38. MONTE CARLO EVENT GENERATORS

Written January 2012 by P. Nason (INFN, Milan) and P.Z. Skands (CERN)

General-purpose Monte Carlo (GPMC) generators like `HERWIG` [142], `HERWIG++` [143], `PYTHIA` 6 [139], `PYTHIA` 8 [140], and `SHERPA` [5], provide fully exclusive modeling of high-energy collisions. They play an essential role in QCD modeling (in particular for aspects beyond fixed-order perturbative QCD), in data analysis, where they are used together with detector simulation to provide a realistic estimate of the detector response to collision events, and in the planning of new experiments, where they are used to estimate signals and backgrounds in high-energy processes. They are built from several components, that describe the physics starting from very short distance scales, up to the typical scale of hadron formation and decay. Since QCD is weakly interacting at short distances (below a femtometer), the components of the GPMC dealing with short-distance physics are based upon perturbation theory. At larger distances, all soft hadronic phenomena, like hadronization and the formation of the underlying event, cannot be computed from first principles, and one must rely upon QCD-inspired models.

The purpose of this review is to illustrate the main components of these generators. It is divided into four sections. The first one deals with short-distance, perturbative phenomena. The basic concepts leading to the simulations of the dominant QCD processes are illustrated here. In the second section, hadronization phenomena are treated. The two most popular hadronization models for the formation of primary hadrons, the string and cluster models, are illustrated. The basics of the implementation of primary-hadron decays into stable ones is also illustrated here. In the third section, models for soft hadron physics are discussed. These include models for the underlying event, and for minimum-bias interactions. Issues of Bose-Einstein and color-reconnection effects are also discussed here. The fourth section briefly introduces the problem of MC tuning.

We use natural units throughout, such that $c = 1$ and $\hbar = 1$, with energy, momenta and masses measured in GeV, and time and distances measured in GeV^{-1}.

38.1. Short-distance physics in GPMC generators

The short-distance components of a GPMC generator deal with the computation of the primary process at hand, with decays of short-lived particles, and with the generation of QCD and QED radiation, on time scales below $1/\Lambda$, with Λ denoting a typical hadronic scale of a few hundred MeV, corresponding roughly to an inverse femtometer. In e^+e^- annihilation, for example, the short-distance physics describes the evolution of the system from the instant when the e^+e^- pair annihilates up to a time when the size of the produced system is just below a femtometer.

In the present discussion we take the momentum scale of the primary process to be $Q \gg \Lambda$, so that the corresponding time and distance scale $1/Q$ is small. Soft- and collinear-safe inclusive observables, such as total decay widths or inclusive cross sections, can be reliably computed in QCD perturbation theory (pQCD), with the perturbative expansion truncated at any fixed order n, and the remainder suppressed by $\alpha_{\mathrm{S}}(Q)^{n+1}$.

Less inclusive observables, however, can receive large enhancements that destroy the convergence of the fixed-order expansion. This is due to the presence of collinear and infrared singularities in QCD. Thus, for example, a correction in which a parton from the primary interaction splits collinearly into two partons of comparable energy, is of order $\alpha_S(Q)\ln(Q/\Lambda)$, where the logarithm arises from an integral over a singularity regulated by the hadronic scale Λ. Since $\alpha_S(Q) \propto 1/\ln(Q/\Lambda)$, the corresponding cross section receives a correction of order unity. Two subsequent collinear splittings yield $\alpha_S^2(Q)\ln^2(Q/\Lambda)$, and so on. Thus, corrections of order unity arise at all orders in perturbation theory. The dominant region of phase space is the one where radiation is strongly ordered in a measure of hardness and/or angle. This means that, from a typical final-state configuration, by clustering together final-state parton pairs with, say, the smallest angle, recursively, we can reconstruct a branching tree, that may be viewed as the splitting history of the event. This history necessarily has some dependence on which measure is used to order the clusterings. However, strong ordering in energy times angle, in virtuality or in transverse momenta are in fact equivalent in the dominant region. In fact, in the small-angle limit, the virtuality t of a parton of energy E, splitting into two on-shell partons is given by

$$t = E^2 z(1-z)(1-\cos\theta) \approx \frac{z(1-z)}{2} E^2\theta^2\,, \tag{38.1}$$

where z and $1-z$ are the energy fractions carried by the produced partons, and θ is their relative angle. The transverse momentum of the final partons relative to the direction of the incoming one is given by

$$p_T^2 \approx z^2(1-z)^2E^2\theta^2. \tag{38.2}$$

Thus, significant differences between these measures only arise in regions with very small z or $1-z$ values. In QCD, because of soft divergences, these regions are in fact important, and the choice of the appropriate ordering variable is very relevant (see Sec. 38.3).

The so called KLN theorem [6,7] guarantees that large logarithmically divergent corrections, arising from final-state collinear splitting and from soft emissions, cancel against the virtual corrections in the total cross section, order by order in perturbation theory. Furthermore, the factorization theorem guarantees that initial-state collinear singularities can be factorized into the parton density functions (PDFs). Therefore, the cross section for the basic process remains accurate up to corrections of higher orders in $\alpha_S(Q)$, provided it is interpreted as an inclusive cross section, rather than as a bare partonic cross section. Thus, for example, the leading order (LO) cross section for $e^+e^- \to q\bar{q}$ is a good LO estimate of the e^+e^- cross section for the production of a pair of quarks accompanied by an arbitrary number of collinear and soft gluons, but is not a good estimate of the cross section for the production of a $q\bar{q}$ pair with no extra radiation.

Shower algorithms are used to compute the cross section for generic hard processes including all leading-logarithmic (LL) corrections. These algorithms begin with the generation of the kinematics of the basic process, performed with a probability proportional to its LO partonic cross section, which is interpreted physically as the inclusive cross section for the basic process, followed by an arbitrary sequence of small-angle splittings. A probability is then assigned to each splitting sequence. Thus,

the initial LO cross section is partitioned into the cross sections for a multitude of final states of arbitrary multiplicity. The sum of all these partial cross sections equals that of the primary process. This property of the GPMCs reflects the KLN cancellation mentioned earlier, and it is often called "unitarity of the shower process", a name that reminds us that the KLN cancellation itself is a consequence of unitarity. The fact that a quantum mechanical process can be described in terms of composition of probabilities, rather than amplitudes, follows from the LL approximation. In fact, in the dominant, strongly ordered region, subsequent splittings are separated by increasingly large times and distances, and this suppresses interference effects.

We now illustrate the basic parton-shower algorithm, as first introduced in Ref. 8. The purpose of this illustration is to give a schematic representation of how shower algorithms work, to introduce some concepts that will be referred to in the following, and to show the relationship between shower algorithms and Feynman-diagram results. For simplicity, we consider the example of e^+e^- annihilation into $q\bar{q}$ pairs. With each dominant (i.e. strongly ordered) final-state configuration one can associate a unique ordered tree diagram, by recursively clustering together final-state parton pairs with the smallest angle, and ending up with the hard production vertex (i.e. the $\gamma^* \to q\bar{q}$). The momenta of all intermediate lines of the tree diagram are then uniquely determined from the final-state momenta. Virtualities in the graph are also strongly ordered. One assigns to each splitting vertex a virtuality t, equal to the invariant mass of the pair of generated partons, the energy fractions z and $1-z$ of the two generated partons, and the azimuth ϕ of the splitting process with respect to the momentum of the incoming parton. For definiteness, we assume that z and ϕ are defined in the center-of-mass (CM) frame of the e^+e^- collision, although other definitions are possible that differ only beyond the LL approximation. The differential cross section for a given final state is given by the product of the differential cross section for the initial $e^+e^- \to q\bar{q}$ process, multiplied by a factor

$$\Delta_i(t,t')\frac{\alpha_{\rm S}(t)}{2\pi}P_{i,jk}(z)\frac{dt}{t}dz\frac{d\phi}{2\pi} \tag{38.3}$$

for each intermediate line ending in a splitting vertex. We have denoted with t' the maximal virtuality that is allowed for the line, with t its virtuality, and z and ϕ refer to the splitting process. $\Delta(t,t')$ is the so-called Sudakov form factor

$$\Delta_i(t,t') = \exp\left[-\int_t^{t'}\frac{dq^2}{q^2}\,\frac{\alpha_{\rm S}(q^2)}{2\pi}\sum_{jk}P_{i,jk}(z)dz\frac{d\phi}{2\pi}\right] . \tag{38.4}$$

The suffixes i and jk represent the parton species of the incoming and final partons, respectively, and $P_{i,jk}(z)$ are the Altarelli-Parisi [9] splitting kernels. Final-state lines that do not undergo any further splitting are associated with a factor

$$\Delta_i(t_0,t') , \tag{38.5}$$

where t_0 is an infrared cutoff defined by the shower hadronization scale (at which the charges are screened by hadronization) or, for an unstable particle, its width (a source cannot emit radiation with a period exceeding its lifetime).

