THE CHARM-QUARK HADRONISATION WITH PYTHIA/ANGANTYR MODEL*

Harsh Shah

Institute of Nuclear Physics Polish Academy of Sciences
Radzikowskiego 152, 31-342 Kraków, Poland
harsh.shah@ifj.edu.pl

Received 10 April 2026, accepted 20 July 2026, published online 4 September 2026

An overview of charm-quark hadronisation within the PYTHIA/Angantyr model is provided. The improvements to the junction formation and fragmentation mechanisms in the context of charm-quark hadronisation are discussed. We discuss these changes in the context of the recently developed global colour reconnection mechanism, which is based on the QCD-inspired colour reconnection model. We show the enhancement in Λ_c baryon yield in pp and pPb collisions at the LHC at $\sqrt{s_{NN}} = 5.02$ TeV collision energy using the PYTHIA/Angantyr model.

DOI:10.5506/APhysPolBSupp.19.5-A6

1. Introduction

Heavy-flavour hadrons (containing a charm or bottom quark) are unique probes to study hadronisation mechanisms in high-energy particle collisions. Their heavy valence quarks, the charm (c) and the bottom (b), can only be produced via perturbative processes such as hard scattering and parton showers. However, the hadronisation of the produced charm and bottom quarks is understood as a non-perturbative process, similar to the hadronisation of light quarks (u, d, and s quarks).

The factorisation approach and Monte Carlo models are used to compare experimental data with theory for heavy-hadron yields in particle collisions. The factorisation approach relies on the three factors: parton distribution functions (PDF) of the beams, parton-level hard scattering cross section, and fragmentation function fitted to e^+e^- and ep data. They give differential probabilities for the charm quarks to produce different charm hadrons. The SPS results [1] show that the predictions from independent fragmentation do

* Presented at the 32$^{\text{nd}}$ Cracow Epiphany Conference on the Recent Results from Heavy Ion Physics, Kraków, Poland, 12–16 January, 2026.

not hold. However, it has been assumed that universal fragmentation can be applied when studying inclusive quantities such as the total charm hadron yield. Moreover, recent work done by the ALICE Collaboration [2–5] shows that the fragmentation function fitted to e^+e^- and ep cannot describe the charm-hadron yield in pp collisions.

Initial states are still defined by PDFs in the PYTHIA 8 [6] Monte Carlo model. The later stages perform hadron production through the perturbative process of heavy-quarks [7, 8] and non-perturbative modelling of hadronisation [9]. Recently, the so-called QCD colour reconnection (CR) model [10] in PYTHIA 8 has been developed. The model has gained a lot of attention due to its ability to correctly reproduce the Λ_c^+ yield and Λ_c^+/D^0 ratio as a function of $p_\perp$ in pp collisions at various collision energies at the LHC [2, 4, 5]. However, it undershoots data for the production rates of Ξ_c and Ω_c baryons [11–13]. Following its success and ability to enhance baryon production, the model has been extended to heavy-ion collisions by introducing spatial constraints on colour dipoles to be colour-reconnected in the Angantyr model [14, 15].

In this paper, I gave an overview of the work done in [16]. In the next section, I will describe the process of charm-quark hadronisation in PYTHIA 8. I will also discuss the changes we have introduced in [16] to improve various charm-hadron yields in pp collisions. In Section 3, I will show the results of Λ_c^+ baryons in pPb collisions using the same model parameters which are used for pp collisions.

2. The charm-quark hadronisation

The heavy-quark production in pp collisions in PYTHIA 8 is constrained to partonic scattering and parton showers due to their heavy mass. There are *three* processes to produce a heavy quark in PYTHIA.

(1) The leading order (LO) scattering of the quarks and gluons producing a quark–antiquark pair of heavy quarks.

(2) The decay of gauge bosons (Z, $W^\pm$, and Higgs bosons decay) and of heavier quarks (top- and bottom-quark decays).

(3) The parton showers produce heavy quarks by gluon splitting, flavour excitation, and pair creation.

