

The secondary vertex finding algorithm with the ATLAS detector

Sebastian Heer^{*†}

On behalf of the ATLAS Collaboration

University of Bonn (DE)

E-mail: heer@uni-bonn.de

A high performance identification of jets, produced via fragmentation of bottom quarks, is crucial for the ATLAS experiment, running at the LHC. These jets can be identified by exploiting the presence of cascade decay vertices from bottom hadrons. A general vertex-finding algorithm is introduced and its application to the search for secondary vertices inside jets is described. Kinematic properties of the reconstructed vertices are used to construct several b -jet identification algorithms. The features and performance of the secondary vertex finding algorithm in a jet, as well as the performance of the jet tagging algorithms, are studied using simulated $pp \rightarrow t\bar{t}$ events at a centre-of-mass energy of 13 TeV.

The European Physical Society Conference on High Energy Physics

5-12 July

Venice, Italy

^{*}Speaker.

[†]Supported by European Research Council grant ERC-CoG-617185

1. Motivation

The ability to identify jets containing b -hadrons (b -jets) is important for the physics program of the ATLAS experiment [1] at the LHC. During the fragmentation process a b -quark forms a b -hadron which decays via the electroweak interaction, and thus a b -jet contains charged tracks coming from the decay of the b -hadron and tracks produced in the b -parton showering. The dominant decay modes of a b -hadron are characterized by a transition of a b -quark to a c -quark, followed by the c -hadron decay. b -hadrons have relatively long lifetimes (e.g. 1.5 ps for the B^0 meson) and can travel a few millimetres before they decay. Charmed hadrons lifetimes are smaller, e.g. 1.0 ps for the D^\pm meson and 0.4 ps for the D^0 meson, but they can also travel significant distances at high energies. The displaced vertices resulting from the b -hadron and subsequent c -hadron decays provide a strong signature for a b -quark jet. These vertices can be reconstructed explicitly and their properties can be used to construct efficient b -jet identification algorithms (b -tagging), that are able to separate jets likely to contain b -hadrons from jets containing light or c -hadrons.

2. The secondary vertex finder algorithm

The ATLAS implementation of the secondary vertex finder (SVF) [2] is primarily designed to be used for jet flavour identification. Therefore only tracks inside a given jet cone are considered when searching for secondary vertices. The inputs of the algorithm are the jet axis direction, the primary vertex position of the event and a list of tracks associated to the jet. If successful, the algorithm returns one or more reconstructed secondary vertices, the list of tracks attached to each of them and a set of secondary vertex kinematic features which can be exploited by the subsequent jet tagging algorithms.

The SVF works as a Single Secondary Vertex Finder (SSVF), trying to reconstruct a single secondary vertex per jet. For a b -jet containing both b - and c -hadron decay vertices, the SSVF merges these vertices into a common single vertex if they are close, or reconstructs the vertex with the largest track multiplicity if they are far apart. The finite resolution of the ATLAS tracking detector does not allow the cascade decays of b - and c -hadrons to be resolved in every jet. Quite often the b - or c -vertex contain just one reconstructed track and cannot be reconstructed alone. Therefore the search for a single secondary vertex in a b -jet is justified and the vertex that is found should preserve most of the information necessary for the b -tagging. In addition the SSVF approach has a small probability of finding a (fake) vertex when there are no real secondary vertices in the jet. This feature is crucial for the rejection of light jets in the b -tagging.

All results are obtained using $t\bar{t}$ simulated events corresponding to a centre-of-mass energy of $\sqrt{s} = 13$ TeV proton-proton collisions, more details are available in Ref. [3]. Only $t\bar{t}$ decays with at least one lepton from a subsequent W -boson decay are included in the simulation.

