IL NUOVO CIMENTO Vol. 35 C, N. 3 Maggio-Giugno 2011
DOI 10.1393/ncc/i2012-11222-9

Colloquia: TOP2011

Boosted objects at the LHC

M. Spannowsky
Institute for Particle Physics Phenomenology, Department of Physics, Durham University DH1 3LE, UK

ricevuto l' 1 Marzo 2012

Summary. — Almost all theoretical extensions of the Standard Model predict heavy TeV-scale resonances which have to couple to electroweak-scale resonances, *e.g.*, top quarks or electroweak gauge bosons. Therefore, boosted electroweak-scale resonances with large branching ratios into jets is a highly probable scenario in many processes probing new physics. Here, jet substructure methods can help to disentangle the sought-after signal from the backgrounds. In this brief review we classify scenarios where jet substructure methods can be beneficial for new physics searches at the LHC and discuss the application of the HEPTopTagger in some of these scenarios.

PACS `14.65.Ha` – Top quarks.

1. – Introduction

The large potential for searches of new electroweak-scale particles by looking inside a fat jet has only been appreciated recently [1, 2]. At the LHC with its 14 TeV center-of-mass energy, particles with masses around the electroweak scale are frequently produced beyond threshold, *i.e.* boosted transverse to the beam direction. Either because they recoil against other energetic resonances or because they arise from decays of even heavier particles, *e.g.*, Z or KK-gluons. If the resonances transverse momentum is bigger than their mass, their decay products tend to be collimated in the lab frame. Thus, only a small part of the detector has to be considered to reconstruct the boosted resonance.

However, at LHC many sources of hadronic radiation exist. Apart from the so-called final state radiation (FSR) off the decay products of an electroweak-scale resonance, proton-bunch crossings give rise to radiation from the initial state (ISR), the underlying event (UE) and pileup. Initial state radiation are soft and collinear jets, arising because the incoming partons have to bridge the gap in scale between the proton and the hard process. Underlying event is additional soft QCD activity arising from a given proton-proton interaction surrounding the hard event. It is caused by semi- or non-perturbative interactions between the proton remnants. Finally, pileup is the effect of multiple proton-proton collisions in one beam crossing.

For an optimal discrimination of a hadronically decaying electroweak resonance from QCD jets, the resonance's FSR has to be disentangled from ISR, UE and pileup.

Sequential jet algorithms [3-5], popular for their infrared safety, allow to associate a recombination history to every jet. Therefore, a jet is not only a massive object with a specific cone size and a three-momentum but has a well defined internal structure. Thus, more information is accessible to discriminate the signal from backgrounds. Over the last few years a plethora of different methods has been proposed to use the internal structure of jets in searches for new physics [2, 6].

In this brief review we identify four different kinematic configurations where it can be beneficial to use subjet techniques, see sect. **2**. We find that these configurations cover a wide range of new physics scenarios. However, the challenges are manifold and different approaches are necessary to disentangle electroweak-scale resonances from the backgrounds.

Soft uncorrelated radiation, *e.g.*, UE or pileup, has a large effect on the mass of a fat jet. Thus, in searches which require a large fat-jet cone size grooming techniques are mandatory. In sect. **3** we discuss the most popular grooming techniques, focusing on their similarities and differences.

In sect. **4** we present one example of a top tagger, the so-called HEPTopTagger (Heidelberg-Eugene-Paris). This tagger is designed to tag top quarks with medium transverse momentum ($p_{T,t} \simeq m_t$). We discuss several examples of new physics searches where this tagger can be used.

2. – Kinematic configurations

Different scenarios have been identified where jet substructure techniques are of value to improve phenomenological analysis. They can be cast into four different classes characterized by the resonances transverse momentum and the overall business of the event.