The default masses of charm and bottom quarks, 1.5 GeV and 4.8 GeV, respectively, used in PYTHIA 8 are effective nominal masses, which are usually fixed with careful considerations and fit to experimental data [6, 7]. The quark mass is always a part of the matrix elements used for perturbative calculations of the above-mentioned three processes, parton showers, and the quark production cross sections in PYTHIA 8. Following the latest work arguing in favour of reducing the charm-quark mass for a better fit to the LHC experiment data in [17], in this work, we decided to reduce the charm- and bottom-quark masses to 1.3 GeV and 4.2 GeV, respectively.

2.1. Colour reconnection

The concept of colour reconnections (CR) [18] in PYTHIA 8 has been introduced to reproduce the observed multiplicity of charged particles and the increase in $\langle p_\perp \rangle$ as a function of the $(N_{\rm ch})$ distribution as observed in experiments. The colour dipoles are produced as a result of the perturbative interactions of the partons and parton showers in PYTHIA 8. The CR rearranges the colour connections between the produced colour dipoles. The CR mechanism reduces the net-string length, which is calculated as λ-measure (it is an indirect representation of the rapidity span of the colour dipole, which can give an approximate number of the produced hadrons) at the colour dipole level. Recently, an $\mathbf{SU(3)}$ colour algebra-based QCD-inspired CR (QCDCR) model [10] has been developed. The QCDCR model acts on the $N_c = 3$ limit, unlike the default CR model, which acts on the $N_c \to \infty$ limit.

2.2. Junction hadronisation

The colour algebra enables the formation of a junction-like topology during two- and three-dipole colour rearrangements in the QCDCR model. These junctions are always produced as junction–anti-junction pairs. Junction points always produce a baryon during hadronisation. Hence, the junction pairs maintain the baryon number conservation. These junction baryons are additional baryons to Pythia's default baryon production mechanisms, diquark production [19], and the "popcorn" mechanism [20] of string fragmentation.

The string fragmentation relies on the quark-pair production rate, which is a mass-dependent quantity. A $q\bar{q}$-pair production rate is given by

$$\mathrm{d}P \simeq \mathrm{d}^2 p_\perp \exp\left(-\pi m_\perp^2/\kappa\right) , \tag{1}$$

where κ is string tension, and $m_\perp$ is transverse mass of the quark with mass m_q and transverse momenta $p_\perp$. It gives an extremely low probability for the production of heavy-quark pairs (*e.g.* "charm" pairs) during the string fragmentation. Therefore, one can safely assume that all of the heavy quarks are produced only in hard scatterings and parton showers in PYTHIA 8.

All heavy quarks (and light quarks) are at the string ends at the time of hadronisation. In addition to the string fragmentation parameters, the formation of a heavy meson or baryon also depends on the CR mechanism, which rearranges the colour connections of the heavy quarks before the hadronisation stage in PYTHIA 8.

The junctions formed due to the QCDCR model influence the heavy-baryon production if the heavy quarks are directly connected to the junction point. The default calculation for the λ-measure uses the energy of the quark in the rest frame of the junction. For the heavy quarks, this assumption gives a larger value because of the invariant mass of the heavy quarks. Hence, we decided to use the rapidity of the heavy quark in the junction rest frame to calculate the λ-measure, which not only gives a lower value, but also a better assumption for the available disposable energy in the dipole. This new method increases the probability of a heavy quark being directly connected to the junction point after CR.

Before this work, the junction fragmentation was either performed in the junction rest frame[1] or in the centre-of-mass frame of the collision. The quarks were treated as massless in this skim. We have realised that if a heavy quark is directly connected to the junction point, the junction system often fails to obtain a junction rest frame. We modified the junction fragmentation algorithm in [16]. We introduced a condition to junction fragmentation, so that if heavy quarks are directly connected to the junction point, the junction fragmentation will be performed in the heaviest quark's rest frame. The highest energy leg fragments last in the junction hadronisation, which is often the leg connected to the heavy quark due to its mass. We also modified the algorithm for the energy calculations of the junction leg [16], which is used to decide the sequence of junction legs to be fragmented. Our modifications use the three momenta of the partons, instead of the energy, to calculate the fragmentation sequence.

