Parton showers with quantum interference

Zoltán Nagy

Theory Division, CERN,
CH-1211 Geneva 23, Switzerland
E-mail: Zoltan.Nagy@cern.ch

Davison E. Soper

Institute of Theoretical Science, University of Oregon,
Eugene, OR 97403-5203, U.S.A.
E-mail: soper@uoregon.edu

ABSTRACT: We specify recursive equations that could be used to generate a lowest order parton shower for hard scattering in hadron-hadron collisions. The formalism is based on the factorization soft and collinear interactions from relatively harder interactions in QCD amplitudes. It incorporates quantum interference between different amplitudes in those cases in which the interference diagrams have leading soft or collinear singularities. It incorporates the color and spin information carried by partons emerging from a hard interaction. One motivation for this work is to have a method that can naturally cooperate with next-to-leading order calculations.

KEYWORDS: Hadronic Colliders, QCD, Jets.

Contents

1. Introduction

Parton shower Monte Carlo event generators, such as HERWIG [1] and PYTHIA [2], have proven to be enormously useful since the development of the main ideas in the 1980s [3–5]. These computer programs perform calculations of cross sections according to an approximation to the standard model or some of its possible extensions. Because of the great success of these programs, it is worthwhile to investigate possible improvements. In this paper, we propose a theoretical structure for event generators that generalizes the structure of current programs and allows the elimination of certain approximations used currently.

Parton showers are mostly reflections of QCD interactions. In order to present a reasonably complete discussion of the QCD issues in a parton shower while keeping the length of this paper within reasonable bounds, we limit the presentation to QCD and omit any discussion of how electroweak and beyond-the-standard-model interactions are to be added to the QCD interactions to make a useful event generator.

What is a parton shower Monte Carlo event generator? Let us consider hadron-hadron collisions, which is the case relevant for the Tevatron and the Large Hadron Collider. An experiment will produce a large number of events f, where one can characterize an event as a list of the momenta and flavors of the final state particles produced. The experiment can measure a cross section $\sigma[F]$ corresponding to an observable[1] that assigns to each event f a number $F(f)$. The relation of the cross section and the function F is

$$\sigma[F] \approx \frac{1}{\mathcal{L}} \sum_{n=1}^{N} F(f_n) \ , \tag{1.1}$$

where $\mathcal{L}$ is the integrated luminosity for an experimental run and the f_n are the observed events. For example, the cross section to produce a Higgs boson and two jets having certain characteristics is specified by setting $F(f) = 1$ if f contains a Higgs boson and two jets having these characteristics and $F(f) = 0$ otherwise.[2]

---

[1]In order to be subject to reliable calculation in QCD perturbation theory, the function F should have the property known as infrared safety. However, a parton shower event generator is also useful for observables that are not infrared safe.

[2]The case in which $F(f)$ takes values 0 or 1 is the most common, but other possibilities are allowed. For instance, the energy-energy correlation function in electron-positron annihilation is of the more general variety.

A parton shower Monte Carlo event generator calculates this cross section by producing a large number N of simulated events f_n, each with an accompanying weight w_n. The calculated cross section is then

$$\sigma[F] \approx \frac{1}{N} \sum_{n=1}^{N} w_n \, F(f_n) \ . \qquad (1.2)$$

Most typically, the weights are all equal, so that $1/w_n$ is the simulated luminosity per point $\mathcal{L}/N$. Our definition of the category of parton shower Monte Carlo event generator includes the possibility that the weights are complex numbers produced for each event. It is always possible to throw away the imaginary parts of the w_n since we know in advance that the imaginary part of the sum in eq. (1.2) vanishes, so having complex weights is equivalent to having real weights that can be positive or negative. This situation occurs in typical event generators [6, 7] that are based on next-to-leading order perturbation theory.[3]

In a typical parton shower event generator, the physics is modeled as a process in classical statistical mechanics. Some number of partons are produced in a hard interaction. Then each parton has a chance to split into two partons, with the probability to split determined from an approximation to the theory. Parton splitting continues in this probabilistic style until a complete parton shower has developed.

The parton splitting probability is biggest when the two daughter partons are almost massless with nearly collinear momenta or when one of their momenta is soft (near $p = 0$), or both. There is a simple underlying approximation used: the amplitude for producing $m + 1$ partons when two of the momenta p_i and p_j are nearly collinear or one is soft factors into a splitting function times the matrix element for producing m partons.

The underlying approximation is the factorization of amplitudes in the soft or collinear limits. However, further approximations are usually added:

1. The interference between a diagram in which a soft gluon is emitted from one hard parton and a diagram in which the same soft gluon is emitted from another hard parton is treated in an approximate way, with the "angular ordering" approximation.

2. Color is treated in an approximate way, valid when $1/N_c^2 \to 0$ where $N_c = 3$ is the number of colors.

3. Parton spin is treated in an approximate way. According to the full quantum amplitudes, when a parton splits, the angular distribution of the daughter partons depends on the mother parton spin and even on the interference between different mother-parton spin states. This dependence is typically ignored.

With the use of these further approximations, one can get to a formalism in which the shower develops according to classical statistical mechanics with a certain evolution operator.

Our purpose in this paper is to investigate whether one can have a formulation of parton showers based on the factorization of amplitudes in the soft or collinear limits in

---

[3]The recent paper [8] provides an exception to this rule.

which one does not make the additional approximations enumerated above. For this, one would have to use quantum statistical mechanics instead of classical statistical mechanics.

It might seem that doing the problem in quantum mechanics is hopelessly complicated. However, within the soft/collinear factorization approximation, the problem is fairly simple because it is almost classical. In fact, if partons did not have color or spin, the problem would be classical (as we discuss in section 2). Thus what we need is a fully quantum treatment of color and spin. We arrange for this by making use of the quantum density operator in color ⊗ spin space.

In the subsequent sections, we define evolution equations for the quantum density matrix within the soft/collinear factorization approximation. The matrix evolves in "shower time" from harder splittings to softer splittings. The iterative solution of these equations gives $\sigma[F]$ in the form of a sum of integrals. To give some idea of the structure, we omit any mention of hadronization and write the result in a notation that is quite abbreviated compared to the notation in the body of the paper,

$$\sigma[F] = \int dP_0 \; f_0 \sum_{N=1}^{\infty} \prod_{j=1}^{N} \left(\int d\zeta_j \; f_j \right) \; F \; . \tag{1.3}$$

There is, first of all, an integration (including sums, for discrete variables) over momenta, flavors, spins, and colors for initial partons that emerge from the hard matrix element and its complex conjugate. Here we call all of these variables collectively P_0, the initial partonic variables. There is a function f_0 that depends on P_0 and represents the hard matrix element at the start of the shower times its complex conjugate. Then there is a sum over how many splittings, N, there are.[4] Next there is an integration over splitting variables ζ_j for the jth splitting. The splitting variables include the label telling which parton split and momentum variables, for which a dimensionless virtuality y, a momentum fraction z, and an azimuthal angle ϕ might be used. There are also discrete flavor, color, and spin variables. At each splitting, there is a set of starting partonic variables, P_{j-1} and a set of new partonic variables P_j that are determined by P_{j-1} and the splitting parameters ζ_j. For each splitting, there is a function f_j that depends on P_{j-1} and ζ_j. We have integrations over the splitting parameters for splittings 1 through N. At the end, there is the measurement function F that depends on the partonic variables P_N reached after all of the splittings.

The structure of this representation is similar to that in conventional parton showers, with the functions f_j made from splitting functions and Sudakov exponentials that express the probability for not splitting. There are, however, some important structural differences that result from not making the approximations 1, 2, and 3 above. Chief among them is the use of the spin and color variables.

What we develop in this paper is an evolution equation that results in a representation of $\sigma[F]$ as integrals of known functions. Of course, one will want to turn the integrals into

---

[4]We have formally iterated the evolution equation an infinite number of times, allowing any number of splittings. However, we imagine that there is a cutoff on splitting hardness, so that very large values of N are seldom encountered. Some of our splittings are $1 \to 1$ self interactions rather than $1 \to 2$ splittings.

numbers. How to do that is a question of numerical integration that we leave for future work. However, it may be useful to sketch how a numerical evaluation might work.

To evaluate $\sigma[F]$ numerically, one has to construct the functions involved as described in the body of this paper and then perform all of the integrations (and sums). In particular, Monte Carlo integration can be used for many of the integrations. In a numerical method that is very, very simple, one can choose random points P_0 first, according to a density ρ_0. Then one would choose the first splitting variables, ζ_1, according to a density ρ_1 that is determined by the variables P_0. This would determine new partonic variables P_1. Continuing, we choose ζ_j according to a density ρ_j that is determined by variables P_{j-1} and we use ζ_j together with P_{j-1} to determine P_j. At the end, we use P_N as input to the measurement function F. This process constitutes a Markov chain that produces "events" with a final state P_N. The probability density for getting a final state P_N with a given shower history is ρ_0 times the product of the ρ_j for $j \geq 1$. We multiply $F(P_n)$ by a weight equal to f_0/ρ_0 times a product of the f_j/ρ_j.

It would be a design goal to choose the ρ_j to be roughly proportional to the absolute values of f_j. This kind of importance sampling would produce weights that do not vary over a wide range.

We may note that in a conventional parton shower, the f_j are everywhere positive and, for $j \geq 1$, integrate to 1. Thus one can choose the ρ_0 to be proportional to f_0 and $\rho_j = f_j$ for $j \geq 1$. Then the weight function is a constant. In our case, the factors in f_j are not everywhere positive, so we expect to need weights, which could have either sign.

What we have described above would generate a very conceptually simple numerical solution to the evolution equation. We expect that one could do much better, particularly by performing the spin sums not by numerical Monte Carlo summation but by exact summation. For this, one could adapt the method proposed by Collins [9] and elaborated by Knowles [10] and by Richardson [11].

We leave issues of the numerical evaluation of the integrals, beyond this simple discussion, for future work.

Some features of the formalism presented here can best be understood by asking what would happen if we kept only the leading $1/N_c^2 \to 0$ limit and averaged over spins everywhere, thus making approximations 2 and 3 above. We would then have a shower based on gluon emission from color dipoles. With such a picture, the imposition of a cut to enforce angular ordering (approximation 1) is not needed: interference between gluon emissions from both halves of a color dipole is already included. The resulting evolution equation would then be similar to what is implemented as $k_\perp$-showers in PYTHIA, as described in ref. [12]. One could also modify the formalism presented here change from k^2 as the evolution variable to the version of $k_\perp^2$ used in PYTHIA. Then there would be two main features that differed between the present formalism and PYTHIA. One difference is in the choice of splitting functions. The PYTHIA choice is the Altarelli-Parisi splitting functions, defined with a certain definition of the momentum fraction z. Our splitting functions are made from the Feynman diagrams for one off-shell parton producing two on-shell partons with only a minimal manipulation to separate this part of the diagram from the hard scattering to which it attaches. The definitions match in limit of collinear splittings, but differ away

from this limit. The other difference is in the momentum mapping that connects the momentum space for two initial state partons and $m+1$ final state partons to that with m final state partons. We have investigated [13] the possibility of using the Catani-Seymour [14] mapping and splitting functions, which are commonly used for next-to-leading order calculations. However, we have here adopted a mapping that avoids the use of designated "spectator" partons that share some of their momenta.[5] PYTHIA uses the Catani-Seymour momentum mapping for final state splittings and something more complicated for initial state splittings. We note here the recent paper [16], which explores other possibilities for both momentum mapping and splitting functions. We also note that S. Schumann and F. Krauss and, separately, M. Dinsdale, M. Temick, and S. Weinzierl have very recently implemented the Catani-Seymour dipole subtraction functions and momentum mappings as the basis for a parton shower. The first results look promising [17].

One of our goals has been to have a formulation that can coexist easily with matching the probabilities generated by showering to known exact tree level matrix elements, as in refs. [18] and with using next-to-leading order hard matrix elements as in refs. [6, 7, 19]. However, we leave for future work the analysis of how one can match the showers to the exact tree level matrix elements or to next-to-leading order hard matrix elements.

One may wonder whether removing the approximations 1, 2, and 3 listed above is numerically important. We do not have a definitive answer. What we would like to do is to set up a formalism that does not make approximations beyond the basic soft/collinear factorization approximation, then (in future work) implement this formalism as a working algorithm and computer code. One could then make the further approximations separately or all together and see what difference they make.

We close this introduction with some comments on whether a parton shower Monte Carlo event generator ought to allow weights for generated events, and in particular negative weights.

We first note that in a real experiment the relation between the measurement function F and the measured cross section is a little more complicated than we indicated in eq. (1.1). Instead, we have

$$\sigma[F] \approx \frac{1}{\mathcal{L}} \sum_{n=1}^{N} \frac{1}{a_n} F(f_n) \ , \tag{1.4}$$

where $\mathcal{L}$ is the integrated luminosity for an experimental run in which N total events n are collected and a_n is the acceptance for the event resulting from the way the detector is triggered. For example, if a fraction 10^{-3} of a certain kind of event is recorded, then for those events a_n is 10^{-3}. Thus the weight factors w_n/N in eq. (1.2) are analogous to $1/(a_n\mathcal{L})$ in the analysis of real data.

We can also examine the effect of weights on the statistical error in eq. (1.2). The

---

[5]One motivation for avoiding a special role for designated "spectator" partons is that if one wants to go to a next-to-leading order splitting kernel or to subtractions for a NNLO perturbative calculation, problems can arise from a third parton becoming collinear with the designated spectator parton [15].

expected error $\mathcal{E}$ is given by

$$\mathcal{E}^2 \sim \frac{1}{N} \left(\frac{1}{N} \sum_{n=1}^{N} \left[w_n\, F(f_n) - \langle wF \rangle \right]^2 \right) \ . \tag{1.5}$$

Here $\langle wF \rangle$ indicates an average. If it is very expensive to use a large N, for instance because calculating $F(f_n)$ requires a full detector simulation, then one would like to make $\mathcal{E}^2$ as small as possible for a fixed N. That suggests not that w_n should be constant, but that $w_n\, F(f_n)$ should be approximately constant for the observables F of most interest.[6] It is never possible to make $w_n\, F(f_n)/\langle wF \rangle = 1$ for all events, but one does not want to have lots of events for which this ratio is much smaller than 1 nor any events for which the ratio is much bigger than 1. Having events for which $w_n\, F(f_n)/\langle wF \rangle \sim -1$ is not a good thing, but, since $|-1|$ is not much larger than 1, it is not really damaging from the point of view of avoiding large statistical errors.

Evidently having weights that are real numbers of either sign, or complex numbers, does not make it impossible to apply eq. (1.2). It does, however, make the analysis a little more complicated. However, we believe that that the added complication does not present a serious problem.

Preview. Since this is a rather lengthy paper, some preview of what is in it may be helpful. Section 2 contains an introduction to the notation we use. This notation is, we think, useful for thinking about a variety of formulations of the parton shower idea. We present it in the context of a simple scalar field theory that is free from a lot of the complications of quantum chromodynamics (QCD). We then turn to QCD, with its complications. We present in section 3 the structure that we propose for a parton shower that contains quantum interference. In order to present this structure in just a few pages, we leave for later sections most of the detailed definitions. The first of these, the momentum and flavor mapping, is covered in section 4. Then section 5 covers spin. This provides enough background to present the splitting functions for the quantum amplitudes in section 6. The description of color, which is rather more complicated than that of spin, is presented in section 7. We are then able to specify the shower evolution operator in sections 8 and 9. We follow this with discussions of two interesting issues, the evolution of color in section 10 and soft gluon coherence in section 11. We analyze the structure of the functions that appear in the Sudakov exponent in section 12. Eventually shower evolution stops and a hadronization model is inserted. We discuss this in section 13. We present some concluding remarks in section 14. There are two appendices that deal with certain technical issues.

2. A notation for parton showers

Starting in the next section, we present a formulation for parton showers in QCD hard scattering events in hadron-hadron collisions, taking into account the complexities introduced by spin and color correlations and by soft, wide angle gluon emissions in addition

---

[6]Thus, if our primary interest were in the high P_T tail of a jet P_T distribution, we would not want to use most of the available computer time to generate low P_T events. Rather, we would want to generate few low P_T events, giving each of them a high weight to compensate.

to collinear splittings. In order to do this, we use a mathematical language that helps to organize the algorithm. Alas, the complexities of the real physical situation make the needed construction a bit subtle. Therefore, in this section we first introduce some of the needed language in a simpler situation. The notation introduced here will be used again for QCD in the subsequent sections.

Consider the process $e^+e^- \to hadrons$ in a world in which hadrons consist of just one kind of massless scalar particle, which has no color. The e^+e^- annihilation produces a pair of virtual scalar particles through an interaction that we need not specify. In the evolution of the hadronic state, we can still have collinear singularities similar to those found in QCD if the theory consists of ϕ^3 theory in six dimensional space-time. The cross section to measure an observable F can be written as

$$\sigma[F] = \sum_m \frac{1}{m!} \int \left[d\{p\}_m\right] |M(\{p\}_m)|^2 F(\{p\}_m) \ . \tag{2.1}$$

Here there is a sum over the number m of produced particles, $\{p\}_m = \{p_1, \dots, p_m\}$, and

$$\int \left[d\{p\}_m\right] \equiv \prod_{i=1}^m \left\{ \int \frac{d^6 p_i}{(2\pi)^6} \, 2\pi \delta_+(p_i^2) \right\} (2\pi)^6 \delta\left(P_0 - \sum_{i=1}^m p_i\right) \ , \tag{2.2}$$

with $P_0 = (\sqrt{s}, \vec{0})$. The function $M(\{p\}_m)$ gives the matrix element to produce m particles with momenta $\{p\}_m$, while $F(\{p\}_m)$ describes the measurement to be done.

We want to describe this using an algorithm that approximates $|M(\{p\}_m)|^2$ based on $|M(\{p\}_m)|^2$ for $m = 2$ and the subsequent generation of the rest of the particles based on a narrow angle approximation for one particle to split into two. This is to be done using a Monte Carlo simulation in which the system evolves from 2 particles to many particles as a simulation time t progresses from $t = 0$ to a large value, at which the simulation is terminated. There are various possibilities for the physical meaning of the time t. We will take it that a splitting $l \to i + j$ occurs at time $t = \log(Q_0^2/(2p_i \cdot p_j))$, where Q_0^2 is the hardness scale of the hard interaction with which we start.

At each stage of this simulation, let the cross section to have m particles with momenta $\{p\}_m$ be $\rho(\{p\}_m, t)$. Summing over the number of particles and integrating over momenta gives the total cross section,

$$\sigma_{\mathrm{T}} = \sum_m \frac{1}{m!} \int \left[d\{p\}_m\right] \rho(\{p\}_m, t) \ . \tag{2.3}$$

At the final time, t_{f}, the value of the measurement function is

$$\sigma[F] = \sum_m \frac{1}{m!} \int \left[d\{p\}_m\right] \rho(\{p\}_m, t_{\mathrm{f}}) \, F(\{p\}_m) \ . \tag{2.4}$$

The possible functions ρ (at a given time t) form a vector space, so that ρ at time t can be considered to be a vector $|\rho(t))$. We use rounded brackets here. The notation $|\psi\rangle$ is reserved for a quantum state, while $|\rho)$ denotes a state in the sense of statistical

mechanics. We therefore call it a statistical state. The inner product is[7]

$$(A|B) = \sum_m \frac{1}{m!} \int [d\{p\}_m] \; A(\{p\}_m) \, B(\{p\}_m) \; . \tag{2.5}$$

We can define basis vectors $\big|\{p\}_m\big)$ in this space so that

$$\rho(\{p\}_m, t) = \big(\{p\}_m\big|\rho(t)\big) \; . \tag{2.6}$$

With these definitions, there is a completeness relation

$$1 = \sum_m \frac{1}{m!} \int [d\{p\}_m] \; \big|\{p\}_m\big)\big(\{p\}_m\big| \; . \tag{2.7}$$

The measurement function F can also be considered to be a vector, $\big(F\big|$. Thus

$$\sigma[F] = \big(F\big|\rho(t_{\mathrm{f}})\big) \; . \tag{2.8}$$

There is a special vector $\big(1\big|$ with

$$\big(1\big|\{p\}_m\big) = 1 \; . \tag{2.9}$$

This vector represents the totally inclusive measurement function corresponding to the total cross section,

$$\sigma_{\mathrm{T}} = \big(1\big|\rho(t)\big) \; . \tag{2.10}$$

Now we are ready to discuss the evolution of the statistical state. We take the evolution to be given by a linear operator $\mathcal{U}(t, t')$, with

$$\big|\rho(t)\big) = \mathcal{U}(t, t')\big|\rho(t')\big) \; . \tag{2.11}$$

Here $\mathcal{U}(t, t) = 1$. These operators have the group composition property

$$\mathcal{U}(t_3, t_2)\mathcal{U}(t_2, t_1) = \mathcal{U}(t_3, t_1) \; . \tag{2.12}$$

The class of evolution operators that we will use is defined by two operators. The first is an infinitesimal generator of evolution or hamiltonian, $\mathcal{H}_{\mathrm{I}}(t)$. We can specify $\mathcal{H}_{\mathrm{I}}(t)$ by giving its action on an arbitrary state $\big|\rho\big)$,

$$\mathcal{H}_{\mathrm{I}}(t)\big|\rho\big) \; . \tag{2.13}$$

In a lowest order shower, the operator $\mathcal{H}_{\mathrm{I}}(t)$ describes parton splitting, changing a state with m particles to one with $m+1$ particles. One of the particles in $\{p\}_m$, say particle l, is removed and replaced by two, with momenta $\hat{p}_l$ and $\hat{p}_{m+1}$. The quantum amplitude after the splitting is approximately

$$M(\{\hat{p}\}_{m+1}) \approx M(\{p\}_m) \times \frac{g}{2\hat{p}_l \cdot \hat{p}_{m+1}} \; . \tag{2.14}$$

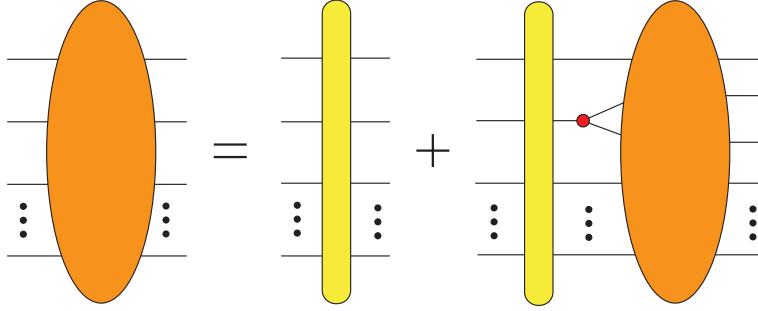---

[7]Note that there is no * here.

Figure 1: Illustration of eq. (2.18). The wide ovals represent the evolution operator $\mathcal{U}$ while the narrow ovals represent the no change operator $\mathcal{N}$, which provides the Sudakov exponentials, and the small circle is a parton splitting. Shower time t runs from left to right.

Here, to precisely define the right-hand side, one needs to redefine the momenta so that the mother parton has a momentum $p_l \approx \hat{p}_l + \hat{p}_{m+1}$ that is nevertheless on-shell, $p_l^2 = 0$. Thus the momenta $\{p\}_m$ are functions of the momenta $\{\hat{p}\}_{m+1}$. We omit a discussion here of the various ways to define this momentum mapping. More important for now is the idea that the quantum matrix element factorizes in the form (2.14) when $2\hat{p}_l \cdot \hat{p}_{m+1}$ is much smaller than any of the dot products among the momenta in $\{p\}_m$. This factorization is at the heart of the reason why parton shower Monte Carlo programs give useful approximations. For the statistical splitting function in eq. (2.13), we need the square of the quantum amplitude. Thus we want

$$\left(\{\hat{p}\}_{m+1}\big|\mathcal{H}_I(t)\big|\rho\right) = \sum_l \delta\left(t - \log\left(\frac{Q_0^2}{2\hat{p}_l \cdot \hat{p}_{m+1}}\right)\right)\left[\frac{g}{2\hat{p}_l \cdot \hat{p}_{m+1}}\right]^2 \left(\{p\}_m\big|\rho\right) \ . \qquad (2.15)$$

We have inserted the definition of the Monte Carlo time, t, that we here imagine using. For our present pedagogical purposes, the details of the definition of $\mathcal{H}_I(t)$ are not so important. What is important is that it reflects the factorization (2.14).

The second operator used in the construction of $\mathcal{U}(t', t)$ is a no-change operator $\mathcal{N}(t, t')$ with $\mathcal{N}(t, t) = 1$ and

$$\mathcal{N}(t_3, t_2)\mathcal{N}(t_2, t_1) = \mathcal{N}(t_3, t_1) \ . \qquad (2.16)$$

The no-change operator leaves the basis states unchanged except for multiplying each of them by an eigenvalue Δ:

$$\mathcal{N}(t', t)\big|\{p\}_m\right) = \Delta(t', t; \{p\}_m)\big|\{p\}_m\right) \ . \qquad (2.17)$$

The evolution operator $\mathcal{U}$ is expressed in terms of the hamiltonian and the no-change operators by (see figure 1)

$$\mathcal{U}(t_3, t_1) = \mathcal{N}(t_3, t_1) + \int_{t_1}^{t_3} dt_2 \, \mathcal{U}(t_3, t_2)\, \mathcal{H}_I(t_2)\, \mathcal{N}(t_2, t_1) \ . \qquad (2.18)$$

This equation is interpreted as saying that either the system evolves without splitting from t_1 to t_3, or else it evolves without splitting until an intermediate time t_2, splits at t_2, and then evolves (possibly with further splittings) from t_2 to t_3.

We need one more ingredient. We wish to construct the evolution so that it leaves the total cross section invariant, $\big(1\big|\mathcal{U}(t',t)\big|\rho\big) = \big(1\big|\rho\big)$. This should hold for every state $\big|\rho\big)$, so

$$\big(1\big|\mathcal{U}(t',t) = \big(1\big| \quad . \tag{2.19}$$

This assumption defines what $\mathcal{N}(t_3, t_1)$ has to be. Multiply eq. (2.18) on the left by $\big(1\big|$ and on the right by $\big|\{p\}_m\big)$. Then, using eqs. (2.19) and (2.17) we have

$$1 = \Delta(t_3, t_1; \{p\}_m) + \int_{t_1}^{t_3} dt_2 \, \big(1\big|\mathcal{H}_{\mathrm{I}}(t_2)\big|\{p\}_m\big) \, \Delta(t_2, t_1; \{p\}_m) \quad . \tag{2.20}$$

If we now differentiate with respect to t_3, we get

$$\frac{d}{dt_3} \, \Delta(t_3, t_1; \{p\}_m) = -\big(1\big|\mathcal{H}_{\mathrm{I}}(t_3)\big|\{p\}_m\big) \, \Delta(t_3, t_1; \{p\}_m) \quad . \tag{2.21}$$

The solution of this with the initial condition $\Delta(t_1, t_1; \{p\}_m) = 1$ is

$$\Delta(t_3, t_1; \{p\}_m) = \exp\left(-\int_{t_1}^{t_3} d\tau \, \big(1\big|\mathcal{H}_{\mathrm{I}}(\tau)\big|\{p\}_m\big) \right) \quad . \tag{2.22}$$

This result has a simple interpretation. The quantity $\big(1\big|\mathcal{H}_{\mathrm{I}}(\tau)\big|\{p\}_m\big)$ is the total probability for one of the partons in the state $\{p\}_m$ to split at time τ. The exponential, known as the Sudakov factor, is the probability that none of these partons splits between t_1 and t_3.

This operator notation provides a convenient way to express the essence of standard shower Monte Carlo algorithms. Different algorithms differ in how the evolution variable t is defined and in exactly what the splitting operator $\mathcal{H}_{\mathrm{I}}(t)$ is.

At its heart, the shower Monte Carlo idea is that soft interactions factor from hard interactions in QCD. After being produced in a hard interaction, partons in QCD travel a long way before undergoing much softer interactions such as splitting. What subsequent splitting does occur does not much change the basic hard matrix element for producing (almost) on-shell partons. There are some complications, however. On shell partons carry both spin and color. Furthermore, soft gluons can transmit color changes over long distances. In the subsequent sections, we will extend the meaning of the symbols used here so as to accommodate spin and color.

3. Structure of the calculation

In this section we introduce notational conventions and a general structure for the calculation that we will use later in the paper.

3.1 The space of quantum parton states

In order to describe showers, we need a notation for the description of quantum states consisting of two initial state partons and m final state partons. The partons are labeled by an index such as i that takes values "a" or "b" for the initial state partons and $1, 2, \ldots, m$

for the final state partons. Each parton is described by a momentum p, a flavor $f \in \{g, u, \bar{u}, d, \bar{d}, \ldots\}$, a spin index s and a color index c. We denote the quantum numbers of such a state by

$$\{p, f, s, c\}_m \equiv \{\eta_a, a, s_a, c_a; \eta_b, b, s_b, c_b; p_1, f_1, s_1, c_1; \ldots; p_m, f_m, s_m, c_m\} . \tag{3.1}$$

Here there is a special notation with respect to the incoming partons. The momentum fractions of the incoming partons are denoted by η_a and η_b, defined below. We call the flavor of parton "a" simply a and we call the flavor of parton "b" simply b. This notation is useful for designating the parton distribution functions, $f_{a/A}(\eta_a, \mu_F^2)$ and $f_{b/B}(\eta_b, \mu_F^2)$. For the purpose of describing backward evolution of the initial state partons, we will often need the antiflavors of the incoming partons. We use f_a and f_b for these,

$$\begin{aligned} f_a &= -a \ , \\ f_b &= -b \ . \end{aligned} \tag{3.2}$$

Here our notation is $-u = \bar{u}$, $-\bar{u} = u$, $-g = g$, etc.

The final state partons are always on-shell. Our kinematics allows parton masses, with

$$p_j^2 = m^2(f_j) . \tag{3.3}$$

To describe the momenta of the initial state partons, we start by defining p_A and p_B to be massless approximations to the momenta of the two incoming hadrons

$$\begin{aligned} p_A^2 &= 0 \ , \\ p_B^2 &= 0 \ , \\ 2p_A \cdot p_B &= s \ . \end{aligned} \tag{3.4}$$

The initial state partons are on-shell. In general, they can have masses but, with a small modification of the notation, their masses could be set to zero.[8] In any case, we take the initial state partons to have zero transverse momentum. Thus

$$\begin{aligned} p_a &= \eta_a p_A + \frac{m^2(f_a)}{\eta_a s} p_B \ , \\ p_b &= \eta_b p_B + \frac{m^2(f_b)}{\eta_b s} p_A \ . \end{aligned} \tag{3.5}$$

This defines the momentum fractions η_a and η_b.

In the event that we include parton masses, the hardness scale Q_0^2 at which the parton shower is initiated should be much bigger than the mass of any parton that is included as a possible constituent of the incoming hadrons. For example, if the hard process were $t + \bar{t}$ production near threshold, then top quarks should not be used as possible initial state partons. Thus we demand that

$$Q_0^2 > 4m_H^2 \ , \tag{3.6}$$

---

[8]There are several possibilities for the treatment of masses of initial state partons and there are some subtle issues associated with the choice. We mention some of these issues in section 14.

where m_H is the mass of the heaviest quark that is included as an initial state parton, typically the b quark. In any reasonable application of the formalism of this paper, the ">" here will be "$\gg$." Equation (3.6) suffices to make certain kinematic formulas in the paper work. At any stage in the shower, we define $p_a + p_b = Q$. At the first step, the starting hard scattering, with a sensible definition of the starting hardness scale we must have $Q^2 \geq Q_0^2$. The values of Q^2 increase as the shower develops. Thus at any stage we will have

$$(p_a + p_b)^2 > 4m_H^2 \ , \tag{3.7}$$

We will impose a kinematic restriction on the momentum fractions,

$$\frac{m^2(f_b)}{\eta_b s} < \eta_a \ , \qquad\qquad \frac{m^2(f_a)}{\eta_a s} < \eta_b \ . \tag{3.8}$$

These limits require that the momentum in the system in the direction of p_A comes mainly from parton "a," and the same for $p_A \leftrightarrow p_B$, a $\leftrightarrow$ b. We can be sure that both conditions hold by requiring

$$\eta_a \eta_b s > m_H^2 \ . \tag{3.9}$$

The reason for imposing this condition is as follows. Given eq. (3.5), there are two choices for $\eta_a \eta_b s$ that yield the same value of $(p_a + p_b)^2$. As long as condition (3.7) holds, the larger of the two choices for $\eta_a \eta_b s$ satisfies eq. (3.9). Imposing eq. (3.9) eliminates the other solution, in which parton "a" moves in approximately the p_B direction and parton "b" moves in approximately the p_A direction, creating a large value for $(p_a + p_b)^2$.

The upper limit on η_a and η_b is 1. We note here that this is an approximation. To discuss this, let n_A and n_B be dimensionless lightlike vectors in the directions of p_A and p_B respectively, normalized to $n_A \cdot n_B = 1$. Then consider, for example, the limit on η_a. The total momentum in the direction of n_A of the final state particles is

$$(p_a + p_b) \cdot n_B = \left[\eta_a + \frac{m^2(f_b)}{\eta_b s} \right] p_A \cdot n_B \ . \tag{3.10}$$

This can be bigger than the available momentum $p_A \cdot n_B$ if η_a is very close to 1. A remedy for this would be to redefine p_A and p_B in our formulas. Suppose that the exact hadron momenta are P_A and P_B. Then if we put

$$
\begin{aligned}
p_A &= \lambda \, P_A \cdot n_B \, n_A \ , \\
p_B &= \lambda \, P_B \cdot n_A \, n_B \ ,
\end{aligned}
\tag{3.11}
$$

then there is a value of λ that makes $(p_A + p_B)^2 = (P_A + P_B)^2$. By taking a value of λ that is a little smaller than this, one can ensure that the momentum in the final state in the directions of n_A and n_B is not more than was present in the initial state. It is this momentum that is available for the "underlying event." With this adjustment, the value of s in our formulas is a little less than the true c.m. squared energy, $(P_A + P_B)^2$.

Our notation with respect to spin and color is meant to be flexible. A standard helicity basis will work for spin. For color, we begin with a straightforward basis in which each parton i has a color index c_i that can take values 1,2,3 for quarks and antiquarks and

1,...,8 for gluons. Later, we will want to consider the subspace of the whole color space in which the parton state is a singlet under the SU(3) color group. We will choose a basis for this subspace. Using this basis, we will still have labels that we can call $\{c\}_m$, but the new labels will describe the color links among the partons rather than individual color indices c_i for the individual partons. The notation $\{p, f, c, s\}_m$ is thus supposed to include the possibility of any representation of the colors of the $m + 2$ partons.

3.2 The density matrix

A matrix element used in the computation of a cross section can be thought of as having the form of a function of the momenta and flavors that carries indices for spin and color,

$$M(\{p, f\}_m)_{s_a, s_b, s_1, \ldots, s_m}^{c_a, c_b, c_1, \ldots, c_m} \quad . \tag{3.12}$$

Here we denote the functions M for different numbers m of final state partons by the same name, simply $M(\{p, f\}_m)$ rather than $M_m(\{p, f\}_m)$. The array M can be thought of as a vector in spin and color space,

$$\big| M(\{p, f\}_m) \big\rangle \quad . \tag{3.13}$$

The inner product $\big\langle M' \big| M \big\rangle$ denotes multiplying M'^* by M and summing over the spins and colors.

An observable F can be specified by giving a set of functions $F(\{p, f\}_m)$ that are linear operators on the color-spin space. (In many important cases, $F(\{p, f\}_m)$ is a function times the unit operator on color-spin space. If F is be simply made from theta functions defining final state cuts, then $\sigma[F]$ is the cross section to find the final state partons within the cuts.) With this notation the cross section for an observable F takes the form[9]

$$\sigma[F] = \sum_m \frac{1}{m!} \int \big[d\{p, f\}_m \big] \frac{f_{a/A}(\eta_a, \mu_F^2) \, f_{b/B}(\eta_b, \mu_F^2)}{4 n_c(a) n_c(b) \, 2 \eta_a \eta_b p_A \cdot p_B} \tag{3.14}$$
$$\times \big\langle M(\{p, f\}_m) \big| F(\{p, f\}_m) \big| M(\{p, f\}_m) \big\rangle \quad .$$

Here the functions f are parton distribution functions while $n_c(a)$ is the number of colors that a parton of flavor a can have, $N_c = 3$ for a quark or antiquark, $N_c^2 - 1 = 8$ for a gluon. The factor $4 n_c(a) n_c(b)$ turns the sum over spins and colors for the initial state partons into an average over spins and colors. We have indicated the appropriate integrations over

---

[9]This formula contains a parton flux factor $2\eta_a \eta_b p_A \cdot p_B$ that corresponds to massless partons. The flux factor for scattering of free massive particles is more complicated. However, $(p_a + p_b)^2$ is always bigger than the initial hard scale Q_0^2 and the formalism of this paper is valid only when Q_0^2 is much larger than the masses of any initial state partons. For this reason, we use the flux factor for massless parton scattering.

momenta by

$$
\begin{aligned}
\int \big[d\{p,f\}_m\big] g(\{p,f\}_m) \equiv \prod_{i=1}^{m} \left\{ \sum_{f_i} \int \frac{d^4 p_i}{(2\pi)^4} \, 2\pi \delta_+(p_i^2 - m^2(f_i)) \right\} \sum_a \int_0^1 d\eta_{\mathrm a} \sum_b \int_0^1 d\eta_{\mathrm b} \\
\times (2\pi)^4 \delta\!\left(p_{\mathrm a} + p_{\mathrm b} - \sum_{i=1}^{m} p_i \right) \theta\big(m_{\mathrm H}^2 < \eta_{\mathrm a}\eta_{\mathrm b}s\big) \\
\times g(\{p,f\}_m) \ .
\end{aligned}
$$
(3.15)

Here $g(\{p,f\}_m)$ is an arbitrary function.

