Available on the CERN CDS information server **CMS PAS GEN-17-002**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-generators@cern.ch 2021/09/07

A new set of CMS tunes for novel colour reconnection models in PYTHIA8 based on underlying-event data

The CMS Collaboration

Abstract

New sets of parameter tunes for two of the colour reconnection models implemented in the PYTHIA8 event generator, QCD-inspired and gluon move, are obtained based on the default CMS PYTHIA8 underlying-event tune, CP5. Measurements sensitive to underlying-event performed at $\sqrt{s}$ =1.96, 7, and 13 TeV are used to constrain the parameters of the colour reconnection models and the multiple parton interactions simultaneously. The new colour reconnection tunes are compared to various measurements at a centre-of-mass energy of 1.96, 7, 8, and 13 TeV including measurements of underlying event, strange particle multiplicities, jet substructure observables, hadron ratios, jet shapes and colour flow in top quark pair events, and top quark mass. The new colour reconnection tunes can be used to estimate systematic uncertainties related to colour reconnection modelling.

1 Introduction

The Monte Carlo (MC) event generators, such as PYTHIA8 [1], are indispensable tools when performing measurements at hadron colliders. In order to provide an accurate description of the high energy collisions the "hard" scattering and the so-called underlying event (UE) are computed for each event. In a hard scattering, two initial partons give rise to an interaction at a large exchanged transverse momentum (p_{T}). The underlying event represents any additional activity occurring at lower scales accompanying the hard scattering. The UE consists of several components, such as initial- and final-state radiation (ISR and FSR, respectively), multiple parton interactions (MPI), and beam-beam remnants (BBR). All the coloured partons produced by these ingredients are finally transformed into colourless hadrons, through the hadronisation process. Particularly relevant for the characterisation of the UE are the MPI, which consist of numerous additional 2-to-2 parton-parton interactions, occurring within the single collision event. Due to the increase of the partonic content at small longitudinal momentum fractions (x), the MPI contribution increases with increasing collision energy. PYTHIA8 regularises the contribution of the primary hard-scattering processes and MPI with respect to the differential cross section through a parameter p_{T0} that depends on the centre-of-mass energy, $\sqrt{s}$. The energy dependence of the p_{T0} parameter in PYTHIA8 is described with a power law function of the form:

$$p_{\mathrm{T0}}(\sqrt{s}) = p_{\mathrm{T0}}^{\mathrm{ref}} \left(\frac{\sqrt{s}}{\sqrt{s_0}} \right)^{\epsilon}, \tag{1}$$

where $p_{\mathrm{T0}}^{\mathrm{ref}}$ is the value of p_{T0} at a reference energy $\sqrt{s_0}$, and ϵ is a tunable parameter which determines the energy dependence. At a given centre-of-mass energy the mean number of additional interactions from MPI depends on p_{T0}, the parton distribution function (PDF), and the overlap of the matter distributions of the two colliding hadrons. In order to track colour information during the development of the parton shower, partons are represented by colour lines. Because each MPI adds coloured partons to the final state, it creates a dense net of colour lines that overlap with the coloured parton fields of the hard scattering and with each other. All the generated colour lines may be connected with each other according to a colour reconnection (CR) model. The CR mechanism allows colour strings originating from different interactions to be connected and exchange colour information. CR was first included in minimum bias simulations in order to reproduce the increase of average transverse momentum of the charged particles ($\langle p_{\mathrm{T}} \rangle$) as a function of the multiplicity of the measured charged particles (n_{ch}), see e.g. [2, 3] and also to describe the $dn_{ch}/d\eta$ distribution [3]. Here, pseudorapidity (η) is defined as $\eta = -\ln[\tan(\theta/2)]$, where the polar angle θ is defined with respect to the anticlockwise-beam direction. By introducing correlations between the various partonic systems, one is able to reduce the number of particles in the final state, which exhibit a larger $\langle p_{\mathrm{T}} \rangle$ than in a scenario without CR. Different phenomenological models for CR have been developed and are included in simulations. In these models, the general idea is to construct a colour potential dependent on the string lengths of the colour field to be minimised. Then, the new partonic configuration is fed to the subsequent hadronisation process. None of the MPI processes or the CR models are completely determined by first principles and they all include free parameters. A specified set of such parameters that has been adjusted to better fit some aspects of the data is referred to as a "tune". The MPI model, which is implemented in PYTHIA8, is documented in Ref. [2]. The recent versions of PYTHIA8 and Herwig++/7 [4, 5] implement different CR models that can be tested in events from proton-proton (pp) collisions. The models implemented in PYTHIA8, referred to as the "MPI-based", "QCD-inspired", and "gluon-move" CR models, are briefly described in the following:

- **MPI-based model:** is the simplest model [6] implemented in MC event generators

and introduces only one tunable parameter. In this model, the partons are classified by the MPI system they belong to. Each MPI is originally a $2 \rightarrow 2$ scattering. A reconnection probability for an MPI with a hardness scale p_{T} of the $2 \rightarrow 2$ interaction is defined as:

$$P = \frac{p^2_{\mathrm{T_{Rec}}}}{(p^2_{\mathrm{T_{Rec}}} + p^2_{\mathrm{T}})} \tag{2}$$

with $p_{\mathrm{T_{Rec}}} = R \cdot p_{\mathrm{T0}}$, where R is a tunable parameter and p_{T0} is the energy-dependent dampening parameter used for MPIs. This parameter regularises the partonic cross section to avoid divergence at low p_{T}. According to this formula, MPI systems at high p_{T} would tend to escape out of the interaction point, without being colour reconnected to the hard scattering system. Colour fields originating from a low-p_{T} MPI system would, instead, more likely to exchange colour. Once the systems to be connected are determined, partons of low-p_{T} systems are added to the strings defined by the highest-p_{T} system to achieve minimal total string length.

- **QCD-inspired model:** the QCD-inspired model [7] implemented in PYTHIA8 adds the QCD colour rules on top of the minimisation of the string length. The model constructs all pair of dipoles that are allowed to be reconnected by QCD colour rules that determines the colour-compatibility of two strings. This is done iteratively until no further allowed reconnection shortens the total string length. It uses a simple space-time picture to connect causally the produced strings through a string-length measure (λ) to decide whether a possible reconnection is actually favoured. The default parametrisation for λ is

$$\lambda = \ln(1 + \sqrt{2}\frac{E_1}{m_0}) + \ln(1 + \sqrt{2}\frac{E_2}{m_0}) \tag{3}$$

  where E_1 and E_2 represent the energies of the coloured partons calculated in the rest-frame of the dipole, and m_0 variable is a constant with dimensions of energy [7]. Additionally, it allows the creation of junction structures that give the possibility to include the effects of higher order effects in CR. Junctions are not simply dipoles but they can be directly produced by three or four dipoles. This new CR model takes the leading-colour strings (in which each MPI is considered independent of each other) and transforms them into a different colour configuration based on the rules described above. Since the model relies purely on the outgoing partons, it is applicable to any type of collision. The main difference compared to the other CR models in PYTHIA8 is the introduction of reconnections that form junction structures. From a pure colour consideration the probability to form a junction topology is three times larger than an ordinary reconnection. However, the junction will introduce additional strings, and is therefore often disfavoured due to a larger λ measure. Given the close connection between junctions and baryons, the new model predicts a baryon enhancement. It was shown that it is able to simultaneously describe the Λ production for both LEP and LHC experiments [7], which neither of the earlier PYTHIA8 tunes have been able to do.

- **Gluon-move model**: in this scheme [8], all final-state gluons are identified, alternatively only a fraction of them, along with all the colour connected pairs of partons. Then, an iterative process starts. For each final-state gluon attached to a string piece of partons, the change in the string length, resulting from moving the gluon to a string piece belonging to different partons, is calculated. The gluon is moved to the

string for which the move gives the smallest change. In this scheme, quarks would not be reconnected, i.e. they would remain in the same position without any colour exchange. To improve this picture, the flip mechanism could be included, which also gives the quarks the possibility to be mixed up. Junctions are allowed to take part in the flip step as well, but no considerable differences are expected because of the limitation of the junction formation in this model. In this work, the flip mechanism is not utilised due to its unperceived effect in single diffractive events. The main free parameters of the gluon-move model account for the lower limit of the string length allowed for colour connection, the fraction of gluons allowed to move or flip, and the lower limit of the allowed reduction of the string lengths.

Usually, MPI and CR effects are investigated and constrained through fits to measurements of the number and energy density of charged particles produced in pp collisions. UE measurements have been performed at various collision energies by different collaborations [9–13]. The measurements are typically performed by studying the multiplicity and the scalar sum of the p_{T} of the charged particles (p_{T}^{sum}), measured as a function of the p_{T} of the leading charged particle in the event. Different regions of the plane transverse to the direction of the beams are generally considered, as defined by the direction of the leading charged particle. A "toward" region mainly includes the products of the hard scattering, two "transverse" regions contain the products of MPI and is affected by contributions from ISR and FSR, and an "away" region comprises the recoiling objects belonging to the hard scattering. More recent measurements [9, 10] subdivide the transverse region into a "transMIN" and a "transMAX" one, which are defined to be the regions with the minimum and maximum number of particles between the two transverse regions. This is done in order to try to disentangle in a better way contributions from MPI, ISR, and FSR. For events with large initial or final-state radiation the transMAX region contains the "transverse-side" jet, while both the transMAX and transMIN regions receive contributions from the MPI and beam-beam remnants. Thus, the transMIN region is sensitive to the MPI and BBR, while the transMAX minus the transMIN is very sensitive to initial and final-state radiation (ISR and FSR). A sketch of the different regions is shown in Figure 1.

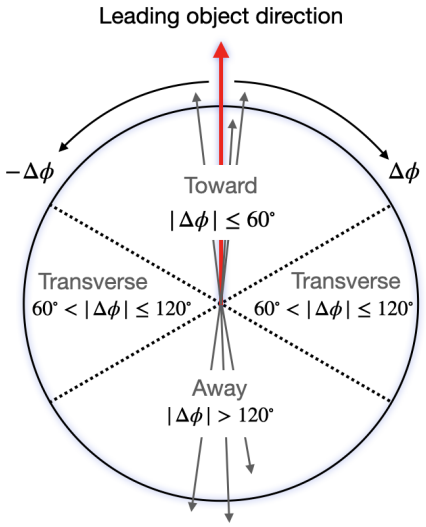

Figure 1: Display of the topology of a hadron-hadron collision in which a "hard" parton-parton collision has occurred. The "toward" region contains the "toward-side" jet, while the "away" region, on the average, contains the "away-side" jet.

It has been shown that a consistent description of the charged particle multiplicity and p_{T} sum

without taking into account the CR effects is not possible using the PYTHIA8 hadronisation model [14]. In particular, the outcome of a tune without CR is a good description of either the amount of charged particles or their p_{T} sum, but not a good simultaneous description of both. In general, the largest tension of predictions from tunes with respect to the data is observed in the soft region ($p_{\mathrm{T}} \sim 2 - 5\,\mathrm{GeV}$), where CR effects are expected to be more relevant. CR effects are also important for other processes occurring at larger scales in pp collisions. In Ref. [15], it is shown that to describe the UE variables at low average p_{T} values in top quark and antiquark pair ($\mathrm{t\bar{t}}$) events, effects of CR need to be included. The effects of CR may become more prominent in precision measurements such as the top quark mass. The uncertainty on the top quark mass because of CR is conventionally estimated by taking the difference in the predictions for a given model with and without CR. As shown in Ref. [8], this procedure might represent an underestimation of the assessed uncertainty. A better way for approaching the uncertainty estimation would be to consider a variety of CR models and variations of their parameters (e.g. see Ref. [16]). These models describe the underlying soft physics of pp collisions, and allow the investigation of their effects on the considered observable of interest, e.g. the top quark mass. The new CR models, QCD-inspired and gluon-move, were implemented in PYTHIA8 after tuning the model parameters to the existing data at 7 TeV and at lower centre-of-mass energies [7, 8]. The model predictions, with their default parameter settings in PYTHIA8.226 and CP5, are given in Figure 2 for charged particle and $p_{\mathrm{T}}^{\mathrm{sum}}$ densities measured by CMS experiment at $\sqrt{s} = 13\,\mathrm{TeV}$ [9] in transMIN and transMAX regions, and in Figure 3 for the charged-particle pseudorapidity (η) distribution, $dN_{\mathrm{ch}}/d\eta$, measured by CMS at $\sqrt{s} = 13\,\mathrm{TeV}$ [17]. The results show that the models have to be re-tuned in order to describe the underlying soft physics of pp collisions at 13 TeV.

This note presents results from two tunes making use of the QCD-inspired and the gluon-move CR models. The new CR tunes presented in this note are based on the default CMS tune CP5 [18]. In combination with the CP5 tune based in the MPI-based CR model, the new CR tunes constitute a valid input for comparing their performance on several observables, deepening the understanding of the CR mechanism, and for evaluating uncertainties due to CR effects.

The note is organised as follows. In Section 2, the tuning strategy is explained in detail and the results are presented. Section 3 shows a selection of validation plots related to observables measured at $\sqrt{s} = 1.96$, 7, 8, and 13 TeV, compared to the predictions of the new tunes. In Section 4, a study of the uncertainty on the top quark mass measurement due to the CR models is presented, before summarising the results in Section 5.

