PROCEEDINGS OF SCIENCE

First measurement of groomed event shape observables in deep-inelastic electron-proton scattering at HERA

Radek Žlebčík[a,*]

[a] *Istituto Nazionale di Fisica Nucleare (INFN), via Valerio 2, Trieste, Italy*

E-mail: radek.zlebcik@ts.infn.it

The first measurement of groomed event shape observables in deep-inelastic electron-proton scattering (DIS) is presented. The study is based on data collected with the H1 detector at HERA in the years 2003-2007 at the centre-of-mass energy $\sqrt{s} = 319\,\text{GeV}$. The grooming techniques have been popular for jet substructure studies in hadron-hadron collisions, but have not been used in DIS before. In this measurement, the grooming is performed in the Breit frame utilizing the Centauro jet clustering algorithm which is designed for DIS event topologies. For events selected in the kinematic region of $Q^2 > 150\,\text{GeV}^2$ and $0.1 < y < 0.7$ the groomed 1-jettines and groomed invariant mass are measured considering three choices of the grooming parameter. Several Monte Carlo predictions and analytical calculations are compared to the data.

42nd International Conference on High Energy Physics (ICHEP2024)
18-24 July 2024
Prague, Czech Republic

*Speaker

PoS(ICHEP2024)487

1. Introduction

The event shape observables have been extensively studied in the e^+e^- [1] and ep collisions [2]. Such observables are calculable to high precision using perturbation quantum chromodynamics (pQCD) which allows, for example, a precise determination of the strong coupling constant α_S [3] or tuning the hadronisation and parton shower parameters of the MC generators [4].

An alternative way of probing pQCD is to study the jets and their substructure. Over the years several jet algorithms have been developed for hadron-hadron collisions including the anti-k_T [5] and the Cambridge-Aachen [6] algorithms. The substructure of the "fat" jets can help to identify boosted objects, such as the Higgs boson or the top quark. To disentangle the effects of underlying event and/or multi-parton interactions (MPI) from the main decay, the grooming techniques have been developed [7]. The grooming techniques such as Soft Drop [8] algorithm are used to remove the soft and wide-angle radiation from the jet, which suppresses the soft QCD contribution and makes the main process (often of the electroweak nature) more pronounced. The Soft Drop algorithm uses a clustering sequence of the Cambridge-Aachen algorithm and iteratively removes the softest particle in the branching until the grooming condition is met. The algorithm operates in reverse order, starting from the point where all constituents are clustered together.

It was proposed to apply the grooming techniques also in the ep interactions [9], where in deep-inelastic scattering (DIS) the hadronic final state can be considered as a single jet. Even though the ep interactions are free of MPI, the grooming techniques are still advantageous as the groomed observables are free of non-global logarithms and the hadronization effects are reduced. Consequently, the groomed observables can be calculated in pQCD with better precision and the grooming parameter z_{cut} controls the magnitude of non-perturbative effects.

The most natural frame in DIS to study the event shape observables is the Breit frame [10], where the proton flies to the positive z direction, the virtual photon to the negative z direction and the photon's 4-momentum has zero time-like component. In the Breit frame we distinguish between the "Current hemisphere" ($\eta < 0$) and the "Remnant hemisphere" ($\eta > 0$), where η is the pseudorapidity. However, the anti-k_T algorithm is not suitable for the Breit frame, as for example, at the lowest order the struck parton is aligned with the z-axis and has no transverse momentum (see Fig. 1). For this reason, the Centauro jet algorithm [11] was developed which is designed especially for the DIS event topologies. We use the Centauro jet algorithm to perform grooming in the Breit frame, the whole hadronic final state is considered as a single jet.

In the presented analysis, which is described in detail in [12], two observables are measured, the groomed 1-jettiness τ_1^b and the groomed invariant mass (GIM). The 1-jettiness can be calculated from the particles in the Current hemisphere of the Breit frame as:

$$\tau_1^b = 1 - \frac{2}{Q^2} \sum_{i \in \text{Current}} P_{z,i}^{\text{Breit}}, \tag{1}$$

where $P_{z,i}^{\text{Breit}}$ is the longitudinal momentum of the i-th particle in the Breit frame. The 1-jettiness is similar to thrust in the e^+e^- collisions, but the axis of the thrust is aligned with the z direction. In older HERA measurements [2] the thrust defined in the exactly same way as in e^+e^- was used, however, it was shown that such definition includes non-global logarithms [13] and is, therefore,