The final state particles carry labels $i \in \{1,\dots,m\}$. Then particle i has momentum, flavor, spin, and color given by $\{p_i, f_i, s_i, c_i\}$. One can arrange the definitions such that the amplitude M is symmetric under interchange of the labels. However, we do not necessarily do so. Instead, the notation allows for a general labeling scheme.[10] A measurement function F must be symmetric under interchange of labels, since the labels are not physical. Two amplitudes M that become the same if the labels are symmetrized are equivalent.

We will find it useful to rewrite $\sigma[F]$ in the form of a trace over the spin and color space,

$$
\sigma[F] = \sum_m \frac{1}{m!} \int \big[d\{p,f\}_m\big] \operatorname{Tr}\{\rho(\{p,f\}_m) F(\{p,f\}_m)\} \ ,
$$
(3.16)

where

$$
\rho(\{p,f\}_m) = \big|M(\{p,f\}_m)\big\rangle \frac{f_{a/A}(\eta_{\mathrm a},\mu_F^2) f_{b/B}(\eta_{\mathrm b},\mu_F^2)}{4 n_{\mathrm c}(a) n_{\mathrm c}(b) \, 2\eta_{\mathrm a}\eta_{\mathrm b} p_{\mathrm A}\cdot p_{\mathrm B}} \big\langle M(\{p,f\}_m)\big| \ .
$$
(3.17)

Thus ρ is the density operator in color $\otimes$ spin space. It is illustrated in figure 2. The density operator, for momentum as well as spin, is widely used as the basis of quantum statistical mechanics. It was introduced for the spin space in parton showers by Collins [9] and is used in HERWIG for the heavy partner particles in supersymmetry [11].

We can expand $\rho(\{p,f\}_m)$ in basis states $\big|\{s,c\}_m\big\rangle$ for the color $\otimes$ spin space,

$$
\rho(\{p,f\}_m) = \sum_{s,c} \sum_{s',c'} \big|\{s,c\}_m\big\rangle \, \rho(\{p,f,s',c',s,c\}_m) \, \big\langle\{s',c'\}_m\big| \ .
$$
(3.18)

Here $\rho(\{p,f,s',c',s,c\}_m)$ is a function that depends on the momenta and flavors $\{p,f\}_m$, the labels $\{s,c\}_m$ for the quantum "ket" state and the labels $\{s',c'\}_m$ for the conjugate quantum "bra" state. We denote the state labels collectively by $\{p,f,s',c',s,c\}_m$. Thus $\rho(\{p,f,s',c',s,c\}_m)$ a function giving the matrix elements of the density matrix. We find it convenient to base our treatment on this function.

---

[10]Just to take a trivial example, in a $u\bar u g$ state, the label 1 might be assigned to the up quark, 2 to the anti-up quark and 3 to the gluon. Of course, it will not work in general to use the flavors as labels because one can have two final state partons with the same label.
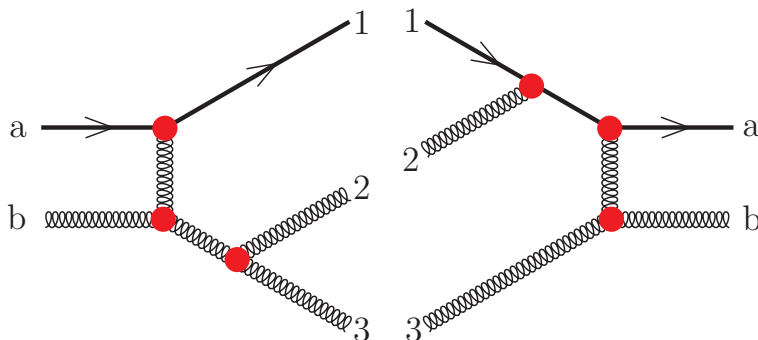
Figure 2: Illustration of a contribution to ρ, eq. (3.17). The Feynman graph on the left is a contribution to $\left|M(\{p,f\}_m)\right\rangle$ and the Feynman graph on the right is a contribution to $\left\langle M(\{p,f\}_m)\right|$. The factor with parton distributions is not illustrated. The momenta and flavors of the labeled partons match between $\left|M(\{p,f\}_m)\right\rangle$ and $\left\langle M(\{p,f\}_m)\right|$, but the partons on the left have colors and spins $\{s,c\}_m$ while those on the right have possibly different colors and spins $\{s',c'\}_m$. Additionally, ρ can contain quantum interference between different graphs, as illustrated here. We imagine that eventually the partons will evolve to form a final state in the middle and an initial state on the left and right.

Later, we will choose an orthonormal basis for the spin space, so that

$$\left\langle\{s'\}_m\big|\{s\}_m\right\rangle = \delta_{\{s\}_m}^{\{s'\}_m} \quad . \tag{3.19}$$

For the color space, we will find it convenient to use a basis in which $\left\langle\{c\}_m\big|\{c\}_m\right\rangle$ is approximately but not exactly 1 and where $\left\langle\{c'\}_m\big|\{c\}_m\right\rangle$ is not generally zero for $\{c\}_m \neq \{c'\}_m$. With a non-orthogonal basis, we need to be a little careful about the notation. Suppose that we expand a vector in such a basis,

$$|\psi\rangle = \sum_{\{c\}_m} \big|\{c\}_m\big\rangle\, a(\{c\}_m) \quad . \tag{3.20}$$

A convenient representation for the expansion coefficients $a(\{c\}_m)$ is obtained by taking matrix elements with elements of the dual basis $\big|\{c\}_m\big\rangle_D$ defined[11] by

$$_D\!\left\langle\{c'\}_m\big|\{c\}_m\right\rangle = \delta_{\{c\}_m}^{\{c'\}_m} \quad . \tag{3.21}$$

Using the dual basis we can write

$$_D\!\left\langle\{c\}_m\big|\psi\right\rangle = a(\{c\}_m) \quad . \tag{3.22}$$

The expansion of any vector in the basis $\big|\{c\}_m\big\rangle$ can conveniently be obtained directly by using the completeness relation

$$1 = \sum_{\{c\}_m} \big|\{c\}_m\big\rangle \; _D\!\left\langle\{c\}_m\big| \quad . \tag{3.23}$$

---

[11]Readers familiar with general relativity can think of $\left\langle\{c\}_m|a\right\rangle$ as the covariant components a_c of a and $_D\!\left\langle\{c\}_m|a\right\rangle$ as the contravariant components a^c. One can lower indices using the metric tensor $g_{cc'}$, analogous to $\left\langle\{c\}_m|\{c'\}_m\right\rangle$. The inverse matrix $_D\!\left\langle\{c\}_m|\{c'\}_m\right\rangle_D$ is analogous to $g^{cc'}$.

If we want to expand a bra vector in the basis $\langle\{c\}_m|$ we can use the completeness relation

$$1 = \sum_{\{c\}_m} |\{c\}_m\rangle_D \langle\{c\}_m| \ . \tag{3.24}$$

This is particularly useful with respect to operators. Let O be a linear operator on the color space. If $|\psi\rangle$ has expansion coefficients $a(\{c\}_m)$ defined by eq. (3.20) and $|\psi'\rangle = O|\psi\rangle$ has expansion coefficients $a'(\{c'\}_m)$ then O is conventionally described by the matrix defined by

$$a'(\{c'\}_m) = \sum_{\{c\}_m} o(\{c'\}_m, \{c\}_m)\, a(\{c\}_m) \ . \tag{3.25}$$

We can write this as

$$o(\{c'\}_m, \{c\}_m) = {}_D\langle\{c'\}_m|O|\{c\}_m\rangle \ . \tag{3.26}$$

3.3 Statistical states

The set of functions $\rho(\{p, f, s', c', s, c\}_m)$ forms a vector space, which we can call the statistical state space (as distinct from the quantum state space). We can call the vector corresponding to this function simply $|\rho)$. Note the rounded brackets instead of angle brackets that we use for quantum states, as in $|\psi\rangle$. We can define basis vectors $(\{p, f, s', c', s, c\}_m|$ for the statistical state space[12] so that

$$\rho(\{p, f, s', c', s, c\}_m) = (\{p, f, s', c', s, c\}_m|\rho) \ . \tag{3.27}$$

There are also ket basis vectors such that the completeness relation for the basis states is

$$1 = \sum_m \frac{1}{m!} \int \left[d\{p, f, s', c', s, c\}_m\right] |\{p, f, s', c', s, c\}_m)(\{p, f, s', c', s, c\}_m| \ , \tag{3.28}$$

where $\left[d\{p, f, s, s', c, c'\}_m\right]$ is an extension of the integration measure eq. (3.15)

$$\int \left[d\{p, f, s', c', s, c\}_m\right] \equiv \int \left[d\{p, f\}_m\right] \sum_{s_a, s'_a, c_a, c'_a} \sum_{s_b, s'_b, c_b, c'_b} \prod_{i=1}^m \left\{ \sum_{s_i, s'_i, c_i, c'_i} \right\} \ . \tag{3.29}$$

The corresponding inner product of basis states is

$$(\{p, f, s', c', s, c\}_m|\{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_{\tilde{m}}) = \delta_{m, \tilde{m}}\, \delta(\{p, f, s', c', s, c\}_m; \{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m) \ , \tag{3.30}$$

where the function δ is a generalization of the ordinary delta-function that is defined by

$$\frac{1}{m!} \int \left[d\{p, f, s', c', s, c\}_m\right] \delta(\{p, f, s', c', s, c\}_m; \{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m)\, h(\{p, f, s', c', s, c\}_m)$$
$$= h(\{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m) \ . \tag{3.31}$$

---

[12]More precisely, the bra vectors $(F|$ are vectors in the dual space to the ket vectors $|\rho)$, that is the space of linear functions on the vectors $|\rho)$.

– 17 –

If we want to expand a bra vector in the basis $\langle\{c\}_m|$ we can use the completeness relation

$$1 = \sum_{\{c\}_m} |\{c\}_m\rangle_D \langle\{c\}_m| \ . \tag{3.24}$$

This is particularly useful with respect to operators. Let O be a linear operator on the color space. If $|\psi\rangle$ has expansion coefficients $a(\{c\}_m)$ defined by eq. (3.20) and $|\psi'\rangle = O|\psi\rangle$ has expansion coefficients $a'(\{c'\}_m)$ then O is conventionally described by the matrix defined by

$$a'(\{c'\}_m) = \sum_{\{c\}_m} o(\{c'\}_m, \{c\}_m)\, a(\{c\}_m) \ . \tag{3.25}$$

We can write this as

$$o(\{c'\}_m, \{c\}_m) = {}_D\langle\{c'\}_m|O|\{c\}_m\rangle \ . \tag{3.26}$$

3.3 Statistical states

The set of functions $\rho(\{p, f, s', c', s, c\}_m)$ forms a vector space, which we can call the statistical state space (as distinct from the quantum state space). We can call the vector corresponding to this function simply $|\rho)$. Note the rounded brackets instead of angle brackets that we use for quantum states, as in $|\psi\rangle$. We can define basis vectors $(\{p, f, s', c', s, c\}_m|$ for the statistical state space[12] so that

$$\rho(\{p, f, s', c', s, c\}_m) = (\{p, f, s', c', s, c\}_m|\rho) \ . \tag{3.27}$$

There are also ket basis vectors such that the completeness relation for the basis states is

$$1 = \sum_m \frac{1}{m!} \int \left[d\{p, f, s', c', s, c\}_m\right] |\{p, f, s', c', s, c\}_m)(\{p, f, s', c', s, c\}_m| \ , \tag{3.28}$$

where $\left[d\{p, f, s, s', c, c'\}_m\right]$ is an extension of the integration measure eq. (3.15)

$$\int \left[d\{p, f, s', c', s, c\}_m\right] \equiv \int \left[d\{p, f\}_m\right] \sum_{s_a, s'_a, c_a, c'_a} \sum_{s_b, s'_b, c_b, c'_b} \prod_{i=1}^m \left\{ \sum_{s_i, s'_i, c_i, c'_i} \right\} \ . \tag{3.29}$$

The corresponding inner product of basis states is

$$(\{p, f, s', c', s, c\}_m|\{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_{\tilde{m}}) = \delta_{m, \tilde{m}}\, \delta(\{p, f, s', c', s, c\}_m; \{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m) \ , \tag{3.30}$$

where the function δ is a generalization of the ordinary delta-function that is defined by

$$\frac{1}{m!} \int \left[d\{p, f, s', c', s, c\}_m\right] \delta(\{p, f, s', c', s, c\}_m; \{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m)\, h(\{p, f, s', c', s, c\}_m)$$
$$= h(\{\tilde{p}, \tilde{f}, \tilde{s}', \tilde{c}', \tilde{s}, \tilde{c}\}_m) \ . \tag{3.31}$$

---

[12]More precisely, the bra vectors $(F|$ are vectors in the dual space to the ket vectors $|\rho)$, that is the space of linear functions on the vectors $|\rho)$.

Here h is any well behaved function of the variables indicated, defined on the integration surface.

Let us define a vector corresponding to a measurement function F using[13]

$$\left(F\middle|\{p,f,s',c',s,c\}_m\right) = \left\langle\{s',c'\}_m\middle|F(\{p,f\}_m)\middle|\{s,c\}_m\right\rangle \ . \tag{3.32}$$

Then, using the completeness relation (3.28), the cross section (3.16) corresponding to a measurement function F can be expressed using eqs. (3.18) and (3.27) as

$$\sigma[F] = \left(F\middle|\rho\right) \ . \tag{3.33}$$

In the case that F consists of a function $F(\{p,f\}_m)$ times a unit operator in spin-color space, the inner product $\left(F\middle|\rho\right)$ is

$$\left(F\middle|\rho\right) = \sum_m \frac{1}{m!} \int \left[d\{p,f,s',c',s,c\}_m\right] F(\{p,f\}_m) \left\langle\{s'\}_m\middle|\{s\}_m\right\rangle\left\langle\{c'\}_m\middle|\{c\}_m\right\rangle$$
$$\times \rho(\{p,f,s',c',s,c\}_m) \ . \tag{3.34}$$

3.4 The resolution scale

We now need to introduce a resolution scale into our equations. We first discuss the resolution scale of the observable. In eq. (3.33), let us suppose that the observable represented by the functions $F(\{p,f\}_m)$ is infrared safe. To be precise about what this means, we first demand that the functions $F(\{p,f\}_m)$ be smooth functions of the momentum variables and that they be invariant under label interchanges. Then we consider a list of parton variables $\{\hat{p},\hat{f}\}_{m+1}$ for $m+1$ partons and suppose that $\hat{p}_{m+1}$ becomes collinear with the momentum $\hat{p}_l$ of parton l. Then we can consider the list of parton momenta $\{p,f\}_m$ where $p_j = \hat{p}_j$ and $f_j = \hat{f}_j$ for $j \neq l$ while $p_l = \hat{p}_m + \hat{p}_l$ and $f_l = \hat{f}_m + \hat{f}_l$ (with the obvious definition of adding flavors). That is, the partons with variables $\{\hat{p},\hat{f}\}_{m+1}$ could have arisen from partons with variables $\{p,f\}_m$ by the collinear splitting of parton l into new partons with labels l and $m+1$. Then infrared safety requires that

$$F(\{\hat{p},\hat{f}\}_{m+1}) \to F(\{p,f\}_m) \tag{3.35}$$

in the limit. This should also hold if parton $m+1$ becomes soft, $\hat{p}_{m+1} \to 0$. It should also hold with a suitable adjustment of the notation for a splitting of one of the initial state partons. This is, so far, just the standard definition of infrared safety. It allows us to have perturbatively calculable cross sections. Now let us extend the definition to include a scale. We can say that the observable is infrared safe at scale μ^2 if

$$F(\{\hat{p},\hat{f}\}_{m+1}) \approx F(\{p,f\}_m) \tag{3.36}$$

when $|2\hat{p}_{m+1}\cdot\hat{p}_l| < \mu^2$. To be really precise, we should specify how good this approximation has to be, but this will not matter for our purposes. Strictly speaking an "infrared safe

---

[13]Note that this equation for the measurement function has a different structure from the equation used to define the statistical state vector $|\rho)$, $\rho(\{p,f\}_m) = |\{s,c\}_m)(\{p,f,s',c',s,c\}_m|\rho)\langle\{s',c'\}_m|$.

observable" without further qualification is one that is infrared safe at any scale, no matter how small. However, what is usually meant is that it is infrared safe at a scale not much smaller than the scale Q_0^2 of the hardest interaction in the problem. What we want to do here is to specify the splitting scale at which the observable is sensitive to the splitting. That scale could be much smaller than Q_0^2.

Now we introduce the concept of the density operator ρ evaluated at resolution scale μ^2. The idea is that interactions with scales greater than μ^2 are included in $|M\rangle\langle M|$ while interactions with scales smaller than μ^2 are integrated out (for final state interactions) or included in the parton distributions (for initial state interactions).[14] We can describe this in a rough way as follows. In each cut Feynman diagram, each integration region for final state partons that produces a collinear or soft divergence can be described as a region in which some group of partons with labels i become collinear to a given direction of a mother parton or some become soft, so that $(\sum_i p_i)^2 \to 0$. Divide this region into subregions with $(\sum_i p_i)^2 < \mu^2$ and $(\sum_i p_i)^2 > \mu^2$. In the $(\sum_i p_i)^2 < \mu^2$ region we can combine the partons i into a single effective parton for purposes of calculating the observable. That is, the partons i are "unresolved." This leads to a free integration over this region using a constant F for fewer partons. Adding these real emission integrals to the corresponding virtual diagrams and counterterms from the parton distributions gives a finite result containing logarithms of μ^2/Q_0^2.

A more intuitive way of thinking about this is to imagine writing the Feynman diagrams in a coordinate space representation, in which we integrate over the positions x_i^μ of the interaction vertices relative to the position of the hard interaction, which is determined to within $1/Q_0^2$. Then we can restrict these integrations to $|x_i^2| < 1/\mu^2$ before integrating over the final state momenta.

In order that the parton distributions in eq. (3.17) include initial state interactions at all scales smaller than μ^2, the factorization scale $\mu_{\rm F}^2$ at which the parton distributions are evaluated should be μ^2.

With this meaning of ρ evaluated at resolution scale μ^2, $(F|\rho)$ is invariant under $\mu^2 \to \mu^2 + \delta\mu^2$ as long as F is infrared safe at a scale equal to μ^2 or larger. However, if one looks at $|\rho)$ with a resolution scale smaller than μ^2 then $(F|\rho)$ does see the effect of changing μ^2.

Our object in this paper is to construct an approximate version of $|\rho)$ as a function of the resolution scale. The idea is to construct $|\rho)$ using a parton shower, starting from $\mu^2 = Q_0^2$ and evolving to smaller values of μ^2, down to a final infrared cutoff. This intuitively appealing idea has been inherent in the idea of a parton shower since the earliest days. It would be very useful to have a precise field theoretic definition of $|\rho)$ as a function of resolution scale. However, this is beyond our scope in this paper.

One finds that the logarithm of the resolution scale μ^2 is more useful as a variable than μ^2 itself. Therefore we define Monte Carlo time t by

$$\mu^2 = Q_0^2 \, e^{-t} \tag{3.37}$$

---

[14]This is the idea of the standard factorization theorem [20]. However, we here go beyond anything that has been proved.

and write the density operator with this resolution scale as $\big|\rho(t)\big)$. Then also the parton distributions are evaluated at factorization scale

$$\mu_{\mathrm{F}}^2 = Q_0^2\, e^{-t} \ . \tag{3.38}$$

3.5 Parton shower evolution

We are now prepared to set up a quite general framework for describing a parton shower. We take the framework as a set of axioms that we hope are reasonably intuitive. Later, we relate the operators that occur to the structure of Green functions at tree level in QCD.

We use the evolution variable t that specifies the resolution scale of ρ according to eq. (3.37). Thus t starts at zero and increasing t corresponds to decreasing virtuality. One stops evolution at an infrared cutoff t_{f} at which the use of an evolution based on perturbation theory is no longer appropriate. For instance, t_{f} might correspond to a 1 GeV^2 virtuality. The evolving shower is represented by a state $\big|\rho(t)\big)$ that begins with an initial state $\big|\rho(0)\big)$. The evolution is given by a linear operator $\mathcal{U}(t,t')$, with

$$\big|\rho(t)\big) = \mathcal{U}(t,t')\big|\rho(t')\big) \ . \tag{3.39}$$

Here $\mathcal{U}(t,t) = 1$. These operators have the group composition property

$$\mathcal{U}(t_3,t_2)\,\mathcal{U}(t_2,t_1) = \mathcal{U}(t_3,t_1) \ . \tag{3.40}$$

The class of evolution operators that will use is defined by two operators, $\mathcal{H}_{\mathrm{I}}(t)$ and $\mathcal{V}(t)$, according to the differential equation

$$\frac{d}{dt}\mathcal{U}(t,t') = [\mathcal{H}_{\mathrm{I}}(t) - \mathcal{V}(t)]\,\mathcal{U}(t,t') \ , \tag{3.41}$$

with initial condition $\mathcal{U}(t,t) = 1$.

The first operator, $\mathcal{H}_{\mathrm{I}}(t)$, represents parton interactions and, in general, changes the number of partons and their momenta. We specify $\mathcal{H}_{\mathrm{I}}(t)$ by giving its matrix elements

$$\big(\{\hat{p},\hat{f},\hat{s}',\hat{c}',\hat{s},\hat{c}\}_{m'}\big|\mathcal{H}_{\mathrm{I}}(t)\big|\{p,f,s',c',s,c\}_m\big) \ . \tag{3.42}$$

In a lowest order shower, which we consider in this paper, the operator $\mathcal{H}_{\mathrm{I}}(t)$ describes $1 \to 2$ parton splitting, changing a state with m final state partons to one with $m+1$ final state partons.

The second operator that controls evolution, $\mathcal{V}(t)$, describes the effect of virtual graphs and the "unresolved" part of real emission graphs. In a lowest order shower, we do not account for the virtual graphs exactly, but rather account for only the infrared singular part of the virtual graphs, which can be deduced from the real emission graphs. That is, $\mathcal{V}(t)$ is determined from $\mathcal{H}_{\mathrm{I}}(t)$. The operator $\mathcal{V}(t)$ does not change the number of partons or their flavors or spins, but can change their color states.

We construct the shower algorithm in such a way that it conserves probability in a certain sense. If we were dealing with $e^+ + e^- \to hadrons$, we would demand that the development of the shower does not change the total cross section. For hadron-hadron collisions,

the total cross section does not have a well defined perturbative expansion. However, if we start with a state $|\rho(0))$ that is "hard" in the sense that $(\{p, f, s', c', s, c\}_m|\rho(0))$ is non-zero only for parton configurations with a large transverse energy, then we can demand that as this state evolves into the shower, the contribution from $|\rho(t))$ to the total cross section does not change. The observable that measures the total cross section is

$$F_1(\{p, f\}_m) = 1 \ . \tag{3.43}$$

We will call the vector corresponding to F_1 simply $(1|$. Using eq. (3.32), the inner product of $(1|$ with a basis state is

$$(1|\{p, f, s', c', s, c\}_m) = \langle\{s'\}_m|\{s\}_m\rangle \langle\{c'\}_m|\{c\}_m\rangle \ . \tag{3.44}$$

The statement that shower evolution leaves contributions to the total cross section invariant is

$$(1|\mathcal{U}(t', t)|\rho) = (1|\rho) \tag{3.45}$$

for any (suitably hard) state $|\rho)$. This requirement leads to a relation between the matrix elements of $\mathcal{V}$, and $\mathcal{H}_{\mathrm{I}}$. To derive this relation we multiply eq. (3.41) on the left by $(1|$ and on the right by $|\rho)$. After using eq. (3.45), we get

$$0 = (1|[\mathcal{H}_{\mathrm{I}}(t) - \mathcal{V}(t)]\mathcal{U}(t, t')|\rho) \ . \tag{3.46}$$

Since this holds for any suitably hard state $|\rho)$ we have

$$0 = (1|[\mathcal{H}_{\mathrm{I}}(t) - \mathcal{V}(t)] \ . \tag{3.47}$$

We multiply on the right by $|\{p, f, s', c', s, c\}_m)$ to obtain

$$(1|\mathcal{V}(t)|\{p, f, s', c', s, c\}_m) = (1|\mathcal{H}_{\mathrm{I}}(t)|\{p, f, s', c', s, c\}_m) \ . \tag{3.48}$$

At this point, we need to discuss the structure of the parton splitting operator $\mathcal{H}_{\mathrm{I}}$. In subsequent sections, we derive the form of $\mathcal{H}_{\mathrm{I}}(t)$ based on the structure of QCD tree level matrix elements in the limit that two of the $m+1$ partons become massless and collinear, one becomes massless and collinear with one of the beam directions, or one (a gluon) becomes soft. In this limit, the matrix elements take a factored form, *hard* $\otimes$ (*soft & collinear*). This factorization leads to the definition[15] of $\mathcal{H}_{\mathrm{I}}(t)$ in section 9. At the moment, what we need is the structure of $(1|\mathcal{H}_{\mathrm{I}}(t)|\{p, f, s', c', s, c\}_m)$, which represents the inclusive splitting probability at splitting scale t. We will find

$$(1|\mathcal{H}_{\mathrm{I}}(t)|\{p, f, s', c', s, c\}_m) = 2\langle\{s'\}_m|\{s\}_m\rangle \langle\{c'\}_m|h(t, \{p, f\}_m)|\{c\}_m\rangle \ , \tag{3.49}$$

where the function $h(t, \{p, f\}_m)$ is given in section 12, eq. (12.21). The important point is that there is a trivial spin structure and a non-trivial color structure.

---

[15]The definition is not unique because there is freedom to choose what to do away from the soft and collinear limits

We take the operator $\mathcal{V}(t)$ to operate only on the color space and define its action in terms of its matrix elements, for which we use the notation $\left(\{\hat{c}', \hat{c}\}_m \big| \mathcal{V}(t; \{p, f\}_m) \big| \{c', c\}_m\right)$. The definition of $\mathcal{V}(t)$ in terms of these matrix elements is

$$\mathcal{V}(t)\big|\{p, f, s', c', s, c\}_m\big) = \sum_{\{\hat{c}', \hat{c}\}_m} \big|\{p, f, s', \hat{c}', s, \hat{c}\}_m\big) \left(\{\hat{c}', \hat{c}\}_m \big| \mathcal{V}(t; \{p, f\}_m) \big| \{c', c\}_m\right) \ .$$

(3.50)

Thus, using eq. (3.44),

$$\big(1\big|\mathcal{V}(t)\big|\{p, f, s', c', s, c\}_m\big) = \sum_{\{\hat{c}', \hat{c}\}_m} \big\langle\{s'\}_m\big|\{s\}_m\big\rangle\big\langle\{\hat{c}'\}_m\big|\{\hat{c}\}_m\big\rangle$$
$$\times \left(\{\hat{c}', \hat{c}\}_m \big| \mathcal{V}(t; \{p, f\}_m) \big| \{c', c\}_m\right) \ .$$

(3.51)

If we insert eq. (3.49) and eq. (3.51) into eq. (3.48) and cancel the spin factors, we get

$$\sum_{\{\hat{c}', \hat{c}\}_m} \big\langle\{\hat{c}'\}_m\big|\{\hat{c}\}_m\big\rangle\big(\{\hat{c}', \hat{c}\}_m\big|\mathcal{V}(t; \{p, f\}_m)\big|\{c', c\}_m\big)$$
$$= 2\big\langle\{c'\}_m\big|h(t, \{p, f\}_m)\big|\{c\}_m\big\rangle \ .$$

(3.52)

There is a simple way to satisfy this equation. We define

$$\big(\{\hat{c}', \hat{c}\}_m\big|\mathcal{V}(t, \{p, f\}_m)\big|\{c', c\}_m\big) = {}_D\big\langle\{\hat{c}\}_m\big|h(t, \{p, f\}_m)\big|\{c\}_m\big\rangle \, \delta^{\{\hat{c}'\}_m}_{\{c'\}_m}$$
$$+ \delta^{\{\hat{c}\}_m}_{\{c\}_m} \, \big\langle\{c'\}_m\big|h(t, \{p, f\}_m)\big|\{\hat{c}'\}_m\big\rangle_D \ .$$

(3.53)

What we have done here is to decompose $\big(\{\hat{c}', \hat{c}\}_m\big|\mathcal{V}(t, \{p, f\}_m)\big|\{c', c\}_m\big)$ into two terms. In the first term, nothing happens on the bra side of the density matrix but there is a virtual correction on the ket side, while in the second term nothing happens on the ket side of the density matrix but there is a virtual correction on the bra side.[16] With this definition, when we perform the sums over $\{\hat{c}\}_m$ and $\{\hat{c}'\}_m$ in eq. (3.52) using eq. (3.23) we see that eq. (3.52) is satisfied.

We note that one could imagine solving eq. (3.41) numerically in the form

$$\mathcal{U}(t + \Delta t, 0)\big|\rho(0)\big) = [1 - \mathcal{V}(t)\Delta t]\mathcal{U}(t, 0)\big|\rho(0)\big) + \mathcal{H}_{\mathrm{I}}(t)\Delta t\,\mathcal{U}(t, 0)\big|\rho(0)\big) \ .$$

(3.54)

That is, one could use small time steps in which either one of the partons splits or else no parton splits and the weights for different color states are readjusted. However, this is not the way that shower evolution is typically constructed.

To proceed down a more traditional path, we define two other operators, $\mathcal{V}_{\mathrm{E}}(t)$ and $\mathcal{V}_{\mathrm{S}}(t)$ with sum

$$\mathcal{V}_{\mathrm{E}}(t) + \mathcal{V}_{\mathrm{S}}(t) = \mathcal{V}(t) \ .$$

(3.55)

---

[16]What $\mathcal{V}$ contains is the singular parts of the virtual corrections, which are related to the collinear and soft singularities of the real emission diagrams, plus an "unresolved" contribution from the real emission diagrams. Thus we obtain $\mathcal{V}$ from $\mathcal{H}$.

The distinction between $\mathcal{V}_{\mathrm{E}}(t)$ and $\mathcal{V}_{\mathrm{S}}(t)$ lies in how we treat them within shower generation: $\mathcal{V}_{\mathrm{E}}(t)$ is exponentiated and $\mathcal{V}_{\mathrm{S}}(t)$ is subtracted. We express the solution of eq. (3.41) in the form

$$\mathcal{U}(t,t') = \mathcal{N}(t,t') + \int_{t'}^{t} d\tau \, \mathcal{U}(t,\tau) \left[\mathcal{H}_{\mathrm{I}}(\tau) - \mathcal{V}_{\mathrm{S}}(\tau) \right] \mathcal{N}(\tau,t') \ , \tag{3.56}$$

where the operator $\mathcal{N}(t,t')$ is the time ordered exponential of the operator $\mathcal{V}_{\mathrm{E}}(t)$,

$$\mathcal{N}(t,t') = \mathbb{T} \exp \left\{ - \int_{t'}^{t} d\tau \, \mathcal{V}_{\mathrm{E}}(\tau) \right\} \ . \tag{3.57}$$

Here $\mathbb{T}$ represents the ordering in evolution time t. The operator $\mathcal{N}(t,t')$ is a generalization of the standard Sudakov exponential in parton shower Monte Carlo programs. It has the group multiplication property

$$\mathcal{N}(t_3,t_2) \, \mathcal{N}(t_2,t_1) = \mathcal{N}(t_3,t_1) \tag{3.58}$$

and satisfies the differential equation

$$\frac{d}{dt} \mathcal{N}(t,t') = -\mathcal{V}_{\mathrm{E}}(t) \, \mathcal{N}(t,t') \ . \tag{3.59}$$

Eq. (3.56) is interpreted as saying that either the system evolves without splitting from t' to t, or else it evolves without splitting until an intermediate time τ, splits or undergoes a color change at τ, and then evolves (possibly with further splittings or color changes) from τ to t. The first term contains a summation of effects from virtual splittings. In the second term we have a parton splitting contribution along with a subtraction that arises from the part of the virtual splitting contribution that was not summed to form part of $\mathcal{N}(t,t')$.

Now the operator $\mathcal{V}(t)$ is completely defined but we still need to define $\mathcal{V}_{\mathrm{E}}(t)$ and $\mathcal{V}_{\mathrm{S}}(t)$. Here we have some freedom. There are at least three obvious choices:

1. We could define $\mathcal{V}_{\mathrm{E}}(t) = 0$. Then $\mathcal{V}_{\mathrm{S}}(t) = \mathcal{V}(t)$. This choice leads to a trivial Sudakov exponential, $\mathcal{N}(t,t') = 1$. This is similar to what one does in fixed order calculations when the singularities of the real emission graphs with $m+1$ final state partons are removed by the subtraction terms with m partons. This is not useful in the context of a parton shower.

2. We could define $\mathcal{V}_{\mathrm{S}}(t) = 0$. Then $\mathcal{V}_{\mathrm{E}}(t) = \mathcal{V}(t)$. This means we exponentiate the whole virtual splitting operator. It is the most "shower way" to organize the parton evolution. The integral of $\mathcal{V}(t)$ over a range of t produces large logarithms and all of these logarithms appear in the Sudakov exponent. Since $\mathcal{V}(t)$ is a non-diagonal matrix in color space, the implementation of this choice may present difficulties.

3. Alternatively we can define $\mathcal{V}_{\mathrm{E}}(t)$ to be the diagonal in color as follows:

$$\begin{aligned}
\left(\{\hat{c}',\hat{c}\}_m \big| \mathcal{V}_{\mathrm{E}}(t,\{p,f\}_m) \big| \{c',c\}_m\right) = \delta_{\{c\}_m}^{\{\hat{c}\}_m} \delta_{\{c'\}_m}^{\{\hat{c}'\}_m} \times \big[&\langle \{c\}_m | h(t,\{p,f\}_m) | \{c\}_m \rangle \\
&+ \langle \{c'\}_m | h(t,\{p,f\}_m) | \{c'\}_m \rangle \big] \ .
\end{aligned} \tag{3.60}$$

Then $\mathcal{V}_{\mathrm{S}}(t)$ is

$$\mathcal{V}_{\mathrm{S}}(t, \{p, f\}_m) = \mathcal{V}(t, \{p, f\}_m) - \mathcal{V}_{\mathrm{E}}(t, \{p, f\}_m) \ . \tag{3.61}$$

With this choice, the computation of the operator $\mathcal{N}(t, t')$ is simple because the basis vectors in the statistical space are eigenvectors of the operator $\mathcal{V}_{\mathrm{E}}(t)$.

With this alternative, we do not exponentiate everything. However, when we study color in the following sections, we will see with this choice that (with the color basis that we will choose) $\mathcal{V}_{\mathrm{S}}(t)$ is small compared to $\mathcal{V}_{\mathrm{E}}(t)$. First, we have

$$\left\langle \{c\}_m \middle| h(t, \{p, f\}_m) \middle| \{c\}_m \right\rangle = \sum_{\{\tilde{c}\}_m} \left\langle \{c\}_m \middle| \{\tilde{c}\}_m \right\rangle {}_D\!\left\langle \{\tilde{c}\}_m \middle| h(t, \{p, f\}_m) \middle| \{c\}_m \right\rangle \ ,$$

$$\left\langle \{c'\}_m \middle| h(t, \{p, f\}_m) \middle| \{c'\}_m \right\rangle = \sum_{\{\tilde{c}\}_m} \left\langle \{c'\}_m \middle| h(t, \{p, f\}_m) \middle| \{\tilde{c}\}_m \right\rangle_D \left\langle \{\tilde{c}\}_m \middle| \{c'\}_m \right\rangle \ .$$

$$\tag{3.62}$$

We will see in eq. (7.14) that the matrix $\left\langle \{c\}_m \middle| \{\tilde{c}\}_m \right\rangle$ is the unit matrix except for $1/N_{\mathrm{c}}^2$ corrections. Thus the difference between using dual basis vectors $\left| \{\tilde{c}\}_m \right\rangle_D$ in eq. (3.53) and ordinary basis vectors $\left| \{c'\}_m \right\rangle$ is not important in the large N_{c} limit. Second, we will see in section 12 that the matrix $\left\langle \{\hat{c}\}_m \middle| h(t, \{p, f\}_m) \middle| \{c\}_m \right\rangle$ is almost diagonal in the sense that its off-diagonal matrix elements are suppressed compared to its diagonal matrix elements by factors of $1/N_{\mathrm{c}}^2$. Thus the part of the virtual contribution that is not exponentiated is small. This small part is not neglected, but we can leave it out of the Sudakov exponent and treat it as a subtraction instead. Effectively, this means that we treat $1/N_{\mathrm{c}}^2$ as a small parameter in addition to α_{s}.

