Measurement of CollinearDrop jet mass and its correlation with substructure observables in pp collisions

(Dated: March 11, 2024)

Jets are collimated sprays of final-state particles produced from initial high-momentum-transfer partonic scatterings in particle collisions. Substructure variables aim to reveal details of the parton fragmentation and hadronization processes that create a jet. By removing collinear radiation while maintaining most of the soft radiation components, one can construct CollinearDrop jet observables, which have enhanced sensitivity to the soft phase space within jets. We present the first CollinearDrop jet measurement, corrected for detector effects with a machine learning method, MultiFold, and its correlation with SoftDrop groomed jet observables. We observe that the amount of grooming affects the angular and momentum scales of the first hard splitting of the jet and is related to the formation time of such splitting. These measurements indicate that the non-perturbative effects are strongly correlated with the perturbative fragmentation process.

Introduction High-energy particle collisions provide opportunities to study experimentally quarks and gluons (partons), the fundamental degree of freedom in the theory of Quantum Chromodynamics (QCD). In some of these collisions, incoming quarks and gluons (partons) interact with each other through the exchange of a high-momentum virtual particle, producing outgoing partons with high transverse momentum (p_{T}). Such outgoing partons are highly virtual and will undergo a series of splitting processes as they come on mass shell. This process is called the parton shower, and can be described perturbatively in terms of the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations [1–3]. When the virtuality of the partons is comparable to the confinement scale Λ_{QCD}, the non-perturbative transition to baryons and mesons (hadrons), known as hadronization, begins. Experimentally, the spray of the final-state hadrons can be measured and clustered into jets. Jets reconstructed with a clustering algorithm [4] can serve as a proxy for the kinematics of the outgoing partons.

While the interaction among partons can be well understood with the principles of perturbative QCD (pQCD), the non-perturbative components of the parton shower and hadronization remain challenging for theoretical calculations and rely mostly on phenomenological models in Monte Carlo event generators. Measurements of observables sensitive to such non-perturbative QCD (npQCD) effects will provide important tests for the theories and constraints on the models. Together with studies of observables calculable from pQCD, investigation of those sensitive to npQCD effects offers an avenue for a comprehensive understanding of the full parton-to-hadron evolution picture.

Beyond the jet p_{T}, or other combinations of the jet four-momentum observables, jet substructure observables [5] are useful tools that can provide insight into the parton shower and hadronization processes. To enhance perturbative contributions, SoftDrop [6] grooming is often used to remove wide-angle soft radiation within the jet. The procedure, detailed in Ref. [6], starts by reclustering the jet with an angular-ordered sequential recombination algorithm called Cambridge/Aachen [7, 8]. Then the last step of the clustering is undone and the softer prong is removed based on the SoftDrop condition:

$$z_{\mathrm{g}} = \frac{\min(p_{\mathrm{T},1}, p_{\mathrm{T},2})}{p_{\mathrm{T},1} + p_{\mathrm{T},2}} > z_{\mathrm{cut}}(R_{\mathrm{g}}/R_{\mathrm{jet}})^{\beta}, \qquad (1)$$

where z_{cut} is the SoftDrop momentum fraction threshold, β is an angular exponent, R_{jet} is the jet resolution parameter, $p_{\mathrm{T},1,2}$ are the transverse momenta of the two subjets, and R_{g} is defined as:

$$R_{\mathrm{g}} = \sqrt{(y_1 - y_2)^2 + (\phi_1 - \phi_2)^2}, \qquad (2)$$

where $y_{1,2}$ and $\phi_{1,2}$ are, respectively, the rapidities and azimuthal angles of the two subjets. z_{g} and R_{g} describe the momentum imbalance and the opening angle of the SoftDrop groomed jet, respectively.

Although the SoftDrop groomed jet substructure observables have been extensively studied both experimentally [9–14] and theoretically [15], the wide-angle and soft radiation which are dominated by npQCD processes, have not yet been explored in detail.

