MIT-CTP 4952

Telescoping Jet Substructure

Yang-Ting Chien[a,d],* Alex Emerman[c,d,e], Shih-Chieh Hsu[b], Samuel Meehan[b], and Zachary Montague[a,b]
[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[b] *Department of Physics, University of Washington, Seattle, WA 98195*
[c] *Department of Physics, Columbia University, New York, NY 10027*
[d] *Theoretical Division, T-2, Los Alamos National Laboratory, Los Alamos, NM 87545*
[e] *Physics Department, Reed College, Portland, OR 97202*

(Dated: November 10, 2017)

We introduce a novel jet substructure calculus, called telescoping jet deconstruction, which exploits the variation of observables with respect to a sampling of phase-space boundaries quantified by the variability. We apply this method to identify boosted W boson and top quark jets using telescoping jet grooming and telescoping subjets, demonstrating its ability to disentangle the information coming from the subjet topology and that coming from the subjet substructure. We find excellent performance of the variability measure, in particular its robustness against finite detector resolution. This method provides a new direction in heavy particle tagging and enables a more complete and systematic approach to the decomposition of jet substructure.

The Large Hadron Collider (LHC) has begun to regularly probe physics above the electroweak scale, where the momenta of massive Standard Model particles are much larger than their invariant masses, resulting in hadronic decays of jets with prong-like substructures. Many jet substructure variables have been designed [1–3] and combined using multivariate techniques [4–7] to identify such jets and increase the sensitivity to beyond the Standard Model physics. However, the ability to accurately reconstruct the features of such jets is obscured by the presence of additional proton-proton interactions (i.e. pileup) as well as the underlying event of the hard collision, both of which cause for additional radiation to fall within the catchment area of the jet. Often, this radiation is removed through one of a number of grooming procedures, i.e. pruning [8] or trimming [9]. These observables and grooming procedures target certain intuitive features of the radiation properties and often have tuneable parameters. For example, in the pruning algorithm, the z_{cut} and D_{cut} parameters control the softness and noncollinearity of a discarded particle. Conventionally, one makes a single choice of parameters deemed optimal by some metric. However, such a choice may neglect the full information the entire observable class contains.

Recently, the Q-jets technique [10] introduced nondeterminism in jet clustering. The procedure probes each jet multiple times and quantifies differences among pruned jets using the variability in the mass of the resulting ensemble of pruned jets. This concept can be refined by examining the radiation pattern surrounding the dominant energy flow with multiple angular resolutions $\{R_i\}$, one can extract the full information contained in jets at all angular scales, analogous to that investigated in [11]. In this Letter we apply this technique, called telescoping jet deconstruction to analyze a set of commonly used jet observables and grooming procedures. We demonstrate the feasibility of the method as applied to the identification of hadronically decaying W bosons and top quarks. Analogous to the Q-jets procedure, where the the mass volatility was used as the final observable, the variability of each observable, under the variation of the parameters of the given algorithm, is used as the fundamental observable.

In hadronic boosted two-body resonance decays, such as a W boson, the resonance mass M introduces a two-prong structure in the jet at an angular scale $\Theta \approx 2M/p_T$ between the two prongs, where p_T is the transverse momentum of the heavy particle. On the other hand, QCD jets initiated by isolated quarks and gluons have such distinct scale. However, when examining jets whose mass is near $M \pm \Delta m$, QCD jets are also two-prong-like but with a more distended radiation pattern when $\Delta m \gg \Gamma$ where Γ is the natural width of the resonance. Besides this nontrivial *subjet topology*, the strong interaction dictates the formation of subjets with *subjet substructures* and *subjet superstructures* [12] which are sensitive to the partonic origins of subjets.

