Telescoping Jet Substructure

Yang-Ting Chien[a], Alex Emerman[b] and Shih-Chieh Hsu[c]

[a] Theoretical Division, T-2, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[b] Physics Department, Reed Colloge, Portland, OR 97202, USA
[c] Department of Physics, University of Washington, Seattle, WA 98195, USA

Telescoping jets (Tjets), which was proposed recently to systematically probe the hadronic event structure at multiple angular scales, is applied to boosted particle tagging by telescoping the subjets. More generally, it is important to realize that jet (substructure) observables and grooming techniques have artificial angular and/or energy scales in the definitions. Depending on the parameters chosen, different regions of phase spaces can be probed and information be extracted. Telescoping jet substructure systematically scans through the whole parameter space of interest and helps construct more powerful observables to discriminate the QCD backgrounds. We demonstrate in W tagging how the idea of telescoping jet substructure can be exploited.

As the Run 2 of the Large Hadron Collider (LHC) is going to start next year, with the center of mass energy of the collisions boosted to 13 TeV, we will be a step further to probe the physics of electroweak symmetry breaking and the physics beyond the standard model. At the energy scale much higher than the electroweak scale, even massive standard model particles (W/Z, higgs bosons and top quarks) that are produced are highly boosted, with their energies much larger than their rest masses. The hadronic decay products of a boosted particle are confined within a small angle around its momentum direction, resulting in a jet with substructure. On the other hand, hadronic signals usually have large QCD backgrounds. Being able to tag boosted particles more efficiently plays an important role not only in the search for new physics but also in understanding the higgs boson in the hadronic sector.

There have been many useful techniques developed for hadronic W/Z [?], higgs and top tagging, and some of the essential differences with the QCD backgrounds are used. For example, the electroweak decay processes at the hard scale give the subjet kinematics distinct from the QCD splittings. Also, W/Z and top jets consist mostly of quark subjets, whereas for QCD jets they usually have a fixture of quark and gluon subjets. So methods with better quark gluon discrimination are helpful. Moreover, the color structure among the subjets is different between signals and backgrounds: W/Z and higgs are color singlet therefore not as color connected to the rest of the event as the QCD backgrounds.

Boosted particle tagging is further complicated with the presence of initial state radiation, underlying events and pileup. Much effort has also been made in developing jet grooming techniques to eliminate radiation contaminations that are more likely to be soft. Some common feature involves getting rid of soft and/or wide angle radiation. Although jet grooming aims at removing radiation contaminations, it does provide an operation on jets that signals and backgrounds may react differently.

Whether it is a jet (substructure) observable, a grooming procedure, a boosted particle tagger as a whole, or an analysis method in general, there are usually artificial free parameters in the definition. For example, a jet is conventionally defined as the set of particles moving within an angular scale R around certain direction of dominant energy flow, or more practically the outcome from running a jet algorithm with a parameter R. In jet pruning the parameters D_{cut} and z_{cut} define how wide angle between each pair of merged particles and how soft a particle should be to prune it away. In these examples, R, D_{cut} and z_{cut} are all free parameters, and there is no a priori reason for favoring some from the others. Nevertheless, the parameters are conventionally chosen to optimize an analysis.

Recently, Qjets was proposed with non-deterministic jet clustering. The probabilistic nature allows a single event or jet to be (re)clustered multiple times with different clustering histories. Each observable constructed from different trees of a single event or jet becomes a distribution. The information from the different interpretations of a single event or jet is combined with a more sophisticated statistical method, and the observable volatility $\mathcal{V}$ is constructed which gives good background discrimination power.

Soon after, telescoping jets was proposed which probed the hadronic event structure with multiple R's in the jet definition. A clear physical picture emerges which reveals the importance of the information contained in the wide angle radiation from the jet axes in an event. Moreover, telescoping jets provides a systematic way to extract more information out of each event at multiple angular scales. This leads to a significant improvement in the search for the associated production of a higgs and a leptonic Z in the $h \to b\bar{b}$ channel.

