11 Beam Remnants and Underlying Events

Each incoming beam particle may leave behind a beam remnant, which does not take active part in the initial-state radiation or hard-scattering process. If nothing else, these remnants need to be put together and colour connected to the rest of the event. In addition, in hadron–hadron collisions, the composite nature of the two incoming hadrons implies the possibility that several pairs of partons can enter into separate but simultaneous scatterings, 'multiple interactions'. In some fraction of events, these additional scatterings can be hard or semi-hard, but due to the infrared peaking of the cross section the bulk of them should normally be fairly soft compared to the primary interaction. This extra component gives a non-negligible contribution to the 'underlying event' structure, and thus to the total multiplicity. Finally, in high-luminosity colliders, it is possible to have several collisions between beam particles in one and the same beam crossing, i.e. pile-up events, which further act to build up the general particle production activity that is to be observed by detectors. These three aspects are described in turn, with emphasis on the middle one, that of multiple interactions within a single hadron–hadron collision.

Work is ongoing to improve the theoretical framework for multiple interactions. Pythia therefore contains two partly separate approaches to this physics, which also leads to two partly separate descriptions of beam remnants. To make matters worse, from a pedagogical point of view, also the description of initial- and final-state radiation has been improved, with the transition from virtuality-ordered to transverse-momentum-ordered evolution, and with the interleaving of multiple interactions and initial-state emissions in one common $p_\perp$-ordered sequence. In total we therefore need to distinguish three main scenarios.

- The 'old model', based on [Sjö87a], contains the old beam-remnant and showering machineries. It remains the default while the new one is being developed, for backwards reference and comparisons. For instance, the 'Tune A' [Fie02] incarnation of this model is commonly used at the Tevatron, and offers a convenient reference against which other models can be compared, without the need to know how to do a full detector simulation. Furthermore, the newer models below are mainly being developed for pp and $\mathrm{p\overline{p}}$ collisions so far, so have barely been tested with meson and fixed-energy resolved-photon beams, and is not at all integrated into a complete consistent set of photon interactions for a spectum of incoming energies and virtualities.
- The 'intermediate model', developed in [Sjö04], attempts a more sophisticated description of correlations in flavour, colour, longitudinal and (primordial) transverse momenta between the beam remnants and the shower initiators than offered by the old one, and also introduces some other improvements. The virtuality-ordered showers are the same as in the old model, but now each interaction can be associated with its own showering activity, while before only the hardest interaction would be showered, for technical reasons. In practice, we do not expect much usage of the intermediate model: either people stay with the old or move right to the new one below. We have kept it mainly so that results from our publications can be reproduced.
- The 'new model', with further aspects described in [Sjö04a], where transverse-momentum-ordered showers are introduced and multiple interactions and initial-state radiation are interleaved. The beam-remnant description introduced in the intermediate model stays essentially unchanged.

Several of the key multiple interactions aspects are common between the models, such as the generation of kinematics of the multiple (semi)hard interactions and the impact-parameter picture. Other aspects are addressed in a more or less similar spirit, but in different ways, such as the desire to reduce the string length relative to naïve expectations. Therefore the current section starts with a description of the old model, and thereafter

outlines where the intermediate and new models differ.

11.1 Beam Remnants — Old Model

The initial-state radiation algorithm reconstructs one shower initiator in each beam. (If initial-state radiation is not included, the initiator is nothing but the incoming parton to the hard interaction.) Together the two initiators delineate an interaction subsystem, which contains all the partons that participate in the initial-state showers, in the hard interaction, and in the final-state showers. Left behind are two beam remnants which, to first approximation, just sail through, unaffected by the hard process. (The issue of additional interactions is covered in the next section.)

A description of the beam-remnant structure contains a few components. First, given the flavour content of a (colour-singlet) beam particle, and the flavour and colour of the initiator parton, it is possible to reconstruct the flavour and colour of the beam remnant. Sometimes the remnant may be represented by just a single parton or diquark, but often the remnant has to be subdivided into two separate objects. In the latter case it is necessary to share the remnant energy and momentum between the two. Due to Fermi motion inside hadron beams, the initiator parton may have a 'primordial $k_\perp$' transverse momentum motion, which has to be compensated by the beam remnant. If the remnant is subdivided, there may also be a relative transverse momentum. In the end, total energy and momentum has to be conserved. To first approximation, this is ensured within each remnant separately, but some final global adjustments are necessary to compensate for the primordial $k_\perp$ and any effective beam-remnant mass.

11.1.1 Hadron Beams

Consider first a proton (or, with trivial modifications, any other baryon or antibaryon).

- If the initiator parton is a u or d quark, it is assumed to be a valence quark, and therefore leaves behind a diquark beam remnant, i.e. either a ud or a uu diquark, in a colour antitriplet state. Relative probabilities for different diquark spins are derived within the context of the non-relativistic **SU(6)** model, i.e. flavour **SU(3)** times spin **SU(2)**. Thus a ud is 3/4 ud_0 and 1/4 ud_1, while a uu is always uu_1.
- An initiator gluon leaves behind a colour octet uud state, which is subdivided into a colour triplet quark and a colour antitriplet diquark. **SU(6)** gives the appropriate subdivision, 1/2 of the time into $\mathrm{u} + \mathrm{ud}_0$, 1/6 into $\mathrm{u} + \mathrm{ud}_1$ and 1/3 into $\mathrm{d} + \mathrm{uu}_1$.
- A sea quark initiator, such as an s, leaves behind a uud$\bar{\mathrm{s}}$ four-quark state. The PDG flavour coding scheme and the fragmentation routines do not foresee such a state, so therefore it is subdivided into a meson plus a diquark, i.e. 1/2 into $\mathrm{u}\bar{\mathrm{s}} + \mathrm{ud}_0$, 1/6 into $\mathrm{u}\bar{\mathrm{s}} + \mathrm{ud}_1$ and 1/3 into $\mathrm{d}\bar{\mathrm{s}} + \mathrm{uu}_1$. Once the flavours of the meson are determined, the choice of meson multiplet is performed as in the standard fragmentation description.
- Finally, an antiquark initiator, such as an $\bar{\mathrm{s}}$, leaves behind a uuds four-quark state, which is subdivided into a baryon plus a quark. Since, to first approximation, the s$\bar{\mathrm{s}}$ pair comes from the branching $\mathrm{g} \to \mathrm{s}\bar{\mathrm{s}}$ of a colour octet gluon, the subdivision uud + s is not allowed, since it would correspond to a colour-singlet s$\bar{\mathrm{s}}$. Therefore the subdivision is 1/2 into $\mathrm{ud}_0\mathrm{s} + \mathrm{u}$, 1/6 into $\mathrm{ud}_1\mathrm{s} + \mathrm{u}$ and 1/3 into $\mathrm{uu}_1\mathrm{s} + \mathrm{d}$. A baryon is formed among the ones possible for the given flavour content and diquark spin, according to the relative probabilities used in the fragmentation. One could argue for an additional weighting to count the number of baryon states available for a given diquark plus quark combination, but this has not been included.

One may note that any u or d quark taken out of the proton is automatically assumed to be a valence quark. Clearly this is unrealistic, but not quite as bad as it might seem. In particular, one should remember that the beam-remnant scenario is applied to the initial-state shower initiators at a scale of $Q_0 \approx 1$ GeV and at an x value usually much

larger than the x at the hard scattering. The sea quark contribution therefore normally is small.

For a meson beam remnant, the rules are in the same spirit, but somewhat easier, since no diquark or baryons need be taken into account. Thus a valence quark (antiquark) initiator leaves behind a valence antiquark (quark), a gluon initiator leaves behind a valence quark plus a valence antiquark, and a sea quark (antiquark) leaves behind a meson (which contains the partner to the sea parton) plus a valence antiquark (quark).

11.1.2 Photon Beams

A resolved photon is similar in spirit to a meson. A VMD photon is associated with either ρ^0, ω, ϕ or J/ψ, and so corresponds to a well-defined valence flavour content. Since the ρ^0 and ω are supposed to add coherently, the $\mathrm{u}\overline{\mathrm{u}} : \mathrm{d}\overline{\mathrm{d}}$ mixing is in the ratio 4 : 1. Similarly a GVMD state is characterized by its $\mathrm{q}\overline{\mathrm{q}}$ classification, in rates according to e_{q}^2 times a mass suppression for heavier quarks.

In the older photon physics options, where a quark content inside an electron is obtained by a numerical convolution, one does not have to make the distinction between valence and sea flavour. Thus any quark (antiquark) initiator leaves behind the matching antiquark (quark), and a gluon leaves behind a quark + antiquark pair. The relative quark flavour composition in the latter case is assumed proportional to e_{q}^2 among light flavours, i.e. 2/3 into $\mathrm{u} + \overline{\mathrm{u}}$, 1/6 into $\mathrm{d} + \overline{\mathrm{d}}$, and 1/6 into $\mathrm{s} + \overline{\mathrm{s}}$. If one wanted to, one could also have chosen to represent the remnant by a single gluon.

11.1.3 Lepton Beams

If no initial-state radiation is assumed, an electron (or, in general, a lepton or a neutrino) leaves behind no beam remnant. Also when radiation is included, one would expect to recover a single electron with the full beam energy when the shower initiator is reconstructed. This does not have to happen, e.g. if the initial-state shower is cut off at a non-vanishing scale, such that some of the emission at low Q^2 values is not simulated. Further, for purely technical reasons, the distribution of an electron inside an electron, $f_{\mathrm{e}}^{\mathrm{e}}(x, Q^2)$, is cut off at $x = 1 - 10^{-10}$. This means that always, when initial-state radiation is included, a fraction of at least 10^{-10} of the beam energy has to be put into one single photon along the beam direction, to represent this not simulated radiation. The physics is here slightly different from the standard beam-remnant concept, but it is handled with the same machinery. Beam remnants can also appear when the electron is resolved with the use of parton distributions, but initial-state radiation is switched off. Conceptually, this is a contradiction, since it is the initial-state radiation that builds up the parton distributions, but sometimes the combination is still useful. Finally, since QED radiation has not yet been included in events with resolved photons inside electrons, also in this case effective beam remnants have to be assigned by the program.

The beam-remnant assignments inside an electron, in either of the cases above, is as follows.

