LU TP 09-31
MCnet/09/17
November 2009

Monte Carlo Tools*

T. Sjöstrand†

Department of Theoretical Physics,
Lund University,
Sölvegatan 14A,
S-223 62 Lund, Sweden

arXiv:0911.5286v1 [hep-ph] 27 Nov 2009

1 Introduction

Given the current landscape in experimental high-energy physics, these lectures are focused on applications of event generators for hadron colliders like the Tevatron and LHC. Much of it would also be relevant for e^+e^- machines like LEP, ILC and CLIC, or $e^\pm p$ machines like HERA, but with some differences not discussed here. Heavy-ion physics is not at all addressed, since it involves rather different aspects, specifically the potential formation of a quark–gluon plasma. Further, within the field of high-energy $pp/p\overline{p}$ collisions, the emphasis will be on the common aspects of QCD physics that occurs in all collisions, rather on those aspect that are specific to a particular physics topic, such as B production or supersymmetry. Many of these topics are instead covered by other lectures at this school.

Section 2 contains a first overview of the physics picture and the generator landscape. Thereafter section 3 describes the usage of matrix elements, section 4 the important topics of initial- and final-state showers, and section 5 how showers can be matched to different hard processes. The issue of multiparton interactions and their role in mimimum-bias and underlying-event physics is introduced in section 6, followed by some comments on hadronization in section 7. The article concludes with an outlook on the ongoing generator-development work in section 8.

Slides for these and other similar lectures (Sjöstrand 2009, MCnet 2007, CTEQ–MCnet 2008, MCnet 2009) are complementary to this writeup in style and contents, including many (colour) illustrations absent here.

*To appear in the proceedings of the 65th Scottish Universities Summer School in Physics: LHC Physics, St Andrews, 16 – 29 August 2009.

†torbjorn@thep.lu.se

2 Overview

In real life, machines produce events that are stored by the data acquisition system of a detector. In the virtual reality, event generators like HERWIG (Corcella et al 2001) and PYTHIA (Sjöstrand et al 2006) play the role of machines like the Tevatron and LHC, and detector simulation programs like GEANT 4 the role of detectors like ATLAS or CMS. The real and virtual worlds can share the same event reconstruction framework and subsequent physics analysis. It is by understanding how an original physics input is distorted step-by-step in the better-controlled virtual world that an understanding can be gained of what may be going on in the real world. For approximate studies the detector simulation and reconstruction steps can be shortcut, so that generators can be used directly in the physics studies.

A number of physics analyses would not be feasible without generators. Specifically, a proper understanding of the (potential) signal and background processes is important to separate the two. The key aspect of generators here is that they provide a detailed description of the final state so that, ideally, any experimental observable or combination of observables can be predicted and compared with data. Thereby generators can be used at various stages of an experiment: when optimizing the detector and its trigger design to the intended physics program, when estimating the feasibility of a specific physics study, when devising analysis strategies, when evaluating acceptance corrections, and so on.

However, it should always be kept in mind that generators are not perfect. They suffer from having to describe a broad range of physics, some of which is known from first principles, while other parts are modelled in different frameworks. (In the latter case, a generator actually acts as a vehicle of ideology, where ideas are disseminated in prepackaged form from theorists to experimentalists.) Given the limited resources, different authors may also have invested more or less time on specific physics topics, and therefore these may be more or less well modelled. It always pays to compare several approaches before drawing too definite conclusions. Blind usage of a generator is not encouraged: then you are the slave rather than the master.

Why then *Monte Carlo* event generators? Basically because Einstein was wrong: God does throw dice! In quantum meachanics, calculations provide the *probability* for different outcomes of a measurement. Event-by-event, it is impossible to know beforehand what will happen: anything that is at all allowed could be next. It is only when averaging over large event samples that the expected probability distributions emerge — provided we did the right calculation to high enough accuracy. In generators, (pseuo)random numbers are used to make choices intended to reproduce the quantum mechanical probabilities for different outcomes at various stages of the process.

The buildup of the structure in an event occurs in several steps, and can be summarized as follows:

- Initially two hadrons are coming in on a collision course. Each hadron can be viewed as a bag of partons — quarks and gluons.

- A collision between two partons, one from each side, gives the hard process of interest, be it for physics within or beyond the Standard Model: $ug \to ug$, $u\overline{d} \to W^+$, $gg \to h^0$, etc. (Actually, the bulk of the cross section results in rather mundane events, with

at most rather soft jets. Such events usually are filtered away at an early stage, however.)

- When short-lived "resonances" are produced in the hard process, such as the top, $W^{\pm}$ or Z^0, their decay has to be viewed as part of this process itself, since e.g. spin correlations are transferred from the production to the decay stages.

- A collision implies accelerated colour (and often electromagnetic) charges, and thereby bremsstrahlung can occur. Emissions that can be associated with the two incoming colliding partons are called Initial-State Radiation (ISR). As we shall see, such emissions can be modelled by so-called space-like parton showers.

- Emissions that can be associated with outgoing partons are instead called Final-State Radiation (FSR), and can be approximated be time-like parton showers. Often the distinction between a hard process and ISR and FSR is ambiguous, as we shall see.

- So far we only extracted one parton from each incoming hadron to undergo a hard collision. But the hadron is made up of a multitude of further partons, and so further parton pairs may collide within one single hadron–hadron collision — multiparton interactions (MPI). (Not to be confused with pileup events, when several hadron pairs collide during a bunch–bunch crossing, but with obvious analogies.)

- Each of these further collisions also may be associated with its ISR and FSR.

- The colliding partons take a fraction of the energy of the incoming hadrons, but much of the energy remains in the beam remnants, which continue to travel essentially in the original directions. These remnants also carry colours that compensate the colour taken away by the colliding partons.

- As the partons created in the previous steps recede from each other, confinement forces become significant. The structure and evolution of these force fields cannot currently be described from first principles, so models have to be introduced. One common approach is to assume that a separate confinement field is stretched between each colour and its matching anticolour, with each gluon considered as a simple sum of a colour and an anticolour, and all colours distinguishable from each other (the $N_C \to \infty$ limit).

- Such fields can break by the production of new quark–antiquark pairs that screen the endpoint colours, and where a quark from one break (or from an endpoint) can combine with an antiquark from an adjacent break to produce a primary hadron. This process is called hadronization.

- Many of those primary hadrons are unstable and decay further at various timescales. Some are sufficiently long–lived that their decays are visible in a detector, or are (almost) stable. Thereby we have reached scales where the event-generator description has to be matched to a detector-simulation framework.

- It is only at this stage that experimental information can be obtained and used to reconstruct back what may have happened at the core of the process.

The Monte Carlo method allows these steps to be considered sequentially, and within each step to define a set of rules that can be used iteratively to construct a more and more complex state, ending with hundreds of particles moving out in different directions. Since each particle contains of the order of ten degrees of freedom (flavour, mass, momentum, production vertex, lifetime, ...), thousands of choices are involved for a typical event. The aim is to have a sufficiently realistic description of these choices that both the average behaviour and the fluctuations around this average are well decribed.

Schematically, the cross section for a range of final states is provided by

$$\sigma_{\text{final state}} = \sigma_{\text{hard process}}\, \mathcal{P}_{\text{tot,hard process}\to\text{final state}} \ ,$$

properly integrated over the relevant phase-space regions and summed over possible "paths" (of showering, hadronization, etc.) that lead from a hard process to the final state. That is, the dimensional quantities are associated with the hard process; subsequent steps are handled in a probabilistic approach.

The spectrum of event generators is very broad, from general-purpose ones to more specialized ones. HERWIG and PYTHIA are the two most commonly used among the former ones, with SHERPA (Gleisberg et al 2009) coming up. Among more specialized programs, many deal with the matrix elements for some specific set of processes, a few with topics such as parton showers or particle decays, but there are e.g. no freestanding programs that handle hadronization. In the end, many of the specialized programs are therefore used as "plugins" to the general-purpose ones.

3 Matrix elements and their usage

From the Lagrangian of a theory the Feynman rules can be derived, and from them matrix elements can be calculated. Combined with phase space it allows the calculation of cross sections. As a simple example consider the scattering of quarks in QCD, say $u(1)\, d(2) \to u(3)\, d(4)$, a process similar to Rutherford scattering but with gluon exchange instead of photon ditto. The Mandelstam variables are defined as $\hat{s} = (p_1 + p_2)^2$, $\hat{t} = (p_1 - p_3)^2$ and $\hat{u} = (p_1 - p_4)^2$. In the cm frame of the collision $\hat{s}$ is the squared total energy and $\hat{t}, \hat{u} = -\hat{s}(1 \mp \cos\hat{\theta})/2$ where $\hat{\theta}$ is the scattering angle. The differential cross section is then

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\hat{t}} = \frac{\pi}{\hat{s}^2}\,\frac{4}{9}\,\alpha_{\mathrm{s}}^2\,\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} \ ,$$

which diverges roughly like $\mathrm{d}p_\perp^2/p_\perp^4$ for transverse momentum $p_\perp \to 0$. We will come back to this issue when discussing multiparton interactions; for now suffice to say that some lower cutoff $p_{\perp\min}$ need to be introduced. Similar cross sections, differing mainly by colour factors, are obtained for $q\, g \to q\, g$ and $g\, g \to g\, g$. A few further QCD graphs, like $g\, g \to q\, \overline{q}$, are less singular and give smaller contributions. These cross sections then have to be convoluted with the flux of the incoming partons i and j in the two incoming hadrons A and B:

$$\sigma = \sum_{i,j} \iiint \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}\hat{t}\, f_i^{(A)}(x_1, Q^2)\, f_j^{(B)}(x_2, Q^2)\, \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}\hat{t}} \ . \tag{1}$$

The parton density functions (PDFs) of gluons and sea quarks are strongly peaked at small momentum fractions $x_1 \approx E_i/E_A$, $x_2 \approx E_j/E_B$. This further enhances the peaking of the cross section at small $p_\perp$ values.

In order to address the physics of interest a large number of processes, both within the Standard Model and in various extensions of it, have to be available in generators. Indeed many can also be found in the general-purpose ones, but by far not enough. Further, often processes are there available only to lowest order, while experimental interest may be in higher orders, with more jets in the final state, either as a signal or as a potential background. So a wide spectrum of matrix-element-centered programs are available, some quite specialized and others more generic.