During the final stage of junction fragmentation, a diquark is formed by combining the two quarks from the low- and middle-energy legs. The spin assignment to the diquark is done by a set of parameters[2]. The *ad hoc* values of the parameters in PYTHIA 8 are 0.5, 0.7, 0.9, and 1, for the cases where the heaviest quark is u/d, s, c, and b, respectively. In the QCDCR model, these were instead all set to 0.0275, which is close to a well-constrained value used for diquark–anti-diquark breakups in a normal string. However, the formation of diquarks in the joining of junction legs is different from the diquark-pair production during the string fragmentation. Moreover, we know that the QCDCR model is unable to reproduce the experimentally observed yields of heavier baryon states. We have tested the effect of raising the values to the default ones in PYTHIA 8 when using QCDCR in this work.

Our modifications increase the probability of the heavy-quarks forming a baryon over a meson during the junction fragmentation. Since the "Rope hadronisation" mechanism [21] did not fully comply with the complex junc-

[1] The junction rest frame defines when all three junction legs in the momenta frame have 120° between them.

[2] The PYTHIA 8 parameter `StringFlav:probQQ1toQQ0join` is used to set these values.

tion systems, we increased the strange quarks production probability from 0.217 to 0.4 [3]. We assume that the highest amount of heavy-quark production would be in the most central collisions. Hence, we consider that the increased probability of strange-quark production in string fragmentation is a reasonable alternative to the Rope hadronisation.

3. Results

The Angantyr model has been developed as an extension of PYTHIA 8 to simulate heavy-ion collisions as a sophisticated superposition of pp-like collisions. Therefore, underlying physics improvements in PYTHIA 8 for pp collisions should also reflect in the results of heavy-ion collisions. Hence, we decided to test the effects of our changes in PYTHIA 8 for pp collisions simultaneously with pPb collisions (a heavy-ion collision system) and without any additional tuning of the Angantyr model parameters.

In this section, we show the effects of the changes described in the previous section. We focus on Λ_c^+ baryons and Λ_c^+/D^0 ratio for pp and pPb collisions at $\sqrt{s} = 5.02$ TeV collision energy. For more results at other collision energies and results related to hyperon productions (*e.g.* Σ, Σ_c, Ξ, and Ξ_c) in pp collisions, refer to [16].

In figure 1, we show the $p_\perp$ distribution of the prompt Λ_c^+ baryons and the Λ_c^+/D^0 ratio is compared with $\sqrt{s} = 5.02$ TeV pp collisions ALICE data [2]. We see that the QCDCR significantly improves the Pythia (def) yield for the Λ_c^+ cross section and Λ_c^+/D^0 ratio distributions. We can also observe that even after all the new changes introduced in this work, we are still able to maintain a good description of the Λ_c^+ baryon production.

Figure 2 shows pPb results from the ALICE experiment at $\sqrt{s_{NN}} = 5.02$ TeV for the Λ_c^+ cross section and the Λ_c^+/D^0 ratio, in comparison with the Angantyr model. The default Angantyr (red lines) severely underestimated the Λ_c^+ yield, while the global CR significantly improves the description of data, although the Λ_c cross section is somewhat underestimated. Adding the changes introduced in [16] does not influence the result. This was to be expected, since in pp collisions, the effect on Λ_c^+ was also minor.

We observe that for low $p_\perp$, the model fails to reproduce the behaviour of the data. This is clearly visible in the ratio to D^0, where our model shows no sign of reduction at small $p_\perp$.

[3] The PYTHIA 8 parameter `StringFlav:probStoUD` is used for this change.