One set of observables that are sensitive to the soft wide-angle radiation are known as CollinearDrop [16]. The general case involves the difference of two different SoftDrop selections $\mathrm{SD}_1 = (z_{\mathrm{cut},1}, \beta_1)$ and $\mathrm{SD}_2 = (z_{\mathrm{cut},2}, \beta_2)$ on the same jet. For nonzero values of SD_1 and SD_2 parameters with $z_{\mathrm{cut},1} \leq z_{\mathrm{cut},2}$ and $\beta_1 \geq \beta_2$, SD_2 aims to reduce the collinear contributions from fragmentation, and SD_1 aims to reduce the wide-angle contributions from initial-state radiation (ISR), underlying event (UE) and pileup.

As the QCD parton shower is angular ordered [17], the soft wide-angle radiation captured by the CollinearDrop jet observables happens on average at an early stage of the shower. Unlike CollinearDrop, SoftDrop then captures the late stage collinear and perturbative splittings. Therefore, a simultaneous measurement of CollinearDrop jet and SoftDrop jet observables can help illustrate the hard-soft dynamics in the parton shower.

The CollinearDrop jet mass is defined in terms of the ungroomed jet mass M and the SoftDrop groomed jet mass M_{g}:

$$M_{(\mathrm{g})} = \left|\sum\nolimits_{i\in(\mathrm{groomed})\ \mathrm{jet}} p_i\right| = \sqrt{E_{(\mathrm{g})}^2 - |\vec{\mathbf{p}}_{(\mathrm{g})}|^2}, \quad (3)$$

where p_i is the four-momentum of the ith constituent in a (groomed) jet, and $E_{(\mathrm{g})}$ and $\vec{\mathbf{p}}_{(\mathrm{g})}$ are the energy and three-momentum vector of the (groomed) jet, respectively. We denote the CollinearDrop groomed jet mass by a:

$$a = \frac{M^2 - M_{\mathrm{g}}^2}{p_{\mathrm{T}}^2}. \quad (4)$$

a is calculable in Soft Collinear Effective Field Theory (SCET) at the parton level [16].

In this paper, we present measurements of the CollinearDrop groomed jet mass, to study the less-explored phase space of soft and wide-angle radiation; we also measure the correlation between the CollinearDrop groomed mass with R_{g} and z_{g}, in pp collisions at $\sqrt{s} =$ 200 GeV at STAR. One notable feature of these measurements is that they are fully corrected for detector effects with MultiFold, a novel machine learning method which preserves the correlations in the multi-dimensional observable phase space on a jet-by-jet basis [18]. We then compare our fully corrected measurements with predictions from event generators and analytical calculations done in the SCET framework.