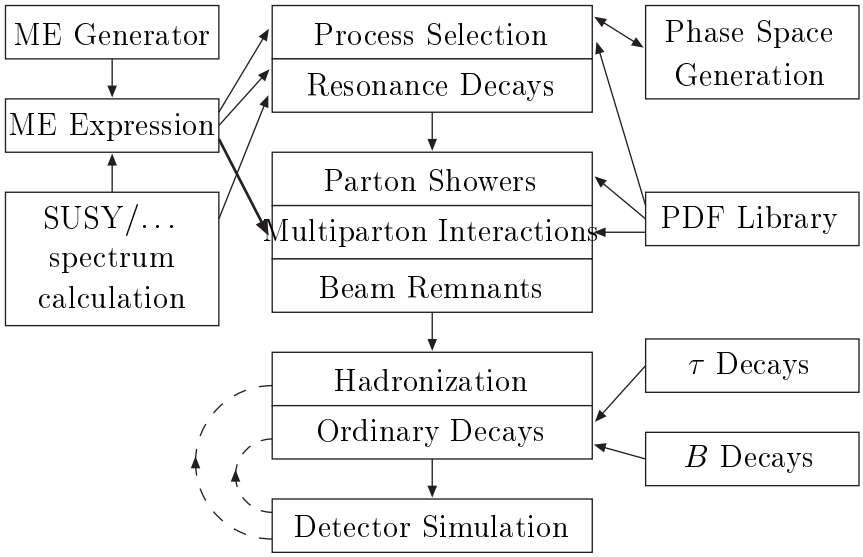

Figure 1: Example how different programs can be combined in the event-generation chain.

The way these programs can be combined with a general-purpose generator is illustrated in Fig. 1. In the study of Supersymmetry (SUSY) it is customary to define a model in terms of a handful parameters, e.g. specified at some large Grand Unification scale. It is then the task of a spectrum calculator to turn this into a set of masses, mixings and couplings for the physical states to be searched for. Separately, the matrix elements can be calculated with these properties as unknown parameters, and only when the two are combined is it possible to speak of physically relevant matrix-element expressions. These matrix elements now need to be combined with PDFs and sampled in phase space, preferable with some preweighting procedure so that regions with high cross sections are sampled more frequently. The primarily-produced SUSY particles typically are unstable and undergo sequential decays down to a lightest supersymmetric particle (LSP), again with branching ratios and angular distributions that should be properly modelled. The LSP would be neutral and escape undetected, while other decay products would be normal quarks and leptons.

It is at this stage that general-purpose programs take over. They describe the showering associated with the above process, the presence of additional interactions in the same hadron–hadron collision, the structure of beam remnants, and the hadronization and decays. They would still rely on the externally supplied PDFs, and potentially make use of programs dedicated to τ and B decays, where spin information and form factors require special encoding. Even after the event has been handed on to the detector-simulation program some parts of the generator may be used in the simulation of secondary interactions and

decays.

Several standards have been developed to further this interoperability. The Les Houches Accord (LHA) for user processes (Boos et al 2001) specifies how parton-level information about the hard process and sequential decays can be encoded and passed on to a general-purpose generator. Originally it was defined in terms of two Fortran commonblocks, but more recently a standard Les Houches Event File format (Alwall et al 2007) offfers a language-independent alternative approach. The Les Houches Accord Parton Density Functions (LHAPDF) library (Borilkov et al 2006) makes different PDF sets available in a uniform framework. The SUSY Les Houches Accord (SLHA) (Skands et al 2004, Allanach et al 2009) allows a standardized transfer of masses, mixings, couplings and branching ratios from spectrum calculators to other programs. The HepMC C++ event record (Dobbs and Hansen 2001) succeeds the HEPEVT Fortran one as a standard way to transfer information from a generator on to the detector-simulation stage. One of the key building blocks for several of these standards is the PDG codes for all the most common particles (Amsler et al 2008), also in some scenarios for physics beyond the Standard Model.

The $2 \to 2$ processes we started out with above are about the simplest one can imagine at a hadron collider. In reality one needs to go on to higher orders. In $\mathcal{O}(\alpha_s^3)$ two new kind of graphs enter. One kind is where one additional parton is present in the final state, i.e. $2 \to 3$ processes. The cross section for such processes is almost always divergent when one of the parton energies vanish (soft singularities) or two partons become collinear (collinear singularities). The other kind is loop graphs, with an additional intermediate parton not present in the final state, i.e. a correction to the $2 \to 2$ processes. This gives negative divergences that exactly cancel the positive ones above, with only finite terms surviving. For inclusive event properties such next-to-leading order (NLO) calculations lead to an improved accuracy of predictions, but for more exclusive studies the mathematical cancellation of singularities has to be supplemented by more physical techniques, which is far from trivial.

The tricky part of the calculations is the virtual corrections. NLO is now state-of-the-art, with NNLO still in its infancy. If one is content with Born-level diagrams only, i.e. without any loops, it is possible to go to quite high orders, with eight or more partons in the final state. These partons have to be kept well separated to avoid the phase-space regions where the divergences become troublesome. In order to cover also regions where partons become soft/colliner we therefore next turn our attention to parton showers.

4 Parton showers

As already noted, the emission rate for a branching such as $q \to qg$ diverges when the gluon either becomes collinear with the quark or when the gluon energy vanishes. The QCD pattern is similar to that for $e \to e\gamma$ in QED, except with a larger coupling, and a coupling that increases for smaller relative $p_\perp$ in a branching, thereby further enhancing the divergence. Furthermore the non-Abelian character of QCD leads to $g \to gg$ branchings with similar divergences, without any correspondence in QED. The third main branching, $g \to q\overline{q}$ with its $\gamma \to e^+e^-$ QED equivalence, does not have the soft divergence and is less important.

Now, if the rate for one emission of a gluon is big, then also the rate for two or more will

be big, and thus the need for high orders. With showers we introduce two new concepts that make like easier,
(1) an iterative structure that allows simple expressions for $q \to qg$, $g \to gg$ and $g \to q\overline{q}$ branchings to be combined to build up complex multiparton final states, and
(2) a Sudakov factor that offers a physical way to handle the cancellation between real and virtual divergences.
Neither of the simplifications is exact, but together they allow us to provide sensible approximate answers for the structure of emissions in soft and collinear regions of phase space.

4.1 The shower approach

The starting point is to "factorize" a complex $2 \to n$ process, where n represents a large number of partons in the final state, into a simple core process, maybe $2 \to 2$, convoluted with showers, Fig. 2. To begin with, in a simple $ud \to ud$ process the incoming and outgoing quarks must be on the mass shell, i.e. satisfy $p^2 = E^2 - \mathbf{p}^2 = m_q^2 \sim 0$, at long timescales. By the uncertainty principle, however, the closer one comes to the hard interaction, i.e the shorter the timescales considered, the more off-shell the partons may be.

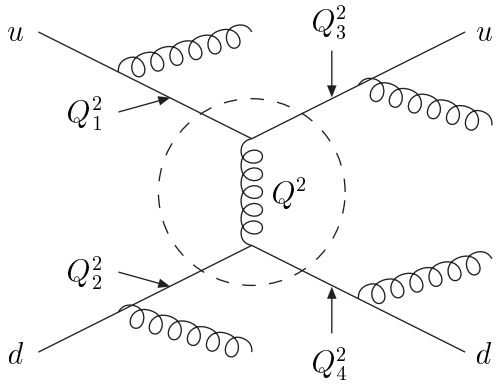

Figure 2: The "factorization" of a $2 \to n$ process.

Thus the incoming quarks may radiate a succession of harder and harder gluons, while the outgoing ones radiate softer and softer gluons. One definition of hardness is how off-shell the quarks are, $Q^2 \sim |p^2| = |E^2 - \mathbf{p}^2|$, but we will encounter other variants later. In the initial-state radiation (ISR) part of the cascade these virtualities are spacelike, $p^2 < 0$, hence the alternative name spacelike showers. Correspondingly the final-state radiation (FSR) is characterized by timelike virtualities, $p^2 > 0$, and hence also called timelike showers. The difference is a consequence of the kinematics in branchings.

The cross section for the whole $2 \to n$ graph is associated with the cross section of the hard subprocess, with the approximation that the other Q_i^2 virtualities can be neglected in the matrix-element expression. In the limit that all the $Q_i^2 \ll Q^2$ this should be a good approximation. In other words, first the hard process can be picked without any reference to showers, and only thereafter are showers added with unit probability. But, of course, the showers do modify the event shape, so at the end of the day the cross section is affected. For instance, the total transverse energy $E_{\perp \mathrm{tot}}$ of an event is increased by ISR, so the cross

sections of events with a given $E_{\perp\mathrm{tot}}$ is increased by the influx of events that started out with a lower $E_{\perp\mathrm{tot}}$ in the hard process.

It is important that the hard-process scale Q^2 is picked to be the largest one, i.e. $Q^2 > Q_i^2$ in Fig. 2. If e.g. $Q_1^2 > Q^2$ then instead the $ug \to ug$ subgraph ought to be chosen as hard process, and the gluon of virtuality Q^2 ought to be part of the ISR off the incoming d. Without such a criterion one might doublecount a given graph.

4.2 Final-state radiation

Let us next turn to a more detailed presentation of the showering approach, and begin with the simpler final-state stage. This is most cleanly studied in the process $e^+e^- \to \gamma^*/Z^0 \to q\overline{q}$. The first-order correction here corresponds to the emission of one additional gluon, by either of the two Feynman graphs in Fig. 3. Neglect quark masses and introduce energy fractions $x_j = 2E_j/E_{\mathrm{cm}}$ in the rest frame of the process. Then the cross section is of the form

$$\frac{\mathrm{d}\sigma_{\mathrm{ME}}}{\sigma_0} = \frac{\alpha_{\mathrm{s}}}{2\pi}\,\frac{4}{3}\,\frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}\,\mathrm{d}x_1\,\mathrm{d}x_2 \ ,$$

where σ_0 is the $q\overline{q}$ cross section, i.e. without the gluon emission.