Notice that the definition of the Sudakov form factor is such that

$$\Delta_i(t_2,t_1)+\int_{t_2}^{t_1}\frac{dt}{t}dz\frac{d\phi}{2\pi}\sum_{jk}\Delta_i(t,t_1)\frac{\alpha_{\rm S}(t)}{2\pi}P_{i,jk}(z)=1\,. \tag{38.6}$$

This implies that the cross section for developing the shower up to a given stage does not depend on what happens next, since subsequent factors for further splitting or not splitting add up to one.

The shower cross section can then be formulated in a probabilistic way. The Sudakov form factor $\Delta_i(t_2,t_1)$ is interpreted as the probability for a splitting *not* to occur, for a parton of type i, starting from a branching vertex at the scale t_1, down to a scale t_2. Notice that $0<\Delta_i(t_2,t_1)\leq 1$, where the upper extreme is reached for $t_2=t_1$, and the lower extreme is approached for $t_2=t_0$. From Eq. (38.4), it seems that the Sudakov form factor should vanish if $t_2=0$. However, because of the presence of the running coupling in the integrand, t_2 cannot be taken smaller than some cutoff scale of the order of Λ, so that at its lower extreme the Sudakov form factor is small, but not zero. Event generation then proceeds as follows. One gets a uniform random number $0\leq r\leq 1$, and seeks a solution of the equation $r=\Delta_i(t_2,t_1)$ as a function of t_2. If r is too small and no solution exists, no splitting is generated, and the line is interpreted as a final parton. If a solution t_2 exists, a branching is generated at the scale t_2. Its z value and the final parton species jk are generated with a probability proportional to $P_{i,jk}(z)$. The azimuth is generated uniformly. This procedure is started with both the initial quark and the antiquark, and is applied recursively to all generated partons, thus producing two shower cascades. It may generate an arbitrary number of partons, and it stops when no final-state partons undergo further splitting.

We emphasize that the shower cross section described above can be derived from perturbative QCD by keeping only the collinear-dominant real and virtual contributions to the cross section. In particular, up to terms that vanish after azimuthal averaging, the product of the cross section for the basic process, times the factors

$$\frac{\alpha_{\rm S}}{2\pi}\frac{dt}{t}dz\frac{d\phi}{2\pi}P_{i,jk}(z) \tag{38.7}$$

at each branching vertex, gives the leading collinear contribution to the tree-level cross section for the same process. The dominant virtual corrections in the same approximation are provided by the running coupling at each vertex and by the Sudakov form factors in the intermediate lines.

38.1.1. *Angular correlations* :

In gluon splitting processes ($g\to q\bar{q}$, $g\to gg$) in the collinear approximation, the distribution of the split pair is not uniform in azimuth, and the Altarelli-Parisi splitting functions are recovered only after azimuthal averaging. This dependence is due to the interference of positive and negative helicity states for the gluon that undergoes splitting. Spin correlations propagate through the splitting process, and determine acausal correlations of the EPR kind [10]. A method to partially account for these effects

was introduced in Ref. 11, in which the azimuthal correlation between two successive splittings is computed by averaging over polarizations. This can then be applied at each branching step. Acausal correlations are argued to be small, and are discarded with this method, that is still used in the PYTHIA code [139]. A method that fully includes spin correlation effects was later proposed by Collins [12], and has been implemented in the fortran HERWIG code [13].

38.1.2. ***Initial-state radiation*** :
Initial-state radiation (ISR) arises because incoming charged particles can radiate before entering the hard-scattering process. In doing so, they acquire a non-vanishing transverse momentum, and their virtuality becomes negative (spacelike). The dominant logarithmic region is the collinear one, where virtualities become larger and larger in absolute value with each emission, up to a limit given by the hardness of the basic process itself. A shower that starts by considering the highest virtualities first would thus have to work backward in time for ISR. A corresponding backwards-evolution algorithm was formulated by Sjöstrand [14], and was basically adopted in all shower models.

The key point in backwards evolution is that the evolution probability depends on the amount of partons that could have given rise to the one being evolved. This is reflected by introducing the ratio of the PDF after the branching to the PDF before the branching in the definition of the backward-evolution Sudakov form factor,

$$\Delta_i^{\mathrm{ISR}}(t,t') = \exp\left[-\int_{t'}^{t}\frac{dt''}{t''}\frac{\alpha_{\mathrm{S}}(t'')}{2\pi}\int_x^1\frac{dz}{z}\sum_{jk}P_{j,ik}(z)\frac{f_j(t'',x/z)}{f_i(t'',x)}\right] . \tag{38.8}$$

Notice that there are two uses of the PDFs: they are used to compute the cross section for the basic hard process, and they control ISR via backward evolution. Since the evolution is generated with leading-logarithmic accuracy, it is acceptable to use two different PDF sets for these two tasks, provided they agree at the LO level.

In the context of GPMC evolution, each ISR emission generates a finite amount of transverse momentum. Details on how the recoils generated by these transverse "kicks" are distributed among other partons in the event, in particular the ones involved in the hard process, constitute one of the main areas of difference between existing algorithms, see Ref. 146. An additional $\mathcal{O}(1\,\mathrm{GeV})$ of "primordial k_T" is typically added, to represent the sum of unresolved and/or non-perturbative motion below the shower cutoff scale.

38.1.3. ***Soft emissions and QCD coherence*** :
In massless field theories like QCD, there are two sources of large logarithms of infrared origin. One has to do with collinear singularities, which arise when two final-state particles become collinear, or when a final-state particle becomes collinear to an initial-state one. The other has to do with the emission of soft gluons at arbitrary angles. Because of that, it turns out that in QCD perturbation theory two powers of large logarithms can arise for each power of α_{S}. The expansion in leading soft and collinear logarithms is often referred to as the double-logarithmic expansion.

Within the conventional parton-shower formalism, based on collinear factorization, it was shown in a sequel of publications (see Ref. 16 and references therein) that the double-logarithmic region can be correctly described by using the angle of the emissions as the ordering variable, rather than the virtuality, and that the argument of α_S at the splitting vertex should be the relative parton transverse momentum after the splitting. Physically, the ordering in angle approximates the coherent interference arising from large-angle soft emission from a bunch of collinear partons. Without this effect, the particle multiplicity would grow too rapidly with energy, in conflict with e^+e^- data. For this reason, angular ordering is used as the evolution variable in both the HERWIG [16] and HERWIG++ [17] programs, and an angular veto is imposed on the virtuality-ordered evolution in PYTHIA 6 [18].

A radical alternative formulation of QCD cascades first proposed in Ref. 19 focuses upon soft emission, rather than collinear emission, as the basic splitting mechanism. It then becomes natural to consider a branching process where it is a parton pair (i.e. a dipole) rather than a single parton, that emits a soft parton. Adding a suitable correction for non-soft, collinear partons, one can achieve in this framework the correct logarithmic structure for both soft and collinear emissions in the limit of large number of colors N_c, without any explicit angular-ordering requirement. The ARIADNE [145] and VINCIA [21] programs are based on this approach. In SHERPA, the default shower [22] is also of a dipole type [23], while the $p_\perp$-ordered showers in PYTHIA 6 and 8 represent a hybrid, combining collinear splitting kernels with dipole kinematics [24].

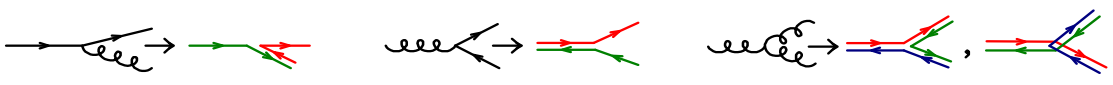
38.1.4. *Massive quarks* :

Quark masses act as cut-off on collinear singularities. If the mass of a quark is below, or of the order of Λ, its effect in the shower is small. For larger quark masses, like in c or b production, it is the mass, rather than the typical hadronic scale, that cuts off collinear radiation. For a quark with energy E and mass m_Q, the divergent behavior $d\theta/\theta$ of the collinear splitting process is regulated for $\theta \leq \theta_0 = m_Q/E$. We thus expect less collinear activity for heavy quarks than for light ones, which in turn is the reason why heavy quarks carry a larger fraction of the momentum acquired in the hard production process.

This feature can be implemented with different levels of sophistication. Using the fact that soft emission exhibits a zero at zero emission angle, older parton shower algorithms simply limited the shower emission to be not smaller than the angle θ_0. More modern approaches are used in both PYTHIA, where mass effects are included using a kind of matrix-element correction method [25], and in HERWIG++ and SHERPA, where a generalization of the Altarelli-Parisi splitting kernel is used for massive quarks [26].

