Groomed Event Shapes in DIS

The H1 collaboration

Abstract:

Event shapes provide incisive probes of QCD, both its perturbative and non-perturbative aspects. Grooming techniques have been developed to separate perturbative from non-perturbative components of jets in a theoretically well-controlled way, and have been applied extensively to jet measurements in hadronic collisions. In this preliminary, the first application of grooming techniques to event shape measurements at HERA is presented, utilizing archived electron-proton Deep Inelastic Scattering data from the H1 experiment. The analysis is based on the novel Centauro jet clustering algorithm, which is designed specifically for the event topologies of ep DIS collisions. Cross-section measurements of groomed event 1-jettiness and groomed invariant jet mass are shown. The measurements are compared to Monte Carlo models, and to a theoretical calculation based on Soft Collinear Effective Theory.

Contents

1 Introduction

In the last decade, theoretical understanding of jets and their substructure has undergone a renaissance, inspired largely by the need for theoretical and experimental precision at the Large Hadron Collider. One of the most powerful techniques to emerge from this flourishing is jet grooming, which has proven highly fruitful in both high-energy and nuclear physics. It is a natural transition then to attempt to apply these grooming techniques in different collisions systems.

Event shapes are some of the earliest hadronic observables, dating back to before QCD was the established theory of the strong interaction. They can be calculated to high precision in perturbation theory and lend themselves nicely to factorization. For this reason they have been very successful in extracting the strong coupling constant α_S and providing valuable input to monte carlo event generators. By combining precisely calculated event shapes with well-understood grooming techniques, one can hope to access some of the wealth of information encoded in a QCD jet.

Deep-inelastic electron proton scattering provides a clean environment in which to study jets. The partonic scattering kinematics can be reconstructed much more precisely than in proton-proton collisions. Electron-positron collisions also provide a clean environment, although the momentum transfer Q is fixed by the center-of-mass energy of the collider, which is not always variable. DIS enables a precise study of jets at many different values of momentum transfer. The HERA collider complex at DESY, the only electron-proton collider thus far to exist, therefore is an excellent tool with which to study jets and their substructure.

Jet grooming has been used extensively and to great success at proton-proton colliders, and while underlying event effects are significantly smaller in DIS, grooming nonetheless has a variety of benefits. A few of these benefits are: the ability to construct observables which are free of non-global logarithms, reduction of hadronization corrections, and the existence of a parameter that can tune the amount of non-perturbative contribution.

Following the prescription laid out in Ref. [1], using the Centauro jet clustering algorithm [2], we therefore begin a study of groomed event shapes in DIS, armed for the first time with tools designed expressly for that purpose.

This analysis based on the 1-jettiness analysis described in [3]. The technical details of the analysis, such as triggers used, event QA, fiducial volume cuts, etc. are the same unless explicitly mentioned in this note.

2 Observables

2.1 Breit Frame

The Breit frame is the frame of reference where the incoming parton (at Born-level) has its momentum reversed after the collision with the virtual boson. The proton remnant-going direction is defined as $\eta = +\infty$ while the struck parton-going direction is defined as $\eta = -\infty$. The region $\eta > 0$ is known as the remnant hemisphere (RH), while the $\eta < 0$ region is known as the current hemisphere (CH). In the Breit frame, the scattered electron tends to fall at midrapidity.

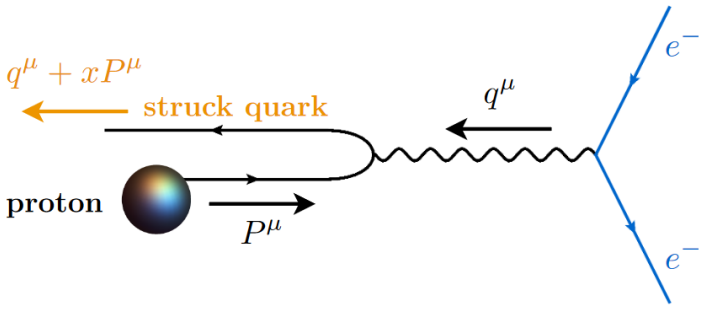


Figure 1: Illustration of the Born-level process in DIS. Image from Ref. [2].