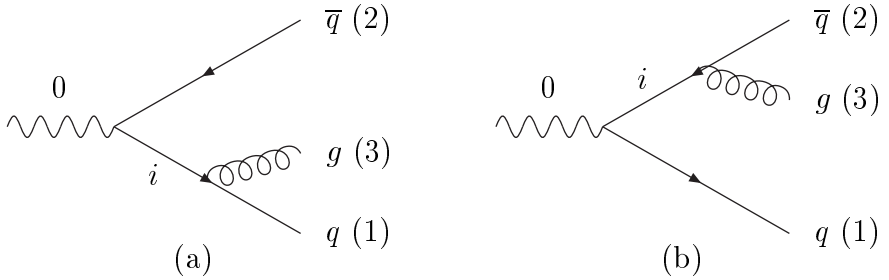

Figure 3: The two Feynman graphs that contribute to $\gamma^*/Z^0(0) \to q(1)\,\overline{q}(2)\,g(3)$

Now study the kinematics in the limit $x_2 \to 1$. Since $1 - x_2 = m_{13}^2/E_{\mathrm{cm}}^2$ we see that this corresponds to the "collinear region", where the separation between the q and g vanishes. Equivalently, the virtuality $Q^2 = Q_i^2 = m_{13}^2$ of the intermediate quark propagator i in Fig. 3a vanishes. Although the full answer contains contributions from both graphs it is obvious that, in this region, the amplitude of the one in Fig. 3a dominates over the one in Fig. 3b. We can therefore view the process as $\gamma^*/Z^0 \to q\overline{q}$ followed by $q \to qg$. Define the energy sharing in the latter branching by $E_q = zE_i$ and $E_g = (1-z)E_i$. The kinematics relations then are

$$\begin{aligned} 1 - x_2 = \frac{m_{13}^2}{E_{\mathrm{cm}}^2} = \frac{Q^2}{E_{\mathrm{cm}}^2} \quad &\Longrightarrow \quad \mathrm{d}x_2 = \frac{\mathrm{d}Q^2}{E_{\mathrm{cm}}^2} \\ x_1 \approx z \quad &\Longrightarrow \quad \mathrm{d}x_1 \approx \mathrm{d}z \\ x_3 \approx 1 - z & \end{aligned}$$

so that

$$\mathrm{d}\mathcal{P} = \frac{\mathrm{d}\sigma_{\mathrm{ME}}}{\sigma_0} = \frac{\alpha_{\mathrm{s}}}{2\pi}\,\frac{\mathrm{d}x_2}{(1-x_2)}\,\frac{4}{3}\,\frac{x_2^2 + x_1^2}{(1-x_1)}\,\mathrm{d}x_1 \approx \frac{\alpha_{\mathrm{s}}}{2\pi}\,\frac{\mathrm{d}Q^2}{Q^2}\,\frac{4}{3}\,\frac{1+z^2}{1-z}\,\mathrm{d}z \qquad (2)$$

Here $\mathrm{d}Q^2/Q^2$ corresponds to the "collinear" or "mass" singularity and $\mathrm{d}z/(1-z) = \mathrm{d}E_g/E_g$ to the soft-gluon singularity.

The interesting aspect of eq. (2) is that it is universal: whenever there is a massless quark in the final state, this equation provides the probability for the same final state except for the quark being replaced by an almost collinear qg pair (plus some other slight kinematics adjustments to conserve overall energy and momentum). That is reasonable: in a general process any number of distinct Feynman graphs may contribute and interfere in a nontrivial manner, but once we go to a collinear region only one specific graph will contribute, and that graphs always has the same structure, in this case with an intermediate quark propagator. Corresponding rules can be derived for what happens when a gluon is replaced by a collinear gg or $q\overline{q}$ pair. These rules are summarized by the DGLAP equations

$$\begin{aligned} \mathrm{d}\mathcal{P}_{a\to bc} &= \frac{\alpha_\mathrm{s}}{2\pi}\frac{\mathrm{d}Q^2}{Q^2}P_{a\to bc}(z)\,\mathrm{d}z \qquad (3)\\ \text{where} \qquad P_{q\to qg} &= \frac{4}{3}\frac{1+z^2}{1-z}\ ,\\ P_{g\to gg} &= 3\,\frac{(1-z(1-z))^2}{z(1-z)}\ ,\\ P_{g\to q\overline{q}} &= \frac{n_f}{2}\left(z^2+(1-z)^2\right) \quad (n_f = \text{no. of quark flavours})\ . \end{aligned}$$

Furthermore, the rules can be combined to allow for the successive emission in several steps, e.g. where a $q \to qg$ branching is followed by further branchings of the daughters. That way a whole shower develops.

Such a picture should be reliable in cases where the emissions are strongly ordered, i.e. $Q_1^2 \gg Q_2^2 \gg Q_3^2 \ldots$. Showers would not be useful if they only could be applied to strongly-ordered parton configurations, however. A further study of the $\gamma^*/Z^0 \to q\overline{q}g$ example shows that the simple sum of the $q \to qg$ and $\overline{q} \to \overline{q}g$ branchings reproduce the full matrix elements, with interference included, to better than a factor of 2 over the full phase space. This is one of the simpler cases, and of course one should expect the accuracy to be worse for more complicated final states. Nevertheless, it is meaningful to use the shower over the whole strictly-ordered, but not necessarily strongly-ordered, region $Q_1^2 > Q_2^2 > Q_3^2 \ldots$ to obtain an approximate answer for multiparton topologies.

We did not yet tame the fact that probabilities blow up in the soft and collinear regions. For sure, perturbation theory will cease to be meaningful at so small Q^2 scales that $\alpha_\mathrm{s}(Q^2)$ diverges; there confimenent effects and hadronization phenomena take over. Typically therefore some lower cutoff at around 1 GeV is used to regulate both soft and collinear divergences: below such a scale no further branchings are simulated. Whatever perturbative effects may remain are effectively pushed into the parameters of the nonperturbative framework. That way we avoid the singularities, but we can still have branching "probabilities" well above unity, which does not seem to make sense.

This brings us to the second big concept of this section, the *Sudakov (form) factor*. In the context of particle physics it has a specific meaning related to the properties of virtual corrections, but more generally we can just see it as a consequence of the conservation of total probability

$$\mathcal{P}(\text{nothing happens}) = 1 - \mathcal{P}(\text{something happens})\ ,$$

where the former is multiplicative in a time-evolution sense:

$$\mathcal{P}_\text{nothing}(0 < t \le T) = \mathcal{P}_\text{nothing}(0 < t \le T_1)\,\mathcal{P}_\text{nothing}(T_1 < t \le T)\ .$$

When these two are combined the end result is

$$\mathrm{d}\mathcal{P}_{\mathrm{first}}(T) = \mathrm{d}\mathcal{P}_{\mathrm{something}}(T)\ \exp\left(-\int_0^T \frac{\mathrm{d}\mathcal{P}_{\mathrm{something}}(t)}{\mathrm{d}t}\mathrm{d}t\right)\ .$$

That is, the probability for something to happen for the *first* time at time T is the naive probability for this to happen, *times* the probability that this did not yet happen.

A common example is that of radioactive decay. If the number of undecayed radioactive nuclei at time t is $\mathcal{N}(t)$, with initial number $\mathcal{N}_0$ at time $t = 0$, then a naive ansatz would be $\mathrm{d}\mathcal{N}/\mathrm{d}t = -c\mathcal{N}_0$, where c parametrizes the decay likelihood per unit of time. This equation has the solution $\mathcal{N}(t) = \mathcal{N}_0(1 - ct)$, which becomes negative for $t > 1/c$, because by then the probability for having had a decay exceeds unity. So what we made wrong was not to take into account that only an undecayed nucleus can decay, i.e. that the equation ought to have been $\mathrm{d}\mathcal{N}/\mathrm{d}t = -c\mathcal{N}(t)$ with the solution $\mathcal{N}(t) = \mathcal{N}_0 \exp(-ct)$. This is a nicely well-behaved expression, where the total probability for decays goes to unity only for $t \to \infty$. If c had not been a constant but varied in time, $c = c(t)$, it is simple to show that the solution instead would have become

$$\mathcal{N}(t) = \mathcal{N}_0 \exp\left(-\int_0^t c(t')\,\mathrm{d}t\right) \quad \Longrightarrow \quad \frac{\mathrm{d}\mathcal{N}}{\mathrm{d}t} = -c(t)\mathcal{N}_0 \exp\left(-\int_0^t c(t')\,\mathrm{d}t\right)\ .$$

For a shower the relevant "time" scale is something like $1/Q$, by the Heisenberg uncertainty principle. That is, instead of evolving to later and later times we evolve to smaller and smaller Q^2. Thereby the DGLAP eq. (3) becomes

$$\mathrm{d}\mathcal{P}_{a\to bc} = \frac{\alpha_{\mathrm{s}}}{2\pi}\frac{\mathrm{d}Q^2}{Q^2}\,P_{a\to bc}(z)\,\mathrm{d}z\ \exp\left(-\sum_{b,c}\int_{Q^2}^{Q^2_{\mathrm{max}}}\frac{\mathrm{d}Q'^2}{Q'^2}\int\frac{\alpha_{\mathrm{s}}}{2\pi}\,P_{a\to bc}(z')\,\mathrm{d}z'\right)\ ,$$

where the exponent (or simple variants thereof) is the Sudakov factor. As for the radioactive-decay example above, the inclusion of a Sudakov ensures that the total probability for a parton to branch never exceeds unity. Then you may have sequential radioactive decay chains, and you may have sequential parton branchings, but that is another story.

It is a bit deeper than that, however. Just as the standard branching expressions can be viewed as approximations to the complete matrix elements for real emission, the Sudakov is an approximation to the complete virtual corrections from loop graphs. The divergences in real and virtual emissions, so strange-looking in the matrix-element language, here naturally combine to provide a physical answer everywhere. What is not described in the shower, of course, is the non-universal finite parts of the real and virtual matrix elements.

The implementation of a cascade evolution now makes sense. Starting from a simple $q\overline{q}$ system the q and $\overline{q}$ are individually evolved downwards from some initial Q^2_{max} until they branch. At a branching the mother parton disappears and is replaced by two daughter partons, which in their turn are evolved downwards in Q^2 and may branch. Thereby the number of partons increases, until the lower cutoff scale is reached.