- An e^- initiator leaves behind a γ remnant.
- A γ initiator leaves behind an e^- remnant.
- An e^+ initiator leaves behind an $\mathrm{e}^- + \mathrm{e}^-$ remnant.
- A q ($\overline{\mathrm{q}}$) initiator leaves behind a $\overline{\mathrm{q}} + \mathrm{e}^-$ ($\mathrm{q} + \mathrm{e}^-$) remnant.
- A g initiator leaves behind a $\mathrm{g} + \mathrm{e}^-$ remnant. One could argue that, in agreement with the treatment of photon beams above, the remnant should be $\mathrm{q} + \overline{\mathrm{q}} + \mathrm{e}^-$. The program currently does not allow for three beam-remnant objects, however.

11.1.4 Primordial $k_\perp$

It is customary to assign a primordial transverse momentum to the shower initiator, to take into account the motion of quarks inside the original hadron, basically as required by the uncertainty principle. A number of the order of $\langle k_\perp \rangle \approx m_\mathrm{p}/3 \approx 300$ MeV could therefore be expected. However, in hadronic collisions much higher numbers than that are often required to describe data, typically of the order of 1 GeV at fixed-target energies[EMC87] and 2 GeV at collider energies [Miu99, Bál01], if a Gaussian parameterization is used. Thus, an interpretation as a purely nonperturbative motion inside a hadron is difficult to maintain.

Instead a likely culprit is the initial-state shower algorithm. This is set up to cover the region of hard emissions, but may miss out on some of the softer activity, which inherently borders on nonperturbative physics. By default, the shower does not evolve down to scales below $Q_0 = 1$ GeV. Any shortfall in shower activity around or below this cutoff then has to be compensated by the primordial $k_\perp$ source, which thereby largely loses its original meaning. One specific reason for such a shortfall is that the current initial-state shower algorithm does not include non-ordered emissions in Q^2, as is predicted to occur especially at small x and Q^2 within the BFKL/CCFM framework [Lip76, Cia87].

11.1.5 Remnant Kinematics

By the hard scattering and the initial-state radiation machinery, the shower initiator has been assigned some fraction x of the four-momentum of the beam particle, leaving behind $1 - x$ to the remnant. If the remnant consists of two objects, this energy and momentum has to be shared, somehow. For an electron in the old photoproduction machinery, the sharing is given from first principles: if, e.g., the initiator is a q, then that q was produced in the sequence of branchings $\mathrm{e} \to \gamma \to \mathrm{q}$, where x_γ is distributed according to the convolution in eq. (54). Therefore the $\overline{\mathrm{q}}$ remnant takes a fraction $\chi = (x_\gamma - x)/(1 - x)$ of the total remnant energy, and the e takes $1 - \chi$.

For the other beam remnants, the relative energy-sharing variable χ is not known from first principles, but picked according to some suitable parameterization. Normally several different options are available, that can be set separately for baryon and meson beams, and for hadron + quark and quark + diquark (or antiquark) remnants. In one extreme are shapes in agreement with naïve counting rules, i.e. where energy is shared evenly between 'valence' partons. For instance, $\mathcal{P}(\chi) = 2\,(1 - \chi)$ for the energy fraction taken by the q in a $\mathrm{q} + \mathrm{qq}$ remnant. In the other extreme, an uneven distribution could be used, like in parton distributions, where the quark only takes a small fraction and most is retained by the diquark. The default for a $\mathrm{q} + \mathrm{qq}$ remnant is of an intermediate type,

$$\mathcal{P}(\chi) \propto \frac{(1-\chi)^3}{\sqrt[4]{\chi^2 + c_\mathrm{min}^2}} \,, \tag{201}$$

with $c_\mathrm{min} = 2\langle m_\mathrm{q} \rangle / E_\mathrm{cm} = (0.6\ \mathrm{GeV})/E_\mathrm{cm}$ providing a lower cut-off. The default when a hadron is split off to leave a quark or diquark remnant is to use the standard Lund symmetric fragmentation function. In general, the more uneven the sharing of the energy, the less the total multiplicity in the beam-remnant fragmentation. If no multiple interactions are allowed, a rather even sharing is needed to come close to the experimental multiplicity (and yet one does not quite make it). With an uneven sharing there is room to generate more of the total multiplicity by multiple interactions [Sjö87a].

In a photon beam, with a remnant $\mathrm{q} + \overline{\mathrm{q}}$, the χ variable is chosen the same way it would have been in a corresponding meson remnant.

Before the χ variable is used to assign remnant momenta, it is also necessary to consider the issue of primordial $k_\perp$. The initiator partons are thus assigned each a $k_\perp$

value, vanishing for an electron or photon inside an electron, distributed either according to a Gaussian or an exponential shape for a hadron, and according to either of these shapes or a power-like shape for a quark or gluon inside a photon (which may in its turn be inside an electron). The interaction subsystem is boosted and rotated to bring it from the frame assumed so far, with each initiator along the $\pm z$ axis, to one where the initiators have the required primordial $k_\perp$ values.

The $p_\perp$ recoil is taken by the remnant. If the remnant is composite, the recoil is evenly split between the two. In addition, however, the two beam remnants may be given a relative $p_\perp$, which is then always chosen as for $\mathrm{q}_i\overline{\mathrm{q}}_i$ pairs in the fragmentation description.

The χ variable is interpreted as a sharing of light-cone energy and momentum, i.e. $E + p_z$ for the beam moving in the $+z$ direction and $E - p_z$ for the other one. When the two transverse masses $m_{\perp 1}$ and $m_{\perp 2}$ of a composite remnant have been constructed, the total transverse mass can therefore be found as

$$m_\perp^2 = \frac{m_{\perp 1}^2}{\chi} + \frac{m_{\perp 2}^2}{1-\chi} \,, \tag{202}$$

if remnant 1 is the one that takes the fraction χ. The choice of a light-cone interpretation to χ means the definition is invariant under longitudinal boosts, and therefore does not depend on the beam energy itself. A χ value close to the naïve borders 0 or 1 can lead to an unreasonably large remnant $m_\perp$. Therefore an additional check is introduced, that the remnant $m_\perp$ be smaller than the naïve c.m. frame remnant energy, $(1-x)E_{\mathrm{cm}}/2$. If this is not the case, a new χ and a new relative transverse momentum is selected.

Whether there is one remnant parton or two, the transverse mass of the remnant is not likely to agree with $1-x$ times the mass of the beam particle, i.e. it is not going to be possible to preserve the energy and momentum in each remnant separately. One therefore allows a shuffling of energy and momentum between the beam remnants from each of the two incoming beams. This may be achieved by performing a (small) longitudinal boost of each remnant system. Since there are two boost degrees of freedom, one for each remnant, and two constraints, one for energy and one for longitudinal momentum, a solution may be found.

Under some circumstances, one beam remnant may be absent or of very low energy, while the other one is more complicated. One example is Deeply Inelastic Scattering in ep collisions, where the electron leaves no remnant, or maybe only a low-energy photon. It is clearly then not possible to balance the two beam remnants against each other. Therefore, if one beam remnant has an energy below 0.2 of the beam energy, i.e. if the initiator parton has $x > 0.8$, then the two boosts needed to ensure energy and momentum conservation are instead performed on the other remnant and on the interaction subsystem. If there is a low-energy remnant at all then, before that, energy and momentum are assigned to the remnant constituent(s) so that the appropriate light-cone combination $E \pm p_z$ is conserved, but not energy or momentum separately. If both beam remnants have low energy, but both still exist, then the one with lower $m_\perp/E$ is the one that will not be boosted.

11.2 Multiple Interactions — Old Model

In this section we present the original model [Sjö87a] to describe the possibility that several parton pairs undergo hard interactions in a hadron–hadron collision, and thereby contribute to the overall event activity, in particular at low $p_\perp$. The same model can also be used to describe the VMD γp events, where the photon interacts like a hadron. Many basic features of this model, for instance the introduction of a $p_\perp$ cutoff corresponding to an inverse colour screening distance, and the options for a non-trivial transverse density structure in the incoming hadrons, carry over to the new scenario. It is therefore recom-

mended first to read this section, even if the objective should be to learn about the new scenario.

It should from the onset be made clear that this is not an easy topic. In fact, in the full event generation process, probably no other area is as poorly understood as this one. The whole concept of multiple interactions has been very controversial, with contradictory experimental conclusions [AFS87], but a CDF study [CDF97] some years ago started to bring more general acceptance, further accelerated by the underlying-event studies of R.D. Field [Fie02].

The multiple interactions scenario presented here [Sjö87a] was the first detailed model for this kind of physics, and is still one of the very few available. We will present two related but separate scenarios, one 'simple' model and one somewhat more sophisticated. In fact, neither of them are all that simple, which may make the models look unattractive. However, the world of hadron physics *is* complicated, and if we err, it is most likely in being too unsophisticated. The experience gained with the model(s), in failures as well as successes, could be used as a guideline in the evolution of yet more detailed and accurate models.

Our basic philosophy will be as follows. The total rate of parton–parton interactions, as a function of the transverse momentum scale $p_\perp$, is assumed to be given by perturbative QCD. This is certainly true for reasonably large $p_\perp$ values, but here we shall also extend the perturbative parton–parton scattering framework into the low-$p_\perp$ region. A regularization of the divergence in the cross section for $p_\perp \to 0$ has to be introduced, however, which will provide us with the main free parameter of the model. Since each incoming hadron is a composite object, consisting of many partons, there should exist the possibility of several parton pairs interacting when two hadrons collide. It is not unreasonable to assume that the different pairwise interactions take place essentially independently of each other, and that therefore the number of interactions in an event is given by a Poisson distribution. This is the strategy of the 'simple' scenario.

Furthermore, hadrons are not only composite but also extended objects, meaning that collisions range from very central to rather peripheral ones. Reasonably, the average number of interactions should be larger in the former than in the latter case. Whereas the assumption of a Poisson distribution should hold for each impact parameter separately, the distribution in number of interactions should be widened by the spread of impact parameters. The amount of widening depends on the assumed matter distribution inside the colliding hadrons. In the 'complex' scenario, different matter distributions are therefore introduced.

