ATLAS Note

KIP-2016

Draft version 0.1

Jet Observables using Subjet-assisted Tracks

The ATLAS Collaboration[1], Oleg Brandt[a], Sascha Dreyer[a], Fabrizio Napolitano[a]

[a]*Heidelberg University*

12th December 2016

This note presents the details of the Monte-Carlo studies on the subjet-assisted oberservables for groomed large-radius jet. In particular the observables for the Energy Correlation Functions and n-Subjettiness variables used by the ATLAS collaboration, C2, D2, τ_{21} and τ_{32} are discussed using subjet-assisted tracks; the mass observable contructed with this technique, m^{TAS}, is presented and discussed with a modified four-momentum prescription. In all the variables studied, large improvement have been found using this novel techniques, the first ones evaluating in terms of QCD event rejection in W/Z boson, top quark and Higgs boson tagging; the second one in terms of precision reconstruction of the large-radius jet mass.

Contents

1 Introduction

Jets are collimated streams of particles resulting from quarks and gluons fragmentation and hadronization. The distribution of energy inside a jet contains information about the initiating particle. When a massive particle such as a top quark, Higgs boson or W/Z bosons is produced with significant Lorentz boost and decays into quarks, the entire hadronic decay may be captured inside a single jet. The mass of such jets (jet mass) is one of the most powerful tools for distinguishing massive particle decays from the continuum multijet background; the Energy Correlation Functions and n-Subjettiness C2, D2, τ_{21} and τ_{32} provide an ad-hoc tool pupusely developed for the multijet background and constitue a fundamental part of many for boson taggers. This note documents the so-called subjet-assisted techniques with the ATLAS detector *insref*. The track-assisted subjet mass m^{TAS} definition is presented and confronted with the standard development in ATLAS, m^{comb} and m^{TA}. Energy Correlation Functions and n-Subjettiness with the modified subjet-assisted technique is presented and confronted with the standard one in ATLAS. The note ends with conclusions and future outlook in *insref*.

2 ATLAS detector

ATLAS (A Toroidal ApparatuS) is a multi-purpose particle detector with nearly 4π coverage in solid angle. A lead/liquid-argon sampling electromagnetic calorimeter is split into barrel ($|\eta| < 1.5$) and end-cap ($1.5 < |\eta| < 3.2$) sections. A steel/scintillating-tile hadronic calorimeter covers the barrel region ($|\eta| < 1.7$)

ATLAS	Description and performance
Magnetic field	2 T solenoid; 0.5 T toroid barrel and 1 T toroid end-cap
Tracker	Inner detector: IBL, Silicon pixel and strips, TRT $\sigma_{p_T}/p_T \simeq 5 \times 10^{-4} p_T \otimes 1\%$
EM calorimeter	EMB, EMEC and pre-sampler (Liquid Argon and lead) $\sigma_E/E \simeq 10\%/\sqrt{E} \otimes 0.7\%$
Hadronic calorimeter	Tile (Fe and scintillating tiles) and HEC (Cu and LAr) $\sigma_E/E \simeq 50\%/\sqrt{E} \otimes 3\%$
Muons	Inner detector and muon spectrometers $\sigma_{p_T}/p_T \simeq 2\%$ at 50 GeV $\sigma_{p_T}/p_T \simeq 10\%$ at 1 TeV
Trigger	L1 and HLT (L2 and EF) Rates from ~40 MHz to ~75 kHz (L1) and to ~200 Hz (HLT)

and two end-cap copper/liquid-argon sections extend to higher pseudo-rapidity ($1.5 < |\eta| < 3.2$). Finally, the forward region ($3.1 < |\eta| < 4.9$) is covered by a liquid-argon calorimeter with Cu (W), absorber in the electromagnetic (hadronic) section. Inside the calorimeters there is a 2 T solenoid that surrounds an inner tracking detector which measures charged particle trajectories covering a pseudo-rapidity range $|\eta| < 2.5$ with pixel and silicon micro-strip detectors (SCT) and additionally which covers the region $|\eta| < 2.0$ with a straw-tube transition radiation tracker (TRT). Outside the calorimeter there is a muon spectrometer: a system of detectors for triggering up to $|\eta| < 2.4$ and precision tracking chambers up to $|\eta| < 2.7$ inside a magnetic field supplied by three large superconducting toroid magnets.

A breakdown of the ATLAS sub-detector performance is shown in Table **??**.