This does not mean that everything is uniquely specified. In particular, the choice of evolving in $Q^2 = |p^2|$ is by no means obvious. Any alternative variable $P^2 = f(z)\,Q^2$ would work equally well, since $\mathrm{d}P^2/P^2 = \mathrm{d}Q^2/Q^2$. Other evolution variables include the transverse momentum, $p_\perp^2 \approx z(1-z)m^2$, and the energy-weighted emission angle $E^2\theta^2 \approx m^2/(z(1-z))$.

Both these two alternative choices are favourable when the issue of *coherence* is introduced. Coherence means that emissions do not occur independently. For instance, consider $g_1 \to g_2\, g_3$, followed by an emission of a gluon either from 2 or 3. When this gluon is soft it cannot resolve the individual colour charges of g_2 and g_3, but only the net charge of the two, which of course is the charge of g_1. Thereby the multiplication of partons in a shower is reduced relative to naive expectations. As it turns out, evolution in $p_\perp$ or angle automatically includes this reduction, while one in virtuality does not.

In the study of FSR, e.g. at LEP, three algorithms have been commonly used. The Herwig angular-ordered and Pythia mass-ordered ones are conventional parton showers as described above, while the Ariadne (Gustafson and Pettersson 1988, Lönnblad 1992) $p_\perp$-ordered one is based on a picture of dipole emissions. That is, instead of considering $a \to b\, c$ one studies $a\, b \to c\, d\, e$. One aspect of this is that, in addition to the branching parton, Ariadne also explicitly includes a "recoil parton" needed for overall energy–momentum conservation. Additionally emissions off a and b are combined in a well-defined manner.

All three approaches have advantages and disadvantages. As already mentioned, Pythia does not inherently include coherence, but has to add that approximately by brute force. Both Pythia and Herwig break Lorentz invariance slightly. The Herwig algorithm cannot cover the full phase space with it emissions, but has to fill in some "dead zones" using higher-order matrix elements. The Ariadne dipole picture does not include $g \to q\overline{q}$ branchings in a natural way.

When all is said and done, it turns out that all three algorithms do quite a decent job of describing LEP data, but typically Ariadne does best. In recent years the $p_\perp$-ordered approach has also gained ground, having been introduced in Pythia (Sjöstrand and Skands 2005) and on its way for Sherpa (Schumann and Krauss 2008, Winter and Krauss 2008). However, also the angular-ordered showers are being further developed (Gieseke et al 2003).

4.3 Initial-state radiation

The structure of initial-state radiation (ISR) is more complicated than that of FSR, since the nontrivial structure of the incoming hadrons enter the game. A proton is made up out of three quarks, uud, plus the gluons that bind them together. This picture is not static, however: gluons are continuously emitted and absorbed by the quarks, and each gluon may in its turn temporarily split into two gluons or into a $q\overline{q}$ pair. Thus a proton is teeming with activity, and much of it in a nonperturbative region where we cannot calculate. We are therefore forced to introduce the concept of a parton density $f_b(x, Q^2)$ as an empirical distribution, describing the probability to find a parton of species b in a hadron, with a fraction x of the hadron energy–momentum when the hadron is probed at a resolution scale Q^2.

While $f_b(x, Q^2)$ itself cannot be predicted, the change of f_b with resolution scale can, once Q^2 is large enough that perturbation theory should be applicable:

$$\frac{\mathrm{d}\, f_b(x, Q^2)}{\mathrm{d}(\ln Q^2)} = \sum_a \int_x^1 \frac{\mathrm{d}z}{z}\, f_a(x', Q^2)\, \frac{\alpha_\mathrm{s}}{2\pi}\, P_{a\to bc}\left(z = \frac{x}{x'}\right) \quad . \tag{4}$$

This is actually nothing but our familiar DGLAP equations. Before they were written in an exclusive manner: given a parton a, what is the probability that it will branch to $b\, c$ during a change $\mathrm{d}Q^2$? Here the formulation is instead inclusive: given that the probability

distributions $f_a(x, Q^2)$ of all partons a are known at a scale Q^2, how is the distribution of partons b changed by the set of possible branchings $a \to b$ ($+c$, here implicit). The splitting kernels $P_{a\to bc}(z)$ are the same to leading order, but differ between ISR and FSR in higher orders. In higher orders also the concept of $f_b(x, Q^2)$ as a positive definite probability is lost, additional complications that we will not consider any further here.

Even though eqs. (3) and (4) are equivalent, the physics context is different. In FSR the outgoing partons have been kicked to large timelike virtualities by the hard process and then cascade downwards towards the mass shell. In ISR we rather start out with a simple proton at early times and then allow more and more spacelike virtualities as we get closer to the hard interaction.

So, when the hard scattering occurs, in some sense the initial-state cascade is already there, as a virtual fluctuation. Had no collision occured the fluctuation would have collapsed back, but now one of the partons of the fluctuation is kicked out in a quite different direction and can no longer recombine with its sister parton from its last branching, nor with other partons in the cascade that lead up to this particular parton. Post facto we therefore see that a chain of branchings with increasing Q^2 values built up an ISR shower, Fig. 4.

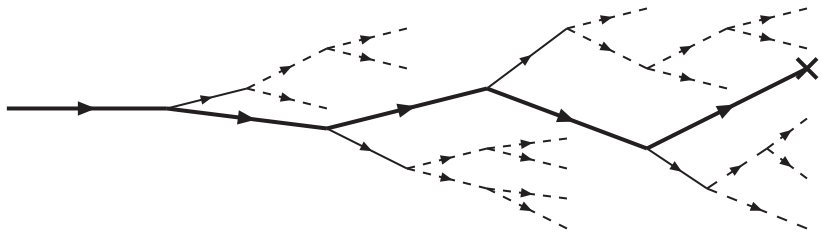

Figure 4: A cascade of successive branchings. The thick line represents the main chain of spacelike partons leading in to the hard interaction (marked by a cross). The thin lines are partons that cannot be recombined, while dashed lines are further fluctuations that may (if spacelike) or may not (if timelike) recombine. In this graph lines can represent both quarks and gluons.

The obvious way to simulate this situation would be to pick partons in the two incoming hadrons from parton densities at some low Q^2 scale, and then use the exclusive formulation of eq. (3) to construct a complete picture of partons available at higher Q^2 scales, event by event. The two sets of incoming partons could then be weighted by the cross section for the process under study. A problem is that this may not be very efficient. We have to evolve for all possible fluctuations, but at best one particular parton will collide and most of the other fluctuations will collapse back. The cost may become prohibitive when the process of interest has a constrained phase space, like a light-mass Higgs which has to have the colliding partons matched up in a very narrow mass bin.

There are ways to speed up this "forwards evolution" approach. However, the most common solution is instead to adopt a "backwards evolution" point of view. Here one starts at the hard interaction and then tries to reconstruct what happened "before". To be more precise, the cross-section formula in eq. (1) already includes the summation over all possible incoming shower histories by the usage of Q^2-dependent parton densities. Therefore what remains is to pick one exclusive shower history from the inclusive set that went into

the Q^2-evolution. To do this, recast eq. (4) as

$$\mathrm{d}\mathcal{P}_b = \frac{\mathrm{d}f_b}{f_b} = |\mathrm{d}(\ln Q^2)| \sum_a \int \mathrm{d}z \, \frac{x' f_a(x',t)}{x f_b(x,t)} \, \frac{\alpha_\mathrm{s}}{2\pi} \, P_{a\to bc}\left(z = \frac{x}{x'}\right) \; .$$

Then we have defined a *conditional probability*: if parton b is present at scale Q^2, what is the probability that it will turn out to have come from a branching $a \to bc$ at some infinitesimally *smaller* scale? (Recall that the original eq. (4) was defined for increasing virtuality.) Like for FSR this expression has to be modified by a Sudakov factor to preserve total probability, and this factor is again the exponent of the real-emission expression with a negative sign, integrated over Q^2 from an upper starting scale Q^2_max down to the Q^2 of the hypothetical branching.

The approach is now clear. First a hard scattering is selected, making use of the Q^2-evolved parton densities. Then, with the hard process as upper maximum scale, a succession of ISR branchings are reconstructed at lower and lower Q^2 scales, going "backwards in time" towards the early low-virtuality initiators of the cascades. Again some cutoff needs to be introduced when the nonperturbative regime is reached.

Unfortunately the story does not end there. For FSR we discussed the need to take into account coherence effects and the possibility to use different variables. Such issues exist here as well, but also additional ones. For instance, evolution need not be strictly ordered in Q^2, and non-ordered chains in some cases can be important. Another issue is that there can be so many partons evolving inside a hadron that they become close-packed, which leads to recombinations.

5 Combining matrix elements and parton showers

As we have seen, both matrix elements (ME) and parton showers (PS) have advantages and disadvantages.

To recall, ME allow a systematic expansion in powers of α_s, and thereby offer a controlled approach towards higher precision. Calculations can be done with several partons in the final state, so long as only Born-level results are asked for, and it is possible to tailor the phase-space cuts for these partons precisely to the experimental needs. Loop calculations are much more difficult, on the other hand, and the mathematically correct cancellation between real- and virtual-emission graphs in the soft/collinear regions is not physically sensible. Therefore ME cannot be used to explore the internal structure of a jet, and are difficult to match to hadronization models, which are supposed to take over in the very soft/collinear region.

PS, on the other hand, clearly are approximate and do not come with a guaranteed level of precision for well separated jets. You cannot steer the probabilistic evolution of a shower too much, and therefore the efficiency for obtaining events in a specific region of phase space can be quite low. On the other hand, PS are universal, so for any new model you only need to provide the basic hard process and then PS will turn that into reasonably realistic multiparton topologies. The use of Sudakov factors ensures a physically sensible behaviour in the soft/collinear regions, and it is also here that the PS formalism is supposed to be as most reliable. It is therefore possible to obtain a good picture of the internal structure of jets, and to provide a good match to hadronization models.