4. Momentum and flavor mapping

In this and the following sections, we explore how to define the splitting operator $\mathcal{H}_I$. The first issue to examine is the momentum mapping. We begin with an m parton state with momenta $\{p\}_m$. One of the partons, with label $l \in \{\mathrm{a}, \mathrm{b}, 1, \ldots, m\}$, splits. After the splitting, we have an $m + 1$ parton state with momenta $\{\hat{p}\}_{m+1}$. Our notation is that parton l splits into partons with labels l and $m + 1$, while the other partons keep their labels. The momenta $\{\hat{p}\}_{m+1}$ after splitting are determined by the momenta $\{p\}_m$ and a momentum splitting variable that we call ζ_{p}, which defines the momenta of the daughter partons. There is also a flavor splitting variable, ζ_{f}, which tells the daughter flavors.

In this section, we first describe the splitting of a final state parton, then move on to the somewhat more complicated splitting of an initial state parton. For the final state splitting we first describe how $\{\hat{p}, \hat{f}\}_{m+1}$ is determined from $\{p, f\}_m$ and $\{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}$. Then we state the inverse transformation, from $\{\hat{p}, \hat{f}\}_{m+1}$ to $\{p, f\}_m$ and $\{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}$. Finally, we deduce the jacobian for this change of variables. We will then be ready to do the same thing for the splitting of an initial state parton.

There are many ways to define the momentum mapping $\{p\}_m \leftrightarrow \{\hat{p}\}_{m+1}$. One of the most successful is that of Catani and Seymour [14]. This may be called a local mapping:

the momenta of most of the partons are left unchanged, while the momenta of two partons are mapped into the momenta of three partons, $(p_l, p_k) \leftrightarrow (\hat{p}_l, \hat{p}_{m+1}, \hat{p}_k)$. Here parton k is a spectator parton, chosen as one of the partons that is color connected to parton l. (There is an exception to this rule in the case of initial state splittings.) The antenna factorization of ref. [21] also uses a local mapping. We use a global mapping, in which all of the partons participate, as in ref. [22]. This way, each parton has to contribute only a little momentum. We also include quark masses in the kinematics.

4.1 Splitting a final state parton

We begin by defining what happens to the parton flavors when a final state parton l splits. The partons with indices other than l and $m + 1$ keep their flavors,

$$\hat{f}_j = f_j \ , \qquad j \notin \{l, m+1\} \ . \tag{4.1}$$

What happens to partons l and $m + 1$ is given by the value of the variable $\zeta_{\mathrm{f}} = (\hat{f}_l, \hat{f}_{m+1})$. The flavor splitting variable takes values in a set $\Phi_l(f_l)$ that depends on the flavor of the mother parton. If parton l is a quark or antiquark, then the set $\Phi_l(f_l)$ has only one element,

$$\Phi_l(f_l) = \{(f_l, \mathrm{g})\} \ , \qquad f_l \neq \mathrm{g} \ . \tag{4.2}$$

Here, we have used the freedom to assign labels in order to assign the label l to the daughter quark or antiquark and the label $m + 1$ to the gluon. If parton l is a gluon, then ζ_{f} can be a pair of gluons or any choice of $(q, \bar{q})$ flavors,

$$\Phi_l(\mathrm{g}) = \{(\mathrm{g}, \mathrm{g}), (\mathrm{u}, \bar{\mathrm{u}}), (\mathrm{d}, \bar{\mathrm{d}}), \dots\} \ . \tag{4.3}$$

In the case of a $\mathrm{g} \rightarrow q + \bar{q}$ splitting, we again use the freedom to assign labels in order to assign the label l to the daughter quark the label $m + 1$ to the daughter antiquark.

We now turn to the momenta. Parton l has momentum p_l with $p_l^2 = m^2(f_l)$ and splits into two partons l and $m+1$ with momenta $\hat{p}_l$ and $\hat{p}_{m+1}$ respectively. The daughter partons are on-shell: $\hat{p}_l^2 = m^2(\hat{f}_l)$ and $\hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1})$. We always have $(\hat{p}_l + \hat{p}_{m+1})^2 \geq m^2(f_l)$.

We need a bit of notation. Let Q be the total momentum of the final state partons,

$$Q \equiv \sum_{j=1}^{m} p_j = p_{\mathrm{a}} + p_{\mathrm{b}} \ . \tag{4.4}$$

Define

$$
\begin{aligned}
a_l &= \frac{Q^2}{2p_l \cdot Q} \ , \\
b_l &= \frac{m^2(f_l)}{2p_l \cdot Q} \ .
\end{aligned}
\tag{4.5}
$$

Note that $a_l + b_l \geq 1$. To see this, let $K = Q - p_l$ be the total momentum of the final state spectator partons, with $K^2 \geq 0$. Then

$$
\begin{aligned}
0 &\leq \frac{K^2}{2p_l \cdot Q} \\
&= \frac{Q^2 - 2p_l \cdot Q + m^2(f_l)}{2p_l \cdot Q} \\
&= a_l + b_l - 1 \ .
\end{aligned}
\tag{4.6}
$$

In order to define the momentum mapping, we first determine the total momentum

$$P_l = \hat{p}_l + \hat{p}_{m+1} \tag{4.7}$$

of the daughters of parton l. We take P_l to be a linear combination of p_l and Q,

$$P_l = \lambda p_l + \frac{1 - \lambda + y}{2a_l} \, Q \; . \tag{4.8}$$

There are two parameters in this definition. The first, y, is a measure of the virtuality, $P_l^2 - m^2(f_l)$, of the splitting. The second, λ, is a function of y that we will determine presently.

For an exactly collinear splitting or the emission of gluon with momentum $\hat{p}_{m+1} = 0$, we have $P_l = p_l$. Away from these limits, the spectator partons will have to donate some momentum in order to allow $P_l \neq p_l$. We elect to leave the momenta of the initial state partons unchanged, $\hat{p}_{\mathrm{a}} = p_{\mathrm{a}}$ and $\hat{p}_{\mathrm{b}} = p_{\mathrm{b}}$. Instead, we choose to obtain the needed momentum from the final state spectator partons by letting the momenta after the splitting be related to the momenta before the splitting by a Lorentz transformation,

$$\hat{p}_j^\mu = \Lambda^\mu_{\;\nu} \, p_j^\nu \; , \qquad j \notin \{l, m+1\} \; . \tag{4.9}$$

With this method of transferring momentum, each parton donates a share of the needed momentum, with low momentum partons donating only a little momentum.

The total momentum of the final state spectator partons before the splitting is

$$K = Q - p_l \; . \tag{4.10}$$

Since the momenta of the initial state partons remains the same, $\hat{Q} = \hat{p}_{\mathrm{a}} + \hat{p}_{\mathrm{b}}$ is the same as Q. The total momentum of the final state spectator partons after the splitting is then

$$\hat{K} = Q - P_l \; . \tag{4.11}$$

Since each final state spectator is changed by a Lorentz transformation, we have

$$\hat{K}^\mu = \Lambda^\mu_{\;\nu} \, K^\nu \; . \tag{4.12}$$

In fact, there is a Lorentz transformation that does this, namely

$$\Lambda(\hat{K}, K)^\mu_{\;\nu} = g^\mu_\nu - \frac{2(\hat{K} + K)^\mu (\hat{K} + K)_\nu}{(\hat{K} + K)^2} + \frac{2\hat{K}^\mu K_\nu}{K^2} \; , \tag{4.13}$$

provided that $\hat{K}^2 = K^2$. Thus P_l must lie on the hyperbola $(Q - P_l)^2 = (Q - p_l)^2$ in the Q-p_l plane, as illustrated in figure 3.

In the case in which the momenta K and $\hat{K}$ are carried by a single massless spectator, $\hat{K}$ is parallel to K. In this case, one can use an alternative representation of the boost in eq. (4.13) that remains well defined when $K^2 = (\hat{K} + K)^2 = 0$,

$$\Lambda(\hat{K}, K)^\mu_{\;\nu} = g^\mu_\nu + \left(\frac{K \cdot n}{\hat{K} \cdot n} - 1 \right) n^\mu \bar{n}_\nu + \left(\frac{\hat{K} \cdot n}{K \cdot n} - 1 \right) \bar{n}^\mu n_\nu \; , \tag{4.14}$$
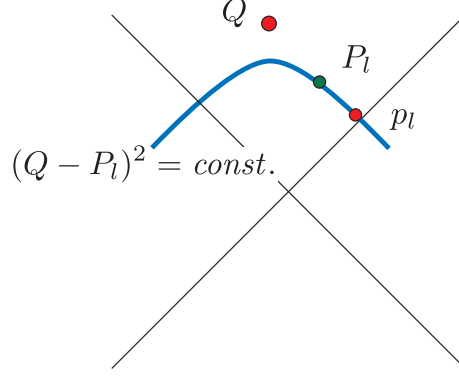
Figure 3: Momentum space mapping for a final state splitting. Coordinate axes in the two lightlike directions in the plane of p_l and Q are shown. Points p_l and Q are shown in the case that p_l is nearly lightlike. Then P_l lies on the hyperbola $(Q - P_l)^2 = const.$ that passes through p_l.

where n and $\bar{n}$ are lightlike vectors in the Q-p_l plane with $n \cdot \bar{n} = 1$ and $(p_l \cdot n/p_l \cdot \bar{n}) < (Q \cdot n/Q \cdot \bar{n})$. That is, these are the vectors along the two coordinate axes in figure 3, with n directed toward the upper right and $\bar{n}$ directed toward the upper left in the diagram.

We can arrange that $\hat{K}^2 = K^2$ by making a proper choice of λ in eq. (4.8). We have

$$K^2 = (a_l + b_l - 1)\, 2Q{\cdot}p_l \ , \tag{4.15}$$

and

$$\hat{K}^2 = \frac{1}{4a_l}\left[(2a_l - 1 - y)^2 - (1 - 4a_l b_l)\lambda^2\right] 2Q{\cdot}p_l \ . \tag{4.16}$$

We obtain $\hat{K}^2 = K^2$ if

$$\lambda = \sqrt{\frac{(1 + y)^2 - 4a_l(y + b_l)}{1 - 4a_l b_l}} \ . \tag{4.17}$$

We note that $\lambda \to 1$ when $y \to 0$. As y increases, λ decreases. There is a maximum possible value of y, namely the value that makes $\lambda = 0$,

$$y_{\max} = \left(\sqrt{a_l} - \sqrt{a_l + b_l - 1}\right)^2 - b_l \ . \tag{4.18}$$

The splitting parameter y has a simple meaning. From eq. (4.8), we find, using eq. (4.17),

$$y = \frac{P_l^2 - m^2(f_l)}{2p_l{\cdot}Q} \ . \tag{4.19}$$

That is, y is a dimensionless measure of the virtuality of the splitting. An alternative relation is

$$y = \frac{P_l^2 - m^2(f_l)}{2P_l{\cdot}Q - (P_l^2 - m^2(f_l))} \ . \tag{4.20}$$

There are also alternative relations for a_l and b_l,

$$
\begin{aligned}
a_l &= \frac{Q^2}{2P_l{\cdot}Q + m^2(f_l) - P_l^2} \ , \\
b_l &= \frac{m^2(f_l)}{2P_l{\cdot}Q + m^2(f_l) - P_l^2} \ .
\end{aligned}
\tag{4.21}
$$

These relations can be derived with the use of the relation

$$2P_l \cdot Q = (1 + y)\, 2p_l \cdot Q \ . \tag{4.22}$$

There is also a minimum value of y. Since $P_l^2 > [m(\hat{f}_l) + m(\hat{f}_{m+1})]^2$, we have $y > y_{\min}$ where

$$y_{\min} = \frac{[m(\hat{f}_l) + m(\hat{f}_{m+1})]^2 - m^2(f_l)}{2p_l \cdot Q} \ . \tag{4.23}$$

Note that if all of the partons are massless, then $y_{\min} = 0$. In addition, if $f_l = \hat{f}_l$ is a massive quark flavor and $\hat{f}_{m+1} = $ g, then also $y_{\min} = 0$.

We have seen how, given the virtuality variable y, we can define $P_l = \hat{p}_l + \hat{p}_{m+1}$ in the plane of p_l and Q such that $P_l^2 - m^2(f_l) = y\, 2p_l \cdot Q$ and so that the needed momentum donation from the spectator partons can be obtained by a Lorentz transformation. It remains to define $\hat{p}_l$ and $\hat{p}_{m+1}$ individually. This is simple. We let $\hat{p}_l$ and $\hat{p}_{m+1}$ be any momenta on the appropriate mass shells that sum to P_l.

It will prove convenient to formulate this rather abstractly. We denote the daughter parton momenta by

$$\zeta_{\mathrm{p}} \equiv (\hat{p}_l, \hat{p}_{m+1}) \ . \tag{4.24}$$

We must have

$$\zeta_{\mathrm{p}} \in \Gamma_l(\{p\}_m, \zeta_{\mathrm{f}}) \ , \tag{4.25}$$

where the set $\Gamma_l(\{p\}_m, \zeta_{\mathrm{f}})$ is described as follows,

$$
\begin{aligned}
\Gamma_l(\{p\}_m, \zeta_{\mathrm{f}}) = \Big\{ (\hat{p}_l, \hat{p}_{m+1}) \ \Big| \ & \hat{p}_l^2 = m^2(\hat{f}_l), \quad \hat{p}_l \cdot Q > 0, \\
& \hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1}), \quad \hat{p}_{m+1} \cdot Q > 0, \\
& \varepsilon_{\mu\nu\alpha\beta}(\hat{p}_l^\nu + \hat{p}_{m+1}^\nu)\, p_l^\alpha\, Q^\beta = 0, \\
& (Q - \hat{p}_l - \hat{p}_{m+1})^2 = (Q - p_l)^2 \Big\} \ .
\end{aligned}
\tag{4.26}
$$

That is, each of $\hat{p}_l$ and $\hat{p}_{m+1}$ lies on the appropriate forward mass shell, their sum lies in the plane of p_l and Q, and $\hat{K}^2 = K^2$. These conditions entail that

$$[m(\hat{f}_l) + m(\hat{f}_{m+1})]^2 \le (\hat{p}_l + \hat{p}_{m+1})^2 \le \left(\sqrt{Q^2} - \sqrt{Q^2 + m^2(f_l) - 2p_l \cdot Q} \right)^2 \ , \tag{4.27}$$

which corresponds to $y_{\min} \le y \le y_{\max}$.

We note that the set Γ_l is a three dimensional surface in the space of momenta $(\hat{p}_l, \hat{p}_{m+1})$. One can choose three coordinates to describe this surface, for instance a virtuality variable, a momentum fraction variable, and an azimuthal angle. Different choices of coordinates may be best for different purposes, so we leave this choice open.

We give a name to this transformation of momenta and flavors:

$$\{\hat{p}, \hat{f}\}_{m+1} = R_l(\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}) \ . \tag{4.28}$$

In parton splittings close to the collinear or soft limits, partons lose energy when they split. It is of interest to see how this property carries over to splittings that are not close

to the limit. Here, the choice of the mapping R_l may be considered to be part of a model for shower evolution that could be sensible or perhaps not so sensible. An investigation of properties of the model is thus of some significance.

Using $K^2 = \hat{K}^2$, we derive

$$Q{\cdot}p_l - Q{\cdot}\hat{p}_l = \hat{p}_{m+1}{\cdot}\hat{K} + \frac{1}{2}\left[m^2(f_l) - m^2(\hat{f}_l) + m^2(\hat{f}_{m+1})\right] \ . \tag{4.29}$$

This is especially interesting in the case that parton l is a quark and parton $m + 1$ is a gluon. Then the term involving masses vanishes. We note that $\hat{p}_{m+1} \cdot \hat{K} \geq 0$ because both $\hat{p}_{m+1}$ and $\hat{K}$ lie inside or on the forward light cone. Thus the energy of the quark, as measured in the frame in which Q is at rest, is bigger before the splitting than after the splitting. By emitting bremsstrahlung, the quark slows down. If evolution of the final state were to continue long enough, the quark would slow to a stop in the $\vec{Q} = 0$ frame. Then radiation from that quark would cease. At any point in the shower evolution, there can be an initial state splitting, discussed below in section 4.4. This changes p_a or p_b and thus $p_a + p_b = Q$. Now final state quarks tend to come to rest in the new $\vec{Q} = 0$ frame. We judge that a tendency for quarks to slow down (and, similarly, for gluons to lose energy) is reasonably sensible. Of course, the partons should not be allowed to shower indefinitely. At some resolution scale, a perturbative model for showering is simply wrong and a process by which partons combine to form hadrons is needed.

4.2 Combining two final state partons

It is significant (and useful) that this transformation has an inverse. Let start with $\{\hat{p}\}_{m+1}$ and determine $\{p\}_m$ and $\{\zeta_p, \zeta_f\}$.

The splitting variable for the momenta is given by the momenta of the daughter partons, $\zeta_p = (\hat{p}_l, \hat{p}_{m+1})$. From $\{\hat{p}\}_{m+1}$ we determine

$$Q = \sum_{j=1}^{m+1} \hat{p}_j = \hat{p}_a + \hat{p}_b \ . \tag{4.30}$$

Then eq. (4.20) gives y, eq. (4.21) gives a_l and b_l, and eq. (4.17) gives λ. Since the calculation of λ involves taking the square root of λ^2, we should check that $\lambda^2 > 0$. For this purpose, we can express λ^2 in terms of dot products of vectors as

$$\lambda^2 = \frac{4[(Q{\cdot}P_l)^2 - Q^2 P_l^2] + m^2(f_l)\left[4P_l{\cdot}\hat{K} + m^2\right]}{4Q{\cdot}\hat{K}\left[P_l{\cdot}\hat{K} + P_l^2 - m^2(f_l)\right] + (P_l^2 - m^2(f_l))^2} \ . \tag{4.31}$$

Since Q, P_l, and $\hat{K} = Q - P_l$ lie inside or on the positive lightcone and $P_l^2 - m^2(f_l) > 0$, both the numerator and the denominator are non-negative. With y, λ, a_l and b_l at hand, one can calculate the lightlike momentum p_l by rearranging eq. (4.8),

$$p_l = \frac{1}{\lambda}(\hat{p}_l + \hat{p}_{m+1}) - \frac{1 - \lambda + y}{2\lambda a_l} Q \ . \tag{4.32}$$

We now have p_l. We define $p_a = \hat{p}_a$ and $p_b = \hat{p}_b$. This leaves the p_j for $j \notin \{l, a, b\}$. For this, we need the inverse Lorentz transformation to eq. (4.9). From $K = Q - p_l$ and

$\hat{K} = Q - \hat{p}_l - \hat{p}_{m+1}$, we construct $\Lambda(K, \hat{K})^\mu_{\ \nu}$ using eq. (4.13) or eq. (4.14) with the roles of $\hat{K}$ and K interchanged. Then

$$p_j^\mu = \Lambda(K, \hat{K})^\mu_{\ \nu}\ \hat{p}_j^\nu\ , \qquad j \notin \{l, \mathrm{a}, \mathrm{b}\}\ . \tag{4.33}$$

The transformation of the flavors is simple. The splitting variable ζ_{f} is given by the flavors of the daughter particles, $(\hat{f}_l, \hat{f}_{m+1})$. The flavor of the mother parton is

$$f_l = \hat{f}_l + \hat{f}_{m+1}\ , \tag{4.34}$$

with the obvious definition of adding flavors, as in $\mathrm{d} + \mathrm{g} = \mathrm{d}$ and $\mathrm{u} + \bar{\mathrm{u}} = \mathrm{g}$. The flavors of the other partons are unchanged

$$f_j = \hat{f}_j\ , \qquad j \notin \{l, m+1\}\ . \tag{4.35}$$

We give a name to this transformation of momenta and flavors,

$$\{\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}\} = Q_l(\{\hat{p}, \hat{f}\}_{m+1})\ . \tag{4.36}$$

This is the inverse transformation to R_l, eq. (4.28).

4.3 The integration measure for final state splitting

With a suitable choice of the integration measure $d\zeta_{\mathrm{p}}$ for integrating over the splitting variables ζ_{p}, we can arrange that

$$\int [d\{\hat{p}, \hat{f}\}_{m+1}] g(\{\hat{p}, \hat{f}\}_{m+1}) = \int [d\{p, f\}_m] \sum_{\zeta_{\mathrm{f}} \in \Phi_l(f_l)} \int d\zeta_{\mathrm{p}} \theta(\zeta_{\mathrm{p}} \in \Gamma_l(\{p\}_m, \zeta_{\mathrm{f}})) g(\{\hat{p}, \hat{f}\}_{m+1}) \tag{4.37}$$

for an arbitrary function $g(\{\hat{p}, \hat{f}\}_{m+1})$. The definition that we need is

$$
\begin{aligned}
d\zeta_{\mathrm{p}} \equiv\ & dy\ \theta(y_{\min} < y < y_{\max})\ \lambda \frac{p_l \cdot Q}{\pi} \\
& \times \frac{d^4 \hat{p}_l}{(2\pi)^4}\, 2\pi \delta_+(\hat{p}_l^2 - m^2(\hat{f}_l))\, \frac{d^4 \hat{p}_{m+1}}{(2\pi)^4}\, 2\pi \delta_+(\hat{p}_{m+1}^2 - m^2(\hat{f}_{m+1})) \\
& \times (2\pi)^4\, \delta\!\left(\hat{p}_l + \hat{p}_{m+1} - \lambda p_l - \frac{1 - \lambda + y}{2a_l}\, Q\right)\ .
\end{aligned}
\tag{4.38}
$$

Here the limits on y are given in eqs. (4.18) and (4.23).

4.4 Splitting an initial state parton

Consider the splitting of an initial state parton, say parton "a." The initial state parton with momentum $p_{\mathrm{a}} \approx \eta_{\mathrm{a}} p_{\mathrm{A}}$ splits to produce a new initial state parton with momentum $\hat{p}_{\mathrm{a}} \approx \hat{\eta}_{\mathrm{a}} p_{\mathrm{A}}$ and a new final state parton with label $m+1$ and momentum $\hat{p}_{m+1}$. We are using the usual backwards evolution here, so that the evolution going forward in time is $\hat{p}_{\mathrm{a}} \to p_{\mathrm{a}} + \hat{p}_{m+1}$. In this subsection, we describe how $\{\hat{p}, \hat{f}\}_{m+1}$ is determined from $\{p, f\}_m$ and splitting variables $\{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}$. The splitting of the other initial state parton is described by the same formulas with $\mathrm{a} \leftrightarrow \mathrm{b}$.

We begin by defining what happens to the parton flavors, recalling our notation that for the initial state partons "a" and "b", f_a, $\hat{f}_\mathrm{a}$, f_b and $\hat{f}_\mathrm{b}$ denote the opposite of the flavor of the physical incoming parton. The partons with indices other than "a" and $m+1$ keep their flavors,

$$\hat{f}_j = f_j \ , \qquad j \notin \{\mathrm{a}, m+1\} \ . \tag{4.39}$$

What happens to partons "a" and $m+1$ is given by the value of the variable $\zeta_\mathrm{f} = (\hat{f}_\mathrm{a}, \hat{f}_{m+1})$. The flavor splitting variable takes values in a set $\Phi_\mathrm{a}(f_\mathrm{a})$ that depends on the flavor of the mother parton. This set is determined by the requirement that $f_\mathrm{a} = \hat{f}_\mathrm{a} + \hat{f}_{m+1}$. If parton "a" is a quark or antiquark, then the set $\Phi_\mathrm{a}(f_\mathrm{a})$ has only two elements,

$$\Phi_\mathrm{a}(f_\mathrm{a}) = \{(f_\mathrm{a}, \mathrm{g}), (\mathrm{g}, f_\mathrm{a})\} \ , \qquad f_\mathrm{a} \neq \mathrm{g} \ . \tag{4.40}$$

If parton "a" is a gluon, then ζ_f can be a pair of gluons or any choice of $(q, \bar{q})$ or $(q, \bar{q})$ flavors,

$$\Phi_\mathrm{a}(\mathrm{g}) = \{(\mathrm{g}, \mathrm{g}), (\mathrm{u}, \bar{\mathrm{u}}), (\bar{\mathrm{u}}, \mathrm{u}), (\mathrm{d}, \bar{\mathrm{d}}), (\bar{\mathrm{d}}, \mathrm{d}), \dots\} \ . \tag{4.41}$$

We now turn to the momenta. Let parton "a" with momentum fraction η_a radiate a parton $m+1$ with momentum $\hat{p}_{m+1}$. We then need to define how to determine $\{\hat{p}\}_{m+1}$ from $\{p\}_m$ and $\hat{p}_{m+1}$.

As discussed in section 3.1, we can include masses for the initial state partons.[17] We take the partons to be on-shell with zero transverse momenta,

$$
\begin{aligned}
p_\mathrm{a} &= \eta_\mathrm{a} p_\mathrm{A} + \frac{m^2(f_\mathrm{a})}{\eta_\mathrm{a} s} p_\mathrm{B} \ , \\
p_\mathrm{b} &= \eta_\mathrm{b} p_\mathrm{B} + \frac{m^2(f_\mathrm{b})}{\eta_\mathrm{b} s} p_\mathrm{A} \ , \\
\hat{p}_\mathrm{a} &= \hat{\eta}_\mathrm{a} p_\mathrm{A} + \frac{m^2(\hat{f}_\mathrm{a})}{\hat{\eta}_\mathrm{a} s} p_\mathrm{B} \ .
\end{aligned}
\tag{4.42}
$$

Recall that we define p_A and p_B to be lightlike approximations to the incoming hadron momenta, with $2 p_\mathrm{A} \cdot p_\mathrm{B} = s$. The radiated parton can have a mass,

$$\hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1}) \ . \tag{4.43}$$

We take the momentum fraction of parton "b" to remain the same,

$$\hat{\eta}_\mathrm{b} = \eta_\mathrm{b} \ . \tag{4.44}$$

The momentum fraction $\hat{\eta}_\mathrm{a}$ after the splitting will be determined by $\hat{p}_{m+1}$. As in the case of a final state splitting, it is not generally possible to have $\hat{p}_\mathrm{a} = p_\mathrm{a} + \hat{p}_{m+1}$ given the mass shell conditions and the possibility that the radiated parton has non-zero transverse momentum. In order to allow the approximation that both p_a and $\hat{p}_\mathrm{a}$ are on-shell with zero transverse momenta, we therefore take some momenta from the final state spectator

---

[17]Recall that this is optional. One could just replace the masses for the initial state partons by zero, $m^2(f_\mathrm{a}) = m^2(\hat{f}_\mathrm{a}) = m^2(f_\mathrm{b}) = 0$. One could also set all quark masses to zero.

partons by letting the momenta after the splitting be related to the momenta before the splitting by a Lorentz transformation,

$$\hat{p}_j^\mu = \Lambda^\mu_{\ \nu}\, p_j^\nu \ , \qquad j \in \{1, \ldots, m\} \ . \tag{4.45}$$

Since each final state spectator is changed by a Lorentz transformation, we have

$$\hat{K}^\mu = \Lambda^\mu_{\ \nu}\, K^\nu \ , \tag{4.46}$$

where K is the momentum of the final state partons before the splitting,

$$K = p_{\rm a} + p_{\rm b} \ , \tag{4.47}$$

and $\hat{K}$ is the momentum of the final state spectators after the splitting,

$$\hat{K} = \hat{p}_{\rm a} + p_{\rm b} - \hat{p}_{m+1} \ . \tag{4.48}$$

In order for K and $\hat{K}$ to be related by a Lorentz transformation, we need $\hat{K}^2 = K^2$. To see what this means, define

$$\hat{Q}(\hat{\eta}_{\rm a}) = \hat{p}_{\rm a} + p_{\rm b} = \left(\hat{\eta}_{\rm a} + \frac{m^2(\hat{f}_{\rm b})}{\eta_{\rm b} s} \right) p_{\rm A} + \left(\eta_{\rm b} + \frac{m^2(\hat{f}_{\rm a})}{\hat{\eta}_{\rm a} s} \right) p_{\rm B} \ . \tag{4.49}$$

Then we demand that

$$(\hat{Q}(\hat{\eta}_{\rm a}) - \hat{p}_{m+1})^2 = K^2 \ . \tag{4.50}$$

The vector $\hat{p}_{m+1}$ determines $\hat{\eta}_{\rm a}$. In the space of $\hat{p}_{m+1}$, a surface of constant $\hat{\eta}_{\rm a}$ is the intersection of the hyperbola $\hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1})$ with the hyperbola given by eq. (4.50), as illustrated in figure 4. Using $\hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1})$ in eq. (4.50), we can write

$$2\, \hat{Q}(\hat{\eta}_{\rm a}) \cdot \hat{p}_{m+1} = \hat{Q}(\hat{\eta}_{\rm a})^2 + m^2(\hat{f}_{m+1}) - K^2 \ . \tag{4.51}$$

Looked at this way, a surface of constant $\hat{\eta}_{\rm a}$ is the intersection of the hyperbola $\hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1})$ with the plane defined by eq. (4.51). We will require $\hat{\eta}_{\rm a} < 1$. Thus the allowed region in $\hat{p}_{m+1}^2$ is the part of the forward mass shell with

$$2\, \hat{Q}(1) \cdot \hat{p}_{m+1} < \hat{Q}(1)^2 + m^2(\hat{f}_{m+1}) - K^2 \ . \tag{4.52}$$

In order to solve for $\hat{\eta}_{\rm a}$ given $\hat{p}_{m+1}$, we write K^2 and $\hat{K}^2$ in the form

$$\begin{aligned} K^2 &= \alpha \eta_{\rm a} - \frac{\beta}{\eta_{\rm a}} - \gamma \ , \\ \hat{K}^2 &= \hat{\alpha} \hat{\eta}_{\rm a} - \frac{\hat{\beta}}{\hat{\eta}_{\rm a}} - \hat{\gamma} \ . \end{aligned} \tag{4.53}$$

Here

$$\begin{aligned} \alpha &= \eta_{\rm b} s \ , \\ \beta &= -\frac{m^2(f_{\rm a})\, m^2(f_{\rm b})}{\eta_{\rm b} s} \ , \\ \gamma &= -m^2(f_{\rm a}) - m^2(f_{\rm b}) \ , \end{aligned} \tag{4.54}$$
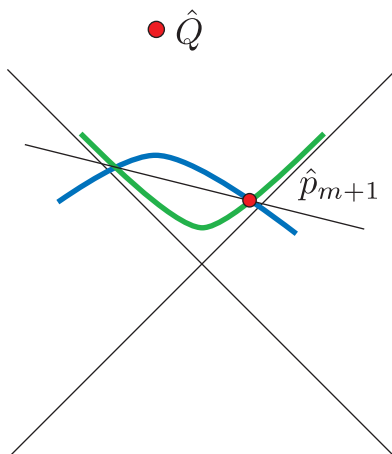
Figure 4: Momentum space for $\hat{p}_{m+1}$ in an initial state splitting. Coordinate axes along p_A and p_B are shown, along with the point $\hat{Q}(\hat{\eta}_a)$. We show the mass shell for $\hat{p}_{m+1}$ and the hyperbola $(\hat{Q}(\hat{\eta}_a) - \hat{p}_{m+1})^2 = K^2$. Points $\hat{p}_{m+1}$ for a given value of $\hat{\eta}_a$ are on the intersection of these two hyperbolas. We illustrate a point $\hat{p}_{m+1}$ with zero momentum transverse to p_A and p_B, but in general $\hat{p}_{m+1}$ can have transverse components. The plane specified in eq. (4.51) is also depicted.

while

$$
\begin{aligned}
\hat{\alpha} &= \eta_b s - 2 p_A \cdot \hat{p}_{m+1} \;, \\
\hat{\beta} &= \frac{m^2(\hat{f}_a)}{s} \left\{ 2 p_B \cdot \hat{p}_{m+1} - \frac{m^2(f_b)}{\eta_b} \right\} \;, \\
\hat{\gamma} &= 2 p_b \cdot \hat{p}_{m+1} - m^2(\hat{f}_a) - m^2(f_b) - m^2(\hat{f}_{m+1}) \;.
\end{aligned}
\tag{4.55}
$$

The condition $\hat{K}^2 = K^2$ now determines $\hat{\eta}_a$,

$$
\hat{\eta}_a = \frac{1}{2\hat{\alpha}} \left\{ K^2 + \hat{\gamma} + \sqrt{(K^2 + \hat{\gamma})^2 + 4\hat{\alpha}\hat{\beta}} \right\} \;.
\tag{4.56}
$$

It is a consequence of eq. (4.56) and the kinematic conditions (3.6) and (3.9) that

$$
\hat{\eta}_a > \eta_a \;.
\tag{4.57}
$$

We prove this in appendix A.

Having fixed $\hat{\eta}_a$ so that $\hat{K}^2 = K^2$, these two momenta will be related by eq. (4.46),

$$
\hat{K}^\mu = \Lambda(\hat{K}, K)^\mu{}_\nu \, K^\nu \;,
\tag{4.58}
$$

where $\Lambda^\mu{}_\nu$ the Lorentz transformation (4.13). This allows us to define the spectator momenta after the splitting to be related to the spectator momenta before the splitting by this same boost, as in eq. (4.45),

$$
\hat{p}_j^\mu = \Lambda(\hat{K}, K)^\mu{}_\nu \, p_j^\nu \;, \qquad j \in \{1, \ldots, m\} \;.
\tag{4.59}
$$

For a final state splitting, we defined a splitting variable ζ_p in a three dimensional set Γ_l so that $\{p\}_m$ together with ζ_p determines $\{\hat{p}\}_{m+1}$. Here we use an analogous notation, with

$$
\zeta_p = (\hat{p}_a, \hat{p}_{m+1}) \;.
\tag{4.60}
$$

– 33 –

We choose

$$\zeta_{\mathrm{p}} \in \Gamma_{\mathrm{a}}(\{p\}_m, \zeta_{\mathrm{f}}) \ , \tag{4.61}$$

where

$$\Gamma_{\mathrm{a}}(\{p\}_m, \zeta_{\mathrm{f}}) = \left\{ (\hat{p}_{\mathrm{a}}, \hat{p}_{m+1}) \ \middle| \ \hat{p}_{m+1}^2 = m^2(\hat{f}_{m+1}), \quad \hat{p}_{m+1} \cdot (p_{\mathrm{a}} + p_{\mathrm{b}}) > 0, \right.$$
$$\hat{p}_{\mathrm{a}} = \hat{\eta}_{\mathrm{a}} p_{\mathrm{A}} + \frac{m^2(\hat{f}_{\mathrm{a}})}{\hat{\eta}_{\mathrm{a}} s} p_{\mathrm{B}},$$
$$\hat{K}^2(\hat{\eta}_{\mathrm{a}}) = K^2(\eta_{\mathrm{a}}), \quad \hat{K} \cdot K > 0,$$
$$\left. \eta_{\mathrm{a}} < \hat{\eta}_{\mathrm{a}} < 1 \right\} \ . \tag{4.62}$$

That is, the radiated parton lies on the forward mass-shell and the new initial state parton is also on its mass shell with has zero transverse momentum. Its momentum fraction $\hat{\eta}_{\mathrm{a}}$ is determined by $\hat{K}^2 = K^2$, with K inside the forward lightcone. The solution of $\hat{K}^2 = K^2$ is given by eq. (4.56). We also need $\hat{\eta}_{\mathrm{a}}$ to be big enough that that the kinematic bound that was given in eq. (3.9) is obeyed. This follows from $\hat{\eta}_{\mathrm{a}} > \eta_{\mathrm{a}}$. Finally, $\hat{\eta}_{\mathrm{a}}$ must be less than 1.

As for a final state splitting, the set Γ_{a} is a three dimensional surface in the space of momenta $(\hat{p}_{\mathrm{a}}, \hat{p}_{m+1})$. One can describe this surface using three coordinates such as a virtuality variable, a momentum fraction variable, and an azimuthal angle. We leave the choice of coordinates open.

We give a name to this transformation of momenta and flavors, R_l with $l = \mathrm{a}$:

$$\{\hat{p}, \hat{f}\}_{m+1} = R_{\mathrm{a}}(\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}) \ . \tag{4.63}$$

4.5 Combining an initial state parton with a final state parton

The transformation for splitting an initial state parton has an inverse. Let us start with $\{\hat{p}, \hat{f}\}_{m+1}$ and determine $\{p, f\}_m$ and $\{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}$.

The momentum splitting variable is simply $\zeta_{\mathrm{p}} = (\hat{p}_{\mathrm{a}}, \hat{p}_{m+1})$. For the momentum fraction of parton "b" before the splitting, we take

$$\eta_{\mathrm{b}} = \hat{\eta}_{\mathrm{b}} \ . \tag{4.64}$$

To determine η_{a}, we simply use the representation (4.53) of K^2 and $\hat{K}^2$ and solve $K^2 = \hat{K}^2$ for η_{a} in terms of $\hat{\eta}_{\mathrm{a}}$. This gives

$$\eta_{\mathrm{a}} = \frac{1}{2\alpha} \left\{ \hat{K}^2 + \gamma + \sqrt{(\hat{K}^2 + \gamma)^2 + 4\alpha\beta} \right\} \ . \tag{4.65}$$

Once we have η_{a}, we can construct $K = p_{\mathrm{a}} + p_{\mathrm{b}}$. Then from K and $\hat{K}$ we can construct $\Lambda(K, \hat{K})$, the inverse Lorentz transformation matrix to $\Lambda(\hat{K}, K)$ defined in eq. (4.13) by simply using the same equation with $K \leftrightarrow \hat{K}$. Then we can construct the remaining momenta using

$$p_j^\mu = \Lambda(K, \hat{K})^\mu{}_\nu \, \hat{p}_j^\nu \ . \tag{4.66}$$

The transformation of the flavors is simple. The splitting variable ζ_{f} is given by the flavors of the (backwards evolution) daughter particles, $(\hat{f}_{\mathrm{a}}, \hat{f}_{m+1})$. The flavor of the mother parton is

$$f_{\mathrm{a}} = \hat{f}_{\mathrm{a}} + \hat{f}_{m+1} \ . \tag{4.67}$$

The flavors of the other partons are unchanged

$$f_j = \hat{f}_j \ , \qquad j \notin \{\mathrm{a}, m+1\} \ . \tag{4.68}$$

We give a name to this transformation of momenta and flavors,

$$\{\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}\} = Q_{\mathrm{a}}(\{\hat{p}, \hat{f}\}_{m+1}) \ . \tag{4.69}$$

This is the inverse transformation to R_{a}, eq. (4.63).

