PROCEEDINGS OF SCIENCE

Systematic exploration of multi-scale jet substructure in p + p collisions at $\sqrt{s}$ = 200 GeV by the STAR experiment

Monika Robotková[a,b,*] **on behalf of the STAR collaboration**

[a] *Nuclear Physics Institute of the CAS,*
Husinec - Řež, 130, Řež 250 68, Czech Republic

[b] *Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague,*
Břehová 7, Praha, 115 19, Czech Republic

E-mail: robotkova@ujf.cas.cz

Jets are multi-scale objects that connect asymptotically free partons to confined hadrons. Jet substructure measurements in vacuum provide essential insight into the parton evolution and the ensuing non-perturbative processes. In this study, we use the SoftDrop grooming technique, based on the angular-ordered Cambridge/Aachen reclustering algorithm, to probe correlations between jet substructure variables. This technique provides a correspondence between experimental observables and QCD splitting functions in vacuum. Corrections for detector effects are carried out utilizing either a three dimensional correction procedure or a machine learning based framework called MultiFold, with the latter retaining the correlations across more jet substructure observables, in an unbinned way.

In these proceedings, we explore ensemble-level and jet-by-jet correlations between variables such as the shared momentum faction (z_g), splitting scale (k_T), groomed mass fraction (μ), jet charge (Q) and groomed jet radius (R_g) for jets of varying momenta and radii in p+p collisions at $\sqrt{s}$ = 200 GeV using the STAR detector. To study the evolution along the jet shower, we present splitting observables at the first, second, and third hard splits in the SoftDrop splitting history for various jet and initiator prong momenta. Finally, the measurements are compared to leading order Monte Carlo models, PYTHIA 6, PYTHIA 8 and HERWIG.

26-31 March 2023
Aschaffenburg, Germany

*Speaker

1. Introduction

Jets are collimated sprays of hadrons created from hard scattered partons via parton showering and hadronization. Study of jet substructure provides insight into perturbative (parton shower) and non-perturbative QCD (hadronization) processes. Access to the jet substructure is allowed by using the SoftDrop [1] grooming technique based on removing soft wide angle radiation. Jets are first clustered with the anti-k_{T} [2] algorithm and then reclustered with the Cambridge/Aachen (C/A) [3] algorithm in order to obtain angular-ordered tree. By undoing the last step of the C/A algorithm the jet is separated into two subjets which are required to pass the SoftDrop condition:

$$z_{\mathrm{g}} = \frac{\min(p_{\mathrm{T},1}, p_{\mathrm{T},2})}{p_{\mathrm{T},1} + p_{\mathrm{T},2}} > z_{\mathrm{cut}} \left(\frac{R_{\mathrm{g}}}{R}\right)^{\beta}, \tag{1}$$

where $p_{\mathrm{T},i}$ is the transverse momentum of the corresponding subjet, R is the resolution parameter of the jet, and R_{g} is the distance between the two subjets. Two free parameters in this condition are set to $(z_{\mathrm{cut}}, \beta) = (0.1, 0)$ in our analysis. Softer subjets are removed and process is repeated until the condition is passed.

SoftDrop technique also produces two observables: the shared momentum fraction (z_{g}) and the groomed radius (R_{g}) which represent momentum and angular scale, respectively. The splitting scale $k_{\mathrm{T}} = z_{\mathrm{g}} p_{\mathrm{T,jet}} \sin R_{\mathrm{g}}$ allows us to connect both these scales and study the kinematic region of Lund Plane [4].

To study the soft component of the jet usually removed by SoftDrop we use a technique called CollinearDrop [5]. This technique is based on the difference of an observable for two SoftDrop settings of z_{cut} and β parameters. In our case, we use $(z_{\mathrm{cut},1}, \beta_1) = (0, 0)$ and $(z_{\mathrm{cut},2}, \beta_2) = (0.1, 0)$, where first setting is jet without grooming and second setting is our standard SoftDrop setting. After this subtraction, what remains is soft radiation that would have been removed by SoftDrop. Using both SoftDrop and CollinearDrop techniques enables study of different stages of the jet.

