Introduction to Monte Carlo Event Generation and Parton Showering

CTEQ School 2018, Mayagüez, Puerto Rico

Silvan Kuttimalai

25th June 2018

In these lecture notes, the basic principles of Monte Carlo event generation and parton showering are explained. The lectures are intended as a primer for a hands-on tutorial, in which the students will have the opportunity to code up their own parton shower Monte Carlo for e^+e^- annihilation entirely from scratch. An effort is made to cover the basics in great detail, rather than to give a comprehensive overview of state-of-the-art methods. For more information on modern full fledged general purpose event generators, I recommend references [1–3] and the references therein.

1 The Basic Problem

Monte Carlo event generation is a technique for deriving theory predictions for observables at collider experiments. Generically, such observables can be described in a somewhat idealized way by a function O of the final state particles detected in a collision. If we denote the final state by f and the corresponding phase space (the momenta of the particles in f) by Φ^f, we can write

$$O(f;\Phi^f)\,. \tag{1}$$

The exact outcome, i.e. the final f state with its momenta $\Phi^f = \{p_1,\ldots,p_{n_f}\}$, cannot be predicted for any one specific particle collision, even if the exact initial conditions are known. As a consequence of the principles of quantum mechanics, only the probability for some final state to be produced during a collision can be predicted. This probability is proportional to the differential cross section

$$\frac{\mathrm{d}\,\sigma(f;\Phi^f)}{\mathrm{d}\,\Phi^f}\,. \tag{2}$$

Provided a theory prediction for the differential cross section, one can make meaningful predictions for the expectation value of some observable

$$\langle O \rangle = \mathcal{L} \times \sum_f \int \frac{\mathrm{d}\,\sigma(f;\Phi^f)}{\mathrm{d}\,\Phi^f} O(f;\Phi^f)\,\mathrm{d}\,\Phi^f\,. \tag{3}$$

Focusing on one specific final state we can drop the corresponding index f for clarity. The total contribution from several final states can be trivially obtained by adding them all up. The task of calculating the expectation value of the observable then boils down to evaluating integrals of the form

$$\langle O \rangle = \mathcal{L} \times \int \frac{\mathrm{d}\,\sigma(\Phi)}{\mathrm{d}\,\Phi} O(\Phi)\,\mathrm{d}\,\Phi\,. \tag{4}$$

For a process in which two particles a and b with momenta p_a and p_b annihilate into n particles with momenta p_i with $i \in \{1,\ldots,n\}$, we have, for example

$$\frac{\mathrm{d}\,\sigma}{\mathrm{d}\,\Phi}\,\mathrm{d}\,\Phi = \frac{1}{F}|\mathcal{M}(\Phi)|^2\,\mathrm{d}\,\Phi \tag{5}$$

$$\mathrm{d}\,\Phi = (2\pi)^4\delta^4(p_a + p_b - \sum_{i=1}^{n} p_i)\prod_{i=1}^{n}\frac{\mathrm{d}^3\,\vec{p}_i}{(2\pi)^3 2E_i} \tag{6}$$

$$F = 4\sqrt{(p_a p_b)^2 - (m_a m_b)^2} \tag{7}$$

Both the observable and the scattering matrix element $\mathcal{M}$ can be arbitrarily complex functions, thus rendering the analytical evaluation of the integral extremely challenging, if at all possible. Numerical methods therefore generally have to be employed. The Monte Carlo method is the most flexible and widely used method in particle physics.

2 Monte Carlo Integration

As stated above, calculating the expectation value of a particle collider observable typically involves complex high dimensional phase space integrals. Let the quantity of interest generically be an integral I of some function f in a multi-dimensional variable x

$$I = \int_G f(x)\,\mathrm{d}^n\,x\,, \tag{8}$$

where G is the integration region. In the Monte Carlo approach, the integral I is approximated by drawing a large number N of random phase space points $x_i \in G$ with $i \in \{1,\ldots,N\}$

$$I_N = \frac{V}{N}\sum_{i=1}^{N} f(x_i)\,, \tag{9}$$

where V is the volume of the integration region G. The law of large numbers guarantees that

$$\lim_{N\to\infty} I_N = I\,. \tag{10}$$

Practically, one can of course only generate a finite number of phase space points and therefore only calculate I_N for some finite N as a prediction for I. The square root of the statistical variance of I_N can then be used as an estimator for the statistical uncertainty Δ of this prediction. It is fairly straightforward to show that this quantity is given by

$$\Delta(I_N) = \sqrt{\mathrm{Var}(I_N)} = \frac{V}{\sqrt{N}}\sqrt{\mathrm{Var}(f)}\,. \tag{11}$$

