PYTHIA8 underlying event tune for RHIC energies

Manny Rosales Aguilar,[1] Zilong Chang,[2] Raghav Kunnawalkam Elayavalli,[3,2] Renee Fatemi,[1] Yang He,[4] Yuanjing Ji,[5] Dmitry Kalinkin,[6] Matthew Kelsey,[7,*] Isaac Mooney,[7] and Veronica Verkest[7]

[1]*University of Kentucky, Lexington, Kentucky 40506, USA*
[2]*Brookhaven National Laboratory, Upton, New York 11973, USA*
[3]*Yale University, New Haven, Connecticut 06517, USA*
[4]*Shandong University, Qingdao, Shandong 266237, China*
[5]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[6]*Indiana University, Bloomington, Indiana 47408, USA*
[7]*Wayne State University, Detroit, Michigan 48202, USA*

(Received 18 October 2021; accepted 14 December 2021; published 11 January 2022)

We report an underlying event tune for the PYTHIA8 Monte Carlo event generator that is applicable for hadron collisions primarily at $\sqrt{s}$ ranges available at the Relativistic Heavy-Ion Collider (RHIC). We compare our new PYTHIA8 tuned predictions to midrapidity inclusive $\pi^{\pm}$ spectra, jet substructure, Drell-Yan production, and underlying event measurements from RHIC and the Tevatron, as well as underlying event data from the Large Hadron Collider. With respect to the default PYTHIA8 Monash Tune, the new "Detroit" tune shows significant improvements in the description of the experimental data. Additionally, we explore the validity of PYTHIA8 predictions for forward rapidity π in $\sqrt{s} = 200$ GeV collisions, where neither tune is able to sufficiently describe the data. We advocate for the new tune to be used for PYTHIA8 studies at current and future RHIC experiments, and discuss future tuning exercises at lower center-of-mass energies, where forward/backward kinematics are essential at the upcoming Electron-Ion Collider.

DOI: 10.1103/PhysRevD.105.016011

I. INTRODUCTION

Monte Carlo (MC) event generators that simulate relativistic lepton-lepton, lepton-hadron, and hadron-hadron collisions are an essential part of high-energy particle and nuclear physics. From a theoretical perspective, MC models are able to test our fundamental understanding of the Standard Model, in particular quantum chromodynamics (QCD), and offer a prescription for the initial and final states of the collision system. From the experimental point of view, MCs are an integral piece in the simulation chain that aims to reproduce realistic spectra that are ultimately used to extract detector acceptance and resolution corrections, and study systematic effects in experimental data. There are several event generators that are currently available, for example PYTHIA [1–3] and Herwig/Herwig++ [4,5], that have been widely used to simulate collisions at the Large Hadron Collider (LHC) and the Relativistic Heavy-Ion Collider (RHIC). The simulation routines include various QCD physics processes that factorize a single collision into two regimes consisting of the perturbative hard scattering and evolution via a parton shower, and various nonperturbative components such as hadronization, underlying event and multiparton interactions.

Physics processes implemented in event generators include multiple parameters which are turned to experimental measurements, often from $e^+ + e^-$ collisions. The PYTHIA event generator studied in this publication has been very successful in describing data at the LHC and has been the subject of many tuning exercises over the past decades [6–13]. While the global tuning of PYTHIA6 and 8 in Refs. [6,7] are in good agreement with data at LHC energies, there are significant discrepancies in describing data from collisions at lower center-of-mass energies [14,15]. These disagreements can mainly be understood as a consequence of incorrect modeling of the soft QCD underlying event (UE) stemming from the center-of-mass energy extrapolation that is used. A PYTHIA6 "STAR" tune was produced in Ref. [16] which updated the value of the parameter controlling the energy extrapolation of the low transverse momentum (p_{T}) cross section, and was able to adequately describe the $\pi^{\pm,0}$ p_{T} and jet p_{T} spectrum in proton-proton ($p + p$) collisions at $\sqrt{s} = 200$ GeV from STAR and PHENIX as demonstrated in Refs. [14,17]. For PYTHIA8, there has been several tuning exercises with LHC

*mkelsey@wayne.edu

data at $\sqrt{s} = 7$ and 13 TeV collisions, and some lower collision energy data either from the LHC or the Tevatron [9,10,12]. The LHC-focused tune produced in [10] utilized LHC data from $\sqrt{s} = 900$ GeV and 7 TeV $p + p$ collisions, and it was argued therein that including Tevatron data within the tuning procedure leads to inadequate tune performance at LHC energies. In contrast, measurements of the UE multiplicity from CDF [18] at $\sqrt{s} = 900$ and 1960 GeV and LHC data at the top energies $\sqrt{s} = 7$ and 13 TeV were used in this [12] tuning exercise to control the impact of energy extrapolation. However, the aforementioned tune was not able to describe the CDF UE data at $\sqrt{s} = 300$ GeV. The energy dependence of low p_{T} regularization was also studied in [9] using the same LHC data and CDF data at $\sqrt{s} = 300$, 900, and 1960 GeV, and concluded that the inclusion of a new term in the extrapolation produced improved agreement at lower energies. All these observations motivate the need for a dedicated PYTHIA8 tune, in addition to the existing PYTHIA6 STAR tune, for the nominal RHIC energy of $\sqrt{s} = 200$ GeV.

The organization of the following sections is as follows: Sec. II describes the general tuning procedure via the Professor toolkit [11]; Sec. III is dedicated to describing the data and RIVET [19] implementation; Sec. IV presents our tuned results; Sec. V compares the new PYTHIA8 predictions to selected data distributions; in Sec. VI we compare the default and tuned distributions at forward rapidity; and Sec. VII summarizes our results. The Appendix presents the comparison of our new tune to all the experimental data.

II. TUNING PROCEDURE

We utilize a parametrization-based approach for the tuning procedure provided by the Professor (v2.3.3) [11] toolkit. In the Professor tuning methodology, PYTHIA parameters of interest are sampled n times across a provided range and for each sampling a MC generation is produced and the resulting prediction is compared to data. In each bin of data, the result of the random samplings are parameterized by a third order polynomial. In the case of N PYTHIA tuning parameters, the corresponding polynomials are N-dimensional. The coefficients of the polynomials are computed numerically within the Professor code. A χ^2 fit of the polynomial parameterizations to the data is performed using Minuit [20,21] to determine the best values of each tuning parameter.

The starting point of our tuning exercise is the default Monash tune [7] and PYTHIA8.303. The NNPDF2.3 parton distribution functions (PDF) [22] used in the Monash tune have since been updated with improved data and methods, and therefore we utilize the recent NNPDF3.1 leading-order PDF set with $\alpha_s(m_Z) = 0.130$ [23]. The settings that are varied in our tune are $p_{T,0}^{Ref}$ and its energy-dependence scaling parameter, the proton matter distribution parameters, and color reconnection range. The $p_{T,0}^{Ref}$ parameters regularizes the low p_T cross section divergence, and is a key parameter in all PYTHIA tunes. The energy-scaling of $p_{T,0}$ follows a power-law function, and is controlled by the ecmPow parameter. We change the reference energy that corresponds to the $p_{T,0}^{Ref}$ parameter to be 200 GeV. We do so for two reasons: (1) using a different PDF set requires a complete retune of $p_{T,0}^{Ref}$; and (2) to control the power-law extrapolation as much as possible at low center-of-mass energies due to the rapidly varying functional form in this region. For the proton shape function, we change it to the double Gaussian matter profile (MultipartonInteractions: bProfile = 2) and vary its two respective parameters, coreRadius and coreFraction. Table I tabulates all the tuning parameters and their respective ranges. From the previous PYTHIA6 "STAR" tune study in [16], the values for $p_{T,0}^{Ref}$ and ecmPow are expected to be smaller than the default values, and therefore their ranges are chosen to cover all values lower than the default with some overlap. The other tuning parameter ranges cover all possible values. A complete description of the above tuning parameters are given in [1–3].

We sample 70 values of the tuning parameters in Table I within the specified ranges using the Professor package to be used as anchor points in the generator response polynomial parametrization. We have checked that using a smaller fraction (but still sufficient for 5 tuning parameters) of the sampled anchor points produces compatible results with respect to the full sample. We generate 10 million events for each simulation run to ensure the MC statistics are sufficiently less than that of the data in the region of interest. The tuned results are determined by minimizing the weighted χ^2,

$$\chi^2 = \sum_i w_i (F[i] - d_i) C^{-1} (F[i] - d_i), \tag{1}$$

where the index i runs over all data points d_i, w_i is a weight for each data point, $F[i]$ is the parametrized PYTHIA prediction for the ith data point, and C^{-1} is the inverse covariance matrix which contains only experimental uncertainties (assumed no bin-by-bin correlations). For the final

TABLE I. PYTHIA8 settings and tuning parameters.

Setting	Default	New
PDF:pSet	13	17
MultipartonInteractions:ecmRef	7 TeV	200 GeV
MultipartonInteractions:bprofile	3	2
Tuning Parameter	**Default**	**Range**
MultipartonInteractions:pT0Ref	2.28 GeV	0.5–2.5 GeV
MultipartonInteractions:ecmPow	0.215	0.0–0.25
MultipartonInteractions:coreRadius	0.4	0.1–1.0
MultipartonInteractions:coreFraction	0.5	0.0–1.0
ColourReconnection:range	1.8	.0–9.0

TABLE II. Midrapidity data used in the tuning procedure.