In the case of boosted top quarks, the top mass (M_t) and the W mass (M_W) are not hierarchically separated **SAM : What does hierarchically separated mean? It is not common jargon in my mind.**. Therefore $\Theta_t \approx 2M_t/p_T^t$ and $\Theta_W \approx 2M_W/p_T^W$ can be comparable. This results in the generic three-prong structure in the hadronic top decay $t \to W + b \to q_1 + q_2 + b$. However, when examining jets with a mass near $M_t \pm \Delta m$ the selected QCD jets are again two-prong-like and so observables which distinguish three-prong jets from two-prong jets will help discriminate the such jets from true top quark jets.

Given an arbitrary jet observable $\mathcal{O}$ with a parameter a, the variation of the observable with respect to the sampling of parameters $\{a_i\}$ within $(a_{\min}, a_{\max})$, or the *variability*, is quantified by the coefficient of variation $v_{\mathcal{O}}^a$ defined as the ratio of the standard deviation and the mean of $\{\mathcal{O}_{a_i}\}$,

$$v^a_{\mathcal{O}} = \frac{\sigma(\mathcal{O}_{a_i})}{\langle \mathcal{O}_{a_i} \rangle} \ . \quad (1)$$

Variations with respect to multiple varied parameters can be studied using the variability matrix. Much like the first derivative in calculus, the variability $v^a_{\mathcal{O}}$ measures the change of the observable $\mathcal{O}$ with respect to the change of the phase space boundary set by the parameter a. Instead of combining observables with different parameters in a multivariate analysis, the variability can give a trend of the observable variation which itself can be used as a distinguishing feature of signal and background jets.

We consider a variety of telescoping deconstruction applications. We focus on the variability of the jet mass with respect to varying the parameters which determine the jet constituents contributing to the jet mass. The sampling of the telescoping parameters is uniform within the range $(a_{\min}, a_{\max})$.

Telescoping pruning: Using the k_T reclustering algorithm, pruning discards soft and noncollinear particles when merging particles i and j if the combination is both soft and wide-angled,

$$\begin{aligned} &\frac{\min(p_{T_i}, p_{T_j})}{|p_{T_i} + p_{T_j}|} < z_{\text{cut}} && \text{(soft)} \\ &\Delta R_{ij} > D_{\text{cut}} \ , && \text{(noncollinear)} \end{aligned} \quad (2)$$

where p_{T_i} are the particle transverse momenta and $\Delta R_{ij} = \sqrt{\Delta y^2_{ij} + \Delta\phi^2_{ij}}$ is the distance between the particles i and j rapidity y and azimuthal angle ϕ. We fix $z_{\text{cut}} = 0.1$ and construct v_{prun}, the variability of the pruned jet mass with the telescoping parameter $a \in (0.1, 2.0)$ in $D_{\text{cut}} = a\ 2m_{\text{jet}}/p_{T_{\text{jet}}}$.

Telescoping trimming: Trimming [9] reclusters jets into subjets using the k_T algorithm with subjet radius R_{sub}. The subjet i is discarded if it is soft, i.e.

$$p_{T_i} < f_{\text{cut}}\ p_{T_{\text{jet}}} \ . \quad \text{(soft)} \quad (3)$$

Here p_{T_i} is the transverse momentum of the i^{th}subjet. We construct v_{trim}, the variability of the trimmed jet mass with the telescoping parameter $a = f_{\text{cut}} \in (0.0, 0.1)$.

Telescoping subjets: N subjets are exclusively reconstructed around dominant energy flows within jets. A similar method using the leading subjets in a reclustered jet was explored in [13]. We choose the subjet axes as the N-subjettiness axes [14] with $\beta = 1$, and build subjets around them with radius R_T [15–18]. Particles are assigned to the nearest axis according to the distance ΔR_{ij} between the axis $\hat{n}_i$ and the particle j,

$$\text{subjet}_i = \{p_j \mid \Delta R_{ij} < R_T \text{ and } \Delta R_{ij} < \Delta R_{kj}\ ,\ \forall k \neq i\}, \quad (4)$$