The Breit frame is defined as the frame where

$$2x\vec{P} + \vec{q} = 0$$

holds. Here $\vec{P}$ and $\vec{q}$ are defined as the four-momenta of the incoming proton beam and the exchanged virtual boson. The boost from the lab frame to the Breit frame thus depends on x and $\vec{q}$, both of which can be reconstructed via any of the standard kinematic reconstruction methods.

In the Breit frame, the maximum available longitudinal momentum is $Q/2$. At Born-level, the Breit frame p_T of the scattered parton is zero. Events with large amounts of transverse momentum in the Breit frame are produced by multi-jet topologies. An example of this is QCD compton scattering, where the quark recoils from an emitted hard gluon and gains a non-zero p_T.

2.2 Grooming Procedure

The Centauro jet algorithm uses an asymmetric clustering measure to preferentially create a jet out of radiation in the current hemisphere of the Breit frame. This allows

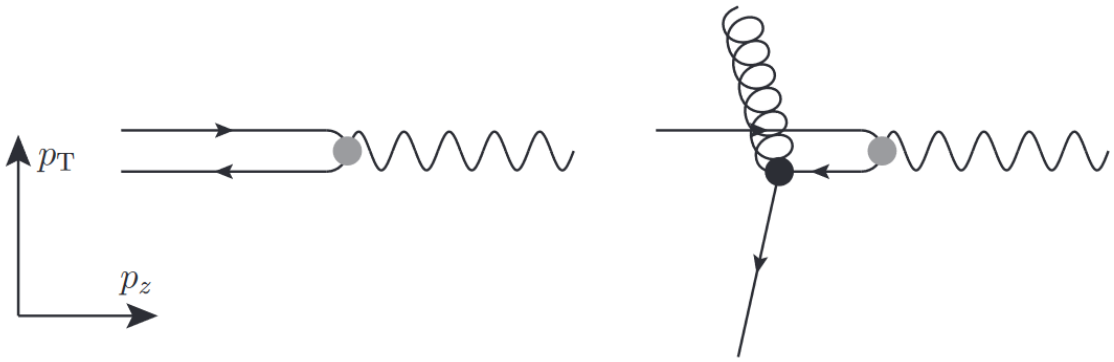


Figure 2: Illustration of jet kinematics in the Breit frame. The presence of an emitted gluon produces a dijet event with non-zero p_T. Image from Ref. [4]

Centauro to cluster the Born-level configuration into a jet more easily than previous lab- and Breit-frame algorithms.

In the HERA convention, the Born-level jet is produced at $\eta = -\infty$ in the Breit frame, and therefore will not be captured by k_t-like algorithms applied in the Breit frame. Lab-frame k_t-like algorithms may capture the born-level jet, but the likelihood that this occurs depends on the event kinematics and equivalently the jet pseudorapidity in the lab-frame. Additionally, lab frame jets can have a large amount of contamination from QCD initial-state radiation and target fragmentation in the target hemisphere, which are uncorrelated with the struck parton.

In addition to being used as a standard jet algorithm, Centauro can also be applied to an entire DIS event (equivalent to setting the jet radius $R \to \infty$) to produce a clustering tree in which the last clustered radiation is that which is farthest from $\eta = -\infty$ in the Breit frame. As laid out in [1], the natural unit for comparison of branches in this tree is the Lorentz invariant momentum fraction z_i,

$$z_i = \frac{P \cdot p_i}{P \cdot q}$$

which in the Breit frame represents the fraction of the virtual boson momentum carried by the object i. Branches of the tree with low z_i are either soft or at wide angle with respect to the virtual boson. Branches with high z_i are likely to be fragments of the struck parton.

The grooming procedure explored in this work is as follows: First cluster all particles in the event into a clustering tree with Centauro. Next, iteratively decluster the

tree in the order in which it was clustered, and compare the z_i of the branches at each step. The grooming is passed when the grooming condition

$$\frac{min(z_i, z_j)}{z_i + z_j} > z_{cut}$$

is met. This is a modified version of the $\beta = 0$ case of the modified MassDrop tagger grooming algorithm [5]. Here z_i plays the role of p_{ti} in standard mMDT. If the grooming condition is not met, the branch with smaller z_i is removed and the remaining branch is subdivided further. This continues until the grooming condition is met.