Besides momentum and angular observables, it is also important to study the jet mass observables and correlations between them in order to understand the jet evolution, due to the connection between the jet mass and parton virtuality. The jet mass observable coming from CollinearDrop $\Delta M/M$ is defined as

$$\Delta M/M = \frac{M - M_{\mathrm{g}}}{M}, \tag{2}$$

where M is the jet mass and M_{g} is the mass of the groomed jet defined as

$$M_{(\mathrm{g})} = \left| \sum_{i \in (\mathrm{groomed})\mathrm{jet}} p_i \right| = \sqrt{E^2_{(\mathrm{g})} - |\mathbf{p}_{(\mathrm{g})}|^2}. \tag{3}$$

The second mass observable called the groomed mass fraction (μ) is defined as

$$\mu \equiv \frac{\max(m_{\mathrm{j},1}, m_{\mathrm{j},2})}{M_{\mathrm{g}}}, \tag{4}$$

where $m_{\mathrm{j},i}$ is the mass of the subjet.

The first observable $\Delta M/M$ provides the information about the amount of mass groomed away and the second observable μ allows study of mass sharing of the hard splitting.

2. Analysis details

The p+p data were collected in 2012 at $\sqrt{s}$ = 200 GeV by the STAR experiment [6]. A detailed description of the data and analysis cuts can be found in Ref. [7]. Since the data are affected by the detector effects, we need to correct them to obtain the true particle-level distribution. For this purpose, we use two unfolding methods. The first one is "(2+1)D" unfolding by RooUnfold [8] and the second one is the machine learning method MultiFold [9].

The first method, (2+1)D unfolding, consists of two parts. In the first part, two observables μ and $\log(k_{\mathrm{T}})$ vs. R_{g} are unfolded using the 2D Iterative Bayesian unfolding [10] for each detector-level $p_{\mathrm{T,jet}}$ bin. Then the correction for the $p_{\mathrm{T,jet}}$ is done. We use a $p_{\mathrm{T,jet}}$ response matrix with axes given by particle-level and detector-level $p_{\mathrm{T,jet}}$ and perform the projections of each particle-level $p_{\mathrm{T,jet}}$ bin onto detector-level $p_{\mathrm{T,jet}}$ bins. These projection are then used as weights for the 2D unfolded spectra and weighted spectra are summed. We also apply additional corrections for jet finding and trigger efficiencies.

The second method called MultiFold allows one to simultaneously unfold multiple observables in an unbinned way. We unfold at the same time six observables, namely $p_{\mathrm{T,jet}}$, R_{g}, z_{g}, M, M_{g} and jet charge $Q^{\kappa=2}$, defined as $Q^{\kappa=2} = \frac{1}{(p_{\mathrm{T,jet}})^{\kappa=2}} \sum_{i \in \mathrm{jet}} q_i \cdot (p_{\mathrm{T},i})^{\kappa=2}$, where q_i and $p_{\mathrm{T},i}$ are electric charge and transverse momentum of the i-th constituent, respectively.

This method uses PYTHIA 6 (STAR tune) [11–13] as a particle-level prior. PYTHIA 6 events pass through GEANT 3 simulation of the STAR detector and are embedded into zero-bias measured data from the same run period as the data being corrected. After the embedding procedure, detector-level jets are reconstructed and matched to the particle-level true jets. To correct detector-level jets emerging from background - so called 'fake jets' - simulations provide us with fake rates which are used as initial weights for input data to MultiFold. An efficiency correction was applied for the particle-level jets missed at detector level.