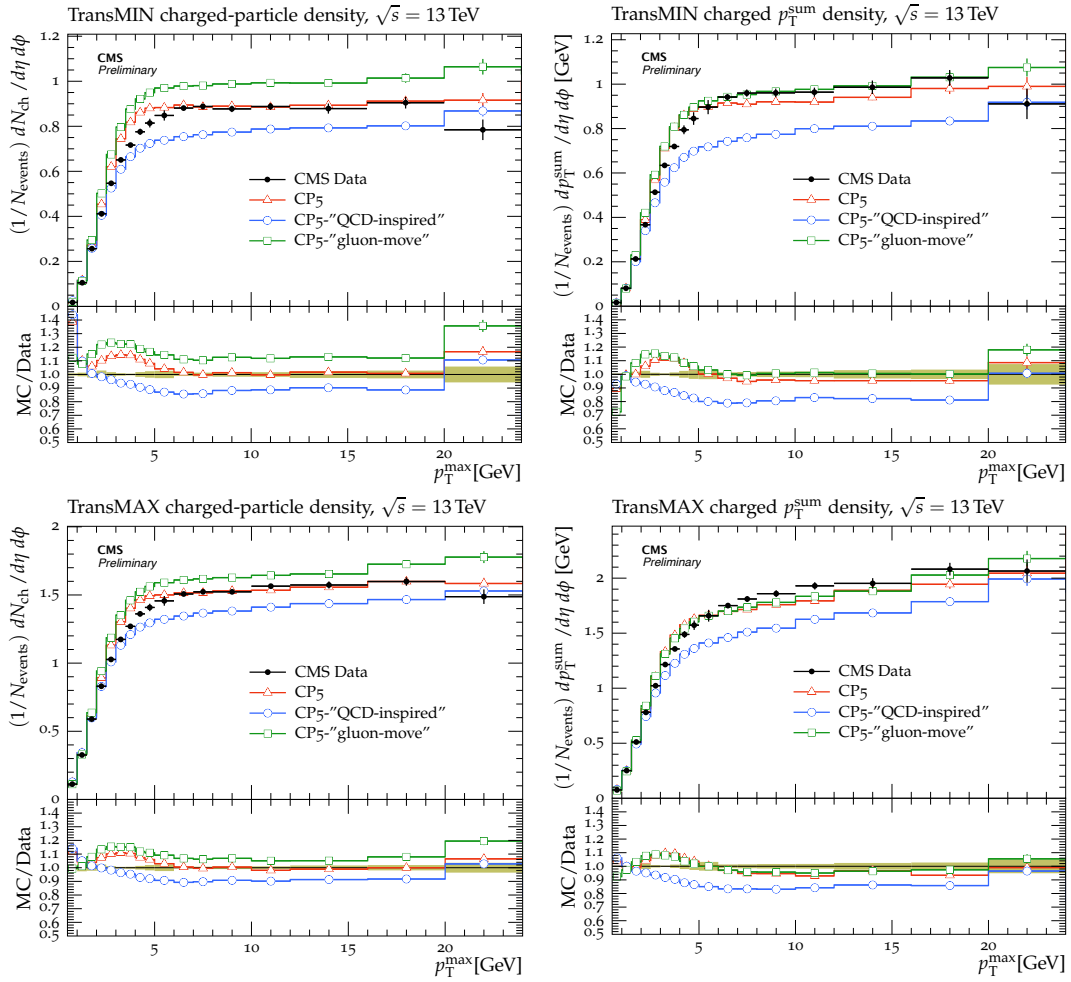

Figure 2: The charged particle (left) and p_T^{sum} densities (right) in the transMIN (upper) and transMAX (lower) regions, as a function of the p_T of the leading charged particle, p_T^{max}, measured by the CMS experiment at $\sqrt{s} = 13$ TeV [9]. The predictions of the tunes CP5, CP5-"QCD-inspired", and CP5-"gluon-move" are compared to data. The coloured band represents the total experimental uncertainty in the data.

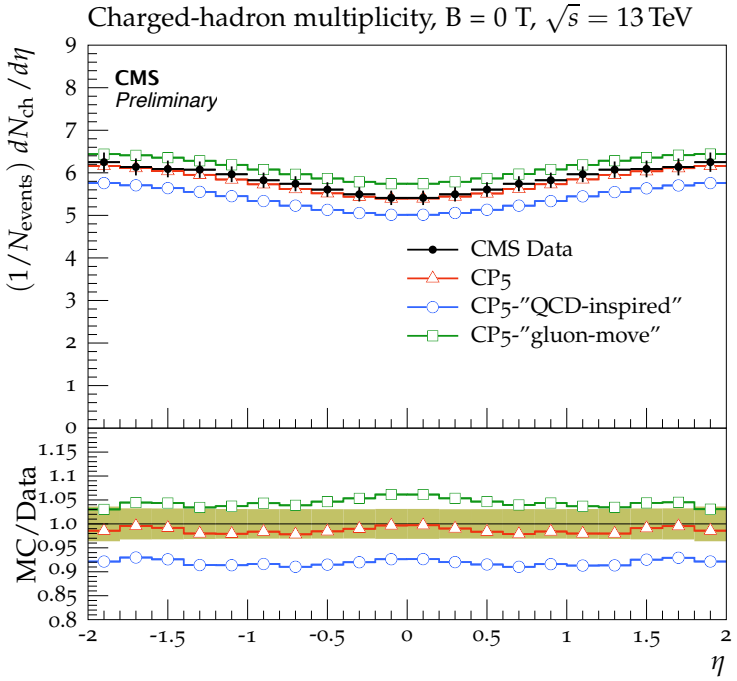

Figure 3: The pseudorapidity of charged hadrons, $dN_{\text{ch}}/d\eta$, measured in $|\eta| < 2$ by the CMS experiment at $\sqrt{s} = 13$ TeV [17]. The predictions of the tunes CP5, CP5-"QCD-inspired", and CP5-"gluon-move" are compared to data. The coloured band represents the total experimental uncertainty in the data.

2 Tuning strategy

A new set of event tunes based on CMS and CDF UE data are derived using the QCD-inspired and the gluon-move CR models, implemented in the PYTHIA 8.226 event generator. In the following, they are referred to as "CP5-CR1", and "CP5-CR2", where CR1 stands for the QCD-inspired model and CR2 stands for the gluon-move CR model. The RIVET routines used as inputs to the tune fits along with the RIVET distribution names, x-axis ranges, relative importance (R) of the input distributions, number of bins, and centre-of-mass energy values are displayed in Table 1 for CP5-CR1 and CP5-CR2. The CP5 tune is used as a baseline for the CR tuning since it has been used as the default PYTHIA8 tune for most of the new CMS analyses using data at 13 TeV starting from 2017, and it has explicitly been tested against a rich number of different final states (UE, QCD, top quark, and vector boson + jets) and observables [18]. Tune CP5 uses the NNPDF31_nnlo_as_0118 [19] PDF set, the strong coupling parameter (α_S) 0.118 for ISR, FSR, and MPI, and the MPI-based CR model. It also uses a double Gaussian functional form with two tunable parameters, MultipartonInteractions:coreFraction and MultipartonInteractions:coreRadius, to model the overlap distribution between the two colliding protons. The tune parameters are documented in Ref. [18] and displayed in Table 3. Having tunes with different CR models, one has a consistent way of evaluating systematic uncertainties due to colour reconnection effects in specific measurements.

The new tunes are obtained by constraining the parameters controlling the contributions of the MPI and of each of the CR models, simultaneously. The strategy followed to obtain the CP5-CR1 and CP5-CR2 tunes is similar to that used for the CP5 tune, i.e. the same observables, which are sensitive to soft and semi-hard MPI, are considered to constrain the parameters. These are the charged particle multiplicity and average p_{T} sum as a function of the

Table 1: List of input RIVET routines, distributions, x-axis ranges, R of the distributions in the fit, number of bins and the centre-of-mass energy used in the fits to derive the CP5-CR1 and CP5-CR2 tunes.

			CP5-CR1			CP5-CR2		
RIVET routine	$\sqrt{s}$ (TeV)	Distribution	Fit range (GeV)	N_{bins}	R	Fit range (GeV)	N_{bins}	R
CMS_2015_I1384119	13	N_{ch} vs η		20	1		20	1
CMS_2015_PAS_FSQ_15_007	13	TransMIN charged p_T^{sum}	2–28	15	1	3–36	15	0.5
		TransMAX charged p_T^{sum}	2–28	15	1	3–36	15	0.5
		TransMIN N_{ch}	2–28	15	1	3–36	15	0.1
		TransMAX N_{ch}	2–28	15	1	3–36	15	0.1
CMS_2012_PAS_FSQ_12_020	7	TransMAX N_{ch}	3–20	10	1	3–20	10	0.1
		TransMIN N_{ch}	3–20	10	1	3–20	10	0.1
		TransMAX charged p_T^{sum}	3–20	10	1	3–20	10	0.1
		TransMIN charged p_T^{sum}	3–20	10	1	3–20	10	0.1
CDF_2015_I1388868	2	TransMIN N_{ch}	2–15	11	1	2–15	11	0.1
		TransMAX N_{ch}	2–15	11	1	2–15	11	0.1
		TransMIN charged p_T^{sum}	2–15	11	1	2–15	11	0.1
		TransMAX charged p_T^{sum}	2–15	11	1	2–15	11	0.1

leading charged-particle transverse momentum, p_T^{max}, measured in transMIN and transMAX regions by the CMS experiment at 13 TeV [9] and at 7 TeV [11], and by the CDF experiment at 1.96 TeV [13]. The charged particle multiplicity as a function of η, measured by CMS at 13 TeV [20] is also used in the fit. Likewise for CP5, the region between $0.5 < p_T^{max} < 3$ GeV is excluded in the fit, since it is affected by diffractive processes whose free parameters are not considered in the tuning procedure.

The MPI-related parameters that are kept free in both CP5-CR1 and CP5-CR2 tunes are:

- `MultipartonInteractions:pT0Ref`, the parameter included in the regularisation of the partonic QCD cross section. It sets the lower scale of the MPI contribution;
- `MultipartonInteractions:ecmPow`, the exponent of the $\sqrt{s}$ dependence;
- `MultipartonInteractions:coreRadius`, the width of the core when a double Gaussian matter profile is assumed for the overlap distribution between the two colliding protons. A double Gaussian form identifies an inner, dense part, which is called core, and an outer less dense part;
- `MultipartonInteractions:coreFraction`, the fraction of quark and gluon content enclosed in the core when a double Gaussian matter profile is assumed.

The tunable CR parameters in the QCD-inspired model that are considered in the fit are:

- `ColourReconnection:m0`, the variable that determines whether a possible reconnection is actually favoured in the λ measure in Eq. 3;
- `ColourReconnection:junctionCorrection`, the multiplicative correction for the junction formation, applied to the `m0` parameter;
- `ColourReconnection:timeDilationPar`, the parameter controlling the time dilation that forbids colour reconnection between strings that are not in causal contact.

For the CP5-CR1 tune, the parameters `StringZ:aLund`, `StringZ:bLund`, `StringFlav:probQQtoQ`, and `StringFlav:probStoUD`, relative to the hadronisation, proposed in [7], are also used in the initial settings. The first two of these parameters control the shape of the longitudinal fragmentation function used in the Lund string model in PYTHIA8,

while the latter two are the probability of diquark over quark fragmentation and the probability of strangeness, respectively.

For the gluon-move scheme, the following parameters are considered for optimisation:

- `ColourReconnection:m2lambda`, an approximate hadronic mass-square scale and the parameter used in the calculation of λ;
- `ColourReconnection:fracGluon`, the probability that a given gluon will be considered for being moved. It thus gives the average fraction of gluons being considered.

The hadronisation parameters are kept the same as in the CP5 tune which uses default Monash parameters [21]. The parameters and their ranges considered in the fits are shown in Table 2.

Table 2: MPI and CR parameter ranges used in the tuning procedure.

PYTHIA8 parameter	Min–Max
MPI parameters	
`MultipartonInteractions:pT0Ref`	1.0 – 3.0
`MultipartonInteractions:ecmPow`	0.0 – 0.3
`MultipartonInteractions:coreRadius`	0.2 – 0.8
`MultipartonInteractions:coreFraction`	0.2 – 0.8
QCD-inspired	
`ColourReconnection:m0`	0.1 – 4.0
`ColourReconnection:junctionCorrection`	0.01– 10
`ColourReconnection:timeDilationPar`	0 – 60
Gluon-move model	
`ColourReconnection:m2lambda`	0.2 – 8.0
`ColourReconnection:fracGluon`	0.8 – 1.0

The fits are performed using both the PROFESSOR 1.4.0 software [22], that takes random values for each parameter in the defined multidimensional parameter space, and RIVET 2.4.0 [23] that produces the individual generator predictions for the considered observables. About 200 different choices of parameters are considered for building the random grid in the parameter space. For each choice of parameters, proton-proton inelastic events, including contributions of single-diffractive dissociation (SD), double diffraction dissociation (DD), central diffraction (CD), and non-diffractive processes (ND), are generated. It has been checked that the bin-by-bin envelopes of the different MC predictions encompass the data points. After building the grid in the parameter space, PROFESSOR performs an interpolation of the bin values for the considered observables in the parameter space, according to a third-order polynomial function. It has been checked that the degree of the polynomial used for the interpolation does not influence the tune results. The obtained function $f^{\mathrm{b}}(\mathbf{p})$ models the MC response of each bin b of the observable $\in O$ as a function of the vector of the parameters $\mathbf{p}$. The final step is the minimisation of the χ^{*2} function given by this formula:

$$\chi^{*2}(\mathbf{p}) = \sum_{\mathrm{b}\in O} \frac{(f^{\mathrm{b}}(\mathbf{p}) - R_{\mathrm{b}})^2}{\Delta_{\mathrm{b}}^2} \tag{4}$$

where R_{b} is the data value for each bin b, Δ_{b}^2 expresses the total bin uncertainty of the data. Note that χ^{*2} is not a true χ^2 function because of the reasons explained below. Treating all distributions that are used as input to the fit for the CP5-CR2 tune results in a tune that describes the data poorly: in particular, it underestimates the $dN_{ch}/d\eta$ data at 13 TeV by about 30%. This

is because the χ^2 definition treats all bins equally and the importance of $dN_{ch}/d\eta$ may be lost because of its relative precision with respect to other observables. The $dN_{ch}/d\eta$ distribution is one of the key observables that is sensitive to the number of processes and therefore increasing the importance of this observable in the fit is reasonable. In PROFESSOR, this is done by using weights with a non-standard χ^2 definition. To keep the standard properties of a χ^2 fit, instead we increase the total uncertainties of the other distributions. For the CP5-CR2 tune, in order to increase the relative importance of $dN_{ch}/d\eta$ at 13 TeV, the total uncertainty in each bin is scaled up by $1/\sqrt{R}$ with R values displayed in Table 1. Therefore, the total uncertainty of each bin of of p_{T} sum in the transMIN and transMAX regions at 13 TeV is scaled up by $\sqrt{2}$ and that of all other distributions by $\sqrt{10}$. These scale factors ensure the distributions are well described after the tuning. No scaling is needed for the input distributions for the fit for the CP5-CR1 tune, which means that all distributions are considered with the same importance. The experimental uncertainties used in the fit, in general, have bin-to-bin correlations. However, some of the bins of the UE distributions that go in the fit, e.g. $p_{\mathrm{T}}^{\mathrm{max}} > 10$ GeV, are dominated by statistical uncertainties and are uncorrelated between bins. In the minimisation procedure, because the correlations between bins are not available for the input measurements, the experimental uncertainties are assumed to be uncorrelated between data points. The minimisation procedure yields the values of the parameters that best fit the considered data. The parameters obtained from the CP5-CR1, and CP5-CR2 fits, as well as the value of the goodness of the fit are shown in the Table 3.

Table 3: The parameters obtained in the fits of the CP5-CR1 and CP5-CR2 tunes, compared with that of the tune CP5. The upper part of the table displays the fixed input parameters of the tune, while the lower part shows the fitted tune parameters. The number of degrees of freedom (N_{dof}) and the goodness of fit divided by N_{dof} are also shown.

