Physics Analyses with Jets

The CMSDAS Jet team

Outline

- Concepts
  - What is a Jet?
    - Jets in CMS
    - Jet calibration
  - Practical Instructions for Beginners
    - How to access jets
    - How to apply JEC
    - Jet Composition
    - Jet Resolution
    - Jet substructure
  - Links to More Information

Why should you care about jets ?

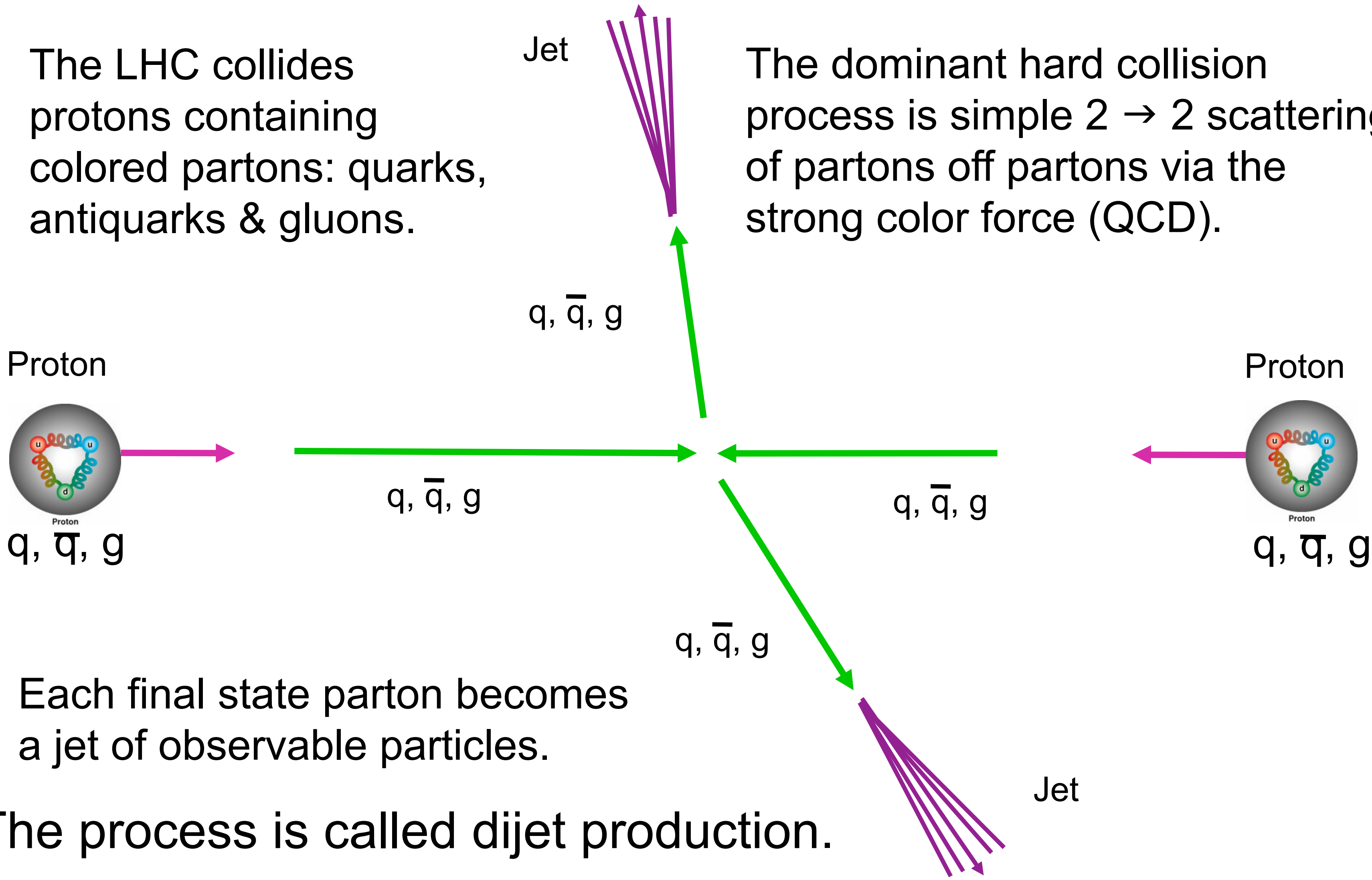
- Jets are everywhere; their cross section is orders of magnitude higher than most other processes.
- Jets can fake as γ, e, μ, τ
  - -Probability of jet faking a γ ~10^{-4}
  - -Probability of faking e/μ ~10^{-5}, but some jets have real lepton, e.g., b-jets
  - -Probability of faking a τ ~10^{-3}
- Light quark or gluon jets can fake b-quark jet at the % level
- Missing Transverse Energy must be corrected for jet energy measurements.

If jets are not your signal they are most certainly your background !

Jets at LHC

The LHC collides protons containing colored partons: quarks, antiquarks & gluons.

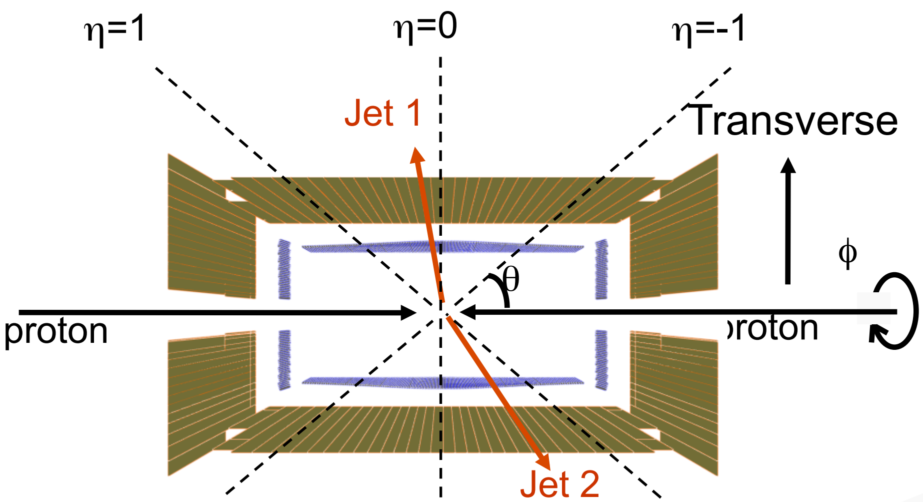
The dominant hard collision process is simple 2 → 2 scattering of partons off partons via the strong color force (QCD).

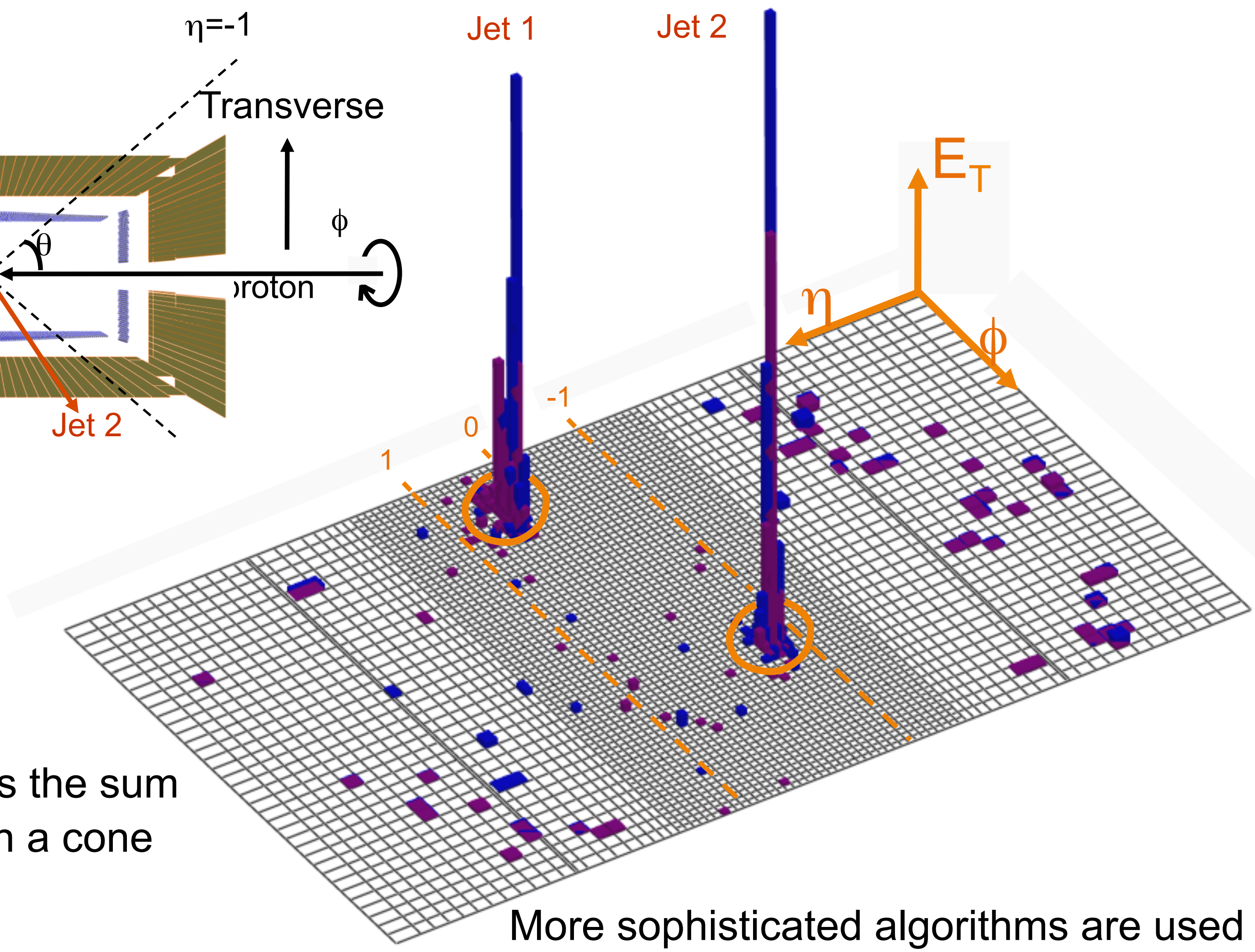


Each final state parton becomes a jet of observable particles.

The process is called dijet production.

