EPJ Web of Conferences **90**, 10001 (2015)
DOI: 10.1051/epjconf/20159010001

Toward the development of more robust jet definitions

Zack Sullivan[1,a] and Daniel Duffty[1,b]

[1] *Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616-3793, USA*

Abstract. When the LHC restarts at 13–14 TeV the luminosity will increase to 50–100 interactions per crossing. Under existing jet definitions, this will lead to multiple fake jets built up out of pileup, and poor energy resolution requiring large cuts on measured transverse momentum. In order to recover access to low energy jets we determine the minimal requirements for sensible jet definitions. We then propose the first of a new class of jet algorithms, called "p-jets," that will be less sensitive to the noise of multiple interactions and underlying event.

1 Introduction

The first runs of the Large Hadron Collider (LHC) at 7 TeV and 8 TeV ran at moderate luminosities and were very successful. As the next runs at 13–14 TeV begin, the instantaneous number of pp interactions per crossing will increase from the average of 20 at 8 TeV to 50–100 at 13 TeV. As a consequence, most reconstructed jets under current algorithms will be nothing but combinations of random particles from different interactions that accidentally passed clustering criteria — otherwise known as pileup jets. In order to avoid this problem, existing jet algorithms will have to change.

During the Community Summer Study of 2013, experimentalists from ATLAS and CMS joined with theorists to examine the effects of pileup on various Standard Model processes [1]. We investigated the reconstruction of W bosons using the anti-k_T algorithm and various pileup subtraction schemes. Among the conclusions was that the dijet mass resolution using anti-k_T jets will be insufficient to identify a W boson mass peak without significant cuts on jet energy [1]. This is evident in Fig. 5a from the working group paper (reproduced in Fig. 1 below) where above 100 interactions per crossing the dijet signal of the W mass has no clear peak, even with no background.

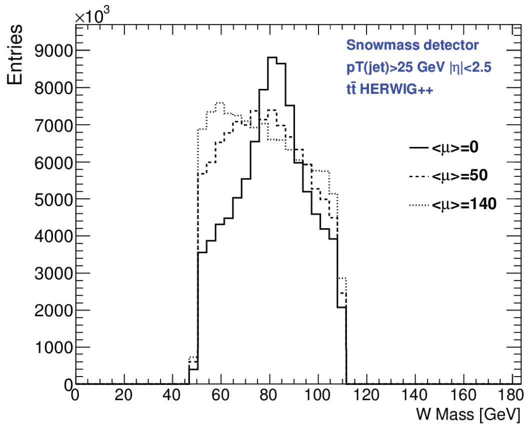

Figure 1. Fig. 5a from Ref. [1] shows the dijet invariant mass for a pure W boson decay in the presence of $<\mu>$ = 0, 50, or 140 interactions per crossing using an anti-k_T algorithm.

The problem extends to all low energy jets and, hence, one of the conclusions of the "Top algorithms and reconstruction" working group was that high luminosity runs of the LHC would at least double the jet energy uncertainty below 100 GeV. In response, there has been a push to raise jet energy thresholds to 100 GeV to avoid this troublesome region. Unfortunately, at this level we would lose access to the production thresholds of all heavy Standard Model particles (Higgs, top quarks, W, and Z).

At the XLIV International Symposium on Multiparticle Dynamics we presented a poster proposing a new class of generalized noise-tolerant priority-based jet algorithms ("p-jets") that will provide stable reconstruction of jets for W and Z thresholds in the high pileup environments about to appear at the LHC.

2 Priority based jets "p-jets"

Historically there have been several jet algorithms used in collider experiments [2–6], where measured calorimeter towers are successively clustered into jets as long as

$$\min(p_{Ti}^{2k}, p_{Tj}^{2k})\Delta R_{ij}^2/D^2 < p_{Ti}^{2k}\,, \tag{1}$$

where p_{Ti} is the transverse momentum of the ith cluster, ΔR_{ij} is a distance in azimuthal angle–rapidity space, and D is a cutoff. The Cambridge-Aachen ($k = 0$) variation forms the basis of most jet substructure [7–9] and boosted jet algorithms [10]. These boosted objects are most useful

[a] e-mail: Zack.Sullivan@IIT.edu
[b] e-mail: dduffty@iit.edu

for high energy $p_T > 300$–500 GeV jets, but are restricted to measuring the small cross section tails of the Standard Model.

For lower energy jets the LHC experiments have settled on the "anti-k_T" ($k = -1$) variation. While relatively simple to calibrate, anti-k_T algorithms will not perform well at low energies [1]. In addition, they do not match well with onto actual QCD radiation spectra: namely, anti-k_T jets become broader as the energy increases, while real QCD radiation becomes more narrowly concentrated with increasing energy due to the boost of the decaying particle.