PYTHIA8 Parameter	CP5 [18]	CP5-CR1	CP5-CR2[1]
PDF set	NNPDF3.1 NNLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
$\alpha_{\mathrm{S}}(m_{\mathrm{Z}})$	0.118	0.118	0.118
`SpaceShower:rapidityOrder`	on	on	on
`MultipartonInteractions:EcmRef` [GeV]	7000	7000	7000
$\alpha_{\mathrm{S}}^{\mathrm{ISR}}(m_{\mathrm{Z}})$ value/order	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_{\mathrm{S}}^{\mathrm{FSR}}(m_{\mathrm{Z}})$ value/order	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_{\mathrm{S}}^{\mathrm{MPI}}(m_{\mathrm{Z}})$ value/order	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_{\mathrm{S}}^{\mathrm{ME}}(m_{\mathrm{Z}})$ value/order	0.118/NLO	0.118/NLO	0.118/NLO
`StringZ:aLund`	–	0.38	–
`StringZ:bLund`	–	0.64	–
`StringFlav:probQQtoQ`	–	0.078	–
`StringFlav:probStoUD`	–	0.2	–
`SigmaTotal:zeroAXB`	off	off	off
`BeamRemnants:remnantMode`	–	1	–
`ColourReconnection:mode`	–	1	2
`MultipartonInteractions:pT0Ref` [GeV]	1.410	1.375	1.454
`MultipartonInteractions:ecmPow`	0.033	0.033	0.054
`MultipartonInteractions:coreRadius`	0.763	0.605	0.649
`MultipartonInteractions:coreFraction`	0.630	0.445	0.489
`ColourReconnection:range`	5.176	–	–
`ColourReconnection:junctionCorrection`	–	0.238	–
`ColourReconnection:timeDilationPar`	–	8.580	–
`ColourReconnection:m0`	–	1.721	–
`ColourReconnection:m2lambda`	–	–	4.917
`ColourReconnection:fracGluon`	–	–	0.993
N_{dof}	183[2]	157	158
$\chi^{*2}/N_{\mathrm{dof}}$	1.04	2.37	0.89

Although they are not used in precision measurements and are beyond the scope of this article, we also derive CR tunes based on the CP1 and CP2 settings in order to study the effect of using an LO PDF set with alternative CR models. CP1 and CP2 are the two tunes in the CPX (X = 1-5) tune family [18] that use an LO PDF set [19]. We found that the predictions of CR tunes based on CP1 and CP2 in the MB and UE observables are similar to the predictions of CR tunes based on CP5. However, CP1-CR1 (i.e. CP1 with QCD-inspired colour reconnection model) has a different trend in particle multiplicity distributions compared to predictions of other tunes discussed in this study. This different trend of CP1-CR1 cannot be attributed to the use of LO PDF set, because both CP1 and CP2 use the same LO PDF set and we do not see a different trend with CP2-CR1. This different trend observed with CP1-CR1 in particle multiplicity distributions could be an input for further tuning and development of the QCD-inspired model. Therefore, in the appendix of this article, we give the tune settings of CR tunes based on CP1 and CP2 along with their predictions in the particle multiplicity distributions.

[1] At the time of writing this note, a preliminary version of the CP5-CR2 tune was derived including several jet substructure observables [24–26] in the fits. The tune, called CP5-CR2-j, has been used in the MC production in the CMS experiment. The CP5-CR2 and CP5-CR2-j tunes have very similar predictions in all final states discussed in this note, because the tunes differ slightly only in the following parameters, where for CP5-CR2-j: `MultipartonInteractions:ecmPow` = 0.056, `MultipartonInteractions:coreRadius` = 0.653, `MultipartonInteractions:coreFraction` = 0.439, `ColourReconnection:m2lambda` = 4.395, `MultipartonInteractions:fracGluon` = 0.990.

[2] The number of degrees of freedom for the tune CP5 is given as 63 in Ref. [18]. However, note that this value is the N_{dof} of the tune uncertainty when only 13 TeV distributions are used. The value of N_{dof} for CP5 consistent with our calculation in this note is 183.

3 Validation of the tunes

In this Section, a set of validation plots is shown for observables measured at a centre-of-mass energy of 13, 8, 7, and 1.96 TeV. In the Figures 4-17, CMS data points are shown in black, compared to predictions obtained from the PYTHIA8 event generator with the tunes CP5 (red), CP5-CR1 (blue), and CP5-CR2 (green). The lower panels show the ratios between each Monte Carlo prediction and the data.

3.1 The underlying-event and minimum-bias observables

"Minimum bias" is a generic term used to describe the class of events that are collected with the loosest event selection possible. Although this type of events generally correspond to inelastic events, including non-diffractive (ND) and diffractive (SD + DD + CD) contributions, the contributions of these processes may vary depending on the trigger requirements used in the experiments. For example, a sample of non-single-diffractive-enhanced (NSD-enhanced) events can be selected by suppressing the SD contribution at the trigger level. For all of the validation plots presented in this section, inelastic events (i.e. ND, SD, DD, and CD) are simulated with PYTHIA8.226 and compared to data at different centre-of-mass energies. Note that an update to the description of the elastic scattering part in PYTHIA8.235 led to a slight decrease (0.37 mb at 13 TeV) in the default non-diffractive cross section. In order to reproduce the conditions of PYTHIA8.226 in PYTHIA8.235 or in a newer version, one should set the non-diffractive cross section manually.

The UE observables, charged-particle multiplicity density and average p_{T} sum in the transMIN and transMAX regions, measured by the CMS experiment at 13 TeV [9] and shown in Figure 4, are well described by all tunes in the plateau region. The rising part of the spectrum, the region up to 3 GeV of $p_{\mathrm{T}}^{\mathrm{max}}$, is highly sensitive to diffractive contributions, and the optimisation of these components is out of the scope of this study. This part of the charged-particle multiplicity density distributions is described similarly by all tunes, whereas in the p_{T} sum density distributions, the predictions of CP5 differ slightly from the predictions of the CR tunes, showing that the CP5 has a harder p_{T} spectrum for low-p_{T} particles. Observables sensitive to the softer part of the MPI spectrum, such as the pseudorapidity distribution of charged hadrons in inelastic pp collisions measured by the CMS experiment at $\sqrt{s} = 13$ TeV [17] and shown in Figure 5, are also well described by the CP5 tunes. A crucial test for the performance of UE tunes, and of the CR simulation in particular, is the description of the average transverse momentum of the charged particles, as a function of the charged particle multiplicity. Comparisons to the measurements by the ATLAS experiment at $\sqrt{s} = 13$ TeV in the transMAX and transMIN regions [10] are displayed in Figure 6. CP5 describes data better than CR tunes due to its harder p_{T} spectrum in low-p_{T} particles, resulting in better correlations between the charged-particle average transverse momentum and charged-particle multiplicity. All CR tunes show a good agreement with the data, confirming the accuracy of the parameters obtained for the new CR models. The improvement in the tuned CR models and their success in describing the data can clearly be seen comparing Figure 4 with Figure 2, and Figure 5 with Figure 3.

In Figure 7, charged particle and $p_{\mathrm{T}}^{\mathrm{sum}}$ densities measured by the CMS experiment at 7 TeV [11] in transMIN and transMAX regions, as a function of the p_{T} of the leading charged particle, $p_{\mathrm{T}}^{\mathrm{max}}$, are compared to predictions from the tunes CP5 and CP5-CR. The central values of the data are well described for $p_{\mathrm{T}}^{\mathrm{max}} > 5$ GeV, except that of $p_{\mathrm{T}}^{\mathrm{sum}}$ by CP5-CR2 in the transMAX region. In Figure 8 charged-particle and $p_{\mathrm{T}}^{\mathrm{sum}}$ densities in the transverse region, as a function of the p_{T} of the leading charged particle, and average transverse momentum in the transverse region as a function of the leading charged particle p_{T} and of the charged particle multiplicity,

measured by the ATLAS experiment at 7 TeV [12], are compared to the predictions from the tunes CP5 and CP5-CR. The central values of the charged-particle and $p_{\mathrm{T}}^{\mathrm{sum}}$ densities above 5 GeV is described well by the tunes. The central values of the average transverse momentum in bins of the leading charged particle p_{T} and of the charged particle multiplicity are consistent with the data points within 10%. A similar level of agreement as observed at 13 TeV is achieved by the new tunes at 7 TeV. The performance of the new tunes is checked at 7 TeV using also inclusive measurements of charged-particle pseudorapidity distributions. The CMS data at $\sqrt{s} = 7$ TeV [27] on the pseudorapidity of charged particles, $dN_{\mathrm{ch}}/d\eta$, with at least one charged particle in $|\eta| < 2.4$, are compared to predictions from the tunes CP5, and CP5-CR, in Figure 9. As can be seen from the figure, CP5 and CP5-CR1 have similar predictions while CP5-CR2 predicts about 4% less charged particles than them in all η bins of the measurement. Although all tunes provide a reasonable description of $dN_{\mathrm{ch}}/d\eta$ with deviations of up to ∼10%, the data and MC show different trends for $|\eta| > 1.2$. The decreasing trend for data in $|\eta| > 1.2$ is not well described by the tunes. In the more central region, i.e. $|\eta| < 1.2$, the shape of the predictions agrees well with the data but there is a difference in normalisation; for example, CP5 and CP5-CR1 predict 3–4% and CP5-CR2 predict about 7% less charged particles in all bins for $|\eta| < 1.2$ compared to the data.

In Figure 10, charged particle and $p_{\mathrm{T}}^{\mathrm{sum}}$ densities measured at $\sqrt{s} = 1.96$ TeV by the CDF experiment [13] in transMIN and transMAX regions, as a function of the p_{T} of the leading charged particle, are compared to predictions from the tunes CP5 and CP5-CR, respectively. All predictions reproduce well the UE observables at $\sqrt{s} = 1.96$, 7, and 13 TeV.

We compare the new CMS tunes with MB and UE data measured at forward pseudorapidities that are not used in the fits. Observables measured at large pseudorapidities, such as the forward energy flow in two different selections, as measured by the CMS experiment at 13 TeV [28], and shown in Figure 11 are well reproduced by the CMS tunes.

3.2 Particle multiplicities

Figure 12 shows the strange particle production, for Λ baryons and K_s^0 mesons, as a function of rapidity, measured by the CMS experiment [29] in non-single diffractive (NSD) events at $\sqrt{s} =$ 7 TeV. The NSD events are simulated with PYTHIA8.226 and the predictions of the tunes are compared to data as shown in the figure. In Ref. [7], it was shown that the new colour reconnection models might be beneficial for describing the ratios of strange particles multiplicities, for example Λ/K_S^0 in pp collisions. We observe that all CP5 tunes, regardless of the colour reconnection model, describe particle production for K_S^0 vs rapidity very well. However, they underestimate particle production for Λ vs rapidity by about 30%. Therefore, the ratio Λ/K_S^0 is not perfectly described, but this could be cured by different hadronisation models [30, 31].

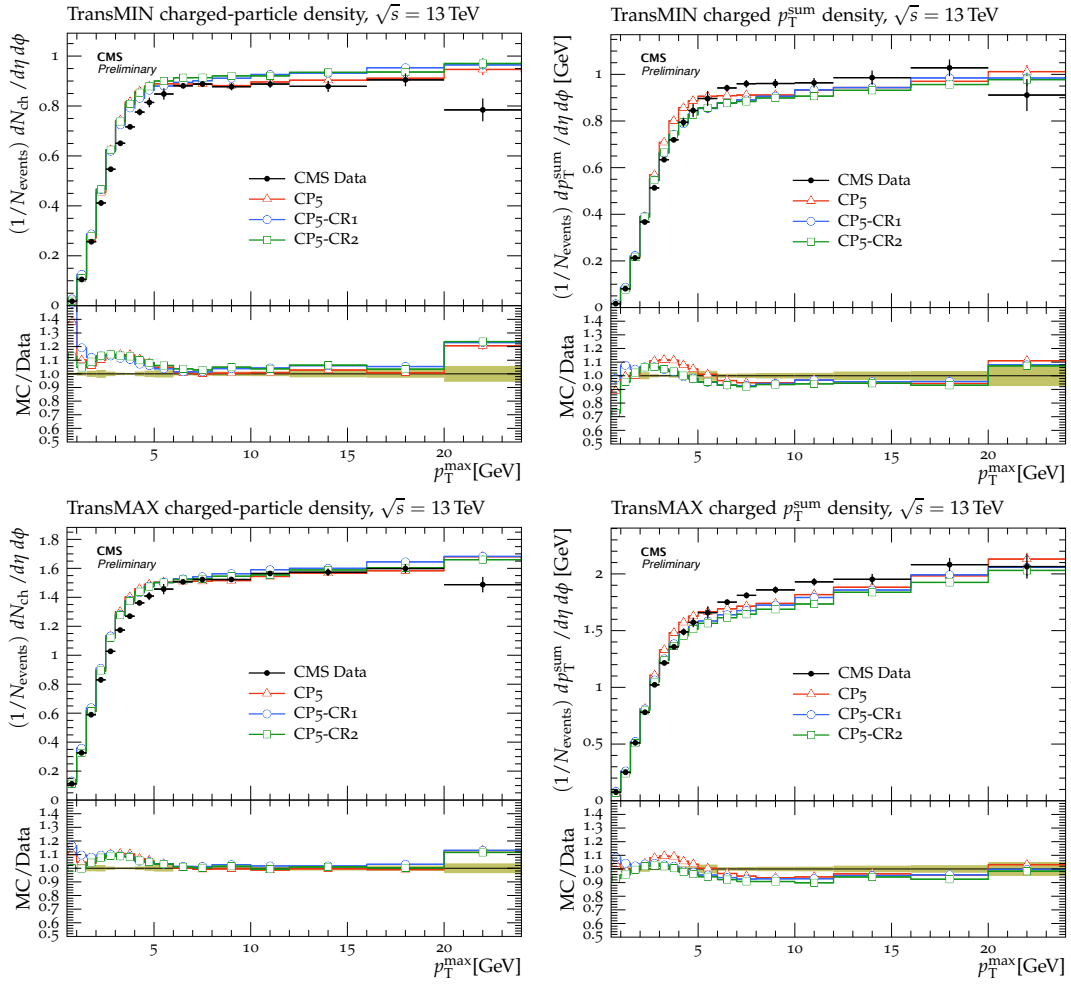

Figure 4: The charged particle (left) and p_T^{sum} (right) densities in the transMIN (upper) and transMAX (lower) regions, as a function of the p_T of the leading charged particle, p_T^{max}, measured by the CMS experiment at $\sqrt{s} = 13$ TeV [9]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