3. Selection and cleaning of two-track vertices

The set of tracks passing the track selection cuts are used to form all possible two-track vertices. However the straightforward secondary vertex reconstruction based on the closeness of two tracks in the space results in a large fraction of fake vertices in an environment with high track

density (e.g. in the core of a jet or in events with very high pileup). Such fake vertices can be suppressed by exploiting the following observation: any secondary track should not have associated detector hits with radii smaller than the radius of the SV found by the algorithm. This observation is used to remove fake vertices by requiring that the hit patterns of the tracks associated to a two-track vertex are consistent with the position of the reconstructed secondary vertex inside the tracking detector.

Significant numbers of secondary vertices are produced by interactions of primary hadrons with detector material, photon conversions and decays of long-lived particles (K_S and Λ). These are real vertices, but they constitute a significant background to the decays of b - and c -hadrons, therefore they need to be eliminated from the b -jet identification process. To achieve the highest purity of the reconstructed b - and c -hadron decay vertices several different methods have been developed.

Hadronic interactions with the detector material are rejected based on comparison of the vertex radius with radii of cylindrical material layers. The nuisance two-track vertices produced by hadronic interactions are removed from further consideration if their radii are consistent within the detector material position. As true b - and c -hadron decay vertices cannot be further distinguished from hadronic interactions, the true decay vertices within the vetoed regions are removed as well, thus reducing the overall reconstruction efficiency.

Figure 1 shows the cleaning effect of applying the material interaction removal procedure and the V^0 cleaning procedure: applying both procedures, and applying each of them separately. The figure shows the radial distance between the final reconstructed secondary vertices and the nearest material layer. Hadronic interaction vertices are concentrated at the low end of these distributions, as can be clearly seen in the light jets plot, where a contribution of such vertices is significant. As expected, the vertices in b -jets are practically unaffected by the combined vertex cleaning procedure, while the number of reconstructed vertices in the light jets is significantly reduced.

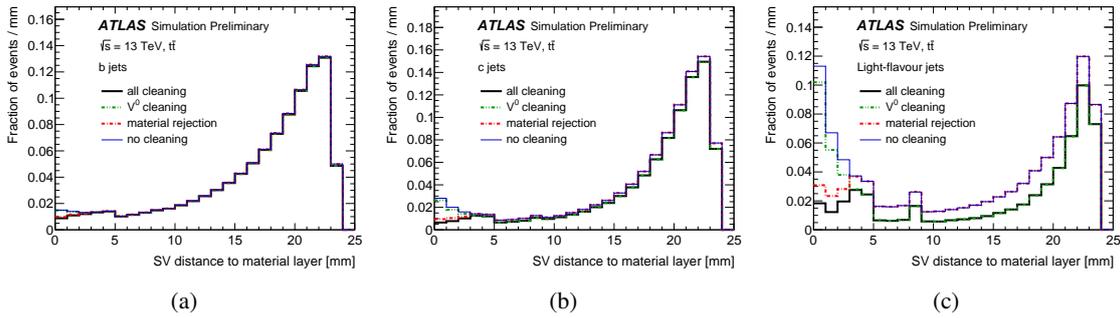


Figure 1: The distribution of the distance between the SV and the nearest material layer for different selection cuts for (a) b -jets, (b) c -jets, (c) light jets. Both of the cleaning algorithms are applied (black), no cleaning is applied (blue), only material rejection is applied (red), only V^0 cleaning is applied (green). The distributions are normalized to the curve where no cleaning is applied [2].

4. Properties of secondary vertices

Several variables are computed by the SVF for the final secondary vertices. The normalized distributions of the SV properties are shown in Figure 2: the SV mass, the energy ratio of the tracks contributing to the SV to the sum of energies of all the tracks in the jet, the number of two-track vertices (after the cleaning of two-track vertices has been applied). Not all hadronic interaction vertices can be removed without reducing the efficiency for real vertices, some enhancements around the material layers positions remain. Some of those variables are used as inputs to the SV based b -tagging algorithms.