The particles which remain after the grooming has passed are typically collimated along the virtual boson direction. QCD ISR, the beam remnant, and wide-angle, soft radiation are groomed away, leaving effectively only the fragments of the struck parton. From the particles left after grooming, event shape observables can be calculated. The

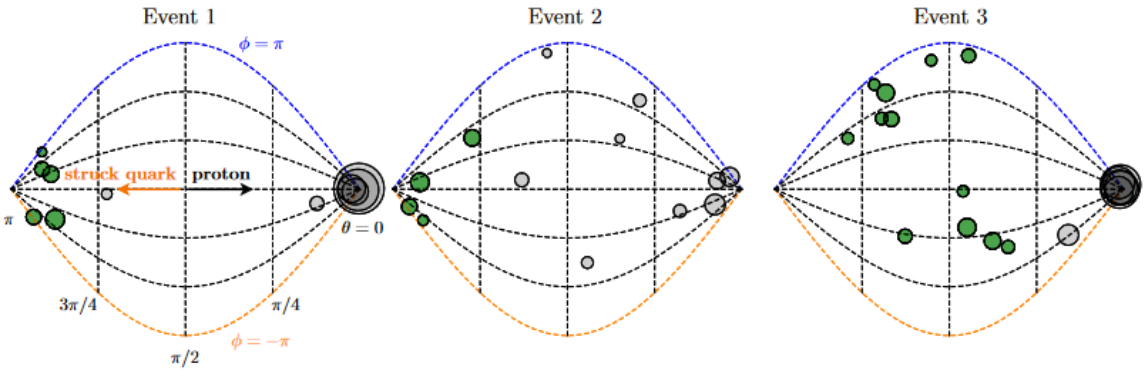


Figure 3: Visualization of groomed events. Grey circles represent particles which did not pass grooming. Note the universal removal of beam-related radiation at $\eta = +\infty$. Image from Ref. [1]

removal of radiation in the proton-going direction means that these event shapes are more tightly correlated to the development of the struck parton jet than standard DIS event shapes. Event shapes groomed according to this procedure can therefore be treated equivalently to jets, to leading order. The exception to this treatment is dijet events. Event 3 in Fig. 2 shows a dijet event, in which both of the jets have survived the grooming.

Occasionally an event will be groomed down to a single particle, at which point the grooming terminates. These events are dropped from the sample and do not contribute to the event shape cross section. The rate at which events never pass grooming increases with z_{cut}.