Analysis details The STAR experiment [19] recorded data from $\sqrt{s} = 200$ GeV pp collisions during the 2012 RHIC run. As energetic charged particles travel from the interaction point to the perimeter of the Time Projection Chamber (TPC), they ionize the gas atoms in the TPC and leave hits, from which we reconstruct tracks. Neutral particles do not interact with the gas in the TPC and instead deposit their energy through the development of electromagnetic showers in Barrel Electro-Magnetic Calorimeter (BEMC) towers. Events are required to pass the jet patch trigger with a minimum transverse energy $E_{\mathrm{T}} > 7.3$ GeV be deposited in a 1×1 patch in $\eta\times\phi$ in the BEMC. Before any run selections, 65M events pass this trigger selection, corresponding to an integrated luminosity of 23 pb^{-1}. In addition, events are required to have primary vertices within ±30 cm from the center of the detector along the beam axis. We apply a 100% hadronic correction to tower energy measurement: if a charged track extrapolates to a tower, then the whole track's p_{T} is removed from the tower E_{T}. The same track and tower selections are applied as in Ref. [11] and [14]. We reconstruct jets from TPC tracks ($0.2 < p_{\mathrm{T}} < 30$ GeV/c, with a charged pion mass assignment) and BEMC towers ($0.2 < E_{\mathrm{T}} < 30$ GeV, assuming massless) using the anti-k_{T} sequential recombination clustering algorithm [4] with a resolution parameter of $R = 0.4$. We apply the selections of $p_{\mathrm{T}} > 15$ GeV/c, $|\eta| < 0.6$, transverse energy fraction of all neutral components < 0.9, and $M > 1$ GeV/c^2 on reconstructed jets, consistent with the selections in Ref. [14]. Similar to Ref. [11] and [14], no background subtraction is done, because the UE contribution to jets is low for STAR kinematics and unfolding can correct for any fluctuation in it. In addition, we select jets that pass SoftDrop grooming with the standard cuts of $(z_{\mathrm{cut}}, \beta) = (z_{\mathrm{cut},2}, \beta_2) = (0.1, 0)$. For this analysis, the less aggressive SoftDrop grooming criteria is set to no grooming, $(z_{\mathrm{cut},1}, \beta_1) = (0, 0)$, so the CollinearDrop groomed observables are the difference in the ungroomed and SoftDrop groomed observables. This simplification can be made since the wide-angle contributions from ISR, UE and pileup are not significant for the dataset used in this analysis. Specifically, the contribution of UE to jet p_{T} for a jet with $20 < p_{\mathrm{T}} < 25$ GeV/c is less than 1% [20].

We measure the following jet observables: p_{T}, z_{g} (defined in Eq. 1), R_{g} (defined in Eq. 2), M (defined in Eq. 3), M_{g} (defined in Eq. 3), and jet charge $Q^{\kappa=2}$. $Q^{\kappa=2}$ is defined as:

$$Q^{\kappa=2} = \frac{1}{p_{\mathrm{Tjet}}^2}\sum\nolimits_{i\in\mathrm{jet}} q_i \cdot p_{\mathrm{T}_i}^2, \quad (5)$$

where q_i and p_{T_i} are the electric charge and p_{T} of the ith jet constituent, respectively.

Experimentally, jet measurements need to be corrected for detector effects to compare with theoretical calculations and model predictions. The traditional correction procedure uses Bayesian inference in as many as three dimensions and requires the observables to be binned based on the resolution [21]. On the other hand, MultiFold [18] is a machine learning technique that is able to correct data at a higher dimensionality in an un-binned fashion. As it preserves the correlation between the input and corrected observables across dimensionality, MultiFold is potentially desirable for this study.

We fully corrected six jet observables simultaneously for detector effects using MultiFold. In addition to jets from data, matched pairs of jets from simulations with (detector-level) and without (particle-level) detector effects are input for MultiFold. The particle-level prior used for unfolding is jets from events generated with PYTHIA6 [22] with the STAR tune [23]. This is a single-parameter modification to the Perugia 2012 tune [24] to better match STAR data. Consistent with [**Dmitri's paper**], at particle-level, hadron weak decays are not enabled while strong and electromagnetic decays are. The PYTHIA events are run through GEANT3 [25] simulation of the STAR detector, and embedded into data from zero-bias events from the same run period as the analyzed data. The detector-level jets are then reconstructed after this embedding procedure. We geometrically match a detector-level jet to a particle-level jet by requiring $\Delta R < 0.4$ between the two in the same event.

MultiFold achieves the goal of unfolding through iteratively reweighting the weights assigned to each jet in simulations [18]. It is naturally unbinned since these weights are per-jet quantities. There are two steps for each iteration. In the first step, a neural network classifier is trained with the binary cross-entropy loss function, to distinguish jets from data and jets from the (reweighted) detector-level simulation. The input to the neural network has as many dimensions as the number of jet observables of interest (in our case, 6), and the output dimension is 2, each of which represents the probabilities that the jet comes from data and from simulation, respectively. It has been shown in Ref. [26] that, the output of such a neural network can be used to estimate a set of new weights to apply to the detector-level simulation (possibly reweighted from the previous iteration). This effectively allows us to convert a high-dimensional reweighting problem to a classification problem. Since the detector-level jets and the particle-level jets are matched, these weights can be applied to the particle-level jets (possibly reweighted from the previous iteration) as well. However, due to the stochastic nature of detector response, identical particle-level jets are likely to match to different detector-level jets. A second step is then needed to convert these "per-instance" [18] (where each instance is a detector-level and particle-level pair) weights to a function that gives a unique prescription to any particle-level jet. These weights obtained from the second step are then either applied to the detector-level and particle-level jets in the next iteration, or quoted as the final prescription to obtain the unfolded jets if it is the last iteration.