Experimental observation of Jets





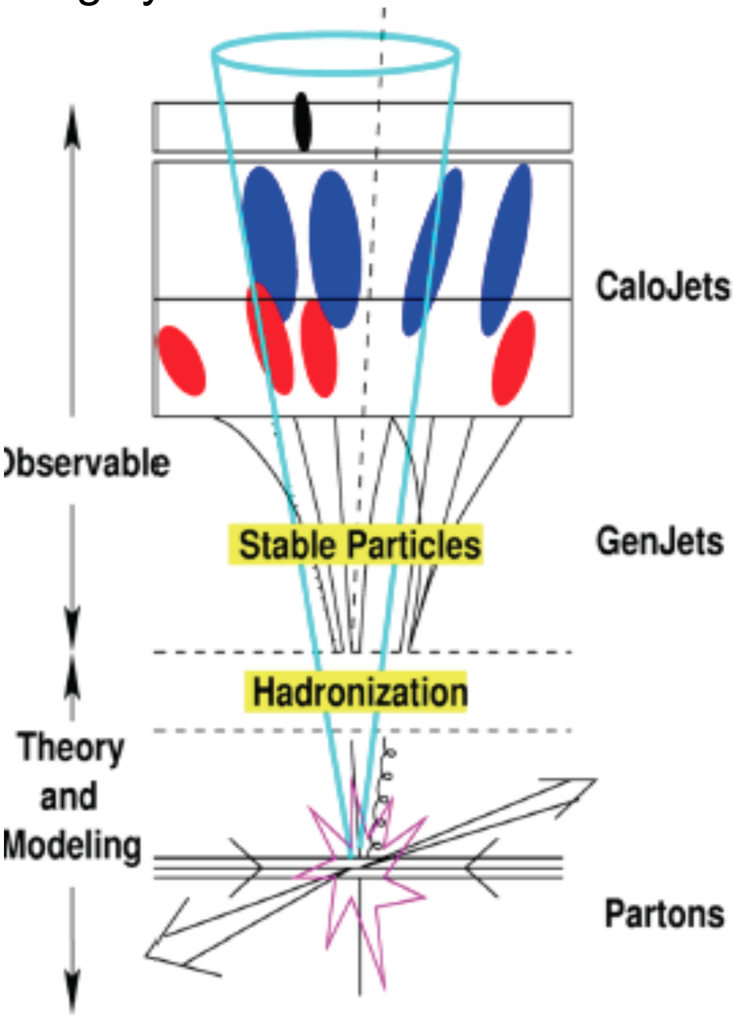
In CMS trigger, a jet is the sum of particle 4-vectors in a cone of radius

$$R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.5$$

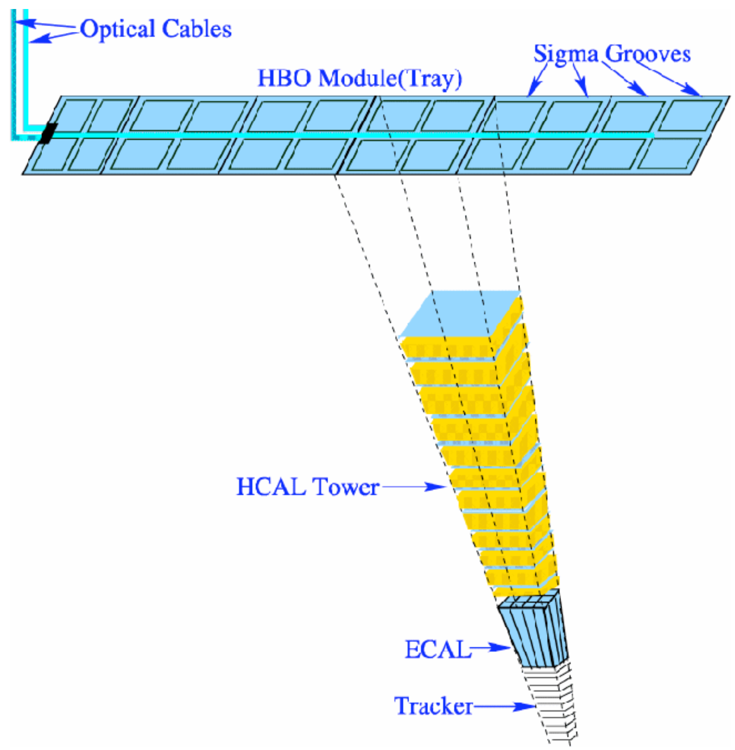
More sophisticated algorithms are used in offline analysis to reconstruct jets.

Jets in CMS

- Jets are the signature of partons, materialized as sprays of highly collimated hadrons.

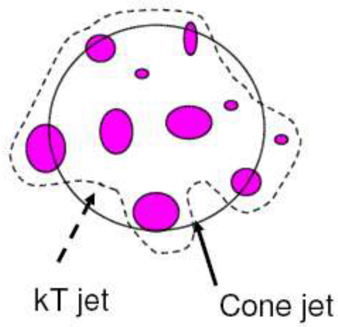


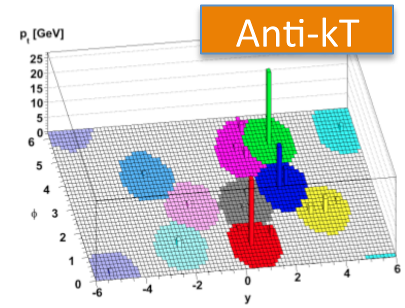
- A particle flow jet is the output of the jet finding algorithm when applied to a set of particle 4-vectors.



How are Jets defined?

- A *jet algorithm* is a set of mathematical rules that reconstruct unambiguously the properties of a jet.





CMS default

- Fixed cone algorithms:
  - ✧ Iterative Cone
  - ✧ Seedless Infrared Safe Cone (SISCone)

- Successive recombination algorithms:

$$d_{ij} = p_{T,i}^{2p} \qquad d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta R_{ij}^2}{D^2}$$

p=1 -> k_T jet algorithm
p=0 -> CA jet algorithm
p=-1 -> "Anti-k_T" jet algorithm

- Different inputs to the jet algorithm lead to different types of jets:

✧ **Calorimeter jets (CaloJets)**: Clustered from CaloTowers

✧ **Track Jets**: Clustered from charged particle tracks

✧ **Jets plus Tracks**: Correct calorimeter jets using momentum of tracks.

✧ **Particle Flow Jets**: Clustered from identified particles, reconstructed using all detector components.

Getting hands dirty: How to access Jets in CMSSSW

https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookJetAnalysis#JetAna

Example code in:
https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideCMSDataAnalysisSchoolJetAnalysis

```
#include "DataFormats/JetReco/interface/PFJetCollection.h"
#include "DataFormats/JetReco/interface/PFJet.h"
.......................................................
JetAlgorithm  = cfg.getParameter<std::string> ("JetAlgorithm");
.......................................................
edm::Handle<edm::View<reco::Jet> > jets;
evt.getByLabel(JetAlgorithm,jets);
edm::View<reco::Jet>::const_iterator i_jet, endpjets = jets->end();
```

configurable jet collection name

(via python config script)

- ak5PFJets
- ak5GenJets
- goodPatJetsPFlow

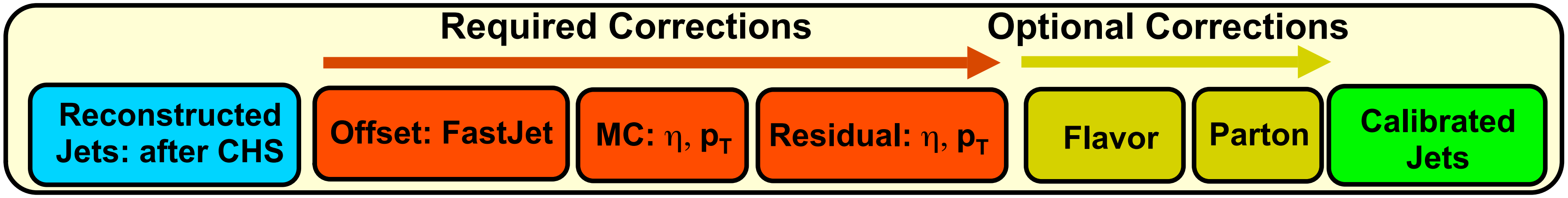
Handle the jet collection

// define an iterator for the jet collection

```
for (i_jet = jets->begin();  i_jet != endpjets ;  ++i_jet) {
      pt= (*i_jet).pt();
      eta = (*i_jet).eta();
      phi = (*i_jet).phi();
      jetArea = (*i_jet).jetArea();
}
```

Basic Jet quantities

Jet energy calibration: overview



✦ Factorization facilitates the use of data-driven corrections

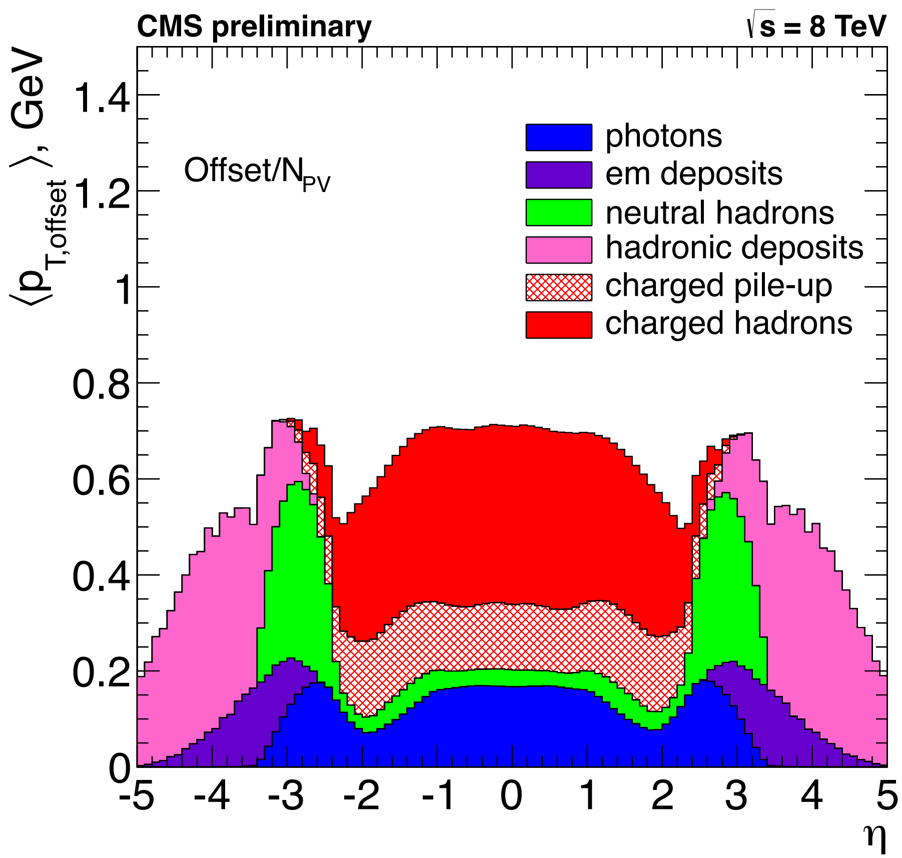
- Breaking the correction into pieces that are naturally measured in collider data:
  - **Offset**: pile-up and noise measured in zero-bias events.
  - **MC**: jet response vs. η, P_T using MC truth.
  - **Residual**: jet response vs. η, P_T using dijet balance and γ/Z +jet in data.

In CMS the most widely used jet is anti-k_T 0.5 (0.7 for QCD measurements). For pileup studies, consider anti-k_T 0.5, 0.7, 0.8 with various grooming techniques: filtering, trimming, pruning.

