Automated Uncertainty Variations in the Antenna Formalism of Parton Showers

Michael Deimetry

Supervisor:
Prof. Peter Skands

An honours thesis presented for the degree of
Bachelor of Science (Honours)

School of Physics and Astronomy
Faculty of Science
Monash University
Australia

December 6, 2022

Abstract

Currently, it is not possible to make predictions applicable to collider phenomenology directly from theories of the standard model. Instead, parton showers, a common approach which has been successful casts the proliferation of a hard process as an iterative Markov Chain Monte Carlo procedure which is progressed through the sampling of kernels derived from the underlying Quantum Field Theory. Traditionally, uncertainties of observables obtained from showering are determined by rerunning the shower with parameters modified at the extremes of the uncertainty bounds. The variation in the output places uncertainties on the observable of interest. Such an approach is computationally expensive, error-prone and unfeasible for many variations. Instead, we present a method for automating uncertainties in the so-called antenna formalism through the assignment of weights to event samples. We obtain formulae for colour-factor, non-singular term and renormalization-scale variations and implement the approach in Vincia. Our results verify the applicability of the procedure to the antenna formalism.

Contents

Chapter 1

Mathematical Prerequisites

An understanding of QCD is in large part understanding SU(N) and representations thereof. The theory presented here will appear in the Quantum Chromodynamics (QCD) Lagrangian since QCD is a Yang-Mills theory. As we will see, representations of $SU(3)$ are used to organize the mesons and baryons and represent the colour wavefunction of quarks and gluons. Additionally, the coupling strength of quarks and gluons is obtained directly from the Lie algebra $su(3)$ which characterize the gluon and quark colour states. To that end, it is instructive to take some time to develop these ideas here. The reader is not obliged to fully understand this chapter to proceed to the remainder of the thesis. However, It is the authors opinion that the additional effort to understand this chapter will add a richness to the remainder of the thesis. Mathematicians and physicists abide by different conventions, all developments here are with the physics convention unless stated otherwise. In passing the corrisponding mathematical convention will be introduced. The discussion presented here will be an amalgamation mainly of the treatments provided by [1] and [2]. For further mathematical or physics details, the reader is encouraged to visit these sources respectively. For a bare-bones approach specific to physics, the reader is referred to [3]. For brevity, proofs will be omitted.

1.1   Lie Groups and Algebras

We may avoid the full language of Lie groups and Algebras if we restrict our discussion to Matrix Lie groups and Algebras with a focus on SU(N).

Definition 1.1.1 *The group SU(N) is given by*

$$SU(N) = \{A \in M_N(\mathbb{C}) | A^\dagger A = I \ and \ \det A = 1\} \tag{1.1.1}$$

with the binary operation of matrix multiplication.

SU(N) is a manifold qualifying it as a Matrix Lie group. To understand Matrix Lie groups, we require the following definition.

Definition 1.1.2 (Convergence) *Given some sequence $\{A_m\}$ where $A_m \in M_N(\mathbb{C})$, the sequence is said to converge to matrix A if*

$$\lim_{m \to \infty} (A_m)_{ij} = A_{ij} \quad \forall \ 1 \leq i,j \leq N \tag{1.1.2}$$

where convergence is in the usual way as defined on $\mathbb{C}$.

A Matrix Lie group is therefore defined as follows.

Definition 1.1.3 (Matrix Lie Groups) *A Matrix Lie group, G, is a subgroup of $GL(N;\mathbb{C})$ where all convergent sequences, if invertible, are in $GL(N;\mathbb{C})$ which is the set of invertible matrices of (possibly) complex entries endowed with the group operation of matrix multiplication.*

Note, the subtle distinction in that, there may be sequences in the Lie group which converge to elements outside $GL(N;\mathbb{C})$ (singular matrices).

In anticipation of the intimate connection between Matrix Lie groups and algebras, the reader is reminded of the matrix exponential.

Definition 1.1.4 *For $X \in M_N(\mathbb{C})$, the matrix exponential is defined as*

$$\exp(X) = \sum_{m=0}^{\infty} \frac{X^m}{m!} \tag{1.1.3}$$

where X^0 is defined as the identity element.

Convergence of the matrix exponential will not be an issue since it can be shown that this is always convergent using the notion of convergence provided in definition 1.1.2.

We are now ready to define Matrix Lie algebras which can be associated with Lie groups.[1]

Definition 1.1.5 (Lie Algebra) *A finite dimensional Lie algebra $\mathfrak{g}$ is a finite dimensional vector space endowed with a bracket $[\cdot, \cdot]$ which is a map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ with the properties*

1. *Bilinearity in both slots*

2. *Skew Symmetry*

3. *The Jacobi Identity holds:*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \qquad \forall X, Y, Z \in \mathfrak{g} \tag{1.1.4}$$

The Lie algebra is called real(complex) if the underlying vector space is real(complex).

To each Lie group a corrisponding Lie algebra is assigned. Studying Lie groups can be aided by the study of the corrisponding Lie algebra. The association is as follows.

Proposition 1.1.1 *Let G be a matrix Lie group, then the corrisponding real Lie algebra, $\mathfrak{g}$, is the set*

$$\mathfrak{g} = \{X \in M_N(\mathbb{C}) | \exp(i\alpha X) \in G \, \forall \, \alpha \in \mathbb{R}\} \tag{1.1.5}$$

The bracket operation is the commutator $[X, Y] = XY - YX$.

In the mathematical literature, the i in the matrix exponential is omitted.

It is straightforward to verify that the matrix Lie algebra of SU(N) is

$$su(N) = \{A \in M_N(\mathbb{C}) | A^\dagger = A \text{ and } \operatorname{tr} A = 0\} \tag{1.1.6}$$

---

[1]A very common confusion often made in the physics literature is that a Matrix Lie algebra need not be an algebra i.e. it is not the commutation relations which make an algebra.

Throughout we will use the Pauli matrices as the generators (basis) for $su(2)$.[2]

A set of generators for $su(3)$ as proposed by [4] are the Gell-Mann matrices. The ordering of the basis is standard and denoted by λ_a where $a = 1, 2, ..., 8$. In the coming discussion, a second set of generators will also be used which are defined as

$$I_\pm = F_1 \pm iF_2 \qquad (1.1.7) \qquad\qquad U_\pm = F_6 \pm iF_7 \qquad (1.1.10)$$

$$I_3 = F_3 \qquad (1.1.8)$$
$$Y = \frac{2}{\sqrt{3}}F_8 \qquad (1.1.11)$$
$$V_\pm = F_4 \pm iF_5 \qquad (1.1.9)$$

where $F_a = \frac{1}{2}\lambda_a$. We will use this generating set in developing the fundamental representation. We finish this section by noting that the Casimirs of $SU(3)$ are given by

$$C_F = \frac{4}{3} \qquad (1.1.12) \qquad\qquad C_A = 3 \qquad (1.1.13)$$

and we note an additional constant which appears frequently, $T_R = \frac{1}{2}$. These arise from the relations

$$\text{Tr}\{F_aF_b\} = T_R\delta_{ab} \quad \text{where } a,b \in \{1, 2, ..., 8\} \qquad (1.1.14)$$

$$\sum_a (F_a)_{ij}(F_a)_{jk} = C_F\delta_{ik} \quad \text{where } a,b \in \{1, 2, ..., 8\} \text{ and } i,j,k \in \{1, 2, 3\} \qquad (1.1.15)$$

$$\sum_{c,d} f_{acd}f_{bcd} = C_A\delta_{ab} \quad \text{where } a,b,c,d \in \{1, 2, ..., 8\} \qquad (1.1.16)$$

where f are the structure constants of the Lie Algebra which arise due to the non-commutivity. Given some basis set $\{X_1, X_2, ..., X_s\}$ where s is the dimension of the vector space, the commutation relation between pairs of elements in the basis is given by[3]

$$[X_j, X_k] = 2i\sum_{l=1}^{s} f_{jkl}X_l \qquad (1.1.17)$$

where f_{jkl} are called structure constants.

1.2 Representations of Lie Groups and Algebras

Definition 1.2.1 (Lie group representation) *A representation of a Lie group is the pair (Π, V) where Π is a Lie group homomorphism* [4]

$$\Pi : G \to GL(V) \qquad (1.2.1)$$

Definition 1.2.2 (Lie algebra representation) *A Lie algebra representation is the pair (π, V) where π is a Lie algebra homomorphism* [5] [6]

$$\pi : \mathfrak{g} \to gl(V) = End(V) \qquad (1.2.3)$$

---

[2]The labelling is as follows: $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

[3]The term $2i$ is often also chosen to be i in some physics conventions. Mathematicians usually omit the i entirely.

[4]For our purposes, it suffices that Π simply be a group homomorphism.

[5]It is left to the reader to verify that the Lie algebra of $GL(V)$ is indeed $End(V)$.

[6]A Lie algebra homomorphism requires that π be linear and preserve the bracket operation

$$\pi([X, Y]) = [\pi(X), \pi(Y)] \qquad (1.2.2)$$

π. *The degree of a representation is given by the dimension of the vector space.*

Definition 1.2.3 (Irreducibility) *A representation is said to be simple or irreducible if there doesn't exist any non-trivial subspace which is itself a valid representation, only with a lower dimensional vector space.*

Little further will be said about irreducibility, however, it can be compared to the fundamental theorem of arithmetic where a representation is made up of irreducible representations. The degree of a representation is the dimension of the underlying vector space. For instance, a degree two representation is a doublet.

The representations of a Lie group and algebra aren't the same. We will consider a few representations of $su(2)$ and $su(3)$ which appear frequently in QCD.

Example 1.2.1 (Adjoint Representation) *For a Lie group G with Lie algebra $\mathfrak{g}$, the adjoint representation is given by $(Ad, \mathfrak{g})$ where $Ad : G \to GL(\mathfrak{g})$ given by*

$$Ad_A(X) = AXA^{-1} \tag{1.2.4}$$

where $A \in G$ and $X \in \mathfrak{g}$.[7] So the vector space in the Lie group representation is the corrisponding Lie algebra.

For some Lie algebra $\mathfrak{h}$, the adjoint representation is given by $(ad, \mathfrak{h})$ where $ad : \mathfrak{h} \to End(\mathfrak{h})$ given by

$$ad_X(Y) = [X, Y] \tag{1.2.5}$$

for $X, Y \in \mathfrak{h}$. Again, the vector space of the representation is the Lie algebra $\mathfrak{h}$.

Example 1.2.2 (The Standard Representation of Lie algebras and its dual) *The standard representation of a Matrix Lie Algebra, $\mathfrak{g}$, which is a subset of $GL(N; \mathbb{C})$ is given by $(\pi, \mathbb{C}^n)$ where $\pi(A) = A$ for some $A \in G$. The dual of the standard representation differs only in the Lie group homomorphism $\pi(A) = -A^T$.*

Example 1.2.3 (Irreducible Representations of $su(2)$) *For every $l = 0, 1/2, 1, ...$ there exists an irreducible representation, (π, V), of degree $2l + 1$. If we conveniently choose the basis given by $T_1 = \sigma_1/2, T_2 = \sigma_2/2, T_3 = \sigma_3/2$, then we may define operators*

$$L^{\pm} = \pi(T_1) \mp i\pi(T_2) \tag{1.2.6}$$

$$L_3 = \pi(T_3) \tag{1.2.7}$$

Then there exists an ordered basis of V, $\{u_0, u_1, ..., u_{2l}\}$, the operators are act as follows

$$L_3 v_j = (l - j)v_j \tag{1.2.8}$$

$$L^- v_j = \begin{cases} v_{j+1} & \text{if } j < 2l \\ 0 & \text{if } j = 2l \end{cases} \tag{1.2.9}$$

$$L^+ v_j = \begin{cases} j(2l + 1 - j)v_{j-1} & \text{if } j > 0 \\ 0 & \text{if } j = 0 \end{cases} \tag{1.2.10}$$

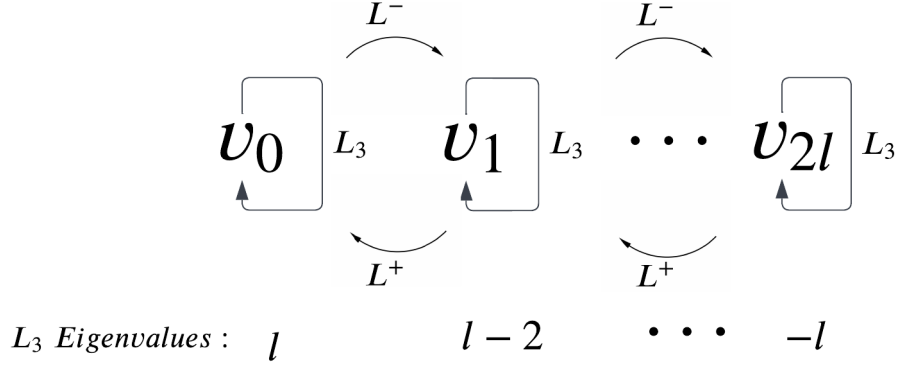[7] For Lie group SU(N), the map becomes $Ad_A(X) = AXA^\dagger$.

Figure 1.1. Irreducible representations of degree $2l + 1$ of $su(2)$ on vector space V with ordered basis $\{v_0, v_1, ..., v_{2l}\}$. $L^{\pm}$ are commonly referred to in physics as ladder operators.

A proof of this result can be found in [1], albeit with a different basis since he proves an analogous result for $sl(2; \mathbb{C})$. For the convention used in physics, the reader is referred to [5]. An illustration of the action of $su(2)$ on V is commonly illustrated using a diagram as shown in figure 1.1.[8] The Casmir operator

$$L^2 = \pi(T_1)^2 + \pi(T_2)^2 + \pi(T_3)^2 \tag{1.2.11}$$

determines the representation being used. That is,

$$L^2 v_j = l(l+1)v_j \tag{1.2.12}$$

The irreducible representations of $su(3)$ require a larger discussion and therefore receive a dedicated subsection.

1.2.1 Irreducible Representations of $su(3)$

Definition 1.2.4 *A weight, $\mu = (m_1, m_2)$ of a representation, (π, V) is a pair of eigenvalues which simultaneously diagonalize $\pi(I_3)$ and $\pi(Y)$. That is, there exists some $v \in V$ where*

$$\pi(I_3)v = m_1 v \tag{1.2.13}$$

$$\pi(Y)v = m_2 v \tag{1.2.14}$$

Definition 1.2.5 *The roots are the non-zero weights of the adjoint representation, that is, $\alpha = (\alpha_1, \alpha_2)$ is a root if there exists some $Z \in su(3)$ where*

$$[I_3, Z] = a_1 Z \tag{1.2.15}$$

$$[Y, Z] = a_2 Z \tag{1.2.16}$$

The roots of the basis elements are provided in table 1.1 and plotted in figure 1.2. All roots must be linear combinations of these. The roots can be used to form new weights as follows: If $\alpha = (a_1, a_2)$

---

[8]The reader may be reminded of angular momentum which is indeed an example of such a representation.
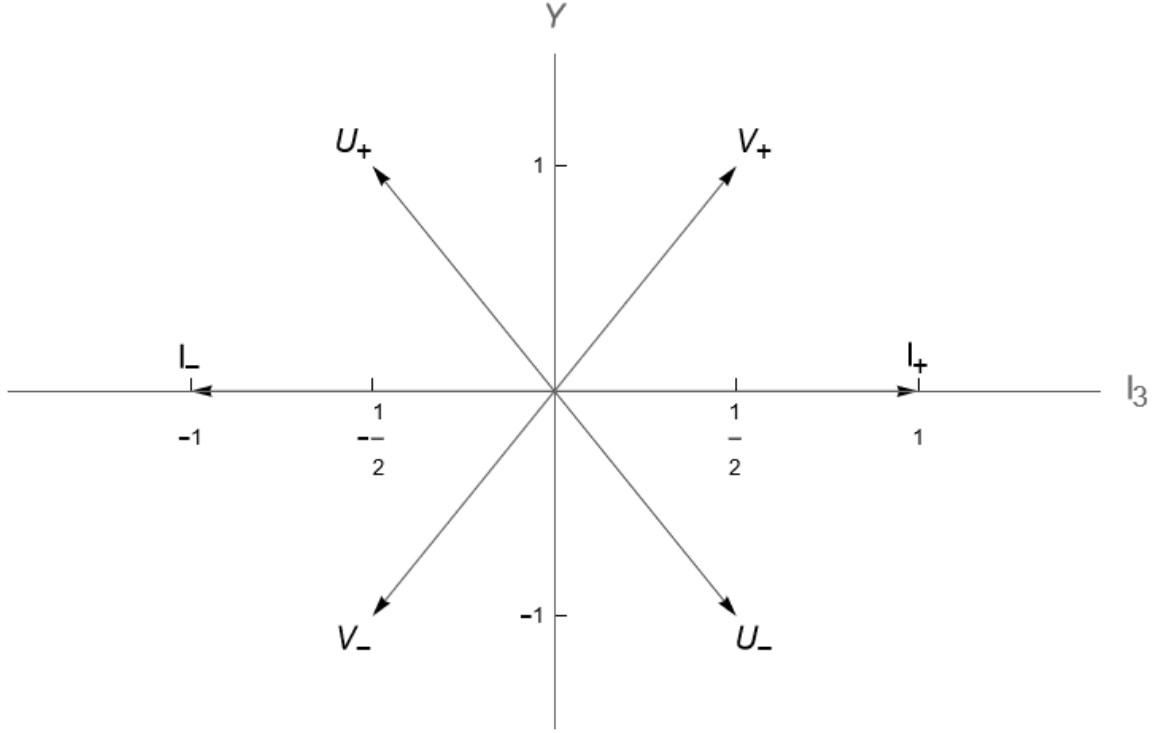
Figure 1.2. Roots of $su(3)$, tabulated in 1.1.

Table 1.1. Roots for generators of $su(3)$ Lie algebra as defined in equations (1.1.7) to (1.1.11)

	I_3	Y
$I_\pm$	± 1	0
$V_\pm$	$\pm 1/2$	± 1
$U_\pm$	$\mp 1/2$	± 1

is a root with root vector Z, then

$$\pi(I_3)\pi(Z)v = (m_1 + a_1)\pi(Z)v \tag{1.2.17}$$

$$\pi(Y)\pi(Z)v = (m_1 + a_2)\pi(Z)v \tag{1.2.18}$$

Definition 1.2.6 *A weight, μ_1, is said to be the highest weight of a representation if $\mu_1 + \alpha = 0$ for all α corrisponding to a root of V_+, U_+ or I_+.*

Theorem 1.2.1 *For every ordered pair $\mu = (m_1, m_2)$ which is the $\mathbb{N}$-span of $(\frac{1}{2}, \frac{1}{3})$ and $(0, \frac{2}{3})$, there exists an irreducible representation, (π_μ, V) with highest weight μ. The representations with highest weights $(\frac{1}{2}, \frac{1}{3})$ and $(0, \frac{2}{3})$ are dubbed the fundamental and anti-fundamental representations respectively.*

In the physics literature, the representations are referred to by their degrees.[9]

---

[9]This leaves room for ambiguity since representations of different highest weight may share the same degree.

(a) 3 Representation   (b) $\overline{3}$ Representation

Figure 1.3. Fundamental and anti-fundamental representations of $su(3)$ weight diagrams. The weights for each basis element is plotted on the $Y - I_3$ plane.

If we denote the standard representation and its dual of some Lie algebra, $\mathfrak{g}$, with (π_1, V_1) and (π_2, V_2) respectively. Then consider the representation given by (π, V) where
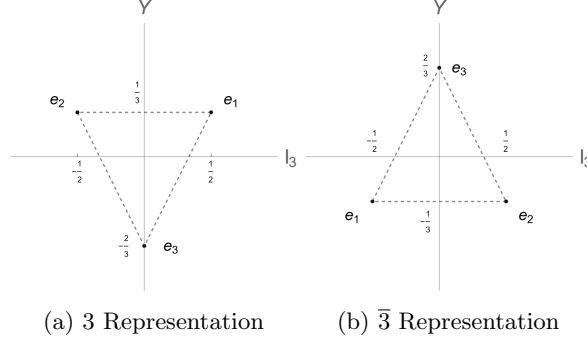
$$V = V_1^{\otimes m_1} \otimes V_2^{\otimes m_2} \tag{1.2.19}$$

and

$$\pi(X) = (\underbrace{\pi_1 \otimes \cdots \otimes \pi_1}_{m_1} \otimes \underbrace{\pi_2 \otimes \cdots \otimes \pi_2}_{m_2})(X) = \underbrace{\pi(X) \otimes I \otimes \cdots \otimes I}_{m_1} \otimes \underbrace{I \otimes \cdots \otimes I}_{m_2}$$
$$+ \underbrace{I \otimes \pi(X) \otimes \cdots \otimes I}_{m_1} \otimes \underbrace{I \otimes \cdots \otimes I}_{m_2}$$
$$\vdots$$
$$+ \underbrace{I \otimes \cdots \otimes I \otimes \pi(X)}_{m_1} \otimes \underbrace{I \otimes \cdots \otimes I}_{m_2}$$
$$\vdots \tag{1.2.20}$$

where $X \in \mathfrak{g}$. The additional terms mimic those shown here but applied to the elements acting on $V^{\otimes m_2}$ and with the homomorphism π_2. If $v_1 \in V_1$ and $v_2 \in V_2$ are the highest weight vectors corrisponding to the weights $a = (a_1, a_2), b = (b_1, b_2)$ respectively, then observe that given $v = \underbrace{v_1 \otimes \cdots \otimes v_1}_{m_1} \otimes \underbrace{v_2 \otimes \cdots \otimes v_2}_{m_2}$ then

$$\pi(I_3)v = (a_1 m_1 + b_1 m_2)\, v \tag{1.2.21}$$
$$\pi(Y)v = (a_2 m_1 + b_2 m_2)\, v \tag{1.2.22}$$

so the weights are the $\mathbb{N}$-span of a and b. Using this construction, if we can find representations with highest weights $(\frac{1}{2}, \frac{1}{3})$ and $(0, \frac{2}{3})$, then we may build up all the representations of $su(3)$, within which we may find the irreducible representations.

We will choose the fundamental representation to be the standard representation. It is left as an exercise to the reader to verify that the weight vectors of this representation are the standard basis e_1, e_2, e_3 where e_i where all entries are zero except the ith which contains one. This representation has weights $(1/2, 1/3)$, $(-1/2, 1/3)$ and $(0, -2/3)$ respectively. The weights for the dual representation are given by $(2/3, 0)$, $(1/2, 1/3)$ and $(-1/2, -1/3)$. The weights for each basis element have been plotted in a so-called weight diagram in figure 1.3.
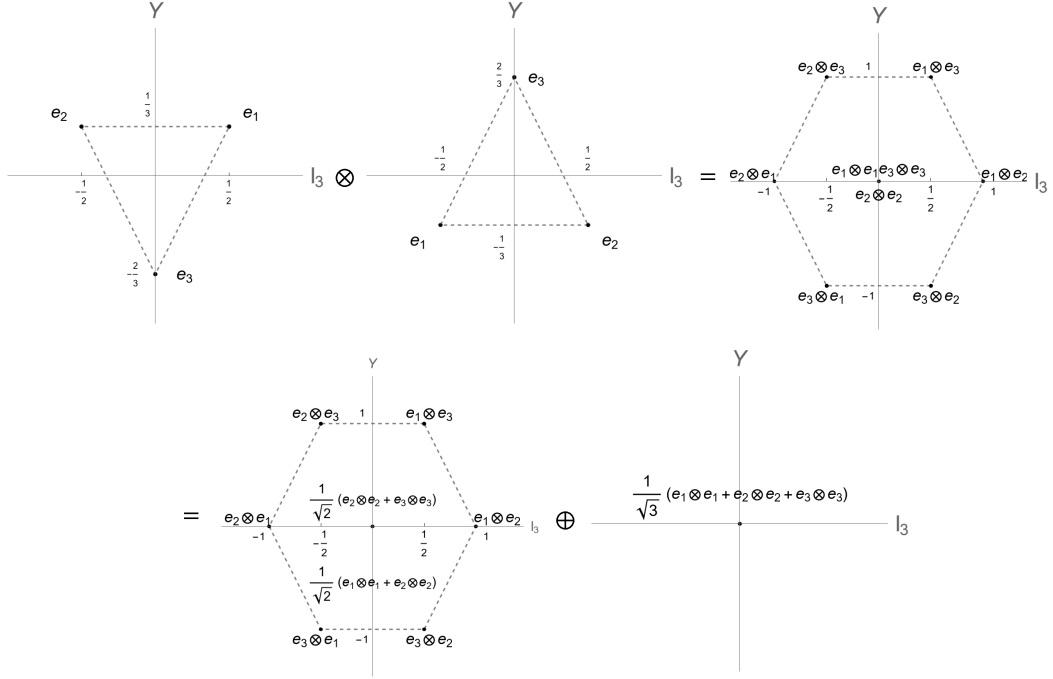
Figure 1.4. Tensor product construction of fundamental and anti-fundamental representations which decomposes into the the adjoint/octet and trivial/singlet representations.