In a nutshell: ME are good for well separated jets, PS for the structure inside jets. Clearly the two complement each other, and a marriage is highly desirable. To do this, without doublecounting or gaps in the phase space coverage, is less trivial, and several alternative approaches have been developed. In the following we will discuss three main options: merging, vetoed parton showers and NLO matching, roughly ordered in increasing complexity. Which of these to use may well depend on the task at hand.

5.1 Merging

The aspiration of merging is to cover the whole phase space with a smooth transition from ME to PS. The typical case would be a process where the lowest-order (LO) ME is known, as well as the next-to-leading-order (NLO) real-emission one, say of an additional gluon. The shower should then reproduce

$$W^{\mathrm{ME}} = \frac{1}{\sigma(\mathrm{LO})} \; \frac{\mathrm{d}\sigma(\mathrm{LO}+g)}{\mathrm{d}(\mathrm{phasespace})} \tag{5}$$

starting from a LO topology. If the shower populates phase space according to W^{PS} this implies that a correction factor $W^{\mathrm{ME}}/W^{\mathrm{PS}}$ need to be applied.

At first glance this does not apper to make sense: if all we do is get back W^{ME}, then what did we gain? However, the trick is to recall that the PS formula comes in two parts: the real-emission answer and a Sudakov factor that ensures total conservation of probability. What we have called W^{PS} above should only be the real-emission part of the story. It is also this one that we know *will* agree with W^{ME} in the soft and collinear regions. Actually, with some moderate amount of effort it is often possible to ensure that $W^{\mathrm{ME}}/W^{\mathrm{PS}}$ is of order unity over the whole phase space, and to adjust the showers in the hard region so that the ratio always is below unity, i.e. so that standard Monte Carlo rejection techniques can be used. What the Sudakov factor then does is introduce some ordering variable Q^2, so that the whole phase space is covered starting from "hard" emissions and moving to "softer" ones. At the end of the day this leads to a distribution over phase space like

$$W^{\mathrm{PS}}_{\mathrm{actual}}(Q^2) = W^{\mathrm{ME}}(Q^2) \exp\left(- \int_{Q^2}^{Q^2_{\mathrm{max}}} W^{\mathrm{ME}}(Q'^2)\, \mathrm{d}Q'^2 \right) \; .$$

That is, we have used the PS choice of evolution variable to provide an exponentiated version of the ME answer. As such it agrees with the ME answer in the hard region, where the Sudakov factor is close to unity, and with the PS in the soft/collinear regions, where $W^{\mathrm{ME}} \approx W^{\mathrm{PS}}$.

In Pythia this approach is used for essentially all resonance decays in the Standard Model and minimal supersymmetric extensions thereof (Norrbin and Sjöstrand 2001): $\gamma^*/Z^0 \to q\overline{q}$, $t \to bW^+$, $W^+ \to u\overline{d}$, $H \to b\overline{b}$, $\chi^0 \to \tilde{q}\overline{q}$, $\tilde{q} \to q\tilde{g}$, ... It is also used in ISR to describe e.g. $q\overline{q} \to \gamma^*/Z^0/W^{\pm}$.

Merging is also used for several processes in Herwig, such as $\gamma^*/Z^0 \to q\overline{q}$, $t \to bW^+$ and $q\overline{q} \to \gamma^*/Z^0/W^{\pm}$. A special problem here is that the angular-ordered algorithms, both for FSR and for ISR, leave some "dead zones" of hard emissions that are kinematically forbidden for the shower to populate. It is therefore necessary to start directly from higher-order matrix elements in these regions. A consistent treatment still allows a smooth joining across the boundary.

5.2 Vetoed parton showers

The objective of vetoed parton showers is again to combine the real-emission behaviour of ME with the emission-ordering-variable-dependent Sudakov factors of PS. While the merging approach only works for combining the LO and NLO expressions, however, the vetoed parton showers offer a generic approach for combining several different orders.

To understand how the algorithm works, consider a lowest-order process such as $q\overline{q} \to W^{\pm}$. For each higher order one additional jet would be added to the final state, so long as only real-emission graphs are considered: in first order e.g. $q\overline{q} \to W^{\pm}g$, in second order e.g. $q\overline{q} \to W^{\pm}gg$, and so on. Call these (differential) cross sections σ_0, σ_1, σ_2, ... It should then come as no surprise that each σ_i, $i \geq 1$, contains soft and collinear divergences. We therefore need to impose some set of ME phase-space cuts, e.g. on invariant masses of parton pairs, or on parton energies and angular separation between them. When these cuts are varied, so that e.g. the mass or energy thresholds are lowered towards zero, all of these σ_i, $i \geq 1$, increase without bounds.

The reason is that in the ME approach without virtual corrections there is no "detailed balance", wherein the addition of cross section to σ_{i+1} is compensated by a depletion of σ_i. That is, if you have an event with i jets at some resolution scale, and a lowering of the minimal jet energy reveals the presence of one additional jet, then you should reclassify the event from being i-jet to being $i+1$-jet. Add one, subtract one, with no net change in $\sum_i \sigma_i$. So the trick is to use the Sudakovs of showers to ensure this detailed balance. Of course, in a complete description the cancellation between real and virtual corrections is not completely exact but leaves a finite net contribution, which is not predicted in this approach.

A few alternative algorithms exist along these lines. All share the three first steps as follows:

1) Pick a hard process within the ME-cuts-allowed phase-space region, in proportions provided by the ME integrated over the respective allowed region, $\sigma_0 : \sigma_1 : \sigma_2 : \ldots$. Use for this purpose a fix $\alpha_{\mathrm{s}0}$ larger than the α_{s} values that will be used below.

2) Reconstruct an imagined shower history that describes how the event could have evolved from the lowest-order process to the actual final state. That provides an ordering of emissions by whatever shower-evolution variable is intended.

3) The "best-bet" choice of α_{s} scale in showers is known to be the squared transverse momentum of the respective branching. Therefore a factor $W_\alpha = \prod_{\mathrm{branchings}}(\alpha_{\mathrm{s}}(p^2_{\perp i})/\alpha_{\mathrm{s}0})$, provides the probability that the event should be retained.

Now the algorithms part way. In the CKKW and L approaches the subsequent steps are:

4) Evaluate Sudakov factors for all the "propagator" lines in the shower history reconstructed in step 2, i.e. for intermediate partons that split into further partons, and also for the evolution of the final partons down to the ME cuts without any further emissions. This provides an acceptance weight $W_{\mathrm{Sud}} = \prod_{\text{“propagators”}} \mathrm{Sudakov}(Q^2_{\mathrm{beg}}, Q^2_{\mathrm{end}})$ where Q^2_{beg} is the large scale where a parton is produced by a branching and Q^2_{end} either is the scale at which the parton branches or the ME cuts, the case being.

4a) In the CKKW approach (Catani et al 2001) the Sudakovs are evaluated by analytical formulae, which is fast.

4b) In the L approach (Lönnblad 2002) trial showers are used to evaluate Sudakovs, which

is slower but allows a more precise modelling of kinematics and phase space than offered by the analytic expression.
5) Now the matrix-element configuration can be evolved further, to provide additional jets below the ME cuts used. In order to avoid doublecounting of emissions, any branchings that might occur above the ME cuts must be vetoed.

The MLM approach (Mangano et al 2007) is rather different. Here the steps instead are:
4') Allow a complete parton shower to develop from the selected parton configuration.
5') Cluster these partons back into a set of jets, e.g. using a cone-jet algorithm, with the same jet-separation criteria as used when the original parton configuration was picked.
6') Try to match each jet to its nearest original parton.
7') Accept the event only if the number of clustered jets agrees with the number of original partons, and if each original parton is sensibly matched to its jet. This would not be the case e.g. if one parton gave rise to two jets, or two partons to one jet, or an original b quark migrated outside of the clustered jet.
The point of the MLM approach is that the probability of *not* generating any additional fatal jet activity during the shower evolution is provided by the Sudakovs used in the step 4'.

The different approaches have been compared both on an experimental and on a theoretical levet, to understand the differences and possible shortcomings (Alwall et al 2008, Lavesson and Lönnblad 2008).

5.3 NLO matching

Matching to next-to-leading order in some respects is the most ambitious approach: it aims to get not only real but also virtual contributions correctly included, so that cross sections are accurate to NLO, and that NLO results are obtained for all observables when formally expanded in powers of α_s. Thus hard emissions should again be generated according to ME, while soft and collinear ones should fall within the PS regime. There are two main approaches on the market: MC@NLO and POWHEG.

For MC@NLO (Frixione and Webber 2002) the scheme works as follows in simplfied terms:
1) Calculate the NLO ME corrections to an n-body process, including $n+1$-body real corrections and n-body virtual ones.
2) Calculate analytically how a first branching in a shower starting from a n-body topology would populate $n+1$-body phase space, excluding the Sudakov factor.
3) Subtract the shower expression from the $n+1$ ME one to obtain the "true" $n+1$ events, and consider the rest as belonging to the n-body class. The PS and ME expressions agree in the soft and collinear limits, so the singularities there cancel, leaving finite cross sections both for the n- and $n+1$-body event classes.
4) Now add showers to both kinds of events.

Several processes have been considered, such as Z^0, $b\overline{b}$, $t\overline{t}$ and W^+W^- production. A technical problem is that, although ME and PS converge in the collinear region, it is not guaranteed that ME is everywhere above PS. This is solved by having a small fraction of events with negative weights.

The POWHEG approach (Nason 2004, Frixione et al 2007) is very closely related to the

merging approach presented earlier, but is more differential in phase space:

$$\mathrm{d}\sigma = \bar{B}(v)\mathrm{d}\Phi_v \left[\frac{R(v,r)}{B(v)} \exp\left(-\int_{p_\perp} \frac{R(v,r')}{B(v)} \mathrm{d}\Phi'_r \right) \mathrm{d}\Phi_r \right] \tag{6}$$

where

$$\bar{B}(v) = B(v) + V(v) + \int \mathrm{d}\Phi_r [R(v,r) - C(v,r)] \ ,$$

and
$v, \mathrm{d}\Phi_v$ are the Born-level n-body variables and differential phase space,
$r, \mathrm{d}\Phi_r$ are extra $n+1$-body variables and differential phase space,
$B(v)$ the Born-level cross section,
$V(v)$ the virtual corrections,
$R(v,r)$ the real-emission cross section, and
$C(v,r)$ the counterterms for collinear factorization of parton densities.
The basic idea is to pick the real emission with the largest transverse momentum according to complete ME's, including the Sudakov factor derived from an exponentiation of the real-emission expression, with the NLO normalization as a prefactor. (Note that the expression inside the square bracket of eq. (6) integrates to unity for any v if real emissions are allowed down to the soft/collinear singularities. With some effective cutoff a few events will not have any emissions at all above this cut.) Thereafter normal showers can be used to do the subsequent evolution downwards from the $p_\perp$ scale picked by the above equations.