We utilize the default settings of MultiFold as in [18], with two dense neural networks, each with three dense layers and 100 nodes per layer. We train the neural networks with TensorFlow [27] and Keras [28] using the Adams optimization algorithm [29]. In addition, we also use the default setting for the choice of activation functions, loss function, fraction of sample size for validation, and maximum number of epochs. To prevent over-training, an early stopping is implemented after 50 consecutive epochs in which the loss value for the validation sample has not improved.

To correct for fake jets, i.e., detector-level jets arising from background, fake rates were obtained from simulations and used as initial weights for the data as an input to MultiFold. For particle-level jets that are missed at detector level due to effects such as tracking inefficiency, an efficiency correction was done post-unfolding in a multi-dimensional fashion.

The correction procedure was validated using a Monte Carlo closure test, which showed good performance of the unfolding among all observables for jets with $20 < p_{\mathrm{T}} < 50$ GeV/c. In addition, we compared the fully corrected jet mass distributions for three different p_{T} bins, using both MultiFold and RooUnfold [14]. The ratios of MultiFold distributions over RooUnfold distributions are confirmed to be consistent with unity. These establish further confidence in application of MultiFold to the data.

The statistical uncertainty is estimated with the bootstrap technique [30]. In particular, 50 pseudo-datasets are created and used to repeat the unfolding procedure, where each jet from data has been resampled from a Poisson distribution with a mean of 1.

The sources of systematic uncertainties are variations of hadronic correction scale (from 100% to 50%), tower energy resolution (varied by 3.8%), tracking efficiency (varied by 4%) and unfolding procedure. The first three sources are treated in the same way as Ref. [11] and [14]. The dominant source for systematic uncertainty is the variation of unfolding procedure, up to x% in the peak region for jets in $20 < p_{\mathrm{T}} < 30$ GeV/c, and y% for jets in $30 < p_{\mathrm{T}} < 50$ GeV/c. The unfolding variation includes variation of the prior and random seed. The prior variation is accounted for through simultaneous reweighting of all six unfolded observables as well as a, based on prior distributions from PYTHIA [31] and HERWIG [32]. The variation of the random seed affects the initialization of the weights of the neural networks, and is accounted for with the standard error on the fully corrected result obtained from 100 different initial seeds.

Different from analyses that use RooUnfold, Ref. [11] and [14], this analysis does not explicitly account for the variation of the number of iterations as a separate source of uncertainty. Going to a higher number of iterations reduces the prior dependence bias; in fact, mathematically, the most correct number of iterations is infinity [18]. However, the statistical limitations would introduce unwanted fluctuations at such high number of iterations [18]. This can manifest through a large uncertainty from the variation of initial seeds, as well as the statistical uncertainty obtained with the bootstrap technique. The deviation of the result due to not able to perform an infinite number of iterations shows up as the prior dependence. Therefore, the prior variation uncertainty effectively accounts for the uncertainty due to the number of iterations not being ideal, and the number of iterations can be selected by considering when a) the prior dependence uncertainty, b) seed uncertainty, and c) statistical uncertainty are low. We select an iteration number of 15, low enough such that the uncertainty due to seed variation and statistical uncertainty are both reasonable, at the cost of a non-negligible prior dependence uncertainty.