11.2.1 The basic cross sections

The QCD cross section for hard $2 \to 2$ processes, as a function of the $p_\perp^2$ scale, is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_\perp^2} = \sum_{i,j,k} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \int \mathrm{d}\hat{t}\, f_i(x_1, Q^2)\, f_j(x_2, Q^2)\, \frac{\mathrm{d}\hat{\sigma}_{ij}^k}{\mathrm{d}\hat{t}}\, \delta\left(p_\perp^2 - \frac{\hat{t}\hat{u}}{\hat{s}}\right) , \tag{203}$$

cf. section 7.2. Implicitly, from now on we are assuming that the 'hardness' of processes is given by the $p_\perp$ scale of the scattering. For an application of the formula above to small $p_\perp$ values, a number of caveats could be made. At low $p_\perp$, the integrals receive major contributions from the small-x region, where parton distributions are poorly understood theoretically (Regge-limit behaviour, dense-packing problems etc. [Lev90]) and not yet measured. Different sets of parton distributions can therefore give numerically rather different results for the phenomenology of interest. One may also worry about higher-order corrections to the jet rates, K factors, beyond what is given by parton-shower corrections — one simple option we allow here is to evaluate α_s of the hard-scattering process at an optimized scale, such as $\alpha_\mathrm{s}(0.075p_\perp^2)$ [Ell86].

The hard-scattering cross section above some given $p_{\perp\min}$ is given by

$$\sigma_{\rm hard}(p_{\perp\min}) = \int_{p^2_{\perp\min}}^{s/4} \frac{{\rm d}\sigma}{{\rm d}p^2_\perp}\,{\rm d}p^2_\perp \; . \tag{204}$$

Since the differential cross section diverges roughly like ${\rm d}p^2_\perp/p^4_\perp$, $\sigma_{\rm hard}$ is also divergent for $p_{\perp\min} \to 0$. We may compare this with the total inelastic, non-diffractive cross section $\sigma_{\rm nd}(s)$ — elastic and diffractive events are not the topic of this section. At current collider energies $\sigma_{\rm hard}(p_{\perp\min})$ becomes comparable with $\sigma_{\rm nd}$ for $p_{\perp\min} \approx$ 2–3 GeV, and at larger energies this occurs at even larger $p_{\perp\min}$. This need not lead to contradictions: $\sigma_{\rm hard}$ does not give the hadron–hadron cross section but the parton–parton one. Each of the incoming hadrons may be viewed as a beam of partons, with the possibility of having several parton–parton interactions when the hadrons pass through each other. In this language, $\sigma_{\rm hard}(p_{\perp\min})/\sigma_{\rm nd}(s)$ is simply the average number of parton–parton scatterings above $p_{\perp\min}$ in an event, and this number may well be larger than unity.

While the introduction of several interactions per event is the natural consequence of allowing small $p_{\perp\min}$ values and hence large $\sigma_{\rm hard}$ ones, it is not the solution of $\sigma_{\rm hard}(p_{\perp\min})$ being divergent for $p_{\perp\min} \to 0$: the average $\hat{s}$ of a scattering decreases slower with $p_{\perp\min}$ than the number of interactions increases, so naïvely the total amount of scattered partonic energy becomes infinite. One cut-off is therefore obtained via the need to introduce proper multi-parton correlated parton distributions inside a hadron. This is not a part of the standard perturbative QCD formalism and is therefore not built into eq. (204). In practice, even correlated parton-distribution functions seems to provide too weak a cut, i.e. one is lead to a picture with too little of the incoming energy remaining in the small-angle beam-jet region [Sjö87a].

A more credible reason for an effective cut-off is that the incoming hadrons are colour neutral objects. Therefore, when the $p_\perp$ of an exchanged gluon is made small and the transverse wavelength correspondingly large, the gluon can no longer resolve the individual colour charges, and the effective coupling is decreased. This mechanism is not in contradiction with perturbative QCD calculations, which are always performed assuming scattering of free partons (rather than partons inside hadrons), but neither does present knowledge of QCD provide an understanding of how such a decoupling mechanism would work in detail. In the simple model one makes use of a sharp cut-off at some scale $p_{\perp\min}$, while a more smooth dampening is assumed for the complex scenario.

One key question is the energy-dependence of $p_{\perp\min}$; this may be relevant e.g. for comparisons of jet rates at different Tevatron/RHIC energies, and even more for any extrapolation to LHC energies. The problem actually is more pressing now than at the time of the original study [Sjö87a], since nowadays parton distributions are known to be rising more steeply at small x than the flat $xf(x)$ behaviour normally assumed for small Q^2 before HERA. This translates into a more dramatic energy dependence of the multiple-interactions rate for a fixed $p_{\perp\min}$.

The larger number of partons should also increase the amount of screening, however, as confirmed by toy simulations [Dis01]. As a simple first approximation, $p_{\perp\min}$ is assumed to increase in the same way as the total cross section, i.e. with some power $\epsilon \approx 0.08$ [Don92] that, via reggeon phenomenology, should relate to the behaviour of parton distributions at small x and Q^2. Thus the default in PYTHIA is

$$p_{\perp\min}(s) = (1.9\ {\rm GeV}) \left(\frac{s}{1\ {\rm TeV}^2}\right)^{0.08} \tag{205}$$

for the simple model, with the same ansatz for $p_{\perp 0}$ in the impact-parameter-dependent approach, except that then 1.9 GeV $\to$ 2.0 GeV. At any energy scale, the simplest criteria to fix $p_{\perp\min}$ is to require the average charged multiplicity $\langle n_{\rm ch} \rangle$ or the height of the (pseudo)rapidity 'plateau' ${\rm d}n_{\rm ch}/{\rm d}\eta|_{\eta=0}$ to agree with the experimentally determined one.

In general, there is quite a strong dependence of the multiplicity on $p_{\perp\min}$, with a lower $p_{\perp\min}$ corresponding to more multiple interactions and therefore a higher multiplicity. This is one of the possible inputs into the 1.9 GeV and 2.0 GeV numbers, making use of UA5 data in the energy range 200–900 GeV [UA584]. The energy dependence inside this range is also consistent with the chosen ansatz [But05]. However, clearly, neither the experimental nor the theoretical precision is high enough to make too strong statements. It should also be remembered that the $p_{\perp\min}$ values are determined within the context of a given calculation of the QCD jet cross section, and given model parameters within the multiple-interactions scenario. If anything of this is changed, e.g. the parton distributions used, then $p_{\perp\min}$ ought to be retuned accordingly.

11.2.2 The simple model

In an event with several interactions, it is convenient to impose an ordering. The logical choice is to arrange the scatterings in falling sequence of $x_\perp = 2p_\perp/E_{\rm cm}$. The 'first' scattering is thus the hardest one, with the 'subsequent' ('second', 'third', etc.) successively softer. It is important to remember that this terminology is in no way related to any picture in physical time; we do not know anything about the latter. (In a simplified picture with the incoming hadrons Lorentz-contracted into flat pancakes, the interactions would tend to have a specelike separation, i.e. without meaningful time ordering.) In principle, all the scatterings that occur in an event must be correlated somehow, naïvely by momentum and flavour conservation for the partons from each incoming hadron, less naïvely by various quantum mechanical effects. When averaging over all configurations of soft partons, however, one should effectively obtain the standard QCD phenomenology for a hard scattering, e.g. in terms of parton distributions. Correlation effects, known or estimated, can be introduced in the choice of subsequent scatterings, given that the 'preceding' (harder) ones are already known.

With a total cross section of hard interactions $\sigma_{\rm hard}(p_{\perp\min})$ to be distributed among $\sigma_{\rm nd}(s)$ (non-diffractive, inelastic) events, the average number of interactions per event is just the ratio $\overline{n} = \sigma_{\rm hard}(p_{\perp\min})/\sigma_{\rm nd}(s)$. As a starting point we will assume that all hadron collisions are equivalent (no impact-parameter dependence), and that the different parton–parton interactions take place completely independently of each other. The number of scatterings per event is then distributed according to a Poisson distribution with mean $\overline{n}$. A fit to Sp$\overline{\rm p}$S collider multiplicity data [UA584] gave $p_{\perp\min} \approx 1.6$ GeV (for parton distributions in use at the time), which corresponds to $\overline{n} \approx 1$. For Monte Carlo generation of these interactions it is useful to define

$$f(x_\perp) = \frac{1}{\sigma_{\rm nd}(s)} \frac{{\rm d}\sigma}{{\rm d}x_\perp} \; , \tag{206}$$

with ${\rm d}\sigma/{\rm d}x_\perp$ obtained by analogy with eq. (203). Then $f(x_\perp)$ is simply the probability to have a parton–parton interaction at $x_\perp$, given that the two hadrons undergo a non-diffractive, inelastic collision.

The probability that the hardest interaction, i.e. the one with highest $x_\perp$, is at $x_{\perp 1}$, is now given by

$$f(x_{\perp 1}) \exp\left\{ - \int_{x_{\perp 1}}^1 f(x'_\perp)\, {\rm d}x'_\perp \right\} \; , \tag{207}$$

i.e. the naïve probability to have a scattering at $x_{\perp 1}$ multiplied by the probability that there was no scattering with $x_\perp$ larger than $x_{\perp 1}$. This is the familiar exponential dampening in radioactive decays, encountered e.g. in parton showers in section 10.1.2. Using the same technique as in the proof of the veto algorithm, section 4.2, the probability to have an i:th scattering at an $x_{\perp i} < x_{\perp i-1} < \cdots < x_{\perp 1} < 1$ is found to be

$$f(x_{\perp i}) \frac{1}{(i-1)!} \left(\int_{x_{\perp i}}^1 f(x'_\perp)\, {\rm d}x'_\perp \right)^{i-1} \exp\left\{ - \int_{x_{\perp i}}^1 f(x'_\perp)\, {\rm d}x'_\perp \right\} \; . \tag{208}$$

The total probability to have a scattering at a given $x_\perp$, irrespectively of it being the first, the second or whatever, obviously adds up to give back $f(x_\perp)$. The multiple interaction formalism thus retains the correct perturbative QCD expression for the scattering probability at any given $x_\perp$.

