Modelling hadronic interactions in HEP MC generators

Peter Skands\textsuperscript{1,2,}\textsuperscript{a}

1 Theory Unit, CERN, CH-1211 Geneva 23, Switzerland
2 School of Physics and Astronomy, Monash University, Melbourne, Australia

Abstract. HEP event generators aim to describe high-energy collisions in full exclusive detail. They combine perturbative matrix elements and parton showers with dynamical models of less well-understood phenomena such as hadronization, diffraction, and the so-called underlying event. We briefly summarise some of the main concepts relevant to the modelling of soft/inclusive hadron interactions in MC generators, in particular PYTHIA, with emphasis on questions recently highlighted by LHC data.

1. Introduction

This summary was written in the context of the 18th International Symposium on Very High Energy Cosmic Ray Interactions (ISVHECRI). It is based on two recent proceedings-style mini-reviews of soft-inclusive event-generator models\cite{1,2}, updated and extended in a hopefully reasonably coherent and useful form. In Sect. 2, we give an overview of soft physics models, focusing on multi-parton interactions (MPI). In Sect. 3, we discuss the physics of colour reconnections (CR) along with some alternative proposed interpretations of observed properties of particle spectra in hadron collisions. Finally in Sect. 4, we give a brief overview of the most recent tuning efforts in the context of the PYTHIA 8 event generator, in particular the so-called Monash 2013 tune.

2. Soft physics models

Soft physics models can essentially be divided into two broad categories. The first starts from perturbative QCD (partons, matrix elements, jets) and uses a factorized perturbative expansion for the hardest parton-parton interaction, combined with parton showers and detailed models of hadronization and (soft and hard) multiparton interactions (MPI). This is the approach taken by general-purpose event generators, like HERWIG\cite{4,5}, PYTHIA\cite{6,7}, and SHERPA\cite{8}. Since they agree with perturbative QCD (pQCD) at high $p_\perp$, they are used extensively by the collider-physics community, see\cite{9,10} for reviews. The price is a typically low predictivity for very soft physics, though the modelling of diffractive and other soft-inclusive phenomena is generally improving, and is an active area of research in all the generators. Collisions involving nuclei with $A \geq 2$ are generally not addressed at all by these generators, though extensions exist\cite{11,12}.

At the other end of the spectrum are tools starting from Regge theory (optical theorem, cut and uncut pomeron), like QGSJET\cite{13}. A priori, there are no jets whatsoever in this formalism, and the dynamical picture is one of purely longitudinal strings breaking up and producing particles. To a high-$p_\perp$ collider physicist, the complete absence of jets may seem a quite radical starting assumption, but recall that the vast majority of the (soft-inclusive) cross section involves very small momentum transfers. These models are typically used e.g. for heavy-ion collisions and cosmic-ray air showers, for which the small fraction of events that contain hard identifiable jets can often be neglected (though obviously not in hard tails, such as jet-quenching studies). The main focus is here on the soft physics, though perturbative contributions can be added in, e.g. by the introduction of a “hard pomeron”. In-between are tools like PHOJET\cite{14}, DPMJET\cite{15}, EPOS\cite{16}, and SIBYLL\cite{17}, which contain elements of both languages (with EPOS adding a further component: hydrodynamics\cite{18}). Note, however, that all of these models rely on string models of hadronization and hence have some overlap with PYTHIA on that aspect of the event modelling.

Regardless of the details, any framework that attempts to combine soft and hard QCD eventually faces the following problem: At some point, the perturbatively calculable (“hard”) parton-parton cross section exceeds the total (“hard+soft”) hadron-hadron cross section. This is illustrated in Fig. 1, for proton-proton collisions at CM energies from 13 TeV (top pane) to 100 TeV (bottom pane). At each CM energy, the total (inelastic) hadron-hadron cross section (based on\cite{20,21}) is shown as a horizontal line with filled black squares. The x axis, labeled $p_{T\text{\_min}}$, represents an arbitrary lower limit of integration for the perturbative QCD $2 \rightarrow 2$ cross sections,

\begin{equation}
\sigma_{2\rightarrow2}(p_{T\text{\_min}}) = \int_{p_{T\text{\_min}}}^{p_{T_{\text{max}}}} d \hat{p}_T^2 \frac{d\hat{\sigma}_{2\rightarrow2}}{d\hat{p}_T^2},
\end{equation}

