EPJ Web of Conferences **291**, 01010 (2024) https://doi.org/10.1051/epjconf/202429101010
MESON 2023

Overview of hadronization of quarks in proton-proton and e^+e^- collisions

Mattia Faggin[1,2,*]

[1]University of Trieste, Italy
[2]INFN Trieste, Italy

Abstract. The so-called hadronization is a non-perturbative QCD phenomenon corresponding to the formation of colourless hadrons from coloured quark constituents. The hadron formation in point-like e^+e^- collisions can be described by string models. According to the Lund model, the $q\bar{q}$ pair production in the scattering is followed by a shower of light partons produced via multiple colour-string breaking, which produce colour singlets in the final states. The probability to obtain a hadron of a given species carrying a certain momentum fraction of the original quark is quantified by the fragmentation functions. They are assumed universal and usually constrained from e^+e^- and e^-p collisions, and they successfully describe the production of mesons in e^+e^- and pp collisions at the colliders. However, recent measurements of pp collision data from the LHC showed a surprising relative enhancement of baryon production compared to mesons, and model predictions based on string fragmentation do not describe the data. In this talk, an overview of the most recent experimental results of hadron production cross section in pp collisions at the LHC compared with e^+e^- results will be provided. A comparison with novel theoretical models implementing hadronization mechanisms different from the Lund string fragmentation will be also discussed, as well as experimental results obtained in larger systems, like p–Pb and Pb–Pb collisions.

1 Introduction

Hadronization is the mechanisms by which quarks and gluons produced in parton scatterings form hadrons. There are no first-principle descriptions of hadron formation, given the non-perturbative nature of the phenomenon. This prevents any perturbative QCD (pQCD) calculation. For this reason, to describe the hadronization it is necessary to resort to models and make use of phenomenological parameters.

Depending on the collision system, hadronization takes place via different mechanisms. The standard description of hadron formation in e^+e^- and pp collisions is based on string models, while in heavy-ion collisions hadrons can be formed via parton coalescence in the quark-gluon plasma (QGP).

The standard description of hadron production in e^+e^- collisions at large Q^2 is based on a factorization approach, according to which the transverse-momentum (p_T) production of a

---

[*] e-mail: mfaggin@cern.ch

hadron of species h can be written as

$$\frac{\mathrm{d}\sigma^{\mathrm{e^+e^-}\to h}}{\mathrm{d}p_{\mathrm{T}}^{h}}(p_{\mathrm{T}};\mu_{\mathrm{F}},\mu_{\mathrm{R}}) = \frac{\mathrm{d}\sigma^{\mathrm{q}}}{\mathrm{q}p_{\mathrm{T}}^{q}}(p_1,p_2;\mu_{\mathrm{F}},\mu_{\mathrm{R}}) \otimes D_{\mathrm{q}\to h}(z = p_h/p_{\mathrm{q}};\mu_{\mathrm{F}})\,, \tag{1}$$

where:

- μ_{F} and μ_{R} indicate the factorization and renormalization scales, respectively;
- $\mathrm{d}\sigma^{q}/\mathrm{d}p_{\mathrm{T}}^{q}(p_1,p_2;\mu_{\mathrm{F}},\mu_{\mathrm{R}})$ is the production cross section of the $\mathrm{q\bar{q}}$ pair in a $\mathrm{e^+e^-}$ scattering, where p_1 and p_2 denote the momentum of the leptons in the initial state;
- $D_{\mathrm{q}\to h}(z = p_h/p_{\mathrm{q}};\mu_{\mathrm{F}})$ is the fragmentation function, which quantifies the probability of a quark q to produce a hadron h carrying a fraction z of the original quark momentum.