3 Monte Carlo Samples

refraseMT The samples used are divided into two main groups: SM background and beyond SM signal. The SM background includes the QCD multijet samples, produced with a falling p_{T} spectrum. The beyond SM signals are $W' \to WZ \to q\bar{q}'q\bar{q}$, $Z' \to t\bar{t}$ (top quarks considered in the full hadronic channel ($t \to W(\to q\bar{q}')b$)) and RS-Graviton $\to hh \to b\bar{b}b\bar{b}$, i.e. final states have only jets in all the samples. The details of the samples are given in Table 1; the masses considered span from 0.5 to 5 TeV to improve and diversify the kinematic space covered.

A set of kinematic distributions for the W' is shown in Figure **??**: on the left the p_{T} distribution where the kinks correspond to the Jacobian peak of the mass considered and the η distribution on the right. The green dots represent the distribution before the selection, which is $p_{\mathrm{T}} > 250$ GeV and $|\eta| < 2.0$ and the red dots after this selection. This selection typical for many searches for BSM physics. All the other samples and the background can be found in the Appendix. In what follows, it will also be used the nomenclature

Process	ME Generator & Fragmentation	ME PDFs	UE Tune	Resonance Masses
QCD multijet	Pythia 8	NNPDF23LO	A14	N/A
$W' \to WZ$	Pythia 8	NNPDF23LO	A14	1.5, 2.5, 3, 4, 5 TeV
$Z' \to t\bar{t}$	Pythia 8	NNPDF23LO	A14	1.5, 1.75, 2.5, 3, 4, 5 TeV
$G_{RS} \to hh(\to b\bar{b})$	Pythia 8	NNPDF23LO	A14	0.5, 1, 1.5, 2, 2.5, 3 TeV
$W' \to \tilde{W}\tilde{W}$ with $m_{\tilde{W}} = m_t$	Pythia 8	NNPDF23LO	A14	1.5, 2.5, 3, 4, 5 TeV

Table 1: Overview of the Monte Carlo Samples used. The first line shows QCD standard model process, the second, the third and the forth the beyond SM samples considered; the last line the "massive W/Z" sample.

boosted W/Z for the W' sample, *boosted tops* for the Z' sample, *boosted Higgs* for the G_{RS} sample and *massive W* for the $W' \to \tilde{W}\tilde{W}$ with $m_{\tilde{W}} = m_t$. *refraseMT*

4 Object Definition

4.1 Large-radius jet mass definitions

4.2 Energy Correlation Functions

Information about the substructure of large-R jets can be used to discriminate between different event topologies. These are one, two and respectively three hard substructures (or prongs) inside the large-R jet. QCD jets are characterized by one hard substructure, jets originated by W or Z bosons feature two and Top quark jets feature three substructures (hadronic decay channels).

The Energy Correlation Functions ECF(N,β) or N-point correlators, described in Reference [**bib:ECF**], explore the substructure of a jet using a sum over the constituents. The correlation between pairs and triples of constituents is considered by the product of their p_{T}, multiplied by the angular weighting, which is defined by the product of the pairwise angular distances of the considered constituents. This angular part can be scaled against the momentum part via an exponent β. The default value for β is 1, corresponding to angular and momentum parts being weighted equally.

$$\begin{aligned} \mathrm{ECF1} &= \sum_{constituents} p_{\mathrm{T}} \\ \mathrm{ECF(2,}\beta\mathrm{)} &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} p_{\mathrm{T},i} p_{\mathrm{T},j} \Delta R_{ij}^{\beta} \\ \mathrm{(ECF(3,}\beta\mathrm{)} &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{k=j+1}^{n} p_{\mathrm{T},i} p_{\mathrm{T},j} p_{\mathrm{T},k} (\Delta R_{ij} \Delta R_{ik} \Delta R_{jk})^{\beta} \end{aligned} \tag{1}$$

The ECF(N) variables can be expanded straightforwardly to larger values of N by considering this definition. With this, ECF(2) uses pairwise correlation and is sensitive to two-prong structures, whereas ECF3 relies on triple-wise correlations to identify three-prong structures. ECF(1) corresponds to the p_{T} of the whole jet by a summation over the constituents p_{T}, thereby serving as normalization to minimize the energy scale dependence.