38.1.5. *Color information* :

Shower MC generators track large-N_c color information during the development of the shower. In the large-N_c limit, a quark is represented by a color line, i.e. a line with an arrow in the direction of the shower development, an antiquark by an anticolor line, with the arrow in the opposite direction, and a gluon by a pair of color-anticolor lines. The rules for color propagation are:



(38.9)

At the end of the shower development, partons are connected by color lines. We can have a quark directly connected by a color line to an antiquark, or via an arbitrary number of intermediate gluons, as shown in Fig. 38.1.

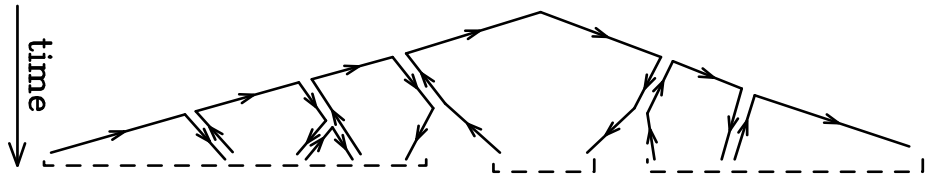


Figure 38.1: Color development of a shower in e^+e^- annihilation. Systems of color-connected partons are indicated by the dashed lines.

It is also possible for a set of gluons to be connected cyclically in color, as e.g. in the decay $\Upsilon \to ggg$.

The color information is used in angular-ordered showers, where the angle of color-connected partons determines the initial angle for the shower development, and in dipole showers, where dipoles are always color-connected partons. It is also used in hadronization models, where the initial strings or clusters used for hadronization are formed by systems of color-connected partons.

38.1.6. ***Electromagnetic corrections*** :
The physics of photon emission from light charged particles can also be treated with a shower MC algorithm. A high-energy electron, for example, is accompanied by bremsstrahlung photons, which considerably affect its dynamics. Also here, similarly to the QCD case, electromagnetic corrections are of order $\alpha_{\rm em} \ln Q/m_e$, or even of order $\alpha_{\rm em} \ln Q/m_e \ln E_\gamma/E$ in the region where soft photon emission is important, so that their inclusion in the simulation process is mandatory. This can be done with a Monte Carlo algorithm. In case of photons emitted by leptons, at variance with the QCD case, the shower can be continued down to values of the lepton virtuality that are arbitrarily close to its mass shell. In practice, photon radiation must be cut off below a certain energy, in order for the shower algorithm to terminate. Therefore, there is always a minimum energy for emitted photons that depends upon the implementations (and so does the MC truth for a charged lepton). In the case of electrons, this energy is typically of the order of its mass. Electromagnetic radiation below this scale is not enhanced by collinear singularities, and is thus bound to be soft, so that the electron momentum is not affected by it.

For photons emitted from quarks, we have instead the obvious limitation that the photon wavelength cannot exceed the typical hadronic size. Longer-wavelength photons are in fact emitted by hadrons, rather than quarks. This last effect is in practice never modeled by existing shower MC implementations. Thus, electromagnetic radiation from quarks is cut off at a typical hadronic scale.

38.1.7. ***Beyond-the-Standard-Model Physics*** :
The inclusion of processes for physics beyond the Standard Model (BSM) in event generators is to some extent just a matter of implementing the relevant hard processes and (chains of) decays, with the level of difficulty depending on the complexity of the model and the degree of automation [27,28]. Notable exceptions are long-lived colored particles [29], particles in exotic color representations, and particles showering under new gauge symmetries, with a growing set of implementations documented in the individual GPMC manuals. Further complications that may be relevant are finite-width effects (discussed in Sec. 38.1.8) and the assumed threshold behavior.

In addition to code-specific implementations [146], there are a few commonly adopted standards that are useful for transferring information and events between codes. Currently, the most important of these is the Les Houches Event File (LHEF) standard [30], normally used to transfer parton-level events from a hard-process generator to a shower generator. Another important standard is the Supersymmetry Les Houches Accord (SLHA) format [31], originally used to transfer information on supersymmetric particle spectra and couplings, but by now extended to apply also to more general BSM frameworks and incorporated within the LHEF standard [32].

38.1.8. ***Decay Chains and Particle Widths*** :
In most BSM processes and some SM ones, an important aspect of the event simulation is how decays of short-lived particles, such as top quarks, electroweak and Higgs bosons, and new BSM resonances, are handled. We here briefly summarize the spectrum of possibilities, but emphasize that there is no universal standard. Users are advised to check whether the treatment of a given code is adequate for the physics study at hand.

The appearance of an unstable resonance as a physical particle at some intermediate stage of the event generation implies that its production and decay processes are treated as being factorized. This is valid up to corrections of order Γ/m_0, with Γ the width and m_0 the pole mass. States whose widths are a substantial fraction of their mass should not be treated as "physical particles," but rather as intrinsically off-shell internal propagator lines.

For states treated as physical particles, two aspects are relevant: the mass distribution of the decaying particle itself and the distributions of its decay products. For the former, matrix-element generators often use a simple δ function at m_0. The next level up, typically used in GPMCs, is to use a Breit-Wigner distribution (relativistic or non-relativistic), which formally resums higher-order virtual corrections to the mass distribution. Note, however, that this still only generates an improved picture for *moderate* fluctuations away from m_0. Similarly to above, particles that are significantly off-shell (in units of Γ) should not be treated as resonant, but rather as internal off-shell propagator lines. In most GPMCs, further refinements are included, for instance by letting Γ be a function of m ("running widths") and by limiting the magnitude of the allowed fluctuations away from m_0.

For the distributions of the decay products, the simplest treatment is again to assign them their respective m_0 values, with a uniform phase-space distribution. A more sophisticated treatment distributes the decay products according to the differential decay

matrix elements, capturing at least the internal dynamics and helicity structure of the decay process, including EPR-like correlations. Further refinements include polarizations of the external states [33] and assigning the decay products their own Breit-Wigner distributions, the latter of which opens the possibility to include also intrinsically off-shell decay channels, like $H \to WW^*$.

During subsequent showering of the decay products, most parton-shower models will preserve their total invariant mass, so as not to skew the original resonance shape.

When computing partial widths and/or modifying decay tables, one should be aware of the danger of double-counting intermediate on-shell particles, see Sec. 38.2.3.

38.1.9. ***Matching with Matrix Elements*** :
Shower algorithms are based upon a combination of the collinear (small-angle) and soft (small-energy) approximations and are thus inaccurate for hard, large-angle emissions. They also lack next-to-leading order (NLO) corrections to the basic process.

Traditional GPMCs, like `HERWIG` and `PYTHIA`, have included for a long time the so called Matrix Element Corrections (MEC), first formulated in Ref. 34 with later developments summarized in Ref. 146. They are available for processes involving two incoming and one outgoing or one incoming and two outgoing particles, like DIS, vector boson and Higgs production and decays, and top decays. The MEC corrects the emission of the hardest jets at large angles, so that it becomes exact at leading order.

In the past decade, considerable progress has taken place in order to improve the parton shower description of hard collisions, in two different directions: the so called Matrix Elements and Parton Shower matching (ME+PS from now on), and the matching of NLO calculations and Parton Showers (NLO+PS).

The ME+PS method allows one to use tree-level matrix elements for hard, large-angle emissions. It was first formulated in the so-called CKKW paper [150], and several variants have appeared, including the CKKW-L, MLM, and pseudoshower methods, see Refs. 151, 146 for summaries. Truncated showers are required [153] in order to maintain color coherence when interfacing matrix-element calculations to angular-ordered parton showers using these methods. It is also important to ensure consistent α_S choices between the real (ME-driven) and virtual (PS-driven) corrections [38].

In the ME+PS method one typically starts by generating exact matrix elements for the production of the basic process plus a certain number $\leq n$ of other partons. A minimum separation is imposed on the produced partons, requiring, for example, that the relative transverse momentum in any pair of partons is above a given cut $Q_{\rm cut}$. One then reweights these amplitudes in such a way that, in the strongly ordered region, the virtual effects that are included in the shower algorithm (i.e. running couplings and Sudakov form factors) are also accounted for. At this stage the generated configurations are tree-level accurate at large angle, and at small angle they match the results of the shower algorithm, except that there are no emissions below the scale $Q_{\rm cut}$, and no final states with more than n partons. These kinematic configurations are thus fed into a GPMC, that must generate all splittings with relative transverse momentum below the scale $Q_{\rm cut}$, for initial events with less than n partons, or below the scale of the smallest pair transverse momentum, for events with exactly n partons. The matching parameter

$Q_{\rm cut}$ must be chosen to be large enough for fixed-order perturbation theory to hold, but small enough so that the shower is accurate for emissions below it. Notice that the accuracy achieved with MEC is equivalent to that of ME+PS with $n = 1$, where MEC has the advantage of not having a matching parameter $Q_{\rm cut}$.

The popularity of the ME+PS method is due to the fact that processes with many jets appear often as background of new physics searches. These jets are typically required to be well separated, and to have large transverse momenta. These kinematical configurations, away from the small-angle region, are precisely those where GPMCs fail to be accurate, and it is thus mandatory to describe them using exact tree-level matrix element calculations.