Results Figure 1 shows the distribution of fully corrected CollinearDrop groomed jet masses for jets with $20 < p_{\mathrm{T}} < 30$ GeV/c and $30 < p_{\mathrm{T}} < 50$ GeV/c. This measurement excludes jets with $M = M_{\mathrm{g}}$ (?% of jets in this $20 < p_{\mathrm{T}} < 30$ GeV/c and ?% of jets in $30 < p_{\mathrm{T}} < 50$ GeV/c) so that the peak in the small but nonzero a region is visible. The $M = M_{\mathrm{g}}$ case corresponds to the jets whose first splittings pass the criterion of $(z_{\mathrm{cut}}, \beta) = (0.1, 0)$ without the need of SoftDrop grooming, because the lower-p_{T} prong of the splitting carries at least 10% of the total jet p_{T}. We observe that the data do not show a p_{T} dependence of a. Comparisons with event generator descriptions are shown in dashed lines, with vertical error bars indicating statistical uncertainties. Both PYTHIA6 STAR tune [23] and HERWIG 7.2.2 [32] capture the qualitative trend of data, although there is some tension with PYTHIA 8.303 with Detroit tune [31] (finalize after systematics are done).

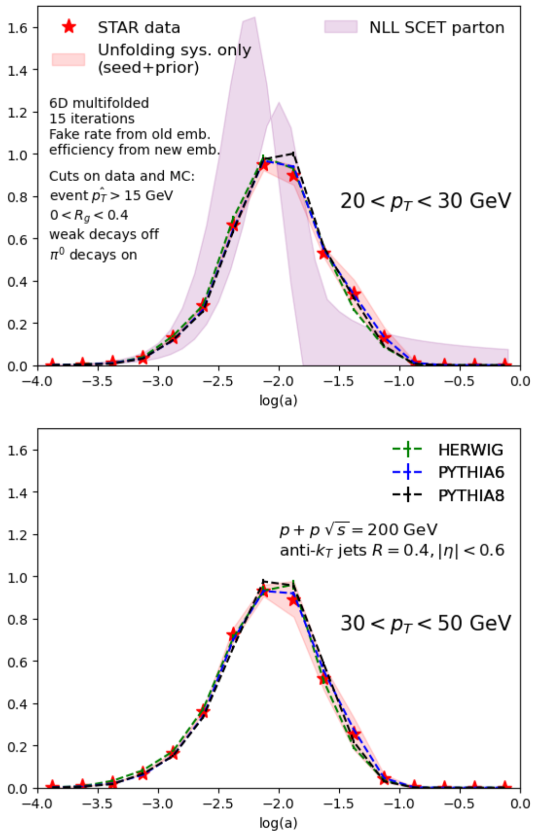

FIG. 1. CollinearDrop jet mass distributions.

Analytic calculation with NLL SCET performed at the parton level shows deviation from both event generator predictions and data, indicating that the CollinearDrop groomed mass is sensitive to hadronization effects. The error band indicates typical scale variations in theoretical calculation.

Figure 2 shows the correlation between a and the SoftDrop groomed shared momentum fraction $z_{\rm g}$ and the SoftDrop groomed jet radius $R_{\rm g}$ in $20 < p_{\rm T} < 30$ GeV/c, where the average value of the CollinearDrop groomed jet mass is indicated by the color of each bin in the $z_{\rm g} - R_{\rm g}$ plane. The $M = M_{\rm g}$ jets are included in this plot. This plane captures the Lund Plane of the first groomed splitting. We see that a is strongly correlated with $R_{\rm g}$ while weakly correlated with $z_{\rm g}$.