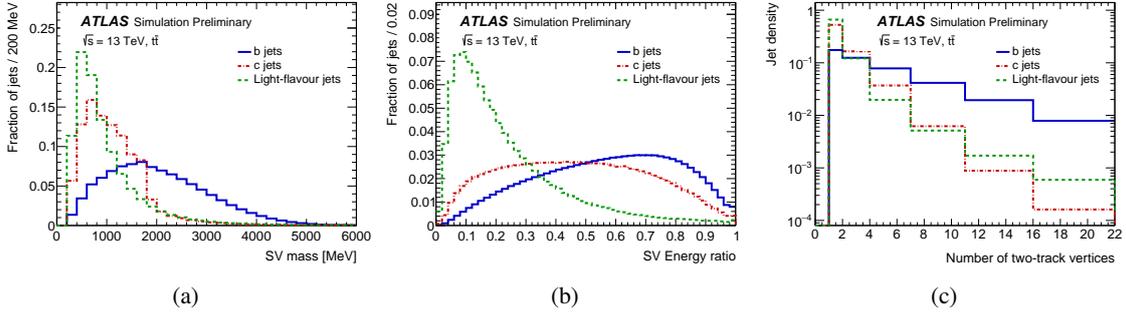


Figure 2: Secondary vertex properties distributions with b -jet (blue), c -jet (red) and light jet (green). All distributions have been normalized to one, including over- and underflow bins [2].

5. Performance of the secondary vertex finder

The secondary vertex reconstruction efficiencies, i.e. the fractions of jets which have a reconstructed secondary vertex, for different jet flavours of the SVF are shown as a function of jet η and jet p_T in Figure 3. An average reconstruction efficiency of $\sim 80\%$ is found in b -jets, with a drop for $|\eta| \geq 1.5$ being observed for both b - and c -jets. The probability to have fake secondary vertices in light jets shows very low values in the central region but an increase at high $|\eta|$ regions. Both effects can be explained by the larger amount of detector material in the endcap region for $|\eta| > 1.5$, resulting in higher hadronic interaction rate and decrease of the track reconstruction efficiency, and decrease of spatial accuracy of the track reconstruction there. In addition, the degraded tracking resolution augments the pileup contribution to the fake secondary vertex rate in the endcap region. For jet p_T , the increase of track density in high- p_T jets leads to degradation of the track reconstruction efficiency and consequently smaller amount of reconstructed tracks from b/c -hadron decays [4]. A larger amount of fragmentation tracks in a jet is responsible for the larger fake rate of reconstructed vertices, as well. As shown in the figure, the number of fake vertices is increasing with jet p_T in light jets, while the SV reconstruction efficiencies for b -jets slightly degrade with jet p_T . Increasing or decreasing, the efficiencies have a direct effect in degrading or improving the b -tagging efficiencies (see Figure 4).

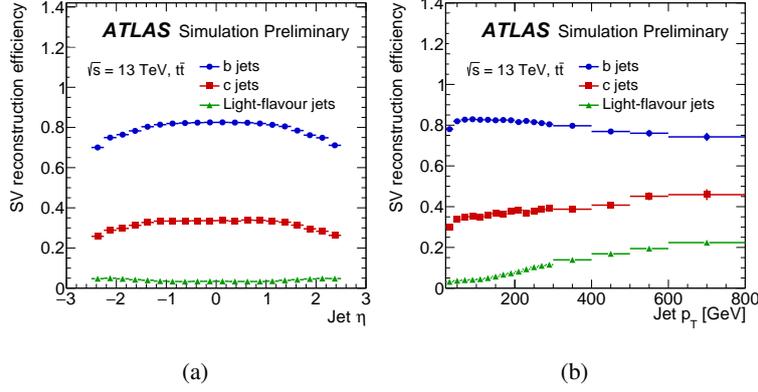


Figure 3: The secondary vertex reconstruction efficiencies for three jet flavours, as a function of jet η (a), as a function of jet p_T (b). b -jets in blue, c -jets in red, and fake vertices are shown in green for light jets [2].