The NLO+PS methods extend the accuracy of the generation of the basic process at the NLO level in QCD. They must thus include the radiation of an extra parton with full tree-level accuracy, since this radiation constitutes a NLO correction to the basic process. They must also include NLO virtual corrections. They can be viewed as an extension of the MEC method with the inclusion of NLO virtual corrections. They are however more general, since they are applicable to processes of arbitrary complexity. Two of these methods are now widely used: `MC@NLO` [152] and `POWHEG` [153,40], with several alternative methods now also being pursued, see Ref. 146 and references therein.

NLO+PS generators should produce NLO accurate distributions for inclusive quantities, and should generate the hardest jet with tree-level accuracy even at large angle. It should be recalled, though, that in $2 \to 1$ processes like Z/W production, GPMCs including MEC and weighted by a constant K factor may perform nearly as well, and, if suitably tuned, may even yield a better description of data. It may thus be wise to consider tuning also the NLO+PS generators for these processes.

ME+PS generators should be preferred over NLO+PS ones when one needs an accurate description of hard, large-angle emissions, beyond the hardest jet. Attempts to merge ME+PS and NLO+PS, in order to get event samples that have the advantage of both methods have appeared, see Refs. 41, 146, and references therein.

Several ME+PS implementations use existing LO generators, like `ALPGEN` [60], `MADGRAPH` [59], and others summarized in Ref. 151, for the calculation of the matrix element, and feed the partonic events to a GPMC like `PYTHIA` or `HERWIG` using the Les Houches Interface for User Processes (LHI/LHEF) [44,30]. `SHERPA` and `HERWIG++` also include their own matrix-element generators and matching algorithms.

Several NLO+PS processes are implemented in the `MC@NLO` program [152], together with the new `AMC@NLO` development [86], and in the `POWHEG BOX` framework [40]. Again, `SHERPA` and `HERWIG++`, also include their own `POWHEG` implementation, suitably adapted with the inclusion of vetoed and truncated showers, for several processes.

38.2. Hadronization Models

In the context of event generators, *hadronization* denotes the process by which a set of colored partons (*after* showering) is transformed into a set of color-singlet *primary* hadrons, which may then subsequently decay further. This non-perturbative transition takes place at the *hadronization scale* $Q_{\rm had}$, which by construction is identical to the infrared cutoff of the parton shower. In the absence of a first-principles solution to the relevant dynamics, event generators use QCD-inspired phenomenological models to describe this transition.

A key difference between MC hadronization models and the fragmentation-function (FF) formalism used to describe inclusive hadron spectra in perturbative QCD (see Chap. 9 of PDG book) is that the former is always defined at the hadronization scale, while the latter can be defined at an arbitrary perturbative scale Q. They can therefore only be compared directly if the perturbative evolution between Q and $Q_{\rm had}$ is taken into account. FFs are calculable in pQCD, given a non-perturbative initial condition obtained by fits to hadron spectra. In the MC context, one can prove that the correct QCD evolution of the FFs arises from the shower formalism, with the hadronization model providing an explicit parametrization of the non-perturbative component. It should be kept in mind, however, that the MC modeling of shower and hadronization includes much more information on the final state since it is fully exclusive (i.e., it addresses all particles in the final state explicitly), while FFs only describe inclusive spectra. This exclusivity also enables MC models to make use of the color-flow information coming from the perturbative shower evolution (see Sec. 38.1.5) to determine between which partons the confining potentials should arise.

If one had an exact hadronization model, its dependence upon the hadronization scale $Q_{\rm had}$ would be compensated by the corresponding scale dependence of the shower algorithm, which stops generating branchings at the scale $Q_{\rm had}$. However, due to their complicated and fully exclusive nature, it is generally not possible to enforce this compensation automatically in MC models of hadronization. One must therefore be aware that the model must be "retuned" by hand if changes are made to the perturbative evolution, in particular if the infrared cutoff is modified. Tuning is discussed briefly in Sec. 38.4.

An important result in "quenched" lattice QCD (see Chap. 17 of PDG book) is that the potential of the color-dipole field between a charge and an anticharge appears to grow linearly with the separation of the charges, at distances greater than about a femtometer. This is known as "linear confinement", and it forms the starting point for the *string model of hadronization*, discussed below in Sec. 38.2.1. Alternatively, a property of perturbative QCD called "preconfinement" is the basis of the *cluster model of hadronization*, discussed in Sec. 38.2.2.

Finally, it should be emphasized that the so-called "parton level" that can be obtained by switching off hadronization in a GPMC, is not a universal concept, since each model defines the hadronization scale differently (e.g. by a cutoff in $p_\perp$, invariant mass, etc., with different tunes using different values for the cutoff). Comparisons to distributions at

this level may therefore be used to provide an idea of the overall impact of hadronization corrections within a given model, but should be avoided in the context of physical observables.

38.2.1. ***The String Model*** :
Starting from early concepts [46], several hadronization models based on strings have been proposed [146]. Of these, the most widely used today is the so-called Lund model [147,48], implemented in `PYTHIA` [139,140]. We concentrate on that particular model here, though many of the overall concepts would be shared by any string-inspired method.

Consider a color-connected quark-antiquark pair with no intermediate gluons emerging from the parton shower (like the $\bar{q}q$ pair in the center of Fig. 38.1), e.g. a red q and an antired $\bar{q}$. As the charges move apart, linear confinement implies that a potential $V(r) = \kappa\, r$ is reached for large distances r. (At short distances, there is a Coulomb term $\propto 1/r$ as well, but this is neglected in the Lund string.) This potential describes a string with tension $\kappa \sim 1\,\mathrm{GeV/fm} \sim 0.2\,\mathrm{GeV}^2$. The physical picture is that of a color flux tube being stretched between the q and the $\bar{q}$.

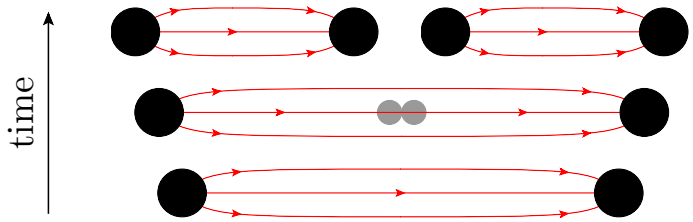


Figure 38.2: Illustration of string breaking by quark pair-creation in the string field.

As the string grows, the non-perturbative creation of quark-antiquark pairs can break the string, via the process $(q\bar{q}) \to (q\bar{q}') + (q'\bar{q})$, illustrated in Fig. 38.2. More complicated color-connected quark-antiquark configurations involving intermediate gluons (like the $\bar{q}gggq$ and $\bar{q}gq$ systems on the left and right part of Fig. 38.1) are treated by representing gluons as transverse "kinks." Thus soft gluons effectively build up a transverse structure in the originally one-dimensional object, with infinitely soft ones smoothly absorbed into the string. For strings with finite-energy kinks, the space-time evolution is slightly more involved [48], but the main point is that there are no separate free parameters for gluon jets. Differences with respect to quark fragmentation arise simply because quarks are only connected to a single string piece, while gluons have one on either side, increasing their relative energy loss (per unit invariant time) by a factor of 2, similar to the ratio of color Casimirs $C_A/C_F = 2.25$.

Since the string breaks are causally disconnected (as can be realized from space-time diagrams [48]) , they do not have to be considered in any specific time-ordered sequence. In the Lund model, the string breaks are generated starting with the leading ("outermost") hadrons, containing the endpoint quarks, and iterating inwards towards the center of the string, alternating randomly between the left and right sides. One can

thereby split off a single on-shell hadron in each step, making it straightforward to ensure that only states consistent with known hadron states are produced.

For each breakup vertex, quantum mechanical tunneling is assumed to control the masses and $p_\perp$ kicks that can be produced, leading to a Gaussian suppression

$$\text{Prob}(m_q^2, p_{\perp q}^2) \;\propto\; \exp\left(\frac{-\pi m_q^2}{\kappa}\right) \exp\left(\frac{-\pi p_{\perp q}^2}{\kappa}\right) \;, \tag{38.10}$$

where m_q is the mass of the produced quark flavor and $p_\perp$ is the non-perturbative transverse momentum imparted to it by the breakup process (the antiquark has the same mass and opposite $p_\perp$), with a universal average value of $\left\langle p_{\perp q}^2 \right\rangle = \kappa/\pi \sim (250\,\text{MeV})^2$. The charm and bottom masses are sufficiently heavy that they are not produced at all in the soft fragmentation. The transverse direction is defined with respect to the string axis, so the $p_\perp$ in a frame where the string is moving will be modified by a Lorentz boost. Note that the effective amount of "non-perturbative" $p_\perp$, in a Monte Carlo model with a fixed shower cutoff Q_{had}, may be larger than the purely non-perturbative κ/π above, to account for effects of additional unresolved soft-gluon radiation below Q_{had}. In principle, the magnitude of this additional component should scale with the cutoff, but in practice it is up to the user to enforce this by retuning the relevant parameter when changing the hadronization scale.