2˙1. *Large transverse momentum, non-busy final state.* – If the event is generated with a large invariant mass and the number of physical objects is small, *e.g.*, a heavy TeV-scale resonance decays into two electroweak-scale resonances ($pp \to X_{\mathrm{TeV}} \to 2Y_{\mathrm{EW}} \to \text{jets}$), the electroweak scale objects Y_{EW} tend to be highly boosted, their decay products are highly collimated and the radiation off the two Y_{EW} are well separated with respect to each other. Due to $p_{T,Y} \gg m_Y$ the fat jets cone size does not need to be very large ($R \simeq 0.8$) to catch all the necessary FSR to reconstruct the boosted resonances. Because of the relatively small cone size and the absence of many sources of hard radiation in the event, the effect of UE, pileup, and ISR is less pronounced and one finds $m_j \simeq m_Y$. The importance of jet grooming methods to remove uncorrelated soft radiation is reduced but jet substructure methods are necessary to discriminate between a resonance jet and a QCD jet.

2˙2. *Large transverse momentum, busy final state.* – SUSY cascade decays are examples where, depending on the mass splitting of the SUSY particles in the decay chain, the boosted resonances can have large transverse momentum accompanied by many hard jets: $pp \to 2X_{\mathrm{TeV}} \to Y_{\mathrm{EW}} + \text{jets} \to \text{jets}$. Considering UE, pileup, and ISR, the situation is similar to the non-busy final state, but the jet substructure methods might need to be adjusted to the many additional sources of hard uncorrelated radiation in the fat jet.

2˙3. *Medium transverse momentum, non-busy final state.* – If two electroweak scale resonances are directly produced from proton collisions, $pp \to Y_{\rm EW} Y'_{\rm EW}$, they are usually produced close to threshold, *e.g.*, $pp \to HW$. Only a small fraction of the events yield $p_{T,Y} > m_Y$. Still, focusing on this fraction of events can be a superior way to disentangle the signal from the backgrounds. Phenomenological studies [7] have shown that, because of the kinematic features, event reconstruction efficiencies, b-tagging efficiencies and the jet energy resolution can be improved and that combinatorial problems in the identification of the decay products of $Y_{\rm EW}$ are reduced. To make as many signal events accessible to jet substructure methods as possible the fat jets cone size has to be much bigger than in the highly boosted scenarios. Therefore, UE, pileup, and ISR affect the fat jets mass much stronger and one should consider using techniques to remove uncorrelated soft radiation, *e.g.*, jet grooming methods.

2˙4. *Medium transverse momentum, busy final state.* – Events with many sources of hard radiation and less collimated decay products of the $Y_{\rm EW}$ are the most difficult scenarios, irrespective if jet substructure techniques are applied or not. Although many of the advantages outlined for the medium transverse momentum non-busy final states can be carried over, combinatorial problems to discriminate between decay products and hard radiation from other sources in the fat jet can occur much more frequently. Taggers for $Y_{\rm EW}$ have to take that into account and combine jet grooming techniques with criteria to discriminate $Y_{\rm EW}$ decay products from other hard radiation.

3. – Jet grooming methods

Hadronic final states of hard interactions resulting form proton-bunch crossings at the LHC are subject to many sources of QCD radiation. Final state radiation, soft and collinear jets radiated off the decay products of an electroweak scale resonance, can be described well using the parton shower.

At the LHC, the amount of transverse momentum of the underlying event radiation and pileup per unit rapidity can be large [8,9] and their effect on the jet mass depends on the cone size of the fat jet [10]. Some sequential jet algorithms, such as the k_T and the C/A algorithm aim to preserve the physical picture of the jet evolution from the hard scale to the hadronization scale in the recombination sequence. Soft uncorrelated radiation, *i.e.* UE and pileup, spoils this picture. An additional complication in identifying events with hadronically decaying electroweak resonances is that splittings of quarks and gluons can geometrically induce a jet mass of the order of the electroweak scale.

Jet grooming methods, like filtering, trimming and pruning, remove soft uncorrelated radiation from a fat jet while retaining final state radiation off the resonance. For QCD jets grooming methods reduce the upper end of the jet mass distribution, whereas for signal events they yield a sharper peak near the true resonance mass $m_j = m_{\rm res}$. To keep these methods generic it is implicitly assumed that for boosted heavy particles $p_{T,\rm FSR} > p_T$, $(\rm ISR, UE, PU)$. Thus, the transverse momentum of the subjets is an important criterion to discriminate between final state radiation and other radiation.