4.6 The integration measure for initial state splitting

With a suitable choice of the integration measure $d\zeta_{\mathrm{p}}$ for integrating over the splitting variables ζ_{p}, we can arrange that

$$\begin{aligned}
\int [d\{\hat{p}, \hat{f}\}_{m+1}] \ &g(\{\hat{p}, \hat{f}\}_{m+1}) \\
&= \int [d\{p, f\}_m] \sum_{\zeta_{\mathrm{f}} \in \Phi_{\mathrm{a}}(f_{\mathrm{a}})} \int d\zeta_{\mathrm{p}} \ \theta(\zeta_{\mathrm{p}} \in \Gamma_{\mathrm{a}}(\{p\}_m, \zeta_{\mathrm{f}})) \ g(\{\hat{p}, \hat{f}\}_{m+1})
\end{aligned} \tag{4.70}$$

for an arbitrary function $g(\{\hat{p}, \hat{f}\}_{m+1})$ with support in $\hat{K}^2 > 4m_{\mathrm{H}}^2$. The definition that we need is

$$d\zeta_{\mathrm{p}} \equiv \frac{d^4 \hat{p}_{m+1}}{(2\pi)^4} \ 2\pi\delta_+(\hat{p}_{m+1}^2 - m^2(\hat{f}_{m+1})) \ \frac{\alpha + \beta/\eta_{\mathrm{a}}^2}{\hat{\alpha} + \hat{\beta}/\hat{\eta}_{\mathrm{a}}^2} \ . \tag{4.71}$$

The factor $(\alpha + \beta/\eta_{\mathrm{a}}^2)/(\hat{\alpha} + \hat{\beta}/\hat{\eta}_{\mathrm{a}}^2)$ here is just $d\hat{\eta}_{\mathrm{a}}/d\eta_{\mathrm{a}}$ calculated from the relation $\hat{K}^2 = K^2$.

5. Spin states

The quantum scattering amplitude $\big|M(\{p, f\}_m)\big\rangle$ is a vector in spin $\otimes$ color space. Thus we can expand it in terms of spin and color basis vectors,

$$\big|M(\{p, f\}_m)\big\rangle = \sum_{\{c\}_m} \big|\{c\}_m\big\rangle \sum_{\{s\}_m} \big|\{s\}_m\big\rangle \ M(\{p, f, s, c\}_m) \ . \tag{5.1}$$

The treatment of color that appears to us to be most useful for parton showers is a bit subtle and, in particular, involves basis vectors that are not exactly conventionally normalized and are not exactly orthogonal to one another. The color basis is described in section 7. In contrast, our spin basis vectors are quite standard, are orthogonal and normalized,

$$\big\langle\{s'\}'_m\big|\{s\}_m\big\rangle = \delta_{m',m} \, \delta_{\{s'\}_m, \{s\}_m} \ . \tag{5.2}$$

The spin labels $\{s_{\mathrm{a}}, s_{\mathrm{b}}, s_1, \ldots, s_m\}$ represent the helicities of the corresponding particles.[18]

The basis states for parton l are represented for quarks or antiquarks by Dirac spinors $U(p_l, s_l)$ or $V(p_l, s_l)$, where $p_l^2 = m^2(f_l)$. For gluons, we need polarization vectors $\varepsilon(p_l, s_l)$ with $p_l^2 = 0$. Our definition makes use an auxiliary vector n_l, chosen along the intersection of the positive lightcone with the plane of p_l and Q, where $Q = \sum_{j=1}^{m} p_j = p_{\mathrm{a}} + p_{\mathrm{b}}$ is the total momentum of the final state particles. The normalization of n_l is not important. We take the solution that is not close to p_l in the case that $m^2(f_l)$ is small. A convenient set of choices is

$$
n_l = \begin{cases} p_{\mathrm{B}} \;, & l = \mathrm{a} \;, \\ p_{\mathrm{A}} \;, & l = \mathrm{b} \;, \\ Q - \dfrac{Q^2}{Q{\cdot}p_l + \sqrt{(Q{\cdot}p_l)^2 - Q^2\, m^2(f_l)}}\, p_l \;, & l \in \{1, \ldots, m\} \;. \end{cases}
\tag{5.3}
$$

For quarks and antiquarks, we use Dirac spinors $U(p, s)$ or $V(p, s)$ with $p^2 = m^2$ and $s = \pm 1/2$. We can take $V(p, s) = [\overline{U}(p, s)C]^T$, where C is the charge conjugation matrix, $i\gamma^2\gamma^0$ in the chiral representation of the gamma matrices, with $C^{-1}\gamma^\mu C = -(\gamma^\mu)^T$ and $C^{-1} = C^\dagger = C^T = -C$. The spinors obey $(\slashed{p} - m)U(p, s) = 0$ and $(\slashed{p} + m)V(p, s) = 0$. They are normalized to

$$
\begin{aligned}
\overline{U}(p, s)\gamma^\mu U(p, s) &= 2p^\mu \;, \\
\overline{V}(p, s)\gamma^\mu V(p, s) &= 2p^\mu \;.
\end{aligned}
\tag{5.4}
$$

We use helicity eigenstates, defined so that

$$
\begin{aligned}
\gamma_5 \slashed{s} U(p, \pm 1/2) &= \pm U(p, \pm 1/2) \;, \\
\gamma_5 \slashed{s} V(p, \pm 1/2) &= \pm V(p, \pm 1/2) \;.
\end{aligned}
\tag{5.5}
$$

Here the spin vector s is

$$
s = \frac{1}{m} p - \frac{m}{p{\cdot}n} n \;,
\tag{5.6}
$$

where n is the auxiliary lightlike vector from eq. (5.3). Thus $s^2 = -1$ and $s \cdot p = 0$.

A convenient definition that defines the phase of $U(p, -s)$ in terms of the phase of $U(p, s)$ is

$$
U(p, -s) = \left(1 + \frac{m}{p{\cdot}n}\, \slashed{n}\right) V(p, s).
\tag{5.7}
$$

For gluons, we need polarization vectors $\varepsilon^\mu(p, s; Q)$, representing a given helicity s and defined with the aid of an auxiliary vector Q. The polarization vectors obey $p{\cdot}\varepsilon(p, s; Q) = 0$ and

$$
Q{\cdot}\varepsilon(p, s; Q) = 0 \;.
\tag{5.8}
$$

We can also write

$$
\varepsilon(p, s; Q) = \varepsilon(p, s; n) \;,
\tag{5.9}
$$

---

[18]Many authors follow the convention that the helicity label for an incoming particle is the negative of the particle's helicity. In contrast, our convention is that s_{a} and s_{b}, as well as $s_1, \ldots, s_m$, represent the physical helicities of the particles.

where $n \cdot \varepsilon(p, s; n) = 0$ and n is a lightlike vector defined in eq. (5.3). We define the phase by using the standard definition in terms of mass zero Dirac spinors [23],

$$\varepsilon^\mu(p, \pm 1; n) = \pm \frac{\overline{U}(n, \pm 1/2) \gamma^\mu U(p, \pm 1/2)}{\sqrt{2}\, \overline{U}(p, \mp 1/2) U(n, \pm 1/2)} \quad . \tag{5.10}$$

With this definition, there is a simple relation between polarization vectors defined with different auxiliary vectors Q and thus different lightlike auxiliary vectors n [23],

$$\varepsilon^\mu(p, \pm 1; n) - \varepsilon^\mu(p, \pm 1; n') = \mp \frac{\sqrt{2}\, \overline{U}(n, \pm 1/2) U(n', \mp 1/2)}{\overline{U}(n, \pm 1/2) U(p, \mp 1/2)\, \overline{U}(p, \pm 1/2) U(n', \mp 1/2)}\, p^\mu \quad . \tag{5.11}$$

If we use these polarization vectors with the exact tree-level Feynman diagrams, gauge invariance of the matrix elements, together with the fact that $\varepsilon^\mu(p, s; n)$ differs from $\varepsilon^\mu(p, s; n')$ by a vector proportional to p^μ, shows that the amplitude is independent of the choice of the auxiliary vector Q. Our matrix elements will be approximate and will be gauge invariant only to the extent that the splittings are close to the soft or collinear limits. Thus some dependence on the auxiliary vector used to define the polarization vectors will result.

6. Splitting functions for the quantum states

Consider for a moment a theory without spin, color, or flavors, say ϕ^3 theory in six dimensions as in section 2. An $(m + 1)$-parton scattering amplitude $M(\{\hat{p}\}_{m+1})$ is simple in the limit in which two of the partons are approximately collinear. Supposing that partons $m + 1$ and l are almost collinear, we have

$$M(\{\hat{p}\}_{m+1}) \approx v(\{\hat{p}\}_{m+1})\, M(\{p\}_m) \quad , \tag{6.1}$$

where $v(\{\hat{p}\}_{m+1}) = g/(2\hat{p}_l \cdot \hat{p}_{m+1})$. Here $\{p\}_m$ is determined from $\{\hat{p}\}_{m+1}$ by the (six-dimensional version of) the transformation described in section 4. This factorization formula for the amplitude becomes exact in the limit that partons $m+1$ and l become collinear. Away from the collinear limit, there is some freedom to choose the momentum mapping and the splitting amplitude $v(\{\hat{p}\}_{m+1})$. One has to make a definite choice based on ease of computation or conceptual simplicity. In the case of QCD, we have soft as well as collinear singularities, we have parton flavors (which are rather trivially treated) and we have color and spin, which are not so trivial. Let us see how to describe splitting in QCD.

It has been known for a long time that QCD amplitudes factor in the soft and collinear limits [24]. Indeed, there are beautiful modern formulas for the factors [23, 25, 21]. We have adopted a more pedestrian approach that has at least the advantage of encompassing the soft and collinear limits at the same time and of including masses. A treatment of the squared amplitude that is rather similar to the approach of this paper, but at higher order, may be found in ref. [26].

6.1 Definition of the splitting functions v_l

The QCD scattering amplitude for $m+1$ partons is a vector $\left|M(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$ in color $\otimes$ spin space. In the limit that two partons, l and $m+1$ are almost collinear, this amplitude takes a certain limiting form,

$$\left|M(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle \sim \left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle \;, \tag{6.2}$$

where $\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$ is to be defined precisely below. When $\hat{p}_{m+1}$ becomes soft, then all of the $\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$ amplitudes contribute to the limit,

$$\left|M(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle \sim \sum_l \left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle \;. \tag{6.3}$$

We arrange the definition so that eqs. (6.2) and (6.3) are exact in the collinear or soft limit respectively. We also arrange that $\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$ is defined for any $\{\hat{p}, \hat{f}\}_{m+1}$. Then these equations are approximate away from the limit. The amplitude $\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$ is then the contribution to the $(m+1)$-parton amplitude from the splitting of parton l in the parton shower approximation.

We now need to define $\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle$. This amplitude factors into a splitting operator times the m-parton matrix element evaluated at momenta and flavors $\{p, f\}_m$ determined from $\{\hat{p}, \hat{f}\}_{m+1}$ according to the transformation $Q_l(\{\hat{p}, \hat{f}\}_{m+1})$, eq. (4.36) or eq. (4.69),

$$\left|M_l(\{\hat{p}, \hat{f}\}_{m+1})\right\rangle = t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1}) \, V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1}) \left|M(\{p, f\}_m)\right\rangle \;. \tag{6.4}$$

In eq. (6.4), $V^\dagger(\{\hat{p}\}_{m+1})$ is the analogue of $v(\{\hat{p}\}_{m+1})$ but is now an operator on the spin part of the color $\otimes$ spin space. There is also an operator $t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1})$ on the color part of the color $\otimes$ spin space. This operator multiplies by the right color matrix. We will not comment further on it in this section, but will turn to the description of color in section 7. The spin dependent splitting operator can be described in terms of its matrix elements,

$$\left\langle \{\hat{s}\}_{m+1}\big| V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1})\big| \{s\}_m \right\rangle. \tag{6.5}$$

This is a simple function of $\{\hat{p}, \hat{f}\}_{m+1}$, $\{\hat{s}\}_{m+1}$, and $\{s\}_m$. Furthermore, we can take it to be diagonal in the spectator spins,

$$\left\langle \{\hat{s}\}_{m+1}\big| V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1})\big| \{s\}_m \right\rangle = \left(\prod_{j \notin \{l, m+1\}} \delta_{\hat{s}_j, s_j}\right) v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) \;. \tag{6.6}$$

Our object in this section is to define the splitting functions v from the QCD vertices. In writing formulas for $v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l)$, we will use the momentum p_l. We understand that this is obtained from $\{\hat{p}, \hat{f}\}_{m+1}$ according to the transformation $Q_l(\{\hat{p}, \hat{f}\}_{m+1})$.
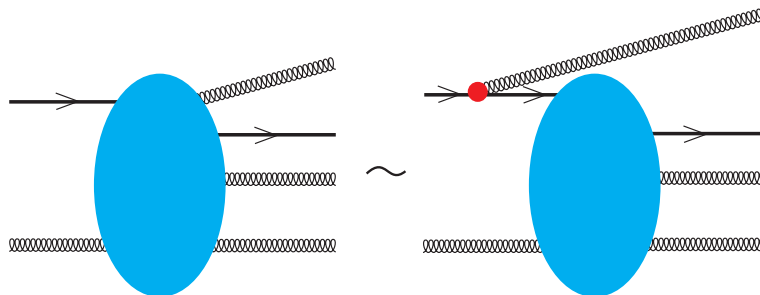
Figure 5: Illustration of factorization at the amplitude level, leading to the definition of the splitting function in eq. (6.19). When the final state gluon becomes collinear with the initial state quark, the full amplitude is approximately the amplitude for one fewer parton, with the incoming quark after the gluon radiation approximated as being on shell, times a splitting function simply made from the QCD qqg vertex and the singular quark propagator. If the final state gluon is soft, then this diagram is one of the possibilities. Then, the soft gluon could have been emitted from any of the external legs of the diagram and we must sum over all of the possible emissions. In this case, we can, however, use a simpler splitting function.

6.2 Initial state $q \to q + \mathrm{g}$ splitting, quark scatters

Consider an initial state $q \to q + \mathrm{g}$ splitting in which the gluon goes into the final state and the quark enters the hard scattering, as illustrated in figure 5. We highlight this case because it exhibits some complications compared to final state $q \to q + \mathrm{g}$ splitting. The parton that splits could be either of the initial state partons. We examine the case that it is parton "a." The kinematics were defined in sections 4.4 and 4.5. In order to keep our notation for this subsection as simple as possible, we will write

$$
\begin{aligned}
p &= p_{\mathrm{a}} \ , & \eta &= \eta_{\mathrm{a}} \ , \\
\hat{p} &= \hat{p}_{\mathrm{a}} \ , & \hat{\eta} &= \hat{\eta}_{\mathrm{a}} \ , \\
q &= \hat{p}_{m+1} \ , & m &= m(f_{\mathrm{a}}) = m(\hat{f}_{\mathrm{a}}) \ , \\
\hat{s} &= \hat{s}_{\mathrm{a}} \ , & \varepsilon_\mu &= \varepsilon_\mu(\hat{p}_{m+1}, \hat{s}_{m+1}; \hat{Q}) \ .
\end{aligned}
\tag{6.7}
$$

Here ε is defined to be orthogonal to $\hat{Q} = \hat{p}_{\mathrm{a}} + \hat{p}_{\mathrm{b}} = \hat{p}_{\mathrm{a}} + p_{\mathrm{b}}$.

The Feynman rules for $\big|M(\{\hat{p}, \hat{f}\}_{m+1})\big\rangle$ give the following structure

$$
M = H \frac{\slashed{P} + m}{P^2 - m^2} \, g t^c \slashed{\varepsilon}^* U(\hat{p}, \hat{s}) \ .
\tag{6.8}
$$

Here there are factors $U(\hat{p}, \hat{s})$ for the initial state quark and ε_μ^* for the final state gluon. There is a vertex $g\gamma_\mu$ with a color matrix t^c. There is a propagator for the off-shell quark that carries momentum

$$
P = \hat{p} - q \ .
\tag{6.9}
$$

The quark propagator has a denominator $P^2 - m^2$ and a numerator $\slashed{P} + m$. Finally, there is the rest of the diagram, H, which contains hard interactions. Thus H carries a Dirac spinor index. We need to relate this to $\big|M(\{p, f\}_m)\big\rangle$, where the momenta and flavors

– 39 –

$\{p, f\}_m$ as well as the splitting variables $\{\zeta_p, \zeta_f\}$ are given by the momentum and flavor mapping $Q_l(\{\hat{p}, \hat{f}\}_{m+1})$, eq. (4.36).

We will be concerned with the behavior of M in the soft and collinear limits. The soft limit is $q \to 0$ (for all four components of q). Since $P^2 - m^2 = -2\hat{p} \cdot q$, the denominator is proportional to a single power of q, while the numerator stays finite in the $q \to 0$ limit. To describe the collinear limit, let $q_\perp$ be the part of q orthogonal to p_A and p_B. The collinear limit is $q_\perp \to 0$ while $q \cdot p_B$ stays finite. Also, $m \to 0$ with $m^2 \lesssim |q_\perp^2|$. Then near the collinear limit, the denominator is

$$P^2 - m^2 = \left(\frac{\hat{\eta}}{\hat{\eta} - \eta} \, q_\perp^2 - \frac{\hat{\eta} - \eta}{\hat{\eta}} \, m^2 \right) (1 + \mathcal{O}(q_\perp^2, m^2)) \ . \tag{6.10}$$

That is, the denominator has two powers of $q_\perp$. To analyze the numerator, we note that

$$
\begin{aligned}
p &= \eta p_A + \mathcal{O}(q_\perp^2, m^2) \ , \\
\hat{p} &= \hat{\eta} p_A + \mathcal{O}(q_\perp^2, m^2) \ , \\
\hat{q} &= (\hat{\eta} - \eta) p_A + q_\perp + \mathcal{O}(q_\perp^2, m^2) \ .
\end{aligned}
\tag{6.11}
$$

The numerator has a factor $(\slashed{P} + m)\slashed{\varepsilon}^* U(\hat{p}, \hat{s})$. Using $q \cdot \varepsilon = 0$ and $(\slashed{p} - m))U(\hat{p}, \hat{s}) = 0$, one can rewrite this factor in the form

$$(\slashed{P} + m)\slashed{\varepsilon}^* U(\hat{p}, \hat{s}) = 2\left(\hat{p} - \frac{\hat{\eta}}{\hat{\eta} - \eta} q \right) \cdot \varepsilon^* \, U(\hat{p}, \hat{s}) + \slashed{\varepsilon}^* \left[\slashed{q} - \frac{\hat{\eta} - \eta}{\hat{\eta}}(\slashed{p} - m) \right] U(\hat{p}, \hat{s}) \ . \tag{6.12}$$

Looking at the numerator in this form, we see that it vanishes in the collinear limit proportionally to one power of $q_\perp$ or m. Thus M is only half as singular in the collinear limit as it first appears.

We now develop an approximation for M. We can insert a factor 1 next to H in eq. (6.8) so that it reads

$$M = H \, \frac{\slashed{n}(\slashed{p} - m) + (\slashed{p} + m)\slashed{n}}{2p \cdot n} \, \frac{\slashed{P} + m}{P^2 - m^2} \, (gt^c \slashed{\varepsilon}^*) U(\hat{p}, \hat{s}) \ . \tag{6.13}$$

Here n is the lightlike vector $n = p_B$.

We now notice that the contribution from the first term, namely

$$M_{\text{ns}} = H \, \frac{\slashed{n}(\slashed{p} - m)}{2p \cdot n} \frac{\slashed{P} + m}{P^2 - m^2} \, (gt^c \slashed{\varepsilon}^*) U(\hat{p}, \hat{s}) \ , \tag{6.14}$$

can be neglected because it does not have a collinear or soft singularity. To see this takes a little analysis. First, we write

$$p = P + (p + q - \hat{p}) \ . \tag{6.15}$$

Since $(\slashed{P} - m)(\slashed{P} + m) = P^2 - m^2$, we have

$$M_{\text{ns}} = H \, \frac{\slashed{n}}{2n \cdot p} \, (gt^a \slashed{\varepsilon}^*) U(\hat{p}, \hat{s}) + H \, \frac{\slashed{n}(\slashed{p} + \slashed{q} - \slashed{\hat{p}})}{2p \cdot n} \frac{\slashed{P} + m}{P^2 - m^2} \, (gt^c \slashed{\varepsilon}^*) U(\hat{p}, \hat{s}) \ . \tag{6.16}$$

The first term is non-singular because the denominator is cancelled. In the second term, the vector $(p + q - \hat{p})$ vanishes in the collinear or soft limit. In the soft limit, it is proportional to one power of q, which cancels the single power of q in the denominator. In the collinear limit, it is proportional to one power of $q_\perp$. As we have just seen, the rest of the numerator contains an additional factor of $q_\perp$. Together, these cancel the two powers of $q_\perp$ from the denominator. Thus no singularity remains.

We are left with

$$M_{\mathrm{sing}} = H \, \frac{(\not{p} + m)\not{n}}{2p \cdot n} \, \frac{\not{P} + m}{P^2 - m^2} \, (gt^c \not{\varepsilon}^*) U(\hat{p}, \hat{s}) \ . \tag{6.17}$$

In eq. (6.17), the factor $\not{p} + m$ is

$$\not{p} + m = \sum_s U(p, s)\overline{U}(p, s) \ . \tag{6.18}$$

The factor $U(p, s)$ is to be associated with H, giving the hard scattering amplitude for an incoming quark with spin s. The remaining factor, $\overline{U}(p, s)$, then becomes part of the splitting function. This calculation leads us to define the splitting function as

$$v_{\mathrm{a}} = -\frac{\sqrt{4\pi\alpha_{\mathrm{s}}}}{(\hat{p} - q)^2 - m^2} \, \varepsilon_\mu^* \, \frac{\overline{U}(p, s)\not{n}(\not{\hat{p}} - \not{q} + m)\gamma^\mu U(\hat{p}, \hat{s})}{2p \cdot n} \ . \tag{6.19}$$

This does not include the color matrix and a factor -1, which will be included in the color operator $t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1})$. As indicated by the derivation, eq. (6.19) is directly given by the factorized structure of QCD Feynman graphs in the soft and collinear limits. There is freedom to choose the form of the splitting function as one moves away from these limits. We have made a simple choice.

In the hard part of the diagram, we can make approximations that are valid for $q_\perp^2 \to 0$. In particular, we can adjust the momenta of the partons with indices other than "a," replacing $\hat{p}_j$ by p_j as defined by the momentum mapping $R_{\mathrm{a}}(\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\})$, eq. (4.63).

6.3 Initial state $q \to q + \mathrm{g}$ splitting, gluon scatters

We consider next the process in which an initial state quark with label "a" splits to make a quark that goes into the final state (with label $m + 1$) and a gluon that enters the hard scattering (with label "a"). We simplify the notation as in the previous subsection by using

$$\begin{aligned}
p &= p_{\mathrm{a}} \ , & \eta &= \eta_{\mathrm{a}} \ , \\
\hat{p} &= \hat{p}_{\mathrm{a}} \ , & \hat{\eta} &= \hat{\eta}_{\mathrm{a}} \ , \\
q &= \hat{p}_{m+1} \ , & m &= m(\hat{f}_{\mathrm{a}}) = m(\hat{f}_{m+1}) \ , \\
\hat{s} &= \hat{s}_{\mathrm{a}} \ , & s' &= \hat{s}_{m+1} \ .
\end{aligned} \tag{6.20}$$

The Feynman rules for $\left| M(\{\hat{p}, \hat{f}\}_{m+1}) \right\rangle$ give the following structure

$$M = H^\mu \, \frac{D_{\mu\nu}(P; n)}{P^2} \, \overline{U}(q, s')(gt^c \gamma^\nu) U(\hat{p}, \hat{s}) \ . \tag{6.21}$$

Here H is the hard part of the graph, now with a vector index, and there is a propagator for the off-shell gluon that carries momentum

$$P = \hat{p} - q \ . \tag{6.22}$$

We have chosen the axial gauge $n \cdot A = 0$, where n is the lightlike vector $n = P_{\mathrm{B}}$. The numerator of the gluon propagator is

$$D^{\mu\nu}(P;n) = -g^{\mu\nu} + \frac{P^\mu n^\nu + n^\mu P^\nu}{P \cdot n} \ . \tag{6.23}$$

As in the previous subsection, M is singular in the collinear limit, in which $q_\perp \to 0$ and $m \to 0$.[19] As in the previous section, the numerator is proportional to $q_\perp$ and the denominator is proportional to $q_\perp^2$ in the collinear limit.

In order to find a suitable approximation for M, we insert 1 next to H, so that it reads

$$M = H_\alpha \left[-D^{\alpha\mu}(p;n) + \frac{p^\alpha n^\mu + n^\alpha p^\mu}{p \cdot n} \right] \frac{D_{\mu\nu}(P;n)}{P^2} \ \overline{U}(q,s')(gt^c\gamma^\nu)U(\hat{p},\hat{s}) \ . \tag{6.24}$$

We can drop the term $(p^\alpha n^\mu + n^\alpha p^\mu)/(n \cdot p)$. To see this, we write

$$\frac{p^\alpha n^\mu + n^\alpha p^\mu}{p \cdot n} = \frac{P^\alpha n^\mu + n^\alpha P^\mu}{P \cdot n} + D^{\alpha\mu}(p;n) - D^{\alpha\mu}(P;n) \ . \tag{6.25}$$

The term $P^\alpha n^\mu$ gives zero when contracted with $D_{\mu\nu}$. When we contract $n^\alpha P^\mu$ with $D_{\mu\nu}$, we get

$$\frac{n^\alpha P^\mu}{P \cdot n} \frac{D_{\mu\nu}(P;n)}{P^2} = \frac{n^\alpha n_\nu}{(P \cdot n)^2} \ , \tag{6.26}$$

which does not have a collinear singularity. Finally, the difference $D^{\alpha\mu}(p;n) - D^{\alpha\mu}(P;n)$ is proportional to $q_\perp$ in the collinear limit because $P - p \propto q_\perp$ in this limit. The remaining numerator factor gives another factor $q_\perp$ in the collinear limit, so that the factor $q_\perp^2$ from the denominator is cancelled.

Thus we are left with $M \sim M_{\mathrm{sing}}$, where

$$M_{\mathrm{sing}} = -H_\alpha \, D^{\alpha\mu}(p;n) \, \frac{D_{\mu\nu}(P;n)}{P^2} \ \overline{U}(q,s')(gt^c\gamma^\nu)U(\hat{p},\hat{s}) \ . \tag{6.27}$$

The factor $D^{\alpha\mu}(p;n)$ is

$$D^{\alpha\mu}(p;n) = \sum_s \varepsilon^\alpha(p,s;n)\varepsilon^\mu(p,s;n)^* = \sum_s \varepsilon^\alpha(p,s;\hat{Q})\varepsilon^\mu(p,s;\hat{Q})^* \ . \tag{6.28}$$

It is equivalent to use $n = p_{\mathrm{B}}$ or $\hat{Q} = \hat{p}_{\mathrm{a}} + \hat{p}_{\mathrm{b}} = \hat{p}_{\mathrm{a}} + p_{\mathrm{b}}$ to define the polarization vectors since they are orthogonal to both $p = p_{\mathrm{a}}$ and p_{b}. The factor $\varepsilon^\alpha(p,s;\hat{Q})$ is to be associated with H, giving the hard scattering amplitude for a gluon with spin s. The remaining factor, $\varepsilon^\mu(p,s;\hat{Q})^*$, then becomes part of the splitting function.

---

[19]We do not need to be concerned with the soft limit, $q \to 0$ with $m \to 0$. Here, there is only a $1/\sqrt{q}$ singularity, which is too weak to create a logarithmically divergent integration over final states. Nevertheless, our approximation to M also matches the behavior of M in the soft limit.

This calculation leads us to define the splitting function as

$$v_{\mathrm{a}} = -\frac{\sqrt{4\pi\alpha_{\mathrm{s}}}}{(\hat{p}-q)^2}\, \varepsilon^{\mu}(p,s;\hat{Q})^*\, D_{\mu\nu}(\hat{p}-q;n)\, \overline{U}(q,s')\gamma^{\nu}U(\hat{p},\hat{s}) \ . \tag{6.29}$$

This does not include the color matrix, which is included in the color operator $t_l^{\dagger}(f_l \to \hat{f}_l + \hat{f}_{m+1})$.

In the hard part of the diagram, we can again make approximations that are valid for $q_{\perp}^2 \to 0$. In particular, we can adjust the momenta of the partons with indices other than "a," replacing $\hat{p}_j$ by p_j as defined by the momentum mapping $R_{\mathrm{a}}(\{p,f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\})$, eq. (4.63). We will also want to use $Q = p_{\mathrm{a}} + p_{\mathrm{b}}$ instead of $\hat{Q} = \hat{p}_{\mathrm{a}} + \hat{p}_{\mathrm{b}}$ to define the polarization vector for parton l as it enters the hard scattering. This is just a change in notation rather than an approximation since Q is in the p-$\hat{Q}$ plane.

6.4 Other qqg splittings

For splittings involving a qqg vertex, there are several other choices for the flavors f_l, $\hat{f}_l$ and $\hat{f}_{m+1}$ and for whether the index l of the parton that splits is a final state index or an initial state index. The results for $v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l)$ are listed in table 1. In constructing this table, we keep track of two sign factors. First, there is a numerator sign that we compute as in the examples in the preceding subsections. This is a minus sign whenever a gluon or an antiquark enters the hard scattering from the initial state or leaves the hard scattering to the final state and a plus sign when a quark enters or leaves the hard scattering. Second, there is a color sign. There is always a color matrix t^a. We count a color **3** line entering the hard scattering as a color $\bar{\mathbf{3}}$ line leaving the hard scattering. Viewed this way, the color matrix is actually $(t^a)^T$. However, the generator of color rotations for the $\bar{\mathbf{3}}$ representation of SU(3) is $-(t^a)^T$. Thus we include $-(t^a)^T$ as part of the color operator $t_l^{\dagger}(f_l \to \hat{f}_l + \hat{f}_{m+1})$ and include the minus sign as part of the splitting function v_l. This leaves a factor -1 in v_l whenever a $\bar{\mathbf{3}}$ line leaves the hard scattering. The sign included in $t_l^{\dagger}(f_l \to \hat{f}_l + \hat{f}_{m+1})$ is indicated in the last column of table 1. The net sign of v_l is shown in the table.

The construction for a final state splitting, as for an initial state splitting, makes use of a lightlike vector n_l that is in the plane of p_l and $\hat{Q} = \hat{p}_{\mathrm{a}} + \hat{p}_{\mathrm{b}}$, which is the same as the plane of p_l and $Q = p_{\mathrm{a}} + p_{\mathrm{b}}$. The normalization of n_l is not significant. Our choice for this vector was defined in eq. (5.3).

6.5 Splitting with a ggg vertex

We construct the splitting function for a $\mathrm{g} \to \mathrm{g} + \mathrm{g}$ splitting in a similar fashion. In the case of a final state splitting, we use the ggg QCD vertex,

$$v^{\alpha\beta\gamma}(p_a, p_b, p_c) = g^{\alpha\beta}(p_a - p_b)^{\gamma} + g^{\beta\gamma}(p_b - p_c)^{\alpha} + g^{\gamma\alpha}(p_c - p_a)^{\beta} \ , \tag{6.30}$$

to define

$$v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) = \frac{\sqrt{4\pi\alpha_{\mathrm{s}}}}{2\hat{p}_{m+1}\cdot\hat{p}_l}\, \varepsilon_{\alpha}(\hat{p}_{m+1}, \hat{s}_{m+1}; \hat{Q})^*\varepsilon_{\beta}(\hat{p}_l, \hat{s}_l; \hat{Q})^*\varepsilon^{\nu}(p_l, s_l; \hat{Q})$$

l	f_l	$\hat{f}_l$	$\hat{f}_{m+1}$	$v_l \times \dfrac{1}{\sqrt{4\pi\alpha_{\rm s}}}$	color
F	q	q	g	$\varepsilon_\mu(\hat{p}_{m+1},\hat{s}_{m+1};\hat{Q})^* \dfrac{\overline{U}(\hat{p}_l,\hat{s}_l)\gamma^\mu[\slashed{\hat{p}}_l + \slashed{\hat{p}}_{m+1} + m(f_l)]\slashed{n}_l U(p_l,s_l)}{2p_l\cdot n_l\,[(\hat{p}_l + \hat{p}_{m+1})^2 - m^2(f_l)]}$	t^a
F	$\bar{q}$	$\bar{q}$	g	$\varepsilon_\mu(\hat{p}_{m+1},\hat{s}_{m+1};\hat{Q})^* \dfrac{\overline{V}(p_l,s_l)\slashed{n}_l[\slashed{\hat{p}}_l + \slashed{\hat{p}}_{m+1} - m(f_l)]\gamma^\mu V(\hat{p}_l,\hat{s}_l)}{2p_l\cdot n_l\,[(\hat{p}_l + \hat{p}_{m+1})^2 - m^2(f_l)]}$	$-t^a$
I	$\bar{q}$	$\bar{q}$	g	$-\varepsilon_\mu(\hat{p}_{m+1},\hat{s}_{m+1};\hat{Q})^* \dfrac{\overline{U}(p_l,s_l)\slashed{n}_l(\slashed{\hat{p}}_l - \slashed{\hat{p}}_{m+1} + m(f_l))\gamma^\mu U(\hat{p}_l,\hat{s}_l)}{2p_l\cdot n_l\,[(\hat{p}_l - \hat{p}_{m+1})^2 - m^2(f_l)]}$	$-t^a$
I	q	q	g	$-\varepsilon_\mu(\hat{p}_{m+1},\hat{s}_{m+1};\hat{Q})^* \dfrac{\overline{V}(\hat{p}_l,\hat{s}_l)\gamma^\mu(\slashed{\hat{p}}_l - \slashed{\hat{p}}_{m+1} - m(f_l))\slashed{n}_l V(p_l,s_l)}{2p_l\cdot n_l\,[(\hat{p}_l - \hat{p}_{m+1})^2 - m^2(f_l)]}$	t^a
F	g	q	$\bar{q}$	$-\varepsilon^\mu(p_l,s_l;\hat{Q})D_{\mu\nu}(\hat{p}_l + \hat{p}_{m+1},n_l)\dfrac{\overline{U}(\hat{p}_l,\hat{s}_l)\gamma^\nu V(\hat{p}_{m+1},\hat{s}_{m+1})}{(\hat{p}_l + \hat{p}_{m+1})^2}$	t^a
I	g	$\bar{q}$	q	$-\varepsilon^\mu(p_l,s_l;\hat{Q})^* D_{\mu\nu}(\hat{p}_l - \hat{p}_{m+1};n_l)\dfrac{\overline{U}(\hat{p}_{m+1},\hat{s}_{m+1})\gamma^\nu U(\hat{p}_l,\hat{s}_l)}{(\hat{p}_l - \hat{p}_{m+1})^2}$	t^a
I	g	q	$\bar{q}$	$-\varepsilon^\mu(p_l,s_l;\hat{Q})^* D_{\mu\nu}(\hat{p}_l - \hat{p}_{m+1};n_l)\dfrac{\overline{V}(\hat{p}_l,\hat{s}_l)\gamma^\nu V(\hat{p}_{m+1},\hat{s}_{m+1})}{(\hat{p}_l - \hat{p}_{m+1})^2}$	t^a
I	q	g	q	$-\varepsilon_\mu(\hat{p}_l,\hat{s}_l;\hat{Q})\dfrac{\overline{U}(\hat{p}_{m+1},\hat{s}_{m+1})\gamma^\mu[\slashed{\hat{p}}_l - \slashed{\hat{p}}_{m+1} - m(f_l)]\slashed{n}_l V(p_l,s_l)}{2p_l\cdot n_l\,[(\hat{p}_l - \hat{p}_{m+1})^2 - m^2(f_l)]}$	t^a
I	$\bar{q}$	g	$\bar{q}$	$-\varepsilon_\mu(\hat{p}_l,\hat{s}_l;\hat{Q})\dfrac{\overline{U}(p_l,s_l)\slashed{n}_l[\slashed{\hat{p}}_l - \slashed{\hat{p}}_{m+1} + m(f_l)]\gamma^\mu V(\hat{p}_{m+1},\hat{s}_{m+1})}{2p_l\cdot n_l\,[(\hat{p}_l - \hat{p}_{m+1})^2 - m^2(f_l)]}$	$-t^a$

Table 1: Splitting functions $v_l(\{\hat{p},\hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l)$ for splittings involving a $q\bar{q}g$ vertex, with a common factor $\sqrt{4\pi\alpha_{\rm s}}$ removed. The values of l are either in the set of initial state indices $I = \{{\rm a,b}\}$ or in the set of final state indices $F = \{1,\dots,m\}$. The flavors f_l, $\hat{f}_l$, and $\hat{f}_{m+1}$ can be $f = {\rm g}$ or can be a quark index, $q \in \{{\rm u,d},\dots\}$, or an antiquark index $\bar{q} \in \{\bar{\rm u},\bar{\rm d},\dots\}$. Recall from eq. (3.2) that in the case of an initial state parton, f_l and $\hat{f}_l$ denote the opposite of the incoming flavors of the parton. For a row in which q denotes a quark index, $\bar{q}$ denotes the corresponding antiquark index. The next column gives the value of v_l corresponding to the values of l, f_l, $\hat{f}_l$, and $\hat{f}_{m+1}$ indicated. The last column indicates the sign of the color matrix that is incorporated into the color operator $t_l^\dagger$. The lightlike vector n_l is defined in eq. (5.3).

$$\times v^{\alpha\beta\gamma}(\hat{p}_{m+1}, \hat{p}_l, -\hat{p}_{m+1} - \hat{p}_l)\, D_{\gamma\nu}(\hat{p}_l + \hat{p}_{m+1};n_l). \quad (6.31)$$

For an initial state splitting, we have

$$v_l(\{\hat{p},\hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) = -\frac{\sqrt{4\pi\alpha_{\rm s}}}{2\hat{p}_{m+1}\cdot\hat{p}_l}\, \varepsilon_\alpha(\hat{p}_{m+1}, \hat{s}_{m+1};\hat{Q})^* \varepsilon_\beta(\hat{p}_l, \hat{s}_l;\hat{Q})\varepsilon^\nu(p_l, s_l;\hat{Q})^*$$
$$\times v^{\alpha\beta\gamma}(\hat{p}_{m+1}, -\hat{p}_l, \hat{p}_l - \hat{p}_{m+1})\, D_{\gamma\nu}(\hat{p}_l - \hat{p}_{m+1};n_l). \quad (6.32)$$

In each case, we have the exact QCD vertex and the exact propagator for the off-shell gluon in $n_l \cdot A = 0$ gauge followed by a projection onto the physical gluon degrees of freedom contained in the on-shell polarization vector.