Since quark masses are difficult to define for light quarks, the value of the strangeness suppression is determined from experimental observables, such as the K/π and K^*/ρ ratios. The parton-shower evolution generates a small amount of strangeness as well, through perturbative $g \to s\bar{s}$ splittings. The optimal value for the non-perturbative $2s/(u+d)$ ratio should therefore exhibit a mild anticorrelation with the amount of quarks produced in the perturbative stage.

Baryon production can also be incorporated, by allowing string breaks to produce pairs of *diquarks*, loosely bound states of two quarks in an overall $\bar{3}$ representation. Again, since diquark masses are difficult to define, the relative rate of diquark to quark production is extracted, e.g. from the p/π ratio, and since the perturbative shower splittings do not produce diquarks, the effective value for this parameter is mildly correlated with the amount of $g \to q\bar{q}$ splittings occurring on the shower side. More advanced scenarios for baryon production have also been proposed, see Ref. 48. Within the `PYTHIA` framework, a fragmentation model including baryon string junctions [49] is also available.

The next step of the algorithm is the assignment of the produced quarks within hadron multiplets. Using a nonrelativistic classification of spin states, the fragmenting q may combine with the $\bar{q}'$ from a newly created breakup to produce a meson — or baryon, if diquarks are involved — of a given valence quark spin S and angular momentum L. The lowest-lying pseudoscalar and vector meson multiplets, and spin-1/2 and -3/2 baryons, are assumed to dominate in a string framework[1], but individual rates are not predicted by the model. This is therefore the sector that contains the largest amount of free parameters.

[1] The `PYTHIA` implementation includes the lightest pseudoscalar and vector mesons,

From spin counting, the ratio V/P of vectors to pseudoscalars is expected to be 3, but in practice this is only approximately true for B mesons. For lighter flavors, the difference in phase space caused by the V–P mass splittings implies a suppression of vector production. When extracting the corresponding parameters from data, it is advisable to begin with the heaviest states, since so-called feed-down from the decays of higher-lying hadron states complicates the extraction for lighter particles, see Sec. 38.2.3. For diquarks, separate parameters control the relative rates of spin-1 diquarks vs. spin-0 ones and, likewise, have to be extracted from data.

With $p_\perp^2$ and m^2 now fixed, the final step is to select the fraction, z, of the fragmenting endpoint quark's longitudinal momentum that is carried by the created hadron, an aspect for which the string model is highly predictive. The requirement that the fragmentation be independent of the sequence in which breakups are considered (causality) imposes a "left-right symmetry" on the possible form of the fragmentation function, $f(z)$, with the solution

$$f(z) \propto \frac{1}{z}(1-z)^a \exp\left(-\frac{b\,(m_h^2 + p_{\perp h}^2)}{z}\right) \quad , \tag{38.11}$$

which is known as the Lund symmetric fragmentation function (normalized to unit integral). The dimensionless parameter a dampens the hard tail of the fragmentation function, towards $z \to 1$, and may in principle be flavor-dependent, while b, with dimension GeV^{-2}, is a universal constant related to the string tension [48] which determines the behavior in the soft limit, $z \to 0$. Note that the explicit mass dependence in $f(z)$ implies a harder fragmentation function for heavier hadrons (in the rest frame of the string).

As a by-product, the probability distribution in invariant time τ of $q'\bar{q}$ breakup vertices, or equivalently $\Gamma = (\kappa\tau)^2$, is also obtained, with $dP/d\Gamma \propto \Gamma^a \exp(-b\Gamma)$ implying an area law for the color flux, and the average breakup time lying along a hyperbola of constant invariant time $\tau_0 \sim 10^{-23}$s [48].

For massive endpoints (e.g. c and b quarks, or hypothetical hadronizing new-physics particles), which do not move along straight lightcone sections, the exponential suppression with string area leads to modifications of the form $f(z) \to f(z)/z^{b\,m_Q^2}$, with m_Q the mass of the heavy quark [50]. Although different forms can also be used to describe inclusive heavy-meson spectra (see Sec 19.9 of PDG book), such choices are not consistent with causality in the string framework and hence are theoretically disfavored in this context, one well-known example being the Peterson formula [51],

$$f(z) \propto \frac{1}{z}\left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right)^{-2} \quad , \tag{38.12}$$

with ϵ_Q a free parameter expected to scale $\propto 1/m_Q^2$.

with the four $L = 1$ multiplets (scalar, tensor, and 2 pseudovectors) available but disabled by default, largely because several states are poorly known and thus may result in a worse overall description when included. For baryons, the lightest spin-1/2 and -3/2 multiplets are included.

38.2.2. ***The Cluster Model*** :
The cluster hadronization model is based on *preconfinement*, i.e., on the observation [52,53] that the color structure of a perturbative QCD shower evolution at any scale Q_0 is such that color-singlet subsystems of partons (labeled "clusters") occur with a universal invariant mass distribution that only depends on Q_0 and on $\Lambda_{\rm QCD}$, not on the starting scale Q, for $Q \gg Q_0 \gg \Lambda_{\rm QCD}$. Further, this mass distribution is power-suppressed at large masses.

Following early models based on this universality [8,54], the cluster model developed by Webber [141] has for many years been a hallmark of the `HERWIG` and `HERWIG++` generators, with an alternative implementation [56] now available in the `SHERPA` generator. The key idea, in addition to preconfinement, is to force "by hand" all gluons to split into quark-antiquark pairs at the end of the parton shower. Compared with the string description, this effectively amounts to viewing gluons as "seeds" for string breaks, rather than as kinks in a continuous object. After the splittings, a new set of low-mass color-singlet clusters is obtained, formed only by quark-antiquark pairs. These can be decayed to on-shell hadrons in a simple manner.

The algorithm starts by generating the forced $g \to q\bar{q}$ breakups, and by assigning flavors and momenta to the produced quark pairs. For a typical shower cutoff corresponding to a gluon virtuality of $Q_{\rm had} \sim 1\,\text{GeV}$, the $p_\perp$ generated by the splittings can be neglected. The constituent light-quark masses, $m_{u,d} \sim 300\,\text{MeV}$ and $m_s \sim 450\,\text{MeV}$, imply a suppression (typically even an absence) of strangeness production. In principle, the model also allows for diquarks to be produced at this stage, but due to the larger constituent masses this would only become relevant for shower cutoffs larger than 1 GeV.

If a cluster formed in this way has an invariant mass above some cutoff value, typically 3–4 GeV, it is forced to undergo sequential $1 \to 2$ cluster breakups, along an axis defined by the constituent partons of the original cluster, until all sub-cluster masses fall below the cutoff value. Due to the preservation of the original axis in these breakups, this treatment has some resemblance to the string-like picture.

Next, on the low-mass side of the spectrum, some clusters are allowed to decay directly to a single hadron, with nearby clusters absorbing any excess momentum. This improves the description of the high-z part of the fragmentation spectrum — where the hadron carries almost all the momentum of its parent jet — at the cost of introducing one additional parameter, controlling the probability for single-hadron cluster decay.

Having obtained a final distribution of small-mass clusters, now with a strict cutoff at 3–4 GeV and with the component destined to decay to single hadrons already removed, the remaining clusters are interpreted as a smoothed-out spectrum of excited mesons, each of which decays isotropically to two hadrons, with relative probabilities proportional to the available phase space for each possible two-hadron combination that is consistent with the cluster's internal flavors, including spin degeneracy. It is important that all the light members (containing only *uds*) of each hadron multiplet be included, as the absence of members can lead to unphysical isospin or SU(3) flavor violation. Typically, the lightest pseudoscalar, vector, scalar, even and odd charge conjugation pseudovector, and tensor multiplets of light mesons are included. In addition, some excited vector multiplets of light mesons may be available. For baryons, usually only the lightest octet, decuplet and

singlet baryons are present, although both the HERWIG++ and SHERPA implementations now include some heavier baryon multiplets as well.

Contrary to the case in the string model, the mechanism of phase-space suppression employed here leads to a natural enhancement of the lighter pseudoscalars, and no parameters beyond the spectrum of hadron masses need to be introduced at this point. The phase space also limits the transverse momenta of the produced hadrons relative to the jet axis.

Note that, since the masses and decays of excited heavy-flavor hadrons in particular are not well known, there is some freedom in the model to adjust these, which in turn will affect their relative phase-space populations.

38.2.3. *Hadron and τ Decays* :

Of the so-called primary hadrons, originating directly from string breaks and/or cluster decays (see above), many are unstable and so decay further, until a set of particles is obtained that can be considered stable on time scales relevant to the given measurement[2]. The decay modeling can therefore have a significant impact on final particle yields and spectra, especially for the lowest-lying hadronic states, which receive the largest relative contributions from decays (feed-down). Note that the interplay between primary production and feed-down implies that the hadronization parameters should be retuned if significant changes to the decay treatment are made.