Experiment	$\sqrt{s}$ (GeV)	Observable	Reference	Figure
STAR	200	$\pi^{\pm}$ cross sections vs p_T	[24]	3, 14
PHENIX	200	Dimuon pairs from Drell-Yan vs di-muon p_T	[25]	16
STAR	200	Average charged particle multiplicities and p_T vs leading jet p_T in the forward, transverse, and away regions	[15]	3, 8, 9
CDF	300, 900, 1960	Charge particle density and $\sum p_T$ vs leading hadron p_T in transverse region	[18]	10, 11, 12
STAR	200	SoftDrop groomed jet substructure (z_g and R_g)	[26]	15
STAR	200	Inclusive and groomed jet mass	[27]	3, 15

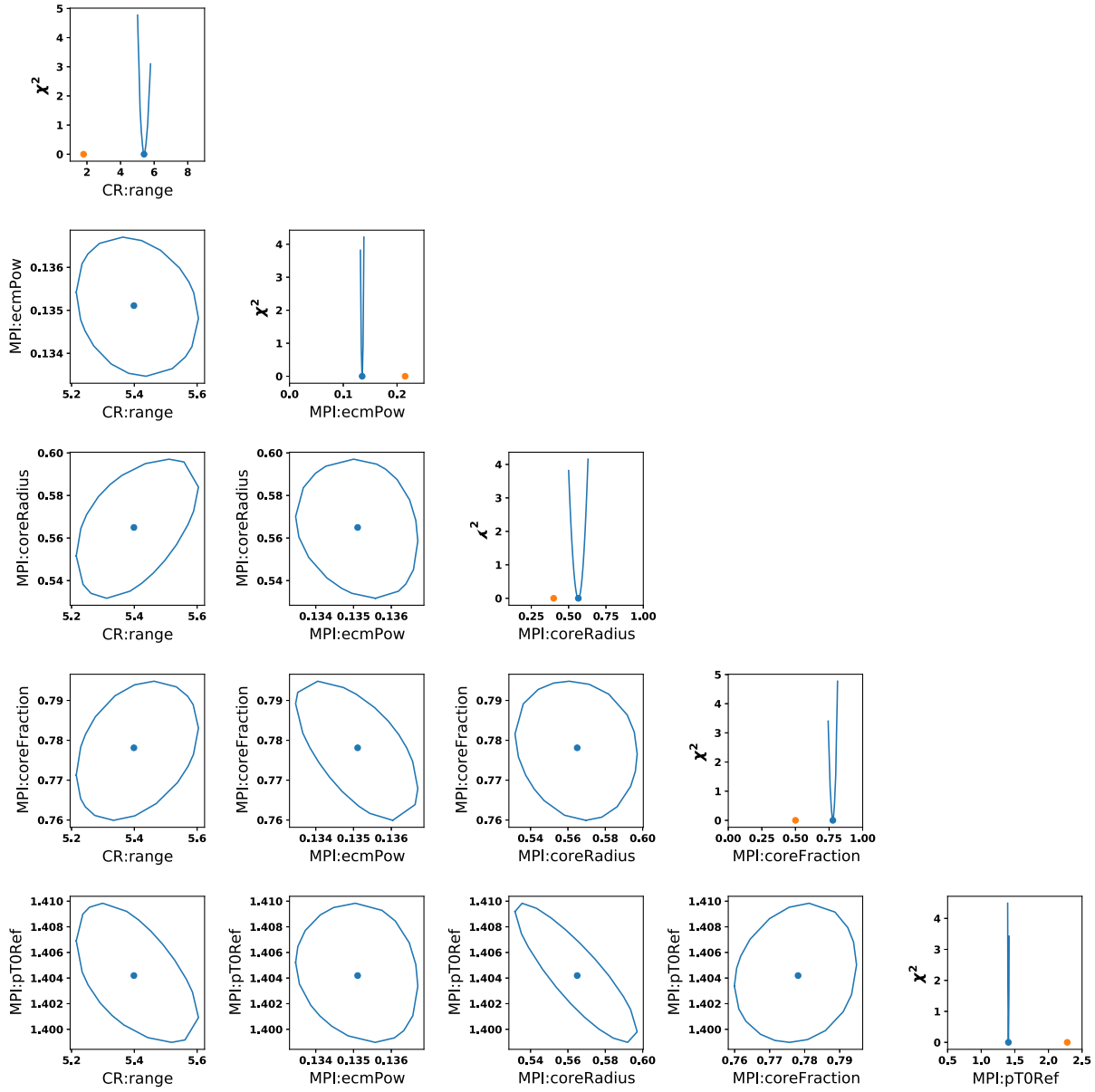

FIG. 1. Minimum-subtracted χ^2 profiles in the vicinity of the best fit value, shown in the right-most diagonal panels, and the one sigma correlation contours between all tuning parameters, shown on the off-diagonal panels. The single (orange) points along the diagonal panels show the default PYTHIA8.303 values tuned at a reference energy of 7 TeV. Additionally, the x-axis scale in the diagonal panels show the allowed ranges used in the tuning procedure.

TABLE III. PYTHIA8 tuned parameters.

Tuning Parameter	Default	Detroit
MultipartonInteractions:pT0Ref	2.28 GeV	1.40 GeV
MultipartonInteractions:ecmPow	0.215	0.135
MultipartonInteractions:coreRadius	0.4	0.56
MultipartonInteractions:coreFraction	0.5	0.78
ColourReconnection:range	1.8	5.4

fit we do not assign any weights beyond unity to any of the data.

III. INPUT DATA AND MC GENERATION

Midrapidity data from STAR, PHENIX, and CDF are utilized in this tuning exercise. The measurements can be grouped in three main categories; identified particle spectra, event multiplicities, and jet substructure as described in Table II. Collectively, they broadly sample the available phase space and test the model's ability to describe the nonperturbative and perturbative parts of a $p+p$ collision. The spectra measurements focus on $\pi^{\pm}$ yields at STAR [24] and Drell-Yan di-muon pair production from PHENIX [25]. We include measurements of the UE event multiplicity from STAR [15] and CDF [18] at various regions such as toward, away, and transverse with respect to the trigger object, be it a reconstructed jet or charged hadron, respectively. Lastly, two sets of jet substructure measurements from STAR [26,27] on the SoftDrop splitting observables at the first split and the invariant and groomed jet mass, all measured differentially as a function of the jet p_{T} and jet resolution parameter R, are included in the tuning exercise. The substantial impact on jet substructure observables from the variation of the MPI and UE parameters is an indication of the multifaceted inner workings of the PYTHIA event generator. The last column of Table II mentions the respective figure in both the paper draft and in the Appendix where the comparisons with the data are shown.

We prepared RIVET analyses for each of the measurements mentioned above (when analyses were not publicly available) and they are all made available here github.com/star-bnl/star-PYTHIA8-tune.

The various MC runs are all analyzed by the RIVET analyses and the resulting output yoda files are processed by the Professor toolkit to determine the minimum χ^2. We exclude the $\pi^{\pm}$ cross section below $p_T < 1$ GeV/c to avoid potential feed-down effects which we studied by turning weak decays on or off. We also remove data from the fit in which the envelope of the PYTHIA predictions are significantly varied compared to the data systematic uncertainties. Similarly, data points where our MC statistical uncertainties are still too large, for example the very high p_{T} bins, are also excluded to optimize the tuning but we note that it is only applicable to a few regions of phase space of our observables. We explicitly turn off long-lived particle decays in our MC generation as all relevant observables have either been corrected for feed-down decays or unfolded to particle level distributions.

IV. NEW PYTHIA8 DETROIT TUNE

Figure 1 show the χ^2 profiles projected onto each tuning parameter in the vicinity of the global minimum (offset to have minimum at zero), and the one sigma correlation contours between all tuning parameters. The central values of the tuning parameters are tabulated in Table III, and define the new "Detroit" tune. In general the errors on the central values from Minuit are much smaller than the eigentune values quoted below, and are therefore not given. The χ^2 per degrees of freedom ($n.d.f.$) for the best fit is $\chi^2/\mathrm{n.d.f.} = 611/493$.

In Fig. 2 we plot the values for $p_{T,0}^{Ref}$, at their respective reference energies, and their extrapolations for the Monash, CMS CP1 [12], and Detroit tunes. The extrapolations are determined as $p_{T,0} = p_{T,0}^{\mathrm{Ref}}(\sqrt{s}/\sqrt{s}_{\mathrm{Ref}})^{\mathrm{ecmPow}}$. We compare to the CMS CP1 tune as it followed a similar tuning strategy utilizing both LHC data and the CDF UE data at $\sqrt{s} = 900$ and 1960 GeV, and the same PDFs. However, in the CMS CP1 tune the reference energy is kept at 7 TeV. The new $p_{T,0}^{\mathrm{Ref}}$ value at a reference energy of $\sqrt{s} = 200$ GeV is about 30% larger compared to the $\sqrt{s} = 200$ GeV extrapolated $p_{T,0}$ using the Monash values, and comparable to the extrapolated values in the CMS CP1 tune. However, in the latter comparison the values of $p_{T,0}$ diverge as you go up or down in collision energy. The value for the energy-depended extrapolation reduced by almost 40% compared to the Monash tune, which significantly reduces how fast $p_{T,0}$ varies. Compared to the CMS CP1 tune, our new extrapolation parameter is roughly 10% smaller.