Using the tensor product construction, we can build up an explicit representations of $su(3)$. We consider a final example which contains the adjoint representation. This is obtained as follows

$$3 \otimes \bar{3} = 9 = 8 \oplus 1 \tag{1.2.23}$$

The adjoint representation/octet, 8, can be explicitly obtained as is done by [1] in section 6.5. These can simply be obtained from lookups of Clebsch–Gordan coefficients.[10] The weight diagram for equation (1.2.23) is shown in figure 1.4.

Our study of the Lie algebras of $SU(2)$ and $SU(3)$ apply to representation of the Lie groups themselves. A representation of a Lie algebra uniquely determines that of the Lie groups by [11]

$$\Pi(e^{i\alpha X}) = e^{i\alpha\pi(X)} \tag{1.2.24}$$

where X is an element of the corrisponding Lie algebra. To be clear the Lie group representations are characterized by the Lie algebra representation.

---

[10]There are certainly many other ways of doing this as well, for instance through the use of character theory.

[11]This is not generally true, this only applies to simply connected groups which $SU(2)$ and $SU(3)$ are a part. This result can be found in [1] in Theorem 5.6 and Corollary 5.7.

Chapter 2

Introduction

Before delving into the project, we cover prerequisites which will be essential for understanding the contents of this thesis as well as providing a general introduction to the various fields intertwined here. We begin with an overview of parton showers and outline the entire process at a high-level. We then examine the experimental context where we will introduce the aims of this these. This is followed by a brief revision of Quantum Field Theory (QFT) followed by a discussion of the components of the standard model which pertain to QCD which is underpinned by QFT. We will conclude this section with a discussion of Markov Chain Monte Carlo (MCMC) methods. The reader is assumed to be familiar with the concepts in particle physics at a rudimentary level. If not, the reader is advised to begin with §2.3 and §2.4 and return to §2.1.

2.1 Parton Showers and Event Generators

There exists a large gap between the Lagrangians of QFT and the phenomenology of particle colliders. It is usually the physicist's intention to study some hard process i.e. a process occurring at some high energy scale. At such scales, the process may be calculable perturbatively up to Leading Order (LO), Next To Leading Order (NLO) or beyond but not far beyond. Comparison of, for instance, cross-sections with data collected at a collider bears no resemblance with the hard interaction on paper. Instead, the experimentalist will see a large variety of particles. The discrepancy is for many reasons; the event taking place at the collider occurs over a large range of energy scales which result in the emission of bremsstrahlung polluting the detector signals. This is the tendency for accelerated charged to emit radiation (derived classically in [6] and quantum mechanically in [7] for Quantum Electrodynamics (QED)). From QCD alone, pollutants will include hadrons containing the once coloured partons due to bremsstrahlung and newly created partons in the Cornell potential. Colour is an additional degree of freedom partons exhibit. These potentials will exist amongst final state partons as well as the beam remnant which may contain unbound coloured states. Possible fates of some of these short lived hadrons may be further decay. This proliferation of final state particles is dubbed showering. While our focus will be on parton showers i.e. showers containing only partons, where possible, multiple QFT's will contribute within a single event. It should be noted that this proliferation isn't limited to the products of the hard process but also prior to the hard interaction. Additionally, passing luminous beams past one another will certainly result in many interactions, each of which giving rise to their own proliferation. The showering due to each interaction will certainly evolve differently resulting in several hard processes within a single event.

While there are several approaches in development to model these effects, none are as developed and successful as showering produced by so-called event generators. For instance, an alternative
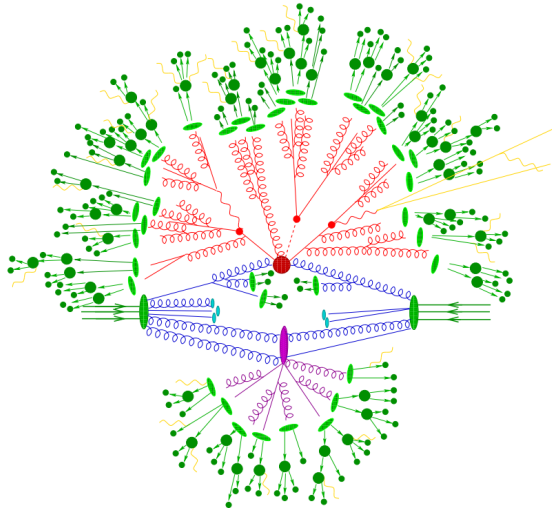
Figure 2.1. Illustration of a possible generated event. The hard process is marked with a red circle at the centre. Blue radiation is ISR while red radiation is FSR, a part of which is QED bremsstrahlung in yellow. The purple blob indicates a second hard process implying a MPI. The light green blobs indicate the hadronization while green emissions blobs subsequent hadron decays. Obtained from [8].

to event generators is lattice QCD. However, lattice QCD isn't compatible with any hadronization model. Hadronization is simply the grouping of quarks into colourless bound states. This thesis will focus on the Pythia event generator developed at Monash. An example of such an event is illustrated in figure 2.1. A parton shower begins with the calculation of a hard process through perturbative means, usually to N^2LO. The showering is parameterized through three variables for each parton, these variables characterize the transverse and longitudinal evolution of the showering while the third is usually an emission angle. Subsequent final states are determined through draws from splitting kernels obtained from the underlying QFT. Each emission redistributes the momentum of the final state to conserve the relevant quantities through so-called recoil schemes to ensure on shell partons. The implication is then that no two runs will produce the same final state. If a run with some hard process is desired, then seldom will a run randomly contain such an event with the many possibilities at hand. Naively, events would be generated and scrapped until this is to occur which is certainly inefficient. Instead, the shower is then evolved forwards and backwards from the hard process producing FSR and ISR respectively. Due to the virtually of the propagators, these showering schemes are commonly referred to as timelike and spacelike respectively. These showers therefore run from the hard process scale downward to the beam energies and the handronization scale. These showers are usually interleaved with one another as well as with those of other hard processes if several hard interactions exist. Finally, the showering is halted at the confinement/hadronization scale where a phenomenological model determines the colourless final state. Confinement is a conjecture where at low energy scales, partons cannot exist alone but must be grouped to have no overall colour. While showering, partons are assigned unique colouring for simplicity, this is taken into account through Colour Reconnections (CR)[9, 10] to restore the would be confining potentials. The interested reader is referred to §3.2.2 for a discussion of the handling of colour. Finally, a handronization model takes hold. Pythia 8.3 currently uses the Lund string model[11, 12].

2.2 Experimental Motivation and Validation of Parton Showers

Before considering the details of event generation via parton showering, we present a compelling experimental validation. In pursuit of this task, we will discuss the experimental considerations necessary for comparing experiment with prediction which will further motivate the interest in event generators such as Pythia. The discussion draws form [13], section 16.

An event generator, to become applicable to data from a collider needs to do more than provide a method for proliferating a hard process as emphasised in §2.1. An event generator must be capable of producing an event from which physically meaningful experimental observables are attainable. For instance, a prediction is meaningless if made in terms of the predictions of a particular theoretical framework e.g. the virtuality of a propagator which is an interpretation in the context of QFT. If a prediction is made in terms of a well-defined observable on detectable particles, kinematic phase-space cuts must be made to match the capability of the detector. In the context of QCD, the final-state must contain only colourless particles. Hence, the necessity of a hadronization model is necessary and provides strong motivations for considering parton showers. In the process illustrated in figure 2.2 of $Z^0/\gamma^* \to \ell^+\ell^-$, making predictions in terms of a specific process is not useful since it is not possible to infer a priori the process responsible for the pair production i.e. Z^0 or γ^*. If the process involving Z^0 is desired, then the prediction can be made around the Z^0-mass resonance where the propagator diverges. Figure 2.2 provides a comparison between the main event generators available; Pythia 8.145[14, 15], Sherpa 1.2.3[16, 17] and Herwig++ 2.5.0[18–20]. The results were collected by the collider at Fermilab during 1992–1995 at scales around m_Z from $p\bar{p}$ collisions at Center-of-Mass (COM) of 1.8TeV. Indeed the event generators seem to resemble the experimental data. However, we encounter a fundamental limitation of event generators currently, namely the absence of error bars, hence, it is difficult to say whether deviations of predictions from experimental data is a failure of the underlying predictive model or negligible. This is the ultimate aim of this thesis and will be addressed in §4.1. We will present a efficient method for estimating uncertainties on predictions obtained from parton showers. We note that this thesis will focus on a plugin for Pythia dubbed Vincia which uses a different showering formalism which we will introduce in §3.2.3.

2.3 Quantum Field Theory (QFT)

Each observable event in nature usually admits multiple theories, each applicable in a limited range. The Lagrangian formulation is simply a restatement of Newtonian mechanics and may describe a system with a finite number of degrees of freedom. Such a theory may describe a non-relativistic system where non-locality can be ignored. A classical field theory formulates the Lagrangian in covariant language in terms of fields which assign some quantity to each spacetime point. At the cost of additional complexity, this resolves the issues encountered with Lagrangian mechanics. Finally, the incorporation of quantum mechanical effects specifies a QFT. A fitting description of QFT is as follows:

> It is literally the language in which the laws of Nature are written.

(David Tong)

The transition from a classical field theory to QFT is somewhat analogous to that from Hamiltonian mechanics to quantum mechanics. In that vein, we begin our discussion with classical field theory.

A classical field theory formulates the degrees of freedom of a system in terms of fields. Rather than a Lagrangian, the useful quantity becomes a Lagrangian density, $\mathcal{L}$, which carries units $[J/V]$ and integrates to the Lagrangian

$$L = \int d^3x \; \mathcal{L} \tag{2.3.1}$$
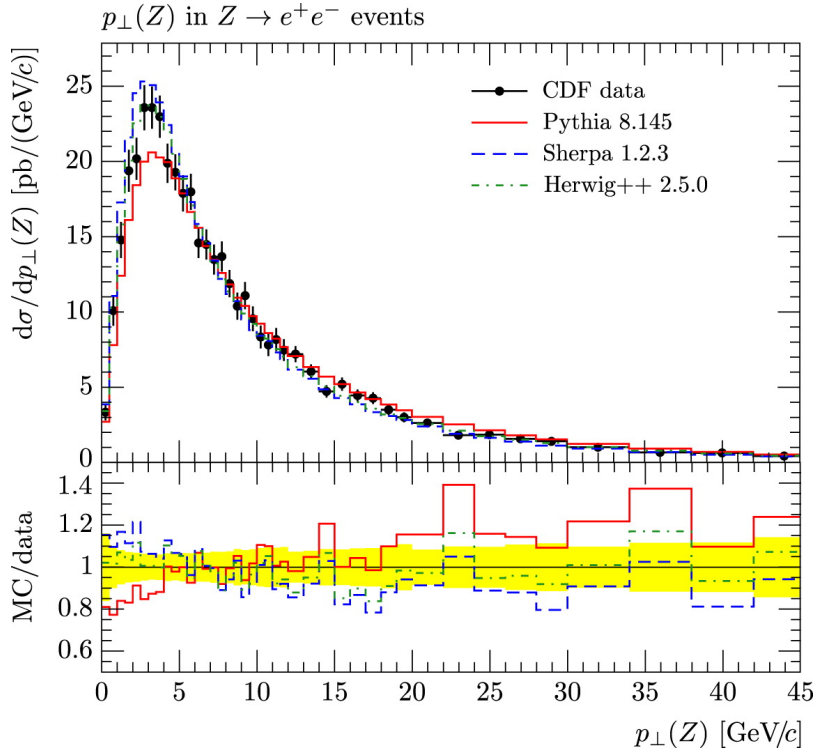
Figure 2.2. Comparison of experimental data obtained from $Z \to e^+e^-$ at Fermilab as compared with various event generators. Obtained from [21].

where $\mathcal{L} = \mathcal{L}(\varphi^a(x), \partial_\mu \varphi^a(x))$. The Lagrangian density is often referred to simply as the Lagrangian. Hamilton's principle then applies to L which states

$$\delta S = \int dt \ \delta L = 0 \tag{2.3.2}$$

which is satisfied if the fields abide by the Euler-Lagrange equations formulated in terms of the fields and their derivatives:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \varphi_a} = 0 \tag{2.3.3}$$

A classical field theoretic equivalent of Noether's Theorem is as follows:

Theorem 2.3.1 (Noether's Theorem) *Every differentiable continuous one-parameter family symmetry of the system gives rise to a conserved current, dubbed the Noether current given by*

$$j^\mu = \left(\sum_a \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} \delta \varphi_a \right) - F^\mu \tag{2.3.4}$$

provided the first variation in the Lagrangian is by a total derivative i.e. $\delta \mathcal{L} = \partial_\mu F^\mu$. The corrisponding (locally) conserved charge is given by

$$Q = \int d^3x \ j^0 \tag{2.3.5}$$

which satisfies the continuity equation

$$\partial_\mu j^\mu = 0 \tag{2.3.6}$$

The proof is widely available (*e.g.* see [22], §3.5.1, [23], §1.3.1). A conserved quantity in a field theory has a continuity equation implying that the conserved current is conserved locally. Symmetries in physics then are of paramount importance and will be explored in the context of QCD in §2.4 where we will find that it is invariant under $SU(3)$.

Local symmetries imply the transformation is dependent on each spacetime point. A subset of these are global symmetries which are independent of each spacetime point. If the transformation is only applied to the fields, $\varphi^a(x) \to \varphi^{a'}(x)$, then the symmetry is internal. An example of which and prominent in QFT's of the standard model are gauge theories which are transformations of the fields specified by a Lie group. A Yang-Mills theory is one in which the Lie group is $SU(N)$ for some positive integer N.

When the transformation acts on the coordinate itself, then the symmetry is one of spacetime. Since each QFT must be relativistic, every set of field equations is required to remain invariant under representations of the Poincaré group.

The Lorentz group is given by all matrices which satisfy

$$\Lambda^{\mu}{}_{\sigma}\eta^{\sigma\tau}\Lambda^{\nu}{}_{\tau} = \eta^{\mu\nu} \tag{2.3.7}$$

which is denoted by $O(3;1)$ while elements with determinant 1 by $SO(3;1)$. Usually, of particular interest is the proper orthochronous subgroup given by

$$SO(3;1)^{\uparrow}_{+} = \{\Lambda | \Lambda \in O(3;1) \text{ where } \det\Lambda = 1 \ \Lambda^0{}_0 \geq 1\} \tag{2.3.8}$$

The Poincaré group is the Lorentz group with spacetime translations. However, we focus on representations of the subgroup of the Poincaré group which contains only the proper orthochronous Lorentz group.

Given some representations of the Lorentz group, (Π, V),[1] then the field transforms under the Poincaré group as as

$$\phi(x) \to \phi'(x) = \Pi(\Lambda)\varphi(\Lambda^{-1}(x - a)) \tag{2.3.9}$$

where $\Lambda \in SO(3;1)^{\uparrow}_{+}$, a defines some spacetime translation i.e. Any element of the Poincaré acts on a spacetime point by $\Lambda x + a$. Recall, a Lie algebra representation is a homomorphism which allows for the representations to be fully characterized by the corresponding Lie Algebra. For the Poincaré group, this is given by[24]

$$[P_{\mu}, P_{\nu}] = 0 \tag{2.3.10}$$

$$[M_{\mu\nu}, P_{\rho}] = i(\eta_{\mu\rho}P_{\nu} - \eta_{\nu\rho}P_{\mu}) \tag{2.3.11}$$

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}) \tag{2.3.12}$$

The generators of the Lorentz group alone are given by M and the addition of the generators of translations, P, completes the Poincaré group Lie algebra. Any element of the Poincaré group can then be generated using

$$\exp(ia_{\mu}P^{\mu})\exp\left(\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right) \tag{2.3.13}$$

An example of generating representations of the Poincaré group can be found in [23] chapter 4.[2]

---

[1] The vector space in this representation are the fields φ.

[2] In some conventions, the i is omitted and the commutation relation as is done in [23].

Analogously, the Hamiltonian density for fields is defined as

$$\mathcal{H} = \pi^a(x)\dot{\varphi}_a(x)|_{\dot{\varphi}_a(x)=\dot{\varphi}_a(\pi^a(x),x)} - \mathcal{L}(x) \tag{2.3.14}$$

where $\pi^a(x) = (\partial/\partial\dot{\varphi}_a(x))\mathcal{L}$ is the conjugate momentum and the Hamiltonian is given by $H = \int d^3x\mathcal{H}$.

We proceed from classical field theory to QFT by promoting poisson brackets of the fields to commutation relations where now the fields are operators. As is in quantum mechanics, there are QFT analogues of the Schrödinger, Heisenberg and interaction picture. Within the interaction picture, the hamiltonian is divided into two terms

$$H = H_0 + H_{\text{per}} \tag{2.3.15}$$

with the operators and states given by[23]

$$|\varphi\rangle = \exp(iH_0t)|\varphi\rangle_S \tag{2.3.16}$$
$$\mathcal{O}_I = \exp(iH_0t)\mathcal{O}_S\exp(-iH_0t) \tag{2.3.17}$$

where I, S subscripts represent quantities in the interaction and Schrödinger pictures respectively. Although the separation of the Hamiltonian in this way is arbitrary, it becomes beneficial when the theory without the perturbing Hamiltonian is known. Within the interaction picture, the Schrödinger equation becomes

$$i\frac{d|\varphi\rangle_I}{dt} = H_I(t)|\varphi\rangle_I \tag{2.3.18}$$

where $H_I = (H_{\text{per}})_I$. Since $|\varphi\rangle_I(t) = U(t,t_0)|\varphi\rangle_I(t_1)$, 2.3.18 is also a differential equation for the time evolution operator U. This result is familiar from quantum mechanics with the important distinction that $[H_I(t_0), H_I(t_1)] \neq 0$ when $t_0 \neq t_1$. Fortunately, the solution is given by Dyson's formula

$$U(t,t_0) = T\exp\left(-i\int_{t_0}^t H_I(t')dt'\right) \tag{2.3.19}$$

where T is the time ordering operator given by

$$T(\mathcal{O}_1(t_1)\mathcal{O}_1(t_2)) = \begin{cases} \mathcal{O}_1(t_1)\mathcal{O}_1(t_2) & t_1 > t_2 \\ \mathcal{O}_1(t_2)\mathcal{O}_1(t_1) & t_2 > t_1 \end{cases} \tag{2.3.20}$$

Unfortunately, the complexity of the result is hidden away by notation. When 2.3.19 is expanded, it reads

$$U(t,t_0) = \mathbb{1} + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!}\int_{t_0}^t dt_1\int_{t_0}^t dt_2\ldots\int_{t_0}^t dt_n T[H_I(t_1)H_I(t_2)\ldots H_I(t_n)] \tag{2.3.21}$$

Given some initial state $|i\rangle$, then it is evolved to some later state using $U(t,t_0)$. So the projection of this later state onto a state of interest $|f\rangle$ is given by

$$\langle f| U(t,t_0)|i\rangle \tag{2.3.22}$$

If we assume that the eigenstates of H_0 are unchanged by the perturbative component when infinitely separated in time, then

$$\lim_{t_0\to-\infty}\lim_{t\to\infty}\langle f| U(t,t_0)|i\rangle \equiv \langle f| S|i\rangle \tag{2.3.23}$$

where S is dubbed the S-matrix. That is, we assume that the initial and final states begin and end infinitely separated with the interaction becoming relevant only when in proximity of each other. 2.3.23 is the essential component we require to obtain predictions from QFT's. Each term within S contains some possible processes by which $|i\rangle$ may evolve, each of these may be represented pictorially in so-called Feynman diagrams. The fields within H_I dictate the "building blocks" of the allowed interactions dubbed vertices. These building blocks can be combined to form arbitrarily complicated processes which will in turn correspond to a later term within Dyson's formula. Each theory is assigned a set of so-called Feynman rules which are applied to a Feynman diagram to obtain so-called matrix amplitudes, denoted by $i\mathcal{A}$ and related to equation (2.3.23) by

$$\mathcal{M} = \langle f | S | i \rangle = i\mathcal{A}(2\pi)^4 \delta^{(4)}(p_f - p_i) \tag{2.3.24}$$

where p_i, p_f denote the incoming and outgoing 4-momenta respectively. We may begin producing predictions from the theory such as cross-sections by

$$\sigma \propto \int d\Phi |\mathcal{M}|^2 \tag{2.3.25}$$

where the integral is over all allowed incoming and outgoing momenta while ensuring on-shell particles, this is a phase space. The first diagram viable diagram with the fewest vertices is dubbed the LO. Diagrams with a single additional vertex are dubbed NLO and so on. Now $H_I \propto g$ where g is a coupling constant which determines the strength of the interactions. So Dyson's formula is simply an expansion in powers of g. If g is sufficiently small, only terms near LO are relevant. If g is large, Dyson's formula will diverge. This regime is dubbed non-perturbative.

2.4 Quantum chromodynamics (QCD)

2.4.1 The Eightfold way and the Beginnings of Colour

The subatomic fermionic particles which participate in the strong interaction are so-called quarks. A quark may be categorized by the charge of QCD it carries and its flavour which dictates properties such as mass and electric charge. The charge of QCD is dubbed color, a misleading name since this is unrelated to the typical definition. As we will see, color and electric charge have fundamental differences despite playing analogous roles in their respective theories. The flavours of quark (to date) are given in table 2.1 as well as some common properties which we will explore in the coming discussion.

Table 2.1. The top quark mass is based on direct measurements which differs slightly when compared obtained using different approaches. All these particles have spin $\frac{1}{2}$, have baryon number $\frac{1}{3}$ and positive parity. All particles have strangeness 0 except the strange quark which has strangeness -1. Values obtained from [25].

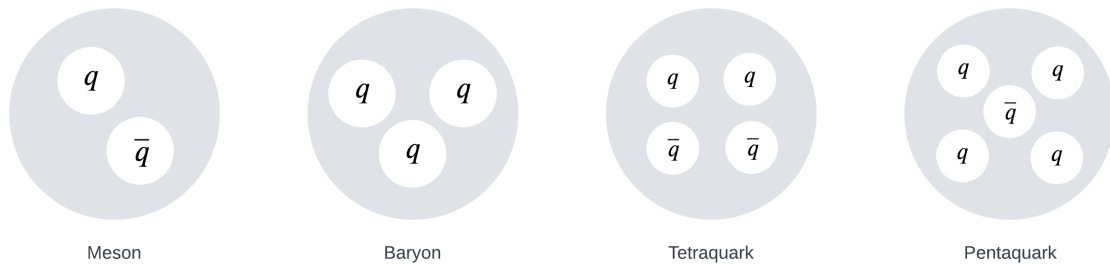| flavour	mass	electric charge (e)
up (u)	$2.16^{+0.49}_{-0.26}$ MeV	2/3
down (d)	$4.67^{+0.48}_{-0.17}$ MeV	$-1/3$
strange (s)	$93.4^{+8.6}_{-3.4}$ MeV	$-1/3$
charm (c)	1.27 ± 0.02 GeV	2/3
bottom (b)	$4.18^{+0.03}_{-0.02}$ GeV	$-1/3$
top (t)	172.69 ± 0.3 GeV	2/3

Figure 2.3. Categorizations of bound states of quarks. All have been observed. Hexaquarks and Heptaquarks are theorised to exist but have not been observed to date.

Combinations of quarks within bound states form hadrons which are categorized based on the number of valance quarks contained, some possibilities are shown in figure 2.3. There are theorised bound states with more valance quarks, however, none have been found.

Hadrons which aren't mesons or baryons are dubbed the exotic sort since they aren't encompassed in the quark model which we develop now.

A meson is composed of a quark-antiquark pair. The need for this will become clear once we discuss color in more detail. Each lepton has a corrisponding antilepton. A convention is chosen to fix which of the pair is the lepton and which is the antilepton. The flavour of the antiquark is usually denoted with a bar overtop, sometimes as with electrons, this is denoted by an opposing electric charge i.e. e^-e^+. The eightfold way as proposed by [4] arranges the mesons and baryons containing only up, down and strange quarks systematically. This is done through the standard representation of $SU(3)$. The corrisponding vector space $\mathbb{C}^3$ is dubbed flavour space. The standard basis elements are assigned the three quark flavours as follows

$$u = e_1 \quad (2.4.1) \qquad d = e_2 \quad (2.4.2) \qquad s = e_3 \quad (2.4.3)$$

The antiquarks are represented by the anti-fundamental representation. So a meson is simply $q \otimes \bar{q}$. Then each weight of the nonet defines each of the mesons, the basis elements corrisponding to $(0,0)$ weights are redefined to make clear the singlet state η' and to corrispond to the π^0/η states. These are given by

$$|\eta'\rangle = \frac{1}{\sqrt{3}}(u \otimes \bar{u} + d \otimes \bar{d} + s \otimes \bar{s}) \tag{2.4.4}$$

$$|\pi^0\rangle = \frac{1}{\sqrt{2}}(u \otimes \bar{u} - d \otimes \bar{d}) \tag{2.4.5}$$

$$|\eta\rangle = \frac{1}{\sqrt{6}}(u \otimes \bar{u} + d \otimes \bar{d} - 2s \otimes \bar{s}) \tag{2.4.6}$$

Particles with the spin aligned will have spin 1 while those with spin anti-aligned will have spin zero. These are presented in a nonet in figure 2.4. Each row corresponds to mesons of fixed strangeness while each diagonal corresponds to a fixed electric charge. This is because $\bar{s}$ contributes positively to the Y-component of the weight while an s contributes negatively. If there are no strange quarks, the Y component of the weight remains 0 since a combination of the other flavours will cancel the Y-component. The right-most diagonal correspond to u-quarks pairs with an anti s or d. The left-most diagonal is the opposite and the centre either contains no up quarks or only up quarks. Since the up quark is the only of these three flavours which has a different charge as seen from table 2.1, the emergence of constant charge along diagonals is therefore clear.

