Soft QCD theory

Torbjörn Sjöstrand[a,*]

[a]*Department of Astronomy and Theoretical Physics, Lund University,*
Sölvegatan 14A, SE-223 62 Lund, Sweden

E-mail: torbjorn.sjostrand@thep.lu.se

Soft QCD is beyond perturbative control, and therefore phenomenological models have to be developed. These are implemented and combined within event generators. Typical aspects considered are multiparton interactions, colour reconnection, and hadronization. Of special interest are non-universal features of hadronization at the LHC, such as the recently observed enhancement of charm and bottom baryons, and the bottom baryon asymmetry, as well as the some years earlier observed strangeness enhancement. Several new production mechanisms, such as junction reconnection, ropes and shove, have been proposed to address some of these issues. Alternatively, an admixture of a quark–gluon plasma component also in pp collisions is introduced in the core–corona approach.

The Tenth Annual Conference on Large Hadron Collider Physics - LHCP2022
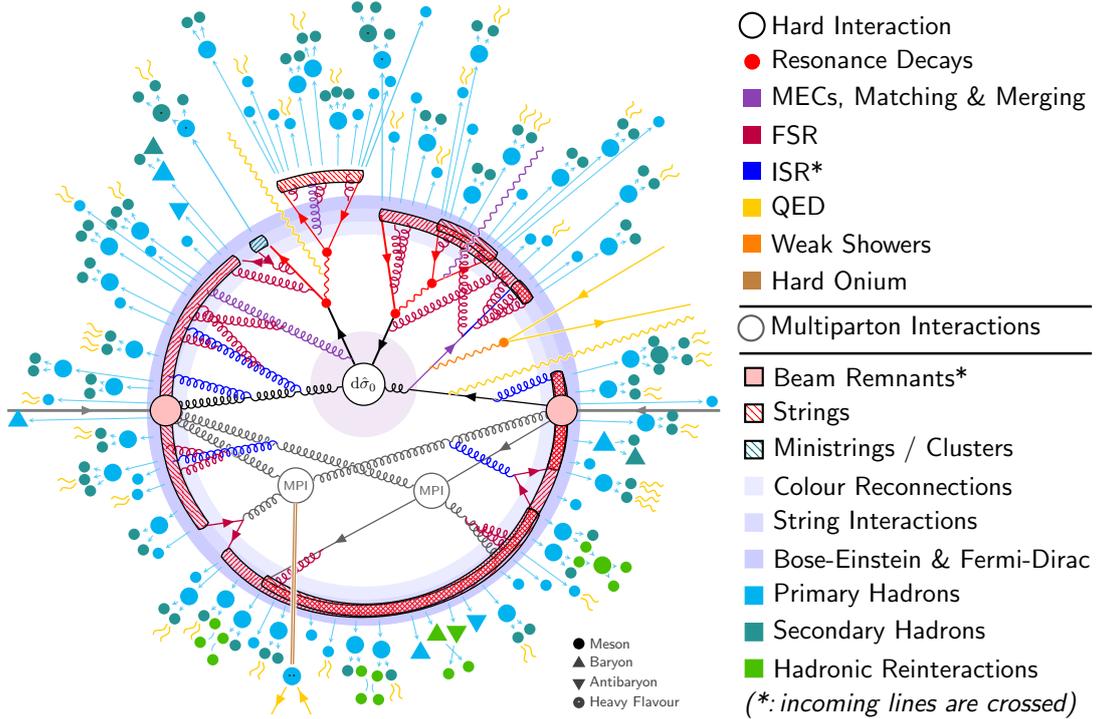*16-20 May 2022*
online

---

[*]Speaker

Figure 1: The main components in the structure of an event, illustrated for the case of top pair production, and approximately ordered from hard to soft in the list. Adapted from [1].

1. Introduction

The structure of a typical pp collision is illustrated in Fig. 1, split into the aspects we need to combine to obtain an overall description.

In the perturbative domain the hard interaction process defines the core, calculated by the combination of matrix elements and parton distributions. Decays of resonances such as W, Z, H and t also should be included in this core. Higher-order matrix element corrections, on the one hand, and parton showers, on the other, contribute to the emergence of multiparton topologies, and can partly overlap. Special matching and merging procedures therefore are devised to avoid doublecounting, and to provide a smooth transition between the two descriptions. The showers can be subdivided into initial- and final-state radiation, and may also include QED and weak emissions, and even the perturbative emission of charmonium and bottomonium states.

Since hadrons are composite, multiple parton–parton interactions may occur inside a single pp one. These straddle the perturbative–nonperturbative borderline.