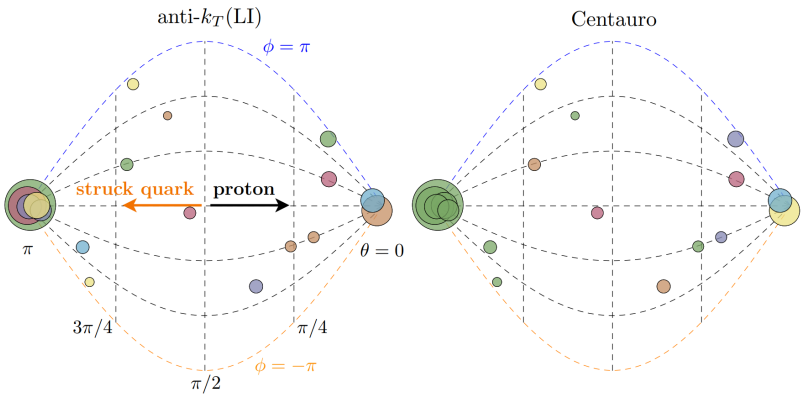

Figure 1: A scheme comparing the anti-k_T jet algorithm (left) and the Centauro jet algorithm (right) in the Breit frame. Particles are depicted as disks with area proportional to the particle energy. Particles with identical color are part of the same jet. Taken from [11].

not suitable for the pQCD calculations. The groomed invariant mass is the invariant mass of the hadronic final state particles which survived grooming.

2. The measurement

The data used in the analysis were collected with the H1 detector at HERA in the years 2003-2007 at the centre-of-mass energy $\sqrt{s} = 319\,\text{GeV}$. Both e^+p and e^-p interactions are used in the analysis. The analyzed dataset has an integrated luminosity of $351\,\text{pb}^{-1}$. A detailed description of the H1 detector components relevant to the analysis can be found in [12]. In the offline analysis, the 4-vectors of the energy flow objects are used, these combine the information from the calorimeters and the tracking detectors such that no double counting of the energy is present. The positive z-axis of the laboratory coordinate system is aligned with the direction of the proton momentum. The events are selected in the photon virtuality range $Q^2 > 150\,\text{GeV}^2$ and with inelasticity $0.1 < y < 0.7$.

The energy flow objects are clustered in the Breit frame by the Centauro algorithm and, consequently, the clustering sequence is used to perform the grooming in a similar way as for Soft Drop. Three values of the grooming strength z_{cut} are used in the analysis, 0.05, 0.1, 0.2 (for brevity 0.1 is omitted in this text). The detector-level groomed observables are corrected for the detector resolution and acceptance using regularized unfolding with Tikhonov regularization on the curvature [14]. Two MC models were used for the unfolding of groomed 1-jettiness and groomed invariant mass, Django [15] and Rapgap [16] which have different types of parton shower. The difference between the unfolded observables from the two MC models is considered as a systematic uncertainty and is the dominant one for most of the measured bins. In addition, the measured cross sections are corrected for the QED radiative effects calculated by HERACLES, the size of the corrections is about 15%. The main systematic uncertainties originate from the unfolding (regularization strength, MC model dependence) and the integrated luminosity. The uncertainties related to the energy scale and angular resolution of the energy flow objects and scattered electron are typically smaller. Together with the single differential measurements of groomed τ_1^b and GIM, these two variables are also measured double-differentially in several bins of Q^2.

3. Comparison with the theoretical predictions

The measured distributions of groomed 1-jettiness are shown in Fig. 2 for two values of the grooming parameter z_{cut}. The groomed 1-jettiness peaks at small τ_1^b around 0.05 which corresponds

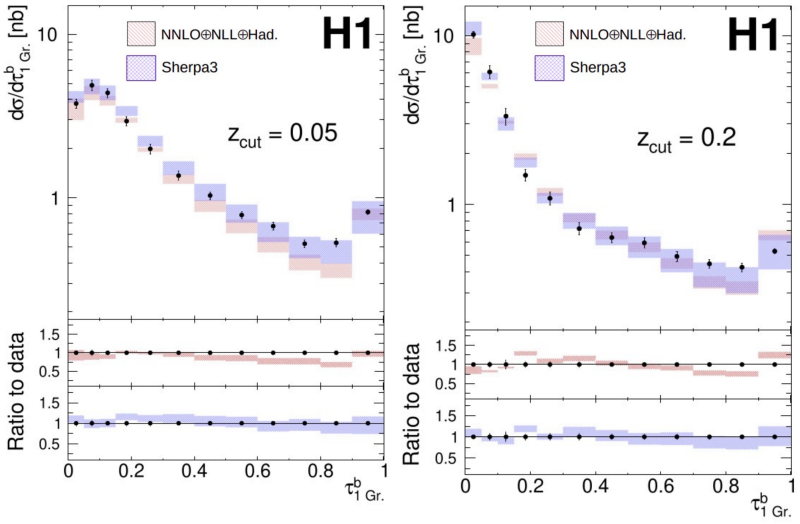

Figure 2: The groomed 1-jettiness distribution for a grooming strenght $z_{\text{cut}} = 0.05$ (left) and $z_{\text{cut}} = 0.2$ (right) as measured by the H1 collaboration. The black error bars represent the total experimental uncertainty. The NNLO+NLL' [17] and the Sherpa 3 [18] MC predictions are compared to the data. For comparison with other theoretical predictions see [12].