6.6 Soft splitting function

These splitting functions enable us to approximate the $(m + 1)$-parton matrix element in the cases that $\hat{p}_{m+1}$ is collinear with $\hat{p}_l$ or else $\hat{p}_{m+1}$ is soft. In the special case that $\hat{p}_{m+1}$ is soft, or possibly soft and collinear with $\hat{p}_l$, a simpler splitting function can be used. When $\hat{p}_{m+1}$ is soft, we have

$$\left| M(\{\hat{p}, \hat{f}\}_{m+1}) \right\rangle \sim \sum_l \left| M_l^{\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \right\rangle \ , \tag{6.33}$$

where

$$\left| M_l^{\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \right\rangle = t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1}) \, V_l^{\dagger,\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \left| M(\{p, f\}_m) \right\rangle \ . \tag{6.34}$$

The matrix elements of $V_l^{\dagger,\text{soft}}$ are specified by a function v_l^{soft},

$$\left\langle \{\hat{s}\}_{m+1} \middle| V_l^{\dagger,\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \middle| \{s\}_m \right\rangle = \left(\prod_{j \notin \{l, m+1\}} \delta_{\hat{s}_j, s_j} \right) v_l^{\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) \ . \tag{6.35}$$

If parton $m + 1$ is a quark or antiquark, $v_l^{\text{soft}} = 0$. When parton $m + 1$ is a gluon,

$$v_l^{\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) = \sqrt{4\pi\alpha_{\text{s}}} \, \delta_{\hat{s}_l, s_l} \, \frac{\varepsilon(\hat{p}_{m+1}, \hat{s}_{m+1}; \hat{Q})^* \cdot \hat{p}_l}{\hat{p}_{m+1} \cdot \hat{p}_l} \ . \tag{6.36}$$

The functions v_l^{soft} are not as powerful as the functions v_l because they provide good approximations only in the soft gluon limit. Nevertheless, we will have occasion to make use of them.

7. Description of color

We will need a description of the quantum color state that is adapted to a description of shower evolution. If we use an index notation, to each parton with label l there is associated a color index a_l, which takes values $1, \ldots, 3$ for a quark or antiquark and takes values $1, \ldots, 8$ for a gluon. There is also a spin index λ_l, which takes values $\pm 1/2$ for quark and ± 1 for a gluon. We can expand M in terms of color and spin basis vectors in the form

$$M(\{p, f\}_m)^{a_{\text{a}}, a_{\text{b}}, a_1, \ldots, a_m}_{\lambda_{\text{a}}, \lambda_{\text{b}}, \lambda_1, \ldots, \lambda_m} = \sum_{\{c\}_m} \Psi(\{c\}_m)^{a_{\text{a}}, a_{\text{b}}, a_1, \ldots, a_m} \sum_{\{s\}_m} \Xi(\{s\}_m)_{\lambda_{\text{a}}, \lambda_{\text{b}}, \lambda_1, \ldots, \lambda_m}$$
$$\times M(\{p, f, s, c\}_m) \ , \tag{7.1}$$

where the $\Psi(\{c\}_m)$ form a basis for the space of color singlet amplitudes with color labels $\{c\}_m$ and the $\Xi(\{s\}_m)$ form a basis for the spin space with spin labels $\{s\}_m$. The quantities $M(\{p, f, s, c\})$ are the expansion coefficients. In a vector notation, this is

$$\left| M(\{p, f\}_m) \right\rangle_{\text{c,s}} = \sum_{\{c\}_m} \left| \{c\}_m \right\rangle_{\text{c}} \otimes \sum_{\{s\}_m} \left| \{s\}_m \right\rangle_{\text{s}} M(\{p, f, s, c\}_m) \ . \tag{7.2}$$

Here $\left|M(\{p,f\}_m)\right\rangle_{\mathrm{c,s}}$ lies in the combined color-spin space while $\left|\{c\}_m\right\rangle_{\mathrm{c}}$ is a vector in color space and $\left|\{s\}_m\right\rangle_{\mathrm{s}}$ is a vector in spin space.

As discussed in section 5, we use a conventional treatment of spin. We assume that the spin labels λ already represent parton helicities using suitable conventions for choosing corresponding Dirac spinors and polarization vectors. Then the basis vector labels can be simply $\{s\}_m = \{s_{\mathrm{a}}, s_{\mathrm{b}}, s_1, \ldots, s_m\}$ and the basis vectors can be

$$\Xi(\{s\}_m)_{\lambda_{\mathrm{a}}, \lambda_{\mathrm{b}}, \lambda_1, \ldots, \lambda_m} = \delta^{\lambda_{\mathrm{a}}}_{s_{\mathrm{a}}} \delta^{\lambda_{\mathrm{b}}}_{s_{\mathrm{b}}} \delta^{\lambda_1}_{s_1} \cdots \delta^{\lambda_m}_{s_m} \ . \tag{7.3}$$

Then we have an orthonormal basis: $\left\langle \{s'\}_m \middle| \{s\}_m \right\rangle$ is 1 if the spin labels are all the same and zero otherwise.

We use a treatment of color that is conventional but more subtle than the treatment for spin. We turn to this subject in this section.

7.1 Color basis

We first note that as far as color is concerned, an initial state quark is equivalent to a final state antiquark and an initial state antiquark is equivalent to a final state quark. Thus, in the prose description in this section we use "quark" and "q" to refer to a final state quark or an initial state antiquark and we use "antiquark" and "$\bar{q}$" to refer to a final state antiquark or an initial state quark.[20]

We next note that the amplitude $\left|M(\{p,f\}_m)\right\rangle$ us always invariant under an overall rotation of all of the parton colors. Thus what we really need is a basis for the space of color singlet amplitudes in the color space. There is a widely used and intuitively appealing way to do this that, furthermore, matches with the idea of color strings forming between outgoing partons [27].

The color basis vectors $\left|\{c\}_m\right\rangle_{\mathrm{c}}$ are labeled by a *color string configurations* $\{c\}_m$. A color string configuration can be described as a set $\{S_1, \ldots, S_n\}$ of one or more *strings* S. There are two types of strings, *open strings* and *closed strings*. An open string is an ordered set of parton indices that we denote by $S = [l_1, l_2, \ldots, l_{n-1}, l_n]$. Here l_1 is the label of a quark, l_n is the label of an antiquark, and $l_2, \ldots, l_{n-1}$ are labels of gluons. A closed string is an ordered set of at least two parton indices that we denote by $S = (l_1, l_2, \ldots, l_{n-1}, l_n)$. Here all of the indices label gluons and we treat sets that differ by a cyclic permutation of the indices as being the same. Thus a complete color string configuration for a quark, an antiquark, and five gluons might be $\{[1,6,3,7], (4,2,5)\}$. This is a notation for a possible $\{c\}_5$.

Now we can define the basis states. We take $\Psi(\{c\}_m)$ to be a product

$$\Psi(\{c\}_m)^{a_1, \ldots, a_m} = \Psi(S_1)^{\{a\}_{[1]}} \Psi(S_2)^{\{a\}_{[2]}} \ldots \Psi(S_K)^{\{a\}_{[K]}} \ . \tag{7.4}$$

Here we have denoted the set of color indices represented in string k by

$$\{a\}_{[k]} = \{a_{l_1}, \ldots, a_{l_n}\} \tag{7.5}$$

if string k is $[l_1, \ldots, l_n]$ or $(l_1, \ldots, l_n)$.

---

[20]Recall from eq. (3.2) that in the case of an initial state parton, f_l and $\hat{f}_l$ denote the opposite of the incoming flavors of the parton.
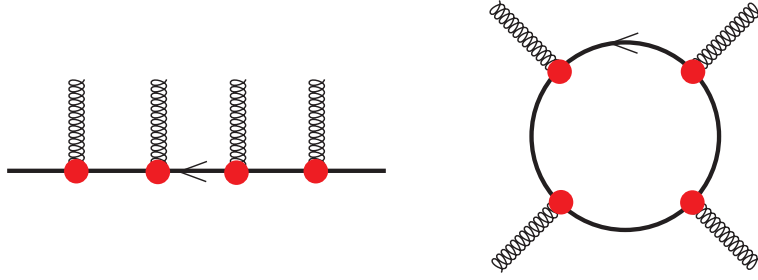
Figure 6: Illustration an open string color basis state (left) and a closed sting color basis state (right).

We can now define the component factors $V(S)$. For notational convenience, we suppose that $l_i = i$ so that the partons along the string are numbered sequentially, $[1, 2, \ldots, n]$ or $(1, 2, \ldots, n)$.

We first consider an open string, as illustrated on the left in figure 6. We define

$$\Psi(S)^{\{a\}} = n(S)^{-1/2} \left[t^{a_2} t^{a_3} \cdots t^{a_{n-1}} \right]_{a_1 a_n} \quad , \tag{7.6}$$

where the t^a are the SU(3) generator matrices for the fundamental representation and we take the a_1, a_n matrix element of the matrix product of the generator matrices (normalized to $T_R = 1/2$). The normalization factor $n(S)$ is

$$n(S) = N_c C_F^{n-2} \quad . \tag{7.7}$$

With this normalization,

$$\langle S|S \rangle \equiv \sum_{\{a\}} |\Psi(S)^{\{a\}}|^2 = 1 \quad . \tag{7.8}$$

For a closed string with the same parton labels (now all gluons) we define, as illustrated on the right in figure 6,

$$\Psi(S)^{\{a\}} = n(S)^{-1/2} \, \mathrm{Tr} \left[t^{a_1} t^{a_2} \cdots t^{a_n} \right] \quad , \tag{7.9}$$

where, again, the t^a are the SU(3) generator matrices for the fundamental representation[21] and where

$$n(S) = C_F^n \quad . \tag{7.10}$$

With this normalization,

$$\langle S|S \rangle \equiv \sum_{\{a\}} |\Psi(S)^{\{a\}}|^2 = 1 - \left(\frac{-1}{2 N_c C_F} \right)^{n-1} \quad . \tag{7.11}$$

This is approximately 1 in the limit of a large number of colors.

---

[21]One could use the adjoint representation here. However, an adjoint representation string is approximately equivalent to two fundamental representation strings and having two strings when one would do makes the description more complicated.
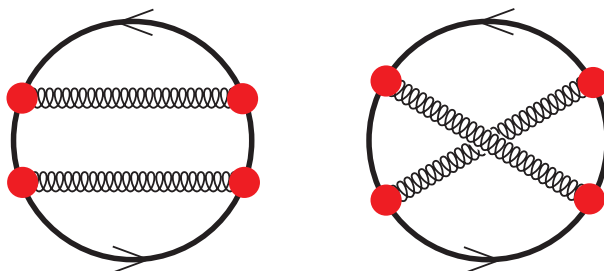
Figure 7: Inner products for color basis states. The left hand picture illustrates the inner product of the $[1,2,3,4]$ with itself. The color diagram shown gives $C_F^2 N_c$, which is just canceled by the normalization factor $1/n(S)$ from eq. (7.7). The right hand picture illustrates the inner product of the $[1,2,3,4]$ with $[1,3,2,4]$. The color diagram gives $-C_F/2$. Multiplying by $1/n(S)$ gives $-1/(N_c^2 - 1)$.

A general color basis state is a product of string states.[22] It includes a normalization factor $n(\{c\}_m)^{-1/2}$,

$$n(\{c\}_m) = n(S_1)\, n(S_2) \cdots n(S_K) \ . \tag{7.12}$$

The normalization of the states is

$$\big\langle \{c\}_m \big| \{c\}_m \big\rangle = \prod_k \big\langle S_k | S_k \big\rangle \ , \tag{7.13}$$

where the factors are given in eqs. (7.8) and (7.11). Thus $\big\langle \{c\}_m \big| \{c\}_m \big\rangle \approx 1$ in the large N_c limit.

The basis vectors $\big| \{c\}_m \big\rangle$ are not exactly normalized and they are not orthogonal. However, the inner product between two different basis vectors is small in the limit of a large number of colors,

$$\big\langle \{c'\}_m \big| \{c\}_m \big\rangle = \mathcal{O}(1/N_c^2) \qquad \{c'\}_m \neq \{c\}_m \ . \tag{7.14}$$

For instance, suppose $\{c\}_m = [1,2,3,4]$ and suppose $\{c'\}_m = [1,3,2,4]$, with the positions of the two gluons reversed. Then $\big\langle \{c'\}_m \big| \{c\}_m \big\rangle = -1/(N_c^2 - 1)$. The calculation is illustrated in figure 7.

7.2 Parton insertion operators

It will prove useful to define certain operators that act on an m parton color state with partons with labels $\{a, b, 1, \ldots, m\}$ and add one parton with label $m + 1$. We let $a_+^\dagger(l)$

---

[22]The building blocks used are the invariant matrices δ_{ij} to connect a quark and an antiquark and t_{ij}^a to connect a quark, antiquark and gluon. One could also use the completely antisymmetric matrix ϵ_{ijk} to connect three quarks or three antiquarks. However, we don't need states made using ϵ_{ijk} because the amplitudes M have net baryon number zero.

insert a gluon just to the right of parton l on whatever string contains parton l,

$$
\begin{aligned}
a_+^\dagger(l)\left|\{\ldots,[a_1,\ldots,l,\ldots,a_n],\ldots\}\right\rangle &= \left|\{\ldots,[a_1,\ldots,l,m+1,\ldots,a_n],\ldots\}\right\rangle \;, \\
a_+^\dagger(l)\left|\{\ldots,(a_1,\ldots,l,\ldots,a_n),\ldots\}\right\rangle &= \left|\{\ldots,(a_1,\ldots,l,m+1,\ldots a_n),\ldots\}\right\rangle \;, \\
a_+^\dagger(l)\left|\{\ldots,[l,\ldots,a_n],\ldots\}\right\rangle &= \left|\{\cdots[l,m+1,\ldots,a_n],\ldots\}\right\rangle \;, \\
a_+^\dagger(l)\left|\{\ldots,[a_1,\ldots,l],\ldots\}\right\rangle &= 0 \;.
\end{aligned}
\tag{7.15}
$$

The first equation applies to the case that l labels a gluon on an open string while the second equation applies to the case that l labels a gluon on an closed string. The third equation applies when l labels a quark at the end of an open string. The fourth equation applies when l labels an antiquark at the end of an open string. In this case, there is no place to put the gluon, so the result is defined to be zero. Similarly, we define an operator $a_-^\dagger(l)$ that inserts a gluon just to the left of parton l on whatever string contains parton l,

$$
\begin{aligned}
a_-^\dagger(l)\left|\{\ldots,[a_1,\ldots,l,\ldots,a_n],\ldots\}\right\rangle &= \left|\{\ldots,[a_1,\ldots,m+1,l,\ldots a_n],\ldots\}\right\rangle \;, \\
a_-^\dagger(l)\left|\{\ldots,(a_1,\ldots,l,\ldots,a_n),\ldots\}\right\rangle &= \left|\{\ldots,(a_1,\ldots,m+1,l,\ldots a_n),\ldots\}\right\rangle \;, \\
a_-^\dagger(l)\left|\{\ldots,[l,\ldots,a_n],\ldots\}\right\rangle &= 0 \;, \\
a_-^\dagger(l)\left|\{\ldots,[a_1,\ldots,l],\ldots\}\right\rangle &= \left|\{\ldots,[a_1,\ldots,m+1,l],\ldots\}\right\rangle \;.
\end{aligned}
\tag{7.16}
$$

We define an operator $a_q^\dagger(l)$ that breaks a string at the position of a gluon with label l, creating a quark that we take to inherit the label l and an antiquark with the new label $m+1$. If l is a gluon on an open string, this creates two open strings. If l is a gluon on a closed string, this turns the closed string into an open string. If l labels a quark or antiquark, we define the result to be zero. Thus

$$
\begin{aligned}
a_q^\dagger(l)\left|\{\ldots,[a_1,\ldots,l,\ldots,a_n],\ldots\}\right\rangle &= \left|\{\ldots,[a_1,\ldots,m+1],[l,\ldots a_n],\ldots\}\right\rangle \;, \\
a_q^\dagger(l)\left|\{\ldots,(a_1,\ldots,l,\ldots,a_n),\ldots\}\right\rangle &= \left|\{\ldots,[l,\ldots a_n,a_1,\ldots,m+1],\ldots\}\right\rangle \;, \\
a_q^\dagger(l)\left|\{\ldots,[l,\ldots,a_n],\ldots\}\right\rangle &= 0 \;, \\
a_q^\dagger(l)\left|\{\ldots,[a_1,\ldots,l],\ldots\}\right\rangle &= 0 \;.
\end{aligned}
\tag{7.17}
$$

Finally, we define an operator $a_s^\dagger(l)$ that removes a gluon with label l from its string, and creates a new open string consisting of just a quark that we take to inherit the label l and an antiquark with the new label $m+1$. If l labels a quark or antiquark, we define the result to be zero. Thus

$$
\begin{aligned}
a_s^\dagger(l)\left|\{\ldots,[a_1,\ldots,l,\ldots,a_n],\ldots\}\right\rangle &= \left|\{\ldots,[a_1,\ldots,a_n],[l,m+1]\cdots\}\right\rangle \;, \\
a_s^\dagger(l)\left|\{\ldots,(a_1,\ldots,l,\ldots,a_n),\ldots\}\right\rangle &= \left|\{\ldots,(a_1,\ldots,a_n),[l,m+1],\ldots\}\right\rangle \;, \\
a_s^\dagger(l)\left|\{\ldots,[l,\ldots,a_n],\ldots\}\right\rangle &= 0 \;, \\
a_s^\dagger(l)\left|\{\ldots,[a_1,\ldots,l],\ldots\}\right\rangle &= 0 \;.
\end{aligned}
\tag{7.18}
$$

We will see the utility of these operators presently, when we study the color flow in parton splitting.

7.3 Color evolution for the quantum states

In this subsection, examine how the color vector changes with successive parton splittings. Consider starting with a color state $|\psi\rangle$ for m partons and letting a final state parton with label l and flavor f_l emit a gluon with label $m + 1$.[23] After the gluon emission, parton l retains its label l and its flavor. This produces a new color state

$$\big|\hat{\psi}\big\rangle = t_l^\dagger(f_l \to f_l + \mathrm{g})\big|\psi\big\rangle \ . \tag{7.19}$$

We can define the color splitting operator $t_l^\dagger(f_l \to f_l + \mathrm{g})$ precisely by writing this out in component notation

$$\hat{\psi}^{a_\mathrm{a},a_\mathrm{b},a_1,\ldots,a_l,\ldots,a_m,a_{m+1}} = \sum_{\tilde{a}_l} T(f_l)^{a_{m+1}}_{a_l,\tilde{a}_l} \psi^{a_\mathrm{a},a_\mathrm{b},a_1,\ldots,\tilde{a}_l,\ldots,a_m} \ . \tag{7.20}$$

The matrix $T(f_l)^{a_{m+1}}$ here is the generator matrix in the **8** representation if $f_l = \mathrm{g}$, in the **3** representation if $f_l \in \{\mathrm{u},\mathrm{d},\ldots\}$, and in the $\bar{\mathbf{3}}$ representation if $f_l \in \{\bar{\mathrm{u}},\bar{\mathrm{d}},\ldots\}$.[24]

This notation applies for an initial state splitting too since, from eq. (3.2), in the case of an initial state parton, f_l and $\hat{f}_l$ denote the opposite of the incoming flavors of the parton. In this case, the arrow in $f_l \to f_l + \mathrm{g}$ and more generally in $f_l \to \hat{f}_l + \hat{f}_{m+1}$ refers to backward evolution.

Now consider what happens when the splitting operator is applied to one of the color basis vectors, $\big|\{c\}_m\big\rangle$. For the case that $f_l = \mathrm{g}$ we have

$$\sum_{a_l} T(\mathrm{g})^{\tilde{a}_{m+1}}_{\tilde{a}_l,a_l} \ t^{a_l} = \sum_{a_l} i f_{\tilde{a}_l,\tilde{a}_{m+1},a_l} \ t^{a_l} = t^{\tilde{a}_l} t^{\tilde{a}_{m+1}} - t^{\tilde{a}_{m+1}} t^{\tilde{a}_l} \ . \tag{7.21}$$

Thus $t_l^\dagger(\mathrm{g} \to \mathrm{g} + \mathrm{g})\big|\{c\}_m\big\rangle$ is a normalization factor times a difference of two basis vectors

$$\left[\frac{n(\{\hat{c}_+\}_{m+1})}{n(\{c\}_m)}\right]^{1/2} \big|\{\hat{c}_+\}_{m+1}\big\rangle - \left[\frac{n(\{\hat{c}_-\}_{m+1})}{n(\{c\}_m)}\right]^{1/2} \big|\{\hat{c}_-\}_{m+1}\big\rangle \ . \tag{7.22}$$

In $\{\hat{c}_\pm\}_{m+1}$, gluon l is replaced by gluons l and $m + 1$. Specifically, if in $\{c\}_m$ gluon l appears in an open string as $[i_1,\ldots,l,\ldots,i_n]$ then in $\{\hat{c}_+\}_{m+1}$ this string is replaced by $[i_1,\ldots,l,m+1,\ldots,i_n]$ and in $\{\hat{c}_-\}_{m+1}$ this string is replaced by $[i_1,\ldots,m+1,l,\ldots,i_n]$. The normalization factor is

$$\frac{n(\{\hat{c}_\pm\}_{m+1})}{n(\{c\}_m)} = C_\mathrm{F} \ . \tag{7.23}$$

Analogous remarks apply for a closed string. We can state this result using the operators defined in eqs. (7.15) and (7.16),

$$t_l^\dagger(\mathrm{g} \to \mathrm{g} + \mathrm{g}) = \sqrt{C_\mathrm{F}}\, a_+^\dagger(l) - \sqrt{C_\mathrm{F}}\, a_-^\dagger(l) \ . \tag{7.24}$$

This is illustrated in figure 8.

---

[23]In [14], the final state labels are called i and j. We identify i with l, and j with $m+1$. The labels for the initial m partons are $\{\mathrm{a},\mathrm{b},1,\ldots,m\}$ and for the subsequent state after splitting are $\{\mathrm{a},\mathrm{b},1,\ldots,m,m+1\}$. The partons not involved in the splitting keep their labels.

[24]Specifically, $T(q)^a_{ij} = t^a_{ij}$, $T(\bar{q})^a_{ij} = -t^a_{ji}$ and $T(\mathrm{g})^a_{bc} = i f_{bac}$, where q here represents any quark flavor and $\bar{q}$ represents any antiquark flavor.
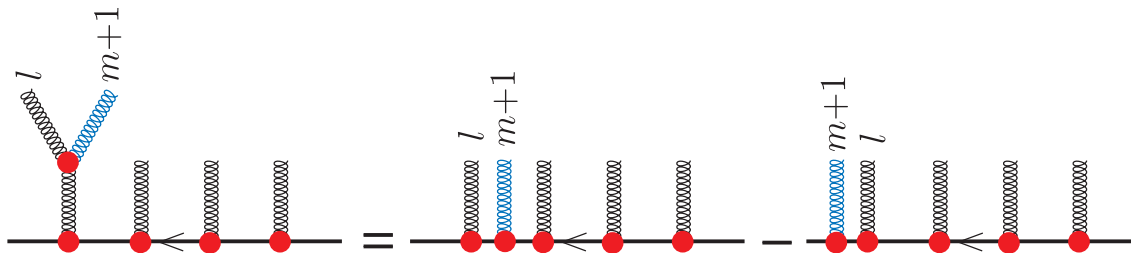
Figure 8: Illustration of eq. (7.24). A gluon l spits into gluon l and gluon $m+1$. The result is to place the new gluon $m+1$ either to the right of gluon l, with insertion operator $a_+^\dagger(l)$ or to the left, with insertion operator $a_-^\dagger(l)$. The normalization factors $\sqrt{C_F}$ are not indicated.

For a quark splitting to a quark plus a gluon we have a similar operator, one more generator matrix is inserted at the quark end of the string, so that

$$t_l^\dagger(q \to q + g) = \sqrt{C_F}\, a_+^\dagger(l) \ . \tag{7.25}$$

For an antiquark splitting to an antiquark plus a gluon, one more generator matrix is inserted at the antiquark end of the string, so that

$$t_l^\dagger(\bar{q} \to \bar{q} + g) = -\sqrt{C_F}\, a_-^\dagger(l) \ . \tag{7.26}$$

In these equations, q represents any quark flavor and $\bar{q}$ represents any antiquark flavor.

The operators for gluon emission obey an important identity. The matrix element M and our approximations to it are always color singlets,

$$\sum_l \sum_{\tilde{a}_l} T(f_l)_{a_l,\tilde{a}_l}^{a_{m+1}} \psi^{a_a,a_b,a_1,...,\tilde{a}_l,...,a_m} = 0 \ . \tag{7.27}$$

Thus

$$\sum_l t_l^\dagger(f_l \to f_l + g)|\psi\rangle = 0 \ . \tag{7.28}$$

We will use this identity in the following sections to rearrange the splitting formulas.

The last possibility for parton splitting is that of a gluon splitting into a quark-antiquark pair with flavors q and $\bar{q}$. The color splitting operator is defined by

$$|\hat{\psi}\rangle = t_l^\dagger(g \to q + \bar{q})|\psi\rangle \ , \tag{7.29}$$

where

$$\hat{\psi}^{a_a,a_b,a_1,...,a_m,a_{m+1}} = \sum_{\tilde{a}_l} t_{a_l,a_{m+1}}^{\tilde{a}_l} \psi^{a_a,a_b,a_1,...,\tilde{a}_l,...,a_m} \ . \tag{7.30}$$

Our convention here is that the quark inherits the gluon label, l, and the antiquark gets the new label, $m+1$. To analyze this, we use

$$t_{a_l,a_{m+1}}^{\tilde{a}_l}\, t_{c,d}^{\tilde{a}_l} = \frac{1}{2}\,\delta_{c,a_{m+1}}\delta_{a_l,d} - \frac{1}{2N_c}\,\delta_{a_l,a_{m+1}}\delta_{c,d} \ . \tag{7.31}$$

Thus $t_l(g \rightarrow q + \bar{q}) \big| \{c\}_m \rangle$ is a combination of two basis vectors,

$$\frac{1}{2} \left[\frac{n(\{\hat{c}_1\}_{m+1})}{n(\{c\}_m)} \right]^{1/2} \big| \{\hat{c}_1\}_{m+1} \rangle - \frac{1}{2N_\mathrm{c}} \left[\frac{n(\{\hat{c}_2\}_{m+1})}{n(\{c\}_m)} \right]^{1/2} \big| \{\hat{c}_2\}_{m+1} \rangle \ . \qquad (7.32)$$

In $\{\hat{c}_1\}_{m+1}$, the string in which gluon l resides is split. Specifically, if in $\{c\}_m$, gluon l appears in an open string as $[i_1, \ldots, i_A, l, i_B \ldots i_n]$, then in $\{\hat{c}_1\}_{m+1}$ this string is replaced by two strings, $[i_1, \ldots, i_A, m+1][l, i_B, \cdots, i_n]$. On the other hand, if in $\{c\}_m$, gluon l appears in a closed string as $(i_1, \ldots, i_A, l, i_B, \ldots i_n)$, then in $\{\hat{c}_1\}_{m+1}$ this string is replaced by an open string, $[l, i_B, \ldots, i_n, i_1 \ldots, i_A, m+1]$. In either case,

$$\frac{n(\{\hat{c}_1\}_{m+1})}{n(\{c\}_m)} = \frac{N_\mathrm{c}}{C_\mathrm{F}} \ . \qquad (7.33)$$

In the second color state in eq. (7.32), $\big| \{\hat{c}_2\}_{m+1} \rangle$, gluon l is simply removed from its string and a new trivial $q\bar{q}$ string is created. Specifically, if in $\{c\}_m$ gluon l appears in an open string as $[i_1, \ldots, i_A, l, i_B \ldots i_n]$, then in $\{\hat{c}_2\}_{m+1}$ this string is replaced by two strings, $[i_1, \ldots, i_A, i_B, \cdots, i_n]$ and $[l, m+1]$. On the other hand, if in $\{c\}_m$ gluon l appears in an closed string as $(i_1, \ldots, i_A, l, i_B, \ldots i_n)$, then in $\{\hat{c}_2\}_{m+1}$ this string is replaced by a closed string $(i_1, \ldots, i_A, i_B, \ldots i_n)$ and the open string $[l, m+1]$. In either case,

$$\frac{n(\{\hat{c}_2\}_{m+1})}{n(\{c\}_m)} = \frac{N_\mathrm{c}}{C_\mathrm{F}} \ . \qquad (7.34)$$

We can state this result using the operators defined in eqs. (7.17) and (7.18),

$$t_l^\dagger(\mathrm{g} \rightarrow q + \bar{q}) = \sqrt{\frac{N_\mathrm{c}}{4C_\mathrm{F}}} \, a_q^\dagger(l) - \sqrt{\frac{1}{4N_\mathrm{c}C_\mathrm{F}}} \, a_s^\dagger(l) \ . \qquad (7.35)$$

We can now see the advantage of this choice of color basis, beyond the fact that it is widely used for NLO calculations. In this basis, the description of parton splitting is very simple. For splittings $q \rightarrow qg$ and $\bar{q} \rightarrow \bar{q}g$, we simply add a gluon to a string. For splittings $\mathrm{g} \rightarrow \mathrm{gg}$, we replace one gluon on a string by two. There are two terms corresponding to the two possible orders for the two gluons on the string. For a splitting $\mathrm{g} \rightarrow \bar{q}q$, we split a string or we remove the gluon from its string and create a new $q\bar{q}$ string. The normalization factors may be considered to be just a matter of bookkeeping. However, one should note that the splittings $q \rightarrow qg$, $\bar{q} \rightarrow \bar{q}g$. and $\mathrm{g} \rightarrow \mathrm{gg}$ come with numerical factors $\sqrt{C_\mathrm{F}}$ in the amplitude. These factors are large for large N_c. The splitting $\mathrm{g} \rightarrow q\bar{q}$ comes with a numerical factor for the first term, $[N_\mathrm{c}/(4C_\mathrm{F})]^{1/2}$ that is *not* large in the large N_c limit. (The numerical factor for the second term is small in the large N_c limit.) Thus the color factor for the splitting $\mathrm{g} \rightarrow q\bar{q}$ is smaller than the color factors for the other splittings. This makes $\mathrm{g} \rightarrow q\bar{q}$ splitting somewhat disfavored even though there are several $q\bar{q}$ flavor choices available.

8. Evolution for the statistical states

We have seen how parton splitting works at the quantum amplitude level. We now need to use these results to formulate the effect of splitting on the density operator ρ and thus

the statistical state $|\rho)$. Suppose that we have the function

$$\left(\{p, f, s', c', s, c\}_m | \rho\right) \tag{8.1}$$

describing m final state partons plus the two initial state partons. After one splitting, we will have $m + 1$ final state partons, described by

$$\left(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1} | \rho\right) \ . \tag{8.2}$$

We need this function within the soft and collinear splitting approximations that we have used for the amplitudes.

Any of the partons can split. In particular, parton i in $\left|M(\{p, f\}_m)\right\rangle$ and parton j in $\left\langle M(\{p, f\}_m)\right|$ can split. In the simplest case (and the only case incorporated into typical parton shower Monte Carlo event generators), $i = j$. However, one can have $i \neq j$ and still get the same momenta and flavors $\{\hat{p}, \hat{f}\}_{m+1}$ from both splittings and still get a logarithmic divergence if one were to integrate over $\{\hat{p}\}_{m+1}$. This happens when parton i emits a gluon $m + 1$ in $\left|M(\{p, f\}_m)\right\rangle$ and parton j emits gluon $m + 1$ in $\left\langle M(\{p, f\}_m)\right|$. An interference graph of this sort has no collinear divergence from $\hat{p}_{m+1}$ being collinear to $\hat{p}_i$ or $\hat{p}_j$. However it does have a soft divergence from $\hat{p}_{m+1} \to 0$. For this reason, we need to include the case $i \neq j$.

For each splitting, we relate $\{\hat{p}, \hat{f}\}_{m+1}$ to the starting momenta and flavors $\{p, f\}_m$ and splitting variables $\{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}$ using one of the mappings

$$\{\hat{p}, \hat{f}\}_{m+1} = R_l(\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}) \ , \tag{8.3}$$

as specified in section 4. In the case $i \neq j$, we can use either the mapping with $l = i$ or the mapping with $l = j$.[25] We find it most useful to average over these two possibilities. We use the mapping with $l = i$ with a weight A_{ij} and the mapping with $l = j$ with weight A_{ji}. In general, the weights can depend on the momenta $\{\hat{p}\}_{m+1}$. In that case, the momentum mapping R_l with $l = i$ is used for A_{ij} and the momentum mapping R_l with $l = j$ is used for A_{ji} . In this paper, the default value is $A_{ij} = 1/2$, but our notation allows for other choices. In the case $i = j$, we use the mapping with $l = i = j$.

We are thus led to write the density operator $\hat{\rho}$ after splitting as a sum of contributions $\hat{\rho}_{ij}^{(l)}$, where the superscript indicates the treatment of the kinematics and the subscripts indicate which partons split,

$$
\begin{aligned}
\hat{\rho}(\{\hat{p}, \hat{f}\}_{m+1}) &= \sum_l \hat{\rho}_{ll}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) \\
&\quad + \sum_{\substack{i,j \\ i \neq j}} \left\{ A_{ij}(\{\hat{p}\}_{m+1}) \, \hat{\rho}_{ij}^{(i)}(\{\hat{p}, \hat{f}\}_{m+1}) + A_{ji}(\{\hat{p}\}_{m+1}) \, \hat{\rho}_{ij}^{(j)}(\{\hat{p}, \hat{f}\}_{m+1}) \right\} \\
&= \sum_l \Big\{ \hat{\rho}_{ll}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) \\
&\quad + \sum_{k \neq l} A_{lk}(\{\hat{p}\}_{m+1}) \left[\hat{\rho}_{lk}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) + \hat{\rho}_{kl}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) \right] \Big\} \ .
\end{aligned}
\tag{8.4}
$$

---

[25]Either mapping suffices because in either case the map becomes the identity map in the soft limit $\hat{p}_{m+1} \to 0$.