Pileup contribution to jet energy

✦Pileup (PU) measured with Zero Bias data and MC

- Random cone allows to separate contribution per detector
- Most charged hadrons can be associated to pileup vertices and removed



•Part that can be removed is labeled "charged hadrons"

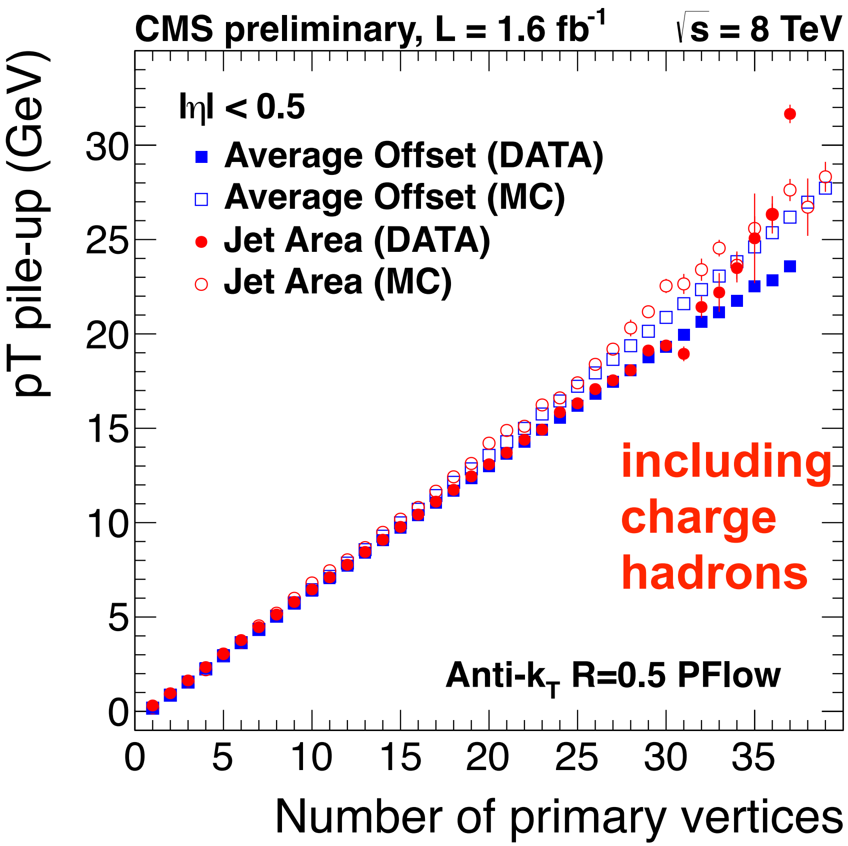
•Part that remains as PU after this needs to be subtracted

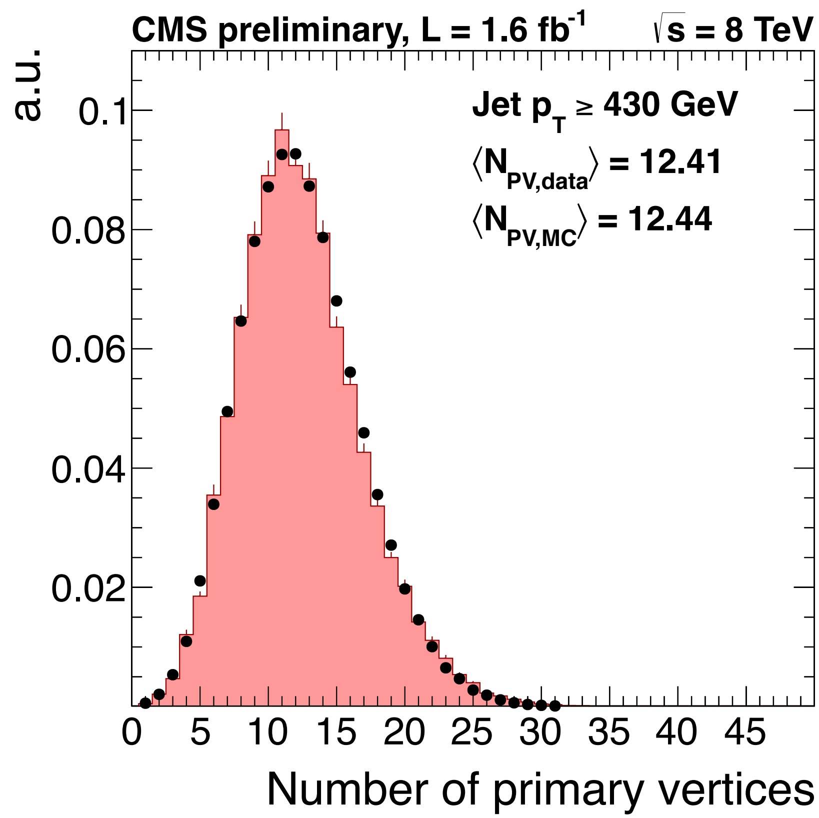
PU density x Effective area
(FastJet-ρ)

•PU density depends on the # of primary vertex in the event

Pileup correction: using area subtraction

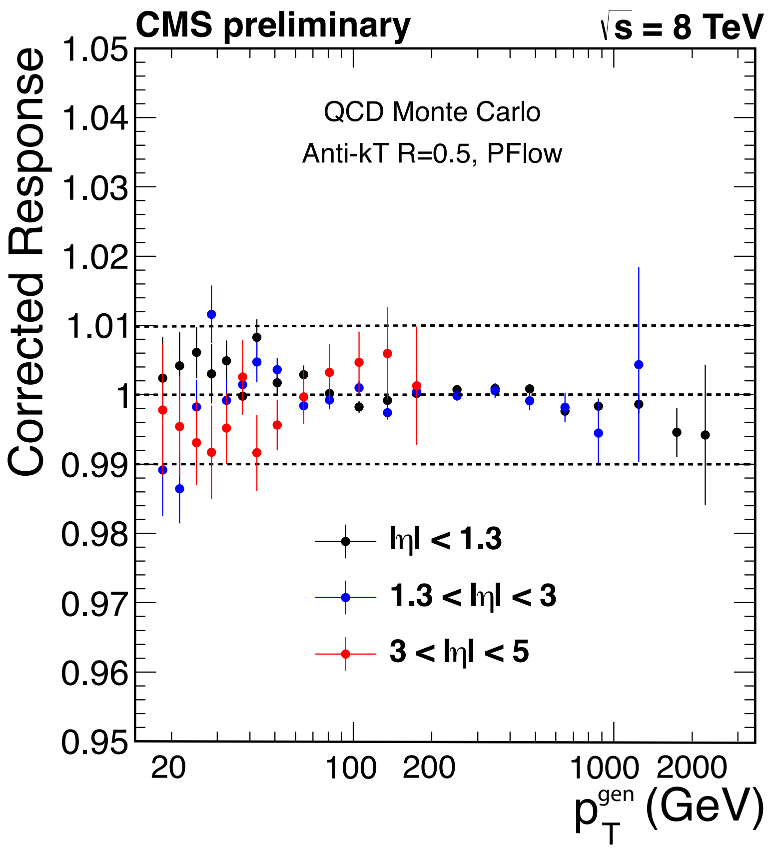
- Both NPV-based and FastJet-ρ-based corrections are in agreement
- Remaining Data/MC difference accounted for with separate PU corrections
  - Reweight pileup Poisson mean in MC to data. Poisson mean determined from measured luminosity and Minimum bias cross section.

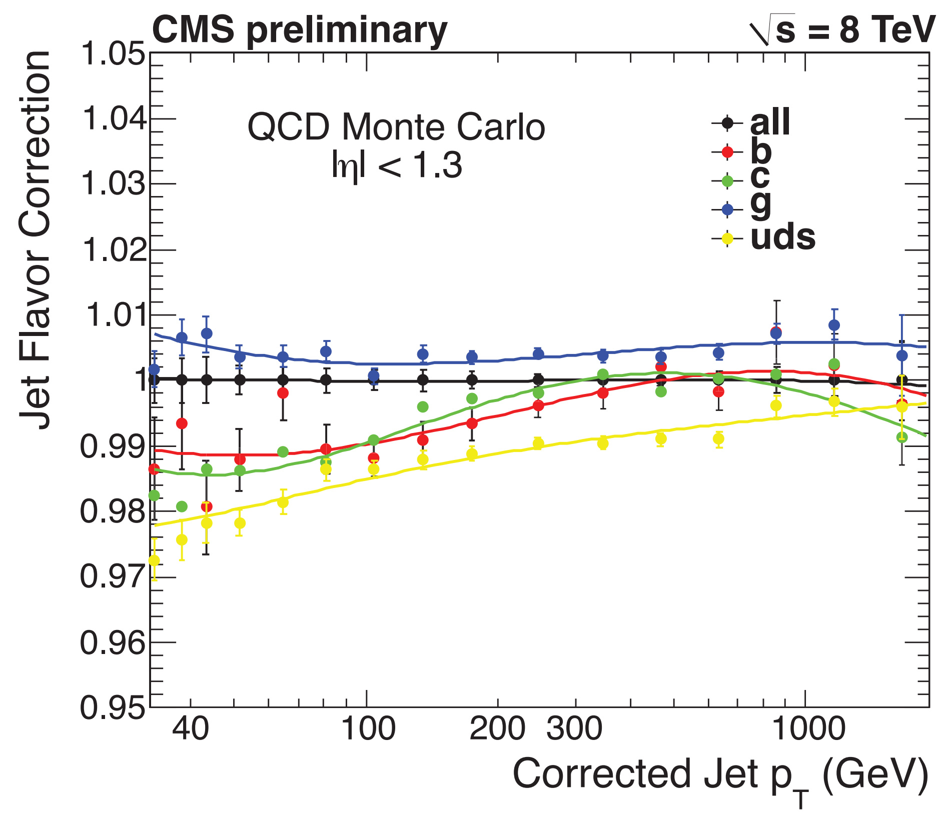




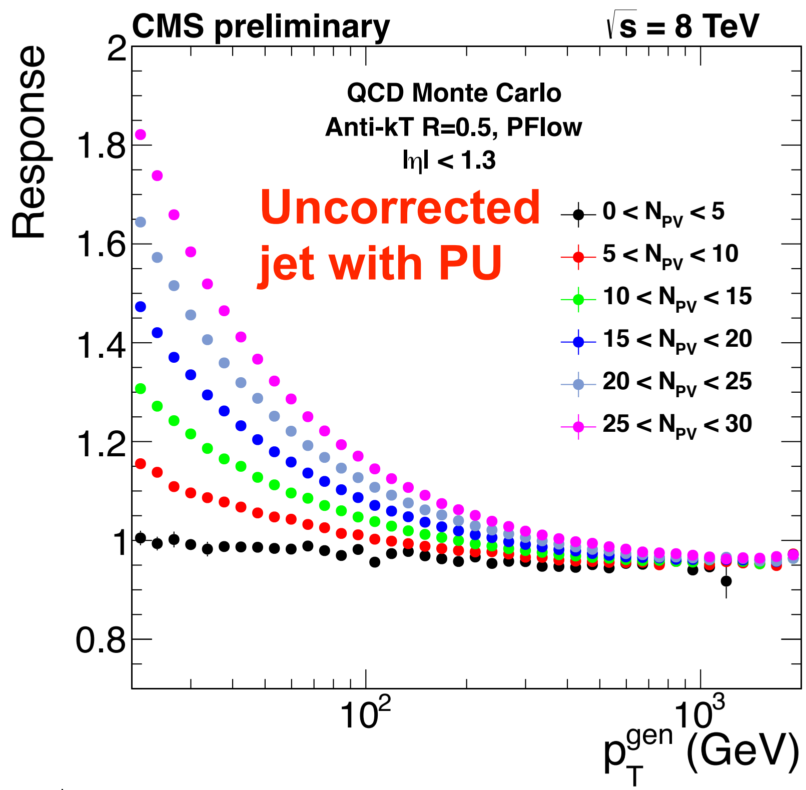
After PU subtraction: start with MC-based correction

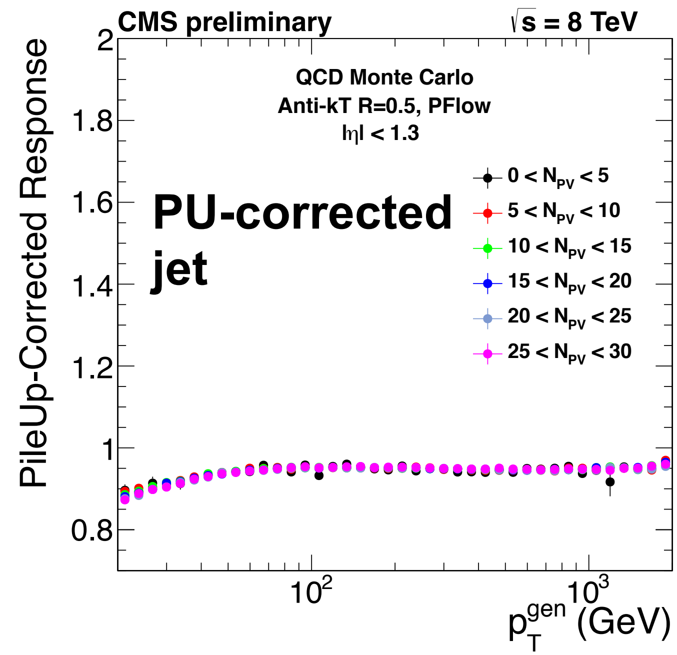
- ✦Eta and p_T corrections derived from QCD MC sample
  - Corrected response closes well in MC.
- ✦Particle flow minimizes flavor response differences
  - Maximum flavor difference within 3% in barrel for p_T > 30 GeV.