to the events where the groomed final state is a single narrow jet (the "peak" region). The region of high τ_1^b, around $\tau_1^b > 0.5$ is characterized by multi-jet topology and is sensitive to the fixed-order calculations while the small τ_1^b region is sensitive to the resummation effects. The data are accompanied by the Sherepa 3 MC predictions [18] at the NLO+PS level, where the uncertainty of the predictions is obtained from the 7-point variation of the renormalization and factorization scales. This theoretical uncertainty is about 10% in the peak region and about 30% for the high τ_1^b values. The predictions were provided directly by the Sherepa 3 authors and are based on the pre-release version of Sherpa 3 which has better performance for DIS due to an extra tuning. The Fig. 2 also includes the NNLO+NLL' prediction [17] where the fixed-order and resummation effects are combined. These predictions have smaller uncertainty from the scale variation, however, the agreement between the data and central values of the predictions is comparable to the Sherpa 3 MC.

The measured spectrum of the groomed invariant mass (GIM), more precisely $\ln(M^2_{\text{Gr.}}/150\,\text{GeV}^2)$, is presented in Fig. 3 for two choices of the z_{cut} parameter. This time, it is compared to the SCET calculations [9] which should be valid in the limit $M^2_{\text{Gr.}}/Q^2 \ll z_{\text{cut}} \ll 1$, i.e. for small value of the grooming parameter but even for much smaller values of $M^2_{\text{Gr.}}/Q^2$. For that reason, the precisions are shown only for GIM < -2 which corresponds to $M^2_{\text{Gr.}}/Q^2 \simeq 0.1$. It can be seen that for small z_{cut} value of $z_{\text{cut}} = 0.05$ the slope of the SCET predictions differs from the data. On the other hand, for $z_{\text{cut}} = 0.2$ the agreement between the data and the SCET predictions is better which may be related to a better fulfilment of $M^2/Q^2 \ll z_{cut}$ criteria. The SCET calculations depend

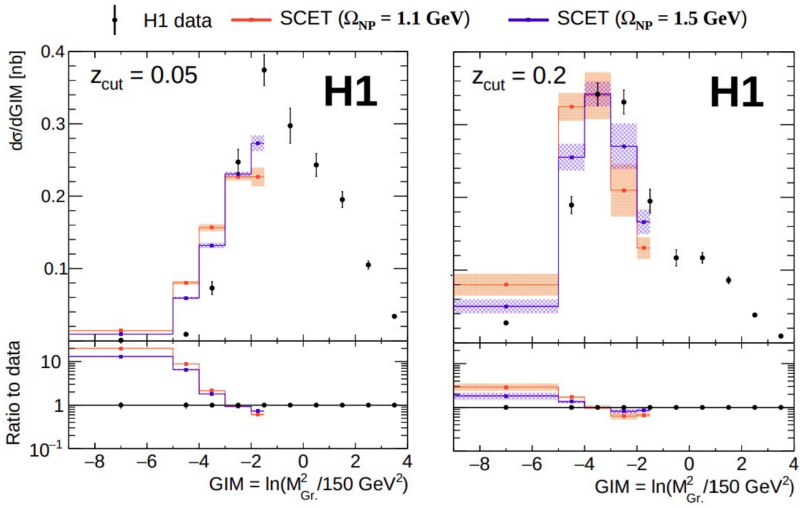

Figure 3: The spectrum of the groomed invariant mass GIM as measured by H1 collaboration for two choices of the groomed parameter z_{cut}. Together with the data points, the SCET predictions for two choices of Ω_{NP} parameter are shown. See [12] for a more detailed description.

on the non-perturbative shape function mean parameter Ω_{NP}, where predictions for two choices, $\Omega_{\text{NP}} = 1.1$ GeV and $\Omega_{\text{NP}} = 1.5$ GeV are shown. Apparently, H1 data prefer the higher value of Ω_{NP}.