With the help of the integral

$$F(x_\perp) = \int_{x_\perp}^1 f(x'_\perp)\,\mathrm{d}x'_\perp = \frac{1}{\sigma_{\mathrm{nd}}(s)} \int_{sx_\perp^2/4}^{s/4} \frac{\mathrm{d}\sigma}{\mathrm{d}p_\perp^2}\,\mathrm{d}p_\perp^2 \tag{209}$$

(where we assume $F(x_\perp) \to \infty$ for $x_\perp \to 0$) and its inverse F^{-1}, the iterative procedure to generate a chain of scatterings $1 > x_{\perp 1} > x_{\perp 2} > \cdots > x_{\perp i}$ is given by

$$x_{\perp i} = F^{-1}(F(x_{\perp i-1}) - \ln R_i)\ . \tag{210}$$

Here the R_i are random numbers evenly distributed between 0 and 1. The iterative chain is started with a fictitious $x_{\perp 0} = 1$ and is terminated when $x_{\perp i}$ is smaller than $x_{\perp\mathrm{min}} = 2p_{\perp\mathrm{min}}/E_{\mathrm{cm}}$. Since F and F^{-1} are not known analytically, the standard veto algorithm is used to generate a much denser set of $x_\perp$ values, whereof only some are retained in the end. In addition to the $p_\perp^2$ of an interaction, it is also necessary to generate the other flavour and kinematics variables according to the relevant matrix elements.

Whereas the ordinary parton distributions should be used for the hardest scattering, in order to reproduce standard QCD phenomenology, the parton distributions to be used for subsequent scatterings must depend on all preceding x values and flavours chosen. We do not know enough about the hadron wave function to write down such joint probability distributions. To take into account the energy 'already' used in harder scatterings, a conservative approach is to evaluate the parton distributions, not at x_i for the i:th scattered parton from hadron, but at the rescaled value

$$x'_i = \frac{x_i}{1 - \sum_{j=1}^{i-1} x_j}\ . \tag{211}$$

This is our standard procedure in the simple model; we have tried a few alternatives without finding any significantly different behaviour in the final physics.

In a fraction $\exp(-F(x_{\perp\mathrm{min}}))$ of the events studied, there will be no hard scattering above $x_{\perp\mathrm{min}}$ when the iterative procedure in eq. (210) is applied. It is therefore also necessary to have a model for what happens in events with no (semi)hard interactions. The simplest possible way to produce an event is to have an exchange of a very soft gluon between the two colliding hadrons. Without (initially) affecting the momentum distribution of partons, the 'hadrons' become colour octet objects rather than colour singlet ones. If only valence quarks are considered, the colour octet state of a baryon can be decomposed into a colour triplet quark and an antitriplet diquark. In a baryon-baryon collision, one would then obtain a two-string picture, with each string stretched from the quark of one baryon to the diquark of the other. A baryon-antibaryon collision would give one string between a quark and an antiquark and another one between a diquark and an antidiquark.

In a hard interaction, the number of possible string drawings are many more, and the overall situation can become quite complex when several hard scatterings are present in an event. Specifically, the string drawing now depends on the relative colour arrangement, in each hadron individually, of the partons that are about to scatter. This is a subject about which nothing is known. To make matters worse, the standard string fragmentation description would have to be extended, to handle events where two or more valence quarks have been kicked out of an incoming hadron by separate interactions. In particular, the position of the baryon number would be unclear. Such issues will be further discussed below, when we go on to describe more recent models, but in the original studies they were sidestepped. Specifically, we assumed that, following the hardest interaction, all subsequent interactions belong to one of three classes.

- Scatterings of the gg → gg type, with the two gluons in a colour-singlet state, such that a double string is stretched directly between the two outgoing gluons, decoupled from the rest of the system.
- Scatterings gg → gg, but colour correlations assumed to be such that each of the gluons is connected to one of the strings 'already' present. Among the different possibilities of connecting the colours of the gluons, the one which minimizes the total increase in string length is chosen. This is in contrast to the previous alternative, which roughly corresponds to a maximization (within reason) of the extra string length.
- Scatterings gg → $q\overline{q}$, with the final pair again in a colour-singlet state, such that a single string is stretched between the outgoing q and $\overline{q}$.

By default, the three possibilities were assumed equally probable. (More recent studies [Fie02] have suggested the minimal string length topology to dominate, an issue well worth studying further.) Note that the total jet rate is maintained at its nominal value, i.e. scatterings such as qg → qg are included in the cross section, but are replaced by a mixture of gg and $q\overline{q}$ events for string drawing issues. Only the hardest interaction is guaranteed to give strings coupled to the beam remnants. One should not take this approach to colour flow too seriously — clearly it is a simplification — but the overall picture does not tend to be very dependent on the particular choice you make.

Since a gg → gg or $q\overline{q}$ scattering need not remain of this character if initial- and final-state showers were to be included (e.g. it could turn into a qg-initiated process), radiation is only included for the hardest interaction. In practice, this need not be a serious problem: except for the hardest interaction, which can be hard because of experimental trigger conditions, it is unlikely for a parton scattering to be so hard that radiation plays a significant rôle.

In events with multiple interactions, the beam-remnant treatment is slightly modified. First the hard scattering is generated, with its associated initial- and final-state radiation, and next any additional multiple interactions. Only thereafter are beam remnants attached to the initiator partons of the hardest scattering, using the same machinery as before, except that the energy and momentum already taken away from the beam remnants also include that of the subsequent interactions.

11.2.3 A model with varying impact parameters

Up to this point, it has been assumed that the initial state is the same for all hadron collisions, whereas in fact each collision also is characterized by a varying impact parameter b. Within the classical framework of the model reviewed here, b is to be thought of as a distance of closest approach, not as the Fourier transform of the momentum transfer. A small b value corresponds to a large overlap between the two colliding hadrons, and hence an enhanced probability for multiple interactions. A large b, on the other hand, corresponds to a grazing collision, with a large probability that no parton–parton interactions at all take place.

In order to quantify the concept of hadronic matter overlap, one may assume a spherically symmetric distribution of matter inside the hadron, $\rho(\mathbf{x})\,\mathrm{d}^3x = \rho(r)\,\mathrm{d}^3x$. For simplicity, the same spatial distribution is taken to apply for all parton species and momenta. Several different matter distributions have been tried, and are available. We will here concentrate on the most extreme one, a double Gaussian

$$\rho(r) \propto \frac{1-\beta}{a_1^3}\,\exp\left\{-\frac{r^2}{a_1^2}\right\} + \frac{\beta}{a_2^3}\,\exp\left\{-\frac{r^2}{a_2^2}\right\}\ . \tag{212}$$

This corresponds to a distribution with a small core region, of radius a_2 and containing a fraction β of the total hadronic matter, embedded in a larger hadron of radius a_1. While

it is mathematically convenient to have the origin of the two Gaussians coinciding, the physics could well correspond to having three disjoint core regions, reflecting the presence of three valence quarks, together carrying the fraction β of the proton momentum. One could alternatively imagine a hard hadronic core surrounded by a pion cloud. Such details would affect e.g. the predictions for the t distribution in elastic scattering, but are not of any consequence for the current topics. To be specific, the values $\beta = 0.5$ and $a_2/a_1 = 0.2$ were picked as default values. It should be noted that the overall distance scale a_1 never enters in the subsequent calculations, since the inelastic, non-diffractive cross section $\sigma_{\rm nd}(s)$ is taken from literature rather than calculated from the $\rho(r)$.

Compared to other shapes, like a simple Gaussian, the double Gaussian tends to give larger fluctuations, e.g. in the multiplicity distribution of minimum-bias events: a collision in which the two cores overlap tends to have a strongly increased activity, while ones where they do not are rather less active. One also has a biasing effect: hard processes are more likely when the cores overlap, thus hard scatterings are associated with an enhanced multiple interaction rate. This provides one possible explanation for the experimental 'pedestal effect' [UA187]. Recent studies of CDF data [Fie02, Mor02] have confirmed that indeed something more peaked than a single Gaussian is required to understand the transition from minimum-bias to underlying-event activity.

For a collision with impact parameter b, the time-integrated overlap $\mathcal{O}(b)$ between the matter distributions of the colliding hadrons is given by

$$\begin{aligned}\mathcal{O}(b) &\propto \int \mathrm{d}t \int \mathrm{d}^3x\, \rho(x,y,z)\, \rho(x+b,y,z+t) \\ &\propto \frac{(1-\beta)^2}{2a_1^2} \exp\left\{-\frac{b^2}{2a_1^2}\right\} + \frac{2\beta(1-\beta)}{a_1^2+a_2^2} \exp\left\{-\frac{b^2}{a_1^2+a_2^2}\right\} + \frac{\beta^2}{2a_2^2} \exp\left\{-\frac{b^2}{2a_2^2}\right\} \end{aligned} \quad (213)$$

The necessity to use boosted $\rho(\mathbf{x})$ distributions has been circumvented by a suitable scale transformation of the z and t coordinates.

The overlap $\mathcal{O}(b)$ is obviously strongly related to the eikonal $\Omega(b)$ of optical models. We have kept a separate notation, since the physics context of the two is slightly different: $\Omega(b)$ is based on the quantum mechanical scattering of waves in a potential, and is normally used to describe the elastic scattering of a hadron-as-a-whole, while $\mathcal{O}(b)$ comes from a purely classical picture of point-like partons distributed inside the two colliding hadrons. Furthermore, the normalization and energy dependence is differently realized in the two formalisms.

The larger the overlap $\mathcal{O}(b)$ is, the more likely it is to have interactions between partons in the two colliding hadrons. In fact, there should be a linear relationship

$$\langle \tilde{n}(b) \rangle = k\mathcal{O}(b) \ , \quad (214)$$

where $\tilde{n} = 0, 1, 2, \ldots$ counts the number of interactions when two hadrons pass each other with an impact parameter b. The constant of proportionality, k, is related to the parton–parton cross section and hence increases with c.m. energy.

For each given impact parameter, the number of interactions is assumed to be distributed according to a Poisson. If the matter distribution has a tail to infinity (as the double Gaussian does), events may be obtained with arbitrarily large b values. In order to obtain finite total cross sections, it is necessary to assume that each event contains at least one semi-hard interaction. The probability that two hadrons, passing each other with an impact parameter b, will actually undergo a collision is then given by

$$\mathcal{P}_{\rm int}(b) = 1 - \exp(-\langle \tilde{n}(b) \rangle) = 1 - \exp(-k\mathcal{O}(b)) \ , \quad (215)$$

according to Poisson statistics. The average number of interactions per event at impact parameter b is now

$$\langle n(b) \rangle = \frac{\langle \tilde{n}(b) \rangle}{\mathcal{P}_{\rm int}(b)} = \frac{k\mathcal{O}(b)}{1 - \exp(-k\mathcal{O}(b))} \ , \quad (216)$$

where the denominator comes from the removal of hadron pairs which pass without colliding, i.e. with $\tilde{n} = 0$.