The variance of f is defined in terms of it's mean $\bar{f} = I/V$ as

$$\mathrm{Var}(f) = \frac{1}{V}\int_G (f(x) - \bar{f})^2\,\mathrm{d}^n\,x\,. \tag{12}$$

This demonstrates that the scaling behavior of the uncertainty is essentially $1/\sqrt{N}$ with no dependence on the number of phase space dimensions. This is one of the advantages compared to other numerical methods, which often exhibit a problematic scaling with the number of phase space dimensions.

3 Monte Carlo Event Generation

Monte Carlo event generation is simply the application of the Monte Carlo method to integrals of the form (4):

$$\langle O\rangle = \lim_{N\to\infty} \mathcal{L}\times\frac{V}{N}\sum_{i=1}^{N}\left.\frac{\mathrm{d}\,\sigma(\Phi)}{\mathrm{d}\,\Phi}\right|_{\Phi=\Phi_i} O(\Phi_i) \tag{13}$$

$$= \lim_{N\to\infty} \mathcal{L}\times\frac{V}{N}\sum_{i=1}^{N} w_i\,O(\Phi_i)\,, \tag{14}$$

where V is the volume of the integration region. The only ingredients needed in order to implement (14) are routines that evaluate the differential cross section $\mathrm{d}\,\sigma/\,\mathrm{d}\,\Phi$ and the observable $O(\Phi)$ for a given phase space point and a routine that randomly generates phase space points Φ_i. A fully generic but simple algorithm for generating valid phase space points (with all particles on shell and with four-momentum conservation respected) is described in [4]. Once a set of phase space points has been generated, it can be stored along with the corresponding values of the differential cross section weights w_i. These phase space points can be interpreted as simulated *Monte Carlo events*. They can then be used to calculate *any* observable O without redoing the most complex part of the calculation, i.e. the generation of phase

space points and the evaluation of the differential cross section. Herein lies the flexibility of the method and one of the most important advantages compared to an analytic observable-specific calculation.

3.1 A simple example

In the tutorials, we will be looking at $e+e- \to q\bar{q}$, where q is any massless quark. In this simple example, one can show that the phase space can be simplified as

$$\int \mathrm{d}\,\Phi = \frac{\mathrm{d}^3\,\vec{p}_q}{(2\pi)^3 2E_q}\frac{\mathrm{d}^3\,\vec{p}_{\bar{q}}}{(2\pi)^3 2E_{\bar{q}}} = \int_0^{2\pi}\mathrm{d}\,\phi\int_{-1}^{+1}\frac{\mathrm{d}[\cos(\theta)]}{16\pi}\,. \tag{15}$$

Combined with the matrix element $\mathcal{M}$, which can be looked up in [5], this gives

$$\int \frac{\mathrm{d}\,\sigma}{\mathrm{d}\,\Phi}\,\mathrm{d}\,\Phi = \int_0^{2\pi}\mathrm{d}\,\phi\int_{-1}^{+1}\mathrm{d}[\cos(\theta)]\frac{|\mathcal{M}(\cos\theta)|^2}{64\pi^2 E_{\mathrm{cm}}^2} \tag{16}$$

In the tutorial toy Monte Carlo, we therefore have a simple two-dimensional integration to solve. We achieve this by drawing ϕ and $\cos\theta$ from a uniform distributions in the intervals $[0, 2\pi]$ and $[-1, 1]$, respectively.

4 Variance Reduction Techniques

Expression (11) for the statistical uncertainty on the Monte Carlo integration result shows that the method will perform poorly if the integrand has a large variance, i.e. if the integrand fluctuates a lot. It is often computationally not feasible to solve this issue by simply increasing the number N of integrand evaluations. Note that in order to improve the uncertainty by a factor of 10, one would have to increase N by a factor of 100.