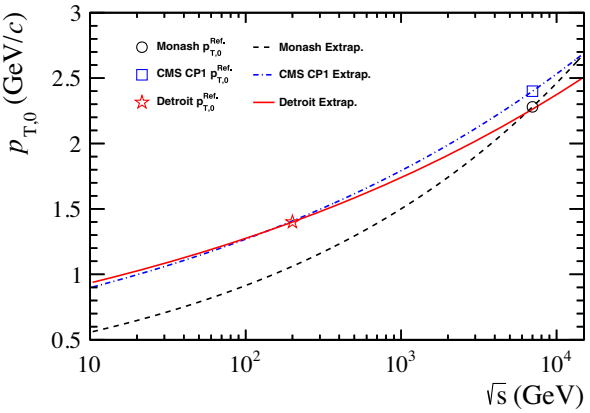

FIG. 2. Values for $p_{T,0}^{\mathrm{Ref}}$ for the Monash (open black circle) [7], CMS CP1 (open blue square) [12], and Detroit (open red star) tunes. The energy extrapolations for the Monash, CMS CP1, and Detroit tunes are shown as the dashed black, dot-dashed blue, and solid red lines, respectively.

TABLE IV. PYTHIA8 tune parameter variations for each eigentune.

Tuning Parameter	1+	1−	2+	2−	3+	3−	4+	4−	5+	5−
MultipartonInteractions:pT0Ref (GeV)	1.37	1.43	1.38	1.42	1.44	1.37	1.41	1.40	1.40	1.41
MultipartonInteractions:ecmPow	0.132	0.138	0.135	0.135	0.119	0.150	0.145	0.126	0.148	0.125
MultipartonInteractions:coreRadius	0.74	0.41	0.77	0.41	0.57	0.56	0.57	0.56	0.51	0.60
MultipartonInteractions:coreFraction	0.84	0.72	0.72	0.82	0.78	0.78	0.78	0.78	0.60	0.90
ColourReconnection:range	7.50	3.61	5.38	5.41	5.40	5.40	5.40	5.40	5.41	5.40

We observe that the proton overlap function parameters coreRadius and coreFraction, and color r econnection range to have slightly increased values in our tune compared to the default PYTHIA8 and CMS CP1 values as presented in Table III. We have also performed the tuning study with the default proton overlap shape function (MultipartonInteractions:bprofile = 3) and find the description of the $\sqrt{s} = 300$ GeV UE data to be inadequate, and had a global χ^2 more than a factor of two larger with respect to the Detroit tune.

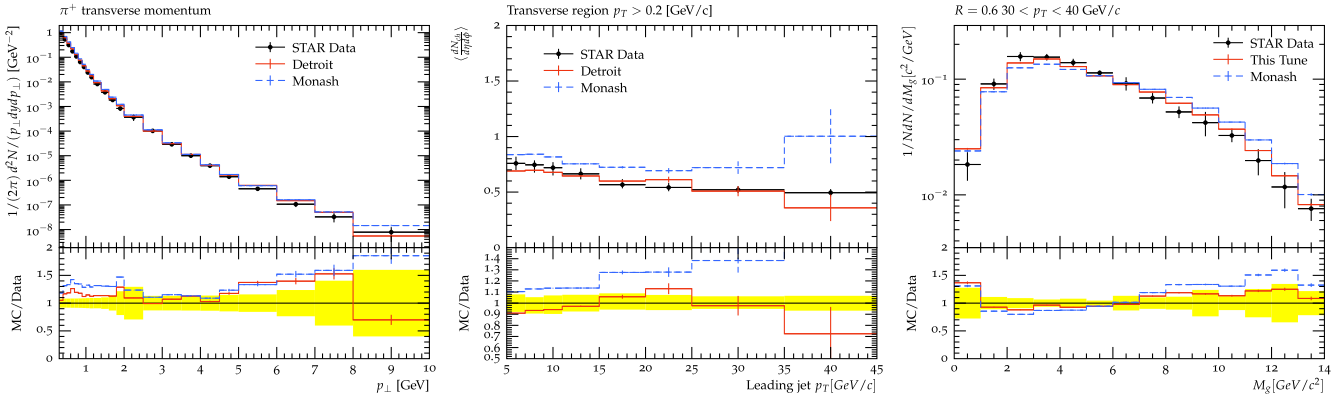

FIG. 3. Comparison of the default (blue dashed) and Detroit PYTHIA8 tunes (red solid) with mid-rapidity π^+ cross sections as a function of p_T (left) [24], UE multiplicity as a function of leading jet p_T (middle) [15], and the SoftDrop groomed jet mass (right) [27] in $p + p$ collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

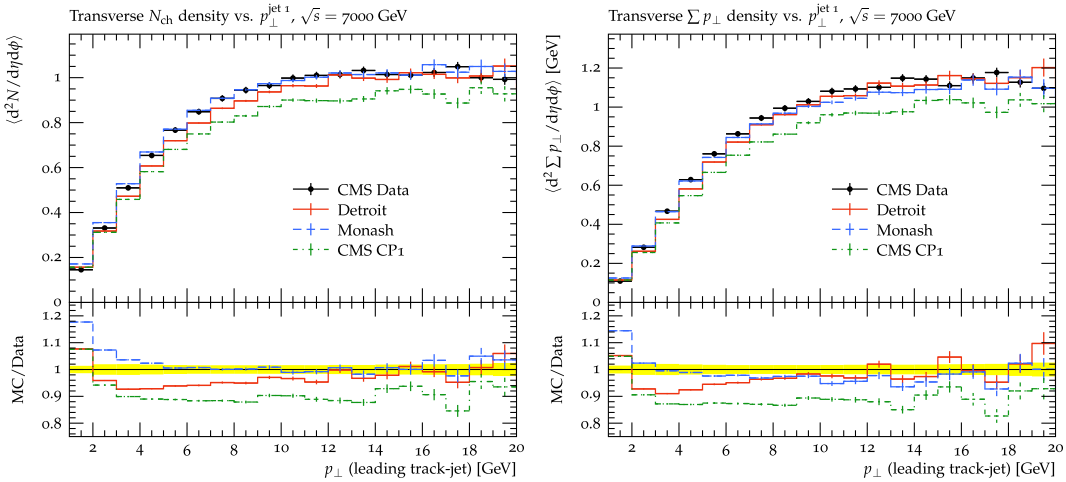

FIG. 4. Underlying event observable in the transverse region as a function of leading track-jet p_T from the CMS measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV [28]. The left shows the charge particle multiplicity and the right shows the p_T sum. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

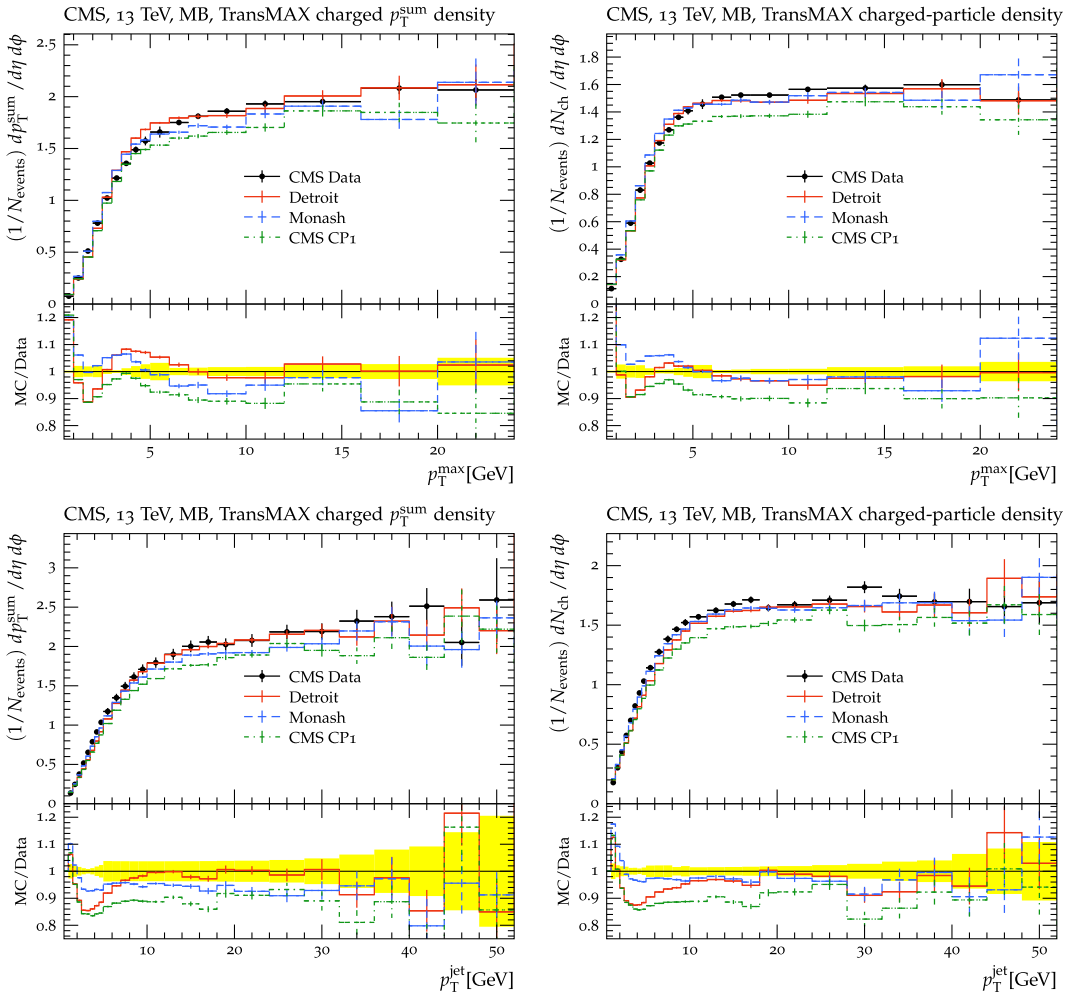

FIG. 5. The p_T sum (left) and charge particle multiplicity (right) as a function of leading track (top) and jet (bottom) p_T in the transMax region from the CMS measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV [29]. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

We also provide a set of "eigentunes" that quantify the MC errors that may be used for systematic studies. This was done using the Professor package which diagonalizes the covariance matrix at the best fit point, and provides values for the tuned parameters that correspond to deviations along the principle directions for a fixed tolerance $\Delta\chi^2$. For our analysis, we choose a heuristic $\Delta\chi^2 =$ n.d.f./2 as done in Ref. [13]. Table IV tabulates the parameter values for each eigentune.