Our goal is to develop new classes of jet definitions that retain the core of the QCD radiation emission (a jet) while throwing away the tails that are dominated by pileup. We require any new jet algorithm to satisfy several conditions:

1. It must be infrared safe, and should match on to theoretical calculations in a manner that matches the shape of the QCD radiation spectrum. The anti-k_T algorithm fails in this second condition, as the size of anti-k_T jets quickly grows in η–ϕ space with energy and generally saturates its maximum cutoff D, while physical boosted jets become narrower.

2. It must treat experimental objects in the same manner as in theoretical calculations so that jet matching makes sense.

3. We must be able to calibrate the jets in different environments, and demonstrate an improvement at high luminosities compared to existing anti-k_T algorithms.

4. Critically, this has to work on an event-by-event basis, and not rely on large statistics to cancel misreconstructions due to fluctuations in the pileup absorbed into the jet.

The primary goal of a jet algorithm is to absorb the final state collinear singularities into an extensive object. The need for this can be seen in the structure of a three-jet final state in relation to the two-jet final state,

$$\frac{\sigma_{ijk}}{\sigma_{ij}} = \int \frac{x_i^2 + x_j^2}{(1-x_i)(1-x_j)} \text{ with } x_i = \frac{2E_i}{\sqrt{s}} . \tag{2}$$

This ratio is proportional to an integral involving the angle between particles i and j of the form

$$\int \frac{1}{1-\cos\theta_{ij}} . \tag{3}$$

The limit $x_i \to 1$ (or $x_j \to 1$) is then equivalent to the limit in which the angle $\theta_{ij} \to 0$, where the ratio of the cross sections diverges. The key feature of a jet definition is to soften this singularity. We observe that we can accomplish this by multiplying by *any* function that contains a power of the angle θ^n with $n > 1$,

$$\int \frac{1}{1-\cos\theta_{ij}} \times \theta^n . \tag{4}$$

If we satisfy this condition, then as long as we have a unique algorithm, we can match onto a theoretical calculation in the small angle limit with our algorithm.

We introduce what we call the hard-edge priority-based jet algorithm (hard "p-jets") as a generalized approach to clustering objects into jets. In order to match the shape of the QCD radiation spectrum, we create a distance-dependent threshold function that must be overcome for clustering to occur. Energy far away from the core of the jet will be suppressed — thereby avoiding absorption of the nearly pure noise due to pileup that dominates the edges of the jet.

The threshold function is set by comparing pair-wise objects (towers or tracks or particles) with an objective function satisfying our general condition. In order to match the leading dipole of the QCD radiation spectrum we first focus on the function

$$A \sin^2\left(\frac{\pi}{4}\frac{\Delta R}{R_{\max}}\right) , \tag{5}$$

where A is the highest transverse energy of the two objects under consideration, and ΔR is the standard distance in η–ϕ space ($(\Delta R)^2 = (\Delta\eta)^2 + (\Delta\phi)^2$). If the lower transverse energy exceeds the threshold then the four-vectors are combined into a new cluster.

Our adaptive jet algorithm consists of first identifying the fraction of energy of the lower transverse energy object that exceeds the threshold function for every pair of objects in the detector. The pair with the highest fractional energy above threshold is given the highest *priority* and is absorbed. (The name "priority-based jets" comes from both this priority ordering and to the setting of the threshold based on the highest energy object.) After the leading pair is combined into a new cluster, we recalculate all priorities and iterate until all thresholds are above nearby clusters. At the end of the process we are left with a unique list of "jets." This procedure can be seen in the cartoon diagram of Fig. 2.

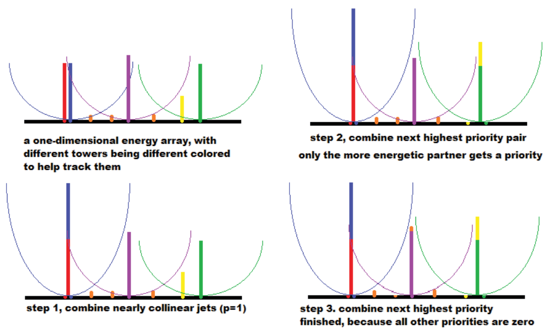

Figure 2. Cartoon of the "p-jet" algorithm. The highest priority objects are combined in step 1 and their threshold curve is recalculated. The new highest priority pair is combined in step 2, and the process is iterated until all energy clusters are below prioritized thresholds.