3. Results

Figure 1 (left) shows the correlation between groomed radius R_{g} and CollinearDrop groomed mass fraction $\Delta M/M$. We can observe the diagonal trend which indicates anti-correlation between R_{g} and $\Delta M/M$. This anti-correlation between these two observables is consistent with expectation of angular ordering of the parton shower. Figure 1 (right) shows the projection of $\Delta M/M$ for three different R_{g} selections. We can see that for the large radius (yellow color) the $\Delta M/M$ peaks at small values which indicates only little or no soft wide angle radiation in the shower. Data were also compared to different Monte Carlo simulations, namely PYTHIA 8 (Detroit tune) [14, 15] and HERWIG 7 (LHC tune) [16]. Both models describe the trend of the data.

Correlation between z_{g} and $\Delta M/M$ is shown in Figure 2. We can see that the more mass is groomed away relative to the ungroomed mass (yellow color), the flatter and more non-perturbative the z_{g} distribution is. This correlation also shows how an early-stage emission affects the momentum imbalance of the later splitting. Monte Carlo models again describe the trend of the data.

Figure 3 shows the μ observable for three different R_{g} selections and two different $p_{\mathrm{T,jet}}$ bins. We observe weaker dependence on R_{g} compared to the $\Delta M/M$, but the shift of μ to smaller values for smaller R_{g} indicates that narrow splits lead to smaller transfer of virtuality or mass. The

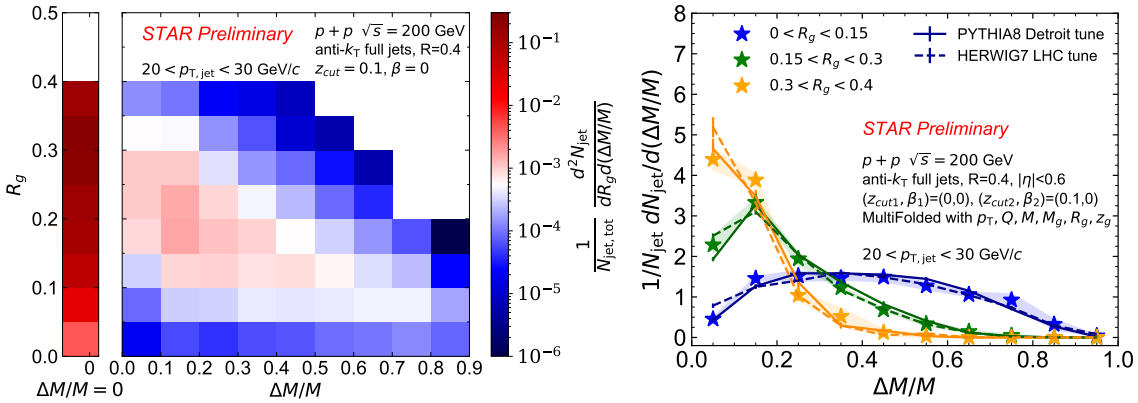

Figure 1: Correlation between $\Delta M/M$ and R_g (left) unfolded with MultiFold method and the projection of $\Delta M/M$ for three different R_g selections (right) for jets with $R = 0.4$ in p+p collisions at $\sqrt{s} = 200$ GeV.

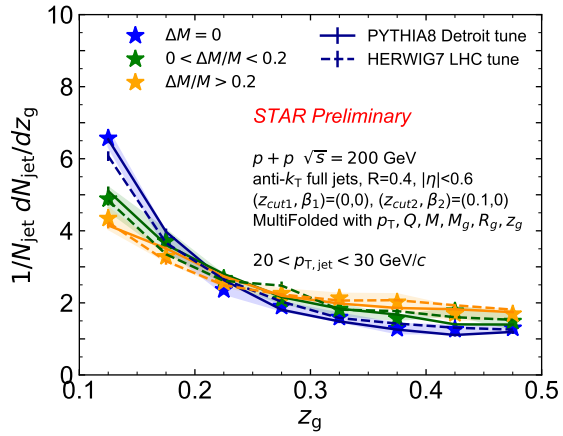

Figure 2: z_g distributions unfolded with MultiFold method for three $\Delta M/M$ bins for jets with $R = 0.4$ in p+p collisions at $\sqrt{s} = 200$ GeV.

dependence on $p_{T,jet}$ is very small. The trend of the data is well described by the PYTHIA 6 and 8 (Monash tune) [17] simulations.