3˙1. *Filtering.* – Filtering, the first proposed jet grooming method, was introduced as part of the so-called BDRS Higgs tagger [1]. Its target application is HW and HZ production with a leptonic decay of the gauge bosons and with the Higgs boson decaying to $b\bar{b}$. A mass drop requirement identifies the vicinity of the Higgs decay products. The procedure called filtering then performs a recombination of the remaining fat jet constituents with a much smaller cone size, R_f. It results in n_f small subjets. This obviously reduces the effective area of the fat jet considered for mass reconstruction and this way tames any QCD effects scaling with R. For the Higgs boson the best mass resolution is achieved by reconstructing the Higgs mass from the $n_f = 3$ hardest filtered subjets. This means we include two b-jets and the hardest wide-angle gluon radiation. Two free parameters, R_f and n_f control the filtering performance.

3˙2. *Trimming.* – Trimming [11] targets very similar effects as filtering. In the first step we reconstruct a fat jet which will be heavily impacted by QCD radiation. Its subjets we recombine with a higher resolution R_{trim}, defining a larger number of smaller subjets. These subjets can be separated into two categories: hard and soft. This discrimination is based on the transverse momentum, so hard subjets obey $p_{T,j} > f_{\text{trim}}\Lambda_{\text{trim}}$, where f_{trim} is an adjustable parameter and Λ_{trim} is an intrinsic scale of the fat jet. It can for example be chosen as its jet mass or its transverse momentum. While we discard all soft subjets the recombined hard subjets define a trimmed (fat) jet. Just like filtering this reduces the effective size of the fat jet entering any kind of jet mass measurement. Because Λ_{trim} can be different for each fat jet the trimming procedure is self-adaptive: for a fat jet with large transverse momentum and/or mass the subjets need to have a larger transverse momentum to stay inside the trimmed jet. Just as the filtering procedure, trimming requires two input parameters.

3˙3. *Pruning.* – Unlike filtering or trimming, pruning [12] removes underlying event and pileup while building the jet, *i.e.* as part of the jet algorithm. In a first step it defines a fat jet which can be based on a sequential recombination algorithm. In a second step its constituents are pruned by checking in every recombination step $\min(p_{T,j_1}, p_{T,j_2})/p_{T,j_{1+2}} < z_{\text{prune}}$ and $\Delta R_{j_1,j_2} > R_{\text{prune}}$. If both conditions are met, the merging $j_1,\ j_2 \to j$ is vetoed. Just as filtering and trimming, pruning depends on two parameters: z_{prune} and R_{prune}. z_{prune} ensures that recombined well separated subjets are not very asymmetric in p_T. R_{prune} can be determined on a jet-by-jet basis.

Unlike filtering, pruning and trimming are self-adaptive procedures, applicable to a multi-jet final state in an unbiased resonance search.

In [13] it has been shown that pruning, trimming and filtering treat QCD jets differently while yielding a strong correlation in the reconstruction of electroweak scale resonances. Thus, by combining different grooming techniques we can improve the signal-to-background ratio in new physics searches.

4. – Phenomenological application of top taggers

The reconstruction of boosted hadronically decaying top quarks was one of the first applications of jet substructure methods in searches for new physics [14,15]. As discussed in sect. **2**˙1, in events where boosted top quarks arise from TeV scale resonances top taggers which make use of the substructure of large jets are necessary to discriminate top jets from QCD jets.

Many different approaches to tag boosted top quarks have been proposed [16]. In [2] it has been shown that they perform similarly on highly boosted top quarks. It is worth noting that it might be possible to combine different top tagging ideas.