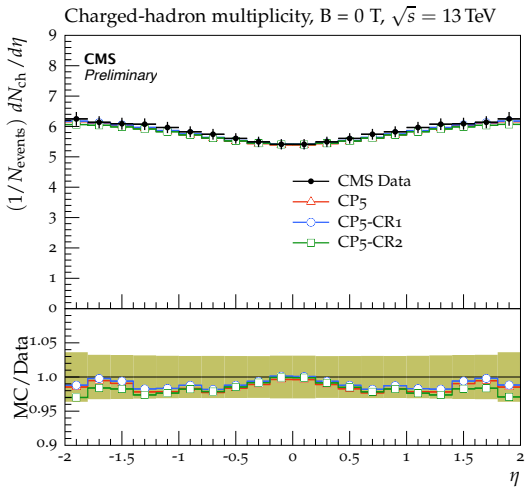

Figure 5: The pseudorapidity of charged hadrons, $dN_{\text{ch}}/d\eta$, measured by the CMS experiment at $\sqrt{s} = 13\,\text{TeV}$ [17]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

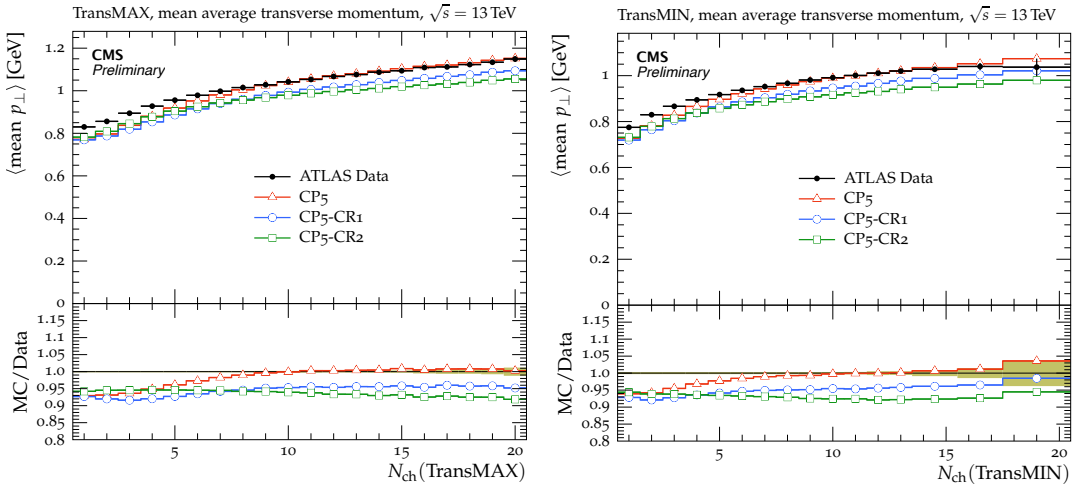

Figure 6: The mean charged-particle average transverse momentum as a function of charged-particle multiplicity in transMAX (left) and transMIN (right) regions, measured by the ATLAS experiment at $\sqrt{s} = 13\,\text{TeV}$ [10]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

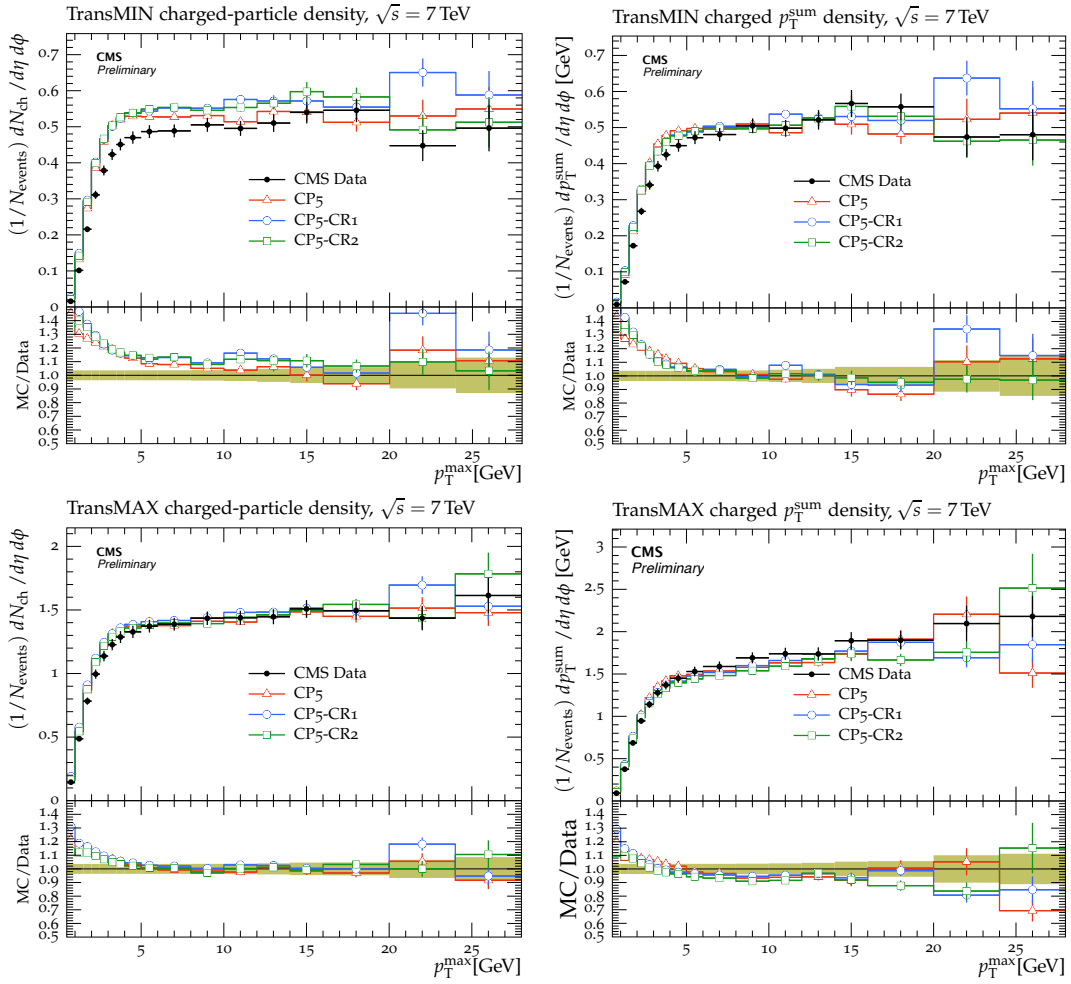

Figure 7: The charged particle (left) and p_T^{sum} (right) densities in the transMIN (upper) and transMAX (lower) regions, as a function of the p_T of the leading charged particle, p_T^{max}, measured by the CMS experiment at $\sqrt{s} = 7$ TeV [11]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

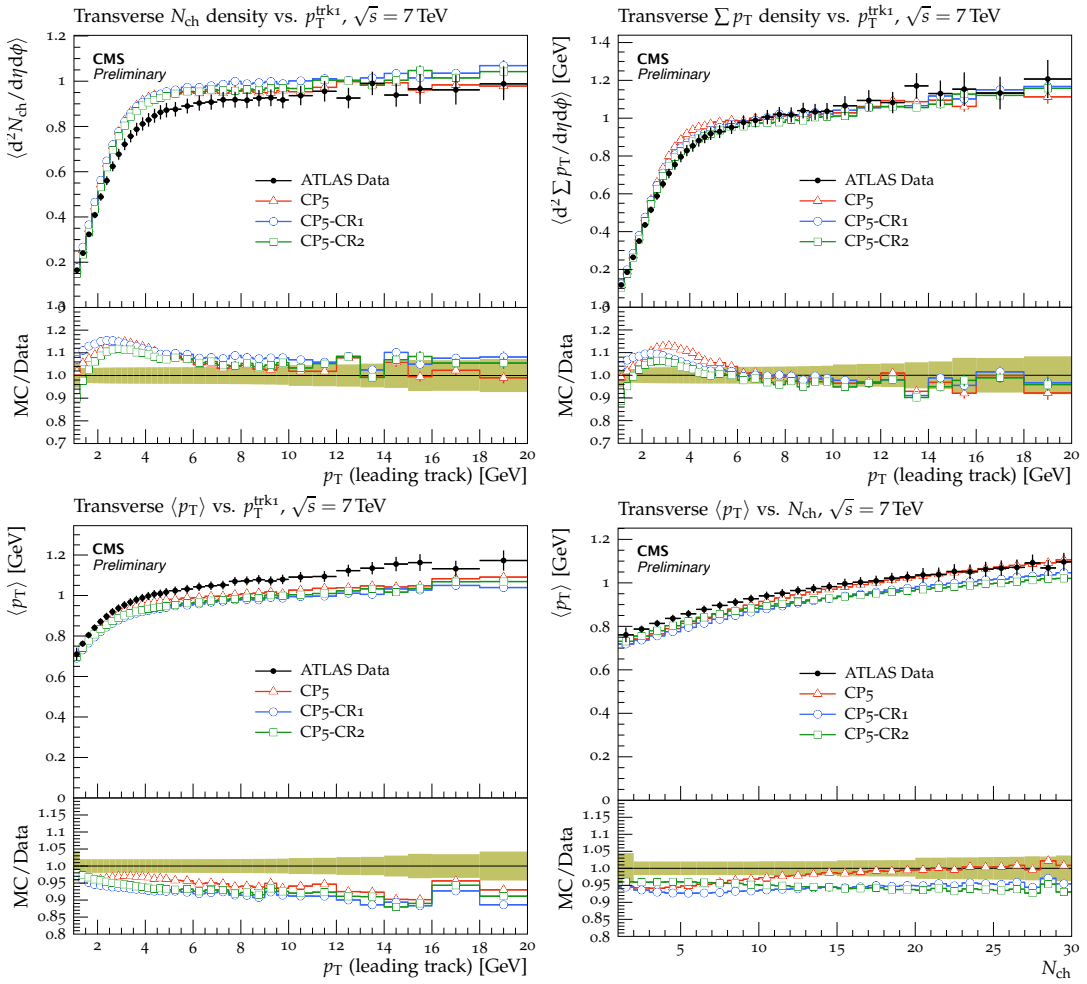

Figure 8: The charged-particle (upper left) and p_T^{sum} densities (upper right) in the transverse region, as a function of the p_T of the leading charged particle, and on average transverse momentum in the transverse region as a function of the leading charged particle p_T (lower left) and of the charged particle multiplicity (lower right), measured by the ATLAS experiment at $\sqrt{s} = 7$ TeV [12]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

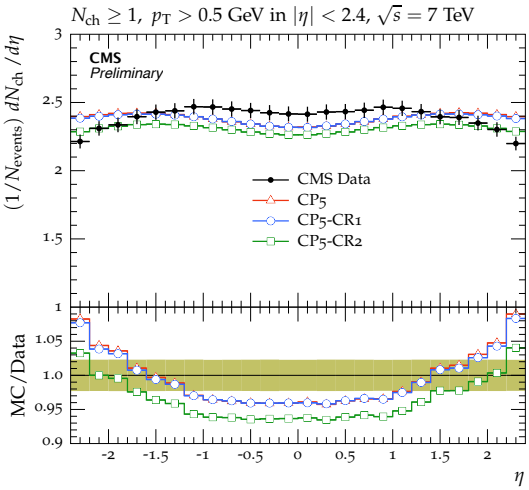

Figure 9: The pseudorapidity of charged particles, $dN_{ch}/d\eta$, with at least one charged particle in $|\eta| < 2.4$, measured by the CMS experiment at $\sqrt{s} = 7$ TeV [27]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

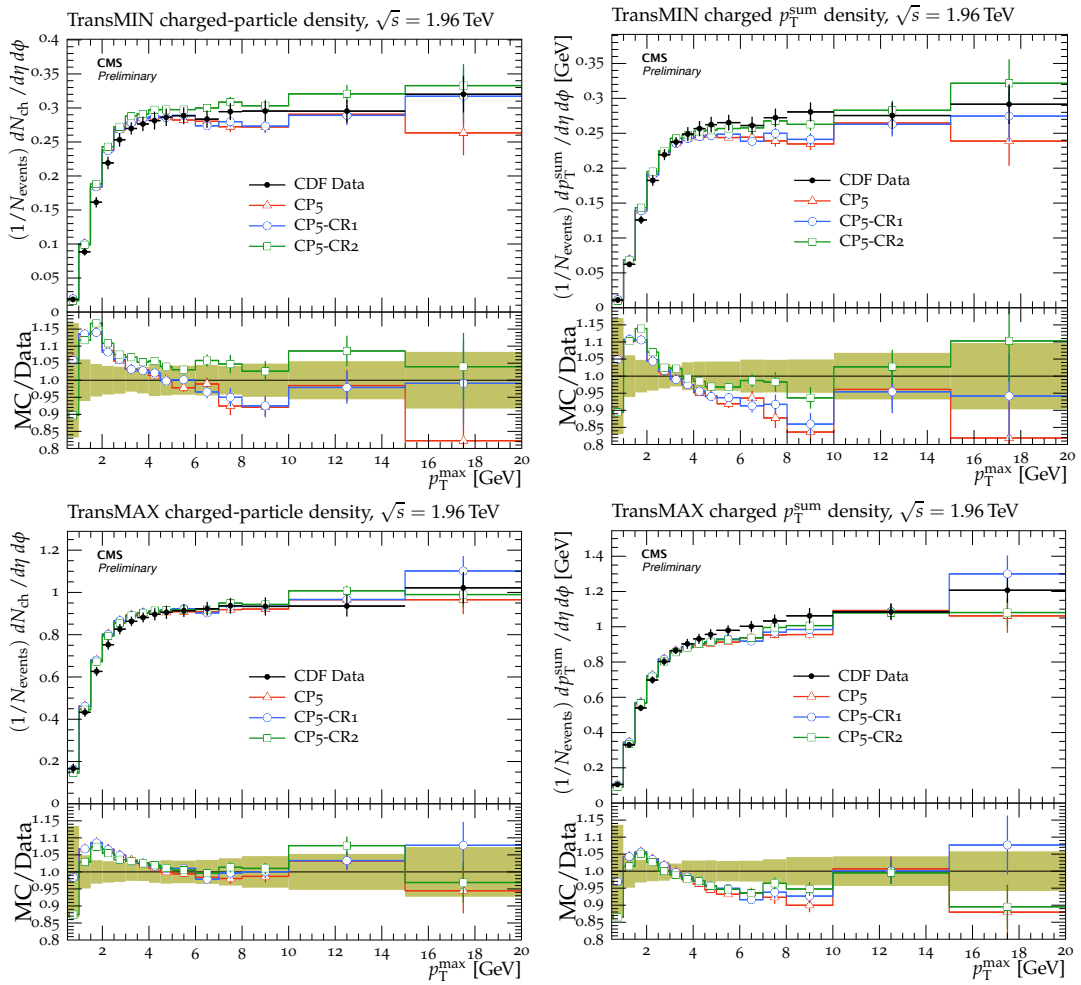

Figure 10: The charged particle (left) and p_T^{sum} densities (right) in the transMIN (upper) and transMAX (lower) regions, as a function of the p_T of the leading charged particle, p_T^{max}, measured by the CDF experiment at $\sqrt{s} = 1.96$ TeV [13]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