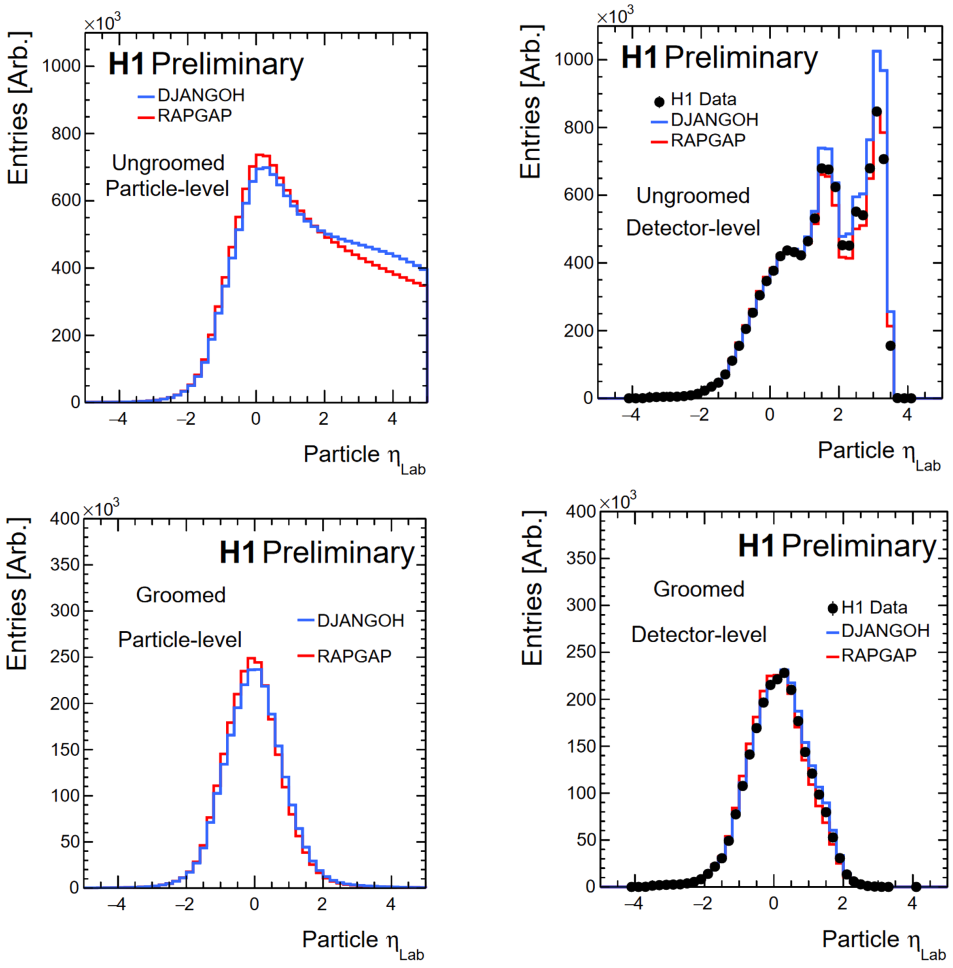


Figure 4: The locations of all particles in the lab frame for the groomed and ungroomed cases at detector- and particle-level

Particles at large forward and backward rapidity are generally less well reconstructed. The distribution of all particles in the lab frame is shown in fig. 4.

It can be seen that a large fraction of all particles are in the forward region, with a large number of particles beyond the highest rapidity accessible by the detector. The particles which survive grooming, on the other hand, tend to be well contained within the central region ($-2 < \eta < 2$) of the detector as shown in fig. 4, and thus are well reconstructed in general. The secondary peak at high pseudorapidity is removed by the grooming, and the number of particles in the remnant hemisphere is reduced by almost an order of magnitude.

2.3 Event Shape Definitions

The two event shapes currently considered are the groomed invariant mass and the groomed 1-jettiness τ_B^1. They are defined as follows:

$$\tau^1_{B_{Gr.}} = \frac{2}{Q^2} \sum_{i \in X} min(q_B \cdot p_i, q_J \cdot p_i) \tag{2.1}$$

$$M^2_{Gr.} = (\sum_i p_i^\mu)^2 \tag{2.2}$$

The 1-Jettiness event shape is described in detail in Ref. [6]. The groomed 1-jettiness $\tau^1_{B_{Gr.}}$ is defined as the sum over four-vectors in the hadronic final state, dotted with the axis describing the beam direction or the virtual boson direction in the Breit frame that produces a smaller value. In the Breit frame, the smaller of the two dot products will always be whichever hemisphere the particle is in. Particles in the current hemisphere will have smaller dot products with q_J, while particles in the remnant hemisphere will have smaller dot products with q_B.

3 Inclusive Kinematics

3.1 Event Selection

The detector-level quantities contributing to the event selection are taken from previous H1 high Q^2 jet analyses. The scattered electron energy $E_{e'}$ is required to be > 11 GeV. The scattered electron polar angle $\vartheta_{e'}$ is required to be < 2.7 Rad. The z-location of the vertex is constrained to -35 cm $< z_{vtx} <$ 35 cm. In addition, a fiducial volume cut is placed on the impact location of the scattered electron. If the electron falls into a crack in the LAr, the event is rejected. Higher-level cuts are also utilized to reduce the background from QED Compton and photoproduction events. The trigger for the accepted subset of events is $> 99\%$ efficient.

An additional cut requires the value of the total event $\sum_i E_i - P_{z_i}$ to be greater than 50 GeV and less that 60 GeV. This cut vetoes photoproduction, QED ISR, and losses due to acceptance effects. This cut is harsher than many previous HERA analyses, but it was found that the observable purities are improved by a harsher cut, and the decrease in statistics is negligible.