The convergence can, however, be significantly improved by variance reduction techniques. The methods employed in Monte Carlo codes typically rely on a biased generation of phase space points x. This can be intuitively understood. Suppose we know that the integrand $f(x)$ is almost zero everywhere except for small region in phase space where $f(x)$ is large and potentially fluctuates wildly. Then we can improve the convergence by drawing x in such a way that most points end up in this region. This bias then has the be accounted for, otherwise the integration result will of course change. To see how this work, we re-write (8) as an expectation value with respect to the constant probability distribution $p(x) = 1/V$

$$I = V \times \int f(x)\frac{1}{V}\,\mathrm{d}\,x \tag{17}$$

$$= V \times \int f(x)p(x)\,\mathrm{d}\,x \tag{18}$$

$$= V \times E[f, p]\,, \tag{19}$$

where $E[f,p]$ denotes the expectation value of f with respect to the probability distribution function p. Now it is straightforward to see that

$$\begin{aligned} V \times E[f,p] &= V \times E[fg/g,p] \\ &= E[f/g,g] \end{aligned} \tag{20}$$

This means that we can evaluate the function f/g instead of f if we distribute the phase space points according to g. If the variance of f/g is smaller than the variance of f, this in turn leads to a better convergence due to (11). We can re-phase (20) also in terms of an integration variable transformation ϕ.

$$\int f(x)\,\mathrm{d}\,x = \int f(\phi(t)) \times \phi'(t)\,\mathrm{d}\,t\,. \tag{21}$$

Now we choose $\phi(t) = G^{-1}(t)$, where G^{-1} is the inverse of the primitive integral of g and find

$$\int f(x)\,\mathrm{d}\,x = \int f(G^{-1}(t)) \times [G^{-1}]'(t)\,\mathrm{d}\,t \tag{22}$$

$$= \int \frac{f(G^{-1}(t))}{G'(G^{-1}(t))}\,\mathrm{d}\,t \tag{23}$$

$$= \int \frac{f(G^{-1}(t))}{g(G^{-1}(t))}\,\mathrm{d}\,t\,. \tag{24}$$

As shown in secion 5, this corresponds to drawing t from the probability distribution g and multiplying with the factor $1/g$ as described above.

The method outlined above of course relies on the function g being constructed in such a way as to minimize the variance of f/g. One can show that, ideally, $g = |f|/I$. This is of course not possible without knowing I. In practice, one therefore constructs an approximate g, either based on some knowledge about the integrand, or algorithmically by probing the integrand. The latter approach is taken in the VEGAS method [6]. Both approches can be combined into very powerful adaptive multichannel algorithms [7].

5 Biased Random Number Generation

Essential ingredients for any Monte Carlo event generator are sequences of random numbers. In practice, one typically relies on pseudorandom numbers generated by some algorithm. We will not discuss such algorithms here but simply assume we have an algorithm at hand that is capable of generating a uniformly distributed random variable $r \in [0,1]$. Given such an algorithm, there are several methods to generate numbers for a variable $\tau \in [\tau_0, \tau_1]$ that is distributed according to some arbitrary probability density function $g(\tau)$.

First some notation: let r be a random variable, then we denote by p_r it's probability distribution function and by P_r its cumulative distribution

function:

$$\text{Prob}(r_0 < r < r_1) = \int_{r_0}^{r_1} p_r(r')\,\mathrm{d}\,r' = P_r(r_1) - P_r(r_0)\,. \tag{25}$$

For a uniformly distributed variable r in $[0;1]$, we have for example

$$p_r(r') = 1 \tag{26}$$

$$P_r(r_1) = r_1\,. \tag{27}$$

We will now discuss the *inverse transform sampling* method for generating values for a random variable that is distributed according to some arbitrary probability density function g. This method will be heavily used in the tutorials. Suppose we know g's primitive integral G as well as it's inverse G^{-1} and suppose that the random variable r is uniformly distributed. Then, as we will show, the random variable

$$\tau = G^{-1}(r) \tag{28}$$

is distributed according to g, as desired. Here it is of course crucial that the primitive G is chosen such that $G(\tau_0) = 0$ and $G(\tau_1) = 1$. One can show that (28) indeed yields a variable that is distributed according to g by calculating the cumulative distribution function of τ:

$$P_\tau(\tau_1) = \text{Prob}(\tau < \tau_1) \tag{29}$$

$$= \text{Prob}(G^{-1}(r) < \tau_1) \tag{30}$$

$$= \text{Prob}(r < G(\tau_1)) \tag{31}$$

$$= P_r(G(\tau_1)) = G(\tau_1) \tag{32}$$

$$\Rightarrow p_\tau(\tau') = g(\tau') \tag{33}$$

The technique outlined above can only be applied if one can construct and invert the function G, which is not always the case. Alternatively, the hit-and-miss technique can be employed. In this method, τ is generated uniformly in $[\tau_0, \tau_1]$. A value of τ_i genrated in this way is then accepted with probability $g(\tau_i)/g_{\max}$, where $g_{\max}$ is the maximum value of g in the interval $[\tau_0, \tau_1]$. This method is computationally expensive in cases where the function evaluations of g are expensive. Every generated point involves a function call to g, regardless of whether it will be eventually accepted or rejected. Using this method in Monte Carlo event generation, one can however perform an unweighting of events in a straightforward way by subjecting the acceptance/rejection procedure to all events according to their weight.