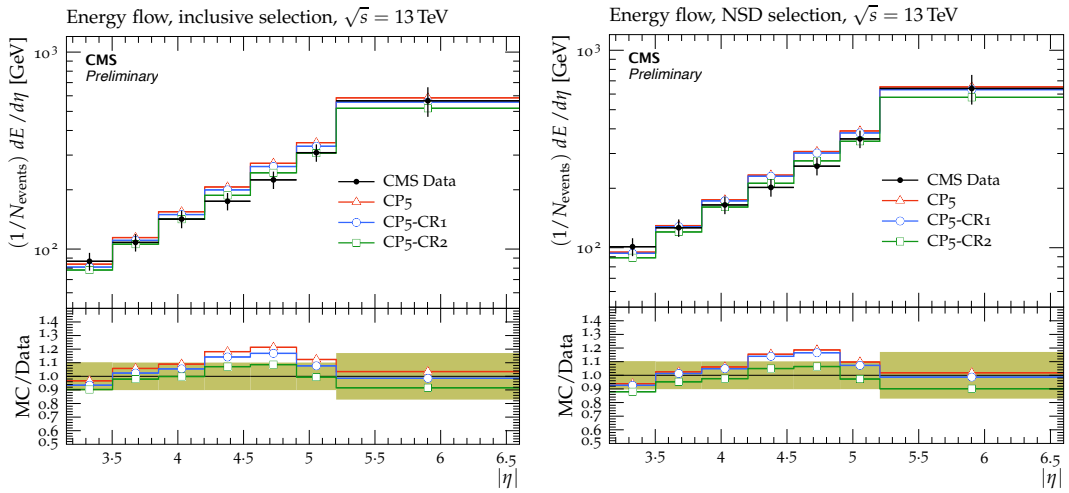

Figure 11: The forward energy flow as a function of pseudorapidity in two different selections, in minimum-bias events (left) and in events with a presence of a hard dijet system (right), measured by the CMS experiment at $\sqrt{s} = 13$ TeV [28]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

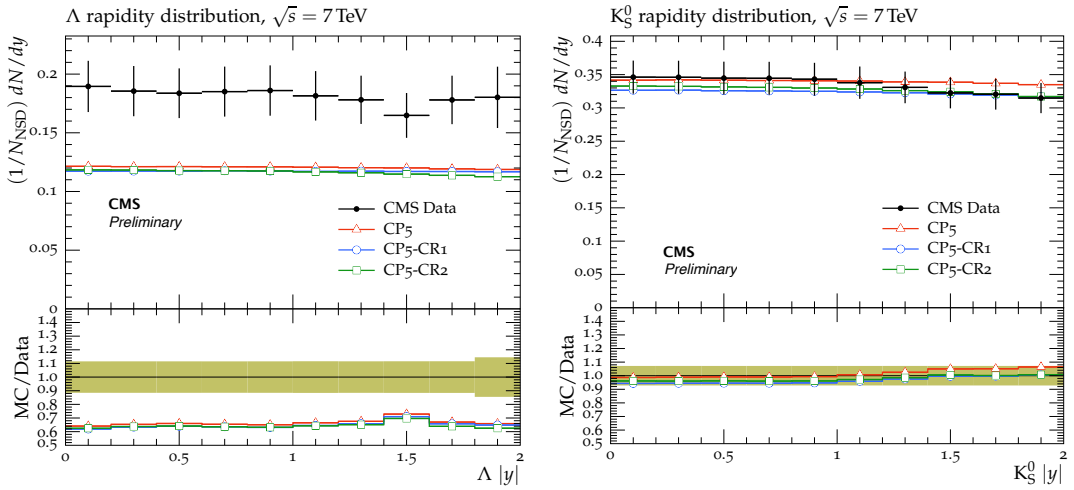

Figure 12: The strange particle production, Λ baryons (left) and K^0_s mesons (right), as a function of rapidity, measured by the CMS experiment at $\sqrt{s} = 7$ TeV [29]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

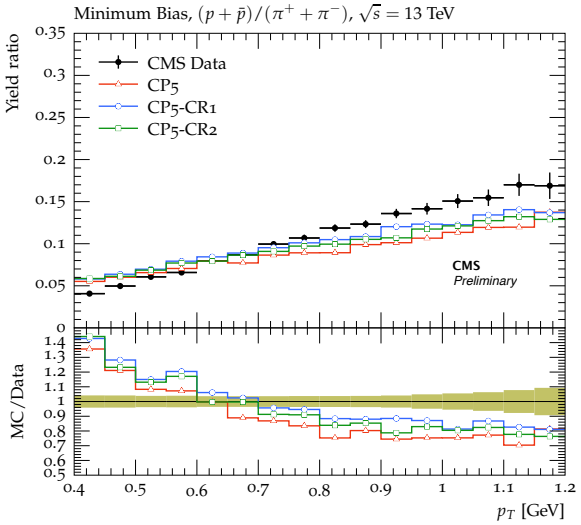

Figure 13: Ratios of particle yields, p/π, as a function of transverse momentum in miminum bias events, measured by the CMS experiment at $\sqrt{s} = 13$ TeV [32]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

We also investigated the multiplicities of identified particles in simulated MB events (ND+SD+DD+CD). Figure 13 shows the ratio of proton over pion production as a function of particle p_{T} [32]. All of the tunes predict a similar trend, showing that the new CR models do not have a significant improvement in the ratio of proton to pion production. However, it is known that this observable is strongly correlated with event particle multiplicity and not only colour reconnections, while hadronisation and multiparton interactions also play a key role in describing the ratios of particle yields. The ratios of all light, charm, and bottom baryons to mesons are shown in Figure 14, together with charm hadron data from e^+e^- colliders [33] and bottom hadron data from the DELPHI collaboration [34]. This can be considered as one of the key distributions in order to correctly obtain the identified particle spectra and it seems that the tunes describe well the baryon-to-meson ratios at different centre-of-mass energies.

3.3 Jet substructure observables

The number of charged particles inside jets is one of the observables which makes it possible to distinguish quark-initiated jets from gluon-initiated jets. The average number of charged hadrons with $p_{\mathrm{T}} > 500$ MeV inside the jets measured by the CMS experiment as a function of the jet p_{T} is given in Figure 15 [24]. The predictions of the CR tunes are comparable and they produce roughly 5% less charged particles than the CP5 one. All predictions have a reasonable description of the data.

In Figure 16, distributions of $F(z) = dN_{ch}/(N_{\mathrm{jet}}dz)$, where z is the longitudinal momentum fraction, and N_{ch} is the charged-particle multiplicity in the jet, are measured. $F(z)$ is a parameter related to fragmentation function, and is presented for $p_{\mathrm{T}}^{\mathrm{jet}} = 25 - 40\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{jet}} = 400 - 500\,\mathrm{GeV}$ [25]. The CR tunes describe low-$p_{\mathrm{T}}^{\mathrm{jet}}$ data better than CP5 and their predictions reasonably agree with the high-$p_{\mathrm{T}}^{\mathrm{jet}}$ data except the last bin. The high-$p_{\mathrm{T}}^{\mathrm{jet}}$ data is well described by the CP5 tune within uncertainties.

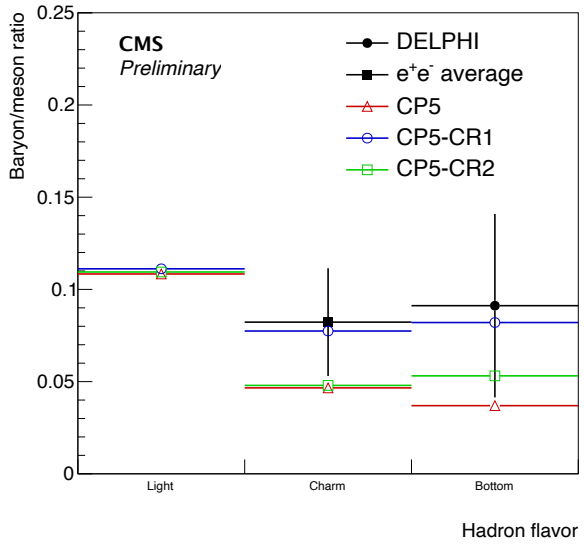

Figure 14: Ratios of particle yields for light, charm, and bottom hadrons predicted by the the CP5 and CP5-CR tunes compared to data from e^+e^- colliders [33] and the DELPHI collaboration [34].

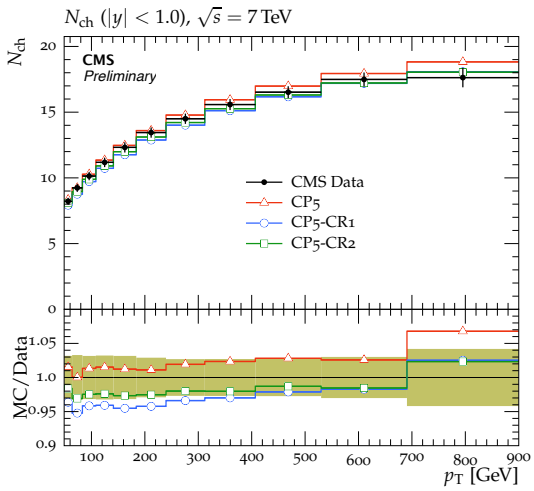

Figure 15: Average charged-hadron multiplicity as a function of the jet p_T for jets with rapidity $|y| < 1$, measured by the CMS experiment at $\sqrt{s} = 7$ TeV [24]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

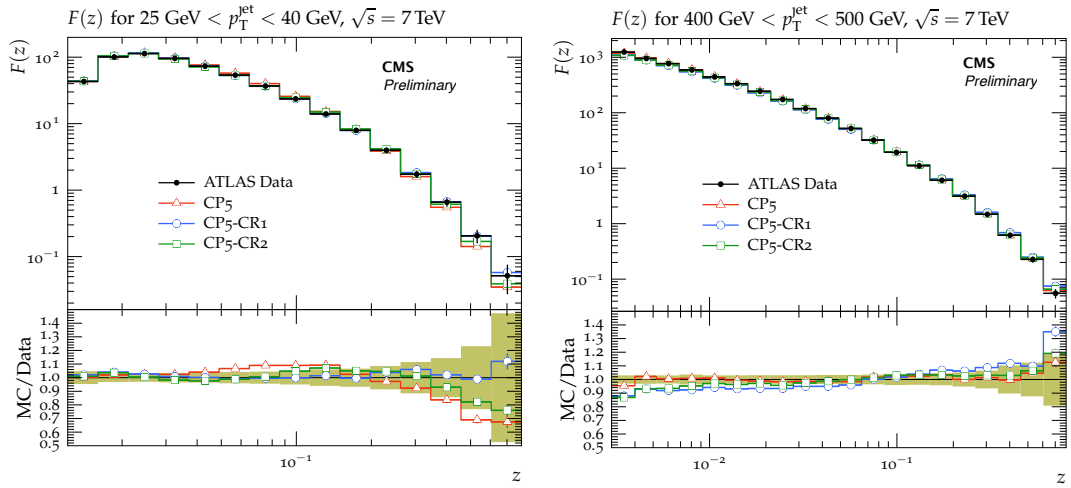

Figure 16: Distributions of $F(z)$ for 25 GeV $< p_T^{\text{jet}} <$ 40 GeV (left) and 400 GeV $< p_T^{\text{jet}} <$ 500 GeV (right) for jets with pseudorapidity $|\eta_{\text{jet}}| < 1.2$, measured by the ATLAS experiment at $\sqrt{s} = 7$ TeV [25]. The predictions of the CP5 and CP5-CR tunes are compared to data. The coloured band represents the total experimental uncertainty in the data.

3.4 Drell–Yan events

Drell–Yan (DY) events with the Z boson decaying to $\mu^+\mu^-$ were generated with PYTHIA8 and compared to CMS data at 13 TeV. Figure 17 shows the charge multiplicity and transverse momentum flow as a function of the Z boson p_{T} (in the invariant $\mu^+\mu^-$ mass window of 81–101 GeV) in the region transverse to the boson momentum [35], which is expected to be dominated by the underlying event.

The CP5 tunes predict up to 15% too many charged particles at low Z boson p_{T}, where additional effects like the intrinsic transverse momentum of the interacting partons (i.e., primordial k_{T}) are expected to play a role. Higher-order corrections as implemented in MADGRAPH5_aMC@NLO [36] with FxFx merging [37] are necessary to describe the total p_{T} flow. The impact of the different CR models is found to be negligible in DY events.