This is an Open Access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

\textsuperscript{a} c-mail: peter.skands@monash.edu

\textsuperscript{1} The first was a Snowmass / FCC-hh study\cite{1–3} focusing on extrapolations of soft-physics models to 100 TeV CM energies. The other was a contribution on the PYTHIA generator(s) to a CERN Yellow Report prepared by the LHC forward physics study group; in progress.

\textsuperscript{2} Are generally not
where the partonic $d\sigma_{2 \to 2}$ cross section is dominated by $t$-channel gluon exchange with a characteristic $1/s^2$ singularity, and we have suppressed integrations over partonic $x$ fractions. The divergence of this partonic cross section, $1/s^2 \sim 1/p_T^2$ for low $p_T$, augmented by running-coupling and low-$x$ parton-distribution effects, implies that at some $p_{T\text{min}}$ value, the red and blue parton-parton cross-section curves in Fig. 1 must exceed the total hadron-hadron one. (The small difference between the two curves represent different PDF and $\alpha_s$ choices.) At the LHC at 13 TeV (top pane), the parton-parton cross section becomes equal to the hadron-hadron one for $p_{T\text{min}}(13 \text{ TeV}) \sim 5 \text{ GeV}$, while the corresponding value at 100 TeV (bottom pane) is $p_{T\text{min}}(100 \text{ TeV}) \sim 10 \text{ GeV}$. Although these are arguably quite low scales in the context of “jets”, the main point is that they are still perturbative.

We do not naively expect that non-perturbative effects significantly reduce the $2 \to 2$ cross section at $p_T$ values as large as 10 GeV. The parton- and hadron-level cross sections can be reconciled by noting that their ratio,

$$\langle n \rangle (p_{T\text{min}}, s) = \frac{\sigma_{2 \to 2}(p_{T\text{min}}, s)}{\sigma_{\text{inel}}(s)} ,$$

counts how big a fraction of all (inelastic) events contain a partonic $2 \to 2$ scattering above a given $p_{T\text{min}}$, as a function of hadron-hadron CM energy, $\sqrt{s}$. If this fraction is greater than one, it simply means that each hadron-hadron collision contains more than one such partonic $2 \to 2$ scattering. Thus the idea is born: multiple perturbative parton-parton interactions (MPI).

As mentioned above, MPI has historically been an essential ingredient in the modelling of hadron-hadron collisions especially in PYTHIA (see [22]). When augmented by impact-parameter dependence (an aspect that goes beyond this mini-review), it allows to describe a number of important phenomenological features, such as the extremely wide multiplicity distributions and significant deviations from KNO scaling [23] observed already eg at the SPS [24, 25]. Modern implementations of partonic MPI are featured in EPOS, HERWIG++, PHOJET, PYTHIA 6.4 & 8, SHERPA, and SIBYLL 2.0, while QGSJET and SIBYLL 1 rely on multiple cut pomerons (i.e., not associated with partonic jets).

In perturbative MPI-based models, one should be aware that the amount of soft MPI is sensitive to the PDFs at low $x$ and $Q^2$, a region which is not especially well controlled. Physically, colour screening and/or saturation effects should be important. In practice, one typically introduces an $E_{CM}$-dependent regularisation scale, $p_{T\text{min}}(\sqrt{s})$ (of the same order and usually slightly smaller than the $p_{T\text{min}}$ scale discussed above), which is assumed to modify the naive LO QCD $2 \to 2$ cross sections in the following way,
In PYTHIA 8, the MPI model extends and improves the $p_{\perp}$-ordered one from PYTHIA 6. The main differences are: full interleaving of final-state showers with ISR and MPI [38]; a richer mix of MPI processes, including electroweak processes and multiple-$J/\psi$ and $\Upsilon$ production (see the HTML manual under “Multiparton Interactions:processLevel”); an option to select the second MPI “by hand” (see the HTML manual under “A Second Hard Process”); and an option for final-state parton-parton rescattering [39] (mimicking a mild collective-flow effect in the context of a dilute parton system, see the HTML manual under “Multiparton Interactions: Rescattering”); colour reconnections are handled somewhat differently (see the HTML manual and [40,41]); and an option for an $x$-dependent transverse proton size [32].