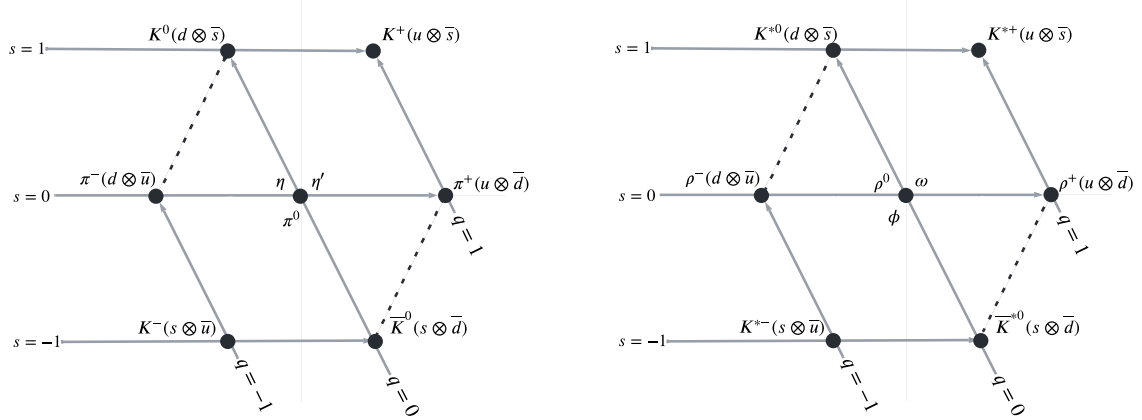The spin one mesons with weights $(0,0)$ aren't accounted for using this formulation. They are

Figure 2.4. Nonet representation of $SU(3)$ of possible meson combinations containing up, down and strange quarks. The left diagram are scalar mesons. The right are vector mesons. Diagonals corrispond to mesons of constant electric charge and along rows are mesns of constant strangeness. The mesons in the centres have flavour wavefunctions equations (2.4.4) to (2.4.6) and equations (2.4.7) to (2.4.9).

still presented in a nonet although this is not strictly true anymore. These mesons are given by

$$\left|\rho^0\right\rangle = \frac{1}{\sqrt{2}}(u \otimes \overline{u} - d \otimes \overline{d}) \tag{2.4.7}$$

$$\left|\omega\right\rangle = \frac{1}{\sqrt{2}}(u \otimes \overline{u} + d \otimes \overline{d}) \tag{2.4.8}$$

$$\left|\phi\right\rangle = s \otimes \overline{s} \tag{2.4.9}$$

It should be noted, these are only the flavour components of the meson wavefunctions.

The same construction can be applied to baryons albeit not every weight corresponds to an observed baryon. Now since we are combining 3 quarks, they decompose into irreducibles as[3]

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1 \tag{2.4.10}$$

Only the decuplet and an octet are used. The octet comprises of baryons with the spins of two quarks aligned and opposite to the third. The decuplet comprises of baryons with spins entirely aligned.

From this we see the need for an additional degree of freedom. Since a fermion must obey fermi-dirac statistics, we require that no two quarks be in exactly the same state. A clear contradiction is exhibited in particular by the Δ^{++} baryon which appears to be symmetric under the exchange of any of the constituent quarks.[4] Indeed, the antisymmetry of the wavefunction is due to an additional degree of freedom which we briefly introduced as colour. Unlike QED, which has a single electric charge, there are three colors which are labelled red, green and blue. QCD is invariant under $SU(3)$ [5] which is applied to the theory in the standard fundamental representation where the three colours are assigned the basis elements

---

[3]In fact, the degree of this representation is 27 since it can be shown that the dimension of the tensor product of vector space is simply the product of the dimensions of the individual spaces.

[4]Observe that $u \otimes u \otimes u = -u \otimes u \otimes u \Rightarrow u \otimes u \otimes u = 0$ implying the state doesn't exist where any two fermions have been exchanged.

[5]This is an exact symmetry of QCD meaning QCD does not discriminate color.
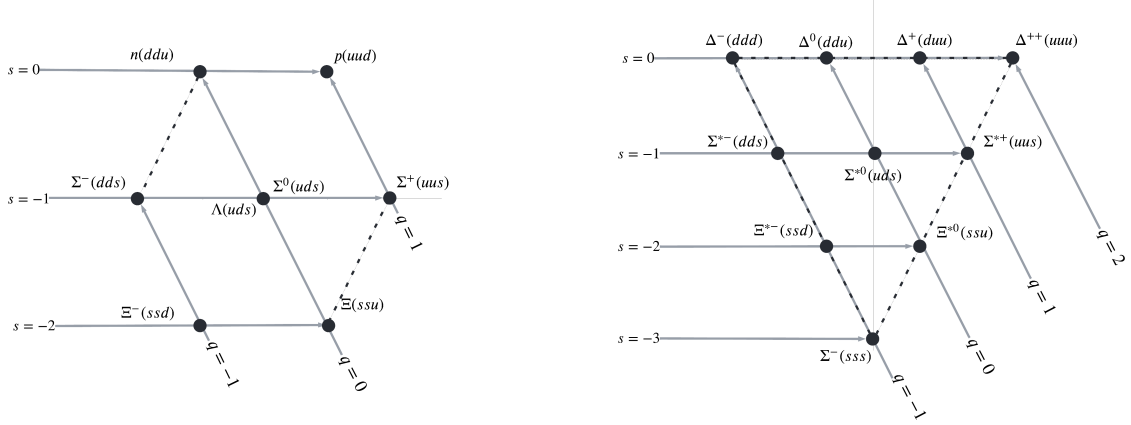
Figure 2.5. A representation with degree 27 corrisponding to the tensor product of the fundamental representation which decomposes as shown in equation (2.4.10). Only the decuplet and an octet of the irreducible decomposition corresponds to observed baryons. The decuplet corresponds to baryons with spin 3/2 and the octet corresponds to baryons with spin 1/2. The tensor product symbol has been dropped for brevity.

$$r = e_1 \qquad (2.4.11) \qquad g = e_2 \qquad (2.4.12) \qquad b = e_3 \qquad (2.4.13)$$

Anti-quarks carry anti-color which is represented by the anti-fundamental representation. The color structure is identical to that of flavour, hence everything said previously is applicable to color. How does this additional degree of freedom make the wavefunction anti-symmetric? This comes from an experimental observation of nature dubbed color confinement which states that hadrons can only exist in colorless states. The process by which this occurs is dubbed hadronization. Hence when combining the color component as we did with flavour, only the singlet state may correspond to a possible state. This breaks the symmetry of the wavefunction. For instance, in the case of baryons, the singlet component is given by

$$|\psi_c\rangle = \frac{1}{\sqrt{6}}(rgb - rbg + gbr - grb + brg - bgr) \tag{2.4.14}$$

As required, no symmetry in sight.

2.4.2 Lagrangian of QCD

Quarks, being fermions mean a natural place to start is at the lagrangian (density) of the Dirac equation given by

$$\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.4.15}$$

As mentioned, the Lagrangian must remain invariant locally under the action of $SU(3)$ on the color space which is through the fundamental representation. That is,

$$\psi \to U\psi \tag{2.4.16}$$

where $\boldsymbol{\alpha} = (\alpha_1(x), \alpha_2(x), ..., \alpha_8(x))$, $\mathbf{T} = (T^1, T^2, ..., T^8)$ and $U = \exp(i\boldsymbol{\alpha} \cdot \mathbf{T})$ with $T^a = \frac{1}{2}\lambda^a$ and λ^a are the Gell-Mann matrices. The locality of the gauge theory is captured in the coefficients, α, which may be a function of spacetime. The Dirac equation, however, does not meet our requirement. We can overcome this by looking to QED for inspiration. In the case of QED, the Lagrangian is
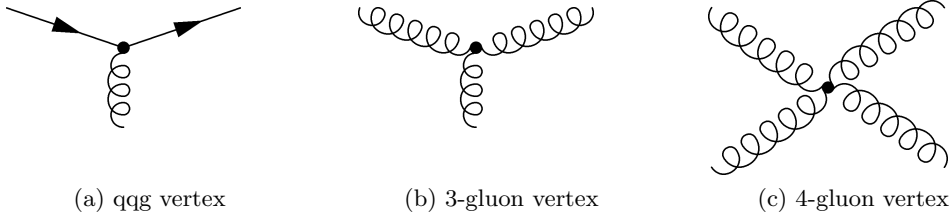
(a) qqg vertex   (b) 3-gluon vertex   (c) 4-gluon vertex

Figure 2.6. Possible gluon vertices in QCD due to the final term in equation (QCD Lagrangian) which couples multiple gluon fields.

invariant under $U(1)$ transformations which is made possible through the introduction of a covariant derivative which contains fields which transform in a particular way. The analogous treatment for QCD is to introduce a covariant derivative with components given by

$$D_{\mu;ij} = \partial_\mu \delta_{ij} - ig_s \sum_a T_{ij}^a A_\mu^a \qquad (2.4.17)$$

where we now have 8 gauge boson fields A_μ^a, $i, j = 1, 2, 3$ and $\mu = 0, 1, 2, 3$. For the covariant derivative introduced to make the Lagrangian invariant, it must transform as

$$\sum_a T^a A_\mu^a \to U \sum_a T^a A_\mu^a U^\dagger + \frac{i}{g_s}(\partial_\mu U) U^\dagger \qquad (2.4.18)$$

From this we see that the gauge fields must be in the adjoint representation. These 8 gauge fields corrispond to 8 now particles dubbed gluons. We now must introduce gluon field tensors into the Lagrangian to account for all possible phenomena involving gluons. These are given by

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f_{abc} A_\mu^b A_\nu^c \qquad (2.4.19)$$

Einstein summation is implied. The additional term not present in the Faraday tensor of QED is to preserve gauge invariance. They appear since $SU(3)$ is non-abelian, hence when moving the generators past each other the commutation relations

$$[T^a, T^b] = if_{abc} T^c \qquad (2.4.20)$$

are introduced. Again, an implicit sum is assumed. This additional term as we will see in a moment has the profound consequence of allowing for vertices with only gluons.

The entire Lagrangian of QCD is then

$$\mathcal{L} = \left(\sum_q \psi_q^i (i\gamma^\mu) D_{\mu;ij} \psi_q^j - m_q \bar\psi_q^i \psi_{iq} \right) - \frac{1}{4} G^{\mu\nu;a} G_{\mu\nu;a} \qquad \text{(QCD Lagrangian)}$$

The Einstein summation convention is implicit except for the quarks involved in the system which we place an explicit sum over to avoid too much clutter. The summation term couples gluons to quarks while the final term in the Lagrangian gives rise to terms which couple three and four gluons. These give rise to vertices as shown in figure 2.6. Observe that as we stated previously, QCD doesn't treat different flavours differently aside from their masses.

Since the Lagrangian of QED is the result of enforcing $U(1)$ guage invariance as is that of QCD, the Feynman rules are similar. It can be argued as [2] does for example that the quark current is given by

$$j^\mu = -ig_s(c_j^\dagger T^a c_i)(\bar u \gamma^\mu u) \qquad (2.4.21)$$
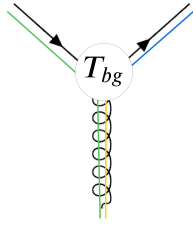
Figure 2.7. Exchange of color of green to blue quark through the exchange of a $g\bar{b}$ quark. This preserves the color at the vertex with $T_{bg} = c_b^\dagger T c_g$. The $\bar{b}$ is representation with the opposite color.

where c is one of the standard basis elements in $\mathbb{C}^3$ depending on the colours of the legs of the vertex. With the implicit knowledge that the Gell-Mann matrices act on the color component of the fermion and the gamma matrices act on the spinor component, we may write the vertex as

$$-ig_s c_j^\dagger T^a c_i \gamma^\mu \tag{2.4.22}$$

We therefore see that legs can only be coupled through a Gell-Mann matrix which contains a non-zero component at the required ji component. The gluon stats are therefore

$$(r\bar{b} + b\bar{r})/\sqrt{2} \tag{2.4.23}$$
$$(r\bar{g} + g\bar{r})/\sqrt{2} \tag{2.4.24}$$
$$(b\bar{g} + g\bar{b})/\sqrt{2} \tag{2.4.25}$$
$$(r\bar{r} - b\bar{b})/\sqrt{2} \tag{2.4.26}$$

$$-i(r\bar{b} - b\bar{r})/\sqrt{2} \tag{2.4.27}$$
$$-i(r\bar{g} - g\bar{r})/\sqrt{2} \tag{2.4.28}$$
$$-i(b\bar{g} - g\bar{b})/\sqrt{2} \tag{2.4.29}$$
$$(r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6} \tag{2.4.30}$$

This amounts to simply defining a new basis for the octet of the $3 \otimes \bar{3}$ representation, of which there are many. With this definition of the 8 gluons, an example of the "exchange" of color at a QCD vertex is illustrated in figure 2.7. A $g\bar{b}$ gluon couples the different colors of the incoming and outgoing quark.

2.4.3 Running Coupling and Renormalization

In QED, it is commonly overlooked that the electric charge measured in the lab is screened due to vacuum polarization. In considering, for example, NLO corrections to the s-channel process, $\ell^+\ell^- \to \gamma^* \to \ell^+\ell^-$, may include loops on external legs or within the photon propagator. The former contributions are guaranteed to cancel by the ward identity while the latter will result in divergent Feynman diagrams. The divergent terms are swept into a redefined coupling constant resulting in an infinite series of corrections from beyond leading order diagrams. Consider an s-channel with vacuum polarization:

$$\tag{2.4.31}$$

For each such loop, a divergent integral is contributed to the matrix amplitude and absorbed into the coupling of the theory with

$$\pi(q^2) = f\left(\frac{q^2}{\mu^2}\right) \quad \text{where } f(x) = \int_0^1 z(1-z)\ln(1 + xz(1-z))dz \tag{2.4.32}$$
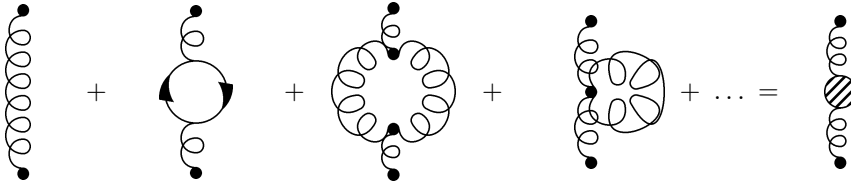
23

Figure 2.8. Additional loop diagram contributions to running coupling. This is in contrast to QED which is shown in equation (2.4.31). Adapted from [2].

which is dependent on some reference scale μ^2 where the coupling is known. We will implicitly assume all four-momenta are timelike by taking the absolute value. The details of how this form is obtained can be found in [2, 26]. If we were to calculate the corrections to the coupling from consider higher order terms, we would obtain

$$\alpha(q^2) = \alpha(\mu^2) \sum_{k=0}^{\infty} \left(\alpha(\mu^2)\pi(q^2)\right)^k \overset{q^2 \gg \mu^2}{=} \frac{\alpha(\mu^2)}{1 - \alpha(\mu^2)\ln\left(\frac{q^2}{\mu^2}\right)/3\pi} \tag{2.4.33}$$

Most importantly from this result, we see that the coupling of electrically charged fermions are is based on energy scale of the interaction. This is a remarkable result which has been observed in particularly sensitive scenarios such as the lamb shift.

An analogous result holds for QCD, albeit having much more profound consequences. Divergent integrals at higher orders may now be contributed by several new types of diagrams as shown in figure 2.8. The running of the strong coupling is governed by

$$q^2 \frac{\partial \alpha_s(q^2)}{\partial q^2} = \beta(\alpha_s(q^2)) \quad \text{where } \beta(\alpha_s(q^2)) = -\alpha_s^2(q^2)(b_0 + \mathcal{O}(\alpha_s(q^2))) \tag{2.4.34}$$

with $b_0 = \frac{11n_c - 2n_f}{12\pi}$ and n_c, n_f the number of colours and flavours respectively. The solution to equation (2.4.34) is given by

$$\alpha_s(q^2) = \frac{\alpha_s(q^2)}{1 + b_0\alpha_s(\mu^2)\ln\left(\frac{q^2}{\mu^2}\right) + \mathcal{O}(\alpha_s^2(\mu^2))} \tag{2.4.35}$$

Unlike QED where the running of the coupling towards increasing energy scales causes "anti-screening", within QCD the opposite occurs. That is, running QED to large enough energy scales would result in strongly coupled leptons. Conversely, at sufficiently large energies, quarks couple with sufficient weakness as to reinstate perturbative QCD (pQCD). While the order of the strong coupling doesn't run to the fine structure constant for energies seen in colliders, pQCD is still applicable albeit with slower convergence; this phenomenon is dubbed asymptotic freedom.[6] Usually equation (2.4.35) is expressed in terms of a characteristic scale Λ_{QCD}

$$\alpha_s(q^2) = \frac{1}{b_0\ln\left(\frac{q^2}{\Lambda_{\text{QCD}}^2}\right)} \quad \text{where } \ln\Lambda_{\text{QCD}}^2 = \ln\mu^2 - \frac{1}{b_0\alpha_s(\mu^2)} \tag{2.4.36}$$

A plot of the running of the coupling of QED and QCD are shown in figure 2.9 for comparison.

---

[6]In fact at sufficiently high energy scales, QCD becomes a conformal theory meaning it becomes scale independent.
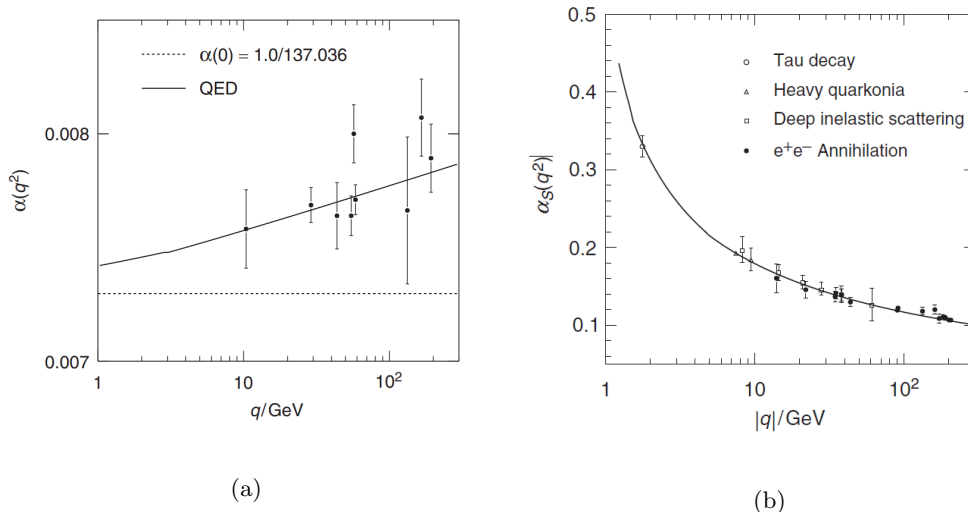
Figure 2.9. Running of the electromagnetic coupling (a) compared with the strong coupling (b) as measured from several experimental sources. The strong coupling drops with energy scale while the electromagnetic coupling grows very slowly. I fitting description for the strong coupling perhaps is sprinting rather than running. Obtained from [2] which are adapted form (a) [27] and (b) [28].

2.5 Markov chain Monte Carlo (MCMC)

Certainly whenever statistical models are implemented computationally, a mention of MCMC is warranted. MCMC are a class of numerical techniques for sampling from intractable distributions. Mainly, this is due to such high dimensional and intractable functional forms. Indeed, in this thesis both of which will certainly be encountered due to the probabilistic picture quantum mechanics paints of the subatomic world. In the coming subsections, we will begin by reminding the reader of random variables and sequences before proceeding to discussing practical methods for sampling. The integrals we will encounter in parton showers will be over spaces of hundreds of dimensions and beyond, for which we will discuss some basic techniques to tackle these. We will follow mainly [29] and [30].

2.5.1 Markov Chains

The discussion here makes reference to [31] in discussing the markovian property and stationarity.

A proper introduction must begin with random processes which in turn rely on random variables and vectors. Given some probability space, $(\Omega, \mathcal{F}, P)$, where Ω denotes the sample space and $\mathcal{F}$ the collection of subsets with well-defined probabilities given by the probability law P, each subset constitutes a so-called event. A random variable, θ, defined on this space is a function[7]

$$\theta : \Omega \to \mathbb{C} \cong \mathbb{R}^2 \tag{2.5.1}$$

then, θ may be equivalently described by a tuple of real random variables (θ_1, θ_2)

$$\theta_1 : \Omega \to \operatorname{Re} \theta \tag{2.5.2}$$

$$\theta_2 : \Omega \to \operatorname{Im} \theta \tag{2.5.3}$$

---

[7]In fact this is not the most general definition, within measure theory, a random variable's co-domain is a measureable space E which we have specialized to $E = \mathbb{R}^2$. For further details, the user is referred to [32].

This concept may be extended to a vector of random variables

$$\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n) \quad \text{where } \theta_i : \Omega \to \mathbb{R} \, \forall \, i \tag{2.5.4}$$

This motivates an alternative viewpoint in which $\boldsymbol{\theta}$ is seen as a mapping

$$\boldsymbol{\theta} : \Omega \to l^\infty \tag{2.5.5}$$

where l^∞ denotes a sequence space where remaining slots are padded with identically zero random variables (see [33] for a revision of sequence spaces). Such a view defines random processes where $\boldsymbol{\theta}(i) = \theta_i$. We may define several random processes on Ω or a single process containing random variables

$$\theta_i : \Omega \to \mathbb{R}^n \tag{2.5.6}$$

for some $n \in \mathbb{N}$. In other words, given N random processes on Ω denoted by $\boldsymbol{\theta}_j$ where $j = 1, 2, ..., N$, then $\boldsymbol{\theta}(i) = (\boldsymbol{\theta}_1(i), \boldsymbol{\theta}_2(i), ..., \boldsymbol{\theta}_1(N))$ i.e. we are keeping track of several random processes on Ω .

For a random process, we may define a transition probability classifying the probability the $(n+1)^{\text{st}}$ random variable admits some value given knowledge of the outcomes of the first n in the sequence. If the transition probability relies on the nth variable only

$$P(\theta^{(n+1)} \leq y | \cap_{l=0}^{n} \theta^{(l)}) = P(\theta^{(n+1)} \leq y | \theta^{(n)} = x) \quad \text{cont. rand. var.} \tag{2.5.7}$$

$$P(\theta^{(n+1)} = y | \cap_{l=0}^{n} \theta^{(l)}) = P(\theta^{(n+1)} = y | \theta^{(n)} = x) \quad \text{disc. rand. var.} \tag{2.5.8}$$

then the chain is said to be markovian.

A parton shower will probabilistically proliferate a state with few partons by successively making draws from quantum mechanical distributions to decide various properties of the partons such as flavour, spin, kinematics etc. Since the evolution of the system is via stochastic draws, this constitutes a probability space. The sample space consists of every possible proliferation allowed by the QFT until the partons reach the confinement scale. Studying the sample space directly is infeasible. Instead, it is most convenient to impose several random processes.

Pythia has an event record which to first approximation is a vector of vectors where each vector carries the information of a particle in the shower [15]. Within Pythia, each slot in the event record is an instance of the class Particle which stores information about the parton and its relation to other particles in the shower. The slots within the event record are not deleted but rather updated as the shower progresses. For instance, if a parton decays, it's status code is updated to reflect this change, however, it is not stricken from the record. This can be used to characterize the markov chain, the event record at some point in the shower represents the current state of the shower and is updated as the shower progresses. This can be cast as a random process where each slot of the random process is an instance of the event record. We can be more precise about this. Let a single parton be given by

$$\textbf{parton} = \mathbb{N} \times \text{status} \times \textbf{mothers} \times \textbf{daughters} \times \textbf{colours} \times \textbf{p} \times E_{\text{CM}} \times m \tag{2.5.9}$$

There are additional properties which have been omitted here for clarity such as spin, helicity, polarization etc. Further details on the event record can be found in [14].