3.5 Top quark observables

3.5.1 UE and jet substructure in $\mathrm{t\bar{t}}$ events

A study of the UE in top quark pair ($\mathrm{t\bar{t}}$) events [15] showed that, even in the $\mathrm{t\bar{t}}$ environment, CR effects in UE observables are subtle. This study utilized events with one electron (e), one muon (μ) and two jets. It is shown that CR is needed to increase the accuracy of predictions for $\overline{p_{\text{T}}} < 3$ GeV, $\overline{p_z} < 5$ GeV, average aplanarity, $\overline{p_{\text{T}}}$ vs $p_{\text{T}}(e^{\pm}\mu^{\mp})$ and $\overline{p_{\text{T}}}$ vs jet multiplicity. Here, $\overline{p_{\text{T}}}$ (or $\overline{p_z}$) stands for the average p_{T} (or p_z) per charged particle calculated using the ratio of the scalar sum of the p_{T} (or p_z) divided by the charged multiplicity. It is also shown in Ref. [15] that when rope hadronisation [38, 39] is used, CR is needed to improve the description of $\overline{p_{\text{T}}}$, $\overline{p_z}$, $p_{\text{T}}(e^{\pm}\mu^{\mp})$, $m(e^{\pm}\mu^{\mp})$, and average $\overline{p_{\text{T}}}$ in different $p_{\text{T}}(e^{\pm}\mu^{\mp})$ and jet multiplicity categories for toward, transverse, and away UE regions. In addition to studying the differences between the QCD-based, gluon move, and rope hadronisation models, the Ref. [15] also estimated the effects of the CR on the top quark decay products, by investigating the differences between predictions using PYTHIA8 with the option early resonance decay (ERD) off and on ("erdOn" in Table 4). The ERD off and on options allow the colour reconnection to take place, respectively, before or after the top quark decay. In particular, the ERD option allows the top

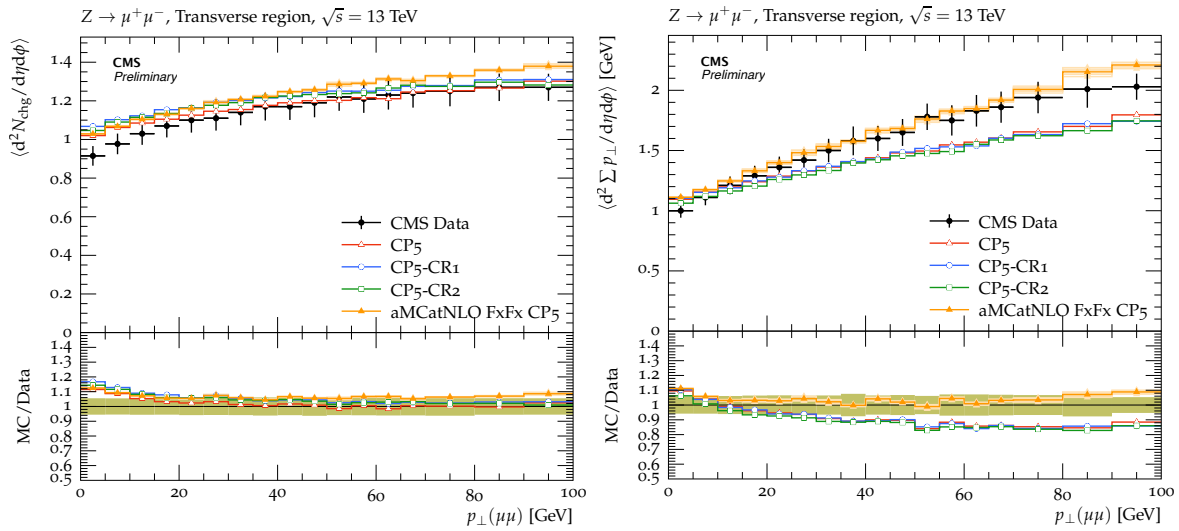

Figure 17: Number of charged particles and p_T flow in the transverse region of DY events, measured by the CMS experiment at $\sqrt{s} = 13$ TeV in bins of Z boson p_T [35]. The plots show the predictions of the CP5 and CP5-CR tunes compared to data. The coloured band represents the total experimental uncertainty in the data.

quark decay products to be colour reconnected with the partons from MPI systems. Ref. [15] showed that these different models and options give similar predictions for UE observables in $t\bar{t}$ events. However, some jet-shape distributions in $t\bar{t}$ events display a more significant effect [40], e.g. in number of charged particles in jets. In the following, we investigate how the new PYTHIA8 CR tunes describe the CMS $t\bar{t}$ jet substructure data [40]. In the CMS measurement, jets reconstructed using the anti-k_T algorithm [41] with a distance parameter of $R = 0.4$ as implemented in FASTJET3.1 [42] are used. Jets with $p_T > 30$ GeV within $|\eta| < 2$ are selected. Jet pairs are required to be far from each other in $\eta - \phi$ space: $\Delta R(jj) > 0.8$. The jet substructure observables are calculated from the jet constituents that have $p_T > 1$ GeV to avoid the rapid changes in tracking efficiency and misidentification rate below 1 GeV. Here we focus on two variables, $\lambda_0^0(N)$ and the angle between two groomed subjets (ΔR_g) which are shown in Figure 18. The variable $\lambda_0^0(N)$ is defined to be the number of charged particles with $p_T > 1$ GeV in the jet. Groomed jets refer to jet from which soft and wide-angle radition are systematically removed [43, 44]. It is observed that none of the tunes describe the data well for these two observables. In Ref. [40], these observables are also presented separately for the samples enriched in quarks and gluons jets and it is observed that the agreement between the data and MC is significantly worse for the bottom quark jets than for the light-quark- and gluons-enriched jet samples. As concluded in the original note, flavor-dependent improvements in the physics modeling or an update in the MC parameter tuning may be required for a better description of the data. The inclusion of observables of the jet substructure as well as different hadronization models in the tuning can help to better understand and describe these observables.

3.5.2 Pull angle in $t\bar{t}$ events

Figure 19 displays the pull angle between jets originating from the decay of a W boson in $t\bar{t}$ events, as measured by the ATLAS experiment [45]. The observable is shown for the case where only the charged constituents of the jet are used in the calculation. The data are compared against predictions from POWHEG+PYTHIA8 using the CP5 tunes or the corresponding CR tunes. This observable is particularly sensitive to the setting of the ERD option. With ERD

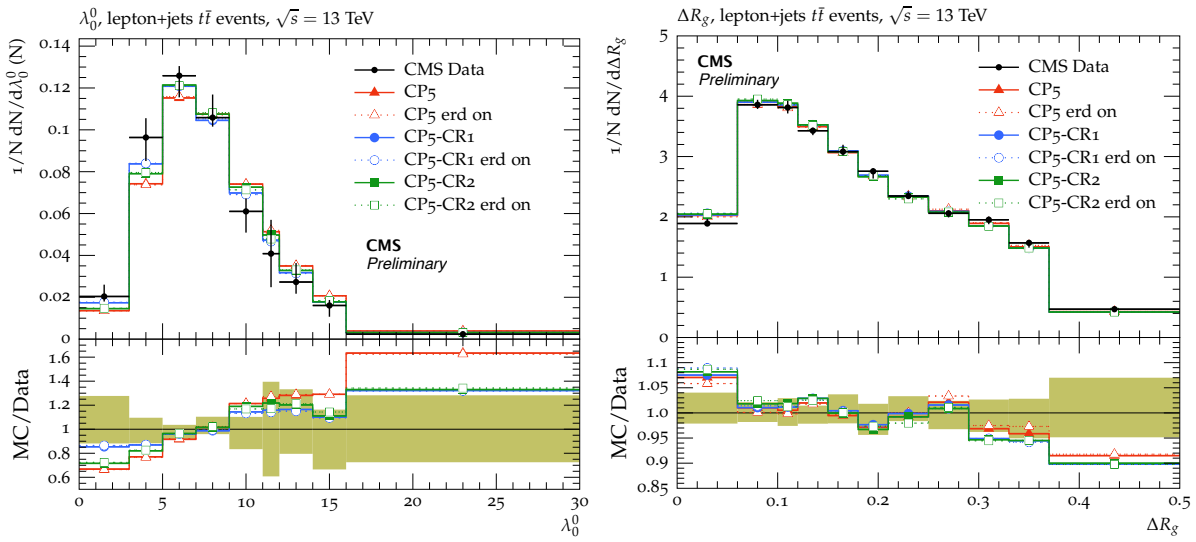

Figure 18: Distributions of the particle multiplicity (left) and the angle between two groomed subjets (ΔR_g) (right) measured by the CMS experiment at $\sqrt{s} = 13$ TeV [40]. The coloured band represents the total experimental uncertainty in the data.

turned off the decay products of the W boson in $t\bar{t}$ events are not included in colour reconnection, and the predictions using the tunes with the various colour reconnection models are similar to each other. With ERD enabled, colour reconnection can now modify the pull angle between the two jets, which is observed in Figure 19. The predictions of each tune also show significant differences between each other when ERD is enabled. For both the nominal and CR1 (QCD-inspired) tunes, the prediction with ERD improves the description of the data, and the difference between the predictions with ERD and no ERD is larger for the CR1-based tunes. We observe the opposite for the CR2-based (gluon-move) tunes. The choice without ERD is preferred here. Note that this picture might be different if the flip mechanism was added in the tuning of the gluon-move model. The move step in the gluon-move model is a bit restrictive because it allows only gluons to move between the string end-points. The inclusion of the flip mechanism would also allow the string end-points to be mixed with each other and therefore could further reduce the total string length in an event. However, as mentioned earlier, the flip mechanism is not included due to its unperceived effect in single diffractive events. Overall, the QCD-inspired model with ERD provides the best description of the jet pull angle. The differences between the predictions using the different tunes observed here indicate that the inclusion of observables such as the jet pull angle and other jet substructure observables could be beneficial in future tune derivations. Extending this study with comparisons to the observables measured in Ref. [46] would be of interest.

4 Uncertainty on the top mass due to colour reconnection

The mass of the top quark, m_t, has been measured with high precision using the 7, 8, and 13 TeV data at the LHC [16, 47–59]. The most precise value of the $m_t = 172.44 \pm 0.13(stat) \pm 0.47(syst)$ GeV was measured by the CMS Collaboration combining 7 and 8 TeV data [53]. In order to improve further the precision of top quark mass measurements, a complete analysis of the systematic uncertainties on the measurement is crucial. One of the dominant systematic model uncertainties is due to the possible reconnections of the colour flow in the top quark decays [53]. The procedure for estimating this uncertainty used for Run I analyses was based on

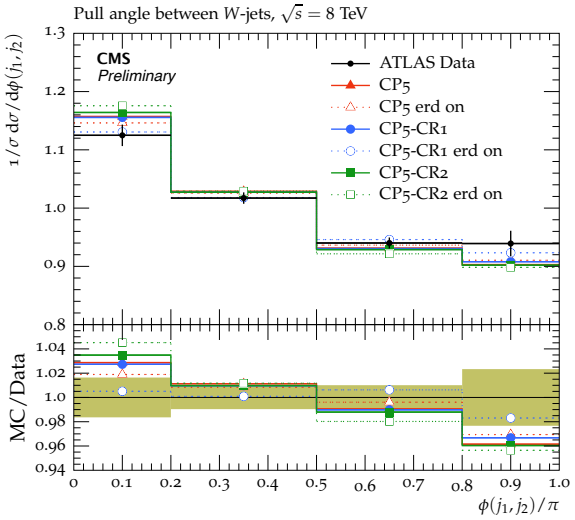

Figure 19: Pull angle between jets from the W boson in top quark decays, calculated from the charged constituents of the jets, measured by the ATLAS experiment using $\sqrt{s} = 8$ TeV data to investigate colour flow. The coloured band represents the total experimental uncertainty in the data.

a comparison of two values of m_t, calculated by using predictions using the same UE tune with and without CR effects. The new CMS tunes, presented in Section 2 and which use different CR models, can be used to give a better evaluation of the CR uncertainty. In particular, the uncertainty can now be calculated by comparing results for m_t values obtained from different realistic CR models, such as the default model in PYTHIA8 and the other ones, QCD-inspired and gluon-move. Additionally, one can also estimate the effects of the CR on the top quark decay products, by investigating the differences between predictions using PYTHIA8 with the option ERD off and on, as e.g. done Ref. [15] for UE event observables. A determination of m_t using a kinematic reconstruction of the decay products in the semi-leptonic (in which one of the W bosons from $t\bar{t}$ decays into a muon or electron and a neutrino, and the other into a quark-antiquark pair) $t\bar{t}$ events at 13 TeV using 35.9 fb^{-1} can be found in Ref. [16]. In this analysis, the top quark mass and the jet energy scale factor are determined simultaneously through a joint-likelihood fit to the selected events. The results with QCD-inspired and gluon-move modes were also compared. The nominal PYTHIA8 UE tune used was CUETP8M2T4 [60], and the parameters of the CR models were tuned to UE and MB data at 13 TeV [16]. It is found that the gluon move model results in a larger shift in the top quark mass value of 0.31 GeV, and is taken as the uncertainty in the modelling of colour reconnection on the measured top quark mass. This is the largest source of uncertainty in the measured top quark mass, where the total uncertainty is 0.62 GeV. Without this source, the total uncertainty would be 0.54 GeV.

We compare the top quark and W boson mass values obtained with different tune configurations based on our new tunes in Table 4. Top quark candidates are constructed by a RIVET routine in a sample of simulated semi-leptonic $t\bar{t}$ events. Events must contain exactly one lepton with $p_T > 30$ GeV, and $|\eta| < 2.1$. Leptons are "dressed" with the surrounding photons within a cone of 0.1, and are required to yield a leptonically decaying W boson candidate with an invariant mass within 5 GeV of 80.4 GeV when combined with a neutrino in the event. The events must also contain at least four jets, reconstructed with the anti-k_T algorithm, with $p_T > 30$ GeV within $|\eta| < 2.4$. At least two of the jets are required to originate from the decay of a bottom quark, whilst at least another two jets, referred to as light-quark jets, must not originate from a

bottom quark. One jet originating from a bottom quark is combined with the lepton and neutrino to form a leptonically-decaying top quark candidate, whilst another jet originating from a bottom quark is combined with two other jets to form a hadronically-decaying top quark candidate. The difference in invariant mass of the two top quark candidates must be less than 20 GeV, and the invariant mass of the two light-quark jets must be within 10 GeV of 80.4 GeV. If more than one combination of jets satisfy these criteria when combined with the lepton and neutrino, then only one combination is chosen based on how similar the invariant masses of the two top quark candidates are to each other, and on how close the invariant mass of the light-quark jets are to 80.4 GeV. The top quark and W boson mass values are then obtained by fitting a Gaussian with an 8 GeV mass window around the corresponding mass peak. The table also contains the differences from the nominal m_t and m_W values, and the difference $m_t - 0.5 \times m_W$. The latter quantity is used in Ref. [16] to obtain an estimation of the m_t measurement uncertainty taking into account the shift in m_W but giving it a certain weight optimised for each measurement. From Table 4, we observe that the largest deviation from the predictions of CP5 is CP5-CR2 erdOn (0.32 GeV) similar to the shift found in Ref. [16] in the hybrid method (which gives the lowest total uncertainty) using CUETP8M2T4.

Table 4: The top quark mass, m_t, and W mass, m_W, extracted by a fit to the predictions of the different PYTHIA8 tunes. The uncertainties in the m_t and m_W values correspond to the uncertainty in the fitted m_t and m_W.