Central in the soft/nonperturbative phase, occurring at later invariant time scales than the hard/perturbative one, is the formation of colour-confining force fields, called strings or clusters, that subsequently fragment to produce the primary set of hadrons. These fields are drawn between the partons produced in the hard phase of the event, plus the beam-remnant partons that passed through the interaction region unharmed. Before the fragmentation step the fields may interact with each other, notably by colour reconnection. After it, primary hadrons may decay over a wide range

of time scales. They may also rescatter against each other in the early stages, when the hadronic density is high. Bose–Einstein and Fermi-Dirac statistics can affect the production of identical particles during the fragmentation, decay and rescattering stages.

Given the need to combine many disparate effects over a wide range of scales, the most convenient approach is that of (Monte Carlo) event generators [2]. Here random numbers are used to emulate the quantum mechanical uncertainty at each step of the construction of the final state, and different code pieces can be called on to perform the required range of tasks. Generators can be used to predict event properties, for detector and trigger design, to correct data, for acceptance and smearing, and to interpret data in terms of underlying physics mechanisms.

The three most commonly used generators that combine hard and soft physics are Herwig [3], Pythia [1] and Sherpa [4]. Generators for matrix elements and their interface to parton showers notably include MadGraph5_aMC@NLO [5] and the PowHeg Box [6]. A wider range of generators exists for applications to other fields, such as heavy ion collisions or cosmic ray cascades, often modelling only soft-physics aspects [7]. Of special note is Epos [8], which spans the range between pp and heavy-ion collisions.

2. Multiparton interactions

Multiparton interactions (MPIs) are essential to explain the bulk properties of events, such as multiplicity distributions. Nevertheless the underlying theory and the modelling thereof is still hotly debated [9]. MPIs must extend down to a transverse momentum $p_\perp$ of 2–3 GeV to generate the observed amount of activity in events, a number that would seem to be in a perturbative regime, but the lower turnoff must come from nonperturbative physics, such as the screening of colour charges within the proton.

A similar dichotomy exists in the study of two hard subcollisions, double parton scattering (DPS), a subset of the MPIs. The standard formula for the cross section of simultaneously having processes A and B in an event is

$$\sigma_{AB}^{\text{DPS}} = \frac{m}{2} \frac{\sigma_A \, \sigma_B}{\sigma_{\text{eff}}} \quad \text{where } m = 1 \text{ if } A = B \text{ and } m = 2 \text{ else.}$$

Here σ_A and σ_B are perturbatively calculable, while σ_{eff} is a nonperturbative number. The latter is related to the nondiffractive cross section, of order 50 mb at LHC energies, but then reduced when the impact-parameter profile is considered. If protons are approximated by Gaussian distributions then $\sigma_{\text{eff}} \approx 25$ mb, and a picture with "hot spots" would give even lower numbers, i.e. more activity. Flavour and momentum conservation have opposite effects. Consider e.g. $W^- W^-$ production in the forward direction, where one W but not both could benefit from a high-x valence d quark.

By now, several σ_{eff} measurements have been done at the LHC, but there is a wide spread of results, in the range $5 - 20$ mb. It is not yet clear where this spread comes from. Numbers at the lower end of the range also would suggest an overall much larger MPI activity than what inclusive measurements do. An interesting CMS study of 4-jet production [10] shows that more than a factor of 2 spread comes from using different perturbative input, say leading vs. next-to-leading order. All data points also come with large error bars, that reflects the difficult separation of signal from background, the uncertain impact of more than two MPIs for the level of the underlying event, and possibly more. In the future it will become important to analyze several different processes under as identical conditions as possible, to better understand the origin of the current σ_{eff} spread.

3. Colour reconnection and baryon enhancement

The MPI activity leads to many colour fields — strings — being pulled out in pp collisions, most of them running essentially parallel with each other at low transverse momenta. Naively these strings would all be drawn out to the beam remnants, i.e. to the colour "hole" left behind by each parton that initiates an MPI. In colour reconnection (CR) scenarios the strings are drawn out less than expected. One reason could be that the starting picture assumes infinitely many colours, such that each string piece is uniquely defined, while the correct QCD picture provides ambiguities between identical colours. Reassignments that lead to shorter strings, i.e. smaller invariant masses, are then likely to be dynamically preferred. The CR picture explains e.g. why the average transverse momentum increases as a function of the charged multiplicity. At LEP2 the hadronic final state of W^+W^- is best described with a ~50% CR rate [11], but is still only 2.2σ away from no-CR.

New CR models continue to appear. One such is the QCDCR one [12]. In it, regular $q_1\bar{q}_2 + q_3\bar{q}_4 \to q_1\bar{q}_4 + q_3\bar{q}_2$ reconnections are supplemented by ones that produce junction topologies, i.e. where three color lines come together in a Y-shaped configuration. In the $q_1\bar{q}_2 + q_3\bar{q}_4$ example, if q_1 and q_3 are nearby on the left side of the event, and $\bar{q}_2$ and $\bar{q}_4$ ditto on the right one, the two long strings can collapse to one in the central region, giving $\succ\!\!\prec$ topologies. Here the left-side junction will be associated with the production of a baryon and the right-side antijunction with that of an antibaryon. Thus an extra source of baryon production is introduced, and one that becomes more important in high-multiplicity events, where more strings run parallel and can reconnect this way.