6 Parton Showers

6.1 Parton Fragmentation

With the methods described above we are in principle ready to construct a Monte Carlo event generator if we have a way to compute differential scat-

tering cross section. Methods to do that for reactions involving elementary particles such as electrons, positrons, quarks, and gluons are well established. In reactions involving QCD partons, the particles eventually hitting the detector are composite hadrons, however. QCD partons do not exist as free particles, instead they form hadrons in a non-perturbative *fragmentation* process. One can introduce fragmentation functions

$$D_i^h(\xi)\,. \tag{34}$$

similarly to parton distribution functions (PDFs) in order to describe this process. Such functions would have to be extracted from experiment. They can be interpreted as the probability for some quark i to fragment into hadron h carrying a fraction ξ of its energy. We will consider only one hadron in what follows and therefore drop index h for clarity.

In terms of the fragmentation functions and scattering processes involving only bare partons, we can for example write down an expression for the differential energy distribution in hadrons of type h. Imagine measuring the energy E_h inside such hadrons or, equivalently, the scaled energy

$$x = \frac{2E_h}{\sqrt{s}} \in [0,1]\,. \tag{35}$$

Now consider the corresponding differential distribution

$$F(x) = \frac{\mathrm{d}\,\sigma^{e^+e^-\to\ h+X}}{\mathrm{d}\,x} = \sum_i \int_0^1 \mathrm{d}\,z \int_0^1 \mathrm{d}\,\xi\ C_i(z,\mu_f^2) D_i(\xi,\mu_f^2)\delta(z\xi - x) \tag{36}$$

$$= \sum_i \int_x^1 \frac{\mathrm{d}\,z}{z} C_i(z,\mu_f^2) D_i(x/z,\mu_f^2)\,, \tag{37}$$

where C_i denotes the partonic inclusive production cross section for $e^+e^- \to i+X$. This quantity can be computed in perturbation theory since it involves only elementary partons. At leading order and we have simply

$$C_i(z,\mu_f^2) = c_i^0 \delta(1-z) \tag{38}$$

$$F(x) = \sum_i c_i^0 D_i(x)\,. \tag{39}$$

where the $c_i^0 = 0$ for $i = g$ and c_q^0 is the leading order cross section $\sigma_{\mathrm{LO}}(e^+e^- \to q\bar{q})$.

6.2 Scale Dependence of Fragmentation Functions

Now consider NLO corrections to C_q where q is a massless quark. One can show that the real-emission corrections from gluon emissions off the quarks

$e^+e^- \to q\bar{q}g$ factorize in the limit where the gluon is collinear to either of them. In the region where the gluon is collinear to the quark, we have

$$\mathrm{d}\,\sigma(e^+e^- \to\ q\bar{q}g) \approx \mathrm{d}\,\sigma(e^+e^- \to\ q\bar{q}) \times \frac{\alpha_s}{2\pi}\hat{P}_{qq}(z)\frac{\mathrm{d}\,t}{t}\,\mathrm{d}\,z \tag{40}$$

where t is the virtuality (invariant mass) of the intermediate quark and z is the energy fraction transferred from the intermediate quark to the final state quark. The splitting kernel $\hat{P}_{qq}$ is given by

$$\hat{P}_{qq}(z) = C_f \frac{1+z^2}{1-z}\,. \tag{41}$$

The splitting kernel is singular in the region $z = 1$, where the gluon becomes soft. This singularity is canceled by virtual corrections, which can be absorbed into a *regularized* splitting kernel P_{qq}:

$$P_{qq}(z) = C_f \left[\frac{1+z^2}{1-z}\right]_+ \tag{42}$$

$$= C_f \left[\frac{1+z^2}{[1-z]_+} + \frac{3}{2}\delta(1-z)\right] \tag{43}$$

The plus distribution is typically defined in terms of its properties as an integration kernel:

$$\int_0^1 [g(z)]_+ f(z)\,\mathrm{d}\,z = \int_0^1 [g(z)f(z) - g(z)f(1)]\,\mathrm{d}\,z\,. \tag{44}$$