V. COMPARISONS WITH MIDRAPIDITY DATA

Figure 3 shows the comparison of the PYTHIA8 Monash and Detroit tunes for a representative sample of measurements in data. All others are shown in Appendix. The midrapidity π^+ cross section on the left, the average UE transverse charge particle multiplicity in the middle, and the groomed jet mass distribution on the right for $\sqrt{s} =$ 200 GeV $p + p$ collisions. We find, in general, the Detroit tune provides a better match to the data than the Monash tune, in particular the underlying event charged particle multiplicity. It is especially relevant to highlight that reducing the π cross section at low momenta directly translates to a significantly better description of the UE multiplicity and the jet mass, especially in the tails of the distribution.

In $\sqrt{s} = 200$ GeV $p + p$ collisions, the π^- spectrum is well described by the Detroit tune across the entire measured p_T range. The π^+ spectrum is consistent except still being slightly over predicted above $p_T = 5$ GeV/c by about 10% considering experimental uncertainties. The charged particle multiplicities in the toward, away, and transverse regions are well described by the Detroit tune across all leading jet p_T. The average charge particle p_T versus leading jet p_T for the toward, away, and transverse regions changed only slightly in the Detroit tune with respect to the Monash tune. In the transverse region the description of the data is slightly improved and in the

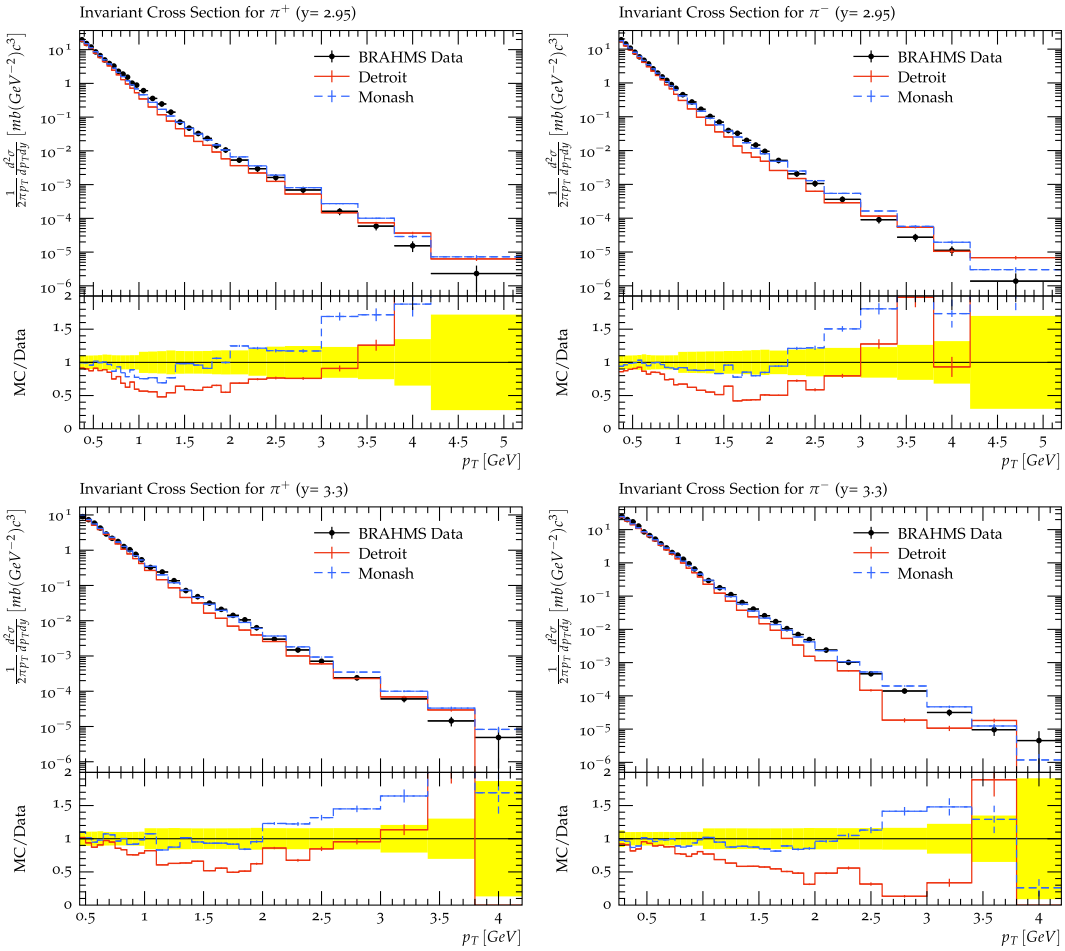

FIG. 6. $\pi^{\pm}$ cross sections at forward rapidity in $p+p$ collisions at $\sqrt{s}=200$ GeV measured by the BRAHMS experiment [30] compared to the default (blue dashed) and Detroit (red solid) PYTHIA8 tunes. The top two figures show π^+ and π^-, at $y=2.95$. The bottom two show the same except for $y=3.3$. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

toward and away regions slightly degraded by a few percent. For the average p_T for tracks with $p_T > 0.5$ GeV/c in the transverse region, we observe both the Detroit and Monash tunes undershoot the data at low leading jet p_T by around 8% and improve at higher leading jet p_T. The Drell-Yan di-muon data is well described by the Detroit tune where as the Monash tune slightly under predicted the data. In all the jet-substructure observables we see an improved agreement with the data using the Detroit tuned PYTHIA, but observe some persistent discrepancies in the tails of the distributions.

We observe that the Detroit tune is able to describe the CDF charge particle multiplicities in all $\sqrt{s}=300$, 900, and 1960 GeV center-of-mass energies, as shown in Appendix. The charged particle p_T sum is well described at $\sqrt{s}=1960$ GeV. In the $\sqrt{s}=300$ and 900 GeV charged particle p_T sum observables the Detroit tune slightly under predicts the data in the leading particle p_T range of around 4 to 8 GeV/c, but is less than 10% considering experimental uncertainties.

We compare our tune with the Monash and CMS CP1 tunes to underlying event observables measured by the CMS experiment in data from 7 [28] and 13 TeV [29] $p+p$ collisions in Figs. 4 and 5, respectively (additional 13 TeV comparisons in Appendix in Fig. 13). We note for all these comparisons, we have explicitly turned off long-lived decays in the MC generation. In the 7 TeV data, the Monash tune is able to describe the data better than the Detroit tune and the CMS CP1 tune. At higher leading track-jet p_T, the Detroit tune is consistent with the data and the CMS CP1 tune slightly under predicts the data. At 13 TeV, our tune is able to describe the data above a leading track or jet p_T of 5 and 10 GeV, respectively. The Monash tune is able to describe the charged particle density in the same regions, but under predicts the p_T sum. The CMS CP1 tune under predicts both. We note that in the low p_T

regions, the predictions from our tune vary more in shape with respect to the Monash tune, and is due to the proton shape function used.

VI. COMPARISONS AT FORWARD RAPIDITY

We compare the Detroit and Monash PYTHIA8 tunes to the measured $\pi^{\pm}$ cross sections at rapidity $y = 2.95$ and 3.3 measured by the BRAHMS experiment [30] and the inclusive π^0 cross sections in $3.4 < \eta < 4.0$ measured by the STAR experiment [31] in $\sqrt{s} = 200$ GeV $p + p$ collisions. As shown in Figs. 6 and 7 for BRAHMS and STAR data, respectively, both the Monash and Detroit tunes undershoot the data in the measured low and mid p_T/E_π ranges, and are consistent or exceed the data at high p_T/E_π. In the case of the Detroit tune, this discrepancy is persistent even considering the envelope covered by the eigentunes in Table IV. An interesting observation is that the Detroit tune does worse compared to the standard tune. We have attempted simultaneous tuning of mid- and forward-rapidity data, but are unable to recover the agreement seen for the midrapidity tune. The forward rapidity data favors larger values of $p_{T,0}^{\mathrm{Ref.}}$ compared to the Monash value to induce a better MC agreement with the data. We additionally have expanded our simultaneous tuning exercise to include the initial state radiation parameters α_s, intrinsic k_T, and (SpaceShower) $p_{T,0}^{\mathrm{Ref.}}$, and utilizing a different proton shape parametrization (MultipartonInteractions:bProfile = 3) with its associated parameter. Even with the inclusion of these additional tuning parameters and an enhanced tunable phase-space, we are still unable to reach a satisfactory agreement with both mid- and forward-rapidity data.