An example where the treatment in PYTHIA 8 already surpasses the one in PYTHIA 6 is hard diffraction (for soft diffraction, the modelling is the same between 6 and 8, though the diffractive and string-fragmentation tuning parameters may of course differ). The default modelling of hard diffraction in PYTHIA 8 is described in [42] and follows an Ingelman-Schlein approach [43] to introduce partonic substructure in high-mass diffractive scattering. (“High-mass” is defined as corresponding to diffractive masses greater than about 10 GeV, though this can be modified by the user, see the HTML manual under “Diffraction”). This gives rise to harder $p_{\perp}$ spectra and diffractive jets. A novel feature of the PYTHIA 8 implementation is that hard diffractive interactions can include MPI (inside the Pomeron-proton system such that the rapidity gap is not destroyed), with a rate governed by the (user-specifiable) Pomeron-proton total cross section, $\sigma_{p\gamma}$. This predicts that there should be an “underlying event” also in hard diffractive events, which could be searched for eg in the region “transverse” to diffractive jets, and/or in association with diffractive heavy quark production, which is currently being implemented in PYTHIA 8.

We should also note that the default parameterization of the $pp$ and $pp$ cross sections in PYTHIA 8 is still based on a fairly old (1992) Donnachie-Landshof fit [44], with an asymptotic behaviour $\sigma_{\text{TOT}} \propto s^{0.08}$. This is combined with Schuler-Sjöstrand parametrizations of the diffractive components [45]. More recent studies based on LHC data [20,46,47] indicate a steeper rise, $\propto s^{0.096}$ [21,48]. In the context of PYTHIA, the difference seems mainly to be reflected in PYTHIA predicting a too small elastic component agrees well with LHC data. Updating the total cross sections is on the “to-do” list for a future version of PYTHIA 8.

Finally, an alternative treatment relying on the min-bias Rockefeller (MBR) model is also available in PYTHIA 8 [49].

3. Colour reconnections

The issue of final-state colour reconnections (CR) is becoming increasingly recognised as one of the main outstanding problems in soft-inclusive hadron-hadron physics [40,41,50,51], with significant potential implications not only for min-bias type physics but also impacting high-$p_{\perp}$ precision measurements such as the top quark mass [51,52].
Physically, CR may reflect a generalisation of soft colour coherence, dense-packing, and/or collective effects (parton-, string-, or hadron-rescattering). Disentangling the causes and effects of CR is likely to be a crucial topic for soft-QCD studies to unravel during the coming years. This will require the definition and study of CR-sensitive observables and a detailed consideration of the interplay between PDFs, MPI, and diffractive physics, with MPI possibly contributing to destroying rapidity gaps in “originally” diffractive events, and CR possibly creating them in “originally” non-diffractive ones [50,53].