Here the sums run over the set $\{a, b, 1, \ldots, m\}$. We use eq. (6.4) for the contributions $\hat{\rho}_{ij}^{(l)}$ in the case $i = j$,

$$\begin{aligned}
\rho_{ll}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) &= t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1}) \, V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1}) \, \rho(\{p, f\}_m) \\
&\quad \times V_l(\{\hat{p}, \hat{f}\}_{m+1}) \, t_l(f_l \to \hat{f}_l + \hat{f}_{m+1}) \\
&\quad \times S_l(\{\hat{f}\}_{m+1}) \ .
\end{aligned} \tag{8.5}$$

Here $S_l(\{\hat{f}\}_{m+1})$ is a counting factor that is determined by the conventions we have used to label the final state partons,

$$S_l(\{\hat{f}\}_{m+1}) = \begin{cases} 1/2 \, , & l \in \{1, \ldots, m\}, \ \hat{f}_l = \hat{f}_{m+1} = g \\ 1 \, , & l \in \{1, \ldots, m\}, \ \hat{f}_l \neq g, \hat{f}_{m+1} = g \\ 0 \, , & l \in \{1, \ldots, m\}, \ \hat{f}_l = g, \hat{f}_{m+1} \neq g \\ 1 \, , & l \in \{1, \ldots, m\}, \ \hat{f}_l = q, \hat{f}_{m+1} = \bar{q} \\ 0 \, , & l \in \{1, \ldots, m\}, \ \hat{f}_l = \bar{q}, \hat{f}_{m+1} = q \\ 1 \, , & l \in \{a, b\} \end{cases} \tag{8.6}$$

This factor is zero for $\hat{f}_l = g$ and $\hat{f}_{m+1} = q$ or $\hat{f}_{m+1} = \bar{q}$, even though flavor conservation allows this combination in a $q \to q + g$ or $\bar{q} \to \bar{q} + g$ splitting, because we have chosen to label the final daughter gluon in these cases as parton $m + 1$. Similarly, this factor is zero for $\hat{f}_l = \bar{q}$ and $\hat{f}_{m+1} = q$, even though flavor conservation allows this combination in a $g \to q + \bar{q}$ splitting, because we have chosen to label the final daughter antiquark in this case as parton $m + 1$. In the case of a final state $g \to g + g$ splitting on line l, the splitting probability is symmetric under interchange of the labels l and $m + 1$, so that integrating over $\hat{p}_l$ and $\hat{p}_{m+1}$ would count the same physical configuration twice. The factor $1/2$ corrects for this symmetry. This issue is discussed in some detail in appendix B.

For $i \neq j$, we can use the simpler splitting operator V^{soft} as in eq. (6.33),

$$\begin{aligned}
\hat{\rho}_{ij}^{(l)}(\{\hat{p}, \hat{f}\}_{m+1}) &= t_i^\dagger(f_i \to \hat{f}_i + \hat{f}_{m+1}) \, V_i^{\dagger,\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \, \rho(\{p, f\}_m) \\
&\quad \times V_j^{\text{soft}}(\{\hat{p}, \hat{f}\}_{m+1}) \, t_j(f_j \to \hat{f}_j + \hat{f}_{m+1}) \ .
\end{aligned} \tag{8.7}$$

Note that $V_i^{\dagger,\text{soft}}$ and V_j^{soft} both vanish if parton $m + 1$ is not a gluon. Thus $\hat{\rho}_{ij}^{(l)} = 0$ for $i \neq j$ unless parton $m + 1$ is a gluon.

This definition for $\hat{\rho}_{ij}^{(l)}$ defines a mapping of the statistical states in which $|\rho)$ becomes

$$\big| \hat{\rho}_{ij}^{(l)} \big) = \mathcal{S}_{ij}^{(l)} |\rho) \ , \tag{8.8}$$

after splitting. Then

$$\begin{aligned}
\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1} \big| \mathcal{S}_{ij}^{(l)} |\rho\big) &= \frac{1}{m!} \int \big[d\{p, f, s', c', s, c\}_m \big] \\
&\quad \times \big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1} \big| \mathcal{S}_{ij}^{(l)} \big| \{p, f, s', c', s, c\}_m \big) \\
&\quad \times \big(\{p, f, s', c', s, c\}_m |\rho \big) \ .
\end{aligned} \tag{8.9}$$

The splitting operator $\mathcal{S}_{ij}^{(l)}$ is defined by giving its matrix elements,

$$
\begin{aligned}
\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{S}_{ij}^{(l)}\big|\{p, f, s', c', s, c\}_m\big) &= \\
&\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(i, j; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big) \\
&\times \big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(i, j; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big) \\
&\times \big(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\big) \\
&\times (m+1) \frac{n_{\mathrm{c}}(a) n_{\mathrm{c}}(b)\, \eta_{\mathrm{a}} \eta_{\mathrm{b}}}{n_{\mathrm{c}}(\hat{a}) n_{\mathrm{c}}(\hat{b})\, \hat{\eta}_{\mathrm{a}} \hat{\eta}_{\mathrm{b}}} \frac{f_{\hat{a}/A}(\hat{\eta}_{\mathrm{a}}, \mu_F^2) f_{\hat{b}/B}(\hat{\eta}_{\mathrm{b}}, \mu_F^2)}{f_{a/A}(\eta_{\mathrm{a}}, \mu_F^2) f_{b/B}(\eta_{\mathrm{b}}, \mu_F^2)} \ .
\end{aligned}
\tag{8.10}
$$

We discuss each factor in turn.

The first factor, $\mathcal{G}(i, j; \{\hat{f}\}_{m+1})$, describes how the splittings change the parton colors,

$$
\begin{aligned}
\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(i, j; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big) &= {}_D\langle\{\hat{c}\}_{m+1}\big|t_i^\dagger(f_i \to \hat{f}_i + \hat{f}_{m+1})\big|\{c\}_m\rangle \\
&\times \langle\{c'\}_m\big|t_j(f_j \to \hat{f}_j + \hat{f}_{m+1})\big|\{\hat{c}'\}_{m+1}\rangle_D \ .
\end{aligned}
\tag{8.11}
$$

In the next factor, $\mathcal{W}(i, j; \{\hat{f}, \hat{p}\}_{m+1})$, contains the spin-dependent splitting functions. The simplest case occurs for $i = j$ with anything other than a final state g $\to$ g+g splitting. For that case

$$
\begin{aligned}
\big(\{\hat{s}', \hat{s}\}_{m+1}\big|&\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big) \\
&= S_l(\{\hat{f}\}_{m+1}) \langle\{\hat{s}\}_{m+1}\big|V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1})\big|\{s\}_m\rangle\langle\{s'\}_m\big|V_l(\{\hat{p}, \hat{f}\}_{m+1})\big|\{\hat{s}'\}_{m+1}\rangle \ .
\end{aligned}
$$

For $i = j$ in general we define

$$
\begin{aligned}
\big(\{\hat{s}', \hat{s}\}_{m+1}\big|&\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big) \\
&= S_l(\{\hat{f}\}_{m+1})\bigg\{ \langle\{\hat{s}\}_{m+1}\big|V_l^\dagger(\{\hat{p}, \hat{f}\}_{m+1})\big|\{s\}_m\rangle\langle\{s'\}_m\big|V_l(\{\hat{p}, \hat{f}\}_{m+1})\big|\{\hat{s}'\}_{m+1}\rangle \\
&\quad + \theta(l \in \{1, \ldots, m\}, \hat{f}_l = \hat{f}_{m+1} = \mathrm{g})\,\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\widetilde{\mathcal{W}}(l, l; \{\hat{p}\}_{m+1})\big|\{s', s\}_m\big) \bigg\} \ .
\end{aligned}
\tag{8.12}
$$

The second term is included in the special case of a final state g $\to$ g + g splitting. In this case, the first term is symmetric under an interchange $l \leftrightarrow m + 1$ between the two final state gluons. It gives a leading singularity when gluon $m + 1$ is soft but also gives a leading singularity when gluon l is soft. We seek to use the freedom to assign labels to ensure that gluon $m + 1$ can be soft but not gluon l. The operator $\widetilde{\mathcal{W}}$ in the added term is antisymmetric under $l \leftrightarrow m + 1$, so that it gives zero contribution after integration over the final state momenta. The simplest choice for $\widetilde{\mathcal{W}}$ would be the first term times $[2\,\theta(\hat{p}_{m+1} \cdot Q < \hat{p}_l \cdot Q) - 1]$ which is antisymmetric under $l \leftrightarrow m+1$. Then the total $\mathcal{W}$ would be just the first term times $2\,\theta(\hat{p}_{m+1} \cdot Q < \hat{p}_l \cdot Q)$. Clearly this eliminates the singularity when gluon l is soft. However, we adopt a slightly more subtle procedure. Write the tensor that defines the three gluon vertex as $v = v_1 + v_2 + v_3$,

$$
\begin{aligned}
v_1^{\alpha\beta\gamma}(p_a, p_b, p_c) &= g^{\alpha\beta}(p_a - p_b)^\gamma \ , \\
v_2^{\alpha\beta\gamma}(p_a, p_b, p_c) &= g^{\beta\gamma}(p_b - p_c)^\alpha \ , \\
v_3^{\alpha\beta\gamma}(p_a, p_b, p_c) &= g^{\gamma\alpha}(p_c - p_a)^\beta \ .
\end{aligned}
\tag{8.13}
$$

Then we can define partial vertex functions analogous to those in eq. (6.31) by

$$v_{J,l}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) = \frac{\sqrt{4\pi\alpha_{\rm s}}}{2\hat{p}_{m+1} \cdot \hat{p}_l} \, \varepsilon_\alpha(\hat{p}_{m+1}, \hat{s}_{m+1}; \hat{Q})^* \varepsilon_\beta(\hat{p}_l, \hat{s}_l; \hat{Q})^* \varepsilon^\nu(p_l, s_l; \hat{Q})$$
$$\times v_J^{\alpha\beta\gamma}(\hat{p}_{m+1}, \hat{p}_l, -\hat{p}_{m+1} - \hat{p}_l) \, D_{\gamma\nu}(\hat{p}_l + \hat{p}_{m+1}; n_l). \quad (8.14)$$

for $J \in \{1, 2, 3\}$. Finally we define partial operators V by the analogue of eq. (6.6),

$$\langle\{\hat{s}\}_{m+1}|V_{J,l}^\dagger(\{\hat{p}, \hat{f}\}_{m+1})|\{s\}_m\rangle = \left(\prod_{j\notin\{l,m+1\}} \delta_{\hat{s}_j, s_j}\right) v_{J,l}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) \ . \quad (8.15)$$

Now the first term in $\mathcal{W}$ comes from $\sum_{J,K=1}^3 v_J \times v_K^*$. The term with a leading singularity when gluon $m + 1$ is soft is $v_2 \times v_2^*$. The term with a leading singularity when gluon l is soft is $v_3 \times v_3^*$. Therefore we need to get rid of $v_3 \times v_3^*$ and double $v_2 \times v_2^*$ to make up for the factor $1/2$ in S_l. We thus define

$$\left(\{\hat{s}', \hat{s}\}_{m+1}\big|\widetilde{\mathcal{W}}(l, l; \{\hat{p}\}_{m+1})\big|\{s', s\}_m\right) =$$
$$\langle\{\hat{s}\}_{m+1}|V_{2,l}^\dagger(\{\hat{p}, \hat{f}\}_{m+1})|\{s\}_m\rangle\langle\{s'\}_m|V_{2,l}(\{\hat{p}, \hat{f}\}_{m+1})|\{\hat{s}'\}_{m+1}\rangle$$
$$- \langle\{\hat{s}\}_{m+1}|V_{3,l}^\dagger(\{\hat{p}, \hat{f}\}_{m+1})|\{s\}_m\rangle\langle\{s'\}_m|V_{3,l}(\{\hat{p}, \hat{f}\}_{m+1})|\{\hat{s}'\}_{m+1}\rangle \ . \quad (8.16)$$

This is computationally very simple even if it takes some time to explain.

For $i \neq j$, we use the simpler splitting operators $V^{\rm soft}$,

$$\left(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(i, j; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\right) = \langle\{\hat{s}\}_{m+1}|V_i^{\dagger,{\rm soft}}(\{\hat{p}, \hat{f}\}_{m+1})|\{s\}_m\rangle$$
$$\times \langle\{s'\}_m|V_j^{\rm soft}(\{\hat{p}, \hat{f}\}_{m+1})|\{\hat{s}'\}_{m+1}\rangle \ . \quad (8.17)$$

In the next factor, the matrix element $\left(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\right)$ is a delta function that enforces the requirement that the momenta and flavors $\{p, f\}_m$ are related to the momenta and flavors after the splitting by the mapping Q_l. The definition is

$$\frac{1}{m!} \int [d\{p, f\}_m] \left(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\right) h(\{p, f\}_m) = h(\{p', f'\}_m) \ , \quad (8.18)$$

where $h(\{p, f\}_m)$ is an arbitrary test function and

$$\{\{p', f'\}_m, \{\zeta_{\rm p}, \zeta_{\rm f}\}\} = Q_l(\{\hat{p}, \hat{f}\}_{m+1}) \quad (8.19)$$

as specified in section 4. Another useful identity for $\mathcal{P}_l$ is[26]

$$\frac{1}{(m+1)!} \int [d\{\hat{p}, \hat{f}\}_{m+1}] \, g(\{\hat{p}, \hat{f}\}_{m+1}) \left(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\right)$$
$$= \frac{1}{(m+1)} \sum_{\zeta_{\rm f}\in\Phi_l(f_l)} \int d\zeta_{\rm p} \, \theta(\zeta_{\rm p} \in \Gamma_l(\{p\}_m, \zeta_{\rm f})) \, g(\{\hat{p}', \hat{f}'\}_{m+1})) \ , \quad (8.20)$$

---

[26]To derive eq. (8.20) from eq. (8.18), we add one more integration to eq. (8.18), in the form $\int [d\{\hat{p}, \hat{f}\}_{m+1}] \, g(\{\hat{p}, \hat{f}\}_{m+1})$. Then on the right hand side, we change integration variables from $\{p', f'\}_{m+1}$ to $\{\{p, f\}_m, \{\zeta_{\rm p}, \zeta_{\rm f}\}\}$ according to eqs. (4.37) and (4.70). This gives a result that is equivalent to eq. (8.20).

where $g\big(\{\hat{p}, \hat{f}\}_{m+1}\big)$ is an arbitrary test function and

$$\{\hat{p}', \hat{f}'\}_{m+1} = R_l(\{\{p, f\}_m, \{\zeta_{\mathrm{p}}, \zeta_{\mathrm{f}}\}\}) \tag{8.21}$$

is the inverse transformation to Q_l, as specified in section 4. In this form, we display an integration over the splitting variables that would occur in an implementation of this formalism as a computer program.

The counting factor $(m + 1)$ is the ratio of the factor $(m + 1)!$ in the normalization integral for $\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{S}_{ij}^{(l)}\big|\rho\big)$ to the factor $m!$ in the normalization integral for $\big(\{p, f, s', c', s, c\}_m\big|\rho\big)$. This factor is derived in appendix B. The factor with parton distributions comes from eqs. (3.17) and (3.18). In the case of a final state splitting, this factor is 1.

We can now assemble our result. From eq. (8.4) we have

$$|\hat{\rho}) = \mathcal{S}|\rho) \ , \tag{8.22}$$

where the total splitting operator is

$$\mathcal{S} = \sum_l \mathcal{S}_l \ , \quad \mathcal{S}_l = \mathcal{S}_{ll}^{(l)} + \sum_{k \neq l} \mathcal{A}_{lk} \left\{ \mathcal{S}_{lk}^{(l)} + \mathcal{S}_{kl}^{(l)} \right\} \ . \tag{8.23}$$

Here the sums run over the set $\{\mathrm{a}, \mathrm{b}, 1, \dots, m\}$ and $\mathcal{A}_{lk}$ is the operator on the space of statistical states that multiplies a basis vector $\big|\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big)$ by the corresponding function $A_{lk}(\{\hat{p}\}_{m+1})$.

We manipulate the result a bit. Because the quantum amplitudes are color singlets, as reflected in eq. (7.28), when $\hat{f}_{m+1} = \mathrm{g}$ the gluon emission operators $\mathcal{G}(l, k; \{\hat{f}\}_{m+1})$ obey

$$\sum_k \mathcal{G}(l, k; \{\hat{f}\}_{m+1}) = 0 \qquad , \qquad \sum_k \mathcal{G}(k, l; \{\hat{f}\}_{m+1}) = 0 \ . \tag{8.24}$$

Then the $\mathcal{G}(l, k; \{\hat{f}\}_{m+1})$ are not independent. For this reason, there can be color coherence cancellations that are always present but are not evident if we use all of the possible $\mathcal{G}(l, k; \{\hat{f}\}_{m+1})$ operators. Accordingly, when $\hat{f}_{m+1} = \mathrm{g}$, we eliminate $\mathcal{G}(l, l; \{\hat{f}\}_{m+1})$ with the replacement

$$\mathcal{G}(l, l; \{\hat{f}\}_{m+1}) = -\frac{1}{2} \sum_{k \neq l} \mathcal{G}(l, k; \{\hat{f}\}_{m+1}) - \frac{1}{2} \sum_{k \neq l} \mathcal{G}(k, l; \{\hat{f}\}_{m+1}) \ . \tag{8.25}$$

We will see in section 11 how this allows cancellations to occur at a low level of a calculation.

With this replacement, the total splitting operator is given for the case $\hat{f}_{m+1} = \mathrm{g}$ by

$$
\begin{aligned}
&\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{S}_l\big|\{p, f, s', c', s, c\}_m\big)\\
&= (m+1)\big(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\big)\frac{n_\mathrm{c}(a)n_\mathrm{c}(b)\,\eta_\mathrm{a}\eta_\mathrm{b}}{n_\mathrm{c}(\hat{a})n_\mathrm{c}(\hat{b})\,\hat{\eta}_\mathrm{a}\hat{\eta}_\mathrm{b}}\frac{f_{\hat{a}/A}(\hat{\eta}_\mathrm{a}, \mu_F^2)f_{\hat{b}/B}(\hat{\eta}_\mathrm{b}, \mu_F^2)}{f_{a/A}(\eta_\mathrm{a}, \mu_F^2)f_{b/B}(\eta_\mathrm{b}, \mu_F^2)}\\
&\quad\times \sum_{\substack{k\in\{\mathrm{a},\mathrm{b},1,\ldots,m\}\\ k\neq l}}\Bigg\{\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(l, k; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)\\
&\qquad\times\bigg[A_{lk}(\{\hat{p}\}_{m+1})\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, k; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)\\
&\qquad\quad -\frac{1}{2}\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)\bigg]\\
&\qquad+\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(k, l; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)\\
&\qquad\times\bigg[A_{lk}(\{\hat{p}\}_{m+1})\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(k, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)\\
&\qquad\quad -\frac{1}{2}\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)\bigg]\Bigg\}\ .
\end{aligned}
\tag{8.26}
$$

In the case that $\hat{f}_{m+1} \neq \mathrm{g}$, that is $\{\hat{f}_l, \hat{f}_{m+1}\} = \{q, \bar{q}\}$, this becomes

$$
\begin{aligned}
\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{S}_l\big|\{p, f, s', c', s, c\}_m\big) &= (m+1)\big(\{\hat{p}, \hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p, f\}_m\big)\\
&\quad\times\frac{n_\mathrm{c}(a)n_\mathrm{c}(b)\,\eta_\mathrm{a}\eta_\mathrm{b}}{n_\mathrm{c}(\hat{a})n_\mathrm{c}(\hat{b})\,\hat{\eta}_\mathrm{a}\hat{\eta}_\mathrm{b}}\frac{f_{\hat{a}/A}(\hat{\eta}_\mathrm{a}, \mu_F^2)f_{\hat{b}/B}(\hat{\eta}_\mathrm{b}, \mu_F^2)}{f_{a/A}(\eta_\mathrm{a}, \mu_F^2)f_{b/B}(\eta_\mathrm{b}, \mu_F^2)}\\
&\quad\times\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(l, l; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)\\
&\quad\times\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)\ .
\end{aligned}
\tag{8.27}
$$

In the following section, we will use $\mathcal{S}$ to define the splitting operator at shower time t, $\mathcal{H}_\mathrm{I}(t)$. Then in section 10 we will study the evolution of the color structure that is expressed in eqs. (8.26) and (8.27). In section 11, we will examine the how quantum coherence for soft gluon emission is contained in eq. (8.26).

9. The operator $\mathcal{H}_\mathrm{I}(t)$

It remains to define the operator $\mathcal{H}_\mathrm{I}(t)$. The integral of this operator over t gives the total probability for a splitting at any scale,

$$
\int_0^\infty dt\ \mathcal{H}_\mathrm{I}(t) = \mathcal{S}\ .
\tag{9.1}
$$

To get the probability for a splitting at the scale corresponding to shower time t, we simply need to insert a delta function that defines t. Our default choice is

$$
t = \log\left(\frac{Q_0^2}{|(\hat{p}_l + (-1)^{\delta_{l,\mathrm{a}}+\delta_{l,\mathrm{b}}}\hat{p}_{m+1})^2 - m^2(f_l)|}\right)\ .
\tag{9.2}
$$

Here the virtuality is defined using $\hat{p}_l + \hat{p}_{m+1}$ for a final state splitting and $\hat{p}_l - \hat{p}_{m+1}$ for an initial state splitting. We take Q_0^2 to be the hardness scale of the initial hard scattering that starts the parton shower, so that the initial value of t is zero. One could take Q_0^2 to be the minimum of the values $2p_i \cdot p_j$ for final state particles i, j from the initial hard scattering. Here we can neglect all quark masses compared to Q_0. Other definitions of t are possible. For instance, many authors use a measure of the transverse momentum in a parton splitting.

When we use the definition (9.2) of t, we obtain the corresponding definition of $\mathcal{H}_\mathrm{I}(t)$.

$$
\begin{aligned}
\big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{H}_\mathrm{I}(t)\big|\{p, f, s', c', s, c\}_m\big) \\
= \sum_{l \in \{\mathrm{a},\mathrm{b},1,\ldots,m\}} \big(\{\hat{p}, \hat{f}, \hat{s}', \hat{c}', \hat{s}, \hat{c}\}_{m+1}\big|\mathcal{S}_l\big|\{p, f, s', c', s, c\}_m\big) \\
\times \delta\left(t - \log\left(\frac{Q_0^2}{|(\hat{p}_l + (-1)^{\delta_{l,\mathrm{a}}+\delta_{l,\mathrm{b}}}\hat{p}_{m+1})^2 - m^2(f_l)|}\right)\right) \ ,
\end{aligned}
\tag{9.3}
$$

where $\mathcal{S}_l$ is defined in eqs. (8.26) and (8.27). In these equations there are parton distribution functions evaluated at a factorization scale μ_F, which we define according to eq. (3.38) in terms of t. In addition, there is a factor α_s, which needs to be evaluated at a scale μ_R that has not been made explicit in the notation. The argument presented in section (3.4) indicates that the momentum scale at the splitting is just the resolution scale that we use for μ_F. Thus we take

$$
\mu_\mathrm{F}^2 = \mu_\mathrm{R}^2 = Q_0^2\, e^{-t} \ .
\tag{9.4}
$$

10. Color evolution of the statistical states

We are now in a position to say something about the color structure of the statistical states and the evolution of this structure. Notice that in general $\big(\{p, f, s', c', s, c\}_m\big|\rho\big)$ can be non-zero for $\{c'\}_m \neq \{c\}_m$. Even if we start with $\{c'\}_n = \{c\}_n$, at some early stage of evolution, splitting will generate states with $\{c'\}_m \neq \{c\}_m$ at later stages. However, in the end we measure something that is color independent. This means taking the color trace of ρ. That is, we multiply $\big(\{p, f, s', c', s, c\}_m\big|\rho\big)$ by $\langle\{c'\}_m|\{c\}_m\rangle$ and sum over the color configurations $\{c'\}_m$ and $\{c\}_m$, as in eq. (3.34). Recall that when $\{c'\}_m \neq \{c\}_m$, the inner product $\langle\{c'\}_m|\{c\}_m\rangle$ is of order $1/N_\mathrm{c}^2$ to some power and is thus small in the large N_c^2 limit and numerically small for $N_\mathrm{c} = 3$. Thus configurations with $\{c'\}_m \neq \{c\}_m$ are not very important. Furthermore, if we start with a state with $\{c'\}_m \neq \{c\}_m$, splitting cannot generate $\{\hat{c}'\}_{m+1} = \{\hat{c}\}_{m+1}$ at the next stage. Thus one can speak of a "leading color approximation" in which contributions with $\{c'\}_m \neq \{c\}_m$ are always dropped. This is what happens in most parton shower Monte Carlo programs. We do not drop terms, but we should understand what happens in the leading color approximation.

In order to understand splitting in the leading color configuration, suppose that we apply $\mathcal{H}_\mathrm{I}(t)$ to a state $\big|\{p, f, s', c', s, c\}_m\big)$ in which $\{c'\} = \{c\}$. Thus we consider the action
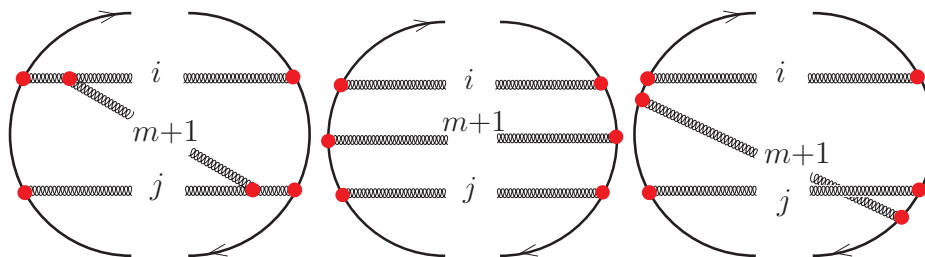
Figure 9: Illustration of eq. (10.2). The first diagram shows the color configuration corresponding to the operator $\mathcal{G}(i, j; \{\hat{f}\}_{m+1})$: the gluon's color is emitted from parton i and absorbed on gluon j. Using eq. (7.24) to write this in terms of color basis states, one contribution is shown in the middle diagram. This is the leading color contribution. Another contribution is shown in the right-hand diagram. This contribution is subleading. If i and j had not been color connected to each other, there would have been no contribution that survives in the leading color approximation.

of the color splitting operator $\mathcal{G}(i, j; \{\hat{f}\}_{m+1})$, on a state $\left|\{c', c\}_m\right)$ with $\{c'\} = \{c\}$,

$$
\begin{aligned}
\left(\{\hat{c}', \hat{c}\}_{m+1}\middle|\mathcal{G}(i, j; \{\hat{f}\}_{m+1})\middle|\{c, c\}_m\right) = {}_D\!\left\langle\{\hat{c}\}_{m+1}\middle|t_i^\dagger(f_i \to \hat{f}_i + \hat{f}_{m+1})\middle|\{c\}_m\right\rangle \\
\times {}_D\!\left\langle\{\hat{c}'\}_{m+1}\middle|t_j^\dagger(f_j \to \hat{f}_j + \hat{f}_{m+1})\middle|\{c\}_m\right\rangle^* .
\end{aligned}
\tag{10.1}
$$

We first consider gluon emission. That is $\hat{f}_{m+1} = \mathrm{g}$. Then, after using eq. (8.25), all of the contributions to $\mathcal{H}_\mathrm{I}(t)|\rho)$ come from $\mathcal{G}(i, j; \{\hat{f}\}_{m+1})$ with $i \neq j$. In order to have a definite case in mind, let us suppose that i and j are gluons. Then, using eq. (7.24) to represent the $t^\dagger$ operators in terms of the operators $a^\dagger$ that insert the added gluon in particular places in the string basis states, our matrix element becomes

$$
\begin{aligned}
&C_\mathrm{F}\left\langle\{\hat{c}\}_{m+1}\middle|a_+^\dagger(i)\middle|\{c\}_m\right\rangle \left\langle\{\hat{c}'\}_{m+1}\middle|a_+^\dagger(j)\middle|\{c\}_m\right\rangle^* + \\
&C_\mathrm{F}\left\langle\{\hat{c}\}_{m+1}\middle|a_-^\dagger(i)\middle|\{c\}_m\right\rangle \left\langle\{\hat{c}'\}_{m+1}\middle|a_-^\dagger(j)\middle|\{c\}_m\right\rangle^* - \\
&C_\mathrm{F}\left\langle\{\hat{c}\}_{m+1}\middle|a_+^\dagger(i)\middle|\{c\}_m\right\rangle \left\langle\{\hat{c}'\}_{m+1}\middle|a_-^\dagger(j)\middle|\{c\}_m\right\rangle^* - \\
&C_\mathrm{F}\left\langle\{\hat{c}\}_{m+1}\middle|a_-^\dagger(i)\middle|\{c\}_m\right\rangle \left\langle\{\hat{c}'\}_{m+1}\middle|a_+^\dagger(j)\middle|\{c\}_m\right\rangle^* .
\end{aligned}
\tag{10.2}
$$

Consider the case that gluon j is just to the right of gluon i along a color string. Then if the new gluon is inserted to the right of gluon i, we get the same state as when the new gluon is inserted to the left of gluon j. That is, the third term, with $a_+^\dagger(i)$ and $a_-^\dagger(j)$, gives a non-zero matrix element when $\{\hat{c}'\}_{m+1} = \{\hat{c}\}_{m+1}$. This contribution is thus kept in the leading color approximation. The other three contributions would be thrown away in this approximation. This is illustrated in figure 9.

Similarly, if gluon j is just to the left of gluon i along a color string, then the contribution with $a_-^\dagger(i)$ and $a_+^\dagger(j)$ makes a leading color contribution, while the other three contributions would be thrown away in the leading color approximation.

Suppose now that gluon j is *not* next to gluon i along a color string. Then all four terms would be thrown away in the leading color approximation.

We can summarize this by saying that there is a term that is kept in the leading color approximation when gluons i and j are "color connected": next to each other along a color
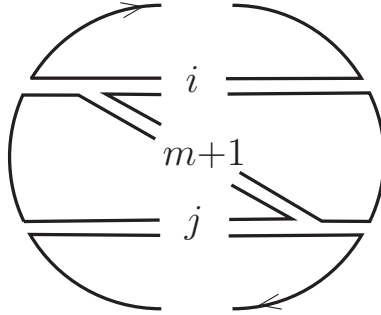
Figure 10: The starting diagram from figure 9 in the leading color approximation, in which gluons are emitted from color dipoles. The i-j color dipole is a unit that emits a gluon and absorbs it on the other side of the diagram. At the next stage there will be an i-$(m + 1)$ color dipole and an $(m + 1)$-j color dipole.

string. An analogous analysis leads to the same conclusion if one or both of partons i and j are quarks.

One may say that in the leading color approximation, the operating units are color dipoles, consisting of partons that are next to each other along a single color string. Within this approximation, a gluon may be considered to carry the $\mathbf{3} \times \bar{\mathbf{3}}$ representation of SU(3) instead of the $\mathbf{8}$ representation. Then, for instance, the $\bar{\mathbf{3}}$ half of a gluon forms a dipole with the $\mathbf{3}$ part of the neighboring gluon or with a neighboring quark. A dipole can emit a gluon. But there is no interference between diagrams in which the gluon is emitted by different dipoles. This is illustrated in figure 10. The color dipole picture was introduced as the basis of the parton shower program ARIADNE [28].

With the formalism presented in this paper, one can easily implement this approximation, but one can also keep corrections to it.

We have considered gluon emission. The case of a g $\rightarrow q+\bar{q}$ splitting is rather different. Then the contributing terms are from $\mathcal{G}(l, l; \{\hat{f}\}_{m+1})$. For the operators $t_l^\dagger(\mathrm{g} \rightarrow q + \bar{q})$ in eq. (10.1), we use eq. (7.35). In the leading term, with operator $a_q^\dagger(l)$, the color string containing the gluon breaks at the position of the gluon, with the creation of two new string ends. There is a second term, with operator $a_s^\dagger(l)$, in which gluon l simply disappears from its string and a new q-$\bar{q}$ string is created. However, the coefficient of $a_s^\dagger(l)$ is small in the $N_\mathrm{c} \rightarrow \infty$ limit and would be dropped in the leading color approximation. In fact, the coefficient of the leading $a_q^\dagger(l)$ is finite in the $N_\mathrm{c} \rightarrow \infty$ limit, instead of growing with N_c. Thus g $\rightarrow q + \bar{q}$ splitting is color suppressed compared to g $\rightarrow$ g + g splitting. Some of the color suppression is cancelled by the number of available $q\bar{q}$ flavors. Parton shower Monte Carlo programs normally include the leading term in $g \rightarrow q + \bar{q}$ splitting. With the formalism presented in this paper, we can keep both terms in g $\rightarrow q + \bar{q}$ splitting.

11. Soft gluon coherence

When a soft gluon is emitted from the partons involved in a hard scattering, the angular distribution of the emitted gluon reflects the structure of the emitting partons as a whole.

In particular, a soft gluon emitted from a pair of colored partons at an angle that is bigger than the angle between the partons effectively sees just one parton with the total color charge of the pair. Since the splitting operator $\mathcal{H}_{\mathrm{I}}(t)$ coherently sums the leading soft gluon singularities for emissions from a hard partonic system, including all of the interference diagrams, soft gluon coherence is automatically included. It is of interest to see how this happens and what form it takes in the color basis used in this paper.

Consider the matrix element of the splitting operator $\mathcal{S}_l$ for the case $\hat{f}_{m+1} = \mathrm{g}$, as given in Eq. (8.26) with the default choice $A_{lk} = 1/2$. There is a sum over indices $k \in \{\mathrm{a}, \mathrm{b}, 1, \ldots, m\}$ with $k \neq l$. There are two terms with different color factors, $\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(l, k; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)$ and $\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(k, l; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)$. These have analogous structures, so it suffices to analyze one of them. The coefficient of the color factor $-\big(\{\hat{c}', \hat{c}\}_{m+1}\big|\mathcal{G}(l, k; \{\hat{f}\}_{m+1})\big|\{c', c\}_m\big)$ is

$$\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}_{\mathrm{tot}}(l, k; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big) \equiv \big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big)$$
$$- \big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}(l, k; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big).$$
(11.1)

The first term is from the square of the amplitude for emitting the gluon from parton l. The second term represents the interference between the emission of gluon $m + 1$ from parton l and the emission of this gluon from parton k. Let us evaluate this in the soft-gluon approximation, $V \to V^{\mathrm{soft}}$, as defined in eqs. (6.35) and (6.36). We can also approximate $\hat{p}_l \sim p_l$ and $\hat{p}_k \sim p_k$. With these replacements, we have

$$\big(\{\hat{s}', \hat{s}\}_{m+1}\big|\mathcal{W}_{\mathrm{tot}}(l, k; \{\hat{f}, \hat{p}\}_{m+1})\big|\{s', s\}_m\big) \sim \left(\prod_{j \neq m+1} \delta_{\hat{s}_j, s_j} \delta_{\hat{s}'_j, s'_j} \right)(4\pi\alpha_{\mathrm{s}})\frac{\varepsilon(\hat{p}_{m+1}, \hat{s}_{m+1}; Q)^* \cdot p_l}{\hat{p}_{m+1} \cdot p_l}$$
$$\times \left\{ \frac{\varepsilon(\hat{p}_{m+1}, \hat{s}_{m+1}; Q) \cdot p_l}{\hat{p}_{m+1} \cdot p_l} - \frac{\varepsilon(\hat{p}_{m+1}, \hat{s}_{m+1}; Q) \cdot p_k}{\hat{p}_{m+1} \cdot p_k} \right\}.$$
(11.2)

There are singularities when $\hat{p}_{m+1}$ is collinear with p_l and when $\hat{p}_{m+1}$ collinear with p_k. The singularity for $\hat{p}_{m+1}$ collinear with p_k is not strong enough to give a logarithmic divergence when we integrate over the direction of $\hat{p}_{m+1}$. The singularity for $\hat{p}_{m+1}$ collinear with p_l *is* strong enough to give a logarithmic divergence. Here, the leading singularity comes from the first term, while the second term, representing the interference graph, gives only an integrable collinear singularity.

When $\hat{p}_{m+1}$ is not close to being collinear with p_l or p_k, both terms are important. However, suppose that the angle between p_l and p_k is small and that $\hat{p}_{m+1}$ makes an angle with either of them that is substantially greater than this angle. Then the two contributions approximately cancel each other. That is, soft gluon radiation associated with this color factor is approximately confined to a cone about the directions of partons l and k with opening angle on the order of the angle between p_l and p_k.