One example of a top tagger is the so-called HEPTopTagger (Heidelberg-Eugene-Paris) [17]. The HEPTopTagger is designed to reconstruct top quarks which are only mildly boosted. To capture the decay products of tops with $p_{T,t} \sim 200\,\text{GeV}$ in one fat jet, it is necessary to increase its cone size, *e.g.*, $R = 1.5$. However, increasing the jet area poses two problems for the tagging algorithm. First, subjet combinatorics will increase and it will get more difficult to identify the top decay products. Second, ISR, UE and pileup will become a huge problem, so the HEPTopTagger includes a jet grooming stage.

The tagging algorithm proceeds along the following steps:

1. Un-doing the last clustering of the jet j the mass drop criterion $\min m_{j_i} < 0.8\, m_j$ determines if we keep j_1 and j_2. Subjets with $m_{j_i} < 30\,\text{GeV}$ are not considered, which eventually ends the iterative unclustering.
2. Apply a filtering stage to construct one three-subjet combination with a jet mass within $m_t \pm 25\,\text{GeV}$.
3. Order these three subjets by p_T. If their jet masses (m_{12}, m_{13}, m_{23}) satisfy one of the following three criteria, accept them as a top candidate:

$$\begin{aligned}
&0.2 < \arctan\frac{m_{13}}{m_{12}} < 1.3 \quad \text{and} \quad R_{\text{min}} < \frac{m_{23}}{m_{123}} < R_{\text{max}}\,, \\
&R_{\text{min}}^2\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\text{max}}^2\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > R_{\text{s}}\,, \\
&R_{\text{min}}^2\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\text{max}}^2\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > R_{\text{s}}\,.
\end{aligned}$$

4. For consistency, require the combined p_T of the three subjets to be above 200 GeV.

The dimensionless mass windows $R_{\text{min}} = 85\% \times m_W/m_t$ and $R_{\text{max}} = 115\% \times m_W/m_t$ are tunable and will be optimized by the experimental collaborations.

The HEPTopTagger has been used in the scenarios of sect. **2**˙2-**2**˙4:

Because scalar top partners can ameliorate the top quarks impact to the hierarchy problem of the Higgs, they are among the most anticipated particles to be found at the LHC. The HEPTopTagger was applied [17] to reconstruct the light top squark of the MSSM $\tilde{t}_1$ in a final-state with only jets and missing transverse energy, $pp \to \tilde{t}_1\tilde{t}_1 \to t\bar{t}\chi_1\bar{\chi}_1$. While a reconstruction with standard techniques yields $S/B \sim 1/7$, a subjet analysis in combination with mT2 can result in $S/B \sim 0.88$ and $S/\sqrt{B} \simeq 6$ after $10\,\text{fb}^{-1}$. If one of the tops decays leptonically a leptonic top tagger can be used [18] to separate the neutrinos MET contribution from the neutralinos MET, which allows an effective use of m_{T2}.

Recently, CDF [19] and D0 [20] measured an unexpectedly large forward-backward asymmetry of the top quarks. Measurements of this quantity are subtle at the LHC, due to its proton-proton initial state. However, one can define a forward/central charge asymmetry which captures the physics. Unfortunately, for the dominating gg initial state at the LHC there is no asymmetry at all. To enhance the subdominant $q\bar{q}$ and qg production processes it is beneficial to require a large invariant mass of the $t\bar{t}$ system, *i.e.* require boosted tops. By reconstructing the momentum of the hadronic top and measuring the charge of the second top's lepton, it is possible to count the number of tops and antitops in the forward or central region. This allows to measure the forward/central charge asymmetry at the LHC [21].

In early ATLAS and CMS reports the $t\bar{t}H$ production channel with subsequent Higgs decay to bottom quarks was one of the major discovery channels for a light Higgs boson. Further studies revealed a very poor signal-to-background ratio of 1/9 [22], making the channel very sensitive to systematic uncertainties which might prevent it from reaching a 5σ significance for any luminosity. However, at high transverse momentum, after reconstructing the boosted, hadronically decaying top quark using the HEPTopTagger as well as the Higgs boson with a modified version of the BDRS method, and requiring 3 b-tags, the signal-to-background ratio can be improved to $\sim 1/2$, while keeping the statistical significance at a similar value to that in ref. [23].