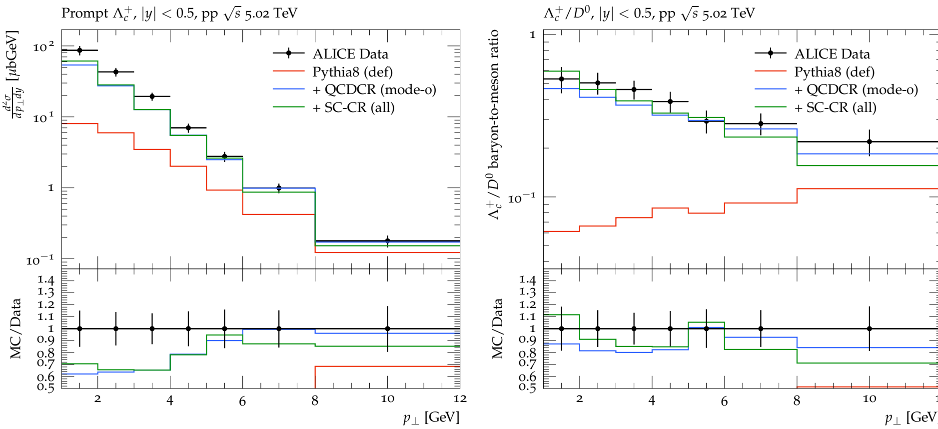

Fig. 1. Left: Prompt Λ_c^+ distribution as a function of $p_\perp$. Right: The Λ_c^+/D^0 ratio as a function of $p_\perp$. The pp collision at $\sqrt{s} = 5.02$ TeV data from the ALICE experiment [2]. The experimental results are compared with PYTHIA 8 default, QCDCR (mode-0), and spatially constrained QCDCR with all the new changes from this work shown as red, blue, and green lines, respectively. From [16].

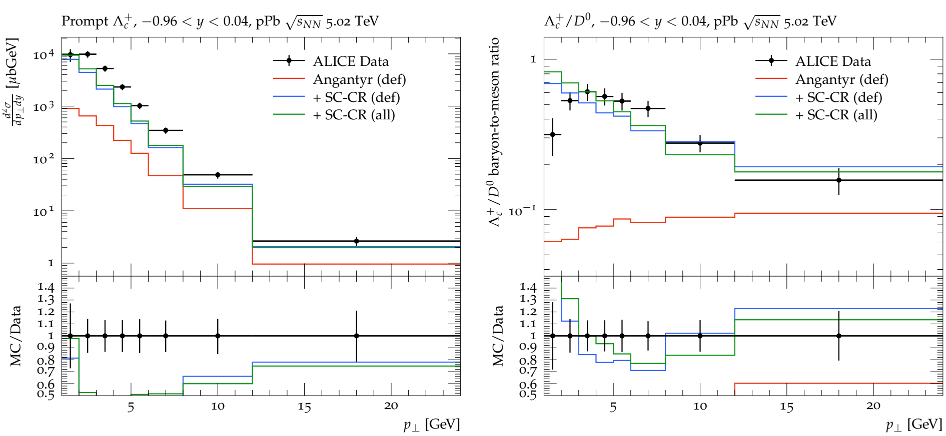

Fig. 2. Left: Prompt Λ_c^+ distribution as a function of $p_\perp$. Right: The Λ_c^+/D^0 ratio as a function of $p_\perp$. The red, blue, and green lines are Angantyr default as Angantyr (def), Angantyr with a new tune with global CR as SC-CR (def), and the changes we have made in this work in the global CR as SC-CR (all), respectively. The pPb collision at $\sqrt{s_{NN}} = 5.02$ TeV data is from the ALICE experiment [2]. (Note that the rapidity region, $-0.96 < y < 0.04$, is given in the collision rest frame.) From [16].

4. Conclusion

We show that applying the QCD-inspired colour reconnection model to heavy-ion collisions provides a better description of heavy-baryon yields. The production rates of Λ_c^+ are improved in pPb collisions using the global CR, which has previously worked well in pp collisions.