A more intuitive understanding can be obtained from a limiting procedure:

$$\int_0^1 [g(z)]_+ f(z)\,\mathrm{d}\,z = \lim_{\varepsilon\to 0}\int_0^1 g_\varepsilon(z) f(z)\,\mathrm{d}\,z\,, \tag{45}$$

where

$$g_\varepsilon(z) = \Theta(1-\varepsilon-z)g(z) - \frac{\Theta\left(z-(1-\varepsilon)\right)}{\varepsilon}\int_0^{1-\varepsilon} g(z')\,\mathrm{d}\,z'\,. \tag{46}$$

Applied to a splitting kernel with a soft singularity at $z = 1$, the first term can be interpreted as the real-emission contribution resolved above energy fractions ε of the gluon. The second term corresponds to virtual loop corrections and unresolved contributions below ε.

In terms of the regularized kernel (42) involving the plus prescription, we have at NLO

$$C_q(z) = c_q^0\left[\delta(1-z) + \int_0^{Q^2}\frac{\mathrm{d}\,t}{t}\frac{\alpha_s}{2\pi}P_{qq}(z)\right]\,, \tag{47}$$

where $Q^2 = s$ is the center-of-mass energy of the collider. The integral over t is logarithmically divergent as $t \to 0$. For now we will regularize this divergence with a lower bound, imposing $t > \kappa^2$.

$$C_q(z) = c_q^0 \left[\delta(1-z) + \frac{\alpha_s}{2\pi} \log\left(\frac{Q^2}{\kappa^2}\right) P_{qq}(z) \right] \tag{48}$$

The divergence at $\kappa^2 = 0$ is not subject to cancellations with virtual corrections. It comes from collinear splittings partons involving small momentum transfers and should thus be considered part of the fragmentation process described by the fragmentation function rather than the partonic cross section. The divergent part up to some arbitrary factorization scale μ_f is therefore absorbed into a renormalized fragmentation function as follows

$$D_q(x, \mu_f^2) = \int_x^1 \frac{\mathrm{d}\, z}{z} \left[\delta(1-z) + \frac{\alpha_s}{2\pi} \log\left(\frac{\mu_f^2}{\kappa^2}\right) P_{qq}(z) \right] D_q(x/z) \tag{49}$$

$$= D_q(x) + \frac{\alpha_s}{2\pi} \log\left(\frac{\mu_f^2}{\kappa^2}\right) \int_x^1 \frac{\mathrm{d}\, z}{z} P_{qq}(z) D_q(x/z)\,. \tag{50}$$

In terms of this renormalized fragmentation function we now have up to terms of order α_s^2

$$F(x) = \int_x^1 \frac{\mathrm{d}\, z}{z} c_q^0 \left[\delta(1-z) + \frac{\alpha_s}{2\pi} \log\left(\frac{Q^2}{\mu_f^2}\right) P_{qq}(z) \right] D_q(x/z, \mu_f^2)\,, \tag{51}$$

which is independent of the regulator κ. This regulator can now be safely removed. The singularity is of course still present in the relation between the renormalized and the unrenormalized fragmentation function, but the former is not observable. The "bare" fragmentation functions without scale dependence disappear entirely from our expressions if written in terms of the renormalized scale dependent fragmentation functions. In that sense, the bare functions are only intermediate objects.

A consequence of this is that we cannot compute the renormalized fragmentation functions themselves using their definition in (50). Only their dependence on the factorization scale can be extracted from (50). It reads

$$t \frac{\partial}{\partial t} D_q(\xi, t) = \frac{\alpha_s}{2\pi} \int_\xi^1 \frac{\mathrm{d}\, z}{z} P_{qq}(z) D_q(\xi/z, t) \tag{52}$$

With $D_q(\xi, t)$ measured at some value of t, the evolution equation (62) can be used to obtain $D_q(\xi, t)$ at any other value of t. This corresponds to changing the amount of radiation that is absorbed into the fragmentation function.