The MC@NLO and POWHEG methods are formally equivalent to NLO, but not beyond, so differences are useful for exploring higher orders. The latter approach may be more appealing, since it eliminates the negative-weights problem and agrees with the concept of $p_\perp$ as the natural hardness scale, also e.g. for showers. However, as of now, MC@NLO has been worked out for more processes.

Note that the real-emission $n+1$-body part is only handled to LO accuracy, and higher-order jet topologies not at all. The NLO methods thus are useful for precision measurements of the total cross section of a process such as Z^0 or top production, but for studies of multiparton topologies the vetoed showers are more appropriate. Each tool to its task.

6 Multiparton interactions

The cross section for $2 \to 2$ QCD parton processes is dominated by t-channel gluon exchange, as we already mentioned, and thus diverges like $\mathrm{d}p_\perp^2/p_\perp^4$ for $p_\perp \to 0$. Introduce a lower cut $p_{\perp\mathrm{min}}$ and integrate the interaction cross section above this, properly convoluted with parton densities. At LHC energies this $\sigma_{\mathrm{int}}(p_{\perp\mathrm{min}})$ reaches around 100 mb for $p_{\perp\mathrm{min}} = 5$ GeV, and 1000 mb at around 2 GeV. Since each interaction gives two jets to lowest order, the jet cross section is twice as big. This should be compared with an expected *total* cross section of the order of 100 mb. In addition, at least a third of the total cross section is related to elastic scattering $pp \to pp$ and low-mass diffractive states $pp \to pX$ that could not contain jets.

So can it really make sense that $\sigma_{\mathrm{int}}(p_{\perp\mathrm{min}}) > \sigma_{\mathrm{tot}}$? Yes, it can! The point is that each incoming hadron is a bunch of partons. You can have several (more or less) independent parton–parton interactions when these two bunches pass through each other. And then an

event with n interactions above $p_{\perp\min}$ counts once for the total cross section but once *for each interaction* when the interaction rate is calculated. That is,

$$\sigma_{\rm tot} = \sum_{n=0}^{\infty} \sigma_n \qquad \text{while} \qquad \sigma_{\rm int} = \sum_{n=0}^{\infty} n\,\sigma_n \ ,$$

where σ_n is the cross section for events with n interactions. Thus $\sigma_{\rm int} > \sigma_{\rm tot}$ is equivalent to $\langle n \rangle > 1$, that each event on the average contains more than one interaction. Furthermore, if interactions do occur independently when two hadron pass by each other, then one would expect a Poissonian distribution, $\mathcal{P}_n = \langle n \rangle^n \exp(-\langle n \rangle)/n!$, so that several interactions could occur occasionally also when $\sigma_{\rm int}(p_{\perp\min}) < \sigma_{\rm tot}$, e.g. for a larger $p_{\perp\min}$ cut. Energy–momentum conservation ensures that interactions never are truly independent, and also other effects enter (see below), but the Poissonian ansatz is still a useful starting point.

Multiparton interactions (MPI) can only be half the solution, however. The divergence for $p_{\perp\min} \to 0$ would seem to imply an infinite average number of interactions. But what one should realize is that, in order to calculate the $\mathrm{d}\hat{\sigma}/\mathrm{d}\hat{t}$ matrix elements within standard perturbation theory, it has to be assumed that free quark and gluon states exist at negative and positive infinity. That is, the confinement of colour into hadrons of finite size has not been taken into account. So obviously perturbation theory has to have broken down by

$$p_{\perp\min} \simeq \frac{\hbar}{r_p} \approx \frac{0.2 \text{ GeV} \cdot \text{fm}}{0.7 \text{ fm}} \approx 0.3 \text{ GeV} \simeq \Lambda_{\rm QCD} \ .$$

The nature of the breakdown is also easy to understand: a small-$p_\perp$ gluon, to be exchanged between the two incoming hadrons, has a large transverse wavelength and thus almost the same phase across the extent of each hadron. The contributions from all the colour charges in a hadron thus add coherently, and that means that they add to zero since the hadron is a colour singlet.

What is then the typical scale of such colour screening effects, i.e. at what $p_\perp$ has the interaction rate dropped to ∼half of what it would have been if the quarks and gluons of a proton had all been free to interact fully independently? That ought to be related to the typical separation distance between a given colour and its opposite anticolour. When a proton contains many partons this characteristic screening distance can well be much smaller than the proton radius. Empirically we need to introduce a $p_{\perp\min}$ scale of the order of 2 GeV to describe Tevatron data, i.e. of the order of 0.1 fm separation. It is not meaningful to take this number too seriously without a detailed model of the space–time structure of a hadron, however.

The 2 GeV number is very indirect and does not really tell exactly how the dampening occurs. One can use a simple recipe, with a step-function cut at this scale, or a physically more reasonable dampening by a factor $p_\perp^4/(p_{\perp 0}^2 + p_\perp)^2$, plus a corresponding shift of the $\alpha_{\rm s}$ argument,

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_\perp^2} \propto \frac{\alpha_{\rm s}^2(p_\perp^2)}{p_\perp^4} \to \frac{\alpha_{\rm s}^2(p_{\perp 0}^2 + p_\perp^2)}{(p_{\perp 0}^2 + p_\perp^2)^2} \ , \tag{7}$$

with $p_{\perp 0}$ a dampening scale that also lands at around 2 GeV. This translates into a typical number of 2–3 interactions per event at the Tevatron and 4–5 at LHC. For events with jets or other hard processes the average number is likely to be higher.

6.1 Multiparton-interactions models

A good description of Tevatron data has been obtained with a simple model (Sjöstrand and van Zijl 1987) based on the following principles:
1) Only address MPI for inelastic nondiffractive events, with cross section $\sigma_{\rm nd}$, which form the bulk of what is triggered as minimum bias.
2) Dampen the perturbative jet cross section using the smooth turnoff of eq. (7), to cover the whole $p_\perp$ range down to $p_\perp = 0$.
3) Hadrons are extended, and therefore partons are distributed in (transverse) coordinates. To allow a flexible parametrization and yet have an easy-to-work-with expression, a double Gaussian $\rho_{\rm matter}(\mathbf{r}) = N_1 \exp\left(-r^2/r_1^2\right) + N_2 \exp\left(-r^2/r_2^2\right)$ is used, where N_2/N_1 and r_2/r_1 are tunable parameters.
4) The matter overlap during a collision, calculated by

$$\mathcal{O}(b) = \int \mathrm{d}^3\mathbf{x}\,\mathrm{d}t\; \rho^{\rm boosted}_{1,\rm matter}(\mathbf{x},t)\rho^{\rm boosted}_{2,\rm matter}(\mathbf{x},t) \;,$$

directly determines the average activity in events at different impact parameter b: $\langle n(b)\rangle \propto \mathcal{O}(b)$. That is, central collisions tend to have more activity, peripheral less.
5) An event has to contain at least one interaction to be an event at all. This provides a natural dampening of the cross section at large impact parameters. Normalizations have to be picked such that the b-integrated probability for having at least one interaction gives $\sigma_{\rm nd}$, while the b-integrated rate of all interactions gives the (dampened) jet cross section. Further, $p_{\perp 0}$ has to be selected sufficiently small that $\sigma_{\rm int} > \sigma_{\rm nd}$.
6) To first approximation the number of interactions at a given impact parameter obeys a Poissonian distribution, with the 0-interaction rate removed. Since central collisions have a larger mean and peripheral ones a smaller, the end result is a distribution broader than a Poissonian.
7) The interactions are generated in an ordered sequence of decreasing $p_\perp$ values: $p_{\perp 1} > p_{\perp 2} > p_{\perp 3} > \ldots$. This is possible with the standard Sudakov kind of trick:

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}p_{\perp i}} = \frac{1}{\sigma_{\rm nd}}\frac{\mathrm{d}\sigma}{\mathrm{d}p_\perp}\;\exp\left[-\int_{p_\perp}^{p_{\perp(i-1)}} \frac{1}{\sigma_{\rm nd}}\frac{\mathrm{d}\sigma}{\mathrm{d}p'_\perp}\mathrm{d}p'_\perp\right] \;,$$

with a starting $p_{\perp 0} = E_{\rm cm}/2$.
8) The ordering of emissions allows parton densities to be rescaled in x after each interaction, so that energy–momentum is not violated. Thereby the tail towards large multiplicities is reduced.
9) For technical reasons the model was simplified after the first interaction, so that there only gg or $q\overline{q}$ outgoing pairs were allowed, and no showers were added to these further $2 \to 2$ interactions.
Already this simple PYTHIA-based model is able to "explain" a large set of experimental data.

More recently a number of improvements have been included (Sjöstrand and Skands 2004 and 2005, Corke and Sjöstrand 2009).
1) The introduction of junction fragmentation, wherein the confinement field between the three quarks in a baryon is described as a Y-shaped topology, now allows the handling of topologies where several valence quarks are kicked out, thus allowing arbitrary flavours and

showering in all interactions in an event.
2) Parton densities are not only rescaled for energy–momentum conservation, but also to take into account the number of remaining valence quarks, or that sea quarks have to occur in $q\overline{q}$ pairs.
3) The introduction of $p_\perp$-ordered showers allows the selection of new ISR and FSR branchings and new interactions to be interleaved in one common sequence of falling $p_\perp$ values. Thereby the competition between especially ISR and MPI, which both remove energy from the incoming beams, is modelled more realistically.
4) Rescattering is optionally allowed, wherein one parton may undergo successive scatterings.