3.2 Kinematic Reconstruction

The phase space chosen for this analysis is $150 < Q^2 < 20000$ and $0.2 < y < 0.7$. No direct cut is placed on the value of x. The kinematic variables x, y, and Q^2 are

reconstructed with the $I\Sigma$ method as described in [7]. Henceforth, the quantity Σ_{HFS} is defined as the sum over $E - P_z$ of all particles in the hadronic final state. This is different than the quantity used in the previous section in that the scattered electron is not included in the sum.

$$E^{e'}_{I\Sigma} = \frac{\Sigma_{HFS} + E_{e'}(1 - \cos\theta_{e'})}{2} \tag{3.1}$$

$$x_{I\Sigma} = \frac{E_{e'}}{E_p}\frac{\cos^2\theta/2}{y_\Sigma} \tag{3.2}$$

$$y_{I\Sigma} = y_\Sigma = \frac{\Sigma}{\Sigma + E(1 - \cos\theta)} \tag{3.3}$$

$$Q^2_{I\Sigma} = \frac{E^2\sin^2\theta}{1 - y_\Sigma} \tag{3.4}$$

After the kinematic variables are determined, the boost to the Breit frame can be reconstructed according to

$$2x\vec{P} + \vec{q} = 0$$

With the $I\Sigma$ method, around 0.2% of events have an unphysical boost, typically because the reconstructed value of $\beta > 1$. These events are rejected.

4 Results

The groomed invariant mass and 1-jettiness spectra are evaluated at $z_{cut} = 0.05$, 0.1, and 0.2. The corrected data are compared to DJANGOH, RAPGAP, default Pythia 8.3, PYTHIA with VINCIA, PYTHIA with DIRE, default Herwig 7.2, Herwig 7.2 with the internal implementation of MC@NLO, Herwig 7.2 with dipole merging, SHERPA with AHADIC++, and SHERPA with Lund string fragmentation. The groomed invariant mass spectrum is compared additionally to a NNLL prediction from Ref. [1]. A more thorough description of the model parameters used can be found in the Appendix.

The binning presented here can be somewhat unintuitive. Table 1 illustrates the values of $\frac{M^2_{Gr.}}{Q^2_{Min.}}$ and $M_{Gr.}$ included in each bin. The range of possible values the GIM can take is bracketed by 0 and W^2, such that $0 < \frac{M^2_{Gr.}}{Q^2} < \frac{1-x}{x}$.

The normalization for the GIM is provided by the minimum Q^2 used in the analysis, which in this case is 150 GeV^2. This choice of normalization is justified in Ref. [1], and is included in the NNLL calculation.

Large invariant masses are produced by multijet events, while low invariant masses are produced by single jet events. As z_{cut} increases, a shoulder representing a highly enhanced sample of dijet events appears in the GIM spectrum. The models which do the

Bin Edge	1	2	3	4	5	6	7	8	9	10	11
$\frac{M^2_{Gr.}}{Q^2_{Min.}}$	.0001	.0067	.018	0.05	0.14	0.37	1.00	2.72	7.39	20.09	54.60
$M_{Gr.}$	0.083	1.01	1.66	2.73	4.506	7.43	12.25	20.19	33.29	54.89	90.50

