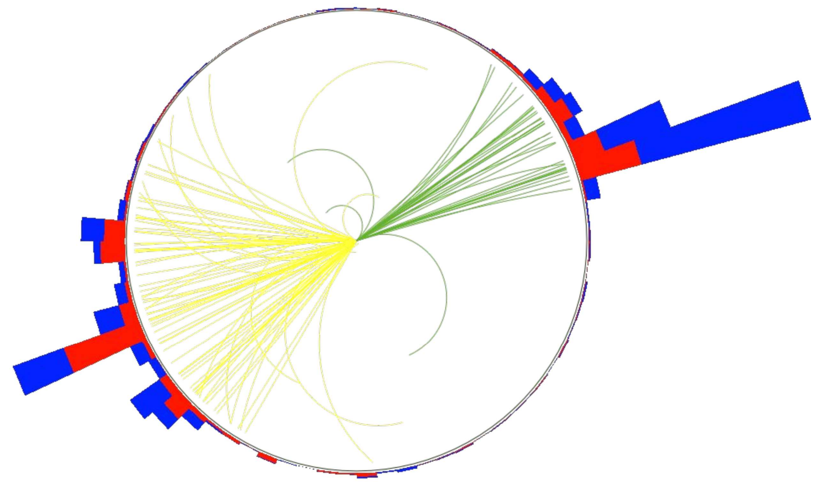


Jets "Short Exercise"
CMSvDAS June 2023

Ashley Parker, Andris Potrebko, Mikael Myllymaki
using slides from Cristina Ana Mantilla Suarez, Christine McLean, Henning Kirschenmann, Raffaele Gerosa, John Hakala

using previous slides from Kevin Pedro, Cristina Ana Mantilla Suarez, Chin Lung Tan, Caleb James Smith, Juska Pekkanen, Alexx Perloff, David Yu, Alejandro Gomez Espinosa

using previous slides from Alexx Perloff, Justin Pilot, Tongguang Cheng

using previous slides from Jim Dolen, Salvatore Rappoccio, et. al.

https://indico.cern.ch/event/1257234/

Exercise Plans

- Introductions
- Setup computing environment and start running scripts
  - Start here: JMEDAS github repository
    - Login to SWAN, checkout code, init (and copy) proxy from lxplus
- 4 blocks in total, 7 remote facilitators:
  - Section 1: Intro, Basics, Jet Types and Algorithms
  - Section 2: Pileup
  - Section 3: JEC and JER
  - Section 4: Jet Substructure

Ashley Parker

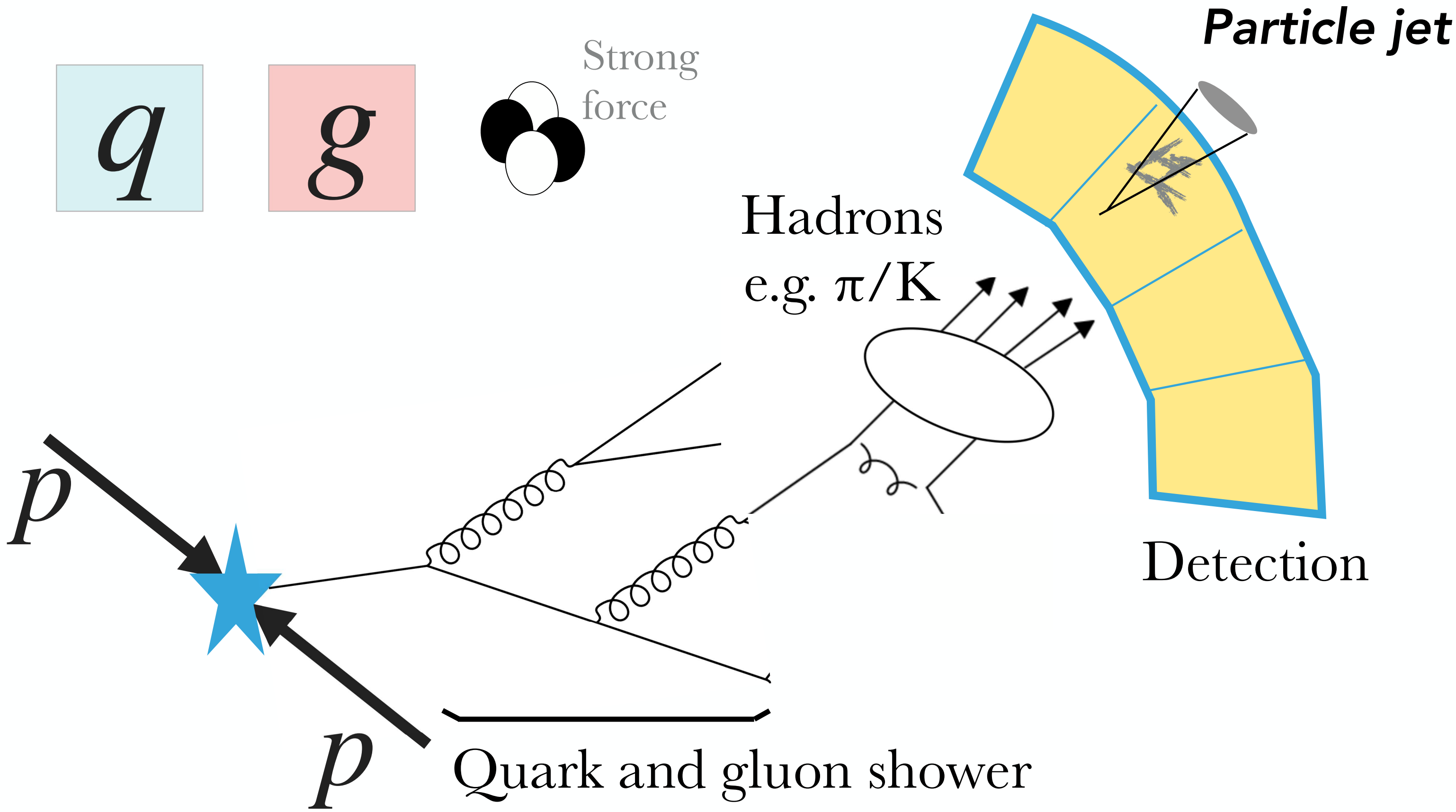
Andris Potrebko

Mikael Myllymaki

- A fuller version of the slides with the presentation recordings can be found at the twiki of the 2022 CMSDAS twiki page.