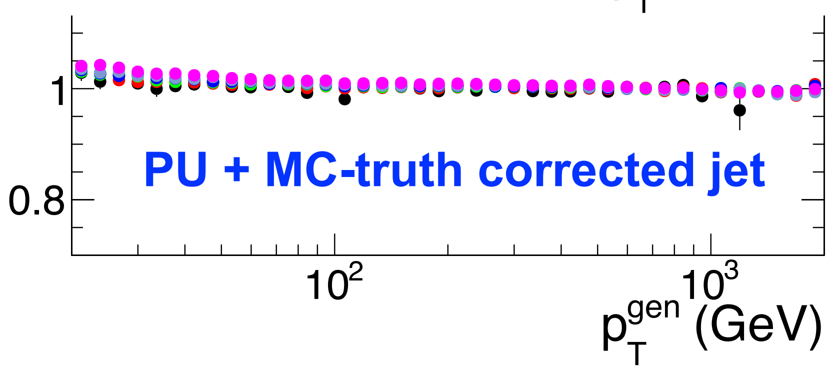




Combined pileup and MC truth effects





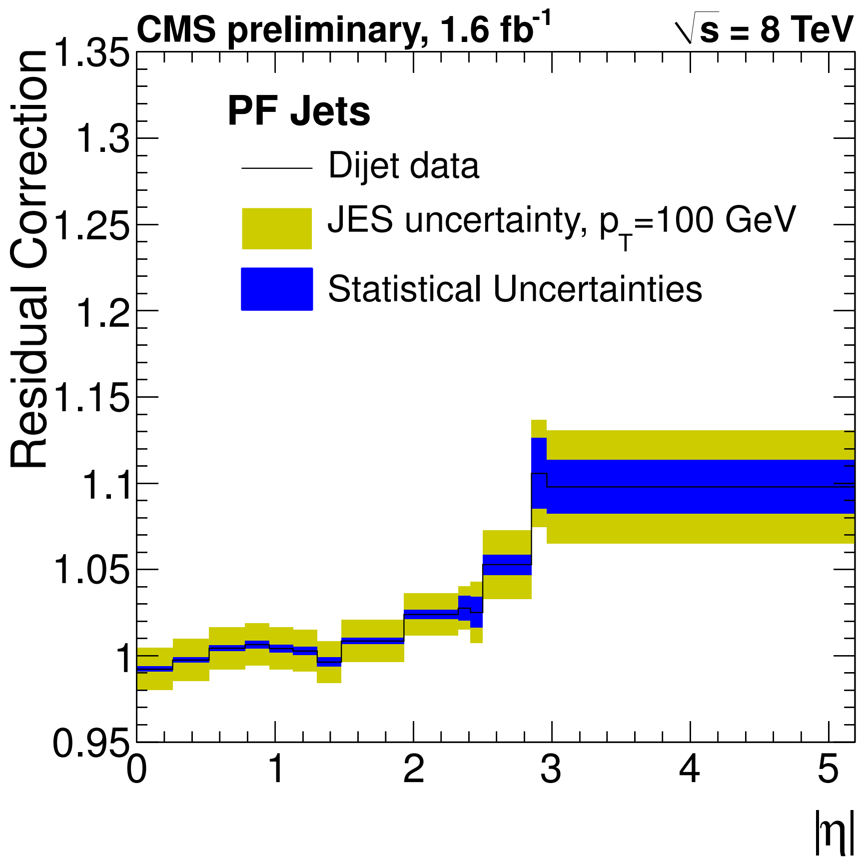


- ✦PU corrections remove response dependance vs N_{PV}.
- ✦MC truth correction brings the closure back to one.

Residual correction in data: η & p_T dependence

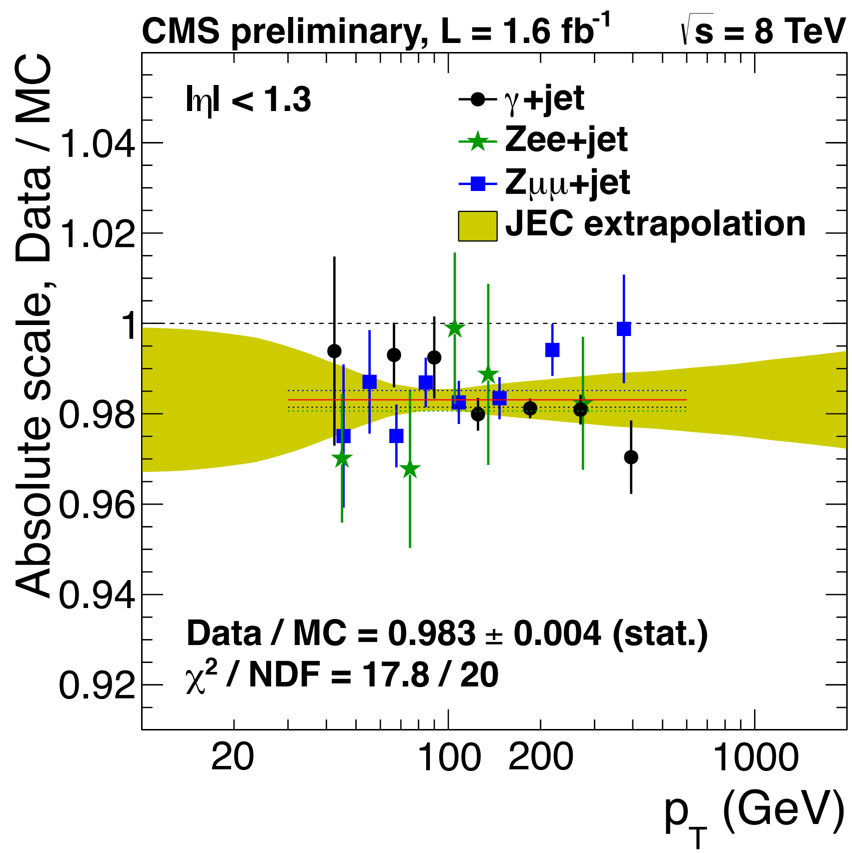
η dependence

- Using dijet events
- < 2.5% for jets in |η|<2.4
- HF modeling requires 5−10% correction



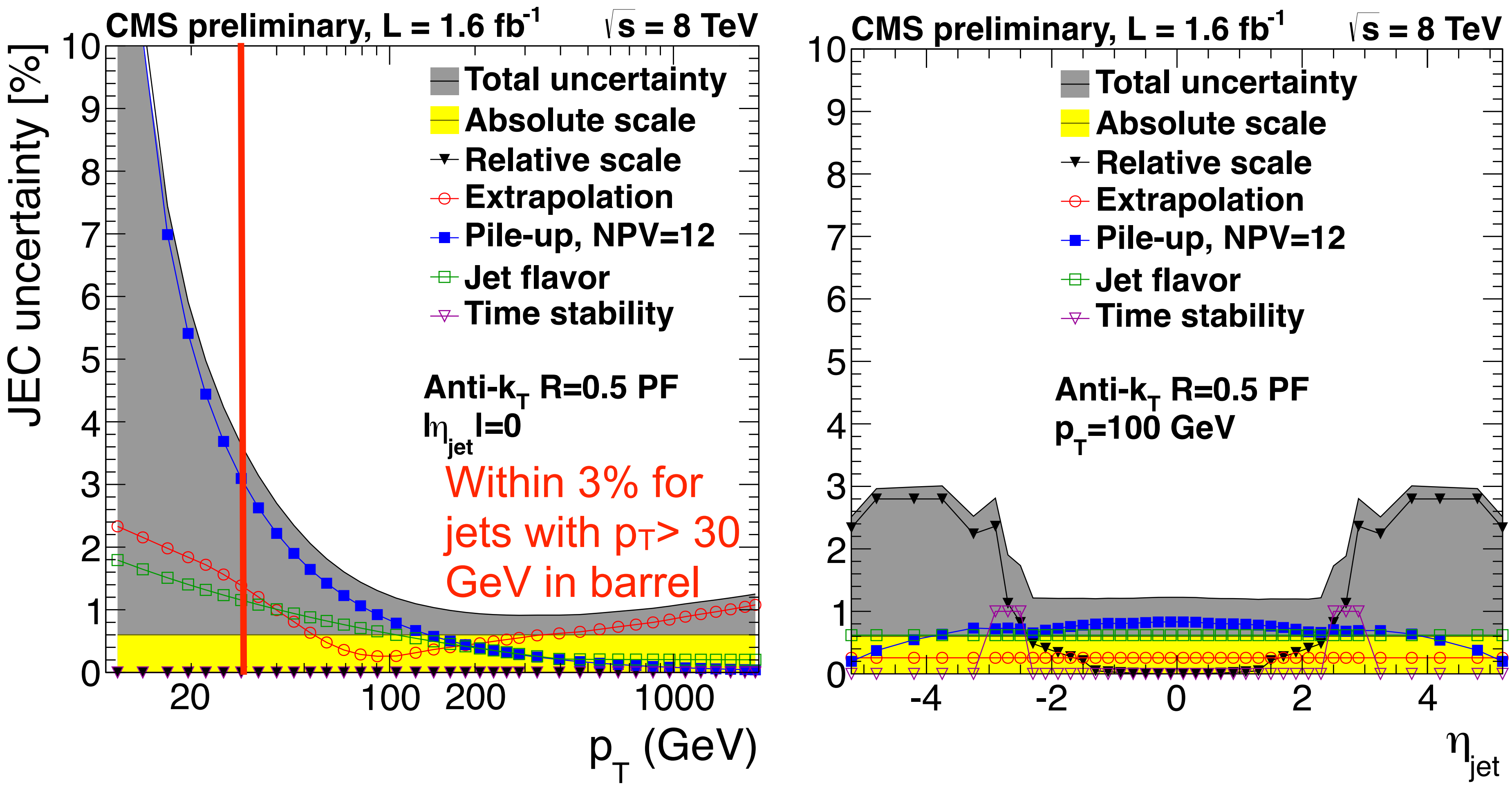
p_T dependence

- Using Zμμ+jet, Zee+jet, γ+jet events
- No significant p_T dependance observed, so a single flat scale factor is used.