In the context of MPI models, the question of CR arises naturally when one considers how the additional parton-parton interactions should be represented in terms of strings or clusters fragmenting into hadrons. An ad-hoc solution could be to represent the additional MPI systems as overall colour singlets, i.e., individual strings or clusters hadronizing separately from the rest of the event. This simple scenario existed as an option in the original PYTHIA implementation [22], was the basis of the initial HERWIG++ MPI modelling [54], and as far as I understand is also the basis of the modelling of the fragmentation of cut pomeron in QGSJET [13]. However, this colour structure physically corresponds to a diffractive colour flow (singlet exchange), which is not consistent with r-channel gluon (colour-octet) or cut-pomeron exchange. Empirically, it also leads to conflict with the data and e.g., produces too large forward peaks in the charged-particle pseudorapidity spectrum (from disconnected “diffractive-looking” MPI systems boosted along the z axis) [55], and predicts that the average transverse momentum is roughly independent of charged multiplicity, in stark contrast to observations [22]. Therefore the MPI models in both PYTHIA 6 and 8 also included an alternative (though still rather ad hoc) option for “inserting” the additional partons from MPI onto the string pieces created by the primary interaction, in a way designed to minimize the overall increase in “string length” [22,38,39]. In studies of Tevatron data, Rick Field in particular found a very strong preference for the minimal-string-length option, resulting in the famous “Tune A” family of tunes [56], which were the first to deliver a satisfactory description of the underlying event at high energies. There remained the physics question of how nature arranged for this preference to be selected. Early attempts at modelling realistic colour flow with octet exchanges did not provide an explanation [37], and the most successful models today are still driven by string-length minimizations [50,51,55], without any particularly deep understanding of the microphysics involved. An alternative scenario is provided by EPOS, which assumes that high string densities triggers a hydrodynamic phase [18]. Other possibilities currently under development include the idea of “colour ropes” [57,58] – strings carrying several units of colour charge (instead of n ordinary strings on top of each other) – and generalised colour coherence applied to the process of string formation [41]. Whatever the case, this is clearly a fertile area for model building today, with potentially important consequences, and for which we are already aware of several sensitive observables, most importantly the evolution of $p_{\perp}$ spectra with charged multiplicity (and particle mass, see e.g. [59]), but also heavy-ion inspired flow-type observables could be revealing, the dependence of particle spectra in the underlying event on underlying-event activity for fixed jet $p_{\perp}$, and the emergence and destruction of rapidity gaps could all carry sensitive additional information.

4. Recent PYTHIA tunes

The most recent PYTHIA 8 tune is currently the Monash 2013 tune [19], which has been selected as the new default tune since version 8.2 [7], replacing the earlier Tune 4C default of Pythia 8.1 [38,60]. A summary of the main properties are as follows: for the final-state fragmentation, it allows 10% more strangeness in string breaks, and has somewhat softer heavy-quark (c and b) fragmentation functions, achieving better agreement with s-, c-, and b-sensitive observables at LEP and SLD. In the context of $pp$ collisions, it is based on a new LO NNPDF 2.3 PDF set [61–63], which has a slightly larger low-x gluon than the previous default CTEQ6L1 set [64], hence the Monash 2013 tune produces more forward activity. There is also a better agreement with the energy scaling of average min-bias multiplicities at LHC energies, from 900 to 7000 GeV [65].

The tuning efforts, however, did not explicitly attempt to retune the diffractive components, and there are still significant discrepancies for identified-particle rates and spectra in $pp$ collisions. Those may point towards a need for better CR models [41] and/or for inclusion of other soft/collective effects, whatever their origin.

Important remaining open questions include dedicated tuning studies in the context of diffraction, for instance to constrain the total Pomeron-proton cross section, $\sigma_{pp}$, which controls the amount of MPI in hard diffractive processes, the sensitivity to the diffractive PDFs, and dedicated tests of string-fragmentation parameters in the specific context of diffractive final states, as compared with LEP-tuned parameters.

For completeness, we note that the most recent author-driven PYTHIA 6 tunes are the so-called Perugia 2012 set of tunes [26], now superseded by the Monash 2013 tune of PYTHIA 8. We hope to provide Perugia-like tune- variations also for PYTHIA 8 tunes in the future, though this was not done in the context of the Monash 2013 tune.

This work was supported in part by the Research Executive Agency (REA) of the European Commission under the Grant Agreements PITN-GA-2012-315877 (MCnet).

References

[1] P.Z. Skands (2013), 1308.2813
[2] J. Campbell, K. Hatakeyama, J. Huston, F. Petriello, J.R. Andersen et al. (2013), 1310.5189
[3] W. Barletta, M. Battaglia, M. Klute, M. Mangano, S. Prestemon et al., Nucl. Instrum. Meth. A764, 352 (2014)