The fragmentation functions encode the nature of hadronization. This is usually described by string models, which are interfaced with the parton shower produced in event generators. For example, in PYTHIA generator a colour-flux tube, called *string*, is created between the $\mathrm{q\bar{q}}$ pair produced in the hard scattering. In this models, gluons represent kinks along the string, and the hadron formation is ruled by the Lund fragmentation [1, 2]. This develops via multiple string breaks into pairs of (di)quark-anti(di)quark pairs, following the tunnelling probability

$$P \propto \exp\left(-\frac{\pi m_{\mathrm{T,q}}^2}{\kappa}\right) = \exp\left(-\frac{\pi m_{\mathrm{q}}^2}{\kappa}\right)\exp\left(-\frac{\pi p_{\mathrm{T,q}}^2}{\kappa}\right), \tag{2}$$

where κ indicates the string tension, m_{q} the quark mass and $p_{\mathrm{T,q}}$ the quark transverse momentum. Given the exponential suppression with the mass, the formation of $\mathrm{c\bar{c}}$ pairs is significantly suppressed compared with that of lighter $\mathrm{q\bar{q}}$ ones: $\mathrm{u:d:s:c} \simeq 1:1:1/3:10^{-11}$. In the Lund model, the string splitting into diquark-antidiquark pairs rules the baryon formation.

Several calculations employ the factorization approach also to describe the hadron formation in pp collisions, where the formula in Eq. 1 must be changed into

$$\frac{\mathrm{d}\sigma^{\mathrm{pp}\to h}}{\mathrm{d}p_{\mathrm{T}}^{h}}(p_{\mathrm{T}};\mu_{\mathrm{F}},\mu_{\mathrm{R}}) = \mathrm{PDF}(\mathrm{x}_1,\mu_{\mathrm{F}})\cdot\mathrm{PDF}(\mathrm{x}_2,\mu_{\mathrm{F}})\otimes\frac{\mathrm{d}\sigma^{\mathrm{q}}}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{q}}}(x_1,x_2;\mu_{\mathrm{F}},\mu_{\mathrm{R}})\otimes D_{\mathrm{q}\to h}(z = p_h/p_{\mathrm{q}};\mu_{\mathrm{F}})\,. \tag{3}$$

The terms $\mathrm{PDF}(x_{1,2},\mu_{\mathrm{F}})$ indicate the parton distribution functions, which quantify the probability to pick-up two partons of given species from the protons in the initial state, carrying a fraction x of the original proton momentum. The main question to be addressed is whether pp collisions can be considered as a pure superimposition of several point-like collisions among partons, or whether the hadronization is changed by the presence of a colour-rich environment. Model calculations based on factorization usually employ fragmentation functions constrained from $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collision, assuming them as universal among collision systems. Such models describe the production of charm hadrons in $\mathrm{e^+e^-}$collisions [4], and they successfully predict the charm meson-to-meson ratios in pp collisions [5, 6]. In such collision systems, the measurements are in agreement with Pythia 8 simulations with Monash. In this tune, the default one in Pythia 8 [7], the string-length minimization is reached also allowing the colour reconnection between partons from different multi-parton interactions (MPIs). The agreement between the experimental results and the model predictions described above support the assumption of independent fragmentation among collision systems.

2 New phenomena in pp collisions at the LHC

Recent experimental results from LHC indicate that pp collisions might be significantly different from $\mathrm{e^+e^-}$ ones.