Also shown in Fig. 2 is curves of constant formation time t, which approximates the time it takes for a parton to radiate a gluon. This can be estimated as the life-time of the parton using the Heisenberg uncertainty principle [17]. It is related to other parton kinematic variables by:

$$t = \frac{1}{2Ez(1-z)(1-\cos(\theta))}, \tag{6}$$

where E is the energy of the parent parton, z is the momentum fraction carried by the lower-$p_{\rm T}$ daughter parton, and θ is the opening angle between the two daughter. E can be approximated by the jet $p_{\rm T}$; for a perturbative hard splitting, z and θ can be approximated by the SoftDrop $z_{\rm g}$ and $R_{\rm g}$, respectively [11]. We obtain the curves shown by replacing the parton variables in Eq. 6 with their (SoftDrop) jet counterparts, so t can be interpreted as the time that the first hard splitting to pass the SoftDrop criterion takes to develop. The strong correlation between a and $R_{\rm g}$ can therefore be understood as how the amount of early-stage radiation affects when the hard splitting happens. Specifically, to shed a significant amount of mass at the early stage of the parton shower, which is predominantly done via soft gluon radiation, the hard splitting needs to happen relatively late on average.

It is worth emphasizing that the measurement shown Fig. 2 showcases the power of MultiFold, which enabled us to make selections in three variables, $p_{\rm T}$, $z_{\rm g}$ and $R_{\rm g}$, and study a fourth one a which itself is composite of a few variables; all of these observables have been fully corrected for detector effects.

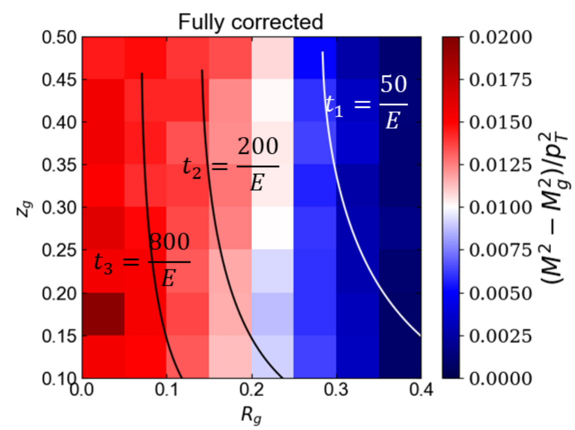

FIG. 2. CollinearDrop groomed mass as a function of $z_g - R_g$

Figure 3 shows the log(a) distributions for specific regions of the $z_{\rm g} - R_{\rm g}$ plane. The leftmost bin includes the $a = 0$ jets, which do not have anything removed by SoftDrop and are therefore possibly dominated by jets whose first splittings in the parton shower are already perturbative. Region 3 ($0.15 < R_{\rm g} < 0.25$ and $0.1 < z_{\rm g} < 0.2$) includes asymmetric and intermediate-angle splittings while Region 2 ($0.15 < R_{\rm g} < 0.25$ and $0.4 < z_{\rm g} < 0.5$) includes symmetric and intermediate-angle splittings. Despite the different $z_{\rm g}$ selections, the fraction of $a = 0$ jets and the distributions in $a > 0$ are similar. The weak dependence of a on $z_{\rm g}$ is consistent with our observation made for Fig. 2.

However, as we continue to scan across the plane, we notice drastic changes in the fraction of jets with $a = 0$ as well as differences in shape in the $a > 0$ region. We first move onto Region 1 ($0 < R_{\rm g} < 0.1$ and $0.4 < z_{\rm g} < 0.5$), which includes symmetric and collinear radiation. Fig. 3 also shows that, compared to Regions 2 and 3, Region 1 is more likely to have soft radiation groomed away by SoftDrop as indicated by the decreased count for $a = 0$, and has a broader tail for the small but nonzero a region. On the other hand, we observe from Fig. 2 that we have on average higher values of a in this region, which can be understood as mostly affected by the slightly higher values in $\log(a) > -1.5$. The distribution of $\log(a)$ is wider in both directions arises from that a selection of narrow hard splitting opens up a large phase space for the amount of radiation preceding the splitting.