The traditional HERWIG soft underlying event (SUE) approach to this issue has its origin in the UA5 Monte Carlo. In it a number of clusters are distributed almost independently in rapidity and transverse momentum, but shifted so that energy–momentum is conserved, and the clusters then decay isotropically. The multiplicity distribution of clusters and their y and $p_\perp$ spectra are tuned to give the observed inclusive hadron spectra. No jets are produced in this approach.

The JIMMY program (Butterworth et al 1996) started as an add-on to HERWIG, but is nowadays an integrated part (Bähr et al 2008b). It replaces the SUE model with a MPI-based one more similar to the PYTHIA ones above, e.g. with an impact-parameter-based picture for the multiparton-interactions rate. Many technical differences exist, e.g. JIMMY interactions are not picked to be $p_\perp$-ordered and thus energy–momentum issues are handled differently.

The DPMJET/DTUJET/PHOJET family of programs (Aurenche et al 1994, Engel and Ranft 1996, Roesler et al 2000) come from the "historical" tradition of soft physics, wherein multiple $p_\perp \approx 0$ "pomeron" exchanges fill a role somewhat similar to the hard MPI above. Jet physics was originally not included, but later both hard and soft interactions have been allowed. One strong point is that this framework also allows diffractive events to be included as part of the same basic machinery.

6.2 Multiparton-interactions studies

How do we know that MPI exist? The key problem is that it is not possible to identify jets coming from $p_\perp \approx 2$ GeV partons. Therefore we either have to use indirect signals for the presence of interactions at this scale or we have to content ourselves with studying the small fraction of events where two interactions occur at visibly large $p_\perp$ values.

An example of the former is the total charged multiplicity distribution in high-energy $pp/p\overline{p}$ collision. This distribution is very broad, and is even getting broader with increasing energy, meaured in terms of the width over the average, $\sigma(n_{\rm ch})/\langle n_{\rm ch}\rangle$. By contrast, recall that for a Poissonian this quantity scales like $1/\sqrt{n_{\rm ch}}$ and thus is getting narrower. Simple models, with at most one interaction and with a fragmentation framework in agreement with LEP data, cannot explain this: they are way to narrow, and have the wrong energy behaviour. If MPI are included the additional variability in the number of interactions per event offers the missing piece. The variable impact parameter improves the description further.

Another related example is forward–backward correlations. Consider the charged multiplicity n_f and n_b in a forward and a backward rapidity bin, each of width one unit,

separated by a central rapidity gap of size Δy. It is not unnatural that n_f and n_b are somewhat correlated in two-jet events, and for small Δy one may also be sensitive to the tails of jets. But the correlation coefficient, although falling with Δy, still is appreciable even out to $\Delta y = 5$, and here again traditional one-interaction models come nowhere near. In a MPI scenario each interaction provides additional particle production over a large rapidity range, and this additional number-of-MPI variability leads to good agreement with data.

Direct evidence comes from the study of four-jet events. These can be caused by two separate interactions, but also by a single one where higher orders (call it ME or PS) has allowed two additional branchings in a basic two-jet topology. Fortunately the kinematics should be different. Assume the four jets are ordered in $p_\perp$, $p_{\perp 1} > p_{\perp 2} > p_{\perp 3} > p_{\perp 4}$. If coming from two separate interactions the jets should pair up into two separately balancing sets, $|\mathbf{p}_{\perp 1} + \mathbf{p}_{\perp 2}| \approx 0$ and $|\mathbf{p}_{\perp 3} + \mathbf{p}_{\perp 4}| \approx 0$. If an azimuthal angle φ is introduced between the two jet axes this also should be flat if the interactions are uncorrelated. By contast the higher-order graph offers no reason why the jets should occur in balanced pairs, and the φ distribution ought to be peaked at small values, corresponding to the familiar collinear singularity. The first to observe an MPI signal this way was the AFS collaboration at ISR (pp at 62 GeV), but with large uncertainties. A more convincing study was made by CDF, who obtained a clear signal in a sample with three jets plus a photon. In fact the deduced rate was almost a factor of three higher than naive expectations, but quite in agreement with the impact-parameter-dependent picture, wherein correlations of this kind are enhanced. Recently also D0 have shown a comparable signal with even higher statistics.

A topic that has been quite extensively studied in CDF is that of the jet pedestal (Field 1999–2009), i.e. the increased activity seen in events with a jet, even away from the jet itself, and away from the recoiling jet that should be there. Some effects come from the showering activity, i.e. the presence of additional softer jets, but much of it rather finds its explanation in MPI, as a kind of "trigger bias" effect, as follows.
(1) Central collisions tend to produce many interactions, peripheral ones few.
(2) If an event has n interactions there are n chances that one of them is hard.
Combine the two and one concludes that events with hard jets are biased towards central collisions and many additional interactions. The rise of the pedestal with triggger-jet energy saturates once $\sigma_{\rm int}(p_{\perp\rm min} = p_{\perp\rm jet}) \ll \sigma_{\rm nd}$, however, because by then events are already maximally biased towards small impact parameter. And this is indeed what is observed in the data: a rapid rise of the pedestal up to $p_{\perp\rm jet} \approx 10$ GeV, and then a slower increase that is mainly explained by showering contributions.

In more detailed studies of this kind of pedestal effects there are also some indications of a jet substructure in the pedestal, i.e. that indeed the pedestal is associated with the production of additional (soft) jet pairs.

In spite of many qualitative successes, and even some quantitative ones, one should not be lead to believe that all is understood. Possibly the most troublesome issue is how colours are hooked up between all the outgoing partons that come from several different interactions. A first, already difficult, question is how colours are correlated between all the partons that are taken out from an incoming hadron. These colours are then mixed up by the respective scattering, in principle (approximately) calculable. But, finally, all the outgoing partons will radiate further and overlap with each other on the way out, and how much that may mess up colours is an open question.

A sensitive quantity is $\langle p_\perp \rangle (n_{\rm ch})$, i.e. how the average transverse momentum of charged

particles varies as a function of their multiplicity. If interactions are uncorrelated in colour this curve tends to be flat: each further interaction adds about as much $p_\perp$ as $n_{\rm ch}$. If colours somehow would rearrange themselves, so that the confinement colour fields would not have to run criss-cross in the event, then the multiplicity would not rise as fast for each further interaction, and so a positive slope would result. The embarrassing part is that the CDF tunes tend to come up with values that are about 90% on the way to being maximally rearranged, which is way more than one would have guessed. Obviously further modelling and tests are necessary here.

Another issue is whether the $p_{\perp 0}$ regularization scale should be energy-dependent. In olden days there was no need for this, but it became necessary when HERA data showed that parton densities rise faster at small x values than had commonly been assumed. This means that the partons become more close-packed and that the colour screening increases faster with increasing collision energy. Therefore an energy-dependent $p_{\perp 0}$ is not unreasonable, but also cannot be predicted. If one assumes that $p_{\perp 0} \propto E_{\rm cm}^p$, with some power p, then the debate has centered on the range $p = 0.16 - 0.26$, with current best tunes (Skands 2009, Buckley et al 2009) leaning towards the higher end of this range. Recall that a larger p implies a larger $p_{\perp 0}$ at LHC energies, and thus a smaller multiplicity. Realistically we must still allow for some range of uncertainty, however.

7 Hadronization

The physics mechanisms discussed so far are mainly being played out on the partonic level, while experimentalists observe hadrons. In between exists the very important hadronization phase, where all the outgoing partons end up confined inside hadrons of a typical 1 GeV mass scale. This phase cannot (so far?) be described from first principles, but has to involve some modelling. The main approaches in use today are string fragmentation (Andersson et al 1983) and cluster fragmentation (Webber 1984).

Hadronization models start from some ideologically motivated principles, but then have to add "cookbook recipes" with free parameters to arrive at a complete picture of all the nitty gritty details. This should come as no surprise, given that there are hundereds of known hadron species to take into account, each with its mass, width, wavefunction, couplings, decay patterns and other properties that could influence the structure of the observable hadronic state, and with many of those properties being poorly or not at all known. In that sense, it is sometimes more surprising that models can work as well as they do than that they fail to describe everything.

While non-perturbative QCD is not solved, lattice QCD studies lend support to a linear confinement picture (in the absence of dynamical quarks), i.e. the energy stored in the colour dipole field between a charge and an anticharge increases linearly with the separation between the charges, if the short-distance Coulomb term is neglected. This is quite different from the behaviour in QED, and is related to the presence of a three-gluon vertex in QCD. The details are not yet well understood, however.

The assumption of linear confinement provides the starting point for the string model, most easily illustrated for the production of a back-to-back $q\overline{q}$ jet pair. As the partons move apart, the physical picture is that of a colour flux tube (or maybe colour vortex line) being stretched between the q and the $\overline{q}$. The transverse dimensions of the tube are of

typical hadronic sizes, roughly 1 fm. If the tube is assumed to be uniform along its length, this automatically leads to a confinement picture with a linearly rising potential. In order to obtain a Lorentz covariant and causal description of the energy flow due to this linear confinement, the most straightforward way is to use the dynamics of the massless relativistic string with no transverse degrees of freedom. The mathematical, one-dimensional string can be thought of as parametrizing the position of the axis of a cylindrically symmetric flux tube. From hadron spectroscopy, the string constant, i.e. the amount of energy per unit length, is deduced to be $\kappa \approx 1$ GeV/fm.

As the q and $\overline{q}$ move apart, the potential energy stored in the string increases, and the string may break by the production of a new $q'\overline{q}'$ pair, so that the system splits into two colour singlet systems $q\overline{q}'$ and $q'\overline{q}$. If the invariant mass of either of these string pieces is large enough, further breaks may occur. In the Lund string model, the string break-up process is assumed to proceed until only on-mass-shell hadrons remain, each hadron corresponding to a small piece of string.

In order to generate the quark–antiquark pairs $q'\overline{q}'$, which lead to string break-ups, the Lund model invokes the idea of quantum mechanical tunnelling. This leads to a flavour-independent Gaussian spectrum for the transverse momentum of $q'\overline{q}'$ pairs. Tunnelling also implies a suppression of heavy quark production, $u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$. Charm and heavier quarks hence are not expected to be produced in the soft fragmentation.

