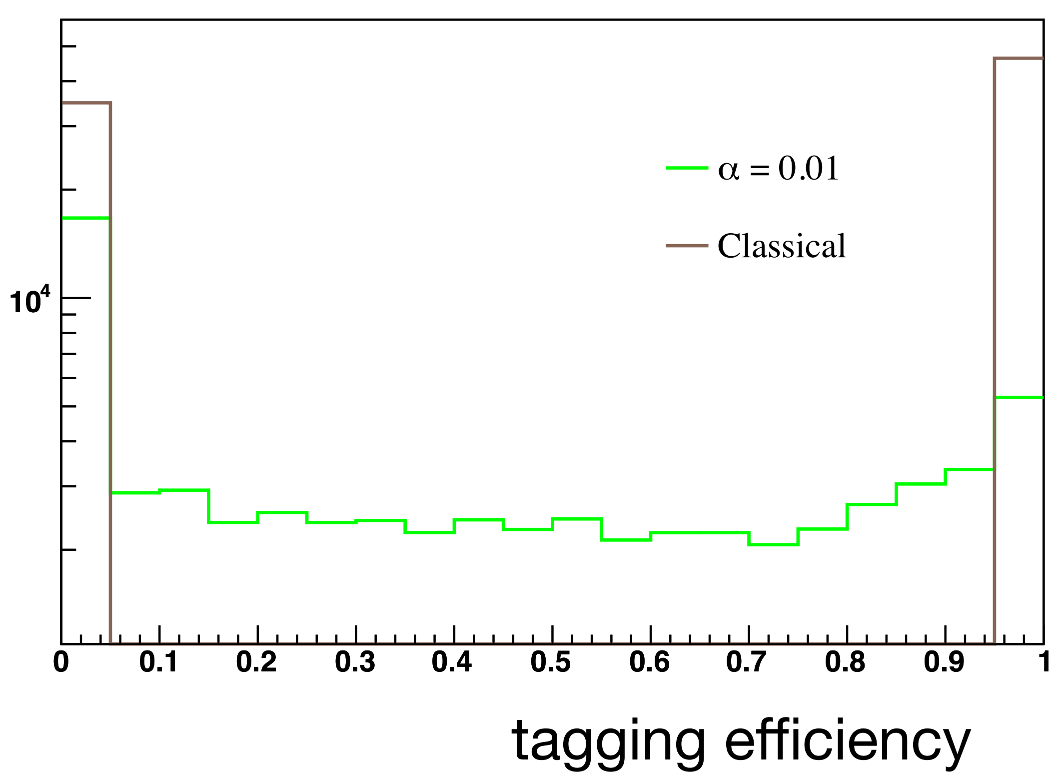


tagging efficiency is a number
between 0 to 1

QPruning vs. Pruning

Consider candidates for a W jet





Pruning -> QPruning

A transition from a discrete (binomial distribution) to a continuous distribution

Statistical aspects of QJets

Consider a master set of jets

In an experiment

- pick N jets at random out of the master set
- Qjet+prune every jet and determine
  1. tagging efficiency (τ_j)
  2. An observable averaged over all qjets (q_j)

     ex. average mass of the distribution (μ_j)

Experimental Observables:

$$N_{\mathrm{expt}} = \sum_j \tau_j \qquad\qquad Q_{\mathrm{expt}} = \frac{\sum_j \tau_j q_j}{\sum_j \tau_j}$$

we want to determine δN_{expt} and δQ_{expt}

Statistical aspects of QJets

we want to determine δN_{expt} and δQ_{expt}

In an experiment

- pick N jets at random out of the master set
- Qjet+prune every jet and determine
  1. tagging efficiency (τ_j)
  2. An observable averaged over all qjets (q_j)

     ex. average mass of the distribution (μ_j)

Standard Procedure:

- repeat the experiment many times and calculate δN_{expt} and δQ_{expt}

one can derive δN_{expt} and δQ_{expt} analytically

Statistical aspects of QJets

one can derive δN_{expt} and δQ_{expt} analytically

$$\delta N_{\text{expt}} = \sqrt{N_{\text{expt}}} \times \sqrt{\langle\tau\rangle + \frac{\text{var}(\tau)}{\langle\tau\rangle}}$$

$$\delta Q_{\text{expt}} = \frac{\langle Q_{\text{expt}}\rangle}{\sqrt{N}} \times \sqrt{\frac{\text{var}(\tau)}{\langle\tau\rangle^2} + \frac{\text{var}(q\tau)}{\langle q\tau\rangle^2} - 2\frac{\text{cov}(\tau, q\tau)}{\langle\tau\rangle\langle q\tau\rangle}} + \mathcal{O}\left(\frac{1}{N}\right)$$