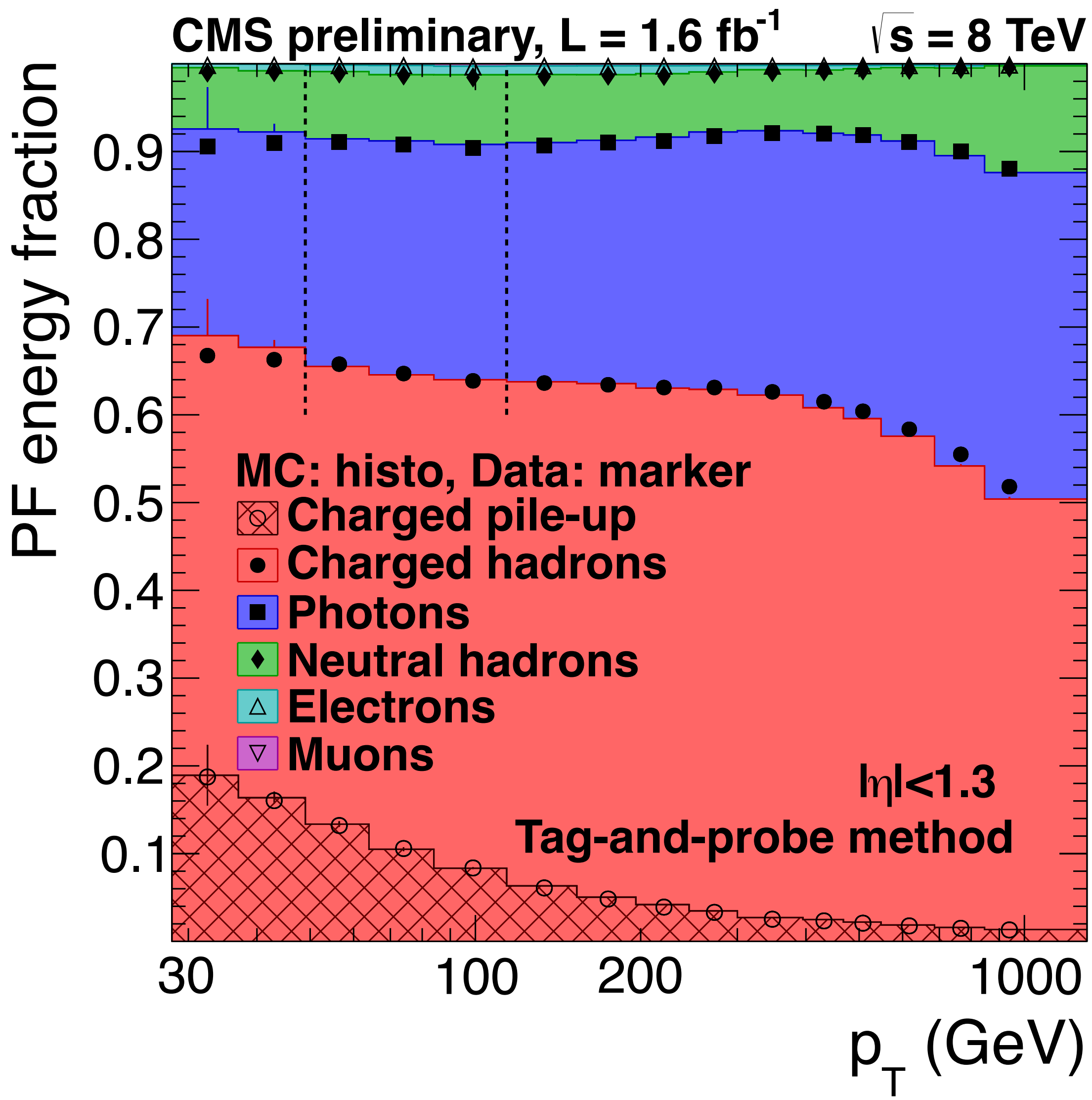


Correction uncertainties

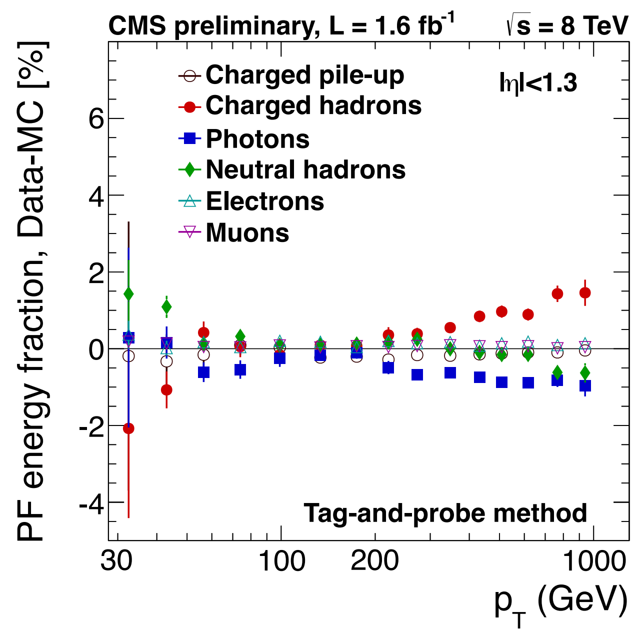
- ✦Uncertainties in 2012 data comparable to 2010, 2011.
  - Pileup uncertainties increasing due to higher average pileup.



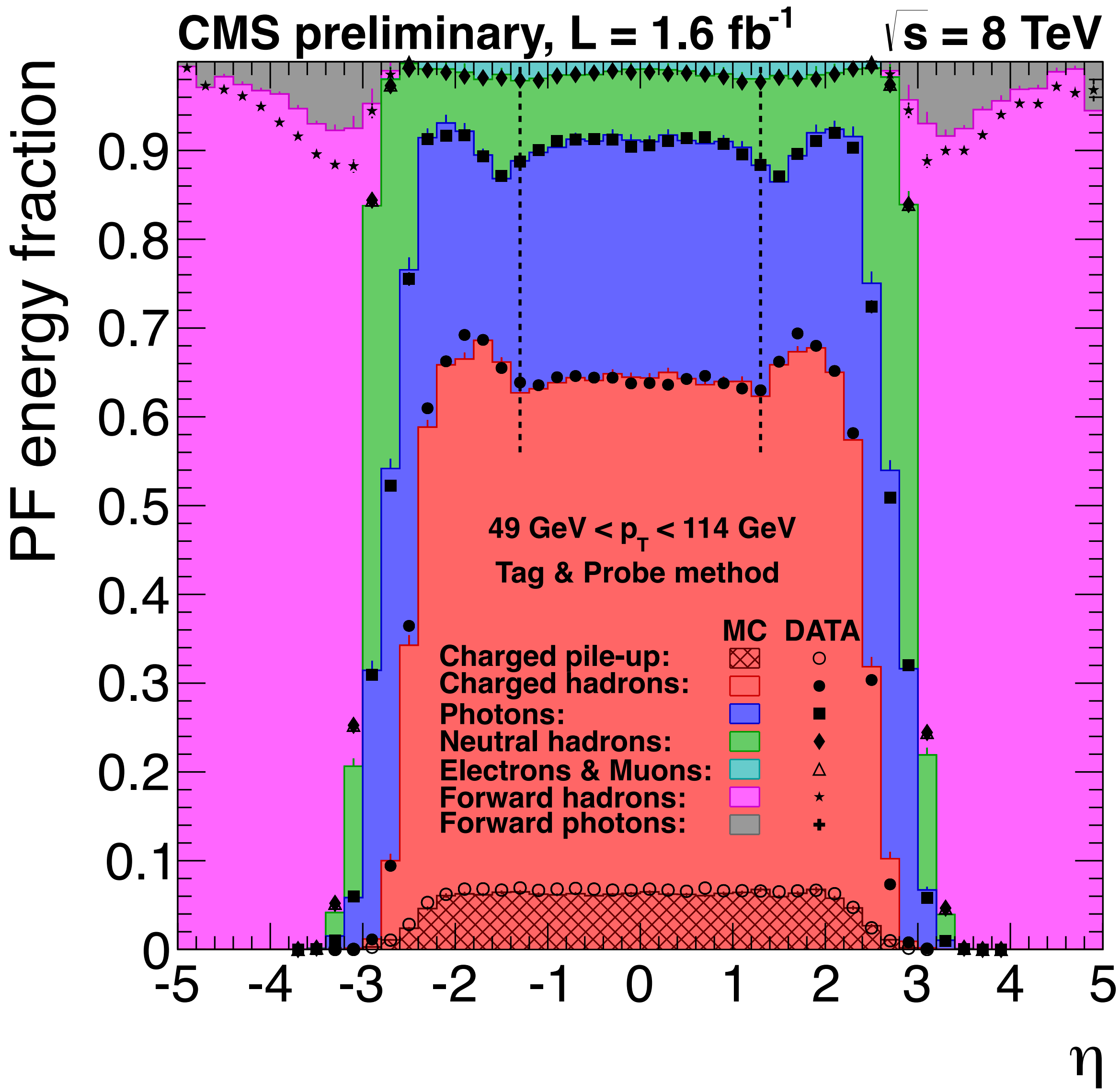
Jet composition vs p_T in barrel



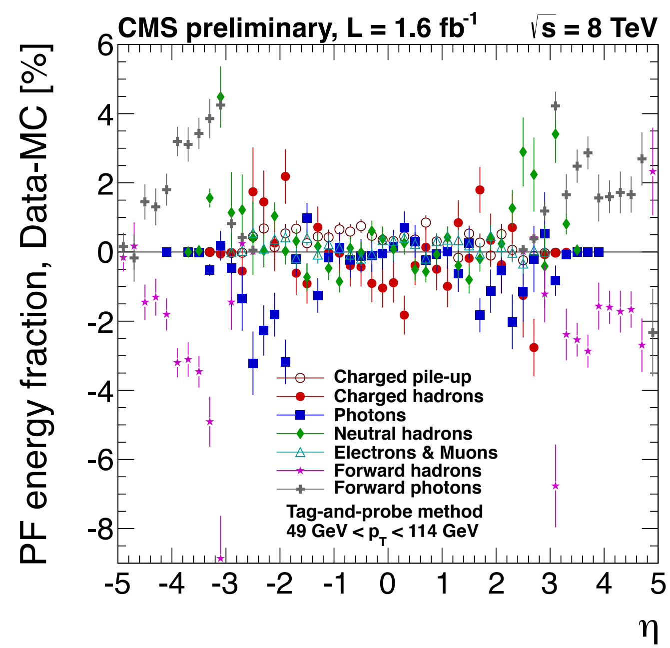
- ✦Jet composition agrees well between Data and MC
  - •consistent with small residual JEC at the 1-2% level.



Jet composition vs η



- ✦Jet composition shows increasing differences in the forward region
  - •consistent with JEC at 2-13% level.



How to apply Jet Corrections in CMSSW

- The "jet correction service" is the software that delivers the correction factor.
- The "correction module" delivers the corrected, re-ordered, jet collection.
- Currently available jet corrections are derived from MC truth.