Extending this approach to three parallel strings, it is possible to obtain two disconnected junction systems, $q_1\bar{q}_2 + q_3\bar{q}_4 + q_5\bar{q}_6 \to q_1q_3q_5 + \bar{q}_2\bar{q}_4\bar{q}_6$. This latter mechanism has also been implemented in the HERWIG cluster fragmentation model [13].

This approach has come into the limelight by the recent ALICE observation of enhanced charm and bottom baryon production. Notably the fraction of charm quarks that become Λ_c^+ is slightly above 20% [14], about a factor 3 more than observed at LEP. This enhancement is concentrated to $p_\perp$ scales below 10 GeV [15], and above that the LEP numbers are approached. The data is well explained by the QCDCR model, where CR is most likely to occur in the packed low-$p_\perp$ region, while high-$p_\perp$ jets mainly evolve in a vacuum. Also some models based on quark recombination, thermodynamics or quark–gluon plasma formation give comparable results. A similar $p_\perp$-dependent behaviour has also been observed for Λ_b by LHCb [16]. One should add that the QCDCR model fails to describe all the relevant distributions, and thus is not the ultimate answer. It seems very likely, however, that baryon production by CR will become a main staple in the further development of CR and hadronization models.

4. Beam drag and forward physics

In fixed-target experiments it has been known since long that the flavour content of the beams influences the production of heavy hadrons. In $\pi^- p$ collisions, e.g., D^- production increasingly dominates over D^+ one at higher momenta in the π^- direction. This is easily understood for the subprocess $\bar{u}u \to \bar{c}c$, where the $\bar{c}$ is colour-connected to the d in the π^- remnant, while the c is connected to the ud proton remnant. Then the drag of the colour fields pulls $\bar{c}$ forwards and c backwards [17]. A low $\bar{c}d$ invariant mass even leads to a collapse directly into a D^-.

Similar effects are predicted both for charm and bottom at the LHC, leading to small asymmetries also at central rapidities. Asymmetries of around 2% have now been observed by LHCb, favouring Λ_b over $\overline{\Lambda}_b$ [18]. The sign and the slowly rising rapidity dependence of the effect agrees with default PYTHIA, but is about a factor of 2 lower. Interestingly, the QCDCR scenario agrees much better with data. Likely the junction mechanism adds a symmetric source of Λ_b and $\overline{\Lambda}_b$ production that dilutes the asymmetry present in the traditional baryon production. Furthermore, in default PYTHIA the asymmetry is peaked at small $p_\perp$, i.e. where the b/$\overline{\text{b}}$ quark is closer to the beam direction, while data and QCDCR show only a modest $p_\perp$ dependence. This is again consistent with a junction mechanism that is most active at small $p_\perp$, and therefore gives the largest dilution effect there.

More generally, our understanding of the forward region is still limited. Traditionally this has been a region of special interest mainly for cosmic-ray people, hence the LHCf experiment. Now new-particle searches and neutrino studies in the forward direction attracts increasing interest at the LHC. Comparisons with models have shown significant disagreements in many cases, and studies are under way to improve the situation. This may require a better modelling of beam remnants and diffraction, among others.

5. Hadronization

The ALICE observation of an increasing strangeness fraction with increasing multiplicity [19] was unexpected. The trend smoothly attaches to results for heavy-ion runs, where the high strangeness rate is coming from thermodynamics at the freezeout temperature, when the quark–gluon plasma (QGP) turns into hadrons. Traditional wisdom held that pp systems would not give sufficiently large and long-lived regions for a QGP to form, but this is now cast in doubt.

The simplest and most appealing solution is offered by core–corona models, notably the EPOS one. In it, initially all colour fields are drawn out from all partonic interactions. But then, in regions where these fields are very closely packed, it is assumed that they together melt into a QGP state. A typical topology therefore would have a dense QGP core surrounded by a normal-state dilute corona. A low-multiplicity pp collision would be all corona, and thus have the same particle composition as in e^+e^-, but with increasing multiplicity the fraction of core increases, and thereby the relative rate of strangeness production. In a nucleus–nucleus (AA) collision the core region would dominate, except for the most peripheral collisions, and the strangeness fraction saturate.