1. $\mathbb{N}$: Particle identification code as specified in [25]. This indicates the flavour and whether it is anti-matter.

2. status: Indicates whether the parton remains by the end of the shower. Positive numbers indicate it remains.
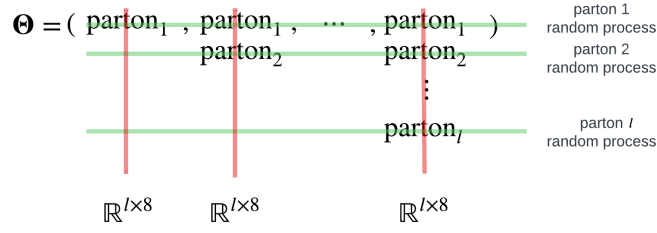
Figure 2.10. Visualization of markov chain of parton shower where a single event is given by $\boldsymbol{\Theta}$. This can be though of as a random process with each step in the chain a snapshot of the shower at some time. Alternatively, this can be seen as a several random processes for each parton in the shower.

3. **mothers** ($\mathbb{N} \times \mathbb{N}$): The id's of the two parent partons of the current parton.

4. **daughters** ($\mathbb{N} \times \mathbb{N}$): The id's of the two daughter partons of the current parton.

5. **p** ($\mathbb{R} \times \mathbb{R} \times \mathbb{R}$): Three momentum of the parton.

6. E_{CM}: COM energy of current parton.

7. m: Rest mass of current parton.

Then

$$\boldsymbol{\Theta}(n) : \Omega \to \mathbf{parton}_1 \times \mathbf{parton}_2 \times \cdots \times \mathbf{parton}_l \tag{2.5.10}$$

where l is the number of partons to have existed at some point in the shower. From this view, it is clear that the the process is markovian since the history of the particle is unimportant, only its current status. As was mentioned, an alternative view is that for each particle in the shower a random process is defined. These two views are illustrated in 2.10. The event is represented by $\boldsymbol{\Theta}$ and each column is a vector in $\mathbb{R}^{l \times 8}$ since there will be l partons by the end of the shower, each has 8 properties as given in 2.5.9. So $\boldsymbol{\Theta}$ can be seen as a sequence of random variables which are all vectors. Alternatively, this may simply be l random processes for each parton.

2.5.2 Accept/Reject Method

Our demands for an algorithm which samples from a complicated distribution are not few. Not only do we require some technique for sampling but that it must be capable of high efficiency and applicable to a non-normalized Probability Density Function (pdf), dubbed a kernel.

If we suppose the kernal we wish to make draws from is denoted by $\pi(x)$, then we can use a so-called blanketing distribution, with pdf denoted q which can be easily sampled. The blanketing property implies $\pi(x) \leq Aq(x)$ for all x given some $A < \infty$. The Accept/Reject Method is as follows:

1. Independently make draws from the blanketing distribution and $Uniform(0,1)$, these are denoted by x and u respectively.

2. If $u(Aq(x)) \leq \pi(x)$, accept x as the drawn element. Otherwise, return to the first step.

The proof for this is widely available (*e.g.* see [30] pg. 28). As required, the pdf which this technique samples is given by

$$f(x|u(Aq(x)) \leq \pi(x)) = \frac{\pi(x)}{\int dx\,\pi(x)} \tag{2.5.11}$$

which is the normalized kernal. Regions where q closely resembles π will often be accepted when trialed since the ratio $\pi(x)/(Aq(x))$ will be close to one. For regions where this isn't the case, rejections will happen in proportion to how poorly Aq resembles π and will therefore divert samples to other more likely regions. The blanketing distribution sets the efficiency. The probability a trial will be accepted is given by

$$P(u(Aq(x)) < \pi(x)) = \frac{1}{A} \int dx \, \pi(x) \tag{2.5.12}$$

So choosing a function q which badly approximates π will require a large value of A which will reduce the acceptance probability leading to more samples being required.

2.5.3 Numerical Integration

This discussion draws mainly from [34–37]. Before we are in a position to properly discuss how MCMC may be used to approximate a numerical integral, we must first specify several definitions and theorems which will aid in verifying MCMC integration and properties thereof. While several numerical integration techniques exist which may be generalized to higher dimensions, all of which become increasingly slower with increasing dimensions and grow in error. For example, the error in Simpson's rule scales as $M^{-4/d}$ with M the number of hypercubes segmenting the region of interest (see [37] for a brief comparison of other methods). Alternatively, MCMC is independent of dimension which is a necessity when discussing integrals of hundreds of dimensions.

When discussing sequences of random variable, we are no longer able to make specific statements on their behaviour but only probabilistic guarantees, as such we must define a new notion of convergence.[8]

Definition 2.5.1 (Convergence in Probability) *A sequence $\{X_n\}_n$ convergences in probability to X if for every $\epsilon > 0$*

$$\lim_{n \to \infty} P(|X_n - X| \geq \epsilon) = 0 \tag{2.5.13}$$

Note, definition 2.5.1 implies that there exists $n \in \mathbb{N}$ such that the $P(|X_n - X| \geq \epsilon) < \delta$ for any $\epsilon, \delta > 0$. A stark difference here compared to the usual definition in real analysis is that there are now two constraints set by ϵ and δ (for a reminder of the usual definition of the convergence of a sequence, see [38]).

We will also require the following theorem:

Theorem 2.5.1 (Law of Large Numbers) *An independent and identically distributed sequence $\{X_n\}$ with sequence $\{S_n/n\}_n$ of partial sums $S_n = \sum_{i=1}^{n} X_i$ converges in probability to $E[X]$.*

Now suppose we wish to integrate $f : \mathbb{R}^d \to \mathbb{R}$ over a $d-$dimensional region

$$I = \int_\Omega dx \, f(x) \quad \text{with } x \in \mathbb{R}^d \tag{2.5.14}$$

Then we may naively do so using the sample average estimator $\hat{\mu}_n$. Having sampled f n times using the sequence of random uniform variables $(x_1, x_2, ..., x_n)$

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \tag{2.5.15}$$

---

[8]There are several notions of probabilistic convergence, we only state the notions which the law of large numbers requires.

This may define sequence $\{\hat{\mu}_n\}$. Observe that this is the partial sum of a sequence of $f(x_i)$ divided by n. Then by theorem 2.5.1

$$\plim_n \hat{\mu}_n = \langle f(x_i) \rangle = \int dx \, f(x)\pi(x) = \frac{I}{V} \tag{2.5.16}$$

where $\pi(x) = 1/V$ is the pdf of a uniform random variable with support $V = \int_\Omega dx$. It is of little use knowing the expectation of the estimator has no bias without having a bound on the uncertainty range. We may define a variance estimator in an analogous fashion with the expectation substituted exactly. In practice this isn't true since the expectation is itself estimated which introduces a bias.

$$\hat{\sigma}^2(\hat{\mu}_n) = \frac{1}{n^2}\text{Var}\left[\sum_i f(x_i)\right] = \frac{\sigma^2(f(x))}{n} \tag{2.5.17}$$

where we have assumed $f(x_i)$ are independent. From this we can immediately see that interestingly, the variance in the estimator is entirely independent of the dimension of the integral. We obtain the famous result that standard deviation is $\sigma(f(x))/\sqrt{n}$.

Chapter 3

Event Generation

Usually, at the heart of an experiment is a (relatively) simple process which is the process under investigation (the hard process). For instance, in a proton-proton beam, the process of interest may be a Drell-Yan process[39]. These processes occur at a sufficiently high energy scales where perturbative methods may be and are employed in programs such as Pythia, Herwig and Sherpa. The hard process products will be highly energetic emitting radiation and eventually decaying. This repeats with each succession, losing energy until the particles are at an energy scale where no further reactions may occur. Beyond the hard process, near the QCD scale $\Lambda \approx 220\text{MeV}$, perturbative methods are no longer applicable. Within this regime, so-called showering algorithms pass the final state contains partons to a hadronization model. The discussion will draw mainly from [13, 14, 40–43] and will focus explicitly on Pythia since this is the program which is the subject of this thesis. In particular, we focus on a plugin to Pythia dubbed Vincia which specifically handles the showering formalism which is the focus of this thesis.

3.1 Cross Section Factorization of Hard Processes

The hard processes which ship with Pythia can be found in [15] in the Process Selection section. The matrix amplitudes for these processes are calculated to some order and input into the program code. The calculation of hard cross-sections is facilitated by factorization formulas which cast the hadron-hadron cross-sections in terms of the constituent partons. These factorizations are only strictly valid in Deep Inelastic Scattering (DIS), however, although no formal proof exists, when applied to hadron-hadron interactions, good agreement with experiment is found. The (inclusive) cross-section for two hadrons to produce a state $n \equiv |p_1, p_2, p_3, ..., p_n\rangle$ where p_i denotes the ith parton is given by ([44])

$$\sigma_{H_1 H_2 \to n} = \sum_{ab} \int dx_a dx_b f_{a/H_1}(x_a, \mu_F^2) f_{b/H_2}(x_b, \mu_F^2) \hat{\sigma}_{ab \to n}(\mu_F^2) \tag{3.1.1}$$

Hence, a hadronic cross-section is simply the weighted contribution of all possible partonic channels. Each channel has associated with it a cross-section depending on the parton 4-momentum fractions, x, carried by the corrisponding parton. σ is multiplied by the number of partons which possess the given momentum fraction. This is captured by $f_{a/H_1} dx_a / f_{b/H_2} dx_b$ which are the number of partons of type a/b in hadrons H_1/H_2 with momentum fractions within the range $x_a + dx_a / x_b + dx_b$. The factorization scale μ_F indicates the energy scale at which the Parton Distribution Function (PDF)'s where determined. PDF's originate from the study of DIS where these functions are inferred at the scale of the four-momentum of virtual photon. Proofs for the various factorization theorems can be
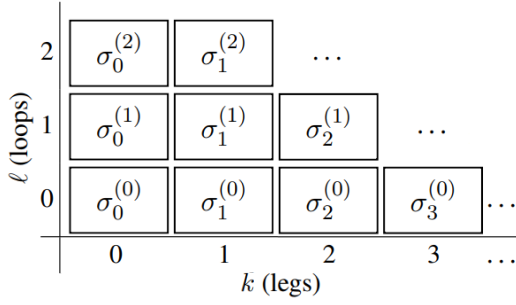
Figure 3.1. Visualization of equation (3.1.3) where each box represents an element the expansion. The superscript corresponds to order above born level given by $\hat{\sigma}_0^{(0)}$ due to virtal corrections. The subscript is due to real corrections. The bottom row corresponds to leading order in $n + k$ elements in the final state.

found in [45]. Often our studies are in terms of some observable of interest which we denote with $\mathcal{O}$. Hence the cross-section with respect to some observable is given by[42]

$$\frac{d\sigma_{H_1 H_2 \to n}}{d\mathcal{O}} = \sum_{ab} \int dx_a dx_b f_{a/H_1}(x_a, \mu_F^2) f_{b/H_2}(x_b, \mu_F^2) \frac{d\hat{\sigma}_{ab \to n}(\mu_F^2)}{d\mathcal{O}} \tag{3.1.2}$$

$\hat{\sigma}$ can be further factorized in terms of virtual and real corrections. Real correction correspond to additional particles in the final state. Loop corrections correspond to off-shell particles which contribute virtual corrections. The expansion is given by

$$\hat{\sigma}_{ab \to n}(\mu_F) = \sum_k^\infty \int d\Phi_{n+k} \left| \sum_{l=0}^\infty \mathcal{M}_{n+k}^{(\ell)} \right|^2 \tag{3.1.3}$$

The sum over k corresponds to real corrections which introduce additional partons in the final state. Embedded in the sum of k are virtual corrections denoted with ℓ which introduce loops within Feynman diagrams i.e. additional orders of the coupling. For some fixed k, the inner expansion is given by

$$\sum_{l=0}^\infty \mathcal{M}_{n+k}^{(\ell)} \sum_{j=0}^\infty \mathcal{M}_{n+k}^{(j)*} = \mathcal{M}_{n+k}^{(0)} \mathcal{M}_{n+k}^{(0)*} + \mathcal{M}_{n+k}^{(0)} \mathcal{M}_{n+k}^{(1)*} + \dots \tag{3.1.4}$$

$$+ \mathcal{M}_{n+k}^{(1)} \mathcal{M}_{n+k}^{(0)*} + \mathcal{M}_{n+k}^{(1)} \mathcal{M}_{n+k}^{(1)*} + \dots \tag{3.1.5}$$

From now on, we will use the shorthand $\mathcal{M}_{n+k}^{(\ell,j)} = \mathcal{M}_{n+k}^{(\ell)} \mathcal{M}_{n+k}^{(j)*}$ for clarity. This expansion is best visualized as presented by [46] and reproduced in figure 3.1. Each term in 3.1.4 will correspond to a term in the overall cross-section. Hence, the overall cross section is given by

$$\hat{\sigma} = \sum_{ij} \hat{\sigma}_j^{(i)} \tag{3.1.6}$$

where $\hat{\sigma}_k^{(i)}$ contains all amplitudes which contain i additional orders of the coupling above the born with k corrections.

If we consider the NLO term, then there are two possibilities:

1. An additional real emission in both matrix amplitudes is present i.e. $\mathcal{M}_1^{(0,0)}$.

2. One of the matrix amplitudes contains an additional virtual correction $\mathcal{M}_0^{(1,0)}, \mathcal{M}_0^{(0,1)}$.

31

So the overall cross-section calculation up to NLO is given by

$$\hat{\sigma}^{\mathrm{NLO}} = \int d\Phi_n \mathcal{M}_0^{(0,0)} + \int d\Phi_n \left(\mathcal{M}_0^{(1,0)} + \mathcal{M}_0^{(0,1)} \right) + \int d\Phi_{n+1} \mathcal{M}_1^{(0,0)} \tag{3.1.7}$$

$$= \hat{\sigma}_0^{(0)} + \hat{\sigma}_0^{(1)} + \hat{\sigma}_1^{(0)} \tag{3.1.8}$$

This corresponds exactly to the second diagonal from the bottom left in figure 3.1. Higher order corrections correspond to subsequent diagonals. Explicit calculation of these amplitudes will show that many of integrals presented in equation (3.1.8) are divergent individually, however by the Kinoshia-Lee-Nauenberg (KLN) theorem as discussed and proven by [47, 48] guarantees that the integrals grouped by order produce finite results. Methods for doing this are non-trivial since not all the integrals are over the same space as we see with the NLO expansion shown in equation (3.1.8). There are several techniques which can be used to solve these integrals which are beyond the scope of this these. If 3.1.8 is evaluated, we obtain

$$\hat{\sigma}^{\mathrm{NLO}} = \hat{\sigma}_0^{(0)} \left(1 + \frac{\alpha_S(E_{\mathrm{COM}})}{\pi} + \mathcal{O}(\alpha_S^2) \right) \tag{3.1.9}$$

The treatment presented so far is not general enough to apply to all observables. An observable where the KLN theorem will apply must have the order-by-order terms correspond to degenerate states so that such a grouping may occur. Such observables which abide by this condition are dubbed infrared-safe observables. Following with [44], given some observable which is a function of $n + 1$ partons with momenta denoted

$$\{p\}_{n+1} = \{p_1, p_2, ..., p_n, p_{n+1}\} \tag{3.1.10}$$

Observe, the momenta are in a set implying the ordering is interchangeable. So we can discuss any two momenta without loss of generality. Then for $\mathcal{O}$ to be infrared-safe, it satisfy two conditions

1. In the collinear limit where two partons with 4-momenta $p_n \to z\tilde{p}$ and $p_{n+1} \to (1-z)\tilde{p}$ where z delineates the 4-momenta fraction carried by each parton, then

$$\mathcal{O}(\{p_n, p_{n+1}\}) \to \mathcal{O}(\{\tilde{p}\}) \tag{3.1.11}$$

2. If a parton is collinear to the beam i.e. possessing momentum $p_n \to \xi p_{\mathrm{beam}}$, then

$$\mathcal{O}(\{p\}_n) \to \mathcal{O}(\{p\}_{n-1}) \tag{3.1.12}$$

In the second condition, if $\xi \to 0$, then the parton is within the soft limit. Indeed, these are in themselves reasonable requirements of our observables since no detector can distinguish between two perfectly collinear particles or a particle with arbitrarily small momentum. Implementation of infrared safe observables requires additional care since algorithms must group jets to cancel singularities, a class which satisfies this requirement are jet clustering algorithms. [49] goes into further detail on jet clustering algorithms and infrared safety.

3.2   Parton Showers

Having obtained the cross-section for the hard process, further corrections are handled by the parton shower. Our task is then to take the cross-section $\sigma_n \equiv \sigma_{h_1 h_2 \to n}$ and incorporate additional loop and real corrections up to the hadronization scale. As we will see, this will be done by casting each correction (virtual or real) as steps of a markov chain. The equations developed here are sufficiently general to apply to electroweak processes with relevant quantities such as Feynman rules substituted.

In fact, Pythia contains such processes (*e.g.* see [14] for a discussion of the theoretical basis or [15] for the available processes). The discussion in this thesis will specialize to QCD.

The hard process in the centre of the shower contains high momentum transfers with outgoing partons which will inevitably radiate through bremsstrahlung predominantly. From §2.4, the radiative processes which may take place are

$$q \to qg \tag{3.2.1}$$

$$g \to gg \tag{3.2.2}$$

$$g \to q\bar{q} \tag{3.2.3}$$

To evolve the hard cross-section through subsequent emissions, we require a scheme which iteratively modifies the cross-section to produce σ_{n+1} due to some real radiative correction. To preserve unitarity, an additional term will need be developed which accounts for loop corrections i.e. no additional emissions. We will come to this detail once we have developed additional terminology, for now we begin with real emissions. Recall, the cross-section for some process is given by [1]

$$\sigma_n = \int d\Phi_n |\mathcal{M}_n|^2 \tag{3.2.4}$$

where

$$d\Phi_n = d\Phi_n(P; p_1, ..., p_n) = (2\pi)^4 \delta^{(4)} \left(P - \sum_{f=1}^{n} p_f \right) \prod_{i=1}^{n} \frac{d^4 p_i}{(2\pi)^3} \delta(p_i^2 - m_i^2) \tag{3.2.5}$$

which is the Lorentz Invariant Phase Space (LIPS) measure. P denotes the total momentum of the initial (or final) state. As shown by [50], the measure may be factorized as

$$d\Phi_n(P; 1, ...n) = d\Phi_{n-(m-1)}(P; \pi, m+1, ..., n) \frac{ds_\pi}{(2\pi)} d\Phi_m(p_\pi; 1, ..., m) \tag{3.2.6}$$

where $d\Phi_{n-(m-1)}$ and $d\Phi_m$ are respectively the LIPS measures containing m and $n - m + 1$ particles. s_π can be interpreted as an interconnecting s-channel virtual propagator denoted by π with mandelstam variable

$$s_\pi = (p_1 + \cdots + p_m)^2 \tag{3.2.7}$$

The factorization can be visualized as shown in figure 3.2. This provides precisely the factorization of the phase-space we require. For example, consider a process with n partons $\to n + 1$ partons due to process $\tilde{i}\tilde{k} \to ijk$. If we wish to write the $n + 1$-parton phase-space in terms of the n-parton state and some additional factor, let $m = 2$ in 3.2.6 to obtain

$$d\Phi_{n+1}(P; p_1, ..., p_i, p_j, p_k, ..., p_{n+1}) = d\Phi_n(P; p_1, ..., p_{\tilde{i}}, p_{\tilde{k}}, ..., p_n) d\Phi_{+1}(p_i, p_j, p_k) \tag{3.2.8}$$

which is precisely as we would hope. $d\Phi_{+1} \equiv \left| \frac{\partial(p_{\tilde{i}}, p_{\tilde{k}})}{\partial(p_\pi, p_k)} \right| \frac{ds_\pi}{(2\pi)} d\Phi_2(p_\pi; p_i, p_j)$ contains a jacobian which changes variables from $p_i, p_j, p_k = p_\pi, p_k \to p_{\tilde{i}}, p_{\tilde{k}}$. Finally, we define a so-called radiation function K which allows us to complete the cross-section factorization by changing the matrix amplitude squared for n particles to $n + 1$ i.e. $K|\mathcal{M}_n|^2 = |\mathcal{M}_{n+1}|^2$. For the process $\tilde{i}\tilde{k} \to ijk$, the factorization becomes exact in the limit that emissions become soft or collinear. If j is the emitted parton, then a soft emission requires $E_j \to 0$ and a collinear emission requires $i \parallel j$ or $j \parallel k$.
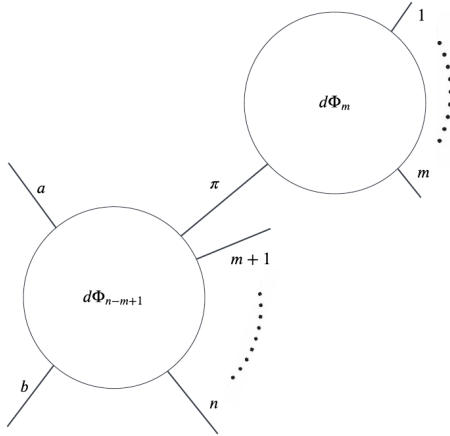
---

[1]The flux factor has been suppressed.

Figure 3.2. Factorization of LIPS into two processes containing m and $n - m + 1$ particles with an interconnecting virtual particle π.

The cross-section is as follows

$$d\sigma_{n+1}^{j/\tilde{i}\tilde{k}} = d\Phi_{n+1}|\mathcal{M}_{n+1}|^2 \xrightarrow{\text{single-unresolved}} d\Phi_n d\Phi_{+1} K_{j/\tilde{i}\tilde{k}}|\mathcal{M}_n|^2 \tag{3.2.9}$$

$$= K_{j/\tilde{i}\tilde{k}}(d\Phi_n|\mathcal{M}_n|^2)d\Phi_{+1} \tag{3.2.10}$$

$$= d\sigma_n^{j/\tilde{i}\tilde{k}} \underbrace{K_{j/\tilde{i}\tilde{k}}d\Phi_{+1}}_{\text{prob. of emission}} \tag{3.2.11}$$

This is a remarkable result and the heart of parton showers. We see that an $n + 1$-parton process cross-section can be written in term of a (known) n-parton cross-section with some correction. We may interpret this correction as a probability which implies K is a kernel to be sampled. Using this prescription, we may calculate the cross-section with as many real emissions by chaining these correction terms forming a markov chain with kernel K.

The overall cross-section is simply the sum over all viable channels which produce the desired final state which in turn are each calculable by factorization in the single-unresolved limit

$$d\sigma_{n+1} = \sum_{ijk} d\sigma_{n+1}^{j/\tilde{i}\tilde{k}} \tag{3.2.12}$$

We will define so-called antenna functions, denoted by a, where $a = |\mathcal{M}_{n+1}^2|/|\mathcal{M}_n|^2 \to K$ i.e. instead of calculating K directly with some limit in mind, one obtains a general expression first with no limits taken. Since the antenna function is independent of the specific $n \to n + 1$ process, antenna functions are calculated using single-vertex processes i.e. $n = 2 \to 3$. Since this factorization is only exact in the limits discussed, antenna functions don't possess universal functional forms. Only in the a singular limit is the kernel fixed. Within the collinear limit, we obtain so-called Dokshitzer Gribov Lipatov Altarelli Parisi (DGLAP) kernels given by

$$a_{j/\tilde{i}\tilde{k}} \xrightarrow{i\|j} K_{\tilde{i}\to ij} = \frac{8\pi\alpha_s}{2p_i \cdot p_j} P_{\tilde{i}\to ij} \tag{3.2.13}$$

while in the soft limit we obtain

$$a_{j/\tilde{i}\tilde{k}} \xrightarrow{E_j\to 0} K_{j/\tilde{i}\tilde{k}} = 8\pi\alpha_s \mathcal{C}_{j/\tilde{i}\tilde{k}} \underbrace{\frac{2p_i \cdot p_k}{(2p_i \cdot p_j)(2p_j \cdot p_k)}}_{\text{soft eikonal}} \tag{3.2.14}$$

where $\mathcal{C}_{j/\tilde{i}\tilde{k}}$ is proportional to a colour factor $\{C_A, C_F, T_R\}$ depending on the process under consideration with $C_F = \frac{N_c^2 - 1}{2N_c}$, $C_A = N_c$, $T_R = \frac{1}{2}$ and $N_c = 3$. Pythia's simple shower is evolved using DGLAP kernels while the plug-in, Vincia, uses the antenna formalism. It is instructive to begin our discussion of parton showers in the DGLAP formalism and generalize to antenna showers.