Particle summary tables, such as those given elsewhere in this *Review*, represent a condensed summary of the available experimental measurements and hence may be incomplete and/or exhibit inconsistencies within the experimental precision. In an MC decay package, on the other hand, all information must be quantified and consistent, with all branching ratios summing to unity. When adapting particle summary information for use in a decay package, a number of choices must therefore be made. The amount of ambiguity increases as more excited hadron multiplets are added to the simulation, about which less and less is known from experiment, with each GPMC making its own choices.

A related choice is how to distribute the decay products differentially in phase space, in particular which matrix elements to use. Historically, MC generators contained matrix elements only for selected (generator-specific) classes of hadron and τ decays, coupled with a Breit-Wigner smearing of the masses, truncated at the edges of the physical decay phase space (the treatment of decay thresholds can be important for certain modes [146]) . A more sophisticated treatment can then be obtained by reweighting the generated events using the obtained particle four-momenta and/or by using specialized external packages such as EVTGEN [57] for hadron decays and TAUOLA [58] for τ decays.

More recently, HERWIG++ and SHERPA include helicity-dependence in τ decays [59,144], with a more limited treatment available in PYTHIA 8 [140]. The HERWIG++ and SHERPA generators have also included significantly improved internal simulations of hadronic decays, which include spin correlations between those decays for which matrix elements are used.

[2] E.g., a typical hadron-collider definition of a "stable particle" is $c\tau \geq 10\,\text{mm}$, which includes the weakly-decaying strange hadrons (K, Λ, $\Sigma^{\pm}$, $\bar{\Sigma}^{\pm}$, Ξ, Ω).

HERWIG++ and PYTHIA include the probability for B mesons to oscillate into $\bar{B}$ ones before decay. SHERPA and EVTGEN also include CP-violating effects and, for common decay modes of the neutral meson and its antiparticle, the interference between the direct decay and oscillation followed by decay.

We end on a note of warning on double counting. This may occur if a particle can decay via an intermediate on-shell resonance. An example is $a_1 \to \pi\pi\pi$ which may proceed via $a_1 \to \rho\pi$, $\rho \to \pi\pi$. If these decay channels of the a_1 are both included, each with their full partial width, a double counting of the on-shell $a_1 \to \rho\pi$ contribution would result. Such cases are normally dealt with consistently in the default MC generator packages, so this warning is mostly relevant for users that wish to edit decay tables on their own.

38.3. Models for Soft Hadron-Hadron Physics

38.3.1. ***Minimum-Bias and Diffraction*** :
The term "minimum bias" (MB) originates from the experimental requirement of a minimal number of tracks (or hits) in a given instrumented region. In order to make MC predictions for such observables, all possible contributions to the relevant phase-space region must be accounted for. There are essentially four types of physics processes, which together make up the total hadron-hadron (hh) cross section: 1) elastic scattering[3]: $hh \to hh$, 2) single diffractive dissociation: $hh \to h + \text{gap} + X$, with X denoting anything that is not the original beam particle, and "gap" denoting a rapidity region devoid of observed activity; 3) double diffractive dissociation: $hh \to X + \text{gap} + X$, and 4) inelastic non-diffractive scattering: everything else. A fifth class may also be defined, called central diffraction ($hh \to h + \text{gap} + X + \text{gap} + h$). Some differences exist between theoretical and experimental terminology [61]. In the experimental setting, diffraction is defined by an observable gap, of some minimal size in rapidity. In the MC context, each diffractive physics process typically produces a whole spectrum of gaps, with small ones suppressed but not excluded.

The inelastic non-diffractive part of the cross section is typically modeled either by smoothly regulating and extending the perturbative QCD scattering cross sections all the way to zero $p_\perp$ [149] (PYTHIA 6, PYTHIA 8, and SHERPA), or by regulating the QCD cross sections with a sharp cutoff [63](HERWIG+JIMMY) and adding a separate class of intrinsically soft scatterings below that scale [64](HERWIG++). See also Sec. 38.3.2. In all cases, the three most important ingredients are: 1) the IR regularization of the perturbative scattering cross sections, including their PDF dependence, 2) the assumed matter distribution of the colliding hadrons, possibly including multi-parton correlations [49] and/or x dependence [65], and 3) additional soft-QCD effects such as color reconnections and/or other collective effects, discussed in Sec. 38.3.3.

Currently, there are essentially three methods for simulating diffraction in the main MC models: 1) in PYTHIA 6, one picks a diffractive mass according to parametrized cross

[3] The QED elastic-scattering cross section diverges and is normally a non-default option in MC models.

sections $\propto \mathrm{d}M^2/M^2$ [66]. This mass is represented as a string, which is fragmented as described in Sec. 38.2.1, though differences in the effective scale of the hadronization may necessitate a (re)tuning of the fragmentation parameters for diffraction; 2) in `PYTHIA` 8, the high-mass tail beyond $M \sim 10\,\mathrm{GeV}$ is augmented by a partonic description in terms of pomeron PDFs [67], allowing diffractive jet production including showers and underlying event [68]; 3) the `PHOJET` and `DPMJET` programs also include central diffraction and rely directly on a formulation in terms of pomerons (color-singlet multi-gluon states) [69–71]. Cut pomerons correspond to exchanges of soft gluons while uncut ones give elastic and diffractive topologies as well as virtual corrections that help preserve unitarity. So-called "hard pomerons" provide a transition to the perturbative regime. Fragmentation is still handled using the Lund string model, so there is some overlap with the above models at the hadronization stage. In addition, a pomeron-based package exists for `HERWIG` [72], and an effort is underway to construct an MC implementation of the "KMR" model [73] within the `SHERPA` generator. Color reconnections (Sec. 38.3.3) may also play a role in creating rapidity gaps and the underlying event (Sec. 38.3.2) in destroying them.

38.3.2. *Underlying Event and Jet Pedestals* :

In the event-generator context, the term underlying event (UE) denotes any additional activity *beyond* the basic process and its associated ISR and FSR activity. The dominant contribution to this is believed to come from additional color exchanges between the beam particles, which can be represented either as multiple parton-parton interactions (MPI) or as so-called cut pomerons (Sec. 38.3.1). The experimentally observed fact that the UE is more active than MB events at the same CM energy is called the "jet pedestal" effect.

The most easily identifiable consequence of MPI is arguably the possibility of observing several hard parton-parton interactions in one and the same hadron-hadron event. This tends to produce largely uncorrelated back-to-back jet pairs, with each pair having a small value of $\mathrm{sum}(\vec{p}_\perp)$. For comparison, jets from bremsstrahlung tend to be aligned with the direction of their parent initial- or final-state partons. The fraction of MPI that give rise to additional reconstructible jets is, however, quite small. Soft interactions that do not give rise to observable jets are much more plentiful, and can give significant corrections to the color flow and total scattered energy of the event. This affects the final-state activity in a more global way, increasing multiplicity and summed E_T distributions, and contributing to the break-up of the beam remnants in the forward direction.

The first detailed Monte Carlo model for perturbative MPI was proposed in Ref. 149, and with some variation this still forms the basis for most modern implementations. Some useful additional references can be found in Ref. 146. The first crucial observation is that the t-channel propagators appearing in perturbative QCD $2 \to 2$ scattering almost go on shell at low $p_\perp$, causing the differential cross sections to become very large, behaving roughly as

$$d\sigma_{2\to 2} \propto \frac{dt}{t^2} \sim \frac{dp_\perp^2}{p_\perp^4} \ . \tag{38.13}$$

This cross section is an inclusive number. Thus, if a single hadron-hadron event contains two parton-parton interactions, it will "count" twice in $\sigma_{2\to 2}$ but only once in σ_{tot}, and so on. In the limit that all the interactions are independent and equivalent, one would

have

$$\sigma_{2\to 2}(p_{\perp\min}) = \langle n\rangle(p_{\perp\min})\ \sigma_{\rm tot}\ , \tag{38.14}$$

with $\langle n\rangle(p_{\perp\min})$ giving the average of a Poisson distribution in the number of parton-parton interactions above $p_{\perp\min}$ per hadron-hadron collision,

$$\mathcal{P}_n(p_{\perp\min}) = (\langle n\rangle(p_{\perp\min}))^n\ \frac{\exp\left(-\langle n\rangle\ (p_{\perp\min})\right)}{n!}\ . \tag{38.15}$$

This simple argument in fact expresses unitarity; instead of the total interaction cross section diverging as $p_{\perp\min} \to 0$ (which would violate unitarity), we have restated the problem so that it is now the *number of MPI per collision* that diverges, with the total cross section remaining finite. At LHC energies, the $2 \to 2$ scattering cross sections computed using the full LO QCD cross section folded with modern PDFs becomes larger than the total pp one for $p_\perp$ values of order 4–5 GeV [74]. One therefore expects the average number of perturbative MPI to exceed unity at around that scale.