The relationship $\langle n \rangle = \sigma_{\rm hard}/\sigma_{\rm nd}$ was earlier introduced for the average number of interactions per non-diffractive, inelastic event. When averaged over all impact parameters, this relation must still hold true: the introduction of variable impact parameters may give more interactions in some events and less in others, but it does not affect either $\sigma_{\rm hard}$ or $\sigma_{\rm nd}$. For the former this is because the perturbative QCD calculations only depend on the total parton flux, for the latter by construction. Integrating eq. (216) over b, one then obtains

$$\langle n \rangle = \frac{\int \langle n(b) \rangle \, \mathcal{P}_{\rm int}(b) \, {\rm d}^2b}{\int \mathcal{P}_{\rm int}(b) \, {\rm d}^2b} = \frac{\int k\mathcal{O}(b) \, {\rm d}^2b}{\int (1 - \exp(-k\mathcal{O}(b))) \, {\rm d}^2b} = \frac{\sigma_{\rm hard}}{\sigma_{\rm nd}} \; . \tag{217}$$

For $\mathcal{O}(b)$, $\sigma_{\rm hard}$ and $\sigma_{\rm nd}$ given, with $\sigma_{\rm hard}/\sigma_{\rm nd} > 1$, k can thus always be found (numerically) by solving the last equality.

The absolute normalization of $\mathcal{O}(b)$ is not interesting in itself, but only the relative variation with impact parameter. It is therefore useful to introduce an 'enhancement factor' $e(b)$, which gauges how the interaction probability for a passage with impact parameter b compares with the average, i.e.

$$\langle \tilde{n}(b) \rangle = k\mathcal{O}(b) = e(b) \, \langle k\mathcal{O}(b) \rangle \; . \tag{218}$$

The definition of the average $\langle k\mathcal{O}(b) \rangle$ is a bit delicate, since the average number of interactions per event is pushed up by the requirement that each event contain at least one interaction. However, an exact meaning can be given [Sjö87a].

With the knowledge of $e(b)$, the $f(x_\perp)$ function of the simple model generalizes to

$$f(x_\perp, b) = e(b) \, f(x_\perp) \; . \tag{219}$$

The naïve generation procedure is thus to pick a b according to the phase space ${\rm d}^2b$, find the relevant $e(b)$ and plug in the resulting $f(x_\perp, b)$ in the formalism of the simple model. If at least one hard interaction is generated, the event is retained, else a new b is to be found. This algorithm would work fine for hadronic matter distributions which vanish outside some radius, so that the ${\rm d}^2b$ phase space which needs to be probed is finite. Since this is not true for the distributions under study, it is necessary to do better.

By analogy with eq. (207), it is possible to ask what the probability is to find the hardest scattering of an event at $x_{\perp 1}$. For each impact parameter separately, the probability to have an interaction at $x_{\perp 1}$ is given by $f(x_{\perp 1}, b)$, and this should be multiplied by the probability that the event contains no interactions at a scale $x'_\perp > x_{\perp 1}$, to yield the total probability distribution

$$\begin{aligned} \frac{{\rm d}\mathcal{P}_{\rm hardest}}{{\rm d}^2b \, {\rm d}x_{\perp 1}} &= f(x_{\perp 1}, b) \, \exp\left\{ - \int_{x_{\perp 1}}^1 f(x'_\perp, b) \, {\rm d}x'_\perp \right\} \\ &= e(b) \, f(x_{\perp 1}) \, \exp\left\{ -e(b) \int_{x_{\perp 1}}^1 f(x'_\perp) \, {\rm d}x'_\perp \right\} \; . \end{aligned} \tag{220}$$

If the treatment of the exponential is deferred for a moment, the distribution in b and $x_{\perp 1}$ appears in factorized form, so that the two can be chosen independently of each other. In particular, a high-$p_\perp$ QCD scattering or any other hard scattering can be selected with whatever kinematics desired for that process, and thereafter assigned some suitable 'hardness' $x_{\perp 1}$. With the b chosen according to $e(b) \, {\rm d}^2b$, the neglected exponential can now be evaluated, and the event retained with a probability proportional to it. From the $x_{\perp 1}$ scale of the selected interaction, a sequence of softer $x_{\perp i}$ values may again be generated as in the simple model, using the known $f(x_\perp, b)$. This sequence may be empty, i.e. the event need not contain any further interactions.

It is interesting to understand how the algorithm above works. By selecting b according to $e(b)\,\mathrm{d}^2b$, i.e. $\mathcal{O}(b)\,\mathrm{d}^2b$, the primary b distribution is maximally biased towards small impact parameters. If the first interaction is hard, by choice or by chance, the integral of the cross section above $x_{\perp 1}$ is small, and the exponential close to unity. The rejection procedure is therefore very efficient for all standard hard processes in the program — one may even safely drop the weighting with the exponential completely. The large $e(b)$ value is also likely to lead to the generation of many further, softer interactions. If, on the other hand, the first interaction is not hard, the exponential is no longer close to unity, and many events are rejected. This pulls down the efficiency for 'minimum bias' event generation. Since the exponent is proportional to $e(b)$, a large $e(b)$ leads to an enhanced probability for rejection, whereas the chance of acceptance is larger with a small $e(b)$. Among events where the hardest interaction is soft, the b distribution is therefore biased towards larger values (smaller $e(b)$), and there is a small probability for yet softer interactions.

To evaluate the exponential factor, the program pretabulates the integral of $f(x_\perp)$ at the initialization stage, and further increases the Monte Carlo statistics of this tabulation as the run proceeds. The $x_\perp$ grid is concentrated towards small $x_\perp$, where the integral is large. For a selected $x_{\perp 1}$ value, the $f(x_\perp)$ integral is obtained by interpolation. After multiplication by the known $e(b)$ factor, the exponential factor may be found.

In this section, nothing has yet been assumed about the form of the $\mathrm{d}\sigma/\mathrm{d}p_\perp$ spectrum. Like in the impact-parameter-independent case, it is possible to use a sharp cut-off at some given $p_{\perp\mathrm{min}}$ value. However, now each event is required to have at least one interaction, whereas before events without interactions were retained and put at $p_\perp = 0$. It is therefore aesthetically more appealing to assume a gradual turn-off, so that a (semi)hard interaction can be rather soft part of the time. The matrix elements roughly diverge like $\alpha_\mathrm{s}(p_\perp^2)\,\mathrm{d}p_\perp^2/p_\perp^4$ for $p_\perp \to 0$. They could therefore be regularized as follows. Firstly, to remove the $1/p_\perp^4$ behaviour, multiply by a factor $p_\perp^4/(p_\perp^2 + p_{\perp 0}^2)^2$. Secondly, replace the $p_\perp^2$ argument in α_s by $p_\perp^2 + p_{\perp 0}^2$. If one has included a K factor by a rescaling of the α_s argument, as mentioned earlier, replace $0.075\,p_\perp^2$ by $0.075\,(p_\perp^2 + p_{\perp 0}^2)$.

With these substitutions, a continuous $p_\perp$ spectrum is obtained, stretching from $p_\perp = 0$ to $E_\mathrm{cm}/2$. For $p_\perp \gg p_{\perp 0}$ the standard perturbative QCD cross section is recovered, while values $p_\perp \ll p_{\perp 0}$ are strongly damped. The $p_{\perp 0}$ scale, which now is the main free parameter of the model, in practice comes out to be of the same order of magnitude as the sharp cut-off $p_{\perp\mathrm{min}}$ did, i.e. 1.5–2 GeV at collider energies, but typically about 10% higher.

Above we have argued that $p_{\perp\mathrm{min}}$ and $p_{\perp 0}$ should only have a slow energy dependence, and even allowed for the possibility of fixed values. For the impact-parameter-independent picture this works out fine, with all events being reduced to low-$p_\perp$ two-string ones when the c.m. energy is reduced. In the variable-impact-parameter picture, the whole formalism only makes sense if $\sigma_\mathrm{hard} > \sigma_\mathrm{nd}$, see e.g. eq. (217). Since σ_nd does not vanish with decreasing energy, but σ_hard would do that for a fixed $p_{\perp 0}$, this means that $p_{\perp 0}$ has to be reduced significantly at low energies, possibly even more than implied by our assumed energy dependence. The more 'sophisticated' model of this section therefore makes sense at collider energies, whereas it may not be well suited for applications at fixed-target energies. There one should presumably attach to a picture of multiple soft Pomeron exchanges.

11.3 Beam Remnants (and Multiple Interactions) — Intermediate Model

A new description of multiple interactions and beam remnants is introduced with Pythia 6.3 [Sjö04]. It is more sophisticated than the old one, but unfortunately as a consequence less robust. Therefore extra caution should be exercised. Currently it should only be used for the study of single minimum-bias events or events underlying

hard processes in pp or p$\overline{\text{p}}$ collisions. It is not fully developed and tested for meson or photon beams, and cannot be used with pile-up events, or with baryon-number-violating SUSY events.

The basic scheme for the generation of the multiple interactions is kept from the old models. That is, based on the standard $2 \to 2$ QCD matrix elements convoluted with standard parton densities, a sequence of $p_\perp$-ordered interactions is generated: $p_{\perp 1} > p_{\perp 2} > p_{\perp 3} > \ldots$. The sequence is stopped at some lower cut-off scale $p_{\perp \min}$, or alternatively the matrix elements are damped smoothly around and below a scale $p_{\perp 0}$. An impact-parameter-dependent picture of hadron-hadron interactions can lead to further elements of variability between events. See above for details on the basic framework.

The main limitation of the old approach is that there was no way to handle complicated beam remnants, e.g. where two valence quarks had been kicked out. Therefore the structure of all interactions subsequent to the first one, i.e. the one with largest $p_\perp$, had to be substantially simplified. The introduction of junction fragmentation in [Sjö03] allowed this restriction to be lifted.