Figure 4 shows $\log(k_T)$ distributions for three R_g selections and two $p_{T,jet}$ bins. We can observe strong dependence on R_g and weak dependence on $p_{T,jet}$. Since the zero value on the x-axis corresponds to 1 GeV, it is evident that the distribution is moving from the non-perturbative to perturbative region as we go from smaller to larger R_g, since formation time $\tau \sim 1/R_g$. This measurement allows us to study a broad kinematic region of the Lund Plane.

4. Conclusion

We presented correlation between jet substructure observables at the first hard split using SoftDrop and CollinearDrop methods. Two unfolding methods for multidimensional measurements were introduced, (2+1)D unfolding and MultiFold. It was shown that selecting on different jet substructure observables and correlations between them allows us to study a broad kinematic

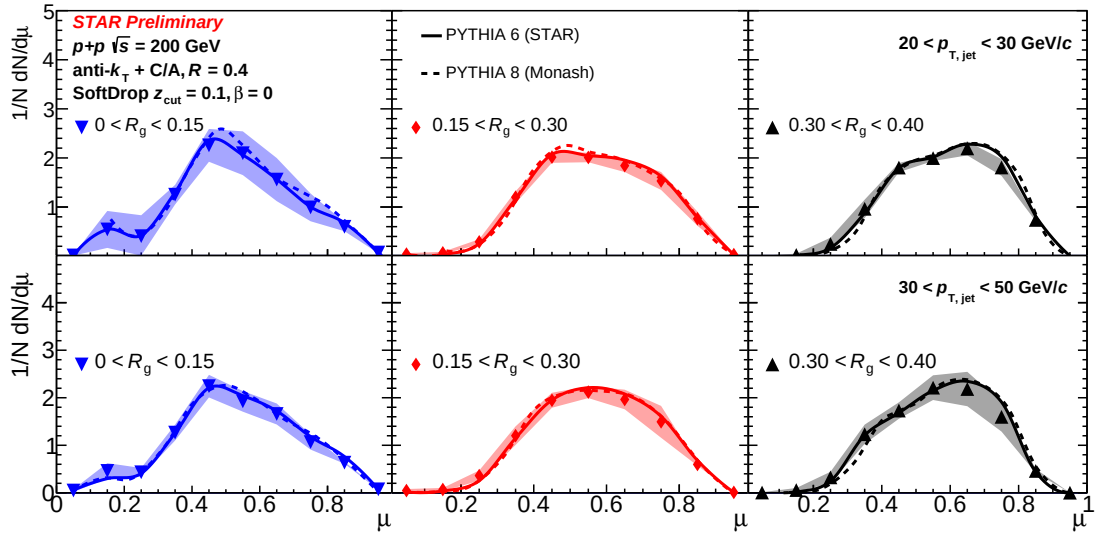

Figure 3: μ distributions unfolded with (2+1)D method for three R_g bins for jets with $R = 0.4$ in p+p collisions at $\sqrt{s} = 200$ GeV. Individual panels correspond to different $p_{T,jet}$ intervals (see legend).