Statistical aspects of QJets

one can derive δN_{expt} and δQ_{expt} analytically

$$\delta N_{\rm expt} \;=\; \sqrt{N_{\rm expt}} \times \sqrt{\langle\tau\rangle + \frac{{\rm var}(\tau)}{\langle\tau\rangle}}$$

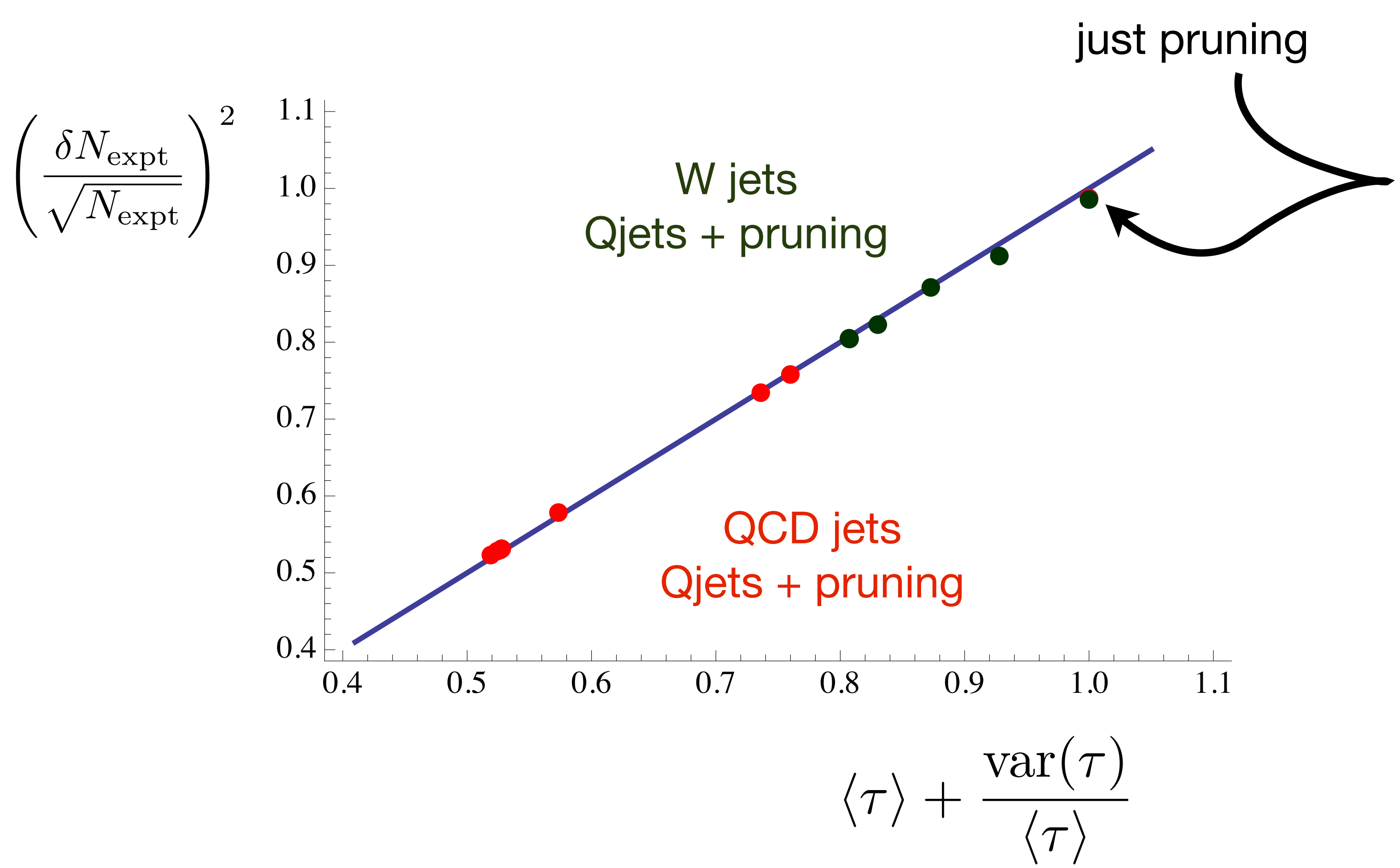
$$\sqrt{\langle\tau\rangle + \frac{{\rm var}(\tau)}{\langle\tau\rangle}} \;\leq\; 1 \text{ for all distributions}$$

$$=\; 1 \text{ for binomial distributions (i.e. for pruning)}$$

Statistical aspects of QJets

one can derive δN_{expt} and δQ_{expt} analytically

$$\delta N_{\rm expt} = \sqrt{N_{\rm expt}} \times \sqrt{\langle\tau\rangle + \frac{{\rm var}(\tau)}{\langle\tau\rangle}}$$