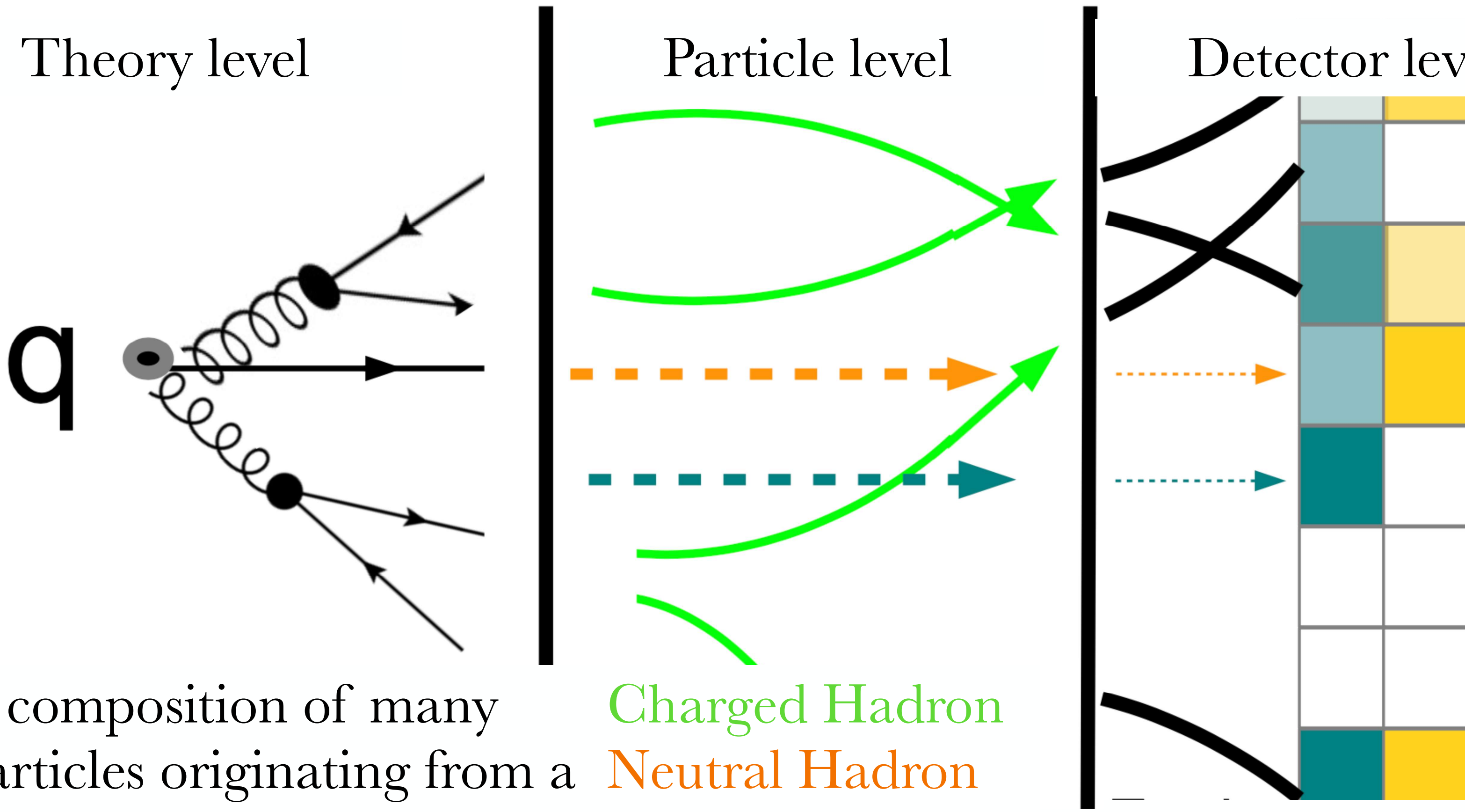
Section 1: Jet algorithms

Jets are signatures of quarks and gluons



We can define jets at different levels:

Theory level | Particle level | Detector level

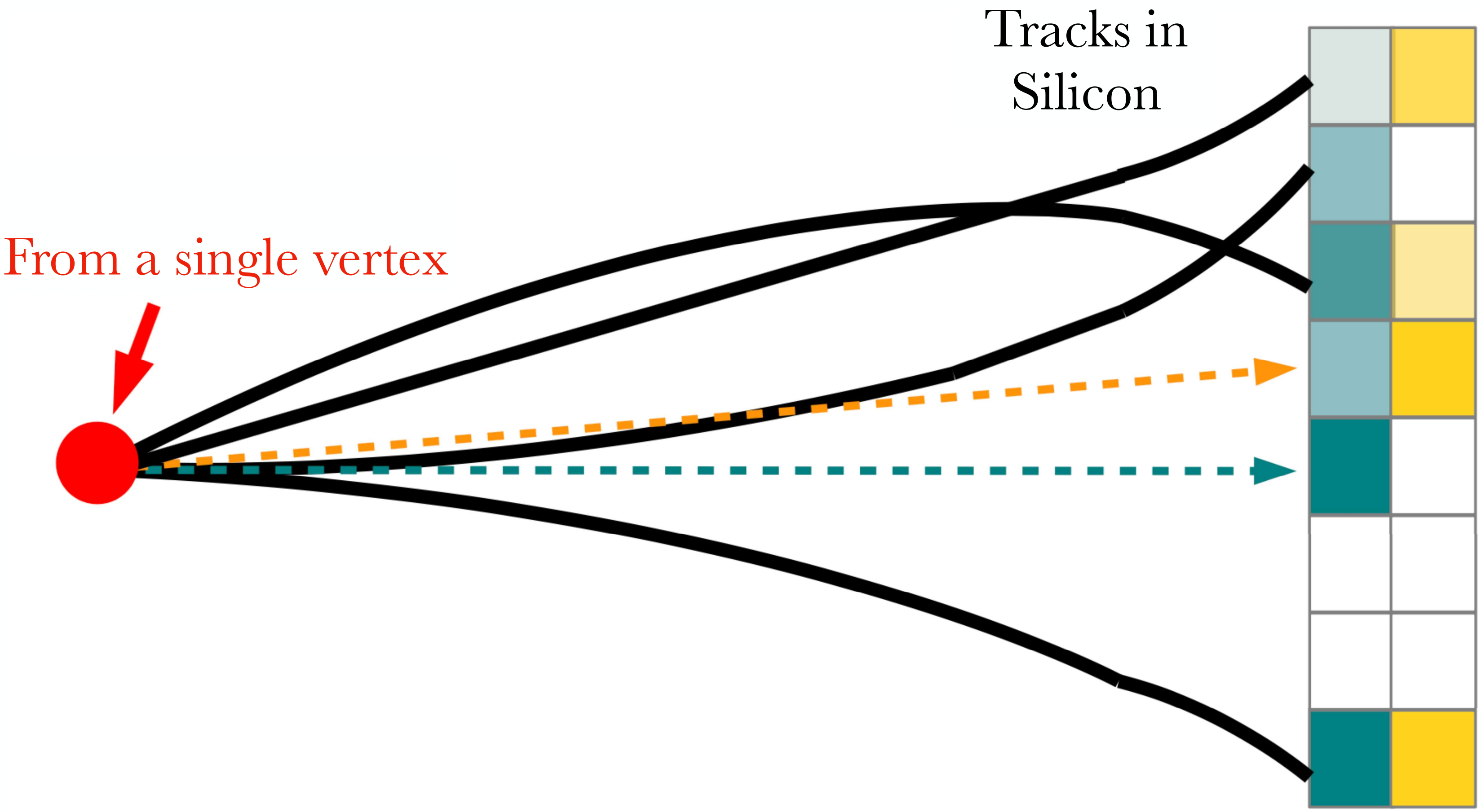


A composition of many particles originating from a quark or gluon

Charged Hadron
Neutral Hadron
Photon

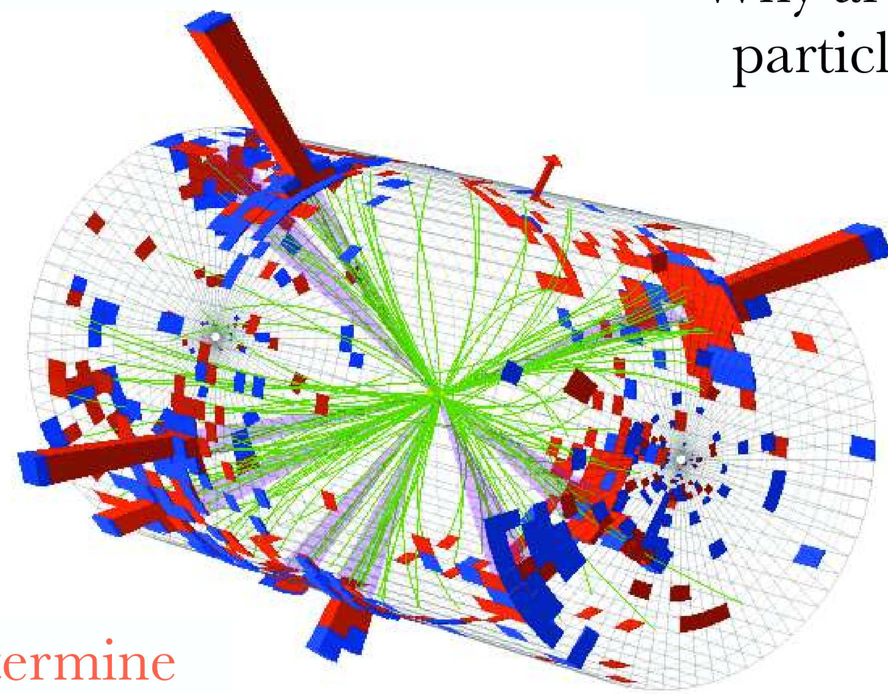
Tracker/ECAL/HCAL

A jet inside the detector



A jet inside the detector

Why are there so many particles inside a jet?

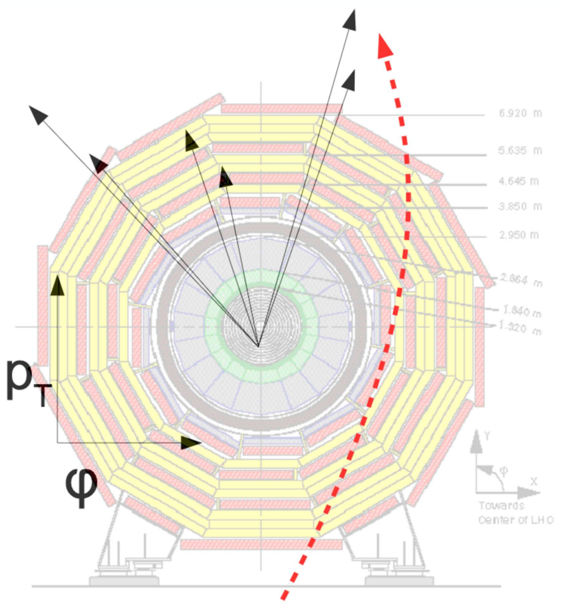


CMS Experiment at LHC, CERN
Data recorded: Mon May 23 21:46:26 2011 EDT
Run/Event: 165567 / 347495624
Lumi section: 280
Orbit/Crossing: 73255853 / 3161

How do you determine which particles should be included?

Jet reconstruction

CMS Transverse view



Highest p_T

- Start from a list of particles (each particle is known as constituent)
- Need a jet clustering algorithm:
  - Iteratively find two particles in the event that are closest in some distance measure

$$\Delta R \min(p_T{}^1, p_T{}^2)^\alpha$$

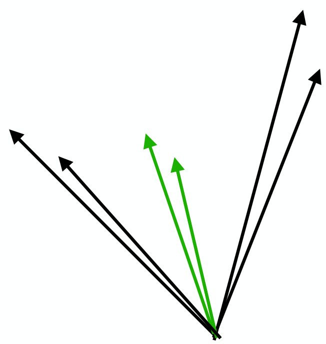
α = 1: kT algorithm

α = 0: Cambridge-Aachen algorithm

α = -1: Anti-kT algorithm

  - New soft particle: $p_T \rightarrow 0$ and new collinear particle $\Delta R \rightarrow 0$

Iterating over two



Start soft (small)

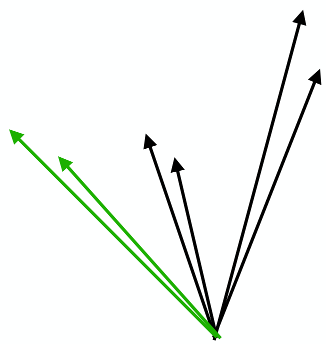
$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha = 1$: kT algorithm

$\alpha = 0$: Cambridge-Aachen algorithm

$\alpha = -1$: Anti-kT algorithm

Iterating over two



Start close

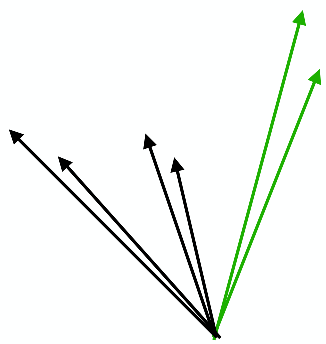
$\Delta R \min(p_T{}^1, p_T{}^2)^\alpha$

$\alpha = 1$: kT algorithm

$\alpha = 0$: Cambridge-Aachen algorithm

$\alpha = -1$: Anti-kT algorithm

Iterating over two



Start hard (big)

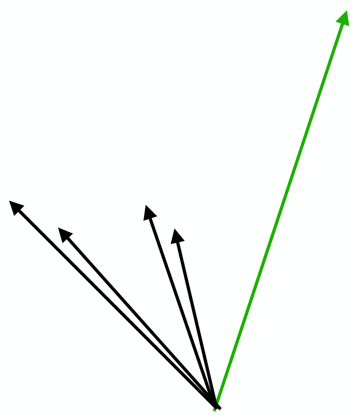
$$\Delta R \, \min(p_T^{\,1}, p_T^{\,2})^{\alpha}$$

α = 1: kT algorithm

α = 0: Cambridge-Aachen algorithm

α = -1: Anti-kT algorithm

Now merge into a particle



Merge into a particle

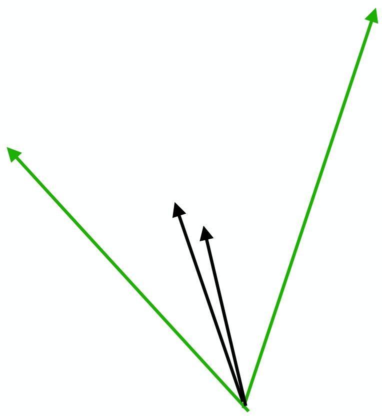
$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha = 1$: kT algorithm

$\alpha = 0$: Cambridge-Aachen algorithm

$\alpha = -1$: Anti-kT algorithm

Merge next set



Merge next

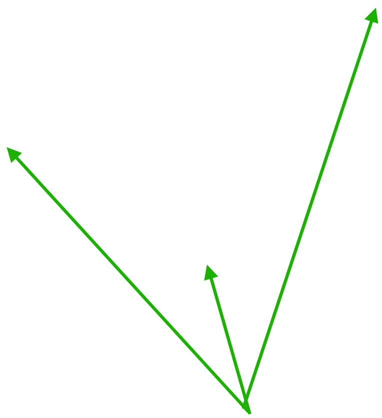
$$\Delta R \min(p_T{}^1, p_T{}^2)^\alpha$$