6.3 The Sudakov Form Factor

Given a measurement of $D_q(\xi, t)$ at some value t, equation (62) allows us to perturbatively predict the value of D_q at some other value t'. As will be shown, it is useful for practical purposes to re-write this equation in terms of a quantity called the Sudakov form factor

$$\Delta_q(t, t_0) = \exp\left[-\int_{t_0}^{t} \frac{\mathrm{d}\,\bar{t}}{\bar{t}} \int_0^{1-\varepsilon} \frac{\alpha_s}{2\pi} \hat{P}_{qq}(\bar{z})\,\mathrm{d}\,\bar{z}\right], \tag{53}$$

where ε is an infrared regulator that precisely corresponds to the ε in the plus-prescription as defined in (46), as will be shown later.

We note that the Sudakov form factor $\Delta_q(t, t_0)$ as a function of the second argument can be interpreted as a probability, since $\Delta(t, t_0) \in [0, 1]$ for $t_0 \in [t, 0]$. We will for now not interpret this probability but merely note that the corresponding probability *density* $p(t')$ can be found by differentiation of the form factor:

$$p(t') = \Delta_q(t, t') \frac{1}{t'} \int_0^{1-\varepsilon} \frac{\alpha_s}{2\pi} P_{qq}(\bar{z})\,\mathrm{d}\,\bar{z}\,. \tag{54}$$

6.4 Parton Shower Evolution

Solving the evolution equation (62) numerically is not entirely straightforward because the splitting kernel P_{qq} is irregular at $z = 1$ due to the divergence which is subtracted at exactly $z = 1$ through the plus prescription that effectively adds contributions from virtual corrections. Writing out the plus prescription explicitly, we have

$$\begin{aligned}
t\frac{\partial}{\partial t} D_q(\xi, t) &= \frac{\alpha_s}{2\pi} \int_\xi^1 \frac{\mathrm{d}\,z}{z} P_{qq}(z) D_q(\xi/z, t) && (55)\\
&= \frac{\alpha_s}{2\pi} \int_0^1 \mathrm{d}\,z \hat{P}_{qq}(z) \left[\frac{1}{z} D_q(\xi/z, t) - D_q(\xi, t)\right] && (56)\\
&= \frac{\alpha_s}{2\pi} \int_x^1 \frac{\mathrm{d}\,z}{z} \hat{P}_{qq}(z) D_q(\xi/z, t) - \frac{\alpha_s}{2\pi} \int_0^1 \mathrm{d}\,z \hat{P}_{qq}(z) D_q(\xi, t) && (57)\\
&\approx \frac{\alpha_s}{2\pi} \int_\xi^{1-\varepsilon} \frac{\mathrm{d}\,z}{z} \hat{P}_{qq}(z) D_q(\xi/z, t) - \frac{\alpha_s}{2\pi} \int_0^{1-\varepsilon} \mathrm{d}\,z \hat{P}_{qq}(z) D_q(\xi, t)\,, && (58)
\end{aligned}$$

where we set $D_q(\xi, t) = 0$ for $\xi > 1$ in order to be able to extend the integration region down to $z = 0$. In the last step we have used the definition of the plus distribution given in (46).

We can now re-write equation (58) in terms of the Sudakov form factor

introduced in (53) as follows

$$t\frac{\partial}{\partial t}D_q(\xi,t)+\frac{\alpha_s}{2\pi}\int_0^{1-\varepsilon}\mathrm{d}\,z\hat{P}_{qq}(z)D_q(\xi,t)=\frac{\alpha_s}{2\pi}\int_\xi^{1-\varepsilon}\frac{\mathrm{d}\,z}{z}\hat{P}_{qq}(z)D_q(\xi/z,t) \tag{59}$$

$$t\frac{\partial}{\partial t}\frac{D_q(\xi,t)}{\Delta_q(t,t_0)}=\frac{\alpha_s}{2\pi}\int_\xi^{1-\varepsilon}\frac{\mathrm{d}\,z}{z}\hat{P}_{qq}(z)D_q(\xi/z,t) \tag{60}$$

This differential equation can now be integrated, giving

$$D_q(\xi,t)=\Delta_q(t,t_0)D_q(\xi,t_0)+\int_{t_0}^{t}\frac{\mathrm{d}\,t'}{t'}\Delta_q(t,t')\int_\xi^{1-\varepsilon}\frac{\mathrm{d}\,z}{z}\frac{\alpha_s}{2\pi}\hat{P}_{qq}(z)D_q(\xi/z,t') \tag{61}$$

Equation (61) now explicitly gives $D_q(x,t)$ in terms of an integral involving $D_q(x,t')$ at all $t'<t$. More importantly, it can be interpreted probabilistically as we will show.