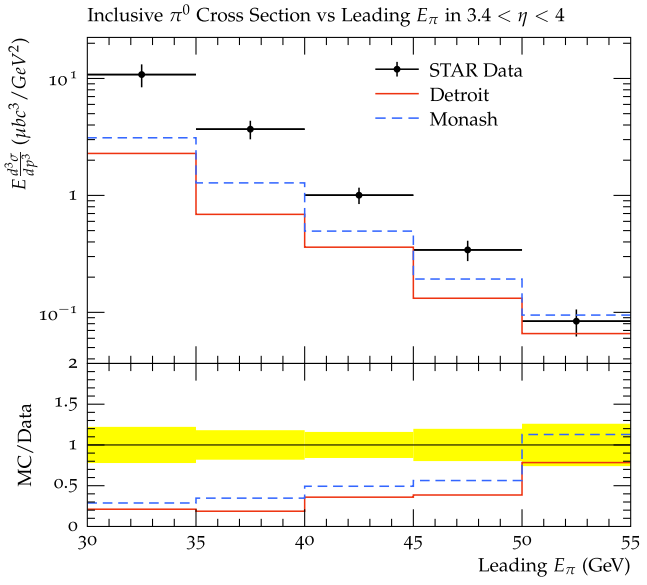

FIG. 7. Inclusive π^0 cross sections versus leading π^0 energy at forward rapidity in $p + p$ collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment [31] compared to the default (blue dashed) and Detroit (red solid) PYTHIA8 tunes. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

Having a PYTHIA tune that is able to simultaneously describe the physics at both mid- and forward-rapidity regions will be essential for the upcoming RHIC program. The STAR experiment has installed a new detector system in the forward region ($2.5 < \eta < 4$) for the 2022 Run [32]. Electromagnetic and hadronic calorimeters, as well as silicon trackers, will allow for full jet reconstruction for the first time in the far forward region. This will open up significantly enhanced opportunities for forward physics and midforward rapidity correlation studies with respect to previous data. Additionally, with the Electron-Ion Collider program on the horizon spanning a broad kinematic range and excellent detector capabilities across rapidities, event generators will need to describe physics at the level of the experimental precision in order to advance the discovery potential. A forward tuning exercise that will aim to address these points will be the focus of future studies.

VII. CONCLUSION

This publication represents the first UE tuning exercise of the PYTHIA8 generator utilizing data in proton-proton collisions at center-of-mass energies down to 200 GeV. We have provided a new UE tune, which we name the Detroit tune, and a set of eigentunes for the PYTHIA8 event generator that utilizes data from $\sqrt{s} = 200$, 300, 900, and 1960 GeV hadron-hadron collisions. We observe that the default Monash tune, that was tuned to data at LHC energies, significantly deviates from STAR, PHENIX, and CDF measurements. This is due to the energy-scaling of the PYTHIA $p_{T,0}$ regularization parameter not being appropriate. Our retuning of this parameter, and others related to the MPI, show a significant improvement in the simultaneous description of the data across all center-of-mass energies that are studied. We additionally compared our tune at LHC energies and find agreement with the UE data at higher p_T. This agreement is either comparable or better than the default Monash tune. At low p_T there are large deviations with respect to the data, and are driven by the different proton shape function used.

We advocate for using this tune for PYTHIA8 studies at both current (STAR) and future (sPHENIX) RHIC experiments for the higher statistics running period planned for 2022–2025. Also, there are possibilities for sequential retunings of PYTHIA8 as more data becomes available at RHIC energies, particularly at forward rapidities, but also to additionally improve jet-substructure observables. Therefore, our tune serves as a natural starting point for these future exercises. Our tuning study encourages the authors of the respective MC models to take into account data at a variety of collision energies toward a universal parameterizations of physics leading to improvements across varied kinematics.

ACKNOWLEDGMENTS

We would like to thank the STAR collaboration and its management for initiating such a tuning exercise and their support and comments during the tuning process has indelibly contributed to its successful completion. We also like to thank Elke-Caroline Aschenauer and Carl Gagliardi in particular for input during the tuning. Special thanks to Flemming Videbaek for access to BRAHMS data. We express our gratitude to Christian Bierlich for useful discussions throughout this work. We thank the RCF facility at Brookhaven National Laboratory and the Wayne State University high performance computing center for the computational resources used for this work. This work was supported in part by the Office of Nuclear Physics within the U.S. DOE Office of Science, the U.S. National Science Foundation, and the National Natural Science Foundation of China.

APPENDIX: COMPARISONS OF THE DETROIT PYTHIA8 TUNE WITH EXPERIMENTAL DATA

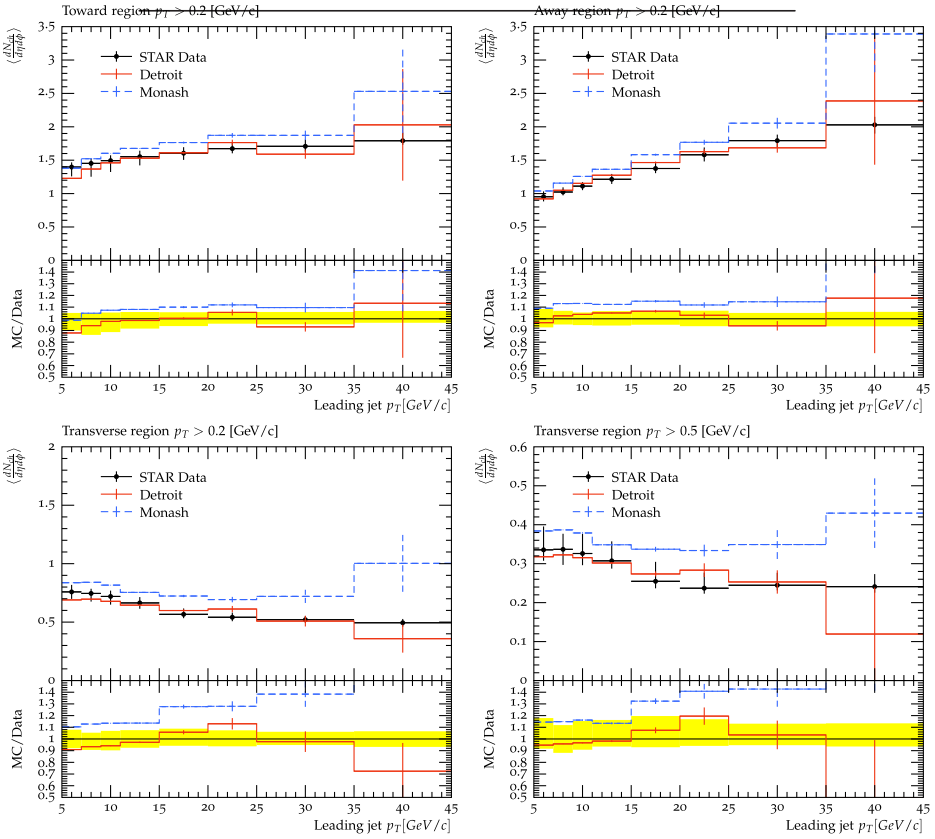

FIG. 8. Average charged particle multiplicity as a function of leading jet in $p + p$ collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment [15]. The top left and right, and bottom left show for the toward, transverse, and away regions, respectively, for charged particles with $p_T > 0.2$ GeV/c. The bottom right shows the transverse region for charged particles with $p_T > 0.5$ GeV/c. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

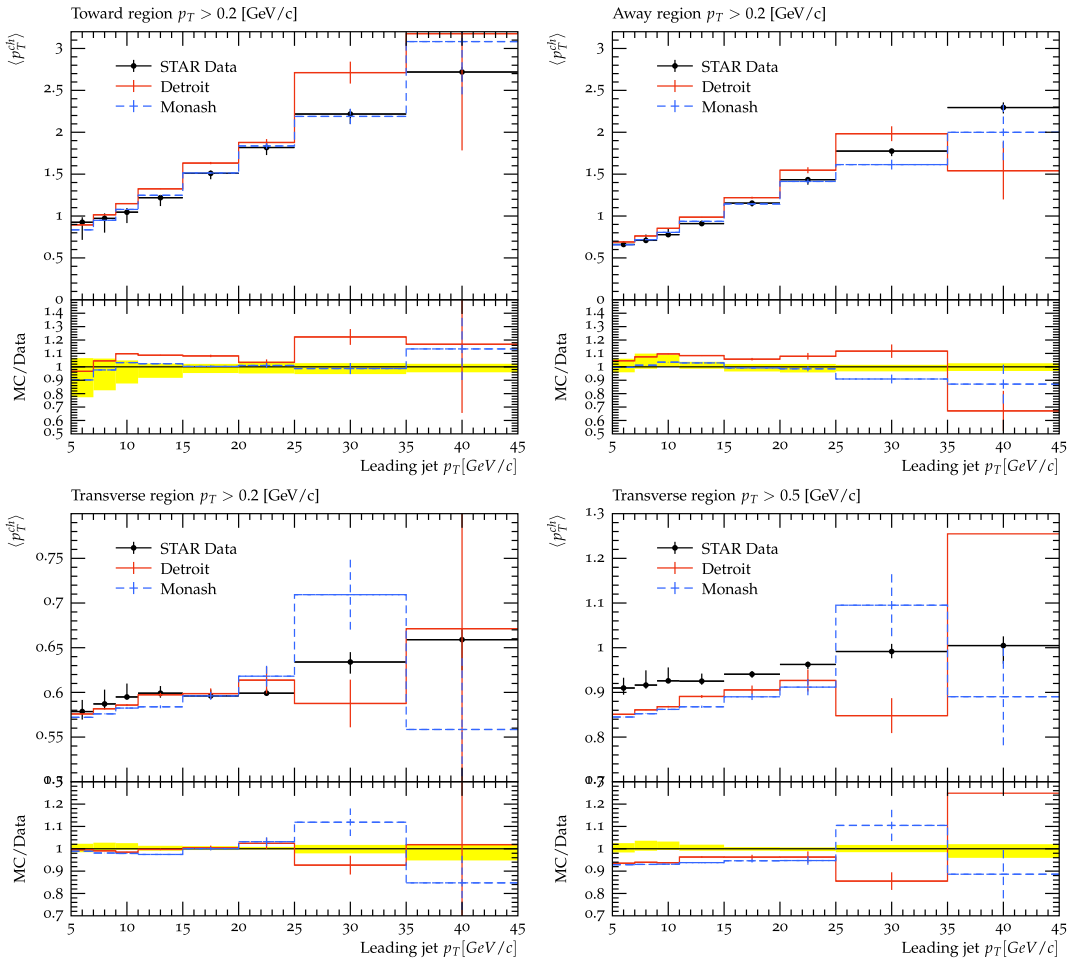

FIG. 9. Average charged particle p_T as a function of leading jet in $p+p$ collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment [15]. The top two and bottom left figures show the toward and transverse, and away regions, respectively, for charged particles with $p_T > 0.2$ GeV/c. The bottom right shows the transverse region for charged particles with $p_T > 0.5$ GeV/c. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