Telescoping jets scans through the artificial parameter R in the jet definition. Each choice of R defines a specific region of phase space to study. More generally, each set of parameters used in an analysis method defines the angular or the energy scale at which an event or jet is probed. By telescoping away from the optimal choice of parameters, we gain extra information about the event or jet (sub)structure at a broader range of scales.

In this paper we apply telescoping jets to boosted particle tagging. We demonstrate in W tagging how the idea can be exploited, and we will leave the applications to higgs and top tagging, as well as new physics searches for our next paper.

The particles from a boosted particle decay are mostly confined within a small angle. The common starting point of tagging a particle is to reconstruct a fat jet with a relatively large R in the jet algorithm so that almost all the relevant radiation is captured. From here on the fat jet is the object of interest to know whether it is initiated from a boosted heavy particle decay or a pure QCD jet. A straightforward application of Tjets is to recluster the fat jet multiple times with different $R_{\rm sub}$'s, which is referred to as **telescoping reclustering**.

In the context of W tagging, a similar idea was previously explored in [**?**] and the authors looked at the mass m and the transverse momentum p_T of the hardest subjet in each reclustering. For W jets, with a small R the leading subjet picks out the radiation initiated from one of the quarks decay from the W. As the value of R becomes large, the hardest subjet can actually include most of the radiation from the W. By scanning through R and looking at the R-cores, some information about the energy profile of the jet can be gained, which is a mixture of the hard scale subjet kinematics and the soft-collinear radiation pattern at a lower scale. The analytic properties of the R-cores are different, and they give an interpolation between the two ends. However, we would like to factorize the information from the hard, electroweak and the soft, QCD scales.

To reveal the jet substructure, we can improve the telescoping reclustering by first identifying the dominant energy flows within the fat jet. For example, for W jets we can recluster the fat jet using the anti-k_T algorithm with a "suitable" R and determine the axes of the two hardest subjets from the reconstructed subjet cores. The proper value of R used should be smaller than the characteristic angular separation $2m_W/p_T$ between subjets. We can also use the two subjettiness axes with a specific choice of β in the definition. Then we telescope around the predetermined subjet axes with multiple R's. This is referred to as **telescoping subjets**. For top tagging with possibly a three-prone jet substructure for the signals, we telescope around the three subjettiness axes as an example.

These are the original idea of Tjets directly applied to boosted particle tagging. However, with the generalized telescoping jets idea, which scans through the artificial parameters used in any analysis method, we can try a lot of interesting extensions. For example, the $D_{\rm cut}$ in pruning sets the angular scale of the separation between particle after the procedure. We can do **telescoping pruning** by telescoping the $D_{\rm cut}$ around the conventional scale $m_{\rm jet}/p_{T\rm jet}$. In trimming, the $f_{\rm cut}$ determines how soft a subjet should be to be trimmed away. We can do **telescoping trimming** by telescoping the $f_{\rm cut}$. In subjettiness, the parameter β determines the weight of the angles between the subjettiness axes and the particles in the subjettiness definition, and it sets the angular scale of probing the jet substructure. We can do **telescoping subjettiness** by telescoping the β. Now, you can see that we can do telescoping X, and X can be any technique.

You may wonder what is the point here. When optimizing the parameters in a conventional analysis, a scan through the parameter space is already the standard procedure. However, instead of choosing the optimal set of parameters with a specific criterium, combining all the information will give a significant improvement in the analysis. We will use volatility as a simple evaluation of how different the information is extracted with different set of parameters. However, it is straightforward to apply the standard multivariate analysis method to process the full information. Another subtlety in telescoping jets is the range of parameter to scan through, similar to the rigidity parameter to choose in Qjets. We will discuss this point in the rest of the paper when we demonstrate specifically how various versions of Tjets is implemented in W tagging.

[] Y. Cui, Z. Han, and M. D. Schwartz, Phys.Rev. **D83**, 074023 (2011), 1012.2077.