We imagine a quark being produced with initial virtuality t of the order of the scale of the hard process (Q^2 in e^+e^- collisions). Then we imagine the quark being able to emit a gluon and thereby split. This gives rise to a virtuality t' of the intermediate quark which would be limited to values below the hard initial scale t. We imagine the value of the virtuality t' in the splitting being distributed according to (54). The probability for the splitting *not* to occur in the interval $[t,t_0]$ i.e. the probability of no splitting above the scale t_0 is then given by the Sudakov form factor (53).

We can then interpret the first term in (61) as contributions to the fragmentation from quarks that were produced at t and then do not produce any emission above t_0 (hence the factor $\Delta_q(t,t_0)$). These quarks deposit their energy into the hadron according to $D_q(x,t_0)$, i.e. the fragmentation function that can be thought of as absorbing all emissions below t_0.

The second term in (61) then adds contribution from quarks which undergo a splitting at some intermediate scale t', distributed according to (54). Correspondingly, the fragmentation function is evaluated at t'. The first argument of the fragmentation function is ξ/z because the energy transferred in the fragmentation of the final quark must be larger by a factor of $1/z$ due to the energy lost in the splitting. Equation (61) then recursively applies again to $D_q(\xi/z,t')$.

Parton shower branching algorithms explicitly simulate this probabilistic process and thus implement (61). Starting at the hard scale of the process, final state partons undergo a branching with the value of virtuality in the next splitting distributed according to (54). The lower bound t_0 on the virtuality is where the parton shower algorithm stops.

In the discussion above we have only considered one type of splitting in which $q\to qg$. The generalization to the case in which all splittings are

considered reads

$$t\frac{\partial}{\partial t}D_i(\xi,t) = \sum_j \frac{\alpha_s}{2\pi}\int_\xi^1 \frac{\mathrm{d}\,z}{z} P_{ji}(z) D_j(\xi/z,t) \tag{62}$$

where the full set of regularized splitting functions is given by

$$P_{qq}(z) = C_F\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \tag{63}$$

$$P_{gq}(z) = C_F\left[\frac{1+(1-z)^2}{z}\right] \tag{64}$$

$$P_{gg}(z) = C_A\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right]\times 2 + \delta(1-z)\frac{11C_A - 4N_f T_R}{6} \tag{65}$$

$$P_{qg}(z) = T_R\left[z^2 + (1-z)^2\right] \tag{66}$$

For parton shower evolusion we use the generalization of (61), which reads

$$D_i(x,t) = \Delta_i(t,t_0)D_i(x,t_0) + \sum_j \int_{t_0}^t \frac{\mathrm{d}\,\bar{t}}{\bar{t}}\Delta_i(t,\bar{t})\int_0^{1-\varepsilon}\frac{\mathrm{d}\,z}{z}\frac{\alpha_s}{2\pi}\hat{P}_{ji}(z)D_j(x/z,\bar{t})\,, \tag{67}$$

with the Sudakov form factor

$$\Delta_i(t,t_0) = \exp\left[-\sum_j\int_{t_0}^t\frac{\mathrm{d}\,\bar{t}}{\bar{t}}\int_0^{1-\varepsilon}\mathrm{d}\,\bar{z}\frac{\alpha_s}{2\pi}\hat{P}_{ji}(\bar{z})\right] \tag{68}$$

and the unregularized splitting kernels

$$\begin{aligned}
P_{qq}(z) &= C_F\left[\frac{1+z^2}{1-z}\right]\\
P_{gq}(z) &= C_F\left[\frac{1+(1-z)^2}{z}\right]\\
P_{gg}(z) &= C_A\left[\frac{z}{(1-z)} + \frac{1-z}{z} + z(1-z)\right]\\
P_{qg}(z) &= T_R\left[z^2+(1-z)^2\right]
\end{aligned} \tag{69}$$

6.5 Soft evolution

A parton shower algorithm as described above does not only express fragmentation functions evaluated at some large scale in terms of fragmentation functions at smaller scales. In addition, it explicitly adds the collinear emissions that are absorbed into the functions at high scales back to the physical picture through the branching process. This process correctly captures the

physics of collinear emissions, which are enhanced due to the $1/t$ singularities of in the corresponding cross section in (40).