11.3.1 Flavour and x Correlations

In the new beam-remnants approach, the flavour content of the remnant is bookkept, and is used to determine possible flavours in consecutive interactions. Thus, the standard parton densities are only used to describe the hardest interaction. Already in the old model, the x scale of parton densities is rescaled in subsequent interactions, such that the new $x' = 1$ corresponds to the remaining momentum rather than the original beam momentum, eq. (211). But now the distributions are not only squeezed in this manner, their shapes are also changed, as follows:

- Whenever a valence quark is kicked out, the number of remaining valence quarks of that species is reduced accordingly. Thus, for a proton, the valence d distribution is completely removed if the valence d quark has been kicked out, whereas the valence u distribution is halved when one of the two is kicked out. In cases where the valence and sea u and d quark distributions are not separately provided from the PDF libraries, it is assumed that the sea is flavour-antiflavour symmetric, so that one can write e.g.

$$u(x, Q^2) = u_{\text{val}}(x, Q^2) + u_{\text{sea}}(x, Q^2) = u_{\text{val}}(x, Q^2) + \overline{u}(x, Q^2). \tag{221}$$

  The parametrized u and $\overline{u}$ distributions are then used to find the relative probability for a kicked-out u quark to be either valence or sea.
- When a sea quark is kicked out, it must leave behind a corresponding antisea parton in the beam remnant, by flavour conservation. We call this a companion quark, and bookkeep it separately from the normal sea. In the perturbative approximation the sea quark q_s and its companion q_c (not to be confused with flavour labels) come from a gluon branching $\text{g} \to \text{q}_\text{s} + \text{q}_\text{c}$, where it is implicitly understood that if q_s is a quark, q_c is its antiquark, and vice versa. This branching often would not be in the perturbative regime, but we choose to make a perturbative ansatz, and also to neglect subsequent perturbative evolution of the q_c distribution. If so, the shape of the companion q_c distribution as a function of x_c, given a sea parton at x_s, becomes

$$q_\text{c}(x_\text{c}; x_\text{s}) \propto \frac{g(x_\text{c} + x_\text{s})}{x_\text{c} + x_\text{s}} P_{\text{g} \to \text{q}_\text{s}\text{q}_\text{c}}(z), \tag{222}$$

  where

$$z = \frac{x_\text{c}}{x_\text{c} + x_\text{s}}, \tag{223}$$

$$P_{\mathrm{g}\to\mathrm{q_s q_c}}(z) \;=\; \frac{1}{2}\left(z^2+(1-z)^2\right), \tag{224}$$

$$g(x) \;\propto\; \frac{(1-x)^n}{x} \quad ;\; (n=\mathrm{MSTP}(87)) \tag{225}$$

is chosen. (Remember that this is supposed to occur at some low Q^2 scale, so $g(x)$ above should not be confused with the high-Q^2 gluon behaviour.)

- The normalization of valence and companion distributions is fixed by the respective number of quarks, i.e. the sum rules

$$\int_0^{x_{\mathrm{rem}}} q_{\mathrm{v}}(x)\,\mathrm{d}x \;=\; n_{\mathrm{q_v}}, \tag{226}$$

$$\int_0^{x_{\mathrm{rem}}} q_{\mathrm{c},i}(x;x_{\mathrm{s},i})\,\mathrm{d}x \;=\; 1 \quad (\text{for each } i), \tag{227}$$

where x_{rem} is the longitudinal momentum fraction left after the previous interactions and $n_{\mathrm{q_v}}$ is the number of q valence quarks remaining. Gluon and sea distributions do not have corresponding requirements. Therefore their normalization is adjusted, up or down, so as to obtain overall momentum conservation in the parton densities, i.e. to fulfil the remaining sum rule:

$$\int_0^{x_{\mathrm{rem}}} \left(\sum_{\mathrm{q}} q(x)+g(x)\right)\mathrm{d}x = x_{\mathrm{rem}} \tag{228}$$

Detailed formulae may be found in [Sjö04]. However, in that article, the sea+gluon rescaling factor, eq. (4.26), was derived assuming the momentum fraction taken by a companion quark scaled like $1/X$, where X is the total remaining momentum. As it happens, the momentum fraction, x_s, taken by the sea quark that originally gave rise to the companion quark, is already taken into account in the definition of the companion distribution. Hence a factor $(X+x_s)/X$ should be introduced for each companion distribution in eq. (4.26). With this change, the companions take up more 'room' in momentum space. In fact, in extreme cases it is possible to create so many companions in the beam remnant that the sum of their momenta is larger than allowed by momentum conservation, i.e. the rescaling factor becomes negative. Note that this occurs extremely rarely, and only when very many multiple interactions have occurred in an event. Since the companions are now taking up more room than allowed, we address the problem by scaling them (all of them) down by precisely the amount needed to restore momentum conservation.

The above parton-density strategy is not only used to pick a succession of hard interactions. Each interaction may now, unlike before, have initial- and final-state shower activity associated with it. The initial-state shower is constructed by backwards evolution, using the above parton densities. Even if the hard scattering does not involve a valence quark, say, the possibility exists that the shower will reconstruct back to one.

11.3.2 Colour Topologies

The second part of the model is to hook up the scattered partons to the beam remnants, in colours, in transverse momenta and in longitudinal momenta. The order in which these choices are made is partly intertwined, and one has to consider several special cases. In this description we only give an outline, evading many of the details. Especially the assignment of colours is physically uncertain and technically challenging.

Each of the two incoming beam hadrons can be viewed — post facto — as having consisted of a set of 'initiator' and 'remnant' partons. An initiator is the 'original' parton that starts the initial-state cascade that leads up to one of the hard interactions. Together the initiators and remnants make up the original proton, so together they carry the

proton net flavour content, and energy and momentum, and are in a colour singlet state. Therefore the remnant partons are constrained to carry 'what is left' when the initiators are removed. The one exception is that we do not attempt to conserve longitudinal momentum (and thereby energy) exactly within each incoming proton, but only for the system as a whole. The reason here is that we have put all initiators and remnants on mass shell, whereas a more proper treatment ought to have included (moderate) space-like virtualities for them. Such virtualities would have dissipated in the hard interactions, and so not survived to the final state. The current choice therefore shortcuts a number of technical details that in the end should not matter.

In an event with n multiple interactions, a corresponding set of n initiator partons is defined for each of the two incoming beams. This also defines the left-behind flavour content of the remnants. At most there could be $n + 3$ remnant partons for a proton with its 3 valence flavours, but there could also be none, if all valence quarks have been kicked out and no unmatched companions have been added. In the latter case a gluon is inserted to represent the beam remnant and carry leftover momentum, but else we do not introduce gluons in the beam remnant.

In the string model, the simplest representation of a baryon is the Y-shape topology, where each of the three valence quarks is connected via its string piece to the central junction (see section 12.2.5). This junction does not carry energy or momentum of its own, but topologically it is the carrier of the baryon number. Under normal circumstances, the legs in the Y are quite short. Interactions may deform the topology, however. The simplest example would be Deeply Inelastic Scattering, where one of the valence quarks is kicked violently, so that one of the three strings of the Y will stretch out and fragment into hadrons. In this topology the other two will remain so close to each other and to the junction that they effectively act as a single unit, a diquark.

In a hadronic interaction, a valence quark is kicked out by a coloured gluon rather than a colourless photon. This colour exchange implies that the string from the junction will no longer attach to the scattered quark but rather to some other quark or gluon, but the other two quarks still effectively form a diquark. When two or three valence quarks are kicked out, the junction motion will become more complicated, and must be considered in full. Equivalently, when all of the valence quarks have large relative momenta, by definition no two of them have a small invariant mass and hence a diquark description would be inappropriate.

Multiple valence quark interactions are rare, however. The bulk of interactions are gluonic. In this case we can imagine that the gluon originates from an unresolved emission off one of the valence quarks, i.e. that its anticolour is attached to one of the quarks and its colour attached to the junction. After the gluon is kicked out, colour lines will then connect this quark and the junction to partons in the final state of the interaction. If a second gluon is kicked out, it can, in colour space, have been located either between the quark and the first gluon, or between this gluon and the junction, or between the junction and one of the other two quarks. (We do not include the possibility that this gluon together with the first one could have been radiated from the system as an overall colour singlet system, i.e. we do not (yet) address diffractive event topologies in this framework.) If all of these possibilities had equal probability, the junction would often have two or all three legs reconnected, and the baryon number could be moved quite dramatically in the event. Empirically, this does not appear to happen (?), and furthermore it could be argued that perturbative and impact-parameter arguments both allow much of the activity to be correlated in 'hot spot' regions that leave much of the rest of the proton unaffected. Therefore a free suppression parameter is introduced, such that further gluons preferably connect to a string piece that has already been disturbed. In this way, gluons preferentially will be found on one of the three colour lines to the junction.

The order in which they appear along that line is more difficult to make statements about. So far in our description, no consideration has been given to the resulting momen-

tum picture, specified after transverse and longitudinal momenta have been picked (see below). This in general implies that strings will be stretched criss-cross in the event, if the initiator gluons are just randomly ordered along the string piece on which they sit. It is unlikely that such a scenario catches all the relevant physics. More likely is that, among all the possible colour topologies for the final state, those that correspond to the smaller total string length are favoured, all other aspects being the same. Several options have been introduced that approach this issue from different directions (see `MSTP(89)` and `MSTP(95)`). One is to attach the initiator gluons preferentially in those places that order the hard-scattering systems in rapidity (the default), another to prefer the attachments that will give rise to the smaller string lengths in the final state. A third option is to rearrange the colour flow between the final-state partons themselves, again giving preference to those rearrangements which minimize the overall string length. So far no preferred scenario has been identified.

A complication is the following. Two $\mathrm{g}+\mathrm{g} \to \mathrm{g}+\mathrm{g}$ scatterings each on their own may have a perfectly sensible colour flow. Still, when the two initial gluons on each side of the event are attached to each other and to the rest of the remnants, the resulting colour flow may become unphysical. Specifically the colour flow may 'loop back' on itself, such that a single gluon comes to form a separate colour singlet system. Such configurations are rejected and alternative colour arrangements are tried.

Another, rare, occurence is that the two junctions of the event can come to be connected to each other via two strings (graphical representation: `-j=j-`, where each dash corresponds to a string piece). Since we have not (yet) programmed a fragmentation scheme for such events, we simply reject them and generate a new event instead.