where k is the index of the other axes $\hat{n}_k$. The variability v_N of the invariant masses of the sum of N subjets is reconstructed with the telescoping parameter $a = R_T \in (0.1, 1.0) \times R$. Note that $a_{\max}$ is chosen to identically be the jet radius R to scan through the entire catchment area of the jet. On the other extreme, the dominant energy features will be lost if a is too small and so $a_{\min}$ is chosen as $0.1 \times R$. We further focus only on N =2 and 3 in W and top tagging, respectively but this could be extended in the case of more exotic boosted topologies [**?**]. The angular directions of the subjet axes encode information about the subjet topology. For instance, in W tagging, the subjet topology is affected by the jet mass cut, but W and QCD jets can still have significantly different distributions for the angle θ_2 between the two prongs. For top tagging with $N = 3$, we consider the minimal angle $\theta_{\min}$ among the subjet axes. For QCD jets, this angle is expected to be small while for top jets where two of the three prongs has been identified, it will be distributed away from zero. We attempt to identify the W inside the top jet [19, 20] by considering m_{W2}, the invariant mass of two of the three exclusive voronoi regions closest to the W mass, and its variability $v_{m_{W2}}$ by scanning within those two regions.

The study is performed using samples generated from the Monte Carlo simulations of proton-proton collisions at $\sqrt{s} = 13$ TeV using Pythia8 [21]. Particles are clustered into jets with FastJet 3 [22] using the anti-k_T algorithm [23] with $R = 1.0$ and are required to be central with a pseudorapidity $|\eta| < 1.2$. We consider two kineamtic regimes, where the jet p_T is between 350 GeV and 500 GeV or 800 GeV and 1 TeV. Signal W boson and top quark jets are generated using decays of heavy Kaluza-Klein gravitons with an invariant mass at 1 or 2 TeV for the two p_T bins in $gg \to G^* \to W^+W^-$ or $t\bar{t} \to$ hadrons. Background QCD jet are generated from the standard model dijet process. To study the impact of finite detector resolution, we compare the results with the particles clustered in pseudo-calorimeter (η, ϕ) cells of size 0.1×0.1, with each cell momentum constructed with zero mass and direction from the primary vertex. **SAM : Is there actually pileup overlaid?** Finally, in the case of telescoping subjets, jets are groomed using the trimming algorithm with $R_{\text{sub}} = 0.3$ and $f_{\text{cut}} = 0.05$ to remove the effects of pileup and a selection on the trimmed jet mass is made between 70 GeV and 90 GeV for W tagging and between 160 GeV and 190 GeV for top tagging.

To examine the complementarity of the information contained in the telescoping subjet variables, subsets of them are inputs for Boosted Decision Trees (BDTs) implemented in TMVA [24]. For top tagging we also consider the ratio v_{32} between v_2 and v_3,