Tune	m_t [GeV]	Δm_t [GeV]	m_W [GeV]	Δm_W [GeV]	$\Delta m_t - 0.5 \times \Delta m_W$ [GeV]
CP5	171.93 ± 0.02	0	79.76 ± 0.02	0	0
CP5 erdOn	172.18 ± 0.03	0.25	80.15 ± 0.02	0.40	0.13
CP5-CR1	171.97 ± 0.02	0.04	79.74 ± 0.02	-0.02	0.05
CP5-CR1 erdOn	172.01 ± 0.03	0.08	79.98 ± 0.02	0.23	-0.04
CP5-CR2	171.91 ± 0.02	-0.02	79.85 ± 0.02	0.10	-0.07
CP5-CR2 erdOn	172.32 ± 0.03	0.39	79.90 ± 0.02	0.14	0.32

5 Summary

New sets of parameters for two of the colour reconnection (CR) models, QCD-inspired and gluon-move, implemented in the PYTHIA8 event generator are obtained, based on the default CMS CP5 PYTHIA8 tune. Measurements sensitive to underlying-event (UE) contributions performed at $\sqrt{s} = 1.96$, 7, and 13 TeV are used to constrain the parameters for the CR and for the multiple parton interactions simultaneously. Different measurements at a centre-of-mass energy of 1.96, 7, 8, and 13 TeV are used to evaluate the performance of the new tunes. The central values predicted by the new CR tunes for the UE and minimum-bias distributions significantly better describe the data than the CR models with their default parameters before tuning. The predictions of the new tunes are able to achieve a very good level of agreement against many underlying-event observables including UE data measured at forward pseudorapidities. However, it should be noted that the models after tuning perform no better than the CP5 tune for the observables presented in this study. The new CR tunes are also tested against measurements of strange particle multiplicities for Λ baryons and K^0_s mesons. It is shown that new CR models alone do not improve the description of the strange particle production versus rapidity distribution for Λ baryons. It is observed that all CP5 tunes, irrespective of the CR model, describe particle production for K^0_S versus rapidity very well. The predictions of the new tunes to jet shapes and colour flow measurements done with top quark pair events are also compared to data. None of the tunes describe the jet shapes distributions considered in this note well, and all tunes have similar predictions. Some differences are also observed with respect to the colour flow data which is particularly sensitive to the ERD option in the CR models. The differences between the predictions using the different tunes observed here indicate that the inclusion of observables such as the jet pull angle and other jet substructure observables could be beneficial in tuning studies. A study of uncertainty on the top quark mass measurement due to CR effects is also presented which shows that CR will continue to be one of the dominating uncertainty sources in the top quark mass measurements.

References

[1] T. Sjöstrand et al., "An introduction to Pythia 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[2] T. Sjöstrand and M. van Zijl, "A multiple interaction model for the event structure in hadron collisions", *Phys. Rev. D* **36** (1987) 2019, doi:10.1103/PhysRevD.36.2019.

[3] T. Sjöstrand, "Colour reconnection and its effects on precise measurements at the LHC", LU-TP 13-37 and MCnet-13-16, 2013. arXiv:1310.8073.

[4] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++", *Eur. Phys. J. C* **72** (2012) 2225, doi:10.1140/epjc/s10052-012-2225-5, arXiv:1206.0041.

[5] S. Gieseke, P. Kirchgaeßer, and S. Plätzer, "Baryon production from cluster hadronisation", *Eur. Phys. J. C* **78** (2018) 99, doi:10.1140/epjc/s10052-018-5585-7, arXiv:1710.10906.

[6] T. Sjöstrand and P. Z. Skands, "Multiple interactions and the structure of beam remnants", *JHEP* **03** (2004) 053, doi:10.1088/1126-6708/2004/03/053, arXiv:hep-ph/0402078.

[7] J. R. Christiansen and P. Z. Skands, "String formation beyond leading colour", *JHEP* **08** (2015) 003, doi:10.1007/JHEP08(2015)003, arXiv:1505.01681.

[8] S. Argyropoulos and T. Sjöstrand, "Effects of color reconnection on $t\bar{t}$ final states at the LHC", *JHEP* **11** (2014) 043, doi:10.1007/JHEP11(2014)043, arXiv:1407.6653.

[9] CMS Collaboration, "Underlying event measurements with leading particles and jets in pp collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-FSQ-15-007, 2015.

[10] ATLAS Collaboration, "Measurement of charged-particle distributions sensitive to the underlying event in $\sqrt{s} = 13$ TeV proton-proton collisions with the ATLAS detector at the LHC", *JHEP* **03** (2017) 157, doi:10.1007/JHEP03(2017)157, arXiv:1701.05390.

[11] CMS Collaboration, "Measurement of the underlying event activity in pp collisions at the LHC at 7 TeV and comparison with 0.9 TeV", CMS Physics Analysis Summary CMS-PAS-FSQ-12-020, 2012.

[12] ATLAS Collaboration, "Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900$ GeV and 7 TeV with the ATLAS detector", *Phys. Rev. D* **83** (2011) 112001, doi:10.1103/PhysRevD.83.112001, arXiv:1012.0791.

[13] CDF Collaboration, "Study of the energy dependence of the underlying event in proton-antiproton collisions", *Phys. Rev. D* **92** (2015) 092009, doi:10.1103/PhysRevD.92.092009, arXiv:1508.05340.

[14] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[15] CMS Collaboration, "Study of the underlying event in top quark pair production in pp collisions at 13 TeV", *Eur. Phys. J. C* **79** (2019) 123, doi:10.1140/epjc/s10052-019-6620-z, arXiv:1807.02810.

[16] CMS Collaboration, "Measurement of the top quark mass with lepton+jets final states using p p collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **78** (2018) 891, doi:10.1140/epjc/s10052-018-6332-9, arXiv:1805.01428.

[17] CMS Collaboration, "Measurement of pseudorapidity distributions of charged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV by the CMS experiment.", CMS Physics Analysis Summary CMS-PAS-FSQ-15-008, 2016.

[18] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[19] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[20] CMS Collaboration, "Pseudorapidity distribution of charged hadrons in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Phys. Lett. B* **751** (2015) 143, doi:10.1016/j.physletb.2015.10.004, arXiv:1507.05915.

[21] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", *Eur. Phys. J. C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[22] A. Buckley et al., "Systematic event generator tuning for the LHC", *Eur. Phys. J. C* **65** (2010) 331, doi:10.1140/epjc/s10052-009-1196-7, arXiv:0907.2973.

[23] A. Buckley et al., "Rivet user manual", *Comput. Phys. Commun.* **184** (2013) 2803, doi:10.1016/j.cpc.2013.05.021, arXiv:1003.0694.

[24] CMS Collaboration, "Shape, transverse size, and charged hadron multiplicity of jets in pp collisions at 7 TeV", *JHEP* **06** (2012) 160, doi:10.1007/JHEP06(2012)160, arXiv:1204.3170.

[25] ATLAS Collaboration, "Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector", *Eur. Phys. J. C* **71** (2011) 1795, doi:10.1140/epjc/s10052-011-1795-y, arXiv:1109.5816.

[26] ATLAS Collaboration, "Measurement of the charged-particle multiplicity inside jets from $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector", *Eur. Phys. J. C* **76** (2016) 322, doi:10.1140/epjc/s10052-016-4126-5, arXiv:1602.00988.

[27] CMS Collaboration, "Pseudorapidity distributions of charged particles in pp collisions at $\sqrt{s} = 7$ TeV with at least one central charged particles", CMS Physics Analysis Summary CMS-PAS-QCD-10-024, 2011.

[28] CMS Collaboration, "Measurement of the energy density as a function of pseudorapidity in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019) 391, doi:10.1140/epjc/s10052-019-6861-x, arXiv:1812.04095.

[29] CMS Collaboration, "Strange particle production in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV", *JHEP* **05** (2011) 064, doi:10.1007/JHEP05(2011)064, arXiv:1102.4282.

[30] C. Bierlich, G. Gustafson, and L. Lönnblad, "A shoving model for collectivity in hadronic collisions", LU-TP 16-64 and MCnet-16-48, 2016. arXiv:1612.05132.

[31] C. Bierlich, "Rope hadronization and strange particle production", *EPJ Web Conf.* **171** (2018) 14003, doi:10.1051/epjconf/201817114003, arXiv:1710.04464.

[32] CMS Collaboration, "Measurement of charged pion, kaon, and proton production in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **96** (2017) 112003, doi:10.1103/PhysRevD.96.112003, arXiv:1706.10194.

[33] L. Gladilin, "Charm hadron production fractions", technical report, 12, 1999. arXiv:hep-ex/9912064.

[34] DELPHI Collaboration, "Tuning and test of fragmentation models based on identified particles and precision event shape data", *Z. Phys. C* **73** (1996) 11, doi:10.1007/s002880050295.

[35] CMS Collaboration, "Measurement of the underlying event activity in inclusive Z boson production in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2018) 032, doi:10.1007/JHEP07(2018)032, arXiv:1711.04299.

[36] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[37] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[38] C. Bierlich, G. Gustafson, L. Lönnblad, and A. Tarasov, "Effects of overlapping strings in pp collisions", *JHEP* **03** (2015) 148, doi:10.1007/JHEP03(2015)148, arXiv:1412.6259.

[39] C. Bierlich and J. R. Christiansen, "Effects of color reconnection on hadron flavor observables", *Phys. Rev. D* **92** (2015) 094010, doi:10.1103/PhysRevD.92.094010, arXiv:1507.02091.

[40] CMS Collaboration, "Measurement of jet substructure observables in $\mathrm{t\bar{t}}$ events from proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **98** (2018) 092014, doi:10.1103/PhysRevD.98.092014, arXiv:1808.07340.

[41] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[42] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[43] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[44] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[45] ATLAS Collaboration, "Measurement of colour flow with the jet pull angle in $\mathrm{t\bar{t}}$ events using the ATLAS detector at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **750** (2015) 475, doi:10.1016/j.physletb.2015.09.051, arXiv:1506.05629.

[46] ATLAS Collaboration, "Measurement of colour flow using jet-pull observables in $\mathrm{t\bar{t}}$ events with the ATLAS experiment at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **78** (2018) 847, doi:10.1140/epjc/s10052-018-6290-2, arXiv:1805.02935.

[47] ATLAS, CDF, CMS, D0 Collaboration, "First combination of Tevatron and LHC measurements of the top-quark mass", atlas-conf-2014-008; cdf-note-11071; cms-pas-top-13-014; d0-note-6416; fermilab-tm-2582-e, 2014. arXiv:1403.4427.

[48] ATLAS Collaboration, "Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **75** (2015) 158, doi:10.1140/epjc/s10052-015-3373-1, arXiv:1409.0832.

[49] ATLAS Collaboration, "Measurement of the top quark mass in the $\mathrm{t\bar{t}} \to$ lepton+jets and $\mathrm{t\bar{t}} \to$ dilepton channels using $\sqrt{s} = 7$ TeV ATLAS data", *Eur. Phys. J. C* **75** (2015) 330, doi:10.1140/epjc/s10052-015-3544-0, arXiv:1503.05427.

[50] ATLAS Collaboration, "Measurement of the top quark mass in the $\mathrm{t\bar{t}} \to$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data", *Phys. Lett. B* **761** (2016) 350, doi:10.1016/j.physletb.2016.08.042, arXiv:1606.02179.

[51] ATLAS Collaboration, "Top-quark mass measurement in the all-hadronic $\mathrm{t\bar{t}}$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **09** (2017) 118, doi:10.1007/JHEP09(2017)118, arXiv:1702.07546.

[52] ATLAS Collaboration, "Measurement of the top quark mass in the $\mathrm{t\bar{t}} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results", *Eur. Phys. J. C* **79** (2019) 290, doi:10.1140/epjc/s10052-019-6757-9, arXiv:1810.01772.

[53] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV", *Phys. Rev. D* **93** (2016) 072004, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044.

[54] CMS Collaboration, "Measurement of the top-quark mass in $\mathrm{t\bar{t}}$ events with lepton+jets final states in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **12** (2012) 105, doi:10.1007/JHEP12(2012)105, arXiv:1209.2319.

[55] CMS Collaboration, "Measurement of the Top-Quark Mass in $\mathrm{t\bar{t}}$ Events with Dilepton Final States in pp Collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **72** (2012) 2202, doi:10.1140/epjc/s10052-012-2202-z, arXiv:1209.2393.

[56] CMS Collaboration, "Measurement of the top-quark mass in all-jets $\mathrm{t\bar{t}}$ events in pp collisions at $\sqrt{s}$=7 TeV", *Eur. Phys. J. C* **74** (2014) 2758, doi:10.1140/epjc/s10052-014-2758-x, arXiv:1307.4617.

[57] CMS Collaboration, "Measurement of the top quark mass using single top quark events in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **77** (2017) 354, doi:10.1140/epjc/s10052-017-4912-8, arXiv:1703.02530.

[58] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019) 368, doi:10.1140/epjc/s10052-019-6863-8, arXiv:1812.10505.

[59] CMS Collaboration, "Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13$ TeV and combination with the lepton+jets channel", *Eur. Phys. J. C* **79** (2019) 313, doi:10.1140/epjc/s10052-019-6788-2, arXiv:1812.10534.

[60] CMS Collaboration, "Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $\mathrm{t\bar{t}}$ at $\sqrt{s} = 8$ and 13 TeV", CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016.

A CR tunes with an LO PDF set

Table 5: List of input RIVET routines, distributions, x-axis ranges, R of the distributions in the fit, number of bins and the centre-of-mass energy used in the fits to derive the CP1-CR1 and CP1-CR2 tunes.

			CP1-CR1			CP1-CR2		
RIVET routine	$\sqrt{s}$ (TeV)	Distribution	Fit range (GeV)	N_{bins}	R	Fit range (GeV)	N_{bins}	R
CMS_2015_I1384119	13	N_{ch} vs η		20	1		20	1
CMS_2015_PAS_FSQ_15_007	13	TransMIN charged p_T^{sum}	3–36	15	1	4–36	13	0.20
		TransMAX charged p_T^{sum}	3–36	15	1	4–36	13	0.20
		TransMIN N_{ch}	3–36	15	1	4–36	13	0.20
		TransMAX N_{ch}	3–36	15	1	4–36	13	0.20
CMS_2012_PAS_FSQ_12_020	7	TransMAX N_{ch}	3–20	10	1	3–20	10	0.10
		TransMIN N_{ch}	3–20	10	1	3–20	10	0.10
		TransMAX charged p_T^{sum}	3–20	10	1	3–20	10	0.10
		TransMIN charged p_T^{sum}	3–20	10	1	3–20	10	0.10
CDF_2015_I1388868	2	TransMIN N_{ch}	2–15	11	1	2–15	11	0.10
		TransMAX N_{ch}	2–15	11	1	2–15	11	0.10
		TransMIN charged p_T^{sum}	2–15	11	1	2–15	11	0.10
		TransMAX charged p_T^{sum}	2–15	11	1	2–15	11	0.10

Table 6: List of input RIVET routines, distributions, x-axis ranges, R of the distributions in the fit, number of bins and the centre-of-mass energy used in the fits to derive the CP2-CR1 and CP2-CR2 tunes.