3.2.1 DGLAP Formalism

Final State Radiation (FSR)

Within the DGLAP regime, the showering is cast in terms of processes of the form $\tilde{i} \to ij$. This is analogous to the monopole approximation of QED i.e. emissions are from monopoles. We may choose to characterize the phase space using (m_{ij}^2, z, ϕ_{ij}) which denote respectively the invariant mass between the daughter partons, the energy fraction carried by the ith parton and the branching angle in the daughter partons rest frame. The invariant mass of the monopole-emission pair is given by

$$m^2 = m_{ij}^2 = (p_i + p_j)^2 \tag{3.2.15}$$

$$= p_i^2 + p_j^2 + 2p_i \cdot p_j \tag{3.2.16}$$

$$\xrightarrow{m_i, m_j \to 0} 2p_i \cdot p_j \tag{3.2.17}$$

Assuming masses are negligible is justified since the energy scale of the showering is large in comparison. So the cross-section becomes

$$d\sigma_{\tilde{i} \to ij} = d\sigma_{\tilde{i}} \frac{8\pi\alpha_s}{2p_i \cdot p_j} P_{\tilde{i} \to ij}(z, \phi) d\Phi_{+1}(p_i, p_j, p_k) \tag{3.2.18}$$

$$= d\sigma_{\tilde{i}} \frac{\alpha_s}{2\pi} P_{\tilde{i} \to ij}(z, \phi) \frac{dm^2}{m^2} dz \, d\phi \tag{3.2.19}$$

where p_k denotes the momentum of the recoiler which will be elaborated upon in subsequent discussion. In changing the integration variables we used $d\Phi_{+1}(p_i, p_j, p_k) = \frac{1}{16\pi^2} dm^2 \, dz \, d\phi$ where we have introduced another jacobian since we are once again changing variables. In the collinear limit, the phase space has several equivalent characterizations. We will choose $\frac{dm^2}{m^2} = \frac{dp_\perp^2}{p_\perp^2}$ since $p_\perp$ is used in Vincia showers. The emission spin-averaged radiation kernels for the processes in equations (3.2.1) to (3.2.3) where originally derived in [51] and are given by

$$P_{q \to qg} = C_F \frac{1 + (1-z)^2}{z} \tag{3.2.20}$$

$$P_{g \to gg} = 2C_A \frac{(1 - z(1-z))^2}{z(1-z)} \tag{3.2.21}$$

$$P_{g \to q\bar{q}} = T_R(z^2(1-z)^2) \tag{3.2.22}$$

To complete the cross-section, a sum over all possible decay channels must be included giving

$$d\sigma_{\tilde{i}+1} = d\sigma_{\tilde{i}} \underbrace{\sum_{ij} \frac{\alpha_s}{2\pi} P_{\tilde{i} \to ij}(z) \frac{dp_\perp^2}{p_\perp^2} dz}_{d\mathcal{P}(p_\perp^2, z, d\Phi_{+1}(p_\perp^2, z))} \tag{3.2.23}$$

where we have identified the probability (at leading order) of an emission from some $\tilde{i}$ is given by $d\mathcal{P}$. This is an inclusive probability since this takes into consideration only parton $\tilde{i}$, this will remain implicit throughout our discussion. Then the probability that $\tilde{i}$ will emit a parton carrying any energy fraction is given by

$$d\mathcal{P}(p_\perp^2, d\Phi_{+1}(p_\perp^2)) = \int_{z_{min}}^{z_{max}} dz \, d\mathcal{P}(p_\perp^2, z, d\Phi_{+1}(p_\perp^2, z)) \tag{3.2.24}$$

where the limits have been placed to restrict a daughter parton being too soft, the upper limit must then be imposed by symmetry. This has several advantages:

1. This avoids the infrared divergence $\int_{z_{min}}^{z_{max}} \frac{dz}{z} = \log \frac{z_{max}}{z_{min}}$.

2. Takes into account the experimental limitations of the detectors which cannot resolve such soft emissions.

3. This is equivalent to incorporating a mass for the partons dubbed the dead cone approximation. It is approximate since the phase space cut-off should be gradual.

In order to make the shower exclusive which is done through a so-called ordering variable which in our instance is $p_\perp^2$. Given a final state containing n partons, the shower will evolve the parton which branches to the hardest scale. However, the probability for parton $\tilde{i}$ to branch at some later scale, $p_{\perp_1}^2$ requires knowledge of the no-emission probability between the scales $p_{\perp_0}^2 \to p_{\perp_1}^2$ customarily denoted by $\Delta(p_{\perp_0}^2, p_{\perp_1}^2)$. This ensures that the parton which is undergoing a decay at some scale hasn't already done so previously i.e.

$$\left[\text{prob. to branch at } p_{\perp_1}^2 \text{ given starting scale } p_{\perp_0}^2\right] = \left[\text{prob. not to branch within range } (p_{\perp_0}^2, p_{\perp_1}^2)\right] \tag{3.2.25}$$

$$\times \left[\text{prob. to branch at } p_{\perp_1}^2\right] \tag{3.2.26}$$

The differential branching probability is therefore given by

$$d\mathcal{P}(p_{\perp_1}^2, d\Phi_{+1}(p_{\perp_1}^2)) = \Delta(p_{\perp_0}^2, p_{\perp_1}^2) d\mathcal{P}(p_{\perp_1}^2, d\Phi_{+1}(p_{\perp_1}^2)) \tag{3.2.27}$$

Since the no-emission probability is $1-$emission probability, in order to preserve unitarity,[2] at some $p_\perp^2$ scale, we require that the no-emission probability change in proportion to the amount that the emission probability changes i.e. if within some region of phase-space, the no-emission probability drops, it must be because the emission probability increased. This leads to the differential equation

$$\frac{d\Delta(p_{\perp_0}^2, p_{\perp_1}^2)}{dp_{\perp_1}^2} = \Delta(p_{\perp_0}^2, p_{\perp_1}^2) \frac{d\mathcal{P}(p_{\perp_1}^2, d\Phi_{+1}(p_{\perp_1}^2))}{dp_{\perp_1}^2} \tag{3.2.28}$$

with solution given by the so-called Sudakov form factor

$$\Delta(p_{\perp_0}^2, p_{\perp_1}^2) = \exp\left(-\int_{p_{\perp_1}^2}^{p_{\perp_0}^2} d\mathcal{P}(p_\perp^2, d\Phi_{+1}(p_\perp^2))\right) \tag{3.2.29}$$

$$= \exp\left(-\int_{p_{\perp_1}^2}^{p_{\perp_0}^2} \int_{z_{min}}^{z_{max}} \int_0^{2\pi} \sum_j K_{j/\tilde{i}\tilde{k}} d\Phi_{+1}(p_\perp^2, z, \phi)\right) \tag{3.2.30}$$

$$= \prod_j \Delta_{j/\tilde{i}\tilde{k}} \quad \text{with } \Delta_{j/\tilde{i}\tilde{k}} = \exp\left(-\iiint K_{j/\tilde{i}\tilde{k}} d\Phi_{+1}(p_\perp^2, z, \phi)\right) \quad \text{(Sudakov Factor)}$$

This result is sufficiently general as to apply to any showering formalism i.e. not specific to the DGLAP formalism. Δ possesses the composition property, easily seen by separating the integral into multiple segments. Hence given some intermediate scale $p_{\perp_0}^2 > p_{\perp_a}^2 > p_{\perp_1}^2$

$$\Delta(p_{\perp_0}^2, p_{\perp_1}^2) = \Delta(p_{\perp_0}^2, p_{\perp_a}^2)\Delta(p_{\perp_a}^2, p_{\perp_1}^2) \tag{3.2.31}$$

---

[2]Once we obtain an explicit form for Δ, the reader may check that unitarity is indeed preserved for every infinitesimal region of the phase-space by Taylor expanding and truncating Δ at first order.

Although beyond the scope of this thesis, the Sudakov factor has the interpretation of incorporating virtual corrections to the shower. Hence, we have a procedure which allows us to build complex cross-sections which incorporates real and virtual corrections. The general expression for a branching is therefore given by [3]

$$\pi(p_{\perp_1}^2, z | p_{\perp_0}^2) = \Delta(p_{\perp_0}^2, p_{\perp_1}^2) \sum_j K_{j/\tilde{i}\tilde{k}} \tag{3.2.32}$$

$$= \exp\left(-\int_{p_{\perp_1}^2}^{p_{\perp_0}^2} \int_{z_{min}}^{z_{max}} \sum_{j'} d\Phi_{+1}(p_\perp^2, z) K_{j'/\tilde{i}\tilde{k}}\right) \sum_j K_{j/\tilde{i}\tilde{k}} \tag{3.2.33}$$

which may be integrated over the phase-space to obtain probabilities. While colloquially dubbed a kernel, this is specifically a transition kernel.

Vincia's implementation is not directly in this naive way we have described but through an equivalent computationally more efficient approach where the individual distributions are summed and sampled, the branching parton is then chosen based on its likelihood to branch at the given point in phase-space relative to all others.

Initial State Radiation (ISR)

As was hinted in §2.1, the showering is evolved backwards from the hard interaction. Such an approach, although more complicated, is essential for practicality and efficiency of MCMC implementations of parton showers. Traditionally, a shower is run and discarded if the hard process isn't found. By evolving backwards from the hard process of interest, we are guaranteed to only generate events of interest. Backwards evolution to generate ISR was first introduced in [52] which we follow here. The reader is referred to [8, 14, 53] for a description by other authors.

The formalism is developed in terms of PDF's and evolution variables $p = (x, t = \ln \mu_F^2)$[4] denoting the hadron momentum fraction and the factorization scale respectively. Again, consider process $\tilde{i} \to ij$. Parton i's PDF at some p is determined by contributions of all processes capable of producing the target parton at some energy fraction, z, the contribution is suppressed by the rarity of the parent parton $\tilde{i}$. In the collinear limit, the PDF evolution is described by the Altarelli-Parisi (AP) equations

$$\frac{df_i(x,t)}{dt} = \frac{\alpha_s(\mu_F^2)}{2\pi} \sum_{\tilde{i}} \int \underbrace{dx' f_{\tilde{i}}(x',t)}_{\substack{\text{\# of } \tilde{i} \text{ carrying momentum} \\ \text{frac. in } [x', x'+dx']}} \times \underbrace{dz P_{i/\tilde{i}}(z)}_{\substack{\text{prob. process } \tilde{i} \to i + \\ \text{anything occurs within } [z, z+dz]}} \times \delta(x - x'z) \tag{3.2.34}$$

The delta function requires parton i acquire the necessary energy fraction z from parton $\tilde{i}$ to carry x of the total hadron momentum i.e. $p_i = xp_h = z(x'p_h)$. The probability that $\tilde{i}$ produce i is given by

$$P_{i/\tilde{i}}(z) = \sum_j P_{\tilde{i} \to ij}(z) \tag{3.2.35}$$

So equation (3.2.34) provides a method for determining the density of parton i due to emissions from parton $\tilde{i}$. In other words, the probability of parton i disappearing in region dt is given by

$$\frac{d\mathcal{P}_i(x,t)}{dt} = \frac{1}{f_i(x,t)} \frac{df_i(x,t)}{dt} \tag{3.2.36}$$

---

[3]The third component ϕ has been suppressed.

[4]p does not denote momentum here, this is simply denotes coordinates in phase-space.

The probability that parton i disappear in some finite interval $[t_0, t_1]$ with $t_0 > t_1$ is given by

$$\mathcal{P}_i(x, t_0, t_1) = \int_{t_1}^{t_0} \frac{df_i(x, t)}{f_i(x, t)} \tag{3.2.37}$$

Now a poisson random variable describes the probability of N emission occurring within (t_0, t_1) is given by

$$P(X = n) = \frac{\lambda^n e^{-\lambda}}{n!} \tag{3.2.38}$$

where $X \sim \text{Poisson}(\lambda)$ with $\lambda = \mathcal{P}_i(x, t_0, t_1)$. Now equation (3.2.38) is nonsensical beyond a single event where the daughter parton is no longer existent. However, this allows us to extract the "no absorption" probability given by

$$S_b(t_0, t_1; x) = \exp\left(-\int dt \, \frac{\alpha_s(\mu_F^2)}{2\pi} \sum_{\tilde{\imath}} \int dx' \frac{f_{\tilde{\imath}}(x', t)}{f_i(x, t)} \int dz \, P_{i/\tilde{\imath}}(z)\delta(x - x'z) \right) \tag{3.2.39}$$

This result is analogous to the Sudakov form factor. This approach is analogous to that used in [54] for deriving the Sudakov factor. [54] are a slide pack developed by Prof. Peter Skands for the purposes of masters lectures at Monash and not publicly available. The reader may get in contact and request access.

3.2.2 Colour in Parton Showers

As we have discussed in chapter 1 and §2.4, the colour algebra of $SU(N_c)$ is incorporated into the QCD Lagrangian through the fundamental representation but characterized through the representation of the corresponding Lie algebra. The gluons enter the Lagrangian through the adjoint representation which is an irreducible component of tensoring the fundamental and anti-fundamental representations. Since no colourless gluons exist in nature, the singlet is excluded from the theory which introduces complications stemming from the forced use of the adjoint representation. The implementation of colour exactly into parton showers is unfeasible. Fortunately, the implementation directly in terms of $N_c \otimes N_c$ possesses several features which make it feasible computationally. A theory which incorporates the singlet describes a world where QCD possesses an analog to the photon of QED which we refer to as the phantom gluon. The Feynman rules enforce a "flow" of colour through vertices which begin and end at external legs owing to the now possible $c\bar{c}$ phantom gluon. The Feynmann rules for this new theory can be found in [55] where it is referred to as colour-flow QCD. This requirement is imposed by kronecker deltas at the vertices. An example of a process with colour-flow drawn overtop is presented in figure 3.3. This diagrammatic representation of colour flow is called t'Hooft's double line notation first presented in [56]. The introduction of a virtual gluon which may possibly be a phantom carries with it the ramification of multiple Feynmann diagrams contributing to a single process which introduces sub-leading interference terms which aren't captured by the showers. An illustration of this is provided in [55] in section 2.5. To mitigate the effects of the phantom gluon, parton showers actually implement a theory where $N_c \to \infty$ referred to as the LC limit introducing an unlimited variety of charges within the shower. Since,

$$N_c \otimes N_c = (N_c^2 - 1) \otimes 1 \tag{3.2.40}$$

then in the LC limit, the effect of the singlet state is negligible. Consequently, colour dipoles are only formed between partons which have a "common ancestor" so to speak. The limitless colours introduced into the shower are assigned Les Houches colour tags which enumerate the charges as they are introduced. Examples of colour tags are given by the numbers next to the flow lines in figure 3.3. An alternative take on the topic can be found in [57].
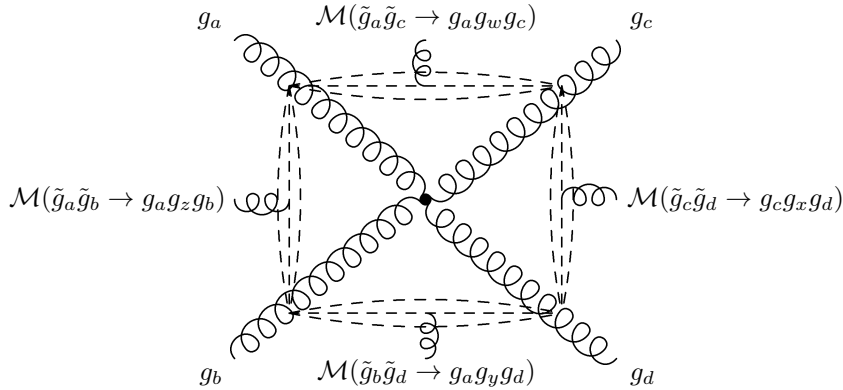
Figure 3.3. An example of a Drell-Yan process with a hard interaction splitting into two jets beginning to shower. In colour-flow QCD the feynman diagrams are represented by colour lines beginning and ending at incoming and outgoing legs. Parton showers implement this theory in the in the LC limit introducing unique charges, each assigned a Les Houches colour tag.

3.2.3 Antenna Formalism

Recall

$$K_{j/\tilde{i}\tilde{k}} = \frac{|\mathcal{M}(\tilde{i}\tilde{k} \to ijk)|^2}{|\mathcal{M}(\tilde{i}\tilde{k})|^2} \tag{3.2.41}$$

The amplitude is calculated by summing all Feynmann diagrams which may contribute to this process, namely a history where the emission is due to $\tilde{i}$ or $\tilde{k}$.

$$\mathcal{M}(\tilde{i}\tilde{k} \to ijk) = \mathcal{M}(\tilde{i} \to ij) + \mathcal{M}(\tilde{k} \to jk) \tag{3.2.42}$$

The treatment in DGLAP showers is justified in the collinear limit where

$$|\mathcal{M}(\tilde{i}\tilde{k} \to ijk)|^2 \xrightarrow{i\|j \text{ or } j\|k} |\mathcal{M}(\tilde{i} \to ij)|^2 + |\mathcal{M}(\tilde{k} \to jk)|^2 + \text{negligible interference} \tag{3.2.43}$$

which considers only so-called dipoles. More generally interference terms within Feynmann diagrams are non-negligible. A more complete description requires $2 \to 3$ kernels as opposed the DGLAP $1 \to 2$ dipole splittings. For instance, the splitting kernel for $q\bar{q} \to qg\bar{q}$ is given by

$$a_{g/\tilde{q}\tilde{\bar{q}}} = \frac{|\mathcal{M}(\tilde{q}\tilde{\bar{q}} \to q\bar{q}g)|^2}{|\mathcal{M}(\tilde{q}\tilde{\bar{q}})|^2} \tag{3.2.44}$$

$$= \frac{\left| \raisebox{-0.5em}{\includegraphics{}} \right|^2}{\left| \raisebox{-0.5em}{\includegraphics{}} \right|^2} = \frac{\left| \raisebox{-0.5em}{\includegraphics{}} \right|^2 + \left| \raisebox{-0.5em}{\includegraphics{}} \right|^2 + 2\mathcal{R}\left(\raisebox{-0.5em}{\includegraphics{}} \raisebox{-0.5em}{\includegraphics{}}^* \right)}{\left| \raisebox{-0.5em}{\includegraphics{}} \right|^2} \tag{3.2.45}$$

Antenna functions, denoted by $a_{j/\tilde{i}\tilde{k}}$, may be used for this purpose reproducing the known forms given in equations (3.2.13) and (3.2.14) in the relevant limit. However, this is not strictly

Figure 3.4. Matrix amplitudes of neighbouring antennae. In the approximation of unique colours (see §3.2.2), an antenna is formed between each quark-antiquark colour pairs and antennae are formed between neighbouring gluons. Hence two matrix amplitudes of antennae contribute to emissions from a single dipole collinearly. Hence, for gluons, neighbouring antenna must be summed to reproduce the collinear limit. It is assumed implicitly that partons before emission are denoted with a tilde and antennae emissions are denoted by g_w, g_x, g_y, g_z.

the case. Instead, neighbouring matrix amplitudes will contain the partons shared between neighbouring antennae. Hence a collinear emission will have contributions from both matrix amplitudes. Antennae are only formed between colour-anticolour pairs. Within a shower where all colours are unique (see §3.2.2), antenna functions reproduce all limits since each parton is only a member of a single antenna. The exception is when considering gluons which will be members of two antennae owing to the colour-anticolour charge they carry. An illustration of these ideas is given in figure 3.4. Derivations of antenna functions are given in [58] and the procedure for obtaining the singular limits from antenna function is provided in [59]. Within the literature, antenna functions may be given in terms of $s_{ab} = 2p_a \cdot p_b$. The phase-space variable is also modified to $d\Phi_{\text{ant}+1}(s_{ij}, s_{jk}, \phi)$.

The consideration of an emitter as being an antenna as opposed to a dipole naturally incorporates so-called colour coherence into the shower. The effect was originally observed in QED where electric dipoles (for example, through pair production from cosmic rays as originally observed by Chudakov) would only radiate within a cone of opening bounded by the opening angle of the pair. Hence, for small opening angles the pair would be seen as neutrally charged to a soft photon at a wide-angle until such an separation where the charges are coherent. This behaviour naturally emerges from the cross-section obtained from the antenna function, a derivation of this result can be found in [8], section 3.3. This effect is equally observed between QCD colour dipoles.[5] For details on the Chudakov effect, the reader is referred to [60]. While this effect is naturally taken into consideration in antenna showers, the effect must be considered explicitly in DGLAP showers through a suitable evolution variable which imposes angular ordering i.e. ever decreasing emission angles.[6] An instructive illustration of the emission cone is shown in figure 3.5a for some process. While the latest version of Vincia evolves timelike showers in decreasing momentum which preserves angular ordering, Pythia 6 evolved showers based on virtually. The effect of vetoing emissions which didn't comply with the Chudakov effect as opposed to no further consideration is shown in figure 3.5b.

---

[5]A subtlety here is that wide-angle emissions may occur but when averaged azimuthally, are zero.

[6]Herwig uses pseudorapidity to incorporate angular ordering.

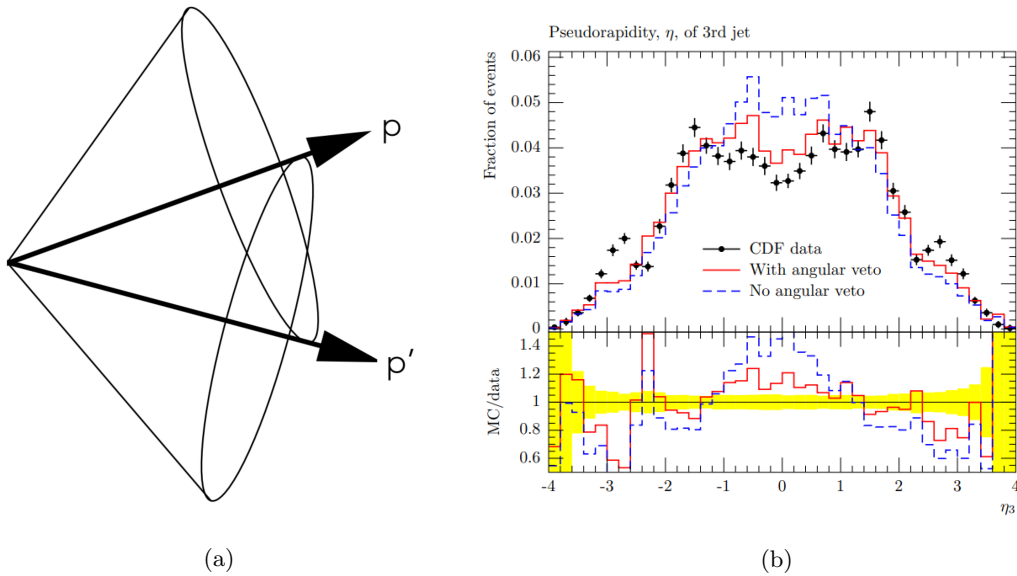(a)   (b)

Figure 3.5. (a) Region of net bremsstrahlung emission due to restriction by Chudakov effect. Obtained from [61]. [61] is a slide pack for masters lectures at Monash. The reader is to contact Prof. Peter Skands for access. (b) Pythia 6 showering with and without angular ordering. Data from Tevatron of $p\bar{p}$ collision at $\sqrt{s} = 1.8$TeV. Obtained from [62].

3.2.4 Vincia Kinematic Maps

(y_{ij}, y_{jk}) **phase-space**

We briefly discuss so-called kinematic maps[7] which determine how the properties such as momentum are assigned to the partons introduced into the final state. As was mentioned in passing in 3.2, in Pythia, all final state partons are given momenta such that the entire final state is on-shell. Mappings depend on the type of radiation at each end of the antenna i.e. initial-initial, initial-final resonance-final and final-final. In developing the theory of this section and the next, we restrict ourselves to final-final recoil maps of massless partons.

A description of the kinematics of the final state, as a matter of convenience, is given in the COM frame of the parent partons $\tilde{i}\tilde{k}$ or equivalently ijk by conservation of momentum. The four-momenta in this frame are characterized by the variables θ_{ij}, θ_{ik} in the x-z plane and $E_a = |\mathbf{p}_a|$ for $a = i, j, k$ i.e. the partons are assumed massless. The system is anchored by aligning parton i with the z-axis. So the four-momenta are given by

$$p_i^\mu = (|\mathbf{p}_i|, 0, 0, |\mathbf{p}_i|) \tag{3.2.46}$$

$$p_j^\mu = (|\mathbf{p}_j|, -|\mathbf{p}_j|\sin\theta_{ij}, 0, |\mathbf{p}_j|\cos\theta_{ij}) \tag{3.2.47}$$

$$p_k^\mu = (|\mathbf{p}_k|, |\mathbf{p}_k|\sin\theta_{ik}, 0, |\mathbf{p}_k|\cos\theta_{ik}) \tag{3.2.48}$$

with

$$E_i = \frac{s_{ij} + s_{ik}}{2m_{\tilde{i}\tilde{k}}^2} \tag{3.2.49}$$

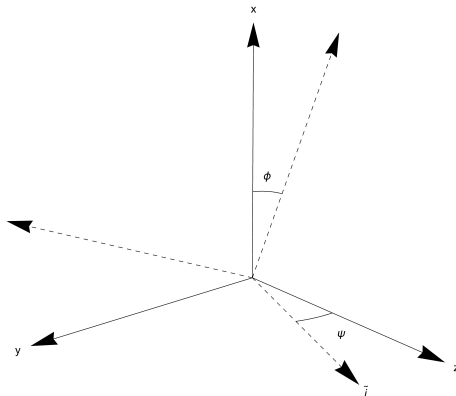[7]This is also referred to as recoil schemes/maps.