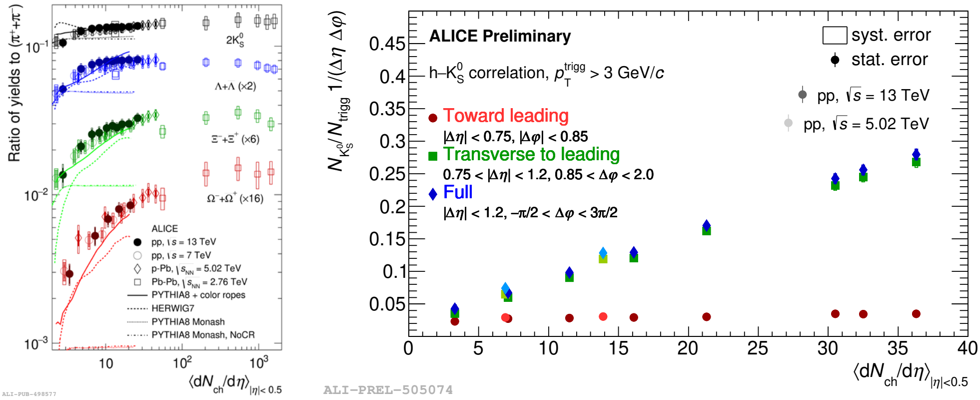

Figure 1. Left: ratios of strange-hadron production yield with that of charged pions in pp, p–Pb and Pb–Pb collisions at the LHC as a function of event multiplicity [8]. Right: K^0_S production, separated in the components in- (toward leading) and out-of- (transverse to leading) jet, as a function of event multiplicity at mid-rapidity in pp collisions at $\sqrt{s}$ = 13 TeV.

2.1 Strange-hadron production in pp collisions

A phenomenon that historically was considered a clear signature of the QGP formation in ultra-relativistic heavy-ion collisions is the strangeness enhancement. This corresponds to the increase of the yield-ratio between strange hadrons and charged pions in heavy-ion collisions compared to minimum-bias pp collisions, due to the enhanced formation of $s\bar{s}$ pairs in the QGP. In the left panel of Fig. 1 the ratios of strange-hadron production yields with that of charged pions in pp, p–Pb and Pb–Pb collisions are shown as a function of event multiplicity at mid-rapidity. These results show that the strangeness enhancement increases with the valence strange-quark content of the produced strange hadron. More interestingly, this comparison shows a sort of continuity among collision systems: if compared as a function of event multiplicity at mid-rapidity, the yield-ratio of strange hadrons to charged pions smoothly increases from pp to Pb–Pb collisions. This might suggest similarities among collision systems, when similar event multiplicities are taken into account.

In the right panel of Fig. 1 the production yield of K^0_S as a function of event multiplicity at mid-rapidity in pp collisions at $\sqrt{s}$ = 13 TeV is reported. The production of K^0_S mesons linearly increases as a function of the event multiplicity, and the experimental results show that the overall production of strange hadrons is dominant outside jets. This result indicates that the underlying event surrounding the partonic hard scattering processes play a crucial role in the production of strange hadrons in pp collisions at the LHC.

2.2 Baryon-to-meson yield-ratios in pp collisions

Recent results on the meson and baryon production at the LHC show significant differences compared to e^+e^- collisions. The prompt Λ_c^+/D^0 baryon-to-meson ratio measured in pp collisions at $\sqrt{s}$ = 5.02 TeV [11, 12] shows a clear dependence as a function of p_T, with a maximum vale at $p_T \sim 2$ GeV/c, and it is significantly larger than the ratio $\Lambda_c^+/D^0 \sim 0.113$ [13] measured in e^+e^- collisions at LEP. The predictions with Pythia Monash tune, and pQCD-based calculations based of factorization and employing fragmentation functions constrained from e^+e^- underestimate the measurement in pp collisions by about a factor of 5 at low p_T. A