The ECF(N) variable tends to very small values for collinear or soft configurations of N constituents and is defined to be zero for jets with less than N constituents. For ECF(2), only pairs of constituents that are angular separated but not soft result in sum terms that are non-negligible, which directly leads to the picture of two hard substructures inside the jet. A similar conclusion can be made for ECF(3) and three hard substructures. Resulting from this, a jet with N or more hard substructures features a high ECFN value while a jet with fewer than N substructures has a lower ECF(N) value. Consequently, one can define ratios of Energy Correlation Functions. Two of them, called C2 and D2 are found to be very powerful to distinguish between one- and two-prong like jets, see e.g. Reference [**bib:power_counting**].

$$\begin{aligned} \mathrm{C2} &= \frac{\mathrm{ECF(3)} \cdot \mathrm{ECF(1)}}{\mathrm{ECF(2)}^2} \\ \mathrm{D2} &= \frac{\mathrm{ECF(3)} \cdot \mathrm{ECF(1)}^3}{\mathrm{ECF(2)}^3} \end{aligned} \tag{2}$$

E.g. a jet originated from a W boson features a small ECF(3) but a high ECF(2) value resulting in small C2/D2, corresponding to a high agreement with the two-prong hypothesis. QCD jets feature a very small ECF(3) and a small ECF(2) value. This results, considering the power of ECF(2) in the definitions, in a higher C2/D2 value as for a W boson jet. These variables are IRC-safe for $\beta > 0$ and theoretically very well understood, see Reference [**bib:analytic_ECF**]. D2 was found to perform slightly better for tagging W boson jets as C2 in Reference [**bib:w_tagging**], most notably due to a more p_{T} robust cut value and a somewhat higher background rejection.

4.3 n-Subjettiness

4.4 N-Subjettiness

The n-Subjettiness variable τ_N, introduced in Reference [**bib:nsub**], quantifies the level of agreement between a given large-R jet and a certain number N of subjet axes. Several possibilities to define the subjet axes exist. Two often used definitions are k_{T}-axes and the k_{T}-WTA (Winner Takes All) definition. In both cases, the jet is reclustered with an exclusive k_{T}-algorithm, that is running the recombination just until N subjets are clustered. The k_{T}-axes are defined by the four-momenta of the k_{T}-subjets, WTA correspond to the four-momentum of the hardest constituent in each k_{T}-subjet.

As C2 and D2, N-Subjettiness is a measure for the whole jet, calculated via a sum over the jets constituents (calorimeter clusters as default).

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k}\, min(\Delta R_{1,k}, \Delta R_{2,k}, ..., \Delta R_{N,k})^{\beta} \tag{3}$$

For each term, the constituents p_{T} is multiplied by the distance to the nearest subjet axes. The overall value is normalized with a sum over the constituents p_{T} times the characteristic radius parameter R of the large jet.

$$d_0 = \sum_k p_{T,k} R_0 \tag{4}$$

Similar to ECF(N,β), the angular measure ΔR_{ij} can be scaled relative to the p_{T} factor via the exponent β. N-Subjettiness is an IRC-safe variable for values of $\beta \geq 0$.

Small values of τ_N correspond to a jet with all constituents more or less aligned or near to the given N subjet axes, hence the jet is compatible with the assumption to be composed of N or fewer subjets. A higher value in contrast indicates a consistency with more than N subjets as a non negligible part is located apart of the N subjet axes. Consequently, W/Z or Higgs boson jets are likely to feature a small τ_2 and a high τ_1 value. QCD jets with their one-prong structure result in a high τ_2 and a small τ_1 value. While τ_1 and τ_2 alone provide only slightly separation, the ratio

$$\tau_{21} = \frac{\tau_2}{\tau_1} \tag{5}$$

is an effective discrimination variable.

The extension to three-prong like jet identification and discrimination from one and two-prong structures follows quite naturally by taking the ratio of τ_3 and τ_2.

$$\tau_{32} = \frac{\tau_3}{\tau_2} \tag{6}$$

Consequently, the hadronic decay of top quarks via $t \to Wb$ and the W decaying into two quarks can be tagged using the τ_{32} variable.

5 Track-assisted subjet mass

6 Energy Correlation Functions and n-Subjettiness

7 Conclusions

Appendix

In a paper, an appendix is used for technical details that would otherwise disturb the flow of the paper. Such an appendix should be printed before the Bibliography.

List of contributions

Auxiliary material

In an ATLAS paper, auxiliary plots and tables that are supposed to be made public should be collected in an appendix that has the title 'Auxiliary material'. This appendix should be printed after the Bibliography. At the end of the paper approval procedure, this information can be split into a separate document – see `atlas-auxmat.tex`.

In an ATLAS note, use the appendices to include all the technical details of your work that are relevant for the ATLAS Collaboration only (e.g. dataset details, software release used). This information should be printed after the Bibliography.