6. Performance of the secondary vertex based b -tagging algorithms

The performance of a secondary vertex based b -tagging algorithm, like SV0 [5, 6] and SV1, is limited by the ability to reconstruct secondary vertices in a b -jet. As shown in previous sections (Figure 3), the SV reconstruction efficiency in a b -jet reaches a maximum $\sim 80\%$ at the jet transverse momentum ~ 80 GeV, then it gradually decreases at higher p_T . A more universal and powerful b -jet tagging algorithm can be constructed by combining the SV based algorithm with the track impact parameter based algorithm IP3D [7]. When a SV is reconstructed in a jet, a combination of the tracking information from IP3D and kinematical information from SV provides a significant improvement of the light jet rejection. When a SV is not reconstructed in a jet, the combined algorithm is still able to identify the b -jet using the IP3D information only, although with higher fake rate. The SV based algorithms and the impact parameter based algorithm IP3D provide jet weights depending on the jet flavour content. These weights are logarithms of the likelihood ratios and therefore can be combined by summing them, corresponding to the multiplication of the corresponding likelihoods.

The light and c -jet rejections versus the b -tagging efficiency for the algorithms SV0, SV1 and the combination of SV1 and IP3D are shown in Figure 4. The b -tagging efficiency is defined as the fraction of jets that are labeled as b -jets and tagged as b -jets by a tagging algorithm, while the light rejection is the inverse of the mis-tagging rate: The mis-tagging rate is the fraction of jets labeled as light jets and tagged as b -jets. At 70% b -tagging efficiency the combined algorithm has a 40% (60%) higher light rejection than the SV1 (SV0) algorithm.

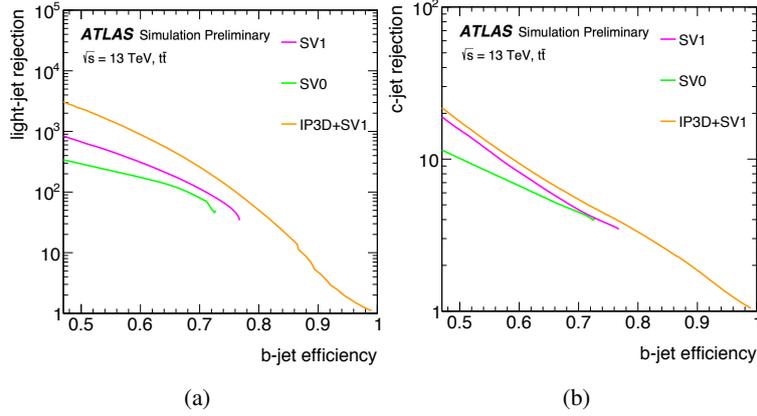


Figure 4: The SV0, SV1 and SV1+IP3D b -tagging efficiency versus the light jet rejection (a) and the c -jet rejection (b) versus b -jet efficiency. [2]

7. Conclusions

The algorithm represents a unified approach to exploit the presence of secondary vertices in a jet for the jet flavour identification. It shows a high performance and has a low fake rate. The kinematic features of the secondary vertices are used to construct several high performance b -jet tagging algorithms which can be used either separately or in combination with other b -tagging algorithms. The SVF algorithm is optimized and tuned to the information provided by the ATLAS detector to achieve the highest reconstruction efficiency and purity of the secondary vertices. In the central region of the detector a typical SV reconstruction efficiency for b -jets of 80% is achieved, with c -jet and light-jet efficiencies of 30% and $<5\%$. This efficiency reaches a maximum for jets with a transverse momentum of 80 GeV.

References

- [1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Inst 3 (2008) S08003
- [2] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, ATL-PHYS-PUB-2017-011
- [3] ATLAS Collaboration, *Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2016-004
- [4] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2*, CERN-EP-2017-045, arXiv: 1704.07983 [hep-ex]
- [5] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 900 GeV Collision Data*, ATLAS-CONF-2010-004
- [6] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 7 TeV Collision Data*, ATLAS-CONF-2010-042
- [7] ATLAS Collaboration, *Performance of b -Jet Identification in the ATLAS Experiment*, J. Inst 11 (2016) P04008, CERN-PH-EP-2015-216, arXiv: 1512.01094 [hep-ex]