Example configuration: RecoJets/JetAnalyzers/test/runL2L3JetCorrectionExample_cfg.py

get the text files from the database:

```
JetMETCorrections/Modules/test/JetCorrectionDBReader_cfg.py
payloadName = cms.untracked.string('AK5PF')
```

corrected jet collection:

```
ak5PFJetsL2L3 = cms.EDProducer('PFJetCorrectionProducer',
src = cms.InputTag('ak5PFJets'),
correctors = cms.vstring('ak5PFL2L3') )
```

Correcting jets "on-the-fly":

```
edm::Handle<PFJetCollection> jets; //define input jet collection
Event.getByLabel (jetCollectionName, jets); //get input jet collection
const JetCorrector* corrector = JetCorrector::getJetCorrector (JetCorrectionService,iSetup);
//Get the jet corrector from the event setup
for (PFJetCollection::const_iterator jet = jets->begin(); jet != jets->end(); jet++) {
PFJet correctedJet = *jet; //copy original jet
int index = jet-jets->begin();
edm::RefToBase<reco::Jet> jetRef(edm::Ref<PFJetCollection>(jets,index));
double jec = corrector->correction(*i_jet,jetRef,iEvent,iSetup);
```

Where to get information on Jets?

How can I learn more about jet energy corrections?

https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections

In the above TWiki you can find examples for applying default/optional corrections.

Where can I find details on the derivation of jet corrections?

"Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6**: P11002, 2011 *and the references there in.*

❑ To mention a few:

Anti-kT algorithm:

M. Cacciari, G.P. Salam and G. Soyez, The anti-kt jet clustering algorithm , JHEP 04 (2008) 063 [arXiv:0802.1189].

Fast Jet:

M. Cacciari, G.P. Salam and G. Soyez, FastJet: a software package for jet finding in pp and $e^{+}e^{-}$ collisions webpage , http://www.fastjet.fr/ , (2011).

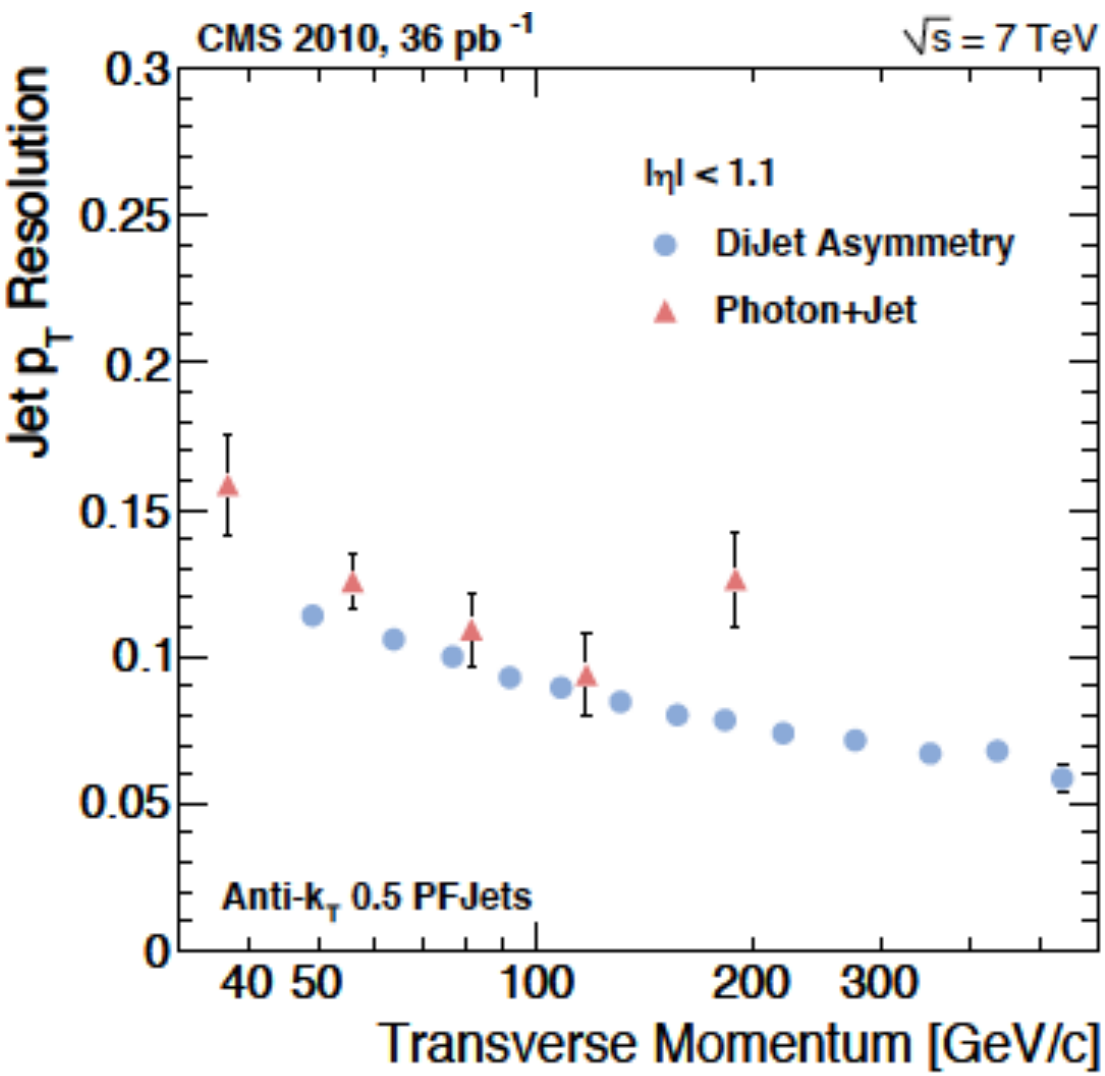
Particle Flow algorithm:

CMS collaboration, Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV, CMS-PAS-PFT-10-002.

All the above documents are available publicly from the CMS Physics Results web page:

https://twiki.cern.ch/twiki/bin/view/CMS/PhysicsResults

Jet resolution



Jet p_T resolution

- Jet resolution improves at higher p_T
- Resolution is slightly better in endcap than barrel
- The core of the measured jet p_T resolution in data is broader than one in simulation

Introduction to Jet Substructure

❑ Jet substructure is currently a very popular tool to analyze jets that are produced from heavy objects such as top quarks, Higgs bosons, W/Z bosons, or any beyond-Standard-Model hadronically-decaying object.

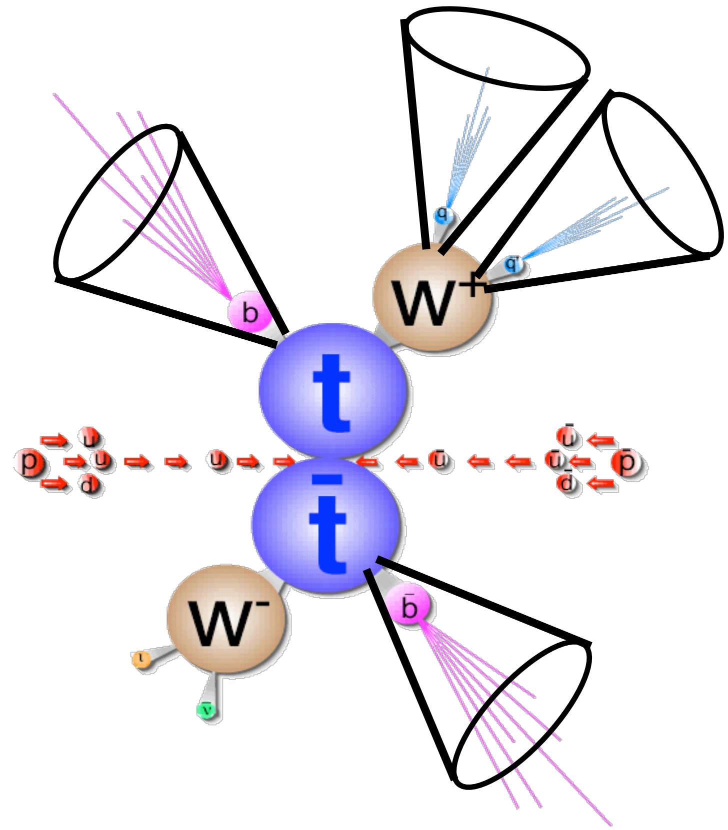
❑ An excellent overview of jet substructure techniques can be found here: Boosted objects: a probe of beyond the Standard Model physics by Abdesselam et al.

❑ Jet pruning algorithm:

Designed to find heavy objects within a jet, and in the process, removes soft and wide-angle clusters from the clustering sequence.

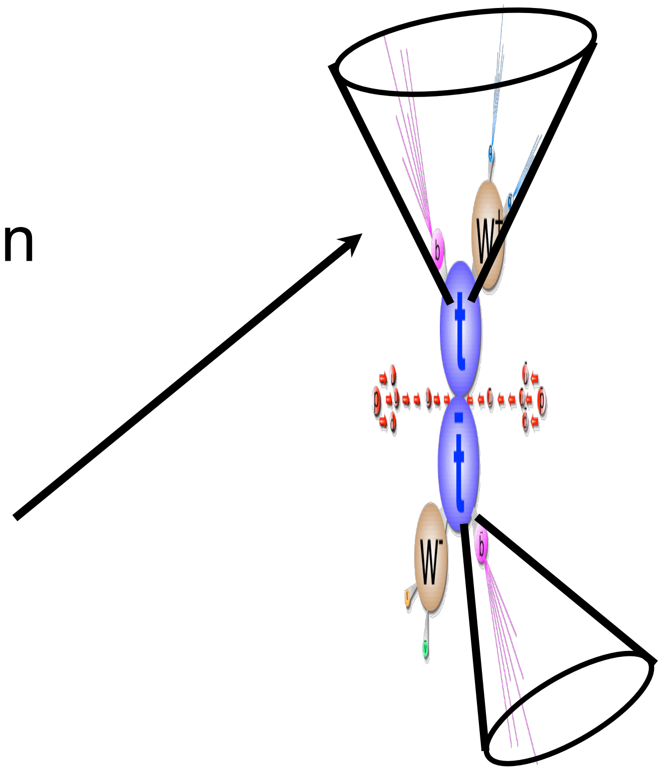
Jet Substructure: Motivation

- Problem! Traditional techniques start to lose sensitivity (in part) due to jet merging at higher masses!
- Cannot rely on methods to assign partons to jets anymore
- Have to consider cases where partons merge into a single jet



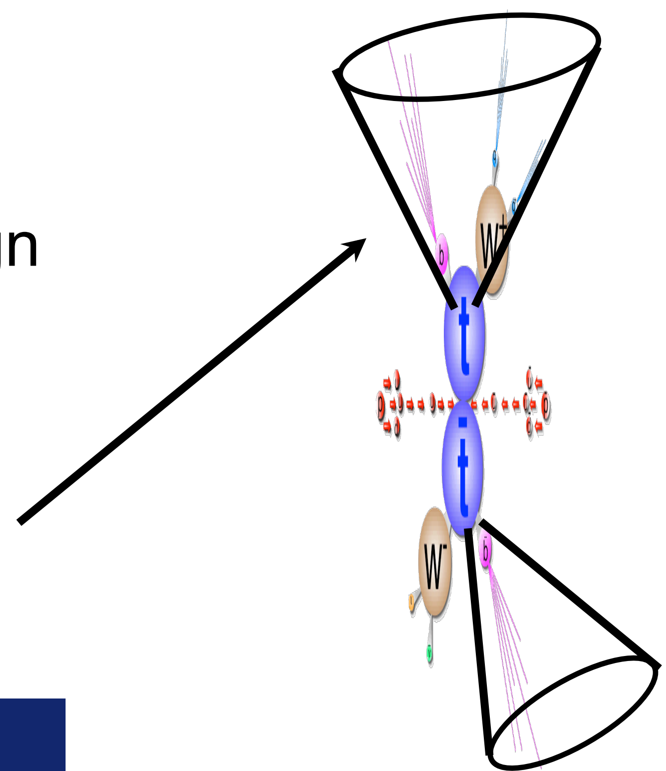
Jet Substructure: Motivation

- Problem! Traditional techniques start to lose sensitivity (in part) due to jet merging at higher masses!
- Cannot rely on methods to assign partons to jets anymore
- Have to consider cases where partons merge into a single jet