$$v_{32} = \frac{v_3}{v_2} \ . \quad (5)$$

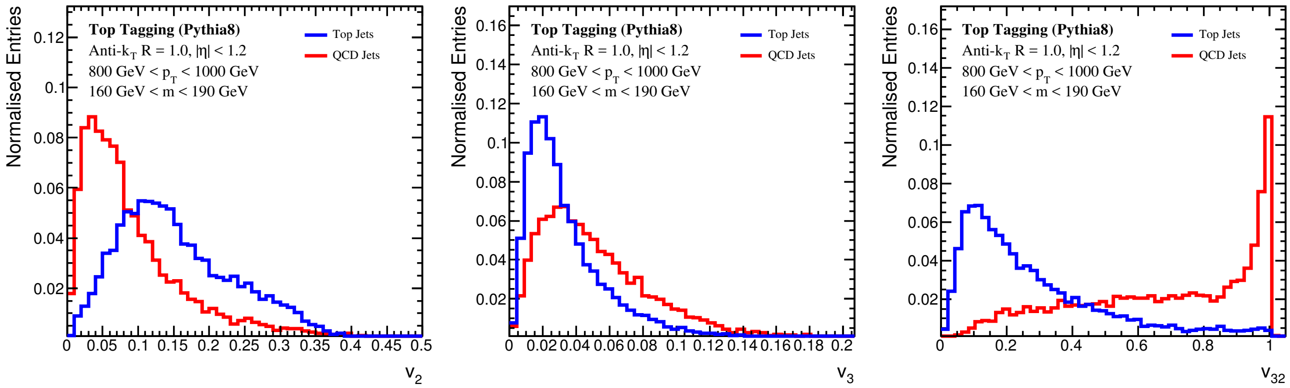

FIG. 1. The distributions of the variabilities v_2 (left panel) and v_3 (middle panel), as well as their ratio v_{32} (right panel) for top and QCD jets with 800 GeV $< p_T <$ 1 TeV and 160 GeV $< m <$ 190 GeV using the truth particle information.

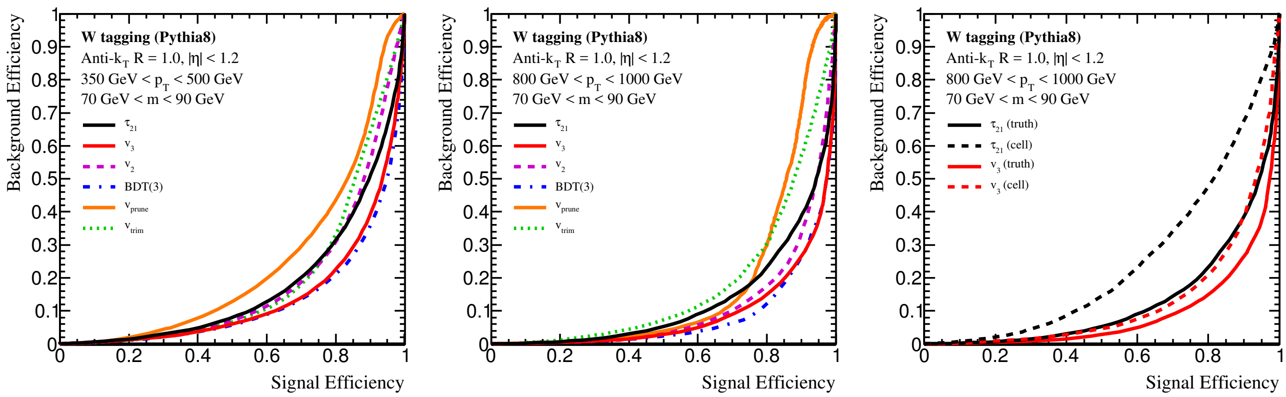

FIG. 2. The W tagging ROC curves of the variabilities v_2, v_3, $v_{\rm trim}$, and $v_{\rm prun}$, the BDT combinations of three telescoping subjets variables $\{v_2, v_3, \theta_2\}$, and the two-prong tagger $\tau_{21} = \tau_2/\tau_1$, in the $(300\ \text{GeV}, 500\ \text{GeV})$ jet p_T bin (left panel) and the $(800\ \text{GeV}, 1\ \text{TeV})$ bin (middle panel). Right panel: ROC curves of v_3 and τ_{21} in the $(800\ \text{GeV}, 1\ \text{TeV})$ jet p_T bin. Solid curves correspond to the ones with the truth particle information, and the dashed curves are the ones using the pseudo-calorimeter cell particle information.