α = 1: kT algorithm

α = 0: Cambridge-Aachen algorithm

α = -1: Anti-kT algorithm

Merge next set



And next

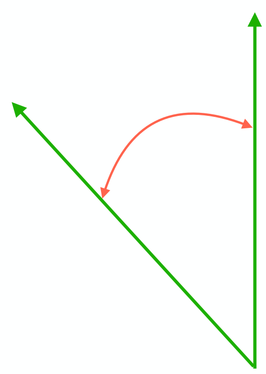
$$\Delta R \min(p_T^{1}, p_T^{2})^{\alpha}$$

$\alpha = 1$: kT algorithm

$\alpha = 0$: Cambridge-Aachen algorithm

$\alpha = -1$: Anti-kT algorithm

Merge next set



If distance > X
(X=0.4,0.8 - stop)

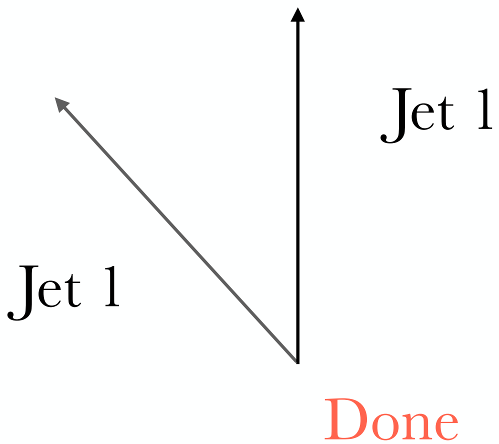
$$\Delta R \min(p_T^{\,1}, p_T^{\,2})^{\alpha}$$

$\alpha = 1$: kT algorithm

$\alpha = 0$: Cambridge-Aachen algorithm

$\alpha = -1$: Anti-kT algorithm

Done



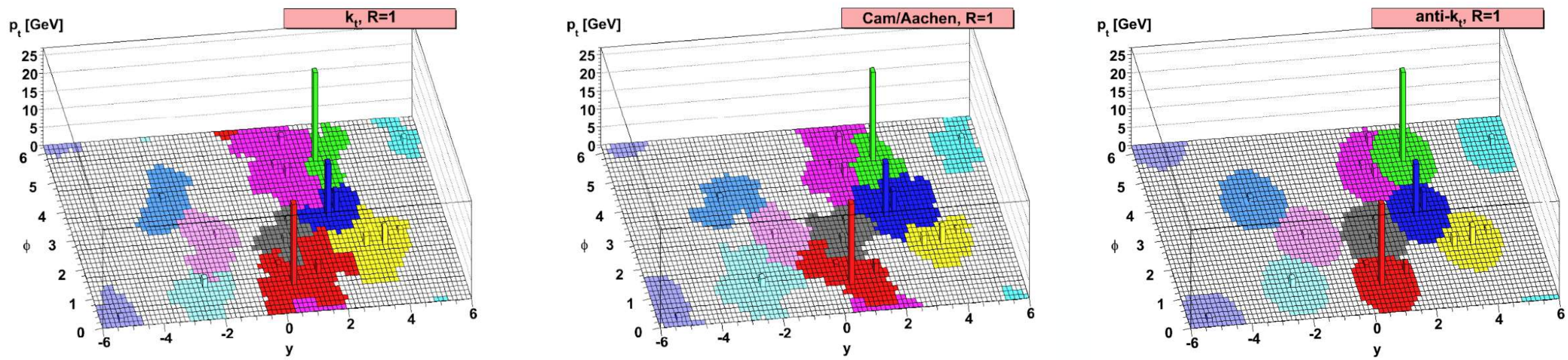
$$\Delta R \, \min(p_T{}^1, p_T{}^2)^\alpha$$

α = 1: kT algorithm

α = 0: Cambridge-Aachen algorithm

α = -1: Anti-kT algorithm

Comparison between jet algorithms



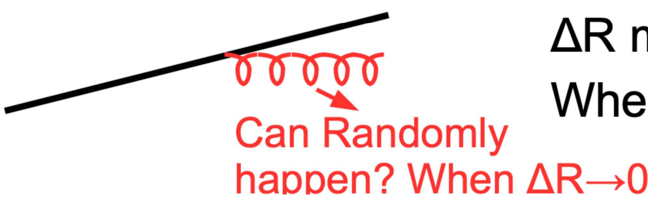
- How to do clustering by yourself?
  - Use FASTJET: http://fastjet.fr/

Infrared and collinear safety

- To calculate anything with a jet we need to observe :
  - Infrared safety : invariance with random particle w/ $E \to 0$

$\Delta R \min(p_T^{1}, p_T^{2})^{\alpha} \to 0$ ($p_T \to 0$)

  - Collinear safety : invariance with random split $\Delta R \to 0$



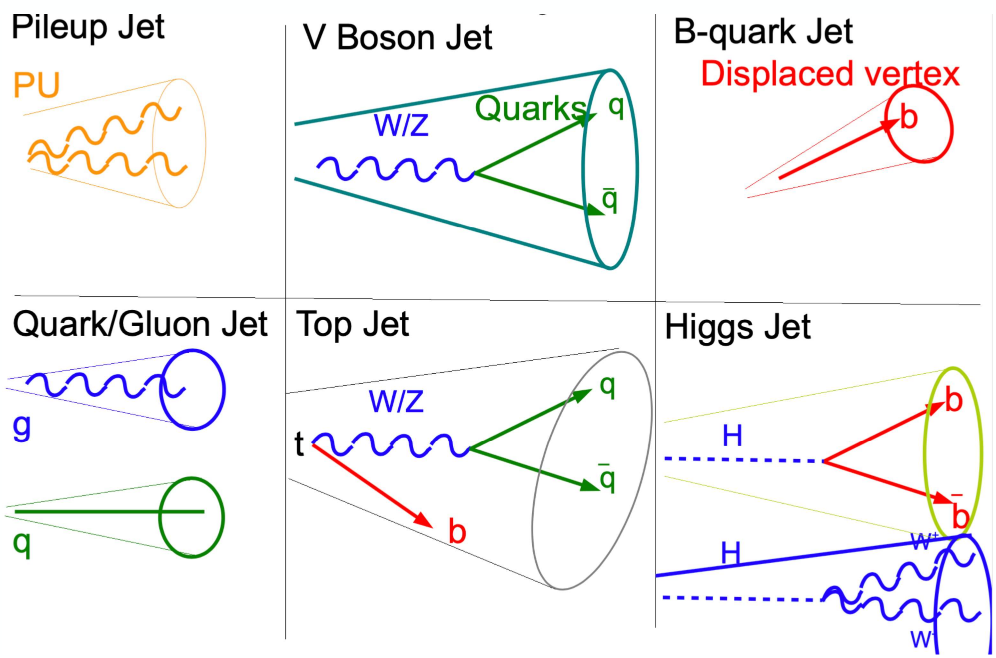
$\Delta R \min(p_T^{1}, p_T^{2})^{\alpha} \to 0$