Region 4 ($0.3 < R_{\rm g} < 0.4$ and $0.1 < z_{\rm g} < 0.2$) includes asymmetric and wide-angle splittings, characteristic of perturbative early emissions. Again compared to Regions 2 and 3, in Region 4, the significant fraction of $a = 0$ jets indicates that it is highly probable that no non-perturbative early emission has happened before the perturbative emission. This is likely the explanation for why the z-axis values are also close to 0 in this region in Fig. 2.

Conclusions In this Letter, we have presented the first CollinearDrop groomed observable measurement, the CollinearDrop groomed mass, and its correlations with groomed jet substructure observables, in pp collisions at $\sqrt{s} = 200$ GeV with the STAR experiment. A machine learning driven method to correct for detector effects, MultiFold, has been applied for the first time to hadronic collision data, which allows for access of multi-dimensional correlations on a jet-by-jet basis. We demonstrate how MultiFold allows us to present measurements in 4 dimensions and shows promising potential for future multi-differential measurements as the community enters high-statistics, precision QCD era.

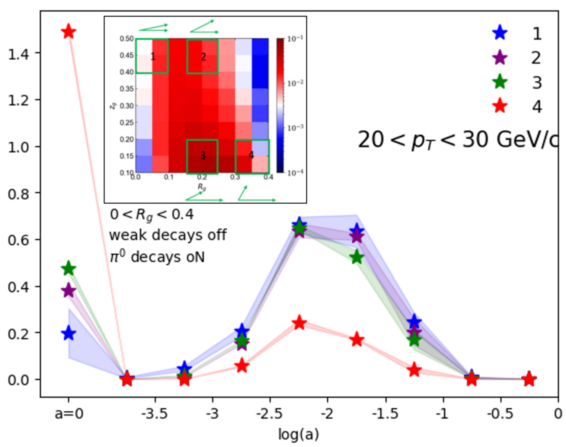

FIG. 3. Distribution of $\log(a)$ with various selections of $R_{\rm g}$ and $z_{\rm g}$.

Event generator predictions and theoretical calculation were shown to qualitatively describe the data for the CollinearDrop groomed mass, which probes the soft radiation within jets. From the investigation of the correlation between the CollinearDrop groomed mass a and the SoftDrop groomed observables $z_{\rm g}$ and $R_{\rm g}$, we observe that on average, a large nonperturbative radiation biases the perturbative splitting to happen late. We also observed a strong correlation between the CollinearDrop groomed mass and $R_{\rm g}$. In particular, a large $R_{\rm g}$ biases toward a higher probability that the jet has no radiation prior to the perturbative splitting, and a small $R_{\rm g}$ biases towards a higher probability that the jet has some radiation prior to the splitting. These measurements demonstrate the interplay between the nonperturbative processes and the perturbative jet fragmentation.

[1] V. N. Gribov and L. N. Lipatov, Deep inelastic ep scattering in perturbation theory, Sov. J. Nucl. Phys. **15**, 438 (1972).

[2] G. Altarelli and G. Parisi, Asymptotic freedom in parton language, Nuclear Physics B **126**, 298 (1977).

[3] Y. L. Dokshitzer, Calculation of the structure functions for Deep Inelastic Scattering and e^+e^- annihilation by perturbation theory in Quantum Chromodynamics, Sov. Phys. JETP **46**, 641 (1977).

[4] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k_t jet clustering algorithm, Journal of High Energy Physics **2008**, 063 (2008).

[5] S. Marzani, G. Soyez, and M. Spannowsky, *Looking Inside Jets* (Springer International Publishing, 2019).

[6] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, Soft Drop, JHEP **05**, 146, arXiv:1402.2657 [hep-ph].

[7] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, Better jet clustering algorithms, JHEP **08**, 001, arXiv:hep-ph/9707323.

[8] M. Wobisch and T. Wengler, Hadronization corrections to jet cross-sections in deep inelastic scattering, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)* (1998) pp. 270–279, arXiv:hep-ph/9907280.

[9] A. M. Sirunyan *et al.* (CMS), Measurement of the splitting function in pp and Pb-Pb collisions at $\sqrt{s_{\rm NN}} =$ 5.02 TeV, Phys. Rev. Lett. **120**, 142302 (2018), arXiv:1708.09429 [nucl-ex].