Two important ingredients remain to fully regulate the remaining divergence. Firstly, the interactions cannot use up more momentum than is available in the parent hadron. This suppresses the large-n tail of the estimate above. In `PYTHIA`-based models, the MPI are ordered in $p_\perp$, and the parton densities for each successive interaction are explicitly constructed so that the sum of x fractions can never be greater than unity. In the `HERWIG` models, instead the uncorrelated estimate of $\langle n\rangle$ above is used as an initial guess, but the generation of actual MPI is stopped once the energy-momentum conservation limit is reached.

The second ingredient invoked to suppress the number of interactions, at low $p_\perp$ and x, is color screening; if the wavelength $\sim 1/p_\perp$ of an exchanged colored parton becomes larger than a typical color-anticolor separation distance, it will only see an *average* color charge that vanishes in the limit $p_\perp \to 0$, hence leading to suppressed interactions. This provides an infrared cutoff for MPI similar to that provided by the hadronization scale for parton showers. A first estimate of the color-screening cutoff would be the proton size, $p_{\perp\min} \approx \hbar/r_p \approx 0.3\,\text{GeV} \approx \Lambda_{\rm QCD}$, but empirically this appears to be far too low. In current models, one replaces the proton radius r_p in the above formula by a "typical color screening distance," i.e., an average size of a region within which the net compensation of a given color charge occurs. This number is not known from first principles [73] and is perceived of simply as an effective cutoff parameter. The simplest choice is to introduce a step function $\Theta(p_\perp - p_{\perp\min})$. Alternatively, one may note that the jet cross section is divergent like $\alpha_{\rm S}^2(p_\perp^2)/p_\perp^4$, cf. Eq. (38.13), and that therefore a factor

$$\frac{\alpha_{\rm S}^2(p_{\perp 0}^2 + p_\perp^2)}{\alpha_{\rm S}^2(p_\perp^2)}\ \frac{p_\perp^4}{(p_{\perp 0}^2 + p_\perp^2)^2} \tag{38.16}$$

would smoothly regulate the divergences, now with $p_{\perp 0}$ as the free parameter. Regardless of whether it is imposed as a smooth (`PYTHIA` and `SHERPA`) or steep (`HERWIG++`) function, this is effectively the main "tuning" parameter in such models.

Note that the numerical value obtained for the cross section depends upon the PDF set used, and therefore the optimal value to use for the cutoff will also depend on this choice.

Note also that the cutoff does not have to be energy-independent. Higher energies imply that parton densities can be probed at smaller x values, where the number of partons rapidly increases. Partons then become closer packed and the color screening distance d decreases. The uncertainty on the energy and/or x scaling of the cutoff is a major concern when extrapolating between different collider energies [75].

We now turn to the origin of the observational fact that hard jets appear to sit on top of a higher "pedestal" of underlying activity than events with no hard jets. This is interpreted as a consequence of impact-parameter-dependence: in peripheral collisions, only a small fraction of events contain any high-$p_\perp$ activity, whereas central collisions are more likely to contain at least one hard scattering; a high-$p_\perp$ triggered sample will therefore be biased towards small impact parameters, b. The ability of a model to describe the shape of the pedestal (e.g. to describe both MB and UE distributions simultaneously) therefore depends upon its modeling of the b-dependence, and correspondingly the impact-parameter shape constitutes another main tuning parameter.

For each impact parameter b, the number of interactions $\tilde{n}(b)$ can then still be assumed to be distributed according to Eq. (38.15), again modulo momentum conservation, but now with the mean value of the Poisson distribution depending on impact parameter, $\langle\tilde{n}(b)\rangle$. This causes the final n-distribution (integrated over b) to be wider than a Poissonian.

Finally, there are two perturbative modeling aspects which go beyond the introduction of MPI themselves: 1) parton showers off the MPI, and 2) perturbative parton-rescattering effects. Without showers, MPI models would generate very sharp peaks for back-to-back MPI jets, caused by unshowered partons passed directly to the hadronization model. However, with the exception of the oldest `PYTHIA6` model, all GPMC models do include such showers [146], and hence should exhibit more realistic (i.e., broader and more decorrelated) MPI jets. On the initial-state side, the main questions are whether and how correlated multi-parton densities are taken into account and, as discussed previously, how the showers are regulated at low $p_\perp$ and/or low x. Although none of the MC models currently impose a rigorous correlated multi-parton evolution, all of them include some elementary aspects. The most significant for parton-level results is arguably momentum conservation, which is enforced explicitly in all the models. The so-called "interleaved" models [24] attempt to go a step further, generating an explicitly correlated multi-parton evolution in which flavor sum rules are imposed to conserve, e.g. the total numbers of valence and sea quarks [49].

Perturbative rescattering in the final state can occur if partons are allowed to undergo several distinct interactions, with showering activity possibly taking place in-between. This has so far not been studied extensively, but a first exploratory model is available [76]. In the initial state, parton rescattering/recombination effects have so far not been included in any of the GPMC models.

38.3.3. ***Bose-Einstein and Color-Reconnection Effects*** :
In the context of e^+e^- collisions, Bose-Einstein (BE) correlations have mostly been discussed as a source of uncertainty on high-precision W mass determinations at LEP [77]. In hadron-hadron (and nucleus-nucleus) collisions, however, BE correlations are used extensively to study the space-time structure of hadronizing matter ("femtoscopy").

In MC models of hadronization, each string break and/or particle/cluster decay is normally factorized from all other ones. This reduces the number of variables that must be considered simultaneously, but also makes the introduction of correlations among particles from different breaks/decays intrinsically difficult to address. In the context of GPMCs, a few semi-classical models are available within the `PYTHIA` 6 and 8 generators [78], in which the BE effect is mimicked by an attractive interaction between pairs of identical particles in the final state, with no higher correlations included. This "force" acts after the decays of very short-lived particles, like ρ, but before decays of longer-lived ones, like π^0. The main differences between the variants of this model is the assumed shape of the correlation function and how overall momentum conservation is handled.

As discussed in Sec. 38.2, leading-color ("planar") color flows are used to set up the hadronizing systems (clusters or strings) at the hadronization stage. If the systems do not overlap significantly in space and time, subleading-color ambiguities and/or non-perturbative reconnections are expected to be small. However, if the density of displaced color charges is sufficiently high that several systems can overlap significantly, full-color and/or reconnection effects should become progressively larger.

In the specific context of MPI, a crucial question is how color is neutralized *between* different MPI systems, including the remnants. The large rapidity differences involved imply large invariant masses (though normally low $p_\perp$), and hence large amounts of (soft) particle production. Indeed, in the context of soft-inclusive physics, it is these "inter-system" strings/clusters that furnish the dominant particle-production mechanism, and hence their modeling is an essential part of the soft-physics description, affecting topics such as MB/UE multiplicity and $p_\perp$ distributions, rapidity gaps, and precision mass measurements. A more comprehensive review of color-reconnection effects can be found in Ref. 146.

38.4. Parameters and Tuning

The achievable accuracy in GPMC models depends both on the inclusiveness of the chosen observable and on the sophistication of the simulation. An important driver for the latter is obviously the development of improved theoretical models, discussed in the preceding sections; but it also depends crucially on the available constraints on the remaining free parameters. Using existing data to constrain these is referred to as generator tuning.

Although MC models may appear to have a bewildering array of adjustable parameters, most of them only control relatively small (exclusive) details of the event generation. The majority of the (inclusive) physics is determined by only a few, very important ones, such as the value of α_S, in the perturbative domain, and the properties of the non-perturbative fragmentation functions, in the non-perturbative one. One may

therefore take a factorized approach, first constraining the perturbative parameters and thereafter the non-perturbative ones, each ordered in a measure of their relative significance to the overall modeling.

At LO×LL, perturbation theory is doing well if it agrees with an IR safe measurement within 10%. It would therefore not make much sense to tune a GPMC beyond roughly 5% (it might even be dangerous, due to overfitting). The advent of NLO Monte Carlos may reduce this number slightly, but only for quantities for which one expects NLO precision. For LO Monte Carlos, distributions should be normalized to unity, since the NLO normalization is not tunable. For quantities governed by non-perturbative physics, uncertainties are larger. For some quantities, e.g. ones for which the underlying modeling is known to be poor, an order-of-magnitude agreement or worse may have to be accepted.

In the context of LO×LL GPMC tuning, subleading aspects of coupling-constant and PDF choices are relevant. In particular, one should be aware that the choice of QCD Λ parameter $\Lambda_{\rm MC} = 1.569\Lambda_{\overline{\rm MS}}$ (for 5 active flavors) improves the predictions of coherent shower algorithms at the NLL level [79], and hence this scheme is typically considered the baseline for shower tuning. The question of LO vs. NLO PDFs is more involved [146], but it should be emphasized that the low-x gluon in particular is important for determining the level of the underlying event in MPI models (Sec. 38.3.2), and hence the MB/UE tuning (and energy scaling [75]) is linked to the choice of PDF in such models. Further issues and an example of a specific recipe that could be followed in a realistic set-up can be found in Ref. 80.