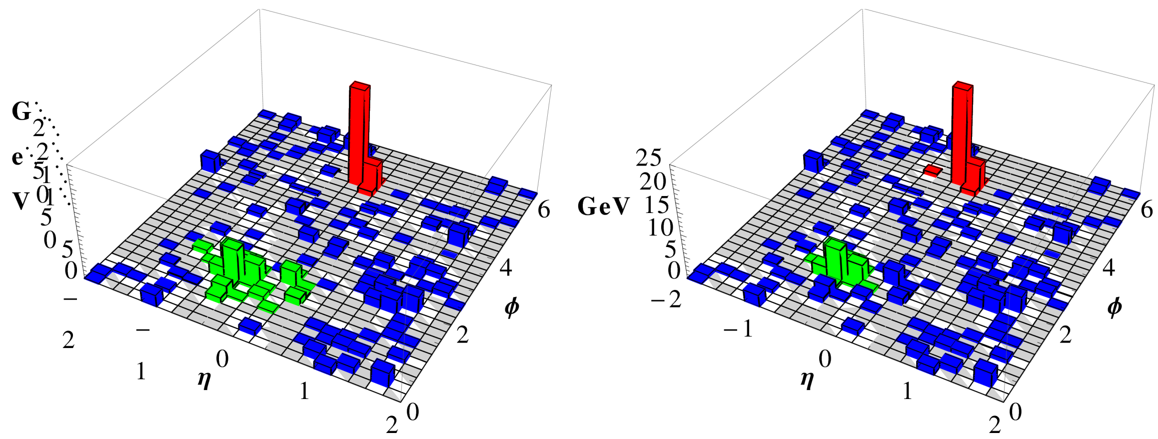

Figure 3. Towers included in the dijet reconstruction of a Z boson in a single interaction per crossing using (left) "p-jets" with $R_{\max} = 1$, and (right) anti-k_T jets with $D = 0.5$.

3 Preliminary comparison to anti-k_T

While our primary goal is to introduce the new generalized class of priority-based jet definitions we call "p-jets," we also must demonstrate that these new algorithms are at least as successful as the existing algorithms under favorable conditions. In Fig. 3 we show lego plots for the reconstruction of a Z boson in a single interaction per crossing for p-jets (left) and anti-k_T jets (right). The "p-jet" reconstruction uses $R_{\max} = 1$, and follows the natural spread energy for each energy cluster — the high energy jet (red) is narrow, while the lower energy jet (green) is broader. Anti-k_T jets with $D = 0.5$, in contrast, carve out similar size objects irrespective of the actual particle flow.

All jet algorithms reconstruct a fraction of the energy contained within the region of the QCD radiation spectrum we identify as a jet, and hence we must apply jet-energy corrections (JEC). These corrections are tiny for p-jets by design — namely the dipole radiation inspired threshold function we use captures most of the jet energy without the surrounding noise. In Fig. 4 we compare the reconstruction of the dijet invariant mass peak for Z decay between p-jets, anti-k_T, and k_T jets before and after a jet-energy correction tuned on Z (to dilepton) + jets samples. We can see that p-jets perform similarly to anti-k_T with no additional pileup, and both reproduce a Z mass peak.

This new class of priority-based jet algorithms is aimed at clustering the core of the primary QCD radiation and excluding the noise surrounding the jet. Anti-k_T jets do not distinguish between any radiation within their cutoff, and in high pileup will absorb all objects regardless of source. We have begun to examine the performance of both algorithms in the presence of 50 minimum-bias events of pileup. Fig. 5 illustrates the nature of the problem, where several calorimeter towers can grow to larger energies than real Z decay purely from statistical fluctuations. Nevertheless, the p-jet algorithm can identify the clusters corresponding to the jets from Z decay. Which towers are clustered into jets differs between anti-k_T (which produces round jets) and p-jets which follows the flow of radiation. The next stage of our analysis involves recalibration in high-luminosity environments, and demonstration an improvement in Z reconstruction.

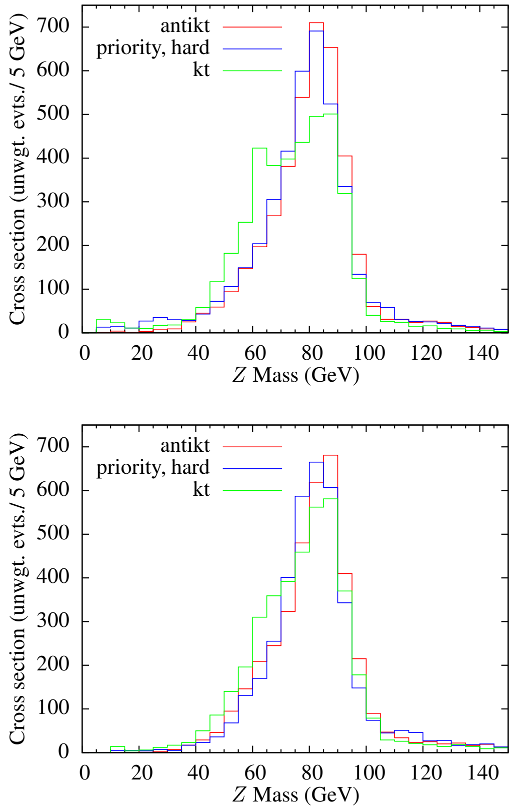

Figure 4. Reconstruction of the Z mass in a dijet sample before (top) and after (bottom) jet energy corrections. P-jets have very small energy corrections, and perform similarly to anti-k_T jets in this simple scenario.