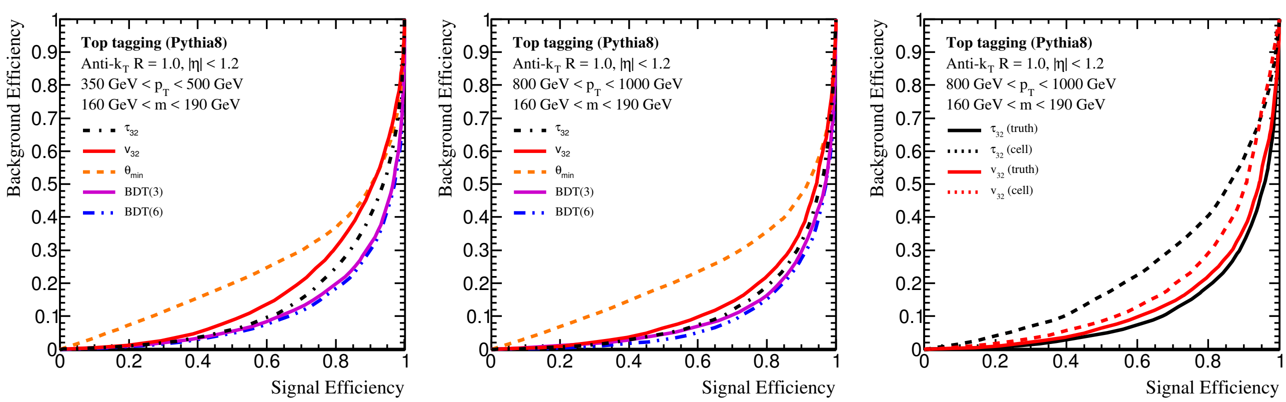

FIG. 3. The top tagging ROC curves of the variability ratio v_{32}, the minimal angle among three subjets $\theta_{\rm min}$, the BDT combinations of three and six telescoping subjets variables $\{m_{W2}, v_2, v_3\}$ and $\{\theta_2, \theta_{\rm min}, m_{W2}, v_2, v_3, v_{m_{W2}}\}$, and the three-prong tagger $\tau_{32} = \tau_3/\tau_2$, in the $(300\ \text{GeV}, 500\ \text{GeV})$ jet p_T bin (left panel) and the $(800\ \text{GeV}, 1\ \text{TeV})$ bin (middle panel). Right panel: ROC curves of v_{32} and τ_{32} in the $(800\ \text{GeV}, 1\ \text{TeV})$ jet p_T bin. Solid curves correspond to the ones with the truth particle information, and the dashed curves are the ones using the pseudo-calorimeter cell particle information.

Shown in Figure 1 are the distributions of v_2, v_3, and v_{32} for top and QCD jets. We find that top jets have a broader v_2 distribution and a narrower v_3 distribution. The large variation of the jet mass when telescoping around the two subjet axes is caused by the transition of the W from being partially reconstructed to fully reconstructed. There is not an intrinsic mass scale dictating the third hard emission for QCD jets. On the other hand, the three prongs inside top jets are quark-initiated subjets, whereas the subjets in QCD jets can have gluonic origins. Quark subjets are narrower than gluon subjets; therefore v_3 of top jets is statistically smaller. v_{32} has almost the same performance as the BDT with input $\{v_2, v_3\}$, suggesting that v_{32} may be the optimal way of combining the two variabilities. **SAM : The ideas being expressed here are jumbled and I cannot sort them out. You need to try again since it comes off as just a laundry list of disjoint observations.**

An interesting feature of v_{32} is that it cuts off naturally and sharply at 1, most clearly seen in QCD jets. Crucially, $v_3 \le v_2$. The two-prong structure in QCD jets implies that v_2 and v_3 collect almost the same information. The third energy flow axis can not be displaced far from the two axes determined at $N = 2$. Hence, little new information is collected by constructing a third subjet and the distribution of v_{32} for QCD jets peaks at 1. In the case where there is a third, semi-hard emission, the emission is captured by all telescoping subjets at $N = 3$ and does not induce the observable variation and so $v_3 < v_2$. In general, for larger N, more particles are and so the variability is expected to decrease $(v_{N+1} \le v_N)$.