[10] G. Aad *et al.* (ATLAS), Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV, Phys. Rev. D **101**, 052007 (2020).

[11] J. Adam, L. Adamczyk, J. Adams, J. Adkins, G. Agakishiev, M. Aggarwal, Z. Ahammed, I. Alekseev, D. Anderson, A. Aparin, *et al.* (STAR Collaboration), Measurement of groomed jet substructure observables in p+p collisions at $\sqrt{s} = 200$ GeV with STAR, Physics Letters B **811**, 135846 (2020).

[12] M. Aaboud *et al.* (ATLAS), Measurement of the Soft-Drop jet mass in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector, Phys. Rev. Lett. **121**, 092001 (2018),

arXiv:1711.08341 [hep-ex].

[13] A. M. Sirunyan *et al.* (CMS), Measurement of the groomed jet mass in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} =$ 5.02 TeV, JHEP **10**, 161, arXiv:1805.05145 [hep-ex].

[14] M. Abdallah *et al.* (STAR), Invariant jet mass measurements in *pp* collisions at $\sqrt{s} = 200$ GeV at RHIC, Phys. Rev. D **104**, 052007 (2021).

[15] A. J. Larkoski, I. Moult, and B. Nachman, Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning, Phys. Rept. **841**, 1 (2020), arXiv:1709.04464 [hep-ph].

[16] Y.-T. Chien and I. W. Stewart, Collinear drop, Journal of High Energy Physics **2020**, 1 (2020).

[17] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian, *Basics of perturbative QCD* (1991).

[18] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman, and J. Thaler, OmniFold: A method to simultaneously unfold all observables, Phys. Rev. Lett. **124**, 182001 (2020).

[19] K. H. Ackermann *et al.* (STAR), STAR detector overview, Nucl. Instrum. Meth. A **499**, 624 (2003).

[20] J. Adam, L. Adamczyk, J. Adams, J. Adkins, G. Agakishiev, M. Aggarwal, Z. Ahammed, I. Alekseev, D. Anderson, A. Aparin, *et al.* (STAR Collaboration), Underlying event measurements in p+p collisions at $\sqrt{s} = 200$ GeV at RHIC, Physical Review D **101**, 052004 (2020).

[21] G. D'Agostini, Improved iterative bayesian unfolding (2010), arXiv:1010.0632 [physics.data-an].

[22] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP **05**, 026, arXiv:hep-ph/0603175.

[23] 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, Kentucky U. (2015), arXiv:1907.11233 [hep-ex].

[24] P. Z. Skands, Tuning Monte Carlo generators: The Perugia tunes, Phys. Rev. D **82**, 074018 (2010), arXiv:1005.3457 [hep-ph].

[25] R. Brun, F. Bruyant, F. Carminati, S. Giani, M. Maire, A. McPherson, G. Patrick, and L. Urban, GEANT detector description and simulation tool 10.17181/CERN.MUHF.DMJ1 (1994).

[26] A. Andreassen and B. Nachman, Neural networks for full phase-space reweighting and parameter tuning, Phys. Rev. D **101**, 091901 (2020).

[27] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.

[28] F. Chollet *et al.*, Keras, `https://keras.io` (2015).

[29] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization (2017), arXiv:1412.6980 [cs.LG].

[30] B. Efron, Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics **7**, 1 (1979).

[31] M. R. Aguilar, Z. Chang, R. K. Elayavalli, R. Fatemi, Y. He, Y. Ji, D. Kalinkin, M. Kelsey, I. Mooney, and V. Verkest, Pythia 8 underlying event tune for rhic energies, Physical Review D **105**, 016011 (2022).

[32] J. Bellm, G. Bewick, S. Ferrario Ravasio, *et al.*, Herwig 7.2 release note, The European Physical Journal C **80**, 452 (2020).