Figure 3.6. Dashed coordinates denote basis of parent partons which are referred to with primed axes labels. The coordinate axes of the daughter partons are obtained through transformation $x^i = R_x(\psi)R_{R_x^{-1}(\psi)}(\phi)x^{i'}$ where $i = 1, 2, 3$.

where $s_{ij} = 2p_ip_j$ with the other momenta obtained through cyclic permutation $i \to j \to k \to i$ etc. The angles are given by

$$\cos\theta_{ib} = 1 - \frac{s_{ib}}{2|\mathbf{p}_i||\mathbf{p}_b|} \qquad (3.2.50)$$

where $b = j, k$. While these variables specify the four-momenta of the daughter partons entirely, two additional angles as measured from the branchers plane are required to complete the mapping. The parent parton, $\tilde{i}$, is defined to point in the z' direction. The z-axis for the daughter partons is determined by rotating by angle ϕ about the z'-axis followed by a rotation by ψ in the x-axis. ψ can be thought as the angle $\theta_{\tilde{i}i}$. The relative orientations of the coordinate axes of the parent and daughter partons is illustrated in figure 3.6.

Given some process $\tilde{i}\tilde{k} \to ijk$, it is customary to describe the phase-space in terms of the fractional invariant masses [8]

$$y_{ab} = \frac{s_{ab}}{m_{\tilde{i}\tilde{k}}^2} \qquad (3.2.51)$$

with

$$y_{ij} + y_{jk} + y_{ik} = 1 \qquad (3.2.52)$$

Although the phase-space may be characterized by any pair of variables in equation (3.2.52), we will use (y_{ij}, y_{jk}) to coincide with the literature. The full phase-space is therefore characterized by the region $y_{ij} + y_{jk} \leq 1$ as shown in figure 3.7. If we let $m_{\tilde{i}\tilde{k}} = 1GeV$, then $y_{ab} = s_{ab}$ (in magnitude only) which can be substituted directly into equations (3.2.49) and (3.2.50) to obtain expressions for equations (3.2.46) to (3.2.48). Given these results we may visualize the level curves for various quantities in the (y_{ij}, y_{jk}) phase-space. Figure 3.7b displays the level curves for θ_{ij}, those of θ_{jk} are simply the reflection of θ_{ij} across the 45° line. Most importantly we can see that the collinear limit $i \parallel j$ is located on the vertical axis (the converse is true for θ_{jk}). The level curves of $E_k = |\mathbf{p}_k|$ are lines of gradient -1 with increasing energy fractions the larger the length of the perpendicular bisector as shown in figure 3.7a. The kinematic configuration can be fully visualized using a Dalitz plot in figure 3.7c. Massless kinematic maps for other radiation configurations are given in [64]. A

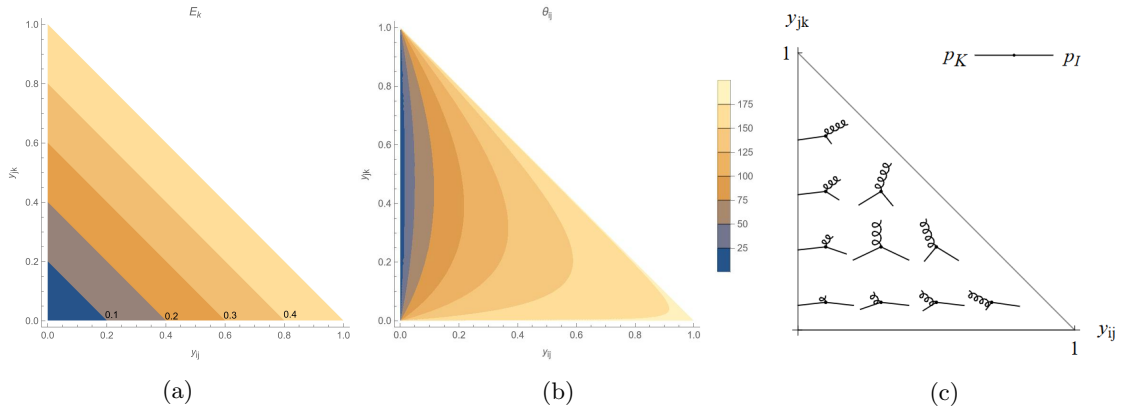[8]This definition differs based on the nature of the radiation.

Figure 3.7. (a) E_k level curves given by lines of gradient -1. (b) θ_{ij} level curves with collinear limit corresponding to $y_{ij} \to 0$. The θ_{jk} is given in level curves are reflected along the $45°$ line. (c) Dalitz plot for visualizing final-final Vincia phase-space. Insets are drawn using equations (3.2.46) to (3.2.48). The inset is oriented so that the gluon points upwards. Obtained from [63].

derivation for massive partons is given in [63].

The Lund Plane

An alternative representation of the phase-space which will prove instructive in §4.3.3 is the Lund plane. The phase-space is now parameterized in terms of $\ln p_\perp$ and η (pseudo-rapidity) which results in a triangular phase-space. Within this parameterization, in a $p_\perp$ ordered shower, subsequent emissions will sweep the Lund plane from the top to the bottom. This parameterization comes with several useful properties, we may easily visualize an extension of the phase-space due to subsequent emissions which will each have their embedded Lund planes. However, it can be shown that the emission probability density throughout plane is given by

$$\rho(\eta, \ln p_\perp) = \frac{dn_{emission}}{d\ln p_\perp \, d\eta} \approx \frac{2\alpha_s(p_\perp)\mathcal{C}}{\pi} \tag{3.2.53}$$

where $\mathcal{C}$ is a colour factor. Hence the density is constant aside from the running of the coupling. To preserve this, subsequent emissions modify this phase-space with additional "leaves" which have the same density of states modified by the colour factor of the emission. The additional triangular phase-space of the emission is mapped to the two sides of the leaf. Hence, emissions from the daughter parton result in additional leaves off the parent parton leaf and multiple emissions from the same parton result in leaves arising in different regions of the phase-space. This is illustrated well in [65].

3.2.5 Full Showering Algorithm

Having established the components of a parton shower, we are ready to outline the procedure entirely. Note, that there are several variations on the one presented here.

The shower begins with the hard cross section calculated using the Large Hadron Collider (LHC) master formula 3.1.1 to some order. The results of these calculations are built into Pythia. If this state contains n-partons, then each colour connected pair will have a corrisponding associated antenna function. A trial branching scale is generated using the veto algorithm (see appendix A) which samples the summed distribution providing a branching scale t. The antenna pair chosen to branch is selected according to the relative proportion of the antenna kernels at the scale t. Since the evolution variable is $p_\perp^2$, then this only fixes the branching to a contour within the phase-space.
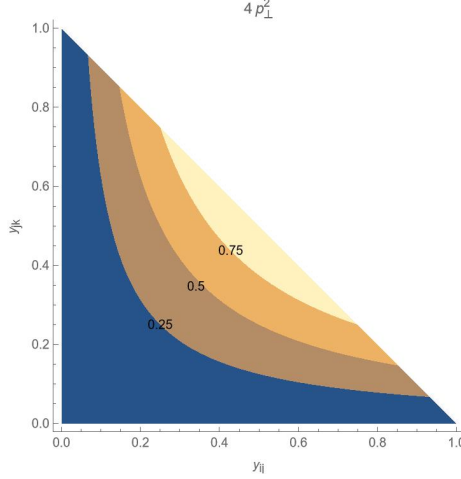
Figure 3.8. Phase-space in terms of antenna formalism variables with contours indicating constant transverse momentum.

A point is selected by fixing the z-parameter in equation (3.2.29) and the azmuthial branching angle ϕ as specified in figure 3.6 in the COM frame. The kinematics are determined as specified in §3.2.4 and boosted back to the lab frame. This process is repeated until a stopping scale is reached, this is usually chosen to be the hadronization scale where a hadronization model and jet clustering algorithms take hold.

The heart of MCMC method can be concisely written. To that end, we rewrite equation (3.1.2) as follows

$$\frac{d\sigma_{H_1 H_2 \to n}}{d\mathcal{O}} = \sum_{ab} \int dx_a dx_b f_{a/H_1}(x_a, \mu_F^2) f_{b/H_2}(x_b, \mu_F^2) \int d\Phi_n \, \hat{\sigma}_{ab \to n}(\mu_F^2) \delta(\mathcal{O} - \mathcal{O}(\{p\}_n))$$

(3.2.54)

where we have simply rewritten the differential cross-section over all final parton momentum configurations with the delta function projecting the cross-section onto observable $\mathcal{O}$ evaluated with the momentum configuration as in equation (3.1.2). Rather than evaluating this expression at the born level, we may simply choose to evaluate this expression after the addition of a parton i.e. after some showering. This will be encapsulated within the showering operator $\mathcal{S}$ which we expect to simplify to equation (3.2.54) when evaluated at the born level. $\mathcal{S}$ must simply evaluate the cross-section at the born and born+1 levels with relative weightings depending on the probability of branching i.e.

$$\mathcal{S}(\{p\}_n, t_0, t_1) = \Delta(t_0, t_1)\delta(\mathcal{O} - \mathcal{O}(\{p\}_n)) - \sum_{\substack{\text{possible branchings} \\ \tilde{i}\tilde{k} \to ijk}} \int_{t_0}^{t_1} dt \frac{d\Delta(t_0, t)}{dt} \mathcal{S}(\{p\}_{n+1}, t, t_1)$$

(3.2.55)

where $t_0 > t_1$. Recall, $\frac{d\Delta}{dt}$ is as given in equation (3.2.28) which is the branching probability. Over many iterations equation (4.2.8) will grow to divide the calculation of the total cross-section through smaller and smaller weighted contributions from the many branchings happening at the various scales. It is clear that $\mathcal{S}$ will preserve unitarity since the embedded showering operator will begin evolving from the chosen branching scale with each iteration adding to the cross-section by weighting the branching and no-branching cross-sections.

Chapter 4

Vincia Variations

Estimation of uncertainties is of paramount importance to properly compare collider data with predictions. However, the complexity of parton showers means that the estimation of uncertainties is highly non-trivial. With a new era approaching which will produce measurements with unprecedented precision, placing uncertainties on theoretical predictions has become urgent. For instance, the upgrades for the LHC can be found in [66, 67]. Traditionally, shower uncertainties are obtained by rerunning the shower with varied parameters. However, this is unfeasible for many variations and so we present an alternative formulation [68]. We begin with a description of the theoretical basis for automated uncertainty variations where we translate the results of [68] for DGLAP showers into the antenna formalism. We then present the results of the implementation via reweighting and rerunning the showers. The reader is referred to [69] for the implementation of variations for an older version of Pythia. For an alternative take on the shower variations as applied to Herwig 7, the reader is referred to [70].

4.1 Vincia Automated Uncertainty Method

As has been established, an event is predicated on probabilities depending on the kernels of the various branchers and channels. Hence, the probability that some particular event emerges relies on the kernels and parameters contained therein. For notational convenience, we will denote the shower scale with $t = p_\perp^2$.

Suppose our interest lies in the likelihood of some event, E, containing n branchings. The sequences $\{j_a\}, \{\tilde{i}_a\}, \{\tilde{k}_a\}$, denote the subsequent daughter and parent partons respectively at scales $\{t_a\}$. Hence, the overall probability density of this event, if written in terms of antenna functions is given by

$$P_E = \prod_{a=1}^{n} \Delta(t_{a-1}, t_a) a_{j_a/\tilde{i}_a \tilde{k}_a}(t_a) \tag{4.1.1}$$

where we assume a strong ordering i.e. $t_{a-1} \geq t_a$. We require that equation (4.1.1) reproduce the appropriate kernels in the soft and collinear limits [1] [2]

$$P_E \xrightarrow{\text{single-unresolved}} \prod_{a=1}^{n} \Delta(t_{a-1}, t_a) K_{j_a/\tilde{i}_a \tilde{k}_a} \tag{4.1.2}$$

---

[1] It is implicit that the antenna function in the Sudakov factor also changes.

[2] It is implicit that in some single unresolved limit, there are collinear and soft limits as in equations (3.2.13) and (3.2.14) despite the suggestive usage here were we have assumed a soft limit.

The particular sampling method used in parton showers differs from that of §2.5.2 slightly. The reader is referred to appendix A although this is not vital to understand the coming analysis. Equation (4.1.1) can be further explicitly written to contain the blanketing functions, $\hat{a}$ and $\hat{\Delta}$.[3] $\hat{\Delta}$ uses the same blanketing function in the exponential. Recall, as specified in §2.5.2, the probabilities of accepting or rejecting a trial sampling are given by

$$P_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t) = \frac{a_{j/\tilde{i}\tilde{k}}(t)}{\hat{a}(t)} \qquad (4.1.3) \qquad P_{\text{rej}}^{j/\tilde{i}\tilde{k}}(t) = 1 - P_{acc}^{j/\tilde{i}\tilde{k}}(t) \qquad (4.1.4)$$

Each branching then is the result of several rejected trials followed by a success. Hence, we introduce an additional sequence $\{l_a\}$ denoting the number of rejected trials per branching. The length of this sequence is the number of branchings which occurred in the shower. For each l_a there is an associated sequence $\{t_b^{(a)}\}$ of length l_a of the intermediate scales where samplings where rejected. Hence, the true probability density is given by

$$\prod_{a=1}^{n} \underbrace{\hat{\Delta}(t_{l_a}^{(a)}, t_a)\hat{a}(t_a)P_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a)}_{\text{prob. of accepting at scale } t_a} \prod_{b=1}^{l_a} dt_b^{(a)} \underbrace{P_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)})\hat{\Delta}(t_{b-1}^{(a)}, t_b^{(a)})\hat{a}(t_b^{(a)})}_{\text{prob. rejecting branching at scale } t_b^{(a)}} \qquad (4.1.5)$$

where $t_0^{(a)}$ is defined as the starting scale of the chosen brancher.[4]

Traditionally, uncertainties on shower predictions are obtained by varying the parameters of the shower manually and rerunning the shower to see the corresponding effect on the prediction. This is computationally inefficient and error-prone. In addition, VINCIA may not offer a setting for varying some parameter of the shower so the user must manipulate the code directly. The method we will present also has the advantage that any number of variations can be obtained from a single shower with little additional computational overhead. For the moment, we will make no direct specification of the source of uncertainty within the kernel. Instead, a modified kernel will be denoted by primed variable.

For every sample of the branching distribution, a reweighting is applied for every accepted and rejected trial sampling from the blanketing distribution given by

$$R_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t) = \frac{a'_{j/\tilde{i}\tilde{k}}(t)}{a_{j/\tilde{i}\tilde{k}}(t)} \qquad (4.1.7) \qquad R_{\text{rej}}^{j/\tilde{i}\tilde{k}}(t) = \frac{1 - P'^{j/\tilde{i}\tilde{k}}_{\text{acc}}(t)}{1 - P_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t)} \qquad (4.1.8)$$

where $P' = a'/a$. We now prove the success of this approach by deriving the kernel of this procedure. The proof is analogous to that found in [68].

Proof: The probability of a branching to occur is composed of the probability that the branching occur after any number of rejected trials i.e.

$$d\mathcal{P} = \sum_n d\mathcal{P}_n \qquad (4.1.9)$$

where $d\mathcal{P}_n$ denotes the probability of an emission after n rejected trials within region $(t, t + dt)$. The first few are shown below to give the reader the general idea of how these probabilities are

---

[3]Although it is possible to use a different blanketing function for each channel. For simplicity, we assume a single function since the generalization is clear.

[4]To help make clear the notation, equation (7) of [68] in our notation is written as

$$\hat{\Delta}(t_1^{(1)}, t_1)P_{\text{acc}}(t_1)\hat{a}(t_1)dt_1^{(1)} P_{\text{rej}}(t_1^{(1)})\hat{\Delta}(t_0^{(1)}, t_1^{(1)})\hat{a}(t_1^{(1)}) \qquad (4.1.6)$$

where in [68], $t_1^{(1)}$ is denoted by t_1, t_1 is denoted by t and $t_0^{(1)}$ is denoted by t_0.

produced.

$$dP_0 = dt\, R_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t) P_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t)\hat{a}(t)\hat{\Delta}(t_0, t) \tag{4.1.10}$$

$$dP_1 = dt \int_t^{t_0} dt_1 \underbrace{\hat{a}(t_1)\hat{\Delta}(t_0, t_1) P_{\text{rej}}^{j/\tilde{i}\tilde{k}}(t_1) R_{\text{rej}}^{j/\tilde{i}\tilde{k}}(t_1)}_{\text{failed branching at } t_1} \underbrace{\hat{a}(t)\hat{\Delta}(t_1, t) P_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t) R_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t)}_{\text{successful branching at } t} \tag{4.1.11}$$

$$= dt\, a_{j/\tilde{i}\tilde{k}} R_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t)\hat{\Delta}(t_0, t) \int_t^{t_0} dt_1 \left(\hat{a}(t_1) - a'_{j/\tilde{i}\tilde{k}}(t_1)\right) \tag{4.1.12}$$

$$\vdots$$

$$dP_n = dt\, a_{j/\tilde{i}\tilde{k}} R_{\text{acc}}^{j/\tilde{i}\tilde{k}}(t)\hat{\Delta}(t_0, t) \int_t^{t_0} dt_1 \left(\hat{a}(t_1) - a'_{j/\tilde{i}\tilde{k}}(t_1)\right) \times \ldots \times \int_t^{t_{n-1}} dt_n \left(\hat{a}(t_n) - a'_{j/\tilde{i}\tilde{k}}(t_n)\right) \tag{4.1.13}$$

$$= dt\, a'_{j/\tilde{i}\tilde{k}}\hat{\Delta}(t_0, t)\frac{1}{n!} \left(\int_t^{t_0} dt_1 \left(\hat{a}(t_1) - a'_{j/\tilde{i}\tilde{k}}(t_1)\right)\right)^n \tag{4.1.14}$$

where the integral in equation (4.1.11) is to sum all possible intermediate scales where the single rejection occurs. Each rejection will result in an additional embedded integral. The final equality was obtained using the result

$$\int_{t_0}^t dt_1\, f(t_1) \int_{t_0}^{t_1} dt_2\, f(t_2) \cdots \int_{t_0}^{t_{n-1}} dt_n\, f(t_n) = \frac{1}{n!} \left(\int_{t_0}^t dt'\, f(t')\right)^n \tag{4.1.15}$$

So the overall kernel is given by

$$\frac{dP}{dt} = \sum_n \frac{dP_n}{dt} \tag{4.1.16}$$

$$= a'_{j/\tilde{i}\tilde{k}}\hat{\Delta}(t_0, t) \sum_n^\infty \frac{1}{n!} \left(\int_t^{t_0} dt_1 \left(\hat{a}(t_1) - a'_{j/\tilde{i}\tilde{k}}(t_1)\right)\right)^n \tag{4.1.17}$$

$$= a'_{j/\tilde{i}\tilde{k}} \exp\left(-\int_t^{t_0} dt_1\, \hat{a}(t_1)\right) \exp\left(\int_t^{t_0} dt_1 \left(\hat{a}(t_1) - a'_{j/\tilde{i}\tilde{k}}(t_1)\right)\right) \tag{4.1.18}$$

$$= a'_{j/\tilde{i}\tilde{k}} \exp\left(-\int_t^{t_0} dt_1\, a'_{j/\tilde{i}\tilde{k}}(t_1)\right) \tag{4.1.19}$$

$$\square$$

Hence, equation (4.1.5) becomes

$$\prod_{a=1}^n \underbrace{\hat{\Delta}(t_{l_a}^{(a)}, t_a)\hat{a}(t_a) R_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a) P_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a)}_{\text{prob. of accepting at scale } t_a} \prod_{b=1}^{l_a} dt_b^{(a)} \underbrace{R_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)}) P_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)})\hat{\Delta}(t_{b-1}^{(a)}, t_b^{(a)})\hat{a}(t_b^{(a)})}_{\text{prob. rejecting branching at scale } t_b^{(a)}} \tag{4.1.20}$$

That is, the shower with modified kernel a' will be reweighted by an overall factor of

$$w = \prod_{a=1}^n R_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a) \prod_{b=1}^{l_a} R_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)}) \tag{4.1.21}$$

We may obtain the pdf of this event by summing all possible intermediate rejected states which

lead to the final scale. We obtain

$$\prod_{a=1}^{n} \hat{a}(t_a) R_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a) P_{\text{acc}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_a) \hat{\Delta}(t_0^{(a)}, t_a) \sum_{l_a=0}^{\infty} \prod_{b=1}^{l_a} \int_{t_a}^{t_{b-1}^{(a)}} dt_b^{(a)} R_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)}) P_{\text{rej}}^{j_a/\tilde{i}_a\tilde{k}_a}(t_b^{(a)}) \hat{a}(t_b^{(a)})$$

$$(4.1.22)$$

$$= \prod_{a=1}^{n} a'_{j_a/\tilde{i}_a\tilde{k}_a}(t_a) \hat{\Delta}(t_0^{(a)}, t_a) \sum_{l_a=0}^{\infty} \frac{1}{l_a!} \left(\int_{t_a}^{t_0^{(a)}} dt\, \hat{a}(t) - a'_{j_a/\tilde{i}_a\tilde{k}_a}(t) \right)^{l_a} \qquad (4.1.23)$$

$$= \prod_{a=1}^{n} \Delta'(t_{a-1}, t_a) a'_{j_a/\tilde{i}_a\tilde{k}_a}(t_a) \qquad (4.1.24)$$

Observe that the primed shower explores the same region of the phase-space as the original shower. Hence, regions of phase-space which are underrepresented by one shower i.e. under sampled will impact the primed shower which may have preference for a suppressed region of the phase-space region.

Recall, as discussed in §3.2 that within the single-unresolved limits the soft-eikonal and/or collinear terms will arise within the appropriate limit independent of the particular process; for such terms, variations aren't applicable since these structures are universal. Instead, there are three valid classes of variations which we now introduce in turn; each of which is of interest for different reasons. If we look at the process in equation (3.2.44), the corresponding antenna function is given by [64]

$$a_{g/\tilde{q}\tilde{\bar{q}}} = \frac{4\pi\alpha_s(\,t\,)\,\mathcal{C}_{g/\tilde{q}\bar{q}}}{s_{\tilde{q}\tilde{\bar{q}}}} \left(\frac{(1-y_{qg})^2 + (1-y_{g\bar{q}})^2}{y_{qg}y_{g\bar{q}}} - \frac{2\mu_{\tilde{q}}^2}{y_{qg}^2} - \frac{2\mu_{\tilde{\bar{q}}}^2}{y_{qg}^2} + \boxed{1} \right) \qquad (4.1.25)$$

where $s_{ij} = 2p_i \cdot p_j$, $s_{\tilde{i}\tilde{k}} = m_{\tilde{i}\tilde{k}}^2$, $y_{ij} = \frac{s_{ij}}{s_{\tilde{q}\tilde{\bar{q}}}}$, $\mu_j^2 = \frac{m_j^2}{s_{\tilde{q}\tilde{\bar{q}}}}$ and the three possible sources of variations considered in this thesis are highlighted. Renormalization-scale variations (red) encapsulate the uncertainty in the coupling measurement at the Z-mass. Constant term variations (yellow) parameterizing nonsingular terms i.e. terms of the form $\propto \frac{1}{s_{ab}}$ allow for the use of antenna functions which differ by only the trailing finite term but will correctly reproduce the singular limits. Since this term is parameterization dependent, fixing this value would be an arbitrary choice and is therefore ambiguous away from the singular-limits. Finally, colour factor variations (green) investigate the variation in the shower due to sub-leading effects in colour. This will be discussed in further detail in §4.3.3. Hence, if we were able to incorporate subleading colour effects, this variation would be of no use.

The reader is reminded that the method as presented in [68] assumes DGLAP kernels while we will cast all equations in terms of antenna functions. It will be convenient to use colour and coupling-stripped antenna functions as follows

$$a_{j/\tilde{i}\tilde{k}} = 4\pi\alpha_s C_{j/\tilde{i}\tilde{k}} \bar{a}_{j/\tilde{i}\tilde{k}} \qquad (4.1.26)$$

The implementation and results are only valid for FSR. The implementation is studied through the pair annihilation $e^+e^- \to Z^0 \to q\bar{q} \to$ hadrons for its simplicity. The e^+e^- are given a combined COM energy of m_Z. Bremsstrahlung from the initial state was explicitly switched off leaving only FSR as required.