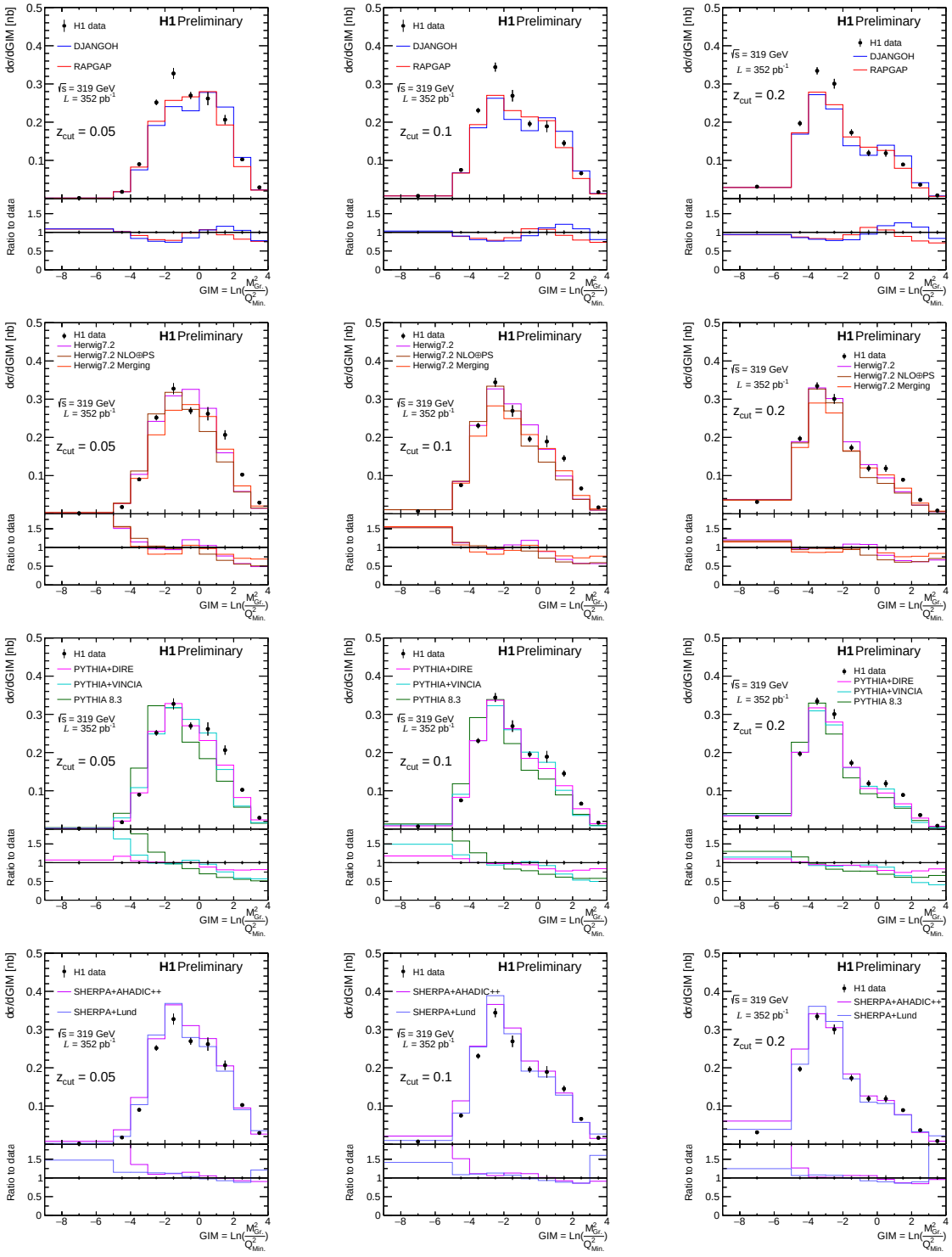
Table 1: Bin edges for the GIM measurement.

best job at modelling this fixed-order behavior are RAPGAP and Herwig + Matchbox. The three implementations of PYTHIA significantly underestimate the large invariant mass region, while DJANGOH and default Herwig tend to overestimate. This failure of PYTHIA may be due to improper matching to a fixed-order calculation. It will be interesting to compare this region with explicit fixed-order predictions, either analytic or numerical.

The middle and low GIM region is sensitive to jet substructure and hadronization. The sensitivity to small angles and hadron masses makes measurement in this region challenging, necessitating the large size of the first bin to maintain an acceptable purity. In this region, the PYTHIA-based models tend to overestimate while DJANGOH and RAPGAP tend to have reasonable agreement. DJANGOH, RAPGAP, and all the PYTHIA models implement Lund string fragmentation, while Herwig utilizes cluster fragmentation. SHERPA offers the possibility to use both cluster and string fragmentation.