Recent years have seen the emergence of automated tools that attempt to reduce the amount of both computer and manpower required for tuning [81]. Automating the human expert input is more difficult. In the tools currently on the market, this question is addressed by a combination of input solicited from the generator authors (e.g., which parameters and ranges to consider, which observables constitute a complete set, etc) and a set of weights determining the relative priority given to each bin in each distribution. The field is still burgeoning, however, and future sophistications are to be expected. Nevertheless, the overall quality of the automated tunes appear to at least be competitive with the manual ones.

References:

1. G. Corcella *et al.*, JHEP **0101**, 010 (2001), `hep-ph/0011363`.
2. M.Bähr *et al.*, Eur. Phys. J. **C58**, 639 (2008), `arXiv:0803.0883`.
3. T. Sjöstrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), `hep-ph/0603175`.
4. T. Sjöstrand, S. Mrenna, and P. Z. Skands, Comp. Phys. Comm. **178**, 852 (2008), `arXiv:0710.3820`.
5. T. Gleisberg *et al.*, JHEP **0402**, 056 (2004), `hep-ph/0311263`.
6. T. Kinoshita, J. Math. Phys. **3**, 650 (1962).
7. T. Lee and M. Nauenberg, Phys. Rev. **133**, 1549 (1964).
8. G.C. Fox and S. Wolfram, Nucl. Phys. **B168**, 285 (1980).
9. G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
10. A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
11. B.R. Webber, Phys. Lett. **B193**, 91 (1987).

12. J.C. Collins, Nucl. Phys. **B304**, 794 (1988).
13. I.G. Knowles, Comp. Phys. Comm. **58**, 271 (1990).
14. T. Sjöstrand, Phys. Lett. **B157**, 321 (1985).
15. A. Buckley *et al.*, Phys. Reports **504**, 145 (2011), arXiv:1101.2599.
16. G. Marchesini and B.R. Webber, Nucl. Phys. **B310**, 461 (1988).
17. S. Gieseke, P. Stephens, and B. Webber, JHEP **0312**, 045 (2003), hep-ph/0310083.
18. M. Bengtsson and T. Sjöstrand, Nucl. Phys. **B289**, 810 (1987).
19. G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).
20. L. Lönnblad, Comp. Phys. Comm. **71**, 15 (1992).
21. W.T. Giele, D.A. Kosower, and P.Z. Skands, Phys. Rev. **D78**, 014026 (2008), arXiv:0707.3652.
22. S. Schumann and F. Krauss, JHEP **0803**, 038 (2008), arXiv:0709.1027.
23. Z. Nagy and D.E. Soper, JHEP **0510**, 024 (2005), hep-ph/0503053.
24. T. Sjöstrand and P.Z. Skands, Eur. Phys. J. **C39**, 129 (2005), hep-ph/0408302.
25. E. Norrbin and T. Sjöstrand, Nucl. Phys. **B603**, 297 (2001), hep-ph/0010012.
26. S. Catani *et al.*, Nucl. Phys. **B627**, 189 (2002), hep-ph/0201036.
27. A. Semenov, Comp. Phys. Comm. **180**, 431 (2009), arXiv:0805.0555.
28. N.D. Christensen and C. Duhr, Comp. Phys. Comm. **180**, 1614 (2009), arXiv:0806.4194.
29. M. Fairbairn *et al.*, Phys. Reports **438**, 1 (2007), hep-ph/0611040.
30. J. Alwall *et al.*, Comp. Phys. Comm. **176**, 300 (2007), hep-ph/0609017.
31. P.Z. Skands *et al.*, JHEP **0407**, 036 (2004), hep-ph/0311123.
32. J. Alwall *et al.*, (2007), arXiv:0712.3311.
33. P. Richardson, JHEP **0111**, 029 (2001), hep-ph/0110108.
34. M. Bengtsson and T. Sjöstrand, Phys. Lett. **B185**, 435 (1987).
35. S. Catani *et al.*, JHEP **11**, 063 (2001), hep-ph/0109231.
36. J. Alwall *et al.*, Eur. Phys. J. **C53**, 473 (2008), arXiv:0706.2569.
37. P. Nason, JHEP **11**, 040 (2004), hep-ph/0409146.
38. B. Cooper *et al.*, (2011), arXiv:1109.5295.
39. S. Frixione and B.R. Webber, JHEP **06**, 029 (2002), hep-ph/0204244.
40. S. Alioli *et al.*, JHEP **1006**, 043 (2010), arXiv:1002.2581.
41. S. Alioli, K. Hamilton, and E. Re, (2001), arXiv:1108.0909.
42. M.L. Mangano *et al.*, JHEP **0307**, 001 (2003), hep-ph/0206293.
43. J. Alwall *et al.*, JHEP **1106**, 128 (2011), arXiv:1106.0522.
44. E. Boos *et al.*, (2007), hep-ph/0109068.
45. V. Hirschi *et al.*, JHEP **1105**, 044 (2011), arXiv:1103.0621.
46. X. Artru and G. Mennessier, Nucl. Phys. **B70**, 93 (1974).
47. B. Andersson *et al.*, Phys. Reports **97**, 31 (1983).
48. B. Andersson, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **7** (1997).
49. T. Sjöstrand and P.Z. Skands, JHEP **0403**, 053 (2004), hep-ph/0402078.
50. M. Bowler, Z. Phys. **C11**, 169 (1981).
51. C. Peterson *et al.*, Phys. Rev. **D27**, 105 (1983).
52. D. Amati and G. Veneziano, Phys. Lett. **B83**, 87 (1979).
53. A. Bassetto, M. Ciafaloni, and G. Marchesini, Phys. Lett. **B83**, 207 (1979).

54. R.D. Field and S. Wolfram, Nucl. Phys. **B213**, 65 (1983).
55. B.R. Webber, Nucl. Phys. **B238**, 492 (1984).
56. J.-C. Winter, F. Krauss, and G. Soff, Eur. Phys. J. **C36**, 381 (2004), hep-ph/0311085.
57. D. Lange, Nucl. Instrum. Methods **A462**, 152 (2001).
58. S. Jadach *et al.*, Comp. Phys. Comm. **76**, 361 (1993).
59. D. Grellscheid and P. Richardson, (2007), arXiv:0710.1951.
60. T. Gleisberg *et al.*, JHEP **0902**, 007 (2009), arXiv:0811.4622.
61. V. Khoze *et al.*, Eur. Phys. J. **C69**, 85 (2010), arXiv:1005.4839.
62. T. Sjöstrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).
63. J.M. Butterworth, J.R. Forshaw, and M.H. Seymour, Z. Phys. **C72**, 637 (1996), hep-ph/9601371.
64. M. Bähr *et al.*, (2009), arXiv:0905.4671.
65. R. Corke and T. Sjöstrand, JHEP **1105**, 009 (2011), 1101.5953.
66. G.A. Schuler and T. Sjöstrand, Phys. Rev. **D49**, 2257 (1994).
67. G. Ingelman and P. Schlein, Phys. Lett. **B152**, 256 (1985).
68. S. Navin, (2010), arXiv:1005.3894.
69. P. Aurenche *et al.*, Comp. Phys. Comm. **83**, 107 (1994), hep-ph/9402351.
70. F.W. Bopp, R. Engel, and J. Ranft, (1998), hep-ph/9803437.
71. S. Roesler, R. Engel, and J. Ranft, p. 1033 (2000), hep-ph/0012252.
72. B.E. Cox and J.R. Forshaw, Comp. Phys. Comm. **144**, 104 (2002), hep-ph/0010303.
73. M. Ryskin, A. Martin, and V. Khoze, Eur. Phys. J. **C71**, 1617 (2011), arXiv:1102.2844.
74. M. Bähr, J.M. Butterworth, and M.H. Seymour, JHEP **01**, 065 (2009), arXiv:0806.2949.
75. H. Schulz and P.Z. Skands, Eur. Phys. J. **C71**, 1644 (2011), arXiv:1103.3649.
76. R. Corke and T. Sjöstrand, JHEP **01**, 035 (2009), arXiv:0911.1901.
77. LEP Electroweak Working Group, (2005), hep-ex/0511027.
78. L. Lönnblad and T. Sjöstrand, Eur. Phys. J. **C2**, 165 (1998), hep-ph/9711460.
79. S. Catani, B. R. Webber, and G. Marchesini, Nucl. Phys. **B349**, 635 (1991).
80. P.Z. Skands, (2011), arXiv:1104.2863.
81. A. Buckley *et al.*, Eur. Phys. J. **C65**, 331 (2010), arXiv:0907.2973.