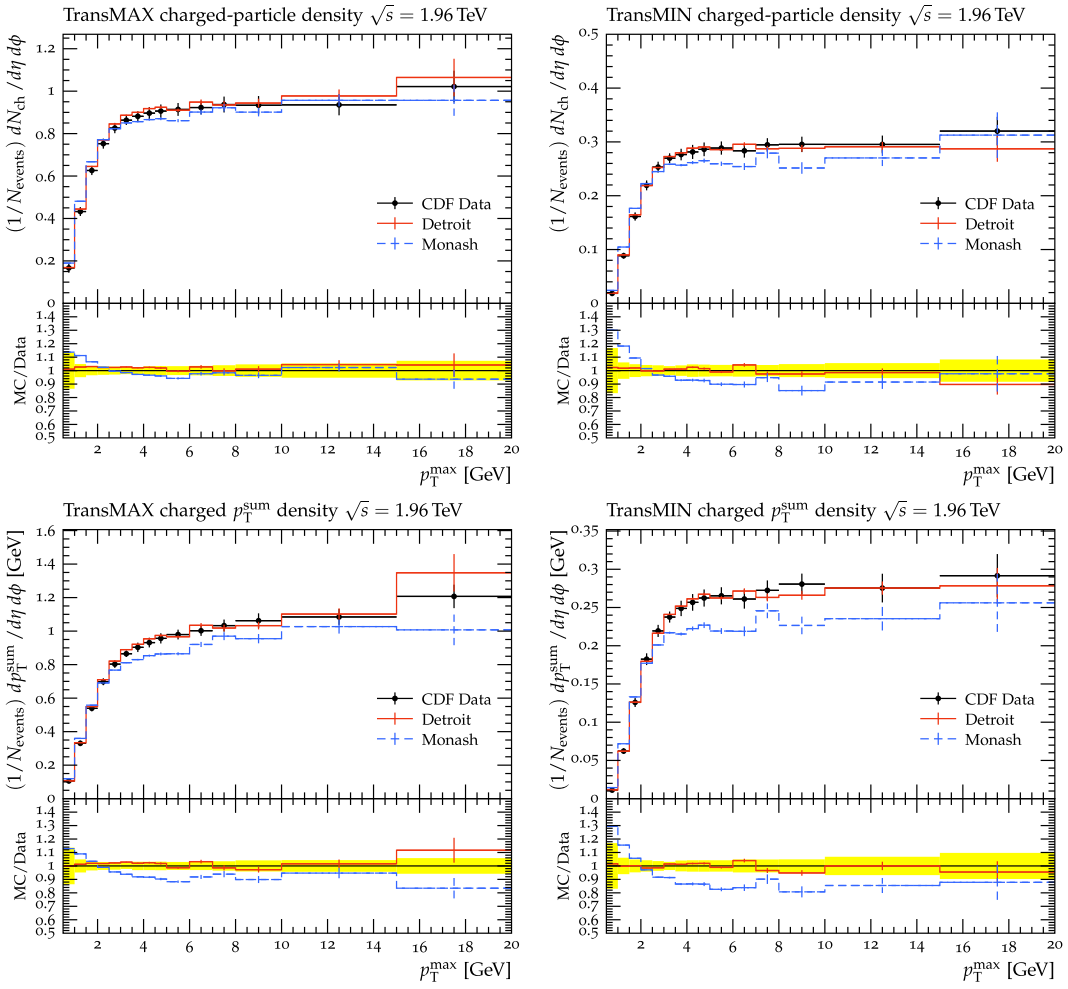

FIG. 10. Underlying event observable as a function of leading hadron p_T from the CDF measurement in proton-antiproton collisions at $\sqrt{s} = 1960$ GeV [18]. The top left and right show the charge particle multiplicity in the transMAX and transMIN regions (see text for definitions), respectively. The bottom left and right figures show the charge particle p_T sum for the transMAX and transMIN regions, respectively. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

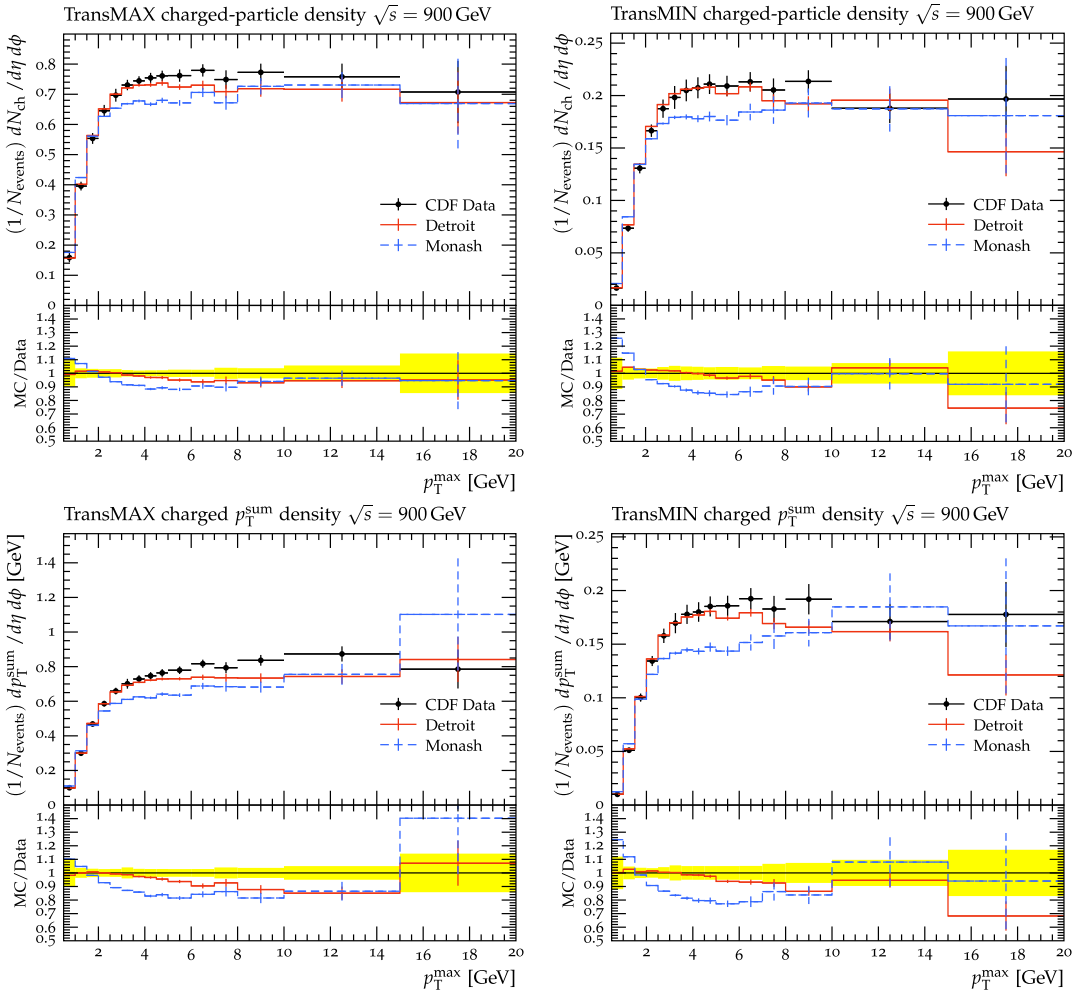

FIG. 11. Underlying event observable as a function of leading hadron p_T from the CDF measurement in proton-antiproton collisions at $\sqrt{s} = 900$ GeV [18]. The top left and right show the charge particle multiplicity in the transMAX and transMIN regions (see text for definitions), respectively. The bottom left and right figures show the charge particle p_T sum for the transMAX and transMIN regions, respectively. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