Statistical aspects of QJets

one can derive δN_{expt} and δQ_{expt} analytically

$$\delta Q_{\rm expt} = \frac{\langle Q_{\rm expt}\rangle}{\sqrt{N}} \times \sqrt{\frac{\mathrm{var}(\tau)}{\langle\tau\rangle^2} + \frac{\mathrm{var}(q\tau)}{\langle q\tau\rangle^2} - 2\frac{\mathrm{cov}(\tau, q\tau)}{\langle\tau\rangle\langle q\tau\rangle}} + \mathcal{O}\left(\frac{1}{N}\right)$$

simplifies for just pruning

$$\sqrt{\frac{\mathrm{var}\,(q\tau)_0}{\langle q\tau\rangle_0^2} - \frac{\mathrm{var}\,(\tau)_0}{\langle\tau\rangle_0^2}}$$

However, depending on the algorithm δQ_{expt} can go either way

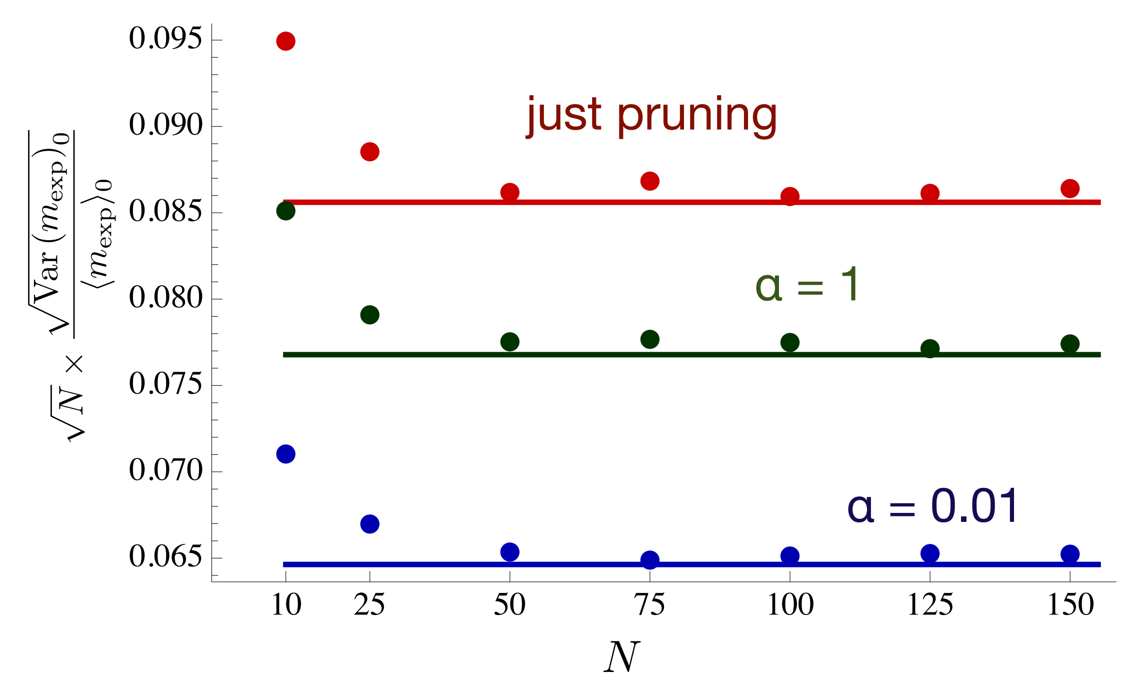
Statistical aspects of QJets

one can derive δN_{expt} and δQ_{expt} analytically

$$\delta Q_{\text{expt}} = \frac{\langle Q_{\text{expt}}\rangle}{\sqrt{N}} \times \sqrt{\frac{\text{var}(\tau)}{\langle\tau\rangle^2} + \frac{\text{var}(q\tau)}{\langle q\tau\rangle^2} - 2\frac{\text{cov}(\tau, q\tau)}{\langle\tau\rangle\langle q\tau\rangle}} + \mathcal{O}\left(\frac{1}{N}\right)$$

For numerical example, we use

q = μ = average jetmass in the bin



Conclusion

Grooming tools (pruning, trimming, filtering) even though designed for boosted search, are useful and essential for non-boosted cases.

We introduced QJets: a non-deterministic jet clustering algorithm.

- QJets Clustering lets us look inside a jet in a new way.
- QJets + pruning renders stability to jet observables and provides new discriminants for the discovery of signal jets.