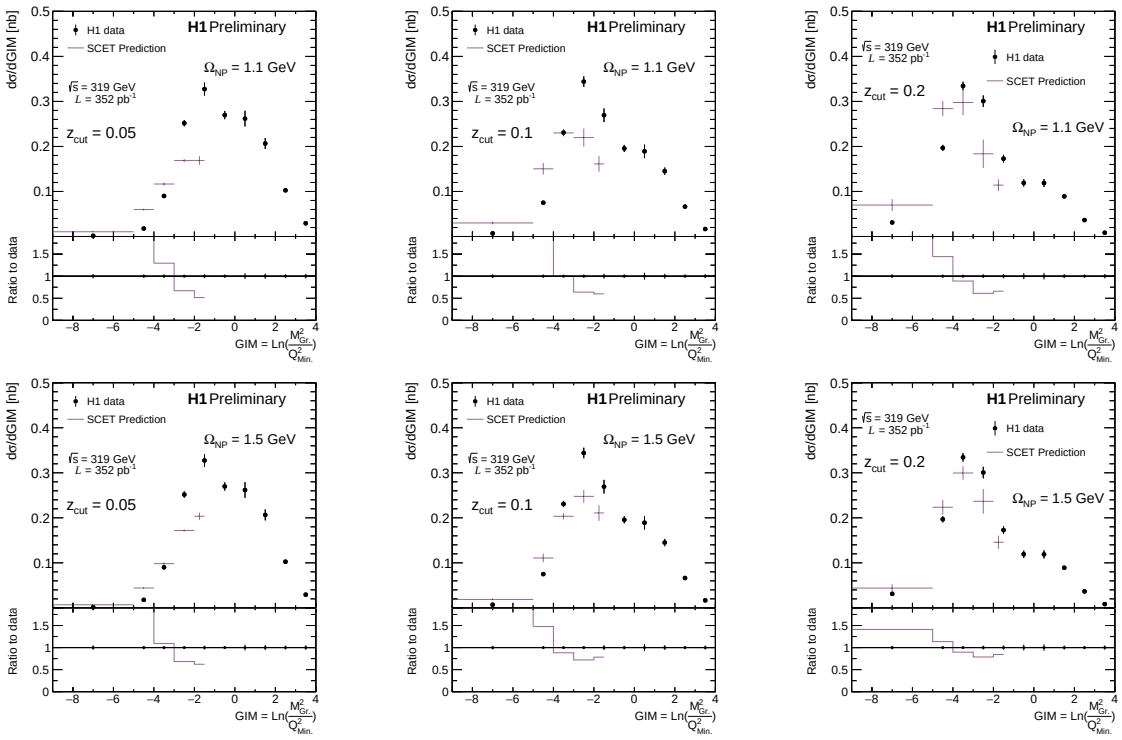
NNLL predictions are available for two different values of the non-perturbative shape function mean Ω, 1.5 GeV and 1.1 GeV. With a measurement differential in z_{cut} and Q it may be possible to extract the preferred value of this parameter.

The groomed τ_1^b spectrum has a similar interpretation to the GIM spectrum. The large values of τ_1^b are dominated by fixed-order behavior while the low values are sensitive to single jet evolution and hadronization. The PYTHIA-based models once again overestimate the peak region while underestimating the fixed-order tail. DJANGOH, RAPGAP, and Herwig+Matchbox do a good job in the tail but underestimate the peak.