A tunnelling mechanism can also be used to explain the production of baryons. This is still a poorly understood area. In the simplest possible approach, a diquark in a colour antitriplet state is just treated like an ordinary antiquark, such that a string can break either by quark–antiquark or antidiquark–diquark pair production. A more complex scenario is the 'popcorn' one, where diquarks as such do not exist, but rather quark–antiquark pairs are produced one after the other.

In general, the different string breaks are causally disconnected. This means that it is possible to describe the breaks in any convenient order, e.g. from the quark end inwards. Results, at least not too close to the string endpoints, should be the same if the process is described from the q end or from the $\overline{q}$ one. This 'left–right' symmetry constrains the allowed shape of fragmentation functions $f(z)$, where z is the fraction of $E + p_{\mathrm{L}}$ that the next particle will take out of whatever remains. Here p_{L} is the longitudinal momentum along the direction of the respective endpoint, opposite for the q and the $\overline{q}$. Two free parameters remain, which have to be determined from data.

If several partons are moving apart from a common origin, the details of the string drawing become more complicated. For a $q\overline{q}g$ event, a string is stretched from the q end via the g to the $\overline{q}$ end, i.e. the gluon is a kink on the string, carrying energy and momentum. As a consequence, the gluon has two string pieces attached, and the ratio of gluon/quark string forces is 2, a number that can be compared with the ratio of colour charge Casimir operators, $N_C/C_F = 2/(1 - 1/N_C^2) = 9/4$. In this, as in other respects, the string model can be viewed as a variant of QCD, where the number of colours N_C is not 3 but infinite. Fragmentation along this kinked string proceeds along the same lines, as sketched for a single straight string piece. Therefore no new fragmentation parameters have to be introduced.

The concept of cluster fragmentation offers the great promise of a simple, local and universal description of hadronization. At the end of the shower evolution all gluons are split into $q\overline{q}$ pairs, and $q\overline{q}'$ colour singlets can be formed from them by keeping track of the colour flow in the event. Typically the q and $\overline{q}'$ of such a singlet were formed in adjacent

shower branches, and therefore tend to have a rather small mass. These so-called clusters are assumed to be the basic units from which the hadrons are produced. A cluster is ideally only characterized by its total mass and total flavour content, i.e. unlike a string it does not possess an internal structure. If the shower-evolution cutoff is chosen such that most clusters have a mass of a few GeV, the cluster mass spectrum may be thought of as a superposition of fairly broad (i.e. short-lived) resonances. Phase-space aspects may then be expected to dominate the decay properties. This applies both for the selection of decay channels, and for the kinematics of the decay. Thus a decay is assumed to be isotropic in the rest frame of the cluster. This gives a compact description with few parameters. This approach has been successful in explaining the particle composition in terms of very few parameters. The momentum distributions and correlations are a bit more tricky, and in practice some string ideas are needed, e.g. to break large-mass clusters into smaller ones along a string direction.

One general conclusion is that neither of the two models is well constrained from first principles. In the string model many parameters are needed for the flavour composition setup, while the energy–momentum (and space–time) picture is very economical. In the cluster model it is the other way around.

8 Summary and outlook

In these lectures we have followed the flow of generators roughly "inwards out", i.e. from short-distance processes to long-distance ones. At the core lies the hard process, described by matrix elements. It is surrounded by initial- and final-state showers, that should be properly matched to the hard process. Multiple parton–parton interactions can occur, and the colour flow is tied up with the structure of beam remnants. At longer timescales the partons turn into hadrons, many of which are unstable and decay further. This basic pattern is likely to remain in the future, but many aspects will change.

One such aspect, that stands a bit apart, is that of languages. The traditional event generators, like Pythia and Herwig, have been developed in Fortran — up until the end of the LEP era this was the main language in high-energy physics. But now the experimental community has switched to C++ for heavy compting. The older generators are still being used, hidden under C++ wrappers, but this can only be a temporary solution, for several reasons. One is that younger experimentalists often need to look into the code of generators and tailor some parts to specific needs of theirs, and if then the code is in an unknown language this will not work. Another is that theory students who apply for non-academic positions are much better off if their resumés say "object-oriented programming guru" rather than "Fortran fan".

A conversion program thus has begun on many fronts. Sherpa, as the youngest of the general-purpose generators, was conceived from the onset as a C++ package and thus is some steps ahead of the other programs in this respect. Herwig++ (Bähr et al 2008a) is a complete reimplementation of Herwig, as is Pythia 8 (Sjöstrand et al 2008) of Pythia 6. Both conversions have taken longer than originally hoped, but by now the new programs are fully operational and starting to be used by the LHC collaborations. ThePEG (Lönnblad 2006) is a generic toolkit for event generators, used in particular by Herwig++.

The authors of these event generators have joined in MCnet, which currently is funded by

the European Union as a Marie Curie Research Training Network. In addition a few other projects are pursued, notably the RIVET package that implements various experimental analyses from the literature, and the PROFESSOR framework (Buckley et al 2009) that takes RIVET input as the starting point for semiautomatic tuning of event generators. MCnet arranges summer schools each year (MCnet 2007, CTEQ–MCnet 2008, MCnet 2009), alone or in collaboration with the CTEQ school. There are also funds to allow graduate students in theory and experiment to come and work with a generator author on a specific project for a few months. If you are interested, have a look at `http://www.montecarlonet.org/`

To summarize these lectures, there are many aspects where we have seen progress in recent years and can hope for more:

- Faster, better and more user-friendly general-purpose matrix-element generators with an improved sampling of phase space.
- New ready-made libraries of physics processes, in particular with full NLO corrections included.
- More precise parton showers.
- Better matching between matrix elements and parton showers.
- Improved models for minimum-bias physics and underlying events.
- Some upgrades of hadronization models and decay descriptions.

In general one would say that generators are getting better all the time, but at the same time the experimental demands are also getting higher, so it is a tight race. However, given that typical hadronic final states at LHC will contain hundreds of particles and quite complex patterns buried in that, it is difficult to see that there are any alternatives to the Monte Carlo generator approach.

Acknowledegments

The author is grateful to the organizers of the School for offering such a pleasant environment. This work was supported in part by the EU Marie Curie Research Training Network "MCnet", under contract number MRTN-CT-2006-035606.

References

Allanach, B et al (2009), *Computer Physics Comm.* **180**, 8.

Alwall, J et al (2007), *Computer Physics Comm.* **176**, 300.

Alwall, J et al (2008), *Eur. Phys. J.* **C53**, 473.

Amsler, C et al (Particle Data Group) (2008), *Phys. Lett.* **B667** (2008) 1.

Andersson, B et al (1983), *Phys. Rep.* **97**, 31.

Aurenche, P et al (1994), *Computer Physics Commun.* **83**, 107.

Bähr, M et al (2008a), *Eur. Phys. J.* **C58**, 639.

Bähr, M et al (2008b), *JHEP* **0807**, 076.

Boos, E et al (2001), in *Proceedings of the Workshop on Physics at TeV Colliders, Les Houches, France, 21 May - 1 Jun 2001*, [hep-ph/0109068].

Bourilkov, D et al (2006), hep-ph/0605240;
`http://projects.hepforge.org/lhapdf/`.

Buckley, A et al (2009), arXiv:0907.2973 [hep-ph].

Butterworth, J M et al (1996), *Z. Phys.* **C72**, 637.

Catani, S et al (2001), *JHEP* **0111**, 063.

Corcella, G et al (2001), *JHEP* **0101**, 010.

Corke, R and Sjöstrand, T (2009), arXiv:0911.1909 [hep-ph].

CTEQ-MCnet (2008), Summer School, Debrecen:
`http://conference.ippp.dur.ac.uk/conferenceOtherViews.py?view=ippp&confId=156`.

Dobbs, M and Hansen, J B (2001) *Computer Physics Comm.* **134**, 41.

Engel, R and Ranft, J (1996), *Phys. Rev.* **D54**, 4244.

Field, R D (1999–2009) presentations partly on behalf of the CDF Collaboration:
`http://www.phys.ufl.edu/~rfield/cdf/rdf_talks.html`

Frixione, S and Webber, B R (2002), *JHEP* **0206**, 029.

Frixione, S et al (2007), *JHEP* **0711**, 070.

Gieseke, S et al (2003), *JHEP* **0312**, 045.

Gleisberg, T et al (2009), *JHEP* **0902**, 007.

Gustafson, G and Pettersson, U (1988) *Nucl. Phys.* **B306**, 746.

Lavesson, N and Lönnblad, L (2008), *JHEP* **0804**, 085.

Lönnblad, L (1992), *Computer Physics Commun.* **71**, 15.

Lönnblad, L (2002), *JHEP* **0205**, 046.

Lönnblad, L (2008), *Nucl. Instrum. Meth.* **A559**, 246.

Mangano, M L et al (2007), *JHEP* **0701**, 013.

MCnet (2007), Summer School, Durham:
`http://conference.ippp.dur.ac.uk/conferenceTimeTable.py?confId=3&detailLevel=contribution&viewMode=parallel`.

MCnet (2009), Summer School, Lund:
`http://conference.ippp.dur.ac.uk/conferenceOtherViews.py?view=ippp&confId=264#2009-07-01`.

Nason, P (2004), *JHEP* **0411**, (2004) 040.

Norrbin, E and Sjöstrand, T (2001), *Nucl. Phys.* **B603**, 297.

Roesler, S et al (2000), hep-ph/0012252.

Schumann, S and Krauss, F (2008), *JHEP* **0803**, 038.

Sjöstrand, T and van Zijl, M (1987), *Phys. Rev.* **D36**, 2019.

Sjöstrand, T and Skands, P (2004), **JHEP 0403**, 053.

Sjöstrand, T and Skands, P (2005), *Eur. Phys. J.* **C39**, 129.

Sjöstrand, T et al (2006), *JHEP* **0605**, 026.

Sjöstrand, T et al (2008), *Computer Physics Comm.* **178**, 852.

Sjöstrand, T (2009), `http://www.thep.lu.se/~torbjorn/` (click on "Talks").

Skands, P et al (2004), *JHEP* **0407**, 036.

Skands, P (2009), arXiv:0905.3418 [hep-ph].

Webber, B R (1984), *Nucl. Phys.* **B238** 492.

Winter, J and Krauss, F (2008), *JHEP* **0807**, 040.