When $\Delta R \to 0$

- The set of hard jets should be unchanged by soft emission and collinear splitting

Different types of jets

- The size of the jet cone allows us to focus on a different object:



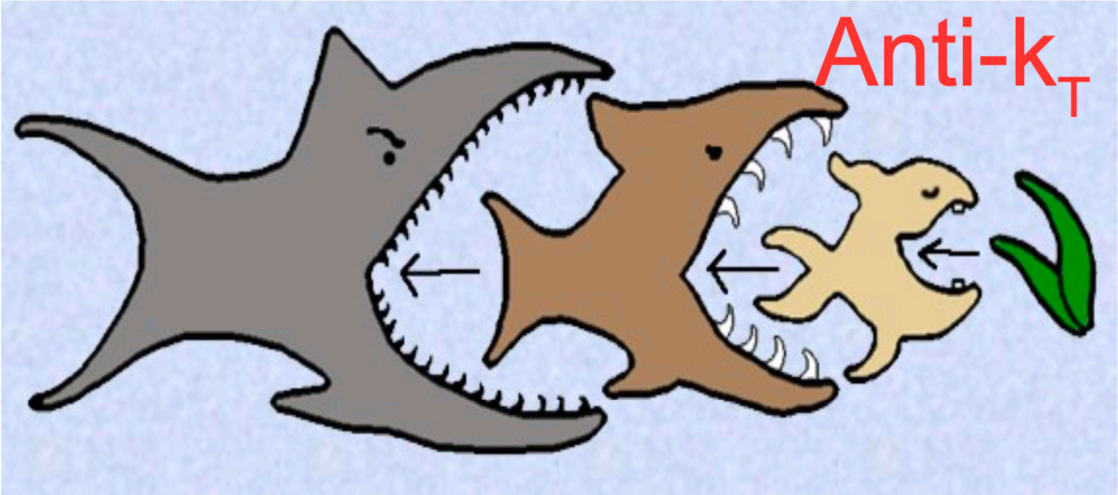
- Larger cone: top jet
- Smaller cone: W jet
- Much smaller: b-jet

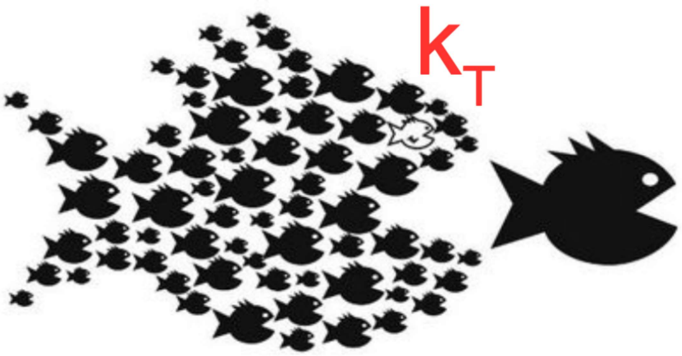
Master formula for heavy object: $\Delta R = \frac{2m}{p_T}$

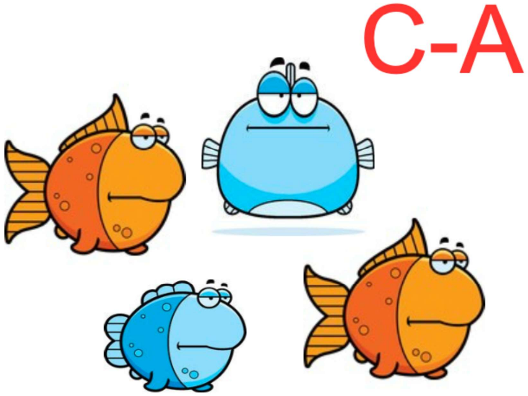
Types of jets in CMS

- We usually use two types of cone-size: 0.4 and 0.8
- Most often we use the anti-kT algorithm : AK4 and AK8
- In MINIAOD:
  - slimmedJets = ak4PFJetsCHS
    p_T >10 GeV (typically analysis cut will be at least p_T>20)
  - slimmedJetsPuppi = ak4PFJetsPUPPI
    p_T >170 GeV with all information, including PF candidate links(typically analysis cut will be at least p_T>200)
- In NanoAOD: https://github.com/cms-nanoAOD/cmssw/blob/master-cmsswmaster/PhysicsTools/NanoAOD/python/jets_cff.py
  - Jet = ak4PFJetsCHS
  - FatJet = ak8PFJetsPUPPI
  - Also see: https://twiki.cern.ch/twiki/bin/viewauth/CMS/JMECustomNanoAOD