The performances of the observables can be illustrated by receiver operating characteristic (ROC) curves, plotting the background efficiency as a function of the signal efficiency, where a lower curve indicates a better tagging performance. Shown in Figure 2 are the ROC curves of v_2, v_3, $v_{\rm trim}$, $v_{\rm prun}$, the BDT combinations of the telescoping subjet variables $\{v_2, v_3, \theta_2\}$, and the two-prong tagger $\tau_{21} = \tau_2/\tau_1$ in W tagging. The left and middle panels correspond respectively to two jet p_T regions of $(350\ \text{GeV}, 500\ \text{GeV})$ and $(800\ \text{GeV}, 1\ \text{TeV})$. Overall the tagging performance increases at higher p_T, demonstrating the general advantage of applying telescoping deconstruction to the boosted regime. We find excellent performance of v_3 and its qualitatively different feature compared to τ_{21}. In the right panel we compare the tagging performance using truth particle information and pseudo-calorimeter clusters, which degrades information about structures smaller than the cell size. We find that v_3 is much more robust against this smearing, especially at high p_T. The v_3 observable utilizes the W isolation and probes the rapid depletion of radiation around the W at larger angles in the boosted regime. This is the manifestation of the fact that the W carries zero color charge which affects the color structure of the subjets. The time dilation that occurs before W hadronically decays can also create a huge difference from QCD jets in the jet formation process**SAM : I have no clue what you mean here or how you come to this conclusion based on this work. It seems like vigorous hand waving.**. On the other hand, the fact that v_3 performs better than v_2 hints at the significance of a third hard emission in W and QCD jets, and v_3 disentangles that effect in the quantification of the W isolation. The physics picture of $v_{\rm trim}$ and $v_{\rm prun}$ is still under study.**SAM : I don't think this is an acceptable way to leave this paragraph. It seems like disjoint observations but to just drop it as "and we don't understand this yet" is unsatisfying and makes the impression that we don't know whats going on.**

Shown in Figure 3 are the ROC curves for top tagging performance including v_{32}, $\theta_{\rm min}$, the BDT combinations of telescoping subjets variables $\{v_2, v_3, m_{W2}\}$ and $\{\theta_2, \theta_{\rm min}, m_{W2}, v_2, v_3, v_{m_{W2}}\}$, and the three-prong tagger $\tau_{32} = \tau_3/\tau_2$. Again the left and middle panels correspond to the two kinematic regimes $p_T = (350\ \text{GeV}, 500\ \text{GeV})$ and $p_T = (800\ \text{GeV}, 1\ \text{TeV})$ and we note tagging performance increases at higher p_T. In the right panel the ROC curves plot both truth-particle and pseudo-calorimeter information. We find the excellent performance of v_{32} and its robustness against smearing, especially at high p_T where the performance of the more conventionally used τ_{32} observable degrades dramatically. This indicates the qualitatively different features of v_{32} and a three-prong tagger**SAM : I don't understand this comment. Please clarify.**. We also find the usefulness of including m_{W2} in the minimal BDT combination which significantly increases the tagging performance. It is clear that the intrinsic mass scale M_W within the top jet is a unique feature distinguishing itself from the QCD background**SAM : But this is an obvious statement and we have attempted to eliminate it by making a mass cut *before* doing this study. This needs clarificiation.**. One may also start to see the W isolation within the top jet in the boosted regime.

To conclude, we introduce a qualitatively new jet substructure calculus using variability to quantify the change of observables with respect to a sampling of the phase-space boundary in the observable definition. This method is general and can be used to analyze arbitrary classes of jet substructure observables and grooming procedures. In this context, when applied to W and top tagging, we find excellent performance, especially in the case of telescoping subjet substructure, quantified via v_3 in W tagging and v_{32} in top tagging. Furthermore, their robustness is found to be significantly better than more widely used N-prong taggers such as N-subjetiness, via a comparison of the performance when reconstructed using truth particles as compared to a pseudo-calorimeter.

The new physics messages we learn include the emergence of the isolation of W jets at high p_T, which is a dominant feature over their two-prong structure. This is true for all other heavy, color-singlet Standard Model particles including the Z and the Higgs boson. The top jet also has features beyond the three-prong structure which can be exploited to increase tagging performance. The telescoping subjets provides a systematic framework within which one can construct qualitatively new jet substructure observables. This paves the road toward complete and systematic jet studies using telescoping deconstruction [25] **SAM : I don't know what is trying to be said with this last paragraph, it seems like an afterthought.**.