			CP2-CR1			CP2-CR2		
RIVET routine	$\sqrt{s}$ (TeV)	Distribution	Fit range (GeV)	N_{bins}	R	Fit range (GeV)	N_{bins}	R
CMS_2015_I1384119	13	N_{ch} vs η		20	0.03		20	0.05
CMS_2015_PAS_FSQ_15_007	13	TransMIN charged p_T^{sum}	5–24	8	1	5–24	8	1
		TransMAX charged p_T^{sum}	5–24	8	0.17	5–24	8	0.25
		TransMIN N_{ch}	5–24	8	1	5–24	8	1
		TransMAX N_{ch}	5–24	8	0.17	5–24	8	0.25
CMS_2012_PAS_FSQ_12_020	7	TransMAX N_{ch}	5–20	7	0.07	5–20	7	0.25
		TransMIN N_{ch}	5–20	7	1	5–20	7	1
		TransMAX charged p_T^{sum}	5–20	7	0.07	5–20	7	0.25
		TransMIN charged p_T^{sum}	5–20	7	1	5–20	7	1
CDF_2015_I1388868	2	TransMIN N_{ch}	2–15	11	0.03	2–15	11	0.05
		TransMAX N_{ch}	2–15	11	0.03	2–15	11	0.05
		TransMIN charged p_T^{sum}	2–15	11	0.03	2–15	11	0.05
		TransMAX charged p_T^{sum}	2–15	11	0.03	2–15	11	0.05

Table 7: The parameters obtained in the fits of the CP1-CR1 and CP1-CR2 tunes, compared to the ones of the tune CP1. The upper part of the table displays the fixed input parameters of the tune, while the lower part shows the fitted tune parameters. The number of degrees of freedom (N_{dof}) and the goodness of fit divided by the number of degrees of freedom are also shown.

PYTHIA8 Parameter	CP1 [18]	CP1-CR1	CP1-CR2
PDF set	NNPDF3.1 LO	NNPDF3.1 LO	NNPDF3.1 LO
$\alpha_{\mathrm{S}}(m_{\mathrm{Z}})$	0.130	0.130	0.130
`SpaceShower:rapidityOrder`	off	off	off
`MultipartonInteractions:EcmRef` [GeV]	7000	7000	7000
$\alpha_{\mathrm{S}}^{\mathrm{ISR}}(m_{\mathrm{Z}})$ value/order	0.1365/LO	0.1365/LO	0.1365/LO
$\alpha_{\mathrm{S}}^{\mathrm{FSR}}(m_{\mathrm{Z}})$ value/order	0.1365/LO	0.1365/LO	0.1365/LO
$\alpha_{\mathrm{S}}^{\mathrm{MPI}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
$\alpha_{\mathrm{S}}^{\mathrm{ME}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
`StringZ:aLund`	–	0.38	–
`StringZ:bLund`	–	0.64	–
`StringFlav:probQQtoQ`	–	0.078	–
`StringFlav:probStoUD`	–	0.2	–
`SigmaTotal:zeroAXB`	off	off	off
`BeamRemnants:remnantMode`	–	1	–
`MultipartonInteractions:bProfile`	2	2	2
`ColourReconnection:mode`	–	1	2
`MultipartonInteractions:pT0Ref` [GeV]	2.400	1.984	2.385
`MultipartonInteractions:ecmPow`	0.154	0.113	0.165
`MultipartonInteractions:coreRadius`	0.544	0.746	0.587
`MultipartonInteractions:coreFraction`	0.684	0.569	0.533
`ColourReconnection:range`	2.633	–	–
`ColourReconnection:junctionCorrection`	–	8.382	–
`ColourReconnection:timeDilationPar`	–	31.070	–
`ColourReconnection:m0`	–	1.845	–
`ColourReconnection:m2lambda`	–	–	2.769
`ColourReconnection:fracGluon`	–	–	0.979
N_{dof}	183	157	150
$\chi^{*2}/N_{\mathrm{dof}}$	0.89	0.73	0.20

Table 8: The parameters obtained in the fits of the CP2-CR1 and CP2-CR2 tunes, compared to the ones of the tune CP2. The upper part of the table displays the fixed input parameters of the tune, while the lower part shows the fitted tune parameters. The number of degrees of freedom (N_{dof}) and the goodness of fit divided by the number of degrees of freedom are also shown.

PYTHIA8 Parameter	CP2 [18]	CP2-CR1	CP2-CR2
PDF set	NNPDF3.1 LO	NNPDF3.1 LO	NNPDF3.1 LO
$\alpha_{\mathrm{S}}(m_{\mathrm{Z}})$	0.130	0.130	0.130
`SpaceShower:rapidityOrder`	off	off	off
`MultipartonInteractions:EcmRef` [GeV]	7000	7000	7000
$\alpha_{\mathrm{S}}^{\mathrm{ISR}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
$\alpha_{\mathrm{S}}^{\mathrm{FSR}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
$\alpha_{\mathrm{S}}^{\mathrm{MPI}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
$\alpha_{\mathrm{S}}^{\mathrm{ME}}(m_{\mathrm{Z}})$ value/order	0.130/LO	0.130/LO	0.130/LO
`StringZ:aLund`	–	0.38	–
`StringZ:bLund`	–	0.64	–
`StringFlav:probQQtoQ`	–	0.078	–
`StringFlav:probStoUD`	–	0.2	–
`SigmaTotal:zeroAXB`	off	off	off
`BeamRemnants:remnantMode`	–	1	–
`MultipartonInteractions:bProfile`	2	2	2
`ColourReconnection:mode`	–	1	2
`MultipartonInteractions:pT0Ref` [GeV]	2.306	2.154	2.287
`MultipartonInteractions:ecmPow`	0.139	0.119	0.146
`MultipartonInteractions:coreRadius`	0.376	0.538	0.514
`MultipartonInteractions:coreFraction`	0.327	0.599	0.525
`ColourReconnection:range`	2.323	–	–
`ColourReconnection:junctionCorrection`	–	0.761	–
`ColourReconnection:timeDilationPar`	–	13.080	–
`ColourReconnection:m0`	–	1.546	–
`ColourReconnection:m2lambda`	–	–	6.186
`ColourReconnection:fracGluon`	–	–	0.978
N_{dof}	183	117	118
$\chi^{*2}/N_{\mathrm{dof}}$	0.54	0.21	0.22

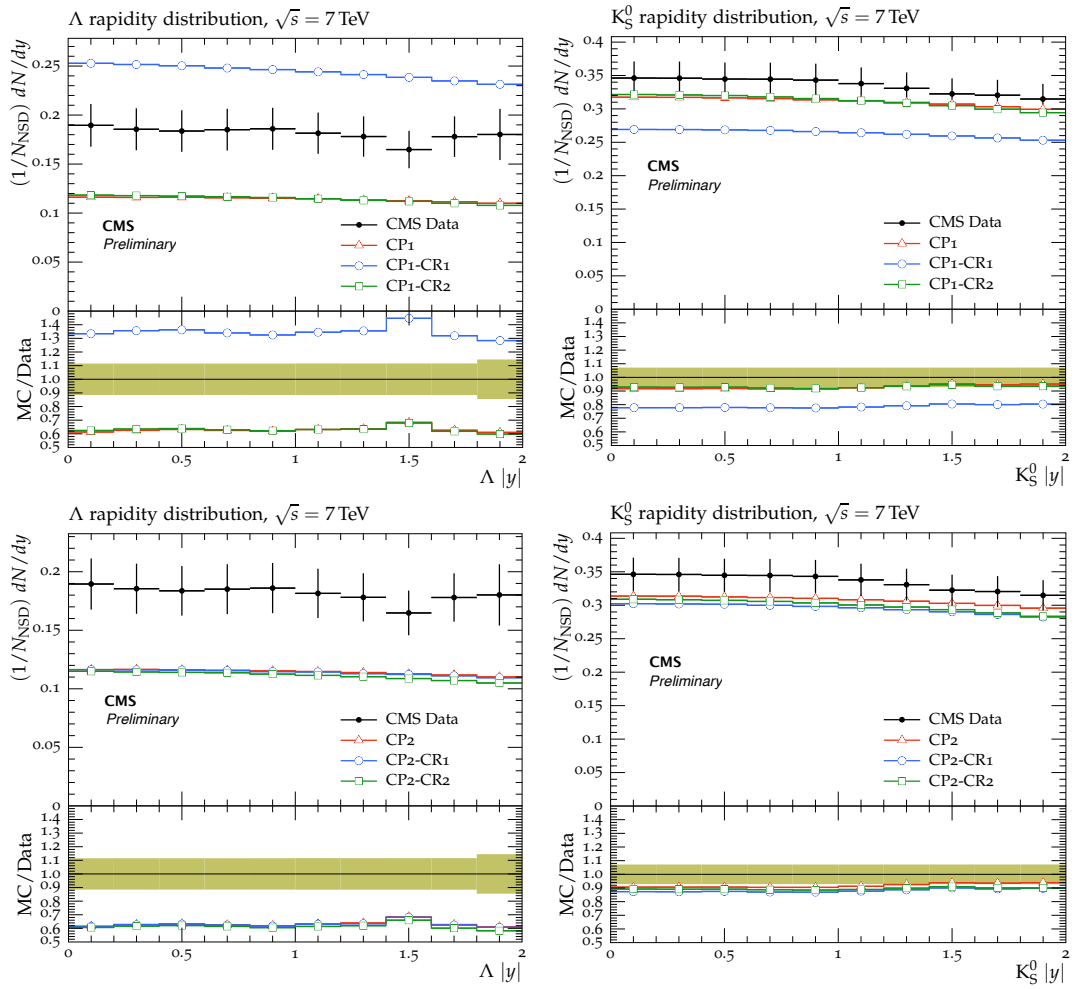

Figure 20: The strange particle production, Λ baryons (left) and K^0_s mesons (right), as a function of rapidity, measured by the CMS experiment at $\sqrt{s} = 7$ TeV [29]. The predictions of the CP1 and CP1-CR tunes (upper) and CP2 and CP2-CR tunes (lower) are compared to data. The coloured band represents the total experimental uncertainty in the data.

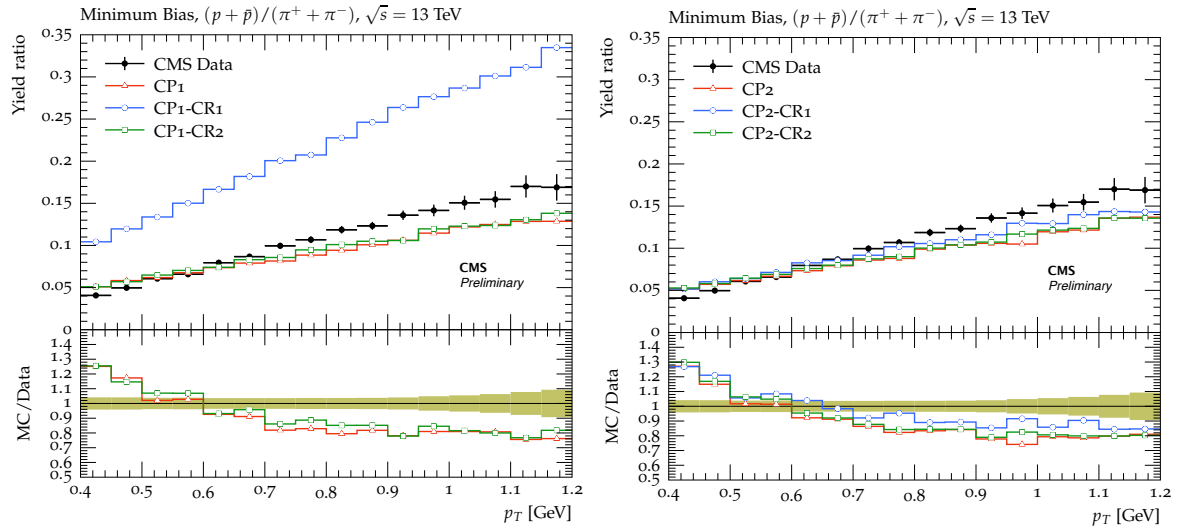

Figure 21: Ratios of particle yields, p/π, as a function of transverse momentum in minimum bias events, measured by the CMS experiment at $\sqrt{s} = 13$ TeV [32]. The predictions of the CP1 and CP1-CR tunes (left) and CP2 and CP2-CR tunes (right) are compared to data. The coloured band represents the total experimental uncertainty in the data.

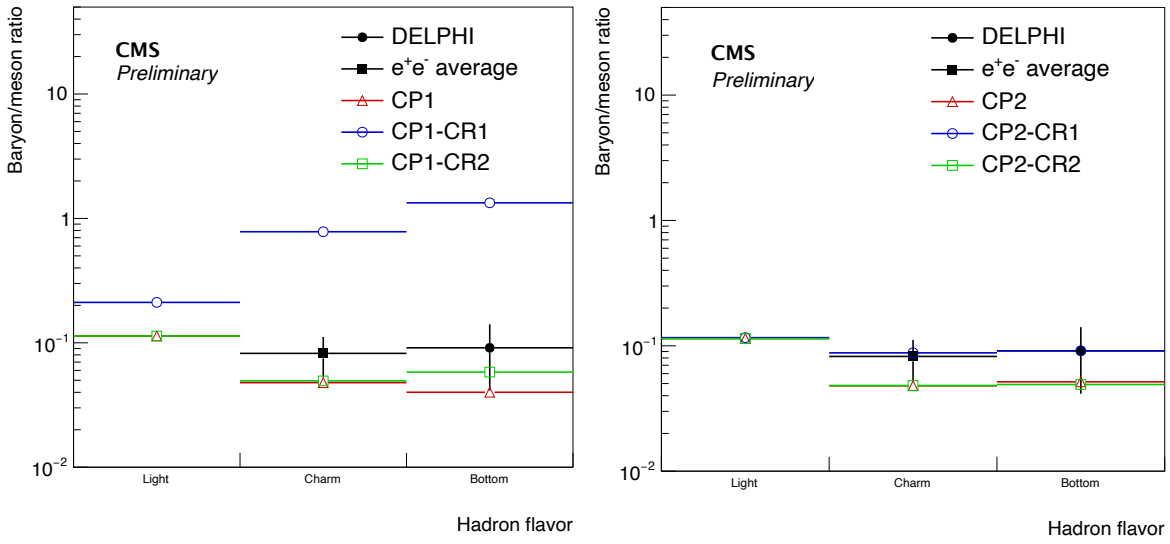

Figure 22: Ratios of particle yields for light, charm, and bottom hadrons predicted by the different PYTHIA8 tunes compared to data. The data are compared to predictions from the CP1 and CP1-CR tunes (left) and CP2 and CP2-CR tunes (right).