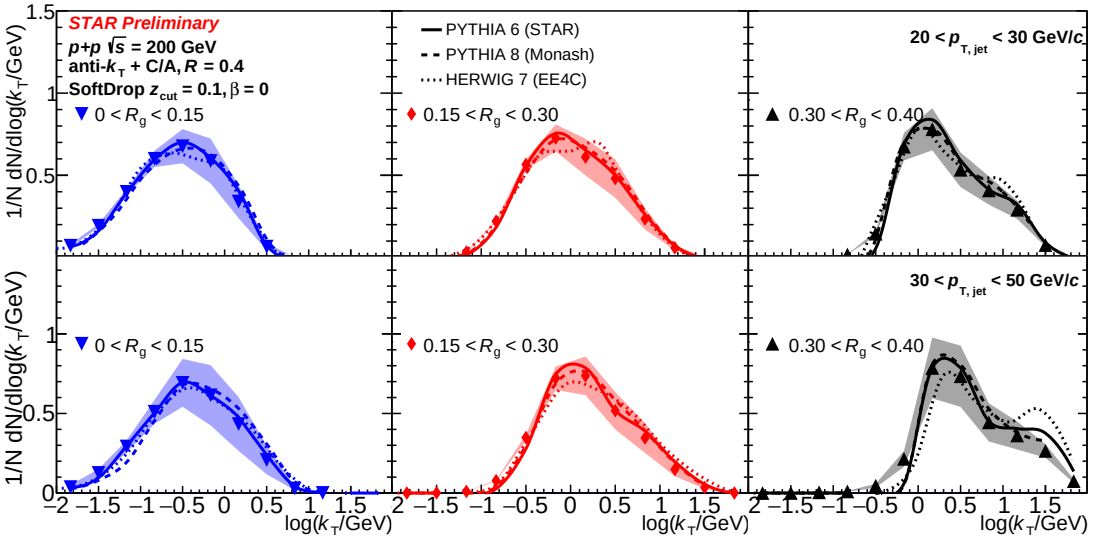

Figure 4: $\log k_T$ distributions unfolded with (2+1)D method for three R_g bins for jets with $R = 0.4$ in p+p collisions at $\sqrt{s} = 200$ GeV. Individual panels correspond to different $p_{T,jet}$ intervals (see legend).

region of the Lund Plane. Our measurements also allow us to disentangle perturbative and mostly non-perturbative dynamics within jet showers at RHIC energies. Data were also compared with different Monte Carlo models and all of them describe the trend of the data well.

Acknowledgments

The work has been supported by the Czech Science Foundation grant 23-07499S.

References

[1] A. J. Larkoski et al. Soft Drop. *JHEP*, 2014(5), 2014.

[2] M. Cacciari et al. The anti-kt jet clustering algorithm. *JHEP*, 2008(04):063, apr 2008.

[3] R. Atkin. Review of jet reconstruction algorithms. *Jour. of Phys: Conf Series*, 645(1):012008, sep 2015.

[4] Frédéric A. Dreyer, Gavin P. Salam, and Grégory Soyez. The Lund jet plane. *Journal of High Energy Physics*, 2018(12), 2018.

[5] Y.-T. Chien, I. Stewart. Collinear drop. *Journal of High Energy Physics*, 2020, 06 2020.

[6] STAR, K. H. Ackermann et al. STAR detector overview. *Nucl. Instrum. Meth. A*, 499:624–632, 2003.

[7] STAR, J. Adam et al. Measurement of Groomed Jet Substructure Observables in p+p Collisions at $\sqrt{s_{NN}}$= 200 GeV with STAR. *Phys. Lett. B*, 811:135846, 2020.

[8] Tim Adye. Unfolding algorithms and tests using RooUnfold. In *PHYSTAT 2011*, Geneva, 5 2011. CERN.

[9] A. Andreassen et al. Omnifold: A method to simultaneously unfold all observables. *Phys. Rev. Lett.*, 124:182001, May 2020.

[10] G. D'Agostini. A multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A*, 362(2-3):487–498, 1995.

[11] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, 2006-05-01.

[12] Peter Z. Skands. Tuning Monte Carlo generators. *Physical Review D*, 82(7), 2010.

[13] J. K. Adkins. Studying transverse momentum dependent distributions in polarized proton collisions via azimuthal single spin asymmetries of charged pions in jets. *Ph.D. Thesis*, 2017.

[14] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, 2015.

[15] M. R. Aguilar et al. Pythia 8 underlying event tune for rhic energies, 2022.

[16] J. Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *The European Physical Journal C*, 76(4), 2016.

[17] P. Skands, S. Carrazza, and J. Rojo. Tuning PYTHIA 8.1. *The European Physical Journal C*, 74(8), 2014.