ACKNOWLEDGEMENTS

Y.-T. Chien would like to thank the organizers of the BOOST2015 conference where telescoping jet substructure was first presented. Y.-T. Chien was supported by the US Department of Energy (DOE), Office of Science under Contract No. DE-AC52-06NA25396, the DOE Early Career Program and the LHC Theory Initiative Postdoctoral Fellowship under the National Science Foundation grant PHY-1419008. A. Emerman was supported by the National Science Foundation under Grant No. PHY-1707971. S.-C. Hsu and S. Meehan were supported by the DOE Office of Science, Office of High Energy Physics Early Career Research program under Award Number DE-SC0015971. Z. Montague was supported by the University of Washington's Ernest M. Henley & Elaine D. Henley Endowed Fellowship.

---

* ytchien@mit.edu

[1] A. Abdesselam *et al.*, *Boost 2010 Oxford, United Kingdom, June 22-25, 2010*, Eur. Phys. J. **C71**, 1661 (2011), arXiv:1012.5412 [hep-ph].

[2] A. Altheimer *et al.*, *BOOST 2011 Princeton , NJ, USA, 22V26 May 2011*, J. Phys. **G39**, 063001 (2012), arXiv:1201.0008 [hep-ph].

[3] A. Altheimer *et al.*, *BOOST 2012 Valencia, Spain, July 23-27, 2012*, Eur. Phys. J. **C74**, 2792 (2014), arXiv:1311.2708 [hep-ex].

[4] D. Adams *et al.*, Eur. Phys. J. **C75**, 409 (2015), arXiv:1504.00679 [hep-ph].

[5] A. J. Larkoski, I. Moult, and B. Nachman, (2017), arXiv:1709.04464 [hep-ph].

[6] A. Collaboration (ATLAS Collaboration), (2017).

[7] C. Collaboration (CMS Collaboration), JHEP **12**, 017. 43 p (2014), comments: Replaced with published version. Added journal reference and DOI.

[8] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys.Rev. **D80**, 051501 (2009), arXiv:0903.5081 [hep-ph].

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

[10] S. D. Ellis, A. Hornig, T. S. Roy, D. Krohn, and M. D. Schwartz, Phys. Rev. Lett. **108**, 182003 (2012), arXiv:1201.1914 [hep-ph].

[11] Y.-T. Chien, D. Farhi, D. Krohn, A. Marantan, D. Lopez Mateos, and M. Schwartz, JHEP **12**, 140 (2014), arXiv:1407.2892 [hep-ph].

[12] J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. **105**, 022001 (2010), arXiv:1001.5027 [hep-ph].

[13] Y. Cui, Z. Han, and M. D. Schwartz, Phys.Rev. **D83**, 074023 (2011), arXiv:1012.2077 [hep-ph].

[14] J. Thaler and K. Van Tilburg, JHEP **03**, 015 (2011), arXiv:1011.2268 [hep-ph].

[15] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. Lett. **105**, 092002 (2010), arXiv:1004.2489 [hep-ph].

[16] Y.-T. Chien, Phys. Rev. **D90**, 054008 (2014), arXiv:1304.5240 [hep-ph].

[17] I. W. Stewart, F. J. Tackmann, J. Thaler, C. K. Vermilion, and T. F. Wilkason, JHEP **11**, 072 (2015), arXiv:1508.01516 [hep-ph].

[18] J. Thaler and T. F. Wilkason, JHEP **12**, 051 (2015), arXiv:1508.01518 [hep-ph].

[19] J. Thaler and L.-T. Wang, JHEP **07**, 092 (2008), arXiv:0806.0023 [hep-ph].

[20] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008), arXiv:0806.0848 [hep-ph].

[21] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008), arXiv:0710.3820 [hep-ph].

[22] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. **C72**, 1896 (2012), arXiv:1111.6097 [hep-ph].

[23] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 063 (2008), arXiv:0802.1189 [hep-ph].

[24] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, PoS **ACAT**, 040 (2007), arXiv:physics/0703039.

[25] Y.-T. Chien, P. T. Komiske, and E. M. Metodiev, to appear soon (2017).