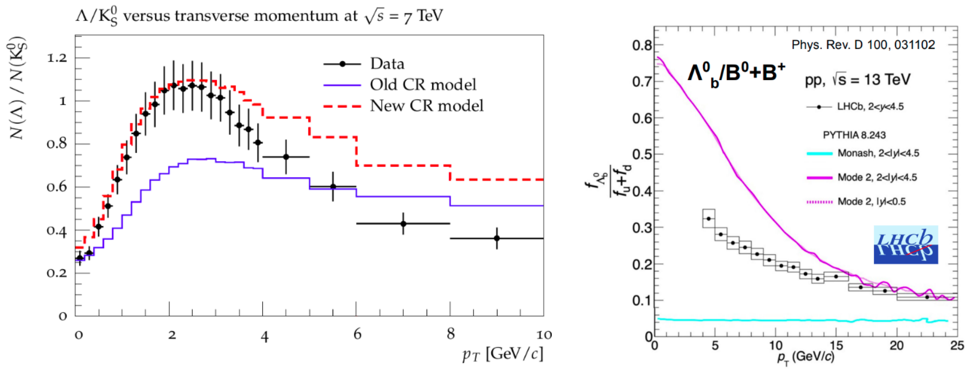

Figure 2. Left: Λ/K^0_S ratio at midrapidity in pp collisions at $\sqrt{s}$ = 7 TeV compared with Pythia predictions [9]. Right: $\Lambda^0_b/(B^0+B^+)$ ratio at forward rapidity in pp collisions at $\sqrt{s}$ = 13 TeV [10].

qualitatively similar p_T dependence is observed for the Λ/K^0_S ratio measured in pp collisions in $\sqrt{s}$ = 7 TeV (Fig. 2 left). Also in this case, the predictions with Pythia Monash tune (blue) underestimates the measured ratio at low p_T. The measurement is better described by Pythia predictions employing a new model (red), which enables the colour reconnection beyond the leading colour topologies (CR-BLC). In this model, a relative baryon enhancement is induced by the formation of junction topologies, and this allows Pythia predictions to quantitatively describe both the strange and charm baryon-to-meson ratios at mid-rapidity in pp collisions at the LHC. Also the measured $\Lambda^0_b/(B^0+B^+)$ baryon-to-meson in pp collisions is significantly underestimated by the predictions obtained with the Pythia Monash tune, as shown in the right panel of Fig. 2. Interestingly, even employing the new colour-reconnection modes the Pythia predictions do not describe correctly the baryon-to-meson ratio in the beauty sector at forward rapidity ($2 < y < 4.5$).

3 The charm-baryon enhancement in pp collisions at the LHC

The recent findings discussed in the previous section question the fact that pp collisions can be considered as a superimposition of several e^+e^--like collisions, as well as the validity of the independent fragmentation assumption. Several model calculations were proposed in the last decade to explain the enhanced baryon-to-meson ratio observed in the charm sector. The statistical hadronisation model with relativistic quark model (SHM+RQM) [14] foresees the presence of a large set of mostly unobserved excited charm baryons, which decay strongly and enrich the abundance of ground-state charm baryons. The abundances of such excited states, predicted by the RQM model [15], are assumed to follow thermal densities ruled only by the mass and spin degeneracy. In the Catania model [16] the charm quark can hadronise either via fragmentation, or by recombining with surrounding light quarks already produced from the event underlying the hard scattering. In the quark (re)combination model (QCM) [17] the charm quarks produced in the hard scatterings hadronise by recombining with surrounding equal-velocity light quarks. Despite the different phenomenology employed in the calculations, the predictions from all the models discussed above agree within uncertainties with the Λ^+_c/D^0 yield-ratios measured in pp collisions at the LHC [18].

The comparison with new experimental results can help validating the mechanisms assumed by these models. The production cross sections in e^+e^- collisions of prompt Λ^+_c and

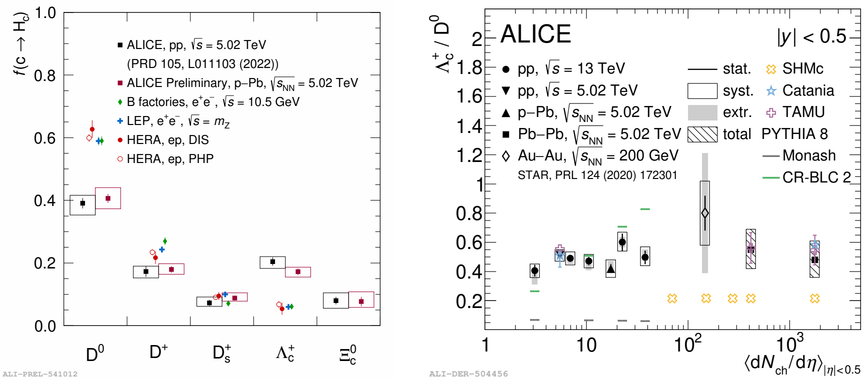

Figure 3. Left: charm fragmentation fractions in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV. Right: p_{T}-integrated Λ_c^+/D^0 ratio as a function of event multiplicity in pp, p–Pb Pb–Pb and Au-Au collisions.