So far, interactions with sea quarks have not been mentioned, either a quark-antiquark pair that both scatter or a single sea scatterer with a leftover companion quark in the remnant. However, we have already argued that each such pair can be viewed as coming from the branching of some initial nonperturbative gluon. This gluon can now be attached to the beam Y topology just like the gluons considered above, and therefore does not introduce any new degrees of freedom. When then the colours are traced, it could well happen that a companion quark together with two remaining valence quarks form a separate colour singlet system. It is then likely that this system will be of low mass and collapse to a single baryon. Such possibilities are optionally allowed (see `MSTP(88)`), and correspondingly a companion antiquark could form a meson together with a single valence quark. As already mentioned, diquarks can be formed from two valence quarks.

11.3.3 Primordial $k_\perp$

Partons are expected to have primordial $k_\perp$ values of the order of a few hundred MeV from Fermi motion inside the incoming hadrons. In reality, one notes a need for a larger input in the context of shower evolution, either in parton showers or in resummation descriptions. This is not yet understood, but could e.g. represent a breakdown of the DGLAP evolution equations at small Q^2. Until a better solution has been found, we therefore have reason to consider an effective 'primordial $k_\perp$', at the level of the initiators, larger than the one above. For simplicity, a parametrized Q-dependent width

$$\sigma(Q) = \max\left(\frac{2.1\ \mathrm{GeV} \times Q}{7\ \mathrm{GeV} + Q}, \texttt{PARJ(21)}\right) \tag{229}$$

is introduced, where σ is the width of the two-dimensional Gaussian distribution of the initiator primordial $k_\perp$ (so that $\langle k_\perp^2 \rangle = \sigma^2$), Q is the scale of the hard interaction and `PARJ(21)` is the standard fragmentation $p_\perp$ width. The remnant partons correspond to $Q = 0$ and thus hit the lower limit. Apart from the selection of each individual $k_\perp$, there is also the requirement that the total $k_\perp$ of a beam adds up to zero. Different strategies

can be used here, from sharing the recoil of one parton uniformly among all other initiator and remnant partons, to only sharing it among those initiator/remnant partons that have been assigned as nearest neighbours in colour space.

11.3.4 Beam-Remnant Kinematics

The longitudinal momenta of the initiator partons have been defined by the x values picked. The remaining longitudinal momentum is shared between the remnants in accordance with their character. A valence quark receives an x picked at random according to a small-Q^2 valence-like parton density, proportional to $(1-x)^a/\sqrt{x}$, where $a = 2$ for a u quark in a proton and $a = 3.5$ for a d quark. A sea quark must be the companion of one of the initiator quarks, and can have an x picked according to the $q_{\mathrm{c}}(x_{\mathrm{c}}; x_{\mathrm{s}})$ distribution introduced above. A diquark would obtain an x given by the sum of its constituent quarks. (But with the possibility to enhance this x, to reflect the extra gluon cloud that could go with such a bigger composite object.) If a baryon or meson is in the remnant, its x is equated with the z value obtainable from the Lund symmetric fragmentation function, again with the possibility of enhancing this x as for a diquark. A gluon only appears in an otherwise empty remnant, and can thus be given $x = 1$. Once x values have been picked for each of the remnants, an overall rescaling is performed such that the remnants together carry the desired longitudinal momentum.

11.4 Multiple Interactions (and Beam Remnants) – New Model

Each multiple interaction is associated with its set of initial- and final-state radiation. In the old model such radiation was only considered for the first, i.e. hardest, interaction. The technical reason had to do with the inability to handle junction string topologies, and therefore the need to simplify the description. In practice, it could be argued that the subsequent interactions would tend to be soft, near the lower $p_{\perp\mathrm{min}}$ or $p_{\perp 0}$ scales, and therefore not be associated with additional hard radiation. Nevertheless it was a limitation.

The new junction and beam-remnant description allows radiation to be associated with each interactions. In the intermediate model, this is done in a disjoint manner: for each interaction, all initial- and final-state radiation activity associated with it is considered in full before the next interaction is selected and the showers are still the old, virtuality-ordered ones.

In the new model, the new, transverse-momentum-ordered showers are introduced. Thus $p_\perp$ becomes the common evolution scale both for multiple interactions (MI), initial-state radiation (ISR) and final-state radiation (FSR), although the technical definition of transverse momentum is slightly different in the three cases.

One can argue that, to a good approximation, the addition of FSR can be deferred until after ISR and MI have been considered in full. Specifically, FSR does not modify the total amount of energy carried by perturbatively defined partons, it only redistributes that energy among more partons. By contrast, both the addition of a further ISR branching and the addition of a further interaction implies more perturbative energy, taken from the limited beam-remnants reservoir. These two mechanisms therefore are in direct competition with each other.

We have already advocated in favour of ordering multiple interactions in a sequence of falling $p_\perp$ values. This does not correspond to an ordering in a physical time, but rather to the requirement that the hardest interaction should be the one for which standard parton densities should apply. Any corrections, e.g. from the kind of flavour correlations already discussed, would have to be introduced by approximate prescriptions, that would become increasingly uncertain as further interactions are considered.

We now advocate that, by the same reasoning, also ISR emissions should be interleaved with the MI chain, in one common sequence of decreasing $p_\perp$ values. That is, a hard second interaction should be considered before a soft ISR branching associated with the hardest interaction. This is made possible by the adoption of $p_\perp$ as common evolution variable. Thus, the standard parton densities are only used to describe the hardest interaction and the ISR branchings that occur *above* the $p_\perp$-scales of any secondary interactions.

In passing, note that the old showers requires two matching parameters, $Q^2_{\mathrm{max,shower}} = f\, p^2_{\perp,\mathrm{MI}}$. These f values, typically in the range 1 to 4, separate for space-like and time-like showers, are there to compensate on the average for the extra z-dependent factors in the relations $p_\perp^2 \approx (1-z)Q^2$ and $p_\perp^2 \approx z(1-z)Q^2$, respectively, so that the showers can start from a $p_\perp$ scale comparable with that of the interaction. In our new model, with $Q^2_{\mathrm{shower}} \approx p_\perp^2$, this matching is automatic, i.e. $f = 1$.

The evolution downwards in $p_\perp$ can now be viewed as a generalization of the backwards evolution philosophy [Sjö85]: given a configuration at some $p_\perp$ resolution scale, what can that configuration have come from at a lower scale? Technically, starting from a hard interaction, a common sequence of subsequent evolution steps — interactions and branchings mixed — can therefore be found. Assuming that the latest step occurred at some $p_{\perp i-1}$ scale, this sets the maximum $p_{\perp\mathrm{max}} = p_{\perp i-1}$ for the continued evolution. What can happen next is then either a new interaction or a new ISR branching on one of the two incoming sides in one of the existing interactions. The probability distribution for $p_\perp = p_{\perp i}$ is given by

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}p_\perp} = \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MI}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p_\perp}\right) \exp\left(-\int_{p_\perp}^{p_{\perp i-1}} \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MI}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p'_\perp}\right) \mathrm{d}p'_\perp\right) \tag{230}$$

in simplified notation. Technically, the $p_{\perp i}$ can be found by selecting a new trial interaction according to $\mathrm{d}\mathcal{P}_{\mathrm{MI}} \exp(-\int \mathrm{d}\mathcal{P}_{\mathrm{MI}})$, and a trial ISR branching in each of the possible places according to $\mathrm{d}\mathcal{P}_{\mathrm{ISR}} \exp(-\int \mathrm{d}\mathcal{P}_{\mathrm{ISR}})$. The one of all of these possibilities that occurs at the largest $p_\perp$ preempts the others, and is allowed to be realized. The whole process is iterated, until a lower cutoff is reached, below which no further interactions or branchings are allowed.

If there were no flavour and momentum constraints linking the different subsystems, it is easy to see that such an interleaved evolution actually is equivalent to considering the ISR of each interaction in full before moving on to the next interaction. Competition is introduced via the correlated parton densities already discussed. Thus distributions are squeezed to be nonvanishing in a range $x \in [0, X]$, where $X < 1$ represents the fraction of the original beam-remnant momentum still available for an interaction or branching. When a trial n'th interaction is considered, $X = 1 - \sum_{i=1}^{n-1} x_i$, where the sum runs over all the already existing interactions. The x_i are the respective momentum fractions of the ISR shower initiators at the current resolution scale, i.e., an x_i is increased each time an ISR branching is backwards-constructed on an incoming parton leg. Similarly, the flavour content is modified to take into account the partons already extracted by the $n-1$ previous interactions, including the effects of ISR branchings. When instead a trial shower branching is considered, the X sum excludes the interaction under consideration, since this energy *is* at the disposal of the interaction, and similarly for the flavour content.

The choice of $p_{\perp\mathrm{max}}$ scale for this combined evolution is process-dependent, as before. For minimum-bias QCD events the full phase space is allowed, while the $p_\perp$ scale of a QCD hard process sets the maximum for the continued evolution, in order not to doublecount. When the hard process represents a possibility not present in the MI/ISR machinery — production of Z^0, top, or supersymmetry, say — again the full (remaining) phase space is available, though for processes that contain explicit jets in the matrix element, such as W+jet, the ISR evolution is restricted to be below the jet cutoff. Note that, when interfacing events from an external matrix element generator, special care has to be taken to ensure that these scales are set consistently.

There is also the matter of a lower $p_{\perp\min}$ scale. Customarily such scales are chosen separately for ISR and MI, and typically lower for the former than the latter. Both cutoffs are related to the resolution of the incoming hadronic wave function, however, and in the current formalism ISR and MI are interleaved, so it makes sense to use the same regularization procedure.Therefore also the branching probability is smoothly turned off at a $p_{\perp 0}$ scale, like for MI, but the ISR suppression factor is the square root of the MI one, since only one Feynman vertex is involved in a shower branching relative to the two of a hard process. Thus the $\alpha_s(p_\perp^2)\,\mathrm{d}p_\perp^2/p_\perp^2$ divergence in a branching is tamed to $\alpha_s(p_{\perp 0}^2 + p_\perp^2)\,\mathrm{d}p_\perp^2/(p_{\perp 0}^2 + p_\perp^2)$. The scale of parton densities in ISR and MI alike is maintained at $p_\perp^2$, however, the argument being that the actual evolution of the partonic content is given by standard DGLAP evolution, and that it is only when this content is to be resolved that a dampening is to be imposed. This also has the boon that flavour thresholds appear where they are expected.