We will investigate the validity of this method through comparison of the distributions of several observables as obtained through a reweighting of a nominal shower and by rerunning the shower with new parameters. Each plot contains three stacked panels, the top panel shows the differential cross-section for the default weight in black and the two variations in red and blue. The bottom

panel overlays the predicted uncertainties i.e. via rerunning the shower with varied parameters to reweighting the shower denoted with a hash pattern. Specifically, the bottom panel plots

$$\frac{d\sigma/d\mathcal{O}\big|_{\text{rerun}}}{d\sigma/d\mathcal{O}\big|_{\text{weighted}}} \tag{4.1.27}$$

The central panel shows the ratio of the prediction to the reweighted shower as to exemplify adherence (or nonconformity) of the two. Variations are turned on in VINCIA through the setting `UncertaintyBands:doVariations = on`. VINCIA allows for a combinations of variations to be applied at once, see [15] for details. Before continuing, we briefly describe the variables which will be used in subsequent plots.

4.2 Event Variables

Our analysis will involve comparisons of the distributions of three variables. Each distribution is built through calculations obtained from several event samples and placed in a histogram with an appropriate weighting factor. The variables under consideration help characterize the event shape through the flow of energy and momentum.

4.2.1 Thrust, Major and Minor Axes

In principal component analysis (PCA) the aim is to find a new basis on some vector space with each basis component pointing in the direction which jointly carries the most information about the data set. This is dubbed the first principal component, subsequent components are dubbed second, third etc. principal components. While the aim of PCA is data compression, here it is to characterize the jet structure of the event.[5] The so-called thrust axis is given by

$$\vec{n}_T = \underset{|\vec{n}|=1}{\arg\max} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} \tag{4.2.1}$$

and the thrust

$$T = \max_{|\vec{n}|=1} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}_T|}{\sum_i |\vec{p}_i|} \tag{4.2.2}$$

The major and minor axes are analogously defined as

$$\vec{n}_{\text{major}} = \underset{\vec{n}-\text{proj}_{W_1}\vec{n},|\vec{n}|=1}{\arg\max} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} \tag{4.2.3} \quad \vec{n}_{\text{minor}} = \underset{\vec{n}-\text{proj}_{W_2}\vec{n},|\vec{n}|=1}{\arg\max} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} \tag{4.2.5}$$

$$\text{major} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}_{\text{major}}|}{\sum_i |\vec{p}_i|} \tag{4.2.4} \qquad \text{minor} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}_{\text{minor}}|}{\sum_i |\vec{p}_i|} \tag{4.2.6}$$

where $W_1 = \text{span}\,\vec{n}_T$ and $W_2 = \text{span}\{\vec{n}_T, \vec{n}_{\text{major}}\}$.

In a 2-jet event as shown in figure 4.1b, all momenta lie along the thrust axis and therefore the inner product does little to reduce the magnitude so $T \to 1$. Conversely, an event with jets

---

[5]The objective function to be maximized also differs slightly. Specifically, thrust doesn't aim to maximize variance and is normalized.
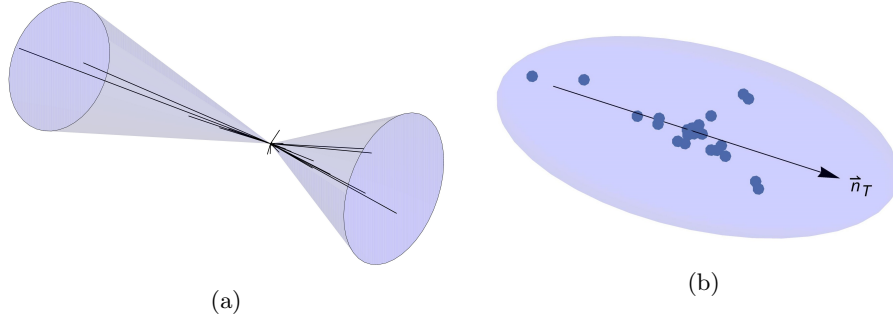
Figure 4.1. An event with two jets in the COM frame (a) and the momentum distribution plotted on p_x, p_y, p_z axes (b) with the thrust axis given by $(0.40, 0.89, -0.21)$. The thrust characterizes how collinear the jet substructure is to the jet axis. For a pencil-like event the thrust is one. Only the stable final state particles are shown.

evenly spread will be insufficiently characterized by the thrust axis, however, since the thrust axis captures the largest component, the thrust must be no less than 0.5. Otherwise, the major or minor components have the potential of being larger than T. So

$$0.5 \leq T \leq 1 \Rightarrow 0 \leq 1 - T \leq 0.5 \tag{4.2.7}$$

A detailed study of thrust can be found in [71].

4.2.2 C- and D-parameters

The discussion draws from [72]. Consider general tensor

$$S^{ab} = \frac{\sum_i p_i^a p_i^b |\vec{p}_i|^{r-2}}{\sum_i |\vec{p}_i|^r} \tag{4.2.8}$$

where p_i^a denotes the ath component of three-momentum and the sum is over all partons in the shower. When $r = 1$,[6] we obtain the linearized momentum tensor which is of interest here since it is collinear-safe. Recall, an observable is collinear-safe if it is unable to distinguish between collinear partons. If we consider a system a system with parton-momenta $\{p\}_{n+1}$ where partons n and $n+1$ are collinear and with momenta related by

$$p = z p_n + (1 - z) p_{n+1} \tag{4.2.9}$$

where z is the momentum fraction given to each parton with $0 \leq z \leq 1$. Then, only for $r = 1$ that

$$\sum_i^{n+1} |\vec{p}_i|^r = \sum_i^{n-1} |\vec{p}_i|^r + |\vec{p}_n|^r + |\vec{p}_{n+1}|^r = \sum_i^{n-1} |\vec{p}_i| + z|p_n| + (1 - z)|p_{n+1}| = \sum_i^{n-1} |\vec{p}_i| + |p| \tag{4.2.10}$$

This therefore implies $S^{ab}(\{p\}_{n+1}) \overset{r \equiv 1}{=} S^{ab}(\{\{p\}_{n-1}, p\})$ as required.

If the three eigenvalues of the linearized momentum tensor are given by $\lambda_1, \lambda_2, \lambda_3$, then

$$C = 3(\lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1) \tag{4.2.11}$$

$$D = 27 \lambda_1 \lambda_2 \lambda_3 \tag{4.2.12}$$

---

[6] $r = 2$ defines the sphericity tensor whose eigenvalues can be used to define sphereicity and aplanarity.

4.3 Vincia Variations Implementation and Results

4.3.1 Renormalization-scale

Since $a \propto \alpha_s$, the weights for renormalization-scale variations are given by

$$R^{j/\tilde{i}\tilde{k}}(t) = \frac{\alpha_s(kt)}{\alpha_s(t)} \tag{4.3.1}$$

For soft-gluon emissions, higher order effects can be absorbed into the LO kernels' typical constants using the so-called Catani-Webber-Marchesini (CMW) scheme [73]. The correction should have full effect in the soft limit and fades linearly further away. Since in the antenna formalism there are three partons in the final state, we force the correction to fade depending on the softest of the three partons

$$\zeta = \min(x_i, x_j, x_k) \tag{4.3.2}$$

where $x_a = 1 - y_{bc}$.

So the overall reweighting factor is given by

$$R^{j/\tilde{i}\tilde{k}}(t) = \frac{\alpha_s(kt)}{\alpha_s(t)} \times \begin{cases} \left(1 + \delta_{jg}(1 - \zeta)\alpha_s(\mu_{\max}^2)b_0 \ln k\right) & \text{gluon emission} \\ 1 & \text{otherwise} \end{cases} \tag{4.3.3}$$

b_0 is as defined in equation (2.4.34), $\mu_{\max}^2 = \max(m_{\tilde{i}\tilde{k}}^2, kt)$, δ_{jg} ensures the correction is only made to gluon emissions and ζ which ensures this correction remains only in the soft limit of gluon emissions. $\mu_{\max}^2$ is chosen to correspond to the largest scale in the problem to ensure the correction is as small as possible. Finally, since α_s will grow near the perturbative cut-off, we impose a limit $|\Delta\alpha_S| \leq \delta$ which can be set using `UncertaintyBands:deltaAlphaSmax`. This variation can be accessed through the setting `fsr:muRfac`. The results are presented in figure 4.2. We show no results within the sector formalism since the result should be and is entirely the same aside from the random fluctuations due to the Monte Carlo process which shrink with increasing statistics. The shower is rerun twice with the settings

$$\texttt{Vincia:renormMultFacEmitF} = 0.5, 2 \tag{4.3.4}$$

$$\texttt{Vincia:renormMultFacSplitF} = 0.5, 2 \tag{4.3.5}$$

The behaviour of the running coupling in the far-infrared regime is modified by shifting the scale by parameter μ_{freeze}^2 at which α_S is evaluated. The variation therefore incorporates this also for comparison and regularization. Hence the running coupling is evaluated at renormalization-scale

$$\mu_{\text{freeze}}^2 + kt \tag{4.3.6}$$

The bottom panel shows the factor by which the rerun shower differs from the default, these are denoted by square (blue) and cross (red) markers. The (forward) hashed region (weighted var.) shows the range over which the shower varies for any variation between renormalization-scales $0.5t$ and $2t$ as determined from the original run by applying weighting factors via the method outlined in §4.1.[7] The remaining plots are as described previously.

Since we are varying the renormalization scale, an observable's dependence on the strong coupling will become evident since there is a large running. Indeed, we can identify the size of the

---

[7]The regularization is implicit.

variations with the order of the coupling. In figure 4.2a, there is a peak in the first bin. This can be understood in terms of the Sudakov factor. Recall, that the distribution that is sampled in parton showers is of the form

$$-\alpha_S \frac{dP}{d\mathcal{O}} \exp(-\alpha_S P) \tag{4.3.7}$$

where P is the shower kernel integrated over some region of phase-space. The the Sudakov factor corresponds to an infinite-order resummation of perturbation theory which introduces various orders of α_S through the exponentiation of the coupling i.e.

$$\exp(-\alpha_S P) = \sum (-1)^n \frac{\alpha_S^n}{n!} P^n \tag{4.3.8}$$

The effect of the Sudakov factor isn't pronounced for the entire range of these observables since the resummed spectrum has good agreement with the first order approximation away from this bin. For sufficiently smaller values of $\mathcal{O}$, $d\sigma/d\mathcal{O}$ causes a peak in the spectrum dubbed the "Sudakov peak" followed by a suppression. The suppression isn't visible at the selected bin widths but the peaked uncertainty at the Sudakov peak is certainly visible. For the majority of the distribution there is a 10% variation corresponding to a dependence on a single order of α_S for $1 - T < 0.33$. Beyond this point, an additional order of α_S is introduced which roughly doubles the variation. The same can be seen in figure 4.2b where there is a peak around the Sudakov peak and a variation of around 20% throughout since the distribution of the C-parameter has second order dependence on the coupling. The reasoning is true for the D-parameter.

It is evident that the weighted shower is converging to the predicted value. The reweighted showers required approximately 20 million event samples for each variation bound. The prediction required only 5 million. We see that in this instance, explicitly rerunning the shower would have been quicker since the predictions required 10 million samples overall while the weighted showers required 20 million overall. With about 4 variations, the utility of reweighting becomes evident.
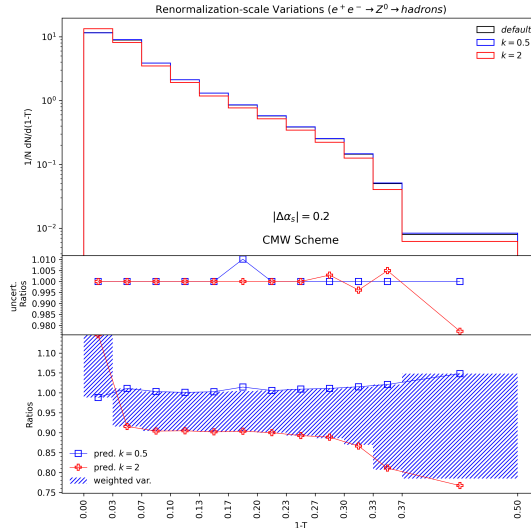
(a)
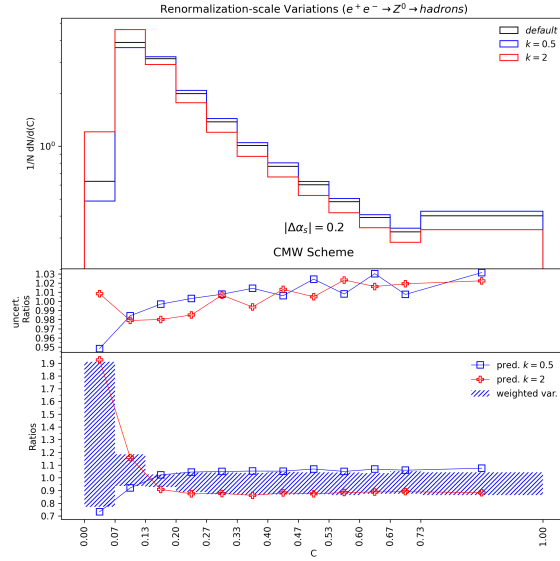
(b)
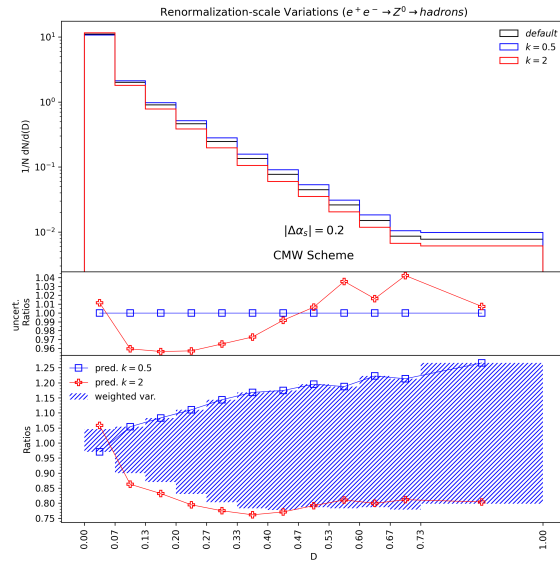
(c)

Figure 4.2. Comparison of renormalization-scale variations as obtained by reweighting the default shower and by rerunning the shower with modified antenna functions. The hashed region is bounded by the variation in the observables (a) 1-T (b) C (c) D over a factor of two of the shower scale. The crosses and squares overlayed are the predictions from rerunning the shower. The top panel shows the variation in the differential cross-section. The bottom panel is the ratio between the reweighted and default shower. The central panel is the ratio between predicted and reweighted values. Bins with fewer counts have been combined as to ensure sufficient statistics. The weighted showers were obtained from 20 million sampled events for each variation and the prediction was obtained from 5 million sampled events for each variation.

4.3.2 Non-singular term

For varying non-singular terms in the antenna function, we note that $\bar{a}_{j/\tilde{i}\tilde{k}} \propto \frac{1}{s_{\tilde{i}\tilde{k}}}$. Hence, we allow for variations of the non-singular terms after factorizing this common term so

$$\bar{a}_{j/\tilde{i}\tilde{k}} \to \bar{a}_{j/\tilde{i}\tilde{k}} + \frac{c_{\mathrm{NS}}}{s_{\tilde{i}\tilde{k}}} \tag{4.3.9}$$

where c_{NS} is dimensionless. Hence, the weight for the variation becomes

$$R^{j/\tilde{i}\tilde{k}}(t) = \frac{\bar{a}'_{j/\tilde{i}\tilde{k}}(t)}{\bar{a}_{j/\tilde{i}\tilde{k}}(t)} = \frac{(\bar{a}_{j/\tilde{i}\tilde{k}}(t) + \frac{c_{\mathrm{NS}}}{s_{\tilde{i}\tilde{k}}})}{\bar{a}_{j/\tilde{i}\tilde{k}}(t)} = 1 + \frac{c_{\mathrm{NS}}}{s_{\tilde{i}\tilde{k}}\bar{a}_{j/\tilde{i}\tilde{k}}(t)} \tag{4.3.10}$$

To ensure modified probabilities aren't negative, the reweighting factor is instead

$$R^{j/\tilde{i}\tilde{k}}(t) = 1 + \max\left(-1 + \mathrm{TINY}, \frac{c_{\mathrm{NS}}}{s_{\tilde{i}\tilde{k}}\bar{a}_{j/\tilde{i}\tilde{k}}(t)}\right) \tag{4.3.11}$$

where TINY is some arbitrary constant which was chosen to be 0.05. We note that this class of variation is also referred to as pdf-variations.

The choice for the finite-term is somewhat arbitrary since different terms can produce the same limiting behaviour. In addition, having this term vanish in one prescription isn't true for another. Hence, setting this term to zero is arbitrary. This variation may be accessed through the setting `fsr:cns`. It is important to note that there is no setting currently available which allows the user to modify the antenna functions in this way directly i.e. without weighting a shower. Instead, as was done for this thesis, the antenna functions must be modified within the code and recompiled. This is certainly not user friendly and error-prone.

The finite-term results are presented in figure 4.4. The shower is rerun twice with variations $c_{NS} = \pm 2$. The weighted variation used 40 million samples per variation while the prediction used 5 million per variation. Despite this, we note that there is a gap between the weighted shower in figure 4.4c and the prediction. Although the variation is ± 2. Away from the singular limits, the singular terms can become quite small which causes the finite-term to have a large influence on the antenna function and consequently the outcome of the shower. This can therefore produce larger variations [63]. Hence, the rejected and accepted reweighting factors must accumulate the desired overall weight using the same number of reweighting factors, so each reweighting factor is larger. Our implementation caps each reweighting factor as well as the overall weight, hence, the desired reweighting is often unable to reach the desired weight. We may remove the caps or include a so-called headroom factor which is a multiplicative factor of the blanketting antenna function. If we expand out the rejection reweighting factor we obtain

$$R_{\mathrm{rej}} = \frac{\hat{a} - a'}{\hat{a} - a} \tag{4.3.12}$$

Hence, if the blanketing kernel approximates well the original kernel then the denominator of the rejection becomes exceedingly small.[8] Since a close blanketing function will result in a large threshold for the accept-reject stage i.e. $a/\hat{a} \approx 1$ and therefore the sampling will likely be accepted. If we study, for instance, the weights due to $c_{\mathrm{NS}} = 2$ variation as shown in figure 4.3. Indeed, we see that increasing the headroom factor results in the distribution of weights to shift to right i.e. larger weights since each reweighting factor is smaller but more are produced. Hence, although we have seen that worse blanketing functions reduce the efficiency of the sampling method, we see that we now have an incentive to do so.
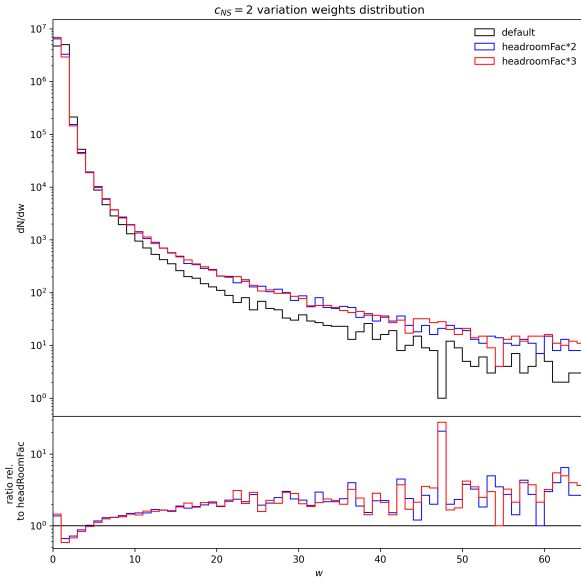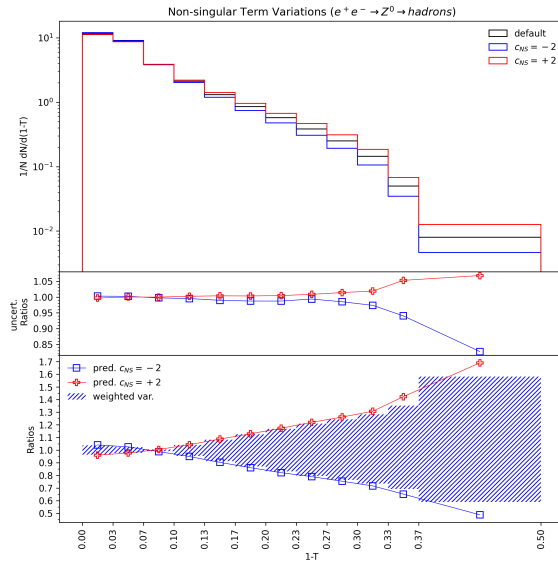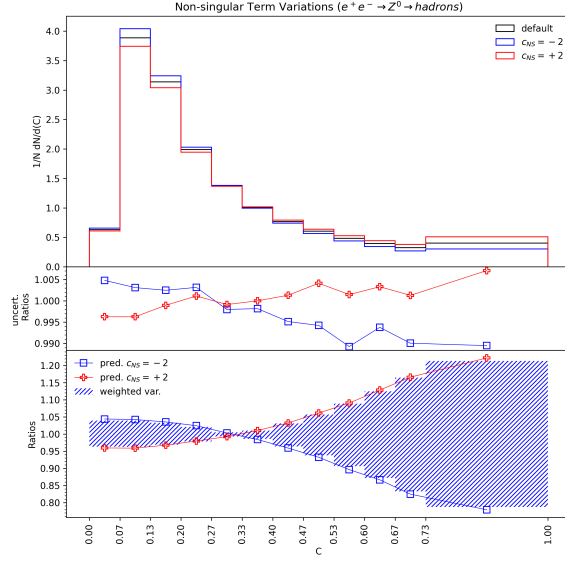

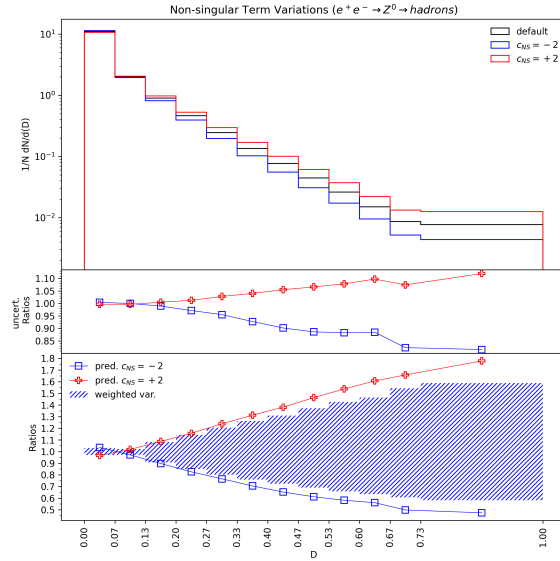[8]This could cause precision problems also.

Figure 4.3. Weight distribution of $c_{\mathrm{NS}} = 2$ for larger headroom factors. The larger headroom factor causes reweighting factors to shrink and not be capped which causes larger weights.

(a)

(b)

(c)

Figure 4.4. Comparison of non-singular term uncertainty variations as obtained by reweighting the default shower and by rerunning the shower with modified antenna functions. The hashed region is bounded by the variation in the observables (a) 1-T (b) C (c) D for range $c_{\mathrm{NS}} = \pm 2$. The crosses and squares overlayed are the predictions from rerunning the shower. The top panel shows the variation in the differential cross-section. The bottom panel is the ratio between the reweighted and default shower. The central panel is the ratio between predicted and reweighted values. The weighted variation used 40 million samples per variation bound while the prediction used 5 million per prediction.