Recap

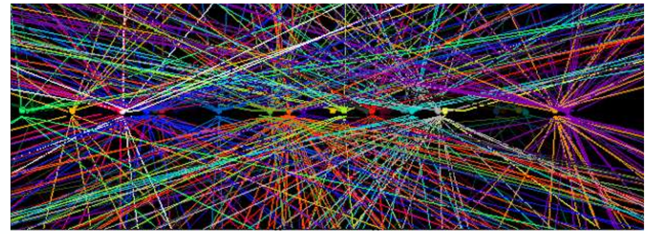




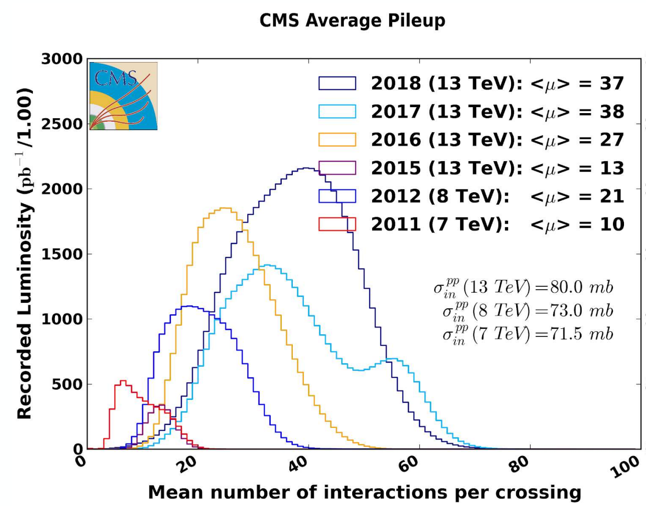


Section 2: Pileup

Pileup - PU

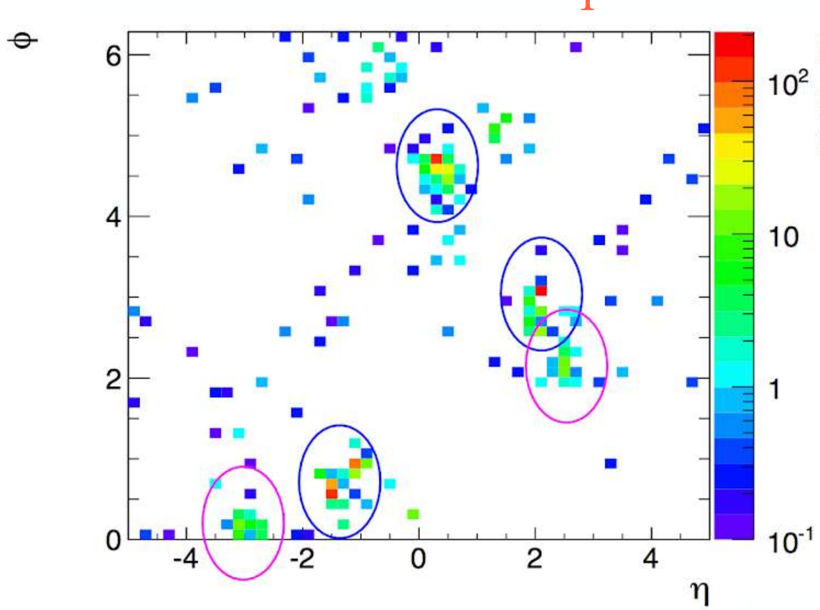


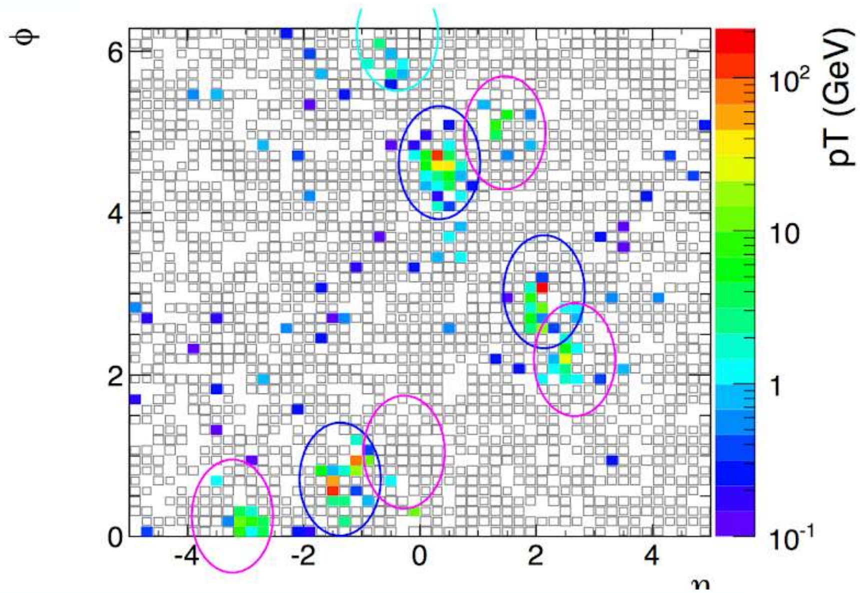
- What is pileup?
  - The **additional interactions** that occur in each bunch crossing because the instantaneous bunch-by-bunch luminosity is very high
    - “additional” implies that there is a hard-scatter interaction that has caused the event to fire the trigger
- Two things to know for your analysis:
  - Need to know amount of pileup - do PU reweighting
  - Need to suppress pileup from jets



Pileup in the detector

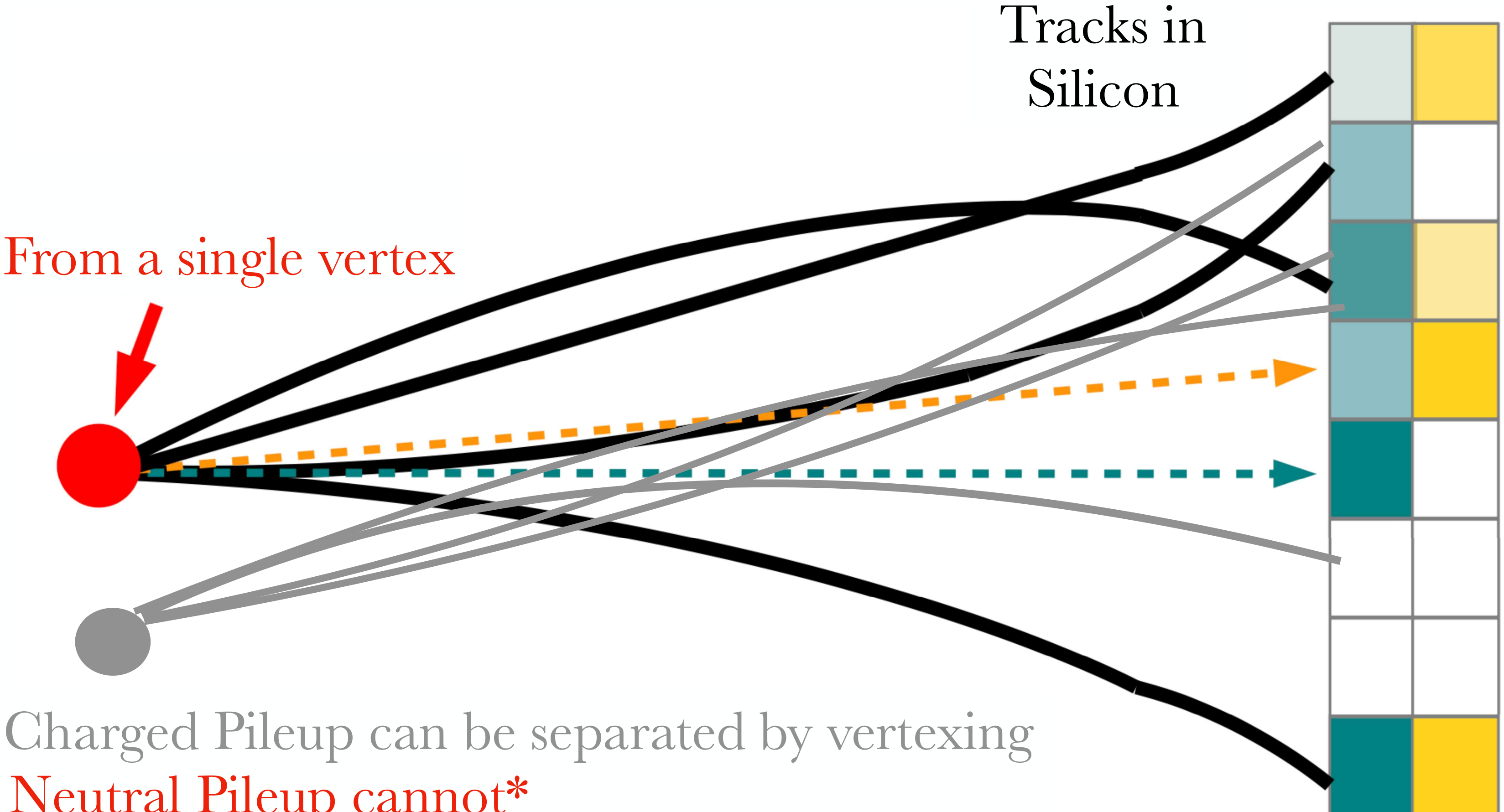
An LHC event at particle level





Now add pileup

Pileup

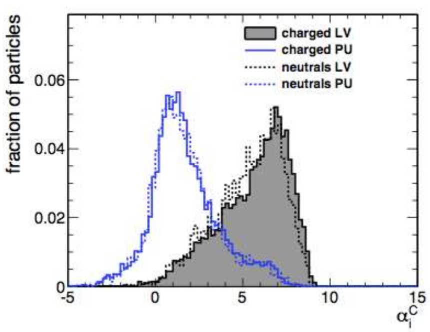


Charged Pileup can be separated by vertexing

Neutral Pileup cannot*

Pileup mitigation algorithms

- Different algorithms: Charged Hadron Subtraction, Δβ correction, etc
- Focus on PUPPI: Pileup per Particle Identification
  - Removes charged particles from pileup
  - Assign weights to neutrals based on the likelihood that they originate from the leading vertex or pileup