$\Sigma_c^{0,+,++}$ states normalized by the total angular-momentum degeneracy (2J + 1) lie on two separate exponential functions [4]. A possible explanation for this behaviour is the larger mass of the constituent spin-1 $(\mathrm{ud})_1$ valence diquark in the Σ_c^+ with respect to the spin-0 $(\mathrm{ud})_0$ one in the Λ_c^+, which penalizes the string fragmentation into $\Sigma_c^{0,+,++}$ baryons. Such penalty factor is not assumed in the hadronization models in pp collisions discussed above. As for the Λ_c^+/D^0 ratio, the prompt $\Sigma_c^{0,+,++}/\mathrm{D}^0$ ratio measured in pp collisions at $\sqrt{s}$ = 13 TeV is significantly underestimated by the Pythia Monash predictions, and it is described within uncertainties by the model predictions discussed above [18].

Further useful insights derive from the measurement of the fraction of prompt Λ_c^+ deriving from the strong decays of $\Sigma_c^{0,+,++}$ baryons ($\Lambda_c^+ \leftarrow \Sigma_c^{0,+,++}$). This quantity in pp collisions was measured to be about 2 times larger than that in e^+e^- [18], and the p_{T}-differential yield-ratio $\Lambda_c^+ \leftarrow \Sigma_c^{0,+,++}/\Lambda_c^+$ is well described by the calculations provided by the QCM, Catania nad SH+RQM models. Interestingly, the predictions obtained with the CR-BLC modes of Pythia event generator overestimate the measurement. This suggests that a further tuning of the colour reconnection parameters in this model might be necessary, especially those ruling the charm-baryon formation by including charm diquarks. Novel measurements of excited Λ_c^+- and $\Sigma_c^{0,+,++}$- baryon production can shed light in this sense.

Several recent measurements of strange-charm baryon production undermine the description of charm-quark hadronization in the models discussed above. Differently from the Λ_c^+/D^0 and $\Sigma_c^{0,+,++}/\mathrm{D}^0$ ratios, the $\Xi^{0,+}/\mathrm{D}^0$ [19] baryon-to-meson ratios in pp collisions at $\sqrt{s}$ = 13 TeV are significantly underestimated by these predictions. Due to the similar underestimation of the $\Xi^{0,+}$- and $\Sigma_c^{0,+,++}$-baryon production, the $\Xi^{0,+}/\Sigma_c^{0,+,++}$ baryon-to-baryon ratio is compatible with the expectations form Pythia Monash. This can be a consequence of similar penalties for colour string to fragment into $\Sigma_c^{0,+,++}$ and $\Xi^{0,+}$ baryons, given the similar masses of the non-strange spin-1 $(\mathrm{uu})_1$, $(\mathrm{ud})_1$ and $(\mathrm{dd})_1$ valence diquarks in the former case, and the strange spin-0 $(\mathrm{us})_0$ diquarks in the latter one.

All the measurements discussed above highlight a different balance between meson and baryon production in the charm sector at the LHC. A summary of these findings is shown in the left panel of Fig. 3, where the charm-quark fragmentation fractions $f(\mathrm{c} \rightarrow \mathrm{H_c})$ are reported. Compared to what measured in e^+e^- collisions, a reduction D^0-meson production of about a factor 1.5 is accompanied by an increase of the Λ_c^+-baryon production of about a factor of 3 in both pp and p–Pb collisions. These results highlight once more the relative baryon enhancement in pp collisions, and do not support the assumption of independent fragmentation discussed above.