The QCDCR model allows for junction-like topologies, which contribute to the additional baryons in PYTHIA 8. For a heavy quark connected to a junction, the λ-measure used in the QCDCR model should be improved. Moreover, during the fragmentation of the junction system, the convention is to calculate the energy in all three legs and fragment from the lowest-energy leg. In this work, we improved these limitations. We introduced a fragmentation of the junction system in the rest frame of the heaviest quark. We introduced the use of momenta of the partons instead of the energy to avoid counting the invariant mass of the quarks as the potential energy available in the junction leg.

We notice that apart from the above-mentioned modifications, we need strange quarks, as many of the heavy baryons contain strange quarks. Hence, we have increased the strange-quark production probability in string fragmentation as a temporary alternative to Rope hadronisation. The charm and bottom quark nominal masses are also reduced in this work.

All these changes improve the PYTHIA 8 description for Ξ, Σ, Ξ_c, and Σ_c baryon production rates [16]. We also managed to keep a reasonable description of the Λ_c^+ and Λ_c^+/D^0 in pp collisions. For the first time, we show Λ_c^+ and Λ_c^+/D^0 results in pPb collisions [16], where we use the same parameters as we have used in pp collisions. The results are generated with the global CR and with the changes introduced in this work, and they show a visible improvement over the default Angantyr setup.

At this stage, it is important to be aware that a lot of work is needed to better understand the heavy quark's hadronisation. We encourage our experimentalist colleagues to test the model against other colliding systems.

The speaker is thankful to the Epiphany conference organisers for providing an opportunity to talk about this novel work at the conference.

REFERENCES

[1] WA82 Collaboration (M. Adamovich *et al.*), *Phys. Lett. B* **305**, 402 (1993).

[2] ALICE Collaboration (S. Acharya *et al.*), *Phys. Rev. Lett.* **127**, 202301 (2021).

[3] ALICE Collaboration (S. Acharya *et al.*), *Phys. Rev. D* **105**, L011103 (2022).

[4] ALICE Collaboration (S. Acharya *et al.*), *Phys. Rev. Lett.* **128**, 012001 (2022).

[5] ALICE Collaboration (S. Acharya *et al.*), *Phys. Rev. D* **109**, 072005 (2024), arXiv:2301.13798 [nucl-ex].

[6] C. Bierlich *et al.*, *SciPost Phys. Codeb.* **2022**, 8 (2022), arXiv:2203.11601 [hep-ph].

[7] E. Norrbin, T. Sjostrand, *Phys. Lett. B* **442**, 407 (1998).

[8] E. Norrbin, T. Sjostrand, *Eur. Phys. J. C* **17**, 137 (2000).

[9] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, *Phys. Rep.* **97**, 31 (1983).

[10] J.R. Christiansen, P.Z. Skands, *J. High Energy Phys.* **2015**, 003 (2015).

[11] Large Ion Collider Experiment Collaboration (S. Acharya *et al.*), *Phys. Rev. Lett.* **127**, 272001 (2021).

[12] ALICE Collaboration (S. Acharya *et al.*), *J. High Energy Phys.* **2021**, 159 (2021).

[13] ALICE Collaboration, *Phys. Lett. B* **846**, 137625 (2023), arXiv:2205.13993 [nucl-ex].

[14] C. Bierlich, G. Gustafson, L. Lönnblad, H. Shah, *J. High Energy Phys.* **2018**, 134 (2018).

[15] L. Lönnblad, H. Shah, *Eur. Phys. J. C* **83**, 575 (2023).

[16] C. Bierlich, G. Gustafson, L. Lönnblad, H. Shah, *Eur. Phys. J. C* **84**, 231 (2024).

[17] C. Bierlich *et al.*, *Eur. Phys. J. Plus* **139**, 593 (2024).

[18] T. Sjöstrand, M. van Zijl, *Phys. Rev. D* **36**, 2019 (1987).

[19] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, *Z. Phys. C* **13**, 361 (1982).

[20] B. Andersson, G. Gustafson, T. Sjöstrand, *Phys. Scr.* **32**, 574 (1985).

[21] C. Bierlich, G. Gustafson, L. Lönnblad, A. Tarasov, *J. High Energy Phys.* **2015**, 148 (2015).