5. – Outlook

In this brief review we have given a categorization of new physics scenarios where searches using jet substructure methods can be beneficial over standard search strategies. Any machine probing the multi-TeV scale will produce electroweak scale resonances which will be highly boosted. In this context top tagging is one of the most prominent applications in searches for new physics. The HEPTopTagger is an example of a top tagger applicable in searches for mildly boosted tops. Currently, many different tagging and jet grooming methods are being evaluated on data to test their validity. Present results indicate a huge potential for new physics searches at the LHC in all discussed kinematic scenarios.

REFERENCES

[1] Butterworth J. M., Davison A. R., Rubin M. and Salam G. P., *Phys. Rev. Lett.*, **100** (2008) 242001.
[2] Abdesselam A. *et al.*, *Eur. Phys. J. C*, **71** (1661) 2011.
[3] Catani S., Dokshitzer Y. L., Seymour M. H. and Webber B. R., *Nucl. Phys. B*, **406** (1993) 187; Ellis S. D. and Soper D. E., *Phys. Rev. D*, **48** (1993) 3160.
[4] Dokshitzer Y. L., Leder G. D., Moretti S. and Webber B. R., *JHEP*, **08** (1997) 001; Wobisch M. and Wengler T., arXiv:hep-ph/9907280.
[5] Cacciari M., Salam G. P. and Soyez G., *JHEP*, **04** (2008) 063.
[6] Almeida L. G., Alon R. and Spannowsky M., arXiv:1110.3684 [hep-ph].
[7] ATLAS Collaboration, ATL-PHYS-PUB-2009-088.
[8] Cacciari M. and Salam G. P., *Phys. Lett. B*, **659** (2008) 119.
[9] Cacciari M., Salam G. P. and Sapeta S., *JHEP*, **04** (2010) 065.
[10] Dasgupta M., Magnea L. and Salam G. P., *JHEP*, **02** (2008) 055.
[11] Krohn D., Thaler J., Wang L.-T., *JHEP*, **02** (2010) 084.
[12] Ellis S. D., Vermilion C. K. and Walsh J. R., *Phys. Rev. D*, **80** (2009) 051501; Ellis S. D., Vermilion C. K. and Walsh J. R., *Phys. Rev. D*, **81** (094023) 2010.
[13] Soper D. E. and Spannowsky M., *JHEP*, **08** (2010) 029.
[14] Agashe K., Belyaev A., Krupovnickas T., Perez G. and Virzi J., *Phys. Rev. D*, **77** (2008) 015003; Lillie B., Randall L. and Wang L.-T., *JHEP*, **09** (2007) 074.
[15] Plehn T. and Spannowsky M., arXiv:1112.4441 [hep-ph].
[16] Brooijmans G., ATL-PHYS-CONF-2008-008; Thaler J. and Wang L.-T., *JHEP*, **07** (2008) 092; Kaplan D. E., Rehermann K., Schwartz M. D. and Tweedie B., *Phys. Rev. Lett.*, **101** (2008) 142001; Almeida L. G., Lee S. J., Perez G., Sung I. and Virzi J., *Phys. Rev. D*, **79** (2009) 074012.
[17] Plehn T., Spannowsky M., Takeuchi M. and Zerwas D., *JHEP*, **10** (2010) 078.
[18] Plehn T., Spannowsky M. and Takeuchi M., arXiv:1111.5034 [hep-ph].
[19] Aaltonen T. *et al.* (CDF Collaboration), *Phys. Rev. D*, **83** (2011) 112003.
[20] Abazov V. M. *et al.* (D0 Collaboration), *Phys. Rev. Lett.*, **100** (2008) 142002.
[21] Hewett J. L. *et al.*, *Phys. Rev. D*, **84** (2011) 054005.
[22] Cammin J. and Schumacher M., ATL-PHYS-2003-024.
[23] Plehn T., Salam G. P. and Spannowsky M., *Phys. Rev. Lett.*, **104** (2010) 111801.