As mentioned before, the event shape observables are a vital input for the MC tuning. To demonstrate the power of the measured data to discriminate between several MC models, the ratio between predicted and measured cross sections of the groomed invariant mass is shown in Fig. 4 for several MC predictions. In Fig. 4 there are legacy MC models Rapgap and Django, but also

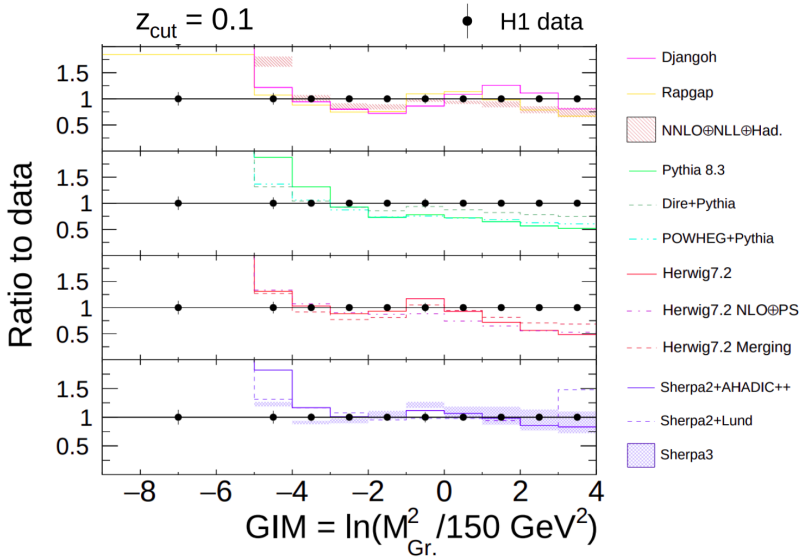

Figure 4: The ratio to the H1 data for the groomed invariant mass distribution for several MC predictions. The black error bars represent the total experimental uncertainty of data, the shaded bands are the uncertainties of the theoretical predictions. A more detailed description can be found in [12].

the more recent generators such as Pythia 8 [19], Herwig 7 [20] and Sherpa [18, 21]. Various

configurations for each generator are shown differing for example in the parton shower model, the hadronization model or usage of NLO POWHEG matching [22] between the matrix element and parton shower. It can be seen that Sherpa 3 and the legacy MC models describe the data best. For the Pythia-based prediction, the DIRE parton shower [23] gives a better data description than the default Pythia 8 parton shower. The description of the data by Herwig 7 is comparable among the three studied options. Sherpa 3 gives superior performance compared to the Sherpa 2 with Cluster hadronization (AHADIC++) [24] and the Sherpa 2 with the Lund hadronization [25].

References

[1] M. Z. Akrawy et al. *Z. Phys. C*, 47:505–522, 1990.
[2] A. Aktas et al. *Eur. Phys. J. C*, 46:343–356, 2006.
[3] Riccardo Abbate, Michael Fickinger, Andre H. Hoang, Vicent Mateu, and Iain W. Stewart. *Phys. Rev. D*, 83:074021, 2011.
[4] Peter Skands, Stefano Carrazza, and Juan Rojo. *Eur. Phys. J. C*, 74(8):3024, 2014.
[5] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. *JHEP*, 04:063, 2008.
[6] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. *JHEP*, 08:001, 1997.
[7] Jonathan M. Butterworth, Adam R. Davison, Mathieu Rubin, and Gavin P. Salam. *Phys. Rev. Lett.*, 100:242001, 2008.
[8] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, and Jesse Thaler. *JHEP*, 05:146, 2014.
[9] Yiannis Makris. *Phys. Rev. D*, 103(5):054005, 2021.
[10] K. H. Streng, T. F. Walsh, and P. M. Zerwas. *Z. Phys. C*, 2:237, 1979.
[11] Miguel Arratia, Yiannis Makris, Duff Neill, Felix Ringer, and Nobuo Sato. *Phys. Rev. D*, 104(3):034005, 2021.
[12] V. Andreev et al. *Eur. Phys. J. C*, 84(7):718, 2024.
[13] Daekyoung Kang, Christopher Lee, and Iain W. Stewart. *Phys. Rev. D*, 88:054004, 2013.
[14] Stefan Schmitt. *JINST*, 7:T10003, 2012.
[15] K. Charchula, G. A. Schuler, and H. Spiesberger. *Comput. Phys. Commun.*, 81:381–402, 1994.
[16] Hannes Jung. *Comput. Phys. Commun.*, 86:147–161, 1995.
[17] Max Knobbe, Daniel Reichelt, and Steffen Schumann. *JHEP*, 09:194, 2023.
[18] Gurpreet Singh Chahal and Frank Krauss. *SciPost Phys.*, 13(2):019, 2022.
[19] Christian Bierlich et al. *SciPost Phys. Codeb.*, 2022:8, 2022.
[20] Johannes Bellm et al. *Eur. Phys. J. C*, 76(4):196, 2016.
[21] Enrico Bothmann et al. *SciPost Phys.*, 7(3):034, 2019.
[22] Stefano Frixione, Paolo Nason, and Carlo Oleari. *JHEP*, 11:070, 2007.
[23] Stefan Höche, Frank Krauss, and Stefan Prestel. *JHEP*, 10:093, 2017.
[24] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. *Eur. Phys. J. C*, 36:381–395, 2004.
[25] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. *Phys. Rept.*, 97:31–145, 1983.

PoS(ICHEP2024)487