A possible explanation of the difference between results in e^+e^- and pp collisions can be related to the similarity of the latter ones with larger collision systems, as discussed in Sec. 2.1. To further investigate this, the prompt D_s/D^0 and Λ_c^+/D^0 yield-ratios have been studied in pp collisions as a function of event multiplicity [20]. If for the former case any significant dependence on neither p_T nor event multiplicity was observed, in the latter one the measurement at high multiplicities is larger than that at low multiplicities of about 5.3σ in the interval $1 \leq p_T < 24$ GeV/c. However, the p_T-integrated Λ_c^+/D^0 ratio does not show any significant multiplicity dependence, and the results in pp, p–Pb, Pb–Pb and Au-Au collisions agree within uncertainties. Model calculations are not able to catch either the magnitude or the multiplicity dependence of the experimental results. New measurements at even lower event-multiplicities in pp collisions might shed light into the possible connection of e^+e^- collisions with other systems.

4 Conclusions

Recent measurements of baryon production in pp collisions at the LHC showed significant difference with those obtained in e^+e^- collisions. Such results highlight how the environment rich of quarks and gluons present in pp collisions can alter hadronization, challenging the validity of independent fragmentation across collision systems. Actually, a full comprehension of the hadronization in pp collisions is still an open point, given that model calculations are not able to describe the measurements. Further effort is needed from both experimental and theoretical sides to finally solve this puzzle.

References

[1] B. Andersson, G. Gustafson, G. Ingelman, T. Sjostrand, Phys. Rept. **97**, 31 (1983)
[2] S. Ferreres-Solé, T. Sjöstrand, Eur. Phys. J. C **78**, 983 (2018), `1808.04619`
[3] J. Bellm et al., Eur. Phys. J. C **76**, 196 (2016), `1512.01178`
[4] M. Niiyama et al. (Belle), Phys. Rev. D **97**, 072005 (2018), `1706.06791`
[5] R. Aaij et al. (LHCb), JHEP **03**, 159 (2016), [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)], `1510.01707`
[6] S. Acharya et al. (ALICE), JHEP **05**, 220 (2021), `2102.13601`
[7] T. Sjostrand, S. Mrenna, P.Z. Skands, Comput. Phys. Commun. **178**, 852 (2008), `0710.3820`
[8] S. Acharya et al. (ALICE), Eur. Phys. J. C **80**, 693 (2020), `2003.02394`
[9] J.R. Christiansen, P.Z. Skands, JHEP **08**, 003 (2015), `1505.01681`
[10] R. Aaij et al. (LHCb), Phys. Rev. D **100**, 031102 (2019), `1902.06794`
[11] S. Acharya et al. (ALICE) (2020), `2011.06079`
[12] S. Acharya et al. (ALICE) (2020), `2011.06078`
[13] L. Gladilin, Eur. Phys. J. C **75**, 19 (2015), `1404.3888`
[14] M. He, R. Rapp, Phys. Lett. B **795**, 117 (2019), `1902.08889`
[15] D. Ebert, R.N. Faustov, V.O. Galkin, Phys. Rev. D **84**, 014025 (2011), `1105.0583`
[16] V. Minissale, S. Plumari, V. Greco, Phys. Lett. B **821**, 136622 (2021), `2012.12001`
[17] J. Song, H.h. Li, F.l. Shao, Eur. Phys. J. C **78**, 344 (2018), `1801.09402`
[18] S. Acharya et al. (ALICE), Phys. Rev. Lett. **128**, 012001 (2022), `2106.08278`
[19] S. Acharya et al. (ALICE), Phys. Rev. Lett. **127**, 272001 (2021), `2105.05187`
[20] S. Acharya et al. (ALICE), Phys. Lett. B **829**, 137065 (2022), `2111.11948`