Jet Substructure: Motivation

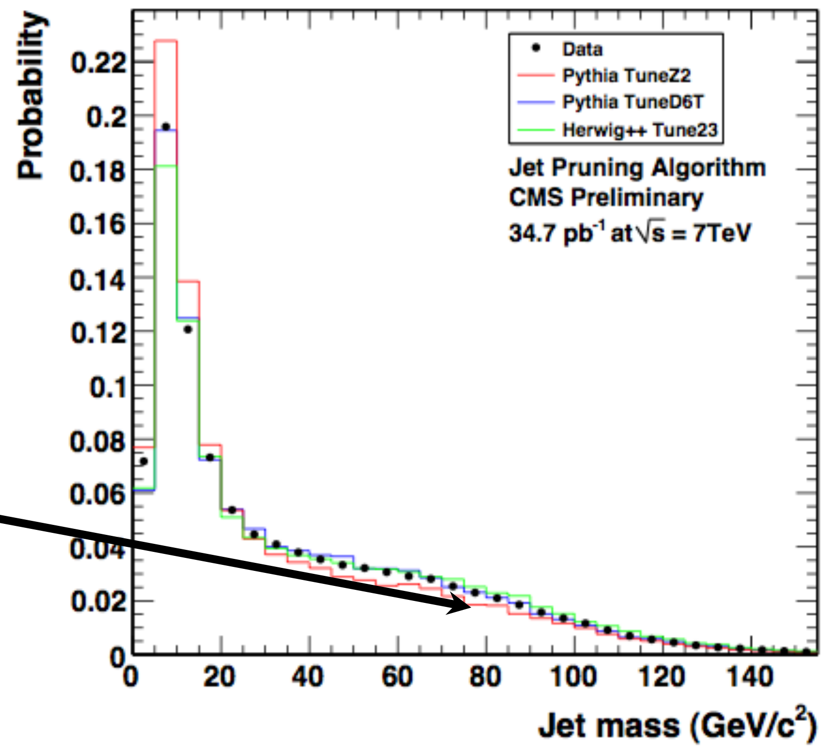
- Problem! Traditional techniques start to lose sensitivity (in part) due to jet merging at higher masses!
- Cannot rely on methods to assign partons to jets anymore
- Have to consider cases where partons merge into a single jet



To understand how to break up a jet, first we need to understand what it is!

Example: Jet Mass: Data vs MC

- Very good agreement between MC and data jet mass!
- Strong dependence on Parton Shower model and tune, but all very good in "signal region" for interesting particles (high jet mass)



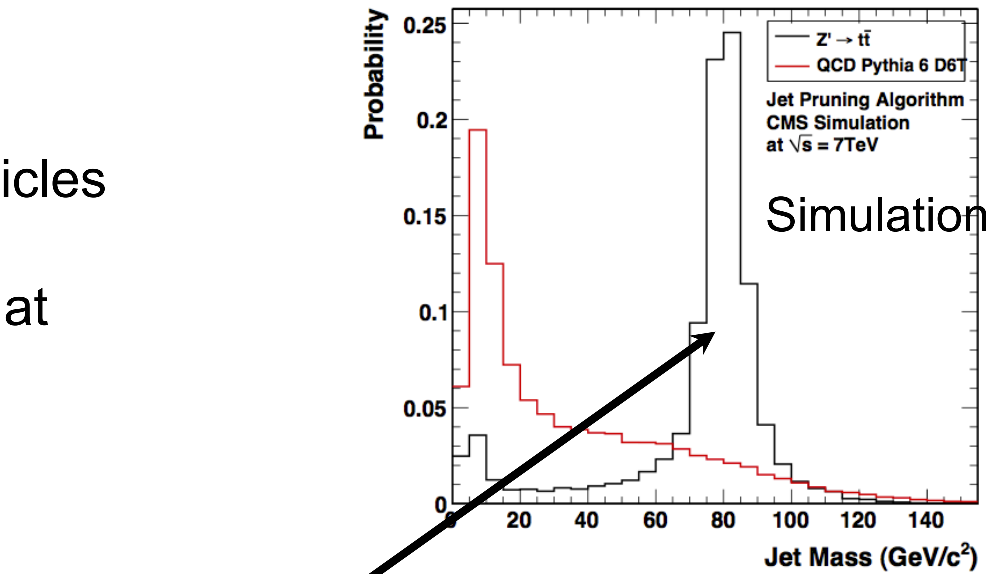
Example: W-tagging with Pruning

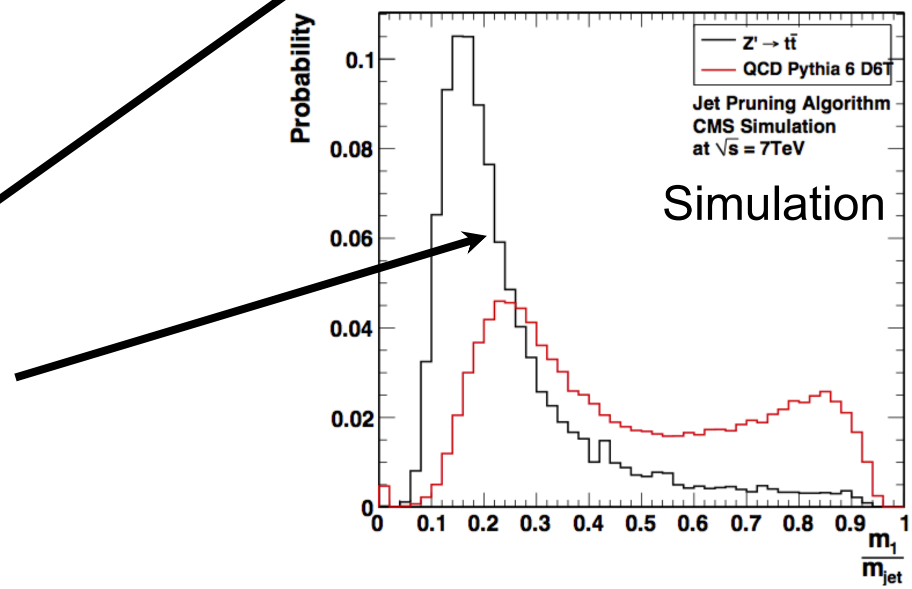
- Ellis et al. (arXiv:0903.5081)
- Improves mass resolution by removing soft, large angle particles from the jet
- Recluster each jet, requiring that each recombination satisfy the following:

$$\frac{\min(p_{T1}, p_{T2})}{p_{Tp}} > 0.1$$

$$\Delta R_{12} < 0.5 \times \frac{m_{\rm jet}}{p_{T}}$$

- For W tagging, require:
  - Jet mass in 60-100 GeV/c^2
  - Mass drop (mu) < 0.4 $\mu = \frac{m_{j1}}{m_j}$





Conway, Erbacher, Dolen, Hu, Maksimovic, Rappoccio, Sierra-Vasquez

Conclusion

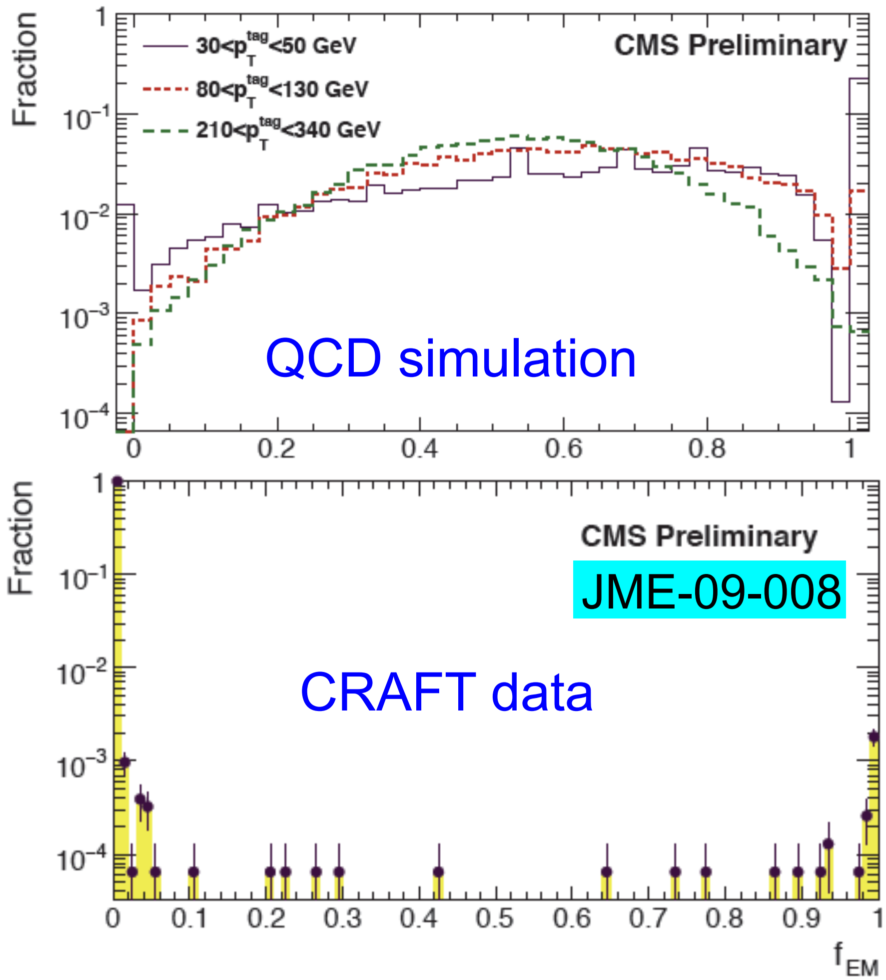
- ✓ **Jet cross section at LHC is several orders of magnitude higher than any other process**
  - We need to understand their performance as much as possible
- ✓ **The baseline jet object at CMS is the Particle Flow Jets**
  - Jets need to be calibrated and cleaned before they can be used for physics analysis
- ✓ CMS uses data driven methods to correct jet energies to the particle level
- ✓ With more data we are probing into new TeV regime
- ✓ The Jet substructure is coming up as a very popular tool to analyze jets that are produced from heavy objects
- ✓ Please find us during the school or email us in case you come up with a question later

  – OR –
- ✓ Write to the CMS Hypernews forum for Jets

Back up

Jet Id

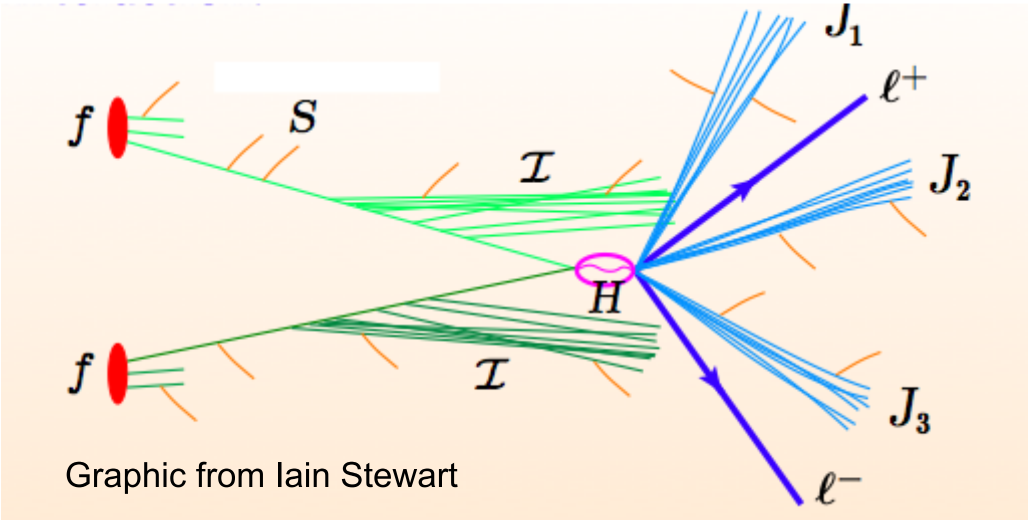
- Studies, based on CRAFT data and MC simulation, indicate that the electromagnetic energy fraction (EMF or f_{EM}) is a powerful criterion to reject fake jets.
- Other quantities are being examined too:
  - #CaloTowers containing 90% of the jet energy
  - RMS of the E_T weighted ϕ distribution
  - fraction of energy from the hottest HPD/ RBX
- Data-driven method methods to measure the jet id efficiency are being developed
- Software implementation of the jet ID is being examined.