- This is done by defining a local metric that differs between pileup (PU) and the leading vertex (LV)

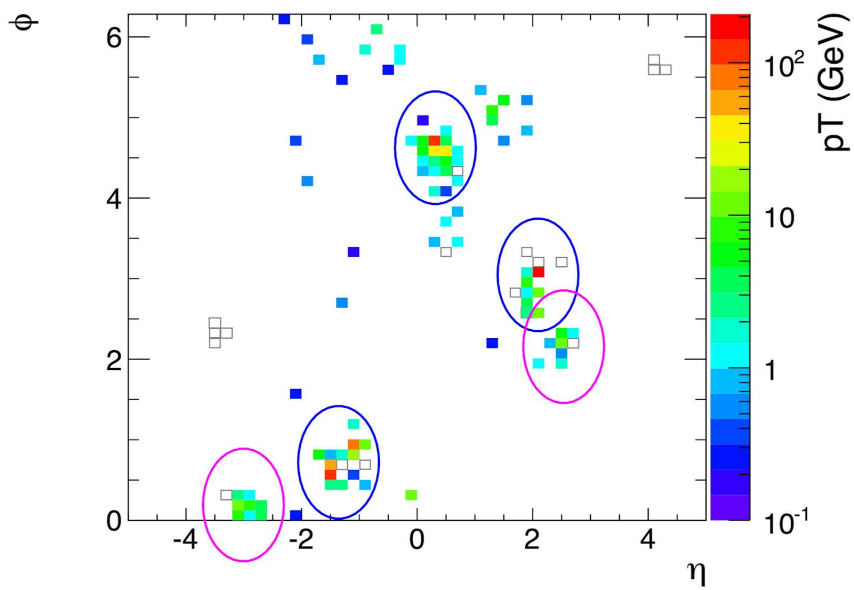
$$\alpha_i^C = \log\left[\sum_{j\in \mathrm{Ch,LV}} \boxed{\frac{p_{T,j}}{\Delta R_{ij}}}\,\Theta(R_0 - \Delta R_{ij})\right]$$

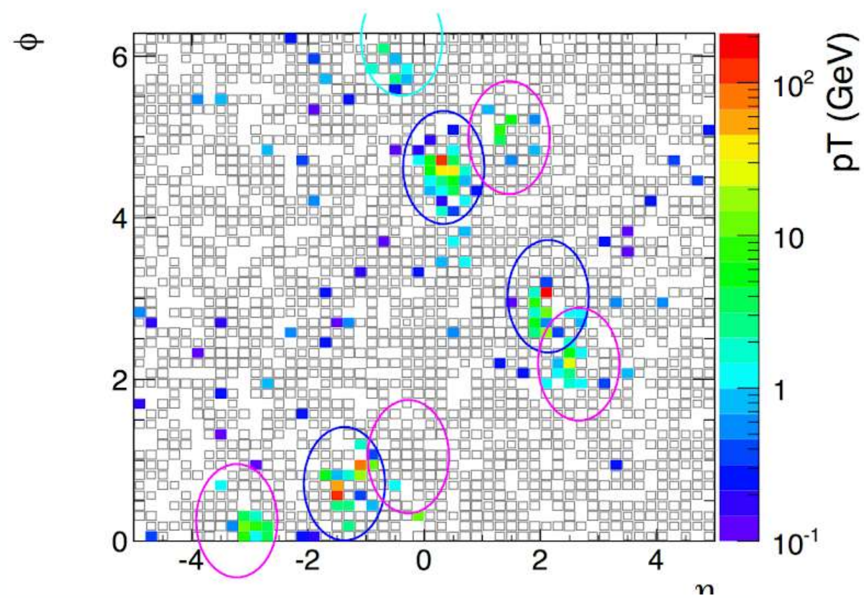
↑ Hard collinear particles

↓ Soft wide angle particles

After PUPPI

Reco with PUPPI



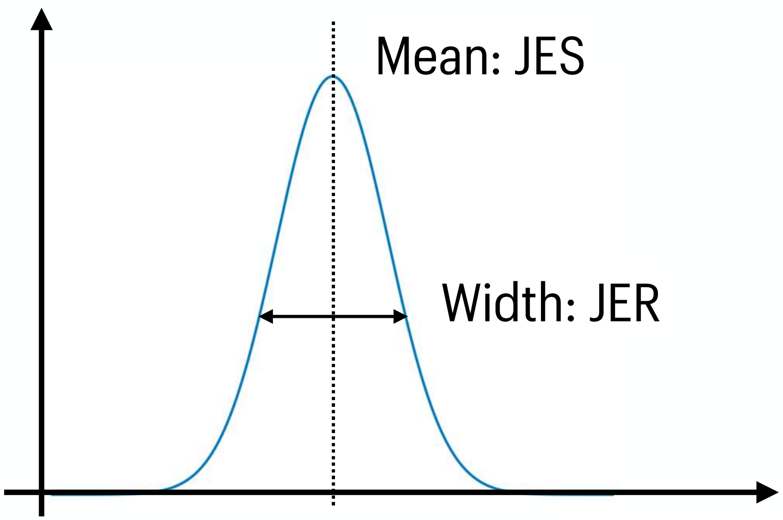


Now add pileup

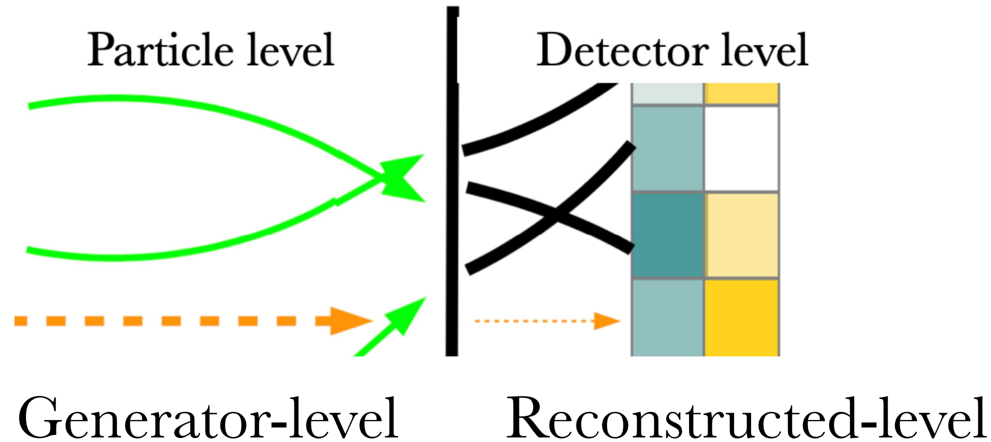
Section 3: Jet Energy Corrections

Jet energy corrections

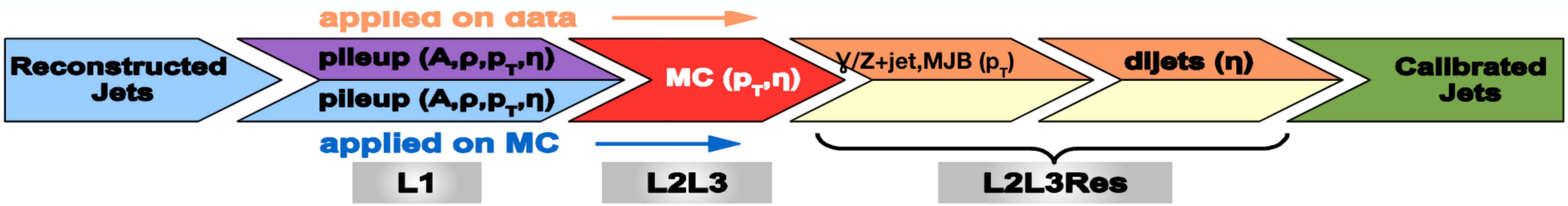
- We expect average Jet p_T response to be different:
    - Pileup adding energy
    - Non-linear calorimeter response