4.3.3 Colour factor

Finally, the colour factor may be varied by reweighting factor

$$R^{j/\tilde{i}\tilde{k}}(t) = \frac{C'_{j/\tilde{i}\tilde{k}}}{C_{j/\tilde{i}\tilde{k}}} \tag{4.3.13}$$

Unlike other variations we have considered, the primary difficulty arises since it is unknown which colour factor should be used. However, the answer, if known, is unambiguous and colour factor variations are obsolete. The choice of colour factor is associated with the shower's treatment of colour at subleading order. Recall, the DGLAP kernels given in equation (3.2.20) where we see the colour factor when collinear to a quark/antiquark is given by C_F and by C_A when collinear to a gluon. This choice is unambiguous within the collinear limit. However, away from this limit where we consider emission from antenna, it is difficult to ascertain the colour factor which should be used. Now we know, regardless of the location of the emission within phase-space, the colour factor cannot exceed the bounds $[2C_F, C_A]$ since the coupling strength cannot be weaker than that of a quark or stronger than a gluon.[9] For antennas with parton-antiparton ends, the decision is unambiguous. For instance, the emission of a gluon from a quark-antiquark pair will have phase-space with quark coupling strength $2C_F$. If the first emission is a quark-gluon antenna, then we are again uncertain. Currently there are several basic suggestions for interpolating between the quark-gluon antenna collinear limits. A rudimentary treatment simply applies colour factor $\mathcal{C}^{(\mathrm{Ave})} = \frac{C_A + 2C_F}{2}$ to all antenna. This however, doesn't produce the known collinear limits. A slightly more sophisticated proposal correctly reproduces the collinear limits and does a simple interpolation

$$\mathcal{C}_{j/\tilde{g}\tilde{q}}(z_q, z_g) = \frac{2C_F z_q}{z_g + z_q} + \frac{C_A z_g}{z_g + z_q} \tag{4.3.14}$$

Indeed,

$$\mathcal{C}_{j/\tilde{g}\tilde{q}}(z_q, z_g) \to 2C_F \quad \text{as} \quad z_q \to 1 \tag{4.3.15}$$

$$\mathcal{C}_{j/\tilde{g}\tilde{q}}(z_q, z_g) \to C_A \quad \text{as} \quad z_g \to 1 \tag{4.3.16}$$

Hence, we refine the shower's subleading colour treatment by explicitly choosing the proper colour factor for the first emission and using $\mathcal{C}^{(\mathrm{Ave})}$ or equation (4.3.14) subsequently. The proper colour factor is easily visualized on a Lund plane since the colour factor is determined by whether the emission occurs on a leaf or otherwise. A discussion of this as well as a deeper analysis of subleading colour in parton showering can be found in [74]. The results of colour factor variations are shown in figure 4.5c. Weighted events used 10 million samples per variation bound while predictions used 5 million per prediction. The variation can be accessed through the setting `fsr:colvar`. The shower is rerun with settings

$$\texttt{Vincia:QQEmitFF:chargeFactor} \tag{4.3.17}$$
$$\texttt{Vincia:QGEmitFF:chargeFactor} \tag{4.3.18}$$
$$\texttt{Vincia:GGEmitFF:chargeFactor} \tag{4.3.19}$$

Once again, the predicted variations appear to coincide with the weighted variations with the features in figure 4.5c following similar reasoning as the renormalization-scale variations. Now we

---

[9]Here we use $2C_F$ in accordance with the normalization used in Pythia which ensures all antenna functions are normalized such that $a \propto \alpha_S/4\pi$.

also include a variation through weighting event samples except with the first emission fixed to $2C_F$ since the process begins with the production of a quark-antiquark pair.[10] This is denoted by hashing and highlighted yellow with legend "fixed first branching". As expected, this reduces the uncertainty bands. For completeness we have also included the shower results with the colour factor given by equation (4.3.14). The middle panel now also includes the ratio of the variation with the first colour factor fixed to that without i.e.

$$\frac{d\sigma/d\mathcal{O}\Big|_{\text{fixed first branching}}}{d\sigma/d\mathcal{O}\Big|_{\text{max. var.}}} \tag{4.3.20}$$
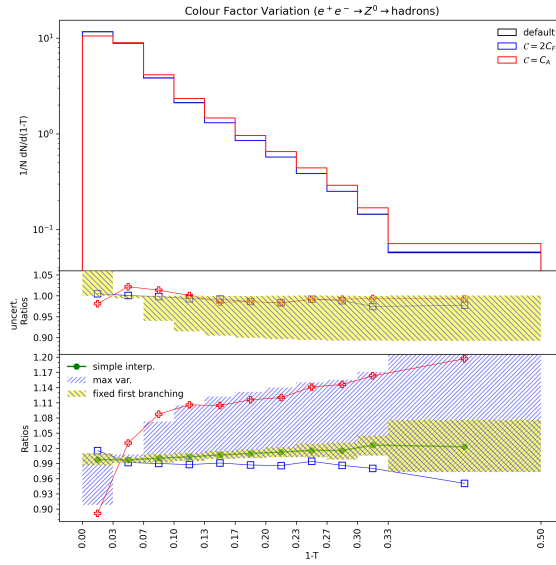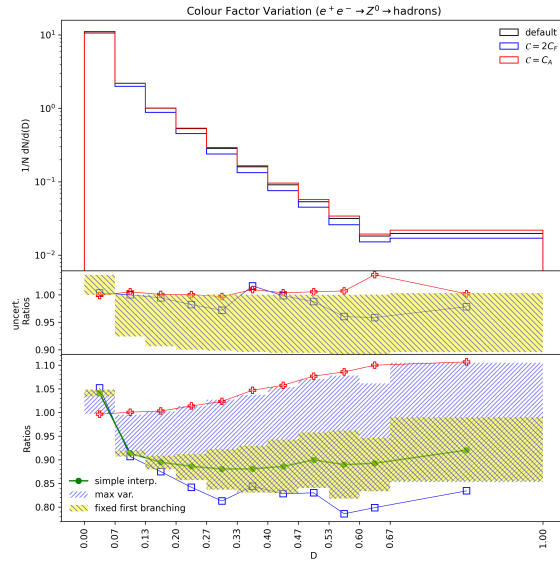

(a)

---

[10]The colour factor emission has not been hard coded. The first emission will be set to C_A if the first emission is from a gluon-gluon antenna.
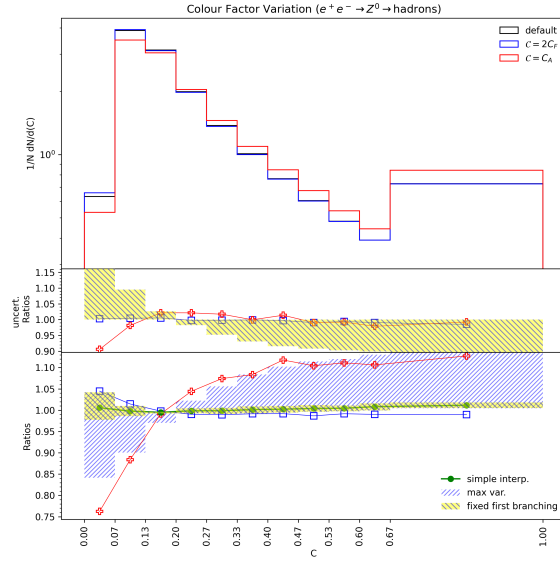
(b)

(c)

Figure 4.5. Comparison of colour factor uncertainty variations as obtained by reweighting the default shower and by rerunning the shower with modified antenna functions. The hashed region (max var.) is bounded by the variation in the observables (a) 1-T (b) C (c) D for range $\mathcal{C} = 2C_F, C_A$. The crosses and squares overlayed are the predictions from rerunning the shower. The top panel shows the variation in the differential cross-section. The bottom panel is the ratio between the reweighted and default shower. The central panel is the ratio between predicted and reweighted values. Weighted events used 10 million samples per variation bound while predictions used 5 million per prediction.

Chapter 5

Concluding Remarks

Parton showers have bridged well the divide between QFT and collider phenomenology by casting event generation in terms of an MCMC process. Tracing uncertainties through such a procedure is unfeasible without rerunning the shower. This can be overcome by assigning showers weights which over many runs amplify or suppress statistics collected as would be by a shower with alternative settings. This approach has been developed and verified for the DGLAP formalism in [68]. We translated this approach to the antenna language and implemented three types of variations in VINCIA. These being, renormalization-scale, pdf and colour factor variations and shown reweighted events indeed produce the correct variations. However, the quality of the result is contingent on several factors. The rate of convergence is quick for smaller scales but slower for larger scales. Additionally, We remind the reader that an event can only approximate an alternative event if the phase-space explored is similar. If all these criteria are met, larger variations will converge slower and possibly not entirely due to large weights which are capped.

Although, this method was only discussed in the context of FSR, it is applicable to ISR, resonance decays and MPI. Hence, future work would translate the formulae for these regimes and implement them in VINCIA. Colour factor variations also require refinement as to better allow parton showers to capture the subleading colour structure. In this thesis, we have only presented interpolations already suggested by others. In principle the variations on renormalization-scale variations may be made smaller. Since b_0 is evaluated depending on the number of flavours with masses below the resolution scale. Hence, there may be a discontinuity which increases the uncertainty. The additional term ([14], pg. 83)

$$\frac{\alpha_s}{6\pi} \ln\left(\frac{m_q}{kp_\perp}\right) \tag{5.0.1}$$

which makes the transition smoother. Hence, currently our estimation is overly conservative. The reader may have noticed that D-parameter variations were not in agreement with the predictions as other variables. This is since the convergence for the D-parameter is much slower. There must then be a to foresee the rate of convergence for different variables or what properties make some variables impact the rate of convergence. In producing the results shown here, the number of samples required before the results were seen to converge varied significantly.

We note that currently the weights distributions are being generated numerically within the Monte-Carlo. However, for some ranges of variations the distributions imply adherence to well known distributions. This is currently unjustified but presents a source of potentially interesting physics to be explored in the future. Figure 5.1 shows the distributions of weights within the first 5 bins corresponding to the lowest observable values i.e. left hand side region of plots. Larger bins have been neglected since sufficient statistics are difficult to obtain. It appears that renormalization-scale
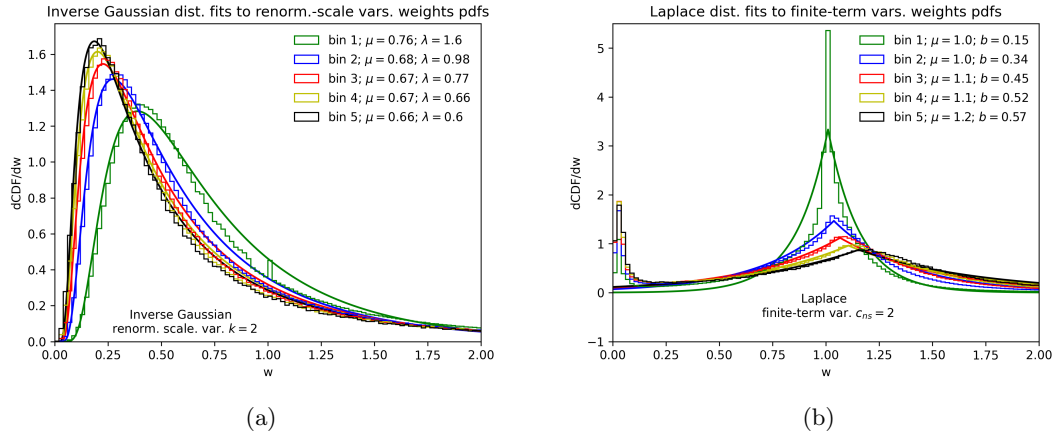
Figure 5.1. Weights distributions obtained from the accumulation of many event samples. The smooth curve is the fitted pdf of a (a) Inverse-Gaussian (b) Laplace distribution to (a) renormalization scale (b) finite-term variations. Only the first five bins are shown since the accumulation of statistics for bins at higher values of the observable is slow.

variations with $k = 2$ are well described by an Inverse-Gaussian distribution with varied parameters. Currently there is no statistical (or first-principles) justification for this, the parameters for the plots presented here have been obtained through best-fits. Perhaps the variation in (mu, λ) can be predicted based on the number of bins and the variation which could be used to better understand these variations. Similarly, figure 5.1b implies the distribution of weights for $c_{ns} = 2$ appear to follow a Laplace distribution, although there appears to be weak adherence for the first bin however. Note that a Laplace distribution has infinite support although weights are only allowed to be positive. However, recall that in equation (4.3.11) we restricted this possibility. Indeed, this would become clear in the context of sampling from a Laplace distribution. If the distribution of weights can be predicted then this may be a first step in calculating weights without the need of MCMC methods.

Appendix A

Veto Algorithm

Practical implementation of equation (3.2.33) can be exceedingly difficult in most circumstances due to unusual phase-space cuts or complicated integrands. This is addressed by a numerical implementation similar in spirit of that presented in §2.5.2. The distribution we wish to sample from is not equation (3.2.33) alone. While showering, a parton is free to branch between two scales with the confinement scale in between. This is remedied through the inclusion of an infra-red cut-off, denoted μ. The unitary of the total inclusive cross-section is then preserved with the addition of a no-branching up to the infra-red cut-off, beyond which no emission is permitted. Finally, a series of step functions throughout the expression ensure branchings only occur within the starting and infra-red scale. This leads to final expression

$$S_p(\mu, p_{\perp_1}^2, z | p_{\perp_0}^2) = \Delta(p_{\perp_0}^2, \mu)\delta(p_{\perp_1}^2 - \mu) + \Theta(p_{\perp_0}^2 - p_{\perp_1}^2)\Theta(p_{\perp_1}^2 - \mu)\pi(p_{\perp_1}^2, z | p_{\perp_0}^2) \qquad (A.0.1)$$

where Θ denote a step function. An illustration of this kernal is shown in figure A.1. equation (A.0.1) can be implemented using the veto algorithm given in algorithm 1. Just as with the accept/reject algorithms, we require a blanketing function $r(p_\perp^2, z) \geq K_{j/\tilde{i}\tilde{k}}(p_\perp^2, z)$ for all $(p_\perp^2, z)$ in within the accessible region of the phase-space. The proof for this algorithm is frequent in the literature *e.g.* see [68, 75]. The validity of the algorithm which is used in the proof is that the blanketing distribution largely over estimates inducing several failed trials. The accuracy of the algorithm relies on the number of failed trials, the more the better. Since algorithm 1 assumes a single branching channel, for a discussion on the inclusion of multiple channels, see [14] pg. 21.
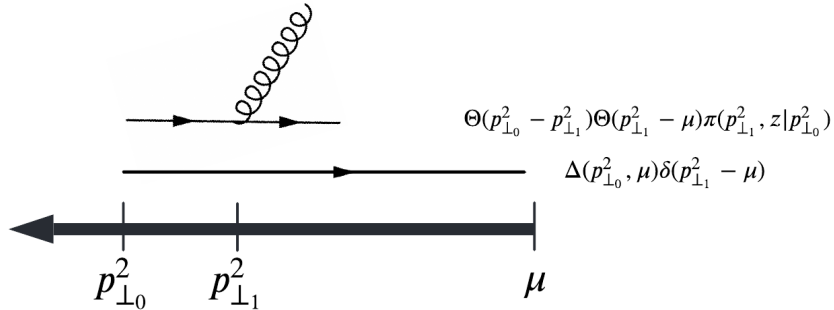

Figure A.1. Visualization of components of equation (A.0.1) which represents the entire branching kernel.

Algorithm 1 Veto Algorithm(sometimes referred to as Sudakov Veto Algorithm) implementing equation (A.0.1). We use the compact notation $\tilde{\Delta}(p_{\perp_0}^2, p_{\perp_1}^2) = \exp\left(-\int r\right)$. r blankets $K_{j/\tilde{i}\tilde{k}}$ in $p_\perp^2$ and z, however we suppress the z dependence for clarity.

1: **while** True **do**
2:     $u \leftarrow \text{Uniform}(0,1)$ solve $u = \tilde{\Delta}(p_{\perp_0}^2, p_{\perp_1}^2)\Theta(p_{\perp_1}^2 - \mu)$ for $p_{\perp_1}^2$
3:     **if** $p_{\perp_1}^2 = \mu$ **then return** μ
4:     **else**   ▷ Accept conditionally to correct for blanketing
5:         $v \leftarrow \text{Uniform}(0,1)$
6:         **if** $v < K_{j/\tilde{i}\tilde{k}}(p_{\perp_1}^2)/r(p_{\perp_1}^2)$ **then return** $p_{\perp_1}^2$
7:         **end if**
8:     **end if**
9: **end while**

Bibliography

[1] B. Hall, *Lie groups, lie algebras, and representations an elementary introduction*. Springer International Publishing, 2015.

[2] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2021.

[3] H. Georgi, *Lie algebras in particle physics*, vol. 54. Perseus Books, Reading, MA, 2nd ed., 1999.

[4] M. Gell-Mann, *The eightfold way: A theory of strong interaction symmetry*, in *The Eightfold Way*, pp. 11–57, CRC Press, (2018).

[5] I. Vuković, "The irreducible representations of the lie algebra su(2)." https://www.phy.uniri.hr/~ivukovic/su(2)IRREPs.pdf, 2015.

[6] J. D. JACKSON, *Classical electrodynamics*. JOHN WILEY & SONS, 2021.

[7] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Perseus Books, 1995.

[8] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 235–295, 2015, DOI [1411.4085].

[9] M. Sandhoff and P. Skands, *Colour annealing-a toy model of colour reconnections*, tech. rep., Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States), 2005.

[10] J. R. Christiansen and T. Sjöstrand, *Color reconnection at future e^+ e^- colliders*, *Eur. Phys. J. C* **75** (2015) 441 [1506.09085].

[11] B. Andersson, *The Lund model*, vol. 7. Cambridge University Press, 7, 2005, 10.1017/CBO9780511524363.

[12] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31.

[13] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Höche, H. Hoeth et al., *General-purpose event generators for LHC physics*, *Physics Reports* **504** (2011) 145.

[14] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2203.11601.

[15] C. B. et. al, "Pythia 8 online manual." https://pythia.org/latest-manual/.

[16] "Sherpa documentation." https://sherpa-team.gitlab.io/.

[17] SHERPA collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[18] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [1512.01178].

[19] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639 [0803.0883].

[20] "The herwig event generator." https://herwig.hepforge.org/.

[21] CDF COLLABORATION collaboration, *Transverse momentum and total cross section of e^+e^- pairs in the z-boson region from $p\overline{p}$ collisions at $\sqrt{s} = 1.8$TeV*, *Phys. Rev. Lett.* **84** (2000) 845.

[22] F. Scheck, *Classical field theory on electrodynamics, Non-Abelian Gauge Theories and gravitation*. Springer Berlin, 2019.

[23] D. Tong, "Quantum field theory." http://www.damtp.cam.ac.uk/user/tong/qft/qft.pdf.

[24] N. N. Bogolûbov, A. A. Logunov, A. Oksak and I. Todorov, *General principles of quantum field theory*, vol. 10. Springer Science & Business Media, 2012.

[25] K. H. et. al, *Review of particle physics: Particle data group*, *Physical Review D* **66** (2002) 100011.

[26] D. Griffiths, *Introduction to elementary particles*. John Wiley & Sons, 2020.

[27] OPAL collaboration, *Tests of the standard model and constraints on new physics from measurements of fermion pair production at 189-GeV to 209-GeV at LEP*, *Eur. Phys. J. C* **33** (2004) 173 [hep-ex/0309053].

[28] S. Bethke, *The 2009 world average of α_s*, *The European Physical Journal C* **64** (2009) 689.

[29] W. R. Gilks, S. Richardson and D. Spiegelhalter, *Markov chain Monte Carlo in practice*. CRC press, 1995.

[30] D. Gamerman and H. F. Lopes, *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC press, 2006.

[31] S. Brooks, A. Gelman, G. Jones and X.-L. Meng, *Handbook of markov chain monte carlo*. CRC press, 2011.

[32] T. Tao, *An introduction to measure theory*, vol. 126. American Mathematical Society Providence, 2011.

[33] E. Kreyszig, *Introductory functional analysis with applications*, vol. 17. John Wiley & Sons, 1991.

[34] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, Inc., USA, 1st ed., 1996.

[35] F. James, *Monte carlo theory and practice*, *Reports on Progress in Physics* **43** (1980) 1145.

[36] S. Weinzierl, *Introduction to Monte Carlo methods*, hep-ph/0006269.

[37] H. G. Katzgraber, *Introduction to monte carlo methods*, *arXiv preprint arXiv:0905.1629* (2009) .

[38] S. Abbott, *Understanding Analysis*. Springer, 2015.

[39] S. D. Drell and T.-M. Yan, *Massive lepton-pair production in hadron-hadron collisions at high energies*, *Phys. Rev. Lett.* **25** (1970) 316.

[40] M. Ritzmann, D. Kosower and P. Skands, *Antenna showers with hadronic initial states*, *Physics Letters B* **718** (2013) 1345.

[41] T. D. Gottschalk, *Backwards evolved initial state parton showers*, *Nuclear Physics B* **277** (1986) 700.

[42] C. T. Preuss, *Improving the Efficiency and Accuracy in Monte-Carlo Event Generation*, Ph.D. thesis, Monash U., 2022. 10.26180/18133637.v1.

[43] T. Sjostrand, *Monte Carlo Generators*, in *2006 European School of High-Energy Physics*, pp. 51–74, 11, 2006, hep-ph/0611247.

[44] Z. Nagy and D. E. Soper, *What is a parton shower?*, *Phys. Rev. D* **98** (2018) 014034.

[45] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [hep-ph/0409313].

[46] P. Z. Skands, *QCD for Collider Physics*, in *2010 European School of High Energy Physics*, 4, 2011, 1104.2863.

[47] T. Kinoshita, *Mass singularities of feynman amplitudes*, *Journal of Mathematical Physics* **3** (1962) 650.

[48] T. D. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, *Phys. Rev.* **133** (1964) B1549.

[49] Y. Dokshitzer, G. Leder, S. Moretti and B. Webber, *Better jet clustering algorithms*, *Journal of High Energy Physics* **1997** (1997) 001.

[50] F. E. James, *Monte Carlo phase space*, CERN Academic Training Lecture. CERN, Geneva, 1968, 10.5170/CERN-1968-015.

[51] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[52] T. Sjostrand, *A Model for Initial State Parton Showers*, *Phys. Lett. B* **157** (1985) 321.

[53] S. Jadach and B. Ward, *Yfs2— the second-order monte carlo program for fermion pair production at lep/slc, with the initial state radiation of two hard and multiple soft photons*, *Computer Physics Communications* **56** (1990) 351.

[54] P. Skands, "Qft beyond fixed order: Introduction to bremsstrahlung and jets."

[55] W. Kilian, T. Ohl, J. Reuter and C. Speckner, *QCD in the Color-Flow Representation*, *JHEP* **10** (2012) 022 [1206.3700].

[56] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl. Phys. B* **72** (1974) 461.

[57] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color-flow decomposition of QCD amplitudes*, hep-ph/0209271.

[58] A. J. Larkoski and M. E. Peskin, *Spin-dependent antenna splitting functions*, 0908.2450 [hep-ph].

[59] N. Fischer, S. Prestel, M. Ritzmann and P. Skands, *Vincia for Hadron Colliders*, *Eur. Phys. J. C* **76** (2016) 589 [1605.06142].

[60] T. Virkus, H. Thomsen, E. Uggerhøj, U. Uggerhøj, S. Ballestrero, P. Sona et al., *Direct measurement of the chudakov effect*, *Physical review letters* **100** (2008) 164802.

[61] P. Z. Skands, "Introduction to bremsstrahlung and jets."

[62] F. Abe, M. Albrow, D. Amidei, C. Anway-Wiese, G. Apollinari, M. Atac et al., *Evidence for color coherence in pp^- collisions at $\sqrt{s}$ =1.8 TeV*, *Phys. Rev. D* **50** (1994) 5562.

[63] A. G.-D. Ridder, M. Ritzmann and P. Skands, *Timelike dipole-antenna showers with massive fermions*, 1108.6172 [hep-ph].

[64] H. Brooks, C. T. Preuss and P. Skands, *Sector showers for hadron collisions*, 2003.00702 [hep-ex, physics:hep-ph].

[65] F. A. Dreyer, G. P. Salam and G. Soyez, *The lund jet plane*, 1807.04758 [hep-ex, physics:hep-ph].

[66] ATLAS, CMS collaboration, *Prospects for SUSY and BSM Physics at the High Luminosity LHC*, .

[67] I. Zurbano Fernandez et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, .

[68] S. Mrenna and P. Skands, *Automated Parton-Shower Variations in Pythia 8*, *Phys. Rev. D* **94** (2016) 074005 [1605.08352].

[69] W. T. Giele, D. A. Kosower and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev. D* **84** (2011) 054003 [1102.2126].

[70] J. Bellm, S. Plätzer, P. Richardson, A. Siódmok and S. Webster, *Reweighting Parton Showers*, *Phys. Rev. D* **94** (2016) 034028 [1605.08256].

[71] JADE collaboration, *Observation of Planar Three Jet Events in e+ e- Annihilation and Evidence for Gluon Bremsstrahlung*, *Phys. Lett. B* **91** (1980) 142.

[72] L. Collaboration, *Studies of hadronic event structure in e+e- annihilation from 30 GeV to 209 GeV with the l3 detector*, `hep-ex/0406049`.

[73] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys. B* **349** (1991) 635.

[74] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni and G. P. Salam, *Logarithmic accuracy of parton showers: a fixed-order study*, `1805.09327 [hep-ex, physics:hep-ph]`.

[75] S. Platzer and M. Sjodahl, *The Sudakov Veto Algorithm Reloaded*, *Eur. Phys. J. Plus* **127** (2012) 26 [`1108.6180`].