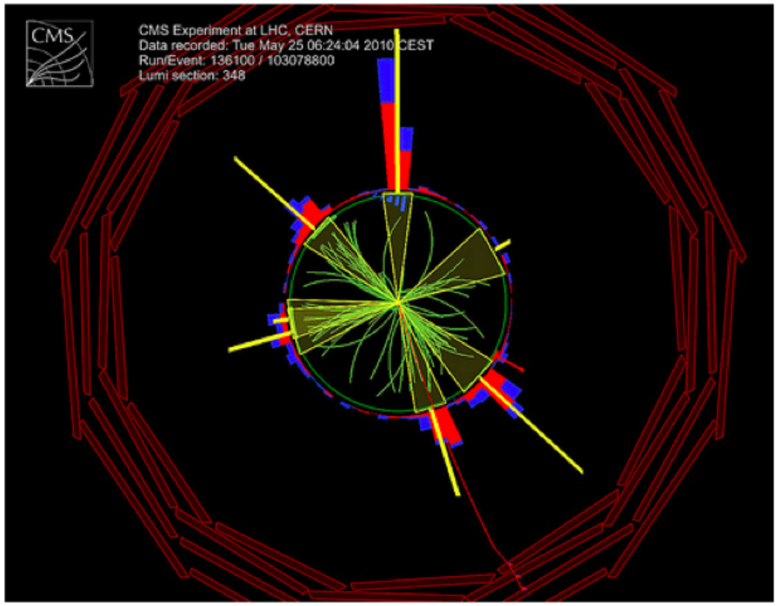
- **Jet Id is defined for *uncorrected* jets only. Never apply jet Id on corrected jets. This means that in your analysis you should apply jet Id first and then apply jet energy correction on those jets that pass jet Id.**

Interlude: Particle Jets

- Nice overview: Ellis et al(arXiv:hep-ph/0412013v1)
- Simple picture of 1->2 decay is complicated in QCD!
- Spray of collimated particles
- A lot of complications!
  - Underlying event
  - Pileup
  - Initial + final state radiation
- The LHC is a VERY jetty place!
  - Messy
  - Complicated to disentangle
  - A huge amount of fun!



Graphic from Iain Stewart



Sequential clustering algorithms

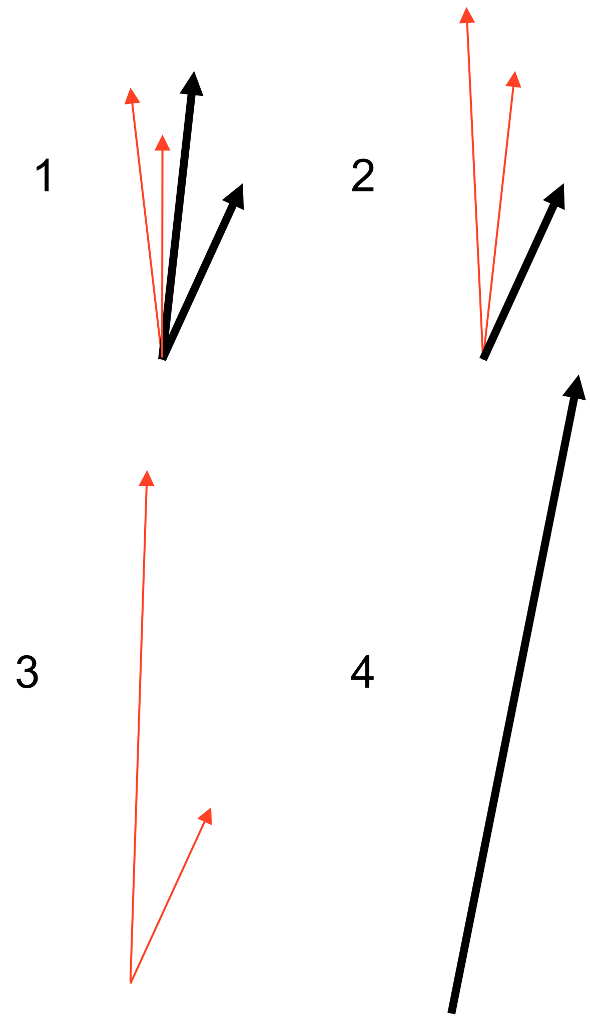
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

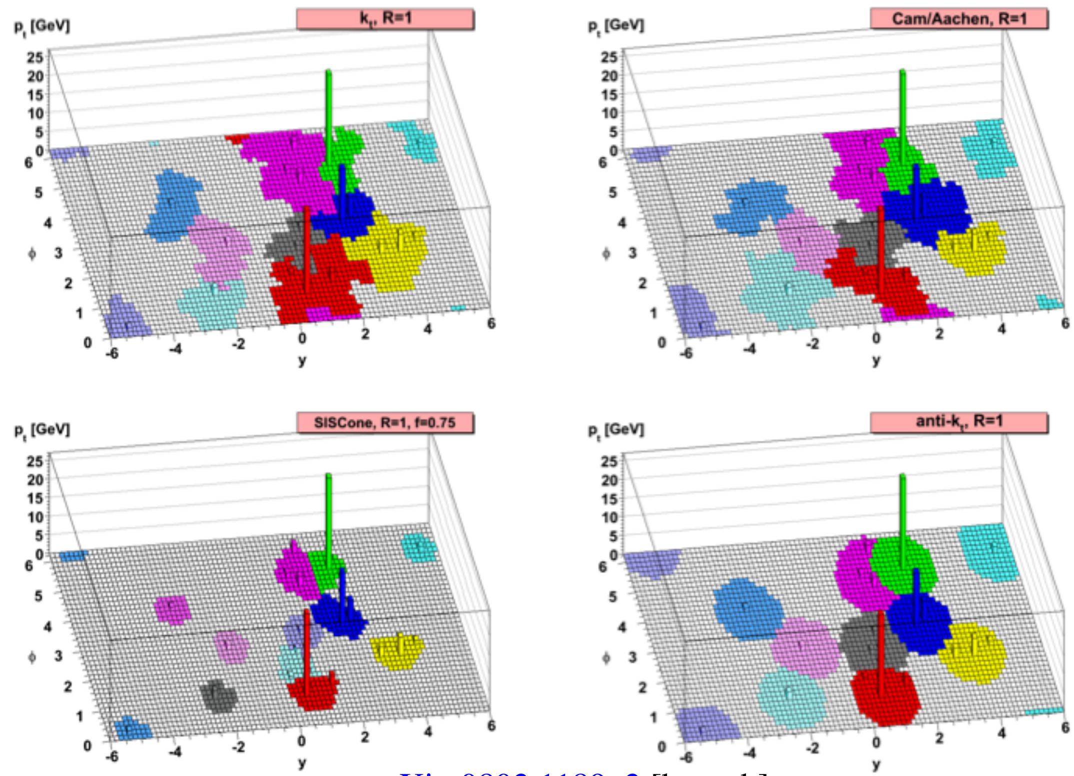
$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - – If min is a dij, merge and iterate
  - – If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential clustering algorithms

- Different types
  - N = 2: “kT”
    - “Irregular” jets, but good for low pt
  - N = 0 : “Cambridge-Aachen” (CA)
    - Also irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Jets in Physics Analysis Toolkit (PAT)

- The pat::Jet is the basic jet object
- Stores internally the jet correction factors that can be retrieved:

jetCorrFactor(string &step, string &flavor="")

- In the default PAT configuration, some jet corrections - relative and absolute - applied (based on MC). Can do correction on the fly.
- Stores jet flavor (from MC) and jet Id vars.
- Can do MC matching with GenJets/partons.
- Can use any kind jet collection: pFlow, JPT.

PAT: jetProducer_cff

```
import FWCore.ParameterSet.Config as cms

# prepare reco information
from PhysicsTools.PatAlgos.recoLayer0.jetTracksCharge_cff import *
from PhysicsTools.PatAlgos.recoLayer0.jetMETCorrections_cff import *

# add PAT specifics
from PhysicsTools.PatAlgos.mcMatchLayer0.jetFlavourId_cff import *
from PhysicsTools.PatAlgos.mcMatchLayer0.jetMatch_cfi import *

# produce object
from PhysicsTools.PatAlgos.producersLayer1.jetProducer_cfi import *

makeAllLayer1Jets = cms.Sequence(
    # reco pre-production
    patJetCharge *
    patJetCorrections *
    # pat specifics
    jetPartonMatch *
    jetGenJetMatch *
    jetFlavourId *
    # object production
    allLayer1Jets
    )
```

```
patJets = cms.EDProducer("PATJetProducer",
    # input
    jetSource = cms.InputTag("ak5CaloJets"),
```

Tools to switch the input Collection

```
    # jet energy corrections
    addJetCorrFactors     = cms.bool(True),
    jetCorrFactorsSource = cms.VInputTag(cms.InputTag("patJetCorrFactors") )
```

JEC are available from within the patJet

```
  # jet energy corrections
  addJetCorrFactors     = cms.bool(True),
  jetCorrFactorsSource = cms.VInputTag(cms.InputTag("patJetCorrFactors") )
```

Jet cleaning and matching

Why do I need to clean jets ?

- Any object which deposits energy in HCAL can get reconstructed as a jet !
- This implies that jet collections also contain electrons, muons, photons, ...
- Before doing analysis with jets, these objects need to be removed from the jet collection.

Jet matching

- For Monte Carlo based analysis one may be interested to know what fraction of the reconstructed jets is matched to generator level quantities: GenJets or partons.
- In data-driven analyses of certain types of events (e.g., dijet, Z+jet, photon +jet) one may want to use p_T balance between a jet and a reference object by doing back-to-back matching in ϕ.

We have standard tools in CMSSW to perform these recurring tasks and to compute efficiency for each step ➜ see next slide