Thus there is approximate angular ordering of soft gluon emissions. One can also make an exact statement about angular ordering [5]. If we sum over the spins of the soft gluon,

we can use

$$
\sum_{\hat{s}_{m+1}} \varepsilon^{\mu}(\hat{p}_{m+1}, \hat{s}_{m+1}; Q) \varepsilon^{\nu}(\hat{p}_{m+1}, \hat{s}_{m+1}; Q)^* = -g^{\mu\nu} + \frac{\hat{p}_{m+1}^{\mu} Q^{\nu} + Q^{\mu} \hat{p}_{m+1}^{\nu}}{\hat{p}_{m+1} \cdot Q} - \frac{Q^2 \, \hat{p}_{m+1}^{\mu} \hat{p}_{m+1}^{\nu}}{(\hat{p}_{m+1} \cdot Q)^2} \ .
\tag{11.3}
$$

Then we can integrate over the azimuthal angle ϕ specifying the rotation of $\hat{p}_{m+1}$ about p_l in the rest frame of Q. This gives

$$
\int \frac{d\phi}{2\pi} \sum_{s_{m+1}, \hat{s}_{m+1}} \delta_{s_{m+1}, \hat{s}_{m+1}} \big(\{\hat{s}', \hat{s}\}_{m+1} \big| \mathcal{W}_{\text{tot}}(l, k; \{\hat{f}, \hat{p}\}_{m+1}) \big| \{s', s\}_m \big) =
$$
$$
\left(\prod_{j \neq m+1} \delta_{\hat{s}_j, s_j} \delta_{\hat{s}'_j, s'_j} \right) (4\pi\alpha_s) \, I_{\text{soft}} \ ,
\tag{11.4}
$$

where

$$
\begin{aligned}
I_{\text{soft}} = \int \frac{d\phi}{2\pi} \, & \frac{p_l^{\mu}}{\hat{p}_{m+1} \cdot p_l} \left\{ \frac{p_l^{\nu}}{\hat{p}_{m+1} \cdot p_l} - \frac{p_k^{\nu}}{\hat{p}_{m+1} \cdot p_k} \right\} \\
& \times \left\{ -g^{\mu\nu} + \frac{\hat{p}_{m+1}^{\mu} Q^{\nu} + Q^{\mu} \hat{p}_{m+1}^{\nu}}{\hat{p}_{m+1} \cdot Q} - \frac{Q^2 \, \hat{p}_{m+1}^{\mu} \hat{p}_{m+1}^{\nu}}{(\hat{p}_{m+1} \cdot Q)^2} \right\} \ .
\end{aligned}
\tag{11.5}
$$

One can multiply this out and perform the integral using

$$
\int \frac{d\phi}{2\pi} \frac{1}{A^2 + B^2 + 2AB \cos\phi} = \frac{1}{|A^2 - B^2|} \ .
\tag{11.6}
$$

This gives

$$
I_{\text{soft}} = \frac{Q \cdot p_l}{p_l \cdot \hat{p}_{m+1} \, Q \cdot \hat{p}_{m+1}} \left\{ \frac{\hat{p}_{m+1} \cdot \omega \cdot p_k \, \beta_l}{\sqrt{(\hat{p}_{m+1} \cdot \omega \cdot p_k)^2 + m^2(f_k) \, (\omega \cdot \hat{p}_{m+1})^2}} + 1 \right\} - \frac{m^2(f_l)}{(p_l \cdot \hat{p}_{m+1})^2} \ .
\tag{11.7}
$$

Here β is the velocity of parton l,

$$
\beta_l = \frac{\sqrt{(p_l \cdot Q)^2 - m^2(f_l) Q^2}}{p_l \cdot Q} \ .
\tag{11.8}
$$

We have written $\omega \cdot v$ for the product $\omega_{\alpha}^{\mu} v^{\alpha}$ of a vector v with the tensor ω defined by

$$
\omega^{\mu\nu} = \frac{Q^{\mu} p_l^{\nu} - p_l^{\mu} Q^{\nu}}{Q \cdot p_l} \ .
\tag{11.9}
$$

The tensor $\beta_l^{-1} \omega^{\mu\nu}$ is the unit antisymmetric tensor in the p_l-Q plane and projects onto this plane. It obeys $\beta_l^{-2} \omega^{\mu}{}_{\alpha} \omega^{\alpha}{}_{\nu} = P^{\mu}_{\nu}$, where P^{μ}_{ν} is the projection operator onto the p_l-Q plane.

In the case that all of the masses are zero, eq. (11.7) becomes

$$
I_{\text{soft}} = \frac{Q \cdot p_l}{p_l \cdot \hat{p}_{m+1} \, Q \cdot \hat{p}_{m+1}} \left\{ \frac{\hat{p}_{m+1} \cdot \omega \cdot p_k}{\sqrt{(\hat{p}_{m+1} \cdot \omega \cdot p_k)^2}} + 1 \right\} \qquad \text{(masses vanish)} \ . \tag{11.10}
$$

The ratio in the first term in eq. (11.10) is either $+1$ or -1. It is $+1$ if

$$\frac{\hat{p}_{m+1}{\cdot}Q}{\hat{p}_{m+1}{\cdot}p_l} > \frac{p_k{\cdot}Q}{p_k{\cdot}p_l} \ . \tag{11.11}$$

That is, the ratio is $+1$ if, in the rest frame of Q, $\hat{p}_{m+1}$ makes a smaller angle with p_l than does p_k. In this small angle region, we have

$$I_{\text{soft}} = \frac{2Q{\cdot}p_l}{p_l{\cdot}\hat{p}_{m+1} \ Q{\cdot}\hat{p}_{m+1}} \qquad \text{(masses vanish, small angle)} \ . \tag{11.12}$$

In the wide angle region, where $\hat{p}_{m+1}$ makes a larger angle with p_l than does p_k, we have

$$I_{\text{soft}} = 0 \qquad \text{(masses vanish, wide angle)} \ . \tag{11.13}$$

Thus, in the massless case, after a sum over spins and an average over the azimuthal angle, the soft radiation outside a cone centered on p_l and extending out to p_k cancels exactly. This phenomenon is known as angular ordering. It can be used to define the angular ordering approximation, in which one simply neglects the radiation outside of this cone even for a fixed set of spins and a fixed azimuthal angle. This angular ordering approximation is commonly used in parton shower Monte Carlo event generators. With the formalism of this paper, one can make this approximation if desired, but it is not required.

In the massive case, it is no longer true that soft radiation outside a cone centered on p_l and extending out to p_k cancels after integrating over the azimuthal angle.

12. Inclusive evolution

As described in section 3.5, the Sudakov exponent is constructed from the operator $\mathcal{V}(t)$. In turn, $\mathcal{V}(t)$ is constructed from $\big(1\big|\mathcal{H}_{\text{I}}(t)\big|\{p,f,s',c',s,c\}_m\big)$, which tells the total probability for the state $\big|\{p,f,s',c',s,c\}_m\big)$ to split at time t. In this section, we determine the structure of $\big(1\big|\mathcal{H}_{\text{I}}(t)\big|\{p,f,s',c',s,c\}_m\big)$.

We know the matrix elements of $\mathcal{H}_{\text{I}}(t)$, so we simply insert the completeness relation (3.28) for the basis states and use eq. (3.44) for the inner product of $\big(1\big|$ with a basis state. Thus

$$\begin{aligned}
\big(1\big|\mathcal{H}_{\text{I}}(t)\big|\{p,f,s',c',s,c\}_m\big) = \frac{1}{(m+1)!} \int & \big[d\{\hat{p},\hat{f},\hat{s}',\hat{c}',\hat{s},\hat{c}\}_{m+1}\big] \\
& \times \big\langle\{\hat{s}'\}_{m+1}\big|\{\hat{s}\}_{m+1}\big\rangle \big\langle\{\hat{c}'\}_{m+1}\big|\{\hat{c}\}_{m+1}\big\rangle \\
& \times \big(\{\hat{p},\hat{f},\hat{s}',\hat{c}',\hat{s},\hat{c}\}_{m+1}\big|\mathcal{H}_{\text{I}}(t)\big|\{p,f,s',c',s,c\}_m\big) \ .
\end{aligned} \tag{12.1}$$

We use eq. (9.3) to express $\mathcal{H}_{\text{I}}(t)$ in terms of a sum of splitting operators $\mathcal{S}_l$ and then use eqs. (8.26) and (8.27) for the matrix elements of $\mathcal{S}_l$. We encounter

$$\begin{aligned}
\frac{1}{(m+1)!} \int \big[d\{\hat{p},\hat{f}\}_{m+1}\big]\big(\{\hat{p},\hat{f}\}_{m+1}\big|\mathcal{P}_l\big|\{p,f\}_m\big)\cdots \\
= \frac{1}{(m+1)} \sum_{\zeta_{\text{f}}\in\Phi_l(f_l)} \int d\zeta_{\text{p}} \ \theta(\zeta_{\text{p}}\in\Gamma_l(\{p\}_m,\zeta_{\text{f}})) \ \cdots \ ,
\end{aligned} \tag{12.2}$$

where we have used eq. (8.20). Thus we are really integrating over the splitting variables. Inside the integral and the sum over l, the variables $\{\hat{p}, \hat{f}\}_{m+1}$ are determined from $\{p, f\}_m$ and the splitting variables $\{\zeta_{\mathrm{f}}, \zeta_{\mathrm{f}}\}$ by the transformation R_l, eqs. (4.28) and (4.63). Then

$$
\begin{aligned}
\big(1\big|\mathcal{H}_{\mathrm{I}}(t)\big|\{p, f, s', c', s, c\}_m\big) &= \sum_l \sum_{\zeta_{\mathrm{f}} \in \Phi_l(f_l)} \int d\zeta_{\mathrm{p}}\ \theta(\zeta_{\mathrm{p}} \in \Gamma_l(\{p\}_m, \zeta_{\mathrm{f}})) \\
&\times \delta\left(t - \log\left(\frac{Q_0^2}{|(\hat{p}_l + (-1)^{\delta_{l,\mathrm{a}}+\delta_{l,\mathrm{b}}} \hat{p}_{m+1})^2 - m^2(f_l)|}\right)\right) \\
&\times \frac{n_{\mathrm{c}}(a) n_{\mathrm{c}}(b)\, \eta_{\mathrm{a}} \eta_{\mathrm{b}}}{n_{\mathrm{c}}(\hat{a}) n_{\mathrm{c}}(\hat{b})\, \hat{\eta}_{\mathrm{a}} \hat{\eta}_{\mathrm{b}}} \frac{f_{\hat{a}/A}(\hat{\eta}_{\mathrm{a}}, \mu_F^2) f_{\hat{b}/B}(\hat{\eta}_{\mathrm{b}}, \mu_F^2)}{f_{a/A}(\eta_{\mathrm{a}}, \mu_F^2) f_{b/B}(\eta_{\mathrm{b}}, \mu_F^2)} \\
&\times \bigg\{ \theta(\hat{f}_{m+1} \neq \mathrm{g})\, \langle\{c'\}_m|g_{ll}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle \langle\{s'\}_m|w_{ll}(\{\hat{f}, \hat{p}\}_{m+1})|\{s\}_m\rangle \\
&+ \theta(\hat{f}_{m+1} = \mathrm{g}) \sum_{k \neq l} \bigg[\langle\{c'\}_m|g_{lk}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle \\
&\times \langle\{s'\}_m|A_{lk}(\{\hat{p}\}_{m+1})w_{lk}(\{\hat{f}, \hat{p}\}_{m+1}) - \frac{1}{2} w_{ll}(\{\hat{f}, \hat{p}\}_{m+1})|\{s\}_m\rangle \\
&+ \langle\{c'\}_m|g_{kl}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle \\
&\times \langle\{s'\}_m|A_{lk}(\{\hat{p}\}_{m+1})w_{kl}(\{\hat{f}, \hat{p}\}_{m+1}) - \frac{1}{2} w_{ll}(\{\hat{f}, \hat{p}\}_{m+1})|\{s\}_m\rangle \bigg] \bigg\} \ .
\end{aligned}
\tag{12.3}
$$

Here the color dependent function is

$$
\begin{aligned}
\langle\{c'\}_m|g_{ij}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle &= \sum_{\{\hat{c}', \hat{c}\}_{m+1}} \langle\{\hat{c}'\}_{m+1}|\{\hat{c}\}_{m+1}\rangle \\
&\times \big(\{\hat{c}', \hat{c}\}_{m+1}|\mathcal{G}(i, j; \{\hat{f}\}_{m+1})|\{c', c\}_m\big)
\end{aligned}
\tag{12.4}
$$

and the spin dependent function is

$$
\begin{aligned}
\langle\{s'\}_m|w_{ij}(\{\hat{f}, \hat{p}\}_{m+1})|\{s\}_m\rangle &= \sum_{\{\hat{s}', \hat{s}\}_{m+1}} \langle\{\hat{s}'\}_{m+1}|\{\hat{s}\}_{m+1}\rangle \\
&\times \big(\{\hat{s}', \hat{s}\}_{m+1}|\mathcal{W}(i, j; \{\hat{f}, \hat{p}\}_{m+1})|\{s', s\}_m\big) \ .
\end{aligned}
\tag{12.5}
$$

Let us look at the color factor in eq. (12.3) first. Using the definition (8.11) of $\mathcal{G}(l, k; \{\hat{f}\}_{m+1})$, we have

$$
\begin{aligned}
\langle\{c'\}_m|g_{ij}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle &= \sum_{\{\hat{c}', \hat{c}\}_{m+1}} \langle\{c'\}_m|t_j(f_j \to \hat{f}_j + \hat{f}_{m+1})|\{\hat{c}'\}_{m+1}\rangle_D \\
&\times \langle\{\hat{c}'\}_{m+1}|\{\hat{c}\}_{m+1}\rangle\,_D\langle\{\hat{c}\}_{m+1}|t_i^\dagger(f_i \to \hat{f}_i + \hat{f}_{m+1})|\{c\}_m\rangle \ .
\end{aligned}
\tag{12.6}
$$

Then using the completeness relations (3.23) and (3.24) we find

$$
\langle\{c'\}_m|g_{ij}(\{\hat{f}\}_{m+1})|\{c\}_m\rangle = \langle\{c'\}_m|t_j(f_j \to \hat{f}_j + \hat{f}_{m+1})\, t_i^\dagger(f_i \to \hat{f}_i + \hat{f}_{m+1})|\{c\}_m\rangle \ . \tag{12.7}
$$

The operator $t_j t_i^\dagger$ is written as $\boldsymbol{T}_j \cdot \boldsymbol{T}_i$ in the work of Catani and Seymour on the dipole subtraction scheme for next-to-leading order calculations [14]. In that work, there is a sum

over final states, while in this paper we follow the evolution of the exclusive final state that comes between the operators t_j and $t_i^\dagger$. We get back to the inclusive case when we form $\left(1\middle|\mathcal{H}_\mathrm{I}(t)\middle|\{p,f,s',c',s,c\}_m\right)$. The operators g_{ij} have some simple properties. From the definition eq. (7.20), we see that

$$\left\langle\{c'\}_m\middle|g_{ij}(\{\hat{f}\}_{m+1})\middle|\{c\}_m\right\rangle = \left\langle\{c'\}_m\middle|g_{ji}(\{\hat{f}\}_{m+1})\middle|\{c\}_m\right\rangle \ . \tag{12.8}$$

Furthermore, when $i = j$, the operators are proportional to the unit operator,

$$\left\langle\{c'\}_m\middle|g_{ll}(\{\hat{f}\}_{m+1})\middle|\{c\}_m\right\rangle = \left\langle\{c'\}_m\middle|\{c\}_m\right\rangle \times \begin{cases} C_\mathrm{F} \ , & \{\hat{f}_l,\hat{f}_{m+1}\} = \{q,\mathrm{g}\} \text{ or } \{\bar{q},\mathrm{g}\} \\ C_\mathrm{A} \ , & \{\hat{f}_l,\hat{f}_{m+1}\} = \{\mathrm{g},\mathrm{g}\} \\ T_\mathrm{R} \ , & \{\hat{f}_l,\hat{f}_{m+1}\} = \{q,\bar{q}\} \end{cases} \ . \tag{12.9}$$

The last of these cases occurs in eq. (12.3).

Let us look next at the spin dependent factor in eq. (12.3). In order to do this, we need to introduce the possibility of averaging over the azimuthal angle of parton splitting. For the splitting of parton l, we define the transverse part, $q_\perp$ of $\hat{p}_{m+1}$ by

$$\hat{p}_{m+1} = a p_l + b n_l + q_\perp \ , \tag{12.10}$$

where n_l is the lightlike vector defined in eq. (5.3) and $q_\perp \cdot p_l = q_\perp \cdot n_l = 0$. We let ϕ be the angle of $q_\perp$ as measured in any convenient coordinate system, so that $\int d\phi$ means integrating over $q_\perp$ at fixed $|q_\perp|$. Thus integrating over ϕ is part of integrating over the splitting variables ζ_p.

Consider first the spin dependent factors for the case $i \neq j$, which arises from interference diagrams. Using the definition (8.17) of $\mathcal{W}(i,j;\{\hat{f},\hat{p}\}_{m+1})$ and the orthogonality of the spin basis vectors, this is

$$\begin{aligned} \left\langle\{s'\}_m\middle|w_{ij}(\{\hat{f},\hat{p}\}_{m+1})\middle|\{s\}_m\right\rangle = {} & \sum_{\{\hat{s}\}_{m+1}} \left\langle\{s'\}_m\middle|V_j^\mathrm{soft}(\{\hat{p},\hat{f}\}_{m+1})\middle|\{\hat{s}\}_{m+1}\right\rangle \\ & \times \left\langle\{\hat{s}\}_{m+1}\middle|V_i^{\dagger,\mathrm{soft}}(\{\hat{p},\hat{f}\}_{m+1})\middle|\{s\}_m\right\rangle \ . \end{aligned} \tag{12.11}$$

Using the definition (6.35) and (6.36) of V^soft, this is

$$\begin{aligned} & \left\langle\{s'\}_m\middle|w_{ij}(\{\hat{f},\hat{p}\}_{m+1})\middle|\{s\}_m\right\rangle \\ & = \sum_{\{\hat{s}\}_{m+1}} \left(\prod_{n\in\{\mathrm{a,b,1,\ldots},m\}} \delta_{\hat{s}_n,s_n}\delta_{\hat{s}_n,s'_n}\right)(4\pi\alpha_\mathrm{s}) \\ & \quad \times \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)^*\cdot\hat{p}_i}{\hat{p}_{m+1}\cdot\hat{p}_i} \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)\cdot\hat{p}_j}{\hat{p}_{m+1}\cdot\hat{p}_j} \\ & = \left(\prod_{n\in\{\mathrm{a,b,1,\ldots},m\}} \delta_{s_n,s'_n}\right)(4\pi\alpha_\mathrm{s})\sum_{\hat{s}_{m+1}} \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)^*\cdot\hat{p}_i}{\hat{p}_{m+1}\cdot\hat{p}_i} \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)\cdot\hat{p}_j}{\hat{p}_{m+1}\cdot\hat{p}_j} \\ & = \left\langle\{s'\}_m\middle|\{s\}_m\right\rangle(4\pi\alpha_\mathrm{s})\sum_{\hat{s}_{m+1}} \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)^*\cdot\hat{p}_i}{\hat{p}_{m+1}\cdot\hat{p}_i} \frac{\varepsilon(\hat{p}_{m+1},\hat{s}_{m+1};Q)\cdot\hat{p}_j}{\hat{p}_{m+1}\cdot\hat{p}_j} \ . \end{aligned} \tag{12.12}$$

Notice that $\langle\{s'\}_m|w_{ij}(\{\hat{f},\hat{p}\}_{m+1})|\{s\}_m\rangle$ is proportional to a unit matrix $\langle\{s'\}_m|\{s\}_m\rangle$ in the spin indices and that it is symmetric under $i\leftrightarrow j$. The coefficient of $\langle\{s'\}_m|\{s\}_m\rangle$ is a product of splitting functions v_i^{soft} and $v_j^{\mathrm{soft}\,*}$ as defined in eq. (6.36), summed over $\hat{s}_i$ and $\hat{s}_j$ and averaged over s_i and s_j. That is

$$\langle\{s'\}_m|w_{ij}(\{\hat{f},\hat{p}\}_{m+1})|\{s\}_m\rangle = \langle\{s'\}_m|\{s\}_m\rangle\,\overline{w}_{ij}(\{\hat{f},\hat{p}\}_{m+1})\ , \tag{12.13}$$

where $\overline{w}_{ij}$ is

$$\begin{aligned}
\overline{w}_{ij}(\{\hat{f},\hat{p}\}_{m+1}) &= \overline{w}_{ji}(\{\hat{f},\hat{p}\}_{m+1})\\
&= \frac{1}{4}\sum_{\hat{s}_{m+1}}\sum_{\hat{s}_i}\sum_{s_i}\sum_{\hat{s}_j}\sum_{s_j} v_j^{\mathrm{soft}}(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_j,s_j)^*\, v_i^{\mathrm{soft}}(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_i,s_i)\ .
\end{aligned} \tag{12.14}$$

Consider next the case $i = j = l \in \{1,\ldots,m\}$ with $\hat{f}_l \neq \mathrm{g}$ or with $l = \mathrm{a}$ or b and any $\hat{f}_l$. Using the definition (8.12) of $\mathcal{W}(l,l;\{\hat{f},\hat{p}\}_{m+1})$ and orthogonality for the spin basis vectors, this is

$$\begin{aligned}
\langle\{s'\}_m|w_{ll}(\{\hat{f},\hat{p}\}_{m+1})|\{s\}_m\rangle &= \sum_{\{\hat{s}\}_{m+1}} S_l(\{\hat{f}\}_{m+1})\,\langle\{s'\}_m|V_l(\{\hat{p},\hat{f}\}_{m+1})|\{\hat{s}\}_{m+1}\rangle\\
&\quad\times \langle\{\hat{s}\}_{m+1}|V_l^\dagger(\{\hat{p},\hat{f}\}_{m+1})|\{s\}_m\rangle\ .
\end{aligned} \tag{12.15}$$

Using the definition of V as given in eq. (6.6), this is

$$\begin{aligned}
\langle\{s'\}_m|w_{ll}(\{\hat{f},\hat{p}\}_{m+1})|\{s\}_m\rangle &= \sum_{\{\hat{s}\}_{m+1}}\left(\prod_{n\notin\{l,m+1\}}\delta_{\hat{s}_n,s_n}\delta_{\hat{s}_n,s'_n}\right)S_l(\{\hat{f}\}_{m+1})\\
&\quad\times v_l^*(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s'_l)\,v_l(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s_l)\\
&= \sum_{\hat{s}_{m+1}}\sum_{\hat{s}_l}\left(\prod_{n\notin\{l,m+1\}}\delta_{s'_n,s_n}\right)S_l(\{\hat{f}\}_{m+1})\\
&\quad\times v_l^*(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s'_l)\,v_l(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s_l)\ .
\end{aligned} \tag{12.16}$$

This is proportional to the unit matrix in the spin indices s_n,s'_n for all n except possibly for s_l,s'_l, the indices that appear in the functions v_l. However, in eq. (12.3) there is an integration over the azimuthal angle ϕ as part of the integration over ζ_{p}. Once we sum over the final state spin indices $\hat{s}_l,\hat{s}_{m+1}$ and integrate over ϕ, the result is invariant under rotations about the p_l axis in the rest frame of Q. Thus the result must vanish for $s_l \neq s'_l$. In addition, the parity invariance of the splitting vertices implies that the result is invariant under a reflection through a plane containing p_l, Q, and any vector transverse to p_l and Q. Under this transformation, $s_l \leftrightarrow -s_l$ and $s'_l \leftrightarrow -s'_l$. Thus the result is proportional to δ_{s_l,s'_l}. The coefficient of the δ_{s_l,s'_l} can be obtained by setting $s'_l \to s_l$, summing over s_l, and

multiplying by 1/2. Thus

$$\int \frac{d\phi}{2\pi} \langle \{s'\}_m | w_{ll}(\{\hat{f}, \hat{p}\}_{m+1}) | \{s\}_m \rangle = \int \frac{d\phi}{2\pi} \left(\prod_{n=1}^{m} \delta_{s'_n, s_n} \right) \frac{1}{2} \sum_{\hat{s}_{m+1}} \sum_{\hat{s}_l} \sum_{\bar{s}_l} S_l(\{\hat{f}\}_{m+1})$$
$$\times v_l^*(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, \bar{s}_l) \, v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, \bar{s}_l)$$
$$= \langle \{s'\}_m | \{s\}_m \rangle \, S_l(\{\hat{f}\}_{m+1})$$
$$\times \frac{1}{2} \sum_{\hat{s}_{m+1}} \sum_{\hat{s}_l} \sum_{\bar{s}_l} \left| v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, \bar{s}_l) \right|^2 \; .$$

$$(12.17)$$

In the last line, we have not written the average over ϕ because, once we have summed over all of the spins, the result is independent of ϕ.

For the special case $i = j = l \in \{1, \ldots, m\}$ with $\hat{f}_l = $ g, there is an extra term in the definition (8.12) of $\mathcal{W}(l, l; \{\hat{f}, \hat{p}\}_{m+1})$. The extra term, involving $\widetilde{\mathcal{W}}$ defined in eq. (8.16), is built from some of the separate terms in the three gluon vertex. Their treatment is essentially the same as the treatment just given for the other vertex functions. The result for $i = j = l$ in general is

$$\int \frac{d\phi}{2\pi} \, \langle \{s'\}_m | w_{ll}(\{\hat{f}, \hat{p}\}_{m+1}) | \{s\}_m \rangle = \langle \{s'\}_m | \{s\}_m \rangle \, \overline{w}_{ll}(\{\hat{f}, \hat{p}\}_{m+1}) \; , \qquad (12.18)$$

where

$$\overline{w}_{ll}(\{\hat{f}, \hat{p}\}_{m+1}) = S_l(\{\hat{f}\}_{m+1}) \frac{1}{2} \sum_{\hat{s}_{m+1}} \sum_{\hat{s}_l} \sum_{s_l} \times \left\{ \left| v_l(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l) \right|^2 \right.$$
$$+ \theta(l \in \{1, \ldots, m\}, \hat{f}_l = \hat{f}_{m+1} = \text{g})$$
$$\times \left[|v_{2,l}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l)|^2 - |v_{3,l}(\{\hat{p}, \hat{f}\}_{m+1}, \hat{s}_{m+1}, \hat{s}_l, s_l)|^2 \right] \right\} \; .$$

$$(12.19)$$

Here $v_{2,l}$ and $v_{3,l}$ are defined in eq. (8.14).

We conclude that $\left(1 | \mathcal{H}_{\mathrm{I}}(t) | \{p, f, s', c', s, c\}_m \right)$ has the form given in eq. (3.49),

$$\left(1 | \mathcal{H}_{\mathrm{I}}(t) | \{p, f, s', c', s, c\}_m \right) = 2 \langle \{s'\}_m | \{s\}_m \rangle \langle \{c'\}_m | h(t, \{p, f\}_m) | \{c\}_m \rangle \; . \qquad (12.20)$$

It is proportional to the unit matrix in spin but is not proportional to the unit matrix, or

even diagonal, in color. The matrix $\langle \{c'\}_m | h(t, \{p, f\}_m) | \{c\}_m \rangle$ is

$$
\langle \{c'\}_m | h(t, \{p, f\}_m) | \{c\}_m \rangle = \frac{1}{2} \sum_l \sum_{\zeta_{\rm f} \in \Phi_l(f_l)} \int d\zeta_{\rm p}\, \theta(\zeta_{\rm p} \in \Gamma_l(\{p\}_m, \zeta_{\rm f}))
$$

$$
\times\, \delta\left(t - \log\left(\frac{Q_0^2}{|(\hat{p}_l + (-1)^{\delta_{l,{\rm a}} + \delta_{l,{\rm b}}} \hat{p}_{m+1})^2 - m^2(f_l)|} \right) \right)
$$

$$
\times\, \frac{n_{\rm c}(a) n_{\rm c}(b)\, \eta_{\rm a} \eta_{\rm b}}{n_{\rm c}(\hat{a}) n_{\rm c}(\hat{b})\, \hat{\eta}_{\rm a} \hat{\eta}_{\rm b}} \, \frac{f_{\hat{a}/A}(\hat{\eta}_{\rm a}, \mu_F^2) f_{\hat{b}/B}(\hat{\eta}_{\rm b}, \mu_F^2)}{f_{a/A}(\eta_{\rm a}, \mu_F^2) f_{b/B}(\eta_{\rm b}, \mu_F^2)}
$$

$$
\times\, \Bigg\{ \theta(\hat{f}_{m+1} \neq {\rm g})\, \langle \{c'\}_m | \{c\}_m \rangle\, T_{\rm R}\, \overline{w}_{ll}(\{\hat{f}, \hat{p}\}_{m+1})
$$

$$
+\, \theta(\hat{f}_{m+1} = {\rm g}) \sum_{k \neq l} \langle \{c'\}_m | g_{lk}(\{\hat{f}\}_{m+1}) | \{c\}_m \rangle
$$

$$
\times\, \big[2\, A_{lk}(\{\hat{p}\}_{m+1})\, \overline{w}_{lk}(\{\hat{f}, \hat{p}\}_{m+1}) - \overline{w}_{ll}(\{\hat{f}, \hat{p}\}_{m+1}) \big] \Bigg\} \;.
$$

(12.21)

Here $\langle \{c'\}_m | g_{ij}(\{\hat{f}\}_{m+1}) | \{c\}_m \rangle$ is given in eq. (12.7) while $\overline{w}_{ij}(\{\hat{f}, \hat{p}\}_{m+1})$ for $i \neq j$ is given in eq. (12.14) and $\overline{w}_{ll}(\{\hat{f}, \hat{p}\}_{m+1})$ is given in eq. (12.19). The scales $\mu_{\rm F}$ used in the parton distribution functions and $\mu_{\rm R}$ used in $\alpha_{\rm s}$ are given by eq. (9.4). Inside the integral and the sum over l, the variables $\{\hat{p}, \hat{f}\}_{m+1}$ are determined from $\{p, f\}_m$ and the splitting variables $\{\zeta_{\rm p}, \zeta_{\rm f}\}$ by the transformation R_l, eqs. (4.28) and (4.63).

Notice that the matrix $\langle \{c'\}_m | h(t, \{p, f\}_m) | \{c\}_m \rangle$ is not diagonal in color. However, the matrix elements with $\{c'\}_m \neq \{c'\}_m$ are suppressed by powers of $1/N_{\rm c}$. For the term with a factor $\langle \{c'\}_m | \{c\}_m \rangle$, this was already noted in eq. (7.14). To see this for the term involving g_{lk}, we should write the matrix element of g_{lk} in the form of eq. (12.6),

$$
\langle \{c'\}_m | g_{lk}(\{\hat{f}\}_{m+1}) | \{c\}_m \rangle = \sum_{\{\hat{c}', \hat{c}\}_{m+1}} {}_D\langle \{\hat{c}'\}_{m+1} | t_k^\dagger(f_k \to \hat{f}_k + \hat{f}_{m+1}) | \{c'\}_m \rangle
$$

$$
\times\, \langle \{\hat{c}'\}_{m+1} | \{\hat{c}\}_{m+1} \rangle\, {}_D\langle \{\hat{c}\}_{m+1} | t_l^\dagger(f_l \to \hat{f}_l + \hat{f}_{m+1}) | \{c\}_m \rangle \;.
$$

(12.22)

The leading contribution for $1/N_{\rm c} \to 0$ comes when $\{c'\}_m = \{c\}_m$ and partons l and k are color connected: they lie next to each other on a string. As discussed in section 10, we can use the representations (7.24), (7.25), (7.26), and (7.35) for the operators $t_l^\dagger$ and $t_k^\dagger$ to see that the action of $t_l^\dagger$ and $t_k^\dagger$ can produce $\{\hat{c}'\}_{m+1} = \{\hat{c}\}_{m+1}$. If $\{c'\}_m \neq \{c\}_m$, then the action of $t_l^\dagger$ and $t_k^\dagger$ always produces $\{\hat{c}'\}_{m+1} \neq \{\hat{c}\}_{m+1}$, so that we get a color suppressed inner product $\langle \{\hat{c}'\}_{m+1} | \{\hat{c}\}_{m+1} \rangle$.

13. End of the shower

As the shower progresses toward smaller and smaller resolution scales $\mu^2 = Q_0^2 e^{-t}$, there must come a point at which the perturbative basis of the evolution equation is no longer valid. Then, at some evolution time $t_{\rm f}$, the shower evolution should be stopped. In the event that the resolution scale of the desired measurement function is larger than $\mu_{\rm f}^2 = Q_0^2 e^{-t_{\rm f}}$,

whatever happens beyond that is not seen by the measurement. Then one could simply apply the measurement function, calculating

$$\big(F\big|\rho(t_{\mathrm{f}})\big) \ . \tag{13.1}$$

Let us suppose that the measurement function does not see the spins or colors of the final state partons. Then, as in eq. (3.34), we need

$$\big(F\big|\rho(t_{\mathrm{f}})\big) = \sum_m \frac{1}{m!} \int \big[d\{p,f,s',c',s,c\}_m\big] F(\{p,f\}_m) \, \big\langle\{s'\}_m\big|\{s\}_m\big\rangle\big\langle\{c'\}_m\big|\{c\}_m\big\rangle$$
$$\times \big(\{p,f,s',c',s,c\}_m\big|\rho(t_{\mathrm{f}})\big) \ . \tag{13.2}$$

It is significant that, although the possibility of $\{c'\}_m \neq \{c\}_m$ is included in the shower evolution presented in this paper, the matrix $\big\langle\{c'\}_m\big|\{c\}_m\big\rangle$ is almost diagonal, with off-diagonal matrix elements being suppressed by factors of $1/N_{\mathrm{c}}^2$. The situations with respect to spins and colors are different. In the end, we must have $\{s'\}_m = \{s\}_m$ exactly. However, the spins get shuffled at each stage of shower evolution and there is no reason that a state with $\{s'\}_m \neq \{s\}_m$ at an earlier stage of evolution cannot evolve into a state with $\{s'\}_m = \{s\}_m$ at the end. On the other hand, color differences between $\{c'\}_m$ and $\{c\}_m$ are a little like entropy. Once $\{c'\}_m \neq \{c\}_m$ at an early stage of evolution, we can never get $\{c'\}_m = \{c\}_m$ at the end. For this reason, the most important part of color evolution is the part that maintains $\{c'\}_m = \{c\}_m$ throughout. This is evolution in the leading color dipole approximation.

Now, what if we wish to use a measurement function with a resolution scale smaller than μ_{f}^2. Then we need a model for what happens at smaller resolution scales (or later and earlier proper times than given by $x^2 \sim 1/\mu_{\mathrm{f}}^2$). Our model should certainly include hadronization. We can easily extend the formalism presented here to encompass hadronization. We have only to replace $\big(F\big|\rho(t_{\mathrm{f}})\big)$ by

$$\big(F_{\mathrm{h}}\big|\mathcal{U}(\infty,t_{\mathrm{f}})\big|\rho(t_{\mathrm{f}})\big) \ . \tag{13.3}$$

Here $\mathcal{U}(\infty,t_{\mathrm{f}})$ represents a model for what happens after Monte Carlo time t_{f}. It starts with partonic states and maps them into the space of hadronic states. Typically the hadronic states are labeled by momenta and hadronic flavors but not spins. Then $\big(F_{\mathrm{h}}\big|$ represents the measurement function in the space of hadronic states. For purposes of discussing the partonic shower, we can denote

$$\big(F_{\mathrm{h}}\big|\mathcal{U}(\infty,t_{\mathrm{f}}) = \big(F_{\mathrm{eff}}\big| \ . \tag{13.4}$$

Thus F_{eff} is the true hadronic measurement function translated back to the partonic level.

Assuming that the hadronization model does not use color or spin information, the measured cross section then takes the form

$$\big(F_{\mathrm{h}}\big|\mathcal{U}(\infty,t_{\mathrm{f}})\big|\rho(t_{\mathrm{f}})\big) = \big(F_{\mathrm{eff}}\big|\rho(t_{\mathrm{f}})\big)$$
$$= \sum_m \frac{1}{m!} \int \big[d\{p,f,s',c',s,c\}_m\big] F_{\mathrm{eff}}(\{p,f\}_m) \tag{13.5}$$
$$\times \big\langle\{s'\}_m\big|\{s\}_m\big\rangle\big\langle\{c'\}_m\big|\{c\}_m\big\rangle\big(\{p,f,s',c',s,c\}_m\big|\rho(t_{\mathrm{f}})\big) \ .$$

Typically, hadronization models *do* use color information. The color field interacting with the outgoing partons is represented as a classical color string. The string then fragments into hadrons. This applies directly in PYTHIA [2] and in a different way in HERWIG [1], where the color strings fragment into color singlet parton clusters immediately. The formalism of this paper is set up with the color string picture in mind. The color states $\{c\}_m$ exactly map onto string configurations, as explained in section 7.1. Thus for terms in eq. (13.5) with $\{c'\}_m = \{c\}_m$, one can use $\{c\}_m$ as the input to the hadronization model. The string model does not tell us what to do with $\{c'\}_m \neq \{c\}_m$. A reasonable suggestion would be to use $\{c'\}_m$ half the time and $\{c\}_m$ half the time. This could be represented as

$$
\begin{aligned}
\big(F_{\mathrm{h}}\big|\mathcal{U}(\infty, t_{\mathrm{f}})\big|\rho(t_{\mathrm{f}})\big) = \sum_m \frac{1}{m!} \int \Big[d\{p, f, s', c', s, c\}_m \Big] \\
\times \frac{1}{2}\Big[F_{\mathrm{eff}}(\{p, f, c\}_m) + F_{\mathrm{eff}}(\{p, f, c'\}_m) \Big] \\
\times \big\langle\{s'\}_m\big|\{s\}_m\big\rangle\big\langle\{c'\}_m\big|\{c\}_m\big\rangle\big(\{p, f, s', c', s, c\}_m\big|\rho(t_{\mathrm{f}})\big) \ .
\end{aligned}
\tag{13.6}
$$

If the hadronization model is based on strings, one needs to do something with the string ends that connect to the initial state partons. Consider, for example, the case that the hard collision at scale μ_{f}^2 is a quark-quark collision and concentrate on one of the initial state quarks. From the point of view of backwards evolution from the scale Q_0^2, this initial state quark appears as a color $\overline{\mathbf{3}}$ line. The initial state quark is part of a colorless hadron, but the net color $\overline{\mathbf{3}}$ is carried by the spectator quarks from this hadron. Thus in a color string model, the color strings from the hard interactions should connect to the spectator partons. Of course, this is not a completely simple problem. In a realistic model, the spectator partons from the two hadrons have many interactions with each other, possibly followed by their own (not very hard) parton showers. After all of these interactions, the spectator partons left over from a quark-quark collision must have $\overline{\mathbf{3}} \otimes \overline{\mathbf{3}}$ color, but its internal color state can be quite complicated. Modeling the spectator interactions is well beyond the scope of this paper but is addressed by Sjöstrand and Skands in ref. [12].