The cutoff for FSR is still kept separate and lower, since that scale deals with the matching between perturbative physics and the nonperturbative hadronization at long time scales, and so has a somewhat different function.

The description of beam remnants offers the same scenarios as outlined for the intermediate model above, and as further described in [Sjö04a]. A more recent option, not found there, is the following 'colour annealing' scenario [San05], available only for the new model, through the options `MSTP(95) = 2 - 5`. It has been constructed to produce similar effects as 'Tune A' and similar tunes of the old model, but is also in some sense intended to be representative of the 'most extreme' case. It starts from the assumption that, at hadronization time, no information from the perturbative colour history of the event is relevant. Instead, what determines how hadronizing strings form between the partons is a minimization of the total potential energy stored in these strings. That is, the partons, regardless of their formation history, will tend to be colour connected to the partons closest to them in momentum space, hence minimizing the total 'string length', as measured by the so-called λ measure [And83a, Sjö03]. Technically, the model implementation starts by erasing the colour tags of all final-state coloured partons. It then begins an iterative procedure (which unfortunately can be quite time-consuming):

1. Loop over all final-state coloured partons
2. For each such parton,
   (i) compute the λ measure for each possible string connection from that parton to other 'colour-compatible' final-state partons which do not already have string pieces connected to them (for `MSTP(95) = 2` and `MSTP(95) = 3` with the extra condition that closed gluon loops are suppressed, or, with options `MSTP(95) = 6` and `MSTP(95) = 7` available from version 6.402, with the condition that connections must be initiated from free triplets), and
   (ii) store the connection with the smallest λ measure for later comparison.
3. Compare all the possible 'minimal string pieces' found, one for each parton. Select the largest of these to be carried out physically. (That parton is in some sense the one that is currently furthest away from all other partons.)
4. If any 'dangling colour tags' are left, repeat from 1.
5. At the end of the iteration, it is possible that the last parton is a gluon, and that all other partons already form a complete colour singlet system. In this case, the gluon is simply attached between the two partons where its presence will increase the total λ measure the least.

Finally, let it be re-emphasized that the issue of colour correllations and reconnections, especially in hadron collisions, is an extremely challenging one, about which very little is known for certain at present (see e.g. the discussions in [San05] and references therein). The present models are therefore in this respect best regarded as vast simplifications of a presumably much more complex physical picture.

11.4.1 Joined Interactions

When the backwards evolution of initial-state radiation traces the prehistory to hard interactions, two partons participating in two separate hard scatterings may turn out to have a common ancestor, *joined interactions* (JI).

The joined interactions are well-known in the context of the evolution of multiparton densities [Kon79], but have not been applied to a multiple-interactions framework. A full implementation of the complete kinematics, intertwining MI, ISR and JI all possible ways, is a major undertaking, worth the effort only if the expected effects are non-negligible. Given the many uncertainties in all the other processes at play, one would otherwise expect that the general tuning of MI/ISR/FSR/... to data would hide the effects of JI.

The current program can simulate the joining term in the evolution equations, and thereby estimate how often and at what $p_\perp$ values joinings should occur. However, the actual kinematics has not been worked out, so the suggested joinings are never performed. Instead the evolution is continued as if nothing had happened. Therefore this facility is more for general guidance than for detailed studies.

To see how it works, define the two-parton density $f_{bc}^{(2)}(x_b, x_c, Q^2)$ as the probability to have a parton b at energy fraction x_b and a parton c at energy fraction x_c when the proton is probed at a scale Q^2. The evolution equation for this distribution is

$$\begin{aligned} \mathrm{d}f_{bc}^{(2)}(x_b, x_c, Q^2) &= \frac{\mathrm{d}Q^2}{Q^2}\frac{\alpha_\mathrm{s}}{2\pi}\iint \mathrm{d}x_a\,\mathrm{d}z\,\Big\{ f_{ac}^{(2)}(x_a, x_c, Q^2)\,P_{a\to bd}(z)\,\delta(x_b - z x_a) \\ &\quad + f_{ba}^{(2)}(x_b, x_a, Q^2)\,P_{a\to cd}(z)\,\delta(x_c - z x_a) \\ &\quad + f_a(x_a, Q^2)\,P_{a\to bc}(z)\,\delta(x_b - z x_a)\,\delta(x_c - (1-z)x_a)\Big\}\ . \end{aligned} \tag{231}$$

As usual, we assume implicit summation over the allowed flavour combinations. The first two terms in the above expression are the standard ones, where b and c evolve independently, up to flavour and momentum conservation constraints, and are already taken into account in the ISR framework. It is the last term that describes the new possibility of two evolution chains having a common ancestry.

Rewriting this into a backwards-evolution probability, the last term gives

$$\begin{aligned} \mathrm{d}\mathcal{P}_{bc}(x_b, x_c, Q^2) &= \left|\frac{\mathrm{d}Q^2}{Q^2}\right| \frac{\alpha_\mathrm{s}}{2\pi}\,\frac{x_a f_a(x_a, Q^2)}{x_b x_c f_{bc}^{(2)}(x_b, x_c, Q^2)}\,z(1-z)P_{a\to bc}(z) \\ &\simeq \left|\frac{\mathrm{d}Q^2}{Q^2}\right| \frac{\alpha_\mathrm{s}}{2\pi}\,\frac{x_a f_a(x_a, Q^2)}{x_b f_b(x_b, Q^2)\,x_c f_c(x_c, Q^2)}\,z(1-z)P_{a\to bc}(z)\ , \end{aligned} \tag{232}$$

introducing the approximation $f_{bc}^{(2)}(x_b, x_c, Q^2) \simeq f_b(x_b, Q^2)\,f_c(x_c, Q^2)$ to put the equation in terms of more familiar quantities. Just like for MI and normal ISR, a form factor is given by integration over the relevant Q^2 range and exponentiation. In practice, eq. (230) is complemented by one more term given by the above probability.

11.5 Pile-up Events

In high-luminosity colliders, there is a non-negligible probability that one single bunch crossing may produce several separate events, so-called pile-up events. This in particular applies to future pp colliders like LHC, but one could also consider e.g. $\mathrm{e^+e^-}$ colliders with high rates of $\gamma\gamma$ collisions. The program therefore contains an option, currently only applicable to hadron–hadron collisions, wherein several events may be generated and put one after the other in the event record, to simulate the full amount of particle production a detector might be facing.

The program needs to know the assumed luminosity per bunch–bunch crossing, expressed in mb^{-1}. Multiplied by the cross section for pile-up processes studied, σ_{pile}, this gives the average number of collisions per beam crossing, $\overline{n}$. These pile-up events are taken to be of the minimum-bias type, with diffractive and elastic events included or not (and a further subdivision of diffractive events into single and double). This means that σ_{pile} may be either σ_{tot}, $\sigma_{\mathrm{tot}} - \sigma_{\mathrm{el}}$ or $\sigma_{\mathrm{tot}} - \sigma_{\mathrm{el}} - \sigma_{\mathrm{diffr}}$. Which option to choose depends on the detector: most detectors would not be able to observe elastic pp scattering, and therefore it would be superfluous to generate that kind of events. In addition, we allow for the possibility that one interaction may be of a rare kind, selected freely by you. There is no option to generate two 'rare' events in the same crossing; normally the likelihood for that kind of occurrences should be small.

If only minimum-bias type events are generated, i.e. if only one cross section is involved in the problem, then the number of events in a crossing is distributed according to a Poisson with the average number $\overline{n}$ as calculated above. The program actually will simulate only those beam crossings where at least one event occurs, i.e. not consider the fraction $\exp(-\overline{n})$ of zero-event crossings. Therefore the actually generated average number of pile-up events is $\langle n \rangle = \overline{n}/(1 - \exp(-\overline{n}))$.

Now instead consider the other extreme, where one event is supposed be rare, with a cross section σ_{rare} much smaller than σ_{pile}, i.e. $f \equiv \sigma_{\mathrm{rare}}/\sigma_{\mathrm{pile}} \ll 1$. The probability that a bunch crossing will give i events, whereof one of the rare kind, now is

$$\mathcal{P}_i = f\, i\, \exp(-\overline{n})\, \frac{\overline{n}^i}{i!} = f\, \overline{n} \exp(-\overline{n})\, \frac{\overline{n}^{i-1}}{(i-1)!}\ . \tag{233}$$

The naïve Poisson is suppressed by a factor f, since one of the events is rare rather than of the normal kind, but enhanced by a factor i, since any one of the i events may be the rare one. As the equality shows, the probability distribution is now a Poisson in $i - 1$: in a beam crossing which produces one rare event, the multiplicity of additional pile-up events is distributed according to a Poisson with average number $\overline{n}$. The total average number of events thus is $\langle n \rangle = \overline{n} + 1$.

Clearly, for processes with intermediate cross sections, $\overline{n}\, \sigma_{\mathrm{rare}}/\sigma_{\mathrm{pile}} \simeq 1$, also the average number of events will be intermediate, and it is not allowed to assume only one event to be of the 'rare' type. We do not consider that kind of situations.

Each pileup event can be assigned a separate collision vertex within the envelope provided by the colliding beams, see `MSTP(151)`. Only simple Gaussian shapes in space and time are implemented internally, however. If this is too restrictive, you would have to assign interaction points yourself, and then shift each event separately by the required amount in space and time.

When the pile-up option is used, one main limitation is that event records may become very large when several events are put one after the other, so that the space limit in the `PYJETS` common block is reached. It is possible to expand the dimension of the common block, see `MSTU(4)` and `MSTU(5)`, but only up to about 20 000 entries, which may not always be enough, especially not for LHC. Simplifications like switching off π^0 decay may help keep down the size, but also has its limits.

For practical reasons, the program will only allow an $\overline{n}$ up to 120. The multiplicity distribution is truncated above 200, or when the probability for a multiplicity has fallen below 10^{-6}, whichever occurs sooner. Also low multiplicities with probabilities below 10^{-6} are truncated.

11.6 Common-Block Variables and Routines

Of the routines used to generate beam remnants, multiple interactions and pile-up events, none are intended to be called directly by the user. The only way to regulate these aspects