In addition to the singularities in the collinear region, there are singularities whenever a gluon becomes soft, i.e. when it's energy becomes small. These singularities are manifest as divergences in the unregularized splitting function, for example in P_{qq} at $z = 1$. If a gluon becomes soft without being collinear to any other parton, then (40) is not valid anymore, however. This is because gluon emissions off *all* hard parton in a process contribute with a soft divergence. The interference terms between emissions off the quark and anti-quark therefore cannot be neglected anymore. The correct soft factorization formula for the cross section reads

$$\mathrm{d}\,\sigma[e^+e^- \to\ q\bar{q}g] \approx \mathrm{d}\,\sigma[e^+e^- \to\ q\bar{q}] \times \frac{\alpha_s}{2\pi}\frac{p_q p_{\bar{q}}}{(p_q k)(p_{\bar{q}} k)} E\,\mathrm{d}\,E\frac{\mathrm{d}\,\Omega}{2\pi}\,, \tag{70}$$

where k denotes the gluon momentum and E it's energy component. The expression above implies three-particle correlations between the quark, the anti-quark, and the emitted gluon which turn out to be important in the description of data. One can still capture these effects in terms of a branching algorithm. In dipole-like showers, not one particle splits into two, but two particles i, j split into three. The two emitting particles i, j can be interpreted as an emitting "color dipole" [8, 9].

In the tutorials, you will have the chance to build a dipole-like shower. It is based on dipole factorization of Catani and Seymour [10]. In their method, the *eikonal factor* in equation (70) is rearranged as follows

$$\frac{p_q p_{\bar{q}}}{(p_q k)(p_{\bar{q}} k)} = \frac{1}{p_q k}\frac{p_q p_{\bar{q}}}{(p_q + p_{\bar{q}})q} + \frac{1}{p_{\bar{q}} k}\frac{p_q p_{\bar{q}}}{(p_q + p_{\bar{q}})q}\,. \tag{71}$$

The factors $\frac{1}{p_q k}$ and $\frac{1}{p_{\bar{q}} k}$ correspond to the factor $\frac{1}{t}$ in (40) and hence to the intermediate virtuality when the gluon splits off the quark or the anti-quark, respectively. The remaining factor is singular only in the region where the gluon becomes soft and can be interpreted as a splitting function. The two terms on the right hand side of (71) therefore have the same structure as (40). One can now construct splitting kernels that reduce to the DLGAP kernels (69) in the collinear limit, but capture correctly the soft correlations implied by (70) away from the collinear limit. The only conceptual difference to the plain collinear evolution algorithm described in section 6.4 is that the splitting functions depend on an additional "spectator" parton, i.e. $p_{\bar{q}}$ in the first term on the right hand side of (71) and p_q in the second term on the right hand side of (71). The explicit form of the splitting functions used in the tutorials is given in [3].

References

[1] A. Buckley et al. "General-Purpose Event Generators for LHC Physics". In: *Physics Reports* 504 (2011), pp. 145–233. arXiv: 1101.2599 [hep-ph].

[2] Torbjorn Sjostrand, Stephen Mrenna and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. arXiv: hep-ph/0603175 [hep-ph].

[3] Stefan Höche. "Introduction to parton-shower event generators". In: *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*. 2015, pp. 235–295. arXiv: 1411.4085 [hep-ph].

[4] R. Kleiss, W. James Stirling and S. D. Ellis. "A New Monte Carlo Treatment of Multiparticle Phase Space at High-energies". In: *Comput. Phys. Commun.* 40 (1986), p. 359.

[5] R. K. Ellis, B.R. Webber and W. J. Stirling. *QCD and Collider Physics*. Cambridge University Press, 1996.

[6] G.P. Lepage. "A New Algorithm for Adaptive Multidimensional Integration". In: *Journal of Computational Physics* 27 (1978), pp. 192–203.

[7] Ronald Kleiss and Roberto Pittau. "Weight optimization in multichannel Monte Carlo". In: *Comput. Phys. Commun.* 83 (1994), pp. 141–146. arXiv: hep-ph/9405257 [hep-ph].

[8] Gosta Gustafson and Ulf Pettersson. "Dipole Formulation of QCD Cascades". In: *Nucl. Phys.* B306 (1988), pp. 746–758.

[9] G. Gustafson. "Dual Description of a Confined Color Field". In: *Phys. Lett.* B175 (1986). [,193(1986)], p. 453.

[10] S. Catani and M.H. Seymour. "A General Algorithm for Calculating Jet Cross-Sections in NLO QCD". In: *Nuclear Physics B* 485 (1997), pp. 291–419. arXiv: hep-ph/9605323 [hep-ph].