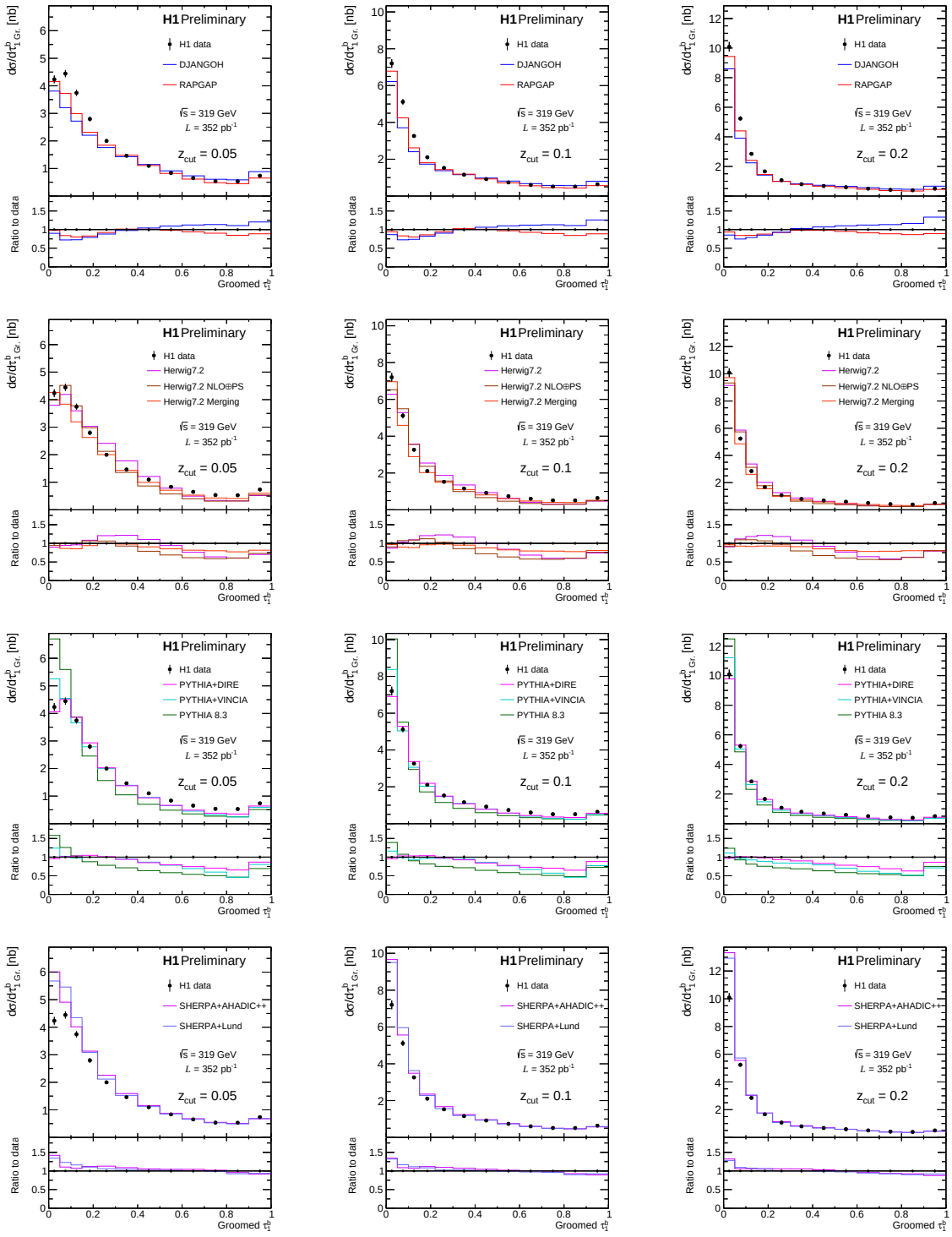
(a) $z_{cut} = 0.05$ (b) $z_{cut} = 0.1$ (c) $z_{cut} = 0.2$

Figure 5: Results from MCEGs for the groomed invariant mass.



(a) $z_{cut} = 0.05$ (b) $z_{cut} = 0.1$ (c) $z_{cut} = 0.2$

Figure 6: Results for the groomed invariant mass from analytic SCET calculations.



(a) $z_{cut} = 0.05$ (b) $z_{cut} = 0.1$ (c) $z_{cut} = 0.2$

Figure 7: Results from MCEGs for the groomed 1-jettiness.

References

[1] Y. Makris, "Revisiting the role of grooming in DIS," *Physical Review D* **103** no. 5, (Mar, 2021) . http://dx.doi.org/10.1103/PhysRevD.103.054005.

[2] M. Arratia, Y. Makris, D. Neill, F. Ringer, and N. Sato, "Asymmetric Jet Clustering in Deep-Inelastic Scattering," arXiv:2006.10751 [hep-ph].

[3] J. Hessler, "Measurement of the 1-jettiness Event Shape Observable in Deep-inelastic Electron-Proton Scattering," other thesis.

[4] D. A. Britzger, *Regularized Unfolding of Jet Cross Sections in Deep-Inelastic ep Scattering at HERA and Determination of the Strong Coupling Constant*. PhD thesis, U. Hamburg, Dept. Phys., 2013.

[5] S. Marzani, L. Schunk, and G. Soyez, "A study of jet mass distributions with grooming," *Journal of High Energy Physics* **2017** no. 7, (Jul, 2017) . http://dx.doi.org/10.1007/JHEP07(2017)132.

[6] D. Kang, C. Lee, and I. W. Stewart, "Using 1-Jettiness to Measure 2 Jets in DIS 3 Ways," *Phys. Rev. D* **88** (2013) 054004, arXiv:1303.6952 [hep-ph].

[7] U. Bassler and G. Bernardi, "On the kinematic reconstruction of deep inelastic scattering at hera," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **361** no. 1-2, (Jul, 1995) 197–208. http://dx.doi.org/10.1016/0168-9002(95)00173-5.