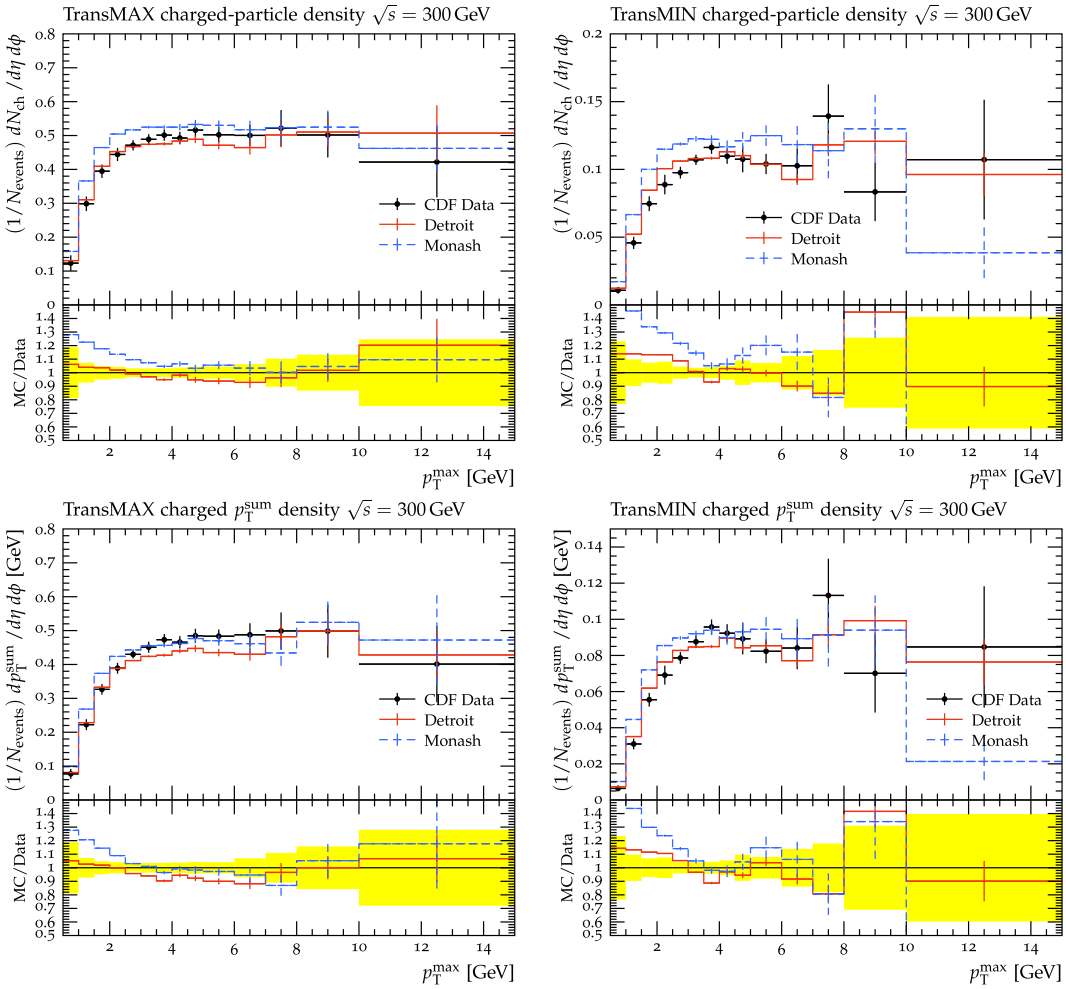

FIG. 12. Underlying event observable as a function of leading hadron p_T from the CDF measurement in proton-antiproton collisions at $\sqrt{s} = 300$ GeV [18]. The top two figures show the charge particle multiplicity in the transMAX and transMIN regions (see text for definitions). The bottom two figures show the charge particle p_T sum for the transMAX and transMIN regions. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

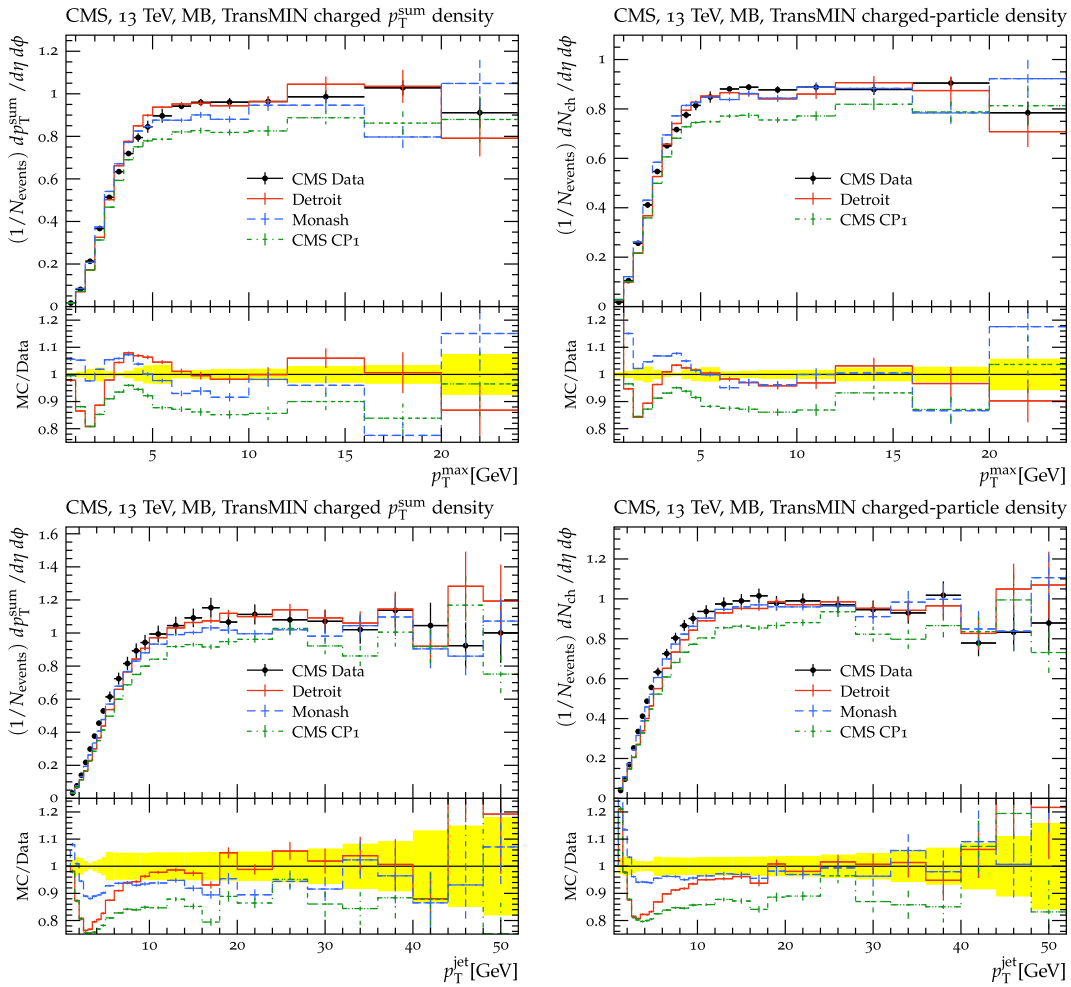

FIG. 13. The p_T sum (left) and charge particle multiplicity (right) as a function of leading track (top) and jet(bottom) p_T in the transMin region from the CMS measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV [29]. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

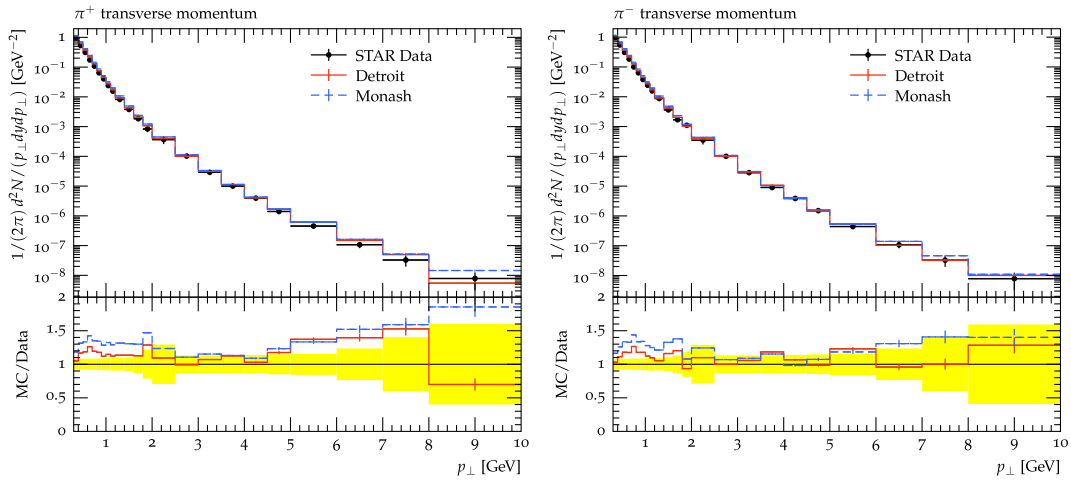

FIG. 14. Midrapidity π cross sections as a function of p_T in $p + p$ collisions at $\sqrt{s} = 200$ GeV [24]. The left figure shows π^+ and the right π^-. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

