1 Jets reconstruction

Quarks produced in hard scattering of partons in pp collisions manifest themselves as hadronic jets. They are reconstructed in CMS by clustering PF candidates using the anti-k_T algorithm [1, 2]: a sequential recombination jet algorithm parametrized by the negative power of the energy scale in the distance measure. To understand how it works, it is useful to introduce the following distances:

$$d_{ij} = \min(k_{Ti}^{-2}, k_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R} \qquad\qquad d_{iB} = k_{Ti}^{-2} \tag{1}$$

where d_{ij} is between entities (particles, pseudojets) i and j and d_{iB} is between entity i and the beam (B). The clustering proceeds by identifying the smallest of the distances. If it is a d_{ij} it combines entities i and j in a new single entity, while if it is d_{iB} it calls i a jet and removes it from the list of entities. The distances are recalculated and the procedure repeated until no entities are left. The jet extension is encoded in our definition of the distance measures: where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, and η_i and ϕ_i are respectively the rapidity and azimuth of particle i. The distance parameter R governs the radius of the clustering. The jet four momentum is computed as the vector sum of the clustered PF candidates four momenta, and a set of Monte Carlo (MC) simulation based corrections are applied to account for contribution from pileup, non-linearity in the detector response and residual data-simulation differences. The jet reconstruction is typically validated using dijet, multijet, γ+jets, and leptonic Z+jets events [3–5]. Typical jets resolutions achieved are of about 15–20% for at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [6].

For the 13 TeV data sets, jets are usually reconstructed with a distance parameter $R = 0.4$ and are denoted as AK4 jets. Analyses exploring Lorentz-boosted hadronic objects also employ large-radius jets clustered using the anti-k_T algorithm with a distance parameter $R = 0.8$. These jets are denoted as AK8 jets.

2 Tagging of b jets

To identify hadronic jets originating from the fragmentation of b quarks, as in this work, the properties of the b-hadrons are used. These hadrons have relatively large masses ($5 - 10$ GeV), long lifetimes ($10^{-12} - 10^{-13}$ s), and daughter particles with hard momentum spectra (typically peaking near the b quark mass, and extending to much higher momenta, dropping by about a decade for every ten GeV) [7]. CMS has developed a variety of algorithms to identify b quarks based on variables such as the impact parameter of charged particle tracks, the properties of reconstructed decay vertices, and the presence of a lepton, or the combination of the above information. This procedure of identification is known as b tagging.

In this work the b tagging algorithm used is the one known as *DeepFlavour* (or sometimes referred to as *DeepJet*): a deep-neural-network (DNN) algorithm based on 16 properties of up to 25 charged and 6 properties of 25 neutral particle-flow jet constituents, as well as 17 properties from up to 4 secondary vertices associated with the jet. More detailed information about the *DeepFlavour* algorithm and its performance can be found in Ref. [8, 9].

References

[1] M. Cacciari et al. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[2] M. Cacciari et al. "FastJet user manual". In: *The European Physical Journal C* 72.3 (Mar. 2012). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-012-1896-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-012-1896-2`.

[3] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *Journal of Instrumentation* 12.02 (Feb. 2017), P02014–P02014. ISSN: 1748-0221. DOI: `10.1088/1748-0221/12/02/p02014`. URL: `http://dx.doi.org/10.1088/1748-0221/12/02/P02014`.

[4] CMS Collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (Nov. 2011), P11002–P11002. ISSN: 1748-0221. DOI: `10.1088/1748-0221/6/11/p11002`. URL: `http://dx.doi.org/10.1088/1748-0221/6/11/P11002`.

[5] CMS Collaboration. "Jet algorithms performance in 13 TeV data". In: *CMS-PAS-JME-16-003* (Mar. 2017). URL: `http://cds.cern.ch/record/2256875`.

[6] CMS Collaboration. "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018". In: *CMS-DP-2020-019* (Apr. 2020). URL: `https://cds.cern.ch/record/2715872`.

[7] M. Tanabashi and others (Particle Data Group). "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: `10.1103/PhysRevD.98.030001`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.98.030001`.

[8] CMS Collaboration. "Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector". In: *CMS-DP-2018-033* (June 2018). URL: `https://cds.cern.ch/record/2627468`.

[9] CMS Collaboration. "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13TeV with Phase 1 CMS detector". In: *CMS-DP-2018-058* (Nov. 2018). URL: `https://cds.cern.ch/record/2646773`.