Nevertheless, it is interesting to consider scenarios that could avoid the introduction of QGP in pp collisions, and maybe even offer alternative explanations for (some of) the signals interpreted as proof of QGP in AA. One such example is rope formation [20], wherein nearby strings can fuse to objects of a larger colour representation. This gives a larger energy release when the ropes break, which translates into a smaller strangeness suppression than in a single string. Higher multiplicities means more strings and more possibilities for rope formation, such that the rope scenario is in reasonable agreement with data. Another possible mechanism is shove, the repulsion of nearby strings, which can give them a transverse motion inherited by the hadrons produced from them, resulting in a flow-like behaviour [21]. While these and other ideas are very interesting, it remains to put it all together in an combined framework that gives a convincing picture at least of pp collisions without any need for a QGP.

6. Summary and outlook

Before LHC, the concept of jet universality was commonly accepted. That is, once the perturbative evolution has run its course, the subsequent transformation from partons to hadrons should follow exactly the same rules, whether in e^+e^-, ep or pp events. We have now seen many LHC analyses that run counter to jet universality. The breaking notably occurs at small transverse momenta, up to order 10 GeV, whereas jet structure above that appears to follow expectations, as far as we can tell. It is a fairly safe bet that the origin of universality breaking is the close-packing of strings in the low-$p_\perp$ region. The new data has been very invigorating; the hadronization field, dormant for many years, now is experiencing the emergence of new ideas. We have highlighted some of them here — QGP on one side, junctions, ropes and shove on another — but it remains to combine the latter ideas into a consistent whole. And, as we have seen, the issues in soft QCD are not limited to the hadronization itself, but is related to the preceding multiparton interactions and colour reconnection descriptions as well, e.g. to set the stage for close-packing effects to occur.

In this article there has only been space to bring up a few of the data and models being discussed today. Several more have been presented in other contributions to this conference. It is clear that continued activity is called for, not only for LHC itself but also to help prepare for future colliders, such as EIC, ILC and FCC, and to strengthen ties to other fields of study, such as fixed-target neutrino interactions and cosmic-ray cascades.

Acknowledgments

References

[1] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2203.11601.

[2] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [1101.2599].

[3] J. Bellm et al., *Herwig 7.2 release note*, *Eur. Phys. J. C* **80** (2020) 452 [1912.06509].

[4] Sherpa collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[5] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer 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 [1405.0301].

[6] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].

[7] J.M. Campbell et al., *Event Generators for High-Energy Physics Experiments*, in *2022 Snowmass Summer Study*, 3, 2022 [2203.11110].

[8] T. Pierog, I. Karpenko, J.M. Katzy, E. Yatsenko and K. Werner, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, *Phys. Rev. C* **92** (2015) 034906 [1306.0121].

[9] P. Bartalini and J.R. Gaunt, eds., *Multiple Parton Interactions at the LHC*, vol. 29, WSP (2019), 10.1142/10646.

[10] CMS collaboration, *Measurement of double-parton scattering in inclusive production of four jets with low transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **01** (2022) 177 [2109.13822].

[11] ALEPH, DELPHI, L3, OPAL, LEP Electroweak Working Group collaboration, *A Combination of preliminary electroweak measurements and constraints on the standard model*, hep-ex/0612034.

[12] J.R. Christiansen and P.Z. Skands, *String Formation Beyond Leading Colour*, *JHEP* **08** (2015) 003 [1505.01681].

[13] S. Gieseke, P. Kirchgaeßer and S. Plätzer, *Baryon production from cluster hadronisation*, *Eur. Phys. J. C* **78** (2018) 99 [1710.10906].

[14] ALICE collaboration, *Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC*, *Phys. Rev. D* **105** (2022) L011103 [2105.06335].

[15] ALICE collaboration, *Measurement of Prompt D^0, Λ_c^+, and $\Sigma_c^{0,++}$(2455) Production in Proton–Proton Collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Rev. Lett.* **128** (2022) 012001 [2106.08278].

[16] LHCb collaboration, *Measurement of b hadron fractions in 13 TeV pp collisions*, *Phys. Rev. D* **100** (2019) 031102 [1902.06794].

[17] E. Norrbin and T. Sjöstrand, *Production and hadronization of heavy quarks*, *Eur. Phys. J. C* **17** (2000) 137 [hep-ph/0005110].

[18] LHCb collaboration, *Observation of a $\Lambda_b^0 - \overline{\Lambda}_b^0$ production asymmetry in proton-proton collisions at $\sqrt{s}$ = 7 and 8 TeV*, *JHEP* **10** (2021) 060 [2107.09593].

[19] ALICE collaboration, *Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions*, *Nature Phys.* **13** (2017) 535 [1606.07424].

[20] C. Bierlich, G. Gustafson, L. Lönnblad and A. Tarasov, *Effects of Overlapping Strings in pp Collisions*, *JHEP* **03** (2015) 148 [1412.6259].

[21] C. Bierlich, G. Gustafson and L. Lönnblad, *A shoving model for collectivity in hadronic collisions*, 1612.05132.