14. Conclusions

We have presented a formulation of parton showering for hadron-hadron collisions. The prediction for a cross section corresponding to an observable F is given by $\big(F\big|\rho(t_{\mathrm{f}})\big)$ or one of the other formulas in section 13, depending on the treatment of hadronization, which is not covered in this paper. The dynamics of the quantum density $\big|\rho(t)\big)$ is given by the evolution operator $\mathcal{U}(t, t')$, so that $\big|\rho(t_{\mathrm{f}})\big) = \mathcal{U}(t_{\mathrm{f}}, 0)\big|\rho(0)\big)$, where $\big|\rho(0)\big)$ is determined from the hard matrix element that starts the shower. Thus the shower dynamics represented in $\mathcal{U}(t, t')$ is based on factorization of soft and collinear singularities from hard scattering. The basic formula is eq. (3.56), which we can rewrite as

$$
\mathcal{U}(t_{\mathrm{f}}, t') = \mathcal{N}(t_{\mathrm{f}}, t') + \int_{t'}^{t_{\mathrm{f}}} d\tau \, \mathcal{U}(t_{\mathrm{f}}, \tau) \left[\mathcal{H}_{\mathrm{I}}(\tau) - \mathcal{V}_{\mathrm{S}}(\tau) \right] \mathcal{N}(\tau, t') \ .
\tag{14.1}
$$

Here $\mathcal{H}_{\mathrm{I}}(t)$ is a parton splitting operator, as defined in the preceding sections and $\mathcal{V}_{\mathrm{S}}(t)$ represents a virtual interaction that interchanges colors. The operator $\mathcal{N}(\tau, t')$ generates

quantity	equation	uses
$\mathcal{U}(t,t')$	Eq. (3.56)	$\mathcal{H}_{\mathrm{I}}$, $\mathcal{N}$, $\mathcal{V}_{\mathrm{S}}$
$\mathcal{H}_{\mathrm{I}}(\tau)$	Eq. (9.3)	$\mathcal{S}_l$
$\mathcal{V}(t,\{p,f\}_m)$	Eq. (3.53)	h
$\mathcal{V}_{\mathrm{S}}(t,\{p,f\}_m)$	Eq. (3.61)	$\mathcal{V}$, $\mathcal{V}_{\mathrm{E}}$
$\mathcal{V}_{\mathrm{E}}(t,\{p,f\}_m)$	Eq. (3.60)	h
$\mathcal{N}(t,t')$	Eq. (3.57)	$\mathcal{V}_{\mathrm{E}}$
$h(t,\{p,f\}_m)$	Eq. (12.21)	g_{ij}, $\overline{w}_{lk}$, A_{lk}
$g_{ij}(\{\hat{f}\}_{m+1})$	Eq. (12.6)	$t_l^\dagger$, color basis
$\overline{w}_{ij}(\{\hat{f},\hat{p}\}_{m+1})$	Eqs. (12.14, 12.19)	v_i^{soft}, v_l, $v_{J,l}$
$t_l^\dagger(f_j \to \hat{f}_j + \hat{f}_{m+1})$	Eqs. (7.24)–(7.26), (7.35)	$a_+^\dagger$, $a_-^\dagger$, $a_q^\dagger$, $a_s^\dagger$
$a_+^\dagger$, $a_-^\dagger$, $a_q^\dagger$, $a_s^\dagger$	Eq. (7.15)–(7.18)	color basis
color basis	Sec. 7.1	
$\mathcal{S}_l$	Eq. (8.23)	$\mathcal{S}_{ij}^{(l)}$, A_{lk}
$\mathcal{S}_{ij}^{(l)}$	Eq. (8.10)	$\mathcal{G}$, $\mathcal{W}$, $\mathcal{P}_l$
$\mathcal{G}(i,j;\{\hat{f}\}_{m+1})$	Eq. (8.11)	$t_i^\dagger$, color basis
$\mathcal{W}(l,l;\{\hat{f},\hat{p}\}_{m+1})$	Eq. (8.12)	S_l, $V_l^\dagger$, $\widetilde{\mathcal{W}}$
$\mathcal{W}(i,j;\{\hat{f},\hat{p}\}_{m+1})$	Eq. (8.17)	$V_i^{\dagger,\mathrm{soft}}$
$\widetilde{\mathcal{W}}(l,l;\{\hat{p}\}_{m+1})$	Eq. (8.16)	$V_{J,l}$
$V_{J,l}^\dagger(\{\hat{p},\hat{f}\}_{m+1})$	Eq. (8.15)	$v_{J,l}$
$v_{J,l}(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s_l)$	Eq. (8.14)	
$S_l(\{\hat{f}\}_{m+1})$	Eq. (8.6)	
$V_l^\dagger(\{\hat{p},\hat{f}\}_{m+1})$	Eq. (6.6)	v_l
$v_l(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s_l)$	Table 1	
	Eqs. (6.31), (6.32)	
$V_l^{\dagger,\mathrm{soft}}(\{\hat{p},\hat{f}\}_{m+1})$	Eq. (6.35)	v_l^{soft}
$v_l^{\mathrm{soft}}(\{\hat{p},\hat{f}\}_{m+1},\hat{s}_{m+1},\hat{s}_l,s_l)$	Eq. (6.36)	
$\mathcal{P}_l$	Eq. (8.18)	Q_l
$Q_l(\{\hat{p},\hat{f}\}_{m+1})$	Sec. 4	
$\int \left[d\{p,f\}_m\right]$	Eq. (3.15)	
$A_{lk}(\{p\}_{m+1})$	Sec. 8	

Table 2: Key formulas.

the standard sort of Sudakov exponential that gives the probability not to have an interaction between shower times t' and τ. We provide table 2 to indicate where the various functions needed to compute $\mathcal{U}$ can be found.

Eq. (14.1) has the proper form to conveniently generate a parton shower. Starting with a state at time t', one would use $\mathcal{N}(\tau,t')$ to determine the time τ for a parton splitting (or color rearrangement). Possibly there is no splitting before the cutoff time t_{f}, as represented by the first term. Otherwise, the operators $\mathcal{H}_{\mathrm{I}}(\tau)$ and $\mathcal{V}_{\mathrm{S}}(\tau)$ give a new partonic state at time τ. Now we operate with $\mathcal{U}(t_{\mathrm{f}},\tau)$, which is to say that we apply this procedure again.

We offer here concluding remarks under several headings.

Implementation. Equation (14.1), together with the formulas in section 13 and the definitions given throughout this paper, represents a certain approximation for the cross section $\sigma[F]$ corresponding to a given observable F. When eq. (14.1) is iterated, the result is expressed in the form of certain integrals and sums. It will be a significant challenge to find ways to implement eq. (14.1) in a manner that allows an efficient calculation of $\sigma[F]$. We expect that there is more than one way to attack this problem. The choice affects the efficiency of calculation, but not the result, $\sigma[F]$. We leave implementation issues to later work.

Evolution variable. We have chosen the evolution variable t to be proportional to the virtuality in the splitting. An alternative would be the transverse momentum in the splitting. The HERWIG choice of the splitting angle does not work well with the formalism presented here since a parton emitted at a fixed angle can be arbitrarily soft, necessitating introducing the final hardness cutoff $\mu_{\rm f}^2$ at each splitting.

Momentum mapping. It is not kinematically possible for an on-shell parton to split into two on-shell partons. However, it is useful to approximate the mother parton as being exactly on-shell in calculating the (relatively) hard scattering in which the mother parton participates. To make this approximation, we need to take the needed momentum from somewhere else. Sometimes this is done by taking momentum from the mother's sister in the previous splitting. In the Catani-Seymour subtraction scheme for doing next-to-leading order calculations, there is a "spectator parton" that donates the required momentum [14]. Rather than taking momentum from a single parton that might not have much to give, we have chosen to take a little momentum from each final state parton, with each donating according to how much momentum it has. Note that we keep momentum in balance at each step, rather than waiting until the end of the shower to make adjustments.

Common evolution. In eq. (14.1), each parton has a chance to split or interchange colors between shower times t and $t + dt$. In the very simplest form of a parton shower, each parton could evolve independently, at least if one ignores momentum conservation and adjusts the momenta only at the end. Then the complete evolution of each parton l could be traced out without keeping track of the other partons. Effectively, there could be a separate time t_l for each parton. However, independent evolution could still be implemented using a common evolution variable t, giving each parton its chance to split between times t and $t + dt$. The physical distinction that characterizes independent evolution is that the evolution is independent if the various functions involved in the splitting of parton l do not involve the states of the other partons. With the definitions of the operators used in eq. (14.1), the use of a common shower time variable is *required*, since each splitting changes the whole partonic state and affects the probabilities for other partons to split at later shower times.

Other choices. Within eq. (14.1), there are quite a number of other choices required. For instance, the splitting functions must have a particular form in the limits of soft and

collinear splittings. However, away from these limits there is a certain freedom to choose. Where choices like this were needed, we exercised the freedom to choose based mostly on conceptual simplicity. Other authors might choose differently.

Interference and angular ordering. In eq. (14.1), quantum interference between emissions of a gluon from different partons is treated exactly in the soft gluon limit. Suppression of wide angle emissions is a result. Typically, parton shower generators make an "angular ordering" approximation to this result. Once one has an implementation of eq. (14.1), it will be interesting to make the standard angular ordering approximation and see how good an approximation it is.

Spin. Partons produced in a hard scattering carry spin, which can affect the angular distribution of their subsequent splittings. Eq. (14.1) includes the full spin information. Typically, parton shower generators average over spins, thus discarding this information. Once one has an implementation of eq. (14.1), it will be interesting to insert spin averages everywhere and see how good an approximation it is.

Color. Partons produced in a hard scattering carry color, which can affect the pattern of future splittings. The formalism presented here includes the full color information. Typically, parton shower generators make use of a leading color approximation that amounts to taking the first term in an expansion about $1/N_c = 0$. Once one has an implementation of eq. (14.1), it will be interesting to make the leading color approximation everywhere and see how good an approximation it is.

Improvements needed. We leave for future work the question of how one could match the parton shower to the exact matrix elements for $2 \to n$ scattering instead of simply starting with $2 \to 2$ scattering. We also leave for future work the question of how one could do this at next-to-leading order. (See, however, ref. [19].) More ambitiously, we would like to extend the whole formalism, including the splitting functions, to next-to-leading order. Current work by others on soft-collinear effective theory may be helpful here [29].

Masses. We gave included quark masses in our momentum mappings and splitting functions. However, there are a number of issues associated with masses that we do not address. Suppose that we start with a hard scattering at a scale Q_0 that is much larger than the mass of the bottom quark, m_b.[27] Suppose additionally that we want to continue evolution down to a scale μ that is less than m_b. Then we need a suitable variable flavor number scheme. At the leading order used in this paper, this is easy enough, but at higher orders of perturbation theory there are some subtle issues. Collins [30] has addressed some of these issues as they arise in deeply inelastic scattering. One should, however, note that for hadron-hadron collisions, the power suppressed terms that are omitted when the cross section is written in a factored form as a hard scattering function convoluted with parton

---

[27]Here we assume that the bottom quark is the heaviest quark counted as initial state parton. One could include production of top quarks, but would not include the top quark as a constituent of the proton unless $Q_0 \gg m_t$.

distributions are not of order $\Lambda_{\rm QCD}^2/Q_0^2$ but rather of order $m_{\rm b}^2/Q_0^2$ [31]. This is not a problem since, in the applications we have in mind, $m_{\rm b}^2/Q_0^2 \ll 1$. We lack a theorem to tell us what to do if we want to keep the hard scattering but lower the resolution scale to $\mu^2 < m_{\rm b}^2$ so as to examine the final state in more detail.

Foundations. The formalism presented here is based on ideas of factorization, both at the amplitude level and at the cross section level, where summation over partonic states that are unresolved at a given scale μ^2 is essential. As the discussion of masses makes clear, more work is needed to make these ideas sufficiently precise to justify the formalism.

Acknowledgments

A. Limit on momentum fraction after splitting

In section (4.4) we defined the kinematics of initial state splitting. For a collinear splitting from an initial state parton, say parton "a,", we have $\hat{\eta}_{\rm a} > \eta_{\rm a}$. It is not exactly evident that this holds in away from the collinear limit, especially with masses. Here we show that this holds under the kinematic conditions (3.6) and (3.9).

Let

$$
\begin{aligned}
\hat{g}(\hat{\eta}_{\rm a}) &= (\hat{p}_{\rm a} + p_{\rm b})^2 \\
&= \hat{\eta}_a \eta_{\rm b} s + m^2(\hat{f}_{\rm a}) + m^2(f_{\rm b}) + \frac{m^2(\hat{f}_{\rm a})\, m^2(f_{\rm b})}{\hat{\eta}_a \eta_{\rm b} s} \quad, \\
g(\eta_{\rm a}) &= (p_{\rm a} + p_{\rm b})^2 \\
&= \eta_a \eta_{\rm b} s + m^2(f_{\rm a}) + m^2(f_{\rm b}) + \frac{m^2(f_{\rm a})\, m^2(f_{\rm b})}{\eta_{\rm a} \eta_{\rm b} s} \quad.
\end{aligned}
\tag{A.1}
$$

Note that $\hat{g}(\hat{\eta}_{\rm a})$ is an increasing function of its argument in the allowed region of momentum fractions:

$$
\frac{1}{\eta_b s} \frac{d\hat{g}(\hat{\eta}_{\rm a})}{d\hat{\eta}_{\rm a}} = 1 - \left(\frac{m(\hat{f}_{\rm a})\, m(f_{\rm b})}{\hat{\eta}_a \eta_{\rm b} s} \right)^2 > 0 \quad,
\tag{A.2}
$$

since $\hat{\eta}_a \eta_{\rm b} s > m_{\rm H}^2$ in our allowed kinematic region according to eq. (3.9). Since $\hat{g}(\hat{\eta}_{\rm a})$ is an increasing function, we just need to show that $\hat{g}(\hat{\eta}_{\rm a}) > \hat{g}(\eta_{\rm a})$.

Recall that the kinematics requires

$$
\begin{aligned}
\hat{g}(\hat{\eta}_{\rm a}) &= (\hat{K} + p_{m+1})^2 \quad, \\
g(\eta_{\rm a}) &= K^2 \quad,
\end{aligned}
\tag{A.3}
$$

where $\hat{K}^2 = K^2$. Thus

$$
\begin{aligned}
\hat{g}(\hat{\eta}_\mathrm{a}) - \hat{g}(\eta_\mathrm{a}) &= \hat{g}(\hat{\eta}_\mathrm{a}) - g(\eta_\mathrm{a}) + g(\eta_\mathrm{a}) - \hat{g}(\eta_\mathrm{a}) \\
&= (\hat{K} + p_{m+1})^2 - K^2 + g(\eta_\mathrm{a}) - \hat{g}(\eta_\mathrm{a}) \\
&= 2\hat{K} \cdot p_{m+1} + m^2(\hat{f}_{m+1}) \\
&\quad + \left(m^2(f_\mathrm{a}) - m^2(\hat{f}_\mathrm{a}) \right) \left[1 + \frac{m^2(f_\mathrm{b})}{\eta_\mathrm{a}\eta_\mathrm{b}s} \right] \ .
\end{aligned}
\tag{A.4}
$$

We need to show that the right hand side of eq. (A.4) is positive. If $m(f_\mathrm{a})^2 \geq m(\hat{f}_\mathrm{a})^2$, this is evident. There is only one case in which $m^2(f_\mathrm{a}) < m^2(\hat{f}_\mathrm{a})$. That is when we have a $q \to qg$ or $\bar{q} \to \bar{q}g$ splitting in which the quark (or antiquark) enters the final state and the gluon enters the hard scattering. Then $m^2(\hat{f}_{m+1}) = m^2(\hat{f}_\mathrm{a})$ and $m^2(f_\mathrm{a}) = 0$. In that case,

$$
\begin{aligned}
\hat{g}(\hat{\eta}_\mathrm{a}) - \hat{g}(\eta_\mathrm{a}) &= 2\hat{K} \cdot p_{m+1} - m^2(\hat{f}_\mathrm{a}) \frac{m^2(f_\mathrm{b})}{\eta_\mathrm{a}\eta_\mathrm{b}s} \\
&> 2\hat{K} \cdot p_{m+1} - m^2(\hat{f}_\mathrm{a}) \ ,
\end{aligned}
\tag{A.5}
$$

where we have used $\eta_a\eta_\mathrm{b}s > m^2(f_\mathrm{b})$ from eq. (3.9). The minimum value of $\hat{K} \cdot p_{m+1}$ occurs when p_{m+1} is proportional to $\hat{K}$. Then $\hat{K} \cdot p_{m+1} = m(\hat{f}_\mathrm{a})\sqrt{K^2}$, so

$$
\hat{g}(\hat{\eta}_\mathrm{a}) - \hat{g}(\eta_\mathrm{a}) > m(\hat{f}_\mathrm{a}) \left(2\sqrt{K^2} - m(\hat{f}_\mathrm{a}) \right) > m(\hat{f}_\mathrm{a}) \left(2\sqrt{Q_0^2} - m(\hat{f}_\mathrm{a}) \right) \ .
\tag{A.6}
$$

This is positive as long as the condition (3.6) holds.

We have seen that

$$
\hat{g}(\hat{\eta}_\mathrm{a}) > \hat{g}(\eta_\mathrm{a}) \ .
\tag{A.7}
$$

Since $\hat{g}(\hat{\eta}_\mathrm{a})$ is an increasing function of $\hat{\eta}_\mathrm{a}$, this implies that $\hat{\eta}_\mathrm{a} > \eta_\mathrm{a}$.

B. Counting factors for the density matrix

In this appendix, we organize the singular contributions to the density matrix starting with the quantum amplitudes defined to be symmetric in the labels of the final state partons (or antisymmetric in the case of identical fermions). Then we introduce the relabelings that define our labeling scheme for parton splittings. This produces the counting factors $S_l(\{\hat{f}\}_{m+1})$ defined in eq. (8.6). The counting factors are related to our parton labeling choices, which in turn related to the singularities of the amplitude in the limit in which masses can be neglected. The counting factors do not depend on parton masses. In order to keep our notation simple, in this appendix we simply take all the parton masses to vanish.

Let $\left| M(\{\hat{p}, \hat{f}\}_{m+1}) \right\rangle$ be the exact tree level matrix element for a final state of $m + 1$ partons, defined to be symmetric under the interchange of the labels for any two of the final state partons, or antisymmetric if the two partons are identical fermions. The matrix element may have a singularity when any of the dot products of two parton momenta, $\hat{p}_i \cdot \hat{p}_j$, approaches zero. Let us define approximate matrix elements $\left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, j) \right\rangle$ that approximate the complete matrix element when $\hat{p}_i \cdot \hat{p}_j$, approaches zero. Here $i \in \{\mathrm{a}, \mathrm{b}\}$

and $j \in \{1, \ldots, m+1\}$ or $i, j \in \{1, \ldots, m+1\}$ with $i < j$. There is more than one way to do this. We can, for instance, use the method of section 6. The approximate matrix elements $\big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle$ thus defined should include a theta function

$$
\theta\left(|\hat{p}_i \cdot \hat{p}_j| < \min_{\substack{\{k,l\} \neq \{i,j\} \\ \{k,l\} \neq \{j,i\}}} |\hat{p}_k \cdot \hat{p}_l| \right)
\tag{B.1}
$$

so that $\big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle$ is not singular when $p_k \cdot p_l$ for some other pair of partons approaches zero. Given the approximate matrix elements, the complete matrix element can be written as

$$
\big|M(\{\hat{p}, \hat{f}\}_{m+1})\big\rangle \sim \sum_{j=1}^{m+1} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; a, j)\big\rangle + \sum_{j=1}^{m+1} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; b, j)\big\rangle
$$
$$
+ \sum_{\substack{i,j=1 \\ i<j}}^{m+1} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle \; .
\tag{B.2}
$$

The right hand side of eq. (B.2) approximates the complete matrix element in any of the singular limits. For our purposes, it is convenient to define

$$
\big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle = \big|M(\{\hat{p}, \hat{f}\}_{m+1}; j, i)\big\rangle
\tag{B.3}
$$

for $i, j \in \{1, \ldots, m+1\}$, $i > j$. Then we can symmetrize the third term in eq. (B.2),

$$
\big|M(\{\hat{p}, \hat{f}\}_{m+1})\big\rangle \sim \sum_{j=1}^{m} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; a, j)\big\rangle + \sum_{j=1}^{m} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; b, j)\big\rangle
$$
$$
+ \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{m+1} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle \; .
\tag{B.4}
$$

Now we construct the density operator from

$$
\rho_0(\{\hat{p}, \hat{f}\}_{m+1}) = \big|M(\{\hat{p}, \hat{f}\}_{m+1})\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1})\big| \; .
\tag{B.5}
$$

Imagine expanding both $\big|M(\{\hat{p}, \hat{f}\}_{m+1})\big\rangle$ and $\big\langle M(\{\hat{p}, \hat{f}\}_{m+1})\big|$ according to eq. (B.4). There are a number of terms, with the general form

$$
const. \times \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i', j')\big| \; .
\tag{B.6}
$$

Not all of these contributions have soft or collinear singularities strong enough to produce a logarithmic divergence if one were to integrate over the momenta $\{\hat{p}\}_{m+1}$.

One contribution that does have a strong enough singularity comes when $i, j, i', j' \in \{1, \ldots, m+1\}$ and $i' = i$, $j' = j$ or $i' = j$, $j' = i$. The sum of these contributions is

$$
\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{m+1} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, j)\big| \; ,
\tag{B.7}
$$

where we have used eq. (B.3).

Another contribution that has a strong enough singularity comes when $i = i' = \mathrm{a}$ and $j, j' \in \{1, \ldots, m+1\}$ with $j' = j$. Similarly, we can have $i = i' = \mathrm{b}$ and $j, j' \in \{1, \ldots, m+1\}$ with $j' = j$. The sum of these contributions is

$$\sum_{j=1}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; \mathrm{a}, j) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; \mathrm{a}, j) \right| + \sum_{j=1}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; \mathrm{b}, j) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; \mathrm{b}, j) \right|.$$

(B.8)

Another contribution that does have a strong enough singularity comes when $i, j, i', j' \in \{1, \ldots, m+1\}$ and $i' = i$, $j' \neq j$ or else $j' = j$, $i' \neq i$ or else $i' = j$, $j' \neq i$ or else $j' = i$, $i' \neq j$. These four cases are really the same, with different labeling. Adding the contributions, we have

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{m+1} \sum_{\substack{k=1 \\ k \neq i, k \neq j}}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right| \ .$$

(B.9)

Here we have a leading singularity only if $\hat{f}_k = \mathrm{g}$. This is the interference between gluon emission from line i and gluon emission from line j.

We can also have interference between gluon emission from an initial state line and gluon emission from a final state line or between gluon emission from one of the initial state lines and gluon emission from the other. When we add all of these cases together and add them to the contribution in eq. (B.9), we get

$$\sum_{\substack{i,j \in \{\mathrm{a}, \mathrm{b}, \ldots, m+1\} \\ i \neq j}} \sum_{\substack{k=1 \\ k \neq i, k \neq j}}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right| \ .$$

(B.10)

There are no more combinations of i, j, i', j' that give leading singular contributions to ρ. Thus the sum of the leading singular contributions is obtained by adding the contributions (B.7), (B.8), and (B.10).

This formulation is fine for constructing an inverse shower, starting from a state with many partons and combining partons to reach a hard scattering with fewer partons. In order to construct a shower starting from the hard scattering, it is convenient to adopt a labeling convention in which $\rho(\{\hat{p}, \hat{f}\}_{m+1})$ is not symmetric under interchanges of the parton labels. This is easy to do.

Consider the contribution (B.7). Using the $i \leftrightarrow j$ symmetry of the matrix elements, this is

$$\sum_{i=1}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1) \right|$$
$$+ \frac{1}{2} \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, j) \right\rangle\!\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, j) \right|.$$

(B.11)

We can now deliberately break the relabeling symmetry in the second term by interchanging the labels j and $m+1$. That is, we choose to label the daughter parton that here caries the label j by $m+1$ instead. Note that this changes $\rho(\{\hat{p}, \hat{f}\}_{m+1})$ for any fixed value of $\{\hat{p}, \hat{f}\}_{m+1}$. However, the result of integrating $\rho(\{\hat{p}, \hat{f}\}_{m+1})$ against any measurement function (which must be symmetric under label interchanges) stays the same. After the interchange $j \leftrightarrow m+1$, the sum over j in the second term simply becomes a factor $m-1$. This gives

$$
\sum_{i=1}^{m} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big|
$$
$$
+ \frac{m-1}{2} \sum_{i=1}^{m} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big| \ .
\tag{B.12}
$$

This is

$$
\frac{m+1}{2} \sum_{i=1}^{m} \big|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\big|; \ .
\tag{B.13}
$$

We now consider the possibilities for flavors: $\{\hat{f}_i, \hat{f}_{m+1}\}$ could be $\{g, g\}$, $\{g, q\}$, $\{g, \bar{q}\}$, $\{q, g\}$, $\{\bar{q}, g\}$, $\{q, \bar{q}\}$, and $\{\bar{q}, q\}$, where q stands for a quark flavor, $\bar{q}$ stands for an antiquark flavor, and a q together with a $\bar{q}$ stands for a quark flavor and its corresponding antiquark flavor. Thus we can insert a factor

$$
\begin{aligned}
1 = &\; \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{g, g\}\big) \\
&+ \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{g, q\}\big) + \theta\big((\{\hat{f}_i, \hat{f}_{m+1}\} = \{g, \bar{q}\}\big) \\
&+ \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{q, g\}\big) + \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{\bar{q}, g\}\big) \\
&+ \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{q, \bar{q}\}\big) + \theta\big(\{\hat{f}_i, \hat{f}_{m+1}\} = \{\bar{q}, q\}\big) \ .
\end{aligned}
\tag{B.14}
$$

In the coefficients of the $\{g, q\}$ and $\{g, \bar{q}\}$ theta functions, we can further define the labeling by interchanging the labels i and $m+1$, so that $m+1$ is the label for the gluon. In the $\{\bar{q}, q\}$ term, we can further define the labeling by interchanging the labels i and $m+1$, so that $m+1$ is the label for the antiquark. With these label choices, our contribution is

$$
(m+1) \sum_{l=1}^{m} S_l(\{\hat{f}\}_{m+1}) \big|M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\big| \ ,
\tag{B.15}
$$

where

$$
S_l(\{\hat{f}\}_{m+1}) = \begin{cases} 1/2 \ , & l \in \{1, \dots, m\}, \ \hat{f}_l = \hat{f}_{m+1} = g \\ 1 \ , & l \in \{1, \dots, m\}, \ \hat{f}_l \neq g, \hat{f}_{m+1} = g \\ 0 \ , & l \in \{1, \dots, m\}, \ \hat{f}_l = g, \hat{f}_{m+1} \neq g \\ 1 \ , & l \in \{1, \dots, m\}, \ \hat{f}_l = q, \hat{f}_{m+1} = \bar{q} \\ 0 \ , & l \in \{1, \dots, m\}, \ \hat{f}_l = \bar{q}, \hat{f}_{m+1} = q \end{cases} \ .
\tag{B.16}
$$

We thus derive the factor $(m+1)$ and the factor $1/2$ for a final state $g \to g + g$ splitting.

Consider now the contribution (B.8). We can break the relabeling symmetry by interchanging the labels j and $m+1$. Then there are $m+1$ equal terms, giving

$$
(m+1) \sum_{l \in \{a, b\}} S_l(\{\hat{f}\}_{m+1}) \big|M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\big\rangle\big\langle M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\big| \ ,
\tag{B.17}
$$

where

$$S_l(\{\hat{f}\}_{m+1}) = 1 \quad l \in \{a, b\} \quad . \tag{B.18}$$

Consider, finally, the interference diagrams, eq. (B.10). We separate this into several terms according to the values of i and j,

$$\sum_{\substack{i,j \in \{a,b\} \\ i \neq j}} \sum_{k=1}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right|$$

$$+ \sum_{i \in \{a,b\}} \sum_{j=1}^{m} \sum_{\substack{k=1 \\ k \neq j}}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right|$$

$$+ \sum_{i=1}^{m} \sum_{j \in \{a,b\}} \sum_{\substack{k=1 \\ k \neq i}}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right|$$

$$+ \sum_{i \in \{a,b\}} \sum_{k=1}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; m+1, k) \right| \tag{B.19}$$

$$+ \sum_{j \in \{a,b\}} \sum_{k=1}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; m+1, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right|$$

$$+ \sum_{j=1}^{m} \sum_{\substack{k=1 \\ k \neq j}}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; m+1, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right|$$

$$+ \sum_{i=1}^{m} \sum_{\substack{k=1 \\ k \neq i}}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; m+1, k) \right|$$

$$+ \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \sum_{\substack{k=1 \\ k \neq i, k \neq j}}^{m+1} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, k) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, k) \right| \quad .$$

We relabel the indices, treating each term separately. In each case, we interchange $k \leftrightarrow m+1$. In the first term, this gives $m+1$ equal terms from the sum over k. In the second and third terms, this gives m equal terms from the sum over k. In the fourth through seventh terms, each term in the sum over k remains as one term. Finally, in the eighth term, there are $m-1$ equal terms from the sum over k. After relabeling, we have

$$(m+1) \sum_{\substack{i,j \in \{a,b\} \\ i \neq j}} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1) \right|$$

$$+ m \sum_{i \in \{a,b\}} \sum_{j=1}^{m} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1) \right|$$

$$+ m \sum_{i=1}^{m} \sum_{j \in \{a,b\}} \left| M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1) \right\rangle \left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1) \right|$$

$$+ \sum_{i \in \{a,b\}} \sum_{k=1}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; k, m+1)\right| \qquad \text{(B.20)}$$

$$+ \sum_{j \in \{a,b\}} \sum_{k=1}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; k, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1)\right|$$

$$+ \sum_{j=1}^{m} \sum_{\substack{k=1 \\ k \neq j}}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; k, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1)\right|$$

$$+ \sum_{i=1}^{m} \sum_{\substack{k=1 \\ k \neq i}}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; k, m+1)\right|$$

$$+ (m-1) \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; i, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; j, m+1)\right| \ .$$

These terms can be combined, after changing the names of some of the summation indices, and added to the contributions (B.15) and (B.17) to give the revised density operator, which we can call ρ_1,

$$\rho_1(\{\hat{p}, \hat{f}\}_{m+1}) =$$
$$(m+1) \sum_{l \in \{a,b,1,\ldots,m\}} S_l(\{\hat{f}\}_{m+1}) \left|M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\right|$$
$$+ (m+1) \sum_{\substack{l,k \in \{a,b,1,\ldots,m\} \\ l \neq k}}^{m} \left|M(\{\hat{p}, \hat{f}\}_{m+1}; l, m+1)\right\rangle\left\langle M(\{\hat{p}, \hat{f}\}_{m+1}; k, m+1)\right| \ .$$

$$\text{(B.21)}$$

The first line here contains the direct terms, eqs. (B.15) and (B.17), while the second line is the interference terms, eq. (B.20).

References

[1] G. Marchesini et al., *HERWIG: a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons*, Version 5.1 - April 1991, *Comput. Phys. Commun.* **67** (1992) 465;
S. Gieseke et al., *HERWIG++ 2.0 release note*, `hep-ph/0609306`.

[2] T. Sjöstrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, *Comput. Phys. Commun.* **82** (1994) 74;
T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [`hep-ph/0603175`].

[3] T. Sjöstrand, *A model for initial state parton showers*, *Phys. Lett.* **B 157** (1985) 321.

[4] T.D. Gottschalk, *Backwards evolved initial state parton showers*, *Nucl. Phys.* **B 277** (1986) 700.

[5] G. Marchesini and B.R. Webber, *Simulation of QCD jets including soft gluon interference*, *Nucl. Phys.* **B 238** (1984) 1;
R.K. Ellis, G. Marchesini and B.R. Webber, *Soft radiation in parton parton scattering*, *Nucl. Phys.* **B 286** (1987) 643 [*Erratum ibid.* **B 294** (1987) 1180].

[6] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244];
S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, *JHEP* **08** (2003) 007 [hep-ph/0305252].

[7] M. Krämer and D.E. Soper, *Next-to-leading order QCD calculations with parton showers. I: collinear singularities*, *Phys. Rev.* **D 69** (2004) 054019 [hep-ph/0306222];
D.E. Soper, *Next-to-leading order QCD calculations with parton showers. II: soft singularities*, *Phys. Rev.* **D 69** (2004) 054020 [hep-ph/0306268];
M. Krämer, S. Mrenna and D.E. Soper, *Next-to-leading order QCD jet production with parton showers and hadronization*, *Phys. Rev.* **D 73** (2006) 014022 [hep-ph/0509127].

[8] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[9] J.C. Collins, *Spin correlations in Monte Carlo event generators*, *Nucl. Phys.* **B 304** (1988) 794.

[10] I.G. Knowles, *A linear algorithm for calculating spin correlations in hadronic collisions*, *Comput. Phys. Commun.* **58** (1990) 271.

[11] P. Richardson, *Spin correlations in Monte Carlo simulations*, *JHEP* **11** (2001) 029 [hep-ph/0110108].

[12] T. Sjöstrand and P.Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, *Eur. Phys. J.* **C 39** (2005) 129 [hep-ph/0408302].

[13] Z. Nagy and D.E. Soper, *A new parton shower algorithm: shower evolution, matching at leading and next-to-leading order level*, hep-ph/0601021.

[14] S. Catani and M.H. Seymour, *A general algorithm for calculating jet cross sections in NLO QCD*, *Nucl. Phys.* **B 485** (1997) 291 [*Erratum ibid.* **B 510** (1998) 503] [hep-ph/9605323].

[15] Z. Trócsányi, private communication.

[16] W.T. Giele, D.A. Kosower and P.Z. Skands, *A simple shower and matching algorithm*, arXiv:0707.3652.

[17] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, arXiv:0709.1027;
M. Dinsdale, M. Ternick and S. Weinzierl, *Parton showers from the dipole formalism*, arXiv:0709.1026.

[18] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231];
L. Lönnblad, *Correcting the colour-dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046 [hep-ph/0112284];
N. Lavesson and L. Lönnblad, *W + jets matrix elements and the dipole cascade*, *JHEP* **07** (2005) 054 [hep-ph/0503293];
M.L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013 [hep-ph/0611129].

[19] Z. Nagy and D.E. Soper, *Matching parton showers to NLO computations*, *JHEP* **10** (2005) 024 [hep-ph/0503053].

[20] J.C. Collins, D.E. Soper and G. Sterman, *Factorization for short distance hadron - hadron scattering*, *Nucl. Phys.* **B 261** (1985) 104;
G.T. Bodwin, *Factorization of the Drell-Yan cross-section in perturbation theory*, *Phys. Rev.* **D 31** (1985) 2616 [*Erratum ibid.* **D 34** (1986) 3932];
J.C. Collins, D.E. Soper and G. Sterman, *Soft gluons and factorization*, *Nucl. Phys.* **B 308** (1988) 833.

[21] D.A. Kosower, *Antenna factorization of gauge-theory amplitudes*, *Phys. Rev.* **D 57** (1998) 5410 [hep-ph/9710213]; *Antenna factorization in strongly-ordered limits*, *Phys. Rev.* **D 71** (2005) 045016 [hep-ph/0311272].

[22] G. Somogyi and Z. Trócsányi, *A new subtraction scheme for computing QCD jet cross sections at next-to-leading order accuracy*, hep-ph/0609041; *A new subtraction scheme for computing QCD jet cross sections at next-to-leading order accuracy*, hep-ph/0609041.

[23] M.L. Mangano and S.J. Parke, *Multiparton amplitudes in gauge theories*, *Phys. Rept.* **200** (1991) 301 [hep-th/0509223].

[24] F.A. Berends and W.T. Giele, *Recursive calculations for processes with n gluons*, *Nucl. Phys.* **B 306** (1988) 759;
M.L. Mangano and S.J. Parke, *Multiparton amplitudes in gauge theories*, *Phys. Rept.* **200** (1991) 301 [hep-th/0509223].

[25] L.J. Dixon, *Calculating scattering amplitudes efficiently*, hep-ph/9601359.

[26] S. Catani and M. Grazzini, *Infrared factorization of tree level QCD amplitudes at the next-to-next-to-leading order and beyond*, *Nucl. Phys.* **B 570** (2000) 287 [hep-ph/9908523].

[27] J.E. Paton and H.-M. Chan, *Generalized Veneziano model with isospin*, *Nucl. Phys.* **B 10** (1969) 516;
F.A. Berends and W. Giele, *The six gluon process as an example of Weyl-Van Der Waerden spinor calculus*, *Nucl. Phys.* **B 294** (1987) 700;
M.L. Mangano, S.J. Parke and Z. Xu, *Duality and multi-gluon scattering*, *Nucl. Phys.* **B 298** (1988) 653;
M.L. Mangano, *The color structure of gluon emission*, *Nucl. Phys.* **B 309** (1988) 461.

[28] L. Lönnblad, *ARIADNE version 4: a program for simulation of QCD cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15.

[29] C.W. Bauer and M.D. Schwartz, *Event generation from effective field theory*, hep-ph/0607296.

[30] J.C. Collins, *Hard-scattering factorization with heavy quarks: a general treatment*, *Phys. Rev.* **D 58** (1998) 094002 [hep-ph/9806259].

[31] R. Doria, J. Frenkel and J.C. Taylor, *Counter example to nonabelian Bloch-Nordsieck theorem*, *Nucl. Phys.* **B 168** (1980) 93;
C.E. Carneiro, M. Day, J. Frenkel, J.C. Taylor and M.T. Thomaz, *Leading noncancelling infrared divergences in perturbative QCD*, *Nucl. Phys.* **B 183** (1981) 445.