Jet p_T response = p_T / p_T^{gen}



Jet energy corrections

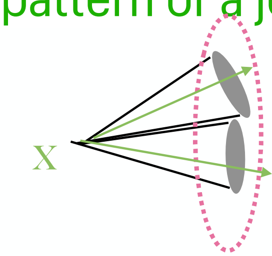


- We take factorized approach:
  - Pileup corrections to correct for offset energy
  - Correction to particle level jet vs. p_T and η from simulation
  - Only for data: Small residual corrections (Pileup/relative and absolute) to correct for differences between data and simulation

Section 4: Jet Substructure

Jet substructure

- A fancy name for an evolving field that studies the internal radiation pattern of a jet:
    - Becomes more interesting at high-Lorentz boost: boosted jets

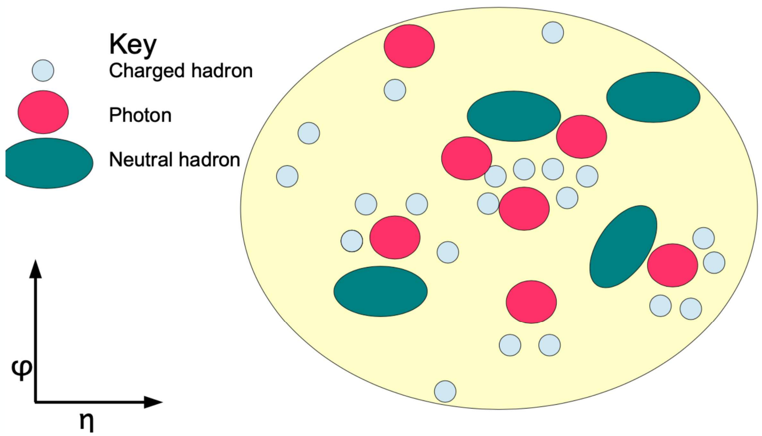


- A couple of useful topics:
    - **Jet grooming and jet mass**
    - **Jet-tagging**

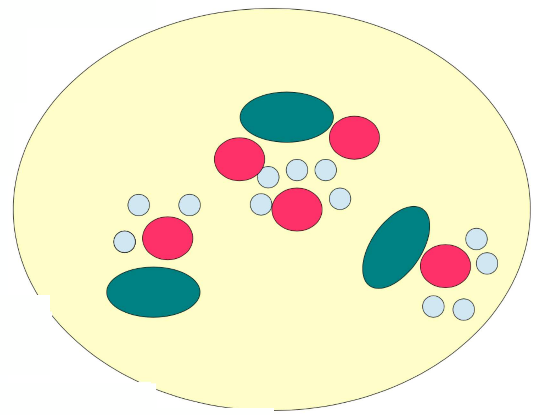
- Some great resources: (beyond CMS)
    - Looking inside jets: https://arxiv.org/pdf/1901.10342.pdf
    - Review in theory and machine learning: https://arxiv.org/abs/1709.04464
    - Experimental review: https://arxiv.org/abs/1803.06991
    - CMS's latest boosted jet paper: https://arxiv.org/abs/2004.08262

Jet tagging

- Grooming **removes a jet's soft and wide-angle radiation**
- All algorithms need to re-cluster to start:



- Soft-drop is widely used in CMS:

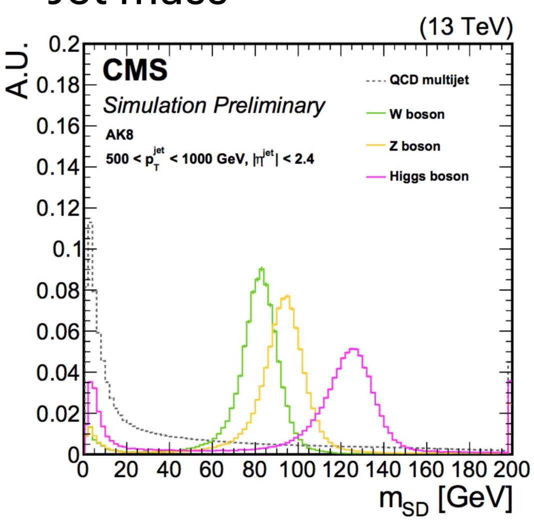


- Iteratively declusters jet removing lowest p_T subjet failing pairwise condition

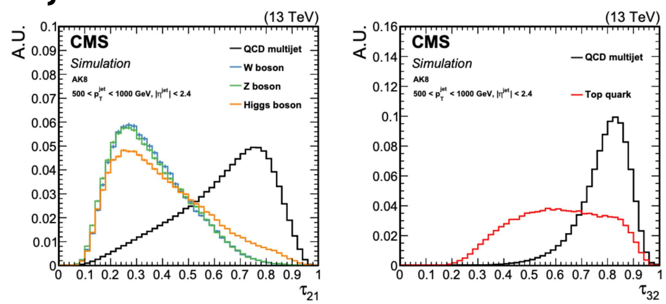
$$z = \frac{\min(p_{\mathrm{T}}^1, p_{\mathrm{T}}^2)}{p_{\mathrm{T}}^1 + p_{\mathrm{T}}^2} \qquad z > z_{cut} \left(\frac{\Delta R_{12}}{R}\right)^{\beta}$$

Boosted Jet tagging

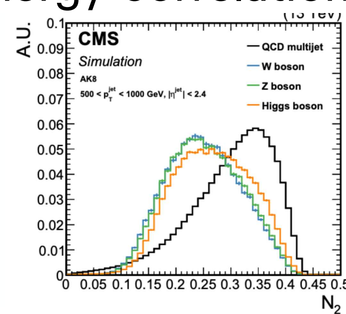
Jet mass



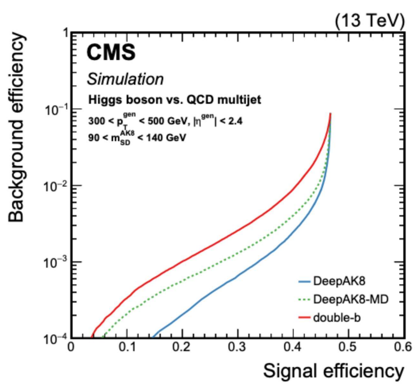
N-subjetiness



Energy-correlation-functions



DeepAK8



Particle-Net