4 Conclusions

Pileup will be a serious problem for the reconstruction of low energy jets in the next run of the LHC. One solu-

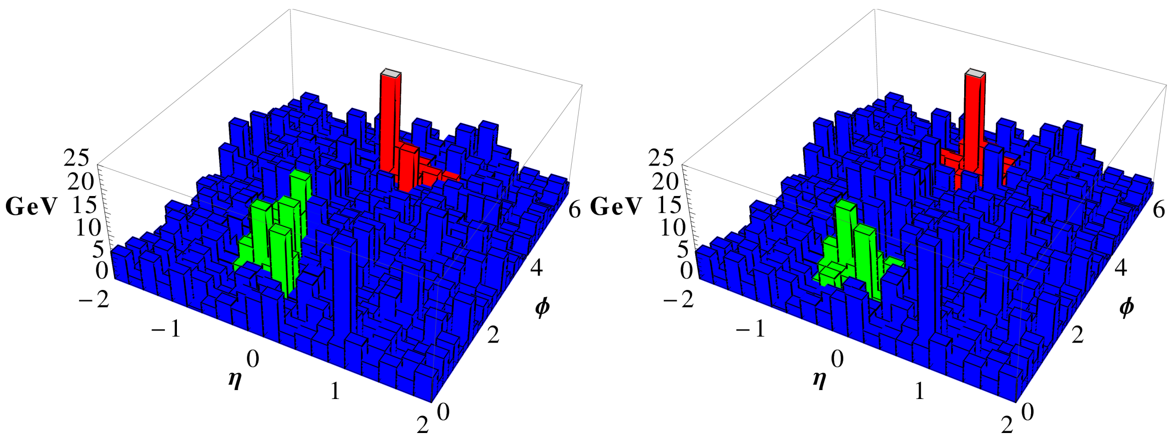

Figure 5. Please write your figure caption here

tion is to ignore low energy jets by performing measurements only on boosted objects, e.g., with energies above 100 GeV. Unfortunately, theoretical predictions of cross sections in the tails of distributions are less well controlled, and so precision measurements of phenomena will require us to look in the low energy regime. In addition, low energy jets play an important role in cascade decays of beyond the Standard Model particles.

In a soon to be released paper [11] we will show that "p-jets" dramatically outperform anti-k_T jets in high pileup scenarios for low energy jets. "P-jets" require significantly smaller jet energy corrections, and provide more robust measurements than anti-k_T jets on an event-by-event basis. With the addition of tracking information we can identify the correct pairing of jets to reconstruct resonant Z production at the LHC in 50 interactions per crossing. "P-jets" herald the beginning of a new era of jet reconstruction in an age of high luminosity.

References

[1] R. Calkins, *et al.*, "Reconstructing top quarks at the upgraded LHC and at future accelerators," in *Planning the Future of U.S. Particle Physics: Report of the 2013 Community Summer Study of the APS Division of Particles and Fields*, edited by Norman A. Graf, Michael E. Peskin, and Jonathan L. Rosner, eConf C1307292, SNOW13-00075 (2014) [arXiv:1307.6908 [hep-ph]].

[2] S. D. Ellis, Z. Kunszt, and D. E. Soper, Phys. Rev. Lett. **62**, 726 (1989).

[3] J. E. Huth *et al.*, in *Proceedings of the Snowmass Summer Study: Research Directions for the Decade*, Snowmass, June 25 – July 13, 1990, p. 134, FNAL-C-90-249-E.

[4] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993) [arXiv:hep-ph/9305266].

[5] M. G. Albrow *et al.* (TeV4LHC QCD Working Group), in *Final Report of the Tev4LHC Workshop*, Fermilab, 2004–2005, edited by M. Carena and S. Mrenna (Fermilab, Batavia, 2006) [arXiv:hep-ph/0610012].

[6] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. **0804**, 063 (2008) [arXiv:0802.1189 [hep-ph]].

[7] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008) [arXiv:0802.2470 [hep-ph]].

[8] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys. Rev. D **80**, 051501 (2009) [arXiv:0903.5081 [hep-ph]].

[9] D. Krohn, J. Thaler, and L. -T. Wang, J. High Energy Phys. **1002**, 084 (2010) [arXiv:0912.1342 [hep-ph]].

[10] A. Abdesselam *et al.*, Eur. Phys. J. C **71**, 1661 (2011) [arXiv:1012.5412 [hep-ph]].

[11] Daniel Duffty and Zack Sullivan, *in production*.