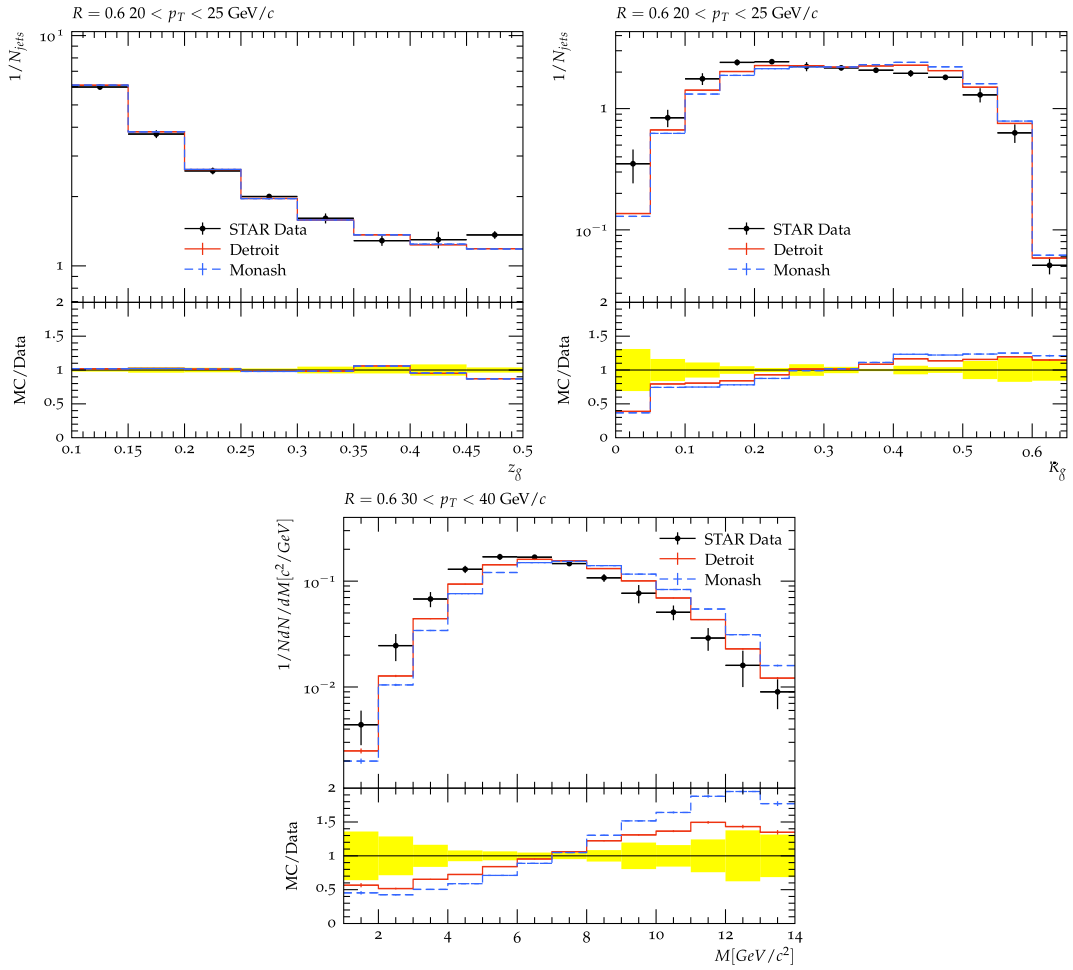

FIG. 15. Example jet substructure distributions in $p + p$ collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment. The top left and right figures show the z_g and R_g observables for anti-kT jets with $R = 0.6$ and jet p_T 20 to 25 GeV/c [26]. The bottom shows the jet mass for anti-kT jets with $R = 0.6$ and p_T 30 to 40 GeV/c [27]. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data and the yellow shaded region shows the data uncertainties.

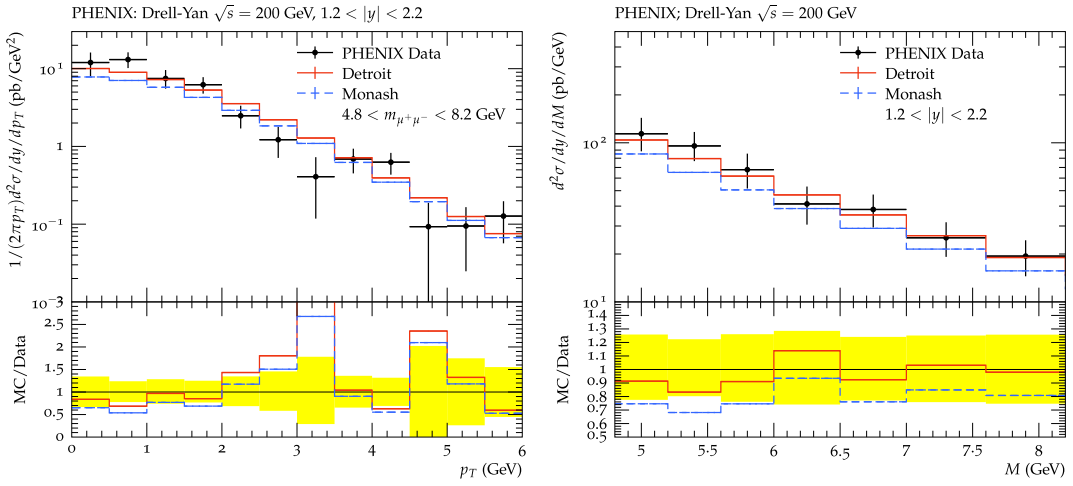

FIG. 16. Di-muon from Drell-Yan p_T (left) and mass (right) in $p + p$ collisions at $\sqrt{s} = 200$ GeV measured by the PHENIX experiment [25]. The bottom panels in each figure show the ratios of the Monte Carlo predictions with respect to the data.

[1] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA6.4 physics and manual, J. High Energy Phys. 05 (2006) 026.

[2] T. Sjostrand, S. Mrenna, and P. Z. Skands, A brief introduction to PYTHIA8.1, Comput. Phys. Commun. **178**, 852 (2008).

[3] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA8.2, Comput. Phys. Commun. **191**, 159 (2015).

[4] J. Bellm *et al.*, Herwig 7.0/Herwig++ 3.0 release note, Eur. Phys. J. C **76**, 196 (2016).

[5] M. Bahr *et al.*, Herwig++ physics and manual, Eur. Phys. J. C **58**, 639 (2008).

[6] P. Z. Skands, Tuning Monte Carlo generators: The Perugia tunes, Phys. Rev. D **82**, 074018 (2010).

[7] P. Skands, S. Carrazza, and J. Rojo, Tuning PYTHIA8.1: The Monash 2013 tune, Eur. Phys. J. C **74**, 3024 (2014).

[8] A. M. Sirunyan *et al.*, Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements, Eur. Phys. J. C **80**, 4 (2020).

[9] P. Gunnellini, H. Jung, and R. Maharucksit, Investigation of the energy dependence of the $p_{T,0}$ parameter in the PYTHIA8 Monte Carlo event generator, Eur. Phys. J. C **78**, 521 (2018).

[10] Summary of ATLAS PYTHIA8 tunes, 2012.

[11] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, Systematic event generator tuning for the LHC, Eur. Phys. J. C **65**, 331 (2010).

[12] V. Khachatryan *et al.* (CMS Collaboration), Event generator tunes obtained from underlying event and multiparton scattering measurements, Eur. Phys. J. C **76**, 155 (2016).

[13] ATLAS PYTHIA8 tunes to 7 TeV data, 2014.

[14] M. S. Abdallah *et al.* (STAR Collaboration), Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 200$ GeV, Phys. Rev. D **103**, L091103 (2021).

[15] J. Adam *et al.* (STAR Collaboration), Underlying event measurements in $p + p$ collisions at $\sqrt{s} = 200$ GeV at RHIC, Phys. Rev. D **101**, 052004 (2020).

[16] J. Adam *et al.* (STAR Collaboration), Longitudinal double-spin asymmetry for inclusive jet and dijet production in pp collisions at $\sqrt{s} = 510$ GeV, Phys. Rev. D **100**, 052005 (2019).

[17] J. Adam *et al.* (STAR Collaboration), Measurement of inclusive charged-particle jet production in Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV, Phys. Rev. C **102**, 054913 (2020).

[18] T. Aaltonen *et al.* (CDF Collaboration), Study of the energy dependence of the underlying event in proton-antiproton collisions, Phys. Rev. D **92**, 092009 (2015).

[19] C. Bierlich *et al.*, Robust independent validation of experiment and theory: Rivet version 3, SciPost Phys. **8**, 26 (2020).

[20] H. Dembinski and P. Ongmongkolkul *et al.*, scikit-hep/iminuit, 2020.

[21] F. James and M. Roos, Minuit: A system for function minimization and analysis of the parameter errors and correlations, Comput. Phys. Commun. **10**, 343 (1975).

[22] R. D. Ball *et al.*, Parton distributions with LHC data, Nucl. Phys. **B867**, 244 (2013).

[23] R. D. Ball *et al.*, Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).

[24] J. Adams *et al.*, Identified hadron spectra at large transverse momentum in p + p and d + Au collisions at $\sqrt{s_{NN}} =$ 200 GeV, Phys. Lett. B **637**, 161 (2006).

[25] C. Aidala *et al.* (PHENIX Collaboration), Measurements of $\mu\mu$ pairs from open heavy flavor and drell-yan in $p + p$ collisions at $\sqrt{s} = 200$ GeV, Phys. Rev. D **99**, 072003 (2019).

[26] J. Adam *et al.*, Measurement of groomed jet substructure observables in $p + p$ collisions at $\sqrt{s} = 200$ GeV with STAR, Phys. Lett. B **811**, 135846 (2020).

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

[28] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV, J. High Energy Phys. 09 (2011) 109.

[29] CMS Collaboration, Underlying event measurements with leading particles and jets in pp collisions at $\sqrt{s} = 13$ TeV, CMS Physics Analysis Summary, 2015, Report No. CMS-PAS-FSQ-15-007.

[30] I. Arsene *et al.* (BRAHMS Collaboration), Production of Mesons and Baryons at High Rapidity and High p_T in Proton-Proton Collisions at $\sqrt{s} = 200$ GeV, Phys. Rev. Lett. **98**, 252001 (2007).

[31] J. Adams *et al.* (STAR Collaboration), Cross Sections and Transverse Single-Spin Asymmetries in Forward Neutral-Pion Production from Proton Collisions at $\sqrt{s} = 200$ GeV, Phys. Rev. Lett. **92**, 171801 (2004).

[32] The STAR forward calorimeter system and forward tracking system, https://drupal.star.bnl.gov/STAR/system/files/Proposal.ForwardUpgrade.Nov_.2018.Review.pdf, 2018.